A Unified Approach to Linear Estimation and Prediction

Athanasios Dermanis

Department of Geodesy and Surveying
University of Thessaloniki

XXth Congress of the International Union of Geodesy and Geophysics
International Association of Geodesy
Vienna, August 11–24, 1991

CONTENTS

Introduction 1

Part 1: The model concept 2

1.1. General and specific model 2
1.2. Characteristics of the general model 2
1.3. The general model in relation to the available data 4
1.4. Alternative specific models corresponding to the same general model 5
1.5. Models with stochastic parameters 7
1.6. Models with prior information 7
1.7. Linearized specific models 8

Part 2: Linear estimation 14

2.1. The alternative types of linear estimation 14
2.2. The general solution to the estimation problem 14
2.3. General solution for simple models 16
2.4. General solution for models with stochastic parameters 17
2.4.1. Models with stochastic parameters - type A 17
2.4.2. Models with stochastic parameters - type B 21
2.5. Summary and discussion 22

Appendix 2A: Derivation of the general estimation solution 28

Part 3: Linear Prediction 32

3.1. The prediction concept 32
3.2. Optimality criteria and prediction types 33
3.3. Best (minimum mean square error) prediction 35
3.4. Inhomogeneous and homogeneous linear prediction 36
3.4.1. Best linear inhomogeneous prediction (inhomBLIP) 37
3.4.2. Best linear homogeneous prediction (homBLIP) 37
3.4.3. Best linear inhomogeneous unbiased prediction (inhomBLUP) 38
3.4.4. Best linear homogeneous unbiased prediction (homBLUP) 39
3.4.5. Uniformly unbiased predictions 40
3.4.5.1. Best Linear inhomogeneous Uniformly Unbiased Prediction (inhomBLUUP) 41
3.4.5.2. Best Linear homogeneous Uniformly Unbiased Prediction (homBLUUP) 42
3.4.6. Vector generalizations and summary 43
3.4.7. MSE matrices and biases 44
3.4.8. Invariance characteristics of the various prediction types 46
3.5. Applications to linear models 48
3.5.1. The Gauss-Markov linear model 48
3.5.2. The random effects linear model 49
3.5.3. The random effects model with error bias 50
3.5.4. The mixed linear model 51
3.5.4.1. General predictions 51
3.5.4.2. BLUE-compatible predictions 53

References 56

Introduction

Geodetic data analysis is dominated by the use of linear techniques, where estimates of unknown deterministic parameters and predictions of stochastic ones are sought in the form of linear functions of the available data. Furthermore, data and model parameters are connected through linear relations (models) resulting from the linearization of original non-linear ones.

Instead of examining separately the various linear models, a unified approach is taken here, where a general solution is given to the estimation or prediction of any parameter, without reference to any particular model. This general solution can be directly applied in order to produce the final solution to any particular model.

Six different types of best (minimum mean square error) estimates and predictions are considered by combining the possibility of having homogeneous or inhomogeneous functions of the data, with no restriction, restriction for unbiased solutions and restriction for uniformly unbiased ones. These include the classical solutions, together with some "robust" alternatives, as well as solutions, which have no direct practical use because they depend on unknown parameters.

In Part 1, the model concept it is investigated and it is shown how the various equivalent specific models may result from the same general model. General models are divided in three classes, the simple ones, without stochastic parameters other than the observational errors, and two classes of models with stochastic parameters, one arising from the use of stochastic model equations, and the other from the consideration of additional stochastic influences on the observations. The case of prior information on model parameters is also considered with two possible treatments: their inclusion as additional pseudo-observations, or as means of stochastic parameters.

In Part 2, a general solution is given for the various types of estimates for any model parameter, which is next applied for the derivation of general solutions for each of the three model classes (simple, stochastic of type A and B). These general solutions are applied to some of the more familiar types of linear specific models: Observation equations, condition equations, mixed equations, observation equations with constraints, mixed equations with constraints, mixed (effects) model, generalized mixed model, random effects model, generalized random effects model, observation equations with prior information, observation equations with partial prior information.

In Part 3, the six different prediction types are derived without inference to any model. Instead, the predictions are depending on the form of information available about the means of both the data (random variables with known outcomes) and any new random variable correlated to the data. Such information is differentiated according to where the mean is known (zero or non-zero), or only partly known (a homogeneous or inhomogeneous linear function of some unknown parameters). The derived general solutions are applied for the prediction for stochastic parameters included or related to those of the various forms of specific linear models. It is shown that any of the six types of prediction is compatible with the corresponding type estimates. The combination of "robust" prediction types with the usual BLUE estimates, as well the invariance characteristics of the various prediction types are critically investigated.

Part 1: The model concept

1.1. General and specific model

The first step in data analysis, in geodesy or in any other applied science, is the choice of a suitable model. The word model has been used in various contexts with different meanings. Here it will be used for two concepts, which are distinct but strongly related. The first is the background physical model, which consists of all assumptions and conventions concerning the physical reality related to the data acquisition and the parameters of interest. Such a model, (which may include assumptions from Newton's laws and Euclidean geometry to the simple information about which angles and distances have been observed in a network,) will be called the *general model*. The second is the *specific model*, actually used in the data analysis, which consists of a set of specific mathematical equations relating fixed (deterministic) parameters with stochastic ones (random variables) together with information about the probabilistic behavior of the latter. In the geodetic literature the specific model is some times separated in two parts: the functional model and the stochastic model.

The choice of a specific model relies on the previous choice of a general model. It is possible to have different specific models resulting from the same general model. The most familiar example in geodesy is the use of either observation or condition equations for the adjustment of a control network with equivalent results.

All specific models must be of course equivalent, i.e., they must lead to identical results for the estimates of fixed parameters and the predictions of random variables, except from the effect of linearization and computational (round off) errors. In this sense the (general) model is the model, which constitutes the basic framework within which the tools of data analysis are to be applied for the assessment of the desired information. It is a prerequisite for this process to be successful, that the general model used, though unavoidably "incorrect", is an appropriate one, i.e. one that succeeds to describe reality in a way that is efficient for a set of particular applications. One of the aims of data analysis is to check the validity of the model with use of statistical inference (hypothesis testing).

A general classification of data analysis procedures can be based on the discrete or continuous character of the observations and the unknowns, as well as on the character of the mathematical relations, which describe their interrelations. The cases of continuous observations, and use of differential, integral or integrodifferential equations will be excluded here, since they are not of primary interest in geodetic applications. It will be assumed in the following that the observable parameters are discrete, and related by algebraic equations to discrete or continuous unknown (not observed) entities. Furthermore, in our cases of interest, continuous entities interact with other parameters through specific discrete parameters, the so-called *signals*. Therefore, at a first stage, the discussion about the general model can be restricted to discrete parameters only.

1.2. Characteristics of the general model

The choice of the general model is based on knowledge about the physical conditions surrounding the particular experiment or similar ones. The term "experiment" should be understood here as a short name for the various phases of the data collection campaign. Essential in the mental process of model construction is the *isolation* of a part of the physical world from interactions with the rest of the world, and the *simplification* of the physical processes taking place within it. Due to this isolation and simplification the model becomes the product of an *abstraction*, which is the basic ingredient that makes science, theoretical or applied, possible. As a result a model is always incorrect, but this poses no limit since the essential question is about its efficiency either for the "explanation" of nature or for the choice of proper actions in our encounter with it.

The general model consists of a set of parameters (represented by real numbers), together with the physical laws describing their interrelations (represented by mathematical equations). The parameters and equations may be infinite in number and, in any way, it may be practically impossible, as well as useless, to enumerate all of them. It is sufficient to identify a finite set of parameters, such that any other model parameter can be expressed as a function of them. Such a finite parameter set will be called a *describing* *set of parameters*. It may be very difficult to explicitly describe some model parameters as functions of the parameters of the chosen describing set, but it is sufficient to demonstrate the existence of such mathematical relations.

The minimum possible number of parameters in a model describing set is a model invariant and it will be called the *parametric rank* or simply the *rank* of the model. A describing set whose parameters are independent, in the sense that none of them can be expressed as a function of some of the others, is called a *fundamental set of parameters*. All fundamental sets have the same number of parameters, which is equal to the parametric rank of the model.

The relation of the model to the parameters, observed in a specific experiment, can be looked upon from two different points of view. In the first approach the model is described with one of its fundamental sets of parameters and the question whether the observations performed in the experiment are sufficient for the determination of all model parameters, in which case we have a *properly designed experiment*. The answer to this question is given by the following proposition:

> *An experiment is properly designed with respect to a given model, when the observed parameters form a model describing set.*

In a properly designed experiment therefore, every model parameter, and in particular the parameters of the fundamental set, can be described (at least in principle) as functions of the observed parameters. If the experiment is *improperly designed*, the parameters, which can be expressed as functions of the observed ones are called *determinable parameters* (or *identifiable* parameters). The determinable parameters are the parameters of a new model, which is a submodel of the original one and represents the part of the physical world, which can be studied with use of the available observations. In such a situation, there are two possible courses of action: either the observations should be extended in a way that a new properly designed experiment is performed for the study of the original model, or, if this is not possible, data analysis should be restrained to the new restricted model, with respect to which the available experiment is properly designed. In the latter case a new fundamental set of parameters must replace the original one, if such a set is needed.

A second approach to the relation between observed parameters and model, which is popular in geodesy, owes its existence to the difficulties in expressing the relations between parameters with simple and readily available mathematical equations. There may exist though a set of parameters, which do not belong to the model but have the advantage that model parameters can be easily expressed as functions of them. The typical example in geodesy is the use of coordinates in models, which involve only shape and size (shape) and not position (position and size).

Such a convenient set of parameters is a fundamental set of a new *extended model* in which the original model is imbedded. For the sake of convenience, instead of working with the original model with respect to which the particular experiment is properly designed, the new extended model, with its easy to use fundamental parameter set, is used instead. The particular experiment is in this case improperly designed with respect to the extended model. A distinction must be made for the parameters of the extended model into those, which are determinable, i.e., they belong to the original model and those, which are indeterminable, i.e., they belong to the extended model but not to the original one. Since our interest lies with the original model, it is possible to reverse our terminology and instead of saying that the experiment is improperly designed with respect to the (extended) model, to say that the (extended) model is an "improper model" with respect to the particular experiment. However, in order to stick to more familiar terminology, the extended model will be characterized as a *model without full rank*, a term related to its rank deficiency, that is the difference between its parametric rank and the smaller parametric rank of the "determinable" original model.

Example:
As an example it is sufficient to consider a horizontal network with only three points, a triangle ABC, where the angles A, B, C and the sides a, b, c have been observed. Specific examples will be given for any of the above abstract definitions.

Properly designed experiment. :

Consists of the observations A, B, C, a, b, c, together with a general model that may be described as the "shape and size of the triangle", as conceived within the framework of planar Euclidean geometry.

Parameters of the model are the angles, the side lengths, the heights, the perimeter, the area, the ratios of any two side lengths, the length of the projection of any side on any other, etc.

Describing set of parameters :

The observed 3 angles and 3 side lengths. The 3 angles and 1 side length. The 3 side lengths and the 3 heights, etc.

Parametric rank:

3, since the shape and size of the triangle can be fully determined by 3 parameters.

Fundamental set of parameters :

The 3 side lengths. 2 side lengths and 1 angle. 1 side length and 2 angles. 2 angles and the perimeter. The area, 1 side length and 1 height. etc.

A special choice: The length of b, the length of the projection of side c on side b and the height h_B (the vertices A, B, C, taken counter-clockwise). Note that this choice is equivalent with the 3 coordinates x_C, x_B, y_B, with respect to a particular reference frame chosen so that its origin is at A and is x-axis passes through C (i.e. such that $x_A = y_A = 0$ and $y_C = 0$).

Improperly designed experiment - model without full rank:

Consists of the observations A, B, C, a, b, c, together with a general model that may be described as the "shape and size and position of the triangle" (extended model). Parameters are those of the previous model as well as the 6 coordinates of the vertices of the triangle, as well as any parameter related to position, e.g. the coordinates of the projection of a vertex on its opposite side, etc.

Determinable parameters:

All parameters of the previous model, related to only the shape and size of the triangle.

Indeterminable parameters:

The coordinates, as well as any parameter related to also the position of the triangle.

1.3. The general model in relation to the available data

A model represents in a way an infinite number of possible similar situations corresponding to the infinite choices of possible numerical values for its parameters. It is the purpose of data analysis to determine the set of values that best represents the corresponding physical reality. One may say that the model gives a qualitative description while the analysis of the observations complements this description of physical reality with a quantitative concretization.

When the number of observations in a properly designed experiment is equal with the parametric rank r of the model, the analysis of the data poses no particular problem. Unique values can be determined for a set of fundamental parameters by solving a set of r algebraic equations with r unknowns. Any other model parameter can be computed from its known mathematical relation to the set of the fundamental parameters. In fact, the simpler choice of fundamental set is the observed parameters themselves, although it is not necessarily the more convenient choice for the computation of all parameters of interest.

When an extended model without full rank is used, its fundamental set has $r + d$ elements (d being the rank deficiency) and no unique solution is possible. However any one of the infinitely possible solutions of the system of r equations in $r + d$ unknowns can be used, since they all finally lead to the same values for the determinable parameters. It must be understood that in this case the values of the parameters of the fundamental set have no real meaning and are simply an information "reservoir" which should be used for the computation of meaningful values of determinable parameters and not of indeterminable ones.

When more observations than the parametric rank are performed the problems do start, as the data analysis if confronted with reality. The outcomes of the observations are the signature of the real world, which is too complicated to conform to the simplicity and abstraction of the model. The inconsistencies can be traced by computing values of the chosen fundamental parameters from different subsets of r out of the $n > r$ observations. The values will be different for each different choice. If r of the observations are used for the computation of values for the remaining $n - r$, the computed values will be different from the observed ones.

At this point instead of discarding the model altogether, it is realized that observations and observables (observed parameters) must be distinguished, and this is done by considering observations as the sum of the observables and new parameters: the observational errors. When the errors are small, i.e., when the inconsistencies from the use of the error-influenced observations in the model are small, the model can be retained, otherwise it must be replaced by a better one. One of the aims of the analysis of the observations becomes the assessment of the validity (sufficiency) of the model. With an admissible model, the errors reflect the combined effect of all aspects of reality ignored in the model.

The invention of observational errors made data analysis from superfluous observations possible in two different ways. The first way is to include the errors in the model as additional parameters. The new model has more parameters than unknowns and infinite solutions, out of which only those with small absolute error values seem reasonable. A compromise out of this lack of uniqueness is the selection of the solution with the smallest errors with the application of the least squares principle, or more generally of the weighted least squares principle, with the problem of choice of weights remaining open.

The second way is to model the observational errors not as fixed but rather as random parameters, with known statistical characteristics, which complement the information present in the observations. As it is well known, fundamental statistics are concerned with the study of the simplest possible model with superfluous observations: The model consists of a single parameter which is repeatedly observed, and the observations are considered as the outcomes of a random variable, out of infinite possible realizations, which are governed by probabilities, specified up to a finite set of parameters (e.g. mean and variance).

The observational errors are assumed to be random variables with known zero means and covariance matrix either known or known up to one, or a limited number of unknown parameters (variance components). Sometimes their distribution is also assumed to be the Gaussian one. This assumption is not needed for the data analysis based on linear inference, but it is necessary for making statistical inferences a posteriori.

1.4. Alternative specific models corresponding to the same general model

With the introduction of the observational errors a specific model must be introduced on the basis of the general model for the performance of data analysis. The target is to obtain numerical values for a set of describing parameters (or preferably a set of fundamental parameters) which can be later on used for the computation of numerical values for any model parameters. The specific model results from the combination of the basic observation equations

$$\mathbf{y}^b = \mathbf{y} + \mathbf{v}$$

($\mathbf{y}^b$ being the observed values, $\mathbf{y}$ the observables and $\mathbf{v}$ the observational errors, or, more generally, a stochastic disturbance) with a set of mathematical relations

$$\mathbf{F}(\mathbf{y},\mathbf{x}) = \mathbf{0}$$

relating the observables $\mathbf{y}$ with a set of other parameters $\mathbf{x}$ of the model. The various possible choices of specific models resulting from the choice of $\mathbf{x}$ and $\mathbf{F}$, can be classified in a number of standard categories corresponding to the various well known special cases of least squares adjustment. Let n be the number of observables $\mathbf{y}$, m the number of parameters $\mathbf{x}$, r the parametric rank of the model, $f = n - r$ the degrees of freedom and t the number of equations in (2). Only full rank models will be considered first.

Observation equations ($m = r$, $t = n$): $\mathbf{y} = \mathbf{f}(\mathbf{x})$

Here $\mathbf{x}$ is a fundamental set for the model, and the observables $\mathbf{y}$ as well as any other model parameter q can be described by unique functions $\mathbf{y} = \mathbf{f}(\mathbf{x})$, $q = q(\mathbf{x})$ of this fundamental set. Different choices of fundamental sets lead to different sets of observation equations.

Condition equations ($m = 0$, $t = f$): $\mathbf{g}(\mathbf{y}) = \mathbf{0}$

The only parameters involved are the observables $\mathbf{y}$ (a describing set), which are more than the minimum needed ($n > r$). They are not therefore independent and they must satisfy $f = n - r$ conditions $\mathbf{g}(\mathbf{y}) = \mathbf{0}$ which are not uniquely defined. Any other model parameter q can be described by a non-unique function $q = q(\mathbf{y})$ of the observables. The condition equations can be viewed as the result of the elimination of all parameters $\mathbf{x}$ from a set of observation equations.

Mixed equations ($0 < m \leq r$, $t = m + f$): $\mathbf{u}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$.

In addition to the observables $\mathbf{y}$, another set of model parameters $\mathbf{x}$ is involved. The parameters $\mathbf{x}$ are independent but they do not form a fundamental set (except when $m = r$). The total set of $n + m$ parameters $\mathbf{x}$ and $\mathbf{y}$ is not independent and they must fulfil $n + m - r = m + f$ conditions $\mathbf{u}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$. Any other model parameter q can be described by a non-unique function, either of the observables $q = q(\mathbf{y})$, or of both $\mathbf{x}$ and $\mathbf{y}$, $q = q(\mathbf{x}, \mathbf{y})$, but not of $\mathbf{x}$ only (except when $m = r$). The mixed equations can be also viewed (for $m < r$) as the result of the elimination of some of the parameters $\mathbf{x}$ from a set of observation equations.

Observation equations with constraints ($r < m < n$, $t = n + k$): $\mathbf{y} = \mathbf{f}(\mathbf{x})$, $\mathbf{h}(\mathbf{x}) = \mathbf{0}$

Here the set $\mathbf{x}$ is a describing set but not a fundamental set. Therefore the m parameters $\mathbf{x}$ are not independent and they must satisfy $k = m - r$ constraints $\mathbf{h}(\mathbf{x}) = \mathbf{0}$. The total set of $n + m$ parameters $\mathbf{x}$ and $\mathbf{y}$ are not independent and they must fulfil $n + m - r = m + f = n + k$ conditions of which the n are the observation equations $\mathbf{y} = \mathbf{f}(\mathbf{x})$ and the k are the constraints $\mathbf{h}(\mathbf{x}) = \mathbf{0}$. Any other model parameter q can be described by a non-unique function $q = q(\mathbf{y})$.

Mixed equations with constraints ($k < m \leq r + k$, $t = m + f = s + k$): $\mathbf{u}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$, $\mathbf{h}(\mathbf{x}) = \mathbf{0}$

In addition to the observables $\mathbf{y}$, another set of model parameters $\mathbf{x}$ is involved. The parameters $\mathbf{x}$ are not independent and they do not form a fundamental set. They also do not form a describing set (except for $m = r + k$). The total set of $n + m$ parameters $\mathbf{x}$ and $\mathbf{y}$ are not independent and they must fulfil $n + m - r = m + f$ conditions of which s are the mixed equations $\mathbf{u}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$ and k the constraints $\mathbf{h}(\mathbf{x}) = \mathbf{0}$ (therefore it must hold that $t = m + f = s + k$). Any other model parameter q can be described by a non-unique function, either of the observables $q = q(\mathbf{y})$, or of both $\mathbf{x}$ and $\mathbf{y}$, $q = q(\mathbf{x}, \mathbf{y})$, but not of $\mathbf{x}$ only (except when $m = r + k$). The mixed equations can be also viewed (for $m < r + k$) as the result of the elimination of some of the parameters $\mathbf{x}$ from a set of observation equations with constraints.

