CHARACTERIZATIONS OF CONDITIONAL COMONOTONICITY

KA CHUN CHEUNG,* *University of Calgary*

Abstract

The notion of conditional comonotonicity was first used implicitly by Kaas, Dhaene, and Goovaerts (2000) and was formally introduced by Jouini and Napp (2004) as a generalization of the classical concept of comonotonicity. The objective of the present paper is to further investigate this relatively new concept. The main result is that a random vector is comonotonic conditional to a certain σ-field if and only if it is almost surely comonotonic locally on each atom of the conditioning σ-field. We also provide a new proof of a distributional representation and an almost sure representation of a conditionally comonotonic random vector.

Keywords: Comonotonicity; conditional comonotonicity; local comonotonicity; regular conditional distribution; projection theorem; measurable graph theorem

2000 Mathematics Subject Classification: Primary 60E99
Secondary 62H20

1. Introduction

The concept of comonotonicity has a long history, starting from Schmeidler [12] and Yaari [13]. It has numerous applications in finance and actuarial science, especially in studying the riskiness of a collection of risks. Dhaene *et al.* [4], [5] provide a good summary of the theory and its various applications. In [7], the notion of conditional comonotonicity was introduced as a natural extension of the classical notion, and was applied to dynamic economics problems. Indeed, this notion has already been used implicitly by [8] to obtain an improved convex upper bound for the sum of random variables. The aim of this paper is to provide further development and to give a more complete account of the theory. In particular, we will show that under some mild conditions, a random vector in $\mathbb{R}^n$ is comonotonic conditional to a certain σ-field if and only if it is almost surely (a.s.) locally comonotonic, i.e. its image is comonotonic in $\mathbb{R}^n$ on each atom of the σ-field, a.s. In the course of deriving this result, we have also produced a new proof on distributional representation and almost sure representation of a conditionally comonotonic random vector.

We end this introduction by giving a brief review of the concept of comonotonicity. For more information on comonotonicity, we refer the reader to [4], [5], and [9]. A subset A of $\mathbb{R}^n$ is called *comonotonic* if, for any $(s_1, \ldots, s_n)$ and $(t_1, \ldots, t_n)$ in A, $(t_i - s_i)(t_j - s_j) \geq 0$ for any $i, j \in \{1, 2, \ldots, n\}$. A random vector $(X_1, \ldots, X_n)$ in $\mathbb{R}^n$ is said to be comonotonic if it has a comonotonic support, i.e. there exists a comonotonic subset C of $\mathbb{R}^n$ such that $\mathrm{P}((X_1, \ldots, X_n) \in C) = 1$. Here, we may replace C by its closure C^- if C is not measurable, since the closure of a comonotonic set is again comonotonic.

Lemma 1. *The following statements are equivalent.*

(a) *The random vector $(X_1, \ldots, X_n)$ is comonotonic.*

(b) $\mathrm{P}(X_1 \leq x_1, \ldots, X_n \leq x_n) = \min_{1 \leq i \leq n} \mathrm{P}(X_i \leq x_i)$.

(c) *There is a P-null set N such that whenever $\omega, \omega' \in N^c$, we have*

$$(X_i(\omega) - X_i(\omega'))(X_j(\omega) - X_j(\omega')) \geq 0.$$

(d) *There exist nondecreasing functions $f_1, \ldots, f_n$ and a random variable Z such that*

$$(X_1, \ldots, X_n) \stackrel{\mathrm{D}}{=} (f_1(Z), \ldots, f_n(Z)).$$

Characterization (b) implies that the joint distribution function of a comonotonic random vector is precisely the Frèchet upper bound. Characterization (c) tells us that the pointwise behavior of a comonotonic random vector, $X_1, \ldots, X_n$, is almost surely moving together in the same direction. Characterization (d) gives us a distributional representation. In fact, we may take Z to be a uniform$(0, 1)$ random variable, and f_i to be $F_{X_i}^{-1}$, where $F_{X_i}^{-1}(s) = \inf\{t \in \mathbb{R} \mid F_{X_i}(t) \geq s\}$. Lemma 1 will be generalized to the conditional case in the sequel.

2. Definitions and notation

We first recall the definition of a probability kernel. Given any two measurable spaces $(A, \mathcal{A})$ and $(B, \mathcal{B})$, a probability kernel from $(A, \mathcal{A})$ to $(B, \mathcal{B})$ is a $[0, 1]$-valued function Q on $A \times \mathcal{B}$ such that $Q(\cdot, F)$ is $\mathcal{A}$-measurable for any fixed $F \in \mathcal{B}$ and $Q(a, \cdot)$ is a probability measure on $(B, \mathcal{B})$ for any fixed $a \in A$.

