Skew $t$-copula and its estimation:
For application to risk aggregation

Toshinao Yoshiba
(Bank of Japan / The Institute of Statistical Mathematics)
E-mail: toshinao.yoshiba@boj.or.jp

Views expressed in this presentation are those of the author and do not necessarily reflect the official views of the Bank of Japan.
Outline

I. Introduction
II. Construction of skew $t$-copula
III. Estimation of skew $t$-copula
IV. Tail dependence of skew $t$-copula
V. Conclusions and open problems
I. Introduction

• Practical needs for risk aggregation
  1. Overall dependence is captured by linear/rank correlation.
  2. Lower tail dependence is captured.
  3. Asymmetry of tail dependence is captured.

Which copula?

– 1 : Gaussian copula
– 1 & 2 : Student’s $t$-copula
– 1 & 2 & 3 : Skew $t$-copula
  • HAC or vine copula may be another way.
I. Introduction (cont.)

- Variety of skew $t$-copula
  - Wide variety for multivariate skew $t$-distribution
  1. Demarta and McNeil (2005) copula (GH skew $t$)
    \[
    X = \gamma V^{-1} + \frac{Z}{\sqrt{V}} \quad \text{V} \sim G(\nu / 2, \nu / 2), \; Z \sim N_d(0, \Sigma)
    \]

- Smith, Gan, and Kohn (2012) copula
  - Implied in Sahu, Dey, and Branco (2003) skew $t$-distribution
  \[
  X = \Gamma |W| + Z \quad \text{V} \sim G(\nu / 2, \nu / 2), \; Z \sim N_d(0, \Sigma)
  \]
  \[
  W \sim N_d(0, I), \; \Gamma = \text{diag}(\gamma_1, \ldots, \gamma_d)
  \]
I. Introduction (cont.)

• Variety of skew t copula (cont.)

   – Kollo and Pettere (2010) tried to construct its copula

\[ X = \frac{Y}{\sqrt{V}} \quad V \sim G(\nu/2, \nu/2), \quad Y \sim SN_d(\lambda, \Psi) \]

\[ Y_j = \frac{\lambda_j |Z_0| + Z_j}{\sqrt{1 + \lambda_j^2}} \quad Z_0 \sim N(0,1), \quad Z \sim N_d(0, \Psi) \]

– \( Z_0 \) is scalar which is different from Sahu, Dey, and Branco (2003).
II. Construction of skew $t$-copula

- Azzalini and Capitanio (2003) $d$-variate skew $t$-distribution $St_d(0,\Omega,\alpha,\nu)$ has density

$$g(x) = 2t_{d,\nu}(x;\Omega)T_{1,\nu+d}\left(\alpha^T x \sqrt{\frac{\nu + d}{x^T \Omega^{-1} x + \nu}}\right),$$

where

$$\Omega = \Lambda^{-1}(\Psi + \lambda\lambda^T)\Lambda^{-1}, \quad \alpha = \frac{\Lambda\Psi^{-1}\lambda}{\sqrt{1 + \lambda^T \Psi^{-1} \lambda}},$$

$$\Lambda = \text{diag}(\sqrt{1 + \lambda_1^2}, \ldots, \sqrt{1 + \lambda_d^2}).$$

- $t_{d,\nu}(\cdot;\Omega)$: $d$-variate Student’s $t$ density with d.f. $= \nu$ and correlation matrix $\Omega$
- $T_{1,\nu+d}(\cdot)$: univariate Student’s $t$ distribution function with d.f. $= \nu + d$
II. Construction of skew $t$-copula (cont.)

- Marginal distribution for $\operatorname{St}_d(0,\Omega,\alpha,\nu)$ is $\operatorname{St}_1(0,1,\lambda_j,\nu)$
  - $j$-th marginal density
    
    $$
    g_j(x) = 2t_{1,\nu}(x)T_{1,\nu+1}\left(\lambda_j x \sqrt{\frac{\nu + 1}{x^2 + \nu}}\right)
    $$

  - $X_j = \frac{\lambda_j |Z_0| + Z_j}{\sqrt{\sqrt{1 + \lambda_j^2}}}$
    
    skewness

  - If $Z$ is correlated, then $\alpha \neq \lambda$

Kollo and Pettere (2010) misspecify as $\alpha_j$
II. Construction of skew $t$-copula (cont.)

- Skew $t$-copula based on Azzalini and Capitanio (2003)

$$C_{st}(u_1, \ldots, u_d ; \Omega, \lambda, \nu)$$

$$= \text{St}_d (\text{St}_1^{-1}(u_1;0,1,\lambda_1,\nu), \ldots, \text{St}_1^{-1}(u_d;0,1,\lambda_d,\nu);0,\Omega,\alpha,\nu)$$

where

$$\alpha = \frac{\Lambda \Psi^{-1} \lambda}{\sqrt{1 + \lambda^T \Psi^{-1} \lambda}} = \frac{\Lambda (\Lambda \Omega \Lambda - \lambda \lambda^T)^{-1} \lambda}{\sqrt{1 + \lambda^T (\Lambda \Omega \Lambda - \lambda \lambda^T)^{-1} \lambda}}$$

$$\Lambda = \text{diag}(\sqrt{1 + \lambda_1^2}, \ldots, \sqrt{1 + \lambda_d^2}).$$
II. Construction of skew $t$-copula (cont.)

