Copula-based Orderings of Multivariate Dependence

Koen Decancq*

March 30, 2010

Abstract

In this paper I investigate the problem of defining a multivariate dependence ordering. First, I provide a characterization of the concordance dependence ordering between multivariate random vectors with fixed margins. Central to the characterization is a multivariate generalization of a well-known bivariate elementary dependence increasing rearrangement. Second, to order multivariate random vectors with non-fixed margins, I impose a scale invariance principle which leads to a copula-based concordance dependence ordering. Finally, a wide family of copula-based measures of dependence is characterized to which Spearman’s rank correlation coefficient belongs.

Keywords: copula, concordance ordering, dependence measures, dependence orderings, multivariate stochastic dominance, supermodular ordering.

JEL Classification: C14

---

*Koen Decancq: Center for Economic Studies, Katholieke Universiteit Leuven, Naamsestraat 69, B-3000 Leuven, Belgium and CORE, Université catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium. E-mail: koen.decancq@econ.kuleuven.be.

This is a revised version of a chapter of my PhD dissertation. I gratefully thank André Decoster, James E. Foster, Luc Lauwers, Erwin Ooghe, Casey Quinn, Erik Schokkaert, Suman Seth, Dirk Van de Gaer, Tom Van Oorti, John A. Weymark and participants at seminars in Leuven, Berlin, Montreal, Verona, Oxford, and Louvain-La-Neuve for very helpful comments and suggestions to this or an earlier version of the paper. Remaining errors are mine.
1 Introduction

A random vector is said to be dependent when its random variables tend to be large or small simultaneously. Starting from this intuitive, but vague notion of dependence, I follow an axiomatic procedure to provide a more precise answer to the question “when is a multivariate random vector $Y$ more dependent than random vector $X$?”

In a wide range of economic problems the notion of multivariate dependence is of particular interest. In financial and actuarial analysis, the dependence between asset returns or claims in a portfolio provides a measure of the degree of its diversification and systemic risk. In welfare economics, the notion of dependence between income distributions at different points in time is inversely related to the (exchange) income mobility, and the dependence between pre- and post-tax income distributions measures horizontal equity (Dardanoni and Lambert 2001). Atkinson and Bourguignon (1982) and Tsui (1999) have argued that an ordering of multidimensional inequality should be dependence-sensitive. Furthermore, the notion of dependence can be used to measure assortativeness in (many-sided) matching problems, the homogeneity of skills in teams and agreement among judges. Hennessy and Lapan (2003) and Meyer and Strulovici (2009) provide an extensive overview of the use of dependence orderings in economic problems.

The question of how to define an attractive dependence ordering has received earlier attention from statisticians and economists. This paper provides a unifying framework in which the earlier work from both disciplines fits naturally. The central role in this framework is played by the elementary rearrangement that leads to a more dependent random vector. In the existing literature, the focus has been almost exclusively on partial dependence orderings on the set of random vectors with fixed marginal distribution functions, which are obtained from bivariate elementary rearrangements. This paper generalizes beyond this particular case in the following three directions.

First, the paper proposes a multivariate generalization of the bivariate elementary rearrangement that has been used to study dependence (or correlation) (Hamada 1974, Epstein and Tanny 1980, Tchen 1980, Atkinson and Bourguignon 1982). The bivariate rearrangement shifts probability mass on the vertices of a rectangle, such that the probability mass on the realizations with both high or both low values is increased and the probability mass on the realizations with one high and one low value is decreased. I show

---

1 In this research question the reader may hear an echo of the seminal paper of Rothschild and Stiglitz (1970), who study the question “when is a random variable $Y$ more variable than random variable $X$?”

2 See, respectively, Campbell et al. (1997) and Denuit et al. (2005) and the references therein.

3 Fields and Ok (1999) provide a survey of the literature on mobility measurement; see in particular D’Agostino and Dardanoni (2009).
that the geometric generalization of this elementary rearrangement which shifts probability mass on the vertices of a hyperbox (rather than a rectangle) leads to a random vector which is more dependent according to the multivariate concordance dependence ordering proposed by Joe (1990). Furthermore, I show that all expected utility maximizers with an even-increasing utility function prefer a random vector that is obtained by a sequence of multivariate dependence-increasing rearrangements to the original random vector. The class of the even-increasing utility functions is a subclass of the class of the 2-increasing (or supermodular or L-superadditive) utility functions.\footnote{The reader finds a detailed treatment on supermodularity by Topkis (1998) and one by Marshall and Olkin (1979) on L-superadditivity.}

Second, the paper proposes an approach to extend the results beyond the case of multivariate random vectors with \textit{fixed} marginal distributions. In many economic applications one is interested in ordering random vectors with respect to their dependence even if the marginal distributions are not fixed. I proceed by imposing a scale invariance principle, which requires that all increasing transformations of the random variables leave the dependence between the random vector unchanged. This principle makes dependence a purely ordinal concept and allows the definition of equivalence classes with respect to the dependence ordering. Imposing this invariance principle, together with sensitivity to the multivariate dependence-increasing rearrangements, leads to a testable criterion in terms of the underlying copula functions of the random vectors. The copula function has recently become increasingly popular as a practical and useful tool to decompose the information in the joint distribution function in information about the distributional profile \textit{within} each dimension and the dependence structure \textit{between} the dimensions. I believe that the result of this paper provides further insights in the meaning and use of the copula function in general.

Third, a wide family of multivariate dependence measures which are consistent with the \textit{partial} copula-based dependence orderings are characterized by imposing two additional requirements to assure its additive representability. The family of measures derived from the “concordance function”, as introduced by Nelsen (2002), belongs to this family. A bivariate member of the family is Spearman’s rank correlation coefficient. Recently, the multivariate generalization of this measure and its statistical properties received attention in the statistical literature (Nelsen 2006, Dolati and Úbeda-Flores 2006, Taylor 2007). However, to the best of my knowledge, no axiomatic characterization of the overarching family of dependence measures has been provided yet.

