Ann. Inst. Statist. Math.
Vol. 47, No. 1, 155–165 (1995)

MINIMAX TESTS FOR CONVEX CONES

LUTZ DÜMBGEN

Institut für Angewandte Mathematik, Universität Heidelberg, Im Neuenheimer Feld 294, D-69120 Heidelberg, Germany

(Received November 12, 1993; revised May 30, 1994)

Abstract. Let $(P_\theta : \theta \in \boldsymbol{R}^p)$ be a simple shift family of distributions on $\boldsymbol{R}^p$, and let $\boldsymbol{K} \subset \boldsymbol{R}^p$ be a convex cone. Within the class of nonrandomized tests of $\boldsymbol{K}$ versus $\boldsymbol{R}^p \backslash \boldsymbol{K}$, whose acceptance region A satisfies $A = A + \boldsymbol{K}$, a test with minimal bias is constructed. This minimax test is compared to a likelihood ratio type test, which is optimal with respect to a different criterion. The minimax test is mimicked in the context of linear regression and one-sided tests for covariance matrices.

Key words and phrases: Bias, convex cone, covariance matrix, duality, linear regression, minimax test, union-intersection principle.

1. Introduction

Let $(P_\theta : \theta \in \Theta)$ be a statistical experiment consisting of distributions P_θ on a measurable space $\boldsymbol{X}$ and an open subset Θ of $\boldsymbol{R}^p$. We consider hypotheses $\Theta \cap \boldsymbol{K}$ with some closed, convex cone $\boldsymbol{K}$ in $\boldsymbol{R}^p$. For example, one often wants to test whether the unknown parameter θ belongs to one of the following cones:

$$\boldsymbol{K}_1 := \left\{ \eta \in \boldsymbol{R}^p : \max_{1 \le i \le p} \eta_i \le 0 \right\},$$

$$\boldsymbol{K}_2 := \left\{ \eta \in \boldsymbol{R}^p : \eta_1 \ge \max_{2 \le i \le p} \eta_i \right\},$$

$$\boldsymbol{K}_3 := \{ \eta \in \boldsymbol{R}^p : \eta_1 \ge \eta_2 \ge \cdots \ge \eta_p \}.$$

There is an extensive literature on such problems; see the book of Robertson *et al.* (1988) or Akkerboom's (1990) lecture notes. In particular, likelihood ratio (LR) tests have received a lot of attention. It is not clear, however, in what sense these tests or its competitors are optimal. The main goal of the present paper is to find nonrandomized tests

$$\boldsymbol{X} \ni x \mapsto \mathbf{1}\{x \notin A\}$$

of $\Theta \cap \boldsymbol{K}$ versus $\Theta \backslash \boldsymbol{K}$ with small risk

$$R(A) := \sup_{\theta \in \Theta \backslash \boldsymbol{K}} P_\theta A$$

under the restriction

$$P_\theta A \geq 1 - \alpha \qquad \forall \theta \in \Theta \cap \boldsymbol{K} \tag{1.1}$$

for some fixed level $\alpha \in \,]0, 1/2[$. In other words, we look for a nonrandomized test of $\Theta \cap \boldsymbol{K}$ versus $\Theta \backslash \boldsymbol{K}$ with level α and small bias $R(A) - (1 - \alpha)$. Typically $R(A) > 1 - \alpha$, because the boundary of $\boldsymbol{K}$ is not smooth; see problem 7 in Chapter 4 of Lehmann (1986).

In Section 2 we consider a simple shift model and minimize the risk R within the class of all acceptance regions $A \subset \boldsymbol{R}^p$ such that

$$A = A + \boldsymbol{K} := \{x + \eta : x \in A, \eta \in \boldsymbol{K}\}$$

and (1.1) holds. The monotonicity constraint $A = A + \boldsymbol{K}$ is a natural requirement, especially when considering the cones $\boldsymbol{K}_j$ mentioned above. It is also mathematically convenient, although there might be decision theoretical arguments against it. It turns out that the corresponding minimax test is constructed according to Roy's (1957) union-intersection (UI) principle, where $\boldsymbol{K}$ is represented as an intersection of a minimal family of halfspaces. In a normal shift model this test is different from the LR-test in general. The latter test is optimal with respect to a different, but weaker criterion. All proofs are deferred to Section 4.

In Section 3 we imitate the minimax test of Section 2 in the context of linear regression and one-sided tests for covariance matrices. In the latter case we consider a cone $\boldsymbol{K}$ which is not polyhedral (i.e. defined by finitely many linear inequalities) as are the examples $\boldsymbol{K}_j$ above.

