Factor Copula Models for Replicated Spatial Data

Supplementary Material

Pavel Krupskii, Raphaël Huser, Marc G. Genton[1]

November 11, 2016

Some key words: copula; heavy tails; non-Gaussian random field; spatial statistics; tail asymmetry.

Short title: Factor Copula Models for Replicated Spatial Data

[1]CEMSE Division, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia. E-mail: pavel.krupskiy@kaust.edu.sa, raphael.huser@kaust.edu.sa, marc.genton@kaust.edu.sa
This research was supported by the King Abdullah University of Science and Technology (KAUST).

S.1 Proof of Proposition 2 in Section 2.1

In the model, we have $F_1^{\mathbf{W}}(z) = \Phi(z) - \exp(\theta_1^2/2 - \theta_1 z)\Phi(z-\theta_1)\theta_2/(\theta_1+\theta_2) + \exp(\theta_2^2/2 + \theta_2 z)\Phi(-z-\theta_2)\theta_1/(\theta_1+\theta_2)$. As $z \to \infty$, we have: $F_1^{\mathbf{W}}(z) = 1 - \theta_2 \exp(\theta_1^2/2 - \theta_1 z)/(\theta_1+\theta_2) + o(\exp(-z^2/2))$. Define z_q^j as the solution to the equation:

$$1 - \theta_2 \exp\left(\theta_1^2/2 - \theta_1 z_q^j\right)/(\theta_1+\theta_2) = 1 - q x_j \quad (j = 1, 2).$$

It implies $z_q^j = \theta_1/2 + (1/\theta_1)\{\log\theta_2 - \log(\theta_1+\theta_2)\} - (\log q + \log x_j)/\theta_1$, and $F_1^{\mathbf{W}}(z_q^j) = 1 - q x_j + o\{q^{-(\log q)/2}\}$. The distribution for the common factor is $F_0(z) = \exp\{-\theta_2(-z)_+\}\{1 - \theta_2\exp(-\theta_1 z_+)/(\theta_1+\theta_2)\}$. Similar to the proof of Proposition 1, we find that

$$F_2^{\mathbf{W}}(z_q^1, z_q^2) = \psi(z_q^1, z_q^2) + \psi(z_q^2, z_q^1),$$

where

$$\begin{aligned}
\psi(z_q^1, z_q^2) &= \int_{\mathbb{R}^1} F_0(z_q^1 - w)\phi(w)\Phi\left\{\left(\frac{1-\rho}{1+\rho}\right)^{1/2}(w-K)\right\}\mathrm{d}w \\
&= \Phi_K - \frac{\theta_2\exp(-\theta_1 z_q^1 - K^2/2 + \theta_1 K + (\theta_1 - K)^2/2)}{\theta_1+\theta_2}\Phi\left\{(\theta_1 - K)\left(\frac{1-\rho}{2}\right)^{1/2}\right\} + o(q^{-\log q/2}) \\
&= \Phi_K - q x_1 \Phi\left\{(\theta_1 - K)\left(\frac{1-\rho}{2}\right)^{1/2}\right\} + o(q^{-\log q/2}) \quad \left(K = \frac{z_q^1 - z_q^2}{1-\rho} = \frac{\log(x_2/x_1)}{\theta_1(1-\rho)}\right),
\end{aligned}$$

where $\Phi_K = \Phi\left(-K\sqrt{\frac{1-\rho}{2}}\right)$. It implies that $\lim_{q\to 0}\ell_q(x_1, x_2) = x_1\Phi\{\lambda/2 + \log(x_1/x_2)/\lambda\} + x_2\Phi\{\lambda/2 + \log(x_2/x_1)/\lambda\}$ where $\lambda = \theta_1\{2(1-\rho)\}^{1/2}$. □

