Math. Operationsforsch. Statist., Ser. Statistics, Vol. 8 (1977) No. 3, 323–331

On Hsu's Model in Regression Analysis

Friedrich Pukelsheim [1]

Summary. The paper exemplifies with Hsu's model a general pattern as how to derive results of variance component estimation from well known results of mean estimation, as far as linear model theory is concerned. This '*dispersion-mean-correspondence*' provides new and short proofs for various theorems from the literature, concerning unbiased invariant quadratic estimators with minimum Bayes risk or minimum variance. For pure variance component models, unbiased non-negative quadratic estimability is characterized in terms of the design matrices.

1. Introduction

The purpose of this communication is to exemplify with Hsu's model a general pattern as how to derive results of variance component estimation from known results of mean estimation.

The dispersion-mean-correspondence (Sect. 2) introduces a derived model such that mean regression in the derived model corresponds to estimating the variance components in the original model. Hsu's model specifies the fourth moments via the kurtosis γ, thus opening the way for minimum variance estimation. Sect. 3 collects some matrix algebra for convenient reference. In Sect. 4, some known results of J. Kleffe and R. Pincus [8], P. L. Hsu [6], H. Drygas [2], and C. R. Rao [14] are proved by persistently applying the dispersion-mean-correspondence. This approach extends insight and understanding of linear model theory, providing short proofs, slight generalizations, and alternative characterizations. Sect. 5 is concerned with the existence of unbiased non-negative definite quadratic estimates of variance components and presents an estimability criterion in terms of the design matrices that specify the model.

Previous Work. S. K. Mitra [10] suggests an approach that is very close to the dispersion-mean-correspondence as presented here or in [12], he, however, stops exploitation at an intermediate stage. J. Seely [16] and other authors, cf., H. Drygas [3], S. Gnot, W. Klonecki and R. Zmyslony [4], J. Kleffe [7], R. Zmyslony [18], reduce variance component estimation to mean estimation in coordinate free terms which seems to provoke ready application somewhat less than the dispersion-mean-correspondence. Most of the examples in Sect. 4 were originally proved by explicitly minimizing the risk function, though the connection to mean estimation is hinted at (H. Drygas [2, p. 382]) or

[1] Institut für Mathematische Stochastik der Universität, Hermann-Herder-Str. 10, D-7800 Freiburg im Breisgau, Federal Republic of Germany.

implicit (C. R. RAO [14, p. 451]). The algebraic notions as introduced below have successfully been utilized by other authors as well, cf., J. KLEFFE and R. PINCUS [8], S. K. MITRA [10], G. P. H. STYAN [17], R. ZMYSLONY [18].

Notations. For a matrix $\boldsymbol{A}$, let $\boldsymbol{A}', \boldsymbol{A}^+, \Re\boldsymbol{A}, \Re^\perp\boldsymbol{A}$ denote its transposed matrix, MOORE-PENROSE inverse [15, p. 26], range (column space), and the range's orthogonal complement, respectively. Let $vec\ \boldsymbol{A}$ be the vector obtained from $\boldsymbol{A}$ by ordering its entries lexicographically. The KRONECKER product [15, p. 29] is denoted by $\otimes$. The function vec is an inner product and tensor product preserving vector space isomorphism:

$$(vec\ \boldsymbol{A})'\ vec\ \boldsymbol{B} = trace\ \boldsymbol{AB}', \tag{1}$$

$$vec\ \boldsymbol{xy}' = \boldsymbol{x} \otimes \boldsymbol{y}, \tag{2}$$

as follows by considering the standard basis vectors $\boldsymbol{e}_\nu$ with ν-th component 1 and zeroes elsewhere, and the basis matrices $\boldsymbol{E}_{\nu\mu} = \boldsymbol{e}_\nu \boldsymbol{e}_\mu'$ with (ν, μ)-entry 1 and zeroes elsewhere. For $(vec\ \boldsymbol{A})'$ we shall also write $vec'\boldsymbol{A}$. $\boldsymbol{I}$ denotes an identity matrix, its order following from the context.

2. Model Set-up and the Dispersion-Mean-Correspondence

The General Linear Model. For an R^n-valued random vector $\boldsymbol{Y}$, a linear model is specified by linear decompositions of both the mean vector $\mathsf{E}\boldsymbol{Y}$ and the dispersion matrix (variance covariance matrix) $\mathsf{D}\boldsymbol{Y}$:

$$\mathsf{E}\boldsymbol{Y} = \sum_{\pi=1}^{p} b_\pi \boldsymbol{x}_\pi = \boldsymbol{Xb}, \quad \mathsf{D}\boldsymbol{Y} = \sum_{\varkappa=1}^{k} t_\varkappa \boldsymbol{V}_\varkappa, \quad \boldsymbol{M} = \boldsymbol{I} - \boldsymbol{XX}^+, \tag{3}$$

where the (n, p)-matrix $\boldsymbol{X} = [\boldsymbol{x}_1 : \ldots : \boldsymbol{x}_p]$ and the k symmetric (n, n)-matrices $\boldsymbol{V}_\varkappa$ are known, whereas $\boldsymbol{b} = (b_1, \ldots, b_p)'$ and $\boldsymbol{t} = (t_1, \ldots, t_k)'$ are to be estimated; $\boldsymbol{M}$ is the orthogonal projector onto $\Re^\perp\boldsymbol{X}$.

For estimating $\boldsymbol{t}$ or linear functions of $\boldsymbol{t}$, we choose, as usual, quadratic estimators $\boldsymbol{Q}(\boldsymbol{Y})$ which, by definition, are derived from bilinear functions $B(.,.)$ by setting both arguments equal to $\boldsymbol{Y}$: $\boldsymbol{Q}(\boldsymbol{Y}) = B(\boldsymbol{Y}, \boldsymbol{Y})$. A maximal invariant statistic with respect to all 'mean translations' $\boldsymbol{y} \to \boldsymbol{y} + \boldsymbol{Xb}$, $\boldsymbol{b} \in \mathrm{R}^p$, is $\boldsymbol{MY}$ (cf., J. SEELY [16, p. 1646], J. KLEFFE [7]). Thus, $\boldsymbol{Q}(\boldsymbol{Y})$ is an *invariant quadratic estimator* (IQE) iff $\boldsymbol{Q}(\boldsymbol{Y}) = \boldsymbol{Q}(\boldsymbol{MY})$.

E.g., when estimating a linear form $\boldsymbol{q}'\boldsymbol{t}$, $\boldsymbol{q} \in \mathrm{R}^k$, the set of all IQEs is $\{\boldsymbol{Y}'\boldsymbol{AY}\}$, where $\boldsymbol{A}$ is an arbitrary symmetric (n, n)-matrix satisfying $\boldsymbol{A} = \boldsymbol{MAM}$, or, equivalently, $\boldsymbol{AX} = \boldsymbol{0}$.

Invariance is a natural statistical requirement: All of $\boldsymbol{Y}$ which is in the range $\Re\boldsymbol{X}$ may be explained by mean regression, leaving the residuals $\boldsymbol{MY}$ for inference on the dispersion parameter $\boldsymbol{t}$. Technically speaking, invariant estimates of $\boldsymbol{t}$ are free from the mean parameter, cf., R. R. CORBEIL and S. R. SEARLE [1], J. KLEFFE [7]. Finally, the expectation of a quadratic estimate $\boldsymbol{Y}'\boldsymbol{AY}$ does not depend on the mean parameter iff $\boldsymbol{X}'\boldsymbol{AX} = \boldsymbol{0}$. In most applications, as in HSU's model, $\boldsymbol{A}$ should be non-negative definite (NND). This, however, and $\boldsymbol{X}'\boldsymbol{AX} = \boldsymbol{0}$ imply $\boldsymbol{AX} = \boldsymbol{0}$, i.e., invariance.

