Aplikace matematiky

Jaroslav Stuchlý

Bayes unbiased estimation in a model with two variance components

Aplikace matematiky, Vol. 32 (1987), No. 2, 120–130

Persistent URL: http://dml.cz/dmlcz/104241

32 (1987) APLIKACE MATEMATIKY No. 2, 120—130

BAYES UNBIASED ESTIMATION IN A MODEL WITH TWO VARIANCE COMPONENTS

JAROSLAV STUCHLÝ

(Received May 17, 1985)

Summary. In the paper an explicit expression for the Bayes invariant quadratic unbiased estimate of the linear function of the variance components is presented for the mixed linear model $\mathbf{t} = \mathbf{X}\boldsymbol{\beta} + \varepsilon$, $\mathrm{E}(\mathbf{t}) = \mathbf{X}\boldsymbol{\beta}$, $\mathrm{D}(\mathbf{t}) = \theta_1\mathbf{U}_1 + \theta_2\mathbf{U}_2$ with the unknown variance components in the normal case. The matrices $\mathbf{U}_1$, $\mathbf{U}_2$ may be singular. Applications to two examples of the analysis of variance are given.

Keywords: mixed linear model, risk function.
AMS Classification: Primary 62 J 99; Secondary 62 C 15.

1. INTRODUCTION

Consider a mixed linear model

(1) $$\mathbf{t} = \mathbf{X}\boldsymbol{\beta} + \varepsilon\,, \quad \mathrm{E}(\mathbf{t}) = \mathbf{X}\boldsymbol{\beta}\,, \quad \mathrm{E}(\varepsilon\varepsilon') = \theta_1\mathbf{U}_1 + \theta_2\mathbf{U}_2 = \mathbf{U}(\boldsymbol{\theta})\,,$$

where $\mathbf{t}$ is an N-dimensional, normally distributed random vector, $\mathbf{X}$ is a known $N \times m$ matrix of rank $\mathrm{R}(\mathbf{X}) = p$, $\boldsymbol{\beta} \in \mathrm{R}^m$ is a vector of unknown parameters, $\mathbf{U}_1,\mathbf{U}_2$ are known, nonnegative definite matrices and $\boldsymbol{\theta} = (\theta_1, \theta_2)'$ is a vector of unknown variance components $\boldsymbol{\theta} \in \mathscr{T}, \mathscr{T} = \{\boldsymbol{\theta}\colon \theta_1 > 0, \theta_2 \geqq 0\}$. .

The problem is to estimate a linear parametric function $\gamma = f_1\theta_1 + f_2\theta_2$ by a quadratic form $\hat{\gamma}(\mathbf{t}) = \mathbf{t}'\mathbf{B}\mathbf{t}$, where $\mathbf{B} \in \mathscr{S}_N$, , $\mathscr{S}_N$ is a class of symmetric $N \times N$ matrices.

We restrict our considerations to quadratic estimates $\hat{\gamma}(\mathbf{t}) = \mathbf{t}'\mathbf{B}\mathbf{t}$ which

(a) are unbiased, i.e. $\mathrm{E}_\theta(\hat{\gamma}) = \gamma$ is satisfied for all $\boldsymbol{\theta}$,

(b) are invariant with respect to translations $\mathbf{t} \to \mathbf{t} - \mathbf{X}\boldsymbol{\beta}$, i.e. $\hat{\gamma}(\mathbf{t}) = \hat{\gamma}(\mathbf{t} - \mathbf{X}\boldsymbol{\beta})$,

(c) minimize the risk function

$$\mathrm{r}(\hat{\gamma}) = \tfrac{1}{2}\int \mathrm{E}_\theta(\hat{\gamma} - \gamma)^2 \,\mathrm{dP}_\theta\,,$$

where P_θ is the prior distribution for the vector parameter $\boldsymbol{\theta}$ having second order moments, i.e.

$$\mathrm{E}(\theta_i\theta_j) = \int \theta_i\theta_j \,\mathrm{dP}_\theta = c_{ij} \geqq 0\,, \quad i, j = 1, 2\,.$$

Such quadratic forms are called the Bayes invariant quadratic unbiased estimates (BIQUE).

It is well known that a quadratic estimator $\hat{\gamma} = \mathbf{t}'\mathbf{B}\mathbf{t}$ is unbiased and invariant if and only if

$$\mathbf{B}\mathbf{X} = \mathbf{0}\,, \quad \operatorname{tr}(\mathbf{B}\mathbf{U}_i) = f_i\,, \quad i = 1, 2.$$

2. SOLUTION

Let $\mathbf{P}$ be an $(N - p) \times N$-matrix satisfying $\mathbf{P}'\mathbf{P} = \mathbf{I} - \mathbf{X}\mathbf{X}^+$, $\mathbf{P}\mathbf{P}' = \mathbf{I}$, where $\mathbf{X}^+$ is the Moore-Penrose inverse of the matrix $\mathbf{X}$ (see [5] or [2]). Consider the random $(N - p)$-vector $\mathbf{y} = \mathbf{P}\mathbf{t}$. Since $\mathbf{P}\mathbf{X} = \mathbf{0}$, the model for $\mathbf{y}$ is

$$\text{(2)} \qquad \mathbf{y} = \mathbf{P}\varepsilon\,, \quad \mathrm{E}(\mathbf{y}) = 0\,, \quad \mathrm{E}(\mathbf{y}\mathbf{y}') = \theta_1\mathbf{V}_1 + \theta_2\mathbf{V}_2 = \mathbf{V}(\theta)\,,$$

