ACADEMIC PRESS

Available at WWW.MATHEMATICSWEB.ORG POWERED BY SCIENCE DIRECT®

Journal of Multivariate Analysis 85 (2003) 40–53

Journal of Multivariate Analysis

http://www.elsevier.com/locate/jmva

# A multivariate dispersion ordering based on quantiles more widely separated

J.M. Fernandez-Ponce[a] and A. Suarez-Llorens[b,*]

[a] *Dpto. Estadística e I.O, Facultad de Matemáticas, Avda. Reina Mercedes s/n, Universidad de Sevilla, 41012-Sevilla, Spain*

[b] *Dpto. Estadística e I.O, Facultad de Ciencias Empresariales, Cl. Duque de Nájera, 8, Universidad de Cádiz., 11002-Cádiz, Spain*

Received 29 June 2000

**Abstract**

A multivariate dispersion ordering based on quantiles more widely separated is defined. This new multivariate dispersion ordering is a generalization of the classic univariate version. If we vary the ordering of the components in the multivariate random variable then the comparison could not be possible. We provide a characterization using a multivariate expansion function. The relationship among various multivariate orderings is also considered. Finally, several examples illustrate the method of this paper.


*AMS 1991 subject classifications:* 60E15; 60E05; 62H99

*Keywords:* Conditional quantile; Expansion; Corrected orthant; Multivariate ordering

## 1. Introduction

Lewis and Thompson [8] introduced a concept of dispersion based on quantiles more widely separated for univariate random variables. That is, $X$ is said to be less dispersed than $Y$ in the Lewis–Thompson (LT) sense, denoted as $X \prec_{\mathrm{Disp}} Y$, if any pair of quantiles of $Y$ are at least more widely separated as corresponding quantiles of $X$. Let $u$ be a real value in $(0,1)$, we use the definition of univariate quantile as

*Corresponding author.

*E-mail addresses:* ferpon@us.es (J.M. Fernandez-Ponce), alfonso.suarez@uca.es (A. Suarez-Llorens).

0047-259X/03/$ - see front matter 
doi:10.1016/S0047-259X(02)00038-6

follows:

$$Q_X(u) = F_X^-(u) = \inf\{x : F_X(x) \geqslant u\}.$$

Then $X \prec_{\text{Disp}} Y$ is defined as $Q_X(v) - Q_X(u) \leqslant Q_Y(v) - Q_Y(u)$ for all $0 < u < v < 1$.

Shaked [10] characterized the univariate dispersive ordering using an expansion function for absolutely continuous distributions (see [10, Theorem 2.3]). Let $F$ and $G$ be two strictly increasing and absolutely continuous distribution functions, then $X \prec_{\text{Disp}} Y$ or $F \prec_{\text{Disp}} G$, if and only if there exists a function $\phi : S_F \to S_G$ (where $S_F$ and $S_G$ are the support of $F$ and $G$, respectively) such that $Y =_{\text{st}} \phi(X)$ and $\phi'(x) \geqslant 1$ for all $x$ in $S_F$. Note that $\phi$ verifies that $\phi(x) - \phi(x') \geqslant x - x'$, for all $x > x'$. Hence, the dispersion ordering in the LT sense is based on the existence of a function which depends on the corresponding distribution functions. Furthermore, in this case $\phi(x) = Q_Y(F_X(x))$ for all $x$ in $S_F$.

Giovagnoli and Wynn [5] extended the concept of dispersion ordering for multivariate distributions in a weak and strong version. Let $\mathbf{X}$ and $\mathbf{Y}$ be random vectors in $\mathbb{R}^n$ with distribution functions $F$ and $G$, respectively. From now on, we denote $\prec_{\text{st}}$ as the classical stochastic ordering for univariate and multivariate distributions and we denote as $F =_{\text{st}} G$ when $F(\mathbf{x}) = G(\mathbf{x})$ for all $\mathbf{x}$ in $\mathbb{R}^n$. The weak dispersion ordering, denoted as $\mathbf{X} \prec_{\text{D}} \mathbf{Y}$, is equivalent to $||\mathbf{X} - \mathbf{X}'||_2 \prec_{\text{st}} ||\mathbf{Y} - \mathbf{Y}'||_2$, where $||\cdot||_2$ corresponds to the Euclidean norm and $\mathbf{X}'$ and $\mathbf{Y}'$ are two independent values for each one, $\mathbf{X}$ and $\mathbf{Y}$, respectively. For instance, this weak dispersion ordering implies that $Tr(\Sigma_{\mathbf{X}}) \leqslant Tr(\Sigma_{\mathbf{Y}})$, where $\Sigma_{\mathbf{X}}$ and $\Sigma_{\mathbf{Y}}$ are the covariance matrices for $\mathbf{X}$ and $\mathbf{Y}$, respectively.

Giovagnoli and Wynn [5] defined the strong ordering if and only if exists a function $k(\cdot)$ such that $\mathbf{X} =_{\text{st}} k(\mathbf{Y})$ and $k(\cdot)$ is a contraction function of $\mathbb{R}^n$, namely

$$||k(\mathbf{y}) - k(\mathbf{x})||_2 \leqslant ||\mathbf{y} - \mathbf{x}||_2 \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$$

A contraction function is characterized through the Loewner ordering for its Jacobian matrix (see [5, Theorem 2]). That is, let $k(\cdot)$ be a continuously differentiable function. Then $k(\cdot)$ is a contraction function if and only if

$$J_k(\mathbf{x})^t J_k(\mathbf{x}) \prec_{\text{L}} I_n \quad \forall \mathbf{x} \in \mathbb{R}^n,$$

where $J_k = \{\frac{\partial k_i}{\partial x_j}\}$ is the Jacobian matrix of $k(\cdot)$, $I_n$ is the identity matrix of order $n$ and $\prec_{\text{L}}$ is the Loewner ordering of matrices such that $A \prec_{\text{L}} B$ if and only if $B - A$ is nonnegative definite.

Giovagnoli and Wynn [5] concluded that the most direct generalization of the one-dimensional case is by taking $k(\cdot)$ a 1–1 function. Thus, the function $g = k^{-1}$ can be termed as "expansion". Then it holds that $||g(\mathbf{y}) - g(\mathbf{x})||_2 \geqslant ||\mathbf{y} - \mathbf{x}||_2$ for all $\mathbf{x}, \mathbf{y}$ in $\mathbb{R}^n$ and the above characterization is replaced by

$$I_n \prec_{\text{L}} J_g(\mathbf{x})^t J_g(\mathbf{x}) \quad \forall \mathbf{x} \in \mathbb{R}^n. \tag{1}$$

To summarize, the strong dispersion ordering generalizes the LT ordering for univariate random variables but it is not so easy, in general, to find out the expression of $k(\cdot)$ and in addition this function has not to be unique.

In Section 2, we introduce a multivariate dispersion ordering based on the strong dispersion ordering which has an interpretation through the quantiles more widely separated. In Section 3, we characterize it using an expansion function for continuous distributions verifying special regularity conditions. We also study the relationship with several classical orderings in dispersion. Finally, we apply this concept to compare various multivariate distributions.

