Ann. Inst. Statist. Math.
Vol. 53, No. 3, 567–575 (2001)

STOCHASTIC ORDERING OF MULTIVARIATE NORMAL DISTRIBUTIONS

ALFRED MÜLLER

Institut für Wirtschaftstheorie und Operations Research, Universität Karlsruhe, Geb. 20.21, D-76128 Karlsruhe, Germany,
email: mueller@wior.uni-karlsruhe.de

(Received December 6, 1999)

Abstract. We show an interesting identity for $Ef(\boldsymbol{Y}) - Ef(\boldsymbol{X})$, where $\boldsymbol{X}, \boldsymbol{Y}$ are normally distributed random vectors and f is a function fulfilling some weak regularity condition. This identity will be used for a unified derivation of sufficient conditions for stochastic ordering results of multivariate normal distributions, some well known ones as well as some new ones. Moreover, we will show that many of these conditions are also necessary. As examples we will consider the usual stochastic order, convex order, upper orthant order, supermodular order and directionally convex order.

Key words and phrases: Multivariate normal distribution, stochastic orders, supermodular order, directionally convex order.

1. Introduction

Stochastic orderings are an important tool for many problems in probability and statistics. This has been demonstrated in the monographs of Stoyan (1983), Shaked and Shanthikumar (1994) and Szekli (1995). Surprisingly, however, in none of these monographs a detailed study of the case of normally distributed vectors can be found. For most of the important stochastic order relations some sufficient conditions are known for the case of normal distributions, but they are scattered in the literature. Some results can be found e.g. in Bäuerle (1997), Mosler (1984), Scarsini (1998) and especially in the books of Tong (1980, 1990). Necessary conditions, however, are very hard to find.

It is the aim of this paper to fill this gap. We will give necessary and sufficient conditions for many important examples of so called *integral stochastic orders* (we refer to Müller (1997) for a general treatment of these stochastic order relations). The main tool in our investigation will be an identity for $Ef(\boldsymbol{Y}) - Ef(\boldsymbol{X})$, where $\boldsymbol{X}, \boldsymbol{Y}$ are normally distributed random vectors. This identity will be derived from an extension of Plackett's identity (see Tong (1990), p. 191) and a double use of partial integration. It should be of interest in its own. The idea of using this method for proving stochastic inequalities for multivariate normal distributions is not new. Tong (1980) has used a similar idea to prove Slepian's inequality, and Joag-Dev*et al.* (1983) also have used a related method to characterize association of normal random variables. But we will show here that many more stochastic ordering results can be derived from that identity. Moreover, we will show that many of the sufficient conditions are also necessary.

2. A useful identity

First we will fix our notation. Let $\boldsymbol{\mu} \in \mathbb{R}^n$, and let $\boldsymbol{\Sigma}$ be a positive semidefinite $n \times n$-matrix. Then a n-dimensional random vector $\boldsymbol{X}$ has a normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ (written as $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$), if it has the characteristic function

$$\Psi_{\boldsymbol{X}}(\boldsymbol{t}) := Ee^{i\boldsymbol{t}^T\boldsymbol{X}} = \exp\left(i\boldsymbol{t}^T\boldsymbol{\mu} - \frac{1}{2}\boldsymbol{t}^T\boldsymbol{\Sigma}\boldsymbol{t}\right), \quad \boldsymbol{t} \in \mathbb{R}^n.$$

If $\boldsymbol{\Sigma}$ is positive definite, and hence invertible, then $\boldsymbol{X}$ has the density

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{\sqrt{(2\pi)^n \det(\boldsymbol{\Sigma})}} \cdot \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right), \quad \boldsymbol{x} \in \mathbb{R}^n.$$

The following identity is a consequence of the well known Fourier Inversion Theorem, see e.g. Lemma 9.5.4 in Dudley (1989).

LEMMA 1. *If $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\boldsymbol{\Sigma}$ is positive definite, then we have the following relation between the characteristic function $\Psi_{\boldsymbol{X}}$ and the density function $f_{\boldsymbol{X}}$:*

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{(2\pi)^n}\int e^{-i\boldsymbol{t}^T\boldsymbol{x}}\Psi_{\boldsymbol{X}}(\boldsymbol{t})d\boldsymbol{t} = \frac{1}{(2\pi)^n}\int \exp\left(-i\boldsymbol{t}^T(\boldsymbol{x}-\boldsymbol{\mu}) - \frac{1}{2}\boldsymbol{t}^T\boldsymbol{\Sigma}\boldsymbol{t}\right)d\boldsymbol{t}.$$

Now we consider the difference $Ef(\boldsymbol{Y}) - Ef(\boldsymbol{X})$ for normally distributed random vectors $\boldsymbol{X}, \boldsymbol{Y}$. We will derive an important identity for this expression under some weak regularity conditions on f. If $f : \mathbb{R}^n \to \mathbb{R}$ is twice continuously differentiable, then we write as usual

$$\nabla f(\boldsymbol{x}) := \left(\frac{\partial}{\partial x_i}f(\boldsymbol{x})\right)_{i=1}^n \quad \text{and} \quad \boldsymbol{H}_f(\boldsymbol{x}) := \left(\frac{\partial^2}{\partial x_i \partial x_j}f(\boldsymbol{x})\right)_{i,j=1}^n$$

for the gradient and the Hesse matrix of f. Recall also that for a matrix $\boldsymbol{A} = (a_{ij})$ the trace is defined as $\text{tr}(\boldsymbol{A}) := \sum_{i=1}^n a_{ii}$, and hence $\text{tr}(\boldsymbol{AB}) = \sum_{i,j=1}^n a_{ij}b_{ij}$, if $\boldsymbol{A}$ and $\boldsymbol{B}$ are symmetric matrices. With these notations we can state the following result.

