Chapter 2
The First Problem of Probabilistic Regression: The Bias Problem

Minimum Bias solution of problems with datum defects. LUMBE of fixed effects.

The *bias problem in probabilistic* regression has been the subject of Sect. 4-37 for simultaneous determination of first moments as well as second central moments by inhomogeneous multilinear, namely bilinear, estimation. Based on the review of the first author "Variance-covariance component estimation: theoretical results and geodetic application" (Statistical and Decision, Supplement Issue No. 2, pages 401–441, 105 references, Oldenbourg Verlag, München 1989), we collected 5 postulates for simultaneous determination of first and second central moments. A first reference is *J. Kleffe* (1978). It forms the basis of Sect. 4-37:

ansatz

$$E\{y\} = \boldsymbol{A}\mu = \boldsymbol{A}\boldsymbol{x}_1, \quad vech\ D\{y\} = \boldsymbol{B}\sigma = \boldsymbol{B}\boldsymbol{x}_2$$

estimation:
ansatz: bilinearity

$$\begin{bmatrix} \widehat{\boldsymbol{x}_1} = \boldsymbol{K}_1 + \boldsymbol{L}_1\boldsymbol{y} + \boldsymbol{M}_1(\boldsymbol{y} \otimes \boldsymbol{y}) \\ \widehat{\boldsymbol{x}_2} = \boldsymbol{K}_2 + \boldsymbol{L}_2\boldsymbol{y} + \boldsymbol{M}_2(\boldsymbol{y} \otimes \boldsymbol{y}) \end{bmatrix} \widehat{\boldsymbol{x}} :\leq \begin{bmatrix} \widehat{\boldsymbol{x}_1} \\ \widehat{\boldsymbol{x}_2} \end{bmatrix}$$

1st postulate

"inhomogeneous, bilinear estimation"

$$\boldsymbol{X} := \begin{bmatrix} \boldsymbol{K}_1 & \boldsymbol{L}_1 & \boldsymbol{M}_1 \\ \boldsymbol{K}_2 & \boldsymbol{L}_2 & \boldsymbol{M}_2 \end{bmatrix}, \quad \boldsymbol{Y} := \begin{bmatrix} \mathbf{1} \\ \mathbf{y} \\ \mathbf{y} \otimes \mathbf{y} \end{bmatrix}$$

2nd postulate

"invariance"
$\tilde{\sigma}$ is invariant in the sense $y \mapsto y - \boldsymbol{A}\mu$

3rd postulate

"unbiased in the mean" versus "minimum bias"
bias vector, bias matrix

E. Grafarend and J. Awange, *Applications of Linear and Nonlinear Models*,
Springer Geophysics, DOI 10.1007/978-3-642-22241-2_2,

$$\boldsymbol{b} := E\{\hat{x}\} - x = \boldsymbol{S}x, \boldsymbol{S} := \boldsymbol{X}\begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{B} \end{bmatrix} - \boldsymbol{I}$$

4th postulate

"$D\{\hat{x}\}$ contains second, third and fourth order moments which are based on a *quasi-Gauss normal* structure."

5th postulate

"best estimation"

$$\|D\{\hat{x}\}\| = min$$

Here we need only two postulates, namely *(i) linear*: $(x_1) = \boldsymbol{L_1 y}$ and *(ii) minimum bias in the mean* when we consider the simple model $E\{\boldsymbol{y}\} = \boldsymbol{A\mu} = \boldsymbol{A}x_1(\mathbf{A} \in \mathcal{R}^{\mathrm{n \times m}},\ \mathrm{n} < m)$ for the moment of first order and $D\{\boldsymbol{y}\} = I\ \sigma^2 = I\ \boldsymbol{x}_2$ for the central moment of the second order ($\sigma^2 > 0$), one parameter of type *variance*. You find the estimate of type *linear* $\widehat{\boldsymbol{x_1}} = \boldsymbol{L_1 y}$ and type *minimum bias* in the mean and of type *FROBENIUS matrix norm* $\|\boldsymbol{B}\|^2$ for the case $m > n$, here $\|\boldsymbol{B}\|^2 = d = m - n$.

