Władysław Milo[*], Zbigniew Wasilewski[**]

ON THE EFFICIENCY OF WEIGHTED LEAST SQUARES ESTIMATORS
IN THE CASE OF A GENERAL LINEAR MODEL

1. Introduction

The concept of "efficient estimates (estimators)" was intro-
duced by R. F i s h e r (cf. [3]) for denoting consistent a-
symptotically normal (CAN) estimators with the asymptotically mi-
nimal variance.

Fisher's reasoning consisted in showing that

a) if $\tilde{\theta}_{(n)}$ is the maximum likelihood estimator (MLE) of the
parameter θ, then under the following regularity conditions:

a_1) the density function $f_Y(y, \theta)$ of the distribution function F_Y
of a random variable Y is two-fold differentiable in O.

a_2) the function $\frac{\partial^2}{\partial\theta^2}$ log $f_Y(y, \theta)$ is uniformly continuous in y
at $0 = 0_o$,

the quantity $\sqrt{n}\,(\tilde{\theta}_{(n)} - \theta)$ has asymptotically normal distribution
with the parameters $(0, I^{-1}(\theta))$, where $I(\theta) = \mathcal{E}\left[\frac{1}{f_Y}\,\frac{\partial f_Y}{\partial\theta}\right]$ denotes
an information quantity given by a sample about θ, and $\mathcal{E}$ denotes
an expectation operator.

b) if $\left\{T_{(n)}\right\}$ denotes a sequence of asymptotically normal e-
stimates, then

$$\lim_{n\to\infty} \mathcal{E}_\theta(\sqrt{n}\,(T_{(n)} - \theta))^2 \geqslant I^{-1}(\theta), \theta \in R^1,$$

where R^1 is the set of real numbers.

---

[*]Dr., Lecturer at the Institute of Econometrics and Sta-
tistics, University of Łódź.
[**]Senior Assistant at the Institute of Econometrics and Sta-
tistics, University of Łódź.

If (a) and (b) hold, then the MLE $\tilde{\theta}_{(n)}$ should be considered as the best asymptotically normal (BAN) estimators in the class of all asymptotically normal estimators.

Unfortunately, there are no estimators with the minimal variance. It results from the following facts. Let $\{T_{(n)}\}$ be any sequence of estimators and the variable $\sqrt{n}(T_{(n)} - \theta)$ be asymptotically normal with the parameters $(0, \sigma^2(\theta))$. Let $\sigma^2(\theta_0) \neq \neq 0$ be for the fixed value of θ_0. For

(1)
$$\tilde{T}_{(n)} = \begin{cases} T_{(n)} & \text{if } |T_{(n)} - \theta_0| > n^{-\frac{1}{4}}, \\ \theta_0 & \text{if } |T_{(n)} - \theta_0| \leqslant n^{-\frac{1}{4}}, \end{cases}$$

one can check that the sequence of $\sqrt{n}(\tilde{T}_{(n)} - \theta)$ is asymptotically normal with the parameters $(\theta, \partial^2(\theta))$, where

$$\partial^2(\theta) = \sigma^2(\theta) \quad \text{if } \theta \neq \theta_0,$$

$$\partial^2(\theta) = 0 < \sigma^2(\theta_0) \quad \text{if } \theta = \theta_0.$$

Using the above way of estimator improvement for the ML estimators in the regularity case, one can construct a sequence $\{T_{(n)}\}$ of asymptotically normal estimators such that

(2)
$$\lim_{n \to \infty} \mathcal{E}_\theta \left(\sqrt{n}(T_{(n)} - \theta)\right)^2 \leqslant I^{-1}(\theta),$$

and, for some θ, even

(2a)
$$\lim_{n \to \infty} \mathcal{E}_\theta \left(\sqrt{n}(T_{(n)} - \theta)\right)^2 < I^{-1}(\theta).$$

The estimators $\{T_{(n)}\}$ that fulfil (2) and/or (2a) for $\theta \in \Theta$ are called superefficient for convex loss functions and these θ for which (2a) holds are called superefficiency points.

The first known to us improvement of the type (1) was presented by J. H o d g e s (cf. [7]) and it was concerned with the case when $Y_1, \ldots, Y_n, \ldots,$ were i.i.d. normal variables with var $(Y_i) = 1$, $\theta \in R^1$, and

$$\tilde{T}_{(n)} = \begin{cases} \alpha \bar{Y}_{(n)} & \text{if } |\bar{Y}_{(n)}| \leqslant n^{-1/4}, \quad |\alpha| < 1, \\[2mm] \bar{Y}_{(n)} & \text{if } |\bar{Y}_{(n)}| > n^{-1/4}, \quad \bar{Y}_{(n)} \text{ is the MLE of } \theta, \end{cases}$$

where $\tilde{T}_{(n)}$ is superefficient at $\theta = 0$.

These new facts have stimulated needs for modifications of Fisher's definition of efficiency.