Consider a random vector $X = \{X_1, \ldots, X_n\}$ defined on a probability space $(\Omega, \mathcal{F}, \mathrm{P})$ and a measurable map Y from $(\Omega, \mathcal{F})$ to another measurable space $(T, \mathcal{T})$. For instance, T could be $\mathbb{R}^m$ and $\mathcal{T}$ the Borel σ-field on $\mathbb{R}^m$. In this case, Y is a random vector. There always exists a probability kernel μ from $(T, \mathcal{T})$ to $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ satisfying

$$\mathrm{P}(X \in A \mid Y) = \mu(Y, A) \quad \text{a.s. for each } A \in \mathcal{B}(\mathbb{R}^n); \tag{1}$$

moreover, this kernel is unique in the sense that if μ' is another kernel satisfying (1), then $\mu(Y, \cdot) = \mu'(Y, \cdot)$ a.s. For a proof of this standard result, we refer the reader to [10, Theorem 6.3] or [11, Section 30]. We call this kernel a regular conditional distribution of X given Y. Since each probability measure $\mu(Y(\omega), \cdot)$ has a unique extension to the universal completion $\widehat{\mathcal{B}(\mathbb{R}^n)}$ of $\mathcal{B}(\mathbb{R}^n)$, and μ remains a probability kernel after this extension, we will hereafter assume that the domain of each $\mu(Y(\omega), \cdot)$ is $\widehat{\mathcal{B}(\mathbb{R}^n)}$.

Note that if $\tilde{X}_1, \ldots, \tilde{X}_n$ are n random variables such that $(X_1, \ldots, X_n) = (\tilde{X}_1, \ldots, \tilde{X}_n)$ a.s., then from the definition of conditional expectation,

$$\mu(Y, A) = \tilde{\mu}(Y, A) \quad \text{a.s.}$$

for any $A \in \mathcal{B}(\mathbb{R}^n)$, where $\tilde{\mu}$ denotes the regular conditional distribution of $(\tilde{X}_1, \ldots, \tilde{X}_n)$ given Y. By uniqueness, we have $\mu(Y, \cdot) = \tilde{\mu}(Y, \cdot)$ a.s.

The next definition is adopted from [7] after a slight modification.

Definition 1. Let X, Y, and μ be defined as above. We say that X is comonotonic conditional to Y (or simply Y-comonotonic) if the probability measure $\mu(Y(\omega), \cdot)$ has a comonotonic support in $\mathbb{R}^n$ for almost all ω.

In [7], the object to be conditioned on is a sub-σ-field of $\mathcal{F}$; in our definition here, it is a random element. While the two approaches are equivalent, the reasons for our choice are twofold:

(i) it seems that conditioning on a random element can make the roles of Y and $\sigma\{Y\}$ clearer;

(ii) in the later part of this paper, the proof demands a certain degree of regularity on the structure of the conditioning σ-field, which can be achieved by assuming that it is generated by a measurable map taking values in a 'nice' space.

Indeed, Kaas *et al.* [8] also implicitly used the random element approach but not the σ-field approach.

Note that Definition 1 generalizes the classical definition of comonotonicity in the sense that if $(X_1, \ldots, X_n)$ is comonotonic in the classical sense, then it is Y-comonotonic whenever $\sigma\{Y\}$ is the trivial σ-field on Ω.

Given any σ-field $\mathcal{H}$ on Ω, we can partition Ω by the following equivalence relation:

$$\omega \sim \omega' \iff \mathbf{1}_H(\omega) = \mathbf{1}_H(\omega') \quad \text{for all } H \in \mathcal{H}.$$

The equivalence classes of this equivalence relation are called $\mathcal{H}$-atoms. In general, an $\mathcal{H}$-atom need not be $\mathcal{H}$-measurable. However, it can be shown that if the σ-field $\mathcal{H}$ is countably generated (i.e. if there exists a countable subset $\mathcal{H}_0 \subseteq \mathcal{H}$ such that $\sigma\{\mathcal{H}_0\} = \mathcal{H}$), then every $\mathcal{H}$-atom is $\mathcal{H}$-measurable. Examples of countably generated σ-fields are those generated by a measurable map taking values in a separable metric space. We remark that any measurable random element ξ from $(\Omega, \mathcal{H})$ into $(M, \mathcal{B})$ would be constant on every $\mathcal{H}$-atom when M is a separable metric space and $\mathcal{B}$ is the Borel σ-field on M.

The following notation will be used throughout: the regular conditional distribution of X given Y is always denoted as μ. For any real number x, we use $(-\infty, x]_i$ to denote the cylinder $\mathbb{R} \times \cdots \times \mathbb{R} \times (-\infty, x] \times \mathbb{R} \times \cdots \times \mathbb{R}$, where the interval $(-\infty, x]$ appears at the ith coordinate. If C is a subset of $\mathbb{R}^n$, then the closure, i.e. the smallest closed set containing C, is denoted as C^-. If ν is a measure on a measurable space $(S, \mathcal{S})$, then the completion of $\mathcal{S}$ with respect to the measure ν is denoted as $\overline{\mathcal{S}}^\nu$, and the universal completion of $\mathcal{S}$ is denoted as $\widehat{\mathcal{S}}$.

3. Two easy characterizations

Let ω be an arbitrary fixed point in Ω. For any $i \in \{1, 2, \ldots, n\}$, the map

$$x \mapsto \mu_i(Y(\omega), x) := \mu(Y(\omega), (-\infty, x]_i)$$

is a distribution function on $\mathbb{R}$. We could define its inverse as follows:

$$\mu_i^{-1}(\omega, u) = \inf\{x \in \mathbb{R} \mid \mu_i(Y(\omega), x) \geq u\}.$$

Obviously, for fixed $\omega \in \Omega$, $u \mapsto \mu_i^{-1}(\omega, u)$ and is nondecreasing and left-continuous.