Real World Problems are nonlinear. We have to divide the solution space into those classes: (a) strong nonlinear, (b) weak nonlinear, and (c) linear. Weak nonlinearity is defined by nonlinear system which allows a *Taylor series approximation*. Linear system equations, for instance the minimum bias problem in the *FROBENIUS matrix norm* with respect to *linear estimation theory*, have the advantage that its *norm of equations* has only one minimum solution. We call that solution "*global*" and "*uniform*". In contrast to linear estimation, weak nonlinear equations produce "*local*" solutions. They experience a great disadvantage: there exist *many local solutions*. A typical example is a geodetic network analysis by *P. LOHSE (1994)*. Another example is also *C.R. RAO and J. KLEFFE (1999, pages 161–180)*.

Minimum bias solutions of rank deficient linear systems have been discussed in detail by *C.R. RAO (1974)*. He introduces the notation of the *substitute matrix* $\boldsymbol{S}$ referring the matrix $\boldsymbol{\xi\xi'}$ by the matrix $\mathbf{S}$ of arbitrary rank, for instance $rk\boldsymbol{S} = m$. Note that the rank of the matrix $\boldsymbol{\xi\xi'}$ is one by the technique of *LAGRANGE multiplyer* the minimization of the S weighted *FROBENIUS norm* $\|(\boldsymbol{I}_m - LA)'\|_L^2$ within *Theorem 2.3*. Obviously, $\mathbf{S}$-weighted homogeneous linear uniform by minimum bias estimator ("hom S-LUMBE") is equivalent to the weighted minimum norm solutions ("**G**-MINOS") subject of the RAO-Theorem. Our example is the special model for $\boldsymbol{S} = \boldsymbol{I}_m$ called "*unity substitute matrix*", obviously $\|\boldsymbol{B}\|^2 = d = m - n = m - rk\boldsymbol{A}$.

In practice, it is difficult to determine the rank of matrix A. For instance, in problems of Configuration Defects in Geodetic Sciences E. Grafarend and E. Livieratos (1979), E. Grafarend and K. Heinz (1978), E. Grafarend and V.

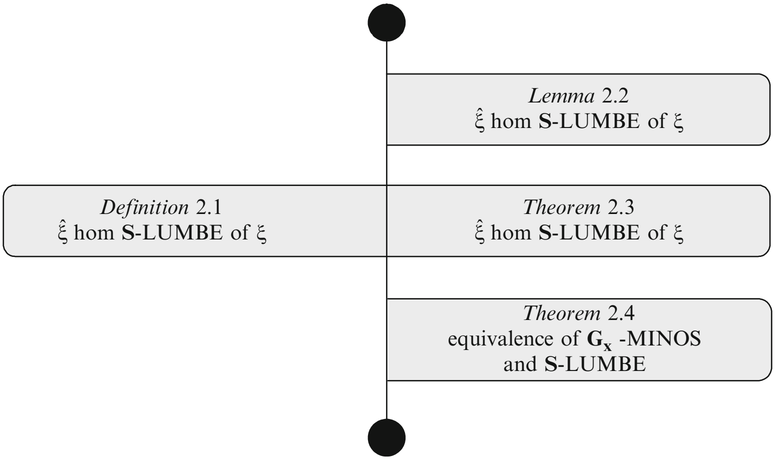

Fig. 2.1 The guideline of chapter two: definition, lemma and theorems

Miller (1985) and E. Grafarend and A. Mader (1989). Critical comments of the RAO bias problem solution have been raised by Arne Bjerhammar and Jujit Kumar Mitra (private communication to the first author). They argue that bias has nothing to do with rank definition of the type m-n. They use the argument that you need PRIOR INFORMATION, for instance of BAYES type: Bayes Estimation x_0. Here we leave this question open.

In summary, the minimum of the trace of the matrix valued bias leads to the rank deficiency d subject to $\mathrm{d} = \mathrm{m} - \mathrm{n}, \mathrm{m} > \mathrm{n}$.

Please pay attention to the guideline of Chap. 2 shown in Fig. 2.1. In particular, mind the structure of definitions, lemma, and theorems. Fast track reading: consult Box 2.1 and read only Theorem 2.3.

