The Annals of Statistics
1981, Vol. 9, No. 5, 1035–1049

TOTAL POSITIVITY PROPERTIES OF ABSOLUTE VALUE MULTINORMAL VARIABLES WITH APPLICATIONS TO CONFIDENCE INTERVAL ESTIMATES AND RELATED PROBABILISTIC INEQUALITIES[1]

BY SAMUEL KARLIN AND YOSEF RINOTT

Stanford University; Stanford University and Hebrew University, Jerusalem

Total positivity properties of multivariate densities are useful in deducing positive dependence of random vector components and related probability inequalities. In this paper we determine necessary and sufficient conditions for total positivity of absolute value multinormal variables. The results are applied to obtain positive dependence and associated inequalities for the multinormal and related distributions, e.g., the multivariate t and Wishart distributions. Inequalities of this type yield bounds for multivariate confidence set probabilities.

1. Introduction. Recent statistical literature encompasses a substantial body of methods and results on multivariate probabilistic inequalities and their applications to multivariate analysis and simultaneous inference. Much of this work concentrates on the normal and related distributions (e.g., multivariate t, F, Wishart). A key to many of these inequalities is the degree of total positivity and/or log concavity properties inherent to these distributions. Total positivity for the multivariate normal distribution, including various applications, were studied recently by Barlow and Proschan (1975), Kemperman (1977), Abdel-Hameed and Sampson (1978), Perlman and Olkin (1978), among others.

The principal theorem of this paper (Theorem 3.1) delineates necessary and sufficient conditions on the convariance matrix of a Gaussian vector random variable $\mathbf{X} = (X_1, X_2, \cdots, X_n)$ implying that the density of the absolute components vector $|\mathbf{X}| = (|X_1|, |X_2|, \cdots, |X_n|)$ is multivariate totally positive of order 2. This result, apart from its intrinsic interest, entails a wide scope of applications. At this point it is useful to review basic information on multivariate total positivity theory essential for our applications. For more details and ramifications of this material, e.g., see Kemperman (1977), and Karlin and Rinott (1980).

A nonnegative function $f(x, y)$ defined for $(x, y) \in \mathscr{X} \times \mathscr{Y} \subseteq R^2$ is totally positive of order 2 (TP_2) if all second order determinants $\det\{f(x_i, y_j)\}$ are nonnegative for every choice, $x_1 < x_2$, $y_1 < y_2$. (In parametric statistics the TP_2 endowment is intimately related to the concept of monotone likelihood ratio, e.g., see Karlin (1968, Chapter 1).)

FACT 1.1. If $f(\mathbf{x}) > 0$ for all $\mathbf{x} = (x_1, \cdots, x_n) \in R^n$ and $f(\mathbf{x}) = f(x_1, x_2, \cdots, x_n)$ is TP_2 in every pair of variables when the remaining variables are kept fixed, then for every $\mathbf{x} = (x_1, \cdots, x_n)$ and $\mathbf{y} = (y_1, \cdots, y_n) \in R^n$

$$f(\mathbf{x} \vee \mathbf{y}) f(\mathbf{x} \wedge \mathbf{y}) \geq f(\mathbf{x}) f(\mathbf{y}) \tag{1.1}$$

where

Received October, 1979; revised October, 1980.
[1] Supported in part by NSF Grant MCS 80624-A02 and NIH Grant GM10452-16.
AMS 1970 *subject classification.* 62H05.
Key words and phrases. Total positivity, association, correlation inequalities, multivariate normal distribution, multivariate t distribution, Wishart distribution.

$\mathbf{x} \vee \mathbf{y} = (\max(x_1, y_1), \cdots, \max(x_n, y_n))$ and $\mathbf{x} \wedge \mathbf{y} = (\min(x_1, y_1), \cdots, \min(x_n, y_n))$.

A function $f(\mathbf{x})$ satisfying (1.1) is henceforth called *multivariate totally positive of order 2* (MTP_2). (Fact 1.1 is essentially due to Lorentz (1953) as noted by Rinott (1973). See also Kemperman (1977).)

The formulation (1.1) provides a concept of MTP_2 functions defined on any Cartesian product of ordered spaces $\mathscr{X}_1 \times \mathscr{X}_2 \times \cdots \mathscr{X}_n$ where each $\mathscr{X}_k$ can be discrete (finite or countable) sets, continuous intervals or more general totally ordered sets. The number of coordinates n can even be infinite or a continuum. Some of these extensions and applications will be dealt with elsewhere.

It lends flexibility and value to the principal theorems of this paper to highlight first a number of basic consequences concomitant to the MTP_2 property. For this objective we present the MTP_2 concept involving two kernels reminiscent of the monotone likelihood ratio property in the univariate context.

FACT 1.2. (Holley (1974), Preston (1974), Kemperman (1977)). Let f_1 and f_2 be two probability densities with respect to Lebesgue measure on R^n. Suppose that for every $\mathbf{x}, \mathbf{y} \in R^n$

$$f_2(\mathbf{x} \vee \mathbf{y}) f_1(\mathbf{x} \wedge \mathbf{y}) \geq f_2(\mathbf{x}) f_1(\mathbf{y}). \tag{1.2}$$

Then for any increasing function φ defined on R^n (that is, $\varphi(\mathbf{x}) \leq \varphi(\mathbf{y})$ provided $\mathbf{x} \leq \mathbf{y}$ where the ordering connotes $x_i \leq y_i$, $i = 1, \cdots, n$, $\mathbf{x} = (x_1, \cdots, x_n)$, $\mathbf{y} = (y_1, \cdots, y_n)$)

$$\int_{R^n} \varphi(\mathbf{x}) f_2(\mathbf{x})\, d\mathbf{x} \geq \int_{R^n} \varphi(\mathbf{x}) f_1(\mathbf{x})\, d\mathbf{x}. \tag{1.3}$$

Another vernacular for (1.3) is the statement

$$f_2 \overset{st}{>} f_1 \tag{1.4}$$

signifying that the *density* f_2 *is stochastically larger than the density* f_1.

COROLLARY 1.1. (Sarkar (1969), Fortuin et al. (1971), Barlow and Proschan (1975)). *Let f be a* MTP_2 *probability density on R^n. Let φ and ψ be both increasing* (*or decreasing*) *functions on R^n. Then*

$$\int_{R^n} \varphi(\mathbf{x})\psi(\mathbf{x}) f(\mathbf{x})\, d\mathbf{x} \geq \left(\int_{R^n} \varphi(\mathbf{x}) f(\mathbf{x})\, d\mathbf{x}\right)\left(\int_{R^n} \psi(\mathbf{x}) f(\mathbf{x})\, d\mathbf{x}\right). \tag{1.5}$$

The inequality (1.5) is commonly referred to as the *multivariate Tchebycheff rearrangement inequality* which in one dimension holds for any probability density $f(x)$.

For $n > 1$ the validity of (1.5) requires some stipulation on f such as the MTP_2 property. Of course, where $\varphi(\mathbf{x})$ and $\psi(\mathbf{x})$ are monotone in opposite directions since $f(\mathbf{x})$ is a density, the inequality sign of (1.5) is reversed.

REMARK 1.1. Let $f(\mathbf{a}, \mathbf{x})$, $\mathbf{a} \in R^m$, $\mathbf{x} \in R^n$ be MTP_2 on R^{n+m} and assume $\int f(\mathbf{a}, \mathbf{x})\, d\mathbf{x} = 1$ for all $\mathbf{a} \in R^m$. Then $\mathbf{a} \geq \mathbf{b}$ implies $f(\mathbf{a}, \mathbf{x} \vee \mathbf{y}) f(\mathbf{b}, \mathbf{x} \wedge \mathbf{y}) \geq f(\mathbf{a}, \mathbf{x}) f(\mathbf{b}, \mathbf{y})$, and Fact 1.2 can be invoked to obtain the stochastic comparison (1.4) between $f_2(\mathbf{x}) = f(\mathbf{a}, \mathbf{x})$ and $f_1(\mathbf{x}) = f(\mathbf{b}, \mathbf{x})$.

REMARK 1.2. Facts 1.1 and 1.2 and Corollary 1.1 continue to hold if R^n is replaced by a product of ordered spaces $\mathscr{X}_1 \times \cdots \times \mathscr{X}_n$ with Lebesgue measure $d\mathbf{x}$ replaced by a product measure $\boldsymbol{\sigma} = \sigma_1 \times \cdots \times \sigma_n$ where σ_i is a σ-finite measure on $\mathscr{X}_i$, $i = 1, \cdots, n$.