2. A minimax result in shift families

In this section let $\boldsymbol{X} = \Theta = \boldsymbol{R}^p$ and $P_\theta := P_o * \delta_\theta$, where the probability distribution P_o is absolutely continuous with respect to Lebesgue measure on $\boldsymbol{R}^p$ and has full support. Now we consider the class $\mathcal{A}(\boldsymbol{K})$ of all Borel sets $A \subset \boldsymbol{R}^p$ such that $A = A + \boldsymbol{K}$. Further let $\mathcal{A}_\alpha(\boldsymbol{K})$ be the set of all $A \in \mathcal{A}(\boldsymbol{K})$ such that $P_o A \geq 1 - \alpha$. One easily verifies that

$$P_\theta A \leq P_{\theta+\eta} A \qquad \forall \theta \in \boldsymbol{R}^p \quad \forall \eta \in \boldsymbol{K} \quad \forall A \in \mathcal{A}(\boldsymbol{K}). \tag{2.1}$$

In particular, any $A \in \mathcal{A}_\alpha(\boldsymbol{K})$ satisfies (1.1).

Let us introduce some notation. The support function of a set $B \subset \boldsymbol{R}^p$ is defined as

$$\sigma(B, z) := \sup_{x \in B} \langle x, z \rangle \qquad (z \in \boldsymbol{R}^p),$$

where $\langle x, z \rangle := x'z$ is the usual inner product on $\boldsymbol{R}^p$, and $\|\cdot\|$ is the corresponding norm. The set

$$B^* := \{z \in \boldsymbol{R}^p : \sigma(B, z) \leq 0\}$$

is the so-called dual cone of B. With the closed halfspaces

$$H_z(r) := \{x \in \boldsymbol{R}^p : \langle x, z \rangle \leq r\} \qquad (z \in \boldsymbol{R}^p, \;\; r \in [-\infty, \infty])$$

one can also write $B^* = \bigcap_{z \in B} H_z(0)$. The convex hull of B is denoted by $\mathrm{conv}(B)$, and $\mathrm{cone}(B) := \{\lambda x : \lambda \geq 0, x \in \mathrm{conv}(B)\}$ is the smallest convex cone containing B. Finally let $\bar{B}$ be the closure of B, and define $\mathrm{dist}(x, B) := \inf_{y \in B} \|x - y\|$.

With the help of Stein's (1956) theorem one can show that the convex sets in $\mathcal{A}(\boldsymbol{K})$ define reasonable tests of $\boldsymbol{K}$ versus $\boldsymbol{R}^p \setminus \boldsymbol{K}$.

PROPOSITION 2.1. *A closed, convex set $C \subset \boldsymbol{R}^p$ belongs to $\mathcal{A}(\boldsymbol{K})$ if, and only if,*

$$\{z \in \boldsymbol{R}^p : \sigma(C, z) < \infty\} \subset \boldsymbol{K}^*. \tag{2.2}$$

In that case the test $\mathbf{1}\{\cdot \notin C\}$ is admissible in the following sense: Let P_o be a nonsingular Gaussian distribution, and let $\phi : \boldsymbol{X} \to [0,1]$ be another test such that $P_o\phi \leq 1 - P_oC$ and $P_\theta\phi \geq 1 - P_\theta C$ for all $\theta \in \boldsymbol{R}^p \setminus \boldsymbol{K}$. Then $\phi(x) = \mathbf{1}\{x \notin C\}$ for P_o-almost all $x \in \boldsymbol{R}^p$.

Now we construct some special sets in $\mathcal{A}_\alpha(\boldsymbol{K})$. Let $\boldsymbol{S}(\boldsymbol{R}^p)$ be the unit sphere in $\boldsymbol{R}^p$, and let $M = M(\boldsymbol{K}) := \boldsymbol{K}^* \cap \boldsymbol{S}(\boldsymbol{R}^p)$. For $x \in \boldsymbol{R}^p$ and $z \in M$ define

$$T_z(x) := P_o H_z(\langle x, z \rangle).$$

We regard T_z as a test statistic for testing the simple hypothesis $H_z(0)$. The distribution $P_o \circ T_z^{-1}$ of T_z under P_o is the uniform distribution on $[0,1]$. For $\emptyset \neq B \subset M$ let

$$T_B(x) := \sup_{z \in B} T_z(x).$$

Then $\mathbf{1}\{T_B(\cdot) > \beta\}$ defines a UI-test of the hypothesis B^* in the sense of Roy (1957). If P_o is a nonsingular Gaussian distribution, then T_M is equivalent to the LR-test statistic. Let us summarize some properties of T_B.

PROPOSITION 2.2. *T_B equals $T_{\bar{B}}$, and the distribution $P_o \circ T_B^{-1}$ is continuous. Let β_B be a minimal number in $]0,1[$ such that*

$$P_o A_B = 1 - \alpha \quad \textit{where} \quad A_B := \{x \in \boldsymbol{R}^p : T_B(x) \leq \beta_B\}.$$

Then A_B is a closed, convex set in $\mathcal{A}_\alpha(\boldsymbol{K})$.

One might wonder, whether there is a smallest closed subset B of M such that $B^* = \boldsymbol{K}$. Let $E = E(\boldsymbol{K})$ be the set of all $e \in M$, which are extremal in the following sense: If $e = \lambda y + \mu z$ for $\lambda, \mu > 0$ and $y, z \in M$, then $y = z = e$.

