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 132, 1993

*Hilmar Drygas**

MINIMAX ESTIMATION IN LINEAR MODELS

Abstract. We consider the linear model $y = X\beta + \sigma\varepsilon$, $E\varepsilon = 0$, $E\varepsilon\varepsilon' = I_n$ under the ellipsoidal constraints $\beta \in \mathcal{B} = \{\beta: \beta'\beta \le 1\}$. We give a review of the problems involved with the determination of the Best Linear Estimator of β in this model.

Key words: Linear models, minimax estimation, BLME.

1. INTRODUCTION

Consider the linear model

$$y = X\beta + \sigma\varepsilon, \quad E\varepsilon = 0, \quad E\varepsilon\varepsilon' = \mathrm{Cov}\,\varepsilon = I_n, \tag{1.1}$$

where X is a known $n \times k$-matrix, the design matrix, β is an unknown $k \times 1$ parameter vector and y, the observation vector, as well as ε, the unobservable disturbance term, are random $n \times 1$-vectors. σ is unknown scalar parameter. The setup implies that

$$\mathrm{Cov}\,y = \sigma^2 \mathrm{Cov}\,\varepsilon = \sigma^2 I_n. \tag{1.2}$$

If instead of (1.1) the model $\tilde{y} = \tilde{X}\beta + \sigma\tilde{\varepsilon}$, $E\tilde{\varepsilon} = 0$, $\mathrm{Cov}\,\tilde{\varepsilon} = V$, V a positive definite (p.d.) $n \times n$-matrix is given then by the transformation $y = V^{-1/2}\tilde{y}$ the model can be brought into the form (1.1) with $X = V^{-1/2}\tilde{X}$. Therefore the restriction to (1.1) is not essential from the theoretical point of view, possibly, however from the numerical point of view. The usual method to estimate β is the

---

* Professor at the University of Kassel.

method of least squares resulting in the estimator $\hat{\beta} \equiv (X'X)^{-}X'y$, where in general A^{-} is a g-inverse of matrix A, i.e. any matrix A^{-} such that $AA^{-}A = A$. However, this method is only a good one in the case that β is really unresticted. If there are some restrictions imposed on β then $\hat{\beta}$ may behave rather badly. We shall assume in the sequel that β obeys the ellipsoidal constraint

$$\beta \in \mathcal{B} = \{\beta: \beta'\beta \leq 1\}. \tag{1.3}$$

An ellipsoidal constraint of the kind

$$\beta \in \mathcal{B} = \{\beta: (\beta - \beta_0)'T(\beta - \beta_0) \leq 1\} \tag{1.4}$$

with a p.d. matrix T can be brought back to the above restriction by the reparametrization

$$\beta_1 = T^{-1/2}(\beta - \beta_0)$$

and the transformation $y \to y - X\beta_0 = \tilde{y}$, $E\tilde{y} = \tilde{X}\beta_1$, $\tilde{X} = XT^{1/2}$, and $\beta = \beta_0 + T^{1/2}\beta_1$. Therefore the restriction to the circle (1.3) is not connected with any loss of generality.

We consider linear estimators $\hat{\beta} = Ly$ of β. In order to define a risk we consider a non negative definite matrix A of order $k \times k$. If A is of rank m, we can write $A = CC'$ with a suitable $k \times m$-matrix C. Then we define the risk $R(\hat{\beta}, \beta)$ as

$$\begin{aligned} R(\hat{\beta}, \beta) &= E((\hat{\beta} - \beta)'A(\hat{\beta} - \beta)) \\ &= E((\hat{\beta} - \beta)'CC'(\hat{\beta} - \beta)) = E\,\|C'(\hat{\beta} - \beta)\|^2. \end{aligned} \tag{1.5}$$

By elementary computations we get

$$R(Ly, \beta) = \beta'(LX - I)'A(LX - I)\beta + \sigma^2 tr(L'AL).$$

The minimax-principle now consists in finding L in such a way that

$$\sup_{\beta \in \mathcal{B}} R(Ly, \beta) = \min_L \sup_{\beta \in \mathcal{B}} R(Ly, \beta).$$

This principle incorporates the prior information given from $\beta \in \mathcal{B}$, i.e., $\|\beta\| \leq 1$.

