PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 18, Fasc. 1 (1998), pp. 33–38

ON ADMISSIBILITY IN ESTIMATING THE MEAN SQUARED ERROR OF A LINEAR ESTIMATOR*

BY

CZESŁAW STĘPNIAK (LUBLIN)

Abstract. Consider a linear estimator of a parametric vector $C\beta$ in the normal Gauss–Markov model $Y \sim \mathcal{N}(X\beta, \sigma V)$. The Mean Squared Error of such an estimator may be presented in the form $k\sigma + \beta' X' KX\beta$ and may be estimated by quadratics in Y. Some basic decision-theoretic questions in the estimation are discussed. Among others, some admissible estimators of the MSE in several classes of the quadratics are given.

1. Introduction. Most of the papers on quadratic estimation refer to the variance components only. Such a problem can be considered with respect to various classes of quadratics. Several authors, among others Klotz, Milton and Zacks [6] and Harville [5], noted that some customarily used ANOVA estimators may be dominated by some biased estimators. Now it is well known (see, e.g., Perlman [9], LaMotte [8], and Stępniak [12]) that no quadratic unbiased estimator of variance components is admissible among all quadratics. The second bad property of the unbiased quadratics was shown by LaMotte [7]. Namely, no unbiased quadratic estimator of a variance component, except the error component, is nonnegative definite. In the past decade some papers on quadratic estimation went beyond this classical framework (e.g. [1]–[4] and [10]).

Let $C\beta$ be an estimable parametric vector in the normal Gauss–Markov model $Y \sim \mathcal{N}(X\beta, \sigma V)$ and let DY be an arbitrary linear estimator of this vector. The Mean Squared Error of DY may be considered as a parametric function of the form

$$\psi = k\sigma + \beta' X' KX\beta.$$

We can be interested in estimating the parameter ψ by quadratic forms in Y. It appears that some other statistical problems such as variable selection or estimating the weighted length of β can be presented in such a form (see, e.g.,

* Supported by KBN grant 2 P03A 044 12.

Gnot et al. [3] and Grządziel [4]). In this context it would be interesting to know whether, as for variance components, any quadratic unbiased estimator of the function ψ is inadmissible among all quadratics.

In this paper we answer both this and other questions of this type. Moreover, some admissible estimators of ψ among all quadratics and among nonnegative quadratics are given.

2. Preliminaries. Let us consider a normal Gauss–Markov model

$$Y \sim \mathcal{N}(X\beta, \sigma V), \tag{1}$$

where X is a known $n \times p$ matrix and V is a symmetric nonnegative definite matrix of order n such that $r(V+XX') > r(X)$, while $\beta \in R^p$ and $\sigma > 0$ are unknown parameters. Then there exists a projector $P = P_X$ on the column space of X and a quadratic form $s = Y'SY$ such that

(A) PY is the Best Linear Unbiased Estimator of the mean vector $\mu = X\beta$;

(B) s is the Best Quadratic Unbiased Estimator of the parameter σ.

Moreover, the estimators $\hat{\mu} = PY$ and $\hat{\sigma} = s$ are stochastically independent and jointly constitute a complete sufficient statistic in the model (1) (see, e.g., Seely [11], Theorem 2.1 and Corollary 2.2). Thus $\hat{\mu}$ and s are the Uniformly Minimum Variance Unbiased Estimators of their expectations.

Now let us consider a parametric vector $\theta = C\beta$ and let $\hat{\theta} = DY$ be an arbitrary linear estimator of θ. The Mean Squared Error of $\hat{\theta}$ is defined by

$$\mathrm{MSE}(\hat{\theta}/\beta, \sigma) = E(\hat{\theta}-\theta)'(\hat{\theta}-\theta). \tag{2}$$

If $C\beta$ is linearly estimable, then the matrix C can be presented in the form $C = BX$ for some B and the MSE can be written as a parametric function

$$\psi = \psi(\beta, \sigma) = k\sigma + \beta' X' KX\beta, \tag{3}$$

where k is a scalar and K is a symmetric matrix of order n. Without loss of generality we may and shall assume that $P'KP = K$.

We are interested in estimating the parameter ψ by quadratic forms $Y'QY$ under the usual risk

$$R(Y'QY/\beta, \sigma) = E(Y'QY-\psi)^2. \tag{4}$$

Since the expression (4) involves only first two moments of the quadratic form $Y'QY$, the initial normality assumption can be weakened to

$$E(Y'QY) = \sigma \operatorname{tr}(VQ) + \beta' X'QX\beta$$

and

$$\operatorname{var}(Y'QY) = 2\sigma^2 \operatorname{tr}(VQVQ) + 4\sigma\beta' X'QVQX\beta$$

for any symmetric matrix Q.