S.2 The gradient of the log-likelihood (5) in Section 3.1

We obtain derivatives of the log-likelihood with respect to model parameters. Let f_{V_0} be the pdf of V_0, $\mathbf{z}_i = (z_{i1}, \ldots, z_{in})^{\mathrm{T}}$, and $z_{ij} = (F_1^{\mathbf{W}})^{-1}(u_{ij}; \boldsymbol{\theta}_F)$, $i = 1, \ldots, n$ and $j = 1, \ldots, d$. The original data can be transformed to uniform scores u_{ij} either non-parametrically or using the integral transform as explained in Section 3.1. We have:

$$\frac{\partial \ell(\mathbf{y}_1, \ldots, \mathbf{y}_N)}{\partial \boldsymbol{\theta}} = \sum_{i=1}^{N} \frac{\partial f_n^{\mathbf{W}}(\mathbf{z}_i; \boldsymbol{\theta}_F, \boldsymbol{\theta}_{\boldsymbol{\Sigma}})/\partial\boldsymbol{\theta}}{f_n^{\mathbf{W}}(\mathbf{z}_i; \boldsymbol{\theta}_F, \boldsymbol{\theta}_{\boldsymbol{\Sigma}})} - \sum_{i=1}^{N}\sum_{j=1}^{n} \frac{\partial f_1^{\mathbf{W}}(z_{ij}; \boldsymbol{\theta}_F)/\partial\boldsymbol{\theta}}{f_1^{\mathbf{W}}(z_{ij}; \boldsymbol{\theta}_F)}, \ \boldsymbol{\theta} = (\boldsymbol{\theta}_F^{\mathrm{T}}, \boldsymbol{\theta}_{\boldsymbol{\Sigma}}^{\mathrm{T}})^{\mathrm{T}}. \quad (1)$$

It follows that $\partial f_n^{\mathbf{W}}(\mathbf{z}_i;\boldsymbol{\theta}_F,\boldsymbol{\theta}_{\boldsymbol{\Sigma}})/\partial\boldsymbol{\theta} = \int_{-\infty}^{\infty}\mathbb{F}_n^{\mathbf{W}}(\mathbf{z}_i,v_0;\boldsymbol{\theta}_F,\boldsymbol{\theta}_{\boldsymbol{\Sigma}})\mathrm{d}v_0$, $\mathbb{F}_n^{\mathbf{W}} = (f_{F,n}^{\mathbf{W}}, f_{\boldsymbol{\Sigma},n}^{\mathbf{W}})^{\mathrm{T}}$, and

$$f_{F,n}^{\mathbf{W}}(\mathbf{z}_i,v_0;\boldsymbol{\theta}_F,\boldsymbol{\theta}_{\boldsymbol{\Sigma}}) = \sum_{j=1}^{n}\partial^j\phi_{\boldsymbol{\Sigma}_{\mathbf{Z}}}(\mathbf{z}_i - v_0)\cdot\frac{\partial(F_1^{\mathbf{W}})^{-1}(u_{ij};\boldsymbol{\theta}_F)}{\partial\boldsymbol{\theta}_F}f_{V_0}(v_0;\boldsymbol{\theta}_F) + \phi_{\boldsymbol{\Sigma}_{\mathbf{Z}}}(\mathbf{z}_i - v_0)\cdot\frac{\partial f_{V_0}(v_0;\boldsymbol{\theta}_F)}{\partial\boldsymbol{\theta}_F},$$

$$f_{\boldsymbol{\Sigma},n}^{\mathbf{W}}(\mathbf{z}_i,v_0;\boldsymbol{\theta}_F,\boldsymbol{\theta}_{\boldsymbol{\Sigma}}) = \frac{\partial\phi_{\boldsymbol{\Sigma}_{\mathbf{Z}}}(\mathbf{z}_i - v_0)}{\partial\boldsymbol{\theta}_{\boldsymbol{\Sigma}}}\cdot f_{V_0}(v_0;\boldsymbol{\theta}_F) = -\frac{1}{2}\tilde{\mathbf{z}}_i^{\mathrm{T}}\frac{\partial\boldsymbol{\Sigma}_{\mathbf{Z}}}{\partial\boldsymbol{\theta}_{\boldsymbol{\Sigma}}}\tilde{\mathbf{z}}_i\cdot\phi_{\tilde{\boldsymbol{\Sigma}}_{\mathbf{Z}}}(\mathbf{z}_i - v_0)\cdot f_{V_0}(v_0;\boldsymbol{\theta}_F) - \frac{1}{2}\mathrm{tr}\left(\boldsymbol{\Sigma}_{\mathbf{Z}}^{-1}\frac{\partial\boldsymbol{\Sigma}_{\mathbf{Z}}}{\partial\boldsymbol{\theta}_{\boldsymbol{\Sigma}}}\right)\cdot\phi_{\boldsymbol{\Sigma}_{\mathbf{Z}}}(\mathbf{z}_i - v_0)\cdot f_{V_0}(v_0;\boldsymbol{\theta}_F),$$

