A Short-Cut Method for Computing Positive Variance Component Estimates

Joachim Hartung

Department of Statistics, SFB 475
University of Dortmund
Germany

Summary

In a general variance component model with positive variance components a short-cut method is presented that yields almost everywhere for these components positive estimators that are invariant with respect to mean value translation and stay near the unbiasedness.

Key Words: Variance components, Minque, positive Minque, Invariance, approximate unbiased positive variance estimates.

Zusammenfassung

In einem allgemeinen Varianz-Komponenten-Modell mit positiven Varianzkomponenten wird eine verkürzte Methode vorgestellt, welche für diese Komponenten fast überall positive Schätzer ergibt, die invariant bzgl. Mittelwerttranslationen sind und nahe der Unverzerrtheit bleiben.

1 Introduction

In a general variance components model there is the problem that unbiased quadratic estimators, or also maximum likelihood estimators if a distributional assumption is made, of the variance components can take on with a positive probability negative values for nonnegative variance components. These estimators are put then in such cases equal to zero, which for a usually strictly positive variance component is an unsatisfactory procedure. Therefore in the following a short-cut procedure is derived that overcomes this deficiency by yielding almost everywhere positive variance component estimators staying near the unbiasedness.

2 The Method

Let us consider the linear variance component model

$$z \sim \left(X\beta,\ \sum_{i=1}^{m} \alpha_i \cdot U_i \right) ,$$

that consists of an n-dimensional random variable z with mean value

$$\mathrm{E}\, z \;=\; X\beta$$

and variance-covariance matrix

$$\mathrm{Cov}(z) \;=\; \sum_{i=1}^{m} \alpha_i \cdot U_i \ ,$$

where the $(n \times k)$-design-matrix X and the m symmetric positive semi definite $(n \times n)$-matrices U_i, $i = 1, \ldots, m$, are known, while the parameter β varies in $\mathbb{R}^k$ and the parameter $\alpha = (\alpha_1, \ldots, \alpha_m)^T$ in $\mathbb{R}^m_+$, the positive orthant of $\mathbb{R}^m$, and we assume $\mathrm{rank}(X) < n$, and $\mathrm{Cov}(z)$ to be positive definite.

The problem considered here is to find quadratic estimates for the variance components $\alpha_1, \ldots, \alpha_m$, which should be positive almost everywhere, i.e. with probability one, and invariant with respect to the group Γ of mean value translations,

$$\Gamma \;=\; \left\{ z \longmapsto z + X\beta \mid \beta \in \mathbb{R}^k \right\} .$$

A maximal invariant linear statistic y with respect to Γ is given by

$$\begin{aligned} y \;&=\; \left(I - XX^+ \right) z \\ &=\; \mathrm{Proj}_{|\mathrm{Range}(X)^\perp}\, z \ , \end{aligned}$$

where I is the $(n \times n)$-identity-matrix and X^+ is the pseudoinverse of X.

We get the reduced (by invariance) linear model

$$y \sim \left(0,\ \sum_{i=1}^{m} \alpha_i \cdot V_i\right), \quad \alpha \in I\!R_+^m ,$$

where $V_i = (I - XX^+)U_i(I - XX^+)$, i=1, ... , m.

Let now A be a symmetric $n \times n$-matrix, then a quadratic (invariant) estimator for a linear form $p^T\alpha > 0$, $p = (p_1, \ldots, p_m)^T$, $p \geq 0$, $p \in I\!R^m$, is given by y^TAy, and denoting "tr" the trace, its bias is given by

$$\begin{aligned} \mathrm{E}\, y^TAy - p^T\alpha &= \mathrm{E}\,\mathrm{tr}\, Ayy^T - p^T\alpha \\ &= \mathrm{tr}\, A\mathrm{E}\,(yy^T) - p^T\alpha \\ &= \mathrm{tr}\, A(\sum_{i=1}^{m} \alpha_i V_i) - p^T\alpha \\ &= \sum_{i=1}^{m} \alpha_i(\mathrm{tr}\, AV_i - p_i) , \end{aligned}$$

such that y^TAy is an unbiased estimator of $p^T\alpha$ if

$$\mathrm{tr}\, AV_i = p_i, \quad \text{for all } i = 1, \ldots, m ,$$

and a solution A_0 of these equations, respectively the corresponding quadratic estimation function y^TA_0y is the (standard-) minimum norm invariant quadratic unbiased estimator (minque) of $p^T\alpha$, if A_0 has the minimum norm among all solutions.

