
ELSEVIER
Volume 141
Issue 9
September 2011
ISSN 0378-3758
journal of statistical planning and inference
For Contents see page iii
Available online at www.sciencedirect.com
ScienceDirect

Journal of Statistical Planning and Inference 141 (2011) 3181–3192

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

Estimating common parameters in heterogeneous random effects models

Andrew L. Rukhin

Statistical Engineering Division, National Institute of Standards and Technology, Gaithersburg, MD 20899-0001, USA

A R T I C L E I N F O

Article history:
Received 17 August 2010
Received in revised form
24 March 2011
Accepted 2 April 2011
Available online 13 April 2011

Keywords:
Almost unbiased variance estimator
DerSimonian–Laird procedure
Estimating equation
Growth curve model
Maximum likelihood
Meta-analysis
Random coefficient model
Variance components

A B S T R A C T

A question of fundamental importance for meta-analysis of heterogeneous multi-dimensional data studies is how to form a best consensus estimator of common parameters, and what uncertainty to attach to the estimate. This issue is addressed for a class of unbalanced linear designs which include classical growth curve models. The solution obtained is similar to the popular DerSimonian and Laird (1986) method for a simple meta-analysis model. By using almost unbiased variance estimators, an estimator of the covariance matrix of this procedure is derived. Combination of these methods is illustrated by two examples and are compared via simulation.

Published by Elsevier B.V.

1. Introduction: the model and its matrix formulation

The use and importance of linear mixed models is well documented (McCulloch and Searle, 2001; Jiang, 2007). One of their important applications is meta-analysis to combine measurements made in several studies which commonly exhibit not only non-negligible between-study variability, but also have different within-study precisions.

Laird and Ware (1982) discussed several statistical methods of fitting linear mixed models by using classical techniques of (empirical) Bayes or maximum likelihood estimation. The algorithms for the maximum likelihood and the restricted maximum likelihood are implemented in R-language (Pinheiro and Bates, 2000).

Consider a mixed effects linear model where several measurements are made in each of p studies with the i-th study performing measurements n_i times, $i=1,\ldots,p$, so that the data vector from study i has the form

$$Y_i = B_i\theta + C_i b_i + e_i. \tag{1}$$

Here B_i is the i-th study $n_i \times q$ design matrix, the $n_i \times r$ matrix C_i is discussed later, and the q-dimensional parameter θ is of interest. The independent r-dimensional vectors b_i represent random *between-study* effects with zero mean and some covariance matrix Ξ, while the errors e_i are independent and normally distributed. The usual motivation of (1) is provided by two-stage modeling with the first stage introducing all parameters and variables for fixed b_i, and the second stage specifying the distribution of these effects.

E-mail addresses: andrew.rukhin@nist.gov, rukhin@ams.jhu.edu

0378-3758/$ - see front matter Published by Elsevier B.V.
doi:10.1016/j.jspi.2011.04.005

For a general growth curve model in (1) $B_i = C_i A_i$ with given $r \times q$ matrices A_i, $i = 1, \ldots, p$, i.e., the column space of the design matrix of fixed effects is contained in that of random effects matrix. If the rank of B_i is q (and we impose this condition), then $q \le r$, or the dimension of random effects must be at least as large as that of the structural parameter. In many applications the reverse inequality is more natural. Indeed in the simplest case of linear regression ($q=2$), the random effects affecting the slope may change the nature of the regression line (i.e., make some regression lines decreasing and some increasing), so that models involving more than two random effects hardly are promising.

In this paper we assume that $C_i = B_i A_i$ with known $q \times r$ matrices A_i, $i = 1, \ldots, p$. The ensuing condition, $r \le q$, may be in better agreement with the idea that simpler models with fairly small between-study effects are most useful. See also Section 5. Thus, we study a heterogeneous random coefficient model of the form

$$Y_i = B_i(\theta + \ell_i) + e_i, \tag{2}$$

with B_i and θ having the same meaning as in (1), and independent normal random vectors $\ell_i = A_i b_i$ have zero mean and the covariance matrix $A_i \Xi A_i^T$. It is also assumed that the errors e_i are independent and normally distributed with the variance depending only on the study (but not on the design row), $e_i \sim \mathbf{N}_{n_i}(0, \sigma_i^2 I)$. The statistical goal is to estimate the parameter θ and to provide a standard error of this estimate leading to a confidence region for this parameter or for a function thereof.

This model extends the well-studied balanced scenario ($n_i \equiv n$, $\sigma_i^2 \equiv \sigma^2$) (Longford, 1994). When $q=r$, it falls into the class of classical growth curve models (Demidenko, 2005; Pan and Fang, 2002), with available results on maximum likelihood and method of moments estimators. In more general balanced growth models, the $n \times m$ matrix data $\mathbf{Y}$ can be represented as $\widetilde{\mathbf{B}}\theta\widetilde{\mathbf{A}} + e$, where θ is a $r \times q$ parametric matrix, $\widetilde{\mathbf{B}}$ and $\widetilde{\mathbf{A}}$ are $n \times r$ and $q \times m$ known matrices, and e is the error term. The model (2) can be considered as a heterogeneous version of reduced-rank growth curve models.

In matrix notation (1) can be written as a particular case of the mixed linear model,

$$\mathbf{Y} = \mathbf{B}\theta + \mathbf{C}\mathbf{b} + e.$$

Here $\mathbf{Y}$ is the total data vector of dimension $N = n_1 + \cdots + n_p$; $\mathbf{B}$ is a matrix of size $N \times q$ formed by stacked in a column matrices $B_1, \ldots, B_p$; $\mathbf{C}$ is the block diagonal matrix of size $N \times rp$ with the i-th block of size $n_i \times r$ consisting of the matrix C_i. The rp-dimensional random vector $\mathbf{b}$ is composed of stacked p components formed by b_i, $i = 1, \ldots, p$. Thus θ is the unknown parameter (fixed effects) vector, and $\mathbf{b}$ is random normal effect vector uncorrelated with the zero mean normal error vector e whose block-diagonal covariance matrix consists of blocks $\sigma_i^2 I_{n_i}$.

When there is just one unknown mean parameter ($q=1$), DerSimonian and Laird (1986) suggested an estimation method of θ, which became immensely popular in biostatistics, especially in analysis of multicenter clinical trials. This popularity is due mainly to the fact that this is a simple non-iterative procedure, which exhibits good properties relative to the classical maximum likelihood estimator of θ (e.g., Jackson et al., 2010), and admits generalizations to non-normal models (Böhning et al., 2002). Our primary goal is to extend this method to a multiparameter situation in the linear model context (2).

Random coefficient models are widely used in econometrics, biological, social and management studies, etc. (see Bliese and Ployhart, 2002 for a survey). Our motivation is collaborative studies of meta-analysis, in particular arising in metrology applications. As a first motivating example, we use a study by Paule and Mandel (1971) in which several laboratories performed (via different techniques) measurements of silver vapor pressure P as a function of the absolute temperature T. According to the heat law, the logarithm of pressure, $Y = \log P$, is a linear function of $1/T$. After removal of obvious outliers there are 298 different temperature points provided by $p=8$ laboratories as given in Table 4, Paule and Mandel (1971). We assume here that the error variance depends only on the laboratory (and not on the temperature value). This study then fits the model (2) with $q=2$, if the matrix B_i is formed by ones and the T-reciprocals employed by i-th laboratory. The slope and intercept estimators derived by different methods are discussed in Section 6.1.

