MONOTONICITY OF CERTAIN MULTIVARIATE NORMAL PROBABILITIES[1]

by

Morris L. Eaton and Michael D. Perlman[2]

Technical Report No. 213

University of Minnesota
Minneapolis, Minnesota

1 This research was partially supported by National Science Foundation Grant NSF-GP-34482.

2 Now at the University of Chicago.

Abstract

Let (X, Y) have a $(p+q)$-dimensional normal distribution and let C, K be convex symmetric sets of dimensions p, q respectively. Under certain restrictions on the mean vector it is shown that $P\{X \in C, Y \in K\}$ is a monotonically increasing function of the first canonical correlation coefficient between X and Y, provided the remaining coefficients are zero.

Key words: multivariate normal distribution, convex symmetric sets, canonical correlations, monotonicity.

1. Introduction and summary.

Consider a family of positive definite matrices $\Sigma(\lambda)$, $|\lambda| \le 1$, defined as follows:

$$\Sigma(\lambda) = \begin{pmatrix} \Sigma_{11} & \lambda\Sigma_{12} \\ \lambda\Sigma_{21} & \Sigma_{22} \end{pmatrix} \tag{1.1}$$

where $\Sigma_{11}: p \times p$, $\Sigma_{22}: q \times q$, $\Sigma_{21}' = \Sigma_{12}: p \times q$. Suppose (X, Y) has a $(p + q)$-dimensional normal distribution $N((\mu, \nu), \Sigma(\lambda))$. Let $C \subseteq R^p$ and $K \subseteq R^q$ be convex symmetric (about 0) sets and define $P(\lambda)$ by

$$P(\lambda) = P_\lambda\{X \in C, Y \in K\} . \tag{1.2}$$

Note that $P(0) = P\{X \in C\}P\{Y \in K\}$ and $P(\lambda) = P(-\lambda)$, so we will assume that $0 \le \lambda \le 1$ throughout the remainder of this paper. The purpose of this paper is to establish the monotonicity (increasing) of $P(\lambda)$ under certain conditions on (μ, ν) and Σ_{12}.

The following results are known.

I. Sidak (1968) and Jogdeo (1970): If $\min\{p,q\} = 1$ and if $(\mu, \nu) = 0$, then $P(\lambda)$ is increasing in λ.

II. Khatri (1967): If the rank of Σ_{12} is 1 and if $(\mu, \nu) = 0$, then $P(\lambda) \ge P(0)$.

III. Das Gupta, Eaton, et al (1972): If the rank of Σ_{12} is 1 and if either (i) $(\mu, \nu) = 0$ or (ii) $\Sigma_{12} = c\mu'\nu$ for a positive scalar c satisfying $c^{-1} \ge \max\{\mu\Sigma_{11}^{-1}\mu', \nu\Sigma_{22}^{-1}\nu'\}$, then $P(\lambda) \ge P(0)$.

Discussion of these results and their statistical applications, primarily in the area of simultaneous confidence intervals, appears in the references cited above. The main result of the present paper is the following:

Theorem 1.1. Under the assumptions in III, $P(\lambda)$ is monotonically increasing in λ.

This theorem contains all previously known results in I, II, III above. The proof is given in section 3, while section 4 contains further conjectures and a counterexample.

2. A preliminary lemma on symmetric unimodal functions.

A function $\varphi: R^n \to [0, \infty)$ is called unimodal (Anderson (1955)) if for each $u > 0$, $\{x | \varphi(x) \geq u\}$ is a convex set in R^n. The following lemma is similar to results given by Anderson (1955) and Sherman (1955).

Lemma 2.1: Suppose $\varphi: R^n \to [0, \infty)$ is a unimodal function, symmetric about 0, and partition $x \in R^n$ as $x = (x_1, \dot{x})$ with $x_1 \in R^1$. Then

$$\Psi(x_1) \equiv \int_{R^{n-1}} \varphi(x_1, \dot{x})\, d\,\dot{x}$$

is a symmetric unimodal function on R^1.

