ADMISSIBLE LINEAR ESTIMATORS IN A LINEAR MODEL WITH THE NATURAL PARAMETER SPACE

Stepniak, Czeslaw
Department of Applied Mathematics, Agricultural University of Lublin

https://doi.org/10.5109/13378

出版情報：Bulletin of informatics and cybernetics. 22 (1/2), pp.71-77, 1986-03. Research Association of Statistical Sciences

Bulletin of Informatics and Cybernetics Vol. 22, No. 1~2, 1986

ADMISSIBLE LINEAR ESTIMATORS IN A LINEAR MODEL WITH THE NATURAL PARAMETER SPACE

By

Czesław STĘPNIAK*

Abstract

It is known that the linear estimation, both with or without unbiasedness, may be reduced to a statistical game with a convex compact parameter set. Then all locally best estimators constitute a complete class and each locally best estimator being unique is admissible. However, if the considered locally best estimator is not unique, then all known sufficient conditions for the admissibility work very hard. We derive a simpler sufficient condition for the admissibility in a linear model with the natural parameter space.

1. Introduction and Summary

The admissibility is a natural way of the selection of statistical rules. For some relevant results in the context of linear estimation see Cohen [1], LaMotte [5], Olsen, Seely and Birkes [6], Rao [7] and Stępniak [8]. The papers [1], [7] and [8] refer to the linear model with the covariance matrix of the form γV, where γ is an unknown scalar; in [5] and [6] the set of the possible covariance matrices may be arbitrary subset of non-negative definite symmetric matrices.

Olsen, Seely and Birkes [6] have shown that any admissible linear unbiased estimator is locally best for some element in the closed convex cone generated by the covariance matrices. A similar result for the linear estimation without unbiasedness was obtained by LaMotte [5]. Conversely, a locally best estimator is admissible provided is unique, but it may be inadmissible in general. LaMotte has presented a procedure by which we can verify, in a finite number of steps, whether a linear estimator is admissible or not.

We restrict our considerations to the linear model with the natural parameter space, i.e. to the model with the covariance matrix of the form $\sum\gamma_i V_i$, where each γ_i is running an open or closed interval or ray with the beginning in zero. In Section 2 the problem of linear unbiased estimation in such a model is reduced to a statistical game with a finite parameter set. In a similar way we may reduce the problem of linear estimation without unbiasedness in the case when the space of possible expectations is one-dimensional, i.e. when the expectation depends on a scalar parameter. This reduction makes it possible for us to derive some sufficient conditions for linear

* Department of Applied Mathematics, Agricultural University of Lublin, Akademicka 13, 20-934 Lublin, Poland.

admissibility by some elementary results in the statistical decision theory.

Our spadework in statistical decision theory is collected in Section 3 and its applications in linear estimation are given in Section 4.

Throughout this paper the usual matrix notation will be used. Among others, if A is a matrix then A' and $R(A)$ will denote, respectively, the transposition and the range (column space) of A. Moreover the symbol R^n stands for the n-dimensional Euclidean space represented by n-dimensional column vectors.

2. The Initial Reduction

The term "linear model" refers usually to the second order model of a random vector X in an Euclidean space R^n. Such a model is defined by the expectation vector EX and by the covariance matrix Cov X. Throughout this paper

$$EX=A\beta$$

and

$$\text{Cov } X=\sum_{i=1}^{q}\gamma_i V_i,$$

where A is a given matrix of $n\times p$, V_i, $i=1, \cdots, q$, are given non-negative definite symmetric matrices of $n\times n$; β and γ_i's are unknown parameters. We shall assume that the linear span of the possible values of β is R^p and each γ_i, $i=1, \cdots, q$, is running an open or closed interval or ray with the beginning in zero. Instead of $\sum\gamma_i V_i$ we shall also write V_γ, where γ is the short of $(\gamma_1, \cdots, \gamma_q)'$. The set of all possible values of γ will be denoted by Γ.