The collaborative studies in the area of acoustics, ultrasound and vibration present many important applications of heterogeneous mixed effect models. A typical example is the accelerometers key comparisons study SIM-AUV.V-K1, Evans et al. (2009) in which the sensitivity (electrical output per unit acceleration input) of three accelerometers was measured as a function of frequency. A quadratic polynomial fit of the combined data of seven National Metrology Institutes for the charge sensitivity as a function of frequency is discussed in Section 6.2. Another metrology oriented application of heterogeneous growth curve models is in round-robin studies.

In the next section we reduce the parameter estimation problem to that of a common vector mean estimation and review the weighted means statistics. Section 3 develops a multivariate version of the meta-analytic method by DerSimonian and Laird. Estimation of the covariance matrix of this statistic is investigated in Section 4, and a feature of best linear unbiased predictors of random effects is discussed in Section 5. The concluding Section 7 reports some Monte Carlo simulation results.

2. Sufficient statistics and matrix weighted means

For fixed ℓ_i, $i = 1, \ldots, p$, the statistics $X_i = (B_i^T B_i)^{-1} B_i^T Y_i$ (the ordinary least squares estimators) and

$$s_i^2 = (Y_i - B_i X_i)^T (Y_i - B_i X_i)$$

are sufficient in the model (2). Indeed

$$\sum_i \sigma_i^{-2}(Y_i - B_i\theta - B_i\ell_i)^T(Y_i - B_i\theta - B_i\ell_i) = \sum_i \sigma_i^{-2}(Y_i - B_iX_i)^T(Y_i - B_iX_i) + \sum_i \sigma_i^{-2}(X_i - \theta - A_i\ell_i)^T B_i^T B_i(X_i - \theta - A_i\ell_i).$$

While X_i has the normal distribution $N_q(\theta, \sigma_i^2(B_i^TB_i)^{-1} + A_i\Xi A_i^T)$, because of our assumptions, s_i^2 has a $\sigma_i^2\chi^2$-distribution with $\nu_i = n_i - q$ degrees of freedom. Thus the problem reduces to estimation of the common vector mean θ on the basis of p independent normal vectors X_i with this mean and covariance matrices $\sigma_i^2(B_i^TB_i)^{-1} + A_i\Xi A_i^T, i = 1, \dots, p$. Notice that $\sigma_i^2(B_i^TB_i)^{-1}$ can be estimated via available s_i^2. Indeed one can use the matrix estimator S_i of $\sigma_i^2(B_i^TB_i)^{-1}$,

$$S_i^{-1} = \hat{\sigma}_i^{-2} B_i^T B_i$$

with $\hat{\sigma}_i^2 = s_i^2/\nu_i$.

Let W_i be non-negative definite symmetric matrix weights such that $\sum_{i=1}^p W_i$ is a non-singular matrix. Consider the class of weighted means $\tilde{X}$ of the form

$$\tilde{X} = \left(\sum_{i=1}^p W_i\right)^{-1} \sum_{i=1}^p W_iX_i. \tag{3}$$

These statistics maintain the optimality property of the weighted mean as the best linear unbiased estimator of the common mean θ for the matrix-valued loss function corresponding to the covariance matrix (see for example, Hall, 2007). Indeed, if independent X_i have a common mean θ and known covariance matrices Σ_i, then the best linear estimator of θ has form (3) with $W_i = \Sigma_i^{-1}$. In this situation the covariance matrix, $Var(\tilde{X})$ has the form

$$Var(\tilde{X}) = \left(\sum_{i=1}^p W_i\right)^{-1}. \tag{4}$$

The "averaging" property of $\tilde{X}$ is expressed by inequalities,

$$\min_i X_i(j) \le \tilde{X}_j \le \max_i X_i(j), \quad j = 1, \dots, q.$$

Here $\tilde{X} = (\tilde{X}_1, \dots, \tilde{X}_q)^T$, $X_i = (X_i(1), \dots, X_i(q))^T$.

Estimators of the form (3) include the traditional estimator of the common vector mean suggested when $q = 1$ by Graybill and Deal (1959),

$$\tilde{X}_0 = \left(\sum_{i=1}^p S_i^{-1}\right)^{-1} \sum_{i=1}^p S_i^{-1}X_i, \tag{5}$$

as well as the vector sample mean,

$$\overline{X} = \frac{1}{p}\sum_{i=1}^p X_i.$$

Since $\Sigma_i = [\sigma_i^2(B_i^TB_i)^{-1} + A_i\Xi A_i^T]$, it makes sense to employ matrix weights of the form

$$W_i = W_i(V) = (S_i + A_iVA_i^T)^{-1}, \tag{6}$$

for some non-negative definite $r \times r$ matrix V designed to estimate Ξ.

An estimator $\tilde{X}$ of θ from the class (3) has the following representation:

$$\tilde{X} = \sum_{i=1}^p \omega_i X_i, \tag{7}$$

where $\omega_i = (\sum_{k=1}^p W_k)^{-1} W_i$, $i = 1, \dots, p$, $\sum_{k=1}^p \omega_k = I$. The matrices ω_i do not have to be symmetric or to commute, but their eigenvalues must be positive.

There are general results on maximum likelihood and restricted maximum likelihood estimation in the variance components setting and algorithms for their calculation (Pinheiro and Bates, 2000). However the following procedure, which has the form (7) with data dependent weights, is tailored to our specific problem, and is much easier to evaluate. It also escapes potential (false or singular) convergence problems.

3. DerSimonian–Laird procedure

If for the matrix weights W_i the symmetric matrix $W = \sum_i W_i$ is non-singular, then with $\tilde{X}$ defined by (3),

$$\sum_i W_i^{1/2} E(X_i - \tilde{X})(X_i - \tilde{X})^T (W_i^T)^{1/2} = \sum_{i=1}^p W_i^{1/2}(I - W^{-1}W_i) Var(X_i)(I - W^{-1}W_i)^T (W_i^T)^{1/2}$$

$$+\sum_{i=1}^{p} W_i^{1/2} W^{-1}\left(\sum_{k:k\neq i} W_k Var(X_k) W_k^T\right)(W^T)^{-1}(W_i^T)^{1/2}. \tag{8}$$

In particular, when $W_i=\sigma_i^{-2}B_i^TB_i$, $Var(X_i)=\sigma_i^2(B_i^TB_i)^{-1}+A_i\Xi A_i^T$,

$$\begin{aligned}&\sum_i \sigma_i^{-2}(B_i^TB_i)^{1/2}(I-\sigma_i^{-2}W^{-1}B_i^TB_i)A_i\Xi A_i^T(I-\sigma_i^{-2}W^{-1}B_i^TB_i)^T(B_i^TB_i)^{1/2}\\ &\quad+\sum_i \sigma_i^{-2}(B_i^TB_i)^{1/2}W^{-1}\left(\sum_{k:k\neq i}\sigma_k^{-4}B_k^TB_kA_i\Xi A_i^TB_k^TB_k\right)W^{-1}(B_i^TB_i)^{-1/2}\\ &\quad=\sum_i \sigma_i^{-2}(B_i^TB_i)^{1/2}E(X_i-\widetilde{X})(X_i-\widetilde{X})^T(B_i^TB_i)^{1/2}-pI+\sum_i \sigma_i^{-2}(B_i^TB_i)^{1/2}W^{-1}(B_i^TB_i)^{1/2}.\end{aligned} \tag{9}$$

