NORTH-HOLLAND

Bounds on Mahalanobis Norms and Their Applications

D. R. Jensen
Department of Statistics
Virginia Polytechnic Institute and State University
Blacksburg, Virginia 24061

Submitted by George P. H. Styan

ABSTRACT

Local and global bounds for ratios of norms, and minimal and maximal norms, are constructed for pairs and ensembles of quadratic norms on $\mathbb{R}^k$, with corresponding results for Mahalanobis distance functions. These support envelopes for distributions of certain quadratic forms in Gaussian variates. Applications are noted in the use of quadratic classification rules and in assessing hit probabilities in ballistic systems. © 1997 Elsevier Science Inc.

1. INTRODUCTION

Let $\mathbf{M}$ be a positive definite real $k \times k$ matrix. Numerous applications utilize norms of the type $\|\mathbf{x}\|_M = (\mathbf{x}'\mathbf{M}\mathbf{x})^{1/2}$, together with the corresponding metric $D_M(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_M$. In the form $\Delta_{\Xi}(\mathbf{x}, \mathbf{y}) = [(\mathbf{x} - \boldsymbol{\theta})'\boldsymbol{\Xi}^{-1}(\mathbf{x} - \boldsymbol{\theta})]^{1/2}$, this is the Mahalanobis (1936) *generalized distance* between an observation $\mathbf{x} \in \mathbb{R}^k$ and the centroid $\boldsymbol{\theta} \in \mathbb{R}^k$ of a k-dimensional Gaussian distribution having information matrix $\mathbf{M}$ and dispersion matrix $\boldsymbol{\Xi} = \mathbf{M}^{-1}$. For further historical details and references see Pillai (1985). In this study we develop local bounds for ratios of norms, and we construct minimal and maximal norms, for pairs and ensembles of quadratic norms on $\mathbb{R}^k$. These findings in turn support envelopes for distributions of quadratic forms in Gaussian ensembles on $\mathbb{R}^k$ and for certain mixtures over these. Applications are noted in the use of zero-mean classification rules, and in evaluating hit probabilities in ballistic systems under nonstandard conditions.

LINEAR ALGEBRA AND ITS APPLICATIONS 264:127–139 (1997)

655 Avenue of the Americas, New York, NY 10010
0024-3795/97/$17.00
PII S0024-3795(97)00235-8

2. PRELIMINARIES

2.1. *Notation*

Designate by $\mathbb{R}^k$ and $\mathbb{R}_+^k$ the Euclidean k-space and its positive orthant, and by $\mathbb{R}^k(\mathbf{0})$ and $\mathbb{R}_+^k(\mathbf{0})$ the corresponding sets excluding $\mathbf{0}$. Let S_k, S_k^0, and S_k^+ comprise the symmetric, the positive semidefinite, and the positive definite real $k \times k$ matrices; $\mathbf{A}'$ and $\mathbf{A}^{-1}$ are the transpose and inverse of $\mathbf{A}$; $\mathbf{A}^{1/2} \in S_k^+$ is the symmetric root of $\mathbf{A} \in S_k^+$; and $\mathrm{Sp}(\mathbf{z}_1, \ldots, \mathbf{z}_r) \subset \mathbb{R}^k$ is the linear span of vectors $\{\mathbf{z}_1, \ldots, \mathbf{z}_r\}$ in $\mathbb{R}^k$. Special arrays include the unit vector $\mathbf{1}_k = [1, 1, \ldots, 1] \in \mathbb{R}^k$, the identity matrix $\mathbf{I}_k$ of order k, and the collection $D_k^+ \subset S_k^+$ comprising diagonal matrices of the type $\mathrm{Diag}(a_1, \ldots, a_k)$.

Take $(\mathbb{R}_+^k, \geqslant)$ to be ordered such that $\mathbf{x} \geqslant \mathbf{y}$ in $\mathbb{R}^k$ if and only if $\{x_i \geqslant y_i;\ 1 \leqslant i \leqslant k\}$. Moreover, $(S_k, \succcurlyeq_L)$ is ordered as in Loewner (1934) such that $\mathbf{A} \succcurlyeq_L \mathbf{B}$ if and only if $\mathbf{A} - \mathbf{B} \in S_k^0$, with $\mathbf{A} \succ_L \mathbf{B}$ whenever $\mathbf{A} - \mathbf{B} \in S_k^+$. Let $\mathbf{A}$ and $\mathbf{B}$ be positive definite, and consider pairs $\{(\gamma_i, \mathbf{q}_i);\ 1 \leqslant i \leqslant k\}$ satisfying $\{(\mathbf{A} - \gamma_i \mathbf{B})\mathbf{q}_i = \mathbf{0};\ 1 \leqslant i \leqslant k\}$. Then with $\mathbf{D}_\gamma = \mathrm{Diag}(\gamma_1, \ldots, \gamma_k)$ containing the ordered roots of $|\mathbf{A} - \gamma \mathbf{B}| = 0$, we have the spectral decomposition $\mathbf{B}^{-1/2}\mathbf{A}\mathbf{B}^{-1/2} = \mathbf{P}\mathbf{D}_\gamma\mathbf{P}'$, whereas the respective vectors are related as $\{\mathbf{q}_i = \mathbf{B}^{-1/2}\mathbf{p}_i;\ 1 \leqslant i \leqslant k\}$. Clearly $\mathbf{A} \succcurlyeq_L \mathbf{B}$ if and only if $\{\gamma_i \geqslant 1;\ 1 \leqslant i \leqslant k\}$, whereas $\mathbf{B} \succcurlyeq_L \mathbf{A}$ corresponds to $\{\gamma_i \leqslant 1;\ 1 \leqslant i \leqslant k\}$. Otherwise at least one of two positive integers (r, s) exists such that

$$\{\gamma_1 \geqslant \cdots \geqslant \gamma_r > \gamma_{r+1} = 1 = \cdots = \gamma_{r+s} > \gamma_{r+s+1} \geqslant \cdots \geqslant \gamma_k > 0\}. \tag{2.1}$$

Corresponding to these are subspaces $R_1 = \mathrm{Sp}(\mathbf{q}_1, \ldots, \mathbf{q}_r)$, $R_2 = \mathrm{Sp}(\mathbf{q}_{r+1}, \ldots, \mathbf{q}_{r+s})$, and $R_3 = \mathrm{Sp}(\mathbf{q}_{r+s+1}, \ldots, \mathbf{q}_k)$, at most two of which may be empty. Specifically, if neither $\mathbf{A} \succcurlyeq_L \mathbf{B}$ nor $\mathbf{B} \succcurlyeq_L \mathbf{A}$, then R_2 alone is empty if and only if there are no unit roots.

2.2. *Spectral Matrix Bounds*

We seek bounds for matrices in $(S_k^+, \succcurlyeq_L)$. But since $(S_k^+, \succcurlyeq_L)$ is not a lattice (Halmos, 1958, p. 142), there is no greatest lower bound (glb) or least upper bound (lub) for $(\mathbf{A}, \mathbf{B})$ in $(S_k^+, \succcurlyeq_L)$. Nonetheless, we may construct *spectral* lower and upper bounds on recalling that $(\mathbb{R}_+^k, \geqslant)$ is a lattice with glb $\mathbf{x} \wedge \mathbf{y}$ and lub $\mathbf{x} \vee \mathbf{y}$ for $(\mathbf{x}, \mathbf{y})$ in $\mathbb{R}^k$, where $\mathbf{x} \wedge \mathbf{y} = [x_1 \wedge y_1, \ldots, x_k \wedge y_k]'$ and $\mathbf{x} \vee \mathbf{y} = [x_1 \vee y_1, \ldots, x_k \vee y_k]'$, with $x_i \wedge y_i = \min(x_i, y_i)$, $1 \leqslant i \leqslant k$, and $x_i \vee y_i = \max(x_i, y_i)$, $1 \leqslant i \leqslant k$. See Vulikh (1967), for example. On imbedding D_k^+ in $(\mathbb{R}_+^k, \geqslant)$, we infer that $(D_k^+, \succcurlyeq_L)$ is itself a lattice with $\mathbf{D}_a \wedge \mathbf{D}_b = \mathrm{Diag}(a_1 \wedge b_1, \ldots, a_k \wedge b_k)$ and $\mathbf{D}_a \vee \mathbf{D}_b = \mathrm{Diag}(a_1 \vee b_1, \ldots, a_k \vee b_k)$, for $\mathbf{D}_a = \mathrm{Diag}(a_1, \ldots, a_k)$ and $\mathbf{D}_b = \mathrm{Diag}(b_1, \ldots, b_k)$.