When a model without full rank is used all the above specific models can be used, except for the condition equations, which do not involve other parameters $\mathbf{x}$. In the other cases the parameters $\mathbf{x}$ are parameters of the extended model and the rank r must be replaced by the rank $r^* = r + d$ of the extended model. The resulting specific models do not have a unique solution for $\mathbf{x}$, and unique solutions (estimates) can be found only for the determinable functions $q(\mathbf{x})$ or $q(\mathbf{x}, \mathbf{y})$, which correspond to parameters of the original full rank model. One way to obtain unique values for any such functions, and therefore also for $\mathbf{x}$, is the introduction of d arbitrary constraints (minimal constraints) which introduce in an arbitrary manner all the necessary information for the part of the extended model which does not belong to the original one. Due to the arbitrariness of this information the resulting values for $\mathbf{x}$ and non-determinable $q(\mathbf{x})$ or $q(\mathbf{x}, \mathbf{y})$ are arbitrary and meaningless. However the values of the determinable parameters $q(\mathbf{x})$ or $q(\mathbf{x}, \mathbf{y})$ are independent of the choice of the minimal constraints, and the arbitrary values of $\mathbf{x}$ can be seen as a useful and convenient information reservoir for their computation. The typical example from geodesy and photogrammetry involves the use of coordinates where the minimal constraints add arbitrary information about the position, orientation (and scale) of the model, or, from a reverse point of view, they introduce an arbitrary frame of reference thus solving the so-called datum problem.

1.5. Models with stochastic parameters

Another class of models results when the introduction of the observational errors, modeled as random variables, proves to be insufficient for the explanation of discrepancies between the observations and the observables as described by the model. Of course it is possible to improve the model in a deterministic way by adding parameters that were previously ignored or assumed to have known fixed values. Yet another approach to model improvement is the introduction of additional stochastic parameters, i.e., parameters modeled as random variables. It is also possible to model as stochastic, parameters that were previously considered fixed (deterministic), or ignored (fixed to zero values).

There are two different ways to introduce stochastic parameters in a model. The first way is to allow stochastic parameters $\mathbf{s}$ to be also included in the model equations, which take then the general form $\mathbf{F}(\mathbf{x},\mathbf{y},\mathbf{s})=\mathbf{0}$. Such models with stochastic parameters will be called of *type A*, in order to distinguish them from the next class of models. They have the general form:

Type A: $\mathbf{y}^b = \mathbf{y}+\mathbf{v}$ $\qquad$ $\mathbf{F}(\mathbf{x},\mathbf{y},\mathbf{s})=\mathbf{0}$ **.**

The other way of including stochastic parameters in a model, is to add to the noise term $\mathbf{v}$ a stochastic term, which is a known function $\phi(\mathbf{s})$ of a set of stochastic parameters $\mathbf{s}$. The resulting class of models will be called of type B, and has the general form

Type B: $\mathbf{y}^b = \mathbf{y}+\phi(\mathbf{s})+\mathbf{v}$ $\qquad$ $\mathbf{F}(\mathbf{y},\mathbf{x})=\mathbf{0}$.

In models with stochastic parameters, the model concept includes any available knowledge about the joint probability distribution of the stochastic parameters, which are the parameters $\mathbf{s}$ and the observational errors $\mathbf{v}$. The same holds true with respect to $\mathbf{v}$ only, for simple models without stochastic parameters. For the case of *linear inference*, where estimates and predictions are restricted to be linear functions of the available data, observations and model misclosures, it is sufficient to know only the mean vector and the covariance matrix of the vector of all the random variables in the model.

Another interesting aspect of models with stochastic parameters, is their natural extension to any other stochastic parameters $\mathbf{s}'$, for which there is available information about their joint distribution with the stochastic parameters $\mathbf{s}$, explicitly included in the model. Within the framework of linear inference it is possible to obtain predictions for $\mathbf{s}'$, from knowledge of only its mean vector, its covariance matrix and its cross-covariance matrix with $\mathbf{s}$. Such a situation arises in models, which implicitly involve an unknown function modeled as a stochastic process, while $\mathbf{s}$ and $\mathbf{s}'$ are values of functionals on the stochastic function (signals).

1.6. Models with prior information

Another class of models results when the observations are not the only available information but prior information also exists for some or all of the model parameters. The prior information is the result of a previous data analysis and it is therefore the outcome of random variables (estimates for fixed parameters, predictions for stochastic ones) which differs from the corresponding "true" values (values of fixed parameters, outcomes for stochastic ones). These models with prior information are not genuine models but they result from the separation of an original model including all observations into two steps, a first step where prior information results from the analysis of a first set of observations, and a second step where a second set of observations is analyzed taking into account the existing prior information. When the original data are available, what is really needed is a sequential solution strategy rather than a new model. However, loss of the original data imposes the consideration of *models with prior information*, when prior estimates are available for all the (non-observable) parameters $\mathbf{x}$, as well as *models with partial prior information*, when prior estimates are available for only some of the elements of $\mathbf{x}$.

Prior information can be treated in two different ways: The first is to treat prior estimates as "pseudo-observations" of the corresponding parameters contaminated by "errors" (errors in the available estimates or predictions). In this approach the observation equations $\mathbf{y}^b = \mathbf{y} + \mathbf{v}$, must be complemented with the additional set $\mathbf{x}^b = \mathbf{x} + \mathbf{v}_\mathbf{x}$, where $\mathbf{x}^b$ is the vector of the prior estimates of $\mathbf{x}$ and $\mathbf{v}_\mathbf{x}$ the (random) errors in those estimates. In this case the model equations $\mathbf{F}(\mathbf{y},\mathbf{x}) = \mathbf{0}$ contain only "observable" parameters. For partial information for some of the parameters, say $\mathbf{x}_2^b$ for $\mathbf{x}_2$ out of $\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix}$, the complementing "pseudo-observations" are $\mathbf{x}_2^b = \mathbf{x}_2 + \mathbf{v}_2$, where $\mathbf{v}_2$ are the corresponding random errors.

Another way for the incorporation of prior information is to eliminate the relevant parameters from the combination of the original model equations $\mathbf{F}(\mathbf{y},\mathbf{x}) = \mathbf{0}$ with the "pseudo-observations" equations ($\mathbf{x}^b = \mathbf{x} + \mathbf{v}_\mathbf{x}$ or $\mathbf{x}_2^b = \mathbf{x}_2 + \mathbf{v}_2$, accordingly). The result is a model with stochastic parameters ($\mathbf{v}_\mathbf{x}$ or $\mathbf{v}$). This is not a genuinely new model, but merely a reformulation of the "pseudo-observations" model into a different specific model corresponding to the same though general model.

The second way of treating prior information, is to directly consider the parameters $\mathbf{x}$ (or $\mathbf{x}_2$) as stochastic parameters with the available values $\mathbf{x}^b$ (or $\mathbf{x}_2^b$) as their mean vectors!

In fact this approach gives numerically the same results as the pseudo-observation approach, for those estimate and prediction types which are invariant under shift transformations. The model is of course a different one but the numerical equivalence is not accidental. There is a kind of duality between the two models. The duality is based in two different but related interpretations of the available prior values:

1. The prior values are outcomes of random variables having the (deterministic) unknown parameters as means and a given covariance matrix.
2. The prior values are the known means of the (unknown) random parameters, which have the given covariance matrix.

For an insight in this duality consider the simpler case of a single random variable with known mean and distribution. Before any outcome is observed it is possible to compute the probability that the outcome to be observed will fall within any fixed set of values. Now consider the same situation with a random variable of the same distribution known except for its mean value, and consider that a single outcome has been observed. Formally the mean value remains a deterministic unknown quantity that is completely unknown. However, most people will agree that this mean value is not completely unknown in view of the available outcome. Values close to the outcome are more probable than others. In fact one can use the known distribution with the known outcome as mean in order to make inferences about the unknown mean, which becomes now as a random variable!

The above two interpretations are two sides of the same coin, and the argument about which is the proper one to use is well known in statistics, especially in the interpretation of confidence intervals. Two prominent statisticians had different points of view in what is now known as the *Neyman-Pearson controversy*.

1.7. Linearized specific models

Data analysis in geodesy and is usually performed not with the original model but with a linearized one, obtained with the help of approximate values $\mathbf{x}^0$, $\mathbf{y}^0$, $\mathbf{s}^0$ for $\mathbf{x} = \mathbf{x}^0 + \delta\mathbf{x}$, $\mathbf{y} = \mathbf{y}^0 + \delta\mathbf{y}$, $\mathbf{s} = \mathbf{s}^0 + \delta\mathbf{s}$ respectively. Occasionally $\mathbf{s}^0 = E\{\mathbf{s}\} = \mathbf{m}_\mathbf{s}$ and only circumstantially $\mathbf{y}^0 = \mathbf{y}^b$, otherwise coefficient matrices depending on the random vector $\mathbf{y}^b$ would be random matrices.

In order to also bring the linearized models in their usual form, the observables $\mathbf{y}$ are eliminated between the "observation" and the "model" equations, while the approximate value $\mathbf{F}^0$ of the model equations is expressed in terms of the misclosure vector $\mathbf{w}$, which results from $\mathbf{F}^0$ when the observations $\mathbf{y}^b$ are used instead of the approximate values $\mathbf{y}^0$.

Models without stochastic parameters (Simple models) : General form: $\mathbf{y}^b = \mathbf{y} + \mathbf{v}$ $\quad \mathbf{F}(\mathbf{y},\mathbf{x}) = \mathbf{0}$
Linearization: $\mathbf{y}^b = \mathbf{y}^0 + \delta\mathbf{y} + \mathbf{v} \quad \Rightarrow \quad \mathbf{b} = \delta\mathbf{y} + \mathbf{v} \quad (\mathbf{b} \equiv \mathbf{y}^b - \mathbf{y}^0)$ $\mathbf{F}(\mathbf{y},\mathbf{x}) = \mathbf{F}(\mathbf{y}^0,\mathbf{x}^0) + \left.\frac{\partial \mathbf{F}}{\partial \mathbf{x}}\right\|_0 \delta\mathbf{x} + \left.\frac{\partial \mathbf{F}}{\partial \mathbf{y}}\right\|_0 \delta\mathbf{y} = \mathbf{0}$ $\mathbf{F}^0 + \mathbf{Q}_x \delta\mathbf{x} + \mathbf{Q}_y \delta\mathbf{y} = \mathbf{0} \quad \Rightarrow \quad \mathbf{t} = \mathbf{Q}_x \delta\mathbf{x} + \mathbf{Q}_y \delta\mathbf{y} \quad (\mathbf{t} \equiv -\mathbf{F}^0)$ $\mathbf{w} \equiv \mathbf{F}(\mathbf{x}^0,\mathbf{y}^b) = \mathbf{F}(\mathbf{x}^0,\mathbf{y}^0) + \left.\frac{\partial \mathbf{F}}{\partial \mathbf{y}}\right\|_0 (\mathbf{y}^b - \mathbf{y}^0) = \mathbf{F}^0 + \mathbf{Q}_y \mathbf{b} \quad \Rightarrow \quad \mathbf{t} = \mathbf{Q}_y \mathbf{b} - \mathbf{w}$
Usual form: $\quad \mathbf{w} = -\mathbf{Q}_x \delta\mathbf{x} + \mathbf{Q}_y \mathbf{v}$

<u>Application to specific models</u>

Observation equations: $\quad \mathbf{y}^b = \mathbf{f}(\mathbf{x}) + \mathbf{v} \quad \Rightarrow \quad \mathbf{b} = \mathbf{A}\delta\mathbf{x} + \mathbf{v}$

$$\mathbf{F}(\mathbf{x},\mathbf{y}) = \mathbf{y} - \mathbf{f}(\mathbf{x}), \qquad \mathbf{Q}_x = -\left.\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right|_0 = -\mathbf{A}, \qquad \mathbf{Q}_y = \mathbf{I}, \qquad \mathbf{t} = \mathbf{f}(\mathbf{x}^0) - \mathbf{y}^0 = \mathbf{0}$$

Condition equations: $\quad \mathbf{g}(\mathbf{y}) = \mathbf{0} \quad \Rightarrow \quad \mathbf{B}\mathbf{v} = \mathbf{w}$

$$\mathbf{F}(\mathbf{x},\mathbf{y}) = \mathbf{g}(\mathbf{y}), \qquad \mathbf{Q}_x = \mathbf{0}, \qquad \mathbf{Q}_y = \left.\frac{\partial \mathbf{g}}{\partial \mathbf{y}}\right|_0 \equiv \mathbf{B}, \quad \mathbf{t} = -\mathbf{g}(\mathbf{y}^0)$$

$$\mathbf{w} = \mathbf{g}(\mathbf{y}^b) = \mathbf{g}(\mathbf{y}^0) + \left.\frac{\partial \mathbf{g}}{\partial \mathbf{y}}\right|_0 (\mathbf{y}^b - \mathbf{y}^0), \qquad \mathbf{t} = -\mathbf{g}(\mathbf{y}^0) = \mathbf{B}\mathbf{b} - \mathbf{w}$$

Mixed equations: $\mathbf{u}(\mathbf{x},\mathbf{y}) = \mathbf{0} \quad \Rightarrow \quad \mathbf{w} + \mathbf{A}\delta\mathbf{x} - \mathbf{B}\mathbf{v} = \mathbf{0}$

$$\mathbf{F}(\mathbf{x},\mathbf{y}) = \mathbf{u}(\mathbf{x},\mathbf{y}), \qquad \mathbf{Q}_x = -\left.\frac{\partial \mathbf{u}}{\partial \mathbf{x}}\right|_0 \equiv \mathbf{A}, \qquad \mathbf{Q}_y = \left.\frac{\partial \mathbf{u}}{\partial \mathbf{y}}\right|_0 \equiv \mathbf{B}, \quad \mathbf{t} = -\mathbf{u}(\mathbf{x}^0,\mathbf{y}^0)$$

$$\mathbf{w} = \mathbf{u}(\mathbf{x}^0,\mathbf{y}^b) = \mathbf{u}(\mathbf{x}^0,\mathbf{y}^0) + \left.\frac{\partial \mathbf{u}}{\partial \mathbf{y}}\right|_0 (\mathbf{y}^b - \mathbf{y}^0), \qquad \mathbf{t} = -\mathbf{u}(\mathbf{x}^0,\mathbf{y}^0) = \mathbf{B}\mathbf{b} - \mathbf{w}$$

Constraints:

$$\mathbf{h}(\mathbf{x}) = \mathbf{0} \quad \Rightarrow \quad \left.\frac{\partial \mathbf{h}}{\partial \mathbf{x}}\right|_0 \delta\mathbf{x} = -\mathbf{h}(\mathbf{x}^0) \quad \Rightarrow \quad \mathbf{H}\mathbf{x} = -\mathbf{h}^0 \tag{14}$$

Constraints can be combined with other models where parameters $\mathbf{x}$ are present, in which case $\mathbf{t}$, $\mathbf{Q}_\mathbf{x}$, $\mathbf{Q}_\mathbf{y}$ are replaced by $\begin{bmatrix} \mathbf{t} \\ -\mathbf{h}^0 \end{bmatrix}$, $\begin{bmatrix} \mathbf{Q}_\mathbf{x} \\ \mathbf{H} \end{bmatrix}$, $\begin{bmatrix} \mathbf{Q}_\mathbf{y} \\ \mathbf{0} \end{bmatrix}$, respectively. Stochastic models with constraints are introduced in a similar way.

Models with stochastic parameters - Type A : General form: $\mathbf{y}^b = \mathbf{y} + \mathbf{v} \qquad \mathbf{F}(\mathbf{x},\mathbf{y},\mathbf{s}) = \mathbf{0}$
Linearization: $\mathbf{y}^b = \mathbf{y}^0 + \delta\mathbf{y} + \mathbf{v} \quad \Rightarrow \quad \mathbf{b} + \delta\mathbf{y} + \mathbf{v} \qquad \mathbf{b} \equiv \mathbf{y}^b - \mathbf{y}^0$ $\mathbf{F}(\mathbf{x},\mathbf{y},\mathbf{s}) = \mathbf{F}(\mathbf{x}^0,\mathbf{y}^0,\mathbf{s}^0) + \left.\frac{\partial \mathbf{F}}{\partial \mathbf{x}}\right\|_0 \delta\mathbf{x} + \left.\frac{\partial \mathbf{F}}{\partial \mathbf{y}}\right\|_0 \delta\mathbf{y} + \left.\frac{\partial \mathbf{F}}{\partial \mathbf{s}}\right\|_0 \delta\mathbf{s} = \mathbf{0}$ $\mathbf{F}^0 + \mathbf{Q}_\mathbf{x}\delta\mathbf{x} + \mathbf{Q}_\mathbf{y}\delta\mathbf{y} = \mathbf{0} \quad \Rightarrow \quad \mathbf{t} = \mathbf{Q}_\mathbf{x}\delta\mathbf{x} + \mathbf{Q}_\mathbf{y}\delta\mathbf{y} + \mathbf{Q}_\mathbf{s}\delta\mathbf{s} = \mathbf{0} \qquad (\mathbf{t} \equiv -\mathbf{F}^0)$ $\mathbf{w} = \mathbf{F}(\mathbf{x}^0,\mathbf{y}^b,\mathbf{s}^0) + \mathbf{F}(\mathbf{x}^0,\mathbf{y}^0,\mathbf{s}^0) + \left.\frac{\partial \mathbf{F}}{\partial \mathbf{y}}\right\|_0 \delta\mathbf{y} = \mathbf{F}^0 + \mathbf{Q}_\mathbf{y}\mathbf{b} \quad \Rightarrow \quad \mathbf{t} = \mathbf{Q}_\mathbf{y}\mathbf{b} - \mathbf{w}$
Usual form: $\mathbf{w} = -\mathbf{Q}_\mathbf{x}\delta\mathbf{x} - \mathbf{Q}_\mathbf{s}\delta\mathbf{s} + \mathbf{Q}_\mathbf{y}\mathbf{v}$

Application to specific models

Mixed model: $\mathbf{y}^b = \mathbf{f}(\mathbf{x},\mathbf{s}) + \mathbf{v} \quad \Rightarrow \quad \mathbf{b} = \mathbf{A}\delta\mathbf{x} + \mathbf{G}\delta\mathbf{s} + \mathbf{v}$:

$$\mathbf{F}(\mathbf{x},\mathbf{y},\mathbf{s}) = \mathbf{y} - \mathbf{f}(\mathbf{x},\mathbf{s}), \qquad \mathbf{Q}_\mathbf{x} = -\left.\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right|_0 = -\mathbf{A}, \qquad \mathbf{Q}_\mathbf{y} = \mathbf{I},\ \mathbf{Q}_\mathbf{s} = -\left.\frac{\partial \mathbf{f}}{\partial \mathbf{s}}\right|_0 = -\mathbf{G}$$

$$\mathbf{t} = \mathbf{f}(\mathbf{x}^0,\mathbf{s}^0) - \mathbf{y}^0 = \mathbf{0} = \mathbf{b} - \mathbf{w}, \qquad \mathbf{w} = \mathbf{b}$$

Generalized mixed model: $\mathbf{u}(\mathbf{x},\mathbf{y},\mathbf{s}) = \mathbf{0} \quad \Rightarrow \quad \mathbf{w} + \mathbf{A}\delta\mathbf{x} + \mathbf{G}\delta\mathbf{s} - \mathbf{B}\mathbf{v} = \mathbf{0}$:

$$\mathbf{F}(\mathbf{x},\mathbf{y},\mathbf{s}) = \mathbf{u}(\mathbf{x},\mathbf{y},\mathbf{s}), \qquad \mathbf{Q}_\mathbf{x} = \left.\frac{\partial \mathbf{u}}{\partial \mathbf{x}}\right|_0 = \mathbf{A},\ \mathbf{Q}_\mathbf{y} = \left.\frac{\partial \mathbf{u}}{\partial \mathbf{y}}\right|_0 = \mathbf{B},\ \mathbf{Q}_\mathbf{s} = \left.\frac{\partial \mathbf{u}}{\partial \mathbf{s}}\right|_0 = \mathbf{G}$$

$$\mathbf{t} = -\mathbf{u}(\mathbf{x}^0,\mathbf{y}^0,\mathbf{s}^0) = \mathbf{B}\mathbf{b} - \mathbf{w}$$

Random effects model: $\mathbf{y}^b = \mathbf{f}(\mathbf{s}) + \mathbf{v} \quad \Rightarrow \quad \mathbf{b} = \mathbf{G}\delta\mathbf{s} + \mathbf{v}$

$$\mathbf{F}(\mathbf{x},\mathbf{y},\mathbf{s}) = \mathbf{y} - \mathbf{f}(\mathbf{s}), \qquad \mathbf{Q}_x = \mathbf{0}, \qquad \mathbf{Q}_y = \mathbf{I},\ \mathbf{Q}_s = -\left.\frac{\partial \mathbf{f}}{\partial \mathbf{s}}\right|_0 = -\mathbf{G}$$

$$\mathbf{t} = \mathbf{f}(\mathbf{s}^0) - \mathbf{y}^0 = \mathbf{0} = \mathbf{Bb} - \mathbf{w}, \qquad \mathbf{w} = \mathbf{Bb}$$