Now, we give some notation and definitions that we will be used later on. Let $\mathbf{x}$ be a vector in $\mathbb{R}^n$ and let $I=\{i_1,\ldots,i_k\}\subseteq\{1,\ldots,n\}$; then we denote $\mathbf{x}_I=(x_{i_1},\ldots,x_{i_k})$. For a random vector $\mathbf{X}$ that takes on values in $\mathbb{R}^n$, the interpretation of $\mathbf{X}_I$ is similar. From now on, we denote $\mathbf{x}\Delta_{(k_1,\ldots,k_n)}\mathbf{y}$ when $x_i\Delta_{k_i}y_i$ for $i=1,\ldots,n$ where $k_i$ is in $\{0,1\}$, $\Delta_0\overset{\text{def}}{=}\leqslant$ and $\Delta_1\overset{\text{def}}{=}\geqslant$. For example, $(x_1,x_2)\Delta_{(0,1)}(y_1,y_2)$ means that $x_1\leqslant y_1$ and $x_2\geqslant y_2$. Also, we denote $\Pi_n$ as the set of all permutations of the elements in $\{1,\ldots,n\}$. That is, $\pi_{i_1,\ldots,i_n}=(i_1,\ldots,i_n)$ is in $\Pi_n$.

## 2. Definition and properties

Let $\mathbf{X}$ be a random vector in $\mathbb{R}^n$ and $\mathbf{u}=(u_1,\ldots,u_n)$ in $[0,1]^n$. The multivariate $\mathbf{u}$-quantile for $\mathbf{X}$, denoted as $\hat{\mathbf{x}}(\mathbf{u})$, is defined as follows:

$$\begin{aligned}\hat{x}_1(u_1)&=Q_{X_1}(u_1),\\&\vdots\\\hat{x}_n(u_n)&=Q_{X_n|\bigcap_{j=1}^{n-1}X_j=\hat{x}_j(u_j)}(u_n).\end{aligned}$$

This known construction is widely used in simulation theory, and it is named the standard construction. The following result, whose proof can be seen in [11], will be used later on:

$$\hat{\mathbf{x}}(\mathbf{U})=_{\text{st}}\mathbf{X}, \tag{2}$$

where $\mathbf{U}$ is a random vector with $n$ independent uniform components in $[0,1]$.

Obviously, this standard construction depends on the choice of the ordering of the marginal distributions. Firstly, we obtain the marginal distribution $X_1$ and we construct $\hat{x}_1(u_1)$ and conditioned on every such possible realization $\hat{x}_1(u_1)$ we next construct $\hat{x}_2$. We thus have constructed so far $(\hat{x}_1,\hat{x}_2)$. Therefore, conditioned on every such possible realization $(\hat{x}_1(u_1),\hat{x}_2(u_2))$ we next construct $\hat{x}_3$. Continuing this procedure, we finally arrive at random vector $\hat{\mathbf{x}}(\mathbf{u})$. We can also consider any other permutation of the components of $\mathbf{X}$. For each permutation $\pi=\pi_{i_1,\ldots,i_n}$ in $\Pi_n$, we can use the well-known orthogonal matrix $A_\pi$ in $M_{n\times n}$, defined by $a_{ji_j}=1$ for $j=1,\ldots,n$ and zero for the rest of components. Consequently, it is trivial to show that $\mathbf{X}_\pi^t=(X_{i_1},\ldots,X_{i_n})=A_\pi\mathbf{X}$. In general, it does not hold $\mathbf{X}_\pi=_{\text{st}}\mathbf{X}$. Thus, we have to define $\hat{\mathbf{x}}_\pi(\mathbf{u})$ as the standard construction for $\mathbf{X}_\pi$. Obviously, in light of result (2), $\hat{\mathbf{x}}_\pi(\mathbf{U})=_{\text{st}}\mathbf{X}_\pi$. Hence, if $\hat{\mathbf{x}}(\mathbf{u},\pi)=A_\pi^t\hat{\mathbf{x}}_\pi(\mathbf{u})$ then $\hat{\mathbf{x}}(\mathbf{U},\pi)=_{\text{st}}\mathbf{X}$. Note that $\hat{\mathbf{x}}(\mathbf{u},\pi)$ and $\hat{\mathbf{x}}(\mathbf{u})$ have different interpretations, they provide different points in $\mathbb{R}^n$. Hereafter, we will only use the permutation $\pi=\pi_{1,\ldots,n}$ but similar results for $\mathbf{X}_\pi$ can be established.

The definition of the multivariate **u**-quantile for **X**, lead us to define the multivariate **x**-rate vector, denoted $\overset{\bigstar}{\mathbf{x}}(\mathbf{x})$, as

$$\begin{aligned} x^{\bigstar}{}_1(x_1) &= P(X_1 \leqslant x_1), \\ &\vdots \\ x^{\bigstar}{}_n(x_n) &= P(X_n \leqslant x_n | \textstyle\bigcap_{j=1}^{n-1} X_j = x_j). \end{aligned}$$

From now on, we establish the regularity conditions such as the distribution function is a continuous function and the corresponding conditional distributions in each component are continuous and strictly increasing functions.

It is easy to show that if the regularity conditions are satisfied then

$$(\hat{x}_i \circ \overset{\bigstar}{x}_i)(x_i) = x_i \quad \forall i = 1, \ldots, n. \tag{3}$$

**Theorem 2.1.** *Let* **X** *be a random vector in* $\mathbb{R}^n$ *verifying the regularity conditions and* **U** *is a random vector with n independent uniform* $[0,1]$ *components. Then* $\overset{\bigstar}{\mathbf{x}}(\mathbf{X}) =_{\text{st}} \mathbf{U}$.

**Proof.** Let $\mathbf{U} = (U_1, \ldots, U_n)$ be the multivariate **X**-rate vector, that is,

$$U_1 = \overset{\bigstar}{x}_1(X_1), \ldots, U_n = \overset{\bigstar}{x}_n(X_n).$$

According to result (3), the density function for the **U** random variable is

$$f_{\overset{\bigstar}{\mathbf{x}}(\mathbf{X})}(u_1, \ldots, u_n) = f_{\mathbf{X}}(\hat{x}_1(u_1), \ldots, \hat{x}_n(u_n)) \times |Det(J)|,$$

where the determinant of the Jacobian matrix is

$$Det(J) = \begin{vmatrix} \dfrac{\partial \hat{x}_1}{\partial u_1} & 0 & \cdots & 0 \\ \dfrac{\partial \hat{x}_2}{\partial u_1} & \dfrac{\partial \hat{x}_2}{\partial u_2} & \cdots & 0 \\ \dfrac{\partial \hat{x}_n}{\partial u_1} & \dfrac{\partial \hat{x}_n}{\partial u_2} & \cdots & \dfrac{\partial \hat{x}_n}{\partial u_n} \end{vmatrix} = \prod_{i=1}^{n} \frac{\partial \hat{x}_i(u_i)}{\partial u_i}.$$

Now, in light of the well-known inverse function theorem we have that

$$\begin{aligned} \prod_{i=1}^{n} \frac{\partial \hat{x}_i(u_i)}{\partial u_i} &= \frac{1}{f_{X_1}(\hat{x}_1(u_1))} \times \frac{1}{f_{X_2|_{X_1=\hat{x}_1(u_1)}}(\hat{x}_2(u_2))} \times \cdots \times \frac{1}{f_{X_n|\bigcap_{j=1}^{n-1} X_j=\hat{x}_j(u_j)}(\hat{x}_n(u_n))} \\ &= \frac{1}{f_{\mathbf{X}}(\hat{x}_1(u_1), \ldots, \hat{x}_n(u_n))} > 0, \end{aligned}$$

thus

$$f_{\overset{\bigstar}{\mathbf{x}}(\mathbf{X})}(u_1, \ldots, u_n) = f_{\mathbf{X}}(\hat{x}_1(u_1), \ldots, \hat{x}_n(u_n))|Det(J)| = 1,$$

where $\hat{\mathbf{x}}(\mathbf{u})$ is in the support of **X** and each value $u_i$ is given by a transformation of a distribution function for conditioned variables. Hence $u_i$ is in $(0,1)$ for $i = 1, \ldots, n$.

