ACTA UNIVERSITATIS LODZIENSIS

FOLIA OECONOMICA 68, 1987

Władysław Milo*

ON BIASED REGULARIZING ESTIMATORS. PART II

1. Introduction

The subject of consideration is the model (see [3], [5] for details) $NM_2 \equiv (R^{n\times k}, \mathcal{S}, Y = x\beta + \Xi, k_0 = k, n_0 = n, P_Y = \mathcal{N}_Y(x\beta, \Omega))$, where the random vector Y has probability distribution P_Y being n dimensional normal distribution N_Y with mean $\mathcal{E}Y = x\beta$, dispersion $\mathcal{D}Y = \Omega$, $\Omega \in R^{n\times n}$, $x \in R^{n\times k}$, $\beta \in R^k$. The purpose of this paper is a comparative analysis of statistical properties of the following three estimators derived from the functionals

$$\varphi_1 = \|\Omega^{-1}(Y - x\beta)\|^2$$

$$\varphi_2 = \varphi_1 + \|\beta\|^2, \quad \varphi_3 = \varphi_1 + \beta'\Gamma\beta$$

$$B_a = (x'\Omega^{-1}x)^{-1}x^{\tau}\Omega^{-1}Y,$$

$$B_b = (x'\Omega^{-1}x + \gamma I)^{-1}x^{\tau}\Omega^{-1}Y,$$

$$B_c = (x'\Omega^{-1}x + \Gamma)^{-1}x^{\tau}\Omega^{-1}Y,$$

and some functions of B_a, B_b, B_c. It would be proved, among others, that (see § 2):

.) the estimators B_a, B_b, the predictors $\hat{Y}_a$, $\hat{Y}_b$, the residuals E_a, E_b are consistent and are having normal distribution;

..) the random quadratic forms $B_a'B_a$, $B_b'B_b$, $\hat{Y}_a'\hat{Y}_a$, $\hat{Y}_b'\hat{Y}_b$, $E_a'E_a$, $E_b'E_b$ do not have χ^2 - distribution,

...) $\operatorname{cov}(B_a, S^2_{E_a}) = 0_{k\times 1}$, $\operatorname{cov}(B_b, S^2_{E_b}) \neq 0_{k\times 1}$, where $S^2_{E_a} = E_a'E_a/(n-k)$, $S^2_{E_b} = E_b'E_b/(n-k)$.

*Lecturer, Institute of Econometrics and Statistics, University of Łódź.

In § 3 there would be proved for B_c some analogous of theorems from § 2.

2. Properties of estimator B_b

From the works [6], [7] it follows that if

$$P_Y = \mathcal{N}_Y(m, \Omega),\ C, \Omega, \mathcal{A} \in R^{n\times n},\ L, m, d \in R^{n\times 1},\ \text{then}$$

I) $\mathcal{E}(Y'\mathcal{A}Y)\ tr(\mathcal{A}\Omega) + m^{\tau}\mathcal{A}m$,

II) $var(Y'\mathcal{A}Y) = 2\,tr(\mathcal{A}\Omega)^2 + 4\,m^{\tau}\mathcal{A}\Omega\mathcal{A}m$,

III) $cov(LY, Y'\mathcal{A}Y) = 2L\Omega\mathcal{A}m$,

IV) $MSE(Y) = tr\,\mathcal{D}(Y) + tr\,ob(Y)\,ob^{\tau}(Y),\ ob(Y) = \mathcal{E}Y - m$,

V) if $Y = d + CU$, $P_U = \mathcal{N}_U(m, \Omega)$, then

$P_Y = \mathcal{N}_Y(Cm + d,\ C\Omega C')$,

VI) if $Q = Y'\mathcal{A}Y + 2a'Y + \bar{\alpha}$, $\mathcal{A} \in R^{n\times n}$, $a \in R^{n\times 1}$, $\bar{\alpha} \in R$, then

$P_Q = \chi^2(s, \lambda)$ iff $\Omega\mathcal{A}\Omega\mathcal{A}\Omega = \Omega\mathcal{A}\Omega$,

$\Omega(a + \mathcal{A}m) = \Omega\mathcal{A}\Omega(a + \mathcal{A}m)$, $\lambda = \bar{\alpha} + 2a^{\tau}m + m^{\tau}\mathcal{A}m$,

$s = tr\,\mathcal{A}\Omega = rz(\Omega\mathcal{A}\Omega)$.

The relations (I) and (IV) also hold for $P_Y \neq \mathcal{N}_Y(\ldots)$ with $\mathcal{E}(Y) = m$, $\mathcal{D}(Y) = \Omega$.

