Available online at www.sciencedirect.com
SCIENCE DIRECT®

Journal of Multivariate Analysis 97 (2006) 1251 – 1261

Journal of Multivariate Analysis
www.elsevier.com/locate/jmva

Total positivity order and the normal distribution☆

Yosef Rinott[a,*,1], Marco Scarsini[b,2]

[a]Department of Statistics, Hebrew University of Jerusalem, Mount Scopus, Jerusalem 91905, Israel
[b]Dipartimento di Statistica e Matematica Applicata, Università di Torino, Piazza Arbarello 8, I–10122 Torino, Italy

Received 29 June 2004
Available online 12 September 2005

Abstract

Unlike the usual stochastic order, total positivity order is closed under conditioning. Here we provide a general formulation of the preservation properties of the order under conditioning; we study certain properties of the order including translation properties and the implications of having equality in the inequality defining the order. Specializing to the multivariate normal distribution, the study of total positivity order leads to new cones defined in terms of covariance M-matrices related to positive dependence, whose properties we study.

AMS 2000 subject classification: Primary 60E15; secondary 62E15; 62H99

Keywords: Likelihood ratio order; M-matrices; MTP_2; Affiliation; FKG condition

1. Introduction

Stochastic order relations have been used to compare different features of random quantities, e.g. location, variability, dependence, etc. The reader is referred to [13, 11] for an introduction to stochastic orders.

☆ Work performed while the authors were fellows of ICER.

* Corresponding author. Fax: +972 2 5883549.
E-mail addresses: rinott@mscc.huji.ac.il (Y. Rinott), marco.scarsini@unito.it (M. Scarsini).

1 Partially supported by a grant of the Israel Science Foundation.

2 Partially supported by MIUR-COFIN.

doi:10.1016/j.jmva.2005.07.008

In this paper we discuss the *multivariate total positivity order*, also known as multivariate likelihood ratio order. This order is stronger than the usual stochastic order. Whereas stochastic order is not preserved under conditioning, the stronger total positivity order does have such a closure property which is often needed in applications. It is this stronger property which makes it interesting and useful (see [15]). This is analogous and connected to the fact that association of random variables is not closed under conditioning, but the stronger condition of MTP_2 (also known as FKG condition, or affiliation) does possess such a closure property. For the FKG condition the reader is referred to [1,3,12,9]. Karlin and Rinott [5] provided a review of MTP_2 distributions. The term affiliation was coined by Milgrom and Weber [10], who studied its properties in the context of auction theory. A dependence ordering involving a TP_2 condition has been introduced by Kimeldorf and Sampson [8] for distributions in the same Fréchet class (see also [4,2]).

The likelihood ratio order for univariate densities f, g is defined by requiring that the ratio f/g be increasing. It is well known that this order is stronger than the usual stochastic order. For multivariate densities, by itself, the condition f/g increasing (in each variable) does not lead to useful results. The condition that f/g is increasing is necessary but not sufficient for the TP order; together with the condition that f is MTP_2 it is also sufficient. We will assume the existence of densities with respect to a product measure on $\mathbb{R}^d$ or a more general lattice, which includes discrete distributions. It is possible to define conditions like the MTP_2 in terms of the associated probability rather than the density. The results are essentially the same.

We will focus our attention on some properties of total positivity order in general, and with more detail when applied to multinormal distributions. In Section 2 we provide a general formulation of the preservation-under-conditioning properties of the order, and applications. The order is defined by an inequality and we study the case of equality. We also characterize the order for translations and convolutions. In Sections 3 we focus on multinormal distributions. We will need some results about covariance M-matrices and a related order, that have some interest of their own, and will be studied in Section 4. We then show that adding a positive deterministic vector to a multinormal random vector produces an increase in the total positivity order if and only if the vector lies in a cone defined by the covariance matrix. This result shows the interaction between the location comparison and the dependence that is captured by the likelihood ratio order.

2. General results

In this section we define the total positivity order and we study some of its properties.

Using the framework of Karlin and Rinott [5] we consider a product lattice $\mathcal{X} := \times_{i=1}^{d} \mathcal{X}_i$ and a product measure σ on this lattice. Densities are taken with respect to $\mathrm{d}\sigma$. For the sake of simplicity we will write $\mathrm{d}\boldsymbol{x}$ instead of $\mathrm{d}\sigma(\boldsymbol{x})$.

Let $J \subseteq \{1, \ldots, d\}$, and $\mathcal{X}_J := \times_{i \in J} \mathcal{X}_i$; for $\boldsymbol{x} \in \mathcal{X}$ let $\boldsymbol{x}_J$ denote a vector in $\mathcal{X}_J$ constructed by using only the coordinates in J of $\boldsymbol{x}$. For $\boldsymbol{x} = (x_1, \ldots, x_d) \in \mathcal{X}$ and $J = \{1, \ldots, d\} \setminus i$ set $\boldsymbol{x}_{-i} := \boldsymbol{x}_J = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_d) \in \mathcal{X}_{-i} := \mathcal{X}_J$. For $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$ we set $\boldsymbol{x} \vee \boldsymbol{y} = (\max\{x_1, y_1\}, \ldots, \max\{x_d, y_d\})$, $\boldsymbol{x} \wedge \boldsymbol{y} = (\min\{x_1, y_1\}, \ldots, \min\{x_d, y_d\})$.

Definition 2.1. Let f and g be densities defined on $\mathcal{X}$.
(a) We say $f \preceq_{\text{TP}} g$ if

$$f(\boldsymbol{x})g(\boldsymbol{y}) \leqslant f(\boldsymbol{x} \wedge \boldsymbol{y})g(\boldsymbol{x} \vee \boldsymbol{y}) \quad \text{for all } \boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}. \tag{2.1}$$

(b) We say that a density f is MTP_2 if $f \preceq_{\text{TP}} f$.