The Dispersion-Mean-Correspondence. For a linear model (3), consider the derived random R^{n^2}-vector $\boldsymbol{MY}\otimes\boldsymbol{MY}$. By (3) and (2), $\mathsf{E}\boldsymbol{MY}\otimes\boldsymbol{MY}=\boldsymbol{M}\otimes\boldsymbol{M}\cdot vec\sum t_\varkappa \boldsymbol{V}_\varkappa$, and $\boldsymbol{MY}\otimes\boldsymbol{MY}$ gives rise to a linear model for mean estimation. Introducing the (n^2, k)-matrices

$$\boldsymbol{D}=[vec\,\boldsymbol{V}_1:\cdots: vec\,\boldsymbol{V}_k],\quad \boldsymbol{D}_{\boldsymbol{M}}=\boldsymbol{M}\otimes\boldsymbol{M}\cdot\boldsymbol{D}\,, \tag{4}$$

we arrive at

$$\mathsf{E}\boldsymbol{MY}\otimes\boldsymbol{MY}=\boldsymbol{D}_{\boldsymbol{M}}\boldsymbol{t}\,. \tag{5}$$

Thus, $\boldsymbol{t}$ may be looked at as the dispersion parameter in the original model (3), or as the mean parameter in the derived model (5); this we call the *dispersion-mean-correspondence.*

It remains to be shown that the class of natural estimators of $\boldsymbol{t}$ is not changed by the dispersion mean-correspondence. Clearly, for any (k, n^2)-matrix $\boldsymbol{L}$,

$$Q(\boldsymbol{Y})=\boldsymbol{L}\cdot\boldsymbol{MY}\otimes\boldsymbol{MY} \tag{6}$$

as a linear estimator of $\boldsymbol{t}$ in the derived model (5), is an IQE of $\boldsymbol{t}$ in the original model (3). Conversely, the KRONECKER product is a tensor product [5, p. 12], i. e., for every bilinear function $B(\boldsymbol{x},\boldsymbol{y})$ there exists a (unique) linear function $\boldsymbol{L}$ such that $B(\boldsymbol{x},\boldsymbol{y})=\boldsymbol{L}\cdot\boldsymbol{x}\otimes\boldsymbol{y}$. This implies that any IQE $Q(\boldsymbol{Y})$ is representable as in (6), and thus is a linear estimator in the derived model. E.g., for an IQE $\boldsymbol{Y}'\boldsymbol{AY}$ of a linear form $\boldsymbol{q}'\boldsymbol{t}$ we get from (1) and (2) $\boldsymbol{Y}'\boldsymbol{AY}=trace\,\boldsymbol{A}\cdot\boldsymbol{MY}(\boldsymbol{MY})'=vec'\boldsymbol{A}\cdot\boldsymbol{MY}\otimes\boldsymbol{MY}$, i.e., $\boldsymbol{L}=vec'\boldsymbol{A}$.

In Sect. 4 we shall be concerned with minimum variance estimation of $\boldsymbol{t}$. To this end, we introduce the (n^2, n^2)-matrices

$$\boldsymbol{F}=\mathsf{D}\boldsymbol{Y}\otimes\boldsymbol{Y},\quad \boldsymbol{F}_{\boldsymbol{M}}=\boldsymbol{M}\otimes\boldsymbol{M}\cdot\boldsymbol{F}\cdot\boldsymbol{M}\otimes\boldsymbol{M},\quad \boldsymbol{N}=\boldsymbol{I}-\boldsymbol{D}_{\boldsymbol{M}}\boldsymbol{D}_{\boldsymbol{M}}^{+}\,, \tag{7}$$

i.e., the matrix of all central mixed fourth moments of $\boldsymbol{Y}$, the dispersion matrix of $\boldsymbol{MY}\otimes\boldsymbol{MY}$, and the orthogonal projector onto $\mathfrak{R}^{\perp}\boldsymbol{D}_{\boldsymbol{M}}$, respectively. The following lemma is stated for later reference. It applies the celebrated LEHMANN-SCHEFFÉ theorem [15, p. 317] to the derived model; the alternative representations in part (ii) follow by (1), (4), and (12).

Lemma 2.1. *Let a linear model be given by* (3) *and* (7); *let* $\boldsymbol{L}$ *be a* (k, n^2)*-matrix,* $\boldsymbol{A}$ *be a symmetric* (n, n)*-matrix, and* $\boldsymbol{q}\in\mathrm{R}^k$. *Then:*
(i) $\boldsymbol{L}\cdot\boldsymbol{MY}\otimes\boldsymbol{MY}$ *is an unbiased* IQE *of* $\boldsymbol{t}$ *with minimum variance under* $\boldsymbol{F}_{\boldsymbol{M}}$ *(among all other unbiased* IQE) *iff* $\boldsymbol{LD}_{\boldsymbol{M}}=\boldsymbol{I}$ *and* $\boldsymbol{LF}_{\boldsymbol{M}}\boldsymbol{N}=\boldsymbol{0}$.
(ii) $\boldsymbol{Y}'\boldsymbol{AY}$ *is an unbiased* IQE *of* $\boldsymbol{q}'\boldsymbol{t}$ *with minimum variance under* $\boldsymbol{F}_{\boldsymbol{M}}$ *iff* $\boldsymbol{AX}=\boldsymbol{0}$, $\boldsymbol{q}=\boldsymbol{D}_{\boldsymbol{M}}'\cdot vec\,\boldsymbol{A}=(trace\;\boldsymbol{V}_1\boldsymbol{A},\ldots, trace\;\boldsymbol{V}_k\boldsymbol{A})'$, *and*

$$\boldsymbol{F}_{\boldsymbol{M}}\cdot vec\,\boldsymbol{A}\in\mathfrak{R}\boldsymbol{D}_{\boldsymbol{M}}=\{vec\sum\lambda_\varkappa \boldsymbol{MV}_\varkappa\boldsymbol{M}/\lambda\in\mathrm{R}^k\}.$$

Hsu's Model. HSU's model specifies the fourth moments $\boldsymbol{F}$ via (quasi-)independence and the kurtosis γ of k random effects $\xi_\varkappa$, cf., H. DRYGAS [2], P. L. HSU [6], J. KLEFFE and R. PINCUS [8], C. R. RAO [14]. Following C. R. RAO [14, p. 446] we assume a linear decomposition according to

$$\boldsymbol{Y}-\mathsf{E}\boldsymbol{Y}=\sum_{\varkappa=1}^{k}\boldsymbol{U}_\varkappa\xi_\varkappa=\boldsymbol{U}\xi\,, \tag{8}$$

where the $\boldsymbol{U}_\varkappa$ are known $(n, c_\varkappa)$-matrices, $c=\sum c_\varkappa$, $\boldsymbol{U}=[\boldsymbol{U}_1:\cdots:\boldsymbol{U}_k]$ is of order (n, c), and $\xi=[\xi_1':\cdots:\xi_k']'$ is a random R^c-vector whose independent subvectors