We recall the following modifications:

m1) an estimator $T_{(n)}$ is said to be asymptotically efficient for a parametric function $g(\theta)$ iff:

$$\lim_{n \to \infty} \mathcal{E}(T_{(n)}) = g(\theta), \qquad \lim_{n \to \infty} \left| \text{var}(T_{(n)}) - \frac{[g'(\theta)]^2}{I(\theta)} \right| = 0,$$

unless $\mathcal{E}(T_{(n)})$, $\text{var}(T_{(n)})$ do not exist;

m2) a family of estimators $\tilde{\theta}_\mathcal{E}$ is said to be $w_\mathcal{E}$ - asymptotically efficient within the subspace $K \subset \theta \subset R^1$ (asymptotically efficient with respect to $\{w_\mathcal{E}\}$) if for each non-void open set $K_0 \subset K$ the following relation

(3) $\lim_{\mathcal{E} \to 0} \left[\inf_{T_\mathcal{E}} \sup_{\theta \in K_0} \mathcal{E}_\theta^{(\mathcal{E})} w_\mathcal{E}(T_\mathcal{E} - \theta) - \sup_{\theta \in K_0} \mathcal{E}_\theta^{(\mathcal{E})} w_\mathcal{E}(\tilde{\theta}_\mathcal{E} - \theta) \right] = 0,$

holds, where $T_\mathcal{E}$ is any estimator, and $\mathcal{E}$ is equal for example $\frac{1}{n}$;

m3) a consistent estimator $T_{(n)}$ of 0 is said to be first order efficient (f.o.e.) iff

(4) $\lim_{n \to \infty} \sqrt{n} \left| T_{(n)} - \theta - \alpha(\theta) z_n \right| = 0,$

in probability, where α does not depend on observations and

$$z_n = \frac{1}{n} \frac{\partial \log f_y(y;\theta)}{\partial \theta}.$$

All presented definitions of efficiency are qualitative in nature and deal with the asymptotic behaviour of estimators. They can be generalized into the case of $\theta \subset R^k$, R^k being the Eu-

clidean space with the dimension k, by including all the restri-
ctions implied by the multi-dimensionality of the considered pro-
blems. Though the whole asymptotics of efficiency is indispensa-
able (both in R^1 and R^k) as a background knowledge for small sam-
ple analysis of efficiency, in practice we need some measures of
efficiency for small samples, and the knowledge how they behave
in dependence on changes of assumptions which underlie the mo-
del generating the observations Y.

The purpose of this paper is to present one of possible
ways óf measuring the efficiency and to analyzé sqme propert-
ies of the presented determinantal efficiency measure.

In § 2 we present an analysis of the properties of weight-
ed estimators in the case of a general linear model.

In § 3 we present a determinantal efficiency measure and
prove that its range belongs to $<0,1> \subset R_+^1$.

In § 4 we derive lower bounds of the determinantal efficie-
ncy measure.

2. Some Properties of Weighted Estimators
in the Case of a General Linear Model

By a general linear model we understand the following model:

$$\mathscr{NM}_\Omega = (\mathscr{R}^{n \times k}, \mathscr{S}, Y = X\beta + \Sigma, k_0 = k, n_0 = n, \mathscr{P}_Y = \mathscr{N}_Y(X\beta, \Omega)),$$

where:
  $R^{n \times k}$ - the set of nxk real matrices,
  $\mathscr{S}$ - a probability space, $\mathscr{S} = (\mathcal{U}, \mathscr{F}, \mathscr{P})$,
  $\mathcal{U}$ - a set of elementary events,
  $\mathscr{F}$ - the σ-Borel field of $\mathcal{U}$ subsets,
  $\mathscr{P}$ - a complete probability measure,
  $X\beta = \mathscr{E}(Y)$, $\Omega = \mathscr{D}(Y) = \mathscr{E}(Y - \mathscr{E}(Y))(Y - \mathscr{E}(y))^* = \mathscr{D}(\Sigma)$,
  $X \in R^{n \times k}$, $\beta \in \mathscr{R}^k$,
"$\mathscr{P}_Y = \mathscr{N}_Y(X\beta, \Omega)$" $\overset{df}{=}$ "the probability distribution of Y is n-
-dimensional normal distribution with $\mathscr{E}(Y) = X\beta$, $\mathscr{D}(Y) = \Omega = \sigma^2 G$".

One of the possible estimation quality functionals for the
model $\mathscr{NM}_\Omega$ is of the form

(5) $$\varphi = \| \Omega^{-1/2} (Y - X\beta) \|^2,$$

where $\| \cdot \|$ denotes Euclidean norm in R^n.

About Ω we assume further

A1) Ω is nonsingular positive definite real matrix, i.e.

$$\Omega \in R^{n \times n}, \qquad \det \Omega \neq 0.$$

By convexity of φ and (A1) it is easy to find that:

(6) $$\arg \min_{\beta} \| \Omega^{-1/2} (Y - X\beta) \|^2 = \hat{\beta} = (X'\Omega^{-1}X)^{-1} X'\Omega^{-1}Y$$

and by the definition of Y and assumptions of $\mathcal{N}\mathcal{M}_\Omega$ we get

(7) $$\mathcal{S}(\hat{\beta}) = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}\mathcal{S}(Y)\Omega^{-1}X(X'\Omega^{-1}X)^{-1} = (X'\Omega^{-1}X)^{-1}$$

Denoting

$$L = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}, \quad L_1 = L + C, \quad C \in R^{k \times n}$$

we have

(8) $$\hat{\beta} = LY,$$

for the Ω-weighted least-squares (Ω-WLS) estimator, and

(8a) $$\hat{\beta}_1 = L_1 Y,$$

for any other weighted least squares estimator.

