ACTA UNIVERSITATIS LODZIENSIS
FOLIA OECONOMICA 152, 2000

*Andrzej Grzybowski**

ON UNCERTAINTY CLASSES AND MINIMAX ESTIMATION IN THE LINEAR REGRESSION MODELS WITH HETEROSCEDASTICITY AND CORRELATED ERRORS[1]

Abstract. The problem of minimax estimation in the linear regression model is considered under the assumption that a prior information about the regression parameter and the covariance matrix of random component (error) is available for the decision-maker. Two models of the uncertainty of the prior knowledge (so called *uncertainty classes*) are proposed. The first one may represent the problem of estimation for heteroscedastic model, the other may reflect the uncertainty connected with the presence of the correlation among errors. Minimax estimators for considered classes are obtained. Some numerical examples are discussed as well.

1. INTRODUCTION

Let us consider the ordinary linear regression model

$$\mathbf{Y} = \mathbf{X}\beta + \mathbf{Z} \tag{1}$$

where $\mathbf{Y}$ is an n-dimensional vector of observations of the dependent variable, $\mathbf{X}$ is a given nonstochastic $(n \times k)$ matrix with the rank k, β is a k-dimensional vector of unknown regression coefficients, $\mathbf{Z}$ is an n-dimensional vector of random errors (random components of the model). We assume $\mathrm{E}(\mathbf{Y}) = \mathbf{X}\beta$ and $\mathrm{cov}(\mathbf{Y}) = \Sigma$.

Various papers deal with the problem of the regression estimation in the presence of prior knowledge about the parameter β. Some of them study the problems where the prior information is of the form of a restricted

* Technical University of Częstochowa, Chair of Applied Mathematics.
[1] Support by KBN. Grant No. 1 HO2B 013 15 is gratefully acknowledged.

parameter space (see e.g.: Drygas, 1996; Drygas, Pilz, 1996; Girko, 1996; Hoffman, 1996; Pilz, 1996). Other papers are devoted to the problems where the prior information is expressed in terms of the probability distribution of the parameter β. In such a case the distribution of the parameter is often assumed to belong to a given class of distributions, (see: Berger, 1982; Berger, Chen, 1987; Berger, 1990; Grzybowski, 1997; Verdu, Poor, 1984). This class models the prior knowledge as well as its uncertainty, so we call it *an uncertainty class*, see Verdu, Poor (1984). Sometimes this approach leads to game theoretic formulation of the original decision problem (see: Verdu, Poor, 1984; Grzybowski, 1997). In this paper we adopt the latter approach. We introduce two uncertainty classes reflecting the uncertainty in two common situations. Section 3 is concerned with the regression estimation in heteroscedastic models. Section 4 deals with the problems associated with the presence of correlation between errors. In each case we solve the game connected with the introduced uncertainty class, i.e. we find minimax estimators and the least favourable states of nature. Some numerical examples are also presented to illustrate important features of the obtained solutions.

2. PRELIMINARY DEFINITIONS AND NOTATION

Let $L(.\,,.)$ be a quadratic loss function, i.e. $L(\beta, a) = (\beta - a)^T\mathbf{H}(\beta - a)$, for a given nonnegative definite $(k \times k)$ matrix $\mathbf{H}$. For a given estimator $\mathbf{d}$ the *risk function* $R(.\,,.)$ is defined by $R(\beta, \mathbf{d}) = \mathrm{E}_\beta L(\beta, \mathbf{d}(\mathbf{Y}))$. Let the parameter β have a distribution π. The *Bayes risk* $r(.\,,.)$ connected with the estimator $\mathbf{d}$ is defined as usual: $r(\pi, \mathbf{d}) = \mathrm{E}_\pi R(\beta, \mathbf{d})$. Let $\mathbf{D}$ denote the set of all allowable for the decision maker estimators. We assume that all of them have the Bayes risk finite. The estimator $\mathbf{d}^\# \in \mathbf{D}$ which satisfies the condition

$$r(\pi, \mathbf{d}^\#) = \inf_{\mathbf{d} \in \mathbf{D}} r(\pi, \mathbf{d})$$

is called a *Bayes estimator* (with respect to π).

