A C T A   U N I V E R S I T A T I S   L O D Z I E N S I S
FOLIA OECONOMICA 123, 1992

*Władysław Milo**

NEW IDEAS IN BIASED ESTIMATION

1. INTRODUCTION

In the theory and practice of econometrics and statistics, the method and the data are all interdependent links in the research process. The econometric and statistical models are seldom correctly specified and the data are seldom free of measurement errors and the choice of methods and models are seldom free of distortions. Out of this very complicated set of problems, in this paper, we are concerned within robustness against instability of solutions with respect to β of the following system of so called normal equations

$$x'x\beta = x'Y \tag{1.1}$$

where the matrix x is n x k real matrix, Y is a random n x 1 vector with values in R^n the Euclidean space, and β is a k x 1 vector of model parameters in the stochastic equation

$$Y = x\beta + U \tag{1.2}$$

where the random vector U has normal distribution with the mathematical expectation EU = 0, the variance-covariance matrix $DU = \sigma^2 I$.

In the case of bad-conditioning of matrix x'x, i.e. in the case of $\nu_{x'x} = \frac{\lambda_{max}}{\lambda_{min}} \geqslant 30$, the solution of (1.1) is unstable with respect to small changes in the elements of vector x'y and in the elements of matrix x'x. In order to stabilize the solu-

* Lecturer at the University of Łódź.

tions of (1.1) there were many efforts to do it. Hoerl and Kennard (1970a, b) proposed the following change in the form of estimation criteria function from Legendre and Gauss' form

$$Q_o(\beta) = ||Y - x\beta||^2 = \sum_{t=1}^{n} (Y_t - x'_{t.}\beta)^2 \tag{1.3}$$

to the ridge-type I form

$$Q_{o1} = Q_o(\beta) + Q_1(\beta) \tag{1.4}$$

where

$$Q_1(\beta) = \gamma ||\beta||^2, \qquad \gamma > 0$$

or to the ridge-type II form

$$Q_{o2}(\beta) = Q_o(\beta) + Q_2(\beta) \tag{1.5}$$

where

$$Q_2(\beta) = \beta'\Gamma\beta,$$

Γ is a diagonal matrix with positive elements.

For many reasons it is good to have our model in reparametrised form, i.e.

$$\left.\begin{array}{l} Y = xVV'\beta + U \\ Y = x_*\beta_* + U \end{array}\right\}, \qquad x_* = xV, \quad \beta_* = V'\beta \tag{1.6}$$

For (1.6) our above criteria functions would take the forms

$$Q_{*o1}(\beta_*) = ||Y - x_*\beta_*||^2 + \gamma\beta'_*\beta_* \tag{1.7}$$

$$Q_{*o2}(\beta_*) = ||Y - x_*\beta_*||^2 + \beta'_*\Gamma\beta_* \tag{1.7a}$$

By using classical differential calculus rules see: Magnus, Neudecker (1988), Dwyer (1967) one can easily find the infimums of the above criteria functions as

$$B_{o1} = (\chi + \gamma I)^{-1}x'Y, \qquad \chi = x'x \tag{1.8}$$

$$B_{o1} = WB_o, \quad W = (\chi + \gamma I)^{-1}x', \quad B_o = \chi^{-1}x'Y \tag{1.8a}$$

$$B_{*o1} = (\Lambda + \gamma I)^{-1}x'_*Y = (\Lambda + \gamma I)^{-1}\Lambda \; B_{*o}, \quad B_{*o} = \Lambda^{-1}x'_*Y \tag{1.8b}$$

$$B_{o2} = (\chi + \Gamma)^{-1}x'Y \tag{1.8c}$$

$$B_{*o2} = (\Lambda + \Gamma)^{-1}x'_*Y = (\Lambda + \Gamma)^{-1} \Lambda B_{*o} \tag{1.8d}$$

The above four forms of ridge type estimators are biased estimators. Each particular form given in (1.8)-(1.8d) defines not one single estimator but a family of estimators. However, for one particular numerical value of γ (or the matrix Γ) we have one particular estimator. All of them belong to the class of regularizing estimators if we impose some conditions of this regularization. The regularization concerns the system (1.1) and stability of solutions of this system.