Using the definition of model NM_2, the relations (I)-(VI) and the definitions of symbols $\mathcal{E}$, $\mathcal{D}$, P, MSE, var, cov (expectation, dispersion, probability distribution, mean square error, variance, covariance) it is easy to see that:

(1a) $B_a = (x'\Omega^{-1}x)^{-1}x'\Omega^{-1}Y = K_aY$, $K_a = (x'\Omega^{-1}x)^{-1}x'\Omega^{-1}$,

(2a) $\mathcal{E}(B_a) = K_a x\beta = \beta$, $ob(B_a) = \mathcal{E}(B_a) - \beta = 0$ (unbiasedness of B_a),

(3a) $\mathcal{D}(B_a) = K_a\Omega K_a = (x\Omega^{-1}x)^{-1}$,

(4a) $P_{B_a} = \mathcal{N}_{B_a}(\beta, x'\Omega^{-1}x)^{-1}$, (normality of B_a)

(5a) $MSE\,(B_a) = tr\,(x'\Omega^{-1}x)^{-1}$,

(6a) $B_a'B_a = Y'K_a'K_aY$,

(7a) $\mathcal{E}(B_a'B_a) = tr\,(K_a'K_a\Omega) + \beta'\beta = tr\,(x'\Omega^{-1}x)^{-1} + \beta'\beta$,

(8a) $var\,(B_a'B_a) = 2tr\,(x'\Omega^{-1}x)^{-2} + 4\beta'(x'\Omega^{-1}x)^{-1}\beta$,

(9a) $P_{B_a'B_a} \neq \chi^2(.,.)$ due to $\Omega K_a'K_a\Omega K_a'K_a \neq \Omega K_a'K_a\Omega$

(10a) $\hat{Y}_a = xB_a = xK_aY$,

(11a) $\mathcal{E}(\hat{Y}_a) = xK_ax\beta = x\beta$, $ob\,(\hat{Y}_a) = x\beta - x\beta = 0$,

(12a) $\mathcal{D}(\hat{Y}_a) = x\,K_a\Omega\,K_a'x = x(x'\Omega^{-1}x)^{-1}x'$,

(13a) $P_{\hat{Y}_A} = \mathcal{N}_{\hat{Y}_a}\,(x\beta, xK_a\Omega K_a'\,x')$,

(14a) $MSE\,(\hat{Y}_a) = tr\,(xK_a\Omega K_a'\,x')$,

(15a) $\hat{Y}_a'\hat{Y}_a = Y'K_a'x'xK_aY$,

(16a) $\mathcal{E}(\hat{Y}_a'\hat{Y}_a) = tr\,(K_a'x'xK_a\Omega) + \beta'x'K_a'x'xK_ax\beta$,

(17a) $var\,(\hat{Y}_a'\hat{Y}_a) = 2tr\,(K_a'x'x\,K_a\Omega)^2 + 4\beta'x'\,K_a'x'xK_a\Omega K_a'\,x'xK_ax\beta$,

(18a) $P_{\hat{Y}_a'\hat{Y}_a/\sigma^2} \neq \chi^2(.,.)$ due to $\Omega K_a'x'x\,K_a\Omega\,K_a'x'xK_a\Omega \neq$
$\neq \Omega K_a'\,x'x\,K_a\Omega$,

(19a) $E_a = (I - xK_a)Y = M_aY$, $M_a = I - xK_a$,

(20a) $\mathcal{E}(E_a) = x\beta - xK_a\,x\beta = 0$, $ob\,(E_a) = 0 - 0 = 0$,

(21a) $\mathcal{D}(E_a) = M_a\Omega M_a'$,

(22a) $P_{E_a} = \mathcal{N}_{E_a}\,(0, M_a\Omega M_a')$,

(23a) $MSE\,(E_a) = tr\,(M_a\Omega M_a')$, (24a) $E'_a E_a = Y'M'_a M_a Y$,

(25a) $\mathcal{E}(E'_a E_a) = tr\,(M'_a M_a \Omega) + \beta' x' M'_a M_a x\beta = tr\,(M'_a M_a \Omega) + 0$,

(26a) $var\,(E'_a E_a) = 2tr\,(M'_a M_a \Omega)^2 + 4\beta' x' M'_a M_a \Omega M'_a M_a x\beta$,

(27a) $P_{E'_a E_a}/\sigma^2 \neq \chi^2(\cdot,\cdot)$, since $\Omega M'_a M_a \Omega M'_a M_a \Omega \neq \Omega M'_a M_a \Omega$,

(28a) $cov\,(B_a, S^2_{E_a}) = cov\,(K_a Y, \frac{1}{n_a} E'_a E_a) = \frac{2}{n_a} K_a \Omega M'_a M_a x\beta = 0$,

$n_a = tr\, M'_a M_a \Phi$, $\Omega = \sigma^2 \Phi$, $tr\,(M'_a M_a = tr[\Phi - x'x(x'\Omega^{-1}x)^{-1}]$.

