Mathematical Geodesy and Positioning Series **No. 26**

Towards a Least-Squares Framework for Adjusting and Testing of both Functional and Stochastic Models

September 2004
(Reprint 1988)

Faculty of Aerospace Engineering

Delft University of Technology

Towards a Least-Squares Framework for Adjusting and Testing of both Functional and Stochastic Models

P.J.G. Teunissen

1988

Delft University of Technology
Faculty of Civil Engineering and Geosciences
Department of MGP

Contents

Introduction

This research memo should be seen as a first attempt to formulate an unified framework for the adjustment and testing of both the functional and stochastic models. In this memo we concentrate on the problem of estimating parts of the stochastic model. The unification is based on the method of least-squares. Our idea, which is worked out in this memo, was to investigate weather it is possible to use the method of least-squares adjustment also for the problem of variance component estimation. This turns out to be the case. As a consequence, we have the possibility of applying one estimation principle, namely our well-known and well understood method of least-squares, to both the problem of estimating the functional model and stochastic model.

Delft, 1988

The present document is a reprint of the original 1988 MGP-report 'Towards a Least-Squares Framework for Adjusting and Testing of both Functional and Stochastic Models'. Since the theory developed in this report is still considered to be relevant for many modern applications, it was decided to produce a more accessible format of the report. The original format turned out to be poorly reproducible electronically using modern day typesetting system. For the reprint we have chosen to use the popular LaTeXtypesetting system. I am greatful to AliReza Amiri-Simkooei who took the painstaking task upon him to transform the original document into a LaTeXversion. This work is greatly acknowledged. To keep the flavor of the original report in tact (including its flaws), the current document is a complete one-to-one reprint of the original version. The current document is thus the LaTeXreprint of the original report.

Delft, 2004

Chapter 1

The Model: $\underline{y} \sim N(Ax, Q_y)$

1.1 Linear Unbiased Estimators (LUE's)

Consider the linear model of observation equations:

$$E\{\underline{y}\} = Ax,\ Q_y \tag{1.1}$$

where A is assumed to have full rank and the covariance matrix of $\underline{y}$ is assumed to be positive definite. Any linear unbiased estimator of x can then be expressed as

$$\underline{\hat{x}} = (L^*A)^{-1}L^*\underline{y}, \tag{1.2}$$

where the $m \times n$ matrix L is arbitrary provided that $(L^*A)^{-1}$ exists. The property of unbiasedness is easily verified with (1.1) and (1.2):

$$E\{\underline{\hat{x}}\} = (L^*A)^{-1}L^*E\{\underline{y}\} = (L^*A)^{-1}L^*Ax = x \tag{1.3}$$

The covariance matrix of $\underline{\hat{x}}$, $Q_{\hat{x}}$, follows from applying the error propagation law to (1.2):

$$Q_{\hat{x}} = (L^*A)^{-1}L^*Q_yL(A^*L)^{-1} \tag{1.4}$$

The results (1.3) and (1.4) are independent of the distribution of $\underline{y}$. Since the estimator $\underline{\hat{x}}$ of (1.2) is a linear estimator, it follows that if $\underline{y}$ is normally distributed then so is $\underline{\hat{x}}$. In this case, the distribution of $\underline{\hat{x}}$ is completely specified by its first two moments, i.e., x and $Q_{\hat{x}}$.

1.2 Least-Squares Estimators (BLUE's)

Consider again model (1.1). The least squares (LSQ) estimator of x reads then:

$$\underline{\hat{x}} = (A^*Q_y^{-1}A)^{-1}A^*Q_y^{-1}\underline{y}, \tag{1.5}$$

Comparison of (1.2) with (1.5) shows that the least-squares estimator is a linear unbiased estimator. The corresponding choice for L is:

$$L = Q_y^{-1}A \tag{1.6}$$

substitution of (1.6) into (1.4) shows that the covariance matrix of the least-squares estimator reads

$$Q_{\hat{x}} = (A^*Q_y^{-1}A)^{-1} \tag{1.7}$$

It can be shown that of all linear unbiased estimators, the LSQ-estimator has minimum variance. It is therefore a minimum variance linear unbiased estimator, also known in the literatures as BLUE (Best Linear Unbiased Estimator). This property of minimum variance is also independent of the distribution of $\underline{y}$.

Chapter 2

The Model: $\underline{y} \sim N(Ax, \sum_{\alpha=1}^{p} \sigma_\alpha^2 Q_\alpha)$

2.1 Least-Squares Estimation of $\sigma_\alpha^2, \ \alpha = 1, 2, \cdots, p$

Consider the linear model of observation equations:

$$E\{\underbrace{\underline{y}}_{m\times 1}\} = \underbrace{A}_{m\times n}\underbrace{x}_{n\times 1}, \ E\{\underbrace{(\underline{y}-Ax)(\underline{y}-Ax)^*}_{m\times m}\} = \sum_{\alpha=1}^{p} \sigma_\alpha^2 \underbrace{Q_\alpha}_{m\times m} \tag{2.1}$$

where A is assumed to have full rank and the matrices Q_α are assumed to be non-negative definite such that the sum $\sum_{\alpha=1}^{p} \sigma_\alpha^2 Q_\alpha$ is non-negative definite. Note that in this case, we have two sets of unknowns: the parameter vector x <u>and</u> the variance components $\sigma_\alpha^2, \ \alpha = 1, 2, \ldots, p$. The idea of our least-squares approach to variance-component estimation is now to interpret the matrix equation of (2.1), which represents the covariance matrix of $\underline{y}$, as a set of m^2-number of observation equations. Thus, just like we interpret the functional model $E\{\underline{y}\} = Ax$ as a set of m-number of observation equations with the observation vector $\underline{y}$, we are going to interpret the stochastic model $E\{(\underline{y}-Ax)(\underline{y}-Ax)^*\} = \sum_{\alpha=1}^{p} \sigma_\alpha^2 Q_\alpha$ as a set of m^2-number of observation equations with the observation matrix $(\underline{y}-Ax)(\underline{y}-Ax)^*$. There is however one complication: the matrix $(\underline{y}-Ax)(\underline{y}-Ax)^*$ is <u>not</u> observable since the vector x is unknown a-priori. This problem can however be circumvented by transforming model (2.1) into a model of *condition equations*. In terms of condition equations, model (2.1) reads

$$B^* E\{\underline{y}\} = 0, \ E\{B^* \underline{y}\underline{y}^* B\} = \sum_{\alpha=1}^{p} \sigma_\alpha^2 B^* Q_\alpha B \tag{2.2}$$

where matrix B satisfies

$$B^* A = 0, \quad \text{with} \ \ rank(B) = b \tag{2.3}$$

Note that the unknown parameter vector x has now been eliminated from the model. If we define the vector of *misclosures*, $\underline{t}$, as

$$B^* \underline{y} = \underline{t}, \tag{2.4}$$

We can write (2.2) more compactly as

$$E\{\underline{t}\} = 0, \quad E\{\underline{t}\underline{t}^T\} = \sum_{\alpha=1}^{p} \sigma_\alpha^2 B^* Q_\alpha B \tag{2.5}$$

Note that there is no adjustment needed for the first part, i.e., the functional part, of model (2.5). There is no redundancy and there are no unknowns. We may therefore concentrate on the second part, i.e., the stochastic part. The matrix equation of (2.5) can be recast into a set of b^2-number of observation equations by stacking the b-number of $b \times 1$ column vectors of $E\{\underline{t}\underline{t}^T\}$ into a $b^2 \times 1$ observation vector. This results in the *linear model of observation equations*:

$$E\{\underbrace{\begin{pmatrix} \underline{t}_1\underline{t} \\ \underline{t}_2\underline{t} \\ \vdots \\ \underline{t}_b\underline{t} \end{pmatrix}}_{b^2\times 1}\} = \underbrace{\begin{pmatrix} (B^*Q_1B)_{01} & \cdots & (B^*Q_pB)_{01} \\ (B^*Q_1B)_{02} & \cdots & (B^*Q_pB)_{02} \\ \vdots & \ddots & \vdots \\ (B^*Q_1B)_{0b} & \cdots & (B^*Q_pB)_{0b} \end{pmatrix}}_{b^2\times p} \underbrace{\begin{pmatrix} \sigma_1^2 \\ \sigma_2^2 \\ \vdots \\ \sigma_p^2 \end{pmatrix}}_{p\times 1} \tag{2.6}$$

The notation $(B^*Q_\alpha B)_{01},\ (B^*Q_\alpha B)_{02}$, etc indicates the first, the second, etc column vector of the matrix $B^*Q_\alpha B$. If we denote the operator which transforms a matrix into a vector by *vec*, i.e.,

$$vec \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{bmatrix} = \begin{bmatrix} x_{11} & \cdots & x_{n1} & x_{12} & \cdots & x_{n2} & \cdots & x_{1n} & \cdots & x_{nn} \end{bmatrix}^* \tag{2.7}$$

Equation (2.6) can be written more compactly as

$$E\{vec(\underline{t}\underline{t}^*)\} = \begin{bmatrix} vec(B^*Q_1B) & vec(B^*Q_2B) & \cdots & vec(B^*Q_pB) \end{bmatrix} \begin{pmatrix} \sigma_1^2 \\ \sigma_2^2 \\ \vdots \\ \sigma_p^2 \end{pmatrix} \tag{2.8}$$

Having established this results, we can now apply the estimation methods of Section I.1 and I.2. That is, we can now compute linear unbiased estimators of the variance components and also, if the covariance matrix of $vec(\underline{t}\underline{t}^*)$ is known, the least squares estimators (BLUE's) of the variance components. If we denote the covariance matrix of $vec(\underline{t}\underline{t}^*)$ by Q_{vec}, the least-squares estimators of the variance components read:

$$\begin{pmatrix} \hat{\underline{\sigma}}_1^2 \\ \hat{\underline{\sigma}}_2^2 \\ \vdots \\ \hat{\underline{\sigma}}_p^2 \end{pmatrix} = \begin{bmatrix} n_{11} & \cdots & n_{1p} \\ n_{21} & \cdots & n_{2p} \\ \vdots & \ddots & \vdots \\ n_{p1} & \cdots & n_{pp} \end{bmatrix}^{-1} \begin{bmatrix} vec\,(B^*Q_1B)^*Q_{vec}^{-1}vec(\underline{t}\underline{t}^*) \\ vec\,(B^*Q_2B)^*Q_{vec}^{-1}vec(\underline{t}\underline{t}^*) \\ \vdots \\ vec\,(B^*Q_pB)^*Q_{vec}^{-1}vec(\underline{t}\underline{t}^*) \end{bmatrix} \tag{2.9}$$

where

$$n_{kl} = vec(B^*Q_kB)^*Q_{vec}^{-1}vec(B^*Q_lB),\ \ k,l = 1,2,\cdots,p \tag{2.10}$$

The above given least squares approach to variance component estimation has a number of attractive features:

1. Since the approach is based on the least squares principle, we know without any additional derivation that the estimators of (2.9) are *unbiased* and of *minimum variance*. These properties are independent of the distribution of $vec(\underline{t}\underline{t}^*)$. Note by the way that if $\underline{t}$ is normally distributed then $vec(\underline{t}\underline{t}^*)$ is certainly not normally distributed.

2. Since the approach is based on the least squares principle, the inverse of the normal matrix in (2.9) automatically gives us the covariance matrix of the variance components.

3. Since the approach is based on the least squares principle, parts of standard software can be used for computing the variance components.

4. Since the approach is based on the least squares principle, parts of our standard quality control theory (unfortunately only a few parts) can be applied to model (2.8) and the result (2.9).

5. The linear model of observation equations (2.8) makes it in principle rather straightforward to apply estimation methods other than least squares. One could in particular think of robust

estimation methods. This may turn out to be an important alternative if one wants to be guarded against misspecifications in the functional part of model (2.8).

In order to insure non negative variance components, one can also incorporate non-negativity constraints $\sigma_\alpha^2 \geq 0,\ \alpha = 1, 2, \cdots, p$ in the model (2.8).

6. Finally, the least squares approach to variance component estimation is also attractive from a didactic point of view.

2.2 The Covariance Matrix of $vec(\underline{t}\underline{t}^*)$

In order to be able to compute the LSQ-estimators of the variance components in (2.9), we need to know the $b^2 \times b^2$ covariance matrix of $vec(\underline{t}\underline{t}^*)$, Q_{vec}. In fact we need its inverse, Q_{vec}^{-1}. In (2.9) we silently assumed that this inverse exist. It is however not difficult to show that the covariance matrix Q_{vec} is singular! Recall that

$$vec(\underline{t}\underline{t}^*) = \begin{pmatrix} \underline{t}_1\underline{t} \\ \underline{t}_2\underline{t} \\ \vdots \\ \underline{t}_b\underline{t} \end{pmatrix} \tag{2.11}$$

Now define a $b^2 \times 1$ vector as:

$$a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_b \end{bmatrix} \tag{2.12}$$

where $a_i,\ i = 1, 2, \cdots, b$ are vectors of order $b \times 1$. Taking the inner product of (2.11) and (2.12) gives

$$a^* vec(\underline{t}\underline{t}^*) = \sum_{i=1}^{b} \underline{t}_i a_i^* \underline{t} = \begin{bmatrix} \underline{t}_1 & \cdots & \underline{t}_b \end{bmatrix} \begin{bmatrix} a_1^*\underline{t} \\ \vdots \\ a_b^*\underline{t} \end{bmatrix} = \underline{t}^* \begin{bmatrix} a_1^* \\ \vdots \\ a_b^* \end{bmatrix} \underline{t} \tag{2.13}$$

If we define

$$A = \begin{bmatrix} a_1^* \\ \vdots \\ a_b^* \end{bmatrix} \tag{2.14}$$

we have

$$a^* vec(\underline{t}\underline{t}^*) = \underline{t}^* A \underline{t} \tag{2.15}$$

It will be clear that the covariance matrix of $vec(\underline{t}\underline{t}^*)$ is singular, if vector a exist such that $a^* vec(\underline{t}\underline{t}^*)$ is zero. From (2.15) follows that such vectors indeed exist. For instance, if we take the $b \times b$ matrix A to be *skew-symmetric*, i.e., $A^* = -A$, then

$$\underline{t}^* A \underline{t} = (\underline{t}^* A \underline{t})^* = \underline{t}^* A^* \underline{t} = -\underline{t}^* A \underline{t} \tag{2.16}$$

and thus [1]

$$a^* vec(\underline{t}\underline{t}^*) = 0. \tag{2.17}$$

It seems that the singularity of Q_{vec} makes things drastically more complicated. We will return to this matter in the next subsection. Let us however first derive the covariance matrix of $vec(\underline{t}\underline{t}^*)$. The elements of the covariance matrix Q_{vec} are by definition given as

$$Q_{vec}^{i\,j\,k\,l} = E\{(\underline{t}^i\underline{t}^j - E\{\underline{t}^i\underline{t}^j\})(\underline{t}^k\underline{t}^l - E\{\underline{t}^k\underline{t}^l\})\},\ i, j, k, l = 1, 2, \cdots, b \tag{2.18}$$

[1]Note that this property is *independent* of the distribution of $\underline{t}$

If we factor the right hand side we get

$$Q_{vec}^{i\,j\,k\,l} = E\{\underline{t}^i\underline{t}^j\underline{t}^k\underline{t}^l\} - E\{\underline{t}^i\underline{t}^j\}E\{\underline{t}^k\underline{t}^l\},\ i,j,k,l = 1,2,\cdots,b \tag{2.19}$$

This result shows that we need the second and the fourth multivariate central moments of the random vector $\underline{t}$. If we assume that $\underline{t}$ is normally distributed with mean zero and covariance matrix Q_t, the first four multivariate central moments read

$$\begin{array}{rcl} E\{\underline{t}^i\} & = & 0 \\ E\{\underline{t}^i\underline{t}^j\} & = & q^{ij} \\ E\{\underline{t}^i\underline{t}^j\underline{t}^k\} & = & 0 \\ E\{\underline{t}^i\underline{t}^j\underline{t}^k\underline{t}^l\} & = & q^{ij}q^{kl} + q^{ik}q^{jl} + q^{jk}q^{il} \\ i,j,k,l & = & 1,2,\cdots,b \end{array} \tag{2.20}$$

where q^{ij} represents Q_t in index notation. For a proof of (2.20) we refer to Appendix A. With (2.20), equation (2.19) can be written as

$$Q_{vec}^{ijkl} = q^{ik}q^{jl} + q^{jk}q^{il} \tag{2.21}$$

From this results follows that the $b^2 \times b^2$ covariance matrix Q_{vec} is composed of b^2-number $b \times b$ submatrices, i.e., as

$$Q_{vec} = \begin{pmatrix} Q^{1.1.} & Q^{1.2.} & \cdots & Q^{1.b.} \\ Q^{2.1.} & Q^{2.2.} & \cdots & Q^{2.b.} \\ \vdots & \vdots & Q^{i.k.} & \vdots \\ Q^{b.1.} & Q^{b.2.} & \cdots & Q^{b.b.} \end{pmatrix} \tag{2.22}$$

where the $b \times b$ submatrix $Q^{i.k.}$ is of the form

$$Q^{i.k.} = e_i^* Q_t e_k Q_t + Q_t e_k e_i^* Q_t \tag{2.23}$$

with $e_i^* = (0\ \cdots\ 0\ 1\ 0\ \cdots\ 0)$.

2.3 The Singularity of Q_{vec} and Its Consequences

The covariance matrix Q_{vec} is singular if non-zero $b^2 \times 1$ vectors x exist such that

$$Q_{vec}x = 0 \tag{2.24}$$

If we partition x as

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_b \end{pmatrix} \tag{2.25}$$

where $x_k,\ k = 1,2,\cdots,b$ are $b \times 1$ vectors, we have with (2.23) that

$$\sum_{k=1}^{b} Q^{i.k.}x_k = \sum_{k=1}^{b}[e_i^* Q_t e_k Q_t + Q_t e_k e_i^* Q_t]x_k \tag{2.26}$$

This can also be written as

$$\sum_{k=1}^{b} Q^{i.k.}x_k = \begin{pmatrix} Q_t x_1 & Q_t x_2 & \cdots & Q_t x_b \end{pmatrix} \begin{bmatrix} e_i^* Q_t e_1 \\ e_i^* Q_t e_2 \\ \vdots \\ e_i^* Q_t e_b \end{bmatrix} + \sum_{k=1}^{b} Q_t e_k (e_i^* Q_t x_k) \tag{2.27}$$

or as

$$\sum_{k=1}^{b} Q^{i.k.} x_k = Q_t \begin{pmatrix} x_1 & x_2 & \cdots & x_b \end{pmatrix} \begin{bmatrix} e_1^* Q_t e_i \\ e_2^* Q_t e_i \\ \vdots \\ e_b^* Q_t e_i \end{bmatrix} + Q_t \begin{pmatrix} e_1 & e_2 & \cdots & e_b \end{pmatrix} \begin{bmatrix} x_1^* Q_t e_i \\ x_2^* Q_t e_i \\ \vdots \\ x_b^* Q_t e_i \end{bmatrix} \quad (2.28)$$

or as

$$\sum_{k=1}^{b} Q^{i.k.} x_k = Q_t \begin{pmatrix} x_1 & x_2 & \cdots & x_b \end{pmatrix} \begin{bmatrix} e_1^* \\ e_2^* \\ \vdots \\ e_b^* \end{bmatrix} Q_t e_i + Q_t \begin{pmatrix} e_1 & e_2 & \cdots & e_b \end{pmatrix} \begin{bmatrix} x_1^* \\ x_2^* \\ \vdots \\ x_b^* \end{bmatrix} Q_t e_i \quad (2.29)$$

or with

$$X = \begin{pmatrix} x_1 & x_2 & \cdots & x_b \end{pmatrix} \text{ and } I = \begin{pmatrix} e_1 & e_2 & \cdots & e_b \end{pmatrix} \quad (2.30)$$

as

$$\sum_{k=1}^{b} Q^{i.k.} x_k = Q_t X Q_t e_i + Q_t X^* Q_t e_i^* \quad (2.31)$$

or finally as

$$\boxed{\sum_{k=1}^{b} Q^{i.k.} x_k = Q_t (X + X^*) Q_t e_i, \; i = 1, 2, \ldots, b} \quad (2.32)$$

This result shows that the vectors $x = vec(X)$ which satisfy (2.24), are those vectors for which the matrix X is *skew-symmetric*. These vectors therefore span the nullspace of the matrix Q_{vec}. Now that we know the nullspace of the matrix Q_{vec}, we can again start from model (2.8) to derive the least squares estimators. The fact that linear functions of the observations have zero variance, implies in general that the original linear model with singular covariance matrix can be reduced to a linear model with constraints and a non-singular covariance matrix. To see this, consider the linear model

$$E\{\underline{y}\} = Ax, \quad Q_y \quad (2.33)$$

If

$$T_{m\times m} = \begin{pmatrix} T_1 \\ T_2 \end{pmatrix} \text{ with } T_1 T_2^* = 0 \quad (2.34)$$

is a square and regular transformation matrix, then model (2.33) is equivalent to

$$E\{\begin{bmatrix} T_1 \underline{y} \\ T_2 \underline{y} \end{bmatrix}\} = \begin{bmatrix} T_1 A \\ T_2 A \end{bmatrix} x, \quad \begin{bmatrix} T_1 Q_y T_1^* & T_1 Q_y T_2^* \\ T_2 Q_y T_1^* & T_2 Q_y T_2^* \end{bmatrix} \quad (2.35)$$

If we assume that the row vectors of the matrix T_2 span the nullspace of Q_y, i.e., $Q_y T_2^* = 0$, then (2.35) reduces to

$$E\{\begin{bmatrix} T_1 \underline{y} \\ T_2 \underline{y} \end{bmatrix}\} = \begin{bmatrix} T_1 A \\ T_2 A \end{bmatrix} x, \quad \begin{bmatrix} T_1 Q_y T_1^* & 0 \\ 0 & 0 \end{bmatrix} \quad (2.36)$$

And this model is indeed of the form of observation equations with constraints on the unknown parameter vector x. It thus seems that for our variance-component estimation problem we are dealing with a model of the form of (2.36). A closer look at our problem shows however that this is only part of the story! Let us go back to the $b^2 \times 1$ vector x that span the nullspace of the covariance matrix Q_{vec}. We know from (2.32) that these vectors are characterized by

$$\boxed{Q_{vec} vec(X) = 0 \text{ with } X^* = -X} \quad (2.37)$$

These vectors are in the formulation of (2.36) the row vectors of the matrix T_2. In (2.36) we need to compute the matrix T_2A. For our variance-component estimation model (2.8) this means that we need to compute the inner products of $vec(X)$ with $vec(B^*Q_\alpha B),\ \alpha = 1,2,\ldots,p$. Thus $vec(X)^*vec(B^*Q_\alpha B),\ \alpha = 1,2,\ldots,p$. Since

$$vec(X)^*vec(B^*Q_\alpha B) = \sum_{i=1}^{b} x_i^*(B^*Q_\alpha B)_{0i} \tag{2.38}$$

it follows that

$$vec(X)^*vec(B^*Q_\alpha B) = trace(X^*B^*Q_\alpha B) \tag{2.39}$$

Using the following two properties of the *trace* operator,

$$trace(AB) = trace(BA), \quad \text{and} \quad trace(A) = trace(A^*), \tag{2.40}$$

it follows that

$$\begin{aligned} trace(X^*B^*Q_\alpha B) &= trace(B^*Q_\alpha BX^*) = trace[(B^*Q_\alpha BX^*)^*] \\ &= trace(XB^*Q_\alpha B) = -trace(X^*B^*Q_\alpha B). \end{aligned} \tag{2.41}$$

Hence, with (2.39) we find that

$$vec(X)^*vec(B^*Q_\alpha B) = 0, \quad \text{if} \quad X^* = -X \tag{2.42}$$

This is an important results, because it implies in the formulation of (2.36) that $T_2A = 0$. With $T_2A = 0$, model (2.36) reduces to

$$E\{T_1\underline{y}\} = T_1Ax, \quad T_1Q_yT_1^* \tag{2.43}$$

which is considerably simpler to solve than model (2.36). In our variance-component estimation problem we are thus in fact dealing with a model of the form (2.43). The least-squares estimator of x in model (2.43) reads:

$$\underline{\hat{x}} = [A^*T_1^*(T_1Q_yT_1^*)^{-1}T_1A]^{-1}A^*T_1^*(T_1Q_yT_1^*)^{-1}T_1\underline{y} \tag{2.44}$$

In our variance-component estimation problem matrix Q_{vec} takes the place of Q_y of (2.44) and the rows of the matrix T_2 are given by a linear independent set of vectors $vec(X)$ for which $X^* = -X$. Since we assumed that $T_1T_2^* = 0$, the rows of matrix T_1 are given by a linear independent set of vectors $vec(S)$ for which $S = S^*$. This follows from the fact that $vec(S)^*vec(X) = 0$ if $S = S^*$ and $X^* = -X$ (Confer also (2.42). Since the subspace spanned by the vectors $vec(X)$ for which $X^* = -X$ has dimension $b(b-1)/2$ if X is of order $b \times b$, it follows that the dimension of the subspace spanned by the vectors $vec(S)$ for which $S = S^*$ is given by $b(b+1)/2$ if S is of order $b \times b$. Thus, in our variance-component estimation problem the matrix T_1 of (2.44) is of order $b(b+1)/2 \times b^2$. The matrix to be inverted, $T_1Q_yT_1^*$, is therefore of order $b(b+1)/2 \times b(b+1)/2$.