Let us consider a game $\langle \Gamma, \mathbf{D}, r \rangle$, where Γ is a given class of distributions of the parameter β. Any distribution $\pi^* \in \Gamma$, satisfying the condition

$$\inf_{\mathbf{d} \in \mathbf{D}} r(\pi^*, \mathbf{d}) = \sup_{\pi \in \Gamma} \inf_{\mathbf{d} \in \mathbf{D}} r(\pi, \mathbf{d})$$

is called the *least favourable* distribution (state of nature). The *minimax estimator* is defined as the estimator $\mathbf{d}^* \in \mathbf{D}$ which satisfies the following condition:

$$\sup_{\pi \in \Gamma} r(\pi, \mathbf{d}^*) = \inf_{\mathbf{d} \in D} \sup_{\pi \in \Gamma} r(\pi, \mathbf{d}).$$

Sometimes the robustness of estimators is described in terms of the *supremum* of the Bayes risks, see e.g. Berger (1982, 1990), Verdu, Poor (1984). Then the above estimator is called *minimax-robust.*

Various problems of minimax-robust regression were discussed e.g. in Berger (1982, 1990), Berger, Chen (1987), Chaturvedi, Srivastawa (1992), Grzybowski (1997), Pinelis (1991). The relations between Bayesian analysis and minimax estimation were examined in Berger, Chen (1987), Drygas, Pilz (1996), Hoffman (1996), Pilz (1986). Review of recent results on robust Bayesian analysis and interesting references can be found in Męczarski (1998). Various classes of stable (robust, minimax and other) estimators are also discussed in Milo (1995).

Let us consider the case where our information about the parameter β is described with the help of the following class $\Gamma_{\vartheta,K}$ of distributions π:

$$\Gamma_{\vartheta,K} = \{\pi: E_\pi \beta = \vartheta, \quad \operatorname{cov}(\beta) = \Delta \in K \subset M_k\},$$

where K is a given subset of a space M_k of all positive definite $(k \times k)$ matrices. The point ϑ, fixed throughout this paper, can be thought of as a prior guess for β, while the set K reflects our uncertainty connected with the guess. Let us assume the covariance matrix Σ of the random disturbance $\mathbf{Z}$ belongs to a given subset Ω of the space M_n. Problems of minimax regression estimation in the presence of such a prior information about Σ were considered in Grzybowski (1997), Hoffman (1996), Pinelis (1991). The set $G = K \times \Omega$ is the uncertainty class in our problem.

Let us consider the situation when the set of allowable estimators $\mathbf{L}$ consists of all affine linear estimators $\mathbf{d}$, i.e. estimators having the form $\mathbf{d}(\mathbf{Y}) = \mathbf{AY} + \mathbf{B}$, with $\mathbf{A}$ and $\mathbf{B}$ being a matrix and a vector of the appropriate dimensions. The original problem of estimation of the parameter β now can be treated as a game $\langle G, \mathrm{L}, r\rangle$. The solution of the game was found in Grzybowski (1997). It was proved that if the uncertainty class G is convex then *any* affine linear estimator $\mathbf{d}^*$ which is Bayes with respect to the *least favourable pair of matrices* is minimax-robust (note that the Bayes risk for affine linear estimators is determined by the first two moments of the prior distribution). The least favourable matrices Δ^* and Σ^* satisfy the condition:

$$\operatorname{tr}(\mathbf{C}(\Delta^*, \Sigma^*)\mathrm{H} = \sup_{(\Delta, \Sigma) \in G} \operatorname{tr}(\mathbf{C}(\Delta, \Sigma)\mathbf{H}) \tag{2}$$

where $\mathbf{C}(\Delta, \Sigma) = (\mathbf{X}^T\Sigma^{-1}\mathrm{X} + \Delta^{-1})^{-1}$, while the estimator $\mathbf{d}^*$ is given by the following familiar formula:

$$\mathbf{d}^*(\Delta^*, \Sigma^*) = \mathbf{C}(\Delta^*, \Sigma^*)\mathbf{X}^T(\Sigma^*)^{-1}\mathbf{Y} + \mathbf{C}(\Delta^*, \Sigma^*)(\Delta^*)^{-1}\vartheta \quad (3)$$