Using (I)-(VI), the assumptions of model NM and the definitions of $\mathcal{E}$, $\mathcal{D}$, var, MSE, P, cov it is easy to find out that

(1b) $B_b = K_b Y$, $K_b = (x'\Omega^{-1}x + \gamma I)^{-1} x'\Omega^{-1}$,

(2b) $\mathcal{E}(B_b) = K_b x\beta$, $ob\,(B_b) = K_b x\beta - \beta = (K_b x - I)\beta$,

(3b) $\mathcal{D}(B_b) = K_b \Omega K'_b$,

(4b) $P_{B_b} = \mathcal{N}_{B_b}(K_b x\beta, K_b \Omega K'_b)$,

(5b) $MSE\,(B_b) = tr\,(K_b \Omega K'_b) + \beta'(K_b x - I)'(K_b x - I)\beta$,

(6b) $B'_b B_b = Y' K'_b K_b Y$,

(7b) $\mathcal{E}(B'_b B_b) = tr\,(K'_b K_b \Omega) + \beta' x' K'_b K_b x\beta$,

(8b) $var\,(B'_b B_b) = 2tr\,(K'_b K_b \Omega)^2 + 4\beta' x' K'_b K_b \Omega K'_b K_b x\beta$,

(9b) $P_{B'_b B_b}/\sigma^2 \neq \chi^2(\cdot,\cdot)$ due to $\Omega K'_b K_b \Omega K'_b K_b \Omega \neq \Omega K'_b K_b \Omega$,

(10b) $\hat{Y}_b = xB_b = xK_b Y$,

(11b) $\mathcal{E}(\hat{Y}_b) = xK_b x\beta$, $ob\,(\hat{Y}_b) = (xK_b - I)\,x\beta$,

(12b) $\mathcal{D}(\hat{Y}_b) = xK_b \Omega K'_b x'$;

(13b) $P_{\hat{Y}_b} = \mathcal{N}_{\hat{Y}_b}(xK_b x\beta,\ xK_b\Omega K_b'x')$

(14b) $\mathrm{MSE}(\hat{Y}_b) = \mathrm{tr}(xK_b\Omega K_b'x') + \beta'x'(xK_b - I)'(xK_b - I)x\beta,$

(15b) $\hat{Y}_b'\hat{Y}_b = Y'K_b'x'xK_bY,$

(16b) $\mathcal{E}(\hat{Y}_b'\hat{Y}_b) = \mathrm{tr}(K_b'x'xK_b\Omega) + \beta'x'K_b'x'xK_bx\beta,$

(17b) $\mathrm{var}(\hat{Y}_b'\hat{Y}_b) = 2\mathrm{tr}(K_b'x'xK_b)^2 + 4\beta'x'K_b'x'xK_b\Omega K_b'x'xK_bx\beta,$

(18b) $P_{\hat{Y}_b'\hat{Y}_b/\sigma^2} \neq \chi^2(.,.)$ since $\Omega K_b'x'xK_b\Omega K_b'x'xK_b\Omega \neq \Omega K_b'x'xK_b\Omega$

(19b) $E_b = (I - xK_b)Y = M_bY,\quad M_b = I - xK_b,$

(20b) $\mathcal{E}(E_b) = M_bx\beta \neq 0,\quad \mathrm{ob}(E_b) = M_bx\beta \neq 0,$

(21b) $\mathcal{D}(E_b) = M_b\Omega M_b',$

(22b) $P_{E_b} = \mathcal{N}_{E_b}(M_bx\beta,\ M_b\Omega M_b'),$

(23b) $\mathrm{MSE}(E_b) = \mathrm{tr}(M_b\Omega M_b') + \beta'x'M_b'M_bx\beta,$

(24b) $E_b'E_b = Y'M_b'M_bY,$

(25b) $\mathcal{E}(E_b'E_b) = \mathrm{tr}(M_b'M_b\Omega) + \beta'x'M_b'M_bx\beta,$

(26b) $\mathrm{var}(E_b'E_b) = 2\mathrm{tr}(M_b'M_b\Omega)^2 + 4\beta'x'M_b'M_b\Omega M_b'M_bx\beta,$

(27b) $P_{E_b'E_b/\sigma^2} \neq \chi^2(.,.)$ since $\Omega M_b'M_b\Omega M_b'M_b\Omega \neq \Omega M_b'M_b\Omega,$