where $\mathbf{V}_1 = \mathbf{P}\mathbf{U}_1\mathbf{P}'$, $\mathbf{V}_2 = \mathbf{P}\mathbf{U}_2\mathbf{P}'$. Since $\{\mathbf{B}\mathbf{X} = \mathbf{0}$ and $\mathbf{B}' = \mathbf{B}\}$ iff $\{\mathbf{B} = \mathbf{P}'\mathbf{P}\mathbf{B}\mathbf{P}'\mathbf{P} = \mathbf{P}'\mathbf{A}\mathbf{P}$ and $\mathbf{A}' = \mathbf{A}\}$, we have $\hat{\gamma} = \mathbf{t}'\mathbf{B}\mathbf{t} = \mathbf{t}'\mathbf{P}'\mathbf{A}\mathbf{P}\mathbf{t} = \mathbf{y}'\mathbf{A}\mathbf{y}$. We see that $\hat{\gamma} = \mathbf{t}'\mathbf{B}\mathbf{t}$ is BIQUE for γ in the model (1) iff $\mathbf{B} = \mathbf{P}'\mathbf{A}\mathbf{P}$ and $\hat{\gamma} = \mathbf{y}'\mathbf{A}\mathbf{y}$ is BQUE for γ in the transformed model (2). The estimate BQUE is defined in the same way as BIQUE only the invariance condition is dropped, which is an irrelevant restriction for the model (2).

Since $\mathbf{y} \sim \mathrm{N}_n(\mathbf{0}, \mathbf{V}(\theta))$, $n = N - p$, the risk function is

$$\text{(3)} \qquad \mathrm{r}(\hat{\gamma}) = \tfrac{1}{2}\int \operatorname{var}_\theta(\mathbf{y}'\mathbf{A}\mathbf{y})\,\mathrm{d}P_\theta = \int \operatorname{tr}\left[\mathbf{A}\mathbf{V}(\theta)\ \mathbf{A}\mathbf{V}(\theta)\right]\mathrm{d}P_\theta = \sum_{i=1}^{2}\sum_{j=1}^{2} c_{ij}\operatorname{tr}(\mathbf{A}\mathbf{V}_i\mathbf{A}\mathbf{V}_j)\,.$$

Our problem is to minimize the risk function (3) uder the unbiasedness conditions

$$\text{(4)} \qquad \operatorname{tr}(\mathbf{A}\mathbf{V}_i) = f_i\,, \quad i = 1, 2\,.$$

Therefore we consider the Lagrangian function

$$\mathrm{L}(\mathbf{A}, \lambda) = \sum_{i=1}^{2}\sum_{j=1}^{2} c_{ij}\operatorname{tr}(\mathbf{A}\mathbf{V}_i\mathbf{A}\mathbf{V}_j) - 2\sum_{i=1}^{2}\lambda_i[\operatorname{tr}(\mathbf{A}\mathbf{V}_i) - f_i]\,,$$

where $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$, are Lagrangian multipliers. Using the symbol $(\partial/\partial\mathbf{A})\,\mathrm{L}(\mathbf{A}, \boldsymbol{\lambda})$ for the $n \times n$-matrix of all partial derivatives of $\mathrm{L}(\mathbf{A}, \boldsymbol{\lambda})$ with respect to n^2 elements of $\mathbf{A} \in \mathscr{S}_n$, we obtain (see [4])

$$\frac{\partial \mathrm{L}}{\partial \mathbf{A}} = \sum_{i=1}^{2}\sum_{j=1}^{2} c_{ij}(2\mathbf{T}_{ij} - \operatorname{diag}\mathbf{T}_{ij}) - 2\sum_{i=1}^{2}\lambda_i(2\mathbf{V}_i - \operatorname{diag}\mathbf{V}_i) = 0\,,$$

where $\mathbf{T}_{ij} = \mathbf{V}_i\mathbf{A}\mathbf{V}_j + \mathbf{V}_j\mathbf{A}\mathbf{V}_i$ and $\operatorname{diag}\mathbf{T}$ is the diagonal matrix which has the same diagonal elements as $\mathbf{T}$.

This equation can be rewritten in the form

$$(5) \qquad \sum_{i=1}^{2} \sum_{j=1}^{2} c_{ij} \mathbf{V}_i \mathbf{A} \mathbf{V}_j = \sum_{i=1}^{2} \lambda_i \mathbf{V}_i .$$

The BIQUE is given by the matrix $\mathbf{A}$ fulfilling (4) and (5).

Since the risk function (3) is linear in $\mathbf{C} = (c_{ij})$ we can put $c_{11} = 1$ without loss of generality. Using the notation $c_{12} = c_{21} = u$, $c_{22} = u^2 + v$ $(u \geqq 0, v \geqq 0)$, the matrix $\mathbf{C}$ takes the form

$$(6) \qquad \mathbf{C} = \begin{pmatrix} 1 & u \\ u & u^2 + v \end{pmatrix} = \begin{pmatrix} 1 \\ u \end{pmatrix} (1, u) + \begin{pmatrix} 0 & 0 \\ 0 & v \end{pmatrix} .$$

Let us first assume $u \neq 0$. We use the notation

$$\mathbf{W} = \mathbf{V}_1 + u\mathbf{V}_2 , \quad \mathbf{V} = \sqrt{(v)}\, \mathbf{V}_2 .$$

We have $\mathscr{M}(\mathbf{V}) \subset \mathscr{M}(\mathbf{W})$. Here $\mathscr{M}(\mathbf{V})$ is the vector space generated by the columns of $\mathbf{V}$. The equation (5) now has the form

$$(7) \qquad \mathbf{VAV} + \mathbf{WAW} = \lambda_1 \mathbf{V}_1 + \lambda_2 \mathbf{V}_2 .$$

Let $\mathbf{W}$ be a positive semi-definite matrix of rank $\mathrm{R}(\mathbf{W}) = \mathrm{r}$. Let λ be a constant and $\mathbf{x}$ a vector such that

$$\mathbf{Vx} = \lambda \mathbf{Wx} , \quad \mathbf{Wx} \neq \mathbf{0} .$$

Than λ is called a proper eigenvalue and $\mathbf{x}$ a proper eigenvector of $\mathbf{V}$ with respect to $\mathbf{W}$.