To be unbiased the estimator $\hat{\beta}_1$ must fulfil $\mathcal{E}(\hat{\beta}_1) = L_1 X \beta = \beta$, i.e. $L_1 X = I_{(k)}$. The last relation holds iff $CX = 0$. By the last condition the estimator $\hat{\beta}_1$ can be written as

(8b) $$\hat{\beta}_1 = \beta + L_1 Z.$$

By (8b) and the properties of the dispersion operator $\mathcal{S}$ we have

$$\mathcal{S}(\hat{\beta}_1) = L_1 \Omega L_1' = L\Omega L' + L\Omega C' + C\Omega L' + C\Omega C'.$$

Since $L\Omega C' = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}\Omega C' = (X'\Omega^{-1}X)^{-1}X'C'$, then by $CX = 0$ we obtain

$$L \cap C' = 0, \text{ and by the same arguments } C \cap L' = 0.$$

Thus

(9) $$\mathcal{B}(\hat{B}_1) = L \cap L' + C \cap C'.$$

By (7) and (9) we obtain

(10) $$\mathcal{B}(\hat{B}_1) - \mathcal{B}(\hat{B}) = C \cap C'.$$

By (A1) and properties of Gram matrices, the matrix $C \cap C'$ is also positive definite iff rank C = rank $X = k_o$.

We have proved therefore

T h e o r e m   1. Under the model $\mathcal{N} \mathcal{M}_\Omega$, the linear esti-mator is efficient in the sense that any other linear estimator $\hat{B}_1 = (L + C)Y$ has the dispersion matrix $\mathcal{B}(\hat{B}_1)$ defined by (9) and such that $\mathcal{B}(\hat{B}_1) > \mathcal{B}(\hat{B})$, i.e. $\mathcal{B}(\hat{B}_1) - \mathcal{B}(\hat{B})$ is a positive definite matrix. ◆

The estimator $\hat{B}$ is the Ω-WLS estimator with the weight ma-trix Ω. It can be used only if this matrix is exactly known in practice. In most practical situations we do not know the matrix but its approximation, i.e. the matrix $\Omega_{\gamma\delta}$, where γ runs the indices of estimators of autocorrelation coefficients of ρ, and δ runs the indices of autocorrelation schemes. The concrete form of $\Omega_{\gamma\delta}$ depends on the assumed autocorrelation sche-mes about the components of the vector Ξ as well as estimation methods for the autocorrelation coefficient ρ in a specified autocorrelation scheme. The most often used schemes in practi-cal econometric and statistical applications are as follows:
  - first order autoregressive schemes,
  - second order autoregressive schemes,
  - fourth order autoregressive schemes (for a quarterly data),
  - first order moving-average scheme,
  - combining autoregressive moving-average scheme.

The above mentioned types of schemes, as well as other sche-mes, form the first criterion of the differentiation of Ω_α. The second criterion is formed by different types of estimation me-thods for the autocorrelation coefficient ρ, i.e. for example, the following estimators:

a) the sample first order autocorrelation coefficient (or the sample correlation between the successive residuals)

$$\hat{\rho}_1 = \frac{\sum\limits_{t=2}^{n} E_t E_{t-1}}{\sum\limits_{t=1}^{n} E_t^2},$$

where Σ_t (unobservable) were replaced by the LS residuals $E_t =$ $= Y_t - X_t^{'} \beta$, $X_t^{'} = (X_{t1}, \ldots, X_{tk})$, $B = (X^{'}X)^{-1}X^{'}Y$, $t = 1,\ldots,$ $n = 1, n;$

b) Theil's modification of $\hat{\rho}_1$:

$$\hat{\rho}_1^* = \frac{(n-k) \sum\limits_{t=2}^{n} E_t E_{t-1}}{(n-1) \sum\limits_{t=1}^{n} E_t^2};$$

c) the estimator

$$\tilde{\rho}_1 = 1 - \frac{1}{d}, \quad d = \sum\limits_{t=2}^{n}(E_t - E_{t-1})^2 / \sum\limits_{t=1}^{n} E_t^2;$$

d) Theil's-Nagar's modification of $\tilde{\rho}_1$, i.e.

$$\tilde{\rho}_1^* = \frac{n^2 \left(1 - \frac{d}{2}\right) + k^2}{n^2 - k^2},$$

e) Dent's adjustment of Theil-Nagar estimator

$$\tilde{\rho}_1^{**} = \frac{1}{\vartheta} (\tilde{\rho}_1^* - u),$$

where:

$$u = \frac{1}{n^2 - k^2}\left[k^2 - \frac{n^2}{n - k}(m - 1)\right], \quad \vartheta = \frac{n^2}{n^2 - k^2}\left(1 + \frac{m - 1}{n - k}\right),$$

$$m = tr\left[(X'HX)(X'X)^{-1}\right],$$

H is the matrix from an approximation of the inverse of $\sigma^2 F =$
$= \Omega$, i.e. $H : F^{-1} = (1 + \rho^2)I_{(n)} - 2\rho H$.