Under some conditions the above biased ridge estimators are shrinking (contracting) least squares estimator B_o or its reparametrized counterpart B_{*o}. These conditions are simply conditions under which the matrices W or W_* have their euclidean norms less than one. In the econometric and statistical literature there are other estimators shrinking the values of Gauss-Legendre's type estimator B_o or B_{*o}.

Out of them two are very well known and simple to use. The first is so called PC-(principle component) estimator

$$B_{o4} = B_{PC} = \sum_{j=\varkappa+1}^{k} \lambda_j^{-1} a_j v_{.j} = \left[\begin{array}{c|c} O_{\varkappa} & O_{k-\varkappa} \\ \hline O'_{k-\varkappa} & I_{k-\varkappa} \end{array}\right] B_{*o} = W_{PC} B_{*o} \qquad (1.9)$$

where $a_j = v'_{.j} X'Y$, and the second is shrunken estimator

$$B_{o5} = B_{SH} = dB_o = \sum_{j=1}^{k} d\, \lambda_j^{-1} a_j v_{.j} \qquad (1.10)$$

where

$$d = \max(0, 1 - \delta\sigma^2(||\beta||^2)^{-1}).$$

Also Sclove estimator is shrinking estimator. It has the following form

$$B_{o6} = B_{SC} = \left[\begin{array}{c|c} d\, I_{\varkappa} & O \\ \hline O & I_{k-\varkappa} \end{array}\right] B_{*o} = W_{SC} B_{*o} \qquad (1.11)$$

where d can be chosen as in B_{o4}.

As a generalisation of B_{o4} W e b s t e r et al. (1974) have defined so called latent root estimator B_{o3}. It is of the following form

$$B_{o3} = B_{LR} = \sum_{j=\varkappa}^{k} \tilde{\lambda}_j^{-1} \tilde{a}_j \tilde{v}_{.j} \qquad (1.12)$$

where

$$\tilde{a}_j = -(n-1)^{-1} s_Y^2 \mathring{v}_{oj} \left(\sum_{i=1}^{k} \tilde{v}_{oi}^{2} \tilde{\lambda}_i^{-1}\right)^{-1}$$

$$\mathring{v}_{.j} = (\mathring{v}_{oj}, \tilde{v}_{1j}, \ldots, \tilde{v}_{kj})', \quad \tilde{v}_{.j} = (\tilde{v}_{1j}, \ldots, \tilde{v}_{kj})'$$

$$s_Y^2 = (n-1)^{-1} Y'(I - n^{-1} \mathbb{1}\mathbb{1}')Y$$

and $\mathring{v}_{.j}$ is the eigen-vector of the matrix corr $[(Y \vdots x), (Y \vdots x)]$ connected with the eigen-value $\tilde{\lambda}_j$ of this correlation matrix.

Another type of shrunking is derived from minimax ideas and has the form

$$B_{o7} = \frac{\beta'\beta}{\beta'\beta + tr(x'x)} B_{*o} = d_7 B_{*o} \tag{1.13}$$

or Hemmerle ideas

$$B_{o8} = W_H B_{*o}, \quad W_H = (I + \Gamma)^{-1} \tag{1.14}$$

where W_H is a diagonal matrix with $\Gamma = \text{diag}\,(\gamma_1^*, \ldots, \gamma_k^*)$ and $0 < \gamma_i^* = (1 - 2\gamma_o - \sqrt{1 - 4\gamma_o})(2\gamma_o)^{-1}$, $\gamma \in (0, 4^{-1})$.

