DEUTSCHE GEODÄTISCHE KOMMISSION
bei der Bayerischen Akademie der Wissenschaften

Reihe C Dissertationen Heft Nr. 577

Jianqing Cai

Statistical Inference of the Eigenspace Components of a Symmetric Random Deformation Tensor

München 2004

Verlag der Bayerischen Akademie der Wissenschaften
in Kommission beim Verlag C. H. Beck

ISSN 0065-5325 ISBN 3 7696 5016 6

DEUTSCHE GEODÄTISCHE KOMMISSION
bei der Bayerischen Akademie der Wissenschaften

Reihe C Dissertationen Heft Nr. 577

Statistical Inference of the Eigenspace Components of a Symmetric Random Deformation Tensor

Doctoral thesis
accepted by the Faculty of Aerospace Engineering and Geodesy of the University of Stuttgart

Von der Fakultät für Luft- und Raumfahrttechnik und Geodäsie der Universität Stuttgart
zur Erlangung der Würde eines Doktor-Ingenieurs (Dr.-Ing.)
genehmigte Dissertation

submitted by / vorgelegt von

Jianqing Cai

München 2004

Verlag der Bayerischen Akademie der Wissenschaften
in Kommission beim Verlag C. H. Beck

ISSN 0065-5325 ISBN 3 7696 5016 6

Adresse der Deutschen Geodätischen Kommission:

Deutsche Geodätische Kommission

Marstallplatz 8 ● D – 80 539 München

Telefon +49 - (0)89 - 23 031 -0 / -1113 ● Telefax +49 - (0)89 - 23 031 -1283 / -1100
E-mail hornik@dgfi.badw.de ● http://dgk.badw.de

Main referee / Hauptberichter: Prof. Dr.-Ing.habil. Dr.tech.h.c.mult. Dr.-Ing.E.h. E. W. Grafarend (Universität Stuttgart)
Co-referees / Mitberichter: Prof. Dr.habil. H. Walk (Universität Stuttgart)
Prof. Dr.-Ing.habil. B. Schaffrin (The Ohio State University, USA)

Day of submission / Tag der Einreichung: 26.02.2004
Day of examination / Tag der mündlichen Prüfung: 05.07.2004

ISSN 0065-5325 ISBN 3 7696 5016 6

Abstract

For the *validation of a symmetric rank-two random tensor*, for instance of strain and stress, the *eigenspace components* (principal components, principal directions) play a key role. They classify deformation and stress patterns in earthquake regions, of plate tectonics and of glacially isostatic rebounds. The main purpose of this study is to develop the proper statistical inference for the eigenspace components of a two- and three-dimensional symmetric deformation tensor. Let us assume that the strain or stress tensor has been directly observed or indirectly determined by other measurements. According to the *Measurement Axiom* such a symmetric rank-two tensor is *random*. For its *statistical inference*, we assume that the random tensor is *tensor-valued Gauss-Laplace normally distributed*. It is proven that the vectorized three-dimensional symmetric random tensor $\mathbf{y}$ = vech $\mathbf{T} \in \mathbb{R}^{6\times1}$ has a BLUUE estimate $\hat{\boldsymbol{\mu}}_{\mathrm{y}} \in \mathbb{R}^{6\times1}$ which is multivariate normally distributed, $\hat{\boldsymbol{\mu}}_{\mathrm{y}} \sim \mathcal{N}_6(\boldsymbol{\mu}_{\mathrm{y}}, n^{-1}\boldsymbol{\Sigma}_{\mathrm{y}}; \hat{\boldsymbol{\mu}}_{\mathrm{y}})$, where n is the number of full tensor observations and $\boldsymbol{\Sigma}_{\mathrm{y}} = D\{\mathrm{vech}\,\mathbf{T}\}$, the variance-covariance matrix of $\mathbf{y}$. The BIQUUE sample variance-covariance matrix $\hat{\boldsymbol{\Sigma}}_{\mathrm{y}}$ is Wishart distributed $\hat{\boldsymbol{\Sigma}}_{\mathrm{y}} \sim \mathcal{W}_6(n-1, (n-1)^{-1}\boldsymbol{\Sigma}_{\mathrm{y}}; \hat{\boldsymbol{\Sigma}}_{\mathrm{y}})$. The *eigenspace synthesis* relates the eigenspace elements to the observations by means of a nonlinear vector-valued function establishing a *special nonlinear multivariate Gauss-Markov model*. For the linearized forms, we have succeeded to construct *BLUUE* (*B*est *L*inear *U*niformly *U*nbiased *E*stimation) *of the eigenspace elements* and *BIQUUE* (*B*est *I*nvariant *Q*uadratic *U*niformly *U*nbiased *E*stimation) of its variance-covariance matrix for the two- and three-dimensional cases. The test statistics, such as *Hotelling's* T^2, *Lawley-Hotelling's trace test, likelihood ratio statistics and Growth-Curve model* are proposed. In two case studies both model and hypothesis tests are applied to the two- and three-dimensional, symmetric rank two strain rate tensor observations in the region of central Mediterranean and Western Europe, which are derived from ITRF92 to ITRF2000 series station positions and velocities. The related *linear hypothesis test* has documented large confidence regions for the eigenspace components, namely *eigenvalues and eigendirections*, based upon real measurement configurations. They lead to the statement *to be cautious* with data of type extension and contraction as well as the orientation of principal stretches.

Numerical tests have documented that the estimate $\hat{\boldsymbol{\xi}}$ of type BLUUE of the parameter vector $\boldsymbol{\xi}$ within a linear *Gauss-Markov model* $\{\mathbf{A}\boldsymbol{\xi} = E\{\mathbf{y}\},\ \boldsymbol{\Sigma}_{\mathrm{y}} = D\{\mathbf{y}\}\}$ is *not* robust against *outliers* in the stochastic observation vector $\mathbf{y}$. It is for this reason that *we give up* the postulate of unbiasedness, but keeping the set-up of a *linear estimation* $\hat{\boldsymbol{\xi}} = \mathbf{L}\mathbf{y}$ of homogeneous type. According to best linear estimators of type homBLE (*Best homogeneously Linear Estimation*), S-homBLE and α-homBLE of the *fixed effects* $\boldsymbol{\xi}$ (*Grafarend and Schaffrin* 1993, *Schaffrin* 2000). We have developed a new method of determining the optimal regularization parameter α in uniform Tykhonov-Phillips regularization (α-weighted BLE) by minimizing the trace of the Mean Square Error matrix $MSE\{\hat{\boldsymbol{\xi}}\}$ (A-optimal design) in the general case. Within two case studies, the new method is tested and analyzed in the univariate and the multivariate case with data which is derived from simulated observations of a random tensor of type strain rate.

Zusammenfassung

Für die Validierung eines symmetrischen Zufallstensors, zum Beispiel der Spannung und Strain, spielen die Eigenkomponenten (Hauptverzerrungen und Orientierungen) eine Schlüsselrolle. Mit ihnen lassen sich Spannung und Strain in erdbebengefährdeten Regionen, bei der Plattentektonik sowie bei isostatisch postglazialen Hebungen klassifizieren. Die Entwicklung geeigneter mathematisch-statistischer Verfahren zur Schätzung der Eigenkomponenten eines zwei- oder dreidimensionalen symmetrischen Deformationstensors ist der Hauptgegenstand der vorliegenden Arbeit. Es wird angenommen, dass der Spannungs- oder Straintensor entweder direkt beobachtet oder aus anderen Beobachtungen abgeleitet wurde. Auf Grund des *Beobachtungsaxioms* ist ein solcher symmetrischer Tensor zweiter Stufe zufällig. Für seine statistische Inferenz nehmen wir an, dass der Zufallstensor Gauß-Laplace normal verteilt ist. Es wird gezeigt, dass der vektorisierte dreidimensionale symmetrische Zufallstensor $\mathbf{y}$ = vech $\mathbf{T} \in \mathbb{R}^{6\times1}$ eine beste lineare erwartungstreue uniforme Schätzung (BLUUE) $\hat{\boldsymbol{\mu}}_{\mathrm{y}} \in \mathbb{R}^{6\times1}$ hat. Diese ist multivariat normalverteilt mit $\hat{\boldsymbol{\mu}}_{\mathrm{y}} \sim \mathcal{N}_6(\boldsymbol{\mu}_{\mathrm{y}}, n^{-1}\boldsymbol{\Sigma}_{\mathrm{y}}; \hat{\boldsymbol{\mu}}_{\mathrm{y}})$. n ist die Anzahl der Tensor-Beobachtungen, $\boldsymbol{\Sigma}_{\mathrm{y}} = D\{\mathrm{vech}\,\mathbf{T}\}$ die Varianz-Kovarianz-Matrix der Beobachtungen $\mathbf{y}$. Die BIQUUE (beste invariante quadratische uniforme erwartungstreue Schätzung) Varianz-Kovarianz-Matrix $\hat{\boldsymbol{\Sigma}}_{\mathrm{y}}$ der Stichprobe ist *Wishart*-verteilt $\hat{\boldsymbol{\Sigma}}_{\mathrm{y}} \sim \mathcal{W}_6(n-1, (n-1)^{-1}\boldsymbol{\Sigma}_{\mathrm{y}}; \hat{\boldsymbol{\Sigma}}_{\mathrm{y}})$. Da die Eigenraumsynthese eines symmetrischen Zufallstensors bezüglich tensorwertiger Beobachtungen nichtlinear ist, müssen die jeweiligen Parameter innerhalb eines speziellen nichtlinearen multivariaten *Gauß-Markoff Modells* geschätzt werden. Zur Stichprobenprüfung der

Eigenraumsynthese wird dessen Linearisierung aus den ursprünglich nichtlinearen Beobachtungsgleichungen abgeleitet. Die Schätzungen ($\boldsymbol{\Sigma}$ -BLUUE) der Eigenraumbestandteile und die Schätzung ihrer Varianz-Kovarianzmatrix der Art BIQUUE werden für den zwei- und dreidimensionalen Fall entwickelt und entsprechende Teststatistiken wie *Hotelling's* T^2, die Likelihood-Verhältnisstatistiken und das „*Growth-Curve* model“ generiert. In zwei Fallstudien werden sowohl Modell- als auch Hypothesentests auf zwei- und dreidimensionale, symmetrische Tensor-Beobachtungen der Strainraten in der zentralen Mittelmeerregion und Westeuropa angewendet, die von Stationspositionen und -geschwindigkeiten der Reihe ITRF92 bis ITRF 2000 abgeleitet werden. Die verwandten Hypothesentests liefern, basierend auf realen Messkonfigurationen, große Konfidenzintervalle für die Eigenwerte und Eigenrichtungen, so dass mit der Interpretation der Größenausdehnung, Kontraktion und Hauptstreckungsrichtung äußerst vorsichtig umgegangen werden muss.

Numerische Tests haben dokumentiert, dass die Schätzung $\hat{\boldsymbol{\xi}}$ des Typs BLUUE des Unbekanntenvektors $\boldsymbol{\xi}$ im linearen *Gauss-Markov-Modell* $\{\mathbf{A}\boldsymbol{\xi} = E\{\mathbf{y}\},\ \boldsymbol{\Sigma}_y = D\{\mathbf{y}\}\}$ nicht gegen Ausreißer im stochastischen Beobachtungsvektor robust ist. Aus diesem Grund geben wir das Postulat der Unverzerrtheit auf, behalten aber den Ansatz der homogenen linearen Schätzung $\hat{\boldsymbol{\xi}} = \mathbf{Ly}$ bei. Auf Grundlage bester linearer Schätzer vom Typ α-homBLE (beste homoge lineare Schätzung), S-homBLE und α-homBLE der fixen Effekte $\boldsymbol{\xi}$ (*Grafarend und Schaffrin* (1993), *Schaffrin* (2000)) haben wir eine neue Methode der Bestimmung des optimalen Regularisierungsparameters α einer uniformen *Tykhonov-Phillips*-Regularisierung (α-gewichtete BLE) für den allgemeinen Fall entwickelt. Das Kriterium ist die Minimierung der Spur der Matrix $MSE\{\hat{\boldsymbol{\xi}}\}$ der mittleren Fehlerquadrate (*A-optimales Design*). Im Rahmen zweier Fallstudien wird die neue Methode für den univariaten und multivariaten Fall mit Daten, die aus simulierten Beobachtungen eines Strainratentensors abgeleitet werden, getestet und analysiert.

Acknowledgments

I would like to express my sincere gratitude to my supervisor, Prof. Dr. -Ing. habil. Dr. tech. h.c. mult. Dr.-Ing. E.h. mult. Erik W. Grafarend, for introducing the classic and modern mathematical statistics topic and their application to the analysis of deformation tensor to me and for his permanent scientific guidance, inspiration and continuous encouragement. I also appreciate very much his constructive critical comments throughout the period of this study, especially during the time after his recovery from severe illness in 2003.

My sincere thanks go to Prof. Dr. habil. Harro Walk and Prof. Dr.-Ing. habil. Burkhard Schaffrin for their voluntary takeover as co-referees. Especially Prof. Schaffrin gave many constructive comments, which led to the result of the optimal parameter for Tykhonov-Phillips regularization by A-optimal design, estimates of eigenspace components of a random tensor. Thanks must also be extended to him for the proof reading and corrections in the review stage.

In addition, I would like to thank Prof. Dr. Peiliang Xu for his many suggestions during the early stage of this study.

Many thanks also go to my colleagues at the Geodetic Institute, University of Stuttgart, especially Dr.-Ing. Friedrich W. Krumm, Dr.-Ing. Johannes Engels and Dr.-Ing. Behzad Voosoghi for the scientific support and discussions in the eigenspace synthesis of a 3-D tensor and the calculating of the strain/stress tensor, as well as Ms. Anita Vollmer for early proof reading.

All in all this work represents the results of the DFG Research Project (GR 323/36-1/2). Therefore I also appreciate very much Prof. Dr. J. Hartung's efforts in initialising this project.

Last but not least, I thank my wife Li Tao and our son Hongxiang for their understanding, patience and continuous support during the time of my research and the writing of the dissertation.

Contents

Chapter 0

Introduction

A central task of Geosciences, in particular of Geodesy and Geophysics, is to determine the temporal change of the Earth's shape by observations and analysis of geodetic and geophysical global, regional or local networks. With the new space geodetic techniques, such as Very Long Baseline Interferometry (VLBI), Satellite Laser Ranging (SLR), and Global Positioning System (GPS), three-dimensional positions and velocities of points in these networks have been determined with high accuracy (~ mm level) from relative regular measurement campaigns, which have become a key tool in plate tectonic studies. These data have improved our knowledge and understanding of (1) regional deformation and strain accumulation related to earthquakes, (2) contemporary relative plate tectonic motions of the North American, Pacific, South American, Eurasian, Australian, Nazca, and Caribbean plate, (3) internal deformation of lithospheric plates, and (4) crustal motion and deformation occurring in the regions of high earthquake activity. These facts suggest that the components of deformation measures such as the symmetric stress or strain tensor can be estimated from the highly accurate geodetic data and analyzed through the proper statistical testing procedures.

Deformation tensors are practically random, since they are either directly measured or indirectly inverted from other geo-measurements. The estimate of random symmetric rank two tensors and associated statistical inference are usually based on the statistics, e.g., sample means and sample variance-covariance. So we should firstly derive the sampling distributions of sample means of the random tensor. The values of any sample statistics depend on a particular samples one happens to obtain. It varies from sample to sample. Thus a statistic is a random variable. As such, it has a probability distribution called sampling distribution. We owe the early development of sampling distributions under normality to *P.S. Laplace* (1812), *Carl Friedrich Gauss* (1816), *Friedrich Robert Helmert* (1876a) for the *Helmert distribution* (which is highly valued as the starting point for modern small sample theory), *Thorvald N. Thiele* (1889, 1903), *Karl Pearson* (1900) for his *Chi square distribution*, *Sealy Gosset* (1908a, b) for his *Student t-distribution*, *Ronald A. Fisher* (1920, 1922) for the *F-distribution* and *John Wishart* (1928) for the *Wishart distribution*.

The hypothesis test of sample mean vector and sample variance-covariance matrix of a symmetric random tensor belongs to multivariate analysis which is the branch of statistics devoted to the study of random variables that are not necessarily independent. Where inference is concerned several (generally correlated) measurements are made on every observed subject. Many current multivariate statistical procedures were developed during the first half of the twentieth century. A reasonable complete list of the developers would be voluminous. However, a few individuals can be cited as having made important initial contributions to the theory and practice of multivariate analysis. *T. Galton* and *K. Pearson* did pioneering work in the area of correlation and regression analysis. *R.A. Fisher's* derivation of the exact distribution of the sample correlation coefficient and related quantities provided the impetus for multivariate distribution theory. *C. Spearman* and *K. Pearson* were amongst the first to work in the area of factor analysis. Significant contributions to multivariate analysis were made during the 1930s by. *S. S. Wilks* (general procedures for testing certain multivariate hypotheses), *H. Hotelling* (*Hotelling's* T^2, *principle component analysis, canonical correlation analysis*), *R. A. Fisher* (*discrimination and classification*), and *P. C. Mahalanobis* (generalized distance, hypothesis testing). *J. Wishart* derived an important joint distribution of sample variance and covariance that bears his name. Later *M. Bartlett and G. E. P. Box* contributed to the large sample theory associated with certain multivariate test statistics. The body of statistical methodology used to analyze simultaneous measurements on many variables is called multivariate analysis. Many multivariate methods are based on an underlying probability model known as the multivariate normal. The objectives of scientific investigations, for which multivariate methods most naturally lend themselves, include the following:

- Data reduction or structural simplification.
- Sorting and grouping.
- Investigation of dependence among variables.
- Predication.
- Hypothesis construction and testing.

In the deformation analysis in geosciences (geodesy, geophysics and geology), we are often confronted with the problem of a two-dimensional (or planar and horizontal), symmetric rank-two deformation tensor. Its *eigenspace components (principal components*, *principal direction*) play an important role in interpreting the geodetic phenomena like earthquakes (seismic deformations), plate motions and plate deformations among others. With the new space geodetic methods three-dimensional positions and velocities of points in these networks have been determined with high accuracy (~ mm level) from relative regular measurement campaigns, which have become a key tool in plate tectonic studies. This fact suggests that the components of a two-dimensional deformation tensor can be estimated from the high accuracy geodetic data and analyzed through the proper statistical testing procedures. According to the *Measurement Axiom* such a two-dimensional, symmetric (2, 0) tensor is a *random tensor* **T** which we assume to be an element of the tensor-valued *Gauss-Laplace* normal distribution over $\mathbb{R}^{2\times 2}$ of type independently, identically distributed (i.i.d.) tensor-valued observations, but with identical off-diagonal elements.

In reality, crustal motions and deformation are of three-dimensional nature and most deformation tensors derived from geodetic, geological and seismological observations are three-dimensional, such as the seismic moment tensors. In the last two decades some efforts have been made to formulate the problem in the three-dimensional space. A curvilinear three-dimensional finite element method has been introduced by *Grafarend* (1986) for the representation of local strain and local rotation tensors in terms of ellipsoidal, *Gauss-Krüger* or *UTM* coordinates. More researches about the three-dimensional strain and strain rate tensor analysis in geodesy are referred to the papers of *Brunner* (1979), *Lichtenegger and Sünkel* (1989), *Dermanis and Grafarend* (1993) and *Wittenburg*(1999). In comparison with the more complete solution and application about two-dimensional deformation tensors on the Earth, there are three aspects that limit the deformational analysis in three dimensions: (1) there is insufficient accuracy of vertical components of point positions due to unresolved modeling errors; (2) the vertical movements of the crust, whether uplift or subsidence, are generally an order of magnitude smaller than horizontal movements; (3) the restriction of the extensive geodetic measurements to the earth's surface make the vertical gradient of the velocity vector generally unobservable. In addition to the analysis of the eigenspace components of three-dimensional deformation tensor it is difficult to uniquely determine the three eigendirections.

Random tensors, also called random matrices, were first analysed in *Nuclear Physics* (*Porter* 1965, *Mehta* 1991), independently in *Mathematical Statistics* (*Anderson* 1984) in the context of multivariate modeling. Given the *probability density function* (pdf) of a random tensor of second order, it has been documented that apart from special cases, the exact *probability density function* of the *random eigenspace components* cannot be found in a closed form. Accordingly, statistical analysts have focussed on approximate and/or limit distributions, for instance, of the products of random matrices and/or the random eigenspace components (*Anderson* 1958; *Mehta* 1991; *Girko* 1979, 1990, 1995, 2000; *Cohen, Kesten & Newman* 1984).

In the Earth Science random tensors have only recently been investigated from the statistical point of view. Since the tensors in the Earth Science are physical quantities and their dimensions are generally low (3 for stress/strain tensors and 6 for elastic material tensors), mathematically approximate/limit distributions of the random eigenspace components are of limited practical value. In fact, the study of random stress/strain has been focused on the following four aspects: (1) the exact distribution of the random principal stress/strain components, since the dimension of stress/strain tensors is not greater than three and since the number of measurements is always finite; (2) the accuracy of the random eigenspace components. The accuracy is generally not investigated in the mathematical literature of rank-two random tensors. It is however a routine indicator that must be attached to any estimated/derived geo-quantity; (3) the biases of the random eigenspace components. Since the mapping between a stress/strain tensor and its eigenspace components is nonlinear, the random eigenspace components are biased. The biases of the eigenspace components, except for some inequality results on the biases of the random eigenvalues (see e.g. Cacoullos 1965), have not been well investigated in the mathematical literature on rank-two random tensors. They can have an important role to play in correctly interpreting the estimated stress/strain field geophysically, however; and (4) the eigendirections. The eigendirections have been almost always treated as nuisance parameters in nuclear physics and multivariate analysis. Geophysically, the eigendirections are very important and thus cannot be ignored (*Xu and Grafarend* 1996b).

The first work on the statistical analysis of random tensors in the Earth Sciences was to compute the first-order accuracy of the principal eigenvalues of a symmetric, rank-two random tensor (*Angelier et al*. 1982 as an appendix, and probably independently, *Soler & van Gelder* 1991). *Kagan & Knopoff* (1985) studied statistically the first two moments of stochastic three-dimensional (3D) seismic moment tensor invariants, which were used to explain complex fault geometry (*Kagan* 1992).

On the assumption that a strain tensor or stress tensor has been directly measured or derived from other observations, such a two-dimensional, symmetric rank-two tensor is a random tensor **T** which we assume to be an element of the tensor-valued *Gauss-Laplace* normal distribution over $\mathbb{R}^{2\times 2}$ of type independently, identically distributed (i.i.d.) tensor-valued observations. The distribution of the *eigenspace components* of the rank-two random tensor (principal components, principal directions) has been investigated by *Xu and Grafarend* (1996a; 1996b), which is significantly different from the commonly used *Gauss-Laplace* normal distribution. The possible bias terms of eigenspace components and the nonlinear error propagation are also studied in these two papers. By means of the numerical analysis with the power series expansion, the marginal probability density function $f_\lambda(\lambda_1, \lambda_2)$ of random eigenvalues λ_1, λ_2 has been approximately computed by *Cai* (2001). In recent years *Xu* (1999a) and *Kagan* (2000) developed the general distribution of the eigenspace components of the symmetric, rank-two random tensor, which can hardly be applied directly to real Engineering and Earth Science problems, since an exact distribution theory of eigenspace components is almost always unavailable.

These reasons give rise to investigate the

> Statistical inference of eigenspace components
> of the two- and three-dimensional symmetric,
> rank-two random tensor ("random matrix")

based upon a linearized multivariate *Gauss-Markov* model which will provide us with the second-order statistics of eigenspace components. Such a statistical inference on a random matrix is completed by the design of a linear hypothesis test.

With the benefit of the development of the space geodesy and the continuous observations of the permanent networks, such as, *International GPS Service* (*IGS*) Network, *International Laser Ranging Service* (*ILRS*) Network, *International VLBI Service for Geodesy and Astrometry* (*IVS*) Network and *International DORIS Service* (*IDS*) Network and their combination *International Terrestrial Reference Frame* (*ITRF*) by IERS, we can now derive the strain rate tensor observation and estimate the eigenspace component parameters of these random tensor samples, which address not only the present-day deformation pattern but also their continuous change of them. In two case studies both BLUUE and BIQUUE models and hypothesis tests are applied to the eigenspace components of two- and three-dimensional strain rate tensor observations in the area of central Mediterranean and Western Europe, which are derived from ITRF92 to ITRF2000 series station positions and velocities in Sections 6.6.and 6.7. The related *linear hypothesis test* has documented large confidence regions for the eigenspace components, namely *eigenvalues and eigendirections*, based upon real measurement configurations. They lead to the statement *to be cautious* with data of type extension and contraction as well as with the orientation of principal stretches.

In the estimate of deformation tensor we often see that the estimate $\hat{\boldsymbol{\xi}}$ of type BLUUE of the parameter vector $\boldsymbol{\xi}$ within a linear *Gauss-Markov model* $\{\mathbf{A}\boldsymbol{\xi} = E\{\mathbf{y}\},\ \boldsymbol{\Sigma}_\mathrm{y} = D\{\mathbf{y}\}\}$ is *not* robust against *outliers* in the stochastic observation vector $\mathbf{y}$. It is for this reason that we give up the postulate of unbiasedness, but keeping the set-up of a *linear estimation* $\hat{\boldsymbol{\xi}} = \mathbf{L}\mathbf{y}$ of homogeneous type. The biased estimation is a special inverse problem, also related to *Tykhonov-Phillips regulator* or *ridge estimator*. Ever since *Tykhonov* (1963) and *Phillips* (1962) introduced the *hybrid minimum norm approximation solution* (HAPS) of a *linear improperly posed problem* there has been left the open problem to evaluate the weighting factor α between the least-squares norm and the minimum norm of the unknown parameters. Since the 1960s this problem has been studied intensively not only in mathematical statistical field but also in industry, see e.g. *Hocking* (1976), *Hoerl* (1985), *Hanke and Hansen* (1993) und *Engl* (1993). In most applications of *Tykhonov-Phillips* type of regularization the weighting factor α is determined by simulation studies, but according to the literature also optimization techniques have been applied. Here we aim at an objective method to determine the *weighting factor* α within α-HAPS.

Alternatively, improperly posed problems, which appear in solving integral equations of the first kind or downward continuation problems in potential theory, depart from observations which are elements of a probability space. Accordingly, estimation techniques of type BLUUE (best linear uniformly unbiased estimation) have been implemented to estimate $\hat{\boldsymbol{\xi}}$ as an unknown parameter vector $\boldsymbol{\xi}$ ("fixed effects") within a linear *Gauss-Markov model*. Such an estimation is *not* robust against *outliers* in the stochastic observation vector $\mathbf{y} \in \mathbb{Y}$.

The second method of regularizing an improperly posed problem offers the possibility to determine the regularization parameter α in an optimal way. For instance, by an A-optimal design of type

> "minimize the trace of the *Mean Square*
> *Error matrix* $\operatorname{tr}\mathrm{MSE}\{\hat{\boldsymbol{\xi}}\}$ of $\hat{\boldsymbol{\xi}}$ (α-hom BLE) to find
> $\hat{\alpha} = \arg\{\operatorname{tr}\mathrm{MSE}\{\hat{\boldsymbol{\xi}}\} = \min\}$"

we are able to construct the regularization parameter α which *balances the trace of the variance-covariance matrix* $\operatorname{tr}\mathrm{D}\{\hat{\boldsymbol{\xi}}\}$ *and the trace of the quadratic bias* $\operatorname{tr}\boldsymbol{\beta}\boldsymbol{\beta}'$ for the bias vector $\boldsymbol{\beta} = -[\mathbf{I} - \mathbf{L}\mathbf{A}]\boldsymbol{\xi}$.

According to the facts and status introduced above this dissertation presents the complete statistical analysis of random deformation tensor (case study: two- and three-dimensional, symmetric rank two strain rate tensor) with emphasis on their eigenspace components. The main contributions of this study are:

- Determination of the sampling distribution of the three-dimensional deformation tensor and development of the univariate and multivariate hypothesis tests, especially the *eigen-inference* and test with a *Growth Curve model*;
- Derivation of the general sampling distribution of the estimate within a Gauss-Markov linear model;
- Derivation of the regularization parameter in uniform Tykhonov-Phillips regularization (α-weighted BLE) by minimizing the trace of the *Mean Square Error matrix MSE*$\{\hat{\xi}\}$ (*A-optimal design*) in the general case for the Gauss-Markov model;
- Linearization of the special nonlinear multivariate *Gauss-Markov model* related the tensor elements and the eigenspace components;
- Development of the BLUUE of the eigenspace elements of two-dimensional random tensor and BIQUUE of its variance-covariance matrix for the linearized model;
- Establishment of the unique *eigenvalue-eigenvector analysis and synthesis* of a three-dimensional symmetric random matrix based on the review and choice of orthogonal similarity transformation matrices, which leads to the generalization of the BLUUE of the eigenspace elements of three-dimensional random tensor and BIQUUE of its variance-covariance matrix in the three-dimensional case.
- The theorems and estimators are in closed form and practical which bring a sound meaning to the statistical analysis of deformation tensor.

In this doctoral thesis the following topics will be presented in detail:

Chapter 1 first discusses the normal distribution property of a three-dimensional, symmetric random tensor. Further it will derive the sampling distribution of the sample mean and sample variance with classical methods. Section 1.3 deals with a matrix method of deriving the sampling distribution directly from the probability density function for the sample mean from the multivariate normal population of a three-dimensional, symmetric rank two random tensor. Based on the *Wishart distribution* the sampling distribution connected with sample variance-covariance of symmetric random tensor is derived and the independence between sample mean and sample variance-covariance is studied in Section 1.4. As a generalization of the sampling distribution theory in the direct observation case for a scalar or vectorized random tensor, Section 1.5 will develop the sampling distribution of the estimate of the linear *Gauss-Markov* model and the sampling distribution of the orthonormal transformed parameters.

Chapter 2 develops the testing hypotheses concerning the sample mean vector and the sample variance covariance matrix, i.e. the estimated parameters (mean vector and covariance matrix) of tensor-valued multivariate normal population of a three-dimensional, symmetric rank-two random tensor, which are (1) Tests on $\boldsymbol{\mu}$ with $\boldsymbol{\Sigma}$ known (χ^2-test); (2) Tests on $\boldsymbol{\mu}$ with $\boldsymbol{\Sigma}$ unknown (*Hotelling's* T^2-test); (3) Test on equality of two mean vectors with common variance-covariance matrix (*Hotelling's two-sample* T^2 test and *Wilks'* Λ test); (4) Test on variance-covariance matrix is equal to a given matrix (*likelihood ratio statistics*); (5) Test on the equality of two variance-covariance matrices (*likelihood ratio statistics*); (6) Tests on the mean vectors and variance-covariance matrices are equal to a given vector and matrix (*likelihood ratio statistics*).

Chapter 3 develops the optimal α for Tykhonov-Phillips regularization by A-optimal design. In Section 3.1 the regularization parameter in uniform Tykhonov-Phillips regularization (α-weighted BLE) is determined by minimizing the trace of the *Mean Square Error matrix MSE*$\{\hat{\xi}\}$ (*A-optimal design*) in the general case for the Gauss-Markov model. With two comparisons it is shown that the optimal ridge parameter *k* in *ridge regression* developed by *Hoerl and Kennard* (1970a, 1970b) and *Hoerl, Kennard and Baldwin* (1975) are just the special case of our general solution by A-optimal design. Based on the introduction of the multivariate $\alpha-$homBLE for the multivariate parameters, the determination of the optimal weight factor α is generalized to the multivariate Gauss-Markov model, which we shall call "*multivariate ridge estimator*". In lieu of two case studies, these models are tested and analyzed with numerical results computed from simulated direct observations of a random tensor of type strain rate in univariate and multivariate cases.

Chapter 4 deals with statistical inference of the eigenspace components of a two-dimensional, symmetric rank-two random tensor. First, the *eigenspace analysis and synthesis* of a symmetric random matrix are reviewed.

Second, the nonlinear function, which relates the tensor elements to the eigenspace components, is linearized with respect to a *special nonlinear multivariate Gauss-Markov model*. *Third*, for its linearized form, *BLUUE* of the eigenspace components and *BIQUUE* of its variance-covariance matrix have been established successfully. *Fourth*, the *sampling distribution* of eigenspace components is derived. The test statistics, such as *Hotelling's* T^2, *likelihood ratio statistics and* the general linear hypothesis test with *growth curve model*, are proposed. Hypothesis tests for the random tensor sample means as well as its one variance component will be used in the case study of validating a given random strain rate tensor in Chapter 6.

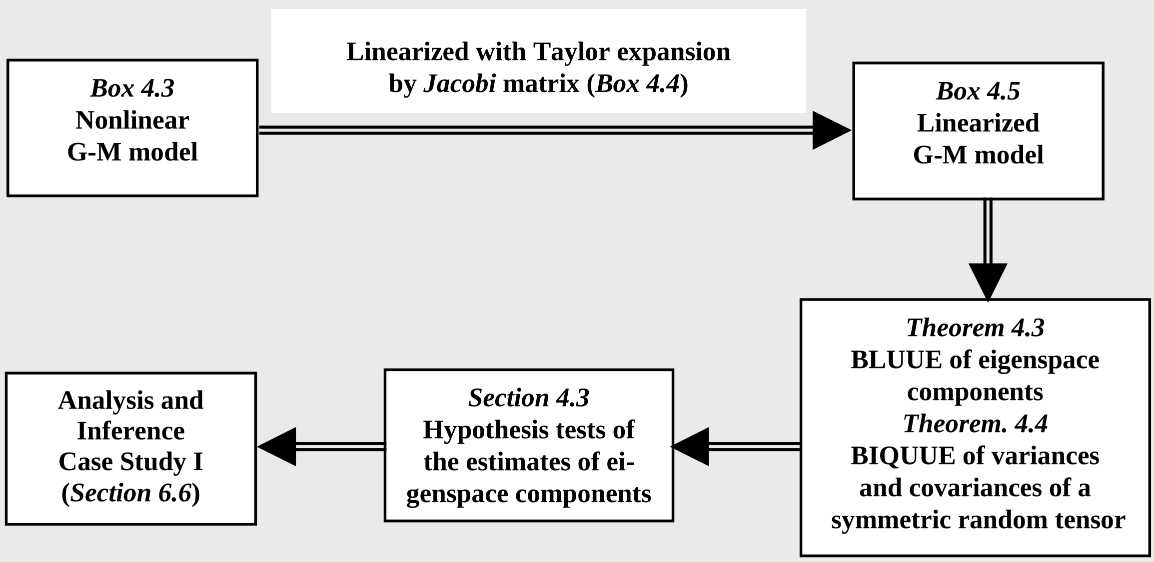
Chapter 5 deals with statistical inference of the eigenspace components of a three-dimensional, symmetric rank-two random tensor. *First*, based on the review and choice of orthogonal similarity transformation matrices the *eigenspace analysis and synthesis* of a three-dimensional symmetric random matrix are established uniquely. *Second*, the nonlinear function that relates the tensor elements to the eigenspace components is linearized with respect to a *special nonlinear multivariate Gauss-Markov model*, which enables the *BLUUE of the eigenspace elements* and *BIQUUE* of its variance-covariance matrix, developed in *Section* 4.2 to be successfully applied in the three-dimensional case. Third, the test statistics, such as *Hotelling's* T^2 and *likelihood ratio statistics,* are generated. Hypothesis tests for the random tensor sample means as well as its one variance component will be used in the case study of validating a given three-dimensional random strain rate tensor in Chapter 6.

Chapter 6 begins with a discussion of the geodynamic setting of the Earth and especially the selected investigated regions: the central Mediterranean and Western Europe. Then the space geodetic observations are introduced. Thirdly the selection of ITRF sites is performed after the history and quality of the ITRF realization series and the related residual velocities of selected ITRF sites are computed. Further the methods of derivation, the two- and three-dimensional geodetic strain rates are introduced and applied to derive the strain rates from the residual velocities, which are based on the *Finite-Element-Method* (FEM). For two case studies both BLUUE and BIQUUE models and hypothesis tests are applied to the eigenspace components of two- and three-dimensional strain rate tensor observations in the area of central Mediterranean and Western Europe, which are derived from ITRF92 to ITRF2000 series station positions and velocities in Sections 6.6. and 6.7. Further detailed analysis of the results is also performed with respect to geodynamical and statistical aspects.

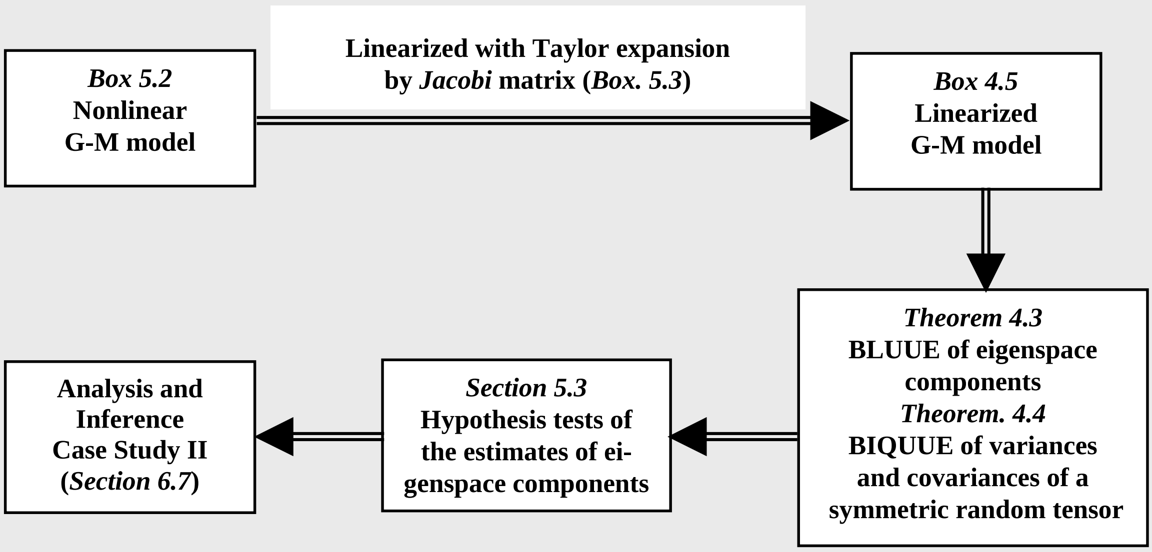
Chapter 7 concludes the main contributions and results in this study and makes a prospect for further applications of the developed theory and methods.

At last we summarize the statistical inference and analysis of two- and three-dimensional, symmetric rank two deformation tensors as developed in Chapter 4, 5 and 6 in the following two schemas :

I. The scheme of inference and analysis of the eigenspace components of a two-dimensional random tensor



II. The scheme of inference and analysis of the eigenspace components of a three-dimensional random tensor



Chapter 1

Sampling distributions of three-dimensional, symmetric rank-two random tensor and the estimate of Gauss-Markov model

In order to make the quality of the estimated random tensors significant, statistical inference has to be applied, which is usually based on the statistics, e.g., sample means and sample variance. So we should derive the sampling distributions of the symmetric rank-two random tensor. The values of any sample statistic depend on the particular sample that one happens to obtain. It varies from sample to sample. Thus a statistic is a random variable. As such, it has a probability distribution called sampling distribution. We owe the early development of sampling distributions under normality to *P.S. Laplace* (1812), *Carl Friedrich Gauss* (1816), *Friedrich Robert Helmert* (1876) for the *Helmert distribution* (which is highly valued as the starting point for modern small sample theory), *Thorvald N. Thiele* (1889,1903), *Karl Pearson* (1900) for his *Chi-square distribution*, *Sealy Gosset* (1908a, b) for his *Student t-distribution, Ronald A. Fisher* (1920, 1922) for the *F-distribution* and *John Wishart* (1928) for the *Wishart distribution*.

In the following sections we will first discuss the normal distribution property of a three-dimensional, symmetric random tensor. Further we will introduce the derivation of the sampling distribution of the sample mean and sample variance with classic methods. Section 1.3 deals with a matrix method of deriving the sampling distribution directly from the probability density function for the sample mean from the multivariate normal population of a three-dimensional, symmetric rank-two random tensor. Based on the *Wishart distribution* the sampling distribution connected with the sample variance-covariance matrix of a symmetric random tensor is derived and the independence between sample mean and sample variance-covariance is studied in Section 1.4. As a generalization of the sampling distribution theory in the direct observation case for a scalar or vectorized random tensor, section 1.5 will develop the sampling distribution of the estimate of the linear *Gauss-Markov* model and the sampling distribution of the orthonormal transformed parameters.

1.1 The normal distribution of a symmetric random tensor

A tensor is a mathematical quantity that can be used to describe the state or the physical properties of a material. We describe a tensor by a set of scalar components referred to a particular coordinate system. A rank-two tensor in three-dimensional space has nine components, the most important examples of these in geophysics are stress, strain and strain rate. Rank-two tensors are used to describe physical quantities that have magnitudes and are associated with three directions. Any rank-two tensor can be defined as a sum of symmetric tensors and an anti-symmetric tensor. Here we will concern ourselves with the statistical properties of the symmetric tensor.

Since basic quantities to infer the stress tensor and strain tensor in Earth sciences are contaminated by random errors, the tensor will be random. Before we discuss the statistical properties of a symmetric random tensor, we will first present the definitions and properties of both random vector and matrix and the multivariate normal distribution of the random vector and matrix.

In multivariate analysis, each observation consists of a vector or matrix. The elements of a random vector or a random matrix are random variables. Formally, a random variable is a function defined for each element of a sample space. We shall generally define a random vector and its moments.

Definition 1.1 (random vector)

A *random* $n\times 1$ *vector* $\mathbf{x}$ is a vector

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

of random variables $x_1, x_2, ..., x_n$, which are jointly distributed.

Definition 1.2 (the first order moment - the mean or expectation)

The first order moment of a random $n \times 1$ vector $\mathbf{x}$ is defined to be the vector of expectations

$$E\{\mathbf{x}\} = \begin{bmatrix} E\{x_1\} \\ \vdots \\ E\{x_n\} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix} := \boldsymbol{\mu} \tag{1.1}$$

More generally, if $\mathbf{Z}=(z_{ij})$ is a $p \times q$ random matrix then E{$\mathbf{Z}$}, the expectation of $\mathbf{Z}$ is the matrix whose *i-j*th element is E{z_{ij}}.

Definition 1.3 (the centralized second order moment – the dispersion matrix, also called variance-covariance matrix)

The centralized second order moment of a random $n \times 1$vector $\mathbf{x}$ is defined to be the $n \times n$ matrix

$$\begin{aligned} D\{\mathbf{x}\} = \boldsymbol{\Sigma}_{\mathbf{x}} &=: E\{[\mathbf{x} - E\{\mathbf{x}\}][\mathbf{x} - E\{\mathbf{x}\}]'\} \\ &= E\{(\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})'\}, \end{aligned} \tag{1.2}$$

the *i*, *j*th off-diagonal element of $\boldsymbol{\Sigma}_{\mathbf{x}}$ is

$$\sigma_{ij} = E\{(x_i - \mu_i)(x_j - \mu_j)\},$$

the covariance between x_i and x_j and the *i*th diagonal element of $\boldsymbol{\Sigma}_{\mathbf{x}}$ is

$$\sigma_{ii} = E\{(x_i - \mu_i)^2\},$$

the variance of x_i. It is proved that $\boldsymbol{\Sigma}_{\mathbf{x}}$ is positive-definite.

The majority of multivariate inferential procedures is based on the assumption that the random vector of interest has a multivariate normal distribution, which is the direct generalization of the univariate normal distribution. Before developing the multivariate normal density function and its properties, we will first review the univariate normal distribution.

A normally distributed random variable *x* with mean μ and variance σ^2 is defined as

Definition 1.4 (univariate normal distribution)

A random variable *x* with mean μ and variance σ^2 is said to have a univariate normal distribution, in symbols $x \sim \mathcal{N}(\mu, \sigma^2)$, if the probability density function of *x* is of the form

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}, \ x \in \mathbb{R}. \tag{1.3}$$

The standardized variable $z = (x - \mu)/\sigma$ with mean 0 and variance 1 is said to have a standard normal distribution with the density

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \ z \in \mathbb{R}. \tag{1.4}$$

The multivariate normal distribution of the random vector $\mathbf{x} = [x_1, x_2, ..., x_n]'$ can be generalized by the univariate normal distribution (1.3) of one random variable as presented in *Definition* 1.5.

Definition 1.5 (multivariate normal distribution)

The $n \times 1$ random vector $\mathbf{x}$ with mean $\boldsymbol{\mu}$ and variance covariance matrix $\boldsymbol{\Sigma}_\mathbf{x}$ is said to have a nonsingular multivariate normal distribution, in symbols

$$\mathbf{x} \sim \mathcal{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}_\mathbf{x}),\ \boldsymbol{\Sigma}_\mathbf{x} > 0,$$

if (i) $\boldsymbol{\Sigma}_\mathbf{x}$ is positive-definite, and (ii) the probability density function of $\mathbf{x}$ is of the form

$$f(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}_\mathbf{x}) = \frac{1}{(2\pi)^{n/2}(\det \boldsymbol{\Sigma}_\mathbf{x})^{1/2}} \exp\{-\tfrac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}_\mathbf{x}^{-1}((\mathbf{x}-\boldsymbol{\mu})\},\ \mathbf{x} \in \mathbb{R}^n \quad (1.5)$$

where $\mathbf{x} = [x_1, x_2, ..., x_n]'$ and $E\{\mathbf{x}\} = \boldsymbol{\mu}$ is the first order moment (1.1) and $D\{\mathbf{x}\} = \boldsymbol{\Sigma}_\mathbf{x} = E\{[\mathbf{x}-E\{\mathbf{x}\}][\mathbf{x}-E\{\mathbf{x}\}]'\}$ is the centralized second order moment (1.2).

The matrix normal distribution is also important in order to express the multivariate normal distribution. *Dawid* (1981), *Mardia* (1979, 1993), *Muirhead* (1982), *Rosen* (1988) and *Brown* (1993) published different expressions, however we prefer that of *Muirhead* (1982).

We write that an $r \times s$ random matrix $\mathbf{Y}$ is normally distributed, say $\mathbf{Y}$ is $\mathbf{Y} \sim \mathcal{N}(\mathbf{M}, \mathbf{C} \otimes \mathbf{D})$ where $E\{\mathbf{Y}\}=\mathbf{M}$ is $r \times s$ mean value matrix, $\mathbf{C}$ and $\mathbf{D}$ are $r \times r$ and $s \times s$ positive-definite matrices and $\mathbf{C} \otimes \mathbf{D}$ is the variance covariance matrix of the vector $\mathbf{y}$=vec($\mathbf{Y}$). The statement "$\mathbf{Y}$ is $\mathbf{Y} \sim \mathcal{N}(\mathbf{M}, \mathbf{C} \otimes \mathbf{D})$ " is equivalent to the statement that " $\mathbf{y}$ is $\mathbf{y} \sim \mathcal{N}_{rs}(\mathbf{m}, \mathbf{C} \otimes \mathbf{D})$," with $\mathbf{m}$=vec($\mathbf{M}$). The following result gives the joint density function of the elements of $\mathbf{Y}$, which we name the matrix normal distribution.

Theorem 1.6 (multivariate matrix normal distribution)

The $r \times s$ random matrix $\mathbf{Y}$ with mean matrix $\mathbf{M}$ and variance covariance matrix $\mathbf{C} \otimes \mathbf{D}$ of the vector of $\mathbf{y}$=vec($\mathbf{Y}$) is said to have a multivariate matrix normal distribution, in symbols,

$$\mathbf{Y} \sim \mathcal{N}(\mathbf{M}, \mathbf{C} \otimes \mathbf{D})$$

if (i) $\mathbf{C}$ and $\mathbf{D}$ are $r \times r$ and $s \times s$ positive-definite matrices and (ii) the probability density function of $\mathbf{Y}$ is of the form

$$f(\mathbf{Y}) = (2\pi)^{-rs/2}(\det \mathbf{C})^{-s/2}(\det \mathbf{D})^{-r/2} \operatorname{etr}\{-\tfrac{1}{2}\mathbf{C}^{-1}(\mathbf{Y}-\mathbf{M})\mathbf{D}^{-1}(\mathbf{Y}-\mathbf{M})'\},\ \mathbf{Y} \in \mathbb{R}^{r \times s} \quad (1.6)$$

where etr(Z):= exp{ tr Z}.

Proof:

Since $\mathbf{y}$=vec($\mathbf{Y}$) is $\mathbf{y} \sim \mathcal{N}_{rs}(\mathbf{m}, \mathbf{C} \otimes \mathbf{D})$, with $\mathbf{m}$=vec($\mathbf{M}$), from (1.5) the joint density function of the element of $\mathbf{y}$ is

$$f(\mathbf{y}) = (2\pi)^{-rs/2}(\det(\mathbf{C} \otimes \mathbf{D}))^{-1/2} \exp\{-\tfrac{1}{2}(\mathbf{y}-\mathbf{m})'(\mathbf{C} \otimes \mathbf{D})^{-1}(\mathbf{y}-\mathbf{m})\},\ \mathbf{y} \in \mathbb{R}^{rs \times 1} \quad (1.7)$$

According to the properties of Kronecker products

$$\det(\mathbf{C} \otimes \mathbf{D}) = (\det \mathbf{C})^s(\det \mathbf{D})^r,\ \text{if } \mathbf{C} \text{ is } r \times r,\ \mathbf{D} \text{ is } s \times s.$$

$$(\mathbf{C} \otimes \mathbf{D})^{-1} = \mathbf{C}^{-1} \otimes \mathbf{D}^{-1},\ \text{if } \mathbf{C} \text{ and } \mathbf{D} \text{ are nonsingular.}$$

$$\operatorname{tr}(\mathbf{PX}'\mathbf{QXR}) = (\operatorname{vec}(\mathbf{X}))'(\mathbf{RP} \otimes \mathbf{Q}')\operatorname{vec}(\mathbf{X}),$$

and with respect to the corresponding items $\mathbf{P} = \mathbf{C}^{-1}$, $\mathbf{X}' = \mathbf{Y} - \mathbf{M}$, $\mathbf{Q} = \mathbf{D}^{-1}$ and $\mathbf{R} = \mathbf{I}$ we can see that (1.7) is the same as (1.6). This completes the proof.

With these definitions and the theorem we will now establish the distribution of the symmetric random tensor.

Let there be given a three-dimensional, symmetric rank-two random tensor $\mathbf{T}$ which is either directly or indirectly estimated from observations by a model adjustment. The components of $\mathbf{T}$ can be expressed in a matrix

$$\mathbf{T} = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{bmatrix}. \quad (1.8)$$

Such a three-dimensional, symmetric (2, 0) tensor is a random tensor **T** which we assume to be an element of the tensor-valued Gauss normal distribution over $\mathbb{R}^{3\times3}$ of type independently, identically distributed (i.i.d.) tensor-valued observations. The probability distribution of this random tensor can be presented in matrix normal form of *theorem* 1.6. In order to derive the sampling distribution of the sample mean and sample variance covariance, noting that $t_{21} = t_{12}$, $t_{31} = t_{13}$ and $t_{32} = t_{23}$ we can make a simplification: the symmetric random tensor (1.8) is vectorized by (vech, vector-half)

$$\mathbf{y} = \text{vech}\,\mathbf{T} = [t_{11}\ t_{12}\ t_{13}\ t_{22}\ t_{23}\ t_{33}]',\ \ \mathbf{y} \in \mathbb{R}^{6\times1} \tag{1.9}$$

According to the *Definition 1.5* we can get directly the joint multivariate normal probability density function (p.d.f.) of the three-dimensional, symmetric rank-two random tensor, which is presented in *Definition* 1.7

Definition 1.7 (normal distribution of a symmetric random tensor)

The vectorized random tensor $\mathbf{y} = \text{vech}\,\mathbf{T} = [t_{11}\ t_{12}\ t_{13}\ t_{22}\ t_{23}\ t_{33}]'$, $\mathbf{y} \in \mathbb{R}^{6\times1}$ with mean vector $\boldsymbol{\mu}$ and variance covariance $\boldsymbol{\Sigma}_{\mathbf{y}}$ is said to have a nonsingular multivariate normal distribution, in symbols

$$\mathbf{y} \sim \mathcal{N}_6(\boldsymbol{\mu}, \boldsymbol{\Sigma}_{\mathbf{y}})$$

if (i) $\boldsymbol{\Sigma}_{\mathbf{y}}$ is positive-definite, and (ii) the probability density function of **y** is of the form

$$f(\mathbf{y}; \boldsymbol{\mu}, \boldsymbol{\Sigma}_{\mathbf{y}}) = (2\pi)^{-6/2}(\det\boldsymbol{\Sigma}_{\mathbf{y}})^{-1/2}\exp\{-\frac{1}{2}[\mathbf{y}-\boldsymbol{\mu}]'\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}[\mathbf{y}-\boldsymbol{\mu}]\}, \tag{1.10}$$

where $E\{\mathbf{y}\} = \boldsymbol{\mu}$ is the first order moment – the mean value vector, $D\{\mathbf{y}\} = \boldsymbol{\Sigma}_{\mathbf{y}} = E\{[\mathbf{y}-E\{\mathbf{y}\}][\mathbf{y}-E\{\mathbf{y}\}]'\}$ is the centralized second order moment- the dispersion matrix, also called variance-covariance matrix

1.2 The sampling distribution of sample mean and sample variance of scalar

In this section we will derive the sampling distribution of the sample mean and sample variance of random scalar with the classical method, which is a direct derivation from its distribution density function. Let us first introduce one *lemma* about the sampling distribution of the statistics of a random scalar

Lemma 1.8 (i.i.d. observation of type Gauss normal, distribution of the sample mean and sample variance)

Let $(y_1, y_2, \ldots, y_n) \in \mathbb{Y}$ be a set of observations, $\mathbf{y} := [y_1, y_2, \ldots, y_n]'$, $\dim\mathbb{Y} = n$, a vector-valued independent, identically distributed (i.i.d.) random variable from a Gauss normal distribution. Its moment of first order as well as its central moments of second order are specified by $\mu := \mu_1 = \mu_2 = \cdots = \mu_n$ and $\sigma^2 := \sigma^2_1 = \sigma^2_2 = \cdots = \sigma^2_n$. Then the sample mean $\hat{\mu}$ of type BLUUE

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{n}\mathbf{1}'\mathbf{y} = \frac{1}{n}\mathbf{y}'\mathbf{1} \tag{1.11}$$

is an element of a specific Gauss normal distribution of type

$$\hat{\mu} \sim \mathcal{N}(\mu, \frac{1}{n}\sigma^2)$$

with the sample statistic $\hat{\sigma}^2$ of type BIQUUE

$$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \hat{\mu})^2 = \frac{1}{n-1}(\mathbf{y}-\mathbf{1}\hat{\mu})'(\mathbf{y}-\mathbf{1}\hat{\mu}). \tag{1.12}$$

The random variable u has a *Chi-square* distribution with n-1 degrees of freedom

$$u = \frac{(n-1)\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-1). \tag{1.13}$$

Before proving this *lemma*, we should introduce the definition of chi-square distribution, which was first found by *Helmert* (1876) and *K. Pearson* (1900, 1931) and plays a very important role in sampling theory.

Definition 1.9 (central *chi-square distribution*)

A random variable x is said to have a ***chi-square distribution,*** and it is referred to as a chi-square random variable, if and only if its probability density function of x is of the form

$$f(x) = \begin{cases} \dfrac{1}{2^{\upsilon/2}\Gamma(\upsilon/2)} x^{(\upsilon-2)/2} e^{-x/2} & \text{for } x > 0 \\ 0 & \text{elsewhere} \end{cases} \quad (1.14)$$

where the parameter υ is referred to as the degrees of freedom, which is a positive integer. Therefore the x is also said to have a chi-square distribution with υ degrees of freedom.

Proof:

The probability density function (p.d.f.) of the i.i.d. Gauss observations is

$$f(y_1,\cdots,y_n) = f(y_1)\cdots f(y_n) = (2\pi)^{-n/2}\sigma^{-n}\exp\{-\frac{1}{2}(\mathbf{y}-\mathbf{1}\mu)'(\mathbf{y}-\mathbf{1}\mu)/\sigma^2\}\,. \quad (1.15)$$

We shall find the p.d.f. of $\hat{\mu}$ and $\hat{\sigma}^2$.

We have

$$\begin{aligned}(\mathbf{y}-\mathbf{1}\mu)'(\mathbf{y}-\mathbf{1}\mu) &= \sum_{i=1}^{n}(y_i-\mu)^2 = \sum_{i=1}^{n}[(y_i-\hat{\mu})-(\mu-\hat{\mu})]^2 = \\ &= (n-1)\hat{\sigma}^2 + n(\hat{\mu}-\mu) + 2(\hat{\mu}-\mu)(\sum_{i=1}^{n} y_i - n\hat{\mu}) = \\ &= (n-1)\hat{\sigma}^2 + n(\hat{\mu}-\mu)\end{aligned} \quad (1.16)$$

So with (1.15) and (1.16) we get

$$f(\mathbf{y})dy_1\cdots dy_n = (2\pi)^{-n/2}(\sigma^2)^{-n/2}\exp\{-\frac{1}{2}n\frac{\hat{\mu}-\mu}{\sigma^2}\}\cdot\exp\{-\frac{1}{2}(n-1)\hat{\sigma}^2/\sigma^2\}dy_1\cdots dy_n \quad (1.17)$$

We now perform the *Helmert transformation* (*Helmert* 1876)

$$\mathbf{X} := \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1}\end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{1\cdot 2}} & -\frac{1}{\sqrt{1\cdot 2}} & 0 & \cdots & 0 \\ \frac{1}{\sqrt{2\cdot 3}} & \frac{1}{\sqrt{2\cdot 3}} & -\frac{2}{\sqrt{2\cdot 3}} & \cdots & \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{1}{\sqrt{(n-1)n}} & \frac{1}{\sqrt{(n-1)n}} & \frac{1}{\sqrt{(n-1)n}} & \cdots & -\frac{n-1}{\sqrt{(n-1)n}} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} := \mathbf{U}\mathbf{y}\,, \quad (1.18)$$

where $\mathbf{U} \in \mathbb{R}^{(n-1)\times n}$ is a right orthogonal matrix, i.e., $\mathbf{U}\mathbf{U}' = \mathbf{I}_{n-1}$.

The corresponding volume element transformation with $\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} y_i$ is

$$dy_1\cdots dy_n = \mathbf{J}\, d\hat{\mu}\, dx_1\cdots dx_{n-1}\,, \quad (1.19)$$

and its inverse

$$dx_1\cdots dx_{n-1}d\hat{\mu} = \mathbf{J}^*\, dy_1\cdots dy_n = \begin{vmatrix} & \mathbf{U} & & \\ \hline \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n}\end{vmatrix}\, dy_1\cdots dy_n =$$

$$= \frac{1}{\sqrt{n}}\begin{vmatrix} & \mathbf{U} & & \\ \hline \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \cdots & \frac{1}{\sqrt{n}}\end{vmatrix} dy_1\cdots dy_n := |\mathbf{U}_2|\; dy_1\cdots dy_n\quad . \quad (1.20)$$

Since $\mathbf{U}_2$ is a $n \times n$ orthogonal matrix, i.e., det $\mathbf{U}_2 = \pm 1$, we get

$$\mathbf{J}^* = \frac{1}{\sqrt{n}} \left| \det \mathbf{U}_2 \right| = \frac{1}{\sqrt{n}} \tag{1.21}$$

and so

$$\mathbf{J} = (\mathbf{J}^*)^{-1} = \sqrt{n}\ . \tag{1.22}$$

From the cumulative distribution function (c.d.f.) of $\hat{\mu},\ x_1,\ \cdots,\ x_{n-1},\ F(\hat{\mu},\ x_1,\ \cdots,\ x_{n-1})$ we have

$$dF(\hat{\mu},\ x_1,\ \cdots,\ x_{n-1}) = f(\hat{\mu},\ x_1,\ \cdots,\ x_{n-1})\, d\hat{\mu}\ dx_1 \cdots dx_{n-1} \tag{1.23}$$

and together with (1.16) we obtain

$$f(\mathbf{y})dy_1 \cdots dy_n = (2\pi)^{-n/2} \sigma^{-n} \exp[-\frac{1}{2} n \frac{(\hat{\mu}-\mu)^2}{\sigma^2}] \exp[-\frac{1}{2}(n-1)\hat{\sigma}^2 / \sigma^2] \sqrt{n}\, d\hat{\mu}\, dx_1 \cdots dx_{n-1}\,. \tag{1.24}$$

We further perform the following transformation (*Cramer*, 1945)

$$x_i = \sqrt{n}\, s\, z_i\,, \quad i=1,2,\ \ldots,\ n\text{-}1, \tag{1.25}$$

where $s^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{\mu})^2$, to replace the n-1 variables x_i by n new variables s and $z_1, \ldots, z_{n-1}$.

Accordingly, there is a relation among the new variables, which is found by squaring and adding the n-1 equations (1.25). We then obtain

$$\sum_{i=1}^{n-1} z_i^{\ 2} = 1\,. \tag{1.26}$$

and thus one of z_i , say z_{n-1}, may be expressed as a function of the n-2 others, so that the old variables $x_1,\ldots, x_{n-1}$ are replaced by the new variables s and $z_1, \ldots, z_{n-2}$. For the Jacobian $\mathbf{J}_z$ of the transformation we have since $\partial z_{n-1} / \partial z_i = -z_i / z_{n-1}$,

$$\mathbf{J}_z = \begin{vmatrix} \sqrt{n}s & 0 & \cdots & 0 & \sqrt{n}z_1 \\ 0 & \sqrt{n}s & \cdots & 0 & \sqrt{n}z_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & \sqrt{n}s & \sqrt{n}z_{n-2} \\ -\sqrt{n}s\frac{z_1}{z_{n-1}} & \cdots & \cdots & -\sqrt{n}s\frac{z_1}{z_{n-1}} & \sqrt{n}z_{n-1} \end{vmatrix} = \frac{n^{(n-1)/2}s^{n-2}}{z_{n-1}} \begin{vmatrix} 1 & 0 & \cdots & 0 & z_1 \\ 0 & 1 & \cdots & 0 & z_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & z_{n-2} \\ -z_1 & -z_2 & \cdots & -z_{n-1} & z_{n-1}^{\ 2} \end{vmatrix} \tag{1.27}$$

$$= \pm \frac{n^{(n-1)/2}s^{n-2}}{\sqrt{1-z_1^2 - \cdots\ - z_{n-2}^2}} \times 1\ \ .$$

Thus we obtain the expression

$$\begin{aligned} f(\mathbf{y})dy_1 \cdots dy_n &= 2 \times (2\pi)^{-n/2} \sigma^{-n} \exp[-\frac{1}{2} n \frac{(\hat{\mu}-\mu)^2}{\sigma^2}] \exp[-\frac{1}{2}(n-1)\hat{\sigma}^2 / \sigma^2] \times \\ &\times n^{1/2} \frac{n^{(n-1)/2}s^{n-2}}{\sqrt{1-z_1^2 - \cdots - z_{n-2}^2}} d\hat{\mu}\, ds\, dz_1 \cdots dz_{n-2}. \end{aligned} \tag{1.28}$$

With the relationship $s^2 = \frac{n-1}{n}\hat{\sigma}^2$, the Jacobian $\mathbf{J}_{\hat{\sigma}}$ of the transformation from the element s to element $\hat{\sigma}^2$ is

$$\mathbf{J}_{\hat{\sigma}} = \frac{1}{2s} \frac{n-1}{n} \quad . \tag{1.29}$$

The right side of (1.28) will be

$$2(2\pi)^{-n/2}\sigma^{-n}\exp[-\frac{1}{2}n\frac{(\hat{\mu}-\mu)^2}{\sigma^2}]\exp[-\frac{1}{2}(n-1)\hat{\sigma}^2/\sigma^2]n^{1/2}\frac{n^{(n-1)/2}s^{n-2}}{\sqrt{1-z_1^2-\cdots-z_{n-2}^2}}\frac{1}{2s}\frac{n-1}{n}d\hat{\mu}\,d\hat{\sigma}^2dz_1\cdots dz_{n-2}=$$
$$=(2\pi)^{-n/2}\sigma^{-n}n^{1/2}n^{(n-1)/2}(\frac{n-1}{n}\hat{\sigma}^2)^{(n-3)/2}\frac{n-1}{n}\exp[-\frac{1}{2}n\frac{(\hat{\mu}-\mu)^2}{\sigma^2}]\exp[-\frac{1}{2}(n-1)\hat{\sigma}^2/\sigma^2]\frac{d\hat{\mu}\,d\hat{\sigma}^2dz_1\cdots dz_{n-2}}{\sqrt{1-z_1^2-\cdots-z_{n-2}^2}}=$$
$$=\frac{1}{\sqrt{2\pi}\sigma}\sqrt{n}\exp[-\frac{1}{2}n\frac{(\hat{\mu}-\mu)^2}{\sigma^2}]d\hat{\mu}\,(\frac{n-1}{2\sigma^2})^{(n-1)/2}(\hat{\sigma}^2)^{(n-3)/2}\exp[-\frac{1}{2}(n-1)\hat{\sigma}^2/\sigma^2]d\hat{\sigma}^2\times$$
$$\times(\pi)^{-(n-1)/2}\frac{dz_1\cdots dz_{n-2}}{\sqrt{1-z_1^2-\cdots-z_{n-2}^2}} \tag{1.30}$$
$$=\frac{1}{\sqrt{2\pi}\sigma}\sqrt{n}\exp[-\frac{1}{2}n\frac{(\hat{\mu}-\mu)^2}{\sigma^2}]d\hat{\mu}\,(\frac{n-1}{2\sigma^2})^{(n-1)/2}\frac{1}{\Gamma(\frac{n-1}{2})}(\hat{\sigma}^2)^{(n-3)/2}\exp[-\frac{1}{2}(n-1)\hat{\sigma}^2/\sigma^2]d\hat{\sigma}^2\times$$
$$\times\frac{\Gamma(\frac{n-1}{2})}{\pi^{(n-1)/2}}\frac{dz_1\cdots dz_{n-2}}{\sqrt{1-z_1^2-\cdots-z_{n-2}^2}}$$

The p.d.f. of (1.15) appears as a product of three factors with the probability elements $\hat{\mu}$, $\hat{\sigma}^2$ and the joint probability elements $z_1,\ldots, z_{n-2}$.

We thus see that $\hat{\mu}$ and $\hat{\sigma}^2$ are independent not only of one another, but also of the combined variable ($z_1,\ldots, z_{n-2}$) and that the distributions of $\hat{\mu}$ and $\hat{\sigma}^2$ are the following:

$$f(\hat{\mu})=\frac{1}{\sqrt{2\pi}\sigma}\sqrt{n}\exp[-\frac{1}{2}n\frac{(\hat{\mu}-\mu)^2}{\sigma^2}] \tag{1.31}$$

$$f(\hat{\sigma}^2)=(\frac{n-1}{2\sigma^2})^{(n-1)/2}\frac{1}{\Gamma(\frac{n-1}{2})}(\hat{\sigma}^2)^{(n-3)/2}\exp[-\frac{1}{2}(n-1)\hat{\sigma}^2/\sigma^2]\,. \tag{1.32}$$

With new variable $u=\frac{n-1}{\sigma^2}\hat{\sigma}^2$ we can get

$$f_u(u)=\frac{1}{2^{(n-1)/2}\Gamma(\frac{n-1}{2})}u^{\frac{n-1}{2}-1}\exp(-\frac{1}{2}u), \tag{1.33}$$

which is the right form of *Chi-square* distribution with (n-1) degree of freedom (1.14).

This completes the proof.

1.3 The sampling distribution of the sample mean of a symmetric random tensor

The sampling distributions of the basic statistics are important for the statistical inference of the symmetric random tensor. In this section we will discuss the sampling distribution of the sample mean of a symmetric random tensor.

Let us use the symmetric random tensor introduced in (1.8), which is vectorized by (vech, vector-half) $\mathbf{y}=\text{vech}\,\mathbf{T}=[t_{11}\ t_{12}\ t_{13}\ t_{22}\ t_{23}\ t_{33}]'$, $\mathbf{y}\in\mathbb{R}^{6\times1}$. This is a random vector which is normally distributed according to *Definition* 1.7. We write it as $\mathbf{y}\sim\mathcal{N}_6(\boldsymbol{\mu},\,\boldsymbol{\Sigma}_\mathbf{y})$.

Suppose that our sample of n observations on $\mathbf{T}$ is $\mathbf{T}_1$, $\mathbf{T}_2,\ldots$, $\mathbf{T}_n$, whose related vectorized forms $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n$ are independently distributed according to $\mathcal{N}_6(\boldsymbol{\mu},\,\boldsymbol{\Sigma}_\mathbf{y})$. We may arrange the vectorized form matrix of observations as

$$\mathbf{Y}=\begin{bmatrix}\mathbf{y}_1'\\ \mathbf{y}_2'\\ \vdots\\ \mathbf{y}_n'\end{bmatrix},\quad \mathbf{Y}\in\mathbb{R}^{n\times6},$$

where $\mathbf{y}_i=\text{vech}\,\mathbf{T}_i$. Then,

$$E\{\mathbf{Y}\}=\begin{bmatrix}\boldsymbol{\mu}_1'\\ \boldsymbol{\mu}_2'\\ \vdots\\ \boldsymbol{\mu}_n'\end{bmatrix}=\mathbf{1}\boldsymbol{\mu}',\quad \text{where}\quad \mathbf{1}=[1,1,\cdots,1]'\in\mathbb{R}^n, \tag{1.34}$$

and for the dispersion we have the transposed form of **Y**

$$\mathbf{Y}' = [\mathbf{y}_1, \ \mathbf{y}_2, \ \cdots \ , \ \mathbf{y}_n],$$

where the columns $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n$ are independent 3×1 random vectors, each with the same covariance matrix $\mathbf{\Sigma}_{\mathrm{y}}$. We then have

$$\operatorname{vec}\mathbf{Y}' = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_n \end{bmatrix}, \quad \operatorname{vec}\mathbf{Y}' \in \ \mathbb{R}^{6n\times 1} \tag{1.35}$$

whose covariance follows that

$$D(\operatorname{vec}\mathbf{Y}') = \begin{bmatrix} \mathbf{\Sigma}_{\mathrm{y}} & 0 & \cdots & 0 \\ 0 & \mathbf{\Sigma}_{\mathrm{y}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \mathbf{\Sigma}_{\mathrm{y}} \end{bmatrix}, \quad D(\operatorname{vec}\mathbf{Y}') \ \in \ \mathbb{R}^{6n\times 6n}$$
$$= \mathbf{I}_n \otimes \mathbf{\Sigma}_{\mathrm{y}} . \tag{1.36}$$

Then the sample mean vector $\hat{\boldsymbol{\mu}}_{\mathrm{y}}$ of type BLUUE is

$$\hat{\boldsymbol{\mu}}_{\mathrm{y}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{y}_i = \frac{1}{n}\mathbf{Y}'\mathbf{1} \tag{1.37}$$

Now we shall find the p.d.f. of $\hat{\boldsymbol{\mu}}_{\mathrm{y}}$.

According to (1.6) of *Theorem 1.6* the joint p.d.f. of the independently identically distributed Gauss sampling observation **Y** can be written as

$$f(\mathbf{Y}) = (2\pi)^{-6n/2}[\det(\mathbf{I}_n \otimes \mathbf{\Sigma}_{\mathrm{y}})]^{-1/2} \operatorname{etr}\{-\frac{1}{2}[\mathbf{Y}-\mathbf{1}\boldsymbol{\mu}']\mathbf{\Sigma}_{\mathrm{y}}^{-1}[\mathbf{Y}-\mathbf{1}\boldsymbol{\mu}']'\}. \tag{1.38}$$

With the properties of Kronecker products introduced in the proof of *Theorem 1.6* we can get

$$[\det(\ \mathbf{I}_n \otimes \mathbf{\Sigma}_{\mathrm{y}})]^{-1/2} = (\det \mathbf{I}_n)^{-6/2}(\det \mathbf{\Sigma}_{\mathrm{y}})^{-n/2},$$

$$\operatorname{tr}(\mathbf{AB}) = \operatorname{tr}(\mathbf{BA}).$$

So we have the reform of (1.38)

$$f(\mathbf{Y}) = (2\pi)^{-6n/2}(\det \mathbf{\Sigma}_{\mathrm{y}})^{-n/2} \operatorname{etr}\{-\frac{1}{2}\mathbf{\Sigma}_{\mathrm{y}}^{-1}[\mathbf{Y}-\mathbf{1}\boldsymbol{\mu}']'[\mathbf{Y}-\mathbf{1}\boldsymbol{\mu}']\}. \tag{1.39}$$

Now we perform the *Helmert transformation* (*Helmert* 1876)

$$\mathbf{X} = \begin{bmatrix} \frac{1}{\sqrt{1\cdot 2}} & -\frac{1}{\sqrt{1\cdot 2}} & 0 & \cdots & 0 \\ \frac{1}{\sqrt{2\cdot 3}} & \frac{1}{\sqrt{2\cdot 3}} & -\frac{2}{\sqrt{2\cdot 3}} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{1}{\sqrt{(n-1)n}} & \frac{1}{\sqrt{(n-1)n}} & \frac{1}{\sqrt{(n-1)n}} & \cdots & -\frac{n-1}{\sqrt{(n-1)n}} \\ \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \cdots & \frac{1}{\sqrt{n}} \end{bmatrix} \begin{bmatrix} \mathbf{y}_1' \\ \mathbf{y}_2' \\ \vdots \\ \mathbf{y}_{n-1}' \\ \mathbf{y}_n' \end{bmatrix} =: \mathbf{HY} \ . \tag{1.40}$$

The $n\times n$ matrix given in (1.40) is not only orthogonal, it also has the property that all the rows sum to zero except for the last (and the last row has common elements). Such an orthogonal matrix is called a *Helmert matrix* (Lancaster, 1965). The Jacobian of this transformation is $\mathbf{J} \ = \ (\det \mathbf{H})^6 = \ 1$. Partition **X** as

$$\mathbf{X} = \begin{bmatrix} \mathbf{Z} \\ \text{--} \\ \mathbf{x}' \end{bmatrix}, \quad \text{where } \mathbf{Z} \in \mathbb{R}^{(n-1)\times 6} \quad \text{and } \mathbf{x} \in \mathbb{R}^{6\times 1}. \tag{1.41}$$

then

$$\mathbf{Y}'\mathbf{Y} = \mathbf{X}'\mathbf{X} = \mathbf{Z}'\mathbf{Z} + \mathbf{x}\mathbf{x}' . \tag{1.42}$$

The term $[\mathbf{Y}-\mathbf{1}\boldsymbol{\mu}']'[\mathbf{Y}-\mathbf{1}\boldsymbol{\mu}']$ of (1.39) can be expanded as

$$\begin{aligned}[\mathbf{Y}-\mathbf{1}\boldsymbol{\mu}']'[\mathbf{Y}-\mathbf{1}\boldsymbol{\mu}'] &= \mathbf{Y}'\mathbf{Y}-\mathbf{Y}'\mathbf{1}\boldsymbol{\mu}'-\boldsymbol{\mu}\mathbf{1}'\mathbf{Y}+n\boldsymbol{\mu}\boldsymbol{\mu}' \\ &= \mathbf{Z}'\mathbf{Z}+\mathbf{x}\mathbf{x}'-\mathbf{Y}'\mathbf{1}\boldsymbol{\mu}'-(\mathbf{Y}'\mathbf{1}\boldsymbol{\mu}')'+n\boldsymbol{\mu}\boldsymbol{\mu}' \quad .\end{aligned} \tag{1.43}$$

Since the first (n -1) rows of **H** are orthogonal to $\mathbf{1}\in\mathbb{R}^n$, i.e.,

$$\mathbf{H}\mathbf{1}=[0,\ \cdots,\ 0,\quad \sqrt{n}]',$$

then

$$\mathbf{Y}'\mathbf{1}\boldsymbol{\mu}'=\mathbf{X}'\mathbf{H}\mathbf{1}\boldsymbol{\mu}'=\left[\mathbf{Z}'\ :\ \mathbf{x}\right]\begin{bmatrix}0\\ \vdots\\ 0\\ \sqrt{n}\end{bmatrix}\boldsymbol{\mu}'=\sqrt{n}\mathbf{x}\boldsymbol{\mu}' \quad . \tag{1.44}$$

Substituting back into (1.43) then gives

$$\begin{aligned}[\mathbf{Y}-\mathbf{1}\boldsymbol{\mu}']'[\mathbf{Y}-\mathbf{1}\boldsymbol{\mu}'] &= \mathbf{Z}'\mathbf{Z}+\mathbf{x}\mathbf{x}'-\sqrt{n}\boldsymbol{\mu}\mathbf{x}'-\sqrt{n}\mathbf{x}\boldsymbol{\mu}'+n\boldsymbol{\mu}\boldsymbol{\mu}' \\ &= \mathbf{Z}'\mathbf{Z}+[\mathbf{x}-\sqrt{n}\boldsymbol{\mu}][\mathbf{x}-\sqrt{n}\boldsymbol{\mu}]'.\end{aligned} \tag{1.45}$$

Hence the joint p.d.f. of **Z** and **x** can be expressed by substituting (1.45) into (1.39)

$$\begin{aligned} f(\mathbf{Y}) = {} & (2\pi)^{-6(n-1)/2}(\det\boldsymbol{\Sigma}_{\mathrm{y}})^{-(n-1)/2}\,\mathrm{etr}\{-\tfrac{1}{2}\boldsymbol{\Sigma}_{\mathrm{y}}^{-1}\mathbf{Z}'\mathbf{Z}\}\times \\ & (2\pi)^{-6/2}(\det\boldsymbol{\Sigma}_{\mathrm{y}})^{-1/2}\exp\{-\tfrac{1}{2}[\mathbf{x}-\sqrt{n}\boldsymbol{\mu}]'\boldsymbol{\Sigma}_{\mathrm{y}}^{-1}[\mathbf{x}-\sqrt{n}\boldsymbol{\mu}]\}.\end{aligned} \tag{1.46}$$

This implies that **Z** is distributed according to $\mathcal{N}_{(n-1),\ 6}(\mathbf{0},\ \mathbf{I}_{n-1}\otimes\boldsymbol{\Sigma}_{\mathrm{y}})$ and independently of **x**, which is distributed according to $\mathcal{N}_6(\sqrt{n}\boldsymbol{\mu},\ \boldsymbol{\Sigma}_{\mathrm{y}})$.

Since $\mathbf{x}'=\dfrac{1}{\sqrt{n}}\mathbf{1}'\mathbf{Y}$, so $\mathbf{x}=\dfrac{1}{\sqrt{n}}\mathbf{Y}'\mathbf{1}$ and with $\hat{\boldsymbol{\mu}}_{\mathrm{y}}=\dfrac{1}{n}\mathbf{Y}'\mathbf{1}$ of (1.37) we get

$$\mathbf{x}=\sqrt{n}\,\hat{\boldsymbol{\mu}}_{\mathrm{y}} \tag{1.47}$$

and the Jacobian of this transformation

$$\mathbf{J}_{\mu}=\sqrt{n}\,. \tag{1.48}$$

So the p.d.f. of $\hat{\boldsymbol{\mu}}_{\mathrm{y}}$, $f(\hat{\boldsymbol{\mu}}_{\mathrm{y}})$, can be derived from the second term of (1.46) and (1.47).

Firstly, the exponential term of (1.46) is expressed by $\hat{\boldsymbol{\mu}}_{\mathrm{y}}$ of (1.47):

$$\begin{aligned}-\tfrac{1}{2}[\mathbf{x}-\sqrt{n}\boldsymbol{\mu}]'\boldsymbol{\Sigma}_{\mathrm{y}}^{-1}[\mathbf{x}-\sqrt{n}\boldsymbol{\mu}] &= -\tfrac{1}{2}[\sqrt{n}\hat{\boldsymbol{\mu}}_{\mathrm{y}}-\sqrt{n}\boldsymbol{\mu}]'\boldsymbol{\Sigma}_{\mathrm{y}}^{-1}[\sqrt{n}\hat{\boldsymbol{\mu}}_{\mathrm{y}}-\sqrt{n}\boldsymbol{\mu}] \\ &= -\tfrac{1}{2}n[\hat{\boldsymbol{\mu}}_{\mathrm{y}}-\boldsymbol{\mu}]'\boldsymbol{\Sigma}_{\mathrm{y}}^{-1}[\hat{\boldsymbol{\mu}}_{\mathrm{y}}-\boldsymbol{\mu}],\end{aligned}$$

then we have

$$\begin{aligned} f(\hat{\boldsymbol{\mu}}_{\mathrm{y}}) &= (2\pi)^{-6/2}(\det\boldsymbol{\Sigma}_{\mathrm{y}})^{-1/2}\exp\{-\tfrac{1}{2}n[\hat{\boldsymbol{\mu}}_{\mathrm{y}}-\boldsymbol{\mu}]'\boldsymbol{\Sigma}_{\mathrm{y}}^{-1}[\hat{\boldsymbol{\mu}}_{\mathrm{y}}-\boldsymbol{\mu}]\}\times|\mathbf{J}_{\mu}| \\ &= (2\pi)^{-6/2}(\det\boldsymbol{\Sigma}_{\mathrm{y}})^{-1/2}\exp\{-\tfrac{1}{2}n[\hat{\boldsymbol{\mu}}_{\mathrm{y}}-\boldsymbol{\mu}]'\boldsymbol{\Sigma}_{\mathrm{y}}^{-1}[\hat{\boldsymbol{\mu}}_{\mathrm{y}}-\boldsymbol{\mu}]\}\times\sqrt{n} \\ &= (2\pi)^{-6/2}\left[\det(n^{-1}\boldsymbol{\Sigma}_{\mathrm{y}})\right]^{-1/2}\exp\{-\tfrac{1}{2}[\hat{\boldsymbol{\mu}}_{\mathrm{y}}-\boldsymbol{\mu}]'(n^{-1}\boldsymbol{\Sigma}_{\mathrm{y}})^{-1}[\hat{\boldsymbol{\mu}}_{\mathrm{y}}-\boldsymbol{\mu}]\}.\end{aligned} \tag{1.49}$$

This shows immediately that the sample mean vector $\hat{\boldsymbol{\mu}}_{\mathrm{y}}$ of the vectorized 3×3 symmetric random tensor is distributed according to $\mathcal{N}_6(\boldsymbol{\mu},\ n^{-1}\boldsymbol{\Sigma}_{\mathrm{y}};\ \ \hat{\boldsymbol{\mu}}_{\mathrm{y}})$.

Now we can characterize these derivations in *Theorem 1.10*.

Theorem 1.10 (the sampling distribution of the sample mean vector of random tensor)

The sampling distribution of the sample mean vector $\hat{\boldsymbol{\mu}}_{\mathrm{y}}$ of the vectorized 3×3 symmetric random tensor

$$\hat{\boldsymbol{\mu}}_{\mathrm{y}}=\frac{1}{n}\sum_{i=1}^{n}\mathbf{y}_i=\frac{1}{n}\mathbf{Y}'\mathbf{1} \tag{1.37}$$

is distributed according to $\mathcal{N}_6(\boldsymbol{\mu},\ n^{-1}\boldsymbol{\Sigma}_{\mathrm{y}};\ \ \hat{\boldsymbol{\mu}}_{\mathrm{y}})$ with the p.d.f.

$$f(\hat{\boldsymbol{\mu}}_{\mathrm{y}})=(2\pi)^{-6/2}\left[\det(n^{-1}\boldsymbol{\Sigma}_{\mathrm{y}})\right]^{-1/2}\exp\{-\tfrac{1}{2}[\hat{\boldsymbol{\mu}}_{\mathrm{y}}-\boldsymbol{\mu}]'(n^{-1}\boldsymbol{\Sigma}_{\mathrm{y}})^{-1}[\hat{\boldsymbol{\mu}}_{\mathrm{y}}-\boldsymbol{\mu}]\}. \tag{1.49}$$

1.4 The sampling distribution of the sample variance-covariance of a symmetric random tensor

Now we shall derive the sampling distribution of the sample variance-covariance matrix of a symmetric random tensor in vectorized form. First of all let us make a review of the *Wishart* distribution. The derivation of *Wishart* distribution, which is very fundamental in multivariate analysis, was a major breakthrough for the development of multivariate analysis. It is a multivariate matrix generalization of the univariate *Chi-square* distribution. The *Wishart* distribution was first derived by *Fisher* (1915) for *p=2*. *Wishart* (1928) gave a geometrical derivation of this distribution for general *p* in the general case. Other proofs were given by *Mahalanobis*, *Bose* and *Roy* (1937), *Hsu* (1939), *James* (1954), *Olkin* and *Roy* (1954) and others. *Anderson* (1958, 1984) and *Muirhead* (1982) present the detailed treatment of both the central and the noncentral *Wishart* distribution.

The following *theorem* shows the density function of the first term of (1.45) $\mathbf{A}=\mathbf{Z}'\mathbf{Z}$. The derivation of *theorem 1.11* is due to *James* (1954), *Olkin* and *Roy* (1954) and *Muirhead* (1982).

Theorem 1.11 (*Wishart* distribution of the 6×6 random matrix $\mathbf{A}=\mathbf{Z}'\mathbf{Z}$)

If $\mathbf{A}=\mathbf{Z}'\mathbf{Z}$, where the (*n*-1)×6 matrix **Z** is $\mathcal{N}_{(n-1),\,6}(\mathbf{0},\ \mathbf{I}_{n-1}\otimes\mathbf{\Sigma}_\mathbf{y})$ $(n-1\geq 6)$, then **A** is said to have the Wishart distribution with *n*-1 degrees of freedom and variance-covariance matrix $\mathbf{\Sigma}_\mathbf{y}$, denoted by

$$\mathcal{W}_6(n-1,\ \mathbf{\Sigma}_\mathbf{y})$$

The density function of **A** is

$$f(\mathbf{A})=\frac{[\det(\mathbf{Z}'\mathbf{Z})]^{(n-1-6-1)/2}}{(2)^{6(n-1)/2}\Gamma_6(\frac{n-1}{2})(\det\mathbf{\Sigma}_\mathbf{y})^{(n-1)/2}}\operatorname{etr}\{-\frac{1}{2}\mathbf{\Sigma}_\mathbf{y}^{-1}\mathbf{Z}'\mathbf{Z}\} \tag{1.50}$$

where $\Gamma_6(\frac{n-1}{2})$ denotes the multivariate gamma function.