Last year, during MSA'97, we examined an uncertainty class with the sets K and Ω defined as follows: $K = \{\Delta\colon \Delta = d\mathbf{I}_k, d \in (0, d_0]\}$, $\Omega = \{\Sigma\colon \Sigma = s\mathbf{I}_n, s \in (0, s_0]\}$, with given real values d_0, s_0 and $\mathbf{I}_k$, $\mathbf{I}_n$ being the identity matrices in the spaces M_k and M_n, respectively, see Grzybowski (1997), Hoffman (1996), Pinelis (1991). Practically, the class may represent the case where the uncertainties connected with each coefficient β_i are independent and *the same* while the regression model satisfies two assumptions: *homoscedasticity* and *independence of errors* $\mathbf{Z}_i$.

In the sequel of the paper we propose uncertainty classes representing the situation when the uncertainties connected with each coefficient of regression parameter β may be different and the above two assumptions about the regression model may not be satisfied.

For convenience we adopt the following notation. For any n-dimensional vectors $\mathbf{a}$, $\mathbf{b}$ we write $\mathbf{a} \geqslant \mathbf{b}$ if $a_i \geqslant b_i$, $i = 1, ..., n$. We write $\mathbf{a} > \mathbf{0}$ if all components of the vector $\mathbf{a}$ are positive. For any matrix $\mathbf{A}$ we write $\mathbf{A} > 0$ ($\mathbf{A} \geqslant 0$) if the matrix is positive (nonnegative) definite. For any vector $\mathbf{a}$ the symbol diag($\mathbf{a}$) stands for a diagonal matrix with the components of $\mathbf{a}$ on the main diagonal.

3. MINIMAX ESTIMATION IN HETEROSCEDASTIC MODELS

Let $\delta \in R^k$ and $\sigma \in R^n$ be given vectors. Let $G(\delta, \sigma) = K \times \Omega$ denote an uncertainty class where the sets K and Ω are defined as follows: $K = \{\Delta \in M_k\colon \Delta = \text{diag}(\mathbf{d}), \mathbf{0} < \mathbf{d} \leqslant \delta\}$, $\Omega = \{\Sigma \in M_n\colon \Sigma = \text{diag}(\mathbf{s}), \mathbf{0} < \mathbf{s} \leqslant \sigma\}$. The class may represent the case where the uncertainty connected with coefficients β_i of the regression parameter are different and the random errors Z_i are independent but they may have different standard deviation. The following proposition provides the minimax estimators for such problems.

Proposition 1. Let $\Delta_\delta = \text{diag}(\delta)$ and $\Sigma_\sigma = \text{diag}(\sigma)$. The estimator $\mathbf{d}^*(\Delta_\delta, \Sigma_\sigma)$ given by (3) is the minimax estimator for the game $\langle G(\delta, \sigma), \mathbf{L}, r$. □

In view of the above mentioned results, in order to prove the Proposition it is sufficient to show that the two matrices $(\Delta_\delta, \Sigma_\sigma)$ satisfy the condition (2). For this purpose we need the following lemmas.

Lemma 1. Let $\mathbf{A} = [a_{ij}]_{k \times k} > \mathbf{0}$ and $\mathbf{H} = [h_{ij}]_{k \times k} \geqslant 0$. Let $\mathbf{A}_{11}$ and $\mathbf{H}_{11}$ be the submatrices of $\mathbf{A}$ and $\mathbf{H}$, respectively, obtained by deleting the first row and the first column. Then

$$\text{tr}(\mathbf{A}^{-1}\mathbf{H}) - \text{tr}(\mathbf{A}_{11}^{-1}\mathbf{H}_{11}) \geqslant 0 \quad (4)$$

Proof of Lemma 1.