2. THE MINIMAX ESTIMATOR

It is well-known that

$$\begin{aligned} \sup_{\|\beta\| <1} \beta'(LX - I)'A(LX - I)\beta &= \max_{\|\beta\| \leq 1} \beta'(LX - I)'A(LX - I)\beta \\ &= \lambda_{max}((LX - I)'A(LX - I)); \end{aligned} \tag{2.1}$$

the largest eigenvalue of the n.n.d. matrix $(LX - I)'A(LX - I) = Q(L)$ and the maximum is attained if β is any unit eigenvector belonging to the largest eigenvalue of $Q(L)$. Therefore the optimization problem consists in

$$\text{minimize } \lambda_{max}(Q(L)) + \sigma^2 tr(L'AL) \text{ subject to } L \in \mathbb{R}_{k\times n}. \quad (2.2)$$

The difficulty now consists in the fact that $\lambda_{max}(Q(L))$ is not a differentiable function of L [G i r k o (1988), p. 74; S t a h l e c k e r (1985); S t a h l e c k e r and L a u t e r b a c h (1989), p. 2758]. Since $Q(L) = (LX - I)'CC'(LX - I)$ and $\lambda_{max}((LX - I)'CC'(LX - I)) = \lambda_{max}(C'(LX - I)(LX - I)'C)$ [see G i r k o (1988), p. 73] the minimization of (2.2) is equivalent to the minimization of

$$Z(L) = \lambda_{max}(\tilde{Q}(L)) + \sigma^2 tr(L'AL), \quad (2.3)$$

where $\tilde{Q}(L) = C'(LX - I)(LX - I)'C$. Clearly, $\tilde{Z}(L) = \lambda_{max}(\tilde{Q}(L))$ is an expression of the kind

$$\tilde{Z}(L) = \max_{\|\alpha\| \le 1} a'W(L)a \quad (2.4)$$

where $W(L)$ is of the form $A_0LBB'L'A_0' + DLE + E'L'D' + F$ with suitable matrices A_0, B, D, E, and F. It is easy to verify (using Cauchy-Schwarz inequality) that $W(L)$ is a convex function of L. Hence $Z(L)$ is a convex function of L, too and a local minimum of $Z(L)$ is a global minimum as well. Thus L is optimal iff for any matrix Θ of order $k \times n$

$$\frac{d}{d\gamma} Z(L + \gamma\Theta)\Big|_{\gamma=0} = 0, \quad (2.5)$$

provided that the derivative exists. The differentiation of $\frac{d}{d\gamma}\sigma^2 tr((L+\gamma\Theta)'A(L+\gamma\Theta))$ is of course not very difficult and yields the value for $\gamma = 0$ equal to $2tr(AL\Theta')$. It is more difficult to find the derivative of $\lambda_{max}(\tilde{Q}(L + \gamma\Theta))$, evaluated at $\gamma = 0$. However, [G i r k o (1988), p. 64 ff. and p. 74] has found the corresponding expression. Let e_1 be a unit eigenvector of $\tilde{Q}(L) = C'(LX - I)(LX - I)'C$ corresponding to the largest eigenvalue of $\tilde{Q}(L)$. Then

$$\frac{d}{d\gamma}\lambda_{max}(\tilde{Q}(L+\gamma\Theta))|_{\gamma=0} = e_1'C'(LX-I)X'\Theta'C\,e_1$$
$$= tr(e_1'C'(LX-I)X'\Theta'C\,e_1) \qquad (2.6)$$
$$= tr(C\,e_q'\,e_1'C'(LX-I)X'\Theta').$$