Our problem may be considered with respect to various classes of quadratics (cf. Stępniak [12]). The following ones are standard:

$$\mathscr{C}_0 = \{Y'QY\colon\ Q \text{ is an arbitrary symmetric matrix of order } n\},$$

$$\mathscr{C}_1 = \{Y'QY\colon\ E(Y'QY) = \psi\}, \quad \mathscr{C}_2 = \{Y'QY\colon\ Q \text{ is nonnegative definite}\}.$$

Considering the quadratic form $\hat{\mu}'K\hat{\mu}$, where $\hat{\mu}$ is the BLUE of μ, we see that

$$E(\hat{\mu}'K\hat{\mu}) = \sigma\,\mathrm{tr}(VK) + \beta'X'KX\beta.$$

Thus, by (B) and by the completeness of $(\hat{\mu}, s)$, the parametric function (3) is quadratically estimable and

$$\hat{\psi}_1 = [k - \mathrm{tr}(VK)]\,s + \hat{\mu}'K\hat{\mu} \tag{5}$$

is its BQUE.

In particular, if DY is an unbiased estimator of θ, then (3) reduces to $k\sigma$ and the BQUE ks of the MSE of DY is dominated by the biased quadratic $d(d+2)^{-1}ks$, where $d = r(V+XX') - r(X)$ is the number of the freedom degrees for error in the model (1). In general, the parametric function (3) involves also a term depending on β. Then the estimator (5) is dominated by

$$\hat{\psi}_0 = \frac{d}{d+2}[k - \mathrm{tr}(VK)]\,s + \hat{\mu}'K\hat{\mu}.$$

We note that the MSE defined by (2) is a nonnegative function of β and σ, while the quadratic form $\hat{\psi}_0$ is not nonnegative definite unless $k \geqslant \mathrm{tr}(VK)$. If $k < \mathrm{tr}(VK)$, then the estimator $\hat{\psi}_0$ can be replaced by a nonnegative definite quadratic

$$\hat{\psi}_2 = \hat{\mu}'K\hat{\mu}.$$

In the next section we shall prove that the quadratic forms $\hat{\psi}_0$ and $\hat{\psi}_2$ are admissible estimators of the Mean Squared Error (2) under the risk (4) in the classes $\mathscr{C}_0$ and $\mathscr{C}_2$, respectively.

3. Results. Let us start with some auxiliary results.

PROPOSITION 1. *Let $Y'QY$ and $Y'TY$ be arbitrary quadratic estimators of a parametric function $\psi = k\sigma + \beta'X'KX\beta$ such that $X'QX = X'KX$ and*

$$R(Y'TY/\beta, \sigma) \leqslant R(Y'QY/\beta, \sigma) \tag{6}$$

for all $\beta \in R^p$ and $\sigma > 0$. Then $X'TX = X'KX$.

Proof. We observe that

$$R(Y'QY/\beta, \sigma) = E(Y'QY - \psi)^2 = \mathrm{var}(Y'QY) + [E(Y'QY) - \psi]^2$$
$$= \mathrm{var}(Y'QY) + [\mathrm{tr}(VQ) - k]^2\sigma^2$$

and

$$R(Y'TY/\beta, \sigma) = \mathrm{var}(Y'TY) + [E(Y'TY) - \psi]^2$$
$$= \mathrm{var}(Y'TY) + \big[(\mathrm{tr}(VT) - k)\sigma + \beta' X'(T-K)X\beta\big]^2.$$

Now for fixed but arbitrary $\beta_0 \in R^p$ and $\sigma_0 > 0$ let us consider a real function

$$f(x) = R(Y'QY/x\beta_0, \sigma_0) - R(Y'TY/x\beta_0, \sigma_0).$$

This function can be written in the form

$$f(x) = ax^4 + bx^2 + c, \quad \text{where } a = -[\beta_0' X'(T-K)X\beta_0]^2.$$

Now it is clear that the condition (6) implies the desired result $X'TX = X'KX$. ∎

As a direct consequence of the proposition we get

COROLLARY 1. *A quadratic form $Y'QY$ is an admissible estimator of its expectation in the class $\mathscr{C}_0$ if and only if it is admissible in the class $\{Y'TY\colon X'TX = X'QX\}$.*

In the light of the above, the estimation of a parametric function $\psi = k\sigma + \beta' X'KX\beta$ is worth considering in an auxiliary class

$$\mathscr{C}_3 = \{Y'QY\colon X'QX = X'KX\}.$$

We shall prove

PROPOSITION 2. *In the class $\mathscr{C}_3$ the best estimator of the parametric function ψ exists and this estimator can be presented in the form*

$$\hat{\psi}_0 = \frac{d}{d+2}[k - \mathrm{tr}(VK)]\, s + \hat{\mu}' K \hat{\mu}, \tag{7}$$

where $d = r(V+XX') - r(X)$.

