Biwm Calculations

Leonardo Uchoa

Conteúdo do documento

Neste documento estão presentes todas as contas feitas para a implementação da estimação por máxima verossimilhança do modelo Matérn Bivariado proposto por Gneiting. Aqui aproveita-se a estrutura de blocos da matriz de covariância cruzada para melhor estimar os parâmetros.

Formulação

Para o modelo (Cressie)

$$\mathbf{Y}^*(\mathbf{s}) = \mu(\mathbf{s}) + \eta(\mathbf{s}) + \epsilon(\mathbf{s}) \tag{1}$$

temos que $\mathbf{s} \in \mathbb{R}^2$ e que

- $\mu(\mathbf{s})$ é a variação determinística (de grande escala);
- $\eta(\mathbf{s})$ é o efeito aleatório espacial (variação de pequena escala, dependente da resolução dos dados);
- $\epsilon(\mathbf{s})$ é a variação em escala micro (não detectável na resolução (condicional) dos dados).

Agora adapta-se esta formulação para o cenário multivariado

$$\mathbf{Y}^*(\mathbf{s}) = \boldsymbol{\mu}(\mathbf{s}) + \boldsymbol{\eta}(\mathbf{s}) + \boldsymbol{\epsilon}(\mathbf{s}) \tag{2}$$

em que novamente $\mathbf{s} \in \mathbb{R}^2$, e

- $\boldsymbol{\mu}(\mathbf{s})$ é o vetor de variações determinísticas (de grande escala);
- $\boldsymbol{\eta}(\mathbf{s})$ é o vetor de efeitos aleatórios espaciais (variações de pequena escala, dependentes da resolução dos dados);
- $\boldsymbol{\epsilon}(\mathbf{s})$ é o vetor de variações em escala micro (não detectáveis na resolução (condicional) dos dados).

de forma que o processo $\mathbf{Y}$ é bivariado com a seguinte estrutura

$$\mathbf{Y}^*(\mathbf{s}) = \begin{pmatrix} Y_{11}(\mathbf{s}) & Y_{12}(\mathbf{s}) \\ \vdots & \vdots \\ Y_{N1}(\mathbf{s}) & Y_{N2}(\mathbf{s}) \end{pmatrix} \tag{3}$$

Neste ponto a estrateǵia é modelar as observações bivariadas do processo espacial por meio de uma Normal Bivariada. Ou seja, cada par de observações $\Big(Y_{i1}(\mathbf{s}), Y_{i2}(\mathbf{s})\Big)$ seria uma realização de $N_2(\boldsymbol{\mu}, \Sigma)$, indexada (de alguma maneira) no espaço. Para isto é interessante reformular a equação (3) em termos da função $\text{Vec}(\cdot)$:

$$\mathbf{Y}(\mathbf{s}) := Vec(\mathbf{Y}^*(\mathbf{s})) = \begin{pmatrix} Y_{11}(\mathbf{s}) \\ Y_{12}(\mathbf{s}) \\ \vdots \\ Y_{N1}(\mathbf{s}) \\ Y_{N2}(\mathbf{s}) \end{pmatrix} \tag{4}$$

em que atribui-se ao processo $\mathbf{Y}(\mathbf{s})$ a distribuição Normal Multivariada com vetor de funções médias $\boldsymbol{\mu}^T(\mathbf{s}) = (\mu_1(\mathbf{s}), \mu_2(\mathbf{s}))$ e função de covariância cruzada $\mathbf{\Sigma}(\mathbf{h})$ com estrutura dada por

$$\mathbf{\Sigma}(\mathbf{h}) = \begin{pmatrix} C_{11}(\mathbf{h}) & C_{12}(\mathbf{h}) \\ C_{21}(\mathbf{h}) & C_{22}(\mathbf{h}) \end{pmatrix} \tag{5}$$

