A Measure of Monotonicity of Two Random Variables

1Farida Kachapova and 2Ilias Kachapov
1School of Computing and Mathematical Sciences, Faculty of Design and Creative Technology, Auckland University of Technology, New Zealand
2Examination Academic Services, University of Auckland, Auckland, New Zealand

Abstract: Problem statement: When analyzing random variables it is useful to measure the degree of their monotone dependence or compare pairs of random variables with respect to their monotonicity. Existing coefficients measure general or linear dependence of random variables. Developing a measure of monotonicity is useful for practical applications as well as for general theory, since monotonicity is an important type of dependence. Approach: Existing measures of dependence are briefly reviewed. The Reimann coefficient is generalized to arbitrary random variables with finite variances. Results: The article describes criteria for monotone dependence of two random variables and introduces a measure of this dependence-monotonicity coefficient. The advantages of this coefficient are shown in comparison with other global measures of dependence. It is shown that the monotonicity coefficient satisfies natural conditions for a monotonicity measure and that it has properties similar to the properties of the Pearson correlation; in particular, it equals 1 (-1) if and only if the pair X, Y is comonotonic (counter-monotonic). The monotonicity coefficient is calculated for some bivariate distributions and the sample version of the coefficient is defined. Conclusion/Recommendations: The monotonicity coefficient should be used to compare pairs of random variables (such as returns from financial assets) with respect to their degree of monotone dependence. In the problems where the monotone relation of two variables has a random noise, the monotonicity coefficient can be used to estimate variance and other central moments of the noise. By calculating the sample version of the coefficient one will quickly find pairs of monotone dependent variables in a big dataset.

Key words: Monotonicity, comonotonic, counter-monotonic, monotone dependence, measure of dependence, Pearson correlation, bivariate distributions, random variables

INTRODUCTION

Dependence of random variables is studied and estimated in various applications (see, for example, Tularam et al., 2010). Two most important types of dependence are linear and monotone dependence. Other types include positive and negative quadrant dependence (see Kimeldorf and Sampson, 1987).

Random variables X and Y are called Positively Quadrant Dependent (PQD) if for any \( x, y \in \mathbb{R} \), \( F_{X,Y}(x, y) \geq F_X(x) \cdot F_Y(y) \). “Negative Quadrant Dependence (NQD) is defined by reversing the concept of PQD” (Kimeldorf and Sampson, 1987).

Suppose \( H_1 \) and \( H_2 \) are the joint distribution functions of \( <X_1, Y_1> \) and \( <X_2, Y_2> \), respectively, with the same marginals; then \( <X_1, Y_1> \) is called more PQD than \( <X_2, Y_2> \) if for any \( x, y \in \mathbb{R} \), \( H_1(x, y) \geq H_2(x, y) \).

Kimeldorf and Sampson (1978) introduced the concept of monotone dependence as follows: two continuous random variables X and Y are called monotone dependent if there exists a monotone function g, in which \( Y = g(X) \) with probability 1.

Global measures of dependence of two random variables include the Pearson correlation \( \rho \), the coefficients of Spearman, Kendall, Schweizer and Wolff and others (see, for example, Scarsini, 1984; Schweizer and Wolff, 1981).

Kimeldorf and Sampson (1978) introduced monotone correlation \( \rho^* \) and showed that if X and Y are monotone dependent, then \( \rho^* (X, Y) = 1 \); but the converse is false. Also \( \rho^* \) does not distinguish between the increasing and decreasing types of monotonicity.

The Spearman, Kendall and Schweizer-Wolff coefficients can be used to measure monotone dependence but they are more appropriate for ordinal
variables rather than continuous variables, since they depend only on the ranks of the observations.

Schweizer and Wolff (1981) used copulas and metrics to generate several nonparametric measures of dependence. Copulas are a useful technique developed earlier for Pearson, Spearman and Kendall correlations.

Reimann (1992) introduced a measure $\lambda^{**}$ for PQD random variables $X$ and $Y$. In this article it is generalized to some random variables $X$ and $Y$ with finite variances; it is shown to measure monotonicity.

**MONOTONICITY CONDITIONS**

In this article we will consider only non-degenerate random variables (a degenerate random variable is constant almost surely).