Let us write down the matrix **A** in the following form:

$$\mathbf{A}=\begin{bmatrix} a_{11} & \mathbf{w}^T \\ \mathbf{w} & \mathbf{A}_{11} \end{bmatrix}.$$

One can verify that

$$\mathrm{tr}(\mathbf{A}^{-1}\mathbf{H})-\mathrm{tr}(\mathbf{A}_{11}^{-1}\mathbf{H}_{11})=\mathrm{tr}(\mathbf{M}\mathbf{H}),$$

where

$$\mathbf{M}=c\begin{bmatrix} 1 & -\mathbf{w}^T\mathbf{A}_{11}^{-1} \\ -\mathbf{A}_{11}^{-1}\mathbf{w} & (\mathbf{A}_{11}^{-1}\mathbf{w})(\mathbf{w}^T\mathbf{A}_{11}^{-1}) \end{bmatrix}$$

with $c=(a_{11}-\mathbf{w}^T\mathbf{A}_{11}^{-1}\mathbf{w})^{-1}$.

Since $a_{11}-\mathbf{w}^T\mathbf{A}_{11}^{-1}\mathbf{w}=\dfrac{\det(\mathbf{A})}{\det(\mathbf{A}_{11})}$ we see that $c>0$.

It appears that $\mathbf{M}\geqslant 0$. Indeed, for *an arbitrary* k-dimensional vector $\mathbf{x}^T=(x_1, \mathbf{x}_2^T)$, $x_1\in\mathrm{R}$, $\mathbf{x}_2\in\mathrm{R}^{k-1}$ we have

$$\mathbf{x}^T\mathbf{M}\mathbf{x}=c(x_1-b)^2,\text{ with } b=\mathbf{w}^T\mathbf{A}_{11}^{-1}\mathbf{x}_2.$$

Now, let $\mathbf{e}^T=(1, 1, .., 1)$ and let $\mathbf{M}*\mathbf{H}$ denote the Hadamard product of the matrices **M**, **H**. Then $\mathrm{tr}(\mathbf{MH})=\mathbf{e}^T(\mathbf{M}*\mathbf{H})\mathbf{e}$, see Rao (1973). On the other hand, from the Schur lemma we know that $(\mathbf{M}*\mathbf{H})\geqslant 0$ for any matrices $\mathbf{M}\geqslant 0$ and $\mathbf{H}\geqslant 0$. It follows that $\mathbf{e}^T(\mathbf{M}*\mathbf{H})\mathbf{e}\geqslant 0$ and the proof is completed. □

Lemma 2. Let $\mathbf{A}>0$, $\mathbf{H}\geqslant 0$ be given $k\times k$ matrices. Let for $i=1, \ldots, k$ and $x>0$ functions f_i be defined as follows: $f_i(x)=\mathrm{tr}(\mathbf{A}_x^{-1}\mathbf{H})$, where $\mathbf{A}_x=[b_{jl}]_{k\times k}$ with $b_{jl}=\begin{cases} a_{ii}+\dfrac{1}{x} & \text{for } j=l=i \\ a_{jl} & \text{otherwise.}\end{cases}$

Then for each $i=1, \ldots, k$ the function f_i is non-decreasing.

Proof of Lemma 2.

Without loss of generality we may consider $i=1$. A little calculation shows that for each $x>0$ the derivative of f_1 does exist and

$$f'_1(x)=\frac{\det(\mathbf{A}_{11})\det(\mathbf{A})[\mathrm{tr}(\mathbf{A}^{-1}\mathbf{H})-\mathrm{tr}(\mathbf{A}_{11}^{-1}\mathbf{H}_{11})]}{[\det(\mathbf{A}_{11})+x\det(\mathbf{A})]^2}.$$

In view of Lemma 1 and our assumptions about the matrices **A** and **B**, the derivative is nonnegative, which completes the proof. □

Proof of the Proposition 1.