$\boldsymbol{\xi}_\varkappa$ have independent components $\xi_{\varkappa,\nu}$ satisfying

$$\mathsf{E}\xi_{\varkappa,\nu}=0, \quad \mathsf{E}\xi^2_{\varkappa,\nu}=\sigma^2_\varkappa, \quad \mathsf{E}\xi^4_{\varkappa,\nu}=(\gamma_\varkappa+3)\,\sigma^4_\varkappa, \quad \nu=1,\ldots,c_\varkappa\,. \tag{9}$$

In general, a random R^c-vector $\boldsymbol{\xi}$ whose independent components satisfy $\mathsf{E}\xi_\nu=0$, $\mathsf{E}\xi^2_\nu=\tilde\sigma^2_\nu$, $\mathsf{E}\xi^4_\nu=(\tilde\gamma_\nu+3)\,\tilde\sigma^4_\nu$, has mixed fourth moments

$$\mathsf{D}\boldsymbol{\xi}\otimes\boldsymbol{\xi}=\sum_{\mu,\nu=1}^{c}\tilde\sigma^2_\mu\tilde\sigma^2_\nu\,(\boldsymbol{E}_{\mu\mu}\otimes\boldsymbol{E}_{\nu\nu}+\boldsymbol{E}_{\mu\nu}\otimes\boldsymbol{E}_{\nu\mu}+\delta_{\mu\nu}\tilde\gamma_\nu\boldsymbol{E}_{\nu\nu}\otimes\boldsymbol{E}_{\nu\nu})\,. \tag{10}$$

Finally, we introduce

$$\boldsymbol{H}=\begin{bmatrix}\boldsymbol{1}'_{c_1} & & 0\\ & \ddots & \\ 0 & & \boldsymbol{1}'_{c_k}\end{bmatrix}, \quad \tilde{\boldsymbol{\sigma}}^2=\boldsymbol{H}'\boldsymbol{\sigma}^2,\ \tilde{\boldsymbol{\gamma}}=\boldsymbol{H}'\boldsymbol{\gamma}\,, \tag{11}$$

where the vectors $\boldsymbol{1}_{c_\varkappa}$ consist of $c_\varkappa$ ones, so that $\boldsymbol{H}$ is of order (k, c). Thus, (10) and (11) yield the mixed fourth moments under assumption (9). We are now ready to precisely define a HSU-model:

Definition. *A* HSU-*model is a linear model as specified by* (3), (4), (7), (10), (11), *where* $\boldsymbol{V}_\varkappa=\boldsymbol{U}_\varkappa\boldsymbol{U}'_\varkappa$, $\varkappa=1,\ldots,k$; $\boldsymbol{U}$ *and the* $\boldsymbol{U}_\varkappa$*'s are as in* (8), $\boldsymbol{t}=\boldsymbol{\sigma}^2=(\sigma_1^2,\ldots,\sigma_k^2)'$, $\boldsymbol{\gamma}=(\gamma_1,\ldots,\gamma_k)'$, *and* $\boldsymbol{F_M}=\boldsymbol{F_M}(\boldsymbol{\sigma}^2,\boldsymbol{\gamma})=\boldsymbol{MU}\otimes\boldsymbol{MU}\cdot(\mathsf{D}\boldsymbol{\xi}\otimes\boldsymbol{\xi})\cdot\boldsymbol{U}'\boldsymbol{M}\otimes\boldsymbol{U}'\boldsymbol{M}$.

In order to appealingly display $\boldsymbol{F_M}(\boldsymbol{\sigma}^2,\boldsymbol{\gamma})$, we now collect some matrix algebra.

3. Some Matrix Algebra

The Separating Property of *vec*. For any 3 matrices $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ of appropriate order one has

$$\mathit{vec}\,\boldsymbol{ABC}=\boldsymbol{A}\otimes\boldsymbol{C}'\cdot\mathit{vec}\,\boldsymbol{B}\,. \tag{12}$$

Taking $\boldsymbol{B}=\boldsymbol{E}_{\nu\mu}$, this follows at once from (2).

Diagonalizer. For a square matrix $\boldsymbol{A}$, *Diag* $\boldsymbol{A}$ denotes the diagonal matrix with diagonal entries copied from $\boldsymbol{A}$. For a vector $\boldsymbol{\gamma}$, *Diag* $\boldsymbol{\gamma}$ is the diagonal matrix with diagonal equal to $\boldsymbol{\gamma}$.

Introducing the (c^2, c)-matrix $\boldsymbol{D}_c=[\mathit{vec}\,\boldsymbol{E}_{11}:\cdots:\mathit{vec}\,\boldsymbol{E}_{cc}]$, where $\boldsymbol{E}_{11}$ etc. are basis (c, c)-matrices, yields, by (2), $\boldsymbol{D}_c\boldsymbol{D}'_c=\sum\mathit{vec}\,\boldsymbol{E}_{\nu\nu}\cdot\mathit{vec}'\,\boldsymbol{E}_{\nu\nu}=\sum(\boldsymbol{e}_\nu\otimes\boldsymbol{e}_\nu)\times\times(\boldsymbol{e}_\nu\otimes\boldsymbol{e}_\nu)'=\sum\boldsymbol{E}_{\nu\nu}\otimes\boldsymbol{E}_{\nu\nu}$, and, by (12), $\boldsymbol{D}_c\boldsymbol{D}'_c\cdot\mathit{vec}\,\boldsymbol{A}=\mathit{vec}\,\mathit{Diag}\,\boldsymbol{A}$. More general, for every diagonal (c, c)-matrix $\boldsymbol{\Delta}$, and for every (c, c)-matrix $\boldsymbol{A}$, we have

$$\boldsymbol{D}_c\boldsymbol{\Delta}\boldsymbol{D}'_c\cdot\mathit{vec}\,\boldsymbol{A}=\mathit{vec}\,\boldsymbol{\Delta}\,\mathit{Diag}\,\boldsymbol{A}, \quad \text{where} \quad \boldsymbol{D}_c=[\mathit{vec}\,\boldsymbol{E}_{11}:\cdots:\mathit{vec}\,\boldsymbol{E}_{cc}]. \tag{13}$$

HADAMARD's Produkt. When diagonalizer are used, HADAMARD's product $\boldsymbol{A}*\boldsymbol{B}=((A_{ij}\cdot B_{ij}))$ is not far. This is due to *Diag* $\boldsymbol{A}=\boldsymbol{I}*\boldsymbol{A}$ and the following lemma.

Lemma 3.1. *Let* $\boldsymbol{D}_c$, $\boldsymbol{D}_n$ *be defined as in* (13) *of order* (c^2, c), (n^2, n), *respectively. Let* $\boldsymbol{A}$, $\boldsymbol{B}$ *two* (c, n)*-matrices, and* $\boldsymbol{a}\in\mathrm{R}^c$, $\boldsymbol{b}\in\mathrm{R}^n$. *Then:*

(i) $\boldsymbol{A}*\boldsymbol{B}=\boldsymbol{D}'_c\cdot\boldsymbol{A}\otimes\boldsymbol{B}\cdot\boldsymbol{D}_n\,.$

(ii) $\boldsymbol{a}'\cdot\boldsymbol{A}*\boldsymbol{B}\cdot\boldsymbol{b}=\mathit{trace}\,\mathit{Diag}\,\boldsymbol{a}\cdot\boldsymbol{A}\cdot\mathit{Diag}\,\boldsymbol{b}\cdot\boldsymbol{B}'\,.$

Proof. Verification is immediate when taking $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{a}, \boldsymbol{b}$ to be basis matrices (vectors). ■

Using Lemma 3.1, many properties of the HADAMARD product, cf., e.g., G. P. H. STYAN [17], may easily be inferred from corresponding properties of the KRONECKER product.