where ∂^j denotes the derivative with respect to the j-th argument ($j = 1,\ldots,n$), $\tilde{\boldsymbol{\Sigma}}_Z := \partial\boldsymbol{\Sigma}_{\mathbf{Z}}^{-1}/\partial\boldsymbol{\theta}_{\boldsymbol{\Sigma}} = -\boldsymbol{\Sigma}_{\mathbf{Z}}\frac{\partial\boldsymbol{\Sigma}_{\mathbf{Z}}}{\partial\boldsymbol{\theta}_{\boldsymbol{\Sigma}}}\boldsymbol{\Sigma}_{\mathbf{Z}}$ and $\tilde{\mathbf{z}}_i = \boldsymbol{\Sigma}_{\mathbf{Z}}^{-1}(\mathbf{z}_i - v_0)^{\mathrm{T}}$. Also,

$$\frac{\partial f_1^{\mathbf{W}}(z_{ij};\boldsymbol{\theta}_F)}{\partial\boldsymbol{\theta}_F} = \partial^1 f_1^{\mathbf{W}}(z_{ij};\boldsymbol{\theta}_F)\cdot\frac{\partial(F_1^{\mathbf{W}})^{-1}(u_{ij};\boldsymbol{\theta}_F)}{\partial\boldsymbol{\theta}_F} + \left.\frac{\partial f_1^{\mathbf{W}}(z_{ij};\boldsymbol{\theta}^*)}{\partial\boldsymbol{\theta}^*}\right|_{\boldsymbol{\theta}^*=\boldsymbol{\theta}_F}.$$

It is easy to see that

$$\partial^j\phi_{\boldsymbol{\Sigma}_{\mathbf{Z}}}(\mathbf{z}_i - v_0) = -\phi_{\boldsymbol{\Sigma}_{\mathbf{Z}}}(\mathbf{z}_i - v_0)\cdot\tilde{z}_{ij}, \quad \frac{\partial(F_1^{\mathbf{W}})^{-1}(u_{ij};\boldsymbol{\theta}_F)}{\partial\boldsymbol{\theta}_F} = -\frac{\partial F_1^{\mathbf{W}}(z_{ij};\boldsymbol{\theta}_F)/\partial\boldsymbol{\theta}_F}{f_1^{\mathbf{W}}(z_{ij};\boldsymbol{\theta}_F)}.$$

Therefore, to calculate the gradient of $l(\mathbf{z}_1,\ldots,\mathbf{z}_n)$, the following derivatives are needed:

$$\frac{\partial f_{V_0}(v;\boldsymbol{\theta}_F)}{\partial\boldsymbol{\theta}_F},\ \frac{\partial F_1^{\mathbf{W}}(z;\boldsymbol{\theta}_F)}{\partial\boldsymbol{\theta}_F},\ \frac{\partial f_1^{\mathbf{W}}(z;\boldsymbol{\theta}_F)}{\partial z},\ \frac{\partial f_1^{\mathbf{W}}(z;\boldsymbol{\theta}_F)}{\partial\boldsymbol{\theta}_F},\ \frac{\partial\boldsymbol{\Sigma}_{\mathbf{Z}}}{\partial\boldsymbol{\theta}_{\boldsymbol{\Sigma}}}.$$

We use these analytical derivatives to calculate the gradient for the likelihood as given by (1). To compute the maximum likelihood estimates for the parameters, $\boldsymbol{\theta}_F, \boldsymbol{\theta}_{\boldsymbol{\Sigma}}$, we use the function `nlm()` from R statistical software to minimize the negative log-likelihood with a given gradient.