The paper is structured as follows. Section two introduces notation. Section three surveys the existing literature on dependence orderings for bivariate random vectors with
fixed marginal distribution functions. In section four the bivariate rearrangements are
generalized into the multivariate setting and I investigate the multivariate dependence
orderings which are consistent with these rearrangements. Section five introduces the
copula function to define copula-based dependence orderings that allow the ordering
of random vectors when the marginal distributions are not fixed. Section six looks at
a class of additive dependence measures that are order-preserving with respect to the
dependence ordering and investigates the properties of one of its prominent members.
Section seven concludes and surveys some possibilities for further research.

2 Notation

Throughout the paper the following notational conventions are adopted. Let $X = (X_1, \ldots, X_m)$ and $Y = (Y_1, \ldots, Y_m)$ be two different $m$-dimensional random vectors
which have the finite set $S$ as support. The number of dimensions $m$ is assumed to be
fixed with $m \geq 2$. The set of the discrete $m$-dimensional random vectors is denoted
$\mathcal{X}$.\footnote{In principle, I will only consider discrete random vectors with finitely
many realizations. However, the results of this paper can be extended to arbitrary random vectors by limiting
arguments as developed by Rothschild and Stiglitz (1970, lemma 2).} The corresponding joint distribution function of $X$ is $F_X$ so that $F_X(x_1, \ldots, x_m) = \Pr [X_1 \leq x_1 \text{ and } \ldots \text{ and } X_m \leq x_m]$ and the joint survival function $\overline{F}_X$ is defined by
$\overline{F}_X(x_1, \ldots, x_m) = \Pr [X_1 > x_1 \text{ and } \ldots \text{ and } X_m > x_m]$. The univariate marginal distribution functions of $F_X$ are denoted by $F_{\cdot 1}, \ldots, F_{\cdot m}$.

I investigate a binary relation $\prec$ on the set of the discrete random vectors $\mathcal{X}$ which is
assumed to be antisymmetric and transitive, hence it is a weak ordering in the sense
of Fishburn (1970). The binary relation $\prec$ should be interpreted as meaning ‘is less
dependent than’ and will be referred to as the ‘dependence ordering’ in the following. The
equivalence dependence ordering is denoted $\sim$.

Following Joe (1997), the subset of the set of discrete random vectors $\mathcal{X}$ with given mar-
ginal distribution functions $F_1, \ldots, F_m$ is referred to as the Fréchet class $\mathcal{F}(F_1, \ldots, F_m)$
or by its shorthand $\mathcal{F}$ when there is no confusion about the underlying marginal distribution functions. Fréchet classes are a natural environment to study dependence, since
their elements only differ in their dependence structure. Yet, an important goal of this
paper is to extend the results so that random vectors from different Fréchet classes can
also be compared with respect to their dependence.

In the following, the class of $k$-increasing functions will play an important role. To intro-
duce such a function, let $x_j < \bar{x}_j$ for all dimensions $j = 1, \ldots, m$ and define for an arbi-
trary function $U$ the first difference operator of the $j$th component $\Delta^j_k \Delta^j_k U(x_1, \ldots, x_m) = U(x_1, \ldots, x_j, \ldots, x_m) - U(x_1, \ldots, x_j, \ldots, x_m)$. A function $U : \mathbb{R}^m \to \mathbb{R}$ is said to be $k$-increasing if it holds that

$$\Delta^j_k \Delta^j_k U(x_1, \ldots, x_m) > 0,$$

for all $(x_1, \ldots, x_m) \in \mathbb{R}^m$, and $1 \leq j_1, \ldots, j_k \leq m$. A function $U$ is said to be $k$-decreasing if and only if $-U$ is $k$-increasing. For any function $U$ which is sufficiently differentiable and $k$-increasing, it holds that $\partial^k U(x)/\partial x_{j_1} \ldots \partial x_{j_k} > 0$. A function which is 2-increasing is also called supermodular (Topkis 1998) or L-superadditive (Marshall and Olkin 1979).

Finally, I define two additional sets, the join-semilattice and meet-semilattice of the set of realizations $S$. Consider $S^n = \{x^1, \ldots, x^n\}$ a subset of $S$ which consists of $n$ realizations. The join and meet operation are defined in the standard way. The join (component-wise maximum) of $S^n$ equals $(\max \{x^1_1, \ldots, x^n_1\}, \ldots, \max \{x^1_m, \ldots, x^n_m\})$ and the meet (component-wise minimum) equals $(\min \{x^1_1, \ldots, x^n_1\}, \ldots, \min \{x^1_m, \ldots, x^n_m\})$. The join-semilattice $J(S)$ of $S$ is the smallest set containing $S$ that is closed under the join operation. It consists of the joins of all subsets of $S$. The meet semi-lattice $M(S)$ of $S$ is the smallest set containing $S$ that is closed under the meet operation. $M(S)$ consists of the meets of all subsets of $S$.

### 3 Bivariate dependence orderings on a Fréchet class

In this section I survey the existing literature on bivariate dependence orderings on a Fréchet class. In a bivariate setting, it is intuitive to say that a random vector is more dependent when its realizations are simultaneously high or low together in both dimensions than when every realization is high in one dimension and low in the other (Hamada 1974). To formalize this intuition, it is useful to define first a positive 2-rearrangement.

**Definition 1.** Let $X$ and $Y$ be in $X$ with support $S$. Consider a rectangle $B_2 = [\underline{x}_{j_1}, \overline{x}_{j_1}] \times [\underline{x}_{j_2}, \overline{x}_{j_2}]$ whose vertices are in $S$, with $\overline{x}_j < \overline{x}_j$ for all $j \in \{j_1, j_2\}$. If $Y$ can be obtained from $X$ by adding a positive probability mass $\varepsilon$ to all vertices of the rectangle $B_2$ with an even⁶ number of components $x_j = \overline{x}_j$ and subtracting $\varepsilon$ from all vertices of the rectangle $B_2$ with an odd number of components $x_j = \underline{x}_j$, then $Y$ is obtained from $X$ by a positive 2-rearrangement.

---

⁶ In the following, I adopt the convention that 0 is even.
The probability mass \( \varepsilon > 0 \) is assumed to be small enough so that the remaining probability mass is non-negative everywhere. The positive 2-rearrangement is also known as a ‘correlation increasing transformation’ (Epstein and Tanny 1980) or a ‘basic rearrangement’ (Boland and Proschan 1988). In the literature on mobility measurement it plays an important role (Atkinson 1981, Dardanoni 1993, D’Agostino and Dardanoni 2009). A negative 2-rearrangement is defined likewise as rearrangement of a negative probability mass \( \varepsilon \). Observe that a 2-rearrangement affects only two of the \( m \) dimensions.