It is convenient to express the inequality (1.5) in random variable notation. If $\mathbf{X} =$

$(X_1, \cdots, X_n)$ is a random vector following a MTP_2 density then by (1.5) for every pair of increasing or decreasing functions φ and ψ, we have

$$E\{\varphi(\mathbf{X})\psi(\mathbf{X})\} \geq \{E\varphi(\mathbf{X})\}\{E\psi(\mathbf{X})\}. \tag{1.6}$$

This implies that if $\varphi_1, \cdots, \varphi_r$ are all nonnegative and increasing (or decreasing) then

$$E\{\varphi_1(\mathbf{X})\varphi_2(\mathbf{X}) \cdots \varphi_r(\mathbf{X})\} \geq \prod_{i=1}^r E\varphi_i(\mathbf{X}). \tag{1.7}$$

Random variables $X_1, \cdots, X_n$ satisfying (1.6) are called *associated* (see Esary, Proschan and Walkup (1967)). In particular, (1.7) entails the relations

$$\mathrm{Cov}(X_i, X_j) \geq 0 \tag{1.8}$$

$$\begin{aligned} P(X_1 \geq c_1, \cdots, X_n \geq c_n) &\geq \prod_{i=1}^n P(X_i \geq c_i) \\ P\{X_1 \leq c_1, \cdots, X_n \leq c_n\} &\geq \prod_{i=1}^n P\{X_i \leq c_i\} \end{aligned} \tag{1.9}$$

The latter is commonly called *positive quadrant dependence*, Lehmann [1966]. For further ramifications on the connections between MTP_2 and other notions of positive dependence see e.g. Esary and Proschan (1972), Barlow and Proschan (1975), Shaked (1975), Jogdeo (1977), Abdel-Hameed and Sampson (1978) and Karlin and Rinott (1980).

With the incentive of Facts 1.1–1.2 in mind, together with the corollaries of (1.5)–(1.9), encompassing many basic probabilistic inequalities, it is germane to ascertain the total positivity character of the normal and related classical densities. Deciding when a multinormal density is MTP_2 is elementary. Specifically,

FACT 1.3. Let $\mathbf{X} = (X_1, \cdots, X_n) \sim N(0, \Sigma)$ i.e., $\mathbf{X}$ is normally distributed with covariance matrix Σ. Then $\mathbf{X}$ has a MTP_2 density iff the off diagonal elements of $-\Sigma^{-1}$ are all nonnegative (e.g., see Sarkar (1969), Barlow and Proschan (1975)).

In particular, a bivariate normal density is totally positive if and only if the correlation coefficient is nonnegative.

Our main theorem (proved in Section 3) concerns the MTP_2 nature of the density for $|\mathbf{X}| = (|X_1|, |X_2|, \cdots, |X_n|)$ where $\mathbf{X} = (X_1, \cdots, X_n)$ follows the distribution of $N(\mathbf{0}, \Sigma)$. Such results would be of interest in establishing bounds for two-sided multivariate confidence intervals and inequalities for functions of $|\mathbf{X}|$ (even functions of $\mathbf{X}$).

THEOREM 3.1. *Let*

$$\mathbf{X} = (X_1, X_2, \cdots, X_n) \quad \textit{be distributed as } N(0, \Sigma). \tag{1.10}$$

Then a necessary and sufficient condition that the density of

$$|\mathbf{X}| = (|X_1|, |X_2|, \cdots, |X_n|) \tag{1.11}$$

be MTP_2 *is that there exists a diagonal matrix D with diagonal elements* ± 1 *such that the off-diagonal elements of*

$$-D\Sigma^{-1}D \quad \textit{are all nonnegative.} \tag{1.12}$$

Our proof of Theorem 3.1, developed in Sections 2 and 3, relies on results of Kelly and Sherman (1968) pertaining to correlation inequalities for particle systems subject to interaction potentials.

REMARK 1.3. Since $D\Sigma D$ is the covariance matrix of $(d_1X_1, \cdots, d_nX_n)$ where $d_i = \pm 1$, $i = 1, \cdots, n$, are the diagonal elements of D, the condition of (1.12) is equivalent in view of Fact 1.3 to the existence of an adjustment of signs yielding $(d_1 X_1, \cdots, d_nX_n)$ with a MTP_2 density.

Thus, the MTP_2 property of $|\mathbf{X}| = (|X_1|, \cdots, |X_n|)$ implies by (1.8) that for an appropriate $D = \text{diag}(d_1, \cdots, d_n)$ the correlations between the components of the adjusted vector $(d_1X_1, \cdots, d_nX_n)$, which coincide with the entries of $D\Sigma D$ are nonnegative. The converse is not true, indeed, a multinormal vector $\mathbf{X} = (X_1, \cdots, X_n)$ with all positive correlations need not have the absolute valued vector $|\mathbf{X}|$, MTP_2.

REMARK 1.4. With a bivariate normal density for $\mathbf{X} = (X_1, X_2)$ the density $g(z_1, z_2)$ of $\mathbf{Z} = (Z_1, Z_2) = (|X_1|, |X_2|)$ has the form $\exp[-bz_1^2 - cz_2^2] \cosh(az_1z_2)$ which is always TP_2 (since $\cosh(az_1z_2)$ is TP_2 for $z_1, z_2 > 0$) without any restriction on the signature of the correlation between X_1 and X_2.

REMARK 1.5. The result of Theorem 3.1 for $n = 3$ is due to Abdel-Hameed and Sampson (1978). They phrased this case as follows. A necessary and sufficient condition that the density of $\mathbf{X} = (|X_1|, |X_2|, |X_3|)$ be MTP_2 is that

(1.13) $$b_{12}b_{13}b_{23} \geq 0 \qquad \text{where} \quad B = \| b_{ij} \| = -\Sigma^{-1}.$$

It is readily checked that (1.13) is equivalent to (1.12) for a covariance matrix of order 3. They conjectured the general Theorem 3.1.

We next inquire concerning accessible criteria that enable us to verify (1.12).

For $n > 3$ the condition of (1.12) is not equivalent to the property that $B = -\Sigma^{-1}$ satisfies $b_{ij}b_{ik}b_{jk} \geq 0$ for all $1 \leq i \neq j \neq k \leq n$ in the presence of zero terms in B (see Section 3). Nevertheless, there is a modification of Theorem 3.1 in the spirit of (1.13) which we now state.

THEOREM 3.1′. *Let* $\mathbf{X}$ *and* $|\mathbf{X}|$ *be as in* (1.10) *and* (1.11), *respectively, and set* $B = \| b_{ij} \| = -\Sigma^{-1}$. *A necessary and sufficient condition that the density of* $|\mathbf{X}|$ *be* MTP_2 *is the existence of a consistent assignment*

$$b_{ij}^* = \begin{cases} b_{ij} & \text{if} \quad b_{ij} \neq 0 \\ +1 \text{ or } -1 & \text{if} \quad b_{ij} = 0 \end{cases}$$

fulfilling

(1.14) $$b_{ij}^* b_{ik}^* b_{jk}^* > 0 \qquad \textit{for all} \quad 1 \leq i \neq j \neq k \leq n.$$

It is useful to exhibit several classes of covariance matrices fulfilling (1.12).

EXAMPLE 4.1. Let $\mathbf{X} \sim N(0, \Sigma)$ where $\Sigma = \Lambda + R$ with $R = \| \alpha_i\alpha_j \|$ of rank 1, $\alpha = (\alpha_1, \cdots, \alpha_n)$ and Λ is a diagonal matrix, $\Lambda = \text{diag}(\gamma_1, \gamma_2, \cdots, \gamma_n)$, $\gamma_i > 0$, $i = 1, \cdots, n$. It is familiar that Σ^{-1} admits the representation $\Sigma^{-1} = \Lambda^{-1} - S$, $S = \| \beta_i\beta_j \|$. Obviously, for $D = \text{diag}(d_1, d_2, \cdots, d_n)$, $-D\Sigma^{-1}D$ has off diagonal elements $c_{ij} = d_i\beta_i d_j\beta_j$, and the choice $d_i = \text{sign } \beta_i$ renders $c_{ij} \geq 0$. Thus, the conditions of Theorems 3.1 are fulfilled and the density of $|\mathbf{X}|$ is MTP_2.

EXAMPLE 4.2. Let

$$\mathbf{Y} = \mathbf{X} + (\lambda_1 U, \cdots, \lambda_k U, \lambda_{k+1} V, \cdots, \lambda_n V)$$

where

$$\mathbf{X} = (X_1, \cdots, X_n) \sim N(0, \Lambda), \Lambda = \text{diag}(\gamma_1, \cdots, \gamma_n), (U, V) \sim N\left(0, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right).$$

Then, for any $\lambda_1, \cdots, \lambda_n$ and correlation ρ the density of

(1.15) $$|\mathbf{Y}| = (|Y_1|, |Y_2|, \cdots, |Y_n|) \qquad \text{is } MTP_2.$$

For details on this example, see Section 4.

We analyze in Section 4 the class of examples

$$\mathbf{Y} = \mathbf{X} + (\lambda_1 U, \lambda_2 U, \cdots, \lambda_k U, \lambda_{k+1} V, \cdots, \lambda_{k+l} V, \lambda_{k+l+1} W, \cdots, \lambda_n W)$$

where $\mathbf{X} \sim N(0, \Lambda)$ and (U, V, W) are correlated multinormal variables and ascertain the conditions when $|\mathbf{Y}|$ is MTP_2. Other cases of covariance structures are also discussed.