Denote Sym the Hilbert space of all symmetric $(n \times n)$-matrices with the inner product of two matrices $A, B \in$ Sym defined by $\mathrm{tr}\, AB$, which then induces the standard norm $\|A\| = \sqrt{\mathrm{tr}\, A^2}$. Furthermore let PSD denote the cone of positive semi definite matrices in Sym.

If the matrices $V_1, \ldots, V_m$ are linearly independent, which for simplicity may be assumed here, then the minque A_0 exists for all $p \in I\!R^m$. Since we only claim $A_0 \in$ Sym, of course we usually get

$$y^TAy < 0 \ \text{ with positive probability} ,$$

and only in rare cases $A_0 \in$ PSD.

Resticting in advance A to be in PSD has the consequence that the equations for unbiasedness are seldom fulfilled, so that these conditions had to be weakened, cf. Seely(1971), Rao(1972), Pukelsheim(1981), Lehmann and Casella(1998), and Hartung(1981), where in section 4 there is also a solution algorithm given, which however needs some numerical effort.

Therefore in the following a short-cut method is presented that yields an approximation in PSD to A_0 with a correction for bias.

Let us introduce the linear operator

$$\text{g:}\quad \begin{array}{l} \text{Sym} \longrightarrow I\!R^m \\ A \longmapsto gA = \begin{pmatrix} \text{tr}\, AV_1 \\ \vdots \\ \text{tr}\, AV_m \end{pmatrix} \end{array} ,$$

then its adjoint g^* is given by

$$g^* : \quad \begin{array}{l} I\!R^m \longrightarrow \text{Sym} \\ a \longmapsto g^* a \;=\; \sum_{i=1}^m \alpha_i V_i, \quad a = (a_1, \ldots a_m)^T \end{array} ,$$

such that gg^* becomes the Gram-matrix G,

$$gg^* = G \quad = \quad \{\text{tr}\, V_i V_j\}_{\substack{i=1,\ldots,m \\ j=1,\ldots,m}} ,$$

of which the inverse G^{-1} exists because of the assumed linear independence of $V_1, \ldots, V_m$.

Denote g^+ the pseudoinverse operator of g, then the minque A_0 is given by

$$A_0 \quad = \quad g^+ p \ ,$$

which because of $g^+ = g^*(gg^*)^+$ permits the computational representation

$$A_0 \quad = \quad \sum_{i=1}^m a_i \cdot V_i, \ \text{with}\ a = (a_1, \ldots, a_m)^T = G^{-1} p.$$

Let us define now the vectors $b = (b_1, \ldots, b_m)^T$ and $c = (c_1, \ldots, c_m)^T$ by

$$b_i = \begin{cases} a_i, & \text{if } a_i > 0 \\ 0, & \text{if } a_i \le 0 \end{cases} , \quad \text{and} \quad c_i = \begin{cases} -a_i, & \text{if } a_i < 0 \\ 0, & \text{if } a_i \ge 0 \end{cases} ,$$

then

$$\begin{aligned} A_0 \quad &= \quad g^* a \\ &= \quad g^* b - g^* c \\ &=: \quad A_1 - A_2 \ , \end{aligned}$$

where $A_1 \in \text{PSD}$, $A_2 \in \text{PSD}$, and

$$gA_0 = p, \quad gA_1 = Gb =: q, \quad gA_2 = Gc =: r \ ,$$

such that

$$p^T\alpha \;=\; q^T\alpha - r^T\alpha\ ,$$

with the estimators

$$\begin{aligned}\widehat{p^T\alpha} &= \widehat{q^T\alpha} - \widehat{r^T\alpha}\\ &= y^TA_1y - y^TA_2y.\end{aligned}$$

We assume $A_2 \neq 0$, otherwise $A_0 \in$ PSD.