Consequently, $\overset{\star}{\mathbf{x}}(\mathbf{X})$ is in $(0,1)^n$ and the components of the $\overset{\star}{\mathbf{x}}(\mathbf{X})$ vector are independent. Thus, $\overset{\star}{\mathbf{x}}(\mathbf{X}) =_{\text{st}} \mathbf{U}$, where $\mathbf{U}$ is a random vector with $n$ independent uniform $(0,1)$ components. □

If the distribution function is strictly increasing and continuous then there exists only one real value for each $u$ in $(0,1)$, denoted as $Q_X(u)$, such that $F(Q_X(u)) = u$. Obviously, it is also verified that $\bar{F}(Q_X(u)) = 1-u$ where $\bar{F} = 1-F$. Therefore, each univariate quantile under the above conditions is also characterized by the survival function. This property is not so trivial to generalize for multivariate distributions. Note that, for example, if $\mathbf{x} = (x_1, x_2)$ is a point in $\mathbb{R}^2$ then it will divide the plane in four orthants. It is trivial that for the bidimensional random variables, there does not exist only one point in $\mathbb{R}^2$ such that $F(x_1, x_2) = u$ with $u$ in $(0,1)$. The solution to the equation $F(x_1, x_2) = u$ can be expressed as a function, that is $x_2 = h(x_1)$. From a stochastic point of view, the $(x_1, h(x_1))$ points could be different. Suppose that $F(x_1, x_2) = u$ and $F(z_1, z_2) = u$. However it may be possible that $P(X_1 \geqslant x_1, X_2 \geqslant x_2) \neq P(X_1 \geqslant z_1, X_2 \geqslant z_2)$. Therefore, we are interested in points in $\mathbb{R}^2$ such that they have the same probabilities in each orthant. This idea lead us to define the corrected orthant concept at points in $\mathbb{R}^n$.

Let $\mathbf{X}$ be a random vector in $\mathbb{R}^n$ with distribution function $F$ and $\mathbf{z} = (z_1, \ldots, z_n)$ a point in $\mathbb{R}^n$. The $\Delta_{(k_1,\ldots,k_n)}$-corrected orthant in $\mathbf{z}$, denoted as $R_{\mathbf{X}}(\mathbf{z}, \Delta_{(k_1,\ldots,k_n)})$, is defined as

$$R_{\mathbf{X}}(\mathbf{z}, \Delta_{(k_1,\ldots,k_n)}) = \{\mathbf{x} \in \mathbb{R}^n : x_1 \Delta_{k_1} Q_{X_1}(\overset{\star}{x}_1(z_1)), \ldots,$$

$$x_n \Delta_{k_n} Q_{X_n|\bigcap_{j=1}^{n-1} X_j = x_j}(\overset{\star}{x}_n(z_n))\}.$$

It is easy to show that if $\mathbf{X}$ is a random vector with independent components then the $\Delta_{(k_1,\ldots,k_n)}$-corrected orthant in $\mathbf{z}$ is the corresponding classic orthant in $\mathbf{z}$ for each $(k_1, \ldots, k_n)$.

**Proposition 2.1.** *Let* $\mathbf{X}$ *be a random vector in* $\mathbb{R}^n$ *with distribution function* $F$ *verifying the regularity conditions, then*

$$P\{\mathbf{X} \in R_{\mathbf{X}}(\hat{\mathbf{x}}(\mathbf{u}), \Delta_{(k_1,\ldots,k_n)})\} = \prod_{i=1}^{n} [(u_i)^{1-k_i}(1-u_i)^{k_i}]. \tag{4}$$

**Proof.** The proof is by mathematical induction. The theorem is clearly true for $n = 1$.

Suppose now that, $i = k-1$ (where $2 \leqslant k \leqslant n$), equality holds and to complete the induction argument, we must show that (4) holds for $i = n$. For this purpose, we denote $I = \{1, \ldots, n-1\}$ and

$$\mathbf{X} = (X_1, \ldots, X_{n-1}, X_n) = (\mathbf{X}_I, X_n)$$

and

$$\hat{\mathbf{x}}(\mathbf{u}) = (\hat{x_1}(u_1), \ldots, \hat{x}_{n-1}(u_{n-1}), \hat{x_n}(u_n)) = (\hat{\mathbf{x}}_I(\mathbf{u}_I), \hat{x_n}(u_n)).$$

Also, let $B(n-1) = \{\mathbf{x}_I \in \mathbb{R}^{n-1} : \mathbf{x}_I \text{ is in } R_{\mathbf{X}_I}(\hat{\mathbf{x}}_I(\mathbf{u}_I), \Delta_{(k_1,\ldots,k_{n-1})})\}$ be a subset in $\mathbb{R}^{n-1}$. Thus, by supposition,

$$\begin{aligned}\underline{P}\{\mathbf{X} \in R_{\mathbf{X}}(\hat{\mathbf{x}}(\mathbf{u}), \Delta_{(k_1,\ldots,k_n)})\} &= \int_{B(n-1)} \left[ \int_{[x_n \Delta_{k_n} Q_{X_n|\mathbf{X}_I=\mathbf{x}_I}(u_n)]} dF_{X_n|\mathbf{X}_I=\mathbf{x}_I}(x_n) \right] dF_{\mathbf{X}_I}(\mathbf{x}_I). \\ &= \left\{ \prod_{i=1}^{n-1} ((u_i)^{1-k_i}(1-u_i)^{k_i}) \right\} [(u_n)^{1-k_n}(1-u_n)^{k_n}] \\ &= \prod_{i=1}^{n} [(u_i)^{1-k_i}(1-u_i)^{k_i}]. \qquad \square\end{aligned}$$

The $\Delta_{(k_1,\ldots,k_n)}$-corrected orthants in the corresponding multivariate $\mathbf{u}$-quantile accumulate the same probability for two random vectors with distribution functions under the regularity conditions. These considerations lead us to define a new dispersive ordering based on conditional quantiles more widely separated.