Lemma 1. *Let* $\mathrm{R}(\mathbf{N}'\mathbf{V}) = \mathrm{R}(\mathbf{N}'\mathbf{VN})$, $\mathbf{N} = \mathbf{W}^{\perp}(\mathscr{M}(\mathbf{W}^{\perp}) = [\mathscr{M}(\mathbf{W})]^{\perp})$, *let* $\mathbf{Q}$ *be the matrix of* $\mathbf{W}$*-ortogonal proper eigenvectors of* $\mathbf{V}$ *with respect to* $\mathbf{W}$ *and let* $\mathbf{D}_1$ *be the corresponding diagonal matrix of proper eigenvalues. Then the transformation*

$$\mathbf{T} = (\mathbf{Q}, \mathbf{N}) \begin{pmatrix} \mathbf{I}_r & \mathbf{O} \\ \mathbf{O} & \mathbf{U} \end{pmatrix} ,$$

where $\mathbf{U}$ *is a unitary matrix such that* $\mathbf{U}'\mathbf{N}'\mathbf{VNU}$ *is diagonal provides the simultaneous reduction of* $\mathbf{V}$ *and* $\mathbf{W}$ *to*

$$\mathbf{T}'\mathbf{VT} = \begin{pmatrix} \mathbf{D}_1 & \mathbf{O} \\ \mathbf{O} & \mathbf{D}_2 \end{pmatrix} , \quad \mathbf{T}'\mathbf{WT} = \begin{pmatrix} \mathbf{I}_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix} ,$$

where $\mathbf{D}_2$ *is the diagonal matrix of the eigenvalues of* $\mathbf{N}'\mathbf{VN}$.

For proof see [5], p. 126.

We have $\mathscr{M}(\mathbf{V}) \subset \mathscr{M}(\mathbf{W})$ which implies $\mathbf{N}'\mathbf{VN} = \mathbf{N}'\mathbf{V} = \mathbf{O}$ and we can take $\mathbf{U} = \mathbf{I}_{n-r}$. By Lemma 1 there exists a regular transformation matrix $\mathbf{T} = (\mathbf{Q}, \mathbf{N})$ that provides the simultaneous reduction to

$$\mathbf{T}'\mathbf{VT} = \begin{pmatrix} \mathbf{D}_1 & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix} , \quad \mathbf{T}'\mathbf{WT} = \begin{pmatrix} \mathbf{I}_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix} .$$

If we denote $\mathbf{R} = (\mathbf{R}_1, \mathbf{R}_2) = (\mathbf{T}^{-1})'$, we can write $\mathbf{R}_1'\mathbf{Q} = \mathbf{Q}'\mathbf{R}_1 = \mathbf{I}_r$ and therefore

$$\mathbf{V} = \mathbf{R}\begin{pmatrix}\mathbf{D}_1 & \mathbf{O}\\ \mathbf{O} & \mathbf{O}\end{pmatrix}\mathbf{R}' = \mathbf{R}_1\mathbf{D}_1\mathbf{R}_1',$$

$$\mathbf{W} = \mathbf{R}\begin{pmatrix}\mathbf{I}_r & \mathbf{O}\\ \mathbf{O} & \mathbf{O}\end{pmatrix}\mathbf{R}' = \mathbf{R}_1\mathbf{R}_1'.$$

Now (7) has the form

$$\mathbf{R}\begin{pmatrix}\mathbf{I}_r & \mathbf{O}\\ \mathbf{O} & \mathbf{O}\end{pmatrix}\mathbf{R}'\mathbf{A}\mathbf{R}\begin{pmatrix}\mathbf{I}_r & \mathbf{O}\\ \mathbf{O} & \mathbf{O}\end{pmatrix}\mathbf{R}' + \mathbf{R}\begin{pmatrix}\mathbf{D}_1 & \mathbf{O}\\ \mathbf{O} & \mathbf{O}\end{pmatrix}\mathbf{R}'\mathbf{A}\mathbf{R}\begin{pmatrix}\mathbf{D}_1 & \mathbf{O}\\ \mathbf{O} & \mathbf{O}\end{pmatrix}\mathbf{R}' = \lambda_1\mathbf{V}_1 + \lambda_2\mathbf{V}_2.$$

Denoting

$$\mathbf{Z} = \mathbf{R}'\mathbf{A}\mathbf{R} = \begin{pmatrix}\mathbf{Z}_{11} & \mathbf{Z}_{12}\\ \mathbf{Z}_{12}' & \mathbf{Z}_{22}\end{pmatrix}$$

we get

$$\mathbf{Z}_{11} + \mathbf{D}_1\mathbf{Z}_{11}\mathbf{D}_1 = \lambda_1(\mathbf{I}_r - 1/\sqrt{(v)}\,\mathbf{D}_1) + \lambda_2\, 1/\sqrt{(v)}\,\mathbf{D}_1,$$

and therefore

$$\mathbf{Z}_{11} = \tfrac{1}{2}\left[\lambda_1\left(\mathbf{I}_r - \frac{u}{\sqrt{v}}\mathbf{D}_1\right) + \lambda_2\frac{1}{\sqrt{v}}\mathbf{D}_1\right](\mathbf{I}_r + \mathbf{D}_1^2)^{-1} +$$

$$+ \tfrac{1}{2}(\mathbf{I}_r + \mathbf{D}_1^2)^{-1}\left[\lambda_1\left(\mathbf{I}_r - \frac{u}{\sqrt{v}}\mathbf{D}_1\right) + \lambda_2\frac{1}{\sqrt{v}}\mathbf{D}_1\right]$$

and $\mathbf{Z}_{12}$, $\mathbf{Z}_{22}$ can be arbitrary. Hence

$$\mathbf{A} = \mathbf{R}'^{-1}\mathbf{Z}\mathbf{R}^{-1} = \mathbf{Q}\mathbf{Z}_{11}\mathbf{Q}' + \mathbf{N}\mathbf{Z}_{12}'\mathbf{Q}' + \mathbf{Q}\mathbf{Z}_{12}\mathbf{N}' + \mathbf{N}\mathbf{Z}_{22}\mathbf{N}'.$$