Generalized random effects model: $\mathbf{g}(\mathbf{y},\mathbf{s}) = \mathbf{0} \quad \Rightarrow \quad \mathbf{w} = \mathbf{Bv} - \mathbf{G}\delta\mathbf{s}$:

$$\mathbf{F}(\mathbf{x},\mathbf{y},\mathbf{s}) = \mathbf{g}(\mathbf{y},\mathbf{s}), \qquad \mathbf{Q}_x = \mathbf{0}, \qquad \mathbf{Q}_y = \left.\frac{\partial \mathbf{g}}{\partial \mathbf{y}}\right|_0 = \mathbf{B},\ \mathbf{Q}_s = \left.\frac{\partial \mathbf{g}}{\partial \mathbf{s}}\right|_0 = \mathbf{G} \quad ???\mathbf{Q}_\mathbf{y} =$$

$$\mathbf{t} = -\mathbf{g}(\mathbf{y}^0,\mathbf{s}^0) = \mathbf{Bb} - \mathbf{w}$$

Models with stochastic parameters - Type B : General form: $\mathbf{y}^b = \mathbf{y} + \boldsymbol{\varphi}(\mathbf{s}) + \mathbf{v} \qquad \mathbf{F}(\mathbf{x},\mathbf{y}) = \mathbf{0}$
Linearization: $\mathbf{y}^b = \mathbf{y}^0 + \delta\mathbf{y} + \boldsymbol{\varphi}^0 + \left.\frac{\partial \boldsymbol{\varphi}}{\partial \mathbf{s}}\right\|_0 \delta\mathbf{s} + \mathbf{v} \quad \Rightarrow \quad \mathbf{b} = \delta\mathbf{y} + \mathbf{Q}_s\delta\mathbf{s} + \mathbf{v} \qquad (\mathbf{b} \equiv \mathbf{y}^b - \mathbf{y}^0 - \boldsymbol{\varphi}^0)$ $\mathbf{F}(\mathbf{x},\mathbf{y}) = \mathbf{F}(\mathbf{x}^0,\mathbf{y}^0) + \left.\frac{\partial \mathbf{F}}{\partial \mathbf{x}}\right\|_0 \delta\mathbf{x} + \left.\frac{\partial \mathbf{F}}{\partial \mathbf{y}}\right\|_0 \delta\mathbf{y} = \mathbf{0}$ $\mathbf{F}^0 + \mathbf{Q}_x\delta\mathbf{x} + \mathbf{Q}_y\delta\mathbf{y} = \mathbf{0} \quad \Rightarrow \quad \mathbf{t} = \mathbf{Q}_x\delta\mathbf{x} + \mathbf{Q}_y\delta\mathbf{y} \qquad (\mathbf{t} \equiv -\mathbf{F}^0)$ $\mathbf{w} = \mathbf{F}(\mathbf{x}^0, \mathbf{y}^b - \boldsymbol{\varphi}^0) = \mathbf{F}(\mathbf{x}^0,\mathbf{y}^0) + \left.\frac{\partial \mathbf{F}}{\partial \mathbf{y}}\right\|_0 \mathbf{b} = \mathbf{F}^0 + \mathbf{Q}_y\mathbf{b} \quad \Rightarrow \quad \mathbf{t} = \mathbf{Q}_y\mathbf{b} - \mathbf{w}$
Usual form: $\mathbf{w} = -\mathbf{Q}_x\delta\mathbf{x} + \mathbf{Q}_y\mathbf{Q}_s\delta\mathbf{s} + \mathbf{Q}_y\mathbf{v}$

<u>Application to specific models</u>

Mixed model: $\mathbf{y}^b = \mathbf{f}(\mathbf{x}) + \boldsymbol{\varphi}(\mathbf{s}) + \mathbf{v} \quad \Rightarrow \quad \mathbf{b} = \mathbf{A}\delta\mathbf{x} + \mathbf{G}\delta\mathbf{s} + \mathbf{v}$:

$$\mathbf{F}(\mathbf{x},\mathbf{y}) = \mathbf{y} - \mathbf{f}(\mathbf{x}), \qquad \mathbf{Q}_x = -\left.\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right|_0 = -\mathbf{A}, \qquad \mathbf{Q}_y = \mathbf{I},\ \mathbf{Q}_s = \left.\frac{\partial \mathbf{f}}{\partial \mathbf{s}}\right|_0 = \mathbf{G} \qquad \textbf{f->}\boldsymbol{\varphi}\textbf{?}$$

$$\mathbf{t} = \mathbf{f}(\mathbf{x}^0) - \mathbf{y}^0 = \mathbf{0} = \mathbf{b} - \mathbf{w}, \qquad \mathbf{w} = \mathbf{b}$$

Generalized mixed model: $\mathbf{u}(\mathbf{x},\mathbf{y}) = \mathbf{0} \quad \Rightarrow \quad \mathbf{w} = -\mathbf{A}\delta\mathbf{x} + \mathbf{G}\delta\mathbf{s} + \mathbf{Bv}$

$$\mathbf{F}(\mathbf{x},\mathbf{y}) = \mathbf{u}(\mathbf{x},\mathbf{y}), \qquad \mathbf{Q}_x = \left.\frac{\partial \mathbf{u}}{\partial \mathbf{x}}\right|_0 = \mathbf{A}, \qquad \mathbf{Q}_y = \left.\frac{\partial \mathbf{u}}{\partial \mathbf{y}}\right|_0 = \mathbf{B}, \qquad \mathbf{Q}_y \left.\frac{\partial \mathbf{u}}{\partial \mathbf{s}}\right|_0 = \mathbf{Q}_y\mathbf{Q}_s = \mathbf{G} \ \ ???$$

$$\mathbf{t} = -\mathbf{u}(\mathbf{x}^0,\mathbf{y}^0) = \mathbf{Bb} - \mathbf{w}$$

Random effects model: $\mathbf{y}^b = \boldsymbol{\varphi}(\mathbf{s}) + \mathbf{v} \quad \Rightarrow \quad \mathbf{b} = \mathbf{G}\delta\mathbf{s} + \mathbf{v}$

$$\mathbf{F}(\mathbf{x},\mathbf{y}) = \mathbf{y}, \qquad \mathbf{Q_x} = \mathbf{0}, \qquad \mathbf{Q_y} = \mathbf{I},\ \mathbf{Q_s} = \left.\frac{\partial \boldsymbol{\varphi}}{\partial \mathbf{s}}\right|_0 = \mathbf{G}$$

$$\mathbf{t} = -\mathbf{y}^0 = \mathbf{0} = \mathbf{b} - \mathbf{w}, \qquad \mathbf{w} = \mathbf{b}$$

Generalized random effects model: $\mathbf{g}(\mathbf{y}) = \mathbf{0} \quad \Rightarrow \quad \mathbf{w} = \mathbf{B}\mathbf{v} + \mathbf{G}\delta\mathbf{s}$:

$$\mathbf{F}(\mathbf{x},\mathbf{y}) = \mathbf{g}(\mathbf{y}), \qquad \mathbf{Q_x} = \mathbf{0}, \qquad \mathbf{Q_y} = \left.\frac{\partial \mathbf{g}}{\partial \mathbf{y}}\right|_0 = \mathbf{B}, \qquad \mathbf{Q_y}\left.\frac{\partial \mathbf{g}}{\partial \mathbf{s}}\right|_0 = \mathbf{Q_y}\mathbf{Q_s} = \mathbf{G} \quad ???$$

$$\mathbf{t} = -\mathbf{g}(\mathbf{y}^0) = \mathbf{B}\mathbf{b} - \mathbf{w}.$$

For the sake of notational simplicity $\delta\mathbf{x}$, $\delta\mathbf{y}$, $\delta\mathbf{s}$, will be replaced from now on by $\mathbf{x}$, $\mathbf{y}$, $\mathbf{s}$, respectively.

The submatrices of the various specific models are summarized in the following tables for easy reference.

Specific model	linear form	$\mathbf{t}$	$\mathbf{Q_x}$	$\mathbf{Q_y}$
Observation equations	$\mathbf{b} = \mathbf{A}\mathbf{x} + \mathbf{v}$	$\mathbf{0}$	$-\mathbf{A}$	$\mathbf{I}$
Condition equations.	$\mathbf{B}\mathbf{v} = \mathbf{w}$	$\mathbf{B}\mathbf{b} - \mathbf{w}$	$\mathbf{0}$	$\mathbf{B}$
Mixed equations	$\mathbf{w} = -\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{v}$	$\mathbf{B}\mathbf{b} - \mathbf{w}$	$\mathbf{A}$	$\mathbf{B}$
Observation equations with constraints	$\mathbf{b} = \mathbf{A}\mathbf{x} + \mathbf{v}$ $\mathbf{H}\mathbf{x} = -\mathbf{h}^0$	$\begin{bmatrix}\mathbf{0}\\ -\mathbf{h}^0\end{bmatrix}$	$\begin{bmatrix}-\mathbf{A}\\ \mathbf{H}\end{bmatrix}$	$\begin{bmatrix}\mathbf{I}\\ \mathbf{0}\end{bmatrix}$
Mixed equations with constraints	$\mathbf{w} = -\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{v}$ $\mathbf{H}\mathbf{x} = -\mathbf{h}^0$	$\begin{bmatrix}\mathbf{B}\mathbf{b} - \mathbf{w}\\ -\mathbf{h}^0\end{bmatrix}$	$\begin{bmatrix}\mathbf{A}\\ \mathbf{H}\end{bmatrix}$	$\begin{bmatrix}\mathbf{B}\\ \mathbf{0}\end{bmatrix}$
Simple models: $\mathbf{b} = \mathbf{y} + \mathbf{v}$ $\quad \mathbf{t} = \mathbf{Q_x}\mathbf{x} + \mathbf{Q_y}\mathbf{y}$				

Specific model	linear form	$\mathbf{t}$	$\mathbf{Q_x}$	$\mathbf{Q_y}$	$\mathbf{Q_s}$
Mixed	$\mathbf{b} = \mathbf{A}\mathbf{x} + \mathbf{G}\mathbf{s} + \mathbf{v}$	$\mathbf{0}$	$-\mathbf{A}$	$\mathbf{I}$	$-\mathbf{G}$
Generalized mixed	$\mathbf{w} = -\mathbf{A}\mathbf{x} - \mathbf{G}\mathbf{s} + \mathbf{B}\mathbf{v}$	$\mathbf{B}\mathbf{b} - \mathbf{w}$	$\mathbf{A}$	$\mathbf{B}$	$\mathbf{G}$
Random effects	$\mathbf{b} = \mathbf{G}\mathbf{s} + \mathbf{v}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{I}$	$-\mathbf{G}$
Generalized random effects	$\mathbf{w} = -\mathbf{G}\mathbf{s} + \mathbf{B}\mathbf{v}$	$\mathbf{B}\mathbf{b} - \mathbf{w}$	$\mathbf{0}$	$\mathbf{B}$	$\mathbf{G}$
Models with stochastic parameters - Type A: $\mathbf{b} = \mathbf{y} + \mathbf{v}$ $\quad \mathbf{t} = \mathbf{Q_x}\mathbf{x} + \mathbf{Q_y}\mathbf{y} + \mathbf{Q_s}\mathbf{s}$					

Specific model	linear form	$\mathbf{t}$	$\mathbf{Q}_x$	$\mathbf{Q}_y$	$\mathbf{Q}_s$	$\mathbf{Q}_y\mathbf{Q}_s$
Mixed	$\mathbf{b} = \mathbf{Ax} + \mathbf{Gs} + \mathbf{v}$	$\mathbf{0}$	$-\mathbf{A}$	$\mathbf{I}$	$\mathbf{G}$	-
Generalized mixed	$\mathbf{w} = -\mathbf{Ax} + \mathbf{Gs} + \mathbf{Bv}$	$\mathbf{Bb} - \mathbf{w}$	$\mathbf{A}$	$\mathbf{B}$	-	$\mathbf{G}$
Random effects	$\mathbf{b} = \mathbf{Gs} + \mathbf{v}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{I}$	$\mathbf{G}$	-
Generalized random effects	$\mathbf{w} = \mathbf{Gs} + \mathbf{Bv}$	$\mathbf{Bb} - \mathbf{w}$	$\mathbf{0}$	$\mathbf{B}$	-	$\mathbf{G}$
Models with stochastic parameters – Type B:	$\mathbf{b} = \mathbf{y} + \mathbf{Gs} + \mathbf{v}$	$\mathbf{t} = \mathbf{Q}_x\mathbf{x} + \mathbf{Q}_y\mathbf{y}$				

For the above models, estimates must be found not only for the parameters $\mathbf{x}$ and $\mathbf{y}$ but also for any parameter q of the original model which can be always be expressed as a known function $q = q(\mathbf{x}, \mathbf{y})$, $q = q(\mathbf{x})$ and $q = q(\mathbf{y})$ been particular cases. When working with linearized equations, $q = q(\mathbf{x}, \mathbf{y})$ must be also linearized to give

$$\delta q = q - q(\mathbf{x}^0, \mathbf{y}^0) = \left.\frac{\partial q}{\partial \mathbf{x}}\right|_0 \delta\mathbf{x} + \left.\frac{\partial q}{\partial \mathbf{y}}\right|_0 \delta\mathbf{y} = \mathbf{c}_{\mathbf{x}}^T \delta\mathbf{x} + \mathbf{c}_{\mathbf{y}}^T \delta\mathbf{y}$$

or by dropping δ for notational simplicity

$$q = \mathbf{c}_{\mathbf{x}}^T \mathbf{x} + \mathbf{c}_{\mathbf{y}}^T \mathbf{y} .$$

The estimate $\hat{q}$ of q resulting from any specific model within the same class, must be the same, except for the small effects due to different linearization or computational (round-off) errors. In order to emphasize the equivalence of the specific models, estimates of different types, i.e. according to different estimation principles, will be derived from a single general solution. The general solution will be at a first step specialized to the general solutions for each of the four classes of models considered above. From the general solution of each class, the solution for each particular specific model will be very easily derived, by simply replacing certain general matrices with their values in the specific model.

Part 2: Linear estimation

2.1. The alternative types of linear estimation

Any estimate $\hat{q}$ of a model parameter $q=q(\mathbf{x},\mathbf{y})$ must be a function of the available data, which are the observations $\mathbf{b}$ and the constant terms $\mathbf{t}$ in the model equations satisfied by the parameters, i.e. the linearized form of (2), or (5) when stochastic parameters are included. The estimate $\hat{q}$ must be a function

$$\hat{q}=\hat{q}(\mathbf{b},\mathbf{t}) \tag{1}$$

Linear estimation refers to the simpler case where the above function has a linear form, and can be distinguished into *inhomogeneous* linear estimation

$$\hat{q}=\mathbf{h}^T\mathbf{b}+\mathbf{g}^T\mathbf{t}+\kappa \tag{2}$$

and *homogeneous* linear estimation

$$\hat{q}=\mathbf{h}^T\mathbf{b}+\mathbf{g}^T\mathbf{t} \tag{3}$$

The best values for the coefficients $\mathbf{h}$, $\mathbf{g}$ and κ must be determined according to a chosen optimality criterion. The usual choice is the minimization of the mean square estimation error

$$R=\{(\hat{q}-q)^2\} \tag{4}$$

Estimates with minimum mean square error are simply called best estimates, and use of the form (8) or (9) leads to the Best Linear inhomogeneous Estimation *(inhomBLE)* or to the Best Linear homogeneous Estimation *(homBLE)*, respectively.

Sometimes it is additionally required that the estimate is unbiased, i.e. that the estimation bias

$$\beta=E\{\hat{q}\}-q \tag{5}$$

vanishes, thus obtaining best linear unbiased estimates, i.e. estimates with minimum mean square among all those who are linear (inhomogeneous or homogeneous) and unbiased. This leads to two more types of estimation, the Best Linear inhomogeneous Unbiased Estimation *(inhomBLUE)* or to the Best Linear homogeneous Unbiased Estimation *(homBLUE)*.

The bias is in general a function $\beta=\beta(\mathbf{x},\mathbf{y})$ of the model parameters, and the term unbiased as used before means that the estimate is unbiased for the true, but unfortunately unknown, values of $\mathbf{x}$ and $\mathbf{y}$. A different notion of unbiasedness results if it is required that the estimate is *uniformly unbiased*, i.e. unbiased for any value of $\mathbf{x}$ and $\mathbf{y}$. In this case the condition $\beta(\mathbf{x},\mathbf{y})=0$ is treated as a Diophantine equation and the coefficients of $\mathbf{x}$, $\mathbf{y}$ and the constant terms must be separately set to zero. This leads to another two types of estimation, the Best Linear inhomogeneous Uniformly Unbiased Estimation *(inhomBLUUE)* or to the Best Linear homogeneous Uniformly Unbiased Estimation *(homBLUUE)*.

2.2. The general solution to the estimation problem

The problem of linear estimation of any model parameter which is a known function $q=\mathbf{c}^T\mathbf{z}$ of the model parameters $\mathbf{z}$, is the construction of an estimate $\hat{q}$, which is a "linear" function $\hat{q}=\mathbf{p}^T\mathbf{d}+\kappa$, of the available data

$$\mathbf{d}=\mathbf{U}\mathbf{z}+\mathbf{e} \tag{6}$$

where $\mathbf{e}$ is a stochastic term. The problem is the determination of the parameters $\mathbf{p}$ and κ for linear inhomogeneous estimation, or only $\mathbf{p}$ for the homogeneous one, which satisfy the appropriate optimality criteria. Introducing the notation

$$\mathbf{m}=E\{\mathbf{e}\}\,, \qquad \mathbf{C}=E\{(\mathbf{e}-\mathbf{m})(\mathbf{e}-\mathbf{m})^T\}\,, \qquad \mathbf{m_d}=E\{\mathbf{d}\}=\mathbf{Uz}+\mathbf{m} \tag{7}$$

the general solution is derived in Appendix A and is summarized in table 2.1.

inhomBLE / inhomBLUE:	$\hat{q}=q+\mathbf{p}^T(\mathbf{d}-\mathbf{m_d})$,	$\mathbf{Cp}=\mathbf{0}$
homBLE:	$\hat{q}=\mathbf{p}^T\mathbf{d}$,	$(\mathbf{C}+\mathbf{m_d}\mathbf{m_d}^T)\mathbf{p}=q\mathbf{m_d}$
homBLUE:	$\hat{q}=\mathbf{p}^T\mathbf{d}$,	$\begin{bmatrix}\mathbf{C} & \mathbf{m_d}\\ \mathbf{m_d}^T & 0\end{bmatrix}\begin{bmatrix}\mathbf{p}\\ -\nu\end{bmatrix}=\begin{bmatrix}\mathbf{0}\\ q\end{bmatrix}$
inhomBLUUE:	$\hat{q}=\mathbf{p}^T(\mathbf{d}-\mathbf{m})$,	$\begin{bmatrix}\mathbf{C} & \mathbf{U}\\ \mathbf{U}^T & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{p}\\ -\mathbf{k}\end{bmatrix}=\begin{bmatrix}\mathbf{0}\\ \mathbf{c}\end{bmatrix}$
homBLUUE ($\mathbf{m}\neq\mathbf{0}$):	$\hat{q}=\mathbf{p}^T(\mathbf{d}-\mathbf{m})$	$\begin{bmatrix}\mathbf{C} & \mathbf{U} & \mathbf{m}\\ \mathbf{U}^T & \mathbf{0} & \mathbf{0}\\ \mathbf{m}^T & \mathbf{0} & 0\end{bmatrix}\begin{bmatrix}\mathbf{p}\\ -\mathbf{k}\\ -\nu\end{bmatrix}=\begin{bmatrix}\mathbf{0}\\ \mathbf{c}\\ 0\end{bmatrix}$
homBLUUE ($\mathbf{m}=\mathbf{0}$):	$\hat{q}=\mathbf{p}^T\mathbf{d}$	$\begin{bmatrix}\mathbf{C} & \mathbf{U}\\ \mathbf{U}^T & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{p}\\ -\mathbf{k}\end{bmatrix}=\begin{bmatrix}\mathbf{0}\\ \mathbf{c}\end{bmatrix}$

Table 2.1: General form of linear estimates $\hat{q}$ for any model parameters $q=\mathbf{c}^T\mathbf{z}$ from available data $\mathbf{d}=\mathbf{Uz}+\mathbf{e}$ with $\mathbf{e}\sim(\mathbf{m},\mathbf{C})$ ($\mathbf{m_d}=\mathbf{Uz}+\mathbf{m}$).