THEOREM 2. *Let $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{Y} \sim \mathcal{N}(\boldsymbol{\mu}', \boldsymbol{\Sigma}')$, where $\boldsymbol{\Sigma}$ and $\boldsymbol{\Sigma}'$ are positive definite, and let ϕ_λ be the density of $\mathcal{N}(\lambda\boldsymbol{\mu}' + (1-\lambda)\boldsymbol{\mu}, \lambda\boldsymbol{\Sigma}' + (1-\lambda)\boldsymbol{\Sigma})$, $0 \le \lambda \le 1$. Moreover, assume that $f : \mathbb{R}^n \to \mathbb{R}$ is twice continuously differentiable, and satisfies the following condition:*

$$\lim_{|x_j|\to\infty} f(\boldsymbol{x}) \cdot \phi_\lambda(\boldsymbol{x}) = 0 \quad \forall \boldsymbol{x} \in \mathbb{R}^n, \quad 0 \le \lambda \le 1, \ 1 \le j \le n, \tag{2.1}$$

$$\lim_{|x_j|\to\infty} f(\boldsymbol{x}) \cdot \frac{\partial}{\partial x_i}\phi_\lambda(\boldsymbol{x}) = 0 \quad \forall \boldsymbol{x} \in \mathbb{R}^n, \quad 0 \le \lambda \le 1, \ 1 \le i, \ j \le n, \tag{2.2}$$

and

$$\lim_{|x_j|\to\infty} \phi_\lambda(\boldsymbol{x}) \cdot \frac{\partial}{\partial x_i}f(\boldsymbol{x}) = 0 \quad \forall \boldsymbol{x} \in \mathbb{R}^n, \quad 0 \le \lambda \le 1, \ 1 \le i, \ j \le n. \tag{2.3}$$

Then

$$Ef(\boldsymbol{Y}) - Ef(\boldsymbol{X}) = \int_0^1\!\!\int \left((\boldsymbol{\mu}'-\boldsymbol{\mu})^T\nabla f(\boldsymbol{x}) + \frac{1}{2}\,\text{tr}((\boldsymbol{\Sigma}'-\boldsymbol{\Sigma})\boldsymbol{H}_f(\boldsymbol{x}))\right) \cdot \phi_\lambda(\boldsymbol{x})\, d\boldsymbol{x}\, d\lambda. \tag{2.4}$$

PROOF. Define $g(\lambda) := \int f(\boldsymbol{x})\phi_\lambda(\boldsymbol{x})d\boldsymbol{x}$. Then $Ef(\boldsymbol{Y}) - Ef(\boldsymbol{X}) = g(1) - g(0)$. Hence it is sufficient to show that g' is equal to the expression inside of the outer integral in (2.4). But this can be seen as follows. From Lemma 1 we can deduce that

$$\begin{aligned}
\frac{\partial}{\partial\lambda}\phi_\lambda(\boldsymbol{x}) &= \frac{\partial}{\partial\lambda}\frac{1}{(2\pi)^n}\int e^{-i\boldsymbol{t}^T\boldsymbol{x}}\Psi_\lambda(\boldsymbol{t})d\boldsymbol{t} \\
&= \frac{1}{(2\pi)^n}\int \frac{\partial}{\partial\lambda}e^{-i\boldsymbol{t}^T\boldsymbol{x}}\Psi_\lambda(\boldsymbol{t})d\boldsymbol{t} \\
&= \frac{1}{(2\pi)^n}\int e^{-i\boldsymbol{t}^T\boldsymbol{x}}\Psi_\lambda(\boldsymbol{t})\cdot\left(-i\boldsymbol{t}^T(\boldsymbol{\mu}'-\boldsymbol{\mu}) - \frac{1}{2}\boldsymbol{t}^T(\boldsymbol{\Sigma}'-\boldsymbol{\Sigma})\boldsymbol{t}\right)d\boldsymbol{t} \\
&= -\sum_{i=1}^n(\mu_i'-\mu_i)\frac{\partial}{\partial x_i}\phi_\lambda(\boldsymbol{x}) + \frac{1}{2}\sum_{i,j=1}^n(\sigma_{ij}'-\sigma_{ij})\frac{\partial^2}{\partial x_i\partial x_j}\phi_\lambda(\boldsymbol{x}).
\end{aligned}$$

