Estimation of Structured t-Copulas

Attilio Meucci[1]
attilio_meucci@symmys.com

This version: April 2008

Last version available at
`www.symmys.com > Research > Working Papers`

Abstract

We describe a simple recursive routine to estimate by maximum likelihood the correlation matrix and the degrees of freedom of the *t*-copula, when structure needs to be imposed on the eigenvalues for dimensionality issues. The code implementing the routines discussed here is available at `www.symmys.com` > `Teaching` > `MATLAB`.

JEL Classification: *C1, G11*

Keywords: *isotropy, shrinkage, structured correlation, estimation-maximization, maximum likelihood, radial generator*

[1] The author is grateful to Rahul Malhotra and Jesus Ruiz-Mata for their feedback

We consider a generic market with a t copula parametrized by the degrees of freedom and a correlation matrix. In Section 1 we setup the model. In Section 2 we write the implicit equation satisfied by the maximum likelihood estimator of the correlation matrix. In Section 3 we propose a recursive routine to solve this equation and estimate correlation and degrees of freedom. In Section 4 we adapt this routine to impose extra structure on the correlation matrix in the estimation process: this situation is important for applications in finance, where the number of variables is large, as compared to the number of observations. In Section 5 we show an application to the par swap interest rate curve.

1 The market model

We consider a market of N risk factors $\mathbf{X}$. We assume fully general marginal distributions, as represented by their cdf's $F_1, \ldots, F_N$ and a Student t copula, parametrized by a correlation matrix $\mathbf{C}$ and the degrees of freedom ν.

In other words, consider the transformed factors $\mathbf{Y}$ defined entry-wise as follows:

$$Y_n \equiv \Phi_\nu^{-1}\left(F_n\left(X_n\right)\right), \quad n = 1, \ldots N, \tag{1}$$

where Φ_ν denotes the cdf of the univariate t distribution. Our market $\mathbf{X}$ is such that $\mathbf{Y}$ is distributed as

$$\mathbf{Y} \sim \mathrm{St}\left(\nu, \mathbf{0}, \mathbf{C}\right), \tag{2}$$

where $\mathrm{St}\left(\nu, \boldsymbol{\mu}, \boldsymbol{\Sigma}\right)$ denotes the multivariate t distribution with ν degrees of freedom, location parameter $\boldsymbol{\mu}$ and scatter parameter $\boldsymbol{\Sigma}$.

We assume that we observe a time series of T independent and identically distributed (across time) joint observations of the factors $\mathbf{X}$, which we collect into a $T \times N$ panel $\mathcal{X}$. In other words, these factors are the "invariants" of the market. We stress that $\mathbf{X}$ need not represent an observable: for instance, in a multivariate GARCH model, $\mathbf{X}$ represent the i.i.d. shocks that drive the dynamics of the whole process.

Our purpose is estimating ν and $\mathbf{C}$ from $\mathcal{X}$.

2 Estimation: theory

First of all, we proceed to estimate the marginal distributions of the N factors. This can be done in terms of the empirical distribution, kernel smoothing, or parametric assumptions. This is not the focus of this note, and therefore we assume that these estimates are given and represented by the cdf's $\widehat{F}_1, \ldots, \widehat{F}_N$.

Assuming that the degrees of freedom ν are known, we can generate the $T \times N$ panel $\mathcal{Y}$ defined entry-wise as follows

$$\mathcal{Y}_{t,n} \equiv \Phi_\nu^{-1}\left(\widehat{F}_n\left(\mathcal{X}_{t,n}\right)\right). \tag{3}$$

The panel $\mathcal{Y}$ represents a time series of joint independent and identically distributed realizations of the multivariate t distribution (2). The ML estimator of the parameters of generic elliptical distributions is discussed in Meucci (2005). Adapting from that source, we show in the appendix that, for given degrees of freedom ν, the ML estimator of $\mathbf{C}$ is a weighted version of the sample estimator:

$$\widehat{\mathbf{C}} = \frac{1}{T} \sum_{t=1}^{T} w_t \mathbf{y}_t \mathbf{y}_t', \tag{4}$$

where $\mathbf{y}_t$ denotes transpose of the t-th row in $\mathcal{Y}$ and the weights w_t are defined as follows:

$$w_t \equiv \frac{\nu + N}{\nu + \mathbf{y}_t' \widehat{\mathbf{C}}^{-1} \mathbf{y}_t}. \tag{5}$$

This is a highly implicit definition of the estimator $\widehat{\mathbf{C}}$, which cannot be solved analytically, except in the normal case: indeed, when the degrees of freedom are very large $w_t \approx 1$ and (4) becomes the standard sample correlation.

3 Estimation: practice

Although (4)-(5) cannot be solved analytically, it is easy to compute $\widehat{\mathbf{C}}$ numerically by means of a recursive routine, similar in nature to the expectation-maximization algorithm by Dempster, Laird, and Rubin (1977).

Step 0. Set $u \equiv 0$ and initialize $\widehat{\mathbf{C}}_{(u)}$ as the sample correlation.