For any $\Delta = \text{diag}(d_1, d_2, ..., d_k)$ and $\Sigma = \text{diag}(s_1, s_2, ..., s_n)$, $\Delta > 0$, $\Sigma > 0$, let the function $g(d_1, ..., d_k, s_1, ..., s_n)$ be defined as follows: $g(d_1, ..., d_k, s_1, ..., s_n) = \mathrm{T}(\Delta, \Sigma) =$

$$\text{tr}\{[\mathbf{X}^T \text{diag}(\frac{1}{s_1}, ..., \frac{1}{s_n})\mathbf{X} + \text{diag}(\frac{1}{d_1}, ..., \frac{1}{d_k})]^{-1}\mathbf{H} \quad (5)$$

It is easy to check that the function g can also be expressed in the following way: $g(d_1, ..., d_k, s_1, ..., s_n) =$

$$\text{tr}[\text{diag}(d_1, ..., d_k)\mathbf{H}] - \text{tr}[\mathbf{X}\,\text{diag}(d_1, ..., d_k)\mathbf{X}^T + \text{diag}(s_1, ..., s_n)]^{-1}\mathbf{N} \quad (6)$$

where $\mathbf{M} = \mathbf{X}\,\text{diag}(d_1, ..., d_k)\,\mathbf{H}\,\text{diag}(d_1, ..., d_k)\,\mathbf{X}^T$. Note that $\mathbf{M} \geqslant 0$.

It can be seen from (5) that for each $i = 1, ..., k$ the function g as a function of d_i is of the same form as the functions f_i from Lemma 2. The relation (6) shows that g as a function of s_i has got the form: const $- f_i\left(\frac{1}{s_i}\right)$. Thus, in view of (5), (6) and Lemma 2, the function g is non-decreasing w.r.t. each variable $d_1, ..., d_k, s_1, ..., s_n$. It results that on the set $\{(\mathbf{d}, \mathbf{s}) \in \mathbf{R}^k \times \mathbf{R}^n \colon 0 < \mathbf{d} \leqslant \delta, \mathbf{0} < \mathbf{s} \leqslant \sigma\}$ it achieves its maximum at the point (δ, σ). Since the set is convex the condition (2) yields the desired result. □

One may notice that in the considered case the least favourable states of nature (and associated with them minimax estimators) are intuitive – the matrices Δ_δ, Σ_σ are connected with greatest values of variances of the regression parameter and error, respectively. In the next section we discuss a problem where there is no such "predictable" value of the least favourable state of nature.

4. MINIMAX ESTIMATION IN SOME PROBLEMS CONNECTED WITH CORRELATION BETWEEN ERRORS

Let $\mathbf{P}(\rho)$ denote a matrix with elements $p_{ij} = \frac{\rho^{|i-j|}}{1-\rho^2}$, $|\rho| < 1$. Such matrices appear in a natural way in the case where the dependence between errors can be described by the following first order autocorrelation process:

$$Z_i = \rho Z_{i-1} + V_i, \quad |\rho| < 1.$$

where $\mathbf{V} = (V_1, V_2, ..., V_n)$ is a random vector with $E(\mathbf{V}) = \mathbf{0}$ and $D^2(\mathbf{V}) = \omega\mathbf{I}_n$, $0 < \omega < \infty$. It is well-known that then $\mathrm{Cov}(\mathbf{Z}) = \omega\mathbf{P}(\rho)$.

For given constants $\omega > 0$, $-1 < \rho_1 < \rho_2 < 1$ let us consider an uncertainty class $G(\delta, \omega, \rho_1, \rho_2) = K \times \Omega$ with the set K defined as previously while $\Omega = \{\Sigma \in M_n\colon \Sigma = \Sigma(w, \alpha) = w[\alpha\mathbf{P}(\rho_1) + (1-\alpha)\mathbf{P}(\rho_2)],\ 0 < w \leqslant \omega,\ 0 \leqslant \alpha \leqslant 1\}$.

The following proposition provides the minimax estimators when the uncertainty class is $G(\delta, \omega, \rho_1, \rho_2)$.

