# *Part 3: Linear Prediction*

## *3.1. The prediction concept*

While estimation is the assignment of numerical values to unknown fixed (non-random) parameters, prediction is the corresponding process for parameters modeled as random variables. More precisely *prediction* is the process of assigning numerical values to the unknown outcomes of random variables, on the basis of the known corresponding outcomes of another set of random variables. In other words prediction is the estimation of unknown outcomes of random variables. The known outcomes are usually the results of measurements, but they can also be known functions of the actually observed values. The term prediction is also used some times to denote not only the process but also the resulting numerical values. Here the term prediction and predicted value will be used for distinction in analogy with the term estimation and estimate.

A random variable $Y$ in its rigorous mathematical definition is a measurable function defined on a probability space $(\Omega, B, P)$, see e.g. **Rao** (1965) for a rigorous exposition. This roughly means that random variables are functions, which assign numerical values $Y(\omega)$ (outcomes) to different "experiments" $\omega$, out of the set of all possible experiments $\Omega$, while (almost) each set of experiments $B$ (event) has a specific probability $P(B)$ of happening. The probabilities of the events themselves are unknown and one works with the probabilities of sets of the outcomes. These probabilities follow from the property that a random variable is a measurable function, and they are not defined for every set of numerical values but only for the sets within a well-defined class. The same is true for the probabilities of the experiments where the events, i.e. the sets of experiments with defined probabilities, are not all possible sets but only those within a well-defined class.

In data analysis one deals with only one experiment and the same terms (observations, estimates, errors, predicted values) to denote both the random variables (functions) and their corresponding outcomes (numerical values). Since this may be a source of confusion the distinction will be emphasized whenever needed in the following.

The prediction process itself may be visualized as a mapping $\mathbf{y} \to \tilde{\mathbf{u}}$ from the known outcomes $\mathbf{y} = \boldsymbol{Y}(\omega)$ of the original random variables $\boldsymbol{Y}$ to the "estimates" $\tilde{\mathbf{u}}$ of the outcomes $\mathbf{u} = \boldsymbol{U}(\omega)$ of the new random variables $\boldsymbol{U}$. Since this mapping is independent of the specific numerical values $\mathbf{y}$ of the known outcomes, i.e., it is the same over all possible outcomes, it is therefore also a mapping $\boldsymbol{Y} \to \tilde{\boldsymbol{U}}$ from the original random variables $\boldsymbol{Y}$ to the predicted ones $\tilde{\boldsymbol{U}}$, defined "pointwise" $\boldsymbol{Y}(\omega) \to \boldsymbol{U}(\omega)$ for each outcome $\omega$. It is this interpretation of prediction as a mapping between sets of random variables, which allows optimal prediction to be defined according to desired stochastic properties of the predicted random variables $\tilde{\boldsymbol{U}}$.

In the standard linear model used in geodetic data analysis the known observations $\mathbf{b}$ are related to the non-random parameters $\mathbf{x}$ and the unknown observational errors $\mathbf{v}$, through

$$\mathbf{b} = \mathbf{A}\mathbf{x} + \mathbf{v}, \tag{1}$$

(fixed effects model), where $\mathbf{A}$ is the known design matrix. Equation (1) is at the same time a relation $\boldsymbol{B} = \mathbf{A}\mathbf{x} + \boldsymbol{V}$, between the random variables $\boldsymbol{V}$ and $\boldsymbol{B}$, as well as between their corresponding outcomes $\mathbf{b} = \boldsymbol{B}(\omega)$, $\mathbf{v} = \boldsymbol{V}(\omega)$. In this model the only random variables under prediction are the errors $\boldsymbol{V}$. The predicted values $\tilde{\mathbf{v}} = \tilde{\boldsymbol{V}}(\omega)$ follow directly from the estimates $\hat{\mathbf{x}}$ ($\tilde{\mathbf{v}} = \mathbf{b} - \mathbf{A}\hat{\mathbf{x}}$), on the basis of the compatibility condition $\mathbf{b} = \mathbf{A}\hat{\mathbf{x}} + \tilde{\mathbf{v}}$. The symbols ( $\wedge$ ) and ( ~ ) are used throughout to denote the known estimates and predicted values of the corresponding unknown fixed parameters and the unknown outcomes of random variables, respectively.

In this approach estimation precedes prediction and priority is given to the optimality of the estimates, without much concern about the optimality of the prediction, since anyway, in data analysis, the estimates $\hat{\mathbf{x}}$ are of primary importance and predictions $\tilde{\mathbf{v}}$ of a secondary one.

Here, on the contrary, the reverse approach will be followed. Optimal predictions will be directly determined for any random variable from the observations $\mathbf{b}$, and estimates will follow on the basis of compatibility with the predictions of the random variables present in the model, such as the random variables $\mathbf{s}$ (stochastic parameters) and $\mathbf{v}$ (errors) in the model (mixed effects model)

$$\mathbf{b} = \mathbf{A}\mathbf{x} + \mathbf{G}\mathbf{s} + \mathbf{v} . \tag{2}$$

Also the compatibility of optimal predictions with existing optimal estimates will be examined. In fact the prediction theory developed here is complementary to the unified linear estimation theory developed in Part 2. Optimality of predictions can be sought without compatibility concerns, in the particular case where only stochastic parameters are present according to the model (random effects model)

$$\mathbf{b} = \mathbf{G}\mathbf{s} + \mathbf{v} . \tag{3}$$

For the prediction to take place both sets of random variables, original and new, must be stochastically interrelated. The maximum possible information about this interrelation is the joint probability function or the joint probability density function of the random variables of both sets. When this information is not available, the interrelation can be partly expressed through the cross-covariance matrix of the two sets, within what is usually called a "second order theory". This cross-covariance matrix, together with the two auto-covariance matrices and the corresponding mean vectors, are then the building tools for the construction of various optimal predictions.

Instead of examining prediction within various linear models, such as (1), (2) and (3) above, a more unified approach will be followed here to obtain results that can be readily applied to any particular linear model.

The model-independent general approach followed here includes various former derivations based on specific linear models, see, e.g. **Bibby** & **Toutenburg** (1977), **Toutenburg** (1982), and **Schaffrin** (1985) where further references to the literature can be found.

## *3.2. Optimality criteria and prediction types*

Let $\mathbf{y}$ be a vector of random variables with known outcomes and $u$ a single random variable whose outcome is to be predicted. If $\tilde{u}$ is the prediction - random variable, the prediction error - random variable is defined as

$$\varepsilon = \tilde{u} - u \tag{4}$$

and the expectation of its square (i.e. the mean value of $\varepsilon^2$ over all possible outcomes) is the mean square error (MSE) of the prediction

$$\mathrm{MSE}(\tilde{u}) = E\{\varepsilon^2\} = E\{(\tilde{u} - u)^2\} \tag{5}$$

Optimality of prediction is relative to an appropriate choice of a risk function $R(\tilde{u}, u)$, the optimal or best prediction $\tilde{u}$ being the one minimizing $R$. For a single random variable the MSE is an appropriate risk function. When a vector of random variables $\mathbf{u}$ are to be simultaneously predicted by $\tilde{\mathbf{u}}$, the mean square error (MSE) is defined as the matrix

$$\mathrm{MSE}(\tilde{u}) = E\{(\tilde{\mathbf{u}} - \mathbf{u})(\tilde{\mathbf{u}} - \mathbf{u})^T\} \tag{6}$$

and a possible choice of risk function is the scalar MSE

$$R(\tilde{\mathbf{u}}, \mathbf{u}) = E\{(\tilde{\mathbf{u}} - \mathbf{u})^T (\tilde{\mathbf{u}} - \mathbf{u})\} = \mathrm{traceMSE}(\tilde{\mathbf{u}}) \tag{7}$$

A somewhat more general form of risk function is

$$R(\tilde{\mathbf{u}},\mathbf{u}) = E\{(\tilde{\mathbf{u}}-\mathbf{u})^T \mathbf{V}(\tilde{\mathbf{u}}-\mathbf{u})\} \tag{8}$$

where $\mathbf{V}$ is a positive-definite weight matrix, (7) being the special case of (8) for $\mathbf{V} = \mathbf{I}$. It can be shown that for the class of linear predictions which will be considered here, the choice of $\mathbf{V}$ has no effect on the results (see, e.g., **Toutenburg**, 1982). When the scalar MSE of equation (7) is used as risk function, its minimization leads to the need for complicated matrix algebra. The derivations of the optimal predictions can be greatly simplified with the use of the following postulate:

The optimal prediction of a single random variable must be independent of which set of random variables needs to be predicted. The prediction of a random variable will be the same when separately predicted, as when predicted together with any other set of random variables. This excludes the use of different weights on different predictions according to their relative significance in a specific application. On the other hand it yields predictions independent of any specific application, by minimizing (5) separately for each component $u_i$ of the vector $\mathbf{u}$. The results $\tilde{u}_i$ are identical to the components of the vector $\tilde{\mathbf{u}}$ resulting from the minimization of (7).

The optimal predictions obtained using the MSE as the risk function to be minimized, are simply called *best* predictions.

It is highly desirable to obtain the prediction which is best, i.e. has minimum MSE among all possible predictions. This is not always possible since the prediction $\tilde{u}$ is a function $\tilde{u} = \tilde{u}(\mathbf{y})$ of the observed random variables $\mathbf{y}$**,** and the set of all possible functions is simply too large to be analytically handled. For this reason the prediction is confined to a smaller class of parametrized functions and the best prediction in the class is obtained through the determination of the values of the parameters which minimize the risk function. One simple choice is the classes of the *homogeneous linear predictions*

$$\tilde{u} = \mathbf{d}^T \mathbf{y} \tag{9}$$

which is a class of linear functions of $\mathbf{y}$, with $\mathbf{d}$ being the vector of parameters defining the members of the class. Another simple choice is the classes of the *inhomogeneous linear predictions*

$$\tilde{u} = \mathbf{d}^T \mathbf{y} + \kappa \, , \tag{10}$$

which is a class of not actually linear functions of $\mathbf{y}$, with $\mathbf{d}$, and $\kappa$ being the class parameters. With these simply defined classes the problem of best prediction becomes an algebraic problem of determination of the values of the parameters $\mathbf{d}$ (or $\mathbf{d}$ and $\kappa$) which minimize the MSE.

The term *Best Linear inhomogeneous Prediction* (*inhomBLIP* for short) refers to the prediction which has the minimum MSE among all possible inhomogeneous linear predictions. The term *Best Linear homogeneous Prediction* (*homBLIP* for short) refers to the prediction which has the minimum MSE among all possible homogeneous linear predictions. Since the class of inhomogeneous linear predictions is larger of the class of homogeneous linear predictions, which it contains, one expects that in general inhomBLIP has a smaller risk than homBLIP. On the other hand, the non-linearity of inhomBLIP as a function of $\mathbf{y}$, has some undesired side effects that will be pointed out in the following.

Another characteristic of prediction which might need consideration in addition to the MSE, is its bias, defined as the difference between the means of the prediction and the predicted random variable

$$\beta = E\{\tilde{u}\} - E\{u\} = E\{\varepsilon\} \, . \tag{11}$$

When the condition $\beta = 0$ is imposed, the corresponding prediction is called unbiased. Combination with the two classes (9) and (10) gives two more types of prediction:

The term *Best Linear inhomogeneous Unbiased Prediction* (*inhomBLUP* for short) refers to the prediction which has the minimum MSE among all possible unbiased inhomogeneous linear predictions. The term *Best Linear homogeneous Unbiased Prediction* (*homBLUP* for short) refers to the prediction which has the minimum MSE among all possible unbiased homogeneous linear predictions. Since the condition of unbiasedness leads to a smaller class of functions, it is expected that inhomBLUP will generally have larger risk than inhomBLIP, and the same holds for homBLUP with respect to homBLIP.

Finally prediction of a random variable $u$ should not be confused with the estimation of its mean value $E\{u\}$. In prediction what is estimated is the outcome of the random variable corresponding to the available outcome of the observations.

### 3.3. Best (minimum mean square error) prediction

When the joint probability function $f(u,\mathbf{y})$ of $u$ and $\mathbf{y}$ is known it is possible to seek the best prediction $\tilde{u} = g(\mathbf{y})$ of $u$, which is the prediction with minimum MSE among all possible functions $g(\mathbf{y})$. It can be shown that $\tilde{u} = g(\mathbf{y})$ is the conditional mean of $u$ for the given $\mathbf{y}$, which is defined as

$$g(\mathbf{y}) = E\{u \mid \mathbf{y}\} = \int_u u\ f_{u|\mathbf{y}}(u,\mathbf{y})\, du \tag{12}$$

where

$$f_{u|\mathbf{y}}(u,\mathbf{y}) = \frac{f(u,\mathbf{y})}{f_u(u)} \tag{13}$$

is the conditional probability density function of $u$ for the given $\mathbf{y}$, and

$$f_u(u) = \int_{\mathbf{y}} f(u,\mathbf{y})\, d\mathbf{y} \tag{14}$$

is the marginal distribution of $u$.

The MSE can be written

$$E\{\varepsilon^2\} = \int_u \int_{\mathbf{y}} \varepsilon^2 f(u,\mathbf{y})\, du\, d\mathbf{y} = \int_u \int_{\mathbf{y}} \varepsilon^2 f_u(u) f_{u|\mathbf{y}}(u,\mathbf{y})\, du\, d\mathbf{y} = \int_u E\{\varepsilon^2 \mid \mathbf{y}\} f_u(u)\, du \tag{15}$$

and it is minimized when the integrant $E\{\varepsilon^2 \mid \mathbf{y}\}$ becomes a minimum. Setting

$$m = E\{u \mid \mathbf{y}\} \tag{16}$$

the integrant can be written

$$\begin{aligned} E\{\varepsilon^2 \mid \mathbf{y}\} &= E\{(u-m+m-g)^2 \mid \mathbf{y}\} = E\{(u-m)^2 \mid \mathbf{y}\} - 2E\{(u-m)(g-m) \mid \mathbf{y}\} + E\{(m-g)^2 \mid \mathbf{y}\} = \\ &= E\{(u-m)^2 \mid \mathbf{y}\} - 2(g-m)\big(E\{u \mid \mathbf{y}\} - m\big) + E\{(m-g)^2 \mid \mathbf{y}\} = \\ &= E\{(u-m)^2 \mid \mathbf{y}\} + E\{(m-g)^2 \mid \mathbf{y}\}\,. \end{aligned} \tag{17}$$

The first term is independent of $g$ and the second term which is non-negative takes its minimum (zero) for $g(\mathbf{y}) = m$ and the proof is completed.