Consider random vectors \( Y \) and \( X \), such that \( Y \) can be obtained from \( X \) by a positive 2-rearrangement. Since any 2-rearrangement leaves the marginal distributions unchanged, \( X \) and \( Y \) belong to the same Fréchet class. A 2-rearrangement only affects the dependence structure between the random variables. More precisely, the following principle requires the dependence ordering to increase after a positive 2-rearrangement.

2-Dependence Principle (2DEP). Let \( X \) and \( Y \) be in \( X \). If \( Y \) can be obtained from \( X \) by a finite sequence of positive 2-rearrangements, then \( X \prec Y \).

Epstein and Tanny (1980) and Tchen (1980) have independently shown the following equivalences.

**Proposition 1.** Let \( X \) and \( Y \) be in \( F \) with support \( S \) and suppose \( m = 2 \). The dependence ordering \( \prec \) on \( F \) satisfies 2DEP if and only if \( X \prec Y \) is equivalent to:

1. \( \int U(x_1, \ldots, x_m) dF_X(x_1, \ldots, x_m) \leq \int U(x_1, \ldots, x_m) dF_Y(x_1, \ldots, x_m) \) for all \( 2 \)-increasing utility functions \( U \),
2. \( F_X(x_1, \ldots, x_m) \leq F_Y(x_1, \ldots, x_m) \) for all \( (x_1, \ldots, x_m) \) in \( J(S) \),
3. \( \overline{F}_X(x_1, \ldots, x_m) \leq \overline{F}_Y(x_1, \ldots, x_m) \) for all \( (x_1, \ldots, x_m) \) in \( M(S) \).

Proof: The equivalence between the dependence ordering satisfying 2DEP and condition 1 is proven by Epstein and Tanny (1980, theorem 6). Epstein and Tanny (1980, theorem 1) show equivalence between condition 1 and \( F_X(x_1, \ldots, x_m) \leq F_Y(x_1, \ldots, x_m) \) for all \( (x_1, \ldots, x_m) \) in \( \mathbb{R}^m \) with \( m = 2 \). Dyckerhoff and Mosler (1997, theorem 1) show that is suffices to check the inequality for all \( (x_1, \ldots, x_m) \) in \( J(S) \). The equivalence between condition 2 and 3 is proven by Epstein and Tanny (1980, theorem 3). Dyckerhoff and Mosler (1997, theorem 3) show that checking for all \( (x_1, \ldots, x_m) \) in \( M(S) \) in this case also suffices.

Note that the equivalence only applies to bivariate random vectors which belong to the same Fréchet class. This result adds three equivalent conditions to the statement that \( Y \)
is more dependent than $X$ according to a dependence ordering $\prec$ which satisfies 2DEP. Each of these conditions provides an alternative perspective on the dependence ordering.

The integral in the first condition can be interpreted as a von Neumann-Morgenstern utility function, so that the condition requires unanimous agreement by all expected utility maximizers with a 2-increasing or supermodular utility function. A utility function is 2-increasing if its arguments are complements in the Auspitz-Lieben-Edgeworth-Pareto (ALEP) sense (Kannai 1980), or in case of a differentiable utility function, if the marginal utility of a dimension is increasing in the level of any other dimension. Richard (1975) refers to a decision maker with a supermodular (and differentiable) utility function as ‘multivariate risk seeking’ and Epstein and Tanny (1980) call such a utility function a ‘correlation-affine’ utility function. In many applications it may seem more natural to consider 2-decreasing (or multivariate risk averse) utility functions instead, in which case the direction of the inequality in condition one reverses. In the statistical literature, the dependence ordering that satisfies the first condition is known as the supermodular dependence ordering (see, for instance, Joe (1990), Müller and Scarsini (2000) and Denuit et al. (2005)).

The second condition is a test of bivariate first-order stochastic dominance as studied by Hadar and Russel (1974) and Atkinson and Bourguignon (1982) who show equivalence between condition one and two under the additional assumption of differentiability of the utility function. The second condition defines the lower orthant dependence ordering in the statistical literature. Condition three presents its mirror-image and defines the upper orthant dependence ordering. Note that the sets $J(S)$ and $M(S)$ are finite sets, so that condition two and three are testable in practice. If we deal with a finite sample from an underlying population, many procedures are developed to test these inequalities in a statistical significant way (see, for instance, Dardanoni and Forcina (1998), Davidson and Duclos (2000) and Cebrián et al. (2004)).

4 Multivariate dependence orderings on a Fréchet class

To study multivariate dependence orderings, I investigate in a first step the multivariate dependence ordering which satisfies the 2-rearrangement principle between any pair of dimensions. This dependence ordering seems a natural extension of the bivariate dependence ordering. It can be shown to be equivalent to the multivariate supermodular

---

7 The second condition offers an additional interpretation of the dependence ordering as an ordering of ‘more vector dominance’. For any possible realization $x$ in $S$, a randomly chosen realization from $Y$ is more likely to be vector dominated by $x$ than a randomly chosen realization from $X$. 

---
dependence ordering which has been studied by Müller and Scarsini (2000) among others (see Denuit et al. (2005, chapter 6) for a survey).

**Proposition 2.** Let $X$ and $Y$ be in $\mathcal{F}$ with support $S$ and suppose $m \geq 2$. The dependence ordering $\prec$ on $\mathcal{F}$ satisfies 2DEP if and only if $X \prec Y$ is equivalent to

1. $\int U(x_1, \ldots, x_m) dF_X(x_1, \ldots, x_m) \leq \int U(x_1, \ldots, x_m) dF_Y(x_1, \ldots, x_m)$
   
   for all 2-increasing utility functions $U$.

Proof: The $\Rightarrow$ part of the proof follows straight from the definition of a 2-increasing function. Boland and Proshan (1988) have proven the complete proposition in their proposition 2.5 (b). Recently, Meyer and Strulovici (2009, theorem 1) have provided a shorter proof using arguments from lattice theory.