Assume-se também estacionariedade de segunda ordem para este processo, onde as funções médias não dependem das localizações e as funções de covariâncias simples e cruzadas só dependem da distância entre as observações. Isto é

$$\begin{aligned} &E(Y_{i1}^*(\mathbf{s} + \mathbf{h})) = E(Y_{i1}(\mathbf{s})) = \mu_1 \\ &E(Y_{i2}^*(\mathbf{s} + \mathbf{h})) = E(Y_{i2}(\mathbf{s})) = \mu_2 \\ &E(Y_{i2}^*(\mathbf{s} + \mathbf{h})Y_{i2}^*(\mathbf{s})) = C_{11}(||\mathbf{s}\text{-}\mathbf{h}||) \\ &E(Y_{i1}^*(\mathbf{s} + \mathbf{h})Y_{i2}^*(\mathbf{s})) = C_{22}(||\mathbf{s}\text{-}\mathbf{h}||) \\ &E(Y_{i2}^*(\mathbf{s} + \mathbf{h})Y_{i1}^*(\mathbf{s})) = C_{12}(||\mathbf{s}\text{-}\mathbf{h}||) \\ &E(Y_{i1}^*(\mathbf{s} + \mathbf{h})Y_{i2}^*(\mathbf{s})) = C_{21}(||\mathbf{s}\text{-}\mathbf{h}||) \end{aligned} \tag{6}$$

Neste momento postulam-se funções para estruturar $C_{ij}(\mathbf{s})$. Aqui as funções serão Matérn, formuladas de forma que a sua composição em $\mathbf{\Sigma}(\mathbf{h})$, dita $\mathbf{\Sigma}_{\boldsymbol{\theta}}(\mathbf{h})$ com $\boldsymbol{\theta} = (\boldsymbol{\sigma}, \mathbf{a}, \boldsymbol{\nu}, \mu)$, formam a função de covariância Matérn Bivariada. Neste momento restringe-se ao modelo reduzido (parsimonioso) que induz às funções de covarâncias e covariâncias cruzadas

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

onde $\boldsymbol{\theta} = (\sigma_1, \sigma_2, a, \mu_1, \mu_2)$. Toda esta formulação nos permite escrever a log-verossimilhança do processo como a de uma normal multivariada

$$l(\boldsymbol{\theta}) = -1/2\big(log(|\mathbf{\Sigma}_{\boldsymbol{\theta}}|) + \mathbf{x}^t\mathbf{\Sigma}_{\boldsymbol{\theta}}^{-1}\mathbf{x} + 2Nlog(2\pi)\big) \tag{7}$$

Derivadas

Fórmula Geral da Derivada da Log-verossimilhança

Ao derivarmos 7 em relação a qualquer elemento, θ, de $\boldsymbol{\theta}$, temos a expressão geral da derivada da log-verossimilhança:

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \theta} = tr\Big(\mathbf{\Sigma}_{\boldsymbol{\theta}}^{-1}\frac{\partial \mathbf{\Sigma}_{\boldsymbol{\theta}}}{\partial \theta}\Big) - \mathbf{x}^t\left[\mathbf{\Sigma}_{\boldsymbol{\theta}}^{-1}\frac{\partial \mathbf{\Sigma}_{\boldsymbol{\theta}}}{\partial \theta}\mathbf{\Sigma}_{\boldsymbol{\theta}}^{-1}\right]\mathbf{x}. \tag{8}$$

Então se $y = \mathbf{x}\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1}$, pela simetria da função de covariância Matérn Bivariada tem-se que

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \theta} = tr\Big(\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1}\frac{\partial \boldsymbol{\Sigma}_{\boldsymbol{\theta}}}{\partial \theta}\Big) - \mathbf{y}^t\left[\frac{\partial \boldsymbol{\Sigma}_{\boldsymbol{\theta}}}{\partial \theta}\right]\mathbf{y}. \tag{9}$$

Derivada das Funções de Covariâncias

Para obter $\partial \boldsymbol{\Sigma}_{\boldsymbol{\theta}}/\partial\theta$, onde $\boldsymbol{\theta} = (\sigma_1^2, \sigma_2^2, a, \rho, \mu_1, \mu_2)$, vamos utilizar a regra da cadeia passo a passo.

Derivada de $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ c.r.a σ_1^2

$$\frac{\partial \boldsymbol{\Sigma}_{\boldsymbol{\theta}}(\mathbf{h})}{\partial \sigma_1^2} = \begin{pmatrix} M(\mathbf{h}|\nu_1, a) & \\ \frac{\rho\sigma_2}{2\sigma_1}M(\mathbf{h}|\frac{\nu_1+\nu_2}{2}, a) & \mathbf{O} \end{pmatrix} \tag{10}$$

Derivada de $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ c.r.a σ_2^2

$$\frac{\partial \boldsymbol{\Sigma}_{\boldsymbol{\theta}}(\mathbf{h})}{\partial \sigma_2^2} = \begin{pmatrix} \mathbf{O} & \\ \frac{\rho\sigma_1}{2\sigma_2}M(\mathbf{h}|\frac{\nu_1+\nu_2}{2}, a) & M(\mathbf{h}|\nu_2, a) \end{pmatrix} \tag{11}$$

