Appendix of High Dimensional Dynamic Stochastic Copula Models*

Drew D. Creal† Ruey S. Tsay
Chicago Booth *Chicago Booth*

January 13, 2014

Abstract

This is an on-line appendix showing some of the details of implementation. The notation used is defined in the paper.

Keywords: state space models; dynamic copulas; Bayesian estimation; particle filters; credit default swaps.

JEL classification codes: C32, G32.

*We thank Cynthia Wu for helpful comments. Both authors gratefully acknowledge financial support from the Fama Miller Center for Research in Finance. Drew Creal gratefully acknowledges additional financial support from the William Ladany Faculty Scholar Fund at the University of Chicago Booth School of Business.

†Corresponding author: The University of Chicago Booth School of Business, 5807 South Woodlawn Avenue, Chicago, IL 60637, USA, dcreal@chicagobooth.edu

1 MCMC and particle filtering algorithms

1.1 MCMC algorithms for copulas

1.1.1 Full conditional distributions for $\zeta_{t,j}$

In Step 2 of the MCMC algorithm, we draw the mixing variables $\zeta_{t,j}$, where we have dropped the i subscript. For each group j, we separate the observations into two groups, those observations belonging to group j and those that do not. We use n_j to denote the number of observations in group j, $x_{t,j}$ is a $n_j \times 1$ vector of all x_{it} in this group. Similarly, $D_{t,j}$ is a $n_j \times n_j$ matrix that is a subset of D_t, $\tilde{\lambda}_{t,j}$ is the $n_j \times p$ subset of $\tilde{\lambda}_t$, and $\tilde{X}_{t,j}$ is the $n_j \times k$ subset of $\tilde{X}_t$.

$$
\begin{aligned}
p\left(\zeta_{t,j}|x_t, \lambda_t, z_t, \zeta_{t,j^-}, \theta\right) &\propto p\left(x_t|\lambda_t, z_t, \zeta_{t,j}, \zeta_{t,j^-}, \theta\right) p\left(\zeta_{t,j}|\nu_j\right) \\
&\propto \zeta_{t,j}^{-(n_j+\nu_j)/2-1} \exp\left(-\frac{1}{2}\zeta_{t,j}^{-1}\left[x_{t,j}' D_{t,j}^{-1} x_{t,j} + \nu_j\right] + \right. \\
&\qquad \left. \zeta_{t,j}^{-\frac{1}{2}} x_{t,j}' D_{t,j}^{-1}\left[\tilde{\lambda}_{t,j}' z_t + \beta_j' \tilde{X}_{t,j}\right]\right)
\end{aligned}
$$

When there is only one group $G = 1$, the model is a standard Student's t copula. Then, we can integrate out z_t and draw from the full conditional distribution exactly. The full conditional distributions for ζ_t are $\zeta_t \sim \text{Inv-Gamma}\left(\frac{\nu+n}{2}, \frac{\nu+x_t' R_t^{-1} x_t}{2}\right)$ for $t = 1, \ldots, T$.

1.2 Equicorrelation models

To draw missing values of u_{it} in the equicorrelation models with Student's t errors, we represent the model as $x_t = \sqrt{\zeta_t} R_t^{1/2} \varepsilon_t$ where $\varepsilon_t \sim \mathrm{N}\left(0, \mathrm{I}_n\right)$ and $\zeta_t \sim \text{Inv-Gamma}\left(\frac{\nu}{2}, \frac{\nu}{2}\right)$. We separate $x_t = \left(\hat{x}_t, x_t^*\right)$ where $\hat{x}_t$ are the $\hat{n}_t \times 1$ vector of missing values. Draw $\hat{x}_t$ from their full conditional distribution $\hat{x}_t \sim \mathrm{N}\left(\hat{\mu}_t, \zeta_t \hat{R}_t\right)$ where $\hat{\mu}_t = R_{t,12} R_{t,22}^{-1} x_t^*$ and $\hat{R}_t = R_{t,11} - R_{t,12} R_{t,22}^{-1} R_{t,21}$. The matrices $R_{t,ij}$ represent the relevant sub-blocks of the correlation matrix R_t. We then take $\hat{u}_{it} = T\left(\hat{x}_{it}|\nu\right)$ as the draw for the missing value.