In the particular case where $u$ and $\mathbf{y}$ have the joint multivariate Gaussian distribution

$$\begin{bmatrix} u \\ \mathbf{y} \end{bmatrix} \sim N\left(\begin{bmatrix} m_u \\ \mathbf{m_y} \end{bmatrix}, \begin{bmatrix} \sigma_u^2 & \mathbf{c}_{u\mathbf{y}} \\ \mathbf{c}_{\mathbf{y}u} & \mathbf{C_{yy}} \end{bmatrix}\right) \tag{18}$$

with

$$f(u,\mathbf{y}) = (2\pi)^{-(n+1)/2} \begin{vmatrix} \sigma_u^2 & \mathbf{c}_{u\mathbf{y}} \\ \mathbf{c}_{\mathbf{y}u} & \mathbf{C_{yy}} \end{vmatrix} \exp\left\{-\frac{1}{2}\begin{bmatrix} m_u \\ \mathbf{m_y} \end{bmatrix}^T \begin{bmatrix} \sigma_u^2 & \mathbf{c}_{u\mathbf{y}} \\ \mathbf{c}_{\mathbf{y}u} & \mathbf{C_{yy}} \end{bmatrix}^{-1} \begin{bmatrix} m_u \\ \mathbf{m_y} \end{bmatrix}\right\} \tag{19}$$

the best prediction becomes (**Graybill**, 1965, p. 106)

$$\tilde{u} = E\{u \mid \mathbf{y}\} = m_u + \mathbf{c}_{u\mathbf{y}} \mathbf{C}_{\mathbf{yy}}^{-1} (\mathbf{y} - \mathbf{m_y}) \tag{20}$$

### *3.4. Inhomogeneous and homogeneous linear prediction*

The best prediction cannot be determined when information about the random variables $\mathbf{y}$ (given) and $u$ (to be predicted) is restricted to the means

$$m = E\{u\}, \qquad \mathbf{m} = E\{\mathbf{y}\} \tag{21}$$

and the variances and covariances

$$\sigma^2 = E\{(u-m)^2\}, \qquad \mathbf{C} = E\{(\mathbf{y}-\mathbf{m})(\mathbf{y}-\mathbf{m})^T\}, \qquad \mathbf{c} = E\{(\mathbf{y}-\mathbf{m})(u-m)\} \tag{22}$$

(note that subscripts have been dropped for notational simplicity). In this case the prediction must be confined to the classes of inhomogeneous and homogeneous linear predictions.

It is also possible to consider predictions where $\mathbf{m}$ and $m$ are not completely known, but they are not completely unknown either, being restricted to be of the form

$$\mathbf{m} = \mathbf{Ax} + \mathbf{t}, \qquad m = \mathbf{a}^T \mathbf{x} + t \tag{23}$$

where $\mathbf{A}$, $\mathbf{t}$, $\mathbf{a}$ and $t$ are known while the parameter vector $\mathbf{x}$ is unknown. Giving consideration to the special cases $\mathbf{m} = \mathbf{0}$, $m = 0$, $\mathbf{t} = \mathbf{0}$, $t = 0$, it is possible to construct 16 different models from the combination of one of the following 4 models for $\mathbf{m}$ with one of the corresponding models for m

| | | | | | |
|---|---|---|---|---|---|
| (a) | $\mathbf{m} = \mathbf{0}$ | | (1) | $m = 0$ | |
| (b) | $\mathbf{m} \neq \mathbf{0} =$ known | | (2) | $m \neq 0 =$ known | |
| (c) | $\mathbf{m} = \mathbf{Ax}$, | $\mathbf{A} =$ known | (3) | $m = \mathbf{a}^T \mathbf{x}$, | $\mathbf{a} =$ known |
| (d) | $\mathbf{m} = \mathbf{Ax} + \mathbf{t}$, | $\mathbf{A}, \mathbf{t} =$ known | (4) | $m = \mathbf{a}^T \mathbf{x} + t$, | $\mathbf{a}, t =$ known |

For the inhomogeneous prediction $\tilde{u} = \mathbf{d}^T \mathbf{y} + \kappa$ the prediction bias becomes

$$\beta_{\text{inhom}} = E\{\mathbf{d}^T \mathbf{y} + \kappa\} - E\{u\} = \mathbf{d}^T \mathbf{m} + \kappa - m \tag{24}$$

and the risk function (mean square error) becomes

$$R_{\text{inhom}} = E\{(\mathbf{d}^T \mathbf{y} + \kappa - u)^2\} = E\{[\mathbf{d}^T (\mathbf{y}-\mathbf{m}) - (u-m) + \beta]^2\} =$$

$$= E\{\mathbf{d}^T (\mathbf{y}-\mathbf{m})(\mathbf{y}-\mathbf{m})^T \mathbf{d} - (u-m)^2 + \beta^2 - 2\mathbf{d}^T (\mathbf{y}-\mathbf{m})(u-m) - 2\beta(u-m) - 2\beta\mathbf{d}^T (\mathbf{y}-\mathbf{m})\} =$$

$$= \mathbf{d}^T \mathbf{C} \mathbf{d} + \sigma^2 + \beta_{\text{inhom}}^2 - 2\mathbf{d}^T \mathbf{c} \tag{25}$$

The bias and risk function of the homogeneous prediction follow directly from those of the inhomogeneous prediction by simply setting $\kappa = 0$

$$\beta_{\text{hom}} = \mathbf{d}^T \mathbf{m} - m \tag{26}$$

$$R_{\text{hom}} = \mathbf{d}^T \mathbf{C} \mathbf{d} + \sigma^2 + \beta_{\text{inhom}}^2 - 2\mathbf{d}^T \mathbf{c} \tag{27}$$

Note that the risk function has the same form for both cases and the subscripts can be drawn for notational simplicity. For the minimization of the risk function the following derivatives are needed

$$\frac{\partial R}{\partial \mathbf{d}} = 2\mathbf{d}^T \mathbf{C} - 2\mathbf{c}^T + 2\beta \frac{\partial \beta}{\partial \mathbf{d}} = 2\mathbf{d}^T \mathbf{C} - 2\mathbf{c}^T + 2\beta \mathbf{m}^T \tag{28}$$

$$\frac{\partial R}{\partial \kappa} = 2\beta \frac{\partial \beta}{\partial \kappa} = 2\beta \tag{29}$$

for the inhomogeneous case.

### *3.4.1. Best linear inhomogeneous prediction (inhomBLIP)*

For the best linear inhomogeneous prediction the parameters $\mathbf{d}$ and $\kappa$, which minimize the inhomogeneous risk function are determined by setting the derivatives of the risk function with respect to $\mathbf{d}$ and $\kappa$ equal to zero. From equations (28) and (29) it follows that the solution is determined by the system of equations

$$\mathbf{C}\mathbf{d} - \mathbf{c} + \beta \mathbf{m} = \mathbf{0} \tag{30}$$

$$\beta = \mathbf{d}^T \mathbf{m} + \kappa - m = 0 \tag{31}$$

with obvious solution (assuming $|\mathbf{C}| \neq 0$)

$$\mathbf{d} = \mathbf{C}^{-1} \mathbf{c} \tag{32}$$

$$\kappa = m - \mathbf{d}^T \mathbf{m} = m - \mathbf{c}^T \mathbf{C}^{-1} \mathbf{m} \,. \tag{33}$$

With these values the best inhomogeneous linear prediction becomes

$$\text{inhomBLIP :} \qquad \tilde{u} = m + \mathbf{c}^T \mathbf{C}^{-1} (\mathbf{y} - \mathbf{m}) \tag{34}$$

This relation is valid for all the special cases for $\mathbf{m}$ and $m$, but it has no practical values for combinations involving cases (c), (d) and (3), (4) where it depends on the unknown parameters $\mathbf{x}$.

### *3.4.2. Best linear homogeneous prediction (homBLIP)*

For the best linear homogeneous prediction the parameters $\mathbf{d}$ which minimize the homogeneous risk function are determined by setting the derivative of the risk function with respect to $\mathbf{d}$ equal to zero. From equation (28) it follows that the solution is determined by

$$\mathbf{Cd} - \mathbf{c} + \beta\mathbf{m} = \mathbf{Cd} - \mathbf{c} + (\mathbf{m}^T\mathbf{d} - m)\mathbf{m} = \mathbf{0} \tag{35}$$

with obvious solution

$$\mathbf{d} = (\mathbf{C} + \mathbf{mm}^T)^{-1}(\mathbf{c} + m\mathbf{m}) \tag{36}$$

so that the homBLIP prediction becomes

$$\tilde{u} = (\mathbf{c} + m\mathbf{m})^T(\mathbf{C} + \mathbf{mm}^T)^{-1}\mathbf{y} \tag{37}$$

With the use of the matrix identity

$$(\mathbf{C} + \mathbf{mm}^T)^{-1} = \mathbf{C}^{-1} - \frac{1}{1 + \mathbf{m}^T\mathbf{C}^{-1}\mathbf{m}}\mathbf{C}^{-1}\mathbf{mm}^T\mathbf{C}^{-1} \tag{38}$$

the prediction takes the form

$$\tilde{u} = \frac{m - \mathbf{m}^T\mathbf{C}^{-1}\mathbf{c}}{1 + \mathbf{m}^T\mathbf{C}^{-1}\mathbf{m}}\mathbf{m}^T\mathbf{C}^{-1}\mathbf{y} - \mathbf{c}^T\mathbf{C}^{-1}\mathbf{y} \tag{39}$$

A more convenient form is

$$\text{homBLIP :} \qquad \tilde{u} = \alpha m + \mathbf{c}^T\mathbf{C}^{-1}(\mathbf{y} - \alpha\mathbf{m}), \qquad \alpha = \frac{\mathbf{m}^T\mathbf{C}^{-1}\mathbf{y}}{1 + \mathbf{m}^T\mathbf{C}^{-1}\mathbf{m}} \tag{40}$$

Again the above equation is valid for all special cases for $\mathbf{m}$ and $m$, but has no practical value for cases involving models (c), (d) and (3), (4).

### *3.4.3. Best linear inhomogeneous unbiased prediction (inhomBLUP)*

When prediction is restricted to be unbiased, the risk function must be minimized under the condition that $\beta = 0$. The Lagrangean function for the minimization is

$$\Phi = R - 2\lambda\beta \tag{41}$$

where is a $\lambda$ is a Lagrange multiplier, with solution determined by

$$\frac{\partial\Phi}{\partial\mathbf{d}} = \frac{\partial R}{\partial\mathbf{d}} - 2\lambda\frac{\partial\beta}{\partial\mathbf{d}} = \frac{\partial R}{\partial\mathbf{d}} - 2\lambda\mathbf{m}^T = \mathbf{0} \tag{42}$$

$$\frac{\partial\Phi}{\partial\kappa} = \frac{\partial R}{\partial\kappa} - 2\lambda\frac{\partial\beta}{\partial\kappa} = \frac{\partial R}{\partial\kappa} - 2\lambda = 0 \tag{43}$$

$$\frac{\partial\Phi}{\partial\lambda} = -2\beta = 0 \tag{44}$$

Using $\beta = 0$ from the last equation and the derivatives of $R$ from equations (28) and (29) the solution system becomes

$$\mathbf{Cd} - \mathbf{c} - \lambda\mathbf{m} = \mathbf{0} \tag{45}$$

$$\lambda = 0 \tag{46}$$

$$\beta = \mathbf{d}^T \mathbf{m} + \kappa - m = 0 \tag{47}$$

with solution

$$\mathbf{d} = \mathbf{C}^{-1}\mathbf{c} \tag{48}$$

$$\kappa = m - \mathbf{d}^T \mathbf{m} = m - \mathbf{c}^T \mathbf{C}^{-1}\mathbf{m} \tag{49}$$

With these values the prediction becomes

$$\text{inhomBLUP :} \qquad \tilde{u} = m + \mathbf{c}^T \mathbf{C}^{-1}(\mathbf{y} - \mathbf{m}) \tag{50}$$

which is identical with the inhomBLIP prediction. The derivation above could have been avoided by noting that the inhomBLIP prediction happens to be unbiased. Since the set of inhomogeneous unbiased predictions is a subset of the set of inhomogeneous predictions and the minimum of $R$ over the larger set happens to fall within the smaller subset of unbiased predictions, it will obviously be the minimum over the smaller subset as well.