Since $\mathbf{A}$ and $\mathbf{B}$ may be recovered from $\mathbf{B}^{-1/2}\mathbf{A}\mathbf{B}^{-1/2} = \mathbf{P}\mathbf{D}_{\gamma}\mathbf{P}'$ as $\mathbf{A} = \mathbf{B}^{1/2}\mathbf{P}\mathbf{D}_{\gamma}\mathbf{P}'\mathbf{B}^{1/2}$ and $\mathbf{B} = \mathbf{B}^{1/2}\mathbf{P}\mathbf{I}_k\mathbf{P}'\mathbf{B}^{1/2}$, we accordingly define matrices $\mathbf{A} \wedge \mathbf{B} = \mathbf{B}^{1/2}\mathbf{P}(\mathbf{D}_{\gamma} \wedge \mathbf{I}_k)\mathbf{P}'\mathbf{B}^{1/2}$ and $\mathbf{A} \vee \mathbf{B} = \mathbf{B}^{1/2}\mathbf{P}(\mathbf{D}_{\gamma} \vee \mathbf{I}_k)\mathbf{P}'\mathbf{B}^{1/2}$ as the *spectral glb* and the *spectral lub*, respectively, for $(\mathbf{A}, \mathbf{B})$ in $(S_k^+, \succcurlyeq_L)$. It follows constructively that $\mathbf{A} \wedge \mathbf{B} \preccurlyeq_L \{\mathbf{A}, \mathbf{B}\} \preccurlyeq_L \mathbf{A} \vee \mathbf{B}$ on $(S_k^+, \succcurlyeq_L)$. Further details are given in Jensen (1993).

2.3. *Basic Distributions*

Probability density function and cumulative distribution function are abbreviated as pdf and cdf; $\mathscr{L}(\mathbf{Y})$ designates the law of distribution of $\mathbf{Y} \in \mathbb{R}^k$; and $N_k(\boldsymbol{\theta}, \boldsymbol{\Sigma})$ denotes the nonsingular Gaussian law on $\mathbb{R}^k$ having mean $\boldsymbol{\theta}$, dispersion matrix $\boldsymbol{\Sigma} \in S_k^+$, and Gaussian measure $G_k(\cdot; \boldsymbol{\theta}, \boldsymbol{\Sigma})$. Let $C(k)$ comprise the compact convex sets in $\mathbb{R}^k$ that are symmetric under reflection through $\mathbf{0} \in \mathbb{R}^k$; let $\{\mu_t(\cdot);\ t \in \tau\}$ be an ensemble of probability measures on $\mathbb{R}^k$; and suppose that measures $\{\nu_m(\cdot), \nu_M(\cdot)\}$ can be found such that $\nu_m(A) \leqslant \{\mu_t(A);\ t \in \tau\} \leqslant \nu_M(A)$ for each $A \in C(k)$. Then $\nu_m(\cdot)$ is called a *stochastic minorant*, and $\nu_M(\cdot)$ a *stochastic majorant*, for the ensemble $\{\mu_t(\cdot);\ t \in \tau\}$. For further details see Jensen (1993).

Standard distributions on $\mathbb{R}_+^1$ include the cdf $G(t; m)$ of the central chi-squared (χ^2) distribution having m degrees of freedom, and the cdf $F(t; r, s)$ of the central Snedecor-Fisher F-distribution having (r, s) degrees of freedom. Further classes arise as follows. Suppose that $\mathscr{L}(\mathbf{U}) = N_k(\mathbf{0}, \mathbf{I}_k)$; let $Z = a_1 U_1^2 + \cdots + a_k U_k^2$; denote by $F_k(z; a_1, \ldots, a_k)$ its cdf; and generate the class $F(k) = \{F_k(\cdot; a_1, \ldots, a_k);\ [a_1, \ldots, a_k] \in \mathbb{R}^k(\mathbf{0})\}$ of all such cdf's on $\mathbb{R}^1$ as $\mathbf{a} = [a_1, \ldots, a_k]$ ranges over $\mathbb{R}^k(\mathbf{0})$. In particular, $F_0(k) \subset F(k)$ is the subclass of distributions on $\mathbb{R}_+^1$ as $\mathbf{a}$ ranges over $\mathbb{R}_+^k(\mathbf{0})$. Series expansions and other properties of distributions in $F_0(k)$ are developed in Johnson and Kotz (1970), and expressions in closed form are given in Bock and Solomon (1988) for the case $k = 2$; see also Mathai and Provost (1992). A basic property is that distributions in $F(k)$ increase stochastically in each $\{a_i;\ 1 \leqslant i \leqslant k\}$, i.e., for each fixed z, $F_k(z; a_1, \ldots, a_k)$ is a decreasing function of each $\{a_i;\ 1 \leqslant i \leqslant k\}$, so that $F_k(z; a_1, \ldots, a_k)$ is order-reversing when considered as a function on $(\mathbb{R}^k, \geqslant)$ with z fixed. Connections to standard distributions follow on noting that $F_k(t/a; a, \ldots, a) = G(t; k)$ for $F_k(z; a, \ldots, a) \in F_0(k)$, with $a > 0$.

3. BOUNDS FOR NORMS

We construct bounds on ratios of norms, and minimal and maximal norms, for pairs $(\|\cdot\|_A, \|\cdot\|_B)$, including (1) $\mathbf{A} \succcurlyeq_L \mathbf{B}$, (2) $\mathbf{B} \succcurlyeq_L \mathbf{A}$, and (3) neither $\mathbf{A} \succcurlyeq_L \mathbf{B}$ nor $\mathbf{B} \succcurlyeq_L \mathbf{A}$, as special cases. Our principal findings follow, where some of R_1, R_2, R_3 may be empty as noted.

THEOREM 1. *Let* $\|\cdot\|_A$ *and* $\|\cdot\|_B$ *be quadratic norms on* $\mathbb{R}^k$, *and let* $\{\gamma_1 \geqslant \cdots \geqslant \gamma_k > 0\}$ *be the ordered roots of* $|\mathbf{A} - \gamma\mathbf{B}| = 0$. *Then the following bounds apply whenever the indicated subspaces are nonempty*:

(i) $\gamma_k^{1/2} \leqslant \|\mathbf{x}\|_A/\|\mathbf{x}\|_B \leqslant \gamma_1^{1/2}$ *for every* $\mathbf{x} \in \mathbb{R}^k$;
(ii) $\gamma_r^{1/2} \leqslant \|\mathbf{x}\|_A/\|\mathbf{x}\|_B \leqslant \gamma_1^{1/2}$ *for every* $\mathbf{x} \in R_1 = \mathrm{Sp}(\mathbf{q}_1, \dots, \mathbf{q}_r)$;
(iii) $\|\mathbf{x}\|_A/\|\mathbf{x}\|_B = 1$ *for every* $\mathbf{x} \in R_2 = \mathrm{Sp}(\mathbf{q}_{r+1}, \dots, \mathbf{q}_{r+s})$; *and*
(iv) $\gamma_k^{1/2} \leqslant \|\mathbf{x}\|_A/\|\mathbf{x}\|_B \leqslant \gamma_{r+s+1}^{1/2}$ *for every* $\mathbf{x} \in R_3 = \mathrm{Sp}(\mathbf{q}_{r+s+1}, \dots, \mathbf{q}_k)$.