Proof. By the completeness of $(\hat{\mu}, s)$ for any quadratic estimator $Y'QY$ belonging to $\mathscr{C}_3$ there exists an estimator $cs + \hat{\mu}'K\hat{\mu}$, where $c \in R$, such that

$$E(cs + \hat{\mu}'K\hat{\mu} - \psi)^2 \leqslant E(Y'QY - \psi)^2.$$

Moreover,

$$E(cs + \hat{\mu}'K\hat{\mu} - \psi)^2 = c^2 \mathrm{var}\, s + [E(cs + \hat{\mu}'K\hat{\mu}) - k\sigma - \beta' X'KX\beta]^2 + \mathrm{var}(\hat{\mu}'K\hat{\mu})$$
$$= \frac{2c^2\sigma^2}{d} + (c-b)^2\sigma^2 + \mathrm{var}(\hat{\mu}'K\hat{\mu})$$
$$= \left(\frac{d+2}{d}c^2 - 2bc\right)\sigma^2 + b^2\sigma^2 + \mathrm{var}(\hat{\mu}'K\hat{\mu}),$$

where $b = k - \mathrm{tr}(VK)$. Let us consider a real function

$$g(x) = \frac{d+2}{d}x^2 - 2bx. \tag{8}$$

Its minimum is attained for

$$x = \frac{bd}{d+2} = \frac{d}{d+2}[k - \mathrm{tr}(VK)].$$

This implies the desired result. ∎

In consequence of the proposition we get

COROLLARY 2. *A quadratic form $Y'QY$ is an admissible estimator of its expectation in the model* (1) *if and only if $Y'QY = \hat{\mu}'K\hat{\mu}$ for some symmetric matrix K. In particular, the BQUE of the MSE of DY is inadmissible unless* $\mathrm{var}(D'Y) = 0$.

COROLLARY 3. *If $\psi = k\sigma$, then*

$$\hat{\psi}_0 = d(d+2)^{-1}ks$$

is its best invariant quadratic estimator.

The main results of this paper are contained in Theorems 1 and 2.

THEOREM 1. *The quadratic form $\hat{\psi}_0$ defined by* (7) *is an admissible estimator of the parametric function $\psi = k\sigma + \beta'X'KX\beta$ in the class $\mathscr{C}_0$.*

The theorem follows directly by Propositions 1 and 2. ∎

Now let us assume that the parametric function ψ is nonnegative (for instance, if ψ is the MSE of a linear estimator). Then it is natural to seek for an estimator of ψ in the class $\mathscr{C}_2$. We note that the quadratic $\hat{\psi}_0$ defined by (7) does not belong to the class if $k < \mathrm{tr}(VK)$.

THEOREM 2. *If ψ is a nonnegative parametric function of the form $\psi = k\sigma + \beta'X'KX\beta$ such that $k < \mathrm{tr}(VK)$, then the quadratic form*

$$\hat{\psi}_2 = \hat{\mu}'K\hat{\mu} \tag{9}$$

is an admissible nonnegative estimator of ψ.

Proof. We need only to observe that if $b = k - \mathrm{tr}(VK) < 0$, then the function $g(x)$ defined by (8) is increasing for $x \geqslant 0$ and, therefore, its conditional minimum under $x \geqslant 0$ is attained for $x = 0$. ∎

Acknowledgement. Thanks are due to a referee who drew my attention to the works [1]–[4] and [10] and contributed some valuable comments which improved presentation.

REFERENCES

[1] J. O. Berger and C. Robert, *Subjective hierarchical Bayes estimation of a multivariate normal mean*, Ann. Statist. 18 (1990), pp. 617–651.

[2] J. R. Brook and T. Moore, *On the expected length of the least squares coefficient vector*, J. Econometrics 12 (1980), pp. 245–246.

[3] S. Gnot, G. Trenkler and R. Zmyślony, *Nonnegative minimum biased quadratic estimation in the linear regression models*, J. Multivariate Anal. 54 (1995), pp. 113–125.

[4] M. Grządziel, *A simulation study of a new variable selection strategy based on minimum biased quadratic estimation*, Discussiones Mathematicae. Algebra and Stochastic Methods 15 (1995), pp. 325–333.

[5] D. A. Harville, *Alternative formulations and procedures for the two-way mixed model*, Biometrics 34 (1978), pp. 441–453.

[6] J. H. Klotz, R. C. Milton and S. Zacks, *Mean square efficiency of estimators of variance components*, J. Amer. Statist. Assoc. 64 (1969), pp. 1383–1402.

[7] L. R. LaMotte, *On non-negative quadratic unbiased estimation of variance components*, ibidem 68 (1973), pp. 728–730.

[8] — *Some results on biased linear estimation applied to variance components estimation*, in: *Mathematical Statistics and Probability: Proceedings. Sixth International Conference, Wisla, Poland*, Springer, New York 1980, pp. 266–274.

[9] M. D. Perlman, *Reduced mean square error estimation for several parameters*, Sankhyā B 34 (1972), pp. 89–92.

[10] A. L. Rukhin, *Quadratic estimators of quadratic functions of normal parameters*, J. Statist. Plann. Inference 15 (1987), pp. 301–310.

[11] J. Seely, *A complete sufficient statistic for the linear model under normality and singular covariance matrix*, Comm. Statist. Theory Methods A 7 (1978), pp. 1465–1473.

[12] C. Stępniak, *On admissibility in various classes of quadratic estimators*, Statist. Probab. Lett. 4 (1986), pp. 17–19.

Department of Mathematics
Agricultural University of Lublin
P.O. Box 158
PL-20-950 Lublin, Poland

Received on 30.9.1996;
revised version on 1.9.1997