Transform MCMC schemes for sampling intractable factor copula models

Rodrigo Targino[1]

School of Applied Mathematics (EMAp), Fundação Getulio Vargas (FGV)

20/May/2021

[1] Joint work with Cyril Bénézet and Emmanuel Gobet

Introduction

- In financial and actuarial risk management, modelling **dependency** within a random vector $\mathcal{X}$ is crucial
- A standard approach is the use of a **copula model**
- **Drawback:** Most parametric copulas are not suitable for high dimensional applications
- Generic statistics of interest, for $\mathcal{X} \sim C(F^{(1)}, \dots, F^{(d)})$

$$\mathbb{E}\left(g(\mathcal{X})\right) \quad \text{and} \quad \mathbb{E}\left(g(\mathcal{X}) \mid \mathcal{X} \in A\right).$$

- If $\{\mathcal{X} \in A\}$ is a rare event (e.g. tail event), i.i.d. MC sampling is **inefficient**
- **MCMC sampling** may be helpful.

Examples

- **Tail dependence**: (McNeil, Frey, Embrechts, ('05))

$$\lambda^u_{i,\mathcal{I}} := \lim_{\alpha \to 1^-} \mathbb{P}\left[\mathcal{X}^{(i)} > \mathrm{VaR}_\alpha\left(\mathcal{X}^{(i)}\right) \mid \forall j \in \mathcal{I}, \mathcal{X}^{(j)} > \mathrm{VaR}_\alpha\left(\mathcal{X}^{(j)}\right)\right]$$

- **Semi-correlation** (Ang and Chen ('02), Gabbi ('05)):

$$\rho^+_{i,j} := \mathrm{Cor}\left(\mathcal{X}^{(i)}, \mathcal{X}^{(j)} \mid \mathcal{X}^{(i)} > 0, \mathcal{X}^{(j)} > 0\right),$$
$$\rho^-_{i,j} = \mathrm{Cor}\left(\mathcal{X}^{(i)}, \mathcal{X}^{(j)} \mid \mathcal{X}^{(i)} < 0, \mathcal{X}^{(j)} < 0\right)$$

- ***k*-expected shortfall** (Oh and Patton ('17)):

$$(k - \mathrm{ES})^{(i)} = \mathbb{E}\left(\mathcal{X}^{(i)} \middle| \left(\sum_{j=1}^{d} \mathbf{1}_{\{\mathcal{X}^{(j)} \geq C\}}\right) > k\right)$$

Factor copulas

- Oh and Patton ('17): use as copula C the copula of an **auxiliary vector** $\mathcal{Y} = \Phi(\mathcal{Z})$, with $\Phi : \mathbb{R}^D \to \mathbb{R}^d$
  - $\mathcal{Z} := (\mathcal{M}^{(1)}, \ldots, \mathcal{M}^{(J)}, \epsilon^{(1)}, \ldots, \epsilon^{(d)})$ with $D = J + d$
  - $\mathcal{M} = (\mathcal{M}^{(1)}, \ldots, \mathcal{M}^{(J)})$ (factors)
  - $\epsilon = (\epsilon^{(1)}, \ldots, \epsilon^{(d)})$ (idiosyncratic errors)
  - $(\mathcal{M}^{(1)}, \ldots, \mathcal{M}^{(J)}, \epsilon^{(1)}, \ldots, \epsilon^{(d)})$ indep. and $(\epsilon^{(i)})_{i=1}^{d}$ i.i.d.
- **Example** (linear factor copula):
  - $\mathcal{Y}^{(i)} = \mathcal{M} + \epsilon^{(i)}$
  - $\mathcal{M} \sim \text{skew t}(\nu, \lambda)$
  - $\epsilon^{(i)} \overset{iid}{\sim} t(\nu)$
- **Notation:**
  - $\mathcal{Y} \sim C(G^{(1)}, \ldots, G^{(d)})$
  - $\mathcal{X} \sim C(F^{(1)}, \ldots, F^{(d)})$

Factor copulas

The problem:

- $\mathcal{X} = (\mathcal{X}_1 \ldots, \mathcal{X}_k) \sim C(F^{(1)}, \ldots, F^{(d)})$
- $\mathbb{E}\left(g(\mathcal{X}) \mid \mathcal{X} \in A\right) \approx \frac{1}{n} \sum_{k=1}^{n} g(\mathcal{X}_k)$
- Convergence rate?