There are also other ideas. F a r e b r o t h e r (1975) proposes the estimator

$B_{o9} = B_{*o} + W_F(\beta - \beta_{*o})$, W_F is a k x k matrix and β_{*o} a given estimator

$$B_{10} = B_T = \frac{\beta' x' Y}{\sigma^2 + \beta' X \beta} \beta \tag{1.15}$$

There are only two of the above estimators that are ready to use in practice. These are B_{PC} and B_{LR}. Other need more or less estimation of additional non-model "parameters" that are supposed to ease stability problems in (1.1). We shall present these estimates in the next section. In Section 3 we present new ideas to find robust estimators against instability and describe their properties.

2. EMPIRICAL BIASED ESTIMATORS AND THEIR PROPERTIES

Let us present a brief list of empirical counterparts of the previously presented estimators. This list is by no means complete. It has the form

$B_{*o1}(C_1) = (\Lambda + C_1 I)^{-1} x_*' Y,$

where $C_1 = S^2_{E_o} (\max_j B^2_{*oj})^{-1}$ is an estimator of γ proposed by Hoerl, Kennard (1970a),

$B_{*o1}(C_2) = (\Lambda + C_2 I)^{-1} x_*' Y,$

where $C_2 = k S^2_{E_{*o}} (\sum_{j=1}^{k} \lambda_j B^2_{*oj})^{-1}$ is Lawless, Wang (1976) proposition,

$B_{*o1}(C_3) = (\Lambda + C_3 I)^{-1} x_*' Y,$

where $C_3 = k\, S^2_{E_{*o}} ||B_{*o}||^{-2}$ see Hoerl et al. (1975)

$B_{*o1}(C_4) = (\Lambda + C_4 I)^{-1} x_*' Y,$

where C_4 = solution with respect to γ of $S = ||B_{*o1}(\gamma)||^2$, and $S = ||B_{*o}||^2 - S^2_{E_{*o}} \operatorname{tr} \Lambda^{-1}$ if $S > 0$ or $C_4 = 0$ otherwise.

$B_{*o1}(C_5) = (\Lambda + C_5 I =)^{-1} x_*' Y,$

where C_5 = solution of $\sum_{j=1}^{k} S_E^{-2} (\lambda_j + C)^{-1} C \lambda_j B^2_{*oj} = k$ is Dempster et al. (1977) proposition of estimating γ,

$B_{*o1}(C_6) = (\Lambda + C_6 I)^{-1} x_*' Y,$

where $C_6 = n\, C_o$, $C_o = \arg\min \operatorname{tr}(I - A(C)))^{-2} ||I - A(C)y||^2$, and where the matrix $A(C) = x_*(\Lambda + nCI)^{-1} x_*'$,

$B_{*o1}(C_7) = (\Lambda + C_7 I)^{-1} x_*' Y,$

with C_7 = solution of $\sum_{j=1}^{k} (\lambda_j + \gamma)^{-2} \lambda_j = k$ proposed by Lee (1979).

There were among others also solutions proposed by Farebrother Baranchik, Dempster and so on. Similarly for the family of estimators B_{*o2}.

It is interesting to make comparisons of small sample behaviour of empirical biased estimators. Such comparisons were made by many people in many works. Some of them are: Hoerl et al., (1975), L a w l e s s, W a n g (1976), L e e (1979), D e m p s t e r et al., (1977), W i c h e r n et al. (1978), G o l u b et al. (1979), K o n a r z e w s k a, M i l o (1979-1983), M i l o (1983, 1984).

One can distinguish the following general features of these studies:

- non-similarity of scopes,
- different lists of estimators,
- different model structures,
- non-comparability of simulations plans,
- different evaluation criteria,
- different ways of choosing β, x, σ^2, corrx,
- approach to linking the correlation matrix of x and bad conditioning of matrix x.

The first group of studies is represented by D e m p s t e r et al. (1977). There were used only two criterias: MSEB, SPE. For each model structure from 160 such structures there was generated only one replication. They considered 57 estimators. The best estimator was B_{*o1} (C_5) with respect to SEB. Accidental choice of parameters values makes comparisons not so valid as others.