### *3.4.4. Best linear homogeneous unbiased prediction (homBLUP)*

The solution for this case follows the lines of the previous derivation, with the same Lagrangean function $\Phi$, but since now $R$ and $\beta$ do not depend on $\kappa$, only the derivatives of $\Phi$ with respect to $\mathbf{d}$ and $\lambda$ should be set equal to zero

$$\frac{\partial \Phi}{\partial \mathbf{d}} = \frac{\partial R}{\partial \mathbf{d}} - 2\lambda \frac{\partial \beta}{\partial \mathbf{d}} = \frac{\partial R}{\partial \mathbf{d}} - 2\lambda \mathbf{m}^T = \mathbf{0} \tag{51}$$

$$\frac{\partial \Phi}{\partial \lambda} = -2\beta = 0 \tag{52}$$

which with the help of equation (28) become

$$\mathbf{Cd} - \mathbf{c} - \lambda \mathbf{m} = \mathbf{0} \tag{53}$$

$$\beta = \mathbf{d}^T \mathbf{m} - m = 0 \tag{54}$$

The first equation gives

$$\mathbf{d} = \mathbf{C}^{-1}\mathbf{c} + \lambda \mathbf{C}^{-1}\mathbf{m} \tag{55}$$

which can be replaced into (54) and solved for $\lambda$ to give (assuming $\mathbf{m} \neq \mathbf{0}$)

$$\lambda = \frac{m - \mathbf{c}^T \mathbf{C}^{-1}\mathbf{m}}{\mathbf{m}^T \mathbf{C}^{-1}\mathbf{m}} \tag{56}$$

With the above values of $\mathbf{d}$ and $\lambda$ the prediction becomes

$$\tilde{u} = \mathbf{c}^T \mathbf{C}^{-1}\mathbf{y} + \frac{m - \mathbf{c}^T \mathbf{C}^{-1}\mathbf{m}}{\mathbf{m}^T \mathbf{C}^{-1}\mathbf{m}} \mathbf{m}^T \mathbf{C}^{-1}\mathbf{y} \tag{57}$$

A more convenient form is

$$\text{homBLUP}: \qquad \tilde{u} = \alpha\, m + \mathbf{c}^T \mathbf{C}^{-1}(\mathbf{y} - \alpha\,\mathbf{m}), \qquad \alpha = \frac{\mathbf{m}^T \mathbf{C}^{-1}\mathbf{y}}{\mathbf{m}^T \mathbf{C}^{-1}\mathbf{m}}, \qquad (\mathbf{m} \neq \mathbf{0}) \tag{58}$$

When $\mathbf{m} = \mathbf{0}$, the bias becomes $\beta = -m$ and therefore homBLUP does not exist when $\mathbf{m} = \mathbf{0}$ and $m \neq 0$ (cases a2, a3, a4). When $\mathbf{m} = \mathbf{0}$ and $m = 0$ (case a1) equation (53) becomes

$$\mathbf{Cd} - \mathbf{c} = \mathbf{0} \tag{59}$$

(independent of $\lambda$) and the prediction takes the form

$$\tilde{u} = \mathbf{c}^T \mathbf{C}^{-1}\mathbf{y}\,. \tag{60}$$

In a more convenient form the prediction for all cases can be summarized as

$$\begin{array}{llll} \text{homBLUP}: & \tilde{u} = \alpha\, m + \mathbf{c}^T \mathbf{C}^{-1}(\mathbf{y} - \alpha\,\mathbf{m}), & \alpha = \dfrac{\mathbf{m}^T \mathbf{C}^{-1}\mathbf{y}}{\mathbf{m}^T \mathbf{C}^{-1}\mathbf{m}}, & (\mathbf{m} \neq \mathbf{0}) \\ & \quad " & \alpha = 1 & (\mathbf{m} = \mathbf{0},\, m = 0) \\ & \tilde{u} = \text{ does not exist} & (\mathbf{m} = \mathbf{0},\, m \neq 0) & \end{array} \tag{61}$$

### 3.4.5. *Uniformly unbiased predictions*

When either $\mathbf{m}$ or $m$ or both depend linearly on a set of unknown parameters $\mathbf{x}$, all the above predictions turn out to depend also on $\mathbf{x}$ and they cannot be used in practice. Some times it is proposed to replace $\mathbf{x}$ with a known approximation $\mathbf{x}_0$ with the hope that the resulting predictions $\tilde{u}(\mathbf{x}_0)$ will not deviate significantly from the correct ones $\tilde{u}(\mathbf{x})$. However such approximations are not always available and the deviation from the optimal predictions cannot be controlled without knowledge of $\mathbf{x}$.

One way to obtain predictions which are independent of the $\mathbf{x}$ is the restriction not to unbiased predictions (which are unbiased for the true but unknown values of $\mathbf{x}$) but to predictions which are unbiased for any $\mathbf{x}$. Such predictions are called uniformly unbiased predictions and their class is more restricted than the class of unbiased predictions.

The conditions for uniform unbiasedness follow by treating $\beta(\mathbf{x}) = 0$ as a Diophantine equation in $\mathbf{x}$, and setting the coefficients of $\mathbf{x}$ and the constant term equal to zero. The general model under consideration is (case d4)

$$\mathbf{m} = \mathbf{Ax} + \mathbf{t}\,, \qquad m = \mathbf{a}^T \mathbf{x} + t \tag{62}$$

All other cases follow by setting $\mathbf{t}$, $t$, $\mathbf{A}$, $\mathbf{a}^T$ equal to zero or $\mathbf{t} = \mathbf{m}$, $t = m$ to obtain any one of the 16 possible subcases. However the concept of uniformly unbiased prediction does not apply to cases (a1), (a2), (b1), (b2) which do not involve $\mathbf{x}$.

In the inhomogeneous case the bias is

$$\beta = \mathbf{d}^T \mathbf{m} + \kappa - m = \mathbf{d}^T (\mathbf{Ax} + \mathbf{t}) - (\mathbf{a}^T \mathbf{x} + t) + \kappa = (\mathbf{A}^T \mathbf{d} - \mathbf{a})^T \mathbf{x} + \mathbf{d}^T \mathbf{t} - t + \kappa \tag{63}$$

and the conditions for uniformly unbiased prediction become

$$\mathbf{A}^T \mathbf{d} - \mathbf{a} = \mathbf{0} \tag{64}$$

$$\mathbf{d}^T\mathbf{t} - t + \kappa = 0\,. \tag{65}$$

For the homogeneous case ( $\kappa = 0$ ) the first condition remains the same and the second becomes

$$\mathbf{d}^T\mathbf{t} - t = 0\,. \tag{66}$$

### *3.4.5.1. Best Linear inhomogeneous Uniformly Unbiased Prediction (inhomBLUUP)*

For the minimization of the risk function $R$ under the two conditions for uniformly unbiased prediction, the Lagrangean function is

$$\Phi = R - 2\mathbf{k}^T(\mathbf{A}^T\mathbf{d} - \mathbf{a}) - 2\nu(\mathbf{d}^T\mathbf{t} - t + \kappa) \tag{67}$$

where $\mathbf{k}$ and $\nu$ are Lagrange multipliers. The solution is determined by the system

$$\frac{\partial\Phi}{\partial\mathbf{d}} = \frac{\partial R}{\partial\mathbf{d}} - 2\mathbf{k}^T\mathbf{A}^T - 2\nu\mathbf{t}^T = 2\mathbf{d}^T\mathbf{C} - 2\mathbf{c}^T + 2\beta\mathbf{m}^T - 2\mathbf{k}^T\mathbf{A}^T - 2\nu\mathbf{t}^T = \mathbf{0} \tag{68}$$

$$\frac{\partial\Phi}{\partial\kappa} = \frac{\partial R}{\partial\kappa} - 2\nu = 2\beta - 2\nu = 0 \tag{69}$$

$$\frac{\partial\Phi}{\partial\mathbf{k}} = -2(\mathbf{A}^T\mathbf{d} - \mathbf{a})^T = \mathbf{0} \tag{70}$$

$$\frac{\partial\Phi}{\partial\nu} = -2(\mathbf{d}^T\mathbf{t} - t + \kappa) = 0 \tag{71}$$

Since $\beta = 0$, equation (69) gives $\nu = 0$ and the remaining ones become

$$\mathbf{C}\mathbf{d} - \mathbf{c} - \mathbf{A}\mathbf{k} = \mathbf{0} \tag{72}$$

$$\mathbf{A}^T\mathbf{d} - \mathbf{a} = \mathbf{0} \tag{73}$$

$$\mathbf{t}^T\mathbf{d} - t + \kappa = 0\,. \tag{74}$$

Solving (72) for $\mathbf{d}$

$$\mathbf{d} = \mathbf{C}^{-1}\mathbf{c} + \mathbf{C}^{-1}\mathbf{A}\mathbf{k} \tag{75}$$

and replacing in (73) and solving for $\mathbf{k}$ it follows that

$$\mathbf{k} = \mathbf{N}^{-1}\mathbf{a} - \mathbf{N}^{-1}\mathbf{A}^T\mathbf{C}^{-1}\mathbf{c} \tag{76}$$

where

$$\mathbf{N} = \mathbf{A}^T\mathbf{C}^{-1}\mathbf{A} \tag{77}$$

and $|\mathbf{N}| \neq 0$ has been assumed. With this value of $\mathbf{k}$, (75) becomes

$$\mathbf{d} = \mathbf{C}^{-1}\mathbf{c} + \mathbf{C}^{-1}\mathbf{A}\mathbf{N}^{-1}\mathbf{a} - \mathbf{C}^{-1}\mathbf{A}\mathbf{N}^{-1}\mathbf{A}^T\mathbf{C}^{-1}\mathbf{c} = \mathbf{H}^T\mathbf{C}^{-1}\mathbf{c} + \mathbf{C}^{-1}\mathbf{A}\mathbf{N}^{-1}\mathbf{a} \tag{78}$$

where

$$\mathbf{H} = \mathbf{I} - \mathbf{A}\mathbf{N}^{-1}\mathbf{A}^T\mathbf{C}^{-1}. \tag{79}$$

From equation (74) and the above value of $\mathbf{d}$ it follows that

$$\kappa = t - \mathbf{t}^T\mathbf{d} = t - \mathbf{a}^T\mathbf{N}^{-1}\mathbf{A}\mathbf{C}^{-1}\mathbf{t} - \mathbf{c}^T\mathbf{C}^{-1}\mathbf{H}\mathbf{t} \tag{80}$$

The above values of $\mathbf{d}$ and $\kappa$ give the best linear inhomogeneous uniformly unbiased prediction

$$\text{inhomBLUUP}: \qquad \begin{gathered} \tilde{u} = t + \mathbf{a}^T\mathbf{N}^{-1}\mathbf{A}\mathbf{C}^{-1}(\mathbf{y}-\mathbf{t}) + \mathbf{c}^T\mathbf{C}^{-1}\mathbf{H}(\mathbf{y}-\mathbf{t}) \\ \mathbf{H} = \mathbf{I} - \mathbf{A}(\mathbf{A}^T\mathbf{C}^{-1}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{C}^{-1}. \end{gathered} \tag{81}$$

This equation refers to case (d4) and other cases follow by using the appropriate values for $\mathbf{A}$, $\mathbf{a}$, $\mathbf{t}$, $t$. InhomBLUUP does not apply to cases (a1), (a2), (b1), (b2) and it does not exist when $\mathbf{A} = \mathbf{0}$ and $\mathbf{a} \neq \mathbf{0}$ (cases a3, a4, b3, b4) since in these cases it is impossible for $\beta$ to vanish for every $\mathbf{x}$.

### *3.4.5.2. Best Linear homogeneous Uniformly Unbiased Prediction (homBLUUP)*

The best linear homogeneous uniformly unbiased prediction is derived in a similar way. The Lagrangean function is

$$\Phi = R - 2\mathbf{k}^T(\mathbf{A}^T\mathbf{d} - \mathbf{a}) - 2\nu(\mathbf{d}^T\mathbf{t} - t) \tag{82}$$

and the solution is determined by the system

$$\frac{\partial \Phi}{\partial \mathbf{d}} = \frac{\partial R}{\partial \mathbf{d}} - 2\mathbf{k}^T\mathbf{A}^T - 2\nu\mathbf{t}^T = 2\mathbf{d}^T\mathbf{C} - 2\mathbf{c}^T + 2\beta\mathbf{m}^T - 2\mathbf{k}^T\mathbf{A}^T - 2\nu\mathbf{t}^T = \mathbf{0} \tag{83}$$

$$\frac{\partial \Phi}{\partial \mathbf{k}} = -2(\mathbf{A}^T\mathbf{d} - \mathbf{a})^T = \mathbf{0} \tag{84}$$

$$\frac{\partial \Phi}{\partial \nu} = -2(\mathbf{d}^T\mathbf{t} - t + \kappa) = 0 \tag{85}$$

Since $\beta = 0$ the system becomes

$$\mathbf{C}\mathbf{d} - \mathbf{c} - \mathbf{A}\mathbf{k} - \nu\,\mathbf{t} = \mathbf{0} \tag{86}$$

$$\mathbf{A}^T\mathbf{d} = \mathbf{a} \tag{87}$$

$$\mathbf{t}^T\mathbf{d} = t \tag{88}$$

Solving (86) for $\mathbf{d}$

$$\mathbf{d} = \mathbf{C}^{-1}\mathbf{c} + \mathbf{C}^{-1}\mathbf{A}\mathbf{k} + \nu\,\mathbf{C}^{-1}\mathbf{t} \tag{89}$$

and replacing in (87) and solving for $\mathbf{k}$ it follows that

$$\mathbf{k} = \mathbf{N}^{-1}\mathbf{a} - \mathbf{N}^{-1}\mathbf{A}^T\mathbf{C}^{-1}\mathbf{c} - \nu\,\mathbf{N}^{-1}\mathbf{A}^T\mathbf{C}^{-1}\mathbf{t} \tag{90}$$

With this value of $\mathbf{k}$, (89) becomes

$$\mathbf{d} = \mathbf{C}^{-1}\mathbf{c} + \mathbf{C}^{-1}\mathbf{A}\mathbf{N}^{-1}\mathbf{a} - \mathbf{C}^{-1}\mathbf{A}\mathbf{N}^{-1}\mathbf{A}^T\mathbf{C}^{-1}\mathbf{c} - \nu\,\mathbf{C}^{-1}\mathbf{A}\mathbf{N}^{-1}\mathbf{A}^T\mathbf{C}^{-1}\mathbf{t} + \nu\,\mathbf{C}^{-1}\mathbf{t} =$$

$$= \mathbf{H}^T\mathbf{C}^{-1}\mathbf{c} + \mathbf{C}^{-1}\mathbf{A}\mathbf{N}^{-1}\mathbf{a} + \nu\,\mathbf{H}^T\mathbf{C}^{-1}\mathbf{t} \tag{91}$$

Replacing the above value of $\mathbf{d}$ in equation (88) and solving for $\nu$, it follows that

$$\nu = \frac{1}{\mathbf{t}^T\mathbf{C}^{-1}\mathbf{H}\mathbf{t}}(t - \mathbf{c}^T\mathbf{C}^{-1}\mathbf{H}\mathbf{t} - \mathbf{a}^T\mathbf{N}^{-1}\mathbf{A}\mathbf{C}^{-1}\mathbf{t})\,. \tag{92}$$

With the above values of $\mathbf{d}$ and $\nu$ the best linear homogeneous uniformly unbiased prediction becomes

$$\tilde{u} = \mathbf{c}^T\mathbf{C}^{-1}\mathbf{H}\mathbf{y} + \mathbf{a}^T\mathbf{N}^{-1}\mathbf{A}\mathbf{C}^{-1}\mathbf{y} + \frac{\mathbf{t}^T\mathbf{C}^{-1}\mathbf{H}\mathbf{y}}{\mathbf{t}^T\mathbf{C}^{-1}\mathbf{H}\mathbf{t}}(t - \mathbf{c}^T\mathbf{C}^{-1}\mathbf{H}\mathbf{t} - \mathbf{a}^T\mathbf{N}^{-1}\mathbf{A}\mathbf{C}^{-1}\mathbf{t}) \tag{93}$$

A more convenient form is

homBLUUP : $\tilde{u} = \alpha\, t + \mathbf{a}^T\mathbf{N}^{-1}\mathbf{A}\mathbf{C}^{-1}(\mathbf{y} - \alpha\,\mathbf{t}) + \mathbf{c}^T\mathbf{C}^{-1}\mathbf{H}(\mathbf{y} - \alpha\,\mathbf{t})$

$$\alpha = \frac{\mathbf{t}^T\mathbf{C}^{-1}\mathbf{H}\mathbf{y}}{\mathbf{t}^T\mathbf{C}^{-1}\mathbf{H}\mathbf{t}} \tag{94}$$

This equation refers to case (d4) and other cases follow by using the appropriate values for $\mathbf{A}$, $\mathbf{a}$, $\mathbf{t}$, $t$. For the same reasons as with inhomBLUUP, homBLUUP does not apply to cases (a1), (a2), (b1), (b2) and it does not exist for cases (a3), (a4), (b3) and (b4). Furthermore it does not exist for cases (c2) and (c4) where also it is impossible for $\beta$ to vanish for every $\mathbf{x}$.