This yields

$$\begin{aligned}
g'(\lambda) &= \int f(\boldsymbol{x})\frac{\partial}{\partial\lambda}\phi_\lambda(\boldsymbol{x})d\boldsymbol{x} \\
&= -\sum_{i=1}^n(\mu_i'-\mu_i)\int f(\boldsymbol{x})\frac{\partial\phi_\lambda(\boldsymbol{x})}{\partial x_i}d\boldsymbol{x} + \frac{1}{2}\sum_{i,j=1}^n(\sigma_{ij}'-\sigma_{ij})\int f(\boldsymbol{x})\frac{\partial^2\phi_\lambda(\boldsymbol{x})}{\partial x_i\partial x_j}d\boldsymbol{x} \\
&= \sum_{i=1}^n(\mu_i'-\mu_i)\int \phi_\lambda(\boldsymbol{x})\frac{\partial f(\boldsymbol{x})}{\partial x_i}d\boldsymbol{x} + \frac{1}{2}\sum_{i,j=1}^n(\sigma_{ij}'-\sigma_{ij})\int \phi_\lambda(\boldsymbol{x})\frac{\partial^2 f(\boldsymbol{x})}{\partial x_i\partial x_j}d\boldsymbol{x} \\
&= \int\left((\boldsymbol{\mu}'-\boldsymbol{\mu})^T\nabla f(\boldsymbol{x}) + \frac{1}{2}\operatorname{tr}((\boldsymbol{\Sigma}'-\boldsymbol{\Sigma})\boldsymbol{H}_f(\boldsymbol{x}))\right)\cdot\phi_\lambda(\boldsymbol{x})\ d\boldsymbol{x}.
\end{aligned}$$

Here the third equality follows from a (double) application of partial integration, taking into account the conditions (2.1)–(2.3). □

Remarks 1. Notice that the conditions (2.1)–(2.3) are very weak regularity conditions, which assure the existence of all occurring integrals. They are always fulfilled, if the function f together with its first derivatives fulfills a polynomial growth condition at infinity.

2. In the course of completing this article, we discovered that a similar result can already be found in Houdré *et al.* (1998). Their proof, however, is quite different. It uses advanced tools from the theory of Lévy processes.

From Theorem 2 we can immediately derive the following sufficient condition for non-negativity of $Ef(\boldsymbol{Y}) - Ef(\boldsymbol{X})$.

COROLLARY 3. *Let $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{Y} \sim \mathcal{N}(\boldsymbol{\mu}', \boldsymbol{\Sigma}')$, and assume that $f : \mathbb{R}^n \to \mathbb{R}$ satisfies the conditions of Theorem 2. Then $Ef(\boldsymbol{Y}) - Ef(\boldsymbol{X}) \geq 0$, if the following two conditions hold:*

$$\sum_{i=1}^n(\mu_i'-\mu_i)\frac{\partial f(\boldsymbol{x})}{\partial x_i} \geq 0 \qquad \text{for all} \quad \boldsymbol{x}\in\mathbb{R}^n, \tag{2.5}$$

and

$$\sum_{i,j=1}^{n} (\sigma'_{ij} - \sigma_{ij}) \frac{\partial^2 f(\boldsymbol{x})}{\partial x_i \partial x_j} \geq 0 \qquad \text{for all} \quad \boldsymbol{x} \in \mathbb{R}^n. \tag{2.6}$$

PROOF. If $\boldsymbol{\Sigma}$ and $\boldsymbol{\Sigma}'$ are positive definite, then the assertion follows immediately from Theorem 2. If one or both of them are only semi-definite, then we can use an approximation argument as follows. Let $\boldsymbol{I}$ be the identity matrix. Then for arbitrary $\varepsilon > 0$ the matrices $\boldsymbol{\Sigma}+\varepsilon\boldsymbol{I}$ and $\boldsymbol{\Sigma}'+\varepsilon\boldsymbol{I}$ are positive definite. Hence we can apply Theorem 2 to this perturbed matrices. Now let ε approach zero to get the assertion. □

3. Stochastic orderings

From Corollary 3 we can derive many sufficient conditions for stochastic ordering of multivariate normal distributions. For a comprehensive treatment of stochastic order relations we refer to Shaked and Shanthikumar (1994). Many of them can be defined as follows: Let $\mathbf{F}$ be some class of measurable functions $f : \mathbb{R}^n \to \mathbb{R}$. Then we say that for two random vectors with values in $\mathbb{R}^n$ the relation

$$\boldsymbol{X} \leq_{\mathbf{F}} \boldsymbol{Y} \qquad \text{holds, if} \quad Ef(\boldsymbol{X}) \leq Ef(\boldsymbol{Y}) \quad \text{for all} \quad f \in \mathbf{F}, \tag{3.1}$$

holds whenever the expectation is well defined. A unified treatment of this type of orderings can be found in Müller (1997). The most important examples are the following ones.

