2019-01-01

Modeling Correlated Data via Copulas

Panfeng Liang
University of Texas at El Paso

MODELING CORRELATED DATA VIA COPULAS

PANFENG LIANG

Master's Program in Computational Science

APPROVED:

Ori Rosen, Ph.D., Chair

Vanessa Lougheed, Ph.D.

Michael Pokojovy, Ph.D.

James Wood, Ph.D.

Stephen Crites, Ph.D.
Dean of the Graduate School

MODELING CORRELATED DATA VIA COPULAS

by

PANFENG LIANG

THESIS

Presented to the Faculty of the Graduate School of

The University of Texas at El Paso

in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE

Computational Science Program

THE UNIVERSITY OF TEXAS AT EL PASO

December 2019

Abstract

Copulas are widely used to model the dependency structure among components of multivariate data sets. Elliptical copulas, such as Gaussian copula, are most popular copulas being used since many data sets follow elliptical distributions or meta-elliptical distributions (Fang et al. (2002)). However, today's approaches and software packages require us to assume the specific category, such as Gaussian or Student's T, of the elliptical copula before estimating it. In this thesis, we will propose a Bayesian method using Markov chain Monte Carlo (MCMC) methods to estimate the density function of elliptical copulas without specifying it is the copula of Gaussian, Student's T or Logistic, etc.

Table of Contents

List of Figures

Chapter 1

Introduction

1.1 Introduction to Copulas

Assume that we are asked to compare the two bivariate data sets in Figure 1.1 in terms of the dependence structure between the underlying two variables. These two data sets look very different in their scatter plots. The $\boldsymbol{X} = (X_1, X_2)$ are more concentrated on the center while the $\boldsymbol{Y} = (Y_1, Y_2)$ has higher density on the bottom left corner. The similarity between their dependence structure can hardly be told by just looking at their scatter plots.

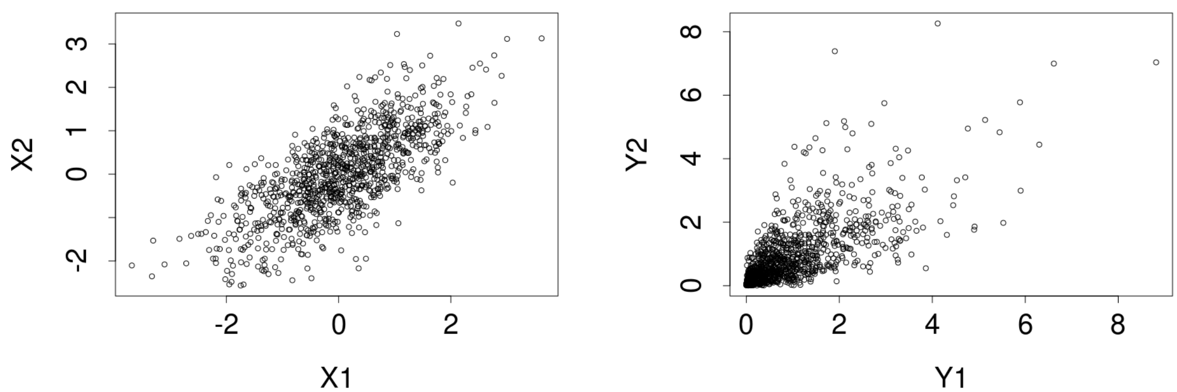

Figure 1.1: Scatter plots of $n = 1000$ independent observations of (X_1, X_2) and of (Y_1, Y_2) with the identical dependence structure

By looking at the estimated marginal kernel densities of (X_1, X_2) and (Y_1, Y_2) in Figure 1.2, it seems they have very different marginal distributions. The marginal distribution of $\boldsymbol{X}$

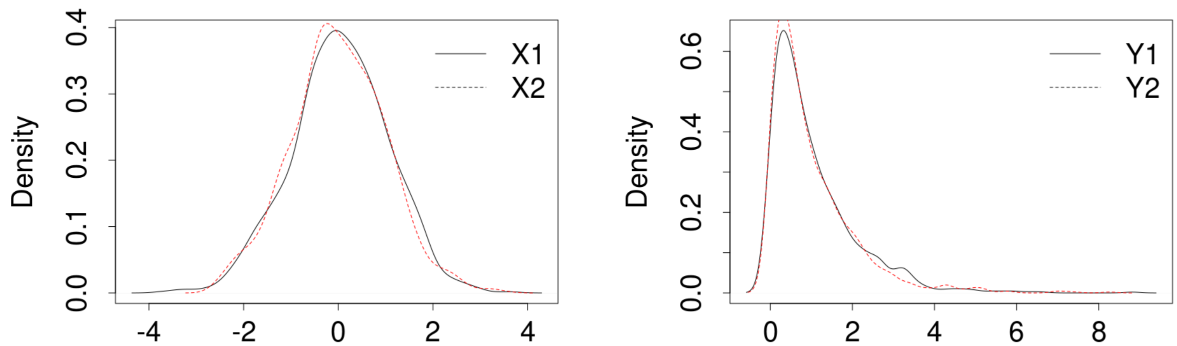

Figure 1.2: Kernel density estimates of (X_1, X_2) and of (Y_1, Y_2)

looks like a normal distribution and the one for $\boldsymbol{Y}$ is more like exponential distribution. The difference in marginal distributions increases the difficulty of the comparison in dependence structures of these two bivariate distributions. If we can use a transformation to transform $\boldsymbol{X}$ and $\boldsymbol{Y}$, making their marginal distributions identical, then the comparison of their dependence structures will be much easier.

For univariate cases, the Probability Integral Transformation (PIT) can achieve this goal. The PIT can standardize any continuous distribution into a standard uniform distribution.

Lemma 1 (Probability Integral Transformation) *Let F be the cumulative distribution function (CDF) of a continuous random variable X, that is, $X \sim F$. Then $U = F(X)$ follows a uniform distribution, that is, $U \sim U(0,1)$. The transformation $U = F(X)$ is called the probability integral transformation.*

The PIT can be generalized to the bivariate or multivariate cases.

Theorem 1 (Copula Transformation) *Let $X_1 \sim F_1, \ldots, X_d \sim F_d$, where $F_1, \ldots, F_d$ are continuous. Then*

$$U_1 = F_1(X_1) \sim U(0,1), \ldots, X_d = F_d(X_d) \sim U(0,1).$$

Hence for all $u_1, \ldots, u_d \in [0, 1]$,

$$Pr(U_1 \le u_1, \ldots, U_d \le u_d) = C(u_1, \ldots, u_d,)$$

where C *is a copula, i.e., a CDF with standard uniform marginals.*

The copula transformation converts the margins of a multivariate data into the standard uniform distribution, making the dependence structure more explicit.
Figure 1.3 is the scatter plots after the copula transformation on $\boldsymbol{X}$ and $\boldsymbol{Y}$. We can easily tell that the dependence structures between their two variables are actually the same.

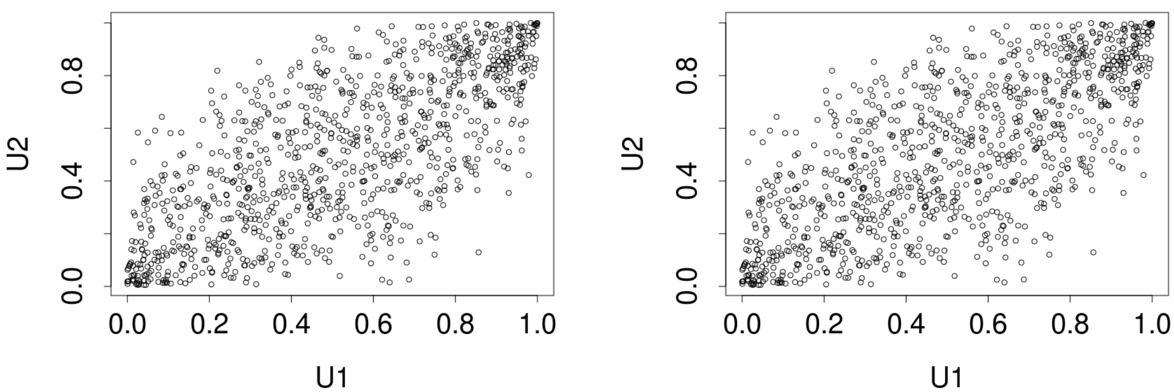

Figure 1.3: The Dependence Structure

In the example above, we can see any multivariate distribution has a underlying copula, which captures the dependence structure among its underlying variables. The copula can be obtained though the copula transformation. Although X and Y have different margins, but their copulas (dependence) are the same.
According to Sklar (1959), a multivariate distribution with continuous marginals can be unique defined by its copula and it marginal distributions. Hence, to estimate a multivariate distribution, we can estimate its copula and its marginal distributions.

Theorem 2 (Sklar's theorem) *For any CDFs $F_1, \ldots, F_d$ and any copula C*

$$H(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d)). \tag{1.1.1}$$

is a valid joint CDF with margins $F_1, \ldots, F_d$. If $F_1, \ldots, F_d$ are continuous, then C is unique.

A copula itself is a joint distribution of a random vector $\boldsymbol{U} = (U_1, \ldots, U_2)$ with uniformly distributed margins. It follows from the probability integral transformation that $U_i = F_i(Y_i)$.

The joint density function of $\boldsymbol{Y}$ is obtained by differentiation on the CDF above through chain rule:

$$\begin{aligned} h(y_1, \ldots, y_n) &= \frac{\partial^n H(y_1, \ldots, y_n)}{\partial y_1 \cdots \partial y_n} \\ &= \frac{\partial^n H(y_1, \ldots, y_n)}{\partial F_1(y_1) \cdots \partial F_1(y_n)} \times \prod_{i=1}^{n} \frac{\partial F_i(y_i)}{\partial y_i} \\ &= c\big((F_1(y_1), \ldots, F_n(y_n)\big) \cdot f_1(y_1) \cdot \cdots \cdot f_n(y_n), \end{aligned}$$

where c is the density of the copula and $f_i, i = 1, \ldots, n$, are the marginal PDFs of Y_i.

1.2 The Inversion Method

A simple way to construct a copula is by the inversion method.

Given a multivariate distribution H and continuous margins F_i, $i = 1, \ldots, p$,

$$C(u_1, \ldots, u_p) = H\big((F_1^{-1}(u_1), \ldots, F_p^{-1}(u_p)\big).$$

The C can be obtained by

$$\begin{aligned} C(u_1, \ldots, u_p) &= C(F_1(x_1), \ldots, F_p(x_p)) \\ &= H(F_1^{-1}(u_1) \ldots F_p^{-1}(u_p)). \end{aligned}$$

The inversion method is for getting the CDF of the underlying copula of a known multivariate distribution H with margins $F_i, \ldots, p$. However, if the H 's model is unknown, it is impossible to get its copula through this way.
By differentiation on C, the pdf of the copula can be obtained :

$$c(u_1, \ldots, u_p) = \frac{h(F_1^{-1}(u_1) \ldots F_p^{-1}(u_p))}{f_1(F_1^{-1}(u_1)) \ldots f_p(F_1^{-1}(u_p))} \tag{1.2.1}$$

The term $\frac{1}{f_1(x_1)\ldots f_p(x_p)}$ is actually the Jacobian of the copula transformation from $\boldsymbol{X}$ to $\boldsymbol{U}$.