From table 2.1 it is directly seen that the estimates of type inhomBLE, homBLE, inhomBLUE and homBLUE, can not be used since they all depend on the unknown parameter q. The only estimates which can be directly used are those of the types inhomBLUUE and homBLUUE.

In general an estimate is better if its mean square error (used as the minimized risk function R in the choice of optimal estimation) is smaller. The minimum of the risk function is smaller when the class of functions where the minimum is sought is larger. Therefore the following inequalities hold

$$\text{R(inhomBLE)} < \text{R(homBLE)} \tag{8}$$

$$\text{R(inhomBLUE)} < \text{R(homBLUE)} \tag{9}$$

$$\text{R(inhomBLUUE)} < \text{R(homBLUUE)} \tag{10}$$

$$\text{R(inhomBLE)} < \text{R(inhomBLUE)} < \text{R(inhomBLUUE)} \tag{11}$$

$$\text{R(homBLE)} < \text{R(homBLUE)} < \text{R(homBLUUE)} \tag{12}$$

In fact R(inhomBLE) = R(inhomBLUE). In general it can be said that the better an estimate, the more useless it is.

Model Class	$\mathbf{d}=\mathbf{U}\mathbf{z}+\mathbf{e}$	$\mathbf{m}$	$\mathbf{C}$
simple	$\begin{bmatrix}\mathbf{b}\\ \mathbf{t}\end{bmatrix}=\begin{bmatrix}\mathbf{0} & \mathbf{I}\\ \mathbf{Q_x} & \mathbf{Q_y}\end{bmatrix}\begin{bmatrix}\mathbf{x}\\ \mathbf{y}\end{bmatrix}+\begin{bmatrix}\mathbf{v}\\ \mathbf{0}\end{bmatrix}$	$\begin{bmatrix}\mathbf{m_v}\\ \mathbf{0}\end{bmatrix}$	$\begin{bmatrix}\mathbf{C_v} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}$
stochastic parameters type A	$\begin{bmatrix}\mathbf{b}\\ \mathbf{t}\end{bmatrix}=\begin{bmatrix}\mathbf{0} & \mathbf{I}\\ \mathbf{Q_x} & \mathbf{Q_y}\end{bmatrix}\begin{bmatrix}\mathbf{x}\\ \mathbf{y}\end{bmatrix}+\begin{bmatrix}\mathbf{v}\\ \mathbf{Q_s s}\end{bmatrix}$	$\begin{bmatrix}\mathbf{m_v}\\ \mathbf{Q_s m_s}\end{bmatrix}$	$\begin{bmatrix}\mathbf{C_v} & \mathbf{0}\\ \mathbf{0} & \mathbf{Q_s C_s Q_s^{\mathit{T}}}\end{bmatrix}$
stochastic parameters type B	$\begin{bmatrix}\mathbf{b}\\ \mathbf{t}\end{bmatrix}=\begin{bmatrix}\mathbf{0} & \mathbf{I}\\ \mathbf{Q_x} & \mathbf{Q_y}\end{bmatrix}\begin{bmatrix}\mathbf{x}\\ \mathbf{y}\end{bmatrix}+\begin{bmatrix}\mathbf{v}+\mathbf{Q_s s}\\ \mathbf{0}\end{bmatrix}$	$\begin{bmatrix}\mathbf{m_v}+\mathbf{Q_s m_s}\\ \mathbf{0}\end{bmatrix}$	$\begin{bmatrix}\mathbf{C_v}+\mathbf{Q_s C_s Q_s^{\mathit{T}}} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}$
with prior information	$\begin{bmatrix}\mathbf{x}^b\\ \mathbf{b}\\ \mathbf{t}\end{bmatrix}=\begin{bmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\\ \mathbf{Q_x} & \mathbf{Q_y}\end{bmatrix}\begin{bmatrix}\mathbf{x}\\ \mathbf{y}\end{bmatrix}+\begin{bmatrix}\mathbf{v_x}\\ \mathbf{v}\\ \mathbf{0}\end{bmatrix}$	$\begin{bmatrix}\mathbf{m_x}\\ \mathbf{m_v}\\ \mathbf{0}\end{bmatrix}$	$\begin{bmatrix}\mathbf{C_x} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{C_v} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{0}\end{bmatrix}$
with partial prior information	$\begin{bmatrix}\mathbf{x}_2^b\\ \mathbf{b}\\ \mathbf{t}\end{bmatrix}=\begin{bmatrix}\mathbf{0} & \mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{I}\\ \mathbf{Q}_{\mathbf{x}_1} & \mathbf{Q}_{\mathbf{x}_2} & \mathbf{Q_y}\end{bmatrix}\begin{bmatrix}\mathbf{x}_1\\ \mathbf{x}_2\\ \mathbf{y}\end{bmatrix}+\begin{bmatrix}\mathbf{v}_2\\ \mathbf{v}\\ \mathbf{0}\end{bmatrix}$	$\begin{bmatrix}\mathbf{m}_2\\ \mathbf{m_v}\\ \mathbf{0}\end{bmatrix}$	$\begin{bmatrix}\mathbf{C}_2 & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{C_v} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{0}\end{bmatrix}$
stochastic parameters and prior information	$\begin{bmatrix}\mathbf{x}^b\\ \mathbf{b}\\ \mathbf{t}\end{bmatrix}=\begin{bmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\\ \mathbf{Q_x} & \mathbf{Q_y}\end{bmatrix}\begin{bmatrix}\mathbf{x}\\ \mathbf{y}\end{bmatrix}+\begin{bmatrix}\mathbf{v_x}\\ \mathbf{v}\\ \mathbf{0}\end{bmatrix}$	$\begin{bmatrix}\mathbf{m_x}\\ \mathbf{m_v}\\ \mathbf{Q_s m_s}\end{bmatrix}$	$\begin{bmatrix}\mathbf{C_x} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{C_v} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{Q_s C_s Q_s^{\mathit{T}}}\end{bmatrix}$
stochastic parameters partial prior information	$\begin{bmatrix}\mathbf{x}_2^b\\ \mathbf{b}\\ \mathbf{t}\end{bmatrix}=\begin{bmatrix}\mathbf{0} & \mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{I}\\ \mathbf{Q}_{\mathbf{x}_1} & \mathbf{Q}_{\mathbf{x}_2} & \mathbf{Q_y}\end{bmatrix}\begin{bmatrix}\mathbf{x}_1\\ \mathbf{x}_2\\ \mathbf{y}\end{bmatrix}+\begin{bmatrix}\mathbf{v}_2\\ \mathbf{v}\\ \mathbf{0}\end{bmatrix}$	$\begin{bmatrix}\mathbf{m}_2\\ \mathbf{m_v}\\ \mathbf{Q_s m_s}\end{bmatrix}$	$\begin{bmatrix}\mathbf{C}_2 & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{C_v} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{Q_s C_s Q_s^{\mathit{T}}}\end{bmatrix}$

Table 2.2: The matrices of the "data equations" $\mathbf{d}=\mathbf{U}\mathbf{z}+\mathbf{e}$, $\mathbf{e}\sim(\mathbf{m},\mathbf{C})$ for the various classes of specific models

From this general solution, the general solutions will be derived for the particular model classes: simple models (i.e. without stochastic parameters, other than observational noise), models with stochastic parameters, models with prior information, models with partial prior information, models with stochastic parameters and prior information, models with stochastic parameters and partial prior information. From these the specialized solution for any specific model will follow in a straightforward way.

The derivation of the general solution for each particular model class follows directly from the general solution, by simply replacing the matrices $\mathbf{d}$, $\mathbf{U}$, $\mathbf{z}$, $\mathbf{m}$, $\mathbf{C}$ with their appropriate values summarized in table 2.2.

The solutions for the specific models within each class are determined from the general solution of the class by specifying the values of the matrices $\mathbf{Q_x}$, $\mathbf{Q_y}$, $\mathbf{Q_s}$ and $\mathbf{t}$ in the linearized model equations. Tables with the values of these matrices have been given in section 1.6.

2.3. General solution for simple models

Using the values of $\mathbf{d}$, $\mathbf{U}$, $\mathbf{m}$ and $\mathbf{C}$ for the simple models from table 2.2, in the estimates of table 2.1, the general solution is obtained for the various estimation types:

inhomBLE / inhomBLUE / homBLE ($\mathbf{t}\neq\mathbf{0}$) */ homBLUE* ($\mathbf{t}\neq\mathbf{0}$)

$$\mathbf{C_v h}=\mathbf{0} \qquad\qquad \hat{q}=q+\mathbf{h}^T(\mathbf{b}-\mathbf{y}-\mathbf{m_v})$$

homBLE ($\mathbf{t}=\mathbf{0}$)

$$[\mathbf{C_v}+(\mathbf{y}-\mathbf{m_v})(\mathbf{y}-\mathbf{m_v})^T]\mathbf{h}=q(\mathbf{y}+\mathbf{m_v}) \qquad \hat{q}=q+\mathbf{h}^T\mathbf{b}$$

homBLUE ($\mathbf{t}=\mathbf{0}$)

$$\begin{bmatrix}\mathbf{C_v} & \mathbf{y}+\mathbf{m_v}\\(\mathbf{y}+\mathbf{m_v})^T & 0\end{bmatrix}\begin{bmatrix}\mathbf{h}\\-\nu\end{bmatrix}=\begin{bmatrix}\mathbf{0}\\q\end{bmatrix} \qquad \hat{q}=q+\mathbf{h}^T\mathbf{b}$$

inhomBLUUE / homBLUUE ($\mathbf{m_v}=\mathbf{0}$)

$$\hat{q}=\mathbf{c_y}^T(\mathbf{b}-\mathbf{m_v})+\mathbf{g}^T[\mathbf{t}-\mathbf{Q_y}(\mathbf{b}-\mathbf{m_v})] \quad / \quad \hat{q}=\mathbf{c_y}^T\mathbf{b}+\mathbf{g}^T(\mathbf{t}-\mathbf{Q_y b})$$

$$\begin{bmatrix}\mathbf{Q_y C_v Q_y^T} & \mathbf{Q_x}\\\mathbf{Q_x^T} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{g}\\-\mathbf{k_x}\end{bmatrix}=\begin{bmatrix}\mathbf{Q_y C_v c_y}\\\mathbf{c_x}\end{bmatrix}$$

homBLUUE ($\mathbf{m_V}\neq\mathbf{0}$)

$$\hat{q}=\mathbf{c_y}^T(\mathbf{b}-\mathbf{m_v})+\mathbf{g}^T[\mathbf{t}-\mathbf{Q_y}(\mathbf{b}-\mathbf{m_v})]$$

$$\begin{bmatrix}\mathbf{Q_y C_v Q_y^T} & \mathbf{Q_x} & \mathbf{Q_y m_v}\\\mathbf{Q_x^T} & \mathbf{0} & \mathbf{0}\\\mathbf{m_v^T Q_y^T} & \mathbf{0} & 0\end{bmatrix}\begin{bmatrix}\mathbf{g}\\-\mathbf{k_x}\\\nu\end{bmatrix}=\begin{bmatrix}\mathbf{Q_y C_v c_y}\\\mathbf{c_x}\\\mathbf{m_v^T c_y}\end{bmatrix}$$

The general solution for simple models, for the most usual case where $|\mathbf{C_v}|\neq 0$, $|\mathbf{Q_y C_v Q_y^T}|\neq 0$ and $|\mathbf{Q_x^T M^{-1} Q_x}|\neq 0$ is summarized in table 2.3.

2.4. General solution for models with stochastic parameters

2.4.1. Models with stochastic parameters - type A

General form of linear estimates for models with stochastic parameters - type A:

$$\mathbf{b}=\mathbf{y}+\mathbf{v}, \qquad \mathbf{v}\sim(\mathbf{m_v},\mathbf{C_v})$$

$$\mathbf{t}=\mathbf{Q_x x}+\mathbf{Q_y y}+\mathbf{Q_s s} \qquad \mathbf{s}\sim(\mathbf{m_s},\mathbf{C_s})$$

$$(\mathbf{m_t}=\mathbf{Q_x x}+\mathbf{Q_y y}+\mathbf{Q_s m_s})$$

inhomBLE / inhomBLUE

$$\hat{q}=q+\mathbf{h}^T(\mathbf{b}-\mathbf{y}-\mathbf{m_v})+\mathbf{g}^T(\mathbf{t}-\mathbf{m_t}), \qquad \mathbf{C_v h}=\mathbf{0}, \qquad \mathbf{Q_s C_s Q_s^T g}=\mathbf{0}$$

Table 2.3: General form of linear estimates for simple models $\mathbf{b}=\mathbf{y}+\mathbf{v}$, $\quad \mathbf{v}\sim(\mathbf{m_v},\mathbf{C_v})$ $\mathbf{t}=\mathbf{Q_x x}+\mathbf{Q_y y}$ $\mathbf{M}=\mathbf{Q_y C_v Q_y}^T \quad \mathbf{N}=\mathbf{Q_x}^T\mathbf{M}^{-1}\mathbf{Q_x}$ for the regular case $\lvert\mathbf{C_v}\rvert\neq 0$, $\lvert\mathbf{M}\rvert\neq 0$, $\lvert\mathbf{N}\rvert\neq 0$
inhomBLE / inhomBLUE / homBLE ($\mathbf{t}\neq\mathbf{0}$) / homBLUE ($\mathbf{t}\neq\mathbf{0}$): $\quad \hat{q}=q$
homBLE ($\mathbf{t}=\mathbf{0}$): $\quad \hat{q}=\dfrac{(\mathbf{y}+\mathbf{m_v})^T\mathbf{C_v}^{-1}\mathbf{b}}{1+(\mathbf{y}+\mathbf{m_v})^T\mathbf{C_v}^{-1}(\mathbf{y}+\mathbf{m_v})}q$
homBLUE ($\mathbf{t}=\mathbf{0}$): $\quad \hat{q}=\dfrac{(\mathbf{y}+\mathbf{m_v})^T\mathbf{C_v}^{-1}\mathbf{b}}{1+(\mathbf{y}+\mathbf{m_v})^T\mathbf{C_v}^{-1}\mathbf{y}}q$
inhomBLUUE / homBLUUE: $\quad \hat{q}=\mathbf{c_x}^T\hat{\mathbf{x}}+\mathbf{c_y}^T\hat{\mathbf{y}}$ $\hat{\mathbf{x}}=\mathbf{N}^{-1}\mathbf{Q_x}^T\mathbf{M}^{-1}[\mathbf{t}-\mathbf{Q_y}(\mathbf{b}-\alpha\,\mathbf{m_v})]$ $\hat{\mathbf{y}}=(\mathbf{b}-\alpha\,\mathbf{m_v})+\mathbf{C_v Q_y}^T\mathbf{M}^{-1}[\mathbf{t}-\mathbf{Q_x}\hat{\mathbf{x}}-\mathbf{Q_y}(\mathbf{b}-\alpha\,\mathbf{m_v})]=\mathbf{b}-\hat{\mathbf{v}}$ inhomBLUUE / homBLUUE ($\mathbf{m_v}=\mathbf{0}$ & $\mathbf{m_s}=\mathbf{0}$): $\alpha=1$ homBLUUE ($\mathbf{m_v}\neq\mathbf{0}$ or $\mathbf{m_s}\neq\mathbf{0}$): $\alpha=\dfrac{\mathbf{m_v}^T\mathbf{Q_y}^T\mathbf{M}^{-1}\mathbf{H}(\mathbf{Q_y b}-\mathbf{t})}{\mathbf{m_v}^T\mathbf{Q_y}^T\mathbf{M}^{-1}\mathbf{H}\mathbf{Q_y m_v}}$, $\quad \mathbf{H}=\mathbf{I}-\mathbf{Q_x N}^{-1}\mathbf{Q_x}^T\mathbf{M}^{-1}$

Case $\lvert\mathbf{C_v}\rvert\neq 0$: $\quad \hat{q}=q+\mathbf{g}^T(\mathbf{t}-\mathbf{m_t}) \quad \mathbf{Q_s C_s Q_s}^T\mathbf{g}=\mathbf{0}$

Case $\lvert\mathbf{Q_s C_s Q_s}^T\rvert\neq 0$: $\quad \hat{q}=q+\mathbf{h}^T(\mathbf{b}-\mathbf{y}-\mathbf{m_v}) \quad \mathbf{C_v h}=\mathbf{0}$

Case $\lvert\mathbf{C_v}\rvert\neq 0$ & $\lvert\mathbf{Q_s C_s Q_s}^T\rvert\neq 0$: $\quad \hat{q}=q$

homBLE

$$\hat{q}=\mathbf{h}^T\mathbf{b}+\mathbf{g}^T\mathbf{t}, \qquad \begin{bmatrix}\mathbf{C_v}+(\mathbf{y}+\mathbf{m_v})(\mathbf{y}+\mathbf{m_v})^T & (\mathbf{y}+\mathbf{m_v})\mathbf{m_t}^T \\ \mathbf{m_t}(\mathbf{y}+\mathbf{m_v})^T & \mathbf{Q_s C_s Q_s}^T+\mathbf{m_t m_t}^T\end{bmatrix}\begin{bmatrix}\mathbf{h}\\ \mathbf{g}\end{bmatrix}=q\begin{bmatrix}\mathbf{y}+\mathbf{m_v}\\ \mathbf{m_t}\end{bmatrix}$$

Case $\lvert\mathbf{C_v}\rvert\neq 0$: $\quad \hat{q}=\alpha\,q+\mathbf{g}^T(\mathbf{t}-\alpha\,\mathbf{m_t}) \qquad \alpha=\dfrac{(\mathbf{y}+\mathbf{m_v})^T\mathbf{C_v}^{-1}\mathbf{b}}{1+(\mathbf{y}+\mathbf{m_v})^T\mathbf{C_v}^{-1}(\mathbf{y}+\mathbf{m_v})}$

$$\{[1+(\mathbf{y}+\mathbf{m_v})^T\mathbf{C_v}^{-1}(\mathbf{y}+\mathbf{m_v})]\mathbf{Q_s C_s Q_s}^T+\mathbf{m_t m_t}^T\}\,\mathbf{g}=q\,\mathbf{m_t}$$

Case $\lvert\mathbf{Q_s C_s Q_s}^T\rvert\neq 0$: $\quad \hat{q}=\alpha\,q+\mathbf{h}^T[\mathbf{b}-\alpha\,(\mathbf{y}-\mathbf{m_v})] \qquad \alpha=\dfrac{\mathbf{m_t}^T(\mathbf{Q_s C_s Q_s}^T)^{-1}\mathbf{t}}{1+\mathbf{m_t}^T(\mathbf{Q_s C_s Q_s}^T)^{-1}\mathbf{m_t}}$

$$\{[1+\mathbf{m_t}^T(\mathbf{Q_s C_s Q_s}^T)^{-1}\mathbf{m_t}]\mathbf{C_v}+(\mathbf{y}+\mathbf{m_v})(\mathbf{y}+\mathbf{m_v})^T\}\,\mathbf{h}=q\,(\mathbf{y}+\mathbf{m_v})$$

Case $\lvert\mathbf{C_v}\rvert\neq 0$ & $\lvert\mathbf{Q_s C_s Q_s}^T\rvert\neq 0$: $\quad \alpha=\dfrac{(\mathbf{y}+\mathbf{m_v})^T\mathbf{C_v}^{-1}\mathbf{b}+\mathbf{m_t}^T(\mathbf{Q_s C_s Q_s}^T)^{-1}\mathbf{t}}{1+(\mathbf{y}+\mathbf{m_v})^T\mathbf{C_v}^{-1}(\mathbf{y}+\mathbf{m_v})+\mathbf{m_t}^T(\mathbf{Q_s C_s Q_s}^T)^{-1}\mathbf{m_t}}$

homBLUE

$$\hat{q}=\mathbf{h}^T\mathbf{b}+\mathbf{g}^T\mathbf{t} \qquad \begin{bmatrix} \mathbf{C_v} & \mathbf{0} & \mathbf{y}+\mathbf{m_v} \\ \mathbf{0} & \mathbf{Q_s C_s Q_s^T} & \mathbf{m_t} \\ (\mathbf{y}+\mathbf{m_v})^T & \mathbf{m_t^T} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{h} \\ \mathbf{g} \\ -\nu \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ q \end{bmatrix}$$