The full conditional distributions for ζ_t are inverse Gamma distributions

$$\zeta_t \sim \text{Inv-Gamma}\left(\frac{\nu+n}{2}, \frac{\nu+x_t' R_t^{-1} x_t}{2}\right)$$

for $t = 1, \ldots, T$. When drawing the degrees of freedom parameter, we marginalize the mixing variables ζ_t out of the Metropolis-Hastings acceptance ratio.

1.3 Marginal distributions

1.3.1 MCMC algorithms

To estimate the univariate SV models for the marginals, we extend the MCMC algorithm of Omori, Chib, Shephard, and Nakajima (2007) to include skewed Student's t errors. For convenience, we drop the i sub-scripts.

- Draw the regression and skewness parameters (β_y, γ_y) jointly. Conditional on the observed data, $\delta_{1:T}$, and $h_{1:T}$, this is simply a regression model with heteroskedasticity.
- Calculate $y_t^* = \log\left(\left(\frac{y_t - W_t\beta_y - \gamma_y\delta_t}{\sqrt{\delta_t}}\right)^2\right)$. Draw the discrete indicator variables $s_{1:T}$ that are part of the 10 component mixture approximation, see Omori et al. (2007).
- Conditional on the indicator variables $s_{1:T}$, draw (ϕ_h, σ^2, ρ) jointly while marginalizing over $(h_{1:T}, \mu_h)$. As in Omori et al. (2007), we form a linear, Gaussian state space model and use the Kalman filter to maximize the log-posterior to find the mode. Using the independence MH algorithm, we draw a proposal $(\phi_h^*, \sigma^{2*}, \rho^*) \sim q\left(\phi_h, \sigma^2, \rho\right)$ from a Student's t distribution with 5 degrees of freedom, mean equal to the mode, and scale equal to the inverse Hessian at the mode. We accept this draw with probability

$$\alpha = \frac{p(y_{1:T}^*|W_{1:T}, s_{1:T}, \delta_{1:T}, \psi^*)p\left(\phi_h^*, \sigma^{2*}, \rho^*\right) q\left(\phi_h, \sigma^2, \rho\right)}{p(y_{1:T}^*|W_{1:T}, s_{1:T}, \delta_{1:T}, \psi)p\left(\phi_h, \sigma, \rho\right) q\left(\phi_h^*, \sigma^{2*}, \rho^*\right)} \tag{1}$$

where $p(y_{1:T}^*|s_{1:T}, \delta_{1:T}, \psi^*)$ is the likelihood calculated by the Kalman filter and $p\left(\phi_h, \sigma, \rho\right)$

is the prior.

- Draw $(h_{1:T}, \mu_h)$ by forming a linear, Gaussian state space model and using the simulation smoothing algorithm of Durbin and Koopman (2002).

- Draw the degrees of freedom ν without conditioning on $\delta_{1:T}$. Conditional on $h_{1:T}$, the conditional likelihood $p\left(y_t|W_t, h_t, \theta\right)$ is a generalized hyperbolic skewed Student's t distribution. In models with no skewness γ_y, this reduces to a Student's t distribution. We maximize the log-posterior to find the mode and inverse Hessian at the mode. Using the independence MH algorithm, we draw a proposal $\nu^* \sim q\left(\nu\right)$ from a Student's t distribution with 5 degrees of freedom, mean equal to the mode, and scale equal to the inverse Hessian at the mode. We accept this draw with probability

$$\alpha = \frac{p(y_{1:T}|W_{1:T}, h_{1:T}, \psi^*)p\left(\nu^*\right)q\left(\nu\right)}{p(y_{1:T}|W_{1:T}, h_{1:T}, \psi)p\left(\nu\right)q\left(\nu^*\right)} \tag{2}$$

where $p\left(\nu\right)$ is the prior. In practice, we optimize under the constraint that $\nu > 3$ and propose values only in this region.

- Draw δ_t for $t = 1, \ldots, T$ from a generalized inverse Gaussian (GIG) distribution, $\mathrm{GIG}\left(\alpha_t, \chi_t, \psi_t\right)$, where $\alpha_t = -\frac{\nu+1}{2}$, $\chi_t = \frac{(y_t - W_t\beta_y)^2}{\exp(h_t)} + \nu$, and $\psi_t = \frac{\gamma_y^2}{\exp(h_t)}$. In models with no skewness when $\gamma_y = 0$, this simplifies to an inverse Gamma distribution.

