4 $\underline{y}^R$ - Variates

4.1 Introduction

For the case that m=2 and n=1 the model with *observation equations* reads:

$$(1)\qquad \underset{2\times1}{E\{\underline{y}\}} = \underset{2\times1}{a}\ \underset{1\times1}{x}\ ;\ \underset{2\times2}{E\{(\underline{y}-ax)(\underline{y}-ax)^*\}} = \underset{2\times2}{Q_y}\ .$$

The corresponding model with *condition equations* reads:

$$(2)\qquad \underset{1\times2}{b^*}\ \underset{2\times1}{E\{\underline{y}\}} = \underset{1\times1}{0}\ ;\ \underset{2\times2}{E\{(\underline{y}-E\{\underline{y}\})(\underline{y}-E\{\underline{y}\})^*\}} = \underset{2\times2}{Q_y}\ .$$

In chapter 1 (see figure 1.8 or figure 1.14 in section 1.3) it was shown that the estimator $\hat{\underline{y}}$ of $E\{\underline{y}\}$ is the oblique projection of $\underline{y}$ onto the line with direction vector a. The direction of projection is parallel to the tangent line of the ellipse $z^*Q_y^{-1}z = a^*Q_y^{-1}a$ at a (see figure 4.1).

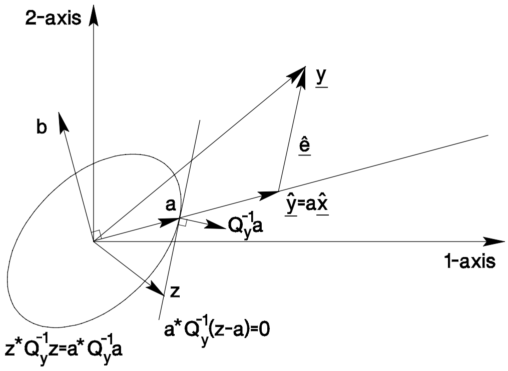

Figure 4.1: Oblique projection of $\underline{y}$

If we rotate the direction vector a so that it becomes parallel to the 1-axis, figure 4.1 transforms into figure 4.2.

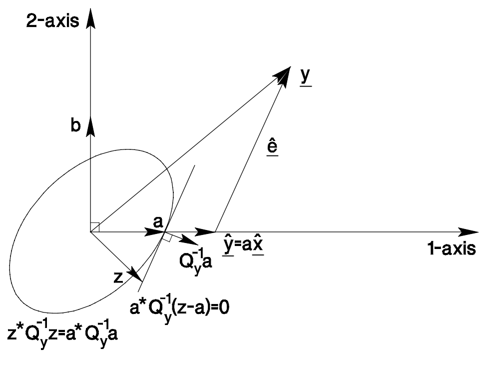

Figure 4.2: Oblique projection of $\underline{y}$ in case $a = (a_1\ 0)^*$

In this case the second component a_2 of the vector a is zero, and the first component b_1 of the vector b is zero. The corresponding two models (1) and (2) then take the form:

$$E\left\{\begin{pmatrix} \underline{y}_1 \\ \underline{y}_2 \end{pmatrix}\right\} = \begin{pmatrix} a_1 \\ 0 \end{pmatrix} x \ ; \ E(\underline{y}-ax)(\underline{y}-ax)^* = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix}$$

(3)

and:

(4) $$\begin{pmatrix} 0 \\ 1 \end{pmatrix}^* E\left\{\begin{pmatrix} \underline{y}_1 \\ \underline{y}_2 \end{pmatrix}\right\} = 0 \quad ; \quad E\{(\underline{y} - E\{\underline{y}\})(\underline{y} - E\{\underline{y}\})^*\} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix} .$$

Equation (4) shows that the coefficient of $E\{\underline{y}_1\}$ equals zero. This implies that the observable $\underline{y}_1$ does *not* appear in the condition equation. Observables that do not appear in the condition equations are called *free variates*. Thus $E\{\underline{y}_1\}$ is a free variate. If one writes (4) as:

(5) $$E\{\underline{y}_2\} = 0 \quad ; \quad E\{(\underline{y}_2 - E\{\underline{y}_2\})(\underline{y}_2 - E\{\underline{y}_2\})^*\} = \sigma_2^2 \ ,$$

one might be inclined to conclude, since $E\{\underline{y}_1\}$ does not appear in (5), that the estimator $\hat{\underline{y}}_1$ of $E\{\underline{y}_1\}$ equals $\underline{y}_1$ and thus that $\hat{\underline{e}}_1=\underline{y}_1-\hat{\underline{y}}_1$ is zero. This however is *not true*. Figure 4.2 shows namely clearly that both $\hat{\underline{e}}_1$ and $\hat{\underline{e}}_2$ are unequal to zero. The reason for this is that the major and minor axes of the ellipse of figure 4.2 make a non-zero angle with the 1-axis and 2-axis. If the major and minor axes of the ellipse are parallel to the 1-axis and the 2-axis, then figure 4.2 transforms into figure 4.3.

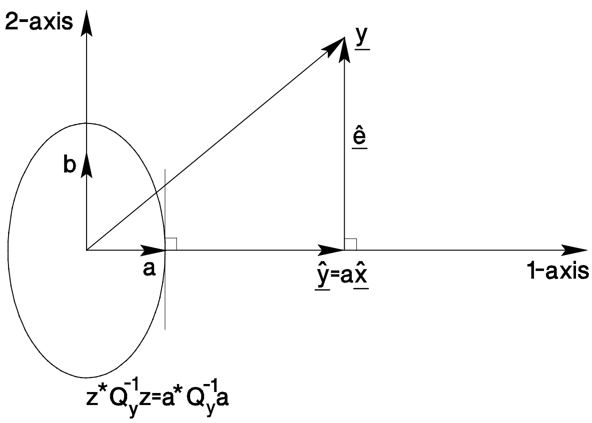

Figure 4.3: The covariance matrix Q_y is diagonal

In this case σ_{12}=0, Q_y is diagonal and $\hat{\underline{e}}_1$=0. It thus seems that the covariance, σ_{12}, between $\underline{y}_1$ and $\underline{y}_2$ determines to what extent $\hat{\underline{e}}_1=\underline{y}_1 - \hat{\underline{y}}_1$ differs from zero. If σ_{12}=0 then $\hat{\underline{e}}_1$=0, and if

$\sigma_{12}{\neq}0$ then also $\underline{\hat{e}}_1{\neq}0$. In order to find out how $\underline{\hat{e}}_1$ depends on σ_{12} consult figure 4.4. Figure 4.4 shows that:

(6) $$\frac{\hat{e}_1}{\hat{e}_2} = \tan\beta \ ,$$

and that β is the angle between the vector $Q_y^{-1}a$ and the 1-axis.

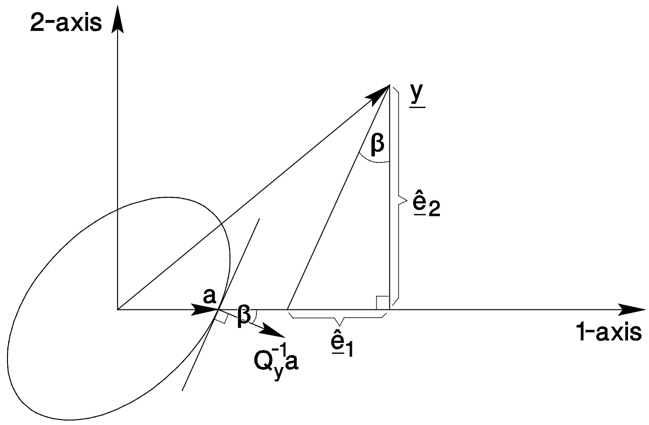

Figure 4.4: $\tan\beta = \frac{\hat{e}_1}{\hat{e}_2}$

Since the inverse of Q_y is:

$$Q_y^{-1} = \left(\sigma_1^2\sigma_2^2 - \sigma_{12}^2\right)^{-1} \begin{pmatrix} \sigma_2^2 & -\sigma_{12} \\ -\sigma_{12} & \sigma_1^2 \end{pmatrix}$$

it follows with $a{=}(a_1\ 0)^*$ that:

$$Q_y^{-1}a = \left(\sigma_1^2\sigma_2^2 - \sigma_{12}^2\right)^{-1} \begin{pmatrix} \sigma_2^2 a_1 \\ -\sigma_{12}a_1 \end{pmatrix} .$$