Case $|\mathbf{C_v}|\neq 0$: $\quad \hat{q}=\alpha\, q+\mathbf{g}^T(\mathbf{t}-\alpha\,\mathbf{m_t}) \qquad \alpha=\dfrac{(\mathbf{y}+\mathbf{m_v})^T\mathbf{C_v^{-1}b}}{(\mathbf{y}+\mathbf{m_v})^T\mathbf{C_v^{-1}}(\mathbf{y}+\mathbf{m_v})}$

$$\{[(\mathbf{y}+\mathbf{m_v})^T\mathbf{C_v^{-1}}(\mathbf{y}+\mathbf{m_v})]\mathbf{Q_s C_s Q_s^T}+\mathbf{m_t m_t^T}\}\mathbf{g}=q\,\mathbf{m_t}$$

Case $|\mathbf{Q_s C_s Q_s^T}|\neq 0$: $\quad \hat{q}=\alpha\, q+\mathbf{h}^T[\mathbf{b}-\alpha\,(\mathbf{y}-\mathbf{m_v})] \qquad \alpha=\dfrac{\mathbf{m_t^T}(\mathbf{Q_s C_s Q_s^T})^{-1}\mathbf{t}}{\mathbf{m_t^T}(\mathbf{Q_s C_s Q_s^T})^{-1}\mathbf{m_t}}$

$$\{[\mathbf{m_t^T}(\mathbf{Q_s C_s Q_s^T})^{-1}\mathbf{m_t}]\mathbf{C_v}+(\mathbf{y}+\mathbf{m_v})(\mathbf{y}+\mathbf{m_v})^T\}\mathbf{h}=q\,(\mathbf{y}+\mathbf{m_v})$$

Case $|\mathbf{C_v}|\neq 0$ & $|\mathbf{Q_s C_s Q_s^T}|\neq 0$: $\quad \hat{q}=\dfrac{(\mathbf{y}+\mathbf{m_v})^T\mathbf{C_v^{-1}b}+\mathbf{m_t^T}(\mathbf{Q_s C_s Q_s^T})^{-1}\mathbf{t}}{(\mathbf{y}+\mathbf{m_v})^T\mathbf{C_v^{-1}}(\mathbf{y}+\mathbf{m_v})+\mathbf{m_t^T}(\mathbf{Q_s C_s Q_s^T})^{-1}\mathbf{m_t}}$

inhomBLUUE / homBLUUE ($\mathbf{m_v}=\mathbf{0}$ & $\mathbf{m_s}=\mathbf{0}$)

$$\hat{q}=\mathbf{c_y^T}(\mathbf{b}-\mathbf{m_v})+\mathbf{g}^T[\mathbf{t}-\mathbf{Q_y}(\mathbf{b}-\mathbf{m_v})-\mathbf{Q_s m_s}]$$

$$\begin{bmatrix} \mathbf{M} & \mathbf{Q_x} \\ \mathbf{Q_x^T} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{g} \\ -\mathbf{k_x} \end{bmatrix} = \begin{bmatrix} \mathbf{Q_y C_v c_y} \\ \mathbf{c_x} \end{bmatrix}, \qquad \mathbf{M}=\mathbf{Q_s C_s Q_s^T}+\mathbf{Q_y C_v Q_y^T}$$

Case $|\mathbf{M}|\neq 0$:

$$\hat{q}=\mathbf{c_y^T}(\mathbf{b}-\mathbf{m_v})+(\mathbf{Q_x k_x}+\mathbf{Q_y C_v c_y})^T\mathbf{M}^{-1}[\mathbf{t}-\mathbf{Q_y}(\mathbf{b}-\mathbf{m_v})-\mathbf{Q_s m_s}]$$

$$(\mathbf{Q_x^T M^{-1} Q_x})\mathbf{k_x}=\mathbf{c_x}-\mathbf{Q_x^T M^{-1} Q_y C_v c_y}$$

Case $|\mathbf{M}|\neq 0$ & $|\mathbf{N}|\neq 0$: $\qquad (\mathbf{N}=\mathbf{Q_x^T M^{-1} Q_x})$

$$\hat{q}=\mathbf{c_x^T\hat{x}}+\mathbf{c_y^T\hat{y}}$$

$$\mathbf{\hat{x}}=\mathbf{N^{-1}Q_x^T M^{-1}}[\mathbf{t}-\mathbf{Q_y}(\mathbf{b}-\mathbf{m_v})-\mathbf{Q_s m_s}]$$

$$\mathbf{\hat{y}}=(\mathbf{b}-\mathbf{m_v})+\mathbf{C_v Q_y^T M^{-1}}[\mathbf{t}-\mathbf{Q_x\hat{x}}-\mathbf{Q_y}(\mathbf{b}-\mathbf{m_v})-\mathbf{Q_s m_s}]=\mathbf{b}-\mathbf{\hat{v}}$$

$$\mathbf{\hat{s}}=\mathbf{m_s}+\mathbf{C_s Q_s^T M^{-1}}[\mathbf{t}-\mathbf{Q_x\hat{x}}-\mathbf{Q_y}(\mathbf{b}-\mathbf{m_v})-\mathbf{Q_s m_s}] \qquad \text{(compatible)}$$

homBLUUE ($\mathbf{m_v}\neq\mathbf{0}$ or $\mathbf{m_s}\neq\mathbf{0}$, or both)

$$\hat{q}=\mathbf{c_y^T}(\mathbf{b}-\mathbf{m_v})+\mathbf{g}^T[\mathbf{t}-\mathbf{Q_y}(\mathbf{b}-\mathbf{m_v})-\mathbf{Q_s m_s}]$$

$$\begin{bmatrix} \mathbf{M} & \mathbf{Q_x} & \mathbf{Q_y m_v}-\mathbf{Q_s m_s} \\ \mathbf{Q_x^T} & \mathbf{0} & \mathbf{0} \\ (\mathbf{Q_y m_v}-\mathbf{Q_s m_s})^T & \mathbf{0} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{g} \\ -\mathbf{k_x} \\ \nu \end{bmatrix} = \begin{bmatrix} \mathbf{Q_y C_v c_y} \\ \mathbf{c_x} \\ \mathbf{m_v^T c_y} \end{bmatrix}$$

Case $|\mathbf{M}|\neq 0$:

$$\hat{q}=\mathbf{c}_{\mathrm{y}}^{T}(\mathbf{b}-\mathbf{m}_{\mathrm{v}})+[\mathbf{Q}_{\mathrm{x}}\mathbf{k}_{\mathrm{x}}-\nu(\mathbf{Q}_{\mathrm{y}}\mathbf{m}_{\mathrm{v}}-\mathbf{Q}_{\mathrm{s}}\mathbf{m}_{\mathrm{s}})+\mathbf{Q}_{\mathrm{y}}\mathbf{C}_{\mathrm{v}}\mathbf{c}_{\mathrm{y}}]^{T}\mathbf{M}^{-1}[\mathbf{t}-\mathbf{Q}_{\mathrm{y}}(\mathbf{b}-\mathbf{m}_{\mathrm{v}})-\mathbf{Q}_{\mathrm{s}}\mathbf{m}_{\mathrm{s}}]$$

$$\begin{bmatrix}\mathbf{Q}_{\mathrm{x}}^{T}\mathbf{M}^{-1}\mathbf{Q}_{\mathrm{x}} & \mathbf{Q}_{\mathrm{x}}^{T}\mathbf{M}^{-1}(\mathbf{Q}_{\mathrm{y}}\mathbf{m}_{\mathrm{v}}-\mathbf{Q}_{\mathrm{s}}\mathbf{m}_{\mathrm{s}})\\ (\mathbf{Q}_{\mathrm{y}}\mathbf{m}_{\mathrm{v}}-\mathbf{Q}_{\mathrm{s}}\mathbf{m}_{\mathrm{s}})^{T}\mathbf{M}^{-1}\mathbf{Q}_{\mathrm{x}} & (\mathbf{Q}_{\mathrm{y}}\mathbf{m}_{\mathrm{v}}-\mathbf{Q}_{\mathrm{s}}\mathbf{m}_{\mathrm{s}})^{T}\mathbf{M}^{-1}(\mathbf{Q}_{\mathrm{y}}\mathbf{m}_{\mathrm{v}}-\mathbf{Q}_{\mathrm{s}}\mathbf{m}_{\mathrm{s}})\end{bmatrix}\begin{bmatrix}\mathbf{k}_{\mathrm{x}}\\ -\nu\end{bmatrix}=$$
$$=\begin{bmatrix}\mathbf{c}_{\mathrm{x}}-\mathbf{Q}_{\mathrm{x}}^{T}\mathbf{M}^{-1}\mathbf{Q}_{\mathrm{y}}\mathbf{C}_{\mathrm{v}}\mathbf{c}_{\mathrm{y}}\\ \mathbf{m}_{\mathrm{v}}^{T}\mathbf{c}_{\mathrm{x}}-(\mathbf{Q}_{\mathrm{y}}\mathbf{m}_{\mathrm{v}}-\mathbf{Q}_{\mathrm{s}}\mathbf{m}_{\mathrm{s}})^{T}\mathbf{M}^{-1}\mathbf{Q}_{\mathrm{y}}\mathbf{C}_{\mathrm{v}}\mathbf{c}_{\mathrm{y}}\end{bmatrix}$$

Case $|\mathbf{M}|\neq 0$ & $|\mathbf{N}|\neq 0$:

$$\hat{q}=\mathbf{c}_{\mathrm{x}}^{T}\hat{\mathbf{x}}+\mathbf{c}_{\mathrm{y}}^{T}\hat{\mathbf{y}}$$

$$\hat{\mathbf{x}}=\mathbf{N}^{-1}\mathbf{Q}_{\mathrm{x}}^{T}\mathbf{M}^{-1}[\mathbf{t}-\mathbf{Q}_{\mathrm{y}}(\mathbf{b}-\alpha\mathbf{m}_{\mathrm{v}})-\alpha\mathbf{Q}_{\mathrm{s}}\mathbf{m}_{\mathrm{s}}]$$

$$\hat{\mathbf{y}}=(\mathbf{b}-\alpha\mathbf{m}_{\mathrm{v}})+\mathbf{C}_{\mathrm{v}}\mathbf{Q}_{\mathrm{y}}^{T}\mathbf{M}^{-1}[\mathbf{t}-\mathbf{Q}_{\mathrm{x}}\hat{\mathbf{x}}-\mathbf{Q}_{\mathrm{y}}(\mathbf{b}-\alpha\mathbf{m}_{\mathrm{v}})-\alpha\mathbf{Q}_{\mathrm{s}}\mathbf{m}_{\mathrm{s}}]=\mathbf{b}-\hat{\mathbf{v}}$$

$$\hat{\mathbf{s}}=\alpha\mathbf{m}_{\mathrm{s}}+\mathbf{C}_{\mathrm{s}}\mathbf{Q}_{\mathrm{s}}^{T}\mathbf{M}^{-1}[\mathbf{t}-\mathbf{Q}_{\mathrm{x}}\hat{\mathbf{x}}-\mathbf{Q}_{\mathrm{y}}(\mathbf{b}-\alpha\mathbf{m}_{\mathrm{v}})-\alpha\mathbf{Q}_{\mathrm{s}}\mathbf{m}_{\mathrm{s}}]\quad\text{(compatible)}$$

$$\alpha=\frac{\mathbf{m}_{\mathrm{s}}^{T}\mathbf{Q}_{\mathrm{s}}^{T}\mathbf{M}^{-1}\mathbf{H}(\mathbf{t}-\mathbf{Q}_{\mathrm{y}}\mathbf{b})}{\mathbf{m}_{\mathrm{s}}^{T}\mathbf{Q}_{\mathrm{s}}^{T}\mathbf{M}^{-1}\mathbf{H}\mathbf{Q}_{\mathrm{s}}\mathbf{m}_{\mathrm{s}}},\quad \mathbf{H}=\mathbf{I}-\mathbf{Q}_{\mathrm{x}}\mathbf{N}^{-1}\mathbf{Q}_{\mathrm{x}}^{T}\mathbf{M}^{-1}$$

Table 2.4:	General form of linear estimates for models with stochastic parameters - type A: $\mathbf{b}=\mathbf{y}+\mathbf{v}$, $\mathbf{t}=\mathbf{Q}_{\mathrm{x}}\mathbf{x}+\mathbf{Q}_{\mathrm{y}}\mathbf{y}+\mathbf{Q}_{\mathrm{s}}\mathbf{s}$ $\quad$ $\mathbf{v}\sim(\mathbf{m}_{\mathrm{v}},\mathbf{C}_{\mathrm{v}})$, $\mathbf{s}\sim(\mathbf{m}_{\mathrm{s}},\mathbf{C}_{\mathrm{s}})$ $\mathbf{m}_{\mathbf{t}}=\mathbf{Q}_{\mathrm{x}}\mathbf{x}+\mathbf{Q}_{\mathrm{y}}\mathbf{y}+\mathbf{Q}_{\mathrm{s}}\mathbf{m}_{\mathrm{s}}$ $\mathbf{M}=\mathbf{Q}_{\mathrm{s}}\mathbf{C}_{\mathrm{s}}\mathbf{Q}_{\mathrm{s}}^{T}+\mathbf{Q}_{\mathrm{y}}\mathbf{C}_{\mathrm{v}}\mathbf{Q}_{\mathrm{y}}^{T}$, $\quad$ $\mathbf{N}=\mathbf{Q}_{\mathrm{x}}^{T}\mathbf{M}^{-1}\mathbf{Q}_{\mathrm{x}}$ for the regular case: $|\mathbf{C}_{\mathrm{v}}|\neq 0$, $|\mathbf{Q}_{\mathrm{s}}\mathbf{C}_{\mathrm{s}}\mathbf{Q}_{\mathrm{s}}^{T}|\neq 0$, $|\mathbf{M}|\neq 0$, $|\mathbf{N}|\neq 0$
inhomBLE / inhomBLUE:	$\hat{q}=q$
homBLE:	$\hat{q}=\dfrac{(\mathbf{y}+\mathbf{m}_{\mathrm{v}})^{T}\mathbf{C}_{\mathrm{v}}^{-1}\mathbf{b}+\mathbf{m}_{\mathbf{t}}^{T}(\mathbf{Q}_{\mathrm{s}}\mathbf{C}_{\mathrm{s}}\mathbf{Q}_{\mathrm{s}}^{T})^{-1}\mathbf{t}}{1+(\mathbf{y}+\mathbf{m}_{\mathrm{v}})^{T}\mathbf{C}_{\mathrm{v}}^{-1}(\mathbf{y}+\mathbf{m}_{\mathrm{v}})+\mathbf{m}_{\mathbf{t}}^{T}(\mathbf{Q}_{\mathrm{s}}\mathbf{C}_{\mathrm{s}}\mathbf{Q}_{\mathrm{s}}^{T})^{-1}\mathbf{m}_{\mathbf{t}}}q$
homBLUE:	$\hat{q}=\dfrac{(\mathbf{y}+\mathbf{m}_{\mathrm{v}})^{T}\mathbf{C}_{\mathrm{v}}^{-1}\mathbf{b}+\mathbf{m}_{\mathbf{t}}^{T}(\mathbf{Q}_{\mathrm{s}}\mathbf{C}_{\mathrm{s}}\mathbf{Q}_{\mathrm{s}}^{T})^{-1}\mathbf{t}}{(\mathbf{y}+\mathbf{m}_{\mathrm{v}})^{T}\mathbf{C}_{\mathrm{v}}^{-1}(\mathbf{y}+\mathbf{m}_{\mathrm{v}})+\mathbf{m}_{\mathbf{t}}^{T}(\mathbf{Q}_{\mathrm{s}}\mathbf{C}_{\mathrm{s}}\mathbf{Q}_{\mathrm{s}}^{T})^{-1}\mathbf{m}_{\mathbf{t}}}q$
inhomBLUUE / homBLUUE:	$\hat{q}=\mathbf{c}_{\mathrm{x}}^{T}\hat{\mathbf{x}}+\mathbf{c}_{\mathrm{y}}^{T}\hat{\mathbf{y}}$ $\hat{\mathbf{x}}=\mathbf{N}^{-1}\mathbf{Q}_{\mathrm{x}}^{T}\mathbf{M}^{-1}[\mathbf{t}-\mathbf{Q}_{\mathrm{y}}(\mathbf{b}-\alpha\mathbf{m}_{\mathrm{v}})-\alpha\mathbf{Q}_{\mathrm{s}}\mathbf{m}_{\mathrm{s}}]$ $\hat{\mathbf{y}}=(\mathbf{b}-\alpha\mathbf{m}_{\mathrm{v}})+\mathbf{C}_{\mathrm{v}}\mathbf{Q}_{\mathrm{y}}^{T}\mathbf{M}^{-1}[\mathbf{t}-\mathbf{Q}_{\mathrm{x}}\hat{\mathbf{x}}-\mathbf{Q}_{\mathrm{y}}(\mathbf{b}-\alpha\mathbf{m}_{\mathrm{v}})-\alpha\mathbf{Q}_{\mathrm{s}}\mathbf{m}_{\mathrm{s}}]=\mathbf{b}-\hat{\mathbf{v}}$ $\hat{\mathbf{s}}=\alpha\mathbf{m}_{\mathrm{s}}+\mathbf{C}_{\mathrm{s}}\mathbf{Q}_{\mathrm{s}}^{T}\mathbf{M}^{-1}[\mathbf{t}-\mathbf{Q}_{\mathrm{x}}\hat{\mathbf{x}}-\mathbf{Q}_{\mathrm{y}}(\mathbf{b}-\alpha\mathbf{m}_{\mathrm{v}})-\alpha\mathbf{Q}_{\mathrm{s}}\mathbf{m}_{\mathrm{s}}]$ (comp.)
inhomBLUUE / homBLUUE ($\mathbf{m}_{\mathrm{v}}=\mathbf{0}$ & $\mathbf{m}_{\mathrm{s}}=\mathbf{0}$):	$\alpha=1$
homBLUUE ($\mathbf{m}_{\mathrm{v}}\neq\mathbf{0}$ or $\mathbf{m}_{\mathrm{s}}\neq\mathbf{0}$):	$\alpha=\dfrac{(\mathbf{Q}_{\mathrm{y}}\mathbf{m}_{\mathrm{v}}-\mathbf{Q}_{\mathrm{s}}\mathbf{m}_{\mathrm{s}})^{T}\mathbf{M}^{-1}\mathbf{H}(\mathbf{Q}_{\mathrm{y}}\mathbf{b}-\mathbf{t})}{(\mathbf{Q}_{\mathrm{y}}\mathbf{m}_{\mathrm{v}}-\mathbf{Q}_{\mathrm{s}}\mathbf{m}_{\mathrm{s}})^{T}\mathbf{M}^{-1}\mathbf{H}(\mathbf{Q}_{\mathrm{y}}\mathbf{m}_{\mathrm{v}}-\mathbf{Q}_{\mathrm{s}}\mathbf{m}_{\mathrm{s}})}$, $\mathbf{H}=\mathbf{I}-\mathbf{Q}_{\mathrm{x}}\mathbf{N}^{-1}\mathbf{Q}_{\mathrm{x}}^{T}\mathbf{M}^{-1}$

2.4.2. Models with stochastic parameters - type B

General form of linear estimates for models with stochastic parameters - type B:

$$\mathbf{b}=\mathbf{y}+\mathbf{v}+\mathbf{Q_s s} \;, \qquad \mathbf{v}\sim(\mathbf{m_v},\mathbf{C_v}) \,, \qquad \mathbf{s}\sim(\mathbf{m_s},\mathbf{C_s})$$

$$\mathbf{t}=\mathbf{Q_x x}+\mathbf{Q_y y}$$

Notation:

$$\mathbf{R}=\mathbf{Q_s C_s Q_s^{\mathit{T}}}+\mathbf{C_v} \qquad \mathbf{M}=\mathbf{Q_y R Q_y^{\mathit{T}}}$$

inhomBLE / inhomBLUE / homBLE $(\mathbf{t}\neq\mathbf{0}\;\; \mathbf{t}\neq\mathbf{0})$ */ homBLUE* $(\mathbf{t}\neq\mathbf{0})$

$$\hat{q}=q+\mathbf{h}^T(\mathbf{b}-\mathbf{y}-\mathbf{m_v}-\mathbf{m_s}) \qquad \mathbf{Rh}=\mathbf{0}\,,$$

Case $|\mathbf{R}|\neq 0$: $\hat{q}=q$

homBLE $(\mathbf{t}=\mathbf{0})$)

$$\hat{q}=\mathbf{h}^T\mathbf{b}$$

$$[\mathbf{R}+(\mathbf{y}+\mathbf{Q_s m_s}+\mathbf{m_v})(\mathbf{y}+\mathbf{Q_s m_s}+\mathbf{m_v})^T]\mathbf{h}=q(\mathbf{y}+\mathbf{Q_s m_s}+\mathbf{m_v})$$