Provided that σ_i^2 are replaced by $\widehat{\sigma}_i^2$, the identity (9) can be used as an estimating equation for the parameters θ and Ξ in the following way. Let $\widetilde{X}_0$ be the Graybill–Deal estimator (5). Put

$$C=\sum_i \widehat{\sigma}_i^{-2}(B_i^TB_i)^{1/2}(X_i-\widetilde{X}_0)(X_i-\widetilde{X}_0)^T(B_i^TB_i)^{1/2}-pI+\sum_i \widehat{\sigma}_i^{-2}(B_i^TB_i)^{1/2}\left(\sum_k \widehat{\sigma}_k^{-2}B_k^TB_k\right)^{-1}(B_i^TB_i)^{1/2},$$

so that this symmetric matrix estimates the right-hand side of (9). Similarly the matrix-valued weights

$$\widetilde{\omega}_i=\left(\sum_k \widehat{\sigma}_k^{-2}B_k^TB_k\right)^{-1}\widehat{\sigma}_i^{-2}B_i^TB_i=\left(\sum_{k=1}^{p}S_k^{-1}\right)^{-1}S_i^{-1}$$

can be employed to approximate the left-hand side of (9) leading to the equation

$$\sum_i \widehat{\sigma}_i^{-2}(B_i^TB_i)^{1/2}(I-\widetilde{\omega}_i)A_iVA_i^T(I-\widetilde{\omega}_i)^T(B_i^TB_i)^{1/2}+\sum_i \widehat{\sigma}_i^{-2}(B_i^TB_i)^{1/2}\left(\sum_{j:j\neq i}\widetilde{\omega}_jA_iVA_i^T\widetilde{\omega}_j^T\right)(B_i^TB_i)^{1/2}=C. \tag{10}$$

This equation allows one to determine a symmetric matrix solution V (an estimator of Ξ). We put $V_{DL}=V_+$ to be the positive part of V, i.e., let V_{DL} have the same spectral decomposition as V, with eigenvalues being positive parts of V eigenvalues. The matrix weights of the estimator $\widetilde{X}_{DL}$ then have the form (6),

$$W_i=(A_iV_{DL}A_i^T+S_i)^{-1}. \tag{11}$$

Eq. (10) extends the procedure suggested by DerSimonian and Laird (1986) when $q=1$. Similar, moment-type estimating equations are considered by Rukhin (2007) when $B_i\equiv B$, $B^TB=I$, and by Demidenko (2005, pp. 192, 288) under condition that all σ_i^2 are known. Notice that the quadratic form in the residuals $X_i-\widetilde{X}$ used here to obtain (10) is different from the one employed by Demidenko (2005) whose moment-type estimator does not coincide with the DerSimonian–Laird solution when $q=1$.

To solve (10), denote by $Vec(F)$ the $q^2\times 1$ vector formed by stacking the columns of the $q\times q$ matrix F under each other, and by $F\otimes G$ the tensor (Kronecker) product of matrices F and G.

Then, according to Lemma 16.1.2 and Theorem 16.2.1 of Harville (1997),

$$Vec(S_i^{-1/2}(I-\widetilde{\omega}_i)A_iVA_i^T(I-\widetilde{\omega}_i)^TS_i^{-1/2})=[S_i^{-1/2}(I-\widetilde{\omega}_i)A_i\otimes S_i^{-1/2}(I-\widetilde{\omega}_i)A_i]Vec(V)$$

and

$$Vec\left(\sum_i S_i^{-1/2}\left(\sum_{j\neq i}\widetilde{\omega}_jA_iVA_i^T\widetilde{\omega}_j^T\right)S_i^{-1/2}\right)=\sum_i\sum_{j\neq i}(S_i^{-1/2}\widetilde{\omega}_jA_i\otimes S_i^{-1/2}\widetilde{\omega}_jA_i)Vec(V).$$

Thus, with

$$R=\sum_i S_i^{-1/2}(I-\widetilde{\omega}_i)A_i\otimes S_i^{-1/2}(I-\widetilde{\omega}_i)A_i+\sum_{i\neq j}S_i^{-1/2}\widetilde{\omega}_jA_i\otimes S_i^{-1/2}\widetilde{\omega}_jA_i, \tag{12}$$

the "vectorized" version of (10) can be written as

$$RVec(V)=Vec(C).$$

Assuming that the matrix R is invertible, one obtains

$$Vec(V)=R^{-1}Vec(C). \tag{13}$$

Eq. (10) can be solved by using matrices of smaller size. Denote by $Vech(F)$ the $q(q+1)/2\times 1$ vector formed by stacking the subdiagonal elements of a $q\times q$ symmetric matrix F under each other, $Vech(F)=(f_{11},f_{21},\ldots,f_{q1},f_{22},\ldots,f_{2q},\ldots,f_{qq})^T$. Then for every symmetric matrix F, $Vec(F)=G_qVech(F)$, with the *duplicating* matrix G_q of size $q^2\times q(q+1)/2$ (see Harville, 1997, Section 16.4). If H_q is a left inverse of G_q, (10) means that

$$H_q R G_q Vech(V) = Vech(C),$$

and $Vech(V)$ can be obtained from this equation.

For example, when $q=2$, the elements r_{ij} of the 4×4 matrix R are such that $r_{12}=r_{13}$, $r_{21}=r_{31}$, $r_{24}=r_{34}$, $r_{42}=r_{43}$, $r_{22}=r_{33}$, $r_{23}=r_{32}$. Indeed R is a sum of tensor squares, so that

$$H_2 R G_2 = \begin{pmatrix} r_{11} & 2r_{12} & r_{14} \\ r_{21} & r_{22}+r_{23} & r_{24} \\ r_{41} & 2r_{42} & r_{44} \end{pmatrix},$$

and (13) reduces to three simultaneous linear equations.

If $A_i \equiv A$ and $B_i^T = (B,\dots,B)^T$ with the same $m\times q$ design matrix B of rank q repeated r_i times ($n_i = r_{im}$), the formulas above simplify dramatically. Indeed put $Y_i = (Y_1^{(i)},\dots,Y_{r_i}^{(i)})^T$, $\overline{Y}_i = \sum_{k=1}^{r_i} Y_k^{(i)}/r_i$. Then $B_i^T B_i = r_i B^T B$, $X_i = (B^TB)^{-1}B^T\overline{Y}_i$, and

$$\widehat{\sigma}_i^2 = \frac{\sum_{k=1}^{r_i}(Y_k^{(i)} - B(B^TB)^{-1}B^T\overline{Y}_i)^T(Y_k^{(i)} - B(B^TB)^{-1}B^T\overline{Y}_i)}{n_i - q}.$$