1.3 Elliptical Distributions

A continuous random vector $\boldsymbol{X} = (X_1, ..., X_p)$ has an elliptical distribution $\epsilon_p(\boldsymbol{\mu}, \Omega, g)$ with location vector $\boldsymbol{\mu}$, covariance matrix Ω and generator function g, if its density function is

$$f(\boldsymbol{x}; \boldsymbol{\mu}, \Omega, g) = c_p|\Omega|^{-1/2} g((\boldsymbol{x} - \boldsymbol{\mu})'\Omega^{-1}(\boldsymbol{x} - \boldsymbol{\mu})), \tag{1.3.1}$$

where

1. $c_p = \Gamma(p/2)\pi^{-p/2} / \int_0^\infty t^{p/2-1} g(t)dt$.
2. $\boldsymbol{\mu} \in \mathbb{R}^p$.
3. Ω is a $p \times p$ positive definite matrix.
4. g is a non-negative function on $[0, \infty)$ such that $\int_0^\infty t^{p/2-1} g(t)dt < \infty$.

If $\boldsymbol{X} \sim \epsilon_p(\boldsymbol{\mu}, \Omega, g)$, then $\boldsymbol{X} \sim \epsilon_p(\boldsymbol{\mu}^*, \Omega^*, g^*)$ if and only there are two positive numbers a and b such that $\mu^* = \mu$, $\Sigma^* = a\Sigma$ and $g^*(t) = bg(at)$, for all $t \in \mathbb{R}$ (see Gómez et al. (2003)).

1.4 Stochastic Form

A p-variate vector $\boldsymbol{X}$ from $\epsilon_p(\boldsymbol{\mu}, \Omega, g)$ can also be expressed in the following stochastic representation (see Fang et al. (2002) and Fang et al. (2005))

$$\boldsymbol{X} = \boldsymbol{\mu} + RA\boldsymbol{U}, \tag{1.4.1}$$

where

1. A is the Cholesky factor of Ω ($\Omega = AA'$).

2. $\boldsymbol{U}$ is a random vector uniformly distributed on the p-dimensional unit sphere of $\mathbb{R}^p$.

3. R is a continuous nonnegative random variable, independent of $\boldsymbol{U}$, whose density function is

$$h(r) = \frac{2}{\int_0^\infty t^{p/2-1} g(t) dt} r^{p-1} g(r^2) \mathbb{1}_{[0,\infty]}(r). \tag{1.4.2}$$

Let $\boldsymbol{Y} = \Omega^{-1/2}(\boldsymbol{X} - \boldsymbol{\mu})$, then $\boldsymbol{Y}'\boldsymbol{Y} = (\boldsymbol{X} - \boldsymbol{\mu})'\Omega^{-1}(\boldsymbol{X} - \boldsymbol{\mu})$. By using the stochastic representation (1.4.1) of $\boldsymbol{X}$, it follows that $\boldsymbol{Y}'\boldsymbol{Y} = R\boldsymbol{U}'A'\Omega^{-1}A\boldsymbol{U}R = R\boldsymbol{U}'\boldsymbol{U}R = R^2$. Also, from (1.4.2), it follows that the density function (1.3.1) can be expressed as

$$f(r; A) = \frac{\Gamma(p/2)}{2\pi^{p/2}} |A|^{-1} r^{1-p} h(r). \tag{1.4.3}$$

In (1.4.3), the Cholesky factor A and the univariate density function $h(r)$ are unknown. We propose to estimate A parametrically and $h(r)$ nonparametrically.

1.5 Elliptical Copulas

By the inversion method mentioned in 1.2, a copula can be constructed from a multivariate distribution H with its margins.

For example, a well known copula, Gaussian copula, is obtained by the CDF of a multivariate Gaussian distribution and the CDF of the standard normal distribution

$$\Phi_n\big((\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_n)\big).$$

The Gaussian distribution belongs to the elliptical family. The copula of other distributions in elliptical family, such as the copula of multivariate Student-t distribution, can also be used to be obtained with the inversion method. This type of copulas are called copulas of elliptical distributions or elliptical copulas.

1.6 Meta-elliptical Distributions

In Sklar's theorem, we know the joint distribution can be expressed in terms of its univariate margins and a copula. If the copula of a distribution is an elliptical copula, this distribution is called a meta-elliptical distribution. See Fang et al. (2002).

The Y in Figure 1.1 is a sample from a meta-elliptical distribution, which is constructed with a Gaussian copula and exponential margins.

Unlike elliptical distributions, meta-elliptical distributions do not require identical margins, so it serves a more general tool to model data sets.

1.7 Naming Convention of Copulas

Copulas are usually named after the distribution used to induce the dependence between the variables.

For example, if a copula capturing the dependence of variables in a Gaussian distribution is called a Gaussian copula. By this naming convention, the underlying copula of an elliptical distribution should be called an elliptical copula.

However, some articles, such as Genest et al. (2007), refer to elliptical copulas as meta-elliptical copulas. This also follows the naming convention, because the name meta-elliptical copula literally means copula describing the dependence of variables in a meta-elliptical distribution, which is an elliptical copula according to Fang et al. (2002).

In other words, an elliptical copula and a meta-elliptical copula mean the same thing. In this thesis, we will use the name elliptical copula.

Chapter 2

Dirichlet Process

2.1 Introduction

The Dirichlet process (DP) was originally defined in Ferguson (1973), as follows.

Definition 1 *Let γ be a non-null finite measure (nonnegative and finitely additive) on measurable space $(\mathscr{X}, \mathscr{A})$. We say P is a Dirichlet process with parameters γ if for every $k = 1, 2, \ldots,$ and measurable partition $(B_1, \cdots, B_k)$ of $\mathscr{X}$, the distribution of $(P(B_1), \cdots, P(B_k))$ is Dirichlet, $\mathscr{D}(\gamma(B_1), \cdots, \gamma(B_k))$.*

2.2 Metaphors

2.2.1 Stick-breaking Process

Sethuraman (1994) gave a constructive definition of the Dirichlet process. The stick-breaking process consists of defining an infinite sequence of mixing proportions by

$$w_1 = \beta_1, \ w_k = \beta_k \prod_{j=1}^{k-1}(1-\beta_j), \quad k = 2, \ 3, \ldots,$$

where $\beta_1, \beta_2 \ldots \overset{iid}{\sim} \text{Beta}(1, \gamma)$. Sethuraman (1994) showed that the DP is actually an infinite sum of the form $P = \sum_{k=1}^{\infty} w_k \delta_{\phi_k}$ that obeys the definition of the stick-breaking process.

Proof of The Equivalence of The Stick-breaking Process and The DP

Three lemmas will be used to prove that the stick-breaking process is a DP.

Lemma 2 *Let* $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \ldots, \gamma_k)$ *and* $\boldsymbol{\delta} = (\delta_1, \delta_2, \ldots, \delta_k)$ *be k-dimensional vectors. Let* $\boldsymbol{U}$, $\boldsymbol{V}$ *be independent k-dimensional random vectors with Dirichlet distribution* $Dir(\boldsymbol{\gamma})$ *and* $Dir(\boldsymbol{\delta})$, *respectively. Let* W *be independent of* $(\boldsymbol{U}, \boldsymbol{V})$ *having a Beta distribution* $Beta(\sum \gamma_j, \sum \delta_j)$. *Then the distribution of* $W\boldsymbol{U} + (1-W)\boldsymbol{V}$ *is the Dirichlet distribution* $Dir(\boldsymbol{\gamma} + \boldsymbol{\delta})$.

Lemma 3 *Let* $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \ldots, \gamma_k)$, *and let* $\beta_j = \gamma_j / \sum \gamma_j$. *Then*

$$\sum \beta_j Dir(\boldsymbol{\gamma} + \boldsymbol{e}_j) = Dir(\boldsymbol{\gamma}),$$

where $\boldsymbol{e}_j$ *is a unit basis vector with the* j*th element equal to* 1 *and the other elements equal to* 0. *The left-hand side can also be written as* $E[Dir(\boldsymbol{\gamma} + \boldsymbol{X})]$, *where* $\boldsymbol{X}$ *is a random vector that takes the value* $\boldsymbol{e}_j$ *with probability* β_j.

Lemma 4 *Let* W,U *be a pair of random variables where* W *takes values in* $[-1, 1]$ *and* U *takes values in a linear space. Suppose that* V *is a random variable taking values in the same linear space as* U *and which is independent of* (W, U) *and satisfies the distributional equation*

$$V \stackrel{st}{=} U + WV.$$

Suppose that $P(|W| = 1) \neq 1$. *Then there is only one distribution for* V *that satisfies Lemma 4.* "$\stackrel{st}{=}$" *means having the same distribution.*

Assume G is a Dirichlet process by Ferguson's definition with concentration parameter α_0 and base distribution G_0, and denote it by $DP(\alpha_0, G_0)$. We generate an infinite list $(\phi_1, \phi_2, \ldots, \phi_k, \ldots) \sim G_0$ and form $G' = \sum_{k=1}^{\infty} \pi_k \delta_k$, where $\pi_1 = \beta_1,\ \pi_k = \beta_k \prod_{j=1}^{k-1}(1 - \beta_j),\quad k = 2,\ 3, \ldots,$ and $\beta_1, \beta_2 \ldots \stackrel{iid}{\sim} \text{Beta}(1, \alpha_0)$. Our goal is to show that $G' = G$ in distribution. The strategy is to show that for all partitions $(A_1, A_2, \ldots, A_k)$ of the sample space of G_0, G' has the finite Dirichlet marginals

$$(G'(A_1), G'(A_2), \ldots, G'(A_k)) \sim Dir(\alpha_0 G_0(A_1), \alpha_0 G_0(A_2), \ldots, \alpha_0 G_0(A_k)).$$

Let f be a deterministic function that transforms the random variables $\phi_k s$ and corresponding $\beta_k s$ into the sum of infinite terms

$$G' = f(\boldsymbol{\phi}, \boldsymbol{\beta}) = \sum_{k=1}^{\infty} \pi_k \delta_k.$$

The function f can be interpreted as the function used to construct a stick-breaking process. Let $\boldsymbol{\beta}^*$ denote $(\beta_2, \beta_3, \ldots)$, with the first element of the infinite sequence removed. Now,

$$G' = \pi_1 \delta_{\phi_1} + (1 - \pi) f(\boldsymbol{\beta}^*, \boldsymbol{\phi}^*) = \pi_1 \delta_{\phi_k} + (1 - \pi_1) G'',$$

where $G' \stackrel{d}{=} G''$ ("$\stackrel{d}{=}$" means equality in distribution).

