Chin. Ann. of Math.
12 B:1(1991), 45—57.

GAMMA-MINIMAX ESTIMATORS FOR THE MEAN VECTOR OF A MULTIVARIATE NORMAL DISTRIBUTION

CHEN LANXIANG (陈兰祥)*

Abstract

Γ-minimax estimators are determined for the mean vector of a multivariate normal distribution under arbitrary squared error loss. Thereby the set Γ consists of all priors whose vector of first moments and matrix of second moments satisfy some given restrictions. Necessary and sufficient conditions are derived which ensure a prior being least favourable in Γ and the unique Bayes estimator with respect to this prior being Γ-minimax. By applying these results the Γ-minimax estimator is explicitly found in some special cases or can be computed by solving a system of non-linear equations or by minimizing a quadratic form on a compact and convex set.

§ 1. Introduction

Eestimating the mean vector of a multivariate normal distribution is a common statistical problem. It arises for example from regression models (see e. g. [1], p. 236). In this paper the problem of determining Γ-minimax estimators under arbitrary squared error loss is considered. Thereby the covariance matrix of the normal distribution is assumed to be known. The subset Γ of priors is fixed by imposing restrictions on the vector of first moments and the matrix of second moments. Similar sets Γ are considered by Solomon (1972)[6]. However in [6] the analysis is restricted to linear estimators. In the univariate case with Γ consisting of all priors with fixed first and second moments the Γ-minimax estimator is already known (see [1], Example 4.29, and [3]).

In the third section the unique Bayes estimators with respect to a class of normal priors in Γ and their risk functions are determined. In the fourth section the basic characterization of the Γ-minimax estimator is proved. This characterization shows that determining the Γ-minimax estimator is equivalent to determining the stationary point of a compact and convex subset of an Euclidean space $\mathbf{R}^p$. By applying this result in two special cases the Γ-minimax estimator is ex-

---

Manuscript received March 14, 1988.

* Department of Applied Mathematics, Tongji University, Shanghai, China.

plicitly found. In the fifth section a geometric characterization of the Γ-minimax estimator is proved. This result shows that the Γ-minimax estimator can often be determined by solving a system of nonlinear equations. The sixth section shows that the Γ-minimax estimator can be calculated by minimizing a quadratic form on a compact and convex subset of $\mathbf{R}^p$ if the loss function, the covariance matrix of the normal distributions, and the subset Γ of prior satisfy a certain condition. Examples are presented at the end of the paper where the Γ-minimax estimator is explicitly found.

§2. Notation and Preliminary Results

The mean vector $\theta\in\mathbf{R}^p$ of a multivariate mormal distribution with known, symmetric, and positive definite covariance matrix Σ is to be estimated under arbitrary squared error loss

$$s(\theta, a)=(\theta-a)^T R(\theta-a), \quad \theta, a\in\mathbf{R}^p, \tag{1}$$

where R denotes a symmetric and positive definite matrix.

Let Π be the set of all priors, i. e. Borel probability measures on $\mathbf{R}^p$, for which

$$m(\pi)=\left(\int\theta_i\pi(d\theta)\right)_{1\leqslant i\leqslant p}\in\mathbf{R}^p,$$

the vector of first moments, and

$$M(\pi)=\left(\int\theta_i\theta_j\pi(d\theta)\right)_{1\leqslant i,j\leqslant p},$$

the symmetric and positive semi-definite matrix of second moments, exist. Let $\leqslant$ denote the partial ordering on the set of symmetric $p\times p$-matrices defined by $A\leqslant B$ if $B-A$ is positive semi-definite.

In the sequel convex subsets of priors of the form

$$\Gamma=\{\pi\in\Pi\,|\,m(\pi)\in V,\ M(\pi)\leqslant M\}\neq\varnothing$$

are considered, where the closed and convex set $V\subset\mathbf{R}^p$ and the positive definite matrix M are fixed. In the whole paper the same results are obtained if subsets

$$\tilde{\Gamma}=\{\pi\in\Pi\,|\,m(\pi)\in V,\ M(\pi)=M\}$$

instead of Γ are considered. Put

$$E_M=\{m\in\mathbf{R}^p\,|\,m\cdot m^T\leqslant M\}$$

and $V_M=V\cap E_M$. The following lemma shows that E_M is an ellipsoid.

Lemma 1. *The set E_M satisfies*

$$E_M=\{m\in\mathbf{R}^p\,|\,m^T M^{-1}m\leqslant 1\}.$$

In particular E_M is compact and convex.

Proof A simple calculation shows

$$\begin{bmatrix}1 & 0\\ m & I\end{bmatrix}\cdot\begin{bmatrix}1 & 0\\ 0 & M-mm^T\end{bmatrix}\cdot\begin{bmatrix}1 & m^T\\ 0 & I\end{bmatrix}=\begin{bmatrix}1 & m^T\\ m & M\end{bmatrix}$$
$$=\begin{bmatrix}1 & m^TM^{-1}\\ 0 & I\end{bmatrix}\cdot\begin{bmatrix}1-m^TM^{-1}m & 0\\ 0 & M\end{bmatrix}\cdot\begin{bmatrix}1 & 0\\ M^{-1}m & I\end{bmatrix},$$

where I denotes the identity matrix. Hence $M-mm^T$ is positive semi-definite if and only if $1-m^TM^{-1}m\geqslant 0$ since M is symmetric and positive definite.