Next we fix $u \in (0, 1)$. For any $s \in \mathbb{R}$, observe that

$$\{\omega \colon \mu_i^{-1}(\omega, u) \leq s\} = \{\omega \colon u \leq \mu_i(Y(\omega), s)\} \in \sigma\{Y\}$$

because μ is a probability kernel. Hence, the map

$$\omega \mapsto \mu_i^{-1}(\omega, u), \qquad i = 1, 2, \ldots, n,$$

is $\sigma\{Y\}$-measurable for any fixed $u \in (0, 1)$. From this, we may represent this map as

$$\omega \mapsto \mu_i^{-1}(\omega, u) = v_i(Y(\omega), u). \tag{2}$$

If U is a uniform$(0, 1)$ random variable, then the distribution of $\mu_i^{-1}(\omega, U)$ is the i-coordinate projection of the measure $\mu(Y(\omega), \cdot)$, that is

$$\mu_i^{-1}(\omega, U) \overset{\text{D}}{=} \mu(Y(\omega), \cdot) \circ \pi_i^{-1},$$

where π_i is the i-coordinate projection map from $\mathbb{R}^n$ to $\mathbb{R}$.

Now we present two easy characterizations of conditional comonotonicity. Characterization (b) is analogous to statement (b) in Lemma 1. Characterization (c) will be used in the sequel and a similar statement can be found in [7]. Here we give a more detailed proof.

Lemma 2. *The following statements are equivalent.*

(a) $(X_1, \ldots, X_n)$ *is Y-comonotonic.*

(b) *For any fixed real numbers $x_1, \ldots, x_n$,*

$$\mathrm{P}(X_1 \le x_1, \ldots, X_n \le x_n \mid Y) = \min_{1 \le i \le n} \mathrm{P}(X_i \le x_i \mid Y) \quad a.s.$$

(c) *If U is a* uniform$(0, 1)$ *random variable then, for* P *almost all ω, the random vector*

$$(\mu_1^{-1}(\omega, U), \ldots, \mu_n^{-1}(\omega, U))$$

has distribution $\mu(Y(\omega), \cdot)$.

Proof. Statement (a) implies statement (b). For any fixed real numbers $x_1, \ldots, x_n$,

$$\begin{aligned}
\mathrm{P}(X_1 \le x_1, \ldots, X_n \le x_n \mid Y) &= \mu(Y, \times_{i=1}^n (-\infty, x_i]) \quad \text{a.s.} \\
&= \min_{1 \le i \le n} \mu(Y, (-\infty, x_i]_i) \quad \text{a.s.} \\
&= \min_{1 \le i \le n} \mathrm{P}(X_i \le x_i \mid Y) \quad \text{a.s.,}
\end{aligned}$$

where the second equality follows from Lemma 1 and the P almost sure comonotonicity of $\mu(Y, \cdot)$.

Statement (b) implies statement (a). For any fixed real numbers $x_1, \ldots, x_n$,

$$\mu(Y, \times_{i=1}^n (-\infty, x_i]) = \min_{1 \le i \le n} \mu(Y, (-\infty, x_i]_i) \quad \text{a.s.,}$$

where the exceptional null set depends on the choice of $x_1, \ldots, x_n$. By considering rational x_i first and then using the continuity of the measure $\mu(Y(\omega), \cdot)$, we have

$$\mu(Y, \times_{i=1}^n (-\infty, x_i]) = \min_{1 \le i \le n} \mu(Y, (-\infty, x_i]_i) \quad \text{for all } x_1, \ldots, x_n, \text{ a.s.}$$

This means that the measure $\mu(Y, \cdot)$ is comonotonic almost surely.

Statement (a) implies statement (c). We calculate the joint distribution of $\mu_1^{-1}(\omega, U), \ldots,$ $\mu_n^{-1}(\omega, U)$ as follows: for any real numbers $s_1, \ldots, s_n$,

$$
\begin{aligned}
\mathrm{P}(\mu_1^{-1}(\omega, U) &\leq s_1, \ldots, \mu_n^{-1}(\omega, U) \leq s_n) \\
&= \mathrm{P}(U \leq \mu_1(Y(\omega), s_1), \ldots, U \leq \mu_n(Y(\omega), s_n)) \\
&= \mathrm{P}(U \leq \min_{1 \leq i \leq n} \mu_i(Y(\omega), s_i)) \\
&= \min_{1 \leq i \leq n} \mu_i(Y(\omega), s_i) \\
&= \mu(Y(\omega), \times_{i=1}^n (-\infty, s_i]) \quad \text{a.s.,}
\end{aligned}
$$

where the last equality follows from Lemma 1. This shows that (c) is true.

Statement (c) implies statement (a). Since each map $\mu_i^{-1}(\omega, \cdot)$ is nondecreasing for fixed ω, the random vector $(\mu_1^{-1}(\omega, U), \ldots, \mu_n^{-1}(\omega, U))$ is comonotonic. Hence, its distribution $\mu(Y(\omega), \cdot)$ is comonotonic for almost all ω by Lemma 1.