Then, we have

$$G' \stackrel{st}{=} \pi_1 \delta_{\phi_k} + (1 - \pi_1) G'.$$

If we have a partition $(A_1, A_2, \ldots, A_k)$ of Φ, the sample space of G_0 in Ferguson's definition, then

$$\begin{pmatrix} G'(A_1) \\ \vdots \\ G'(A_k) \end{pmatrix} \stackrel{st}{=} \pi_1 \begin{pmatrix} \delta_{\phi_1}(A_1) \\ \vdots \\ \delta_{\phi_1}(A_k) \end{pmatrix} + (1 - \pi_1) \begin{pmatrix} G'(A_1) \\ \vdots \\ G'(A_k) \end{pmatrix}. \tag{2.2.1}$$

Note that

$$\delta_{\phi_1}(A_k) = \begin{cases} 1 & \phi_1 \in A_1 \\ 0 & \text{otherwise.} \end{cases}$$

By Lemma 4, if there is a distribution of $(G'(A_1), \ldots, G'(A_k)$ that satisfies (2.2.1), then it must be unique. If we can prove the $Dir(\alpha_0 G_0(A_1), \alpha_0 G_0(A_2), \ldots, \alpha_0 G_0(A_k))$ satisfies the Equation (2.2.1), then it must be the unique distribution of $(G'(A_1), \ldots, G'(A_k))$,

$$(G'(A_1), \ldots, G'(A_k) \sim Dir(\alpha_0 G_0(A_1), \alpha_0 G_0(A_2), \ldots, \alpha_0 G_0(A_k)).$$

This is a proof that $G' = DP(\alpha_0, G_0)$.

Let $\boldsymbol{Z} \sim Dir(\alpha_0 G_0(A_1), \alpha_0 G_0(A_2), \ldots, \alpha_0 G_0(A_k))$ and $\boldsymbol{X} = (\delta_{\phi_1}(A_1), \ldots, \delta_{\phi_1}(A_k))$, then $\boldsymbol{X} \sim Multinomial(G_0(A_1), G_0(A_2), \ldots, G_0(A_k))$. So $P(\delta_{\phi_1}(A_k) = 1) = G_0(A_k)$. We want

to show the distribution of $\pi_1 \boldsymbol{X} + (1-\pi_1)\boldsymbol{Z}$ is still $Dir(\alpha_0 G_0(A_1), \alpha_0 G_0(A_2), \ldots, \alpha_0 G_0(A_k))$. In other words, (2.2.1) is satisfied by $Dir(\alpha_0 G_0(A_1), \alpha_0 G_0(A_2), \ldots, \alpha_0 G_0(A_k))$.
Conditioning on $\boldsymbol{X} = \boldsymbol{e}_j$

$$(\pi_1 \boldsymbol{X} + (1-\pi_1)\boldsymbol{Z} | \boldsymbol{X} = \boldsymbol{e}_j) \stackrel{d}{=} \pi_1 Dir(\boldsymbol{e}_j) + (1-\pi_1)\boldsymbol{Z}.$$

Since, $\boldsymbol{Z} \sim Dir(\alpha_0 G_0(A_1), \alpha_0 G_0(A_2), \ldots, \alpha_0 G_0(A_k))$, denoted as $Dir(\boldsymbol{\alpha})$, where $\boldsymbol{\alpha} = (\alpha_0 G_0(A_1), \alpha_0 G_0(A_2), \ldots, \alpha_0 G_0(A_k))$ by Lemma 2,

$$(\pi_1 \boldsymbol{X} + (1-\pi_1)\boldsymbol{Z} | \boldsymbol{X} = \boldsymbol{e}_j) \stackrel{d}{=} Dir(\boldsymbol{e}_j + \boldsymbol{\alpha}).$$

$P(\boldsymbol{X} = \boldsymbol{e}_j) = G_0(A_j) = \alpha_j / \sum \alpha_j$, where α_j is the jth element of $\boldsymbol{\alpha}$ and is equal to $\alpha_0 G_0(A_j)$. To obtain the distribution of $(\pi_1 \boldsymbol{X} + (1-\pi_1)\boldsymbol{Z})$, we sum over all the values of $\boldsymbol{X}$,

$$\sum_{j=1}^{K} \frac{\alpha_j}{\sum \alpha_j} Dir(\boldsymbol{e}_j + \boldsymbol{\alpha}) = Dir(\boldsymbol{\alpha}), \text{by Lemma 3.}$$

Thus, $\pi_1 \boldsymbol{X} + (1-\pi_1)\boldsymbol{Z} \sim Dir(\boldsymbol{\alpha})$. The distribution $Dir(\boldsymbol{\alpha})$ satisfies (2.2.1) and it is the distribution of $(G'(A_1), \ldots, G'(A_k))$. Hence, $G' \stackrel{d}{=} G$.

2.2.2 The Chinese Restaurant Process

Imagine a restaurant with countably infinitely many tables. Customers walk in and sit down at some table. The tables are chosen according to the following random process:

1. The first customer always chooses an unoccupied table.

2. The nth customer chooses an unoccupied table with probability $\frac{\alpha}{n-1+\alpha}$, and an occupied table with probability $\frac{c}{n-1+\alpha}$ where c is the number of people sitting at the table.

After infinitely many customers have entered the restaurant, the stochastic process underlying the CRP is a Dirichlet process.

2.2.3 The Pólya Urn Scheme

Another way to visualize the Dirichlet process is called the modified Pólya urn scheme. We start with an empty urn. Then we proceed as follows:

1. Generate a random number from $U(0,1)$. If it is less than $\frac{\alpha}{\alpha+n}$, where n is the number of balls in the urn, then add a ball with new color into the urn.

2. Otherwise, pick out a ball from the urn and add it back with a new ball of the same color.

After infinitely many balls have been put into the urn, the distribution over infinitely many colors is the same with the distribution over tables in the Chinese restaurant process.

2.3 Mixture of Dirichlet Process

The model

$$f(x) = \sum_{j=1}^{K} w_j f_j(x \mid \boldsymbol{\phi}_j) \tag{2.3.1}$$

is a mixture with K components, where f_j is the pdf of a distribution with parameters $\boldsymbol{\phi}_j$. Augmenting the data with random variables $Z_i, Z_i \in \{1, \ldots, K\}$, indicating which component gives rise to $\boldsymbol{X}_i$, we can write (2.3.1) in a hierarchical form as follows:

$$\begin{aligned}
(\boldsymbol{X}_i \mid Z_i, \boldsymbol{\Phi}) &\sim f(x_i \mid \boldsymbol{\phi}_{z_i}) \\
Z_i \mid \boldsymbol{w} &\sim \text{Discrete}(w_1, \ldots, w_K),
\end{aligned}$$

where $\boldsymbol{\Phi} = (\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_K)$ and $\boldsymbol{w} = (w_1, \ldots, w_K)$.
The priors on $\boldsymbol{\Phi}$ and $\boldsymbol{w}$ are

$$\begin{aligned}
\boldsymbol{\phi}_j &\sim G_0 \\
\boldsymbol{w} &\sim \text{Dir}(\gamma/K, \ldots, \gamma/K).
\end{aligned}$$

We show that this is a mixture of Dirichlet process (MDP) when $K \to \infty$.

$$P(z_1, \ldots, z_n \mid \boldsymbol{w}) = \prod_{j=1}^{K} w_j^{n_j},$$

where n_j is the number of observations assigned to component j.

$$\begin{aligned}
P(\boldsymbol{Z}) &= \int P(z_1, \ldots, z_n \mid \boldsymbol{w}) \times P(\boldsymbol{w}) d\boldsymbol{w} \\
&= \int \prod_{j=1}^{K} w_j^{n_j} \times \frac{\Gamma(\gamma)}{[\Gamma(\frac{\gamma}{K})]^K} w_1^{\frac{\gamma}{K}-1} \cdots w_K^{\frac{\gamma}{K}-1} d\boldsymbol{w} \\
&= \frac{\Gamma(\gamma)}{\Gamma(n+\gamma)} \prod_{j+1}^{K} \frac{\Gamma(n_j + \frac{\gamma}{K})}{\Gamma(\frac{\gamma}{K})}
\end{aligned}$$

$$P(\boldsymbol{w} \mid \boldsymbol{Z}) = \frac{P(\boldsymbol{Z} \mid \boldsymbol{w}) \times P(\boldsymbol{w})}{P(\boldsymbol{Z})} = \mathrm{Dir}(\gamma/K + n_{i1}, \ldots, \gamma/K + n_{ik}),$$

where n_{ik} is the number of observations previously (before the ith) assigned to component k.

$$\begin{aligned}
P(z_i = k \mid z_1, \ldots, z_{i-1}) &= \int P(z_i = k \mid \boldsymbol{w}) \times P(\boldsymbol{w} \mid z_1, \ldots, z_{i-1}) d\boldsymbol{w} \\
&= \frac{\prod_{j=1}^{K} \Gamma(\gamma/K + n_{ij})}{\Gamma(\gamma + i)} \times \frac{\Gamma(\gamma + i - 1)}{\prod_{j=1}^{K} \Gamma(\gamma/K + n_{(i-1)j})} \times \\
&\quad \int \frac{\Gamma(\gamma + i)}{\prod_{j=1}^{K} \Gamma(\gamma/K + n_{ij})} \prod_{j=1}^{K} w_j^{\gamma/K + n_{ij} - 1} d\boldsymbol{w} \\
&= \frac{n_{ik} + \gamma/K}{i - 1 + \gamma}.
\end{aligned}$$

By taking the limit $K \to \infty$, for components with $n_{ik} > 0$,

$$P(z_i = k \mid z_1, \ldots, z_{i-1}) = \frac{n_{ik}}{i - 1 + \gamma},$$

which is identical to the probability of the ith customer sitting at an occupied table in the Chinese restaurant with n_{ik} customers already sitting at this table.

The probability of any particular unoccupied cluster approaches zero as $K \to \infty$. However, the total probability assigned to all unoccupied clusters is still positive:

$$
\begin{aligned}
P(z_i \neq z_j \ \forall\, j < i \mid z_1, \ldots, z_{i-1}) &= 1 - \sum_{k: n_{ik} > 0} P(z_i = k \mid z_1, \ldots, z_{i-1}) \\
&= 1 - \sum_{k: n_{ik} > 0} \frac{n_{ik}}{i - 1 + \gamma} \\
&= \frac{\gamma}{i - 1 + \gamma}.
\end{aligned}
$$

Applying this result, the conditional probability of $(\boldsymbol{\theta}_i \mid \boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_{i-1})$ where $\boldsymbol{\theta}_i = \boldsymbol{\phi}_{z_i}$ becomes

$$
P(\boldsymbol{\theta}_i \mid \boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_{i-1}) = \begin{cases} \frac{1}{i-1+\gamma} \sum_{j=1}^{i-1} \delta(\boldsymbol{\theta}_j = \boldsymbol{\theta}_i), & \text{if } \exists\, \boldsymbol{\theta}_j = \boldsymbol{\theta}_i \text{ for } j = 1, \ldots, i-1 \\ \frac{\gamma}{i-1+\gamma} G_0(\boldsymbol{\theta}_i), & \text{otherwise,} \end{cases} \tag{2.3.2}
$$

where $\delta(\boldsymbol{\theta}_j = \boldsymbol{\theta}_i) = 1$ when $\boldsymbol{\theta}_j = \boldsymbol{\theta}_i$ and zero otherwise.