4. On Estimates in the Hsu-Model

The Mixed Fourth Moment. In the next two lemmas, we study the mixed fourth moments as a linear operator.

Lemma 4.1. *Let* $\boldsymbol{\xi}$ *be a random* R^c*-vector whose independent components* ξ_ν *satisfy* $\mathsf{E}\xi_\nu = 0$, $\mathsf{E}\xi_\nu^2 = \tilde{\sigma}_\nu^2$, *and* $\mathsf{E}\xi_\nu^4 = (\tilde{\gamma}_\nu + 3)\, \tilde{\boldsymbol{\sigma}}_\nu^4$. *For a fixed* $\tilde{\sigma}_0^2$, *put* $\boldsymbol{\Delta}_1 = \mathsf{D}\boldsymbol{\xi} = Diag\ \tilde{\boldsymbol{\sigma}}_0^2$, $\boldsymbol{\Gamma} = Diag\ \tilde{\boldsymbol{\gamma}}$, $\boldsymbol{\Delta}_2 = \boldsymbol{\Delta}_1, \boldsymbol{\Gamma}\boldsymbol{\Delta}_1$, *and let* $\boldsymbol{D}_c$ *be defined as in* (13). *Then, for every* (n, n)*-matrix* $\boldsymbol{A}$,

$$(\mathsf{D}\boldsymbol{\xi} \otimes \boldsymbol{\xi}) \cdot vec\ \boldsymbol{A} = (2\boldsymbol{\Delta}_1 \otimes \boldsymbol{\Delta}_1 + \boldsymbol{D}_c \boldsymbol{\Delta}_2 \boldsymbol{D}_c') \cdot vec \frac{\boldsymbol{A} + \boldsymbol{A}'}{2}. \tag{14}$$

Proof. The assertion is a consequence of formulae (10), (12), and (13). ■

Lemma 4.2. *Assume a* Hsu*-model. For a fixed* $\boldsymbol{\sigma}_0^2$, *put* $\boldsymbol{V}_* = \sum \sigma_{0,\varkappa}^2 \boldsymbol{V}_\varkappa = \boldsymbol{U}\boldsymbol{\Delta}_1\boldsymbol{U}'$. *Then:*

(i) *For every* (n, n)*-matrix* $\boldsymbol{A}$,

$$\boldsymbol{F}_{\boldsymbol{M}}(\boldsymbol{\sigma}_0^2, \boldsymbol{\gamma}) \cdot vec\ \boldsymbol{A} = (2\boldsymbol{M}\boldsymbol{V}_*\boldsymbol{M} \otimes \boldsymbol{M}\boldsymbol{V}_*\boldsymbol{M} + \boldsymbol{M}\boldsymbol{U} \otimes \boldsymbol{M}\boldsymbol{U} \cdot \boldsymbol{D}_c \boldsymbol{\Delta}_2 \boldsymbol{D}_c' \tag{15}$$
$$\boldsymbol{U}'\boldsymbol{M} \otimes \boldsymbol{U}'\boldsymbol{M}) \cdot vec \frac{\boldsymbol{A} + \boldsymbol{A}'}{2}.$$

(ii) *For every symmetric* (n, n)*-matrix* $\boldsymbol{A}$ *satisfying* $\boldsymbol{A} = \boldsymbol{M}\boldsymbol{A}\boldsymbol{M}$,

$$\boldsymbol{F}_{\boldsymbol{M}}(\boldsymbol{\sigma}_0^2, \boldsymbol{\gamma}) \cdot vec\ \boldsymbol{A} = vec\,(2\boldsymbol{M}\boldsymbol{V}_*\boldsymbol{A}\boldsymbol{V}_*\boldsymbol{M} + \boldsymbol{M}\boldsymbol{U}\boldsymbol{\Delta}_2 \cdot Diag\ \boldsymbol{U}'\boldsymbol{A}\boldsymbol{U} \cdot \boldsymbol{U}'\boldsymbol{M}). \tag{16}$$

Proof. Part (i) follows from Lemma 4.1, and, by (12), implies part (ii). ■

With these preparations we are now ready to characterize optimal estimates in the Hsu-model.

BAYES estimates. J. KLEFFE and R. PINCUS [8, Th. 3.8] consider unbiased IQEs with minimum BAYES risk:

Theorem 4.1. *Assume a* Hsu*-model with a priori distribution* P *for* $\boldsymbol{\sigma}^2$, *put*

$$\bar{\boldsymbol{R}} = \mathsf{E}_P \boldsymbol{\sigma}^2 \boldsymbol{\sigma}^{2\prime}, \quad \bar{\boldsymbol{S}} = Diag\ \boldsymbol{H}'\bar{\boldsymbol{R}}\boldsymbol{H}, \tag{17}$$

and let $\boldsymbol{q} \in \mathrm{R}^k$. *Then* $\boldsymbol{Y}'\boldsymbol{A}\boldsymbol{Y}$ *is an unbiased* IQE *of* $\boldsymbol{q}'\boldsymbol{t}$ *with minimum* BAYES *risk at* $\boldsymbol{\gamma}$ *iff* $\boldsymbol{A}\boldsymbol{X} = \boldsymbol{O}$, $\boldsymbol{q} = (trace\ \boldsymbol{V}_1\boldsymbol{A}, \ldots, trace\ \boldsymbol{V}_k\boldsymbol{A})'$, *and* $\boldsymbol{M}\boldsymbol{U} \cdot (2\boldsymbol{H}'\bar{\boldsymbol{R}}\boldsymbol{H} + \bar{\boldsymbol{S}}\boldsymbol{\Gamma}) * (\boldsymbol{U}'\boldsymbol{A}\boldsymbol{U}) \cdot \boldsymbol{U}'\boldsymbol{M}$ *is a linear combination of* $\boldsymbol{M}\boldsymbol{V}_1\boldsymbol{M}, \ldots, \boldsymbol{M}\boldsymbol{V}_k\boldsymbol{M}$.

Proof. By Lemma 4.2.ii, the risk operator is $\Phi(\boldsymbol{A}) = vec\ \mathsf{E}_P\,(2\boldsymbol{M}\boldsymbol{U}\boldsymbol{\Delta}_1\boldsymbol{U}'\boldsymbol{A}\boldsymbol{U}\boldsymbol{\Delta}_1 \cdot \boldsymbol{U}'\boldsymbol{M} + \boldsymbol{M}\boldsymbol{U}\boldsymbol{\Delta}_1^2\boldsymbol{\Gamma} \cdot Diag\ \boldsymbol{U}'\boldsymbol{A}\boldsymbol{U} \cdot \boldsymbol{U}'\boldsymbol{M})$. But, cf., [17], $\boldsymbol{\Delta}_1 = Diag\ \boldsymbol{H}'\boldsymbol{\sigma}^2$ implies $\boldsymbol{\Delta}_1\boldsymbol{U}'\boldsymbol{A}\boldsymbol{U}\boldsymbol{\Delta}_1 = (\boldsymbol{H}'\boldsymbol{\sigma}^2\boldsymbol{\sigma}^{2\prime}\boldsymbol{H}) * (\boldsymbol{U}'\boldsymbol{A}\boldsymbol{U})$ and $\boldsymbol{\Delta}_1^2 = (\boldsymbol{H}'\boldsymbol{\sigma}^2\boldsymbol{\sigma}^{2\prime}\boldsymbol{H}) * \boldsymbol{I} = Diag\ \boldsymbol{H}'\boldsymbol{\sigma}^2\boldsymbol{\sigma}^{2\prime}\boldsymbol{H}$. Thus $\mathsf{E}_P\boldsymbol{\Delta}_1^2 = \bar{\boldsymbol{S}}$, and we get $\Phi(\boldsymbol{A}) = vec\ \boldsymbol{M}\boldsymbol{U} \cdot (2\boldsymbol{H}'\bar{\boldsymbol{R}}\boldsymbol{H} + \bar{\boldsymbol{S}}\,\boldsymbol{\Gamma}) * (\boldsymbol{U}'\boldsymbol{A}\boldsymbol{U}) \cdot \boldsymbol{U}'\boldsymbol{M}$. The assertion now follows by Lemma 2.1.ii, mutatis mutandis. ■