We close the introduction by describing a number of applications of Theorem 3.1 forthcoming with the assistance of Facts 1.1–1.3, and the inequalities of (1.3)–(1.7). By these means inequalities of the type

$$P\{\mathbf{X} \in A \cap B\} \geq P\{\mathbf{X} \in A\} P\{\mathbf{X} \in B\} \tag{1.16}$$

are shown to hold for sets A, B in a class described below, where $\mathbf{X} \sim N(0, \Sigma)$ and Σ satisfies (1.12). Here A and B are "even" sets, i.e., if $\mathbf{x} = (x_1, x_2, \cdots, x_n) \in A$ then $(\epsilon_1 x_1, \epsilon_2 x_2, \cdots, \epsilon_n x_n) \in A$ for all choices of $\epsilon_i = \pm 1$, and the same holds for B. Also, for any $\mathbf{0} \leq \mathbf{x} \leq \mathbf{y} \in R^n$, $\mathbf{y} \in A$ implies $\mathbf{x} \in A$, and $\mathbf{y} \in B$ entails $\mathbf{x} \in B$.

The result of (1.16) relates to a challenging conjecture that (1.16) holds for any $\mathbf{X} \sim N(0, \Sigma)$ for balanced convex sets (i.e., where $A = -A$, $B = -B$, and A and B are convex. (See Das Gupta et al. (1972), Pitt (1977); see also Sidak (1973), Jogdeo (1977), Dykstra (1979)). In certain respects the conclusion of (1.16) covers a more general situation as we treat it in that A, B are not necessarily convex, but in other respects less general as we demand more symmetry (evenness) than mere reflection symmetry (= balanced sets) along with the MTP_2 condition.

For another illustration of the power of Theorem 3.1 consider

$$|\mathbf{X}| = (|X_1|, \cdots, |X_n|) \text{ possessing a joint MTP}_2 \text{ density where } \mathbf{X} \sim N(0, \Sigma). \tag{1.17}$$

Define

$$S_i = \textstyle\sum_{\nu=1}^{p} X_{i\nu}^2, \qquad i = 1, 2, \cdots, n, \tag{1.18}$$

where $\mathbf{X}_\nu = (X_{1\nu}, X_{2\nu}, \cdots, X_{n\nu})$, $\nu = 1, \cdots, p$ are i.i.d. vectors such that (1.17) holds. Clearly, $S_1, S_2, \cdots, S_n$ have the distribution of the diagonal elements of a random positive definite $n \times n$ matrix $\mathbf{S}$ where $\mathbf{S}$ is governed by the distribution $W_n(\text{p}, \Sigma)$ (Wishart distribution with p degrees of freedom). We will establish (see Theorem 6.1) the following result: *Under the conditions and constructions of* (1.17)–(1.18),

$$\Pr\{S_1 \geq c_1, S_2 \geq c_2, \cdots, S_n \geq c_n\} \geq \textstyle\prod_{i=1}^{n} P\{S_i \geq c_i\}, \tag{1.19}$$

for any positive c_i. Actually the variables $\{S_1, S_2, \cdots, S_n)$ are associated. In particular, the inequality (1.19) *applies if Σ has the form described in Examples* 4.1–4.2.

The finding of (1.19) generalizes a result in Das Gupta et al. (1972), and Jogdeo (1977). Other inequalities and bounds on the probability content of symmetric sets for the multivariate t and F distributions emerge from judicious applications of Theorem 3.1, (1.3) and (1.7) coupled with appropriate conditioning arguments and devices of the theory of associated or positively dependent random vectors.

2. Some correlation inequalities for binary variables. We begin by introducing the Kelly-Sherman inequalities in a formulation convenient for our main theorem.

Let $\boldsymbol{\delta} = (\delta_1, \cdots, \delta_n)$ be a random vector such that each component assumes the possible values ± 1 governed by the joint probability density

$$p(\boldsymbol{\delta}) = p(\delta_1, \cdots, \delta_n) = c \exp\left[\textstyle\sum_{1 \leq i < j \leq n} a_{ij} \delta_i \delta_j\right]. \tag{2.1}$$

Here $c > 0$ is a normalizing constant and $a_{ij} \geq 0$, $1 \leq i < j \leq n$. Given a function $f(\delta_1, \cdots, \delta_n)$ we denote the expectation of f with respect to the density $p(\boldsymbol{\delta})$ by $E_{\mathscr{S}} f = \sum_{\boldsymbol{\delta}} f(\boldsymbol{\delta}) p(\boldsymbol{\delta})$ where the sum extends over $\boldsymbol{\delta}$ in

$$\Delta = \{\boldsymbol{\delta} = (\delta_1, \cdots, \delta_n) \colon \delta_i = \pm 1, i = 1, \cdots, n\}. \tag{2.2}$$

Let $F_A(\boldsymbol{\delta})$ denote the function

$$f_A(\delta_1, \cdots, \delta_n) = \prod_{i\in A} \delta_i \qquad \text{where} \quad A \subseteq \{1, \cdots, n\} = N.$$

Using properties of exponential polynomials Kelly and Sherman (1968) developed an extensive theory of correlation inequalities of which the following is an important case.

THEOREM 2.1. (Kelly-Sherman (1968)) *For any two index sets A, $B \subseteq N$ and $a_{ij} \geq 0$, $1 \leq i < j \leq n$*

$$E_{\mathscr{S}}(f_A f_B) \geq (E_{\mathscr{S}} f_A)(E_{\mathscr{S}} f_B). \tag{2.3}$$

REMARK 2.1. The density $p(\boldsymbol{\delta})$ is MTP_2 and therefore the component random variables $\delta_1, \delta_2, \cdots, \delta_n$ are associated, so that $E_{\mathscr{S}}(fg) \geq (E_{\mathscr{S}} f)(E_{\mathscr{S}} g)$ when f and g are both increasing (or decreasing). However, the functions f_A and f_B appearing in (2.3) are not monotone in general.

PROPOSITION 2.1. *Let $A_1, \cdots, A_r$; $B_1, \cdots, B_s$; and $C_1, \cdots, C_t$ be any subsets of N and let $a_1, \cdots, a_r$; $b_1, \cdots, b_s$; and $c_1, \cdots, c_t$ be any nonnegative constants. Then,*

$$E_{\mathscr{S}}\{\exp[\textstyle\sum_{\nu=1}^r a_\nu f_{A_\nu}]\exp[\sum_{\mu=1}^s b_\mu f_{B_\mu}]\} \geq E_{\mathscr{S}}\{\exp[\sum_{\nu=1}^r a_\nu f_{A_\nu}]\} E_{\mathscr{S}}\{\exp[\sum_{\mu=1}^s b_\mu f_{B_\mu}]\} \tag{2.4}$$

and

$$E_{\mathscr{S}}\{\exp[\textstyle\sum_{\nu=1}^r a_\nu f_{A_\nu}]\exp[\sum_{\mu=1}^s b_\mu f_{B_\mu}]\exp[\sum_{\lambda=1}^t c_\lambda f_{C_\lambda}]\} \tag{2.5}$$

$$\geq E_{\mathscr{S}}\{\exp[\textstyle\sum_{\nu=1}^r a_\nu f_{A_\nu}]\} E_{\mathscr{S}}\{\exp[\sum_{\mu=1}^s b_\mu f_{B_\mu}]\} E_{\mathscr{S}}\{\exp[\sum_{\lambda=1}^t c_\lambda f_{C_\lambda}]\}$$

and similarly for any number of products.

PROOF. We deal with (2.4) since (2.5) plainly results by twice application of (2.4). We expand the exponential terms on the left of (2.4) and rearrange to obtain

$$E_{\mathscr{S}}\{\exp[\textstyle\sum_{\nu=1}^r a_\nu f_{A_\nu}] \exp[\sum_{\mu=1}^s b_\mu f_{B_\mu}]\}$$

$$= E_{\mathscr{S}}\left\{\sum_{m=0}^{\infty} \frac{1}{m!} (\textstyle\sum_{\nu=1}^r a_\nu f_{A_\nu})^m \sum_{l=0}^{\infty} \frac{1}{l!} (\sum_{\mu=1}^s b_\mu f_{B_\mu})^l\right\}$$

$$= \sum_{m=0}^{\infty} \sum_{l=0}^{\infty} \frac{1}{m!}\frac{1}{l!} E_{\mathscr{S}}\{(\textstyle\sum_{\nu=1}^r a_\nu f_{A_\nu})^m (\sum_{\mu=1}^s b_\mu f_{B_\mu})^l\}$$

$$= \sum_{m=0}^{\infty} \sum_{l=0}^{\infty} \frac{1}{m!}\frac{1}{l!} E_{\mathscr{S}}\left\{\sum_{m_1+\cdots+m_r=m} \binom{m}{m_1 \cdots m_r} \prod_{\nu=1}^r (a_\nu f_{A_\nu})^{m_\nu} \right. \tag{2.6}$$

$$\left. \cdot \sum_{l_1+\cdots+l_s=l} \binom{l}{l_1 \cdots l_s} \prod_{\mu=1}^s (b_\mu f_{B_\mu})^{l_\mu}\right\}$$

$$= \sum_{m=0}^{\infty} \sum_{l=0}^{\infty} \frac{1}{m!}\frac{1}{l!} \sum \binom{m}{m_1 \cdots m_r} (\textstyle\prod_{\nu=1}^r a_\nu^{m_\nu}) \sum \binom{l}{l_1 \cdots l_s}$$

$$\cdot (\textstyle\prod_{\mu=1}^s b_\mu^{l_\mu}) E_{\mathscr{S}}\{(\prod_{\nu=1}^r f_{A_\nu}^{m_\nu})(\prod_{\mu=1}^s f_{B_\mu}^{l_\mu})\}.$$

Define the sets of indices C and D by the specification $i \in C(D)$ if and only if $\sum_{\nu=1}^r m_\nu I_{A_\nu}(i)$ is odd ($\sum_{\mu=1}^s l_\mu I_{B_\mu}(i)$ is odd) where

$$I_{A_\nu}(i) = \begin{cases} 1 & \text{if } i \in A_\nu \\ 0 & \text{if } i \notin A_\nu \end{cases}, \qquad I_{B_\mu}(i) = \begin{cases} 1 & \text{if } i \in B_\mu \\ 0 & \text{if } i \notin B_\mu \end{cases},$$

and then a little reflection confirms that $f_C(\boldsymbol{\delta}) = \Pi_{\nu=1}^r[f_{A_\nu}(\boldsymbol{\delta})]^{m_\nu}$ and $f_D(\boldsymbol{\delta}) = \Pi_{\mu=1}^s[f_{B_\mu}(\boldsymbol{\delta})]^{l_\mu}$.