The Graybill–Deal weights are

$$w_i = \left(\sum_k \frac{r_k}{\widehat{\sigma}_k^2}\right)^{-1} \frac{r_i}{\widehat{\sigma}_i^2}, \quad i=1,\dots,p,$$

so that

$$\widetilde{X}_0 = \sum_i w_i X_i = (B^TB)^{-1}B^T\sum_i w_i \overline{Y}_i.$$

Put $\widetilde{Y}_0 = \sum w_i \overline{Y}_i$ to get

$$C = (B^TB)^{-1/2}B^T\sum_i \frac{r_i}{\widehat{\sigma}_i^2}(\overline{Y}_i - \widetilde{Y}_0)(\overline{Y}_i - \widetilde{Y}_0)^T B(B^TB)^{-1/2} - (p-1)I.$$

The solution $V=V_{DL}$ in (10),

$$AV_{DL}A^T = \left(\sum_i \frac{r_i}{\widehat{\sigma}_i^2} - \left[\sum_i \frac{r_i}{\widehat{\sigma}_i^2}\right]^{-1}\sum_i \frac{r_i^2}{\widehat{\sigma}_i^4}\right)^{-1} \times \left[\sum_i \frac{r_i}{\widehat{\sigma}_i^2}(X_i - \widetilde{X}_0)(X_i - \widetilde{X}_0)^T - (p-1)B^TB\right]_+$$

$$= \left(\left[\sum_i \frac{r_i}{\widehat{\sigma}_i^2}\right]^2 - \sum_i \frac{r_i^2}{\widehat{\sigma}_i^4}\right)^{-1} \times \left[\sum_{i<j} \frac{r_i r_j}{\widehat{\sigma}_i^2\widehat{\sigma}_j^2}(X_i - X_j)(X_i - X_j)^T - (p-1)\sum_i \frac{r_i}{\widehat{\sigma}_i^2}(B^TB)^{-1}\right]_+, \tag{14}$$

is the direct extension of DerSimonian and Laird (1986, p 183) formula. Notice that

$$\sum_{i<j}\frac{r_i r_j}{\sigma_i^2\sigma_j^2}E(X_i - X_j)(X_i - X_j)^T = \sum_{i<j}\frac{r_i r_j}{\sigma_i^2\sigma_j^2}\left[2AYA^T + \left(\frac{r_i}{\sigma_i^2} + \frac{r_j}{\sigma_j^2}\right)(B^TB)^{-1}\right] = \left[\left(\sum_i \frac{r_i}{\sigma_i^2}\right)^2 - \frac{r_i^2}{\sigma_i^4}\right]AYA^T + (p-1)\sum_i \frac{r_i}{\sigma_i^2}(B^TB)^{-1},$$

so that

$$\frac{A(EV-Y)A^T}{p-1} = E\frac{\sum_i r_i\left(\frac{1}{\sigma_i^2} - \frac{1}{\widehat{\sigma}_i^2}\right)}{\left[\sum_i \frac{r_i}{\widehat{\sigma}_i^2}\right]^2 - \sum_i \frac{r_i^2}{\widehat{\sigma}_i^4}}(B^TB)^{-1}.$$

When $p=2$,

$$AV_{DL}A^T = \frac{1}{2}(B^TB)^{-1/2}\left[(B^TB)^{1/2}(X_1 - X_2)(X_1 - X_2)^T(B^TB)^{1/2} - \left(\frac{\widehat{\sigma}_1^2}{r_1} + \frac{\widehat{\sigma}_2^2}{r_2}\right)I\right]_+(B^TB)^{-1/2},$$

which shows that $V_{DL}=0$, if

$$(X_1 - X_2)^T B^T B(X_1 - X_2) \le \frac{\widehat{\sigma}_1^2}{r_1} + \frac{\widehat{\sigma}_2^2}{r_2}.$$

Otherwise the rank of V_{DL} is one. More generally, if Q is the Moore–Penrose generalized inverse of the matrix $\sum r_i \widehat{\sigma}_i^{-2}(X_i - \widetilde{X}_0)(X_i - \widetilde{X}_0)^T$, then the matrix in the right-hand side of (14) is

$$\sum \frac{r_i}{\widehat{\sigma}_i^2}(B^TB)^{1/2}(X_i-\widetilde{X}_0)[(p-1)I-Q]_+(X_i-\widetilde{X}_0)^T(B^TB)^{1/2}.$$

It follows that the rank of V_{DL} cannot exceed $\min(q,p-1)$. When $p>q$, there are solutions V_{DL} of full rank q.

Proposition 3.1. *In the notation of Section2, the DerSimonian–Laird estimator $\widetilde{X}_{DL}$ of the parameter θ is the weighted means statistic* (3) *whose matrix weights are*

$$W_i = [A_i V_{DL} A_i^T + \widehat{\sigma}_i^2 (B_i^T B_i)^{-1}]^{-1}, \quad i=1,\ldots,p,$$

with the non-negative matrix V_{DL} satisfying Eq. (10). *If the matrix* (12) *is non-singular, V_{DL} can be found from* (13). *In the balanced case ($A_i \equiv A$, $B_i^T=(B,\ldots,B)^T, n_i=r_i m$), the matrix $AV_{DL}A^T$ is determined by* (14).

In the next section we discuss a method to estimate the variance of a matrix weighted estimators.

4. Estimation of the covariance matrix: almost unbiased statistic

Here we suggest an estimator of the covariance matrix $Var(\widetilde{X})$ similar to the one advocated in a more general setting of linear models (such as (1)) by Horn et al. (1975).

Let ω_i be fixed normalized matrix weights, $\sum_i \omega_i = I$. To estimate the matrix $Var(\widetilde{X}) = \sum_i \omega_i Var(X_i)\omega_i^T$, for the (unbiased) weighted means statistic $\widetilde{X}$, one can use the almost unbiased estimate of $Var(X_i)$. For a fixed $i=1,\ldots,p$,

$$Var(X_i-\widetilde{X}) = (I-\omega_i)Var(X_i)(I-\omega_i)^T + \sum_{k\neq i}\omega_k Var(X_k)\omega_k^T = \sum_k \omega_k Var(X_k)\omega_k^T + Var(X_i) - \omega_i Var(X_i) - Var(X_i)\omega_i^T.$$

When $\omega_i = [\sum_k Var(X_k)^{-1}]^{-1} Var(X_i)^{-1}$, the first term in the right-hand side simplifies to

$$\sum_k \omega_k Var(X_k)\omega_k^T = \left[\sum_k Var(X_k)^{-1}\right]^{-1} = \frac{1}{2}\omega_i Var(X_i) + \frac{1}{2}Var(X_i)\omega_i^T,$$

which holds for all $i=1,\ldots,p$. By substituting this expression in the previous formula, one obtains

$$Var(X_i-\widetilde{X}) = Var(X_i) - \tfrac{1}{2}[\omega_i Var(X_i) + Var(X_i)\omega_i^T].$$

Horn et al. (1975, p. 382) (see also Mandel, 1964) argue that by continuity if the weights are only approximately correct, this is an approximate identity. Thus, an *almost unbiased* estimator V_i of $Var(X_i)$ is derived by solving the following equation:

$$(X_i-\widetilde{X})(X_i-\widetilde{X})^T = V_i - \tfrac{1}{2}(\omega_i V_i + V_i \omega_i^T).$$

To find this solution, let

$$Vech(\widehat{Var}(X_i)) = [H_q\{(I-\tfrac{1}{2}\omega_i)\otimes(I-\tfrac{1}{2}\omega_i) - \tfrac{1}{4}\omega_i\otimes\omega_i\}G_q]^{-1} \times Vech((X_i-\widetilde{X})(X_i-\widetilde{X})^T). \tag{15}$$

As in Section 3, taking the positive part of a non-positively defined symmetric matrix, makes sense here too. Actually, as in our situation, $Var(X_i) \geq \sigma_i^2(B_i^TB_i)^{-1}$, and an unbiased estimate $\widehat{\sigma}_i^2$ of σ_i^2 is available, it seems reasonable to use as the final estimate of $Var(X_i)$,

$$\widehat{Var}(X_i) = \max[V_i, \widehat{\sigma}_i^2(B_i^TB_i)^{-1}] = \widehat{\sigma}_i^2(B_i^TB_i)^{-1} + [V_i - \widehat{\sigma}_i^2(B_i^TB_i)^{-1}]_+$$

with V_i determined from (15) (cf. Rukhin, 2007).

