Tatra Mt. Math. Publ. **17** (1999), 311–318

Tatra Mountains
Mathematical Publications

LINEAR MINIMAX ESTIMATION IN THE THREE PARAMETERS CASE

Hilmar Drygas

ABSTRACT. We consider the linear model $Ey = X\beta$, cov $y = I$ under the restriction $\|\beta\| \le 1$. The Minimax-$B'B$-estimator is the estimator $\hat{\beta} = Cy$ minimizing the maximal $B'B$-mean square error. The solution heavily depends on the dimension j of maximal eigenspace of $B(CX-I)(CX-I)'B'$. If $k = 3$, only the case $j = 2$ poses essential problems. It is shown that if $X'X$ and $B'B$ possess a joint eigenvector, a simple solution of the Linear Minimax Problem may be possible.

Consider the standard linear model

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

and the additional information that

$$\beta \in \mathfrak{B} = \{\beta\colon \|\beta\|^2 = \beta'\beta \le 1\}\,. \tag{1.2}$$

An intuitive way of incorporating this restriction is to determine the Minimax Linear Estimator (MILE) of β, i.e., the minimizer of

$$\begin{aligned} \psi(C) &= \sup_{\beta \in B} E\big(\|B(Cy-\beta)\|^2\big) \\ &= \lambda_{\max}(CX-I)'B'B(CX-I) + \operatorname{tr}\big((BC)(BC)'\big) \\ &= \lambda_{\max}\big(B(CX-I)(CX-I)'B'\big) + \operatorname{tr}\big((BC)(BC)'\big)\,, \end{aligned} \tag{1.3}$$

where $\hat{\beta} = Cy$ is the estimator and $B \in \mathbb{R}_{m\times k}$ is of rank m. $A = B'B$ is called the loss-matrix. The minimization of $\psi(C)$ is a convex optimization problem with a unique solution BC, but $\psi(C)$ is in general not differentiable.

Let us introduce some notation. Let $B(CX-I) = F$ be called the bias-matrix. It follows from the admissibility of the minimax-estimator that $BC = (B+F)X^+$, where X^+ is the Moore-Penrose generalized inverse of X. Because of this relation, the problem is solved if the bias-matrix F is computed.

2000 Mathematics Subject Classification: 62J05.
Keywords: linear model, minimax estimation, eigenspace, eigenvector.

Denote by the maximal eigenspace the eigenspace belonging to the maximal eigenvalue of FF'. Christopeit and Helmes [1], Drygas and Pilz [3] and Girko [5] have proved that a necessary and sufficient condition for F to be optimal is that there is an ON-basis $\{u_i, \dots, u_j\}$ of the maximal eigenspace whose dimension is supposed to be j, and numbers $p_i, \dots, p_j$; $p_i \geq 0$, $\sum_{j=1}^{j} p_i = 1$ such that

$$\sum_{i=1}^{j} p_i u_i u_i' F X' + (F + B) X^{+} = 0 . \tag{1.4}$$

This equation is also equivalent to

$$\sum_{i=1}^{j} p_i u_i u_i' F X' X + (F + B) = 0 \tag{1.5}$$

which is more convenient to work with. If we introduce the matrices $U_1 = (u_i, \dots, u_j)$ and $P = \operatorname{diag}(p_i, \dots, p_j)$ then $P \geq 0$, $\operatorname{tr} P = 1$, $U_1' U_1 = I_j$ and (1.5) can be written as

$$U_1 P U_1' F X' X + (F + B) = 0 . \tag{1.6}$$

Equations (1.4), (1.5) or (1.6) are called left spectral equations. If we use the singular value decomposition

$$F = \lambda_1 U_1 V_1' + U_2 \Lambda_2 V_2' , \tag{1.7}$$

$U_1' U_2 = 0$, $V_1' V_1 = I_j$, $V_1' V_2 = 0$, $U_2' U_2 = I_{k-j}$, $V_2' V_2 = I_{k-j}$, $\Lambda_2 \geq 0$, Λ_2 diagonal, then it is easily found that

$$F V_1 P V_1' X' X + (B + F) = 0 . \tag{1.8}$$

This equation is called right spectral equation. From it easily follows that Cy is a Bayesian estimator of β such that $E(\beta\beta') = V_1 P V_1' = M$, $\operatorname{tr}(M) = 1$.