Except for very restrictive cases there is no sufficient knowledge about behaviour of the bias-robustness, mean square error (MSE)-robustness, efficiency robustness of $\Omega_{\alpha\delta}$-WLS estimators on the changes of assumptions defining $\mathcal{M}_\Omega$ for each known (α, δ). This calls for extensive studies on numerical simulation robustness. In this paper, however, we will not analyze the problems of studying robustness. Instead of this we will study some properties of $\Omega_{\gamma\delta}$-WLS estimator with respect to Ω-WLS--estimator. One of such properties concerns the conditions of e-quality of dispersion matrices of $\Omega_{\gamma\delta}$-WLSE and Ω-WLSE, i.e. for dispersion of $\hat{B}_{\gamma\delta} = \hat{B}(\Omega_{\gamma\delta}) = (X'\Omega_{\gamma\delta}^{-1} X)^{-1} X'\Omega_{\gamma\delta}^{-1} Y$, $\hat{B} = \hat{B}(\Omega) =$
$= (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}Y$. For the brevity of notation we denote $\alpha =$
$= (\gamma, \delta)$, $\Omega_{\gamma\delta} = \Omega_\alpha$.

By (7) and the properties of the operator $\mathcal{S}$ we have

$$\mathcal{S}(\hat{B}_\alpha) = (X'\Omega_\alpha^{-1}X)^{-1}X'\Omega_\alpha^{-1}\Omega\Omega_\alpha^{-1}X(X'\Omega_\alpha^{-1}X)^{-1}$$

and

$$\mathcal{S}(\hat{B}) = (X'\Omega^{-1}X)^{-1}$$

We shall fix the conditions of equality $\mathcal{S}(\hat{B}_\alpha) = \mathcal{S}(\hat{B})$. Such conditions are given in

    T h e o r e m   2. If.

    a) $\Omega\Omega_\alpha = \Omega_\alpha\Omega$;

    b) V is the matrix of the eigen vectors of Ω_α and Ω, i.e.
$V'V = VV' = I$;

    c) $V_0 \in R^{n\times k}$ is of the form $V_0 = (v_{.1} + v_{.n}, \ldots, v_{.k} +$
$+ v_{.,n-k+1})$, where $v_{.i}$ $(i = \overline{1, n})$ is the i-th column of V and
$V_0'V_0 = 2I_{(k)}$, $V_0V_0' = I_{(n)}$;

    d) the matrix X satisfies, alternatively, one of the fol-
lowing conditions:

    d1) $X = V_0$,

    d2) $X = V_0 G$, det $G \neq 0$, $G \in R^{k\times k}$,

    d3) $X = V_0 G$, $G'G = GG' = I_{(k)}$, $G \in R^{k\times k}$;

then

(11) $$\mathcal{B}(\hat{B}_\alpha) = \mathcal{B}(\hat{B}). \quad \blacklozenge$$

To prove this theorem one needs to use conditions (a)-(d), the properties of symmetry, positive definiteness of Ω_α, Ω, theorem of S t o l l and W o n g [6, p. 227] and the resulting facts that $\Omega V = V\Lambda$, $\Omega^{-1}V = V\Lambda^{-1}$, $V'\Omega^{-1} = \Lambda^{-1}V'$, $\Omega_\alpha V = VF$, $\Omega_\alpha^{-1}V = VF^{-1}$, $V'\Omega_\alpha^{-1} = F^{-1}V$, where Λ, F are the diagonal matrices of eigenvalues of Ω and Ω_α, and Λ^{-1}, F^{-1} are the eigenvalues matrices for Ω^{-1} and Ω_α^{-1}. A detailed proof is given in the work [5].

Another trivial situation when $\mathcal{B}(\hat{B}_\alpha) = \mathcal{B}(\hat{B})$ holds, is determined by the condition $\Omega_\alpha = \Omega$, i.e. then $(X'\Omega^{-1}X)^{-1}X'\Omega^{-1}\Omega\Omega^{-1}X(X'\Omega^{-1}X)^{-1} = (X'\Omega^{-1}X)^{-1}$.

3. A Determinantal Efficiency Measure

We recall that the most efficient estimator in the case of $\mathcal{N}\mathcal{M}_\Omega$ is the estimator $\hat{B} = \hat{B}(\Omega)$ and any other estimator including $\hat{B}_\alpha$ is less efficient. In order to arrive at this conclusion we represent the matrix Ω_α as $\Omega_\alpha = \Omega + A$. It is easy to find that

$$(\Omega + A)^{-1} = \Omega^{-1} - \Omega^{-1}(\Omega^{-1} + A^{-1})^{-1}\Omega^{-1} = \Omega^{-1} - G,$$

where $G = \Omega^{-1}(\Omega^{-1} + A^{-1})^{-1}\Omega^{-1}$.

By definition of $\hat{B}$ we now have

$$\hat{B}_\alpha = (X'(\Omega^{-1} - G)X)^{-1}X'(\Omega^{-1} - G)Y,$$

and since

$$(X'(\Omega^{-1} - G)X)^{-1} = (X'\Omega^{-1}X - X'GX)^{-1} = (X'\Omega^{-1}X)^{-1} -$$

$$- (X'\Omega^{-1}X)^{-1}[(X'\Omega^{-1}X)^{-1} - (X'GX)^{-1}]^{-1} (X'\Omega^{-1}X)^{-1},$$

$$X'(\Omega^{-1} - G)Y = X'\Omega^{-1}Y - X'GY.$$

therefore

(12) $\hat{B}_\alpha = (L + C)Y$,

where:

(13) $L = (X' \Omega^{-1}X)^{-1} X' \Omega^{-1}$,

(13a) $C = -(X'\Omega^{-1}X)^{-1} X'G - (X'\Omega^{-1}X)^{-1} [(X' \Omega^{-1}X)^{-1} -$

             $- (X'GX)^{-1}]^{-1} (X' \Omega^{-1}X)^{-1} (X'\Omega^{-1} - X'G)$.