By Theorem 2.1 we know

$$E_{\mathscr{S}}\{\textstyle\prod_{\nu=1}^r f_{A_\nu}^{m_\nu})(\prod_{\mu=1}^s f_{B_\mu}^{l_\mu})\} \geq (E_{\mathscr{S}}(\prod_{\nu=1}^r f_{A_\nu}^{m_\nu}))(E_{\mathscr{S}}(\prod_{\mu=1}^s f_{B_\mu}^{l_\mu})), \tag{2.7}$$

and this inequality can be applied to the final expression of (2.6). With this done and then reversing the order of steps involved in (2.6) we obtain the lower estimate

$$\left\{\sum_{m=0}^\infty \frac{1}{m!}\sum\binom{m}{m_1 \cdots m_r}(\textstyle\prod_{\nu=1}^r a_\nu^{m_\nu})E_{\mathscr{S}}(\prod_{\nu=1}^r f_{A_\nu}^{m_\nu})\right\}$$
$$\cdot\left\{\sum_{l=0}^\infty \frac{1}{l!}\sum\binom{l}{l_1 \cdots l_s}(\textstyle\prod_{\mu=1}^s b_\mu^{l_\mu})E_{\mathscr{S}}(\prod_{\mu=1}^s f_{B_\mu}^{l_\mu})\right\}$$
$$= E_{\mathscr{S}}\left\{\sum_{m=0}^\infty \frac{1}{m!}(\textstyle\sum_{\nu=1}^r a_\nu f_{A_\nu})^m\right\}E_{\mathscr{S}}\left\{\sum_{l=0}^\infty \frac{1}{l!}(\sum_{\mu=1}^s b_\mu f_{B_\mu})^l\right\}$$
$$= \{E_{\mathscr{S}}(\exp[\textstyle\sum_{\nu=1}^r a_\nu f_{A_\nu}])\}\{E_{\mathscr{S}}(\exp[\sum_{\mu=1}^s b_\mu f_{B_\mu}])\}.$$

The proof of Proposition 2.1 is complete.

Theorem 2.2. *Let* $a_{ij} \geq 0$; $1 \leq i < j \leq n$; *and* $x_i \geq 0$, $i = 1, \cdots, n$. *Define the function*

$$g(a_{12}, a_{13}, \cdots, a_{n-1,n}, x_1, \cdots, x_n) = \textstyle\sum_{\delta\in\Delta}\exp\left[\sum_{1\leq i<j\leq n} a_{ij}\delta_i\delta_j x_i x_j\right] \tag{2.8}$$

with the set Δ *given in* (2.2). *Then* g *is* MTP_2 *over the positive orthant in all its variables.*

PROOF. The TP_2 property with respect to the variables x_1 and x_2 is tantamount to the inequality

$$\begin{aligned} &g(a_{12}, \cdots, a_{n-1,n}, x_1, \cdots, x_n)g(a_{12}, \cdots, a_{n-1,n}, x_1 + h, x_2 + k, x_3, \cdots, x_n) \\ &\geq g(a_{12}, \cdots, a_{n-1,n}, x_1 + h, x_2, \cdots, x_n)g(a_{12}, \cdots, a_{n-1,n}, x_1, x_2 + k, x_3, \cdots, x_n)\end{aligned} \tag{2.9}$$

with h and k positive. By continuity we may assume x_1 and x_2 positive and without loss of generality we may take $x_i = 1$, $i = 1, \cdots, n$ after substituting $a_{ij}x_i x_j$ for a_{ij} and h/x_1 and k/x_2 for h and k, respectively. Under the convention $a_{21} = a_{12}$ canceling common factors, (2.9), reduces to

$$\begin{aligned} &E_{\mathscr{S}}\{\exp[\delta_1 h \textstyle\sum_{j\neq 1}^n a_{1j}\delta_j + \delta_2 k \sum_{j\neq 2}^n a_{2j}\delta_j + \delta_1\delta_2 hk a_{12}]\} \\ &\geq E_{\mathscr{S}}\{\exp[\delta_1 h \textstyle\sum_{j\neq 1}^n a_{1j}\delta_j]\}E_{\mathscr{S}}\{\exp[\delta_2 k \sum_{j\neq 2}^n a_{2j}\delta_j]\}.\end{aligned} \tag{2.10}$$

Identifying $f_{A_\nu}(\boldsymbol{\delta}) = \delta_1\delta_\nu$, $\nu = 2, 3, \cdots, n$; $f_{B_\mu}(\boldsymbol{\delta}) = \delta_2\delta_\mu$, $\mu = 1, 3, \cdots, n$; $f_{C_1}(\boldsymbol{\delta}) = \delta_1\delta_2$ and applying (2.5) to the left expression of (2.10) the desired inequality results except for the factor $E_{\mathscr{S}}\{\exp[hka_{12}\delta_1\delta_2]\}$.

To deal with this, set $hka_{12} = v \geq 0$. Then,

$$\begin{aligned} E_{\mathscr{S}}\{\exp[hka_{12}\delta_1\delta_2]\} &= E_{\mathscr{S}}\{\exp[\delta_1\delta_2 v]\} \\ &= \textstyle\sum_{k=0}^\infty \frac{1}{k!}E_{\mathscr{S}}[(\delta_1\delta_2 v)^k] = 1 + \sum_{k=1}^\infty \frac{v^k}{k!}E_{\mathscr{S}}[(\delta_1\delta_2)^k].\end{aligned} \tag{2.11}$$

For k even $E_{\mathscr{S}}(\delta_1\delta_2)^k = 1$ whereas for k odd $E(\delta_1\delta_2)^k = E(\delta_1\delta_2) \geq (E\delta_1)(E\delta_2) = 0$. Therefore the quantity in (2.11) exceeds 1 and (2.9) follows.

The proof of the TP_2 property in any of the pairings $(x_i, x_j)(a_{ij}, a_{kl})$ and (x_i, a_{kl}), $1 \leq i < j \leq n$, $1 \leq k < l \leq n$ follows the same lines. The proof of Theorem 2.2 is complete.

THEOREM 2.3. *Let* $\mathbf{X} = (X_1, \cdots, X_n) \sim N(0, \Sigma_1)$ *and* $\mathbf{Y} = (Y_1, \cdots, Y_n) \sim N(0, \Sigma_2)$ *and suppose* $\Sigma_1^{-1} = \Lambda - A$, $\Sigma_2^{-1} = \Lambda - B$ *where* Λ *is a diagonal matrix with elements* $\lambda_1, \cdots, \lambda_n > 0$ *and* $A = \| a_{ij} \|_{i,j=1}^n$, $B = \| b_{ij} \|_{i,j=1}^n$ *satisfy* $a_{ij} \geq b_{ij} \geq 0$, $1 \leq i, j \leq n$ *and* $a_{ii} = b_{ii}$. *Let* $f_{|\mathbf{X}|}$ *and* $f_{|\mathbf{Y}|}$ *denote the densities of* $|\mathbf{X}|$ *and* $|\mathbf{Y}|$, *respectively. Then* (*see* (1.1))

$$f_{|\mathbf{X}|}(\mathbf{x} \vee \mathbf{y}) f_{|\mathbf{Y}|}(\mathbf{x} \wedge \mathbf{y}) \geq f_{|\mathbf{X}|}(\mathbf{x}) f_{|\mathbf{Y}|}(\mathbf{y}) \tag{2.12}$$

for all $\mathbf{x} = (x_1, \cdots, x_n)$, $\mathbf{y} = (y_1, \cdots, y_n)$ *satisfying* $x_i, y_i \geq 0$, $i = 1, \cdots, n$.