From this it follows that:

(7) $$\tan\beta = \frac{\sigma_{12}}{\sigma_2^2} \ ,$$

which with (6) gives that:

(8) $$\boxed{\underline{\hat{e}}_1 = \sigma_{12}\sigma_2^{-2}\underline{\hat{e}}_2 .}$$

This result shows how $\underline{\hat{e}}_1{=}\underline{y}_1 - \underline{\hat{y}}_1$, and thus the estimator $\underline{\hat{y}}_1{=}\underline{y}_1 - \underline{\hat{e}}_1$ of free variates, can be determined from $\underline{\hat{e}}_2{=}\underline{y}_2 - \underline{\hat{y}}_2$. In the next section we will generalize (8) to the multi-dimensional case.

Free variates are examples of so-called $\underline{y}^R$-variates. The *definition* of $\underline{y}^R$-variates reads as follows:

$\underline{y}^R$-variates are observables that are either stochastically or functionally <u>R</u>elated to another set of observables $\underline{y}$.

There are three types of $\underline{y}^R$-variates:

1. $\underline{y}^R$-variates that correlate with $\underline{y}$-variates. These are the *free variates* (vrije grootheden);
2. $\underline{y}^R$-variates that are functions of $\underline{y}$-variates. These are the *derived variates* (afgeleide grootheden);
3. $\underline{y}^R$-variates of which the $\underline{y}$-variates are functions. These are the *constituent variates* (samenstellende grootheden).

ad 1. These variates occur for instance when new measurements need to be included in existing geodetic networks. Consider for instance the levelling network of figure 3.1 of Chapter 3. Assume that the five heights have been estimated with the available nine levelled height differences. This gives the five estimators of the heights, $\hat{\underline{x}}_1$, $\hat{\underline{x}}_2$, $\hat{\underline{x}}_3$, $\hat{\underline{x}}_4$, $\hat{\underline{x}}_5$. Now assume that a new height difference has been measured, say between the points 1 and 5. If we denote this height difference by y_{10}, we can write the condition equation as:

(9) $$E\{\underline{y}_{10}\} - E\{\hat{\underline{x}}_5 - \hat{\underline{x}}_1\} = 0 \ .$$

On the basis of this condition equation we can compute $\hat{\underline{y}}_{10}$ and $\hat{\hat{\underline{x}}}_5$ $\hat{\hat{\underline{x}}}_1$, the improved estimators for the heights of the points 1 and 5. However, since the estimators $\hat{\underline{x}}_2$, $\hat{\underline{x}}_3$ and $\hat{\underline{x}}_4$ are correlated with $\hat{\underline{x}}_1$ and $\hat{\underline{x}}_5$, also the estimators $\hat{\underline{x}}_2$, $\hat{\underline{x}}_3$ and $\hat{\underline{x}}_4$ can be improved. The variates $\hat{\underline{x}}_2$, $\hat{\underline{x}}_3$ and $\hat{\underline{x}}_4$ are in this case the free variates since they do not appear in condition equation (9).

ad 2. Examples of derived variates in the case of the levelling network are height differences which are not directly measured, but which can be computed as functions of the estimated heights.

ad 3. Consider the levelling network of figure 4.5. Assume that the height difference between the points 1 and 2 has not been measured directly, but that instead it equals the sum of a number of measured height differences.

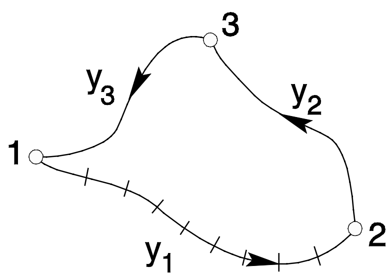

Figure 4.5: Levelling network

One can now include either the individual measured height differences in the condition equation or their sum $\underline{y}_1$. In the last case the individual measured height differences are the constituent variates.

