

AN EMPIRICAL BAYES DERIVATION OF BEST LINEAR UNBIASED PREDICTORS

ROHANA J. KARUNAMUNI

Received 8 October 2001

Let $(Y_1, \theta_1), \dots, (Y_n, \theta_n)$ be independent real-valued random vectors with Y_i , given θ_i , is distributed according to a distribution depending only on θ_i for $i = 1, \dots, n$. In this paper, best linear unbiased predictors (BLUPs) of the θ_i 's are investigated. We show that BLUPs of θ_i 's do not exist in certain situations. Furthermore, we present a general empirical Bayes technique for deriving BLUPs.

2000 Mathematics Subject Classification: 62C12.

1. Introduction. Let $(Y_1, \theta_1), \dots, (Y_n, \theta_n)$ be independent real-valued random vectors satisfying the following:

- (i) conditional on θ_i , Y_i is distributed according to a distribution depending only on θ_i , $E(Y_i | \theta_i) = \theta_i$, and $\text{Var}(Y_i | \theta_i) = \mu_2(\theta_i)$ with independence over θ_i , $i = 1, \dots, n$;
- (ii) θ_i 's are independent with $E(\theta_i) = \mu_i$ and $\text{Var}(\theta_i) = \tau^2$, $i = 1, \dots, n$, where μ_i 's and τ are fixed numbers;
- (iii) $0 < D_i = E\mu_2(\theta_i) < \infty$, $i = 1, \dots, n$, where D_i 's are fixed numbers.

A special case of the above model is the so-called mixed linear model given by

$$y_i = \mathbf{x}_i' \boldsymbol{\beta} + \mathbf{u}_i + e_i, \quad i = 1, \dots, n, \quad (1.1)$$

where $\mathbf{x}_i = (x_{i1}, \dots, x_{ik})'$, $\boldsymbol{\beta}$ is a k vector of unknown parameters, sampling errors e_i and the random effects \mathbf{u}_i are independently distributed with $E(e_i) = 0$, $E(\mathbf{u}_i) = 0$, $\text{Var}(e_i) = D_i$, and $\text{Var}(\mathbf{u}_i) = \tau^2$, $i = 1, \dots, n$. The mixed linear model can also be written as

$$y_i = \theta_i + e_i, \quad \theta_i = \mathbf{x}_i' \boldsymbol{\beta} + \mathbf{u}_i, \quad i = 1, \dots, n. \quad (1.2)$$

The aim of this paper is about the best linear unbiased predictors (BLUPs) of θ_i , $i = 1, \dots, n$. BLUPs are estimates of the realized value of the random variable θ_i and are *linear* in the sense that they are linear functions of the data, y_i ; *unbiased* in the sense that the average value of the estimate is equal to the average of the quantity being estimated; *best* in the sense that they have minimum mean squared error within the class of linear unbiased estimators; and *predictors* to distinguish them from the estimators.

The first derivation of BLUPs seems to have been given by Henderson [8] who studied a more general version of the mixed linear model, namely, $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$, where \mathbf{Z} is a known design matrix, while \mathbf{e} is a vector of errors which is uncorrelated with random vector \mathbf{u} . Henderson has described BLUPs as being “joint maximum likelihood estimates,” and has assumed that \mathbf{u} and \mathbf{e} are normally distributed. A number of other derivations have been given over the years. Within the classical school of thought, BLUPs have been shown to have minimum mean squared error within the class of linear unbiased estimators (see Henderson [9] and Harville [5, 6]); in the Bayesian framework, BLUPs have been derived as the posterior mean of the parameter of interest with a noninformative prior for $\boldsymbol{\beta}$ (see Dempfle [2], and Lindley and Smith [13]); and another derivation of BLUPs has been given by Jiang [10] showing a connection between BLUP and restricted maximum likelihood. In an interesting review article, Robinson [14] has given a wide-ranging account of BLUPs in the mixed model with examples and applications. However, his discussion on empirical Bayes methods and their connection with BLUPs is rather limited—he only states that BLUPs are equivalent to one of the techniques of parametric empirical Bayes methodology, see Robinson [14, Section 5.7]. Commenting on Robinson’s paper, Harville [7] has demonstrated a connection between BLUP and empirical Bayes estimators for a one-way random effects model given by $y_{ij} = \mu + a_i + e_{ij}$ ($i = 1, \dots, n$; $j = 1, \dots, J_i$). The purpose of this paper is to investigate the connection between BLUPs and empirical Bayes estimators more closely. In this paper, our discussion is focused on the general model described by (i), (ii), and (iii) at the beginning of the introduction. We first show the nonexistence of BLUPs under certain situations, and then we present a general empirical Bayes technique for deriving BLUPs. Briefly, our claim is as follows: for $i = 1, \dots, n$, suppose $\delta_i(Y_i, \mu_i)$ denotes the linear Bayes estimator of θ_i with respect to squared error loss, and suppose that $\hat{\mu}_i$ denotes the best linear unbiased estimator (BLUE) of μ_i based on Y_1, \dots, Y_n . Then $\hat{\delta}_i(Y_i, \hat{\mu}_i)$ is the BLUP of θ_i , $i = 1, \dots, n$, whenever BLUPs exist. This general argument is in line with Robinson’s statement that BLUPs are equivalent to one of the techniques of parametric empirical Bayes methods. Our argument, however, gives a clear process of derivation and would be quite useful when deriving BLUPs in applications.

The main results of the paper are given in [Section 2](#). The proofs are deferred to [Section 3](#).