Proof. Conclusion (i) restates standard variational properties of the generalized Rayleigh quotient $\mathbf{x}'\mathbf{A}\mathbf{x}/\mathbf{x}'\mathbf{B}\mathbf{x}$. As proofs for (ii)–(iv) run parallel, we focus on conclusion (ii). Beginning with $\mathbf{B}^{-1/2}\mathbf{A}\mathbf{B}^{-1/2} = \mathbf{P}\mathbf{D}_\gamma\mathbf{P}' = \sum_{i=1}^k \gamma_i \mathbf{p}_i\mathbf{p}_i'$, let $L_1 = \mathrm{Sp}(\mathbf{p}_1, \dots, \mathbf{p}_r)$, $L_2 = \mathrm{Sp}(\mathbf{p}_{r+1}, \dots, \mathbf{p}_{r+s})$, and $L_3 = \mathrm{Sp}(\mathbf{p}_{r+s+1}, \dots, \mathbf{p}_k)$, corresponding to (2.1). If $\mathbf{u} \in L_1$ has unit length, then $\mathbf{u}$ can be represented as $\mathbf{u} = c_1\mathbf{p}_1 + \cdots + c_r\mathbf{p}_r$ such that $\mathbf{c}'\mathbf{c} = 1$ with $\mathbf{c}' = [c_1, \dots, c_r]$. We now infer that $\mathbf{u}'\mathbf{B}^{-1/2}\mathbf{A}\mathbf{B}^{-1/2}\mathbf{u} = (\sum_{i=1}^r c_i\mathbf{p}_i)'(\sum_{j=1}^k \gamma_j\mathbf{p}_j\mathbf{p}_j')(\sum_{i=1}^r c_i\mathbf{p}_i) = \sum_{i=1}^r c_i^2\gamma_i$ from the orthonormality of $\mathbf{P}$. Clearly we have $\gamma_r \leqslant \mathbf{u}'\mathbf{B}^{-1/2}\mathbf{A}\mathbf{B}^{-1/2}\mathbf{u} \leqslant \gamma_1$; moreover, equality is achieved on the right at $\mathbf{c}' = [1, 0, \dots, 0]$ and on the left at $\mathbf{c}' = [0, \dots, 0, 1]$. Conclusion (ii) now follows because the spanning vectors are related through $\{\mathbf{q}_i = \mathbf{B}^{-1/2}\mathbf{p}_i;\ 1 \leqslant i \leqslant k\}$ as noted, so that $\mathrm{Sp}(\mathbf{p}_1, \dots, \mathbf{p}_r) = \mathrm{Sp}(\mathbf{q}_1, \dots, \mathbf{q}_r)$. Conclusions (iii) and (iv) follow similarly, to complete our proof. ∎

To construct minimal and maximal norms for $(\|\cdot\|_A, \|\cdot\|_B)$, and eventually for ensembles of such norms, define $\eta_m(\mathbf{x}) = [\mathbf{x}'(\mathbf{A} \wedge \mathbf{B})\mathbf{x}]^{1/2}$ and $\eta_M(\mathbf{x}) = [\mathbf{x}'(\mathbf{A} \vee \mathbf{B})\mathbf{x}]^{1/2}$. The following result is basic.

THEOREM 2. *Let* $\mathbf{A} \wedge \mathbf{B}$ *and* $\mathbf{A} \vee \mathbf{B}$ *be the spectral glb and lub for* $(\mathbf{A}, \mathbf{B})$ *in* $(S_k^+, \succcurlyeq_L)$, *and consider the norms* $\eta_m(\cdot) = \|\cdot\|_{\mathbf{A} \wedge \mathbf{B}}$ *and* $\eta_M(\cdot) = \|\cdot\|_{\mathbf{A} \vee \mathbf{B}}$. *Then the following bounds apply*:

(i) $\eta_m(\mathbf{x}) \leqslant \{\|\mathbf{x}\|_A, \|\mathbf{x}\|_B\} \leqslant \eta_M(\mathbf{x})$ *for each* $\mathbf{x} \in \mathbb{R}^k$;
(ii) $\eta_m(\mathbf{x}) = \|\mathbf{x}\|_B$ *and* $\|\mathbf{x}\|_A = \eta_M(\mathbf{x})$ *for each* $\mathbf{x} \in R_1$;
(iii) $\eta_m(\mathbf{x}) = \|\mathbf{x}\|_B = \|\mathbf{x}\|_A = \eta_M(\mathbf{x})$ *for each* $\mathbf{x} \in R_2$; *and*
(iv) $\eta_m(\mathbf{x}) = \|\mathbf{x}\|_A$ *and* $\|\mathbf{x}\|_B = \eta_M(\mathbf{x})$ *for each* $\mathbf{x} \in R_3$.

Proof. Conclusion (i) follows constructively, since $\|\mathbf{x}\|_G \geqslant \|\mathbf{x}\|_H$ for every $\mathbf{x} \in \mathbb{R}^k$ if and only if $\mathbf{G} \succcurlyeq_L \mathbf{H}$ on $(S_k^+, \succcurlyeq_L)$. As arguments supporting conclusions (ii)–(iv) are similar, we focus on (ii), inferring as in the proof for Theorem 1(ii) that $\mathbf{x} \in R_1$ if and only if $\mathbf{x} = c_1\mathbf{q}_1 + \cdots + c_r\mathbf{q}_r = \mathbf{B}^{-1/2}(c_1\mathbf{p}_1 + \cdots + c_r\mathbf{p}_r)$ for some $\mathbf{c}' = [c_1, \ldots, c_r]$. It follows from the orthonormality of $\{\mathbf{p}_1, \ldots, \mathbf{p}_k\}$ that for $\mathbf{x} \in R_1$, $\mathbf{x}'(\mathbf{A} \wedge \mathbf{B})\mathbf{x}$ can be written as

$$\begin{aligned} \mathbf{x}'(\mathbf{A} \wedge \mathbf{B})\mathbf{x} &= \left(\sum_{i=1}^{r} c_i\mathbf{p}_i\right)' \mathbf{B}^{-1/2}\mathbf{B}^{1/2}\mathbf{P}(\mathbf{D}_\gamma \wedge \mathbf{I}_k)\mathbf{P}'\mathbf{B}^{1/2}\mathbf{B}^{-1/2}\left(\sum_{i=1}^{r} c_i\mathbf{p}_i\right) \\ &= \sum_{i=1}^{r} (\gamma_i \wedge 1)c_i^2 = \sum_{i=1}^{r} c_i^2 = \mathbf{x}'\mathbf{B}\mathbf{x} \end{aligned} \tag{3.1}$$

to prove the first assertion of conclusion (ii). The second assertion follows similarly with $\mathbf{x}'(\mathbf{A} \vee \mathbf{B})\mathbf{x} = \sum_{i=1}^{r} c_i^2\gamma_i = \mathbf{x}'\mathbf{A}\mathbf{x}$ for $\mathbf{x} \in R_1$. Parallel arguments establish conclusions (iii) and (iv), to complete our proof. ∎

We have actually proved much more. Let $\{\boldsymbol{\Gamma}(t);\ t \in \tau\}$ be a bounded set in $(S_k^+, \succcurlyeq_L)$ such that $\mathbf{0} \prec_L \boldsymbol{\Gamma}_m \preccurlyeq_L \boldsymbol{\Gamma}(t) \preccurlyeq_L \boldsymbol{\Gamma}_M$ for every $t \in \tau$. If we now generate the corresponding ensemble $\{\|\cdot\|_{\boldsymbol{\Gamma}(t)};\ t \in \tau\}$ of norms on $\mathbb{R}^k$, then without further proof we have the following bounds.

COROLLARY 2.1. *Let* $\{\|\cdot\|_{\boldsymbol{\Gamma}(t)};\ t \in \tau\}$ *be an ensemble of norms on* $\mathbb{R}^k$ *such that* $\mathbf{0} \prec_L \boldsymbol{\Gamma}_m \preccurlyeq_L \{\boldsymbol{\Gamma}(t);\ t \in \tau\} \preccurlyeq_L \boldsymbol{\Gamma}_M$, *and let* $\eta_m(\cdot) = \|\cdot\|_{\boldsymbol{\Gamma}_m}$ *and* $\eta_M(\cdot) = \|\cdot\|_{\boldsymbol{\Gamma}_M}$. *Then the inequalities*

$$\eta_m(\mathbf{x}) \leqslant \left\{\|\mathbf{x}\|_{\boldsymbol{\Gamma}(t)};\ t \in \tau\right\} \leqslant \eta_M(\mathbf{x}) \tag{3.2}$$

hold uniformly for every $\mathbf{x} \in \mathbb{R}^k$.