By the Schwarz inequality $\pi\in\Gamma$ implies $m(\pi)\in E_M$ and therefore $m(\pi)\in V_M$. Hence

$$\Gamma=\{\pi\in\Pi\,|\,m(\pi)\in V_M,\ M(\pi)\leqslant M\}. \tag{2}$$

If $m\in V_M$ then the normal distribution $\pi=N(m,\ M-mm^T)$ is a prior in Γ. This implies that $\Gamma\neq\varnothing$ if and only if $V_M\neq\varnothing$. In particular Lemma 1 shows that the set V_M is compact and convex.

Let Δ be the set of all (non-randomized) estimators, i. e. Borel measurable functions δ; $\mathbf{R}^p\to\mathbf{R}^p$. The Bayes risk of an estimator $\delta\in\Delta$ with respect to a prior $\pi\in\Pi$ is defined by

$$r(\pi,\ \delta)=\int R(\theta,\ \delta)\,\pi(d\theta), \tag{3}$$

where $R(\cdot,\ \delta)$ denotes the risk function of δ given by

$$R(\theta,\ \delta)=\frac{1}{\sqrt{(2\pi)^p\det\Sigma}}\int_{\mathbf{R}^p}(\theta-\delta(x))^TR(\theta-\delta(x))\cdot e^{-\frac{1}{2}(x-\theta)^T\Sigma^{-1}(x-\theta)}dx,\quad \theta\in\mathbf{R}^p. \tag{4}$$

An estimator $\delta^*\in\Delta$ with

$$\sup_{\pi\in\Gamma} r(\pi,\ \delta^*)=\inf_{\delta\in\Delta}\sup_{\pi\in\Gamma} r(\pi,\ \delta)$$

is called Γ-minimax estimator, i. e. a Γ-minimax estimator minimizes the maximum Bayes risk with respect to the elements of Γ.

§ 3. A Class of Bayes Estimators with Respect to Normal Priors

In the sequel the normal priors

$$\pi_M=N(m,\ M-mm^T)\in\Gamma$$

are considered where $m\in V_M$. It is well known (see e. g. [1], Example 4.9 and p. 162, and [7], Theorem 2.2.7 and Remark 2.2.2) that under squared error loss (1) the linear estimator $\delta_M\in\Delta$ with

$$\delta_M(x)=(I-\Sigma(\Sigma+M-mm^T)^{-1})x+\Sigma(\Sigma+M-mm^T)^{-1}m,\ x\in\mathbf{R}^p,$$

is the unique Bayes estimator with respect to π_M (except a set of Lebesgue measure zero). The Sherman-Morrison formula (see e. g. [5], 2.3.11)

$$(\Sigma+M-mm^T)^{-1}=(\Sigma+M)^{-1}+\frac{(\Sigma+M)^{-1}mm^T(\Sigma+M)^{-1}}{1-m^T(\Sigma+M)^{-1}m}$$

and a short calculation yield

$$\delta_m(x)=\left[M(\Sigma+M)^{-1}-\frac{\Sigma(\Sigma+M)^{-1}mm^T(\Sigma+M)^{-1}}{1-m^T(\Sigma+M)^{-1}m}\right]x$$
$$+\frac{\Sigma(\Sigma+M)^{-1}}{1-m^T(\Sigma+M)^{-1}m}m,\ x\in\mathbf{R}^p. \tag{5}$$

The risk function (4) of a linear estimator $\delta(x)=Ax+b$, $x\in\mathbf{R}^p$, A being a $p\times p$-matrix, $b\in\mathbf{R}^p$, has the form

$$R(\theta,\delta)=((I-A)\theta-b)^TR((I-A)\theta-b)+\operatorname{tr}(A^TRA\Sigma),\quad \theta\in\mathbf{R}^p,$$

where $\operatorname{tr}(\cdot)$ denotes the trace of a matrix. In case of the linear Bayes estimator δ_m in (5) this yields

$$R(\theta,\delta)=(\theta-m)^TB(m)(\theta-m)+c(m),\ \theta\in\mathbf{R}^p,$$

where

$$B(m)=\left[I+\frac{(\Sigma-M)^{-1}mm^T}{1-m^T(\Sigma+M)^{-1}m}\right](\Sigma+M)^{-1}\Sigma R\Sigma(\Sigma+M)^{-1}$$
$$\cdot\left[I+\frac{mm^T(\Sigma+M)^{-1}}{1-m^T(\Sigma+M)^{-1}m}\right]$$

is a symmetric and positive definite matrix and

$$c(m)=\operatorname{tr}(C(m))$$

denotes the trace of the matrix

$$C(m)=(\Sigma+M)^{-1}\left[M-\frac{mm^T(\Sigma+M)^{-1}\Sigma}{1-m^T(\Sigma+M)^{-1}m}\right]R$$
$$\cdot\left[M-\frac{\Sigma(\Sigma+M)^{-1}mm^T}{1-m^T(\Sigma+M)^{-1}m}\right](\Sigma+M)^{-1}\Sigma.$$

Hence the Bayes risk (3) of δ_m with respect to a prior $\pi\in\Pi$ is given by

$$r(\pi,\delta_m)=\operatorname{tr}(B(m)M(\pi))-2m^TB(m)m(\pi)+m^TB(m)m+c(m). \tag{6}$$