Proof: First note that by Fubini's Theorem,

$$\Psi(x_1) = \int_0^\infty [\int_{R^{n-1}} I_{\{\omega | \varphi(\omega) \geq u\}}(x_1, \dot{x})\, d\,\dot{x}]\, d\, u,$$

where I_A denotes the indicator function of the set A, so it suffices to prove the lemma when $\varphi(x_1, \dot{x}) = I_D(x_1, \dot{x})$ where D is a symmetric convex set in R^n. We now have

$$\Psi(x_1) = \int_{R^{n-1}} I_D(x_1, \dot{x})\, d\,\dot{x}$$

and it must be shown that $A_t = \{x_1 | \Psi(x_1) \geq t\}$ is a convex symmetric set in R^1, for each $t > 0$. The symmetry of A_t is clear. Let $D_{x_1} = \{\dot{x} | (x_1, \dot{x}) \in D\} \subseteq R^{n-1}$ and let m denote Lebesgue measure on R^{n-1}, so that

$$\Psi(x_1) = m(D_{x_1}).$$

Now, if v_1 and v_2 are in A_t, we must show that

$$\Psi(\alpha\, v_1 + (1 - \alpha)\, v_2) \geq t$$

for $0 < \alpha < 1$, or equivalently that

$$\Psi^{1/(n-1)}(\alpha\, v_1 + (1 - \alpha)\, v_2) \geq t^{1/(n-1)} .$$

Since $D_{\alpha v_1+(1-\alpha)v_2} \supseteq \alpha\, D_{v_1} + (1 - \alpha)\, D_{v_2}$, we have

$$\begin{aligned}
\Psi^{1/(n-1)}(\alpha\, v_1 + (1 - \alpha)\, v_2) &= m^{1/(n-1)}(D_{\alpha v_1+(1-\alpha)v_2}) \\
&\geq m^{1/(n-1)}(\alpha\, D_{v_1} + (1 - \alpha)\, D_{v_2}) \\
&\geq \alpha\, m^{1/(n-1)}(D_{v_1}) + (1 - \alpha)\, m^{1/(n-1)}(D_{v_2}) \\
&= \alpha\, \Psi^{1/(n-1)}(v_1) + (1 - \alpha)\, \Psi^{1/(n-1)}(v_2) \\
&\geq t^{1/(n-1)}
\end{aligned}$$

where the second inequality is an application of the Brunn-Minkowski inequality (see Anderson (1955)).

Remark:

The above argument also shows that if D is a (symmetric) convex set in R^n and if

$$\Psi(x_1) = \int_{R^r} I_D(x_1, \dot{x})\, d\,\dot{x}$$

with $x_1 \in R^{n-r}$, then Ψ is a (symmetric) unimodal function for all $1 \le r \le n-1$.

3. Proof of the main theorem.

The proof of Theorem 1.1 is essentially based on the bivariate case. We first treat the bivariate case ($p = q = 1$) and assume that $\Sigma(\lambda)$ is a correlation matrix, i.e.,

$$\Sigma_{11} = \Sigma_{22} = 1, \quad \lambda\Sigma_{12} = \lambda\rho, \quad \text{where} \quad 0 < |\rho| < 1.$$

The assumptions in III now take the following form: either (i) $(\mu, \nu) = 0$ or (ii') $\mu\nu$ has the same sign as ρ and $|\rho| \le \min\{|\mu\,\nu^{-1}|, |\nu\,\mu^{-1}|\}$.

Theorem 3.1. Under assumption III in the bivariate case, $P_\lambda\{|X| \le a, |Y| \le b\}$ is monotonically increasing in λ for all $a > 0$, $b > 0$.