- Student’s $t$-copula $\Rightarrow$ Skew $t$-copula

- $\rho = 0.5$, $\nu = 3$
- Marginal distribution: Standard Normal
III. Estimation of skew $t$-copula

- Maximum likelihood estimation (MLE)
  - For $n$ independent observations $u_1, \ldots, u_n$ unif. distributed on $[0,1]^d$, log-likelihood for parameter $\theta=(\Omega, \lambda, \nu)$ is

$$l(\theta) = \sum_{i=1}^{N} l_i(\theta),$$

$$l_i(\theta) = \ln \left( \frac{2\Gamma((\nu+d)/2)}{(\pi\nu)^{d/2}\Gamma(\nu/2)} \right) - \frac{1}{2} \ln |\Omega| - \frac{\nu+d}{2} \ln \left( 1 + \frac{x_i^T \Omega^{-1} x_i}{\nu} \right)$$

$$+ \ln \left[ T_{1,\nu+d} \left( \alpha^T x_i \sqrt{\frac{\nu+d}{x_i^T \Omega^{-1} x_i + \nu}} \right) \right] - \sum_{j=1}^{d} \ln g_j(x_{ij}; \lambda_j, \nu),$$

where

$$x_i = (x_{i1}, \ldots, x_{id}), \ u_i = (u_{i1}, \ldots, u_{id}), \ x_{ij} = St_1^{-1}(u_{ij};0,1, \lambda_j, \nu)$$

- Accurate quantile function is time consuming.
  - e.g. $q_{St}$ function in $sn$ library for $R$
III. Estimation of skew \( t \)-copula (cont.)

- MLE (cont.)
  - Log-likelihood calculation time for \( n=1,000 \) bivariate data, using accurate quantile function (\( qst \) in \texttt{sn ver. 0.4-17} takes 4.4 seconds on Windows Vista with Intel Core 2 Duo CPU 2.4GHz.
  - Using empirical quantile with 100,000 random numbers 20 times faster.
  - However, because empirical quantile may have large error in the tail, it is an issue whether the empirical quantile function is enough accurate or not.
III. Estimation of skew $t$-copula (cont.)

- **Method of Moments**
  - Using some moments which does not depend on marginal distribution is alternative estimation to MLE and intuitive.
  - Correlation matrix $\Omega$ (or $\Psi$) is estimated from sample rank correlation. Degree of freedom parameter $\nu$ is estimated from average tail dependence. Skewness parameter $\lambda$ is estimated from difference in upper and lower tail dependence.
  - To estimate correlation matrix, **closed-form of rank correlation** is needed.
IV. Tail dependence of skew $t$-copula

- Tail dependence of bivariate copula
  - Tail conditional probabilities with some threshold $u$
    
    $$
    \eta_L(u) = \Pr[F_2(X_2) < u \mid F_1(X_1) < u] = \frac{C(u, u)}{u} \quad \text{lower}
    $$
    
    $$
    \eta_U(u) = \Pr[F_2(X_2) > u \mid F_1(X_1) > u] = \frac{1 - 2u + C(u, u)}{1 - u} \quad \text{upper}
    $$

- Tail dependence (definition)
  $$
  \eta_L = \lim_{u \to 0^+} \eta_L(u), \eta_U = \lim_{u \to 1^-} \eta_U(u)
  $$

- Lower and upper tail dependence of Student’s $t$-copula
  $$
  \eta_L = \eta_U = 2T_{1, \nu+1} \left( - \sqrt{\frac{(1 - \rho)(\nu + 1)}{(1 + \rho)}} \right) \equiv \eta_t
  $$
IV. Tail dependence of skew $t$-copula (cont.)

• Tail dependence of bivariate skew $t$-copula
  
  – Bortot (2010) derives tail dependences as

  \[
  \eta_L = \frac{1-T_{1,v+2}(2\alpha \sqrt{(v+2)(1+\rho)/2})}{1-T_{1,v+1}(\lambda \sqrt{v+1})} \eta_t
  \]

  \[
  \eta_U = \frac{T_{1,v+2}(2\alpha \sqrt{(v+2)(1+\rho)/2})}{T_{1,v+1}(\lambda \sqrt{v+1})} \eta_t \quad \lambda_1 = \lambda_2 = \lambda
  \]

  \[
  \alpha = \frac{\lambda}{\sqrt{(1+\rho)^2 - \lambda^2 (1-\rho^2)}}
  \]

IV. Tail dependence of skew $t$-copula (cont.)

- Lower and upper tail dependence depend on skewness parameter $\lambda$

\[
\lambda > (\langle 0 \iff \frac{T_{1,v+2}(2\alpha \sqrt{(v + 2)(1 + \rho)} / 2)}{T_{1,v+1}(\lambda \sqrt{v + 1})} > (\langle 1
\]

- $\lambda < 0 \implies \eta_L > \eta_t > \eta_U$
- $\lambda > 0 \implies \eta_L < \eta_t < \eta_U$

- Bortot (2010) uses different approximation from Fung and Seneta (2010). It is also an issue whether both tail dependence coincide.

\[
\int_{-\infty}^{-\sqrt{(1-\rho)(v+1)/(1+\rho)}} T_{1,v+1}(z) \frac{\partial}{\partial z} T_{1,v+2} \left( \frac{\alpha \sqrt{v + 2}\sqrt{(1-\rho^2)/(v+1)} z - (1+\rho)}{\sqrt{1+z^2/(v+1)}} \right) dz = 0?
\]
V. Conclusions and open problems

• We construct skew $t$-copula implied in Azzalini and Capitanio (2003) multivariate skew $t$-distribution

• Estimation of skew $t$-copula have several open problems
  – For MLE, accurate and fast quantile function for univariate skew $t$-distribution is needed.
  – For method of moments, closed-form of rank correlation is needed.
  – Which way should practitioners go including some other robust methods than MLE, method of moments?