Derivada de $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ c.r.a ρ

$$\frac{\partial \boldsymbol{\Sigma}_{\boldsymbol{\theta}}(\mathbf{h})}{\partial \sigma_2^2} = \begin{pmatrix} \mathbf{O} & \\ \sigma_1\sigma_2 M(\mathbf{h}|\frac{\nu_1+\nu_2}{2}, a) & \mathbf{O} \end{pmatrix} \tag{12}$$

Derivada de $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ c.r.a a

Neste caso temos que

$$\frac{\partial \boldsymbol{\Sigma}_{\boldsymbol{\theta}}(\mathbf{h})}{\partial a} = \begin{pmatrix} \sigma_1^2\psi_1 & \rho\sigma_1^2\sigma_2^2\psi_3 \\ \rho\sigma_1^2\sigma_2^2\psi_3 & \sigma_2^2\psi_2 \end{pmatrix} \tag{13}$$

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

$$\frac{\partial M(\mathbf{h}|\nu, a)}{\partial a} = \frac{2^{1-\nu}d^\nu}{\Gamma(\nu)}\left[\nu a^{\nu-1}K_\nu(ad) + a^\nu\frac{\partial K_\nu(ad)}{\partial a}\right] \tag{14}$$

e, como

$$\frac{\partial K_\nu(ad)}{\partial a} = d\left[\frac{\nu}{ad}K_\nu(ad) - K_{\nu+1}(ad)\right] \tag{15}$$

então

$$\frac{\partial M(\mathbf{h}|\nu, a)}{\partial a} = \frac{2^{1-\nu}d^\nu}{\Gamma(\nu)}\left[\nu a^{\nu-1}K_\nu(ad) + a^\nu d\left(\frac{\nu}{ad}K_\nu(ad) - K_{\nu+1}(ad)\right)\right]. \tag{16}$$

Ao simplificar a última equação, obtem-se que

$$\psi_i = \frac{2^{1-\nu_i}d^\nu a^{\nu_i-1}}{\Gamma(\nu_i)}\left[2\nu_i K_{\nu_i}(ad) - adK_{\nu_i+1}(ad)\right] \tag{17}$$

Estrutura de Blocos para Contas na Log-Verossimilhança

Log do Determinante

Todas as contas a seguir tem o intuito de melhorar problemas de instabilidade numérica, reduzir o custo computacional de inversas.

Se

$$\mathbf{\Sigma}_{\boldsymbol{\theta}}(\mathbf{h}) = \begin{pmatrix} C_{11}(\mathbf{h}) & C_{12}(\mathbf{h}) \\ C_{21}(\mathbf{h}) & C_{22}(\mathbf{h}) \end{pmatrix} \tag{18}$$

Então

$$det(\Sigma) = |\Sigma| = \left| C_{11}(\mathbf{h}) - C_{12}(\mathbf{h})C_{22}^{-1}(\mathbf{h})C_{21}(\mathbf{h}) \right| * \left| C_{22}(\mathbf{h}) \right| \tag{19}$$

Portanto, com o logarítmo fica

$$log(det(\Sigma)) = log\left| C_{11}(\mathbf{h}) - C_{12}(\mathbf{h})C_{22}^{-1}(\mathbf{h})C_{21}(\mathbf{h}) \right| + log\left| C_{22}(\mathbf{h}) \right| \tag{20}$$

Inversa

Aqui existem duas formas. Eu preferi usar a que irei apresentar a seguir pois ela encaixa bem com o resultado (20) do log do determinante[1].

Se

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

Então

$$\begin{aligned}
C_{11}^{*}(\mathbf{h}) &= \left[C_{11}(\mathbf{h}) - C_{12}(\mathbf{h})C_{22}^{-1}(\mathbf{h})C_{21}(\mathbf{h}) \right]^{-1} \\
C_{12}^{*}(\mathbf{h}) &= -\left[C_{11}(\mathbf{h}) - C_{12}(\mathbf{h})C_{22}^{-1}(\mathbf{h})C_{21}(\mathbf{h}) \right]^{-1} C_{12}(\mathbf{h})C_{22}^{-1}(\mathbf{h}) \\
C_{21}^{*}(\mathbf{h}) &= -C_{22}^{-1}(\mathbf{h})C_{21}(\mathbf{h}) \left[C_{11}(\mathbf{h}) - C_{12}(\mathbf{h})C_{22}^{-1}(\mathbf{h})C_{21}(\mathbf{h}) \right]^{-1} \\
C_{22}^{*}(\mathbf{h}) &= C_{22}^{-1}(\mathbf{h}) + C_{22}^{-1}(\mathbf{h})C_{21}(\mathbf{h}) \left[C_{11}(\mathbf{h}) - C_{12}(\mathbf{h})C_{22}^{-1}(\mathbf{h})C_{21}(\mathbf{h}) \right]^{-1} C_{12}(\mathbf{h})C_{22}^{-1}(\mathbf{h})
\end{aligned} \tag{22}$$