MV estimates. The rest of this section is concerned with *minimum variance unbiased* IQE (MV – UB – IQE) of a linear form $\boldsymbol{q}'\boldsymbol{t}$.

P. L. Hsu was the first in this area, and his problem [6, Th. 2] reads in terms of the derived model: When is the simple least squares estimate of minimum variance? Hsu assumes independent components of $\boldsymbol{\gamma}$ with equal variances σ^2 and possibly unequal kurtosis $\gamma_1, \ldots, \gamma_n$.

Theorem 4.2. *Assume a* Hsu*-model with* $c=n$, $\boldsymbol{U}=\boldsymbol{I}$, $\boldsymbol{V}_*=\sigma^2\boldsymbol{I}$, *and* $s=\mathrm{rank}\,X<n$. *Put* $\boldsymbol{M}_2=\boldsymbol{M} * \boldsymbol{M}$, *and* $\boldsymbol{m}=(M_{11}, \ldots, M_{nn})'$. *Then* $(n-s)^{-1}\boldsymbol{Y}'\boldsymbol{M}\boldsymbol{Y}$ *is a* MV – UB – IQE *of* σ^2 *iff*

$$\boldsymbol{M}_2\boldsymbol{\Gamma}\boldsymbol{m}=\varrho\boldsymbol{m}, \quad \varrho=(n-s)^{-1}\,\boldsymbol{m}'\boldsymbol{\Gamma}\boldsymbol{m}\,. \tag{18}$$

Proof. Here, $\boldsymbol{D_M}=vec\,\boldsymbol{M}$, so that the simple least squares estimate is $(\boldsymbol{D}'_{\boldsymbol{M}}\boldsymbol{D}_{\boldsymbol{M}})^{-1}\boldsymbol{D}'_{\boldsymbol{M}}\cdot\boldsymbol{MY}\otimes\boldsymbol{MY}=(trace\,\boldsymbol{M})^{-1}\;trace\;\boldsymbol{M}\cdot\boldsymbol{MY}(\boldsymbol{MY})'=(n-s)^{-1}\,\boldsymbol{Y}'\boldsymbol{MY}$, by (1) and (2). This is of minimum variance [11, p. 148] iff $\Re\boldsymbol{D}_{\boldsymbol{M}}$ is invariant under $\boldsymbol{F}_{\boldsymbol{M}}(\sigma_0^2, \boldsymbol{\gamma})$. Lemma 4.2.ii yields $\boldsymbol{F}_{\boldsymbol{M}}(\sigma_0^2, \boldsymbol{\gamma})\cdot vec\,\boldsymbol{M}=\sigma_0^4\,vec\,(2\boldsymbol{M}+\boldsymbol{M\Gamma}\cdot Diag\;\boldsymbol{M}\cdot\boldsymbol{M})$. Thus, a necessary and sufficient condition is

$$\boldsymbol{M}\cdot Diag\;\boldsymbol{\Gamma m}\cdot\boldsymbol{M}=\varrho\boldsymbol{M}, \quad \varrho=(n-s)^{-1}\,\boldsymbol{m}'\boldsymbol{\Gamma}\boldsymbol{m}\,. \tag{19}$$

By considering the diagonal elements, (19) implies (18). The converse follows, with Lemma 3.1.ii, from $\|\boldsymbol{M}\cdot Diag\;\boldsymbol{\Gamma m}\cdot\boldsymbol{M}-\varrho\boldsymbol{M}\|^2=trace\;Diag\;\boldsymbol{\Gamma m}\cdot\boldsymbol{M}\cdot Diag\,\boldsymbol{\Gamma m}\cdot\boldsymbol{M}-2\varrho\;trace\;Diag\;\boldsymbol{m}\cdot\boldsymbol{\Gamma}\cdot Diag\;\boldsymbol{m}+\varrho^2\,trace\;\boldsymbol{M}=\boldsymbol{m}'\boldsymbol{\Gamma}\boldsymbol{M}_2\boldsymbol{\Gamma}\boldsymbol{m}-2\varrho\boldsymbol{m}'\boldsymbol{\Gamma}\boldsymbol{m}+\varrho\boldsymbol{m}'\boldsymbol{\Gamma}\boldsymbol{m}=\boldsymbol{m}'\boldsymbol{\Gamma}(\boldsymbol{M}_2\boldsymbol{\Gamma}\boldsymbol{m}-\varrho\boldsymbol{m})$. ■
The next result, due to H. DRYGAS [2, Th. 3.5.a], does also lead to eq. (19).

Theorem 4.3. *Let* $\boldsymbol{Y}'\boldsymbol{AY}$ *be an* IQE *of* σ^2 *for the* Hsu*-model in Th. 4.2. Then* $\boldsymbol{Y}'\boldsymbol{AY}$ *is a* MV – UB – IQE *of* σ^2 *iff trace* $\boldsymbol{A}=1$ *and* $2\boldsymbol{A}+\boldsymbol{M\Gamma}\cdot Diag\;\boldsymbol{A}\cdot\boldsymbol{M}$ *is a scalar multiple of* $\boldsymbol{M}$.

Proof. Using Lemmas 2.1.ii, 4.2.ii, we get $1=\boldsymbol{D}'_{\boldsymbol{M}}\,vec\,\boldsymbol{A}=trace\,\boldsymbol{A}$, and $\boldsymbol{F}_{\boldsymbol{M}}(\sigma_0^2, \boldsymbol{\gamma})\,vec\,\boldsymbol{A}=\sigma_0^4\,vec\,(2\boldsymbol{MAM}+\boldsymbol{M\Gamma}\cdot Diag\;\boldsymbol{A}\cdot\boldsymbol{M})$. ■
For the general Hsu-model, C. R. RAO [14, Th. 1] derives the following estimates. J. KLEFFE [7, Th. 2] gives similar representations.