In the second group (see the works of Hoerl, Golub, Galarneau, Wichern, Lin, Kmenta, Trenkler) studies are made with the use of some important ways of reducing parameter space. The estimators lists are less pro-authors but rather small. There are considered a few values of n, k, det Λ, (for example n = 30, k = 2, 5, det Λ = 1.84, 12.15, 0.113, 0.0001). The vectors β are taken from the paralell direction to the eigenvectors connected with the smallest and largest eigenvalues. In Galarneau-Gibbons studies the best estimator for $\beta = v_{.k}$ is $B_{*o1}(C_5)$, and for $\beta = v_{.1}$ is $B_{*o1}(C_3)$ in the sense of MSE. The estimator $B_{*o1}(C_{GM})$ with C_{GM} defined by Galarneau-MacDonald is also good in MSE and B_{SC} and $B_{*o1}(C_5)$ are best in MSE, MAE in Lin, Kmenta studies.

The third group (see: Lawless, Wang, Galarneau) has: small list of estimators (less than 4 estimators) non-empirically oriented, excluding Lee criteria of evaluations, proauthor criterias, nonhomogenous estimators belonging to the different families. However, the choice of β is very good since it is taking directions of $v_{.1}$, $v_{.k}$, the values of σ^2 and the degree of bad-conditioning. According to these studies the best estimators are B_{PC} and B_{LR}.

The fourth group of studies i.e. Konarzewska, Milo studies is characterized by the use of 10 criterias of evaluations, the use of two principles of reduction of parameter space, and different levels of bad conditioning. The effects of correlation structure in X and bad conditioning in x were not separated. For $\beta = v_{.1}$ the best performance had $B_{*o1}(C_6)$ and next $B_{*o1}(C_3)$ with respect to all criteria. For β from the close neighbourhood of $v_{.k}$ the best was $B_{*o1}(C_5)$, $B_{*o1}(C_2)$.

It is hard to draw decisive conclusions till someone makes a series of studies that would take into account the vast list of estimators, the valid list of citeria of comparisons, good principles of reduction of parameters space, relationships between correlation structures of X and bad-conditioning of x. It is worthwhile to use also another type of biased estimators. In the next section we shall derive them.

3. NEW IDEAS IN BIASED ESTIMATION

Suppose that we want to regularize the system (1.1). One of the ways to do it is

$$(\chi + \frac{\sigma^2}{||\beta||^2} I)\beta = x'Y, \quad \chi = x'x \tag{3.1}$$

Let us denote

$$\underset{\sim}{\chi} = \chi + \frac{\sigma^2}{||\beta||^2} I, \qquad \nu_\chi = \frac{\lambda_k}{\lambda_1}$$

where λ_k is the largest and λ_1 is the smallest eigenvalue of the matrix $\chi = x'x$. The appropriate largest and smallest eigenvalues of the matrix $\underset{\sim}{\chi}$ are $\underset{\sim}{\lambda}_k$ and $\underset{\sim}{\lambda}_1$ that are, by the known facts from linear algebra, equal

$$\underset{\sim}{\lambda}_1 = \lambda_1 + \frac{\sigma^2}{||\beta||^2}, \quad \underset{\sim}{\lambda}_k = \lambda_k + \frac{\sigma^2}{||\beta||^2}$$

and hence, by simple algebraic operations, we obtain

$$\nu_{\underset{\sim}{\chi}} = \frac{\sigma^2 + \lambda_k||\beta||^2}{\sigma^2 + \lambda_1||\beta||^2}.$$

By standard calculus rules it is easy to find that

$$\frac{\partial \nu_{\underset{\sim}{\chi}}}{\partial\beta} = \frac{2\sigma^2}{m^2}(\lambda_k - \lambda_1)\beta, \quad m = \lambda_1||\beta||^2 + \sigma^2.$$