The parameters $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_{i-1}$ may not be unique since $\boldsymbol{\theta}_i = \boldsymbol{\phi}_{z_i}$. Since the observations are assumed to be exchangeable, we can regard any observation i as the last one and write the conditional probability of $\boldsymbol{\theta}_i$ given other $\boldsymbol{\theta}_j$ for $j \neq i$ as in (2.3.2).

Chapter 3

Modeling Dependence Structure via Elliptical Copulas

3.1 Statistical Modeling Using an Elliptical Copula

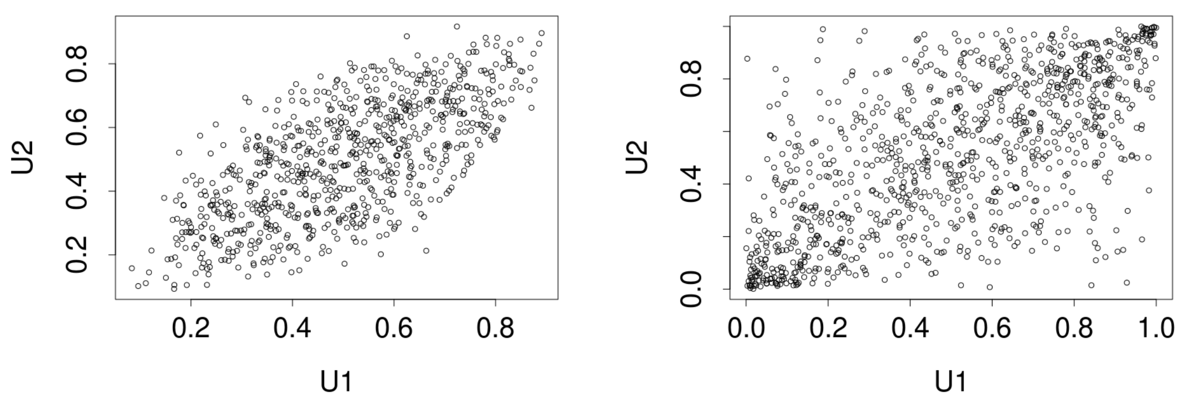

Figure 3.1: Examples of Elliptical Copulas: Logistic copula (left) and Student's t copula (right)

From Section 1.3, we know the PDF of an elliptical distribution is

$$f_E(\mathbf{x}; \Omega, g) = c_p |\Omega|^{-1/2} g((\mathbf{x})' \Omega^{-1} (\mathbf{x})).$$

By using the inversion method of Section 1.2, the density of an elliptical copula can be written as

$$c_E(u_1, \dots, u_p) = \frac{f_E(Q^{-1}(u_1) \dots Q^{-1}(u_p))}{q(Q^{-1}(u_1)) \dots q(Q^{-1}(u_p))}, \tag{3.1.1}$$

where

- f_E is the joint pdf of the elliptical distribution;
- Q is the marginal cdf of the elliptical distribution;
- q is the marginal pdf of the elliptical distribution.

Note that elliptical distributions have identical margins. The copula in Figure 1.3 of Chapter 1 is a Gaussian copula, which is the most common elliptical copula. Figure 3.1 shows two samples from other elliptical copulas.

3.2 Modeling Q and q

The PDF c_E of Equation (3.1) contains f_E, Q and q. From Section 1.4, the f_E can be expressed based on the stochastic form (1.4.1)

$$f(r; A) = \frac{\Gamma(p/2)}{2\pi^{p/2}} |A|^{-1} r^{1-p} h(r),$$

where $r = \sqrt{\mathbf{x}'\Omega^{-1}\mathbf{x}}$ and A is the Cholesky factor of the correlation matrix.
We propose the following infinite gamma mixture to express $h(r)$,

$$h(r) = \sum_{j=1}^{\infty} w_j f_g(r \mid \alpha_j, \beta_j), \tag{3.2.1}$$

where $f_g(r \mid \alpha_j, \beta_j)$ is the pdf of a gamma distribution with shape and rate parameters α_j, β_j respectively and w_j are mixing weights satisfying $\sum_{j=1}^{\infty} w_j = 1$ and $0 \leq w_j \leq 1$.
According to Gómez et al. (2003), the relationship between the generator function of an elliptical distribution and the PDF $h(r)$ is

$$g(t) = \frac{\Gamma(p/2)}{2\pi^{p/2}} t^{(1-p)/2} h(t^{1/2}). \tag{3.2.2}$$

According to Genest et al. (2007), the marginal PDF and CDF can be derived from the generator g as follows.

$$
\begin{aligned}
Q(z) &= 1/2 + \frac{\pi^{(p-1)/2}}{\Gamma((p-1)/2)} \int_0^z \int_{u^2}^{\infty} (y-u^2)^{(p-1)/2-1} g(y)dydu, \\
q(z) &= \frac{\pi^{(p-1)/2}}{\Gamma(\frac{p-1}{2})} \int_{z^2}^{\infty} (y-z^2)^{(p-1)/2-1} g(y)dy.
\end{aligned}
$$

3.3 The Case $p = 3$

When $p = 3$, the function Q is much simpler, since $(y-u^2)^{(p-1)/2-1}$ is equal to 1. Then we have

$$
\begin{aligned}
Q(z) &= 1/2 + \int_0^z \int_{u^2}^{\infty} g(y)dydu, \\
q(z) &= \pi \int_{z^2}^{\infty} g(y)dy.
\end{aligned}
$$

$\int_{u^2}^{\infty} g(y)dy$ can be expanded as

$$
\begin{aligned}
\int_{u^2}^{\infty} g(y)dy &= \frac{\Gamma(3/2)}{2\pi^{3/2}} \int_{u^2}^{\infty} y^{-1} h(y^{1/2})dy \\
&= \frac{\Gamma(3/2)}{2\pi^{3/2}} \int_{u^2}^{\infty} y^{-1} \sum w_j f_g(y^{1/2} \mid \alpha_j, \beta_j)dy \\
&= \frac{\Gamma(3/2)}{2\pi^{3/2}} \sum w_j \int_{u^2}^{\infty} y^{-1} f_g(y^{1/2} \mid \alpha_j, \beta_j)dy \qquad (3.3.1) \\
&= \frac{\Gamma(3/2)}{2\pi^{3/2}} \sum w_j \int_{u^2}^{\infty} y^{-1} \frac{\beta_j^{\alpha_j}}{\Gamma(\alpha_j)} y^{(\alpha_j-1)/2} e^{-\beta_j y^{1/2}} dy \\
&= \frac{\Gamma(3/2)}{2\pi^{3/2}} \sum w_j \frac{\beta_j^{\alpha_j}}{\Gamma(\alpha_j)} \int_{u^2}^{\infty} y^{(\alpha_j-3)/2} e^{-\beta_j y^{1/2}} dy
\end{aligned}
$$

In (3.3), $y^{(\alpha_j-3)/2} e^{-\beta_j y^{1/2}}$ looks like the kernel of a Gamma distribution with respect to $y^{1/2}$. Making the substitution $t = y^{1/2}$,

$$
\begin{aligned}
\int_{u^2}^{\infty} g(y)dy &= \frac{\Gamma(3/2)}{2\pi^{3/2}} \sum w_j \frac{\beta_j^{\alpha_j}}{\Gamma(\alpha_j)} \int_{u^2}^{\infty} y^{(\alpha_j-3)/2} e^{-\beta_j y^{1/2}} dy \\
&= \frac{\Gamma(3/2)}{2\pi^{3/2}} \sum w_j \frac{\beta_j^{\alpha_j}}{\Gamma(\alpha_j)} \int_{|u|}^{\infty} t^{(\alpha_j-3)} e^{-\beta_j t} 2t dt \\
&= \frac{\Gamma(3/2)}{\pi^{3/2}} \sum w_j \frac{\beta_j^{\alpha_j}}{\Gamma(\alpha_j)} \int_{|u|}^{\infty} t^{(\alpha_j-2)} e^{-\beta_j t} dt \\
&= \frac{\Gamma(3/2)}{\pi^{3/2}} \sum w_j \frac{\beta_j^{\alpha_j}}{\Gamma(\alpha_j)} \frac{\alpha_j - 1)}{\beta_j^{\alpha_j-1}} \\
&\int_{|u|}^{\infty} \frac{\beta_j^{\alpha_j-1}}{\Gamma(\alpha_j-1)} t^{(\alpha_j-2} e^{-\beta_j t} dt \\
&= \frac{\Gamma(3/2)}{\pi^{3/2}} \sum w_j \frac{\beta_j}{\alpha_j - 1} \int_{|u|}^{\infty} \frac{\beta_j^{\alpha_j-1}}{\Gamma(\alpha_j-1)} t^{(\alpha_j-1-1)} e^{-\beta_j t} dt \\
&= \frac{\Gamma(3/2)}{\pi^{3/2}} \sum w_j \frac{\beta_j}{\alpha_j - 1} [1 - F(|u|)]
\end{aligned}
$$

where F is the CDF of Gamma$(\alpha_j - 1, \beta_j)$.

Therefore, when $p = 3$, $Q$and q can be reduced to

$$
\begin{aligned}
Q(z) &= 1/2 + \pi^{-1/2}\Gamma(3/2) \int_0^z \sum w_j \frac{\beta_j}{\alpha_j - 1} [1 - F(|u|)] du \\
q(z) &= \pi^{-1/2}\Gamma(3/2) \sum w_j \frac{\beta_j}{\alpha_j - 1} [1 - F(|u|)]
\end{aligned}
$$

The CDF Q can be further simplified by using the incomplete gamma function.