4. Local comonotonicity

In [7], the following example was considered. Suppose that $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ and the random variables X_1 and X_2 are defined by

	ω_1	ω_2	ω_3	ω_4
X_1	1	2	3	4
X_2	3	4	1	2

If P assigns $\frac{1}{4}$ to each $\{\omega_i\}$, it is obvious that X_1 and X_2 are not comonotonic in the classical sense. However, as stated (without proof) by Jouini and Napp [7], X_1 and X_2 are comonotonic conditional to the σ-field $\mathcal{G} = \{\phi, \{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}, \Omega\}$ (say generated by some Y). The intuitive idea is that if X_1 and X_2 are comonotonic locally on every $\sigma\{Y\}$-atom, then they are Y-comonotonic. Inspired by this example, we introduce the following definition.

Definition 2. Let X, Y, and μ be defined as above. We say that X is locally comonotonic on Y (or simply locally Y-comonotonic) if there is a P-null set N such that, for every $\sigma\{Y\}$-atom A,

$$
C_A := \{X(\omega) \mid \omega \in A \cap N^c\} \tag{3}
$$

is a comonotonic subset in $\mathbb{R}^n$.

The example used by Jouini and Napp [7] suggests that local comonotonicity implies conditional comonotonicity. The aim of this section is to give a formal proof of this assertion; moreover, we shall show that these two concepts are actually equivalent. In the mean time, we shall also provide a different proof of the distribution representation and an almost sure representation of a Y-comonotonic random vector.

Before we present the theorem which deals with the general case, it would be beneficial to note (see Theorem 1, below) that the above questions can be answered in an elementary way when $\sigma\{Y\}$ is generated by a countable partition. Under this special situation, there is no need to assume any particular structure on $(T, \mathcal{T})$ and $(\Omega, \mathcal{F})$.

Theorem 1. *When $\sigma\{Y\}$ is generated by a countable partition, conditional comonotonicity is equivalent to local comonotonicity.*

Proof. Suppose that X is Y-comonotonic. Let A be an $\sigma\{Y\}$-atom with $P(A) > 0$. Then there exists $\omega_A \in A$ such that $\mu(Y(\omega_A), \cdot)$ has a comonotonic support in $\mathbb{R}^n$, i.e. $\mu(Y(\omega_A), \hat{C}_A) = 1$ for some measurable comonotonic subset $\hat{C}_A$ of $\mathbb{R}^n$. As $\mu(Y(\cdot), \hat{C}_A)$ is $\sigma\{Y\}$-measurable, it is constant on A. Hence, $\mu(Y(\omega), \hat{C}_A) = 1$ for all $\omega \in A$. Since A is $\sigma\{Y\}$-measurable, we have

$$\int_A \mathbf{1}_{\{X \in \hat{C}_A\}}(\omega) \, P(d\omega) = \int_A P(X \in \hat{C}_A \mid Y)(\omega) \, P(d\omega) = \int_A \mu(Y(\omega), \hat{C}_A) \, P(d\omega) = P(A).$$

This means that $X \in \hat{C}_A$ a.s. on A. Since there are only countably many $\sigma\{Y\}$-atoms, we may conclude that outside a certain P-null set N, X is comonotonic on $A \cap N^c$ for each $\sigma\{Y\}$-atom A, i.e. X is locally Y-comonotonic.

Conversely, suppose that there is a P-null set N such that X is comonotonic on $A \cap N^c$ for each $\sigma\{Y\}$-atom A. Let A be any $\sigma\{Y\}$-atom A with $P(A) > 0$, and let C_A be the image of X on $A \cap N^c$ as defined in (3). Then C_A is a comonotonic subset of $\mathbb{R}^n$. Since $\mu(Y(\cdot), C_A^-)$ is constant on the $\sigma\{Y\}$-atom A (to be denoted as k), then

$$\begin{aligned}
k \, P(A) &= \int_A \mu(Y(\omega), C_A^-) \, P(d\omega) \\
&= \int_A P(X \in C_A^- \mid Y)(\omega) \, P(d\omega) \\
&= \int_A \mathbf{1}_{\{X \in C_A^-\}}(\omega) \, P(d\omega) \\
&= P(A),
\end{aligned}$$

hence $k = 1$. This implies that $\mu(Y(\omega), \cdot)$ has a comonotonic support for almost all ω, i.e. X is Y-comonotonic.

We can see that the above argument breaks down when there are uncountably many $\sigma\{Y\}$-atoms with zero probability because their union may have a strictly positive probability. To overcome this difficulty, it seems that extra regularity is needed on our underlying probability space and $\sigma\{Y\}$. In particular, we shall see in the next theorem that it is sufficient to assume that both $(T, \mathcal{T})$ and $(\Omega, \mathcal{F})$ are Polish spaces, i.e. T and Ω are Polish spaces, $\mathcal{T}$ and $\mathcal{F}$ are the respective Borel σ-fields on them. For example, we may model our underlying probability space $(\Omega, \mathcal{F}, P)$ to be $([0, 1], \mathcal{B}([0, 1]), \lambda)$, where λ is the Lebesgue measure on $\mathcal{B}([0, 1])$.

To prepare for the theorem, we need the next two results. A Borel space is a measurable space that is Borel isomorphic to a Borel subset in $[0, 1]$. For example, every Polish space is Borel.