- Usual stochastic order: $\boldsymbol{X} \leq_{st} \boldsymbol{Y}$, if $Ef(\boldsymbol{X}) \leq Ef(\boldsymbol{Y})$ for all increasing functions $f : \mathbb{R}^n \to \mathbb{R}$.
- Convex order: $\boldsymbol{X} \leq_{cx} \boldsymbol{Y}$, if $Ef(\boldsymbol{X}) \leq Ef(\boldsymbol{Y})$ for all convex functions $f : \mathbb{R}^n \to \mathbb{R}$.
- Increasing convex order: $\boldsymbol{X} \leq_{icx} \boldsymbol{Y}$, if $Ef(\boldsymbol{X}) \leq Ef(\boldsymbol{Y})$ for all increasing convex functions $f : \mathbb{R}^n \to \mathbb{R}$.

For these classical orderings, which we will consider first, sufficient conditions for ordering normal random vectors are well known, but we give here new purely analytical proofs, based on Corollary 3. Previously the proofs have always been based on almost sure representations. We admit that these proofs are more elementary for the classical orderings. The usefulness of our approach will become much more obvious later, when we consider orderings, for which an almost sure representation is not available. Moreover, we show that these conditions are also necessary. This is mostly easy to see, but we could not find these results in the literature. Therefore we state them here with proofs. For completeness we will first state the results in the one-dimensional case, since we will need them later on in the proofs of the necessity parts.

THEOREM 4. *Let $X \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and $Y \sim \mathcal{N}(\mu_2, \sigma_2^2)$ be univariate normally distributed. Then*
a) $X \leq_{st} Y$, *if and only if* $\mu_1 \leq \mu_2$ *and* $\sigma_1^2 = \sigma_2^2$;
b) $X \leq_{icx} Y$, *if and only if* $\mu_1 \leq \mu_2$ *and* $\sigma_1^2 \leq \sigma_2^2$;
c) $X \leq_{cx} Y$, *if and only if* $\mu_1 = \mu_2$ *and* $\sigma_1^2 \leq \sigma_2^2$.

PROOF. a) The sufficiency of $\mu_1 \leq \mu_2$ and $\sigma_1^2 = \sigma_2^2$ follows from the fact that in this case $Y \sim X + \mu_2 - \mu_1$. On the other hand, $X \leq_{st} Y$ can only hold, if the ratio f_Y / f_X of their densities fulfills

$$\lim_{t \to -\infty} \frac{f_Y(t)}{f_X(t)} \leq 1 \qquad \text{and} \qquad \lim_{t \to +\infty} \frac{f_Y(t)}{f_X(t)} \geq 1.$$

But this is only possible if $\mu_1 \leq \mu_2$ and $\sigma_1^2 = \sigma_2^2$.

b) For the if-part observe that Y has the same distribution as $X + Z$, where $Z \sim \mathcal{N}(\mu_2 - \mu_1, \sigma_2^2 - \sigma_1^2)$ is independent of X. Since $EZ \geq 0$, Jensen's inequality implies for any increasing convex function f that

$$Ef(Y) = Ef(X + Z) = E(E[f(X + Z) \mid X]) \geq Ef(X + EZ) \geq Ef(X).$$

To show the converse, let us assume $X \leq_{icx} Y$. Then $EX \leq EY$, hence $\mu_1 \leq \mu_2$. Let us assume that $\sigma_1^2 > \sigma_2^2$. Then $\lim_{t\to+\infty} f_Y(t)/f_X(t) = 0$, and therefore

$$E(X-t)^+ = \int_t^\infty (1 - F_X(x))dx > \int_t^\infty (1 - F_Y(x))dx = E(Y-t)^+$$

for sufficiently large t, a contradiction to $X \leq_{icx} Y$.

c) The proof of the if-part is similar to b), and the only-if-part follows from b), taking into account that $X \leq_{cx} Y$ implies $X \leq_{icx} Y$ as well as $-X \leq_{icx} -Y$. □

Now we will start our investigation of the multivariate case. To begin with, we consider the usual stochastic order.

THEOREM 5. *Let $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\boldsymbol{X}' \sim \mathcal{N}(\boldsymbol{\mu}', \boldsymbol{\Sigma}')$ be n-dimensional normally distributed random vectors. Then $\boldsymbol{X} \leq_{st} \boldsymbol{X}'$ if and only if $\mu_i \leq \mu_i'$ for all $1 \leq i \leq n$ and $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}'$.*

PROOF. It is well known that it is sufficient to consider twice differentiable increasing functions. But then the if-part follows immediately from Corollary 3, since a differentiable function is increasing, if and only $\nabla f(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x} \in \mathbb{R}^n$, and hence (2.5) holds. Equation (2.6) holds trivially, since we have $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}'$. To show the converse, let us assume that $\boldsymbol{X} \leq_{st} \boldsymbol{X}'$ holds. This implies that $X_i \leq_{st} X_i'$ holds for all marginals. Hence we can deduce from Theorem 4 a) that we must have $\mu_i \leq \mu_i'$ and $\sigma_{ii} = \sigma_{ii}'$. Moreover, we have $X_i + X_j \leq_{st} X_i' + X_j'$ for all $1 \leq i < j \leq n$, and since $X_i + X_j \sim \mathcal{N}(\mu_i + \mu_j, \sigma_{ii}^2 + \sigma_{jj}^2 + 2\sigma_{ij})$, we must also have $\sigma_{ii}^2 + \sigma_{jj}^2 + 2\sigma_{ij} = {\sigma_{ii}'}^2 + {\sigma_{jj}'}^2 + 2\sigma_{ij}'$ and thus it is necessary that $\sigma_{ij} = \sigma_{ij}'$. □