(28b) $\mathrm{cov}(B_b, S^2_{E_b}) = \mathrm{cov}(K_bY, \frac{1}{n_b}E_b'E_b) = \frac{2}{n_b}K_b\Omega M_b'M_bx\beta \neq 0,$

Under the assumptions

(29) $\mathrm{plim}\,\frac{1}{\sqrt{n}}(X_{(n)}'\Omega_{(n)}^{-1}X_{(n)})^{-1} = Q_a^{-1} \neq 0,$

(or $\mathrm{plim}\,\frac{1}{\sqrt{n}}X_{(n)}'\Omega_{(n)}^{-1}X = Q_a$ by nonsingularity of Q_a^{-1}),

$$\infty > |x_{i,l,(n)}|, \ |\omega_{i,l,(n)}| < \infty, \ \forall i,l,n \in N = \{1,\ldots,\}$$

(30) $\operatorname{plim} \frac{1}{\sqrt{n}} ((X'_{(n)}\Omega^{-1}_{(n)} X_n) + \gamma I)^{-1} = Q_b^{-1} \neq 0,$

(31) $\operatorname{plim} \frac{1}{\sqrt{n}} (X'_{(n)}\Omega^{-1}_{(n)} X_n + \Gamma)^{-1} = Q_c^{-1} \neq 0,$

(32) $\operatorname{plim} \frac{1}{\sqrt{n}} (X'_{(n)}\Omega^{-1}_{(n)} \Xi_{(n)}) = 0,$

by the same arguments as in A n d e r s o n, T a y l o r work [1], M i l o's works [4] the estimators B_a, B_b are consistent, that is, $\operatorname{plim} B_a = \operatorname{plim} B_b = \beta$

The relations (1a) - (28a) prove.

Theorem 1. Let the assumptions of model NM_1 and (29), (32) hold. Then

a) B_a is unbiased, consistent, efficient and normally distributed; the quadratic form $B_a' B_a$ does not have χ^2 - distribution;

b) $\hat{Y}_a$ is unbiased, consistent, normally distributed predictor; the quadratic form $\hat{Y}_a'\hat{Y}_a$ does not have χ^2 - distribution;

c) E_a is unbiased, consistent, normally distributed residual vector; the quadratic form $E_a'E_a$ does not have χ^2 - distribution;

d) $\operatorname{cov}(B_a, S^2_{E_a}) = 0.$ ◆

[Note: consistency of $\hat{Y}_a$ follows from consistency of B_a and $\hat{Y} = xB_a$; consistency of E_a follows from consistency of B_a and the fact that $E_a = M_a Y$; efficiency of B_a follows from the fact that for each K_a^*, $K_a^* = K_a + C$ it is $\mathcal{D}(K_a^*) > \mathcal{D}(K_a Y)$, where $C \in R^{k \times n}$].

The relations (1b)-(28b) prove.

Theorem 2. Let the assumptions of model NM_2 and (29), (30), (32) be fulfilled. Then

a) the estimator B_b is biased, consistent and it has multivariate normal distribution but $\|B_b\|^2$ does not have χ^2 - distribution;

b) the ex-post predictor $\hat{Y}_b$ is biased, consistent, and it has n-variate normal distribution but $\|\hat{Y}_b\|^2$ does not have χ^2 - distribution;

c) the ex-post residual vector E_b is biased, consistent, and it has multivariate singular normal distribution but $\|E_b\|^2$ does not have χ^2 - distribution;

d) the covariance of B_b and S_b^2 is different from the zero vector. ◆

Denoting

$$A_1 = (x'\Omega^{-1}x)^{-1}, \quad A_2 = (x'\Omega^{-1}x + \gamma I)^{-1},$$

it is easily seen that $A_1 = A_1'$, $A_2 = A_2'$, and A_1A_2 are positive definite matrices. Hence det $(A_1) > 0$, det $(A_2) > 0$, $A_1A_2 = A_2A_1$. Therefore, an orthogonal matrix T diagonalizes simultanously matrices A_1 and A_2, i.e.

(33) $$T'A_1T = \Lambda_1 = \mathrm{diag}\left(\frac{1}{\lambda_1}, \ldots, \frac{1}{\lambda_k}\right),$$

where λ_i, $i = 1, \ldots, k$ are given eigen values of the matrix $x'\Omega^{-1}x$, and

$$T'A_2T = \Lambda_2 = \mathrm{diag}\left(\frac{1}{\lambda_1 + \gamma}, \ldots, \frac{1}{\lambda_k + \gamma}\right).$$

Recall that $K_a = A_1x'\Omega^{-1}$, $K_b = A_2x'\Omega^{-1}$, $\mathcal{D}(B_a) = A_1$, $\mathcal{D}(B_b) = K_b\Omega K_b' = A_2A_1^{-1}A_2' = A_2A_1^{-1}A_2$, $K_bx = A_2A_1^{-1}$.