This holds in the case that $\lambda_{max}(\tilde{Q}(L))$ is a simple eigenvalue. It is still unclear what happens if λ_{max} is a multiple eigenvalue of $\tilde{Q}(L)$. Thus

$$\frac{d}{d\gamma}(\lambda_{max}(\tilde{Q}(L+\gamma\Theta)) + \sigma^2\, tr((L+\gamma\Theta)'A(L+\gamma\Theta)))|_{\gamma=0} \qquad (2.7)$$
$$= 2tr([C_{e_1\,e_1'}C'(LX-I)X' + \sigma^2 Al]\Theta') = 0$$

is the necessary and sufficient optimality condition. Since this is to hold for Θ, (2.7) is equivalent to the equation

$$C_{e_1e_1'}C'(LX-I)X' + \sigma^2CC'L = 0. \qquad (2.8)$$

The solution of this equation is very simple if $C = b$, a $k \times 1$-vector. Then $C'(LX-I)(LX-I)'C = b'(LX-I)(LX-I)'\,b$ is a number and trivially the largest eigenvalue. Thus $e_1 = 1$ (or $e_1 = -1$) and (2.8) becomes

$$CC'[(LX-I)X' + \sigma^2L)] = 0. \qquad (2.9)$$

A solution independent of b is given by the solution of the equation

$$(LX-I)X' + \sigma^2L = 0 \qquad (2.10)$$

which gives the ridge estimator

$$Ly = \hat{\beta} = (X'X + \sigma^2I)^{-1}X'y.$$

It seems that this approach should be followed further and it should be compared with the results obtained by other methods.

An alternative approach was given by Läuter [L ä u t e r (1975), H o f f m a n n (1979), S t a h l e c k e r (1985), p. 111)]. Let

$$S = \sigma^{-2}X'X = (\sigma^2(X'X)^{-1})^{-1}, \qquad (2.11)$$

$$F = S^{-1}\,AS^{-1}, \qquad (2.12)$$

then necessary and sufficient for the existence of a solution Ly of the Minimax-problem is the existence of a n.n.d. matrix V and a positive real number v such that

$$D_{v,V} = \frac{1}{\sqrt{v}} (F + V)^{1/2} - S^{-1} \text{ is n.n.d.,} \quad \text{(i)}$$

$$\frac{1}{\sqrt{v}}(F + V)^{-1/2}V = SV, \quad \text{(ii)}$$

$$\frac{1}{\sqrt{v}} \operatorname{tr}((F + V)^{1/2}) = 1 + \operatorname{tr}(S^{-1}). \quad \text{(iii)}$$

If these conditions are met then a minimax-estimator is given by the ridge estimator

$$\hat{\beta} = (X'X + D_{v,V})^{-1} X'y. \quad (2.13)$$

If $A = I$, i.e. $C = I$, then $V = 0$ and $v = (\operatorname{tr}(S^{-1})(1 + \operatorname{tr} S^{-1})^{-1})^2$ meets the conditions (i), (ii), and (iii). The minimax-estimator is the shrunken estimator

$$\hat{\beta} = \frac{1}{1 + \sigma^2 \operatorname{tr}((X'X)^{-1})} (X'X)^{-1}X'y, \quad (2.14)$$

$$L = (1 + \sigma^2\operatorname{tr}((X'X)^{-1}))^{-1}(X'X)^{-1}X', \quad \text{i.e.}$$

$$LX - I = -(1 + \sigma^2\operatorname{tr}((X'X)^{-1}))^{-1}\sigma^2\operatorname{tr}(X'X)^{-1} I,$$

$$C(LX - I)(LX - I)'C' = vI.$$

v is the maximal eigenvalue, but a multiple one. The Girko-equation (2.8) is then equivalent to

$$X'X\, e_1e_1'\, X'X = \frac{1}{(\operatorname{tr}(X'X)^{-1})} X'X \quad (2.15)$$

and means that e_1e_1' is proportional to a generalized inverse of X'X. This can only happen if Rank(X) is equal to zero or one. But this case is ruled out by the assumption of regularity of X'X.