THEOREM 6. *Let $\boldsymbol{X}, \boldsymbol{X}'$ be n-dimensional random vectors with normal distributions $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\mathcal{N}(\boldsymbol{\mu}', \boldsymbol{\Sigma}')$ respectively. Then the following conditions are equivalent:*

1. $\boldsymbol{X} \leq_{cx} \boldsymbol{X}'$.
2. $\boldsymbol{\mu} = \boldsymbol{\mu}'$ *and* $\boldsymbol{\Sigma}' - \boldsymbol{\Sigma}$ *is positive semi-definite.*

PROOF. a) We again apply Corollary 3. Here Equation (2.5) holds trivially, and (2.6) can be shown as follows: Since $\boldsymbol{\Sigma}' - \boldsymbol{\Sigma}$ is positive semi-definite, it has the canonical representation $\boldsymbol{\Sigma}' - \boldsymbol{\Sigma} = \sum_{k=1}^n \lambda_k \boldsymbol{a}^{(k)} {\boldsymbol{a}^{(k)}}^T$, where $\boldsymbol{a}^{(k)}$ are the eigenvectors, and $\lambda_k \geq 0$ are the corresponding eigenvalues. Hence

$$\sum_{i,j=1}^n (\sigma_{ij}' - \sigma_{ij}) \frac{\partial^2 f(\boldsymbol{x})}{\partial x_i \partial x_j} = \sum_{k=1}^n \lambda_k \sum_{i,j=1}^n a_i^{(k)} a_j^{(k)} \frac{\partial^2 f(\boldsymbol{x})}{\partial x_i \partial x_j} = \sum_{k=1}^n \lambda_k \cdot {\boldsymbol{a}^{(k)}}^T \boldsymbol{H}_f(\boldsymbol{x}) \boldsymbol{a}^{(k)} \geq 0,$$

since a twice differentiable function f is convex, if and only if its Hesse matrix $\boldsymbol{H}_f$ is positive semi-definite.

b) Now assume that $\boldsymbol{X} \leq_{cx} \boldsymbol{X}'$ holds. Since all the functions $f(\boldsymbol{x}) = x_i$ and $f(\boldsymbol{x}) = -x_i$, $1 \leq i \leq n$ are convex, it is clear that the condition $\boldsymbol{\mu} = \boldsymbol{\mu}'$ is necessary. Therefore let us now assume that $\boldsymbol{\Sigma}' - \boldsymbol{\Sigma}$ is not positive semi-definite, i.e. there is some $\boldsymbol{a} \in \mathbb{R}^n$ such that $\boldsymbol{a}^T(\boldsymbol{\Sigma}' - \boldsymbol{\Sigma})\boldsymbol{a} < 0$. Then

$$E(\boldsymbol{a}^T(\boldsymbol{X} - \boldsymbol{\mu}))^2 = \boldsymbol{a}^T E((\boldsymbol{X} - \boldsymbol{\mu})(\boldsymbol{X} - \boldsymbol{\mu})^T)\boldsymbol{a} = \boldsymbol{a}^T\boldsymbol{\Sigma}\boldsymbol{a} > \boldsymbol{a}^T\boldsymbol{\Sigma}'\boldsymbol{a} = E(\boldsymbol{a}^T(\boldsymbol{X}' - \boldsymbol{\mu}))^2.$$

Hence $\boldsymbol{X} \leq_{cx} \boldsymbol{X}'$ does not hold, since the function $f(\boldsymbol{x}) = (\boldsymbol{a}^T(\boldsymbol{x} - \boldsymbol{\mu}))^2$ is obviously convex. □

Remark. In this case the only-if part can also be found in Scarsini ((1998), Theorem 4), but the proof given there is much more complicated.

For the case of $\leq_{icx}$ an if-and-only-if characterization seems to be unknown. There is still a gap between the necessary and the sufficient conditions in this case. We only can show the following result.