We will now show how, without explicitly inverting the matrix $T_1Q_yT_1^*$, the matrix $A^*T_1^*$ $(T_1Q_yT_1^*)^{-1}T_1A$ and the vector $A^*T_1^*(T_1Q_yT_1^*)^{-1}T_1\underline{y}$ of (2.44) can be computed. Consider the system of linear equations:

$$Q_yu = v \tag{2.45}$$

We will assume that the system is consistent, i.e., that

$$v \in R(Q_y) = \text{range- or column-space of } Q_y \tag{2.46}$$

If we reparameterize u as

$$u = T_1^*\alpha + T_2^*\beta, \tag{2.47}$$

and substitute into (2.45) we get

$$Q_yT_1^*\alpha = v, \tag{2.48}$$

since $Q_y T_2^* = 0$. Premultiplying (2.48) with T_1 and inverting the results gives

$$\alpha = (T_1 Q_y T_1^*)^{-1} T_1 v \tag{2.49}$$

Substitution into (2.47) gives then

$$u = T_1^* (T_1 Q_y T_1^*)^{-1} T_1 v + T_2^* \beta \tag{2.50}$$

This is the *general solution* of the consistent system (2.45). The first part on the right hand side of (2.50) represents a *particular solution* of (2.45) and the second part represents the *homogeneous solution*, i.e., the solution of $Q_y u = 0$. When we premultiply (2.50) with A^*, the homogeneous part disappears since $A^* T_2^* = 0$ and we get

$$A^* u = A^* T_1^* (T_1 Q_y T_1^*)^{-1} T_1 v \tag{2.51}$$

From this result we can conclude that *any* particular solution (2.45) when premultiplied with A^*, equals the righthand side of (2.51). This implies that if we are allowed to take v as one of the column vectors of A, say the i^{th} column vector, then the i^{th} column vector of $A^* T_1^* (T_1 Q_y T_1^*)^{-1} T_1 A$ is obtained from premultiplying an *arbitrary* particular solution of (2.45) with $v = A e_i$ by A^*. Similarly, if we are allowed to take v equal to $\underline{y}$, then $A^* T_1^* (T_1 Q_y T_1^*)^{-1} T_1 \underline{y}$ is obtained from premultiplying an *arbitrary* particular solution of (2.45) with $v = \underline{y}$ by A^*. What remains to be shown is therefore whether $R(A) \subset R(Q_y)$ and $\underline{y} \in R(Q_y)$. We will first proof $R(A) \subset R(Q_y)$. If $v \in R(A)$ then v can be written as $v = A\lambda$ for some λ. Since $T_2 A = 0$ it follows that $T_2 v = 0$. Since T is square and regular, and $T_1 T_2^* = 0$ it follows that $v = T_1^* \delta$ for some δ. In order to continue our proof we first proof that

$$Q_y = T_1^* (T_1 T_1^*)^{-1} T_1 Q_y T_1^* (T_1 T_1^*)^{-1} T_1 \tag{2.52}$$

clearly

$$Q_y = T^* T^{-*} Q_y T^{-1} T \tag{2.53}$$

with

$$T^{-1} = \begin{bmatrix} (T_1 T_1^*)^{-1} T_1 \\ (T_2 T_2^*)^{-1} T_2 \end{bmatrix}^* \tag{2.54}$$

this gives

$$Q_y = T^* \begin{bmatrix} (T_1 T_1^*)^{-1} T_1 Q_y T_1^* (T_1 T_1^*)^{-1} & 0 \\ 0 & 0 \end{bmatrix} T \tag{2.55}$$

Since $Q_y T_2^* = 0$, from (2.55) equation (2.52) follows. We now know that if $v \in R(A)$ then $v = T_1^* \delta$ for some δ. But with (2.52) this implies that $v \in R(Q_y)$. We have therefore shown that indeed $R(A) \subset R(Q_y)$. The proof that $\underline{y} \in R(Q_y)$ goes along the same line. We know from (2.43) that $T_2 \underline{y} = 0 =$ constant. Therefore $\underline{y} = T_1^* \delta$ for some δ. And again with (2.52) this implies that $\underline{y} \in R(Q_y)$.

We are now ready to apply the above to our problem of variance-component estimation. That is, in analogy with (2.45) we consider the consistent system

$$Q_{vec} vec(U) = vec(V) \tag{2.56}$$

where V is chosen as (see (2.8))

$$V = B^* Q_\alpha B, \quad \alpha = 1, 2, \ldots, p \;\text{ and }\; V = \underline{tt}^* \tag{2.57}$$

According to (2.32) we can write $Q_{vec} vec(U_\alpha) = vec(B^* Q_\alpha B)$ as

$$Q_t (U_\alpha + U_\alpha^*) Q_t e_i = B^* Q_\alpha B e_i, \quad i = 1, 2, \ldots, b \tag{2.58}$$

or as

$$Q_t (U_\alpha + U_\alpha^*) Q_t = B^* Q_\alpha B \tag{2.59}$$

or as

$$U_\alpha + U_\alpha^* = Q_t^{-1} B^* Q_\alpha B Q_t^{-1} \tag{2.60}$$

From our previous discussion we know that any particular solution may be taken. One such particular solution is

$$U_\alpha = \frac{1}{2} Q_t^{-1} B^* Q_\alpha B Q_t^{-1} \tag{2.61}$$

The (β, α)-element of the normal matrix of our LSQ-solution of the variance-component estimation problem reads therefore

$$vec(B^* Q_\beta B)^* vec(U_\alpha) = \frac{1}{2} vec(B^* Q_\beta B)^* vec(Q_t^{-1} B^* Q_\alpha B Q_t^{-1}) \tag{2.62}$$

If we denote this element as $n_{\beta\alpha}$ we have

$$\begin{aligned} n_{\beta\alpha} &= \frac{1}{2} vec(B^* Q_\beta B)^* vec(Q_t^{-1} B^* Q_\alpha B Q_t^{-1}) \\ &= \frac{1}{2} trace(B^* Q_\beta B Q_t^{-1} B^* Q_\alpha B Q_t^{-1}) \end{aligned} \tag{2.63}$$

In a similar way as above we can write $Q_{vec} vec(U) = vec(\underline{t}\underline{t}^*)$ with the help of (2.32) as

$$U + U^* = Q_t^{-1} \underline{t}\underline{t}^* Q_t^{-1} \tag{2.64}$$

One particular solution is

$$U = \frac{1}{2} Q_t^{-1} \underline{t}\underline{t}^* Q_t^{-1} \tag{2.65}$$

Therefore

$$vec(B^* Q_\beta B)^* vec(U) = \frac{1}{2} vec(B^* Q_\beta B)^* vec(Q_t^{-1} \underline{t}\underline{t}^* Q_t^{-1}) \tag{2.66}$$

If we denote this element as $\underline{l}_\beta$ we have

$$\begin{aligned} \underline{l}_\beta &= \frac{1}{2} vec(B^* Q_\beta B)^* vec(Q_t^{-1} \underline{t}\underline{t}^* Q_t^{-1}) \\ &= \frac{1}{2} trace(B^* Q_\beta B Q_t^{-1} \underline{t}\underline{t}^* B Q_t^{-1}) \\ &= \frac{1}{2} trace(\underline{t}^* Q_t^{-1} B^* Q_\beta B Q_t^{-1} \underline{t}) \\ &= \frac{1}{2} \underline{t}^* Q_t^{-1} B^* Q_\beta B Q_t^{-1} \underline{t} \end{aligned} \tag{2.67}$$

With (2.63) and (2.68) we are now able to compute the least-squares solution of the linear model (2.8) as:

$$\begin{pmatrix} \underline{\hat{\sigma}}_1^2 \\ \underline{\hat{\sigma}}_2^2 \\ \vdots \\ \underline{\hat{\sigma}}_p^2 \end{pmatrix} = \begin{bmatrix} n_{11} & \cdots & n_{1p} \\ n_{21} & \cdots & n_{2p} \\ \vdots & \ddots & \vdots \\ n_{p1} & \cdots & n_{pp} \end{bmatrix}^{-1} \begin{bmatrix} \frac{1}{2} \underline{t}^* Q_t^{-1} B^* Q_1 B Q_t^{-1} \underline{t} \\ \frac{1}{2} \underline{t}^* Q_t^{-1} B^* Q_2 B Q_t^{-1} \underline{t} \\ \vdots \\ \frac{1}{2} \underline{t}^* Q_t^{-1} B^* Q_p B Q_t^{-1} \underline{t} \end{bmatrix} \tag{2.68}$$

with

$$n_{kl} = \frac{1}{2} trace(B^* Q_k B Q_t^{-1} B^* Q_l B Q_t^{-1}) \tag{2.69}$$

This solution thus replaces (2.9) where it was assumed that Q_{vec} was invertible. Note that while we took care of the singularity of Q_{vec}, we also reduced the order of the matrices which need to be inverted. In (2.9) we had to invert an $b^2 \times b^2$ matrix Q_{vec}, while in (2.68) we have to invert the $b \times b$ matrix Q_t. It should also be noted that, since we assumed $\underline{t}$ to be normally distributed when deriving the covariance matrix Q_{vec}, the BLUE's property of (2.68) is restricted to the class of normal distributions. The LUE's property of course still holds in general. Finally we note that while the inverse of the normal matrix gives the covariance matrix of the variance-components, the normal matrix itself is the covariance matrix of the $p \times 1$ vector on the right hand side of (2.68).

Solution (2.68) can be used directly if the matrix B is available. In practice however one will usually have the design matrix A available, instead of B. We shall therefore have to rewrite (2.68) in terms of A. From (2.63) follows that

$$n_{\beta\alpha} = \frac{1}{2} trace(B^* Q_\beta B Q_t^{-1} B^* Q_\alpha B Q_t^{-1}) = \frac{1}{2} trace(Q_\beta B Q_t^{-1} B^* Q_\alpha B Q_t^{-1} B^*) \tag{2.70}$$

with

$$Q_y B Q_t^{-1} B^* = I - A(A^* Q_y^{-1} A)^{-1} A^* Q_y^{-1} = P_A^\perp \tag{2.71}$$

follows therefore

$$n_{\beta\alpha} = \frac{1}{2} trace(Q_\beta Q_y^{-1} P_A^\perp Q_\alpha Q_y^{-1} P_A^\perp) \tag{2.72}$$

Similarly, it follows with

$$\underline{\hat{e}} = Q_y B Q_t^{-1} B^* \underline{y} = Q_y B Q_t^{-1} \underline{t} = P_A^\perp \underline{y} \tag{2.73}$$

from (2.68) that

$$\underline{l}_\beta = \frac{1}{2} \underline{\hat{e}}^* Q_y^{-1} Q_\beta Q_y^{-1} \underline{\hat{e}} = \frac{1}{2} \underline{y}^* P_A^\perp Q_y^{-1} Q_\beta Q_y^{-1} P_A^\perp \underline{y} \tag{2.74}$$

As we mentioned earlier, (2.72) is the covariance matrix of (2.74). With (2.72) and (2.74), solution (2.68) can also be written as

$$\begin{pmatrix} \underline{\hat{\sigma}}_1^2 \\ \underline{\hat{\sigma}}_2^2 \\ \vdots \\ \underline{\hat{\sigma}}_p^2 \end{pmatrix} = \begin{bmatrix} n_{11} & \cdots & n_{1p} \\ n_{21} & \cdots & n_{2p} \\ \vdots & \ddots & \vdots \\ n_{p1} & \cdots & n_{pp} \end{bmatrix}^{-1} \begin{bmatrix} \frac{1}{2} \underline{\hat{e}}^* Q_y^{-1} Q_1 Q_y^{-1} \underline{\hat{e}} \\ \frac{1}{2} \underline{\hat{e}}^* Q_y^{-1} Q_2 Q_y^{-1} \underline{\hat{e}} \\ \vdots \\ \frac{1}{2} \underline{\hat{e}}^* Q_y^{-1} Q_p Q_y^{-1} \underline{\hat{e}} \end{bmatrix} \tag{2.75}$$

with

$$n_{\beta\alpha} = \frac{1}{2} trace(Q_\beta Q_y^{-1} P_A^\perp Q_\alpha Q_y^{-1} P_A^\perp) \tag{2.76}$$

Let us as a simple application of (2.75), assume that there is only one variance component, i.e., $p = 1$. From (2.75) follows then

$$\underline{\hat{\sigma}}^2 = \frac{\frac{1}{2} \underline{\hat{e}}^* Q_y^{-1} Q_1 Q_y^{-1} \underline{\hat{e}}}{\frac{1}{2} trace(Q_1 Q_y^{-1} P_A^\perp Q_1 Q_y^{-1} P_A^\perp)} \tag{2.77}$$

with

$$E\{\underline{\hat{\sigma}}_1^2\} = \sigma_1^2 \quad \text{and} \quad \sigma^2_{\hat{\sigma}_1^2} = \frac{2}{trace(Q_1 Q_y^{-1} P_A^\perp Q_1 Q_y^{-1} P_A^\perp)} \tag{2.78}$$

with $Q_y = \sigma_1^2 Q_1$, $P_A^\perp P_A^\perp = P_A^\perp$, and $trace(P_A^\perp) = rank(P_A^\perp) = m - n$, the above simplifies to:

$$\boxed{\underline{\hat{\sigma}}_1^2 = \frac{\underline{\hat{e}}^* Q_1^{-1} \underline{\hat{e}}}{m-n}, \quad E\{\underline{\hat{\sigma}}_1^2\} = \sigma_1^2 \quad \text{and} \quad \sigma^2_{\hat{\sigma}_1^2} = \frac{2\sigma_1^4}{m-n}} \tag{2.79}$$

These are the well-known results for the estimator of the variance factor of unit weight. Our least-squares approach implies that the above estimator is optimal in the sense that it is *unbiased* and has *minimum variance*! With our least-squares approach we now also have a unified framework in which the well-known estimator of the variance-factor of unit weight finds its logical place. That is, contrary to most lecture notes, we now do not have to introduce the estimator of the variance factor of unit weight in an ad hoc way!

2.4 Estimation of the Covariance Matrix from Repeated Measurements

In our least-squares approach we so far considered only the estimation of the variance-components σ_α^2 of $Q_y = \sum_{\alpha=1}^p \sigma_\alpha^2 Q_\alpha$. The whole procedure applies however equally well to the estimation of *covariance components*. In fact, the least squares approach can also be used to estimate the *covariance matrix* from repeated measurements.

From our formulae (2.68) and (2.75) we see that we need $Q_y = \sum_{\alpha=1}^p \sigma_\alpha^2 Q_\alpha$ in order to compute the estimators $\underline{\hat{\sigma}}_\alpha^2$. But the components σ_α^2 of $\sum_{\alpha=1}^p \sigma_\alpha^2 Q_\alpha$ are unknown a-priori! One way out of this dilemma is to perform *iterations*. One starts with an initial guess for the σ_α^2. Using these values, one computes with either (2.68) or (2.75) estimates for the σ_α^2, which in the next cycle are considered the improved initial guess for σ_α^2. And so on. The estimators obtained in each cycle are unbiased estimators of the σ_α^2. However, they are not of minimum variance, not even after convergence of the iterations. Convergence is achieved if the initial guess for σ_α^2 equals the computed estimate $\underline{\hat{\sigma}}_\alpha^2$. But since the computed estimate $\underline{\hat{\sigma}}_\alpha^2$ is not necessarily equal to σ_α^2, the property of minimum variance may not necessarily be achieved. Hence, in practice one usually will have to be satisfied with *almost* minimum variance unbiased estimators. It will be clear that the amount in which the computed estimates lack the property of minimum variance, depends on the initial guess and the number of iterations performed. The above discussion presupposes that the variance components σ_α^2 are needed in order to compute the estimators $\underline{\hat{\sigma}}_\alpha^2$. Indeed, formulae (2.68) or (2.75) tell us that we need $Q_y = \sum_{\alpha=1}^p \sigma_\alpha^2 Q_\alpha$ and thus σ_α^2. There are however special cases where the σ_α^2 are *not* needed a-priori! One such case we already met when discussing the estimator for the variance-factor of unit weight. Another important case where this holds true occurs when one wants to estimate the covariance matrix from repeated measurements.

consider the following model:

$$\underbrace{E\{\underline{y}_i\}}_{m\times 1} = \underbrace{A}_{m\times n}\underbrace{x_i}_{n\times 1}, \quad \underbrace{E\{(\underline{y}_i - E\{\underline{y}_i\})(\underline{y}_j - E\{\underline{y}_j\})^*\}}_{m\times m} = \sigma_{ij} I_m, \quad i,j = 1,2,\ldots,r \tag{2.80}$$

Written out in full, this model reads

$$\underbrace{E\{\underline{y}\}}_{mr\times 1} = E\{\begin{bmatrix} \underline{y}_1 \\ \underline{y}_2 \\ \vdots \\ \underline{y}_r \end{bmatrix}\} = \underbrace{\begin{bmatrix} A & & & \\ & A & & \\ & & \ddots & \\ & & & A \end{bmatrix}}_{mr\times nr} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \end{bmatrix}}_{nr\times 1}, \quad Q_y = \underbrace{\begin{bmatrix} \sigma_1^2 I & \sigma_{12} I & \cdots & \sigma_{1r} I \\ \sigma_{12} I & \sigma_2^2 I & \cdots & \sigma_{2r} I \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1r} I & \sigma_{2r} I & \cdots & \sigma_r^2 I \end{bmatrix}}_{mr\times mr} \tag{2.81}$$

The unknowns in this model are the $nr \times 1$-number of elements of the vector x

$$x = \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \end{bmatrix}}_{nr\times 1} \tag{2.82}$$

and the $r(r+1)/2$ number of elements σ_i^2 and σ_{ij} of the symmetric matrix

$$Q = \underbrace{\begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1r} \\ \sigma_{12} & \sigma_2^2 & \cdots & \sigma_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1r} & \sigma_{2r} & \cdots & \sigma_r^2 \end{bmatrix}}_{r\times r} \tag{2.83}$$

Using the Kronecker product $\otimes$, we can write (2.81) with (2.82) and (2.83) as

$$\underbrace{E\{\underline{y}\}}_{mr\times 1} = \underbrace{(I\otimes A)}_{mr\times nr}\underbrace{x}_{nr\times 1}, \quad Q_y = \underbrace{Q\otimes I}_{mr\times mr} \tag{2.84}$$

We shall now apply (2.75) to model (2.84) in order to find unbiased and minimum variance estimators for the elements of the matrix Q of (2.83). With appropriate matrices Q_α, matrix Q can be written as

$$Q = \sum_{\alpha=1}^{r(r+1)/2} \sigma_\alpha^2 Q_\alpha \tag{2.85}$$

where σ_α^2 is respectively $\sigma_1^2,\ \sigma_{12},\ \sigma_{13},\ \ldots,\ \sigma_r^2$. Equation (2.74) reads then for the model (2.84)

$$\underline{l}_\beta = \frac{1}{2}\underline{y}^* P_{I\otimes A}^{\perp *}.Q^{-1}\otimes I.Q_\beta \otimes I.Q^{-1}\otimes I.P_{I\otimes A}^{\perp}\underline{y} \tag{2.86}$$

with

$$\begin{aligned} P_{I\otimes A}^{\perp} &= I\otimes I - P_{I\otimes A} \\ &= I\otimes I - I\otimes A[I\otimes A^*.Q^{-1}\otimes I.I\otimes A]^{-1} I\otimes A^*.Q^{-1}\otimes I \\ &= I\otimes [I - A(A^*A)^{-1}A^*] = I\otimes P_A^{\perp} \end{aligned} \tag{2.87}$$

and

$$\underline{y} = \sum_{i=1}^{r} e_i \otimes \underline{y}_i, \quad \text{with} \quad e_{i[r\times 1]} = (0\cdots 0\,1\,0\cdots 0)^* \tag{2.88}$$

this gives

$$\underline{l}_\beta = \frac{1}{2}\sum_{i=1}^{r} e_i^* \otimes \underline{y}_i^*.I\otimes P_A^{\perp *}.Q^{-1}\otimes I.Q_\beta\otimes I.Q^{-1}\otimes I.I\otimes P_A^{\perp}.\sum_{j=1}^{r} e_j\otimes \underline{y}_j \tag{2.89}$$

or

$$\underline{l}_\beta = \frac{1}{2}\sum_{i=1}^{r} e_i^*\otimes \underline{y}_i^*.Q^{-1}Q_\beta Q^{-1}\otimes P_A^{\perp}.\sum_{j=1}^{r} e_j\otimes \underline{y}_j \tag{2.90}$$

or

$$\underline{l}_\beta = \frac{1}{2}\sum_{i=1}^{r}\sum_{j=1}^{r} e_i^* Q^{-1}Q_\beta Q^{-1} e_j\ \underline{y}_i^* P_A^{\perp}\underline{y}_j. \tag{2.91}$$

Because of the symmetry of the matrices $Q^{-1}Q_\beta Q^{-1}$ this result can also be written as

$$\underline{l}_\beta = \frac{1}{2}\sum_{i=1}^{r}(e_i^* Q^{-1}Q_\beta Q^{-1} e_i\ \underline{y}_i^* P_A^{\perp}\underline{y}_i) + \frac{1}{2}.2.\sum_{i=1}^{r}\sum_{j=i+1}^{r}(e_i^* Q^{-1}Q_\beta Q^{-1} e_j\ \underline{y}_i^* P_A^{\perp}\underline{y}_j) \tag{2.92}$$

Let us now turn our attention to equation (2.72). This equation reads, for our model (2.84):

$$N_{\beta\alpha} = \frac{1}{2} trace(Q_\beta\otimes I.Q^{-1}\otimes I.I\otimes P_A^{\perp}.Q_\alpha\otimes I.Q^{-1}\otimes I.I\otimes P_A^{\perp}) \tag{2.93}$$

or

$$N_{\beta\alpha} = \frac{1}{2} trace(Q_\beta Q^{-1}Q_\alpha Q^{-1}\otimes P_A^{\perp}) \tag{2.94}$$

or

$$N_{\beta\alpha} = \frac{1}{2} trace(Q_\beta Q^{-1} Q_\alpha Q^{-1}) trace(P_A^{\perp}) \tag{2.95}$$

or

$$N_{\beta\alpha} = \frac{1}{2}(m-n) trace(Q_\beta Q^{-1}Q_\alpha Q^{-1}) \tag{2.96}$$

since $trace(P_A^\perp) = rank(P_A^\perp) = m - n$. Since the matrices $Q_\alpha,\ \alpha = 1, 2, \ldots, r(r+1)/2$, of (2.85) are of the form

$$Q_\alpha = \begin{cases} e_i e_j^* & \text{for} \quad \sigma_\alpha^2 := \sigma_i^2 \quad i = j \\ \\ e_i e_j^* + e_j e_i^* & \text{for} \quad \sigma_\alpha^2 := \sigma_{ij} \quad i \neq j \end{cases} \tag{2.97}$$

We may write, with the help of (2.96):

$$\begin{aligned} \sum_{\alpha=1}^{r(r+1)/2} N_{\beta\alpha}\underline{\hat{\sigma}}_\alpha^2 &= \frac{1}{2}(m-n)\sum_{i=1}^{r} trace(Q_\beta Q^{-1} e_i e_i^* Q^{-1})\underline{\hat{\sigma}}_i^2 \\ &+ \frac{1}{2}(m-n)\sum_{i=1}^{r}\sum_{j=i+1}^{r} trace(Q_\beta Q^{-1}(e_i e_j^* + e_j e_i^*)Q^{-1})\underline{\hat{\sigma}}_{ij} \end{aligned} \tag{2.98}$$

This can also be written as:

$$\begin{aligned} \sum_{\alpha=1}^{r(r+1)/2} N_{\beta\alpha}\hat{\sigma}_\alpha^2 &= \frac{1}{2}(m-n)\{\sum_{i=1}^{r}(e_i^* Q^{-1} Q_\beta Q^{-1} e_i\, \hat{\sigma}_i^2) \\ &+ 2\sum_{i=1}^{r}\sum_{j=i+1}^{r}(e_i^* Q^{-1} Q_\beta Q^{-1} e_j\, \hat{\sigma}_{ij})\} \end{aligned} \tag{2.99}$$

Since $\sum_{\alpha=1}^{r(r+1)/2} N_{\beta\alpha}\underline{\hat{\sigma}}_\alpha^2 = \underline{l}_\beta$, it follows from (2.92) and (2.99) that the unbiased and minimum variance estimators of the elements σ_{ij} of the matrix Q of (2.83) are given by:

$$\boxed{\underline{\hat{\sigma}}_{ij} = \frac{\underline{y}_i^* P_A^\perp \underline{y}_j}{m-n}, \quad \text{and} \quad \underline{\hat{\sigma}}_i^2 = \frac{\underline{y}_i^* P_A^\perp \underline{y}_i}{m-n}} \tag{2.100}$$

Note that we need not know Q in order to compute these estimates! If we denote $\underline{\hat{y}}_i = P_A \underline{y}_i$, then (2.100) can be written as

$$\underline{\hat{\sigma}}_{ij} = \frac{1}{m-n}\sum_{k=1}^{m}(\underline{y}_{ki} - \underline{\hat{y}}_{ki})(\underline{y}_{kj} - \underline{\hat{y}}_{kj}). \tag{2.101}$$

From this follows that the covariance matrix Q is estimated as

$$\boxed{\underbrace{\underline{Q}}_{r\times r} = \frac{1}{m-n}\sum_{k=1}^{m}\begin{pmatrix} \underline{y}_{k1} - \underline{\hat{y}}_{k1} \\ \underline{y}_{k2} - \underline{\hat{y}}_{k2} \\ \vdots \\ \underline{y}_{kr} - \underline{\hat{y}}_{kr} \end{pmatrix}\begin{pmatrix} \underline{y}_{k1} - \underline{\hat{y}}_{k1} \\ \underline{y}_{k2} - \underline{\hat{y}}_{k2} \\ \vdots \\ \underline{y}_{kr} - \underline{\hat{y}}_{kr} \end{pmatrix}^*} \tag{2.102}$$

If we define the matrices $\underline{Y} = [\underline{y}_1\, \underline{y}_2 \cdots \underline{y}_i \cdots \underline{y}_r]$ and $\underline{\hat{Y}} = [\underline{\hat{y}}_1\, \underline{\hat{y}}_2 \cdots \underline{\hat{y}}_i \cdots \underline{\hat{y}}_r]$ then (2.102) can alternatively be written as

$$\boxed{\underbrace{\underline{Q}}_{r\times r} = \frac{1}{m-n}\underbrace{(\underline{Y} - \underline{\hat{Y}})^*}_{r\times m}\underbrace{(\underline{Y} - \underline{\hat{Y}})}_{m\times r}} \tag{2.103}$$

In order to exemplify the theory, we consider two examples:

2.4.1 Example 1

We want to estimate the variance σ^2 of a distomat by measuring an unknown distance x an m-number of times. We assume that the observations are normally distributed. Model (2.81) reads then for our case:

$$E\{\underline{y}\} = E\{\begin{bmatrix} \underline{y}_1 \\ \underline{y}_2 \\ \vdots \\ \underline{y}_m \end{bmatrix}\} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} x, \quad Q_y = \sigma^2 I_m \tag{2.104}$$

Thus $r = 1,\ n = 1$ and $A = [1\ 1\ \ldots\ 1]^*$. Hence, $P_A\underline{y}$ is

$$\underline{\hat{y}} = P_A\underline{y} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \frac{1}{m}\sum_{i=1}^{m} \underline{y}_i = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \underline{\overline{y}} \tag{2.105}$$

With (2.102) the result reads then

$$\boxed{\underline{\hat{\sigma}}^2 = \frac{1}{m-1}\sum_{k=1}^{m}(\underline{y}_k - \underline{\overline{y}})^2,\ E\{\underline{\hat{\sigma}}^2\} = \sigma^2,\ \sigma^2_{\hat{\sigma}^2} = \frac{2\sigma^4}{m-1}} \tag{2.106}$$

Note that this result can also be obtained from (2.79), the estimator of the variance factor of unit weight.