PROPOSITION 2.3. *Suppose that $\boldsymbol{K}$ has nonvoid interior. Then $\boldsymbol{K}^* = \mathrm{cone}(E)$ and $\boldsymbol{K} = E^*$. If B is any closed subset of M such that $\boldsymbol{K} = B^*$, then $E \subset B$.*

Thus $\bar{E}$ has the above minimality property. More important is that the corresponding set A_E minimizes R over $\mathcal{A}_\alpha(\boldsymbol{K})$ and defines a consistent test.

THEOREM 2.1. *Suppose that $\boldsymbol{K}$ has nonvoid interior. Then*

$$R(A_E) = \min_{A \in \mathcal{A}_\alpha(\boldsymbol{K})} R(A).$$

For $E \subset B \subset M$,

$$R(A_B) = \beta_B,$$

and the test $\mathbf{1}\{\cdot \notin A_B\}$ is consistent in that

$$P_\theta A_B \to 0 \quad \text{as} \quad \operatorname{dist}(\theta, \boldsymbol{K}) \to \infty.$$

For the special cones $\boldsymbol{K}_j$ mentioned in the introduction, one can easily deduce from Proposition 2.3 that

$$\begin{aligned} E(\boldsymbol{K}_1) &= \{e_{(i)} : 1 \le i \le p\}, \\ E(\boldsymbol{K}_2) &= \{2^{-1/2}(e_{(i)} - e_{(1)}) : 2 \le i \le p\}, \\ E(\boldsymbol{K}_3) &= \{2^{-1/2}(e_{(i+1)} - e_{(i)}) : 1 \le i \le p-1\}, \end{aligned}$$

where $e_{(1)}, e_{(2)}, \ldots, e_{(p)}$ is the standard basis of $\boldsymbol{R}^p$. Let P_o be the standard normal distribution $\mathcal{N}(0, I_p)$. Then $T_z(x) = \Phi(\langle x, z \rangle)$ with the standard normal distribution function Φ. Hence one can also write $A_B = \{x \in \boldsymbol{R}^p : \sup_{z \in B} \langle x, z \rangle \le \tilde{\beta}_B\}$, and $\beta_B = \Phi(\tilde{\beta}_B)$. For instance, the set $A_{E(\boldsymbol{K}_1)}$ equals $\{x \in \boldsymbol{R}^p : \max_{1 \le i \le p} x_i \le \tilde{\beta}_E\}$. Using the standard expansion $\Phi(r) = 1 - \exp(-r^2/2 + o(r^2))$ as $r \to \infty$, one can show that $\tilde{\beta}_{E(\boldsymbol{K}_1)} = \sqrt{2 \log p}(1 + o(1))$ and

$$R(A_{E(\boldsymbol{K}_1)}) = \Phi(\tilde{\beta}_{E(\boldsymbol{K}_1)}) = 1 - p^{-1+o(1)} \quad \text{as} \quad p \to \infty.$$

On the other hand, $A_{M(\boldsymbol{K}_1)} = \{x \in \boldsymbol{R}^p : \sum_{i=1}^p (x_i^+)^2 \le \tilde{\beta}^2_{M(\boldsymbol{K}_1)}\}$. The Law of Large Numbers for $p^{-1} \sum_{i=1}^p (x_i^+)^2$ yields $\tilde{\beta}_{M(\boldsymbol{K}_1)} = \sqrt{p/2}(1 + o(1))$ and

$$R(A_{M(\boldsymbol{K}_1)}) = 1 - \exp(-p/4 + o(p)) \quad \text{as} \quad p \to \infty.$$

Hence the risk of A_E can be considerably smaller than the risk of A_M. Similar arguments apply to $\boldsymbol{K}_2$ and $\boldsymbol{K}_3$.

In the standard Gaussian shift model, $\operatorname{dist}(\theta, \boldsymbol{K})$ is a measure of how good a test ϕ of $\boldsymbol{K}$ versus $\{\theta\}$ can be. One might argue that instead of $R(1 - \phi)$ one should consider the risk

$$R_\delta(1 - \phi) := \sup_{\theta \in \boldsymbol{R}^p : \operatorname{dist}(\theta, \boldsymbol{K}) \ge \delta} P_\theta(1 - \phi)$$

for some (but what?) $\delta > 0$. This is an interesting open problem. Presumably neither A_E nor A_M are optimal with respect to this criterion. So far we can only show that A_M is approximately optimal as $\delta \to \infty$ by modifying Stein's

(1956) arguments. However, this is admittedly a weak optimality result, because $R_\delta(1-\phi) \to 0$ as $\delta \to \infty$ for most reasonable tests ϕ.

THEOREM 2.2. *Let $P_o = \mathcal{N}(0, I_p)$, and let ϕ be any test such that $P_o\phi = \alpha$. Then*

$$\frac{R_\delta(1-\phi)}{R_\delta(A_M)} \to \infty \qquad as \quad \delta \to \infty$$

unless $\phi(x) = \mathbf{1}\{x \notin A_M\}$ for P_o-almost all $x \in \boldsymbol{R}^p$.

Since the two criteria $R(\cdot)$ and '$R_\infty(\cdot)$' lead to different answers, one could combine the two tests A_E and A_M via the UI-principle or use A_B for some set B strictly between E and M.