Example: (Oh and Patton ('17))

- Model for the **losses** of the stocks in the S&P 100:
  - $\mathcal{Y}^{(i)} = \beta_{S(i)} \mathcal{M}^{(0)} + \gamma_{S(i)} \mathcal{M}^{(S(i))} + \epsilon^{(i)}$
  - $\mathcal{M}^{(0)} \sim$ skew $\mathrm{t}(\nu, \lambda)$,
  - $\mathcal{M}^{(S)} \overset{iid}{\sim} t(\nu),\ S = 1, \ldots, J-1,$ with $\mathcal{M}^{(S)} \perp\!\!\!\perp \mathcal{M}^{(0)}$,
  - $\epsilon^{(i)} \overset{iid}{\sim} t(\nu),\ i = 1, \ldots, d,\ \epsilon^{(i)} \perp\!\!\!\perp \mathcal{M}^{(j)}$
- Compute the $(k - ES)^{(i)}$:

$$\mathbb{E}\left(\mathcal{X}^{(i)} \,\middle|\, \mathcal{X}^{(1)} > 1\%, \ldots, \mathcal{X}^{(d)} > 1\%\right)$$

How to sample $\mathcal{X}$?

Algorithm 1: Usual sampling of $\mathcal{X}$ through sampling of $\mathcal{Z}$

1 Sample $\mathcal{Z} = (\mathcal{M}, \epsilon)$

2 Compute $\mathcal{Y} = \Phi(\mathcal{Z})$

3 Get $U = (U_1, \ldots, U_d) = (G^{(1)}(\mathcal{Y}^{(1)}), \ldots, G^{(d)}(\mathcal{Y}^{(d)}))$

4 Set $\mathcal{X}^{(i)} = (F^{(i)})^{-1}(U_i)$

$$\mathcal{Z} = \begin{bmatrix} \mathcal{Z}^{(1)} \\ \vdots \\ \mathcal{Z}^{(D)} \end{bmatrix} \overset{\Phi}{\rightarrow} \mathcal{Y} = \begin{bmatrix} \mathcal{Y}^{(1)} = \Phi^{(1)}(\mathcal{Z}) \\ \vdots \\ \mathcal{Y}^{(d)} = \Phi^{(d)}(\mathcal{Z}) \end{bmatrix} \rightarrow U = \begin{bmatrix} U^{(1)} = G^{(1)}\left(\mathcal{Y}^{(1)}\right) \\ \vdots \\ U^{(d)} := G^{(d)}\left(\mathcal{Y}^{(d)}\right) \end{bmatrix} \rightarrow$$

$$\rightarrow \begin{bmatrix} \mathcal{X}^{(1)} = \left(F^{(1)}\right)^{-1}\left(U^{(1)}\right) \\ \vdots \\ \mathcal{X}^{(d)} = \left(F^{(d)}\right)^{-1}\left(U^{(d)}\right) \end{bmatrix} = \mathcal{X}$$

- **Infeasible** if $G^{(i)}$ is not known!

A feasible algorithm to compute $\mathbb{E}(g(\mathcal{X}))$

Algorithm 2: Sampling of $\mathcal{X}$ through approximate sampling of $\mathcal{Z}$ and approximation of $G^{(i)}$

Input: $(F^{(i)})^{-1}$ the quantile of $\mathcal{X}^{(i)}$, $\mathcal{Z}_0 \in \mathbb{R}^D$

Output: $\mathcal{X}_k = \left(\mathcal{X}_k^{(1)}, \ldots, \mathcal{X}_k^{(d)}\right)$ for $1 \leq k \leq n$.

for $k \leftarrow 1$ **to** n **do**

Sample $\mathcal{Z}_k$ from $\mathcal{P}(\mathcal{Z}_{k-1}, \cdot)$.

Compute $\mathcal{Y}_k = \Phi(\mathcal{Z}_k)$.

Approximate and mollify $G^{(i)}$ by

$$\tilde{G}_k^{(i)}(y) := \frac{1}{2\sqrt{k}} + \left(1 - \frac{1}{\sqrt{k}}\right)\left(\frac{1}{k}\sum_{\ell=1}^{k} \mathbf{1}_{\mathcal{Y}_\ell^{(i)} \leq y}\right).$$

Set $V_k^{(i)} := \tilde{G}_k^{(i)}(\mathcal{Y}_k^{(i)})$ and $V_k := (V_k^{(i)})_{i=1}^d$.