**Definition 2.1.** Let $\mathbf{X}$ and $\mathbf{Y}$ be two random vectors in $\mathbb{R}^n$. We say that $\mathbf{X}$ is less than $\mathbf{Y}$ in dispersion sense, denoted as $\mathbf{X} \prec_{\text{Disp}} \mathbf{Y}$, if

$$||\hat{\mathbf{x}}(\mathbf{v}) - \hat{\mathbf{x}}(\mathbf{u})||_2 \leqslant ||\hat{\mathbf{y}}(\mathbf{v}) - \hat{\mathbf{y}}(\mathbf{u})||_2,$$

for all $\mathbf{u}$ and $\mathbf{v}$ in $(0,1)^n$.

It is easy to prove that this ordering is a generalization of dispersive ordering in the LT sense for any two random variables.

Note that this new ordering depends on the chosen permutation. If $\mathbf{X} \prec_{\text{Disp}} \mathbf{Y}$, then it could not be held that $A_\pi \mathbf{X} \prec_{\text{Disp}} A_\pi \mathbf{Y}$ for any $\pi = \pi_{i_1,\ldots,i_n}$ in $\Pi_n$. We provide a counter-example in Section 4 in this sense (see Remark 4.1). Thus we define that $\mathbf{X}$ is less than $\mathbf{Y}$ in dispersion sense, under a permutation $\pi = \pi_{i_1,\ldots,i_n}$ in $\Pi_n$, if $A_\pi \mathbf{X} \prec_{\text{Disp}} A_\pi \mathbf{Y}$. All considerations for the rate vector and for the corrected orthants are possible for the dispersion ordering under a permutation no more to take a permutation of the components.

We finish this section introducing two lemmas which we will be used in the next results.

**Lemma 2.1.** *Let $A \in M_{n\times n}$ be a lower triangular matrix, such that $A^t A - Id_{nxn}$ is nonnegative definite. Then the diagonal elements verify that $a_{ii}^2 \geqslant 1$, for $i = 1, \ldots, n$.*

**Proof.** By hypothesis $\mathbf{x}^t A^t A \mathbf{x} \geqslant \mathbf{x}^t \mathbf{x}$ for all $\mathbf{x}$. Let $\mathbf{u}_i$ be the normalized eigenvector associated to the eigenvalue $a_{ii}$ for $A$. Then, $\mathbf{u}_i^t A^t A \mathbf{u}_i \geqslant \mathbf{u}_i^t \mathbf{u}_i$. Consequently, it holds $a_{ii}^2 \geqslant 1$. $\square$

**Lemma 2.2.** *Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *be any two random vectors in* $\mathbb{R}^n$. *If* $\mathbf{X} \prec_{\text{st}} \mathbf{Y}$ *then* $\mathbf{X}_I \prec_{\text{st}} \mathbf{Y}_I$ *for each* $I \subset \{1, \ldots, n\}$.

**Proof.** See [11, Theorem 4.B.10(c)]. □

## 3. The main theorem

**Theorem 3.1.** *Let* $F$ *and* $G$ *be distribution functions under the regularity conditions. Assume that the supports* $S_F$ *and* $S_G$ *are intervals in* $\mathbb{R}^n$. *Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *be random variables distributed, respectively, according to* $F$ *and* $G$. *Then* $\mathbf{X} \prec_{\text{Disp}} \mathbf{Y}$, *if and only if there exists a function* $\Phi : S_F \to S_G$ *such that*

$$\Phi(\mathbf{X}) =_{\text{st}} \mathbf{Y} \quad \text{with } [\Phi(\mathbf{x})]_i = \Phi_i(x_1, \ldots, x_i) \quad (i = 1, \ldots, n) \tag{5}$$

*and the Jacobian matrix of* $\Phi$ *satisfies*

$$I_n \prec_{\text{L}} J_\Phi(\mathbf{x})^t J_\Phi(\mathbf{x}) \quad \text{for all } \mathbf{x} \in \mathbb{R}^n \tag{6}$$

*and*

$$\frac{\partial \Phi_i(x_1, \ldots, x_i)}{\partial x_i} \geqslant 0. \tag{7}$$

*Moreover, if this is the case then*

$$\Phi_i(x_1, \ldots, x_i) = (\hat{y}_i \circ \overset{\star}{x}_i)(x_i) \quad (i = 1, \ldots, n). \tag{8}$$

**Proof.** Firstly, we will show that the only one function which verifies conditions (5)–(7) is the function given by (8). The proof is by mathematical induction. If $\mathbf{Y} =_{\text{st}} \Phi(\mathbf{X})$, according to Lemma 2.2 then $Y_1 =_{\text{st}} \Phi_1(X_1)$. On the other hand, in light of conditions (6), (7) and Lemma 2.1, then $\partial \Phi_1 / \partial x_1 \geqslant 1$. Thus, it follows from Theorem 2.3 in [10] that $\Phi_1(x_1) = (\hat{y}_1 \circ \overset{\star}{x}_1)(x_1)$. Now, assume that (8) is verified for $i = 1, \ldots, n-1$. Now, we must show that condition (8) holds for $i = n$. Using Lemma 2.2, we obtain that $Y_n =_{\text{st}} \Phi_n(X_1, \ldots, X_n)$ and $(\Phi_1(X_1), \ldots, \Phi_{n-1}(X_1, \ldots, X_{n-1})) =_{\text{st}} \mathbf{Y}_I$. Consequently, for a given vector $\mathbf{y}_I = (y_1, \ldots, y_{n-1})$ it holds

$$Y_n|_{\mathbf{Y}_I = \mathbf{y}_I} =_{\text{st}} \Phi_n(X_1, \ldots, X_n)|_{(\Phi_1(X_1) = y_1, \ldots, \Phi_{n-1}(\mathbf{X}_I) = y_{n-1})}.$$

By the induction hypothesis, we know that the system

$$\begin{aligned} \Phi_1(x_1) &= y_1, \\ \Phi_2(x_1, x_2) &= y_2, \\ &\vdots \\ \Phi_{n-1}(x_1, \ldots, x_{n-1}) &= y_{n-1} \end{aligned}$$

has a only one solution in a vector $\mathbf{x}_I = (x_1, \ldots, x_{n-1})$ for each $\mathbf{y}_I$. Therefore,

$$\Phi_n(X_1, \ldots, X_n)|_{(\Phi_1(X_1) = y_1, \ldots, \Phi_{n-1}(X_1, \ldots, X_{n-1}) = y_{n-1})} =_{\text{st}} \Phi_n(\mathbf{x}_I, X_n|_{\mathbf{X}_I = \mathbf{x}_I}).$$

Thus we have that there exists a transformation which maps $X_n|_{\mathbf{X}_I=\mathbf{x}_I}$ to $Y_n|_{\mathbf{Y}_I=\mathbf{y}_I}$. Using (6), (7) and Lemma 2.1, it holds that $\frac{\partial \Phi_n}{\partial x_n} \geqslant 1$. Hence, according to Theorem 2.3 in [10] for a fixed vector $\mathbf{x}_I$ we obtain

$$\Phi_n(\mathbf{x}_I, x_n) = Q_{Y_n|_{\mathbf{Y}_I=\mathbf{y}_I}}(F_{X_n|_{\mathbf{X}_I=\mathbf{x}_I}}(x_n)) = (\hat{y}_n \circ \overset{\star}{x}_n)(x_n).$$

We conclude the induction argument with this last result. Now, we will show that $\mathbf{X} \prec_{\text{Disp}} \mathbf{Y}$ is a necessary condition for (5)–(7). In light of expression of $\Phi$ in (8), it is easy to show that this function maps conditional quantiles for $\mathbf{X}$ to conditional quantiles for $\mathbf{Y}$. Making use of condition (6) it holds that $\Phi$ is an expansion function, (see condition (1)). Consequently,

$$\begin{aligned} ||\hat{\mathbf{x}}(\mathbf{v}) - \hat{\mathbf{x}}(\mathbf{u})||_2 &\leqslant ||\Phi(\hat{\mathbf{x}}(\mathbf{v})) - \Phi(\hat{\mathbf{x}}(\mathbf{u}))||_2 \\ &= ||\hat{\mathbf{y}}(\mathbf{v}) - \hat{\mathbf{y}}(\mathbf{u})||_2 \end{aligned}$$

for all $\mathbf{u}, \mathbf{v}$ in $(0,1)^n$, then $\mathbf{X} \prec_{\text{Disp}} \mathbf{Y}$.