Proof: Without loss of generality we may assume $\rho > 0$. We have

$$P_\lambda\{|X| \le a, |Y| \le b\} = \int_{-a-\mu}^{a-\mu} \int_{-b-\nu}^{b-\nu} f_{\lambda\rho}(x, y)\, dx\, dy$$

where

$$f_{\lambda\rho}(x,y) = (2\pi)^{-1}(1 - \lambda^2\rho^2)^{-\frac{1}{2}}\exp\{-\tfrac{1}{2}(1 - \lambda^2\rho^2)^{-1}(x^2 - 2\lambda\rho xy + y^2)\}$$

is the bivariate normal density. By a result of Plackett (1954)

$$\frac{\partial}{\partial\lambda}[f_{\lambda\rho}(x, y)] = \rho\frac{\partial^2}{\partial x\,\partial y}[f_{\lambda\rho}(x, y)],$$

so that

$$\rho^{-1} \frac{d}{d\lambda} P_\lambda\{|X| \le a, \ |Y| \le b\}$$

$$= f_{\lambda\rho}(a-\mu, \ b-\nu) - f_{\lambda\rho}(a-\mu, \ -b-\nu) - f_{\lambda\rho}(-a-\mu, \ b-\nu) + f_{\lambda\rho}(-a-\mu, \ -b-\nu)$$

$$= M\{\exp[\lambda\rho ab(1-\lambda^2\rho^2)^{-1}]\cosh[(a(\mu-\lambda\rho\nu) + b(\nu-\lambda\rho\mu))(1-\lambda^2\rho^2)^{-1}]$$

$$- \exp[-\lambda\rho ab(1-\lambda^2\rho^2)^{-1}]\cosh[(a(\mu-\lambda\rho\nu) - b(\nu-\lambda\rho\mu))(1-\lambda^2\rho^2)^{-1}]\}$$

where

$$M = \pi^{-1}(1-\lambda^2\rho^2)^{-\frac{1}{2}}\exp\{-\tfrac{1}{2}(1-\lambda^2\rho^2)^{-1}(a^2+b^2+\mu^2+\nu^2-2\lambda\rho\mu\nu)\} \ .$$

Since $a, b, \rho > 0$,

$$\exp[\lambda\rho ab(1-\lambda^2\rho^2)^{-1}] > \exp[-\lambda\rho ab(1-\lambda^2\rho^2)^{-1}]$$

when $\lambda > 0$. Furthermore, under assumption (ii'), $\mu - \lambda\rho\nu$ and $\nu - \lambda\rho\mu$ have the same sign for all $0 \le \lambda \le 1$, so

$$\cosh[(a(\mu-\lambda\rho\nu) + b(\nu-\lambda\rho\mu))(1-\lambda^2\rho^2)^{-1}]$$

$$\ge \cosh[(a(\mu-\lambda\rho\nu) - b(\nu-\lambda\rho\mu))(1-\lambda^2\rho^2)^{-1}] \ ,$$

while equality holds under (i). Thus

$$\frac{d}{d\lambda} P_\lambda\{|X| \le a, \ |Y| \le b\} > 0$$

if $0 < \lambda \le 1$, which completes the proof.

<u>Corollary 3.2.</u> Let g, h be nonnegative, symmetric (about 0), unimodal functions on R^1. Under assumption III in the bivariate case, $E_\lambda[g(X)h(Y)]$ is increasing in λ.

<u>Proof:</u> The functions g and h can be approximated by increasing sequences of nonnegative symmetric unimodal step functions, say g_n and h_n. By Theorem 3.1, $E_\lambda[g_n(X)\, h_n(Y)]$ is increasing in λ. The result now follows from the Monotone Convergence Theorem.

We now return to the $(p+q)$-dimensional case and apply Corollary 3.2 to obtain our main result in the special case where $\Sigma(\lambda)$ is a "canonical correlation" matrix. That is, $\Sigma_{11} = I_p$ (the p-dimensional identity matrix), $\Sigma_{22} = I_q$, and

$$(3.1) \qquad \lambda\Sigma_{12} = \begin{pmatrix} \lambda\rho_1 & 0 & \cdots & 0 \\ 0 & \lambda\rho_2 & & \\ \vdots & & \ddots & \\ 0 & & & \end{pmatrix},$$

where $1 \geq \rho_1 \geq \rho_2 \geq \ldots \geq \rho_s \geq 0$ are the "canonical correlations" between X and Y when $\lambda = 1$, and $s = \min(p, q)$. The assumptions in III take the following form: $\rho_2 = \ldots = \rho_s = 0$ and either (i) $(\mu, \nu) = 0$ or (ii'') $\mu = (\mu_1, 0, \ldots, 0)$ and $\nu = (\nu_1, 0, \ldots, 0)$ where $\mu_1\nu_1$ has the same sign as ρ_1 and $|\rho_1| \leq \min\{|\mu_1\nu_1^{-1}|, |\nu_1\mu_1^{-1}|\}$.