It is easy to verify that

$$\mathbf{W}^- = \mathbf{Q}\mathbf{Q}',$$

$$(\mathbf{W} + \mathbf{V}\mathbf{W}^-\mathbf{V})^- = [\mathbf{R}_1(\mathbf{I}_r + \mathbf{D}_1^2)\,\mathbf{R}_1']^- = \mathbf{Q}(\mathbf{I}_r + \mathbf{D}_1^2)\,\mathbf{Q}'$$

and therefore

$$\mathbf{Q}\mathbf{Z}_{11}\mathbf{Q}' = \tfrac{1}{2}\,\mathbf{W}^-\left[\lambda_1\left(\mathbf{W} - \frac{u}{\sqrt{v}}\mathbf{V}\right) + \lambda_2\frac{1}{\sqrt{v}}\mathbf{V}\right](\mathbf{W} + \mathbf{V}\mathbf{W}^-\mathbf{V})^- +$$

$$+ \tfrac{1}{2}(\mathbf{W} + \mathbf{V}\mathbf{W}^-\mathbf{V})^-\left[\lambda_1\left(\mathbf{W} - \frac{u}{\sqrt{v}}\mathbf{V}\right) + \lambda_2\frac{1}{\sqrt{v}}\mathbf{V}\right]\mathbf{W}^-.$$

Hence

$$\mathbf{A} = \tfrac{1}{2}(\mathbf{V}_1 + u\mathbf{V}_2)^-(\lambda_1\mathbf{V}_1 + \lambda_2\mathbf{V}_2)\,[\mathbf{V}_1 + u\mathbf{V}_2 + v\mathbf{V}_2(\mathbf{V}_1 + u\mathbf{V}_2)^-\mathbf{V}_2]^- +$$

$$+ \tfrac{1}{2}[\mathbf{V}_1 + u\mathbf{V}_2 + v\mathbf{V}_2(\mathbf{V}_1 + u\mathbf{V}_2)^-\mathbf{V}_2]^-(\lambda_1\mathbf{V}_1 + \lambda_2\mathbf{V}_2)(\mathbf{V}_1 + u\mathbf{V}_2)^- +$$

$$+ \mathbf{N}\mathbf{Z}_{12}'\mathbf{Q}' + \mathbf{Q}\mathbf{Z}_{12}\mathbf{N}' + \mathbf{N}\mathbf{Z}_{22}\mathbf{N}',$$

where λ_1 and λ_2 satisfy the unbiasedness conditions (4). Since $\mathbf{NW} = \mathbf{O}$ and $\mathcal{M}(\mathbf{V}_1) \subset \mathcal{M}(\mathbf{W})$, $\mathcal{M}(\mathbf{V}_2) \subset \mathcal{M}(\mathbf{W})$, the value of the risk function (3) does not depend on the choise of the matrices $\mathbf{Z}_{12}$, $\mathbf{Z}_{22}$ and we can take them equal to $\mathbf{O}$. So far we have considered the prior distribution with $u \neq 0$.

Now let $u = 0$. Since $\theta_2 \geqq 0$, we have $\mathrm{P}(\theta_2 = 0) = 1$ which implies $v = 0$. The equation (7) has the form

$$\mathbf{V}_1\mathbf{A}\mathbf{V}_1 = \lambda_1\mathbf{V}_1 + \lambda_2\mathbf{V}_2 .$$

A necessary and sufficient condition for this equation to have a solution for all real $\lambda_1\lambda_2$ is $\mathcal{M}(\mathbf{V}_2) \subset \mathcal{M}(\mathbf{V}_1)$, in which case the general solution is

$$\mathbf{A} = \lambda_1\mathbf{V}_1^-\mathbf{V}_1\mathbf{V}_1^- + \lambda_2\mathbf{V}_1^-\mathbf{V}_2\mathbf{V}_1^- + \mathbf{H} - \mathbf{V}_1^-\mathbf{V}_1\mathbf{H}\mathbf{V}_1\mathbf{V}_1^-$$

where $\mathbf{H}$ is an arbitrary matrix (see Theorem 2.3.2 in [5]).

The value of the risk function (3) does not depend on the choice of the matrix $\mathbf{H}$ and we can take $\mathbf{H} = \mathbf{O}$.

The value of the risk function (3) is invariant for any choice of the g-inverses in the expressions for $\mathbf{A}$ (see Lemma 2.4.4 in [5]) and we can use the Moore-Penrose inverses.

We have established the following theorem.

Theorem 1. *For all $u \geqq 0$ let $\mathcal{M}(\mathbf{V}_2) \subset \mathcal{M}(\mathbf{V}_1 + u\mathbf{V}_2)$.*

(i) *A* BQUE *for the parametric function $\gamma = f_1\theta_1 + f_2\theta_2$ in the model (2) exists if and only if*

$$\begin{pmatrix} f_1 \\ f_2 \end{pmatrix} \in \mathcal{M}\begin{pmatrix} \operatorname{tr}(\mathbf{M}_1\mathbf{V}_1), & \operatorname{tr}(\mathbf{M}_2\mathbf{V}_1) \\ \operatorname{tr}(\mathbf{M}_1\mathbf{V}_2), & \operatorname{tr}(\mathbf{M}_2\mathbf{V}_2) \end{pmatrix},$$

where

$$\mathbf{M}_i = \tfrac{1}{2}(\mathbf{W}^+\mathbf{V}_i\mathbf{K}^+ + \mathbf{K}^+\mathbf{V}_i\mathbf{W}^+)\,, \quad i = 1, 2,,$$

$$\mathbf{W} = \mathbf{V}_1 + u\mathbf{V}_2\,, \quad \mathbf{K} = \mathbf{W} + v\mathbf{V}\mathbf{W}^+\mathbf{V}_2\,.$$