In the first chapter, we have solved a special algebraic regression problem, namely the inversion of a system of consistent linear equations which is classified as underdetermined. By means of the postulate of a minimum norm solution, namely $||\mathbf{x}||^2 = \min$, we were able to determine m unknowns ($m > n$, say $m = 10^6$) from n observations (more unknowns m than equations n, say $n = 10$). Indeed, such a mathematical solution may surprise the analyst: in the example, MINOS produces one million unknowns from 10 observations. Though MINOS generates a rigorous solution, we are left with some doubts.

How can we conveniently interpret such an "unbelievable solution"? The key for an evaluation of MINOS is handed over to us by treating the special algebraic regression problem with datum defect. The bias generated by any solution of an underdetermined or ill-posed problem will be introduced as a decisive criterion for evaluating MINOS, now in the context of a probabilistic regression problem. In particular, a special form of LUMBE, the *linear uniformly minimum bias estimator* $||\mathbf{LA} - \mathbf{I}||^2 =$ min leads us to a solution which is equivalent to "MINOS". Alternatively we may say that in the various classes of solving an underdetermined problem "LUMBE" generates a solution of minimal bias.

What is a *underdetermined regression problem* of type MINOS? By means of a certain objective function, here of type "*minimum bias*", we solve the inverse problem of linear and nonlinear equations with fixed effects. In particular, in order to minimize the bias vector $\boldsymbol{E}\{\hat{\boldsymbol{\xi}}\} - \boldsymbol{\xi}$ we meet the problem that its minimum norm $\|\boldsymbol{E}\{\hat{\boldsymbol{\xi}}\} - \boldsymbol{\xi}\|^2 = tr[(\boldsymbol{E}\{\hat{\boldsymbol{\xi}}\} - \boldsymbol{\xi})(\boldsymbol{E}\{\hat{\boldsymbol{\xi}}\} - \boldsymbol{\xi})']$ subject to $\hat{\boldsymbol{\xi}} = \mathbf{Ly}$ *depends on the product* $\boldsymbol{\xi}\boldsymbol{\xi}'$ where $\boldsymbol{\xi}$ denotes the *unknown parameter vector*. In this section *C.R. Rao* proposed the use instead of the matrix $\boldsymbol{\xi}\boldsymbol{\xi}'$, $rk\boldsymbol{\xi}\boldsymbol{\xi}' = 1$, the *substitute matrix* $\mathbf{S}$ with full rank, for instance.

2-1 Linear Uniformly Minimum Bias Estimator (LUMBE)

Let us introduce the special consistent linear Gauss–Markov model specified in Box 2.1, which is given form of a *consistent system* of linear equations relating non-stochastic, real-valued vector $\boldsymbol{\xi}$ of unknowns to the observation vector $\mathbf{y}$. Here, the rank rk$\mathbf{A}$ of the matrix $\mathbf{A}$ connecting the observations and unknown parameters equals the number n of observations $\mathbf{y} \in \mathbb{R}^n$.

Box 2.1. (Special consistent linear Gauss–Markov model of fixed effects. Bias vector, bias matrix, vector and matrix bias norms).

$$\{\mathbf{y} = \mathbf{A}\boldsymbol{\xi} | \mathbf{A} \in R^{n \times m}, \mathrm{rk}\mathbf{A} = n, n < m\}$$

ξ *unknown*

"*ansatz*"

$$\hat{\boldsymbol{\xi}} = \mathbf{Ly} \tag{2.1}$$

"*bias vector*"

$$\boldsymbol{\beta} := \boldsymbol{E}\{\hat{\boldsymbol{\xi}}\} - \boldsymbol{\xi} = -[\mathbf{I_m} - \mathbf{LA}]\,\boldsymbol{\xi}, \ \forall \boldsymbol{\xi} \in \mathbb{R}^m \tag{2.2}$$