Case $|\mathbf{R}|\neq 0$: $\hat{q}=\dfrac{(\mathbf{y}+\mathbf{Q_s m_s}+\mathbf{m_v})^T\mathbf{R}^{-1}\mathbf{b}}{1+(\mathbf{y}+\mathbf{Q_s m_s}+\mathbf{m_v})^T\mathbf{R}^{-1}(\mathbf{y}+\mathbf{Q_s m_s}+\mathbf{m_v})}\,q$

homBLUE $(\mathbf{t}=\mathbf{0})$

$$\hat{q}=\mathbf{h}^T\mathbf{b} \qquad \begin{bmatrix}\mathbf{R} & \mathbf{y}+\mathbf{Q_s m_s}+\mathbf{m_v}\\ (\mathbf{y}+\mathbf{Q_s m_s}+\mathbf{m_v})^T & 0\end{bmatrix}\begin{bmatrix}\mathbf{h}\\ \nu\end{bmatrix}=\begin{bmatrix}\mathbf{0}\\ q\end{bmatrix}$$

Case $|\mathbf{R}|\neq 0$: $\hat{q}=\dfrac{(\mathbf{y}+\mathbf{Q_s m_s}+\mathbf{m_v})^T\mathbf{R}^{-1}\mathbf{b}}{(\mathbf{y}+\mathbf{Q_s m_s}+\mathbf{m_v})^T\mathbf{R}^{-1}(\mathbf{y}+\mathbf{Q_s m_s}+\mathbf{m_v})}\,q$

inhomBLUUE / homBLUUE $(\mathbf{Q_s m_s}-\mathbf{m_v}=\mathbf{0})$

$$\hat{q}=\mathbf{c_y}^T(\mathbf{b}-\mathbf{Q_s m_s}-\mathbf{m_v})+\mathbf{g}^T[\mathbf{t}-\mathbf{Q_y}(\mathbf{b}-\mathbf{Q_s m_s}-\mathbf{m_v})]$$

$$\begin{bmatrix}\mathbf{Q_y R Q_y^{\mathit{T}}} & \mathbf{Q_x}\\ \mathbf{Q_x^{\mathit{T}}} & 0\end{bmatrix}\begin{bmatrix}\mathbf{g}\\ -\mathbf{k_x}\end{bmatrix}=\begin{bmatrix}\mathbf{Q_y R c_y}\\ \mathbf{c_x}\end{bmatrix}$$

Case $|\mathbf{M}|\neq 0$: $\qquad (\mathbf{M}=\mathbf{Q_y R Q_y^{\mathit{T}}})$

$$\hat{q}=\mathbf{c_y}^T(\mathbf{b}-\mathbf{Q_s m_s}-\mathbf{m_v})+[\mathbf{Q_x k_x}+\mathbf{Q_y R c_y}]^T\mathbf{M}^{-1}[\mathbf{t}-\mathbf{Q_y}(\mathbf{b}-\mathbf{Q_s m_s}-\mathbf{m_v})]$$

$$(\mathbf{Q_x^{\mathit{T}} M^{-1} Q_x})\mathbf{k_x}=\mathbf{c_x}-\mathbf{Q_x^{\mathit{T}} R^{-1} Q_y R c_y}$$

Case $|\mathbf{M}|\neq 0$ & $|\mathbf{N}|\neq 0$: $\qquad (\mathbf{N}=\mathbf{Q_x^{\mathit{T}} M^{-1} Q_x})$

$$\hat{q}=\mathbf{c_x}^T\hat{\mathbf{x}}+\mathbf{c_y}^T\hat{\mathbf{y}}$$

$$\hat{\mathbf{x}}=\mathbf{N}^{-1}\mathbf{Q_x^{\mathit{T}} M^{-1}}[\mathbf{t}-\mathbf{Q_y}(\mathbf{b}-\mathbf{m_v}-\mathbf{Q_s m_s})]$$

$$\hat{\mathbf{y}} = (\mathbf{b} - \mathbf{m}_{\mathbf{v}} - \mathbf{Q}_{\mathbf{s}}\mathbf{m}_{\mathbf{s}}) + \mathbf{C}_{\mathbf{v}}\mathbf{Q}_{\mathbf{y}}^{T}\mathbf{M}^{-1}[\mathbf{t} - \mathbf{Q}_{\mathbf{x}}\hat{\mathbf{x}} - \mathbf{Q}_{\mathbf{y}}(\mathbf{b} - \mathbf{m}_{\mathbf{v}} - \mathbf{Q}_{\mathbf{s}}\mathbf{m}_{\mathbf{s}})]$$

homBLUUE $(\mathbf{Q}_{\mathbf{s}}\mathbf{m}_{\mathbf{s}} + \mathbf{m}_{\mathbf{v}} \neq \mathbf{0})$

$$\hat{q} = \mathbf{c}_{\mathbf{y}}^{T}(\mathbf{b} - \mathbf{m}_{\mathbf{v}} - \mathbf{Q}_{\mathbf{s}}\mathbf{m}_{\mathbf{s}}) + \mathbf{g}^{T}[\mathbf{t} - \mathbf{Q}_{\mathbf{y}}(\mathbf{b} - \mathbf{m}_{\mathbf{v}} - \mathbf{Q}_{\mathbf{s}}\mathbf{m}_{\mathbf{s}})]$$

$$\begin{bmatrix} \mathbf{M} & \mathbf{Q}_{\mathbf{x}} & \mathbf{Q}_{\mathbf{y}}(\mathbf{Q}_{\mathbf{s}}\mathbf{m}_{\mathbf{s}} + \mathbf{m}_{\mathbf{v}}) \\ \mathbf{Q}_{\mathbf{x}}^{T} & \mathbf{0} & \mathbf{0} \\ (\mathbf{Q}_{\mathbf{s}}\mathbf{m}_{\mathbf{s}} + \mathbf{m}_{\mathbf{v}})^{T}\mathbf{Q}_{\mathbf{y}}^{T} & \mathbf{0} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{g} \\ -\mathbf{k}_{\mathbf{x}} \\ \nu \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_{\mathbf{y}}\mathbf{R}\mathbf{c}_{\mathbf{y}} \\ \mathbf{c}_{\mathbf{x}} \\ (\mathbf{Q}_{\mathbf{s}}\mathbf{m}_{\mathbf{s}} + \mathbf{m}_{\mathbf{v}})^{T}\mathbf{c}_{\mathbf{y}} \end{bmatrix}$$

$$\hat{q} = \mathbf{c}_{\mathbf{y}}^{T}(\mathbf{b} - \mathbf{m}_{\mathbf{v}} - \mathbf{Q}_{\mathbf{s}}\mathbf{m}_{\mathbf{s}}) + [\mathbf{Q}_{\mathbf{x}}\mathbf{k}_{\mathbf{x}} - \nu\,\mathbf{Q}_{\mathbf{y}}(\mathbf{Q}_{\mathbf{s}}\mathbf{m}_{\mathbf{s}} + \mathbf{m}_{\mathbf{v}}) + \mathbf{Q}_{\mathbf{y}}\mathbf{R}\mathbf{c}_{\mathbf{y}}]^{T}\mathbf{R}^{-1}[\mathbf{t} - \mathbf{Q}_{\mathbf{y}}(\mathbf{b} - \mathbf{m}_{\mathbf{v}} - \mathbf{Q}_{\mathbf{s}}\mathbf{m}_{\mathbf{s}})]$$

$$\begin{bmatrix} \mathbf{Q}_{\mathbf{x}}^{T}\mathbf{M}^{-1}\mathbf{Q}_{\mathbf{x}} & \mathbf{Q}_{\mathbf{x}}^{T}\mathbf{M}^{-1}\mathbf{Q}_{\mathbf{y}}(\mathbf{Q}_{\mathbf{s}}\mathbf{m}_{\mathbf{s}} + \mathbf{m}_{\mathbf{v}}) \\ (\mathbf{Q}_{\mathbf{s}}\mathbf{m}_{\mathbf{s}} + \mathbf{m}_{\mathbf{v}})^{T}\mathbf{Q}_{\mathbf{y}}^{T}\mathbf{M}^{-1}\mathbf{Q}_{\mathbf{x}} & (\mathbf{Q}_{\mathbf{s}}\mathbf{m}_{\mathbf{s}} + \mathbf{m}_{\mathbf{v}})^{T}\mathbf{Q}_{\mathbf{y}}^{T}\mathbf{M}^{-1}\mathbf{Q}_{\mathbf{y}}(\mathbf{Q}_{\mathbf{s}}\mathbf{m}_{\mathbf{s}} + \mathbf{m}_{\mathbf{v}}) \end{bmatrix} \begin{bmatrix} \mathbf{k}_{\mathbf{x}} \\ -\nu \end{bmatrix} =$$

$$= \begin{bmatrix} \mathbf{c}_{\mathbf{x}} - \mathbf{Q}_{\mathbf{x}}^{T}\mathbf{M}^{-1}\mathbf{R}\mathbf{C}_{\mathbf{v}}\mathbf{c}_{\mathbf{y}} \\ (\mathbf{Q}_{\mathbf{s}}\mathbf{m}_{\mathbf{s}} + \mathbf{m}_{\mathbf{v}})^{T}(\mathbf{c}_{\mathbf{y}} - \mathbf{Q}_{\mathbf{y}}^{T}\mathbf{M}^{-1}\mathbf{Q}_{\mathbf{y}}\mathbf{R}\mathbf{c}_{\mathbf{y}}) \end{bmatrix}$$

Case $|\mathbf{M}| \neq 0$ & $|\mathbf{N}| \neq 0$:

$$\hat{q} = \mathbf{c}_{\mathbf{x}}^{T}\hat{\mathbf{x}} + \mathbf{c}_{\mathbf{y}}^{T}\hat{\mathbf{y}}$$

$$\hat{\mathbf{x}} = \mathbf{N}^{-1}\mathbf{Q}_{\mathbf{x}}^{T}\mathbf{M}^{-1}[\mathbf{t} - \mathbf{Q}_{\mathbf{y}}(\mathbf{b} - \alpha\,\mathbf{m}_{\mathbf{v}} - \alpha\,\mathbf{Q}_{\mathbf{s}}\mathbf{m}_{\mathbf{s}})]$$

$$\hat{\mathbf{y}} = (\mathbf{b} - \alpha\,\mathbf{m}_{\mathbf{v}} - \alpha\,\mathbf{Q}_{\mathbf{s}}\mathbf{m}_{\mathbf{s}}) + \mathbf{R}\mathbf{Q}_{\mathbf{y}}^{T}\mathbf{M}^{-1}[\mathbf{t} - \mathbf{Q}_{\mathbf{x}}\hat{\mathbf{x}} - \mathbf{Q}_{\mathbf{y}}(\mathbf{b} - \alpha\,\mathbf{m}_{\mathbf{v}} - \alpha\,\mathbf{Q}_{\mathbf{s}}\mathbf{m}_{\mathbf{s}})]$$

$$\alpha = \frac{(\mathbf{Q}_{\mathbf{s}}\mathbf{m}_{\mathbf{s}} + \mathbf{m}_{\mathbf{v}})^{T}\mathbf{Q}_{\mathbf{y}}^{T}\mathbf{M}^{-1}\mathbf{H}(\mathbf{t} - \mathbf{Q}_{\mathbf{y}}\mathbf{b})}{(\mathbf{Q}_{\mathbf{s}}\mathbf{m}_{\mathbf{s}} + \mathbf{m}_{\mathbf{v}})^{T}\mathbf{Q}_{\mathbf{y}}^{T}\mathbf{M}^{-1}\mathbf{H}\mathbf{Q}_{\mathbf{y}}(\mathbf{Q}_{\mathbf{s}}\mathbf{m}_{\mathbf{s}} + \mathbf{m}_{\mathbf{v}})}, \qquad \mathbf{H} = \mathbf{I} - \mathbf{Q}_{\mathbf{x}}\mathbf{N}^{-1}\mathbf{Q}_{\mathbf{x}}^{T}\mathbf{M}^{-1}$$

2.5. Summary and discussion

It has been shown here that an estimate $\hat{q}$ of any deterministic model parameter q, i.e., any parameter which is a function of the general form $q = \mathbf{c}_{\mathbf{x}}^{T}\mathbf{x} + \mathbf{c}_{\mathbf{y}}^{T}\mathbf{y}$ of the parameters **x** used in the model and the observables **y**, has the form $\hat{q} = \mathbf{c}_{\mathbf{x}}^{T}\hat{\mathbf{x}} + \mathbf{c}_{\mathbf{y}}^{T}\hat{\mathbf{y}}$, where $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are the corresponding estimates of **x** and **y** respectively. The estimates $\hat{q}$, $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ vary according to the principle of estimation used (inhomogeneous or homogeneous, biased, unbiased or uniformly unbiased).

The equations for the computation of the estimates $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ vary also according to which specific model is used for the description of the same general model. In fact the parameters x may be different in different models or they may be completely absent (condition equations). However the estimate according to the chosen estimation principle, is the same for any model parameter q (e.g., for the elements of **y**), no matter which specific model has been chosen for the description of the same general model, except for the effect of different linearization errors or computational errors.

Among the various types of estimates, only the uniformly unbiased ones (inhomBLUUE and homBLUUE) can be directly used. These two types are identical in models without stochastic parameters where the observational errors are always assumed to have zero means, as well as in models with stochastic parameters which have zero means. For this reason the distinction is usually dropped in the literature, and the term unbiased is used with the meaning of uniformly unbiased since no distinction need to be made with the useless simply unbiased estimates. The common inhomBLUUE-homBLUUE estimate is therefore referred as the BLUE estimate. Here the term BLUUE will be preferred.

Table 2.5: General form of linear estimates for models with stochastic parameters - type B:

$\mathbf{b}=\mathbf{y}+\mathbf{Q_s}\mathbf{s}+\mathbf{v}$, $\mathbf{v}\sim(\mathbf{m_v},\mathbf{C_v})$, $\mathbf{s}\sim(\mathbf{m_s},\mathbf{C_s})$

$\mathbf{t}=\mathbf{Q_x}\mathbf{x}+\mathbf{Q_y}\mathbf{y}$

$\mathbf{R}=\mathbf{Q_s}\mathbf{C_s}\mathbf{Q_s}^T+\mathbf{C_v}$, $\mathbf{M}=\mathbf{Q_y}\mathbf{R}\mathbf{Q_y}^T$, $\mathbf{N}=\mathbf{Q_x}^T\mathbf{M}^{-1}\mathbf{Q_x}$

For the regular case: $|\mathbf{C_v}|\neq 0$, $|\mathbf{R}|\neq 0$, $|\mathbf{M}|\neq 0$, $|\mathbf{N}|\neq 0$

inhomBLE / inhomBLUE / homBLE ($\mathbf{t}\neq\mathbf{0}$) / homBLUE ($\mathbf{t}\neq\mathbf{0}$):

$$\hat{q}=q$$

homBLE ($\mathbf{t}=\mathbf{0}$):

$$\hat{q}=\frac{(\mathbf{y}+\mathbf{Q_s}\mathbf{m_s}+\mathbf{m_v})^T\mathbf{R}^{-1}\mathbf{b}}{1+(\mathbf{y}+\mathbf{Q_s}\mathbf{m_s}+\mathbf{m_v})^T\mathbf{R}^{-1}(\mathbf{y}+\mathbf{Q_s}\mathbf{m_s}+\mathbf{m_v})}q$$

homBLUE ($\mathbf{t}=\mathbf{0}$):

$$\hat{q}=\frac{(\mathbf{y}+\mathbf{Q_s}\mathbf{m_s}+\mathbf{m_v})^T\mathbf{R}^{-1}\mathbf{b}}{(\mathbf{y}+\mathbf{Q_s}\mathbf{m_s}+\mathbf{m_v})^T\mathbf{R}^{-1}(\mathbf{y}+\mathbf{Q_s}\mathbf{m_s}+\mathbf{m_v})}q$$

inhomBLUUE / homBLUUE

$$\hat{q}=\mathbf{c_x}^T\hat{\mathbf{x}}+\mathbf{c_y}^T\hat{\mathbf{y}} \qquad \hat{\mathbf{x}}=\mathbf{N}^{-1}\mathbf{Q_x}^T\mathbf{M}^{-1}[\mathbf{t}-\mathbf{Q_y}(\mathbf{b}-\alpha\,\mathbf{m_v}-\alpha\,\mathbf{Q_s}\mathbf{m_s})]$$

$$\hat{\mathbf{y}}=(\mathbf{b}-\alpha\,\mathbf{m_v}-\alpha\,\mathbf{Q_s}\mathbf{m_s})+\mathbf{R}\mathbf{Q_y}^T\mathbf{M}^{-1}[\mathbf{t}-\mathbf{Q_x}\hat{\mathbf{x}}-\mathbf{Q_y}(\mathbf{b}-\alpha\,\mathbf{m_v}-\alpha\,\mathbf{Q_s}\mathbf{m_s})]$$

inhomBLUUE / homBLUUE ($Q_S\, m_S + m_V = 0$):

$$\alpha=1$$

homBLUUE ($\mathbf{Q_s}\mathbf{m_s}+\mathbf{m_v}\neq\mathbf{0}$):

$$\alpha=\frac{(\mathbf{Q_s}\mathbf{m_s}+\mathbf{m_v})^T\mathbf{Q_y}^T\mathbf{M}^{-1}\mathbf{H}(\mathbf{t}-\mathbf{Q_y}\mathbf{b})}{(\mathbf{Q_s}\mathbf{m_s}+\mathbf{m_v})^T\mathbf{Q_y}^T\mathbf{M}^{-1}\mathbf{H}\mathbf{Q_y}(\mathbf{Q_s}\mathbf{m_s}+\mathbf{m_v})}, \qquad \mathbf{H}=\mathbf{I}-\mathbf{Q_x}\mathbf{N}^{-1}\mathbf{Q_x}^T\mathbf{M}^{-1}$$

In models with stochastic parameters, which have non-zero means, inhomBLUUE differs from homBLUUE. Since inhomBLUUE has mean square error which is smaller or equal to that of homBLUUE, it seems reasonable to prefer the inhomogeneous estimate over the homogeneous one, which in addition is computationally more complicated. It may be argued however, that homBLUUE has a nice property of robustness against a wrong mean vector $\mathbf{m_s}$ of the stochastic parameters $\mathbf{s}$. Indeed in the homBLUUE estimates $\mathbf{m_s}$ is controlled by a factor α which is a function of the observations, while inhomBLUUE has the same form as homBLUUE with the specific value $\alpha=1$.

When the original model with stochastic parameters $\mathbf{s}$ is non-linear, the mean values $\mathbf{m_s}$ are an obvious choice for the approximate values $\mathbf{s}^0$ needed for the linearization, and the linearized model has reduced stochastic parameters $\delta\mathbf{s}=\mathbf{s}-\mathbf{s}^0$ with zero means. When $\mathbf{s}^0\neq\mathbf{m_s}$, the reduced stochastic parameters $\delta\mathbf{s}$ have means $\mathbf{m}_{\delta\mathbf{s}}=\mathbf{m_s}-\mathbf{s}^0$. As a consequence the homBLUUE estimates depend on the approximate values used in the linearization!

In the following the BLUUE estimates for the various specific models are summarized for easy reference.