Step 1. Compute the weights:

$$w_t^{(u)} \equiv \frac{\nu + N}{\nu + \mathbf{y}_t' \widehat{\mathbf{C}}_{(u)}^{-1} \mathbf{y}_t}. \tag{6}$$

Step 2. Compute the scatter matrix:

$$\widehat{\mathbf{\Sigma}}_{(u+1)} = \frac{1}{T} \sum_{t=1}^{T} w_t^{(u)} \mathbf{y}_t \mathbf{y}_t'. \tag{7}$$

Step 3. Extract the correlation

$$\widehat{\mathbf{\Sigma}}_{(u+1)} \equiv \operatorname{diag}(\widehat{\boldsymbol{\sigma}}) \, \widehat{\mathbf{C}}_{(u+1)} \operatorname{diag}(\widehat{\boldsymbol{\sigma}}). \tag{8}$$

Step 4. Check for convergence

$$d \equiv \sqrt{\frac{1}{N} \operatorname{tr} \left(\widehat{\mathbf{C}}_{(u+1)} - \widehat{\mathbf{C}}_{(u)} \right)^2}. \tag{9}$$

If d is less than a desired threshold, say 10^{-15}, stop. Otherwise, set $u \equiv u+1$ and go back to Step 1.

The estimate $\widehat{\mathbf{C}}(\nu)$ is a function of the number of degrees of freedom we assumed in the beginning. To estimate ν, consider the log-likelihood

$$\mathcal{L}(\nu) \equiv \frac{1}{T} \sum_{t=1}^{T} \ln \left(f_{\nu, \mathbf{0}, \widehat{\mathbf{C}}(\nu)} (\mathbf{y}_t) \right), \tag{10}$$

where $f_{\nu,\boldsymbol{\mu},\boldsymbol{\Sigma}}$ denotes the pdf of the multivariate t distribution, see (20). We select a grid $\mathcal{G} \equiv \{\nu_1, \ldots, \nu_J\}$ of J significative values for ν and we set the estimator of the degrees of freedom as:

$$\widehat{\nu} \equiv \underset{\nu \in \mathcal{G}}{\operatorname{argmax}}\, \mathcal{L}(\nu)\,. \tag{11}$$

The above routine is extremely fast and convergence is reached within a few seconds even for very large problems.

4 Imposing structure

In many applications in finance the number of observations T is large compared to the number of factors N. In this situation it is desirable to impose structure on the correlation matrix that defines the copula. Here we consider one such example. The present discussion can easily be generalized to different correlation structures.

Consider the principal component decomposition of the correlation

$$\mathbf{C} \equiv \mathbf{E}\boldsymbol{\Lambda}\mathbf{E}'. \tag{12}$$

In this expression

$$\boldsymbol{\Lambda} \equiv diag\,(\lambda_1, \ldots, \lambda_N) \tag{13}$$

is the diagonal matrix of the eigenvectors of $\mathbf{C}$ sorted in decreasing order; and

$$\mathbf{E} \equiv \left(\mathbf{e}^{(1)}, \ldots, \mathbf{e}^{(N)}\right)$$

is the orthogonal matrix whose columns represent the respective eigenvectors of $\mathbf{C}$.

We impose that the last eigenvalues be equal and strictly larger than zero. In other words, we impose that (13) read

$$\boldsymbol{\Lambda} \equiv diag\,(\lambda_1, \ldots, \lambda_K, \underline{\lambda}, \ldots, \underline{\lambda})\,, \tag{14}$$

for some $K \leq N$ and some lower bound $\underline{\lambda}$. The choice of K is data driven.

This structure provides a particular form of shrinkage, see Meucci (2005). It assumes isotropy in the hyperplane spanned by the eigenvectors relative to the lowest eigenvalues. Indeed, from an estimation perspective, the directions defined by the lowest eigenvectors are not well defined.

To impose (14) in the estimation process, we add a few steps to the routine in Section 3

Step 2b. Perform the principal component decomposition of the estimated correlation:

$$\widehat{\boldsymbol{\Sigma}}_{(u+1)} \equiv \widehat{\mathbf{E}}\widehat{\boldsymbol{\Lambda}}\widehat{\mathbf{E}}'. \tag{15}$$

Step 2c. Redefine the eigenvalues:

$$\widehat{\boldsymbol{\Lambda}} \equiv diag\left(\widehat{\lambda}_1, \ldots, \widehat{\lambda}_K, \underline{\lambda}, \ldots, \underline{\lambda}\right), \tag{16}$$

where

$$\underline{\lambda} \equiv \frac{1}{N-K+1} \sum_{k=K+1}^{N} \widehat{\lambda}_k. \tag{17}$$

Step 2d. Recompose the scatter matrix:

$$\widehat{\mathbf{\Sigma}}_{(u+1)} \equiv \widehat{\mathbf{E}}\widehat{\mathbf{\Lambda}}\widehat{\mathbf{E}}. \tag{18}$$

We remark that the subsequent steps of the routine in Section 3 can distort the exact structure (14) of the eigenvalues. However, extensive experiments show that this distortion is minimal.