2. Main results. Throughout this section, we assume that $(Y_1, \theta_1), \dots, (Y_n, \theta_n)$ are independent random vectors, where the Y_i ’s are observed whereas the θ_i ’s are not. We first derive expressions for BLUPs of θ_i , $i = 1, \dots, n$, under various conditions on the prior parameters of the θ_i ’s. Recall that $\hat{\theta}_i$ is called a BLUP of θ_i under squared error loss if $E(\hat{\theta}_i - \theta_i) = 0$, $\hat{\theta}_i$ is a linear combination of the observations (Y_1, \dots, Y_n) , and $\hat{\theta}_i$ has the minimum mean squared error $E(\hat{\theta}_i - \theta_i)^2$, where E denotes expectation with respect to all the random variables involved; see, for example, Searle et al. [15, Chapter 7]. In other words, we investigate predictors of the form $\hat{\theta}_i = \sum_{j=1}^n c_{ij} Y_j$, where c_{ij} (≥ 0) are some constants. In order for $E(\hat{\theta}_i - \theta_i) = 0$ to be satisfied, it is required that c_{ij} satisfy the restriction that $\sum_{j=1}^n c_{ij} \mu_j = \mu_i$, $i = 1, \dots, n$, where $\mu_k = E(\theta_k)$, $k = 1, \dots, n$. When the θ_i ’s are i.i.d. then the preceding condition reduces to $\sum_{j=1}^n c_{ij} = 1$, $i = 1, \dots, n$.

THEOREM 2.1. Let $(Y_1, \theta_1), \dots, (Y_n, \theta_n)$ be independent real-valued random vectors satisfying the following:

- (i) conditional on θ_i , Y_i is distributed according to a distribution depending only on θ_i and $E(Y_i | \theta_i) = \theta_i$, and $\text{Var}(Y_i | \theta_i) = \mu_2(\theta_i)$ with independence over $i = 1, \dots, n$;
- (ii) θ_i 's are independent with $E(\theta_i) = \mu_i$ and $\text{Var}(\theta_i) = \tau^2$ for $i = 1, \dots, n$, where μ_i 's and τ are fixed numbers;
- (iii) $0 < D_i = E\mu_2(\theta_i) < \infty$, $i = 1, \dots, n$, where D_i 's are fixed numbers.

Let c_{ij} (≥ 0) be constants such that $\sum_{j=1}^n c_{ij}\mu_j = \mu_i$, $i = 1, \dots, n$. Then the mean squared error (Bayes risk) of $\hat{\theta}_i = \sum_{j=1}^n c_{ij}Y_j$ as an estimator of θ_i is given by

$$r_i(\mathbf{c}) = E \left(\sum_{j=1}^n c_{ij}Y_j - \theta_i \right)^2 = \sum_{j=1}^n c_{ij}^2 D_j + \tau^2 \left(\sum_{j=1}^n c_{ij}^2 + 1 - 2c_{ii} \right). \quad (2.1)$$

The values of c_{ij} that minimize $r_i(\mathbf{c})$ subject to the restriction $\sum_{j=1}^n c_{ij}\mu_j = \mu_i$ is c_{ij}^* such that for $i = 1, \dots, n$,

$$c_{ij}^* = \frac{\rho \mu_j}{2(D_j + \tau^2)}, \quad j = 1, \dots, i-1, i+1, \dots, n; \quad c_{ii}^* = \frac{2\tau^2 + \rho \mu_i}{2(D_i + \tau^2)}, \quad (2.2)$$

where $\rho = \mu_i [D_i(D_i + \tau^2)^{-1}] / \sum_{j=1}^n \mu_j^2 [2(D_j + \tau^2)]^{-1}$.

Note that the BLUP $\sum_{j=1}^n c_{ij}^* Y_j$ in [Theorem 2.1](#) depends on the μ_i 's as well. Thus, if μ_i 's are completely unknown then there is no BLUP for θ_i . However, if μ_i 's are partially known then a BLUP can be developed as a function of the X_j 's alone, provided that D_i 's and τ^2 are all known. These possibilities are discussed in the following corollaries.

COROLLARY 2.2. Assume that the conditions of [Theorem 2.1](#) hold. Further, suppose that $\mu_i = \mu$ for $i \geq 1$. Then $\sum_{j=1}^n c_{ij}^* Y_j$ reduces to

$$\left(\frac{D_i}{D_i + \tau^2} \right) \frac{\sum_{j=1}^n Y_j (D_j + \tau^2)^{-1}}{\sum_{j=1}^n (D_j + \tau^2)^{-1}} + \left(\frac{\tau^2}{D_i + \tau^2} \right) Y_i. \quad (2.3)$$

COROLLARY 2.3. Assume that the conditions of [Theorem 2.1](#) hold. Further, suppose that $\mu_i = t_i \beta$, where t_i are some known constants, $i = 1, \dots, n$, and β is an unknown parameter. Then $\sum_{j=1}^n c_{ij}^* Y_j$ reduces to

$$\left(\frac{D_i}{D_i + \tau^2} \right) t_i \hat{\beta} + \left(\frac{\tau^2}{D_i + \tau^2} \right) Y_i, \quad (2.4)$$

where $\hat{\beta} = [\sum_{j=1}^n (D_j + \tau^2)^{-1} t_j^2]^{-1} [\sum_{j=1}^n (D_j + \tau^2)^{-1} t_j X_j]$.

COROLLARY 2.4. Assume that the conditions of [Theorem 2.1](#) hold. Further, suppose that $\mu_i = \mathbf{x}'_i \boldsymbol{\beta}$, where $\mathbf{x}_i = (x_{i1}, \dots, x_{ik})'$ are known vectors and $\boldsymbol{\beta}$ is a k vector of unknown parameters, $i = 1, \dots, n$. Then $\sum_{j=1}^n c_{ij}^* Y_j$ reduces to

$$\left(\frac{D_i}{D_i + \tau^2} \right) \mathbf{x}'_i \hat{\boldsymbol{\beta}} + \left(\frac{\tau^2}{D_i + \tau^2} \right) Y_i, \quad (2.5)$$

where $\hat{\boldsymbol{\beta}} = [\sum_{j=1}^n (D_j + \tau^2)^{-1} \mathbf{x}'_j \mathbf{x}_j]^{-1} [\sum_{j=1}^n (D_j + \tau^2)^{-1} \mathbf{x}_j Y_j]$.