Consider a parameter $\Phi=\Phi(\beta, \gamma)$. Under the above assumptions this parameter possesses an unbiased estimator of the form $\hat{\Phi}=h'X$, $h\in R^n$, if and only if $\Phi=k'\beta$ for some $k\in R(A')$. Denote by $\mathfrak{M}$ the set of all $h\in R^n$ such that $E(h'X)=k'\beta$. The set may be written in the form $\mathfrak{M}=\{h\in R^n : A'h=k\}$.

The estimators $h'X$, $h\in\mathfrak{M}$, are compared, as in Olsen, Seely and Birkes [6], according to their possible variances. For $h_1, h_2\in\mathfrak{M}$, h_1 is said to be as good as h_2 if $h_1'V_\gamma h_1\le h_2'V_\gamma h_2$ for all $\gamma\in\Gamma$; h_1 is better than h_2 if h_1 is as good as h_2 and h_2 is not as good as h_1. An $h\in\mathfrak{M}$ is said to be admissible if no vector in $\mathfrak{M}$ is better than h.

It follows from our assumptions about Γ that the minimal closed convex cone generated by $\{V_\gamma : \gamma\in\Gamma\}$ is identical with one generated by the set $\{V_1, \cdots, V_q\}$. Thus, by some arguments in [6] the problem of linear unbiased estimation for the parameter Φ is reduced to the statistical game (Θ, D, R) with the parameter set $\Theta=\{\theta_1, \cdots, \theta_q\}$, the decision set $D=\{h\in R^n : A'h=k\}$ and the risk function $R(\theta_i, d)=d'V_i d$, $i=1, \cdots, q$, $d\in D$.

3. Bayes and Admissible Rules

Consider a statistical game (Θ, D, R), where Θ is the parameter set, D is the set of the decision rules and $R=R(\theta, d)$ is the risk of a rule $d\in D$ under a parameter $\theta\in\Theta$.

Let $\mathfrak{T}$ be the minimal σ-field in Θ such that all one-element subsets of Θ are members of $\mathfrak{T}$ and the function $R(\cdot, d)$ is $\mathfrak{T}$-measurable for all $d \in D$. Consider the set of all distributions τ on $(\Theta, \mathfrak{T})$ such that the value

$$r(\tau, d)=\int_{\Theta} R(\theta, d) d\tau(\theta) \tag{1}$$

exists for all $d \in D$. The value (1) is called the Bayes risk of the rule d with respect to a prior distribution τ on Θ.

A rule $d_0 \in D$ is called τ-Bayes if $r(\tau, d_0)=\inf_{d \in D} r(\tau, d)$ and Bayes if is τ-Bayes for some prior distribution τ on Θ.

In this Section we shall show some limit properties of the set of Bayes and admissible rules and give some sufficient conditions for the admissibility.

THEOREM 1. *Let (Θ, D, R) be a statistical game with a finite set Θ and let $d_0, d_1, d_2, \cdots$ be statistical rules such that d_n is τ_n-Bayes for some prior distribution τ_n, $n=1, 2, \cdots$, and d_0 satisfies the condition*

$$\lim_{n \to \infty} [r(\tau_n, d_n)-r(\tau_n, d_0)]=0. \tag{2}$$

Then d_0 is Bayes. Moreover, if τ_n weak converges to some prior τ_0 then d_0 is τ_0-Bayes.

PROOF. Let d_n be τ_n-Bayes rule, $n=1, 2, \cdots$. Because the set of all distributions on a finite set is compact, we can choose a subsequence of $\{\tau_n\}$ being weak convergent to some prior τ_0 on Θ. Without loss of generality assume that $\tau_n \Rightarrow \tau_0$. We are ready to show that any d_0 satisfying the condition (2) is τ_0-Bayes.