2.4.2 Example 2

We want to estimate the 2×2 variance-covariance matrix of a *digitizer* by measuring the coordinates of an unknown point an m-number of times. We assume that the observations are normally distributed. Model (2.81) reads then for our case

$$E\{\underline{y}\} = E\{\begin{bmatrix} \underline{y}_1 \\ \underline{y}_2 \end{bmatrix}\} = E\{\begin{bmatrix} \underline{y}_{11} \\ \vdots \\ \underline{y}_{m1} \\ \underline{y}_{12} \\ \vdots \\ \underline{y}_{m2} \end{bmatrix}\} = \begin{bmatrix} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \\ 0 & 1 \\ \vdots & \vdots \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad Q_y = \begin{bmatrix} \sigma_1^2 I_m & \sigma_{12} I_m \\ \sigma_{12} I_m & \sigma_2^2 I_m \end{bmatrix} \tag{2.107}$$

Thus $r = 2,\ n = 1$ and $A = \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix}^*$. Hence, $P_A\underline{y}_i$ is

$$P_A\underline{y}_i = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \frac{1}{m}\sum_{l=1}^{m} \underline{y}_{li} = \underline{\overline{y}}_i, \quad i = 1, 2 \tag{2.108}$$

With (2.102) the results read then

$$\boxed{\begin{bmatrix} \underline{\hat{\sigma}}_1^2 & \underline{\hat{\sigma}}_{12} \\ \underline{\hat{\sigma}}_{12} & \underline{\hat{\sigma}}_2^2 \end{bmatrix} = \frac{1}{m-1}\sum_{k=1}^{m} \begin{bmatrix} \underline{y}_{k1} - \underline{\overline{y}}_1 \\ \underline{y}_{k2} - \underline{\overline{y}}_2 \end{bmatrix} \begin{bmatrix} \underline{y}_{k1} - \underline{\overline{y}}_1 \\ \underline{y}_{k2} - \underline{\overline{y}}_2 \end{bmatrix}^*} \tag{2.109}$$

The corresponding covariance matrix is given by

$$D\{\begin{bmatrix} \underline{\hat{\sigma}}_1^2 \\ \underline{\hat{\sigma}}_{12} \\ \underline{\hat{\sigma}}_2^2 \end{bmatrix}\} = \frac{2}{m-1}(\sigma_1^2\sigma_2^2 - \sigma_{12}^2)^2 \begin{bmatrix} \sigma_2^4 & -2\sigma_2^2\sigma_{12} & \sigma_{12}^2 \\ & 2(\sigma_1^2\sigma_2^2 + \sigma_{12}^2) & -2\sigma_1^2\sigma_{12} \\ & & \sigma_1^4 \end{bmatrix}^{-1} \tag{2.110}$$

In case $\sigma_1 = \sigma_2 = \sigma$ and $\sigma_{12} = 0$, it follows:

$$\boxed{D\{\begin{bmatrix} \underline{\hat{\sigma}}_1^2 \\ \underline{\hat{\sigma}}_{12} \\ \underline{\hat{\sigma}}_2^2 \end{bmatrix}\} = \frac{2\sigma^4}{m-1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 1 \end{bmatrix}} \tag{2.111}$$

Chapter 3

On the Distribution of Variance Components

3.1 Quadratic Forms in Normal Variables

If we denote the inverse of the normal matrix in (2.68) as $N_{\beta\alpha}^{-1}$, it follows that the least-squares estimator of σ_β^2 is given as

$$\underline{\hat{\sigma}}_\beta^2 = \frac{1}{2}\sum_{\alpha=1}^{p} N_{\beta\alpha}^{-1}\underline{t}^* Q_t^{-1} B^* Q_\alpha B Q_t^{-1}\underline{t} \tag{3.1}$$

or as

$$\underline{\hat{\sigma}}_\beta^2 = \underline{t}^*(Q_t^{-1} B^* \frac{1}{2}\sum_{\alpha=1}^{p} N_{\beta\alpha}^{-1} Q_\alpha B Q_t^{-1})\underline{t} \tag{3.2}$$

Hence, each least-squares estimator of a variance-component can be written as a *quadratic form* in the normal vector $\underline{t}$:

$$\underline{\hat{\sigma}}_\beta^2 = \underline{t}^* A_\beta \underline{t} \tag{3.3}$$

with

$$A_\beta = Q_t^{-1} B^* \frac{1}{2}\sum_{\alpha=1}^{p} N_{\beta\alpha}^{-1} Q_\alpha B Q_t^{-1} \tag{3.4}$$

In the following, we shall *assume* that the symmetric matrix A_β is non-negative definite. In practice, this may not be the case, since, as we know, negative estimates of the variance-component are possible. In order to derive the distribution of $\underline{\hat{\sigma}}_\beta^2$ for non-negative matrices A_β, we need the distribution of $\underline{t}^* A_\beta \underline{t}$. The following theorem gives a general representation of the distribution of $\underline{t}^* A \underline{t}$.

Theorem: Let the $b \times 1$ vector $\underline{t}$ be normally distributed with mean $E\{\underline{t}\} = t$ and positive definite covariance matrix Q_t. Let A be a symmetric non-negative definite matrix of order b. Then there exists a positive-definite diagonal matrix $\Lambda_r = diag(\lambda_1, \lambda_2, \cdots, \lambda_r)$ and a vector $u \in \mathbb{R}^r$ such that

$$\boxed{\underline{t}^* A \underline{t} = (\underline{z} + u)^* \Lambda_r (\underline{z} + u) = \sum_{i=1}^{r} \lambda_i (\underline{z}_i + u_i)^2} \tag{3.5}$$

where $\underline{z}$ has the *standard normal distribution*, i.e. $\underline{z} \sim N(0, I_r)$. The number r is the rank of AQ_t or $Q_t A$. The diagonal elements of Λ_r are the r *positive eigenvalues* of AQ_t or $Q_t A$. And if $U_r \Lambda_r U_r^*$ is the *singular value decomposition* of $Q_t^{1/2} A Q_t^{1/2}$, i.e., $Q_t^{1/2} A Q_t^{1/2} = U_r \Lambda_r U_r^*$, with $Q_t^{1/2}$ a square-root of Q_t, i.e., $Q_t = Q_t^{1/2} Q_t^{1/2}$, then the $r \times 1$ vector u can be computed as

$$u = U_r^* Q_t^{-1/2} t. \tag{3.6}$$

Proof: If we define the random vector $\underline{x} = Q_t^{-1/2}(\underline{t} - t)$, then clearly $\underline{x}$ has a standard normal distribution, i.e., $\underline{x} \sim N(0, I_b)$. Substitution of $\underline{t} = t + Q_t^{1/2}\underline{x}$ in $\underline{t}^* A \underline{t}$ gives

$$\underline{t}^* A \underline{t} = t^* A t + 2t^* A Q_t^{1/2} \underline{x} + \underline{x}^* Q_t^{1/2} A Q_t^{1/2} \underline{x} \tag{3.7}$$

Since the matrix $Q_t^{1/2} A Q_t^{1/2}$ is *symmetric* and *non-negative definite* it has real-valued non-negative eigenvalues and corresponding orthonormal eigenvectors. If we collect the b-number of eigenvalues in the $b \times b$ diagonal matrix Λ and the corresponding orthonormal eigenvectors as columns in the $b \times b$ matrix U then

$$Q_t^{1/2} A Q_t^{1/2} = U \Lambda U^* \tag{3.8}$$

with

$$U^* U = U U^* = I_b \tag{3.9}$$

If $rank(Q_t^{1/2} A Q_t^{1/2}) = r$, then r-number of eigenvalues are positive and $(b-r)$-number of eigenvalues are zero. We may therefore partition (3.8) as

$$\begin{aligned} Q_t^{1/2} A Q_t^{1/2} &= [U_r \, U_{b-r}] \begin{bmatrix} \Lambda_r & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} U_r^* \\ U_{b-r}^* \end{bmatrix} \\ &= U_r \Lambda_r U_r^* \end{aligned} \tag{3.10}$$

with

$$U_r^* U_r = I_r \tag{3.11}$$

Substitution of (3.10) into (3.7) gives

$$\underline{t}^* A \underline{t} = t^* A t + 2t^* Q_t^{-1/2} U_r \Lambda_r U_r^* \underline{x} + \underline{x}^* U_r \Lambda_r U_r^* \underline{x} \tag{3.12}$$

or

$$\begin{aligned} \underline{t}^* A \underline{t} &= t^* A t - t^* Q_t^{-1/2} U_r \Lambda_r U_r^* Q_t^{-1/2} t \\ &+ (U_r^* \underline{x} + U_r^* Q_t^{-1/2} t)^* \Lambda_r (U_r^* \underline{x} + U_r^* Q_t^{-1/2} t) \end{aligned} \tag{3.13}$$

or

$$\begin{aligned} \underline{t}^* A \underline{t} &= t^* A t - t^* Q_t^{-1/2} Q_t^{1/2} A Q_t^{1/2} Q_t^{-1/2} t \\ &+ (U_r^* \underline{x} + U_r^* Q_t^{-1/2} t)^* \Lambda_r (U_r^* \underline{x} + U_r^* Q_t^{-1/2} t) \end{aligned} \tag{3.14}$$

or

$$\underline{t}^* A \underline{t} = (\underline{z} + u)^* \Lambda_r (\underline{z} + u) \tag{3.15}$$

with

$$\begin{aligned} \underline{z} &= U_r^* \underline{x} = U_r^* Q_t^{-1/2} (\underline{t} - t) \\ u &= U_r^* Q_t^{-1/2} t \end{aligned} \tag{3.16}$$

Since $\underline{x}$ is distributed as $\underline{x} \sim N(0, I_b)$, and $U_r^* U_r = I_r$, it follows that $\underline{z} = U_r^* \underline{x}$ is distributed as $\underline{z} \sim N(0, I_r)$. Note that since $|Q_t^{1/2} A Q_t^{1/2} - \lambda I_b| = |A Q_t - \lambda I_b| = |Q_t A - \lambda I_b|$, the eigenvalues of $Q_t^{1/2} A Q_t^{1/2}$, $A Q_t$ and $Q_t A$ are the same. *End of proof.*

The above theorem says that $\underline{t}^* A \underline{t}$ is distributed as a linear combination of r independent non-central χ^2-distribution with 1 degree of freedom and non-centrality parameters u_i^2, $i = 1, 2, \cdots, r$, i.e.,

$$\boxed{\underline{t}^* A \underline{t} \sim \sum_{i=1}^{r} \lambda_i \chi^2 (1, u_i^2)} \tag{3.17}$$

From this follows that if all the positive eigenvalues of AQ_t *equal* 1, then $\underline{t}^*A\underline{t}$ is distributed as a non-central χ^2-distribution with r degrees of freedom and non-centrality parameter $u^*u = t^*Q_t^{-1/2}U_rI_rU_r^*Q_t^{-1/2}t = t^*At$:

$$\underline{t}^*A\underline{t} \sim \chi^2(r, t^*At) \quad \text{if} \quad \Lambda_r = I_r. \tag{3.18}$$

Since the mean of a central χ^2-distribution with 1 degree of freedom is 1, the mean of $\underline{t}^*A\underline{t}$ follows from (3.5) as

$$\boxed{E\{\underline{t}^*A\underline{t}\} = \sum_{i=1}^{r}\lambda_i + \sum_{i=1}^{r}\lambda_i u_i^2 = trace(AQ_t) + t^*At} \tag{3.19}$$

Since the variance of the non-central χ^2-distribution with 1 degree of freedom and non-centrality parameter u_i^2 is $2(1+2u_i^2)$, the variance of $\underline{t}^*A\underline{t}$ follows from (3.5) as

$$\boxed{\sigma^2_{\underline{t}^*A\underline{t}} = 2\sum_{i=1}^{r}\lambda_i^2 + 4\sum_{i=1}^{r}\lambda_i^2 u_i^2 = 2trace(AQ_tAQ_t) + 4t^*AQ_tA\,t} \tag{3.20}$$

3.2 The Distribution of $\underline{\hat{\sigma}}^2_\beta$

We shall assume that the estimate $\underline{\hat{\sigma}}^2_\beta$ of σ^2_β are non-negative. With the theorem of the previous section and (3.3) and (3.4) we then have the following result:

Corollary: The variance-component estimator $\underline{\hat{\sigma}}^2_\beta$ is distributed as

$$\underline{\hat{\sigma}}^2_\beta \sim \sum_{i=1}^{r}\lambda_i\chi_i^2(1,0) \tag{3.21}$$

where the χ_i^2 are mutually independent and the λ_i are the r positive eigenvalues of

$$\left|B^*[\sum_{\alpha=1}^{p}(\frac{1}{2}N_{\beta\alpha}^{-1} - \lambda\sigma_\alpha^2)Q_\alpha]B\right| = 0 \tag{3.22}$$

Note that since the matrix B^* is of the order $b\times m$, the number of positive eigenvalues, r, never exceed b.
The result (3.22) is expressed in terms of the matrix B which however is often not explicitly available. We shall therefore reexpress (3.22) in terms of $Q_y = \sum_{\alpha=1}^{p}\sigma_\alpha^2 Q_\alpha$ and $Q_{\hat{e}}$. In order to do this, we need the following two properties of the determinant of a matrix:

1. Let X and Y be two arbitrary matrices of order $n\times n$. Then

$$|XY| = |X|.|Y| \tag{3.23}$$

2. Let X and Y^* be any two matrices of order $m\times n$ and suppose $m \geq n$. Then

$$|XY - \lambda I_m| = (-\lambda)^{m-n}|YX - \lambda I_n| \tag{3.24}$$

The determinant of (3.22) can be written as

$$\begin{aligned}\left|B^*[\sum_{\alpha=1}^{p}(\frac{1}{2}N_{\beta\alpha}^{-1} - \lambda\sigma_\alpha^2)Q_\alpha]B\right| &= \left|B^*[\sum_{\alpha=1}^{p}\frac{1}{2}N_{\beta\alpha}^{-1}Q_\alpha]B - \lambda B^*\sum_{\alpha=1}^{p}\sigma_\alpha^2 Q_\alpha B\right| \\ &= \left|B^*[\sum_{\alpha=1}^{p}\frac{1}{2}N_{\beta\alpha}^{-1}Q_\alpha]B - \lambda B^* Q_y B\right|\end{aligned}$$

with (3.23) and (3.24) one gets

$$\begin{aligned}\left|B^*[\sum_{\alpha=1}^{p}(\frac{1}{2}N_{\beta\alpha}^{-1}-\lambda\sigma_\alpha^2)Q_\alpha]B\right| &= \left|B^*[\sum_{\alpha=1}^{p}\frac{1}{2}N_{\beta\alpha}^{-1}Q_\alpha]B(B^*Q_yB)^{-1}-\lambda I_b\right|.\,|B^*Q_yB| \\ &= (-\lambda)^{b-m}\left|\sum_{\alpha=1}^{p}\frac{1}{2}N_{\beta\alpha}^{-1}Q_\alpha B(B^*Q_yB)^{-1}B^*-\lambda I_m\right|.\,|B^*Q_yB|\end{aligned}$$

if $\lambda \neq 0$. From this follows that for non-zero eigenvalues, (3.22) is equivalent to

$$\left|\sum_{\alpha=1}^{p}\frac{1}{2}N_{\beta\alpha}^{-1}Q_\alpha B(B^*Q_yB)^{-1}B^*-\lambda I_m\right|=0 \tag{3.25}$$

With $Q_{\hat{e}} = Q_yB(B^*Q_yB)^{-1}B^*Q_y$ this gives

$$\left|\sum_{\alpha=1}^{p}\frac{1}{2}N_{\beta\alpha}^{-1}Q_\alpha Q_y^{-1}Q_{\hat{e}}Q_y^{-1}-\lambda I_m\right|=0 \tag{3.26}$$

or with (3.24)

$$\left|Q_y^{-1}(\sum_{\alpha=1}^{p}\frac{1}{2}N_{\beta\alpha}^{-1}Q_\alpha)Q_y^{-1}Q_{\hat{e}}-\lambda I_m\right|=0 \tag{3.27}$$

The result (3.22) can therefore be rephrased as:

Final Result: The variance-component estimator $\underline{\hat{\sigma}}_\beta^2$ is distributed as

$$\underline{\hat{\sigma}}_\beta^2 \sim \sum_{i=1}^{r}\lambda_i\chi_i^2(1,0) \tag{3.28}$$

where the χ_i^2 are mutually independent and the λ_i are the r positive eigenvalues of

$$\left|Q_y^{-1}(\sum_{\alpha=1}^{p}\frac{1}{2}N_{\beta\alpha}^{-1}Q_\alpha)Q_y^{-1}Q_{\hat{e}}-\lambda I_m\right|=0 \tag{3.29}$$

To see this result at work let us derive the distribution of the variance-factor of unit weight. In this case, we have $p=1,\ Q_y=\sigma_1^2Q_1$ and $N_{11}=\frac{1}{2}(m-n)\sigma_1^{-4}$. The above eigenvalue problem becomes then

$$\left|Q_y^{-1}\frac{1}{2}[\frac{1}{2}(m-n)\sigma_1^{-4}]^{-1}Q_1\sigma_1^{-2}Q_1^{-1}Q_{\hat{e}}-\lambda I_m\right|=0 \tag{3.30}$$

or

$$\left|[\sigma_1^{-2}(m-n)]^{-1}Q_{\hat{e}}Q_y^{-1}-\lambda I_m\right|=0 \tag{3.31}$$

or with $Q_{\hat{e}}Q_y^{-1}=P_A^\perp$

$$\left|P_A^\perp-\lambda\sigma_1^{-2}(m-n)I_m\right|=0 \tag{3.32}$$

Since the eigenvalues of a projector are 1 or 0, it follows since $rank(P_A^\perp)=m-n$ that the positive eigenvalues are

$$\lambda_1=\lambda_2=\cdots=\lambda_{m-n}=[\sigma_1^{-2}(m-n)]^{-1} \tag{3.33}$$

From (3.29) follows then that the variance-factor of unit weight $\underline{\hat{\sigma}}_1^2$ is distributed as

$$\boxed{\underline{\hat{\sigma}}_1^2 \sim \frac{\sigma_1^2\chi^2(m-n,0)}{m-n}} \tag{3.34}$$

This is a well-known result and very simple indeed. For general case of more than one variance-component, the eigenvalues λ_i of (3.29) will usually differ mutually and consequently the distribution of $\underline{\hat{\sigma}}^2_\beta$ will be a very complicated one. As far as I know no practical closed form expression for the *cumulative distribution* function of $\underline{\hat{\sigma}}^2_\beta$ is available. This function is needed to perform hypothesis testing, to compute critical values and to compute the power (reliability). Fortunately, however, asymptotic expansions which can be used for computer calculation are available[1]. Also suitable (and may be practical useful) approximations are available. Once the distribution of $\underline{\hat{\sigma}}^2_\beta$ is available one can think of testing hypotheses. One possible approach would be the following: Assume the null hypothesis as

$$H_0: \ E\{\underline{t}\} = 0, \ \ Q_t = B^* \sum_{\alpha=1}^{p} \sigma^2_\alpha Q_\alpha B, \ \begin{cases} \sigma^2_\alpha \neq 1 & \text{for} \quad \alpha = 1, 2, \cdots, p, \ \alpha \neq i \\ \sigma^2_\alpha = 1 & \text{for} \quad \alpha = i \end{cases} \tag{3.35}$$

Assume the alternative hypothesis as

$$H_A: \ E\{\underline{t}\} = 0, \ \ Q_t = B^* \sum_{\alpha=1}^{p} \sigma^2_\alpha Q_\alpha B, \ \sigma^2_\alpha \neq 1 \ \text{ for } \ \alpha = 1, 2, \cdots, p \tag{3.36}$$

Compute the estimator of σ^2_i under H_A. This estimator depends however on the unknown $\sigma^2_\alpha, \ \alpha = 1, 2, \cdots, p$. Approximate the estimator $\underline{\hat{\sigma}}^2_i$ of σ^2_i under H_A therefore by assuming that $\sigma^2_\alpha = 1, \ \alpha = 1, 2, \cdots, p$, and call this *approximate estimator* $\underline{\hat{\sigma}}'^2_i$. As we know, this approximate estimator is still unbiased. Then derive the distribution of $\underline{\hat{\sigma}}'^2_i$. This distribution depends however under H_0 still on the unknown $\sigma^2_\alpha, \ \alpha = 1, 2, \cdots, p, \ \alpha \neq i$. One can approximate this distribution by replacing the unknown σ^2_α by the estimates $\hat{\sigma}'^2_\alpha, \ \alpha = 1, 2, \cdots, p, \ \alpha \neq i$. After this one can perform the *significance test*: $\sigma^2_i = 1$ or $\sigma^2_i \neq 1$.
Another approach would be the following: Assume the null hypothesis as

$$H_0: \ E\{\underline{t}\} = 0, \ Q_t = B^* \sum_{\alpha=1}^{p} Q_\alpha B \tag{3.37}$$

and the the alternative hypothesis as

$$H_{A_i}: \ E\{\underline{t}\} = 0, \ Q_t = B^* (\sum_{\alpha=1 \, \alpha \neq i}^{p} Q_\alpha + \sigma^2_i Q_i) B, \ \sigma^2_i \neq 1 \tag{3.38}$$

This approach parallels the *data snooping* approach and it has some distinct advantages over the above first approach. First of all, the null hypothesis is completely specified, it is a so-called *simple hypothesis*. Secondly, there is only *one unknown*, namely σ^2_i, in the alternative hypothesis. This is advantageous from a computational point of view. In the following section we will consider the case of least-squares estimation under the above H_{A_i}.