We now derive BLUPs using an empirical Bayes technique. Let $\delta_i(Y_i, \mu_i) = a_i^* Y_i + b_i^*$ denote the linear Bayes estimator of θ_i with respect to squared error loss based on the observation Y_i , $i = 1, \dots, n$. Then a_i^* and b_i^* can be obtained by solving the equations $\partial \Delta_i / \partial a_i = 0$ and $\partial \Delta_i / \partial b_i = 0$, where $\Delta_i = E(\theta_i - a_i Y_i - b_i)^2$, $i = 1, \dots, n$.

First, assume that the conditions of [Theorem 2.1](#) hold. Then observe that

$$\frac{\partial \Delta_i}{\partial a_i} = -2E(\theta_i - a_i Y_i - b_i)(Y_i), \quad \frac{\partial \Delta_i}{\partial b_i} = -2E(\theta_i - a_i Y_i - b_i). \quad (2.6)$$

By setting $\partial \Delta_i / \partial b_i = 0$, we obtain $b_i^* = \mu_i D_i (D_i + \tau^2)^{-1}$. Now substituting the preceding value in $\partial \Delta_i / \partial a_i = 0$, we obtain $a_i^* = \tau^2 (\tau^2 + D_i)^{-1}$. Thus, the Bayes estimator $\delta_i(Y_i, \mu_i) = a_i^* Y_i + b_i^*$ is given by

$$\delta_i(Y_i, \mu_i) = \tau^2 (\tau^2 + D_i)^{-1} Y_i + \mu_i D_i (D_i + \tau^2)^{-1}. \quad (2.7)$$

Suppose now that $\mu_i = \mu$ for all $i \geq 1$. Then, the BLUE of the common μ is given by

$$\hat{\mu} = \frac{\sum_{j=1}^n Y_j (D_j + \tau^2)^{-1}}{\sum_{j=1}^n (D_j + \tau^2)^{-1}}. \quad (2.8)$$

Substituting $\hat{\mu}$ for μ in (2.7) yields the empirical Bayes estimator

$$\hat{\delta}_i = \tau^2 (\tau^2 + D_i)^{-1} Y_i + D_i (\tau^2 + D_i)^{-1} \frac{\sum_{j=1}^n Y_j (D_j + \tau^2)^{-1}}{\sum_{j=1}^n (D_j + \tau^2)^{-1}}. \quad (2.9)$$

Note that (2.9) is the same as the BLUP (2.3) derived in [Corollary 2.2](#).

Suppose now that $\mu_i = t_i \beta$, where t_i is some known constant, $i = 1, \dots, n$, and β is an unknown parameter. Then, the BLUE of μ_i is given by

$$\hat{\mu}_i = t_i \hat{\beta}, \quad (2.10)$$

where

$$\hat{\beta} = \frac{\sum_{j=1}^n t_j Y_j (D_j + \tau^2)^{-1}}{\sum_{j=1}^n t_j^2 (D_j + \tau^2)^{-1}} \quad (2.11)$$

is the weighted least squares estimator of β . Substituting $\hat{\mu}_i$ of (2.10) for μ_i in (2.7) yields the empirical Bayes estimator

$$\hat{\delta}_i = \tau^2(\tau^2 + D_i)^{-1}Y_i + D_i(\tau^2 + D_i)^{-1}t_i\hat{\beta}, \quad (2.12)$$

which is the same as the BLUP (2.4) given in Corollary 2.3.

Finally, suppose that $\mu_i = \mathbf{x}_i'\beta$, where \mathbf{x}_i and β are as defined in Corollary 2.4. Then, the BLUE of μ_i is given by

$$\hat{\mu}_i = \mathbf{x}_i'\hat{\beta}, \quad (2.13)$$

where

$$\hat{\beta} = \left[\sum_{j=1}^n (D_j + \tau^2)^{-1} \mathbf{x}_j' \mathbf{x}_j \right]^{-1} \left[\sum_{j=1}^n (D_j + \tau^2)^{-1} \mathbf{x}_j' Y_j \right], \quad (2.14)$$

the weighted least squares estimator of β . Substituting $\hat{\mu}_i$ of (2.13) for μ_i in (2.7) yields the empirical Bayes estimator

$$\hat{\delta}_i = \tau^2(\tau^2 + D_i)^{-1}Y_i + D_i(\tau^2 + D_i)^{-1}\mathbf{x}_i'\hat{\beta} = \mathbf{x}_i'\hat{\beta} + y_i(Y_i - \mathbf{x}_i'\hat{\beta}), \quad (2.15)$$

where $y_i = \tau^2(D_i + \tau^2)^{-1}$. Again, observe that (2.15) is the same as the BLUP (2.5) derived in Corollary 2.4.

We now present a multivariate extension of Theorem 2.1 to the case that \mathbf{Y}_i 's and θ_i 's are real-valued random vectors. As a generalization of the univariate case, we assume that $\theta_1, \dots, \theta_n$ are i.i.d. $p \times 1$ random vectors with a common distribution G having a second moment. For given θ_i , assume that \mathbf{Y}_i have a distribution F_{θ_i} , which also has a finite second moment. For a given estimator $\delta(\mathbf{y})$, the mean squared error of $\delta(\mathbf{y})$ is defined as

$$R(\delta) = \text{tr}[E(\delta(\mathbf{Y}) - \theta)(\delta(\mathbf{Y}) - \theta)'], \quad (2.16)$$

where $\text{tr}(\cdot)$ denotes the trace of the corresponding matrix and expectation E is with respect to all the random variables involved.