(ii) *The* BQUE *is given by*

$$\hat{\gamma} = \mathbf{y}'\mathbf{A}\mathbf{y} = \lambda_1\mathbf{y}'\mathbf{M}_1\mathbf{y} + \lambda_2\mathbf{y}'\mathbf{M}_2\mathbf{y}\,,$$

where λ_1, λ_2 satisfy the conditions

$$\lambda_1 \operatorname{tr}(\mathbf{M}_1\mathbf{V}_1) + \lambda_2 \operatorname{tr}(\mathbf{M}_2\mathbf{V}_1) = f_1\,. \tag{8}$$

$$\lambda_1 \operatorname{tr}(\mathbf{M}_1\mathbf{V}_2) + \lambda_2 \operatorname{tr}(\mathbf{M}_2\mathbf{V}_2) = f_2\,.$$

Corollary 1. *If $\mathbf{W}$ is a positive definite matrix and $\mathbf{V}_1\mathbf{V}_2 = \mathbf{V}_2\mathbf{V}_1$, than we get $\mathbf{A} = \lambda_1\mathbf{M}_1 + \lambda_2\mathbf{M}_2$ with $\mathbf{M}_i = \mathbf{V}_i(\mathbf{W}^2 + v\mathbf{V}_2^2)^{-1}$, $i = 1, 2$.*

Proof. Let $\mathbf{V}_1\mathbf{V}_2 = \mathbf{V}_2\mathbf{V}_1$ and let $\mathbf{W}$ be a positive definite matrix. Then for $i = 1, 2$ we have $\mathbf{W}(\mathbf{V}_i\mathbf{W}^{-1} - \mathbf{W}^{-1}\mathbf{V}_i)\,\mathbf{W} = \mathbf{W}\mathbf{V}_i - \mathbf{V}_i\mathbf{W} = \mathbf{O}$ which implies $\mathbf{V}_i\mathbf{W}^{-1} = \mathbf{W}^{-1}\mathbf{V}_i$. Hence $\mathbf{M}_i = \frac{1}{2}(\mathbf{V}_i\mathbf{W}^{-1}\mathbf{K}^{-1} + \mathbf{V}_i\mathbf{K}^{-1}\mathbf{W}^{-1}) = \mathbf{V}_i(\mathbf{W}^2 + v\mathbf{V}_2^2)^{-1}$.

Remark 1. The limit case $c_{11} = 0$ requires separate treatment. Alternatively it may be viewed as the limit case of (6) as v tends to infinity.

Remark 2. Using the spectral decomposition, Gnot and Kleffe [1] have derived explicit expressions for BQUE of $\gamma = f_1\theta_1 + f_2\theta_2$ in the case $\mathbf{V}_1 = \mathbf{I}$. Their results are $\gamma^{\sim} = \mathbf{y}'\mathbf{A}^{\sim}\mathbf{y}$, $\mathbf{A}^{\sim} = \lambda_1\Sigma^{-1} + \lambda_2\mathbf{V}_2\Sigma^{-1}$, $\Sigma = \mathbf{I} + 2u\mathbf{V}_2 + (u^2 + v)\,\mathbf{V}_2^2$, where λ_1 and λ_2 satisfy the conditions

$$\lambda_1 \operatorname{tr}(\Sigma^{-1}) + \lambda_2 \operatorname{tr}(\Sigma^{-1}\mathbf{V}_2) = f_1\,,$$
$$\lambda_1 \operatorname{tr}(\Sigma^{-1}\mathbf{V}_2) + \lambda_2 \operatorname{tr}(\Sigma^{-1}\mathbf{V}_2^2) = f_2\,.$$

It is easy to verify that our estimate $\hat{\gamma}$ is equal to $\gamma^{\sim}$ in the case $\mathbf{V}_1 = \mathbf{I}$.

Remark 3. In particular if $v = 0$ we get the local best estimate $\mathbf{y}'\mathbf{A}\mathbf{y}$ at the point $(\theta_1, \theta_2) = (1, u)$ with

$$\mathbf{A} = \lambda_1\mathbf{W}^+\mathbf{V}_1\mathbf{W}^+ + \lambda_2\mathbf{W}^+\mathbf{V}_2\mathbf{W}^+\,,$$

where λ_1 and λ_2 satisfy the conditions

$$\lambda_1 \operatorname{tr}(\mathbf{V}_1\mathbf{W}^+\mathbf{V}_1\mathbf{W}^+) + \lambda_2 \operatorname{tr}(\mathbf{V}_1\mathbf{W}^+\mathbf{V}_2\mathbf{W}^+) = f_1\,,$$
$$\lambda_1 \operatorname{tr}(\mathbf{V}_1\mathbf{W}^+\mathbf{V}_2\mathbf{W}^+) + \lambda_2 \operatorname{tr}(\mathbf{V}_2\mathbf{W}^+\mathbf{V}_2\mathbf{W}^+) = f_2$$

and therefore this BQUE is equal to MINQUE (see [3]) in the case that $\mathbf{W}$ is regular.

We shall now rewrite Theorem 1 for the model (1). Since for every matrix $\mathbf{A} \in \mathscr{S}_N$ we have

$$\mathbf{P}'(\mathbf{PAP}')^+\,\mathbf{P} = (\mathbf{MAM})^+\,,$$

where $\mathbf{M} = \mathbf{P}'\mathbf{P} = \mathbf{I} - \mathbf{XX}^+$, we get the following theorem.