Given an $\mathcal{X}$-valued random quantity $\boldsymbol{X}$, we denote its density by $f_{\boldsymbol{X}}$ and write $\boldsymbol{X} \preceq_{\text{TP}} \boldsymbol{Y}$ whenever $f_{\boldsymbol{X}} \preceq_{\text{TP}} f_{\boldsymbol{Y}}$. Shaked and Shantikumar [13] denote the order $\preceq_{\text{TP}}$ by $\leqslant_{\text{lr}}$.
We mention briefly the following facts that can be found in [11]:

Lemma 2.2.
(a) *The relation $f \preceq_{\text{TP}} g$ implies that $h(\boldsymbol{x}) := g(\boldsymbol{x})/f(\boldsymbol{x})$ is increasing in $\boldsymbol{x}$. The converse is not true in general.*
(b) *If either f or g is MTP_2 then $f \preceq_{\text{TP}} g$ if and only if $h(\boldsymbol{x}) := g(\boldsymbol{x})/f(\boldsymbol{x})$ is increasing.*
(c) *For MTP_2 densities the relation $\preceq_{\text{TP}}$ is a partial order, that is, it is reflexive, antisymmetric and transitive.*

Definition 2.3.
(a) Given two random quantities $\boldsymbol{X}, \boldsymbol{Y}$, we say that $\boldsymbol{X} \leqslant_{\text{st}} \boldsymbol{Y}$ if $\mathbb{E}[\phi(\boldsymbol{X})] \leqslant \mathbb{E}[\phi(\boldsymbol{Y})]$ for all nondecreasing functions ϕ.
(b) A random vector $\boldsymbol{X}$ is associated if $\text{Cov}[\phi(\boldsymbol{X}), \psi(\boldsymbol{X})] \geqslant 0$ for all nondecreasing functions ϕ, ψ.

An important consequence of the total positivity order is the following result by Holley [3] which says that it implies stochastic order, and the FKG inequality which says that the FKG or MTP_2 condition implies association. For a proof of these results see, e.g., [11]. It is well known that part (a) implies part (b) easily.

Proposition 2.4. *If $\boldsymbol{X} \preceq_{\text{TP}} \boldsymbol{Y}$ then $\boldsymbol{X} \leqslant_{\text{st}} \boldsymbol{Y}$.* (b) *If $f_{\boldsymbol{X}}$ is MTP_2 then $\boldsymbol{X}$ is associated.*

Stochastic order is not preserved under conditioning. As we shall see, total positivity order is preserved under certain conditioning, and thus, if the total positivity order holds, we are guaranteed stochastic order also after conditioning. The following proposition generalizes and unifies various previous results on conditioning with a very simple proof. In order to state it we need some notation. Given $\boldsymbol{A}, \boldsymbol{B} \subseteq \mathcal{X}$ we denote $\boldsymbol{A} \vee \boldsymbol{B} = \{\boldsymbol{a} \vee \boldsymbol{b} : \boldsymbol{a} \in \boldsymbol{A}, \boldsymbol{b} \in \boldsymbol{B}\}$, and $\boldsymbol{A} \wedge \boldsymbol{B} = \{\boldsymbol{a} \wedge \boldsymbol{b} : \boldsymbol{a} \in \boldsymbol{A}, \boldsymbol{b} \in \boldsymbol{B}\}$. We write $\boldsymbol{A} \leqslant \boldsymbol{B}$ if $\boldsymbol{a} \leqslant \boldsymbol{b}$ for all $\boldsymbol{a} \in \boldsymbol{A}$ and $\boldsymbol{b} \in \boldsymbol{B}$.

Theorem 2.5.
(a) *Let $\boldsymbol{A}, \boldsymbol{B} \subseteq \mathcal{X}$ satisfy $\boldsymbol{A} \vee \boldsymbol{B} \subseteq \boldsymbol{B}$, $\boldsymbol{A} \wedge \boldsymbol{B} \subseteq \boldsymbol{A}$. Then $\boldsymbol{X} \preceq_{\text{TP}} \boldsymbol{Y}$, implies*

$$[\boldsymbol{X} | \boldsymbol{X} \in \boldsymbol{A}] \preceq_{\text{TP}} [\boldsymbol{Y} | \boldsymbol{Y} \in \boldsymbol{B}]. \tag{2.2}$$

(b) *Conversely, if* (2.2) *holds for all $\boldsymbol{A}, \boldsymbol{B}$ as above then $\boldsymbol{X} \preceq_{\text{TP}} \boldsymbol{Y}$.*

Proof.
(a) For all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$, the assumptions imply, with $\mathbb{I}$ denoting indicator function,

$$\mathbb{I}_{\boldsymbol{A}}(\boldsymbol{x})\mathbb{I}_{\boldsymbol{B}}(\boldsymbol{y}) \leqslant \mathbb{I}_{\boldsymbol{A}}(\boldsymbol{x} \wedge \boldsymbol{y})\mathbb{I}_{\boldsymbol{B}}(\boldsymbol{x} \vee \boldsymbol{y}) \quad \text{and} \quad f_{\boldsymbol{X}}(\boldsymbol{x}) f_{\boldsymbol{Y}}(\boldsymbol{y}) \leqslant f_{\boldsymbol{X}}(\boldsymbol{x} \wedge \boldsymbol{y}) f_{\boldsymbol{Y}}(\boldsymbol{x} \vee \boldsymbol{y}).$$

Since $f_{X|A}(x) := f_{X|X\in A}(x) = f_X(x)\mathbb{I}_A(x)/P(X \in A)$, we have

$$f_{X|A}(x) f_{Y|B}(y) \leqslant f_{X|A}(x \wedge y) f_{Y|B}(x \vee y) \quad \text{for all} \quad x, y \in \mathcal{X}.$$

(b) Just take (2.2) with $A = B = \mathcal{X}$. □

Remark 2.6. The following are examples where Theorem 2.5 applies. Some special cases were proved in the literature in various ways, including arguments which require differentiation.

(i) Shaked and Shanthikumar [13, Theorem 4E1, p. 133] consider the special case where $A = B$ are rectangular sets. In fact any sublattice will do.
(ii) Another special case is when A and B are any sets such that $A \leqslant B$. Singletons are of special interest.
(iii) An example where we do not necessarily have $A \leqslant B$ is the rectangles defined by $A = [a, b]$, $B = [c, d]$ with $a \leqslant c$, $b \leqslant d$.
(iv) If A is a decreasing set and B an increasing set (namely, their indicators are, respectively, decreasing and increasing functions), then the conditions hold.
(v) Often there is interest in conditioning on a subset of the variables being in some suitable set. In this case Theorem 2.5 is applied with A and B in $\mathcal{X}_J$ where J is the set of indices of the variables on which we condition. See (2.4) below.