Hence

$$\mathcal{E}(B_b) \leq \mathcal{E}(B_a) \Longleftrightarrow K_bx\beta \leq \beta \Longleftrightarrow A_2A_1^{-1}\beta \leq \beta \overset{\beta>0}{\Longleftrightarrow}$$

$$\beta'(A_2A_1^{-1})\beta \leq \beta'\beta \Longleftrightarrow \det(A_2A_1^{-1}) \leq \det(I) \Longleftrightarrow$$

(34a) $$\det(A_2A_1^{-1}) \leq 1.$$

Similarly,

(34b) $$\mathcal{E}(B_b) > \mathcal{E}(B_a) \Longleftrightarrow \det(A_2A_1^{-1}) > 1 \quad \text{under } \beta > 0.$$

By definition of $\mathcal{D}(B_a)$, $\mathcal{D}(B_b)$ we have

(34c) $$\mathcal{D}(B_b) \geq \mathcal{D}(B_a) \Leftrightarrow A_2A_1^{-1}A_2 \geq A_1 \Leftrightarrow \det(A_2A_1^{-1})^2 \geq 1,$$

(34d) $\mathcal{D}(B_b) < \mathcal{D}(B_a) \iff A_2A_1^{-1}A_2 < A_1 \iff \det(A_2A_1^{-1})^2 < 1.$

By definition of MSE (B_a), MSE (B_b) we get MSE (B_b) = $= \mathrm{tr}(A_2A_1^{-1}A_2) + \beta'(A_2A_1^{-1} - I)'(A_2A_1^{-1} - I)\beta$, MSE $(B_a) = \mathrm{tr}(A_1)$.

Hence the condition MSE $(B_b) \lesseqgtr$ (MSE (B_a) holds iff

(34c) $\mathrm{tr}(A_2A_1^{-1}A_2 - A_1) \lesseqgtr -\beta'(A_2A_1^{-1} - I)'(A_2A_1^{-1} - I)\beta.$

The collected conditions (34a)-(34c) give

Theorem 3. Let the assumptions of theorems 1 and 2 be satisfied. Then

a) $\mathcal{E}(B_b) \leqslant \mathcal{E}(B_a) \iff \det(A_2A_1^{-1}) \leqslant 1$ under $\beta > 0$,

b) $\mathcal{E}(B_b) > \mathcal{E}(B_b) \iff \det(A_2A_1^{-1}) > 1$ under $\beta > 0$,

c) $\mathcal{D}(B_b) < \mathcal{D}(B_a)$ in a sense that $\mathcal{D}(B_b) - \mathcal{D}(B_a)$ is negative definite matrix $\iff \det(A_2A_1^{-1})^2 < 1$,

d) $\mathcal{D}(B_b) \geqslant \mathcal{D}(B_a)$, i.e. $\mathcal{D}(B_b) - \mathcal{D}(B_a)$ is non-negative definite matrix $\iff \det(A_2A_1^{-1})^2 \geqslant 1$,

e) MSE $(B_b) \lesseqgtr$ MSE $(B_a) \iff \mathrm{tr}(A_2A_1^{-1}A_2 - A_1) \lesseqgtr -\beta'(A_2A_1^{-1} - I)'(A_2A_1^{-1} - I)\beta$. ◆

Due to the fact $A_1, A_2 \in R^{k\times k}$, by Cauchy's theorem, it follows

$$\det(A_2A_1^{-1}) = \det(A_2)\det(A_1^{-1}),$$

and by (33)

$$\det(A_2) = \det(T'A_2T) = \det\Lambda_2 = \prod_{i=1}^{k}\frac{1}{(\lambda_i + \gamma)}$$

$$\det(A_1) = \det(T'A_1T) = \det\Lambda_1 = \prod_{i=1}\frac{1}{\lambda_i}.$$

Hence the necessary and sufficient conditions (a)-(d) from the theorem 3 can be replaced as follows

$$\det(A_2A_1^{-1}) \leqslant 1 \quad \text{to} \quad \prod_{i=1}^{k} \lambda_i(\lambda_i+\gamma)^{-1} \leqslant 1,$$

$$\det(A_2A_1^{-1}) > 1 \quad \text{to} \quad \prod_{i=1}^{k} \lambda_i(\lambda_i+\gamma)^{-1} > 1,$$

$$\det(A_2A_1^{-1})^2 \lesseqgtr 1 \quad \text{to} \quad \prod_{i=1}^{k} \left(\frac{\lambda_i}{\lambda_i+\gamma}\right)^2 \lesseqgtr 1.$$

[Note: for practical purposes one can use instead of an orthogonal matrix T an orthonormal matrix T calculated by the use of Jacobi algorithm].