***Proof*:**

Since **Z** is distributed according to $\mathcal{N}_{(n-1),\,6}(\mathbf{0},\ \mathbf{I}_{n-1}\otimes\mathbf{\Sigma}_\mathbf{y})$ from the *Theorem 1.6* the first derivation of the cumulative distribution function (c.d.f.) of *F*(**Z**) can be written

$$dF(\mathbf{Z})\ =\ f(\mathbf{Z})d\mathbf{Z}=(2\pi)^{-6(n-1)/2}(\det\mathbf{\Sigma}_\mathbf{y})^{-(n-1)/2}\operatorname{etr}\{-\frac{1}{2}\mathbf{\Sigma}_\mathbf{y}^{-1}\mathbf{Z}'\mathbf{Z}\}d\mathbf{Z}, \tag{1.51}$$

where the volume element $d\mathbf{Z}=\Lambda_{i=1}^{n-1}\Lambda_{j=1}^{6}dz_{ij}$ has been included to facilitate the calculation of Jacobians of (1.40). Since $n-1\geq 6$, **Z** has a rank of 6 with the probability 1. Put $\mathbf{Z}=\mathbf{H}_1\mathbf{T}_1$, where $\mathbf{H}_1$ is (*n*-1)×6 with $\mathbf{H}_1'\mathbf{H}_1=\mathbf{I}_{n-1}$(i.e., $\mathbf{H}_1\in\mathbf{V}_{n-1,6}$, the *Stiefel manifold* consisting of (*n*-1)×6 matrix with orthonomal columns) and $\mathbf{T}_1$ is 6×6 upper-triangular. Then $\mathbf{A}=\mathbf{Z}'\mathbf{Z}=\mathbf{T}_1'\mathbf{T}_1$ and the volume elements *d***Z** become

$$d\mathbf{Z}=(2)^{-6}(\det\mathbf{A})^{(n-1-6-1)/2}d\mathbf{A}\mathbf{H}_1'd\mathbf{H}_1, \tag{1.52}$$

so that the joint density of **A** and $\mathbf{H}_1$ is

$$(2\pi)^{-6(n-1)/2}(\det\mathbf{\Sigma}_\mathbf{y})^{-(n-1)/2}\operatorname{etr}\{-\frac{1}{2}\mathbf{\Sigma}_\mathbf{y}^{-1}\mathbf{A}\}(2)^{-6}(\det\mathbf{A})^{(n-1-6-1)/2}d\mathbf{A}\mathbf{H}_1'd\mathbf{H}_1. \tag{1.53}$$

The marginal density function of **A** then follows from this by integrating with respect to $\mathbf{H}_1$ over the Stiefel manifold $\mathbf{V}_{n\text{-}1,\,6,}$ using

$$\int_{V_{n-1,6}}\mathbf{H}_1'd\mathbf{H}_1=\frac{2^6\pi^{6(n-1)/2}}{\Gamma_6(\frac{n-1}{2})} \tag{1.54}$$

With (1.54) in (1.53) we have

$$\begin{aligned}(2\pi)^{-6(n-1)/2}(\det\mathbf{\Sigma}_\mathbf{y})^{-(n-1)/2}\operatorname{etr}\{-\frac{1}{2}\mathbf{\Sigma}_\mathbf{y}{}^{-1}\mathbf{A}\}2^{-6}(\det\mathbf{A})^{(n-1-6-1)/2}\frac{2^6\pi^{6(n-1)/2}}{\Gamma_6(\frac{n-1}{2})}d\mathbf{A}=\\ =\frac{(\det\mathbf{A})^{(6-1-6-1)/2}}{(2)^{6(n-1)/2}\Gamma_6(\frac{n-1}{2})(\det\mathbf{\Sigma})^{(n-1)/2}}\operatorname{etr}\{-\frac{1}{2}\mathbf{\Sigma}_\mathbf{y}{}^{-1}\mathbf{A}\}d\mathbf{A}\quad.\end{aligned} \tag{1.55}$$

This shows (1.50).

Further we show the density function of the sample variance-covariance matrix $\hat{\mathbf{\Sigma}}_\mathbf{y}$ of type BIQUUE

$$\hat{\mathbf{\Sigma}}_\mathbf{y}=\frac{1}{n-1}[\mathbf{Y}-\mathbf{1}\hat{\boldsymbol{\mu}}_\mathbf{y}']'[\mathbf{Y}-\mathbf{1}\hat{\boldsymbol{\mu}}_\mathbf{y}']=\frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{y}_i-\hat{\boldsymbol{\mu}}_\mathbf{y})(\mathbf{y}_i-\hat{\boldsymbol{\mu}}_\mathbf{y})' \tag{1.56}$$

Since

$$\begin{aligned}[\mathbf{Y}-\mathbf{1}\boldsymbol{\mu}']'[\mathbf{Y}-\mathbf{1}\boldsymbol{\mu}']=[(\mathbf{Y}-\mathbf{1}\hat{\boldsymbol{\mu}}_\mathbf{y}')+\mathbf{1}(\hat{\boldsymbol{\mu}}_\mathbf{y}'-\boldsymbol{\mu}')]'[(\mathbf{Y}-\mathbf{1}\hat{\boldsymbol{\mu}}_\mathbf{y}')+\mathbf{1}(\hat{\boldsymbol{\mu}}_\mathbf{y}'-\boldsymbol{\mu}')]=\\ =(\mathbf{Y}-\mathbf{1}\hat{\boldsymbol{\mu}}_\mathbf{y}')'(\mathbf{Y}-\mathbf{1}\hat{\boldsymbol{\mu}}_\mathbf{y}')+(\hat{\boldsymbol{\mu}}_\mathbf{y}-\boldsymbol{\mu})\mathbf{1}'\mathbf{1}(\hat{\boldsymbol{\mu}}_\mathbf{y}-\boldsymbol{\mu})'+(\mathbf{Y}-\mathbf{1}\hat{\boldsymbol{\mu}}_\mathbf{y}')'\mathbf{1}(\hat{\boldsymbol{\mu}}_\mathbf{y}-\boldsymbol{\mu})'+(\hat{\boldsymbol{\mu}}_\mathbf{y}-\boldsymbol{\mu})\mathbf{1}'(\mathbf{Y}-\mathbf{1}\hat{\boldsymbol{\mu}}_\mathbf{y}'),\end{aligned}$$

the last two terms of the equality being

$$(\mathbf{Y}-\mathbf{1}\hat{\boldsymbol{\mu}}_y')'\mathbf{1}(\hat{\boldsymbol{\mu}}_y-\boldsymbol{\mu})'+(\hat{\boldsymbol{\mu}}_y-\boldsymbol{\mu})\mathbf{1}'(\mathbf{Y}-\mathbf{1}\hat{\boldsymbol{\mu}}_y')=$$
$$=\mathbf{Y}'\mathbf{1}\hat{\boldsymbol{\mu}}_y'-\hat{\boldsymbol{\mu}}_y\mathbf{1}'\mathbf{1}\hat{\boldsymbol{\mu}}_y'-\mathbf{Y}'\mathbf{1}\boldsymbol{\mu}'+\hat{\boldsymbol{\mu}}_y\mathbf{1}'\mathbf{1}\boldsymbol{\mu}'+\hat{\boldsymbol{\mu}}_y\mathbf{1}'\mathbf{Y}-\hat{\boldsymbol{\mu}}_y\mathbf{1}'\mathbf{1}\hat{\boldsymbol{\mu}}_y'-\boldsymbol{\mu}\mathbf{1}'\mathbf{Y}+\boldsymbol{\mu}\mathbf{1}'\mathbf{1}\hat{\boldsymbol{\mu}}_y',$$

with $\hat{\boldsymbol{\mu}}_y=(1/n)\mathbf{Y}'\mathbf{1}$, so $\mathbf{Y}'\mathbf{1}=n\hat{\boldsymbol{\mu}}_y$ and $\mathbf{1}'\mathbf{Y}=n\hat{\boldsymbol{\mu}}_y'$ and $\mathbf{1}'\mathbf{1}=n$ bring them into above, then we get

$$n\hat{\boldsymbol{\mu}}_y\hat{\boldsymbol{\mu}}_y'-\hat{\boldsymbol{\mu}}_y n\hat{\boldsymbol{\mu}}_y'-n\hat{\boldsymbol{\mu}}_y\boldsymbol{\mu}'+\hat{\boldsymbol{\mu}}_y n\boldsymbol{\mu}'+\hat{\boldsymbol{\mu}}_y n\hat{\boldsymbol{\mu}}_y'-\hat{\boldsymbol{\mu}}_y n\hat{\boldsymbol{\mu}}_y'-\boldsymbol{\mu}n\hat{\boldsymbol{\mu}}_y'+\boldsymbol{\mu}n\hat{\boldsymbol{\mu}}_y' = \mathbf{0}.$$

So

$$[\mathbf{Y}-\mathbf{1}\boldsymbol{\mu}']'[\mathbf{Y}-\mathbf{1}\boldsymbol{\mu}']=(\mathbf{Y}-\mathbf{1}\hat{\boldsymbol{\mu}}_y')'(\mathbf{Y}-\mathbf{1}\hat{\boldsymbol{\mu}}_y')+(\hat{\boldsymbol{\mu}}_y-\boldsymbol{\mu})\mathbf{1}'\mathbf{1}(\hat{\boldsymbol{\mu}}_y-\boldsymbol{\mu})'$$
$$=\mathbf{Z}'\mathbf{Z}+n(\hat{\boldsymbol{\mu}}_y-\boldsymbol{\mu})(\hat{\boldsymbol{\mu}}_y-\boldsymbol{\mu})'.$$

Compared with (1.56) we have the relationship of $\mathbf{A}=\mathbf{Z}'\mathbf{Z}=(n-1)\hat{\boldsymbol{\Sigma}}_y$. By making the transformation $\mathbf{A}=(n-1)\hat{\boldsymbol{\Sigma}}_y$ in (1.50), whose Jacobian is $\mathbf{J}=(n-1)^{6(6+1)/2}$ (*Deemer and Olkin* 1951, also *Press* 1972, p.45) it follows

$$\frac{(\det\mathbf{A})^{(n-1-6-1)/2}}{(2)^{6(n-1)/2}\Gamma_6(\frac{n-1}{2})(\det\boldsymbol{\Sigma}_y)^{(n-1)/2}}\operatorname{etr}\{-\frac{1}{2}\boldsymbol{\Sigma}_y^{-1}\mathbf{A}\}d\mathbf{A}=$$
$$=\frac{(\det(n-1)\hat{\boldsymbol{\Sigma}}_y)^{(n-1-6-1)/2}}{(2)^{6(n-1)/2}\Gamma_6(\frac{n-1}{2})(\det\boldsymbol{\Sigma}_y)^{(n-1)/2}}\operatorname{etr}\{-\frac{1}{2}\boldsymbol{\Sigma}_y^{-1}\mathbf{A}\}\mathbf{J}d\hat{\boldsymbol{\Sigma}}_y$$
$$=\frac{(N-1)^{6(n-1-6-1)/2}(N-1)^{6(6+1)/2}(\det\hat{\boldsymbol{\Sigma}}_y)^{(n-1-6-1)/2}}{(2)^{6(n-1)/2}\Gamma_6(\frac{n-1}{2})(\det\boldsymbol{\Sigma}_y)^{(n-1)/2}}\operatorname{etr}\{-\frac{1}{2}\boldsymbol{\Sigma}_y^{-1}\mathbf{A}\}d\hat{\boldsymbol{\Sigma}}_y \quad (1.57)$$
$$=\frac{(N-1)^{6(n-1)/2}(\det\hat{\boldsymbol{\Sigma}}_y)^{(n-1-6-1)/2}}{(2)^{6(n-1)/2}\Gamma_6(\frac{n-1}{2})(\det\boldsymbol{\Sigma}_y)^{(n-1)/2}}\operatorname{etr}\{-\frac{1}{2}\boldsymbol{\Sigma}_y^{-1}\mathbf{A}\}d\hat{\boldsymbol{\Sigma}}_y$$
$$=\frac{1}{\Gamma_6(\frac{n-1}{2})(\det\boldsymbol{\Sigma}_y)^{(n-1)/2}}\left(\frac{n-1}{2}\right)^{6(n-1)/2}\operatorname{etr}\{-\frac{1}{2}\boldsymbol{\Sigma}_y^{-1}\mathbf{A}\}(\det\hat{\boldsymbol{\Sigma}}_y)^{(n-1-6-1)/2}d\hat{\boldsymbol{\Sigma}}_y$$

So the density function of the sample covariance matrix $\hat{\boldsymbol{\Sigma}}_y$ is

$$f(\hat{\boldsymbol{\Sigma}}_y)=\frac{1}{\Gamma_6(\frac{n-1}{2})(\det\boldsymbol{\Sigma}_y)^{(n-1)/2}}(\frac{n-1}{2})^{6(n-1)/2}\operatorname{etr}\{-\frac{1}{2}(n-1)\boldsymbol{\Sigma}_y^{-1}\hat{\boldsymbol{\Sigma}}_y\}(\det\hat{\boldsymbol{\Sigma}}_y)^{(n-1-6-1)/2}. \quad (1.58)$$

which is the right form of the *Wishart* distributed $\hat{\boldsymbol{\Sigma}}_y \sim \mathcal{W}_6(n-1,\ (n-1)^{-1}\boldsymbol{\Sigma}_y;\ \hat{\boldsymbol{\Sigma}}_y)$.

Note that from (1.50) we have $\hat{\boldsymbol{\Sigma}}_y=\mathbf{A}/(n-1)=\mathbf{Z}'\mathbf{Z}/(n-1)=\tilde{\mathbf{Z}}'\tilde{\mathbf{Z}}$, where

$$\tilde{\mathbf{Z}}=(n-1)^{-1/2}\mathbf{Z}\ \sim\ \mathcal{N}_{(n-1),\,6}(\mathbf{0},\ \mathbf{I}_{n-1}\otimes(n-1)^{-1}\boldsymbol{\Sigma}_y) \quad (1.59)$$

so that $\hat{\boldsymbol{\Sigma}}_y$ is distributed with Wishart distributed $\hat{\boldsymbol{\Sigma}}_y \sim \mathcal{W}_6(n-1,\ (n-1)^{-1}\boldsymbol{\Sigma}_y;\ \hat{\boldsymbol{\Sigma}}_y)$.

This leads to another direct derivation of the sampling distribution of a symmetric random tensor's sample variance-covariance matrix.

From (1.46) and (1.47) we can find that $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}_y$ are independent. Now we can characterize these derivations in *Theorem 1.12*.

Theorem 1.12 (the sampling distribution of the sample variance-covariance of a symmetric random tensor)

The sampling distribution of the sample variance-covariance $\hat{\boldsymbol{\Sigma}}_y$ of the vectorized 3×3 symmetric random tensor

$$\hat{\boldsymbol{\Sigma}}_y=\frac{1}{n-1}[\mathbf{Y}-\mathbf{1}\hat{\boldsymbol{\mu}}_y']'[\mathbf{Y}-\mathbf{1}\hat{\boldsymbol{\mu}}_y']=\frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{y}_i-\hat{\boldsymbol{\mu}}_y)(\mathbf{y}_i-\hat{\boldsymbol{\mu}}_y)' \quad (1.56)$$

is distributed according to $\hat{\boldsymbol{\Sigma}}_y \sim \mathcal{W}_6(n-1,\ (n-1)^{-1}\boldsymbol{\Sigma}_y;\ \hat{\boldsymbol{\Sigma}}_y)$ and independent of $\hat{\boldsymbol{\mu}}$. The p.d.f. of $\hat{\boldsymbol{\Sigma}}_y$ is

$$f(\hat{\boldsymbol{\Sigma}}_y)=\frac{1}{\Gamma_6(\frac{n-1}{2})(\det\boldsymbol{\Sigma}_y)^{(n-1)/2}}(\frac{n-1}{2})^{6(n-1)/2}\operatorname{etr}\{-\frac{1}{2}(n-1)\boldsymbol{\Sigma}_y^{-1}\hat{\boldsymbol{\Sigma}}_y\}(\det\hat{\boldsymbol{\Sigma}}_y)^{(n-1-6-1)/2}. \quad (1.58)$$

1.5 Sampling distributions of the estimates within the Gauss-Markov model

The derivations of the sampling distribution about the sample mean with the decomposition method are mostly discussed in the direct observation case. Here we will develop the sampling theory in a more general case, which includes (1) the sampling distribution within the special *Gauss-Markov* model; (2) the sampling distribution within the linear *Gauss-Markov* model and (3) the sampling distribution of the orthonormally transformed parameters.

1.5.1 The sampling distribution of the estimates within a special *Gauss-Markov* model

In this section we shall derive the sampling distribution of the estimates within a special *Gauss-Markov* model. We first introduce *Theorem 1.13*

Theorem 1.13 (marginal probability distributions, special linear Gauss-Markov model):

$$\begin{aligned} E\{\mathbf{y}\} &= \mathbf{A}\xi \\ D\{\mathbf{y}\} &= \mathbf{I}_n\sigma^2 \end{aligned} \quad subject\ \ to \left[\begin{array}{l} \mathbf{A}\in\mathbb{R}^{n\times m}, \text{rk}\,\mathbf{A}=m, E\{\mathbf{y}\}\in\mathcal{R}(\mathbf{A}) \\ \sigma^2\in\mathbb{R}^+ \end{array}\right.$$

defines a special Gauss-Markov model based upon independent, identically distributed (i.i.d.), *Gauss* normally distributed observations $\mathbf{y} := [y_1, y_2, \ldots, y_n]'$.
$\hat{\xi}$ is BLUUE of ξ in the *special linear Gauss-Markov model*

$$\hat{\xi} = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{y}\left[\begin{array}{c} E\{\hat{\xi}\} = \xi \\ D\{\hat{\xi}\} = (\mathbf{A}'\mathbf{A})^{-1}\sigma^2 \end{array}\right. \tag{1.60}$$

$\hat{\xi}$ is an element of a specific Gauss normal distribution of type $\hat{\xi} \sim \mathcal{N}\{\xi, (\mathbf{A}'\mathbf{A})^{-1}\sigma^2\}$ with the marginal probability density function

$$f_1(\hat{\xi};\, \xi,\, (\mathbf{A}'\mathbf{A})^{-1}\sigma^2) = (2\pi)^{-m/2}\sigma^{-m}\,|\,\mathbf{A}'\mathbf{A}\,|^{1/2}\, \exp\{-\frac{1}{2}(\hat{\xi}-\xi)'\mathbf{A}'\mathbf{A}(\hat{\xi}-\xi)/\sigma^2\}. \tag{1.61}$$

$\hat{\sigma}^2$ is the estimate of the only variance component of type BIQUUE

$$\hat{\sigma}^2 = \frac{1}{n-\text{rk}\,\mathbf{A}}(\mathbf{y}-\mathbf{A}\hat{\xi})'(\mathbf{y}-\mathbf{A}\hat{\xi}) \tag{1.62}$$

with the marginal probability density function

$$f_2(\hat{\sigma}^2) = \frac{1}{\sigma^p 2^{p/2}\Gamma(p/2)}\, p^{p/2}\hat{\sigma}^{p-2}\exp\{-\frac{1}{2}p\frac{\hat{\sigma}^2}{\sigma^2}\}, \tag{1.63}$$

where $p := n - \text{rk}\,\mathbf{A}$. The random variable x has a chi-square distribution with p degrees of freedom

$$x := (n-\text{rk}\,\mathbf{A})\frac{\hat{\sigma}^2}{\sigma^2} = \frac{p}{\sigma^2}\hat{\sigma}^2 = \frac{1}{\sigma^2}(\mathbf{y}-\mathbf{A}\hat{\xi})'(\mathbf{y}-\mathbf{A}\hat{\xi}) \tag{1.64}$$

with the probability density function

$$f_2(x) = \frac{1}{2^{p/2}\Gamma(p/2)}x^{\frac{p}{2}-1}\exp(-\frac{1}{2}x)\,. \tag{1.65}$$

Before proving *Theorem 1.13* we should introduce the *Lemma 1.14* about the transformation *of polar coordinates* $[\phi_1, \phi_2, \cdots, \phi_{n-1}, r] \in \mathbb{Y}$ as parameters of *an Euclidian observation space* to *Cartan coordinates* $[y_1, \cdots, y_n] \in \mathbb{Y}$. In addition we introduce the hypervolume element of a sphere $\mathbb{S}^{n-1} \subset \mathbb{Y}, \dim\mathbb{Y} = n$. First of all, we give three examples. *Second*, we summarize the general results in *Lemma 1.14*.

Lemma 1.14 (polar coordinates, hypervolume element, hypersurface element):

Let

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ \cdots \\ y_{n-3} \\ y_{n-2} \\ y_{n-1} \\ y_n \end{bmatrix} = r \begin{bmatrix} \cos\phi_{n-1}\cos\phi_{n-2}\cos\phi_{n-3}\cdots\cos\phi_2\cos\phi_1 \\ \cos\phi_{n-1}\cos\phi_{n-2}\cos\phi_{n-3}\cdots\cos\phi_2\sin\phi_1 \\ \cos\phi_{n-1}\cos\phi_{n-2}\cos\phi_{n-3}\cdots\cos\phi_2 \\ \cos\phi_{n-1}\cos\phi_{n-2}\cos\phi_{n-3}\cdots\sin\phi_2 \\ \cdots \\ \cos\phi_{n-1}\cos\phi_{n-2}\sin\phi_{n-3} \\ \cos\phi_{n-1}\cos\phi_{n-2} \\ \cos\phi_{n-1}\sin\phi_{n-2} \\ \sin\phi_{n-1} \end{bmatrix} \tag{1.66}$$

be a transformation of polar coordinates $(\phi_1,\phi_2,\cdots,\phi_{n-2},\phi_{n-1},r)$ to Cartesian coordinates $(y_1,y_2,\cdots,y_{n-1},y_n)$, their domain and range given by

$$(\phi_1,\phi_2,\cdots,\phi_{n-2},\phi_{n-1},r)\in[0,2\pi]\times]-\frac{\pi}{2},+\frac{\pi}{2}[\times\cdots\times]-\frac{\pi}{2},+\frac{\pi}{2}[\times]-\frac{\pi}{2},+\frac{\pi}{2}[\times]0,\infty[\ ,$$

then the *local hypervolume elements* are

$$dy_1\cdots dy_n = r^{n-1}dr\cos^{n-2}\phi_{n-1}\cos^{n-3}\phi_{n-2}\cdots\cos^2\phi_3\cos\phi_2 d\phi_{n-1}d\phi_{n-2}\cdots d\phi_2 d\phi_1 \tag{1.67}$$

as well as the *global hypersurface element*

$$\omega_{n-1} = \frac{2\cdot\pi^{(n-1)/2}}{\Gamma(\frac{n-1}{2})} := \int_{-\pi/2}^{+\pi/2}\cos^{n-2}\phi_{n-1}d\phi_{n-1}\cdots\int_{-\pi/2}^{+\pi/2}\cos\phi_2 d\phi_2\int_0^{2\pi}d\phi_{n-1}\,, \tag{1.68}$$

where $\Gamma(x)$ is the *gamma function*.

Proof:

The cumulative pdf of the multidimensional *Gauss-Laplace* probability distribution of the observation vector $\mathbf{y}=[y_1,\cdots,y_n]'\in\mathbb{Y}$ is

$$f(\mathbf{y}\,|\,E\{\mathbf{y}\},D\{\mathbf{y}\}=\mathbf{I}_n\sigma^2)dy_1\cdots dy_n = \frac{1}{(2\pi)^{n/2}\sigma^n}\exp[-\frac{1}{2\sigma^2}(\mathbf{y}-E\{\mathbf{y}\})'(\mathbf{y}-E\{\mathbf{y}\})]dy_1\cdots dy_n \tag{1.69}$$

We aim at splitting it into two *marginal pdfs* $f_1(\hat{\xi})$ of $\hat{\xi}$, BLUUE of ξ, and $f_2(\hat{\sigma}^2)$ of $\hat{\sigma}^2$, BIQUUE of σ^2, i.e.

$$f(\mathbf{y}\,|\,E\{\mathbf{y}\},D\{\mathbf{y}\}=\mathbf{I}_n\sigma^2)dy_1\cdots dy_n = f_1(\hat{\xi})f_2(\hat{\sigma}^2)d\hat{\xi}_1\cdots d\hat{\xi}_m d\hat{\sigma}^2 \tag{1.70}$$

First, *let us decompose the quadratic form* $\|\mathbf{y}-E\{\mathbf{y}\}\|^2$ *into estimates* $\widehat{E\{\mathbf{y}\}}$ *of* $E\{\mathbf{y}\}$.

$$\mathbf{y}-E\{\mathbf{y}\} = \mathbf{y}-\widehat{E\{\mathbf{y}\}}+(\widehat{E\{\mathbf{y}\}}-E\{\mathbf{y}\})$$

$$\mathbf{y}-E\{\mathbf{y}\} = \mathbf{y}-\mathbf{A}\hat{\xi}+\mathbf{A}(\hat{\xi}-\xi)$$

and

$$\begin{aligned}(\mathbf{y}-E\{\mathbf{y}\})'(\mathbf{y}-E\{\mathbf{y}\}) &= (\mathbf{y}-\widehat{E\{\mathbf{y}\}})'(\mathbf{y}-\widehat{E\{\mathbf{y}\}})+(\widehat{E\{\mathbf{y}\}}-E\{\mathbf{y}\})'(\widehat{E\{\mathbf{y}\}}-E\{\mathbf{y}\})+\\ &+(\mathbf{y}-\widehat{E\{\mathbf{y}\}})'(\widehat{E\{\mathbf{y}\}}-E\{\mathbf{y}\})+(\widehat{E\{\mathbf{y}\}}-E\{\mathbf{y}\})'(\mathbf{y}-\widehat{E\{\mathbf{y}\}})=\\ &=(\mathbf{y}-\widehat{E\{\mathbf{y}\}})'(\mathbf{y}-\widehat{E\{\mathbf{y}\}})+(\widehat{E\{\mathbf{y}\}}-E\{\mathbf{y}\})'(\widehat{E\{\mathbf{y}\}}-E\{\mathbf{y}\})\end{aligned}$$

$$\|\mathbf{y}-E\{\mathbf{y}\}\|^2 = \|\mathbf{y}-\widehat{E\{\mathbf{y}\}}\|^2+\|\widehat{E\{\mathbf{y}\}}-E\{\mathbf{y}\}\|^2$$

$$(\mathbf{y}-E\{\mathbf{y}\})'(\mathbf{y}-E\{\mathbf{y}\}) = (\mathbf{y}-\mathbf{A}\hat{\xi})'(\mathbf{y}-\mathbf{A}\hat{\xi}) + (\hat{\xi}-\xi)'\mathbf{A}'\mathbf{A}(\hat{\xi}-\xi) + \\ + (\mathbf{y}-\mathbf{A}\hat{\xi})'\mathbf{A}(\hat{\xi}-\xi) + (\hat{\xi}-\xi)'\mathbf{A}'(\mathbf{y}-\mathbf{A}\hat{\xi}) = \\ = (\mathbf{y}-\mathbf{A}\hat{\xi})'(\mathbf{y}-\mathbf{A}\hat{\xi}) + (\hat{\xi}-\xi)'\mathbf{A}'\mathbf{A}(\hat{\xi}-\xi)$$

$$\| \mathbf{y}-E\{\mathbf{y}\} \|^2 = \| \mathbf{y}-\mathbf{A}\hat{\xi} \|^2 + \| \hat{\xi}-\xi \|^2_{\mathbf{A}'\mathbf{A}} \tag{1.71}$$

Here, we took advantage of the *orthogonality relation*.

$$(\hat{\xi}-\xi)'\mathbf{A}'(\mathbf{y}-\mathbf{A}\hat{\xi}) = (\hat{\xi}-\xi)'\mathbf{A}'(\mathbf{I}_n - \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}')\mathbf{y} =$$

$$= (\hat{\xi}-\xi)'(\mathbf{A}' - \mathbf{A}'\mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}')\mathbf{y} = 0\,.$$

Second, we implement $\hat{\sigma}^2$ BIQUUE of σ^2 into the decomposed quadratic form.

$$\| \mathbf{y}-\mathbf{A}\hat{\xi} \|^2 = (\mathbf{y}-\mathbf{A}\hat{\xi})'(\mathbf{y}-\mathbf{A}\hat{\xi}) = \mathbf{y}'(\mathbf{I}_n - \mathbf{A}(\mathbf{A}\mathbf{A})^{-1}\mathbf{A}')\mathbf{y} =$$

$$= \mathbf{y}'\mathbf{M}\mathbf{y} = (n - \operatorname{rk}\mathbf{A})\hat{\sigma}^2$$

$$\| \mathbf{y}-E\{\mathbf{y}\} \|^2 = (n - \operatorname{rk}\mathbf{A})\hat{\sigma}^2 + (\hat{\xi}-\xi)'\mathbf{A}'\mathbf{A}(\hat{\xi}-\xi)$$

$$\| \mathbf{y}-E\{\mathbf{y}\} \|^2 = \mathbf{y}'\mathbf{M}\mathbf{y} + (\hat{\xi}-\xi)'\mathbf{N}(\hat{\xi}-\xi) \tag{1.72}$$

The matrix of the *normal equations* $\mathbf{N}:=\mathbf{A}'\mathbf{A}, \operatorname{rk}\mathbf{N} = \operatorname{rk}\mathbf{A}'\mathbf{A} = \operatorname{rk}\mathbf{A} = m$ and the matrix of the *variance component estimation* $\mathbf{M} := \mathbf{I}_n - \mathbf{A}(\mathbf{A}\mathbf{A})^{-1}\mathbf{A}'$, $\operatorname{rk}\mathbf{M} = \; = n - \operatorname{rk}\mathbf{A} = n - m$ have been introduced since their rank forms the basis of the *general forward and backward Helmert transformation*.

$$\mathbf{H}\mathbf{H}' = \mathbf{I}_n$$

$$\mathbf{z} = \sigma^{-1}\mathbf{H}(\mathbf{y}-E\{\mathbf{y}\}) = \sigma^{-1}\mathbf{H}(\mathbf{y}-\mathbf{A}\xi) \tag{1.73}$$

and

$$\mathbf{y}-E\{\mathbf{y}\} = \sigma\mathbf{H}'\mathbf{z} \tag{1.74}$$

$$\frac{1}{\sigma^2}(\mathbf{y}-E\{\mathbf{y}\})'(\mathbf{y}-E\{\mathbf{y}\}) = \mathbf{z}'\mathbf{H}'\mathbf{H}\mathbf{z} = \mathbf{z}'\mathbf{z}$$

$$\frac{1}{\sigma^2}\| \mathbf{y}-E\{\mathbf{y}\} \|^2 = \| \mathbf{z} \|^2,$$

where $\mathbf{H} \in \mathbb{R}^{n\times n}$ is the *quadratic Helmert matrix*, also called *extended Helmert matrix* or *augmented Helmert matrix* (Lancaster, 1965):

$$\mathbf{H} :=$$

$$\begin{bmatrix} \frac{1}{\sqrt{1\cdot 2}} & -\frac{1}{\sqrt{1\cdot 2}} & 0 & 0 & \cdots & 0 & 0 \\ \frac{1}{\sqrt{2\cdot 3}} & \frac{1}{\sqrt{2\cdot 3}} & -\frac{2}{\sqrt{2\cdot 3}} & 0 & \cdots & 0 & 0 \\ \frac{1}{\sqrt{3\cdot 4}} & \frac{1}{\sqrt{3\cdot 4}} & \frac{1}{\sqrt{3\cdot 4}} & -\frac{3}{\sqrt{3\bullet 4}} & \cdots & 0 & 0 \\ \cdots & & & & \cdots & & \\ \frac{1}{\sqrt{(n-1)(n-2)}} & \frac{1}{\sqrt{(n-1)(n-2)}} & \frac{1}{\sqrt{(n-1)(n-2)}} & \frac{1}{\sqrt{(n-1)(n-2)}} & \cdots & -\frac{n-1}{\sqrt{(n-1)(n-2)}} & 0 \\ \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{n(n-1)}} & \cdots & \frac{1}{\sqrt{n(n-1)}} & -\frac{n}{\sqrt{n(n-2)}} \\ \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \cdots & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} \end{bmatrix} \tag{1.75}$$

Since the *quadratic Helmert matrix* is *orthonormal*, the absolute value of the *Jacobian* of the *general backward Helmert* transformation $\mathbf{z} \mapsto \mathbf{y} - \mathbf{A}\boldsymbol{\xi} = \sigma \mathbf{H}'\mathbf{z}$ is

$$J_{\mathbf{y}-\mathbf{A}\boldsymbol{\xi}\to\mathbf{z}} = |\det \mathbf{J}| = \sigma^n |\mathbf{H}'| = \sigma^n |\mathbf{H}| = \sigma^n . \tag{1.76}$$

Therefore we have the transformation of the *volume element* in (1.69)

$$dy_1 dy_2 \cdots dy_n = \sigma^n dz_1 dz_2 \cdots dz_n \tag{1.77}$$

which generates the *cumulative probability* (1.69)

$$\begin{aligned} dF &= f(y_1, y_2, \cdots, y_n) dy_1 \cdots dy_n = J_{\mathbf{y}-\mathbf{A}\boldsymbol{\xi}\to\mathbf{z}} f(z_1, z_2, \cdots, z_n) dz_1 dz_2 \cdots dz_n = \\ &= \frac{1}{(2\pi)^{n/2}} \exp\{-\frac{1}{2}\mathbf{z}'\mathbf{z}\} = \frac{1}{(2\pi)^{n/2}} \exp\{-\frac{1}{2}(z_1^2 + z_2^2 \cdots + z_n^2)\} dz_1 dz_2 \cdots dz_n \end{aligned} \tag{1.78}$$

Third, the standard canonical variable $\mathbf{z} \in \mathbb{R}^n$ has to be associated with norms $\| \mathbf{y} - \mathbf{A}\hat{\boldsymbol{\xi}} \|$ and $\| \hat{\boldsymbol{\xi}} - \boldsymbol{\xi} \|_{\mathbf{A}'\mathbf{A}}$. We take advantage of the eigenspace representation of the matrices (**M**, **N**) and their associated norms.

$$\mathbf{y}'\mathbf{M}\mathbf{y} = \mathbf{y}\mathbf{V}\boldsymbol{\Lambda}_{\mathbf{M}}\mathbf{V}'\mathbf{y} \;\; \textit{versus} \;\; (\hat{\boldsymbol{\xi}} - \boldsymbol{\xi})'\mathbf{N}(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}) = (\hat{\boldsymbol{\xi}} - \boldsymbol{\xi})'\mathbf{U}\boldsymbol{\Lambda}_{\mathbf{N}}\mathbf{U}'(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi})$$

$$\underset{\in \mathbb{R}^n = \mathbb{R}^{n-m} \times \mathbb{R}^m}{\boldsymbol{\Lambda}_{\mathbf{M}} = Diag(\mu_1, \cdots, \mu_{n-m}, 0, \cdots, 0)} \;\; \textit{versus} \;\; \underset{\in \mathbb{R}^m}{\boldsymbol{\Lambda}_{\mathbf{N}} = Diag(\nu_1, \cdots, \nu_m)}$$

m eigenvalues of the matrix **M** are zero, but $n - \operatorname{rk} \mathbf{A} = n - m$ is the number of its non-vanishing eigenvalues which we denote by $(\mu_1, \cdots, \mu_{n-m})$. In contrast, $m = \operatorname{rk} \mathbf{A}$ is the number of eigenvalues of the matrix **N**, all non-zero. The *canonical random variables*

$$\mathbf{V}'\mathbf{y} = \mathbf{y}^* \Leftrightarrow \mathbf{y} = \mathbf{V}\mathbf{y}^* \;\; \textit{and} \;\; \mathbf{U}'(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}) = \hat{\boldsymbol{\eta}} - \boldsymbol{\eta}$$

lead to

$$\frac{1}{\sigma^2}(\mathbf{y} - E\{\mathbf{y}\})'(\mathbf{y} - E\{\mathbf{y}\}) = \frac{1}{\sigma^2}(\mathbf{y}^*)\boldsymbol{\Lambda}_{\mathbf{M}}\mathbf{y}^* + \frac{1}{\sigma^2}(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta})\boldsymbol{\Lambda}_{\mathbf{N}}(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta})$$

$$\frac{1}{\sigma^2}(\mathbf{y} - E\{\mathbf{y}\})'(\mathbf{y} - E\{\mathbf{y}\}) = \frac{1}{\sigma^2}\sum_{j=1}^{n-m}(y_j^*)^2 \mu_j + \frac{1}{\sigma^2}\sum_{i=1}^{m}(\hat{\eta}_i - \eta_i)^2 \nu_i$$

$$\frac{1}{\sigma^2}(\mathbf{y} - E\{\mathbf{y}\})'(\mathbf{y} - E\{\mathbf{y}\}) = z_1^2 + \cdots + z_{n-m}^2 + z_{n-m+1}^2 + \cdots + z_n^2$$

subject to

$$z_1^2 + \cdots + z_{n-m}^2 = = \frac{1}{\sigma^2}\sum_{j=1}^{n-m}(y_j^*)^2 \mu_j \;\; \text{and} \;\; z_{n-m+1}^2 + \cdots + z_n^2 = \frac{1}{\sigma^2}\sum_{i=1}^{m}(\hat{\eta}_i - \eta_i)^2 \nu_i$$

$$\begin{aligned} \| \mathbf{z} \|^2 = \mathbf{z}'\mathbf{z} &= z_1^2 + \cdots + z_{n-m}^2 + z_{n-m+1}^2 + \cdots + z_n^2 = \frac{1}{\sigma^2}\mathbf{y}'\mathbf{M}\mathbf{y} + \frac{1}{\sigma^2}(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi})'\mathbf{N}(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}) = \\ &= \frac{1}{\sigma^2}\| \mathbf{y} - E\{\mathbf{y}\} \|^2 = \frac{1}{\sigma^2}(\mathbf{y} - E\{\mathbf{y}\})'(\mathbf{y} - E\{\mathbf{y}\}) \end{aligned}$$

Obviously, the *eigenspace synthesis* of the matrices $\mathbf{N} = \mathbf{A}'\mathbf{A}$ and $\mathbf{M} = \mathbf{I}_n - \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'$ has guided us to the proper structure synthesis of the *generalized Helmert transformation*.

Fourth, the norm decomposition enables us to split the *cumulative probability* (1.78) into the *pdf* of the *Helmert random variable* $x := z_1^2 + \cdots + z_{n-m}^2 = \sigma^{-2}(n - \operatorname{rk} \mathbf{A})\hat{\sigma}^2 = \sigma^{-2}(n - m)\hat{\sigma}^2$ and the *pdf* of the *difference random parameter vector* $z_{n-m+1}^2 + \cdots + z_n^2 = \sigma^{-2}(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi})'\mathbf{A}'\mathbf{A}(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi})$.

$$dF = f(z_1, \cdots, z_{n-m}, z_{n-m+1}, \cdots, z_n) dz_1 \cdots dz_{n-m} dz_{n-m+1} dz_n$$

$$\begin{aligned} f(z_1, \cdots, z_n) &= \frac{1}{(2\pi)^{n/2}} \exp\{-\frac{1}{2}\mathbf{z}'\mathbf{z}\} = \\ &= (\frac{1}{2\pi})^{\frac{n-m}{2}} \exp\{-\frac{1}{2}(z_1^2 + \cdots + z_{n-m}^2)\}(\frac{1}{2\pi})^{\frac{m}{2}} \exp\{-\frac{1}{2}(z_{n-m+1}^2 + \cdots + z_n^2)\} \end{aligned} \tag{1.79}$$

The partitioned vector of the standard random variable z is associated with the norm $\| z_{n-m} \|^2$ and $\| z_m \|^2$, namely

$$\| z_{n-m} \|^2 + \| z_m \|^2 = z_1^2 + \cdots + z_{n-m}^2 + z_{n-m+1}^2 + \cdots + z_n^2. \tag{1.80}$$

Part A

Let us introduce *Helmert's polar coordinates* $(\phi_1, \cdots, \phi_{n-m-1}, r)$ which represent the *Cartesian coordinates*

$$\begin{aligned} z_1 &= r\cos\phi_{n-m-1}\cos\phi_{n-m-2}\cdots\cos\phi_2\cos\phi_1 \\ z_2 &= r\cos\phi_{n-m-1}\cos\phi_{n-m-2}\cdots\cos\phi_2\sin\phi_1 \\ &\cdots \\ z_{n-m-1} &= r\cos\phi_{n-m-1}\sin\phi_{n-m-2} \\ z_{n-m} &= r\sin\phi_{n-m-1} \end{aligned} \tag{1.81}$$

The representation of the local (*n-m*)-dimensional *hypervolume element* in terms of *polar coordinates* $(\phi_1, \phi_2, \cdots, \phi_{n-m-1}, r)$ has already been given by *Lemma 1.14.*

$$\begin{aligned} dz_1 dz_2 \cdots dz_{n-m-1} dz_{n-m} &= r^{n-m-1} dr (\cos\phi_{n-m-1})^{n-m-1} (\cos\phi_{n-m-2})^{n-m-2} \cdots \\ &\cdots \cos^2\phi_3 \cos\phi_2 d\phi_{n-m-1} d\phi_{n-m-2} \cdots d\phi_3 d\phi_2 d\phi_1 \end{aligned} \tag{1.82}$$

Here, we only transform the new random variable r into *Helmert's* random variable x.

$$x := r^2 \Rightarrow dx = 2rdr, \; dr = \frac{dx}{2\sqrt{x}}, \; r^{n-m-1} = x^{(n-m-1)/2}$$

$$r^{n-m-1} dr = \frac{1}{2} x^{(n-m-1)/2} dx \tag{1.83}$$

Part A concludes with the representation of the *left pdf* in terms of *Helmert's polar coordinates*

$$\begin{aligned} dF_\ell &= \left(\frac{1}{2\pi}\right)^{\frac{n-m}{2}} \exp\{-\frac{1}{2}(z_1^2 + \cdots + z_{n-m}^2)\} dz_1 \cdots dz_{n-m} = \\ &= \frac{1}{2}\left(\frac{1}{2\pi}\right)^{\frac{n-m}{2}} \exp(-\frac{1}{2}x)\, x^{\frac{n-m-2}{2}} dx (\cos\phi_{n-m-1})^{n-m-1} (\cos\phi_{n-m-2})^{n-m-2} \cdots \cos^2\phi_3 \cos\phi_2 d\phi_{n-m-1} d\phi_{n-m-2} \cdots d\phi_3 d\phi_2 d\phi_1 \end{aligned} \tag{1.84}$$

Part B

Part B focuses on the representation of the *right pdf* in terms of the random variables.

$$dF_r = \left(\frac{1}{2\pi}\right)^{\frac{m}{2}} \exp\{-\frac{1}{2}(z_{n-m+1}^2 + \cdots + z_n^2)\} dz_{n-m+1} \cdots dz_n \tag{1.85}$$

$$z_{n-m+1}^2 + \cdots + z_n^2 = \frac{1}{\sigma^2} (\hat{\xi} - \xi)' \mathbf{A}' \mathbf{A} (\hat{\xi} - \xi) \tag{1.86}$$

The computation of the local m-dimensional hypervolume element $dz_{n-m+1} \cdots dz_n$ could be derived in the following way, which is based upon the matrix of the metric, $\sigma^{-2}\mathbf{A}'\mathbf{A}$. Since the metric $\sigma^{-2}\mathbf{A}'\mathbf{A}$ is positive-definite there exists a nonsingular $m \times m$ matrix $\mathbf{B}$, such that

$$\sigma^{-2}\mathbf{A}'\mathbf{A} = \mathbf{B}'\mathbf{B} \tag{1.87}$$

and we put

$$\mathbf{z}_m = \mathbf{B}(\hat{\xi} - \xi). \tag{1.88}$$

So (1.86) is equivalent to

$$\mathbf{z}_m' \mathbf{z}_m = (\hat{\xi} - \xi)' \mathbf{B}' \mathbf{B} (\hat{\xi} - \xi). \tag{1.89}$$

The Jacobian of the transformation (1.88) is

$$J_{\mathbf{z}_m \to \hat{\boldsymbol{\xi}}} = \det[\frac{\partial(z_{n-m+1},\cdots,z_n)}{\partial(\hat{\xi}_1,\cdots,\hat{\xi}_m)}] = \det \mathbf{B} = [\det(\mathbf{B}'\mathbf{B})]^{1/2} =$$
$$= [\det(\sigma^{-2}\mathbf{A}'\mathbf{A})]^{1/2} = \sigma^{-m}[\det(\mathbf{A}'\mathbf{A})]^{1/2} = \tag{1.90}$$
$$= \sigma^{-m} \mid \mathbf{A}'\mathbf{A} \mid^{1/2}$$

which brings

$$dz_{n-m+1}\cdots dz_n = \frac{1}{\sigma^m} \mid \mathbf{A}'\mathbf{A} \mid^{1/2} d\hat{\xi}_1 \cdots d\hat{\xi}_m . \tag{1.91}$$

The *first* representation of the *right pdf* is given by

$$dF_r = (\frac{1}{2\pi})^{\frac{m}{2}} \exp\{-\frac{1}{2}(z^2_{n-m+1} + \cdots + z^2_n)\} dz_{n-m+1}\cdots dz_n =$$
$$= (\frac{1}{2\pi})^{\frac{m}{2}} \frac{\mid \mathbf{A}'\mathbf{A} \mid^{1/2}}{\sigma^m} \exp\{-\frac{1}{2\sigma^2}(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi})'\mathbf{A}'\mathbf{A}(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi})\} d\hat{\xi}_1 \cdots d\hat{\xi}_m . \tag{1.92}$$

Part C

Part C is an attempt to merge the *left and right pdf* according to

$$dF = dF_\ell dF_r = \frac{1}{2}(\frac{1}{2\pi})^{\frac{n-m}{2}} \exp(-\frac{1}{2}x)\, x^{(n-m-2)/2} dx d\omega_{n-m-1} \times$$
$$\times (\frac{1}{2\pi})^{\frac{m}{2}} \frac{\mid \mathbf{A}'\mathbf{A} \mid^{1/2}}{\sigma^m} \exp\{-\frac{1}{2}(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi})'\mathbf{A}'\mathbf{A}(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi})\} d\hat{\xi}_1 \cdots d\hat{\xi}_m .$$

The local (*n*-*m*-1)-dimensional hypersurface element has been denoted by $d\omega_{n-m-1}$ according to *Lemma 1.14*.

Fifth, we are going to compute the *marginal pdf* of $\hat{\boldsymbol{\xi}}$ BLUUE of $\boldsymbol{\xi}$.

$$dF_1 = f_1(\hat{\boldsymbol{\xi}}) d\hat{\xi}_1 \cdots d\hat{\xi}_m$$

includes the *first marginal pdf* $f_1(\hat{\boldsymbol{\xi}})$

$$f_1(\hat{\boldsymbol{\xi}}) := \int_0^\infty dx \oint d\omega_{n-m-1} \frac{1}{2}(\frac{1}{2\pi})^{\frac{n-m}{2}} \exp(-\frac{1}{2}x)\, x^{(n-m-2)/2} (\frac{1}{2\pi})^{\frac{m}{2}} \frac{\mid \mathbf{A}'\mathbf{A} \mid^{1/2}}{(\sigma^2)^{m/2}} \exp\{-\frac{1}{2}(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi})'\mathbf{A}'\mathbf{A}(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi})\}$$

subject to

$$\int_0^\infty dx \oint d\omega_{n-m-1} \frac{1}{2}(\frac{1}{2\pi})^{\frac{n-m}{2}} \exp(-\frac{1}{2}x)\; x^{(n-m-2)/2} = 1.$$

This leads us to

$$f_1(\hat{\boldsymbol{\xi}}) = (\frac{1}{2\pi})^{\frac{m}{2}} \frac{\mid \mathbf{A}'\mathbf{A} \mid^{1/2}}{\sigma^m} \exp\{-\frac{1}{2}(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi})'\mathbf{A}'\mathbf{A}(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi})\} . \tag{1.93}$$

Unfortunately, such a general *multivariate Gauss-Laplace normal distribution* cannot be tabulated. An alternative is offered by introducing canonical unknown parameters $\hat{\boldsymbol{\eta}}$ as random variables, which will be discussed in Section 4.3.3.

Sixth, we shall compute the *marginal pdf* of *Helmert's random variable* $x = (n - \mathrm{rk}\mathbf{A})\hat{\sigma}^2 / \sigma^2 = (n-m)\hat{\sigma}^2 / \sigma^2$, with $\hat{\sigma}^2$ as BIQUUE of σ^2, and with

$$dF = f_2(x) dx$$

including the *second marginal pdf* $f_2(x)$.

The definition

$$f_2(x) := \oint d\omega_{n-m-1} \frac{1}{2}(\frac{1}{2\pi})^{\frac{n-m}{2}} \exp(-\frac{1}{2}x)\; x^{(n-m-2)/2} \int_{-\infty}^{+\infty} d\hat{\xi}_1 \cdots \int_{-\infty}^{+\infty} d\hat{\xi}_m (\frac{1}{2\pi})^{\frac{m}{2}} \frac{\mid \mathbf{A}'\mathbf{A} \mid^{1/2}}{\sigma^m} \exp\{-\frac{1}{2\sigma^2}(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi})'\mathbf{A}'\mathbf{A}(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi})\}$$

subject to

$$\omega_{n-m-1} = \oint d\omega_{n-m-1} = \frac{2\pi^{(n-m-1)/2}}{\Gamma(\frac{n-m-1}{2})}$$

according to Lemma 1.14

$$\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} d\hat{\xi}_1\cdots d\hat{\xi}_m (\frac{1}{2\pi})^{\frac{m}{2}} \frac{|\mathbf{A}'\mathbf{A}|^{1/2}}{\sigma^m} \exp\{-\frac{1}{2\sigma^2}(\hat{\xi}-\xi)'\mathbf{A}'\mathbf{A}(\hat{\xi}-\xi)\} =$$

$$= \int_{-\infty}^{+\infty} dz_{n-m+1}\cdots\int_{-\infty}^{+\infty} dz_n (\frac{1}{2\pi})^{\frac{m}{2}} \exp\{-\frac{1}{2}(z_{n-m+1}^2+\cdots+z_n^2)\} = 1$$

leads us to

$$p := n - \operatorname{rk}\mathbf{A} = n - m$$

$$\sqrt{\pi}\Gamma(\frac{n-m-1}{2}) = \Gamma(\frac{1}{2})\Gamma(\frac{n-m-1}{2}) = \Gamma(\frac{n-m}{2}) = \Gamma(\frac{p}{2})$$

$$f_2(x) = \frac{1}{2^{p/2}\Gamma(p/2)} x^{\frac{p}{2}-1} \exp(-\frac{1}{2}x), \tag{1.94}$$

which is the standard pdf of the normalized sample variance, known as *Helmert's Chi-Square* χ^2 with $p = n - \operatorname{rk}\mathbf{A} = n - m$ "*degrees of freedom*". If we substitute (1.64) $x=(n-rk\mathbf{A})\hat{\sigma}^2/\sigma^2=(n-m)\hat{\sigma}^2/\sigma^2$ and $dx = (n-\operatorname{rk}\mathbf{A})\sigma^{-2}d\hat{\sigma}^2 = (n-m)\sigma^{-2}d\hat{\sigma}^2$, we arrive at the *pdf* of the sample variance $\hat{\sigma}^2$, in particular

$$dF_2 = f_2(\hat{\sigma}^2)d\hat{\sigma}^2$$

$$f_2(\hat{\sigma}^2) = \frac{1}{\sigma^p 2^{p/2}\Gamma(p/2)} p^{p/2}\hat{\sigma}^{p-2} \exp(-\frac{1}{2}p\frac{\hat{\sigma}^2}{\sigma^2}). \tag{1.95}$$

This concludes the proof.

1.5.2 The sampling distribution of the estimates within a linear Gauss-Markov model

In this section we shall derive the sampling distribution of the estimates within a linear Gauss-Markov model. First of all we introduce *Theorem 1.15*

Theorem 1.15

Let $E\{\mathbf{y}\} = \mathbf{A}\xi,\ \mathbf{A} \in \mathbb{R}^{n\times m},\ E\{\mathbf{y}\} \in \mathcal{R}(\mathbf{A}),\ \operatorname{rk}\mathbf{A} = m$
$D\{\mathbf{y}\} = \mathbf{\Sigma}_y \in \mathbb{R}^{n\times n}$, $\mathbf{\Sigma}_y$ positive-definite, $\operatorname{rk}\mathbf{\Sigma}_y = n$,
be a linear Gauss-Markov model based upon independent, identically distributed (i.i.d.) *Gauss* normally distributed observations $\mathbf{y} := [y_1, y_2,\ldots, y_n]'$, $\hat{\xi}$ is $\mathbf{\Sigma}_y$ - BLUUE of ξ in the *linear Gauss-Markov model*

$$\hat{\xi} = (\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A})^{-1}\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{y} \text{ subject to } \left[\begin{array}{l} E\{\hat{\xi}\} = \xi \\ D\{\hat{\xi}\} = (\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A})^{-1} \end{array}\right. \tag{1.96}$$

$\hat{\xi}$ is an element of a specific Gauss normal distribution of type $\hat{\xi} \sim \mathcal{N}\{\xi, (\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A})^{-1}\}$ with the probability density function

$$f(\hat{\xi}; \xi, (\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A})^{-1}) = (2\pi)^{-m/2} |(\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A})^{-1}|^{-1/2} \exp\{-\frac{1}{2}(\hat{\xi}-\xi)'\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}(\hat{\xi}-\xi)\}. \tag{1.97}$$

Proof:

The probability density function (p.d.f.) of the i.i.d. Gauss normally distributed observations $\mathbf{y}$ is

$$f(\mathbf{y}; E\{\mathbf{y}\}, \mathbf{\Sigma}_y) = (2\pi)^{-n/2}(\det\mathbf{\Sigma}_y)^{-1/2} \exp\{-\frac{1}{2}(\mathbf{y}-E\{\mathbf{y}\})'\mathbf{\Sigma}_y^{-1}(\mathbf{y}-E\{\mathbf{y}\})\}. \tag{1.98}$$

We aim at deriving the marginal distributions of $\hat{\xi}$.

A first decomposition of $\mathbf{y} - E\{\mathbf{y}\}$ is

$$\mathbf{y}-E\{\mathbf{y}\}=\mathbf{y}-\mathbf{A}\xi=\mathbf{y}-\mathbf{A}\hat{\xi}+\mathbf{A}(\hat{\xi}-\xi)$$

i.e.

$$(\mathbf{y}-E\{\mathbf{y}\})'\boldsymbol{\Sigma}_y^{-1}(\mathbf{y}-E\{\mathbf{y}\})=(\mathbf{y}-\mathbf{A}\hat{\xi})'\boldsymbol{\Sigma}_y^{-1}(\mathbf{y}-\mathbf{A}\hat{\xi})+(\hat{\xi}-\xi)'\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}(\hat{\xi}-\xi) \tag{1.99}$$

So with (1.98) and (1.99) we get

$$f(\mathbf{y};E\{\mathbf{y}\},\boldsymbol{\Sigma}_y)dy_1\cdots dy_n=$$
$$(2\pi)^{-n/2}(\det\boldsymbol{\Sigma}_y)^{-1/2}\exp\{-\frac{1}{2}((\mathbf{y}-\mathbf{A}\hat{\xi})'\boldsymbol{\Sigma}_y^{-1}(\mathbf{y}-\mathbf{A}\hat{\xi})\}\cdot\underbrace{\exp\{-\frac{1}{2}(\hat{\xi}-\xi)'\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}(\hat{\xi}-\xi)\}}_{\hat{\xi}-statistic}dy_1\cdots dy_n. \tag{1.100}$$

Because of the general variance-covariance matrix $\boldsymbol{\Sigma}_y$ in (1.100), the methods used in the proof of *Theorem 1.13* can be hardly applied in derivation of the probability density function of $\hat{\xi}$. First we will simplify this problem by the transformation of the linear Gauss-Markov model

$$\mathbf{z}=\boldsymbol{\Sigma}_y^{-1/2}\mathbf{y},\quad \mathbf{B}=\boldsymbol{\Sigma}_y^{-1/2}\mathbf{A}. \tag{1.101}$$

The Jacobian of the transformation (1.101) is

$$J_{\mathbf{y}\to\mathbf{z}}=d\mathbf{y}/d\mathbf{z}=|\boldsymbol{\Sigma}_y^{1/2}|=|\boldsymbol{\Sigma}_y|^{1/2},$$

so that the differential elements are connected by the relation

$$dy_1dy_2\cdots dy_n=|\boldsymbol{\Sigma}_y|^{1/2}\ dz_1dz_2\cdots dz_n,$$

and the probability density function of **y** (1.98) transforms to

$$f(\mathbf{y};E\{\mathbf{y}\},\boldsymbol{\Sigma}_y)dy_1dy_2\cdots dy_n=(2\pi)^{-n/2}\,|\boldsymbol{\Sigma}_y|^{-1/2}\ \exp\{-\frac{1}{2}(\mathbf{z}-E\{\mathbf{z}\})'\boldsymbol{\Sigma}_y'^{-1/2}\boldsymbol{\Sigma}_y^{-1}\boldsymbol{\Sigma}_y^{-1/2}(\mathbf{z}-E\{\mathbf{z}\})\}\,|\boldsymbol{\Sigma}_y|^{1/2}\ dz_1dz_2\cdots dz_n$$
$$=(2\pi)^{-n/2}\exp\{-\frac{1}{2}(\mathbf{z}-E\{\mathbf{z}\})'(\mathbf{z}-E\{\mathbf{z}\})\}dz_1dz_2\cdots dz_n\,. \tag{1.102}$$

So we get the probability density function (p.d.f.) of **z** as

$$f(\mathbf{z};E\{\mathbf{z}\},\boldsymbol{\Sigma}_z)=(2\pi)^{-n/2}\exp\{-\frac{1}{2}(\mathbf{z}-E\{\mathbf{z}\})'(\mathbf{z}-E\{\mathbf{z}\})\} \tag{1.103}$$

which is a standard multivariate Gauss normal distribution.

Then the linear Gauss-Markov model will be simplified as

$$E\{\mathbf{z}\}=\mathbf{B}\xi,\ \mathbf{B}\in\mathbb{R}^{n\times m},\ E\{\mathbf{z}\}\in\mathcal{R}(\mathbf{B}),\ \text{rk}\,\mathbf{B}=m$$
$$D\{\mathbf{z}\}=\mathbf{I}_n\in\mathbb{R}^{n\times n}. \tag{1.104}$$

The $\boldsymbol{\Sigma}_y$ - BLUUE of ξ , $\hat{\xi}$, will be represented as

$$\hat{\xi}=(\mathbf{B}'\mathbf{B})^{-1}\mathbf{B}'\mathbf{z}\left[\begin{array}{c}E\{\hat{\xi}\}=\xi\\ D\{\hat{\xi}\}=(\mathbf{B}'\mathbf{B})^{-1}\end{array}\right. \tag{1.105}$$

which reduces the derivation of the probability density function of $\hat{\xi}$ for the general variance-covariance matrix $\boldsymbol{\Sigma}_y$ to a special simple case that $\sigma^2=1$ in (1.61) of *Theorem 1.13* and with respect to the probability density function (p.d.f.) of the vector **z** (1.103)**,** i.e.

$$f(\hat{\xi};\,\xi,\,(\mathbf{B}'\mathbf{B})^{-1})=(2\pi)^{-m/2}\,|\mathbf{B}'\mathbf{B}|^{1/2}\ \exp\{-\frac{1}{2}(\hat{\xi}-\xi)'\mathbf{B}'\mathbf{B}(\hat{\xi}-\xi)\} \tag{1.106}$$

With the relation of (1.101) and from the probability density function (p.d.f.) of **z** (1.103) we get the probability density function of $\hat{\xi}$ for the general variance-covariance matrix $\boldsymbol{\Sigma}_y$

$$f(\hat{\xi};\,\xi,\,(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A})^{-1})=(2\pi)^{-m/2}\,|(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A})^{-1}|^{-1/2}\ \exp\{-\frac{1}{2}(\hat{\xi}-\xi)'\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}(\hat{\xi}-\xi)\}. \tag{1.107}$$

This completes the proof that $\hat{\xi}\sim\mathcal{N}\{\xi,(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A})^{-1}\}$ of *Theorem 1.15* .

1.5.3 The sampling distribution of the orthonormally transformed parameters

While deriving (1.93) for the proof of *Theorem 1.13* we have mentioned that $\hat{\xi}_1,\cdots,\hat{\xi}_m$ are dependently distributed. In order to make the hypothesis tests about the distinct elements more efficient and uncorrelated, we could naturally transform the original parameters to canonical parameters η_i of uncorrelated linear combinations of ξ_i's. This method uses a similar technique to the well known *principal component analysis*, which was introduced by *K. Pearson* (1901) as a tool of fitting planes to a system of points in space and were later generalized by *Hotelling* (1931) for analyzing correlation structures and the *canonical form of a linear model*. In fact principal components analysis is concerned fundamentally with the eigenstructure of covariance matrices, i.e., with their eigenvalues and eigenvectors. Therefore we will firstly make an orthonormal transformation of the original parameters, then derive the p.d.f. of the transformed parameters and the related the variance-covariance matrix of them, and perform hypothesis test for them, which we name *eigen-inference*.

Let us introduce the *canonical random variables* $(\hat{\eta}_1,\cdots,\hat{\eta}_m)$ which are generated by decorrelating the quadratic form $\|\hat{\boldsymbol{\xi}}-\boldsymbol{\xi}\|^2_{\mathbf{A}'\mathbf{A}}$.

$$(\hat{\boldsymbol{\xi}}-\boldsymbol{\xi})'\mathbf{A}'\mathbf{A}(\hat{\boldsymbol{\xi}}-\boldsymbol{\xi}) = (\hat{\boldsymbol{\xi}}-\boldsymbol{\xi})'\mathbf{U}Diag(\frac{1}{\lambda_1},\cdots,\frac{1}{\lambda_m})\mathbf{U}'(\hat{\boldsymbol{\xi}}-\boldsymbol{\xi}) \tag{1.108}$$

Here, we took advantage of the *eingespace synthesis* of the matrix $\mathbf{A}'\mathbf{A} =: \mathbf{N}$ and $(\mathbf{A}'\mathbf{A})^{-1} =: \mathbf{N}^{-1}$. Such an inverse normal matrix is representing the *dispersion matrix* $D\{\hat{\boldsymbol{\xi}}\} = (\mathbf{A}'\mathbf{A})^{-1}\sigma^2 = \mathbf{N}^{-1}\sigma^2$.

$$\mathbf{U}\mathbf{U}' = \mathbf{I}_m \sim \mathbf{U} \in \mathrm{SO}(m) := \{\mathbf{U} \in \mathbb{R}^{n\times m} \mid \mathbf{U}\mathbf{U}' = \mathbf{I}_m, |\mathbf{U}| = +1\}$$

$$\mathbf{N} := \mathbf{A}'\mathbf{A} = \mathbf{U}Diag(\nu_1,\cdots,\nu_m)\mathbf{U}'$$

versus

$$\mathbf{N}^{-1} := (\mathbf{A}'\mathbf{A})^{-1} = \mathbf{U}Diag(\lambda_1,\cdots,\lambda_m)\mathbf{U}'$$

subject to

$$\nu_1 = \lambda_1^{-1},\cdots,\nu_m = \lambda_m^{-1} \;\; or \;\; \lambda_1 = \nu_1^{-1},\cdots,\lambda_m = \nu_m^{-1}$$

$$|\mathbf{A}'\mathbf{A}|^{1/2} = \sqrt{\nu_1\cdots\nu_m} = \frac{1}{\sqrt{\lambda_1\cdots\lambda_m}}$$

$$\hat{\boldsymbol{\eta}}-\boldsymbol{\eta} := \mathbf{U}'(\hat{\boldsymbol{\xi}}-\boldsymbol{\xi}) \Leftrightarrow \hat{\boldsymbol{\xi}}-\boldsymbol{\xi} := \mathbf{U}(\hat{\boldsymbol{\eta}}-\boldsymbol{\eta}) \tag{1.109}$$

$$\|\hat{\boldsymbol{\xi}}-\boldsymbol{\xi}\|^2_{\mathbf{A}'\mathbf{A}} =: (\hat{\boldsymbol{\xi}}-\boldsymbol{\xi})'\mathbf{A}'\mathbf{A}(\hat{\boldsymbol{\xi}}-\boldsymbol{\xi}) = (\hat{\boldsymbol{\eta}}-\boldsymbol{\eta})'Diag(\frac{1}{\lambda_1},\cdots,\frac{1}{\lambda_m})(\hat{\boldsymbol{\eta}}-\boldsymbol{\eta}) \tag{1.110}$$

The local m-dimensional hypervolume element $d\hat{\xi}_1\cdots d\hat{\xi}_m$ is transformed to the local m-dimensional hypervolume element $d\hat{\eta}_1\cdots d\hat{\eta}_m$

$$d\hat{\xi}_1\cdots d\hat{\xi}_m = J_{\hat{\boldsymbol{\xi}}\to\hat{\boldsymbol{\eta}}}\, d\hat{\eta}_1\cdots d\hat{\eta}_m, \tag{1.111}$$

in which the Jocobian of the orthonormal transformation $J_{\hat{\boldsymbol{\xi}}\to\hat{\boldsymbol{\eta}}} = |\mathbf{U}| = 1$.

Accordingly, with the orthonormal transformation (1.109) we get the *cumulative probability* of the canonical parameters $\hat{\boldsymbol{\eta}}(\hat{\eta}_1,\hat{\eta}_2,\cdots,\hat{\eta}_m)$ from (1.93)

$$dF = f(\hat{\xi}_1,\hat{\xi}_2,\cdots,\hat{\xi}_m)d\hat{\xi}_1 d\hat{\xi}_2\cdots d\hat{\xi}_m = J_{\hat{\boldsymbol{\xi}}\to\hat{\boldsymbol{\eta}}}\, f(\hat{\eta}_1,\hat{\eta}_2,\cdots,\hat{\eta}_m)d\hat{\eta}_1 d\hat{\eta}_2\cdots d\hat{\eta}_m$$

i.e.,

$$dF = (\frac{1}{2\pi})^{\frac{m}{2}}\frac{1}{\sigma^m\sqrt{\lambda_1\cdots\lambda_m}}\exp\{-\frac{1}{2\sigma^2}(\hat{\boldsymbol{\eta}}-\boldsymbol{\eta})'Diag(\frac{1}{\lambda_1},\cdots,\frac{1}{\lambda_m})(\hat{\boldsymbol{\eta}}-\boldsymbol{\eta})\}d\hat{\eta}_1\cdots d\hat{\eta}_m \tag{1.112}$$

which alternatively leads us to

$$f(\hat{\boldsymbol{\eta}}) = (\frac{1}{2\pi})^{\frac{m}{2}}\frac{1}{\sigma^m}\frac{1}{\sqrt{\lambda_1\cdots\lambda_m}}\exp\{-\frac{1}{2\sigma^2}\sum_{i=1}^{m}\frac{(\hat{\eta}_i-\eta)^2}{\lambda_i}\} \tag{1.113}$$

$$f(\hat{\eta}_1,\cdots,\hat{\eta}_m) = f(\hat{\eta}_1)\cdots f(\hat{\eta}_m)$$

$$f(\hat{\eta}_i) := \frac{1}{\sqrt{2\pi}\sigma\sqrt{\lambda_i}}\exp\{-\frac{1}{2\sigma^2}\frac{(\hat{\eta}_i-\eta)^2}{\lambda_i}\} \text{ for all } i\in\{1,\cdots,m\} \tag{1.114}$$

Obviously the transformed random variables $(\hat{\eta}_1,\cdots,\hat{\eta}_m)$ represent BLUUE of $(\eta_1,\cdots,\eta_m)$ and are mutually independent following a *Gauss-Laplace normal distribution* $\hat{\eta}_i \sim \mathcal{N}(\eta_i, \sigma^2\lambda_i)$; in particular $E\{\hat{\boldsymbol{\eta}}\} = \boldsymbol{\eta}$ and $D\{\hat{\boldsymbol{\eta}}\} = \sigma^2 Diag(\lambda_1,\cdots\lambda_m)$. Furthermore $z_i = (\hat{\eta}_i-\eta_i)/(\sigma^2\lambda_i)^{1/2}$ are independently distributed as $\mathcal{N}(0, 1)$ and z_i^2 has the *Helmert Chi-square* distribution with 1 degree of freedom, $z_i^2 \sim \chi_1^2$.

In summary of these derivations we can formulate the following theorem which is complementary to *Theorem 1.16*.

Theorem 1.16 (marginal probability distributions of the orthonormally transformed parameters, special linear Gauss-Markov model):

By means of *Principal Component Analysis* (PCA), based on the called *Singular* Value Decomposition *(SVD) or* Eigenvalue Analysis *(EIGEN)* of $(\mathbf{A}'\mathbf{A})^{-1}$,

$$\begin{aligned}\boldsymbol{\eta} &= \mathbf{U}'_\xi \boldsymbol{\xi} \\ \hat{\boldsymbol{\eta}} &= \mathbf{U}'_\xi \hat{\boldsymbol{\xi}}\end{aligned} \quad \text{subject to} \quad \begin{bmatrix}\mathbf{U}'_\xi (\mathbf{A}'\mathbf{A})^{-1}\mathbf{U}_\xi = \boldsymbol{\Lambda} = Diag(\lambda_1, \lambda_2, \cdots, \lambda_{m-1}, \lambda_m) \\ \mathbf{U}'_\xi \mathbf{U}_\xi = \mathbf{I}_m, \det \mathbf{U}_\xi = +1\end{bmatrix}$$

the canonical *fixed effects* $(\hat{\eta}_1, \cdots, \hat{\eta}_m)$ become BLUUE of $(\eta_1, \cdots, \eta_m)$ can be orthonormally transformed from the BLUUE $(\hat{\xi}_1, \cdots, \hat{\xi}_m)$ of the *fixed effects* $(\xi_1, \cdots, \xi_m)$ within the special *Gauss-Markov* model of *Theorem 1.11*. The marginal pdf of the canonical parameters $\hat{\boldsymbol{\eta}}(\hat{\eta}_1, \hat{\eta}_2, \cdots, \hat{\eta}_m)$ is represented by

$$f_1(\hat{\boldsymbol{\eta}}) = \frac{1}{(2\pi)^{m/2}(\sigma^2)^{m/2}}(\lambda_1\lambda_2\cdots\lambda_{m-1}\lambda_m)^{-1/2}\exp\{-\frac{1}{2\sigma^2}\sum_{i=1}^m \frac{(\hat{\eta}_i - \eta)^2}{\lambda_i}\} \tag{1.115}$$

$$f_1(\hat{\boldsymbol{\eta}} \mid \boldsymbol{\eta}, \boldsymbol{\Lambda}\sigma^2) = f(\hat{\eta}_1) f(\hat{\eta}_2) \cdots f(\hat{\eta}_{m-1}) f(\hat{\eta}_m)$$

$$f(\hat{\eta}_i) = \frac{1}{\sigma\sqrt{\lambda_i}\sqrt{2\pi}}\exp\{-\frac{1}{2\sigma^2}\frac{(\hat{\eta}_i - \eta_i)^2}{\lambda_i}\} \text{ for all } i \in \{1, \cdots, m\} \tag{1.116}$$

and the *transformed fixed effects* $(\hat{\eta}_1, \cdots, \hat{\eta}_m)$ are mutually independent, following a *Gauss-Laplace normal distribution*

$$\hat{\eta}_i \sim \mathcal{N}(\eta_i \mid \sigma^2 \lambda_i) \text{ for all } i \in \{1, \cdots, m\}$$

$$z_i := \frac{\hat{\eta}_i - \eta}{\sqrt{\sigma^2 \lambda_i}} \Rightarrow f_1(z_i) dz_i = \frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}z_i^2) dz_i \text{ for all } i \in \{1, \cdots, m\}.$$

Chapter 2

Hypothesis tests of sample mean vector and sample variance-covariance matrix of a three-dimensional, symmetric rank-two random tensor

The statistical inference includes the point estimation derived in Chapter 1 and the hypothesis test. On the basis of the sampling distributions derived in Chapter 1 we will develop the distribution of multivariate test statistics for the testing of hypotheses concerning the sample mean vector and the sample variance covariance matrix, i.e. the estimated parameters (mean vector and variance-covariance matrix) of a tensor-valued multivariate normal population of a three-dimensional, symmetric rank-two random tensor, which include:

(1) Tests on $\boldsymbol{\mu}$ with $\boldsymbol{\Sigma}$ known (χ^2-test);
(2) Tests on $\boldsymbol{\mu}$ with $\boldsymbol{\Sigma}$ unknown (Hotelling's T^2-test);
(3) Tests on the equality of two mean vectors with common variance-covariance matrix (*Hotelling's two-sample* T^2 test and *Wilks'* Λ test);
(4) Tests if the variance-covariance matrix is equal to a given matrix (*likelihood ratio statistics*);
(5) Tests on the equality of two variance-covariance matrices (*likelihood ratio statistics*);
(6) Tests if the mean vectors and variance-covariance matrices are equal to a given vector and matrix (*likelihood ratio statistics*);
(7) Tests on the equality of two mean vectors and the respective variance-covariance matrices (*likelihood ratio statistics*).

2.1 Hypothesis test of the sample mean vector of a symmetric random tensor

There is one major area of statistical inference, the testing of hypotheses, which relates to the moments of a probability distribution. In experimental research one may wish to compare the yield of the new line with that of a standard line, and perhaps recommend the new line to replace the standard line, if it appears superior; this is a common situation in research.

Definition 2.1 (statistical hypothesis test of a statistical hypothesis)

A statistical hypothesis $\mathcal{H}$ is an assertion or conjecture about the distribution of one or more random variables. A test of a statistical hypothesis H is a rule or procedure for deciding whether to reject $\mathcal{H}$.

Concerning the testing, two hypotheses are discussed: The first, the hypothesis to be tested, is called the *null hypothesis*, denoted by $\mathcal{H}_0$, and the second is called the *alternative hypothesis*, denoted by $\mathcal{H}_1$. The thinking is that if the null hypothesis is *false*, then the alternative hypothesis is true, and vice versa. We often say that $\mathcal{H}_0$ is treated against, or versus, $\mathcal{H}_1$. If the null hypothesis is not rejected we say that $\mathcal{H}_0$ is accepted. With this kind of thinking, two types of errors can be made.

Definition 2.2 (types of errors, size of error)

Rejection of $\mathcal{H}_0$ when it is true is called a *Type I error*, and acceptance of $\mathcal{H}_0$ when it is false and $\mathcal{H}_1$ is true instead, is called a *Type II error*. The size of a *Type I error* is defined to be the probability α that a *Type I error* is made, and similarly the size of a *Type II error* is the probability 1-β that a *Type II error* is made in regards of $\mathcal{H}_1$.

The point of departure for hypotheses testing is the definition of a *test quantity*. For instance, for a random sample $\{y_1, ..., y_n\}$ of size n from a known probability distribution characterised by parameters like the non-centralized statistical moments $\mu_1, ..., \mu_m$ or the centralized statistical moments $\pi_1 ..., \pi_m$ of order m, we may choose a functional of the sample mean $\hat{\mu}_1 = \hat{\mu}$ or of the sample variance $\hat{\pi}_2 = \hat{\sigma}^2$ as a test quantity which is in general a function

$$
\begin{aligned}
&t = t(y_1,...,y_n;\mu_1,...,\mu_m) \in \mathbb{T} \text{ or} \\
&t = t(y_1,...,y_n;\pi_1,...,\pi_m) \in \mathbb{T} \text{ or} \\
&t = t(y_1,...,y_n;\mu_1,\pi_2 ...,\mu_{m-1},\pi_m) \in \mathbb{T}
\end{aligned}
\tag{2.1}
$$

of the set $\{y_1, ..., y_n\}$ of observations and the set $\{\mu_1, ..., \mu_m\}$, $\{\pi_1 ..., \pi_m\}$, or $\{\mu_1, \pi_2 ..., \mu_{m-1}, \pi_m\}$ of parameters of the probability distribution. *The probability distribution of the test functions*

$$f(t, \mu_1, ..., \mu_m),$$

in short the *test statistics*, is known. The test quantity is called a pivotal quantity if its probability distribution does not depend on the elements of the set $\{\mu_1, ..., \mu_m\}$ or $\{\pi_1 ..., \pi_m\}$ of parameters.

$$\begin{aligned} &\mathcal{H}_{01} : \upsilon \le \upsilon^0 \text{ versus } \mathcal{H}_{11} : \upsilon = \upsilon^1 > \upsilon^0 \\ &\mathcal{H}_{02} : \upsilon \ge \upsilon^0 \text{ versus } \mathcal{H}_{12} : \upsilon = \upsilon^1 < \upsilon^0 \\ &\mathcal{H}_{03} : \upsilon = \upsilon^0 \text{ versus } \mathcal{H}_{13} : \upsilon = \upsilon^1 \ne \upsilon^0. \end{aligned} \tag{2.2}$$

The *null hypothesis* $\mathcal{H}_0$ is formulated by means of a choice of the non-observable parameters, say υ^0, representing *either* an element of the set $\{\mu_1, ..., \mu_m\}$ of parameters (non-centralized first order statistical moments) *or* an element of the set $\{\{\pi_1 ..., \pi_m\}$ of parameters (centralized second order statistical moments). In short, we write

$$\mathcal{H}_{01} : \upsilon \le \upsilon^0,\ \mathcal{H}_{02} : \upsilon \ge \upsilon^0,\ \mathcal{H}_{03} : \upsilon = \upsilon^0 .$$

Alternatively we choose

$$\mathcal{H}_{11} : \upsilon = \upsilon^1 > \upsilon^0,\ \mathcal{H}_{12} : \upsilon = \upsilon^1 < \upsilon^0,\ \mathcal{H}_{13} : \upsilon = \upsilon^1 \ne \upsilon^0 .$$

Accordingly, such a test is called a *right one-sided test*, a *left one-sided test* and a *two-sided test*, respectively. This notion will become more obvious when we determine the probabilities α and 1-β of the Type I errors and Type II errors, for example for the two-sided test

$$\mathcal{H}_{03} : \upsilon = \upsilon^0 \text{ versus } \mathcal{H}_{13} : \upsilon = \upsilon^1 \ne \upsilon^0 .$$

The null hypothesis $\mathcal{H}_{03}$ is accepted if the test quantity t is an element of the *acceptance region* $c_1 \le t \le c_2$, $c_1 = c_{\alpha/2}$, $c_2 = c_{1-\alpha/2}$. The *critical values* c_1 and c_2 are determined from the probability identity

$$\begin{aligned} P\{c_1 \le t \le c_2; \upsilon^0\} &= P\{t \le c_2; \upsilon^0\} - P\{t \le c_1; \upsilon^0\} = \\ &= \int_{-\infty}^{c_2} f_{\upsilon^0}(t)dt - \int_{-\infty}^{c_1} f_{\upsilon^0}(t)dt = \\ &= (1 - \frac{\alpha}{2}) - \frac{\alpha}{2} = 1 - \alpha \end{aligned} \tag{2.3}$$

with respect to a significance level α, namely by a linear *Volterra integral equation of the first kind*. In contrast, $\mathcal{H}_{03}$ is rejected if the test quantity is "*out of the acceptance interval*" $t \notin [c_1, c_2]$, in particular if t is an element of the *rejection region or critical region* $\mathbb{C}_{lr} := \{t \mid -\infty < t \le c_{\alpha/2}, c_{1-\alpha/2} \le t < +\infty\}$. The probability to reject a true null hypothesis, that is to make a *Type I error*, is measured by the *error probability* α, e.g. 1%, 5%, or 10%, respectively.

$$P\{\text{Type I error }\} = 1 - P\{c_1 \le t \le c_2; \upsilon^0\} = \alpha \tag{2.4}$$

The specific *alternative hypothesis* $\mathcal{H}_{13}$ is validated by the probability identity

$$\begin{aligned} P\{t < c_1; \upsilon^1\} + P\{t > c_2; \upsilon^1\} &= 1 + P\{t < c_1; \upsilon^1\} - P\{t \le c_2; \upsilon^1\} = \\ &= 1 + \int_{-\infty}^{c_1} f_{\upsilon^1}(t)dt - \int_{-\infty}^{c_2} f_{\upsilon^1}(t)dt = \\ &= 1 + F_{\upsilon^1}(c_1) - F_{\upsilon^1}(c_2) \\ &=: \beta \end{aligned} \tag{2.5}$$

which is called the power of the two-sided test. β is a measure of the probability to reject the null hypothesis $\mathcal{H}_{03}$ in favor of the specific alternative hypothesis $\mathcal{H}_{13}$. In contrast,

$$P\{\text{Type II error }\} = P\{t \le c_2; \upsilon^1\} - P\{t \le c_1; \upsilon^1\} = 1 - \beta \tag{2.6}$$

is a measure for the probability to reject the specific hypothesis $\mathcal{H}_{13}$ in favor of the false null hypothesis $\mathcal{H}_{03}$, that is to commit a Type II error.

The following table illustrates the rationale of hypothesis testing.

	$\mathcal{H}_0$ is true	Specific $\mathcal{H}_1$ is true
Acceptance of $\mathcal{H}_0$	correct decision P\{correct decision\}=1-α	Type II error P\{ Type II error\}=1-β
Rejection of $\mathcal{H}_0$	Type I error P \{Type I error\} = α significance level	correct decision P\{correct decision\}=β (power of test)

While P\{Type I error \} measures the probability to reject a true *null hypothesis* $\mathcal{H}_0$, P\{Type II error \} measures the probability to reject the true alternative hypothesis $\mathcal{H}_1$.

The hypothesis test of sample mean vector and sample variance-covariance matrix of a symmetric random tensor belongs to multivariate analysis which is the branch of statistics devoted to the study of random variables that are not necessarily independent. Where inference is concerned, several (generally correlated) measurements are made on every observed subject.

Many current multivariate statistical procedures were developed during the first half of the twentieth century. A reasonably complete list of the developers would by voluminous. However, a few individuals can be cited as having made important initial contributions to the theory and practice of multivariate analysis.

T. Galton and *K. Pearson* did pioneering work in the areas of correlation and regression analysis. *R.A. Fisher's* derivation of the exact distribution of the sample correlation coefficient and related quantities provided the impetus for multivariate distribution theory. *C. Spearman* and *K. Pearson* were among the first to work in the area of factor analysis. Significant contributions to multivariate analysis were made during the 1930s by. *S. S. Wilks* (general procedures for testing certain multivariate hypotheses), *H. Hotelling* (*Hotelling's* T^2, *principle component analysis, canonical correlation analysis*), *R. A. Fisher* (*discrimination and classification*), and *P. C. Mahalanobis* (generalized distance, hypothesis testing). *J. Wishart* derived an important joint distribution of sample variance and covariance that bears his name. Later *M. Bartlett* and *G. E. P. Box* contributed to the large sample theory associated with certain multivariate test statistics.

The body of statistical methodology used to analyze simultaneous measurements on many variables is called multivariate analysis. Many multivariate methods are based on an underlying probability model known as the multivariate normal distribution.

The objectives of scientific investigations, for which multivariate methods most naturally lend themselves, include the following:

- Data reduction or structural simplification.
- Sorting and grouping.
- Investigation of the dependence among variables.
- Predication.
- Hypothesis construction and testing.

One of the central messages of multivariate analysis is that p correlated variables must be analyzed jointly. This principle is exemplified by the methods presented in *Section* 2.1 and 2.2. Inference, that is, reaching valid conclusions on the basis of sample information. While real data are never exactly multivariate normal, the normal density is often a useful approximation to the "true" population distribution.

On the basis of the sampling distributions derived in *Section* 1.4 and 1.5 the distribution of multivariate test statistics needed for testing hypotheses concerning the parameters (covariance matrix and mean vector) of a tensor-valued multivariate normal population, such as *Hotelling's* T^2 and likelihood ratio statistics, are developed.

At this point we shall concentrate on inferences about a population mean vector and its component parts in this section, although we introduce statistical analysis of the component means based on simultaneous confidence statements. In *Section* 2.2 we shall discuss the hypothesis test for the sample variance-covariance matrix of a symmetric random tensor. From the many books about these subjects, we refer to Grafarend (2000) for the univariate hypothesis test and *Giri* (1977), *Muirhead* (1982), *Rencher* (1995, 1998), *Anderson* (1958, 1984) and *Srivastava* (2002).

2.1.1 Tests on $\boldsymbol{\mu}$ with $\boldsymbol{\Sigma}$ known

Hypothesis testing has started with the test on the mean with known or unknown variance, in particular for the random sample of a *Gauß-Laplace normal distribution*. The test on a sample mean vector assuming a known $\boldsymbol{\Sigma}$ is introduced to illustrate the issues involved in multivariate testing and serve as foundation for the unknown $\boldsymbol{\Sigma}$ case. We do not consider one-sided alternative hypotheses because they do not readily generalize to multivariate tests.

For the univariate case the hypothesis test of interest is that the mean is equal to a given value μ_0, versus the alternative that it is not equal to μ_0.

$$\mathcal{H}_0 : \mu = \mu_0 \;\; versus \;\; \mathcal{H}_1 : \mu = \mu_1 \neq \mu_0$$

Let $\{y_1, ..., y_n\}$ be a set of independently identically distributed (i.i.d.) observations from a normal sample of size n with the unknown mean value μ and the known variance σ^2.

$$t(y_1,...,y_n;\mu_0) := \frac{\hat{\mu}-\mu_0}{\sigma/\sqrt{n}} = \frac{\sqrt{n}}{\sigma}[\frac{y_1+...+y_n}{n} - \mu_0] \tag{2.7}$$

with respect to the sample mean $\hat{\mu}$ of type BLUUE. $t(\hat{\mu};\mu_0)$ is Gauß-Laplace standard normally distributed with mean zero and variance one. The probability identity

$$P\{-c \le t \le +c\} = P\{\hat{\mu} - c\frac{\sigma}{\sqrt{n}} \le \mu_0 \le \hat{\mu} + c\frac{\sigma}{\sqrt{n}}\} = 1-\alpha = \gamma \tag{2.8}$$

relates the error probability α of the two-sided test to the confidence level γ. If μ_0 is an element of the confidence interval $\hat{\mu} - c\sigma/\sqrt{n} \le \mu_0 \le \hat{\mu} + c\sigma/\sqrt{n}$, the null hypothesis $\mathcal{H}_0 : \mu = \mu_0$ is accepted. We eject $\mathcal{H}_0$ if the confidence interval does not contain μ_0.

Equivalently, we can use the statistic $t^2(y_1,...,y_n;\mu_0)$

$$t^2(y_1,...,y_n;\mu_0) := n[\hat{\mu}-\mu_0](\sigma^2)^{-1}[\hat{\mu}-\mu_0], \tag{2.9}$$

which is distributed as χ^2 with one degree of freedom. The probability identity

$$P\{t^2 \le +c\} = P\{t^2 \le \chi^2_{1,1-\alpha}\} = 1-\alpha = \gamma . \tag{2.10}$$

In the multivariate analysis of the sample mean vector for a symmetric random tensor in vector form we wish to hypothesize the value of the mean vector jointly when the variance-covariance matrix is known :

$$\mathcal{H}_{01}: \; \boldsymbol{\mu} = \boldsymbol{\mu}_0, \mathcal{H}_{11}: \; \boldsymbol{\mu} \neq \boldsymbol{\mu}_0 \quad \text{with } \boldsymbol{\Sigma}_{\mathbf{y}} \text{ known}$$

More explicitly, for the three-dimensional, symmetric rank-two random tensor discussed in *Section 1.3* we have

$$\mathcal{H}_{01}: \begin{bmatrix}\mu_1\\ \vdots \\ \mu_6\end{bmatrix} = \begin{bmatrix}\mu_{01}\\ \vdots \\ \mu_{06}\end{bmatrix}, \quad \mathcal{H}_{11}: \begin{bmatrix}\mu_1\\ \vdots \\ \mu_6\end{bmatrix} \neq \begin{bmatrix}\mu_{01}\\ \vdots \\ \mu_{06}\end{bmatrix} \quad \text{with } \boldsymbol{\Sigma}_{\mathbf{y}} \text{ known}.$$

The vector equality in $\mathcal{H}_{01}$ implies $\mu_i = \mu_{0i}$ for all $i = 1, \cdots, 6$. The vector inequality in $\mathcal{H}_{11}$ implies $\mu_i \neq \mu_{0i}$ for at least one $i \in \{1,\cdots,6\}$.