### *3.4.6. Vector generalizations and summary*

When the prediction is required not for a single but for a vector of random variables $\mathbf{u}$, the prediction equations for $\tilde{u}_i$ follow by applying the previous results for every component $u_i$ separately and combining them in a single vector $\tilde{\mathbf{u}}$. Using the notation

$$\mathbf{m}_\mathbf{y} = E\{\mathbf{y}\} \qquad \mathbf{m}_\mathbf{u} = E\{\mathbf{u}\} \qquad \mathbf{C}_{\mathbf{yy}} = E\{(\mathbf{y}-\mathbf{m}_\mathbf{y})(\mathbf{y}-\mathbf{m}_\mathbf{y})^T\} \qquad \mathbf{C}_{\mathbf{uy}} = E\{(\mathbf{u}-\mathbf{m}_\mathbf{u})(\mathbf{y}-\mathbf{m}_\mathbf{y})^T\} \tag{95}$$

for the mean vectors and the covariance matrices the following results are obtained

<u>A. inhomBLIP = inhomBLUP, homBLIP and homBLUP</u>

$$\tilde{\mathbf{u}} = \alpha\,\mathbf{m}_\mathbf{u} + \mathbf{C}_{\mathbf{uy}}\mathbf{C}_{\mathbf{uy}}^{-1}(\mathbf{y} - \alpha\,\mathbf{m}_\mathbf{y}) \tag{96}$$

inhomBLIP = inhomBLUP : $\alpha = 1$ (97)

homBLIP : $$\alpha = \frac{\mathbf{m}_\mathbf{y}^T\mathbf{C}_{\mathbf{yy}}^{-1}\mathbf{y}}{1 + \mathbf{m}_\mathbf{y}^T\mathbf{C}_{\mathbf{yy}}^{-1}\mathbf{m}_\mathbf{y}} \tag{98}$$

homBLUP : $$\alpha = \frac{\mathbf{m}_{\mathbf{y}}^T \mathbf{C}_{\mathbf{yy}}^{-1} \mathbf{y}}{\mathbf{m}_{\mathbf{y}}^T \mathbf{C}_{\mathbf{yy}}^{-1} \mathbf{m}_{\mathbf{y}}} \tag{99}$$

(does not exist for $\mathbf{m}_{\mathbf{u}} \neq \mathbf{0}$, $\mathbf{m}_{\mathbf{y}} = \mathbf{0}$, cases a2, a3, a4)
(In case a1: $\alpha =$ undefined, $\tilde{\mathbf{u}} =$ independent of $\alpha$)

For predictions depending on a set of unknown parameters $\mathbf{x}$, the following values of $\mathbf{m}_{\mathbf{y}}$ and $\mathbf{m}_{\mathbf{u}}$ should be used in the above equations

$$\mathbf{m}_{\mathbf{y}} = \mathbf{Ax} + \mathbf{t} \qquad \mathbf{m}_{\mathbf{u}} = \mathbf{A}^* \mathbf{x} + \mathbf{t}^* \tag{100}$$

(where in particular $\mathbf{t}$, or $\mathbf{t}^*$, or both may be zero), leading to predictions which depend on the unknown $\mathbf{x}$ and they cannot be used.

<u>B. Uniformly unbiased predictions</u>

$$\tilde{\mathbf{u}} = \alpha\, \mathbf{t}^* + \mathbf{A}^* \mathbf{N}^{-1} \mathbf{A}^T \mathbf{C}_{\mathbf{yy}}^{-1} (\mathbf{y} - \alpha\, \mathbf{t}) + \mathbf{C}_{\mathbf{uy}} \mathbf{C}_{\mathbf{yy}}^{-1} \mathbf{H} (\mathbf{y} - \alpha\, \mathbf{t}) \tag{101}$$

$$\mathbf{N} = \mathbf{A}^T \mathbf{C}_{\mathbf{yy}}^{-1} \mathbf{A} \tag{102}$$

$$\mathbf{H} = \mathbf{I} - \mathbf{ANA}^T \mathbf{C}_{\mathbf{yy}}^{-1} \tag{103}$$

inhomBLUUP : $$\alpha = 1 \tag{104}$$

(does not apply to cases a1, a2, b1, b2)
(does not exist for cases a3, a4, b3, b4)

homBLUUP : $$\alpha = \frac{\mathbf{t}^T \mathbf{C}_{\mathbf{yy}}^{-1} \mathbf{Hy}}{\mathbf{t}^T \mathbf{C}_{\mathbf{yy}}^{-1} \mathbf{Ht}} \tag{105}$$

(does not apply to cases a1, a2, b1, b2)
(does not exist for cases a3, a4, b3, b4, c2, c4)

### *3.4.7. MSE matrices and biases*

The bias vector for any inhomogeneous predictions $\tilde{\mathbf{u}} = \mathbf{Dy} + \mathbf{k}$, $\mathbf{D}$ and $\mathbf{k}$ being the coefficient matrices, is given by

$$\boldsymbol{\beta} = E\{\tilde{\mathbf{u}}\} - E\{\mathbf{u}\} = \mathbf{Dm}_{\mathbf{y}} + \mathbf{k} - \mathbf{m}_{\mathbf{u}} \tag{106}$$

The mean square error matrix for inhomogeneous predictions $\tilde{\mathbf{u}} = \mathbf{Dy} + \mathbf{k}$, is given by

$$\mathrm{MSE}(\tilde{\mathbf{u}}) = E\{(\tilde{\mathbf{u}} - \mathbf{u})(\tilde{\mathbf{u}} - \mathbf{u})^T\} = E\{[\mathbf{D}(\mathbf{y} - \mathbf{m}_{\mathbf{y}}) - (\mathbf{u} - \mathbf{m}_{\mathbf{u}}) + \boldsymbol{\beta}][\mathbf{D}(\mathbf{y} - \mathbf{m}_{\mathbf{y}}) - (\mathbf{u} - \mathbf{m}_{\mathbf{u}}) + \boldsymbol{\beta}]^T\} =$$

$$= \mathbf{DC}_{\mathbf{yy}} \mathbf{D}^T - \mathbf{DC}_{\mathbf{uy}}^T - \mathbf{C}_{\mathbf{uy}} \mathbf{D}^T + \mathbf{C}_{\mathbf{uu}} + \boldsymbol{\beta\beta}^T \tag{107}$$

The same relations hold for homogeneous predictions if $\mathbf{k} = \mathbf{0}$ is set.

Identifying $\mathbf{D}$ and $\mathbf{k}$ for each particular type of prediction and replacing in the above formulas, the following results are derived (note that the coefficient $\alpha$ is a random variable when $\alpha \neq 1$):

A. inhomBLIP = inhomBLUP, homBLIP and homBLUP

$\mathrm{MSE}(\tilde{\mathbf{u}}) = \mathbf{C}_{\mathbf{uu}} - \mathbf{C}_{\mathbf{uy}}\mathbf{C}_{\mathbf{yy}}^{-1}\mathbf{C}_{\mathbf{uy}}^{T} + \rho\,(\mathbf{C}_{\mathbf{uy}}\mathbf{C}_{\mathbf{yy}}^{-1}\mathbf{m}_{\mathbf{y}} - \mathbf{m}_{\mathbf{u}}\,(\mathbf{C}_{\mathbf{uy}}\mathbf{C}_{\mathbf{yy}}^{-1}\mathbf{m}_{\mathbf{y}} - \mathbf{m}_{\mathbf{u}})^{T}$ MISSING “**)**” !!!

inhomBLIP = inhomBLIP :

$$\boldsymbol{\beta} = \mathbf{0} \tag{108}$$

$$\rho = 0 \tag{109}$$

homBLIP :

$$\boldsymbol{\beta} = \frac{1}{1 + \mathbf{m}_{\mathbf{y}}^{T}\mathbf{C}_{\mathbf{yy}}^{-1}\mathbf{m}_{\mathbf{y}}}(\mathbf{C}_{\mathbf{uy}}\mathbf{C}_{\mathbf{yy}}^{-1}\mathbf{m}_{\mathbf{y}} - \mathbf{m}_{\mathbf{u}}) \tag{110}$$

$$\rho = \frac{1}{1 + \mathbf{m}_{\mathbf{y}}^{T}\mathbf{C}_{\mathbf{yy}}^{-1}\mathbf{m}_{\mathbf{y}}} \tag{111}$$

homBLUP :

$$\boldsymbol{\beta} = \mathbf{0} \tag{112}$$

$$\rho = \frac{1}{\mathbf{m}_{\mathbf{y}}^{T}\mathbf{C}_{\mathbf{yy}}^{-1}\mathbf{m}_{\mathbf{y}}} \tag{113}$$

B. InhomBLUUP and homBLUUP

$$\mathrm{MSE}(\tilde{\mathbf{u}}) = \mathbf{C}_{\mathbf{uu}} - \mathbf{C}_{\mathbf{uy}}\mathbf{C}_{\mathbf{yy}}^{-1}\mathbf{C}_{\mathbf{uy}}^{T} + (\mathbf{C}_{\mathbf{uy}}\mathbf{C}_{\mathbf{yy}}^{-1}\mathbf{A} - \mathbf{A}^{*})\mathbf{N}^{-1}(\mathbf{C}_{\mathbf{uy}}\mathbf{C}_{\mathbf{yy}}^{-1}\mathbf{A} - \mathbf{A}^{*})^{T} +$$

$$+ \rho\,(\mathbf{t}^{*} - \mathbf{A}^{*}\mathbf{N}^{-1}\mathbf{A}^{T}\mathbf{C}_{\mathbf{yy}}^{-1}\mathbf{t} - \mathbf{C}_{\mathbf{uy}}\mathbf{C}_{\mathbf{yy}}^{-1}\mathbf{Ht})(\mathbf{t}^{*} - \mathbf{A}^{*}\mathbf{N}^{-1}\mathbf{A}^{T}\mathbf{C}_{\mathbf{yy}}^{-1}\mathbf{t} - \mathbf{C}_{\mathbf{uy}}\mathbf{C}_{\mathbf{yy}}^{-1}\mathbf{Ht})^{T} \tag{114}$$

inhomBLUUP :

$$\boldsymbol{\beta} = \mathbf{0} \tag{115}$$

$$\rho = 0 \tag{116}$$

homBLUUP :

$$\boldsymbol{\beta} = \mathbf{0} \tag{117}$$

$$\rho = \frac{1}{\mathbf{t}^{T}\mathbf{C}_{\mathbf{yy}}^{-1}\mathbf{m}_{\mathbf{y}}\mathbf{Ht}} \tag{118}$$

### 3.4.8. Invariance characteristics of the predictions

The random vector $\mathbf{y}$ with known outcome, which is the basis of the predictions for other random vectors $\mathbf{u}$, is not always uniquely defined. In the most straightforward case where the known outcomes are those originally observed, they are replaced in the linearization process by reduced observations. These follow by subtracting from the original observations approximate values of the observables which are not uniquely defined: a different set of the, more or less arbitrary, approximate values for the parameters leads to different approximate values of the observables and consequently a different random vector of reduced observations. In other situations where the formulation of the "adjustment with observation equations" cannot be directly applied, e.g. in the method of condition equations, the known vector is a vector of misclosures, which are nonlinear functions of the original observations.

It is therefore necessary to examine the behavior of the prediction process when a transformation of the known random vector is used instead. If $\mathbf{a}$ and $\mathbf{b}$ are random vectors, $\tilde{\mathbf{a}}(\mathbf{b})$ denotes the prediction of $\mathbf{a}$, based on the known $\mathbf{b}$. Let $\mathcal{T}$ be a class of transformations and $T$ a member of the class.