Unfortunately, there is no constructive way of finding V, v besides some special cases as that one discussed above. Therefore it is necessary to apply numerical methods for the determination of the minimax-estimator Ly. Such numerical methods were developed by Stahlecker [S t a h l e c k e r (1985), p. 141 ff.; S t a h l e c k e r-L a u t e r b a c h (1989), p. 2757 ff.]. He uses a p-norm-approximation of the largest eigenvalue. Let Q be any k × k-matrix. Then

$$(\operatorname{tr}(1/kQ^p))^{1/p} \le \lambda_{max}(Q) \le (\operatorname{tr} Q^p)^{1/p} \quad (2.16)$$

for any positive integer p and

$$\lim_{p\to\infty} (\operatorname{tr}(1/kQ^p))^{1/p} = \lim_{p\to\infty} (\operatorname{tr}(Q^p))^{1/p} = \lambda_{max}(Q). \quad (2.17)$$

Therefore it seems convenient to replace the minimization problem arising from Minimax-Estimation by either

$$\text{Minimize} \quad 1/k(\mathrm{tr}((\tilde{Q}(L))^p))^{1/p} + \sigma^2 \,\mathrm{tr}(L'AL) \tag{2.18}$$

or by

$$(\mathrm{tr}((\tilde{Q}(L))^p))^{1/p} + \sigma^2 \,\mathrm{tr}(L'AL). \tag{2.19}$$

Stahlecker proves that there are $k \times n$-matrices L_p and $\bar{L}_p$ which solve the minimization problems (2.18) and (2.19), respectively. There are numerical procedures known from optimization theory [the authors quote Dennis and Schnabel (1983) and Göpfert (1973)]. Moreover, they can show that there exists a subsequence $\{\bar{L}_{pj}\}$, $pj \in \mathbb{N}$ such that $\lim_{j\to\infty} \bar{L}_{pj} = L$ and L is a minimax-estimator. They are also able to obtain estimates on $||L_{pj} - L||$. ($||A|| = (\mathrm{tr}(A'A))^{1/2}$). These estimates allow to stop computations as soon as a required accuracy is obtained.

3. AFFINE AND ELLIPSOIDAL RESTRICTIONS

Consider the linear model $Ey = X\beta$, $\mathrm{Cov}\; y = \sigma^2 I$ and assume that besides the ellipsoidal restrictions $||\beta|| \leq 1$ also affine restrictions $R\beta = r$ are given, where R is a given $s \times k$-matrix (without restricting generality $\mathrm{Rank}(R) = s$ can be assumed and $r \in \mathbb{R}^s$). Let therefore

$$\mathcal{A} = \{\beta \in \mathbb{R}^k : R\beta = r\} \tag{3.1}$$

At a first glance it may seem that these additional restrictions do not produce any new statistical problem what minimax-estimation is concerned. The reason is that we could consider r as additional observations with zero covariance. Thus we could build up the linear regression model

$$E\begin{pmatrix} y \\ r \end{pmatrix} = \begin{pmatrix} R \\ R \end{pmatrix} \beta, \quad \mathrm{Cov}\begin{pmatrix} y \\ r \end{pmatrix} = \sigma^2 \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \tag{3.2}$$

under the ellipsoidal constraint $||\beta|| \leq 1$. A minimax-estimator can easily be calculated, the corresponding formula can be found in [Drygas (1991), section 1]. We call this estimator naive minimax-estimator. It coincides with the minimax-estimator to be discussed later if $r = 0$. The main shortcoming of the naive minimax-estimator is that he does not react on the possible situation $\mathcal{A} \cap \mathcal{B} = \emptyset$ or the situation $\mathcal{A} \cap \mathcal{B} = \{t_*\}$, a single point in which case the naive minimax-estimator is not equal to t_* with probability one.