From the necessary condition of extremum we get the following form of theoretical estimator $\underset{\sim}{B}_1$ that is minimizing the criteria function

$$\underset{\sim}{Q}(\beta) = Q_o(\beta) + \nu_{\underset{\sim}{\chi}}(||\beta||^2, \sigma^2, \nu_\chi).$$

$$\underset{\sim}{B}_1 = (\chi + \frac{(\nu_\chi - 1)}{\sigma^2\lambda_1(\lambda_1 r + 1)^2} I)^{-1} x'Y,$$

where the ratio $r = \frac{||\beta||^2}{\sigma^2}$ is the ratio of the parametric part of the signal $x\beta$ and the noise variance σ^2.

In the case when the value r and its nominator and denominator is not known we shall estimate them by the use of simple estimator

$$\hat{\sigma}^2 = (n - k)^{-1} E_o' E_o, \quad E_o = M_o Y = (I - xx^+) Y, \quad \text{and} \quad B = x^+Y,$$

where the matrix $x^+ = (x'x)^{-1}x$.

So the empirical estimator connected with $\underset{\sim}{B}_1$ is the following

$$\hat{\underset{\sim}{B}}_1 = (\chi + \frac{\nu_\chi - 1}{\hat{\sigma}^2\lambda_1(\lambda_1\hat{r} + 1)^2} I)^{-1} x'Y.$$

Under the conditions of uniform boundedness of columns of x, stochastic orthogonality of U and rows of x, infinite value of the smallest eigenvalue of the matrix $A = \sum_{t=1}^{n} x_{t:}x'_{t}$.

and under

$$(\sum_{i=1}^{k} \lambda_i)^2 \geqslant k(\sum_{i=1}^{k} \lambda_i^2 - \sum_{i=1}^{k} \lambda_i),$$

$$\sum_{i=1}^{k} \frac{\lambda_i^2\beta_{*i}^2}{(\lambda_i + \gamma)^2} \to 0,$$

by the use of Chebyshev inequality the estimator $\underset{\sim}{B}_1$ is strongly convergent to the vector β.

In a very similar way by using the above assumptions and conditions one can deduce that this estimator is asymptotically normal (only one additional assumption is needed).

By similar arguments one can derive the second new family of estimators. It is based on the following regulatization

$$(\chi + \frac{\sigma^2}{\beta'\chi\beta} I)\beta = x'Y. \tag{3.2}$$

Here the ratio $r = \frac{\beta'\chi\beta}{\sigma^2}$ is the ratio of the squared length of the signal $x\beta$ to the noise variance σ^2. By similar arguments as in the previous regularization we obtain the following degree of bad conditioning of regularized matrix i.e.

$$\nu_{\underset{\sim}{\chi}} = \frac{\sigma^2 + \lambda_k \beta'\chi\beta}{\sigma^2 + \lambda_1 \beta'\chi\beta} = \frac{1}{m}.$$

It is easy to find that

$$\frac{\partial\nu_{\underset{\sim}{\chi}}}{\partial\beta} = \frac{2\sigma^2(\lambda_k - \lambda_1)}{m^2}\chi\beta.$$

Solving the necessary conditions equations we get the estimator $\underset{\sim}{B}_2$

$$\underset{\sim}{B}_2 = \frac{m^2}{m^2 + \sigma^2(\lambda_k - \lambda_1)} B_o, \qquad B_o = x^+Y.$$

For any matrix x that has $\lambda_k \neq \lambda_1$, and any model with $\sigma^2 \neq 0$, $\beta \neq 0$ the estimator B_2 is shrunking estimator with respect to the idol estimator B_o. In fact one needs to remember that $\underset{\sim}{B}_2$ is a pencil of estimators that are contracting the values of classical estimator B_o.