An alternative form of the estimator in (15) comes from the formula

$$V_i = (I-\tfrac{1}{2}\omega_i)^{-1}(X_i-\widetilde{X})(X_i-\widetilde{X})^T(I-\tfrac{1}{2}\omega_i^T)^{-1} + \tfrac{1}{4}(I-\tfrac{1}{2}\omega_i)^{-1}\omega_i V_i \omega_i^T (I-\tfrac{1}{2}\omega_i^T)^{-1},$$

leading to the iteration scheme

$$V_i^{(n+1)} = (I-\tfrac{1}{2}\omega_i)^{-1}(X_i-\widetilde{X})(X_i-\widetilde{X})^T(I-\tfrac{1}{2}\omega_i^T)^{-1} + \tfrac{1}{4}(I-\tfrac{1}{2}\omega_i)^{-1}\omega_i \max[V_i^{(n)}, \widehat{\sigma}_i^2(B_i^TB_i)^{-1}]\omega_i^T(I-\tfrac{1}{2}\omega_i^T)^{-1},$$

$n=0,1,\ldots, V_i^{(0)}=0$, which typically converges very fast.

The resulting formula for the $Var(\widetilde{X})$ estimator has the form

$$\widehat{Var}(\widetilde{X}) = \sum_i \omega_i \widehat{Var}(X_i)\omega_i^T, \tag{16}$$

which can be recovered from its vectorized version via (15),

$$Vech(\widehat{Var}(\widetilde{X})) = \sum_i H_q(\omega_i\otimes\omega_i)G_q \times \left[H_q\left\{\left(I-\frac{\omega_i}{2}\right)\otimes\left(I-\frac{\omega_i}{2}\right) - \frac{\omega_i\otimes\omega_i}{4}\right\}G_q\right]^{-1} Vech((X_i-\widetilde{X})(X_i-\widetilde{X})^T). \tag{17}$$

The statistic (16) gives an estimate of the covariance matrix of any weighted means statistic for fixed weights ω_i. This estimate is uniquely defined when $I-2^{-1}\omega_i \otimes I-2^{-1} I \otimes \omega_i$ is invertible (which holds provided that $I-\omega_i$ is non-singular.)

Proposition 4.1. *The almost unbiased estimator* (16) *of* $Var(\widetilde{X})$ *satisfies* (15), *and its vectorized version is given by* (17).

If $\omega_i = w_i I$ with positive scalar weights $w_i, \sum w_k = 1$, then the formula for $\widehat{Var}(\widetilde{X})$ simplifies

$$\widehat{Var}(\widetilde{X}) = \sum_i w_i^2 \max\left[\frac{1}{1-w_i}(X_i-\widetilde{X})(X_i-\widetilde{X})^T, \widehat{\sigma}_i^2 (B_i^T B_i)^{-1}\right].$$

For example if $w_i \equiv p^{-1}$, i.e. $\widetilde{X} = \overline{X}$ is the sample mean, $\widetilde{Var}(\widetilde{X}) = [p(p-1)]^{-1} \sum_i (X_i-\overline{X})(X_i-\overline{X})^T$ is the classical sample covariance matrix.

An alternative estimator (4) of $Var(\widetilde{X})$ for a weighted means statistic (3) is commonly used, in particular when $\widetilde{X}$ is the maximum likelihood estimator or $\widetilde{X} = \widetilde{X}_0$. Simulations reported in Section 7 indicate that (16) is a better estimate of the covariance matrix than (4) in terms of the coverage probability of resulting confidence regions.

An approximate $(1-\alpha)$ confidence ellipsoid for θ has the form

$$(\widetilde{X}-\theta)^T [\widehat{Var}(\widetilde{X})]^{-1} (\widetilde{X}-\theta) \le q F_{q,p-q}(\alpha),$$

where $F_{q,p-q}(\alpha)$ denotes the critical point of F-distribution with indicated degrees of freedom. This suggestion also parallels the case $q=1$ (Follmann and Proshan, 1999).

A confidence interval for a linear function of θ, say, $a^T\theta$, follows as $\widehat{Var}(a^T\widetilde{X}) = a^T \widehat{Var}(\widetilde{X}) a$. An approximate $(1-\alpha)$ confidence interval has the form

$$a^T\widetilde{X} \pm t_{\alpha/2}(p-q)\sqrt{a^T \widehat{Var}(\widetilde{X}) a}. \tag{18}$$

Simultaneous confidence intervals for several linear functions can be derived similarly. For example, when $a = X^T b$, these simultaneous intervals are

$$a^T\widetilde{X} \pm \sqrt{q F_{q,p-q}(\alpha) a^T \widehat{Var}(\widetilde{X}) a}.$$

5. Estimation of random effects

There are situations when the random effects b_i are of interest (see Robinson, 1991; Beran, 1995). In animal breeding they are used to estimate generic merits. In metrology applications the random effects can be interpreted as the deviations of individual lab measurements from the consensus mean θ, which are to be estimated according to the protocol in many collaborative studies.

The formulas for the best linear unbiased predictors (BLUP) $\widehat{b}_i$ are well known. In the model (1) they are found as minimizers in

$$\begin{aligned}
&\min_{\theta, b_1,\ldots,b_p}\left[\sum_i \sigma_i^{-2}(Y_i-B_i\theta-C_ib_i)^T(Y_i-B_i\theta-C_ib_i)+\sum_i b_i^T \Xi^{-1} b_i\right] \\
&= \min_{\theta, b_1,\ldots,b_p, c}\left[\sum_i b_i^T \Xi^{-1} b_i + \sum_i \sigma_i^{-2}(Y_i-B_i\theta-B_ic-C_ib_i)^T(Y_i-B_i\theta-B_ic-C_ib_i)\right] \\
&= \min_{\substack{\theta, b_1,\ldots,b_p, \\ c: B_i c = C_i d_i, i=1,\ldots,p}}\left[\sum_i b_i^T \Xi^{-1} b_i + \sum_i \sigma_i^{-2}(Y_i-B_i\theta-C_i(b_i+d_i))^T(Y_i-B_i\theta-C_i(b_i+d_i))\right] \\
&= \min_{\theta, b_1,\ldots,b_p}\left[\sum_i \sigma_i^{-2}(Y_i-B_i\theta-C_ib_i)^T(Y_i-B_i\theta-C_ib_i)+\min_{d_i: C_i d_i = B_i c, i=1,\ldots,p}\sum_i (b_i+d_i)^T \Xi^{-1}(b_i+d_i)\right].
\end{aligned}$$

If $b_i = \widehat{b}_i$, the last minimum is attained when $d_i = 0$. If for $i = 1,\ldots,p$, $C_i d_i = B_i c$, then one must have

$$\sum_i d_i^T \Xi^{-1} \widehat{b}_i = 0.$$

Therefore BLUP of some linear combinations of random effects are identically equal to zero. This phenomenon is discussed by Searle (1997) who uses the explicit formula for the BLUP. If $B_i = C_i A_i$, then one can put $d_i = A_i c$, so that in the traditional growth curve models there are always non-trivial linear combinations of random effects whose best linear predictor is zero. Our model, in which $C_i = B_i A_i$, leads to a smaller dimension of the subspace formed by such combinations.