4. STOCHASTIC BOUNDS

Our developments support stochastic bounds for distributions of quadratic forms in Gaussian ensembles and mixtures. These arise in linear statistical inference, in large-sample theory, in tests for categorical data, in the ballistic analysis of weapons systems, in the detection of signals from noise, in the study of bone lengths determined *in vivo* using x-ray stereography, and elsewhere. For further details and references see Johnson and Kotz (1970), Jensen and Solomon (1972, 1994), and Mathai and Provost (1992), for example.

We next construct bounds for certain cdf's in $F_0(k)$ in terms of other members of the class. Consider the forms $\{\mathbf{Y}'\mathbf{A}\mathbf{Y}, \mathbf{Y}'\mathbf{B}\mathbf{Y}\}$ such that $\mathscr{L}(\mathbf{Y}) = N_k(\mathbf{0}, \mathbf{\Sigma})$ with $\mathbf{\Sigma} \in S_k^+$; equivalently consider $\{V = \mathbf{U}'\mathbf{\Xi}\mathbf{U}, W = \mathbf{U}'\mathbf{\Omega}\mathbf{U}\}$ such that $\mathbf{\Xi} = \mathbf{\Sigma}^{1/2}\mathbf{A}\mathbf{\Sigma}^{1/2}$, $\mathbf{\Omega} = \mathbf{\Sigma}^{1/2}\mathbf{B}\mathbf{\Sigma}^{1/2}$, and $\mathscr{L}(\mathbf{U}) = N_k(\mathbf{0}, \mathbf{I}_k)$; and observe that $\mathscr{L}(V) = \mathscr{L}(\xi_1 U_1^2 + \cdots + \xi_k U_k^2)$ and $\mathscr{L}(W) = \mathscr{L}(\omega_1 U_1^2 + \cdots + \omega_k U_k^2)$, with $\xi_1 \geqslant \cdots \geqslant \xi_k > 0$ and $\omega_1 \geqslant \cdots \geqslant \omega_k > 0$ as the ordered eigenvalues of $\mathbf{\Xi}$ and $\mathbf{\Omega}$, respectively. Accordingly, designate their cdf's as $F_V(\cdot; \xi_1, \ldots, \xi_k)$ and $F_W(\cdot; \omega_1, \ldots, \omega_k)$, and observe that both belong to $F_0(k)$. Finally let $\boldsymbol{\delta}_m$ and $\boldsymbol{\delta}_M$ be vectors in $\mathbb{R}^k$ comprising the ordered eigenvalues $\{\delta_{m1} \geqslant \cdots \geqslant \delta_{mk} > 0\}$ and $\{\delta_{M1} \geqslant \cdots \geqslant \delta_{Mk} > 0\}$ of $\mathbf{\Xi} \wedge \mathbf{\Omega}$ and $\mathbf{\Xi} \vee \mathbf{\Omega}$, respectively. The following envelopes for cdf's in $F_0(k)$ are basic.

THEOREM 3. *Consider definite forms* $V = \mathbf{Y}'\mathbf{A}\mathbf{Y}$ *and* $W = \mathbf{Y}'\mathbf{B}\mathbf{Y}$ *such that* $\mathscr{L}(\mathbf{Y}) = N_k(\mathbf{0}, \mathbf{\Sigma})$; *let* $\boldsymbol{\gamma} = [\gamma_1, \ldots, \gamma_k]$ *be the ordered roots of* $|\mathbf{A} - \gamma\mathbf{B}| = 0$; *and let* $\boldsymbol{\xi} = [\xi_1, \ldots, \xi_k]$, $\boldsymbol{\omega} = [\omega_1, \ldots, \omega_k]$, $\boldsymbol{\delta}_m = [\delta_{m1}, \ldots, \delta_{mk}]$, *and* $\boldsymbol{\delta}_M = [\delta_{M1}, \ldots, \delta_{Mk}]$ *comprise the ordered eigenvalues of* $\mathbf{\Xi} = \mathbf{\Sigma}^{1/2}\mathbf{A}\mathbf{\Sigma}^{1/2}$, $\mathbf{\Omega} = \mathbf{\Sigma}^{1/2}\mathbf{B}\mathbf{\Sigma}^{1/2}$, $\mathbf{\Xi} \wedge \mathbf{\Omega}$, *and* $\mathbf{\Xi} \vee \mathbf{\Omega}$, *respectively. Then the following bounds apply.*

(i) $G(t/\omega_1\gamma_1; k) \leqslant F_W(t/\gamma_1; \omega) \leqslant F_V(t; \xi) \leqslant F_W(t/\gamma_k; \omega) \leqslant G(t/\omega_k\gamma_k; k)$ *for every* $t \in \mathbb{R}_+^1$;

(ii) $F_k(t; \omega \vee \xi) \leqslant \{F_V(t; \xi), F_W(t; \omega)\} \leqslant F_k(t; \omega \wedge \xi)$ *for every* $t \in \mathbb{R}_+^1$; *and*

(iii) $G(t/\delta_{M1}; k) \leqslant F_k(t; \delta_M) \leqslant \{F_V(t; \xi), F_W(t; \omega)\} \leqslant F_k(t; \delta_m) \leqslant G(t/\delta_{mk}; k)$ *for every* $t \in \mathbb{R}_+^1$.

Proof. Let $\mathbf{U} = \mathbf{\Sigma}^{-1/2}\mathbf{Y}$, so that $\mathscr{L}(\mathbf{U}) = N_k(\mathbf{0}, \mathbf{I}_k)$. To see conclusion (i), use variational properties of the generalized Rayleigh quotient $\mathbf{z}'\mathbf{\Xi}\mathbf{z}/\mathbf{z}'\mathbf{\Omega}\mathbf{z}$ as in Theorem 1(i) to infer that $\gamma_k\mathbf{z}'\mathbf{\Omega}\mathbf{z} \leqslant \mathbf{z}'\mathbf{\Xi}\mathbf{z} \leqslant \gamma_1\mathbf{z}'\mathbf{\Omega}\mathbf{z}$ holds pointwise in $\mathbf{z} \in \mathbb{R}^k$. Then $P(\mathbf{U}'\mathbf{\Xi}\mathbf{U} \leqslant t) \leqslant P(\mathbf{U}'\mathbf{\Omega}\mathbf{U} \leqslant t/\gamma_k)$ follows by inclusion, so that $F_V(t; \boldsymbol{\xi}) \leqslant F_W(t/\gamma_k; \boldsymbol{\omega})$, and similarly for $F_W(t/\gamma_1; \boldsymbol{\omega}) \leqslant F_V(t; \boldsymbol{\xi})$, giving the three inner inequalities of conclusion (i). The outer inequalities on the left and right of (i) follow on replacing $\boldsymbol{\omega}$ by $\omega_1\mathbf{1}_k$ and $\omega_k\mathbf{1}_k$, respectively, then using the monotonicity of $F_W(t; \omega_1, \ldots, \omega_k)$ and the fact that $F_W(t; \omega, \ldots, \omega) = G(t/\omega; k)$, to complete a proof for conclusion (i). Conclusion (ii) follows directly from the monotonicity of $F_k(t; a_1, \ldots, a_k)$, together with the ordering $\boldsymbol{\omega} \wedge \boldsymbol{\xi} \leqslant \{\boldsymbol{\omega}, \boldsymbol{\xi}\} \leqslant \boldsymbol{\omega} \vee \boldsymbol{\xi}$ on $(\mathbb{R}_+^k, \geqslant)$. Conclusion (iii) follows on noting that $\mathbf{z}'(\mathbf{\Xi} \wedge \mathbf{\Omega})\mathbf{z} \leqslant \{\mathbf{z}'\mathbf{\Xi}\mathbf{z}, \mathbf{z}'\mathbf{\Omega}\mathbf{z}\} \leqslant \mathbf{z}'(\mathbf{\Xi} \vee \mathbf{\Omega})\mathbf{z}$ holds pointwise for $\mathbf{z} \in \mathbb{R}^k$, and then proceeding by inclusion as in the proof for (i). That $P(\mathbf{U}'(\mathbf{\Xi} \wedge \mathbf{\Omega})\mathbf{U} \leqslant t) = F_k(t; \delta_m)$ follows as usual on diagonalizing $\mathbf{\Xi} \wedge \mathbf{\Omega}$, and similarly that $P(\mathbf{U}'(\mathbf{\Xi} \vee \mathbf{\Omega})\mathbf{U} \leqslant t) = F_k(t; \delta_M)$. The outer inequalities in (iii) follow from steps used in the proof for (i), to complete our proof. ∎