As usual, for a random variable $X$ with distribution function $F_X$, the quantile function $F_X^{-1}$ is defined by:

$$F_X^{-1}(u) = \inf \{ x \in \mathbb{R} : F_X(x) \geq u \}, \quad u \in [0, 1],$$

with $\inf \emptyset = +\infty$ by convention.

The quantile function $F_X^{-1}$ is non-decreasing and left-continuous. Other simple properties of the quantile function are listed in the following lemma.

**Lemma 1:**

1) For any $u \in [0, 1]$, $x \in \mathbb{R}$: $F_X^{-1}(u) \leq x \Leftrightarrow u \leq F_X(x)$

2) $F_{X,Y}^{-1}(u) = F_X^{-1}(u) + \alpha$ for all $u \in (0, 1)$

3) If $\beta > 0$, then $F_{X,Y}^{-1}(u) = \beta F_X^{-1}(u)$ for all $u \in (0, 1)$

4) If $\beta < 0$, then $F_{X,Y}^{-1}(u) = \beta F_X^{-1}(1-u)$ for all $u \in (0, 1)$, except a countable or finite set of points

The concepts of comonotonicity and counter-monotonicity were studied by Bauerle and Muller (1998); Denuit and Dhaene (2003); Dhaene et al. (2002); Dempster (2002); Rachev (2003). Comonotonicity of a pair of random variables $X$ and $Y$ means their monotone increasing dependence, i.e., their values change in the same direction. Counter-monotonicity of the pair of $X$ and $Y$ means their monotone decreasing dependence, i.e., their values change in opposite directions. Dhaene et al. (2002) give a mathematically accurate definition of comonotonicity for $n$ random variables, which we reproduce here for $n = 2$.

1) A set $A \subseteq \mathbb{R}^2$ is called comonotonic if for any of its elements $\langle x_1, y_1 \rangle$ and $\langle x_2, y_2 \rangle$: either ($x_1 \leq x_2$ and $y_1 \leq y_2$) or ($x_1 \geq x_2$ and $y_1 \geq y_2$) holds.

2) A pair $\langle X, Y \rangle$ of random variables is said to be comonotonic if it has a comonotonic support.

A counter-monotonic pair $\langle X, Y \rangle$ is defined by changing the second row in 1) to:

either ($x_1 \leq x_2$ and $y_1 \geq y_2$) or ($x_1 \geq x_2$ and $y_1 \leq y_2$) holds.

Three criteria for comonotonicity were proven in (Dhaene et al., 2002). In case $n = 2$ they have the following form.

**Theorem 1:** A pair $\langle X, Y \rangle$ of random variables is comonotonic if and only if one of the following equivalent conditions holds:

1) $(\forall x, y \in \mathbb{R}) [F_{X,Y}(x, y) = \min \{F_X(x), F_Y(y)\}]$

2) For $U \sim \text{Uniform}(0, 1)$, $\langle X, Y \rangle = \langle F_X^{-1}(U), F_Y^{-1}(U) \rangle$

3) There exist a random variable $Z$ and non-decreasing functions $g, h$, such that

$$\langle X, Y \rangle = \langle g(Z), h(Z) \rangle$$

Here $F_{X,Y}$ denotes the joint distribution function of $X, Y$ and $\ = \ $ denotes equality in distribution. A similar theorem for counter-monotonicity follows.

**Theorem 2:** The pair $\langle X, Y \rangle$ is counter-monotonic if and only if one of the following equivalent conditions holds:

4) $(\forall x, y \in \mathbb{R}) [F_{X,Y}(x, y) = \max \{F_X(x), F_Y(y)\} - 1, 0)]$

5) For $U \sim \text{Uniform}(0, 1)$

$$\langle X, Y \rangle = \langle F_X^{-1}(U), F_Y^{-1}(1-U) \rangle$$

6) There exist a random variable $Z$, a non-decreasing function $g$ and a non-increasing function $h$, such that:

$$\langle X, Y \rangle = \langle g(Z), h(Z) \rangle$$

In the following theorem we prove other criteria for comonotonicity and counter-monotonicity. We assume that the random variables $X$ and $Y$ are defined on the same probability space $\langle \Omega, \Sigma, P \rangle$, where $\Sigma$ is the collection of all events in this space.