Note that the sets in (ii) and (iv) above need not be sublattices. It is easy to see that for $[X|X \in A]$ to be MTP_2 it is necessary that A be a lattice. Hence (ii) and (iv) can provide natural examples where the TP order holds without existence the MTP_2 condition in one or both of the ordered variables.

In the special case that $X = Y$ and therefore X is MTP_2 we obtain the following result:

Proposition 2.7.
(a) *Let $A, B \subseteq \mathcal{X}$ satisfy $A \vee B \subseteq B$, $A \wedge B \subseteq A$. Then X MTP_2 implies*

$$[X|X \in A] \preceq_{\mathrm{TP}} [X|X \in B].$$

(b) *If for some $i \in \{1, \ldots, d\}$ and for all $a \leqslant b$ we have*

$$X_{-i}|X_i = a \preceq_{\mathrm{TP}} X_{-i}|X_i = b,$$

then X is MTP_2.

Proof. Part (a) *is a special case of Theorem* 2.5. To prove part (b) note that for $s_{-i}, t_{-i} \in \mathcal{X}_{-i}$ the condition implies after simple cancellation

$$f_X(s_{-i}, a) f_X(t_{-i}, b) \leqslant f_X(s_{-i} \wedge t_{-i}, a) f_X(s_{-i} \vee t_{-i}, b).$$

From this we get easily $f_X(x) f_X(y) \leqslant f_X(x \wedge y) f_X(x \vee y)$, for all x, y. □

Remark 2.8. We *may* say $f \approx_{\mathrm{TP}} g$ if

$$f(x)g(y) = f(x \wedge y)g(x \vee y) \quad \text{for all } x, y \in \mathcal{X}. \tag{2.3}$$

However, as we shall see this condition implies $g = f$ so the notation $g \approx_{\mathrm{TP}} f$ is used only temporarily. Furthermore, as we shall see, in this case f is MTP_2, and represents independent variables.

The following lemma is only partly new (see, e.g. [5] and references therein). It is important, and so we prove it, also in order to verify the case of equality, which we need later.

Lemma 2.9. *Let $f \preceq_{\mathrm{TP}} g$, and define $f_{-i}(\boldsymbol{x}_{-i}) = \int f(\boldsymbol{x})\mathrm{d}x_i$. Then $f_{-i} \preceq_{\mathrm{TP}} g_{-i}$. If $f \approx_{\mathrm{TP}} g$, then $f_{-i} \approx_{\mathrm{TP}} g_{-i}$.*

Proof. Write $(\boldsymbol{x}_{-i}, z_i) = (x_1, \dots, x_{i-1}, z_i, x_{i+1}, \dots, x_d)$. Let

$$\begin{aligned} a &:= f(\boldsymbol{x}_{-i}, x_i)g(\boldsymbol{y}_{-i}, y_i), & b &:= f(\boldsymbol{x}_{-i}, y_i)g(\boldsymbol{y}_{-i}, x_i), \\ c &:= f(\boldsymbol{x}_{-i} \wedge \boldsymbol{y}_{-i}, y_i)g(\boldsymbol{x}_{-i} \vee \boldsymbol{y}_{-i}, x_i), & d &:= f(\boldsymbol{x}_{-i} \wedge \boldsymbol{y}_{-i}, x_i)g(\boldsymbol{x}_{-i} \vee \boldsymbol{y}_{-i}, y_i). \end{aligned}$$

Note that

$$f_{-i}(\boldsymbol{x}_{-i})g_{-i}(\boldsymbol{y}_{-i}) = \int\int_{x_i<y_i} [a+b]\,\mathrm{d}x_i\,\mathrm{d}y_i + \frac{1}{2}\int\int_{x_i=y_i} [a+b]\,\mathrm{d}x_i\,\mathrm{d}y_i.$$

With a similar expression holding for $f_{-i}(\boldsymbol{x}_{-i} \wedge \boldsymbol{y}_{-i})g_{-i}(\boldsymbol{x}_{-i} \vee \boldsymbol{y}_{-i})$ it is easy to see that we have an inequality in the right direction between the integrals on the regions $\{x_i = y_i\}$ where $a = b \leqslant c = d$, with equality in the case $f \approx_{\mathrm{TP}} g$, and it remains to show that $f \preceq_{\mathrm{TP}} g$ implies

$$\int\int_{x_i<y_i} [a+b]\,\mathrm{d}x_i\,\mathrm{d}y_i \leqslant \int\int_{x_i<y_i} [c+d]\,\mathrm{d}x_i\,\mathrm{d}y_i$$

and $f \approx_{\mathrm{TP}} g$ implies

$$\int\int_{x_i<y_i} [a+b]\,\mathrm{d}x_i\,\mathrm{d}y_i = \int\int_{x_i<y_i} [c+d]\,\mathrm{d}x_i\,\mathrm{d}y_i.$$

Now

$$c + d - (a+b) = \frac{1}{d}[(d-a)(d-b) + (cd - ab)].$$

For $x_i < y_i$, the condition $f \preceq_{\mathrm{TP}} g$ implies $ab \leqslant cd$ and $a, b \leqslant d$ with equalities when $f \approx_{\mathrm{TP}} g$, and the result follows. □

The next result now follows readily from Lemma 2.9 and Theorem 2.5.