§4. The Basic Characterization of the Γ-Minimax Estimator

If a parameter $\tilde{m}\in V_M$ satisfies condition (7) in the following first theorem then the normal prior $\pi_{\tilde{m}}=N(\tilde{m},M-\tilde{m}\tilde{m}^T)$ is least favourable in Γ, i. e.

$$\inf_{\delta\in\Delta}r(\pi_{\tilde{m}},\delta)=\sup_{\pi\in\Gamma}\inf_{\delta\in\Delta}r(\pi,\delta),$$

and the Bayes estimator $\delta_{\tilde{m}}$ with respect to $\pi_{\tilde{m}}$ is Γ-minimax. The second theorem shows that there exists exactly one parameter $\tilde{m}\in V_M$ which fulfils condition (7) in Theorem 1. The proof of Theorem 2 uses a stationary point result of Browder and Karamardian which follows from a minimax inequality of Ky Fan.

Theorem 1. *If $\tilde{m}\in V_M$ satisfies*

$$d(\tilde{m})^T\tilde{m}=\inf\{d(\tilde{m})^Tm\mid m\in V_M\}, \tag{7}$$

then $\delta_{\tilde{m}}$ *according to* (5) *is the unique* Γ*-minimax estimator (except a set of Lebesgue measure zero) and the prior* $\pi_{\tilde{m}}=N(\tilde{m}, M-\tilde{m}\tilde{m}^T)$ *is least favourable in* Γ, *whereby*

$$d(m)=[(\Sigma+M)^{-1}mm^T+(1-m^T(\Sigma+M)^{-1}m)I]$$
$$\cdot(\Sigma+M)^{-1}\Sigma R\Sigma(\Sigma+M)^{-1}m,\ m\in V_M.$$

Proof First note that

$$B(m)m=(1-m^T(\Sigma+M)^{-1}m)^{-2}d(m),\ m\in V_M. \tag{8}$$

Since $B(\tilde{m})$ is symmetric and positive definite, there exists a non-singular matrix A such that $B(\tilde{m})=AA^T$. Therefore

$$\mathrm{tr}(B(\tilde{m})(M-M(\pi)))=\mathrm{tr}(A^T(M-M(\pi))A)\geqslant 0,$$

since $M-M(\pi)$ and thus $A^T(M-M(\pi))A$ are positive semi-definite for every $\pi\in\Gamma$ by $M(\pi)\leqslant M$. Hence

$$\mathrm{tr}(B(\tilde{m})M(\pi))\leqslant \mathrm{tr}(B(\tilde{m})M),\ \pi\in\Gamma. \tag{9}$$

Now (2), (6), (7), (8) and (9) yield

$$\begin{aligned}\sup_{\pi\in\Gamma} r(\pi,\delta_{\tilde{m}})&=\sup_{\pi\in\Gamma}[\mathrm{tr}(B(\tilde{m})M(\pi))-2\tilde{m}^TB(\tilde{m})m(\pi)+\tilde{m}^TB(\tilde{m})\tilde{m}+c(\tilde{m})]\\&\leqslant \mathrm{tr}(B(\tilde{m})M)-2(1-\tilde{m}^T(\Sigma+M)^{-1}\tilde{m})^{-2}\inf_{m\in V_M} d(\tilde{m})^Tm\\&\quad+\tilde{m}^TB(\tilde{m})\tilde{m}+c(\tilde{m})\\&=\mathrm{tr}(B(\tilde{m})M)-\tilde{m}^TB(\tilde{m})\tilde{m}+c(\tilde{m})=r(\pi_{\tilde{m}},\delta_{\tilde{m}}).\end{aligned}$$

Therefore $(\pi_{\tilde{m}},\delta_{\tilde{m}})\in\Gamma\times\Delta$ is a saddle point in the statistical game (Γ,Δ,r) which proves the theorem by a well known result in game theory, whereby the Γ-minimax estimator is uniquely determined since $\delta_{\tilde{m}}$ is the unique Bayes estimator with respect to $\pi_{\tilde{m}}$.

Theorem 2. *There exists a unequely determined* $\tilde{m}\in V_M$ *which fulfils condition* (7) *stated in Theorem* 1, *i. e.*

$$d(\tilde{m})^T\tilde{m}=\inf\{d(\tilde{m})^Tm\,|\,m\in V_M\}.$$

Proof It is obvious that the function $d: V_M\to\mathbf{R}^p$ defined as in Theorem 1 is continuous. The set $V_M\neq\varnothing$ is compact and convex. Therefore a result of Browder and Karamardian (see [2], Lemma 3.3.1), which is a special case of a minimax inequality due to Ky Fan([4], Corollary 1), yields the existence of a parameter $\tilde{m}\in V_M$ which satisfies (7) stated in Theorem 1, i. e. $\tilde{m}$ is a so-called stationary point of V_M.