3. Modifications

3.1 *Linear regression*

Let us describe briefly how one can modify the tests A_B of the preceding section in the context of linear regression: Let

$$Y = D\theta + E,$$

where $\theta \in \boldsymbol{R}^p$ is an unknown parameter, $D \in \boldsymbol{R}^{n\times p}$ is a given design matrix with rank $p < n$, and $E \in \boldsymbol{R}^n$ is an unobserved vector having independent, Gaussian components with mean zero and unknown standard deviation $\sigma > 0$. As in Section 2 let $\boldsymbol{K}$ be a closed, convex cone in $\boldsymbol{R}^p$ such that interior$(\boldsymbol{K}) \neq \emptyset$. With $V := (D'D)^{-1}$ let

$$\hat\theta = \hat\theta(Y) := VD'Y, \qquad \hat\sigma = \hat\sigma(Y) := \sqrt{\|Y - DVD'Y\|^2/(n-p)}$$

be the usual estimators for θ and σ. The distribution of $\hat\gamma = \hat\gamma(Y) := \hat\sigma^{-1}\hat\theta$ depends only on the parameter $\gamma := \sigma^{-1}\theta$, and $\theta \in \boldsymbol{K}$ if, and only if, $\gamma \in \boldsymbol{K}$. For $z \in M$ and $E \subset B \subset M$ let

$$\tilde T_z(x) := (z'Vz)^{-1/2}\langle\hat\gamma(x), z\rangle \qquad \text{and} \qquad \tilde T_B(x) := \sup_{z\in B} \tilde T_z(x).$$

All random variables $\tilde T_z(Y)$, $z \in M$, have a student distribution with $n-p$ degrees of freedom if $\gamma = 0$. One easily verifies that

$$\tilde T_B(x + D\eta) \le \tilde T_B(x) \qquad \forall x \in \boldsymbol{R}^n \backslash D\boldsymbol{R}^p \quad \forall \eta \in \boldsymbol{K}.$$

Hence, if $\tilde\beta_B > 0$ is chosen such that

$$\mathbb{P}_{\gamma=0}\{\tilde T_B(Y) > \tilde\beta_B\} = \alpha,$$

then $\mathbf{1}\{\tilde T_B(\cdot) > \tilde\beta_B\}$ defines a test of $\boldsymbol{K}$ at level α. For $B = M$ this is just the $\bar E^2$ test as defined in Robertson *et al.* (1988). In view of Theorem 2.1, however, we favor the test $\mathbf{1}\{\tilde{T}_E(\cdot) > \tilde{\beta}_E\}$. In fact one can easily modify the proof of Theorem 2.1 in order to show that this test has minimal bias among all tests of the form $\mathbf{1}\{\hat{\gamma}(\cdot) \notin A\}$, where A is a set in $\mathcal{A}(\boldsymbol{K})$ such that $\mathbb{P}_{\gamma=0}\{\hat{\gamma} \in A\} \geq 1-\alpha$. In case of $\boldsymbol{K} = \boldsymbol{K}_2$ we end up with Dunnet's (1955) test (extended to arbitrary design matrices D), which rejects the hypothesis if

$$\max_{2\leq i\leq p} \frac{\hat{\theta}_i - \hat{\theta}_1}{\hat{\sigma}\sqrt{V_{ii} - 2V_{i1} + V_{11}}}$$

is too large.

3.2 *One-sided tests for covariance matrices*

Let $\boldsymbol{X}$ be the space of all symmetric matrices in $\boldsymbol{R}^{d\times d}$ equipped with inner product $\langle x, z\rangle := \operatorname{trace}(xz)$ and norm $\|x\| := \langle x, x\rangle^{1/2}$. It can be identified with $\boldsymbol{R}^{d(d+1)/2}$. Let Θ be the set of all positive definite $\theta \in \boldsymbol{X}$. Suppose that one observes a random matrix $S \in \Theta$ having Wishart distribution $\mathcal{W}(\Sigma, n)$ with unknown matrix parameter $\Sigma \in \Theta$ and $n \geq d$ degrees of freedom.

There are various test hypotheses in multivariate analysis involving closed, convex cones in $\boldsymbol{X}$. For instance consider the hypothesis $I + \bar{\Theta}$, where I is the identity matrix in $\boldsymbol{R}^{d\times d}$, and $\bar{\Theta}$ is the closed, convex cone of nonnegative definite matrices in $\boldsymbol{X}$. In other words one wants to test, whether $u'\Sigma u \geq u'u$ for all $u \in \boldsymbol{S}(\boldsymbol{R}^d)$. A natural test statistic for such a simple hypothesis is $u'u/u'Su$, and Roy's (1957) UI-principle leads to the test statistic

$$\max_{u\in \boldsymbol{S}(\boldsymbol{R}^d)} (u'u/u'Su) = \lambda_{\min}(S)^{-1},$$

where $\lambda_{\min}(x)$ stands for the smallest eigenvalue of $x \in \boldsymbol{X}$. Kuriki (1993) considered the LR-test for a similar testing problem.