In the case of full rank B and X, Läuter [7] and Hofmann [6] have found that M can be represented in the form

$$M = a_0^{\frac{1}{2}} (F + A)^{\frac{1}{2}} - S^{-1}, \quad S = X' X , \quad F_0 = S^{-1} B' B S^{-1}, \tag{1.9}$$

where $a_0 > 0$, $A \geq 0$ and, of course, $M \geq 0$. Drygas and Läuter [4] have found the relation with the spectral equation and the singular value decomposition:

$$a_0^{\frac{1}{2}} = \lambda_1^{-1}, \qquad A = S^{-1} V_2 \big(\lambda_1^2 I_{k-j} - \Lambda_2^2 \big) V_2' S^{-1}, \tag{1.10}$$

where, of course, since $\operatorname{tr}(M) = 1$, the relation

$$\lambda_1 = \frac{\operatorname{tr}(A + F)^{\frac{1}{2}}}{1 + \operatorname{tr} S^{-1}} \tag{1.11}$$

must hold. Moreover, $A = CC'$, where

$$C = S^{-1} V_2 \big(\lambda_1^2 I_{k-j} - \Lambda_2^2\big)^{\frac{1}{2}} \tag{1.12}$$

and $C \in \mathbb{R}_{k \times k-j}$. There is a necessary and sufficient condition-equation that A must fulfill, in order to yield the solution, namely

$$SA = A_1 (A + F_0)^{-\frac{1}{2}} A \tag{1.13}$$

(see Läuter [7], Hofmann [6], Drygas and Läuter [4]).

From (1.12) it follows that (1.13) is equivalent to

$$SC = \lambda_1 (F_0 + CC')^{-\frac{1}{2}} C\,, \tag{1.14}$$

$C \in \mathbb{R}_{k \times k-j}$.

In Drygas [2] it was shown that a solution in the case $j = 1$ and $j = k$ can easily be computed. If $j = 1$, then

$$Cy = \rho\big(B^{-1} u_1\big)\big(B^{-1} u_1\big)' X' y\,, \tag{1.15}$$

where u_1 is a unit-eigenvector of $Q = B(I + X'X)^{-1} B'$ and ρ is the corresponding eigenvalue. Moreover,

$$\lambda_1^2 = \rho^2 \big(u_1' (BB')^{-1} u_1\big) \geq \lambda_{\max} (I - u_1 u_1') BB' (I - u_1 u_1') \tag{1.16}$$

must hold. If there are no multiple eigenvalues of Q and none of its eigenvectors meets (1.16), the case $j = 1$ can be ruled out.

The case $j = k$ holds iff $A \equiv 0$ solves the Hofmann–Läuter equation, i.e., iff

$$\frac{1}{\lambda_1} (F_0)^{\frac{1}{2}} - S^{-1} \geq 0\,. \tag{1.17}$$

If $k = 3$ and $j = 1$, $j = 3$ can be ruled out, then $j = 2$ is valid. In this case the non-linear eigenvalue problem (1.14)

$$\mathrm{S}c = \lambda_1 (F_0 + cc')^{-\frac{1}{2}} c\,, \qquad \lambda_1 = \frac{\operatorname{tr}(cc' + F_0)^{\frac{1}{2}}}{1 + \operatorname{tr} S^{-1}}\,, \tag{1.18}$$

must be solved, perhaps iteratively, $c \in \mathbb{R}_{k \times 1}$.

2. The solution in the case $k = 3$, $j = 2$

A solution of the non-linear eigenvalue problem (1.18) can be found if there is a common eigenvector of $S = X'X$ and $F_0 = (X'X)^{-1} B'B (X'X)^{-1} = S^{-1} A S^{-1}$. This is equivalent to finding a common eigenvector of $B'B = A$ and $S = X'X$. We remark that if $B'BX'X = X'XB'B$ then $B'B = A$ and

$S = X'X$ are jointly diagonalizable. In this case the solution of the Linear Minimax Problem was found by Hofmann [6].

Let c_0, $\|c_0\| = 1$ be a joint eigenvector of S and F_0 and

$$F_0 = \sum_{i=1}^{k-1} \lambda_i u_i u_i' + \lambda_0 c_0 c_0' \tag{2.1}$$

be a spectral decomposition of F_0. We want to determine $\lambda \in \mathbb{R}$ in such a way that $c = \lambda c_0$, $\lambda \neq 0$, meets the Hofmann–Läuter equation (1.18). Since then

$$F_0 + cc' = F_0 + \lambda^2 c_0 c_0' = \sum_{i=1}^{k-1} \lambda_i u_i u_i' + (\lambda^2 + \lambda_0) c_0 c_0' \tag{2.2}$$

it follows that

$$(F_0 + cc')^{-\frac{1}{2}} = \left(\sum_{i=1}^{k-1} \lambda_i^{-\frac{1}{2}} u_1 u_i' + (\lambda^2 + \lambda_0)^{-\frac{1}{2}} c_0 c_0' \right), \tag{2.3}$$

$$\lambda_1 (F_0 + cc')^{-\frac{1}{2}} c = \lambda_1 (\lambda^2 + \lambda_0)^{-\frac{1}{2}} \lambda c_0 . \tag{2.4}$$