To test $\mathcal{H}_{01}$, we use a sample of N observations on $\mathbf{t}$, namely $\mathbf{t}_1, \mathbf{t}_2, \ldots , \mathbf{t}_n,$ whose related vectorized forms $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n$ are distributed according to $\mathcal{N}_6(\boldsymbol{\mu}, \boldsymbol{\Sigma}_{\mathbf{y}})$ of *Section* 1.3. The test statistic is

$$Z^2(\mathbf{y}_1,...,\mathbf{y}_n;\boldsymbol{\mu}_0) := n[\hat{\boldsymbol{\mu}}_{\mathbf{y}} - \boldsymbol{\mu}_0]\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}[\hat{\boldsymbol{\mu}}_{\mathbf{y}} - \boldsymbol{\mu}_0], \tag{2.11}$$

which is distributed as χ^2_6 by *Lemma 1.8* and *Theorem 1.10*. We reject $\mathcal{H}_{01}$ if $Z^2 > \chi^2_{6,\,1-\alpha}$. Thus, for one variable, it will refer to (2.10), whereas for the case of a 3×3 symmetric random tensor, Z^2 of (2.11) has a chi-square distribution with six degrees of freedom.

Since we cannot get the expectation of the variance-covariance matrix from the observation of deformation measures in our real experience, this test statistic is not very practical in our case.

2.1.2 Tests on $\boldsymbol{\mu}$ with $\boldsymbol{\Sigma}$ unknown

Firstly let us review briefly the familiar one–sample t-test in the univariate case.

The hypothesis test of interest is that the mean is equal to a given value μ_0, versus the alternative that it is not equal to μ_0.

$$\mathcal{H}_0 : \mu = \mu_0 \;\; versus \;\; \mathcal{H}_1 : \mu = \mu_1 \neq \mu_0$$

Let $\{y_1, ..., y_n\}$ be a set of independently identically distributed (i.i.d.) observation from a normal sample of size n with the unknown mean value μ and unknown variance σ^2. The test statistic

$$t(y_1,...,y_n;\mu_0) := \frac{\hat{\mu}-\mu_0}{\hat{\sigma}/\sqrt{n}} = \frac{\sqrt{n}}{\hat{\sigma}}[\frac{y_1+...+y_n}{n}-\mu_0] \tag{2.12}$$

where $\hat{\mu}$, $\hat{\sigma}^2$ represent the sample mean, and sample variance of type BLUUE and type BIQUUE, respectively. (2.12) has a *Student t*-distribution with n-1 degrees of freedom. The probability identity

$$P\{-c \leq t \leq +c\} = P\{\hat{\mu} - c\frac{\hat{\sigma}}{\sqrt{n}} \leq \mu_0 \leq \hat{\mu} + c\frac{\hat{\sigma}}{\sqrt{n}}\} = 1-\alpha = \gamma$$

relates the error probability α of the two-sided test to the confidence level γ. If μ_0 is an element of the confidence interval $\hat{\mu} - c\hat{\sigma}/\sqrt{n} \leq \mu_0 \leq \hat{\mu} + c\hat{\sigma}/\sqrt{n}$, the null hypothesis $\mathcal{H}_0 : \mu = \mu_0$ is accepted. We reject $\mathcal{H}_0$ if the confidence interval does not contain μ_0.

Secondly in the multivariate analysis of the sample mean vector for a symmetric random tensor in vector form we wish to hypothesize the value of the mean vector jointly when the variance-covariance matrix is unknown :

$$\mathcal{H}_{02}: \; \boldsymbol{\mu} = \boldsymbol{\mu}_0, \; \mathcal{H}_{12}: \; \boldsymbol{\mu} \neq \boldsymbol{\mu}_0 \quad \text{with } \boldsymbol{\Sigma}_{\mathrm{y}} \text{ unknown.}$$

More explicitly, for the three-dimensional, symmetric rank-two random tensor discussed in *Section 1.3* we have

$$\mathcal{H}_{02}: \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_6 \end{bmatrix} = \begin{bmatrix} \mu_{01} \\ \vdots \\ \mu_{06} \end{bmatrix}, \quad \mathcal{H}_{12}: \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_6 \end{bmatrix} \neq \begin{bmatrix} \mu_{01} \\ \vdots \\ \mu_{06} \end{bmatrix} \quad \text{with } \boldsymbol{\Sigma}_{\mathrm{y}} \text{ unknown}.$$

The vector equality in $\mathcal{H}_{02}$ implies $\mu_i = \mu_{0i}$ for all $i = 1, \cdots, 6$. The vector inequality in $\mathcal{H}_{12}$ implies $\mu_i \neq \mu_{0i}$ for at least one $i \in \{1, \cdots, 6\}$.

To test $\mathcal{H}_{02}$, we use a sample of N observations on **t**, namely $\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_n$, whose related vectorized forms $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n$ are distributed according to $\mathcal{N}_6(\boldsymbol{\mu}, \boldsymbol{\Sigma}_{\mathrm{y}})$ giving rise to the sample mean vector $\hat{\boldsymbol{\mu}}_{\mathrm{y}}$ of *Theorem 1.10* to the sample covariance matrix $\hat{\boldsymbol{\Sigma}}_{\mathrm{y}}$ of *Theorem 1.12*. Hotelling's T^2 statistic (*Hotelling* 1931) is defined as

$$T^2 := n[\hat{\boldsymbol{\mu}}_{\mathrm{y}} - \boldsymbol{\mu}_0]'\hat{\boldsymbol{\Sigma}}_{\mathrm{y}}^{-1}[\hat{\boldsymbol{\mu}}_{\mathrm{y}} - \boldsymbol{\mu}_0]. \tag{2.13}$$

which is distributed as $T^2_{6,\, n-1}$. Note, that for one variable T^2 is the square of the usual *Student t-statistic* (2.12). In general, it is clear that $T^2 \geq 0$ and if $\boldsymbol{\mu}_0 = \mathbf{0}$ then $\hat{\boldsymbol{\mu}}_{\mathrm{y}}$ should be close to **0**, and so should be T^2. This characteristic is one of the most important properties of the *Wishart* distribution (*Theorem 1.11*). Now we should derive its relationship with the *F* distribution.

With the sampling distribution of $\hat{\boldsymbol{\mu}}_{\mathrm{y}}$ and $\hat{\boldsymbol{\Sigma}}_{\mathrm{y}}$, and the independence of them , we write T^2/(n-1) as

$$\frac{T^2}{n-1} = \frac{n[\hat{\boldsymbol{\mu}}_{\mathrm{y}} - \boldsymbol{\mu}_0]'\hat{\boldsymbol{\Sigma}}_{\mathrm{y}}^{-1}[\hat{\boldsymbol{\mu}}_{\mathrm{y}} - \boldsymbol{\mu}_0]}{(n-1)[\hat{\boldsymbol{\mu}}_{\mathrm{y}} - \boldsymbol{\mu}_0]'\boldsymbol{\Sigma}_{\mathrm{y}}^{-1}[\hat{\boldsymbol{\mu}}_{\mathrm{y}} - \boldsymbol{\mu}_0]}[\hat{\boldsymbol{\mu}}_{\mathrm{y}} - \boldsymbol{\mu}_0]'\boldsymbol{\Sigma}_{\mathrm{y}}^{-1}[\hat{\boldsymbol{\mu}}_{\mathrm{y}} - \boldsymbol{\mu}_0]$$
$$= \frac{n[\hat{\boldsymbol{\mu}}_{\mathrm{y}} - \boldsymbol{\mu}_0]'\boldsymbol{\Sigma}_{\mathrm{y}}^{-1}[\hat{\boldsymbol{\mu}}_{\mathrm{y}} - \boldsymbol{\mu}_0]}{(n-1)[\hat{\boldsymbol{\mu}}_{\mathrm{y}} - \boldsymbol{\mu}_0]'\boldsymbol{\Sigma}_{\mathrm{y}}^{-1}[\hat{\boldsymbol{\mu}}_{\mathrm{y}} - \boldsymbol{\mu}_0]/\{[\hat{\boldsymbol{\mu}}_{\mathrm{y}} - \boldsymbol{\mu}_0]'\hat{\boldsymbol{\Sigma}}_{\mathrm{y}}^{-1}[\hat{\boldsymbol{\mu}}_{\mathrm{y}} - \boldsymbol{\mu}_0]\}}$$

Note that a key assumption in the T^2 distribution is the independence of $\hat{\boldsymbol{\mu}}_{\mathrm{y}}$ and $\hat{\boldsymbol{\Sigma}}_{\mathrm{y}}$, which holds when sampling from a multivariate normal population (*Theorem 1.11*).

For the three-dimensional, symmetric rank-two random tensor case discussed in *Section 1.1* and after *Muirhead* (1982, p.96) we have

$$n[\hat{\boldsymbol{\mu}}_{\mathrm{y}} - \boldsymbol{\mu}_0]'\boldsymbol{\Sigma}_{\mathrm{y}}^{-1}[\hat{\boldsymbol{\mu}}_{\mathrm{y}} - \boldsymbol{\mu}_0] \sim \chi^2_6, \quad (n-1)\frac{[\hat{\boldsymbol{\mu}}_{\mathrm{y}} - \boldsymbol{\mu}_0]'\boldsymbol{\Sigma}_{\mathrm{y}}^{-1}[\hat{\boldsymbol{\mu}}_{\mathrm{y}} - \boldsymbol{\mu}_0]}{[\hat{\boldsymbol{\mu}}_{\mathrm{y}} - \boldsymbol{\mu}_0]'\hat{\boldsymbol{\Sigma}}_{\mathrm{y}}^{-1}[\hat{\boldsymbol{\mu}}_{\mathrm{y}} - \boldsymbol{\mu}_0]} \sim \chi^2_{n-1-6+1}$$

Dividing them each by their respect degree of freedom and using the definition of the *F* distribution shows that

$$\frac{T^2}{n-1}\frac{n-1-6+1}{6} \sim F_{6,\ n-1-6+1}. \tag{2.14}$$

This is of great practical importance in testing hypotheses about the mean vector of the vectorized random tensor when the covariance matrix is unknown.

If the observed T^2 is too large – that is, $\hat{\boldsymbol{\mu}}_{\mathbf{y}}$ is "too far" from $\boldsymbol{\mu}_0$ - the hypothesis $H_{02} : \hat{\boldsymbol{\mu}}_{\mathbf{y}} = \boldsymbol{\mu}_0$ is rejected. Due to the relationship (2.14) we can calculate the probability identity

$$\begin{aligned} &P\{T^2 \le \frac{(n-1)\cdot 6}{n-1-6+1} F_{6,\ n-1-6+1}(1-\alpha)\} = \\ &= P\{n[\hat{\boldsymbol{\mu}}_{\mathbf{y}} - \boldsymbol{\mu}_0]\hat{\boldsymbol{\Sigma}}_{\mathbf{y}}^{-1}[\hat{\boldsymbol{\mu}}_{\mathbf{y}} - \boldsymbol{\mu}_0] \le \frac{(n-1)\cdot 6}{n-1-6+1} F_{6,\ n-1-6+1}(1-\alpha)\} = 1-\alpha = \gamma \end{aligned} \tag{2.15}$$

where $F_{6,\ n-1-6+1}(1-\alpha)$ is the upper $(100\alpha)th$ percentile of the $F_{6,\ n-1-6+1}$ distribution. (2.15) leads immediately to a test of the hypothesis $\mathcal{H}_{02} : \hat{\boldsymbol{\mu}}_{\mathbf{y}} = \boldsymbol{\mu}_0$ versus $\mathcal{H}_{12} : \hat{\boldsymbol{\mu}}_{\mathbf{y}} \ne \boldsymbol{\mu}_0$. At the error probability α, reject $\mathcal{H}_{02}$ in favor of $\mathcal{H}_{12}$ if

$$T^2 = n[\hat{\boldsymbol{\mu}}_{\mathbf{y}} - \boldsymbol{\mu}_0]\hat{\boldsymbol{\Sigma}}_{\mathbf{y}}^{-1}[\hat{\boldsymbol{\mu}}_{\mathbf{y}} - \boldsymbol{\mu}_0] > \frac{(n-1)\cdot 6}{n-1-6+1} F_{6,\ n-1-6+1}(1-\alpha)\} \tag{2.16}$$

It has been shown that Hotelling's T^2 test is a uniformly most powerful invariant test (Anderson, 1984, p.183). Further it is also the likelihood ratio test.

This test is just our case of the repeated observations of deformation measures in one place or network with the same technique and the same conditions, in which we have only the estimates of the sample mean vector and the sample variance-covariance matrix. So we may use Hotelling's T^2 statistic (2.13) to test the sample mean vector.

2.1.3 Tests on equality of two mean vectors with common variance-covariance matrix

A T^2 test for testing the equality of the mean vectors from two multivariate populations can be developed by analogy with univariate procedure. See *Johnson and Wichern* (1988, p.221) and *Anderson* (1984, p.167) in detail.

In the multivariate case, we wish to compare the mean vectors from two population. This is also called as "*Hotelling's two-sample T^2 test*". We assume that two independent random samples $\mathbf{y}_{11}, \mathbf{y}_{12}, \cdots, \mathbf{y}_{1n_1}$ are distributed according to $\mathcal{N}_6(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$ and $\mathbf{y}_{21}, \mathbf{y}_{22}, \cdots, \mathbf{y}_{2n_2}$ are distributed according to $\mathcal{N}_6(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$ giving rise to the sample mean vectors $\hat{\boldsymbol{\mu}}_1$ and $\hat{\boldsymbol{\mu}}_2$ respectively, where $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ are unknown. In order to obtain a T^2 test, we must assume $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2$, which is of importance for the small sample size n_1 and n_2. From these two samples a pooled estimator of the common covariance matrix $\boldsymbol{\Sigma}$ is calculated as

$$\mathbf{S}_{pl} = \frac{(n_1-1)\hat{\boldsymbol{\Sigma}}_1 + (n_2-1)\hat{\boldsymbol{\Sigma}}_2}{n_1+n_2-2} \tag{2.17}$$

for which $E\{\mathbf{S}_{pl}\} = \boldsymbol{\Sigma}$. To test

$$\mathcal{H}_{03}: \ \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2,\ \mathcal{H}_{13}: \ \boldsymbol{\mu}_1 \ne \boldsymbol{\mu}_2 \quad \text{with } \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2.$$

we use the test statistic

$$T^2 := \frac{n_1 n_2}{n_1+n_2}[\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2]'\mathbf{S}_{pl}^{-1}[\hat{\boldsymbol{\mu}}_1 - \hat{\boldsymbol{\mu}}_2], \tag{2.18}$$

which is distributed as $T^2_{6,\ n_1+n_2-2}$ and its relationship with the *F-Statistic* is:

$$\frac{n_1+n_2-6-1}{3(n_1+n_2-2)}T^2 = F_{6,\ n_1+n_2-6-1}. \tag{2.19}$$

We can use the same procedure of *Section* 2.1.2 to make hypothesis test $\mathcal{H}_{03}: \ \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$.

Now we discuss the second method of deriving Hotelling's two-sample T^2 test, which is based on likelihood ratio test. We develop the decomposition (1.57) of *Section* 1.4 for the sample of one population to our two populations test. First let us note that

$$n = n_1 + n_2, \quad \hat{\boldsymbol{\mu}}_1 = \frac{1}{n_1}\sum_{j=1}^{n_1}\mathbf{y}_{1j}, \quad \hat{\boldsymbol{\mu}}_2 = \frac{1}{n_2}\sum_{j=1}^{n_2}\mathbf{y}_{2j}, \quad \hat{\boldsymbol{\mu}}_{12} = \frac{1}{n}\sum_{i=1}^{2} n_i\hat{\boldsymbol{\mu}}_i = \frac{n_1\hat{\boldsymbol{\mu}}_1 + n_2\hat{\boldsymbol{\mu}}_2}{n}$$

$$\mathbf{A}_1 = \sum_{j=1}^{n_1} (\mathbf{y}_{1j} - \hat{\boldsymbol{\mu}}_1)(\mathbf{y}_{1j} - \hat{\boldsymbol{\mu}}_1)', \quad \mathbf{A}_2 = \sum_{j=1}^{n_2} (\mathbf{y}_{2j} - \hat{\boldsymbol{\mu}}_2)(\mathbf{y}_{2j} - \hat{\boldsymbol{\mu}}_2)'$$

$$\mathbf{W} = \sum_{i=1}^{2} \sum_{j=1}^{n_i} (\mathbf{y}_{ij} - \hat{\boldsymbol{\mu}}_{12})'(\mathbf{y}_{ij} - \hat{\boldsymbol{\mu}}_{12})$$

Since

$$\mathbf{W} = \sum_{i=1}^{2} \sum_{j=1}^{n_i} (\mathbf{y}_{ij} - \hat{\boldsymbol{\mu}}_{12})(\mathbf{y}_{ij} - \hat{\boldsymbol{\mu}}_{12})' = \sum_{i=1}^{2} \sum_{j=1}^{n_i} (\mathbf{y}_{ij} - \hat{\boldsymbol{\mu}}_i)(\mathbf{y}_{ij} - \hat{\boldsymbol{\mu}}_i)' + \sum_{i=1}^{2} n_i (\hat{\boldsymbol{\mu}}_i - \hat{\boldsymbol{\mu}}_{12})(\hat{\boldsymbol{\mu}}_i - \hat{\boldsymbol{\mu}}_{12})' \tag{2.20}$$

we note that for

$$\mathbf{A} = \sum_{i=1}^{2} \sum_{j=1}^{n_i} (\mathbf{y}_{ij} - \hat{\boldsymbol{\mu}}_i)(\mathbf{y}_{ij} - \hat{\boldsymbol{\mu}}_i)', \quad \mathbf{B} = \sum_{i=1}^{2} n_i (\hat{\boldsymbol{\mu}}_i - \hat{\boldsymbol{\mu}}_{12})(\hat{\boldsymbol{\mu}}_i - \hat{\boldsymbol{\mu}}_{12})' \tag{2.21}$$

we have

$$\mathbf{W} = \mathbf{A} + \mathbf{B} \tag{2.22}$$

Under $\mathcal{H}_{03}: \ \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$, the corresponding random matrices **A, B** are independently distributed as $\mathcal{W}_6(n-2, \boldsymbol{\Sigma})$, $\mathcal{W}_6(1, \boldsymbol{\Sigma})$, and **W** is distributed as $\mathcal{W}_6(n-1, \boldsymbol{\Sigma})$, from which we obtain the famous test statistic-*Wilks' Lambda statistic*

$$\Lambda(6, \ n-2, 1) = \frac{\det \mathbf{A}}{\det(\mathbf{A}+\mathbf{B})} = \frac{\det \mathbf{A}}{\det \mathbf{W}} \tag{2.23}$$

which was first proposed by *Wilks* (1932) and later by *Hsu* (1941). We reject $\mathcal{H}_{03}$ if the ratio of generalized variances (2.23) is too small.

For the two sample case we have the likelihood function

$$L(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \ \boldsymbol{\Sigma}) = (2\pi)^{-6n/2} (\det \boldsymbol{\Sigma})^{-n/2} \operatorname{etr}\{-\frac{1}{2} \boldsymbol{\Sigma}^{-1} \sum_{i=1}^{2} \sum_{j=1}^{n_1} (\mathbf{y}_{ij} - \hat{\boldsymbol{\mu}}_i)(\mathbf{y}_{ij} - \hat{\boldsymbol{\mu}}_i)'\}.$$

The likelihood ratio statistic is

$$\Lambda_1 = \frac{\max_{\boldsymbol{\mu}, \boldsymbol{\Sigma}} L(\boldsymbol{\mu}, \boldsymbol{\mu}, \ \boldsymbol{\Sigma})}{\max_{\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\Sigma}} L(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2 \ \boldsymbol{\Sigma})}$$

in which

$$\max_{\boldsymbol{\mu}, \boldsymbol{\Sigma}} L(\boldsymbol{\mu}, \boldsymbol{\mu}, \ \boldsymbol{\Sigma}) = (2\pi)^{-3n/2} (\det \mathbf{W})^{-n/2} n^{-3n/2} \ \mathrm{e}^{-3n/2}$$

$$\max_{\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\Sigma}} L(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2 \ \boldsymbol{\Sigma}) = (2\pi)^{-3n/2} (\det \mathbf{A})^{-n/2} n^{-3n/2} \ \mathrm{e}^{-3n/2},$$

then for testing $\mathcal{H}_{03}: \ \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$ the likelihood ratio statistic is

$$\Lambda_1 = \frac{(\det \mathbf{A})^{n/2}}{(\det \mathbf{W})^{n/2}} \tag{2.24}$$

This test is equivalent to *Wilks'* Λ test of (2.23).

This test covers our case of repeated observations of deformation measures of two places or two networks with the same technique and the same conditions, in which we have only the estimates of the sample mean vectors and the sample variance-covariance matrices. So we have to use Hotelling's T^2 two-sample statistic (2.18) to test the sample mean vector.

When the variance-covariance matrices of two populations are not equal, the two-sample T^2 statistic in (2.18) does not have a T^2 distribution, which leads to the Behrens-Fisher problem (*Behrens* 1929, *Fisher* 1939). Since this situation often takes place in the case of repeated observations of deformation measures of two places or two networks with different techniques and under difference conditions. An optimal approximate solution of this problem remains under investigation.

In this section we have so far discussed three hypothesis tests about the sample mean of a symmetric random tensor, which are summarised in Box 2.1.

***Box* 2.1** (Hypothesis tests of the sample mean)

$$\mathcal{H}_{01}: \ \boldsymbol{\mu} = \boldsymbol{\mu}_0, \ \mathcal{H}_{11}: \ \boldsymbol{\mu} \neq \boldsymbol{\mu}_0 \quad \text{with } \boldsymbol{\Sigma}_{\mathbf{y}} \text{ known}$$
$$\mathcal{H}_{02}: \ \boldsymbol{\mu} = \boldsymbol{\mu}_0, \ \mathcal{H}_{12}: \ \boldsymbol{\mu} \neq \boldsymbol{\mu}_0 \quad \text{with } \boldsymbol{\Sigma}_{\mathbf{y}} \text{ unknown}$$
$$\mathcal{H}_{03}: \ \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2, \ \mathcal{H}_{13}: \ \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2 \quad \text{with } \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2$$

2.2 Hypothesis test of the sample variance-covariance matrix of a symmetric random tensor

Now we consider the hypothesis test of the sample variance-covariance matrix of a symmetric random tensor. These tests are often carried out to check the assumptions pertaining to other tests. We will cover four types of hypotheses: (1) the variance-covariance matrix is equal to a given matrix, (2) two variance-covariance matrices are equal, (3) the mean vector and the variance-covariance matrix are equal to a given vector and matrix, respectively, which is obviously the combination of the hypothesis test about the mean vector of *Section 2.1.3* and (1) of this section for the variance-covariance matrix, and more generally (4) several mean vectors and the variance-covariance matrices from several normal populations are equal. In most cases we use the likelihood ratio approach. The resulting test statistics are often determined by the ratio of determinants of the sample variance-covariance matrix.

In the univariate test of variance we are interested in (1) test on the variance-covariance matrix with mean known or unknown, which are related to the Chi-square distribution, and (2) test on the mean difference from two independent normal samples, which are related to the *Fisher F*-distribution. For more detail we referred to Koch (1997, 1999) and Grafarend (2000).

2.2.1 Tests if the variance-covariance matrix is equal to a given matrix

We are interested in testing if the variance-covariance matrix is equal to a given matrix

$$\mathcal{H}_{04}: \ \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0, \ \mathcal{H}_{14}: \ \boldsymbol{\Sigma} \neq \boldsymbol{\Sigma}_0$$

Let $\mathbf{y}_1, \cdots, \mathbf{y}_n$ be independent $\mathcal{N}_6(\boldsymbol{\mu}, \boldsymbol{\Sigma})$-distributed random vectors of vectorized random tensors with unknown mean vector $\boldsymbol{\mu}$ and variance-covariance matrix $\boldsymbol{\Sigma}$, and consider testing the null hypothesis $\mathcal{H}_{04}: \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0$, where $\boldsymbol{\Sigma}_0$ is a specified positive-definite matrix, against $\mathcal{H}_{14}: \boldsymbol{\Sigma} \neq \boldsymbol{\Sigma}_0$.

At first we assume that $\boldsymbol{\Sigma}_0 = \mathbf{I}_3$, then transform to the general $\boldsymbol{\Sigma}_0$. According to the *Definition 1.5* the likelihood function is

$$\begin{aligned} L(\boldsymbol{\mu}, \ \boldsymbol{\Sigma}) &= (2\pi)^{-6n/2} (\det \boldsymbol{\Sigma})^{-n/2} \exp\{-\frac{1}{2}\sum_{i=1}^{n} (\mathbf{y}_i - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} ((\mathbf{y}_i - \boldsymbol{\mu})\} = \\ &= (2\pi)^{-6n/2} (\det \boldsymbol{\Sigma})^{-n/2} \operatorname{etr}\{-\frac{1}{2} \boldsymbol{\Sigma}^{-1} [\mathbf{Y} - \mathbf{1}\boldsymbol{\mu}']' [\mathbf{Y} - \mathbf{1}\boldsymbol{\mu}']\}. \end{aligned} \tag{2.25}$$

in which $\mathbf{Y}' = [\mathbf{y}_1, \ \mathbf{y}_2, \ \cdots \ , \mathbf{y}_n]$. With the decomposition introduces in *Section 1.4*

$$\begin{aligned} [\mathbf{Y} - \mathbf{1}\boldsymbol{\mu}']' [\mathbf{Y} - \mathbf{1}\boldsymbol{\mu}'] &= (\mathbf{Y} - \mathbf{1}\hat{\boldsymbol{\mu}}_{\mathbf{y}}')' (\mathbf{Y} - \mathbf{1}\hat{\boldsymbol{\mu}}_{\mathbf{y}}') + (\hat{\boldsymbol{\mu}}_{\mathbf{y}} - \boldsymbol{\mu}) \mathbf{1}' \mathbf{1} (\hat{\boldsymbol{\mu}}_{\mathbf{y}} - \boldsymbol{\mu})' \\ &= \mathbf{Z}'\mathbf{Z} + n(\hat{\boldsymbol{\mu}}_{\mathbf{y}} - \boldsymbol{\mu})(\hat{\boldsymbol{\mu}}_{\mathbf{y}} - \boldsymbol{\mu})' \\ &= \mathbf{A} + n(\hat{\boldsymbol{\mu}}_{\mathbf{y}} - \boldsymbol{\mu})(\hat{\boldsymbol{\mu}}_{\mathbf{y}} - \boldsymbol{\mu})' \end{aligned}$$

formula (2.25) will become

$$L(\boldsymbol{\mu}, \ \boldsymbol{\Sigma}) = (2\pi)^{-6n/2} (\det \boldsymbol{\Sigma})^{-n/2} \operatorname{etr}\{-\frac{1}{2} \boldsymbol{\Sigma}^{-1} \mathbf{A}\} \exp\{[\hat{\boldsymbol{\mu}}_{\mathbf{y}} - \boldsymbol{\mu}]' \boldsymbol{\Sigma}^{-1} [\hat{\boldsymbol{\mu}}_{\mathbf{y}} - \boldsymbol{\mu}]\}. \tag{2.26}$$

The likelihood ratio statistic is

$$\Lambda_2 = \frac{\max_{\boldsymbol{\mu}} L(\boldsymbol{\mu}, \mathbf{I}_6)}{\max_{\boldsymbol{\mu}, \boldsymbol{\Sigma}} L(\boldsymbol{\mu}, \boldsymbol{\Sigma})}$$

in which

$$\max_{\boldsymbol{\mu}} L(\boldsymbol{\mu}, \mathbf{I}_6) = (2\pi)^{-6n/2} \operatorname{etr}\{-\frac{1}{2} \mathbf{A}\}$$

$$\max_{\boldsymbol{\mu}, \boldsymbol{\Sigma}} L(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = (2\pi)^{-6n/2} (\det \mathbf{A})^{-n/2} n^{-6n/2} \mathrm{e}^{-6n/2}.$$

where $\mathbf{A}=(\mathbf{Y}-\mathbf{1}\hat{\boldsymbol{\mu}}_{\mathbf{y}}')'(\mathbf{Y}-\mathbf{1}\hat{\boldsymbol{\mu}}_{\mathbf{y}}')=(n-1)\hat{\boldsymbol{\Sigma}}_{\mathbf{y}}$, then for testing $\mathcal{H}_{04}\colon \boldsymbol{\Sigma}=\boldsymbol{\Sigma}_0=\mathbf{I}_6$ the likelihood ratio statistic is

$$\Lambda_2=\left(\frac{e}{n}\right)^{6n/2}(\det\mathbf{A})^{n/2}\operatorname{etr}\{-\frac{1}{2}\mathbf{A}\} \tag{2.27}$$

For the general $\boldsymbol{\Sigma}_0\neq\mathbf{I}_6$ let $\mathbf{B}$ be a 6×6 nonsingular matrix such that $\mathbf{B}\boldsymbol{\Sigma}_0\mathbf{B}'=\mathbf{I}_6$ and put $\mathbf{x}_i=\mathbf{B}^{-1}\mathbf{y}_i$, $i=1,\cdots,6$, so $\mathbf{x}_i\sim\mathcal{N}_6(\mathbf{B}^{-1}\boldsymbol{\mu},\ \mathbf{B}^{-1}\boldsymbol{\Sigma}(\mathbf{B}^{-1})')$. To hypothesize the $\mathcal{H}_{04}:\boldsymbol{\Sigma}=\boldsymbol{\Sigma}_0$ is equivalent to testing $\mathcal{H}_{04}\colon\mathbf{B}^{-1}\boldsymbol{\Sigma}(\mathbf{B}^{-1})'=\mathbf{I}_6$, and now we have

$$\mathbf{A}_{\mathbf{x}}=\sum_{i=1}^{6}(\mathbf{x}_i-\hat{\boldsymbol{\mu}}_{\mathbf{x}})(\mathbf{x}_i-\hat{\boldsymbol{\mu}}_{\mathbf{x}})'=\mathbf{B}^{-1}\mathbf{A}(\mathbf{B}^{-1})'$$

By substituting $\mathbf{A}_{\mathbf{x}}$ into (2.27) we get the likelihood ratio statistic

$$\Lambda_2=\left(\frac{e}{n}\right)^{6n/2}(\det\boldsymbol{\Sigma}_0^{-1}\mathbf{A})^{n/2}\operatorname{etr}\{-\frac{1}{2}\boldsymbol{\Sigma}_0^{-1}\mathbf{A}\},\ \text{where }\mathbf{A}=(n-1)\hat{\boldsymbol{\Sigma}}_{\mathbf{y}} \tag{2.28}$$

So we have *Theorem 2.3*

Theorem 2.3

The likelihood ratio test of $\mathcal{H}_{04}:\boldsymbol{\Sigma}=\boldsymbol{\Sigma}_0$ with unknown mean vector $\boldsymbol{\mu}$ and unknown variance-covariance matrix $\boldsymbol{\Sigma}$ rejects $\mathcal{H}_{04}$ whenever

$$\Lambda_2=\left(\frac{e}{n}\right)^{3n/2}(\det\boldsymbol{\Sigma}_0^{-1}\mathbf{A})^{n/2}\operatorname{etr}\{-\frac{1}{2}\boldsymbol{\Sigma}_0^{-1}\mathbf{A}\}\leq C_\alpha$$

where the constant C_α is chosen in such a way that the test has size α.

To evaluate the constant C_α, we need the distribution of Λ_2 under null hypothesis $\mathcal{H}_{04}$, which are given, e.g., by *Anderson* (1958), *Giri* (1977): When the null hypotheses are true, $-2\log\Lambda_2$ is distributed as $\chi^2_{6(6+1)/2}$, when $n\to\infty$. *Das Gupta* (1969) has proved that the likelihood ratio test of (2.28) is biased, see also *N. Sugiura and H. Nagao* (1968). *Muirhead* (1982) has discussed the unbiased modified likelihood ratio statistic Λ_2^* in detail

$$\Lambda_2^*=\left(\frac{e}{n-1}\right)^{3(n-1)/2}(\det\boldsymbol{\Sigma}_0^{-1}\mathbf{A})^{(n-1)/2}\operatorname{etr}\{-\frac{1}{2}\boldsymbol{\Sigma}_0^{-1}\mathbf{A}\},\ \text{where }\mathbf{A}=(n-1)\hat{\boldsymbol{\Sigma}}_{\mathbf{y}}, \tag{2.29}$$

and provided the tables of its asymptotic distribution. We reject the null hypothesis $\mathcal{H}_{04}$ for small enough Λ_2^*.

2.2.2 Tests on the equality of two variance-covariance matrices

We consider testing the null hypothesis that the variance-covariance matrices of two normal distributions are equal, given independent samples from the two populations. Let $\mathbf{y}_{i1},\cdots,\mathbf{y}_{in}$ be independent $\mathcal{N}_6(\boldsymbol{\mu}_i,\boldsymbol{\Sigma}_i)$ -distributed random vectors of vectorized random tensors with unknown mean vector $\boldsymbol{\mu}_i$ and variance-covariance matrix $\boldsymbol{\Sigma}_i$ and consider testing the hypothesis

$$\mathcal{H}_{05}:\ \boldsymbol{\Sigma}_1=\boldsymbol{\Sigma}_2,\ \mathcal{H}_{15}:\ \boldsymbol{\Sigma}_1\neq\boldsymbol{\Sigma}_2$$

Let $\hat{\boldsymbol{\mu}}_{\mathbf{y}i}$ and $\mathbf{A}_i$ be, respectively, the mean vector and the matrix of sums of squares and products formed from the ith sample; that is

$$\hat{\boldsymbol{\mu}}_{\mathbf{y}i}=\frac{1}{n_i}\sum_{j=1}^{n_i}\mathbf{y}_{ij},\quad \mathbf{A}_i=\sum_{j=1}^{n_i}(\mathbf{y}_{ij}-\hat{\boldsymbol{\mu}}_{\mathbf{y}i})(\mathbf{y}_{ij}-\hat{\boldsymbol{\mu}}_{\mathbf{y}i})'$$

and denote $\mathbf{A}=\mathbf{A}_1+\mathbf{A}_2$, $n=n_1+n_2$.

The likelihood ratio test of $\mathcal{H}_{05}$, first derived by Wilks (1932), is given in the following theorem for the case of two populations.

Theorem 2.4

The likelihood ratio test of $\mathcal{H}_{05}:\boldsymbol{\Sigma}_1=\boldsymbol{\Sigma}_2$ with unknown mean vector $\boldsymbol{\mu}$ and unknown variance-covariance matrix $\boldsymbol{\Sigma}$ rejects $\mathcal{H}_{05}$ whenever

$$\Lambda_3=\frac{(\det\mathbf{A}_1)^{n_1/2}(\det\mathbf{A}_2)^{n_2/2}n^{6n/2}}{(\det\mathbf{A})^{n/2}n_1^{6n_1/2}n_2^{6n_2/2}}=\frac{(\det\mathbf{A}_1)^{n_1/2}(\det\mathbf{A}_2)^{n_2/2}n^{6n/2}}{(\det\mathbf{A}_1+\mathbf{A}_2)^{(n_1+n_2)/2}n_1^{6n_1/2}n_2^{6n_2/2}}\leq C_\alpha \tag{2.30}$$

where the constant C_α is chosen in such a way that the test has size α.

When the null hypotheses are true, $-2\log\Lambda_3$ is distributed as $\chi^2_{6(6+1)/2}$, when $n\to\infty$. For the unbiasedness and modified likelihood ratio statistic about $\mathcal{H}_{05}$ we refer to *Giri* (1977) and *Murihead* (1982) in detail.

2.2.3 Tests if the mean vector and variance-covariance matrix are equal to a given vector and matrix

We consider testing if the mean vector and the variance-covariance matrix are equal to a given vector and matrix, respectively, which is obviously the combination of the hypothesis test $\mathcal{H}_{01}$ about the mean vector of *Section 2.1.1* and $\mathcal{H}_{04}$ of *Section 2.2.1* for the variance-covariance matrix:

$$\mathcal{H}_{06}: \ \boldsymbol{\mu} = \boldsymbol{\mu}_0, \ \boldsymbol{\Sigma}_{\mathbf{y}} = \boldsymbol{\Sigma}_0 \quad \mathcal{H}_{16}: \ \boldsymbol{\mu} \neq \boldsymbol{\mu}_0 \ \text{or} \ \boldsymbol{\Sigma}_{\mathbf{y}} \neq \boldsymbol{\Sigma}_0$$

The likelihood ratio test is given in the following theorem from Anderson (1958)

Theorem 2.5

Given the n independent 3×1 samples vectors $\mathbf{y}_1, \cdots, \mathbf{y}_n$, all distributed as $\mathcal{N}_6(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, the likelihood ratio test of size α of $\mathcal{H}_{06}: \ \boldsymbol{\mu} = \boldsymbol{\mu}_0, \ \boldsymbol{\Sigma}_{\mathbf{y}} = \boldsymbol{\Sigma}_0$ is based on

$$\Lambda_4 = \left(\frac{e}{n}\right)^{6n/2} (\det \boldsymbol{\Sigma}_0^{-1}\mathbf{A})^{n/2} \operatorname{etr}\{-\frac{1}{2}\boldsymbol{\Sigma}_0^{-1}\mathbf{A}\} \exp\{-\frac{1}{2} n[\hat{\boldsymbol{\mu}}_{\mathbf{y}} - \boldsymbol{\mu}_0]' \boldsymbol{\Sigma}_0^{-1} [\hat{\boldsymbol{\mu}}_{\mathbf{y}} - \boldsymbol{\mu}_0]\} \tag{2.31}$$

and reject $\mathcal{H}_{06}$ if $\Lambda_4 \leq C_\alpha$, where the constant C_α is chosen in such a way that the test has size α. When the null hypothesis is true, $-2\log \Lambda_4$ is asymptotically distributed as $\chi^2_{6(6+1)/2+6}$.

The likelihood ratio test (2.31) is unbiased, which has been established by *Sugiura and Nagao* (1968) and *Das Gupta* (1969). *Muirhead* (1982) provided the tables of its asymptotic distribution. We reject the null hypothesis $\mathcal{H}_{06}$ for small enough Λ_4.

2.2.4 Tests on the equality of two mean vectors and two variance-covariance matrices

Now we consider simultaneous testing for the equality of the mean vectors and the variance-covariance matrices from two populations, which is obviously the combination of the hypothesis test $\mathcal{H}_{03}$ about the mean vector of *Section 2.1.3* and $\mathcal{H}_{05}$ of *Section 2.2.2* for the variance-covariance matrix, which is our seventh hypothesis test:

$$\mathcal{H}_{07}: \ \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2, \ \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 \quad \mathcal{H}_{17}: \ \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2 \ \text{or} \ \boldsymbol{\Sigma}_1 \neq \boldsymbol{\Sigma}_2$$

The likelihood ratio statistic for hypothesis $\mathcal{H}_{07}$ is the product of the Λ_1 of $\mathcal{H}_{03}$ and Λ_3 of $\mathcal{H}_{05}$.

$$\begin{aligned} \Lambda_5 &= \Lambda_1 \Lambda_3 = \\ &= \frac{(\det \mathbf{A})^{n/2}}{(\det \mathbf{W})^{n/2}} \frac{(\det \mathbf{A}_1)^{n_1/2} (\det \mathbf{A}_2)^{n_2/2} n^{3n/2}}{(\det \mathbf{A})^{n/2} n_1^{3n_1/2} n_2^{3n_2/2}} = \\ &= \frac{(\det \mathbf{A}_1)^{n_1/2} (\det \mathbf{A}_2)^{n_2/2} n^{3n/2}}{(\det \mathbf{W})^{n/2} n_1^{3n_1/2} n_2^{3n_2/2}} \end{aligned} \tag{2.32}$$

and the likelihood ratio test rejects $\mathcal{H}_{07}$ whenever

$$\Lambda_5 \leq C_\alpha,$$

where C_α depends on the error probability α.

Box (1949) has derived the distribution of the modified likelihood ratio statistic Λ_5^*

$$\Lambda_5^* = \frac{(\det \mathbf{A}_1)^{(n_1-1)/2} (\det \mathbf{A}_2)^{(n_2-1)/2} n^{3(n-2)/2}}{(\det \mathbf{W})^{(n-1)/2} (n_1-1)^{3(n_1-1)/2} (n_2-1)^{3(n_2-1)/2}} \tag{2.33}$$

If the null hypothesis is true, $-2\rho \log \Lambda_5^*$ is distributed as $\chi^2_{6(6+1)/2}$, when $n \to \infty$, where ρ is a numerical value related to the sampling number, the dimension of the sampling vector, and the total number of populations.

In this section we have discussed four hypothesis tests about the sample variance-covariance matrix of a symmetric random tensor and the combination of mean vectors, which are summarised in *Box* 2.2.

***Box* 2.2**. (Hypothesis tests of the sample variance-covariance)

$\mathcal{H}_{04}: \ \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0, \ \mathcal{H}_{14}: \ \boldsymbol{\Sigma} \neq \boldsymbol{\Sigma}_0$

$\mathcal{H}_{05}: \ \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2, \ \mathcal{H}_{15}: \ \boldsymbol{\Sigma}_1 \neq \boldsymbol{\Sigma}_2$

$\mathcal{H}_{06}: \ \boldsymbol{\mu} = \boldsymbol{\mu}_0, \ \boldsymbol{\Sigma}_{\mathbf{y}} = \boldsymbol{\Sigma}_0 \quad \mathcal{H}_{16}: \ \boldsymbol{\mu} \neq \boldsymbol{\mu}_0 \ \text{or} \ \boldsymbol{\Sigma}_{\mathbf{y}} \neq \boldsymbol{\Sigma}_0$

$\mathcal{H}_{07}: \ \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2, \ \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 \quad \mathcal{H}_{17}: \ \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2 \ \text{or} \ \boldsymbol{\Sigma}_1 \neq \boldsymbol{\Sigma}_2$

Chapter 3

Optimal α for Tykhonov-Phillips regularization by A-optimal design - α-weighted BLE, and a simulated case study for 2-D strain rate tensors

Numerical tests have documented that the estimate $\hat{\boldsymbol{\xi}}$ of type BLUUE of the parameter vector $\boldsymbol{\xi}$ within a linear *Gauss-Markov model* $\{\mathbf{A}\boldsymbol{\xi} = E\{\mathbf{y}\},\ \boldsymbol{\Sigma}_y = D\{\mathbf{y}\}\}$ is *not* robust against *outliers* in the stochastic observation vector **y**. It is for this reason that *we give up* the postulate of unbiasedness, but keep the set-up of a *linear estimation* $\hat{\boldsymbol{\xi}} = \mathbf{Ly}$ of homogeneous type. Ever since *Tykhonov* (1963) and *Phillips* (1962) introduced the *hybrid minimum norm approximation solution* (HAPS) of a *linear improperly posed problem* there has been left the open problem to evaluate the weighting factor α between the least-squares -norm and the minimum length norm of the unknown parameters. In most applications of *Tykhonov-Phillips* type of regularization the weighting factor α is determined by simulation studies, but according to the literature listed below also optimization techniques have been applied. Here we aim at an objective method to determine the *weighting factor* α within α-HAPS.

Alternatively, improperly posed problems which appear in solving integral equations of the first kind or downward continuation problems in potential theory depart from observations which are elements of a probability space. Accordingly, estimation techniques of type BLUUE (best linear uniformly unbiased estimation) have been implemented to estimate $\hat{\boldsymbol{\xi}}$ as an unknown parameter vector $\boldsymbol{\xi}$ ("fixed effects") within a linear *Gauss-Markov model*, such an estimation is *not* robust against *outliers* in the stochastic observation vector $\mathbf{y} \in \mathbb{Y}$. Here we assume that the observation vector **y** is an element of the *observation space* $\mathbb{Y}$, dim $\mathbb{Y}$= *n*, namely an observation space $\mathbb{Y} = \mathbb{R}^n$ equipped with a Euclidean metric. Due to possibly unstable solutions of type BLUUE with respect to the "*fixed effects*" linear Gauss-Markov model *we give up* the postulate of *unbiasedness*, but keep the set-up of a *linear estimation* $\hat{\boldsymbol{\xi}} = \mathbf{Ly}$ of homogeneous type. According to Grafarend and Schaffrin (1993), updated by Schaffrin (2000), the best linear estimation of type α-homBLE (*α-weighted Best homogeneously Linear Estimation*) which is based on *hybrid norm optimization* of type (i) minimum variance *and* (ii) minimum bias leads us to the *equivalence* of α-homBLE *and* α-HAPS under the following condition. *If* we choose the weight matrix in the least squares norm as the inverse matrix of the variance covariance matrix of the observations *as well as* the weight matrix in the minimum norm acting on the unknown parameter vector as the inverse substitute bias weight matrix, *then* α-homBLE *and* α-HAPS are *equivalent*.

The second method of regularizing an improperly posed problem offers the possibility to determine the regularization parameter α in an optimal way. For instance, by an A-optimal design of type

"minimize the trace of the *Mean Square*
Error matrix $\operatorname{tr}\operatorname{MSE}\{\hat{\boldsymbol{\xi}}\}$ of $\hat{\boldsymbol{\xi}}$ (α-hom BLE) to find
$\hat{\alpha} = \arg\{\operatorname{tr}\operatorname{MSE}\{\hat{\boldsymbol{\xi}}\} = \min\}$"

we are able to construct the regularization parameter α which *balances the trace of the variance-covariance matrix* $\operatorname{tr}\operatorname{D}\{\hat{\boldsymbol{\xi}}\}$ *and the trace of the quadratic bias* $\operatorname{tr}\boldsymbol{\beta\beta}'$ for the bias vector $\boldsymbol{\beta} = -[\mathbf{I} - \mathbf{LA}]\boldsymbol{\xi}$.

The biased estimation solves a special inverse problem, and is also known as *Tykhonov-Phillips regulator* or *ridge estimator*. For a comprehensive discussion and review about the methods of solving the inverse problem we refer to Allen, (1971, 1974), Arslan and Billor (2000), Bouman (1998), Chaturvedi and Singh (2000), Donatos and Michailidis (1990), Draper, et al. (1979), Droge (1993), Engels, et al. (1993), Engl (1993), EL-Sayed (1996), Farebrother (1975, 1976, 1978), Firinguetti (1996), Firinguetti and Rubio (2000), Gibbons (1981), Golub Heath and Wahba (1979), Grafarend and Schaffrin (1993), Gui, et al. (1998a, b, 2000, 2001), Gunst and Mason (1977), Gunst and Mason (1980), Hanke and Hansen (1993), Hansen (1992,1993), Hansen (1993), Hemmerle (1975), Hemmerle and Brantle (1978), Hocking (1976), Hoerl and Kennard (1970a, 1970b), Hoerl, Kennard and Baldwin (1975), Hoerl (1985), Hoerl, Schuenemeyer and Hoerl (1985), Ilk (1986), Kacirattin, Sakalloglu and Akdeniz (1998), Lawless and Wang (1976), Liu (1993), Louis, Maass and Lowerre (1974), Mallows (1973), Markiewicz (1996), Marquardt (1970,1974), Marquardt and Snee (1975), Mayer and Willke (1973), McDonald and Galarneau (1975), Nomura (1988, 1998), Ohtani (1986, 1998), Phillips (1962), Rao (1975, 1976), Schaffrin, Heidenreich and Grafarend (1977), Schaffrin and Middel (1990), Schaffrin (1995), Schaffrin (2000), Shalabh (1998), Smith and Campbell (1980), Srivastava, et al. (1983), Theobald (1974), Tykhonov (1963), Tykhonov, et al. (1977), Tykhonov and Arsenin (1977), Vinod and Ullah (1981), Xu (1992a, b, 1998), Xu and Rummel (1994a, b), Wang and Xiao (2001), Wenchenko (2000).

Of those quoted references, Grafarend and Schaffrin (1993) as well as Schaffrin (2000) have systematically derived the best linear estimators of type homBLE (*Best homogeneously Linear Estimation*), S-homBLE and α-homBLE of the *fixed effects* ξ, which turn out to enhance the best linear uniformly unbiased estimator of type BLUUE, but suffer from the effect being biased. In this chapter the regularization parameter in uniform Tykhonov-Phillips regularization (α-weighted BLE) is determined by minimizing the trace of the *Mean Square Error matrix MSE*$\{\hat{\xi}\}$ (*A-optimal design*) in the general case for the Gauss-Markov model. Through two comparisons it is shown that the optimal ridge parameter *k* in *ridge regression* developed by *Hoerl* and *Kennard* (1970a, 1970b) and *Hoerl, Kennard* and *Baldwin* (1975) is just the special case of our general solution by A-optimal design. Based on the introduction of the multivariate $\alpha-$homBLE for the multivariate parameters, the determination of the optimal weight factor α are generalized to the multivariate Gauss-Markov model, which we shall call "*multivariate ridge estimator*". In lieu of a case study, both model and estimators are tested and analyzed with numerical results computed from simulated direct observations of a random tensor of type strain rate.

3.1 The optimal regularization parameter α in uniform Tykhonov-Phillips regularization by A-optimal design (α-weighted BLE)

Let us first introduce the *special Gauss-Markov model* $\mathbf{y} = \mathbf{A}\xi + \mathbf{e}$ specified in *Box 3.1*, which is given for the *first order moments* in the form of a *consistent system of linear equations relating* the first *non-stochastic* ("fixed"), *real-valued vector* ξ of *unknown* parameters to the expectation $E\{\mathbf{y}\}$ of the *stochastic*, real-valued vector $\mathbf{y}$ of observations, $\mathbf{A}\xi = E\{\mathbf{y}\}$, since $E\{\mathbf{y}\} \in \mathcal{R}(\mathbf{A})$ is an element of the column space $\mathcal{R}(\mathbf{A})$ of the real-valued, *non-stochastic* ("fixed") "*first order design matrix*" $\mathbf{A} \in \mathbb{R}^{n\times m}$. The rank of the fixed matrix $\mathbf{A}$, rk $\mathbf{A}$, equals the number *m* of unknowns, $\xi \in \mathbb{R}^m$. In addition, the *second order central moments* in the *regular* variance-covariance matrix $\mathbf{\Sigma}_\mathrm{y}$, also called *dispersion matrix* $D\{\mathbf{y}\}$, constitute the second matrix $\mathbf{\Sigma}_\mathrm{y} \in \mathbb{R}^{n\times n}$ of *unknowns* to be specified in a linear model furtheron.

***Box 3.1*:**

Special Gauss–Markov model

$$\mathbf{y} = \mathbf{A}\xi + \mathbf{e}$$

1st moments

$$\mathbf{A}\xi = E\{\mathbf{y}\},\ \mathbf{A} \in \mathbb{R}^{n\times m},\ E\{\mathbf{y}\} \in \mathcal{R}(\mathbf{A}),\ \text{rk}\,\mathbf{A} = m \tag{3.1}$$

2nd moments

$$\mathbf{\Sigma}_\mathrm{y} = D\{\mathbf{y}\} \in \mathbb{R}^{n\times n}, \mathbf{\Sigma}_\mathrm{y} \text{ positive definite, rk } \mathbf{\Sigma}_\mathrm{y} = n$$
$$\xi,\ \mathbf{y} - E\{\mathbf{y}\} = \mathbf{e}\ \textit{unknown},\ \mathbf{\Sigma}_\mathrm{y}\ \textit{unknown, but structured}. \tag{3.2}$$

Obviously a homogeneously linear form $\hat{\xi} = \mathbf{L}\mathbf{y}$ is sufficient to generate $\mathbf{\Sigma}$ - BLUUE (Best Linear Uniformly Unbiased Estimation with respect to the $\mathbf{\Sigma}$ - norm) for the special Gauss-Markov model (3.1), (3.2). Explicit representations of $\mathbf{\Sigma}$ - BLUUE of type $\hat{\xi}$ as well as of its dispersion matrix $D\{\hat{\xi} \mid \hat{\xi}\ \mathbf{\Sigma}\text{-BLUUE}\}$ generated by solving the normal equations derived from the minimum of the quadratic constraint *Lagrangean* are collected in

Theorem 3.1 ($\hat{\xi}$ BLUUE of ξ):

Let $\hat{\xi} = \mathbf{L}\mathbf{y}$ be $\mathbf{\Sigma}_\mathrm{y}$ - BLUUE $\hat{\xi}$ of ξ in the *special linear Gauss-Markov model* (3.1), (3.2). Then

$$\hat{\xi} = \mathbf{L}\mathbf{y} = (\mathbf{A}'\mathbf{\Sigma}_\mathrm{y}^{-1}\mathbf{A})^{-1}\mathbf{A}'\mathbf{\Sigma}_\mathrm{y}^{-1}\mathbf{y} \tag{3.3}$$

$$\hat{\xi} = \mathbf{\Sigma}_{\hat{\xi}}\mathbf{A}'\mathbf{\Sigma}_\mathrm{y}^{-1}\mathbf{y} \tag{3.4}$$

subject to the related dispersion matrix

$$D\{\hat{\xi}\} := \mathbf{\Sigma}_{\hat{\xi}} = (\mathbf{A}'\mathbf{\Sigma}_\mathrm{y}^{-1}\mathbf{A})^{-1}. \tag{3.5}$$

Apparently $\hat{\xi}$ of type $\mathbf{\Sigma}_\mathrm{y}$ - BLUUE of ξ is *not* robust against *outliers* in the stochastic vector $\mathbf{y}$ of observations. It is for this reason that *we give up* the postulate of unbiasedness, but keep the set-up of a *linear estimation* $\hat{\xi} = \mathbf{L}\mathbf{y}$ of homogeneous type, which turns out to better than the best linear uniformly unbiased estimator of type hom BLUUE, but suffers from the effect to be biased. Here we will focus on best linear estimators of type α-homBLE of the *fixed effects* ξ. At first let us begin with a discussion of the *bias vector* and the *bias matrix* as well as of the *Mean Square Error matrix* $MSE\{\hat{\boldsymbol{\xi}}\}$ with respect to a *homogeneously linear estimate* $\hat{\boldsymbol{\xi}} = \mathbf{Ly}$ of fixed effects $\boldsymbol{\xi}$ based upon *Box 3.2*.

***Box 3.2*:**

Bias vector, bias matrix, Mean Square Error matrix
in the special Gauss–Markov model with fixed effects

$$E\{\mathbf{y}\} = \mathbf{A}\boldsymbol{\xi} \tag{3.6}$$

$$D\{\mathbf{y}\} = \boldsymbol{\Sigma}_y \tag{3.7}$$

"ansatz"

$$\hat{\boldsymbol{\xi}} = \mathbf{Ly} \tag{3.8}$$

bias vector

$$\boldsymbol{\beta} := E\{\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}\} = E\{\hat{\boldsymbol{\xi}}\} - \boldsymbol{\xi} \tag{3.9}$$

$$\boldsymbol{\beta} = \mathbf{L}E\{\mathbf{y}\} - \boldsymbol{\xi} = -[\mathbf{I}_m - \mathbf{LA}]\boldsymbol{\xi} \tag{3.10}$$

bias matrix

$$\mathbf{B} := \mathbf{I}_m - \mathbf{LA} \tag{3.11}$$

decomposition

$$\hat{\boldsymbol{\xi}} - \boldsymbol{\xi} = (\hat{\boldsymbol{\xi}} - E\{\hat{\boldsymbol{\xi}}\}) + (E\{\hat{\boldsymbol{\xi}}\} - \boldsymbol{\xi}) \tag{3.12}$$

$$\hat{\boldsymbol{\xi}} - \boldsymbol{\xi} = \mathbf{L}(\mathbf{y} - E\{\mathbf{y}\}) - [\mathbf{I}_m - \mathbf{LA}]\boldsymbol{\xi} \tag{3.13}$$

Mean Square Error matrix

$$MSE\{\hat{\boldsymbol{\xi}}\} := E\{(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi})(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi})'\} \tag{3.14}$$

$$MSE\{\hat{\boldsymbol{\xi}}\} = \mathbf{L}D\{\mathbf{y}\}\mathbf{L}' + [\mathbf{I}_m - \mathbf{LA}]\boldsymbol{\xi}\boldsymbol{\xi}'[\mathbf{I}_m - \mathbf{LA}]' \tag{3.15}$$

$$(E\{\hat{\boldsymbol{\xi}} - E\{\hat{\boldsymbol{\xi}}\}\} = 0)$$

S-*modified Mean Square Error*

$$MSE_S\{\hat{\boldsymbol{\xi}}\} := \mathbf{L}D\{\mathbf{y}\}\mathbf{L}' + [\mathbf{I}_m - \mathbf{LA}]\mathbf{S}[\mathbf{I}_m - \mathbf{LA}]' \tag{3.16}$$

S - nonnegative definite *substitute* matrix.

Frobenius matrix norm

$$\| MSE\{\hat{\boldsymbol{\xi}}\} \| := \operatorname{tr} E\{(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi})(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi})'\} \tag{3.17}$$

$$\begin{aligned} &\| MSE\{\hat{\boldsymbol{\xi}}\} \| = \\ &= \operatorname{tr} \mathbf{L}D\{\mathbf{y}\}\mathbf{L}' + + \operatorname{tr}[\mathbf{I}_m - \mathbf{LA}]\boldsymbol{\xi}\boldsymbol{\xi}'[\mathbf{I}_m - \mathbf{LA}]' \\ &= \| \mathbf{L}' \|^2_{\Sigma_y} + \| (\mathbf{I}_m - \mathbf{LA})' \|^2_{\xi\xi'} \end{aligned} \tag{3.18}$$

$$\begin{aligned} &\| MSE_S\{\hat{\boldsymbol{\xi}}\} \| := \\ &:= \operatorname{tr} \mathbf{L}D\{\mathbf{y}\}\mathbf{L}' + \operatorname{tr}[\mathbf{I}_m - \mathbf{LA}]\mathbf{S}[\mathbf{I}_m - \mathbf{LA}]' \\ &= \| \mathbf{L}' \|^2_{\Sigma_y} + \| (\mathbf{I}_m - \mathbf{LA})' \|^2_S \end{aligned} \tag{3.19}$$

α-weighted hybrid minimum variance – minimum bias norm

$$\begin{aligned} &\| MSE_{\alpha,S}\{\hat{\boldsymbol{\xi}}\} \| := \\ &= \operatorname{tr} \mathbf{L}D\{\mathbf{y}\}\mathbf{L}' + \operatorname{tr}[\mathbf{I}_m - \mathbf{LA}]\frac{1}{\alpha}\mathbf{S}[\mathbf{I}_m - \mathbf{LA}]' \\ &= \| \mathbf{L}' \|^2_{\Sigma_y} + \frac{1}{\alpha}\| (\mathbf{I}_m - \mathbf{LA})' \|^2_S \end{aligned} \tag{3.20}$$

special assumption

$$\dim \mathcal{R}(\mathbf{SA}') = \operatorname{rk} \mathbf{SA}' = \operatorname{rk} \mathbf{A} = m \Rightarrow \operatorname{rk} \mathbf{S} \geq m \Rightarrow \mathbf{S}^{-1} \text{ exists.} \tag{3.21}$$

The *bias vector* $\boldsymbol{\beta}$ is conventionally defined by $E\{\hat{\boldsymbol{\xi}}\} - \boldsymbol{\xi}$ subject to the homogeneous estimation form $\hat{\boldsymbol{\xi}} = \mathbf{Ly}$. Accordingly, the bias vector can be represented by (3.10) $\boldsymbol{\beta} = -[\mathbf{I}_m - \mathbf{LA}]\boldsymbol{\xi}$. Since the vector $\boldsymbol{\xi}$ of *fixed effects* is unknown, there has been made the proposal to instead use the matrix $\mathbf{I}_m - \mathbf{LA}$ as a *matrix-valued measure of bias*. A measure of the estimation error is the *Mean Square Error matrix* $MSE\{\hat{\xi}\}$ of type (3.14). $MSE\{\hat{\xi}\}$ can be decomposed into two *basic parts*:

- the dispersion matrix $D\{\hat{\xi}\} = \mathbf{L}D\{\mathbf{y}\}\mathbf{L}'$
- the dyadic bias product $\boldsymbol{\beta\beta}'$

Indeed the vector $\hat{\xi} - \xi$ can be decomposed as well into two parts of type (3.12), (3.13), namely (i) $\hat{\xi} - E\{\hat{\xi}\} = \mathbf{Le}$ and (ii) $E\{\hat{\xi}\} - \xi = \boldsymbol{\beta}$ which may be called random estimation error (due to observation noise) and bias, respectively. The double decomposition of the vector $\hat{\xi} - \xi$ leads straightforwardly to the double representation of the matrix $MSE\{\hat{\xi}\}$ of type (3.15). Such a representation suffers from two effects: *Firstly* the vector ξ of *fixed effects* is unknown, *secondly* the matrix $\xi\xi'$ has only rank 1. In consequence, the matrix $[\mathbf{I}_m - \mathbf{LA}]\xi\xi'[\mathbf{I}_m - \mathbf{LA}]'$ has only rank 1, too. In this situation the proposal has been made to *modify* $MSE\{\hat{\xi}\}$ with respect to $\xi\xi'$ by a *higher rank matrix* **S.** A homogeneously linear α-weighted hybrid minimum variance-minimum bias estimation (α-homBLE) is presented in *Definition 3.1* which is based upon the weighted sum of two norms of type (3.20), namely

- *average variance* $\|\mathbf{L}'\|^2_{\boldsymbol{\Sigma}_\mathbf{y}} = \mathrm{tr}\,\mathbf{L}\boldsymbol{\Sigma}_\mathbf{y}\mathbf{L}'$
- ***S***-*weighted average bias* $\|(\mathbf{I}_m - \mathbf{LA})'\|^2_\mathbf{S} = \mathrm{tr}[\mathbf{I}_m - \mathbf{LA}]\mathbf{S}[\mathbf{I}_m - \mathbf{LA}]'$

  where we expect ξ to belong to the column space $\mathcal{R}(\mathbf{S})$.

The very important estimator α-homBLE *is balancing* variance and bias by the weighting factor α which is illustrated by *Figure 3.1*.

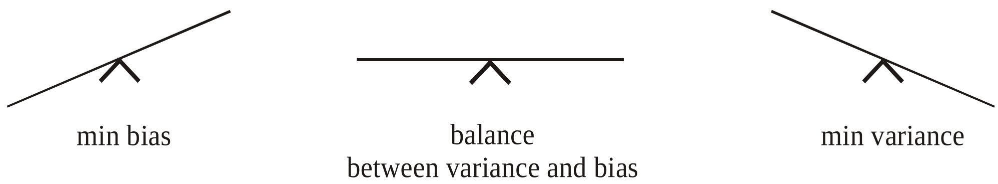


Figure 3.1 Balance of variance and bias by the weighting factor α

Definition 3.2 ($\hat{\xi}$ homBLE of ξ):

A $m\times 1$ vector $\hat{\xi}$ is called homogeneous *BLE* of ξ in the *special linear Gauss-Markov model with fixed effects* of *Box 3.1*, if and only if

(1*st*) $\hat{\xi}$ is a homogeneously linear form

$$\hat{\xi} = \mathbf{Ly}\,, \tag{3.22}$$

(2*nd*) in comparison to all other linear estimations $\hat{\xi}$ has the minimum *M*ean *S*quare *E*rror in the sense of

$$\begin{aligned} &\| MSE\{\hat{\xi}\} \| = \\ &= \mathrm{tr}\,\mathbf{L}D\{\mathbf{y}\}\mathbf{L}' + \mathrm{tr}[\mathbf{I}_m - \mathbf{LA}]\xi\xi'[\mathbf{I}_m - \mathbf{LA}]' = \\ &= \| \mathbf{L}' \|^2_{\Sigma_\mathbf{y}} + \| (\mathbf{I}_m - \mathbf{LA})' \|^2_{\xi\xi'} = \min_\mathbf{L} \end{aligned} \tag{3.23}$$

Definition 3.3 ($\hat{\xi}$ **S**-homBLE of ξ):

A $m\times 1$ vector $\hat{\xi}$ is called homogeneous ***S****-homBLE* of ξ in the *special linear Gauss-Markov model with fixed effects* of *Box 3.1*, if and only if

(1*st*) $\hat{\xi}$ is a homogeneously linear form

$$\hat{\xi} = \mathbf{Ly}\,, \tag{3.24}$$

(2*nd*) in comparison to all other linear estimations $\hat{\xi}$ has the minimum **S**-modified *M*ean *S*quare *E*rror in the sense of

$$\begin{aligned} &\| MSE_\mathbf{S}\{\hat{\xi}\} \| = \\ &= \mathrm{tr}\,\mathbf{L}D\{\mathbf{y}\}\mathbf{L}' + \mathrm{tr}[\mathbf{I}_m - \mathbf{LA}]\mathbf{S}[\mathbf{I}_m - \mathbf{LA}]' \\ &= \| \mathbf{L}' \|^2_{\Sigma_\mathbf{y}} + \| (\mathbf{I}_m - \mathbf{LA})' \|^2_\mathbf{S} = \min_\mathbf{L} \end{aligned} \tag{3.25}$$

Definition 3.4 ($\hat{\xi}$ homlinear α-weighted hybrid min var-min bias solution, or α- homBLE):

A $m\times 1$ vector $\hat{\xi}$ is called homogeneously linear α-weighted hybrid minimum variance-minimum bias estimate (α-homBLE) of ξ in the *special linear Gauss-Markov model with fixed effects* of *Box 3.1*, if and only if

(1*st*) $\hat{\xi}$ is a homogeneously linear form

$$\hat{\xi} = \mathbf{Ly}\ , \tag{3.26}$$

(2*nd*) in comparison to all other homogeneously linear estimates $\hat{\xi}$ has the minimum variance-minimum bias property in the sense of the α-weighted hybrid norm

$$\begin{aligned} &\| MSE_{\alpha,\mathbf{S}}\{\hat{\xi}\}\|^2 = \\ &= \operatorname{tr}\mathbf{L}D\{\mathbf{y}\}\mathbf{L}' + \frac{1}{\alpha}\operatorname{tr}(\mathbf{I}_m - \mathbf{LA})\mathbf{S}(\mathbf{I}_m - \mathbf{LA})' \\ &= \|\mathbf{L}'\|^2_{\Sigma_y} + \frac{1}{\alpha}\|(\mathbf{I}_m - \mathbf{LA})'\|^2_{\mathbf{S}} = \min_{\mathbf{L}} \end{aligned} \tag{3.27}$$

in particular with respect to the special assumption

$$\alpha \in \mathbb{R}^+,\ \dim\mathcal{R}(\mathbf{SA}') = \operatorname{rk}\mathbf{SA}' = \operatorname{rk}\mathbf{A} = m \ \Rightarrow\ \mathbf{S}^{-1} \text{ exists.}$$

The estimates $\hat{\xi}$ of type *homBLE*, ***S*** *-homBLE* and α-homBLE can be characterized by the following *Lemma*.

Lemma 3.5 (homBLE, **S**-homBLE and α- homBLE):

A $m\times 1$ vector $\hat{\xi}$ is *homBLE*, ***S*** *-homBLE* and *α - homBLE* of ξ in the special linear Gauss-Markov model with fixed effects of *Box 3.1*, if and only if the matrix $\hat{\mathbf{L}}$ fulfils the normal equations

(1*st*) *homBLE*:

$$(\mathbf{\Sigma}_y + \mathbf{A}\xi\xi'\mathbf{A}')\hat{\mathbf{L}}' = \mathbf{A}\xi\xi' \tag{3.28}$$

(2nd) **S** -homBLE:

$$(\mathbf{\Sigma}_y + \mathbf{ASA}')\hat{\mathbf{L}}' = \mathbf{AS} \tag{3.29}$$

(3rd) α- homBLE:

$$(\mathbf{SA}'\mathbf{\Sigma}_y^{-1}\mathbf{A} + \alpha\mathbf{I}_m)\hat{\mathbf{L}} = \mathbf{SA}'\mathbf{\Sigma}_y^{-1}$$

or, if **S** is non-singular, (It is, if $\operatorname{rk}\mathbf{SA}' = \operatorname{rk}\mathbf{A} = m$) (3.30)

$$(\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A} + \alpha\mathbf{S}^{-1})\hat{\mathbf{L}} = \mathbf{A}'\mathbf{\Sigma}_y^{-1}$$

Proof :

(i) homBLE:

The hybrid norm $\| MSE\{\hat{\xi}\}\|^2$ establishes the Lagrangean

$$\mathcal{L}_1(\mathbf{L}) := \operatorname{tr}\mathbf{L\Sigma}_y\mathbf{L}' + \operatorname{tr}(\mathbf{I}_m - \mathbf{LA})\xi\xi'(\mathbf{I}_m - \mathbf{LA})' = \min_{\mathbf{L}}$$

for $\hat{\xi}$ as *homBLE* of ξ. The *necessary conditions* for the minimum of the *quadratic Lagrangean* $\mathcal{L}_1(\mathbf{L})$ are

$$\frac{\partial\mathcal{L}_1}{\partial\mathbf{L}}(\hat{\mathbf{L}}) := 2[\mathbf{\Sigma}_y\hat{\mathbf{L}}' + \mathbf{A}\xi\xi'\mathbf{A}'\hat{\mathbf{L}}' - \mathbf{A}\xi\xi']' = 0$$

which agrees with the normal equations (3.28). The theory of the derivative of a scalar-valued function with respect to a matrix is reviewed in *Appendix A* of the book by Grafarend and Schaffrin (1993).

The second derivatives, namely

$$\frac{\partial^2\mathcal{L}_1}{\partial(\operatorname{vec}\mathbf{L})\partial(\operatorname{vec}\mathbf{L})'}(\hat{\mathbf{L}}) > 0$$

at the “point” $\hat{\mathbf{L}}$, constitute the *sufficient conditions*.

In order to compute such a $mn\times mn$ matrix of second derivatives we have to vectorize the matrix normal equation

$$\frac{\partial \mathcal{L}_1}{\partial \mathbf{L}}(\hat{\mathbf{L}}) := 2\hat{\mathbf{L}}(\mathbf{\Sigma}_y + \mathbf{A}\xi\xi'\mathbf{A}') - 2\xi\xi'\mathbf{A}',$$

$$\frac{\partial \mathcal{L}_1}{\partial (\text{vec}\mathbf{L})}(\hat{\mathbf{L}}) := \text{vec}[2\hat{\mathbf{L}}(\mathbf{\Sigma}_y + \mathbf{A}\xi\xi'\mathbf{A}') - 2\xi\xi'\mathbf{A}'],$$

where vec denotes the vec-operation (vectorization of an array). With the property of the vec-opertation, $\text{vec}\,\mathbf{BC} = (\mathbf{C}' \otimes \mathbf{I}_n)\,\text{vec}\,\mathbf{B}$ for suitable matrices **B** and **C**, $\forall \mathbf{B} \in \mathbb{R}^{n\times m}$, $\forall \mathbf{C} \in \mathbb{R}^{m\times q}$ and the *Kronecker-Zehfuss product* $\mathbf{B} \otimes \mathbf{C}$ of two arbitrary matrices as well as $(\mathbf{B}+\mathbf{C}) \otimes \mathbf{D} = \mathbf{B} \otimes \mathbf{D} + \mathbf{C} \otimes \mathbf{D}$ of three arbitrary matrices subject to size **B** = size **C** we have

$$\frac{\partial \mathcal{L}_1}{\partial (\text{vec}\mathbf{L})}(\hat{\mathbf{L}}) = 2[(\mathbf{\Sigma}_y + \mathbf{A}\xi\xi'\mathbf{A}') \otimes \mathbf{I}_m]\,\text{vec}\,\hat{\mathbf{L}} - 2\,\text{vec}(\xi\xi'\mathbf{A}').$$

With the theory of matrix derivatives: Derivatives of a matrix-valued function of a matrix, namely $\partial f(\text{vec}\mathbf{X})/\partial(\text{vec}\mathbf{X})'$, we are now prepared to compute the second derivatives as

$$\frac{\partial^2 \mathcal{L}_1}{\partial(\text{vec}\,\mathbf{L})\partial(\text{vec}\,\mathbf{L})'}(\hat{\mathbf{L}}) = 2[(\mathbf{\Sigma}_y + \mathbf{A}\xi\xi'\mathbf{A}') \otimes \mathbf{I}_m].$$

Since $\mathbf{\Sigma}_y + \mathbf{A}\xi\xi'\mathbf{A}'$ is a *positive-definite matrix* the second derivatives constitute the sufficient conditions

$$\frac{\partial^2 \mathcal{L}_1}{\partial(\text{vec}\,\mathbf{L})\partial(\text{vec}\,\mathbf{L})'}(\hat{\mathbf{L}}) = 2[(\mathbf{\Sigma}_y + \mathbf{A}\xi\xi'\mathbf{A}') \otimes \mathbf{I}_m] > 0.$$

The *vec operation*, the *Kronecker-Zehfuss product* and the *derivatives of a matrix-valued function with respect to a matrix* are also reviewed in *Appendix A* of the book by Grafarend and Schaffrin (1993).

(ii) **S**-homBLE:

The hybrid norm $\| MSE_{\mathbf{S}}\{\hat{\xi}\} \|^2$ establishes the Lagrangean

$$\mathcal{L}_2(\mathbf{L}) := \text{tr}\,\mathbf{L}\mathbf{\Sigma}_y\mathbf{L}' + \text{tr}\,(\mathbf{I}_m - \mathbf{LA})\mathbf{S}(\mathbf{I}_m - \mathbf{LA})' = \min_{\mathbf{L}}$$

for $\hat{\xi}$ as ***S**- homBLE* of ξ . Following the first part of the proof we are led to the *necessary conditions* for the minimum of the *quadratic Lagrangean* $\mathcal{L}_2(\mathbf{L})$

$$\frac{\partial \mathcal{L}_2}{\partial \mathbf{L}}(\hat{\mathbf{L}}) := 2[\mathbf{\Sigma}_y\hat{\mathbf{L}}' + \mathbf{ASA}'\hat{\mathbf{L}}' - \mathbf{AS}]' = 0$$

as well as to the *sufficient conditions*

$$\frac{\partial^2 \mathcal{L}_2}{\partial(\text{vec}\mathbf{L})\partial(\text{vec}\mathbf{L})'}(\hat{\mathbf{L}}) = 2[(\mathbf{\Sigma}_y + \mathbf{ASA}') \otimes \mathbf{I}_m] > 0.$$

The *normal equations* of ***S**-homBLE* $\partial\mathcal{L}_2/\partial\mathbf{L}(\hat{\mathbf{L}}) = 0$ agree with (3.29)

(iii) α-homBLE:

The hybrid norm $\| MSE_{\alpha,\mathbf{S}}\{\hat{\xi}\} \|^2$ establishes the Lagrangean

$$\mathcal{L}_3(\mathbf{L}) := \text{tr}\,\mathbf{L}\mathbf{\Sigma}_y\mathbf{L}' + \frac{1}{\alpha}\text{tr}\,(\mathbf{I}_m - \mathbf{LA})\mathbf{S}(\mathbf{I}_m - \mathbf{LA})' = \min_{\mathbf{L}}$$

for $\hat{\xi}$ as *α-homBLE* of ξ . Following the first part of the proof we are led to the *necessary conditions* for the minimum of the *quadratic Lagrangean* $\mathcal{L}_3(\mathbf{L})$

$$\frac{\partial \mathcal{L}_3}{\partial \mathbf{L}}(\hat{\mathbf{L}}) = 2[\frac{1}{\alpha}\mathbf{ASA}'\hat{\mathbf{L}}' + \mathbf{\Sigma}_y\hat{\mathbf{L}}' - \frac{1}{\alpha}\mathbf{AS}]' = 0$$

as well as to the *sufficient conditions*

$$\frac{\partial^2 \mathcal{L}_3}{\partial(\text{vec}\mathbf{L})\partial(\text{vec}\mathbf{L})'}(\hat{\mathbf{L}}) = 2[(\mathbf{\Sigma}_y + \frac{1}{\alpha}\mathbf{ASA}') \otimes \mathbf{I}_m] > 0.$$

The *normal equations* of *α-homBLE* $\partial\mathcal{L}_3/\partial\mathbf{L}(\hat{\mathbf{L}}) = 0$ agree with (3.30) after the following transformation:

$$\hat{\mathbf{L}}(\alpha\mathbf{\Sigma}_y + \mathbf{ASA}') = \mathbf{SA}' \Leftrightarrow \hat{\mathbf{L}} = \mathbf{SA}'(\alpha\mathbf{\Sigma}_y + \mathbf{ASA}')^{-1} = \mathbf{SA}'\mathbf{\Sigma}_y^{-1}(\alpha\mathbf{I} + \mathbf{ASA}'\mathbf{\Sigma}_y^{-1})^{-1} = (\alpha\mathbf{I} + \mathbf{ASA}'\mathbf{\Sigma}_y^{-1})^{-1}\mathbf{SA}'\mathbf{\Sigma}_y^{-1}$$
$$\Leftrightarrow (\alpha\mathbf{I} + \mathbf{ASA}'\mathbf{\Sigma}_y^{-1})\hat{\mathbf{L}} = \mathbf{SA}'\mathbf{\Sigma}_y^{-1}.$$

♣

For an *explicit representation of* $\hat{\xi}$ as α- homBLE of ξ in the *special Gauss–Markov model with fixed effects* of *Box 3.1*, we solve the normal equations (3.28), (3.29) and (3.30) for

$$\hat{\mathbf{L}} = \arg\{\mathcal{L}(\mathbf{L}) = \min_{\mathbf{L}}\}.$$

Beside the *explicit representation* of $\hat{\xi}$ of type homBLE, $\mathbf{S}$ -homBLE and α- homBLE we present the related dispersion matrix $D\{\hat{\xi}\}$, the *Mean Square Error matrix* $MSE\{\hat{\xi}\}$, the modified *Mean Square Error matrices* $MSE_S\{\hat{\xi}\}$ and $MSE_{\alpha,S}\{\hat{\xi}\}$ in

Theorem 3.6 ($\hat{\xi}$ homBLE):

Let $\hat{\xi} = \mathbf{Ly}$ be *homBLE* of ξ in the *special linear Gauss-Markov model with fixed effects* of *Box 3.1*. Then equivalent representations of the solutions of the normal equations (3.28) are

$$\hat{\xi} = \xi\xi'\mathbf{A}'[\mathbf{\Sigma}_y + \mathbf{A}\xi\xi'\mathbf{A}']^{-1}\mathbf{y} = \xi[\xi'\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}\xi + \mathbf{I}]^{-1}\xi'\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{y} \tag{3.31}$$

complemented by the dispersion matrix

$$\begin{aligned} D\{\hat{\xi}\} &= \xi\xi'\mathbf{A}'[\mathbf{\Sigma}_y + \mathbf{A}\xi\xi'\mathbf{A}']^{-1}\mathbf{\Sigma}_y[\mathbf{\Sigma}_y + \mathbf{A}\xi\xi'\mathbf{A}']^{-1}\mathbf{A}\xi\xi' = \\ &= \xi[\xi'\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}\xi + \mathbf{I}]^{-1}\xi'\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}\xi[\xi'\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}\xi + \mathbf{I}]^{-1}\xi', \end{aligned} \tag{3.32}$$

by the *bias vector* (3.10)

$$\begin{aligned} \boldsymbol{\beta} &:= E\{\hat{\xi}\} - \xi = -[\mathbf{I}_m - \xi\xi'\mathbf{A}'(\mathbf{A}\xi\xi'\mathbf{A}' + \mathbf{\Sigma}_y)^{-1}\mathbf{A}]\xi = \\ &= -[\mathbf{I}_m - \xi[\xi'\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}\xi + \mathbf{I}]^{-1}\xi'\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}]\xi \end{aligned} \tag{3.33}$$

and by the *Mean Square Error matrix* $MSE\{\hat{\xi}\}$:

$$\begin{aligned} MSE\{\hat{\xi}\} &:= E\{(\hat{\xi} - \xi)(\hat{\xi} - \xi)'\} = D\{\hat{\xi}\} + \boldsymbol{\beta}\boldsymbol{\beta}' = \\ &= D\{\hat{\xi}\} + [\mathbf{I}_m - \xi\xi'\mathbf{A}'(\mathbf{A}\xi\xi'\mathbf{A}' + \mathbf{\Sigma}_y)^{-1}\mathbf{A}]\xi\xi'[\mathbf{I}_m - \mathbf{A}'(\mathbf{A}\xi\xi'\mathbf{A}' + \mathbf{\Sigma}_y)^{-1}\mathbf{A}\xi\xi']. \end{aligned} \tag{3.34}$$

At this point we have to comment what *Theorem 3.6* tells us. *homBLE* has generated the estimation $\hat{\xi}$ of type (3.31), the dispersion matrix $D\{\hat{\xi}\}$ of type (3.32), the bias vector of type (3.33) and the *Mean Square Error matrix* $MSE\{\hat{\xi}\}$ of type (3.34) *which all depend on the vector* ξ *and the matrix* $\xi\xi'$, respectively. We already mentioned that ξ and the matrix $\xi\xi'$ are *not* accessible from measurements. The situation is similar to the one in the *theory of hypothesis testing*. As shown later in this section we can produce only an estimator $\hat{\xi}$ and consequently can setup a *hypothesis* $\mathcal{H}_0$ of the "*fixed effects*" ξ. Indeed, a similar argument applies to the *second central moment* $D\{\mathbf{y}\} = \mathbf{\Sigma}_y$ of the "*random effect*" $\mathbf{y}$, the observation vector. Such a dispersion matrix has to be known in order to be able to compute $\hat{\xi}$, $D\{\hat{\xi}\}$, and $MSE\{\hat{\xi}\}$. Again we have to apply the argument that we are only able to construct an estimate $\hat{\mathbf{\Sigma}}_y$ and to set-up a *hypothesis* about $D\{\mathbf{y}\} = \mathbf{\Sigma}_y$.

Theorem 3.7 ($\hat{\xi}$ **S**-homBLE):

Let $\hat{\xi} = \mathbf{Ly}$ be ***S****-homBLE* of ξ in the *special linear Gauss- Markov model with fixed effects* of *Box 3.1*. Then equivalent representations of the solutions of the normal equations (3.29) are

$$\hat{\xi} = \mathbf{SA}'(\mathbf{\Sigma}_y + \mathbf{ASA}')^{-1}\mathbf{y} \tag{3.35}$$

$$\hat{\xi} = (\mathbf{I}_m + \mathbf{SA}'\mathbf{\Sigma}_y^{-1}\mathbf{A})^{-1}\mathbf{SA}'\mathbf{\Sigma}_y^{-1}\mathbf{y} \tag{3.36}$$

$$\hat{\xi} = (\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A} + \mathbf{S}^{-1})^{-1}\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{y} \tag{3.37}$$

complemented by the dispersion matrices

$$D\{\hat{\xi}\} = \mathbf{SA}'(\mathbf{ASA}' + \mathbf{\Sigma}_y)^{-1}\mathbf{\Sigma}_y(\mathbf{ASA}' + \mathbf{\Sigma}_y)^{-1}\mathbf{AS} \tag{3.38}$$

$$D\{\hat{\xi}\} = (\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A} + \mathbf{S}^{-1})^{-1}\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}(\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A} + \mathbf{S}^{-1})^{-1} \tag{3.39}$$

by the *bias vector* (3.10)

$$\begin{aligned}\boldsymbol{\beta} &:= E\{\hat{\xi}\} - \xi = \\ &= -[\mathbf{I}_m - \mathbf{SA}'(\mathbf{ASA}' + \boldsymbol{\Sigma}_y)^{-1}\mathbf{A}]\xi = \\ &= -[\mathbf{I}_m - (\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A} + \mathbf{S}^{-1})^{-1}\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}]\xi = \\ &= -(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A} + \mathbf{S}^{-1})^{-1}\mathbf{S}^{-1}\xi = \\ &= -(\mathbf{SA}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A} + \mathbf{I}_m)^{-1}\xi\end{aligned} \tag{3.40}$$

and by the *Mean Square Error* matrix $MSE\{\hat{\xi}\}$:

$$\begin{aligned}&MSE\{\hat{\xi}\} := E\{(\hat{\xi}-\xi)(\hat{\xi}-\xi)'\} = D\{\hat{\xi}\} + \boldsymbol{\beta\beta}' = \\ &= \mathbf{SA}'(\mathbf{ASA}'+\boldsymbol{\Sigma}_y)^{-1}\boldsymbol{\Sigma}_y(\mathbf{ASA}'+\boldsymbol{\Sigma}_y)^{-1}\mathbf{AS} + [\mathbf{I}_m - \mathbf{SA}'(\mathbf{ASA}'+\boldsymbol{\Sigma}_y)^{-1}\mathbf{A}]\xi\xi'[\mathbf{I}_m - \mathbf{A}'(\mathbf{ASA}'+\boldsymbol{\Sigma}_y)^{-1}\mathbf{AS}] = \\ &= (\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\mathbf{S}^{-1})^{-1}\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\mathbf{S}^{-1})^{-1} + [(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\mathbf{S}^{-1})^{-1}\mathbf{S}^{-1}]\xi\xi'[\mathbf{S}^{-1}(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\mathbf{S}^{-1})^{-1}].\end{aligned} \tag{3.41}$$

But the **S**- modified *Mean Square Error* matrix $MSE_s\{\hat{\xi}\}$:

$$MSE_s\{\hat{\xi}\} = (\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\mathbf{S}^{-1})^{-1}. \tag{3.42}$$

The interpretation of ***S****-homBLE* is even more complex. In extension of the comments to *homBLE* we have to live with another matrix-valued degree of freedom, $\hat{\xi}$ of type (3.35), (3.36), (3.37) and $D\{\hat{\xi}\}$ of type (3.38), (3.39) do no longer depend on the inaccessible matrix $\xi\xi'$, $\mathrm{rk}(\xi\xi') = 1$, but on the "*weight of the bias matrix*" **S**, rk **S** = *m*. Indeed we can associate any element of the bias matrix with a particular weight which can be "*designed*" by the analyst. Again the *bias vector* $\boldsymbol{\beta}$ of type (3.40) as well as the *Mean Square Error* of type (3.41) depend on the vector ξ which is inaccessible. Beside the dependence on the "*weight of the bias matrix* **S**", the quantities $\hat{\xi}$, $D\{\hat{\xi}\}$, $\boldsymbol{\beta}$ and $MSE\{\hat{\xi}\}$ are vector-valued or matrix-valued functions of the *dispersion matrix* $D\{\mathbf{y}\} = \boldsymbol{\Sigma}_y$ which is inaccessible. By *hypothesis testing* we may decide upon the construction of $D\{\mathbf{y}\} = \boldsymbol{\Sigma}_y$ from an estimate $\hat{\boldsymbol{\Sigma}}_y$.