"*bias matrix*"

$$\mathbf{B}' = \mathbf{I_m} - \mathbf{LA} \tag{2.3}$$

"*bias norms*″

$$||\boldsymbol{\beta}||^2 = \boldsymbol{\beta}'\boldsymbol{\beta} = \boldsymbol{\xi}'[\mathbf{I_m} - \mathbf{LA}]'[\mathbf{I_m} - \mathbf{LA}]\boldsymbol{\xi} \tag{2.4}$$

$$||\boldsymbol{\beta}||^2 = \mathrm{tr}\boldsymbol{\beta}\boldsymbol{\beta}' = \mathrm{tr}[\mathbf{I_m} - \mathbf{LA}]\boldsymbol{\xi}\boldsymbol{\xi}'[\mathbf{I_m} - \mathbf{LA}]' = \ ||\mathbf{B}||^2_{\xi\xi'} \tag{2.5}$$

$$||\boldsymbol{\beta}||^2_{\mathbf{S}} : = \mathrm{tr}[\mathbf{I_m} - \mathbf{LA}]\mathbf{S}[\mathbf{I_m} - \mathbf{LA}]' =: \ ||\mathbf{B}||^2_{\mathbf{S}} \tag{2.6}$$

Since we deal with a linear model, it is “a natural choice” to setup a homogeneous linear form to estimate the parameters $\boldsymbol{\xi}$ of fixed effects, at first, namely (2.1) where is a matrix-valued fixed unknown. In order to determine the real-valued $m \times n$ matrix $\mathbf{L}$, the homogeneous linear estimation $\hat{\boldsymbol{\xi}}$ of the vector $\boldsymbol{\xi}$ of fixed effects has to fulfil a certain optimality condition, which we will outline. Let us try to analyze the bias in solving an underdetermined system of linear equations. Take reference to Box 2.1, where we systematically introduce (i) the bias vector $\boldsymbol{\beta}$, (ii) the bias matrix, (iii) the $\mathbf{S}$-modified bias matrix norm as a weighted Frobenius norm. In detail, let us discuss the bias terminology: For a homogeneous linear estimation (2.1), the vector-valued bias $\boldsymbol{\beta} := [\boldsymbol{E}\{\hat{\boldsymbol{\xi}}\} - \boldsymbol{\xi}]$ takes over the special form (2.2), which led us to the definition of the bias matrix $(\mathbf{I} - \mathbf{LA})'$. The norm of the bias vector $\boldsymbol{\beta}$, namely $||\boldsymbol{\beta}||^2 := \boldsymbol{\beta}'\boldsymbol{\beta}$, coincides with the $\boldsymbol{\xi}\boldsymbol{\xi}'$ weighted Frobenius norm of the bias matrix $\mathbf{B}$, namely $||\mathbf{B}||^2_{\xi\xi'}$. Here, we meet the central problem that the weight matrix $\boldsymbol{\xi}\boldsymbol{\xi}'$, $\mathrm{rk}\boldsymbol{\xi}\boldsymbol{\xi}' = 1$, has rank one. In addition, $\boldsymbol{\xi}\boldsymbol{\xi}'$ is not accessible since $\boldsymbol{\xi}$ is unknown. In this problematic case we replace the matrix $\boldsymbol{\xi}\boldsymbol{\xi}'$ by a fixed positive-definite $m \times m$ matrix $\mathbf{S}$, $\mathrm{rk}\mathbf{S} = m$, C.R. Rao’s substitute matrix and define the $\mathbf{S}$-weighted Frobenius matrix norm (2.6) Indeed, the substitute matrix $\mathbf{S}$ constitutes the matrix of the metric of the bias space.

Being prepared for optimality criteria we give a precise definition of $\hat{\boldsymbol{\xi}}$ of type hom $\mathbf{S}$-LUMBE in Definition 2.1. The estimations $\hat{\boldsymbol{\xi}}$ of type hom $\mathbf{S}$-LUMBE can be characterized by Lemma 2.2.

Definition 2.1. ($\hat{\xi}$ hom **S**-LUMBE of ξ)

An $m \times 1$ vector $\hat{\boldsymbol{\xi}}$ is called hom S-LUMBE (homogeneous Linear Uniformly Minimum Bias Estimation) of $\boldsymbol{\xi}$ in the special consistent linear Gauss–Markov model of fixed effects of Box 2.1, if (1st) $\hat{\boldsymbol{\xi}}$ is a homogeneous linear form (2.7) (2nd) – in comparison to all other linear estimation $\hat{\boldsymbol{\xi}}$ has the minimum bias in the sense of (2.8).

$$\hat{\boldsymbol{\xi}} = \mathbf{Ly}, \tag{2.7}$$

$$||\mathbf{B}||^2_{\mathbf{S}} := \ ||(\mathbf{I_m} - \mathbf{LA})'||^2_{\mathbf{S}}. \tag{2.8}$$

Lemma 2.2. ($\hat{\xi}$ hom **S**-LUMBE of ξ).