In the following sections it will be shown that formula (8) holds for all three types of $\underline{y}^R$-variates.

4.2 Free variates

In this section we will generalize formula (8) to the multi-dimensional case. We will give two derivations. One based on the model with *condition equations*, and one based on the model with *observation equations*.

We know that the solution of the model with condition equations

(10) $$B^*E\{\underline{y}\} = 0 \quad ; \quad E\{(\underline{y}-E\{\underline{y}\})(\underline{y}-E\{\underline{y}\})^*\} = Q_y \ ,$$

reads:

(11) $$\begin{cases} \hat{\underline{y}} = [I - Q_y B(B^*Q_yB)^{-1}B^*]\ \underline{y} \\ \hat{\underline{e}} = \underline{y} - \hat{\underline{y}} = Q_yB(B^*Q_yB)^{-1}B^*\underline{y}. \end{cases}$$

Now let us assume that instead of (10) we have the model:

(12) $$\begin{pmatrix} B \\ 0 \end{pmatrix}^* E\{\begin{pmatrix} \underline{y} \\ \underline{y}^R \end{pmatrix}\} = 0; E\{\left[\begin{pmatrix} \underline{y} \\ \underline{y}^R \end{pmatrix} - E\{\begin{pmatrix} \underline{y} \\ \underline{y}^R \end{pmatrix}\}\right]\left[\begin{pmatrix} \underline{y} \\ \underline{y}^R \end{pmatrix} - E\{\begin{pmatrix} \underline{y} \\ \underline{y}^R \end{pmatrix}\}\right]^*\} = \begin{pmatrix} Q_y & Q_{yR} \\ Q_{Ry} & Q_R \end{pmatrix}.$$

In this case the coefficients of $\underline{y}^R$ are zero; thus $\underline{y}^R$ is a vector of *free variates*. When we compare (12) with (10) we see that B^* of (10) is replaced by (B^* 0) and Q_y of (10) by:

$$\begin{pmatrix} Q_y & Q_{yR} \\ Q_{Ry} & Q_R \end{pmatrix} .$$

Using the second equation of (11), we therefore get for model (12):

$$\begin{pmatrix} \hat{\underline{e}} \\ \hat{\underline{e}}^R \end{pmatrix} = \begin{pmatrix} \underline{y} \\ \underline{y}^R \end{pmatrix} - \begin{pmatrix} \hat{\underline{y}} \\ \hat{\underline{y}}^R \end{pmatrix} = \begin{pmatrix} Q_y & Q_{yR} \\ Q_{Ry} & Q_R \end{pmatrix}\begin{pmatrix} B \\ 0 \end{pmatrix}(B^*Q_yB)^{-1}B^*\underline{y}$$

or:

(13) $$\begin{pmatrix} \hat{\underline{e}} \\ \hat{\underline{e}}^R \end{pmatrix} = \begin{pmatrix} Q_yB(B^*Q_yB)^{-1}B^*\underline{y} \\ Q_{Ry}B(B^*Q_yB)^{-1}B^*\underline{y} \end{pmatrix} .$$

From the first equation of (13) it follows that:

$$Q_y^{-1}\hat{\underline{e}} = B(B^*Q_yB)^{-1}B^*\underline{y} \ .$$

If we substitute this into the second equation of (13) we finally get:

(14)
$$\boxed{\underline{\hat{e}}^R = Q_{Ry} Q_y^{-1} \underline{\hat{e}}}$$

This is the multi-dimensional generalization of equation (8) of section 4.1.