BLUUE:

<u>Simple models</u> $\mathbf{v}\sim(\mathbf{m_v},\mathbf{C_v})$

Observation equations $\mathbf{b}=\mathbf{Ax}+\mathbf{v}$

$$\hat{\mathbf{x}}=(\mathbf{A}^T\mathbf{C}_\mathbf{v}^{-1}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{C}_\mathbf{v}^{-1}(\mathbf{b}-\alpha\,\mathbf{m}_\mathbf{v}), \qquad \hat{\mathbf{y}}=\mathbf{A}\hat{\mathbf{x}}=\mathbf{b}-\hat{\mathbf{v}}$$

$$\alpha=1 \quad \text{or} \quad \alpha=\frac{\mathbf{m}_\mathbf{v}^T\mathbf{C}_\mathbf{v}^{-1}\mathbf{Hb}}{\mathbf{m}_\mathbf{v}^T\mathbf{C}_\mathbf{v}^{-1}\mathbf{Hm}_\mathbf{v}} \qquad \mathbf{H}=\mathbf{I}-\mathbf{A}(\mathbf{A}^T\mathbf{C}_\mathbf{v}^{-1}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{C}_\mathbf{v}^{-1}$$

Condition equations $\mathbf{Bv}=\mathbf{w}$

$$\mathbf{M}=\mathbf{BC}_\mathbf{v}\mathbf{B}^T$$

$$\hat{\mathbf{y}}=(\mathbf{b}-\alpha\,\mathbf{m}_\mathbf{v})-\mathbf{C}_\mathbf{v}\mathbf{B}^T\mathbf{M}^{-1}[\mathbf{w}-\mathbf{B}(\alpha\,\mathbf{m}_\mathbf{v})]=\mathbf{b}-\hat{\mathbf{v}}$$

$$\alpha=1 \quad \text{or} \quad \alpha=\frac{\mathbf{m}_\mathbf{v}^T\mathbf{B}^T\mathbf{M}^{-1}\mathbf{b}}{\mathbf{m}_\mathbf{v}^T\mathbf{B}^T\mathbf{M}^{-1}\mathbf{m}_\mathbf{v}}$$

Mixed equations $\mathbf{w}=-\mathbf{Ax}+\mathbf{Bv}$

$$\mathbf{M}=\mathbf{BC}_\mathbf{v}\mathbf{B}^T \qquad \mathbf{N}=\mathbf{A}^T\mathbf{M}^{-1}\mathbf{A}$$

$$\hat{\mathbf{x}}=-\mathbf{N}^{-1}\mathbf{A}^T\mathbf{M}^{-1}[\mathbf{w}-\mathbf{B}(\alpha\,\mathbf{m}_\mathbf{v})]$$

$$\hat{\mathbf{y}}=(\mathbf{b}-\alpha\,\mathbf{m}_\mathbf{v})-\mathbf{C}_\mathbf{v}\mathbf{B}^T\mathbf{M}^{-1}[\mathbf{w}-\mathbf{B}(\alpha\,\mathbf{m}_\mathbf{v})+\mathbf{A}\hat{\mathbf{x}}]=\mathbf{b}-\hat{\mathbf{v}}$$

$$\alpha=1 \quad \text{or} \quad \alpha=\frac{\mathbf{m}_\mathbf{v}^T\mathbf{B}^T\mathbf{M}^{-1}\mathbf{Hw}}{\mathbf{m}_\mathbf{v}^T\mathbf{B}^T\mathbf{M}^{-1}\mathbf{HBm}_\mathbf{v}} \qquad \mathbf{H}=\mathbf{I}-\mathbf{AN}^{-1}\mathbf{A}^T\mathbf{M}^{-1}$$

Observation equations with constraints $\mathbf{b}=\mathbf{Ax}+\mathbf{v}$, $\mathbf{Hx}=-\mathbf{h}^0$

$$\begin{bmatrix}\hat{\mathbf{x}}\\ \mathbf{k}\end{bmatrix}=\begin{bmatrix}\mathbf{A}^T(\mathbf{C}_\mathbf{v}+\mathbf{m}_\mathbf{v}\mathbf{m}_\mathbf{v}^T)^{-1}\mathbf{A} & \mathbf{H}^T\\ \mathbf{H} & \mathbf{0}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{A}^T(\mathbf{C}_\mathbf{v}+\mathbf{m}_\mathbf{v}\mathbf{m}_\mathbf{v}^T)^{-1}(\mathbf{b}-\alpha\,\mathbf{m}_\mathbf{v})\\ -\mathbf{h}^0\end{bmatrix}$$

$$\hat{\mathbf{y}}=\mathbf{A}\hat{\mathbf{x}}=\mathbf{b}-\hat{\mathbf{v}}$$

$\mathbf{N}=\mathbf{A}^T\mathbf{C}_\mathbf{v}^{-1}\mathbf{A}$ non-singular:

$$\hat{\mathbf{x}}=\hat{\mathbf{x}}_0+\mathbf{N}^{-1}\mathbf{H}^T(\mathbf{HN}^{-1}\mathbf{H}^T)^{-1}(\mathbf{H}\hat{\mathbf{x}}_0+\mathbf{h}^0) \qquad \hat{\mathbf{x}}_0=\mathbf{N}^{-1}\mathbf{A}^T\mathbf{C}_\mathbf{v}^{-1}(\mathbf{b}-\alpha\,\mathbf{m}_\mathbf{v})$$

$$\alpha=1 \quad \text{or} \quad \alpha=\frac{\mathbf{m}_\mathbf{v}^T\mathbf{C}_\mathbf{v}^{-1}\mathbf{Rb}-\mathbf{m}_\mathbf{v}^T\mathbf{C}_\mathbf{v}\mathbf{AN}^{-1}\mathbf{H}^T(\mathbf{HN}^{-1}\mathbf{H}^T)^{-1}(\mathbf{Hx}^*+\mathbf{h}^0)}{\mathbf{m}_\mathbf{v}^T\mathbf{C}_\mathbf{v}^{-1}\mathbf{Rm}_\mathbf{v}}$$

$$\mathbf{R}=\mathbf{I}-\mathbf{AN}^{-1}\mathbf{A}^T\mathbf{C}_\mathbf{v}^{-1} \qquad \mathbf{x}^*=\mathbf{N}^{-1}\mathbf{A}^T\mathbf{C}_\mathbf{v}^{-1}\mathbf{b}$$

<u>Models with prior information</u> $\mathbf{v}\sim(\mathbf{m}_\mathbf{v},\mathbf{C}_\mathbf{v})$ $\mathbf{x}^b\sim(\mathbf{x}+\mathbf{m}_\mathbf{x},\mathbf{C}_\mathbf{x})$

Observation equations

$$\hat{\mathbf{x}}=(\mathbf{x}^b-\alpha\,\mathbf{m}_\mathbf{x})+\mathbf{C}_\mathbf{x}\mathbf{A}^T(\mathbf{AC}_\mathbf{x}\mathbf{A}^T+\mathbf{C}_\mathbf{v})^{-1}[(\mathbf{b}-\alpha\,\mathbf{m}_\mathbf{v})-\mathbf{A}(\mathbf{x}^b-\alpha\,\mathbf{m}_\mathbf{x})]=$$

$$=(\mathbf{A}^T\mathbf{C}_\mathbf{v}^{-1}\mathbf{A}+\mathbf{C}_\mathbf{x}^{-1})^{-1}[\mathbf{A}^T\mathbf{C}_\mathbf{v}^{-1}(\mathbf{b}-\alpha\,\mathbf{m}_\mathbf{v})+\mathbf{C}_\mathbf{x}^{-1}(\mathbf{x}^b-\alpha\,\mathbf{m}_\mathbf{x})]$$

$$\hat{\mathbf{y}}=\mathbf{A}\hat{\mathbf{x}}=\mathbf{b}-\hat{\mathbf{v}}$$

$\alpha=1$ or

$$\alpha = \frac{\mathbf{m}_\mathbf{v}^T \mathbf{C}_\mathbf{v}^{-1}\mathbf{b} + \mathbf{m}_\mathbf{x}^T \mathbf{C}_\mathbf{x}^{-1}\mathbf{x}^b - (\mathbf{A}^T \mathbf{C}_\mathbf{v}^{-1}\mathbf{m}_\mathbf{v} + \mathbf{C}_\mathbf{x}^{-1}\mathbf{m}_\mathbf{x})^T \mathbf{N}^{-1}(\mathbf{A}^T \mathbf{C}_\mathbf{v}^{-1}\mathbf{b} + \mathbf{C}_\mathbf{x}^{-1}\mathbf{x}^b)}{\mathbf{m}_\mathbf{v}^T \mathbf{C}_\mathbf{v}^{-1}\mathbf{m}_\mathbf{v} + \mathbf{m}_\mathbf{x}^T \mathbf{C}_\mathbf{x}^{-1}\mathbf{m}_\mathbf{x} - (\mathbf{A}^T \mathbf{C}_\mathbf{v}^{-1}\mathbf{m}_\mathbf{v} + \mathbf{C}_\mathbf{x}^{-1}\mathbf{m}_\mathbf{x})^T \mathbf{N}^{-1}(\mathbf{A}^T \mathbf{C}_\mathbf{v}^{-1}\mathbf{m}_\mathbf{v} + \mathbf{C}_\mathbf{x}^{-1}\mathbf{m}_\mathbf{x})}$$

Mixed equations

$$\mathbf{M} = \mathbf{BCB}^T$$

$$\hat{\mathbf{x}} = (\mathbf{x}^b - \alpha\, \mathbf{m}_\mathbf{x}) + \mathbf{C}_\mathbf{x}\mathbf{A}^T(\mathbf{AC}_\mathbf{x}\mathbf{A}^T + \mathbf{M})^{-1}[(\mathbf{w} - \alpha\, \mathbf{Bm}_\mathbf{v}) - \mathbf{A}(\mathbf{x}^b - \alpha\, \mathbf{m}_\mathbf{x})] =$$

$$= -(\mathbf{A}^T\mathbf{M}^{-1}\mathbf{A} + \mathbf{C}_\mathbf{x}^{-1})^{-1}[\mathbf{A}^T\mathbf{M}^{-1}(\mathbf{w} - \alpha\, \mathbf{Bm}_\mathbf{v}) + \mathbf{C}_\mathbf{x}^{-1}(\mathbf{x}^b - \alpha\, \mathbf{m}_\mathbf{x})]$$ ### CHECK! ###

$$\hat{\mathbf{y}} = (\mathbf{b} - \alpha\, \mathbf{m}_\mathbf{v} - \mathbf{C}_\mathbf{v}\mathbf{B}^T\mathbf{M}^{-1}[\mathbf{w} - \mathbf{B}(\alpha\, \mathbf{m}_\mathbf{v}) + \mathbf{A}\hat{\mathbf{x}} = \mathbf{b} - \hat{\mathbf{v}}$$

$\alpha = 1$ or

$$\alpha = \frac{\mathbf{m}_\mathbf{v}^T\mathbf{B}^T\mathbf{M}^{-1}\mathbf{b} + \mathbf{m}_\mathbf{x}^T\mathbf{M}^{-1}\mathbf{x}^b - (\mathbf{A}^T\mathbf{M}^{-1}\mathbf{Bm}_\mathbf{v} - \mathbf{C}_\mathbf{x}^{-1}\mathbf{m}_\mathbf{x})^T\mathbf{N}^{-1}(\mathbf{A}^T\mathbf{M}^{-1}\mathbf{w} - \mathbf{C}_\mathbf{x}^{-1}\mathbf{x}^b)}{\mathbf{m}_\mathbf{v}^T\mathbf{B}^T\mathbf{M}^{-1}\mathbf{Bm}_\mathbf{v} + \mathbf{m}_\mathbf{x}^T\mathbf{C}_\mathbf{x}^{-1}\mathbf{m}_\mathbf{x} - (\mathbf{A}^T\mathbf{M}^{-1}\mathbf{Bm}_\mathbf{v} - \mathbf{C}_\mathbf{x}^{-1}\mathbf{m}_\mathbf{x})^T\mathbf{N}^{-1}(\mathbf{A}^T\mathbf{M}^{-1}\mathbf{Bm}_\mathbf{v} - \mathbf{C}_\mathbf{x}^{-1}\mathbf{m}_\mathbf{x})}$$

Models with partial prior information $\mathbf{v} \sim (\mathbf{m}_\mathbf{v}, \mathbf{C}_\mathbf{v})$ $\mathbf{x}_2^b \sim (\mathbf{x} + \mathbf{m}_2, \mathbf{C}_2)$

Observation equations $\mathbf{b} = \mathbf{A}_1\mathbf{x}_1 + \mathbf{A}_2\mathbf{x}_2 + \mathbf{v}$ $\mathbf{b} = \mathbf{A}_1\mathbf{x}_1 + \mathbf{A}_2\mathbf{x}_2 + \mathbf{v}$

$$\mathbf{M}_2 = \mathbf{A}_2\mathbf{C}_2\mathbf{A}_2^T + \mathbf{C}_\mathbf{v}$$

$$\hat{\mathbf{x}}_1 = (\mathbf{A}_1^T\mathbf{M}_2^{-1}\mathbf{A}_1)^{-1}\mathbf{A}_1^T\mathbf{M}_2^{-1}[(\mathbf{b} - \alpha\, \mathbf{m}_\mathbf{v}) - \mathbf{A}_2(\mathbf{x}_2^b - \alpha\, \mathbf{m}_\mathbf{x})]$$

$$\hat{\mathbf{x}}_2 = (\mathbf{x}_2^b - \alpha\, \mathbf{m}_\mathbf{x}) + \mathbf{C}_2\mathbf{A}_2^T\mathbf{M}_2^{-1}[(\mathbf{b} - \alpha\, \mathbf{m}_\mathbf{v}) - \mathbf{A}_2(\mathbf{x}_2^b - \alpha\, \mathbf{m}_\mathbf{x}) - \mathbf{A}_1\hat{\mathbf{x}}_1]$$

$$\hat{\mathbf{y}} = \mathbf{A}_1\hat{\mathbf{x}}_1 + \mathbf{A}_2\hat{\mathbf{x}}_2 = \mathbf{A}\hat{\mathbf{x}} = \mathbf{b} - \hat{\mathbf{v}}$$

$\alpha = 1$ or

$$\alpha = \frac{(\mathbf{m}_\mathbf{v} - \mathbf{A}_2\mathbf{m}_2)^\mathrm{T}\mathbf{M}_2^{-1}\mathbf{H}_2(\mathbf{b} - \mathbf{A}_2\mathbf{x}_2^b)}{(\mathbf{m}_\mathbf{v} - \mathbf{A}_2\mathbf{m}_2)^\mathrm{T}\mathbf{M}_2^{-1}\mathbf{H}_2(\mathbf{m}_\mathbf{v} - \mathbf{A}_2\mathbf{m}_2)}, \qquad \mathbf{H}_2 = \mathbf{I} - \mathbf{A}_1(\mathbf{A}_1^T\mathbf{M}_2^{-1}\mathbf{A}_1)^{-1}\mathbf{A}_1^T\mathbf{M}_2^{-1}$$

Models with stochastic parameters - type A $\mathbf{v} \sim (\mathbf{m}_\mathbf{v}, \mathbf{C}_\mathbf{v})$ $\mathbf{s} \sim (\mathbf{m}_\mathbf{s}, \mathbf{C}_\mathbf{s})$

Mixed model $\mathbf{b} = \mathbf{Ax} + \mathbf{Gs} + \mathbf{v}$

$$\mathbf{M} = \mathbf{GC}_\mathbf{s}\mathbf{G}^T + \mathbf{C}_\mathbf{v} \qquad \mathbf{N} = \mathbf{A}^T\mathbf{M}^{-1}\mathbf{A}$$

$$\hat{\mathbf{x}} = \mathbf{N}^{-1}\mathbf{A}^T\mathbf{M}^{-1}(\mathbf{b} - \alpha\mathbf{m}_\mathbf{v} - \alpha\mathbf{Gm}_\mathbf{s})$$

$$\hat{\mathbf{y}} = (\mathbf{b} - \alpha\mathbf{m}_\mathbf{v}) - \mathbf{C}_\mathbf{v}\mathbf{M}^{-1}(\mathbf{b} - \mathbf{A}\hat{\mathbf{x}} - \alpha\mathbf{m}_\mathbf{v} - \alpha\mathbf{Gm}_\mathbf{s}) = \mathbf{b} - \hat{\mathbf{v}}$$

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

$$\alpha = 1 \quad \text{or} \quad \alpha = \frac{(\mathbf{Gm}_\mathbf{s} + \mathbf{m}_\mathbf{v})^T\mathbf{M}^{-1}\mathbf{Hb}}{(\mathbf{Gm}_\mathbf{s} + \mathbf{m}_\mathbf{v})^T\mathbf{M}^{-1}\mathbf{H}(\mathbf{Gm}_\mathbf{s} + \mathbf{m}_\mathbf{v})} \qquad \mathbf{H} = \mathbf{I} - \mathbf{AN}^{-1}\mathbf{A}^T\mathbf{M}^{-1}$$

Generalized mixed model $\mathbf{w} = -\mathbf{Ax} - \mathbf{Gs} + \mathbf{Bv}$

$$\mathbf{M} = \mathbf{GC}_\mathbf{s}\mathbf{G}^T + \mathbf{BC}_\mathbf{v}\mathbf{B}^T \qquad \mathbf{N} = \mathbf{A}^T\mathbf{M}^{-1}\mathbf{A}$$

$$\hat{\mathbf{x}} = \mathbf{N}^{-1}\mathbf{A}^T\mathbf{M}^{-1}(\mathbf{w} - \alpha\mathbf{B}\mathbf{m}_\mathbf{v} + \alpha\mathbf{G}\mathbf{m}_\mathbf{s})$$

$$\hat{\mathbf{y}} = (\mathbf{b} - \alpha\mathbf{m}_\mathbf{v}) - \mathbf{C}_\mathbf{v}\mathbf{B}^T\mathbf{M}^{-1}(\mathbf{w} + \mathbf{A}\hat{\mathbf{x}} - \alpha\mathbf{B}\mathbf{m}_\mathbf{v} + \alpha\mathbf{G}\mathbf{m}_\mathbf{s}) = \mathbf{b} - \hat{\mathbf{v}}$$

$$\hat{\mathbf{s}} = \alpha\mathbf{m}_\mathbf{s} - \mathbf{C}_\mathbf{s}\mathbf{G}^T\mathbf{M}^{-1}(\mathbf{w} + \mathbf{A}\hat{\mathbf{x}} - \alpha\mathbf{B}\mathbf{m}_\mathbf{v} + \alpha\mathbf{G}\mathbf{m}_\mathbf{s})$$

$$\alpha = 1 \quad \text{or} \quad \alpha = \frac{(\mathbf{B}\mathbf{m}_\mathbf{v} - \mathbf{G}\mathbf{m}_\mathbf{s})^T\mathbf{M}^{-1}\mathbf{H}\mathbf{w}}{(\mathbf{B}\mathbf{m}_\mathbf{v} - \mathbf{G}\mathbf{m}_\mathbf{s})^T\mathbf{M}^{-1}\mathbf{H}(\mathbf{B}\mathbf{m}_\mathbf{v} - \mathbf{G}\mathbf{m}_\mathbf{s})} \qquad \mathbf{H} = \mathbf{I} - \mathbf{A}\mathbf{N}^{-1}\mathbf{A}^T\mathbf{M}^{-1}$$

Random effects $\mathbf{b} = \mathbf{G}\mathbf{s} + \mathbf{v}$

$$\mathbf{M} = \mathbf{G}\mathbf{C}_\mathbf{s}\mathbf{G}^T + \mathbf{C}_\mathbf{v}$$

$$\hat{\mathbf{y}} = (\mathbf{b} - \alpha\mathbf{m}_\mathbf{v}) - \mathbf{C}_\mathbf{v}\mathbf{M}^{-1}(\mathbf{b} - \alpha\mathbf{m}_\mathbf{v} - \alpha\mathbf{G}\mathbf{m}_\mathbf{s}) = \mathbf{b} - \hat{\mathbf{v}}$$

$$\hat{\mathbf{s}} = \alpha\mathbf{m}_\mathbf{s} + \mathbf{C}_\mathbf{s}\mathbf{G}^T\mathbf{M}^{-1}(\mathbf{b} - \alpha\mathbf{m}_\mathbf{v} - \alpha\mathbf{G}\mathbf{m}_\mathbf{s})$$

$$\alpha = 1 \quad \text{or} \quad \alpha = \frac{(\mathbf{G}\mathbf{m}_\mathbf{s} + \mathbf{m}_\mathbf{v})^T\mathbf{M}^{-1}\mathbf{b}}{(\mathbf{G}\mathbf{m}_\mathbf{s} + \mathbf{m}_\mathbf{v})^T\mathbf{M}^{-1}(\mathbf{G}\mathbf{m}_\mathbf{s} + \mathbf{m}_\mathbf{v})}$$

Generalized random effects $\mathbf{w} = -\mathbf{G}\mathbf{s} + \mathbf{B}\mathbf{v}$

$$\mathbf{M} = \mathbf{G}\mathbf{C}_\mathbf{s}\mathbf{G}^T + \mathbf{B}\mathbf{C}_\mathbf{v}\mathbf{B}^T$$

$$\hat{\mathbf{y}} = (\mathbf{b} - \alpha\mathbf{m}_\mathbf{v}) - \mathbf{C}_\mathbf{v}\mathbf{M}^{-1}(\mathbf{w} - \alpha\mathbf{B}\mathbf{m}_\mathbf{v} + \alpha\mathbf{G}\mathbf{m}_\mathbf{s}) = \mathbf{b} - \hat{\mathbf{v}}$$

$$\hat{\mathbf{s}} = \alpha\mathbf{m}_\mathbf{s} - \mathbf{C}_\mathbf{s}\mathbf{G}^T\mathbf{M}^{-1}(\mathbf{w} - \alpha\mathbf{B}\mathbf{m}_\mathbf{v} + \alpha\mathbf{G}\mathbf{m}_\mathbf{s})$$

$$\alpha = 1 \quad \text{or} \quad \alpha = \frac{(\mathbf{B}\mathbf{m}_\mathbf{v} - \mathbf{G}\mathbf{m}_\mathbf{s})^T\mathbf{M}^{-1}\mathbf{w}}{(\mathbf{B}\mathbf{m}_\mathbf{v} - \mathbf{G}\mathbf{m}_\mathbf{s})^T\mathbf{M}^{-1}(\mathbf{B}\mathbf{m}_\mathbf{v} - \mathbf{G}\mathbf{m}_\mathbf{s})}$$