If, moreover, $\mathrm{S}c_0 = \mu_0 c_0$ then

$$\mathrm{S}c = \lambda \mu_0 c_0 = \lambda_1 (\lambda^2 + \lambda_0)^{-\frac{1}{2}} \lambda c_0 \tag{2.5}$$

or

$$(\lambda^2 + \lambda_0)^{-\frac{1}{2}} \lambda_1 = \mu_0 \tag{2.6}$$

must hold. Finally

$$\begin{aligned} \operatorname{tr}(cc' + F_0)^{\frac{1}{2}} &= \operatorname{tr} \left(\sum_{i=1}^{k-1} \lambda_i^{\frac{1}{2}} u_i u_i' + (\lambda^2 + \lambda_0)^{\frac{1}{2}} c_0 c_0' \right) \\ &= \sum_{i=1}^{k-1} \lambda_i^{\frac{1}{2}} + (\lambda^2 + \lambda_0)^{\frac{1}{2}} \\ &= \operatorname{tr} F_0^{\frac{1}{2}} + (\lambda^2 + \lambda_0)^{\frac{1}{2}} - \lambda_0^{\frac{1}{2}} \\ &= \operatorname{tr} F_0^{\frac{1}{2}} + \frac{\lambda^2}{(\lambda^2 + \lambda_0)^{\frac{1}{2}} + \lambda_0^{\frac{1}{2}}} . \end{aligned} \tag{2.7}$$

Thus from (2.6) we get the non-linear equation

$$(\lambda^2 + \lambda_0)^{-\frac{1}{2}} \left(\operatorname{tr} F_0^{\frac{1}{2}} + \left(\lambda_0^2 + \lambda^2 \right)^{\frac{1}{2}} - \lambda_0^{\frac{1}{2}} \right) = \mu_0 \left(1 + \operatorname{tr} S^{-1} \right) . \tag{2.8}$$

If this equation in λ is solved, then λc_0 is a solution of the Hofmann-Läuter equation. If, moreover, $\frac{1}{\lambda_1}(F_0 + \lambda^2 c_0 c_0')^{\frac{1}{2}} - S^{-1} \geq 0$ we have found a solution of the Linear Minimax Problem in the case $j = k - 1$.

2.1 Theorem. *Let c_0 be a common eigenvector of $S = X'X$ and $F_0 = S^{-1}B'BS^{-1}$ and $Sc_0 = \mu_0 c_0$, $F_0 c_0 = \lambda_0 c_0$. If λ is a solution of the non-linear equation*

$$(\lambda^2 + \lambda_0)^{\frac{1}{2}} \mu_0 (1 + \operatorname{tr} S^{-1}) = \operatorname{tr} F_0^{\frac{1}{2}} + (\lambda^2 + \lambda_0)^{\frac{1}{2}} - \lambda_0^{\frac{1}{2}} \tag{2.9}$$

and

$$\left(F_0 + \lambda^2 c_0 c_0'\right)^{\frac{1}{2}} \geq \lambda_1 S^{-1} \tag{2.10}$$

then $A = \lambda^2 c_0 c_0'$ solves the Läuter-Hofmann equation, where

$$\lambda_1 = \frac{\operatorname{tr}\left(F_0 + \lambda^2 c_0 c_0'\right)^{\frac{1}{2}}}{1 + \operatorname{tr} S^{-1}} . \tag{2.11}$$

In this case the Bayes-estimator with respect to

$$E(\beta\beta') = M = \frac{1}{\lambda_1}\left(F_0 + \lambda^2 c_0 c_0'\right)^{\frac{1}{2}} - S^{-1} \tag{2.12}$$

is the Linear Minimax Estimator of β.