By (12) and Theorem 1 we obtain that $\mathfrak{B}(\hat{B}_\alpha) \geqslant \mathfrak{B}(\hat{B})$, i.e. $\hat{B}$ is less efficient than $\hat{B}$. For fixed finite sample sizes (especially small sample sizes) it is very important to have a measure of efficiency of the given estimator $\hat{B}_\alpha$. Such a measure is defined in

D e f i n i t i o n   1. The determinantal efficiency measure of the estimator $\hat{B}_\alpha$ for β in $\mathcal{NM}_\Omega$ is the quotient of the determinant of the matrix $\mathfrak{B}(\hat{B})$ to the determinant of the matrix $\mathfrak{B}(\hat{B}_\alpha)$, i.e.

(14) $\Psi_{\hat{B}_\alpha} = \text{eff}_{\hat{B}_\alpha}(X, \Omega, \Omega_\alpha) = \dfrac{\det(\mathfrak{B}(\hat{B}))}{\det(\mathfrak{B}(\hat{B}_\alpha))} =$

                  $= \dfrac{\det^2(X' \Omega_\alpha^{-1}X)}{\det (X' \Omega^{-1}X) \det (X' \Omega_\alpha^{-1}\Omega\, \Omega_\alpha^{-1}X)}$,

where $\Psi_{\hat{B}_\alpha} = \text{eff}_{\hat{B}}(X, \Omega, \Omega_\alpha)$ denotes the determinantal efficiency measure of $\hat{B}_\alpha$ as the function of X, Ω, Ω_α. ◆

Let us assume that:

A2) Ω_α is nonsingular positive definite real matrix (rank $\Omega_\alpha = n_o = n$),

A3) $\Omega \Omega_\alpha = \Omega_\alpha \Omega$.

By (A1), (A2), (A3) the facts that the inverse of positive definite matrix is also a positive definite matrix and the product of positive definite matrices is also a positive definite matrix - (for proofs of this statements see [1], chapters 4, 6) we obtain that

(15) Ω^{-1}, Ω_α^{-1} are positive definite (p.d.) real matrices,

(16) $$\Omega_\alpha^{-1}\,\Omega\,\Omega_\alpha^{-1} \text{ is p.d. matrix.}$$

and by the properties of Gram matrices

(17) $$X'\,\Omega^{-1}X,\ X'\,\Omega_\alpha^{-1}X,\ X'\,\Omega_\alpha^{-1}\Omega\Omega_\alpha^{-1}X \text{ are p.d. matrices.}$$

By the properties of Gram determinants we get

(18) $$\det(X'\Omega^{-1}X)>0,\ \ \det(X'\,\Omega_\alpha^{-1}X)>0,\ \ |X'\Omega_\alpha^{-1}\Omega\,\Omega_\alpha^{-1}X|>0.$$

From (18) and (14) we have therefore

(19) $$\Psi_{\hat{B}_\alpha} > 0$$

and

(20) $\Psi_{\hat{B}_\alpha} = 0$ if $\det(X'\Omega^{-1}X) \neq 0$, $\det(X'\Omega_\alpha^{-1}\Omega\Omega_\alpha^{-1}X) \neq 0$,

$$\det(X'\,\Omega_\alpha^{-1}X) = 0.$$

For $k_o < k$, $\Psi_{\hat{B}_\alpha} = \frac{0}{0}$, so the measure makes no sense and there arises a need for modifications — one of such modifications can be based on taking

$$\Psi_{\hat{B}_\alpha}^{*} = \frac{\text{tr}(\text{\textcyrillic{Ƃ}}(\hat{B}))}{\text{tr}(\text{\textcyrillic{Ƃ}}(\hat{B}_\alpha))}.$$

Now we establish that $\Psi_{\hat{B}_\alpha} \leqslant 1$.

Because for positive definite matrix the determinant is equal to the product of its positive eigenvalues, therefore by the fact that $\text{\textcyrillic{Ƃ}}(\hat{B}_\alpha) \geqslant \text{\textcyrillic{Ƃ}}(\hat{B})$ we come to the conclusion

(21) $$\det(\text{\textcyrillic{Ƃ}}(\hat{B}_\alpha)) \ \geqslant \ \det(\text{\textcyrillic{Ƃ}}(\hat{B})).$$

Hence $\text{eff}_{\hat{B}_\alpha} = \Psi_{\hat{B}_\alpha} \leqslant 1$.

We have proved the following theorem.

T h e o r e m   3.   Under the assumptions of $\mathcal{Mol}_\Omega$, (A1), (A2) the range of the determinantal efficiency measure $\Psi_{\hat{B}_\alpha}$ is the interval $\langle 0,1 \rangle$.