Now, we will show that $\mathbf{X} \prec_{\text{Disp}} \mathbf{Y}$ is a sufficient condition for (5)–(7). According to Theorem 2.1 and result (2) we find that $\Phi(\mathbf{X}) = (\hat{\mathbf{y}} \circ \overset{\star}{\mathbf{x}})(\mathbf{X}) =_{\text{st}} \mathbf{Y}$.

We have only to prove (6) and (7). Since condition (1), we must show that $\Phi$ is an expansion function. From the definition of dispersive ordering it holds

$$\begin{aligned} ||\Phi(\mathbf{x}_1) - \Phi(\mathbf{x}_2)||_2 &= ||(\hat{\mathbf{y}} \circ \overset{\star}{\mathbf{x}})(\mathbf{x}_1) - (\hat{\mathbf{y}} \circ \overset{\star}{\mathbf{x}})(\mathbf{x}_2)||_2 \\ &\geqslant ||(\hat{\mathbf{x}} \circ \overset{\star}{\mathbf{x}})(\mathbf{x}_1) - (\hat{\mathbf{x}} \circ \overset{\star}{\mathbf{x}})(\mathbf{x}_2)||_2 = ||\mathbf{x}_1 - \mathbf{x}_2||_2 \quad \forall \mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n. \end{aligned}$$

Thus, $\Phi$ is an expansion function and making use of condition (1) we obtain (5). Finally, $\partial \Phi_i / \partial x_i$ can be regarded as the ordinary derivative of the function of one variable obtained from $\Phi_i(x_1, \ldots, x_i)$ by fixing $(x_1, \ldots, x_{i-1})$ at $x_i$. If $b \geqslant a$ then $\overset{\star}{x}_i(b) \geqslant \overset{\star}{x}_i(a)$. Thus,

$$\Phi_i(x_1, \ldots, x_{i-1}, b) - \Phi_i(x_1, \ldots, x_{i-1}, a) = (\hat{y}_i \circ \overset{\star}{x}_i)(b) - (\hat{y}_i \overset{\star}{x}_i)(a) \geqslant 0,$$

because the quantile function is an increasing function. Therefore, we have that (7) holds. □

Note that the jacobian matrix of $\Phi$ is a lower triangular matrix. If we are interested to study the dispersion ordering under the permutation $\pi_{i_1,\ldots,i_n}$ in $\Pi_n$ then we have to compare in dispersion $A_\pi \mathbf{X}$ versus $A_\pi \mathbf{Y}$. If $A_\pi \mathbf{X} \prec_{\text{Disp}} A_\pi \mathbf{Y}$ then there exists an expansion function which maps $A_\pi \mathbf{X}$ to $A_\pi \mathbf{Y}$. That is $\Phi(A_\pi \mathbf{X}) =_{\text{st}} A_\pi \mathbf{Y}$. Considering the function $\Upsilon = A_\pi^t \circ \Phi \circ A_\pi$ we observe that $\Upsilon(\mathbf{X}) =_{\text{st}} \mathbf{Y}$. Since $A_\pi$ is orthogonal, it is

easy to show that $\Upsilon$ is also an expansion function. Obviously, the $J_\Phi^t J_\Phi - I_n$ and $J_\Upsilon^t J_\Upsilon - I_n$ matrices have the same eigenvalues. To summarize, if $\mathbf{X}$ and $\mathbf{Y}$ are ordered in dispersion under a permutation $\pi_{i_i,\ldots,i_n}$ in $\Pi_n$, then there exists an unique expansion function based on the permutation $\pi_{i_1,\ldots,i_n}$ which maps $\mathbf{X}$ to $\mathbf{Y}$.

**Corollary 3.1.** *Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *be two random vectors in* $\mathbb{R}^n$. *If* $\mathbf{X} \prec_{\text{Disp}} \mathbf{Y}$, *then*

$$X_j|_{\mathbf{X}_I=\hat{\mathbf{x}}_I(\mathbf{u}_I)} \prec_{\text{Disp}} Y_j|_{\mathbf{Y}_I=\hat{\mathbf{y}}_I(\mathbf{u}_I)}, \tag{9}$$

*where* $I = (1, \ldots, j-1), \mathbf{u}_I = (u_1, \ldots, u_{j-1})$ *in* $(0,1)^{j-1}$ *and* $j = 2, \ldots, n$.

**Proof.** Suppose that $\mathbf{X} \prec_{\text{Disp}} \mathbf{Y}$. Then according to Theorem 3.1 we find that

$$\Phi_j(\hat{\mathbf{x}}_I(\mathbf{u}_I), X_j|_{\mathbf{X}_I=\hat{\mathbf{x}}_I(\mathbf{u}_I)}) =_{\text{st}} Y_j|_{\mathbf{Y}_I=\hat{\mathbf{y}}_I(\mathbf{u}_I)},$$

with $\partial\Phi_j/\partial x_j \geqslant 1$. Consequently there exists an expansion function which maps $X_j|_{\mathbf{X}_I=\hat{\mathbf{x}}_I(\mathbf{u}_I)}$ to $Y_j|_{\mathbf{Y}_I=\hat{\mathbf{y}}_I(\mathbf{u}_I)}$, for $j = 2, \ldots, n$. It follows from Theorem 2.3 in [10] that the result holds. □

The multivariate dispersive ordering implies the univariate dispersive ordering for conditional distributions. However, it is not always true in the other direction. Further on, we will show a counter-example in Section 4, (see Remark 4.3).

Theorem 3.1 has the following implication.