S.3 The log-likelihood for the exponential common factor model

We assume model (2) with $V_0 = V_1 - V_2$ where $V_1 \sim \mathrm{Exp}(\theta_1)$, $V_2 \sim \mathrm{Exp}(\theta_2)$ are independent exponential random factors. The likelihood function and its derivatives can be obtained in closed form in this case. One can check that, for any $\mathbf{z} = (z_1,\ldots,z_n)^{\mathrm{T}}$,

$$f_d^{\mathbf{W}}(z_1,\ldots,z_n;\boldsymbol{\theta}_F,\boldsymbol{\theta}_{\boldsymbol{\Sigma}}) = M(\mathbf{z};\boldsymbol{\theta}_F,\boldsymbol{\theta}_{\boldsymbol{\Sigma}}) \times \left[\exp\left\{\frac{-m_1+(m_1^*)^2}{2}\right\}\Phi(m_1^*) + \exp\left\{\frac{-m_1+(m_2^*)^2}{2}\right\}\Phi(-m_2^*)\right]$$

where

$$M(\mathbf{z};\boldsymbol{\theta}_F,\boldsymbol{\theta}_{\boldsymbol{\Sigma}}) = \frac{\theta_1\theta_2}{\theta_1+\theta_2}\cdot\frac{m_3^{-1/2}}{(2\pi)^{n-1}\det(\boldsymbol{\Sigma}_{\mathbf{Z}})^{1/2}},$$

$$m_1 = \mathbf{z}^{\mathrm{T}}\boldsymbol{\Sigma}_{\mathbf{Z}}^{-1}\mathbf{z}, \quad m_2 = \sum_{j=1}^{n}(\boldsymbol{\Sigma}_{\mathbf{Z}}^{-1}\mathbf{z})_j, \quad m_3 = \sum_{j_1,j_2=1}^{n}(\boldsymbol{\Sigma}_{\mathbf{Z}}^{-1})_{j_1,j_2}, \quad m_1^* = \frac{m_2-\theta_1}{(m_3)^{1/2}}, \quad m_2^* = \frac{m_2+\theta_2}{(m_3)^{1/2}}.$$

This implies that

$$\log f_d^{\mathbf{W}}(z_1,\ldots,z_n;\boldsymbol{\theta}_F,\boldsymbol{\theta}_{\boldsymbol{\Sigma}}) = M^*(\mathbf{z};\boldsymbol{\theta}_F,\boldsymbol{\theta}_{\boldsymbol{\Sigma}}) + \log\left[\exp\left\{\frac{(m_1^*)^2}{2}\right\}\Phi(m_1^*) + \exp\left\{\frac{(m_2^*)^2}{2}\right\}\Phi(-m_2^*)\right],$$

where the function $M^*(\mathbf{z};\boldsymbol{\theta}_F,\boldsymbol{\theta}_{\boldsymbol{\Sigma}}) = -(n-1)\log(2\pi) - 0.5\log\{\det(\boldsymbol{\Sigma}_{\mathbf{Z}})\} - 0.5\log m_3 - 0.5m_1 + \log\theta_1 + \log\theta_2 - \log(\theta_1+\theta_2)$. It is easy to see that for any parameter $\boldsymbol{\theta}$,

$$\frac{\partial \log f_d^{\mathbf{W}}(z_1,\ldots,z_n;\boldsymbol{\theta}_F,\boldsymbol{\theta}_{\boldsymbol{\Sigma}})}{\partial\boldsymbol{\theta}} = \frac{\partial M^*(\mathbf{z};\boldsymbol{\theta}_F,\boldsymbol{\theta}_{\boldsymbol{\Sigma}})}{\partial\boldsymbol{\theta}} + \frac{N(\mathbf{z};\boldsymbol{\theta}_F,\boldsymbol{\theta}_{\boldsymbol{\Sigma}})}{\exp\left\{\frac{(m_1^*)^2}{2}\right\}\Phi(m_1^*) + \exp\left\{\frac{(m_2^*)^2}{2}\right\}\Phi(-m_2^*)},$$

$$N(\mathbf{z};\boldsymbol{\theta}_F,\boldsymbol{\theta}_{\boldsymbol{\Sigma}}) = \left[m_1^*\frac{\partial m_1^*}{\partial\theta}\exp\left\{\frac{(m_1^*)^2}{2}\right\}\Phi(m_1^*) + m_2^*\frac{\partial m_2^*}{\partial\theta}\exp\left\{\frac{(m_2^*)^2}{2}\right\}\Phi(-m_2^*)\right] + (2\pi)^{-1/2}\left[\frac{\partial m_1^*}{\partial\theta} - \frac{\partial m_2^*}{\partial\theta}\right].$$