Theorem 3 as stated pertains to different quadratic forms defined on the same probability space, namely, $N_k(\mathbf{0}, \boldsymbol{\Sigma})$. A dual problem considers a given quadratic form under different Gaussian measures as follows. Let $V = \mathbf{Y}'\mathbf{C}\mathbf{Y}$ and $W = \mathbf{Z}'\mathbf{C}\mathbf{Z}$ such that $\mathscr{L}(\mathbf{Y}) = N_k(\mathbf{0}, \boldsymbol{\Psi})$ and $\mathscr{L}(\mathbf{Z}) = N_k(\mathbf{0}, \boldsymbol{\Gamma})$. Using routine arguments we now infer that $\mathscr{L}(V) = \mathscr{L}(\mathbf{U}'\boldsymbol{\Xi}\mathbf{U})$ and $\mathscr{L}(W) = \mathscr{L}(\mathbf{U}'\boldsymbol{\Omega}\mathbf{U})$, with $\boldsymbol{\Xi} = \boldsymbol{\Psi}^{1/2}\mathbf{C}\boldsymbol{\Psi}^{1/2}$ and $\boldsymbol{\Omega} = \boldsymbol{\Gamma}^{1/2}\mathbf{C}\boldsymbol{\Gamma}^{1/2}$, such that $\mathscr{L}(\mathbf{U}) = N_k(\mathbf{0}, \mathbf{I}_k)$. This is in the form of Theorem 3 with $\boldsymbol{\Sigma} = \mathbf{I}_k$, where $\mathbf{A} = \boldsymbol{\Psi}^{1/2}\mathbf{C}\boldsymbol{\Psi}^{1/2} = \boldsymbol{\Xi}$ and $\mathbf{B} = \boldsymbol{\Gamma}^{1/2}\mathbf{C}\boldsymbol{\Gamma}^{1/2} = \boldsymbol{\Omega}$. Theorem 3 now applies verbatim with $\{\gamma_1 \geqslant \cdots \geqslant \gamma_k\}$ as the ordered roots of $|\mathbf{A} - \gamma\mathbf{B}| = 0 = |\boldsymbol{\Psi}^{1/2}\mathbf{C}\boldsymbol{\Psi}^{1/2} - \gamma\boldsymbol{\Gamma}^{1/2}\mathbf{C}\boldsymbol{\Gamma}^{1/2}|$.

Theorem 3(iii) may be extended to include Gaussian ensembles on $\mathbb{R}^k$ and mixtures over these. Let $\{G_k(\cdot; \boldsymbol{\Gamma}(t)); t \in \tau\}$ consist of zero-mean Gaussian measures on $\mathbb{R}^k$ bounded in dispersion such that $\boldsymbol{\Gamma}_m \preccurlyeq_L \{\boldsymbol{\Gamma}(t); t \in \tau\} \preccurlyeq_L \boldsymbol{\Gamma}_M$ for some $(\boldsymbol{\Gamma}_m, \boldsymbol{\Gamma}_M)$ in S_k^+, and let $M(\tau)$ be the collection of probability measures over τ. Under measurability conditions we consider mixtures of the type

$$\nu(\cdot; \mu) = \int_\tau G_k(\cdot; \boldsymbol{\Gamma}(t))\, d\mu(t) \tag{4.1}$$

over $\{G_k(\cdot; \boldsymbol{\Gamma}(t)); t \in \tau\}$ with mixing distribution $\mu(\cdot) \in M(\tau)$.

For fixed $\mathbf{A} \in S_k^+$, let $\boldsymbol{\alpha}(t) = [\alpha_1(t), \ldots, \alpha_k(t)]$ comprise the ordered roots of $|\mathbf{A} - \alpha(t)[\boldsymbol{\Gamma}(t)]^{-1}| = 0$ for each $t \in \tau$, and let $V(t) = \mathbf{Y}'\mathbf{A}\mathbf{Y}$ be such that $\mathscr{L}(\mathbf{Y}) = N_k(\mathbf{0}, \boldsymbol{\Gamma}(t))$. We infer as before that the cdf of $V(t)$ is $F_k(\nu; \boldsymbol{\alpha}(t))$ for each $t \in \tau$. For the case that $\mathscr{L}(\mathbf{Y})$ is a mixture as in (4.1), we argue conditionally to conclude that the unconditional cdf of $V = \mathbf{Y}'\mathbf{A}\mathbf{Y}$ is the mixture

$$F(\nu; \mu) = \int_\tau F_k(\nu; \boldsymbol{\alpha}(t))\, d\mu(t). \tag{4.2}$$

Bounds for the ensemble $\{F_k(\nu; \boldsymbol{\alpha}(t)); t \in \tau\}$, and for mixtures over these of the type $\{F(\nu; \mu); \mu \in M(\tau)\}$ as in (4.2), are considered next. To these ends let $\boldsymbol{\alpha}(t) = [\alpha_1(t), \ldots, \alpha_k(t)]$ comprise the ordered roots of $|\mathbf{A} - \alpha(t)[\boldsymbol{\Gamma}(t)]^{-1}| = 0$ as before, and let $\boldsymbol{\delta}_m = [\delta_{m1}, \ldots, \delta_{mk}]$ and $\boldsymbol{\delta}_M = [\delta_{M1}, \ldots, \delta_{Mk}]$ comprise the ordered roots of $|\mathbf{A} - \delta[\boldsymbol{\Gamma}_m]^{-1}| = 0$ and $|\mathbf{A} - \delta[\boldsymbol{\Gamma}_M]^{-1}| = 0$, respectively. The following bounds are basic.

THEOREM 4. *Consider the definite form* $V(t) = \mathbf{Y}'\mathbf{A}\mathbf{Y}$ *with* $\mathscr{L}(\mathbf{Y}) \in \{N_k(\mathbf{0}, \boldsymbol{\Gamma}(t)); t \in \tau\}$ *such that* $\boldsymbol{\Gamma}_m \preccurlyeq_L \{\boldsymbol{\Gamma}(t); t \in \tau\} \preccurlyeq_L \boldsymbol{\Gamma}_M$ *for some* $(\boldsymbol{\Gamma}_m, \boldsymbol{\Gamma}_M)$ *in* S_k^+, *and let* $\{F_k(\nu; \boldsymbol{\alpha}(t)); t \in \tau\}$ *be the corresponding ensemble of cdf's for* $\{\mathscr{L}(V(t)); t \in \tau\}$.

(i) *Then bounds for the ensemble of cdf's for* $\{\mathscr{L}(V(t));\, t \in \tau\}$ *are given by*

$$F_k(\nu; \boldsymbol{\delta}_M) \leqslant \{F_k(\nu; \boldsymbol{\alpha}(t));\, t \in \tau\} \leqslant F_k(\nu; \boldsymbol{\delta}_m) \tag{4.3}$$

for each $\nu \in \mathbb{R}^1_+$.