It is well known that the equivalences between the supermodular ordering and the orthant dependence orderings no longer hold once $m > 2$. (For counterexamples, see Joe (1990) in case $m \geq 4$ and Müller and Scarsini (2000) for $m = 3$).

The supermodular dependence ordering has two drawbacks. First, from a practical perspective, it is difficult to test since it involves a procedure which tests all possible supermodular utility functions.\(^8\) Second and more substantially, there seems to be no a priori reason to confine our attention in a multivariate setting to rearrangements which affect only two of the $m$ dimensions. An interesting geometric multivariate generalization of the 2-rearrangements seems therefore to consider the rearrangement of probability mass on the vertices of a $k$-dimensional hyperbox instead of considering only the vertices of a 2-dimensional rectangle.

**Definition 2.** Let $X$ and $Y$ be in $\mathcal{X}$ with support $S$ and suppose $m \geq k \geq 2$. Consider a hyperbox $B_k = [x_{j_1}, \bar{x}_{j_1}] \times \cdots \times [x_{j_k}, \bar{x}_{j_k}]$ whose vertices are in $S$, with $x_{j_1} < \bar{x}_{j_1}$ for all $j \in \{j_1, \ldots, j_k\}$. If $Y$ can be obtained from $X$ by adding positive probability mass $\varepsilon$ to all vertices of the rectangle $B_k$ with an even number of components $x_j = \bar{x}_j$ and subtracting $\varepsilon$ from all vertices of the rectangle $B_k$ with an odd number of components $x_j = \bar{x}_j$, then $Y$ is obtained from $X$ by a positive $k$-rearrangement.

To the best of my knowledge, the properties of this type of rearrangements have never been studied before. Observe that a $k$-rearrangement affects $k$ of the $m$ dimensions. For $k = 2$, the definition of a positive 2-rearrangement is obtained, so that indeed definition 2 generalizes definition 1.\(^9\)

---

\(^8\) See, however, Athey (2000) who shows that class of supermodular functions may be constructed from a finite 'test function set'.

\(^9\) Note that any positive $k$-rearrangement can also be defined recursively as a combination of two
Depending on whether $k$ is even or odd two different cases can be distinguished. For notational simplicity, I therefore refer to an even-rearrangement, when $k$ is an even number, and otherwise to an odd-rearrangement. From the definition it follows that a positive even-rearrangement shifts probability mass to the vertex with all high values $(x_j, \ldots, x_k)$ as well to the vertex with all low values $(x_{\tilde{j}1}, \ldots, x_{\tilde{j}k})$. According to the intuitive starting point of the paper, dependence increases. Things become more intricate for a positive odd-rearrangement: probability mass is shifted to the vertex with all high values, but it is shifted away from the vertex with all low values. A priori it seems hard to say whether dependence increased or decreased after such-rearrangement. Therefore, two different multivariate dependence principles can be imposed to the dependence ordering. Both principles agree that a positive even-rearrangement increases dependence, but they reflect a different position when $k$ is odd. The first principle requires the dependence ordering to be increasing after a positive odd-rearrangement and the second (denoted with an asterisk), requires the dependence ordering to be decreasing after a positive odd-rearrangement.

$k$-Dependence Principle ($k$DEP). Let $X$ and $Y$ be in $\mathcal{X}$ and suppose $m \geq k \geq 2$. If $Y$ can be obtained from $X$ by a finite sequence of positive $k$-rearrangements, then $X \prec Y$.

$k$-Dependence Principle* ($k$DEP*). Let $X$ and $Y$ be in $\mathcal{X}$ and suppose $m \geq k \geq 2$. If $Y$ can be obtained from $X$ by a finite sequence of positive even-rearrangements or an elementary negative odd-rearrangements, then $X \prec Y$.

The next proposition shows that imposing both principles leads to different equivalence results.

**Proposition 3.** Let $X$ and $Y$ be in $\mathcal{F}$ with support $S$ and suppose $m \geq 2$.

The dependence ordering $\prec$ on $\mathcal{F}$ satisfies $k$DEP if and only if $X \prec Y$ is equivalent to:

1. $\int U(x_1, \ldots, x_m) dF_X(x_1, \ldots, x_m) \leq \int U(x_1, \ldots, x_m) dF_Y(x_1, \ldots, x_m)$, for all $k$--increasing utility functions $U$,
2. $F_X(x_1, \ldots, x_m) \leq F_Y(x_1, \ldots, x_m)$ for all $(x_1, \ldots, x_m)$ in $M(S)$.

The dependence ordering $\prec$ on $\mathcal{F}$ satisfies $k$DEP* if and only if $X \prec Y$ is equivalent to:

$(k - 1)$--rearrangements. In particular, for any $k$ with $m \geq k > 2$, a positive $k$-rearrangement is a composite rearrangement of a negative and positive $(k - 1)$--rearrangement on the vertices of two hyperboxes $B_{k-1}^{\tilde{j}_k} = [x_{\tilde{j}_1}, x_{\tilde{j}_k}] \times \cdots \times [x_{\tilde{j}_{k-2}}, x_{\tilde{j}_{k-1}}]$ and $B_{k-1}^{\tilde{j}_k} = [x'_{\tilde{j}_1}, x'_{\tilde{j}_k}] \times \cdots \times [x'_{\tilde{j}_{k-2}}, x'_{\tilde{j}_{k-1}}]$, which coincide for all entries except for the $j_k$--th, so that $x_{\tilde{j}_k} < x'_{\tilde{j}_k}$.

From this alternative recursive definition, it is clear that there is no positive $k$-rearrangement which can be obtained as a sequence of only positive $(k - 1)$-rearrangements.
1. \( \int U(x_1, \ldots, x_m) dF_X(x_1, \ldots, x_m) \leq \int U(x_1, \ldots, x_m) dF_Y(x_1, \ldots, x_m) \)
   for all even-increasing and odd-decreasing utility functions \( U \),

2. \( F_X(x_1, \ldots, x_m) \leq F_Y(x_1, \ldots, x_m) \) for all \( (x_1, \ldots, x_m) \) in \( J(S) \).

Proof: The proof of both parts of the proposition is very similar, so here I only focus on the first part involving \( k \text{DEP} \) only. The equivalence between the statements is proven by a cycle of implications. The proof proceeds in three steps.