3.4 Incomplete Gamma Function

The Gamma function is defined as

$$
\Gamma(s) = \int_0^{\infty} t^{s-1} e^{-t} dt.
$$

The upper incomplete Gamma function is defined as

$$\Gamma(s,x) = \int_x^\infty t^{s-1}e^{-t}dt.$$

The lower incomplete Gamma function is defined as

$$\gamma(s,x) = \int_0^x t^{s-1}e^{-t}dt.$$

The CDF of a Gamma distribution with shape α and rate β is $\frac{1}{\Gamma(\alpha)}\gamma(\alpha,\beta x)$. when $z > 0$,

$$\begin{aligned}
\int_0^z F(|u|)du &= \int_0^z F(u)du \\
&= \int_0^z \frac{1}{\Gamma(\alpha_j)}\gamma(\alpha_j,\beta_j u)du \\
&= \frac{1}{\Gamma(\alpha_j)}\int_0^z [\Gamma(\alpha_j) - \Gamma(\alpha_j,\beta_j u)]du \\
&= z - \frac{1}{\Gamma(\alpha_j)}\int_0^z [\Gamma(\alpha_j,\beta_j u)du \\
&= z - \frac{1}{\Gamma(\alpha_j)}\left[\frac{\Gamma(1+\alpha_j) + \beta_j z\Gamma(\alpha_j,\beta_j z) - \Gamma(1+\alpha_j,\beta_j z)}{\beta_j}\right]
\end{aligned}$$

when $z < 0$, according to Mathematica,

$$\begin{aligned}
\int_0^z F(|u|)du &= \int_0^z F(-u)du \\
&= \int_0^z \frac{1}{\Gamma(\alpha_j)}\gamma(\alpha_j,-\beta_j u)du \\
&= \frac{1}{\Gamma(\alpha_j)}\int_0^z [\Gamma(\alpha_j) - \Gamma(\alpha_j,-\beta_j u)]du \\
&= z - \frac{1}{\Gamma(\alpha_j)}\int_0^z [\Gamma(\alpha_j,-\beta_j u)du \\
&= z - \frac{1}{\Gamma(\alpha_j)}\left[\frac{-\Gamma(1+\alpha_j) + \beta_j z\Gamma(\alpha_j,-\beta_j z) + \Gamma(1+\alpha_j,-\beta_j z)}{\beta_j}\right]
\end{aligned}$$

They can be combined as

$$\int_0^z F(|u|)du = z - \frac{1}{\Gamma(\alpha_j)}\left[\frac{z/|z| * \Gamma(1+\alpha_j) + \beta_j z\Gamma(\alpha_j,\beta_j|z|) - z/|z| * \Gamma(1+\alpha_j,\beta_j|z|)}{\beta_j}\right]$$

Q can be written as

$$
\begin{aligned}
Q(z) &= 1/2 + \pi^{-1/2}\Gamma(3/2)\int_0^z \sum w_j \frac{\beta_j}{\alpha_j - 1}\big[1 - F(|u|)\big]du \\
&= 1/2 + \pi^{-1/2}\Gamma(3/2)\int_0^z \left[\sum w_j \frac{\beta_j}{\alpha_j - 1} - \sum w_j \frac{\beta_j}{\alpha_j - 1} F(|u|)\right]du \\
&= 1/2 + \pi^{-1/2}\Gamma(3/2)\left[\int_0^z \sum w_j \frac{\beta_j}{\alpha_j - 1} du - \int_0^z \sum w_j \frac{\beta_j}{\alpha_j - 1} F(|u|)du\right] \\
&= 1/2 + \pi^{-1/2}\Gamma(3/2)\left[\sum w_j \frac{\beta_j z}{\alpha_j - 1} - \sum \left[w_j \frac{\beta_j}{\alpha_j - 1}\int_0^z F(|u|)du\right]\right] \\
&= 1/2 + \\
&\pi^{-1/2}\Gamma(3/2)\sum \frac{w_j\beta_j}{(\alpha_j - 1)\Gamma(\alpha_j - 1)}\left[\frac{z/|z| * \Gamma(\alpha_j) + \beta_j z\Gamma(\alpha_j - 1, \beta_j|z|) - z/|z| * \Gamma(\alpha_j, \beta_j|z|)}{\beta_j}\right]\Bigg] \\
&= 1/2 + \\
&\pi^{-1/2}\Gamma(3/2)\sum \frac{w_j}{(\alpha_j - 1)\Gamma(\alpha_j - 1)}\Big[z/|z| * \Gamma(\alpha_j) + \beta_j z\Gamma(\alpha_j - 1, \beta_j|z|) - z/|z| * \Gamma(\alpha_j, \beta_j|z|)\Big]\Bigg]
\end{aligned}
$$

3.5 Elliptical Copula When $p = 3$

An elliptical copula can be modeled as

$$
c_E(u_1, \ldots, u_p) = \frac{f_E(Q^{-1}(u_1)\ldots Q^{-1}(u_p))}{q(Q^{-1}(u_1))\ldots q(Q^{-1}(u_p))}, \tag{3.5.1}
$$

where

$$
f_E(\mathbf{x}; \Omega, g) = c_p|\Omega|^{-1/2} g((\mathbf{x})'\Omega^{-1}(\mathbf{x})),
$$

$$
\begin{aligned}
&Q = 1/2 + \pi^{-1/2}\Gamma(3/2)\sum \frac{w_j}{(\alpha_j - 1)\Gamma(\alpha_j - 1)} \\
&\Big[z/|z| * \Gamma(\alpha_j) + \beta_j z\Gamma(\alpha_j - 1, \beta_j|z|) - z/|z| * \Gamma(\alpha_j, \beta_j|z|)\Big]\Bigg],
\end{aligned}
$$

$$
q(z) = \pi^{-1/2}\Gamma(3/2)\sum w_j \frac{\beta_j}{\alpha_j - 1}[1 - F(|u|)].
$$

Chapter 4

Methodology

Assume we are given a p-dimensional data set of size n as follows

$$\begin{bmatrix} x_{11}, x_{12}, \ldots, x_{1p} \\ x_{21}, x_{22}, \ldots, x_{2p} \\ \ldots\ldots \\ x_{n1}, x_{n2}, \ldots, x_{np} \end{bmatrix} \tag{4.0.1}$$

which is drawn from an unknown meta-elliptical distribution.

To estimate the underlying meta-elliptical distribution, we need to estimate its marginal distributions and its copula. Marginal distributions are easy to estimate since they are univariate. They can be estimated with a normal mixture, which has done extensively in the literature (Behboodian (1970)). Therefore, we focus on the estimation of elliptical copulas. Our goal is to estimate an elliptical copula without assuming a specific type, such as Gaussian, Student's t, logistic, etc.

To obtain a sample from an elliptical copula, we apply the probability integral transform to the sample (4.0.1), resulting in

$$\begin{bmatrix} u_{11}, u_{12}, \ldots, u_{1p} \\ u_{21}, u_{22}, \ldots, u_{2p} \\ \ldots\ldots \\ u_{n1}, u_{n2}, \ldots, u_{np} \end{bmatrix}. \tag{4.0.2}$$

According to the elliptical copula model in Equation (3.1), the likelihood of (4.0.2) can be written as

$$
\begin{aligned}
L(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{w}, A; \mathbf{U}) &= \prod_{i=1}^{n} c(u_{i1}, u_{i2}, \ldots, u_{ip}) \\
&= |A|^{-n} \prod_{i=1}^{n} \frac{(r_i^2)^{(1-p)/2} h((r_i^2)^{1/2})}{\prod_{k=1}^{p} q(Q^{-1}(u_{ip}))} \\
&= |A|^{-n} \prod_{i=1}^{n} \frac{{r_i}^{1-p} h(r_i)}{\prod_{k=1}^{p} q(Q^{-1}(u_{ip}))} \\
&= |A|^{-n} \prod_{i=1}^{n} \frac{{r_i}^{1-p} \sum_{j=1}^{J} (w_j f(r_i|\alpha_j, \beta_j))}{\prod_{k=1}^{p} q(Q^{-1}(u_{ip}))},
\end{aligned}
\tag{4.0.3}
$$

where

$$r_i^2 = \boldsymbol{x}' \Omega^{-1} \boldsymbol{x},$$

and

$$\boldsymbol{x} = (Q^{-1}(u_{i1}), Q^{-1}(u_{i2}), \ldots, Q^{-1}(u_{ip})).$$

We augment the data with latent indicators z_{ij}, such that

$$z_{ij} = \begin{cases} 1 & \text{if } r_i \text{ is from component } j \\ 0 & \text{otherwise.} \end{cases}$$

The augmented likelihood

$$L(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{w}, A; \mathbf{U}, Z) = |A|^{-n} \prod_{i=1}^{n} \frac{{r_i}^{(1-p)} \prod_{j=1}^{J} (w_j f(r_i|\alpha_j, \beta_j)^{z_{ij}})}{\prod_{k=1}^{p} q(Q^{-1}(u_{ip}))}. \tag{4.0.4}$$

4.1 The Case $p = 3$

Now, we focus on the three-dimensional case, since the evaluation of Q and q is easier when $p = 3$, according to Section 3.3.

4.1.1 Prior distributions

The following priors refer to the mixture model (3.2.1).

1. $\alpha_1,\ \alpha_2,\ \cdots \stackrel{iid}{\sim} \text{Pareto}(1, c)$, where c is fixed, with pdf

$$\frac{c}{\alpha_j^{c+1}}, \quad \alpha_j > 1.$$

2. $\beta_1,\ \beta_2,\ \cdots \stackrel{iid}{\sim} \text{Gamma}(a, b)$, where a (shape) and b (rate) are fixed.

3. To model the infinite mixture, we use the stick-breaking representation of the Dirichlet process (Sethuraman (1994)), which means that the w_j satisfy

$$w_1 = v_1,\ w_j = v_j \prod_{l=1}^{j-1} (1 - v_l), \quad j = 2,\ 3, \ldots, \tag{4.1.1}$$

   where $v_1, v_2 \ldots \stackrel{iid}{\sim} \text{Beta}(1, \gamma)$. In practice, we truncate the number of components and set it to k, so v_k is set to 1, such that $\sum_{j=1}^{k} w_j = 1$. Another possibility would be to use the slice sampler (Kalli et al. (2011)).

4. $\gamma \sim \text{Gamma}(\eta_1, \eta_2)$, where η_1 and η_2 are fixed hyperparameters.

5. When $p = 2$, only one correlation parameter has to be estimated, and we place the $\text{U}(-1, 1)$ prior on it. But we are dealing with case $p = 3$. When $p > 2$ the positive definiteness of Ω makes it difficult to work directly with its individual entries. We discuss the prior of Ω in Subsection 4.1.2.

4.1.2 The Prior for Ω

To maintain the positive definiteness of Ω, we parameterize the Cholesky factor of Ω using hyperspherical coordinates (Rapisarda et al. (2007)), described also in Pourahmadi and Wang (2015). The Cholesky factor is expressed as:

$$A = \begin{bmatrix} 1 & c_{12} & c_{13} \\ 0 & s_{12} & c_{23}s_{13} \\ 0 & 0 & s_{23}s_{13} \end{bmatrix},$$

where $c_{ij} = \cos(\theta_{ij})$, $s_{ij} = \sin(\theta_{ij})$, and the θ_{ij}'s are some angles. The matrix A is unique if its diagonal entries are positive, or equivalently if the θ_{ij}'s are in the interval $(0, \pi)$. Therefore, we assume a priori that $\theta_{ij} \stackrel{iid}{\sim} \mathrm{U}(0, \pi)$.