Theorem 2. *For all $u \geqq 0$ let $\mathscr{M}(\mathbf{U}_2) \subset \mathscr{M}(\mathbf{U}_1 + u\mathbf{U}_2)$.*

(i) *A* BIQUE *for the parametric function $\gamma = f_1\theta_1 + f_2\theta_2$ in the model* (1) *exists if and only if*

$$\begin{pmatrix} f_1 \\ f_2 \end{pmatrix} \in \mathscr{M}\begin{pmatrix} \operatorname{tr}(\mathbf{MN}_1\mathbf{MU}_1), & \operatorname{tr}(\mathbf{MN}_2\mathbf{MU}_1) \\ \operatorname{tr}(\mathbf{MN}_1\mathbf{MU}_2), & \operatorname{tr}(\mathbf{MN}_2\mathbf{MU}_2) \end{pmatrix},$$

where

$$N_i = \tfrac{1}{2}[(\mathbf{MW}^*\mathbf{M})^+\,\mathbf{U}_i(\mathbf{MK}^*\mathbf{M})^+ + (\mathbf{MK}^*\mathbf{M})^+\,\mathbf{U}_i(\mathbf{MW}^*\mathbf{M})^+]\,,\quad i = 1, 2\,,$$
$$\mathbf{W}^* = \mathbf{U}_1 + u\mathbf{U}_2\,,\quad \mathbf{K}^* = \mathbf{W}^* + v\mathbf{U}_2\mathbf{W}^{*+}\mathbf{U}_2\,.$$

(ii) *The* BIQUE *is given by*

$$\mathbf{t}'\mathbf{Bt} = \lambda_1\mathbf{t}'\mathbf{MN}_1\mathbf{Mt} + \lambda_2\mathbf{t}'\mathbf{MN}_2\mathbf{Mt}\,,$$

where λ_1, λ_2 satisfy the conditions

$$\lambda_1 \operatorname{tr}(\mathbf{MN}_1\mathbf{MU}_1) + \lambda_2 \operatorname{tr}(\mathbf{MN}_2\mathbf{MU}_1) = f_1\,,$$
$$\lambda_1 \operatorname{tr}(\mathbf{MN}_1\mathbf{MU}_2) + \lambda_2 \operatorname{tr}(\mathbf{MN}_2\mathbf{MU}_2) = f_2\,.$$

3. EXAMPLES

Example 1. Let us take the one-way classification model

$$y_{ij} = \alpha_i + h_i\varepsilon_{ij}\,, \quad i, j = 1, 2, \ldots, n\,.$$

Here α_i and ε_{ij} are assumed to be independent random samples from two normal populations with zero means and variances θ_2 and θ_1, respectively, $h_1, h_2, \ldots, h_n \in \mathrm{R}^1$ are known. Suppose the vector of observations is written in the lexicographic order as

$$\mathbf{y} = (y_{11}, \ldots, y_{1n}, y_{21}, \ldots, y_{2n}, y_{31}, \ldots, y_{nn})'\,,$$

and in the same way

$$\boldsymbol{\varepsilon} = (h_1\varepsilon_{11}, \ldots, h_1\varepsilon_{1n}, h_2\varepsilon_{21}, \ldots, h_2\varepsilon_{2n}, h_3\varepsilon_{31}, \ldots, h_n\varepsilon_{nn})'\,.$$

Then we have a model

$$\mathbf{y} = \mathbf{D}\boldsymbol{\alpha} + \boldsymbol{\varepsilon}$$

with

$$\mathbf{D} = \mathscr{D}(\mathbf{1}_n)\,, \quad \operatorname{cov}(\mathbf{y}) = \theta_1\mathbf{V}_1 + \theta_2\mathbf{V}_2\,, \quad \mathbf{V}_1 = \mathscr{D}(h_i^2\mathbf{I}_n)\,, \quad \mathbf{V}_2 = \mathscr{D}(\mathbf{J}_n)\,,$$

where $\mathbf{1}_n = (1, \ldots, 1)'$, $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)'$, $\mathbf{J}_n$ is the $n \times n$ matrix with all entries equal to one and

$$\mathscr{D}(k_i\mathbf{G}_n) = \begin{pmatrix} k_1\mathbf{G}_n & \mathbf{O} & \ldots & \mathbf{O} \\ \mathbf{O} & k_2\mathbf{G}_n & \ldots & \mathbf{O} \\ \ldots & \ldots & \ldots & \ldots \\ \mathbf{O} & \mathbf{O} & \ldots & k_n\mathbf{G}_n \end{pmatrix}$$

for any $n \times n$-matrix $\mathbf{G}_n$.

We have $\mathbf{V}_1\mathbf{V}_2 = \mathbf{V}_2\mathbf{V}_1$. If $u \neq 0$ or $\mathscr{M}(\mathbf{V}_2) \subset \mathscr{M}(\mathbf{V}_1)$, according to Theorem 1 and Corollary 1 the BQUE of $\gamma = f_1\theta_1 + f_2\theta_2$ is $\hat{\gamma} = \mathbf{y}'\mathbf{A}\mathbf{y}$, where

$$\mathbf{A} = \lambda_1\mathbf{M}_1 + \lambda_2\mathbf{M}_2\,,$$

$$\mathbf{M}_1 = \mathbf{V}_1(\mathbf{W}^2 + v\mathbf{V}_2^2)^{-1}\,, \quad \mathbf{M}_2 = \mathbf{V}_2(\mathbf{W}^2 + v\mathbf{V}_2^2)^{-1}\,.$$

In our case we get

$$\mathbf{W} = \mathscr{D}\{h_i^2(\mathbf{I}_n + u_i\mathbf{J}_n)\}\,, \quad \mathbf{W}^2 + v\mathbf{V}_2^2 = \mathscr{D}\{h_i^4(\mathbf{I}_n + \phi_i\mathbf{J}_n)\}\,,$$

$$(\mathbf{W}^2 + v\mathbf{V}_2^2)^{-1} = \mathscr{D}\left\{h_i^{-4}\left(\mathbf{I}_n - \frac{\phi_i}{\omega_i}\mathbf{J}_n\right)\right\}$$

where

$$u_i = u/h_i^2\,, \quad \phi_i = 2u_i + nu_i^2 + nvh_i^{-4}\,, \quad \omega_i = 1 + n\phi_1\,,$$

so that

$$\mathbf{M}_1 = \mathscr{D}\left\{h_i^{-2}\left(I_n - \frac{\phi_i}{\omega_i}\mathbf{J}_n\right)\right\}, \quad \mathbf{M}_2 = \mathscr{D}\left\{h_i^{-4}\frac{1}{\omega_i}\mathbf{J}_n\right\}.$$