Now A_1 is an approximation in PSD to A_0, with the estimate

$$\begin{aligned}y^TA_1y = \widehat{q^T\alpha} &= \widehat{p^T\alpha} + \widehat{r^T\alpha}\\ &< \widehat{r^T\alpha},\ \text{with positive probability}\ ,\end{aligned}$$

although

$$q^T\alpha \;=\; p^T\alpha + r^T\alpha > r^T\alpha.$$

An additive correction of A_1, however, would lead again to possibly negative estimates. So the idea is now to work with a multiplicative correction term ψ, such that $\psi \cdot A_1$ replaces A_0 in $\widehat{p^T\alpha}$, and $\psi \cdot A_2$ replaces A_2 in $\widehat{r^T\alpha}$.

As determination equation for ψ we thus get

$$E\left\{\psi\cdot\ \widehat{q^T\alpha}\ + \psi\cdot\ \widehat{r^T\alpha}\right\} \;\stackrel{!}{=}\; q^T\alpha\ ,$$

yielding

$$\psi \;=\; \frac{q^T\alpha}{q^T\alpha + r^T\alpha}\ ,$$

and so we define our approximate solution in PSD as

$$A_{\text{PSD, appr.}(\psi)} \;:=\; \psi \cdot A_1\ ,$$

where ψ can be estimated by

$$\widehat{\psi} \;=\; \frac{y^TA_1y}{y^TA_1y + y^TA_2y}\ ,$$

which gives for $p^T\alpha$ the desired, approximate estimate

$$\left(\widehat{p^T\alpha}\right)_{\text{PSD, appr.}(\widehat{\psi})} = \widehat{\psi}\cdot y^TA_1y.$$

We can remark that first simulation results show a good performance of this estimator.

Instead of our easily obtainable decomposition of A_0, we may also use the spectral decomposition of A_0 into $A_{1^*} - A_{2^*}$ as follows:

$$\begin{aligned} A_{1^*} &:= \sum_{\lambda \in \sigma(A_0)} \max\{0, \lambda\} \cdot P_\lambda \ , \\ A_{2^*} &:= A_{1^*} - A_0 \ , \end{aligned}$$

where $\sigma(A_0)$ is the spectrum of A_0 and P_λ is the projection onto the eigenspace associated with λ, which needs a higher computational effort.

In more specified models also quite different and more detailed approximations may be derived, cf. e.g. Hartung(1999).

References

HARTUNG, J. (1981). Nonnegative minimum biased invariant estimation in variance component models. *The Annals of Statistics* **9**, 278–292.

HARTUNG, J. (1999). Ordungserhaltende positive Varianzschätzer bei gepaarten Messungen ohne Wiederholungen. *Allgemeines Statistisches Archiv* **83**, erscheint.

LEHMANN, E.L. AND CASELLA, G. (1998). *Theory of Point Estimation.* 2nd. ed., Springer, New York

PUKELSHEIM, F. (1981). On the existence of unbiased nonnegative estimates of variance covariance components. *The Annals of Statistics* **9**, 293–299.

RAO, C.R. (1972). Estimation of Variance and Covariance Components in Linear Models. *Journal of the American Statistical Association* **67**, 112–115.

SEELY, J. (1971). Quadratic subspaces and completeness. *The Annals of Mathematical Statistics* **42**, 710–721.

PROF. DR. JOACHIM HARTUNG
Lehrstuhl Statistik mit Anwendungen in den Ingenieurwissenschaften
Fachbereich Statistik der Universität Dortmund
Vogelpothsweg 87
D–44221 Dortmund
Germany