Compatibility of matrices for correlation-based measures of concordance

Takaaki Koike

<tkoike@uwaterloo.ca>
<https://uwaterloo.ca/scholar/tkoike/home>

Department of Statistics and Actuarial Science
University of Waterloo

Joint work with Marius Hofert
Based on the paper Hofert and Koike (2019), ASTIN Bulletin.

September 9th, 2019

A motivating example

- Given a 3×3 matrix

$$P = \begin{pmatrix} 1 & -0.95 & 0.5 \\ -0.95 & 1 & -0.4 \\ 0.5 & -0.4 & 1 \end{pmatrix},$$

  how to check whether P is a correlation matrix?
- For a correlation matrix P, one can always find a r.v. $\boldsymbol{X}$ (for e.g., $\mathrm{N}(\mathbf{0}, P)$) s.t. $\rho(\boldsymbol{X}) = P$.
- What about matrices of pairwise Spearman's rho, Kendall's tau... or other pairwise measures of concordance (MOC)?

Definitions

Definition 1.1 (κ-compatibility)

For a given $d \times d$ matrix R and a bivariate MOC

$$\kappa : (X, Y) \mapsto [-1, 1],$$

R is called κ-compatible if there exists a continuous d-random vector $\boldsymbol{X} = (X_1, \dots, X_d)$ such that

$$\kappa_d(\boldsymbol{X}) := (\kappa(X_i, X_j))_{i,j=1,\dots,d} = R.$$

Definition 1.2 (κ-compatible set)

A set of all κ-compatible matrices is called a κ-compatible set.

Why do we study compatibility?

- Matrices of pairwise MOCs are often provided as estimates from real data.
- In practice of risk management...
  - the amount of data available is sometimes limited, and
  - risk managers may opt to incorporate important scenarios into the dependence parameters of the models.
- For such cases, some entries of the estimated matrix of pairwise MOC can be determined exogenously by expert opinions.

Main questions

1. Does there exist a class of MOCs whose compatibility is easy to study?

   $\Rightarrow$ We introduce a correlation-based transformed rank measures of concordance.

2. Can we characterize κ-compatible sets for some paticular κ, such as Spearman's rho and Kendall's tau?

   $\Rightarrow$ Positive answers for **Spearman's rho, Blomqvist's beta and van der Waerden's coefficient**.

   $\Rightarrow$ For **Kendall's tau and Gini's gamma**, their characterizations are left open problems.

For ρ: Pearson's linear correlation and two functions g_1, g_2, consider the bivariate measure

$$\kappa_{g_1,g_2}(X_1, X_2) = \rho(g_1(X_1), g_2(X_2)).$$

Definition 2.1 (Seven axioms for MOC; Scarsini, 1984)

1. **Domain**: $\kappa(X, Y)$ is defined for any continuous random variables X, Y.
2. **Symmetry**: $\kappa(X, Y) = \kappa(Y, X)$.
3. **Coherence**: if $C_{X,Y} \preceq C_{X',Y'}$, then $\kappa(X, Y) \leq \kappa(X', Y')$.
4. **Range**: $-1 \leq \kappa(X, Y) \leq 1$.
5. **Independence**: if X and Y are independent, then $\kappa(X, Y) = 0$.
6. **Change of sign**: $\kappa(-X, Y) = -\kappa(X, Y)$.
7. **Continuity**: $\lim_{n\to\infty} \kappa(X_n, Y_n) = \kappa(X, Y)$ if $\lim_{n\to\infty} H_n = H$ pointwise for $(X_n, Y_n) \sim H_n$ and $(X, Y) \sim H$.

What are admissible g_1, g_2?

- The seven axioms imply that (c.f. Scarsini, 1984)

$$\kappa(X_1, X_2) = \kappa(f_1(X_1), f_2(X_2))$$

for any f_1, f_2: strictly increasing (or decreasing) functions.