**Theorem 3:**

1) The pair $\langle X, Y \rangle$ is comonotonic if and only if there exists $B \in \Sigma$, such that $P(B) = 1$ and

7) $$(\forall \omega_1, \omega_2 \in B) [X(\omega_1) < X(\omega_2) \Rightarrow Y(\omega_1) \leq Y(\omega_2)]$$

2) The pair $\langle X, Y \rangle$ is counter-monotonic if and only if there exists $B \in \Sigma$, such that $P(B) = 1$ and
(8) \((\forall \omega_1, \omega_2 \in B) [X (\omega_1) < X (\omega_2) \Rightarrow Y (\omega_2) \geq Y (\omega_1)]\)

Proof:
1) \(\Rightarrow\) Suppose \(<X, Y>\) is comonotonic. Then it has a comonotonic support \(A\). Denote:

\[B = \{\omega \in \Omega: < X (\omega), Y (\omega)> \in A\}\]

Consider \(\omega_1, \omega_2 \in B\) with \(X (\omega_1) < X (\omega_2)\). This implies \(Y (\omega_2) \geq Y (\omega_1)\), since both pairs \(<X (\omega_1), Y (\omega_1)>\) and \(<X (\omega_2), Y (\omega_2)>\) belong to \(A\).

\(\Leftarrow\) Suppose there exists \(B \subseteq \Omega\) of probability 1, such that the condition (7) holds. Denote:

\[A = \{< X (\omega), Y (\omega)>: \omega \in B\}\]

Then \(P (<X, Y> \in A) = P (B) = 1\), so \(A\) is a support of \(<X, Y>\).

Suppose \(<x_1, y_1>\) \(\in A\) and \(<x_2, y_2>\) \(\in A\). Then for some \(\omega_1, \omega_2 \in B\), \(x_1 = X (\omega_1), y_1 = Y (\omega_1), x_2 = X (\omega_2)\) and \(y_2 = Y (\omega_2)\). By (7), \(x_1 < x_2\) implies \(y_1 \geq y_2\) and \(x_1 > x_2\) implies \(y_1 \leq y_2\). Hence \(A\) is comonotonic.

2) \(<X, Y>\) is counter-monotonic if and only if \(<X, -Y>\) is comonotonic. Hence part 2) of the Theorem follows from part 1)

Note: Clearly \(X\) and \(Y\) can be interchanged with each of the conditions (7) and (8).

Sometimes the condition (1) in Theorem 1 is taken as the definition for comonotonicity and the condition (4) in Theorem 2 for counter-monotonicity. The criteria in Theorem 3 are more suitable for the definitions of comonotonicity and counter-monotonicity, since they reflect their meaning and are similar to the definitions of increasing and decreasing functions.

In the case of continuous marginals \(F_X, F_Y\), the second inequality in the formulas (7) and (8) can be made more strict. Also in the following two theorems the conditions (2), (3), (5) and (6) are made stronger (the proofs follow from Theorem 3).

Theorem 4: Suppose the marginal distribution functions \(F_X\) and \(F_Y\) are continuous. The pair \(<X, Y>\) is comonotonic if and only if one of the following equivalent conditions holds:

- \(Y = F_Y^\prime (F_X (X))\) with probability 1.
- There is a non-decreasing function \(g\) such that \(Y = g (X)\) with probability 1.

Theorem 5: Suppose the marginal distribution functions \(F_X\) and \(F_Y\) are continuous. The pair \(<X, Y>\) is counter-monotonic if and only if one of the following equivalent conditions holds:

- \(Y = F_Y^\prime [1 - F_X (X)]\) with probability 1.
- There is a non-increasing function \(h\) such that \(Y = h (X)\) with probability 1.

According to Theorems 4 and 5, in the case of continuous marginals, comonotonicity (counter-monotonicity) is equivalent to monotone increasing (decreasing) dependence defined by Kimeldorf and Sampson (1978), which was described in our introduction.

MONOTONICITY COEFFICIENT

We will fix a random variable \(U\) with the uniform distribution on \((0, 1)\). For a random variable \(X\) with distribution function \(F_X\) denote:

\[X^* = F_X^{-1} (U)\]

By the quantile transfer theorem, \(X^* \equiv^d X\) and \(X' \equiv^d X\) (see Dhaene et al., 2002).