Set $\mathcal{X}_k^{(i)} := \left(F^{(i)}\right)^{-1}\left(V_k^{(i)}\right)$ and $\mathcal{X}_k := (\mathcal{X}_k^{(i)})_{i=1}^d$.

- We also define $W_k^{(i)} := G^{(i)}\left(\mathcal{Y}_k^{(i)}\right)$ and $W_k = \left(W_k^{(i)}\right)_{i=1}^d$.

Assumptions

1. The marginal c.d.f. $G^{(i)}$ of $\mathcal{Y}^{(i)}$ is **continuous**.
2. The transition kernel $\mathcal{P}$ defines a **geometrically ergodic** Markov Chain $(\mathcal{Z}_k : k \geq 0)$ with Lyapunov function $\mathcal{L}$
3. The initial point of $\mathcal{Z}_0 \in \mathcal{A}^Z$ is deterministic.
4. There exists $\boldsymbol{q_{\max} \in [-1, 0)}$ s.t. $\forall \boldsymbol{q} > q_{\max}$, the map $\left(G^{(i)} \circ \Phi^{(i)}\right)^q + \left(1 - G^{(i)} \circ \Phi^{(i)}\right)^q$ is **bounded in $\mathcal{L}$-norm**
5. The function $\varphi := g \circ \left((F^{(1)})^{-1}, \ldots, (F^{(i)})^{-1}, \ldots, (F^{(d)})^{-1}\right)$ is **locally Lipschitz**: there exists a **slowly varying** function $\ell : (0, 1] \to (0, \infty)$ at 0, and a parameter $\boldsymbol{0 \leq \alpha < -q_{\max}}$ s.t.

$$|\varphi(u) - \varphi(v)| \leq \sum_{i=1}^{d} \frac{\ell(u_i \wedge v_i)|u_i - v_i|}{(u_i \wedge v_i)^{\alpha+1}} + \sum_{i=1}^{d} \frac{\ell(1 - u_i \vee v_i)|u_i - v_i|}{(1 - u_i \vee v_i)^{\alpha+1}},$$

$$|\varphi(u)| \leq \sum_{i=1}^{d} \frac{\ell(u_i)}{u_i^{\alpha}} + \sum_{i=1}^{d} \frac{\ell(1 - u_i)}{(1 - u_i)^{\alpha}}.$$

Main results I

Theorem (Uniform convergence of the c.d.f. of $\mathcal{Y}$ in L_p-norm)

For any $p \geq 1$, $n \geq 1$ and $i \in \{1, \ldots, d\}$, we have

$$\left| \sup_{y \in \mathbb{R}} |\tilde{G}_n^{(i)}(y) - G^{(i)}(y)| \right|_p \leq C_p \; n^{-\frac{p}{2(p+1)}},$$

for some finite constant C_p.

Theorem (Strong approximation)

For all $\iota > 0$ and any $p \in [1, \frac{-q_{\max}}{\alpha})$, there exists a constant $C_{\iota,p} > 0$ such that, for any $n \geq 1$,

$$|\varphi(V_n) - \varphi(W_n)|_p \leq C_{p,\iota} n^{-\frac{1}{2p} + \frac{\alpha}{2|q_{\max}|} + \iota}.$$

Main results II

Corollary (Weak convergence)

For all $\iota > 0$*, there exists a constant* $C_\iota > 0$ *such that, for any* $n \geq 1$*,*

$$\left|\mathbb{E}\left(\varphi(V_n)\right) - \mathbb{E}\left(\varphi(U)\right)\right| \leq C_\iota n^{-\frac{1}{2}+\frac{\alpha}{2|q_{\max}|}+\iota}.$$

Corollary (Convergence of Monte Carlo averages)

For all $\iota > 0$ *and for any* $p \geq 1$ *satisfying* $p \vee 2 < \frac{|q_{\max}|}{\alpha}$*, there exists a positive constant* $C_{p,\iota}$ *such that for any* $n \geq 1$*,*

$$\left|\frac{1}{n}\sum_{k=1}^{n}\varphi(V_k) - \mathbb{E}\left(\varphi(U)\right)\right|_p \leq C_{p,\iota} n^{-\frac{1}{2p}+\frac{\alpha}{2|q_{\max}|}+\iota}.$$