$\Rightarrow$ $\kappa(X_1, X_2)$ is forced to be independent of the marginal distributions of X_1, X_2 but be dependent only on the copula of (X_1, X_2), which is the joint distribution of

$$(U_1, U_2) := (F_1(X_1), F_2(X_2)) \sim C_{X_1,X_2}.$$

- Therefore, we consider the following form of κ_{g_1,g_2};

$$\begin{aligned}\kappa_{g_1,g_2}(X_1, X_2) &= \rho\big(g_1(F_1(X_1)), g_2(F_2(X_2))\big)\\ &= \rho\big(g_1(U_1), g_2(U_2)\big) =: \kappa_{g_1,g_2}(C_{X_1,X_2}).\end{aligned}$$

- For κ_{g_1,g_2} to satisfy the coherence axiom, we want

$$C_{X,Y} \preceq C_{X',Y'} \Rightarrow C_{g_1(X),g_2(Y)} \preceq C_{g_1(X'),g_2(Y')}$$

since its (RHS) implies $\kappa_{g_1,g_2}(X,Y) \leq \kappa_{g_1,g_2}(X',Y')$ by coherence of ρ.

Theorem 2.1 (Monotonicity of g_1 and g_2)

Let g_1, g_2 be two continuous functions. If κ_{g_1,g_2} satisfies the seven axioms, then

$$(g_1(x) - g_1(y))(g_2(x) - g_2(y)) \geq 0 \text{ for any } x > y \in [0,1].$$

- Without the loss of generality, we can assume g_1, g_2 are both increasing functions by invariance of ρ under linear transform.

- Under the assumption of left-continuity of g_1, g_2, they are quantiles of some distribution functions. Consequently, we consider the following class:

Definition 2.2 ((G_1, G_2)–transformed rank correlations)

For two distribution functions G_1 and G_2,
(G_1, G_2)-transformed rank correlation coefficient is defined by

$$\kappa_{G_1,G_2}(X_1, X_2) = \rho\big(G_1^{-1}(F_1(X_1)), G_2^{-1}(F_2(X_2))\big),$$

where G_j^{-1} is a generalized inverses of G_j for $j = 1, 2$. We call the pair (G_1, G_2) concordance inducing if κ_{G_1,G_2} is a measure of concordance (i.e., κ_{G_1,G_2} satisfies the seven Scarsini's axioms).

Examples of κ_{G_1,G_2}

1. Spearman's rho: Let $G_1 = G_2 = G$ for G being the cdf of the uniform distribution on $[0,1]$. Then κ_{G_1,G_2} is called the Spearman's rho ρ_S:

$$\rho_S(C) \propto \iint_{[0,1]^2} (C(u,v) - \Pi(u,v))\mathrm{d}u\mathrm{d}v.$$

2. Blomqvist's beta: Let $G_1 = G_2 = G$ for G being the cdf of $\mathrm{Bern}(1/2)$. Then κ_{G_1,G_2} yields the Blomqvist's beta β:

$$\beta(C) = 4C(1/2, 1/2) - 1.$$

3. Van der Waerden's coefficient: Let $G_1 = G_2 = G$ for G being the cdf of $\mathrm{N}(0,1)$. Then κ_{G_1,G_2} is called the van der Waerden's ζ.

Example of Lognormal G-functions

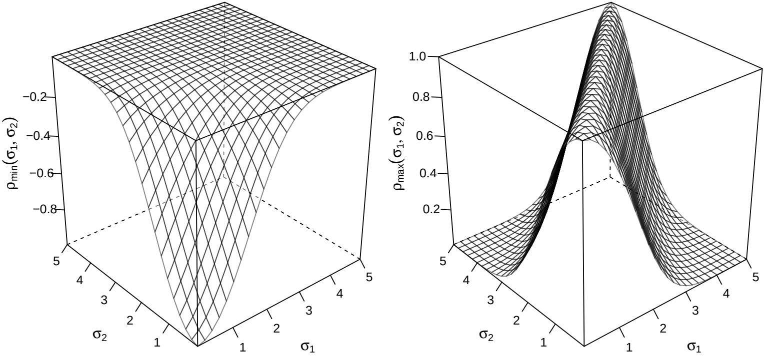

Figure: Minimal (left) and maximal (right) correlations attained by the (G_1, G_2)-transformed rank correlation coefficient κ_{G_1,G_2} where G_j is the distribution function of $\mathrm{LN}(0, \sigma_j)$, $j = 1, 2$.