Theorem 4.4. *Let a* Hsu*-model be given. For a fixed* $\boldsymbol{\sigma}_0^2$, *introduce the* (n, n)*-matrices* $\boldsymbol{V}_*=\sum\sigma_{0,\varkappa}^2\boldsymbol{V}_\varkappa=\boldsymbol{U}\boldsymbol{\Delta}_1\boldsymbol{U}'$, *and* $\boldsymbol{R}_*=(\boldsymbol{MV}_*\boldsymbol{M})^+$, *and the* (c, c)*-matrices* $\boldsymbol{M}_1=\boldsymbol{U}'\boldsymbol{R}_*\boldsymbol{U}$, $\boldsymbol{M}_2=\boldsymbol{M}_1 * \boldsymbol{M}_1$. *Let* $\boldsymbol{q}\in\mathrm{R}^k$. *If* $\boldsymbol{V}_*$ *is positive definite, and the* R^c*-vector* $\boldsymbol{\vartheta}$ *and the* R^k*-vector* $\boldsymbol{\lambda}$ *satisfy*

$$\boldsymbol{H\vartheta}=\boldsymbol{q}, \quad 2\boldsymbol{\vartheta}+\boldsymbol{M}_2\boldsymbol{\Delta}_2\boldsymbol{\vartheta}=\boldsymbol{M}_2\boldsymbol{H}'\boldsymbol{\lambda}\,, \tag{20}$$

then an unbiased IQE *of* $\boldsymbol{q}'\boldsymbol{t}$ *with minimum variance at* $(\boldsymbol{\sigma}_0^2, \boldsymbol{\gamma})$ *is given by*

$$\boldsymbol{Y}'\boldsymbol{R}_*\boldsymbol{U}\cdot Diag\,\frac{1}{2}\,(\boldsymbol{H}'\boldsymbol{\lambda}-\boldsymbol{\Delta}_2\boldsymbol{\vartheta})\cdot\boldsymbol{U}'\boldsymbol{R}_*\boldsymbol{Y}\,. \tag{21}$$

Proof. The assertion follows from Lemma 2.1.ii. Firstly, since $\boldsymbol{R}_*=\boldsymbol{R}_*\boldsymbol{M}$, we have $\boldsymbol{A}=\frac{1}{2}\boldsymbol{R}_*\boldsymbol{U}\cdot Diag\;(\boldsymbol{H}'\boldsymbol{\lambda}-\boldsymbol{\Delta}_2\boldsymbol{\vartheta})\cdot\boldsymbol{U}'\boldsymbol{R}_*$. Now, $vec\;\boldsymbol{U}\cdot Diag\;\boldsymbol{H}'\boldsymbol{\lambda}\cdot\boldsymbol{U}'=vec\sum\lambda_\varkappa\boldsymbol{V}_\varkappa=\boldsymbol{D\lambda}$, and, by (13), $vec\;\boldsymbol{\Delta}_2\,Diag\,\boldsymbol{\vartheta}=\boldsymbol{D}_c\boldsymbol{\Delta}_2\boldsymbol{\vartheta}$. Thus, from (12), $\boldsymbol{L}'=vec\boldsymbol{A}=\frac{1}{2}\boldsymbol{R}_*\otimes\boldsymbol{R}_*\cdot\boldsymbol{D\lambda}-\frac{1}{2}\boldsymbol{R}_*\boldsymbol{U}\otimes\boldsymbol{R}_*\boldsymbol{U}\cdot\boldsymbol{D}_c\boldsymbol{\Delta}_2\boldsymbol{\vartheta}$. Using $\boldsymbol{D}=\boldsymbol{U}\otimes\boldsymbol{U}\cdot\boldsymbol{D}_c\boldsymbol{H}'$ and Lemma 3.1, this yields $\boldsymbol{D}'_{\boldsymbol{M}}\boldsymbol{L}'=\frac{1}{2}\boldsymbol{H}\boldsymbol{D}'_c\cdot\boldsymbol{U}'\otimes\boldsymbol{U}'\cdot\boldsymbol{R}_*\otimes\boldsymbol{R}_*\cdot\boldsymbol{U}\otimes\boldsymbol{U}\cdot\boldsymbol{D}_c\boldsymbol{H}'\boldsymbol{\lambda}-\frac{1}{2}\boldsymbol{H}\boldsymbol{D}'_c.\,\boldsymbol{M}_1\otimes$

$\otimes \boldsymbol{M}_1 \cdot \boldsymbol{D}_c\boldsymbol{\Delta}_2\boldsymbol{\vartheta} = \frac{1}{2}\boldsymbol{H}\boldsymbol{M}_2\boldsymbol{H}'\boldsymbol{\lambda} - \frac{1}{2}\boldsymbol{H}\boldsymbol{M}_2\boldsymbol{\Delta}_2\boldsymbol{\vartheta}$, which equals $\boldsymbol{q}$ under the assumption (20).

Secondly, the positive definiteness of $\boldsymbol{V}_*$ implies $\boldsymbol{R}_*^+\boldsymbol{R}_* = \boldsymbol{M}$. Lemma 4.2.i, then, yields $\boldsymbol{F}_{\boldsymbol{M}}(\sigma_0^2, \gamma) \cdot \boldsymbol{L}' = \boldsymbol{M}\otimes\boldsymbol{M} \cdot \boldsymbol{D}\boldsymbol{\lambda} - \boldsymbol{M}\boldsymbol{U}\otimes\boldsymbol{M}\boldsymbol{U} \cdot \boldsymbol{D}_c\boldsymbol{\Delta}_2\boldsymbol{\vartheta} + \frac{1}{2}\boldsymbol{M}\boldsymbol{U}\otimes\boldsymbol{M}\boldsymbol{U} \cdot \boldsymbol{D}_c\boldsymbol{\Delta}_2\boldsymbol{D}_c' \cdot \boldsymbol{U}'\boldsymbol{R}_*\otimes\boldsymbol{U}'\boldsymbol{R}_* \cdot \boldsymbol{D}\boldsymbol{\lambda} - \frac{1}{2}\boldsymbol{M}\boldsymbol{U}\otimes\boldsymbol{M}\boldsymbol{U} \cdot \boldsymbol{D}_c\boldsymbol{\Delta}_2\boldsymbol{D}_c' \cdot \boldsymbol{U}'\boldsymbol{R}_*\boldsymbol{U}\otimes\boldsymbol{U}'\boldsymbol{R}_*\boldsymbol{U} \cdot \boldsymbol{D}_c\boldsymbol{\Delta}_2\boldsymbol{\vartheta} = \boldsymbol{D}_{\boldsymbol{M}}\boldsymbol{\lambda} - \frac{1}{2}\boldsymbol{M}\boldsymbol{U}\otimes\boldsymbol{M}\boldsymbol{U} \cdot \boldsymbol{D}_c\boldsymbol{\Delta}_2\,(2\boldsymbol{\vartheta} - \boldsymbol{M}_2\boldsymbol{H}'\boldsymbol{\lambda} + \boldsymbol{M}_2\boldsymbol{\Delta}_2\boldsymbol{\vartheta})$. Under the assumption (20), this is in $\Re\boldsymbol{D}_{\boldsymbol{M}}$. ■

Uniformity Criteria. Essentially, the last theorem of this section is also due to C. R. Rao [14, p. 453–454]. The first part characterizes those situations when the estimates are independent of the kurtosis γ; the second part assumes quasi normality, i.e., $\gamma = \boldsymbol{0}$, and investigates independence from σ_0^2.