Theorem 3.8 ($\hat{\xi}$ α-homBLE, also known as: ridge estimator or Tykhonov-Phillips regulator)

Let $\hat{\xi} = \mathbf{Ly}$ be *α- homBLE* of ξ in the *special linear Gauss-Markov model with fixed effects* of *Box 3.1*. Then equivalent representations of the solutions of the normal equations (3.30) are

$$\begin{aligned}\hat{\xi} &= (\mathbf{SA}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A} + \alpha\mathbf{I}_m)^{-1}\mathbf{SA}'\boldsymbol{\Sigma}_y^{-1}\mathbf{y} \\ &= (\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A} + \alpha\mathbf{S}^{-1})^{-1}\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{y}\end{aligned} \tag{3.43}$$

complemented by the dispersion matrix

$$\begin{aligned}&D\{\hat{\xi}\} = \\ &= (\mathbf{SA}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A} + \alpha\mathbf{I}_m)^{-1}\mathbf{SA}'\boldsymbol{\Sigma}_y^{-1}\mathbf{AS}'(\mathbf{SA}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A} + \alpha\mathbf{I}_m)^{-1} \\ &= (\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A} + \alpha\mathbf{S}^{-1})^{-1}\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A} + \alpha\mathbf{S}^{-1})^{-1}\end{aligned} \tag{3.44}$$

by the *bias vector* (3.10)

$$\begin{aligned}\boldsymbol{\beta} &:= E\{\hat{\xi}\} - \xi = \\ &= -[\mathbf{I}_m - (\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A} + \alpha\mathbf{S}^{-1})^{-1}\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}]\xi \\ &= -\alpha(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A} + \alpha\mathbf{S}^{-1})^{-1}\mathbf{S}^{-1}\xi \\ &= -\alpha(\mathbf{SA}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A} + \alpha\mathbf{I}_m)^{-1}\xi\end{aligned} \tag{3.45}$$

and by the *Mean Square Error matrix* $MSE\{\hat{\xi}\}$

$$\begin{aligned}&MSE\{\hat{\xi}\} := E\{(\hat{\xi}-\xi)(\hat{\xi}-\xi)'\} = D\{\hat{\xi}\} + \boldsymbol{\beta\beta}' = \\ &= (\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1} + [(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\alpha\mathbf{S}^{-1}]\xi\xi'[\alpha\mathbf{S}^{-1}(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}] = \\ &= (\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}[\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+(\alpha\mathbf{S}^{-1})\xi\xi'(\alpha\mathbf{S}^{-1})](\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}.\end{aligned} \tag{3.46}$$

But the hybrid α-weighted variance-bias norm $MSE_{\alpha,s}\{\hat{\xi}\}$

$$MSE_{\alpha,s}\{\hat{\xi}\} = (\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A} + \alpha\mathbf{S}^{-1})^{-1}. \tag{3.47}$$

The interpretation of the very important estimator α-homBLE $\hat{\boldsymbol{\xi}}$ of $\boldsymbol{\xi}$ is as follows: $\hat{\boldsymbol{\xi}}$ of type (3.43), also called *ridge estimator* or *Tykhonov-Phillips regulator*, contains the *Cayley inverse* of the normal equation matrix *which is additively composed* of $\mathbf{A}'\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\mathbf{A}$ and $\alpha\mathbf{S}^{-1}$. The weight factor α balances the first observational weight and the second bias weight within the inverse. While the experiment informs us of the variance-covariance matrix $\boldsymbol{\Sigma}_{\mathbf{y}}$, say $\hat{\boldsymbol{\Sigma}}_{\mathbf{y}}$, the *weight of the bias weight matrix* and the *weight factor* α are at the disposal of the analyst. For instance, by the choice **S=** Diag$(s_1,...,s_m)$ we may emphasize an increase or decrease of certain bias matrix elements. The choice of an equally weighted bias matrix is $\mathbf{S}=\mathbf{I}_m$. In contrast, the weight factor α can be alternatively determined by the *A-optimal design* of type

- $\operatorname{tr} D\{\hat{\boldsymbol{\xi}}\} = \min_{\alpha}$, or
- $\operatorname{tr} \boldsymbol{\beta}\boldsymbol{\beta}' = \min_{\alpha}$, or
- $\operatorname{tr} MSE\{\hat{\boldsymbol{\xi}}\} = \min_{\alpha}$.

In the *first case* we optimize the *trace of the variance-covariance matrix* $D\{\hat{\boldsymbol{\xi}}\}$ of type (3.44) . Alternatively by means of $\operatorname{tr}\boldsymbol{\beta}\boldsymbol{\beta}' = \min_{\alpha}$ we optimize the *quadratic bias* where the bias vector $\boldsymbol{\beta}$ of type (3.45) is chosen, regardless of the dependence on $\boldsymbol{\xi}$. Finally for the *third case* – the most meaningful one – we minimize the trace of the *Mean Square Error matrix MSE*$\{\hat{\boldsymbol{\xi}}\}$ of type (3.46), despite of the dependence on $\boldsymbol{\xi}\boldsymbol{\xi}'$. Here we concentrate on the third case and the main result is summarized in

Theorem 3.9 (A-optimal design of α):

Let the *Mean Square Error matrix MSE*$\{\hat{\boldsymbol{\xi}}\}$ of α- homBLE $\hat{\boldsymbol{\xi}}$ with respect to the linear Gauss Markov model be given by

$$\operatorname{tr} MSE\{\hat{\boldsymbol{\xi}}\} =$$
$$= \operatorname{tr}(\mathbf{A}'\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\mathbf{A}'\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\mathbf{A}(\mathbf{A}'\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}+$$
$$+\operatorname{tr}[(\mathbf{A}'\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\alpha\mathbf{S}^{-1}]\boldsymbol{\xi}\boldsymbol{\xi}'[\alpha\mathbf{S}^{-1}(\mathbf{A}'\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}],$$

then $\hat{\alpha}$ follows by A-optimal design in the sense of

$$\operatorname{tr} MSE\{\hat{\boldsymbol{\xi}}\} = \min$$

if and only if

$$\hat{\alpha} = \frac{\operatorname{tr}\mathbf{A}'\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\mathbf{A}(\mathbf{A}'\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-2}\mathbf{S}^{-1}(\mathbf{A}'\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}}{\boldsymbol{\xi}'\mathbf{S}^{-1}(\mathbf{A}'\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-2}\mathbf{A}'\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\mathbf{A}(\mathbf{A}'\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\mathbf{S}^{-1}\boldsymbol{\xi}} \tag{3.48}$$

The proof of *Theorem 3.9* is given in the *Appendix 3-A*. The subject of *optimal design* within Mathematical Statistics has been studied since the nineteen sixties. For more detail let us refer to *R.B. Bapat* (1999), *D.R. Cox* and *N. Reid* (2000), *E. P. Liski, et al*. (2002), *A. Pazman* (1986) and *F. Pukelsheim* (1993).

For the independent, identically distributed (i.i.d) observations *Theorem 3.9* will be simplified as:

Corollary 3.10 (A-optimal design of α for the special Gauss-Markov model with i.i.d. observations):

For the special Gauss-Markov model

$$\mathbf{A}\boldsymbol{\xi} = E\{\mathbf{y}\},\ \mathbf{I}_n\sigma^2 = \boldsymbol{\Sigma}_{\mathbf{y}} = D\{\mathbf{y}\},\ \mathbf{I}_m\sigma^2 = \mathbf{S} \tag{3.49}$$

of independent, identically distributed (i.i.d.) observations with variance σ^2 and an analogous substitute weight matrix **S** scaled by the variance σ^2, an A-optimal choice of the weighting factor α is

$$\hat{\alpha} = \frac{\operatorname{tr}\mathbf{A}'\mathbf{A}(\mathbf{A}'\mathbf{A}+\hat{\alpha}\mathbf{I}_m)^{-3}\sigma^2}{\boldsymbol{\xi}'(\mathbf{A}'\mathbf{A}+\hat{\alpha}\mathbf{I}_m)^{-2}\mathbf{A}'\mathbf{A}(\mathbf{A}'\mathbf{A}+\hat{\alpha}\mathbf{I}_m)^{-1}\boldsymbol{\xi}} \tag{3.50}$$

For the case of i.i.d. observations of a random scalar parameter case (direct observations) the α-homBLE of ξ and BIQUUE of σ^2 are summarized in *Box 3.3* and *Corollary 3.11*.

***Box 3.3*:**

Special Gauss-Markov model: direct observations

$$\boldsymbol{\tau}\xi = E\{\mathbf{y}\},\ \mathbf{I}_n\sigma^2 = \boldsymbol{\Sigma}_y = D\{\mathbf{y}\},\ \boldsymbol{\tau} = [1,...,1]',\ \mathrm{S} = 1\sigma^2 = [\sigma^2]$$

"α-homBLE of ξ"

$$\hat{\xi} = \frac{1}{n+\alpha}\boldsymbol{\tau}'\mathbf{y} = \frac{n}{n+\alpha}\left(\frac{1}{n}\sum_i y_i\right) = \frac{n}{n+\alpha}\bar{\xi} \tag{3.51}$$

"dispersion matrix"

$$D\{\hat{\xi}\} = \sigma^2 \frac{n}{(n+\alpha)^2} \tag{3.52}$$

"bias"

$$\beta = -\frac{\alpha}{n+\alpha}\xi \tag{3.53}$$

$$\begin{aligned} MSE\{\hat{\xi}\} &= D\{\hat{\xi}\} + \boldsymbol{\beta}\boldsymbol{\beta}' = \\ &= \sigma^2 \frac{n}{(n+\alpha)^2} + \frac{\alpha^2}{(n+\alpha)^2}\xi^2 = \\ &= \gamma_1(\alpha) + \gamma_2(\alpha). \end{aligned} \tag{3.54}$$

"BIQUUE of σ^2"

$$\bar{\sigma}^2 = \frac{1}{n-1}\mathbf{y}'(\mathbf{I}_n - \frac{1}{n}\mathbf{J}_n)\mathbf{y},\ \mathbf{J}_n = \boldsymbol{\tau}\boldsymbol{\tau}'. \tag{3.55}$$

In the case of the *special Gauss-Markov model of direct observations* the first order design matrix **A** is of full rank 1. Accordingly, an estimation $\bar{\xi}$ of type *BLUUE* (Best Linear Uniformly Unbiased Estimation) exists and may be used to replace ξ. Although we have so far treated $\mathrm{S} = [\sigma^2]$ as known, we note that, in this particular case, we may treat the variance factor σ^2 as a common unknown and resort to a classical estimation $\bar{\sigma}^2$ of type *BIQUUE* (Best Invariant Quadratic Uniformly Unbiased Estimation), which is a useful substitute of σ^2 in computing the weight α.

Corollary 3.11 (A-optimal design of α for the special Gauss-Markov model with direct i.i.d. observations):

Let us replace (i) ξ by $\bar{\xi}$(BLUUE) and (ii) σ^2 by $\bar{\sigma}^2$(BIQUUE) within the A-optimal choice of the weighting factor $\hat{\alpha}$, Eq. (3.50), with respect to the special Gauss- Markov model Eq. (3.49). Then an approximation $\tilde{\alpha}$ of the A-optimal choice $\hat{\alpha}$, namely

(3.56) $$\tilde{\alpha} = \frac{\bar{\sigma}^2}{\bar{\xi}^2} \qquad \lim_{n\to\infty}\tilde{\alpha} = \hat{\alpha} = \frac{\sigma^2}{\xi^2} \tag{3.57}$$

is obtained with

$$\mathrm{tr}\, MSE\{\hat{\xi}\}|_{\bar{\sigma}^2,\bar{\xi}} = MSE\{\hat{\xi}\}|_{\bar{\sigma}^2,\bar{\xi}} = \bar{\sigma}^2 \frac{n}{(n+\alpha)^2} + \frac{\alpha^2}{(n+\alpha)^2}\bar{\xi}^2 = \bar{\sigma}^2 \frac{n+\alpha^2\bar{\xi}^2}{(n+\alpha)^2}\ . \tag{3.58}$$

Now we would like to compare our solution with the famous *ridge regression* developed by *Hoerl and Kennard* (1970a, 1970b). First, the A-optimal design of α derived by (3.57) is just the same as the optimum *ridge parameter k* (Hoerl and Kennard 1970a, b). Second, *Hoerl, Kennard and Baldwin* (1975) have suggested that if $\mathbf{A}'\mathbf{A} = \mathbf{I}_m$, then a minimum meansquare error (MSE) is obtained if *ridge parameter* $k = m\sigma^2/\boldsymbol{\xi}'\boldsymbol{\xi}$ for multiple linear regression model $\mathbf{y} = \mathbf{A}\boldsymbol{\xi} + \mathbf{e}$ where $\mathbf{A} \in \mathbb{R}^{n\times m}$, rk $\mathbf{A} = m$, $E\{\mathbf{e}\} = \mathbf{0}$ and $D\{\mathbf{y}\} = \mathbf{I}_n\sigma^2$ with σ^2 is chosen unknown. This is just the special case of our general solution (3.50) by A-optimal design under $\mathbf{A}'\mathbf{A} = \mathbf{I}_m$, yielding

$$\begin{aligned} \hat{\alpha} &= \frac{\mathrm{tr}\,\mathbf{A}'\mathbf{A}(\mathbf{A}'\mathbf{A} + \hat{\alpha}\mathbf{I}_m)^{-3}\sigma^2}{\boldsymbol{\xi}'(\mathbf{A}'\mathbf{A} + \hat{\alpha}\mathbf{I}_m)^{-2}\mathbf{A}'\mathbf{A}(\mathbf{A}'\mathbf{A} + \hat{\alpha}\mathbf{I}_m)^{-1}\boldsymbol{\xi}} = \frac{\mathrm{tr}\,\mathbf{I}_m(\mathbf{I}_m + \hat{\alpha}\mathbf{I}_m)^{-3}\sigma^2}{\boldsymbol{\xi}'(\mathbf{I}_m + \hat{\alpha}\mathbf{I}_m)^{-2}\mathbf{I}_m(\mathbf{I}_m + \hat{\alpha}\mathbf{I}_m)^{-1}\boldsymbol{\xi}} = \\ &= \frac{m(1+\hat{\alpha})^{-3}\sigma^2}{\boldsymbol{\xi}'\boldsymbol{\xi}(1+\hat{\alpha})^{-3}} = \frac{m\sigma^2}{\boldsymbol{\xi}'\boldsymbol{\xi}}. \end{aligned}$$

3.2 The special multivariate *Gauss-Markov* model and the multivariate α-BLE

For the case of a 2-D symmetric rank-two random strain rate tensor we have to estimate as multivariate parameters, the three vectorized elements from the direct observations of them, namely via BLUUE of t_{11}, t_{12}, t_{22} and BIQUUE of the related variances, which are estimated with the multivariate *Gauss-Markov* model. We shall first generalize the special *Gauss-Markov* model to the multivariate *Gauss-Markov* model and derive the $\alpha-$BLE for the multivariate estimation, which we call "*multivariate ridge type estimation*".

Let us introduce the multivariate *Gauss-Markov* model $\mathbf{Y}=\mathbf{A\Xi}+\mathbf{E}$ in *Box 3.4*, which is given for the *first order moments* in the form of a *consistent system of linear equations relating* the first *non-stochastic* ("fixed"), *real-valued matrix* $\mathbf{\Xi}$ of *unknowns* to the expectation $E\{\mathbf{Y}\}$ of the *stochastic*, real-valued matrix $\mathbf{Y}$ of observations, $\mathbf{A\Xi}=E\{\mathbf{Y}\}$, since $\mathcal{R}(E\{\mathbf{Y}\})\subset\mathcal{R}(\mathbf{A})$ is a subspace of the column space $\mathcal{R}(\mathbf{A})$ of the real-valued, *non-stochastic* ("fixed") "*first order design matrix*" $\mathbf{A}\in\mathbb{R}^{n\times m}$, here, the symbols denote:

$$\mathbf{A}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix},\quad \mathbf{A}\in\mathbb{R}^{n\times m},\quad \mathbf{\Xi}=\begin{bmatrix} \xi_{11} & \xi_{12} & \cdots & \xi_{1p} \\ \xi_{21} & \xi_{22} & \cdots & \xi_{2p} \\ \vdots & \vdots & \cdots & \vdots \\ \xi_{m1} & \xi_{m2} & \cdots & \xi_{mp} \end{bmatrix}=\begin{bmatrix} \boldsymbol{\xi}_1 & \boldsymbol{\xi}_2 & \cdots & \boldsymbol{\xi}_p \end{bmatrix},$$

$$\mathbf{Y}=\begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1p} \\ y_{21} & y_{22} & \cdots & y_{2p} \\ \vdots & \vdots & \cdots & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{np} \end{bmatrix}=\begin{bmatrix} \mathbf{Y}_1 & \mathbf{Y}_2 & \cdots & \mathbf{Y}_p \end{bmatrix}=\begin{bmatrix} \mathbf{y}_1' \\ \mathbf{y}_2' \\ \vdots \\ \mathbf{y}_n' \end{bmatrix},\qquad \mathbf{Y}\in\mathbb{R}^{n\times p}.$$

In addition, the *second order central moments* $D\{\mathrm{vec}\,\mathbf{Y}\}$, the *regular* variance-covariance matrix $\mathbf{\Sigma}_{\mathrm{vec}\,\mathbf{y}}$, also called *dispersion matrix* should be defined in the multivariate case as follows: Let $\mathbf{Y}_\mathrm{i}$ be the $n\times1$ random vectors of observations of p characteristics and let $\mathrm{Cov}\{\mathbf{Y}_i,\mathbf{Y}_j\}=\sigma_{ij}\mathbf{I}_n$, where the covariance matrix with $\mathbf{\Sigma}=(\sigma_{ij})$, $\mathbf{\Sigma}\in\mathbb{R}^{p\times p}$ is unknown and positive-definite. Since $D\{\mathbf{Y}_i\}=\mathrm{Cov}\{\mathbf{Y}_i,\mathbf{Y}_i\}=\sigma_i^2\mathbf{I}_n$, the components of the vectors $\mathbf{Y}_\mathrm{i}$ of observations are uncorrelated and have equal variance (i.i.d.). We have the vectorized form of $\mathbf{Y}$

$$\mathrm{vec}\,\mathbf{Y}=\begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \\ \vdots \\ \mathbf{Y}_p \end{bmatrix},\quad \mathrm{vec}\,\mathbf{Y}\in\ \mathbb{R}^{np\times1},$$

whose varoiance-covariance matrix follows as

$$D\{\mathrm{vec}\,\mathbf{Y}\}=\begin{bmatrix} \sigma_1^2\mathbf{I}_n & \sigma_{12}\mathbf{I}_n & \cdots & \sigma_{1p}\mathbf{I}_n \\ \sigma_{21}\mathbf{I}_n & \sigma_2^2\mathbf{I}_n & \cdots & \sigma_{2p}\mathbf{I}_n \\ \cdots & \cdots & \ddots & \vdots \\ \sigma_{p1}\mathbf{I}_n & \sigma_{p2}\mathbf{I}_n & \cdots & \sigma_p^2\mathbf{I}_n \end{bmatrix}=\mathbf{\Sigma}\otimes\mathbf{I}_n,\quad D\{\mathrm{vec}\,\mathbf{Y}\}\in\mathbb{R}^{pn\times pn}.$$

To better understand the meaning of the variance-covariance matrix of the observation matrix we study the transposed form of $\mathbf{Y}$

$$\mathbf{Y}'=[\mathbf{y}_1,\ \mathbf{y}_2,\ \cdots,\ \mathbf{y}_n],$$

where the columns $\mathbf{y}_1,\mathbf{y}_2,\ldots,\mathbf{y}_n$ are independent $p\times1$ random vectors, each with the same variance-covariance matrix $\mathbf{\Sigma}_\mathbf{y}=\mathbf{\Sigma}$. We then have the $pn\times1$ vectorized form of observations

$$\mathrm{vec}\,\mathbf{Y}'=\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_n \end{bmatrix},\quad \mathrm{vec}\,\mathbf{Y}'\in\ \mathbb{R}^{pn\times1}$$

whose variance-covariance matrix follows as

$$D\{\mathrm{vec}\,\mathbf{Y}'\}=\begin{bmatrix} \mathbf{\Sigma}_\mathbf{y} & 0 & \cdots & 0 \\ 0 & \mathbf{\Sigma}_\mathbf{y} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \mathbf{\Sigma}_\mathbf{y} \end{bmatrix}=\mathbf{I}_n\otimes\mathbf{\Sigma},\quad D\{\mathrm{vec}\,\mathbf{Y}'\}\ \in\ \mathbb{R}^{pn\times pn} \tag{3.59}$$

Box 3.4:
Multivariate Gauss-Markov model

$$\mathbf{Y} = \mathbf{A\Xi} + \mathbf{E}$$

1st moments

$$\mathbf{A\Xi} = E\{\mathbf{Y}\},\ \mathbf{A} \in \mathbb{R}^{n\times m}, \mathbf{Y} \in \mathbb{R}^{n\times p}\ E\{\mathbf{Y}\} \in \mathcal{R}(\mathbf{A}),\ \text{rk}\,\mathbf{A} = m \tag{3.60}$$

2nd moments

$$D\{\text{vec}\,\mathbf{Y}\} = \mathbf{\Sigma} \otimes \mathbf{I}_n \in \mathbb{R}^{pn\times np},\ \mathbf{\Sigma}\ \text{positive-definite, rk}\,\mathbf{\Sigma} = p$$
$$\mathbf{\Xi},\ E\{\mathbf{Y}\},\ \mathbf{Y} - E\{\mathbf{Y}\} = \mathbf{E}\ \textit{unknown},\ \mathbf{\Sigma}\ \textit{unknown}. \tag{3.61}$$

With the notation introduced above the multivariate *Gauss-Markov* model (3.60) can be presented as

$$\mathbf{A}[\xi_1,\ \xi_2, \cdots, \xi_p] = E\{\mathbf{Y}_1,\ \mathbf{Y}_2, \cdots, \mathbf{Y}_p\},\ D\{\text{vec}\,\mathbf{Y}\} = \mathbf{\Sigma} \otimes \mathbf{I}_n$$

which can be further described as

$$\mathbf{A}\xi_i = E\{\mathbf{Y}_i\},\ \ D\{\mathbf{Y}_i\} = \sigma_i^2 \mathbf{I}_n,\ \text{Cov}(\mathbf{Y}_i, \mathbf{Y}_j) = \sigma_{ij}\mathbf{I}_n,\ \text{for}\ i, j \in \{1,..., p\}.$$

If the observations are normally distributed, we find with *Theorem 3.12*

$$\text{vec}\,\mathbf{Y} \sim \mathcal{N}(\text{vec}(\mathbf{A\Xi}),\ \mathbf{\Sigma} \otimes \mathbf{I}_n)$$

and the vectors $\mathbf{Y}_\text{i}$ of observations are distributed as

$$\mathbf{Y}_i \sim \mathcal{N}(\mathbf{A}\xi_i),\ \sigma_i^2 \mathbf{I}_n),\ \ \text{for}\ i \in \{1,..., p\}\ ,$$

for which the *theorem* 3.1 is valid. To study the estimation of $\mathbf{\Xi}$ and $\mathbf{\Sigma}$, we will represent model (3.60) in the form of a vector. Since $\text{vec}(\mathbf{A\Xi}) = (\mathbf{I}_p \otimes \mathbf{A})\,\text{vec}\,\mathbf{\Xi}$, (3.60) will become

$$(\mathbf{I}_p \otimes \mathbf{A})\,\text{vec}\,\mathbf{\Xi} = E\{\text{vec}\,\mathbf{Y}\},\ \ \text{with}\ D\{\text{vec}\,\mathbf{Y}\} = \mathbf{\Sigma} \otimes \mathbf{I}_n. \tag{3.62}$$

Applying the result (3.3) of the univariate *Gauss-Markov* model and the Kronecker-Zehfuß product yields the BLUUE of $\text{vec}\,\mathbf{\Xi}$ as

$$\begin{aligned}\text{vec}\,\hat{\mathbf{\Xi}} &= [(\mathbf{I}_p \otimes \mathbf{A})'(\mathbf{\Sigma} \otimes \mathbf{I}_n)^{-1}(\mathbf{I}_p \otimes \mathbf{A})]^{-1}(\mathbf{I}_p \otimes \mathbf{A})'(\mathbf{\Sigma} \otimes \mathbf{I}_n)^{-1}\,\text{vec}\,\mathbf{Y} = \\ &= (\mathbf{I}_p\mathbf{\Sigma}^{-1}\mathbf{I}_p \otimes \mathbf{A}'\mathbf{I}_n\mathbf{A})^{-1}(\mathbf{I}_p\mathbf{\Sigma}^{-1} \otimes \mathbf{A}\mathbf{I}_n)'\,\text{vec}\,\mathbf{Y} \\ &= (\mathbf{\Sigma}^{-1} \otimes \mathbf{A}'\mathbf{A})^{-1}(\mathbf{\Sigma}^{-1} \otimes \mathbf{A})'\,\text{vec}\,\mathbf{Y} \\ &= [\mathbf{I}_p \otimes (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}']\,\text{vec}\,\mathbf{Y} \\ &= \text{vec}[(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{Y}],\end{aligned} \tag{3.63}$$

and with (3.5) the covariance matrix of $D(\text{vec}\,\hat{\mathbf{\Xi}})$ is

$$D(\text{vec}\,\hat{\mathbf{\Xi}}) = [\mathbf{I}_p \otimes (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'](\mathbf{\Sigma} \otimes \mathbf{I}_n)[\mathbf{I}_p \otimes (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}']' = \mathbf{\Sigma} \otimes (\mathbf{A}'\mathbf{A})^{-1}. \tag{3.64}$$

Thus the BLUUE of $\mathbf{\Xi}$ is

$$\hat{\mathbf{\Xi}} = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{Y}, \tag{3.65}$$

Hence

$$\hat{\mathbf{\Xi}} = [\hat{\xi}_1,\ \hat{\xi}_2, \cdots, \hat{\xi}_p] = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'[\mathbf{Y}_1,\ \mathbf{Y}_2, \cdots, \mathbf{Y}_p]\ . \tag{3.66}$$

These results are collected in

***Theorem* 3.12** ($\hat{\mathbf{\Xi}}$ BLUUE of $\mathbf{\Xi}$):

The BLUUE $\hat{\mathbf{\Xi}}$ of $\mathbf{\Xi}$ in the *special multivariate linear Gauss-Markov model* (3.60) and (3.61) is

$$\text{vec}\,\hat{\mathbf{\Xi}} = [\mathbf{I}_p \otimes (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}']\,\text{vec}\,\mathbf{Y} \tag{3.63}$$

or

$$\hat{\mathbf{\Xi}} = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{Y}, \tag{3.65}$$

with the related dispersion matrix

$$D(\text{vec}\,\hat{\mathbf{\Xi}}) = \mathbf{\Sigma} \otimes (\mathbf{A}'\mathbf{A})^{-1}\ . \tag{3.64}$$

Note that, when we use the expression (3.59) of variance-covariance matrix, formula (3.62) will be replaced by

$$(\mathbf{A}\otimes\mathbf{I}_p)\operatorname{vec}\boldsymbol{\Xi}' = E\{\operatorname{vec}\mathbf{Y}'\},\ \text{ with } D\{\operatorname{vec}\mathbf{Y}'\} = \mathbf{I}_n\otimes\boldsymbol{\Sigma}, \tag{3.67}$$

since the Kronecker-Zehfuß product $\operatorname{vec}\boldsymbol{\Xi}'\mathbf{A}' = (\mathbf{A}\otimes\mathbf{I}_p)\operatorname{vec}\boldsymbol{\Xi}'$ for the transposed form of (3.60) is

$$\boldsymbol{\Xi}'\mathbf{A}' = E\{\mathbf{Y}'\},\ \mathbf{A}\in\mathbb{R}^{n\times m}, \mathbf{Y}\in\mathbb{R}^{n\times p}\ \mathcal{R}(E\{\mathbf{Y}\})\in\mathcal{R}(\mathbf{A}),\ \operatorname{rk}\mathbf{A} = m. \tag{3.68}$$

From (3.67) we can get the BLUUE of $\operatorname{vec}\boldsymbol{\Xi}'$ is

$$\begin{aligned}\operatorname{vec}\hat{\boldsymbol{\Xi}}' &= [(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\otimes\mathbf{I}_p]\operatorname{vec}\mathbf{Y}'\\ &= \operatorname{vec}[\mathbf{Y}'\mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}],\end{aligned} \tag{3.69}$$

with the related dispersion matrix

$$\begin{aligned}D(\operatorname{vec}\hat{\boldsymbol{\Xi}}') &= [(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\otimes\mathbf{I}_p](\mathbf{I}_n\otimes\boldsymbol{\Sigma})[(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\otimes\mathbf{I}_p]' =\\ &= (\mathbf{A}'\mathbf{A})^{-1}\otimes\boldsymbol{\Sigma}.\end{aligned} \tag{3.70}$$

Since the variance-covariance matrix $\boldsymbol{\Sigma}$ of the individual observation vectors $\mathbf{y}_i$ is unknown, it is to be estimated empirically. We have therefore derived the sample variance-covariance matrix $\hat{\boldsymbol{\Sigma}}$ of type BIQUUE (*Best Invariant Quadratic Uniformly Unbiased Estimation*) which is collected in *Theorem* 3.13 (without proof).

Theorem 3.13 (The sample variance-covariance matrix $\hat{\boldsymbol{\Sigma}}$ of type BIQUUE**):**

The sample variance-covariance matrix $\hat{\boldsymbol{\Sigma}}$ of type BIQUUE for the *special multivariate linear Gauss-Markov model* (3.60) and (3.61) is

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{n-\operatorname{rk}\mathbf{A}}\mathbf{Y}'(\mathbf{I}_n - \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}')\mathbf{Y}. \tag{3.71}$$

For the proof, we refer to *K. R. Koch* (1997, 1999)

Now we shall derive the $\alpha-$BLE for the multivariate estimation of $\boldsymbol{\Xi}$, which we call "multivariate ridge type estimation". Following the derivation of the $\alpha-$BLE of *Theorem 3.8* for the univariate model and comparing the generalization results (3.66) for the multivariate model with (3.3) of the univariate model, we have the $\alpha-$BLE of the multivariate parameters readily in

Theorem 3.14 (*multivariate* α-homBLE, or *multivariate ridge estimator*)

The $\alpha-$BLE of the multivariate parameters in the *special multivariate linear Gauss-Markov model* (3.60) and (3.61) is

$$\operatorname{vec}\hat{\boldsymbol{\Xi}} = [\mathbf{I}_p\otimes(\mathbf{A}'\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\mathbf{A}']\operatorname{vec}\mathbf{Y} \tag{3.72}$$

or

$$\hat{\boldsymbol{\Xi}} = [\hat{\xi}_1,\ \hat{\xi}_2,\dots,\ \hat{\xi}_p] = (\mathbf{A}'\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\mathbf{A}'[\mathbf{Y}_1,\ \mathbf{Y}_2,\dots,\mathbf{Y}_p]. \tag{3.73}$$

complemented by the dispersion matrix

$$\begin{aligned}D(\operatorname{vec}\hat{\boldsymbol{\Xi}}) &= [\mathbf{I}_p\otimes(\mathbf{A}'\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\mathbf{A}'](\boldsymbol{\Sigma}\otimes\mathbf{I}_n)[\mathbf{I}_p\otimes(\mathbf{A}'\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\mathbf{A}']' =\\ &= \boldsymbol{\Sigma}\otimes(\mathbf{A}'\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\mathbf{A}'\mathbf{A}(\mathbf{A}'\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1},\end{aligned} \tag{3.74}$$

by the *bias vector*

$$\begin{aligned}\operatorname{vec}\mathbf{B} &:= E\{\operatorname{vec}\hat{\boldsymbol{\Xi}}\} - \operatorname{vec}\boldsymbol{\Xi} =\\ &= [\mathbf{I}_p\otimes(\mathbf{A}'\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\mathbf{A}']\operatorname{vec}\mathbf{A}\boldsymbol{\Xi} - \operatorname{vec}\boldsymbol{\Xi}\\ &= [\mathbf{I}_p\otimes(\mathbf{A}'\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\mathbf{A}'](\mathbf{I}_p\otimes\mathbf{A})\operatorname{vec}\boldsymbol{\Xi} - \operatorname{vec}\boldsymbol{\Xi}\\ &= -(\mathbf{I}_{mp} - \mathbf{I}_p\otimes(\mathbf{A}'\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\mathbf{A}'\mathbf{A})\operatorname{vec}\boldsymbol{\Xi}\\ &= -[\mathbf{I}_p\otimes(\mathbf{A}'\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\alpha\mathbf{S}^{-1}]\operatorname{vec}\boldsymbol{\Xi}\\ &= -[\mathbf{I}_p\otimes\alpha(\mathbf{S}\mathbf{A}'\mathbf{A}+\mathbf{I}_m)^{-1}]\operatorname{vec}\boldsymbol{\Xi}\end{aligned}$$

or its bias matrix:

$$\mathbf{B} := E\{\hat{\boldsymbol{\Xi}}\} - \boldsymbol{\Xi} = -[\mathbf{I}_m - (\mathbf{A}'\mathbf{A} + \alpha\mathbf{S}^{-1})^{-1}\mathbf{A}'\mathbf{A}]\boldsymbol{\Xi} \tag{3.75}$$

and the matrix of the *Mean Square Estimation Errors* :

$$\begin{aligned} MSE\{\mathrm{vec}\,\hat{\boldsymbol{\Xi}}\} &:= E\{(\mathrm{vec}\,\hat{\boldsymbol{\Xi}} - \mathrm{vec}\,\boldsymbol{\Xi})(\mathrm{vec}\,\hat{\boldsymbol{\Xi}} - \mathrm{vec}\,\boldsymbol{\Xi})'\} \\ &= D\{\mathrm{vec}\,\hat{\boldsymbol{\Xi}}\} + \mathrm{vec}\,\mathbf{B}(\mathrm{vec}\,\mathbf{B})' \end{aligned} \tag{3.76}$$

$$\begin{aligned} MSE\{\mathrm{vec}\,\hat{\boldsymbol{\Xi}}\} = \boldsymbol{\Sigma} \otimes (\mathbf{A}'\mathbf{A} + \alpha\mathbf{S}^{-1})^{-1}\mathbf{A}'\mathbf{A}(\mathbf{A}'\mathbf{A} + \alpha\mathbf{S}^{-1})^{-1} + \\ +[\mathbf{I}_{mp} - \mathbf{I}_p \otimes (\mathbf{A}'\mathbf{A} + \alpha\mathbf{S}^{-1})^{-1}\mathbf{A}'\mathbf{A}]\,\mathrm{vec}\,\boldsymbol{\Xi}(\mathrm{vec}\,\boldsymbol{\Xi})'[\mathbf{I}_{mp} - \mathbf{I}_p \otimes (\mathbf{A}'\mathbf{A} + \alpha\mathbf{S}^{-1})^{-1}\mathbf{A}'\mathbf{A})]' \end{aligned} \tag{3.77}$$

With the results of the multivariate *Gauss-Markov* model and the multivariate $\alpha-\mathrm{BLE}$ derived above, we shall now apply them in our special case: Direct observations of a two-dimensional, symmetric rank-two random tensor,

$$\boldsymbol{\mu} := E\{\mathbf{y}\} = \mathrm{vech}\,\mathbf{t} = [t_{11} \;\; t_{12} \;\; t_{22}]', \;\; \mathbf{y} \in \mathbb{R}^{3\times 1}.$$

This is a random vector which is normally distributed according to *Definition* 1.7. We denote it as $\mathbf{y} \sim \mathcal{N}_3(\boldsymbol{\mu}, \boldsymbol{\Sigma}_{\mathbf{y}})$.

Suppose our sampled n independent observation vectors $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n$ are all distributed according to $\mathcal{N}_3(\boldsymbol{\mu}, \boldsymbol{\Sigma}_{\mathbf{y}})$. We have the vectorized form matrix as

$$\mathbf{Y} = \begin{bmatrix} \mathbf{y}_1' \\ \mathbf{y}_2' \\ \vdots \\ \mathbf{y}_n' \end{bmatrix}, \quad \mathbf{Y} \in \mathbb{R}^{n\times 3},$$

which is just the case in (3.60) with $p = 3$, $m = 1$ and

$$\mathbf{A} = \mathbf{1} = [1, 1, \cdots, 1]', \; \mathbf{1}'\mathbf{1} = n, \; \boldsymbol{\Xi} = [\xi_1, \xi_2, \xi_3], \; \boldsymbol{\Sigma}_{\mathbf{y}} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 \end{bmatrix}, \; \mathbf{S} := \mathbf{I}_1.$$

From (3.65) we have the BLUUE of $\boldsymbol{\Xi}$:

$$\hat{\boldsymbol{\Xi}} = \hat{\boldsymbol{\xi}}' = [\hat{\xi}_1, \hat{\xi}_2, \hat{\xi}_3] = (\mathbf{1}'\mathbf{1})^{-1}\mathbf{1}'\mathbf{Y} = \frac{1}{n}\mathbf{1}'\mathbf{Y} \tag{3.78}$$

and in transposed form

$$\hat{\boldsymbol{\Xi}}' = \hat{\boldsymbol{\xi}} = \begin{bmatrix} \hat{\xi}_1 \\ \hat{\xi}_2 \\ \hat{\xi}_3 \end{bmatrix} = \frac{1}{n}\mathbf{Y}'\mathbf{1} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{y}_i$$

which represents the same estimate as that of (1.37) in S*ection 1.3*.

The related dispersion matrix

$$D(\mathrm{vec}\,\hat{\boldsymbol{\Xi}}) = \boldsymbol{\Sigma}_{\mathbf{y}} \otimes (\mathbf{A}'\mathbf{A})^{-1} = \boldsymbol{\Sigma}_{\mathbf{y}} \otimes \frac{1}{n} = \frac{1}{n}\boldsymbol{\Sigma}_{\mathbf{y}} = \frac{1}{n}\begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 \end{bmatrix}. \tag{3.79}$$

With (3.72) and (3.73) we obtain the $\alpha-\mathrm{BLE}$

$$\mathrm{vec}\,\hat{\boldsymbol{\Xi}} = \hat{\boldsymbol{\xi}} = \begin{bmatrix} \hat{\xi}_1 \\ \hat{\xi}_2 \\ \hat{\xi}_3 \end{bmatrix} = [\mathbf{I}_3 \otimes (\mathbf{1}'\mathbf{1} + \alpha\mathbf{I}_1^{-1})^{-1}\mathbf{1}']\,\mathrm{vec}\,\mathbf{Y} = [\mathbf{I}_3 \otimes (n+\alpha)^{-1}\mathbf{1}']\,\mathrm{vec}\,\mathbf{Y} = \mathrm{vec}[(n+\alpha)^{-1}\mathbf{1}'\mathbf{Y}] \tag{3.80}$$

and

$$\hat{\boldsymbol{\Xi}} = \hat{\boldsymbol{\xi}}' = [\hat{\xi}_1, \hat{\xi}_2, \hat{\xi}_3] = (\mathbf{1}'\mathbf{1} + \alpha\mathbf{I}_1^{-1})^{-1}\mathbf{1}'[\mathbf{Y}_1, \mathbf{Y}_2, \mathbf{Y}_3] = (n+\alpha)^{-1}\mathbf{1}'[\mathbf{Y}_1, \mathbf{Y}_2, \mathbf{Y}_3] \tag{3.81}$$

subject to the dispersion matrix with (3.74)

$$D(\operatorname{vec}\hat{\boldsymbol{\Xi}}) = \boldsymbol{\Sigma}_{\mathbf{y}} \otimes (n+\alpha)^{-1} n (n+\alpha)^{-1} = \boldsymbol{\Sigma}_{\mathbf{y}} \otimes \frac{n}{(n+\alpha)^2} = = \frac{n}{(n+\alpha)^2}\boldsymbol{\Sigma}_{\mathbf{y}} = \frac{n}{(n+\alpha)^2}\begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 \end{bmatrix}. \tag{3.82}$$

In comparison with the results of the univariate case (3.51) and (3.52) we can see that (3.81) is just the generalized from (3.51) and the dispersion matrix (3.82) is of importance since the variance and covariance components of the three elements of the random tensor are rescaled together. With (3.71) we further have the sample variance-covariance matrix of type BIQUUE

$$\hat{\boldsymbol{\Sigma}}_{\mathbf{y}} = \frac{1}{n-1}\mathbf{Y}'(\mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}')\mathbf{Y} = \frac{1}{n}\mathbf{Y}'\begin{bmatrix} 1 & (n-1)^{-1} & \cdots & (n-1)^{-1} \\ (n-1)^{-1} & 1 & \cdots & (n-1)^{-1} \\ \vdots & \vdots & \ddots & \vdots \\ (n-1)^{-1} & (n-1)^{-1} & \cdots & 1 \end{bmatrix}\mathbf{Y}. \tag{3.83}$$

Now we shall be able to estimate the *Mean Square Estimation Error of the* multivariate $\alpha-$BLE of random tensor with (3.77)

$$\begin{aligned} MSE\{\operatorname{vec}\hat{\boldsymbol{\Xi}}\} &= D\{\operatorname{vec}\hat{\boldsymbol{\Xi}}\} + \operatorname{vec}\mathbf{B}(\operatorname{vec}\mathbf{B})' = \\ &= D\{\operatorname{vec}\hat{\boldsymbol{\Xi}}\} + [\mathbf{I}_3 - \mathbf{I}_3 \otimes (n+\alpha)^{-1} n]\operatorname{vec}\boldsymbol{\Xi}(\operatorname{vec}\boldsymbol{\Xi})'[\mathbf{I}_3 - \mathbf{I}_3 \otimes (n+\alpha)^{-1} n)]' \\ &= \frac{n}{(n+\alpha)^2}\boldsymbol{\Sigma}_{\mathbf{y}} + \frac{\alpha}{(n+\alpha)}\mathbf{I}_3 \operatorname{vec}\boldsymbol{\Xi}(\operatorname{vec}\boldsymbol{\Xi})'\mathbf{I}_3 \frac{\alpha}{(n+\alpha)} \\ &= \frac{n}{(n+\alpha)^2}\boldsymbol{\Sigma}_{\mathbf{y}} + \frac{\alpha^2}{(n+\alpha)^2}\operatorname{vec}\boldsymbol{\Xi}(\operatorname{vec}\boldsymbol{\Xi})' \\ &= \frac{n}{(n+\alpha)^2}\begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 \end{bmatrix} + \frac{\alpha^2}{(n+\alpha)^2}\begin{bmatrix} \xi_1^2 & \xi_1\xi_2 & \xi_1\xi_3 \\ \xi_2\xi_1 & \xi_2^2 & \xi_2\xi_3 \\ \xi_3\xi_1 & \xi_3\xi_2 & \xi_3^2 \end{bmatrix}. \end{aligned} \tag{3.84}$$

The trace of it is

$$\begin{aligned} \operatorname{tr} MSE\{\hat{\boldsymbol{\xi}}\} &= \operatorname{tr} MSE\{\operatorname{vec}\hat{\boldsymbol{\Xi}}\} = \frac{n}{(n+\alpha)^2}\operatorname{tr}\boldsymbol{\Sigma}_{\mathbf{y}} + \frac{\alpha^2}{(n+\alpha)^2}\operatorname{tr}\boldsymbol{\xi}\boldsymbol{\xi}' = \\ &= \frac{n}{(n+\alpha)^2}(\sigma_1^2 + \sigma_2^2 + \sigma_3^2) + \frac{\alpha^2}{(n+\alpha)^2}(\xi_1^2 + \xi_2^2 + \xi_3^2) = \gamma_1(\alpha) + \gamma_2(\alpha). \end{aligned} \tag{3.85}$$

With the same condition that led to (3.57), we achieve the optimal weight factor $\hat{\alpha}$ for the multivariate $\alpha-$BLE by minimizing the $MSE\{\operatorname{vec}\hat{\boldsymbol{\Xi}}\}$ of type (3.77), yielding

$$\hat{\alpha} = \frac{\operatorname{tr}\boldsymbol{\Sigma}_{\mathbf{y}}}{\operatorname{tr}\boldsymbol{\xi}\boldsymbol{\xi}'} = \frac{\sigma_1^2 + \sigma_2^2 + \sigma_3^2}{\xi_1^2 + \xi_2^2 + \xi_3^2}. \tag{3.86}$$

Since $\boldsymbol{\xi}$ and $\boldsymbol{\Sigma}_{\mathbf{y}}$ are not themselves available, we could achieve the first approximation to α with the BLUUE $\bar{\xi}_i$ and BIQUUE $\bar{\sigma}_i^2$, namely

$$\tilde{\alpha} = \frac{\operatorname{tr}\bar{\boldsymbol{\Sigma}}_{\mathbf{y}}}{\operatorname{tr}\bar{\boldsymbol{\xi}}\bar{\boldsymbol{\xi}}'} = \frac{\bar{\sigma}_1^2 + \bar{\sigma}_2^2 + \bar{\sigma}_3^2}{\bar{\xi}_1^2 + \bar{\xi}_2^2 + \bar{\xi}_3^2}. \tag{3.87}$$

Alternatively, we can use an iterative procedure where $\hat{\hat{\alpha}}(k+1) = \operatorname{tr}\hat{\boldsymbol{\Sigma}}_{\mathbf{y}_k} / \operatorname{tr}\hat{\boldsymbol{\xi}}(k)\hat{\boldsymbol{\xi}}'(k)$ with $\hat{\boldsymbol{\xi}}(k) = \hat{\boldsymbol{\xi}}(\hat{\hat{\alpha}}(k))$ as the $\alpha-$BLE from (3.81) and $\hat{\boldsymbol{\Sigma}}_{\mathbf{y}_k}$ in modified form

$$\hat{\boldsymbol{\Sigma}}_{\mathbf{y}_k} = \frac{1}{n-1}\mathbf{Y}'(\mathbf{I}_n - \frac{1}{n+\alpha(k)}\mathbf{1}\mathbf{1}')(\mathbf{I}_n - \frac{1}{n+\alpha(k)}\mathbf{1}\mathbf{1}')\mathbf{Y} \tag{3.88}$$

The iteration can be continued until there is stability achieved in $\operatorname{tr} MSE\{\hat{\boldsymbol{\xi}}(k+1)\}$ from (3.86).

3.3 Case study: 2-dimensional strain rate tensor

With these models developed above we are able to successfully perform the α-homBLE estimation to determine the weighting factor α for the univariate and multivariate case. In lieu of a case study, the model is applied to simulated observations of a random tensor of type strain rate based on the real estimate of one station in the Finnish Primary Geodetic Network (Kakkuri and Chen 1992). We will apply the three observation sets firstly to the univariate case under the assumption that the three sets are independent, identically distributed (i.i.d) observations. Secondly, the three set observations will be applied together in the multivariate case and therefore the correlations among them are considered.

Box 3.5 provides the real estimated random strain rate tensor with the related standard deviations and the simulated observations in 11 epochs. We use the notation of $\mathbf{y}_{(1)}, \mathbf{y}_{(2)}, \mathbf{y}_{(3)}$ for the three i.i.d. univariate observations, which are related to the multivariate notation introduced in section 3-2, i.e.

$$\begin{bmatrix} \mathbf{y}_{(1)} = [y_{11}, y_{21}, \cdots, y_{n1}]' \\ \mathbf{y}_{(2)} = [y_{12}, y_{22}, \cdots, y_{n2}]', \\ \mathbf{y}_{(3)} = [y_{13}, y_{23}, \cdots, y_{n3}]' \end{bmatrix} \quad \mathbf{Y} = \begin{bmatrix} y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} \\ \vdots & \vdots & \vdots \\ y_{n1} & y_{n2} & y_{n3} \end{bmatrix} = \begin{bmatrix} \mathbf{y}_{(1)} & \mathbf{y}_{(2)} & \mathbf{y}_{(3)} \end{bmatrix}, \text{ where } n = 11.$$

***Box 3.5*:**

Observations of a random tensor of type strain rate
(epoch 0: Kakkuri and Chen (1992))

$$\mathbf{t}_0 = \begin{bmatrix} t_{11\cdot 0} & t_{12\cdot 0} \\ t_{12.0} & t_{22.0} \end{bmatrix} = \begin{bmatrix} 0.236 & -0.049 \\ -0.049 & 0.148 \end{bmatrix} \ (\mu \text{strain/year})$$

$$\bar{\sigma}_{11.0} = 0.094, \ \bar{\sigma}_{12.0} = 0.054 \ \text{ and } \ \bar{\sigma}_{22.0} = 0.065$$

"*the vectorized form* "

$$\mathbf{y}'_0 = [y_1 \ \ y_2 \ \ y_3]_0 = [t_{11.0} \ \ t_{12.0} \ \ t_{22.0}] \text{ with}$$
$$\bar{\sigma}_{1.0} = 0.094, \ \bar{\sigma}_{2.0} = 0.054, \ \bar{\sigma}_{3.0} = 0.065$$

"Observations of the distinct elements $\mathbf{y}_{(1)}, \mathbf{y}_{(2)}, \mathbf{y}_{(3)}$ in 11 epochs"

epoch i	$\mathbf{y}_{(1)}$ (μstrain/yr)	$\mathbf{y}_{(2)}$ (μstrain/yr)	$\mathbf{y}_{(3)}$ (μstrain/yr)
1	0.1513	-0.0305	0.1615
2	0.2913	-0.0081	0.1624
3	0.2881	-0.0864	0.0826
4	0.1970	-0.0123	0.1186
5	0.2418	-0.1069	0.2390
6	0.2790	-0.0004	0.1180
7	0.2547	-0.1636	0.1501
8	0.2602	-0.0336	0.1999
9	0.4316	-0.0886	0.2063
10	0.0219	-0.0908	0.1570
11	0.2679	-0.0408	0.0428

In the univariate case we assume that the three observation sets $\mathbf{y}_{(1)}, \mathbf{y}_{(2)}, \mathbf{y}_{(3)}$ in Box 3.5 of the distinct elements of strain rate tensors are independent and the sample mean (BLUUE) $\bar{\xi}_1$, $\bar{\xi}_2$ and $\bar{\xi}_3$, and the sample variances (BIQUUE) $\bar{\sigma}^2_{y_1}$, $\bar{\sigma}^2_{y_2}$, $\bar{\sigma}^2_{y_3}$ are estimated by (3.3) and (3.55), respectively:

$$\bar{\xi}_1 = \bar{t}_{11} = \ 0.2441 \ (\mu \text{strain/y}) \ \text{ and } \ \bar{\sigma}^2_{y_1} = 0.010168 (\mu \text{strain/y})^2;$$

$$\bar{\xi}_2 = \bar{t}_{12} = -0.0602 \ (\mu \text{strain/y}) \ \text{ and } \ \bar{\sigma}^2_{y_2} = 0.002585 \ (\mu \text{strain/y})^2;$$

$$\bar{\xi}_3 = \bar{t}_{22} = \ 0.1489 \ (\mu \text{strain/y}) \ \text{ and } \ \bar{\sigma}^2_{y_3} = 0.003195 \ (\mu \text{strain/y})^2.$$

In the multivariate case the BLUUE of t_{11}, t_{12}, t_{22} from (3.78) is identical, namely

$$\begin{bmatrix} \bar{\xi}_1 \\ \bar{\xi}_2 \\ \bar{\xi}_3 \end{bmatrix} = \begin{bmatrix} \bar{t}_{11} \\ \bar{t}_{12} \\ \bar{t}_{22} \end{bmatrix} = \begin{bmatrix} 0.2441\ (\mu\text{strain/y}) \\ -0.0602\ (\mu\text{strain/y}) \\ 0.1489\ (\mu\text{strain/y}) \end{bmatrix},$$

and the full sample variance-covariance matrix $\hat{\mathbf{\Sigma}}_{\mathbf{y}}$ of type BIQUUE from (3.83) is

$$\hat{\mathbf{\Sigma}}_{\mathbf{y}} = \begin{bmatrix} 0.010168 & -0.000024 & 0.000396 \\ -0.000024 & 0.002585 & -0.000740 \\ 0.000396 & -0.000740 & 0.003195 \end{bmatrix} (\mu\text{strain/y})^2$$

while the related dispersion matrix of $\hat{\boldsymbol{\xi}}$ with (3.79) is

$$D\{\hat{\boldsymbol{\xi}}\} = \mathbf{\Sigma}_{\hat{\xi}} = \begin{bmatrix} 0.000924 & -0.000002 & 0.000036 \\ -0.000002 & 0.000235 & -0.000067 \\ 0.000036 & -0.000067 & 0.000290 \end{bmatrix} (\mu\text{strain/y})^2$$

Now we shall be able to analyze the α-BLE estimate and the determination of the weight factor α in both the univariate and multivariate case as explained in *Section* 3.3.1 and 3.3.2.

3.3.1 The univariate α-BLE and the determination of the weight factor α by *A-optimal design*

By means of *Figure 3*.2 we compare α-homBLE (*dashed line*) and BLUUE estimates (*full line*), in particular we document the dependence on the uniform regularization parameter α. Based upon the results summarized in *Box 3.3* we have computed *at first* the trace of the *Mean Square Error* (*MSE*) *matrix*, in particular the functions $\gamma_1(\alpha)$ as variance term and $\gamma_2(\alpha)$ as bias term squared. Within $\{\gamma_1(\alpha), \gamma_2(\alpha)\}$ we have substituted $\{\sigma, \xi\}$ by $\{\bar{\sigma}(\text{BIQUUE}), \bar{\xi}\ (\text{BLUUE})\}$, i.e. the true value σ and ξ with their estimates $\bar{\sigma}$ and $\bar{\xi}$ and plotted them in *Figure* 3.3. The interrelation between the variances, squared biases and the weighting factor is evident. The variance term $\gamma_1(\alpha)$ (*dashed line*) decreases as α increases while the squared bias term $\gamma_2(\alpha)$ (*dotted line*) increases with α. The *dash-dotted line* which represents $\gamma_1(\tilde{\alpha}) + \gamma_2(\tilde{\alpha})$ as $\text{trMSE}\{\hat{\boldsymbol{\xi}} : \tilde{\alpha}\text{-homBLE}\}$ is under the level of $\text{tr}\,MSE\{\bar{\boldsymbol{\xi}} : \text{BLUUE}\}$ as expected. In summary, these estimates determine approximate values of the weighting factor α of type A-optimum with respect to $MSE\{\hat{\boldsymbol{\xi}}\}$ as in (3.56) of *Corollary 3.11*, in particular

$$\tilde{\alpha}_1(1) = 0.171,\ \tilde{\alpha}_2(1) = 0.714 \ \text{ and }\ \tilde{\alpha}_3(1) = 0.144 .$$

Figure 3.4 is a "zoom-in" version of *Figure 3.3*, which illustrates the optimal values $\tilde{\alpha}$ consistent with these curves, respectively $\text{trMSE}\{\hat{\boldsymbol{\xi}}\}$ at minimal points.

Secondly, by an iterative procedure we have updated every element of the first approximate $\{\tilde{\alpha}_1, \tilde{\alpha}_2, \tilde{\alpha}_3\}$ by means of

$$\tilde{\alpha}_j(k+1) = \frac{\hat{\sigma}_j^2(k)}{\hat{\xi}_j^2(k)}, \quad j = 1, 2, 3 \qquad (3.89)$$

$$\text{where } \hat{\xi}_j(k) = \hat{\xi}_j(\tilde{\alpha}(k)),\ \ \hat{\sigma}_j(k) = \hat{\sigma}_j(\hat{\xi}_j(\tilde{\alpha}_j(k)))$$

and

$$\begin{aligned} \hat{\sigma}_j^2(k) &= \frac{1}{n-1}\mathbf{y}'_{(j)}(\mathbf{I}_n - \frac{1}{n+\tilde{\alpha}_j(k)}\mathbf{J}_n)^2\,\mathbf{y}_{(j)} \\ &= \frac{1}{n-1}\sum_{i=1}^{n}(y_{ji} - \frac{n}{n+\tilde{\alpha}_j(k)}\bar{\xi}_j)^2 . \end{aligned} \qquad (3.90)$$

The sequential optimization ends at the *reproducing point* $\tilde{\alpha}_j(k+1) \doteq \tilde{\alpha}_j(k)$ in computer arithmetic where $\text{tr}\,\text{MSE}\{\hat{\xi}_k\}$ reaches its minimum. Such an iterative procedure as illustrated in *Figure 3.5* supports the optimization procedure to generate $\hat{\alpha} = \arg\{\text{tr}\,MSE\{\hat{\boldsymbol{\xi}}\} = \min\}$.

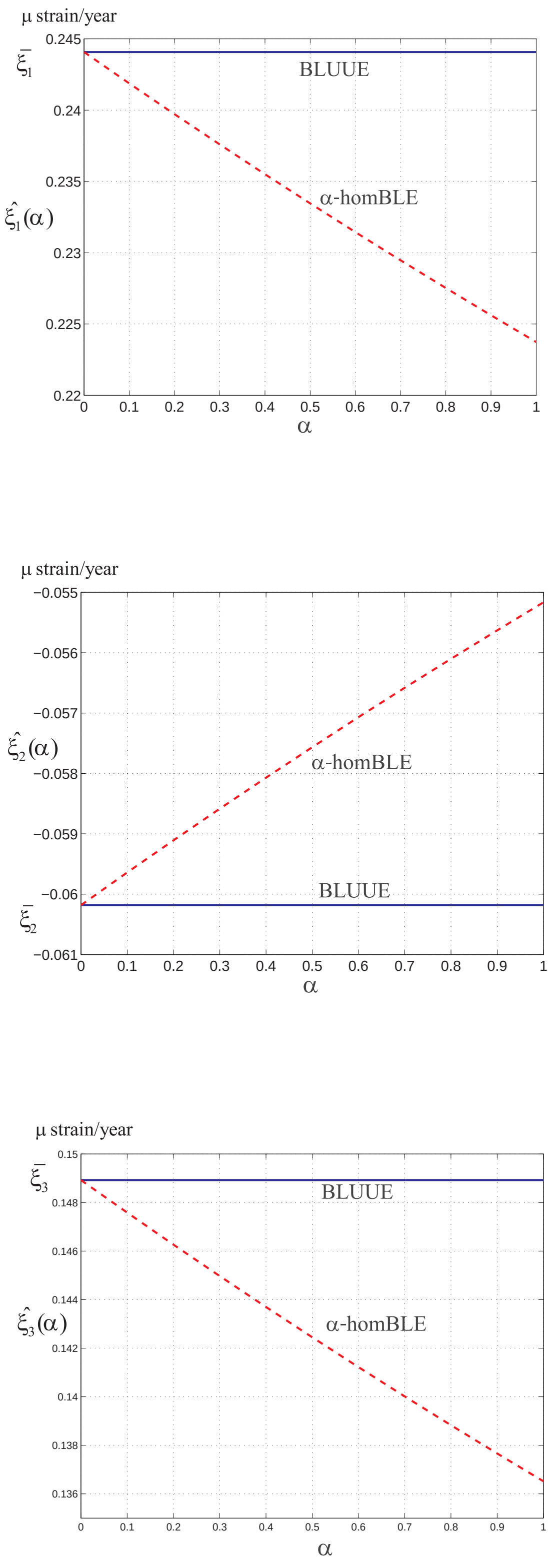


Figure 3.2 $\bar{\xi}_j$(BLUUE) versus $\hat{\xi}_j$(α-homBLE), (j=1, 2, 3) of the direct observations { t_{11}, t_{12}, t_{22} }, symmetric random tensor of type strain rate, as functions of the balancing parameter α.

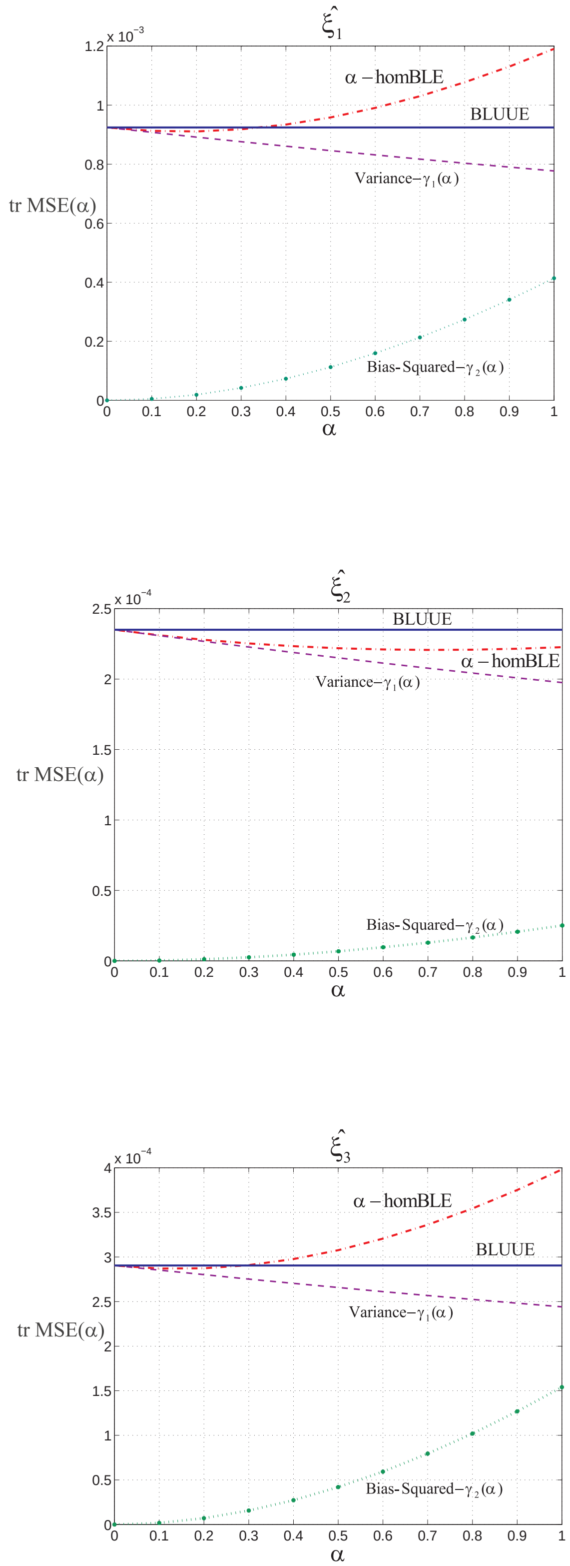


Figure 3.3 The trace of the *Mean Square Error* (*MSE*) functions for the BLUUE and the α-homBLE estimates of the unknowns (ξ_1, ξ_2, ξ_3) as functions of α. Depicted are $\operatorname{tr} MSE\{\bar{\xi}\} = \hat{\sigma}^2_{\bar{\xi}}$ for the BLUUE and $\operatorname{tr} MSE\{\hat{\xi}\} = \gamma_1(\alpha) + \gamma_2(\alpha)$ for the α-homBLE, as well as the separate parts for variance $\gamma_1(\alpha)$ and bias squared $\gamma_2(\alpha)$.

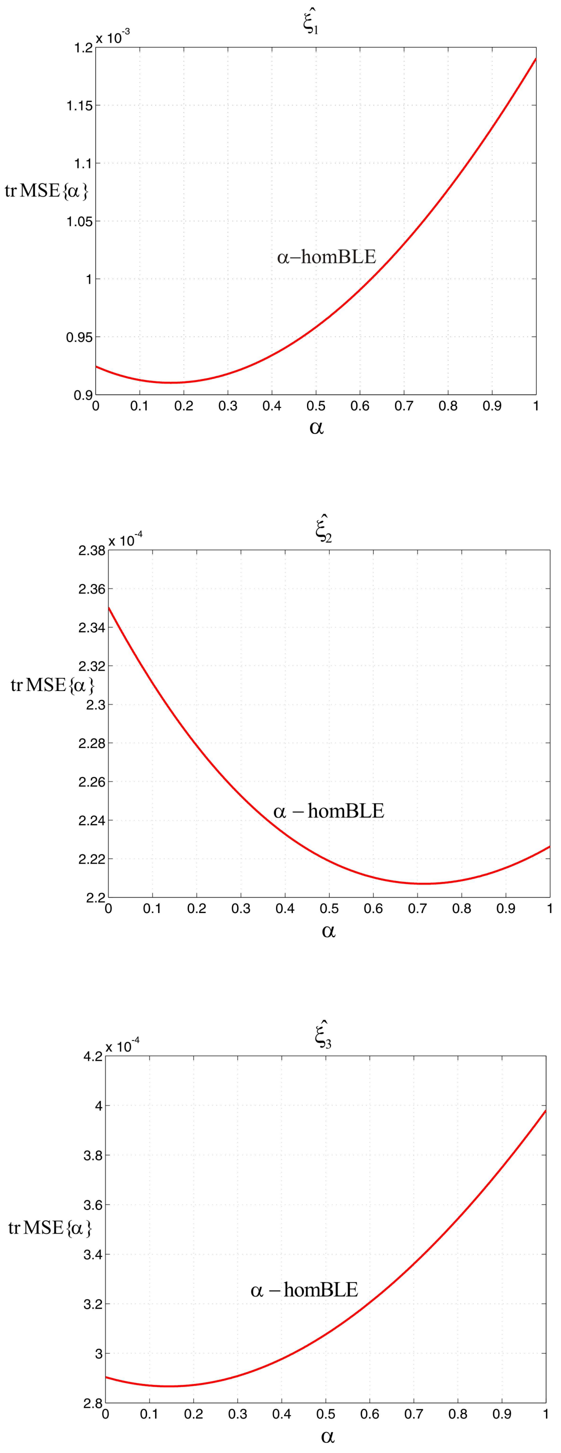


Figure 3.4 The trace of the *Mean Square Error* (*MSE*) functions $\text{tr}\, MSE\{\hat{\xi}\}$ for the α-homBLE estimates of (ξ_1, ξ_2, ξ_3), also depicted in Figure 3.3. Here the function graph is "zoomed-in", in order to emphasize the behavior of the function around its minimal point.

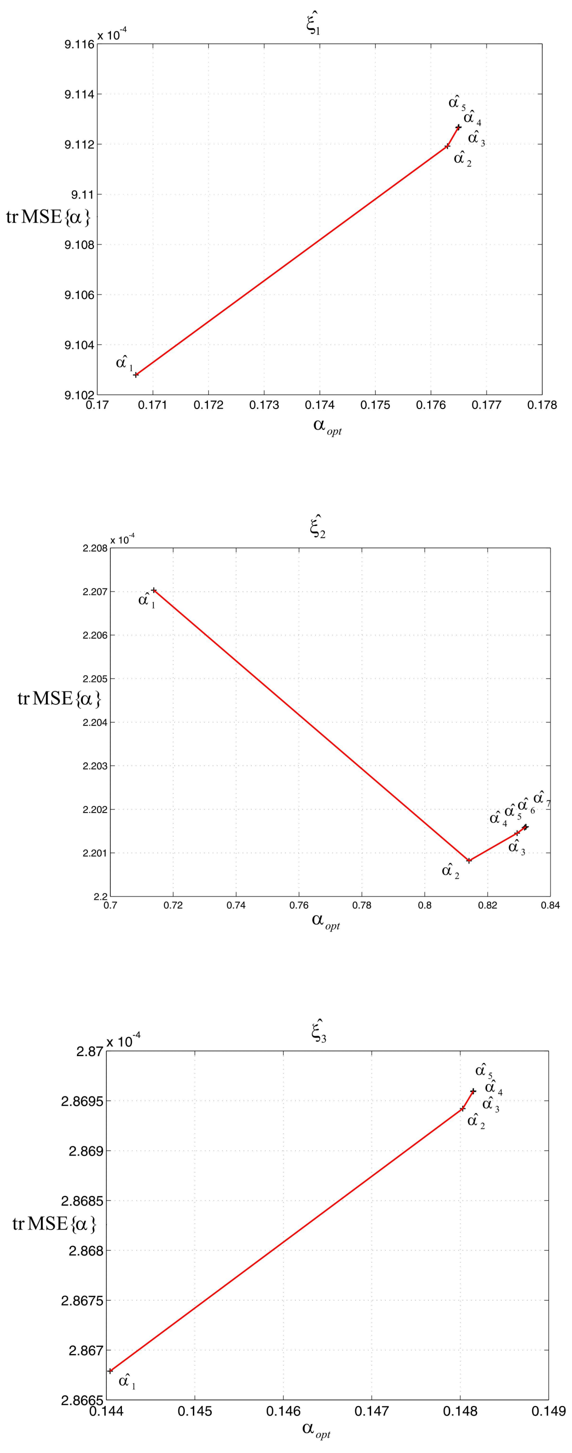


Figure 3.5 Iteration steps for A-optimal α, three sets of direct observations

(i) upper graph: the trace of $MSE\{\alpha\}$ for the estimate $\hat{\xi}_1(\alpha)$ of type α-homBLE of the first set of direct observations;

(ii) middle graph: the trace of $MSE\{\alpha\}$ for the estimate $\hat{\xi}_2(\alpha)$ of type α-homBLE of the second set of direct observations;

(iii) lower graph: the trace of $MSE\{\alpha\}$ for the estimate $\hat{\xi}_3(\alpha)$ of type α-homBLE of the third set of direct observations.

3.3.2 The multivariate α-BLE and the determination of the weight factor α by *A-optimal design*

We apply the multivariate $\alpha-$BLE from (3.87) to calculate the trace of *Mean Square Estimation Error* (*MSE*) and their two divided terms: variance term $\gamma_1(\alpha)$ and bias-squared term $\gamma_2(\alpha)$ after substituting $\boldsymbol{\xi}$ and $\boldsymbol{\Sigma}_\mathbf{y}$ with their estimates $\hat{\boldsymbol{\xi}}(k)=\hat{\boldsymbol{\xi}}(\alpha(k))$ of (3.81) and $\hat{\boldsymbol{\Sigma}}_{\mathbf{y}_k}=\hat{\boldsymbol{\Sigma}}_\mathbf{y}(\alpha(k))$ of (3.88) and plotted them in *Figure 3.6*, which shows in qualitative form the relationship between the variances and the squared bias, and the weight factor α. The variance term $\gamma_1(\alpha)$ decreases as α increases, while the bias-squared term $\gamma_2(\alpha)$ increases with α; both are plotted by the blue and green curves, respectively. As is indicated by the red curve, the trace of *MSE* of $\alpha-$BLE which is the sum of $\gamma_1(\alpha)$ and $\gamma_2(\alpha)$, there exist several values of α for which the trace of *MSE* of $\alpha-$BLE is less than *MSE* of BLUUE .

With these estimates we are interested in the approximate optimal weight factor $\tilde{\alpha}$ from (3.87) for the multivariate model by the *A-optimal design* that minimizes of the trace of $MSE\{\hat{\boldsymbol{\xi}}\}$, which is

$$\tilde{\alpha}=0.187$$

and generates the $\alpha-$BLE of t_{11}, t_{12}, t_{22} as

$$\hat{\boldsymbol{\xi}}(\tilde{\alpha})=\begin{bmatrix}\hat{\xi}_1\\ \hat{\xi}_2\\ \hat{\xi}_3\end{bmatrix}_{\tilde{\alpha}}=\begin{bmatrix}\hat{t}_{11}\\ \hat{t}_{12}\\ \hat{t}_{22}\end{bmatrix}_{\tilde{\alpha}}=\begin{bmatrix}0.2400\ (\mu\,\text{strain/y})\\ -0.0592\ (\mu\,\text{strain/y})\\ 0.1464\ (\mu\,\text{strain/y})\end{bmatrix}$$

As is shown in *Figure* 3.7 this optimal value for the weight factor α is consistent with the curve of the trace of $MSE\{\hat{\boldsymbol{\xi}}\}$, where it reaches its minimum value.

With the iterative procedure introduced above we use the above listed optimal weight factor $\tilde{\alpha}$ as the initial value, i.e. $\hat{\alpha}(1)$ and iterates with $\hat{\alpha}(k+1)=\operatorname{tr}\hat{\boldsymbol{\Sigma}}_{\mathbf{y}_k}/\operatorname{tr}\hat{\boldsymbol{\xi}}(k)\hat{\boldsymbol{\xi}}'(k)$. The iteration can be continued until there is stability achieved in $\hat{\alpha}(k+1)$ or the minimum of $\operatorname{tr}MSE\{\hat{\boldsymbol{\xi}}(k+1)\}$. These iteration results are shown in *Figure* 3.8, from which we can see that the optimal weight for α is at the first iteration for the direct observation set. This supports that our optimal estimate of the weight factor α from (3.87) within the *multivariate model* by the *A-optimal design*, that minimizes the trace of $MSE\{\hat{\boldsymbol{\xi}}\}$ is also reasonable and meaningful in practical data analysis.

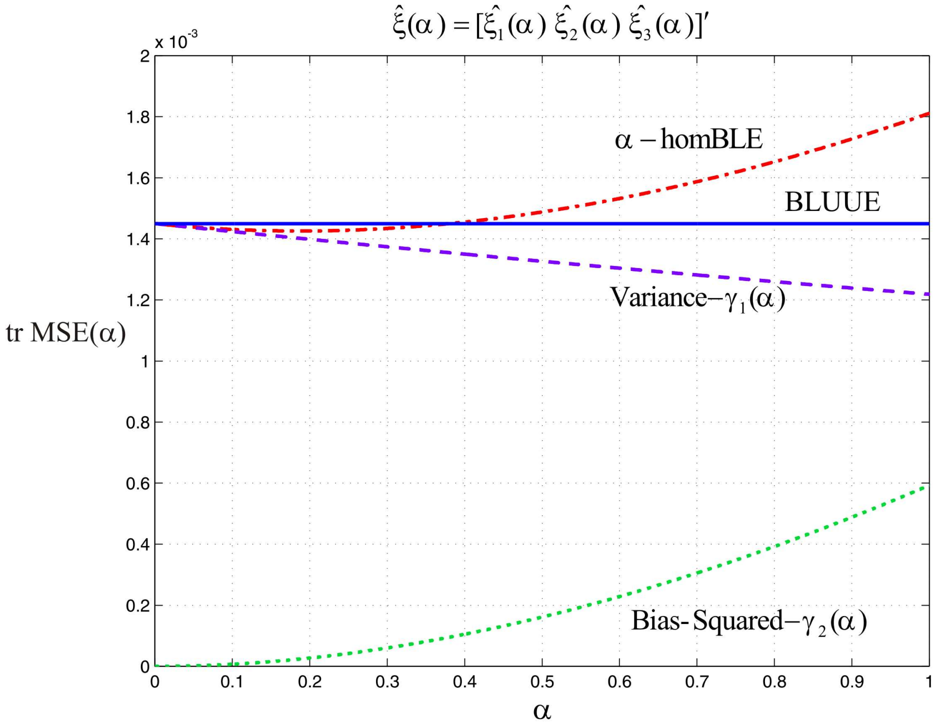


Figure 3.6 The trace of the *MSE* functions for multivariate α-BLE $\hat{\boldsymbol{\xi}}(\alpha)=[\hat{\xi}_1,\hat{\xi}_2,\hat{\xi}_3]'$, their variance term $\gamma_1(\alpha)$ and bias-squared term $\gamma_2(\alpha)$ with the choice of α.

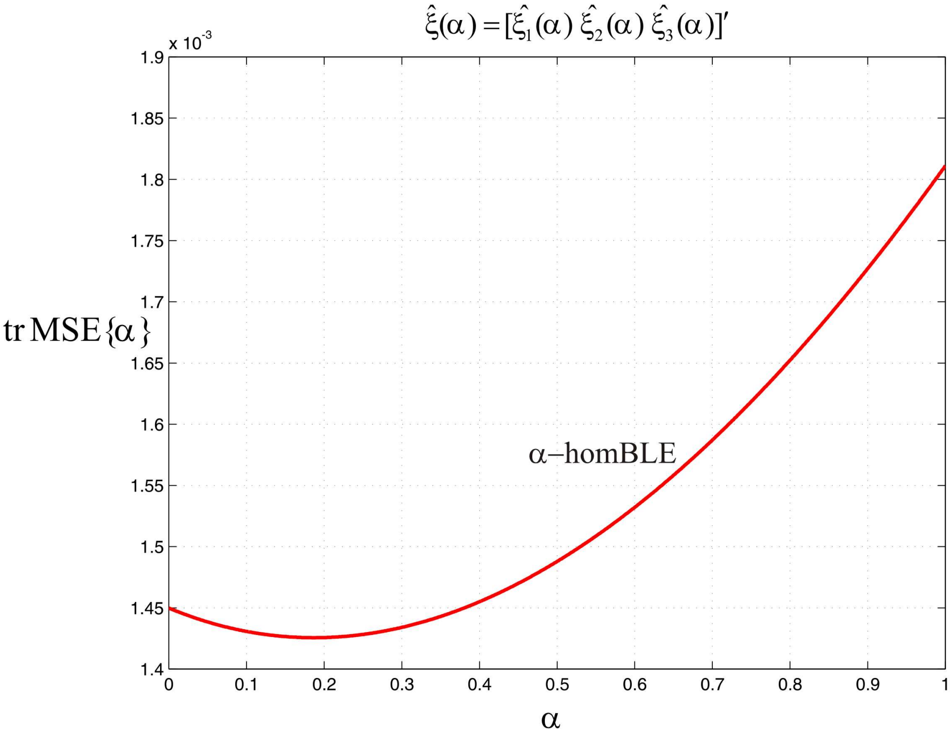


Figure 3.7 The trace of the $MSE\{\hat{\xi}\}$ of the multivariate α-BLE $\hat{\xi}(\alpha) = [\hat{\xi}_1, \hat{\xi}_2, \hat{\xi}_3]'$ of the elements t_{11}, t_{12}, t_{22} of with the choice of α. The optimal value $\hat{\alpha} = 0.187$ of weight factor α is consistent with the minimum value of the curve for the trace of $MSE\{\hat{\xi}\}$.

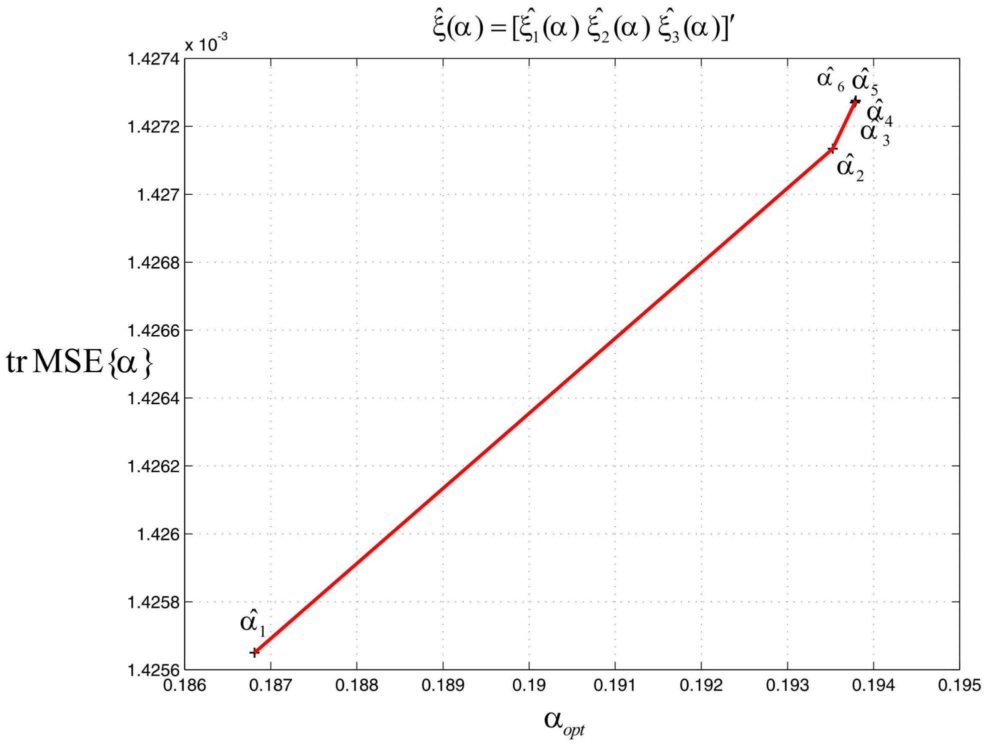


Figure 3.8 The iteration of optimal estimates of the weight factor α for the direct observation set, where the index *i* of $\hat{\alpha}_i$ represents the iterative steps.

Appendix 3-A: Proof of *Theorem 3.9*

Before we prove *Theorem 3.9* let us introduce auxiliary results which are used subsequently.

Lemma 3-A1 (Cayley matrix inverse differentiation):

$$d(\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1} = \\ = -(\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\mathbf{S}^{-1}(\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}d\alpha \quad \text{(A.1)}$$

Proof:

$$\mathbf{M}\mathbf{M}^{-1}=\mathbf{I} \Rightarrow (d\mathbf{M})\mathbf{M}^{-1}+\mathbf{M}(d\mathbf{M}^{-1})=0 \Rightarrow$$

$$d\mathbf{M}^{-1} = -\mathbf{M}^{-1}d\mathbf{M}\,\mathbf{M}^{-1}$$

Example 3-A1: $\mathbf{M} := \mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1} \quad \Rightarrow$

$$\Rightarrow \quad d\mathbf{M} = \mathbf{S}^{-1}d\alpha$$

$$d\mathbf{M}^{-1} = -(\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\mathbf{S}^{-1}(\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}d\alpha$$

q.e.d.

Lemma 3-A2 (differentiation of a scalar function of a matrix, such as the trace):

$$\begin{aligned} &\operatorname{tr}(\mathbf{A}+\mathbf{B}) = \operatorname{tr}\mathbf{A}+\operatorname{tr}\mathbf{B} \\ &d(\operatorname{tr}(\mathbf{A}+\mathbf{B})) = \operatorname{tr}d\mathbf{A}+\operatorname{tr}d\mathbf{B} \\ &d(\operatorname{tr}(\mathbf{X}\mathbf{A}\mathbf{X}')) = \operatorname{tr}(\mathbf{A}+\mathbf{A}')\mathbf{X}'d\mathbf{X} \end{aligned} \quad \text{(A.2)}$$

Lemma 3-A3 (Cayley inverse: sum of two matrices)

$$\begin{aligned} &(\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A} = \\ &= [\mathbf{I}_m+\alpha(\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A})^{-1}\mathbf{S}^{-1}]^{-1} = \\ &= \mathbf{I}_m-\alpha(\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\mathbf{S}^{-1} = \\ &= \mathbf{I}_m-\alpha(\mathbf{S}\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{I}_m)^{-1} \end{aligned} \quad \text{(A.3)}$$

Lemma 3-A4 :

$$\begin{aligned} &\operatorname{tr}(\boldsymbol{\beta}\boldsymbol{\beta}') = \boldsymbol{\beta}'\boldsymbol{\beta} \\ &d(\boldsymbol{\beta}'\boldsymbol{\beta}) = (d\boldsymbol{\beta})'\boldsymbol{\beta}+\boldsymbol{\beta}'(d\boldsymbol{\beta}) = 2\boldsymbol{\beta}'(d\boldsymbol{\beta}) \end{aligned} \quad \text{(A.4)}$$

Example 3-A2:

$$\begin{aligned} \boldsymbol{\beta} &:= -[\mathbf{I}_m-(\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}]\boldsymbol{\xi} = \\ &= -\alpha(\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\mathbf{S}^{-1}\boldsymbol{\xi} = \\ &= -\alpha(\mathbf{S}\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{I}_m)^{-1}\boldsymbol{\xi} \end{aligned}$$

$$\begin{aligned} d\boldsymbol{\beta} &= -d\alpha(\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\mathbf{S}^{-1}\boldsymbol{\xi}-\alpha d[(\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}]\mathbf{S}^{-1}\boldsymbol{\xi} \\ &= -d\alpha(\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\mathbf{S}^{-1}\boldsymbol{\xi}+ \\ &\quad +\alpha(\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\mathbf{S}^{-1}(\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}d\alpha\mathbf{S}^{-1}\boldsymbol{\xi} = \\ &= -(\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}[\mathbf{I}_m-\alpha\mathbf{S}^{-1}(\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}]\mathbf{S}^{-1}\boldsymbol{\xi}\,d\alpha \\ &= -(\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}(\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\mathbf{S}^{-1}\boldsymbol{\xi}\,d\alpha \\ &= -(\mathbf{S}\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{I}_m)^{-1}\mathbf{S}\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}(\mathbf{S}\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{I}_m)^{-1}\boldsymbol{\xi}\,d\alpha \\ &= -(\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}\boldsymbol{\beta}\,d\alpha/\alpha \end{aligned}$$

$$\begin{aligned} d(\boldsymbol{\beta}'\boldsymbol{\beta}) &= 2\boldsymbol{\beta}'(d\boldsymbol{\beta}) = 2[-\alpha(\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\mathbf{S}^{-1}\boldsymbol{\xi}]'[-(\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}(\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\mathbf{S}^{-1}\boldsymbol{\xi}\,d\alpha] = \\ &= 2\alpha\boldsymbol{\xi}'\mathbf{S}^{-1}(\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-2}\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}(\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\mathbf{S}^{-1}\boldsymbol{\xi}\,d\alpha = \frac{2\boldsymbol{\beta}'(\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}\boldsymbol{\beta}}{\alpha}d\alpha = \\ &= 2\alpha\boldsymbol{\xi}'(\mathbf{S}\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{I}_m)^{-2}\mathbf{S}\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}(\mathbf{S}\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{I}_m)^{-1}\boldsymbol{\xi}\,d\alpha = 2[\frac{\boldsymbol{\beta}'\boldsymbol{\beta}}{\alpha}-\boldsymbol{\beta}'(\mathbf{S}\mathbf{A}'\mathbf{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{I}_m)^{-1}\boldsymbol{\beta}]\,d\alpha \end{aligned}$$

***Lemma 3-A5*:** (A-optimum):

$$\operatorname{tr} MSE\{\hat{\boldsymbol{\xi}}\} = extr$$

$$\Leftrightarrow$$

$$\frac{d}{d\alpha}(\operatorname{tr} MSE\{\hat{\boldsymbol{\xi}}\}) =$$
$$= -2\operatorname{tr}[\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-2}\mathbf{S}^{-1}(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}]+ \tag{A.5}$$
$$+2\alpha\boldsymbol{\xi}'\mathbf{S}^{-1}(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-2}\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\mathbf{S}^{-1}\boldsymbol{\xi}$$
$$= 0$$

$$\Leftrightarrow$$

$$\hat{\alpha} =$$
$$= \frac{\operatorname{tr}[\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\hat{\alpha}\mathbf{S}^{-1})^{-2}\mathbf{S}^{-1}(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\hat{\alpha}\mathbf{S}^{-1})^{-1}]}{\boldsymbol{\xi}'\mathbf{S}^{-1}(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\hat{\alpha}\mathbf{S}^{-1})^{-2}\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\hat{\alpha}\mathbf{S}^{-1})^{-1}\mathbf{S}^{-1}\boldsymbol{\xi}} \tag{A.6}$$

***Proof*:**

$$d\operatorname{tr} MSE\{\hat{\boldsymbol{\xi}}\} =$$
$$= \operatorname{tr}\{d[(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}]\} + d(\boldsymbol{\beta}'\boldsymbol{\beta})$$

"the first term"

$$\operatorname{tr}\{d[(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}]\} =$$
$$= \operatorname{tr} 2\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}d(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}$$
$$= -2\operatorname{tr}[\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-2}\mathbf{S}^{-1}(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}]d\alpha$$

"the second term"

$$d(\boldsymbol{\beta}'\boldsymbol{\beta}) =$$
$$= 2\alpha\boldsymbol{\xi}'\mathbf{S}^{-1}(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-2}\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\mathbf{S}^{-1}\boldsymbol{\xi}d\alpha.$$

"differentiation"

$$\frac{d}{d\alpha}(\operatorname{tr} MSE\{\hat{\boldsymbol{\xi}}\}) =$$
$$-2\operatorname{tr}[\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-2}\mathbf{S}^{-1}(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}]+$$
$$+2\alpha\boldsymbol{\xi}'\mathbf{S}^{-1}(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-2}\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\alpha\mathbf{S}^{-1})^{-1}\mathbf{S}^{-1}\boldsymbol{\xi}$$
$$\frac{d}{d\alpha}(\operatorname{tr} MSE\{\hat{\boldsymbol{\xi}}\}) = 0 \quad \Rightarrow$$

$$\hat{\alpha} =$$
$$= \frac{\operatorname{tr}[\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\hat{\alpha}\mathbf{S}^{-1})^{-2}\mathbf{S}^{-1}(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\hat{\alpha}\mathbf{S}^{-1})^{-1}]}{\boldsymbol{\xi}'\mathbf{S}^{-1}(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\hat{\alpha}\mathbf{S}^{-1})^{-2}\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}(\mathbf{A}'\boldsymbol{\Sigma}_y^{-1}\mathbf{A}+\hat{\alpha}\mathbf{S}^{-1})^{-1}\mathbf{S}^{-1}\boldsymbol{\xi}}$$

where $\hat{\alpha} = \hat{\alpha}(\boldsymbol{\xi}, \alpha, \mathbf{S})$ reaches a minimum indeed for $\operatorname{tr} MSE\{\hat{\boldsymbol{\xi}}\}$.

q.e.d

Chapter 4

Statistical inference of the eigenspace components of a two-dimensional, symmetric rank-two random tensor

In the deformation analysis in geosciences (geodesy, geophysics and geology), we are often confronted with the problem of a two-dimensional (or planar and horizontal), symmetric rank-two deformation tensor. The *eigen-space components* (*principal components*, *principal direction*) of it play an important role in interpreting the geodetic phenomena like earthquakes (seismic deformations), plate motions and plate deformations among others. With the new space geodetic methods three-dimensional positions and velocities of points in these networks have been determined with high accuracy (~ mm level) from relative regular measurement campaigns, which have become a key tool in plate tectonic studies. This fact suggests that the components of a two-dimensional deformation tensor can be estimated from the high accuracy geodetic data and analyzed through the proper statistical testing procedures. According to the *Measurement Axiom* such a two-dimensional, symmetric rank-two tensor is a *random tensor* **T** which we assume to be an element of the tensor-valued *Gauss-Laplace* normal distribution over $\mathbb{R}^{2\times 2}$ of type independently, identically distributed (i.i.d.) tensor-valued observations, but with identical off-diagonal elements. In this chapter, first, the *eigenspace analysis and synthesis* of a symmetric random matrix are reviewed. *Second*, the nonlinear function, which relates the tensor elements to the eigenspace components, is linearized with respect to a *special nonlinear multivariate Gauss-Markov model*. *Third*, for its linearized form *BLUUE of the eigenspace elements* and *BIQUUE* of its variance-covariance matrix have been established successfully. *Fourth*, the *sampling distribution* of eigenspace components is derived. The test statistics, such as *Hotelling's* T^2, *likelihood ratio statistics and* the general linear hypothesis test with *growth curve model,* are proposed. Hypothesis tests for the random tensor sample means as well as its one variance component are used in the case study of validating a given random strain rate tensor in Chapter 6.