1.3.2 Particle filtering algorithms for SV models

We run the particle filter to calculate the probability integral transforms $u_t = F\left(y_t|y_{1:t-1}, W_{1:t-1}, \psi\right)$, where we have dropped the i subscripts for convenience.

For $t = 1, \ldots, T$, run:

- For $m = 1, \ldots, M$, draw from a proposal distribution: $h_t^{(m)} \sim q\left(h_t|h_{t-1}^{(m)}, y_{1:t}, W_{1:t}, \psi\right)$.

- For $m = 1, \ldots, M$, calculate the PIT: $\hat{u}_t^{(m)} = \hat{w}_{t-1}^{(m)} \frac{F\left(Y_t < y_t|W_t, h_t^{(m)}, \psi\right) f\left(h_t^{(m)}|h_{t-1}^{(m)}, \psi\right)}{q\left(h_t|h_{t-1}^{(m)}, y_{1:t}, W_{1:t}, \psi\right)}$.

- Calculate the PIT: $\hat{u}_t = \sum_{m=1}^{M} \hat{u}_t^{(m)}$
- For $m = 1, \ldots, M$, calculate the importance weight: $w_t^{(m)} \propto \hat{w}_{t-1}^{(m)} \frac{f\left(y_t|W_t, h_t^{(m)}, \psi\right) f\left(h_t^{(m)}|h_{t-1}^{(m)}, \psi\right)}{q\left(h_t^{(m)}|h_{t-1}^{(m)}, y_t, W_t, \psi\right)}$
- For $m = 1, \ldots, M$, normalize the weights: $\hat{w}_t^{(m)} = \frac{w_t^{(m)}}{\sum_{m=1}^{M} w_t^{(m)}}$.
- Calculate the effective sample size: $\mathrm{ESS}_t = \frac{1}{\sum_{m=1}^{M}\left(\hat{w}_t^{(m)}\right)^2}$
- If $\mathrm{ESS}_t < 0.5M$ resample $\left\{h_t^{(m)}\right\}_{m=1}^{M}$ with probabilities $\left\{\hat{w}_t^{(m)}\right\}_{m=1}^{M}$ and set $\hat{w}_t = 1/M$.

At time $t = 1$, the initial proposal distribution $q\left(h_1|y_1, W_1, \psi\right)$ does not depend on any previous particles.

2 Models and priors

2.1 Priors for the univariate SV models

For the univariate SV models, we use the same priors for all series and drop the i subscript for convenience. We let $\mu_h \sim \mathrm{N}\left(0.5, 2\right)$, $\sigma_h^2 \sim \text{Inv-Gamma}\left(5, 0.5\right)$, $\phi_h \sim \text{beta}\left(50, 2\right)$, ρ is uniform over $(-1, 1)$, β_y and γ_y are jointly multivariate normal with mean zero and covariance matrix equal to the identity matrix. Finally, the degrees of freedom is $\nu_y = \bar{\nu}_y + 3$ where $\bar{\nu}_y \sim \text{Gamma}\left(3.5, 2\right)$.

2.2 Conditional copula densities

The density of the Gaussian copula is

$$p\left(u_t|\Lambda_t, X_t, \theta\right) = \left|R_t\right|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} x_t' \left[R_t^{-1} - I\right] x_t\right),$$

where $x_{it} = \Phi^{-1}\left(u_{it}\right)$. The density of the Student's t copula is

$$p\left(u_t|\Lambda_t, X_t, \theta\right) = \left|R_t\right|^{-\frac{1}{2}} \frac{\Gamma\left(\frac{\nu+n}{2}\right)\Gamma\left(\frac{\nu}{2}\right)^{n-1}}{\Gamma\left(\frac{\nu+1}{2}\right)^{n}} \left(1 + \frac{x_t' R_t^{-1} x_t}{\nu}\right)^{-\frac{(\nu+n)}{2}} \prod_{i=1}^{n} \left(1 + \frac{x_{it}^2}{\nu}\right)^{\frac{\nu+1}{2}},$$

where $x_{it} = T^{-1}\left(u_{it}|\nu\right)$. The grouped Student's t copula does not have a closed from conditional density even when $R_t = R$.