A considerable simplification is obtained if $B = B' = A^{\frac{1}{2}}$. Then $A^{\frac{1}{2}} c_0 = \lambda_0^{\frac{1}{2}} \mu_0 c_0$ and

$$u_3 = c_0 , \quad v_3 = -\frac{B' u_3}{\| B' u_3 \|} = -\frac{\lambda_0^{\frac{1}{2}} \mu_0 c_0}{\lambda_0^{\frac{1}{2}} \mu_0} = -c_0 , \quad \lambda_2^2 = \lambda_0 \mu_0^2 \tag{2.13}$$

(see D r y g a s ([2], section 3)).

$$c = \left(\lambda_1^2 - \lambda_2^2\right)^{\frac{1}{2}} S^{-1} v_3 = -\left(\lambda_1^2 - \lambda_2^2\right)^{\frac{1}{2}} \mu_0^{-1} c_0 = -\left(\frac{\lambda_1^2}{\mu_0^2} - \lambda_0\right)^{\frac{1}{2}} c_0 . \tag{2.14}$$

Thus define

$$\lambda = \left(\frac{\lambda_1^2}{\mu_0^2} - \lambda_0\right)^{\frac{1}{2}} . \tag{2.15}$$

Then $(\lambda^2 + \lambda_0)^{\frac{1}{2}} \mu_0 = \lambda_1$ and from (2.9) we get $\lambda_1\left((1 + \operatorname{tr} S^{-1}) - \mu_0^{-1}\right) = \operatorname{tr} F_0^{\frac{1}{2}} - \lambda_0^{\frac{1}{2}}$, i.e.,

$$\lambda_1 = \frac{\operatorname{tr} F_0^{\frac{1}{2}} - \lambda_0^{\frac{1}{2}}}{(1 + \operatorname{tr} S^{-1}) - \mu_0^{-1}} . \tag{2.16}$$

3. An example

Let

$$(X'X) = \begin{pmatrix} \frac{15}{44} & \frac{9}{11} & 0 \\ \frac{9}{11} & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix} , \qquad B = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \mu \end{pmatrix} , \tag{3.1}$$

$\mu > 0$. We assume that $\mu \leq 0.51048$. In this case $j = 1$ and $j = 3$ can be ruled out. We get

$$B'^{-1}(I + X'X)B^{-1} = \begin{pmatrix} \frac{59}{11} & \frac{18}{11} & 0 \\ \frac{18}{11} & 4 & 0 \\ 0 & 0 & 2\mu^{-2} \end{pmatrix}. \tag{3.2}$$

The eigenvalues are

$$2\mu^{-2}, \quad \frac{71}{11}, \quad \frac{32}{11}, \tag{3.3}$$

and the corresponding unit-eigenvectors are

$$(0,0,1)', \quad \frac{1}{\sqrt{13}}(3,2,0)' \quad \text{and} \quad \frac{1}{\sqrt{13}}(-2,3,0)'. \tag{3.4}$$

Hence the eigenvalues ρ_1, ρ_2, ρ_3 of $B(I + X'X)^{-1}B'$ are

$$\mu^2/2, \quad \frac{11}{71} \quad \text{and} \quad \frac{11}{32}. \tag{3.5}$$

It is easily checked that for none of these eigenvalues and the corresponding eigenvectors the inequality (1.16) is met if $\mu \leq 0.51048$. (Indeed, $\mu > 2$ is a necessary and sufficient condition in order that $j = 1$ is valid.) Thus the case $j = 1$ can be ruled out.

It follows from straightforward computations that

$$S^{-1} = (X'X)^{-1} = \frac{11}{57}\begin{pmatrix} 44 & -12 & 0 \\ -12 & 5 & 0 \\ 0 & 0 & \frac{57}{11} \end{pmatrix} = \begin{pmatrix} 8.491228 & -2.315789 & 0 \\ -2.315789 & 0.964912 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$F_0 = S^{-1}B'BS^{-1} = \left(\frac{11}{57}\right)^2 \begin{pmatrix} 628 & -192 & 0 \\ -192 & 61 & 0 \\ 0 & 0 & \left(\frac{57}{11}\right)^2\mu^2 \end{pmatrix}. \tag{3.6}$$

Hence $c_0 = (0,0,1)'$ is a common eigenvector of S and F_0 and $\lambda_0 = \mu^2$, $\mu_0 = 1$, $1 + \mathrm{tr}(S^{-1}) = 11.45614\ldots$ The eigenvalues of F_0 are $\left(\frac{11}{57}\right)^2 686.89779 = 25.58160\ldots$, $\left(\frac{11}{57}\right)^2 2.10221\cdots = 0.07829\ldots$ and μ^2. It roots are $5.05783\ldots$, $0.27981\ldots$ and μ. Hence $\mathrm{tr}\big(F_0^{\frac{1}{2}}\big) = \mu + 5.33764$ and in the case $j = 3$ would obtain

$$\lambda_1 = \frac{\mu}{11.45614} + \frac{5.33764}{11.45614} = \frac{\mu}{11.45614} + 0.46592. \tag{3.7}$$

It is, however, easily checked that

$$-S^{-1} + \frac{1}{\lambda_1}F_0^{\frac{1}{2}} \tag{3.8}$$

is not n. n. d. if $\mu \leq 0.51098$. Thus the case $j = 3$ can be ruled out as well. (Indeed, $\mu \in (0.51048,\, 0.87525)$ is a necessary and sufficient condition for $j = 3$ to hold.)