Now assume that $\tilde{m}$ and $\bar{m}$ are two stationary points of V_M. Then the Bayes estimators $\delta_{\tilde{m}}$ and $\delta_{\bar{m}}$ are both Γ-minimax estimators because of Theorem 1 and satisfy

$$\delta_{\tilde{m}}(\tilde{m})-\delta_{\bar{m}}(\tilde{m})=\Sigma\Big[(\Sigma+M)^{-1}+\frac{(\Sigma+M)^{-1}\bar{m}\bar{m}^T(\Sigma+M)^{-1}}{1-\bar{m}^T(\Sigma+M)^{-1}\bar{m}}\Big](\tilde{m}-\bar{m}). \tag{10}$$

The uniqueness of the Γ-minimax estimator (except a set of Lebesgue measure zero) and the linearity of the estimators $\delta_{\tilde{m}}$ and $\delta_{\bar{m}}$ yield $\delta_{\tilde{m}}(\tilde{m})=\delta_{\bar{m}}(\tilde{m})$. This and (10) show that $\tilde{m}=\bar{m}$.

In the following two corollaries the Γ-minimax estimator is explicitly determined by applying Theorem 1 for two special forms of the set $V\subset\mathbf{R}^p$ in the definition of the subset Γ of priors.

Corollary 1. *If $V=\{m\}$, $m\in\mathbf{R}^p$, $mm^T\leq M$, consists of exactly one point then $V_M=\{m\}$ and $\tilde{m}=m$ satisfies (7) stated in Theorem 1.*

In the univariate case with fixed second moment, i. e. $p=1$, $\Sigma=\sigma^2>0$, $\omega^2=M-m^2\geqslant0$, and

$$\tilde{\Gamma}=\left\{\pi\in\Pi\,\Big|\int\theta\pi(d\theta)=m,\ \int\theta^2\pi(d\theta)=M\right\},$$

the $\tilde{\Gamma}$-minimax estimator

$$\delta_m(x)=\frac{\omega^2x+\sigma^2m}{\omega^2+\sigma^2},\quad x\in\mathbf{R},$$

obtained by Corollary 1 is already known (see [1], Example 4.29, p. 216, and[3]),

Corollary 2. *If $0\in V$ then $0\in V_M$ and $\tilde{m}=0$ satisfies (7) sated in Theorem 1, i. e.*

$$\delta_0(x)=M(\Sigma+M)^{-1}x,\ x\in\mathbf{R}^p,$$

is the unique Γ-minimax estimator and

$$\pi_0=N(0,\ M)$$

is least favourable in Γ.

In view of Corollary 2 it is subsequently assumed that $0\notin V$. In particular Corollary 2 can be applied to $V=\mathbf{R}^p$, i. e. in the case

$$\Gamma=\{\pi\in\Pi\,|\,M(\pi)\leqslant M\},$$

where restrictions are imposed only on the matrix $M(\pi)$ of second moments for priors $\pi\in\Pi$.

§ 5. A Geometric Characterization of the Γ-Minimax Estimator

The following lemma shows that in the case $0\notin V$ the according to Theorem 2 uniquely determined parameter $\tilde{m}\in V_M$ which fulfils condition (7) stated in Theorem 1 is a boundary point of the set V_M. This makes it feasible to give in Theorem 3 a geometric characterization of the parameter $\tilde{m}\in V_M$.

Lemma 3. *If $0\notin V$ then the parameter $\tilde{m}\in V_M$ which fulfils condition (7) stated in Theorem 1 is an element of the boundary ∂V_M of V_M.*

Proof Assume that $\tilde{m}\neq0$ is an inner point of V_M and satisfies (7). Then $d(\tilde{m})\neq0$ because of (8), $\tilde{m}\neq0$, and $B(\tilde{m})$ being positive definite. Hence without loss of generality $\tilde{d}_i>0$ for some $1\leqslant i\leqslant p$ where $d(\tilde{m})=(\tilde{d}_i)_{1\leqslant i\leqslant p}$.

Since $\tilde{m}$ is an inner point of V_M there exists an $\varepsilon>0$ such that $\overline{m}=\tilde{m}-\varepsilon e_i\in V_M$ where e_i denotes the i-th unit vector in $\mathbf{R}^p$. Then

$$d(\tilde{m})^T\overline{m}=d(\tilde{m})^T\tilde{m}-\varepsilon\tilde{d}_i<d(\tilde{m})^T\tilde{m},$$

which contradicts (7).

Before the geometric characterization of the Γ-minimax estimator is given in Theorem 3 some further notation is necessary.

Since V_M is closed and convex there exists at least one outer unit normal vector $n(m)\in\mathbf{R}^p$ belonging to a supporting hyperplane of V_M for every boundary point $m\in\partial V_M$, i. e.

$$n(m)^Tn(m)=1$$

and

$$n(m)^Tv\leqslant n(m)^Tm$$

for all $v\in V_M$. Let

$$H(n(m))=\{v\in\mathbf{R}^p\,|\,v=m+h,\ n(m)^Th=0,\ h\in\mathbf{R}^p\},\ m\in\partial V_M,$$

be the supporting hyperplane corresponding to the normal vector $n(m)$ which contains the point $m\in\partial V_M$, but no inner point of V_M, i. e. $\{m\}\subset H(m)\cap V_M\subset\partial V_M$.

Theorem 3. *A papameter $\tilde{m}\in\partial V_M$ fulfils condition (7) stated in Theorem 1 if and only if there exists a $\tilde{\lambda}>0$ and an outer unit normal vector $n(\tilde{m})$ such that*

$$n(\tilde{m})+\tilde{\lambda}d(\tilde{m})=0,$$

i. e. the normal vector $n(\tilde{m})$ and the vector $d(\tilde{m})$ defined as in Theorem 1 are parallel but have different directions.

