李丰：瑞典斯德哥尔摩大学统计系

研究兴趣:

贝叶斯理论，计量经济学，预测方法，多元模型

博士论文:

Flexible Bayesian Regression Density Estimation

(柔性贝叶斯回归密度估计)

教过课程:

回归分析，时间序列，统计计算，贝叶斯方法

Introduction to covariate-dependent copula modeling [1]

Feng Li

Department of Statistics, Stockholm University

March, 2013

[1] Based on paper **Li, F. (2012), Modeling covariate-contingent correlation and tail-dependence with copulas**.

What is a copula?

- The word "**copula**" means **linking**.
- **Sklar's theorem (1959)**

  Let H be a multi-dimensional distribution function with marginal distribution functions $F_1(x_1), ..., F_M(x_M)$. Then there exists a function C (**copula function**) such that

$$H(x_1, ..., x_M) = C(F_1(x_1), ..., F_M(x_M))$$
$$= C\left(\int_{-\infty}^{x_1} f(z_1)dz_1, ..., \int_{-\infty}^{x_M} f(z_M)dz_M\right) = C(\mathbf{u}_1, ..., \mathbf{u}_M).$$

Some arbitrary examples

- If $X_1, \ldots, X_M$ are independent, and iff C is a product copula, then

$$C(F_1(x_1), \ldots, F_M(x_M)) = \prod_{i=1}^{M} F_i(x_i)$$

- The bivariate Gaussian copula

$$C(u_1, u_2, \rho) = \mathbf{\Phi}_2(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \rho)$$

- The multivariate probit model is a simple Gaussian copula model, with univariate probit regressions as the marginals.

Correlation and dependence concepts

- The **Kendall's τ** can be written in terms of copula function:

$$\tau = 4 \int\int C(u_1, u_2) dC(u_1, u_2) - 1.$$

- As well as the bivariate lower and upper **tail dependences**

$$\lambda_L = \lim_{u \to 0^+} Pr(X_1 < F_1^{-1}(u) | X_2 < F_2^{-1}(u)) = \lim_{u \to 0^+} \frac{C(u, u)}{u}$$

$$\lambda_U = \lim_{u \to 1^-} Pr(X_1 > F_1^{-1}(u) | X_2 > F_2^{-1}(u)) = \lim_{u \to 1^-} \frac{1 - C(u, u)}{1 - u}$$

The covariate-contingent copula model

- **The marginal models**
  - In principle, any combination of univariate marginal models can be used.
- **The log likelihood**

$$\begin{aligned}
&\log \mathcal{L}(\mathbf{Y}|\mathbf{X}, \lambda_L, \tau, \boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_M) \\
&= \sum\nolimits_{i=1}^{n} \log c(\mathbf{u}_1, \ldots, \mathbf{u}_m, \lambda_L, \tau) \\
&\quad + \sum\nolimits_{m=1}^{M} \log \mathcal{L}_m(\mathbf{Y}_m|\mathbf{X}_m, \boldsymbol{\beta}_m)
\end{aligned}$$

- **Covariate-dependent structure**

$$\boldsymbol{\beta}_m = \varphi^{-1}_{\beta_m}(\mathbf{X}_m \boldsymbol{\alpha}_m) \quad \tau = \varphi^{-1}_{\tau}(X\alpha_\tau)$$

The covariate-contingent copula model

- The priors
  - The priors for the copula functions are easy to specify due to our reparameterization.
  - The priors for the marginal distributions are specified in their usual ways.
- The posterior

$$p(\boldsymbol{\alpha}|\mathbf{Y}) \propto \mathcal{L}(\mathbf{Y}|\alpha) \times \prod_{i \in \{1,..,M,C\}} p(\alpha_i)$$

- The posterior inference is straightforward although the model is very complicated.

The covariate-contingent copula model

- We update all the parameters jointly by using Metropolis-Hastings within Gibbs.
- The proposal density for each parameter vector α is a multivariate t-density with $\mathrm{df} > 2$,

$$\boldsymbol{\alpha}_p | \boldsymbol{\alpha}_c \sim \mathbf{MVT}\left[\hat{\boldsymbol{\alpha}}, \ -\left(\frac{\partial^2 \ln p(\boldsymbol{\alpha}|\mathbf{Y})}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\alpha}'}\right)^{-1}\Bigg|_{\boldsymbol{\alpha}=\hat{\boldsymbol{\alpha}}}, \ \mathrm{df}\right],$$

where $\hat{\boldsymbol{\alpha}}$ is obtained by R steps $(\mathrm{R} \leqslant 3)$ Newton's iterations during the proposal with analytical gradients.
- Variable selections are carried out simultaneously.
- **The key:** The analytical gradients require the derivative

Thank you!

For further details, see

Li, F. (2012), Modeling covariate-contingent correlation and tail-dependence with copulas.