<u>Theorem 3.3.</u>

Under assumption III when $\Sigma(\lambda)$ is a canonical correlation matrix, $P(\lambda)$ defined in (1.2) is increasing in λ.

Proof: Partition $X = (X_1, \dot{X})$ and $Y = (Y_1, \dot{Y})$ where $X_1, Y_1 \in R^1$. Since $\Sigma(\lambda)$ is a canonical correlation matrix, (X_1, Y_1), $\dot{X}$, and $\dot{Y}$ are independent random vectors, so

$$P(\lambda) = E_\lambda[P\{(X_1, \dot{X}) \in C | X_1\} \, P\{(Y_1, \dot{Y}) \in K | Y_1\}]$$

$$\equiv E_\lambda[g(X_1)\, h(Y_1)] \; ,$$

where

$$g(x_1) = (2\pi)^{-(p-1)/2} \int_{R^{p-1}} I_C(x_1, \dot{x}) \, e^{-\frac{1}{2}\|\dot{x}\|^2} \, d\dot{x}$$

$$h(y_1) = (2\pi)^{-(q-1)/2} \int_{R^{q-1}} I_K(y_1, \dot{y}) \, e^{-\frac{1}{2}\|\dot{y}\|^2} \, d\dot{y} \; .$$

Applying Lemma 2.1 with $\varphi(x_1, \dot{x}) = I_C(x_1, \dot{x}) \, e^{-\frac{1}{2}\|\dot{x}\|^2}$, we have that $g(x_1)$ is a symmetric unimodal function on R^1, and similarly for $h(y_1)$. Furthermore, (X_1, Y_1) has a bivariate normal distribution which satisfies the assumptions of Corollary 3.2. Thus by that corollary, $P(\lambda)$ is increasing in λ.

Finally, we prove our main Theorem 1.1 in the general case by reducing $\Sigma(\lambda)$ given by (1.1) to the canonical correlation form and applying Theorem 3.3.

Proof of Theorem 1.1: Consider new variables $U = XA$ and $V = YB$ where $A{:}p \times p$ and $B{:}q \times q$ are non-singular matrices defined by

$$A = \Sigma_{11}^{-\frac{1}{2}} \Gamma, \qquad B = \Sigma_{22}^{-\frac{1}{2}} \Delta$$

where $\Gamma: p \times p$, $\Delta: q \times q$ are orthogonal matrices.

In case (ii), choose the first column of Γ to be $\tau_1^{-1}\Sigma_{11}^{-\frac{1}{2}}\mu'$ and the first column of Δ to be $\tau_2^{-1}\Sigma_{22}^{-\frac{1}{2}}\nu$, where $\tau_1 = (\mu\Sigma_{11}^{-1}\mu')^{\frac{1}{2}}$ and $\tau_2 = (\nu\Sigma_{22}^{-1}\nu')^{\frac{1}{2}}$. Then U and V have mean vectors $(\tau_1, 0, \ldots, 0)$ and $(\tau_2, 0, \ldots, 0)$ respectively, and the covariance matrix of (U, V) has canonical correlation form, i.e., $\mathrm{Cov}(U) = I_p$ $\mathrm{Cov}(V) = I_q$ and

$$\mathrm{Cov}_\lambda(U, V) = \begin{pmatrix} \lambda c \tau_1 \tau_2 & 0 & \cdots & 0 \\ 0 & 0 & & \\ \vdots & & \ddots & \\ 0 & & & \end{pmatrix} .$$

The assumption (ii) that $c > 0$ and $c^{-1} \geq \max\{\tau_1^2, \tau_2^2\}$ implies that assumption (ii'') preceding Theorem 3.3 is satisfied. Let $\hat{C}$ and $\hat{K}$ be the images of C and K under A and B respectively. Then by Theorem 3.3,

$$P_\lambda\{X \in C, Y \in K\} = P_\lambda\{U \in \hat{C}, V \in \hat{K}\}$$

is increasing in λ.