4.1 The eigenspace analysis versus eigenspace synthesis of a two-dimensional, symmetric rank-two random tensor

Let there be given a two-dimensional, symmetric rank-two random strain tensor $\mathbf{T}\in\mathbb{T}_0^2$ which is represented in a commutative left or right orthonormal basis $\{\mathbf{e}^1,\mathbf{e}^2\}$, in short $\mathbf{e}^i\otimes\mathbf{e}^j$ for all $i,j\in\{1,2\}$. "$\otimes$" denotes the tensor product. According to (4.1), $[t_{ij}]\in\mathbb{R}^{2\times 2}$ is called the *matrix representation* of the two-dimensional rank-two tensor. t_{ij} for all $i,j\in\{1,2\}$, establishes the covariant coordinates of the rank-two tensor **T**. The matrix, due to $t_{ij}=t_{ji}$, is symmetric and of full rank two. $\mathbf{T}'$ denotes the transpose of **T**, (4.2). By means of an orthonormal matrix $\mathbf{U}\in\mathbb{SO}(2):=\{\mathbf{U}\in\mathbb{R}^{2\times 2}\mid \mathbf{U}'\mathbf{U}=\mathbf{I}_2, |\mathbf{U}|=+1\}$ the symmetric matrix $\mathbf{T}\in\text{SYM}:=\{\mathbf{T}\in\mathbb{R}^{2\times 2}\mid\mathbf{T}'=\mathbf{T}\}$ can be transformed into the canonical form $\mathbf{\Lambda}=\text{Diag}\{\lambda_1,\lambda_2\}$, also called "*spectral form*".

$$\mathbf{T}=\sum_{i,j=1}^{2}\mathbf{e}^i\otimes\mathbf{e}^j t_{ij}=\sum_{i,j=1}^{2}t_{ij}\mathbf{e}^i\otimes\mathbf{e}^j \tag{4.1}$$

$$\mathbf{T}=[t_{ij}]=\begin{bmatrix} t_{11} & t_{12}\\ t_{21} & t_{22}\end{bmatrix}=[t_{ji}]=\mathbf{T}' \tag{4.2}$$

$$\mathbf{U}:\ \mathbf{T}\mapsto\mathbf{\Lambda}=\text{Diag}(\lambda_1,\lambda_2)=\mathbf{U}'\mathbf{T}\mathbf{U} \tag{4.3}$$

$$\Leftrightarrow$$

$$\mathbf{T}\mathbf{U}-\mathbf{\Lambda}\mathbf{U}=0\ \text{ subject to }\ \mathbf{U}'\mathbf{U}=\mathbf{I}_2 \tag{4.4}$$

$$\Leftrightarrow$$

$$(\mathbf{T}-\lambda_i\mathbf{I}_2)\mathbf{u}_i=0\ \text{ for } i\in\{1,2\} \text{ subject to} \tag{4.5}$$

$$\begin{aligned}\mathbf{u}_1'\mathbf{u}_1=1,\ \mathbf{u}_2'\mathbf{u}_2=1, &\quad:\quad \mathbf{u}_1\in\mathbb{S}^1,\ \mathbf{u}_2\in\mathbb{S}^1\\ \mathbf{u}_1'\mathbf{u}_2=\mathbf{u}_2'\mathbf{u}_1=0 &\quad:\quad \mathbf{u}_1\perp\mathbf{u}_2\end{aligned} \tag{4.6}$$

where

$$\mathbf{U}=\left[\begin{array}{c|c} u_{11} & u_{12}\\ u_{21} & u_{22}\end{array}\right]=[\mathbf{u}_1,\mathbf{u}_2]=\left[\begin{array}{c|c}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{array}\right] \tag{4.7}$$

$$\tan\alpha=\frac{u_{21}}{u_{11}},\ \text{for } \alpha\in\,]-\frac{\pi}{2},+\frac{\pi}{2}],\quad \mathbf{u}_1=\begin{bmatrix}\cos\alpha\\ \sin\alpha\end{bmatrix},\ \mathbf{u}_2=\begin{bmatrix}-\sin\alpha\\ \cos\alpha\end{bmatrix}. \tag{4.8}$$

The formulae (4.3)-(4.8) establish the *eigenspace analysis*. The diagonal matrix $\mathbf{\Lambda}$ contains the *eigenvalues* λ_1, λ_2, the orthonormal matrix **U** the *eigencolumns*, also called coordinates of the *eigenvectors*, namely $[u_{11}, u_{21}]'$ and $[u_{12}, u_{22}]'$. Since $\mathbf{U} \in \mathbb{R}^{2\times 2}$ is an orthonormal matrix, it enjoys the *trigonometric representation* $u_{11} = \cos\alpha$, $u_{21} = \sin\alpha$, and $u_{12} = -\sin\alpha$, $u_{22} = \cos\alpha$. The angular parameter α establishes the *eigenorientation*, namely the orientation of the eigendirections. The solution of the eigenvalue-eigencolumn equation is not unique: There are *four solutions* in general, generated by the quadratic equations (4.6). If we assume that the first element of the eigencolumns *has to be positive* (*Girko* 1995, *Metha* 1991), we arrive at (4.9) and (4.10), respectively. Note, that we have defined the angular parameter α in a *half-open domain* in order to avoid any singularity.

***Corollary* 4.1** (**eigenvalue-eigenvector analysis**)

For a symmetric tensor $\mathbf{T} \in \mathbb{R}^{2\times 2}$ the eigenvalues λ_1, and λ_2 as well as the orientation parameter α, which constitutes the orthonormal matrix $\mathbf{U} \in \mathbb{R}^{2\times 2}$ of eigenvectors are analytically represented by

$$\begin{aligned} &\lambda_1 \in \mathbb{R}, \quad \lambda_1 = \frac{1}{2}\left(t_{11} + t_{22} + \sqrt{(t_{11} - t_{22})^2 + 4t_{21}{}^2}\right) \\ &\lambda_2 \in \mathbb{R}, \quad \lambda_2 = \frac{1}{2}\left(t_{11} + t_{22} - \sqrt{(t_{11} - t_{22})^2 + 4t_{21}{}^2}\right) \\ &\alpha \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right], \ \alpha = \frac{1}{2}\arctan 2t_{12} / (t_{11} - t_{22}) \end{aligned} \tag{4.9}$$

***Corollary* 4.2** (eigenvalue-eigenvector synthesis)

Given the eigenvalues λ_1, and λ_2 as well as the orientation parameter α, which constitute the orthonormal matrix $\mathbf{U} \in \mathbb{R}^{2\times 2}$ of eigencolumns, the symmetric tensor $\mathbf{T} \in \mathbb{R}^{2\times 2}$ is synthetically represented by

$$\begin{aligned} t_{11} &= \lambda_1 \cos^2\alpha + \lambda_2 \sin^2\alpha \\ t_{21} &= t_{12} = \tfrac{1}{2}(\lambda_1 - \lambda_2)\sin 2\alpha \\ t_{22} &= \lambda_1 \sin^2\alpha + \lambda_2 \cos^2\alpha \end{aligned} \tag{4.10}$$

On the basis of *Corollary* 4.1 (eigenspace analysis) and *Corollary* 4.2 (eigenspace synthesis) we are able to portray the symmetric strain tensor **T**, which can be visualized as *strain ellipse*, if $\operatorname{sign}\lambda_1 = \operatorname{sign}\lambda_2$, but as the *strain hyperbola*, if $\operatorname{sign}\lambda_1 \neq \operatorname{sign}\lambda_2$. Figure 4.1 illustrates the *strain* ellipse, Figure 4.2 illustrates the *strain hyperbola*. In the *first case*, the axes of the strain ellipse are directed along the eigenvectors of the strain tensor; the *semi-major axes* of the strain ellipse are identified with the *maximum principal strain* as well as with the *minimum principal strain*, constrained by $\operatorname{sign}\lambda_1 = \operatorname{sign}\lambda_2$. If $\operatorname{sign}\lambda_1 = \operatorname{sign}\lambda_2 = +1$ we speak of *extension*, if $\operatorname{sign}\lambda_1 = \operatorname{sign}\lambda_2 = -1$ of *contraction* instead. Alternatively, in the *second case*, the axes of the strain hyperbola are directed along the eigenvectors of the strain tensor, indicated by the “real axis” showing λ_1 and the “*imaginary axis*” with $|\lambda_2|$, for instance. The notation used in describing the two-dimensional strain tensor are defined in *Box* 4.1.

Box 4.1 (two-dimensional strain tensor)

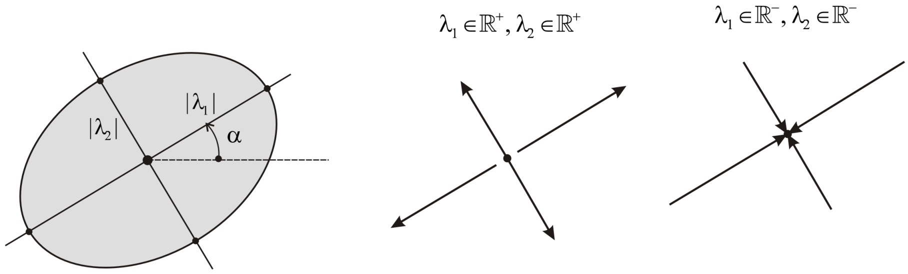
Two-dimensional strain tensor components:

a) t_{11} the normal strain along the 1-axis, positive for extension, negative for contraction.
b) t_{22} the normal strain along the 2-axis, positive for extension, negative for contraction.
c) t_{21} the shear strain ($= t_{12}$), positive for right lateral shear.

The principal components:

d) λ_1 maximum principal strain, the greatest change of length per unit length.
e) λ_2 minimum principal strain, the smallest change of length per unit length.
f) α bearing, or the direction of the maximum principal axis, counterclockwise from the 1-axis (East).

Figure 4.1: Strain ellipse,
sign $\lambda_1 = $ sign λ_2,
extension or contraction



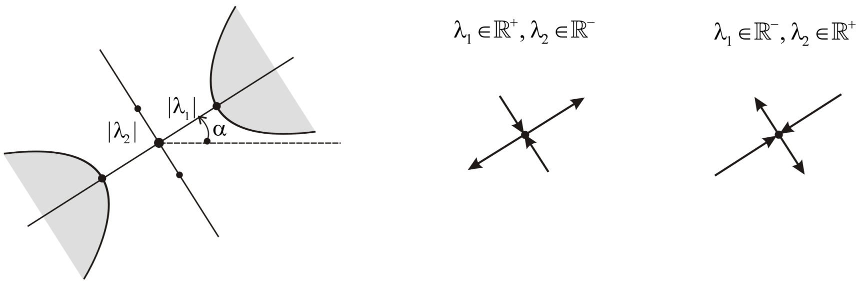


Figure 4.2: Strain hyperbola,
sign $\lambda_1 \neq$ sign λ_2,
extension in one direction, contraction in the other direction *or* vice versa

4.2 The linearized multivariate *Gauss-Markov* model for the estimation of eigenspace components of a two-dimensional, symmetric rank-two random tensor

Chapter 4.1 has documented that the eigenspace synthesis of a symmetric random tensor is nonlinear in terms of the tensor-valued observations, and there is no simple probability density function of the distribution of random eigenspace components. Accordingly, we are unable to derive the exact sampling distribution directly. Here, we will derive the linearized counterpart for sampling the eigenspace synthesis parameters from the originally nonlinear observation equations. The $\mathbf{\Sigma}$ -BLUUE of eigenspace components and their variance-covariance matrix estimate of type BIQUUE will be developed in accordance with the formulas presented earlier by *J. Cai, E. Grafarend and B. Schaffrin* (2001b).

Let us first review the eigenspace analysis *versus* eigenspace synthesis of a symmetric rank-two random tensor as discussed in *Chapter* 4.1. Here we have added the parameter q as a *reduced quaternion element* allowing an algebraic representation of **U**, replacing α which is trigonometric.

***Box* 4.2:**
Eigenspace analysis *versus* eigenspace synthesis
of a two-dimensional, symmetric rank-two random tensor

$$\mathbf{T} = [t_{ij}] \in \mathbb{R}^{2\times 2}$$

$$\text{vech}\,\mathbf{T} = \begin{bmatrix} t_{11} \\ t_{21} \\ t_{22} \end{bmatrix} =: \mathbf{y} \in \mathbb{R}^{3\times 1},$$

(*read* : vec half)

***Box* 4.2 (cont.)**

analysis	*synthesis*
1st parameterization	*1st parameterization*
$\lambda_1 = \frac{1}{2}(t_{11}+t_{22}+\sqrt{(t_{11}-t_{22})^2+4t_{21}^{\ 2}})$	$t_{11} = \lambda_1 \cos^2\alpha + \lambda_2 \sin^2\alpha$
$\lambda_2 = \frac{1}{2}(t_{11}+t_{22}-\sqrt{(t_{11}-t_{22})^2+4t_{21}^{\ 2}})$	$t_{21} = \frac{1}{2}(\lambda_1-\lambda_2)\sin 2\alpha$
$\tan 2\alpha = 2t_{21}/(t_{11}-t_{22})$	$t_{22} = \lambda_1 \sin^2\alpha + \lambda_2 \cos^2\alpha$

(4.11)

2nd parameterization	*2nd parameterization*
$\lambda_1 = \frac{1}{2}(t_{11}+t_{22}+\sqrt{(t_{11}-t_{22})^2+4t_{21}^{\ 2}})$	$t_{11} = \frac{1}{(1+q^2)^2}[\lambda_1(1-q^2)^2+4\lambda_2 q^2]$
$\lambda_2 = \frac{1}{2}(t_{11}+t_{22}-\sqrt{(t_{11}-t_{22})^2+4t_{21}^{\ 2}})$	$t_{21} = \frac{1}{(1+q^2)^2}[2(\lambda_1-\lambda_2)q(1-q^2)]$
$q = \tan\frac{\alpha}{2} = \tan\frac{1}{4}\arctan\frac{2t_{21}}{t_{11}-t_{22}}$,	$t_{22} = \frac{1}{(1+q^2)^2}[4\lambda_1 q^2+\lambda_2(1-q^2)^2]$

(4.12)

$$1+q^2 = \frac{1}{\cos^2(\alpha/2)}, \quad \frac{1-q^2}{1+q^2} = \cos\alpha,$$

$$1-q^2 = \frac{\cos\alpha}{\cos^2(\alpha/2)}, \quad \frac{2q}{1+q^2} = \sin\alpha.$$

Suppose a sample of n observations of $\mathbf{T}$, namely $\mathbf{T}_1, \mathbf{T}_2, \ldots, \mathbf{T}_n$ whose related vectorized forms are $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n$. Here we design an array of vectorized tensor coordinates $y_1 := t_{11}, y_2 := t_{21}, y_3 := t_{22}$ indexed to the number of the sample. *For instance*, $y_{2.3}$ denotes the tensor coordinate $y_2 = t_{21}$ in the *third sample*

$$\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_n] = \begin{bmatrix} y_{1.1} & \cdots & y_{1.n} \\ y_{2.1} & \cdots & y_{2.n} \\ y_{3.1} & \cdots & y_{3.n} \end{bmatrix}, \quad \mathbf{Y} \in \mathbb{R}^{3\times n} \quad , \tag{4.13}$$

whose variance-covariance matrix follows when $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n$ are independent 3×1 random vectors, each with the 3×3 variance-covariance matrix $\mathbf{\Sigma}_{\mathrm{y}i}$, as

$$D(\mathrm{vec}\,\mathbf{Y}) = \begin{bmatrix} \mathbf{\Sigma}_{\mathrm{y}1} & 0 & \cdots & 0 \\ 0 & \mathbf{\Sigma}_{\mathrm{y}2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \mathbf{\Sigma}_{\mathrm{y}n} \end{bmatrix}, \quad D\{\mathrm{vec}\,\mathbf{Y}\} \in \mathbb{R}^{3n\times3n}; \tag{4.14}$$

and when the $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n$ are i.i.d. 3×1 random vectors, each with the same variance-covariance matrix $\mathbf{\Sigma}_{\mathrm{y}}$, we have

$$D(\mathrm{vec}\,\mathbf{Y}) = \begin{bmatrix} \mathbf{\Sigma}_{\mathrm{y}} & 0 & \cdots & 0 \\ 0 & \mathbf{\Sigma}_{\mathrm{y}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \mathbf{\Sigma}_{\mathrm{y}} \end{bmatrix} = \mathbf{I}_n \otimes \mathbf{\Sigma}_{\mathrm{y}}, \quad D\{\mathrm{vec}\mathbf{Y}\} \in \mathbb{R}^{3n\times3n}, \tag{4.15}$$

where $\otimes$ now denotes the *Kronecker –Zehfuss* product of matrices (see *Henderson, Pukelsheim and Searle* 1981; or *Grafarend and B. Schaffrin*, 1993).

Using the eigenspace analysis *versus* eigenspace synthesis we can define the nonlinear *Gauss-Markov* model which is presented by (4.16) - (4.19), where **1** denotes the $n\times1$ "*summation vector*" with all its entries being 1.

***Box* 4.3:**
Special nonlinear multivariate *Gauss-Markov*
model for sampling the eigenspace synthesis

$$\mathbf{Y} = \mathbf{F}(\boldsymbol{\xi})\mathbf{1}' + \bar{\mathbf{E}} \tag{4.16}$$

1st moments

$$E\{\mathbf{Y}\} = E\{\begin{bmatrix} y_{1.1} & \cdots & y_{1.n} \\ y_{2.1} & \cdots & y_{2.n} \\ y_{3.1} & \cdots & y_{3.n} \end{bmatrix}\} = \mathbf{F}\mathbf{1}'$$

$$\mathbf{F} := \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix} = \begin{bmatrix} \lambda_1 \cos^2\alpha + \lambda_2 \sin^2\alpha \\ \frac{1}{2}(\lambda_1 - \lambda_2)\sin 2\alpha \\ \lambda_1 \sin^2\alpha + \lambda_2 \cos^2\alpha \end{bmatrix} = \begin{bmatrix} \xi_1 \cos^2\xi_3 + \xi_2 \sin^2\xi_3 \\ \frac{1}{2}(\xi_1 - \xi_2)\sin 2\xi_3 \\ \xi_1 \sin^2\xi_3 + \xi_2 \cos^2\xi_3 \end{bmatrix}, \tag{4.17}$$

$$\left.\begin{matrix} \xi_1 := \lambda_1 \\ \xi_2 := \lambda_2 \\ \xi_3 := \alpha \end{matrix}\right\} \text{ and } \left\{\begin{matrix} y_1 := t_{11} \\ y_2 := t_{21} \\ y_3 := t_{22} \end{matrix}\right.$$

2nd moments

"*independent between observations*"

$$D(\operatorname{vec}\mathbf{Y}) = \begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{y}1} & 0 & \cdots & 0 \\ 0 & \boldsymbol{\Sigma}_{\mathbf{y}2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \boldsymbol{\Sigma}_{\mathbf{y}n} \end{bmatrix}, \quad D\{\operatorname{vec}\mathbf{Y}\} \in \mathbb{R}^{3n\times 3n}, \tag{4.18}$$

"*i.i.d. observations*"

$$D(\operatorname{vec}\mathbf{Y}) = \begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{y}} & 0 & \cdots & 0 \\ 0 & \boldsymbol{\Sigma}_{\mathbf{y}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \boldsymbol{\Sigma}_{\mathbf{y}} \end{bmatrix} = \mathbf{I}_n \otimes \boldsymbol{\Sigma}_{\mathbf{y}}, \quad D\{\operatorname{vec}\mathbf{Y}\} \in \mathbb{R}^{3n\times 3n}, \tag{4.19}$$

$D\{\operatorname{vec}\mathbf{Y}\} = \boldsymbol{\Sigma} \in \mathbb{R}^{3n\times 3n}$, $\boldsymbol{\Sigma}$ positive-definite, rk $\boldsymbol{\Sigma} = 3n$,
$\boldsymbol{\xi}$, $E\{\mathbf{Y}\}$, $\mathbf{Y} - E\{\mathbf{Y}\} = \bar{\mathbf{E}}$ unknown, $\boldsymbol{\Sigma}$ unknown (but patterned).

In order to estimate the eigenspace components of a symmetric rank-two random tensor, the nonlinear observation equations will be linearized. The linearization process of nonlinear observation equations is applied to the nonlinear mapping $\boldsymbol{\xi} \mapsto \mathbf{F}(\boldsymbol{\xi})$. *The Taylor expansion*

$$\begin{aligned} \mathbf{F}(\boldsymbol{\xi}) = {} & \mathbf{F}(\boldsymbol{\xi}_0) + \mathbf{J}(\boldsymbol{\xi}_0)(\boldsymbol{\xi} - \boldsymbol{\xi}_0) + \mathbf{H}(\boldsymbol{\xi}_0)(\boldsymbol{\xi} - \boldsymbol{\xi}_0) \otimes (\boldsymbol{\xi} - \boldsymbol{\xi}_0) + \\ & + \mathcal{O}[(\boldsymbol{\xi} - \boldsymbol{\xi}_0) \otimes (\boldsymbol{\xi} - \boldsymbol{\xi}_0) \otimes (\boldsymbol{\xi} - \boldsymbol{\xi}_0)] \end{aligned} \tag{4.20}$$

is truncated to the order $\mathcal{O}[(\boldsymbol{\xi} - \boldsymbol{\xi}_0) \otimes (\boldsymbol{\xi} - \boldsymbol{\xi}_0) \otimes (\boldsymbol{\xi} - \boldsymbol{\xi}_0)]$; $\mathbf{J}(\boldsymbol{\xi}_0)$ and $\mathbf{H}(\boldsymbol{\xi}_0)$ represent the *Jacobi matrix* of the first partial derivatives, and the *Hesse matrix* of second derivatives, respectively, of the vector-valued function $\mathbf{F}(\boldsymbol{\xi})$ with respect to the coordinates of the vector $\boldsymbol{\xi}$, both taken at the evaluation point $\boldsymbol{\xi}_0$. In our study the linearized nonlinear model is generated by truncating the vector-valued function $\mathbf{F}(\boldsymbol{\xi})$ to the order $\mathcal{O}[(\boldsymbol{\xi} - \boldsymbol{\xi}_0) \otimes (\boldsymbol{\xi} - \boldsymbol{\xi}_0)]$, namely

$$\begin{aligned} \mathbf{F}(\boldsymbol{\xi}) - \mathbf{F}(\boldsymbol{\xi}_0) &= \mathbf{J}(\boldsymbol{\xi}_0)(\boldsymbol{\xi} - \boldsymbol{\xi}_0) + \mathcal{O}[(\boldsymbol{\xi} - \boldsymbol{\xi}_0) \otimes (\boldsymbol{\xi} - \boldsymbol{\xi}_0)] = \\ &= \mathbf{J}(\boldsymbol{\xi}_0)\Delta\boldsymbol{\xi} + \mathcal{O}[(\boldsymbol{\xi} - \boldsymbol{\xi}_0) \otimes (\boldsymbol{\xi} - \boldsymbol{\xi}_0)]. \end{aligned} \tag{4.21}$$

The linearization of the nonlinear observation equations (4.16) is presented in detail in *Box* 4.4:

***Box* 4.4**:
Linearization of nonlinear observation equations

First set of observation: $\begin{bmatrix} y_{1.1} \\ y_{2.1} \\ y_{3.1} \end{bmatrix} = \begin{bmatrix} t_{11.1} \\ t_{21.1} \\ t_{22.1} \end{bmatrix}$

Define: $$\xi_0 := \begin{bmatrix} \lambda_{1.1} \\ \lambda_{2.1} \\ \alpha_{.1} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} t_{11.1} + t_{22.1} + \sqrt{(t_{11.1} - t_{22.1})^2 + 4t_{21.1}{}^2} \\ t_{11.1} + t_{22.1} - \sqrt{(t_{11.1} - t_{22})^2 + 4t_{21.1}{}^2} \\ \arctan 2t_{21.1} / (t_{11.1} - t_{22.1}) \end{bmatrix} \tag{4.22}$$

"*Linearized nonlinear model*"

$$\mathbf{F}(\xi) - \mathbf{F}(\xi_0) = \mathbf{J}(\xi_0)\Delta\xi + \mathcal{O}[(\xi - \xi_0) \otimes (\xi - \xi_0)] \tag{4.23}$$

"*Jacobi matrix*"

$$\mathscr{A} = \mathbf{J}(\xi_0) = \begin{bmatrix} \frac{\partial f_1}{\partial \lambda_1} & \frac{\partial f_1}{\partial \lambda_2} & \frac{\partial f_1}{\partial \alpha} \\ \frac{\partial f_2}{\partial \lambda_1} & \frac{\partial f_2}{\partial \lambda_2} & \frac{\partial f_2}{\partial \alpha} \\ \frac{\partial f_3}{\partial \lambda_1} & \frac{\partial f_3}{\partial \lambda_2} & \frac{\partial f_3}{\partial \alpha} \end{bmatrix}_{\xi=\xi_0} = \begin{bmatrix} \cos^2 \alpha_{.1} & \sin^2 \alpha_{.1} & (\lambda_{2.1} - \lambda_{1.1}) \sin 2\alpha_{.1} \\ \frac{1}{2} \sin 2\alpha_{.1} & -\frac{1}{2} \sin 2\alpha_{.1} & -(\lambda_{2.1} - \lambda_{1.1}) \cos 2\alpha_{.1} \\ \sin^2 \alpha_{.1} & \cos^2 \alpha_{.1} & -(\lambda_{2.1} - \lambda_{1.1}) \sin 2\alpha_{.1} \end{bmatrix}. \tag{4.24}$$

Based upon the *Taylor series expansion* for $\mathbf{F}(\xi)$ we shall apply the *Gauss-Newton iteration scheme* with $\xi_0 = [\lambda_1, \lambda_2, \alpha]'_1$ as the starting point. ξ_0 is determined by solving once the eigenvalue analysis equations as indicated by (4.22) for the *sample* one. In this way, we have established the design matrix of the first kind $\mathscr{A} = \mathbf{J}(\xi_0)$ as the *Jacobi matrix* $\mathbf{J}$ at the point ξ_0. The special linearized multivariate *Gauss-Markov model* for sampling the eigenspace of a symmetric random matrix is summarized by (4.25) ~ (4.30).

***Box* 4.5**:

Special linearized multivariate *Gauss-Markov* model
for sampling the eigenspace synthesis

$$\mathbf{Y} = \mathbf{F}(\xi_0)\mathbf{1}' + [\mathscr{A}(\xi - \xi_0)]\mathbf{1}' + \mathbf{E} \tag{4.25}$$

"*vectorized version*"

$$\operatorname{vec}\mathbf{Y} = \mathbf{1} \otimes \mathbf{F}(\xi_0) + (\mathbf{1} \otimes \mathscr{A})(\xi - \xi_0) + \operatorname{vec}\mathbf{E} \tag{4.26}$$

with denotations:

$$\operatorname{vec}\mathbf{Y}_0 = \mathbf{1} \otimes \mathbf{F}(\xi_0),\ \mathbf{A} = (\mathbf{1} \otimes \mathscr{A})$$

"*1st moments*"

$$\mathbf{A}(\xi - \xi_0) + \operatorname{vec}\mathbf{Y}_0 = E\{\operatorname{vec}\mathbf{Y}\},\ \operatorname{vec}\mathbf{Y} \in \mathbb{R}^{3n\times 1} \tag{4.27}$$

"*2nd moments*"

$$\begin{gathered} D\{\operatorname{vec}\mathbf{Y}\} = \mathbf{I}_n \otimes \mathbf{\Sigma}_\mathrm{y} := \mathbf{\Sigma} \in \mathbb{R}^{3n\times 3n},\ \mathbf{\Sigma}_\mathrm{y} \text{ positive-definite, rk } \mathbf{\Sigma} = 3n; \\ \xi,\ E\{\mathbf{Y}\},\ \mathbf{Y} - E\{\mathbf{Y}\} + \mathcal{O}\{[(\xi - \xi_0) \otimes (\xi - \xi_0)]\mathbf{1}'_n\} = \mathbf{E} \text{ unknown, } \mathbf{\Sigma}_\mathrm{y} \text{ unknown.} \end{gathered} \tag{4.28}$$

With these definitions and the observations of a random tensor we can *first* estimate the eigenspace components ξ of type $\mathbf{\Sigma}$ - BLUUE (*Best Linear Uniformly Unbiased Estimation*) which are collected in

***Theorem* 4.3** ($\hat{\xi}$ $\boldsymbol{\Sigma}$ – BLUUE of ξ , the eigenspace components of a symmetric random tensor):

The $\boldsymbol{\Sigma}$ - BLUUE $\hat{\xi}$ of ξ in the *special linearized multivariate Gauss-Markov Model* is

$$\boxed{\hat{\xi} = \xi_0 + \Delta\hat{\xi}} \text{ with}$$

$$\begin{aligned}
\Delta\hat{\xi} &= (\hat{\xi} - \xi_0) = \mathbf{L}(\operatorname{vec}\mathbf{Y} - \operatorname{vec}\mathbf{Y}_0) = \\
&= (\mathbf{A}'\boldsymbol{\Sigma}^{-1}\mathbf{A})^{-1}\mathbf{A}'\boldsymbol{\Sigma}^{-1}(\operatorname{vec}\mathbf{Y} - \operatorname{vec}\mathbf{Y}_0) \\
&= [\frac{1}{n}\mathbf{1}' \otimes (\mathcal{A}'\boldsymbol{\Sigma}_y^{-1}\mathcal{A})^{-1}\mathcal{A}'\boldsymbol{\Sigma}_y^{-1}](\operatorname{vec}\mathbf{Y} - \operatorname{vec}\mathbf{Y}_0) \\
&= \frac{1}{n}(\mathcal{A}'\boldsymbol{\Sigma}_y^{-1}\mathcal{A})^{-1}\mathcal{A}'\boldsymbol{\Sigma}_y^{-1}(\mathbf{Y} - \mathbf{Y}_0)\mathbf{1})
\end{aligned} \tag{4.29}$$

subject to the related dispersion matrix

$$D\{\hat{\xi}\} := D\{\Delta\hat{\xi}\} = \boldsymbol{\Sigma}_{\hat{\xi}} = \frac{1}{n}(\mathcal{A}'\boldsymbol{\Sigma}_y^{-1}\mathcal{A})^{-1}. \tag{4.30}$$

Since the variance-covariance matrix $\boldsymbol{\Sigma}_y$ of the observation vector is unknown we have to estimate such a dispersion matrix empirically. $\hat{\boldsymbol{\Sigma}}_y$ as the BIQUUE (*Best Invariant Quadratic Uniformly Unbiased Estimate*) of $\boldsymbol{\Sigma}_y$ is summarized in *Theorem* 4.4, e.g. proven in Koch (1987; 1999).

***Theorem* 4.4** (The sample variance-covariance matrix $\hat{\boldsymbol{\Sigma}}_y$ of type BIQUUE of a symmetric random tensor):

The sample variance-covariance matrix $\hat{\boldsymbol{\Sigma}}_y$ of type BIQUUE for the vectorized observations of a symmetric rank-two random tensor is

$$\hat{\boldsymbol{\Sigma}}_y = \frac{1}{n-1}\mathbf{Y}(\mathbf{I}_n - \frac{1}{n}\mathbf{1}\mathbf{1}')\mathbf{Y}' = \frac{1}{n-1}\boldsymbol{\Omega}. \tag{4.31}$$

4.3 Hypothesis testing for the estimates of eigenspace components of a two-dimensional, symmetric rank-two random tensor

In order for the estimated tensors to be significant, statistical inference has to be applied. Based on the three elements of validation, namely sampling distribution, parameter estimation (point estimate and interval estimation) and hypothesis testing, we refer to *Kendall* and *Stuart* (1958) for the univariate hypothesis test and *Giri* (1977), *Rencher* (1995, 1998), *Anderson* (1958, 1984) and *Muirhead* (1982) for the multivariate hypothesis test.

The sampling distribution of the symmetric rank-two random tensor has been derived in Chapter 1 and 2. On the basis of such a sampling distribution, the distributions of multivariate test statistics needed for testing hypotheses concerning the parameters (mean vector and covariance matrix) for a tensor-valued multivariate *Gauss-Laplace* normal population of a two-dimensional symmetric rank-two random tensor, such as *Hotelling's* T^2 , the *likelihood ratio statistics and* the general linear hypothesis test with the *growth curve model*, are proposed, too.

With the estimates of eigenspace components of random strain rate tensor and their dispersion matrix the following multivariate hypothesis tests will be suggested:

- Test for the eigenspace parameter vector $\xi = \xi_0$ with $\boldsymbol{\Sigma}_y$ unspecified;
- Test for a distinct element of the eigenspace parameter vector with *Student t- test*;
- *Eigen inference* about the orthonormally transformed parameters $\boldsymbol{\eta}$;
- Test for the variance-covariance matrix $\boldsymbol{\Sigma}_y = \boldsymbol{\Sigma}_0$;
- Test for the eigenspace parameter vector and variance-covariance matrix $\xi = \xi_0$, $\boldsymbol{\Sigma}_y = \boldsymbol{\Sigma}_0$;
- The general linear hypothesis test with the growth curve model for eigenspace parameters.

4.3.1 Test for the eigenspace parameter vector $\xi = \xi_0$ with $\mathbf{\Sigma}_y$ unspecified

***Box* 4.6**:

Multivariate hypothesis test about the eigenspace parameter vector ξ assuming *Gauss-Laplace* normally distributed observations of a two-dimensional, symmetric rank-two random tensor

First Test for $\mathcal{H}_{01}: \ \xi = \xi_0, \mathcal{H}_{11}: \ \xi \neq \xi_0$ with $\mathbf{\Sigma}_y$ unspecified;

$$\Leftrightarrow$$

$$\mathcal{H}_{01}: \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \alpha \end{bmatrix} = \begin{bmatrix} \lambda_{10} \\ \lambda_{20} \\ \alpha_0 \end{bmatrix}, \ \mathcal{H}_{11}: \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \alpha \end{bmatrix} \neq \begin{bmatrix} \lambda_{10} \\ \lambda_{20} \\ \alpha_0 \end{bmatrix} \quad \text{with } \mathbf{\Sigma}_y \text{ unspecified}$$

"*Hotelling's* T^2 statistic"

(*Hotelling* 1931, *Muirhead* 1982, *Rencher* 1998)

$$T^2 := [\hat{\xi} - \xi_0]' \hat{\mathbf{\Sigma}}_{\hat{\xi}}^{-1} [\hat{\xi} - \xi_0] \tag{4.32}$$

Note that

$$\frac{n-3}{(n-1)\cdot 3} T^2$$

is an element of *Fisher's F*-distribution $F_{3,\ n-1-3+1}$ (*Rencher* 1998) and

$$P\{T^2 \leq \frac{(n-1)\cdot 3}{n-3} F_{3,\ n-3}(1-\alpha)\} =$$

$$= P\{[\hat{\xi} - \xi_0]' \hat{\mathbf{\Sigma}}_{\hat{\xi}}^{-1} [\hat{\xi} - \xi_0] \leq \frac{(n-1)\cdot 3}{n-3} F_{3,\ n-3}(1-\alpha)\} = 1 - \alpha = \gamma$$

where $F_{3,\ n-3}(1-\alpha)$ is the upper $(100\alpha)th$ percentile of *Fisher's F*-distribution. This immediately leads to a test of the hypothesis $\mathcal{H}_{01}: \xi = \xi_0$ versus $\mathcal{H}_{11}: \xi \neq \xi_0$. At the error probability α we reject $\mathcal{H}_{01}$ in favor of $\mathcal{H}_{11}$ if

$$T^2 = [\hat{\xi} - \xi_0]' \hat{\mathbf{\Sigma}}_{\hat{\xi}}^{-1} [\hat{\xi} - \xi_0] > \frac{(n-1)\cdot 3}{n-3} F_{3,\ n-3}(1-\alpha)\} = T^2_{1-\alpha}$$

4.3.2 Test for a distinct element of the eigenspace parameter vector with *Student t- test*

***Box* 4.7:**

Separate Student t-tests about the eigenspace parameters in ξ

$$\begin{array}{lll|l|l} \textit{Second Test} \text{ for} & \mathcal{H}_{02}: & \lambda_1 = \lambda_{10} & \lambda_2 = \lambda_{20} & \alpha = \alpha_0 \\ \text{(separately)} & \mathcal{H}_{12}: & \lambda_1 \neq \lambda_{10} & \lambda_2 \neq \lambda_{20} & \alpha \neq \alpha_0 \end{array}$$

"two-sided tests with *the test quantities*"

$$t_1 := \frac{\hat{\lambda}_1 - \lambda_{10}}{\hat{\sigma}_1}, \quad t_2 := \frac{\hat{\lambda}_2 - \lambda_{20}}{\hat{\sigma}_2}, \quad t_3 := \frac{\hat{\alpha} - \alpha_0}{\hat{\sigma}_3} \tag{4.33}$$

with respect to $\hat{\lambda}_1, \hat{\lambda}_2, \hat{\alpha}$ of type $\mathbf{\Sigma}$-BLUUE and their variances. *t_1, t_2 and t_3* are elements of the *Student t-distribution with n-1 degrees of freedom*.

The probability identity

$$P\{c_1 \leq t \leq c_2\} = P\{c_1\hat{\sigma} + \mu_0 \leq \hat{\mu} \leq c_2\hat{\sigma} + \mu_0\} = 1 - \alpha = \gamma$$

relates the error probability α of the two-sided test to the confidence level γ. If $\hat{\mu}$ is an element of the confidence interval $c_1\hat{\sigma} + \mu_0 \leq \hat{\mu} \leq c_2\hat{\sigma} + \mu_0$, the null hypothesis $\mathcal{H}_0: \mu = \mu_0$ is accepted. We reject $\mathcal{H}_0$ if the confidence interval does not contain $\hat{\mu}$.

As an example the 95% *confidence interval* for the eigenvalues λ_1, λ_2 and the eigendirection α $[c_1\hat{\sigma}_1 + \lambda_{10}, c_2\hat{\sigma}_1 + \lambda_{10}]$, $[c_1\hat{\sigma}_2 + \lambda_{20}, c_2\hat{\sigma}_2 + \lambda_{20}]$ and $[c_1\hat{\sigma}_3 + \lambda_{30}, c_2\hat{\sigma}_3 + \lambda_{30}]$ are illustrated in *Figure* 4.3, respectively.

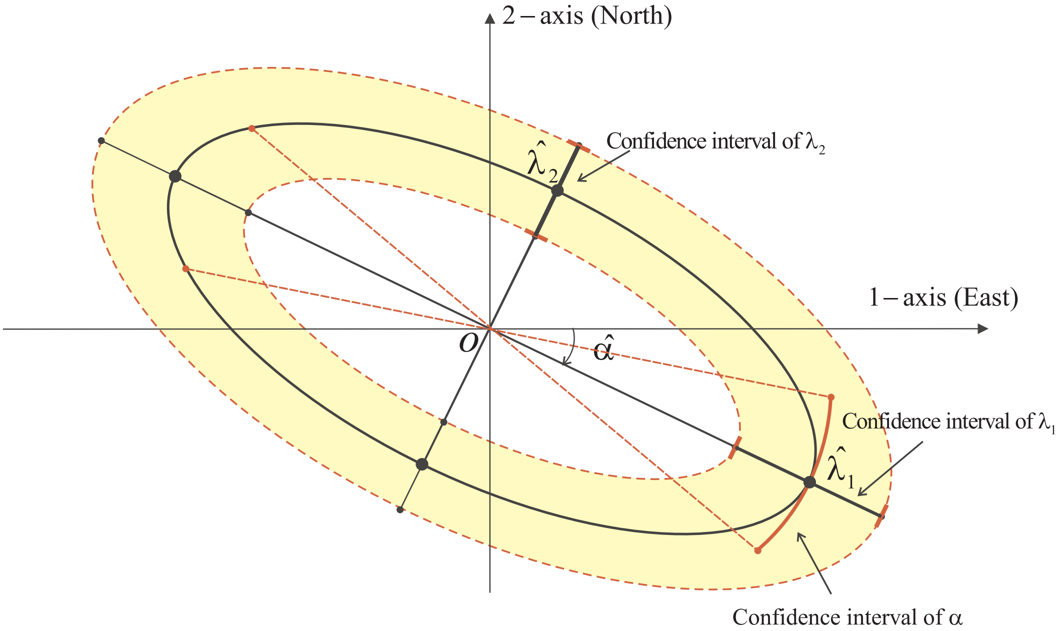


Figure 4.3 The 95% confidence interval for the eigenvalues λ_1, λ_2 and eigendirection α

4.3.3 *Eigen-inference* about the orthonormally transformed parameters $\boldsymbol{\eta}$

From the dispersion matrix $\hat{\boldsymbol{\Sigma}}_{\hat{\xi}}$, the variance-covariance matrix of the eigenspace component parameter vector $\hat{\boldsymbol{\xi}}$ estimated by (4.30), we can see that these eigenspace component parameters are correlated. In order to make the hypothesis tests about the distinct elements more efficient and uncorrelated, we could transform the original parameters into new parameters η_i of uncorrelated linear combinations of ξ_i's. This method uses a similar technique as the well known *principal component analysis*, which was introduced by *K. Pearson* (1901) as a tool of fitting planes to a system of points in space and later generalized by *Hotelling* (1931) for analyzing correlation structures. In fact principal component analysis is concerned fundamentally with the eigenstructure of covariance matrices, i.e., with their eigenvalues and eigenvectors. Therefore in our study we will firstly make an orthonormal transformation of the original parameters, then derive the covariance matrix of the transformed parameters and perform a hypothesis test for them, which we call *eigen-inference*.

From *Theorem 4.3* we have the $\boldsymbol{\Sigma}$-BLUUE of eigenspace components of a symmetric random tensor

$$\hat{\boldsymbol{\xi}} = [\hat{\lambda}_1 \ \hat{\lambda}_2 \ \hat{\alpha}]',$$

and the related dispersion matrix of $\hat{\boldsymbol{\xi}}$ of type BIQUUE

$$D\{\hat{\boldsymbol{\xi}}\} = \hat{\boldsymbol{\Sigma}}_{\hat{\xi}},$$

with the spectral decomposition of the dispersion matrix $\hat{\boldsymbol{\Sigma}}_{\hat{\xi}}$

$$\hat{\boldsymbol{\Sigma}}_{\hat{\xi}} = \mathbf{U}_{\hat{\xi}} \boldsymbol{\Lambda}_{\hat{\eta}} \mathbf{U}'_{\hat{\xi}}, \tag{4.34}$$

where the orthogonal transformation matrix $\mathbf{U}_{\hat{\xi}}$ contains normalized eigenvectors as column vectors (i.e., orthonormal basis, orthonormal matrix)

$$\mathbf{U}_{\hat{\xi}} \text{ with } \mathbf{U}'_{\hat{\xi}}\mathbf{U}_{\hat{\xi}} = \mathbf{I}, \ \det(\mathbf{U}_{\hat{\xi}}) = 1. \tag{4.35}$$

Then we set the transformed parameter vector from the original parameter vector and find their $\boldsymbol{\Sigma}$-BLUUE estimates, respectively, as

$$\boldsymbol{\eta} := \mathbf{U}'_{\hat{\xi}}\boldsymbol{\xi} \ \text{ and } \ \hat{\boldsymbol{\eta}} = \mathbf{U}'_{\hat{\xi}}\hat{\boldsymbol{\xi}} \tag{4.36}$$

which transforms the null hypothesis values of Test 2 in *Box* 4.7:

$$\boldsymbol{\eta}_0 = \mathbf{U}'_{\hat{\xi}}\boldsymbol{\xi}_0 \tag{4.37}$$

Then, from (4.34), we get

$$\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\eta}}} = \mathbf{U}'_{\hat{\xi}}\hat{\boldsymbol{\Sigma}}_{\hat{\xi}}\mathbf{U}_{\hat{\xi}} = \boldsymbol{\Lambda}_{\hat{\boldsymbol{\eta}}} = \begin{bmatrix} \hat{\sigma}^2_{\hat{\eta}_1} & 0 & 0 \\ 0 & \hat{\sigma}^2_{\hat{\eta}_2} & 0 \\ 0 & 0 & \hat{\sigma}^2_{\hat{\eta}_3} \end{bmatrix}, \tag{4.38}$$

from which we can see that the transformed parameters η_i are mutually independent and their standard deviations are:

$$\hat{\sigma}_{\hat{\eta}_1} = \sqrt{\hat{\lambda}_{\eta_1}},\ \hat{\sigma}_{\hat{\eta}_2} = \sqrt{\hat{\lambda}_{\eta_2}},\ \hat{\sigma}_{\hat{\eta}_3} = \sqrt{\hat{\lambda}_{\eta_3}}. \tag{4.39}$$

With these orthonormally transformed results we can now perform the *eigen-inference*. Note that the orthonormally transformed parameters η_i are mutually independently, normally distributed. *Student t-tests* could also be used for every element of the transformed parameters $\hat{\eta}_i$ separately .

The second hypothesis test performed in *Box* 4.7 will be equivalent to the new hypothesis test for the orthonormally transformed parameters, i.e.,

$$\begin{array}{rl} \text{Second Test for } \mathcal{H}_{02}: & \lambda_1 = \lambda_{10} \,\big|\, \lambda_2 = \lambda_{20} \,\big|\, \alpha = \alpha_0 \\ \mathcal{H}_{12}: & \lambda_1 \neq \lambda_{10} \,\big|\, \lambda_2 \neq \lambda_{20} \,\big|\, \alpha \neq \alpha_0 \\ & \Leftrightarrow \\ \text{Eigen-Test for } \mathcal{H}_{02}: & \eta_1 = \eta_{10} \,\big|\, \eta_2 = \eta_{20} \,\big|\, \eta_3 = \eta_{30} \\ \mathcal{H}_{12}: & \eta_1 \neq \eta_{10} \,\big|\, \eta_2 \neq \eta_{20} \,\big|\, \eta_3 \neq \eta_{30} \end{array}$$

which means that, when we accept or reject the new hypothesis tests (*eigen-test*), we will accept or reject the second hypothesis tests accordingly.

These procedures will be summarized in *Box* 4.8.

Box 4.8:

***Eigen-inference* about the transformed parameters $\boldsymbol{\eta}$**

Eigen tests (alternative to the second tests)

$$\begin{array}{rl} \mathcal{H}'_{02}: & \eta_1 = \eta_{10} \,\big|\, \eta_2 = \eta_{20} \,\big|\, \eta_3 = \eta_{30} \\ \mathcal{H}'_{12}: & \eta_1 \neq \eta_{10} \,\big|\, \eta_2 \neq \eta_{20} \,\big|\, \eta_3 \neq \eta_{30} \end{array}$$

"two-sided tests with the *test quantities*"

$$t_1 := \frac{\hat{\eta}_1 - \eta_{10}}{\hat{\sigma}_{\eta_1}},\ t_2 := \frac{\hat{\eta}_2 - \eta_{20}}{\hat{\sigma}_{\eta_2}},\ t_3 := \frac{\hat{\eta}_3 - \eta_{30}}{\hat{\sigma}_{\eta_3}} \tag{4.40}$$

with respect to $\hat{\eta}_1, \hat{\eta}_2, \hat{\eta}_3$ and their related variances . t_1, t_2 *and* t_3 are elements of the *Student t-distribution with n-1 degrees of freedom*.

The probability identity

$$P\{c_1 \leq t \leq +c_2\} = P\{c_1\hat{\sigma} + \eta_0 \leq \hat{\eta} \leq c_2\hat{\sigma} + \eta_0\} = 1 - \alpha = \gamma$$

relates the error probability α of the two-sided test to the confidence level γ. If $\hat{\eta}$ is an element of the confidence interval $c_1\hat{\sigma} + \eta_0 \leq \hat{\eta} \leq c_2\hat{\sigma} + \eta_0$, the null hypothesis $\mathcal{H}_0: \eta = \eta_0$ is accepted. We reject $\mathcal{H}_0$ if the confidence interval does not contain $\hat{\eta}$. Accordingly, we accept the original null hypothesis about the eigenspace components.

This completes the development of *eigen-inference*.

4.3.4 Test for the variance-covariance matrix $\boldsymbol{\Sigma}_\mathbf{y} = \boldsymbol{\Sigma}_0$

***Box* 4.9**

Multivariate hypothesis tests about the variance-covariance matrix $\boldsymbol{\Sigma}_\mathbf{y}$

Third Test for $\mathcal{H}_{03}: \ \boldsymbol{\Sigma}_\mathbf{y} = \boldsymbol{\Sigma}_0, \quad \mathcal{H}_{13}: \ \boldsymbol{\Sigma}_\mathbf{y} \neq \boldsymbol{\Sigma}_0$

"unbiased modified likelihood ratio statistic Λ_1"

(*Giri* 1977, *Muirhead* 1982, Koch 1999, Koch 2001)

$$\Lambda_1 = \left(\frac{e}{n-1}\right)^{3(n-1)/2} (\det(n-1)\hat{\boldsymbol{\Sigma}}_\mathbf{y}\boldsymbol{\Sigma}_0^{-1})^{(n-1)/2} \operatorname{etr}\{-\frac{1}{2}(n-1)\hat{\boldsymbol{\Sigma}}_\mathbf{y}\boldsymbol{\Sigma}_0^{-1}\} \tag{4.41}$$

with respect to the sample variance-covariance matrix $\hat{\boldsymbol{\Sigma}}_\mathbf{y}$ of type BIQUUE .
Since our sample size is relatively small we have to use the exact distribution of $-2\log\Lambda_1$, whose upper 5 and 1 percentage points have been provided by *Muirhead* (1982, p.360).

4.3.5 Test for the eigenspace parameter vector and variance-covariance matrix $\boldsymbol{\xi} = \boldsymbol{\xi}_0, \ \boldsymbol{\Sigma}_\mathbf{y} = \boldsymbol{\Sigma}_0$

***Box* 4.10**

Multivariate hypothesis tests about the eigenspace parameter vector $\boldsymbol{\xi}$ and the variance-covariance matrix $\boldsymbol{\Sigma}_\mathbf{y}$

Fourth Test for $\mathcal{H}_{04}: \ \boldsymbol{\xi} = \boldsymbol{\xi}_0, \ \boldsymbol{\Sigma}_\mathbf{y} = \boldsymbol{\Sigma}_0, \quad \mathcal{H}_{14}: \ \boldsymbol{\xi} \neq \boldsymbol{\xi}_0 \ \textit{or} \ \boldsymbol{\Sigma}_\mathbf{y} \neq \boldsymbol{\Sigma}_0$

"unbiased likelihood ratio statistic Λ_2"

(*Anderson* 1984, *Muirhead* 1982)

$$\Lambda_2 = \left(\frac{e}{n}\right)^{3n/2} (\det(n-1)\hat{\boldsymbol{\Sigma}}_\mathbf{y}\boldsymbol{\Sigma}_0^{-1})^{n/2} \operatorname{etr}\{-\frac{1}{2}(n-1)\hat{\boldsymbol{\Sigma}}_\mathbf{y}\boldsymbol{\Sigma}_0^{-1}\} \exp\{-\frac{1}{2}[\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}_0]' \boldsymbol{\Sigma}_{\hat{\xi}0}^{-1} [\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}_0]\} \tag{4.42}$$

with respect to the eigenspace components of type $\boldsymbol{\Sigma}$-BLUUE and variance-covariance matrix $\hat{\boldsymbol{\Sigma}}_\mathbf{y}$ of type BIQUUE and $\boldsymbol{\Sigma}_{\hat{\xi}0} = (1/n)(\mathcal{A}'\boldsymbol{\Sigma}_0^{-1}\mathcal{A})^{-1}$.
Since our sample size is relatively small we have to use the exact distribution of $-2\log\Lambda_2$, whose upper 5 and 1 percentage points have been provided by *Muirhead* (1982, p.371).

4.3.6 The general linear hypothesis test with growth curve model for eigenspace parameters

Consider a $p \times n$ matrix of observations whose columns follow independent *p-variate Gauss-Laplace multivariate normal distributions* with the same unknown covariance matrix. Each column may represent an individual observation, each row a time when observations were taken. The traditional *special multivariate Gauss-Markov model* $\mathbf{Y} = \mathbf{A}\boldsymbol{\xi} + \mathbf{e}$ is not adequate for dealing with polynomial trends in time. The more general *growth curve* model, introduced by *Potthof and Roy* (1964), may be written as

$$\mathbf{Y} = \mathbf{A}\boldsymbol{\Xi}\mathbf{B} + \mathbf{E} \tag{4.43}$$

where $\mathbf{A}$ is a known $p \times q$ non-random matrix of full rank $q \leq p$; $\boldsymbol{\Xi}$, a $q \times r$ matrix of unknown parameters, $\mathbf{B}$ a $r \times n$ design matrix of rank $r \leq n$; $\mathbf{E}$ denotes a random error matrix, the columns being independently distributed $\mathcal{N}_p(\mathbf{0}, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma}$ is positive-definite. *Khatri* (1966) obtained the maximum likelihood estimate of $\boldsymbol{\Xi}$ in the form

$$\hat{\boldsymbol{\Xi}} = (\mathbf{A}'\boldsymbol{\Omega}^{-1}\mathbf{A})^{-1}\mathbf{A}'\boldsymbol{\Omega}^{-1}\mathbf{Y}\mathbf{B}'(\mathbf{B}\mathbf{B}')^{-1}, \tag{4.44}$$

where

$$\boldsymbol{\Omega} = \mathbf{Y}(\mathbf{I}_n - \mathbf{B}'(\mathbf{B}\mathbf{B}')^{-1}\mathbf{B})\mathbf{Y}', \tag{4.45}$$

is the usual error sum of squares and products matrix, which could be considered proportional to an unbiased estimate of $\boldsymbol{\Sigma}$. For testing the general hypothesis

$$\mathcal{H}_0: \ \mathbf{P}\boldsymbol{\Xi}\mathbf{Q} = \mathbf{0} \quad \textit{versus} \quad \mathcal{H}_{01}: \ \mathbf{P}\boldsymbol{\Xi}\mathbf{Q} \neq \mathbf{0} \tag{4.46}$$

where the $c \times q$ matrix $\mathbf{P}$ has rank $c \le q$, while $\mathbf{Q}$ has dimensions $r \times g$ and rank $g \le r$. The test consists of a multivariate analysis of variance based on the error and hypothesis matrices (*Morrison* 1976)

$$\begin{aligned} \mathbf{V}_{\mathrm{H}} &= (\mathbf{P\Xi Q})(\mathbf{Q'RQ})^{-1}(\mathbf{P\Xi Q})', \\ \mathbf{V}_{\mathrm{E}} &= \mathbf{P}(\mathbf{A'\Omega}^{-1}\mathbf{A})^{-1}\mathbf{P}', \end{aligned} \tag{4.47}$$

where

$$\mathbf{R} = (\mathbf{BB}')^{-1} + (\mathbf{BB}')^{-1}\mathbf{BY'\Omega}^{-1}\mathbf{YB}'(\mathbf{BB}')^{-1} - \mathbf{\Xi}'(\mathbf{A'\Omega}^{-1}\mathbf{A})\mathbf{\Xi} \ . \tag{4.48}$$

Four tests of the hypothesis (4.46) under the *growth curve* model (4.43) could be applied (*Potthof and Roy* 1964, *Siotani, M. et al.* 1985): (1) *Roy's largest root test* with test statistic $\theta_s = \lambda_{\max}/(1+\lambda_{\max})$, where $\lambda_{\max}$ is determined by $|\mathbf{V}_{\mathrm{H}} - \lambda \mathbf{V}_{\mathrm{E}}| = 0$; (2) *Lawley-Hotelling's trace test* with test statistic $T_0^2 = \mathrm{tr}(\mathbf{V}_{\mathrm{H}}\mathbf{V}_{\mathrm{E}}^{-1})$; (3) *Wilks' likelihood ratio tes*t with test statistic $\Lambda = |\mathbf{V}_{\mathrm{H}}| / |\mathbf{V}_{\mathrm{H}} + \mathbf{V}_{\mathrm{E}}|$; and (4) *Bartlett-Nanda-Pillai's trace test* with test statistic $V = \mathrm{tr}\,\mathbf{V}_{\mathrm{H}}(\mathbf{V}_{\mathrm{H}} + \mathbf{V}_{\mathrm{E}})^{-1}$. However, *Roy's test* has the advantage that the distribution of the test statistic under the null hypothesis $\mathcal{H}_0$: $\mathbf{P\Xi Q} = \mathbf{0}$ is known exactly, and has been tabulated (see *Heck* 1960, *Pillai* 1960 and *Kres* 1983); also, associated confidence bounds are available only for *Roy's test*.

To test $\mathcal{H}_0$: $\mathbf{P\Xi Q} = \mathbf{0}$ we calculate the greatest eigenvalues $\lambda_{\max}$ of $\mathbf{V}_{\mathrm{H}}\mathbf{V}_{\mathrm{E}}^{-1}$ and refer $\theta_s = \lambda_{\max}/(1+\lambda_{\max})$ to the approximated *Heck* chart and *Pillai* table with parameters

$$s^* = \min\{c, g\}, \ m^* = \frac{|c-g|-1}{2}, \ n^* = \frac{n-r-p+q-c-1}{2} \ . \tag{4.49}$$

The $100(1-\alpha)$ simultaneous confidence intervals on all bi-linear components $\mathbf{a'P\Xi Qb}$ are given by (*Morrison* 1976)

$$\mathbf{a'P\hat{\Xi}Qb} - [\frac{x_\alpha}{1+x_\alpha}(\mathbf{a'V}_{\mathrm{E}}\mathbf{a})(\mathbf{b'Q'RQb})]^{1/2} \le \mathbf{a'P\Xi Qb} \le \mathbf{a'P\hat{\Xi}Qb} + [\frac{x_\alpha}{1+x_\alpha}(\mathbf{a'V}_{\mathrm{E}}\mathbf{a})(\mathbf{b'Q'RQb})]^{1/2} \tag{4.50}$$

where $x_\alpha \equiv x_{\alpha; s^*, m^*, n^*}$ is the 100α percent Heck or Pillai critical value. If $s^* = 1$, $x_\alpha/(1+x_\alpha)$ should be replaced by the critical value $[(m^*+1)/(n^*+1)]F_{\alpha; 2m^*+2, 2n^*+2}$.

It is worth mentioning that the special linearized multivariate *Gauss-Markov* model for sampling the eigenspace synthesis (4.25) is also a growth curve model corresponding with $\mathscr{A} = \mathbf{A}$, $\mathbf{1}' = \mathbf{B}$ and $\boldsymbol{\xi} = \mathbf{\Xi}$. This fact suggested that the hypothesis (4.46) under the *growth curve* model can be applied to the testing for the estimates of eigenspace parameters directly.

Chapter 5

Statistical inference of the eigenspace components of a three-dimensional, symmetric rank-two random tensor

In Chapter 4 we have achieved the complete solution to the statistical inference of *eigenspace components* of a two-dimensional random tensor. The models are closed and practical. In this chapter we will develop continually this solution for the three-dimensional case. In reality, crustal motions and deformations are of three-dimensional nature and most deformation tensor derived from geodetic, geological and seismological observations are three-dimensional, such as the seismic moment tensors. In the last two decades some efforts have been made to formulate the problem in the three-dimensional space. A curvilinear three-dimensional finite element method has been introduced by *Grafarend* (1986) for the representation of local strain and local rotation tensors in terms of ellipsoidal, *Gauss-Krüger* or *UTM* coordinates. More papers about the three-dimensional strain and strain rate tensor analysis in geodesy are those by *Brunner* (1979), *Lichtenegger and Sünkel* (1989), *Dermanis and Grafarend* (1993) and *Wittenburg*(1999).

The random principal eigenvalues and random eigenvector parameters are of special importance for the prediction of seismic activity. In recent years *Xu* (1999a) and *Kagan* (2000) developed the general distribution of the eigenspace components of the three-dimensional symmetric random tensor of second order, which can hardly be applied directly to real life engineering and Earth science problems, because an exact distribution theory of eigenspace components is almost always unavailable. This reason gives rise herewith to the subject of eigenspace components of a three-dimensional, rank-two symmetric random tensor on the basis of a linearized multivariate *Gauss-Markov* model, which will provide the statistical properties of these eigenspace components. With them we can continue performing the hypothesis tests about the deformation measures. On the assumption that a strain tensor or stress tensor has been directly measured or derived from other observations, such a three-dimensional, symmetric random tensor of second order is a random tensor $\mathbf{T}$ which we assume to be a realization of the tensor-valued Gauss normal distribution over $\mathbb{R}^{3\times 3}$ with independently, identically distributed (i.i.d.) tensor-valued observations, but with identical off-diagonal elements. Since the *eigenspace synthesis* of a symmetric random tensor is nonlinear in terms of the tensor-valued observations, the respective parameters have to be estimated within a special nonlinear multivariate *Gauss-Markov* model.

In this chapter, *first*, based on the review and choice of orthogonal similarity transformation matrices, the *eigenspace analysis and synthesis* of a three-dimensional symmetric random matrix are established uniquely. *Second,* the nonlinear function that relates the tensor elements to the eigenspace components is linearized with respect to a *special nonlinear multivariate Gauss-Markov model*, which enables the *BLUUE of the eigenspace elements* and *BIQUUE* of its variance-covariance matrix, as developed in *Chapter* 4.2 to be successfully applied in the three-dimensional case. Third, the test statistics, such as *Hotelling's* T^2 and *likelihood ratio statistics*, are generated. Hypothesis tests for the random tensor sample means as well as its one variance component are used in the case study of validating a given three-dimensional random strain rate tensor in Chapter 6.

5.1 The eigenspace analysis versus eigenspace synthesis of a three-dimensional, symmetric rank-two random tensor

Let there be given a symmetric three-dimensional rank-two random strain tensor $\mathbf{T}\in\mathbb{T}^2_0$ which is represented in a commutative left or right orthonormal basis $\{\mathbf{e}^1,\mathbf{e}^2,\mathbf{e}^3\}$, in short $\mathbf{e}^i\otimes\mathbf{e}^j$ for all $i,j\in\{1,2,3\}$. "$\otimes$" denotes the tensor product. According to (5.1), $[t_{ij}]\in\mathbb{R}^{3\times 3}$ is called the *matrix representation* of the 3-D rank-two tensor. t_{ij} for all $i,j\in\{1,2,3\}$, establishes the covariant coordinates of the rank-two tensor $\mathbf{T}$. The matrix, due to $t_{ij}=t_{ji}$ is symmetric and of full rank three. $\mathbf{T}'$ denotes the transpose of $\mathbf{T}$ in (5.2). By means of an orthonormal matrix $\mathbf{U}\in\mathbb{SO}(3):=\{\mathbf{U}\in\mathbb{R}^{3\times 3}\mid \mathbf{U}'\mathbf{U}=\mathbf{I}_3, |\mathbf{U}|=+1\}$ the symmetric matrix $\mathbf{T}\in$ SYM:=$\{\mathbf{T}\in\mathbb{R}^{3\times 3}\mid\mathbf{T}'=\mathbf{T}\}$ can be transformed into the canonical form $\mathbf{\Lambda}=\mathrm{Diag}\{\lambda_1,\lambda_2,\lambda_3\}$, also called "*spectral form*".

$$\mathbf{T}=\sum_{i,j=1}^{3}\mathbf{e}^i\otimes\mathbf{e}^j t_{ij}=\sum_{i,j=1}^{3}t_{ij}\mathbf{e}^i\otimes\mathbf{e}^j \tag{5.1}$$

$$\mathbf{T}=[t_{ij}]=\begin{bmatrix} t_{11} & t_{12} & t_{13}\\ t_{21} & t_{22} & t_{23}\\ t_{31} & t_{32} & t_{33}\end{bmatrix}=[t_{ji}]=\mathbf{T}' \tag{5.2}$$

$$\mathbf{U}: \ \mathbf{T} \mapsto \mathbf{\Lambda} = \mathrm{Diag}(\lambda_1, \lambda_2, \lambda_3) = \mathbf{U}'\mathbf{T}\mathbf{U} \tag{5.3}$$

$$\Leftrightarrow$$

$$\mathbf{TU} - \mathbf{\Lambda U} = 0 \ \text{ subject to } \ \mathbf{U}'\mathbf{U} = \mathbf{I}_3 \tag{5.4}$$

$$\Leftrightarrow$$

$$(\mathbf{T} - \lambda_i \mathbf{I}_3)\mathbf{u}_i = 0 \ \text{ for } i \in \{1, 2, 3\} \text{ subject to} \tag{5.5}$$

$$\begin{aligned} &\mathbf{u}_1'\mathbf{u}_1 = 1, \ \mathbf{u}_2'\mathbf{u}_2 = 1, \ \mathbf{u}_3'\mathbf{u}_3 = 1 \quad : \quad \mathbf{u}_1 \in \mathbb{S}^1, \ \mathbf{u}_2 \in \mathbb{S}^1, \ \mathbf{u}_3 \in \mathbb{S}^1 \\ &\mathbf{u}_1'\mathbf{u}_2 = \mathbf{u}_1'\mathbf{u}_3 = \mathbf{u}_2'\mathbf{u}_3 = 0 \qquad : \quad \mathbf{u}_1 \perp \mathbf{u}_2, \ \mathbf{u}_1 \perp \mathbf{u}_3, \ \mathbf{u}_2 \perp \mathbf{u}_3 \end{aligned} \tag{5.6}$$

where

$$\mathbf{U} = \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ u_{21} & u_{22} & u_{23} \\ u_{31} & u_{32} & u_{33} \end{bmatrix} = [\mathbf{u}_1, \ \mathbf{u}_2, \ \mathbf{u}_3]. \tag{5.7}$$

There are many methods to determine the orthonormal matrix **U** for the spectral decomposition or the eigenvalue-eigenvector synthesis (5.3) of the three-dimensional, rank-two strain tensor. Most of them are constructed through three successive rotations, which will be discussed herewith.

5.1.1 The choice of orthogonal similarity transformation matrices

5.1.1.1 Euler angles

Rotation and transformation with *Euler* angles are the commonly used method.

$$\mathbf{R}_z(\chi) = \begin{bmatrix} \cos\chi & \sin\chi & 0 \\ -\sin\chi & \cos\chi & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{5.8}$$

$$\mathbf{R}_y(\varepsilon) = \begin{bmatrix} \cos\varepsilon & 0 & -\sin\varepsilon \\ 0 & 1 & 0 \\ \sin\varepsilon & 0 & \cos\varepsilon \end{bmatrix} \tag{5.9}$$

$$\mathbf{R}_z(\theta) = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{5.10}$$

The total rotation is described by the triple matrix product.

$$\mathbf{A}(\chi, \varepsilon, \theta) = \mathbf{R}_z(\theta)\mathbf{R}_y(\varepsilon)\mathbf{R}_z(\chi) \tag{5.11}$$

Since **A** is an orthogonal matrix, the transformation **y=Ax** is an orthogonal transformation (Rotation). The column and row vectors of **A** are orthonormal, that is, when we represent **A** in form

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = [\mathbf{a}_1 \ \mathbf{a}_2 \ \mathbf{a}_3] \tag{5.12}$$

$$\mathbf{a}_j' \bullet \mathbf{a}_k = \mathbf{a}_k' \bullet \mathbf{a}_j = \begin{cases} 0 & if \ \ j \neq k \\ 1 & if \ \ j = k, where \ \ j, k = 1, \ 2, \ 3. \end{cases} \tag{5.13}$$

Note the order : $\mathbf{R}_z(\chi)$ operates first, then $\mathbf{R}_y(\varepsilon)$ and finally $\mathbf{R}_z(\theta)$. Direct multiplication gives

$$\mathbf{A}(\chi, \varepsilon, \theta) = \begin{bmatrix} \cos\chi\cos\varepsilon\cos\theta - \sin\chi\sin\theta & \sin\chi\cos\varepsilon\cos\theta + \cos\chi\sin\theta & -\sin\varepsilon\cos\theta \\ -\cos\chi\cos\varepsilon\sin\theta - \sin\chi\cos\theta & -\sin\chi\cos\varepsilon\sin\theta + \cos\chi\cos\theta & \sin\varepsilon\sin\theta \\ \cos\chi\sin\varepsilon & \sin\chi\sin\varepsilon & \cos\varepsilon \end{bmatrix} \tag{5.14}$$

Equation $\mathbf{A}(a_{ij})$ with $\mathbf{A}(\chi, \varepsilon, \theta)$, element by element, yields the direction cosines in terms of the three Euler angles.

The *Euler angles* lose their uniqueness for $\varepsilon = 0$; χ and θ are then undetermined. In order to avoid this non-uniqueness *Cardan angles* have been introduced.

5.1.1.2 Cardan angles

Cardan angles α, β, γ , related to the xyz-axis, are uniquely determined as follows:

$$\mathbf{R}_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha \\ 0 & -\sin\alpha & \cos\alpha \end{bmatrix} \tag{5.15}$$

$$\mathbf{R}_y(\beta) = \begin{bmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{bmatrix} \tag{5.16}$$

$$\mathbf{R}_z(\gamma) = \begin{bmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{5.17}$$

$$\mathbf{R}{=}\mathbf{R}_x(\alpha)\mathbf{R}_y(\beta)\mathbf{R}_z(\gamma) = \begin{bmatrix} \cos\beta\cos\gamma & \cos\beta\sin\gamma & -\sin\beta \\ \sin\alpha\sin\beta\cos\gamma - \cos\alpha\sin\gamma & \sin\alpha\sin\beta\sin\gamma + \cos\alpha\cos\gamma & \sin\alpha\cos\beta \\ \cos\alpha\sin\beta\cos\gamma + \sin\alpha\sin\gamma & \cos\alpha\sin\beta\sin\gamma - \sin\alpha\cos\gamma & \cos\alpha\cos\beta \end{bmatrix} \tag{5.18}$$

which consists of successive rotations by: γ about the *z*-axis, β about the new *y*-axis, α about the new *x*-axis. The order of rotations is a matter of convention and the one used here is known as the *xyz* convention. The main reason for the popularity of this *xyz* convention is that it does successive rotations about three different axes.

The meaning of **R** is that any vector **x** given with respect to axes fixed which are in space, is then represented by **Rx** with respect to the rotated axes. Essentially the elements of $\mathbf{R}'$, therefore, give the directional cosines of the rotated axes relative to the fixed axes.

Since **R** is an orthogonal matrix, the transformation **y=Rx** is an orthogonal transformation (rotation). The column and row vectors of **R** are orthonormal, that is, when we represent **R** in the form

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} = [\mathbf{r}_1 \;\; \mathbf{r}_2 \;\; \mathbf{r}_3] \tag{5.19}$$

$$\mathbf{r}_j' \bullet \mathbf{r}_k = \mathbf{r}_k' \bullet \mathbf{r}_j = \begin{cases} 0 & if \;\; j \neq k \\ 1 & if \;\; j = k, where \;\; j,k = 1,\, 2,\, 3. \end{cases} \tag{5.20}$$

The Cardan angles can be obtained from the given rotation matrix $\mathbf{R} \in \mathbb{R}^{3\times 3}$.

$$\text{if } r_{11} = 0 \text{ and } r_{12} = 0, \text{ or equivalently } \cos\beta = 0, \text{ then } \begin{cases} \alpha := 0 \\ \beta = \dfrac{\pi}{2} \\ \gamma = \arctan(\dfrac{r_{32}}{r_{22}}) \end{cases},$$

$$\text{otherwises } \begin{cases} \alpha = \arctan(\dfrac{r_{23}}{r_{33}}) \\ \beta = \arctan(\dfrac{-r_{13}}{\sqrt{r_{11}^2 + r_{12}^2}}) \text{ or } \beta = \arctan(\dfrac{-r_{13}}{\sqrt{r_{23}^2 + r_{33}^2}}) \\ \gamma = \arctan(\dfrac{r_{12}}{r_{11}}) \end{cases} \tag{5.21}$$

The relationship between the three Euler angles and the three Cardan angles has been given in Grafarend (1982).

5.1.1.3 The alternative choice

The simplest way to define three orthonormed directions, i.e., an orthonormed basis of vectors in 3-D is to define the three 3-D rotations that connect the given basis with the natural basis {1, 0, 0}, {0, 1, 0} and {0, 0, 1}. These could, for instance, be the three angles, but these angles do not generalize up to higher dimensions. Instead, we choose three following rotations (*Xu* 1999b, *Tarantola, et al.* 2000). These rotation matrices are also called *Givens matrices* and the operation of going from **x** to **Ux** is called a *Givens transformation* (*Searle* 1982, p.72).

The *Givens matrices* with the angle θ_{32} θ_{31} θ_{21} related to the x, y and z-axes are

$$\mathbf{U}_{32}(\theta_{32}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_{32} & \sin\theta_{32} \\ 0 & -\sin\theta_{32} & \cos\theta_{32} \end{bmatrix} \tag{5.22}$$

$$\mathbf{U}_{31}(\theta_{31}) = \begin{bmatrix} \cos\theta_{31} & 0 & \sin\theta_{31} \\ 0 & 1 & 0 \\ -\sin\theta_{31} & 0 & \cos\theta_{31} \end{bmatrix} \tag{5.23}$$

$$\mathbf{U}_{21}(\theta_{21}) = \begin{bmatrix} \cos\theta_{21} & \sin\theta_{21} & 0 \\ -\sin\theta_{21} & \cos\theta_{21} & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{5.24}$$

$$\mathbf{U}=\mathbf{U}_{32}(\theta_{32})\mathbf{U}_{31}(\theta_{31})\mathbf{U}_{21}(\theta_{21}) = \begin{bmatrix} \cos\theta_{31}\cos\theta_{21} & \cos\theta_{31}\sin\theta_{21} & \sin\theta_{31} \\ -\sin\theta_{32}\sin\theta_{31}\cos\theta_{21} - \cos\theta_{32}\sin\theta_{21} & -\sin\theta_{32}\sin\theta_{31}\sin\theta_{21} + \cos\theta_{32}\cos\theta_{21} & \sin\theta_{32}\cos\theta_{31} \\ -\cos\theta_{32}\sin\theta_{31}\cos\theta_{21} + \sin\theta_{32}\sin\theta_{21} & -\cos\theta_{32}\sin\theta_{31}\sin\theta_{21} - \sin\theta_{32}\cos\theta_{21} & \cos\theta_{32}\cos\theta_{31} \end{bmatrix} \tag{5.25}$$

which consists of successive rotations by: θ_{21} about the z-axis, $-\theta_{31}$ about the new y-axis, θ_{32} about the new x-axis and is presented in *Figure* 5.1. The order of rotations is matter of convention and the one used here is known as the *xyz* convention. The main reason for the popularity of this *xyz* convention is that it does successive rotations about three different axes.

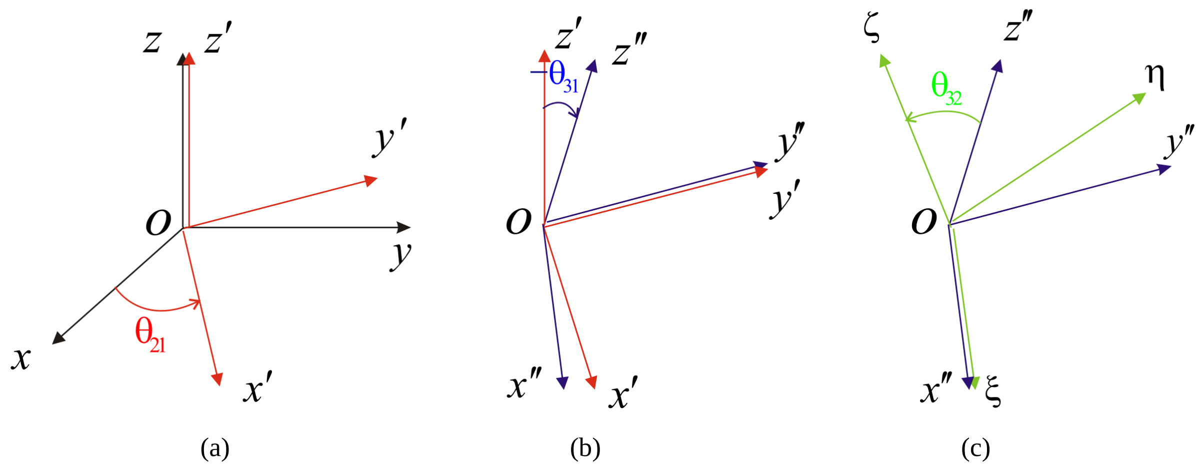


Figure 5.1 (a) Rotation about z through angle θ_{21}; (b) Rotation about y' through angle $-\theta_{31}$; (c) Rotation about x'' through angle θ_{32}.

The meaning of **U** is that any vector **x**, given with respect to axes which are fixed in space is again represented by **Ux** with respect to the rotated axes. As above the elements of $\mathbf{U}'$, therefore, give the direction cosines of the rotated axes relative to the fixed axes.

Since **U** is an orthogonal matrix, the transformation **y=Ux** is an orthogonal transformation (rotation). The column and row vectors of **U** are orthonormal, that is, when we represent **U** in form

$$\mathbf{U} = \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ u_{21} & u_{22} & u_{23} \\ u_{31} & u_{32} & u_{33} \end{bmatrix} = [\mathbf{u}_1 \ \mathbf{u}_2 \ \mathbf{u}_3] \tag{5.26}$$

$$\mathbf{u}_j' \bullet \mathbf{u}_k = \mathbf{u}_k' \bullet \mathbf{u}_j = \begin{cases} 0 & if \ j \neq k \\ 1 & if \ j = k, where \ j,k = 1,\, 2,\, 3. \end{cases} \tag{5.27}$$

The rotation angles can be determined from the given rotation matrix $\mathbf{U} \in \mathbb{R}^{3\times 3}$

$$\text{if } u_{11} = 0 \text{ and } u_{12} = 0 \text{ or equivalently } \cos\theta_{31} = 0, \text{ then } \begin{cases} \theta_{32} := 0 \\ \theta_{31} = \dfrac{\pi}{2} \\ \theta_{21} = \arctan(\dfrac{-u_{32}}{u_{22}}) \end{cases},$$

$$\text{otherwise } \begin{cases} \theta_{32} = \arctan(\dfrac{u_{23}}{u_{33}}) \\ \theta_{31} = \arctan(\dfrac{u_{13}}{\sqrt{u_{11}^2 + u_{12}^2}}) \text{ or } \theta_{31} = \arctan(\dfrac{u_{13}}{\sqrt{u_{23}^2 + u_{33}^2}}) \\ \theta_{21} = \arctan(\dfrac{u_{12}}{u_{11}}) \, . \end{cases} \tag{5.28}$$

In order that the spectral decomposition is unique, the three angles θ_{32}, θ_{31} and θ_{21} are all defined between $-\pi/2$ and $\pi/2$. Thus the two element u_{11} and u_{33} of the orthogonal matrix **U** should be positive.

With (5.22)~(5.28) we can establish the relationship of the strain tensor with its eigenvalues and eigendirections uniquely.

All the three slightly different representations of the orthogonal matrix **U** are mathematically equivalent. Since it is not convenient to generalize the *Euler* (5.8) ~ (5.14) and *Cardan* (5.15) ~ (5.18) representations to the *n*-dimensional case, and in order to take the advantage of *Givens* representation (5.22) ~ (5.25) that does successive rotations about three different axes, we will confine ourselves to the *Givens representation* in the study of a three-dimensional rank-two random tensor.

5.1.2 The eigenspace analysis versus eigenspace synthesis of a three-dimensional, symmetric rank-two random tensor

The formulae (5.3) ~ (5.7) and (5.25) ~ (5.28) establish the *eigenspace analysis*. The diagonal matrix $\mathbf{\Lambda}$ contains the *eigenvalues* $\lambda_1, \lambda_2, \lambda_3,$ the orthonormal matrix **U** the *eigencolumns*, also called coordinates of the *eigenvectors*, namely $[u_{11}, u_{21}, u_{31}]'$, $[u_{12}, u_{22}, u_{32}]'$ and $[u_{13}, u_{23}, u_{33}]'$. Since $\mathbf{U} \in \mathbb{R}^{3\times 3}$ is an orthonormal matrix, it is constructed by the *trigonometric representation* with three rotational angular parameters. These angular parameters establish the *eigenorientation*, namely the orientation of the eigendirections. The solution of the eigenvalue-eigencolumn equation is not unique, as it is generated by the quadratic equations (5.6). If we assume that the elements u_{11}, and u_{33} of the eigencolumns *have to be positive* when the three angles θ_{32}, θ_{31} and θ_{21} are all defined between $-\pi/2$ and $\pi/2$. we arrive at (5.30), (5.32) and (5.33), respectively. Note that we have defined the angular parameters θ_{32}, θ_{31} and θ_{21} in a half *open domain* in order to avoid any singularity.

***Corollary* 5.1 (eigenvalue-eigenvector analysis)**

For a symmetric tensor $\mathbf{T} \in \mathbb{R}^{3\times 3}$ the eigenvalues λ_1, λ_2 and λ_3 as well as the rotational parameters θ_{32}, θ_{31} and θ_{21}, which constitutes the orthonormal matrix $\mathbf{U} \in \mathbb{R}^{3\times 3}$ of eigenvectors are analytically represented by

the characteristic equation

$$|\mathbf{T} - \lambda \mathbf{I}_3| = 0. \tag{5.29}$$

which is a cubic in λ, namely:

$$-\lambda^3 + I_\lambda \lambda^2 - II_\lambda \lambda + III_\lambda = 0 \tag{5.30}$$

where

$$\begin{aligned} I_\lambda &= t_{11} + t_{22} + t_{33} \\ II_\lambda &= t_{22}t_{33} + t_{33}t_{11} + t_{11}t_{22} - t_{23}^2 - t_{13}^2 - t_{12}^2 \\ III_\lambda &= t_{11}t_{22}t_{33} + 2t_{12}t_{23}t_{13} - t_{11}t_{23}^2 - t_{22}t_{13}^2 - t_{33}t_{12}^2 \end{aligned}$$

The three roots $\lambda_1, \lambda_2, \lambda_3$ are called principal components (eigenvalues) with

$$\begin{aligned} I_\lambda &= \lambda_1 + \lambda_2 + \lambda_3 \\ II_\lambda &= \lambda_2\lambda_3 + \lambda_3\lambda_1 + \lambda_1\lambda_2 \\ III_\lambda &= \lambda_1\lambda_2\lambda_3 \end{aligned} \tag{5.31}$$

The related eigenvectors $\mathbf{U}_i \ (i = 1,2,3)$ are solutions of the homogeneous equations:

$$(\mathbf{T} - \lambda_i \mathbf{I}_3)\mathbf{U}_i = \mathbf{0} \tag{5.32}$$

The rotation angles can be determined from the given rotations matrix $\mathbf{U} \in \mathbb{R}^{3\times3}$ (5.25):

$$\mathbf{U}{=}\mathbf{U}_{32}(\theta_{32})\mathbf{U}_{31}(\theta_{31})\mathbf{U}_{21}(\theta_{21}) = \begin{bmatrix} \cos\theta_{31}\cos\theta_{21} & \cos\theta_{31}\sin\theta_{21} & \sin\theta_{31} \\ -\sin\theta_{32}\sin\theta_{31}\cos\theta_{21} - \cos\theta_{32}\sin\theta_{21} & -\sin\theta_{32}\sin\theta_{31}\sin\theta_{21} + \cos\theta_{32}\cos\theta_{21} & \sin\theta_{32}\cos\theta_{31} \\ -\cos\theta_{32}\sin\theta_{31}\cos\theta_{21} + \sin\theta_{32}\sin\theta_{21} & -\cos\theta_{32}\sin\theta_{31}\sin\theta_{21} - \sin\theta_{32}\cos\theta_{21} & \cos\theta_{32}\cos\theta_{31} \end{bmatrix}$$

following (5.28):

$$\text{if } u_{11} = 0 \text{ and } u_{12} = 0 \text{ or equivalently } \cos\theta_{31} = 0, \text{ then } \begin{cases} \theta_{32} := 0 \\ \theta_{31} = \dfrac{\pi}{2} \\ \theta_{21} = \arctan(\dfrac{-u_{32}}{u_{22}}) \end{cases},$$

$$\text{otherwise } \begin{cases} \theta_{32} = \arctan(\dfrac{u_{23}}{u_{33}}) \\ \theta_{31} = \arctan(\dfrac{u_{13}}{\sqrt{u_{11}^2 + u_{12}^2}}) \text{ or } \theta_{31} = \arctan(\dfrac{u_{13}}{\sqrt{u_{23}^2 + u_{33}^2}}) \\ \theta_{21} = \arctan(\dfrac{u_{12}}{u_{11}}) \ . \end{cases}$$

In order that the spectral decomposition is unique, the three angles θ_{32}, θ_{31} and θ_{21} are all $\in \ ]-\pi/2, \pi/2]$. Thus, the two element u_{11} and u_{33} of the orthogonal matrix **U** should be positive.