Proposition 2. There exists a number $\alpha_0 \in [0, 1]$ such that the pair of matrices Δ_δ, $\Sigma(\omega, \alpha_0)$ is the least favourable state of nature and the estimator $\mathbf{d}^*[\Delta_\delta, \Sigma(\omega, \alpha_0)]$ is the minimax estimator in the game $\langle G(\delta, \omega, \rho_1, \rho_2), \mathbf{L}, r\rangle$. The number $\alpha_0 \in [0, 1]$ depends on the matrices $\mathbf{X}$ and $\mathbf{H}$. □

Proposition 2 states that, as in the previous case, the least favourable matrix Δ is associated with the greatest variances of β_i. On the other hand the proposition asserts that the least favourable value of the parameter α depends on the matrices $\mathbf{X}$, $\mathbf{H}$ and, in that sense, the least favourable covariance matrix of the vector $\mathbf{Z}$ is "unpredictable".

The proof of Proposition 2 is based on Lemma 2 and will be omitted. In the sequel we present some numerical examples to show how α_0 depends on the matrix $\mathbf{H}$ determining the loss. The dependence yields that our solution does not have the feature: "minimax prediction" equals "prediction based on minimax estimate of the regression parameter". The solution of the problems considered in the previous Section has got such a property.

In our examples we consider the model (1) with the following fixed values of its characteristics:

$$k = 3,\ n = 7,\ \mathbf{X} = \begin{bmatrix} 5 & 10 & 15 \\ 10 & 5 & 15 \\ -5 & 10 & -15 \\ 5 & 15 & 10 \\ -15 & 10 & 5 \\ -5 & 15 & 5 \\ 10 & -15 & -5 \end{bmatrix}.$$

In all examples we consider the class $G(\delta, \omega, 0, 0.5)$ with fixed values $\delta = (10, ..., 10)$ and $\omega = 10$. This is because we already know how the estimators depend upon these values. To simplify the notation we write $\mathrm{Tr}(\alpha)$ instead of $\mathrm{tr}\{\mathbf{C}[\Delta_\delta, \Sigma(\omega, \alpha)] \cdot \mathbf{H}\}$.

Example 1. The classical problem of estim ation of the regression parameter. In this example we consider the case where $\mathbf{H} = \mathbf{I}$, i.e. the classical problem of estimation of the parameter β. Figure 1 shows the graph of the function Tr. We can see that it has got one maximum.

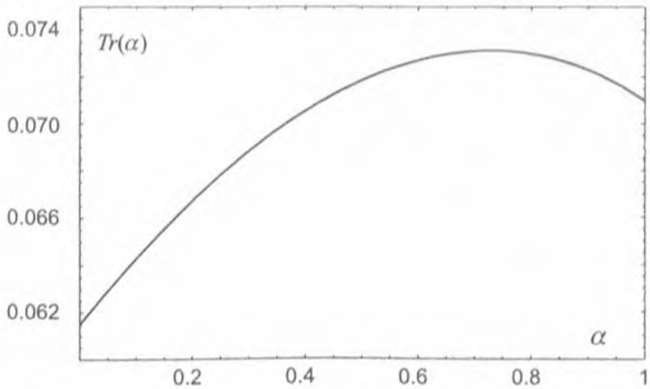

Fig. 1. The function Tr for the problem of estimation of β.

It can be numerically verified that the maximum is taken on for $\alpha_{max} = 0,728304$.

Example 2. The prediction of the dependent variable.

Now let us consider a problem of prediction of the value of the dependent variable **Y** when the independent variables take on the following values: $x_1 = 1$, $x_2 = 5$, $x_3 = 9$. The corresponding matrix determining the loss function is of the following form: $\mathbf{H} = \begin{bmatrix} 1 & 5 & 9 \\ 5 & 25 & 45 \\ 9 & 45 & 81 \end{bmatrix}$. The graph of the function Tr in this case is presented in Fig. 2. Numerical calculation shows that the only maximum of Tr is achieved for $\alpha_{max} = 0,812536$.