PROOF. The joint density of $|\mathbf{X}| = (|X_1|, \cdots, |X_n|)$ at $\mathbf{x} = (x_1, \cdots, x_n)$, $x_i \geq 0$, $i = 1, \cdots, n$ is given by $f_{|\mathbf{X}|}(x_1, \cdots, x_n) = c(\boldsymbol{\lambda}, A) h(\boldsymbol{\lambda}, A, \mathbf{x})$ where

$$h(\boldsymbol{\lambda}, A, \mathbf{x}) = \exp\left\{\frac{1}{2}[-\sum_{i=1}^n \lambda_i x_i^2 + \sum_{i=1}^n a_{ii} x_i^2]\right\} \sum_{\delta\epsilon\Delta} \exp[\sum_{1\leq i<j\leq n} a_{ij}\delta_i\delta_j x_i x_j]. \tag{2.13}$$

By Theorem 2.2 $h(\boldsymbol{\lambda}, A, \mathbf{x})$ *is* MTP_2 *in all components of* A *and* $\mathbf{x}$ ($\boldsymbol{\lambda}$ fixed), implying

$$h(\boldsymbol{\lambda}, A, \mathbf{x} \vee y) h(\boldsymbol{\lambda}, B, \mathbf{x} \wedge \mathbf{y}) \geq h(\boldsymbol{\lambda}, A, \mathbf{x}) h(\boldsymbol{\lambda}, B, \mathbf{y}). \tag{2.14}$$

Multiplying both sides above by $c(\boldsymbol{\lambda}, A) \cdot c(\boldsymbol{\lambda}, B)$, the inequality passes into (2.12).

Theorem 5.2 later presents an extension of this result, in which the assumption $a_{ij} \geq 0$ and $b_{ij} \geq 0$ is relaxed.

3. The Principal Theorem for Absolute Value Multinormal Variables. We are now prepared to prove

THEOREM 3.1 *Let* $\mathbf{X} \sim N(0, \Sigma)$. *A necessary and sufficient condition for the density of* $(|X_1|, \cdots, |X_n|)$ *to be* MTP_2 *is that there exists a diagonal matrix* D *with elements* ± 1 *such that the off-diagonal elements of* $-D\Sigma^{-1}D$ *are all nonnegtive.*

REMARK 3.1 Notice the invariance of the distribution of $|\mathbf{X}| = (|X_1|, \cdots, |X_n|)$ under the linear transformation $\mathbf{X}' = D\mathbf{X}$ where D is a diagonal matrix $\mathrm{diag}(d_1, d_2, \cdots, d_n)$ with each $d_i = +1$ or -1. Indeed, such transformations merely change signs in some of the $X_1, \cdots, X_n$. Thus, the densities of $(|X_1'|, \cdots, |X_n'|)$ for the whole class of covariance matrices $D\Sigma D$ with D diagonal as above are identical.

PROOF OF THEOREM 3.1. *Sufficiency*: On account of Remark 3.1, we can assume that the off diagonal elements of $-\Sigma^{-1}$ are nonnegtive. Sufficiency then follows on the basis of Theorem 2.3 (taking $A = B$).

Necessity. It is required to prove that if $(|X_1|, \cdots, |X_n|)$ follow a joint MTP_2 density then we can construct $D = \mathrm{diag}(\epsilon_1, \cdots, \epsilon_n)(\epsilon_i = \pm 1)$ such that the off-diagonal terms of $-D\Sigma^{-1}D$ are all nonnegative. Scrutiny of the construction reveals that the determination of D is essentially unique.

By relabeling indices when necessary or equivalently permuting suitably rows and columns we can assume for reasons given below that for some $r < n$ the first $r + 1$ rows of $-\Sigma^{-1}$ are of the form $A = \| a_{ij} \|$, $i = 1, \cdots, r + 1$, $j = 1, \cdots, n$; where

$$\begin{aligned} &a_{12}, \cdots, a_{1k_1} \neq 0 \quad \text{and} \quad a_{1i} = 0 \quad \text{for} \quad i > k_1 \\ &a_{2,k_1+1}, \cdots, a_{2,k_2} \neq 0 \quad \text{and} \quad a_{2i} = 0 \quad \text{for} \quad i > k_2 \\ &\vdots \\ &a_{r+1,k_r+1}, \cdots, a_{r+1,n} \neq 0 \quad \text{for some} \quad r < k_r < n. \end{aligned} \tag{3.1}$$

We may assume $k_i > i$ since if $k_i = i$, then $X_1, \cdots, X_i$ and $X_{i+1}, \cdots, X_n$ are independent and can be treated separately by means of induction.

We now construct $D = \mathrm{diag}(\epsilon_1, \cdots, \epsilon_n)$. Since D and $-D$ yield the same transformation we can fix $\epsilon_1 = 1$. Then the *unique* possible choice of the remaining ϵ_j's brings the first nonzero term of each column of $-D\Sigma^{-1}D$ to be positive explicitly as follows

$$\begin{aligned} &\epsilon_j = \mathrm{sgn}\ a_{1j}, j = 2, \cdots, k_1; \quad \epsilon_j = \mathrm{sgn}\ \epsilon_2 a_{2j}, \quad j = k_1 + 1, \cdots, k_2; \\ &\epsilon_j = \mathrm{sgn}\ \epsilon_3 a_{3j}, j = k_2 + 1, \cdots, k_3; \cdots; \epsilon_j = \mathrm{sgn}\ \epsilon_{r+1} a_{r+1,j}, \quad j = k_r + 1, \cdots, n. \end{aligned} \tag{3.2}$$

It is required to show that the off-diagonal elements of $-D\Sigma^{-1}D$, i.e., $a_{ij}\epsilon_i\epsilon_j$, $1 \leqslant i < j \leqslant n$ are all nonnegative. Assume to the contrary, that $a_{ij}\epsilon_i\epsilon_j < 0$ for some (i, j) and consider the "first" such negative term (in lexicographic ordering), i.e., let $l < k$ be such that

$$a_{lk}\epsilon_l\epsilon_k < 0,\ a_{ij}\epsilon_i\epsilon_j \geqslant 0 \quad \text{for all } i < j \text{ such that either } i \leqslant l-1 \text{ or } i = l, j \leqslant k-1. \tag{3.3}$$

Now set $x_1 = \cdots = x_{l-1} = v$, $x_l > 0$, $x_k > 0$, and $x_i = 0$ for $l < i \neq k$. We claim that for large enough v the function

$$\sum_{\delta\epsilon\Delta} \exp[\sum_{1\leqslant i<j\leqslant n} a_{ij}\delta_i\delta_j x_i x_j] \tag{3.4}$$

(where we use the notation $\Delta = \{\boldsymbol{\delta} = (\delta_1, \cdots, \delta_n), \delta_i = \pm 1\}$) is not TP_2 in the variables x_l, x_k which violates the MTP_2 hypothesis for the density of $(|X_1|, \cdots, |X_n|)$. This contradiction can only be averted provided $a_{ij}\epsilon_i\epsilon_j \geqslant 0$ for all $i < j$ as needed to be shown. Since $x_i = 0$ for $l < i \neq k$ the quantity in (3.4) is equal to

$$g_v(x_l, x_k) = \sum_{\delta\in\Delta} \exp[\sum_{1\leqslant i<j\leqslant l} a_{ij}\delta_i\delta_j x_i x_j + \sum_{1\leqslant i\leqslant l-1} a_{ik}\delta_i\delta_k x_i x_k + a_{lk}\delta_l\delta_k x_l x_k] \tag{3.5}$$

and it suffices to show that for some $x_l < x_l'$, $x_k < x_k'$ (and large enough v)

$$g_v(x_l, x_k)\, g_v(x_l', x_k') < g_v(x_l, x_k')\, g_v(x_l', x_k). \tag{3.6}$$

Letting $v \to \infty$, the leading terms in (3.5) are those where all coefficients involving v are nonnegative i.e.,

$$a_{ij}\delta_i\delta_j \geqslant 0 \quad \text{for all} \quad 1 \leqslant i < j \leqslant l \quad \text{and} \quad a_{ik}\delta_i\delta_k \geqslant 0 \quad \text{for all} \quad 1 \leqslant i \leqslant l-1. \tag{3.7}$$

It follows that the ratio of any other term in the summation of (3.5) and a term satisfying (3.7) tends to zero as $v \to \infty$. By the construction in (3.2) and (3.3), the case $\boldsymbol{\delta} = (\epsilon_1, \cdots, \epsilon_n)$ obeys (3.7).

In view of the foregoing discussion, by dividing $g_v(\cdot, \cdot)$ by the term $\exp[\sum_{1\leqslant i<j\leqslant l} a_{ij}\epsilon_i\epsilon_j x_i x_j + \sum_{1\leqslant i\leqslant l-1} a_{ik}\epsilon_i\epsilon_k x_i x_k]$ the inequality (3.6) reduces to

$$(\alpha(v) + qe^{cx_lx_k})(\beta(v) + qe^{cx_l'x_k'}) < (\gamma(v) + qe^{cx_lx_k'})(\lambda(v) + qe^{cx_l'x_k}) \tag{3.8}$$

where $c = a_{lk}\epsilon_l\epsilon_k < 0$, q denotes the number of $\boldsymbol{\delta} \in \Delta$ satisfying (3.7), and $\alpha(v)$, $\beta(v)$,, $\gamma(v)$, $\lambda(v) \to 0$ as $v \to \infty$.

Since $c < 0$, e^{cxy} is SRR_2 (strict reverse rule, Karlin (1968)) in x, y, i.e., $e^{cxy}e^{cx'y'} < e^{cxy'}e^{cx'y}$ for $x < x'$, $y < y'$ so that (3.8) indeed prevails for sufficiently large v.