3.3 LSQ Estimation in Case $Q_t = B^*(\sum_{\alpha=1 \, \alpha \neq i}^{p} Q_\alpha + \sigma^2_i Q_i) B$

If the covariance matrix of $\underline{t}$ is assumed to take the form

$$E\{\underline{t}\underline{t}^*\} = \sum_{\alpha=1 \, \alpha \neq i}^{p} B^* Q_\alpha B + \sigma^2_i B^* Q_i B \tag{3.39}$$

the observation equations of the linear model take the form

$$E\{vec(\underline{t}\underline{t}^*)\} - vec(\sum_{\alpha=1 \, \alpha \neq i}^{p} B^* Q_\alpha B) = vec(B^* Q_i B) \sigma^2_i \tag{3.40}$$

Thus instead of (2.8), we now have (3.40). Note that since a *constant* vector is subtracted from $vec(\underline{t}\underline{t}^*)$, the covariance matrix of $vec(\underline{t}\underline{t}^*)$ can still be used. With (2.63) we get for the above model

$$N = \frac{1}{2} trace(B^* Q_i B Q_t^{-1} B^* Q_i B Q_t^{-1}) \tag{3.41}$$

[1]N. Johnson & S. Kotz: Continuous Univariate Distributions, Vol 2, 1970

And with (2.68) we get for the above model

$$\underline{l} = \frac{1}{2} trace(B^* Q_i B Q_t^{-1} [\underline{t}\underline{t}^* - \sum_{\alpha=1\,\alpha\neq i}^{p} B^* Q_\alpha B] Q_t^{-1}) \tag{3.42}$$

or

$$\underline{l} = \frac{1}{2} \underline{t} Q_t^{-1} B^* Q_i B Q_t^{-1} \underline{t} - \frac{1}{2} trace(B^* Q_i B Q_t^{-1} \sum_{\alpha=1\,\alpha\neq i}^{p} B^* Q_\alpha B Q_t^{-1}) \tag{3.43}$$

or with $Q_t = \sum_{\alpha=1\,\alpha\neq i}^{p} B^* Q_\alpha B + \sigma_i^2 B^* Q_i B$:

$$\underline{l} = \frac{1}{2} \underline{t} Q_t^{-1} B^* Q_i B Q_t^{-1} \underline{t} - \frac{1}{2} trace(B^* Q_i B Q_t^{-1}) + \frac{1}{2} \sigma_i^2\, trace(B^* Q_i B Q_t^{-1} B^* Q_i B Q_t^{-1}). \tag{3.44}$$

With (3.41) the estimator $\underline{\hat{\sigma}}_i^2 = N^{-1} \underline{l}$ reads therefore:

Final Result:

$$\boxed{\underline{\hat{\sigma}}_i^2 = \sigma_i^2 + \frac{\underline{t} Q_t^{-1} B^* Q_i B Q_t^{-1} \underline{t} - trace(B^* Q_i B Q_t^{-1})}{trace(B^* Q_i B Q_t^{-1} B^* Q_i B Q_t^{-1})}} \tag{3.45}$$

with $E\{\underline{\hat{\sigma}}_i^2\} = \sigma_i^2$ and

$$\sigma_{\hat{\sigma}_i^2}^2 = 2[trace(B^* Q_i B Q_t^{-1} B^* Q_i B Q_t^{-1})]^{-1} \tag{3.46}$$

With (3.5) of the theorem of section one, the distribution of $\underline{\hat{\sigma}}_i^2$ follows as:

$$\boxed{\underline{\hat{\sigma}}_i^2 \sim \sigma_i^2 + \frac{\sum_{j=1}^{r} \lambda_j \chi_j^2(1,0) - \sum_{j=1}^{r} \lambda_j}{\sum_{j=1}^{r} \lambda_j^2}} \tag{3.47}$$

where $\lambda_j,\ j = 1, 2, \cdots, r$ are the r positive eigenvalues of

$$|B^* Q_i B - \lambda Q_t| = 0 \tag{3.48}$$

or of

$$|Q_y^{-1} Q_i Q_y^{-1} Q_{\hat{e}} - \lambda I_m| = 0 \tag{3.49}$$

The problem of hypothesis testing may now be tackled as follows: Assume the null hypothesis as

$$H_0 :\ E\{\underline{t}\} = 0,\ Q_t = \sum_{\alpha=1}^{p} B^* Q_\alpha B \tag{3.50}$$

and the the alternative hypothesis as

$$H_{A_i} :\ E\{\underline{t}\} = 0,\ Q_t = \sum_{\alpha=1\,\alpha\neq i}^{p} B^* Q_\alpha B + \sigma_i^2 B^* Q_i B \tag{3.51}$$

Note that although the estimator $\underline{\hat{\sigma}}_i^2$ of (3.45) can not be computed in practice because of the unknown σ_i^2, its distribution *is known under* H_0! Instead of computing $\underline{\hat{\sigma}}_i^2$ we therefore approximate this estimator by an estimator $\underline{\hat{\sigma}}_i'^2$, which is obtained by setting $\sigma_i^2 = 1$ in (3.45). The approximate estimator reads therefore

$$\boxed{\underline{\hat{\sigma}}_i'^2 = 1 + \frac{\underline{t} \bar{Q}_t^{-1} B^* Q_i B \bar{Q}_t^{-1} \underline{t} - trace(B^* Q_i B \bar{Q}_t^{-1})}{trace(B^* Q_i B \bar{Q}_t^{-1} B^* Q_i B \bar{Q}_t^{-1})}} \tag{3.52}$$

with

$$\bar{Q}_t = \sum_{\alpha=1}^{p} B^* Q_\alpha B \tag{3.53}$$

We know from our theory that this approximate estimator is still an *unbiased estimator* of σ_i^2 (however not of minimum variance anymore). Let us verify this for (3.52). With $A = \bar{Q}_t^{-1} B^* Q_i B \bar{Q}_t^{-1}$ and $E\{\underline{t}\} = 0$, we have with (3.19):

$$E\{\underline{t}^* \bar{Q}_t^{-1} B^* Q_i B \bar{Q}_t^{-1} \underline{t}\} = trace(\bar{Q}_t^{-1} B^* Q_i B \bar{Q}_t^{-1} Q_t) \tag{3.54}$$

or with

$$Q_t = \bar{Q}_t + (\sigma_i^2 - 1) B^* Q_i B \tag{3.55}$$

$$\begin{aligned} E\{\underline{t}^* \bar{Q}_t^{-1} B^* Q_i B \bar{Q}_t^{-1} \underline{t}\} &= trace(\bar{Q}_t^{-1} B^* Q_i B) \\ &+ (\sigma_i^2 - 1) trace(\bar{Q}_t^{-1} B^* Q_i B \bar{Q}_t^{-1} B^* Q_i B) \end{aligned} \tag{3.56}$$

Substituting into (3.52) shows indeed that $E\{\underline{\hat{\sigma}}_i'^2\} = \sigma_i^2$. Since the distribution of $\underline{\hat{\sigma}}_i'^2$ is known under H_0 we can now perform the test: $\sigma_i^2 = 1$ versus $\sigma_i^2 \neq 1$. Note by the way that $\underline{\hat{\sigma}}_i^2$ and $\underline{\hat{\sigma}}_i'^2$ have *identical distributions* under H_0. By letting i range from 1 to p, we can like in *data snooping* test whether additional variance-components are needed. They can also be done in an *iterated way* like in the *iterated data snooping* approach. In this context it is also interesting to investigate the form of the *shifting variate* of the linear models (2.8) and (3.40). We will return to this matter later on.

3.4 On the Connection of Two Point Fields

As an interesting application of the theory we have the problem of estimating and testing of the levels of precision of two pointfields which are to be connected. Let the coordinates of the two pointfields be collected in the vectors $\underline{x}_i,\ i = 1, 2$, of order $n \times 1$. We assume the $\underline{x}_i$ to be normally distributed with covariance matrices $\sigma_i^2 Q_i,\ i = 1, 2$. We also assume that $\underline{x}_1$ is independent of $\underline{x}_2$. The model reads then

$$E\{\begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \end{bmatrix}\} = \begin{bmatrix} I_n \\ I_n \end{bmatrix} x, \quad \begin{bmatrix} \sigma_1^2 Q_1 & 0 \\ 0 & \sigma_2^2 Q_2 \end{bmatrix} \tag{3.57}$$

From this follows that matrix B^* takes the form

$$B^* = [I_n \ \ -I_n] \tag{3.58}$$

and matrix Q_t takes the form

$$Q_t = \sigma_1^2 Q_1 + \sigma_2^2 Q_2 \tag{3.59}$$

Since we have two unknowns σ_1^2 and σ_2^2, the normal matrix $N_{\beta\alpha}$ is of order 2×2. With (2.63) we get for our case:

$$\begin{aligned} N_{11} &= \frac{1}{2} trace(Q_1 [\sigma_1^2 Q_1 + \sigma_2^2 Q_2]^{-1} Q_1 [\sigma_1^2 Q_1 + \sigma_2^2 Q_2]^{-1}) \\ N_{12} &= \frac{1}{2} trace(Q_1 [\sigma_1^2 Q_1 + \sigma_2^2 Q_2]^{-1} Q_2 [\sigma_1^2 Q_1 + \sigma_2^2 Q_2]^{-1}) \\ N_{22} &= \frac{1}{2} trace(Q_2 [\sigma_1^2 Q_1 + \sigma_2^2 Q_2]^{-1} Q_2 [\sigma_1^2 Q_1 + \sigma_2^2 Q_2]^{-1}) \end{aligned} \tag{3.60}$$

To simplify things, let us assume that $Q_1 = Q_2$. The result (3.60) simplifies then to:

$$\begin{aligned} N_{11} &= \frac{1}{2} n (\sigma_1^2 + \sigma_2^2)^{-2} \\ N_{12} &= \frac{1}{2} n (\sigma_1^2 + \sigma_2^2)^{-2} \\ N_{22} &= \frac{1}{2} n (\sigma_1^2 + \sigma_2^2)^{-2} \end{aligned} \tag{3.61}$$

Hence, the normal matrix becomes singular! Thus if $Q_1 = Q_2$ the two components σ_1^2 and σ_2^2 of model (3.57) are *not separately estimable* (this also makes sense). Later on we will consider the *estimability problem* for the general model (2.8). For the moment, let us change model (3.57) to overcome the estimability problem. Instead of (3.57) we take

$$E\{\begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \end{bmatrix}\} = \begin{bmatrix} I_n \\ I_n \end{bmatrix} x, \quad \begin{bmatrix} Q & 0 \\ 0 & \sigma^2 Q \end{bmatrix} \tag{3.62}$$

Instead of (3.59) we then get

$$Q_t = (1 + \sigma^2)Q \tag{3.63}$$

We now have one unknown, σ^2, and are in the situation as described in Section 3. We therefore can apply formula (3.45). For our case we have:

$$\begin{aligned} \underline{t}^* Q_t^{-1} B^* Q_i B Q_t^{-1} \underline{t} &= (1+\sigma^2)^{-2} (\underline{x}_1 - \underline{x}_2)^* Q^{-1} (\underline{x}_1 - \underline{x}_2) \\ trace(B^* Q_i B Q_t^{-1}) &= (1+\sigma^2)^{-1} n \\ trace(B^* Q_i B Q_t^{-1} B^* Q_i B Q_t^{-1}) &= (1+\sigma^2)^{-2} n \end{aligned} \tag{3.64}$$

Substituting (3.64) into (3.45) gives

$$\underline{\hat{\sigma}}^2 = \sigma^2 + \frac{(1+\sigma^2)^{-2}(\underline{x}_1 - \underline{x}_2)^* Q^{-1} (\underline{x}_1 - \underline{x}_2) - (1+\sigma^2)^{-1} n}{(1+\sigma^2)^{-2} n} \tag{3.65}$$

or

$$\underline{\hat{\sigma}}^2 = \frac{(\underline{x}_1 - \underline{x}_2)^* Q^{-1} (\underline{x}_1 - \underline{x}_2)}{n} - 1 \tag{3.66}$$

Application of (3.47) shows that $r = n$ and

$$\lambda_1 = \lambda_2 = \cdots = \lambda_n = (1+\sigma^2)^{-1} \tag{3.67}$$

Hence,

$$\underline{\hat{\sigma}}^2 \sim \sigma^2 + \frac{(1+\sigma^2)^{-1}\chi^2(n,0) - (1+\sigma^2)^{-1} n}{(1+\sigma^2)^{-2} n} \tag{3.68}$$

or

$$\boxed{\underline{\hat{\sigma}}^2 \sim \frac{1+\sigma^2}{n}\chi^2(n,0) - 1} \tag{3.69}$$

Note that $1 + \underline{\hat{\sigma}}^2$ is the estimator of the variance factor of unit weight in the model

$$E\{\underline{x}_1 - \underline{x}_2\} = 0, \quad (1+\sigma^2)Q \tag{3.70}$$

Although the above example is a rather trivial one, it is of interest to elaborate the theory for the case of digitizing and connecting maps.

3.5 VCE and the $\underline{w}_i$-Test Statistics

Let us assume that the matrix Q_i of section 3 takes the form

$$Q_i = c_i c_i^*, \quad \text{with} \quad c_i = [0 \cdots 0\,1\,0 \cdots 0]^* \tag{3.71}$$

This implies that we want to estimate the variance σ_i^2 of *one single observation*. Since we have only one unknown variance-component, we can apply the result (3.45). Before doing this, we first note that in our case

$$Q_t = B^* \sum_{\alpha=1\, \alpha\neq i}^{p} Q_\alpha B + \sigma_i^2 B^* c_i c_i^* B \tag{3.72}$$

This we write as

$$Q_t = \bar{Q}_t + (\sigma_i^2 - 1) B^* c_i c_i^* B \tag{3.73}$$

with

$$\bar{Q}_t = \sum_{\alpha=1}^{p} B^* Q_\alpha B \tag{3.74}$$

In (3.45) we need the inverse of Q_t. Using the matrix-identity

$$(A - BD^{-1}C)^{-1} = A^{-1} + A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1} \tag{3.75}$$

the inverse of (3.73) follows as

$$Q_t^{-1} = \bar{Q}_t^{-1} - \frac{\bar{Q}_t^{-1} B^* c_i c_i^* B \bar{Q}_t^{-1}}{(\sigma_i^2 - 1)^{-1} + c_i^* B \bar{Q}_t^{-1} B^* c_i} \tag{3.76}$$

Using this together with (3.71) enables us to write

$$\underline{t}^* Q_t^{-1} B^* Q_i B Q_t^{-1} \underline{t} = [(1 - \frac{c_i^* B \bar{Q}_t^{-1} B^* c_i}{(\sigma_i^2 - 1)^{-1} + c_i^* B \bar{Q}_t^{-1} B^* c_i})(c_i^* B \bar{Q}_t^{-1} \underline{t})]^2 \tag{3.77}$$

In a similar way, we find

$$trace(B^* Q_i B Q_t^{-1}) = [(1 - \frac{c_i^* B \bar{Q}_t^{-1} B^* c_i}{(\sigma_i^2 - 1)^{-1} + c_i^* B \bar{Q}_t^{-1} B^* c_i})(c_i^* B \bar{Q}_t^{-1} B^* c_i)] \tag{3.78}$$

and

$$trace(B^* Q_i B Q_t^{-1} B^* Q_i B Q_t^{-1}) = [(1 - \frac{c_i^* B \bar{Q}_t^{-1} B^* c_i}{(\sigma_i^2 - 1)^{-1} + c_i^* B \bar{Q}_t^{-1} B^* c_i})(c_i^* B \bar{Q}_t^{-1} B^* c_i)]^2 \tag{3.79}$$

Substitution of (3.77), (3.78) and (3.79) into (3.45) gives

$$\begin{aligned} \underline{\hat{\sigma}}_i^2 &= \sigma_i^2 + \frac{(c_i^* B \bar{Q}_t^{-1} \underline{t})^2}{(c_i^* B \bar{Q}_t^{-1} B^* c_i)^2} \\ &- [(1 - \frac{c_i^* B \bar{Q}_t^{-1} B^* c_i}{(\sigma_i^2 - 1)^{-1} + c_i^* B \bar{Q}_t^{-1} B^* c_i})(c_i^* B \bar{Q}_t^{-1} B^* c_i)]^{-1} \end{aligned} \tag{3.80}$$

With

$$[(1 - \frac{c_i^* B \bar{Q}_t^{-1} B^* c_i}{(\sigma_i^2 - 1)^{-1} + c_i^* B \bar{Q}_t^{-1} B^* c_i})(c_i^* B \bar{Q}_t^{-1} B^* c_i)]^{-1} = (\sigma_i^2 - 1) + (c_i^* B \bar{Q}_t^{-1} B^* c_i)^{-1} \tag{3.81}$$

equation (3.80) simplifies to

$$\boxed{\underline{\hat{\sigma}}_i^2 = 1 + \frac{\dfrac{(c_i^* B \bar{Q}_t^{-1} \underline{t})^2}{c_i^* B \bar{Q}_t^{-1} B^* c_i} - 1}{c_i^* B \bar{Q}_t^{-1} B^* c_i}} \tag{3.82}$$

Note that this estimator is *independent* of the unknown σ_i^2. We also note, that since our well-known $\underline{w}_i$-test statistics reads

$$\underline{w}_i = \frac{c_i^* B \bar{Q}_t^{-1} \underline{t}}{(c_i^* B \bar{Q}_t^{-1} B^* c_i)^{1/2}} \tag{3.83}$$

the result (3.82) can be written as

$$\boxed{\underline{\hat{\sigma}}_i^2 = 1 + \frac{\underline{w}_i^2 - 1}{c_i^* B \bar{Q}_t^{-1} B^* c_i}} \tag{3.84}$$

This result also makes clear the sensitivity of the variance-component estimation for misspecifications in the *functional model*; a fact which also follows from the last theorem. With this theorem follows namely that if $E\{\underline{t}\} \neq 0$, then the variance-component estimators are distributed as a linear combination of *non-central* χ^2-distributions. Finally note that we did not make use in the above derivation of the fact that $c_i = [0 \cdots 0\, 1\, 0 \cdots 0]^*$.

3.6 CoVCE and the $\underline{w}_i$-Test Statistics

Let us assume that we want to estimate the covariance between two observations, say observation k and observation l. Matrix Q_i of section 3 takes then the form

$$Q_i = c_k c_l^* + c_l c_k^*, \quad \text{with} \quad c_k = [0 \cdots 0\, 1\, 0 \cdots 0]^* \tag{3.85}$$

The unknown covariance σ_{kl} can then be estimated according to (3.45) as

$$\hat{\underline{\sigma}}_{kl} = \sigma_{kl} + \frac{\underline{t}^* Q_t^{-1} B^* Q_i B Q_t^{-1} \underline{t} - trace(B^* Q_i B Q_t^{-1})}{trace(B^* Q_i B Q_t^{-1} B^* Q_i B Q_t^{-1})} \tag{3.86}$$

where

$$Q_t = \bar{Q}_t + \sigma_{kl} B^* (c_k c_l^* + c_l c_k^*) B \tag{3.87}$$

and $\bar{Q}_t$ is the covariance matrix of $\underline{t}$ in case $\sigma_{kl} = 0$. Note that σ_{kl} is allowed to be *positive and negative*. Hence the problem of *negative variance components* does *not* occur here. Also note that σ_{kl} need not be the total covariance. That is, if a covariance between observations k and l is included in $\bar{Q}_t$, then σ_{kl} of (3.87) should be interpreted as a perturbation or increment. The estimator $\hat{\underline{\sigma}}_{kl}$ follows if we substitute (3.85) and (3.87) into (3.86). As it turns out however the result unfortunately *depends* on the unknown σ_{kl}. Instead of the optimal estimator $\hat{\underline{\sigma}}_{kl}$, we therefore take an approximate, *but still unbiased*, $\hat{\underline{\sigma}}'_{kl}$ by choosing $\sigma_{kl} = 0$ in (3.86). This gives

$$\hat{\underline{\sigma}}'_{kl} = \frac{\underline{t}^* \bar{Q}_t^{-1} B^* Q_i B \bar{Q}_t^{-1} \underline{t} - trace(B^* Q_i B \bar{Q}_t^{-1})}{trace(B^* Q_i B \bar{Q}_t^{-1} B^* Q_i B \bar{Q}_t^{-1})} \tag{3.88}$$

With (3.85) we have

$$\begin{aligned}
\underline{t}^* \bar{Q}_t^{-1} B^* Q_i B \bar{Q}_t^{-1} \underline{t} &= 2(c_k^* B \bar{Q}_t^{-1} \underline{t})(c_l^* B \bar{Q}_t^{-1} \underline{t}) \\
trace(B^* Q_i B \bar{Q}_t^{-1}) &= 2 c_k^* B \bar{Q}_t^{-1} B^* c_l \\
trace(B^* Q_i B \bar{Q}_t^{-1} B^* Q_i B \bar{Q}_t^{-1}) &= 2(c_k^* B \bar{Q}_t^{-1} B^* c_l)^2 + 2(c_k^* B \bar{Q}_t^{-1} B^* c_k)(c_l^* B \bar{Q}_t^{-1} B^* c_l)
\end{aligned}$$

Substituting this into (3.88) gives

$$\boxed{\hat{\underline{\sigma}}'_{kl} = \frac{(c_k^* B \bar{Q}_t^{-1} \underline{t})(c_l^* B \bar{Q}_t^{-1} \underline{t}) - c_k^* B \bar{Q}_t^{-1} B^* c_l}{(c_k^* B \bar{Q}_t^{-1} B^* c_l)^2 + (c_k^* B \bar{Q}_t^{-1} B^* c_k)(c_l^* B \bar{Q}_t^{-1} B^* c_l)}} \tag{3.89}$$

Let us verify the unbiasedness of the estimator $\hat{\underline{\sigma}}'_{kl}$. With (3.87) we have

$$\begin{aligned}
E\{(c_k^* B \bar{Q}_t^{-1} \underline{t})(c_l^* B \bar{Q}_t^{-1} \underline{t})\} &= E\{(c_k^* B \bar{Q}_t^{-1} \underline{t}\underline{t}^* \bar{Q}_t^{-1} B^* c_l)\} = c_k^* B \bar{Q}_t^{-1} Q_t \bar{Q}_t^{-1} B^* c_l \\
&= c_k^* B \bar{Q}_t^{-1} B^* c_l + c_k^* B \bar{Q}_t^{-1} [\sigma_{kl} B^* (c_k c_l^* + c_l c_k^*) B] \bar{Q}_t^{-1} B^* c_l \\
&= c_k^* B \bar{Q}_t^{-1} B^* c_l + \sigma_{kl} [(c_k^* B \bar{Q}_t^{-1} B^* c_l)^2 \\
&+ \ (c_k^* B \bar{Q}_t^{-1} B^* c_k)(c_l^* B \bar{Q}_t^{-1} B^* c_l)]
\end{aligned}$$

With this result and (3.89) it follows that indeed $E\{\hat{\underline{\sigma}}'_{kl}\} = \sigma_{kl}$. If we use the abbreviation

$$n_{kl} = c_k^* B \bar{Q}_t^{-1} B^* c_l \tag{3.90}$$

and remember that

$$\underline{w}_k = \frac{c_k^* B \bar{Q}_t^{-1} \underline{t}}{\sqrt{n_{kk}}} \tag{3.91}$$

we can write (3.89) also as

$$\boxed{\hat{\underline{\sigma}}'_{kl} = \frac{\underline{w}_k \underline{w}_l - \dfrac{n_{kl}}{\sqrt{n_{kk}}\sqrt{n_{ll}}}}{\dfrac{n_{kl}^2}{\sqrt{n_{kk}}\sqrt{n_{ll}}} + \sqrt{n_{kk}}\sqrt{n_{ll}}}} \tag{3.92}$$

Note that under the hypothesis that $\sigma_{kl} = 0$, we have

$$E\{\underline{w}_k \underline{w}_l\} = Cov\{\underline{w}_k\,,\,\underline{w}_l\} = \frac{n_{kl}}{\sqrt{n_{kk}}\sqrt{n_{ll}}}, \quad \text{if} \quad \sigma_{kl} = 0. \tag{3.93}$$

This term equals the *cosine* of the angle between the two vectors c_k and c_l when projected with $P_A^{\perp}$. It is closely related to the *error of the third kind*. That is, if (3.93) is too large one will have difficulty in discriminating between two hypotheses $E\{\underline{t}\} = B^* c_k \nabla_k$ and $E\{\underline{t}\} = B^* c_l \nabla_l$.