THEOREM 2.5. *Let $(\mathbf{Y}_1, \theta_1), \dots, (\mathbf{Y}_n, \theta_n)$ be independent random vector pairs satisfying the following:*

- (i) $\theta_1, \dots, \theta_n$ are i.i.d. according to G with $D(\theta_i) = \nabla < \infty$ for $i = 1, \dots, n$;
- (ii) \mathbf{Y}_i , given θ_i , is distributed according to F_{θ_i} with $D(\mathbf{Y}_i | \theta_i) < \infty$ for $i = 1, \dots, n$;
- (iii) $E(\mathbf{Y}_i | \theta_i) = \theta_i$ for $i = 1, \dots, n$,

where $D(\cdot)$ stands for dispersion matrix (i.e., variance-covariance matrix) of the corresponding vector. Then, as an estimator of θ_i , the mean squared error of the estimator $\hat{\delta}_i = \sum_{j=1}^n c_{ij}Y_j$, with $\sum_{j=1}^n c_{ij} = 1$ and $c_{ij} \geq 0$, as defined by (2.16) is given by

$$R(\mathbf{c}_i, \hat{\delta}_i) = \sum_{j=1, j \neq i}^n c_{ij}^2 \text{tr}[E(D(\mathbf{Y}_j | \theta_i)) + \nabla] + c_{ii}^2 \text{tr}[E(D(\mathbf{Y}_i | \theta_i))] + (1 - c_{ii})^2 \text{tr}(\nabla), \quad (2.17)$$

where $i = 1, \dots, n$. Denote $a_j = \text{tr}[E(D(\mathbf{Y}_j | \theta_j)) + \nabla]$, $j = 1, \dots, i-1, i+1, \dots, n$; $a_i = \text{tr}[E(D(\mathbf{Y}_i | \theta_i))]$ and $V = \text{tr}(\nabla)$. Then the value of $\mathbf{c}_i = (c_{i1}, \dots, c_{in})$ which minimizes

$R(\mathbf{c}_i, \hat{\boldsymbol{\delta}}_i)$ is $\mathbf{c}_i^* = (c_{i1}^*, \dots, c_{in}^*)$ such that

$$\begin{aligned} c_{ij}^* &= \frac{a_i}{a_j \left[1 + (a_i + V) \sum_{j=1, j \neq i}^n a_i^{-1} \right]}, \quad j = 1, \dots, i-1, i+1, \dots, n; \\ c_{ii}^* &= \frac{1 + V \left(\sum_{j=1, j \neq i}^n a_j^{-1} \right)}{\left[1 + (a_n + V) \sum_{j=1, j \neq i}^n a_j^{-1} \right]}. \end{aligned} \quad (2.18)$$

We now give three examples to illustrate BLUPs obtained using the empirical Bayes (EB) method described above.

EXAMPLE 2.6 (normal hierarchy). Consider estimation of θ_i in the model $X_i | \theta_i \sim N(\theta_i, \sigma^2)$ for $i = 1, \dots, n$, independent, and $\theta_i \sim N(\mu, \tau^2)$ for $i = 1, \dots, n$, independent, where μ is unknown. Then the linear Bayes estimator of θ_i based on X_i is

$$\delta_i(X_i, \mu) = \frac{\sigma^2}{\sigma^2 + \tau^2} \mu + \frac{\tau^2}{\sigma^2 + \tau^2} X_i. \quad (2.19)$$

The BLUE of μ is $\hat{\mu} = \bar{X}$. Now, replacing μ by \bar{X} in the preceding expression yields the EB estimator

$$\delta_i(X_i, \hat{\mu}) = \frac{\sigma^2}{\sigma^2 + \tau^2} \bar{X} + \frac{\tau^2}{\sigma^2 + \tau^2} X_i. \quad (2.20)$$

It is easy to show that the preceding EB estimator of θ_i is also the BLUP (2.3) of θ_i given in Corollary 2.2 under the present setup. It is interesting to note that this EB estimator (2.20) can also be obtained as a hierarchical Bayes estimator with an additional (improper) prior, Uniform $(-\infty, \infty)$, on μ . If τ^2 is unknown we can estimate $(\sigma^2 + \tau^2)^{-1}$ by the unbiased estimator $(n-3)/\sum (X_i - \bar{X})^2$ and obtain the EB estimator (see Lindley [12]; Efron and Morris [3, 4])

$$\bar{X} + \left[1 - \frac{(n-1)\sigma^2}{2\sum (X_i - \bar{X})^2} \right] (X_i - \bar{X}), \quad (2.21)$$

which is no longer a best linear predictor; indeed, it is not even linear in X_i 's.

EXAMPLE 2.7 (Poisson hierarchy). Suppose that $X_i | \theta_i \sim \text{Poisson}(\theta_i)$, $i = 1, \dots, n$, independent, and $\theta_i \sim \text{Gamma}(\alpha, \beta)$, $i = 1, \dots, n$, independent. Then the linear Bayes estimator of θ_i based on X_i is

$$\delta_i(X_i, \mu) = \frac{\alpha\beta}{\beta+1} + \frac{\beta}{\beta+1} X_i, \quad (2.22)$$

where $\mu = E(\theta_i) = \alpha\beta$. The BLUE of μ is $\hat{\mu} = \bar{X}$. Thus, replacing $\alpha\beta$ in the equation for the Bayes estimator, we obtain the EB estimator

$$\delta_i(X_i, \hat{\mu}) = \frac{1}{\beta+1} \bar{X} + \frac{\beta}{\beta+1} X_i, \quad (2.23)$$

which is also the BLUP (2.3) of θ_i obtained in Corollary 2.2 under the present setup. Note that if β is also unknown, then there is no BLUP of θ_i .