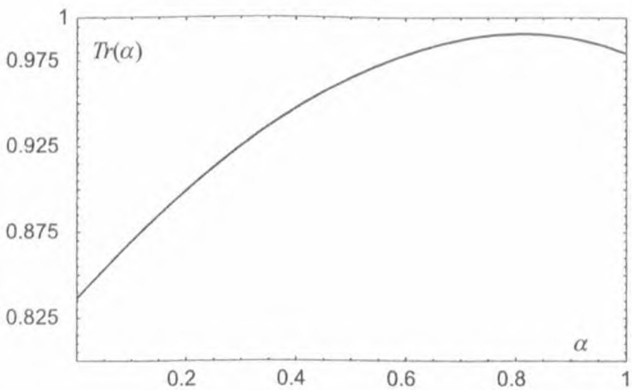

Fig. 2. The function Tr for the problem of prediction for $x_1 = 1$, $x_2 = 5$, $x_3 = 9$

The two above examples show that the least favourable value of the parameter α (and the associated covariance matrix Σ) can hardly be considered as intuitive. The value changes for different matrices H. It seems that in such situations we have different minimax estimators for various purposes (such as the estimation of regression parameter, prediction for different values of independent variables etc.) even in the same model. So, in the case of correlated errors one should be partiularly aware of the purpose of the minimax estimation.

REFERENCES

Berger J. (1982), *Bayesian Robustness and the Stein Effect*, J. Amer. Stat. Assoc. 358–368.

Berger J., Chen S. (1987), *Minimaxity of Empirical Bayes Estimators Derived from Subjective Hyperpriors*, [in:] Advances of the Multivariate Statistical Analysis, ed. A. K. Gupta, 1–12.

Berger J. (1990), *Robust Bayesian Analysis: Sensitivity to the Prior*, J. Statist. Plann. Inference, **25**, 303–328.

Chaturvedi A., Srivastawa A. K. (1992), *Families of Minimax Estimators in Linear Regression Model*, "Sankhya", **54**, Ser. **B**, 278–288.

Drygas H. (1996), *Spectral Methods in Linear Minimax Estimation*, Acta Appl. Math., **43**, 17–42.

Drygas H., Pilz J. (1996), *On the Equivalence of the Spectral Theory and Bayesian Analysis in Minimax Linear Estimation*, Acta Appl. Math., 43, 43–57.

Girko V. L. (1996), *Spectral Theory of Minimax Estimation*, Acta Appl. Math., **43**, 59–69.

Grzybowski A. (1997), *Minimaksowo-odporna estymacja parametrów regresji*, "Przegląd Statystyczny", **3**, 427–435.

Grzybowski A. (1997), *On Some Model of Uncertainty in the Minimax-Robust Multivariate Linear Regression*, Proceedings of 16th International Conference on Multivariate Statistical Analysis MSA'97, 203–212.

Hoffman K. (1996), *A Subclass of Bayes Linear Estimators that are Minimax*, Acta Appl. Math., **43**, 87–95.

Męczarski M. (1998), *Problemy odporności w bayesowskiej analizie statystycznej*, ser. "Monografie i Opracowania", **446**, SGH, Warszawa.

Milo W. (1995), *Stabilność i wrażliwość metod ekonometrycznych*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź.

Pilz J. (1986), *Minimax Linear Regression Estimation with Symmetric Parameter Restrictions*, J. Statist. Plann. Inference, **13**, 297–318.

Pinelis I. F. (1991), *O minimaksnom ocenivanii regressii*, Teorija Verojatnostej i ee primenenija, **35** (3), s. 494–505.

Rao C. R. (1973), *Linear Statistical Inference and its Applications*, John Wiley and Sons, New York. (Polish translation: *Modele liniowe statystyki matematycznej*, PWN, Warszawa 1984).

Verdu S., Poor H. V. (1984), *On Minimax Robustness: A General Approach and Applications*, IEEE Trans. Inform. Theory IT–30, 328–340.