Chapter 4

Estimating and Testing σ_1^2 and σ_2^2 in $\sigma_{0i}^2 = \sigma_1^2 + \sigma_2^2 x_i^q$ for EDM's

4.1 VCE from Repeated Measurements

According to the Least-Squares Estimators of the variance-components, $\sigma_\alpha^2,\ \alpha = 1, 2, \cdots, p$ in the model

$$E\{\underline{y}\} = Ax, \quad E\{(\underline{y} - Ax)(\underline{y} - Ax)^*\} = \sum_{\alpha=1}^{p} \sigma_\alpha^2 Q_\alpha \tag{4.1}$$

are given as

$$\begin{pmatrix} \hat{\underline{\sigma}}_1^2 \\ \hat{\underline{\sigma}}_2^2 \\ \vdots \\ \hat{\underline{\sigma}}_p^2 \end{pmatrix} = \begin{bmatrix} n_{11} & n_{12} & \cdots & n_{1p} \\ n_{21} & n_{22} & \cdots & n_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ n_{p1} & n_{p2} & \cdots & n_{pp} \end{bmatrix}^{-1} \begin{bmatrix} \frac{1}{2}\hat{\underline{e}}^* Q_y^{-1} Q_1 Q_y^{-1} \hat{\underline{e}} \\ \frac{1}{2}\hat{\underline{e}}^* Q_y^{-1} Q_2 Q_y^{-1} \hat{\underline{e}} \\ \vdots \\ \frac{1}{2}\hat{\underline{e}}^* Q_y^{-1} Q_p Q_y^{-1} \hat{\underline{e}} \end{bmatrix} \tag{4.2}$$

with:

$$n_{\beta\alpha} = \frac{1}{2} trace(Q_\beta Q_y^{-1} P_A^\perp Q_\alpha Q_y^{-1} P_A^\perp) \tag{4.3}$$

and with

$$Q_y = \sum_{\alpha=1}^{p} \sigma_\alpha^2 Q_\alpha; \quad P_A^\perp = I - A(A^* Q_y^{-1} A)^{-1} A^* Q_y^{-1}; \quad \hat{\underline{e}} = P_A^\perp \underline{y} \tag{4.4}$$

We will apply the above results to the model

$$E\{\underline{y}\} = E\{\underbrace{\begin{bmatrix} \underline{y}_1 \\ \underline{y}_2 \\ \vdots \\ \underline{y}_r \end{bmatrix}}_{m\,r\times 1}\} = \underbrace{\begin{bmatrix} e & 0 & \cdots & 0 \\ 0 & e & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e \end{bmatrix}}_{m\,r\times r} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \end{bmatrix}}_{r\times 1}, \quad Q_y = \underbrace{\begin{bmatrix} \sigma_{01}^2 I_m & 0 & \cdots & 0 \\ 0 & \sigma_{02}^2 I_m & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{0r}^2 I_m \end{bmatrix}}_{m\,r\times m\,r} \tag{4.5}$$

with:

$$e = [1\ 1\ \cdots\ 1]^*, \quad \sigma_{0i}^2 = \sigma_1^2 + \sigma_2^2 x_i^q,\ i = 1, 2, \cdots, r \tag{4.6}$$

Model (4.5) is valid for the case where one measures an r-number of unknown distances $x_i,\ i = 1, 2, \cdots, r$, each an m-number of times. It is assumed that all the observations are *uncorrelated*.

Furthermore it is assumed that all the precision of the measurements is constant for a constant distance, but that it varies with the distance according to the *law*

$$\sigma_{0i}^2 = \sigma_1^2 + \sigma_2^2 x_i^q, \ \ i = 1, 2, \cdots, r \tag{4.7}$$

where q is an exponent of the distance which one can choose and σ_1^2 and σ_2^2 are the unknown variance-components which need to be estimated. Thus in our case we have two unknowns and the matrices $Q_\alpha, \ \alpha = 1, 2$ of (4.1) take the form

$$Q_1 = \underbrace{\begin{bmatrix} I_m & 0 & \cdots & 0 \\ 0 & I_m & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & I_m \end{bmatrix}}_{mr \times mr}, \quad Q_2 = \underbrace{\begin{bmatrix} x_1^q I_m & 0 & \cdots & 0 \\ 0 & x_2^q I_m & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_r^q I_m \end{bmatrix}}_{mr \times mr} \tag{4.8}$$

In order to apply (4.2) we need $P_A^\perp$ of (4.4). This matrix takes in our case a very simple form:

$$P_A^\perp = \underbrace{\begin{bmatrix} P_e^\perp & 0 & \cdots & 0 \\ 0 & P_e^\perp & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & P_e^\perp \end{bmatrix}}_{mr \times mr} \tag{4.9}$$

with $P_e^\perp = I_m - \frac{1}{m} ee^*$. Note that $P_A^\perp$ is *independent* of the $\sigma_{0i}^2, \ i = 1, 2, \cdots, r$. Also note that the block matrices $P_e^\perp$ of $P_A^\perp$ correspond to the separate adjustment of each unknown distance. That is, per unknown distance we have an adjustment-problem with m-number of observations, one unknown distance and one variance-factor of unit weight σ_{0i}^2. From adjustment theory we know that the variance factor of unit weight can be estimated rather straightforward. In our case the separate estimators of $\sigma_{0i}^2, \ i = 1, 2, \cdots, r$ become

$$\boxed{\underline{\hat{\sigma}}_{0i}^2 = \frac{\underline{y}_i^* P_e^\perp \underline{y}_i}{m-1} \ \text{ with } \ E\{\underline{\hat{\sigma}}_{0i}^2\} = \sigma_{0i}^2, \ \ \sigma_{\hat{\sigma}_{0i}^2}^2 = \frac{2\sigma_{0i}^4}{m-1}, \ i = 1, 2, \cdots, r} \tag{4.10}$$

This result may be used to perform a *global* test for each distance separately. It may also be used for obtaining a reasonable value for m, i.e., the number of measurements. Parallel to (4.10) we may also perform *data snooping* for each distance separately. The w-test statistics for the k^{th}-observation in the i^{th}-distance reads

$$\boxed{\underline{w}_{ki} = \frac{\underline{y}_{ki} - \frac{1}{m}\sum_{l=1}^m \underline{y}_{li}}{\sigma_{0i}\sqrt{1 - \frac{1}{m}}}, \quad i = 1, 2, \cdots, r, \quad k = 1, 2, \cdots, m} \tag{4.11}$$

Once the r-number of estimates $\hat{\sigma}_{0i}^2$ of (4.10) are available, they may be used to get a first indication of whether law (4.7) holds or not. This may be done by plotting the $\hat{\sigma}_{0i}^2$ against the x_i^q. Of course, x_i is unknown, but here one can use the *mean* of the m-number of observed distances. The plot should then look something like:

By interpreting the estimates $\hat{\sigma}_{0i}^2$ of (4.10) as observations, we can now with the help of (4.7) construct the following linear model of observation equations:

$$\boxed{E\{\begin{bmatrix} \underline{\hat{\sigma}}_{01}^2 \\ \underline{\hat{\sigma}}_{02}^2 \\ \vdots \\ \underline{\hat{\sigma}}_{0r}^2 \end{bmatrix}\} = \begin{bmatrix} 1 & x_1^q \\ 1 & x_2^q \\ \vdots & \vdots \\ 1 & x_r^q \end{bmatrix} \begin{bmatrix} \sigma_1^2 \\ \sigma_2^2 \end{bmatrix}, \ \ Q_{\hat{\sigma}_{0i}^2} = \frac{2}{m-1} \underbrace{\begin{bmatrix} \sigma_{01}^4 & 0 & \cdots & 0 \\ 0 & \sigma_{02}^4 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{0r}^4 \end{bmatrix}}_{r \times r}} \tag{4.12}$$

Note that because of our assumptions in (4.5), the $\hat{\underline{\sigma}}^2_{0i}$ are distributed as *independent* χ^2-variables. The matrix $Q_{\hat{\sigma}^2_{0i}}$ is therefore diagonal! In order to find estimators for σ_1^2 and σ_2^2 we can now apply all sorts of estimation principles (robust methods, maximum likelihood, least squares etc). We will solve (4.12) using the *least-squares principle*. The normal matrix of (4.12) reads

$$N = \frac{m-1}{2}\begin{bmatrix} \sum_{i=1}^r \sigma_{0i}^{-4} & \sum_{i=1}^r \sigma_{0i}^{-4} x_i^q \\ \sum_{i=1}^r \sigma_{0i}^{-4} x_i^q & \sum_{i=1}^r \sigma_{0i}^{-4} x_i^{2q} \end{bmatrix} \tag{4.13}$$

Its inverse reads

$$N^{-1} = \frac{2}{m-1}[(\sum_{i=1}^r \sigma_{0i}^{-4})(\sum_{i=1}^r \sigma_{0i}^{-4} x_i^{2q}) - (\sum_{i=1}^r \sigma_{0i}^{-4} x_i^q)^2]^{-1}\begin{bmatrix} \sum_{i=1}^r \sigma_{0i}^{-4} x_i^{2q} & -\sum_{i=1}^r \sigma_{0i}^{-4} x_i^q \\ -\sum_{i=1}^r \sigma_{0i}^{-4} x_i^q & \sum_{i=1}^r \sigma_{0i}^{-4} \end{bmatrix} \tag{4.14}$$

The right-hand side of the normal equations reads:

$$\underline{l} = \frac{m-1}{2}\begin{bmatrix} \sum_{i=1}^r \sigma_{0i}^{-4} \hat{\underline{\sigma}}^2_{0i} \\ \sum_{i=1}^r \sigma_{0i}^{-4} x_i^q \hat{\underline{\sigma}}^2_{0i} \end{bmatrix} \tag{4.15}$$

With (4.14) and (4.15) the solution of (4.12) follows as

$$\boxed{\begin{bmatrix} \hat{\underline{\sigma}}^2_1 \\ \hat{\underline{\sigma}}^2_2 \end{bmatrix} = N^{-1}\underline{l}, \quad E\{\begin{bmatrix} \hat{\underline{\sigma}}^2_1 \\ \hat{\underline{\sigma}}^2_2 \end{bmatrix}\} = \begin{bmatrix} \sigma_1^2 \\ \sigma_2^2 \end{bmatrix}, \quad D\{\begin{bmatrix} \hat{\underline{\sigma}}^2_1 \\ \hat{\underline{\sigma}}^2_2 \end{bmatrix}\} = N^{-1}} \tag{4.16}$$

We have now devised a *two-step or phased procedure* for estimating the variance-components σ_1^2 and σ_2^2. First (4.10) is used to compute the $\hat{\underline{\sigma}}^2_{0i}$, $i = 1, 2, \cdots, r$. Then in a second step the variance-components are computed according to (4.16). The solution so obtained is <u>*identical*</u> to the solution one gets when applying (4.2) and (4.5)!! Note that in the second step *iterations* are needed since the variances $\frac{2}{m-1}\sigma^2_{0i}$ of (4.12) are unknown a-priori. In fact also the x_i^q in the design matrix of (4.12) are unknown, but here it probably suffices to take the mean of the observed distances.

In order that the two estimators $\hat{\underline{\sigma}}^2_1$ and $\hat{\underline{\sigma}}^2_2$ are well-separated their *correlation coefficient* should be small enough. From (4.14) this correlation coefficient follows as

$$\boxed{\rho_{12} = \frac{-\sum_{i=1}^r \sigma_{0i}^{-4} x_i^q}{\sqrt{\sum_{i=1}^r \sigma_{0i}^{-4} x_i^{2q}}\sqrt{\sum_{i=1}^r \sigma_{0i}^{-4}}}} \tag{4.17}$$

This correlation coefficient depends on the angle between the two column vectors of the design matrix of (4.12). More precisely: the correlation coefficient ρ_{12} is small if the distances x_i, $i = 1, 2, \cdots, r$ are chosen such that the angle between the two vectors

$$\begin{bmatrix} \sigma_{01}^{-2} \\ \sigma_{02}^{-2} \\ \vdots \\ \sigma_{0r}^{-2} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \sigma_{01}^{-2} x_1^2 \\ \sigma_{01}^{-2} x_2^2 \\ \vdots \\ \sigma_{01}^{-2} x_r^2 \end{bmatrix} \tag{4.18}$$

is large. Once the estimates $\hat{\sigma}_1^2$ and $\hat{\sigma}_1^2$ are computed, one can try to test their *significance* with respect to the values given by the manufacturer. If we denote the values given by the manufacturer as $\bar{\sigma}_1^2$ and $\bar{\sigma}_2^2$, the test statistic may take the form:

$$\boxed{\underline{\nu}_i = \frac{\hat{\underline{\sigma}}^2_i - \bar{\sigma}^2_i}{\sigma_{\hat{\sigma}^2_i}}, \quad i = 1, 2} \tag{4.19}$$

Although the distribution of $\underline{\nu}_i$ is *unknown*, we may try the *standard normal distribution* as a crude approximation. With this approximation the test can be performed.

Final Remarks

1. From the structure of (4.12) follows that it is not necessary to assume that the number of observation per unknown distance is constant.

2. The structure of model (4.12) resembles the *1-D Helmert-transformation* $E\{\underline{\hat{\sigma}}^2_{0i}\} = \sigma_1^2 + x_i^q \sigma_2^2$. If x_i^q is considered to be stochastic, the model can be written in the form of the *1-D symmetric Helmert transformation* $E\{\underline{\hat{\sigma}}^2_{0i}\} = \sigma_1^2 + E\{\underline{x}_i^q\}\sigma_2^2$. The solution method of (Teunissen: The 1- and 2-D symmetric Helmert transformation, report 87.1, Delft) can then be applied.

3. Note that (4.12) may also be solved *recursively*. This may be of use if one wants to update the estimates of σ_1^2 and σ_2^2 if a new unknown distance is measured.

Chapter 5

Estimation and Testing of Covariance Matrices

5.1 Introduction

Consider the following two hypotheses:

$$\begin{cases} H_0: \quad E\{\underline{y}\} = Ax, \quad B^* x = b, \quad D\{\underline{y}\} = Q_y \\ \\ H_A: \quad E\{\underline{y}\} = Ax, \quad D\{\underline{y}\} = Q_y \end{cases} \tag{5.1}$$

We assume that $\underline{y}$ is normally distributed. The appropriate test statistic is then given by [see lecture notes MGII]:

$$\begin{cases} \underline{T} = [B^* \underline{\hat{x}}_A - b]^* (B^* Q_{\hat{x}_A} B)^{-1} [B^* \underline{\hat{x}}_A - b], \text{ with} \\ \\ \underline{\hat{x}}_A = Q_{\hat{x}_A} A^* Q_y^{-1} \underline{y}, \quad Q_{\hat{x}_A} = (A^* Q_y^{-1} A)^{-1} \end{cases} \tag{5.2}$$

$\underline{T}$ has the following distribution:

$$\begin{cases} H_0: \quad \underline{T} \sim \chi^2(b, 0) \\ \\ H_A: \quad \underline{T} \sim \chi^2(b, \lambda), \quad \text{with} \quad \lambda = [B^* x - b]^* (B^* Q_y B)^{-1} [B^* x - b] \end{cases} \tag{5.3}$$

The test statistic $\underline{T}$ also follows from the *Generalized Likelihood Ratio Test.*
Note: b= number of parameters under H_A minus number of parameters under H_0.

5.2 The Model and Its Solution

As model we consider

$$E\{\underbrace{\underline{y}}_{rm \times 1}\} = \underbrace{(I_r \otimes A)}_{rm \times rn} \underbrace{x}_{rn \times 1}, \quad D\{\underline{y}\} = \underbrace{Q \otimes I_m}_{rm \times rm} \tag{5.4}$$

According to [Teunissen, 1988]:

$$\underbrace{\frac{1}{2} tr(Q_\alpha Q_y^{-1} P_A^\perp Q_\beta Q_y^{-1} P_A^\perp)}_{N_{\alpha\beta}} \hat{\sigma}_\beta^2 = \underbrace{\frac{1}{2} \hat{e}^* Q_y^{-1} Q_\alpha Q_y^{-1} \hat{e}}_{l_\alpha} \tag{5.5}$$

From (5.4) follows that

$$Q_y = Q \otimes I_m, \quad P_A = I_r \otimes A(A^* A)^{-1} A^*, \quad Q_\alpha = Q_\alpha \otimes I_m \tag{5.6}$$

This gives

$$\begin{aligned} N_{\alpha\beta} &= \frac{1}{2} tr(Q_\alpha \otimes I_m Q^{-1} \otimes I_m I_r \otimes P_A^\perp Q_\beta \otimes I_m Q^{-1} \otimes I_m I_r \otimes P_A^\perp) \\ &= \frac{1}{2} tr(Q_\alpha Q^{-1} Q_\beta Q^{-1} \otimes P_A^\perp) \\ &= \frac{1}{2} tr(Q_\alpha Q^{-1} Q_\beta Q^{-1}) tr(P_A^\perp) \end{aligned}$$

or

$$\boxed{N_{\alpha\beta} = \frac{1}{2}(m-n) tr(Q_\alpha Q^{-1} Q_\beta Q^{-1})} \tag{5.7}$$

In our case

$$\hat{\sigma}^2 = (\hat{\sigma}_\beta^2) = \begin{pmatrix} \hat{\sigma}_{11} & \hat{\sigma}_{21} & \ldots & \hat{\sigma}_{r1} & \hat{\sigma}_{22} & \ldots & \hat{\sigma}_{r2} & \ldots & \hat{\sigma}_{rr} \end{pmatrix}^* \tag{5.8}$$

We *define* the matrix L as:

$$\boxed{\underbrace{L}_{r^2 \times r(r+1)/2} \underbrace{\nabla(X)}_{r(r+1)/2 \times 1} = \underbrace{vec(X)}_{r^2 \times 1} \quad \text{for any symmetric} \quad X_{r \times r}}. \tag{5.9}$$

Since the matrix L has full rank $r(r+1)/2$, it follows from (5.9) that

$$\nabla(X) = (L^* L)^{-1} L^* vec(X) \quad \text{for any symmetric} \quad X \tag{5.10}$$

With the projector P [$P\, vec(X) = vec(X)$ for any $X = X^*$] and (5.9) follows that

$$P\, L\, \nabla(X) = P\, vec(X) = vec(X) = L\, \nabla(X) \quad \text{for any} \quad X^* = X \tag{5.11}$$

and thus

$$PL = L \quad \text{or} \quad R(L) \subset R(P) \tag{5.12}$$

Since $rank(P) = rank(L) = r(r+1)/2$ or $dim(R(P)) = dim(R(L))$ it follows with (5.12) that

$$R(P) = R(L) \tag{5.13}$$

This shows, since P is a projector that

$$\boxed{P = L(L^* L)^{-1} L^*} \tag{5.14}$$

We will now derive the inverse of $(N_{\alpha\beta})$. From (5.7) and (5.9) follows that

$$\begin{aligned} N_{\alpha\beta} &= \frac{1}{2}(m-n)(vec(Q_\alpha))^* Q^{-1} \otimes Q^{-1} (vec(Q_\beta)) \\ &= \frac{1}{2}(m-n) \nabla(Q_\alpha)^* L^* (Q^{-1} \otimes Q^{-1}) L \nabla(Q_\beta) \end{aligned}$$

Note that $\nabla(Q_\beta)$ is the *identity matrix* of order $r(r+1)/2$. Hence:

$$Q_{\nabla(\hat{Q})} = (N_{\alpha\beta})^{-1} = [L^*(Q^{-1} \otimes Q^{-1}) L]^{-1} \frac{2}{m-n} \tag{5.15}$$

or

$$L^*(Q^{-1} \otimes Q^{-1}) L Q_{\nabla(\hat{Q})} = \frac{2}{m-n} I \tag{5.16}$$

or

$$\underbrace{L(L^* L)^{-1} L^*}_{P} (Q^{-1} \otimes Q^{-1}) L Q_{\nabla(\hat{Q})} = \frac{2}{m-n} L(L^* L)^{-1} \tag{5.17}$$

or

$$(Q^{-1} \otimes Q^{-1})PLQ_{\nabla(\hat{Q})} = \frac{2}{m-n} L(L^*L)^{-1} \tag{5.18}$$

or

$$PLQ_{\nabla(\hat{Q})} = \frac{2}{m-n}(Q \otimes Q)L(L^*L)^{-1} \tag{5.19}$$

or

$$\boxed{\underbrace{Q_{\nabla(\hat{Q})}}_{r(r+1)/2 \times r(r+1)/2} = \frac{2}{m-n}(L^*L)^{-1}L^*(Q \otimes Q)L(L^*L)^{-1}} \tag{5.20}$$

From (5.5) and (5.6) follows that

$$\begin{aligned} l_\alpha &= \frac{1}{2} y^* I_r \otimes P_A^\perp Q^{-1} \otimes I_m Q_\alpha \otimes I_m Q^{-1} \otimes I_m I_r \otimes P_A^\perp y \\ &= \frac{1}{2} y^* (Q^{-1} Q_\alpha Q^{-1} \otimes P_A^\perp) y \end{aligned} \tag{5.21}$$

With $y = \sum_{i=1}^r e_i \otimes y_i = \sum_{i=1}^r vec(y_i e_i^*)$ this gives

$$\begin{aligned} l_\alpha &= \frac{1}{2}(vec \sum_{i=1}^r y_i e_i^*)^* (Q^{-1} Q_\alpha Q^{-1} \otimes P_A^\perp)(vec \sum_{j=1}^r y_j e_j^*) \\ &= \frac{1}{2} trace[Q^{-1} Q_\alpha Q^{-1} (\sum_{j=1}^r y_j e_j^*)^* P_A^\perp (\sum_{i=1}^r y_i e_i^*)] \end{aligned} \tag{5.22}$$

With $\underbrace{Y}_{m \times r} = [y_1, y_2, ..., y_r] = \sum_{i=1}^r y_i e_i^*$ this gives

$$\begin{aligned} l_\alpha &= \frac{1}{2} trace[Q^{-1} Q_\alpha Q^{-1} Y^* P_A^\perp Y] \\ &= \frac{1}{2}[vec(Q^{-1} Q_\alpha Q^{-1})]^* [vec Y^* P_A^\perp Y] \\ &= \frac{1}{2}[Q^{-1} \otimes Q^{-1} vec Q_\alpha]^* [vec Y^* P_A^\perp Y] \end{aligned} \tag{5.23}$$

or

$$\boxed{l_\alpha = \frac{1}{2} vec(Q_\alpha)^* Q^{-1} \otimes Q^{-1} vec(Y^* P_A^\perp Y)} \tag{5.24}$$

with $vec(Q_\alpha) = L\nabla(Q_\alpha) = L$.Identity, this gives

$$\boxed{l = \frac{1}{2} L^* (Q^{-1} \otimes Q^{-1}) L \nabla(Y^* P_A^\perp Y)} \tag{5.25}$$

From (5.20) and (5.25) follows that:

$$\begin{aligned} \nabla(\hat{Q}) &= \frac{2}{m-n}(L^*L)^{-1}L^*(Q \otimes Q)L(L^*L)^{-1}.\frac{1}{2}L^*(Q^{-1} \otimes Q^{-1})L\nabla(Y^* P_A^\perp Y) \\ &= \frac{2}{m-n}.\frac{1}{2}(L^*L)^{-1}L^*(Q \otimes Q)P(Q^{-1} \otimes Q^{-1})L\nabla(Y^* P_A^\perp Y) \\ &= \frac{1}{m-n}(L^*L)^{-1}L^*PL\nabla(Y^* P_A^\perp Y) \end{aligned} \tag{5.26}$$

or

$$\boxed{\underbrace{\hat{Q}}_{r \times r} = \frac{1}{m-n} Y^* P_A^\perp Y} \tag{5.27}$$

$\hat{Q}$ has a *Wishart distribution*. For later use it is important to know *when* $\nabla(\hat{Q})$ can be approximated by a normal distribution.