Now it is shown that $\lambda_{\min}(S)^{-1}$ is indeed a reasonable test criterion. First of all one can easily show that

$$\lambda_{\min}(S)^{-1} \leq \lambda_{\min}(\tilde{S})^{-1} \quad \text{if} \quad \Sigma \in I + \bar{\Theta},$$

where $\tilde{S}$ is the unobserved random matrix $\Sigma^{-1/2}S\Sigma^{-1/2}$ having a standard Wishart distribution with n degrees of freedom. Thus the test $\mathbf{1}\{\lambda_{\min}(S)^{-1} > \beta_n\}$ with the $(1-\alpha)$-quantile β_n of $\mathcal{L}(\lambda_{\min}(\tilde{S})^{-1})$ has level α. Note also that this test leads to confidence 'intervals'

$$\{H \in \Theta : \lambda_{\min}(H^{-1}S)^{-1} \leq \beta_n\} = \beta_n S - \bar{\Theta}$$

for Σ with coverage probability $1-\alpha$.

As n tends to infinity, $n^{1/2}(\tilde{S}-I)$ converges in distribution to a random matrix $\tilde{X}$ such that $2^{-1/2}\tilde{X}$ has a standard normal distribution on $\boldsymbol{X}$, and $n^{1/2}(\beta_n - 1)$ converges to the $(1-\alpha)$-quantile β of $\mathcal{L}(-\lambda_{\min}(\tilde{X}))$. Suppose that

$$\Sigma = I + n^{-1/2}\theta_n,$$

where $\theta_n \in \boldsymbol{X}$ converges to some $\theta \in \boldsymbol{X}$. Then $n^{1/2}(S-I)$ converges in distribution to $\theta + \tilde{X}$, and

$$\mathbb{P}\{\lambda_{\min}(S)^{-1} > \beta_n\} \to \mathbb{P}\{\theta + \tilde{X} \notin A\},$$

where

$$A := \{x \in \boldsymbol{X} : -\lambda_{\min}(x) \geq \beta\} \in \mathcal{A}(\bar{\Theta}).$$

Thus Roy's test behaves asymptotically as the test $\mathbf{1}\{\cdot \notin A\}$ of the hypothesis $\bar{\Theta}$ in the shift model $(\mathcal{L}(\theta + \tilde{X}) : \theta \in \boldsymbol{X})$. It is optimal in that

$$A = A_{E(\bar{\Theta})}.$$

For one can easily show that $A = A_B$ with the set $B := \{-uu' : u \in \boldsymbol{S}(\boldsymbol{R}^d)\} \subset \boldsymbol{S}(\boldsymbol{X})$. Further, one can deduce from the spectral representation of points in $\boldsymbol{X}$ that $\bar{\Theta}^* = -\bar{\Theta}$ and $B = E(\bar{\Theta})$.

4. Proofs

Before proving the results of Section 2 let us recall some well-known facts from convex analysis. The support function $\sigma(B, \cdot)$ of $B \subset \boldsymbol{R}^p$ coincides with $\sigma(\overline{\mathrm{conv}(B)}, \cdot)$, and

$$\overline{\mathrm{conv}(B)} = \bigcap_{z \in \boldsymbol{R}^p} H_z(\sigma(B, z)).$$

More generally,

$$\mathrm{dist}(x, \mathrm{conv}(B)) = \sup_{z \in \boldsymbol{S}(\boldsymbol{R}^p)} (\langle x, z \rangle - \sigma(B, z)) \vee 0$$

for all $x \in \boldsymbol{R}^p$. Similarly, $B^* = \overline{\mathrm{cone}(B)}^*$, and $\overline{\mathrm{cone}(B)} = B^{**}$.

PROOF OF PROPOSITION 2.1. Suppose that $x + \eta \notin C$ for some pair $(x, \eta) \in C \times \boldsymbol{K}$. Then there exists a $z \in \boldsymbol{R}^p$ such that $\sigma(C, z) < \langle x + \eta, z \rangle$. Since $\langle x, z \rangle \leq \sigma(C, z)$, this implies that $z \in \{\sigma(C, \cdot) < \infty\} \backslash \boldsymbol{K}^*$.

On the other hand, if $C \in \mathcal{A}(\boldsymbol{K})$, then $\sigma(C, \cdot) \equiv \sigma(C, \cdot) + \sigma(\boldsymbol{K}, \cdot)$, and (2.2) follows from the fact that $\sigma(\boldsymbol{K}, \cdot) \in \{0, \infty\}$.

The admissibility of $\mathbf{1}\{\cdot \notin C\}$ is a direct consequence of (2.2) and Stein's (1956) theorem. □

PROOF OF PROPOSITION 2.2. Since $T_z(x) = \int \mathbf{1}\{\langle y, z \rangle \leq \langle x, z \rangle\} P_o(dy)$ and

$$\lim_{(x,z) \to (x_o, z_o)} \mathbf{1}\{\langle y, z \rangle \leq \langle x, z \rangle\} = \mathbf{1}\{\langle y, z_o \rangle \leq \langle x_o, z_o \rangle\} \quad \text{if} \quad \langle y, z_o \rangle \neq \langle x_o, z_o \rangle,$$

it follows from dominated convergence that $T_z(x)$ is a continuous function of $(x, z) \in \boldsymbol{R}^p \times M$. Thus $T_B = T_{\bar{B}} = \max_{z \in \bar{B}} T_z$.