Example of Bernoulli G-functions

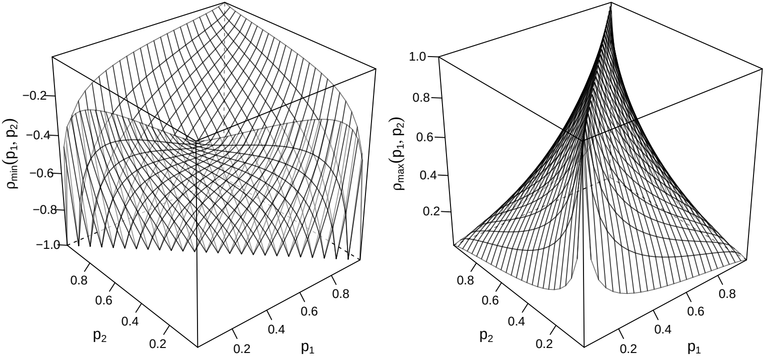

Figure: Minimal (left) and maximal (right) correlations attained by the (G_1, G_2)-transformed rank correlation coefficient κ_{G_1,G_2} where G_j is the distribution function of $\mathrm{B}(1, p_j)$, $j = 1, 2$.

Theorem 2.2 (Characterization of concordance-inducing G)

Let G_1 and G_2 be distribution functions. The (G_1, G_2)-transformed rank correlation coefficient κ_{G_1,G_2} is a measure of concordance if and only if

1. G_1 and G_2 are of the same type with G, where
2. G is a distribution function of a (i) non-degenerated (ii) radially symmetric distribution with (iii) finite second moment.

Remark: If G_1, G_2, G are all of the same type, then

$$\kappa_{G_1,G_2}(X_1, X_2) = \kappa_{G,G}(X_1, X_2) =: \kappa_G(X_1, X_2),$$

by invariance of ρ under location-scale transform. Therefore, w.l.o.g., we can assume $G_1 = G_2 = G$.

Properties of κ_G

Proposition 2.1 (Properties of κ_G)

1. **Uniqueness**: Let G and G' be two continuous concordance-inducing functions. If $\kappa_G(C) = \kappa_{G'}(C)$ for all 2-copulas, then G and G' are of the same type.
2. **Linearity**: For $n \in \mathbb{N}$, let $C_1, \dots, C_n$ be 2-copulas and $\alpha_1, \dots, \alpha_n$ be non-negative numbers such that $\alpha_1 + \cdots + \alpha_n = 1$. Then

$$\kappa_G\left(\sum_{i=1}^{n} \alpha_i C_i\right) = \sum_{i=1}^{n} \alpha_i \kappa_G(C_i).$$

Properties of the compatible set $\mathcal{K}_G$

- Recall the notation of the κ_G-compatible set:

$$\mathcal{K}_G = \{R \in \mathcal{M}^{d\times d} : \exists \boldsymbol{X}: \text{ a continuous } d\text{-r.v. s.t. } \kappa_G(\boldsymbol{X}) = R\}.$$

Proposition 3.1 (Properties of $\mathcal{K}_G$)

1. **Convexity**: $\mathcal{K}_G$ is convex,
2. **Bounds**: For any concordance inducing G, we have

$$\mathcal{P}_d^{\mathsf{B}}(1/2) \subseteq \mathcal{K}_G \subseteq \mathcal{P}_d,$$

where $\mathcal{P}_d$ is the set of all $d \times d$ correlation matrices, and $\mathcal{P}_d^{\mathsf{B}}(1/2)$ is the symmetric Bernoulli compatible set:

$$\mathcal{P}_d^{\mathsf{B}}(1/2) = \{\rho(\boldsymbol{B}) : B_j \sim \mathrm{Bern}(1/2),\ j = 1, \dots, d\}.$$

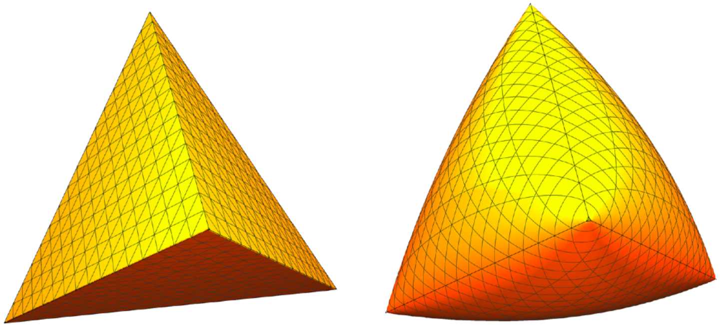

Figure: The set $\mathcal{P}_d^{\mathsf{B}}(1/2)$ (left, cut polytope) and $\mathcal{P}_d$ (right, elliptope) when $d = 3$. $d(d-1)/2 = 3$ off-diagonal entries are projected onto the Euclidean space (The figure is retrieved from Tropp, 2018).