<u>Models with stochastic parameters - type B</u> $\mathbf{v} \sim (\mathbf{m}_\mathbf{v}, \mathbf{C}_\mathbf{v}) \qquad \mathbf{s} \sim (\mathbf{m}_\mathbf{s}, \mathbf{C}_\mathbf{s})$

Mixed model $\mathbf{b} = \mathbf{A}\mathbf{x} + \mathbf{G}\mathbf{s} + \mathbf{v}$

$$\mathbf{M} = \mathbf{G}\mathbf{C}_\mathbf{s}\mathbf{G}^T + \mathbf{C}_\mathbf{v} \qquad \mathbf{N} = \mathbf{A}^T\mathbf{M}^{-1}\mathbf{A} \qquad \mathbf{R} = \mathbf{Q}_\mathbf{s}\mathbf{C}_\mathbf{s}\mathbf{Q}_\mathbf{s}^T + \mathbf{C}_\mathbf{v}$$

$$\hat{\mathbf{x}} = \mathbf{N}^{-1}\mathbf{A}^T\mathbf{M}^{-1}(\mathbf{w} - \alpha\mathbf{m}_\mathbf{v} - \alpha\mathbf{G}\mathbf{m}_\mathbf{s}) \qquad \hat{\mathbf{y}} = \mathbf{A}\hat{\mathbf{x}}$$

$$\alpha = 1 \quad \text{or} \quad \alpha = \frac{(\mathbf{G}\mathbf{m}_\mathbf{s} + \mathbf{m}_\mathbf{v})^T\mathbf{M}^{-1}\mathbf{H}\mathbf{b}}{(\mathbf{G}\mathbf{m}_\mathbf{s} + \mathbf{m}_\mathbf{v})^T\mathbf{M}^{-1}\mathbf{H}(\mathbf{G}\mathbf{m}_\mathbf{s} + \mathbf{m}_\mathbf{v})} \qquad \mathbf{H} = \mathbf{I} - \mathbf{A}\mathbf{N}^{-1}\mathbf{A}^T\mathbf{M}^{-1}$$

Generalized mixed model $\mathbf{w} = -\mathbf{A}\mathbf{x} + \mathbf{G}\mathbf{s} + \mathbf{B}\mathbf{v}$

$$\mathbf{M} = \mathbf{G}\mathbf{C}_\mathbf{s}\mathbf{G}^T + \mathbf{B}\mathbf{C}_\mathbf{v}\mathbf{B}^T \qquad \mathbf{N} = \mathbf{A}^T\mathbf{M}^{-1}\mathbf{A} \qquad \mathbf{R} = \mathbf{Q}_\mathbf{s}\mathbf{C}_\mathbf{s}\mathbf{Q}_\mathbf{s}^T + \mathbf{C}_\mathbf{v}$$

$$\hat{\mathbf{x}} = -\mathbf{N}^{-1}\mathbf{A}^T\mathbf{M}^{-1}(\mathbf{w} - \alpha\mathbf{B}\mathbf{m}_\mathbf{v} - \alpha\mathbf{G}\mathbf{m}_\mathbf{s})$$

$$\hat{\mathbf{y}} = (\mathbf{b} - \alpha\mathbf{m}_\mathbf{v} - \alpha\mathbf{Q}_\mathbf{s}\mathbf{m}_\mathbf{s}) - \mathbf{R}\mathbf{B}^T\mathbf{M}^{-1}(\mathbf{w} + \mathbf{A}\hat{\mathbf{x}} - \alpha\mathbf{B}\mathbf{m}_\mathbf{v} - \alpha\mathbf{G}\mathbf{m}_\mathbf{s})$$

$$\alpha = 1 \quad \text{or} \quad \alpha = \frac{(\mathbf{B}\mathbf{m}_\mathbf{v} + \mathbf{G}\mathbf{m}_\mathbf{s})^T\mathbf{M}^{-1}\mathbf{H}\mathbf{w}}{(\mathbf{B}\mathbf{m}_\mathbf{v} + \mathbf{G}\mathbf{m}_\mathbf{s})^T\mathbf{M}^{-1}\mathbf{H}(\mathbf{B}\mathbf{m}_\mathbf{v} + \mathbf{G}\mathbf{m}_\mathbf{s})} \qquad \mathbf{H} = \mathbf{I} - \mathbf{A}\mathbf{N}^{-1}\mathbf{A}^T\mathbf{M}^{-1}$$

For the sake of completeness also the other types of estimates, which depend on the unknown values of the parameters and cannot be directly used, will be summarized next. For this purpose specific models must be separated into two classes, class A which contains those with homogeneous model equations ($\mathbf{t}=\mathbf{0}$) and class B with the inhomogeneous ones ($\mathbf{t}\neq\mathbf{0}$).

Class A: Observation equations, observation equations with homogeneous constraints ($\mathbf{Hx}=\mathbf{0}$), mixed model.

Class B: Condition equations, mixed equations, observation equations with inhomogeneous constraints ($\mathbf{Hx}=-\mathbf{h}^0\neq\mathbf{0}$), mixed equations with constraints.

All estimates have the common form: $\mathbf{x}=\alpha\hat{\mathbf{x}}$, $\mathbf{y}=\alpha\hat{\mathbf{y}}$
where the parameter α takes the following values:

inhomBLE, inhomBLUE, homBLE (class B), homBLUE (class B): $\alpha=1$

homBLE (class A): $\alpha=\dfrac{\beta}{1+\gamma}$ *homBLUE (class A):* $\alpha=\dfrac{\beta}{\gamma}$

where the parameters β and γ are as follows:

Simple models

$$\beta=(\mathbf{y}+\mathbf{m_v})^T\mathbf{C_v}^{-1}\mathbf{b}$$

$$\gamma=(\mathbf{y}+\mathbf{m_v})^T\mathbf{C_v}^{-1}(\mathbf{y}+\mathbf{m_v})$$

Model with prior information

$$\beta=(\mathbf{y}+\mathbf{m_v})^T\mathbf{C_v}^{-1}\mathbf{b}+(\mathbf{x}+\mathbf{m_x})^T\mathbf{C_v}^{-1}\mathbf{x}^b$$

$$\gamma=(\mathbf{y}+\mathbf{m_v})^T\mathbf{C_v}^{-1}(\mathbf{y}+\mathbf{m_v})+(\mathbf{x}+\mathbf{m_x})^T\mathbf{C_v}^{-1}(\mathbf{x}+\mathbf{m_x})$$

Model with partial prior information

$$\beta=(\mathbf{y}+\mathbf{m_v})^T\mathbf{C_v}^{-1}\mathbf{b}+(\mathbf{x}_2+\mathbf{m}_2)^T\mathbf{C}_2^{-1}\mathbf{x}_2^b$$

$$\gamma=(\mathbf{y}+\mathbf{m_v})^T\mathbf{C_v}^{-1}(\mathbf{y}+\mathbf{m_v})+(\mathbf{x}_2+\mathbf{m}_2)^T\mathbf{C}_2^{-1}(\mathbf{x}_2+\mathbf{m}_2)$$

Models with stochastic parameter - type A:

Mixed model:

$$\beta=(\mathbf{y}+\mathbf{m_v})^T\mathbf{C_v}^{-1}\mathbf{b}$$

$$\gamma=(\mathbf{y}+\mathbf{m_v})^T\mathbf{C_v}^{-1}(\mathbf{y}+\mathbf{m_v})+(\mathbf{y}-\mathbf{Ax}-\mathbf{Gm_s})^T(\mathbf{GC_sG})^{-1}(\mathbf{y}-\mathbf{Ax}-\mathbf{Gm_s})$$

Generalized mixed model:

$$\beta=(\mathbf{y}+\mathbf{m_v})^T\mathbf{C_v}^{-1}\mathbf{b}+(\mathbf{Ax}+\mathbf{By}+\mathbf{Gm_s})^T(\mathbf{GC_sG})^{-1}(\mathbf{Bb}-\mathbf{w})$$

$$\gamma=(\mathbf{y}+\mathbf{m_v})^T\mathbf{C_v}^{-1}(\mathbf{y}+\mathbf{m_v})+(\mathbf{Ax}+\mathbf{By}+\mathbf{Gm_s})^T(\mathbf{GC_sG})^{-1}(\mathbf{Ax}+\mathbf{By}+\mathbf{Gm_s})$$

Random effects model:

$$\beta=(\mathbf{y}+\mathbf{m_v})^T\mathbf{C_v}^{-1}\mathbf{b}$$

$$\gamma=(\mathbf{y}+\mathbf{m_v})^T\mathbf{C_v}^{-1}(\mathbf{y}+\mathbf{m_v})+(\mathbf{y}-\mathbf{Gm_s})^T(\mathbf{GC_sG})^{-1}(\mathbf{y}-\mathbf{Gm_s})$$

Generalized random effects model:

$$\beta = (\mathbf{y}+\mathbf{m_v})^T \mathbf{C_v^{-1}b} + (\mathbf{By}+\mathbf{Gm_s})^T (\mathbf{GC_sG})^{-1} (\mathbf{Bb}-\mathbf{w})$$

$$\gamma = (\mathbf{y}+\mathbf{m_v})^T \mathbf{C_v^{-1}} (\mathbf{y}+\mathbf{m_v}) + (\mathbf{By}+\mathbf{Gm_s})^T (\mathbf{GC_sG})^{-1} (\mathbf{By}+\mathbf{Gm_s})$$

Models with stochastic parameters - type B:

Mixed model:

$$\beta = (\mathbf{y}+\mathbf{Gm_s}+\mathbf{m_v})^T \mathbf{M}^{-1}\mathbf{b}$$

$$\gamma = (\mathbf{y}+\mathbf{Gm_s}+\mathbf{m_v})^T \mathbf{M}^{-1}(\mathbf{y}+\mathbf{Gm_s}+\mathbf{m_v})$$

Generalized mixed model:

$$\beta = (\mathbf{y}+\mathbf{Q_s m_s}+\mathbf{m_v})^T (\mathbf{Q_s C_s Q_s}+\mathbf{C_v})^{-1}\mathbf{b}$$

$$\gamma = (\mathbf{y}+\mathbf{Q_s m_s}+\mathbf{m_v})^T (\mathbf{Q_s C_s Q_s}+\mathbf{C_v})^{-1}(\mathbf{y}+\mathbf{Q_s m_s}+\mathbf{m_v})$$

Appendix 2A: Derivation of the general estimation solution

The problem is to estimate a single parameter $q = \mathbf{c}^T \mathbf{z}$, function of the parameters $\mathbf{z}$, from available data

$$\mathbf{d} = \mathbf{Uz} + \mathbf{e} \tag{A1}$$

where

$$E\{\mathbf{e}\} = \mathbf{m}, \qquad E\{(\mathbf{e}-\mathbf{m})(\mathbf{e}-\mathbf{m})^T\} = \mathbf{C} \tag{A2}$$

Considering only linear estimation, an estimate $\hat{q}$ of q has the form

$$\hat{q} = \mathbf{p}^T\mathbf{d} + \kappa = \mathbf{p}^T\mathbf{Uz} + \mathbf{p}^T\mathbf{e} + \kappa \tag{A3}$$

for linear inhomogeneous prediction, the homogeneous one results with $\kappa = 0$. Since

$$E\{\hat{q}\} = \mathbf{p}^T\mathbf{m_d} + \kappa \qquad \mathbf{m_d} \equiv E\{\mathbf{d}\} = \mathbf{Uz} + \mathbf{m} \tag{A4}$$

the estimation bias becomes

$$\beta = E\{\hat{q}\} - q = \mathbf{p}^T\mathbf{m_d} + \kappa - q = \mathbf{p}^T(\mathbf{Uz}+\mathbf{m}) + \kappa - \mathbf{c}^T\mathbf{z} = (\mathbf{U}^T\mathbf{p}-\mathbf{c})^T\mathbf{z} + \mathbf{p}^T\mathbf{m} + \kappa \tag{A5}$$

The estimation error takes the form

$$\hat{q} - q = \mathbf{p}^T\mathbf{d} + \kappa - q = \mathbf{p}^T[\mathbf{m_d} + (\mathbf{e}-\mathbf{m})] + \kappa - q = \beta + \mathbf{p}^T(\mathbf{e}-\mathbf{m}) \tag{A6}$$

and the mean square error (risk function) becomes

$$R = E\{(\hat{q}-q)^2\} = \beta^2 + \mathbf{p}^T\mathbf{Cp} \tag{A7}$$

The following derivatives are needed for later use

$$\frac{\partial \beta}{\partial \mathbf{p}} = \mathbf{m}_\mathbf{d}^T \qquad\qquad \frac{\partial \beta}{\partial \kappa} = 1 \tag{A8}$$

$$\frac{\partial R}{\partial \mathbf{p}} = 2\beta \frac{\partial \beta}{\partial \mathbf{p}} + 2\mathbf{p}^T \mathbf{C} = 2(\mathbf{Cp} + \beta \mathbf{m}_\mathbf{d})^T \tag{A9}$$

$$\frac{\partial R}{\partial \kappa} = 2\beta \frac{\partial \beta}{\partial \kappa} = 2\beta \tag{A10}$$

inhomBLE

For the determination of the best linear inhomogeneous estimation the values of $\mathbf{p}$ and κ which minimize the risk function (19) must be found. The solution to this minimization problem is determined by setting the derivatives of R with respect to $\mathbf{p}$ and κ equal to zero

$$\frac{\partial R}{\partial \mathbf{p}} = 2(\mathbf{Cp} + \beta \mathbf{m}_\mathbf{d})^T = \mathbf{0} \tag{A11}$$

$$\frac{\partial R}{\partial \kappa} = 2\beta = 2(\mathbf{p}^T \mathbf{m}_\mathbf{d} + \kappa - q) = 0 \tag{A12}$$

so that

$$\mathbf{Cp} = \mathbf{0} \qquad\qquad \kappa = q - \mathbf{p}^T \mathbf{m}_\mathbf{d} \qquad\qquad \hat{q} = q + \mathbf{p}^T (\mathbf{d} - \mathbf{m}_\mathbf{d}) \tag{A13}$$

homBLE

In the case of homogeneous estimation $\kappa = 0$, so that according to (A5)

$$\beta = \mathbf{p}^T \mathbf{m}_\mathbf{d} - q \tag{A14}$$

which combined with (A11) gives

$$(\mathbf{C} + \mathbf{m}_\mathbf{d} \mathbf{m}_\mathbf{d}^T)\mathbf{p} = q\mathbf{m}_\mathbf{d} \qquad\qquad \hat{q} = \mathbf{p}^T \mathbf{d} \tag{A15}$$

inhomBLUE

For the determination of the best linear inhomogeneous unbiased estimation the values of $\mathbf{p}$ and κ which minimize the risk function must be found and at the same time satisfy the condition $\beta = 0$. The solution to this minimization problem is determined by setting the derivatives of the Lagrangean function $\Phi = R - 2\nu\beta$, (ν being a Lagrange multiplier,) with respect to $\mathbf{p}$, κ and ν equal to zero

$$\frac{\partial \Phi}{\partial \mathbf{p}} = \frac{\partial R}{\partial \mathbf{p}} - 2\nu \frac{\partial \beta}{\partial \mathbf{p}} = 2[\mathbf{Cp} + (\beta - \nu)\mathbf{m}_\mathbf{d}]^T = \mathbf{0} \tag{A16}$$

$$\frac{\partial \Phi}{\partial \kappa} = \frac{\partial R}{\partial \kappa} - 2\nu \frac{\partial \beta}{\partial \kappa} = 2(\beta - \nu) = 0 \tag{A17}$$

$$\frac{\partial \Phi}{\partial \nu} = -2\beta = -2(\mathbf{p}^T \mathbf{m}_\mathbf{d} + \kappa - q) = 0 \tag{A18}$$

It follows that $\nu = \beta = 0$, so that (A16) gives $\mathbf{Cp} = \mathbf{0}$, and with κ from (A18) the solution becomes

$$\mathbf{Cp} = \mathbf{0} \qquad \hat{q} = q + \mathbf{p}^T (\mathbf{d} - \mathbf{m_d}) \tag{A19}$$

which is identical with the solution in the inhomBLE case.

<u>homBLUE</u>

In the homogeneous case the Lagrangean function Φ has the same form as in the inhomogeneous case, while β is given by equation (A14) since $\kappa = 0$. The solution is obtained by setting the derivatives of Φ with respect to $\mathbf{p}$ and ν equal to zero

$$\frac{\partial \Phi}{\partial \mathbf{p}} = 2[\mathbf{Cp} + (\beta - \nu)\mathbf{m_d}]^T = \mathbf{0} \tag{A20}$$

$$\frac{\partial \Phi}{\partial \nu} = -2\beta = -2(\mathbf{p}^T \mathbf{m_d} - q) = 0 \tag{A21}$$

which give

$$\mathbf{Cp} - \nu \mathbf{m_d} = \mathbf{0} \qquad \mathbf{m_d}^T \mathbf{p} = q \qquad \hat{q} = \mathbf{p}^T \mathbf{d} \tag{A22}$$

<u>inhomBLUUE</u>

For the bias to vanish uniformly, i.e., for any $\mathbf{z}$, the coefficient of $\mathbf{z}$ and the constant term in (A5)

$$\beta = (\mathbf{U}^T \mathbf{p} - \mathbf{c})^T \mathbf{z} + \mathbf{p}^T \mathbf{m} + \kappa \tag{A23}$$

must be zero, which leads to the conditions

$$\mathbf{U}^T \mathbf{p} - \mathbf{c} = \mathbf{0} \tag{A24}$$

$$\mathbf{p}^T \mathbf{m} + \kappa = 0 \tag{A25}$$

The Lagrangean function becomes

$$\Phi = R - 2\mathbf{k}^T (\mathbf{U}^T \mathbf{p} - \mathbf{c}) - 2\nu(\mathbf{m}^T \mathbf{p} + \kappa) \tag{A26}$$

where $\mathbf{k}$ and ν are Lagrange multipliers. The derivatives of Φ with respect to $\mathbf{p}$, κ, $\mathbf{k}$ and ν must vanish. The last two give (A24) and (A25), which must be combined with

$$\frac{\partial \Phi}{\partial \mathbf{p}} = \frac{\partial R}{\partial \mathbf{p}} - 2\mathbf{k}^T \mathbf{U}^T - 2\nu \mathbf{m}^T =$$

$$= 2[\mathbf{Cp} + \beta \mathbf{m_d} - \mathbf{Uk} - \nu \mathbf{m}]^T = 2[\mathbf{Cp} - \mathbf{Uk} - \nu \mathbf{m}]^T = \mathbf{0} \tag{A27}$$

$$\frac{\partial \Phi}{\partial \kappa} = 2\beta \frac{\partial \beta}{\partial \kappa} - 2\nu = -2\nu = 0 \tag{A28}$$

where $\beta = 0$ has been taken into account. With $\nu = 0$, and κ from (A25), the solution is determined from

$$\mathbf{Cp} - \mathbf{Uk} = \mathbf{0} \qquad \mathbf{U}^T \mathbf{p} = \mathbf{c} \qquad \hat{q} = \mathbf{p}^T (\mathbf{d} - \mathbf{m}) \tag{A29}$$

<u>homBLUUE</u>

In the homogeneous case the bias β is given by equation (A14) since $\kappa = 0$. The conditions for uniformly unbiased estimates become

$$\mathbf{U}^T \mathbf{p} - \mathbf{c} = \mathbf{0} \qquad \mathbf{m}^T \mathbf{p} = 0 \tag{A30}$$

the Lagrangean function becomes

$$\Phi = R - 2\mathbf{k}^T (\mathbf{U}^T \mathbf{p} - \mathbf{c}) - 2\nu(\mathbf{m}^T \mathbf{p}) \tag{A31}$$

The derivatives of Φ with respect to $\mathbf{p}$ must vanish, which gives again (A27), since the last does not depend on κ. Combination of (A27) with (A30) gives the solution

$$\mathbf{Cp} - \mathbf{Uk} - \nu\mathbf{m} = \mathbf{0} \qquad \mathbf{U}^T \mathbf{p} = \mathbf{c} \qquad \mathbf{m}^T \mathbf{p} = 0 \tag{A32}$$

$$\hat{q} = \mathbf{p}^T (\mathbf{d} - \mathbf{m}) \tag{A33}$$