Corollary 2.10. *Let $\boldsymbol{X} \preceq_{\mathrm{TP}} \boldsymbol{Y}$, then for any $I \subset \{1, \dots, d\}$ we have $\boldsymbol{X}_I \preceq_{\mathrm{TP}} \boldsymbol{Y}_I$, and more generally, for any $I, J \subset \{1, \dots, d\}$ and $\boldsymbol{A}_J, \boldsymbol{B}_J \subseteqq \mathcal{X}_J$ satisfying $\boldsymbol{A}_J \vee \boldsymbol{B}_J \subseteq \boldsymbol{B}_J$, $\boldsymbol{A}_J \wedge \boldsymbol{B}_J \subseteq \boldsymbol{A}_J$, we have*

$$[\boldsymbol{X}_I | \boldsymbol{X}_J \in \boldsymbol{A}_J] \preceq_{\mathrm{TP}} [\boldsymbol{Y}_I | \boldsymbol{Y}_J \in \boldsymbol{B}_J]. \tag{2.4}$$

For the case of equality in the TP order condition, i.e. $f \approx_{\mathrm{TP}} g$, we have the following results. In some sense they indicate that a small gap in the inequality defining the order relations suggests near equality in the TP order and near independence in the MTP_2 case. One may expect the opposite if the gaps are large. These results are also useful for the normal distribution studied in the next section. The proofs are given below.

Proposition 2.11. (a) *If* $\boldsymbol{X} \approx_{\mathrm{TP}} \boldsymbol{X}$ *then the components of* $\boldsymbol{X}$ *are independent.*
(b) *If* $\boldsymbol{X} \approx_{\mathrm{TP}} \boldsymbol{Y}$, *then* $\boldsymbol{X}$ *and* $\boldsymbol{Y}$ *have the same distribution and independent components.*

Remark 2.12. Note that $\boldsymbol{X} \approx_{\mathrm{TP}} \boldsymbol{Y}$ is different from the statement that $\boldsymbol{X} \preceq_{\mathrm{TP}} \boldsymbol{Y}$ and $\boldsymbol{Y} \preceq_{\mathrm{TP}} \boldsymbol{X}$. The latter clearly implies that $\boldsymbol{X}$ and $\boldsymbol{Y}$ are identically distributed which shows antisymmetry of the order relation $\boldsymbol{X} \preceq_{\mathrm{TP}} \boldsymbol{Y}$.

Proof of Proposition 2.11. Part (a): The proof is by induction.

For $d = 2$, $f_{\boldsymbol{X}}(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{y}) = f_{\boldsymbol{X}}(\boldsymbol{x} \wedge \boldsymbol{y}) f_{\boldsymbol{X}}(\boldsymbol{x} \vee \boldsymbol{y})$ holds for all $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$ iff for all s_1, s_2, t_1, t_2 we have $f_{\boldsymbol{X}}(s_1, s_2) f_{\boldsymbol{X}}(t_1, t_2) = f_{\boldsymbol{X}}(s_1, t_2) f_{\boldsymbol{X}}(t_1, s_2)$. Integrating with respect to t_1 and t_2 we obtain

$$f_{\boldsymbol{X}}(s_1, s_2) = f_{X_1}(s_1) f_{X_2}(s_2).$$

For $d = 3$, we have that $f_{\boldsymbol{X}} \approx_{\mathrm{TP}} f_{\boldsymbol{X}}$ implies, that $[X_1, X_2|X_3] \approx_{\mathrm{TP}} [X_1, X_2|X_3]$. Hence X_1 and X_2 are conditionally independent, given X_3 by the case $d = 2$ just treated.

Using these facts, and the independence of all pairs of variables again by Lemma 2.9 and the case $d = 2$, we have

$$\begin{aligned} f_{X_1,X_2,X_3}(s_1, s_2, s_3) &= f_{X_1,X_2|X_3}(s_1, s_2|s_3) f_{X_3}(s_3) \\ &= f_{X_1|X_3}(s_1|s_3) f_{X_2|X_3}(s_2|s_3) f_{X_3}(s_3) = f_{X_1}(s_1) f_{X_2}(s_2) f_{X_3}(s_3), \end{aligned}$$

that is, X_1, X_2, X_3 are independent.

Assume for $d = n$ independence of any $(n - 1)$ of the components and conditional independence of any $(n - 1)$ components, given the remaining one. Then

$$\begin{aligned} f_{\boldsymbol{X}}(\boldsymbol{s}) &= f_{X_1,\ldots,X_{n-1}|X_n}(s_1, \ldots, s_{n-1}|s_n) f_{X_n}(s_n) \\ &= f_{X_1|X_n}(s_1|s_n) \cdots f_{X_{n-1}|X_n}(s_{n-1}|s_n) f_{X_n}(s_n) = f_{X_1}(s_1) \cdots f_{X_{n-1}}(s_{n-1}) f_{X_n}(s_n). \end{aligned}$$

Part (b): The relation $\boldsymbol{X} \approx_{\mathrm{TP}} \boldsymbol{Y}$ implies readily that $h(\boldsymbol{x}) := f(\boldsymbol{x})/g(\boldsymbol{x})$ satisfies $h(\boldsymbol{x}) = h(\boldsymbol{y})$ whenever $\boldsymbol{x} \leqslant \boldsymbol{y}$. On a lattice this implies that h is constant since for any $\boldsymbol{u}$, $\boldsymbol{v}$ we have $h(\boldsymbol{u}) = h(\boldsymbol{u} \vee \boldsymbol{v}) = h(\boldsymbol{v})$. It follows that $f = g$ and the rest follows from Proposition 2.11. □

It is natural to study the possibility of having TP order by positive translation, since it is the simplest model which implies stochastic order. We will see that this occurs for *all* positive translations only in a very special case.

In the rest of the section we consider the case that our space $\mathcal{X}$ is $\mathbb{R}^d$. The following result unifies Theorems 1.C.5 and 1.C.22 of [13] and extends them to the multivariate case needed here.

