B-Spline Copula and Its Estimation

Xiaoling Dou[1*], Satoshi Kuriki[2], Gwo Dong Lin[3], Donald Richards[4]

[1]*Department of Natural Sciences, International Christian University*

[2]*The Institute of Statistical Mathematics*

[3]*Institute of Statistical Science, Academia Sinica*

[4]*Department of Statistics, Pennsylvania State University*

ABSTRACT

The B-spline copula is defined by a linear combination of elements of the normalized B-spline basis functions. The B-spline copula includes the Bernstein copulas as a special case. We examine the dependence properties of the B-spline copula and develop an EM algorithm to estimate the parameters of the copula. The EM algorithm is designed to maximize the penalized pseudo-likelihood function, wherein we use the smoothly clipped absolute deviation (SCAD) penalty function for the penalization term. We conduct simulation studies to demonstrate the stability of the proposed numerical procedure, show that penalization yields estimates with smaller mean-square errors when the true parameter matrix is sparse, and provide methods for determining tuning parameters and for model selection.

Keywords: B-spline basis functions; B-spline copula; EM algorithm; Model selection; SCAD penalty

Measuring Multivariate Regression Association via Spatial Sign

Jia-Han Shih[1,*], Yi-Hau Chen[2]

[1]*Department of Applied Mathematics, National Sun Yat-sen University*

[2]*Institute of Statistical Science, Academia Sinica*

ABSTRACT

In this talk, we propose a regression association measure aiming at the predictability of a multivariate outcome $\boldsymbol{Y}$ $\boldsymbol{Y} = (Y_1, \ldots, Y_p)$ from a multivariate covariate $\boldsymbol{X}$. We first generalize the conventional Kendall's tau to assess association between two random vectors. Then, we apply the generalized Kendall's tau to two independent replications from the conditional distribution of $\boldsymbol{Y}$ given $\boldsymbol{X}$, where $\boldsymbol{Y}$ and $\boldsymbol{Y'}$ share the same conditional distribution and conditionally independent given $\boldsymbol{X}$. The proposed measure can be expressed as the proportion of the variance of some function of $\boldsymbol{Y}$ $\boldsymbol{Y}$ that can be explained by $\boldsymbol{X}$, indicating that the measure has an interpretation in terms of predictability. Based on the proposed regression association measure, we further define a conditional regression association measure, which can be utilized to perform variable selection. Since our measure is based on two independent replications from the conditional distribution, a simple nonparametric estimation method based on nearest neighbor is available. Simulations are carried out to examine the performance of the proposed variable selection algorithm and real data examples are analyzed for illustration.

Keywords: Cosine similarity; Functional association; Kendall's tau; Variable selection

The Trivariate Wrapped Cauchy Copula

Shogo Kato[1], Christophe Ley[2], Sophia Loizidou[2], Kanti V. Mardia[3]

[1] *Institute of Statistical Mathematics*

[2] *Department of Mathematics, University of Luxembourg*

[3] *Department of Statistics, University of Leeds*

ABSTRACT

Toroidal data consist of observations comprising multiple angles, commonly found in environmental sciences such as wind directions and wave directions. In this talk, we propose a new distribution for trivariate toroidal data which we call a trivariate wrapped Cauchy copula. The proposed copula has the following advantages: (i) a simple form of density, (ii) an adjustable degree of dependence between every pair of variables, (iii) parameters with clear interpretation, (iv) well-known marginal and conditional distributions, (v) a straightforward data generating mechanism, (vi) unimodality, (vii) a closed-form expression for trigonometric moments, and (viii) a simple implementation procedure for obtaining maximum likelihood estimates. As is the case with general copula models, the proposed copula can be extended to have any specific marginal distributions and hence can be utilized for flexible modeling. Moreover, our construction allows for linear marginals, implying that our copula can also model cylindrical data, which consist of both angular and linear observations. As an application of the extended copula model, we consider a dataset of trivariate dihedral angles of amino acids in bioinformatics. Finally, we discuss how the proposed trivariate copula can be extended to multivariate copulas.

Keywords: Angular data; Directional statistics; Flexible modelling; Wrapped Cauchy distribution