

Derivadas Matriciais - Modelo Reduzido

Leonardo Uchoa

Conteúdo do documento

Aqui estão as derivadas matriciais para que se obtenha o gradiente da log-verossimilhança do modelo em questão.

Formulação

Primeiramente é importante dizer que o modelo proposto é o que segue a seguinte função de covariância

$$\mathbf{\Sigma}_{\boldsymbol{\theta}}(\mathbf{h}) = \begin{pmatrix} C_{11}(\mathbf{h}) & C_{12}(\mathbf{h}) \\ C_{21}(\mathbf{h}) & C_{22}(\mathbf{h}) \end{pmatrix} \quad (1)$$

onde $\boldsymbol{\theta} = (\boldsymbol{\sigma}, \mathbf{a}, \boldsymbol{\nu}, \mu)$. Neste momento restringe-se ao modelo reduzido (parsimonioso) e portanto

$$\begin{aligned} C_{11}(\mathbf{h}) &= \sigma_1 M(\mathbf{h}|a, \nu_1) \\ C_{22}(\mathbf{h}) &= \sigma_2 M(\mathbf{h}|a, \nu_2) \\ C_{12}(\mathbf{h}) &= \rho_{12} \sigma_1 \sigma_2 M(\mathbf{h}|a, (\nu_1 + \nu_2)/2) \\ M(\mathbf{h}|\nu, a) &= \frac{2^{1-\nu} (ad)^\nu K_\nu(ad)}{\Gamma(\nu)} \end{aligned}$$

onde $d = \|\mathbf{h}\|$. Assim, ao assumirmos que o modelo é $\text{vec}(Y) \sim N(0, \mathbf{\Sigma})$, temos que a log-verossimilhança será

$$l(\boldsymbol{\theta}) = -1/2 (\log(|\mathbf{\Sigma}_{\boldsymbol{\theta}}|) + \mathbf{x}^t \mathbf{\Sigma}_{\boldsymbol{\theta}}^{-1} \mathbf{x}) \quad (2)$$

Derivadas

Fórmula Geral da Derivada da Log-verossimilhança

Ao derivarmos 2 em relação a qualquer elemento, θ , de $\boldsymbol{\theta}$, temos a expressão geral da derivada da log-verossimilhança:

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \theta} = \text{tr} \left(\mathbf{\Sigma}_{\boldsymbol{\theta}}^{-1} \frac{\partial \mathbf{\Sigma}_{\boldsymbol{\theta}}}{\partial \theta} \right) - \mathbf{x}^t \left[\mathbf{\Sigma}_{\boldsymbol{\theta}}^{-1} \frac{\partial \mathbf{\Sigma}_{\boldsymbol{\theta}}}{\partial \theta} \mathbf{\Sigma}_{\boldsymbol{\theta}}^{-1} \right] \mathbf{x}. \quad (3)$$

Então se $y = \mathbf{x} \mathbf{\Sigma}_{\boldsymbol{\theta}}^{-1}$, pela simetria da função de covariância Wittle-Matern (suposta no artigo de Gneiting & Kleiber), tem-se que

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \theta} = \text{tr} \left(\mathbf{\Sigma}_{\boldsymbol{\theta}}^{-1} \frac{\partial \mathbf{\Sigma}_{\boldsymbol{\theta}}}{\partial \theta} \right) - \mathbf{y}^t \left[\frac{\partial \mathbf{\Sigma}_{\boldsymbol{\theta}}}{\partial \theta} \right] \mathbf{y}. \quad (4)$$

Derivada das Funções de Covariâncias

Para obter $\partial \Sigma_{\theta} / \partial \theta$, onde $\theta = (\sigma_1^2, \sigma_2^2, a, \rho, \mu, \nu_1, \nu_2)$ (pois estamos no caso reduzido), vamos continuar a regra da cadeia passo a passo.

Derivada de Σ_{θ} c.r.a σ_1^2

$$\frac{\partial \Sigma_{\theta}(\mathbf{h})}{\partial \sigma_1^2} = \begin{pmatrix} M(\mathbf{h}|\nu_1, a) & \\ \frac{\rho \sigma_2}{2 \sigma_1} M(\mathbf{h}|\frac{\nu_1 + \nu_2}{2}, a) & \mathbf{O} \end{pmatrix} \quad (5)$$

Derivada de Σ_{θ} c.r.a σ_2^2

$$\frac{\partial \Sigma_{\theta}(\mathbf{h})}{\partial \sigma_2^2} = \begin{pmatrix} \mathbf{O} & \\ \frac{\rho \sigma_1}{2 \sigma_2} M(\mathbf{h}|\frac{\nu_1 + \nu_2}{2}, a) & M(\mathbf{h}|\nu_2, a) \end{pmatrix} \quad (6)$$

Derivada de Σ_{θ} c.r.a ρ

$$\frac{\partial \Sigma_{\theta}(\mathbf{h})}{\partial \sigma_2^2} = \begin{pmatrix} \mathbf{O} & \\ \sigma_1 \sigma_2 M(\mathbf{h}|\frac{\nu_1 + \nu_2}{2}, a) & \mathbf{O} \end{pmatrix} \quad (7)$$

Derivada de Σ_{θ} c.r.a a

Neste caso temos

$$\frac{\partial \Sigma_{\theta}(\mathbf{h})}{\partial a} = \begin{pmatrix} \sigma_1^2 \psi_1 & \rho \sigma_1^2 \sigma_2^2 \psi_3 \\ \rho \sigma_1^2 \sigma_2^2 \psi_3 & \sigma_2^2 \psi_2 \end{pmatrix} \quad (8)$$

em que ψ_k é a derivada de $\partial M(\mathbf{h}|\nu_k, a)$ c.r.a a para $k = 1, 2, 3$, onde $\nu_3 = (\nu_1 + \nu_2)/2$. Em seguida, temos que

$$\frac{\partial M(\mathbf{h}|\nu, a)}{\partial a} = \frac{2^{1-\nu} d^{\nu}}{\Gamma(\nu)} \left[\nu a^{\nu-1} K_{\nu}(ad) + a^{\nu} \frac{\partial K_{\nu}(ad)}{\partial a} \right] \quad (9)$$

e, como

$$\frac{\partial K_{\nu}(ad)}{\partial a} = d \left[\frac{\nu}{ad} K_{\nu}(ad) - K_{\nu+1}(ad) \right] \quad (10)$$

então

$$\frac{\partial M(\mathbf{h}|\nu, a)}{\partial a} = \frac{2^{1-\nu} d^{\nu}}{\Gamma(\nu)} \left[\nu a^{\nu-1} K_{\nu}(ad) + a^{\nu} d \left(\frac{\nu}{ad} K_{\nu}(ad) - K_{\nu+1}(ad) \right) \right]. \quad (11)$$

Ao simplificar a última equação, obtem-se

$$\frac{2^{1-\nu} d^{\nu} a^{\nu-1}}{\Gamma(\nu)} \left[2\nu K_{\nu}(ad) - ad K_{\nu+1}(ad) \right] \quad (12)$$