Lemma 2.13. *A random vector* $\boldsymbol{X}$ *in* $\mathbb{R}^d$ *with independent components satisfies* $\boldsymbol{X} \preceq_{\mathrm{TP}} \boldsymbol{X} + \boldsymbol{\mu}$ *for all* $\boldsymbol{\mu} \geqslant \boldsymbol{0}$ *if and only if the marginal densities* f_{X_i} *of the components of* $\boldsymbol{X}$ *are log-concave.*

Proof. Note that by independence $\boldsymbol{X}$ is MTP_2, and in this case by Lemma 2.2 we just need to prove that $\boldsymbol{X} \preceq_{\mathrm{TP}} \boldsymbol{X} + \boldsymbol{\mu}$ is equivalent to monotonicity of the likelihood ratio. Thus we show that $h(\boldsymbol{t}) := f_{\boldsymbol{X}+\boldsymbol{\mu}}(\boldsymbol{t})/f_{\boldsymbol{X}}(\boldsymbol{t})$ is increasing in $\boldsymbol{t}$ for all $\boldsymbol{\mu} \geqslant \boldsymbol{0}$ if and only if the densities of the independent components of $\boldsymbol{X}$ are log-concave. Note that log-concavity of the marginals f_{X_i} is equivalent to having all $f_{X_i}(t-\mu)/f_{X_i}(t)$ increasing in t, for all $\mu \geqslant 0$. Writing $h(\boldsymbol{t}) = f_{\boldsymbol{X}}(\boldsymbol{t}-\boldsymbol{\mu})/f_{\boldsymbol{X}}(\boldsymbol{t})$ where $f_{\boldsymbol{X}}$ is a product, we see that log-concavity implies the monotonicity of h. Conversely, by taking $\boldsymbol{\mu}$ which vanishes in all but the ith coordinate, monotonicity of h implies log-concavity of f_{X_i}. □

Theorem 2.14.

(a) *A random vector $\boldsymbol{X} \in \mathbb{R}^d$ satisfies $\boldsymbol{X} \preceq_{\mathrm{TP}} \boldsymbol{X} + \boldsymbol{\mu}$ for all $\boldsymbol{\mu} \geqslant \boldsymbol{0}$ if and only if $\boldsymbol{X}$ has independent components with log-concave marginals.*

(b) *If $\boldsymbol{X}$ has independent components with log-concave marginals then $\boldsymbol{X} \preceq_{\mathrm{TP}} \boldsymbol{X} + \boldsymbol{Y}$ for any random variable $\boldsymbol{Y} \geqslant \boldsymbol{0}$ independent of $\boldsymbol{X}$.*

Proof.

(a) First we show that under the conditions of the proposition

$$f_{\boldsymbol{X}}(\boldsymbol{a}) f_{\boldsymbol{X}}(\boldsymbol{b}) = f_{\boldsymbol{X}}(\boldsymbol{a} \wedge \boldsymbol{b}) f_{\boldsymbol{X}}(\boldsymbol{a} \vee \boldsymbol{b}) \tag{2.5}$$

for all $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^d$.

Note that $\boldsymbol{X}$ satisfies $\boldsymbol{X} \preceq_{\mathrm{TP}} \boldsymbol{X} + \boldsymbol{\mu}$ if and only if

$$f_{\boldsymbol{X}}(\boldsymbol{s}) f_{\boldsymbol{X}}(\boldsymbol{t} - \boldsymbol{\mu}) \leqslant f_{\boldsymbol{X}}(\boldsymbol{s} \wedge \boldsymbol{t}) f_{\boldsymbol{X}}((\boldsymbol{s} \vee \boldsymbol{t}) - \boldsymbol{\mu}) \tag{2.6}$$

for all $\boldsymbol{s}, \boldsymbol{t} \in \mathbb{R}^d$. The choice $\boldsymbol{\mu} = \boldsymbol{0}$ in (2.6) gives inequality in (2.5) in one direction. To get the reverse direction, given any $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^d$, choose in (2.6) $\boldsymbol{s} = \boldsymbol{a} \vee \boldsymbol{b}$, $\boldsymbol{t} = \boldsymbol{a}$ and $\boldsymbol{\mu}$ which vanishes in the coordinates where $\boldsymbol{b} \geqslant \boldsymbol{a}$, and equals $\boldsymbol{a} - \boldsymbol{b}$ on the other coordinates. It is then easy to see that $\boldsymbol{t} - \boldsymbol{\mu} = \boldsymbol{a} \wedge \boldsymbol{b}$, $\boldsymbol{b} = (\boldsymbol{s} \vee \boldsymbol{t}) - \boldsymbol{\mu}$ and $\boldsymbol{a} = \boldsymbol{s} \wedge \boldsymbol{t}$, and with $\boldsymbol{s} = \boldsymbol{a} \vee \boldsymbol{b}$ inequality (2.6) implies

$$f_{\boldsymbol{X}}(\boldsymbol{a} \wedge \boldsymbol{b}) f_{\boldsymbol{X}}(\boldsymbol{a} \vee \boldsymbol{b}) \leqslant f_{\boldsymbol{X}}(\boldsymbol{b}) f_{\boldsymbol{X}}(\boldsymbol{a}).$$

Independence now follows from Proposition 2.11, and the rest of part (a) follows from Lemma 2.13.

(b) This follows by setting $\boldsymbol{y} = \boldsymbol{\mu}$ in (2.6) and integrating the condition with respect to $F_{\boldsymbol{Y}}(\boldsymbol{dy})$ on both sides. □

In the following result part (a) is well known and very useful in verifying the MTP_2 property (see [7,5]). Part (b), whose proof is similar to that of (a), shows that the same holds in the MTP_2 case for the TP order.

Proposition 2.15.

(a) *Suppose $f(\boldsymbol{x})f(\boldsymbol{y}) \neq$ implies $f(\boldsymbol{u})f(\boldsymbol{v}) \neq 0$ for all $\boldsymbol{x} \wedge \boldsymbol{y} \leqslant \boldsymbol{u}, \boldsymbol{v} \leqslant \boldsymbol{x} \vee \boldsymbol{y}$. Then the condition f TP_2 in every pair of variables, with other variables held fixed, implies f MTP_2.*

(b) *Let either f or g be MTP_2 satisfying the above condition on the support, and suppose also that $f(\boldsymbol{x})g(\boldsymbol{y}) \neq 0$ implies $f(\boldsymbol{u})g(\boldsymbol{v}) \neq 0$ for all $\boldsymbol{x} \wedge \boldsymbol{y} \leqslant \boldsymbol{u}, \boldsymbol{v} \leqslant \boldsymbol{x} \vee \boldsymbol{y}$. The relation*

$f \preceq_{\mathrm{TP}} g$ holds if the required inequality is verified for every pair of variables with the others fixed.