The proof of Theorem 3.1 is complete.

REMARK 3.2. Set $\Sigma^{-1} = \Lambda - A$, Λ diagonal and $A = \| a_{ij} \|_{i,j=1}^n$.

(a) If there exists D as in Theorem 3.1, i.e., the off-diagonal elements of $-D\Sigma^{-1}D$ are nonnegative then for all $1 \leqslant i < j < k \leqslant n$, $(a_{ij}a_{ik}a_{jk}) \geqslant 0$.

PROOF. Let $\epsilon_1, \cdots, \epsilon_n$ be the diagonal elements of D. Then

$$(a_{ij}a_{ik}a_{jk}) = (\epsilon_i^2\epsilon_j^2\epsilon_k^2)(a_{ij}a_{ik}a_{jk}) = (\epsilon_i\epsilon_j a_{ij})(\epsilon_i\epsilon_k a_{ik})(\epsilon_j\epsilon_k a_{jk}) \geqslant 0.$$

(b) Suppose Σ^{-1} has (at least) one row in which none of the entries vanish, say, row 1 for definiteness. Then the condition

$$(a_{ij}a_{ik}a_{jk}) \geqslant 0, \quad \text{for all} \quad 1 \leqslant i < j < k \leqslant n \tag{3.9}$$

implies that there exists a diagonal matrix D with diagonal entries ± 1 such that $-D\Sigma^{-1}D$ has nonnegative off-diagonal entries.

PROOF. Consultation of the construction of the necessity part of Theorem 3.1 shows that the elements of $D = \text{diag}(d_1, d_2, \cdots, d_n)$ are explicitly $d_1 = 1$, $d_i = \text{sign } a_{1i}$, $i = 2, \cdots, n$.

(c) Theorem 3.1′ (see Section 1) follows by a similar construction with b_{ij}^* replacing a_{ij}.

The matrix

$$\Sigma^{-1} = \begin{pmatrix} \lambda_1 & -b & -b & 0 \\ -b & \lambda_2 & 0 & -b \\ -b & 0 & \lambda_3 & b \\ 0 & -b & b & \lambda_4 \end{pmatrix} \tag{3.10}$$

provides an example where (3.9) holds, however, there does not exist a suitable D such that $-D\Sigma^{-1}D$ has nonnegative off-diagonal entries. Therefore, for the inverse covariance matrix of (3.10), $|\mathbf{X}| = (|X_1|, |X_2|, |X_3|, |X_4|)$ is not MTP_2.

An interesting matrix consequence of the necessity assertion of Theorem 3.1 is the following

COROLLARY 3.1. *Let* $\mathbf{X} \sim N(0, \Sigma)$ *have* $|\mathbf{X}| = (|X_1|, |X_2|, \cdots, |X_n|)$ MTP_2. *Suppose* Σ *has all nonnegative elements, then for any partitioning of*

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \quad \textit{where} \quad \Sigma_{11} \textit{ is } k \times k, \tag{3.11}$$

the $k \times k$ *matrix*

$$\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} \tag{3.12}$$

exhibits only nonnegative elements and D *yielding* (1.12) *in this case is the identity matrix.*

In the more general case that $D = \text{diag}(d_1, \cdots, d_n)$, $d_i = \pm 1$, renders $-D\Sigma^{-1}D$ with nonnegative off-diagonal elements then for $D_k = \text{diag}(d_1, \cdots, d_k)$ the reduction of D to the first k rows and columns, the matrix

$$D_k(\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})D_k \tag{3.13}$$

exhibits only nonnegative elements.

PROOF. We partition the inverse matrix

$$\Sigma^{-1} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

corresponding to (3.11). When $-D\Sigma^{-1}D$ exhibits nonnegative off-diagonal elements, then plainly

$$-D_k A_{11} D_k \quad \text{has nonnegative off-diagonal entries} \tag{3.14}$$

where D_k is the contraction of D to the first k rows and columns. But A_{11} is the inverse covariance matrix of $X_1, X_2, \cdots, X_k$ conditioned on $X_{k+1}, \cdots, X_n$, whose covariance matrix is (3.12). The property (3.14) implies that the vector of absolute values of these conditional normal variables is MTP_2. We refer to Remark 1.3 to deduce that (3.13) involves only nonnegative elements.

4. Classes of examples of multinormal densities whose absolute components induce a MTP_2 density. We develop classes of examples of covariance matrices Σ which satisfy the conditions of Theorem 3.1. We first discuss Example 4.2 which extends Example 4.1 (both examples are described in the introduction).

ANALYSIS OF EXAMPLE 4.2. A standard analysis reveals that $\mathbf{Y} \sim N(0, \Sigma)$ with

$$\Sigma = \Lambda + \begin{pmatrix} R_{11} & \rho R_{12} \\ \rho R_{21} & R_{22} \end{pmatrix} \tag{4.1}$$

exhibiting the second part in block form composed from the rank 1 matrices

$$R_{11} = \| \lambda_i \lambda_j \|_{i,j=1}^{k} \qquad R_{12} = \| \lambda_i \lambda_\beta \|_{i=1,\ \beta=k+1}^{k,\ n} \tag{4.2}$$

$$R_{21} = R'_{12} = \| \lambda_\alpha \lambda_j \|_{\alpha=k+1,\ j=1}^{n,\ k}, \qquad R_{22} = \| \lambda_\alpha \lambda_\beta \|_{\alpha,\beta=k+1}^{n}.$$

It is convenient for later purposes to use the notation

$$\Lambda = \begin{pmatrix} \Lambda^* & 0 \\ 0 & \Lambda_* \end{pmatrix} \text{ where } \Lambda^* \text{ is } k \times k \text{ and } \Lambda_* \text{ is } n-k \times n-k$$

(4.3) and

$$\lambda^* = \textstyle\sum_{i=1}^{k} \gamma_i^{-1} \lambda_i^2, \qquad \lambda_* = \sum_{i=k+1}^{n} \gamma_i^{-1} \lambda_i^2.$$

It is easy to verify that Σ^{-1} has the form

$$\Sigma^{-1} = \Lambda^{-1} - \begin{pmatrix} a\Lambda^{*-1} R_{11} \Lambda^{*-1} & b\Lambda^{*-1} R_{12} \Lambda_*^{-1} \\ b\Lambda_*^{-1} R_{21} \Lambda^{*-1} & c\Lambda_*^{-1} R_{22} \Lambda_*^{-1} \end{pmatrix} \tag{4.4}$$

where we determine a and b from the pair of linear equations: $1 = a(1 + \lambda^*) + b\rho\lambda_*$, $\rho = a\rho\lambda^* + b(1 + \lambda_*)$ yielding

$$a = \frac{(1 + (1-\rho^2)\lambda_*)}{(1+\lambda^*)(1+\lambda_*) - \rho^2 \lambda^* \lambda_*}, \qquad b = \frac{\rho}{(1+\lambda^*)(1+\lambda_*) - \rho^2 \lambda^* \lambda_*}.$$

Similarly, we have

$$c = \frac{1 + (1-\rho^2)\lambda^*}{(1+\lambda^*)(1+\lambda_*) - \rho^2 \lambda^* \lambda_*}.$$

These determinations reveal that

$$a > 0, \quad c > 0 \quad \text{and} \quad b \text{ has the sign of } \rho. \tag{4.5}$$

We now construct the diagonal matrix D, $D = \operatorname{diag}(d_1, d_2, \cdots, d_n)$ with

$$d_i = \begin{cases} \operatorname{sign} \lambda_i, & i = 1, 2, \cdots, k \\ \operatorname{sign} b\lambda_i, & i = k+1, \cdots, n. \end{cases} \tag{4.6}$$

By virtue of the representation (4.4), and the fact of (4.5), we find that $-D\Sigma^{-1}D$ has nonnegative off-diagonal elements.