The equations (8) for λ_1, λ_2 have now the form

$$\lambda_1 n \sum_{i=1}^{n} \left(1 - \frac{\phi_i}{\omega_i}\right) + \lambda_2 n \sum_{i=1}^{n} \frac{h_i^{-2}}{\omega_i} = f_1 ,$$

$$\lambda_1 n \sum_{i=1}^{n} \frac{h_i^{-2}}{\omega_i} + \lambda_2 n^2 \sum_{i=1}^{n} \frac{h_i^{-2}}{\omega_i} = f_2 .$$

If

$$d = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{1}{\omega_i \omega_j} \left[(\omega_i - \phi_i) h_j^{-4} n - h_i^{-2} h_j^{-2}\right] \neq 0$$

the last equations have a unique solution

$$\lambda_1 = \lambda_1^* = \frac{1}{nd} \sum_{i=1}^{n} \frac{1}{\omega_i} h_i^{-2} (f_1 n h_i^{-2} - f_2) ,$$

$$\lambda_2 = \lambda_2^* = \frac{1}{nd} \sum_{i=1}^{n} \frac{1}{\omega_i} \left[(\omega_i - \phi_i) f_2 - h_i^{-2} f_1\right]$$

and the BIQUE for γ has the form

$$\hat{\gamma} = \lambda_1^* \sum_{i=1}^{n} \sum_{j=1}^{n} h_i^{-2} y_{ij}^2 +$$

$$+ \frac{1}{d} \sum_{i=1}^{n} \sum_{j=1}^{n} \{f_2[(\omega_j - \phi_j) h_i^{-4} + h_j^{-2} h_i^{-2} \phi_i] - f_1[n h_j^{-4} h_i^{-2} \phi_i + h_j^{-2} h_i^{-4}]\} \frac{n}{\omega_i \omega_j} y_{i.}^2 ,$$

where

(9) $$y_{i\cdot} = \frac{1}{n} \sum_{j=1}^{n} y_{ij} .$$

The simplest case occurs if $h_1 = h_2 = \ldots = h_n = 1$. Then

$$\hat{\gamma} = \frac{f_1 n - f_2}{n^2(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} y_{ij}^2 + \frac{f_2 - f_1}{n-1} \sum_{i=1}^{n} y_{i.}^2 .$$

and so we see that in this case BIQUE does not depend on the prior distribution.

If v tends to infinity it can be checked that

$$\lim_{v \to \infty} \hat{\gamma} = \frac{1}{n^3(n-1)} \{(f_1 n^2 - f_2 h^2) \sum_{i=1}^{n} \sum_{j=1}^{n} h_i^{-2} y_{ij}^2 + n^2 \sum_{i=1}^{n} \sum_{j=1}^{n} [f_2(n - 1 + h_j^2 h_i^{-2}] y_{i.}^2\} ,$$

where $h^2 = \sum_{i=1}^{n} h_i^2$.

Example 2. We consider the model

$$y_{ij} = \alpha_i + \beta_j , \quad i, j = 1, 2, \ldots, n ,$$

where α_i and β_j are assumed to be independent with zero means and variances θ_1 and θ_2, respectively.

Using the same method as in Example 1, we get the model

$$\mathbf{y} = (\mathbf{I}_n \otimes \mathbf{1}_n)\,\boldsymbol{\alpha} + (\mathbf{1}_n \otimes \mathbf{I}_n)\,\boldsymbol{\beta}\,,$$

$$\operatorname{cov}(\mathbf{y}) = \theta_1 \mathbf{V}_1 + \theta_2 \mathbf{V}_2 = \theta_1(\mathbf{I}_n \otimes \mathbf{J}_n) + \theta_2(\mathbf{J}_n \otimes \mathbf{I}_n)\,,$$

where the sign $\otimes$ stands for the Kronecker product of matrices.

To get BIQUE for $\gamma = f_1\theta_1 + f_2\theta_2$ by Theorem 1 we must check the Moore-Penrose inverse of the singular matrices

$$\mathbf{W} = \mathbf{V}_1 + u\mathbf{V}_2 = (\mathbf{I}_n \otimes \mathbf{J}_n) + u(\mathbf{J}_n \otimes \mathbf{I}_n)\,,$$

$$\mathbf{K} = \mathbf{W} + v\mathbf{V}_2 \mathbf{W}^+ \mathbf{V}_2\,.$$

We get

$$\mathbf{W}^+ = \frac{1}{n^2 u}\left[u(\mathbf{I}_n \otimes \mathbf{J}_n) + (\mathbf{J}_n \otimes \mathbf{I}_n) - \frac{1 + u(1+u)}{n(1+u)}(\mathbf{J}_n \otimes \mathbf{J}_n)\right],$$

$$\mathbf{K} = (\mathbf{I}_n \otimes \mathbf{J}_n) + \frac{u^2 + v}{u}(\mathbf{J}_n \otimes \mathbf{I}_n) - \frac{v}{nu(1+u)}(\mathbf{J}_n \otimes \mathbf{J}_n)\,,$$

$$\mathbf{K}^+ = \frac{1}{n^2(u^2+v)}\left[(u^2+v)(\mathbf{I}_n \otimes \mathbf{J}_n) + u(\mathbf{J}_n \otimes \mathbf{I}_n) - \frac{u^2(1+u) - uv + (u^2+v)\,\delta}{n\delta}(\mathbf{J}_n \otimes \mathbf{J}_n)\right],$$

where

$$\delta = (1+u)(u^2+u+v) - v = u[(1+u)^2 + v]\,,$$

$$\mathbf{M}_1 = \mathbf{W}^+\mathbf{V}_1\mathbf{K}^+ = \frac{1}{n^3}\left[n(\mathbf{I}_n \otimes \mathbf{J}_n) - \left(1 - \frac{u}{\delta}\right)(\mathbf{J}_n \otimes \mathbf{J}_n)\right],$$

$$\mathbf{M}_2 = \mathbf{W}^+\mathbf{V}_2\mathbf{K}^+ = \frac{1}{n^3(u^2+v)}\left[n(\mathbf{J}_n \otimes \mathbf{I}_n) - \frac{u(1+2u)}{\delta}(\mathbf{J}_n \otimes \mathbf{J}_n)\right].$$