EXAMPLE 2.8 (a regression model). Consider estimation of θ_i in the following regression model: $X_i \mid \theta_i \sim N(\theta_i, \sigma^2)$ for $i = 1, \dots, n$, independent, and $\theta_i \sim N(\alpha + \beta t_i, \tau^2)$ for $i = 1, \dots, n$, where α and β are unknown parameters. Then the linear Bayes estimator of θ_i is

$$\delta_i(X_i, \mu_i) = \frac{\sigma^2}{\sigma^2 + \tau^2} \mu_i + \frac{\tau^2}{\sigma^2 + \tau^2} X_i, \quad (2.24)$$

where $\mu_i = \alpha + \beta t_i$. Recall that the BLUEs of α and β are the least squares estimators given by

$$\hat{\alpha} = \bar{X} - \hat{\beta} \bar{t}, \quad \hat{\beta} = \frac{\sum (X_i - \bar{X})(t_i - \bar{t})}{\sum (t_i - \bar{t})^2}, \quad (2.25)$$

where $\bar{t} = n^{-1} \sum t_i$. Therefore, the BLUE of μ_i is $\hat{\mu}_i = \hat{\alpha} + \hat{\beta} t_i$. Substituting $\hat{\mu}_i$ for μ_i in the Bayes estimator yields the EB estimator

$$\delta_i(X_i, \hat{\mu}_i) = \frac{\sigma^2}{\sigma^2 + \tau^2} (\hat{\alpha} + \hat{\beta} t_i) + \frac{\tau^2}{\sigma^2 + \tau^2} X_i. \quad (2.26)$$

It is easy to see that the EB estimator (2.26) is also the BLUP (2.4) of θ_i under the present setup. Again, the EB estimator (2.26) can also be obtained as a hierarchical Bayes predictor, by appending the specification $(\alpha, \beta) \sim \text{Uniform}(-\infty, \infty) \times (-\infty, \infty)$ to the hierarchy in the present example. If τ^2 is unknown, we can use the fact that $E[\sum (X_i - \hat{\alpha} + \hat{\beta} t_i)^2]^{-1} = (n-4)(\sigma^2 + \tau^2)$ to construct the EB estimator

$$\hat{\alpha} + \hat{\beta} t_i + \left[1 - \frac{(n-4)\sigma^2}{\sum (X_i - \hat{\alpha} + \hat{\beta} t_i)^2} \right] (X_i - \hat{\alpha} - \hat{\beta} t_i), \quad (2.27)$$

which is, again, neither linear nor unbiased for θ_i .

REMARK 2.9. Generally speaking, an empirical Bayes estimator can be thought of as a two-stage estimator. Specifically, consider the Bayes model $Y_i \mid \theta \sim f(\mathbf{y} \mid \theta)$, $i = 1, \dots, n$ and $\theta \mid \mu \sim \pi(\theta \mid \mu)$, where $E(\theta) = \mu$. We, first, obtain an estimate of μ , $\hat{\mu}(\mathbf{y})$, based on the marginal distribution of $\mathbf{Y} = (Y_1, \dots, Y_n)$ with density $f(\mathbf{y} \mid \mu) = \int (\prod_{i=1}^n f(y_i \mid \theta)) \pi(\theta \mid \mu) d\theta$. Then, we substitute $\hat{\mu}(\mathbf{y})$ for μ in $\pi(\theta \mid \mu)$ and determine the estimator that minimizes the empirical posterior loss $\int L(\theta, \delta(\mathbf{y})) \pi(\theta \mid \mathbf{y}, \hat{\mu}(\mathbf{y})) d\theta$. This minimizing estimator is the empirical Bayes estimator. This argument is mathematically equivalent to first obtaining the Bayes estimator $\delta(Y, \mu)$ and then substituting μ by $\hat{\mu}(\mathbf{y})$ (see, e.g., Lehmann and Casella [11, Chapter 4]).

In the present context, our proposed empirical Bayes estimators can be obtained by minimizing the posterior loss $\int (\theta - a_i - b_i Y_i)^2 \pi(\theta \mid \mathbf{y}, \hat{\mu}(\mathbf{y})) d\theta$. Our empirical Bayes argument is somewhat equivalent to the two-stage approach suggested by Bulmer [1, pages 208–209] for the prediction problem of v in the mixed model of $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$: first form a vector of the data \mathbf{Y} corrected for the fixed effects by $\mathbf{Y}_c = \mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}$, where $\hat{\boldsymbol{\beta}}$ is the BLUE of $\boldsymbol{\beta}$, and then, under normality assumptions, v is predicted by $E(\mathbf{u} \mid \mathbf{Y}_c)$. Note that our argument, however, make no reference to normality assumptions and can be applied for more general models.

3. Proofs. In this section, we provide proofs of the results presented in [Section 2](#).

PROOF OF THEOREM 2.1. Write

$$\begin{aligned}
 r_i(\mathbf{c}) &= E \left[\sum_{j=1}^n c_{ij} X_j - \theta_i \right]^2 \\
 &= E \left[\sum_{j=1}^n c_{ij} (X_j - \theta_j) + \sum_{j=1}^n c_{ij} \theta_j - \theta_i \right]^2 \\
 &= E \left[\sum_{j=1}^n c_{ij} (X_j - \theta_j) \right]^2 + E \left[\sum_{j=1}^n c_{ij} \theta_j - \theta_i \right]^2 + CPT,
 \end{aligned} \tag{3.1}$$

where

$$\begin{aligned}
 CPT &= 2E \left[\sum_{j=1}^n c_{ij} (X_j - \theta_j) \right] \left[\sum_{j=1}^n c_{ij} \theta_j - \theta_i \right] \\
 &= 2E \left[\sum_{j=1}^n c_{ij} (\theta_j - \theta_j) \right] \left[\sum_{j=1}^n c_{ij} \theta_j - \theta_i \right] \\
 &= 0.
 \end{aligned} \tag{3.2}$$