(ii) *Under mixtures of type* (4.1) *yielding cdf's of type* (4.2), *bounds for the ensemble of cdf's for* $\mathscr{L}(V)$ *are given by*

$$F_k(\nu; \boldsymbol{\delta}_M) \leqslant \{F(\nu; \mu);\, \mu(\cdot) \in M(\tau)\} \leqslant F_k(\nu; \boldsymbol{\delta}_m) \tag{4.4}$$

for each $\nu \in \mathbb{R}^1_+$ *and* $\mu(\cdot) \in M(\tau)$.

Proof. Given two Gaussian measures $G_k(\cdot; \boldsymbol{\Xi})$ and $G_k(\cdot; \boldsymbol{\Omega})$ on $\mathbb{R}^k$, it is known that $G_k(A; \boldsymbol{\Xi}) \geqslant G_k(A; \boldsymbol{\Omega})$ for every $A \in C(k)$ if and only if $\boldsymbol{\Omega} \succcurlyeq_L \boldsymbol{\Xi}$. For sufficiency see Anderson (1955), and for necessity see Jensen (1984). It follows directly from the boundedness of $\{\boldsymbol{\Gamma}(t);\, t \in \tau\}$ that $G_k(\cdot; \boldsymbol{\Gamma}_m)$ is a Gaussian majorant, and $G_k(\cdot; \boldsymbol{\Gamma}_M)$ is a Gaussian minorant, for the bounded ensemble $\{G_k(\cdot; \boldsymbol{\Gamma}(t));\, t \in \tau\}$. The set $A(\nu) = \{\mathbf{Y} \in \mathbb{R}^k;\, \mathbf{Y}'\mathbf{A}\mathbf{Y} \leqslant \nu\}$ clearly belongs to $C(k)$ for each $\nu \geqslant 0$, since $\mathbf{A}$ is positive definite, whereas the cdf's for $V = \mathbf{Y}'\mathbf{A}\mathbf{Y}$ under various Gaussian measures are given by $F_k(\nu; \boldsymbol{\alpha}(t)) = G_k(A(\nu); \boldsymbol{\Gamma}(t))$, $F_k(\nu; \boldsymbol{\delta}_m) = G_k(A(\nu); \boldsymbol{\Gamma}_m)$, and $F_k(\nu; \boldsymbol{\delta}_M) = G_k(A(\nu); \boldsymbol{\Gamma}_M)$. Conclusion (i) follows immediately. Since these bounds hold pointwise for each $t \in \tau$ under the integral (4.2) independently of $\mu(\cdot) \in M(\tau)$, conclusion (ii) now follows also, to complete our proof. ■

5. APPLICATIONS

We next consider selected topics in applications.

5.1. *Classification*

An object $\mathbf{x} \in \mathbb{R}^k$ is to be assigned to one of two populations $P_{\boldsymbol{\Xi}}$ and $P_{\boldsymbol{\Omega}}$ characterized as $N_k(\boldsymbol{\mu}_1, \boldsymbol{\Xi})$ and $N_k(\boldsymbol{\mu}_2, \boldsymbol{\Omega})$, having prior probabilities π_1 and π_2, respectively. Using Mahalanobis (1936) distances, the standard rule assigns $\mathbf{x}$ to $P_{\boldsymbol{\Xi}}$ whenever

$$\Delta^2_{\boldsymbol{\Omega}}(\mathbf{x}, \boldsymbol{\mu}_2) - \Delta^2_{\boldsymbol{\Xi}}(\mathbf{x}, \boldsymbol{\mu}_1) > c(\boldsymbol{\pi}, \boldsymbol{\gamma}) \tag{5.1}$$

and to $P_{\boldsymbol{\Omega}}$ otherwise, where $\boldsymbol{\pi} = [\pi_1, \pi_2]$ and $c(\boldsymbol{\pi}, \boldsymbol{\gamma}) = \ln(\pi_2|\boldsymbol{\Xi}|/\pi_1|\boldsymbol{\Omega}|) = \ln(\pi_2|\mathbf{D}_{\boldsymbol{\gamma}}|/\pi_1)$, with $\mathbf{D}_{\boldsymbol{\gamma}} = \mathrm{Diag}(\gamma_1, \ldots, \gamma_k)$ containing the ordered roots of $|\boldsymbol{\Xi} - \gamma\boldsymbol{\Omega}| = 0$. It is instructive to examine the difference $[\Delta^2_{\boldsymbol{\Omega}}(\mathbf{x}, \boldsymbol{\mu}_2) - \Delta^2_{\boldsymbol{\Xi}}(\mathbf{x}, \boldsymbol{\mu}_1)]$ as $\mathbf{x}$ varies over $\mathbb{R}^k$. For further details regarding classification problems of these types, see McLachlan (1992), especially Section 3.2.

We examine properties of the foregoing rule under one or more of the conditions $C_1 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$, $C_2 : \pi_1 = \pi_2$, and $C_3 : |\boldsymbol{\Xi}| = |\boldsymbol{\Omega}|$, corresponding respectively to equal means, equal prior probabilities, and equimodular dispersion matrices. Classification under C_1 is called the zero-mean discrimination problem, since $\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2 = \mathbf{0}$. In particular, under C_1–C_3 we see that $c(\boldsymbol{\pi}, \boldsymbol{\gamma}) = 0$, so that classification then rests exclusively on the Mahalanobis (1936) distances from $\mathbf{x}$ to the common centroid in $\mathbb{R}^k$ of the two populations.

Now assume condition $C_1 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \boldsymbol{\mu}$, say, and let $Q(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Xi}, \boldsymbol{\Omega}) = \Delta^2_{\boldsymbol{\Omega}}(\mathbf{x}, \boldsymbol{\mu}) - \Delta^2_{\boldsymbol{\Xi}}(\mathbf{x}, \boldsymbol{\mu})$. To examine its behavior as $\mathbf{x}$ varies, write $Q(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Xi}, \boldsymbol{\Omega}) = \Delta^2_{\boldsymbol{\Omega}}(\mathbf{x}, \boldsymbol{\mu})[1 - \Delta^2_{\boldsymbol{\Xi}}(\mathbf{x}, \boldsymbol{\mu})/\Delta^2_{\boldsymbol{\Omega}}(\mathbf{x}, \boldsymbol{\mu})]$, then apply Theorem 1 to infer that bounds of the type

$$\Delta^2_{\boldsymbol{\Omega}}(\mathbf{x}, \boldsymbol{\mu})(1 - \gamma_U) \leqslant Q(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Xi}, \boldsymbol{\Omega}) \leqslant \Delta^2_{\boldsymbol{\Omega}}(\mathbf{x}, \boldsymbol{\mu})(1 - \gamma_L) \quad (5.2)$$

apply with (γ_L, γ_U) and $\mathbf{x} \in \mathbb{R}^k$ chosen suitably. In particular, Theorem 1(i) applies uniformly for every $\mathbf{x} \in \mathbb{R}^k$ with $(\gamma_L, \gamma_U) = (\gamma_k, \gamma_1)$. Whether $\mathbf{x}$ is closer to $P_{\boldsymbol{\Xi}}$ or to $P_{\boldsymbol{\Omega}}$ in their respective metrics depends on where $\mathbf{x}$ falls in $\mathbb{R}^k$. If $\mathbf{x} \in R_1$, Theorem 1(ii) shows that (5.2) applies uniformly for $\mathbf{x} \in R_1$ with $(\gamma_L, \gamma_U) = (\gamma_r, \gamma_1)$. The difference is negative, since both bounds are negative, so that $\mathbf{x} \in R_1$ is closer to $P_{\boldsymbol{\Omega}}$ than to $P_{\boldsymbol{\Xi}}$. In a similar manner Theorem 1(iv) shows that (5.2) holds uniformly for $\mathbf{x} \in R_3$ with $(\gamma_L, \gamma_U) = (\gamma_k, \gamma_{r+s+1})$. Since both bounds are positive, we conclude that $\mathbf{x} \in R_3$ is closer to $P_{\boldsymbol{\Xi}}$ than to $P_{\boldsymbol{\Omega}}$. Finally, Theorem 1(iii) asserts that $\mathbf{x} \in R_2$ is equidistant from $P_{\boldsymbol{\Xi}}$ and $P_{\boldsymbol{\Omega}}$ for any $\mathbf{x} \in R_2$. In practice, such outcomes of themselves cannot distinguish between the populations $P_{\boldsymbol{\Xi}}$ and $P_{\boldsymbol{\Omega}}$ under the condition $C_1 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$.