6. Examples

In a typical application of heterogeneous growth curve models in meta-analysis there are p studies (laboratories, medical centers, etc.). The j-th study makes its measurements at n_j settings (frequencies, temperatures, dose or treatment levels, etc.). With the matrix B_j formed by q columns, each representing the values of the corresponding regression function at given settings, one can employ model (2) for the n_j-dimensional vector Y_j which is formed by the j-th study data and whose independent components are assumed to have the same variance. With the grouped data set, data.dat, formed by the columns: lab, setting, response, the following R-program evaluates the maximum likelihood estimates and their uncertainties for a linear regression model ($q=2$).

```
require(nlme)
data=read.table("data.dat", header=TRUE)
GD <- groupedData(response setting|lab, data)
mlest <- lme(response setting, GD, random=(1+setting)|lab,
weights=varIdent(form = 1|lab), control= lmeControl(returnObject=TRUE))
summary(mlest)
```

There are situations when lab j makes r_j runs of measurements, with k-th run made at $m_k^{(j)}$ settings. Then the total number of observations for j-th study is $n_j = m_1^{(j)} + \cdots + m_{r_j}^{(j)}$, and the grouped data set is formed by four columns: lab, setting, run, response.

This generic setup is illustrated by the following real life examples.

6.1. Silver vapor pressure study

In the silver vapor pressure study (Paule and Mandel, 1971) several laboratories performed via different technique measurements of silver vapor pressure as a function of the absolute temperature T in the (individual for each laboratory) range from 800 to 1600 K. After removal of dubious results of one laboratory, there are a total of 298 different temperature points T_{ij}, $i=1,\ldots,8$, $j=1,\ldots,n_i$, $n_1+\cdots+n_8=298$ given in Table 4 in Paule and Mandel (1971) which employs $1/T10^4$ in K^{-1} units.

As the logarithm of pressure is supposed to be a linear function of $1/T$, the design matrix B_i is formed by pairs $(1,1/T_{ij})$, $j=1,\ldots,n_i$. A natural assumption is that the error variance depends only on the individual laboratory (and not on the temperature value). This study then fits the model (2) with $p=8$ and $q=2$. Fig. 1 displays the data set.

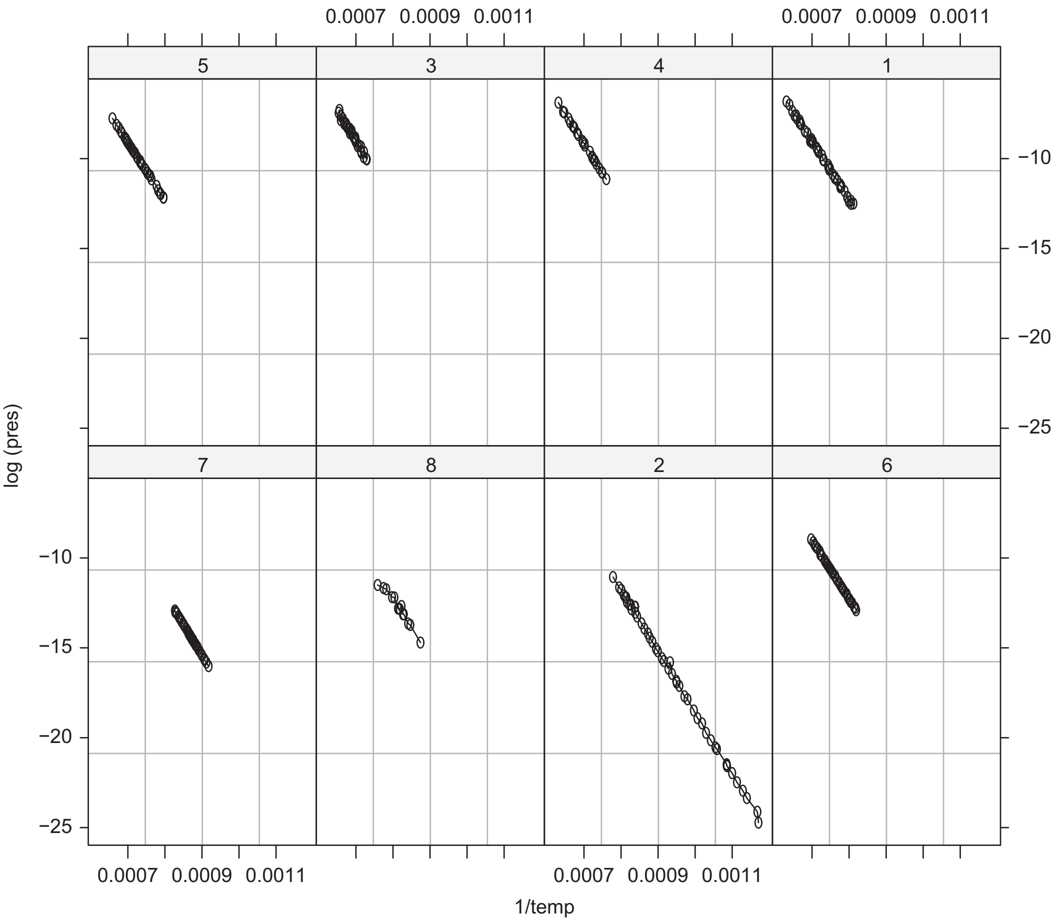

Fig. 1. Silver vapor pressure data.

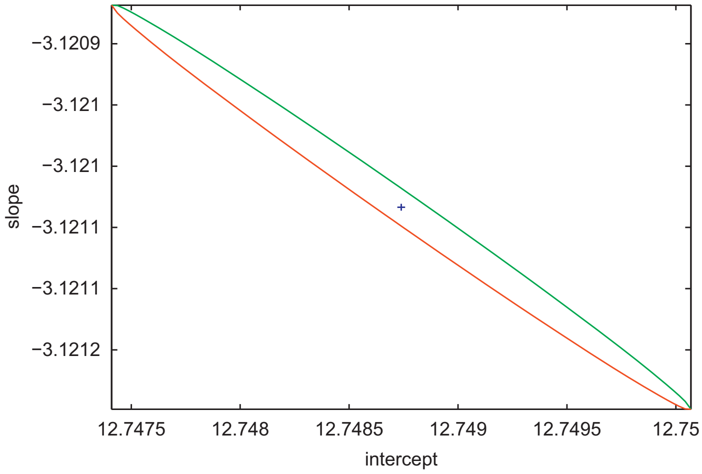

Fig. 2. Confidence ellipsoid for the slope and the intercept in the silver vapor pressure study.