THEOREM 7. *Let $\boldsymbol{X}$, $\boldsymbol{X}'$ be n-dimensional random vectors with normal distributions $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\mathcal{N}(\boldsymbol{\mu}', \boldsymbol{\Sigma}')$ respectively. Then the following conditions hold:*
a) *If $\boldsymbol{\mu} \leq \boldsymbol{\mu}'$ and $\boldsymbol{\Sigma}' - \boldsymbol{\Sigma}$ is positive semi-definite, then $\boldsymbol{X} \leq_{icx} \boldsymbol{X}'$.*
b) *If $\boldsymbol{X} \leq_{icx} \boldsymbol{X}'$, then $\boldsymbol{\mu} \leq \boldsymbol{\mu}'$ and $\boldsymbol{a}^T(\boldsymbol{\Sigma}' - \boldsymbol{\Sigma})\boldsymbol{a} \geq 0$ for all $\boldsymbol{a} \geq \boldsymbol{0}$.*

PROOF. a) In this case (2.5) follows as in Theorem 5 and (2.6) follows as in Theorem 6.

b) If $\boldsymbol{X} \leq_{icx} \boldsymbol{X}'$, then $\boldsymbol{\mu} \leq \boldsymbol{\mu}'$ since the functions $f(\boldsymbol{x}) = x_i$, $i = 1, \ldots, n$, are increasing convex. Now let $\boldsymbol{a} \geq \boldsymbol{0}$. Then the function $f_{\boldsymbol{a}}(\boldsymbol{x}) = f(\boldsymbol{a}^T\boldsymbol{x})$ is increasing convex for all increasing convex functions $f : \mathbb{R} \to \mathbb{R}$. Hence $\boldsymbol{X} \leq_{icx} \boldsymbol{X}'$ implies $\boldsymbol{a}^T\boldsymbol{X} \leq_{icx} \boldsymbol{a}^T\boldsymbol{X}'$. According to Theorem 4 b) this yields $\mathrm{Var}(\boldsymbol{a}^T\boldsymbol{X}) = \boldsymbol{a}^T\boldsymbol{\Sigma}\boldsymbol{a} \leq \boldsymbol{a}^T\boldsymbol{\Sigma}'\boldsymbol{a} = \mathrm{Var}(\boldsymbol{a}^T\boldsymbol{X}')$. □

Now we turn our attention to stochastic order relations, which are suited for the comparison of dependence structures. We first need some notations.

DEFINITION 8. a) For a function $f : \mathbb{R}^n \to \mathbb{R}$ define the difference operators

$$\Delta_i^\varepsilon f(\boldsymbol{x}) := f(\boldsymbol{x} + \varepsilon \boldsymbol{e}_i) - f(\boldsymbol{x}),$$

where $\boldsymbol{e}_i$ is the i-th unit vector and $\varepsilon > 0$.

b) A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be supermodular, if $\Delta_i^\varepsilon \Delta_j^\delta f(\boldsymbol{x}) \geq 0$ holds for all $\boldsymbol{x} \in \mathbb{R}^n$, $1 \leq i < j \leq n$ and all $\varepsilon, \delta \geq 0$.

c) A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be directionally convex, if $\Delta_i^\varepsilon \Delta_j^\delta f(\boldsymbol{x}) \geq 0$ holds for all $\boldsymbol{x} \in \mathbb{R}^n$, $1 \leq i, j \leq n$ and all $\varepsilon, \delta > 0$.

d) A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be $\boldsymbol{\Delta}$-monotone, if for any subset $J = \{i_1, \ldots, i_k\} \subset \{1, \ldots, n\}$ and every $\varepsilon_1, \ldots, \varepsilon_k > 0$

$$\Delta_{i_1}^{\varepsilon_1} \ldots \Delta_{i_k}^{\varepsilon_k} f(\boldsymbol{x}) \geq 0 \qquad \text{for all} \qquad \boldsymbol{x} \in \mathbb{R}^n.$$

The stochastic order relation generated by $\boldsymbol{\Delta}$-monotone functions is called *upper orthant order* (written $\leq_{uo}$), since it can be defined alternatively by comparing upper orthants, i.e. $\boldsymbol{X} \leq_{uo} \boldsymbol{Y}$ holds, if and only if $P(\boldsymbol{X} > \boldsymbol{t}) \leq P(\boldsymbol{Y} > \boldsymbol{t})$ for all $\boldsymbol{t} \in \mathbb{R}^n$. This has been shown by Rüschendorf (1980). But the comparison of upper orthants of normally distributed vectors is well known as Slepian's inequality, which can be found e.g. in Tong (1980), p. 8ff.. It states the following result.

THEOREM 9. *Let* $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$ *and* $\boldsymbol{X}' \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma')$ *be n-dimensional normally distributed random vectors with* $\sigma_{ii} = \sigma'_{ii}$, $1 \leq i \leq n$. *If* $\sigma_{ij} \leq \sigma'_{ij}$, $1 \leq i < j \leq n$, *then* $\boldsymbol{X} \leq_{uo} \boldsymbol{Y}$.

Combining this theorem with Theorem 5 we get the following result.

THEOREM 10. *Let* $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$ *and* $\boldsymbol{X}' \sim \mathcal{N}(\boldsymbol{\mu}', \Sigma')$ *be n-dimensional normally distributed random vectors.*

a) *If* $\mu_i \leq \mu'_i$ *for all* $1 \leq i \leq n$, $\sigma_{ii} = \sigma'_{ii}$, $1 \leq i \leq n$ *and* $\sigma_{ij} \leq \sigma'_{ij}$, $1 \leq i < j \leq n$, *then* $\boldsymbol{X} \leq_{uo} \boldsymbol{Y}$.

b) *If* $\boldsymbol{X} \leq_{uo} \boldsymbol{Y}$, *then* $\mu_i \leq \mu'_i$ *for all* $1 \leq i \leq n$ *and* $\sigma_{ii} = \sigma'_{ii}$.