**Step 1. If the dependence ordering \( \prec \) satisfies \( k \text{DEP} \), then condition 1 holds.**

If \( Y \) is obtained from \( X \) by a positive \( k \)-rearrangement, we can write

\[
\int U(x_1, \ldots, x_m) d[F_X(x_1, \ldots, x_m) - F_Y(x_1, \ldots, x_m)] = \\
\varepsilon \left[ \Delta_{x_{i_1}^j} \cdots \Delta_{x_{i_k}^j} U(x_1, \ldots, x_m) \right].
\]

Since \( \varepsilon > 0 \), \( \int U(x_1, \ldots, x_m) d[F_X(x_1, \ldots, x_m) - F_Y(x_1, \ldots, x_m)] \geq 0 \) if and only if \( \Delta_{x_{i_1}^j} \cdots \Delta_{x_{i_k}^j} U(x_1, \ldots, x_m) \geq 0 \).

**Step 2. If condition 1 holds, then condition 2 holds.**

Rüschendorf (1980) proves the bi-implication between condition 1 and 2 for \( k \)-increasing functions defined on the unit hypercube. The implication needed here can easily be proven by selecting an appropriate function. The indicator function of the upper orthant of the point \( (x_1, \ldots, x_m) \), denoted \( T(z_1 > x_1, \ldots, z_m > x_m) \), returns 1 if the logical condition \( (z_1 > x_1, \ldots, z_m > x_m) \) is fulfilled and 0 otherwise. For all \( (x_1, \ldots, x_m) \) in \( \mathbb{R}^m \) this indicator function is a \( k \)-increasing function (for all \( m \geq k \geq 2 \)), so that it holds that

\[
\int T(z_1 > x_1, \ldots, z_m > x_m) dF_X(z_1, \ldots, z_m) \leq \int T(z_1 > x_1, \ldots, z_m > x_m) dF_Y(z_1, \ldots, z_m),
\]

and hence also that

\[
F_X(x_1, \ldots, x_m) \leq F_Y(x_1, \ldots, x_m) \text{ for all } (x_1, \ldots, x_m) \text{ in } M(S).
\]

**Step 3. If condition 2 holds, then \( \prec \) satisfies \( k \text{DEP} \).**

I provide a constructive proof. The joint survival functions \( F_X \) and \( F_Y \) are discrete survival functions, so that \( \overline{R} = F_Y - F_X \) is a multivariate step-function. It can be written as

\[
\overline{R}(x_1, \ldots, x_m) = \sum_{i=1}^{n} \overline{R}_i(x_1, \ldots, x_m),
\]
whereby every $\overline{R}^i$ is a function which returns a constant value $\varepsilon^i$ on the interior of a hyperbox $(B_k)^i$ and 0 otherwise. These functions $\overline{R}^i$ are obtained from a positive $k$-rearrangement as follows. A positive $k$-rearrangement $i$ on the vertices of hyperbox $(B_k)^i$ can be summarized by its joint rearrangement function $\overline{R}^i : S \rightarrow \{0, \varepsilon^i\}$ which is the integral of the rearranged probability mass $r^i$ for every point from the support $S$:

$$\overline{R}^i(x_1, \ldots, x_m) = \int_x \varepsilon^i(z_1, \ldots, z_m) d(z_1, \ldots, z_m),$$

so that $\overline{R}^i(x_1, \ldots, x_m) = \varepsilon^i$ if $x$ is in the interior of $(B_k)^i$ and 0 otherwise. Hence $F_Y = F_X + n \sum_{i=1}^n \overline{R}^i(x_1, \ldots, x_m)$ so that $Y$ can be reached from $X$ by at most $n$ positive $k$-rearrangements. □

The respective second conditions of the above equivalences are commonly used to define the multivariate upper orthant and lower orthant dependence ordering. The above proposition connects them to the expected utility literature. Under the additional assumption of differentiability, the equivalence between condition 1 and 2 of the second part of the proposition goes back to the early work on multivariate stochastic dominance by Hadar and Russel (1974). The authors show equivalence between multivariate first-order stochastic dominance and unanimity within the class of utility functions with alternating signs of their cross derivatives. Scarsini (1988) calls such a utility function an $m$-variate risk averse utility function, thereby generalizing the earlier work of Richard (1975). Furthermore, the proposition identifies the underlying multivariate elementary rearrangements of both orthant dependence orderings. A priori there seems no reason to prioritize $kDEP^*$ to $kDEP$, yet most economic results on multivariate stochastic dominance focus exclusively on the results obtained by imposing $kDEP^*$. Given the hard choice between both principles, a natural approach seems to impose only what both principles agree about, being that positive even-rearrangements increase dependence. This leads to the following principle.

**Even-Dependence Principle (eDEP).** Let $X$ and $Y$ be in $\mathcal{X}$ and suppose $m \geq k \geq 2$. If $Y$ can be obtained from $X$ by a finite sequence of positive even-rearrangements, then $X \prec Y$.

Imposing the even-dependence principle leads to the following result.

**Corollary 1.** Let $X$ and $Y$ be in $\mathcal{F}$ with support $S$ and suppose $m \geq 2$. The dependence ordering $\prec$ on $\mathcal{F}$ satisfies eDEP if and only if $X \prec Y$ is equivalent to:

1. $\int U(x_1, \ldots, x_m) dF_X(x_1, \ldots, x_m) \leq \int U(x_1, \ldots, x_m) dF_Y(x_1, \ldots, x_m)$, for all even-increasing utility functions $U$,
2. \( F_X(x_1, \ldots, x_m) \leq F_Y(x_1, \ldots, x_m) \) and \( \bar{F}_X(x_1, \ldots, x_m) \leq \bar{F}_Y(x_1, \ldots, x_m) \) for all \((x_1, \ldots, x_m)\) in \( M(S) \cup J(S) \).

Proof: This result follows from the previous one and its proof is omitted.

The second part of the result shows that a dependence ordering which satisfies \( eDEP \) is equivalent to a double test of multivariate stochastic dominance in terms of the joint distribution function and the joint survival function. Joe (1990) uses the double test to define the so-called concordance dependence ordering.