2.3 Stochastic equicorrelation models

In a copula with a stochastic equicorrelation matrix, parameterizing the model directly in terms of the factors driving the matrix imposes the minimal number of restrictions on the flexibility of the correlation matrices. This model is defined as

$$u_{it} = P\left(\varepsilon_{it}|\theta\right), \qquad x_t = R_t^{-\frac{1}{2}}\varepsilon_t, \qquad \varepsilon_t \sim p\left(\varepsilon_t|\theta\right),$$

where the distribution of ε_t can be Gaussian, Student's t, skewed Student's t, etc. In the one-block equicorrelation model, the state variable Λ_t is one dimensional and the correlation matrix evolves as

$$\begin{aligned} R_t &= \left(1-\rho_t\right) I_n + \rho_t J_{n\times n}, \qquad \rho_t = a + \frac{(1-a)}{1+\exp\left(-\Lambda_t\right)}, \qquad a = \frac{-1}{n-1}, \\ \Lambda_{t+1} &= \mu + \Phi_\lambda\left(\Lambda_t - \mu\right) + Q\ \eta_t, \qquad \eta_t \sim \mathrm{N}\left(0, \Sigma\right), \end{aligned}$$

where $J_{n\times n}$ is a matrix of ones. The restriction that $\rho_t \in \left(-\frac{1}{n-1}, 1\right)$ ensures that R_t is positive definite.

In the two-block equicorrelation model, the state vector is has dimension 3×1 with elements $\Lambda_t = (\lambda_{1t}, \lambda_{2t}, \lambda_{3t})$. The correlation matrix evolves as

$$
\begin{aligned}
R_t &= \begin{pmatrix} (1-\rho_{11,t})\, I_{n_1} & 0 \\ 0 & (1-\rho_{22,t})\, I_{n_2} \end{pmatrix} + \begin{pmatrix} \rho_{11,t} J_{n_1 \times n_1} & \rho_{12,t} J_{n_1 \times n_2} \\ \rho_{12,t} J_{n_2 \times n_1} & \rho_{22,t} J_{n_2 \times n_2} \end{pmatrix}, \\
\rho_{11,t} &= a_{11} + \frac{(1-a_{11})}{1+\exp(-\lambda_{1t})}, \qquad a_{11} = -\frac{1}{n_1 - 1}, \\
\rho_{22,t} &= a_{22} + \frac{(1-a_{22})}{1+\exp(-\lambda_{2t})}, \qquad a_{22} = -\frac{1}{n_2 - 1}, \\
\rho_{12,t} &= a_{12} + \frac{(b_{12}-a_{12})}{1+\exp(-\lambda_{3t})}, \qquad a_{12} = -\sqrt{\frac{(\rho_{11,t}(n_1-1)+1)(\rho_{22,t}(n_2-1)+1)}{n_1 n_2}}, \\
& \qquad b_{12} = \sqrt{\frac{(\rho_{11,t}(n_1-1)+1)(\rho_{22,t}(n_2-1)+1)}{n_1 n_2}}, \\
\Lambda_{t+1} &= \mu + \Phi_\lambda (\Lambda_t - \mu) + Q\, \eta_t, \qquad \eta_t \sim \mathrm{N}(0, \Sigma),
\end{aligned}
$$

where n_i are the number of observations in block i. The restrictions on $\rho_{11,t}, \rho_{12,t}$, and $\rho_{22,t}$ ensure that R_t is positive definite. Expressions for the determinant $|R_t|$, inverse R_t^{-1}, and quadratic forms $x_t' R_t^{-1} x_t$ can be found in Engle and Kelly (2012).

References

Durbin, J. and S. J. Koopman (2002). A simple and efficient simulation smoother for state space time series analysis. *Biometrika 89*(3), 603–616.

Engle, R. F. and B. Kelly (2012). Dynamic equicorrelation. *Journal of Business and Economic Statistics 30*(2), 212–228.

Omori, Y., S. Chib, N. Shephard, and J. Nakajima (2007). Stochastic volatility with leverage: fast and efficient likelihood inference. *Journal of Econometrics 140*(2), 425–449.