A prediction is said to be invariant with respect to the known random vector in the class $\mathcal{T}$, if for every $T \in \mathcal{T}$, it holds that

$$\tilde{\mathbf{u}}(T\mathbf{y}) = \tilde{\mathbf{u}}(\mathbf{y}) \tag{119}$$

Also of interest is invariance with respect to the predicted instead of the known random vector:

A prediction is said to be invariant with respect to the predicted random vector in the class $\mathcal{T}$, if for every $T \in \mathcal{T}$, it holds that

$$\widetilde{T\mathbf{u}}(\mathbf{y}) = T\tilde{\mathbf{u}}(\mathbf{y}) \tag{120}$$

$\mathbf{u}_1, \mathbf{u}_2$

In order to understand the significance of invariance with respect to the predicted random vector consider the case where $\mathbf{u} = [\mathbf{u}_1^T \mathbf{u}_2^T]^T$, and $T\mathbf{u} = T(\mathbf{u}_1, \mathbf{u}_2) = \mathbf{u}_1 + \mathbf{u}_2 \equiv \mathbf{w}$. Lack of invariance in this particular case means that, although $\mathbf{w} = \mathbf{u}_1 + \mathbf{u}_2$, $\tilde{\mathbf{w}}(\mathbf{y}) \neq \tilde{\mathbf{u}}_1(\mathbf{y}) + \tilde{\mathbf{u}}_2(\mathbf{y})$, which poses a serious incompatibility problem.

Among the various classes of transformations, of particular interest are the shift transformations $T\mathbf{a} = \mathbf{a} + \boldsymbol{\delta}$, $\boldsymbol{\delta}$ being a known fixed vector, the regular linear transformations $T\mathbf{a} = \mathbf{R}\mathbf{a}$, $\mathbf{R}$ being a regular matrix ($|\mathbf{R}| \neq 0$), and the affine transformations $T\mathbf{a} = \mathbf{R}\mathbf{a} + \boldsymbol{\delta}$, resulting from the combination of the previous two. Considering the affine transformation

$$\mathbf{y}' = \mathbf{R}\mathbf{y} + \boldsymbol{\delta} \tag{121}$$

it holds that

$$\mathbf{m}_{\mathbf{y}'} = E\{\mathbf{y}'\} = \mathbf{R}\mathbf{m}_{\mathbf{y}} + \boldsymbol{\delta}, \qquad \mathbf{C}_{\mathbf{y}\mathbf{y}'} = \mathbf{R}\mathbf{C}_{\mathbf{y}\mathbf{y}}\mathbf{R}^T, \qquad \mathbf{C}_{\mathbf{u}\mathbf{y}'} = \mathbf{C}_{\mathbf{u}\mathbf{y}}\mathbf{R}^T \tag{122}$$

and application of equations (96), (97), (98) and (99) gives

$$\tilde{\mathbf{u}}(\mathbf{y}') = \tilde{\mathbf{u}}(\mathbf{R}\mathbf{y} + \boldsymbol{\delta}) = \alpha' \mathbf{m}_{\mathbf{u}} + \mathbf{C}_{\mathbf{u}\mathbf{y}}\mathbf{R}^T \mathbf{R}\mathbf{C}_{\mathbf{y}\mathbf{y}}^{-1}\mathbf{R}^T [\mathbf{R}\mathbf{y} + \boldsymbol{\delta} - \alpha'(\mathbf{R}\mathbf{m}_{\mathbf{y}} + \boldsymbol{\delta})] =$$

$$= \alpha' \mathbf{m}_{\mathbf{u}} + \mathbf{C}_{\mathbf{u}\mathbf{y}}\mathbf{C}_{\mathbf{y}\mathbf{y}}^{-1}[\mathbf{y} - \alpha' \mathbf{m}_{\mathbf{y}} + (1-\alpha')\mathbf{R}^T \boldsymbol{\delta}] \tag{123}$$

which is to be compared with (96)

$$\tilde{\mathbf{u}} = \alpha\,\mathbf{m_u} + \mathbf{C_{uy}}\mathbf{C_{uy}^{-1}}(\mathbf{y} - \alpha\,\mathbf{m_y}) \tag{124}$$

with $\alpha$ given by (97), (98) or (99). According to the various types of prediction the random variable $\alpha'$ takes the values

inhomBLIP = inhomBLUP : $\quad \alpha' = 1 = \alpha$ (125)

homBLIP : $$\alpha' = \frac{(\mathbf{m_y} + \mathbf{R}^T\boldsymbol{\delta})^T\mathbf{C_{yy}^{-1}}(\mathbf{y} + \mathbf{R}^T\boldsymbol{\delta})}{1 + (\mathbf{m_y} + \mathbf{R}^T\boldsymbol{\delta})^T\mathbf{C_{yy}^{-1}}(\mathbf{m_y} + \mathbf{R}^T\boldsymbol{\delta})} \tag{126}$$

homBLUP : $$\alpha' = \frac{(\mathbf{m_y} + \mathbf{R}^T\boldsymbol{\delta})^T\mathbf{C_{yy}^{-1}}(\mathbf{y} + \mathbf{R}^T\boldsymbol{\delta})}{(\mathbf{m_y} + \mathbf{R}^T\boldsymbol{\delta})^T\mathbf{C_{yy}^{-1}}(\mathbf{m_y} + \mathbf{R}^T\boldsymbol{\delta})} \tag{127}$$

An obvious result is the following:

InhomBLIP = inhomBLUP is invariant with respect to the known random vector $\mathbf{y}$ in the class of affine transformations $T\mathbf{y} = \mathbf{Ry} + \boldsymbol{\delta}$, i.e.,

$$\tilde{\mathbf{u}}(\mathbf{Ry} + \boldsymbol{\delta}) = \tilde{\mathbf{u}}(\mathbf{y}) \qquad \text{(inhomBLIP = inhomBLUP)} \tag{128}$$

Setting $\boldsymbol{\delta} = \mathbf{0}$ in (126) and (127) it follows that $\alpha' = \alpha$, also for homBLIP and homBLUP and (123) becomes identical with (96).

homBLIP and homBLUP are invariant with respect to the known random vector $\mathbf{y}$ in the class of linear transformations $T\mathbf{y} = \mathbf{Ry}$, i.e.,

$$\tilde{\mathbf{u}}(\mathbf{Ry}) = \tilde{\mathbf{u}}(\mathbf{y}) \qquad \text{(homBLIP \& homBLUP)} \tag{129}$$

On the contrary homBLIP and homBLUP are not invariant with respect to the known random vector $\mathbf{y}$, neither in the class of affine transformations $T\mathbf{y} = \mathbf{Ry} + \boldsymbol{\delta}$, nor in the subclass of shift transformations $T\mathbf{y} = \mathbf{y} + \boldsymbol{\delta}$.

For the uniformly unbiased predictions it holds with respect to the affine transformation (121) that

$$\mathbf{m_{y'}} = \mathbf{Rm_y} + \boldsymbol{\delta} = \mathbf{RAx} + \mathbf{Rt} + \boldsymbol{\delta} \equiv \mathbf{A'x} + \mathbf{t'}, \qquad \mathbf{A'} = \mathbf{RA}, \qquad \mathbf{t'} = \mathbf{Rt} + \boldsymbol{\delta} \tag{130}$$

$$\mathbf{N'} = (\mathbf{A'})^T\mathbf{C_{y'y'}^{-1}}\mathbf{A'} = \mathbf{A}^T\mathbf{R}^T\mathbf{RC_{yy}^{-1}}\mathbf{R}^T\mathbf{RA} = \mathbf{N} \tag{131}$$

$$\mathbf{H'} = \mathbf{I} - \mathbf{A'}(\mathbf{N'})^{-1}(\mathbf{A'})^T\mathbf{C_{y'y'}^{-1}} = \mathbf{I} - \mathbf{RAN}^{-1}\mathbf{A}^T\mathbf{R}^T\mathbf{RC_{yy}^{-1}}\mathbf{R}^T = \mathbf{RHR}^T \tag{132}$$

The prediction based on the transformed known vector becomes

$$\begin{aligned}\tilde{\mathbf{u}}(\mathbf{Ry} + \boldsymbol{\delta}) &= \alpha'\,\mathbf{t}^* + \mathbf{A}^*\mathbf{N}^{-1}\mathbf{A}^T\mathbf{R}^T\mathbf{RC_{yy}^{-1}}\mathbf{R}^T[\mathbf{Ry} + \boldsymbol{\delta} - \alpha'(\mathbf{Rt} + \boldsymbol{\delta})] + \\ &\qquad + \mathbf{C_{uy}}\mathbf{R}^T\mathbf{RC_{yy}^{-1}}\mathbf{R}^T\mathbf{RHR}^T[\mathbf{Ry} + \boldsymbol{\delta} - \alpha'(\mathbf{Rt} + \boldsymbol{\delta})] = \\ &= \alpha'\,\mathbf{t}^* + \mathbf{A}^*\mathbf{N}^{-1}\mathbf{A}^T\mathbf{C_{yy}^{-1}}[\mathbf{y} - \alpha'\,\mathbf{t} + (1-\alpha')\mathbf{R}^T\boldsymbol{\delta}] + \mathbf{C_{uy}}\mathbf{C_{yy}^{-1}}\mathbf{H}[\mathbf{y} - \alpha'\,\mathbf{t} + (1-\alpha')\mathbf{R}^T\boldsymbol{\delta}],\end{aligned} \tag{133}$$

which is to be compared with (101)

$$\tilde{\mathbf{u}} = \alpha\,\mathbf{t}^* + \mathbf{A}^*\mathbf{N}^{-1}\mathbf{A}^T\mathbf{C_{yy}^{-1}}(\mathbf{y} - \alpha\,\mathbf{t}) + \mathbf{C_{uy}}\mathbf{C_{yy}^{-1}}\mathbf{H}(\mathbf{y} - \alpha\,\mathbf{t}) \tag{134}$$

with $\alpha$ given by (104) or (105). According to the particular type of uniformly unbiased prediction the variable $\alpha'$ takes the values

inhomBLUUP : $$\alpha' = 1 \tag{135}$$

homBLUUP : $$\alpha' = \frac{(\mathbf{t}+\mathbf{R}^T\boldsymbol{\delta})^T\mathbf{C}_{\mathbf{yy}}^{-1}\mathbf{H}(\mathbf{y}+\mathbf{R}^T\boldsymbol{\delta})}{(\mathbf{t}+\mathbf{R}^T\boldsymbol{\delta})^T\mathbf{C}_{\mathbf{yy}}^{-1}\mathbf{H}(\mathbf{t}+\mathbf{R}^T\boldsymbol{\delta})} \tag{136}$$

An obvious result is the following:

InhomBLUUP is invariant with respect to the known random vector $\mathbf{y}$ in the class of affine transformations $T\mathbf{y} = \mathbf{R}\mathbf{y} + \boldsymbol{\delta}$, i.e.,

$$\tilde{\mathbf{u}}(\mathbf{R}\mathbf{y}+\boldsymbol{\delta}) = \tilde{\mathbf{u}}(\mathbf{y}) \qquad \text{(inhomBLUUP)} \tag{137}$$

Setting $\boldsymbol{\delta} = \mathbf{0}$ in (136) it follows that $\alpha' = \alpha$, also for homBLUUP and (133) becomes identical with (101).

HomBLUUP is invariant with respect to the known random vector $\mathbf{y}$ in the class of linear transformations $T\mathbf{y} = \mathbf{R}\mathbf{y}$, i.e.,

$$\tilde{\mathbf{u}}(\mathbf{R}\mathbf{y}) = \tilde{\mathbf{u}}(\mathbf{y}) \qquad \text{(homBLUUP)} \tag{138}$$

On the contrary homBLUUP is not invariant with respect to the known random vector $\mathbf{y}$, neither in the class of affine transformations $T\mathbf{y} = \mathbf{R}\mathbf{y} + \boldsymbol{\delta}$, nor in the subclass of shift transformations $T\mathbf{y} = \mathbf{y} + \boldsymbol{\delta}$.

## *3.5. Applications to linear models*

### *3.5.1. The Gauss-Markov linear model*

This is the familiar model from the adjustment with the method of observation equations

$$\mathbf{b} = \mathbf{A}\mathbf{x} + \mathbf{v}, \qquad E\{\mathbf{v}\} = \mathbf{0}, \qquad E\{\mathbf{v}\mathbf{v}^T\} = \mathbf{C} = \sigma^2\mathbf{Q} \tag{139}$$

which in the context of linear regression is sometimes called the fixed effects model, to emphasize that the unknown parameters $\mathbf{x}$ are fixed (i.e. non-random). The residuals $\mathbf{v}$ with $\mathbf{m_v} = \mathbf{0}$ and $\mathbf{C_{vv}} = \mathbf{C} = \sigma^2\mathbf{Q}$ must be predicted from the known observations $\mathbf{b}$ with $\mathbf{m_b} = \mathbf{A}\mathbf{x}$ and $\mathbf{C_{bb}} = \mathbf{C} = \sigma^2\mathbf{Q} = \mathbf{C_{vb}}$. This is a prediction of case (c1), which with the help of the previous results (now $\mathbf{u} \to \mathbf{v}$, $\mathbf{y} \to \mathbf{b}$) admits the following predictions

inhomBLIP = inhomBLUP:

$$\tilde{\mathbf{v}} = \mathbf{b} - \mathbf{A}\mathbf{x} \tag{140}$$

homBLIP:

$$\tilde{\mathbf{v}} = \mathbf{b} - \frac{\mathbf{x}^T\mathbf{A}^T\mathbf{P}\mathbf{b}}{\sigma^2 + \mathbf{x}^T\mathbf{N}\mathbf{x}}\mathbf{A}\mathbf{x} \qquad (\mathbf{P} = \mathbf{Q}^{-1}, \quad \mathbf{N} = \mathbf{A}^T\mathbf{P}\mathbf{A}) \tag{141}$$

homBLUP :

$$\tilde{\mathbf{v}} = \mathbf{b} - \frac{\mathbf{x}^T \mathbf{A}^T \mathbf{P}\mathbf{b}}{\mathbf{x}^T \mathbf{N}\mathbf{x}} \mathbf{A}\mathbf{x} \tag{142}$$

inhomBLUUP = homBLUUP :

$$\tilde{\mathbf{v}} = \mathbf{b} - \mathbf{A}\hat{\mathbf{x}}, \qquad \hat{\mathbf{x}} = \mathbf{N}^{-1}\mathbf{A}^T \mathbf{P}\mathbf{b} \qquad \text{(BLUE)} \tag{143}$$

With respect to compatibility, comparison with estimates in Part 2 shows that inhomBLUUP = homBLUUP is compatible with the unknown true $\mathbf{x}$, which is the inhomBLE = inhomBLUE estimate, homBLIP and homBLUP are compatible with the shrunk estimates

$$\frac{\mathbf{x}^T \mathbf{A}^T \mathbf{P}\mathbf{b}}{\sigma^2 + \mathbf{x}^T \mathbf{N}\mathbf{x}} \mathbf{x} \qquad \text{and} \qquad \frac{\mathbf{x}^T \mathbf{A}^T \mathbf{P}\mathbf{b}}{\mathbf{x}^T \mathbf{N}\mathbf{x}} \mathbf{x} \tag{144}$$

which are the homBLE and homBLUE estimates, respectively. Finally inhomBLIP = inhomBLUP, which is the only prediction that can be actually used, is compatible with the BLUUE (inhomBLUUE) estimate $\hat{\mathbf{x}}$, which is usually referred as BLUE in the literature.