Lemma 3. *Suppose that ξ and η are two random elements in Borel spaces S and T, respectively, and f is a measurable map from T to S. If $\xi \overset{\mathrm{D}}{=} f(\eta)$, then there exists $\tilde{\eta} \overset{\mathrm{D}}{=} \eta$ such that $\xi = f(\tilde{\eta})$ a.s.*

Lemma 4. *For any random elements η, $\tilde{\eta}$, ξ, and $\tilde{\xi}$, $P(\xi \in B \mid \eta) = P(\tilde{\xi} \in B \mid \tilde{\eta})$ a.s. for any measurable set B if and only if $(\eta, \xi) \overset{\mathrm{D}}{=} (\tilde{\eta}, \tilde{\xi})$.*

Lemma 3 is taken from [10, Corollary 6.11]; Lemma 4 follows from the definition of conditional expectation.

Theorem 2. *Suppose that both $(T, \mathcal{T})$ and $(\Omega, \mathcal{F})$ are Polish spaces, i.e. T and Ω are Polish spaces, $\mathcal{T}$ and $\mathcal{F}$ are the Borel σ-fields on them. Then the following statements are equivalent.*

(a) *$(X_1, \ldots, X_n)$ is Y-comonotonic.*

(b) *There exist real-valued functions $(y, u) \mapsto v_i(y, u)$ on $T \times (0, 1)$ ($i = 1, 2, \ldots, n$) with the property that $v_i(Y(\omega), \cdot)$ is nondecreasing for fixed $\omega \in \Omega$, and $v_i(\cdot, u)$ is $\mathcal{T}$-measurable for fixed $u \in (0, 1)$, such that*

$$(X_1, \ldots, X_n, Y) \overset{\mathrm{D}}{=} (v_1(Y, U), \ldots, v_n(Y, U), Y), \tag{4}$$

where U is any uniform$(0, 1)$ random variable on $(\Omega, \mathcal{F}, \mathrm{P})$ that is independent of Y.

(c) *There exist a random element $\hat{Y} : (\Omega, \mathcal{F}) \to (T, \mathcal{T})$, a random variable $\hat{U}$, and real-valued functions $(y, u) \mapsto v_i(y, u)$ on $T \times (0, 1)$ ($i = 1, 2, \ldots, n$), with the property that $v_i(Y(\omega), \cdot)$ is nondecreasing for fixed $\omega \in \Omega$, $v_i(\cdot, u)$ is $\mathcal{T}$-measurable for fixed $u \in (0, 1)$, and $Y = \hat{Y}$ a.s., such that*

$$(X_1, \ldots, X_n) = (v_1(\hat{Y}, \hat{U}), \ldots, v_n(\hat{Y}, \hat{U})) \quad a.s.$$

(d) *$(X_1, \ldots, X_n)$ is locally Y-comonotonic, i.e. there is a P-null set N such that, for every $\sigma\{Y\}$-atom A,*

$$C_A = \{ \boldsymbol{X}(\omega) \mid \omega \in A \cap N^c \} \tag{5}$$

is a comonotonic subset in $\mathbb{R}^n$.

Proof. Statement (a) implies statement (b). Let f be a bounded measurable function on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ and g a bounded measurable function on $(T, \mathcal{T})$. Let U be any uniform$(0, 1)$ random variable on $(\Omega, \mathcal{F}, \mathrm{P})$ that is independent of Y. Consider functions v_i ($i = 1, 2, \ldots, n$) as defined in (2). They have the properties required in statement (b). Moreover,

$$
\begin{aligned}
\mathrm{E}[&f(v_1(Y, U), \ldots, v_n(Y, U))g(Y)] \\
&= \mathrm{E}[g(Y) \, \mathrm{E}[f(v_1(Y, U), \ldots, v_n(Y, U)) \mid Y]] \\
&= \mathrm{E}^\omega[g(Y(\omega)) \, \mathrm{E}^U[f(v_1(Y(\omega), U), \ldots, v_n(Y(\omega), U))]] \\
&= \mathrm{E}^\omega[g(Y(\omega)) \, \mathrm{E}^U[f(\mu_1^{-1}(\omega, U), \ldots, \mu_n^{-1}(\omega, U))]] \\
&= \mathrm{E}^\omega\!\left[g(Y(\omega)) \int_{\mathbb{R}^n} f(x_1, \ldots, x_n) \mu(Y(\omega), \mathrm{d}x_1 \times \cdots \times \mathrm{d}x_n) \right] \\
&= \mathrm{E}^\omega[g(Y(\omega)) \, \mathrm{E}[f(X_1, \ldots, X_n) \mid Y](\omega)] \\
&= \mathrm{E}[f(X_1, \ldots, X_n)g(Y)],
\end{aligned}
$$

where the fourth equality follows from Lemma 2. Here, E^U means that the variable being integrated over is U; similarly, E^ω means that the integration is taken over ω. We can conclude from this calculation that

$$(X_1, \ldots, X_n, Y) \overset{\mathrm{D}}{=} (v_1(Y, U), \ldots, v_n(Y, U), Y).$$

Hence, (b) is true.