Note: It is easy to check that for $\Omega_\alpha = \Omega$, $\Psi_{\hat{B}_\alpha} = 1$.

In practice it is worth knowing the values of lower bounds of $\Psi_{\hat{B}_\alpha}$ determined for some typical $\alpha = (\gamma, \delta)$ combinations of auto-

correlation schemes and estimators of ρ. We derive these bounds in the next section.

4. Lower Bounds of $\Psi_{\hat{B}_\alpha}$

One of the possible ways of deriving lower bounds of $\Psi_{\hat{B}_\alpha}$ in the case of $X = V_o$, i.e. $X'X = 2I_{(k)}$ can be based on the method of Lagrange multipliers. Let

(22) $$g = \ln (\Psi_{\hat{B}_\alpha})^{-1} - 2 \, \mathrm{tr} \, L(V_o'V_o - 2I_{(k)}),$$

where:

$$\ln (\Psi_{\hat{B}_\alpha})^{-1} = \ln \det(V_o' \Omega^{-1} V_o) + \ln \det (V_o' \Omega_\alpha^{-1} \Omega \Omega_\alpha^{-1} V_o) -$$

$$- 2 \ln \det (V_o' \Omega_\alpha^{-1} V_o)$$

and L is the upper triangular matrix of $k(k+1)/2$ of Lagrange's multipliers. From (A1)-(A3) it follows

(22a) $$\Omega_\alpha^{-1} \Omega^{-1} = \Omega^{-1} \Omega_\alpha^{-1}, \quad \Omega_\alpha^{-1}\Omega = \Omega \Omega_\alpha^{-1} = H^{-1},$$

$$\Omega^{-1}\Omega_\alpha = \Omega_\alpha\Omega^{-1} = H, \quad H = H',$$

(22b) $$(\Omega_\alpha^{-1} \Omega \Omega_\alpha^{-1}) \Omega^{-1} = \Omega_\alpha^{-1} (\Omega_\alpha^{-1} \Omega \Omega_\alpha^{-1}),$$

(22c) $$(\Omega_\alpha^{-1} \Omega \Omega_\alpha^{-1}) \Omega_\alpha^{-1} = \Omega_\alpha^{-1} (\Omega_\alpha^{-1} \Omega \Omega_\alpha^{-1}),$$

(22d) $$(\Omega_\alpha^{-1} \Omega \Omega_\alpha^{-1})H = H (\Omega_\alpha^{-1} \Omega \Omega_\alpha^{-1}),$$

$$(\Omega_\alpha^{-1} \Omega \Omega_\alpha^{-1})H^2 = H^2 (\Omega_\alpha^{-1} \Omega \Omega_\alpha^{-1}),$$

(22e) $$(\Omega_\alpha^{-1} \Omega \Omega_\alpha^{-1})H^{-1} = H^{-1} (\Omega_\alpha^{-1} \Omega \Omega_\alpha^{-1}).$$

We can formulate now

Theorem 4. If $V_o'V_o = 2I_{(k)}$ and (22a)-(22e) hold and the function g is continuous twofold differentiable, then

(23) $$\inf (ef_{\hat{B}_\alpha}(V_o, \Omega, \Omega_\alpha)) \geqslant \prod_{i=1}^{k} 2h_i^* h_{n-i+1}^* (h_i^* + h_{n-i+1}^*)^{-2},$$

(24) $$V_0 : V_0' V_0 = 2I_{(k)}$$

where $2H^* = V_0' H V_0$ is the matrix of the form $H^* = \text{diag }(h_1^*, \ldots, h_k^*)$. $\blacklozenge$

P r o o f. Differentiating g with respect to X and L (cf. [4], p. 616) and putting $\partial g/\partial X = 0$ and $g/L = 0$, where 0 denotes zero matrix, we have

(24a) $$\Omega^{-1} V_0 (V_0' \Omega^{-1} V_0)^{-1} + \Omega_\alpha^{-1} \Omega \Omega_\alpha^{-1} V_0 (V_0' \Omega_\alpha^{-1} \Omega \Omega_\alpha^{-1} V_0)^{-1} -$$
$$- 2\Omega_\alpha^{-1} V_0 (V_0' \Omega_\alpha^{-1} V_0)^{-1} - 2V_0(L + L') = 0,$$

(24b) $$V_0' V_0 - 2I_{(k)} = 0.$$

Premultiplying (24a) by V_0' and using (24b) we have

(24c) $$L + L' = 0$$

and hence

(24d) $$\Omega^{-1} V_0 (V_0' \Omega^{-1} V_0)^{-1} + \Omega_\alpha^{-1} \Omega \Omega_\alpha^{-1} V_0 (V_0' \Omega_\alpha^{-1} \Omega \Omega_\alpha^{-1} V_0)^{-1} =$$
$$= 2\Omega_\alpha^{-1} V_0 (V_0' \Omega_\alpha^{-1} V_0)^{-1}.$$

Premultiplying (24d) by Ω_α and using the definition of H we obtain

(24e) $$H V_0 (V_0' \Omega^{-1} V_0)^{-1} + H^{-1} V_0 (V_0' \Omega_\alpha^{-1} \Omega \Omega_\alpha^{-1} V_0)^{-1} = 2 V_0 (V_0' \Omega_\alpha^{-1} V_0)^{-1}.$$