The first term of the right-hand side of (3.1) can be evaluated as

$$\begin{aligned}
 E \left[\sum_{j=1}^n c_{ij} (X_j - \theta_j) \right]^2 &= \sum_{j=1}^n c_{ij}^2 E(X_j - \theta_j)^2 + \sum_{j \neq k} c_{ij} c_{ik} E(X_j - \theta_j)(X_k - \theta_k) \\
 &= \sum_{j=1}^n c_{ij}^2 E(X_j - \theta_j)^2 \\
 &= \sum_{j=1}^n c_{ij}^2 V(X_j - \theta_j) \\
 &= \sum_{j=1}^n c_{ij}^2 [EV((X_j - \theta_j)/\theta_j) + VE((X_j - \theta_j)/\theta_j)] \\
 &= \sum_{j=1}^n c_{ij}^2 D_j.
 \end{aligned} \tag{3.3}$$

The second term on the right-hand side of (3.1) is equal to

$$\begin{aligned}
 E \left[\sum_{j=1}^n c_{ij} \theta_j - \theta_i \right]^2 &= E \left[\sum_{j=1}^n c_{ij} \theta_j \right]^2 + E(\theta_i^2) - 2E \left(\theta_i \sum_{j=1}^n c_{ij} \theta_j \right) \\
 &= E \left[\sum_{j=1}^n c_{ij} \theta_j \right]^2 + E(\theta_i^2) - 2E \left(\theta_i \sum_{j=1}^n c_{ij} \theta_j \right).
 \end{aligned} \tag{3.4}$$

Using the independence of θ_i 's and by the restriction $\sum_{j=1}^n c_{ij}\mu_j = \mu_i$, we have

$$\begin{aligned}
 E\left(\theta_i \sum_{j=1}^n c_{ij}\theta_j\right) &= \sum_{j=1}^n c_{ij}E(\theta_j\theta_i) \\
 &= \sum_{j \neq i}^n c_{ij}E(\theta_j\theta_i) + c_{ii}E(\theta_i^2) \\
 &= \sum_{j \neq i}^n c_{ij}\mu_i\mu_j + c_{ii}(V(\theta_i) + \mu_i^2) \\
 &= (\mu_i - c_{ii}\mu_i)\mu_i + c_{ii}(\tau^2 + \mu_i^2) \\
 &= \mu_i^2 + c_{ii}\tau^2,
 \end{aligned} \tag{3.5}$$

$$\begin{aligned}
 E\left[\sum_{j=1}^n c_{ij}\theta_j\right]^2 &= V\left[\sum_{j=1}^n c_{ij}\theta_j\right] + \left[E\left(\sum_{j=1}^n c_{ij}\theta_j\right)\right]^2 \\
 &= \sum_{j=1}^n c_{ij}^2\tau^2 + \left(\sum_{j=1}^n c_{ij}\mu_j\right)^2 \\
 &= \tau^2 \sum_{j=1}^n c_{ij}^2 + \mu_i^2.
 \end{aligned} \tag{3.6}$$

Now combining (3.4) to (3.6), we obtain

$$\begin{aligned}
 E\left[\sum_{j=1}^n c_{ij}\theta_j - \theta_i\right]^2 &= \tau^2 \sum_{j=1}^n c_{ij}^2 + \mu_i^2 + (\tau^2 + \mu_i^2) - 2(\mu_i^2 + c_{ii}\tau^2) \\
 &= \tau^2 \sum_{j=1}^n c_{ij}^2 + \tau^2 - 2c_{ii}\tau^2.
 \end{aligned} \tag{3.7}$$

The proof of (2.1) is now completed by combining (3.1), (3.3), and (3.7).

In order to find the values of c_{ij} that minimize $r_i(\mathbf{c})$ subject to the restriction $\sum_{j=1}^n c_{ij}\mu_j = \mu_i$, we use the Lagrange multiplier method. Write

$$r_i(\mathbf{c}) = \sum_{j=1}^n c_{ij}^2 D_j + \tau^2 \left(\sum_{j=1}^n c_{ij}^2 + 1 - 2c_{ii} \right) + \rho \left(\mu_i - \sum_{j=1}^n c_{ij}\mu_j \right), \tag{3.8}$$

and we find solutions to the equations $\partial r_i / \partial c_{ij} = 0$ subject to the condition $\sum_{j=1}^n c_{ij}\mu_j = \mu_i$. For $j \neq i$ and $j = 1, \dots, n$,

$$\begin{aligned}
 \frac{\partial r_i}{\partial c_{ij}} &= 2D_j c_{ij} + 2\tau^2 c_{ij} - \rho \mu_j, \\
 \frac{\partial r_i}{\partial c_{ii}} &= 2D_i c_{ii} + 2\tau^2 c_{ii} - 2\tau^2 - \rho \mu_i.
 \end{aligned} \tag{3.9}$$

Now $\partial r_i / \partial c_{ij} = 0$ implies for $j \neq i$ and $j = 1, \dots, n$,

$$c_{ij}^* = \frac{\rho \mu_j}{2D_j + 2\tau^2}, \tag{3.10}$$

and $\partial r_i / \partial c_{ii} = 0$ implies

$$c_{ii}^* = \frac{2\tau^2 + \rho\mu_i}{2D_i + 2\tau^2}. \quad (3.11)$$

Using these facts together with $\sum_{j=1}^n c_{ij}\mu_j = \mu_i$ implies that

$$\rho = \frac{\mu_i D_i (D_i + \tau^2)^{-1}}{\sum_{j=1}^n \mu_j^2 [2(D_j + \tau^2)]^{-1}}. \quad (3.12)$$