4.1.3 Posterior Distribution

Let $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_k)'$, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_k)'$, $\boldsymbol{v} = (v_1, \ldots, v_k)'$, $\boldsymbol{\theta} = (\theta_{12}, \theta_{13}, \theta_{23})'$, then the joint posterior distribution is

$$\begin{aligned} P(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{V}, \gamma, \rho | \boldsymbol{X}, Z) \propto |A|^{-n} \prod_{i=1}^{n} & \frac{r_i^{(-2)} \prod_{j=1}^{J} (w_j f(r_i | \alpha_j, \beta_j)^{z_{ij}})}{q(Q^{-1}(u_{i1})) q(Q^{-1}(u_{i2})) q(Q^{-1}(u_{i3}))} \\ & \times \prod_{j=1}^{k} \frac{c}{\alpha_j^{c+1}} \\ & \times \prod_{j=1}^{k} \frac{b^a}{\Gamma(a)} \beta_j^{a-1} e^{-b\beta_j} \\ & \times \prod_{j=1}^{k-1} \gamma (1 - v_j)^{\gamma - 1} \\ & \times \gamma^{\eta_1 - 1} e^{-\eta_2 \gamma} \\ & \times \prod_{i=2}^{p} \prod_{j=1}^{i-1} \mathbb{1}_{(0 \le \theta_{ij} \le \pi)} \\ & \times \left| \frac{\partial \boldsymbol{\rho}}{\partial \boldsymbol{\theta}} \right| \end{aligned} \quad . \tag{4.1.2}$$

where $\boldsymbol{\rho} = (\rho_{12}, \rho_{13}, \rho_{23})'$ and $\left| \frac{\partial \boldsymbol{\rho}}{\partial \boldsymbol{\theta}} \right|$ is the Jacobian of the transformation from $\boldsymbol{\rho}$ to $\boldsymbol{\theta}$ (Pourahmadi and Wang (2015)). The Jacobian is given by

$$\left| \frac{\partial \boldsymbol{\rho}}{\partial \boldsymbol{\theta}} \right| = \prod_{i=2}^{3} \prod_{j=1}^{i-1} \frac{\partial \rho_{ij}}{\partial \theta_{ij}} = -\prod_{j=1}^{2} \left(\prod_{i=j+1}^{3} s_{ij} \right)^{3-j} . \tag{4.1.3}$$

4.1.4 Conditional Posterior Distributions

In the likelihood (4.0.3), Q and q are functions of the parameters $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, as well as the weights w, according to Equation (3.2.1). Therefore, the formulas for conditional posterior distributions all depend on these parameters and are not standard distribution that are easy to sample from.

The conditional posterior distributions are as follows.

$$P(\alpha_j|\beta_j,\rho,\boldsymbol{r},\boldsymbol{z}_j) \propto |A|^{-n} \prod_{i=1}^{n} \frac{{r_i}^{(1-p)} \prod_{j=1}^{J}(w_j f(r_i|\alpha_j,\beta_j)^{z_{ij}})}{\prod_{k=1}^{p} q(Q^{-1}(u_{ip}))} \times \frac{1}{\alpha_j^{c+1}}. \tag{4.1.4}$$

$$P(\beta_j|\alpha_j,\rho,\boldsymbol{r},\boldsymbol{z}_j) \propto |A|^{-n} \prod_{i=1}^{n} \frac{{r_i}^{(1-p)} \prod_{j=1}^{J}(w_j f(r_i|\alpha_j,\beta_j)^{z_{ij}})}{\prod_{k=1}^{p} q(Q^{-1}(u_{ip}))} \times \frac{b^a}{\Gamma(a)} \beta_j^{a-1} e^{-b\beta_j}. \tag{4.1.5}$$

$$P(v_j|Z) \propto |A|^{-n} \prod_{i=1}^{n} \frac{{r_i}^{(1-p)} \prod_{j=1}^{J}(w_j f(r_i|\alpha_j,\beta_j)^{z_{ij}})}{\prod_{k=1}^{p} q(Q^{-1}(u_{ip}))} \times (1-v_j)^{\gamma-1}. \tag{4.1.6}$$

$$P(\gamma|\boldsymbol{v}) \sim \text{Gamma}(k+\eta_1-1, \eta_2 - \sum_{j=1}^{k-1} \log(1-v_j)). \tag{4.1.7}$$

$$P(\boldsymbol{\theta}|\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{r},Z,A) \propto |A|^{-n} \prod_{i=1}^{n} \frac{{r_i}^{(1-p)} \prod_{j=1}^{J}(w_j f(r_i|\alpha_j,\beta_j)^{z_{ij}})}{\prod_{k=1}^{p} q(Q^{-1}(u_{ip}))} \left|\frac{\partial \boldsymbol{\rho}}{\partial \boldsymbol{\theta}}\right|. \tag{4.1.8}$$

4.1.5 Sampling Scheme

1. Calculate $\boldsymbol{x}_i = (Q^{-1}(u_{i1}), Q^{-1}(u_{i2}), Q^{-1}(u_{i3}))$ and $r_i^2 = \boldsymbol{x}_i'\Omega^{-1}\boldsymbol{x}_i$ using the current estimates.

2. The conditional distribution of α_j (Equation (4.1.4)) is not a standard distribution, so we use a Metropolis step to sample from it. The proposal distribution is normal with mean $\alpha_j^{(c)}$ and variance δ_1, where δ_1 is a fixed tuning parameter and $\alpha_j^{(c)}$ is the current value of α_j. The proposed value $\alpha_j^{(p)}$ is accepted with probability $\min\left\{1, \frac{p(\alpha_j^{(p)}|\beta_j^{(c)},\Omega^{(c)},\boldsymbol{r}^{(c)},\boldsymbol{z}_j^{(c)})}{p(\alpha_j^{(c)}|\beta_j^{(c)},\Omega^{(c)},\boldsymbol{r}^{(c)},\boldsymbol{z}_j^{(c)})}\right\}$, provided $\alpha_j^{(p)} > 1$. Update $\boldsymbol{x}$ and $\boldsymbol{r}$ for all i.

3. Use a Metropolis step to sample $\boldsymbol{\beta}$ from (4.1.5). The proposal distribution is nor-

mal with mean $\beta_j^{(c)}$ and variance δ_1, where δ_1 is a fixed tuning parameter and $\beta_j^{(c)}$ is the current value of β_j. The proposed value $\beta_j^{(p)}$ is accepted with probability $\min\left\{1, \frac{p(\beta_j^{(p)}|\alpha_j^{(c)},\Omega^{(c)},\boldsymbol{r}^{(c)},\boldsymbol{z}_j^{(c)})}{p(\beta_j^{(c)}|\alpha_j^{(c)},\Omega^{(c)},\boldsymbol{r}^{(c)},\boldsymbol{z}_j^{(c)})}\right\}$, provided $\beta_j^{(p)} > 0$. Update $\boldsymbol{x}$ and $\boldsymbol{r}$ for all i.

4. Use a Metropolis step to sample v_j from (4.1.6), $j = 1, \ldots, k-1$, and calculate w_j's from the v_j's, according to Equation (4.1.1). Update $\boldsymbol{x}$ and $\boldsymbol{r}$.

5. Sample γ from Gamma$(k + \eta_1 - 1, \eta_2 - \sum_{j=1}^{k-1} \log(1 - v_j))$. Update $\boldsymbol{x}$ and $\boldsymbol{r}$ for all i.

6. Since the pdf (4.1.8) is not standard, we sample the θ_{ij} via independent Metropolis steps. The proposal distribution for θ_{ij} is $\mathrm{U}(\theta_{ij}^{(c)} - \delta_3, \theta_{ij}^{(c)} + \delta_3)$, where $\theta_{ij}^{(c)}$ is the current value of θ_{ij}. The proposed value $\theta_{ij}^{(p)}$ is accepted with probability $\min\left\{1, \frac{p(\theta_{ij}^{(p)}|\boldsymbol{\alpha}^{(c)},\boldsymbol{\beta}^{(c)},\boldsymbol{r}^{(c)},Z,A)}{p(\theta_{ij}^{(c)}|\boldsymbol{\alpha}^{(c)},\boldsymbol{\beta}^{(c)},\boldsymbol{r}^{(c)},Z,A)}\right\}$, where $p(\theta_{ij}|\boldsymbol{\alpha}^{(c)}, \boldsymbol{\beta}^{(c)}, \boldsymbol{r}^{(c)}, Z, A)$ is given in Equation (4.1.8). Update $\boldsymbol{x}$ and $\boldsymbol{r}$ for all i.

4.1.6 Latent Variables

The conditional distribution of the latent variables

$$
\begin{aligned}
P(z_{ij} = 1|\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{w}, \boldsymbol{U}) &= \frac{P(r_i|z_{ij} = 1)P(z_{ij} = 1)}{P(r_i)}, \\
&= \frac{f(r_i|\alpha_j, \beta_j)w_j}{\sum_{j=1}^{J}(f(r_i|\alpha_j, \beta_j)w_j)},
\end{aligned}
\tag{4.1.9}
$$

for $i = 1, \ldots, n$ and $j = 1, \ldots, J$.

Recompute latent variables at each sweep of MCMC, for $i = 1, \ldots, n$ and $j = 1, \ldots, K$.

Chapter 5

Results Based on Simulated Data

5.1 Simulating a Sample from Copulas

As is shown in the motivating example, a sample from a copula can be obtained by two steps. Assume we need to get a sample from the underlying copula C of a p dimensional multivariate distribution H with margins denoted by F_i, $i = 1, \ldots, p$.

Algorithm 1:

- Draw a sample $\boldsymbol{X} = (X_1, \ldots, X_p)$ from H;
- Apply PIT (Probability Integral Transformation)

$$U_i = F_i(X_i), i = 1, \ldots, p$$

.

Then $\boldsymbol{U} = (U_1, \ldots, U_p)$ is a sample from the copula C.

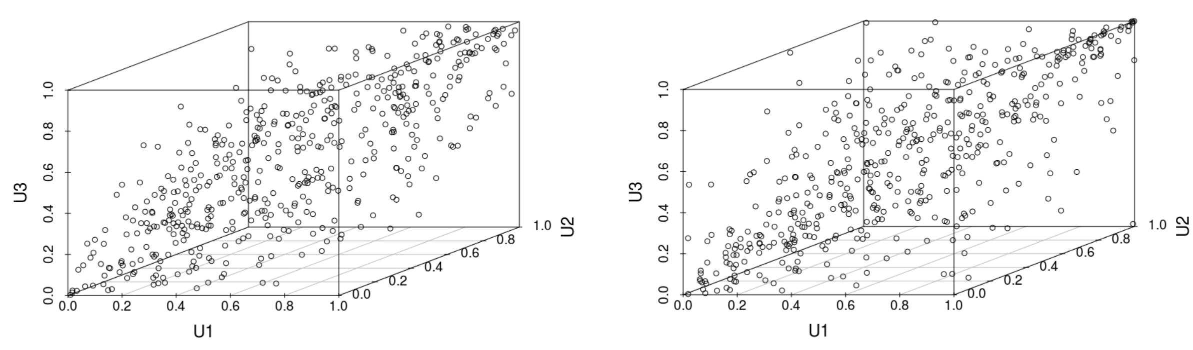

Figure 5.1: A Sample from a Gaussian Copula (left) and a Student's t Copula (right)

Now, assume we have a multivariate Gaussian with $\boldsymbol{\mu} = \mathbf{0}$ and the correlation matrix

$$\Sigma = \begin{bmatrix} 1 & 0.5 & 0.6 \\ 0.5 & 1 & 0.7 \\ 0.6 & 0.7 & 1 \end{bmatrix}. \tag{5.1.1}$$

We draw a sample from its copula, which is a Gaussian copula. According to Algorithm 1, we draw a sample $\boldsymbol{Z} = (Z_1, \ldots, Z_p)$ from $N(\boldsymbol{\mu}, \Sigma)$ first and then transform it into $\boldsymbol{U}$ by applying the formula $U_i = \Phi(Z_i), i = 1, \ldots, p$, where Φ is the standard normal CDF. Figure 5.1 is the scatter plot of simulated data from the Gaussian copula and from the copula of Student's t with degree of freedom equal to 3.