Já que[2] $C_{12}(\mathbf{h}) = C_{21}^{T}(\mathbf{h})$, se definirmos

$$\begin{aligned}
V_1 &:= C_{22}^{-1}(\mathbf{h})C_{21}(\mathbf{h}) \\
V_2 &:= C_{11}(\mathbf{h}) - C_{12}(\mathbf{h})C_{22}^{-1}(\mathbf{h})C_{21}(\mathbf{h})
\end{aligned} \tag{23}$$

[1] Apesar de eu achar que não tem como o alimentar o optim com as inversas obtidas, para ganhar tempo

[2] No artigo Σ é definida assim

podemos obter uma forma simplificada para o sistema (22), já que como $C_{11}, C_{12}, C_{21}, C_{22}$ são simétricas

$$
\begin{aligned}
C_{11}^*(\mathbf{h}) &= V_2^{-1} \\
C_{12}^*(\mathbf{h}) &= -V_2^{-1}V_1^T \\
C_{21}^*(\mathbf{h}) &= -V_1V_2^{-1} \\
C_{22}^*(\mathbf{h}) &= C_{22}^{-1}(\mathbf{h}) - V_1V_2^{-1}V_1^T
\end{aligned}
\tag{24}
$$

pois $V_1 = C_{22}^{-1}(\mathbf{h})C_{21}(\mathbf{h})$ fornece $V_1^T = C_{12}(\mathbf{h})C_{22}^{-1}(\mathbf{h})$.

Note ainda que $C_{12}^*(\mathbf{h}) = C_{21}^{*T}(\mathbf{h})$.

<u>**Prova:**</u>

Primeiramente

$$
\begin{aligned}
V_2^T &= C_{11}(\mathbf{h}) - (C_{12}(\mathbf{h})C_{22}^{-1}(\mathbf{h})C_{12}(\mathbf{h}))^T \\
&= C_{11}(\mathbf{h}) - C_{12}(\mathbf{h})(C_{12}(\mathbf{h})C_{22}^{-1}(\mathbf{h}))^T \\
&= C_{11}(\mathbf{h}) - C_{12}(\mathbf{h})C_{22}^{-1}(\mathbf{h})C_{12}(\mathbf{h}) \\
&= V_2
\end{aligned}
\tag{25}
$$

Agora

$$
\begin{aligned}
C_{12}^{*T}(\mathbf{h}) &= -(V_2^{-1}V_1^T)^T \\
&= V_1^T(V_2^{-1})^T \\
&= V_1^T(V_2^{-1}) \\
&= C_{21}^*(\mathbf{h})
\end{aligned}
\tag{26}
$$

Avaliação de $\mathbf{\Sigma}^{-1}$ na log-verossimilhança

Neste caso precisamos avaliar $y^T\mathbf{\Sigma}^{-1}y$. A estrutura de blocos também pode nos ajudar a deixar o algorítmo um pouco mais estável e rápido.

Novamente a inversa de $\mathbf{\Sigma}$ é denotada por

$$
\mathbf{\Sigma}_{\boldsymbol{\theta}}^{-1}(\mathbf{h}) = \begin{pmatrix} C_{11}^*(\mathbf{h}) & C_{12}^*(\mathbf{h}) \\ C_{21}^*(\mathbf{h}) & C_{22}^*(\mathbf{h}) \end{pmatrix}
\tag{27}
$$