3. Multinormal vectors

In this section we study some properties of the TP order for multinormal distributions. Let $\boldsymbol{X} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$. The mean is arbitrary so we set it to be zero. We also assume that $\boldsymbol{\Sigma}$ is a nonsingular correlation matrix. Henceforth when we write expressions like $\boldsymbol{A} \geqslant \boldsymbol{B}$ for vectors or matrices of the same dimension, we mean entry-wise inequalities.

For properties of the multinormal distribution the reader is referred to [14].

In order to state our results we need to define some cones through M-matrices.

Definition 3.1. A symmetric square matrix is called an M-matrix if its off-diagonal elements are nonpositive.

Using Proposition 2.15 it is easy to see [6] that a multinormal $\boldsymbol{X}$ is MTP_2 if and only if $\boldsymbol{\Sigma}^{-1}$ is an M-matrix. Then $\boldsymbol{\Sigma}$ has nonnegative elements. This can be proved either directly, or resorting to the fact that MTP_2 implies association.

It is easy to see, e.g., [11], that for multinormal vectors, $\boldsymbol{X} \preceq_{\mathrm{TP}} \boldsymbol{Y}$ can hold only if $\boldsymbol{X}$ and $\boldsymbol{Y}$ have the same covariance matrix and therefore one is a translation of the other. Thus below we fully characterize the TP order of multinormal vectors by considering the translation case.

Given a matrix $\boldsymbol{\Sigma}$ whose inverse $\boldsymbol{\Sigma}^{-1}$ is an M-matrix, consider the cone

$$\mathcal{C}_{\boldsymbol{\Sigma}} = \{\boldsymbol{\mu} \in \mathbb{R}^d : \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1} \geqslant \mathbf{0}\} = \{\boldsymbol{\mu} \in \mathbb{R}^d : \text{ there exists } \boldsymbol{a} \geqslant \mathbf{0} \text{ such that } \boldsymbol{\mu} = \boldsymbol{a}^{\mathrm{T}}\boldsymbol{\Sigma}\}.$$

Since for such a $\boldsymbol{\Sigma}$ all entries are nonnegative, it follows that $\boldsymbol{\mu} \geqslant \mathbf{0}$. For $\boldsymbol{\Sigma} = \boldsymbol{I}$ the cone $\mathcal{C}_{\boldsymbol{\Sigma}}$ is the whole positive orthant, and it is the largest possible cone.

Theorem 3.2. *Let $\boldsymbol{X}$ be a multinormal random vector.*
(a) *If for some $\boldsymbol{\mu} \in \mathbb{R}^d$ we have $\boldsymbol{X} \preceq_{\mathrm{TP}} \boldsymbol{X} + \boldsymbol{\mu}$, then $\boldsymbol{X}$ is MTP_2.*
(b) *For $\boldsymbol{X}$ MTP_2, we have $\boldsymbol{X} \preceq_{\mathrm{TP}} \boldsymbol{X} + \boldsymbol{\mu}$ iff $\boldsymbol{\mu} \in \mathcal{C}_{\boldsymbol{\Sigma}}$.*

Lemma 3.3. *Let $\boldsymbol{X} = (X_1, X_2)$ be a bivariate normal random vector. If for some $\boldsymbol{\mu} \in \mathbb{R}^2$ we have $\boldsymbol{X} \preceq_{\mathrm{TP}} \boldsymbol{X} + \boldsymbol{\mu}$, then $\rho[X_1, X_2] \geqslant 0$.*

Proof. Set $\rho = \rho[X_1, X_2]$. Without loss of generality we can assume that $\mathbb{E}[\boldsymbol{X}] = \mathbf{0}$. In general $\boldsymbol{X} \preceq_{\mathrm{TP}} \boldsymbol{X} + \boldsymbol{\mu}$ iff $f_{\boldsymbol{X}}(\boldsymbol{s}) f_{\boldsymbol{X}+\boldsymbol{\mu}}(\boldsymbol{t}) \leqslant f_{\boldsymbol{X}}(\boldsymbol{s} \wedge \boldsymbol{t}) f_{\boldsymbol{X}+\boldsymbol{\mu}}(\boldsymbol{s} \vee \boldsymbol{t})$, which in the normal case is equivalent to

$$\begin{aligned}&(\boldsymbol{s} \vee \boldsymbol{t} - \boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{s} \vee \boldsymbol{t} - \boldsymbol{\mu}) + (\boldsymbol{s} \wedge \boldsymbol{t})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{s} \wedge \boldsymbol{t})\\ &\quad -(\boldsymbol{t} - \boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{t} - \boldsymbol{\mu}) - \boldsymbol{s}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{s} \leqslant 0.\end{aligned} \tag{3.1}$$

Define $M(\boldsymbol{s}, \boldsymbol{t}) := (\boldsymbol{s} \vee \boldsymbol{t})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{s} \vee \boldsymbol{t}) + (\boldsymbol{s} \wedge \boldsymbol{t})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{s} \wedge \boldsymbol{t}) - \boldsymbol{s}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{s} - \boldsymbol{t}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{t}$.

Then the left-hand side of (3.1) becomes

$$M(\boldsymbol{s}, \boldsymbol{t}) - 2\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{s} \vee \boldsymbol{t}) + 2\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{t} = M(\boldsymbol{s}, \boldsymbol{t}) - 2\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{s} - \boldsymbol{t})_{+}.$$

Let now $s_1 < t_1$ and $s_2 > t_2$. Then, up to a positive constant, the latter expression becomes

$$\rho(s_1 s_2 + t_1 t_2 - s_1 t_2 - t_1 s_2) - (\mu_2 - \rho\mu_1)(s_2 - t_2) = (s_2 - t_2)(\rho(s_1 - t_1 + \mu_1) - \mu_2).$$

If ρ is negative, it is always possible to find s_1, t_1 that make the above expression positive. $\square$

Call $\rho_{ij\cdot K}$ the partial correlation coefficient

$$\rho_{ij\cdot K} = \frac{\mathrm{Cov}[X_i, X_j | \boldsymbol{X}_K]}{(\mathrm{Var}[X_i|\boldsymbol{X}_K]\mathrm{Var}[X_j|\boldsymbol{X}_K])^{1/2}}.$$

Lemma 3.4. *For $K \subset \{1, \ldots, d\}$ and $i, j \notin K$ we have*

$$\boldsymbol{X} \preceq_{\mathrm{TP}} \boldsymbol{X} + \boldsymbol{\mu},$$

implies $\rho_{ij\cdot K} \geqslant 0$.