Since P_o has full support, the latter representation of T_B implies that for any $\beta \in [0, 1]$ the set $\{T_B(\cdot) = \beta\}$ has nonvoid interior. Moreover, since both sets

$\{T_B(\cdot) \le \beta\}$ and $\{T_B(\cdot) < \beta\}$ are convex, their boundaries have Lebesgue measure zero. Thus $P_o\{T_B(\cdot) = \beta\} = 0$, whence $P_o \circ T_B^{-1}$ is continuous.

The set A_B is closed and convex, and for all $x \in \boldsymbol{R}^p$ and $\eta \in \boldsymbol{K}$,

$$T_B(x+\eta) = \sup_{z\in B} P_o H_z(\langle x, z\rangle + \langle \eta, z\rangle) \le T_B(x),$$

because $B \subset \boldsymbol{K}^*$. Consequently $A_B \in \mathcal{A}(\boldsymbol{K})$. □

PROOF OF PROPOSITION 2.3. Let η_o be an interior point of $\boldsymbol{K}$. Then $\langle \eta_o, z\rangle < 0$ for all $z \in \boldsymbol{K}^*\backslash\{0\}$, and $\pi(z) := |\langle \eta_o, z\rangle|^{-1} z$ defines a homeomorphism from M onto the compact, convex set $\pi(M) = \boldsymbol{K}^* \cap P$, where $P := \{x \in \boldsymbol{R}^p : \langle x, \eta_o\rangle = -1\}$. One can easily show that $\pi(E)$ is the set of all extreme points of $\pi(M)$. Therefore $\pi(M)$ equals $\operatorname{conv}(\pi(E))$; see Corollary 18.5.1 of Rockafellar (1970). Consequently,

$$\begin{aligned}\boldsymbol{K}^* &= \{\lambda x : \lambda \ge 0, x \in \pi(M)\} = \operatorname{cone}(\pi(E)) = \operatorname{cone}(E),\\ \boldsymbol{K} &= \boldsymbol{K}^{**} = \operatorname{cone}(E)^* = E^*.\end{aligned}$$

Now let $B \subset M$ be closed such that $B^* = \pi(B)^* = \boldsymbol{K}$. Then $\pi(M)$ equals $\overline{\operatorname{cone}(\pi(B))} \cap P$. Since $\pi(B)$ is a compact subset of the hyperplane P, and since $0 \notin P$, one can write

$$\overline{\operatorname{cone}(\pi(B))} = \operatorname{cone}(\pi(B)) = \{\lambda x : \lambda \ge 0, x \in \operatorname{conv}(\pi(B))\}.$$

Consequently, $\pi(M) = \operatorname{conv}(\pi(B))$. But this implies that $\pi(E) \subset \pi(B)$, because $\pi(E)$ is the set of extreme points of $\pi(M)$. □

PROOF OF THEOREM 2.1. Let A be any set in $\mathcal{A}_\alpha(\boldsymbol{K})$. We first prove the following expression for $R(A)$, where Δ_o is a dense subset of the boundary $\partial\boldsymbol{K}$ of $\boldsymbol{K}$ to be specified later:

$$R(A) = \sup_{\theta\in\Delta_o} \lim_{r\to\infty} P_{r\theta} A. \tag{4.1}$$

It follows from the absolute continuity of P_o that $\boldsymbol{R}^p \ni \theta \mapsto P_\theta$ is continuous with respect to total variation. In particular,

$$R(A) \ge \sup_{\theta\in\partial\boldsymbol{K}} P_\theta A = \sup_{\theta\in\Delta_o} P_\theta A.$$

But $r\theta \in \partial\boldsymbol{K}$ for all $\theta \in \partial\boldsymbol{K}$ and $r \ge 0$, and $P_{r\theta}A$ is nondecreasing in r by (2.1). Hence

$$\sup_{\theta\in\Delta_o} P_\theta A = \sup_{\theta\in\Delta_o} \lim_{r\to\infty} P_{r\theta} A.$$

On the other hand, let η_o be a fixed interior point of $\boldsymbol{K}$, and let θ be any point in $\boldsymbol{R}^p\backslash\boldsymbol{K}$. Then $\theta + r\eta_o = r(r^{-1}\theta + \eta_o) \in \partial\boldsymbol{K}$ for some $r = r(\theta) > 0$, and (2.1) implies that

$$P_\theta A \le P_{\theta + r\eta_o} A \le \sup_{\theta\in\Delta_o} P_\theta A,$$

which completes the proof of (4.1).