Statement (b) implies statement (c). Suppose that $(X_1, \ldots, X_n)$ has the distribution representation (4). By Lemma 3, there exists $(\hat{Y}, \hat{U}) \stackrel{\mathrm{D}}{=} (Y, U)$ such that

$$(X_1, \ldots, X_n, Y) = (\nu_1(\hat{Y}, \hat{U}), \ldots, \nu_n(\hat{Y}, \hat{U}), \hat{Y}) \quad \text{a.s.}$$

In particular, this implies that $Y = \hat{Y}$ a.s. and

$$(X_1, \ldots, X_n) = (\nu_1(\hat{Y}, \hat{U}), \ldots, \nu_n(\hat{Y}, \hat{U})) \quad \text{a.s.}$$

Statement (c) implies statement (d). There exist P-null sets N_1 and N_2 such that $Y = \hat{Y}$ on N_1^c, and $(X_1, \ldots, X_n) = (\nu_1(\hat{Y}, \hat{U}), \ldots, \nu_n(\hat{Y}, \hat{U}))$ on N_2^c. Let $N = N_1 \cup N_2$, then N is also a P-null set. Let A be an arbitrary $\sigma\{Y\}$-atom. Then, for any $\omega, \omega' \in A \cap N^c$, we have $\nu_i(Y(\omega), \cdot) = \nu_i(Y(\omega'), \cdot)$ (call this common value $\nu_i(Y(A), \cdot)$); hence, for any $i, j \in \{1, 2, \ldots, n\}$, we have

$$\begin{aligned}
&(X_i(\omega) - X_i(\omega'))(X_j(\omega) - X_j(\omega')) \\
&= (\nu_i(Y(A), \hat{U}(\omega)) - \nu_i(Y(A), \hat{U}(\omega')))(\nu_j(Y(A), \hat{U}(\omega)) - \nu_j(Y(A), \hat{U}(\omega'))) \\
&\geq 0,
\end{aligned}$$

where the last inequality follows because each map ν_i is nondecreasing in its second argument when the first argument is fixed. This shows that the set C_A, as defined in (5), is comonotonic.

Statement (d) implies statement (a). We break down the proof into the following steps.

(i) The construction of $\tilde{X}$. Since the exceptional null set N in statement (d) may not be $\sigma\{Y\}$-measurable, there can be three different types of $\sigma\{Y\}$-atom. Let $\mathcal{N}_1$ be the collection of all $\sigma\{Y\}$-atoms, A, such that $A \subseteq N$; $\mathcal{N}_2$ the collection of all $\sigma\{Y\}$-atoms, A, such that $A \subseteq N^c$; and $\mathcal{N}_3$ be the collection of all $\sigma\{Y\}$-atoms that are not in $\mathcal{N}_1$ or $\mathcal{N}_2$. Let $\boldsymbol{x}^*$ be a point in $\mathbb{R}^n$ such that

$$\mathrm{P}((X_1, \ldots, X_n) \in \{\boldsymbol{x}^*\}) = 0.$$

If such a point does not exist, then $\sigma\{Y\}$ must be generated by a countable partition and, hence, the result follows from Theorem 1. Then there exists a P-null set N_3 such that $\mu(Y(\omega), \{\boldsymbol{x}^*\}) = 0$ when $\omega \in N_3^c$. Then we define the following random vector:

$$\tilde{X}(\omega) = (\tilde{X}_1, \ldots, \tilde{X}_n)(\omega) = \begin{cases} \boldsymbol{x}^*, & \omega \in N, \\ (X_1, \ldots, X_n)(\omega), & \omega \in N^c. \end{cases}$$

By construction, $(\tilde{X}_1, \ldots, \tilde{X}_n) = (X_1, \ldots, X_n)$ a.s. Therefore, if $\tilde{\mu}$ denotes the regular conditional distribution of $\tilde{X}$ given Y, then $\mu(Y, \cdot) = \tilde{\mu}(Y, \cdot)$ a.s. (say outside P-null set N_4). Moreover, for any $\sigma\{Y\}$-atom A, the set

$$\tilde{C}_A = \{(\tilde{X}_1, \ldots, \tilde{X}_n)(\omega) \mid \omega \in A\}$$

is a comonotonic subset of $\mathbb{R}^n$ when $A \in \mathcal{N}_1 \cup \mathcal{N}_2$; when $A \in \mathcal{N}_3$, $\tilde{C}_A$ can be expressed as $C_A \cup \{\boldsymbol{x}^*\}$, where C_A is the comonotonic set defined in (5).

(ii) The measurability of sections of the image of $(\tilde{X}(\omega), Y)$. Since $\mathbb{R}^n \times T$ is a Polish space and $\mathcal{B}(\mathbb{R}^n) \otimes \mathcal{T}$ is the Borel σ-field, by the measurable graph theorem (see, for example, [2, Chapter 2]), the graph of $(\tilde{X}, Y)$, i.e.

$$\mathrm{Gr}(\tilde{X}, Y) = \{(\omega, \tilde{X}(\omega), Y(\omega)) \mid \omega \in \Omega\},$$

is $\mathcal{F} \otimes \mathcal{B}(\mathbb{R}^n) \otimes \mathcal{T}$ measurable in $\Omega \times \mathbb{R}^n \times T$. Since the set

$$G = \{(\tilde{X}(\omega), Y(\omega)) \mid \omega \in \Omega\}$$

is the projection of $\mathrm{Gr}(\tilde{X}, Y)$ on $\mathbb{R}^n \times T$, it is measurable with respect to $\widehat{\mathcal{B}(\mathbb{R}^n) \otimes \mathcal{T}}$, the universal completion of $\mathcal{B}(\mathbb{R}^n) \otimes \mathcal{T}$, by the well-known projection theorem (see [3, Chapter 3]).