Algorithm 3: sampling of $\mathcal{X} \mid A$ via sampling of $\mathcal{Z} \mid \mathcal{Z} \in \mathcal{A}^Z$ and $\mathcal{Z} \mid \mathcal{Z} \in (\mathcal{A}^Z)^c$

Input: $(F^{(i)})^{-1}$ the quantile of $\mathcal{X}^{(i)}$, $\mathcal{Z}_{0,A} \in \mathcal{A}^Z$, $\mathcal{Z}_{0,A^c} \in (\mathcal{A}^Z)^c$

Output: $\mathcal{X}_k = \left(\mathcal{X}_k^{(1)}, \ldots, \mathcal{X}_k^{(d)}\right)$ for $1 \leq k \leq n$.

for $k \leftarrow 1$ **to** n **do**

Sample $\mathcal{Z}_{k,A}$ from $\mathcal{P}(\mathcal{Z}_{k-1,A}, \cdot)$ and **accept if in** $\mathcal{A}^Z$.

Compute $\mathcal{Y}_{k,A} = \Phi(\mathcal{Z}_{k,A})$.

Sample $\mathcal{Z}_{k,A^c}$ from $\mathcal{P}(\mathcal{Z}_{k-1,A^c}, \cdot)$ and **accept if in** $(\mathcal{A}^Z)^c$.

Compute $\mathcal{Y}_{k,A^c} = \Phi(\mathcal{Z}_{k,A^c})$.

Approximate and mollify $G^{(i)}$ by

$$\tilde{G}_k^{(i)}(y) := \frac{1}{2\sqrt{k}} + \left(1 - \frac{1}{\sqrt{k}}\right)\left(\left(\frac{1}{k}\sum_{\ell=1}^{k} \mathbf{1}_{\mathcal{Y}_{\ell,A}^{(i)} \leq y}\right)\mathbb{P}[A] + \left(\frac{1}{k}\sum_{\ell=1}^{k} \mathbf{1}_{\mathcal{Y}_{\ell,A^c}^{(i)} \leq y}\right)\mathbb{P}[A^c]\right).$$

Set $V_k^{(i)} := \tilde{G}_k^{(i)}(\mathcal{Y}_{k,A}^{(i)})$ and $V_k := (V_k^{(i)})_{i=1}^d$.

Set $\mathcal{X}_k^{(i)} := \left(F^{(i)}\right)^{-1}\left(V_k^{(i)}\right)$ and $\mathcal{X}_k := (\mathcal{X}_k^{(i)})_{i=1}^d$.

Example: The statistic

k-Expected Shortfall

- $\mathcal{X}^{(i)}$ denote the **losses** of the i-th stock
- We model the assets in the **S&P 100** index ($d = 100$)
- Our interest is to compute the **k-ES**

$$(k - \mathrm{ES})^{(i)} = \mathbb{E}\left(\mathcal{X}^{(i)} \middle| \mathcal{X}^{(1)} > C, \ldots, \mathcal{X}^{(d)} > C\right).$$

  with $C = 1\%$.
- We estimate $\mathbb{P}[A] \approx 1.42 \times 10^{-4}$ using a **crude MC** procedure for $\mathcal{Y}$, with sample size 10^6

Example: The model

Linear Factor Copula

- $\mathcal{X}^{(i)} \sim t_{\nu_i}(m_i, s_i)$, (marginal stock loss)
- $\mathcal{Y}^{(i)} = \beta_{S(i)}\mathcal{M}^{(0)} + \gamma_{S(i)}\mathcal{M}^{(S(i))} + \epsilon^{(i)}$ with

$$S(i) \in \{1, \ldots, 7\} \quad \text{(industry group)}$$

$$\mathcal{M}^{(0)} \sim \text{skew t}(\nu, \lambda) \quad \text{(market-wide factor)}$$

$$\mathcal{M}^{(S)} \overset{iid}{\sim} t(\nu), \quad \text{(sector specific factor)}$$

$$\epsilon^{(i)} \overset{iid}{\sim} t(\nu), \quad \text{(idiosyncratic noise)}$$

and $\mathcal{M}^{(0)}, \mathcal{M}^{(S)}, \epsilon^{(i)}$ are independent.