EXAMPLE 4.3. We extend Example 4.2 to the case of three groups

$$\mathbf{Y} = \mathbf{X} + (\lambda_1 U, \cdots, \lambda_k U, \lambda_{k+1} V, \cdots, \lambda_{k+l} V, \lambda_{k+l+1} W, \cdots, \lambda_n W) \tag{4.7}$$

where $\mathbf{X} \sim N(0, \Lambda)$ as in Example 4.2, and the 3-component random vector (U, V, W) is distributed following $N(0, S)$;

$$S = \begin{pmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_3 \\ \rho_2 & \rho_3 & 1 \end{pmatrix}. \tag{4.8}$$

The covariance matrix of $\mathbf{Y}$ is simply

$$\Sigma = \Lambda + \begin{pmatrix} R_{11} & \rho_1 R_{12} & \rho_2 R_{13} \\ \rho_1 R_{21} & R_{22} & \rho_3 R_{23} \\ \rho_2 R_{31} & \rho_3 R_{32} & R_{33} \end{pmatrix} \tag{4.9}$$

with

$$R_{11} = \| \lambda_i \lambda_j \|_{i,j=1}^{k} \qquad R_{12} = \| \lambda_i \lambda_\beta \|_{i=1,\beta=k+1}^{k,\ k+l} \qquad R_{13} = \| \lambda_i \lambda_b \|_{i=1,b=k+l+1}^{k,\ n} \tag{4.10}$$

$$R_{21} = R'_{12}, \qquad R_{22} = \| \lambda_\alpha \lambda_b \|_{\alpha,\beta=k+1}^{k+l} \qquad R_{23} = \| \lambda_\alpha \lambda_b \|_{\alpha=k+1,b=k+l+1}^{k+l,\ n}$$

$$R_{31} = R'_{13} \qquad R_{32} = R'_{23} \qquad R_{33} = \| \lambda_a \lambda_b \|_{a,b=k+l+1}^{n}.$$

Some direct manipulations establish the inverse of Σ of the form

$$\Sigma^{-1} = \Lambda^{-1} - \begin{pmatrix} aR^*_{11} & bR^*_{12} & cR^*_{13} \\ bR^*_{21} & dR^*_{22} & eR^*_{23} \\ cR^*_{31} & eR^*_{32} & fR^*_{33} \end{pmatrix} \tag{4.11}$$

where $R^*_{ij} = \Lambda_i^{-1} R_{ij} \Lambda_j^{-1}$ and $\Lambda = \operatorname{diag}(\Lambda_1, \Lambda_2, \Lambda_3)$ entailing also

$$a > 0, \qquad d > 0, \qquad f > 0. \tag{4.12}$$

Apart from a positive factor we routinely determine

$$\operatorname{sign}(b \cdot c \cdot e) = \operatorname{sign}[\rho_1 + (\rho_1 - \rho_2\rho_3)\mu_3][\rho_2 + (\rho_2 - \rho_1\rho_3)\mu_2][\rho_3 + (\rho_3 - \rho_1\rho_2)\mu_1] \tag{4.13}$$

with

$$\mu_1 = \textstyle\sum_1^k \gamma_i^{-1}\lambda_i^2, \qquad \mu_2 = \sum_{k+1}^{k+l} \gamma_i^{-1}\lambda_i^2, \qquad \mu_3 = \sum_{k+l+1}^{n} \gamma_i^{-1}\lambda_i^2.$$

The evaluations in (4.11) lead us to the proposition:

PROPOSITION 4.1. *Consider the multinormal random vector* $\mathbf{Y}$ *as in* (4.7). *Then* $|\mathbf{Y}| = (|Y_1|, |Y_2|, \cdots, |Y_n|)$ *is* MTP_2 *if and only if the* sign (4.13) *is* +1.

PROOF. Following the dictates of Theorem 3.1 we need to find a diagonal matrix D implying that $-D\Sigma^{-1}D$ exhibits nonnegative off-diagonal elements. When λ_i are all nonzero focusing on row 1 of $-D\Sigma^{-1}D$, the diagonal entries of D are mandated to be

$$d_i = \begin{cases} \operatorname{sign} \lambda_i & i = 1, 2, \cdots, k \\ \operatorname{sign} \lambda_i b & i = k+1, \cdots, k+l \\ \operatorname{sign} \lambda_i c & i = k+l+1, \cdots, n. \end{cases}$$

With the information of (4.12) the condition $b \cdot c \cdot e > 0$ of (4.13) maintains consistently the property that all off-diagonal elements in $-D\Sigma^{-1}D$ are nonnegative. The argument works in both directions.

REMARK 4.1. Observe that for μ_1, μ_2, and μ_3 small, the positivity criterion (4.13) essentially reduces to $\rho_1\rho_2\rho_3 \geqslant 0$, while for μ_1, μ_2, and μ_3 all large, the positivity criterion is that the inverse matrix S^{-1} (of (4.8)) carries a negative product for off-diagonal elements.

REMARK 4.2. The recipe of examples 4.2 and 4.3 can be routinely extended to any number of groups in the formation of $\mathbf{Y}$. The analog of Proposition 4.1 essentially states that if Σ has the representation $\Sigma^{-1} = \Lambda^{-1} - \| a_{ij} R^*_{ij} \|_{i,j=1}^m$ where m is the number of groups (compare with (4.11)) then Σ corresponds to an MTP_2 absolute value normal distribution provided that there exists an $m \times m$ diagonal matrix D such that the off-diagonal elements of $D\| a_{ij} \|D$ are nonnegative.

EXAMPLE 4.4. Consider an $n \times n$ covariance matrix of the form $\Sigma = \| \sigma_{ij} \|$ where $\sigma_{ij} = c_{|i-j|}$. It is elementary to check that the inverse matrix $\Sigma^{-1} = A = \| a_{ij} \|$ inherits the same form $a_{ij} = a_{|i-j|}$.

Suppose $c_0 \geqslant c_1 \geqslant c_2 \geqslant \cdots \geqslant c_{n-1}$. An examination of cases reveals that in order that Σ satisfies (1.12) in the cases $n = 3$ and 4 we need for $n = 3$, $c_1^2 \leqslant c_0c_2$, and for $n = 4$, $c_0c_2 - c_1^2 > c_1c_3 - c_2^2 > 0$.

An explicit condition for general n remains unresolved.

EXAMPLE 4.5. Let $\Sigma = I + \rho B$ where B is a positive definite matrix of positive elements. Then for ρ *positive and small enough* $\Sigma^{-1} \approx I - \rho B$ and all off-diagonal elements are negative. Here the criterion of (1.12) works with $D = I$.

EXAMPLE 4.6. The matrix $\Sigma = \| \sigma_{ij} \|$, with $\sigma_{ij} = a_{\min(i,j)} b_{\max(i,j)}$, $i, j = 1, \cdots, n$ where $a_i b_j > 0$ and $a_1/b_1 < a_2/b_2 < \cdots < a_n/b_n$ is called a Green's matrix. Its inverse is a Jacobi (tridiagonal) matrix

$$\Sigma^{-1} = J = \begin{pmatrix} c_1 & d_1 & 0 & \cdots & & 0 \\ d_1 & c_2 & d_2 & & & \vdots \\ 0 & d_2 & c_3 & \ddots & & 0 \\ \vdots & & \ddots & \ddots & \ddots & d_{n-1} \\ 0 & 0 & 0 \cdots & 0 & d_{n-1} & c_n \end{pmatrix} \tag{4.14}$$

displaying zeros off the two contiguous diagonals to the main diagonal. It is known (see Karlin (1968, pages 110–114)) that under the above conditions $d_i < 0$, $i = 1, \cdots, n-1$. Therefore, the covariance matrix Σ corresponds to a random vector $\mathbf{X}$ having $|\mathbf{X}|$ MTP_2.

5. Some association inequalities for absolute value multinormal random variables. Corollary 1.1 implies:

THEOREM 5.1. *Suppose* $\mathbf{X} \sim N(0, \Sigma)$ *satisfying the conditions of Theorem* 3.1. *Then*

$$E[\varphi(|\mathbf{X}|)\psi(|\mathbf{X}|)] \geq E[\varphi(|\mathbf{X}|)]E[\psi(|\mathbf{X}|)] \tag{5.1}$$

for any pair of increasing (*or decreasing*) *functions* φ *and* ψ *of* $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_n)$, $\xi_i \geq 0$.

REMARK 5.1. In particular, the inequalities (1.16) and (5.1) hold if Σ conforms with one of the classes of Section 4. A special case of (5.1) is a result of Khatri (1967) who stipulated $\Sigma = \Lambda + R$ as in Example 4.1. Then

$$P\{|X_1| \geq c_1, |X_2| \geq c_2, \cdots, |X_n| \geq c_n\} \geq \textstyle\prod_{k=1}^{n} P(|X_k| \geq c_k). \tag{5.2}$$

Inequality (1.16) clearly also applies in the case $A = \{\mathbf{x}: |x_1| \leq c_1, |x_2| \leq c_2, \cdots, |x_k| \leq c_k\}$ and $B = \{\mathbf{x}: |x_{k+1}| \leq c_{k+1}, \cdots, |x_n| \leq c_n\}$ (cf. Das Gupta et al. (1972)). Pitt (1977) confirmed (1.16) in dimension $n = 2$ provided A and B are balanced and convex, for any covariance matrix Σ.

The next application exploits the MTP_2 property simultaneously in the covariance parameters and the random variables, i.e., with respect to the components of $A = \| a_{ij} \|_{i,j=1}^{n}$ and $\{x_i\}_1^n$, as in Theorem 2.3.