An $m \times 1$ vector $\hat{\boldsymbol{\xi}}$ is hom S-LUMBE of $\boldsymbol{\xi}$ in the special consistent linear Gauss–Markov model with fixed effects of Box 2.1, if and only if the matrix $\hat{\boldsymbol{L}}$ fulfils the normal equations

$$\mathbf{ASA}'\hat{\mathbf{L}}' = \mathbf{AS}. \tag{2.9}$$

Proof.

The **S**-weighted Frobenius norm $||\,(\mathbf{I_m} - \mathbf{LA})'\,||^2_{\mathbf{S}}$ establishes the Lagrangean (2.10) for **S**-LUMBE. The necessary conditions for the minimum of the quadratic Lagrangean $\mathcal{L}(\mathbf{L})$ are provided by (2.11).

$$\mathcal{L}(\mathbf{L}) := \mathrm{tr}\,(\mathbf{I_m} - \mathbf{LA})\,\mathbf{S}\,(\mathbf{I_m} - \mathbf{LA})' = \min \mathbf{L}, \tag{2.10}$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{L}}\left(\hat{L}\right) := 2\left[\mathbf{ASA}'\hat{\mathbf{L}}' - \mathbf{AS}\right]' = 0. \tag{2.11}$$

The second derivatives (2.12) at the "point" $\hat{\mathbf{L}}$ constitute the sufficiency conditions. In order to compute such a $mn \times mn$ matrix of second derivatives we have to vectorize the matrix normal equation by (2.13)

$$\frac{\partial^2 \mathcal{L}}{\partial(\mathrm{vec}\mathbf{L})\partial(\mathrm{vec}\mathbf{L})'}\left(\hat{\mathbf{L}}\right) > 0, \tag{2.12}$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{L}}\left(\hat{\mathbf{L}}\right) = 2\left[\hat{\mathbf{L}}\mathbf{ASA}' - \mathbf{SA}'\right], \tag{2.13}$$

$$\frac{\partial \mathcal{L}}{\partial(\mathrm{vec}\mathbf{L})}\left(\hat{\mathbf{L}}\right) = \mathrm{vec}2\left[\hat{\mathbf{L}}\mathbf{ASA}' - \mathbf{SA}'\right], \tag{2.14}$$

$$\frac{\partial \mathcal{L}}{\partial(\mathrm{vec}\mathbf{L})}\left(\hat{\mathbf{L}}\right) = 2\left[\mathbf{ASA}' \otimes \mathbf{I_m}\right]\mathrm{vec}\hat{\mathbf{L}} - 2\mathrm{vec}\left(\mathbf{SA}'\right). \tag{2.15}$$

The Kronecker–Zehfuss product $\mathbf{A} \otimes \mathbf{B}$ of two arbitrary matrices as well as Kronecker-Zehfuss product $(\mathbf{A} + \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes \mathbf{C} + \mathbf{B} \otimes \mathbf{C}$ of three arbitrary matrices subject to the dimension condition $\dim \mathbf{A} = \dim \mathbf{B}$ is introduced in Appendix A. The vec operation (vectorization of an array) is reviewed in Appendix A as well as $\mathrm{vec}\mathbf{AB} = (\mathbf{B}' \otimes \mathbf{I}'_\ell)\mathrm{vec}\mathbf{A}$ for suitable matrices **A** and **B**. Now, we are prepared to compute (2.16) as a positive-definite matrix.

$$\frac{\partial^2 \mathcal{L}}{\partial(\mathrm{vec}\mathbf{L})\partial(\mathrm{vec}\mathbf{L})'}\left(\mathbf{L}'\right) = 2(\mathbf{ASA}') \otimes \mathbf{I_m} > 0 \tag{2.16}$$