Suppose not. Then there is a rule $d \in D$ such that

$$r(\tau_0, d)<r(\tau_0, d_0)-\varepsilon$$

for some $\varepsilon>0$. On the other hand, by the weak convergence of τ_n and by (2), there exists an integer n such that

$$|r(\tau_n, d_0)-r(\tau_0, d_0)|<\frac{\varepsilon}{4},$$

$$|r(\tau_n, d)-r(\tau_0, d)|<\frac{\varepsilon}{4}$$

and

$$|r(\tau_n, d_n)-r(\tau_n, d_0)|<\frac{\varepsilon}{4}.$$

Thus

$$\begin{aligned} r(\tau_n, d_n)&=[r(\tau_n, d_n)-r(\tau_n, d_0)]+[r(\tau_n, d_0)-r(\tau_0, d_0)]+r(\tau_0, d_0) \\ &>r(\tau_0, d)+\frac{\varepsilon}{2}>r(\tau_n, d). \end{aligned}$$

This contradicts the condition that d_n is τ_n-Bayes and completes the proof. □

A consequence of this Theorem is

COROLLARY 1. *Let (Θ, D, R) be a statistical game, where $\Theta=\{\theta_1, \cdots, \theta_q\}$, D is a closed subset of an Euclidean space R^n and $R(\theta_i, \cdot)$ is a continuous function of d, $i=1, \cdots, q$. Then the class of all Bayes rules in the game is closed in R^n.*

It would be nice to have a similar result for the admissible rules. Unfortunately this is impossible without some additional assumptions.

THEOREM 2. *Let* (Θ, D, R) *be a statistical game, where* $\Theta=\{\theta_1, \theta_2\}$, D *is a closed subset of* R^n *and* $R(\theta_i, d)$ *is a continuous function of* $d, i=1, 2$. *Then the set of the admissible rules in the game is closed in* R^n.

PROOF. Let d_n be admissible, $n=1, 2, \cdots$, and let

$$\lim_{n\to\infty} d_n=d_0. \tag{3}$$

By Ferguson ([4], Th. 2.10.1) there exists a sequence of prior distributions τ_n on Θ such that d_n is τ_n-Bayes, $n=1, 2, \cdots$. It follows from Corollary 1 that the rule d_0 is τ_0-Bayes for some prior distribution τ_0.

Suppose d_0 is inadmissible. Then, by Ferguson ([4], Th. 2.3.2) $\tau_0(\theta_i)=0$ for some $i\in\{1, 2\}$, say for $i=1$. Thus there exists a rule d such that $R(\theta_2, d)=R(\theta_2, d_0)$ and

$$R(\theta_1, d)=R(\theta_1, d_0)-\varepsilon \tag{4}$$

for some $\varepsilon>0$. As the rule d is τ_0-Bayes and $\tau_0(\theta_2)=1$, we get

$$R(\theta_2, d_n)\geq R(\theta_2, d), \qquad n=1, 2, \cdots. \tag{5}$$

On the other hand, by (3), there exists an integer n such that $R(\theta_1, d_n)>R(\theta_1, d_0)-\frac{\varepsilon}{2}$. Thus, by (4), $R(\theta_1, d_n)>R(\theta_1, d)$ and, via (5), $r(\tau, d_n)>r(\tau, d)$ for any prior distribution τ. In particular $r(\tau_n, d_n)>r(\tau_n, d)$. This contradicts the condition that d_n is τ_n-Bayes, completing the proof. □

REMARK 1. The assumption that Θ containes not more than 2 elements is essential to this Theorem as shown by the example. Let $\Theta=\{\theta_1, \theta_2, \theta_3\}$, D be the convex hull of the sets $S_1=\{x=(x_1, x_2, x_3): x_1^2+x_2^2\leq 1, x_3=1\}$ and $S_2=\{(-1, 0, 0)\}$ and let $R(\theta_i, d)=d^{(i)}$, $i=1, 2, 3$, for any $d=(d^{(1)}, d^{(2)}, d^{(3)})\in D$. Then the rule $d_n=\left(-\sqrt{\frac{n}{n+1}}, -\sqrt{\frac{1}{n+1}}, 1\right)$, $n=1, 2, \cdots$, is admissible but $\lim_{n\to\infty} d_n=(-1, 0, 1)$ is inadmissible because is dominated by $d=(-1, 0, 0)$.