3. Properties of estimator B_c

Using the relations (I)-(VI) from § 2, the assumptions of model NM_2 and the definitions of symbols: $\mathcal{E}$, $\mathcal{D}$, var, MSE, P, cov, one can find that

(1c) $B_c = K_cY, \quad K_c = (x'\Omega^{-1}x+\Gamma)^{-1}x'\Omega^{-1},$

(2c) $\mathcal{E}(B_c) = K_cx\beta, \quad ob(B_c) = (K_cx - I)\beta \neq 0,$

(3c) $\mathcal{D}(B_c) = K_c\Omega K_c',$

(4c) $P_{B_c} = \mathcal{N}_{B_c}(K_cx\beta, K_c\Omega K_c'),$

(5c) $MSE(B_c) = tr(K_c\Omega K_c') + \beta'(K_cx - I)'(K_cx - I)\beta,$

(6c) $B_c'B_c = Y'K_c'K_cY,$

(7c) $\mathcal{E}(B_c'B_c) = tr(K_c'K_c\Omega) + \beta'x'K_c'K_cx\beta,$

(8c) $\text{var}(B_c'B_c) = 2\text{tr}(K_c'K_c\Omega)^2 + 4\beta'x'K_c'K_c\Omega K_c'K_c x\beta,$

(9c) $P_{B_c'B_c/\sigma^2} \neq \chi^2(\cdot,\cdot)$ since $\Omega K_c'K_c\Omega K_c'K_c\Omega \neq \Omega K_c'K_c\Omega,$

(10c) $\hat{Y}_c = xB_c = xK_cY,$

(11c) $\mathcal{E}(Y_c) = xK_cx\beta,\quad \text{ob}(\hat{Y}_c) = (xK_c - I)x\beta \neq 0,$

(12c) $\mathcal{D}(Y_c) = xK_c\Omega K_c'x',$

(13c) $P_{\hat{Y}} = \mathcal{N}_{\hat{Y}_c}(xK_cx\beta,\ xK_c\Omega K_c'x'),$

(14c) $\text{MSE}(\hat{Y}_c) = \text{tr}(xK_c\Omega K_c'x') + \beta'x'(xK_c - I)'(xK_c - I)x\beta,$

(15c) $\hat{Y}_c'\hat{Y}_c = Y'K_c'x'xK_cY,$

(16c) $\mathcal{E}(\hat{Y}_c'\hat{Y}_c) = 2\text{tr}(K_c'x'xK_c\Omega) + \beta'x'K_c'x'xK_cx\beta,$

(17c) $\text{var}(\hat{Y}_c'\hat{Y}_c) = 2\text{tr}(K_c'x'xK_c\Omega)^2 + 4\beta'x'K_c'x'xK_c\Omega K_c'x'xK_cx\beta,$

(18c) $P_{\hat{Y}_c'\hat{Y}_c/\sigma^2} \neq \chi^2(\cdot,\cdot)$ since $\Omega K_c'x'xK_c\Omega K_c'x'xK_c\Omega \neq$
$\neq \Omega K_c'x'xK_c\Omega,$

(19c) $E_c = (I - xK_c)Y = M_cY,\quad M_c = I - xK_c,$

(20c) $\mathcal{E}(E_c) = M_cx\beta \neq 0,\quad \text{ob}(E_c) = M_cx\beta \neq 0,$

(21c) $\mathcal{D}(E_c) = M_c\Omega M_c',$

(22c) $P_{E_c} = \mathcal{N}_{E_c}(M_cx\beta,\ M_c\Omega M_c'),$

(23c) $\text{MSE}(E_c) = \text{tr}(M_c\Omega M_c') + \beta'x'M_c'M_cx\beta,$

(24c) $E_c'E_c = Y'M_c'M_cY,$

(25c) $\mathcal{E}(E_c'E_c) = \text{tr}(M_c'M_c\Omega) + \beta'x'M_c'M_cx\beta,$

(26c) $\mathrm{var}\,(E_c'E_c) = 2\,\mathrm{tr}\,(M_c'M_c\,\Omega)^2 + 4\beta'x'M_c'M_c \quad M_c'M_c x\beta,$

(27c) $P_{E_c'E_c/\sigma^2} \neq \chi^2(.,.)$, since $\Omega\, M_c'M_c\,\Omega\, M_c'M_c\,\Omega \neq \Omega\, M_c'M_c\,\Omega$,

(28c) $\mathrm{cov}\,(B_c, S^2_{E_c}) = \frac{2}{n_c} K_c\,\Omega\, M_c'M_c x\beta \neq 0$, $n_c = \mathrm{tr}\,(M_c'M_c\Phi)$,

$$\Omega = \sigma^2\Phi.$$

The relations (1c)-(28c) prove the following.