5.3 The Teststatistic & Restrictions $B^*QB = \phi_0$

The following two hypotheses will be considered:

$$\begin{cases} H_0: & E\{\underline{y}\} = (I_r \otimes A)x, \quad D\{\underline{y}\} = Q \otimes I_m, \quad B^*QB = \phi_0 \\ H_A: & E\{\underline{y}\} = (I_r \otimes A)x, \quad D\{\underline{y}\} = Q \otimes I_m, \end{cases} \tag{5.28}$$

The restrictions $B^*QB = \phi_0$ can be written as $vec(B^*QB) = vec(\phi_0)$ or as the *linear* restrictions

$$B^* \otimes B^* vec(Q) = vec(\phi_0) \tag{5.29}$$

Since Q is symmetric, (5.29) does not contain *independent* restrictions. Therefore

$$(B^* \otimes B^*)L_r \nabla Q = L_p \nabla \phi_0 \tag{5.30}$$

or

$$\boxed{(L_p^* L_p)^{-1} L_p^* (B^* \otimes B^*) L_r \nabla Q = \nabla \phi_0} \tag{5.31}$$

The teststatistic [see (5.2)] follows then as:

$$\underline{T} \;=\; L_p^+(B^* \otimes B^*)L_r \nabla(\hat{Q}) - \nabla\phi_0]^*[L_p^+(B^* \otimes B^*)L_r Q_{\nabla(\hat{Q})} L_r (B \otimes B) L_p^{+*}]^{-1} \tag{5.32}$$

$$[L_p^+(B^* \otimes B^*)L_r \nabla(\hat{Q}) \nabla\phi_0] \tag{5.33}$$

This gives with (5.20):

$$\underline{T} \;=\; \frac{m-n}{2}[L_p^+(B^* \otimes B^*)L_r \nabla(\hat{Q}) - \nabla(\phi_0)]^*[L_p^+(B^* \otimes B^*)P_r Q \otimes Q P_r (B \otimes B) L_p^{+*}]^{-1} [L_p^+(B^* \otimes B^*)L_r \nabla(\hat{Q}) \nabla(\phi_0)] \tag{5.34}$$

with

$$\begin{aligned} P_r &= \frac{1}{2}[I_{r^2} + K_{rr}] \\ K_{rr}(B \otimes B) &= (B \otimes B)K_{pp} \\ P_r(B \otimes B) &= (B \otimes B)P_p \end{aligned}$$

follows from (5.35) that

$$\underline{T} \;=\; \frac{m-n}{2}[L_p^+(B^* \otimes B^*)L_r \nabla(\hat{Q}) - \nabla(\phi_0)]^*[L_p^+(B^*QB \otimes B^*QB)L_p^{+*}]^{-1} [L_p^+(B^* \otimes B^*)L_r \nabla(\hat{Q}) \nabla(\phi_0)] \tag{5.35}$$

Since

$$[L_p^+(B^*QB \otimes B^*QB)L_p^{+*}]^{-1} = L_p^*(B^*QB)^{-1} \otimes (B^*QB)^{-1} L_p \tag{5.36}$$

and

$$\nabla(\phi_0) = L_p^+(B^* \otimes B^*)L_r \nabla(Q_0) \tag{5.37}$$

Equation (5.36) can be written as

$$\underline{T} = \frac{m-n}{2}[\nabla(\hat{Q} - Q_0)^* L_r (B \otimes B) L_p^{+*} L_p^* (B^*QB)^{-1} \otimes (B^*QB)^{-1} L_p L_p^+ (B^* \otimes B^*) L_r \nabla(\hat{Q} - Q_0)] \tag{5.38}$$

or as

$$\begin{aligned} \underline{T} &= \frac{m-n}{2}[\nabla(\hat{Q} - Q_0)^* L_r^* [B(B^*QB)^{-1}B^*] \otimes [B(B^*QB)^{-1}B^*] L_r \nabla(\hat{Q} - Q_0) \\ &= \frac{m-n}{2}[vec(\hat{Q} - Q_0)]^* [B(B^*QB)^{-1}B^*] \otimes [B(B^*QB)^{-1}B^*][vec(\hat{Q} - Q_0)] \\ &= \frac{m-n}{2} trace[B(B^*QB)^{-1}B^* \, (\hat{Q} - Q_0) \, B(B^*QB)^{-1}B^* \, (\hat{Q} - Q_0)] \\ &= \frac{m-n}{2} trace[(B^*QB)^{-1}(B^*\hat{Q}B - B^*Q_0B) \, (B^*QB)^{-1}(B^*\hat{Q}B - B^*Q_0B)] \end{aligned} \tag{5.39}$$

Under H_0 we have $B^*QB = B^*Q_0B = \phi_0$. Thus *under* H_0 we get

$$\boxed{T_{H_0} = \frac{m-n}{2} trace[[(B^*Q_0B)^{-1}(B^*\hat{Q}B) - I_p][(B^*Q_0B)^{-1}(B^*\hat{Q}B) - I_p]]} \tag{5.40}$$

with

$$|B^*\underline{\hat{Q}}B - \underline{\hat{\lambda}_i} \overbrace{B^*Q_0B}^{\phi_0}| = 0, \; i = 1, 2, ..., p \tag{5.41}$$

We may write (5.40) also as

$$\boxed{\underline{T}_{H_0} = \frac{m-n}{2} \sum_{i=1}^{p} (\underline{\hat{\lambda}}_i - 1)^2} \tag{5.42}$$

Since the difference in the number of parameters between H_A and H_0 is $r(r+1)/2 - \{r(r+1)/2 - p(p+1)/2\}$, the form of the teststatistic *suggests* that I_{H_0} is approximately distributed as

$$\boxed{H_0: \; \underline{T}_{H_0} \sim \chi^2(p(p+1)/2, 0)} \tag{5.43}$$

5.4 A Comparison with the Restricted Generalized Likelihood Ratio Test

If we define $\underline{t} = (I_r \otimes B^*)\underline{y}$ [do not confuse this B, with B of section 3], with $B^*A = 0$, it follows from (5.4) that

$$\underbrace{E\{\underline{t}\}}_{r(m-n)\times 1} = 0, \quad D\{\underline{t}\} = \underbrace{Q}_{r\times r} \otimes \underbrace{B^*B}_{(m-n)\times(m-n)} \tag{5.44}$$

The following hypotheses will be considered

$$H_0': \; D\{\underline{t}\} = Q_0 \otimes B^*B, \quad H_A': \; D\{\underline{t}\} = Q \otimes B^*B \tag{5.45}$$

The restricted likelihood function reads then

$$\log \rho_{\underline{t}}(t/Q) = -\frac{1}{2}\log 2\pi - \frac{1}{2}\log|Q \otimes B^*B| - \frac{1}{2}t^*Q^{-1} \otimes (B^*B)^{-1}t \tag{5.46}$$

From this follows that

$$\begin{aligned} -2\log\frac{\rho_{\underline{t}}(t/Q_0)}{\rho_{\underline{t}}(t/\hat{Q})} &= \log|Q_0\hat{Q}^{-1} \otimes I_{m-n}| + t^*[(Q_0^{-1} - \hat{Q}^{-1}) \otimes (B^*B)]t \\ &= (m-n)\log|Q_0\hat{Q}^{-1}| + y^*[(Q_0^{-1} - \hat{Q}^{-1}) \otimes P_B]y \end{aligned} \tag{5.47}$$

With $y = \sum_{i=1}^r e_i \otimes y_i = \sum_i^r vec(y_ie_i^*)$ this gives

$$\begin{aligned} -2\log\frac{\rho_{\underline{t}}(t/Q_0)}{\rho_{\underline{t}}(t/\hat{Q})} &= (m-n)\log|Q_0\hat{Q}^{-1}| + (vec\sum_{i=1}^r y_ie_i^*)^*[(Q_0^{-1} - \hat{Q}^{-1}) \otimes P_B](vec\sum_{i=1}^r y_ie_i^*) \\ &= (m-n)\log|Q_0\hat{Q}^{-1}| + trace[(Q_0^{-1} - \hat{Q}^{-1})\sum_{j=1}^r y_je_j^*P_B\sum_i^r y_ie_i^*] \\ &= (m-n)\log|Q_0\hat{Q}^{-1}| + trace[(Q_0^{-1} - \hat{Q}^{-1})Y^*P_A^{\perp}Y] \\ &= (m-n)\{\log|Q_0\hat{Q}^{-1}| + trace[Q_0^{-1}\hat{Q} - I]\} \\ &= -(m-n)\{\log|Q_0^{-1}\hat{Q}| + trace[I - Q_0^{-1}\hat{Q}]\} \end{aligned} \tag{5.48}$$

Hence

$$\boxed{-2\log\frac{\rho_{\underline{t}}(t/Q_0)}{\rho_{\underline{t}}(t/\hat{Q})} = -(m-n)\{\sum_{i=1}^r \log\underline{\hat{\lambda}}_i + \sum_{i=1}^r (1 - \underline{\hat{\lambda}}_i)\}} \tag{5.49}$$

Substitution of

$$\log \underline{\hat{\lambda}}_i = \log 1 + (\underline{\hat{\lambda}}_i - 1) - \frac{1}{2}(\underline{\hat{\lambda}}_i - 1)^2 + ... \tag{5.50}$$

into (5.49) shows that

$$\boxed{-2\log\frac{\rho_{\underline{t}}(t/Q_0)}{\rho_{\underline{t}}(t/\hat{Q})} \doteq \frac{m-n}{2}\sum_{i=1}^{r}(\underline{\hat{\lambda}}_i - 1)^2 \;=\; I_{H_0}} \tag{5.51}$$

5.5 The Teststatistic and Restrictions $B^*QC = 0$

The following two hypotheses will be considered:

$$\begin{cases} H_0: \quad E\{\underline{y}\} = (I_r \otimes A)x \quad D\{\underline{y}\} = Q \otimes I_m \quad B^*QC = 0 \\ \\ H_A: \quad E\{\underline{y}\} = (I_r \otimes A)x \quad D\{\underline{y}\} = Q \otimes I_m \end{cases} \tag{5.52}$$

The restriction $B^*QC = 0$ can be written as $vec(B^*QC) = vec(Q) = 0$ or as the *linear* restrictions:

$$(C^* \otimes B^*)vec(Q) = 0 \;\text{ or }\; (C^* \otimes B^*)L_r\nabla(Q) = 0 \tag{5.53}$$

The teststatistic [see (5.2)] follows then as:

$$\underline{T} = \frac{m-n}{2}[C^* \otimes B^* vec(\hat{Q})]^*[C^* \otimes B^* P_r Q \otimes Q P_r C \otimes B]^{-1}[C^* \otimes B^* vec(\hat{Q})] \tag{5.54}$$

with

$$\begin{cases} P_r = \frac{1}{2}(I_{r^2 + K_{rr}}) \qquad\qquad P_r Q \otimes Q P_r = Q \otimes Q P_r \\ \\ K_{rr}(C \otimes B) = (B \otimes C)K_{pq} \end{cases} \tag{5.55}$$

follows from (5.54) that

$$\underline{T} = \frac{m-n}{2}[C^* \otimes B^* vec(\hat{Q})]^*[C^* \otimes B^* Q \otimes Q\{\frac{1}{2}C \otimes B + \frac{1}{2}B \otimes C K_{pq}\}]^{-1}[C^* \otimes B^* vec(\hat{Q})] \tag{5.56}$$

or

$$\underline{T} = (m-n)[C^* \otimes B^* vec(\hat{Q})]^*[C^*QC \otimes B^*QB + C^*QB \otimes B^*QCK_{pq}]^{-1}[C^* \otimes B^* vec(\hat{Q})] \tag{5.57}$$

Therefore, under H_0:

$$\begin{aligned} \underline{T}_{H_0} &= (m-n)vec(\hat{Q})^* C \otimes B(C^*QC)^{-1} \otimes (B^*QB)^{-1}C^* \otimes B^* vec(\hat{Q}) \\ &= (m-n)vec(\hat{Q})^* C(C^*QC)^{-1}C^* \otimes B(B^*QB)^{-1}B^* vec(\hat{Q}) \\ &= (m-n)trace[C(C^*QC)^{-1}C^*\hat{Q}B(B^*QB)^{-1}B^*\hat{Q}] \end{aligned} \tag{5.58}$$

or

$$\boxed{\underline{T}_{H_0} = (m-n)trace[(C^*QC)^{-1}C^*\hat{Q}B\,(B^*QB)^{-1}B^*\hat{Q}C]} \tag{5.59}$$

The form of (5.54) suggests that

$$H_0: \; \underline{T}_{H_0} \sim \chi^2(pq, 0) \tag{5.60}$$

Chapter 6

A New Method for Estimating and Testing the Substitute Matrix H

6.1 The Generalized Eigenvalue Problem

Let $\underline{x}$ be a random n-vector with variance matrix Q_x. Let H_x be a substitute (or criterion) matrix. The precision of $\underline{x}$ is then said to satisfy the criterion if

$$a^* Q_x a \leq a^* H_x a \;\; \forall \; a \in R^n \tag{6.1}$$

or if

$$\frac{a^* Q_x a}{a^* H_x a} \leq 1, \; \forall \; a \in R^n \tag{6.2}$$

since

$$\max_{a \in R^n} \frac{a^* Q_x a}{a^* H_x a} = \lambda_{max} \tag{6.3}$$

where $\lambda_{\max}$ is the maximum eigenvalue of the generalized eigenvalue problem

$$|Q_x - \lambda H_x| = 0 \tag{6.4}$$

it follows that (6.2) is equivalent to

$$\lambda_{max} \leq 1 \tag{6.5}$$

6.2 Invariance of λ

An advantage of the generalized eigenvalue problem approach is that the eigenvalues of (6.4) are *independent* of the chosen S-system. We will prove the following theorem:

Theorem:

The non-zero eigenvalues of $|H_x A^* Q_y^{-1} A - \mu I| = 0$ are the reciprocals of the non-zero eigenvalues of $|Q_x^S - \lambda H_x^S| = 0$, where

$$\begin{aligned} Q_x^S &= S[S^* A^* Q_y^{-1} A S]^{-1} S^*, \quad R^n = R(S) \oplus N(A) \\ H_x^S &= P_{R(S),N(A)} H_x P^*_{R(S),N(A)} \\ P_{R(S),N(A)} &= S(V_0^* S)^{-1} V_0^* = I - V_1[(S^\perp)^* V_1]^{-1}(S^\perp)^* \end{aligned}$$

and

$$R(V_1) = N(A), \quad R(V_0) = N(A)^\perp, \quad R(S^\perp) = R(S)^\perp$$

Proof:

$$|H_x A^* Q_y^{-1} A - \mu I| = |H_x V_0 (S^* V_0)^{-1} S^* A^* Q_y^{-1} A - \mu I| = 0 \tag{6.6}$$

Since

$$\begin{bmatrix} (V_0^* S)^{-1} V_0^* \\ [(S^\perp)^* V_1]^{-1} (S^\perp)^* \end{bmatrix} \begin{bmatrix} S \vdots V_1 \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \tag{6.7}$$

we get

$$\begin{aligned} 0 &= \left| \begin{bmatrix} [(V_0^* S)^{-1} V_0^* H_x V_0 (S^* V_0)^{-1} S^* A^* Q_y^{-1} A S - \mu I] & 0 \\ [(S^\perp)^* V_1]^{-1} (S^\perp)^* H_x V_0 (S^* V_0)^{-1} S^* A^* Q_y^{-1} A S & -\mu I \end{bmatrix} \right| & (6.8) \\ &= \left| [(V_0^* S)^{-1} V_0^* H_x V_0 (S^* V_0)^{-1} S^* A^* Q_y^{-1} A S - \mu I] \right| |-\mu I| \quad \text{for} \quad \mu \neq 0 & (6.9) \end{aligned}$$

This gives with $\lambda = \frac{1}{\mu}$:

$$\begin{aligned} 0 &= \left| [S^* A^* Q_y^{-1} A S]^{-1} - \lambda (V_0^* S)^{-1} V_0^* H_x V_0 (S^* V_0)^{-1} \right| \\ &= \left| \begin{bmatrix} V_0^* S [S^* A^* Q_y^{-1} A S]^{-1} S^* V_0 - \lambda V_0^* H_x V_0 & 0 \\ 0 & \lambda I \end{bmatrix} \right| \\ &= \left| \begin{bmatrix} V_0^* \\ S^{\perp *} \end{bmatrix} \left[S [S^* A^* Q_y^{-1} A S]^{-1} S^* - \lambda S (V_0^* S)^{-1} V_0^* H_x V_0 (S^* V_0)^{-1} S^* \right] \begin{bmatrix} V_0 \vdots S^\perp \end{bmatrix} \right| \\ &= \left| Q_x^S - \lambda H_x^S \right| \end{aligned}$$

6.3 The Alberda-Baarda Substitute Martix

For a two dimensional planar geodetic network, the Alberda-Baarda substitute matrix takes the form

$$H_x = \begin{bmatrix} \begin{bmatrix} d^2 + \Delta d_1^2 & d^2 - d_{12}^2 & \cdots & d^2 - d_{1n}^2 \\ d^2 - d_{21}^2 & d^2 + \Delta d_2^2 & \cdots & d^2 - d_{2n}^2 \\ \vdots & \vdots & \ddots & \vdots \\ d^2 - d_{n1}^2 & d^2 - d_{n2}^2 & \cdots & d^2 + \Delta d_n^2 \end{bmatrix} & 0 \\ 0 & \begin{bmatrix} d^2 + \Delta d_1^2 & d^2 - d_{12}^2 & \cdots & d^2 - d_{1n}^2 \\ d^2 - d_{21}^2 & d^2 + \Delta d_2^2 & \cdots & d^2 - d_{2n}^2 \\ \vdots & \vdots & \ddots & \vdots \\ d^2 - d_{n1}^2 & d^2 - d_{n2}^2 & \cdots & d^2 + \Delta d_n^2 \end{bmatrix} \end{bmatrix}$$

where

- d^2 disappears when H_x is formulated in an S-system
- Δd_i^2 is a parameter per point i
- d_{ij}^2 is a function of the relative positions of points i and j, e.g. $d_{ij}^2 = c_0 + c_1 l_{ij}$

Remark: c_0, c_1, and Δd_i^2 can be considered parameters θ in $H_x(\theta)$.

6.4 A LSQ-Approach for Estimating θ in $H_x(\theta)$

Our objective is to estimate θ such that the difference $Q_x - H_x(\theta)$ is minimal in a least-squares sense. We formulate the following linearized model of observation equations:

$$E\{vec[Q_x - H_x(\theta_0)]\} = vec(\partial_\alpha H_x(\theta_0) \Delta\theta^\alpha) \tag{6.10}$$

Note: both Q_x and $H_x(\theta)$ should be defined in the same S-system.

We will take the *unit matrix* as weight matrix. Then, according to Teunissen [1988] the system of normal equations reads:

$$\boxed{\frac{1}{2}tr[\partial_\beta H_x(\theta_0)]\Delta\hat{\theta}^\beta = \frac{1}{2}tr[\partial_\alpha H_x(\theta_0)\partial_\beta H_x(\theta_0)[Q_x - H_x(\theta_0)]]} \tag{6.11}$$

If the model is *linear*:

$$H_x(\theta) = \sum_\alpha H_\alpha \theta^\alpha \tag{6.12}$$

We get instead of (6.11):

$$\boxed{\frac{1}{2}tr[H_\alpha H_\beta]\hat{\theta}^\beta = \frac{1}{2}tr[H_\alpha Q_x]} \tag{6.13}$$

A disadvantages of the above procedure is that $\hat{\theta}^\beta$ is *not independent of the chosen* S-system.

6.5 Our Proposal

Let $\underline{\hat{x}}$ be the least-squares solution to

$$E\{\underline{y}\} = Ax \quad D\{\underline{y}\} = Q_y \tag{6.14}$$

and define

$$\Delta\underline{x} = \underline{\hat{x}} - E\{\underline{\hat{x}}\} \tag{6.15}$$

Then:

$$E\{\Delta\underline{x}\Delta\underline{x}^*\} = Q_{\hat{x}} \tag{6.16}$$

Although (6.16) holds we will consider the model

$$E\{\Delta\underline{x}\Delta\underline{x}^*\} = H_x(\theta) \tag{6.17}$$

Note: both $Q_{\hat{x}}$ and $H_x(\theta)$ in the same S-system.
This gives after linearization

$$E\{vec[\Delta\underline{x}\Delta\underline{x}^* - H_x(\theta_0)]\} = vec(\partial_\alpha H_x(\theta_0))\Delta\theta^\alpha \tag{6.18}$$

Taking the inverse of (6.16) as weight matrix, application of our theory gives:

$$\frac{1}{2}tr[\partial_\alpha H_x(\theta_0)Q_{\hat{x}}^{-1}\partial_\beta H_x(\theta_0)Q_{\hat{x}}^{-1}]\Delta\underline{\hat{\theta}}^\beta = \frac{1}{2}tr[\partial_\alpha H_x(\theta_0)Q_{\hat{x}}^{-1}[\Delta\underline{x}\Delta\underline{x}^* - H_x(\theta_0)]Q_{\hat{x}}^{-1}] \tag{6.19}$$

Unfortunately this result cannot be used since $\Delta\underline{x}\Delta\underline{x}^*$ is *unknown* in general. However, its expectation $E\{\Delta\underline{x}\Delta\underline{x}^*\}$ is known [see (6.16)]. We therefore propose to replace $\Delta\underline{x}\Delta\underline{x}^*$ in (6.19) by its expectation $Q_{\hat{x}}$ [This is not an unusual procedure; think of eccentrincity errors and Kalman filtering]. We then get instead of (6.19):

$$\boxed{\frac{1}{2}tr[\partial_\alpha H_x(\theta_0)Q_{\hat{x}}^{-1}\partial_\beta H_x(\theta_0)Q_{\hat{x}}^{-1}]\Delta\underline{\hat{\theta}}^\beta = \frac{1}{2}tr[\partial_\alpha H_x(\theta_0)Q_{\hat{x}}^{-1}[Q_{\hat{x}} - H_x(\theta_0)]Q_{\hat{x}}^{-1}]} \tag{6.20}$$

If the model is *linear*:

$$H_x(\theta) = \sum_\alpha H_\alpha \theta^\alpha \tag{6.21}$$

we get instead of (6.20):

$$\boxed{\frac{1}{2}tr[H_\alpha Q_{\hat{x}}^{-1} H_\beta Q_{\hat{x}}^{-1}]\hat{\theta}^\beta = \frac{1}{2}tr[H_\alpha Q_{\hat{x}}^{-1}]} \tag{6.22}$$

Compare this results with (6.13). The above described method has the following advantages:

- All the least-squares diagnostics can be applied.
- No generalized eigenvalue problem needs to be solved for.
- The estimate $\hat{\theta}^{\beta}$ is *independent* of the chosen S-system. This can be seen as follows. Let R be a square and regular matrix. Then:

$$tr[RH_\alpha R^*[RQ_{\hat{x}}R^*]^{-1}RH_\beta R^*[RQ_{\hat{x}}R^*]^{-1}] = tr[H_\alpha Q_{\hat{x}}^{-1}H_\beta Q_{\hat{x}}^{-1}]$$

$$tr[RH_\alpha R^*[RQ_{\hat{x}}R^*]^{-1}] = tr[H_\alpha Q_{\hat{x}}^{-1}]$$

The normal matrix $\frac{1}{2}tr[H_\alpha NH_\beta N]$ is *singular* if and only if there exist $x^\alpha,\ \alpha = 1,2,...,n$ such that:

$$\frac{1}{2}x^\alpha tr[H_\alpha NH_\beta N]x^\beta = 0 \tag{6.23}$$

or

$$\frac{1}{2}tr[H_\alpha x^\alpha NH_\beta x^\beta N] = 0 \tag{6.24}$$

If $\lambda_i,\ i = 1,2,...,n$ are the eigenvalues of $H_\alpha x^\alpha N$ then

$$\frac{1}{2}tr[H_\alpha x^\alpha NH_\beta x^\beta N] = \sum_{i=1}^{n}\lambda_i^2 \tag{6.25}$$

Hence, (6.24) can only be true if

$$\lambda_i = 0\ \ \forall\ i = 1,2,...,n \tag{6.26}$$

or if

$$\boxed{H_\alpha x^\alpha N = 0} \tag{6.27}$$

For instance: $\exists\ x^\alpha$ such that $R(H_\alpha x^\alpha) \subset N(N)$

Example A closed levelling loop with 3 observations.

$$N = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}\quad \Delta d^2 H_1 = \begin{pmatrix} \Delta d^2 & 0 & 0 \\ 0 & \Delta d^2 & 0 \\ 0 & 0 & \Delta d^2 \end{pmatrix} c_1H_2 = \begin{pmatrix} 0 & -c_1\ell & -c_1\ell \\ -c_1\ell & 0 & -c_1\ell \\ -c_1\ell & -c_1\ell & 0 \end{pmatrix} \tag{6.28}$$

Remark: This singularity *does not* occur in the set up of 6.4.

No S-transformation needs to be applied a priori. This can be seen as follows: Substitution of

$$H_\alpha := (V_0^*S)^{-1}V_0^*H_\alpha V_0(S^*V_0)^{-1}\ \text{ and }\ Q_{\hat{x}}^{-1} := S^*A^*Q_y^{-1}AS \tag{6.29}$$

into (15) gives

$$\frac{1}{2}tr[(V_0^*S)^{-1}V_0^*H_\alpha V_0(S^*V_0)^{-1}.S^*A^*Q_y^{-1}AS(V_0^*S)^{-1}V_0^*H_\beta V_0(S^*V_0)^{-1}.S^*A^*Q_y^{-1}AS]\hat{\theta}_\beta = \frac{1}{2}tr[(V_0^*S)^{-1}V_0^*H_\alpha V_0(S^*V_0)^{-1}.S^*A^*Q_y^{-1}AS] \tag{6.30}$$

or

$$\frac{1}{2}tr[(V_0^*S)^{-1}V_0^*H_\alpha A^*Q_y^{-1}AH_\beta A^*Q_y^{-1}AS]\hat{\theta}_\beta = \frac{1}{2}tr[(V_0^*S)^{-1}V_0^*H_\alpha A^*Q_y^{-1}AS] \tag{6.31}$$

or

$$\boxed{\boxed{\frac{1}{2}tr[H_\alpha A^*Q_y^{-1}AH_\beta A^*Q_y^{-1}A]\hat{\theta}_\beta = \frac{1}{2}tr[H_\alpha A^*Q_y^{-1}A]}} \tag{6.32}$$

Note that $Q_{\hat{x}}$ is not needed explicitly. Only the (reduced) normal matrix $A^*Q_y^{-1}A$ is needed.