We will now derive formula (14) using the model with observation equations. The equivalent of model (12) in terms of observation equations reads:

(15)
$$E\{\begin{pmatrix} \underline{y} \\ \underline{y}^R \end{pmatrix}\} = \begin{pmatrix} A & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} x \\ E\{\underline{y}^R\} \end{pmatrix} \;;\; E\{\left[\begin{pmatrix} \underline{y} \\ \underline{y}^R \end{pmatrix} - E\{\begin{pmatrix} \underline{y} \\ \underline{y}^R \end{pmatrix}\}\right]\left[\begin{pmatrix} \underline{y} \\ \underline{y}^R \end{pmatrix} - E\{\begin{pmatrix} \underline{y} \\ \underline{y}^R \end{pmatrix}\}\right]^*\} = \begin{pmatrix} Q_y & Q_{yR} \\ Q_{Ry} & Q_R \end{pmatrix} .$$

Note that:

$$\begin{pmatrix} B \\ 0 \end{pmatrix}^* \begin{pmatrix} A & 0 \\ 0 & I \end{pmatrix} = (0 \;\; 0)$$

holds.

If we denote the inverse of the covariance matrix of the observables by:

(16)
$$\begin{pmatrix} Q_y & Q_{yR} \\ Q_{Ry} & Q_R \end{pmatrix}^{-1} = \begin{pmatrix} G_y & G_{yR} \\ G_{Ry} & G_R \end{pmatrix} ,$$

the *normal equations* for model (15) take the form:

(17)
$$\begin{pmatrix} A^*G_yA & A^*G_{yR} \\ G_{Ry}A & G_R \end{pmatrix} \begin{pmatrix} \underline{\hat{x}} \\ \underline{\hat{y}}^R \end{pmatrix} = \begin{pmatrix} A^*G_y & A^*G_{yR} \\ G_{Ry} & G_R \end{pmatrix} \begin{pmatrix} \underline{y} \\ \underline{y}^R \end{pmatrix} .$$

From the last equation of (17) it follows that:

$$G_R\underline{\hat{y}}^R = G_R\underline{y}^R + G_{Ry}(\underline{y} - A\underline{\hat{x}})$$

or:

$$\underline{\hat{e}}^R = \underline{y}^R - \underline{\hat{y}}^R = -G_R^{-1} G_{Ry}(\underline{y} - A\underline{\hat{x}})$$

or:

(18)
$$\underline{\hat{e}}^R = -G_R^{-1} G_{Ry}\underline{\hat{e}} .$$

This results looks already very similar to (14). However, equation (14) is expressed in terms of covariance matrices, whereas (18) is expressed in terms of weight matrices.
Since:

$$\begin{pmatrix} G_y & G_{yR} \\ G_{Ry} & G_R \end{pmatrix} \begin{pmatrix} Q_y & Q_{yR} \\ Q_{Ry} & Q_R \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}$$

(see (16)), it follows that $G_{Ry}\, Q_y + G_R\, Q_{Ry} = 0$ and thus that:

(19) $$-G_R^{-1} G_{Ry} = Q_{Ry} Q_y^{-1} \; .$$

This, together with (18) proves (14).

4.3 Derived variates

In this section we will prove that formula (14) also holds for *derived variates*. $\underline{y}^R$-variates are derived variates if they are functions of the observables $\underline{y}$. Let us assume that this functional relationship takes the form:

(20) $$\underset{p\times 1}{\underline{y}^R} = \underset{p\times m}{\Lambda} \; \underset{m\times 1}{\underline{y}} \; .$$

Then also:

(21) $$\hat{\underline{y}}^R = \Lambda \; \hat{\underline{y}} \; .$$

Subtraction of (21) from (20) gives:

(22) $$\hat{\underline{e}}^R = \Lambda \; \hat{\underline{e}} \; .$$

What remains to be shown is now that Λ equals $Q_{Ry}\, Q_y^{-1}$. Application of the *propagation law of covariances* to (20) gives:

$$Q_{Ry} = \Lambda \; Q_y$$

or:

$$\Lambda = Q_{Ry} Q_y^{-1} \; .$$

This, together with (22) shows that formula (14) also holds for derived variates.