### *3.5.2. The random effects linear model*

This a model with only random parameters $\mathbf{s}$

$$\mathbf{b} = \mathbf{G}\mathbf{s} + \mathbf{v}, \tag{145}$$

with

$$E\{\mathbf{v}\} = \mathbf{0}, \qquad E\{\mathbf{v}\mathbf{v}^T\} = \mathbf{C}, \tag{146}$$

$$E\{\mathbf{s}\} = \mathbf{m}_{\mathbf{s}}, \qquad E\{(\mathbf{s}-\mathbf{m}_{\mathbf{s}})(\mathbf{s}-\mathbf{m}_{\mathbf{s}})^T\} = \mathbf{C}_{\mathbf{ss}}, \qquad E\{(\mathbf{s}-\mathbf{m}_{\mathbf{s}})\mathbf{v}^T\} = \mathbf{C}_{\mathbf{sv}} = \mathbf{0} \tag{147}$$

As a consequence it holds that

$$\mathbf{m}_{\mathbf{b}} = E\{\mathbf{b}\} = \mathbf{G}\mathbf{m}_{\mathbf{s}}, \qquad \mathbf{C}_{\mathbf{bb}} = \mathbf{G}\mathbf{C}_{\mathbf{ss}}\mathbf{G}^T + \mathbf{C} \equiv \mathbf{M} \tag{148}$$

$$\mathbf{C}_{\mathbf{sb}} = \mathbf{C}_{\mathbf{ss}}\mathbf{G}^T, \qquad \mathbf{C}_{\mathbf{vb}} = \mathbf{C}. \tag{149}$$

In addition to $\mathbf{v}$ and $\mathbf{s}$ it is also required to predict a random vector $\mathbf{s}^*$ with

$$E\{\mathbf{s}^*\} = \mathbf{m}_{\mathbf{s}^*}, \qquad E\{(\mathbf{s}^* - \mathbf{m}_{\mathbf{s}^*})(\mathbf{s}-\mathbf{m}_{\mathbf{s}})^T\} = \mathbf{C}_{\mathbf{s}^*\mathbf{s}}, \qquad E\{(\mathbf{s}^* - \mathbf{m}_{\mathbf{s}^*})\mathbf{v}^T\} = \mathbf{C}_{\mathbf{s}^*\mathbf{v}} = \mathbf{0} \tag{150}$$

so that

$$\mathbf{C}_{\mathbf{s}^*\mathbf{b}} = \mathbf{C}_{\mathbf{s}^*\mathbf{s}}\mathbf{G}^T. \tag{151}$$

The predictions follow from the application of the general results (cases b1, b2). Alternative expressions are derived using the matrix identity

$$\mathbf{M}^{-1} = \mathbf{C}^{-1} - \mathbf{C}^{-1}\mathbf{G}\mathbf{N}_G^{-1}\mathbf{G}^T\mathbf{C}^{-1}, \qquad \mathbf{N}_G = \mathbf{G}^T\mathbf{C}^{-1}\mathbf{G} + \mathbf{C}_{\mathbf{ss}}^{-1} \tag{152}$$

inhomBLIP = inhomBLUP, homBLIP and homBLUP

$$\tilde{\mathbf{s}} = \alpha\,\mathbf{m_s} + \mathbf{C_{ss}}\mathbf{G}^T\mathbf{M}^{-1}(\mathbf{b} - \alpha\,\mathbf{Gm_s}) = \mathbf{N}_G^{-1}[\mathbf{G}^T\mathbf{C}^{-1}\mathbf{b} + \mathbf{C_{ss}^{-1}}(\alpha\,\mathbf{m_s})] \tag{153}$$

$$\tilde{\mathbf{v}} = \mathbf{CM}^{-1}(\mathbf{b} - \alpha\,\mathbf{Gm_s}) = [\mathbf{I} - \mathbf{GN}_G^{-1}\mathbf{G}^T\mathbf{C}^{-1}](\mathbf{b} - \alpha\,\mathbf{m_s}) =$$

$$= \mathbf{b} - \mathbf{GN}_G^{-1}[\mathbf{G}^T\mathbf{C}^{-1}\mathbf{b} + C_{\mathbf{ss}}^{-1}(\alpha\,\mathbf{m_s}) = \mathbf{b} - \mathbf{G}\tilde{\mathbf{s}} \tag{154}$$

$$\tilde{\mathbf{s}}^* = \alpha\,\mathbf{m_{s^*}} + \mathbf{C_{s^*s}}\mathbf{G}^T\mathbf{M}^{-1}(\mathbf{b} - \alpha\,\mathbf{Gm_s}) = \alpha\,\mathbf{m_{s^*}} + \mathbf{C_{s^*s}}\mathbf{C_{ss}^{-1}}\mathbf{N}_G^{-1}\mathbf{G}^T\mathbf{C}^{-1}(\mathbf{b} - \alpha\,\mathbf{Gm_s}) =$$

$$= \alpha\,\mathbf{m_{s^*}} + \mathbf{C_{s^*s}}\mathbf{C_{ss}^{-1}}\{\mathbf{N}_G^{-1}[\mathbf{G}^T\mathbf{C}^{-1}\mathbf{b} + \mathbf{C_{ss}^{-1}}(\alpha\,\mathbf{m_s})] - \alpha\,\mathbf{m_s}\} = \alpha\,\mathbf{m_{s^*}} + \mathbf{C_{s^*s}}\mathbf{C_{ss}^{-1}}(\tilde{\mathbf{s}} - \alpha\,\mathbf{m_s}) \tag{155}$$

The parameter $\alpha$ takes the values

$$\alpha = 1 \qquad \text{(inhomBLIP = inhomBLUP)} \tag{156}$$

$$\alpha = \frac{\mathbf{m_s}^T\mathbf{G}^T\mathbf{M}^{-1}\mathbf{b}}{1 + \mathbf{m_s}^T\mathbf{G}^T\mathbf{M}^{-1}\mathbf{Gm_s}} \qquad \text{(homBLIP)} \tag{157}$$

$$\alpha = \frac{\mathbf{m_s}^T\mathbf{G}^T\mathbf{M}^{-1}\mathbf{b}}{\mathbf{m_s}^T\mathbf{G}^T\mathbf{M}^{-1}\mathbf{Gm_s}} \qquad \text{(homBLUP)} \tag{158}$$

### *3.5.3. The random effects model with error bias*

The Gauss-Markov and the random effects linear models can be generalized to the case where the error vector $\mathbf{v}$ has a non-zero mean vector $\mathbf{m_v}$. Only the random effects linear model with such "error bias" will be considered here, because it is needed for later application. The only differences with the previous case are that

$$E\{\mathbf{v}\} = \mathbf{m_v} \neq \mathbf{0} \tag{159}$$

$$E\{\mathbf{b}\} = \mathbf{m_b} = \mathbf{Gm_s} + \mathbf{m_v}\,. \tag{160}$$

The predictions follow from the application of the general results (cases b1, b2). Alternative expressions are derived using the matrix identities (152).

inhomBLIP = inhomBLUP, homBLIP and homBLUP

$$\tilde{\mathbf{s}} = \alpha\,\mathbf{m_s} + \mathbf{C_{ss}}\mathbf{G}^T\mathbf{M}^{-1}(\mathbf{b} - \alpha\,\mathbf{m_v} - \alpha\,\mathbf{Gm_s}) = \mathbf{N}_G^{-1}[\mathbf{G}^T\mathbf{C}^{-1}(\mathbf{b} - \alpha\,\mathbf{m_v}) + \mathbf{C_{ss}^{-1}}(\alpha\,\mathbf{m_s})] \tag{161}$$

$$\tilde{\mathbf{v}} = \alpha\,\mathbf{m_v} + \mathbf{CM}^{-1}(\mathbf{b} - \alpha\,\mathbf{m_v} - \alpha\,\mathbf{Gm_s}) = \mathbf{b} - \mathbf{GN}_G^{-1}[\mathbf{G}^T\mathbf{C}^{-1}(\mathbf{b} - \alpha\,\mathbf{m_v}) + \mathbf{C_{ss}^{-1}}(\alpha\,\mathbf{m_s})] = \mathbf{b} - \mathbf{G}\tilde{\mathbf{s}} \tag{162}$$

$$\tilde{\mathbf{s}}^* = \alpha\,\mathbf{m_{s^*}} + c\mathbf{G}^T\mathbf{M}^{-1}(\mathbf{b} - \alpha\,\mathbf{m_v} - \alpha\,\mathbf{Gm_s}) =$$

$$= \alpha\,\mathbf{m_{s^*}} + \mathbf{C_{s^*s}}\mathbf{C_{ss}^{-1}}\{\mathbf{N}_G^{-1}[\mathbf{G}^T\mathbf{C}^{-1}(\mathbf{b} - \alpha\,\mathbf{m_v}) + \mathbf{C_{ss}^{-1}}(\alpha\,\mathbf{m_s})] - \alpha\,\mathbf{m_s}\} = \alpha\,\mathbf{m_{s^*}} + \mathbf{C_{s^*s}}\mathbf{C_{ss}^{-1}}(\tilde{\mathbf{s}} - \alpha\,\mathbf{m_s}) \tag{163}$$

The parameter $\alpha$ takes the values

$$\alpha = 1 \qquad \text{(inhomBLIP = inhomBLUP)} \tag{164}$$

$$\alpha = \frac{(\mathbf{Gm_s} + \mathbf{m_v})^T \mathbf{M}^{-1}\mathbf{b}}{1 + (\mathbf{Gm_s} + \mathbf{m_v})^T \mathbf{M}^{-1}(\mathbf{Gm_s} + \mathbf{m_v})} \qquad \text{(homBLIP)} \qquad (165)$$

$$\alpha = \frac{(\mathbf{Gm_s} + \mathbf{m_v})^T \mathbf{M}^{-1}\mathbf{b}}{(\mathbf{Gm_s} + \mathbf{m_v})^T \mathbf{M}^{-1}(\mathbf{Gm_s} + \mathbf{m_v})} \qquad \text{(homBLUP)} \qquad (166)$$

### *3.5.4. The mixed linear model*

This is a model with both stochastic and deterministic parameters

$$\mathbf{b} = \mathbf{Ax} + \mathbf{Gs} + \mathbf{v}, \qquad (167)$$

with

$$E\{\mathbf{v}\} = \mathbf{0}, \qquad E\{\mathbf{vv}^T\} = \mathbf{C}, \qquad (168)$$

$$E\{\mathbf{s}\} = \mathbf{m_s}, \qquad E\{(\mathbf{s}-\mathbf{m_s})(\mathbf{s}-\mathbf{m_s})^T\} = \mathbf{C_{ss}}, \qquad E\{(\mathbf{s}-\mathbf{m_s})\mathbf{v}^T\} = \mathbf{C_{sv}} = \mathbf{0}. \qquad (169)$$

As a consequence it holds that

$$\mathbf{m_b} = E\{\mathbf{b}\} = \mathbf{Ax} + \mathbf{Gm_s}, \qquad \mathbf{C_{bb}} = \mathbf{GC_{ss}G}^T + \mathbf{C} \equiv \mathbf{M} \qquad (170)$$

$$\mathbf{C_{sb}} = \mathbf{C_{ss}G}^T, \qquad \mathbf{C_{vb}} = \mathbf{C}. \qquad (171)$$

In addition to $\mathbf{v}$ and $\mathbf{s}$ it is also required to predict a random vector $\mathbf{s}^*$ with

$$E\{\mathbf{s}^*\} = \mathbf{m_{s^*}}, \qquad E\{(\mathbf{s}^* - \mathbf{m_{s^*}})(\mathbf{s}-\mathbf{m_s})^T\} = \mathbf{C_{s^*s}}, \qquad E\{(\mathbf{s}^* - \mathbf{m_{s^*}})\mathbf{v}^T\} = \mathbf{C_{s^*v}} = \mathbf{0} \qquad (172)$$

so that

$$\mathbf{C_{s^*b}} = \mathbf{C_{s^*s}G}^T \qquad (173)$$

as well as a random parameter $\mathbf{u} = \mathbf{A}^*\mathbf{x} + \mathbf{s}^* + \mathbf{h}$, $\mathbf{h}$ being constant and known so that

$$\mathbf{m_u} = E\{\mathbf{u}\} = \mathbf{A}^*\mathbf{x} + \mathbf{m_{s^*}} + \mathbf{h}, \qquad \mathbf{C_{ub}} = \mathbf{C_{s^*s}G}^T. \qquad (174)$$

#### *3.5.4.1. General predictions*

The predictions follow from the application of the general results (cases d1, d2, d3, d4), with $\mathbf{t} = \mathbf{Gm_s}$, using again the matrix identities (152).