Proof (i) Let $\tilde{m}\in\partial V_M$ satisfy (7) stated in Theorem 1. Let

$$H_\mu=\{v\in\mathbf{R}^p\,|\,v=\mu d(\tilde{m})+h,\ d(\tilde{m})^Th=0,\ h\in\mathbf{R}^p\},\ \mu\in\mathbf{R},$$

denote the hyperplane which is orthogonal to $d(\tilde{m})$ and contains $\mu d(\tilde{m})$. Then the hyperplanes H_μ, $\mu\in\mathbf{R}$, form a partition of $\mathbf{R}^p$ and $d(\tilde{m})^Tv=\mu d(\tilde{m})^Td(\tilde{m})$ if and only if $v\in H_\mu$. Since V_M is compact and convex there exists a number $\mu_0\in\mathbf{R}$ with

$$\mu_0=\inf\{\mu\in\mathbf{R}\,|\,H_\mu\cap V_M\neq\varnothing\}. \tag{11}$$

Therefore

$$\inf_{m\in V_M} d(\tilde{m})^Tm=\inf\{\mu d(\tilde{m})^Td(\tilde{m})\,|\,\mu\in\mathbf{R},\ H_\mu\cap V_M\neq\varnothing\}=\mu_0d(\tilde{m})^Td(\tilde{m}).$$

This and $\tilde{m}\in\partial V_M$ satisfying (7) yield

$$d(\tilde{m})^T\tilde{m}=\mu_0d(\tilde{m})^Td(\tilde{m}),$$

and hence $\tilde{m}\in H_{\mu_0}$. Now (11) implies $H_{\mu_0}\cap V_M\subset\partial V_M$. Since $d(\tilde{m})$ is orthogonal to H_{μ_0} the vector $n_0=-\tilde{\lambda}d(\tilde{m})$ with $\tilde{\lambda}=(d(\tilde{m})^Td(\tilde{m}))^{-1/2}$ which satisfies

$$n_0^Tv\leqslant n_0^T\tilde{m}\quad\text{for all } v\in V_M$$

because of (7) is an outer unit normal vector and obviously

$$n_0+\tilde{\lambda}d(\tilde{m})=0.$$

(ii) Let $\tilde{\lambda}>0$ and $\tilde{m}\in\partial V_M$ such that

$$n(\tilde{m})+\tilde{\lambda}d(\tilde{m})=0,$$

where $n(\tilde{m})$ is an outer unit normal vector. Then

$$n(\tilde{m})^T v \leqslant n(\tilde{m})^T \tilde{m} \quad \text{for all } v \in V_M$$

implies

$$d(\tilde{m})^T v \geqslant d(\tilde{m})^T \tilde{m} \quad \text{for all } v \in V_M,$$

i. e. (7) is satisfied.

Of special interest is the case where $0 \notin V$, $V \subset E_M$, and the boundary of V is smooth such that a uniquely determined outer unit normal vector $n(m)$ exists for every boundary point $m \in \partial V$. Then Theorem 2, Lemma 3, and Theorem 3 show that the parameter $\tilde{m} \in \partial V$ which fulfils condition (7) stated in Theorem 1 satisfies

$$n(\tilde{m}) + \tilde{\lambda} d(\tilde{m}) = 0$$

for some $\tilde{\lambda} > 0$. Therefore $(\tilde{\lambda}, \tilde{m}) \in (0, \infty) \times \partial V$ is the uniquely determined solution of the system of p non-linear equations

$$n(m) + \lambda d(m) = 0,$$

where $(\lambda, m) \in (0, \infty) \times \partial V$.

In the case $0 \notin V$ and $V \not\subset M_M$ the following corollary shows with the add of the geometric characterization given in Theorem 3 that the parameter $\tilde{m} \in V_M$ which fulfils condition (7) stated in Theorem 1 and which is an element of ∂V_M according to Lemma 3 is in fact an element of the genuine subset $\partial V \cap E_M$ of ∂V_M.

Corollary 3. *If $0 \notin V$ then the parameter $\tilde{m} \in V_M$ which fulfils condition (7) stated in Theorem 1 is an element of the set $\partial V \cap E_M$, i. e. $\tilde{m} \notin \partial E_M \cap \mathring{V}$ where $\mathring{V}$ denotes the interior of V.*

Proof Lemma 3 and $0 \notin V$ yield $\tilde{m} \in \partial V_M$. Assume that $\tilde{m} \in \partial E_M \cap \mathring{V}$. Then a uniquely determined outer unit normal vector $n(\tilde{m})$ exists because of Lemma 1. By Theorem 3 there exists a $\tilde{\lambda} > 0$ such that

$$n(\tilde{m}) + \tilde{\lambda} d(\tilde{m}) = 0. \tag{12}$$

Now consider the subset

$$\Gamma^* = \{\pi \in \Pi \mid M(\pi) \leqslant M\}$$

of priors, i. e. $V_M^* = E_M$. Then (12) keeps valid and Theorem 3 and Theorem 2 show that $\tilde{m} \in V_M^*$ is the uniquely determined parameter which satisfies (7) stated in Theorem 1. This and $\tilde{m} \neq 0$ contradict Corollary 2.