Now we show that, for any $y \in T$, the section $G_y = \{\tilde{X}(\omega) \mid (\tilde{X}(\omega), y) \in G\}$ belongs to $\widehat{\mathcal{B}(\mathbb{R}^n)}$. To see this, let ν and $\mathbb{Q}$ be arbitrary probability measures on $(T, \mathcal{T})$ and $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, respectively. Then

$$G \in \widehat{\mathcal{B}(\mathbb{R}^n) \otimes \mathcal{T}} \subseteq \overline{\mathcal{B}(\mathbb{R}^n) \otimes \mathcal{T}}^{\mathbb{Q} \otimes \nu}$$

and, hence, there exist $G^{(1)}, G^{(2)} \in \mathcal{B}(\mathbb{R}^n) \otimes \mathcal{T}$ such that

$$G^{(1)} \subseteq G \subseteq G^{(2)} \quad \text{and} \quad (\mathbb{Q} \otimes \nu)(G^{(2)} \setminus G^{(1)}) = 0.$$

Then the cross-sections $G_y^{(i)} = \{\tilde{X}(\omega) \mid (\tilde{X}(\omega), y) \in G^{(i)}\} \in \mathcal{B}(\mathbb{R}^n)$, $i = 1, 2$, will satisfy

(a) $G_y^{(1)} \subseteq G_y \subseteq G_y^{(2)}$, and

(b) $\mathbb{Q}(G_y^{(2)} \setminus G_y^{(1)}) = 0$ for ν-almost all y by Fubini's theorem.

As ν can be chosen arbitrarily, $G_y \in \overline{\mathcal{B}(\mathbb{R}^n)}^{\mathbb{Q}}$ for all $y \in T$. As $\mathbb{Q}$ is also arbitrary, we conclude that $G_y \in \widehat{\mathcal{B}(\mathbb{R}^n)}$.

(iii) The concentration property of conditional measures. Having shown that each section $G_y \in \widehat{\mathcal{B}(\mathbb{R}^n)}$, we can now develop the concentration property of regular conditional probability. The idea is borrowed from [6]. Let $\tilde{\mu}$ be the regular conditional distribution of $\tilde{X}$ given Y that has been extended to the universal completion $\widehat{\mathcal{B}(\mathbb{R}^n)}$. Observe that

$$\begin{aligned}
\int_\Omega \tilde{\mu}(Y(\omega), G_{Y(\omega)}) \, \mathrm{P}(\mathrm{d}\omega) &= \int_\Omega [\tilde{\mu}(y, G_y)]_{y=Y(\omega)} \, \mathrm{P}(\mathrm{d}\omega) \\
&= \int_\Omega \left[\int_{\mathbb{R}^n} \mathbf{1}_{\{\boldsymbol{x} \in G_y\}} \, \tilde{\mu}(y, \mathrm{d}\boldsymbol{x}) \right]_{y=Y(\omega)} \mathrm{P}(\mathrm{d}\omega) \\
&= \int_\Omega \mathrm{P}[\tilde{X} \in G_Y \mid Y](\omega) \, \mathrm{P}(\mathrm{d}\omega) \\
&= \mathrm{P}(\tilde{X} \in G_Y) \\
&= \mathrm{P}((\tilde{X}, Y) \in G) \\
&= 1,
\end{aligned}$$

hence there is a P-null set $N_5 \in \sigma\{Y\}$ such that whenever $\omega \in N_5^c$, $\tilde{\mu}(Y(\omega), G_{Y(\omega)}) = 1$. This P-null set N_5 can be expressed as a union of $\sigma\{Y\}$-atoms because $\sigma\{Y\}$ is countably generated. If A is a $\sigma\{Y\}$-atom that lies outside N_5, then we must have $\tilde{\mu}(Y(\omega), B) = 0$ whenever $\omega \in A$, $B \in \widehat{\mathcal{B}(\mathbb{B}^n)}$, and $\tilde{X}^{-1}(B) \cap A = \varnothing$, because in this case we have $B \subseteq G_{Y(A)}^c$.

(iv) Now let ω be an arbitrary point in $N_3^c \cap N_3^c \cap N_5^c$ and A be the $\sigma\{Y\}$-atom containing ω. If $A \in \mathcal{N}_1 \cup \mathcal{N}_3$ then

$$\mu(Y(\omega), \tilde{C}_A^-) = \tilde{\mu}(Y(\omega), \tilde{C}_A^-) = 1,$$

where the last equality follows from the fact that $\tilde{X}^{-1}((\tilde{C}_A^-)^c) \cap A = \varnothing$ and the definition of the null set N_5. Similarly, when $A \in \mathcal{N}_2$, then

$$\mu(Y(\omega), C_A^-) = \mu(Y(\omega), C_A^- \cup \{\boldsymbol{x}^*\}) = \tilde{\mu}(Y(\omega), C_A^- \cup \{\boldsymbol{x}^*\}) = 1.$$

Hence, $\mu(Y(\omega), \cdot)$ has a comonotonic support whenever we have $\omega \in N_3^c \cap N_4^c \cap N_5^c$. Hence, $(X_1, \ldots, X_n)$ is Y-comonotonic.

