Modifying copulas for improved dependence modelling

Colette le Roux[1] and Alta de Waal[1 2]

[1] Department of Statistics, University of Pretoria
[2] Center for Artificial Intelligence Research (CAIR)

1 Introduction

In 2007 and 2008, underestimation of correlations and risks, as well as the misuse of dependence models, lead to the financial crisis [5]. This highlighted the need to improve dependence modelling through both the correlation parameter and choice of model used. Copulas are useful for modelling dependence patterns in multivariate data, as well as prediction in regression analysis [1].

The problem is that most traditional methods for dealing with complex dependency structures assume a parametric or Gaussian distribution and linear correlation structure, but these assumptions are often violated in practical applications [3, 8]. Furthermore, the two main approaches to handling outliers or missing data are to either remove them, or replace them with some other appropriate value, but there are instances, such as in risk-management, where these anomalous observations are of key importance and cannot be eliminated. In these cases, appropriate methods are needed to model tail dependencies.

Uncertainty from volatilities, heteroskedasticity, extreme values and missing observations all contribute to the difficulty of dependency estimation and prediction. While ignoring underlying covariates might yield reasonably accurate models in some instances, time (as a covariate) has been found to have an influence on copula parameters when modelling financial data, and could therefore lead to improved prediction and estimation when taken into account [2]. Vine copulas can be applied to address these problems in the multidimensional case, where assumptions to deal with the model complexity are relaxed. A vine copula is a hierarchical factorisation of a high-dimensional copula into the product of bivariate copula densities.

The first problem in high-dimensional dependence structure models is that the computational cost of approximation and parameter estimation increases as the dimension increases [4], making traditional bivariate copula methods, such as MLE [7] and MCMC [6] practically infeasible. The second problem is that vine copulas allow for the analysis of multivariate copulas, but due to the complexity of calculating conditional copulas, the restrictive truncation and simplification assumptions are often applied. A vine copula is proposed to relax assumptions

and simplify computation, ultimately leading to a more flexible and reliable model.

2 Methodology

Given the wide variety of available bivariate copula families specifically designed to model symmetric and asymmetric distributions with central and tail dependencies, copulas are capable of modelling extreme dependencies between variables. A copula can be improved by allowing for underlying variables that influence the strength of dependencies by use of a conditional copula. As a non-parametric approach, a copula process combines a copula and a Gaussian Process (GP) to allow for non-Gaussian distributions. The GP in turn uses a Bayesian framework to deal with missing observations, adding extra flexibility to the copula density. A Gaussian process conditional copula can now be built to improve on the conditional copula [2], using Bayesian non-parametrics to learn the latent functions that specify the shape of the conditional copulas given the conditioning variables and thereby simplifying computation.

When working with the multivariate case, the Gaussian copula can easily capture high-dimensional dependence structures, but is unable to capture asymmetric tail dependencies, making it less appropriate for complex dependence structures. A vine-copula can be applied to model complex dependency structures between multivariate data by decomposing a multivariate copula into a hierarchy of bivariate copulas. This model provides flexibility in that the bivariate copulas can come from any parametric or non-parametric family and can be either conditional or unconditional. The decomposition further avoids the computational cost of the dependence optimisation problem in approximations.

The importance of improving the accuracy of dependency modelling in applications such as finance, econometrics, insurance and meteorology is self-evident, considering the potential risks involved in erroneous estimation and prediction results. In this work, we investigate the advantages, limitations and differences of copulas and vine-copulas in complex dependence structures. Prediction and estimation of complicated dependence structures is expected to improve when modifying a copula to a vine copula. It is also expected that relaxing the assumptions commonly applied to the vine copula in applications with high-dimensional dependency structures, such as independence between the conditional copula and its conditioning variable, will improve model accuracy, since underlying covariates (time in particular) has been found to have an effect on the dependency structure between the main variables. The investigation of conditional copulas and copula processes is reserved for future work.

Keywords: Copula · Copula processes · Gaussian processes · Vine copulas · Bayesian methods.

References

1. Acar, E.F., Azimaee, P., Hoque, M.E.: Predictive assessment of copula models. Canadian Journal of Statistics **47**(1), 8–26 (2019)
2. Hernández-Lobato, J.M., Lloyd, J.R., Hernández-Lobato, D.: Gaussian process conditional copulas with applications to financial time series. In: Advances in Neural Information Processing Systems. pp. 1736–1744. MIT Press (2013)
3. Jaimungal, S., Ng, E.K.: Kernel-based copula processes. In: Joint European Conference on Machine Learning and Knowledge Discovery in Databases. pp. 628–643. Springer, Springer, Berlin, Heidelberg (Sep 2009)
4. Sabeti, A., Wei, M., Craiu, R.V.: Additive models for conditional copulas. Stat **3**(1), 300–312 (Mar 2014)
5. Salmon, F.: The formula that killed wall street. Significance **9**(1), 16–20 (Feb 2012)
6. Tekumalla, L.S., Rajan, V., Bhattacharyya, C.: Vine copulas for mixed data: multi-view clustering for mixed data beyond meta-gaussian dependencies. Machine Learning **106**(9-10), 1331–1357 (2017)
7. Xu, D., Wei, Q., Elsayed, E.A., Chen, Y., Kang, R.: Multivariate degradation modeling of smart electricity meter with multiple performance characteristics via vine copulas. Quality and Reliability Engineering International **33**(4), 803–821 (Jun 2017)
8. Zheng, W., Ren, X., Zhou, N., Jiang, D., Li, S.: Mixture of d-vine copulas for chemical process monitoring. Chemometrics and Intelligent Laboratory Systems **169**, 19–34 (Oct 2017)