5.2 Results

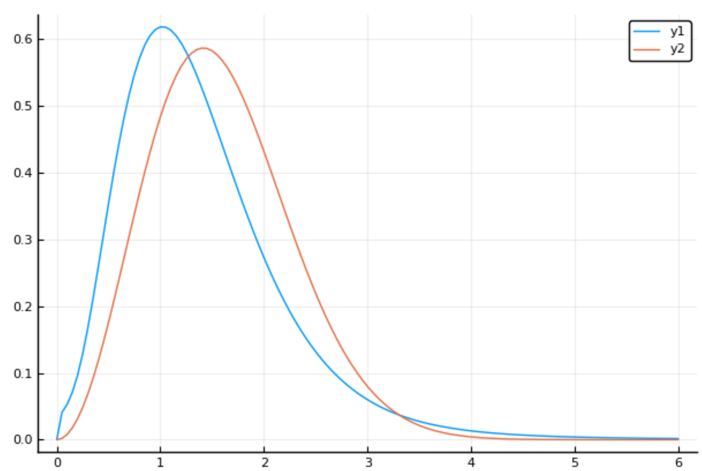

Figure 5.2: Estimated and theoretical $h(r)$ corresponding to a trivariate Gaussian copula. Y_1 is the estimated one and Y_2 is the theoretical one..

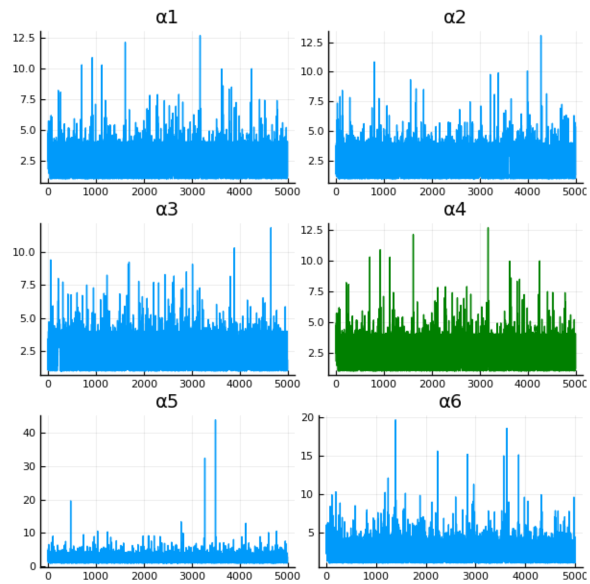

Figure 5.3: Estimated α's (First Six) corresponding to a trivariate Gaussian copula.

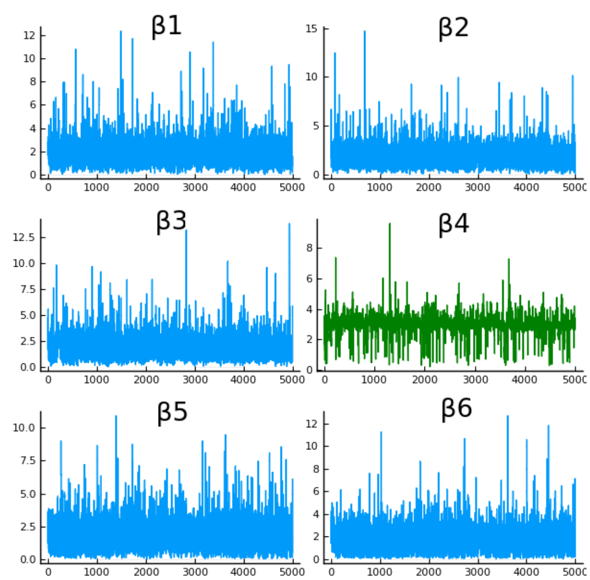

Figure 5.4: Estimated β's (First Six) corresponding to a trivariate Gaussian copula.

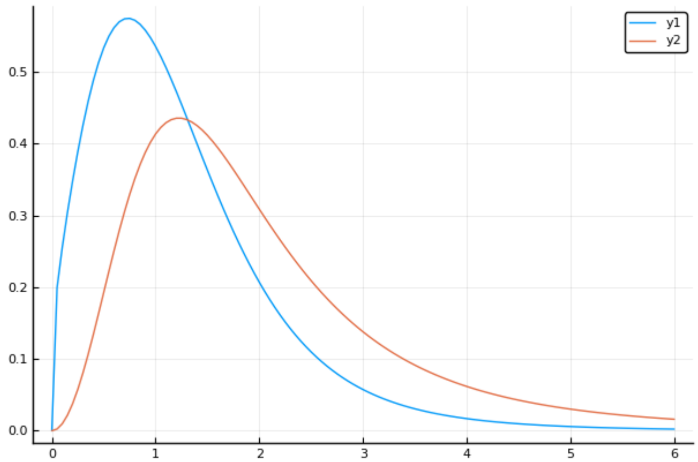

Figure 5.5: Estimated and theoretical $h(r)$ corresponding to a trivariate Student's T copula. Y_1 is the estimated one and Y_2 is the theoretical one.

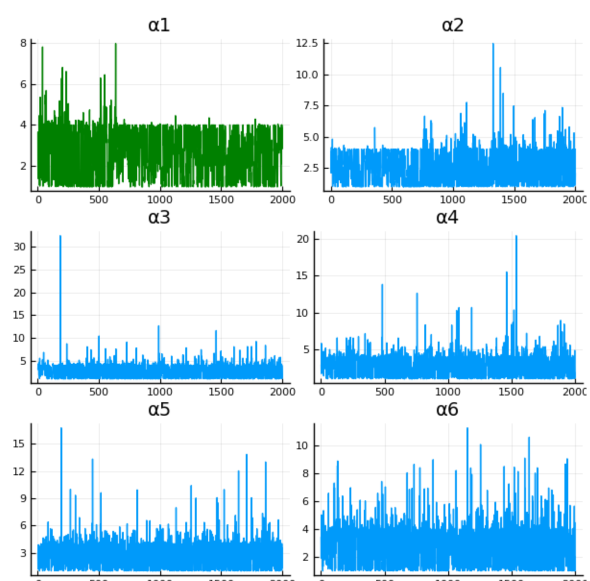

Figure 5.6: Estimated α's (First Six) corresponding to a trivariate Student's T copula.

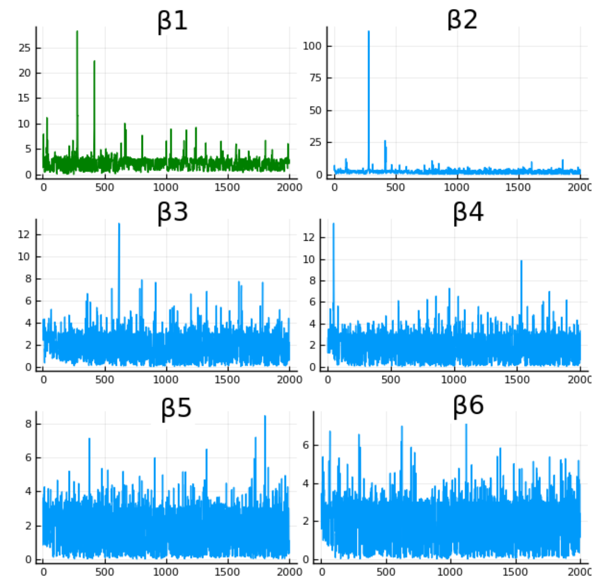

Figure 5.7: Estimated β's (First Six) corresponding to a trivariate Student's T copula.

Chapter 6

Conclusion and Future work

6.1 Conclusion

By using the Inversion Method, and MCMC approach, we developed a copula approach, which could estimate all 3-dimensional elliptical copulas without assuming it is a Gaussian copula, t copula, etc.

6.2 Future Aims

1. As is seen in the Figure (5.2) and (5.5) of Chapter 5, the estimated h's are not close enough to the theoretical ones. There might be hidden problems with the sampling scheme. One of future goals is to improve accuracy.

2. The whole MCMC process takes more than two weeks to finish since the evaluation of quantile function Q^{-1} involves numerical root finding and it is time-consuming. Another future goal is to increase efficiency.

3. The current posterior distributions and sampling schemes are only for 3-dimensional case. After solving the problem in the evaluation the quantile function, our final goal is to generalize this method to higher dimensional case to make it more useful in practice.

6.3 Potential Solution

Since the main problem is caused by the evaluation of Q^{-1}, perhaps it should be modeled differently instead of deriving Q from the generator and finding inverse function.

6.3.1 Linear Combination of Basis Functions

Klein and Smith (2018) introduced a model called copula smoother, allowing nonparametric regression smoothing through the use of a copula.

Let $\boldsymbol{Y} = (Y_1, \ldots, Y_n)'$ be n observations on a continuous response, with covariate vector $\boldsymbol{x}$, satisfying two conditions:

1. $Y_i|x_i$ has the same distribution function F_Y and density function p_Y for $i = 1, \ldots, n$;
2. The dependence structure of the distribution of $\boldsymbol{Y}$ is modeled using a copula which can be expressed with a mixture of Gaussian copulas.

According to Sklar's Theorem, the joint density of $\boldsymbol{Y}|\boldsymbol{x}$ can be written as

$$p(\boldsymbol{Y}|\boldsymbol{x}) = c^{\dagger}(F_Y(y_1), \ldots, F_Y(y_n)|\boldsymbol{x}) \prod_{i=1}^{n} p_Y(y_i), \tag{6.3.1}$$

where $c^{\dagger}$ is the unknown copula of the distribution of $\boldsymbol{Y}$.

Klein and Smith (2018) used the Inversion Method to model the copula $c^{\dagger}$,

$$C_{\pi}(\boldsymbol{u}|\boldsymbol{x}) = F_{\boldsymbol{Z}}(F_{z_1}^{-1}(u_1|\boldsymbol{x}), \ldots, F_{z_n}^{-1}(u_n|\boldsymbol{x})|\boldsymbol{x}), \tag{6.3.2}$$

which is itself a function of $\boldsymbol{x}$. $\boldsymbol{Z}$ is the vector of unobserved latent variables.

The model for the copula of $\boldsymbol{Y}$ is derived from a Bayesian regularized regression model.

For simplicity, Klein and Smith (2018) assumed a single covariate .

Consider the regression model

$$\tilde{Z}_i = \tilde{m}_(x_i) + \varepsilon_i, \text{for } i = 1, \ldots, n \tag{6.3.3}$$

for a pseudo-response $\tilde{Z}_i$, where $\tilde{m}$ is an unknown univariate function, x_i is a covariate value, and the $\varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$.