Similar distributional representation and the almost sure representation of a conditionally comonotonic vector, as described in statements (b) and (c), were obtained in [7]. The proof presented here is new. Note that when proving that (a) implies (b), (b) implies (c), and (c) implies (d), the assumption that $(\Omega, \mathcal{F})$ is Polish is not needed. In proving that (b) implies (c), we need $(T, \mathcal{T})$ to be Polish to guarantee that Lemma 3 is applicable; in proving that (d) implies (a), we need both $(T, \mathcal{T})$ and $(\Omega, \mathcal{F})$ to be Polish spaces so that the image of $(\tilde{X}, Y)$, which is denoted as G in the proof, is measurable with respect to $\mathcal{B}(\widehat{\mathbb{R}^n}) \otimes \mathcal{T}$. It is also remarked that (c) implying (a) can be proved directly without the need for $(\Omega, \mathcal{F})$ to be Polish if $\sigma\{Y\}$ is generated by a countable partition, by using the technique in Theorem 1.

We remark that when $\sigma\{Y\}$ is taken to be the trivial σ-field (for example, when Y is a constant function), the above theorem reduces to the equivalence of statements (a), (c), and (d) in Lemma 1.

The following corollaries are direct consequences of the above theorem.

Corollary 1. *Any random vector $X = (X_1, \ldots, X_n)$ is Y-comonotonic whenever X is Y-measurable.*

Corollary 2. *Suppose that W is another random element in a Borel space $(S, \mathcal{S})$ such that $\sigma\{Y\} \subseteq \sigma\{W\}$. Then $(X_1, \ldots, X_n)$ is W-comonotonic whenever it is Y-comonotonic. In particular, if $(X_1, \ldots, X_n)$ is comonotonic in the classical sense, then it is Y-comonotonic for any random vector Y.*

Corollary 3. *If $(X_1, \ldots, X_n)$ is Y-comonotonic and $f_1, \ldots, f_n$ are n increasing functions, then $(f_1(X_1), \ldots, f_n(X_n))$ is also Y-comonotonic.*

Corollary 1 implies that the concept of conditional comonotonicity is only of interest when the conditioning random element Y is not too complicated. If the structure of Y is so complicated that it contains all the information of X (i.e. $X \subseteq \sigma\{Y\}$), then X is automatically Y-comonotonic and this concept becomes vacuous. From Corollary 2, we can also see that the simpler the conditioning random element is, the more comonotonic is the random vector. In particular, given that X is Y-comonotonic and Z is W-comonotonic, we may be able to determine whether X or Z is more comonotonic by comparing the size of $\sigma\{Y\}$ and $\sigma\{W\}$. If in addition X and Z have the same conditional marginal distributions, it can be proved that the sum of the one that is more comonotonic has a higher convex order than the sum of the other one. For a precise statement and more information in relation to convex ordering, we refer the reader to [1].

5. Conclusion

In this paper, we have studied several characterizations of conditional comonotonicity. It is shown that many results in the classical setting can be extended to the conditional case. The main result describes the local behavior of a conditionally comonotonic random vector on atoms of the conditioning σ-field. In the development, some regularity assumptions are made concerning the structure of the state space $(T, \mathcal{T})$ and the underlying probability space $(\Omega, \mathcal{F})$. It would be interesting to see whether these assumptions can be relaxed.

Acknowledgements

References

[1] CHEUNG, K. C. (2007). Improved convex upper bound via conditional comonotonicity. To appear in *Insurance Math. Econom.*

[2] CHRISTENSEN, J. P. R. (1974). *Topology and Borel structure*. North-Holland, Amsterdam.

[3] DELLACHERIE, C. AND MEYER, P. A. (1978). *Probabilities and Potential* (North-Holland Math. Studies **29**). North-Holland, Amsterdam.

[4] DHAENE, J. *et al.* (2002). The concept of comonotonicity in actuarial science and finance: theory. *Insurance Math. Econom.* **31,** 3–33.

[5] DHAENE, J. *et al.* (2002). The concept of comonotonicity in actuarial science and finance: applications. *Insurance Math. Econom.* **31,** 133–161.

[6] HOFFMANN-JØRGENSEN, J. (1971). Existence of conditional probabilities. *Math. Scand.* **28,** 257–264.

[7] JOUINI, E. AND NAPP, C. (2004). Conditional comonotonicity. *Decis. Econom. Finance* **27,** 153–166.

[8] KAAS, R., DHAENE, J. AND GOOVAERTS, M. J. (2000). Upper and lower bounds for sums of random variables. *Insurance Math. Econom.* **27,** 151–168.

[9] KAAS, R. *et al.* (2002). A simple geometric proof that comonotonic risks have the convex-largest sum. *Astin Bull.* **32,** 71–80.

[10] KALLENBERG, O. (2002). *Foundations of Modern Probability*, 2nd edn. Springer, New York.

[11] LOÈVE, M. (1978). *Probability Theory. II*, 4th edn. Springer, New York.

[12] SCHMEIDLER, D. (1986). Integral representation without additivity. *Proc. Amer. Math. Soc.* **97,** 255–261.

[13] YAARI, M. E. (1987). The dual theory of choice under risk. *Econometrica* **55,** 95–115.