In comparison to the supermodular dependence ordering, the concordance dependence ordering has the practical advantage that testing for it can be done by looking at a finite number of inequalities. It is clear however, that in terms of underlying rearrangements, the supermodular dependence ordering and the concordance dependence ordering are very different. The supermodular dependence ordering focuses on elementary rearrangements which involve only a pair of dimensions (or four realizations). This restriction seems arbitrary. The concordance dependence ordering builds on a wider class of rearrangements, but this might be less appealing to one’s intuitions since these rearrangements involve \( 2^k \) realizations at the same time. In terms of the class of utility functions involved, the conditions on the even cross-derivatives seem rather demanding and may be difficult to interpret economically.

From the logical implications between the dependence principles, it follows that the concordance dependence ordering implies both orthant dependence orderings and is itself implied by the supermodular dependence ordering once \( m > 2 \) (see also Joe (1990), Müller and Scarsini (2000)).

5 Copula-based multivariate dependence orderings

In this section I investigate how the dependence orderings can be extended to order random vectors that belong to different Fréchet classes. I proceed by imposing an additional principle to the dependence ordering that requires that the dependence of a random vector remains invariant when its random variables are transformed by a strictly increasing transformation function.

**Scale Invariance Principle (INV).** Let \( X \) and \( Y \) be in \( \mathcal{X} \) and let \( T_1, \ldots, T_m \) be strictly increasing transformation functions. If \( Y = (T_1(X_1), \ldots, T_m(X_m)) \), then \( X \sim Y \).
This principle goes back to the early work of Hoeffding (1940, 1941) who has studied scale invariant measures of dependence (for more recent references, see Scarsini (1984) and Embrechts et al. (2000)). The principle defines an equivalence relation on the set of random vectors and partitions it in various equivalence classes. All random vectors belonging to the same equivalence class are judged to be equally dependent. Scale invariance is a strong principle and imposing it makes the notion of dependence a purely ordinal one. Schweizer and Wolff (1981) note: “... it is precisely the copula which captures those properties of the joint distribution which are invariant under strictly increasing transformations”. Such a copula function can be defined as follows (Schweizer and Sklar 1983, Nelsen 2006).

**Definition 3.** A copula function is any function $C : [0,1]^m \to [0,1]$ which has the following properties:

1. $C(p_1, \ldots, p_m) = 0$ if at least one of its components $p_j = 0$.
2. $C(1, \ldots, 1, p_j, 1, \ldots, 1) = p_j$ for all $j = 1, \ldots, m$.
3. $C$ is a weakly $m$-increasing function.

Alternatively, a copula function $C$ can be defined as a joint distribution function of a random vector with standard uniform marginal distribution functions (see Embrechts et al. (2000) and Nelsen (2006) for the equivalence between both definitions). Similarly, $C$ can be defined as a joint survival function with standard uniform marginal distribution functions.

In an important theorem, Sklar (1959) has shown that each joint distribution function $F_X$ can be written as a function of its $m$ one-dimensional marginal distribution functions $F_j$ and a copula function $C_X$. For all $x$ in $\mathbb{R}^m$, we have:

$$F_X(x_1, \ldots, x_m) = C_X (F_1(x_1), \ldots, F_m(x_m)).$$

If the marginal distribution functions are all continuous, $C_X$ is uniquely defined. Otherwise, $C_X$ is defined uniquely on $I = \text{Range}(F_1) \times \cdots \times \text{Range}(F_m)$ (Nelsen 2006). When it is uniquely defined, we refer to $C_X$ as the copula function of random vector $X$. Sklar’s result permits a decomposition of the information in the joint distribution function $F_X$ in information about the distributional profiles within each dimension (captured by the marginal distributions $F_1, \ldots, F_m$) and information about the dependence structure between the dimensions (captured by the copula $C_X$). Two important copulas are the independence copula $C_\perp(p_1, \ldots, p_m) = p_1 \cdots p_m$ and the comonotonic copula $C_+(p_1, \ldots, p_m) = \min \{p_1, \ldots, p_m\}$, which reflect respectively the dependence structure in case of no dependence and maximal dependence.
For the purpose of this paper, we are interested in ordering the dependence of random vectors belonging to different Fréchet classes. We are, in other words, interested in the dependence structure of a random vector, irrespective of its marginal distributions. An analysis of dependence based on its copula function, rather than on its joint distribution function, seems therefore a natural approach. The following proposition shows that imposing INV in addition to eDEP leads exactly to such a copula-based concordance dependence ordering.

**Proposition 4.** Let $X$ and $Y$ be in $\mathcal{X}$ with support $S$ and suppose $m \geq 2$. The dependence ordering $\prec$ on $\mathcal{X}$ with support $S$ satisfies eDEP and INV if and only if $X \prec Y$ is equivalent to:

1. $\int U(p_1, \ldots, p_m) dC_X(p_1, \ldots, p_m) \leq \int U(p_1, \ldots, p_m) dC_Y(p_1, \ldots, p_m)$, for all even-increasing utility functions $U$,

2. $C_X(p_1, \ldots, p_m) \leq C_Y(p_1, \ldots, p_m)$ and $\overline{C}_X(p_1, \ldots, p_m) \leq \overline{C}_Y(p_1, \ldots, p_m)$ for all $(p_1, \ldots, p_m)$ in $M(I) \cup J(I)$.

Proof: The proof proceeds in three steps.

**Step 1.** If the dependence ordering $\prec$ satisfies eDEP and INV, then condition 1 holds. Let $X$ and $Y$ have different marginal distribution functions $F_1, \ldots, F_m$ and $F'_1, \ldots, F'_m$ respectively. Define

$$P = (F_1(X_1), \ldots, F_m(X_m)) \text{ and } P' = (F'_1(Y_1), \ldots, F'_m(Y_m)),$$

so that by INV: $P \prec P' \Leftrightarrow X \prec Y$. The random vectors $P$ and $P'$ have standard uniformly distributed (discrete) random variables (by the probability integral transform) and hence they belong to the same Fréchet class. Corollary 1 applies, so that by eDEP its condition 1 holds. Observe that $F_P$ and $F_{P'}$, the joint distribution functions of $P$ and $P'$, can be written as $C_X$ and $C_Y$ (by the definition of the copula function).