Remark: It has been assumed that the normal matrix is invertible.

Remark: In practice one will have $H_x(\theta) = H_0 + \sum H_\alpha \theta^\alpha$ instead of $H_x(\theta) = \sum H_\alpha \theta^\alpha$.

Remark: In the special case that $H_x(\theta) = H\theta$ we get

$$\boxed{\hat{\theta} = \frac{\sum_i \lambda_i}{\sum_i \lambda_i^2}} \tag{6.33}$$

If $\hat{\theta} = 1$ then the precision test is accepted. with:

$$|HA^* Q_y^{-1} A - \lambda_i I| = 0 \tag{6.34}$$

If we write this as $|(A^* Q_y^{-1} A)^- - \mu_i H| = 0$, (6.33) becomes

$$\hat{\theta} = \frac{\sum_i \frac{1}{\mu_i}}{\sum_i \frac{1}{\mu_i^2}} \tag{6.35}$$

From this follows that:

$$\hat{\theta} = \frac{\frac{1}{\mu_{max}}[\sum_i \frac{\mu_{max}}{\mu_i}]}{\frac{1}{\mu_{max}^2}[\sum_i \frac{\mu_{max}^2}{\mu_i^2}]} < \mu_{max} \tag{6.36}$$

since

$$[\sum_i \frac{\mu_{max}^2}{\mu_i^2}] > [\sum_i \frac{\mu_{max}}{\mu_i}] \tag{6.37}$$

Thus:

$$\boxed{\hat{\theta} < \mu_{max}} \tag{6.38}$$

Remark: With (6.32) one can still study partial networks instead of the total network. In this case one needs the reduced normal matrix.

Remark: If one considers instead of $\Delta \underline{x}$ the (estimable) linear functions $B^* \Delta \underline{x}$, then (6.22) should be replaced by:

$$\frac{1}{2} tr[B^* H_\alpha B [B^* Q_{\hat{x}} B]^{-1} B^* H_\beta B [B^* Q_{\hat{x}} B]^{-1}] \hat{\theta}^\beta = \frac{1}{2} tr[B^* H_\alpha B [B^* Q_{\hat{x}} B]^{-1}] \tag{6.39}$$

6.6 On the Teststatistics $\underline{T}_{m-n}$ & $\underline{w}$

According to previous section, the solution of the minimization problem

$$\min_\theta T(\theta) \tag{6.40}$$

with

$$T(\theta) = [vec(Q_x - H_\alpha \theta^\alpha)]^* Q_x^{-1} \otimes Q_x^{-1} [vec(Q_x - H_\alpha \theta^\alpha)] \tag{6.41}$$

is given by

$$\boxed{\hat{\theta}^\beta = [tr(H_\alpha Q_x^{-1} H_\beta Q_x^{-1})]^{-1} tr(H_\alpha Q_x^{-1})} \tag{6.42}$$

If we use the notation

$$\hat{H} = H_\alpha \hat{\theta}^\alpha \tag{6.43}$$

it follows with (6.41) that:

$$T(\hat{\theta}) = [vec(Q_x - \hat{H})]^* Q_x^{-1} \otimes Q_x^{-1} [vec(Q_x - \hat{H})] \tag{6.44}$$

with the property:

$$tr(ABCD) = vec(D)^* (A \otimes C^*) vec(B^*) \tag{6.45}$$

we may write (6.44) as

$$T(\hat{\theta}) = tr[Q_x^{-1}(Q_x - \hat{H}) Q_x^{-1}(Q_x - \hat{H})] \tag{6.46}$$

or as

$$\boxed{T(\hat{\theta}) = tr[(I - Q_x^{-1}\hat{H})(I - Q_x^{-1}\hat{H})] = \sum_{i=1}^{n} (1 - \hat{\lambda}_i)^2} \tag{6.47}$$

where $\hat{\lambda}_i, \; i = 1, 2, ..., n$ are the eigenvalues of

$$|Q_x - \hat{\lambda}_i \hat{H}| = 0 \tag{6.48}$$

Although expression (6.47) looks already rather simple, it can be simplified still a bit further. From (6.47) follows that

$$T(\hat{\theta}) = tr[I_n - 2Q_x^{-1}\hat{H} + Q_x^{-1}\hat{H}Q_x^{-1}\hat{H}] = n - 2tr(Q_x^{-1}\hat{H}) + tr(Q_x^{-1}\hat{H}Q_x^{-1}\hat{H}) \tag{6.49}$$

Substitution of (6.43) gives

$$T(\hat{\theta}) = n - 2tr(Q_x^{-1}H_\alpha)\hat{\theta}^\alpha + \hat{\theta}^\alpha tr(Q_x^{-1}H_\alpha Q_x^{-1}H_\beta)\hat{\theta}^\beta \tag{6.50}$$

But according to (6.42):

$$tr(Q_x^{-1}H_\alpha Q_x^{-1}H_\beta)\hat{\theta}^\beta = tr(H_\alpha Q_x^{-1}) \tag{6.51}$$

Substitution of (6.51) into (6.50) gives:

$$T(\hat{\theta}) = n - 2tr(Q_x^{-1}H_\alpha)\hat{\theta}^\alpha + tr(H_\alpha Q_x^{-1})\hat{\theta}^\alpha \tag{6.52}$$

From this follows that:

$$\begin{aligned} T(\hat{\theta}) &= n - tr(Q_x^{-1}H_\alpha)\hat{\theta}^\alpha \\ &= n - tr(Q_x^{-1}\hat{H}) \qquad (6.53) \\ &= n - \sum_{i=1}^{n} \hat{\lambda}_i \qquad (6.54) \end{aligned}$$

Remark: For leveling networks one should take n of (6.54) equal to $n-1$, and for 2D planer networks one should take n of (6.54) equal to $2n-4$.

Remark: In the special case that $H(\theta) = H\theta$, we have $\hat{\theta} = \sum_i \lambda_i / \sum_i \lambda_i^2$, with $|Q_x^{-1}H - \lambda I| = 0$, and therefore

$$T(\hat{\theta}) = n - \frac{(\sum_i \lambda_i)^2}{(\sum_i \lambda_i^2)} \tag{6.55}$$

Remark: Note that we may write

$$tr[Q_x^{-1}\hat{H}] = tr[A^* Q_y^{-1} A\hat{H}] = tr[A\hat{H}A^* Q_y^{-1}] \tag{6.56}$$

If $\hat{H}$ is *close* to Q_x, then $A\hat{H}A^*Q_y^{-1}$ is *close* to P_A, and we know that $tr(P_A) = n$.

Remark: A comparison of (6.47) and (6.54) shows that $\sum_{i=1}^{n} \hat{\lambda}_i = \sum_{i=1}^{n} \hat{\lambda}_i^2$ or $tr(Q_x^{-1}\hat{H}) = tr(Q_x^{-1}\hat{H}Q_x^{-1}\hat{H})$.

Remark: Expression (6.47) [but not (6.54)] can probably be used for *testing the precision.* Let $H = H_\alpha \theta^\alpha$ be the criterion matrix with $\alpha = 1, 2, ..., p$. The following criterion may then be useful:

$$\frac{tr[(I - Q_x^{-1}H)(I - Q_x^{-1}H)]}{n(n+1)/2 - p} \doteq 1 \tag{6.57}$$

We will now derive the equivalent of the w-teststatistic. We have

$$\begin{aligned} c_{y_i}^* Q_y^{-1} Q_{\hat{e}} Q_y^{-1} c_{y_i} = [vec(e_i e_j^* + e_j e_i^*)]^* [Q_x^{-1} \otimes Q_x^{-1} - (Q_x^{-1} \otimes Q_x^{-1}) vec(H_\alpha) \\ (tr[Q_x^{-1} H_\beta Q_x^{-1} H_\alpha])^{-1} vec(H_\beta)^* (Q_x^{-1} \otimes Q_x^{-1})][vec(e_i e_j^* + e_j e_i^*)] \end{aligned} \tag{6.58}$$

or

$$\begin{aligned} c_{y_i}^* Q_y^{-1} Q_{\hat{e}} Q_y^{-1} c_{y_i} = tr[Q_x^{-1}(e_i e_j^* + e_j e_i^*) Q_x^{-1}(e_i e_j^* + e_j e_i^*)] - tr[Q_x^{-1} H_\alpha Q_x^{-1}(e_i e_j^* + e_j e_i^*)] \\ [tr(Q_x^{-1} H_\beta Q_x^{-1} H_\alpha)]^{-1} tr[Q_x^{-1}(e_i e_j^* + e_j e_i^*) Q_x^{-1} H_\beta] \end{aligned} \tag{6.59}$$

or

$$\begin{aligned} c_{y_i}^* Q_y^{-1} Q_{\hat{e}} Q_y^{-1} c_{y_i} = 2(e_j^* Q_x^{-1} e_i)^2 + 2 e_i^* Q_x^{-1} e_i e_j^* Q_x^{-1} e_j - \\ 4\,[e_j^* Q_x^{-1} H_\alpha Q_x^{-1} e_i][tr(Q_x^{-1} H_\beta Q_x^{-1} H_\alpha)]^{-1} [e_j^* Q_x^{-1} H_\beta Q_x^{-1} e_i] \end{aligned} \tag{6.60}$$

and

$$\begin{aligned} c_{y_i}^* Q_y^{-1} \hat{e} &= [vec(e_i e_j^* + e_j e_i^*)]^* Q_x^{-1} \otimes Q_x^{-1} [vec(Q_x - H_\alpha \hat{\theta}^\alpha)] \\ &= tr[Q_x^{-1}(Q_x - H_\alpha \hat{\theta}^\alpha) Q_x^{-1}(e_i e_j^* + e_j e_i^*)] \\ &= 2 e_j^* (I - H_\alpha \hat{\theta}^\alpha Q_x^{-1}) e_i \\ &= 2\delta_{ij} - 2 e_i^* Q_x^{-1} H_\alpha \hat{\theta}^\alpha e_j \end{aligned} \tag{6.61}$$

From (6.60) and (6.62) follows that:

$$\boxed{w = \frac{\delta_{ij} - e_i^* Q_x^{-1} H_\alpha \hat{\theta}^\alpha e_j}{s}} \tag{6.62}$$

with

$$\begin{aligned} s = [(e_j^* Q_x^{-1} e_i)^2 + e_i^* Q_x^{-1} e_i e_j^* Q_x^{-1} e_j - 2[e_j^* Q_x^{-1} H_\alpha Q_x^{-1} e_i] \\ [tr(Q_x^{-1} H_\beta Q_x^{-1} H_\alpha)]^{-1} [e_j^* Q_x^{-1} H_\beta Q_x^{-1} e_i]]^{1/2} \end{aligned} \tag{6.63}$$

Remark: This result can possibly be used for testing whether *individual elements* of Q_x are close enough to the corresponding elements of $H_\alpha \hat{\theta}^\alpha$.

Wrong! The correct answer is:

$$\boxed{w = \frac{a^*[Q_x^{-1} - Q_x^{-1}\hat{H}Q_x^{-1}]a}{[(a^* Q_x^{-1} a)^2 - (a^* Q_x^{-1} H_\alpha Q_x^{-1} a)(tr[Q_x^{-1} H_\beta Q_x^{-1} H_\alpha])^{-1}(a^* Q_x^{-1} H_\beta Q_x^{-1} a)]^{1/2}}} \tag{6.64}$$

or with $\bar{a} = Q_x^{-1} a$:

$$\boxed{w = \frac{\bar{a}^*[Q_x - \hat{H}]\bar{a}}{[(\bar{a}^* Q_x \bar{a})^2 - (\bar{a}^* H_\alpha \bar{a})(tr[Q_x^{-1} H_\beta Q_x^{-1} H_\alpha])^{-1}(\bar{a}^* H_\beta \bar{a})]^{1/2}}} \tag{6.65}$$

Note that the equation (6.65) can also be written as:

$$w = \frac{\frac{\bar{a}^* Q_x \bar{a}}{\bar{a}^* \hat{H} \bar{a}} - 1}{\left[(\frac{\bar{a}^* Q_x \bar{a}}{\bar{a}^* \hat{H} \bar{a}})^2 - \frac{\bar{a}^* H_\alpha \bar{a}}{\bar{a}^* \hat{H} \bar{a}} (tr[Q_x^{-1} H_\beta Q_x^{-1} H_\alpha])^{-1} \frac{\bar{a}^* H_\alpha \bar{a}}{\bar{a}^* \hat{H} \bar{a}} \right]^{1/2}} \tag{6.66}$$

If we use the *approximation*

$$w = \frac{c_y^* Q_y^{-1} \hat{e}}{(c_y^* Q_y^{-1} Q_{\hat{e}} Q_y^{-1} c_y})^{1/2} \doteq \frac{c_y^* Q_y^{-1} \hat{e}}{(c_y^* Q_y^{-1} c_y})^{1/2} \tag{6.67}$$

Then equation (6.66) reduces to

$$w \doteq 1 - \frac{\bar{a}^* \hat{H} \bar{a}}{\bar{a}^* Q_x \bar{a}} \tag{6.68}$$

Hence:

$$w_{\max} \doteq 1 - \frac{1}{\lambda_{\max}} \tag{6.69}$$

6.7 On the Choice of a Scaled Unit Weight Matrix

According to previous subsections, the following holds:

$$\frac{1}{2} tr[H_\alpha Q_x^{-1} H_\beta Q_x^{-1}] \hat{\underline{\theta}}^\beta = \frac{1}{2} tr[H_\alpha Q_x^{-1} \Delta \underline{x} \Delta \underline{x}^* Q_x^{-1}] \tag{6.70}$$

where both H and Q_x are in the same S-system *excluding* the S-basis. Thus for an H and a Q_x in an arbitrary S-system *including* the S-basis, equation (6.70) should be read as:

$$\begin{aligned} &\frac{1}{2} tr[(V_0^* S)^{-1} V_0^* H_\alpha V_0 (S^* V_0)^{-1} . S^* A^* Q_y^{-1} AS . (V_0^* S)^{-1} V_0^* H_\beta V_0 (S^* V_0)^{-1} . S^* A^* Q_y^{-1} AS] \hat{\underline{\theta}}^\beta \\ &= \frac{1}{2} tr[(V_0^* S)^{-1} V_0^* H_\alpha V_0 (S^* V_0)^{-1} . S^* A^* Q_y^{-1} AS . \Delta \underline{x} \Delta \underline{x}^* . S^* A^* Q_y^{-1} AS] \end{aligned} \tag{6.71}$$

In this equation matrix $S^* A^* Q_y^{-1} AS$ plays the role of weight matrix. Thus $(S^* A^* Q_y^{-1} AS)^{-1}$ plays the role of variance matrix. We will now investigate the consequences if one replaces or approximates $(S^* A^* Q_y^{-1} AS)^{-1}$ by

$$[(V_0^* S)^{-1} V_0^* \sigma^2 I V_0 (S^* V_0)^{-1}] \tag{6.72}$$

If we replace $S^* A^* Q_y^{-1} AS$ in (6.71) by the inverse of (6.72) we get:

$$\begin{aligned} &\frac{1}{2} tr[(V_0^* S)^{-1} V_0^* H_\alpha V_0 (S^* V_0)^{-1} . (S^* V_0)(\sigma^2 V_0^* V_0)^{-1} (V_0^* S) ... \\ &... (V_0^* S)^{-1} V_0^* H_\beta V_0 (S^* V_0)^{-1} . (S^* V_0)(\sigma^2 V_0^* V_0)^{-1} (V_0^* S)] \hat{\underline{\theta}}^\beta = \\ &\frac{1}{2} tr[(V_0^* S)^{-1} V_0^* H_\alpha V_0 (S^* V_0)^{-1} . (S^* V_0)(\sigma^2 V_0^* V_0)^{-1} (V_0^* S) . \Delta \underline{x} \Delta \underline{x}^* . (S^* V_0)(\sigma^2 V_0^* V_0)^{-1} (V_0^* S)] \end{aligned} \tag{6.73}$$

or

$$\begin{aligned} &\frac{1}{2} \sigma^{-4} tr[V_0^* H_\alpha V_0 (V_0^* V_0)^{-1} V_0^* H_\beta V_0 (V_0^* V_0)^{-1}] \hat{\underline{\theta}}^\beta = \\ &\frac{1}{2} \sigma^{-4} tr[V_0^* H_\alpha V_0 (V_0^* V_0)^{-1} V_0^* S . \Delta \underline{x} \Delta \underline{x}^* S^* V_0 (V_0^* V_0)^{-1}] \end{aligned} \tag{6.74}$$

Note that the projector

$$P_{R(A^*),N(A)} = V_0 (V_0^* V_0)^{-1} V_0^* = I - V_1 (V_1^* V_1)^{-1} V_1^* \tag{6.75}$$

is the S-transformation that corresponds with the *minimum-norm* solution. With (6.75), equation (6.74) may be written as

$$tr[H_\alpha P H_\beta P] \hat{\underline{\theta}}^\beta = tr[H_\alpha P S . \Delta \underline{x} \Delta \underline{x}^* S^* P] \tag{6.76}$$

If we replace $\Delta \underline{x} \Delta \underline{x}^*$ by its expectation $(S^* A^* Q_y^{-1} AS)^{-1}$, we get

$$tr[H_\alpha P H_\beta P] \hat{\underline{\theta}}^\beta = tr[H_\alpha P S (S^* A^* Q_y^{-1} AS)^{-1} S^* P] \tag{6.77}$$

or with $Q_x = S(S^*A^*Q_y^{-1}AS)^{-1}S^*$,

$$\boxed{tr[H_\alpha P H_\beta P]\underline{\hat{\theta}}^\beta = tr[H_\alpha P Q_x P]} \tag{6.78}$$

This shows that:

1. The solution is independent of the chosen S-system, because of the occurrence of P in (6.78),
2. The solution corresponds to the case that H and Q_x are defined in the *minimum-norm* S-system.

Remark: Because of the structure of the *normal* matrix in (6.78), it will be possible for some cases to solve (6.78) *analytically*. For example consider a leveling network of n points with the simplest substitute matrix H. Then

$$H = \Delta d^2 I_n, \quad \text{and} \quad P = I_n - e(e^*e)^{-1}e^* \tag{6.79}$$

with

$$e = (1, 1, ..., 1)^* \tag{6.80}$$

Substitution of (6.79) into (6.78) gives

$$tr[P]\Delta\hat{d}^2 = tr[PQ_xP] \tag{6.81}$$

or

$$\boxed{\begin{array}{l} \Delta\hat{d}^2 = \dfrac{tr[PQ_xP]}{n-1} \\ \\ \sigma^2_{\Delta\hat{d}^2} = \dfrac{2\sigma^4}{n-1} \end{array}} \tag{6.82}$$

Note 1: Here we have a link with the *minimum trace*.

Note 2: This $\Delta\hat{d}^2$ is related to *Helmert's mittler punktfehler*.
We will now derive the with (6.78) corresponding *teststatistic* T_{m-n}. We have

$$\begin{aligned} T(\hat{\theta}) = {} & [vec((V_0^*S)^{-1}V_0^*(Q_x - \hat{H})V_0(S^*V_0)^{-1})]^*[(V_0^*S)^{-1}V_0^*\sigma^2 IV_0(S^*V_0)^{-1}]^{-1} \otimes \\ & ...[(V_0^*S)^{-1}V_0^*\sigma^2 IV_0(S^*V_0)^{-1}]^{-1}[vec((V_0^*S)^{-1}V_0^*(Q_x - \hat{H})V_0(S^*V_0)^{-1}] \end{aligned} \tag{6.83}$$

or

$$\begin{aligned} T(\hat{\theta}) = tr \Big\{ & [(V_0^*S)^{-1}V_0^*\sigma^2 IV_0(S^*V_0)^{-1}]^{-1}[(V_0^*S)^{-1}V_0^*(Q_x - \hat{H})V_0(S^*V_0)^{-1}] \\ & ...\ [(V_0^*S)^{-1}V_0^*\sigma^2 V_0(S^*V_0)^{-1}]^{-1}[(V_0^*S)^{-1}V_0^*(Q_x - \hat{H})V_0(S^*V_0)^{-1}] \Big\} \end{aligned} \tag{6.84}$$

or

$$\boxed{T(\hat{\theta}) = \sigma^{-4} tr[P(Q_x - \hat{H})P(Q_x - \hat{H})]} \tag{6.85}$$

We may write (6.85) also as

$$\begin{aligned} T(\hat{\theta}) &= \sigma^{-4}\Big\{tr[PQ_xPQ_x] - tr[PQ_xP\hat{H}] - tr[P\hat{H}PQ_x] + tr[P\hat{H}P\hat{H}]\Big\} \\ &= \sigma^{-4}\Big\{tr[PQ_xPQ_x] - 2tr[PQ_xP\hat{H}] + tr[P\hat{H}P\hat{H}]\Big\} \end{aligned} \tag{6.86}$$

And with (6.78), this simplifies to

$$\boxed{T(\hat{\theta}) = \sigma^{-4} tr[PQ_x(PQ_x - P\hat{H})]} \tag{6.87}$$

We will now derive with (6.78) the corresponding *teststatistic* $\underline{w}$: Let us assume that we want to test how well the variance $a^*Q_x a$ of an estimable function a^*x fits the *model* value $a^*\hat{H}a$. Then, the following two *hypotheses* should be considered:

$$\begin{cases} H_0: & E\{vec(\Delta\underline{x}\Delta\underline{x}^*)\} = vec(H_\alpha)\theta^\alpha \\ \\ H_A: & E\{vec(\Delta\underline{x}\Delta\underline{x}^*)\} = vec(H_\alpha)\theta^\alpha + vec(aa^*) \end{cases} \tag{6.88}$$

These *hypotheses* should actually be read as:

$$\begin{cases} H_0: & E\{vec[S^*A^*Q_y^{-1}AS]\} &=& vec[(V_0^*S)^{-1}V_0^*H_\alpha V_0(S^*V_0)^{-1}]\theta^\alpha \\ \\ H_A: & E\{vec[S^*A^*Q_y^{-1}AS]\} &=& vec[(V_0^*S)^{-1}V_0^*H_\alpha V_0(S^*V_0)^{-1}]\theta^\alpha \\ & &+& vec[(V_0^*S)^{-1}V_0^*aa^*V_0(S^*V_0)^{-1}] \end{cases} \tag{6.89}$$

As covariance matrix, we take

$$[(V_0^*S)^{-1}V_0^*\sigma^2 I V_0(S^*V_0)^{-1}] \otimes [(V_0^*S)^{-1}V_0^*\sigma^2 I V_0(S^*V_0)^{-1}] \tag{6.90}$$

with (6.89) and (6.90), it follows that:

$$\begin{aligned} c_y^*Q_y^{-1}Q_{\hat{e}}Q_y^{-1}c_y = vec[(V_0^*S)^{-1}V_0^*aa^*V_0(S^*V_0)^{-1}]^*[[(V_0^*S)^{-1}V_0^*\sigma^2 V_0(S^*V_0)^{-1}]^{-1} \otimes \\ [(V_0^*S)^{-1}V_0^*\sigma^2 V_0(S^*V_0)^{-1}]^{-1} - [(V_0^*S)^{-1}V_0^*\sigma^2 V_0(S^*V_0)^{-1}]^{-1} \otimes [(V_0^*S)^{-1}V_0^*\sigma^2 V_0(S^*V_0)^{-1}]^{-1} \\ vec[(V_0^*S)^{-1}V_0^*H_\alpha V_0(S^*V_0)^{-1}][tr(H_\alpha PH_\beta P)]^{-1}vec[(V_0^*S)^{-1}V_0^*H_\beta V_0(S^*V_0)^{-1}]^* \\ [(V_0^*S)^{-1}V_0^*\sigma^2 V_0(S^*V_0)^{-1}]^{-1} \otimes [(V_0^*S)^{-1}V_0^*\sigma^2 V_0(S^*V_0)^{-1}]^{-1}]vec[(V_0^*S)^{-1}V_0^*aa^*V_0(S^*V_0)^{-1}] \end{aligned}$$

or

$$\begin{aligned} c_y^*Q_y^{-1}Q_{\hat{e}}Q_y^{-1}c_y = \frac{1}{2}tr\left\{(S^*V_0\sigma^{-2}(V_0^*V_0)^{-1}V_0^*S)(V_0^*S)^{-1}V_0^*aa^*V_0(S^*V_0)^{-1} \ldots\right. \\ \left.(S^*V_0)\sigma^{-2}(V_0^*V_0)^{-1}V_0^*S)(V_0^*S)^{-1}V_0^*aa^*V_0(S^*V_0)^{-1}\right\} \\ -\frac{1}{2}tr\left\{(S^*V_0)\sigma^{-2}(V_0^*V_0)^{-1}V_0^*S.(V_0^*S)^{-1}V_0^*H_\alpha V_0(S^*V_0)^{-1} \ldots\right. \\ \left.(S^*V_0)\sigma^{-2}(V_0^*V_0)^{-1}V_0^*S.(V_0^*S)^{-1}V_0^*aa^*V_0(S^*V_0)^{-1}\right\} 2\sigma^4[tr(H_\alpha PH_\beta P)]^{-1}\ldots \\ \frac{1}{2}tr\left\{(S^*V_0)\sigma^{-2}(V_0^*V_0)^{-1}V_0^*S.(V_0^*S)^{-1}V_0^*aa^*V_0(S^*V_0)^{-1} \ldots\right. \\ \left.(S^*V_0)\sigma^{-2}(V_0^*V_0)^{-1}V_0^*S.(V_0^*S)^{-1}V_0^*H_\beta V_0(S^*V_0)^{-1}\right\} \end{aligned} \tag{6.91}$$

or

$$\begin{aligned} c_y^*Q_y^{-1}Q_{\hat{e}}Q_y^{-1}c_y &= \frac{1}{2}\sigma^{-4}(a^*Pa)^2 - \frac{1}{2}\sigma^{-4}(a^*PH_\alpha Pa)2\sigma^4(tr[H_\alpha PH_\beta P])^{-1}\frac{1}{2}\sigma^{-4}(a^*PH_\beta Pa) \\ &= \frac{1}{2}\sigma^{-4}\left\{(a^*a)^2 - (a^*H_\alpha a)(tr[H_\alpha PH_\beta P])^{-1}(a^*H_\beta a)\right\} \end{aligned} \tag{6.92}$$

since a is an *estimable function*, i.e., $Pa = a$.