Lemma 2: For the random variable \(Y\) and \(\alpha, \beta \in R\), the following holds:

1) \((Y + \alpha)^* = Y^* + \alpha\).
2) If \(\beta > 0\), then \((\beta Y)^* = \beta Y^*\) and \((\beta Y)^\prime = \beta Y^\prime\).
3) If \(\beta < 0\), then \((\beta Y)^* = \beta Y^*\) with probability 1 and \((\beta Y)^\prime = \beta Y^\prime\) with probability 1.

Lemma 2 follows from Lemma 1.

In the rest of the article we will consider only non-degenerate random variables with finite variances. The following theorem presents some well-known results (see, for example, Denuit and Dhaene, 2003) using the aforementioned notations.

Theorem 6: For the random variables \(X\) and \(Y\) the following holds:

1) \(<X^*, Y^*>\) is comonotonic and \(\text{Cov} (X^*, Y^*) > 0\).
2) \(<X^*, Y^*>\) is counter-monotonic and \(\text{Cov} (X^*, Y^*') = \text{Cov} (X^*, Y^*) < 0\).
3) \(\text{Cov} (X^*, Y^*) \leq \text{Cov} (X, Y) \leq \text{Cov} (X^*, Y^*)\).
4) \(<X, Y>\) is comonotonic \(\iff\) \(\text{Cov} (X, Y) = \text{Cov} (X^*, Y^*)\).
5) \(<X, Y>\) is counter-monotonic \(\iff\) \(\text{Cov} (X, Y) = \text{Cov} (X^*, Y^*)\).
The monotonicity coefficient $\rho_m$ of random variables $X$ and $Y$ is defined by:

$$\rho_m(X, Y) = \begin{cases} \frac{\text{Cov}(X, Y)}{\text{Cov}(X^*, Y^*)} & \text{if } \text{Cov}(X, Y) > 0, \\ 0 & \text{if } \text{Cov}(X, Y) = 0, \\ -\frac{\text{Cov}(X, Y)}{\text{Cov}(X^*, Y^*)} & \text{if } \text{Cov}(X, Y) < 0. \end{cases}$$

The definition is valid due to Theorem 6.1), 2).

When the coefficient $\rho_m(X, Y)$ is closer to 1, then Cov $(X, Y)$ is closer to the covariance Cov $(X^*, Y^*)$ of a comonotonic pair, hence the pair $<X, Y>$ is more comonotonic. Similarly, when the coefficient $\rho_m(X, Y)$ is closer to $-1$, then Cov $(X, Y)$ is closer to the covariance Cov $(X^*, Y^*)$ of a counter-monotonic pair, hence the pair $<X, Y>$ is more counter-monotonic. Thus, $\rho_m$ can be used to measure the degree of monotonicity (monotone dependence) between $X$ and $Y$.

**Example 1:** Suppose $X \sim \text{Uniform}(0, 1)$, $Y = g(X)$ and $Z = h(X)$, where:

$$g(x) = \begin{cases} 2x & \text{if } 0 \leq x \leq \frac{1}{3}, \\ 1-x & \text{if } \frac{1}{3} < x \leq \frac{2}{3}, \\ 2x-1 & \text{if } \frac{2}{3} < x \leq 1, \end{cases}$$

and:

$$h(x) = \begin{cases} 1.8x & \text{if } 0 \leq x \leq \frac{1}{3}, \\ -0.6x + 0.8 & \text{if } \frac{1}{3} < x \leq \frac{2}{3}, \\ 2.8x - 0.8 & \text{if } \frac{2}{3} < x \leq 1 \end{cases}$$

The monotonicity coefficients for these bivariate distributions are:

$$\rho_m(X, Y) = \frac{13}{10} = 0.7692 \text{ and } \rho_m(X, Z) = \frac{293}{261} = 0.8908.$$  

So the second pair is more comonotonic, which is also obvious from the graphs in Fig. 1.

**Theorem 7: Properties of the monotonicity coefficient:** For the random variables $X$ and $Y$ the following holds:

1) $\rho_m(X, Y) > 0 \Leftrightarrow \rho(X, Y) > 0$,

where $\rho(X, Y)$ is the Pearson correlation of $X$ and $Y$.

2) $\rho_m(X, Y) < 0 \Leftrightarrow \rho(X, Y) < 0$.

3) $|\rho(X, Y)| \leq |\rho_m(X, Y)| \leq 1$.

The degree of linear dependence is not greater than the degree of monotone dependence.