Theorem 4.5. *Assume the* Hsu-*model and notation of Th.* 4.4. *Then:*
(i) *All* MV – UB – IQE*s at* $(\sigma_0^2, \boldsymbol{0})$ *are of minimum variance at* (σ_0^2, γ) *iff* $\Re\boldsymbol{D}_{\boldsymbol{M}}$ *is invariant under* $\boldsymbol{M}\boldsymbol{U}\otimes\boldsymbol{M}\boldsymbol{U} \cdot \boldsymbol{D}_c\boldsymbol{\Delta}_1^2 \cdot \mathit{Diag}\ \boldsymbol{H}'\gamma \cdot \boldsymbol{D}_c' \cdot \boldsymbol{U}'\boldsymbol{R}_*\otimes\boldsymbol{U}'\boldsymbol{R}_*$*, or, equivalently, iff* $\Re\boldsymbol{M}_2\boldsymbol{H}'$ *is invariant under* $\boldsymbol{M}_2\boldsymbol{\Delta}_1^2\ \mathit{Diag}\ \boldsymbol{H}'\gamma$.
(ii) *All* MV – UB – IQE*s at* $(\boldsymbol{1}_k, \boldsymbol{0})$ *are of minimum variance at* $(\sigma_0^2, \boldsymbol{0})$ *iff* $\Re\boldsymbol{D}_{\boldsymbol{M}}$ *is invariant under* $\boldsymbol{M}\boldsymbol{V}_*\boldsymbol{R}\otimes\boldsymbol{M}\boldsymbol{V}_*\boldsymbol{R}$*, or, equivalently, iff for every* $\boldsymbol{\lambda}\in\mathrm{R}^k$ *there exist a* $\boldsymbol{\mu}\in\mathrm{R}^k$ *such that* $\sum\lambda_\varkappa\boldsymbol{R}\boldsymbol{V}_*\boldsymbol{R}\boldsymbol{V}_\varkappa\boldsymbol{R}\boldsymbol{V}_*\boldsymbol{R} = \sum\mu_\varkappa\boldsymbol{R}\boldsymbol{V}_\varkappa\boldsymbol{R}$. *Here,* $\boldsymbol{V} = \sum\boldsymbol{V}_\varkappa$ *is assumed positive definite,* $\boldsymbol{R} = (\boldsymbol{M}\boldsymbol{V}\boldsymbol{M})^+$*, and* $\boldsymbol{V}_*$ *is assumed non-negative definite.*

Proof. (i) We have to check [11, p. 147] when $\Re\boldsymbol{F}_{\boldsymbol{M}}(\sigma_0^2, \gamma) \cdot \boldsymbol{N} \subset \Re\boldsymbol{F}_{\boldsymbol{M}}(\sigma_0^2, \boldsymbol{0}) \cdot \boldsymbol{N}$. This is the case, by Lemma 4.2.i and after premultiplying with $\boldsymbol{R}_*\otimes\boldsymbol{R}_*$, iff the range of $\boldsymbol{A} = \boldsymbol{R}_*\boldsymbol{U}\otimes\boldsymbol{R}_*\boldsymbol{U} \cdot \boldsymbol{D}_c\boldsymbol{\Delta}_2\boldsymbol{D}_c' \cdot \boldsymbol{U}'\boldsymbol{M}\otimes\boldsymbol{U}'\boldsymbol{M} \cdot \boldsymbol{N}$ is contained in $\Re\boldsymbol{M}\otimes\boldsymbol{M} \cdot \boldsymbol{N}$. But $\boldsymbol{M}\otimes\boldsymbol{M} \cdot \boldsymbol{N} = \boldsymbol{N} \cdot \boldsymbol{M}\otimes\boldsymbol{M}$ is a projector, so $\boldsymbol{N} \cdot \boldsymbol{M}\otimes\boldsymbol{M} \cdot \boldsymbol{A} = \boldsymbol{A}$ yields $\boldsymbol{D}_{\boldsymbol{M}}' \cdot \boldsymbol{R}_*\boldsymbol{U}\otimes\boldsymbol{R}_*\boldsymbol{U} \cdot \boldsymbol{D}_c\boldsymbol{\Delta}_2\boldsymbol{D}_c' \cdot \boldsymbol{U}'\boldsymbol{M}\otimes\boldsymbol{U}'\boldsymbol{M} \cdot \boldsymbol{N} = \boldsymbol{0}$. This is true iff

$$\Re\boldsymbol{M}\boldsymbol{U}\otimes\boldsymbol{M}\boldsymbol{U} \cdot \boldsymbol{D}_c'\boldsymbol{\Delta}_2\boldsymbol{D}_c \cdot \boldsymbol{U}'\boldsymbol{R}_*\otimes\boldsymbol{U}'\boldsymbol{R}_* \cdot \boldsymbol{D}_{\boldsymbol{M}} \subset \Re\boldsymbol{D}_{\boldsymbol{M}}, \tag{22}$$

which is the first characterization. Using $\boldsymbol{D} = \boldsymbol{U}\otimes\boldsymbol{U} \cdot \boldsymbol{D}_c\boldsymbol{H}'$ and premultiplying (22) with $\boldsymbol{D}_c \cdot \boldsymbol{U}'\boldsymbol{R}_*\otimes\boldsymbol{U}'\boldsymbol{R}_*$ yields the second characterization, [14, eq. 5.8].

For part (ii), check $\Re\boldsymbol{F}_{\boldsymbol{M}}(\sigma_0^2, \boldsymbol{0}) \cdot \boldsymbol{N} \subset \Re\boldsymbol{F}_{\boldsymbol{M}}(\boldsymbol{1}_k, \boldsymbol{0})$ analogously to get the first characterization. Then premultiply with $\boldsymbol{R}\otimes\boldsymbol{R}$, use $\boldsymbol{D}_{\boldsymbol{M}}\boldsymbol{\lambda} = \mathit{vec}\ \sum\lambda_\varkappa\boldsymbol{M}\boldsymbol{V}_\varkappa\boldsymbol{M}$ and (12) to get the second characterization [14, eq. 6.4]. ■

As indicated in the proof of part (ii), formulae (12) and (13) may be used to reformulate the above criteria in matrix space. The formulation above parallels that of mean estimation: The BLUE under $\boldsymbol{V}$ is BLUE under $\boldsymbol{V}_*$ iff $\Re\boldsymbol{X}$ is invariant under $\boldsymbol{V}_*\boldsymbol{V}^{-1}$ (cf., [11, p. 149]).

For further results on uniform MV – UB – IQE see H. Drygas [3] and S. Gnot, W. Klonecki and R. Zmyslony [4]. The dispersion-mean-correspondence may also be used to get MINQUE, weighted least squares, or Ridge-type estimates of variance components, cf., F. Pukelsheim [12], [13]. Maximum likelihood estimates are considered, e.g., by R. R. Corbeil and S. R. Searle [1].

5. Unbiased NND Quadratic Estimability

Though one should not dispense with requiring that a quadratic estimate $\boldsymbol{Y'AY}$ of a single variance component $\sigma_\varkappa^2$ be non-negative definite (NND), one does, and there are only few investigations of this subject, cf., H. Drygas [2], L. R. La Motte [9]. The next lemma is implicitly given by L. R. La Motte [9, p. 728], the formulation below is to stress that NND estimates shift the problem from linearity into convexity.

Lemma 5.1. *Assume a linear model as given by* (3), *let* $\boldsymbol{q} \in \mathrm{R}^k$. *Then*:
(i) *There exists an unbiased* IQE *of* $\boldsymbol{q't}$ *iff*
$\boldsymbol{q} \in$ *linear hull* $\{\boldsymbol{D}_M' \cdot \boldsymbol{y} \otimes \boldsymbol{y} \mid \boldsymbol{y} \in \mathrm{R}^n\}$.
(ii) *There exists an unbiased* NND *quadratic estimator of* $\boldsymbol{q't}$ *iff*
$\boldsymbol{q} \in$ *convex hull* $\{\boldsymbol{D}_M' \cdot \boldsymbol{y} \otimes \boldsymbol{y} \mid \boldsymbol{y} \in \mathrm{R}^n\}$.