Here are the estimates of the intercept θ_0 and the slope θ_1 with the (restricted) maximum likelihood estimator $\widehat{X}$ found from the R-language function *lme*,

$$\begin{array}{lcccc} & \widetilde{X} & \widetilde{X}_0 & \overline{X} & \widehat{X} \\ \theta_0 & 13.46 & 15.68 & 15.02 & 14.07\,. \\ \theta_1 & -3.21 & -3.42 & -3.43 & -3.28 \end{array}$$

Numerical evaluation of $\widehat{X}$ would not be possible without the removal of the mentioned outliers, as with the full data, an error message in *lme* function indicated false convergence.

The estimated between lab variance Ξ is

$$\begin{pmatrix} 2.00 & -0.26 \\ -0.26 & 0.03 \end{pmatrix},$$

the *lme* estimator is

$$\begin{pmatrix} 0.79 & -0.07 \\ -0.07 & 0.01 \end{pmatrix}.$$

Quantification of heterogeneity effect in interlaboratory studies is quite important. Paule and Mandel (1971) (who were unaware of the size of Ξ) write "..a typical laboratory's ability to reproduce its own vapor pressure measurements exceeds its ability to reproduce other laboratories' measurements".

The almost unbiased estimator (16) of $Var(\widetilde{X})$ of $\widetilde{X}$ ($1/T10$ in K^{-1} units) is

$$\begin{pmatrix} 1.43 & -0.19 \\ -0.19 & 0.02 \end{pmatrix}.$$

The covariance matrix of $\widehat{x}$ of the restricted maximum likelihood estimator reported by the *lme* procedure is much smaller,

$$\begin{pmatrix} 0.11 & -0.01 \\ -0.01 & 0.00 \end{pmatrix}.$$

The formula (18) leads to an approximate confidence ellipsoid for θ based on a F-distribution with $q=2$ and $p-q=6$ degrees of freedom. It is portrayed in Fig. 2. This ellipsoid provides useful information about the joint nature of the slope and the intercept which was not available in the original study.

6.2. SIM-AUV.V-K1 key comparison study

The goal of the accelerometers study SIM.AUV.V-K1, Evans et al. (2009) was to compare the results of measurements of sensitivity of uniaxial linear accelerometers over a range of frequencies using sinusoidal input signals. The sensitivity of each accelerometer was determined in terms of electrical charge output as a function of acceleration input at each frequency.

We consider determining the growth curve model for charge sensitivity as a function of the frequency for single-ended accelerometers. Only the results for the frequencies 40, 50, 63, 80, 100, 125, 160, 200, 250, 315, 400, 500, 630 and 800 Hz and single-ended accelerometers are considered here. In this example we took $q=3$ with θ corresponding to the coefficients of a quadratic function based on data from $p=7$ National Metrology Institutes (PTB, BNM-CESTA, CSIRO-NML,

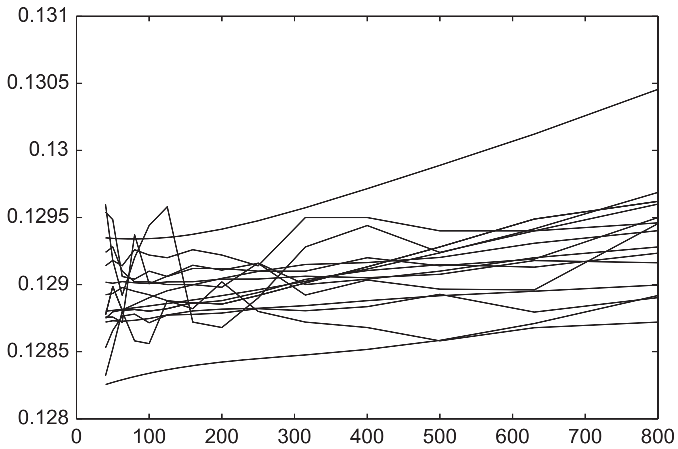

Fig. 3. The SIM.AUV.V-K1 data set and the estimated common curve along with 95% confidence band.

CMI, CENAM, KRISS, NMI-VSL) which used different methods. The quadratic fit is commonly used in piezoelectric accelerometer studies (Serridge and Licht, 1987).

The model (2) was used with the matrices $B_i=(B,\ldots,B)^T$, where the same 14×2 design B formed by the coefficients of functions $1,f,f^2$ at the common frequencies f above is repeated r_i times, $B_i^T B_i = r_i B^T B$ ($r_1=9, r_i=5, i=2,\ldots,7$.)

Here are the estimates of the parameters in /pC/(m/s^2) units,

$$\begin{array}{cccc} & \widetilde{X} & \widetilde{X}_0 & \overline{X} \\ \theta_0 & 0.128778128169481 & 0.128980657524115 & 0.128951225634616 \\ \theta_1 & 0.000000550710309 & 0.000000229182584 & 0.000000157035304 \\ \theta_2 & 0.000000000719480 & 0.000000000080217 & 0.000000000310173 \end{array}.$$

The almost unbiased estimator (16) of $Var(\widetilde{X})$ of $\widetilde{X}$ is

$$\begin{pmatrix} 0.559466287614520 & -0.001042337415196 & -0.000001435646429 \\ -0.001042337415196 & 0.000002112899641 & 0.000000002533601 \\ -0.000001435646429 & 0.000000002533601 & 0.000000000003997 \end{pmatrix},$$

while the estimated within lab variance Ξ is

$$\begin{pmatrix} 0.356768587940128 & -0.000737691375583 & 0.000000759918668 \\ -0.000737691375583 & 0.000004532689216 & -0.000000004683996 \\ 0.000000759918668 & -0.000000004683996 & 0.000000000005507 \end{pmatrix}.$$

The units for these matrices are 10^{-7}/pC/(m/s^2); $\widetilde{Var}(\widetilde{X})$ mentioned in the end of Section 4 is of an order smaller. Determination of the (restricted) maximum likelihood estimators is prohibited by iteration limit reached without convergence. The simultaneous confidence bands are portrayed in Fig. 3. It is believed that this study provided "robust reference values" for charge sensitivity. Moreover, behavior of accelerometer deviations as modeled in (2) for high frequencies has assisted the laboratories in investigating and improving their calibration facilities.

7. Simulation results

The results of a Monte Carlo simulation study for $p=7$, $\theta=(0,1)^T$ are reported here for a linear regression model with the same 5×2 design matrix B whose first column is formed by ones and the second is (2, 2.5, 3, 5, 7.5). The number of repeats r_i is a random permutation of integers from 1 to 7, so that the sample sizes are $n_i=7r_i$. The covariance matrix of the between studies effect was chosen to be $\Xi=[0.5, 0.05; 0.05, 0.015]$. The error variances σ_i^2 were taken to be different multiples ρ of χ^2-distribution with 2 degrees of freedom. The sample variances s_i^2 were realizations of multiples of χ^2-random variables, $s_i^2 \sim \sigma_i^2\chi^2(n_i-2)/(n_i-2)$.