One may argue that the problem of a non-consistent statistical model may occur in many situations. Let us mention a few examples. Consider the linear model $A\,Ey = b$, $\mathrm{Cov}\;y = V$, where A is some linear mapping. If the equation $Ax = b$ is consistent, then a Best Linear Unbiased Estimator (BLUE) of Ey is given by

$$Gy = (I - V\,A'(AV\,A')^{-}A)\,y + V\,A'(AV\,A')^{-}b. \tag{3.3}$$

But what does this formula mean if the equation $Ax = b$ is inconsistent, i.e., contradictory? Under the assumption $\mathrm{im}((AV\,A')^{-}AV) \subseteq \subseteq \mathrm{im}(A)$ - $\mathrm{im}(A) = \{y : y = Ax\}$ - (which is correct if the g-inverse is the Moore-Penrose inverse) Gy is the BLUE of Ey in the model $A\,Ey = AA^{+}b$, $\mathrm{Cov}\;y = V$. Thus Gy is not the BLUE in the given (inconsistent) model but in a different consistent model.

Consider the model $r = R\beta$, $\mathrm{Cov}\;r = 0$. Then $R^{-}r$ has the property that $a'R^{-}r$ is BLUE of $a'\beta$ in this model for any estimable $a'\beta$ provided the equation $r = R\beta$ is consistent. If the equation $R\beta = r$ is not consistent and R^{-} is a reflexive g-inverse of R (i.e., $R^{-}RR^{-} = R^{-}$), then $a'R^{-}r = a'R^{-}RR^{-}r$ is BLUE of $a'\beta$, $a'\beta$ estimable, in the model $RR^{-}r = R\beta$, $\mathrm{Cov}\;RR^{-}r = 0$. Again, the meaningless estimator has changed to a meaningful estimator in a different model.

Let us now consider the consistent equation $R\beta = r$, $\mathrm{Cov}\;r = 0$. If in addition to this, ellipsoidal restrictions $\beta'T\beta \le 1$ are given then Ar is approximate minimax-estimator of β iff $ART^{-1}R' = T^{-1}R'$ [see D r y g a s (1991), formula (1.14) in section 1]. This implies $RART^{-1}R' = RT^{-1}R'$, which is again equivalent to $RAR = R$. It shows that A is a g-inverse of R. Therefore we realize that the additional information $\beta'T\beta \le 1$ reduces the choice of g-inverses R^{-}. However, the information may be destructive and contradictory, leading to an inconsistent model. Inspection of the equation $ART^{-1}R' = T^{-1}R'$ shows that it is not changed if T is replaced by αT, $\alpha > 0$. Since there will be at least one $\alpha > 0$ such that both restrictions are simultaneously consistent, again the approximate minimax-estimator is an approximate minimax-estimator in a meaningful model. This consideration can readily be carried over to the general model $Ey = X\beta$, $\mathrm{Cov}\, y = \sigma^2 V$: Just replace V by αV and T by $\alpha^{-1}T$ (the restriction is then $\beta' T\beta \leq 1$).

The naive minimax-estimator can nevertheless be used if in advance it has been checked that neither $\mathcal{A}\cap\mathcal{B}$ is empty (inconsistent model) nor $\mathcal{A}\cap\mathcal{B}$ consists of a single point (in which case there is no estimation problem at all). This check is not necessary if one uses the minimax-estimator defined by [S t a h l e c k e r and T r e n k l e r (1988)] as follows: C + d is called Minimax-Estimator if

$$\sup_{\beta\in\mathcal{A}\cap\mathcal{B}} R(Cy + d, \beta) = \min_{\tilde{C},\tilde{d}} \sup_{\beta\in\mathcal{A}\cap\mathcal{B}} R(\tilde{C}y + \tilde{d}, \beta). \tag{3.4}$$

In the course of the computation the quantity

$$\alpha = 1 - r'(R'R)^{-}r$$

is computed. $\alpha \geq 0$ is the necessary and sufficient condition that $\mathcal{A}\cap\mathcal{B} \neq \emptyset$. If $\alpha = 0$, then $\mathcal{A}\cap\mathcal{B}$ consists of a single point t_* and the minimax-estimator is just this point. If $\alpha = 1$, i.e., $r = 0$, then the naive minimax-estimator and the minimax-estimator coincide.