If we known that a rule d_0 is τ-Bayes for some prior τ then the class of all candidates for domination of d_0 may be reduced by the following

LEMMA 1. *Let* (Θ, D, R) *be a statistical game with a finite set* Θ *and let* τ *be a prior distribution on* Θ *with the support* $\Theta_0\subseteq\Theta$. *Then a* τ-*Bayes rule* d_0 *is admissible if and only if it is admissible in the class of the rules satisfying the condition* $R(\theta, d)=R(\theta, d_0)$ *for all* $\theta\in\Theta_0$.

The proof of this Lemma is similar to the proof of Theorem 2.3.2 in Ferguson [4] and is omitted.

Now for given but arbitrary prior distributions τ and $\bar{\tau}$ such that $\operatorname{supp}(\bar{\tau})=\Theta$ and for a given sequence $\{c_n\}$ of scalars such that $0<c_n\leq 1$, $n=1, 2, \cdots$, and $\lim_{n\to\infty} c_n=0$ we define a sequence $\{\tau_n\}$ of prior distributions on Θ by

$$\tau_n=\frac{\tau+c_n\bar{\tau}}{1+c_n}, \qquad n=1, 2, \cdots. \tag{6}$$

THEOREM 3. *Let* (Θ, D, R) *be a statistical game, where* $\Theta=\{\theta_1, \cdots, \theta_q\}$, D *is a subset of* R^n *and* $R(\theta_i, \cdot)$ *is a continuous function of* d, $i=1, \cdots, q$. *Moreover let* d_n, $n=1, 2, \cdots$, *be* τ_n-Bayes, where τ_n *is defined by* (6). *Then* $\lim_{n\to\infty} d_n$, *if exists, is admissible.*

REMARK 2. The priors τ_n in this Theorem may be replaced by generalized priors $\pi_n=\bar{\tau}+c_n^{-1}\tau$, $n=1, 2, \cdots$.

REMARK 3. Perhaps Theorem 3 may be derived from the original works by Farrell [2, 3]. However we do not see how to do it in a simple way.

PROOF. Let

$$\lim_{n\to\infty} d_n=d_0. \tag{7}$$

Then by Theorem 1 the rule d_0 is τ-Bayes.

Suppose d_0 is inadmissible. Then there exists a τ-Bayes rule d such that $R(\theta_i, d) \leqq R(\theta_i, d_0)$ with the strict inequality for some $i=i_0$. Thus $R(\theta_{i_0}, d)<R(\theta_{i_0}, d_0)-\frac{\varepsilon}{\bar{\tau}(\theta_{i_0})}$ for some $\varepsilon>0$. This implies

$$r(\bar{\tau}, d)<r(\bar{\tau}, d_0)-\varepsilon. \tag{8}$$

On the other hand, by (7) and by continuity of R with respect to d we get $r(\bar{\tau}, d_n)> r(\bar{\tau}, d_0)-\varepsilon$ for some integer n. This, by (8), implies $r(\bar{\tau}, d_n)>r(\bar{\tau}, d)$. Therefore, as d is τ-Bayes, we get

$$\begin{aligned} r(\tau_n, d_n)&=\frac{1}{1+c_n}r(\tau, d_n)=\frac{c_n}{1+c_n}r(\bar{\tau}, d_n) \\ &>\frac{1}{1+c_n}r(\tau, d)+\frac{c_n}{1+c_n}r(\bar{\tau}, d) \\ &=r(\tau_n, d). \end{aligned}$$

This contradicts the assumption that d_n is τ_n-Bayes and completes the proof. □

4. Applications in Linear Estimation

Return to the problem of linear unbiased estimation of a parameter $\Phi=k'\beta$ in a linear model $EX=A\beta$ and $\operatorname{Cov} X=\sum_{i=1}^{q}\gamma_i V_i$, considered in the Section 2. This problem was reduced to the statistical game (Θ, D, R), where $\Theta=\{\theta_1, \cdots, \theta_q\}$, $D=\{d\in R^n: A'd=k\}$ and $R(\theta_i, d)=d'V_i d$, $i=1, \cdots, q$. A direct consequence of Theorems 2 and 3 is