THEOREM 5.2. *Suppose* $\mathbf{X} \sim N(0, \Sigma_1)$ *and* $\mathbf{Y} \sim N(0, \Sigma_2)$ *where* $\Sigma_1^{-1} = \Lambda - A$, $\Sigma_2^{-1} = \Lambda - B$ *both satisfy the conditions of Theorem* 3.1. *Suppose* $|a_{ij}| \geq |b_{ij}|$, $1 \leq i < j \leq n$ *and* $a_{ii} = b_{ii}$. *Then*

$$f_{|\mathbf{X}|}(\mathbf{x} \vee \mathbf{y}) f_{|\mathbf{Y}|}(\mathbf{x} \wedge \mathbf{y}) \geq f_{|\mathbf{X}|}(\mathbf{x}) f_{|\mathbf{Y}|}(\mathbf{y}) \tag{5.3}$$

for all $\mathbf{x} = (x_1, \cdots, x_n)$, $\mathbf{y} = (y_1, \cdots, y_n)$, $x_i, y_i \geq 0$, $i = 1, \cdots, n$ *where* $f_{|\mathbf{X}|}$ *and* $f_{|\mathbf{Y}|}$ *denote the densities of* $|\mathbf{X}|$ *and* $|\mathbf{Y}|$ *respectively. Let* $g(x_1, \cdots, x_n)$ *be an increasing function of* $x_1, \cdots, x_n$, $x_i \geq 0$, $i = 1, \cdots, n$. *Then*

$$E[g(|X_1|, \cdots, |X_n|)] \geq E[g(|Y_1|, \cdots, |Y_n|)]. \tag{5.4}$$

The inequality (5.4) *is reversed if* g *is decreasing.*

PROOF. By Remark 3.1 we can assume without loss of generality that $a_{ij} \geq b_{ij} \geq 0$, $1 \leq i < j \leq n$. The relation (5.3) is a consequence of Theorem 2.3, and (5.4) follows from Fact 1.2.

6. Some Applications to the Wishart and t distributions. If $(|Z_1|, \cdots, |Z_n|)$ have a joint MTP_2 density function, then so does $(Z_1^2, \cdots, Z_n^2)$, since applying a separate increasing function to the arguments, and adjusting the density by a product of factors as $\prod_{i=1}^n g(z_i)$ does not affect the MTP_2 property.

We know that the MTP_2 property for the random variables $(Z_1^2, \cdots, Z_n^2)$ entails that they are associated random variables as well (see (1.6)).

Let $(Z_{1\nu}, \cdots, Z_{n\nu})$ $\nu = 1, \cdots, p$ be a random sample such that each vector $(|Z_{1\nu}|, \cdots, |Z_{n\nu}|)$ possesses a MTP_2 density. Recall that independent random variables are associated (property P_2 of Esary et al. (1967)) and increasing functions of associated random variables are again associated (property P_4 of [loc. cit.]). Accordingly, if we define

$$S_i = (\textstyle\sum_{\nu=1}^p Z_{i\nu}^2)^{1/2} \qquad i = 1, \cdots, n \tag{6.1}$$

it follows that $S_1, \cdots, S_n$ are associated.

Recall that if $(Z_{1\nu}, \cdots, Z_{n\nu}) \sim N(0, \Sigma)$ then $S_1, \cdots, S_n$ have the distribution of the diagonal elements of a random matrix $\mathbf{S}$, where $\mathbf{S} \sim W_n(p, \Sigma)$. (A Wishart distribution with p degrees of freedom and parameter Σ). With these facts in hand, we have

THEOREM 6.1. *If* $\mathbf{S} \sim W_n(p, \Sigma)$ *and* Σ *corresponds to a* MTP_2 *absolute value multinormal vector then the diagonal elements of* $\mathbf{S}$, *that is* $S_1, \cdots, S_n$ *are associated.*

In particular,

$$P(S_i \geqslant c_i, i = 1, \cdots, n) \geqslant \textstyle\prod_{i=1}^n P(S_i \geqslant c_i). \tag{6.2}$$

This result was obtained by Jogdeo (1977) for the cases that Σ is as in Examples 4.1 and 4.2. For Σ as in Example 4.1 the inequality (6.2) was previously obtained by Das Gupta et al. (1972).

Application to a multivariate t distribution. Let $\mathbf{X} = (X_1, \cdots, X_n) \sim N(0, C)$. For any covariance matrix C, Sidak (1967) proved

$$P(|X_1| \leqslant c_1, \cdots, |X_n| \leqslant c_n) \geqslant \textstyle\prod_{i=1}^n P(|X_i| \leqslant c_i). \tag{6.3}$$

Let $S_1, \cdots, S_n$ be positive, associated random variables and independent of $\mathbf{X}$. Then

$$\begin{aligned} &P(|X_1|/S_1 \leqslant c_1, \cdots, |X_n|/S_n \leqslant c_n) \\ &= E\{P(|X_1|/S_1 \leqslant c_1, \cdots, |X_n|/S_n \leqslant c_n \mid S_1, \cdots, S_n\} \\ &\geqslant E\{\textstyle\prod_{i=1}^n P(|X_i|/S_i \leqslant c_i \mid S_i)\} \\ &= E\left[\textstyle\prod_{i=1}^n \frac{1}{\sigma_i}\left(\frac{2}{\pi}\right)^{1/2} \int_0^{c_i S_i} \exp(-\xi^2/2\sigma_i^2)\, d\xi\right] = E[\textstyle\prod_{i=1}^n h_i(S_i)] \end{aligned}$$

where h_i are obviously monotone increasing in S_i. Now since $\{S_i\}$ are associated by assumption (see (1.7)), the final quantity exceeds or equals

$$\textstyle\prod_{i=1}^n E\{P(\,|X_i|\,/S_i \leqslant c_i \mid S_i)\} = \prod_{i=1}^n P(\,|X_i|\,/S_i \leqslant c_i).$$

We thus obtain

$$P(\,|X_i|\,/S_i \leqslant c_i, i = 1, 2, \cdots, n) \geqslant \textstyle\prod_{i=1}^n P(\,|X_i|\,/S_i \leqslant c_i). \tag{6.4}$$

The case where $S_1, \cdots, S_n$ are generated by (6.1) with $(Z_{1\nu}, \cdots, Z_{n\nu}) \sim N(0, \Sigma)$ and Σ as in Example 4.1, and some further extensions, were given by Sidak (1971).

Applying similar methods we can obtain results on association and related inequalities regarding other versions of the multivariate t distribution, and for various classes of the multivariate F distribution.

REFERENCES

ABDEL-HAMEED, M. and SAMPSON, A. R. (1978). Positive dependence of the bivariate and trivariate absolute normal, t, χ^2 and F distributions. *Ann. Statist.* **6** 1360–1368.

BARLOW, R. E. and PROSCHAN, F. (1975). *Statistical Theory of Reliability and Life Testing.* Holt, Rinehart and Winston, New York.

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 eliptically contoured distributions. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* **2** 241-265, Univ. of Calif., Berkeley.

DYKSTRA, R. L. (1979). Product inequalities involving the multivariate normal distlribution. Tech. Report No. 85, Dept. Statist., Univ. Missouri-Columbia.

ESARY, J. D. and PROSCHAN, F. (1972). Relationships among some notions of bivariate dependence. *Ann. Math. Statist.* **43** 651-655.

ESARY, J. D., PROSCHAN, F. and WALKUP, D. W. (1967). Association of random variables with applications. *Ann. Math. Statist.* **38** 1466-1474.

FORTUIN, C. M., GINIBRE, J. and KASTELEYN, P. W. (1971). Correlation inequalities on some partially ordered sets. *Comm. Math. Phys.* **22** 89-103.

HOLLEY, R. (1974). Remarks on the FKG inequalities. *Comm. Math. Phys.* **36** 227-231.

JOGDEO, K. (1977). Association and probability inequalities. *Ann. Statist.* **5** 495-504.

KARLIN, S. (1968). *Total Positivity.* Stanford Univ. Press.

KARLIN, S. and RINOTT, Y. (1980). Classes of orderings of measures and related correlation inequalities. I. Multivariate totally positive distributions. *J. Multivariate Anal.* **10** 467-498.

KELLY, D. G. and SHERMAN, S. (1968). General Griffiths' inequalities on correlations in Ising ferromagnets. *J. Mathematical Phys.* **9** 466-484.

KEMPERMAN, J. H. B. (1977). On the FKG inequality for measures on partially ordered space. *Indag. Math.* **39** 313-331.

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

LEHMANN, E. L. (1966). Some concepts of dependence. *Ann. Math. Statist.* **37** 1137-1153.

LORENTZ, G. G. (1953). An inequality for rearrangement. *Amer. Math. Monthly* **60** 176-179.

PERLMAN, M. D. AND OLKIN, I. (1978). Unbiasedness of invariant tests for MANOVA and other multivariate problems. *Ann. Statist.* **8** 1326-1341.

PITT, L. D. (1977). A Gaussian correlation inequality for symmetric convex sets. *Ann. Probability* **5** 470-474.

PRESTON, C. J. (1974). A generalization of the FKG inequalities. *Comm. Math. Phys.* **36** 233-241.

RINOTT, Y. (1973). Multivariate majorization and rearrangement inequalities with some applications to probability and statistics. *Israel J. Math.* **15** 60-77.

SARKAR, T. K. (1969). Some lower bounds of reliability. Tech. Report No. 124, Dept. of Operations Res. Statist., Stanford Univ.

SHAKED, M. (1975). On concepts of dependence for multivariate distributions. Ph.D. dissertation, Univ. Rochester.

SIDAK, Z. (1971). On probabilities of rectangles in multivariate student distributions: Their dependence on correlations. *Ann. Math. Statist.* **42** 169-175.

SIDAK, Z. (1973). On probabilities in certain multivariate distributions: Their dependence on correlations. *Appl. Mat.* **18** 128-135.

DEPARTMENT OF MATHEMATICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305

DEPARTMENT OF MATHEMATICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305
AND
DEPARTMENT OF STATISTICS
HEBREW UNIVERSITY
JERUSALEM, ISRAEL