**Step 2.** If condition 1 holds, then condition 2 holds. This implication follows from the equivalence between condition 1 and 2 of corollary 1 which applies to all joint distribution functions/joint survival functions with the same marginal distribution functions, hence also to joint distribution functions/joint survival functions which have standard uniform marginal distribution functions (i.e. to copulas).

**Step 3.** If condition 2 holds, then $\prec$ satisfies eDEP and INV.
The dependence ordering $\prec$ satisfies $e\text{DEP}$ by corollary 1. The fact that $\prec$ satisfies $\text{INV}$ follows from the scale invariance property of the copula function (Schweizer and Wolff 1981, theorem 2).

Alternatively, the reader may want to impose 2DEP or one of the $k$-dimensional dependence principles ($k\text{DEP}$ or $k\text{DEP}^*$). The results are in line with the investigations in the previous section and are not reported for reasons of brevity.

The second condition of the Proposition, the double test of multivariate stochastic dominance in terms of the copula function $C$ and joint survival function $\overline{C}$, has recently been proposed by Nelsen (2002, 2006) as an alternative definition of the concordance dependence ordering. The powerful idea of ordering random vectors from different Fréchet classes by investigating stochastic dominance in terms of their underlying copula functions, goes back to Scarsini (1984). The above Proposition crystallizes the implicit assumptions imposed on the dependence ordering to do so (being $e\text{DEP}$ and $\text{INV}$). Furthermore, the first condition provides a wide class of functions which are order-preserving of the dependence ordering, an observation which will become useful in the next section.

6 Measures of multivariate dependence

The dependence orderings defined in the preceding sections are incomplete. Some pairs of random vectors can be ordered, but for some other pairs the dependence ordering is ‘indecisive’. Although the indecisiveness might be informative in its own right, in many situations one may prefer a measure of multivariate dependence which leads to a complete dependence ordering so that all pairs of random vectors can be ordered. To obtain a family of measures with an attractive and tractable general structure, I impose the following two principles to the dependence ordering. The first adds a flavor of continuity and imposes that if $X \prec Y \prec Z$ holds, then there exists at least one linear convex combination of $X$ and $Z$ that is ordered on either side of $Y$.

**Continuity Principle (CONT).** For all $X, Y$ and $Z$ in $\mathcal{X}$, if $X \prec Z \prec Y$, then there exist an $\alpha$ and $\beta$ in $(0,1)$, such that $\alpha X + (1 - \alpha) Y \prec Z$ and $Z \prec \beta X + (1 - \beta) Y$.

The continuity principle is also known as the ‘Archimedean Axiom’. The second principle

\footnote{Dardanoni and Lambert (2001) use a similar copula-based partial dependence ordering to compare pre and post tax income distributions with respect to their horizontal inequity. The rank-based mobility orderings proposed by D’Agostino and Dardanoni (2009) are implicitly copula-based and Decancq (2009) provides another application, by looking at dependence orderings between the dimensions of well-being in contemporaneous Russia.}
imposes that the dependence ordering of two random vectors $\prec$ is consistent with the dependence ordering of the random vectors after a linear convex combination is taken with a common random vector $Z$.

**Independence Principle (IND).** For all $X, Y$ and $Z$ in $\mathcal{X}$, if $X \prec Y$ and $\lambda$ in $(0, 1)$, then $\lambda X + (1 - \lambda) Z \prec \lambda Y + (1 - \lambda) Z$.\(^{11}\)

These two principles together imply additive representability. Furthermore, the previously introduced multivariate dependence principle ($eDEP$) and the scale invariance principle (INV) will be imposed to come to an additive representation of the copula-based dependence ordering.

**Proposition 5.** Let $X$ and $Y$ be in $\mathcal{X}$ and suppose $m \geq 2$. The dependence ordering $\prec$ on $\mathcal{X}$ satisfies CONT, IND, $eDEP$ and INV if and only if there exist a real-valued even-increasing utility function $U$ such that $X \prec Y \iff D_m(X) < D_m(Y)$ with

$$D_m(X) = \int U(p_1, \ldots, p_m) dC_X(p_1, \ldots, p_m). \quad (1)$$

Proof: It can be verified that the proposed function $D_m$ satisfies all principles. I only prove necessity. The dependence ordering is a weak order in the sense of Fishburn (1970). According to theorem 8.2 of Fishburn (1970), any weak order on the set of discrete random vectors $\mathcal{X}$ which satisfies CONT and IND can be represented by $D_m(X) = \int U(x_1, \ldots, x_m) dF_X(x_1, \ldots, x_m)$ with $U$ a real-valued function.

By Proposition 4 it follows that imposing $eDEP$ restricts the real-valued function $U$ to be even-increasing and that imposing INV leads to a copula-based representation $D_m(X) = \int U(p_1, \ldots, p_m) dC_X(p_1, \ldots, p_m)$. \(\blacksquare\)

Any of the functions $D_m$ is order-preserving of the (concordance) dependence ordering, which makes them very suitable as measures of dependence. Therefore I will refer to $D_m$ as a measure of dependence in the remainder of the paper.\(^{12}\) The above proposition determines the functional form of the dependence measure up to an increasing transformation of $D_m$. For practical purposes it is convenient that the dependence of the benchmark random vectors $X_\perp$ (with independent dimensions and copula $C_\perp$) and $X_+$

\(^{11}\)The meaning of ‘independence’ in this principle is different from the meaning of ‘dependence’ used in the rest of the paper, yet, I prefer to stick to the standard terminology for this principle in the literature on utility theory.

\(^{12}\)Observe that this axiomatic characterization of a dependence measure differs from the axiomatic approach followed in mathematical statistics by Scarsini (1984), Joe (1997), Dolati and Ubeda-Flores (2006) and Taylor (2007), who check a series of attractive statistical properties for a given class of dependence measures.
(with comonotonic dimensions and copula $C_+$) is normalized to 0 and 1 respectively. To achieve that, a linear normalization of $D_m$ can be employed. This leads to the normalized dependence measure $\hat{D}_m$:

$$\hat{D}_m(X) = \frac{D_m(X) - D_m(X_\perp)}{D_m(X_+) - D_m(X_\perp)}.$$ 

The normalized dependence measure $\hat{D}_m$ can be interpreted as the relative position between the benchmarks of independence $X_\perp$ and maximal dependence $X_+$.\(^{13}\) Observe that the dependence $\hat{D}_m$ can take negative values as well, for instance in the bivariate countermonotonic case where high values in one dimension tend to come with low values in the other dimension.