Specifically, let Δ_o be the set of all $\theta \in \partial \boldsymbol{K}$ such that

$$(\boldsymbol{K} - \theta)^* = \{re(\theta) : r \geq 0\}$$

for a unique $e(\theta) \in \boldsymbol{S}(\boldsymbol{R}^p)$, where $\boldsymbol{K} - \theta := \{\eta - \theta : \eta \in \boldsymbol{K}\}$. The fact that Δ_o is dense in $\partial \boldsymbol{K}$ can be seen as follows: For $\theta \in \partial \boldsymbol{K}$ and $\epsilon > 0$ let $\eta \in \text{interior}(\boldsymbol{K}) \cap B(\theta, \epsilon)$, where $B(\theta, \epsilon)$ denotes the closed ball around θ with radius ϵ. Let $R = R(\eta)$ be the maximum of all $r \in (0, \epsilon]$ such that $B(\eta, r) \subset \boldsymbol{K}$. Then there exists a $\tilde{\theta} \in \partial \boldsymbol{K}$ such that $\|\tilde{\theta} - \eta\| = R$; in particular, $\|\tilde{\theta} - \theta\| \leq 2\epsilon$. But

$$(\boldsymbol{K} - \tilde{\theta})^* \subset (B(\eta, R) - \tilde{\theta})^* = \{r(\tilde{\theta} - \eta) : r \geq 0\}.$$

Since $(\boldsymbol{K} - \tilde{\theta})^*$ necessarily contains a point different from 0, this implies that $(\boldsymbol{K} - \tilde{\theta})^*$ equals $\{r(\tilde{\theta} - \eta) : r \geq 0\}$, whence $\tilde{\theta} \in \Delta_o$.

An important fact is that

$$E_o := \{e(\theta) : \theta \in \Delta_o\} \text{ is a dense subset of } E. \tag{4.2}$$

For one can easily show that $(\boldsymbol{K} - \theta)^* = \boldsymbol{K}^* \cap \{\theta\}^\perp$ for all $\theta \in \partial \boldsymbol{K}$. This implies that $E_o \subset M$. Further, for $\theta \in \Delta_o$ let $e(\theta) = \lambda y + \mu z$ with $\lambda, \mu > 0$ and $y, z \in M$. Since $\langle \theta, e(\theta) \rangle = 0$ and $\langle \theta, y \rangle \vee \langle \theta, z \rangle \leq 0$, it follows that $y = z = e(\theta)$. Thus $E_o \subset E$. According to Proposition 2.3 it suffices to show that $E_o^* = \boldsymbol{K}$. Obviously $E_o^* \supset \boldsymbol{K}$, and

$$\partial E_o^* \supset \{x \in E_o^* : \langle x, e \rangle = 0 \text{ for some } e \in E_o\} \supset \Delta_o.$$

Consequently $\partial \boldsymbol{K} \subset \partial E_o^*$. This implies that $E_o^* \subset \boldsymbol{K}$. For if $\theta \in E_o^* \backslash \boldsymbol{K}$ and $\eta_o \in \text{interior}(\boldsymbol{K}) \subset \text{interior}(E_o^*)$, then there would exist a $\lambda = \lambda(\theta) \in]0, 1[$ such that $(1 - \lambda)\theta + \lambda \eta_o \in \partial \boldsymbol{K} \cap \text{interior}(E_o^*)$.

Next we deduce the crucial formula

$$R(A) = \sup_{x \in A} T_E(x). \tag{4.3}$$

For any fixed $\theta \in \Delta_o$ and $r > 0$,

$$P_{r\theta} A = P_o(A - r\theta) = P_o(A + r(\boldsymbol{K} - \theta)).$$

The set $\boldsymbol{K} - \theta$ is convex and contains 0. Hence $r(\boldsymbol{K} - \theta) \subset s(\boldsymbol{K} - \theta)$ for $0 < r < s$, $\text{cone}(\boldsymbol{K} - \theta) = \bigcup_{r>0} r(\boldsymbol{K} - \theta)$, and it follows from monotone convergence that

$$\lim_{r \to \infty} P_{r\theta} A = P_o(A + \text{cone}(\boldsymbol{K} - \theta)).$$

But $\text{interior}(H_{e(\theta)}(0)) \subset \text{cone}(\boldsymbol{K} - \theta) \subset H_{e(\theta)}(0)$, whence

$$P_o(A + \text{cone}(\boldsymbol{K} - \theta)) = P_o H_{e(\theta)}(\sigma(A, e(\theta))) = \sup_{x \in A} T_{e(\theta)}(x).$$

Consequently (4.3) follows from (4.1) and (4.2) together with the first statement of Proposition 2.2.

Formula (4.3) shows that A can be replaced with the larger set $\{x \in \boldsymbol{R}^p : T_E(x) \leq R(A)\}$ without increasing $R(A)$. Since $P_o A \geq 1 - \alpha$, it follows from the definition of β_E that

$$R(A) \geq \beta_E = \sup_{x \in A_E} T_E(x) = R(A_E).$$

For $E \subset B \subset M$ it follows from (4.3) and $T_B \geq T_E$ that $R(A_B)$ is not greater than $\sup_{x \in A_B} T_B(x) = \beta_B$. On the other hand, for $\theta \in \Delta_o$ and $\lambda \in \boldsymbol{R}$,

$$T_B(r\theta + \lambda e(\theta)) \to T_{e(\theta)}(\lambda e(\theta)) \qquad \text{as} \quad r \to \infty.$$

This follows straightforwardly from the fact that $\langle \theta, z \rangle \leq 0$ for all $z \in M$ with equality if, and only if, $z = e(\theta)$. Consequently, if $T_{e(\theta)}(\lambda e(\theta)) < \beta_B$, then $r\theta + \lambda e(\theta) \in A_B$ for sufficiently large $r > 0$. Since $T_E(r\theta + \lambda e(\theta)) \geq T_{e(\theta)}(r\theta + \lambda e(\theta)) = T_{e(\theta)}(\lambda e(\theta))$, this shows that $\sup_{x \in A_B} T_E(x) \geq \beta_B$.