§ 6. Special Quadratic Loss Functions

In this last section the matrix R in the definition (1) of the loss function, the co-variance matrix Σ of the normal distributions, and the matrix M in the definition of the subset Γ of priors satisfy a certain relation given in Theorem 4 below. Then the parameter $\tilde{m} \in V_M$ which fulfils condition (7) stated in Theorem 1 is the minimum of a quadratic form on the compact and convex set V_M. If the matrices R, Σ, and M satisfy a second relation the parameter $\tilde{m}$ is simply the vector of shortest length in V_M as it is shown in Corollary 4. In some examples at the end of this section the Γ-minimax estimator is explicitly determined.

First a technical lemma is proved where a condition is given which is equivalent to (7) stated in Theorem 1.

Lemma 4. *Let $A(m)$ be a symmetric and positive definite matrix such that*

$$d(m)=A(m)m \quad \textit{for every } m\in V_M,$$

where the vector $d(m)$ is defined as in Theorem 1. Then a parameter $\tilde{m}\in V_M$ satisfies condition (7) stated in Theorem 1 if and only if

$$\tilde{m}^T A(\tilde{m})\tilde{m}=\inf\{m^T A(\tilde{m})m \mid m\in V_M\}. \tag{13}$$

Proof If $0\in V$ the assertion follows at once since $A(m)$ is positive definite for every $m\in V_M$. Now consider the case $0\notin V$.

(i) Let $\tilde{m}\in V_M$ fulfil condition (7) stated in Theorem 1. Since $A(\tilde{m})$ is symmetric and positive definite there exists a $p\times p$-matrix $D(\tilde{m})$ such that $A(\tilde{m})=D(\tilde{m}^T)D(\tilde{m})$. Now (4), $\tilde{m}\neq 0$, and the Schwarz inequality yield

$$\tilde{m}^T A(\tilde{m})\tilde{m}\leqslant(\tilde{m}^T A(\tilde{m}))\tilde{m}^{-1}\cdot(\tilde{m}^T D(\tilde{m})^T D(\tilde{m})m)^2\leqslant m^T A(\tilde{m})m$$

for every $m\in V_M$, i. e. (13) is valid.

(ii) Let $\tilde{m}\in V_M$ fulfil (13) and assume that condition (7) stated in Theorem 1 is not satisfied. Then there exists a parameter $\overline{m}\in V_M$ such that

$$\tilde{m}^T A(\tilde{m})\tilde{m}>\tilde{m}^T A(\tilde{m})\overline{m}. \tag{14}$$

Since V_M is convex $m_\alpha=\alpha\overline{m}+(1-\alpha)\tilde{m}\in V_M$ for every $\alpha\in[0, 1]$. A short calculation yields

$$m_\alpha^T A(\tilde{m})m_\alpha=\tilde{m}^T A(\tilde{m})\tilde{m}+\alpha[2\tilde{m}^T A(\tilde{m})(\overline{m}-\tilde{m})+\alpha(\overline{m}-\tilde{m})^T A(\tilde{m})(\overline{m}-\tilde{m})].$$

Therefore (14) shows that there exists an $\alpha^*\in(0, 1)$ such that

$$m_\alpha^T A(\tilde{m})m_\alpha<\tilde{m}^T A(\tilde{m})\tilde{m} \quad \text{for all } \alpha\in(0, \alpha^*),$$

which contradicts (13).

Note that by (8) in the proof of Theorem 1 the maxtrix

$$A(m)=(1-m^T(\Sigma+M)^{-1}m)^2 B(m),\ m\in V_M,$$

satisfies the hypothesis $d(m)=A(m)m$, $m\in V_M$, in Lemma 4.

Theorem 4. *Let the matrix R in the definition (1) of the loss function be given by*

$$R=\nu\Sigma^{-1}(\Sigma+M)\Sigma^{-1} \quad \textit{for some } \nu>0.$$

Then a parameter $\tilde{m}\in V_M$ satisfies condition (7) stated in Theorem 1 if and only if

$$\tilde{m}^T(\Sigma+M)^{-1}\tilde{m}=\inf[m^T(\Sigma+M)^{-1}m \mid m\in V_M\}. \tag{15}$$

Proof The vector $d(m)$ defined as in Theorem 1 satisfies

$$\begin{aligned}d(m)&=\nu[(\Sigma+M)^{-1}mm^T+(1-m^T(\Sigma+M)^{-1}m)I](\Sigma+M)^{-1}m\\&=\nu(\Sigma+M)^{-1}m,\quad m\in V_M.\end{aligned}$$

Therefore the symmetric and positive definite matrix $A(m)=\nu(\Sigma+M)^{-1}$, $m\in V_M$, fulfils the hypothesis of Lemma 4, which proves the theorem since (13) and (15) are obviously equivalent.

Since Σ and M are symmetric and positive definite there exists a non-singular matrix L with $L^T L=(\Sigma+M)^{-1}$. Then the set

$$L_M=\{w\in \mathbf{R}^p \mid w=Lm,\ m\in V_M\}$$

is compact and convex, and the condition (15) in Theorem 4 is obviously equivalent to

$$\tilde{w}^T\tilde{w}=\inf\{w^T w \mid w\in L_M\} \tag{16}$$

whereby $\tilde{w}=L\tilde{m}\in L_M$ is the vector of shortest length in L_M.