Every even-increasing utility function $U$ in Proposition 5 leads to a new dependence measure $\hat{D}_m$. In his pioneering contribution to the copula-based measurement of dependence, Nelsen (1991) focuses on ‘concordance functions’, which are the normalized dependence measures for which the utility function $U$ can be expressed as the sum of a copula function and joint survival function of a reference random vector. By definition all these utility functions are multivariate even-increasing. In particular, the properties of one dependence measure have been studied in detail. This dependence measure is denoted $\hat{D}_m^\perp(X)$ and is obtained by selecting the reference random vector to be one with independent marginal distributions (and hence with copula function $C_\perp$ and joint survival function $C_\perp$), so that (after some algebraic manipulations) one obtains:

$$\hat{D}_m^\perp(X) = \frac{(m + 1) [2^{m-1} \int [C_\perp(p) + C_\perp(p)] dC_X(p) - 1]}{2^m - (m + 1)}.$$ 

Interestingly, the bivariate $\hat{D}_2^\perp$ equals the Spearman rank correlation coefficient $\rho$. The measure $\hat{D}_m^\perp$ is the mean of the two multidimensional rank correlation coefficients which are discussed in Joe (1990), and its properties have been studied by Nelsen (1991, 2002), Dolati and Úbeda-Flores (2006), Schmid and Schmidt (2006, 2007) and Taylor (2007).\(^{14}\)

\(^{13}\) Note that also the normalized dependence measures are order-preserving of the (concordance) dependence ordering, since it holds that $D_m(X_\perp) = D_m(Y_\perp)$ and $D_m(X_+) = D_m(Y_+)$.

\(^{14}\) Some other multivariate generalizations of well-known bivariate dependence measures also belong to this family. If $U$ in expression (1) is set equal to the sum of the copula function and joint survival function of the random vector $X$ itself, the normalized dependence measure $\hat{D}_m$ turns out to be a multidimensional generalization of the Kendall’s bivariate measure of association $\tau$. A multivariate version of Spearman’s footrule coefficient can be derived when selecting the reference random vector equal to the comonotonic random vector $X_\perp$ (Úbeda Flores 2005). A multivariate version of Blomqvist’s $\beta$ is obtained by making use of a utility function $U$ which is the sum of the indicator-functions of the lower and upper orthant of the point $(1/2, \ldots, 1/2)$ (see Úbeda Flores (2005) for details). Behboodian et al. (2007) derive another weakly even-increasing utility function $U$ from the average of the copula of the comonotonic random vector $X_\perp$ and the function describing the countermonotonic random vector $X_-$, which leads to a multivariate generalization of Gini’s measure of association $\gamma$. 

17
Finally, the principle eDEP could be replaced by the more standard 2-rearrangement principle 2DEP in the statement of proposition 5. Following a similar argument, one can show that the complete dependence ordering can be represented by an additive dependence measure based on a 2-increasing utility function, which I denote by $D_{0m}$. Hence, for any of the $m(m-1)/2$ pairs of dimensions $j_1$ and $j_2$ with $1 \leq j_1 < j_2 \leq m$, it holds that $\overline{D}'_2(X_{j_1}, X_{j_2}) \leq \overline{D}'_2(Y_{j_1}, Y_{j_2})$, when $Y$ is obtained from $X$ by a positive 2-rearrangement. The average of the $\overline{D}'_2$'s of all bivariate pairs of dimensions will evaluate $Y$ to be more dependent than $X$, so that

$$\frac{2}{m(m-1)} \sum_{1 \leq j_1 < j_2 \leq m} \overline{D}'_2(X_{j_1}, X_{j_2}) < \frac{2}{m(m-1)} \sum_{1 \leq j_1 < j_2 \leq m} \overline{D}'_2(Y_{j_1}, Y_{j_2}).$$ (2)

This observation provides an axiomatic justification for the approach taken by many practitioners who compute the average of all pair wise bivariate dependence measures (such as Spearman’s rank correlation coefficient) rather than to compute a “genuine” multivariate dependence measure $\overline{D}_m$. It is clear that by construction such an approach can only be sensitive to 2-rearrangements and not to its multivariate generalizations. Again, the choice of the most appropriate rearrangement principle for the dependence ordering drives the choice between a “genuine” multivariate dependence measure $\overline{D}_m$ or an average pair wise bivariate measure as given in expression (2).

7 Conclusion

In this paper the existing results on dependence orderings based on elementary bivariate dependence increasing rearrangements, as studied by Epstein and Tanny (1980) and many others, are extended in three directions.

First, I have argued that in a multivariate setting it is interesting to move beyond a sole focus on rearrangements involving only two dimensions simultaneously. I propose a geometric multivariate generalization of the bivariate dependence increasing rearrangements and show how they are underlying the orthant and concordance dependence orderings. This provides new insights in these orderings. In the statistical literature a plethora of alternative dependence orderings exists (see Joe (1997) and Colangelo et al. (2006) for extensive surveys), and it remains an open question for further research to define the rearrangements that are underlying these alternative dependence orderings.

Second, I showed how a scale invariance principle naturally leads to the copula-based dependence orderings. Copula functions are becoming increasingly popular as a tool to measure and model dependence. Yet, the (implicit) assumption of scale invariance is
a strong one, since it makes the dependence concept a purely ordinal one. An avenue for further research is to consider possible weakenings of this principle (as is standard in Arrovian social choice theory, see d’Aspremont and Gevers (2002) and Bossert and Weymark (2004) for detailed overviews).

Third, I characterized a wide family of dependence measures consistent with the partial dependence orderings. Thereby I addressed an open question raised by Dardanoni and Lambert (2001): “The interesting question, to identify the class of copula-based indices which accord with our [concordance] partial ordering, remains open and is clearly an important one for future research.” Interestingly, this family contains some multivariate generalizations of existing bivariate measures. From a practical perspective, however, the family of dependence measures still remains wide. There seems room for some additional attractive principles to narrow down the family further.

References