4. ELIMINATION AND REPARAMETRIZATION METHODS

Let us consider as simple example the regression model $Ey = x_1\beta_1 + x_2\beta_2$, $\mathrm{Cov}\, y = \sigma^2 I$ with β_1 and β_2 scalars. Assume that we have the ellipsoidal restriction $\beta_1^2 + \beta_2^2 \leq 1$ and the affine restriction $\beta_1 + \beta_2 = 1$. Of course you would try to eliminate β_2 by $\beta_2 = 1 - \beta_1$ implying $\beta_1^2 + \beta_2^2 = \beta_1^2 + (1-\beta_1)^2 = 2\beta_1^2 - 2\beta_1 + 1 = 2(\beta_1 - \frac{1}{2})^2 + 1 - \frac{1}{2} = 2(\beta_1 - \frac{1}{2})^2 + \frac{1}{2} \leq 1$ or equivalently $(\beta_1 - \frac{1}{2})^2 \leq \frac{1}{4}$, which again is equivalent to $|\beta_1 - \frac{1}{2}| \leq 1$ or $\beta_1 \in [0,1]$. So we could write the regression model as well as

$$E(y - x_2) = (x_1 - x_2)\beta_1, \quad \mathrm{Cov}(y - x_2) = \sigma^2 I, \tag{4.1}$$

and the ellipsoidal constraint - here indeed an interval-constraint - $|\beta_1 - \frac{1}{2}|^2 \leq 1$. Indeed in this model one would estimate β_1 by the usual minimax-method resulting in an estimator $\hat{\beta}_1$ · β_2 will then be estimated by $\hat{\beta}_2 = 1 - \hat{\beta}_1$. The question, however, is: What are the properties of this estimator? When computing the estimator $\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2)'$ it turns out that it coincides with the minimax-estimator of β in the Stahlecker-Trenkler sense. This result gives rise to an investigation of elimination and reparametrization methods in linear models with affine and ellipsoidal restrictions.

Let us discuss at first elimination methods. We can assume that $\text{Rang}(R) = s$ (as already done earlier) because otherwise there are redundant or contradictory restrictions. Let R_0 be determined in such a way that

$$H = \begin{pmatrix} R \\ \cdots \\ R_0 \end{pmatrix} \tag{4.2}$$

is a regular $k \times k$-matrix (R_0 is then a $(k-s) \times k$-matrix). If $R = (R_1 \vdots R_2)$, $R_1 \in \mathbb{R}_{s\times s}$, $\text{Rang}(R_1) = s$, then $R_0 = (0 : I_{k-s})$ may be a special choice of R_0 leading to the elimination of β_1.

Let $\gamma = (\gamma_1', \gamma_2')' = H\beta = (r', (R_0\beta)')'$. If $H^{-1} = (H_1 \vdots H_2)$ with $H_1 \in \mathbb{R}_{k\times s}$, $H_2 \in \mathbb{R}_{k\times(k-s)}$, then