Proposition 3.2 (Characterizations of some compatible sets)

1. **Normal variance mixture**: If $\sqrt{W}Z \sim G$ with $W \geq 0$, $\mathbb{E}W = 1$ and $Z \sim \mathrm{N}(0,1)$, then

$$\mathcal{K}_G = \mathcal{P}_d.$$

2. **Spearman's rho**: For the ρ_{S}-compatible set $\mathcal{S}_d$,

$$\mathcal{S}_d \begin{cases} = \mathcal{P}_d & d \leq 9, \\ \subset \mathcal{P}_d & d \geq 12. \end{cases}$$

3. **Blomqvist's beta**: For the β-compatible set $\mathcal{B}_d$, we have

$$\mathcal{B}_d = \mathcal{P}_d^{\mathsf{B}}(1/2) = \mathrm{conv}\{\boldsymbol{c}\boldsymbol{c}^\top : \boldsymbol{c} \in \{\pm 1\}^d\}.$$

Remark: (2) is shown in Devroye & Letac (2015) and Wang et al. (2018), and (3) is in Devroye & Letac (2015).

Our other achievements

We studied more in the paper Hofert and Koike (2019):

- we investigate the attainability problem, that concerns whether, for a given $d \times d$ matrix R, we can construct a random vector $\boldsymbol{X}$ s.t. $\kappa_G(\boldsymbol{X}) = R$.
- compatibility and attainability for block matrices are also studied to solve the problem that checking compatibility and attainability is challenging for high-dimensional matrices.

Future work

- Compatibility and attainability for Kendall's tau and other MOCs.
- Comparison among MOCs... which is the best?
- MOC for non-continuous margins... copulas are not unique but we could define a range MOC via generalized distributional transform.
- Study compatibility of measures of association, such as pairwise maximum mean discrepancy (MMD)... potentialy applications for generative moment matching neural network.

Thank you for your listening!

Email: tkoike@uwaterloo.ca

Website: https://uwaterloo.ca/scholar/tkoike/home

(the presentation slide is also available here.)

References I

1. Devroye, L. and Letac, G. (2015), Copulas with prescribed correlation matrix, *In Memoriam Marc Yor-Seminaire de Probabilites XLVII, Springer*, 585-601.
2. Edwards, H., Mikusinski, P. and Taylor, M. (2005). Measures of concordance determined by D_4-invariant measures on $(0,1)^2$, *Proceedings of the American Mathematical Society*, **133**(5), 1505-1513.
3. Huang, Jinggang and Yang, Liming (2010), Correlation matrix with block structure and efficient sampling methods, *Journal of Computational Finance*, **14**(1), 81.

References II

4. Hofert, M. and Koike, T. (2019). Compatibility and attainability of matrices of correlation-based measures of concordance. *arXiv preprint arXiv:1810.07126*.
5. Huber, M. and Maric, N. (2017), Bernoulli Correlations and Cut Polytopes, *arXiv preprint arXiv:1706.06182*.
6. Joe, Harry (1996), Families of m-variate distributions with given margins and $m(m-1)/2$ bivariate dependence parameters, *Lecture Notes-Monograph Series*, 120141.

References III

7. Roustant, O and Deville, Y (2017), On the validity of parametric block correlation matrices with constant within and between group correlations, *arXiv preprint arXiv:1705.09793*.
8. Scarsini, M. (1984), On measures of concordance, *Stochastica*, **8**(3), 201218.
9. Tropp, J. A. (2018). Simplicial faces of the set of correlation matrices. *Discrete & Computational Geometry*, **60**(2), 512-529.
10. Wang, B., Wang, R. and Wang, Y. (2018), Compatible Matrices of Spearmans Rank Correlation, *http://arxiv.org/abs/1810.03477 (10/15/2018)*.