We next study problems of misclassification. Let $P(1|2)$ be the probability of misclassifying $\mathbf{x}$ into $P_{\boldsymbol{\Xi}}$ when in fact $\mathbf{x} \in P_{\boldsymbol{\Omega}}$, and similarly let $P(2|1)$ be the probability of misclassifying $\mathbf{x} \in P_{\boldsymbol{\Xi}}$ into $P_{\boldsymbol{\Omega}}$. It suffices to consider a canonical form through an elementary transformation $\mathbf{x} \to \mathbf{z}$ taking $P_{\boldsymbol{\Xi}}$ and $P_{\boldsymbol{\Omega}}$ into $P_{\mathbf{D}_\gamma}$ and $P_{\mathbf{I}_k}$ as characterized by $N_k(\boldsymbol{\theta}_1, \mathbf{D}_\gamma)$ and $N_k(\boldsymbol{\theta}_2, \mathbf{I}_k)$, respectively, where $\mathbf{D}_\gamma = \mathrm{Diag}(\gamma_1, \ldots, \gamma_k)$ contains the ordered roots of $|\boldsymbol{\Xi} - \gamma\boldsymbol{\Omega}| = 0$ as before. Under condition C_1 where $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_2 = [\theta_1, \ldots, \theta_k]'$ $= \boldsymbol{\theta}$, say, the expression $Q(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Xi}, \boldsymbol{\Omega})$ in (5.2) transforms into

$$Q(\mathbf{z}; \boldsymbol{\theta}, \boldsymbol{\gamma}) = \sum_{i=1}^{k} \frac{\gamma_i - 1}{\gamma_i}(z_i - \theta_i)^2. \quad (5.3)$$

The classification rule assigning $\mathbf{x}$ to $P_{\boldsymbol{\Xi}}$ equivalently assigns $\mathbf{z}$ to $P_{\mathbf{D}_\gamma}$ whenever $Q(\mathbf{z}; \boldsymbol{\theta}, \boldsymbol{\gamma}) > c(\boldsymbol{\pi}, \boldsymbol{\gamma})$ with $c(\boldsymbol{\pi}, \boldsymbol{\gamma})$ as defined following (5.1). From earlier developments it follows that $P(1|2) = P(Q(\mathbf{z}; \boldsymbol{\theta}, \boldsymbol{\gamma}) > c(\boldsymbol{\pi}, \boldsymbol{\gamma}) \mid$

$\mathbf{V}(\mathbf{z}) = \mathbf{I}_k$) and that $P(2|1) = P(Q(\mathbf{z}; \boldsymbol{\theta}, \boldsymbol{\gamma}) < c(\boldsymbol{\pi}, \boldsymbol{\gamma}) \,|\, \mathbf{V}(\mathbf{z}) = \mathbf{D}_\gamma)$. To proceed we apply distribution theory for quadratic forms in Gaussian variates, starting with the vector $[\alpha_1, \ldots, \alpha_k] = [(\gamma_1 - 1)/\gamma_1, \ldots, (\gamma_k - 1)/\gamma_k]$. Observe that $\{\alpha_1 \geqslant \cdots \geqslant \alpha_k\}$. Generally, as in (2.1), there are r values of $\{\gamma_1 \geqslant \gamma_2 \geqslant \cdots \geqslant \gamma_k\}$ greater than, s values equal to, and $h = k - r - s$ values less than unity. Accordingly, $Q(\mathbf{z}; \boldsymbol{\theta}, \boldsymbol{\gamma})$ is an indefinite form of rank $d = k - s$. Now let $\boldsymbol{\alpha}_1 = [\alpha_1, \ldots, \alpha_r]$; let $\boldsymbol{\alpha}_2 = [\alpha_{r+s+1}, \ldots, \alpha_k]$; and write $\boldsymbol{\alpha} = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2]$. Similarly, with $[\beta_1, \ldots, \beta_k] = [(\gamma_1 - 1), \ldots, (\gamma_k - 1)]$, let $\boldsymbol{\beta}_1 = [\beta_1, \ldots, \beta_r]$; let $\boldsymbol{\beta}_2 = [\beta_{r+s+1}, \ldots, \beta_k]$; and write $\boldsymbol{\beta} = [\boldsymbol{\beta}_1, \boldsymbol{\beta}_2]$. From developments in Section 2.3 it is now clear that

$$P(1|2) = 1 - F_d(c(\boldsymbol{\pi}, \boldsymbol{\gamma}); \boldsymbol{\alpha}) \tag{5.4}$$

and

$$P(2|1) = F_d(c(\boldsymbol{\pi}, \boldsymbol{\gamma}); \boldsymbol{\beta}). \tag{5.5}$$

These cdf's clearly belong to $F(d)$ with $d = k - s$, where s is the number of unit roots in (2.1). Indefinite forms of these types have been studied by Gurland (1955) and others; see Johnson and Kotz (1970) and Mathai and Provost (1992) for further details and references.

Further bounds may be constructed using monotonicity of functions in $F(d)$. Owing to the monotonicity of $F_d(t; \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2)$ as $(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2)$ vary with t fixed, and since $\{\alpha_1 \geqslant \cdots \geqslant \alpha_k\}$, we infer that

$$F_d(t; \alpha_1 \mathbf{1}_r, \alpha_{r+s+1} \mathbf{1}_h) \leqslant F_d(t; \alpha_1, \alpha_2) \leqslant F_d(t; \alpha_r \mathbf{1}_r, \alpha_k \mathbf{1}_h), \tag{5.6}$$

so that lower and upper bounds on $P(1|2)$ are given by

$$1 - F_d(t^*; \alpha_r \mathbf{1}_r, \alpha_k \mathbf{1}_h) \leqslant P(1|2) \leqslant 1 - F_d(t^*; \alpha_1 \mathbf{1}_r, \alpha_{r+s+1} \mathbf{1}_h) \tag{5.7}$$

when evaluated at $t^* = c(\boldsymbol{\pi}, \boldsymbol{\gamma})$. Similar arguments apply for $P(2|1)$. Since the values of β are also ordered as $\beta_1 \geqslant \cdots \geqslant \beta_k$, we obtain bounds on $P(2|1)$ as given by

$$F_d(t^*; \beta_1 \mathbf{1}_r, \beta_{r+s+1} \mathbf{1}_h) \leqslant P(2|1) \leqslant F_d(t^*; \beta_r \mathbf{1}_r, \beta_k \mathbf{1}_h) \tag{5.8}$$

when evaluated at $t^* = c(\boldsymbol{\pi}, \boldsymbol{\gamma})$.

Observe that these bounds all have the structure $P(a_1 \mathbf{U}_1' \mathbf{U}_1 - a_2 \mathbf{U}_2' \mathbf{U}_2 \leqslant t^*)$ or its complement, where $\mathbf{U} = [\mathbf{U}_1', \mathbf{U}_2']'$ such that $\mathbf{U}_1 \in \mathbb{R}^r$, $\mathbf{U}_2 \in \mathbb{R}^h$, and $\mathscr{L}(\mathbf{U}) = N_d(\mathbf{0}, \mathbf{I}_d)$. In what follows we invoke conditions C_1–C_3, under which $c(\boldsymbol{\pi}, \boldsymbol{\gamma}) = 0$. Then we find that $P(a_1\mathbf{U}_1'\mathbf{U}_1 - a_2\mathbf{U}_2'\mathbf{U}_2 \leqslant 0) = P(h\mathbf{U}_1'\mathbf{U}_1/r\mathbf{U}_2'\mathbf{U}_2 \leqslant ha_2/ra_1) = F(ha_2/ra_1;\, r, h)$ using the appropriate Snedecor-Fisher distribution. In short, we have shown that $F_d(0; a_1\mathbf{1}_r, a_2\mathbf{1}_h) = F(ha_2/ra_1; r, h)$. These facts may be substituted into (5.7) and (5.8) to give lower and upper bounds for $P(1|2)$ and $P(2|1)$ in terms of standard F-distributions. In particular, we obtain

$$1 - F(-h\alpha_k/r\alpha_r; r, h) \leqslant P(1|2) \leqslant 1 - F(-h\alpha_{r+s+1}/r\alpha_1; r, h) \quad (5.9)$$

and

$$F(-h\beta_{r+s+1}/r\beta_1; r, h) \leqslant P(2|1) \leqslant F(-h\beta_k/r\beta_r; r, h). \quad (5.10)$$