4) $-1 \leq \rho_m(X, Y) \leq 1$.

5) $\rho_m(Y, X) = \rho_m(X, Y)$.

6) If $X$ and $Y$ are independent, then $\rho_m(X, Y) = 0$.

7) For any $\alpha \in \mathbb{R}$: $\rho_m(X + \alpha Y, Y) = \rho_m(X, Y)$.

8) For any $\beta \in \mathbb{R}$:

\[ \text{if } \beta > 0, \text{ then } \rho_m(X, \beta Y) = \rho_m(X, Y); \]

\[ \text{if } \beta < 0, \text{ then } \rho_m(X, \beta Y) = -\rho_m(X, Y). \]

9) $\rho_m(X, Y) = 1$ if and only if the pair $<X, Y>$ is comonotonic.

10) $\rho_m(X, Y) = -1$ if and only if the pair $<X, Y>$ is counter-monotonic.

11) If $X$ and $Y$ are PQD, then $\rho_m(X, Y) \geq 0$.

12) If $X$ and $Y$ are NQD, then $\rho_m(X, Y) \leq 0$.

13) $\rho_m(X, Y) = \begin{cases} \frac{\rho(X, Y)}{\rho(X^*, Y^*)} & \text{if } \rho(X, Y) > 0, \\ 0 & \text{if } \rho(X, Y) = 0, \\ -\frac{\rho(X, Y)}{\rho(X^*, Y^*)} & \text{if } \rho(X, Y) < 0 \end{cases}$
Proof: 1)-10) follows from Lemma 2 and Theorem 6. 11) and 12) follow from the formula of Hoeffding:

\[
\text{Cov}(X,Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[ F_{X,Y}(x,y) - F_X(x)F_Y(y) \right] \, dx \, dy
\]

13) Obvious.

The linear properties in Theorem 7.7), 8) are stated only for one argument but they also hold for the other argument due to symmetry.

Thus, the properties of the monotonicity coefficient \( \rho_m \) are similar to the properties of the Pearson correlation \( \rho \) but with respect to monotone dependence. The measure \( \rho_m \) is not entirely new. For PQD random variables \( X \) and \( Y \), Reimann (1992) defined a measure \( \lambda^* \) by the formula:

\[
\lambda^*(X,Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[ \min(F_{X,Y}(x,y)) - F_X(x)F_Y(y) \right] \, dx \, dy
\]

The coefficient \( \rho_m \) is a generalization of \( \lambda^* \) to any random variables with finite variances; for PQD variables \( X \) and \( Y \), Reimann (1992) did not study the properties of \( \lambda^* \) except the property \( \lambda^* \geq \rho \). He described \( \lambda^* \) as a measure of association of two random variables rather than a measure of monotone dependence. He defined \( \lambda^* \) in terms of double integrals and \( \rho_m \) has a simpler definition in terms of covariances.

COMPARISON TO OTHER COEFFICIENTS AND APPLICATIONS

The relation between \( \rho_m \) and \( \rho \) is described in the following theorem.

Theorem 8: Relation to the Pearson correlation:
Suppose \( \text{Cov}(X,Y) \neq 0 \). Then:

\[
\rho_m(X,Y) = \rho(X,Y) \iff \text{there exist numbers } a \text{ and } b \text{ such that } Y = a + bX.
\]

In the following examples we compare the values of \( \rho_m \) and \( \rho \) for some bivariate distributions.

Example 2: Suppose the joint density function of \( X \) and \( Y \) is given by:

\[
f(x,y) = \begin{cases} 
2 & \text{if } 0 \leq x \leq y \leq 1, \\
0 & \text{otherwise}
\end{cases}
\]

Then the monotonicity coefficient is:

\[
\rho_m(X,Y) = \frac{2}{32 - 9\pi} \approx 0.5368
\]

versus the Pearson correlation \( \rho(X, Y) = 0.5 \).

Example 3: Suppose \( X \sim \text{normal}(0,1) \), \( \varepsilon \sim \text{normal}(\mu, \tau) \), \( X \) and \( \varepsilon \) are independent and \( Y = (X+\varepsilon)^3 \). Then the monotonicity coefficient is:

\[
\rho_m(X,Y) = \frac{1}{\sqrt{1 + \tau^2}}
\]

As expected, the monotonicity coefficient decreases with increase of \( \tau \) (the spread of the noise \( \varepsilon \)) and does not depend on \( \mu \) (the constant shift of \( X \)).