Premultiplying (24e) by H and using the rules of the addition of the matrices we obtain

(24f) $$H^2 V_0 (V_0' \Omega^{-1} V_0)^{-1} - 2 H V_0 (V_0' \Omega_\alpha^{-1} V_0)^{-1} + V_0 (V_0' \Omega_\alpha^{-1} \Omega \Omega_\alpha^{-1} V_0)^{-1} = 0.$$

From Theorem 2 and corollary 1 (from [7], § 6.6, p. 228) it follows that matrices H^2, H, $\Omega_\alpha^{-1}, \Omega^{-1}, \Omega_\alpha^{-1} \Omega \Omega_\alpha^{-1}$ are simultaneously diagonalized by the matrix V_0 of the semi-orthogonal transformation V_0'. Hence,

(24g) $$H^2 V_{.i}^0 (V_{.i}^{0'} \Omega^{-1} V_{.i}^0)^{-1} - 2 H V_{.i}^0 (V_{.i}^{0'} \Omega_\alpha^{-1} V_{.i}^0)^{-1} +$$

$$+ v_{.i}^{o} (v_{.i}^{o'} \Omega_{\alpha}^{-1} \Omega \Omega_{\alpha}^{-1} v_{.i}^{o})^{-1} = 0, \quad i = 1, \ldots, k$$

and by virtue of the theorem, that the determinant of the diagonal matrix is equal to the product of its main diagonal elements we get

$$(25) \quad ef_{\hat{\theta}_{\alpha}}(V_{o}, \Omega, \Omega_{\alpha}) = \prod_{i=1}^{k} (v_{.i}^{o'} \Omega_{\alpha}^{-1} v_{.i}^{o})^{2} (v_{.i}^{o'} \Omega^{-1} v_{.i}^{o})^{-1} (v_{.i}^{o'} \Omega_{\alpha}^{-1} \Omega \Omega_{\alpha}^{-1} v_{.i}^{o})^{-1}.$$

Premultiplying (24g) by $v_{.i}^{o'} (v_{.i}^{o'} \Omega^{-1} v_{.i}^{o})$ we obtain .

$$(25a) \qquad 2h_{i}^{*2} - 2h_{i}^{*}(v_{.i}^{o'} \Omega^{-1} v_{.i}^{o})(v_{.i}^{o'} \Omega_{\alpha}^{-1} v_{.i}^{o})^{-1} +$$

$$+ (v_{.i}^{o'} \Omega^{-1} v_{.i}^{o})(v_{.i}^{o'} \Omega_{\alpha}^{-1} \Omega \Omega_{\alpha}^{-1} v_{.i}^{o})^{-1} = 0 \qquad i=1,\ldots,k.$$

When Δ of this squared-equation in respect to h^{*} is greater than 0, i.e.

$$4(v_{.i}^{o'} \Omega^{-1} v_{.i}^{o})^{2}(v_{.i}^{o'} \Omega_{\alpha}^{-1} v_{.i}^{o})^{-2} - 8(v_{.i}^{o'} \Omega^{-1} v_{.i}^{o})(v_{.i}^{o'} \Omega_{\alpha}^{-1} \Omega \Omega_{\alpha}^{-1} v_{.i}^{o})^{-1} > 0$$

equation (25a) has two roots, different from zero, the sum of them being of the form

$$(25b) \qquad h_{i1}^{*} + h_{i2}^{*} = (v_{.i}^{o'} \Omega^{-1} v_{.i}^{o})(v_{.i}^{o'} \Omega_{\alpha}^{-1} v_{.i}^{o})^{-1},$$

and their product has the form

$$(25c) \qquad h_{i1}^{*} h_{i2}^{*} = \frac{1}{2}(v_{.i}^{o'} \Omega^{-1} v_{.i}^{o})(v_{.i}^{o'} \Omega_{\alpha}^{-1} \Omega \Omega_{\alpha}^{-1} v_{.i}^{o})^{-1}.$$

From (25b) and (25c) it follows directly

$$(25d) \qquad (v_{.i}^{o'} \Omega_{\alpha}^{-1} v_{.i}^{o})^{2} = (v_{.i}^{o'} \Omega^{-1} v_{.i}^{o})^{2}(h_{i1}^{*} + h_{i2}^{*})^{-2}$$

$$(25\ e) \qquad (v_{.i}^{o'} \Omega_{\alpha}^{-1} \Omega \Omega_{\alpha}^{-1} v_{.i}^{o})^{-1} = 2(v_{.i}^{o'} \Omega^{-1} v_{.i}^{o})^{-1} h_{i1}^{*} h_{i2}^{*}$$

Using (25d) and (25e) we can rewrite (25) in the form

$$(26) \qquad ef_{\hat{\theta}_{\alpha}}(V_{o}, \Omega, \Omega_{\alpha}) = \prod_{i=1}^{k} 2h_{i1}^{*} h_{i2}^{*}(h_{i1}^{*} + h_{i2}^{*})^{-2}.$$