5 Case study

Consider the daily changes in the 1y, 2y, 5y, 7y,10y, 15y, 30y points of the par swap rate curve. In first approximation, we can consider them as invariants, see Meucci (2005). We collect in the panel $\mathcal{X}$ ten years of data for a total of approximately 2,500 joint observations.

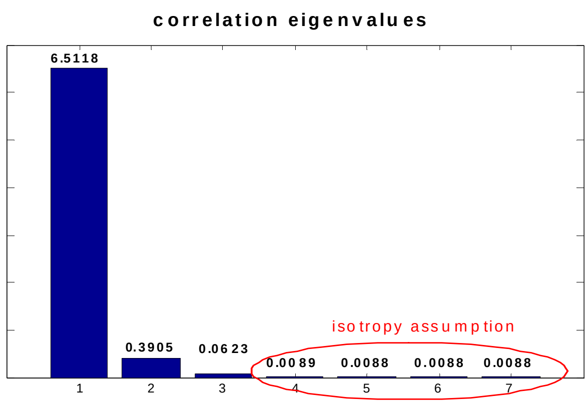

Figure 1: Structured t correlation matrix

We estimate the marginal distributions non–parametrically in terms of their empirical cdf's. In other words, the panel $\mathcal{U}$ defined entry-wise as

$$\mathcal{U}_{t,n} \equiv \widehat{F}_n\left(\mathcal{X}_{t,n}\right)$$

is the empirical copula of the interest rate changes. To extract this empirical copula $\mathcal{U}$ we use the methodology in Meucci (2006). With $\mathcal{U}$ we can compute as in (3) the panel of joint t realizations:

$$\mathcal{Y}_{t,n} \equiv \Phi_{\nu}^{-1}\left(\mathcal{U}_{t,n}\right). \tag{19}$$

Now we can apply the routine discussed in Section 4 to fit a structured t copula to the rate changes. The movements of the curve are described almost in full by the first three principal components, see Litterman and Scheinkman (1991). Therefore, we set $K \equiv 3$ in (14). In Figure 1 we display the results. The code for this case study is available at `www.symmys.com` > `Teaching` > `MATLAB`.

References

Dempster, A. P., M. N. Laird, and D. B. Rubin, 1977, Maximum likelihood from incomplete data via the EM algorithm, *Journal of the Royal Statistical Society* 39, 1–22.

Litterman, R., and J. Scheinkman, 1991, Common factors affecting bond returns, *Journal of Fixed Income* 1, 54–61.

Meucci, A., 2005, *Risk and Asset Allocation* (Springer).

———, 2006, Beyond Black-Litterman in practice: A five-step recipe to input views on non-normal markets, *Risk* 19, 114–119.

6 Appendix

The the probability density function of a variable $\mathbf{Z} \sim \mathrm{St}\left(\nu, \boldsymbol{\mu}, \boldsymbol{\Sigma}\right)$ reads

$$f_{\nu,\boldsymbol{\mu},\boldsymbol{\Sigma}}\left(\mathbf{x}\right) \equiv \frac{1}{\sqrt{|\boldsymbol{\Sigma}|}} g_{N,\nu}\left(\mathrm{Ma}^2\left(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}\right)\right), \tag{20}$$

where the square Mahalanobis distance is defined as

$$\mathrm{Ma}^2\left(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}\right) \equiv \left(\mathbf{x} - \boldsymbol{\mu}\right)' \boldsymbol{\Sigma}^{-1}\left(\mathbf{x} - \boldsymbol{\mu}\right) \tag{21}$$

and $g_{N,\nu}$ is the generator:

$$g_{N,\nu}\left(z\right) \equiv \frac{\Gamma\left(\frac{\nu+N}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\left(\nu\pi\right)^{\frac{N}{2}}}\left(1 + \frac{z}{\nu}\right)^{-\frac{\nu+N}{2}}. \tag{22}$$

As in Meucci (2005), the maximum likelihood estimators $\widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\Sigma}}$ of the location-dispersion parameters $\boldsymbol{\mu}, \boldsymbol{\Sigma}$ are defined implicitly by

$$\widehat{\boldsymbol{\mu}} = \sum_{t=1}^{T} \frac{w\left(\mathrm{Ma}^2\left(\mathbf{x}_t, \widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\Sigma}}\right)\right)}{\sum_{s=1}^{T} w\left(\mathrm{Ma}^2\left(\mathbf{x}_s, \widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\Sigma}}\right)\right)} \mathbf{x}_t \tag{23}$$

$$\widehat{\boldsymbol{\Sigma}} = \frac{1}{T}\sum_{t=1}^{T}\left(\mathbf{x}_t - \widehat{\boldsymbol{\mu}}\right)\left(\mathbf{x}_t - \widehat{\boldsymbol{\mu}}\right)' w\left(\mathrm{Ma}^2\left(\mathbf{x}_t, \widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\Sigma}}\right)\right), \tag{24}$$

where

$$w\left(z\right) \equiv -2\frac{g'_{N,\nu}\left(z\right)}{g_{N,\nu}\left(z\right)} = \frac{\nu + N}{\nu + z}. \tag{25}$$