The chain rule can then be applied to calculate the gradient using the following equalities:

$$\frac{\partial m_1}{\partial\boldsymbol{\theta}_F} = 2\left(\frac{\partial\mathbf{z}}{\partial\boldsymbol{\theta}_F}\right)^{\mathrm{T}}\boldsymbol{\Sigma}_{\mathbf{Z}}^{-1}\mathbf{z}, \quad \frac{\partial m_1}{\partial\boldsymbol{\theta}_{\boldsymbol{\Sigma}}} = \mathbf{z}^{\mathrm{T}}\tilde{\boldsymbol{\Sigma}}_{\mathbf{Z}}\mathbf{z}, \quad \frac{\partial m_2}{\partial\boldsymbol{\theta}_F} = \sum_{i=1}^{n}\left(\boldsymbol{\Sigma}_Z^{-1}\frac{\partial\mathbf{z}}{\partial\boldsymbol{\theta}_F}\right)_i, \quad \frac{\partial m_2}{\partial\boldsymbol{\theta}_{\boldsymbol{\Sigma}}} = \sum_{i=1}^{n}\left(\tilde{\boldsymbol{\Sigma}}_{\mathbf{Z}}\mathbf{z}\right)_i,$$

$$\frac{\partial m_3}{\partial\boldsymbol{\theta}_{\boldsymbol{\Sigma}}} = \sum_{j_1,j_2=1}^{n}\left(\tilde{\boldsymbol{\Sigma}}_{\mathbf{Z}}\right)_{j_1,j_2}, \quad \frac{\partial z}{\partial\boldsymbol{\theta}_F} = \frac{\partial(F_1^{\mathbf{W}})^{-1}(\mathbf{u};\boldsymbol{\theta}_F)}{\partial\boldsymbol{\theta}_F} = -\frac{\partial F_1^{\mathbf{W}}(\mathbf{z};\boldsymbol{\theta}_F)/\partial\boldsymbol{\theta}_F}{f_1^{\mathbf{W}}(\mathbf{z};\boldsymbol{\theta}_F)}.$$

With $\boldsymbol{\theta}_F = (\theta_1,\theta_2)^{\mathrm{T}}$, we get:

$$\frac{\partial f_{V_0}(v;\boldsymbol{\theta}_F)}{\partial\theta_i} = \left(\frac{\theta_{3-i}}{\theta_1+\theta_2}\right)^2\exp\{-\theta_1 v^+ - \theta_2(-v)^+\} - \{(-1)^{i+1}v\}^+ f_{V_0}(v;\boldsymbol{\theta}_F) \quad (i=1,2).$$

Also denote $\xi_1 = \exp(\theta_1^2/2 - z\theta_1)\cdot\Phi(z-\theta_1)$, $\xi_2 = \exp(\theta_2^2/2 + z\theta_2)\cdot\Phi(-z-\theta_2)$. We have:

$$\frac{\partial F_1^{\mathbf{W}}(z;\boldsymbol{\theta}_F)}{\partial\theta_i} = \frac{\theta_{3-i}}{\theta_1+\theta_2}\left[\{z+(-1)^i\theta_i\}\xi_1 + (-1)^{i+1}\phi(z)\right] + \frac{\theta_{3-i}}{(\theta_1+\theta_2)^2}(\xi_1+\xi_2) \quad (i=1,2),$$

$$\frac{\partial f_1^{\mathbf{W}}(z;\boldsymbol{\theta}_F)}{\partial\theta_i} = \left(\frac{\theta_{3-i}}{\theta_1+\theta_2}\right)^2(\xi_1+\xi_2) + \frac{\theta_1\theta_2}{\theta_1+\theta_2}[\{\theta_i+(-1)^i z\}\xi_i - \phi(z)] \quad (i=1,2),$$

$$\frac{\partial f_1^{\mathbf{W}}(z;\boldsymbol{\theta}_F)}{\partial z} = \frac{\theta_1\theta_2}{\theta_1+\theta_2}(\theta_2\xi_2 - \theta_1\xi_1).$$