By using the estimators

$$\hat{m} = \hat{\sigma}^2 + \lambda_1 B_o'\chi B_o = (n - k)^{-1} Y'M_oY + \lambda_1 Y'(I - M_o)Y,$$

$$M_o = I - xx^+$$

we can define single estimator

$$\underset{\sim}{\hat{B}}_2 = \frac{\hat{m}^2}{\hat{m}^2 + \hat{\sigma}^2(\lambda_k - \lambda_1)} B_o.$$

Both the estimator $\hat{m}^2$ and $\hat{\sigma}^2$ are expressible in terms of projection matrix M_o what makes all the last estimators attractive in terms of easiness of getting analytical results referring to the properties of $\underset{\sim}{B}_2$ and $\hat{\underset{\sim}{B}}_2$.

Under suitable conditions the proposed estimators are consistent and asymptotically normally distributed. The last estimator is attractive also in other respects. Due to its convenient form it is easy tractable from the point of view of sensitivity studies of this estimator. These studies are based on derivatives of this estimator with respect to the elements of matrix x, x'x, and the index of bad-conditioning of matrix x'x.

The above estimators were invented in order to diminish negative effects of bad-conditioning. In the case of underestimation, i.e.

$$B < \beta,$$

we will use spheric metric estimators.

Under (1.2) underestimation will take place iff

$$\tilde{u}_{.j}U > 0 \quad \text{and} \quad v_{ij} < 0$$

or

$$\tilde{u}_{.j}U < 0 \quad \text{and} \quad v_{ij} > 0, \quad i, j = \overline{1, k},$$

where $\tilde{u}_{.j}$ is the eigenvector corresponding to λ_j the eigenvalue of x'x. The spheric metric from which the estimator will be derived has the form

$$\frac{||Y - x\beta||^2}{(1 + ||Y||^2)(1 + ||x\beta||^2)}$$

and its minimum is reached at

$$B_{sf} = \frac{m_2}{m_2 - 1} B, \quad B = x^+Y, \quad m_2 = 1 + ||x\beta||^2, \quad 1 = ||Y - x\beta||^2,$$

$$x^+ = (x'x)^{-1}x'.$$

Replacing m_2 and 1 with $\hat{m}_2 = 1 + ||x\beta||^2$ and $\hat{1} = ||MY||^2$, $M = I - xx^+$ we obtain the sample analogue of B_{sf}, i.e.

$$B_{sf} = \frac{\hat{m}_2}{\hat{m}_2 - \hat{1}} B = \frac{1 + Y'(I - M)Y}{1 + Y'(I - 2M)Y} B.$$

Both estimators are biased and these biases are easy to express in terms of $\beta'\chi\beta$ for nonrandom x or $\beta'\Sigma_x\beta$ for random X. One can find that under

$$\beta'\beta \geqslant \frac{(m_2 - \sigma^2 n)(2m_2 - \sigma^2(n+2))\sum_i \lambda_i}{\sigma^2 n(n+2) + m_2(2-n)}$$

we have MSE $B_{sf} \leqslant$ MSE B.

REFERENCES

Baranchik A. (1970), *A Family of Minimax Estimators of the Mean*, "Annals of Mathematical Statistics", No. 2, p. 642-645.

Belsley D. et al. (1980), *Regression Diagnostics*, Wiley, New York.

Białas J., Milo W. (1987), *Wariacyjne metody badania pęku estymatorów regularyzujących*, Work under CPBP 10.09.III.6.

Dempster A. et al. (1977), *A Simulation Study of Alternatives to ols*, "Journal of the American Statistical Association", No. 72, p. 77-106.

Dwyer P. (1967), *Some Applications of Matrix Derivatives in Multivariate Analysis*, "Journal of the American Statistical Association", p. 607-625.

Farebrother R. W. (1975), *The Minimum MSE Linear Estimator and Ridge Regression*, "Technometrics", No. 1, p. 127-138.

Fourgeaud C. et al. (1982), *Some Theoretical Results for Generalised Ridge Estimators*, CEPREMAP, No. 8207, p. 27.