$$\begin{aligned} c_y^*Q_y^{-1}\hat{e} = \frac{1}{2}vec[(V_0^*S)^{-1}V_0^*aa^*V_0(S^*V_0)^{-1}]^*[(S^*V_0)\sigma^{-2}(V_0^*V_0)^{-1}(V_0^*S)] \otimes \\ [(S^*V_0)\sigma^{-2}(V_0^*V_0)^{-1}(V_0^*S)][vec[(V_0^*S)^{-1}V_0^*(Q_x - \hat{H})V_0(S^*V_0)^{-1}]] \end{aligned} \tag{6.93}$$

or

$$\begin{aligned} c_y^*Q_y^{-1}\hat{e} = \frac{1}{2}tr[(S^*V_0)(V_0^*V_0)^{-1}(V_0^*S).(V_0^*S)^{-1}V_0^*(Q_x - \hat{H})V_0(S^*V_0)^{-1}. \\ (S^*V_0)(V_0^*V_0)^{-1}(V_0^*S).(V_0^*S)^{-1}V_0^*aa^*V_0(S^*V_0)^{-1}]\sigma^{-4} \end{aligned} \tag{6.94}$$

or

$$c_y^* Q_y^{-1} \hat{e} = \frac{1}{2}\sigma^{-4} a^* P(Q_x - \hat{H})Pa = \frac{1}{2}\sigma^{-4} a^*(Q_x - \hat{H})a \tag{6.95}$$

From (6.92) and (6.95) follows that:

$$\boxed{w = \frac{1}{\sqrt{2}\sigma^2}.\frac{a^*(Q_x - \hat{H})a}{[(a^*a)^2 - (a^*H_\alpha a)(tr[H_\alpha P H_\beta P])^{-1}(a^* H_\beta a)]^{1/2}}} \tag{6.96}$$

Example: If $H = \Delta d^2 I_n$ and $P = I_n - e(e^*e)e^*$, then

$$\boxed{w = \frac{1}{\sqrt{2}\sigma^2}.\frac{a^*(Q_x - \hat{H})a}{a^*a}.\left(\frac{n-1}{n-2}\right)^{1/2}} \tag{6.97}$$

Note that $w_{\max}$ in this case is related to the generalized eigenvalue problem $|Q_x - \lambda\hat{H}| = 0$. Hence, we have established a like with the ordinary procedure of the generalized eigenvalue problem

$$w_{\max} = \frac{1}{\sqrt{2}\sigma^2}.\left(\frac{n-1}{n-2}\right)^{1/2} (\lambda_{\max} - 1)\Delta\hat{d}^2 \tag{6.98}$$

Appendix A

Backgrounds

A.1 The Moments of $\underline{t} \sim N(0, Q_t)$

The *moment generating function* of $\underline{t}$ is defined as

$$\Phi(s) = E\{\exp[s^*\underline{t}]\} = \int p_{\underline{t}}(t)\exp[s^*t]dt, \tag{A.1}$$

with

$$p_{\underline{t}}(t) = (2\pi)^{-1/2}|Q_t|^{-1/2}\exp[-\frac{1}{2}t^*Q_t^{-1}t] \tag{A.2}$$

Substitution of (A.2) into (A.1) gives

$$\Phi(s) = \int (2\pi)^{-1/2}|Q_t|^{-1/2}\exp[-\frac{1}{2}\{t^*Q_t^{-1}t - 2s^*t\}]dt, \tag{A.3}$$

Substitution of

$$\exp[-\frac{1}{2}\{t^*Q_t^{-1}t - 2s^*t\}] = \exp[-\frac{1}{2}(t - Q_ts)^*Q_t^{-1}(t - Q_ts)]\exp[-\frac{1}{2}s^*Q_ts] \tag{A.4}$$

into (A.3) gives

$$\boxed{\Phi(s) = \exp[\frac{1}{2}s^*Q_ts]} \tag{A.5}$$

From (A.1) follows that:

$$\partial^\alpha_{i_1,\ldots,i_\alpha}\Phi(s) = \int p_{\underline{t}}(t)\partial^\alpha_{i_1,\ldots,i_\alpha}[\exp(s^*t)]dt, \quad i_1,...,i_\alpha = 1,2,...,b \tag{A.6}$$

Substitution of

$$\partial^\alpha_{i_1,\ldots,i_\alpha}[\exp(s^*t)] = t_{i_1}t_{i_2}...t_{i_\alpha}\exp(s^*t) \tag{A.7}$$

into (A.6) gives:

$$\partial^\alpha_{i_1,\ldots,i_\alpha}\Phi(s) = \int p_{\underline{t}}(t)t_{i_1}t_{i_2}...t_{i_\alpha}\exp(s^*t)dt \tag{A.8}$$

Evaluation at $s = 0$, shows that

$$\partial^\alpha_{i_1,\ldots,i_\alpha}\Phi(s)\big|_{s=0} = E\{\underline{t}_{i_1}\underline{t}_{i_2}...\underline{t}_{i_\alpha}\} \tag{A.9}$$

If we write (A.5) in index notation like

$$\Phi(s) = \exp[\frac{1}{2}q^{ij}s_is_j], \quad i,j = 1,2,...,b \tag{A.10}$$

it follows that

$$\begin{array}{rclcrcl} \partial_i \Phi(s)\big|_{s=0} & = & 0, & \partial^2_{ij}\Phi(s)\big|_{s=0} & = & q^{ij} \\ \partial^3_{ijk}\Phi(s)\Big|_{s=0} & = & 0, & \partial^4_{ijkl}\Phi(s)\Big|_{s=0} & = & q^{ij}q^{kl}+q^{ik}q^{jl}+q^{il}q^{jk} \end{array} \tag{A.11}$$

This, together with (A.9) shows that

$$\boxed{\begin{array}{rclcrcl} E\{\underline{t}_i\} & = & 0, & E\{\underline{t}_i\underline{t}_j\} & = & q^{ij} \\ E\{\underline{t}_i\underline{t}_j\underline{t}_k\} & = & 0, & E\{\underline{t}_i\underline{t}_j\underline{t}_k\underline{t}_l\} & = & q^{ij}q^{kl}+q^{ik}q^{jl}+q^{il}q^{jk} \end{array}} \tag{A.12}$$

A.2 The Variance Matrix of $vec(\underline{t}\underline{t}^*)$, with $\underline{t} \sim N(0, Q_t)$

First some standard results on the vec-operator and the Kronecker product. Consider a matrix $A = [a_{ij}]$ of order $m \times n$ and a matrix $B = [b_{ij}]$ of order $r \times s$. The Kronecker product of the two matrices, denoted by $A \otimes B$ is defined as the partitioned matrix

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix} \tag{A.13}$$

$A \otimes B$ is seen to be a matrix of order $mr \times ns$. It has mn blocks, the ij^{th} block is the matrix $a_{ij}B$ of the order $r \times s$. The following properties hold for the Kronecker product:

$$\begin{array}{rcl} vec(ABC) & = & (C^* \otimes A)vec(B) \\ vec(A)^*vec(B) & = & trace(A^*B) \\ trace(ABCD) & = & vec(D^*)^*(C^* \otimes A)vec(B) \\ & = & vec(D)^*(A \otimes C^*)vec(B^*) \\ vec(ab^*) & = & b \otimes a \\ (A \otimes B)^* & = & A^* \otimes B^* \\ rank(A \otimes B) & = & rank(A)rank(B) \\ (A \otimes B)^{-1} & = & A^{-1} \otimes B^{-1} \\ trace(A \otimes B) & = & trace(A)trace(B) \\ (A_1 + A_2) \otimes B & = & A_1 \otimes B + A_2 \otimes B \\ A \otimes (B_1 + B_2) & = & A \otimes B_1 + A \otimes B_2 \\ (A_1A_2) \otimes (B_1B_2) & = & (A_1 \otimes B_1)(A_2 \otimes B_2) \end{array} \tag{A.14}$$

The variance matrix of $vec(\underline{t}\underline{t}^*)$ consists of terms like

$$E\{[\underline{t}_i\underline{t}_j - E\{\underline{t}_i\underline{t}_j\}][\underline{t}_k\underline{t}_l - E\{\underline{t}_k\underline{t}_l\}]\} \quad i,j,k,l = 1,2,...,b \tag{A.15}$$

Since

$$E\{[\underline{t}_i\underline{t}_j - E\{\underline{t}_i\underline{t}_j\}][\underline{t}_k\underline{t}_l - E\{\underline{t}_k\underline{t}_l\}]\} = E\{\underline{t}_i\underline{t}_j\underline{t}_k\underline{t}_l\} - E\{\underline{t}_i\underline{t}_j\}E\{\underline{t}_k\underline{t}_l\} \tag{A.16}$$

substitution of (A.12) gives

$$\boxed{E\{[\underline{t}_i\underline{t}_j - E\{\underline{t}_i\underline{t}_j\}][\underline{t}_k\underline{t}_l - E\{\underline{t}_k\underline{t}_l\}]\} = q^{ik}q^{jl} + q^{il}q^{jk}} \tag{A.17}$$

From (A.17) follows that

$$\boxed{E\{[\underline{t}_i\underline{t} - E\{\underline{t}_i\underline{t}\}][\underline{t}_k\underline{t} - E\{\underline{t}_k\underline{t}\}]^*\} = e_i^*Q_te_kQ_t + Q_te_ke_i^*Q_t} \tag{A.18}$$

with

$$e_i = (0, ..., 0, 1, 0, ..., 0)^* \tag{A.19}$$

From (A.18) follows that

$$\boxed{Q_{vec(\underline{t}\underline{t}^*)} = \sum_i \sum_k e_i e_k^* \otimes e_i^* Q_t e_k Q_t + \sum_i \sum_k e_i e_k^* \otimes Q_t e_k e_i^* Q_t} \tag{A.20}$$

Since

$$\sum_i \sum_k e_i e_k^* \otimes e_i^* Q_t e_k Q_t = \sum_i \sum_k (e_i^* Q_t e_k) e_i e_k^* \otimes Q_t \tag{A.21}$$

and

$$Q_t = \sum_i \sum_k (e_i^* Q_t e_k) e_i e_k^* \tag{A.22}$$

it follows that

$$\sum_i \sum_k e_i e_k^* \otimes e_i^* Q_t e_k Q_t = Q_t \otimes Q_t \tag{A.23}$$

With (A.18), it follows that

$$\begin{aligned}
\sum_i \sum_k e_i e_k^* \otimes Q_t e_k e_i^* Q_t &= \sum_i \sum_k e_i e_k^* I \otimes Q_t e_k e_i^* Q_t \\
&= \sum_i \sum_k (e_i e_k^* \otimes Q_t)(I \otimes e_k e_i^* Q_t) \\
&= \sum_i \sum_k (I e_i e_k^* \otimes Q_t I)(II \otimes e_k e_i^* Q_t) \\
&= \sum_i \sum_k (I \otimes Q_t)(e_i e_k^* \otimes I)(I \otimes e_k e_i^*)(I \otimes Q_t)
\end{aligned}$$

or

$$\sum_i \sum_k e_i e_k^* \otimes Q_t e_k e_i^* Q_t = (I \otimes Q_t)(\sum_i \sum_k e_i e_k^* \otimes e_k e_i^*)(I \otimes Q_t) \tag{A.24}$$

Matrix $\sum_i \sum_k e_i e_k^* \otimes e_k e_i^*$ has the following properties:

1.

$$\sum_i \sum_k e_i e_k^* \otimes e_k e_i^* = \text{symmetric} \tag{A.25}$$

2.

$$(\sum_i \sum_k e_i e_k^* \otimes e_k e_i^*)(\sum_j \sum_l e_j e_l^* \otimes e_l e_j^*) = I \tag{A.26}$$

3.

$$(\sum_i \sum_k e_i e_k^* \otimes e_k e_i^*)(a \otimes b) = b \otimes a \tag{A.27}$$

4.

$$(\sum_i \sum_k e_i e_k^* \otimes e_k e_i^*)(A \otimes B) = (B \otimes A)(\sum_i \sum_k e_i e_k^* \otimes e_k e_i^* \tag{A.28}$$

Proof of (1): Trivial

Proof of (2):

$$\begin{aligned}(\sum_i \sum_k e_i e_k^* \otimes e_k e_i^*)(\sum_j \sum_l e_j e_l^* \otimes e_l e_j^*) &= \sum_i \sum_k \sum_j \sum_l (e_i e_k^* e_j e_l^* \otimes e_k e_i^* e_l e_j^*) \\ &= \sum_i \sum_k \sum_j \sum_l (\delta_{kj}\delta_{il} e_i e_l^* \otimes e_k e_j^*) \\ &= \sum_i \sum_k e_i e_i^* \otimes e_k e_k^* \\ &= (\sum_i e_i e_i^* \otimes \sum_k e_k e_k^*) \\ &= I \otimes I = I \end{aligned} \tag{A.29}$$

Proof of (3):

$$\begin{aligned}(\sum_i \sum_k e_i e_k^* \otimes e_k e_i^*)(a \otimes b) &= \sum_i \sum_k e_i e_k^* a \otimes e_k e_i^* b \\ &= \sum_i \sum_k a_k e_i \otimes b_i e_k \\ &= (\sum_i b_i e_i) \otimes (\sum_k a_k e_k) = b \otimes a \end{aligned} \tag{A.30}$$

Proof of (4): $A = \sum_\alpha a_\alpha e_\alpha^*$; $B = \sum_\beta b_\beta e_\beta^*$;

$$\begin{aligned}(\sum_i \sum_k e_i e_k^* \otimes e_k e_i^*)(A \otimes B) &= (\sum_i \sum_k e_i e_k^* \otimes e_k e_i^*)(\sum_\alpha \sum_\beta a_\alpha e_\alpha^* \otimes b_\beta e_\beta^*) \\ &= (\sum_i \sum_k e_i e_k^* \otimes e_k e_i^*)(\sum_\alpha \sum_\beta (a_\alpha \otimes b_\beta)(e_\alpha^* \otimes e_\beta^*)) \\ &= \sum_\alpha \sum_\beta b_\beta \otimes a_\alpha (e_\alpha^* \otimes e_\beta^*) \\ &= \sum_\alpha \sum_\beta b_\beta \otimes a_\alpha [(\sum_i \sum_k e_i e_k^* \otimes e_k e_i^*)(e_\beta \otimes e_\alpha)]^* \\ &= \sum_\alpha \sum_\beta (b_\beta \otimes a_\alpha)(e_\beta^* \otimes e_\alpha^*) \sum_i \sum_k e_i e_k^* \otimes e_k e_i^* \\ &= \sum_\alpha \sum_\beta (b_\beta e_\beta^* \otimes a_\alpha e_\alpha^*)(\sum_i \sum_k e_i e_k^* \otimes e_k e_i^*) \\ &= (B \otimes A)(\sum_i \sum_k e_i e_k^* \otimes e_k e_i^*) \end{aligned} \tag{A.31}$$

Using (A.28) we may write (A.24) as

$$\sum_i \sum_k e_i e_k^* \otimes Q_t e_k e_i^* Q_t = (\sum_i \sum_k e_i e_k^* \otimes e_k e_i^*)(Q_t \otimes Q_t) \tag{A.32}$$

Substitution of (A.23) and (A.32) into (A.20) finally gives

$$\begin{aligned} Q_{vec(\underline{tt}^*)} &= [I + \sum_i \sum_k e_i e_k^* \otimes e_k e_i^*][Q_t \otimes Q_t] \\ &= [Q_t \otimes Q_t][I + \sum_i \sum_k e_i e_k^* \otimes e_k e_i^*] \\ &= \frac{1}{2}[I + \sum_i \sum_k e_i e_k^* \otimes e_k e_i^*][Q_t \otimes Q_t][I + \sum_i \sum_k e_i e_k^* \otimes e_k e_i^*] \end{aligned} \tag{A.33}$$

A.3 The Singularity of Q_{vec}

It will be clear that the matrix Q_{vec} *has to be* singular. Since $Q_t \otimes Q_t$ is regular, the matrix

$$[I + \sum_i \sum_k e_i e_k^* \otimes e_k e_i^*] \tag{A.34}$$

has to be singular. Since (see also (A.25, A.26, A.27 and A.28))

$$[I + \sum_i \sum_k e_i e_k^* \otimes e_k e_i^*][I + \sum_i \sum_k e_i e_k^* \otimes e_k e_i^*] = 2[I + \sum_i \sum_k e_i e_k^* \otimes e_k e_i^*] \tag{A.35}$$

it follows that the matrix

$$\boxed{P_{b^2 \times b^2} = \frac{1}{2}[I + \sum_i \sum_k e_i e_k^* \otimes e_k e_i^*]} \tag{A.36}$$

is a *projector* (*idempotent*). We will know derive some properties of the projector P. Since the rank of a projector equals its trace, we have:

$$\begin{aligned} rank(P) &= \frac{1}{2} trace[I_{b^2} + \sum_i \sum_k e_i e_k^* \otimes e_k e_i^*] \\ &= \frac{1}{2} b^2 + \frac{1}{2} trace[\sum_i \sum_k e_i e_k^* \otimes e_k e_i^*] \\ &= \frac{1}{2} b^2 + \frac{1}{2} \sum_i \sum_k trace[e_i e_k^*] trace[e_k e_i^*] \\ &= \frac{1}{2} b^2 + \frac{1}{2} \sum_i \sum_k (trace[e_i e_k^*])^2 \end{aligned} \tag{A.37}$$

or

$$\boxed{rank(P) = \frac{1}{2} b(b+1)} \tag{A.38}$$

From this follows that the dimension of the range space and null space of P are

$$\boxed{\begin{aligned} \dim R(P) &= \tfrac{1}{2} b(b+1) \\ \dim N(P) &= b^2 - \tfrac{1}{2} b(b+1) = \tfrac{1}{2} b(b-1) \end{aligned}} \tag{A.39}$$

Since

$$P(a \otimes b) = \frac{1}{2}(a \otimes b + b \otimes a) \tag{A.40}$$

and

$$a \otimes b = vec(ba^*) \tag{A.41}$$

it follows that

$$Pvec(ba^*) = \frac{1}{2} vec(ba^* + ab^*) \tag{A.42}$$

Let X be an arbitrary matrix of order $b^2 \times b^2$ with column vectors x_i, $i = 1, 2, ..., b^2$. Then $X = \sum_i x_i e_i^*$ and thus $vec(X) = \sum_i vec(x_i e_i^*) = \sum_i e_i \otimes x_i$. This shows with (A.42) that

$$Pvec(X) = \frac{1}{2} \sum_i vec(x_i e_i^* + e_i x_i^*) = \frac{1}{2} vec[\sum_i x_i e_i^* + (\sum_i x_i e_i^*)^*] \tag{A.43}$$

or

$$\boxed{Pvec(X) = \frac{1}{2}(X + X^*)} \tag{A.44}$$

From this follows that

$$\boxed{\begin{array}{ll} Pvec(X) = vec(X) & \text{if} \quad X = X^* \\ \\ Pvec(X) = 0 & \text{if} \quad X = -X^* \end{array}} \tag{A.45}$$

Thus the *range space* of P is spanned by vectors $vec(X)$ with X symmetric, and the *null space* of P is spanned by vectors $vec(X)$, with X skew-symmetric. It will be clear from the above that the null spaces of P and Q_{vec} are identical. Thus

$$\boxed{N(P) = N(Q_{vec})} \tag{A.46}$$

Proof:

If $x \in N(P) \;\rightarrow Px = 0 \;\rightarrow (Q_t \otimes Q_t)Px = 0 \;\rightarrow Q_{vec}x = 0 \rightarrow x \in N(Q_{vec})$

If $x \in N(Q_{vec}) \;\rightarrow (Q_t \otimes Q_t)Px = 0 \;\rightarrow Px = 0 \;\rightarrow x \in N(P)$

A.4 The Solution

Consider the linear model

$$E\{\underline{y}\} = Ax, \quad D\{\underline{y}\} = Q_y \tag{A.47}$$

Let $T = [T_1^* T_2^*]^*$ be a square and full rank matrix. Then with (A.47):

$$E\{\begin{bmatrix} T_1\underline{y} \\ T_2\underline{y} \end{bmatrix}\} = \begin{bmatrix} T_1 A \\ T_2 A \end{bmatrix} x, \quad D\{\begin{bmatrix} T_1\underline{y} \\ T_2\underline{y} \end{bmatrix}\} = \begin{bmatrix} T_1 Q_y T_1^* & T_1 Q_y T_2^* \\ T_2 Q_y T_1^* & T_2 Q_y T_2^* \end{bmatrix} \tag{A.48}$$

Now assume that

$$N(Q_y) = R(T_2^*), \qquad R(T_2^*) \subset N(A^*) \tag{A.49}$$

Then with (A.48)

$$E\{\begin{bmatrix} T_1\underline{y} \\ T_2\underline{y} \end{bmatrix}\} = \begin{bmatrix} T_1 A x \\ 0 \end{bmatrix}, \quad D\{\begin{bmatrix} T_1\underline{y} \\ T_2\underline{y} \end{bmatrix}\} = \begin{bmatrix} T_1 Q_y T_1^* & 0 \\ 0 & 0 \end{bmatrix} \tag{A.50}$$

or

$$E\{T_1\underline{y}\} = T_1 A x, \quad D\{T_1\underline{y}\} = T_1 Q_y T_1^* \tag{A.51}$$

The solution of this model reads:

$$\begin{array}{rcl} \underline{\hat{x}} & = & [A^* T_1^* (T_1 Q_y T_1^*)^{-1} T_1 A]^{-1} A^* T_1^* (T_1 Q_y T_1^*)^{-1} T_1 \underline{y} \\ Q_{\hat{x}} & = & [A^* T_1^* (T_1 Q_y T_1^*)^{-1} T_1 A]^{-1} \end{array} \tag{A.52}$$

If we *translate* the above model to our situation, then

$$R(T_1^*) = R(P), \quad R(T_2^*) = N(P) = N(Q_{vec}) \tag{A.53}$$

Since $R(T_1^*) = R(P)$, we have

$$\boxed{P = T_1^* (T_1 T_1^*)^{-1} T_1} \tag{A.54}$$

Hence, with

$$Q_{vec} = PQ_t \otimes Q_t = Q_t \otimes Q_t P \tag{A.55}$$

we get

$$T_1^*(T_1T_1^*)^{-1}T_1Q_{vec}T_1^* = Q_t \otimes Q_t T_1^* \tag{A.56}$$

From this follows that

$$Q_t^{-1} \otimes Q_t^{-1}T_1^*(T_1T_1^*)^{-1} = T_1^*[T_1Q_{vec}T_1^*]^{-1} \tag{A.57}$$

or

$$Q_t^{-1} \otimes Q_t^{-1}T_1^*(T_1T_1^*)^{-1}T_1 = T_1^*[T_1Q_{vec}T_1^*]^{-1}T_1 \tag{A.58}$$

Hence

$$\boxed{T_1^*[T_1Q_{vec}T_1^*]^{-1}T_1 = PQ_t^{-1} \otimes Q_t^{-1}P} \tag{A.59}$$