PROOF. Part a) is an immediate consequence of Theorems 5 and 9. Part b) follows from the fact that $\boldsymbol{X} \leq_{uo} \boldsymbol{Y}$ (resp. $\boldsymbol{X} \geq_{lo} \boldsymbol{Y}$) implies $X_i \leq_{st} Y_i$ for all $1 \leq i \leq n$. □

Unfortunately we are not able to give an if-and-only-if characterization of the upper orthant order for multinormal distributions. It is clear, however, that if $\boldsymbol{\mu} = \boldsymbol{\mu}'$, then $\boldsymbol{X} \leq_{uo} \boldsymbol{Y}$ implies $\sigma_{ij} \leq \sigma'_{ij}$ for all i, j, but we do not know, if this still holds in the case $\boldsymbol{\mu} \leq \boldsymbol{\mu}'$.

Supermodular functions are also called quasimonotone or L-superadditive. The stochastic order relation generated by these functions is called *supermodular order* (written $\leq_{sm}$). Since it is obvious from the definition that every Δ-monotone function is supermodular, it is clear that supermodular order is stronger than the upper orthant order. For some properties and applications of supermodular order we refer to Bäuerle (1997), Bäuerle and Müller (1998), and Shaked and Shanthikumar (1997). In Bäuerle (1997) a sufficient condition for normal distributions has been derived, which has recently been extended to a necessary and sufficient condition by Müller and Scarsini ((2000), Theorem 4.1). They have shown that supermodular ordering holds, if the covariances can be compared. This generalizes Slepian's inequality from indicator functions of rectangles to a much larger class of functions. For a related result we refer to Block and Sampson (1988). Moreover, Huffer (1986) implicitly contains a similar result. He describes an alternative way to proof such results by using the central limit theorem. We will state the result here with proof, since the only complete proof can be found in Müller and Scarsini (2000), as far as we know. Using Corollary 3, however, the proof given there can be simplified considerably.

THEOREM 11. *Let* $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$ *and* $\boldsymbol{Y} \sim \mathcal{N}(\boldsymbol{\mu}', \Sigma')$. *Then the following conditions are equivalent.*

(i) $\boldsymbol{X} \leq_{sm} \boldsymbol{Y}$;

(ii) $\boldsymbol{X}$ *and* $\boldsymbol{Y}$ *have the same marginals and* $\sigma_{ij} \leq \sigma'_{ij}$ *for all* i, j.

PROOF. a) The implication (i) ⇒ (ii) follows immediately from Slepian's inequality and the well known fact, that supermodular order can only hold, if the random vectors have the same marginals.

b) It follows from the results in Müller and Scarsini (2000) that it is sufficient to consider twice differentiable supermodular functions. However, a twice differentiable function is supermodular, if and only if

$$\frac{\partial^2 f(\boldsymbol{x})}{\partial x_i \partial x_j} \geq 0 \qquad \text{for all} \quad \boldsymbol{x} \in \mathbb{R}^n, \;\; 1 \leq i < j \leq n.$$

Hence the implication (ii) ⇒ (i) follows from Corollary 3. □

The *directionally convex order* (denoted as $\leq_{dcx}$), which is generated by the directionally convex functions, also found increasing interest recently, see e.g. Shaked and Shanthikumar (1990), Meester and Shanthikumar (1993), Bäuerle and Rolski (1998), and Müller and Scarsini (1999), but sufficient and necessary conditions for $\leq_{dcx}$ in the case of normal distributions seem to be new. We can show the following result.

THEOREM 12. *Let* $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$ *and* $\boldsymbol{Y} \sim \mathcal{N}(\boldsymbol{\mu}', \Sigma')$. *Then we have* $\boldsymbol{X} \leq_{dcx} \boldsymbol{Y}$, *if and only if* $\boldsymbol{\mu} = \boldsymbol{\mu}'$ *and* $\sigma_{ij} \leq \sigma'_{ij}$ *for all* $1 \leq i, j \leq n$.

PROOF. A twice differentiable function f is directionally convex, if and only if

$$\frac{\partial^2 f(\boldsymbol{x})}{\partial x_i \partial x_j} \geq 0 \qquad \text{for all} \quad \boldsymbol{x} \in \mathbb{R}^n, \;\; 1 \leq i, \; j \leq n.$$

Hence the sufficiency of the mentioned condition follows immediately from Corollary 3. But this condition is also necessary, since all of the following functions are directionally convex: $f(\boldsymbol{x}) = x_i$, $f(\boldsymbol{x}) = -x_i$ and $f(\boldsymbol{x}) = x_i x_j$, $1 \leq i, j \leq n$. □