Example: The sampler

Random Walk Metropolis

- We sample $\mathcal{Z} = (\mathcal{Z}^{(1)}, \ldots, \mathcal{Z}^{(D)})$ using a Markov Chain whose stationary distribution $\pi_{\mathcal{A}^{\mathcal{Z}}}(z)\mathrm{d}z$ is Gaussian restricted to $\mathcal{A}^{\mathcal{Z}}$
- We use the RWM sampler with **Gaussian proposals** for $\mathcal{Z}$
- Moreover,

$$\begin{aligned}
\mathcal{M}^{(0)} &:= G_{\nu,\lambda}^{-1} \circ F_{\mathcal{N}}(\mathcal{Z}^{(1)}) \\
\mathcal{M}^{(i)} &:= G_{\nu}^{-1} \circ F_{\mathcal{N}}(\mathcal{Z}^{(i+1)}), \quad \text{for } i = 1, \ldots, J-1 \\
\epsilon^{(i)} &:= G_{\nu}^{-1} \circ F_{\mathcal{N}}(\mathcal{Z}^{(i+J)}), \quad \text{for } i = 1, \ldots, d
\end{aligned}$$

- Also, $\mathcal{Y} = \Phi(\mathcal{Z}) = \left(\Phi^{(i)}(\mathcal{Z})\right)_{i=1}^{d}$ with, for $1 \leq i \leq d$,

$$\begin{aligned}
\Phi^{(i)} &: \mathbb{R}^D \to \mathbb{R}, \\
&z \mapsto \beta_{S(i)} G_{\nu,\lambda}^{-1} \circ F_{\mathcal{N}}(z^{(1)}) + \gamma_{S(i)} G_{\nu}^{-1} \circ F_{\mathcal{N}}(z^{(S(i)+1)}) \\
&\qquad + G_{\nu}^{-1} \circ F_{\mathcal{N}}(z^{(i+J)}) \, .
\end{aligned}$$

Example: Results

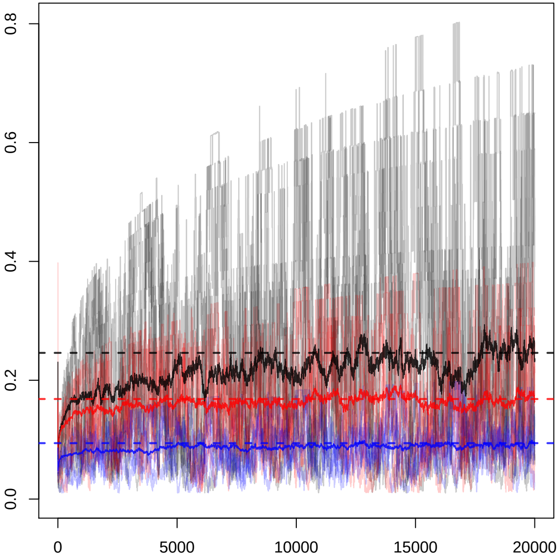

Figure: Black, red and blue: different marginals. Solid colors: average across M chains. Light colors: individual chains.

Conclusion

- We studied the **theoretical** and **numerical** properties of a **transform MCMC** scheme
- This scheme is developed to efficiently compute expectations, conditional to rare events, in which the unconditional distribution is given by an **factor copula**
- Under mild and natural hypotheses, we are able to derive the **convergence rates** for our proposed estimators
- We also revisit the computation of a challenging statistic originated in the **financial risk management** literature.

Thank you for the attention!

Bibliography

[1] Ang, A. and Chen, J. (2002). Asymmetric correlations of equity portfolios. Journal of financial Economics, 63(3):443–494.

[2] Gabbi, G. (2005). Semi-correlations as a tool for geographical and sector asset allocation. European Journal of Finance, 11(3):271–281.

[3] McNeil, A. J., Frey, R., and Embrechts, P. (2010). Quantitative Risk Management: Concepts, Techniques, and Tools. Princeton University Press.

[4] Oh, D. and Patton, A. (2017). Modeling dependence in high dimensions with factor copulas. Journal of Business & Economic Statistics, 35(1):139–154.