<u>inhomBLIP = inhomBLUP, homBLIP and homBLUP</u>

$$\tilde{\mathbf{s}} = a\mathbf{m_s} + \mathbf{C_{ss}G}^T\mathbf{M}^{-1}[\mathbf{b} - \alpha(\mathbf{Ax} + \mathbf{Gm_s})] = \mathbf{N}_G^{-1}[\mathbf{G}^T\mathbf{C}^{-1}(\mathbf{b} - \alpha\,\mathbf{Ax}) + \mathbf{C_{ss}}^{-1}(\alpha\,\mathbf{m_s})] \qquad (175)$$

$$\tilde{\mathbf{v}} = \mathbf{CM}^{-1}[\mathbf{b} - \alpha(\mathbf{Ax} + \mathbf{Gm_s})] =$$

$$= \mathbf{b} - \alpha\,\mathbf{Ax} - \mathbf{G}\mathbf{N}_G^{-1}[\mathbf{G}^T\mathbf{C}^{-1}(\mathbf{b} - \alpha\,\mathbf{Ax}) + \mathbf{C}_{\mathbf{ss}}^{-1}(\alpha\,\mathbf{m}_{\mathbf{s}})] = \mathbf{b} - \alpha\,\mathbf{Ax} - \mathbf{G}\tilde{\mathbf{s}} \tag{176}$$

$$\tilde{\mathbf{s}}^* = a\mathbf{m}_{\mathbf{s}^*} + \mathbf{C}_{\mathbf{s}^*\mathbf{s}}\mathbf{G}^T\mathbf{M}^{-1}[\mathbf{b} - \alpha\,(\mathbf{Ax} + \mathbf{Gm}_{\mathbf{s}})] =$$

$$= a\mathbf{m}_{\mathbf{s}^*} + \mathbf{C}_{\mathbf{s}^*\mathbf{s}}\mathbf{C}_{\mathbf{ss}}^{-1}\{\mathbf{N}_G^{-1}[\mathbf{G}^T\mathbf{C}^{-1}(\mathbf{b} - \alpha\,\mathbf{Ax}) + \mathbf{C}_{\mathbf{ss}}^{-1}(\alpha\,\mathbf{m}_{\mathbf{s}})] - \alpha\,\mathbf{m}_{\mathbf{s}}\} = a\mathbf{m}_{\mathbf{s}^*} + \mathbf{C}_{\mathbf{s}^*\mathbf{s}}\mathbf{C}_{\mathbf{ss}}^{-1}(\tilde{\mathbf{s}} - \alpha\,\mathbf{m}_{\mathbf{s}}) \tag{177}$$

$$\tilde{\mathbf{u}} = \alpha\,(\mathbf{A}^*\mathbf{x} + \mathbf{m}_{\mathbf{s}^*} + \mathbf{h}) + \mathbf{C}_{\mathbf{s}^*\mathbf{s}}\mathbf{G}^T\mathbf{M}^{-1}[\mathbf{b} - \alpha\,(\mathbf{Ax} + \mathbf{Gm}_{\mathbf{s}})] =$$

$$= \alpha\,(\mathbf{A}^*\mathbf{x} + \mathbf{m}_{\mathbf{s}^*} + \mathbf{h}) + \mathbf{C}_{\mathbf{s}^*\mathbf{s}}\mathbf{C}_{\mathbf{ss}}^{-1}\{\mathbf{N}_G^{-1}[\mathbf{G}^T\mathbf{C}^{-1}[\mathbf{b} - \alpha\,\mathbf{Ax}) + \mathbf{C}_{\mathbf{ss}}^{-1}(\alpha\,\mathbf{m}_{\mathbf{s}})] - \alpha\,\mathbf{m}_{\mathbf{s}}\} =$$

$$= \alpha\,(\mathbf{A}^*\mathbf{x} + \mathbf{m}_{\mathbf{s}^*} + \mathbf{h}) + \mathbf{C}_{\mathbf{s}^*\mathbf{s}}\mathbf{C}_{\mathbf{ss}}^{-1}(\tilde{\mathbf{s}} - \alpha\,\mathbf{m}_{\mathbf{s}}) = \mathbf{A}^*(\alpha\,\mathbf{x}) + \tilde{\mathbf{s}}^* + \alpha\,\mathbf{h} \tag{178}$$

The parameter $\alpha$ takes the values

$$\alpha = 1 \qquad \text{(inhomBLIP = inhomBLUP)} \tag{179}$$

$$\alpha = \frac{(\mathbf{Ax} + \mathbf{Gm}_{\mathbf{s}})^T\mathbf{M}^{-1}\mathbf{b}}{1 + (\mathbf{Ax} + \mathbf{Gm}_{\mathbf{s}})^T\mathbf{M}^{-1}(\mathbf{Ax} + \mathbf{Gm}_{\mathbf{s}})} \qquad \text{(homBLIP)} \tag{180}$$

$$\alpha = \frac{(\mathbf{Ax} + \mathbf{Gm}_{\mathbf{s}})^T\mathbf{M}^{-1}\mathbf{b}}{(\mathbf{Ax} + \mathbf{Gm}_{\mathbf{s}})^T\mathbf{M}^{-1}(\mathbf{Ax} + \mathbf{Gm}_{\mathbf{s}})} \qquad \text{(homBLUP)} \tag{181}$$

From the compatibility condition $\mathbf{b} = \mathbf{Ax} + \mathbf{G}\tilde{\mathbf{s}} + \tilde{\mathbf{v}}$ it follows after comparison with the estimates in Part 2, that the inhomBLIP = inhomBLUP prediction is compatible with the true unknown value $\mathbf{x}$, which is the inhomBLE=inhomBLUE estimate, while homBLIP and homBLUP are compatible with the shrunk estimate $\mathbf{x}_\alpha = \alpha\,\mathbf{x}$, which are the homBLE and homBLUE estimates, respectively.

<u>inhomBLUUP and homBLUUP</u>

The following auxiliary notation will be used

$$\mathbf{N} = \mathbf{A}^T\mathbf{M}^{-1}\mathbf{A} \tag{182}$$

$$\mathbf{H} = \mathbf{I} - \mathbf{A}\mathbf{N}^{-1}\mathbf{A}^T\mathbf{M}^{-1} \tag{183}$$

$$\mathbf{x}_\alpha = \mathbf{N}^{-1}\mathbf{A}^T\mathbf{M}^{-1}(\mathbf{b} - \alpha\,\mathbf{Gm}_{\mathbf{s}})\,. \tag{184}$$

The predictions are

$$\tilde{\mathbf{s}} = a\mathbf{m}_{\mathrm{s}} + \mathbf{C}_{\mathrm{ss}}\mathbf{G}^T\mathbf{M}^{-1}\mathbf{H}(\mathbf{b} - \alpha\,\mathbf{Gm}_{\mathbf{s}}) = a\mathbf{m}_{\mathrm{s}} + \mathbf{C}_{\mathrm{ss}}\mathbf{G}^T\mathbf{M}^{-1}(\mathbf{b} - \mathbf{Ax}_\alpha - \alpha\,\mathbf{Gm}_{\mathbf{s}}) =$$

$$= \mathbf{N}_G^{-1}[\mathbf{G}^T\mathbf{C}^{-1}(\mathbf{b} - \mathbf{Ax}_\alpha) + \mathbf{C}_{\mathbf{ss}}^{-1}(\alpha\,\mathbf{m}_{\mathbf{s}})] \tag{185}$$

$$\tilde{\mathbf{v}} = \mathbf{CM}^{-1}\mathbf{H}(\mathbf{b} - \alpha\,\mathbf{Gm}_{\mathbf{s}}) = \mathbf{CM}^{-1}(\mathbf{b} - \mathbf{Ax}_\alpha - \alpha\,\mathbf{Gm}_{\mathbf{s}}) =$$

$$= \mathbf{b} - \mathbf{Ax}_\alpha - \mathbf{G}\mathbf{N}_G^{-1}[\mathbf{G}^T\mathbf{C}^{-1}(\mathbf{b} - \mathbf{Ax}_\alpha) + \mathbf{C}_{\mathbf{ss}}^{-1}(\alpha\,\mathbf{m}_{\mathbf{s}})] = \mathbf{b} - \mathbf{Ax}_\alpha - \mathbf{G}\tilde{\mathbf{s}} \tag{186}$$

$$\tilde{\mathbf{s}}^* = \alpha\,\mathbf{m}_{\mathbf{s}^*} + \mathbf{C}_{\mathbf{s}^*\mathbf{s}}\mathbf{G}^T\mathbf{M}^{-1}\mathbf{H}(\mathbf{b} - \alpha\,\mathbf{G}\mathbf{m}_\mathbf{s}) = a\mathbf{m}_{\mathbf{s}^*} + \mathbf{C}_{\mathbf{s}^*\mathbf{s}}\mathbf{G}^T\mathbf{M}^{-1}(\mathbf{b} - \mathbf{A}\mathbf{x}_\alpha - \alpha\,\mathbf{G}\mathbf{m}_\mathbf{s}) =$$

$$= \alpha\,\mathbf{m}_{\mathbf{s}^*} + \mathbf{C}_{\mathbf{s}^*\mathbf{s}}\mathbf{C}_{\mathbf{ss}}^{-1}\{\mathbf{N}_G^{-1}[\mathbf{G}^T\mathbf{C}^{-1}(\mathbf{b} - \alpha\,\mathbf{A}\mathbf{x}) + \mathbf{C}_{\mathbf{ss}}^{-1}(\alpha\,\mathbf{m}_\mathbf{s})] - \alpha\,\mathbf{m}_\mathbf{s}\} =$$

$$= \alpha\,\mathbf{m}_{\mathbf{s}^*} + \mathbf{C}_{\mathbf{s}^*\mathbf{s}}\mathbf{C}_{\mathbf{ss}}^{-1}(\tilde{\mathbf{s}} - \alpha\,\mathbf{m}_\mathbf{s}) \tag{187}$$

$$\tilde{\mathbf{u}} = \alpha\,(\mathbf{m}_{\mathbf{s}^*} + \mathbf{h}) + \mathbf{A}^*\mathbf{N}^{-1}\mathbf{A}^T\mathbf{M}^{-1}(\mathbf{b} - \alpha\,\mathbf{G}\mathbf{m}_\mathbf{s}) + \mathbf{C}_{\mathbf{s}^*\mathbf{s}}\mathbf{G}^T\mathbf{M}^{-1}\mathbf{H}(\mathbf{b} - \alpha\,\mathbf{G}\mathbf{m}_\mathbf{s}) =$$

$$= \alpha\,\mathbf{m}_{\mathbf{s}^*} + \alpha\,\mathbf{h} + \mathbf{A}^*\mathbf{x}_\alpha + \mathbf{C}_{\mathbf{s}^*\mathbf{s}}\mathbf{G}^T\mathbf{M}^{-1}(\mathbf{b} - \mathbf{A}\mathbf{x}_\alpha - \alpha\,\mathbf{G}\mathbf{m}_\mathbf{s}\} =$$

$$= \alpha\,\mathbf{m}_{\mathbf{s}^*} + \alpha\,\mathbf{h} + \mathbf{A}^*\mathbf{x}_\alpha + \mathbf{C}_{\mathbf{s}^*\mathbf{s}}\mathbf{C}_{\mathbf{ss}}^{-1}\{\mathbf{N}_G^{-1}[\mathbf{G}^T\mathbf{C}^{-1}(\mathbf{b} - \alpha\,\mathbf{A}\mathbf{x}) + \mathbf{C}_{\mathbf{ss}}^{-1}(\alpha\,\mathbf{m}_\mathbf{s})] - \alpha\,\mathbf{m}_\mathbf{s}\} =$$

$$= \mathbf{A}^*\mathbf{x}_\alpha + \tilde{\mathbf{s}}^* + \alpha\,\mathbf{h} \tag{188}$$

The parameter $\alpha$ takes the values

$$\alpha = 1 \qquad \text{(inhomBLUUP)} \tag{189}$$

$$\alpha = \frac{\mathbf{m}_\mathbf{s}^T\mathbf{G}^T\mathbf{M}^{-1}\mathbf{H}\mathbf{b}}{\mathbf{m}_\mathbf{s}^T\mathbf{G}^T\mathbf{M}^{-1}\mathbf{H}\mathbf{G}\mathbf{m}_\mathbf{s}} \qquad \text{(homBLUUP)} \tag{190}$$

From the compatibility condition it follows that the predictions are compatible with the estimate $\mathbf{x}_\alpha$. Comparison with the estimates in Part 2 shows that inhomBLUUP is compatible with the familiar (inhom)BLUUE estimate of $\mathbf{x}$

$$\mathbf{x}_\alpha = \mathbf{N}^{-1}\mathbf{A}^T\mathbf{M}^{-1}(\mathbf{b} - \mathbf{G}\mathbf{m}_\mathbf{s}) = \hat{\mathbf{x}}\,. \tag{191}$$

while homBLUUP is compatible with the homBLUUE estimate.