***Corollary* 5.2 (eigenvalue-eigenvector synthesis)**

Given the eigenvalues λ_1, λ_2 and λ_3 as well as the rotational parameters θ_{32}, θ_{31} and θ_{21}, which constitute the orthonormal matrix $\mathbf{U} \in \mathbb{R}^{3\times3}$ of eigencolumns, the symmetric tensor $\mathbf{T} \in \mathbb{R}^{3\times3}$ is synthetically represented by

$$\begin{aligned}
t_{11} &= \lambda_1 \cos^2\theta_{31}\cos^2\theta_{21} + \lambda_2\cos^2\theta_{31}\sin^2\theta_{21} + \lambda_3\sin^2\theta_{31} \\
t_{12} &= \lambda_1\cos\theta_{31}\cos\theta_{21}(-\sin\theta_{32}\sin\theta_{31}\cos\theta_{21} - \cos\theta_{32}\sin\theta_{21}) + \\
&\quad + \lambda_2\cos\theta_{31}\sin\theta_{21}(-\sin\theta_{32}\sin\theta_{31}\sin\theta_{21} - \cos\theta_{32}\cos\theta_{21}) + \\
&\quad + \lambda_3\sin\theta_{32}\sin\theta_{31}\cos\theta_{31} \\
t_{13} &= \lambda_1\cos\theta_{31}\cos\theta_{21}(-\cos\theta_{32}\sin\theta_{31}\cos\theta_{21} + \sin\theta_{32}\sin\theta_{21}) + \\
&\quad + \lambda_2\cos\theta_{31}\sin\theta_{21}(-\cos\theta_{32}\sin\theta_{31}\sin\theta_{21} - \sin\theta_{32}\cos\theta_{21}) + \\
&\quad + \lambda_3\cos\theta_{32}\sin\theta_{31}\cos\theta_{31} \\
t_{22} &= \lambda_1(-\sin\theta_{32}\sin\theta_{31}\cos\theta_{21} - \cos\theta_{32}\sin\theta_{21})^2 + \lambda_2(-\sin\theta_{32}\sin\theta_{31}\sin\theta_{21} + \cos\theta_{32}\cos\theta_{21})^2 + \\
&\quad + \lambda_3\sin^2\theta_{32}\cos^2\theta_{31} \\
t_{23} &= \lambda_1(-\sin\theta_{32}\sin\theta_{31}\cos\theta_{21} - \cos\theta_{32}\sin\theta_{21})(-\cos\theta_{32}\sin\theta_{31}\cos\theta_{21} + \sin\theta_{32}\sin\theta_{21}) + \\
&\quad + \lambda_2(-\sin\theta_{32}\sin\theta_{31}\sin\theta_{21} + \cos\theta_{32}\cos\theta_{21})(-\cos\theta_{32}\sin\theta_{31}\sin\theta_{21} - \sin\theta_{32}\cos\theta_{21}) + \\
&\quad + \lambda_3\sin\theta_{32}\cos\theta_{32}\cos^2\theta_{31} \\
t_{33} &= \lambda_1(-\cos\theta_{32}\sin\theta_{31}\cos\theta_{21} + \sin\theta_{32}\sin\theta_{21})^2 + \lambda_2(-\cos\theta_{32}\sin\theta_{31}\sin\theta_{21} - \sin\theta_{32}\cos\theta_{21})^2 + \\
&\quad + \lambda_3\cos^2\theta_{32}\cos^2\theta_{31}
\end{aligned} \tag{5.33}$$

On the basis of *Corollary* 5.1 (eigenspace analysis) and *Corollary* 5.2 (eigenspace synthesis) we are able to portray the three-dimensional symmetric strain tensor **T**, which can be visualized as *strain ellipsoid*, if λ_1, λ_2 and λ_3 are all positive, but as the *strain hyperboloid*, if $\operatorname{sign}\lambda_1 \neq \operatorname{sign}\lambda_2$ or $\operatorname{sign}\lambda_1 \neq \operatorname{sign}\lambda_3$ or $\operatorname{sign}\lambda_2 \neq \operatorname{sign}\lambda_3$. *Figure* 5.2 and *Figure* 5.3(a) illustrate the *strain ellipsoid*, *Figure* 5.3 (b) and (c) illustrate the *strain hyperboloid*. In the *first case* (Figure 5.2 or *Figure* 5.3(a)), the axes of the strain ellipsoid are directed along the eigenvectors of the strain tensor; the *semi-major axes* of the strain ellipse are identified with the *maximum principal strain*, *intermediate principal strain* as well as the *minimum principal strain*, constrained by $\operatorname{sign}\lambda_1 = \operatorname{sign}\lambda_2 = \operatorname{sign}\lambda_3 = +1$. Alternatively, in the *second case* (*Figure* 5.3 (b)), the axes of the strain hyperboloids of one sheet are directed along the eigenvectors of the strain tensor, indicated by the "real axes" showing λ_1, λ_2 and the "*imaginary axis*" with $|\lambda_3|$; and in the *third case* (*Figure* 5.3 (c)), the axes of the strain hyperboloids of two sheet are directed along the eigenvectors of the strain tensor, indicated by the "real axis" showing λ_2 and the "*imaginary axes*" with $|\lambda_1|$ and $|\lambda_3|$, for instance. If $\operatorname{sign}\lambda_i = +1$ we speak of *extension*, if $\operatorname{sign}\lambda_i = -1$ of *contraction* instead.

The notations used in describing the three-dimensional strain tensor will be defined in *Box* 5.1.

***Box* 5.1**: (three-dimensional strain tensor)

The representation of three-dimensional strain tensor

Three-dimensional strain tensor components:

t_{11}, t_{22}, t_{33} the normal strain along the 1-, 2- and 3- axis (xyz-axis), respectively;

t_{12}, t_{13}, t_{23} the shear strain between the respective pairs of axes.

The principal components:

λ_1 maximum principal strain, the greatest change of length per unit length;

λ_2 intermediate principal strain, the intermediate change of length per unit length;

λ_3 minimum principal strain, the smallest change of length per unit length;

$\theta_{32}, \theta_{31}, \theta_{21}$, the orientation of the three principal strain axes, respectively.

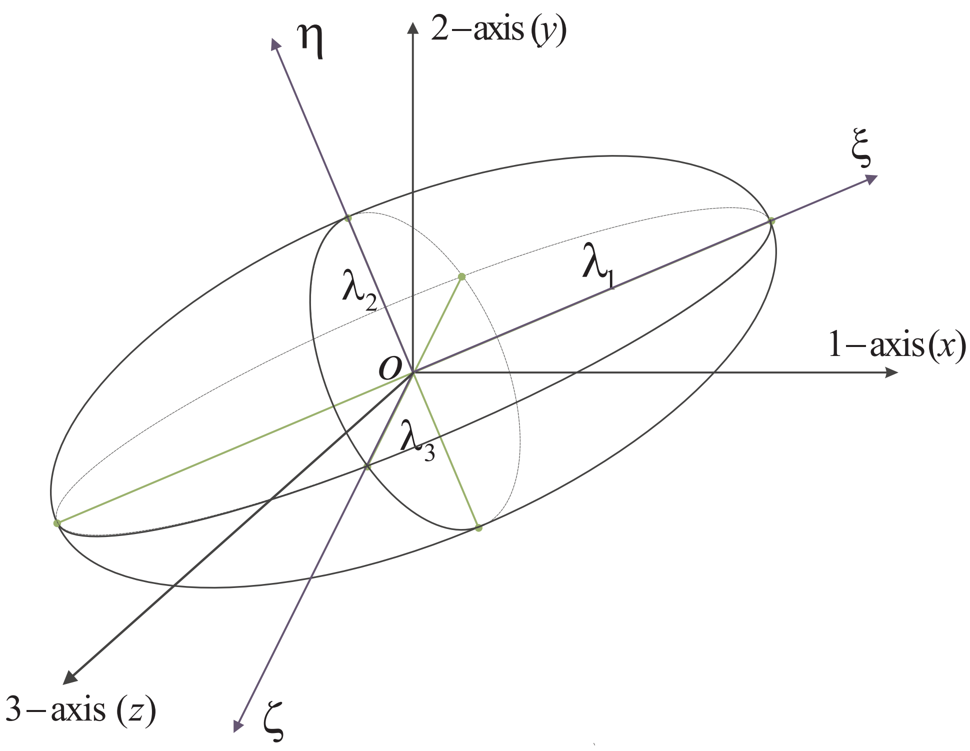


Figure 5.2. The strain ellipsoid of a three-dimensional strain tensor

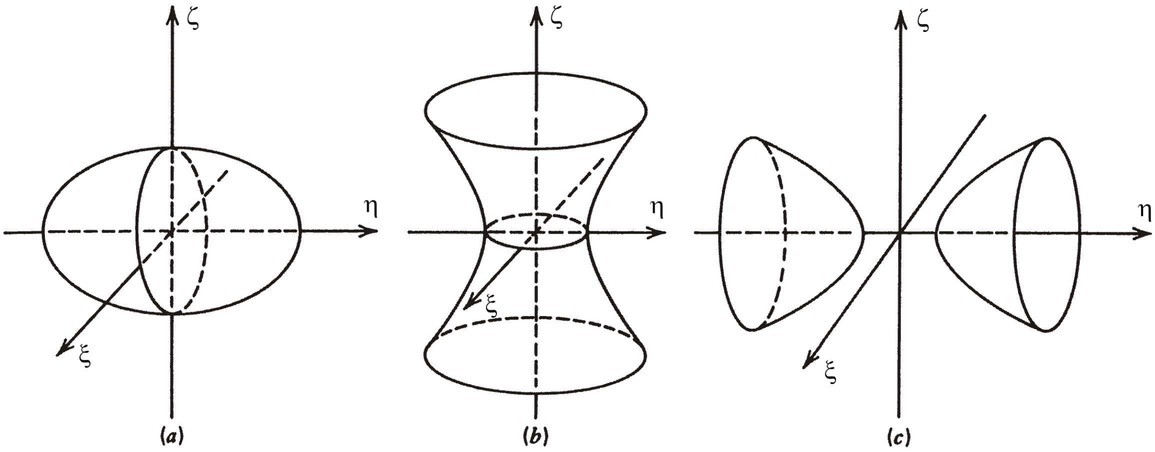


Figure 5.3. The strain ellipsoid and hyperboloid of a three-dimensional strain tensor:
(a) $\lambda_1 > 0$, $\lambda_2 > 0$ and $\lambda_3 > 0$, an ellipsoid;
(b) $\lambda_1 > 0$, $\lambda_2 > 0$ and $\lambda_3 < 0$, a hyperboloid of one sheet; and
(c) $\lambda_1 < 0$, $\lambda_2 > 0$ and $\lambda_3 < 0$, a hyperboloid of two sheet.

5.2 The linearized multivariate *Gauss-Markov* model for the estimation of eigenspace components of a three-dimensional, symmetric rank-two random tensor

Chapter 5.1 has documented that the eigenspace synthesis of a symmetric random tensor is nonlinear in terms of the tensor-valued observations, and there is no simple probability density function of the distribution of random eigenspace components. Accordingly we are unable to derive the exact sampling distribution directly. Here, we will derive the linearized counterpart for sampling the eigenspace synthesis parameters from the originally nonlinear observation equations. The $\boldsymbol{\Sigma}$-BLUUE of eigenspace components and their variance-covariance matrix estimate of type BIQUUE will be developed in accordance with the formulas presented earlier by *J. Cai, E. Grafarend and B. Schaffrin* (2001b). Using the eigenspace analysis *versus* eigenspace synthesis presented in *Chapter* 5.1 and the same notation of *n* observations of **T,** namely $\mathbf{T}_1, \mathbf{T}_2, \ldots, \mathbf{T}_n$ whose related vectorized forms are $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n$, and $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_n] \in \mathbb{R}^{6\times n}$, we can define the nonlinear *Gauss-Markov* model which is presented by (5.34) ~ (5.37) , where **1** denotes the $n\times 1$ "*summation vector*" with all its entries being 1.

Box* 5.2*:

Special nonlinear multivariate *Gauss-Markov* model for sampling the eigenspace synthesis

$$\mathbf{Y} = \mathbf{F}(\xi)\mathbf{1}' + \bar{\mathbf{E}} \tag{5.34}$$

1st moments

$$E\{\begin{bmatrix} y_{1.1} & \cdots & y_{1.n} \\ \vdots & \cdots & \vdots \\ y_{6.1} & \cdots & y_{6.n} \end{bmatrix}\} = \mathbf{F}\mathbf{1}' \quad or \quad E\{\mathbf{Y}\} = \mathbf{F}\mathbf{1}' \tag{5.35}$$

$$\mathbf{F} := \begin{bmatrix} f_1 \\ \vdots \\ f_6 \end{bmatrix} = \begin{bmatrix} f_1(\lambda_1,\lambda_2,\lambda_3,\theta_{32},\theta_{31},\theta_{21}) \\ \vdots \\ f_6(\lambda_1,\lambda_2,\lambda_3,\theta_{32},\theta_{31},\theta_{21}) \end{bmatrix} = \begin{bmatrix} f_1(\xi_1,\xi_2,\xi_3,\xi_4,\xi_5,\xi_6) \\ \vdots \\ f_6(\xi_1,\xi_2,\xi_3,\xi_4,\xi_5,\xi_6) \end{bmatrix}$$

$$\xi_1 := \lambda_1, \xi_2 := \lambda_2, \xi_3 := \lambda_3, \qquad y_1 := t_{11}, y_2 := t_{12}, y_3 := t_{13}$$
$$\xi_4 := \theta_{32}, \xi_5 := \theta_{31}, \xi_6 := \theta_{21} \quad and \quad y_4 := t_{22}, y_5 := t_{23}, y_6 := t_{33}$$

with the eigenvalue-eigenvector synthesis (5.33)

$$\begin{aligned}
f_1(\lambda_1,\lambda_2,\lambda_3,\theta_{32},\theta_{31},\theta_{21}) &= \lambda_1\cos^2\theta_{31}\cos^2\theta_{21}+\lambda_2\cos^2\theta_{31}\sin^2\theta_{21}+\lambda_3\sin^2\theta_{31}\\
f_2(\lambda_1,\lambda_2,\lambda_3,\theta_{32},\theta_{31},\theta_{21}) &= \lambda_1\cos\theta_{31}\cos\theta_{21}(-\sin\theta_{32}\sin\theta_{31}\cos\theta_{21}-\cos\theta_{32}\sin\theta_{21})+\\
&+\lambda_2\cos\theta_{31}\sin\theta_{21}(-\sin\theta_{32}\sin\theta_{31}\sin\theta_{21}-\cos\theta_{32}\cos\theta_{21})+\\
&+\lambda_3\sin\theta_{32}\sin\theta_{31}\cos\theta_{31}\\
f_3(\lambda_1,\lambda_2,\lambda_3,\theta_{32},\theta_{31},\theta_{21}) &= \lambda_1\cos\theta_{31}\cos\theta_{21}(-\cos\theta_{32}\sin\theta_{31}\cos\theta_{21}+\sin\theta_{32}\sin\theta_{21})+\\
&+\lambda_2\cos\theta_{31}\sin\theta_{21}(-\cos\theta_{32}\sin\theta_{31}\sin\theta_{21}-\sin\theta_{32}\cos\theta_{21})+\\
&+\lambda_3\cos\theta_{32}\sin\theta_{31}\cos\theta_{31}\\
f_4(\lambda_1,\lambda_2,\lambda_3,\theta_{32},\theta_{31},\theta_{21}) &= \lambda_1(-\sin\theta_{32}\sin\theta_{31}\cos\theta_{21}-\cos\theta_{32}\sin\theta_{21})^2+\\
&+\lambda_2(-\sin\theta_{32}\sin\theta_{31}\sin\theta_{21}+\cos\theta_{32}\cos\theta_{21})^2+\\
&+\lambda_3\sin^2\theta_{32}\cos^2\theta_{31}\\
f_5(\lambda_1,\lambda_2,\lambda_3,\theta_{32},\theta_{31},\theta_{21}) &= \lambda_1(-\sin\theta_{32}\sin\theta_{31}\cos\theta_{21}-\cos\theta_{32}\sin\theta_{21})(-\cos\theta_{32}\sin\theta_{31}\cos\theta_{21}+\sin\theta_{32}\sin\theta_{21})+\\
&+\lambda_2(-\sin\theta_{32}\sin\theta_{31}\sin\theta_{21}+\cos\theta_{32}\cos\theta_{21})(-\cos\theta_{32}\sin\theta_{31}\sin\theta_{21}-\sin\theta_{32}\cos\theta_{21})+\\
&+\lambda_3\sin\theta_{32}\cos\theta_{32}\cos^2\theta_{31}\\
f_6(\lambda_1,\lambda_2,\lambda_3,\theta_{32},\theta_{31},\theta_{21}) &= \lambda_1(-\cos\theta_{32}\sin\theta_{31}\cos\theta_{21}+\sin\theta_{32}\sin\theta_{21})^2+\\
&+\lambda_2(-\cos\theta_{32}\sin\theta_{31}\sin\theta_{21}-\sin\theta_{32}\cos\theta_{21})^2+\\
&+\lambda_3\cos^2\theta_{32}\cos^2\theta_{31}
\end{aligned}$$

2nd moments

"*independent between observations*", "*i.i.d. observations*"

(5.36)
$$D(\operatorname{vec}\mathbf{Y})=\begin{bmatrix}\boldsymbol{\Sigma}_{y1} & 0 & \cdots & 0\\ 0 & \boldsymbol{\Sigma}_{y2} & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \boldsymbol{\Sigma}_{yn}\end{bmatrix},\quad D(\operatorname{vec}\mathbf{Y})=\begin{bmatrix}\boldsymbol{\Sigma}_{y} & 0 & \cdots & 0\\ 0 & \boldsymbol{\Sigma}_{y} & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \boldsymbol{\Sigma}_{y}\end{bmatrix}=\mathbf{I}_n\otimes\boldsymbol{\Sigma}_y, \tag{5.37}$$

$$D\{\operatorname{vec}\mathbf{Y}\}=\boldsymbol{\Sigma}\in\mathbb{R}^{6n\times 6n},\ \boldsymbol{\Sigma}\text{ positive-definite, rk }\boldsymbol{\Sigma}=6n,$$

$$\boldsymbol{\xi},\ E\{\mathbf{Y}\},\ \mathbf{Y}-E\{\mathbf{Y}\}=\bar{\mathbf{E}}\ \text{unknown},\ \boldsymbol{\Sigma}\text{ unknown (but patterned).}$$

With the same linearization procedure as in (4.21) the linearized observation equations (5.35) are presented in *Box* 5.3 in detail:

Box 5.3:

Linearization of nonlinear observation equation

First set of observation: $\begin{bmatrix}y_{1.1}\\ \vdots\\ y_{6.1}\end{bmatrix}=\begin{bmatrix}t_{11.1}\\ \vdots\\ t_{33.1}\end{bmatrix}$

Define: $\boldsymbol{\xi}_0=[\lambda_{1.1},\lambda_{2.1},\lambda_{3.1},\theta_{32.1},\theta_{31.1},\theta_{21.1}]$

"Linearized nonlinear model"

$$\mathbf{F}(\boldsymbol{\xi})-\mathbf{F}(\boldsymbol{\xi}_0)=\mathbf{J}(\boldsymbol{\xi}_0)\Delta\boldsymbol{\xi}+\mathscr{O}[(\boldsymbol{\xi}-\boldsymbol{\xi}_0)\otimes(\boldsymbol{\xi}-\boldsymbol{\xi}_0)] \tag{5.38}$$

"Jacobi matrix"

$$\mathbf{J}(\boldsymbol{\xi})=\begin{bmatrix}\frac{\partial f_1}{\partial\lambda_1} & \frac{\partial f_1}{\partial\lambda_2} & \frac{\partial f_1}{\partial\lambda_3} & \frac{\partial f_1}{\partial\theta_{32}} & \frac{\partial f_1}{\partial\theta_{31}} & \frac{\partial f_1}{\partial\theta_{21}}\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ \frac{\partial f_6}{\partial\lambda_1} & \frac{\partial f_6}{\partial\lambda_2} & \frac{\partial f_6}{\partial\lambda_3} & \frac{\partial f_6}{\partial\theta_{32}} & \frac{\partial f_6}{\partial\theta_{31}} & \frac{\partial f_6}{\partial\theta_{21}}\end{bmatrix} \tag{5.39}$$

$$\mathscr{A}=\mathbf{J}(\boldsymbol{\xi}_0)=\mathbf{J}(\lambda_{1.1},\lambda_{2.1},\lambda_{3.1},\theta_{32.1},\theta_{31.1},\theta_{21.1}). \tag{5.40}$$

It is important to remark that the absolute value of the Jacobian is:

$$|\det \mathbf{J}| = (\lambda_2 - \lambda_3)(\lambda_1 - \lambda_3)((\lambda_1 - \lambda_2)\cos\theta_{31}, \text{ where } \lambda_1 \geq \lambda_2 \geq \lambda_3 \text{ always,} \tag{5.41}$$

which is useful in derivation of the probability density function of the three-dimensional random tensor spectrum (*Xu*, 1999a, b).

Based upon the *Taylor series expansion* for $\mathbf{F}(\xi)$ we shall again apply the *Gauss-Newton iteration scheme* with $\xi_0 = [\lambda_{1.1}, \lambda_{2.1}, \lambda_{3.1}, \theta_{32.1}, \theta_{31.1}, \theta_{21.1}]$ as the starting point. ξ_0 is determined by solving once the eigenvalue analysis equations as indicated by *Corollary* 5.1 for the *sample* one. In this way, we have established the design matrix of the first kind $\mathcal{A} = \mathbf{J}(\xi_0)$ as the *Jacobi matrix* $\mathbf{J}$ at the point ξ_0. The special linearized multivariate *Gauss-Markov model* for sampling the eigenspace of a three-dimensional, symmetric random matrix is identical with (4.25) ~ (4.30) in the two-dimensional case of Chapter 4.2.

With these definitions and the observations of a random tensor we can *again* estimate the eigenspace components ξ of type Σ - BLUUE (*Best Linear Uniformly Unbiased Estimation*) and $\hat{\Sigma}_{\mathbf{y}}$ as the BIQUUE (*Best Invariant Quadratic Uniformly Unbiased Estimate*) of the variance-covariance matrix $\Sigma_{\mathbf{y}}$, as was summarized in *Theorem* 4.3 and 4.4.

5.3 Hypothesis testing for the estimates of eigenspace components of a three-dimensional, symmetric rank-two random tensor

With the estimates of eigenspace components of a random strain rate tensor and their dispersion matrix the following multivariate hypothesis tests are suggested:

- Test for the eigenspace parameter vector $\xi = \xi_0$ with $\Sigma_{\mathbf{y}}$ unspecified (*Box* 5.4);
- Test for a distinct element of the eigenspace parameter vector with *Student t- test* (*Box* 5.5);
- *Eigen-inference* about the orthonormally transformed parameters $\boldsymbol{\eta}$ (*Box* 5.6);
- Test for the variance-covariance matrix $\Sigma_{\mathbf{y}} = \Sigma_0$ (*Box* 5.7);
- Test for the eigenspace parameter vector and variance-covariance matrix $\xi = \xi_0$, $\Sigma_{\mathbf{y}} = \Sigma_0$ (*Box* 5.8).

***Box* 5.4**:

Multivariate hypothesis test about the eigenspace parameter vector ξ assuming *Gauss-Laplace* normally distributed observations of a three-dimensional, symmetric rank-two random tensor

First Test for $\mathcal{H}_{01}: \ \xi = \xi_0, \ \mathcal{H}_{11}: \ \xi \neq \xi_0$ with $\Sigma_{\mathbf{y}}$ unspecified;

$\Leftrightarrow$

$$\mathcal{H}_{01}: \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \theta_{32} \\ \theta_{31} \\ \theta_{21} \end{bmatrix} = \begin{bmatrix} \lambda_{10} \\ \lambda_{20} \\ \lambda_{30} \\ \theta_{320} \\ \theta_{310} \\ \theta_{210} \end{bmatrix}, \ \mathcal{H}_{11}: \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \theta_{32} \\ \theta_{31} \\ \theta_{21} \end{bmatrix} \neq \begin{bmatrix} \lambda_{10} \\ \lambda_{20} \\ \lambda_{30} \\ \theta_{320} \\ \theta_{310} \\ \theta_{210} \end{bmatrix}, \text{ with } \Sigma_{\mathbf{y}} \text{ unspecified}$$

"*Hotelling's* T^2 statistic"

(*Hotelling* 1931, *Muirhead* 1982, *Rencher* 1998)

$$T^2 := [\hat{\xi} - \xi_0]' \hat{\Sigma}_{\hat{\xi}}^{-1} [\hat{\xi} - \xi_0] \tag{5.42}$$

Note that

$$\frac{n-6}{(n-1)\cdot 6} T^2$$

is an element of *Fisher's F*-distribution $F_{6,\,n-6}$ (Rencher 1998) and

$$P\{T^2 \le \frac{(n-1)\cdot 6}{n-6} F_{6,\ n-6}(1-\alpha)\} =$$

$$= P\{[\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}_0]' \hat{\boldsymbol{\Sigma}}_{\xi}^{-1} [\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}_0] \le \frac{(n-1)\cdot 6}{n-6} F_{6,\ n-6}(1-\alpha)\} = 1-\alpha = \gamma$$

where $F_{6,\ n-6}(1-\alpha)$ is the upper $(100\alpha)th$ percentile of *Fisher's F*-distribution. This leads immediately to a test of the hypothesis $\mathcal{H}_{01}: \boldsymbol{\xi} = \boldsymbol{\xi}_0$ versus $\mathcal{H}_{11}: \boldsymbol{\xi} \ne \boldsymbol{\xi}_0$. At the α error probability, we reject $\mathcal{H}_{01}$ in favor of $\mathcal{H}_{11}$ if

$$T^2 = [\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}_0]' \hat{\boldsymbol{\Sigma}}_{\xi}^{-1} [\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}_0] > \frac{(n-1)\cdot 6}{n-6} F_{6,\ n-6}(1-\alpha)\} = T^2_{1-\alpha}.$$

***Box* 5.5**

Separate Student t-tests about the eigenspace parameters in $\boldsymbol{\xi}$

Second Test for (separately)

$$\begin{array}{l|l|l|l|l|l} \mathcal{H}_{02}: \lambda_1 = \lambda_{10} & \lambda_2 = \lambda_{20} & \lambda_3 = \lambda_{30} & \theta_{32} = \theta_{320} & \theta_{31} = \theta_{310} & \theta_{21} = \theta_{210} \\ \mathcal{H}_{12}: \lambda_1 \ne \lambda_{10} & \lambda_2 \ne \lambda_{20} & \lambda_3 \ne \lambda_{30} & \theta_{32} \ne \theta_{320} & \theta_{31} \ne \theta_{310} & \theta_{21} \ne \theta_{210} \end{array}$$

"two-sided tests with *the test quantities*"

$$t_1 := \frac{\hat{\lambda}_1 - \lambda_{10}}{\hat{\sigma}_1},\ t_2 := \frac{\hat{\lambda}_2 - \lambda_{20}}{\hat{\sigma}_2},\ t_3 := \frac{\hat{\lambda}_3 - \lambda_{30}}{\hat{\sigma}_3},\ t_4 := \frac{\hat{\theta}_{32} - \theta_{320}}{\hat{\sigma}_4},\ t_5 := \frac{\hat{\theta}_{31} - \theta_{310}}{\hat{\sigma}_5},\ t_6 := \frac{\hat{\theta}_{21} - \theta_{210}}{\hat{\sigma}_6} \quad (5.43)$$

with respect to $\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\theta}_{32}, \hat{\theta}_{31}, \hat{\theta}_{21}$ of type $\boldsymbol{\Sigma}$-BLUUE and their variances. t_1, t_2, t_3, t_4, t_5 and t_6 are elements of the *Student t-distribution with n-1 degrees of freedom*.

The probability identity

$$P\{c_1 \le t \le c_2\} = P\{c_1\hat{\sigma} + \mu_0 \le \hat{\mu} \le c_2\hat{\sigma} + \mu_0\} = 1-\alpha = \gamma$$

relates the error probability α of the two-sided test to the confidence level γ. If $\hat{\mu}$ is an element of the confidence interval $c_1\hat{\sigma} + \mu_0 \le \hat{\mu} \le c_2\hat{\sigma} + \mu_0$, the null hypothesis $\mathcal{H}_0: \mu = \mu_0$ is accepted. We reject $\mathcal{H}_0$ if the confidence interval does not contain $\hat{\mu}$.

***Box* 5.6:**

***Eigen-inference* about the transformed parameters $\boldsymbol{\eta}$**

Eigen-tests (alternative to the second tests)

$$\begin{array}{l|l|l|l|l|l} \mathcal{H}'_{02}: \eta_1 = \eta_{10} & \eta_2 = \eta_{20} & \eta_3 = \eta_{30} & \eta_4 = \eta_{40} & \eta_5 = \eta_{50} & \eta_6 = \eta_{60} \\ \mathcal{H}'_{12}: \eta_1 \ne \eta_{10} & \eta_2 \ne \eta_{20} & \eta_3 \ne \eta_{30} & \eta_4 \ne \eta_{40} & \eta_5 \ne \eta_{50} & \eta_6 \ne \eta_{60} \end{array}$$

"two-sided tests with the *test quantities*"

$$t_1 := \frac{\hat{\eta}_1 - \eta_{10}}{\hat{\sigma}_{\eta_1}},\ t_2 := \frac{\hat{\eta}_2 - \eta_{20}}{\hat{\sigma}_{\eta_2}},\ t_3 := \frac{\hat{\eta}_3 - \eta_{30}}{\hat{\sigma}_{\eta_3}},\ t_4 := \frac{\hat{\eta}_4 - \eta_{40}}{\hat{\sigma}_{\eta_4}},\ t_5 := \frac{\hat{\eta}_5 - \eta_{50}}{\hat{\sigma}_{\eta_5}},\ t_6 := \frac{\hat{\eta}_6 - \eta_{60}}{\hat{\sigma}_{\eta_6}} \quad (5.44)$$

with respect to $\hat{\eta}_1, \hat{\eta}_2, \hat{\eta}_3, \hat{\eta}_4, \hat{\eta}_5$ and $\hat{\eta}_6$ and their related variances. t_1, t_2, t_3, t_4, t_5 and t_6 are elements of the *Student t-distribution with n-1 degrees of freedom*.

The probability identity

$$P\{c_1 \le t \le +c_2\} = P\{c_1\hat{\sigma} + \eta_0 \le \hat{\eta} \le c_2\hat{\sigma} + \eta_0\} = 1-\alpha = \gamma$$

relates the error probability α of the two-sided test to the confidence level γ. If $\hat{\eta}$ is an element of the confidence interval $c_1\hat{\sigma} + \eta_0 \le \hat{\eta} \le c_2\hat{\sigma} + \eta_0$, the null hypothesis $\mathcal{H}_0: \eta = \eta_0$ is accepted. We reject $\mathcal{H}_0$ if the confidence interval does not contain $\hat{\eta}$. Accordingly, we accept the original null hypothesis about the eigenspace components.

***Box* 5.7**

Multivariate hypothesis tests about the variance-covariance matrix $\boldsymbol{\Sigma}_{\mathbf{y}}$

Third Test for $\mathcal{H}_{03}: \; \boldsymbol{\Sigma}_{\mathbf{y}} = \boldsymbol{\Sigma}_0, \quad \mathcal{H}_{13}: \; \boldsymbol{\Sigma}_{\mathbf{y}} \neq \boldsymbol{\Sigma}_0$

"unbiased modified likelihood ratio statistic Λ_1"

(Giri 1977, Muirhead 1982, Koch 1999, Koch 2001)

$$\Lambda_1 = \left(\frac{e}{n-1}\right)^{6(n-1)/2} (\det(n-1)\hat{\boldsymbol{\Sigma}}_{\mathbf{y}}\boldsymbol{\Sigma}_0^{-1})^{(n-1)/2} \operatorname{etr}\{-\frac{1}{2}(n-1)\hat{\boldsymbol{\Sigma}}_{\mathbf{y}}\boldsymbol{\Sigma}_0^{-1}\} \tag{5.45}$$

with respect to the sample variance-covariance matrix $\hat{\boldsymbol{\Sigma}}_{\mathbf{y}}$ of type BIQUUE . Since our sample size is relatively small we have to use the exact distribution of $-2\log\Lambda_1$,whose upper 5 and 1 percentage points have been provided by *Muirhead* (1982, p.360).

***Box* 5.8**

Multivariate hypothesis tests about the eigenspace parameter vector $\boldsymbol{\xi}$ and the variance-covariance matrix $\boldsymbol{\Sigma}_{\mathbf{y}}$

Fourth Test for $\mathcal{H}_{04}: \; \boldsymbol{\xi} = \boldsymbol{\xi}_0, \; \boldsymbol{\Sigma}_{\mathbf{y}} = \boldsymbol{\Sigma}_0, \quad \mathcal{H}_{14}: \; \boldsymbol{\xi} \neq \boldsymbol{\xi}_0 \; \text{or} \; \boldsymbol{\Sigma}_{\mathbf{y}} \neq \boldsymbol{\Sigma}_0$

"unbiased likelihood ratio statistic Λ_2"

(*Anderson* 1984, *Muirhead* 1982)

$$\Lambda_2 = \left(\frac{e}{n}\right)^{6n/2} (\det(n-1)\hat{\boldsymbol{\Sigma}}_{\mathbf{y}}\boldsymbol{\Sigma}_0^{-1})^{n/2} \operatorname{etr}\{-\frac{1}{2}(n-1)\hat{\boldsymbol{\Sigma}}_{\mathbf{y}}\boldsymbol{\Sigma}_0^{-1}\} \exp\{-\frac{1}{2}[\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}_0]' \boldsymbol{\Sigma}_{\hat{\xi}0}^{-1} [\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}_0]\} \tag{5.46}$$

with respect to the eigenspace components of type $\boldsymbol{\Sigma}$ -BLUUE and variance-covariance matrix $\hat{\boldsymbol{\Sigma}}_{\mathbf{y}}$ of type BIQUUE and $\boldsymbol{\Sigma}_{\hat{\xi}0} = (1/n)(\mathcal{A}'\boldsymbol{\Sigma}_0^{-1}\mathcal{A})^{-1}$.
Since our sample size is relatively small we have to use the exact distribution of $-2\log\Lambda_2$, whose upper 5 and 1 percentage points have been provided by *Muirhead* (1982, p.371).

Chapter 6

The analysis of the eigenspace components of the strain rate tensor in central Mediterranean and Western Europe, 1992-2000

We have achieved the complete solution to the statistical inference of *eigenspace components* of the deformation tensors. The models developed in the last two chapters are closed and practical. The results bring a sound meaning to the deformation analysis. With these models we could successfully perform the statistical inference of the *eigenspace components* vector and the variance-covariance matrix of the *Gauss-Laplace* normally distributed observations of a random deformation tensor.

With the new space geodetic techniques, such as GPS, VLBI, SLR and DORIS, three-dimensional positions and change rates of network stations can be accurately determined from the regular measurement campaign, which is acknowledged as an accurate and reliable source of information in Earth deformation studies. This fact suggests that the components of deformation measures (such as the stress or strain tensor, etc.) can be estimated from the highly accurate geodetic data and analyzed by means of the proper statistical testing procedures. While station velocity diagrams demonstrate relative motions among stations, strain rate diagrams show the in-situ strain concentration rate which is directly connected to local stress concentration rates and possibly also to seismic hazard potentials (Ward, 1994). In strain analysis the displacements are considered as continuously differentiable according to the surface coordinates. The strain tensor components determined by means of the positional changes of the observation stations can be used for the computation of the stress tensors' components, taking into account the properties of the available materials within the investigation area. Therefore, the strain analysis can be considered as a basis of a dynamic model whereas the classical deformation analysis is similar to a kinematic model (see, e.g. *Flügge* 1972, *Means* 1976, *Grafarend* 1977, *Brunner*, 1979 and *Altiner* 1999).

The first geodetic deformation strain analysis based on the geodetic horizontal displacement was published by *Tsuboi* (1932), who computed the strain pattern using the horizontal displacement of control points in the Tango area of Japan during the period 1900 to 1930, which contains the Tango earthquake of magnitude 7.4 in 1927. The classic strain calculation methods from geodetic observations (distance, direction etc.) are contributed to *Frank* (1966), *Savage and Hastie* (1966) and *Prescott* (1976). Until now, more and more papers are published dealing with the stress, strain or strain rate deformation on the Earth's surface, such as *Angelier* (2002), *Haines and Holt* (1993), *Kahle et al*. (1995), *Kreemer et al*. (2000), *Savage et al*. (2001), *Scherneck et al*. (2002) and *Shen et al.* (1996).

The eigenspace components parameters (eigenvalues and principal directions) are of special importance in the deformation tensor analysis, for instance, the prediction of seismic activity. Due to the nonlinear functional relationship between the eigenvalues, the principal direction and the random tensor **T**, the variance-covariance of the eigenspace components is commonly calculated using a first-order approximate (*Angelier et al.* 1982, *Soler and van Gelder* 1991 and *Feigl et al.* 1990). With the benefit of the development of space geodesy and the continuous observations of the permanent networks including the *International GPS Service* (IGS) Network, *International Laser Ranging Service (ILRS)* Network, *International VLBI Service for Geodesy and Astrometry* (IVS) Network and *International DORIS Service (IDS)* Network and their combination *International Terrestrial Reference Frame* (ITRF) by IERS, we can now derive the strain rate tensor observations and estimate the eigenspace component parameters of these random tensor samples with our developed theory in the last two chapters, which addresses not only the present-day deformation pattern but also their continuous change of them.

In this chapter we begin with the discussion of the geodynamic setting of the Earth and especially the selected investigated region- the central Mediterranean and Western Europe. Then the space geodetic observations are introduced. Thirdly, the ITRF sites are selected according to the history and quality of the ITRF realization series, and the related residual velocities of selected ITRF sites are computed. Further, the methods of derivation for the two- and three-dimensional geodetic stain rates are introduced and applied to derive these strain rates from the residual velocities. In two case studies both BLUUE and BIQUUE models and hypothesis tests are applied to the eigenspace components of the two- and three-dimensional strain rate tensor observations in the area of the central Mediterranean and Western Europe, as derived from ITRF92 to ITRF2000 series station positions and velocities in Sections 6.6.and 6.7. The related *linear hypothesis test* has documented large confidence regions for the eigenspace components, namely *eigenvalues and eigendirections*, based upon real measurement configurations. They lead to the statement *to be cautious* with data of type extension and contraction as well as with the orientation of principal stretches.

6.1 Geodynamic setting of the investigated region

Planet Earth is a dynamic system that evolved within 4.6 billion years and continues this evolution. It depends on the way how heat - *"the geological lifeblood of planets"* is transferred out of the cooling Earth by thermal convection (hot stuff rises). Thermal convection causes plate tectonics: plates of the earth's surface move relatively to each other at a few mm/yr, which causes earthquakes, volcanoes, mountain building at plate boundaries. As the Earth's most important tectonic process, plate motion was first quantitatively described in the early 1960s. Conventional relative plate motion models are derived from combining rates of plate motion, inferred from magnetic anomalies at mid ocean ridges, with directions of plate motion, inferred from the azimuths of transform faults, and earthquake slip vectors at plate boundaries. These data are systematically inverted to yield a global model of the geologically "instantaneous" (covering the past few million years) motion between plates. Such a model is described by a set of angular velocities (*Euler* vectors) specifying the motion of each plate to one arbitrarily fixed plate. The first plate motion models were presented by *Minster and Jordan* (1978) and *Chase* (1978). Many new high-quality plate motion models have become available since the publications of these models. The new data have been used to determine improved global models, for example *NUVEL-1* (*DeMets et al.* (1990), (*Argus and Gordon* 1991) and its successor *NUVEL-1A* (*DeMets et al.* 1994). These models can explain the large-scale features of plate kinematics. Major deformations only take place in the comparatively narrow zones near the plate boundaries. Consequently, a large number of intense earthquakes occur near these zones. On the other hand, there is a low level of seismicity in the interior of plates. Figure 6.1 shows the boundaries of the major plates and the tectonic activity of the Earth (*Davidson, Reed and Davis* 2002).

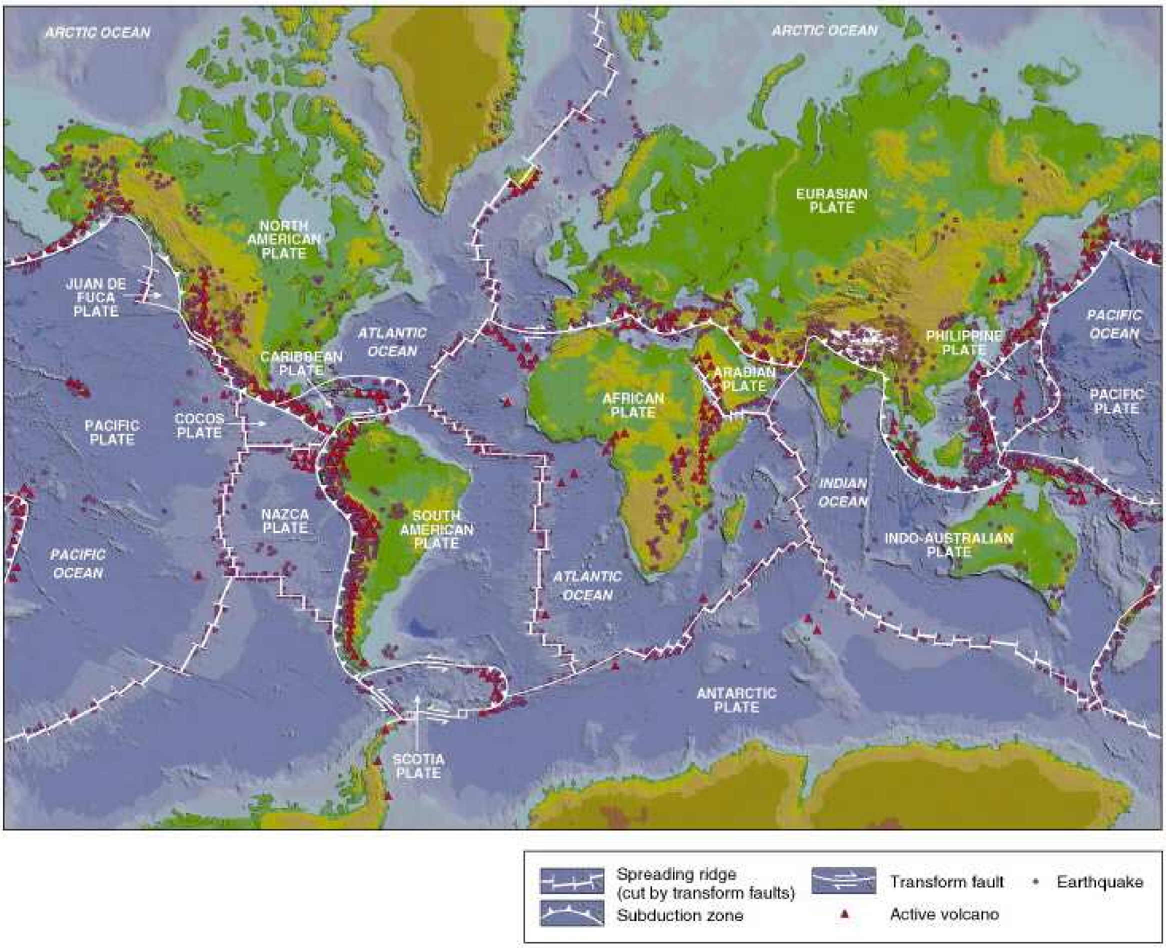


Figure 6.1. Boundaries of the major plates and the tectonic activity of the Earth (*Davidson, Reed & Davis* 2002)

From Figure 6.1 we can see that the recent major tectonic processes occur within the large-scale kinematic framework of active seafloor spreading (divergence) in the Atlantic Ocean and the African–Eurasian convergence (subduction) boundary in the Mediterranean. The spreading rate in the South Atlantic (~ 40 mm/yr), higher than that of the North Atlantic (~ 20 mm/yr) leads to a gradual counterclockwise rotation of the African plate, resulting in a NNW-directed push against Eurasia, which in turn leads to a lithospheric shortening of 5–6 mm/yr, increasing to 40 mm/yr in active subduction zones (*Argus et al.* 1989). With NW–SE-oriented spreading in the North Atlantic, the whole region is expected to be under compression, particularly the Mediterranean area. It is widely recognized that the Mediterranean area represents the collision zone between the African/Arabian and Eurasian plates, but the deformation pattern of this region is characterized by a complex space-time distribution of compressional and tensional events. The active volcanoes and earthquakes in this region show evidently a high seismic activity due to relatively strong tectonic forces that govern the compression zone. Moreover, a widespread intraplate seismicity occurs in the region which illustrates that the plate collision zone is complex and not sharply defined.

Therefore we would choose the Western European and Mediterranean areas to perform the statistical analysis of the eigenspace components of two- and three-dimensional strain rate tensor observations.

As pointed out by *Grafarend and Voosoghi* (2003), the European and Mediterranean area, is known as an extraordinary natural laboratory for the study of seismotectonic processes. This area is geologically, geophysically, and geodetically, one of the best-studied regions on the Earth's surface. The research interest encompasses the past 100 years and consequently a huge number of publications exists addressing local and regional geodynamical processes. A list of sample references are also provided in some papers, such as, *McKenzie* (1970), *Cross et al.* (1987), *Jackson and Mckenzie* (1988), *Argus et al.* (1989), *Smith et al.* (1990), *Castellarin et al.* (1992), *Mueller et al.* (1992), *James and Lambert* (1993), *Ward* (1994), *Reilinger et al.* (1997a, b), *Clarke et al.* (1998), *Kahle et al.*(1998), *DeMets and Dixon* (1999), *Caporali et al.* (2001). More recently, the present-day crustal motions in central Mediterranean area and in Western Europe, are studied by *Anzidei et al.* (2001), *Devoti et al.* (2002a, b), *Grenerczy* (2002), *Caporali* (2003a, b), *Jimenez-Munt et al.* (2003) and *Nocquet and Calais* (2003) with the newly developed continuous observation data from space geodetic networks, such as permanent GPS networks.

The European and Mediterranean area can be divided into three main subregions with distinct geodynamic features, namely Western Europe, Northern Europe and the Alpine-Mediterranean sub-regions. Based on the space geodetic observations history (10 or more years) we will focus on the behaviour of significant active deformation in the North of the Western Mediterranean with respect to Europe, i.e. two of the subregions, Alpine-Mediterranean and Western Europe. Within Western Europe, weak seismic activity is observed. The area is characterized as a field of compressional tectonics. A generalized stress map of Europe (*Mueller et al.* 1992) indicates a generally NW-SE uniform orientation for the maximum compressive horizontal principal stress in Western Europe. The Alpine-Mediterranean region marks a broad transformation zone between the African, Arabian, and Eurasian plates. The region is expected to be largely under compression. It is characterized as a region of intensive seismic activity. The tectonic evolution of this region is strongly affected by the convergence of the microplates (*Voosoghi* 2000). In order to get reliable observation on the ongoing tectonic processes in this area we should select the sites (see Chapter 6.3) that are not affected by local geophysical phenomena.

6.2 Space geodetic data

According to *F.R. Helmert* (1880), the classic assignment of geodesy is the surveying and mapping of the earth's surface and also of the gravitational field, enlarged by the requirements of accuracy. Space geodesy is geodesy by means of satellites, moon, planets, radio stars and quasars, which has been developed since end 1960s. At present there are four widely used techniques in space geodesy, *Very Long Baseline Interferometry* (VLBI), *Satellite Laser Ranging* (SLR), the *Global Positing System* (GPS) and the *Doppler Orbitography and Radiopositioning Integrated Satellite System* (DORIS). The strengths of the different observing techniques include, for example: VLBI has relationship to the inertial reference frame; SLR has relationship to the geocenter and the Earth's gravity field; GPS is a highly operational system for the densification of the terrestrial reference frame; and DORIS has homogeneous global distribution of the tracking stations.

These techniques combine precise satellite-based timing, ranging, and orbit estimation to measure the positions and velocities of geodetic sites to centimeter and centimeter/yr or better accuracy. In the past 10 years the scope and accuracy of space geodetic techniques has expanded greatly. In some regions geodetic measurements are probably more accurate than conventional global plate motion models, which gave the first in-situ measurements of plate motion (*Beutler* 2000). The present accuracy of geodetic VLBI has arrived at ±5-20 mm for session coordinates, the annual coordinate and velocity accuracy is ±1-4mm and ±0.1-1 mm/ yr, respectively (*Schuh et al.* 2002). The new generation of SLR ranging accuracy has reached mm level, which supports maintenance of a centimenter accuracy position. Nowadays the estimation accuracy form ten years of accurate global SLR observations is ±6 mm for coordinates and ±2 mm/year for the velocities (*Angermann et al*. 2001). With geodetic GPS techniques the station coordinates can be determined with achievable accuracy, in general ±1 cm and an annual velocity accuracy can be reached of about ±1 mm/yr. There are two services of DORIS: in operational geodesy with dedicated location beacons, any point on Earth at any time can be determined with about ±20 cm accuracy after a one-day measurement time and ±10 cm after 5 days; the permanent beacon network delivers high precision 3D coordinates for geodetic and geodynamic applications. Positions and motions are available to better than ±1cm and ±1mm/yr, respectively (*Seeber* 2003).

With the capacity of accuracy and the wide distributions of stations on the Earth, space geodetic techniques provide measurements that can be used to infer crustal deformation over global scales and can be compared with predictions from conventional global plate motion models. It is interesting to explore cases where significant differences exist, to determine whether they reflect uncertainties and errors in one or both approaches, or instead reflect real differences in plate motions over different scales. However, space geodetic velocities within boundary zones often differ from the predictions of plate motion models because geodetic velocities over a few years include the effects of transient elastic deformation associated with the cycle of strain accumulation and release in plate boundary earthquake (*Savage* 1983, *Scholz* 1990), which is averaged out over the millions of years used in global plate motion models. There are many comparisons about the different approaches, for instance, *Drewes* (1999) derived an *Actual Plate Kinematic and Deformation Model* (APKIM) from present day geodetic observations, such as VLBI, SLR and GPS. A series of such models has been developed. In the latest version, APKIM2000, about 280 site velocities were used to estimate 12 plate rotation vectors. In general, the agreement between APKIM2000 and NNR-NUVEL-1A (*Argus and Gordon* 1991, *DeMets et al.* 1990) is very high. Significant differences, however, are visible in plate boundary zones (Figure 6.2), where other plate kinematic models based on the 405 ITRF 2000 geodetic site velocities are also illustrated (*Drewes and Angermann* 2001).

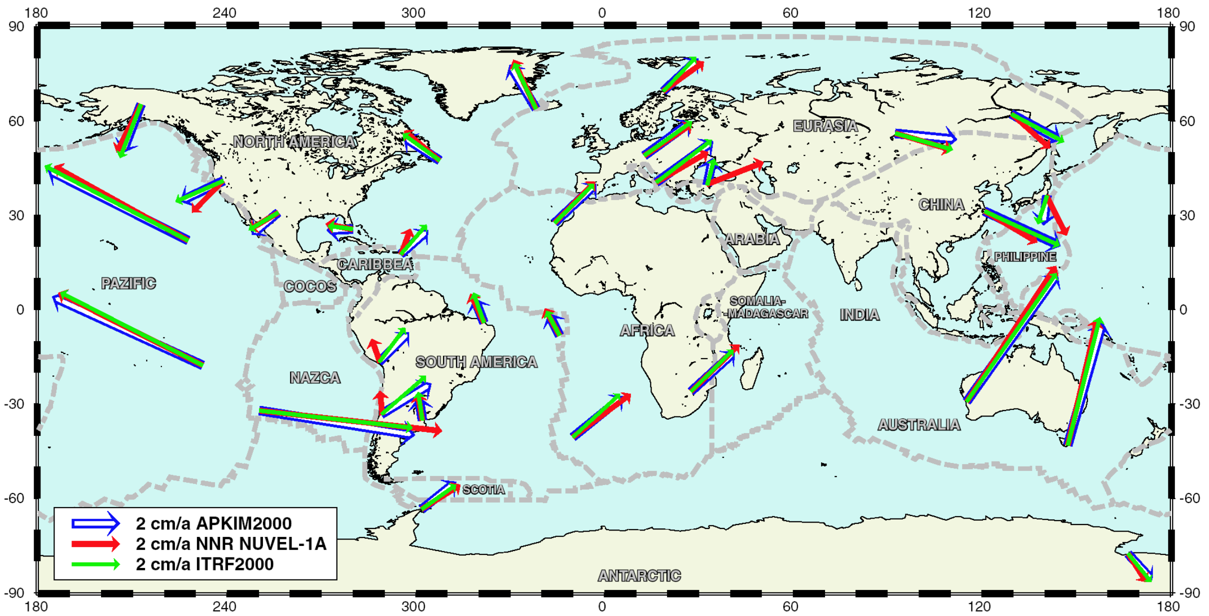


Figure 6.2. Station motions derived from the APKIM2000 model in comparison with ITRF2000 and the NNR NUVEL-1A model (*Drewes and Angermann* 2001)

Eventually, however, deformation becomes slow and diffuse enough that it is more usefully regarded as intraplate. Space geodesy is ideal for addressing this issue because it can measure motions of a few mm/yr. As a result, there has been considerable interest in using space geodesy to quantify the rigidity of the major plates and investigate how deviations from plate rigidity give rise to intraplate deformations and earthquakes. The first step to estimate the rigidity of a plate using space-based geodesy is to find the motions of sites within it, and to compare these motions to those predicted, assuming that the plate is rigid and so can be described by a single Euler vector. The next step is to use geodetic velocities to estimate the intraplate strain rate fields in various areas, and compare it to present-day seismicity, paleoseismicity, and geologic data. This is in principle straightforward; the strain field can be derived by forming least-squares estimates of the velocity gradient in various regions. The studies with GPS have not yet detected intraplate strain accumulation, but show that intraplate strain accumulation rates are slow, which yields useful insight into seismic hazards. Nonetheless, given the rapidly growing number of continuous GPS sites and the longer spans of GPS data, it seems likely that the strain accumulation signal will soon "climb" above the noise and provide a valuable signal for investigation of intraplate tectonics (*Stein and Sella* 2002).

Space geodetic data greatly simplify identification and study of continental microplates. The data are sufficient to estimate an Euler vector describing the micoplate's motion relative to the major plates. During the last few years, precision and accuracy of the space geodesy techniques have been improved, and special efforts have been dedicated to the combination of their results, such as those produced by the IERS (International Earth Rotation Service) and IGS (International GPS Service), which are fortunately dense and accurate in Europe. Many fixed stations in this region are fiducial stations of IGS and IERS global networks, and consequently the tracking history is remarkable for quality and quantity of data. Since the longer history of IERS and with benefit the combination of most of the space geodetic techniques, we will review it and choose the appreciated sites for our study regions.

6.3 The selection of ITRF Sites and data preparation

Before we begin to choose from the ITRF series results and the stations in the studied regions, let us make a review of the history of IERS and the realization of ITRF.

6.3.1 The history of ITRF

The IERS was established as the International Earth Rotation Service in 1987 by the International Astronomical Union and the International Union of Geodesy and Geophysics and it began operation on 1 January 1988. In 2003 it was renamed to *International Earth Rotation and Reference Systems Service* (*IERS*, 2003). The primary objectives of the IERS are to serve the astronomical, geodetic and geophysical communities by providing the International Celestial Reference System (ICRS) and its realization, the International Celestial Reference Frame (ICRF); the International Terrestrial Reference System (ITRS) and its realization, the International Terrestrial Reference Frame (ITRF); Earth orientation parameters required to study earth orientation variations and to transform between the ICRF and the ITRF; Geophysical data to interpret time/space variations in the ICRF, ITRF or earth orientation parameters, and model such variations and the standards, constants and models (i.e., conventions) encouraging international adherence.

The Conventional Terrestrial Reference System (CTRS), established and maintained by the IERS, and nearly exclusively used for today's scientific and practical purposes is the International Terrestrial Reference System (ITRS), which constitutes a set of prescriptions and conventions together with the modelling required to define origin, scale, orientation and time evolution of a Conventional Terrestrial Reference System (CTRS). The ICRS is an ideal reference system, as defined by the *IUGG Resolution No. 2* adopted in Vienna, 1991. The system is realised by the International Terrestrial Reference Frame *(ITRF)* based upon estimated three-dimensional coordinates and velocities of a set of stations observed by *VLBI*, *LLR*, *GPS*, *SLR*, and *DORIS*. The ITRS can be connected to the International Celestial Reference System *(ICRS)* by use of the IERS Earth Orientation Parameters *(EOP)*. The ITRS is defined as follows (McCarthy 2003)

- It is geocentric, the center of mass being defined for the whole Earth, including oceans and atmosphere;
- The unit of length is the meter (SI). This scale is consistent with the TCG time coordinate for a geocentric local frame, in agreement with IAU and IUGG (1991) resolutions. This is obtained by appropriate relativistic modeling;
- Its orientation was initially given by the Bureau International de l'Heure (BIH) orientation at 1984.0;
- The time evolution of the orientation is ensured by using a no-net-rotation condition with regards to horizontal tectonic motions over the whole Earth.

Realizations of the ITRS are produced by the IERS's ITRS Product Center (ITRS-PC) under the name International Terrestrial Reference Frame (ITRF). The current procedure is to combine individual TRF solutions computed by IERS analysis centers using observations of space geodesy techniques: VLBI, LLR, SLR, GPS and DORIS. These individual ITRF solutions currently contain 3-dimensional Cartesian station positions and velocities together with full variance-covariance matrices. Currently, ITRF solutions are published nearly annually by the ITRS-PC in the Technical Notes (cf. *Boucher et al.*, 1999). The numbers (yy) following the designation ITRF specify the last year whose data were used in the formation of the frame. Hence ITRF97 designates the frame of station positions and velocities constructed in 1999 using all of the IERS data available until 1998.

Until now, 10 successive realizations of the ITRF have been published, starting with ITRF88 and ending with ITRF2000, each of which superseded its predecessor.

From ITRF88 till ITRF93, the ITRF Datum Definition is summarized as follows:

- Origin and Scale: defined by an average of selected SLR solutions;
- Orientation: defined by successive alignment since BTS87 whose orientation was aligned to the BIH EOP series. Note that the ITRF93 orientation and its rate were again realigned to the IERS EOP series;
- Orientation Time Evolution: No global velocity field was estimated for ITRF88 and ITRF89 and so the AM0-2 model of (Minster and Jordan, 1978) was recommended. Starting with ITRF91 and till ITRF93, combined velocity fields were estimated. The ITRF91 orientation rate was aligned to that of the NNR-

NUVEL-1 model, and ITRF92 to NNR-NUVEL-1A (*Argus and Gordon*, 1991), while ITRF93 was aligned to the IERS EOP series.

Since the ITRF94, full variance-covariance matrices of the individual solutions incorporated in the ITRF combination have been used. At that time, the ITRF94 datum was achieved as follows (*Boucher et al.*, 1996):

- Origin: defined by a weighted mean of some SLR and GPS solutions;
- Scale: defined by a weighted mean of VLBI, SLR and GPS solutions, corrected by 0.7 ppb to meet the IUGG and IAU requirement to be in the TCG (Geocentric Coordinate Time) time-frame instead of TT (Terrestrial Time) used by the analysis centers;
- Orientation: aligned to the ITRF92;
- Orientation time evolution: aligned the velocity field to the model NNR-NUVEL-1A, over the 7 rates of the transformation parameters.

The ITRF96 was then aligned to the ITRF94, and the ITRF97 to the ITRF96 using the 14 transformation parameters (*Boucher et al*., 1998; 1999).

The ITRF network has improved with time in terms of the number of sites and collocations as well as their distribution over the globe. The ITRF88 network, having about 100 sites and 22 collocations (VLBI/SLR/LLR), and the ITRF2000 network containing about 500 sites and 101 collocations (VLBI/SLR/GPS/DORIS), see Figure 6.3 (*McCarthy* 2003). With the improvements of the analysis strategy by the IERS Analysis Centers and the ITRF combination the ITRF position and velocity precisions have also improved with time.

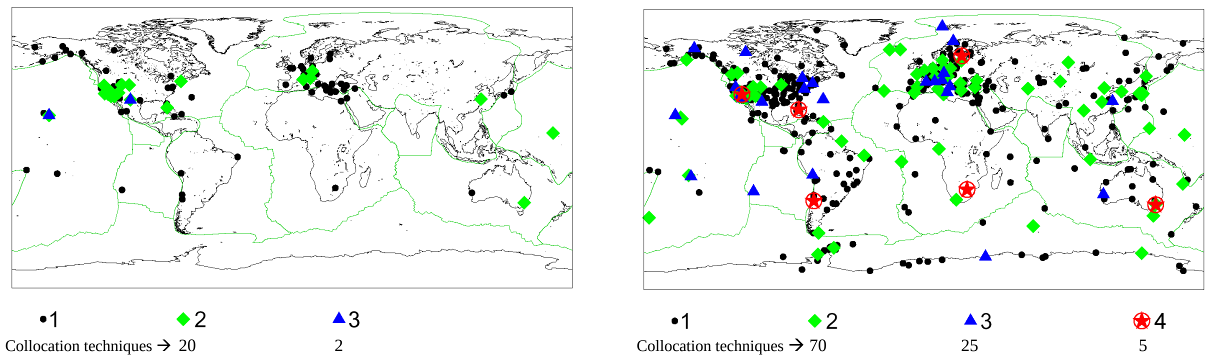


Figure 6.3. The ITRF88 (left) and ITRF2000 (right)sites and collocated techniques. (McCarthy 2003)

As the current Reference Realization of the ITRS, the ITRF2000 is intended to be a standard solution for georeferencing and all Earth science applications. In addition to primary core stations observed by VLBI, LLR, SLR, GPS and DORIS, the ITRF2000 is densified by regional GPS networks in Alaska, Antarctica, Asia, Europe, North and South America, and the Pacific. The individual solutions used in the ITRF2000 combination are generated by the IERS analysis centers using removable, loose or minimum constraints. In terms of datum definition, the ITRF2000 is characterized by the following properties (*McCarthy* 2003)

- the scale is realized by setting to zero the scale and scale rate parameters between ITRF2000 and a weighted average of VLBI and most consistent SLR solutions. Unlike the ITRF97 scale expressed in the TCG-frame, that of the ITRF2000 is expressed in the TT-frame;
- the origin is realized by setting to zero the translation components and their rates between ITRF2000 and a weighted average of most consistent SLR solutions;
- the orientation is aligned to that of the ITRF97 at 1997.0 and its rate is aligned, conventionally, to that of the geological model NNR-NUVEL-1A. This is an implicit application of the no-net-rotation condition, in agreement with the ITRS definition. The ITRF2000 orientation and its rate were established using a selection of ITRF sites with high geodetic quality, satisfying the following criteria:
  - continuous observation for at least 3 years;
  - locations far from plate boundaries and deforming zones;
  - velocity accuracy (as a result of the ITRF2000 combination) better than ±3 mm/y;
  - velocity residuals less than ±3 mm/y for at least 3 different solutions.

The ITRF2000 results show significant disagreement with the geological model NNR-NUVEL-1A in terms of relative plate motions (*Altamimi et al.* 2002). Although the ITRF2000 orientation rate alignment to NNR-NUVEL-1A is ensured at the ±1 mm/y level, regional site velocity differences between the two may exceed ±3 mm/y. Meanwhile it should be emphasized that these differences do not at all disrupt the internal consistency of the ITRF2000, simply because the alignment defines the ITRF2000 orientation rate and nothing more. Moreover, angular velocities of tectonic plates which would be estimated using ITRF2000 velocities may significantly differ from those predicted by the NNR-NUVEL-1A model, more details can be seen in Figure 6.3.

6.3.2 The realization of ITRF

The construction of ITRF is based on the combination of sets of station coordinates (SSCs) and velocities derived from observations of space geodetic techniques, such as VLBI, SLR, and LLR by various analysis centers. In 1991, the IERS added GPS to this list of techniques; and in 1994, it added DORIS. For the determination of a station's position in an ITRF, the station is assigned to a specific tectonic plate. The point position of the station at time, *t*, on the surface of the solid earth, is expressed as (*Boucher and Altimini* 1993):

$$\mathbf{X}(t) = \mathbf{X}_0 + \mathbf{V}_0(t - t_0) + \sum_i \Delta\mathbf{X}_i(t) \tag{6.1}$$

where

- $\Delta\mathbf{X}_i$: corrections to the various time changing effects;
- $\mathbf{X}_0$: position at epoch t_o;
- $\mathbf{V}_0$: velocity at epoch t_o;
- t_o : initial reference epoch (i.e. 1988.0).

The coordinates of sites on the earth's surface slowly change (by up to 10cm per year,or so) due to the motion of the tectonic plates – a component which is familiarly known as "continental drift". The velocity $\mathbf{V}_0$ should be expressed as

$$\mathbf{V}_0 = \mathbf{V}_{plate} + \mathbf{V}_{ice} + \mathbf{V}_r \tag{6.2}$$

where

- $\mathbf{V}_{plate}$: is the horizontal velocity due to plate tectonic motion, which can be described by a geophysical and geological angular velocity vector $\boldsymbol{\omega}_{YY}$ (a Cartesian rotation vector with components ω_X, ω_Y, ω_Z) of the absolute plate motion models, such as the more recent NNR-NUVEL-1A;
- $\mathbf{V}_{ice}$: is the vertical velocity due to post glacial rebound, to be computed from models such as ICE-4G (*Peltier* 1995);
- $\mathbf{V}_r$: is the residual velocity.

In the data analysis $\mathbf{X}_0$, and $\mathbf{V}_r$ should be estimated parameters. When adjusting parameters, in particular velocities, the IERS orientation should be kept at all epochs, which means to ensure the alignment at a reference epoch and the time evolution through a no net rotation condition with regards to horizontal tectonic motion over the whole Earth.

6.3.3. The selection of ITRF series results and the stations

With the review of the history of IERS and the realization of ITRF we will further discuss the selection of ITRF series results and the stations in our studied region.

Prior to ITRF91, no velocity field had been derived so the *AMO-2 model* (*Minister & Jordan* 1978) is applied to account for the time evolution of ITRS. ITRF91 was the first realization of ITRS to derive a global velocity field by combining site velocities estimated by SLR and VLBI analysis centres (*Boucher and Altamimi* 1993). To ensure the condition of no-net-rotation of ITRS with respect to the earth's crust, NNR-NUVEL1 was selected as the reference motion model of ITRF92. NNR-NUVEL1 is a horizontal motion model only. For the consistency of the three-dimensional nature of ITRS, the vertical velocity is set to zero with an assumed error of 1 cm/year (*Boucher and Altamimi* 1993). The ITRF is a dynamic datum which was introduced in ITRF88 meaning that every year there is a change. The change between ITRF91 and ITRF92 was less than 2cm, and as more observation became available and computational techniques improved, revised reference systems were produced, generally on an annual basis (ITRF93, ITRF94, ITRF96, ITRF97 and ITRF2000). However, the change between it and subsequent ITRF's is only of the order of a couple of centimeters. To the early ITRF realizations there are some facts that should be considered in our selection:

- The ITRF92 site velocities seem to be more realistic with respect to ITRF91 (*Boucher and Altamini* 1993) and ITRF93 is consistent with NNR-NUVEL-1A;.
- ITRF91 has less stations than ITRF92 and the following series;
- The early realizations of ITRF are in lower accuracy, for instance, the standard formal error of station coordinates: $\sigma_{ITRF88} \le \pm 20cm$, $\sigma_{ITRF91} \le \pm 10cm$ and since $\sigma_{ITRF92} \le \pm 5cm$. The standard formal error of station velocities: $\sigma_{VITRF91} \ge \sigma_{VITRF92}$;

- Early ITRFs accommodate the horizontal velocity of sites on plate boundaries by assigning a larger a-priori standard deviation (±10 cm/year) to the site's velocity than for sites located on the rigid part of a tectonic plate (±3 mm/year) (*Boucher et al*. 1994);
- Although the station coordinates are changed between different epochs, there is no difference between the velocity field of ITRF92 at epoch 88.0 and epoch 1994.0;
- Nonetheless, these are small differences which demonstrate the excellent reliability of the ITRF velocity fields and the uniform motion hypothesis even over a fairly long period of time.

In order to fulfil the quality requirements (in determination of successful strain calculation in Western Europe and central Mediterranean in our study), the ITRF sites are selected with high geodetic quality, satisfying the following criteria:

- it is a primary permanent station after the standards of IERS since 1991 or collocated with several observational techniques
- has continuous observation for at least 3 years in 1992;
- with station coordinate accuracy better than ±8 mm;
- and velocity accuracy better than ±3 mm/y.

Therefore we have chosen the 8 primary stations and one secondary station (Noto) with all collocated VLBI and SLR techniques from ITRF92 and their following series in Western Europe and central Mediterranean in our study, which are listed in Table 6.1

Table 6.1 Catalogue of selected IERS Sites based on ITRF92 in Europe

DOMES NB.	SITE NAME	TECH. ID.	Country	Long. d m	Lat. d m	Plate	(*)
10002S001	GRASSE	SLR 7835	France	6 55	43 45	EURA	P C
10402S002	ONSALA	VLBI 7213	Sweden	11 55	57 24	EURA	P C
11001S002	GRAZ	SLR 7839	Austria	15 30	47 04	EURA	P C
12711S001	BOLOGNA	VLBI 7230	Italy	11 21	44 29	EURA	P C
12717S001	NOTO	VLBI 7547	Italy	14 59	36 53	AFRC/EURA	S C
12734S001	MATERA	SLR 7939	Italy	16 37	40 42	EURA	P C
13407S010	MADRID	VLBI 1565	Spain	355 44	40 26	EURA	P C
14001S001	ZIMMERWALD	SLR 7810	Switzerland	7 28	46 53	EURA	P C
14201S004	WETTZELL	VLBI 7224	Germany	12 53	49 09	EURA	P C

(*) P: Primary S: Secondary C: Collocation

6.3.4 The computation of residual velocities of ITRF stations

Since we need the residual velocities in determination of the strain rate, the published ITRF velocity $\mathbf{V}_0$ should be converted to residual velocities $\mathbf{V}_r$ with respect to e.g. the Eurasian fixed plate by subtracting the rigid motion of Eurasia, which is computed by the angular velocity vector $\boldsymbol{\omega}_{YY}$ of the Eurasian plate in the absolute plate motion models of every ITRF realizations:

$$\mathbf{V}_r = \mathbf{V}_0 - \mathbf{V}_{plate} = \mathbf{V}_0 - \boldsymbol{\omega}_{YY} \times \mathbf{X} = \begin{bmatrix} V_{X0} \\ V_{Y0} \\ V_{Z0} \end{bmatrix} - \begin{bmatrix} 0 & Z & -Y \\ -Z & 0 & X \\ Y & -X & 0 \end{bmatrix} \begin{bmatrix} \omega_X \\ \omega_Y \\ \omega_Z \end{bmatrix}, \tag{6.3}$$

It should be noted that although the station Noto is located on the African plate, the motion of Matera was recognized as neither purely Euro-Asiatic nor African. For the calculation of strain rate filed in the selected region we have to compute the residual velocity with respect to the same rigid plate, e.g. Eurasia plate.

In the three-dimensional case the vertical velocity due to post glacial rebound should be considered in deriving the residual velocities $\mathbf{V}_r$:

$$\mathbf{V}_r = \mathbf{V}_0 - \mathbf{V}_{plate} - \mathbf{V}_{ice} = \mathbf{V}_0 - \boldsymbol{\omega}_{YY} \times \mathbf{X} - \mathbf{V}_{ice} = \begin{bmatrix} V_{X0} \\ V_{Y0} \\ V_{Z0} \end{bmatrix} - \begin{bmatrix} 0 & Z & -Y \\ -Z & 0 & X \\ Y & -X & 0 \end{bmatrix} \begin{bmatrix} \omega_X \\ \omega_Y \\ \omega_Z \end{bmatrix} - \begin{bmatrix} V_{Xice} \\ V_{Yice} \\ V_{Zice} \end{bmatrix}, \tag{6.4}$$

In ITRF92 the orientation time evolution was ensured by aligning the corresponding velocity fields to NNR-NUVEL-1 model (*Argus and Gordon* 1991, *DeMets et al*. 1990). So for ITRF92, $\boldsymbol{\omega}_{YY}$ corresponds conventionally to the angular velocity of the Eurasian plate in NNR-NUVEL-1 model.

More recently, the geophysical model NNR-NUVEL-1A (*DeMets et al*. 1994) has been used as a reference in the ITRF93 velocity field computation. It should be noted that there is a rotation rate between the ITRF93 velocity field and the NNR-NUVEL-1A model (*Boucher et al*, 1994). Consequently for ITRF93, $\boldsymbol{\omega}_{YY}$ corresponds to the angular velocity of the Eurasian plate in the NNR-NUVEL-1A model to which we added the rotation rate between the ITRF93 velocity field and the NNR-NUVEL-1A model. As time evolution of ITRF94 is consistent with the model NNR-NUVEL-1A (*Boucher et al*, 1996), thus $\boldsymbol{\omega}_{YY}$ corresponds conventionally to the angular velocity of the Eurasian plate in this model.

The reference frame definition (origin, scale, orientation and time evolution) of the ITRF96 is achieved in such a way that ITRF96 is in the same system as ITRF94 (*Boucher et al.* 1998). Consequently, $\boldsymbol{\omega}_{YY}$ is the same as for ITRF94. This same statement is also valid for ITRF97.

For the first time ITRF2000 combines individual solutions that are free from any plate motion model. Its origin is defined by a weighted average of most consistent SLR solutions. Its scale is defined by most consistent SLR and VLBI solutions. Its orientation is aligned to ITRF97 at epoch 1997.0 and its orientation rate follows, conventionally, that of NNR-NUVEL-1A model. The ITRF2000 velocity field was used to estimate angular velocities of 6 major plates, including Eurasia, showing significant disagreement with NNR-NUVEL-1A predictions. It is therefore recommended for $\boldsymbol{\omega}_{YY}$ to use the components of the Eurasian angular velocity estimated from ITRF2000 velocities of 19 European sites with higher geodetic quality. For more details, see *Altamimi et al.* (2002). Table 6.2 summarizes the component values of $\boldsymbol{\omega}_{YY}$:

Table 6.2 The estimation of $\boldsymbol{\omega}_{YY}$

ITRF	ω_X (mas/y)	ω_Y (mas/y)	ω_Z (mas/y)
92	0.21	0.52	-0.68
93	0.32	0.78	-0.67
94	0.20	0.50	-0.65
96	0.20	0.50	-0.65
97	0.20	0.50	-0.65
2000	0.081	0.490	-0.792

In the two-dimensional case we need the surface (horizontal) residual velocities. Since the residual velocity from (6.3) is relatively small they can approximately be transformed from the global geocentric Cartesian coordinate system to the local geodetic system (*Seeber* 2003) using

$$\begin{bmatrix} V_{Er} \\ V_{Nr} \\ V_{Ur} \end{bmatrix} = \begin{bmatrix} -\sin L & \cos L & 0 \\ -\sin B \cos L & -\sin B \sin L & \cos B \\ \cos B \cos L & \cos B \sin L & \sin B \end{bmatrix} \begin{bmatrix} V_{Xr} \\ V_{Yr} \\ V_{Zr} \end{bmatrix} \tag{6.5}$$

where (*B*, *L*) are the geodetic ellipsoidal coordinate of a discussed ITRF site, which are converted from the global Cartesian coordinates of the ITRF sites (X, Y, Z) with respect to the GRS80 reference ellipsoid. The local geodetic system is defined as follows (*Vanicek and Krakiwsky* 1986): it is topocentric (T) ; the *U*-axis is the outward ellipsoid normal passing through T; the *E*-axis is directed towards *geodetic east*; the *N*-axis is directed towards *geodetic north*; Therefore two-dimensional surface (horizontal) velocities are the first two elements of $(V_E \ V_N)$.

Based on the procedures ((6.3) and (6.5)) derived above we can compute the horizontal residual velocities of every ITRF realizations with respect to the "Eurasian fixed' plate motion model (see Table 6.3). They are listed in Table 6.3 together with the ITRF92 to ITRF2000 velocity solutions of the selected stations in the central Mediterranean and Western Europe. The residual velocities will be used to compute the strain rates in the next section. Therefore, together with the principal strain rates for every epoch they are also illustrated in the next section.

For the three-dimensional case we would like to apply directly the three dimensional Cartesian residual velocities in deriving the strain rate tensor for the six series of ITRF realizations in the selected sub-network including four sites : 1 - 4 - 8 - 9 (Grasse - Bologna - Zimmerwald - Wettzell). The three dimensional Cartesian residual velocities of the four selected sites are computed, based on (6.4) and listed in Table 6.4.

Table 6.3 Horizontal station velocities

Site	Velocity (mm/yr) East	Velocity (mm/yr) North	Residual velocity (mm/yr) East	Residual velocity (mm/yr) North	Velocity (mm/yr) East	Velocity (mm/yr) North	Residual velocity (mm/yr) East	Residual velocity (mm/yr) North	Velocity (mm/yr) East	Velocity (mm/yr) North	Residual velocity (mm/yr) East	Residual velocity (mm/yr) North
	ITRF92				ITRF93				ITRF94			
Grasse	20.52	12.05	-0.49	-2.99	25.77	19.99	2.02	-2.73	20.72	12.76	0.62	-1.63
Onsala	16.12	15.20	-3.39	0.97	21.05	23.11	-2.44	1.60	16.95	13.88	-1.69	0.26
Graz	22.04	16.37	-0.02	2.76	28.75	22.78	2.97	2.21	22.15	14.88	1.06	1.86
Bologna	22.23	17.23	0.49	2.95	26.37	25.21	1.41	3.62	22.38	16.14	1.60	2.47
Noto	20.91	21.91	-2.16	8.20	25.62	29.06	-0.40	8.35	20.88	19.39	-1.19	6.27
Matera	22.62	19.37	-0.40	5.98	28.78	26.88	2.41	6.65	24.44	16.92	2.43	4.10
Madrid	19.71	16.82	0.23	0.44	21.09	25.49	0.09	0.73	18.68	16.12	0.06	0.44
Zimmerwald	21.66	16.90	1.04	1.95	25.94	25.15	2.35	2.56	19.70	16.32	-0.02	2.01
Wettzell	19.46	15.77	-1.81	1.69	24.24	23.41	-0.65	2.14	19.66	14.16	-0.68	0.69
	ITRF96				ITRF97				ITRF2000			
Grasse	20.21	14.32	0.11	-0.07	19.79	13.06	-0.31	-1.33	20.34	14.71	-0.35	-0.01
Osala	17.34	13.11	-1.31	-0.51	17.19	13.18	-1.45	-0.44	17.25	13.59	-0.66	-0.69
Graz	22.13	14.29	1.04	1.27	21.89	13.46	0.80	0.43	22.14	14.46	0.72	0.55
Bologna	23.83	15.14	3.04	1.46	23.12	15.32	2.34	1.64	23.36	16.14	2.03	1.83
Noto	22.38	19.09	0.30	5.97	21.96	17.73	-0.11	4.61	21.28	18.05	-2.12	4.07
Matera	23.82	18.52	1.81	5.70	23.53	17.36	1.52	4.55	23.70	18.09	0.71	4.32
Madrid	19.65	15.05	1.02	-0.63	19.13	14.62	0.51	1.06	18.98	15.68	-0.58	0.40
Zimmerwald	18.60	15.01	-1.13	0.70	19.04	13.96	-0.68	-0.35	20.14	15.07	0.13	0.39
Wettzell	20.51	13.37	0.17	-0.10	20.19	13.41	-0.15	-0.06	20.27	14.37	-0.17	0.18

Table 6.4 Three-dimensional Cartesian station velocities

Site	Velocity (mm/yr) V_X	V_Y	V_Z	Residual velocity (mm/yr) Vr_X	Vr_Y	Vr_Z	Velocity (mm/yr) V_X	V_Y	V_Z	Residual velocity (mm/yr) Vr_X	Vr_Y	Vr_Z
	ITRF92						ITRF93					
Grasse	-17.90	18.50	1.80	-5.08	-1.11	-9.09	-17.80	23.80	13.50	0.61	2.11	-2.96
Bologna	-18.40	18.90	10.20	-4.24	-0.38	-0.02	-26.30	21.50	14.30	-6.49	0.11	-1.14
Zimmerwald	-16.40	19.70	10.10	-2.93	0.66	-0.16	-21.10	23.40	17.70	-1.74	2.14	2.20
Wettzell	-16.90	16.10	9.20	-1.81	-2.27	-0.04	-25.20	19.10	12.30	-4.02	-1.59	-1.67
	ITRF94						ITRF96					
Grasse	-12.10	19.40	8.40	0.17	0.65	-2.03	-12.00	18.90	10.60	0.27	0.15	0.17
Bologna	-18.70	19.00	8.40	-5.14	0.57	-1.38	-14.20	21.40	11.80	-0.64	2.97	2.02
Zimmerwald	-13.50	18.10	12.10	-0.61	-0.10	2.28	-11.10	17.30	12.60	1.79	-0.90	2.78
Wettzell	-16.90	16.30	6.80	-2.46	-1.26	-2.05	-15.90	17.40	7.00	-1.46	-0.16	-1.85
	ITRF97						ITRF2000					
Grasse	-11.80	18.50	9.00	0.47	-0.25	-1.43	-13.10	18.90	10.10	-0.54	-0.42	-0.57
Bologna	-18.00	19.90	8.10	-4.44	1.47	-1.68	-18.70	20.00	8.60	-4.60	1.12	-1.64
Zimmerwald	-11.50	17.70	10.70	1.39	-0.50	-1.75	-13.80	18.50	10.00	-0.61	0.05	-0.07
Wettzell	-15.80	17.10	7.10	-1.36	-0.46	-1.75	-15.70	17.20	8.70	-0.72	-0.33	-0.62

6.4 The computation of geodetic strain rate tensor

With the prepared residual velocity of every ITRF series realization in the central Mediterranean and Western Europe we can now calculate and analyse the strain rate tensor in the two- and three-dimensional case with the methods introduced in the following sections.

6.4.1 The two-dimensional geodetic stain rate case

The main objective of this study is to analyze the eigenspace component parameters of the two-dimensional strain rate tensor, which are derived from the two-dimensional horizontal residual velocities on the selected sites of the ITRF92 to ITRF 2000.

When we select geodetic sites as vertices of convex polygons we can evaluate the strain tensor of the polygon by using the horizontal velocities. Let $[V_{E_i}\ V_{N_i}]'$ be the known horizontal residual velocity vector (6.5) of the polygon vertex '*i*' along the East and North directions on the local geodetic coordinate system; the following approximation can be written (*Devoti* et al. 2002a)

$$\begin{bmatrix} V_{E_i} \\ V_{N_i} \end{bmatrix} = \begin{bmatrix} V_{E_B} \\ V_{N_B} \end{bmatrix} + \begin{bmatrix} \frac{\partial V_E}{\partial E} & \frac{\partial V_E}{\partial N} \\ \frac{\partial V_N}{\partial E} & \frac{\partial V_N}{\partial N} \end{bmatrix} \begin{bmatrix} \Delta E_i \\ \Delta N_i \end{bmatrix}, \quad \Leftrightarrow \quad \mathbf{V}_{EN_i} = \mathbf{V}_{EN_B} + \mathbf{L}\,\Delta\mathbf{X}_{EN} \tag{6.6}$$

where $[V_{E_B}\ V_{N_B}]'$ is the unknown velocity vector for a reference internal point '*B*', $\mathbf{L}$ the velocity gradient tensor, $[\Delta E_i\ \Delta N_i]'$, the coordinate difference between the site '*i*' and the reference point '*B*', computed, respectively, as parallel and meridian arc length. For the infinitesimal strain rate we have the strain rate tensor $\mathbf{T} = (\mathbf{L} + \mathbf{L}^T)/2$ and the rotation rate tensors $\mathbf{R} = (\mathbf{L} - \mathbf{L}^T)/2$ (*Dermanis* 2001).

The approximation assumes a linear variation of the velocity components with respect to their coordinate differences. This holds true as long as the polygons are properly chosen, not only in terms of area, but also of expected tectonic behavior. The constant space gradients assumption is just a first order approximation of the underlying tectonic setting. *Savage et al.* (2001) give the formulation for estimating strain and rotation rates in a spherical coordinate system. The spherical solution gives insignificantly different results compared to the Cartesian approximation (6.6) for networks, such as a triangulation, which is several hundred kilometres in aperture.

Since the continuous of velocity field $\mathbf{V}_r$ is unknown, but only the discrete values at points '*i*' are known, we have to use an interpolation method to obtain the velocity field $\mathbf{V}_r$ at any other point. There are many interpolation methods, such as (*Dermanis* 2001): (a) *Finite Element Method (FEM)*-Linear interpolation within each triangle (*Grafarend* 1986, *Straub* 1996); (b) *Interpolation using basis functions* (e.g. *Haines and Holt* 1993) and (c) *Collocation* (minimum norm interpolation with infinite basis functions) (e.g. *Straub and Kahle* 1997). Here we will apply the Finite Element Method to do the linear interpolation within each triangle, which is optimally generated by the *Delaunay-triangulation method* among our selected 9 stations. The characteristics of Delaunay-triangulation are that: (1) no triangle side is cut by another; and (2) no points are contained in any other triangle's circumscribed circle. The Delaunay-triangulation of our selected ITRF site is plotted in Figure 6.4.

For every triangle we select the centroid as the reference point, from which it is very easy to compute the velocity gradient tensor at the centroid in a very straightforward way: in fact, dealing with three velocity vectors, the problem is solved by inverting a system of linear equations with six unknowns (four tensor components plus two velocity components):

$$\begin{bmatrix} V_{E_1} \\ V_{N_1} \\ V_{E_2} \\ V_{N_2} \\ V_{E_3} \\ V_{N_3} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \Delta E_1 & \Delta N_1 & 0 & 0 \\ 0 & 1 & 0 & 0 & \Delta E_1 & \Delta N_1 \\ 1 & 0 & \Delta E_2 & \Delta N_2 & 0 & 0 \\ 0 & 1 & 0 & 0 & \Delta E_2 & \Delta N_2 \\ 1 & 0 & \Delta E_3 & \Delta N_3 & 0 & 0 \\ 0 & 1 & 0 & 0 & \Delta E_3 & \Delta N_3 \end{bmatrix} \begin{bmatrix} V_{E_B} \\ V_{N_B} \\ \frac{\partial V_E}{\partial E} \\ \frac{\partial V_E}{\partial N} \\ \frac{\partial V_N}{\partial E} \\ \frac{\partial V_N}{\partial N} \end{bmatrix} \tag{6.7}$$

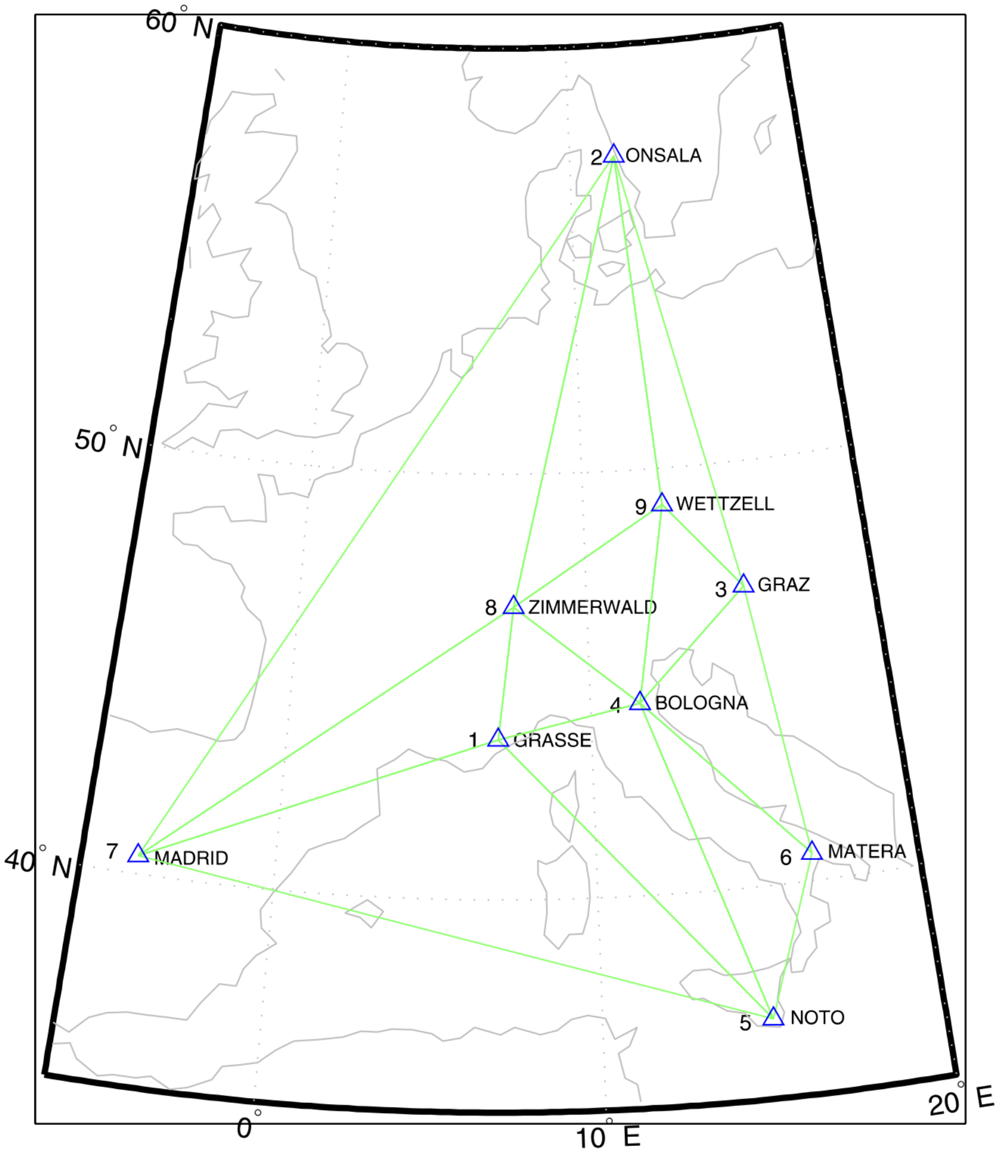


Figure 6.4 The Delaunay-triangulation of the selected ITRF sites

This approach was first proposed by *Terada and Miyabe* (1929). With the velocity gradient tensor we can derive the two-dimensional symmetric strain rate tensor **T** and the antisymmetric rotation rate tensor **R** at the centroid of the discussed Delaunay-triangle network.

$$\mathbf{T} = \frac{1}{2}(\mathbf{L}+\mathbf{L}') = \begin{bmatrix} \frac{\partial V_E}{\partial E} & \frac{1}{2}(\frac{\partial V_E}{\partial N}+\frac{\partial V_N}{\partial E}) \\ \frac{1}{2}(\frac{\partial V_N}{\partial E}+\frac{\partial V_E}{\partial N}) & \frac{\partial V_N}{\partial N} \end{bmatrix}, \tag{6.8}$$

$$\mathbf{R} = \frac{1}{2}(\mathbf{L}-\mathbf{L}') = \begin{bmatrix} 0 & \frac{1}{2}(\frac{\partial V_E}{\partial N}-\frac{\partial V_N}{\partial E}) \\ \frac{1}{2}(\frac{\partial V_N}{\partial E}-\frac{\partial V_E}{\partial N}) & 0 \end{bmatrix}. \tag{6.9}$$

With (6.7), (6.8) and (6.9) we could compute the geodetic strain rates of every Delaunay-triangle for six ITRF relations, and successively the eigenspace components (eigenvalues and eigendirection) together with the maximum shear strain rate $\varepsilon_1 - \varepsilon_2$ and the second strain rate invariant $(\varepsilon_1^2 + \varepsilon_2^2)^{1/2}$. With reference to the continued discussion in Section 6.6, we have only listed the results for two of the 11 triangles in Table 6.4.