Theorem 4. Let the assumptions of model NM_2 and the assumptions (29), (31), (32) be satisfied. Then

a) the estimator B_c is biased, consistent, it has multivariate normal distribution but $\|B_c\|^2$ does not have χ^2 - distribution;

b) the predictor $\hat{Y}_c$ is biased, consistent, it has n-variate normal distribution, but the square of lenght of it does not have χ^2 distribution;

c) the residual vector E_c is biased, consistent, it has n-variate normal distribution with $\|E_c\|^2$ not χ^2 - distributed;

d) the covariance of B_c and $S^2_{E_c}$ is different from the zero vector. ◆

Since the matrix A_3, $A_3 = (x'\,\Omega^{-1}x + \Gamma)^{-1}$, is symmetric and positive definite, therefore $\det\,(A_3) > 0$. Simultaneous diagonalization is possible only in the case of diagonality of matrix Γ. For non-diagonal matrices Γ it holds $A_1A_3 \neq A_3A_1$. There, however, exists (see th. 6 in ch. 4 of Bellman book [2]) a non-singular matrix $\tilde{T}$ such that, due to symmetry and positive definiteness of A_1, A_3,

(35) $\tilde{T}'A_1\tilde{T} = I,\quad \tilde{T}'A_3\tilde{T} = \tilde{\Lambda},\ \tilde{\Lambda} = \mathrm{diag}\,(\tilde{\lambda}_1, \ldots, \tilde{\lambda}_k).$

Noticing that $K_c = A_3X'\,\Omega^{-1}$, $K_c\Omega K_c' = A_3A_1^{-1}A_3 = \mathcal{D}(B_c)$, $K_cx = A_3A_1^{-1}$, $\det\,(A_3A_1^{-1}) > 0$, $\det\,(A_3^2A_1^{-1}) > 0$ it is seen that

(36a) $\mathcal{E}(B_c) \leq \mathcal{E}(B_a) \overset{\beta>0}{\Longleftrightarrow} A_3A_1^{-1}\beta \leq \beta \Longleftrightarrow \det\,(A_3A_1^{-1}) \leq 1,$

(36b) $\mathcal{E}(B_c) > \mathcal{E}(B_a) \overset{\beta>0}{\Longleftrightarrow} \det\,(A_3A_1^{-1}) > 1,$

(36c) $\mathcal{D}(B_c) \lesseqgtr \mathcal{D}(B_a) \Longleftrightarrow A_3A_1^{-1}A_3 \lesseqgtr A_1 \Longleftrightarrow \det (A_3A_1^{-1})^2 \lesseqgtr 1$

in the sense that the matrix $\mathcal{D}(B_c) - \mathcal{D}(B_a)$ is non-positive (positive) definite,

(36d) $\text{MSE } (B_c) \lesseqgtr \text{MSE } (B_a) \Longleftrightarrow \text{tr } (A_3A_1^{-1}A_3 - A_1) \lesseqgtr -\beta'(A_3A_1^{-1} - I)'$

$(A_3A_1^{-1} - I)\beta.$

In the case of diagonal matrix $\Gamma = \text{diag } (\gamma_1, \ldots, \gamma_k)$ the conditions of th. 4 can be rewritten as.

$$\det (A_3A_1^{-1}) \gtreqless 1 \quad \text{to} \quad \prod_{i=1}^{k} \frac{\lambda_i}{\lambda_i + \gamma_i} \gtreqless 1,$$

$$\det (A_3A_1^{-1})^2 \lesseqgtr 1 \quad \text{to} \quad \prod_{i=1}^{k} \left(\frac{\lambda_i}{\lambda_i + \gamma_i}\right)^2 \lesseqgtr 1,$$

where

λ_i, $i = \overline{1, k}$ denote eigen values of A_1.