$$X\beta = X_1 r + X_2\gamma_2, \; X_1 = XH_1, \; X_2 = XH_2. \tag{4.3}$$

Moreover, [see D r y g a s (1991), section 1], it can be shown that $\beta \in \mathcal{B} \cap \mathcal{A}$ iff $(\gamma_2 - \gamma_{2,0})'T_{22}(\gamma_2 - \gamma_{2,0}) \le \alpha$ for some $\alpha \in \mathbb{R}$ and some p.d. T_{22}. $\alpha \ge 0$ holds iff $\mathcal{B} \cap \mathcal{A} \ne \emptyset$ and $\alpha = 0$ iff $\mathcal{B} \cap \mathcal{A}$ consists of a single point t_*. If $\alpha \ge 0$ then the minimax-estimator $\hat{\gamma}_2$ of γ_2 in the model $E(y - X_1 r) = X_2\gamma_2$, $\text{Cov}(y - X_1 r) = \sigma^2 I$ under the ellipsoidal restriction $(\gamma_2 - \gamma_{2,0})'T_{22}(\gamma_2 - \gamma_{2,0}) \le \alpha$ has the property that

$$\hat{\beta} = H_1 r + H_2\hat{\gamma}_2 \tag{4.4}$$

is a minimax-estimator of β. Thus, in a sloppy way, we can formulate:

<u>Theorem</u>. Elimination methods do not destroy minimax-estimators.

Another method to get estimators is the method of reparametrization. Since $R\beta = r$ we get from $\beta = R^-R\beta + (I - R^-R)\beta$, where R^- is a g-inverse of R, i.e., $RR^-R = R$, that

$$\beta = R^-r + (I - R^-R)\beta. \tag{4.5}$$

Indeed $\{x : x = R^-r + (I - R^-R)\beta_1\} = \{x : Rx = r\}$. Let us therefore consider the model

$$E(y - XR^{-}r) = X(I - R^{-}R)\beta_1, \; Cov\,(y - XR^{-}r) = \sigma^2 I \tag{4.6}$$

under the ellipsoidal restrictions $\beta_1'\beta_1 \le 1$. We can estimate β_1 by the minimax-method. Thus we will get an estimator $\hat{\beta}_1$. We estimate β by

$$\hat{\beta} = R^{-}r + (I - R^{-}R)\hat{\beta}_1. \tag{4.7}$$

Then $R\hat{\beta} = r$. This is a continuum of estimators for β. Unless $r = 0$ neither of these estimators coincides with the minimax-estimator. If $R^{-} = R'(RR')^{-}$ $(= R'(RR')^{-1} = R^{+}$ if $\operatorname{Rank}(R) = s)$, then $\hat{\beta}$ coincides with the naive minimax-estimator [D r y g a s (1988)].

Instead of reparametrizing the regression function also the elliposidal constraint could be reparametrized by (4.5). This leads to a model with singular ellipsoidal constraints (see next section). A minimax-estimator in this model exists (β is minimax-estimable) and one version of it is a minimax-estimator in the model $Ey = X\beta$, $Cov\; y = \sigma^2 I$, $\beta \in \mathcal{A} \cap \mathcal{B}$, too.

5. SINGULAR RESTRICTIONS

We consider the linear model

$$Ey = X_1\beta_1 + X_2\beta_2 + \varepsilon, \quad Cov\;\varepsilon = Cov\; y = \sigma^2 I \tag{5.1}$$

under the constraint

$$\beta_1\beta_1 \le 1. \tag{5.2}$$

This is the canonical form of a linear model with singular ellipsoidal constraints, i.e., any linear model with singular ellipsoidal restrictions can be brought into this form after some reparametrization. Consider estimators of β of the form

$$Cy = \begin{pmatrix} C_1 y \\ \cdots \\ C_2 y \end{pmatrix} = \begin{pmatrix} C_1 \\ \cdots \\ C_2 \end{pmatrix} y, \tag{5.3}$$

where $C_i \in \mathbb{R}_{k,xn}$, $X_i \in \mathbb{R}_{nxk}$, $i = 1,2$; $k_1 + k_2 = k$. There are several approaches available in the literature to cope with such problems. We will only discuss our own results [D r y g a s (1985)]. Unless $C_2X_2 = I$ it follows that