Table 6.5 Strain rate tensor components, eigenspace components, max. shear strain rates and second invariant

Epochs	Strain rate tensor components (nanostrain/yr)			Eigenspace components (nanostrain/yr)		(degree)	max. shear strain rate (nanostrain/yr)	second invariant (nanostrain/yr)
	t_{11}	t_{12}	t_{22}	ε_1	ε_2	α_1	$\varepsilon_1 - \varepsilon_2$	$\sqrt{\varepsilon_1^2 + \varepsilon_2^2}$
	Triangle 4 - 5 - 6 (Bologna – Noto - Matera) (No.10)							
ITRF92	1.6621	2.4987	-5.7591	2.4250	-6.5220	16.9781	8.9470	6.9582
ITRF93	6.9357	3.3247	-4.8325	7.8100	-5.7068	14.7339	13.5168	9.6728
ITRF94	8.1550	2.4147	-4.8161	8.5900	-5.2511	10.2108	13.8411	10.0679
ITRF96	0.5796	5.2419	-3.0502	4.3120	-6.7825	35.4514	11.0945	8.0371
ITRF97	1.5487	4.2150	-1.8819	4.3840	-4.7172	33.9280	9.1012	6.4398
ITRF2000	2.8365	5.3361	-1.0826	6.5615	-4.8076	34.9179	11.3691	8.1343
	Triangle 1 - 4 - 8 (Grass – Bologna - Zimmerwald) (No.6)							
ITRF92	1.6460	8.5479	12.4127	17.1312	-3.0725	61.1010	20.2037	17.4045
ITRF93	-1.8824	7.5296	13.2853	16.3883	4.9854	67.6029	11.4029	17.1298
ITRF94	3.0908	3.2370	9.2711	10.6561	1.7058	66.8353	8.9503	10.7918
ITRF96	8.8276	-0.5745	1.6884	8.8735	1.6424	-4.5712	7.2311	9.0242
ITRF97	7.4760	2.6775	1.7626	8.5346	0.7040	21.5726	7.8306	8.5636
ITRF2000	6.1801	2.6301	0.4793	7.2081	-0.5488	21.3493	7.7569	7.2290

6.4.2 The three-dimensional geodetic stain rate case

In fact, most tensors in Geodesy and Geophysics are three-dimensional and have been derived from geodetic, geological and seismological data. As most popular example is the in-situ measurements of the strain tensor by a strain meter and seismic moment tensor by the seismometer. Here we would like to directly apply the three dimensional Cartesian residual velocities of the six ITRF series realizations in determining the three-dimensional strain rate tensor and rotation rate tensor.

When we select geodetic sites as vertices of convex polygons we can evaluate the strain tensor of the polygon by using the residual velocities. Let $[V_{X_i} \; V_{Y_i} \; V_{Z_i}]'$ be the known residual velocity vector of the polygon vertex 'i' along the Cartesian coordinate directions (6.3); the following approximation can be written:

$$\begin{bmatrix} V_{X_i} \\ V_{Y_i} \\ V_{Z_i} \end{bmatrix} = \begin{bmatrix} V_{X_B} \\ V_{Y_B} \\ V_{Z_B} \end{bmatrix} + \begin{bmatrix} \frac{\partial V_X}{\partial X} & \frac{\partial V_X}{\partial Y} & \frac{\partial V_X}{\partial Z} \\ \frac{\partial V_Y}{\partial X} & \frac{\partial V_Y}{\partial Y} & \frac{\partial V_Y}{\partial Z} \\ \frac{\partial V_Z}{\partial X} & \frac{\partial V_Z}{\partial Y} & \frac{\partial V_Z}{\partial Z} \end{bmatrix} \begin{bmatrix} \Delta X_i \\ \Delta Y_i \\ \Delta Z_i \end{bmatrix}, \quad \Leftrightarrow \quad \mathbf{V}_{\mathbf{X}_i} = \mathbf{V}_{\mathbf{X}_B} + \mathbf{L}\,\Delta\mathbf{X}_i \tag{6.10}$$

where $[V_{X_B} \; V_{Y_B} \; V_{Z_B}]'$ is the unknown velocity vector for a reference internal point 'B', $\mathbf{L}$ the velocity gradient tensor, $[\Delta X_i \; \Delta Y_i \; \Delta Z_i]'$, the Cartesian coordinate difference vector between the site 'i' and the reference point 'B'. For the infinitesimal strain rate we have the strain rate tensor $\mathbf{T} = (\mathbf{L} + \mathbf{L}')/2$ and the rotation rate tensors $\mathbf{R} = (\mathbf{L} - \mathbf{L}')/2$ (*Dermanis* 2001).

The approximation assumes a linear variation of the velocity components with respect to their coordinate differences. This holds true as long as the polygons are properly chosen, not only in term of area but also of expected tectonic behavior. The constant space gradients assumption is only a first order approximation of the underlying tectonic setting.

If we choose to work with a tetragon, and we elect as the reference point the barycenter, it is very easy to compute the velocity gradient tensor at the barycenter in a very straightforward way: in fact, dealing with four velocity vectors, the problem is solved by inverting a system of linear equations with twelve unknowns (nine tensors components plus three velocity components):

$$\begin{bmatrix} V_{X_1} \\ V_{Y_1} \\ V_{Z_1} \\ \vdots \\ V_{X_4} \\ V_{Y_4} \\ V_{Z_4} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \Delta X_1 & \Delta Y_1 & \Delta Z_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & \Delta X_1 & \Delta Y_1 & \Delta Z_1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & \Delta X_1 & \Delta Y & \Delta Z_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & \Delta X_4 & \Delta Y_4 & \Delta Z_4 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & \Delta X_4 & \Delta Y_4 & \Delta Z_4 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & \Delta X_4 & \Delta Y_4 & \Delta Z_4 \end{bmatrix} \begin{bmatrix} V_{X_B} \\ V_{Y_B} \\ V_{Z_B} \\ \mathrm{vec}\,\mathbf{L}' \end{bmatrix}, \quad (6.11)$$

where

$$\mathrm{vec}\,\mathbf{L}' = [\frac{\partial V_X}{\partial X}\ \frac{\partial V_X}{\partial Y}\ \frac{\partial V_X}{\partial Z}\ \frac{\partial V_Y}{\partial X}\ \frac{\partial V_Y}{\partial Y}\ \frac{\partial V_Y}{\partial Z}\ \frac{\partial V_Z}{\partial X}\ \frac{\partial V_Z}{\partial Y}\ \frac{\partial V_Z}{\partial Z}]'.$$

With the velocity gradient tensor we can compute the three-dimensional symmetric strain rate tensor **T** and the antisymmetric rotation rate tensor **R** at the barycenter of the discussed tetragonal network.

$$\mathbf{T} = \frac{1}{2}(\mathbf{L}+\mathbf{L}') = \begin{bmatrix} \frac{\partial V_X}{\partial X} & \frac{1}{2}(\frac{\partial V_X}{\partial Y}+\frac{\partial V_Y}{\partial X}) & \frac{1}{2}(\frac{\partial V_X}{\partial Z}+\frac{\partial V_Z}{\partial X}) \\ \frac{1}{2}(\frac{\partial V_Y}{\partial X}+\frac{\partial V_X}{\partial Y}) & \frac{\partial V_Y}{\partial Y} & \frac{1}{2}(\frac{\partial V_Y}{\partial Z}+\frac{\partial V_Z}{\partial Y}) \\ \frac{1}{2}(\frac{\partial V_Z}{\partial X}+\frac{\partial V_X}{\partial Z}) & \frac{1}{2}(\frac{\partial V_Z}{\partial Y}+\frac{\partial V_Y}{\partial Z}) & \frac{\partial V_Z}{\partial Z} \end{bmatrix}, \quad (6.12)$$

$$\mathbf{R} = \frac{1}{2}(\mathbf{L}-\mathbf{L}') = \begin{bmatrix} 0 & \frac{1}{2}(\frac{\partial V_X}{\partial Y}-\frac{\partial V_Y}{\partial X}) & \frac{1}{2}(\frac{\partial V_X}{\partial Z}-\frac{\partial V_Z}{\partial X}) \\ \frac{1}{2}(\frac{\partial V_Y}{\partial X}-\frac{\partial V_X}{\partial Y}) & 0 & \frac{1}{2}(\frac{\partial V_Y}{\partial Z}-\frac{\partial V_Z}{\partial Y}) \\ \frac{1}{2}(\frac{\partial V_Z}{\partial X}-\frac{\partial V_X}{\partial Z}) & \frac{1}{2}(\frac{\partial V_Z}{\partial Y}-\frac{\partial V_Y}{\partial Z}) & 0 \end{bmatrix}. \quad (6.13)$$

With (6.11), (6.12) and (6.13) we have computed the strain rate tensors for the six series of ITRF realizations in the selected sub-network of four sites: 1 - 4 - 8 - 9 (Grasse - Bologna - Zimmerwald - Wettzell) which are listed in Table 6.6.

Table 6.6 Three-dimensional strain rate tensor components of the sub-network of site 1 – 4 – 9 – 8 (Grass – Bologna - Wettzell - Zimmerwald)

Epochs	Strain rate tensor components (1×10^{-7} strain/yr)					
	t_{11}	t_{12}	t_{13}	t_{22}	t_{23}	t_{33}
ITRF92	0.1746	0.9012	2.6401	0.2946	1.4380	5.4648
ITRF93	-2.4174	0.0574	0.2518	0.0558	0.6080	3.3212
ITRF94	-1.6029	-0.0626	0.4812	0.0478	0.3324	2.8452
ITRF96	1.1197	0.3782	2.0597	0.1845	0.5407	3.0706
ITRF97	-0.7771	0.1652	0.2373	0.1466	0.3868	1.3365
ITRF2000	-1.6829	0.2205	-0.9884	0.1735	0.3919	-0.2520

6.5 The representation of the numerical results of 2-D geodetic strain rate and its interpretation

Now we can present the horizontal residual velocities and the principal strain rates of every triangle for six ITRF realizations derived in the second phase. The interpretation of the geodetic strain also results from these six ITRF realizations, and a comparison with the geodynamical setting will follow.

The residual velocities and principal strain rates have to be represented in an appropriate way for any further interpretations and comparisons. We used the MATLAB Mapping Toolbox (*MathWorkd* 2000) to map the sur-face deformation information. The *equidistant conic projection* was described by the Alexandrian astronomer, mathematician and geographer *Claudius Ptolemy* about A.D. 150. Improvements were developed by *Johannes Ruysch* in 1508, *Gerardus Mercator* in the late 16th century, and *Nicolas de l'Isle* in 1745. It is also known as the *Simple Conic* or *Conic projection*. The scale is true along all meridians and the standard parallels. It is constant along any parallel. This projection is free of distortion along the two standard parallels. Distortion is constant along any other parallel. This projection provides a compromise in distortion between conformal and equal-area conic projections, of which it is neither.

The pattern of the principal strain rates (eigenvalues and eigendirections of the 2-D strain rate tensors) of 11 Delaunay-triangles and the associated residual velocities of the selected ITRF92 to ITRF2000 sites in the study region are illustrated in Figures 6.5 to 6.10. *Extension* is represented by a *red* symmetric arrow, and *contraction* is represented by a *blue* symmetric arrow. The residual velocities are also represented by *black* arrows.

First, let us shortly analyze the horizontal movements of these selected sites with respect to the Eurasian fixed plate. The selected six stations Grass, Onsala, Wettzell, Madrid, Graz and Zimmerwald belong to the stable European plate. They have smaller residual velocities with respect to the Eurasia plate described by, e.g., the NNR-NUVEL-1A model, i.e. the residual velocities of these six sites in ITRF97 are below the level of 1.5 mm/yr; for more detail see Figure 6.9. The three Italian stations, Bologna (Medicina), Matera and Noto, all show motions with respect to stable Europe. The two mainland sites, Bologna (Medicina) and Matera are east of the Apennine mountain chain and have north-east-trending velocities with rates increasing southward from, e.g. in ITRF97, 2.9 mm/yr to 4.8 mm/yr. In contrast, Noto, Sicily, just a few hundred km south-west of Matera is moving with rates of 4.6 mm/yr in north-west direction at an apparent angle to the other Italian sites. These significant residual velocities of the three Italy sites reflect the fact that the movements of these sites located in the plate boundary zone between Eurasia and Africa don't agree with NNR NUVEL-1A.

Secondly let us analyze the strain rate solutions of the six epochs. From the derived strain rates from (6.7) and the visual patterns in Figures 6.5 ~ 6.10, we can see that the magnitude and direction of the strain rates in most triangles are nearly consistent with each epoch, except for the triangles 1-4-8 and 1-7-8. As we have explained above, the deformation pattern of this region is characterized by a complex space-time distribution of compressional and tensional events.

We have to compare them with the geodynamic features in detail. Since this study is concentrated on the statistical inference of the eigenspace components of strain rate tensors, we limit our comparison of the intensive seismic activity in the Alpine-Mediterranean regions to the result of ITRF2000.

From Figure 6.10 we can learn that the Bologna-Matera-Noto triangle (4-5-6) suffers from ENE-WSW extension strain with a rate of 5.29 nanostrain/yr, which consists of the seismic strain rate derived from *Centroid Moment Tensor* (CMT) solutions (*Pondrelli et al.* 1995) not only in the maximum principal seismic strain rate of about 5 nanostrain/yr but also in the principal direction. Furthermore, the direction of the extensional strain rate is in accordance with the tectonic evolution of this region (Apenninicis), which is strongly affected by the convergence of the microplate (*Ward* 1994). Our geodetic strain rate results in this subregion are also in accordance with other geodynamic solutions deduced from a fault plane solution of earthquakes that occurred during the last century (*Jackson & McKenzie* 1988), historical seismicity (*Selvaggi* 1998) and the newly published geodetic results by *Anzidei* (2001), *Devoti* (2002a, b), *Caporali* (2003) and *Jimenez-Munt* (2003).

In the Western Mediterranean area, which is covered by the triangle Grass-Noto-Madrid (1-5-7), compression predominates in the NNW direction, which is in good agreement with the observed stress data (*Ward* 1994, *Montone et al*. 1999, *Jimenez-Munt et al*. 2001). This compression is consistent with the view that it is induced by the relative motion between Africa and Eurasia (*DeMets et al*.1994). The geodetic E-W extension is in accordance with the extensional tectonics perpendicular to the Apenninic chain, indicated by the normal fault events. The observed extension which is perpendicular to the chain could indicate that the subduction is also active undernearth the central Apennines. This pattern is in accordance with the radial stress regime proposed by *Rebai* et al. (1992).

The triangle Bologna-Matera-Graz (4-6-3) represents the strain across the Adriatic microplate. The geodetically observed North and East contractions are in good agreement with the northward motion of the Adriatic microplate with respect to Europe (*Devoti et al.* 2002a), and they are also consistent with the compressional stress pattern in this region (*Müller et al.* 1992, *Montone et al.* 1999). The northern triangle Bologna (Medicina)-Graz-Wettzell (3-4-9) represents the strain across the Eastern Alps with a smaller strain rate. The observed NNE contraction is in agreement with the main geological structures, i.e., the subduction of Adriatic microplate to the Alpine front.

At the end of our comparison in the Alpine-Mediterranean regions, we present two published figures, namely in Figure 6.11 which is a general geological map (*Devoti et al.* 2002a), and Figure 6.12 which is the average of active stress map (*Montone et al*. 1999) of the Alpine-Mediterranean regions; both are in agreement with our geodetic strain rate results. On the other hand it is shown that our results are good reproductions of the geodynamical setting in this regions.

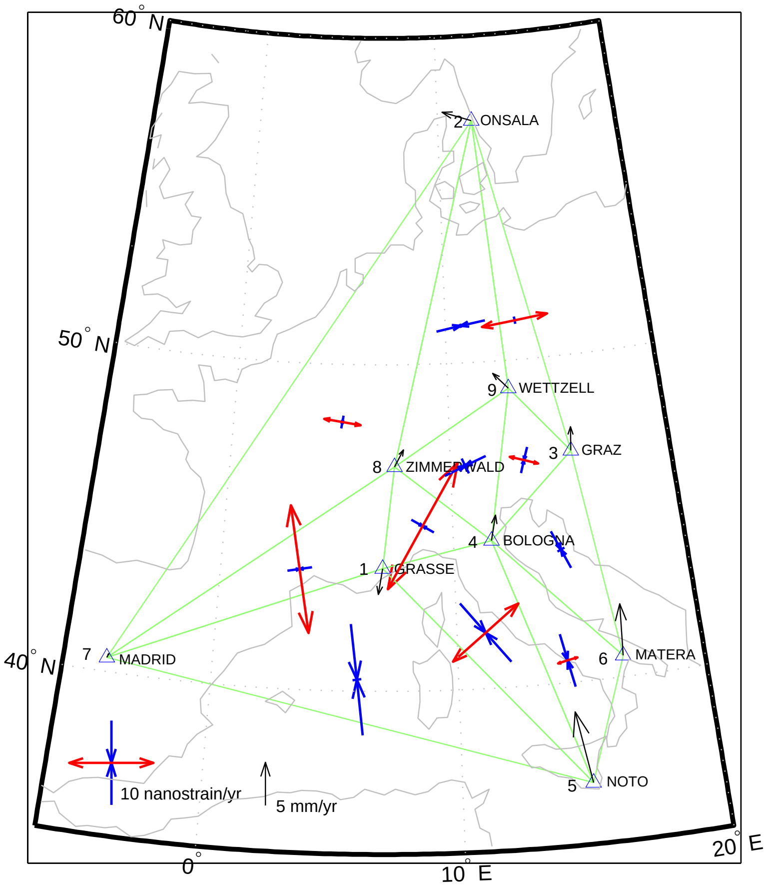


Figure 6.5. Pattern of the principal strain rates of nine triangles and the associated residual velocities of the selected ITRF92 sites in the studying region. *Extension* is represented by symmetric arrows pointing out and *contraction* is represented by symmetric arrows pointing in. The residual velocities are represented by *single* arrows.

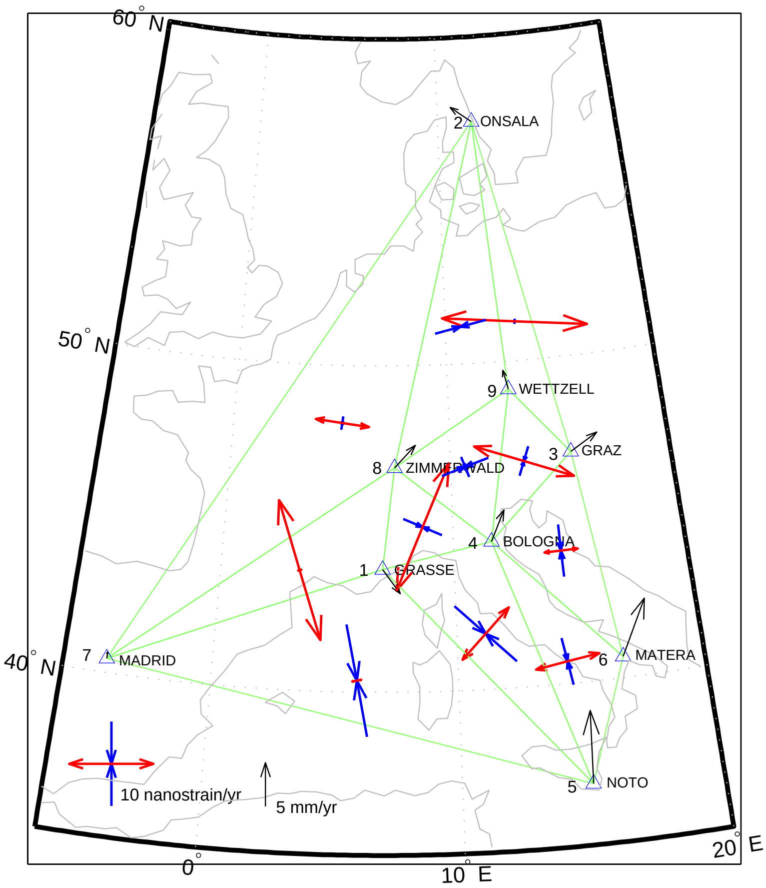


Figure 6.6. Pattern of the principal strain rates of nine triangles and the associated residual velocities of the selected ITRF93 sites in the studying region. *Extension* is represented by symmetric arrows pointing out and *contraction* is represented by symmetric arrows pointing in. The residual velocities are represented by *single* arrows.

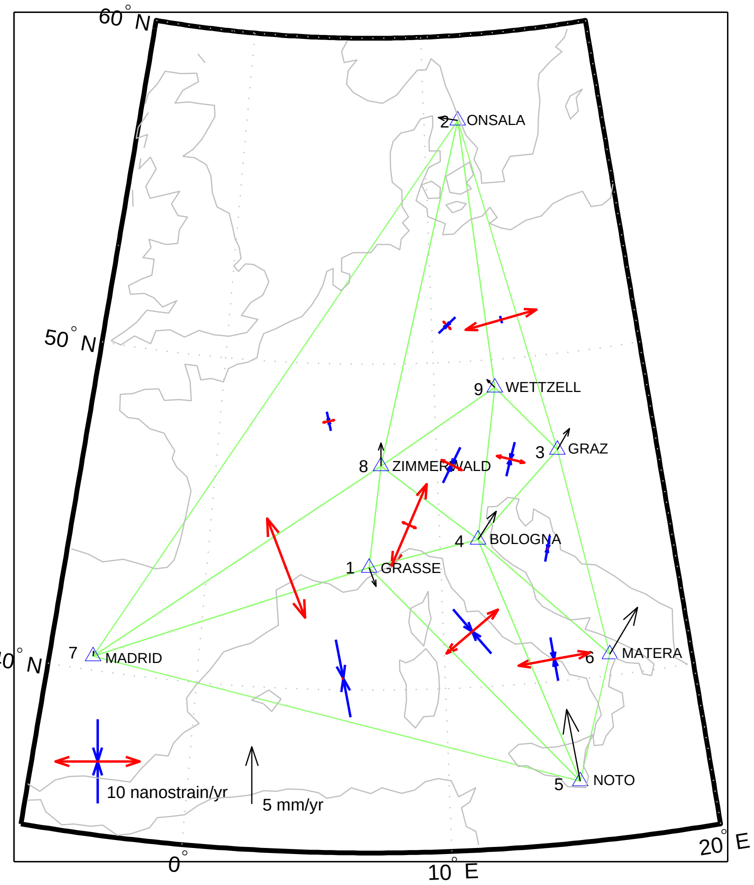


Figure 6.7. Pattern of the principal strain rates of nine triangles and the associated residual velocities of the selected ITRF94 sites in the studying region. *Extension* is represented by symmetric arrows pointing out and *contraction* is represented by symmetric arrows pointing in. The residual velocities are represented by *single* arrows.

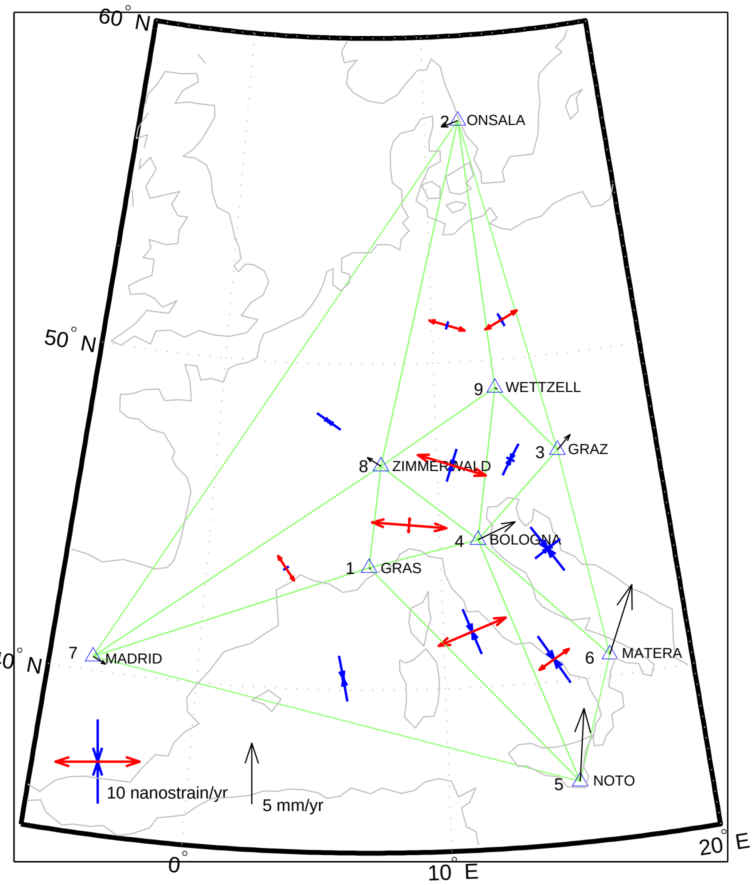


Figure 6.8. Pattern of the principal strain rates of nine triangles and the associated residual velocities of the selected ITRF96 sites in the studying region. *Extension* is represented by symmetric arrows pointing out and *contraction* is represented by symmetric arrows pointing in. The residual velocities are represented by *single* arrows.

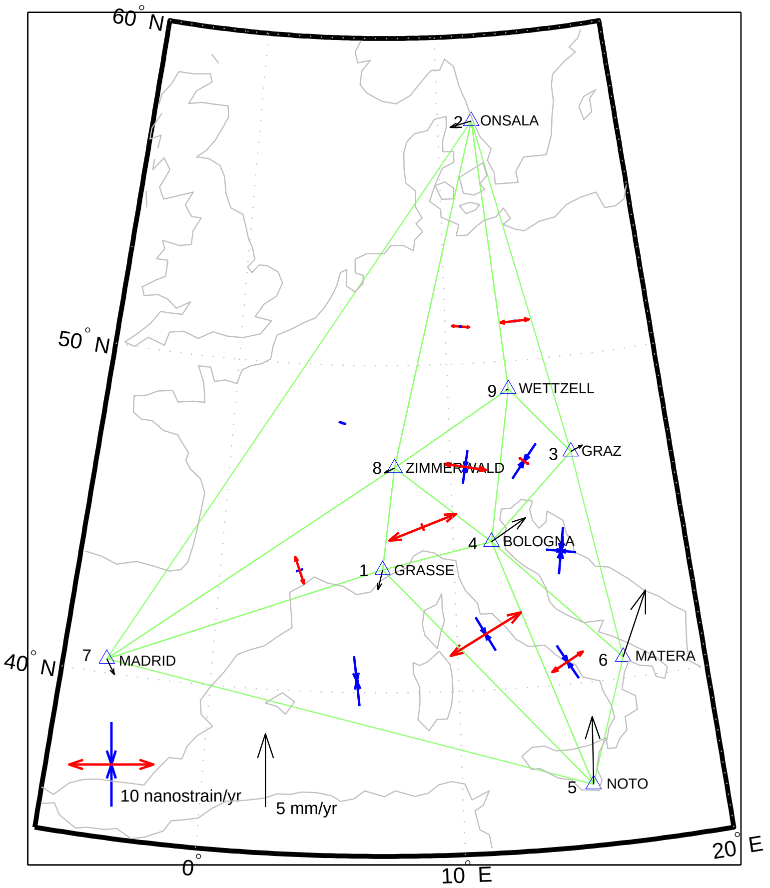


Figure 6.9. Pattern of the principal strain rates of nine triangles and the associated residual velocities of the selected ITRF97 sites in the studying region. *Extension* is represented by symmetric arrows pointing out and *contraction* is represented by symmetric arrows pointing in. The residual velocities are represented by *single* arrows.

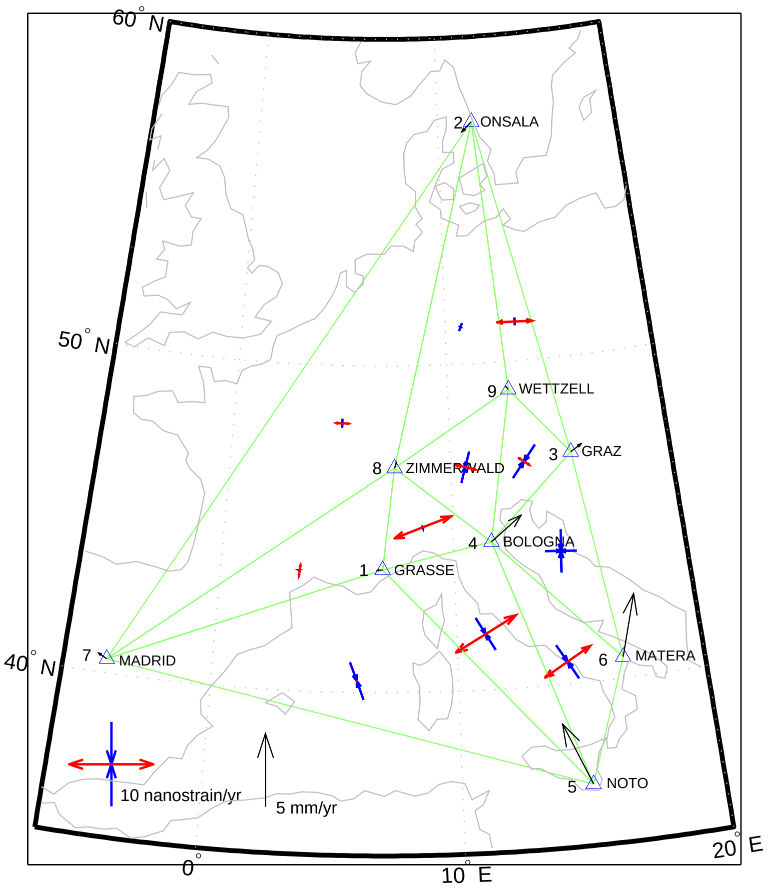


Figure 6.10. Pattern of the principal strain rates of nine triangles and the associated residual velocities of the selected ITRF2000 sites in the studying region. *Extension* is represented by symmetric arrows pointing out and *contraction* is represented by symmetric arrows pointing in. The residual velocities are represented by *single* arrows.

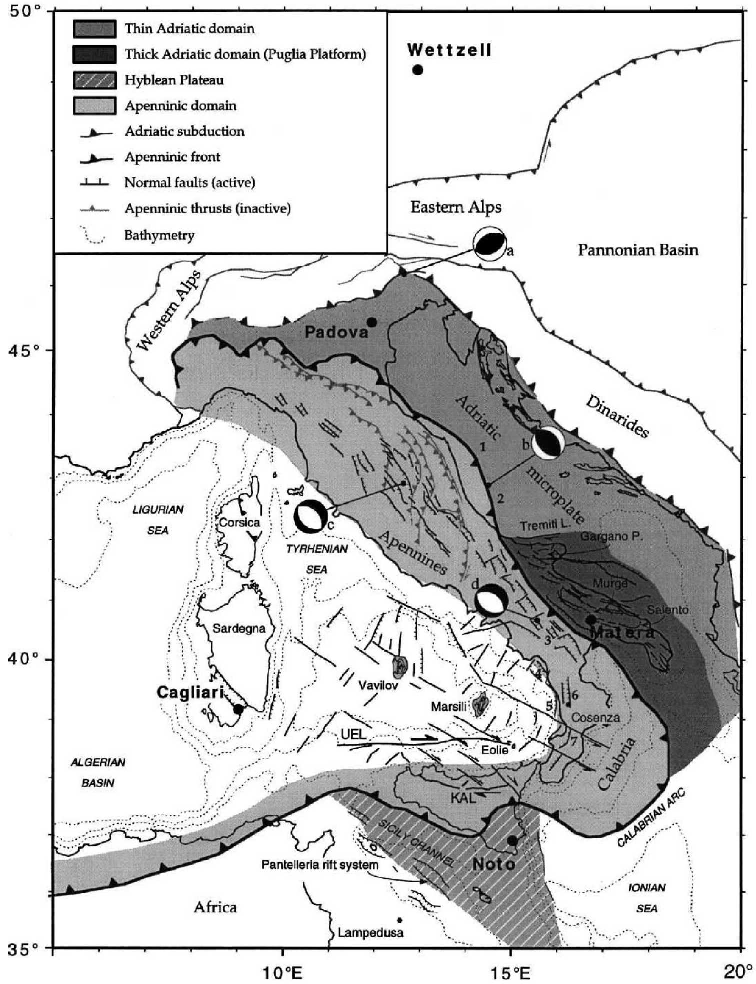


Figure 6.11 Simplified geological structural map of the central Mediterranean region showing the main tectonic domains (*Devoti et al*. 2002b)

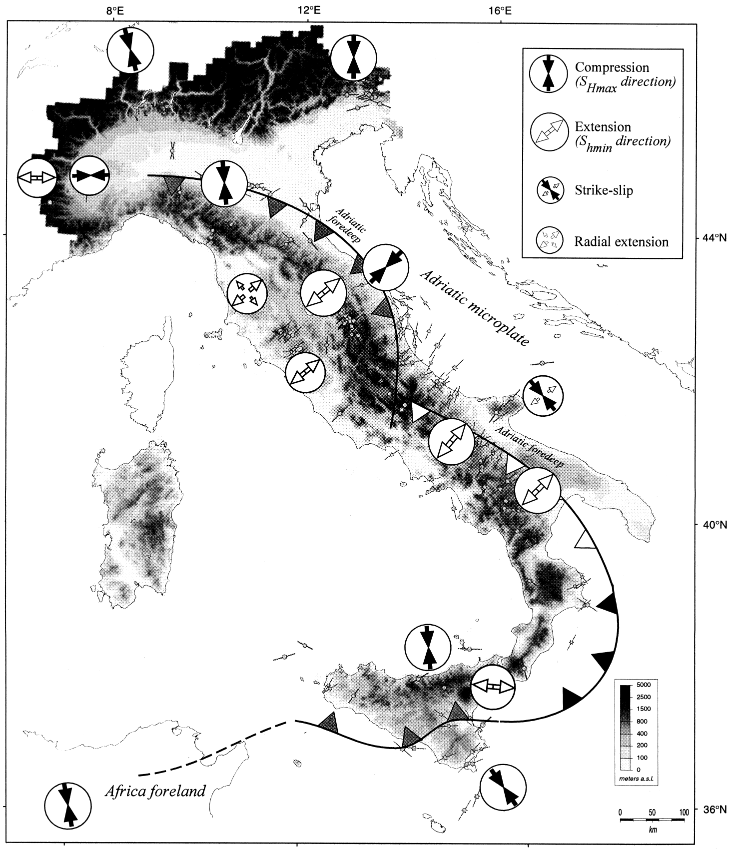


Figure 6.12 The average active stress map of Italy. The structural arcs with shaded triangles indicate the active compressional fronts; solid triangles show active oceanic subduction; open triangles delineate the location of Pilo-Pleistocene thrust front, presently affected by prevalent extension (*Montone et al.* 1999)

6.6 Statistical inference of the eigenspace components of 2-D strain rate tensor

In this section, as a case study, both model and hypothesis tests developed in Chapter 4 will be applied to the observations of random strain rate tensors, derived above for every Delaunay-triangles of the selected ITRF sites at six epochs.

6.6.1 The estimates of the eigenspace components from the strain rates observations of six epochs

With the two-dimensional strain rate tensor observations, calculated by (6.7) with the six epoch ITRF residual velocities, we can now estimate the eigenspace components (eigenvalues and eigendirections) of the two-dimensional strain rate tensors, variance-covariance component matrix of type BIQUUE, and their estimated dispersion matrix with (4.29), (4.31) and (4.30), and successively make hypothesis tests. The detailed results of all 11 Delaunay-triangles in the study region are illustrated in Figure 6.13 together with their 95% confidence intervals. The estimates of the eigenspace components (eigenvalues and eigendirection) and their standard deviations together with the maximum shear strain rate $\hat{\lambda}_1 - \hat{\lambda}_2$ and the second strain rate invariant $(\hat{\lambda}_1^2 + \hat{\lambda}_2^2)^{1/2}$ for these triangles are listed in Table 6.7.

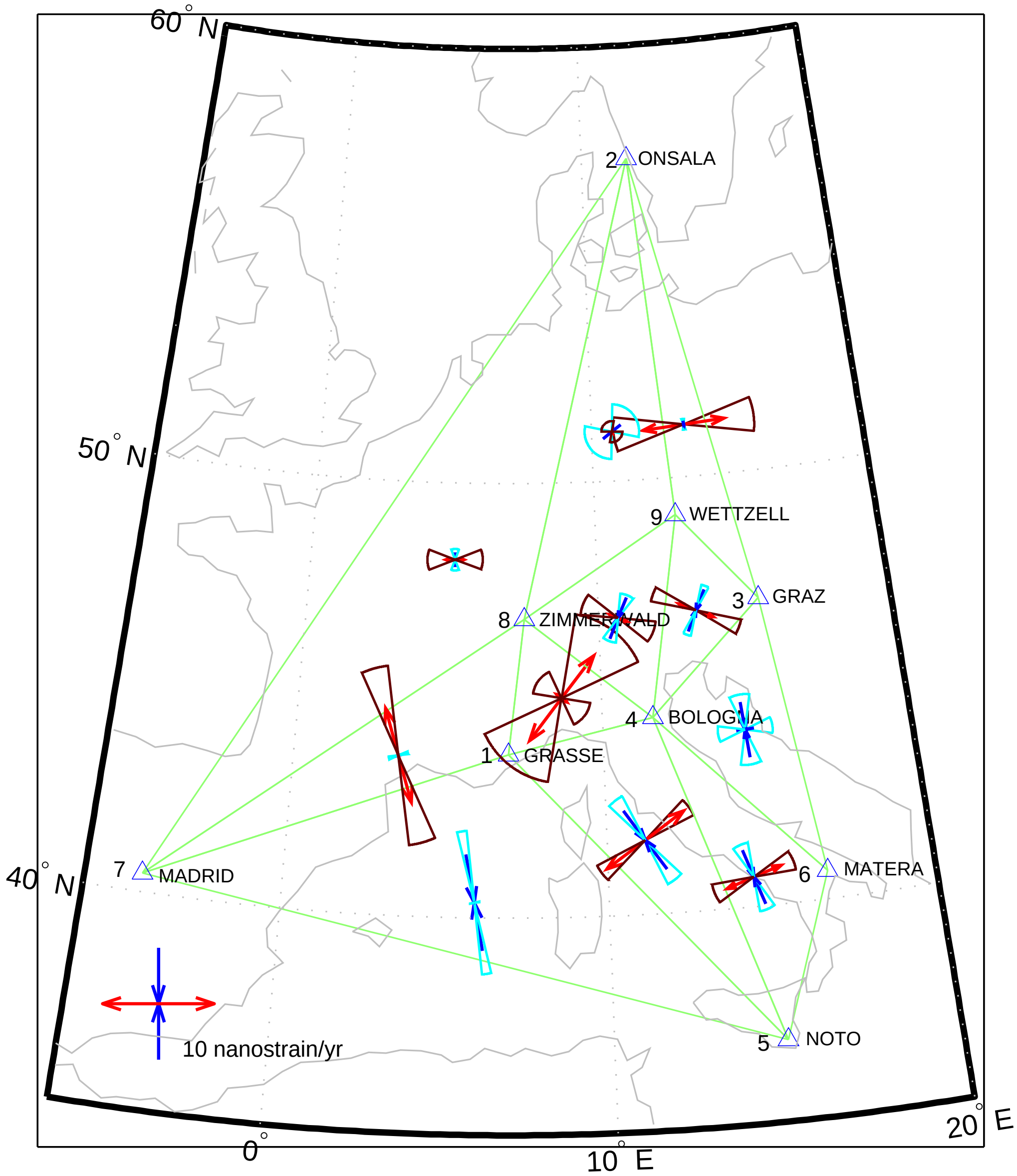


Figure 6.13. Eigenspace components (eigenvalues and eigendirections) of the two-dimensional strain rate tensors and their 95% confidence intervals, estimated from the strain rate observations of ITRF92 to ITRF2000 series in the nine triangle sites in the study region. *Extension* is represented by *red* symmetric arrow and *contraction* is represented by *blue* symmetric arrow.

Table 6.7 The estimates of the eigenspace components and their standard deviations together with the maximum shear strain rates and the second invariant

Triangles	Eigenspace components and standard deviations (nanostrain/yr)				(degree)		max. shear strain rate (nanostrain/yr)	second invariant (nanostrain/yr)
	$\hat{\lambda}_1$	$\hat{\sigma}_{\hat{\lambda}_1}$	$\hat{\lambda}_2$	$\hat{\sigma}_{\hat{\lambda}_2}$	$\hat{\alpha}_1$	$\hat{\sigma}_{\hat{\alpha}_1}$	$\hat{\lambda}_1 - \hat{\lambda}_2$	$\sqrt{\hat{\lambda}_1^2 + \hat{\lambda}_2^2}$
No. 1: 8 - 2 – 7 (Zim - Ons - Mad)	1.6176	±1.3012	-1.2824	±0.2501	0.1681	±8.2284	2.9000	2.0642
No. 2: 1 - 5 – 7 (Gras - Not - Mad)	-0.1637	±0.2973	-8.7116	±1.6321	9.9241	±1.4835	8.5479	8.7131
No. 3: 1 - 8 - 7 (Gras - Zim - Mad)	8.7144	±2.8800	-0.6217	±0.4370	-74.6070	±3.3523	9.3361	8.7366
No. 4: 9 - 8 - 2 (Wet - Zim- Ons)	0.2244	±0.6497	-2.0148	±1.1151	-50.7307	±19.5934	2.2392	2.0273
No. 5: 9 - 3 - 2 (Wet - Graz - Ons)	7.3502	±2.0722	-0.6027	±0.1422	8.5718	±5.4747	7.9529	7.3749
No. 6: 4 - 1 - 8 (Bol - Gras - Zim)	9.5173	±2.1984	1.1889	±1.5589	52.8733	±10.8772	8.3284	9.5913
No. 7: 4 - 9 - 8 (Bol - Wet - Zim)	1.6742	±1.9740	-3.9692	±0.1628	-21.6686	±6.4597	5.6434	4.3079
No. 8: 4 - 9 - 3 (Bol - Wet - Graz)	3.3182	±1.9163	-4.0195	±0.2443	-20.7389	±3.7226	7.3376	5.2121
No. 9: 4 - 3 – 6 (Bol - Graz - Mat)	-1.6041	±1.2931	-5.0069	±0.5130	10.1413	±6.5735	3.4028	5.2575
No. 10: 4 - 5 – 6 (Bol - Not - Mat)	5.2837	±0.9064	-5.2345	±0.3867	23.4382	±5.1932	10.5182	7.4376
No. 11: 4 - 1 - 5 (Bol - Gras - Not)	8.6068	±0.4504	-6.4977	±0.9156	37.1679	±3.8370	15.1045	10.7841

The 95% confidence intervals for the estimates of eigenvalues $\hat{\lambda}_1$, $\hat{\lambda}_2$ and eigendirection $\hat{\alpha}_1$ illustrated in Figure 6.13 provide us with a visual presentation of the possible magnitude and the directions of the extension and contraction of the strain rate. This is important for the prediction of the tectonic activity, including the possible deformation trend and its directions. For example the larger error (confidence interval) of eigenvalues and eigendirection of strain rates, detected obviously in the triangle 1-4-8 (Grasse-Bologna-Zimmerwald), results from the variety of strain rates observations of the six epochs. As illustrated in Figures 6.5 to 6.10, the principal direction of strain rate among the six epochs changes from NNE in ITRF92, 93 and to W-E in ITRF96 and to ENE in ITRF97/ITRF2000, more detailed results are also found in Table 6.5. This fact reflects that the deformation pattern in this triangle area which proves is not stable during the six epochs from 1992 to 2000.

It is necessary to note that, although the strain rate tensor observations are derived from the nine ITRF sites according to the criterion discussed above, in reality they don't satisfy all the conditions of i.i.d. observations, since we have not yet found the right i.i.d. strain tensor observation sets, we apply strain rate tensor observations derived from the nine ITRF stations in six series realizations, assuming approximately that they are i.i.d. observations in our study.

When we repeat the comparison of our strain rate pattern in Figure 6.13 with the geodynamic map, e.g. Figures 6.11 and 6.12, it can be concluded in general that our estimates of eigenspace components of a two-dimensional strain rate tensors are consistent with the tectonic setting. Furthermore we can benefit from the statistical information derived from the estimation procedure developed in Chapter 4, which is presented in the next section.

6.6.2 Statistical inference of the estimates of eigenspace component parameters

The estimates of the eigenspace component parameters and their related dispersion matrix from the strain rate observations of six epochs reflect the statistical average information of the random strain rate tensor, utilize the advantage of the longer time span. With them we can successively perform the statistical inference, i.e.

Statistical Inference = Estimate +Hypothesis test.

Since we are interested in testing our statistical method, we would like to perform all the hypothesis tests discussed in Section 4.3 in detail for just one triangle 4-5-6 (Bologna-Noto-Matera). Using the strain rate observations in Table 6.5 we have estimated the eigenspace component parameters of a rank-two symmetric random tensor, their related dispersion matrix with *Theorem* 4.3, and the sample variance-covariance matrix $\hat{\mathbf{\Sigma}}_{\mathbf{y}}$ of type BIQUUE with *Theorem* 4.4; they are summarized in the following *Box 6.1*.

Box 6.1

Case study: Hypothesis test with a 2-dimensional strain rate tensor in Triangles 4-5-6 of sites Bologna-Matera-Noto

"the $\mathbf{\Sigma}$-BLUUE of eigenspace components of a two-dimensional, symmetric rank-two random tensor with *Theorem* 4.3"

$$\hat{\boldsymbol{\xi}} = \begin{bmatrix} \hat{\lambda}_1 \\ \hat{\lambda}_2 \\ \hat{\alpha} \end{bmatrix} = \begin{bmatrix} 5.2837\ (n\text{strain/y}) \\ -5.2345\ (\mu\text{strain/y}) \\ 0.409074 \quad (\text{arc}) \end{bmatrix} = \begin{bmatrix} 5.2837\ (n\text{strain/y}) \\ -5.2345\ (\mu\text{strain/y}) \\ +23°.4382 \end{bmatrix}$$

"the sample variance-covariance matrix $\hat{\mathbf{\Sigma}}_{\mathbf{y}}$ of type BIQUUE with *Theorem* 4.4"

$$\hat{\mathbf{\Sigma}}_{\mathbf{y}} = \begin{bmatrix} +9.908347 & -2.351484 & -2.515398 \\ -2.351484 & +1.688078 & +2.089224 \\ -2.515398 & +2.089224 & +3.449229 \end{bmatrix} (n\text{strain/y})^2$$

"the related dispersion matrix of $\hat{\boldsymbol{\xi}}$ with *Theorem* 4.3"

$$D\{\hat{\boldsymbol{\xi}}\} = \mathbf{\Sigma}_{\hat{\xi}} = \begin{bmatrix} +0.821618 & +0.208317 & -0.040791 \\ +0.208317 & +0.149544 & +0.000599 \\ -0.040791 & +0.000599 & +0.008215 \end{bmatrix}$$

With these estimates of the eigenspace components of the random strain rate tensor and their dispersion matrix, and under the assumption that the observations of a symmetric rank-two random strain rate tensor are *Gauss-Laplace* normally distributed, the following univariate and multivariate hypothesis tests, discussed in Section 4.3, will be performed:

(1) Test for the eigenspace parameter vector $\boldsymbol{\xi} = \boldsymbol{\xi}_0$ with $\mathbf{\Sigma}_{\mathbf{y}}$ unspecified (see *Box 6.2*);
(2) Test for a distinct element of the eigenspace parameter vector with *Student t- test* (see *Box 6.3*);
(3) *Eigen inference* about the orthonormal transformed parameters $\boldsymbol{\eta}$ (see *Box 6.4*);
(4) Test for the variance-covariance matrix $\mathbf{\Sigma}_{\mathbf{y}} = \mathbf{\Sigma}_0$ (see *Box 6.5*);
(5) Test for the eigenspace parameter vector and variance-covariance matrix $\boldsymbol{\xi} = \boldsymbol{\xi}_0, \mathbf{\Sigma}_{\mathbf{y}} = \mathbf{\Sigma}_0$ (see *Box 6.6*);
(6) The general linear hypothesis test with the growth curve model for eigenspace parameters (see *Box 6.7*).

(1) Test for the eigenspace parameter vector $\boldsymbol{\xi} = \boldsymbol{\xi}_0$ with $\mathbf{\Sigma}_{\mathbf{y}}$ unspecified with *Hotelling's* T^2-test

Box 6.2:

Multivariate hypothesis test about the eigenspace parameter vector $\boldsymbol{\xi}$ assuming *Gauss-Laplace* normally distributed observations of a symmetric rank-two random strain rate tensor

First test for $\mathcal{H}_{01}: \ \boldsymbol{\xi} = \boldsymbol{\xi}_0,\ \mathcal{H}_{11}: \ \boldsymbol{\xi} \neq \boldsymbol{\xi}_0$ with $\mathbf{\Sigma}_{\mathbf{y}}$ unspecified;

$$\Leftrightarrow$$

$$\mathcal{H}_{01}: \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \alpha_1 \end{bmatrix} = \begin{bmatrix} \lambda_{10} \\ \lambda_{20} \\ \alpha_{10} \end{bmatrix} = \begin{bmatrix} 4.3840 \\ -4.7172 \\ 0.5922 \end{bmatrix},\ \mathcal{H}_{11}: \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \alpha_1 \end{bmatrix} \neq \begin{bmatrix} \lambda_{10} \\ \lambda_{20} \\ \alpha_{10} \end{bmatrix} = \begin{bmatrix} 4.3840 \\ -4.7172 \\ 0.5922 \end{bmatrix} \text{ with } \mathbf{\Sigma}_{\mathbf{y}} \text{ unspecified}$$

"*Hotelling's* T^2 statistic" (*Hotelling* 1931, *Muirhead* 1982, *Rencher* 1998)

$$T^2 := [\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}_0]' \hat{\mathbf{\Sigma}}_{\hat{\xi}}^{-1} [\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}_0]$$

with respect to the eigenspace components of type $\mathbf{\Sigma}$-BLUUE and the dispersion matrix $\hat{\mathbf{\Sigma}}_{\hat{\xi}}$

$$\hat{\boldsymbol{\xi}} = \begin{bmatrix} \hat{\lambda}_1 \\ \hat{\lambda}_2 \\ \hat{\alpha}_1 \end{bmatrix} = \begin{bmatrix} 5.2837\ (n\text{strain/y}) \\ -5.2345\ (n\text{strain/y}) \\ 0.409074 \quad (\text{arc}) \end{bmatrix},\quad \hat{\mathbf{\Sigma}}_{\hat{\xi}} = \begin{bmatrix} +0.821618 & +0.208317 & -0.040791 \\ +0.208317 & +0.149544 & +0.000599 \\ -0.040791 & +0.000599 & +0.008215 \end{bmatrix}.$$

According to the design of the test for the eigenspace parameter vector in Section 4.3.1 at the error probability α we reject $\mathcal{H}_{01}$ in favor of $\mathcal{H}_{11}$ if

$$T^2 = [\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}_0]' \hat{\boldsymbol{\Sigma}}_{\hat{\xi}}^{-1} [\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}_0] > \frac{(n-1)\cdot 3}{n-3} F_{3,\ n-3}(1-\alpha)\} = T^2_{1-\alpha}.$$

With error probability α =5% and n=6 the *critical value*

$$F_{3,\ n-3}(1-\alpha) = 9.28,\ \ T^2_{1-\alpha} = 46.38.$$

Since *Hotelling's* statistic $T^2 = 7.32 < T^2_{1-\alpha} = 46.38$, accordingly we accept the null hypothesis $\mathcal{H}_{01} : \boldsymbol{\xi} = \boldsymbol{\xi}_0$ with the risk of α =5% of a *Type I error.*

(2) Test for a distinct element of the eigenspace parameter vector with *Student t- test*

Box 6.3

Separate Student t-tests about the eigenspace parameters in $\boldsymbol{\xi}$

Second test for $\mathcal{H}_{02} :\ \lambda_1 = \lambda_{10} = 4.3840 \,|\, \lambda_2 = \lambda_{20} = -4.7172 \,|\, \alpha_1 = \alpha_{10} = 0.5922$

(separately) $\mathcal{H}_{12} :\ \lambda_1 \neq \lambda_{10} = 4.3840 \,|\, \lambda_2 \neq \lambda_{20} = -4.7172 \,|\, \alpha_1 \neq \alpha_{10} = 0.5922$

"two-sided tests with *the test quantities*"

$$t_1 := \frac{\hat{\lambda}_1 - \lambda_{10}}{\hat{\sigma}_1},\ \ t_2 := \frac{\hat{\lambda}_2 - \lambda_{20}}{\hat{\sigma}_2},\ \ t_3 := \frac{\hat{\alpha}_1 - \alpha_{10}}{\hat{\sigma}_3}$$

with respect to $\hat{\lambda}_1, \hat{\lambda}_2, \hat{\alpha}_1$ of type $\boldsymbol{\Sigma}$ -BLUUE and their variances from *Box 6.1*.
t_1, t_2 and t_3 are elements of the *Student t-distribution with n-1 degree of freedom*.

With error probability α =5% we derive

$$t_{1,1-\alpha/2} = t_{2,1-\alpha/2} = t_{3,1-\alpha/2} = +2.57$$
$$t_{1,\alpha/2} = t_{2,\alpha/2} = t_{3,\alpha/2} = -2.57.$$

The critical values

$$c_{\lambda_1,\alpha/2} = \hat{\sigma}_{\lambda_1} t_{1,\alpha/2} + \lambda_{10} = 2.0539,\ \ c_{\lambda_1,1-\alpha/2} = \hat{\sigma}_{\lambda_1} t_{1,1-\alpha/2} + \lambda_{10} = 6.7141$$
$$c_{\lambda_2,\alpha/2} = \hat{\sigma}_{\lambda_2} t_{1,\alpha/2} + \lambda_{20} = \text{-}5.7113,\ \ c_{\lambda_2,1-\alpha/2} = \hat{\sigma}_{\lambda_2} t_{2,1-\alpha/2} + \lambda_{20} = \text{-}3.7231$$
$$c_{\alpha_1,\alpha/2} = \hat{\sigma}_{\alpha_1} t_{1,\alpha/2} + \alpha_{10} = 0.3592,\ \ c_{\alpha_1,1-\alpha/2} = \hat{\sigma}_{\alpha_1} t_{3,1-\alpha/2} + \alpha_{10} = 0.8251$$

indicate the confidence intervals

$$c_{\lambda_1,\alpha/2} = 2.0539 < \hat{\lambda}_1 = 5.2837 < c_{\lambda_1,1-\alpha/2} = 6.7141,$$
$$c_{\lambda_2,\alpha/2} = \text{-}5.7113 < \hat{\lambda}_2 = \text{-}5.2345 < c_{\lambda_2,1-\alpha/2} = \text{-}3.7231,$$
$$c_{\alpha_1,\alpha/2} = 0.3592 < \hat{\alpha}_1 = 0.409074 < c_{\alpha_1,1-\alpha/2} = 0.8251,$$
$$(c_{\alpha_1,\alpha/2} = 20^\circ 34' 50''.32 < \hat{\alpha}_1 = 23^\circ 26' 17''.52 < c_{\alpha_1,1-\alpha/2} = 47^\circ 16' 29''.09),$$

thereby suggesting the acceptance of all three null hypotheses.

$$\mathcal{H}_{02} :\ \lambda_1 = \lambda_{10} = 4.3840,\ \lambda_2 = \lambda_{20} = -4.7172,\ \alpha_1 = \alpha_{10} = 0.5922$$

with the risk of α =5% of a *Type I error*.

The 95% *confidence interval* for the eigenvalues λ_1, λ_2 and the eigendirection α_1 are

$$[2.0539,\ \ 6.7141]\ (\mu\text{strain/y});$$
$$[\text{-}5.7113,\ \text{-}3.7231]\ (\mu\text{strain/y});$$
$$[\ 20^\circ 34' 50''.32,\ \ 47^\circ 16' 29''.09]$$

respectively.

(3) *Eigen inference* about the orthonormally transformed parameters $\boldsymbol{\eta}$ with *Student t- test*

From the dispersion matrix $\hat{\mathbf{\Sigma}}_{\xi}$ -the variance covariance matrix of the eigenspace component parameter vector $\hat{\boldsymbol{\xi}}$ given in *Box 6.1*, we can see that these eigenspace component parameters are correlated. In order to make the hypothesis tests about the distinct elements more efficient and uncorrelated, we perform the *eigen-inference* derived in Section 4.3.3.

Using the orthogonal transformation (4.34) with normalized eigenvectors as column vectors (i.e., orthonormal basis, orthogonal matrix)

$$\mathbf{U}_{\hat{\xi}} = \begin{bmatrix} 0.079763 & -0.264891 & -0.960974 \\ -0.11894 & 0.954630 & -0.273014 \\ 0.989693 & 0.136071 & 0.044639 \end{bmatrix}, \text{ with } \mathbf{U}'_{\hat{\xi}}\mathbf{U}_{\hat{\xi}} = \mathbf{I}, \det(\mathbf{U}_{\hat{\xi}}) = 1$$

we set the transformed parameter vector (4.36) from the original parameter vectors and their $\mathbf{\Sigma}$-BLUUE estimates, respectively, which transform the null hypothesis values of Test 2 in *Box* 6.3 and the estimates into new values :

$$\boldsymbol{\eta}_0 = \mathbf{U}'_{\hat{\xi}}\boldsymbol{\xi}_0 = \begin{bmatrix} 1.4968 \\ -5.5839 \\ -2.8986 \end{bmatrix} \quad \text{and} \quad \hat{\boldsymbol{\eta}} = \mathbf{U}'_{\hat{\xi}}\hat{\boldsymbol{\xi}} = \begin{bmatrix} 1.4489 \\ -6.3409 \\ -3.6301 \end{bmatrix},$$

then we get

$$\hat{\mathbf{\Sigma}}_{\hat{\eta}} = \mathbf{U}'_{\hat{\xi}}\hat{\mathbf{\Sigma}}_{\hat{\xi}}\mathbf{U}_{\hat{\xi}} = \mathbf{\Lambda}_{\hat{\eta}} = \begin{bmatrix} 0.0049 & 0 & 0 \\ 0 & 0.0918 & 0 \\ 0 & 0 & 0.8827 \end{bmatrix},$$

from which we can see that the transformed parameters η_i are mutually independent and their estimated standard deviation are:

$$\hat{\sigma}_{\hat{\eta}_1} = \sqrt{\lambda_{\eta_1}} = 0.0697, \ \hat{\sigma}_{\hat{\eta}_2} = \sqrt{\lambda_{\eta_2}} = 0.3030, \ \hat{\sigma}_{\hat{\eta}_3} = \sqrt{\lambda_{\eta_3}} = 0.9395.$$

With these orthonormally transformed results we can now perform the *eigen-inference*. Note, that the orthonormally transformed parameters η_i are mutually independently normally distributed. *Student t-tests* could also be used for every transformed parameter $\hat{\eta}_i$.

The second hypothesis test performed in *Box 6.3* will be equivalent to the new hypothesis test for the orthonormally transformed parameters, i.e.,

$$\begin{array}{rl|l|l} \textit{Second test} \text{ for } \mathcal{H}_{02}: & \lambda_1 = \lambda_{10} = 4.3840 & \lambda_2 = \lambda_{20} = -4.7172 & \alpha_1 = \alpha_{10} = 0.5922 \\ \mathcal{H}_{12}: & \lambda_1 \neq \lambda_{10} = 4.3840 & \lambda_2 \neq \lambda_{20} = -4.7172 & \alpha_1 \neq \alpha_{10} = 0.5922 \end{array}$$

$$\Leftrightarrow$$

$$\begin{array}{rl|l|l} \textit{Third Test} \text{ for } \mathcal{H}_{03}: & \eta_1 = \eta_{10} = 1.4968 & \eta_2 = \eta_{20} = -5.5839 & \eta_3 = \eta_{30} = -2.8986 \\ \mathcal{H}_{13}: & \eta_1 \neq \lambda_{10} = 1.4968 & \eta_2 \neq \lambda_{20} = -5.5839 & \eta_3 \neq \eta_{30} = -2.8986 \end{array}$$

which means that when we accept or reject the new third hypothesis tests, we will accept or reject the second hypothesis tests accordingly.

These procedures will be summarized in Box *6.4.*

Box 6.4:

Eigen inference about the transformed parameters $\boldsymbol{\eta}$

$$\begin{array}{rl|l|l} \textit{Third Test} \text{ for } \mathcal{H}_{02}: & \eta_1 = \eta_{10} = 1.4968 & \eta_2 = \eta_{20} = -5.5839 & \eta_3 = \eta_{30} = -2.8986 \\ \mathcal{H}_{12}: & \eta_1 \neq \lambda_{10} = 1.4968 & \eta_2 \neq \lambda_{20} = -5.5839 & \eta_3 \neq \eta_{30} = -2.8986 \end{array}$$

"two-sided test with *test quantities*"

$$t_1 := \frac{\hat{\eta}_1 - \eta_{10}}{\hat{\sigma}_{\eta_1}}, \quad t_2 := \frac{\hat{\eta}_2 - \eta_{20}}{\hat{\sigma}_{\eta_2}}, \quad t_3 := \frac{\hat{\eta}_3 - \eta_{30}}{\hat{\sigma}_{\eta_3}}$$

with respect to $\hat{\eta}_1, \hat{\eta}_2, \hat{\eta}_3$ and their related variances . t_1, t_2 *and* t_3 are elements of the *Student t-distribution with n-1 degree of freedom*.

With error probability α =5%

$$t_{1,1-\alpha/2} = t_{2,1-\alpha/2} = t_{3,1-\alpha/2} = +2.57$$
$$t_{1,\alpha/2} = t_{2,\alpha/2} = t_{3,\alpha/2} = -2.57.$$

The *critical values*

$$c_{\eta_1,\alpha/2} = \hat{\sigma}_{\eta_1} t_{1,\alpha/2} + \eta_{10} = 1.3177, \quad c_{\eta_1,1-\alpha/2} = \hat{\sigma}_{\eta_1} t_{1,1-\alpha/2} + \eta_{10} = 1.6759,$$
$$c_{\eta_2,\alpha/2} = \hat{\sigma}_{\eta_2} t_{2,\alpha/2} + \eta_{20} = -6.3628, \quad c_{\eta_2,1-\alpha/2} = \hat{\sigma}_{\eta_2} t_{2,1-\alpha/2} + \eta_{20} = -4.8049,$$
$$c_{\eta_3,\alpha/2} = \hat{\sigma}_{\eta_3} t_{3,\alpha/2} + \eta_{30} = -5.3137, \quad c_{\eta_3,1-\alpha/2} = \hat{\sigma}_{\eta_3} t_{3,1-\alpha/2} + \eta_{30} = -0.4835,$$

indicate

$$c_{\eta_1,\alpha/2} = 1.3177 < \hat{\eta}_1 = 1.4968 < c_{\eta_1,1-\alpha/2} = 1.6759$$
$$c_{\eta_2,\alpha/2} = -6.3628 < \hat{\eta}_2 = -5.5839 < c_{\eta_2,1-\alpha/2} = -4.8049$$
$$c_{\eta_3,\alpha/2} = -5.3137 < \hat{\eta}_3 = -2.8986 < c_{\eta_3,1-\alpha/2} = -0.4835$$

a result which leads us to *accept the null hypothesis*

$$\mathcal{H}_{03}: \ \eta_1 = \eta_{10} = 1.4968, \ \eta_2 = \eta_{20} = -5.5839, \ \eta_3 = \eta_{30} = -2.8986$$

with the risk of α =5% of a *Type I error*.

Accordingly we accept the original null hypothesis about the eigenspace components

$$\mathcal{H}_{02}: \ \lambda_1 = \lambda_{10} = 4.3840, \ \lambda_2 = \lambda_{20} = -4.7172, \ \alpha_1 = \alpha_{10} = 0.5922.$$

This completes the example of *eigen inference.*

(4) Test for the variance-covariance matrix $\mathbf{\Sigma}_{\mathrm{y}} = \mathbf{\Sigma}_0$ with the *likelihood ratio test*

Box 6.5

Multivariate hypothesis tests about the variance-covariance matrix $\mathbf{\Sigma}_{\mathrm{y}}$

Fourth test for $\mathcal{H}_{04}: \ \mathbf{\Sigma}_{\mathrm{y}} = \mathbf{\Sigma}_0, \quad \mathcal{H}_{14}: \ \mathbf{\Sigma}_{\mathrm{y}} \neq \mathbf{\Sigma}_0$

"unbiased modified likelihood ratio statistic Λ_1"

(Giri 1977, Muirhead 1982, Koch 1999, Koch 2001)

$$\Lambda_1 = \left(\frac{e}{n-1}\right)^{3(n-1)/2} (\det(n-1)\hat{\mathbf{\Sigma}}_{\mathrm{y}} \mathbf{\Sigma}_0^{-1})^{(n-1)/2} \operatorname{etr}\{-\tfrac{1}{2}(n-1)\hat{\mathbf{\Sigma}}_{\mathrm{y}} \mathbf{\Sigma}_0^{-1}\}$$

with respect to the sample variance-covariance matrix $\hat{\mathbf{\Sigma}}_{\mathrm{y}}$ of type BIQUUE in *Table 2.* Since our sample size is relatively small we have to use the exact distribution of $-2\log\Lambda_1$,whose upper 5 and 1 percentage points have been provided by *Muirhead* (1982, p.360).

With α =5% the *critical value is*

$$L_{1-\alpha} = -2\log\Lambda_1(1-\alpha) = 15.508$$

Let us perform this test with three different variance-covariance matrices $\mathbf{\Sigma}_{01}$, $\mathbf{\Sigma}_{02}$ and $\mathbf{\Sigma}_{03}$, namely

$$\mathbf{\Sigma}_{01} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$\mathbf{\Sigma}_{02} = \begin{bmatrix} 2.033987 & 0.444505 & 0.033115 \\ 0.444505 & 0.982429 & 0.585175 \\ 0.033115 & 0.585175 & 1.02023 \end{bmatrix},$$

$$\mathbf{\Sigma}_{03} = \begin{bmatrix} 9.808347 & -2.351484 & -2.515398 \\ -2.351484 & 1.588078 & 2.089224 \\ -2.515398 & 2.089224 & 3.349229 \end{bmatrix},$$

which are chosen according to the hypotheses that the *variance-covariance component matrix* $\mathbf{\Sigma_y}$ is equal to (1) a unit matrix $\mathbf{\Sigma}_{01}$; (2) $\mathbf{\Sigma}_{02}$ - the variance-covariance matrix of the strain rate observations derived from ITRF97 residual velocities; and (3) a matrix $\mathbf{\Sigma}_{03}$ whose diagonal elements are 0.1 smaller than the estimated variance-covariance component matrix $\hat{\mathbf{\Sigma}}_\mathbf{y}$ as given in *Box 6.1*.

With respect to the likelihood ratio statistics Λ_1 and the related $-2\log\Lambda_1 =: L_1$ we find:

$$\Lambda_{11} = 2.264853\times10^{-11} \quad L_{11} = -2\log\Lambda_{11} = 49.021852$$
$$\Lambda_{12} = 1.117952\times10^{-6} \quad L_{12} = -2\log\Lambda_{12} = 27.408023$$
$$\Lambda_{13} = 0.694227 \quad L_{13} = -2\log\Lambda_{13} = 0.729913$$

Since $L_{11} = 49.021852 > L_{1-\alpha} = 15.805$, we reject the *null hypothesis* $\mathcal{H}_{03}$: $\mathbf{\Sigma_y} = \mathbf{\Sigma}_{01}$ with the risk of α =5% of a *Type I error*.

Since $L_{12} = 27.408023 > L_{1-\alpha} = 15.805$, we reject the *null hypothesis* $\mathcal{H}_{03}$: $\mathbf{\Sigma_y} = \mathbf{\Sigma}_{02}$ with the risk of α =5% of a *Type I error*.

Since $L_{13} = 0.729913 < L_{1-\alpha} = 15.805$, we accept the *null hypothesis* $\mathcal{H}_{04}$: $\mathbf{\Sigma_y} = \mathbf{\Sigma}_{03}$ with the risk of α =5% of a *Type I error*.

(5) Test for the eigenspace parameter vector and variance-covariance matrix $\xi = \xi_0, \mathbf{\Sigma_y} = \mathbf{\Sigma}_0$ *likelihood ratio test*

Box 6.6

Multivariate hypothesis tests about the eigenspace parameter vector ξ and the variance-covariance matrix $\mathbf{\Sigma_y}$

Fifth test for $\mathcal{H}_{05}$: $\xi = \xi_0,\ \mathbf{\Sigma_y} = \mathbf{\Sigma}_0$, $\mathcal{H}_{15}$: $\xi \neq \xi_0$ *or* $\mathbf{\Sigma_y} \neq \mathbf{\Sigma}_0$

"unbiased likelihood ratio statistic Λ_2"

(*Anderson* 1984, *Murihead* 1982)

$$\Lambda_2 = \left(\frac{e}{n}\right)^{3n/2} (\det(n-1)\hat{\mathbf{\Sigma}}_\mathbf{y}\mathbf{\Sigma}_0^{-1})^{n/2}\, \mathrm{etr}\{-\frac{1}{2}(n-1)\hat{\mathbf{\Sigma}}_\mathbf{y}\mathbf{\Sigma}_0^{-1}\}\exp\{-\frac{1}{2}[\hat{\xi}-\xi_0]'\mathbf{\Sigma}_{\hat{\xi}0}^{-1}[\hat{\xi}-\xi_0]\}$$

with respect to the eigenspace components of type $\mathbf{\Sigma}$ -BLUUE and variance-covariance matrix $\hat{\mathbf{\Sigma}}_\mathbf{y}$ of type BIQUUE in *Box 6.1* and $\mathbf{\Sigma}_{\hat{\xi}0}^{-1} = (1/n)(\mathcal{A}'\mathbf{\Sigma}_0^{-1}\mathcal{A})^{-1}$.
Since our sample size is relatively small we have to use the exact distribution of $-2\log\Lambda_2$, whose upper 5 and 1 percentage points have been provided by *Murihead* (1982, p.371).

With α =5% the *critical value* is derived

$$L_{1-\alpha} = -2\log\Lambda_2(1-\alpha) = 24.431.$$

Choose ξ_0 as in the hypothesis test one and $\mathbf{\Sigma}_0$ as $\mathbf{\Sigma}_{03}$ in the fourth hypothesis test, namely

$$\xi_0 = \begin{bmatrix}\lambda_{10}\\ \lambda_{20}\\ \alpha_{10}\end{bmatrix} = \begin{bmatrix}4.3840\\ -4.7172\\ 0.5922\end{bmatrix},\ \mathbf{\Sigma}_0 = \begin{bmatrix}9.808347 & -2.351484 & -2.515398\\ -2.351484 & 1.588078 & 2.089224\\ -2.515398 & 2.089224 & 3.349229\end{bmatrix}.$$

With respect to the likelihood ratio statistics Λ_2 and related $-2\log\Lambda_2 = L_2$ data we are led to

$$\Lambda_2 = 0.008752 \quad L_2 = -2\log\Lambda_2 = 9.476949$$

Since $L_2 = 9.476949 < L_{1-\alpha} = 24.431$, we accept the *null hypothesis* $\mathcal{H}_{05}$: $\xi = \xi_0,\ \mathbf{\Sigma_y} = \mathbf{\Sigma}_0$ with the risk of α = 5% of a *Type I error*.

(6) The general linear hypothesis test with a *growth curve model* for eigenspace parameters

As it is mentioned in Section 4.3.6, the special linearized multivariate *Gauss-Markov* model for sampling the eigenspace synthesis in Box 4.5

$$\mathbf{Y} = \mathbf{F}(\xi_0)\mathbf{1}' + [\mathcal{A}(\xi-\xi_0)]\mathbf{1}' + \mathbf{E} \tag{4.25}$$

is also a growth curve model

$$\mathbf{Y} = \mathbf{A\Xi B} + \mathbf{E} \tag{4.43}$$

with the correspondences $\mathscr{A} = \mathbf{A}$, $\mathbf{1}' = \mathbf{B}$ and $\xi = \mathbf{\Xi}$. This fact suggests that the hypothesis (4.46) under the *growth curve* model can be applied to the testing of the eigenspace parameter directly.

For the general linear hypothesis test with a growth curve model for eigenspace parameters we test three cases in Box 6.7. The second case is just for testing the difference of the two eigenvalue parameters, since $\lambda_2 = -5.2345 < 0$ we have to rewrite it as $\lambda_1 + \lambda_2 = 0$.

Box 6.7

The general linear hypothesis test with a growth curve model

$$\textit{Sixth test} \quad \mathcal{H}_{06}: \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \textit{versus} \quad \mathcal{H}_{16}: \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\mathcal{H}'_{06}: \lambda_1 + \lambda_2 = 0 \quad \textit{versus} \quad \mathcal{H}'_{16}: \lambda_1 + \lambda_2 \neq 0$$

$$\mathcal{H}''_{06}: \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \alpha_1 \end{bmatrix} = \begin{bmatrix} \xi_{10} \\ \xi_{20} \\ \xi_{30} \end{bmatrix} \quad \textit{versus} \quad \mathcal{H}''_{16}: \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \alpha_1 \end{bmatrix} \neq \begin{bmatrix} \xi_{10} \\ \xi_{20} \\ \xi_{30} \end{bmatrix}$$

For the first case,

$$\mathcal{H}_{06}: \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \iff \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

i.e.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{bmatrix} \times 1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

which is corresponding to (4.46):

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \ \mathbf{Q} = 1 \ \text{ and } \ \mathbf{\Xi} = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{bmatrix}.$$

With (4.47) and (4.48) we get:

$$\mathbf{R} = (\mathbf{BB}')^{-1} + (\mathbf{BB}')^{-1}\mathbf{BY}'\mathbf{\Omega}^{-1}\mathbf{YB}'(\mathbf{BB}')^{-1} - \mathbf{\Xi}'(\mathbf{A}'\mathbf{\Omega}^{-1}\mathbf{A})\mathbf{\Xi} = 0.1667\,.$$

$$\mathbf{V}_{\mathrm{H}} = (\mathbf{P\Xi Q})(\mathbf{Q}'\mathbf{RQ})^{-1}(\mathbf{P\Xi Q})' = \begin{bmatrix} 167.5049 & -165.9452 \\ -165.9452 & 164.3999 \end{bmatrix}$$

$$\mathbf{V}_{\mathrm{E}} = \mathbf{P}(\mathbf{A}'\mathbf{\Omega}^{-1}\mathbf{A})^{-1}\mathbf{P}' = \begin{bmatrix} 24.6486 & 6.2495 \\ 6.2495 & 4.4863 \end{bmatrix}$$

and the greatest eigenvalue of $\mathbf{V}_{\mathrm{H}}\mathbf{V}_{\mathrm{E}}^{-1}$

$$\lambda_{\max} = (\mathbf{V}_{\mathrm{H}}\mathbf{V}_{\mathrm{E}}^{-1}) = 96.1606\,.$$

Since p=3, q=3, r=1, c=2, g=1, with (4.49) we get $s^* = 1$, $m^* = 0$, $n^* = 1$, and the test statistic

$$t = \frac{n^* + 1}{m^* + 1}\lambda_{\max} = 192.3212$$

can be compared with the critical values of the *F-distribution* with 2 and 4 degrees of freedom. The hypothesis of zero would of course, be rejected at any reasonable level. For example, at 95% confidence level, $F_{0.95,\,2,9} = 6.9443$, thus $t > F_{0.95,2,4}$, and we reject the null hypothesis of

$$\mathcal{H}_{06}:\begin{bmatrix}\lambda_1\\ \lambda_2\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}.$$

For the second case,

$$\mathcal{H}'_{06}:\lambda_1+\lambda_2=0 \quad \Leftrightarrow \quad \xi_1+\xi_2=0$$

i.e.

$$\begin{bmatrix}1 & 1 & 0\end{bmatrix}\begin{bmatrix}\xi_1\\ \xi_2\\ \xi_3\end{bmatrix}\times 1=0$$

which is corresponding to (4.46):

$$\mathbf{P}=\begin{bmatrix}1 & 1 & 0\end{bmatrix},\ \mathbf{Q}=1 \quad \text{and} \quad \mathbf{\Xi}=\begin{bmatrix}\xi_1\\ \xi_2\\ \xi_3\end{bmatrix}$$

and with (4.47) and (4.48), we get $\mathbf{R}=0.1667$, $\mathbf{V}_{\mathrm{H}}=0.0145$ and $\mathbf{V}_{\mathrm{E}}=41.6339$ and the “greatest eigenvalues” of $\mathbf{V}_{\mathrm{H}}\mathbf{V}_{\mathrm{E}}^{-1}$ as $\mathbf{V}_{\mathrm{H}}\mathbf{V}_{\mathrm{E}}^{-1}$ itself:

$$\lambda_{\max}=(\mathbf{V}_{\mathrm{H}}\mathbf{V}_{\mathrm{E}}^{-1})=\frac{\mathbf{V}_{\mathrm{H}}}{\mathbf{V}_{\mathrm{E}}}=0.000348.$$

Since p=3, q=3, r=1, c=1, g=1, with (4.49) we get $s^*=1$, $m^*=-1/2$, $n^*=3/2$ and the test statistic

$$t=\frac{n^*+1}{m^*+1}\mathbf{V}_{\mathrm{H}}\mathbf{V}_{\mathrm{E}}^{-1}=0.0017$$

which is smaller than the quantile at 95% confidence level, $F_{0.95,\,1,5}=6.6079$; accordingly, we accept the null hypothesis of $\mathcal{H}'_{06}:\lambda_1+\lambda_2=0$.

Here we note that the 95% simultaneous confidence interval for $(\lambda_1+\lambda_2)$ can be computed from (4.50) with **a = b =**1 and $x_\alpha/(1+x_\alpha)=[(m^*+1)/(n^*+1)]F_{\alpha;2m^*+2,\,2n^*+2}=[(-1/2+1)/(5+1)]F_{0.95;1,\,5}$ to be

$$c_1=\mathbf{P}\hat{\mathbf{\Xi}}\mathbf{Q}-(\frac{m^*+1}{n^*+1}F_{\alpha;2m^*+2,\,2n^*+2}\mathbf{V}_{\mathrm{E}}\mathbf{Q}'\mathbf{R}\mathbf{Q})^{1/2}=-1.8060$$

$$c_2=\mathbf{P}\hat{\mathbf{\Xi}}\mathbf{Q}+(\frac{m^*+1}{n^*+1}F_{\alpha;2m^*+2,\,2n^*+2}\mathbf{V}_{\mathrm{E}}\mathbf{Q}'\mathbf{R}\mathbf{Q})^{1/2}=\ 1.9044,$$

i.e.

$$-1.8060\le\lambda_1+\lambda_2\le 1.9044.$$

From the actual estimates of $\hat{\lambda}_1$ and $\hat{\lambda}_2$ we have

$$\hat{\lambda}_1+\hat{\lambda}_2=5.2837-5.2345=0.0492$$

which is an element of the confidence interval [c_1, c_2]; the null hypothesis of the second case is therefore accepted.

For the third case of the sixth test

$$\mathcal{H}''_{06}:\begin{bmatrix}\lambda_1\\ \lambda_2\\ \alpha 1\end{bmatrix}=\begin{bmatrix}\xi_{10}\\ \xi_{20}\\ \xi_{30}\end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix}\xi_1\\ \xi_2\\ \xi_3\end{bmatrix}=\begin{bmatrix}\xi_{10}\\ \xi_{20}\\ \xi_{30}\end{bmatrix}$$

i.e.

$$\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}\xi_1\\ \xi_2\\ \xi_3\end{bmatrix}\times 1=\begin{bmatrix}\xi_{10}\\ \xi_{20}\\ \xi_{30}\end{bmatrix}=\begin{bmatrix}4.3840\\ -4.7172\\ 0.5922\end{bmatrix}$$

which is the same as in our first hypothesis test in Box 6.2 and corresponding to (4.46):

$$\mathbf{P}=\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix},\ \mathbf{Q}=1 \quad\text{and}\quad \boldsymbol{\Xi}=\begin{bmatrix}\xi_1\\\xi_2\\\xi_3\end{bmatrix}$$

With (4.47) and (4.48) we get:

$$\mathbf{R}=(\mathbf{BB}')^{-1}+(\mathbf{BB}')^{-1}\mathbf{BY}'\boldsymbol{\Omega}^{-1}\mathbf{YB}'(\mathbf{BB}')^{-1}-\boldsymbol{\Xi}'(\mathbf{A}'\boldsymbol{\Omega}^{-1}\mathbf{A})\boldsymbol{\Xi}=0.1667\,.$$

$$\mathbf{V}_{\mathrm{H}}=(\mathbf{P\Xi Q})(\mathbf{Q}'\mathbf{RQ})^{-1}(\mathbf{P\Xi Q})'=\begin{bmatrix}4.8568 & -2.7925 & -0.9883\\-2.7925 & 1.6056 & 0.5682\\-0.9883 & 0.5682 & 0.2011\end{bmatrix}$$

$$\mathbf{V}_{\mathrm{E}}=\mathbf{P}(\mathbf{A}'\boldsymbol{\Omega}^{-1}\mathbf{A})^{-1}\mathbf{P}'=\begin{bmatrix}24.6486 & 6.2495 & -1.2237\\6.2495 & 4.4863 & 0.0180\\-1.2237 & 0.0180 & 0.2465\end{bmatrix}$$

and the greatest eigenvalue of $\mathbf{V}_{\mathrm{H}}\mathbf{V}_{\mathrm{E}}^{-1}$

$$\lambda_{\max}=(\mathbf{V}_{\mathrm{H}}\mathbf{V}_{\mathrm{E}}^{-1})=0.6695\,.$$

Since p=3, q=3, r=1, c=3, g=1, with (4.49) we get $s^*=1,\ m^*=1/2,\ n^*=1/2$ and the test statistic

$$t=\frac{n^*+1}{m^*+1}\lambda_{\max}=0.6695$$

which is smaller than the quantile 95% confidence level $F_{0.95,\,3,3}=9.2766$; accordingly we accept the null hypothesis of

$$\mathcal{H}''_{06}:\begin{bmatrix}\lambda_1\\\lambda_2\\\alpha 1\end{bmatrix}=\begin{bmatrix}\xi_{10}\\\xi_{20}\\\xi_{30}\end{bmatrix}=\begin{bmatrix}4.3840\\-4.7172\\0.5922\end{bmatrix}.$$

Finally we would like to briefly discuss the statistical property of the trace of $\mathbf{V}_{\mathrm{H}}\mathbf{V}_{\mathrm{E}}^{-1}$. *Lawley* (1938) and *Hotelling* (1947, 1951) have proposed the trace of $\mathbf{V}_{\mathrm{H}}\mathbf{V}_{\mathrm{E}}^{-1}$ as a test criterion, the so called *Lawley –Hotelling Trace Test*. The exact distribution of

$$T_0^2=\mathrm{tr}(\mathbf{V}_{\mathrm{H}}\mathbf{V}_{\mathrm{E}}^{-1})$$

was obtained by Hotelling (1951) for $p = 2$ in the central case. For $p\geq 3$, the distribution of T_0^2 is generally quite complicated. Several authors, *Constantine* (1966), *Davis*(1968,1970a, b) *Muirhead* (1972), *Pillai and Sampson* (1959), *Pillar and Sudjana* (1974), *Pillai and Young* (1971) and *Siotani*, *Hayakawa and Fujikoshi* (1985) considered the central and noncentral distribution of T_0^2. Some tables of the approximate percentage point of T_0^2 are available in *Davis* (1972), *Pillai* (1960) and *Kres* (1983). Its relationship with *Hotelling's* T^2 *test* used in the first test is given by the expression

$$T_0^2=\frac{1}{n-1}T^2.$$

In the third case we have $T_0^2=\mathbf{V}_{\mathrm{H}}\mathbf{V}_{\mathrm{E}}^{-1}=1.4641$, which is equal to T^2/(6-1)=7.32/5 in the first hypothesis test. Since we have accepted the null hypothesis $\mathcal{H}''_{06}$, this comparison also supports our first hypothesis test- *Hotelling's* T^2 *test.*

From the analysis above we can summarize that our estimation theory about the two-dimensional, symmetric rank-two random tensor, developed in Chapter 4 is practical to be applied and produces not only estimates consistent with the tectonic setting, but also successive hypothesis tests which complete the statistical inference of eigenspace component parameters of a two-dimensional, symmetric rank-two strain rate tensor.

6.7 Statistical inference of the eigenspace components of 3-D strain rate tensor

As second case study, both model and hypothesis tests developed in Chapter 5 will be applied to the observations of three-dimensional strain rate tensors derived in Section 6.4 for the sub-network of sites 1 – 4 – 9 – 8 (Grass – Bologna - Wettzell - Zimmerwald) in the studying region, see Figure 6.14, at six epochs.