Corollary 4. *Let the matrices R, Σ, and M satisfy the relations*

$$R=\nu_1\Sigma^{-2} \text{ and } \Sigma+M=\nu_2 I \text{ for some } \nu_1,\ \nu_2>0.$$

Then a parameter $\tilde{m}\in V_M$ satisfies condition (7) stated in Theorem 1 if and only if

$$\tilde{m}^T\tilde{m}=\inf\{m^T m \mid m\in V_M\}, \tag{17}$$

i. e. $\tilde{m}$ is the vector of shortest length in V_M.

Note that in particular the hypothesis of Corollary 4 is satisfied if the matrices R, Σ, and M are multiples of the identity matrix I.

In the following first three examples for different subsets Γ of priors the Γ-minimax estimator is explicitly found by applying Theorem 4 and Corollary 4. In all these examples the least favourable prior $\pi_{\tilde{m}}$ is always a non-singular normal distribution, i. e. the mean vector $\tilde{m}$ is always an inner point of E_M. The fourth example shows that this is generally not valid. Although the subset Γ of priors contains non-singular normal distributions, i. e. V_M contains inner points of E_M, a singular normal distribution is least favourable in Γ.

Example 1. Assume that $\Sigma=\sigma I$, $R=\rho I$, and $M=\mu I$ for some σ, ρ, $\mu>0$. Let

$$V=\{m\in \mathbf{R}^p \mid (m-c)^T(m-c)\leqslant r^2\}$$

be a p-dimensional sphere with centre $c\in \mathbf{R}^p$ and radius $r>0$, where

$$r<\sqrt{c^Tc}<r+\sqrt{\mu},$$

such that $0\notin V$ and such that $V_M=V\cap E_M$ contains more than one point. A short calculation shows that

$$\tilde{m}=\left(1-\frac{r}{\sqrt{c^Tc}}\right)c\in \partial V\cap E_M$$

is the vector of shortest length in V_M, i. e. condition (17) stated in Corollary 4 is satisfied. Therefore Corollary 4, Theorem 1, and (5) show that

$$\delta_{\tilde{m}}(x)=\frac{1}{\sigma+\mu}\left(\mu I-\frac{\sigma(\sqrt{c^Tc}-r)^2}{(\sigma+\mu-(\sqrt{c^Tc}-r)^2)c^Tc}cc^T\right)x$$
$$+\frac{\sigma(\sqrt{c^Tc}-r)}{(\sigma+\mu-(\sqrt{c^Tc}-r)^2)\sqrt{c^Tc}}c,\ x\in\mathbf{R}^p,$$

is the unique Γ-minimax estimator (except a set of Lebesgue measure zero) and that the prior

$$w_{\tilde{m}}=N\left(\left(1-\frac{r}{\sqrt{c^Tc}}\right)c,\ \mu I-\frac{(\sqrt{c^Tc}-r)^2}{c^Tc}cc^T\right)$$

is least favourable in Γ.

Example 2. Assume that $\Sigma=\sigma I$, $R=\rho I$, and $M=\mu I$ for some $\sigma,\ \rho,\ \mu>0$. Let

$$V=\{m\in\mathbf{R}^p \mid m^Tc\geqslant 1\}$$

be a semi-space of $\mathbf{R}^p$, where $c\in\mathbf{R}^p$ satisfies $c^Tc>1/\mu$, such that $0\notin V$ and such that $V_M=V\cap E_M$ contains more than one point. A short calculation shows that

$$\tilde{m}=\frac{1}{c^Tc}c\in\partial V\cap E_M$$

is the vector of shortest length in V_M, i. e. condition (17) stated in Corollary 4 is satisfied. Therefore Corollary 4, Theorem 1, and (5) show that

$$\delta_{\tilde{m}}(x)=\frac{1}{\sigma+\mu}\left(\mu I-\frac{\sigma}{((\sigma+\mu)c^Tc-1)c^Tc}cc^T\right)x$$
$$+\frac{\sigma}{(\sigma+\mu)c^Tc-1}c,\ x\in\mathbf{R}^p,$$

is the unique Γ-minimax estimator (except a set of Lebesgue measure zero) and that the prior

$$\pi_{\tilde{m}}=N\left(\frac{1}{c^Tc}c,\ \mu I-\frac{1}{(c^Tc)^2}cc^T\right)$$

is least favourable in Γ.

Exumple 3. Assume that $\Sigma=\mathrm{diag}(\sigma_1,\ \cdots,\ \sigma_p)$, $M=\mathrm{diag}(\mu_1,\ \cdots\mu_p)$, and

$$R=\lambda\,\mathrm{diag}\left(\frac{\sigma_1+\mu_1}{\sigma_1^2},\ \cdots,\ \frac{\sigma_p+\mu_p}{\sigma_p^2}\right)$$

are diagonal-matrices for some $\lambda,\ \sigma_1,\ \cdots,\ \sigma_p,\ \mu_1,\ \cdots,\ \mu_p>0$. Let

$$V=\{m\in\mathbf{R}^p \mid \alpha_i\leqslant m_i\leqslant\beta_i,\ 1\leqslant i\leqslant p\}$$

be a p-dimenssonal cube, where $\alpha,\ \beta\in\mathbf{R}^p$, $\alpha_i\leqslant\beta_i$, $1\leqslant i\leqslant p$. Define $\tilde{m}\in\mathbf{R}^p$ by

$$\tilde{m}_i=\begin{cases}\alpha_i & \text{for } \alpha_i>0,\\ 0 & \text{for } \alpha_i\leqslant 0\leqslant\beta_i,\ 1\leqslant i\leqslant p,\\ \beta_i & \text{for } \beta_i<0.\end{cases}$$