The proof is now completed by combining (3.10), (3.11), and (3.12). \square

PROOF OF COROLLARY 2.2. When $\mu_j = \mu$ for all $j \geq 1$, the restriction $\sum_{j=1}^n c_{ij}\mu_j = \mu_i$ reduces to $\sum_{j=1}^n c_{ij} = 1$ for all $i \geq 1$. We now minimize

$$r_i(\mathbf{c}) = \sum_{j=1}^n c_{ij}^2 D_j + \tau^2 \left(\sum_{j=1}^n c_{ij}^2 + 1 - 2c_{ii} \right) + \rho_0 \left(1 - \sum_{j=1}^n c_{ij} \right), \quad (3.13)$$

with respect to c_{ij} subject to $\sum_{j=1}^n c_{ij} = 1$. The solutions of $\partial r_i / \partial c_{ij} = 0$ are

$$c_{ij}^* = \frac{\rho_0}{2D_j + 2\tau^2}, \quad \text{for } j \neq i; \quad c_{ii}^* = \frac{2\tau^2 + \rho_0}{2D_i + 2\tau^2}, \quad (3.14)$$

where

$$\rho_0 = \frac{D_i (D_i + \tau^2)^{-1}}{\sum_{j=1}^n [2(D_j + \tau^2)]^{-1}}. \quad (3.15)$$

Now, by rearranging the terms in $\sum c_{ij}^* Y_j$ with the above choices of the c_{ij}^* follows the desired result. \square

PROOF OF COROLLARY 2.3. We minimize

$$r_i(\mathbf{c}) = \sum_{j=1}^n c_{ij}^2 D_j + \tau^2 \left(\sum_{j=1}^n c_{ij}^2 + 1 - 2c_{ii} \right) + \rho_1 \left(t_i - \sum_{j=1}^n c_{ij} t_j \right), \quad (3.16)$$

with respect to c_{ij} subject to $\sum_{j=1}^n c_{ij} t_j = t_i$. The rest of the proof is now similar to the proof of Theorem 2.1. \square

PROOF OF COROLLARY 2.4. Similar to the proof of Corollary 2.3. \square

PROOF OF THEOREM 2.5. First, we write

$$\begin{aligned} R(\mathbf{c}_i, \hat{\boldsymbol{\delta}}_i) &= \text{tr} \left[E \left(\left(\sum_{j=1}^n c_{ij} (\mathbf{Y}_j - \boldsymbol{\theta}_i) \right) \left(\sum_{j=1}^n c_{ij} (\mathbf{Y}_j - \boldsymbol{\theta}_i) \right)' \right) \right] \\ &= \sum_{j=1}^n c_{ij}^2 \text{tr} [E((\mathbf{Y}_j - \boldsymbol{\theta}_i)(\mathbf{Y}_j - \boldsymbol{\theta}_i)')] + \sum_{j \neq k} c_{ij} c_{ik} \text{tr} [E((\mathbf{Y}_j - \boldsymbol{\theta}_i)(\mathbf{Y}_k - \boldsymbol{\theta}_i)')]. \end{aligned} \quad (3.17)$$

Note that

$$\begin{aligned} E((\mathbf{Y}_j - \boldsymbol{\theta}_i)(\mathbf{Y}_j - \boldsymbol{\theta}_i)') &= E((\mathbf{Y}_j - \boldsymbol{\theta}_j)(\mathbf{Y}_j - \boldsymbol{\theta}_j)') + E((\boldsymbol{\theta}_j - \boldsymbol{\theta}_i)(\boldsymbol{\theta}_j - \boldsymbol{\theta}_i)') \\ &\quad + E((\boldsymbol{\theta}_j - \boldsymbol{\theta}_i)(\mathbf{Y}_j - \boldsymbol{\theta}_i)') + E((\mathbf{Y}_j - \boldsymbol{\theta}_j)(\boldsymbol{\theta}_j - \boldsymbol{\theta}_i)') \\ &= E(D(\mathbf{Y}_j | \boldsymbol{\theta}_j)) + D(\boldsymbol{\theta}_j - \boldsymbol{\theta}_i) \\ &= \begin{cases} E(D(\mathbf{Y}_j | \boldsymbol{\theta}_j)) + 2\boldsymbol{\nabla}, & \text{for } j \neq i, \\ E(D(\mathbf{Y}_i | \boldsymbol{\theta}_i)), & \text{for } j = i. \end{cases} \end{aligned} \quad (3.18)$$

Also, for $1 \leq j \neq k \leq n-1$, we have

$$\begin{aligned} E((\mathbf{Y}_j - \boldsymbol{\theta}_i)(\mathbf{Y}_k - \boldsymbol{\theta}_i)') &= E[E((\mathbf{Y}_j - \boldsymbol{\theta}_i)(\mathbf{Y}_k - \boldsymbol{\theta}_i)' | \boldsymbol{\theta}_j, \boldsymbol{\theta}_k)] \\ &= E((\boldsymbol{\theta}_j - \boldsymbol{\theta}_i)(\boldsymbol{\theta}_k - \boldsymbol{\theta}_i)') \\ &= E[E((\boldsymbol{\theta}_j - \boldsymbol{\theta}_i)(\boldsymbol{\theta}_k - \boldsymbol{\theta}_i)' | \boldsymbol{\theta}_i)] \\ &= D(\boldsymbol{\theta}_i) \\ &= \boldsymbol{\nabla}. \end{aligned} \quad (3.19)$$

For all other cases, that is, for $j \neq k$ and either $j = n$ or $k = n$, we have

$$\begin{aligned} E((\mathbf{Y}_j - \boldsymbol{\theta}_j)(\boldsymbol{\theta}_j - \boldsymbol{\theta}_i)') &= E[E((\mathbf{Y}_j - \boldsymbol{\theta}_j)(\boldsymbol{\theta}_j - \boldsymbol{\theta}_i)' | \boldsymbol{\theta}_j, \boldsymbol{\theta}_i)] \\ &= \mathbf{0}. \end{aligned} \quad (3.20)$$