$$\sup_{\beta \in \mathcal{B}} E(Cy - \beta)'A(Cy - \beta) = \infty, \tag{5.4}$$

where $\mathcal{B} = \{\beta = (\beta_1', \beta_2')' : \beta_1'\beta_1 \le 1\}$. Thus Cy has to be a partial unbiased estimator. If $C_2X_2 = I$ is not consistent, then only functions of the kind $D_1\beta_1 + D_2X_2\beta_2$ are minimax-estimable. In the case $A = aa'$ the minimax-estimators are obtained by computing the BLUE of $D_1\beta_1 + D_2X_2\beta_2$ in the artifical linear model

$$E\begin{pmatrix} y \\ \dots \\ 0 \end{pmatrix} = \begin{pmatrix} X_1 & X_2 \\ I & 0 \end{pmatrix}\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}, \quad \mathrm{Cov}\begin{pmatrix} y \\ \dots \\ 0 \end{pmatrix} = \begin{pmatrix} \sigma^2 I & 0 \\ 0 & I \end{pmatrix} \tag{5.5}$$

0 is considered as an artifical observation with expectation β_1, uncorrelated with y and with covariance-matrix I_{k_1}.

REFERENCES

D e n n i s J. E. (jr), S c h n a b e l (1983): *Numerical methods for Unconstrained Optimization and Nonlineare Equations*, Prentice-Hall, New Jersey.

D r y g a s H. (1985): *Minimax Prediction in Linear Models*, [in:] *Linear Statistical Inference* (T. Caliński and W. Klonecki, eds.), Proceedings of the International Conference held at Poznań, Poland, June 4-8, 1984, p. 48-60.

D r y g a s H. (1988): *A note on Minimax-Estimation in Regression Models with Affine Restrictions*, "Kasseler Mathematische Schriften", 2/88 (June).

D r y g a s H., *Linear Minimax-Estimation in Linear Models with affine and Ellipsoidal Restrictions*, 1991 [in:] *Computational Statistics and Data Analysis*.

G i r k o V. L. (1988): *Multidimensional Statistical Analysis*, Wyča Szkola, Kiew (Russian); an English translation, (Kluwer, Doredrecht; 1993) is in preparation.

G ö p f e r t A. (1973): *Mathematische Optimierung in allgemeinen Vektorräumen*, BSG B. G. Teubner Verlagsgesellschaft, Leipzig.

H o f f m a n n K. (1979): *Characterization of Minimax Linear Estimators in Linear Regression*, Math. Operationsf. u. Statistik, Series Statistics, 10, p. 19-26.

L ä u t e r H. (1975): *A minimax Linear Estimator for Linear Parameters Under Restrictions in Form of Inequalities*, Math. Operationsf. u. Statistik, Ser. B, Statistics, 6, p. 689-695.

S t a h l e c k e r P. (1985): *A priori Information und Minimax-Schätzung im Linearen Regressions- modell*, Habilitationsschrift an der Universität Hannover, appeared also in 1987 as No. 108 of Mathematical Systems in Economics, Frankfurt am Main.

S t a h l e c k e r P. and L a u t e r b a c h J. (1989): *Approximate Lineare Minimax Estimation in Regression Analysis with Ellipsoidal Constraints*, "Communications in Statistics; Theory and Methods", 18, no. 8, p. 2755-2784.

S t a h l e c k e r P. and T r e n k l e r G. (1988): *Full and Partial Minimax Estimation with additional Linear Constraints*, "Linear Algebra and its Applications", 111, p. 279-292.

Hilmar Drygas

ESTYMACJA MINIMAX W MODELACH LINIOWYCH

Rozważamy model liniowy $y = X\beta + \sigma\varepsilon$, $E\varepsilon = 0$, $E\varepsilon\varepsilon' = I_n$ przy ograniczeniach elipsoidalnych $\beta \in \mathfrak{B} = \{\beta : \beta'\beta \leq 1\}$. Podajemy przegląd problemów związanych z determinacją Najlepszego Liniowego Estymatora Minimaxowego β w tym modelu.