Galarneau-Gibbons D. (1981), *A Simulation Study of Some Ridge Estimators*, "Journal of the American Statistical Association", p. 131-139.

Golub G. et al. (1979), *Generalized Cross-validation as a Method Chosing a Good Ridge Estimator*, "Technometrics", No. 2, p. 215-222.

Hemmerle W. et al. (1978), *Explicit and Constrained Generalized Ridge Estimation*, "Technometrics", No. 2, p. 109-120.

Hoerl A., Kennard R. (1970a), *Ridge Regression: Biased Estimation*, "Technometrics", No. 1, p. 55-67.

Hoerl A., Kennard R. (1970b), *Ridge Regression: Applications*, "Technometrics", No. 1, p. 69-82.

Hoerl A. et al. (1975), *Ridge Regression*, Communications in Statistics. Theory and Methods", No. 1, p. 105-123.

Konarzewska I., Milo W. (1979, 1980, 1983), *Własności estymatorów obciążonych*, Part 1-3, Work under R.III.9.

L e e W. (1979), *Model Estimation*, Interim Report, No. 2, The Education and Inequality in Canada Project.

L a w l e s s J., W a n g P. (1976), *A Simulation Study of Ridge and Others Regression Estimators*, "Cómmunications in Statistics. Theory and Methods", p. 307-323.

L i n K., K m e n t a J. (1980), *Some New Results on Ridge Regression Estimators*, Discussion Paper University of Michigan Ann Arbor.

M a g n u s J., N e u d e c k e r H. (1988), *Matrix Differential Calculus*, Wiley, New York.

M i l o W. (1983), *Analiza szeregów czasowych*, Łódź University Press, Łódź.

M i l o W. (1984), *On Biased Regularasing Estimators, Part 1*, "Acta Universitatis Lodziensis", No. 34, p. 47-74.

N e u d e c k e r H. (1980), *Best Quadratic Unbiased Estimation*, Report AE, University od Amsterdam.

P o l l o c k D. S. G. (1979), *The Algebra of Econometrics*, Wiley, New York.

T r e n k l e r D., T r e n k l e r G., S c h i p p B. (1985), *Updating the Ridge Estimator Computational Statistics Quarterly*, p. 135-141.

T e r a s v i r t a T. (1982), *Superiority Comparisons of Homogenous Linear Estimators*, Communications in Statistics. Theory and Methods, p. 1595-1601.

W e b s t e r J. et al. (1974), *Latent Root Regression Analysis*, "Technometrics", p. 513-522.

W i c h e r n D. et al. (1978), *A Comparison of Ridge Estimators*, "Technometrics", No. 3, p. 301-311.

Władysław Milo

NOWE IDEE OBCIĄŻONEJ ESTYMACJI I PREDYKCJI

W referacie rozważa się rodziny estymatorów, które są odporne na takie zjawiska, jak złe uwarunkowanie i niedoszacowanie parametrów w modelu $Y = X\beta + U$. Są one wyprowadzone z następujących kryteriów estymacji:

$$\varphi(\beta, \sigma, \nu_\chi) \equiv ||Y - X_\beta||^2 + \nu_{\underset{\sim}{\chi}}(\beta, \sigma, \nu_\chi)$$

oraz

$$\varphi(\beta) \equiv \frac{||Y - X_\beta||^2}{(1 + ||Y||^2)(1 + ||X_\beta||^2)}$$

gdzie:

$$\chi = X'X, \quad \underset{\approx}{\chi} = (\chi + \frac{\sigma^2}{||\beta||^2} I), \quad \nu_{\chi} = \frac{\lambda_k}{\lambda_1},$$

λ_1, λ_k to odpowiednio największa i najmniejsza wartość własna macierzy $\underset{\approx}{\chi}$;

$\nu_{\underset{\approx}{\chi}}$ jest indeksem złego uwarunkowania macierzy $\underset{\approx}{\chi}$.

Dowodzi się zgodności wyprowadzonych estymatorów.