Now, combining (3.18), (3.19), and (3.20) and then substituting in (3.17), we obtain after some simplification $R(\mathbf{c}_i, \hat{\boldsymbol{\delta}}_i) = \sum_{j=1, j \neq i}^n c_{ij}^2 \text{tr}[E(D(\mathbf{Y}_j | \boldsymbol{\theta}_j)) + \boldsymbol{\nabla}] + c_{ii}^2 \text{tr}[E(D(\mathbf{Y}_i | \boldsymbol{\theta}_i))] + (1 - c_{ii})^2 \text{tr}(\boldsymbol{\nabla})$. This proves (2.17). The derivation of (2.18) easily follows from (2.17) subject to the restriction that $\sum_{j=1}^n c_{ij} = 1$. This completes the proof of Theorem 2.5. \square

REFERENCES

- [1] M. G. Bulmer, *The Mathematical Theory of Quantitative Genetics*, Oxford University Press, New York, 1980.
- [2] L. Dempfle, *Realtion entre BLUP (best linear unbiased predictor) et estimateurs Bayesiens*, Ann. Genetique Selection Animale **9** (1977), 27-32.
- [3] B. Efron and C. Morris, *Limiting the risk of Bayes and empirical Bayes estimators. II. The empirical Bayes case*, J. Amer. Statist. Assoc. **67** (1972), 130-139.
- [4] ———, *Stein's estimation rule and its competitors—an empirical Bayes approach*, J. Amer. Statist. Assoc. **68** (1973), 117-130.

- [5] D. Harville, *Extension of the Gauss-Markov theorem to include the estimation of random effects*, Ann. Statist. **4** (1976), no. 2, 384-395.
- [6] ———, *BLUP (best linear unbiased prediction) and beyond*, Advances in Statistical Methods for Genetic Improvement of Livestock (D. Gianola and K. Hammond, eds.), Springer, New York, 1990, pp. 239-276.
- [7] ———, *Comment of the paper by Robinson*, Statist. Sci. **6** (1991), 35-39.
- [8] C. R. Henderson, *Estimation of genetic parameters (abstract)*, Ann. Math. Statist. **21** (1950), 309-310.
- [9] ———, *Selection index and expected genetic advance*, Statistical Genetics and Plant Breeding, Nat. Acad. Sci. Nat. Res., no. 982, Council Publication, Washington, DC, 1963, pp. 141-163.
- [10] J. Jiang, *A derivation of BLUP—best linear unbiased predictor*, Statist. Probab. Lett. **32** (1997), no. 3, 321-324.
- [11] E. L. Lehmann and G. Casella, *Theory of Point Estimation*, Springer-Verlag, New York, 1998.
- [12] D. V. Lindley, *Discussion of the paper by Stein*, J. Roy. Statist. Soc. Ser. B **24** (1962), 265-296.
- [13] D. V. Lindley and A. F. M. Smith, *Bayes estimates for the linear model*, J. Roy. Statist. Soc. Ser. B **34** (1972), 1-41.
- [14] G. K. Robinson, *That BLUP is a good thing: the estimation of random effects*, Statist. Sci. **6** (1991), no. 1, 15-51.
- [15] S. R. Searle, G. Casella, and C. E. McCulloch, *Variance Components*, John Wiley & Sons, New York, 1992.

ROHANA J. KARUNAMUNI: DEPARTMENT OF MATHEMATICAL AND STATISTICAL SCIENCES, UNIVERSITY OF ALBERTA, EDMONTON, ALBERTA, CANADA T6G 2G1
E-mail address: rkurunam@ualberta.ca

Special Issue on Modeling Experimental Nonlinear Dynamics and Chaotic Scenarios

Call for Papers

Thinking about nonlinearity in engineering areas, up to the 70s, was focused on intentionally built nonlinear parts in order to improve the operational characteristics of a device or system. Keying, saturation, hysteretic phenomena, and dead zones were added to existing devices increasing their behavior diversity and precision. In this context, an intrinsic nonlinearity was treated just as a linear approximation, around equilibrium points.

Inspired on the rediscovering of the richness of nonlinear and chaotic phenomena, engineers started using analytical tools from “Qualitative Theory of Differential Equations,” allowing more precise analysis and synthesis, in order to produce new vital products and services. Bifurcation theory, dynamical systems and chaos started to be part of the mandatory set of tools for design engineers.

This proposed special edition of the *Mathematical Problems in Engineering* aims to provide a picture of the importance of the bifurcation theory, relating it with nonlinear and chaotic dynamics for natural and engineered systems. Ideas of how this dynamics can be captured through precisely tailored real and numerical experiments and understanding by the combination of specific tools that associate dynamical system theory and geometric tools in a very clever, sophisticated, and at the same time simple and unique analytical environment are the subject of this issue, allowing new methods to design high-precision devices and equipment.

Authors should follow the Mathematical Problems in Engineering manuscript format described at <http://www.hindawi.com/journals/mpe/>. Prospective authors should submit an electronic copy of their complete manuscript through the journal Manuscript Tracking System at <http://mts.hindawi.com/> according to the following timetable:

Manuscript Due	December 1, 2008
First Round of Reviews	March 1, 2009
Publication Date	June 1, 2009

Guest Editors

José Roberto Castilho Piqueira, Telecommunication and Control Engineering Department, Polytechnic School, The University of São Paulo, 05508-970 São Paulo, Brazil; piqueira@lac.usp.br

Elbert E. Neher Macau, Laboratório Associado de Matemática Aplicada e Computação (LAC), Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 12227-010 São Paulo, Brazil ; elbert@lac.inpe.br

Celso Grebogi, Center for Applied Dynamics Research, King's College, University of Aberdeen, Aberdeen AB24 3UE, UK; grebogi@abdn.ac.uk