Assume that $V_M=V\cap E_M$ contains more than one point, which is obviously equivalent to

$$\tilde{m}^TM^{-1}\tilde{m}=\sum_{i=1}^{p}\frac{\tilde{m}_i^2}{\mu_i}<1$$

because of Lemma 1. The non-singular matrix

$$L=\mathrm{diag}\left(\frac{1}{\sqrt{\sigma_1+\mu_1}},\ \cdots,\ \frac{1}{\sqrt{\sigma_p+\mu_p}}\right)$$

satisfies $L^TL=(\Sigma+M)^{-1}$ and the set L_M as defined after Theorem 4 is given by

$$L_M=\{w\in\mathbf{R}^p \mid w=Lm,\ m\in V_M\}$$
$$=\left\{w\in\mathbf{R}^p \mid \frac{\alpha_i}{\sqrt{\sigma_i+\mu_i}}\leqslant w_i\leqslant\frac{\beta_i}{\sqrt{\sigma_i+\mu_i}},\ 1\leqslant i\leqslant p,\ \sum_{i=1}^{n}\frac{\sigma_i+\mu_i}{\mu_i}w_i^2\leqslant 1\right\}.$$

Hence $\tilde{w}=L\tilde{m}\in L_M$ fulfils condition (16) and therefore $\tilde{m}\in V_M$ satisfies condition

(15) stated in Theorem 4. Thus Theorem 4 and Theorem 1 show that $\delta_{\tilde{m}}\in\Delta$ as defined in (5) is the unique Γ-minimax estimator (except a set of Lebesgue measure zero) and that the prior

$$\pi_{\tilde{m}}=N(\tilde{m},\ M-\tilde{m}\tilde{m}^T)$$

is least favourable in Γ.

Example 4 Assume that $p=2$, $\Sigma=\begin{bmatrix}1&0\\0&4\end{bmatrix}$, $R=\begin{bmatrix}1&0\\0&1/16\end{bmatrix}$, and $M=\begin{bmatrix}4&0\\0&1\end{bmatrix}$. Then

$$E_M=\left\{(m_1,\ m_2)\in\mathbf{R}^2\,\middle|\,\frac{1}{4}m_1^2+m_2^2\leqslant 1\right\}$$

is an ellipse with semi-axes (2, 0) and (0, 1). Let

$$V=\{(m_1,\ m_2)\in\mathbf{R}^2\,|\,m_1+2m_2\geqslant c\}$$

be a semi-plane whereby $10/\sqrt{17}\leqslant c<2\sqrt{2}$. The condition $c<2\sqrt{2}$ ensures that $V_M=V\cap E_M$ contains inner points of E_M. Then

$$V_M=\left\{(m_1,\ m_2)\in\mathbf{R}^2\,\middle|\,\left|\frac{c}{2}-m_1\right|\leqslant\sqrt{2-\frac{1}{4}c^2},\ \frac{1}{2}(c-m_1)\leqslant m_2\leqslant\sqrt{1-\frac{1}{4}m_1^2}\right\}.$$

Therefore

$$\tilde{m}=\left(\frac{c}{2}-\sqrt{2-\frac{1}{4}c^2},\ \frac{c}{4}+\frac{1}{2}\sqrt{2-\frac{1}{4}c^2}\right)^T\in\partial V\cap\partial E_M.$$

The condition $c\geqslant 10/\sqrt{17}$ ensures that $\tilde{m}$ is the vector of shortest length in V_M, i. e. condition (17) stated in Corollary 4 is satisfied. Therefore Corollary 4 and Theorem 1 show that the prior

$$\pi_{\tilde{m}}=N(\tilde{m},\ M-\tilde{m}\tilde{m}^T)$$

is least favourable in Γ, whereby the matrix

$$M-\tilde{m}\tilde{m}^T=\begin{pmatrix}2+c\sqrt{2-\frac{1}{4}c^2} & 1-\frac{1}{4}c^2\\ 1-\frac{1}{4}c^2 & \frac{1}{4}\left(2-c\sqrt{2-\frac{1}{4}c^2}\right)\end{pmatrix}$$

is singular.

The author would like to express his gratitude to Doctor Eichenauer, Professor Kindler, Pro fessor Lehn, and Professor Wegmann for their valuable hints.

References

[1] Berger, J. O., Statistical decision theory, 2d ed., Berlin-Heidelberg-New York, Springer, 1985.

[2] Ichiishi, T., Game theory for economic analysis, New York-London, Academic Press, 1983.

[3] Jackson, D. A., O'Donovan, T. M., Zimmer, W. J. & Deely, J. J., G_2-minimax estimators in the exponential family, *Biometrika*, **57** 1970, 439—443.
[4] Ky Fan, A minimax inequality and applications, in O. Shisha (Ed.) Inequalities III, New York-London, Academic Press, 1972, 103—113.
[5] Ortega, J. M. & Rheinboldt, W. C., Iterative solution of nonlinear equations in several variables, Orlando, Academic Press, 1970.
[6] Soloman, D. L., Λ-minimax estimation of a multivariate location paramenter, *J. Amer. Statist. Assoc.*, **67** 1972, 641—646.
[7] Srivastava, M. S. & Khatri, C. G., An introduction to multivariate statistics, New York, North Holland, 1979.