The studied estimators are $\widetilde{X}$, $\widetilde{X}_0$, and $\overline{X}$. Inclusion of $\widehat{X}$ turned out to be prohibitively difficult in view of frequent false or singular convergence problem. These problems are due to possible multimodality of the (restricted) likelihood function or to ill-conditioned covariance matrices. Out of considered procedures, $\widetilde{X}$ or $\overline{X}$ did not experience serious computational difficulties although occasionally the matrix R in (12) happened to be almost singular in which case its Moore–Penrose pseudoinverse was used in (13). There was numerical instability in calculation of $\widetilde{X}_0$ for very small $\widehat{\sigma}_i^2$, and this is one of the reasons why the Graybill–Deal estimator $\widetilde{X}_0$ is not recommended.

Fig. 4 portrays the coverage probability of the confidence ellipsoids with a nominal confidence coefficient of 95% based on $\widetilde{X}$, $\widetilde{X}_0$, and $\overline{X}$. For the estimator $\widetilde{X}$ we used two different confidence regions, one based on (16), another on (4). The

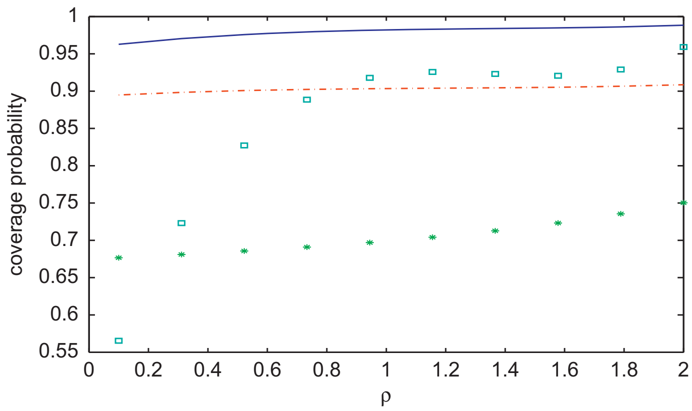

Fig. 4. Plot of actual coverage probabilities for confidence ellipsoids with the nominal 95%-level based on $\tilde{X}$ with (16) (continuous line), $\tilde{X}$ with (4) (line marked by $*$), $\overline{X}$ (dash-dotted line), and $\tilde{X}_0$ (line marked by squares).

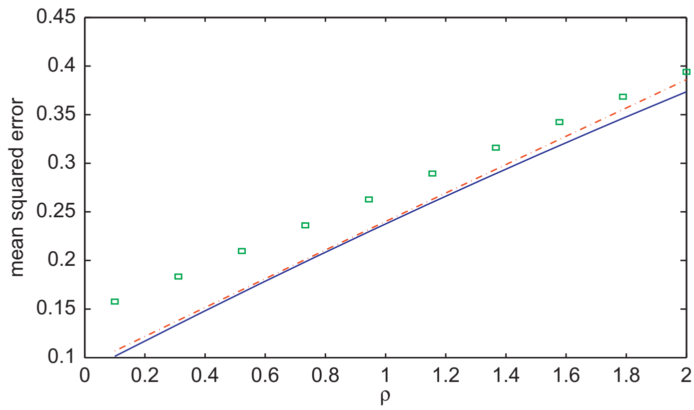

Fig. 5. Plot of mean squared errors of $\tilde{X}$ (continuous line), $\overline{X}$ (dash-dotted line), and $\tilde{X}_0$ (line marked by squares).

confidence region based on $\tilde{X}$ exhibits the best performance. The sample mean provides a good alternative, but $\tilde{X}$ based on (4), and $\tilde{X}_0$ are unsatisfactory.

Fig. 5 displays the mean squared errors of these estimators (i.e., the sum of squares of coordinatewise errors). This characteristic is quite close for $\tilde{X}$ and $\overline{X}$, but the Graybill–Deal estimator $\tilde{X}_0$ performs very poorly. A similar pattern was observed in other simulations for different error variance distributions and design matrices.

References

Beran, R., 1995. Prediction in random coefficient regression. Journal of Statistical Planning and Inference 43, 205–213.

Bliese, P.D., Ployhart, R.E., 2002. Growth modeling using random coefficient models: model building, testing and illustrations. Organizational Research Methods 5, 362–387.

Böhning, D., Malzahn, U., Dietz, E., Schlattmann, P., Viwatwongkasem, C., Biggeri, A., 2002. Some general points in estimating heterogeneity variance with the DerSimonian–Laird estimator. Biostatistics 3, 445–457.

Demidenko, E., 2005. Mixed Models. Wiley, New York.

DerSimonian, R., Laird, N., 1986. Meta-analysis in clinical trials. Controlled Clinical Trials 7, 177–188.

Evans, D., Hornikova, A., Leigh, S., Rukhin, A.L., Strawderman, W., 2009. Report on Key Comparison SIM.AUV.V-K1, Metrologia 46 Technical Supplement 09002.

Follmann, D.A., Proshan, M.A., 1999. Valid inference in random effects meta-analysis. Biometrics 55, 732–737.

Graybill, F.A., Deal, R.D., 1959. Combining unbiased estimators. Biometrics 15, 543–550.

Hall, W.J., 2007. Efficiency of weighted averages. Journal of Statistical Planning and Inference 137, 3548–3556.

Harville, D.A., 1997. Matrix Algebra from a Statistician's Perspective. Springer, New York.

Horn, R.A., Horn, S.A., Duncan, D.B., 1975. Estimating heteroscedastic variance in linear models. Journal of the American Statistical Association 70, 380–385.

Jackson, D., Bowden, J., Baker, R., 2010. How does the DerSimonian and Laird procedure for random effects meta-analysis compare with its more efficient but harder to compute counterparts? Journal of Statistical Planning and Inference 140, 961–970

Jiang, J., 2007. Linear and Generalized Linear Mixed Models and their Applications. Springer, New York.

Laird, N., Ware, J.H., 1982. Random-effects models for longitudinal data. Biometrics 38, 963–974.

Longford, N., 1994. Random Coefficient Models. Oxford University Press, UK.

Mandel, J., 1964. Estimation of weighting factors in linear regression and analysis of variance. Technometrics 6, 1–25.

McCulloch, C.E., Searle, S., 2001. Generalized, Linear and Mixed Models. Wiley, New York.

Pan, J.X., Fang, K.T., 2002. Growth Curve Models and Statistical Diagnostics. Springer, New York.

Paule, R.C., Mandel, J. 1971. Analysis of interlaboratory measurements on the vapor pressure of cadmium and silver. National Bureau of Standards Special Publication, 260-21.
Pinheiro, J., Bates, D., 2000. Mixed Effects Models in S and S-Plus. Springer, New York.
Robinson, G.K., 1991. That BLUP is a good thing: the estimation of random effects. Statistical Science 6, 15–32.
Rukhin, A.L., 2007. Estimating common vector mean in interlaboratory studies. Journal of Multivariate Analysis 98, 434–435.
Searle, S., 1997. Built-in-restrictions on best linear predictors (BLUP) of random effects in mixed models. American Statistician 51, 19–21.
Serridge, M., Licht, T.R., 1987. Piezoelectric Accelerometer and Vibration Preamplifier Handbook. Larsen&Son, Denmark.