Agora se $y^T = (y_1, y_2)^T$, em que y_1 e y_2 são, respectivamente, os vetores resposta associados às primeira e segunda variáveis multivariadas, temos que

$$
\begin{aligned}
y^T\mathbf{\Sigma}_{\boldsymbol{\theta}}^{-1}(\mathbf{h})y &= y^T \begin{pmatrix} C_{11}^*(\mathbf{h}) & C_{12}^*(\mathbf{h}) \\ C_{21}^*(\mathbf{h}) & C_{22}^*(\mathbf{h}) \end{pmatrix} y \\
&= (y_1, y_2)^T \begin{pmatrix} C_{11}^*(\mathbf{h}) & C_{12}^*(\mathbf{h}) \\ C_{21}^*(\mathbf{h}) & C_{22}^*(\mathbf{h}) \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \\
&= (y_1, y_2)^T \begin{pmatrix} C_{11}^*(\mathbf{h})y_1 + C_{12}^*(\mathbf{h})y_2 \\ C_{21}^*(\mathbf{h})y_1 + C_{22}^*(\mathbf{h})y_2 \end{pmatrix} \\
&= y_1^TC_{11}^*(\mathbf{h})y_1 + y_1^TC_{12}^*(\mathbf{h})y_2 + y_2^TC_{21}^*(\mathbf{h})y_1 + y_2^TC_{22}^*(\mathbf{h})y_2 \\
&= y_1^TC_{11}^*(\mathbf{h})y_1 + 2y_1^TC_{12}^*(\mathbf{h})y_2 + y_2^TC_{22}^*(\mathbf{h})y_2
\end{aligned}
\tag{28}
$$

pois como $C_{12}^{*T}(\mathbf{h}) = C_{21}^*(\mathbf{h})$ obtemos que

$$y_1^T C_{12}^*(\mathbf{h}) y_2 = tr(y_1^T C_{12}^*(\mathbf{h}) y_2) = tr(y_1^T C_{12}^*(\mathbf{h}) y_2)^T = tr((C_{12}^*(\mathbf{h}) y_2)^T y_1) = tr(y_2^T C_{12}^{*T}(\mathbf{h}) y_1) = y_2^T C_{21}^*(\mathbf{h}) y_1$$

Também é possível escrever o resultado anterior em termos de (28), o que fornece

$$y_1^T V_2^{-1} y_1 - 2 y_1^T V_2^{-1} V_1^T y_2 + y_2^T C_{22}^{-1} y_2 - y_2^T V_1 V_2^{-1} V_1^T y_2$$

Estrutura de Blocos para Contas no Gradiente

Aqui a idéia é mitigar o custo computacional de $tr\left(\mathbf{\Sigma}_{\boldsymbol{\theta}}^{-1}(\mathbf{h}) \frac{\partial \Sigma}{\partial \theta_k}\right) + y^T \left[\frac{\partial \Sigma(\mathbf{h})}{\partial \theta_k}\right] y$ já que vão aparecer muitas matrizes nulas bem grandes em $\partial \Sigma / \partial \boldsymbol{\theta}_k$

Relembrar é viver: Produto de Matrizes

Se

$$\mathbf{\Sigma}_{\boldsymbol{\theta}}(\mathbf{h}) = \begin{pmatrix} C_{11}(\mathbf{h}) & C_{12}(\mathbf{h}) \\ C_{21}(\mathbf{h}) & C_{22}(\mathbf{h}) \end{pmatrix} \tag{29}$$

e

$$\boldsymbol{M} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \tag{30}$$

Então desde as matrizes $\boldsymbol{M}$ e $\mathbf{\Sigma}_{\boldsymbol{\theta}}(\mathbf{h})$ sejam conformes

$$\mathbf{\Sigma}_{\boldsymbol{\theta}}(\mathbf{h}) \boldsymbol{M} = \begin{pmatrix} C_{11}(\mathbf{h})A + C_{12}(\mathbf{h})C & C_{11}(\mathbf{h})B + C_{12}(\mathbf{h})D \\ C_{21}(\mathbf{h})A + C_{22}(\mathbf{h})C & C_{21}(\mathbf{h})B + C_{22}(\mathbf{h})D \end{pmatrix} \tag{31}$$