### *3.5.4.2. BLUUE-compatible predictions*

From all general types of predictions, those with the less restrictions (inhomBLIP = inhomBLUP, homBLIP, homBLUP) are useless since they depend on the unknown $\mathbf{x}$, while from the uniformly unbiased ones, which are independent of $\mathbf{x}$, only inhomBLUUP is compatible with the familiar BLUUE (inhomBLUUE) estimate $\hat{\mathbf{x}}$ of $\mathbf{x}$. If the BLUUE estimate is to be used, one must search for predictions which are BLUUE-compatible, i.e. such that the compatibility condition

$$\mathbf{b} = \mathbf{A}\hat{\mathbf{x}} + \mathbf{G}\tilde{\mathbf{s}} + \tilde{\mathbf{v}} \tag{192}$$

is satisfied. One way to treat this problem (Schaffrin, 1985) is to enforce $\mathbf{x} = \hat{\mathbf{x}}$ in the mixed model, which with the use of

$$\hat{\mathbf{x}} = \mathbf{N}^{-1}\mathbf{A}^T\mathbf{M}^{-1}(\mathbf{b} - \mathbf{G}\mathbf{m}_\mathbf{s}) \tag{193}$$

becomes

$$(\mathbf{I} - \mathbf{A}\mathbf{N}^{-1}\mathbf{A}^T\mathbf{M}^{-1})\mathbf{b} = -\mathbf{A}\mathbf{N}^{-1}\mathbf{A}^T\mathbf{M}^{-1}\mathbf{G}\mathbf{m}_\mathbf{s} + \mathbf{G}\mathbf{s} + \mathbf{v} \tag{194}$$

Using the notation $\mathbf{H} = \mathbf{I} - \mathbf{A}\mathbf{N}^{-1}\mathbf{A}^T\mathbf{M}^{-1}$ the model becomes

$$\mathbf{r} \equiv \mathbf{Hb} = (\mathbf{H}-\mathbf{I})\mathbf{Gm}_{\mathbf{s}} + \mathbf{Gs} + \mathbf{v} \tag{195}$$

with

$$\mathbf{m}_{\mathbf{r}} = E\{\mathbf{r}\} = \mathbf{HGm}_{\mathbf{s}}, \qquad \mathbf{C}_{\mathbf{rr}} = \mathbf{GC}_{\mathbf{ss}}\mathbf{G}^T + \mathbf{C} = \mathbf{M} \tag{196}$$

$$\mathbf{C}_{\mathbf{sr}} = \mathbf{C}_{\mathbf{ss}}\mathbf{G}^T, \qquad \mathbf{C}_{\mathbf{vr}} = \mathbf{C}, \qquad \mathbf{C}_{\mathbf{s}^*\mathbf{r}} = \mathbf{C}_{\mathbf{s}^*\mathbf{s}}\mathbf{G}^T \tag{197}$$

Application of the general results leads directly to the following predictions, where the notation

$$\mathbf{m} = \mathbf{N}^{-1}\mathbf{A}^T\mathbf{M}^{-1}\mathbf{Gm}_{\mathbf{s}} \tag{198}$$

has been used.

<u>inhomBLIP = inhomBLUP, homBLIP and homBLUP</u>

$$\tilde{\mathbf{s}} = \alpha\,\mathbf{m}_{\mathbf{s}} + \mathbf{C}_{\mathbf{ss}}\mathbf{G}^T\mathbf{M}^{-1}\mathbf{H}(\mathbf{b} - \alpha\,\mathbf{Gm}_{\mathbf{s}}) = \alpha\,\mathbf{m}_{\mathbf{s}} + \mathbf{C}_{\mathbf{ss}}\mathbf{G}^T\mathbf{M}^{-1}[\mathbf{b} - \mathbf{A}\hat{\mathbf{x}} - \alpha\,\mathbf{Gm}_{\mathbf{s}} - (1-\alpha)\mathbf{Am}] =$$

$$= \mathbf{N}_G^{-1}\{\mathbf{G}^T\mathbf{C}^{-1}[\mathbf{b} - \mathbf{A}\hat{\mathbf{x}} - (1-\alpha)\mathbf{Am}] + \mathbf{C}_{\mathbf{ss}}^{-1}(\alpha\,\mathbf{m}_{\mathbf{s}})\}. \tag{199}$$

$$\tilde{\mathbf{v}} = \mathbf{CM}^{-1}\mathbf{H}(\mathbf{b} - \alpha\,\mathbf{Gm}_{\mathbf{s}}) = \mathbf{CM}^{-1}[\mathbf{b} - \mathbf{A}\hat{\mathbf{x}} - \alpha\,\mathbf{Gm}_{\mathbf{s}} - (1-\alpha)\mathbf{Am}] =$$

$$= \mathbf{b} - \mathbf{A}\hat{\mathbf{x}} - (1-\alpha)\mathbf{Am} - \mathbf{GN}_G^{-1}\{\mathbf{G}^T\mathbf{C}^{-1}[\mathbf{b} - \mathbf{A}\hat{\mathbf{x}} - (1-\alpha)\mathbf{Am}] + \mathbf{C}_{\mathbf{ss}}^{-1}(\alpha\,\mathbf{m}_{\mathbf{s}})\} =$$

$$= \mathbf{b} - \mathbf{A}\hat{\mathbf{x}} - (1-\alpha)\mathbf{Am} - \mathbf{G}\tilde{\mathbf{s}}, \tag{200}$$

$$\tilde{\mathbf{s}}^* = \alpha\,\mathbf{m}_{\mathbf{s}^*} + \mathbf{C}_{\mathbf{s}^*\mathbf{s}}\mathbf{G}^T\mathbf{M}^{-1}\mathbf{H}(\mathbf{b} - \alpha\,\mathbf{Gm}_{\mathbf{s}}) =$$

$$= \alpha\,\mathbf{m}_{\mathbf{s}^*} + \mathbf{C}_{\mathbf{s}^*\mathbf{s}}\mathbf{G}^T\mathbf{M}^{-1}[\mathbf{b} - \mathbf{A}\hat{\mathbf{x}} - \alpha\,\mathbf{Gm}_{\mathbf{s}} - (1-\alpha)\mathbf{Am}] =$$

$$= \alpha\,\mathbf{m}_{\mathbf{s}^*} + \mathbf{C}_{\mathbf{s}^*\mathbf{s}}\mathbf{C}_{\mathbf{ss}}^{-1}\{\mathbf{N}_G^{-1}[\mathbf{G}^T\mathbf{C}^{-1}(\mathbf{b} - \mathbf{A}\hat{\mathbf{x}} - (1-\alpha)\mathbf{Am}) + \mathbf{C}_{\mathbf{ss}}^{-1}(\alpha\,\mathbf{m}_{\mathbf{s}})] - \alpha\,\mathbf{m}_{\mathbf{s}}\} =$$

$$= \alpha\,\mathbf{m}_{\mathbf{s}^*} + \mathbf{C}_{\mathbf{s}^*\mathbf{s}}\mathbf{C}_{\mathbf{ss}}^{-1}(\tilde{\mathbf{s}} - \alpha\,\mathbf{m}_{\mathbf{s}}), \tag{201}$$

$$\tilde{\mathbf{u}} = \alpha\,(\mathbf{A}^*\mathbf{x} + \mathbf{m}_{\mathbf{s}^*} + \mathbf{h}) + \mathbf{C}_{\mathbf{s}^*\mathbf{s}}\mathbf{G}^T\mathbf{M}^{-1}\mathbf{H}(\mathbf{b} - \alpha\,\mathbf{Gm}_{\mathbf{s}}) =$$

$$= \alpha\,(\mathbf{A}^*\mathbf{x} + \mathbf{m}_{\mathbf{s}^*} + \mathbf{h}) + \mathbf{C}_{\mathbf{s}^*\mathbf{s}}\mathbf{C}_{\mathbf{ss}}^{-1}\{\mathbf{N}_G^{-1}[\mathbf{G}^T\mathbf{C}^{-1}(\mathbf{b} - \mathbf{A}\hat{\mathbf{x}} - (1-\alpha)\mathbf{Am}) + \mathbf{C}_{\mathbf{ss}}^{-1}(\alpha\,\mathbf{m}_{\mathbf{s}})] - \alpha\,\mathbf{m}_{\mathbf{s}}\} =$$

$$= \alpha\,(\mathbf{A}^*\mathbf{x} + \mathbf{m}_{\mathbf{s}^*} + \mathbf{h}) + \mathbf{C}_{\mathbf{s}^*\mathbf{s}}\mathbf{C}_{\mathbf{ss}}^{-1}(\tilde{\mathbf{s}} - \alpha\,\mathbf{m}_{\mathbf{s}}) = \mathbf{A}^*(\alpha\,\mathbf{x}) + \tilde{\mathbf{s}}^* + \alpha\,\mathbf{h}. \tag{202}$$

The parameter $\alpha$ takes the values

$$\alpha = 1 \qquad \text{(BLUUE-inhomBLIP = BLUUE-inhomBLUP)} \tag{203}$$

$$\alpha = \frac{\mathbf{m}_{\mathbf{s}}^T\mathbf{G}^T\mathbf{M}^{-1}\mathbf{Hb}}{1 + \mathbf{m}_{\mathbf{s}}^T\mathbf{G}^T\mathbf{M}^{-1}\mathbf{HGm}_{\mathbf{s}}} \qquad \text{(BLUUE-homBLIP)} \tag{204}$$

$$\alpha = \frac{\mathbf{m}_\mathbf{s}^T \mathbf{G}^T \mathbf{M}^{-1} \mathbf{Hb}}{\mathbf{m}_\mathbf{s}^T \mathbf{G}^T \mathbf{M}^{-1} \mathbf{HGm}_\mathbf{s}} \qquad \text{(BLUUE-homBLUP)} \qquad (205)$$

These results, and especially those for $\hat{\mathbf{x}}$, $\tilde{\mathbf{s}}$ and $\tilde{\mathbf{v}}$ which are identical with those of **Schaffrin** (1985), are not compatible, despite the use of the modified "compatibility model" (195). It can be directly seen from the last of (200) that the original model (167) is not satisfied by $\hat{\mathbf{x}}$, $\tilde{\mathbf{s}}$ and $\tilde{\mathbf{v}}$, due to the presence of the additional "incompatibility term" $(1-\alpha)\mathbf{Am}$. Only for the inhomBLIP = inhomBLUP prediction where $\alpha = 1$ the incompatibility term vanishes and $\hat{\mathbf{x}}$, $\tilde{\mathbf{s}}$ and $\tilde{\mathbf{v}}$ are BLUUE compatible.

In order to obtain BLUUE compatible solutions the model (195) will be modified introducing

$$\mathbf{e} = (\mathbf{H} - \mathbf{I})\mathbf{Gm}_\mathbf{s} + \mathbf{v} \qquad (207)$$

so that a new model of type "random effects with error bias" results

$$\mathbf{r} \equiv \mathbf{Hb} = \mathbf{Gs} + \mathbf{e} \,. \qquad (208)$$

All predictions from this model are identical with those from model (195) with the exception of $\tilde{\mathbf{v}}$, which will be derived from $\tilde{\mathbf{e}}$ by applying (207) with $\tilde{\mathbf{v}}$ and $\tilde{\mathbf{e}}$ in place of $\mathbf{v}$ and $\mathbf{e}$. Since

$$\mathbf{m}_\mathbf{e} = E\{\mathbf{e}\} = (\mathbf{H} - \mathbf{I})\mathbf{Gm}_\mathbf{s} \,, \qquad \mathbf{C}_\mathbf{er} = \mathbf{C} \qquad (209)$$

the prediction of **e** according to the general results becomes

$$\tilde{\mathbf{e}} = \alpha\,(\mathbf{H} - \mathbf{I})\mathbf{Gm}_\mathbf{s} + \mathbf{CM}^{-1}\mathbf{H}(\mathbf{b} - \alpha\,\mathbf{Gm}_\mathbf{s}) \qquad (210)$$

and the corresponding prediction of $\mathbf{v}$ is

$$\begin{aligned} \tilde{\mathbf{v}} = \tilde{\mathbf{e}} - (\mathbf{H} - \mathbf{I})\mathbf{Gm}_\mathbf{s} &= \mathbf{CM}^{-1}\mathbf{H}(\mathbf{b} - \alpha\,\mathbf{Gm}_\mathbf{s}) + (1-\alpha)(\mathbf{I} - \mathbf{H})\mathbf{Gm}_\mathbf{s} = \\ &= \mathbf{CM}^{-1}[\mathbf{b} - \mathbf{A}\hat{\mathbf{x}} - \alpha\,\mathbf{Gm}_\mathbf{s} - (1-\alpha)\mathbf{Am}] + (1-\alpha)\mathbf{Am} = \\ &= \mathbf{b} - \mathbf{A}\hat{\mathbf{x}} - \mathbf{GN}_G^{-1}\{\mathbf{G}^T\mathbf{C}^{-1}[\mathbf{b} - \mathbf{A}\hat{\mathbf{x}} - (1-\alpha)\mathbf{Am}] + \mathbf{C}_\mathbf{ss}^{-1}(\alpha\,\mathbf{m}_\mathbf{s})\} = \mathbf{b} - \mathbf{A}\hat{\mathbf{x}} - \mathbf{G}\tilde{\mathbf{s}} \end{aligned} \qquad (211)$$

With (211) replacing (200), a set of BLUUE compatible predictions has been obtained also for the homBLIP and homBLUP predictions. In the inhomBLIP = inhomBLUP case (200) and (211) are identical.

Comparison of the results of this section with those of the previous one shows that BLUUE-inhomBLIP = BLUUE-inhomBLUP gives identical results with inhomBLUUP, and BLUUE-homBLUP with homBLUUP. The BLUUE-homBLIP predictions are not identical with any previous results.

It must be emphasized that it has not been shown that the predictions of this section are those optimal in the classes which result from the intersection of the inhomogeneous, inhomogeneous-unbiased, homogeneous and homogeneous-unbiased, with the class of BLUUE compatible predictions. The latter can be determined only for $[\mathbf{s}^T \mathbf{v}^T]^T$, by minimizing the trace of its MSE matrix under the simultaneous satisfaction of the BLUUE compatibility condition and the condition for unbiasedness when applicable. The compatibility condition in this case depends on the unknown true values of $\mathbf{x}$ and $\mathbf{v}$. The concept of uniform compatibility for every value of $\mathbf{v}$ (or both $\mathbf{v}$ and $\mathbf{x}$) must be introduced if predictions independent of $\mathbf{v}$ (or both $\mathbf{v}$ and $\mathbf{x}$) are to be obtained, in the same way that the concept of uniform unbiasedness leads to predictions independent of $\mathbf{x}$. This approach will not be followed here because of its limitation to only $\mathbf{s}$ and $\mathbf{v}$. On the contrary our unified approach yields predictions for any random variable which is stochastically related to the known ones (observations) and therefore also of $\mathbf{s}$ and $\mathbf{v}$.

## *References*

Bibby, J. & H. Toutenburg (1977): Prediction and Improved Estimation in Linear Models. Wiley.

Graybill, F.A. (1976): Theory and Application of the Linear Model. Duxbury Press, North Scituate, Mass.

Rao, C.R. (1965): Linear Statistical Inference and its Applications. Wiley.

Schaffrin, B. (1985): Das geodätische Datum mit stochastischer Vorinformation. Habilitationsschrift, DGK, Reihe C, Heft Nr. 313.

Toutenburg, H. (1982): Prior Information in Linear Models. Wiley.