The system (8) has now the form

$$\lambda_1\left(n - 1 + \frac{u}{\delta}\right) + \lambda_2\, u/\delta = f_1\,,$$

$$\lambda_1 \frac{u}{\delta} + \lambda_2 \frac{1}{u^2+v}\left[n - \frac{u(1+2u)}{\delta}\right] = f_2\,,$$

and the solution is

$$\lambda_1 = \lambda_1^* = f_1\psi[n\delta - u(1+2u)] - f_2\phi u\,,$$

$$\lambda_2 = \lambda_2^* = f_2\phi[(n-1)\,\delta + u] - f_1\phi u\,,$$

where

$$\phi = \frac{u^2 + v}{(n-1)(n\delta - 2u^2)}, \quad \psi = \frac{1}{(n-1)(n\delta - 2u^2)},$$

so that BIQUE is

$$\hat{\gamma} = \mathbf{y}'(\lambda_1^* \boldsymbol{M}_1 + \lambda_2^* \boldsymbol{M}_2)\,\mathbf{y} =$$

$$= \{f_1\psi[n\delta - u(1 + 2u)] - f_2\phi u\}\left\{\sum_{i=1}^{n} y_{i.}^2 - \left(1 - \frac{u}{\delta}\right) n y_{..}^2\right\} +$$

$$+ \{f_2\psi[(n-1)\delta + u] - f_1\psi u\}\left\{\sum_{j=1}^{n} y_{.j}^2 - \frac{u(1+2u)}{\delta} n y_{..}^2\right\} =$$

$$= \frac{\psi n}{\delta}\{[u^2(1 + 2u) - [n\delta - u(1 + 2u)](\delta - u)]f_1 + [u(\delta - u)(u^2 + v) -$$

$$- [(n-1)\delta + u]\,u(1 + 2u)]f_2\}\,y_{..}^2 + \psi\{[n\delta - u(1 + 2u)]f_1 -$$

$$- u(u^2 + v)f_2\}\sum_{i=1}^{n} y_{i.}^2 + \psi\{[(n-1)\delta + u]f_2 - uf_1\}\sum_{j=1}^{n} y_{.j}^2,$$

where $y_{i.}$ is given by (9) and

$$y_{.j} = \frac{1}{n}\sum_{i=1}^{n} y_{ij}, \quad y_{..} = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} y_{ij}.$$

If $v \to \infty$ we get

$$\lim_{v\to\infty} \hat{\gamma} = \frac{nf_1 + f_2}{n-1} y_{..}^2 + \frac{nf_1 - f_2}{n(n-1)}\sum_{i=1}^{n} y_{i.}^2 + \frac{f_2}{n}\sum_{j=1}^{n} y_{.j}^2.$$

References

[1] *S. Gnot, J. Kleffe:* Quadratic estimation in mixed linear models with two variance components. Journal of statistical planning and Inference 8 (1983) 267—279.

[2] *L. Kubáček:* Fundaments of the theory of estimates (Slovak). Veda, Publishing House of Slovak Acad. Sc., Bratislava 1983.

[3] *C. R. Rao:* Minimum variance quadratic unbiased estimation of variance components. J. Multivariate Anal. (1971) I, 445—456.

[4] *C. R. Rao:* Linear statistical inference and its applications, J. Wiley, New York 1973.

[5] *C. R. Rao, S. K. Mitra:* Generalized inverse of matrices and its applications. J. Wiley, New York 1972.

Souhrn

BAYESOVSKÉ NEVYCHÝLENÉ ODHADY V MODELU S DVĚMA VARIANČNÍMI KOEFICIENTY

V článku je odvozen explicitní vzorec pro Bayesův invariantní kvadratický nevychýlený odhad lineární funkce variančních koeficientů v lineárním smíšeném modelu $\mathbf{t} = \boldsymbol{X\beta} + \varepsilon$, $E(\mathbf{t}) = \boldsymbol{X\beta}$,

$D(\mathbf{t}) = \theta_1 \mathbf{U}_1 + \theta_2 \mathbf{U}_2$ s dvěma neznámými variančními koeficienty θ_1 a θ_2 v normálním případě. Na rozdíl od [1] je uvažována obecnější situace, kdy ani jedna z matic $\mathbf{U}_1$, $\mathbf{U}_2$ nemusí být jednotková a obě mohou být singulární. V závěru jsou odvozené výsledky aplikovány na dva příklady z analýzy rozptylu.

Резюме

НЕСМЕЩЕННАЯ ОЦЕНКА БАЙЕСА В МОДЕЛИ С ДВУМЬЯ ДИСПЕРСИОННЫМИ КОМПОНЕНТАМИ

JAROSLAV STUCHLÝ

В этой статье приводится выражение для инвариантной квадратической несмещенной оценки Байеса линейной функции параметров ковариационной матрицы в случае линейной модели $\mathbf{t} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, $E(\mathbf{t}) = \mathbf{X}\boldsymbol{\beta}$, $D(\mathbf{t}) = \theta_1 \mathbf{U}_1 + \theta_2 \mathbf{U}_2$. Предполагается нормальное расределение вектора $\mathbf{t}$ и вообще вырожденность матриц $\mathbf{U}_1$ и $\mathbf{U}_2$.

Статья заканчивается двумья примерами применения изложенной теории к дисперсионному анализу.

Author's address: RNDr. *Jaroslav Stuchlý*, Vysoká škola dopravy a spojov, Katedra matematiky, Marxa-Engelsa 25, 010 88 Žilina.