**Corollary 3.2.** *Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *be two random vectors in* $\mathbb{R}^n$. *Let* $f_{\mathbf{X}}$ *and* $f_{\mathbf{Y}}$ *their density functions, respectively. If* $\mathbf{X} \prec_{\text{Disp}} \mathbf{Y}$, *then*

$$f_{\mathbf{X}}(\hat{\mathbf{x}}(\mathbf{u})) \geqslant f_{\mathbf{Y}}(\hat{\mathbf{y}}(\mathbf{u})) \quad \forall \mathbf{u} \in (0,1)^n.$$

In the special case of independent components, we obtain the following corollaries

**Corollary 3.3.** *Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *be two random vectors with independent components. Then*

$$\mathbf{X} \prec_{\text{Disp}} \mathbf{Y} \text{ if and only if } X_i \prec_{\text{Disp}} Y_i \quad \text{for } i = 1, \ldots, n.$$

**Corollary 3.4.** *Let* $\mathbf{X}_1$ *and* $\mathbf{Y}_1$ *be two random vectors in* $\mathbb{R}^{n_1}$ *and let* $\mathbf{X}_2$ *and* $\mathbf{Y}_2$ *in* $\mathbb{R}^{n_2}$, *where* $\mathbf{X}_1, \mathbf{X}_2$ *are independent and* $\mathbf{Y}_1, \mathbf{Y}_2$ *are too independent. If* $X_1 \prec_{\text{Disp}} Y_1$ *and* $X_2 \prec_{\text{Disp}} Y_2$ *then*

$$(X_1^t, X_2^t)^t \prec_{\text{Disp}} (Y_1^t, Y_2^t)^t.$$

### 3.1. The ordering $\prec_{\text{Disp}}$ compared to other multivariate orderings

It is easy to show that the dispersive ordering implies the strong dispersion ordering (see [5]). We can say that the multivariate dispersion ordering is a case of the strong ordering where the expression of the expansion function is known and it has an interpretation under the corrected orthants more widely separated. Oja [9] interpreted the Jacobian matrix of the expansion in terms of local volume elements.

**Definition 3.1.** Let $\mathbf{X}$ and $\mathbf{Y}$ two random vectors in $\mathbb{R}^n$. We say that $\mathbf{Y}$ is more scattered than $\mathbf{X}$, denoted $\prec_{\Delta}$, if there is a function $g : \mathbb{R}^n \mapsto \mathbb{R}^n$ such that $g(\mathbf{X}) \sim_{\text{st}} \mathbf{Y}$ and for all $\{x_1, \ldots, x_{n+1}\} \subset \mathbb{R}^n$ it holds that

$$\Delta(g(x_1), \ldots, g(x_{n+1})) \geqslant \Delta(x_1, \ldots, x_{n+1}),$$

where $\Delta(x_1, \ldots, x_{n+1})$ is the volume of the "simplex" with vertices at $x_1, \ldots, x_{n+1}$.

From Lemma 3 in [5] we obtain that if $\mathbf{X} \prec_{\text{Disp}} \mathbf{Y}$ then $\mathbf{X} \prec_{\Delta} \mathbf{Y}$.

Block and Sampson [3] introduced the concept of conditionally more dispersed. Fixed $1 \leqslant i \leqslant n$, denote $I(i) = \{1, \ldots, i-1, i+1, \ldots, n\}$. Let denote the number of sign changes of a function $a$ as $S^-(a(x))$.

**Definition 3.2.** Let $\mathbf{X}$ and $\mathbf{Y}$ be two random vectors in $\mathbb{R}^n$ with distribution functions $F$ and $G$, respectively. Fixed $1 \leqslant i \leqslant n$, and suppose the following conditions are satisfied:

1. $F_{\mathbf{X}_{I(i)}}(\mathbf{t}) = G_{\mathbf{Y}_{I(i)}}(\mathbf{t})$ for all $\mathbf{t}$.
2. $E(X_i|_{\mathbf{X}_{I(i)}=\mathbf{t}}) = E(Y_i|_{\mathbf{Y}_{I(i)}=\mathbf{t}})$ for all $\mathbf{t}$.
3. For all $\mathbf{t}$, both conditional distributions are degenerated, or

(i) $S(F(x_i|_{\mathbf{X}_{I(i)}=\mathbf{t}}) - G(x_i|_{\mathbf{Y}_{I(i)}=\mathbf{t}})) = 1$, and
(ii) the sign sequence is $-, +$.

Then $\mathbf{X}$ is said to be conditionally less $i$-dispersed than $\mathbf{Y}$, denoted as $\mathbf{Y} \overset{D(i)}{\rightarrow} \mathbf{X}$.

Condition 3 in Definition 3.2 is known as the criterion of Karlin–Novikoff for univariate variables. Conditions given in Definition 3.2 are interpreted as the conditional distribution $X_i|_{\mathbf{X}_{I(i)}=\mathbf{t}}$ is less in residual life than $Y_i|_{\mathbf{Y}_{I(i)}=\mathbf{t}}$ for all $\mathbf{t}$ (see [12]).

Let $\mathbf{X}$ and $\mathbf{Y}$ be two random vectors in $\mathbb{R}^n$. Assume conditions 1 and 2 in Definition 3.2, if $\mathbf{X} \prec_{\text{Disp}} \mathbf{Y}$ then

$$X_n|_{\mathbf{X}_{I(n)}=\mathbf{t}} \prec_{\text{Disp}} Y_n|_{\mathbf{Y}_{I(n)}=\mathbf{t}},$$

for all $\mathbf{t}$. Condition 1 in Definition 3.3 implies that $\mathbf{X}_{I(n)}$ and $\mathbf{Y}_{I(n)}$ have the same distribution. Therefore it is easy to show that $\hat{\mathbf{x}}_{I(n)} = \hat{\mathbf{y}}_{I(n)}$. Moreover, the ordering in LT sense implies the convex ordering, (see [10,12]). Therefore, according to Corollary 3.1, we have that

$$X_n|_{\mathbf{X}_{I(n)}=\mathbf{t}} \prec_{\mathrm{c}} Y_n|_{\mathbf{Y}_{I(n)}=\mathbf{t}},$$

where we denote $\prec_{\mathrm{c}}$ as the convex ordering. It is not always true that $\prec_{\mathrm{c}}$ implies the criterion of Karlin–Novikoff. Under symmetry conditions, both orderings are equivalent (see [2]).

On the other hand, Block and Sampson [3] interpreted the conditionally dispersive ordering under the convexity in the one relevant component for all other values of the remaining component. The concept of conditionally more dispersed ordering has not sense when there are more than one relevant components.

## 4. Examples

**Example 4.1.** Let $\mathbf{X} \rightsquigarrow N_n(\mu_1, \Sigma_1)$ and $\mathbf{Y} \rightsquigarrow N_n(\mu_2, \Sigma_2)$ be two multivariate normal distributions. In this case, it is easy to show by mathematical induction that the function $\Phi$ defined in (8) satisfies $J_\Phi = AB$ where $A$ and $B$ are two lower triangular matrices with $AA^t = \Sigma_2$ and $B^tB = \Sigma_1^{-1}$. Furthermore, according to Theorem 14.5.11 in [6], we have that

$$A^t = \mathbf{D}_A^{1/2}\mathbf{U}, \quad B = \mathbf{D}_B^{-1/2}(\mathbf{V}^{-1})^t \quad \text{and} \quad \Sigma_1 = \mathbf{V}^t\mathbf{D}_B\mathbf{V},$$

with $\mathbf{U}$ be the unique unit upper triangular matrix and $\mathbf{D}_A = \{d_i\}$ be the unique diagonal matrix such that

$$\Sigma_2 = \mathbf{U}^t\mathbf{D}_A\mathbf{U} \quad \text{and} \quad \mathbf{D}_A^{1/2} = \{\sqrt{d_i}\}.$$

Similarly for the $B$ matrix. The $\mathbf{U}$ and $\mathbf{V}$ matrices can be calculated using the Cholesky decomposition (see [6]). Consequently, in light of Theorem 3.1, $\mathbf{X} \prec_{\mathrm{Disp}} \mathbf{Y}$ if and only if $I_n \preccurlyeq_{\mathrm{L}} (AB)^t AB$.