The Pearson correlation is:

\[
\rho(X,Y) = \frac{3(1 + \tau^2 + \mu^2)}{\sqrt{15(1 + \tau^2)^2 + 36\mu^2(1 + \tau^2)^2 + 9\mu^4(1 + \tau^2)}}
\]

In the case when \( \varepsilon \) has the standard normal distribution (\( \mu = 0, \tau^2 = 1 \)), the monotonicity coefficient is \( \rho_m(X,Y) = \frac{1}{\sqrt{2}} = 0.7071 \) versus the Pearson correlation \( \rho(X, Y) = \frac{3}{\sqrt{30}} \approx 0.5477 \).

Example 4: Suppose \( X \) and \( \varepsilon \) are the same as in Example 3 and \( Y = (\varepsilon - X)^3 \). Then:

\[
\rho_m(X,Y) = -\rho_m(-X,Y) = -\frac{1}{\sqrt{1 + \tau^2}}
\]

Scarsini (1984) introduced some conditions for a measure of dependence of two random variables. Theorems 7 and 8 shows that \( \rho_m \) satisfies a reasonable modification of these conditions. In particular, if each of random variables \( X \) and \( Y \) has a normal distribution, then \( \rho_m(X,Y) = \rho(X,Y) \) by Theorem 8.

One of the Scarsini’s conditions is the invariance of a measure of concordance under increasing transformations of \( X \) and \( Y \). This condition might be useful for a measure of general dependence but not monotone dependence, since the result of increasing transformations of two variables can get closer to or further from a monotonic relation than the original pair, i.e., increasing transformations can change the
degree of monotone dependence. The Pearson correlation does not satisfy this condition; it measures linear dependence, which is a particular case of monotone dependence. The Spearman, Kendall and Schweizer-Wolff coefficients satisfy this condition; they depend only on the ranks of the observations. The coefficient $\rho_m$ is only invariant under changes of scale and location in $X$ and $Y$. We believe that $\rho_m$ is a more appropriate measure of monotone dependence of two variables and illustrate this with the following two examples.

**Example 5.** Table 1 defines random variables $X$, $Y$ and $Z$ on the sample space $\Omega = \{\varepsilon_1, \varepsilon_2, \varepsilon_3\}$ with $P(\varepsilon_k) = \frac{1}{3}$ ($k = 1, 2, 3$).

Since the pairs $<X, Y>$ and $<X, Z>$ have the same ranks, their Kendall coefficients are equal: $\tau(X, Y) = \tau(X, Z) = \frac{1}{3}$ and so are their Spearman coefficients:

$\rho_S(X, Y) = \rho_S(X, Z) = \frac{1}{2}$. But the second pair is more counter-monotonic: it is closer to a decreasing relation as the graphs in Fig. 2 shows; this is reflected by its lower monotonicity coefficient:

$\rho_m(X, Y) = -\frac{2}{3} \approx -0.667$ and

$\rho_m(X, Z) = -\frac{10}{11} \approx -0.909$

**Example 6:** Table 2 defines random variables $X$, $Y$ and $Z$ on the sample space $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$ with $P(\omega_k) = \frac{1}{6}$ ($k = 1, 2, ..., 6$).

Since the pairs $<X, Y>$ and $<X, Z>$ have the same ranks, their Kendall coefficients are equal: $\tau(X, Y) = \tau(X, Z) = \frac{11}{15}$ and so are their Spearman coefficients:

$\rho_S(X, Y) = \rho_S(X, Z) = \frac{31}{35}$. But the second pair is more comonotonic: it is closer to an increasing relation as the graphs in Fig. 3 shows; this is reflected by its higher monotonicity coefficient:

$\rho_m(X, Y) = 0.766$ and $\rho_m(X, Z) = 0.991$.

The coefficient $\rho_m$ can be applied to the problems where $Y$ is a monotone function of $X$ with a random noise $\varepsilon$ included; the monotonicity coefficient can be used to estimate some characteristics of the noise. It is natural to assume that $\varepsilon$ has a normal distribution, as usual. The following two examples illustrate some cases when $\rho_m(X, Y)$ is used to estimate the variance and central moments of the noise. Clearly, $\rho_m$ cannot be used to estimate the mean of $\varepsilon$, since this mean is a constant shift of $X$ or $Y$ and it does not affect the degree of their monotone dependence. In the following examples we also assume the normality of $X$ or $Y$, which simplifies the calculations.