Remark. It follows from the above theorem that for normal distributions with the same marginals there is no difference between supermodular order and directional convex order. Both orderings hold, if the off-diagonal elements of the covariance matrix are ordered. The advantage of $\leq_{dcx}$ lies in the fact that we can also compare random vectors with different marginals. In fact, it follows from Theorem 12 that for normal distributions directional convex order can be decomposed in a supermodular ordering part (i.e. a dependence ordering part, where the marginals are fixed) and a convex ordering part (i.e. a variability ordering part). Indeed, assume that $\sigma_{ij} \leq \sigma'_{ij}$ for all $1 \leq i, j \leq n$. Define the matrix $\tilde{\Sigma}$ by

$$\tilde{\sigma}_{ij} := \begin{cases} \sigma_{ij}, & \text{if } i = j \\ \sigma'_{ij}, & \text{if } i \neq j. \end{cases}$$

Then $\Sigma' - \tilde{\Sigma}$ is positive semidefinite, since it is a diagonal matrix with non-negative entries. Moreover, $\tilde{\Sigma} - \Sigma$ has non-negative entries, and the two matrices have the same diagonal elements. Hence we have shown that $\mathcal{N}(\boldsymbol{\mu}, \Sigma) \leq_{dcx} \mathcal{N}(\boldsymbol{\mu}, \Sigma')$ holds, if and only if there is some $\tilde{\Sigma}$ such that

$$\mathcal{N}(\boldsymbol{\mu}, \Sigma) \leq_{sm} \mathcal{N}(\boldsymbol{\mu}, \tilde{\Sigma}) \leq_{cx} \mathcal{N}(\boldsymbol{\mu}, \Sigma').$$

Acknowledgements

I would like to thank M. Shaked and Y. L. Tong for pointing out some useful references to me.

REFERENCES

Bäuerle, N. (1997). Inequalities for stochastic models via supermodular orders, *Comm. Statist. Stochastic Models*, **13**, 181–201.

Bäuerle, N. and Müller, A. (1998). Modeling and comparing dependencies in multivariate risk portfolios, *ASTIN Bulletin*, **28**, 59–76.

Bäuerle, N. and Rolski, T. (1998). A monotonicity result for the work-load in Markov-modulated queues, *J. Appl. Probab.*, **35**, 741–747.

Block, H. W. and Sampson, A. R. (1988). Conditionally ordered distributions, *J. Multivariate Anal.*, **27**, 91–104.

Dudley R. M. (1989). *Real Analysis and Probability*, Wadsworth & Brooks, Pacific Grove.

Houdré, C., Pérez-Abreu, V. and D. Surgailis (1998). Interpolation, correlation identities and inequalities for infinitely divisible variables, *J. Fourier Anal. Appl.*, **4**, 651–668.

Huffer, F. W. (1986). Slepian's inequality via the central limit theorem, *Canad. J. Statist.*, **14**, 367–370.

Joag-Dev, K., Perlman, M. D. and Pitt, L. D. (1983). Association of normal random variables and Slepian's inequality, *Ann. Probab.*, **11**, 451–455.

Meester, L. E. and Shanthikumar, J. G. (1993). Regularity of stochastic processes: A theory of directional convexity, *Probab. Engrg. Inform. Sci.*, **7**, 343–360.

Mosler, K. (1984). Characterization of some stochastic orderings in multinormal and elliptic distributions, *Operations research, Proceedings of the 12th Annual Meeting in Mannheim 1983*, 520–527.

Müller, A. (1997). Stochastic orders generated by integrals: A unified study, *Advances in Applied Probability*, **29**, 414–428.

Müller, A. and Scarsini, M. (1999). Stochastic comparison of random vectors with fixed copulae, WIOR-Report 554, University Karlsruhe.

Müller, A. and Scarsini, M. (2000). Some remarks on the supermodular order, *J. Multivariate Anal.*, **73**, 107–119.

Rüschendorf, L. (1980). Inequalities for the expectation of Δ-monotone functions, *Z. Wahrsch. Verw. Gebiete*, **54**, 341–349.

Scarsini, M. (1998). Multivariate convex orderings, dependence, and stochastic equality, *J. Appl. Probab.*, **35**, 93–103.

Shaked, M. and Shanthikumar, J. G. (1990). Parametric stochastic convexity and concavity of stochastic processes, *Ann. Inst. Statist. Math.*, **42**, 509–531.

Shaked, M. and Shanthikumar, J. G. (1994). *Stochastic Orders and Their Applications*, Academic Press, London.

Shaked, M. and Shanthikumar, J. G. (1997). Supermodular stochastic orders and positive dependence of random vectors, *J. Multivariate Anal.*, **61**, 86– 101.

Stoyan, D. (1983). *Comparison Methods for Queues and Other Stochastic Models*, Wiley, Chichester.

Szekli, R. (1995). *Stochastic Ordering and Dependence in Applied Probability*, Lecture Notes in Statistics, Vol. 97, Springer, New York.

Tong, Y. L. (1980). *Probability Inequalities in Multivariate Distributions*, Academic Press, New York.

Tong, Y. L. (1990). *The Multivariate Normal Distribution*, Springer, New York.