Proof. By Corollary 2.10 we have that $[(X_i, X_j)|\boldsymbol{X}_K] \preceq_{\mathrm{TP}} [(X_i + \mu_i, X_j + \mu_j)|\boldsymbol{X}_K]$. The result follows now from Lemma 3.3. $\square$

Lemma 3.5 (*Karlin and Rinott [6]*). *Let $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, with $\boldsymbol{\Sigma}$ positive definite. The partial correlation coefficient $\rho_{ij\cdot K}$ is nonnegative for all (i, j) and $K = \{1, \ldots, d\} \setminus \{i, j\}$ if and only if $\boldsymbol{X}$ is MTP$_2$.*

Proof. Let $\boldsymbol{\Sigma}^{-1} = \boldsymbol{\Gamma} = [\gamma_{ij}]$. It suffices to show that $\rho_{ij\cdot K} \geqslant 0$ implies $\gamma_{i,j} \leqslant 0$ for $i \neq j$. Taking $i = 1$, $j = 2$, let $\boldsymbol{\Gamma}_{12}$ be the upper corner 2×2 submatrix of $\boldsymbol{\Gamma}$. Then the covariance matrix of (X_1, X_2) conditioned on the remaining variables is $\boldsymbol{\Gamma}_{12}^{-1}$. If its off diagonal element, $\rho_{12\cdot\{3,\ldots,d\}}$ is nonnegative then $\gamma_{21} \leqslant 0$. $\square$

Proof of Theorem 3.2.

(a) The result follows from Lemmas 3.4 and 3.5.

(b) Use Lemma 2.2 and the fact that $f_{\boldsymbol{X}+\boldsymbol{\mu}}(\boldsymbol{x})/f_{\boldsymbol{X}}(\boldsymbol{x}) = c\exp\{\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{x}\}$ is increasing in $\boldsymbol{x}$ iff $\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1} \geqslant \boldsymbol{0}$. $\square$

4. Properties of the cone $\mathcal{C}_{\boldsymbol{\Sigma}}$

We now study some properties of the cones introduced in Section 3.

Proposition 4.1. *We have $\mathcal{C}_{\boldsymbol{\Sigma}_1} \subseteq \mathcal{C}_{\boldsymbol{\Sigma}_2}$ if and only if $\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1} \geqslant \boldsymbol{0}$.*

Proof. The relation $\mathcal{C}_{\boldsymbol{\Sigma}_1} \subseteq \mathcal{C}_{\boldsymbol{\Sigma}_2}$ implies that for $\boldsymbol{\mu} = \boldsymbol{a}\boldsymbol{\Sigma}_1$ with $\boldsymbol{a} \geqslant \boldsymbol{0}$ there exists $\boldsymbol{b} \geqslant \boldsymbol{0}$ such that $\boldsymbol{\mu} = \boldsymbol{b}\boldsymbol{\Sigma}_2$. Thus for all $\boldsymbol{a} \geqslant \boldsymbol{0}$ there exists $\boldsymbol{b} \geqslant \boldsymbol{0}$ such that $\boldsymbol{a}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1} = \boldsymbol{b}$. This is equivalent to saying that $\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1} \geqslant \boldsymbol{0}$. $\square$

For

$$\boldsymbol{\Sigma}_\rho = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

it is easy to see that the extreme rays of the cone $\mathcal{C}_{\boldsymbol{\Sigma}_\rho}$ are generated by the points $(1, \rho)$ and $(1, 1/\rho)$.

In this particular case Proposition 4.1 yields $\mathcal{C}_{\boldsymbol{\Sigma}_{\rho'}} \subseteq \mathcal{C}_{\boldsymbol{\Sigma}_\rho}$ if and only if $\rho \leqslant \rho'$. This suggests that as dependence increases, the cones decrease. This agrees with the finding of the previous section that the cone is maximal and contains all $\boldsymbol{\mu} \geqslant \mathbf{0}$ for independent variables.

Proposition 4.2. *The relation $\boldsymbol{\Sigma}_2 \preceq \boldsymbol{\Sigma}_1$ defined as $\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1} \geqslant \mathbf{0}$ defines a partial order on correlation matrices.*

Proof. Reflexivity: If $\boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}_1$ then $\boldsymbol{\Sigma}_2 \preceq \boldsymbol{\Sigma}_1$ since $\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1} = \boldsymbol{I} \geqslant \mathbf{0}$.

Antisymmetry: Suppose that $\boldsymbol{S} := \boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1} \geqslant \mathbf{0}$ and $\boldsymbol{T} := \boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1^{-1} \geqslant \mathbf{0}$. Since $\boldsymbol{ST} = \boldsymbol{I}$, it follows that no column of $\boldsymbol{S}$ contains more than one positive element. Since $\boldsymbol{S}$ is nonsingular, each of its rows contains at least one positive element. This implies that $\boldsymbol{S}$ can be written as the product of a diagonal matrix and a permutation matrix. By reversing rows and columns the same argument leads to the fact that $\boldsymbol{T}$ has a similar structure.

We have $\boldsymbol{\Sigma}_1 = \boldsymbol{S}\boldsymbol{\Sigma}_2$. Given the structure of $\boldsymbol{S}$, if one of the positive elements of $\boldsymbol{S}$ is equal to $\alpha \neq 1$, then one of the row of $\boldsymbol{\Sigma}_1$ will be equal to α times a row of $\boldsymbol{\Sigma}_2$. This prevents $\boldsymbol{\Sigma}_1$ from being a correlation matrix. Hence $\boldsymbol{S}$ must be a permutation matrix. But, if it is not the identity, then a row of $\boldsymbol{\Sigma}_1$ will be equal to a different row of $\boldsymbol{\Sigma}_2$, which again is not compatible with the fact that $\boldsymbol{\Sigma}_1$ is a nonsingular correlation matrix. This implies that $\boldsymbol{S}$ is the identity.