For an explicit representation of $\hat{\boldsymbol{\xi}}$ of type hom LUMBE in the special *consistent* linear Gauss–Markov model of fixed effects of Box 2.1, we solve the normal equations for (2.17). Beside the explicit representation of $\hat{\boldsymbol{\xi}}$ of type hom LUMBE in (2.15) we attempt to calculate the bias by (2.19). Of course, the bias computation depends on C.R. Rao's substitute matrix **S**, $\mathrm{rk}\mathbf{S} = m$. Indeed we can associate any element of the solution vector, the dispersion matrix as well as the bias vector with a particular weight which can be "designed" by the analyst.

$$\hat{\mathbf{L}} = \arg\{\mathcal{L}(\mathbf{L}) = \min_{\mathbf{L}}\} \tag{2.17}$$

Theorem 2.3. ($\hat{\boldsymbol{\xi}}$ hom LUMBE of $\boldsymbol{\xi}$)

Let $\hat{\boldsymbol{\xi}} = \mathbf{Ly}$ be hom LUMBE in the special consistent linear Gauss–Markov model of fixed effects of Box 2.1. Then the solution of the normal equation is completed by bias vector

$$\hat{\boldsymbol{\xi}} = \mathbf{SA}'(\mathbf{ASA}')^{-1}\mathbf{y} \tag{2.18}$$

$$\boldsymbol{\beta} := \boldsymbol{E}\{\hat{\boldsymbol{\xi}}\} - \boldsymbol{\xi} = -\left[\mathbf{I_m} - \mathbf{SA}'(\mathbf{ASA}')^{-1}\mathbf{A}\right]\boldsymbol{\xi},\ \ \forall\boldsymbol{\xi} \in \mathbb{R}^m. \tag{2.19}$$

Indeed, the bias computation is problematic since we have no *prior information* of the parameter vector $\boldsymbol{\xi}$. We could indeed use an a posteriori *i*nformation, for instance $\boldsymbol{\xi}_0 = \hat{\boldsymbol{\xi}}$.

2-2 The Equivalence Theorem of $\mathbf{G_x}$-MINOS and S-LUMBE

Of course, we have included the second chapter on hom **S**-LUMBE in order to interpret $\mathbf{G_x}$-MINOS of the first chapter. When are hom **S**-LUMBE and $\mathbf{G_x}$-MINOS equivalent is answered will by Theorem 2.4.

Theorem 2.4. (equivalence of $\mathbf{G}_x$-MINOS and **S**-LUMBE).

With respect to the special consistent linear Gauss–Markov model (2.1), (2.2) $\hat{\boldsymbol{\xi}} = \mathbf{Ly}$ is hom **S**-LUMBE for a positive-definite matrix **S** if $\boldsymbol{\xi}_\mathbf{m} = \mathbf{Ly}$ is $\mathbf{G_x}$-MINOS of the underdetermined system of linear equations (1.47) for

$$\mathbf{G_x} = \mathbf{S}^{-1} \sim \mathbf{G_x^{-1}} = \mathbf{S} \tag{2.20}$$

The proof is straight forward if we compare directly the solution (1.55) of $\mathbf{G_x}$-MINOS and (2.20) of hom **S**-LUMBE. Obviously the inverse matrix of the metric of the parameter space $\mathbb{X}$ is equivalent to the matrix of the metric of the bias space $\mathbb{B}$. Or conversely, the inverse matrix of the metric of the bias space $\mathbb{B}$ determines the matrix of the metric of the parameter space $\mathbb{X}$. In particular, the bias vector $\boldsymbol{\beta}$ of type (2.19) depends on the vector $\boldsymbol{\xi}$ which is *inaccessible*. The situation is similar to the one in hypothesis testing. We can produce only an estimation $\hat{\boldsymbol{\beta}}$ of the bias vector $\boldsymbol{\beta}$ if we identify $\boldsymbol{\xi}$ by the hypothesis $\boldsymbol{\xi}_\mathbf{0} = \hat{\boldsymbol{\xi}}$.