In Klein and Smith (2018), the function $\tilde{m}$ is modeled as a linear combination of p basis functions $b_1, \ldots, b_p$ with coefficients $\beta_1, \ldots, \beta_p$, so that $\tilde{m}(x) = \sum_{j=1}^{p} \beta_j b_j(x)$. (6.3.3) then can be written as

$$\tilde{\boldsymbol{Z}} = B\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{6.3.4}$$

expanded as

$$\begin{bmatrix} \tilde{Z}_1 \\ \tilde{Z}_2 \\ \vdots \\ \tilde{Z}_n \end{bmatrix} = \begin{bmatrix} b_1(x_1) & b_2(x_1) & \ldots & b_p(x_1) \\ b_1(x_2) & b_2(x_2) & \ldots & b_p(x_2) \\ \ldots & \ldots & \ldots & \ldots \\ b_1(x_n) & b_2(x_n) & \ldots & b_p(x_n) \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}. \tag{6.3.5}$$

By using Equation (6.3.4), the evaluation of the quantile functions $z_i = F_{Z_i}^{-1}(u_i|\boldsymbol{x})$ is avoided, which can significantly improve the speed of computing. Besides, this remodeling of the quantile function can also make it easier to generalize this approach to higher dimensional cases.

6.3.2 Mixture of Experts

Beside of modeling the quantile function $z_i = F_{z_i}^{-1}(u_i)$ with a linear combination of basis functions, we can also use a mixture of experts,

$$p(z_i|\boldsymbol{x_i}, \boldsymbol{p}_i, \beta, \boldsymbol{\sigma}) = \sum_{j=1}^{K} p_{ij} \frac{1}{\sigma_j} \phi\left(\frac{z_i - \boldsymbol{\beta}_j' \boldsymbol{x_i}}{\sigma_j}\right), \tag{6.3.6}$$

where

- ϕ is the standard normal pdf,
- $\boldsymbol{x}_i = (1, x_i)$,
- $\boldsymbol{p}_i = (p_{i1}, \ldots, p_{iK})$,

- $\beta = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_K)$,
- $\boldsymbol{\sigma^2} = (\sigma_1, \ldots, \sigma_K)$,
- p_{ij} is the mixing proportion

$$p_{ij} = \frac{\exp(\delta_j' \boldsymbol{x}_i)}{\sum_{j=1}^{K} \exp(\delta_j' \boldsymbol{x}_i)},$$

6.4 Time Schedule of Future Research

Table 6.1: Time Schedule

Tasks to complete project	Approximate time period
Deriving the combination of basis functions	January - March 2020
Coding the combination of basis functions	April 2020
Deriving the mixture of experts	May - August 2020
Coding the mixture of experts	September 2020
Generalize to higher dimensional copulas	October 2020 - January 2021
Coding the generalized method	February - May 2021
Thesis writing: Literature review	June 2021
Thesis writing: Methodology	July 2021
Thesis writing: Experiment	August
Thesis writing: Results (Generating plot)	September 2021
Thesis writing: First draft	October 2021
Thesis writing: Revising	October 2021
Defense	November 2021
Submission and publication	December 2021

References

Behboodian, J. (1970), "On a Mixture of Normal Distributions," *Biometrika*, 57, 215–217.

Fang, H.-B., Fang, K.-T., and Kotz, S. (2002), "The meta-elliptical distributions with given marginals," *Journal of Multivariate Analysis*, 82, 1–16.

— (2005), "Correction to "The meta-elliptical distributions with given marginals" (2002V82 p1-16)," *Journal of Multivariate Analysis*, 94, 222–223.

Ferguson, T. S. (1973), "A Baysian analysis of some nonparametric problems," *The Annals of Statistics*, 1, 209–230.

Genest, C., Favre, A.-C., Béliveau, J., and Jacques, C. (2007), "Metaelliptical copulas and their use in frequency analysis of multivariate hydrological data," *Water Rresources Research*, 43, 1–12.

Gómez, E., Gómez-Villegas, M. A., and Marín, J. M. (2003), "A survey on continuous elliptical vector distributions," *Revista Matematica Complutense*, 16, 345–361.

Kalli, M., Griffin, J. E., and Walker, S. G. (2011), "Slice Sampling Mixture Models," *Statistics and Computing*, 21, 93–105.

Klein, N. and Smith, M. S. (2018), "Implicit Copulas from Bayesian Regularized Regression Smoothers," *Bayesian Analysis*.

Pourahmadi, M. and Wang, X. (2015), "Distribution of random correlation matrices: hyperspherical parameterization of the Cholesky factor," *Statistics and Probability Letters*, 106, 5–12.

Rapisarda, F., Brigo, D., and Mercurio, F. (2007), "Parameterizing correlations: a geometric interpretation," *IMA Journal of Management Mathematics*, 18, 55–73.

Sethuraman, J. (1994), "A constructive definition of Dirichlet priors," *Statistica Sinica*, 639–650.

Sklar, A. (1959), *Fonctions de répartition à n dimensions et leurs marges*, Inst. Statist. Univ. Paris, pp. 229–231.

Appendix A

Julia Code for Estimating Elliptical Copulas

```
using Cubature
using Roots
using Distributions
using IncGammaBeta
using Optim
using LinearAlgebra

using DynamicHMC
using ContinuousTransformations
using Parameters
using MCMCDiagnostics
using DiffWrappers

using SpecialFunctions
using Random
using MATLAB
###############################################

n=500#Sample size
dimention=3
```

```
#set up correlation matrix
rho12=.5
rho13=.6
rho23=.7
omega=[1 rho12 rho13; rho12  1  rho23; rho13  rho23 1]

Random.seed!(1234)
Z=transpose(rand(MvNormal([0,0,0],omega), 500))
#Get samples from the Gaussian copula
U=cdf.(Normal(0,1), Z)

#For estimate the copula of Student's T distribution,
#replace the data with samples from copula for Student's T
#U=readdlm("U_meta_T3.txt")

U=round.(U; digits=2)
#To avoid 1 and 0 in probability
U[(U.>0.99)].=0.99
U[U.<0.01].=0.01

#set up gamma mixture
k=30 # Indicating infinite number of components
# eta1=1
# eta2=1
# hgamma=rand(Gamma(eta1,1/eta2))
# betaDists=Beta(1,hgamma)
# v=Matrix{Float64}(1,k)
```

```
# for j in 1:(k-1)
#     v[:,j]=rand(betaDists,1)
# end
# v[:,k]=1

#Initialized mixing proportion
pEst=Matrix{Float64}(undef,k,1)
pEst[:,1]=repeat(1/k:1/k, inner=k)

#initialized alphas and lamdas for gamma kernels
alphaEst=Matrix{Float64}(undef,k,1)
alphaEst[:,1]=repeat(1.1:1.1, inner=k)

lamdaEst=Matrix{Float64}(undef,k,1)
lamdaEst[:,1]=repeat(1:1, inner=k)

#Hyperparameters for the prior of lamda, which is gamma dist
alphaNull=1
lamdaNull=1

# Marginal CDF of elliptical distribution
z=2
function Q(z)
if z!=0
sign_z=sign(z)
abs_z=abs(z)

1/2+pi^(-1/2)*gamma(3/2).*
```

```
sum(pEst./((alphaEst.-1).*
gamma.(alphaEst.-1)).*
(sign_z.*gamma.(alphaEst)+
z.*lamdaEst.*inc_gamma_upper.(alphaEst.-1, lamdaEst.*abs_z)-
sign_z.*inc_gamma_upper.(alphaEst, lamdaEst.*abs_z)
)
)
else
1/2
end

end
#Quantile function
function Qinv(p)
if p==0.5
0
else
f(x)=abs(Q(x)-p)
op=optimize(f, -10,10)
# optim_error=optim_error+(op.converged==false)
op.minimizer
end
end

#Quantile function of elliptical distribution----"Translator"
function Qinv(p)
quantile.(Normal(0,1), p)
end
```

```
######################################################3
nIteration= 2000
pRecord=Matrix{Float64}(undef,k, nIteration)
alphaRecord=Matrix{Float64}(undef,k, nIteration)
lamdaRecord=Matrix{Float64}(undef,k, nIteration)
eRecord=Matrix{Float64}(undef,k, nIteration)

#m is lower triangular matrix with 0s in diangonal,
#this is a matrix used in calculation of parameter
m=LowerTriangular(ones(k-1,k-1))+Diagonal(repeat(-1:-1, inner=k-1))

densityIndividual=Matrix{Float64}(undef,k,n)
zn=Matrix{Int64}(undef,k ,n)
v=Matrix{Float64}(undef,1,k)
alphaEstPropose=Matrix{Float64}(undef,k,1)
lamdaEstPropose=Matrix{Float64}(undef,k,1)

#Initialized "translation" into elliptical data
X=Qinv.(U)
#By so far, correlation matrix is not being inferenced
omegaEst=omega
omegaEstInv=inv(omegaEst)
RSqr=diag(X*omegaEstInv*transpose(X))
REst=sqrt.(RSqr)
#replace 0 in REst with a random value from nonzero elements
#otherwise the densityMixture will be zero and in denominator
REst[REst.==0].=0.001
#i=1
```

```
@time begin
for i in 1:nIteration
#f(x| mu_i, phi_i) k x n matrix
for j in 1:k
densityIndividual[j,:]=pdf.(Gamma(alphaEst[j],1/lamdaEst[j]), REst)
end

#whole mixture pdf f(x)=p1*f(x|mu1,phi1) + p2*f(x|mu2,phi2)
densityMixture=transpose(pEst)* densityIndividual
#probability vector of Zi vector
prob=densityIndividual.*pEst./densityMixture
for j in 1: n
zn[:,j]=rand(Multinomial(1, prob[:,j]), 1)
end
#########################
r=sum(zn,dims=2)
# hgamma=rand(Gamma(k+eta1-1,1/(eta2-sum(log.(1-v[1:k-1])))),1)
#+hgamma
betaDists=Beta.(1 .+r[1:k-1],(transpose((transpose(r[1:k-1]))*m .+1)))
# v=Matrix{Float64}(1,k)
for j in 1:(k-1)
v[:,j]=rand(betaDists[j],1)
end
v[:,k].=1
summ=0 #do not define variable with a name "sum"
for j in 1:k
pEst[j]=(1-summ)*v[j]
summ=summ+pEst[j]
```

```
end
pRecord[:, i]=pEst
X=Qinv.(U)
RSqr=diag(X*omegaEstInv*transpose(X))
REst=sqrt.(RSqr)
REst[REst.==0].=0.001
#draw alphas and lamdas from NUTS
mat"""
logpdf = @(Parameters)logposterior(Parameters, double($zn),
$REst,1,1, 3,500 );
samples= NUTS_wrapper(logpdf ,repelem(1.1,60).' , 3, 1);
$alphaEst=1+exp(samples(1:30));
$alphaEst($alphaEst==1)=1.001;
$lamdaEst=exp(samples(31:60));
$lamdaEst($lamdaEst==0)=0.001;
"""

alphaRecord[:, i]=alphaEst;
lamdaRecord[:, i]=lamdaEst;
X=Qinv.(U)
RSqr=diag(X*omegaEstInv*transpose(X))
REst=sqrt.(RSqr)
REst[REst.==0].=0.001
end
########################################
end
```

Curriculum Vitae

Panfeng Liang was from People's Republic of China. He went to Minzu University of China in the fall of 2009 and received his bachelor's degree in Computer Science in the Fall of 2013. He entered The University of Texas at El Paso in August 2015. While pursuing his master's degree in Statistics he worked as a Teaching Assistant and Research Assistant. In the Spring of 2018, he started the phd program in computational science in The University of Texas at El Paso.

Contact Information: pliang@miners.utep.edu