4.4 Constituent variates

In this section we will prove that formula (14) holds for *constituent variates*. Constituent variates are $\underline{y}^R$-variates of which the $\underline{y}$-variates are functions. Let us assume that this functional relationship takes the form:

(23) $$\underset{m\times 1}{\underline{y}} = \underset{m\times p}{\Lambda} \; \underset{p\times 1}{\underline{y}^R} \; .$$

Application of the propagation law of variances gives:

(24) $$Q_y = \Lambda \; Q_R \; \Lambda^* \; ,$$

and application of the propagation law of covariances gives:

(25) $$Q_{Ry} = Q_R \Lambda^* \ .$$

As we know $\underline{\hat{e}} = \underline{y} - \underline{\hat{y}}$ follows from solving the model:

(26) $$B^* E\{\underline{y}\} = 0 \ ; \ E\{(\underline{y} - E\{\underline{y}\})(\underline{y} - E\{\underline{y}\})^*\} = Q_y$$

as:

(27) $$\underline{\hat{e}} = Q_y B (B^* Q_y B)^{-1} B^* \underline{y} \ .$$

With (23) we may formulate the model with condition equations also as:

(28) $$B^* \Lambda E\{\underline{y}^R\} = 0 \ ; \ E\{(\underline{y}^R - E\{\underline{y}^R\})(\underline{y} - E\{\underline{y}^R\})^* = Q_R \ .$$

When compared with (26) this means that B of (26) is replaced by $\Lambda^* B$, that $\underline{y}$ is replaced by $\underline{y}^R$ and that Q_y is replaced by Q_R. Instead of (27) we therefore get for model (28):

(29) $$\underline{\hat{e}}^R = Q_R \Lambda^* B \left(B^* \Lambda Q_R \Lambda^* B \right)^{-1} B^* \Lambda \underline{y}^R \ .$$

With (23), (24) and (25) this can also be written as:

$$\underline{\hat{e}}^R = Q_{Ry} B \left(B^* Q_y B \right)^{-1} B^* \underline{y}$$

or with (27) as:

$$\underline{\hat{e}}^R = Q_{Ry} Q_y^{-1} \underline{\hat{e}} \ ,$$

which is identical to (14).

Let us, as an example of the above, "rederive" the solution of the model with condition equations. The model with condition equations reads:

(30) $$B^* E\{\underline{y}\} = 0 \ ; \ E\{(\underline{y} - E\{\underline{y}\})(\underline{y} - E\{\underline{y}\})^*\} = Q_y \ .$$

If we define the misclosure vector as:

(31) $$\underline{t} = B^* \underline{y} \ ,$$

then (30) can also be written as:

(32) $$I^* \ E\{\underline{t}\} = 0 \ ; \ E\{(\underline{t} - E\{\underline{t}\})(\underline{t} - E\{\underline{t}\})^*\} = Q_t \ .$$

This model is also in the form of condition equations. The coefficient matrix of the condition equations is in this case the unit matrix I. If we solve for model (32) we get for $\underline{\hat{e}}_t = \underline{t} - \underline{\hat{t}}$:

(33) $$\underline{\hat{e}}_t = Q_t I (I^* Q_t I)^{-1} I^* \underline{t} = \underline{t} \ .$$

We can now use our formula (14) in order to derive $\hat{\underline{e}} = \underline{y} - \hat{\underline{y}}$. This is done by interpreting $\underline{y}$ of (31) as an $\underline{y}^R$-variate and $\underline{t}$ of (31) as an $\underline{y}$-variate. Application of formula (14) gives then:

(34) $$\hat{\underline{e}} = \boldsymbol{Q}_{yt}\boldsymbol{Q}_t^{-1}\hat{\underline{e}}_t \; .$$

Application of the propagation law of variances to (31) gives:

(35) $$\boldsymbol{Q}_t = \boldsymbol{B}^*\boldsymbol{Q}_y\boldsymbol{B} \; ,$$

and application of the propagation law of covariances to (31) gives:

(36) $$\boldsymbol{Q}_{yt} = \boldsymbol{Q}_y\boldsymbol{B} \; .$$

Substitution of (33), (35) and (36) into (34) gives with (31) then finally:

$$\hat{\underline{e}} = \boldsymbol{Q}_y\boldsymbol{B}(\boldsymbol{B}^*\boldsymbol{Q}_y\boldsymbol{B})^{-1}\boldsymbol{B}^*\underline{y} \; .$$

Hence, we have obtained our well-known and familiar result again.