To illustrate the bounds in (5.9) and (5.10), we consider cases where $k = 5$, $s = 1$, and $r = h = 2$. The natural parameters of the problem are solutions $\boldsymbol{\gamma} = [\gamma_1, \ldots, \gamma_k]$ of the equation $|\boldsymbol{\Xi} - \gamma_i \boldsymbol{\Omega}| = 0$. A variety of choices for these are listed in Table 1 along with lower and upper bounds for the misclassification probabilities $P(1|2)$ and $P(2|1)$ using (5.9) and (5.10). It is clear that the spread between the bounds for $P(1|2)$ narrows as the elements of $[\alpha_1, \ldots, \alpha_r]$ and $[\alpha_{r+s+1}, \ldots, \alpha_k]$ become more homogeneous within each pair of brackets, and similarly for bounds on $P(2|1)$ in terms of the homogeneity of $[\beta_1, \ldots, \beta_r]$ and $[\beta_{r+s+1}, \ldots, \beta_k]$. Moreover, the misclassification probabilities themselves become smaller with greater spread between $[\alpha_1, \ldots, \alpha_r]$ and $[\alpha_{r+s+1}, \ldots, \alpha_k]$ for the case of $P(1|2)$, and between $[\beta_1, \ldots, \beta_r]$ and $[\beta_{r+s+1}, \ldots, \beta_k]$ for the case of $P(2|1)$.

TABLE 1

BOUNDS FOR $P(1|2)$ AND $P(2|1)$ UNDER CONDITIONS C_1–C_3 FOR CASES WHERE $k = 5$, $s = 1$, AND $r = h = 2$

					$P(1\|2)$		$P(2\|1)$	
γ_1	γ_2	γ_3	γ_4	γ_5	Lower	Upper	Lower	Upper
2	2	1	0.50	0.50	0.3333	0.3333	0.3333	0.3333
4	2	1	0.50	0.50	0.3333	0.4286	0.1429	0.3333
4	2	1	0.50	0.40	0.2500	0.4286	0.1429	0.3750
3	2	1	0.50	0.10	0.0526	0.4000	0.2000	0.4737
5	2	1	0.50	0.20	0.1111	0.4444	0.1111	0.4444
6	5	1	0.20	0.10	0.0816	0.1724	0.1379	0.1837
4	4	1	0.10	0.10	0.0769	0.0769	0.2308	0.2308
5	4	1	0.20	0.10	0.0769	0.1667	0.1667	0.2308
7	6	1	0.10	0.05	0.0420	0.0870	0.1304	0.1597

5.2. *Ballistics*

In ballistics let $\mathbf{X} \in \mathbb{R}^3$ be the point of impact of a projectile subject to chance disturbances owing to errors in initial velocity, local turbulence, precipitation, variations in air pressure, topography at point of impact, and other extraneous circumstances. Often $\mathbf{X}$ is modeled as a Gaussian vector having some mean point of impact $\boldsymbol{\mu} \in \mathbb{R}^3$ and dispersion matrix $\boldsymbol{\Xi}$. If r is the effective radius of the weapon on impact and if $B(r)$ is the ball of radius r centered at $\mathbf{0} \in \mathbb{R}^3$, then the probability of a kill is given by $G_3(B(r); \boldsymbol{\Xi})$ on translation to the origin. Details are given in Eckler and Burr (1972), for example. Noting that such models are often overly simplistic, Gilliland (1968) considered more general distributions having convex level sets as in Anderson (1955).

We suppose instead that dispersion characteristics of the trajectory and ultimate point of impact may vary with weather, range, and topography at impact. For example, increasing the range alone may serve to dilate the dispersion matrix when other factors are held fixed. Even if impacts are scattered spherically in a plane normal to the trajectory, the pattern tends to elongate along the path as the plane at impact deviates from normal to the path. Dispersion characteristics thereby may vary with local topography at impact. In short, in practice it often is realistic to suppose that dispersion parameters $\boldsymbol{\Xi}$ belong to some bounded ensemble $\{\boldsymbol{\Gamma}(t);\ t \in \tau\}$. Bounds on kill probabilities for such ensembles derive from Theorem 4 as in (4.3) with $\mathbf{A} = \mathbf{I}_k$. If the actual errors tend to behave stochastically as a mixture over some Gaussian ensemble $\{G_3(\cdot; \boldsymbol{\Gamma}(t));\ t \in \tau\}$ to mimic disturbances over a random environment, then Theorem 4(ii) gives bounds for kill probabilities in such mixtures as in (4.4). Such bounds may be useful in practice in determining whether those probabilities exceed a threshold value germane to assessing further courses of action. It is of interest to note that the Gaussian bounds of Theorem 4 apply for certain mixtures having star-shaped contours, in contrast with distributions having convex contours as in Gilliland (1968). Properties of such mixtures are largely unknown, and no doubt the distributions themselves are quite complicated. Nonetheless, stochastic bounds as in (4.4) apply for all such bounded mixtures.

REFERENCES

Anderson, T. W. 1955. The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities, *Proc. Amer. Math Soc.* 6:170–176.

Bock, M. E. and Solomon, H. 1988. Distributions of quadratic forms, *Austral. J. Statist.* 30A:139–149.

Eckler, A. R. and Burr, S. A. 1972. *Mathematical Models for Target Coverage and Missile Allocation*, Military Operations Research Soc., Washington.
Gilliland, D. C. 1968. On maximization of the integral of a bell-shaped function over a symmetric set, *Naval Res. Logist. Quart.* 15:507–515.
Gurland, J. 1955. Distribution of definite and of indefinite quadratic forms, *Ann. Math. Statist.* 26:122–127.
Halmos, P. R. 1958. *Finite-Dimensional Vector Spaces*, 2nd ed., Van Nostrand, New York.
Jensen, D. R. 1984. Ordering ellipsoidal measures: Scale and peakedness orderings, *SIAM J. Appl. Math.* 44:1226–1231.
Jensen, D. R. 1993. Majorants and minorants for elliptical measures on $\mathbb{R}^k$, *J. Multivariate Anal.* 47:269–282.
Jensen, D. R. and Solomon, H. 1972. A Gaussian approximation to the distribution of a definite quadratic form, *J. Amer. Statist. Assoc.* 67:898–902.
Jensen, D. R. and Solomon, H. 1994. Approximations to joint distributions of definite quadratic forms, *J. Amer. Statist. Assoc.* 89:480–486.
Johnson, N. L. and Kotz, S. 1970. *Distributions in Statistics: Continuous Univariate Distributions*—2, Houghton Mifflin, Boston.
Loewner, C. 1934. Über monotone Matrixfunktionen, *Math. Z.* 38:177–216.
Mahalanobis, P. C. 1936. On the generalized distance in statistics, *Proc. Nat. Inst. Sci. India* 2:49–55.
Mathai, A. M. and Provost, S. B. 1992. *Quadratic Forms in Random Variables*, Marcel Dekker, New York.
McLachlan, G. J. 1992. *Discriminant Analysis and Statistical Pattern Recognition*, Wiley, New York.
Pillai, K. C. S. 1985. Mahalanobis D^2, in *Encyclopedia of Statistical Sciences*, Vol. 5, S. Kotz, N. L. Johnson, and C. B. Read, Eds., Wiley, New York.
Vulikh, B. Z. 1967. *Introduction to the Theory of Partially Ordered Spaces*, Wolters-Noordhoff, Groningen.

Received 11 November 1995; final manuscript accepted 11 December 1996