It remains to investigate the case $j = 2$. $(0,0,1)' = c_0$ is a joint eigenvector of F_0 and S. Therefore possibly

$$M = \lambda_1^{-1}(F_0 + \lambda^2 c_0 c_0')^{\frac{1}{2}} - S^{-1}\,. \tag{3.9}$$

Since $B = B' = A^{\frac{1}{2}}$ we can apply the remarks at the end of the previous section. From $\mu_0 = 1$ and $\lambda_0 = \mu^2$ it follows that $\lambda_2 = \lambda_0^{\frac{1}{2}}\mu_0 = \mu$ and

$$\lambda_1 = \frac{\operatorname{tr} F_0^{\frac{1}{2}} - \lambda_0^{\frac{1}{2}}}{\operatorname{tr} S^{-1}} = \frac{5.33764}{10.45614} = 0.510479\,. \tag{3.10}$$

So it is already clear that our approach cannot work unless $\mu = \lambda_2 < \lambda_1 = 0.510479$. We get $\lambda_1^2 = 0.260589$ and $\lambda = -(0.260589 - \mu^2)^{\frac{1}{2}}$.

It remains to check whether M as defined by (2.12) is n. n. d. We get

$$\begin{aligned}
-S^{-1} + \frac{1}{\lambda_1}(F_0 + cc')^{\frac{1}{2}} &= M \\
&= V_1 P V_1' \\
&= -\begin{pmatrix} 8.49123 & -2.31579 & 0 \\ -2.31579 & 0.96491 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad\qquad (3.11)\\
&\quad + \frac{1}{0.510479}\begin{pmatrix} 4.64687 & -1.33964 & 0 \\ -1.33964 & 0.69077 & 0 \\ 0 & 0 & 0.51048 \end{pmatrix} \\
&= \begin{pmatrix} 0.61173 & -0.30849 & 0 \\ -0.30849 & 0.38827 & 0 \\ 0 & 0 & 0 \end{pmatrix} \geq 0\,,
\end{aligned}$$

which again follows from $(\lambda^2 + \lambda_0)^{\frac{1}{2}} = \frac{\lambda_1}{\mu_0} = \lambda_1$. Thus indeed we can use M to compute the Bayes and hence Minimax-Estimator.

Acknowledgement. I thank the organizers of the conference (Kalas, Pázman, Volaufová, Witkowsky, Wimmer) for their invitation. Moreover, I thank the Institute of Applied Mathematics in Heidelberg and the Dean, Prof. Dahlhaus, for the opportunity to do this work at their institute.

REFERENCES

[1] CHRISTOPEIT, N.—HELMES, K.: *Linear minimax estimation with ellipsoidal constraints*, Acta Appl. Math. **43** (1996), 1–12.

[2] DRYGAS, H.: *Spectral methods in linear minimax estimation*, Acta Appl. Math. **43** (1996), 17–42.

[3] DRYGAS, H.—PILZ, J.: *On the equivalence of spectral theory and Bayesian analysis in minimax linear estimation*, Acta Appl. Math. **43** (1996), 43–57.

[4] DRYGAS, H.—LÄUTER, J.: *On the representation of the minimax linear estimator in the convex linear model*, in: Proceedings of the International Conference on Linear Statistical Inference (Linstat '93) (T. Caliński and R. Kala, eds.), Kluwer Academic Publishers, Dordrecht, 1996, pp. 13–26.

[5] GIRKO, V. L.: *Spectral theory of minimax estimation*, Acta Appl. Math. **43** (1996), 59–69.

[6] HOFMANN, K.: *Characterization of minimax linear estimation in linear regression*, in: Mathematische Operationsforschung und Statistik, Series Statistics, Vol 10, 1979, pp. 19–26.

[7] LÄUTER, H.: *A minimax linear estimator for linear parameters under restrictions in the form of inequalities*, in: Mathematische Operationsforschung und Statistik, Series Statistics, Vol 6, 1975, pp. 689–695.

Received May 11, 1998

Universität Gesamthochschule Kassel
Fachbereich 17
Mathematik-Informatik
FB 17 - Postf. 1011380
Heinrich-Plett-Straße 40
D–34132 Kassel
GERMANY

E-mail: drygas@mathematik.uni-kassel.de