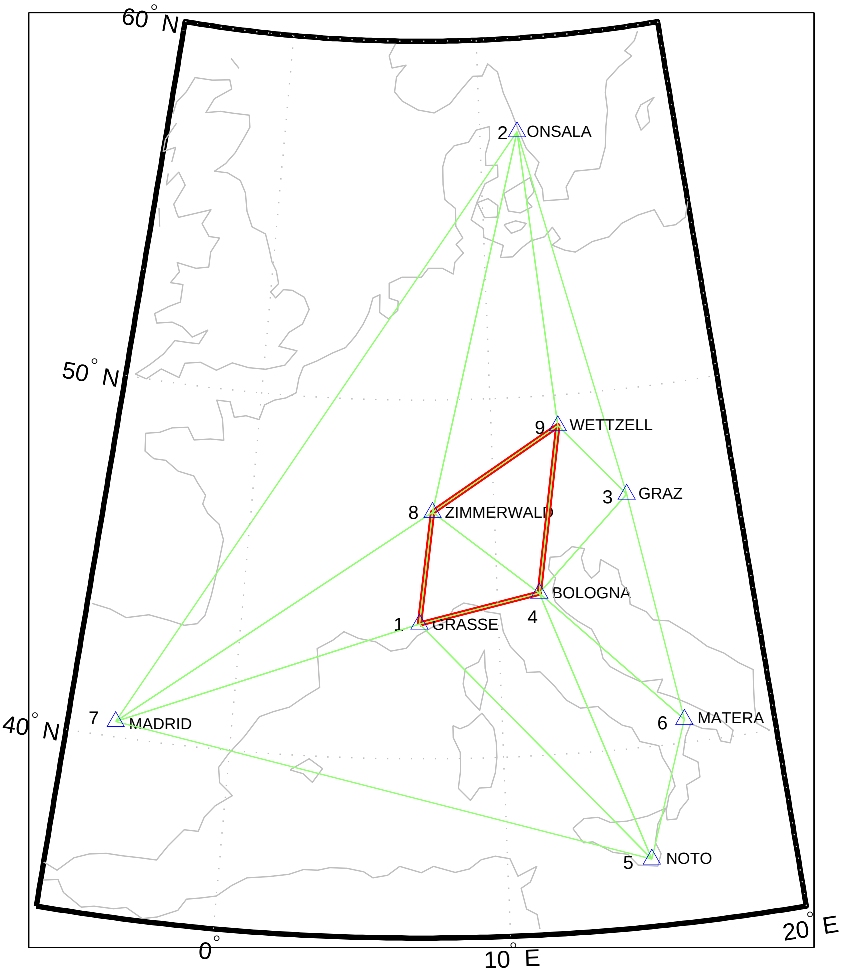


Figure 6.14 The sub-network of sites 1- 4 - 9 - 8 (Grasse-Bologna-Wettzell-Zimmerwald) in the study region

6.7.1 Estimation of the eigenspace components of the 3-D strain rate tensor

With the three-dimensional strain rate tensor observations (Table 6.6) calculated by (6.11) with the six epoch ITRF residual velocities (Table 6.4), we can now estimate the eigenspace components (eigenvalues and eigendirections) of the three-dimensional strain rate tensor, variance-covariance component matrix of type BIQUUE, and their dispersion matrix with (4.29), (4.30) and (4.31), and successively make hypothesis tests. The detailed estimates of eigenspace components (eigenvalues and eigendirections) and the associated standard deviations of the sub-network of the sites 1 – 4 – 9 – 8 are listed in Table 6.8.

Table 6.8. The estimates of the eigenspace components and their standard deviations

Eigenvalues and standard deviation (10^{-7} strain/yr)					
$\hat{\lambda}_1$	$\hat{\sigma}_{\hat{\lambda}_1}$	$\hat{\lambda}_2$	$\hat{\sigma}_{\hat{\lambda}_2}$	$\hat{\lambda}_3$	$\hat{\sigma}_{\hat{\lambda}_3}$
-1.049206	±0.3611	0.013154	±0.0213	2.953239	±0.9910
Orthonormal orientation parameters and standard deviation (degree)					
$\hat{\theta}_{32}$	$\hat{\sigma}_{\hat{\theta}_{32}}$	$\hat{\theta}_{31}$	$\hat{\sigma}_{\hat{\theta}_{31}}$	$\hat{\theta}_{21}$	$\hat{\sigma}_{\hat{\theta}_{21}}$
13.592942	±2.6163	12.172197	±6.7498	4.739774	±3.4609

In order to get a visual presentation of the possible magnitude and the directions of the extension and contraction of the strain rate that is important to predict the tectonic activity, including possible deformation trend and directions, let us first introduce the determination of the eigendirections and their 95% confidence intervals from the orthonormal transformation (rotation) matrix **U**, whose elements are functions of the three rotation angles θ_{32}, θ_{31} and θ_{21}, see (5.25). With the Jacobi matrix for the transformation of three rotation angles θ_{32}, θ_{31} and θ_{21} to the eigenvectors $\mathbf{u}_1$, $\mathbf{u}_2$ and $\mathbf{u}_3$ of **U**, the variance-covariance matrix of $\mathbf{u}_i$ is transformed from the variance-covariance matrix of the orthonormal orientation parameter vector $[\hat{\theta}_{32}\ \hat{\theta}_{31}\ \hat{\theta}_{21}]'$:

$$\hat{\mathbf{\Sigma}}_{\mathbf{u}_i} = \mathbf{J}_{vec_i} \hat{\mathbf{\Sigma}}_{\hat{\boldsymbol{\theta}}} \mathbf{J}'_{vec_i},\ i = 1,\ 2,\ 3,$$

with

$$\mathbf{J}_{vec_i} = \begin{bmatrix} \frac{\partial u_{i1}}{\partial \theta_{32}} & \frac{\partial u_{i1}}{\partial \theta_{31}} & \frac{\partial u_{i1}}{\partial \theta_{21}} \\ \frac{\partial u_{i2}}{\partial \theta_{32}} & \frac{\partial u_{i2}}{\partial \theta_{31}} & \frac{\partial u_{i2}}{\partial \theta_{21}} \\ \frac{\partial u_{i3}}{\partial \theta_{32}} & \frac{\partial u_{i3}}{\partial \theta_{31}} & \frac{\partial u_{i3}}{\partial \theta_{21}} \end{bmatrix}.$$

The detailed results are:

$$D\{\hat{\boldsymbol{\theta}}\} = \mathbf{\Sigma}_{\hat{\boldsymbol{\theta}}} = \begin{bmatrix} 0.002085 & -0.000535 & 0.002642 \\ -0.000535 & 0.013878 & 0.001192 \\ 0.002642 & 0.001192 & 0.003649 \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} 0.974175 & 0.080773 & 0.210850 \\ -0.129701 & 0.964571 & 0.229739 \\ -0.184824 & -0.251153 & 0.950138 \end{bmatrix} = [\mathbf{u}_1\ \mathbf{u}_2\ \mathbf{u}_3],$$

$$\hat{\mathbf{\Sigma}}_{\mathbf{u}_1} = \begin{bmatrix} 0.000677 & 0.001234 & 0.002702 \\ 0.001234 & 0.005617 & 0.002564 \\ 0.002702 & 0.002564 & 0.012445 \end{bmatrix}, \hat{\mathbf{\Sigma}}_{\mathbf{u}_2} = \begin{bmatrix} 0.003426 & -0.001125 & -0.003217 \\ -0.001125 & 0.000371 & 0.001062 \\ -0.003217 & 0.001062 & 0.003046 \end{bmatrix}, \hat{\mathbf{\Sigma}}_{\mathbf{u}_3} = \begin{bmatrix} 0.013261 & -0.001169 & -0.002660 \\ -0.001169 & 0.001967 & -0.000216 \\ -0.002660 & -0.000216 & 0.000643 \end{bmatrix}.$$

Second, we are able to illustrate the eigenspace and the confidence region of the three-dimension strain rate tensor. The three eigendirections determined by $\mathbf{u}_1$, $\mathbf{u}_2$ and $\mathbf{u}_3$ of the orthonormal transformation (rotation) matrix **U** are presented on the unit sphere in *Figure 6.15*, together with their 95% confidence regions.

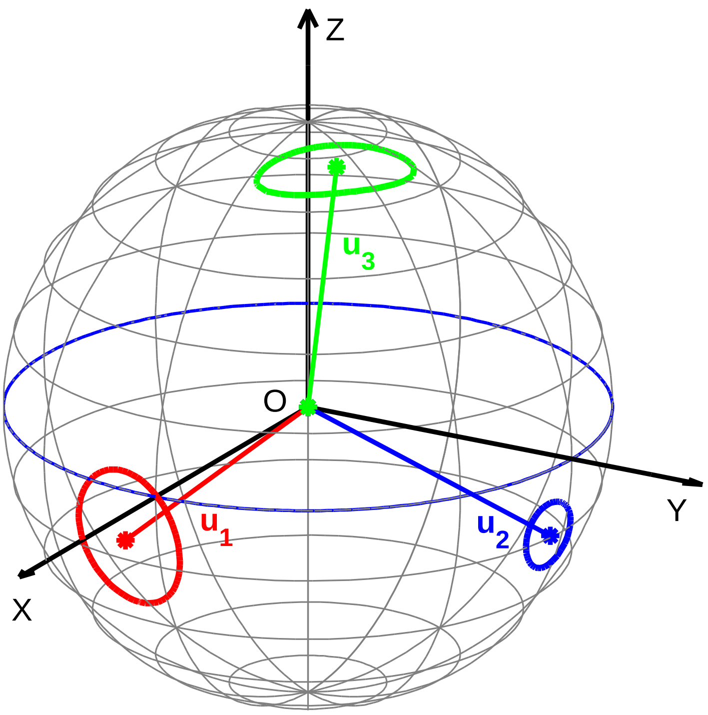


Figure 6.15. The eigendirections of the 3-D strain rate tensor, with their 95% confidence regions

The 95% confidence regions are determined as follows: (1) The variance-covariance matrix $\hat{\mathbf{\Sigma}}_{\mathbf{u}_i}$ with respect to the Cartesian coordinates XYZ is transformed to the eigenspace by $\mathbf{u}_1$, $\mathbf{u}_2$ and $\mathbf{u}_3$ with orthonormal transformation (rotation): $\hat{\mathbf{\Sigma}}_{\mathbf{u}_i Eig} = \mathbf{U}\hat{\mathbf{\Sigma}}_{\mathbf{u}_i}\mathbf{U}'$; (2) We consider the standard deviations $[\hat{\sigma}_{u_{1i}}\ \hat{\sigma}_{u_{2i}}\ \hat{\sigma}_{u_{3i}}]'$, the square roots of the diagonal elements of $\hat{\mathbf{\Sigma}}_{\mathbf{u}_i Eig}$, as the total angular displacement of the *i*th eigendirection and they are projected on the tangent surface at the end of the unit eigen vector $\mathbf{u}_i$ on the unit sphere. For the vector $\mathbf{u}_1$ the possible error regions is determined only by the two standard deviations $[\hat{\sigma}_{u_{21}}\ \hat{\sigma}_{u_{31}}]'$, similarly $\mathbf{u}_2$ by $[\hat{\sigma}_{u_{12}}\ \hat{\sigma}_{u_{32}}]'$ and $\mathbf{u}_3$ by $[\hat{\sigma}_{u_{13}}\ \hat{\sigma}_{u_{23}}]'$. (3) Using the α/2 and (1-α/2) quantiles $t_{\alpha/2}$ and $t_{1-\alpha/2}$ of the *Student distribution* with α=5% and these standard deviations, the *95*% confidence region of the eigendirections of 3-D strain rate are determined by

$$c_{\alpha/2} = \hat{\sigma}_{u_{ij}} t_{\alpha/2},\ c_{1-\alpha/2} = \hat{\sigma}_{u_{ij}} t_{1-\alpha/2}.$$

This brings us the 95% confidence region (with $t_{\alpha/2}$=2.57) for every eigenvector of the strain rate tensor in arc length and angle on the unit sphere in *Table 6.9*, and is projected onto the surface of a unit sphere in Figure 6.15.

Table 6.9. *95*% confidence regions for the eigenvectors of the strain rate tensor

Principal axes (*i*, *j*)	Length ($c_{\alpha/2} = \hat{\sigma}_{u_{ji}} t_{\alpha/2}$)	Angle on the unit sphere ($c_{\alpha/2} = \hat{\sigma}_{u_{ji}} t_{\alpha/2}$ in degree)
1, 2	0.207756	11°.903555
1, 3	0.250465	14°.350605
2, 1	0.112777	6°.461661
2, 3	0.149905	8°.588900
3, 1	0.126935	15°.626844
3, 2	0.117323	7°.272837

6.7.2 Statistical inference of the estimates of eigenspace component parameters of the 3-D strain rate tensor

The estimates of the eigenspace component parameters and their related dispersion matrix from the three-dimensional, symmetric rank-two strain rate observations of six epochs reflect the statistical average information of the random strain rate tensor, utilizing the advantage of the longer time span. With them we can successively perform the statistical inference, i.e.

Statistical Inference = Estimate +Hypothesis test.

The estimates of type BLUUE of the eigenspace component parameters of a three-dimensional, symmetric rank-two random strain rate tensor in the sub-network, their related dispersion matrix with *Theorem* 4.3 and the sample variance-covariance matrix $\hat{\mathbf{\Sigma}}_{\mathbf{y}}$ of type BIQUUE with *Theorem* 4.4 are summarized in the following *Box 6.8.*

Box 6.8

***Case study*: Hypothesis test with a 3-dimensional strain rate tensor in sub-network of sites 1– 4 – 9 – 8 (Grass-Bologna-Wettzell-Zimmerwald)**

"the $\mathbf{\Sigma}$-BLUUE of eigenspace components of a three-dimensional, symmetric rank-two random tensor with *Theorem* 4.3"

$$\hat{\boldsymbol{\xi}} = \begin{bmatrix} \hat{\lambda}_1 \\ \hat{\lambda}_2 \\ \hat{\lambda}_3 \\ \hat{\theta}_{32} \\ \hat{\theta}_{31} \\ \hat{\theta}_{21} \end{bmatrix} = \begin{bmatrix} -1.049206\ (10^{-7}\text{strain/y}) \\ 0.013154(10^{-7}\text{strain/y}) \\ 2.953239(10^{-7}\text{strain/y}) \\ 0.237242\quad (\text{arc}) \\ 0.212445\quad (\text{arc}) \\ 0.082725\quad (\text{arc}) \end{bmatrix} = \begin{bmatrix} -1.049206\ (10^{-7}\text{strain/y}) \\ 0.013154(10^{-7}\text{strain/y}) \\ 2.953239(10^{-7}\text{strain/y}) \\ 13^\circ.592942 \\ 12^\circ.172197 \\ 4^\circ.739774 \end{bmatrix}$$

"the sample variance-covariance matrix $\hat{\boldsymbol{\Sigma}}_{\mathbf{y}}$ of type BIQUUE with *Theorem* 4.4"

$$\hat{\boldsymbol{\Sigma}}_{\mathbf{y}} = \begin{bmatrix} 1.7301 & 0.2955 & 1.3825 & 0.0842 & 0.2180 & 0.9666 \\ 0.2955 & 0.1158 & 0.3337 & 0.0296 & 0.1284 & 0.3793 \\ 1.3825 & 0.3337 & 1.7775 & 0.0707 & 0.4085 & 2.2412 \\ 0.0842 & 0.0296 & 0.0707 & 0.0084 & 0.0283 & 0.0549 \\ 0.2180 & 0.1284 & 0.4085 & 0.0283 & 0.1729 & 0.6346 \\ 0.9666 & 0.3793 & 2.2412 & 0.0549 & 0.6346 & 3.7467 \end{bmatrix} (10^{-7}\ \text{strain/y})^2$$

"the related dispersion matrix of $\hat{\boldsymbol{\xi}}$ with *Theorem* 4.3"

$$D\{\hat{\boldsymbol{\xi}}\} = \boldsymbol{\Sigma}_{\hat{\xi}} = \begin{bmatrix} 0.1304 & 0.0054 & 0.0872 & 0.0022 & 0.0346 & 0.0063 \\ 0.0054 & 0.0005 & -0.0077 & -0.0001 & 0.0006 & -0.0001 \\ 0.0872 & -0.0077 & 0.9820 & -0.0160 & 0.0889 & -0.0050 \\ 0.0022 & -0.0001 & -0.0160 & 0.0021 & -0.0005 & 0.0026 \\ 0.0346 & 0.0006 & 0.0889 & -0.0005 & 0.0139 & 0.0012 \\ 0.0063 & -0.0001 & -0.0050 & 0.0026 & 0.0012 & 0.0036 \end{bmatrix}$$

With these estimates of the eigenspace components of the random strain rate tensor and their dispersion matrix the following multivariate hypothesis tests discussed in Section 5.3 can be performed:

- Test for the eigenspace parameter vector $\boldsymbol{\xi} = \boldsymbol{\xi}_0$ with $\boldsymbol{\Sigma}_{\mathbf{y}}$ unspecified;
- Test for a distinct element of the eigenspace parameter vector with *Student t- test* (see *Box.6.9*);
- *Eigen-inference* about the orthonormally transformed parameters $\boldsymbol{\eta}$;
- Test for the variance-covariance matrix $\boldsymbol{\Sigma}_{\mathbf{y}} = \boldsymbol{\Sigma}_0$;
- Test for the eigenspace parameter vector and variance-covariance matrix $\boldsymbol{\xi} = \boldsymbol{\xi}_0$, $\boldsymbol{\Sigma}_{\mathbf{y}} = \boldsymbol{\Sigma}_0$;
- The general linear hypothesis test with a *growth curve model* for eigenspace parameters.

Here we just make the second one - *the Student t- test* - for the distinct element of the eigenspace parameter vector.

Box 6.9

Separate Student t-tests about the eigenspace parameters in $\boldsymbol{\xi}$

Second test for (separately)

$$\mathcal{H}_{02}: \ \lambda_1 = \lambda_{10} = -0.8199,\ \lambda_2 = \lambda_{20} = 0.0414,\ \lambda_3 = \lambda_{30} = 1.4846\ (10^{-7}\ \text{strain/yr})$$
$$\theta_{32} = \theta_{320} = 0.2958,\ \theta_{31} = \theta_{310} = 0.1211,\ \theta_{21} = \theta_{210} = 0.1047\ (\text{arc})$$
$$\mathcal{H}_{12}: \ \lambda_1 \neq \lambda_{10} = -0.8199,\ \lambda_2 \neq \lambda_{20} = 0.0414,\ \lambda_3 \neq \lambda_{30} = 1.4846\ (10^{-7}\ \text{strain/yr})$$
$$\theta_{32} \neq \theta_{320} = 0.2958,\ \theta_{31} \neq \theta_{310} = 0.1211,\ \theta_{21} \neq \theta_{210} = 0.1047\ (\text{arc})$$

"two-sided tests with *the test quantities*"

$$t_1 := \frac{\hat{\lambda}_1 - \lambda_{10}}{\hat{\sigma}_1},\quad t_2 := \frac{\hat{\lambda}_2 - \lambda_{20}}{\hat{\sigma}_2},\quad t_3 := \frac{\hat{\lambda}_3 - \lambda_{30}}{\hat{\sigma}_3}$$
$$t_4 := \frac{\hat{\theta}_{32} - \theta_{320}}{\hat{\sigma}_4},\quad t_5 := \frac{\hat{\theta}_{31} - \theta_{310}}{\hat{\sigma}_5},\quad t_6 := \frac{\hat{\theta}_{21} - \lambda_{30}}{\hat{\sigma}_6}$$

with respect to $\hat{\lambda}_1$, $\hat{\lambda}_2$, $\hat{\lambda}_3$, $\hat{\theta}_{32}$, $\hat{\theta}_{31}$ and $\hat{\theta}_{21}$ of type $\boldsymbol{\Sigma}$ -BLUUE and their variances from *Box 6.8*. t_1, t_2 t_3, t_4 t_5 *and* t_6 are elements of the *Student t-distribution with n-1 degrees of freedom*.

The probability identity

$$P\{c_1 \leq t \leq c_2\} = P\{c_1\hat{\sigma} + \mu_0 \leq \hat{\mu} \leq c_2\hat{\sigma} + \mu_0\} = 1 - \alpha = \gamma$$

relates the *error probability* α of the two-sided test to the confidence level γ. If $\hat{\mu}$ is an element of the confidence interval $c_1\hat{\sigma}+\mu_0 \le \hat{\mu} \le c_2\hat{\sigma}+\mu_0$, the null hypothesis $\mathcal{H}_0 : \mu = \mu_0$ is accepted. We reject $\mathcal{H}_0$ if the confidence interval does not contain $\hat{\mu}$.

With *error probability* α =5% we derive

$$t_{1,1-\alpha/2} = t_{2,1-\alpha/2} = \cdots = t_{6,1-\alpha/2} = +2.57$$
$$t_{1,\alpha/2} = t_{2,\alpha/2} = \cdots = t_{6,\alpha/2} = -2.57.$$

The critical values

$$c_{\lambda_1,\alpha/2} = \hat{\sigma}_{\lambda_1} t_{1,\alpha/2} + \lambda_{10} = \text{-}1.7482, \quad c_{\lambda_1,1-\alpha/2} = \hat{\sigma}_{\lambda_1} t_{1,1-\alpha/2} + \lambda_{10} = 0.1083$$
$$c_{\lambda_2,\alpha/2} = \hat{\sigma}_{\lambda_2} t_{2,\alpha/2} + \lambda_{20} = \text{-}0.0134, \quad c_{\lambda_2,1-\alpha/2} = \hat{\sigma}_{\lambda_2} t_{2,1-\alpha/2} + \lambda_{20} = 0.0962$$
$$c_{\lambda_3,\alpha/2} = \hat{\sigma}_{\lambda_3} t_{3,\alpha/2} + \lambda_{30} = \text{-}1.0627, \quad c_{\lambda_3,1-\alpha/2} = \hat{\sigma}_{\lambda_3} t_{3,1-\alpha/2} + \lambda_{30} = 4.0319$$
$$c_{\theta_{32},\alpha/2} = \hat{\sigma}_{\theta_{32}} t_{4,\alpha/2} + \theta_{320} = 0.1785, \quad c_{\theta_{32},1-\alpha/2} = \hat{\sigma}_{\theta_{32}} t_{4,1-\alpha/2} + \theta_{320} = 0.4132$$
$$c_{\theta_{31},\alpha/2} = \hat{\sigma}_{\theta_{31}} t_{5,\alpha/2} + \theta_{310} = \text{-}0.1818\,, \quad c_{\theta_{31},1-\alpha/2} = \hat{\sigma}_{\theta_{31}} t_{5,1-\alpha/2} + \theta_{310} = 0.4239$$
$$c_{\theta_{21},\alpha/2} = \hat{\sigma}_{\theta_{21}} t_{6,\alpha/2} + \theta_{210} = \text{-}0.0506, \quad c_{\theta_{21},1-\alpha/2} = \hat{\sigma}_{\theta_{21}} t_{6,1-\alpha/2} + \theta_{210} = 0.2599$$

indicate the confidence intervals

$$c_{\lambda_1,\alpha/2} = \text{-}1.7482 < \hat{\lambda}_1 = \text{-}1.049206 < c_{\lambda_1,1-\alpha/2} = 0.1083,$$
$$c_{\lambda_2,\alpha/2} = \text{-}0.0134 < \hat{\lambda}_2 = 0.013154 < c_{\lambda_2,1-\alpha/2} = 0.0962$$
$$c_{\lambda_3,\alpha/2} = \text{-}1.0627 < \hat{\lambda}_3 = 2.953239 < c_{\lambda_3,1-\alpha/2} = 4.0319$$
$$c_{\theta_{32},\alpha/2} = 0.1785 < \hat{\theta}_{32} = 0.237242 < c_{\theta_{32},1-\alpha/2} = 0.4132$$
$$c_{\theta_{31},\alpha/2} = \text{-}0.1818 < \hat{\theta}_{31} = 0.212445 < c_{\theta_{31},1-\alpha/2} = 0.4239$$
$$c_{\theta_{21},\alpha/2} = \text{-}0.0506 < \hat{\theta}_{21} = 0.082725 < c_{\theta_{21},1-\alpha/2} = 0.2599$$

thereby suggesting the acceptance of all six null hypotheses.

$$\mathcal{H}_{02}: \ \lambda_1 = \lambda_{10} = \text{-}0.8199, \ \lambda_2 = \lambda_{20} = 0.0414, \ \lambda_3 = \lambda_{30} = 1.4846 \ (10^{-7} \text{ strain/yr})$$
$$\theta_{32} = \theta_{320} = 0.2958, \ \theta_{31} = \theta_{310} = 0.1211, \ \theta_{21} = \theta_{210} = 0.1047 \ (\text{arc})$$

with the risk of α =5% of a *Type I error*.

The 95% *confidence intervals* for the eigenvalues λ_1, λ_2, λ_3 and the three rotation angles θ_{32}, θ_{31} and θ_{21} are

[-1.7482, 0.1083] (10^{-7} strain/y);
[-0.0134, 0.0962] (10^{-7} strain/y);
[-1.0627, 4.0319] (10^{-7} strain/y)
[10°.2245 23°.6754]
[-10°.4145 24°.2872]
[-2°.9003 14°.8930]

respectively.

Chapter 7

Conclusions

This chapter will conclude the main contributions and results in this study and makes a prospect of more possible applications of the developed theory, methods and further investigations.

With the new space geodetic techniques such as GPS, VLBI, SLR and DORIS, three-dimensional position and change rate of network stations can be highly accurately determined from regular measurement campaigns, which have became an accurate and reliable source of information in Earth deformation studies. This fact suggests that the components of deformation measures, for instance the stress or strain tensor, can be estimated from the highly accurate geodetic data and analyzed through the proper statistical testing procedures. The *eigenspace components* of these random deformation tensors (principal components, principal directions) are of focal interest in geodesy, geology and geophysics. They play an important role in interpreting the geodetic-geological-geophysical phenomena such as earthquakes (seismic deformations), plate motions, and plate deformations among others.

Having recognized the facts that an exact distribution theory of eigenspace components of a symmetric random tensor is almost always unavailable, i.e. the distributions of the eigenvalues and eigendirections of a symmetric random tensor is different from the normal distribution, and a direct statistical inference of them in real Engineering and Earth Science problems can hardly be performed. We have investigated the statistical inference of eigenspace components of a 2-D and 3-D symmetric rank-two random tensor based upon a linearized multivariate *Gauss-Markov* model, which could provide us with the second-order statistics of eigenspace components. Such a statistical inference on the estimates of eigenspace components of a random tensor is completed by the design of a linear hypothesis test. For this purpose first in Chapter 1 we have systematically studied the sampling distribution of the sample mean vector and sample variance-covariance of the direction observation of a random tensors, which proves that the vectorized three-dimensional symmetric random tensor $\mathbf{y} = \text{vech}\,\mathbf{T} \in \mathbb{R}^{6\times1}$ has a BLUUE estimate $\hat{\boldsymbol{\mu}}_\text{y} \in \mathbb{R}^{6\times1}$ which is multivariate normally distributed, $\hat{\boldsymbol{\mu}}_\text{y} \sim \mathcal{N}_6(\boldsymbol{\mu}_\text{y}, n^{-1}\boldsymbol{\Sigma}_\text{y}; \hat{\boldsymbol{\mu}}_\text{y})$, where n is the number of full tensor observations and $\boldsymbol{\Sigma}_\text{y} = D\{\text{vech}\,\mathbf{T}\}$, the variance-covariance matrix of $\mathbf{y}$. The BIQUUE sample variance-covariance component matrix $\hat{\boldsymbol{\Sigma}}_\text{y}$ is Wishart distributed $\hat{\boldsymbol{\Sigma}}_\text{y} \sim \mathcal{W}_6(n-1,(n-1)^{-1}\boldsymbol{\Sigma}_\text{y};\hat{\boldsymbol{\Sigma}}_\text{y})$. Further in Chapter 2 we have proposed the multivariate testing of hypotheses concerning the sample mean vector and the sample variance-covariance component matrix, i.e. the estimated parameters (mean vector and covariance matrix) of tensor-valued multivariate normal population of a two and 3-D, symmetric rank-two random tensor.

For its linearized form of a special nonlinear multivariate Gauss-Markov model for sampling the eigenspace *synthesis* of a two-dimensional, symmetric rank-two random tensor, the *BLUUE of the eigenspace elements* and *BIQUUE* of its variance-covariance component matrix have been established successfully in *Theorem* 4.3 and 4.4. The proper test statistics, such as *Hotelling's* T^2, *likelihood ratio statistics and* the general linear hypothesis test with a *growth curve model*, are proposed. For the three-dimensional symmetric random tensor we have uniquely established the *eigenspace analysis and synthesis* in *Corollary* 5.2 of a three-dimensional symmetric random tensor based on the choice of the orthogonal similarity transformation matrices in (5.22) to (5.28). This leads to the generalization of the BLUUE of the eigenspace elements of three-dimensional random tensor and BIQUUE of its variance-covariance component matrix in three-dimensional case.

As two case studies both estimates BLUUE and BIQUUE and hypothesis tests have been applied successfully to the eigenspace components of 2-D and 3-D strain rate tensor observations in the area of the central Mediterranean and Western Europe, which are derived from ITRF92 to ITRF2000 series station positions and velocities in Sections 6.6.and 6.7. The analysis with respect to geodynamical and statistical aspects shows that, in general, our estimates of the eigenspace components of a two-dimensional strain rate tensors is consistent with the tectonic setting in the area of the central Mediterranean and Western Europe. Furthermore we can benefit from the statistical information derived from the estimation procedure. For example the 95% confidence intervals for the estimates of eigenvalues $\hat{\lambda}_1$, $\hat{\lambda}_2$ and eigendirection $\hat{\alpha}_1$ illustrated in Figure 6.13, provides us with a visual presentation of the possible magnitude and the directions of the extension and contraction of the strain rate, which is important for the prediction of the tectonic activity including the possible deformation trend and directions.

Thanks to the uniquely established *eigenspace analysis and synthesis* we have estimated the eigenspace component parameters and their dispersion matrix from the 3-D, symmetric rank two strain rate observations of six epochs. These estimates reflect the statistical average information of the random strain rate tensor, utilizing the advantage of the longer time spanner. With them we can successively perform the statistical inference. The 95%

confidence regions illustrated on the unit sphere in Figure 6.15 provides us with the three-dimensional visual presentation of the possible magnitude and the directions of the extension and contraction of the strain rate, which is also important for the prediction of the three-dimensional deformation.

The related *linear hypothesis tests* in these two case studies in Section 6.6 and 6.7 have documented large confidence regions for the eigenspace components, namely *eigenvalues and eigendirections*, based upon real measurement configurations. They lead to the statement *to be cautious* with data of type extension and contraction as well as with the orientation of principal stretches.

It is necessary to note that, although the strain rate tensor observations are derived from the nine ITRF sites according to the criterion discussed above, in reality they don't satisfy the all conditions of i.i.d. observations. Since we have not yet found the right i.i.d. strain tensor observation sets, we apply strain rate tensor observations derived from the nine ITRF stations in six series realizations, assuming approximately that they are i.i.d. observations in our study.

Since numerical tests have documented that the estimate $\hat{\xi}$ of type BLUUE of the parameter vector ξ within a linear *Gauss-Markov model* $\{\mathbf{A}\xi = E\{\mathbf{y}\},\ \mathbf{\Sigma}_\mathbf{y} = D\{\mathbf{y}\}\}$ is *not* robust against *outliers* in the stochastic observation vector $\mathbf{y}$, w*e give up* the postulate of unbiasedness, but keep the set-up of a *linear estimation* $\hat{\xi} = \mathbf{Ly}$ of homogeneous type. According to best linear estimators of type homBLE (*Best homogeneously Linear Estimation*), S-homBLE and α-homBLE of the *fixed effects* ξ (*Grafarend and Schaffrin* 1993, *Schaffrin* 2000), we have developed a new method of determining the optimal regularization parameter α in uniform Tykhonov-Phillips regularization (α-weighted BLE) by minimizing the trace of the Mean Square Error matrix $MSE\{\hat{\xi}\}$ (A-optimal design) in the general case. This estimation formula is closed, which provides us not only with the optimal regularization parameter but also with more quicker and more practical solutions than by other methods such as by means of *L*-Curve (*Hansen* 1992) or the C_p-Plot (*Mallows* 1973). Further, it has been shown that the optimal ridge parameter *k* in *ridge regression* as developed by *Hoerl and Kennard* (1970a, 1970b) and *Hoerl, Kennard and Baldwin* (1975) is just the special case of our general solution by A-optimal design. Based on the introduction of the multivariate $\alpha-$homBLE for the multivariate parameters, the determination of the optimal weight factor α has also been generalized to the multivariate Gauss-Markov model, which we shall call "*multivariate ridge estimator*".

Through the above six chapters of this dissertation we have achieved the complete solution to the statistical inference of *eigenspace components* of the deformation tensors. The models developed in the last two chapters are closed and practical. The results bring a sound meaning to the deformation analysis. With these models we could successfully perform the statistical inference of the *eigenspace components* and the variance-covariance matrix of the *Gauss-Laplace* normally distributed observations of a random deformation tensor (case study: two- and three-dimensional, symmetric rank two strain rate tensor).

Beyond the two case studies in Chapter 6 there are surely further applications and investigations of statistical inference for the eigenspace components of a deformation tensor.

As we have mentioned before, for the estimation of eigenspace components of type BLUUE, we need the observations of random tensors. Except for our case study with strain rate tensor observations derived from the ITRF station coordinates and their velocities in six series realizations, there are other types of symmetric rank-two random tensors such as stress and strain. Especially with benefit to the new development of space geodetic techniques such as GPS, the time series observations of a network can be achieved with higher accurate. The time series observations as samples enable us to apply the estimate BLUUE of the eigenspace components of 2-D or 3-D random tensors and BIQUUE of their variance-covariance components matrix in a more practical way and could be more realistic results.

Regarding the facts in reality, crustal motions and deformations are of three-dimensional nature. Most tensors in Geodesy and Geophysics are three-dimensional having been derived from geodetic, geological and seismological data. Our estimation theory as developed in Chapter 5 can also be applied to the statistical analysis of the estimates of the eigenspace components of the three-dimensional stress/strain or seismic moment tensors with in situ measurements by strain meter and by seismometer, respectively, which are of focal interest in geophysics and seismology.

The new method of determining the optimal regularization parameter α in uniform Tykhonov-Phillips regularization (α-weighted BLE) by minimizing the trace of the Mean Square Error matrix $MSE\{\hat{\xi}\}$ (*A-optimal design*) can be applied in the general case not only to replace the simple BLUUE of direct observations but also the general BLUUE in a linear Gauss-Markov model. It can bring also a sound solution to the improperly posed problems which appear in solving the downward continuation problems in potential theory.

References

Abraham R., Marsden J. and Ratiu T. (1983): Manifolds, Tensor Analysis and Applications, Addison-Wesley Publ. Comp., London

Aki K. and Richards P.G. (1980): Quantitative Seismology, W.H. Freeman, San Francisco

Allen D.M.(1971): Mean square error of prediction as a criterion for selecting variables, Technometrics 13 (1971) 469-481

Allen D.M.(1974): The relationship between variable selection and data argumentation and a method for prediction, Technometrics 16 (1974) 125-127

Altamimi Z., P. Sillard and C. Boucher (2002): ITRF2000 : A New Release of the International Terrestrial Reference Frame for Earth Science Applications, JGR, 107(B10), 2214, 2-17-2-18, 2002.

Altiner Y. (1996): Geometrische Modellierung innerer und äußerer Deformationen der Erdoberfläche, Deutsche Geod. Kommission, Bayer. Akad. der Wissenschaften, Report C 462, München

Altiner Y. (1999): Analytical surface deformation theory for detection of the Earth's Crust Movement, Springer Berlin, Heidelberg, 1999

Anderson T.W. (1958): An Introduction to Multivariate Statistical Analysis, John Wiley & Sons, New York

Anderson T.W. (1984): An Introduction to Multivariate Statistical Analysis (second edition), John Wiley & Sons, New York

Angelier J., Tarantola A., Vallete B. and Manoussis S. (1982): Inversion of Field Data in Fault Tectonics to Obtain the Regional Stress - I. Single Phase Fault Populations: A New Method of Computing the Stress Tensor, Geophys. J. R. Astr. Soc. 69 (1982) 607-621

Angelier, J. (2002): Inversion of earthquake focal mechanisms to obtain the seismotectonic stress IV – a new method free of choice among nodal planes, Geophys. J. Int. 150 (2002) 588-609

Angermann D., Müller H., Gerstl M., Kelm R., Seemüller W. and Vei M. (2001): Laserentfernungsmessungen zu LAGEOS-1 und -2 und ihr Beitrag zu globalen Referenzsystemen, Zeitschrift für Vermessungswesen 126 (2001) 250-256

Anzidei, M., et al (2001): Insights into present-day crustal motion in the central Mediterranean area from GPS surveys, Geophys. J. Int. 146 (2001) 98-110

Argus D. F., Gordon R. G., DeMets C. and Stein S. (1989): Closure of the Africa-Eurasia North America plate motion circuit and tectonics of the Gloria fault. J. Geophys. Res., 44(1989) 5585-5B02.

Argus D.F., R.G. Gordon (1991): No-net-rotation model of current plate velocities incorporating plate motion model NUVEL-1. Geophys. Res. Lett. (18) 2039-2042

Arslan O. and Billor, B.(2000): Robust Liu estimator for regression based on an M-estimator, Journal of Applied Statistics 27 (2000) 39-47

Baarda W. (1968). A Testing Procedure for Use in Geodetic Networks, Neth. Geod. Comm., Publ. Geod., New Ser. 4, Delft

Bapat R.B. (1999): Linear Algebra and Linear Models, 2nd Edition, Springer, New York, Berlin, 1999

Bartlett, M. S. (1939): A note on tests of significance in multivariate analysis, Pro. Of the Cambridge Philosophical Society 35 (1939) 180-185

Bartlett, M. S. (1947): Multivariate analysis, Journal of the Royal Statistical Society, Supplement, 9 (1947) 176-197

Behrens T. (1929): Ein Beitrag zur Fehlerberechnung bei wenigen Beobachtungen, Landwirtsch. Jahrb. 68 (1929) 807-837

Beutler G. (2000): The impact of modern space geodetic observation techniques on geodesy, geodynamics, and on atmosphere sciences (Solicited Paper), EGS XXV General Assembly, Nice, France, April 2000

Bigi G., D. Cosentino, M. Parotto, R. Sartori and P. Scandone, Structural model of Italy, Sheet n. 1 Edited by Castellarin A., Coli M., Dal Piaz G.V., Sartori R., Scandone P. and Vai G.B. CNR, progetto finalizzato Geodinamica, Roma 1990

Boucher C. and Z. Altamimi(1993): Development of a Conventional Terrestrial Reference Frame, in Contributions of Space geodesy to geodynamics: Earth dynamics, D. E. Smith and D. L. Turcotte (Eds), AGU, 89-98

Boucher, C., Z. Altamimi and L. Duhem (1994): Results and Analysis of the ITRF 93, IERS Technical Note 18, Observatoire de Paris, 1994

Boucher C., Z. Altamimi, M. Feissel, P. Sillard, (1996): Results and analysis of the ITRF94, IERS Techn. Note 20, Central Bureau of IERS Observatoire de Paris, 1996

Boucher C., Z. Altamimi, P. Sillard (1998): Results and analysis of the ITRF96, IERS Technical Note 24, Central Bureau of IERS Observatoire de Paris, May 1998

Boucher C., Z. Altamimi, P. Sillard (1999): The 1997 International Terrestrial Reference Frame (ITRF 97), IERS Technical Note 27, Central Bureau of IERS Observatoire de Paris, May 1999

Bouman, J. (1998): Quality of regularization methods, DEOS Report no. 98.2, Delft University Press 1998

Box G.E.P.(1949): A general distribution for a class of likelihood criteria. Biometrika 36, 317-346

Brown P.(1993) Measurement, regression, and calibration, Oxford Science publications, 1993

Brunner F.K. (1979): On the analysis of geodetic networks for the determination of the incremental strain tensor, Survey Review XXV 192 (1979) 56-67

Cacoullos T. (1965): On the bias of functions of characteristic roots of a random matrix, Biometrica, 52(1965) 87-94

Cai J. (2001): The Hypothesis tests and sampling statistics of the eigenvalues and eigendirections of a random tensor of type deformation tensor, Diploma thesis, Department of Geodesy and GeoInformatics, University of Stuttgart, October 2001

Cai J. and Grafarend E. (2000): The sample distribution of the sample means and sample variance-covariance of a three dimensional, symmetric rank two random tensor (strain, stress), Geodetic Week, Potsdam, 10.-12. October 2000

Cai J. and Grafarend E. (2001a): Inference on Deformation Measures, Proc. of the First International Symposium on Robust Statistics and Fuzzy Techniques in Geodesy and GIS, IAG - SSG 4.190, 12-16 March, 2001. Zurich. Eds. A. Carosio and H. Kutterer, Institute of Geodesy and Photogrammetry, ETH, p.85, Zurich 2001

Cai J. and Grafarend E. (2001b): Statistical Inference on Deformation Measures of the Earth Surface Derived from Space Geodetic Measurements, IAG 2001 Scientific Assembly, Budapest, Hungary, 2-7 September, 2001

Cai J., Grafarend E. and Schaffrin B. (2001a): The α-weighted BLE of the direct observations of a symmetric random tensor, Simulated case study: 2-dimensional strain, Geodetic Week, Köln, 18.-21. September 2001

Cai J., Grafarend E. and Schaffrin B. (2001b): The $\mathbf{\Sigma}$-BLUUE of eigenspace components of a symmetric random tensor of second order, Simulated case study: 2-dimensional strain, Geodetic Week, Köln, 18.-21. September 2001

Caporali A., S. Martin, Massironi, M. and Baccini S. (2001): State of Strain Rate in the italian crust from geodetic data, International Association of Geodesy / Section I – Positioning; Subcommission for Europe (EUREF), Publication No. 10, Report on the Symposium of the IAG Subcommission for Europe (EUREF) in Dubrobnik, Croatia, 16 – 18 May 2001

Caporali, A. (2003a): Average strain rate in the Italian crust inferred from a permanent GPS-network – I. Statistical analysis of the time-series of permanent GPS stations, Geophys. J. Int. 155 (2003) 241-253

Caporali, A., Martin S. and Massironi M. (2003b): Average strain rate in the Italian crust inferred from a permanent GPS-network – II. Strain rate versus seismicity and structural geology, Geophys. J. Int. 155 (2003) 254-268

Castellarin A, Cantelli L, Fesce AM, Mercier JL, Picotti V, Pinni GA, Prosser G, Selli L (1992): Alpine compressional tectonics in the Southern Alps: Relationships with the Apennines. Ann. Tecton 6 (1992) 62–94

Chase, C.G. (1978): Plate kinematics: The Americas, East Africa, and the rest of the world. Earth and Planetary Science Letters, 37(1978) 355-368..

Chaturvedi A. and Singh, S.P.(2000): Stein rule prediction of the composite target function in a general linear regression model, Statistical Papers *41* (2000) 359-367

Chen Y.Q. and Chrzanowski A. (1996): Identification of Deformation Models in Space and Time Domain, Survey Review 33 (1996) 518-528

Chrzanowski A. and Chen Y.Q. (1990): Deformation Monitoring, Analysis, and Prediction - Status Report, Proc. FIG XIX Int. Congress, Helsinki, Vol. 6 (1990) 84-97

Chrzanowski A., Chen Y.Q. and Secord M. (1983): On the Strain Analysis of Tectonic Movements Using Fault Crossing Geodetic Surveys, Tectonophysics 97 (1983) 297-315

Ciarlet P.G. (1988): Mathematical Elasticity, Vol.1: Three-Dimensional Elasticity, North Holland Publ. Comp., Amsterdam

Clarke PJ, Davies RR, England PC, Parsons B, Billiris H, Paradissis D, Veis G, Cross PA, Denys PH, Ashkenazi V, Bingley R, Khale HG, Muller MV, Briole P (1998): Crustal strain in central Greece from repeated GPS measurements in the interval 1989–1997. Geophys J Int 135(1998)195–214

Cohen J., Kesten H. and Newman C. (eds.) (1984): Random Matrices and their applications, American Mathematical society, Providence, Rhode Island, 1984

Constantine, A.G. (1966): The distribution of Hotelling's generalized T_0^2, Ann. Math. Statist. 37 (1966) 215-225

Cox D. R. and N. Reid (2000): The Theory of the Design of Experiments, Chapman & Hall/CRC, 2000

Cramer H. (1946): Mathematical Methods of statistics, Princeton Uni. Press. Princeton, NJ

Cross PA, Hannah J, Hradilek L, Kelm R, Ma¨ kinen J, Merry CL, Sjoeberg LE, Steeves RR, Vanicek P, Zilkoski DB (1987): Four-dimensional geodetic positioning. Manuscr Geod 12 (1987)147–222

Crowther N. and Young D. (1974): Notes on the distribution of characteristic roots and functions of characteristic roots of certain matrices in multivariate analysis, Technical Report No. 92, Department of Statistics, Standford University, 1974

Das Gupta S. (1969): Properties of power functions of some test concerning dispersion matrices of multivariate normal distributions, Ann. Math. Statist., 40, 697-701

Davidson, Jon P., Reed, Walter E., Davis, Paul M (2002): Exploring Earth: An introduction to physical geology 2nd Ed., Prentice Hall, Upper Saddle River, NJ

Davis A.W.(1968): A System of Linear Differential Equations for the Distribution of Hotelling's Generalized T_0^2, The Annals of Mathematical Statistics, 39(1968) 815-832.

Davis A.W.(1970a): Further applications of a differential equation for the distribution of Hotelling's generalized T_0^2, Ann. Inst. Statist. Math. 22 77-87

Davis A.W. (1970b): Exact distribution of Hotelling's generalized T_0^2, Biometrika 57 (1970b) 187-191

Davis A.W. (1972): On the Marginal Distributions of the Latent Roots of the Multivariate Beta Matrix, The Annals of Mathematical Statistics, 43(1972) 1664-1670

Dawe D.J. (1984): Matrix and Finite Element Displacement Analysis of Structures, Clarendon Press, Oxford

Dawid A. P. (1981): Some matrix variate distribution theory: Notational considerations and a Bayesian application, Biometrika 68 (1981) 265-274

Deemer W. L. and Olkin I. (1951) The Jacobians of certain matrix transformations useful in multivariate analysis, Biometrika, 345-367

DeMets C and Dixon TH (1999): New kinematic models for Pacific North America motion from 3 Ma to present, I: evidence for steady motion and biases in the NUVEL-1A model. Geophys Res Lett 26, 1921–1924

DeMets, C., Gordon, R. G., Argus, D. F. and Stein, S. (1990): Current Plate Motions, Geophys. J. Int *101* (1990) 425-478

DeMets, C., Gordon, R. G., Argus, D. F. and Stein, S. (1994): Effect of recent revisions to the geomagnetic reversal time scale on estimates of current plate motions, Geophysical Research Letters, *21* (1994) 2191-2194

Dermanis A. and Liveratos E. (1983): Application of deformation analysis in Geodesy and geodynamics, Rev. of Geophysics and Space Physics Vol. 21 No. 1 (1983) 41-50

Dermanis, A. and E.W. Grafarend (1993): The finite element approach to the geodetic computation of two-and threedimensional deformation parameters: a study of frame invariance and parameter estimability, Proc. Of the Int. Conf. On Cartography-Geodesy, Maracaibo, 1992, 66-85

Dermanis, A. (2001): private communication, July 2001

Devoti R., Ferraro C., ., Lanotte R., Luceri V., Nardi A., Pacione R., Rutigliano P., Sciarretta C. Gueguen E., Biaoco G. and Vespe F. (2002a): Geophysical interpretations of geodetic deformations in the central

Mediterranean area, in Plate Boundary Zones, (Eds.) Stein S. and Freymuller J. T., Geodynamics Series Vol. 30, AGU, Washington D.C 57-66

Devoti R., Ferraro C., Gueguen E., Lanotte R., Luceri V., Nardi A., Pacione R., Rutigliano P., Sciarretta C. and Vespe F. (2002b): Geodetic control on recent tectonic movements in the central Mediterranean area, Tectonophysics 346 (2002) 151-167

Donatos G.S. and Michailidis, G.C.(1990): Small sample properties of ridge estimators with normal and non-normal disturbances, Commun. Statist.-Simula. 19 (1990) 935-950

Draper R.N. and Van Nostrand, C.R.(1979): Ridge regression and James- Stein estimation : Review and comments, Technometrics *21* (1979) 451-466

Drewes H. (1999): Realisierung des kinematischen terrestrischen Referenzsystems ohne globale Rotation ("no net rotation"), Mitt. Bundesamt f. Kart. u. Geod. 5, 120-125, Frankfurt a. M.

Drewes H. and Angermann D (2001): The Actual Plate Kinematic and Crustal Deformation Model 2000 (APKIM2000) as a Geodetic Reference, IAG 2001 Scientific Assembly, Budapest, 2.-8.9.2001

Droge B.(1993): On finite-sample properties of adaptative least squares regression estimates, Statistics 24 (1993) 181-203

Dvorak G. and Benveniste Y. (1992): On Transformation Strains and Uniform Fields in Multiphase Elastic Media, Proc. R. Soc. Lond., A437 (1992) 291-310

Edelman A (1989): Eigenvalues and Condition Numbers of Random Matrices, Doctoral thesis, Department of Mathematics of Massachusetts Institute of Technology, 1989

Edelman A (1999): Random Eigenvalues, Work in progress, 1999

EL-Sayed, S.M.(1996): The sampling distribution of Ridge parameter estimator, The Egyptian Statistical Journal 40 (1996) 211-219

Engels J., Grafarend E. Keller W. Martinec Z. Sanso and F. Vanicek P. (1993): The geoid as an inverse problem to be regularized, Proceedings of the international conference "Inverse Problems: Principles and Applications in Geophysics, Technology, and Medicine" Anger, G. et al. (eds). Akademie Verlag, Berlin 1993

Engl H. (1993) Regularization methods for the stable solution of inverse problems, Surveys on Math. For Industry, 3 (1993) 71-143

Eringen A.C. (1962): Nonlinear Theory of Continuous Media, McGraw Hill Book Comp., New York

Farebrother R.W.(1975): The minimum mean square error linear estimator and ridge regression, Technometrics 17 (1975) 127-128

Farebrother R.W.(1976): Further results on the mean square error of ridge regression, Journal of the Royal Statistical Society B.38 (1976) 248-250

Farebrother R.W.(1978): Partitioned ridge regression, Technometrics 20 (1978) 121-122

Feigl K., King R. and Jordan T. (1990) : Geodetic measurement of tectonic deformation in the Santa Maria fold and thrust belt, California. J. Geophysical research. 95 (1990) 2679-2699

Firinguetti L.(1996): A simulation study of ridge regression in an incorrectly specified model, Estadistica 48 (1996) 115-154

Firinguetti L. and Rubio, H. (2000): A note on the moments of stochastic shrinkage parameters in ridge regression, Commun. Statist.-Simula. 29 (2000) 955-970

Fisher R.A.(1915): Frequency distribution of the values of the correlation coefficient in samples from an infinitely large population, Biometrika 10, 507-521

Fisher R. A. (1920): A mathematical examination of the methods of determining the accuracy of an observation by the mean error, and by the mean square error, Mon. Not. R. Astron. Soc. 80, 758-770

Fisher R. A. (1922): On the mathematical foundations of theoretical statistics, Philos Trans., R. Soc. London, A 222, 309-368

Fisher R.A. (1939): The sampling distribution of some statistics obtained from non-linear equation, Ann. Eugenics, Volume 9, pp. 238-249

Flesch L., Holt W, Haines J, and Shen-Tu B. (2000): Dynamics of the Pacific-North American Plate Boundary in the Western United States, Science Vol. 287 (2000) 834-836

Flügge W. (1972): Tensor Analysis and Continuum Mechanics, Springer-Verlag Berlin New York 1972

Frank F. C. (1966): Deduction of Earth Strains from Survey Data, Bulletin of the Seismological Society of America 56 (1966) 35-42

Galton F. (1877): Typical laws of heredity, Nature 15 (1877) 492-495, 512-514, 532-533

Galton F. (1885): Presidential address, Section H, Anthropology, Nature 32(1885) 507-510

Galton F. (1886): Regression Towards Mediocrity in Hereditary Stature, Journal of the Anthropological Institute 15 (1986) 246-263 (1886)

Gauss C. F. (1816): Bestimmung der Genauigkeit der Beobachungen, Z. Astron. Verwandte Wiss. 1. (1816) 185-196

Gephart J. and Forsyth D. (1984): An Improved Method for Determining the Regional Stress Tensor using Earthquake Focal Mechanism Data: Application to the San Fernando Earthquake Sequence, J. Geophys. Res. B89 (1984) 9305-9320

Gibbons D.G.(1981): A simulation study of some ridge estimators, Journal of the American Statistical Association 7 (1981) 131-139

Gillard D., Wyss M. and Nakata J. (1992): A Seismotectonic Model for Western Hawaii Based on Stress Tensor Inversion from Fault Plane Solutions, J. Geophys. Res. B97 (1992) 6629-6641

Giri N. (1977): Multivariate statistical inference, AP, New York

Girko V.L. (1979): Distribution of Eigenvalues and Eigenvectors of Hermitian Stochastic Matrices, Ukrain. Math. J. 31 (1979) 421-424

Girko V.L. (1990): Theory of random Determinants, Kluwer Academic Publishers, 1990

Girko V.L. (1995): An introduction to statistical analysis of random arrays, VSP, Utrecht, The Netherlands, 1995

Girko V.L. (2000): 25 years of stochastic canonical equation for normalized spectral functions of random symmetric matrices, Random Oper. and Stoch. Equ., Vol. 8 No.3 (2000) 275-304

Golub, G., Heath, M. and Wahba, G. (1979): Generalized cross-validation as a methods for choosing a good ridge parameter, Technometrics 21 (1979) 215-223

Gordon R. and Stein S. (1992): Global Tectonics and Space Geodesy, Science 256 (1992) 333-342

Gosset S.(1908a): The probable error of a mean, Biometrika 6 (1908) 1-25

Gosset S.(1908b): Probable error of a correlation coefficient, Biometrika 6 (1908) 302-310

Gower J. C. (1966): Some distance properties of latent root and vector methods used in multivariate analysis, Biometrika 53 (1966) 325-338.

Grafarend E. (1977): Strain-stress relations in geodetic network, publ. Geodetic Inst. Uppsala Uni. No. 16, 1977

Grafarend E. (1982): Six Lectures on geodesy and global Geodynamics, Lecture III : Geodetic Reference Frame, Pro. of the third int. Summer School in the Mountains, eds. H. Moritz and H. Sünkel, Mitt. Geod. Inst. TU Graz, Folge 41, 531-685, Graz, 1982

Grafarend E. (1985): Criterion Matrices for Deforming Networks, In: Optimization and Design of Geodetic Networks. Eds.: E. Grafarend and F. Sanso, Springer Verlag, Berlin New York Tokyo, 363-428

Grafarend E. (1986): Three Dimensional Deformation Analysis: Global Vector Spherical Harmonic and Local Element Representation, Tectonophysics 130 (1986) 337-359

Grafarend E. (2000): Hypothesis tests, Lecture Note, University of Stuttgart, 2000

Grafarend E. and Schaffrin B. (1974): Unbiased Free Net Adjustment, Survey Review 22 (1974) 200-218

Grafarend E. and Schaffrin B. (1993): Ausgleichungsrechnung in linear Modellen, Wissenschaftsverlag, Mannheim 1993

Grafarend E. and Dermanis S. (1993): The finite element approach to the geodetic computation of two- and three-dimensional deformation parameters: A study of frame and parameter estimability, Proc. Int. Conf. on Cartography-Geodesy, Maracaibo (Venezuela) 24. November - 3. December 1992, eds. J. Sevilla and H. Henneberg, Instituto de Astronomiay Geodesia, 66-86, Madrid 1993

Grafarend E. and Voosoghi B. (2003): Intrinsic deformation analysis of the Earth's surface based on displacement fields derived from space geodetic measurements. Case studies: Present-day deformation patterns of Europe and of the Mediterranean area (ITRF data sets), Journal of Geodesy 77(2003) 303-326

Grenerczy G.(2002): Tectontic procresses in the Eurasian-African Plate Boundary Zone Revealed by Space Geodesy, in Plate Boundary Zones, (Eds.) Stein S. and Freymuller J. T., , Geodynamics Series Vol. 30, AGU, Washington D.C 67-86

Gui Q., Huang, W., and Zhang, J.(1998): Robust universal ridge estimation, Arta geodetica et cartographica sinica 3 (1998)

Gui Q. and Zhang, J.(1998): Robust biased estimation and its applications in geodetic adjustments, Journal of Geodesy 72 (1998) 430-435

Gui Q. and Liu, J.(2000): Biased estimation in Gauss-Markov model, AVN *3* (2000) 104-107

Gui Q., Yang, Y. and Guo, J.(2001): Biased estimation in Gauss-Markov model with Constraints, AVN *1* (2001) 28-30

Gunst R.F. and Mason R.L.(1977): Biased Estimation in Regression: An Evaluation Using Mean Squared Error, Journal of the American Statistical Association 72 (1977) 616-628

Gunst R.F. and Mason, R.L.(1980): Regression Analysis and its Application, Marcel Dekker, New York 1980

Gupta A. K and Girko V.L. (Eds.) (1996): Multivariate statistical analysis and theory of random matrices, Proceedings of the Sixth Eugene Lukacs Symposium, Ohio, USA, 29-30 March 1996, VSP, 1996

Gurtin M.E. (1981): An Introduction to Continuum Mechanics, Academic Press, New York

Haines, A. J., W. E. Holt (1993) : A procedure for obtaining the complete horizontal motions within zones of distributed deformation from the inversion of strain rate data, Journal of Geophysical Research 98 (1993) 12,057-12,082

Han Y.Y. (1992): Geometric Analysis of Finite Strain Tensors in Tectonic deformation Mechanics, Sc. China B35 (1992) 102-116

Hanke M. and P. Hansen (1993): Regularization methods for large scale problem, Surveys on Math. For Industry, 3 253-315

Hansen P.(1992): Analysis of discrete ill-posed problems by means of the *L*-curve, SIAM Review *34* 561-580

Hansen P.(1993): The use of the *L*-curve in the regularization of discrete ill-posed problems, SIAM J. Sci. Comput. 14 (1993) 1487-1503

Hartung J. and Werner, H.-J. (1980): Zur Verwendung der restringierten Moore-Penrose Inversen beim Testen von linearen Hypothesen, ZAMM 60 (1980) T344-T346

Hartung J. (1981): Nonnegative Minumum Biased Invariant Estimation in Variance Component Models, The Annals of Statistics 9 (1981) 278-292

Hartung J. and Voet B. (1992): On Planning and Inference in Linear Models with General Covariance Structure, In: Schach S. and Trenkler G. (eds.): Festschrift in Honour of Friedhelm Eicker, Eul-Verlag, Bergisch-Gladbach, 435-452

Hartung J., Elpelt B. and Voet B (1997): Modellkatalog Varianzanalyse, München 1997

Heck, D.L. (1960), Charts on some upper percentage points of the distribution of the largest characteristic root, Ann. Math. Stat., 31, pp. 625-642.

Helmert F. R. (1876a): Über die Wahrscheinlichkeit der Potenzsummen der Beobachtungsfehler und über einig damit im Zusammenhange Stehende Fragen, Z. Math. Phys. 21 (1876) 192-218

Helmert F. R. (1876b): Die Genauigkeit der Formel von Peters zur Berechnung des wahrscheinlichen Beobachtungsfehlers direkter Beobachtungen gleicher Genauigkeit, Astron. Nachr. 88 (1876) 113-132

Helmert F. R. (1880): Die mathematischen und physikalischen Theorien der höheren Geodäsie, Leipzig, B. C. Teubner, 1880

Hemmerle W.J.(1975): An explicit solution for generalized ridge regression, Technometrics 17 (1975) 309-314

Hemmerle W.J. and Brantle, T.F.(1978): Explicit and constrained generalized ridge estimation, Technometrics 20 (1978) 109-120

Henderson H. V., Pukelsheim F. and Searle, S. R. (1981):On the history of the Kronecker product, Paper on BU-733-M, Biometrics Unit, Cornell Universty

Hirvonen R.A. (1971): Adjustment by Least Squares in Geodesy and Photogrammetry, Frederick Ungar Publ. Co., New York

Hocking R. R. (1976) The analysis selection of variables in linear regression, Biometrics 32 (1976) 1-49

Hoerl R. W. (1985): Ridge Analysis 25 Years Later, The American Statistician , 39 (1985) 186-192

Hoerl A. E. and Kennard, R.W.(1970a): Ridge regression: Biased estimation for nonorthogonal problems, Technometrics 12 (1970) 55-67

Hoerl A.E. and Kennard, R.W.(1970b): Ridge regression: Applications to nonorthogonal problems, Technometrics 12 (1970) 69-82

Hoerl A.E., Kennard, R. W. and Baldwin, K.F.(1975): Ridge regression: Some simulations Analysis, Communications in Statistics 4 (1975) 105-123

Hoerl R. W., Schuenemeyer, J. H. and Hoerl, A. E. (1985): A simulation of biased estimation and subset selection regression techniques, Technometrics 28 369-380

Hotelling (1931): The generalization of student's ratio, Ann. Math. Statis. 2, 360-378

Hotelling (1947): Multivariate quality control illustrated by the air testing of sample bombsights, Techniques of statistical analysis (C. Eisenhart, M.W. Hastay, and W.A. Wallis eds.), McGraw-Hill, New York, 113-184

Hotelling (1951): A generalized T-test and measure of multivariate dispersion, Proc. 2nd. Berkeley Symp. Math. Statist. Prob. (J. Neyman, ed.), Univ. California Press, Los Angeles and Berkeley 23-42 (245, 251)

Hsu P. (1938): Note on Hotelling's generalized, Annals of Math. Statistics 9 (1938) 231-243

Hsu P. (1939): A new proof of the joint product moment distribution, Proc. of the Cambridge Phil. Society, 1939

Hsu P. (1941): Canonical Reduction of the general regression problem, Ann. Eugenics, 11 (1941) 42-46

Huang J. Z. (1998): Projection Estimation in Multiple Regression with Application to Functional ANOVA Models, The Annals of Statistics 26 (1998) 242-272

IERS (2003): from http://www.iers.org/iers/

Ilk K. H. (1986): On the regularization of ill-posed problems, Proc. Int. symp. Figure and Dynamics of the Earth, Moon and Planets, Prague 1986, 365-383

Jackson, J. and McKenzie, D. (1988): The relationship between plate motions and seismic moment tensors, and the rates of active deformation in the Mediterranean and Middle East, Geophysical Journal *93* (1988) 45-73

James A. (1954) Normal multivariate analysis and the orthogonal group, Annals of Math. Statistics 25 (1954) 40-75

James TS, Lambert A (1993): A comparison of VLBI data with ICE-3G glacial rebound model. Geophys Res Lett 20:871–874

Jiménez-Munt, I., Sabadini, R. and A. Gardi (2003) : Active deformation in the Mediterranean from Gibraltar to Anatolia inferred from numerical modeling and geodetic and seismological data, Journal of Geophysical Research 108 (2003) 2-1 – 2-24

Jeffreys H. (1961): Cartesian Tensors, Cambridge University Press

Johnson R. A. and Wichern D. W. (1988): Applied Multivariate Statistical Analysis, Prentice Hall, Englewood Cliffs, New Jersey

Kacirattin, S. Sakalloglu, S. and Akdeniz, F. (1998): Mean squared error comparisons of the modified ridge regression estimator and the restricted ridge regression 27 (1998) 131-138

Kagan Y.Y. (1990): Random Stress and Earthquake Statistics: Spatial Dependence, Geophys. J. Int. 102 (1990) 573-583

Kagan Y.Y. (1991): Likelihood Analysis of Earthqualke Catalogues, Geophys. J. Int. 106 (1991) 135-148

Kagan Y.Y. (1992): Correlations of Earthquake Focal Mechanisms, Geophys. J. Int. 110 (1992) 305-320

Kagan Y.Y. and Knopoff L. (1985): The First Order Statistical Moment of the Seismic Moment Tensor, Geophys. J. R. Astr. Soc. 81 (1985) 429-444

Kagan Y.Y. (2000): Temporal correlations of earthquake focal mechanisms, Geophys. J. Int. 143 (2000) 881-897

Kahle, H.-G., Müller, M. V., Geiger, A., Danuser, G., Mueller, S., Veis, G., Billiris, H. and Paradissis, D. (1995): The strain field in northwestern Greece and the Ionian Islands: Results inferred from GPS measurements, Tectonophysics *249* (1995) 41-52

Kahle HG, Straub C, Reilinger R, McClusky S, King R, Hurst K, Veis G, Kasten K, Cross P (1998) The strain field in the eastern Mediterranean region, estimated by repeated GPS measurements. Tectonophysics 294 (1998) 237–252

Kakkuri J. and Chen R. (1992): On Horizontal Crustal Strain in Finland, Bull. Géod. 66 (1992) 12-20

Kendall M.G. and Stuart A. (1958): The Advanced Theory of Statistics, Vol.1, Distribution Theory, Charles Griffin Company Ltd, London

Khatri C. G. (1966): A note on a MANOVA model applied to problems ingrowth curve Ann. Inst. Statist. Math, 18(1966) 75-86

Kirkwood B. H. and Chang T. (1998): Combining Estimates of Tectonic Plate Rotations: An Extension of Welch's Method to Spherical Regression, J. Multivariate Analysis 65 (1998) 71-108

Koch K.R. (1978): Hypothesentests bei singulären Ausgleichungsproblemen, Zeitschrift für Vermessungswesen 103 (1978) 1-10

Koch K.R. (1985): Ein statistisches Auswerteverfahren für Deformationsmessungen, Allgemeine Vermessungs-nachrichten 92 (1985) 97-108

Koch K.R. (1987): Parameterschätzung und Hypothesentests in linearen Modellen, Dümmler, Bonn

Koch K.R. (1991): Moment Tensor Inversion of Local Earthquake Data - I. Investigation of the Method and its Numerical Stability with Model Calculations, Geophys. J. Int. 106 (1991) 305-319

Koch K. (1999): Parameter estimation and hypothesis testing in linear models, Springer-Verlag Berlin Heidel-berg 1999

Koch K. (2001): private communication, 11. 2001

Kreemer C., J. Haines, W. E. Holt, G. Blewitt and D. Lavallee (2000): On the determination of a global strain rate model, Earth Planets Space 52 (2000) 765-770

Kres H. (1983) : Statistical tables for multivariate analysis, Springer-Verlag, New York

Lancaster H. O. (1965): The Helemrt Matrices, Amer. Math. Monthly, 72, 4-13

Laplace P.S (1812) : Theorie analytique des probabilities, Courcier Parise

Lawless J.F. and Wang, P.(1976): A simulation study of ridge and other regression estimations, Communications in Statistics A5 (1976) 307-322

Lawley D.N. (1938): A generalization of Fisher’s z test. Biometrika 30 180-187

Lichtenegger H. and Sünkel H. (1989): Mathematisch-Geophysikalische Model, in Österreichische Beträge zum Wegener-Medlas-Projekt, Mitteilungen der geodätischen Institute der Technischen Universität Graz 65, 61-80

Liski E. P., Mandal, N.K., Shah, K.R. and Sinha, B.K.. (2002) Topic in Optimal Design, Springer, New York and Berlin, 2002

Lisle R. (1992): New Method of Estimating Regional Stress Orientations: Application to Focal Mechanism Data of Recent British Earthquakes, Geophys. J. Int. 110 (1992) 276-282

Liu K (1993): A new class of biased estimates in linear regression, Communications in Statistics A22 (1993) 393-402

Louis A. K. and Maass, P. (1990): A mollifier method for linear operator equation of the first kind, Inverse Prob-lems, 6 (1990) 427-440

Lowerre J. M.(1974): On the mean square error of parameter estimates for some biased estimators, Technomet-rics 16 (1974) 461-464

Ma X.Q. and Kusznir N.J. (1992): 3-D Subsurface Displacement and Strain Fields for Faults and Fault Arrays in a Layered Elastic Half-Space, Geophys. J. Int. 111 (1992) 542-558

Mahalanobis P.C. , Bose R. C. and Roy S. N.(1937): Normalization of statistical variates and the use of rectan-gular coordinates in the theory of sampling distributions, Sankhya 3 (1937) 1-40:

Mallows C. L. (1959): Latent roots and vectors of random matrices, Technical Report 35, Statistical Techniques Research Group, Section of Mathematical Statistics, Department of Mathematics, Princeton Univer-sity, Princeton 1959

Mallows C.L.(1973): Some comments on Cp, Technometrics 15 (1973) 661-675

Mardia K. Kent, J. T., and Bibby, J. M.(1979): Multivariate analysis, Academic Press, Inc., London, 1979

Mardia K. and Goodall, CR.(1993): Spatial-temporal analysis of multivariate environmental monitoring data, In: Multivariate Environmental Statistics Patil G. and Rao C.R. (editors), Elsevier Science publishers. B. V. 1993

Markiewicz A.(19 Marquardt, W.D.(1970): Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation, Technometrics *12* (1970) 591-612

Marquardt W.D.(1974): Discussion No.2, Technometrics 16 (1974) 189-192

Marquardt W.D. and Snee, R.D.(1975): Ridge regression in practice, The American Statistician 29 (1975) 3-2096): Characterization of general ridge estimators, Statistics & Probability Letters 27 145-148

MathWorks (2000): Mapping Toolbox Ver. 1.2, MATLAB V 6.5, 2000

Mayer L.S. and Willke, T.A.(1973): On biased estimation in linear models, Technometrics 15 (1973) 497-508

McCarthy D. D. (2003): IERS Conventions (2003), IERS Technical Note No.32 (to be published)

McDonald G.C. and Galarneau, D.I.(1975): A Monte Carlo evaluation of some ridge-type estimators, Journal of the American Statistical Association 70 (1975) 407-416

McKenzie DP (1970): Plate tectonics of the Mediterranean region, Nature 226, 239–243

Means W.D. (1976): Stress and Strain: Basic concepts of continuum mechanics for geologists. Springer-Verlag New York Heidlberg Berlin

Mehta M.L. (1967): Random Matrices and the Statistical Theory of Energy Levels, Academic Press, New York

Mehta M.L. (1991): Random Matrices (Revised and Enlarged) Second Edition, 1991 Academic Press, New York

Minster, J. B. and Jordan, T. H. (1978): Present-Day Plate Motions, Journal of Geophysical Research *83* (1978) 5331-5354

Montone P., Amato, A., Pondrelli, S. (1999): Active stress map of Italy, JGR. 104 (B11) 25595–25610.

Morrison D. F. (1976): Multivariate statistical methods, 2. ed.. - New York : McGraw-Hill, 1976

Mueller B, Zoback ML, Fuchs K, Mastin L, Gregersen S, Pavoni N, Stephansson O, Ljunggren C (1992) Regional patterns of tectonic stress in Europe, J. Geophys Res 97 (1992)11,783–11,803

Muirhead R.J. (1972a): On the test of independence between two sets of variates, Ann. Math. Statist. 43 (1972a) 1491-1497

Muirhead R.J. (1972b): The asymptotic noncentral distribution of Hotelling's generalized T^2_0, Ann. Math. Statist. 43 (1972b) 1671-1677

Muirhead R. J. (1982): Aspects of multivariate statistical theory, Wiley , New York, 1982

Nocquet J.-M. and E. Calais (2003): Crustal velocity field of western Europe from permanent GPS array solutions, 1996-2001, Geophys. J. Int. 154 (2003) 72-88

Nomura M.(1988): On exact small sample properties of ordinary ridge estimators, Economics Letters 27 (1988) 251-255

Nomura M.(1998): On exact small sample properties of the minimax generalized ridge regression estimators, Commun. Statist.-Simula. 17 (1988) 1213-1229

Ohtani K.(1986): On smalll sample properties of the almost unbiased generalized ridge estimator, Commun. Statist.-Simula. 15 (1986) 1571-1578

Ohtani K.(1998): An MSE comparison of the restricted Stein-rule and minimum mean squared error estimators in regression, Sociedad de Estadistica e Investigacion Operativa Test 7 (1998) 361-376

Olkin and Roy (1954) On Multivariate distribution theory, Annals of Math. Statistics 25 (1954) 329-339

Olkin I. (2000): The 70th Anniversary of the distribution of random Matrices: A Survey, Technical Report No. 2000-06, Department of Statistics, Standford University, 2000

Papazachos C. and Kiratzi A. (1992): A Formulation for Reliable Estimation of Active Crustal Deformation and its Application to Central Greece, Geophys. J. Int. 111 (1992) 424-432

Pazman A. (1986): Foundations of Optimum Experimental Design, Kluwer Academic, Dordrecht, 1986

Pearson K. (1900): On the criterion that a given system of derivation from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling, Philos. Mag. 50 (1900) 157-175

Pearson K. (1901): On lines and planes of closest fit to systems of points in space, The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science, 2(1901) 559- 572

Person K. (1931): Historical Note on the distribution of the standards deviations of samples of any size drawn from an indefinitely large normal parent population. Biometrika, 23, 416-418

Peltier WR (1995): VLBI baseline variations for the ICE-4G model of post glacial rebound, Geophys. Res. Lett. 22 (1995) 465-469

Phillips D.L.(1962): A technique for the numerical solution of certain integral equations of the first kind, J. Ass. Comput. Mach. 9 (1962) 84-96

Pillai K.C.S. and P. Sampson (1959): On Hotelling's generalization of T^2, Biometrika 46 (1959) 160-168

Pillai K.C.S. and Sudjana (1974): On the distribution of Hotelling's trace and power comparisons, Comm. Statist. 3 (1974) 773-779

Pillai K.C.S. and D.C. Young (1971): On the exact distribution of Hotelling's generalized T_0^2, J. Multiv. Anal. 1 (1971) 90-107

Pillai K.C.S. (1960) : Statistical tables for tests of multivariate hypotheses, Manila Statist. Center, Univ. of Philipines

Pondrelli S., Morelli A., and Boschi E., (1995): Seismic deformation in the Mediterranean area estimated by moment tensor summation. *Geophys. J. Int.*, 122, 938-952

Porter C. E. (1965): Statistical Theories of Spectra: Fluctuations (Perspectives in Physics) Academic Press Inc.: London, New York. 1965

Potthoff, R. F. and Roy, S. N. (1964). A generalized multivariate analysis of variance model useful especially for growth curve problems. *Biometrika* 51, 313-326

Prescott W. H. (1976): An extension of Frank's method for obtaining crustal shear strains from survey data, Bulletin of the Seismological Society of America *66* (1976) 1847-1853

Press S. J. (1972): Applied Multivariate Analysis, Holt, Rinehart and Winston, INC. New York, 1972

Pukelsheim F. (1993): Optimal Design of Experiments, Wiley, New York 1993

Rao C.R.(1975): Simultaneous estimation of parameters in different linear models and applications to biometric problems, Biometrics 31 (1975) 545-554

Rao C.R.(1976): Estimation of parameters in a linear model (The 1975 Wald Memorial Lectures), The Annals of Statistics 4 (1976) 1023- 1037

Rebai S., Philip H. and Taboada A. (1992): Modern Tectonic Stress Field in the Mediterranean Region: Evidence for Variation in Stress Directions at Different Scales, Geophys. J. Int. 110 (1991) 106-140

Reilinger RE, McClusky SC, Oral MB, King RW, Toksoz MN (1997a): Global positioning system measurements of present-day crustal movements in the Arabia–Africa–Eurasia plate collision zone. J Geophys Res 102(1997) 9983–9999

Reilinger RE, McClusky SC, Souter BJ (1997b): Preliminary estimates of plate convergence in the Caucasus collision zone from global positioning system measurements. Geophys Res Lett 24, 1815–1818

Reilly W., Fredrich G., Hein G., Landau H., Almazan J. and Caturla J. (1992): Geodetic Determination of Crustal Deformation Across the Strait of Gibraltar, Geophys. J. Int. 111 (1992) 391-398

Rencher A. (1995) : Methods of multivariate analysis, Wiley, New York, 1995

Rencher A. (1998) : Multivariate statistical inference and applications, Wiley, New York, 1998

Rosen D. (1988): Moments for Matrix Normal variables, Statistics 19 (1988) 575-583

Savage J. C. (1983): A Dislocation Model of Strain Accumulation and Release at a Subduction Zone, Journal of Geophysical Research 88 (1983) 4984-4996

Savage J. C. and Hastie L. M. (1966): Surface Deformation Associated with Dip-Slip Faulting, Journal of Geophysical Research *71* (1966) 4897-4904

Savage J.C. , Gan W. and Svarc J.L. (2001): Strain accumulation and rotation in the Eastern California Shear Zone, JGR 106 (2001) 21995-22007

Schneider D. (1982): Complex Crustal Strain Approximation, Ph.D. Dissertation, Reports of the Department of Surveying Engineering, University of New Brunswick No.91, Canada

Schaffrin B. (1995) A comparision of inverse techniques: regularization, weight estimation and homBLUP, IUGG General Association 1995

Schaffrin B. (2000) Minimum means square error adjustment, part I: The emperirical BLE and the repro-BLE for direct Observation, Journal of the Geodetic Society of Japan 46 (2000) 21-30

Schaffrin B., Heidenreich, E. and Grafarend, E.(1977): A representation of the standard gravity field, Manuscipta geodaetica *2* (1977) 135-174

Schaffrin B. and Middel B.(1990): Robust predictors and an alternative iteration scheme for ill-posed problem, in Geophysical Data Inversion-Methods and Applications, A. Vogel, Ch. O. Ofoegbu, R. Gorenflo and Bj. Ursin (Eds.), Vieweg, Braunschweig (1990) 33-51

Scherneck H.-G, Johansson J. M., Haas R., Bergstrand S., Lidberg M. and Koivula H. (2002): BIFROST Project: From geodetic positions to strain rates, in Proc. of the 14^{th} General meeting of the Nordic Geodetic Comission, Poutanen M. and Suurmäki (Eds), 62-65, Nordic Geodetic Comission

Schuh H. Charlot, P. Hase, H. Himwich E., Kingham K., Klatt C., Ma C., Malkin Z., Niell A., Nothnagel A., Schlutter W., Takashima K., Vandenberg N.(2002): IVS Working Group 2 for Product Specification and observing Programs- Final Report. In IVS Annual Report 2001 ed. By N. R. Vandenberg and K. D. Baver, NASA, 2002

Scholz C. H. (1990): The mechanics of earthquakes and faulting, Cambridge University Press, Cambridge

Searle, S. (1982): Matrix Algebra Useful for Statistics, John Wiley & Sons, New York

Seeber G. (2003): Satellite Geodesy, 2nd Edition, de Gruyter, Berlin New York 2003

Selvaggi, G. (1998). Spatial distribution of horizontal seismic strain in the Apennines from historical earthquakes, *Ann. Geofis.*, 41(1998) 241-251

Shalabh (1998): Improved estimation in measurement error models through Stein rule procedure, Journal of Multivariate Analysis 67 (1998) 35-48

Shen, Z.-K., Jackson, D. D. and Ge, B. X. (1996): Crustal deformation across and beyond the Los Angeles basin from geodetic measurements, Journal of Geophysical Research, 101 (1996) 27,957-27,980

Siotani M., Hayakawa T. and Fujikoshi Y. (1985): Modern multivariate statistical analysis: A graduate course and handbook, American Sciences Press, Inc, Columbus, Ohio

Smith G. and Campbell, F.(1980): A Critique of Some Ridge Regression Methods, Journal of the American Statistical Association 75 (1980) 74-81

Smith DE, Kolenkiewicz R, Dunn PJ, Torrance MH, Klosko SM,Robbins JW, Williamson RG, Pavlis EC, Douglas NB, Fricke SK (1990) The determination of present-day tectonic motions from laser ranging to LAGEOS. In: Brunner FK, Rizos C (eds) Developments in Four-Dimensional Geodesy. Ron S. Mather Symposium on Four-Dimensional Geodesy, Sydney, Australia, March 28–31, 1989. Lecture notes in Earth sciences 29, Springer-Verlag, New York pp 221–240

Sokolnikoff I.S. (1956): Mathematical Theory of Elasticity, McGraw Hill Book Comp., New York

Soler T. and van Gelder B. (1991): On Covariances of Eigenvalues and Eigenvectors of Second Rank Symmetric Tensors, Geophys. J. Int. 105 (1991) 537-546

Spearman, C. (1904): “General-intelligence,” objectively determined and measured, American Journal of Psychology 15 (1904) 201-293

Srivastava, M. S.(2002): Methods of Multivariate Statistics, John Wiley & Sons, New York

Srivastava V.K. and A Chaturvedi(1983): Some properties of the distribution of an operational ridge estimator, Metron 30 (1983) 227-237

Stein S. and Sella G. (2002): Plate Boundary Zones : Concept and Approaches, in Plate Boundary Zones , (Eds.) Stein S. and Freymuller J. T., Geodynamics Series Vol. 30, AGU, Washington D.C 1-26

Straub, C. (1996): Recent crustal deformation and strain accumulation in the Marmara SeaRegion, N.W. Anatolia, inferred from GPS Measurements**,** Zürich : Institut für Geodäsie und Photogrammetrie, 1996

Straub, C. and Kahle, H.-G. (1997): GPS and geologic estimates of the tectonic activity in the Marmara Sea region, NW Anatolia, Journal of the Geophysical Research *102* (1997) 27,587-27,601

Sugiura N. and Nagao H. (1968): Unbiasedness of some test criteria for the quality of one or two covariance matrices, Ann. Math. Statist. 39, 1686-1692

Tarantola A., Coll B., Mosegaard K and P. Xu(2000): The null information probability density for anisotropic tensors, from http://web.ccr.jussieu.fr/tarantola/Files/Professional/Papers_PDF/AnisotropicTensors.pdf

Terada T and Miyabe N (1929): Deformation of earth crust in Kiransai District and its relation to the orographic feature, Bull. Earth Res. Inst. 7 (1929) 223–241

Theobald C.M.(1974): Generalizations of mean square error applied to ridge regression, Journal of the Royal Statistical Society B.36 (1974) 103-106

Thiele T. N (1889): Almindelig Iagttagelselar: Sandsynlighedsregning og mindste Kvadraters Methode (The general theory of Observations: Probability calculus and the Method of least squares), Reitzel, Kobenhavn

Thiele T. N (1903): Theory of Observations, Layton, London, Reprinted in Ann. Math. Statist. 2.(1931) 165-307

Tsuboi, C. (1932) Investigation on the deformation of the Earth’s crust found by precise geodetic means, *Jap. J. Astron. Geophys***.** 10 (1933) 93–248,

Tumura Y. (1965): The Distributions of Latent Roots and Vectors, TRU Math 1 (1965) 1-16

Tykhonov A.N.(1963): The regularization of incorrectly posed problem, Soviet Math. Doklady 4 (1963) 1624-1627

Tykhonov A.N., Bolshakov V.D. Byvshev, V. A. Iinskiy A. S. and Neyman, Yu. M. . (1977): A variational method of regularization in the adjustment of free geodetic nets, Geodesy, Mapping and Photogrammetry 19 (1977) 127-139

Tykhonov A.N. and Arsenin, V.Y.(1977): Solutions of Ill-Posed Problems, Wiley, New York 1977

Vanicek P. and Krakiwsky E. (1986): Geodesy - The Concepts (Second Edition), Elsevier Science publishers, Amsterdam, 1986

Vinod H.D. and Ullah, A.(1981): Recent advances in regression methods, Marcel Dekker, New York 1981

Voosoghi B. (2000): Intrinsic Deformation Analysis of the Earth Surface Based on 3-Dimensional Displacement Fields Derived from Space Geodetic Measurements, Doctoral thesis of the Faculty of Civil Engineering and Surveying Science of the University of Stuttgart, Technical reports of Department of Geodesy and Geoinformatics, Universität Stuttgart, Report Nr. 2000.3

Wang Y. and Xiao, T.(2001): Fast realization algorithms for determining regularization parameters in linear inverse problems, Inverse Problems 17 (2001) 281-291

Ward S. N. (1994): Constraints on the seismotectonics of the central Mediterranean from Very Long Baseline Interfermetry, Geophys. J. Int. 117(1994) 441-452

Wenchenko E. (2000): Estimation of the signal-to-noise in the linear regression model, Stat. Papers 41 327-343

Wilks S.S. (1932): Certain generalizations in the analysis of variance, Biometrika 24, 471-494

Wishart J. (1928): The general product moment distribution in samples from a normal multivariate population, Biometrika, 20A, 35-52

Wishart J. (1947): Tests of significance in the simple regression problem, J. of the Institute of Actuaries Students Society 7 (1947) 99-103

Wittenburg R. (1999): Zur Beschreibung von Deformationsprozessen im 3D-Bereich, In: Quo vadis geodesia ...? Festschrift for Erik W. Grafarend on the occasion of his 60th birthday, Schriftenreihe der Institute des Studiengangs Geodäsie und Geoinformatik; Technical Reports Department of Geodesy and GeoInformatics, (1999) 515-523, later collected in: Geodesy - the Challenge of Third Millennium, (Eds.) Grafarend E., Krumm F. and Schwarz V., Springer-Verlag: Berlin, Heidelberg 2003

Wolf H. (1961): Das Fehlerfortpflanzungsgesetz mit Gliedern II. ZfV 86 (1961) 86-88

Wu P. and Hasegawa H.S. (1996): Induced Stresses and Fault Potential in Eastern Canada due to a Realistic Load: A Preliminary Analysis, Geophys. J. Int. 127 (1996) 215-229

Wyss M., Liang B.Y., Tanagawa W. and Wu X. (1992): Comparison of Orientations of Stress and Strain Tensors Based on Fault Plane Solutions in Kaoiki, Hawaii, J. Geophys. Res. 97 (1992) 4769-4790

Xu P.(1992a): Determination of surface gravity anomalies using gradiometric observables, Geophys. J. Int. 110 (1992) 321-332

Xu P.(1992b): The value of minimum norm estimation of geopotential fields, Geophys.J.Int. 111 (1992) 170-178

Xu P.(1998): Truncated SVD methods for discrete linear ill-posed problems, Geophys.J.Int. 135 (1998) 505-514

Xu P. and Rummel, R(1994a): A simulation study of smoothness methods in recovery of regional gravity fields, Geophys.J.Int. 117 (1994) 472-486

Xu P. and Rummel, R(1994b): Generalized ridge regression with applications in determination of potential fields, manuscripta geodaetica 20 (1994) 8-20

Xu P. (1986): Variance-Covariance Propagation for a Nonlinear Function, J. Wuhan Techn. Uni. Surv. Mapping 11 (1986) 92-99

Xu P. and Grafarend E. (1996a): Probability Distribution of Eigenspectra and Eigendirections of a Twodimensional, Symmetric Rank two Random Tensor, J. of Geodesy 70 (1996) 419-430

Xu P. and Grafarend E. (1996b): Statistics and Geometry of the Eigenspectra of Three-Dimensional Second-Rank Symmetric Random Tensors, Geophysical J. Int. 127 (1996) 744-756

Xu P. and Shimada S. (1997): Inferring the relative measure of principal stress components, Geophysical J. Int. 130 (1997) F1-F4

Xu P. (1999a): Spectral theory of constrained second-rank symmetric random tensors, Geophysical J. Int. 138 (1999) 1-24

Xu P.(1999b): Random stress/strain tensors and beyond, In: Quo vadis geodesia ...? Festschrift for Erik W. Grafarend on the occasion of his 60th birthday, Schriftenreihe der Institute des Studiengangs Geodäsie und Geoinformatik; Technical Reports Department of Geodesy and GeoInformatics, (1999) 525-547, later collected in: Geodesy - the Challenge of Third Millennium, (Eds.) Grafarend E., Krumm F. and Schwarz V., Springer-Verlag: Berlin, Heidelberg 2003