It is convenient to put

$$\tilde{A}_1 = \tilde{T}'A_1\tilde{T} = I, \quad \tilde{A}_3 = \tilde{T}'A_3\tilde{T} = \Lambda.$$

The matrix T can be obtained as, f.e., $\tilde{T} = S_1S_2S_3$, where S_1 is an orthogonal matrix that diagonalizes A_1 (but does not diagonalize A_3 except in the case of diagonal matrix Γ), S_2 is a diagonal matrix, i.e. $S_2 = \text{diag}\left(\frac{1}{\sqrt{\lambda_1}}, \ldots, \frac{1}{\sqrt{\lambda_k}}\right)$, λ_i, $i = \overline{1, k}$, are eigen values, of A_1, S_3 is an orthogonal matrix that diagonalizes the matrix $S_2'S_1'A_3S_1S_2$. By (35) and Cauchy's theorem it follows

$$\det (\tilde{A}_3) = \det (\tilde{T}'A_3\tilde{T}) = \det (\tilde{\Lambda}) = \prod_{i=1}^{k} \tilde{\lambda}_i, \quad \det (\tilde{A}_1) = 1.$$

Thus

(36e) $\left(\det (A_3A_1^{-1}) \gtreqless 1\right) \equiv \left(\prod_{i=1}^{k} \tilde{\lambda}_i \gtreqless 1\right)$

(36f) $(\det(A_3A_1^{-1})^2 \lessgtr 1) \equiv (\Pi\tilde{\lambda}_i^2 \lessgtr 1)$,

where $\tilde{\lambda}_i$, $i = \overline{1, k}$, are eigen values of the matrix $S_2'.S_1'\, A_3S_1S_2$. It was proved.

Theorem 5. Let the assumptions of th. 1 and 4 be satisfied. Then

a) $\mathcal{E}(B_c) \lessgtr \mathcal{E}(B_a) \iff \prod_{i=1}^{k} \tilde{\lambda}_i \lessgtr 1$, under $\beta > 0$,

b) $\mathcal{B}(B_c) \lessgtr \mathcal{B}(B_a) \iff \prod_{i=1}^{k} \tilde{\lambda}_i^2 \lessgtr 1$, under $\beta > 0$.

c) $MSE\,(B_c) \lessgtr MSE\,(B_a) \iff tr\,(A_3A_1^{-1}A_3 - A_1) \lessgtr -\beta'(A_3A_1^{-1} - I)'(A_3A_1^{-1} - I)\beta$. ◆

Bibliography

[1] Anderson T., Taylor J. (1976): Strong Consistency of Least Squares Estimates in Normal Linear Regression, "Annals of Statistics", 4, p. 788-790.

[2] Bellman R. (1960): Introduction to Matrix Analysis, New York, McGraw-Hill Co.

[3] Humak K.M.S. (1977): Statistische Methoden der Modellbildung, Band I, Berlin Akademie Verlag.

[4] Milo W. (1976, 1977): Estymacja parametrów ogólnych modeli liniowych. Cz. I, Cz. II, Technical Reports within R. III.9. Grant, Łódź University.

[5] Milo W. (1984): On Biased Regularizing Estimators. Part I, "Acta Universitatis Lodziensis", Folia oeconomica, 34, p. 47--74.

[6] Searle S. (1971): Linear Models, New York, Wiley.

[7] Srivastava M., Khatri C. (1979): An Introduction to Multivariate Statistics, New York, North Holland Co.

Władysław Milo

O OBCIĄŻONYCH ESTYMATORACH REGULARYZUJĄCYCH. CZĘŚĆ II

Celem pracy jest:

- analiza niektórych własności statystycznych regularyzujących estymatorów postaci $B_b = (x'\Omega^{-1}x + \gamma I)^{-1}x\Omega^{-1}Y$ oraz $B_c = (x'\Omega^{-1}x + \Gamma)^{-1}x'\Omega^{-1}Y$,
- opis zmodyfikowanych planów eksperymentów, których celem będzie rozszerzona analiza własności numeryczno-statystycznych pewnej rodziny regularyzujących estymatorów.

Samodzielnie udowodniono 5 nowych twierdzeń. Orzekają one:

a) o obciążoności (nieobciążoności), zgodności wielowymiarowej normalności rozkładów estymatorów B_b, B_c i B_a, predyktorów ex post wg B_a, B_b, B_c, wektorów reszt ex post wg B_a, B_b, B_c;

b) o rozkładach kwadratów długości wektorów B_a, B_b, B_c, $\hat{Y}_a$, $\hat{Y}_b$, $\hat{Y}_c$, E_a, E_b, E_c, które nie są rozkładami χ^2;

c) o kowariancjach par $(B_a, S^2_{E_a})$, $(B_b, S^2_{E_b})$, $(B_c, S^2_{E_c})$, które są odpowiednio wektorem zerowym i wektorami różnymi od zera;

d) o warunkach koniecznych i dostatecznych nierówności lub równości między $\mathcal{E}(B_{(j)})$ a $\mathcal{E}(B_a)$, $\mathcal{D}(B_{(j)})$ a $\mathcal{D}(B_a)$, MSE $(B_{(j)})$ a MSE (B_a), gdzie $j = b, c$.