2-3 Example

Due to the *Equivalence Theorem* $\mathbf{G_x}$-MINOS is equivalent to **S**-LUMBE the only new item is the *bias matrix* $\mathbf{B}\{\hat{\boldsymbol{\xi}}\,|\,\text{hom}\,LUMBE\}$. For our example, let us use the simple assumption $\mathbf{S} = \mathbf{I_m}$. Such an assumption is called "u.s." or "unity substituted" or unity substitute matrix. For our case the matrix $\mathbf{I_m} - \mathbf{A}'(\mathbf{AA}')^{-1}\mathbf{A}$ is idempotent. Note the fact that $\text{tr}\mathbf{A} = \text{rk}\mathbf{A}$ is idempotent. Indeed the Frobenius norm of the u.s. bias matrix $\mathbf{B}\,(\text{hom}\,LUMBE)$ equalizes the square root $\sqrt{m-n} = \sqrt{d}$ of the right complementary index of the matrix **A**

$$||\mathbf{B}||^2 = \text{tr}[\mathbf{I_m} - \mathbf{A}'(\mathbf{AA}')^{-1}\mathbf{A}] \tag{2.21}$$

$$||\mathbf{B}||^2 = d = m - n = m - \text{rk}\mathbf{A} \tag{2.22}$$

Box 2.2 summarizes those data outputs of the front examples of the first chapter relating to $||\mathbf{B}(\text{hom}\,BLUMBE)||$.

Box 2.2. Simple matrix of type u.s., Frobenius norm of the simple bias matrix, front page example.

$$\mathbf{A} \in \mathbb{R}^{2\times 3},$$

$$n = 2,\ m = 3,$$

$$\mathbf{A} := \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix}, \mathbf{AA}' = \begin{bmatrix} 3 & 7 \\ 7 & 21 \end{bmatrix}, (\mathbf{AA}')^{-1} = \frac{1}{14}\begin{bmatrix} 21 & -7 \\ -7 & 3 \end{bmatrix}$$

$$\text{rk}\mathbf{A} = 2$$

$$\mathbf{A}'(\mathbf{AA}')^{-1} = \frac{1}{14}\begin{bmatrix} 14 & -4 \\ 7 & -1 \\ -7 & 5 \end{bmatrix}, \mathbf{A}'(\mathbf{AA}')^{-1}\mathbf{A} = \frac{1}{14}\begin{bmatrix} 10 & 6 & -2 \\ 6 & 5 & 3 \\ -2 & 3 & 13 \end{bmatrix}$$

$$(\mathbf{AA}')^{-2} = \frac{1}{98}\begin{bmatrix} 245 & -84 \\ -84 & 29 \end{bmatrix}, \mathbf{A}'(\mathbf{AA}')^{-2}\mathbf{A} = \frac{1}{98}\begin{bmatrix} 106 & 51 & -59 \\ 51 & 25 & -27 \\ -59 & -27 & 37 \end{bmatrix}$$

$$\boldsymbol{\Sigma}_{\hat{\boldsymbol{\xi}}} = \mathbf{A}'(\mathbf{AA}')^{-2}\mathbf{A}\sigma_{\mathbf{y}}^2 = \frac{1}{98}\begin{bmatrix} 106 & 51 & -59 \\ 51 & 25 & -27 \\ -59 & -27 & 37 \end{bmatrix}\sigma_{\mathbf{y}}^2$$

$$||\mathbf{B}||^2 = \text{tr}\left[\mathbf{I_m} - \mathbf{A}'(\mathbf{AA}')^{-1}\mathbf{A}\right] = \text{tr}\mathbf{I_3} - \text{tr}\mathbf{A}'(\mathbf{AA}')^{-1}\mathbf{A}$$

$$||\mathbf{B}||^2 = 3 - \tfrac{1}{14}(10 + 5 + 13) = 3 - 2 = 1 = d$$

$$||\mathbf{B}|| = 1 = \sqrt{d}.$$

http://www.springer.com/978-3-642-22240-5

Applications of Linear and Nonlinear Models
Fixed Effects, Random Effects, and Total Least Squares
Grafarend, E.; Awange, J.
2012, XXI, 1016 p. 111 illus., 8 illus. in color., Hardcover
ISBN: 978-3-642-22240-5