In order to find $\inf ef_{\hat{B}_\alpha}(V_o, \Omega, \Omega_\alpha)$ we must choose two disjoint subgroups $\{h_{11}^*, \ldots, h_{k1}^*\}$, $\{h_{12}^*, \ldots, h_{k2}^*\}$ from the set $\{h_1^*, \ldots, h_k^*\}$ in such a way that the relation defining $ef_{\hat{B}_\alpha}(V_o, \Omega, \Omega_\alpha)$ in (26) reaches its minimum. In accordance with the combinatorical discussion (similar to that of B l o o m f i e l d-W a t-s o n [2]) we have that the expression $ef_{\hat{B}_\alpha}$ reaches its minimum for $h_{11}^* = h_1^*$, $h_{12}^* = h_{n-i+1}^*$, i.e.

$$\inf ef_{\hat{B}_\alpha}(V_o, \Omega, \Omega_\alpha) \geq \prod_{i=1}^{k} 2h_1^* h_{n-i+1}^* (h_1^* + h_{n-i+1}^*)^{-2},$$

which completes the proof.

Relation (23) may be used for examining the runs of range of the lower bound of the determinantal efficiency measure of the estimator $\hat{B}_\alpha$.

It is interesting to determine other lower bounds rather than these from Theorem 4 for $X = V_o$. To fix them on the basis of stationarity conditions for $\Psi_{\hat{B}_\alpha}$ one has to find solutions of the following equations:

(27) $\dfrac{\partial \Psi_{\hat{B}_\alpha}}{\partial X} = (2 \det(X'\Omega^{-1}X) \det^2(X'\Omega_\alpha^{-1}X))\Omega_\alpha^{-1}X(X'\Omega_\alpha^{-1}X)^{-1} - (\det(X'\Omega^{-1}X)),$

$\Omega^{-1}X(X'\Omega^{-1}X)^{-1} - (\det(X'\Omega_\alpha^{-1}X))\Omega_\alpha^{-1} \Omega \Omega_\alpha^{-1}X(X'\Omega_\alpha^{-1}\Omega \Omega_\alpha^{-1}X)^{-1} = 0,$

(28) $\dfrac{\partial \Psi_{\hat{B}_\alpha}}{\partial \Omega_\alpha} = (\det^2(X'\Omega_\alpha^{-1}X)) \Omega_\alpha^{-1}X(X'\Omega_\alpha^{-1}X)^{-1}X'\Omega_\alpha^{-1} - \Omega\Omega_\alpha^{-1}X,$

$(X'\Omega_\alpha^{-1}\Omega \Omega_\alpha^{-1}X)^{-1}X' = 0,$

(29) $\dfrac{\partial \Psi_{\hat{B}_\alpha}}{\partial \Omega} = \Omega^{-1}X(X'\Omega^{-1}X)^{-1}X'\Omega^{-1} + \Omega_\alpha^{-1}X(X'\Omega_\alpha^{-1}\Omega\Omega_\alpha^{-1}X)^{-1}X'\Omega_\alpha^{-1} = 0.$

Whether there are general solutions of (27)-(29)? If they e-xist, they involve additional assumptions about the extensions of det functions with respect to Ω or Ω_α or X or the restrictive form of X. Due to space limitations we will not continue the discussion of this problem and leave the problem open.

5. Final Remarks

The results obtained in this paper will be used in our stu-
dies on robustness of WLS estimators as well as for the constru-
ction of efficiency tables for different combinations of pairs
(γ, δ).

The analysis presented in the paper does not include the ca-
se when $k_o < k$ and $n_o < n$ in the model $\mathcal{N} \mathcal{M}_\Omega$.

BIBLIOGRAPHY

[1] B e l l m a n   R. (1960), Introduction to Matrix Analysis,
    N.Y., McGraw-Hill.
[2] B l o o m f i e l d P., W a t s o n   G. (1975), The In-
    efficiency of Least Squares, "Biometrika", Vol. 62, p. 121-
    -128.
[3] C r a m e r   H. (1946), Mathematical Methods of Statistics,
    Princeton N.J., Princeton University Press.
[4] D w y e r   P. (1967), Some Applications of Matrix Deriva-
    tives in Multivariate Analysis, J. Amer. Statist.Assoc., Vol.
    62, p. 607-625.
[5] M i l o W., W a s i l e w s k i   Z. (1979), Efektywność
    estymatorów parametrów ogólnych modeli   liniowych. Cz. 1.
    Work under contract R.III.9.5.7.
[6] S t o l l R., W o n g   E. (1969), Linear Algebra, N. Y.,
    Academic Press.
[7] Z a c k s   S. (1971), The Theory of Statistical Inference,
    N. Y., Wiley.

Władysław Milo, Zbigniew Wasilewski

O EFEKTYWNOŚCI WAŻONYCH ESTYMATORÓW
NAJMNIEJSZYCH KWADRATÓW
W PRZYPADKU OGÓLNEGO MODELU LINIOWEGO

Głównym celem pracy jest zaprezentowanie jednego z możliwych sposobów mierzenia efektywności w małych próbach i zanalizowanie niektórych własności ważonych estymatorów najmniejszych kwadratów i przedstawionej wyznacznikowej miary efektywności. W szczególności przedstawiono:
    a) analizę własności estymatorów ważonych w przypadku ogólnego modelu liniowego,
    b) dowód, że miara efektywności znajduje się w przedziale <0,1>,
    c) wyznaczanie dolnego kresu wyznacznikowej miary efektywności.