THEOREM 4. *For the problem of linear unbiased estimation of a parameter* $\Phi=k'\beta$ *in the linear model* $EX=A\beta$ *and* $\operatorname{Cov} X=\sum_{i=1}^{q}\gamma_i V_i$ *with the natural parameter set*

(a) *Any unbiased estimator* $d_0=d_0'X$ *minimizing* $\sum_{i=1}^{q}\gamma_i d'V_i d$ *for some* $\gamma_i>0$, $i=1, \cdots, q$, *is admissible.*

(b) *Let* $\gamma^{(n)}=c_n\bar{\theta}+(1-c_n)\theta$, $n=1, 2, \cdots$, *for some* $\bar{\theta}=(\bar{\theta}_1, \cdots, \bar{\theta}_q)'$ *and* $\theta=(\theta_1, \cdots, \theta_q)'$ *satisfying the conditions* $\bar{\theta}_i>0$ *and* $\theta_i\geq 0$, $i=1, \cdots, q$, *and for some positive* $c_1, c_2, \cdots$ *going to zero. Moreover let* d_n, $n=1, 2, \cdots$, *minimize* $d'V_{\gamma(n)}d$ *over* $d\in D$. *Then* $\lim_{n\to\infty} d_n$, *if exists, is admissible.*

(*c*) *If* $q=2$ *and* d_n, $n=1, 2, \cdots$, *is admissible then* $\lim_{n\to\infty} d_n$, *if exists, is admissible.*

A similar result may be obtained for the problem of linear estimation without unbiasedness in a linear model with the natural parameter space and the expectation $EX=1_n\mu$, where 1_n is the column of n ones and μ is an unknown scalar.

EXAMPLE. Linear unbiased estimation in the unbalanced 1-way random linear model.

Suppose n experimental units are submitted to 1-way classification with k subclasses, where the number of units in the i-th subclass is n_i, $i=1, \cdots, k$. Let X_{ij} be the observation corresponding to the j-th experimental unit in the i-th subclass. Then we may write

$$X_{ij}=\mu+a_i+e_{ij}, \tag{9}$$

where μ is the general mean, a_i, $i=1, \cdots, k$, is the effect of the i-th subclass and e_{ij}, $i=1, \cdots, k$, $j=1, \cdots, n_i$, is the effect of error. Assuming all these effects are independent random variables with the expectations zero and with the variances $\mathrm{Var}\,(a_i)=\gamma_1$ and $\mathrm{Var}\,(e_{ij})=\gamma_2$, $i=1, \cdots, k$, $j=1, \cdots, n_i$, we reach the model

$$\begin{aligned} &EX=\mu 1_n \\ &\mathrm{Cov}\, X=\gamma_1 I_n+\gamma_2\, \mathrm{diag}\,(1_{n_1}1'_{n_1}, \cdots, 1_{n_k}1'_{n_k}), \end{aligned} \tag{10}$$

where $X=(X_{11}, \cdots, X_{1n_1}; \cdots; X_{k1}, \cdots, X_{kn_k})'$, $\sum_{i=1}^k n_i=n$, $\mu\in R$, $\gamma_1>0$ and $\gamma_2\geq 0$. The problem of linear unbiased estimation in such a model reduces to estimation of the parameter μ.

For given $\rho\geq 0$ let $d'X$ be a locally best linear unbiased estimator of μ at the point $\gamma_1=1$ and $\gamma_2=\rho$. It is well known that such estimator is unique and one is determined by the conditions

$$d'1_n=1 \quad \text{and} \quad V_\rho d=c1_n \qquad \text{for some } c\in R, \tag{11}$$

where $V_\rho=I_n+\rho\, \mathrm{diag}\,(1_{n_1}1'_{n_1}, \cdots, 1_{n_k}1'_{n_k})$. Denote this solution of (11) by d_ρ. By Theorem 4(*b*) any d_ρ, $\rho\geq 0$, is admissible.