As for the consistency of A_B it suffices to show that $T_E(x)$ tends to one as $\operatorname{dist}(x, \boldsymbol{K}) \to \infty$. But

$$T_E(x) \geq P_o B\left(0, \sup_{e \in E} \langle x, e \rangle \vee 0\right) \qquad \text{and} \qquad \operatorname{dist}(x, \boldsymbol{K}) = \sup_{z \in M} \langle x, z \rangle \vee 0.$$

With $\pi(z) := |\langle \eta_o, z \rangle|^{-1} z$ as in the proof of Proposition 2.3 the assertion follows from the inequalities

$$\begin{aligned}
\sup_{e \in E} \langle x, e \rangle \vee 0 &\geq \min_{z \in M} |\langle \eta_o, z \rangle| \sup_{e \in E} \langle x, \pi(e) \rangle \vee 0 \\
&= \min_{z \in M} |\langle \eta_o, z \rangle| \sup_{z \in M} \langle x, \pi(z) \rangle \vee 0 \\
&\geq \left(\min_{z \in M} |\langle \eta_o, z \rangle| \Big/ \max_{z \in M} |\langle \eta_o, z \rangle| \right) \sup_{z \in M} \langle x, z \rangle \vee 0. \qquad \square
\end{aligned}$$

PROOF OF THEOREM 2.2. Since $P_o(1-\phi) = P_o A_M = 1-\alpha$, one may assume that $P_o\{x \in \boldsymbol{R}^p \backslash A_M : \phi(x) < 1\} > 0$. But A_M can be written as $\bigcap_{z \in B} H_z(\tilde{\beta}_M)$ for some $\tilde{\beta}_M > 0$ and a countable, dense subset B of M. Hence

$$\omega := \int \mathbf{1}\{x \notin H_z(\tilde{\beta}_M)\}(1 - \phi(x))dx > 0$$

for some $z \in M$, and for $\delta > \tilde{\beta}_M$,

$$\begin{aligned}
&\operatorname{dist}(\delta z, \boldsymbol{K}) = \delta, \\
&P_{\delta z}(1 - \phi) \geq \int \mathbf{1}\{x \in B(\delta z, \delta - \tilde{\beta}_M)\}(1 - \phi(x))P_{\delta z}(dx) \\
&\qquad \geq (2\pi)^{-p/2} \exp(-(\delta - \tilde{\beta}_M)^2/2) \\
&\qquad\qquad \cdot \int \mathbf{1}\{x \in B(\delta z, \delta - \tilde{\beta}_M)\}(1 - \phi(x))dx \\
&\qquad = (2\pi)^{-p/2} \exp(-(\delta - \tilde{\beta}_M)^2/2)(\omega + o(1)) \qquad \text{as} \quad \delta \to \infty.
\end{aligned}$$

On the other hand, if θ is any parameter with $\mathrm{dist}(\theta, \boldsymbol{K}) \geq \delta$, then there is a $z(\theta) \in M$ such that $\langle \theta, z(\theta) \rangle \geq \delta$. Hence

$$P_\theta A_M \leq P_\theta H_{z(\theta)}(\tilde{\beta}_M) \leq \Phi(\tilde{\beta}_M - \delta),$$

and the assertion follows from the well-known fact that $\exp(r^2/2)\Phi(-r) \to 0$ as $r \to \infty$. □

References

Akkerboom, J. C. (1990). Testing problems with linear or angular inequality constraints, *Lecture Notes in Statist.*, **62**, Springer, Berlin.

Dunnett, C. W. (1955). A multiple comparisons procedure for comparing several treatments with a control, *J. Amer. Statist. Assoc.*, **50**, 1096–1121.

Kuriki, S. (1993). Likelihood ratio tests for covariance structure in random effects models, *J. Multivariate Anal.*, **46**, 175–197.

Lehmann, E. L. (1986). *Testing Statistical Hypotheses*, 2nd ed., Wiley, New York.

Robertson, T., Wright, F. T. and Dykstra, R. L. (1988). *Order Restricted Statistical Inference*, Wiley, New York.

Rockafellar, R. T. (1970). *Convex Analysis*, Princeton University Press, New Jersey.

Roy, S. N. (1957). *Some Aspects of Multivariate Analysis*, Wiley, New York.

Stein, C. (1956). The admissibility of Hotelling's T^2-test, *Ann. Math. Statist.*, **27**, 616–623.