Proof. Unbiasedness and non-negative definiteness imply invariance, see above Sect. 2. $\boldsymbol{Y'AY}$ is unbiased for $\boldsymbol{q't}$ iff $\boldsymbol{q} = \boldsymbol{D}_M' \cdot \mathit{vec}\, \boldsymbol{A}$. Assertions i, ii then follow from the spectral representation [15, p. 39] of symmetric matrices, NND matrices, respectively. ■

What has the model to look like such that a single component $t_\varkappa$ be unbiasedly NND estimable? For a pure variance components model, e.g., a Hsu-model, we finally prove as a necessary and sufficient condition: The $\varkappa$-th dispersion design must properly contribute to the explanation of the error space $\Re \boldsymbol{M} \sum \boldsymbol{V}_\varkappa \boldsymbol{M}$, (cf., [15, p. 297]).

Theorem 5.1. *For a linear model* (3), *let all* $\boldsymbol{V}_\varkappa$, $\varkappa = 1, \ldots, k$, *be* NND, *and fix* $\varkappa$. *Then there exists an unbiased* NND *quadratic estimator of* $t_\varkappa$ *iff*
$\Re \boldsymbol{M} \boldsymbol{V}_\varkappa \boldsymbol{M} \not\subset \Re \boldsymbol{M} \sum_{\lambda \neq \varkappa} \boldsymbol{V}_\lambda \boldsymbol{M}$.

Proof. Put $\mathfrak{K} = \mathit{convex\ hull}\ \{\boldsymbol{D}_M' \cdot \boldsymbol{y} \otimes \boldsymbol{y} \mid \boldsymbol{y} \in R^n\}$. Clearly, $\mathfrak{K} = \{\sum \boldsymbol{D}_M' \cdot \boldsymbol{y}_\nu \otimes \boldsymbol{y}_\nu / \boldsymbol{y}_1, \ldots, \boldsymbol{y}_n \in R^n\}$. With all $\boldsymbol{V}_\varkappa$ NND, it is easily shown that $\boldsymbol{e}_\varkappa \in \mathfrak{K}$ iff $\boldsymbol{e}_\varkappa \in \{\boldsymbol{D}_M' \cdot \boldsymbol{y} \otimes \boldsymbol{y} / \boldsymbol{y} \in R^n\}$. The latter means that the nullspace of $\boldsymbol{M} \sum_{\lambda \neq \varkappa} \boldsymbol{V}_\lambda \boldsymbol{M}$ be not contained in the nullspace of $\boldsymbol{M} \boldsymbol{V}_\varkappa \boldsymbol{M}$. This is the orthogonal dual of the assertion. ■

Acknowledgement

The author would like to thank Prof. H. Witting for his encouragement and permanent support. He is also grateful to Prof. H. Drygas and the referee for suggesting improvements and extensions.

References

[1] Corbeil, R. R., and S. R. Searle (1976). Restricted maximum likelihood (REML) estimation of variance components in the mixed model. *Technometrics* **18**, 31–38.
[2] Drygas, H. (1972). The estimation of residual variance in regression analysis. *Math. Operationsforsch. Statist.* **3**, 373–388.

[3] DRYGAS, H. (1977). Best quadratic unbiased estimation in variance-covariance component models. *Math. Operationsforsch. Statist. Ser. Statistics* **8**, 211–231.

[4] GNOT, S., W. KLONECKI and R. ZMYSLONY (1977). Uniformly minimum variance unbiased estimation in various classes of estimators, I. *Math. Operationsforsch. Statist. Ser. Statistics* 8, 199–210.

[5] GREUB, W. H. (1967). *Multilinear Algebra.* Springer-Verlag, Berlin.

[6] HSU, P. L. (1938). On the best quadratic estimate of variance. *Statist. Res. Mem.* **2**, 91–104.

[7] KLEFFE, J. (1977). Invariant methods for estimating variance components in mixed linear models. *Math. Operationsforsch. Statist. Ser. Statistics* 8, 233–250.

[8] KLEFFE, J., and R. PINCUS (1974). BAYES and best quadratic unbiased estimators for variance components and heteroscedastic variances in linear models. *Math. Operationsforsch. Statist.* **5**, 147–159.

[9] LA MOTTE, L. R. (1973). On non-negative quadratic unbiased estimation of variance components. *J. Amer. Statist. Assoc.* **68**, 728–730.

[10] MITRA, S. K. (1971). Another look at RAO's MINQUE of variance components. *Bull. Inst. Internat. Statist.* **44**, 2, 279–283.

[11] MITRA, S. K., and B. J. MOORE (1973). GAUSS-MARKOV estimation with an incorrect dispersion matrix. *Sankhyá Ser. A* **35**, 139–152.

[12] PUKELSHEIM, F. (1977). Estimating variance components in linear models. *J. Multivariate Anal.* **6**, 626–629.

[13] PUKELSHEIM, F. (1977). Equality of two BLUEs and Ridge-type estimates. *Comm. Statist.-Theory Methods* A **6**, 603–610.

[14] RAO, C. R. (1971). Minimum variance quadratic unbiased estimation of variance components. *J. Multivariate Anal.* **1**, 445–456.

[15] RAO, C. R. (1973). *Linear Statistical Inference and its Applications, Second Edition.* John Wiley & Sons, New York.

[16] SEELY, J. (1972). Completeness for a family of multivariate normal distribution. *Ann. Math. Statist.* **43**, 1644–1647.

[17] STYAN, G. P. H. (1973). HADAMARD products and multivariate statistical analysis. *Linear Algebra and Appl.* **6**, 217–240.

[18] ZMYSLONY, R. (1973). Estimation of variance components in random models. *Zastos. Mat.* **13**, 521–527.

Zusammenfassung

In der Theorie der linearen Modelle kann man die Schätzung von Varianzkomponenten vollständig aus der Schätztheorie für den Mittelwert herleiten; diese '*Streuungs-Mittelwert-Korrespondenz*' wird am Beispiel des HSU-Modells verdeutlicht. Sie ergibt kurze neue Beweise für verschiedene in der Literatur vorkommende Sätze über erwartungstreue invariante quadratische Schätzer mit kleinstem BAYES-Risiko bzw. mit kleinster Varianz. Bei reinen Varianzkomponenten-Modellen wird schließlich erwartungstreue positiv-semidefinite quadratische Schätzbarkeit charakterisiert an Hand der Designmatrizen, die das Modell definieren.

Résumé

Dans ce travail on démontre de nouveau quelques théorèmes concernant l'estimation des composants de la variance dans le modèle linéaire de HSU: Par l'application d'une '*correspondance dispersion-moyenne*' la question posée se trouve réduite aux problèmes déjà connus et résolus dans la théorie d'estimation de la moyenne. Dans les théorèmes cités on étudie des estimateurs quadratiques, invariants, sans biais, qui minimisent soit le risque bayesien soit la variance. En outre, l'existence d'un estimateur quadratique, non-negatif, sans biais d'un seul composant de la variance est characterisée par des matrices qui, elles, déterminent le modèle.

Received July 1975. Revised May 1976.