**Example 7:** Suppose $Y = g(X) + \varepsilon$, where $g$ is a monotone function, noise $\varepsilon$ has a normal distribution with variance $\tau^2$, variables $X$ and $\varepsilon$ are independent and $g(X)$ has a normal distribution with variance $\beta^2$. Then:

$$\rho_m(X, Y) = -\frac{\beta}{\sqrt{\beta^2 + \tau^2}}$$

if $g$ is increasing and

$$\rho_m(X, Y) = -\frac{\beta}{\sqrt{\beta^2 + \tau^2}}$$

if $g$ is decreasing.
The result holds for the particular case when \( g(x) = \ln x \), so \( X \) has a lognormal distribution with the second parameter \( \beta \). The following example is a generalization of Example 3.

**Example 8:** Suppose \( Y = g(X + \varepsilon) \), where \( g \) is a monotone function, noise \( \varepsilon \) has a normal distribution with variance \( \tau^2 \), variables \( X \) and \( \varepsilon \) are independent and \( X \) has a normal distribution with variance \( \beta^2 \). Then

\[
\rho_m(X, Y) = \frac{\beta}{\sqrt{\beta^2 + \tau^2}} \quad \text{if } g \text{ is increasing and}
\]

\[
- \frac{\beta}{\sqrt{\beta^2 + \tau^2}} \quad \text{if } g \text{ is decreasing.}
\]

For the random variables from Examples 7 and 8, the variance \( \tau^2 \) of the noise \( \varepsilon \) can be expressed in terms of the monotonicity coefficient \( \rho_m = \rho_m(X, Y) \):

\[
\tau^2 = \beta^2 \left( \frac{1}{\rho_m^2} - 1 \right)
\]

Since \( \varepsilon \) has a normal distribution, this also defines its central moments:

\[
\sigma^2_{2k} = \beta^2 \left( \frac{1}{\rho_m^2} - 1 \right)(2k-1)!!, \quad k = 1, 2, 3, \ldots
\]

Thus, in applications the monotonicity coefficient can be used to estimate the central moments of the noise.

The coefficient \( \rho_m \) naturally generates a monotonicity measure \( r_m \) for a two-dimensional sample:

\[
r_m(x, y) = \begin{cases} 
  \frac{s(x, y)}{s(x', y')} & \text{if } s(x, y) > 0, \\
  0 & \text{if } s(x, y) = 0, \\
  -\frac{s(x, y)}{s(x', y')} & \text{if } s(x, y) < 0
\end{cases}
\]

where \( s(x, y) \) is the sample covariance, \( x^* \) is the sample \( x \) with its values in ascending order and \( y' \) is the sample \( y \) with its values in descending order. The properties of \( r_m \) are similar to the properties of \( \rho_m \). Details are given in (Kachapova and Kachapov, 2010).

**CONCLUSION**

This article introduced the monotonicity coefficient \( \rho_m \), a new measure of the monotone dependence of random variables with finite variances. It was proven that \( \rho_m \) satisfies reasonable conditions for such a measure:

- \( \rho_m \) has linear properties
- it is invariant under changes of scale and location
- \( \rho_m(X, Y) = 0 \) for independent random variables \( X, Y \)
- \( \rho_m(X, Y) = 1 \) for a comonotonic pair \( X, Y \)
- \( \rho_m(X, Y) = -1 \) for a counter-monotonic pair \( X, Y \)

The coefficient \( \rho_m \) is a more sensitive measure of monotonicity than the coefficients depending only on the ranks of observations.

The sample version \( r_m \) of the monotonicity coefficient was defined.

We recommend using \( \rho_m \) to compare pairs of random variables with respect to their degree of monotonicity. For example, in portfolio analysis the monotonicity coefficient can be used to assess the degree of increasing or decreasing monotone dependence between two asset returns and to do respective comparison of pairs of assets. In the problems where the monotone relation of two variables has a random noise, the coefficient \( \rho_m \) can be used to estimate variance and other central moments of the noise.

We recommend to use the sample monotonicity coefficient \( r_m \) to find monotonic relationships in big datasets.
REFERENCES