Now, we show an example in the bidimensional case. We will attend to the well-known expression for the density function in [7]. Let $\mathbf{Z} \rightsquigarrow N(\mu_{\mathbf{Z}}, \Sigma_{\mathbf{Z}})$ be a normal distribution in $\mathbb{R}^2$. The density function is given by

$$f_{\mathbf{Z}}(z_1, z_2) = (2\pi\sqrt{(1-\rho_{\mathbf{Z}}^2)})^{-1} \exp\left\{\frac{-1}{2(1-\rho_{\mathbf{Z}}^2)}\left[\left(\frac{z_1-\mu_{1\mathbf{Z}}}{\sigma_{1\mathbf{Z}}}\right)^2 - 2\rho_{\mathbf{Z}}\left(\frac{z_1-\mu_{1\mathbf{Z}}}{\sigma_{1\mathbf{Z}}}\right)\left(\frac{z_2-\mu_{2\mathbf{Z}}}{\sigma_{2\mathbf{Z}}}\right) + \left(\frac{z_2-\mu_{2\mathbf{Z}}}{\sigma_{2\mathbf{Z}}}\right)^2\right]\right\},$$

where $\rho_{\mathbf{Z}}$ is the linear correlation coefficient between $Z_1$ and $Z_2$.

It is widely known that the conditional distributions for multivariate normal distribution are also normal distributions. Then, the distribution of $Z_2$ conditioned

to $Z_1$ is given by

$$Z_2|_{Z_1=z_1} \rightsquigarrow N\left(\mu_{2\mathbf{Z}} + (z_1 - \mu_{1\mathbf{Z}})\frac{\sigma_{12\mathbf{Z}}}{\sigma^2_{1\mathbf{Z}}}, \sigma_{2\mathbf{Z}}\sqrt{(1-\rho^2_{\mathbf{Z}})}\right).$$

Let $\mathbf{X} \rightsquigarrow N(\mu_{\mathbf{X}}, \Sigma_{\mathbf{Y}})$ and $\mathbf{Y} \rightsquigarrow N(\mu_{\mathbf{Y}}, \Sigma_{\mathbf{Y}})$ be normal distributions. The expression of $\Phi$ is given by

$$\Phi_1(x_1) = \sigma_{1\mathbf{Y}}\left(\frac{x_1 - \mu_{1\mathbf{X}}}{\sigma_{1\mathbf{X}}}\right) + \mu_{1\mathbf{Y}},$$

$$\Phi_2(x_1, x_2) = \sigma_{2\mathbf{Y}}\sqrt{(1-\rho^2_{\mathbf{Y}})}\left\{\frac{x_2 - [\mu_{2\mathbf{X}} + (x_1 - \mu_{1\mathbf{X}})\frac{\sigma_{12\mathbf{X}}}{\sigma^2_{1\mathbf{X}}}]}{\sigma_{2\mathbf{X}}\sqrt{(1-\rho^2_{\mathbf{X}})}}\right\} + \mu_{2\mathbf{Y}} + \left(\frac{x_1 - \mu_{1\mathbf{X}}}{\sigma_{1\mathbf{X}}}\right)\frac{\sigma_{12\mathbf{Y}}}{\sigma_{1\mathbf{Y}}},$$

where $\Phi_1(x_1) = Q_{Y_1}(F_{X_1}(x_1))$ and $\Phi_2(x_1, x_2) = Q_{Y_2|_{Y_1=\Phi_1(x_1)}}(F_{X_2/X_1=x_1}(x_2))$.

The Jacobian matrix of $\Phi$ is given by:

$$J_\Phi = \begin{pmatrix} \dfrac{\sigma_{1\mathbf{Y}}}{\sigma_{1\mathbf{X}}} & 0 \\ \dfrac{\sigma_{2\mathbf{Y}}}{\sigma_{1\mathbf{X}}}\left(\rho_{\mathbf{Y}} - \rho_{\mathbf{X}}\dfrac{\sqrt{(1-\rho^2_{\mathbf{Y}})}}{\sqrt{(1-\rho^2_{\mathbf{X}})}}\right) & \dfrac{\sigma_{2\mathbf{Y}}\sqrt{(1-\rho^2_{\mathbf{Y}})}}{\sigma_{2\mathbf{X}}\sqrt{(1-\rho^2_{\mathbf{X}})}} \end{pmatrix}.$$

As a particular case, if $\rho_{\mathbf{X}} = \rho_{\mathbf{Y}}$, $\sigma_{1\mathbf{Y}} \geqslant \sigma_{1\mathbf{X}}$ and $\sigma_{2\mathbf{Y}} \geqslant \sigma_{2\mathbf{X}}$ then $\mathbf{X} \prec_{\text{Disp}} \mathbf{Y}$.

**Remark 4.1.** Let $\mathbf{X}$ and $\mathbf{Y}$ be as the above example. If $\mathbf{X} \prec_{\text{Disp}} \mathbf{Y}$ then it could not be held $A_\pi \mathbf{X} \prec_{\text{Disp}} A_\pi \mathbf{Y}$ for any $\pi$ in $\Pi_n$. Let $\sigma_{1\mathbf{Y}} = 8$, $\sigma_{2\mathbf{Y}} = 2$, $\rho_{\mathbf{Y}} = 0$, $\sigma_{1\mathbf{X}} = 1$, $\sigma_{2\mathbf{X}} = 3$ and $\rho^2_{\mathbf{X}} = 0,75$, then it holds $\mathbf{X} \prec_{\text{Disp}} \mathbf{Y}$. However, since $\sigma_{2\mathbf{X}} > \sigma_{2\mathbf{Y}}$, it does not hold the ordering $(X_2, X_1) \prec_{\text{Disp}} (Y_2, Y_1)$. In addition, it does not hold either the ordering given by Eaton and Pearlmon [4] which expresses concentration in the following way $\Sigma_1 \preccurlyeq_{\text{L}} \Sigma_2$.

**Remark 4.2.** The expansion function $\Phi$ is not always a linear function. For example, let $(X_1, X_2)$ be a bidimensional random variable with normal distribution and independent components and let $(U_1, U_2)$ be a bidimensional random variable with uniform distribution and independent components in $(0,1)^2$. It is easy to show that $U_i \prec_{\text{Disp}} X_i$ for $i = 1, 2$. Then, according to Corollary 3.3, we have that $(U_1, U_2) \prec_{\text{Disp}} (X_1, X_2)$. Consequently, $\hat{\mathbf{x}}(\mathbf{U}) =_{\text{st}} \mathbf{X}$ with $\hat{\mathbf{x}}$ is an expansion function and obviously it is not a linear function.