In case (i), write $\Sigma_{12} = a'b$ where $a \in R^p$, $b \in R^q$ (since rank $\Sigma_{12} = 1$). Now, choose the first column of Γ to be $\Sigma_{11}^{-\frac{1}{2}} a'/(a\Sigma_{11}^{-1}a')^{\frac{1}{2}}$ and the first column of Δ to be $\Sigma_{22}^{-\frac{1}{2}} b'/(b\Sigma_{22}^{-1}b')^{\frac{1}{2}}$. Here again the covariance matrix of (U, V) has canonical correlation form, and Theorem 3.3 again yields the desired conclusion.

Corollary 3.4: If $\varphi_1 : R^p \to [0, \infty)$ and $\varphi_2 : R^q \to [0, \infty)$ are both symmetric unimodal functions and the assumptions of Theorem 1.1 hold, then $E_\lambda[\varphi_1(X)\varphi_2(Y)]$ is an increasing function of λ.

Proof: The proof is similar to the proof of Corollary 3.2.

Remark: Theorem 1.1 and Corollary 3.4 remain valid if it is only assumed that $\Sigma(\lambda)$ is positive semi-definite for $0 \le \lambda \le 1$, provided that assumption (ii) is replaced by the assumption that $\Sigma_{11} - c\,\mu'\mu$ and $\Sigma_{22} - c\,\nu'\nu$ are both positive semi-definite.

4. Discussion

Consider $(X, Y) \sim N_{p+q}((0, 0), \Sigma(\lambda))$ as in section 1 and define $P(\lambda)$ by (1.2). It has been conjectured (Das Gupta, Eaton, et al (1972)) that $P(\lambda) \geq P(0)$ with no restriction on the rank of Σ_{12}. We here state the stronger conjecure that $P(\lambda)$ is increasing in λ. Theorems 1.1 and 3.3 yield this fact when only one canonical correlation between X and Y is nonzero, i.e., Σ_{12} has rank 1. If rank $(\Sigma_{12}) > 1$, note that as λ increases, all nonzero canonical correlations increase at the same rate. The following example shoes that $P\{X \in C, Y \in K\}$ does not necessarily increase as each canonical correlation coefficient increases separately.

Example: Take $p = q = 2$, $\mu = \nu = 0$, $\Sigma_{11} = \Sigma_{22} = I_2$, and $\Sigma_{12} = \text{diag}\{\rho_1, \rho_2\}$, with $1 > \rho_1 \geq \rho_2 \geq 0$. Then $\text{cov}(X_1 - X_2, Y_1 + Y_2) = \rho_1 - \rho_2$, so $P\{|X_1 - X_2| \leq a, |Y_1 + Y_2| \leq b\}$ is increasing in ρ_1 but decreasing in ρ_2.

References

[1] 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.

[2] Das Gupta, S., Eaton, M. L., Olkin, I., Perlman, M. D., Savage, L. J., and Sobel, M. (1972). Inequalities on the probability content of convex regions for elliptically contoured distributions, Proc. Sixth Berk. Symp. on Math. Stat. and Prob., Vol. II, 241-265.

[3] Jogdeo, K. (1970). A simple proof of an inequality for multivariate normal probabilities of rectangles, Ann. Math. Statist., 41 , 1357-1359.

[4] Khatri , C. G. (1967). On certain inequalities for normal distributions and their applications to simultaneous confidence bounds, Ann. Math. Statist., 38 , 1853-1867.

[5] Sherman, S. (1955). A theorem on convex sets with applications, Ann. Math. Statist., 26 , 763-766.

[6] Sidak, Z. (1968). On multivariate normal probabilities of rectangles: their dependence on correlations., Ann. Math. Statist., 39 , 1425-1434.