Define also a set

$$D_0=\{d\in R^n : d'1_n=1 \quad \text{and} \quad \mathrm{diag}\,(1_{n_1}1'_{n_1}, \cdots, 1_{n_k}1'_{n_k})d=c1_n,\ c\in R\}.$$

Then $D_0\cup\{d_\rho : \rho\geq 0\}$ is the class of all Bayes rules in our problem. It is well known (cf. Olsen, Seely and Birkes [6]) that this class is complete but, perhaps, not minimal complete. We shall show that the sequence $\{d_m, m\geq 1\}$, is convergent. This implies that its limit is admissible.

An explicit solution of (11) is

$$d_\rho=\frac{1}{\sum_{i=1}^k \frac{n_i}{1+n_i\rho}}\left(1_n-\rho\left[\frac{n_1}{1+n_1\rho}1'_{n_1}, \cdots, \frac{n_k}{1+n_k\rho}1'_{n_k}\right]'\right).$$

Let $d_\rho^{(1)}$ be the subvector of d_ρ consisting of the first n_1 components. It can be shown that $d_\rho^{(1)}=\left[\sum_{i=1}^{k}\frac{n_i(1+n_1\rho)}{1+n_i\rho}\right]^{-1}1_{n_1}$. Thus $\lim_{m\to\infty} d_m^{(1)}=\frac{1}{kn_1}1_{n_1}$. Applying this procedure to the other subvectors of d_ρ we get

$$d_\infty=\lim_{m\to\infty} d_m=\frac{1}{k}\left(\frac{1}{n_1}1'_{n_1}, \cdots, \frac{1}{n_k}1'_{n_k}\right)'.$$

By Theorem 4(c) (or 4(b)), the rule d_∞ is admissible.

It can be shown that the rules d_ρ, $\rho\in[0, \infty]$, constitute the minimal complete class for linear unbiased estimation of the parameter μ in the model (10).

We note that the estimator $d_0'X$, corresponding to $\rho=0$, may be presented in the form

$$d_0'X=\frac{1}{n}\sum_{i=1}^{k} n_i\bar{X}_i,$$

where $\bar{X}_i=\frac{1}{n_i}\sum_{i=1}^{k} X_{ij}$ is the mean in the i-th subclass, while

$$d_\infty'X=\frac{1}{k}\sum_{i=1}^{k}\bar{X}_i.$$

Thus any d_ρ, $\rho\in(0, \infty)$, is a compromis between d_0 and d_∞.

Acknowledgment

I wish to thank the referee for his comments which led to an improvement of the presentation.

References

[1] COHEN, A.: *All admissible estimators of the mean vector*, Ann. Math. Statist., **37** (1966) 458-463.
[2] FARRELL, R. H.: *Towards a theory of generalized Bayes tests*, Ann. Math. Statist. **39** (1968) 1-22.
[3] FARRELL, R. H.: *On a necessary and sufficient condition for admissibility of estimators when strictly convex loss is used*, Ann. Math. Statist. **39** (1968) 23-28.
[4] FERGUSON, T. S.: *Mathematical Statistics: A Decision Theoretic Approach*, Academic *Press*, New York (1967).
[5] LAMOTTE, L. R.: *Admissibility in linear estimation*, Ann. Statist., **10** (1982) 245-255.
[6] OLSEN, A., SEELY, J. and BIRKES. D.: *Invariant quadratic estimation for two variance components*, Ann. Statist., **4** (1976) 878-890.
[7] RAO, C. R.: *Estimation of parameters in a linear model*, Ann. Statist., **4** (1976) 1023-1037.
[8] STEPNIAK, C.: *On admissible estimators in a linear model*, Biom. J. **26** (1984) 815-816.

Communicated by S. Kano
Received October 17, 1983
Revised February 14, 1985