Transitivity: This is trivial since $\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1} \geqslant \mathbf{0}$ and $\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_3^{-1} \geqslant \mathbf{0}$ implies $\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_3^{-1} \geqslant \mathbf{0}$. □

The next result shows that as the cones decrease, the correlations increase, and we see the connection between the cone size and the strength of positive dependence.

Proposition 4.3. *If $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ are correlation matrices such that their inverses are M-matrices, and $\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1} \geqslant \mathbf{0}$, then entry-wise $\boldsymbol{\Sigma}_1 \geqslant \boldsymbol{\Sigma}_2$.*

Proof. The relation $\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2^{-1} \geqslant \mathbf{0}$ can be written as $\boldsymbol{\Sigma}_1 = \boldsymbol{A}\boldsymbol{\Sigma}_2$ with $\boldsymbol{A} \geqslant \mathbf{0}$. Set $\boldsymbol{\Sigma}_1 = [\tau_{ij}]$, $\boldsymbol{\Sigma}_2 = [\rho_{ij}]$, and $\boldsymbol{A} = [a_{ij}]$.

We have to show that

$$\tau_{ij} = \sum_{\ell} a_{i\ell}\rho_{\ell j} \geqslant \rho_{ij}. \tag{4.1}$$

Note that $\tau_{ii} = \sum_\ell a_{i\ell}\rho_{\ell i} = 1$, and so inequality (4.1) becomes $\sum_\ell a_{i\ell}\rho_{\ell j} \geqslant \sum_\ell a_{i\ell}\rho_{\ell i}\rho_{ij}$. Using the fact that $a_{i\ell} \geqslant 0$, we now prove that the latter inequality holds term by term.

For $\ell = j$ we have $1 = \rho_{jj} \geqslant \rho_{ji}\rho_{ij}$ by the fact that $\boldsymbol{\Sigma}_2$ is positive definite. For $\ell \neq j$ we have to show $\rho_{\ell j} \geqslant \rho_{\ell i}\rho_{ij}$. We form the covariance matrix of (X_i, X_ℓ, X_j)

$$\begin{pmatrix} 1 & \rho_{i\ell} & \rho_{ij} \\ \rho_{\ell i} & 1 & \rho_{\ell j} \\ \rho_{ji} & \rho_{j\ell} & 1 \end{pmatrix}.$$

Since the inverse of the above matrix is an M-matrix, by computing minors, we see that indeed $\rho_{\ell j} \geqslant \rho_{\ell i}\rho_{ij}$. We comment that it suffices to assume that every 3-subvector has the

MTP_2 property. Note that we used the fact that MTP_2 is preserved under taking marginals (in this case of order 3), and that in the normal case MTP_2 is equivalent to the inverse of the covariance being an M-matrix. □

Remark 4.4. The converse of Proposition 4.3 is not true, that is, there exist correlation matrices $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ such that their inverses are M-matrices and $\mathbf{\Sigma}_1 \geqslant \mathbf{\Sigma}_2$, that do not satisfy the inequality $\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1} \geqslant \mathbf{0}$.

For instance take

$$\mathbf{\Sigma}_1 = \begin{pmatrix} 1.0 & 0.55 & 0.6 \\ 0.55 & 1.0 & 0.8 \\ 0.6 & 0.8 & 1.0 \end{pmatrix}, \qquad \mathbf{\Sigma}_2 = \begin{pmatrix} 1.0 & 0.5 & 0.6 \\ 0.5 & 1.0 & 0.8 \\ 0.6 & 0.8 & 1.0 \end{pmatrix}.$$

Simple algebra shows that they are indeed correlation matrices whose inverses are M-matrices, and $\mathbf{\Sigma}_1 \geqslant \mathbf{\Sigma}_2$. Nevertheless $\mathbf{\Sigma}_1\mathbf{\Sigma}_2^{-1}$ is not nonnegative.

Acknowledgements

We thank the referees for their useful comments.

References

[1] C.M. Fortuin, P.W. Kasteleyn, J. Ginibre, Correlation inequalities on some partially ordered sets, Comm. Math. Phys. 22 (1971) 89–103.
[2] C. Genest, F. Verret, The TP_2 ordering of Kimeldorf and Sampson has the normal-agreeing property, Statist. Probab. Lett. 57 (2002) 387–391.
[3] R. Holley, Remarks on the FKG inequalities, Comm. Math. Phys. 36 (1974) 227–231.
[4] H. Joe, Multivariate Models and Dependence Concepts, Chapman & Hall, London, 1997.
[5] S. Karlin, Y. Rinott, Classes of orderings of measures and related correlation inequalities, I. Multivariate totally positive distributions, J. Multivariate Anal. 10 (1980) 467–498.
[6] S. Karlin, Y. Rinott, M-matrices as covariance matrices of multinomial distributions, Linear Algebra Appl. 52/53 (1983) 419–438.
[7] J.H.B. Kemperman, On the FKG-inequality for measures on a partially ordered space, Nederl. Akad. Wetensch, Proc. Ser. A 39 (1977) 313–331.
[8] G. Kimeldorf, A.R. Sampson, Positive dependence orderings, Ann. Inst. Statist. Math. 39 (1987) 113–128.
[9] T.M. Liggett, Interacting Particle Systems, Springer, New York, 1985.
[10] P.R. Milgrom, R.J. Weber, A theory of auctions and competitive bidding, Econometrica 50 (1982) 1089–1122.
[11] A. Müller, D. Stoyan, Comparison Methods for Stochastic Models and Risks, Wiley, Chichester, 2002.
[12] C.J. Preston, A generalization of the FKG inequalities, Comm. Math. Phys. 36 (1974) 233–241.
[13] M. Shaked, J.G. Shanthikumar, Stochastic Orders and their Applications, Academic Press Inc., Boston MA, 1994.
[14] Y.L. Tong, The Multivariate Normal Distribution, Springer, New York, 1990.
[15] W. Whitt, Multivariate monotone likelihood ratio and uniform conditional stochastic order, J. Appl. Probab. 19 (1982) 695–701.