**Example 4.2.** Let $\mathbf{S}_1 \rightsquigarrow W_n(v, \Sigma_1)$ and $\mathbf{S}_2 \rightsquigarrow W_n(v, \Sigma_2)$ be two Wishart distributions defined through $v$ normal distributions $N(\mathbf{0}, \Sigma_1)$ and $N(\mathbf{0}, \Sigma_2)$, respectively. Let

consider $\Phi(\mathbf{S}_1) = AB\mathbf{S}_1(AB)^t$ where $AA^t = \Sigma_2$ and $B^tB = \Sigma_1^{-1}$ as the previous example. Anderson [1] showed that $\Phi(\mathbf{S}_1) =_{\text{st}} \mathbf{S}_2$.

Furthermore, according to Theorem 16.2.1 in [6], we have that

$$\Phi(vec(\mathbf{S}_1)) = [(AB)\otimes(AB)]vec(\mathbf{S}_1).$$

Consequently, $W_n(v,\Sigma_1) \prec_{\text{Disp}} W_n(v,\Sigma_2)$ if and only if

$$I_{n^2} \preccurlyeq_{\text{L}} [(AB)\otimes(AB)]^t[(AB)\otimes(AB)],$$

that is (in light of [6, Theorem 16.1.2])

$$I_{n^2} \preccurlyeq_{\text{L}} [(AB)^t(AB)]\otimes[(AB)^t(AB)].$$

Therefore, if $N(0,\Sigma_1) \prec_{\text{Disp}} N(0,\Sigma_2)$, (since $I_n \preccurlyeq_{\text{L}} (AB)^t(AB)$) then $W_n(v,\Sigma_1) \prec_{\text{Disp}} W_n(v,\Sigma_2)$.

**Example 4.3.** We introduce an extension of the exponential multivariate of Freund (see [7]). Suppose that a system has $m$ identical components, and times to failure $X_1,\ldots,X_m$. All components come from an exponential distribution

$$f_X(x) = \theta_0^{-1}\exp(-x/\theta_0), \quad x>0, \quad \theta_0>0.$$

If $k$ components have failed and they have not been replaced, then the conditional joint distribution of the lifetimes of the remaining $(m-k)$ components is easily obtained. The joint density of the ordered variables $X_1' \leqslant \cdots \leqslant X_m'$ is

$$f(x_1,\ldots,x_m) = m!\prod_{j=0}^{m-1}[\theta_j^{-1}\exp\{-(m-j)\theta_j^{-1}(x_{j+1}-x_j)\}], \tag{10}$$

where $x_0 = 0$ and $x_1 \leqslant x_2 \leqslant \cdots \leqslant x_m$.

Let $\mathbf{X}$ and $\mathbf{Y}$ two random vectors with density function according to (10), with parameter vectors $\theta = (\theta_0,\ldots,\theta_{m-1})$ and $\theta' = (\theta_0',\ldots,\theta_{m-1}')$ respectively. The dispersive ordering let us choose a permutation. Because of the nature of this exponential distribution, it is especially interesting $\pi_{1,\ldots,m}$. Then

$$\Phi_1(x_1) = \frac{\theta_0'}{\theta_0}x_1,$$
$$\vdots$$
$$\Phi_j(x_1,\ldots,x_j) = \frac{\theta_{j-1}'}{\theta_{j-1}}(x_j - x_{j-1}) + \Phi_{j-1}(x_1,\ldots,x_{j-1}),$$

for $j = 1,\ldots,m$.

Let $A(\theta)$ be a lower triangular matrix such that $a_{ij} = \theta_{j-1}/\sqrt{m-j+1}$, $i \leqslant j$ and zero for the rest of components. Similarly for $A(\theta')$. It is easy to show that $A(\theta)A^t(\theta) = \Sigma_{\mathbf{X}}$, $A(\theta')A^t(\theta') = \Sigma_{\mathbf{Y}}$. From this one, the Jacobian matrix of $\Phi$ is

expressible as

$$J_{\Phi}(x_1,\ldots,x_m)=\begin{pmatrix} \frac{\theta_0'}{\theta_0} & 0 & 0 & \cdots & 0 \\ \frac{\theta_0'}{\theta_0}-\frac{\theta_1'}{\theta_1} & \frac{\theta_1'}{\theta_1} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \cdots & \vdots \\ \frac{\theta_0'}{\theta_0}-\frac{\theta_1'}{\theta_1} & \frac{\theta_1'}{\theta_1}-\frac{\theta_2'}{\theta_2} & \frac{\theta_2'}{\theta_2}-\frac{\theta_3'}{\theta_3} & \cdots & \frac{\theta_{m-1}'}{\theta_{m-1}} \end{pmatrix},$$

where $J_{\Phi}=A(\theta')A^{-1}(\theta)$.

If $\theta_i'/\theta_i=a>1$, for $i=0,\ldots,m-1$ then $\mathbf{X}\prec_{\mathrm{Disp}}\mathbf{Y}$.

**Remark 4.3.** We provide a counter-example for the result obtained in Corollary 3.1. Assume that $m=2$, $\theta_0'/\theta_0=1.1$ and $\theta_1'/\theta_1=1.9$ in Example 4.3. It is easy to show that one eigenvalue of the $J_{\Phi}^t J_{\Phi}$ matrix is lower than 1. Thus, $\mathbf{X}\not\prec_{\mathrm{Disp}}\mathbf{Y}$. However, it holds that

$$X_1\prec_{\mathrm{Disp}}Y_1 \quad\text{and}\quad X_2|_{X_1=\hat{x}_1(u_1)}\prec_{\mathrm{Disp}}Y_2|_{Y_1=\hat{y}_1(u_1)}.$$

## Acknowledgments

We thank referees whose careful reading and marking in the first version of this manuscript led to a number of corrections.

## References

[1] T.W. Anderson, An Introduction to Multivariate Statistical Analysis, Wiley, New York, 1958.
[2] R.E. Barlow, F. Proshan, Statistical Theory of Reliability and Life Testing: Probability Models, Hold, Rinehart and Winston, New York, 1975.
[3] H.W. Block, A.R. Sampson, Conditionally ordered distributions, J. Multivariate Anal. 27 (1988) 91–104.
[4] M.L. Eaton, M.D. Perlman, Concentration and inequalities for multivariate distributions: multivariate normal distributions, Statist. Probab. Lett. 12 (1991) 487–504.
[5] A. Giovagnoli, H.P. Wynn, Multivariate dispersion orderings, Statist. Probab. Lett. 22 (1995) 325–332.
[6] D.A. Harville, Matrix Algebra from a Statistician's Perspective, Springer, New York, 1997.
[7] N. Johnson, S. Kotz, Distributions in Statistics: Continuous Multivariate Distributions, Wiley, New York, 1976.
[8] T. Lewis, J.W. Thompson, Dispersive distributions and the connection between dispersivity and strong unimodality, J. Appl. Probab. 18 (1981) 76–90.
[9] H. Oja, Descriptive statistics for multivariate distributions, Statist. Probab. Lett. 1 (1983) 327–332.
[10] M. Shaked, Dispersive ordering of distributions, J. Appl. Probab. 19 (1982) 310–320.
[11] M. Shaked, J.G. Shanthikumar, Stochastic Orders and their Applications, Academic Press, New York, 1994.
[12] D. Stoyan, Comparison Methods for Queues and Other Stochastic Models, Wiley, New York, 1983.