Caso em que $\theta_k = \sigma_1^2$

Traço

$$\begin{aligned} \mathbf{\Sigma}_{\boldsymbol{\theta}}^{-1}(\mathbf{h}) \frac{\partial \Sigma}{\partial \sigma_1^2} &= \begin{pmatrix} C_{11}^*(\mathbf{h}) & C_{12}^*(\mathbf{h}) \\ C_{21}^*(\mathbf{h}) & C_{22}^*(\mathbf{h}) \end{pmatrix} \begin{pmatrix} M(\mathbf{h}|\nu_1, a) & \\ \frac{\rho\sigma_2}{2\sigma_1} M(\mathbf{h}|\nu_3, a) & \mathbf{O} \end{pmatrix} \\ &= \begin{pmatrix} C_{11}^*(\mathbf{h}) M(\mathbf{h}|\nu_1, a) + C_{12}^*(\mathbf{h}) \frac{\rho\sigma_2}{2\sigma_1} M(\mathbf{h}|\nu_3, a) & C_{11}^*(\mathbf{h}) \frac{\rho\sigma_2}{2\sigma_1} M(\mathbf{h}|\nu_3, a) \\ C_{21}^*(\mathbf{h}) M(\mathbf{h}|\nu_1, a) + C_{22}^*(\mathbf{h}) \frac{\rho\sigma_2}{2\sigma_1} M(\mathbf{h}|\nu_3, a) & C_{21}^*(\mathbf{h}) \frac{\rho\sigma_2}{2\sigma_1} M(\mathbf{h}|\nu_3, a) \end{pmatrix} \end{aligned} \tag{32}$$

Portanto como $C_{12}^*(\mathbf{h}) = C_{21}^{*T}(\mathbf{h})$

$$\begin{aligned} tr\left(\mathbf{\Sigma}_{\boldsymbol{\theta}}^{-1}(\mathbf{h}) \frac{\partial \Sigma}{\partial \sigma_1^2}\right) &= tr\left(C_{11}^*(\mathbf{h}) M(\mathbf{h}|\nu_1, a)\right) + 2tr\left(C_{12}^*(\mathbf{h}) \frac{\rho\sigma_2}{2\sigma_1} M(\mathbf{h}|\nu_3, a)\right) \\ &= tr\left(C_{11}^*(\mathbf{h}) M(\mathbf{h}|\nu_1, a)\right) + \frac{\rho\sigma_2}{\sigma_1} tr\left(C_{12}^*(\mathbf{h}) M(\mathbf{h}|\nu_3, a)\right) \\ &= \frac{1}{\sigma_1^2} tr\left(C_{11}^*(\mathbf{h}) C_{11}(\mathbf{h})\right) + \frac{1}{\sigma_1^2} tr\left(C_{12}^*(\mathbf{h}) C_{12}(\mathbf{h})\right) \\ &= \frac{1}{\sigma_1^2}\left[tr\left(C_{11}^*(\mathbf{h}) C_{11}(\mathbf{h})\right) + tr\left(C_{12}^*(\mathbf{h}) C_{12}(\mathbf{h})\right)\right] \end{aligned} \tag{33}$$

Forma Quadrática

$$
\begin{aligned}
y^T\left[\frac{\partial \Sigma(\mathbf{h})}{\partial \theta_k}\right] y &= \begin{pmatrix} y_1^T & y_2^T \end{pmatrix} \begin{pmatrix} M(\mathbf{h}|\nu_1,a) & \\ \rho\frac{\sigma_2}{2\sigma_1}M(\mathbf{h}|\nu_3,a) & \mathbf{O} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \\
&= \begin{pmatrix} y_1^T & y_2^T \end{pmatrix} \begin{pmatrix} M(\mathbf{h}|\nu_1,a)y_1 + \rho\frac{\sigma_2}{2\sigma_1}M(\mathbf{h}|\nu_3,a)y_2 \\ \rho\frac{\sigma_2}{2\sigma_1}M(\mathbf{h}|\nu_3,a)y_1 \end{pmatrix} \\
&= y_1^T M(\mathbf{h}|\nu_1,a)y_1 + y_1^T\rho\frac{\sigma_2}{2\sigma_1}M(\mathbf{h}|\nu_3,a)y_2 + y_2^T\rho\frac{\sigma_2}{2\sigma_1}M(\mathbf{h}|\nu_3,a)y_1 \\
&= y_1^T M(\mathbf{h}|\nu_1,a)y_1 + \rho\frac{\sigma_2}{\sigma_1}y_1^T M(\mathbf{h}|\nu_3,a)y_2 \\
&= \frac{1}{\sigma_1^2}y_1^T\left[C_{11}(\mathbf{h}|\nu_1,a)y_1 + C_{12}(\mathbf{h}|\nu_3,a)y_2\right]
\end{aligned} \tag{34}
$$

Caso em que $\theta_k = \sigma_2^2$

Traço

$$
\begin{aligned}
\mathbf{\Sigma}_{\boldsymbol{\theta}}^{-1}(\mathbf{h})\frac{\partial \Sigma}{\partial \sigma_2^2} &= \begin{pmatrix} C_{11}^*(\mathbf{h}) & C_{12}^*(\mathbf{h}) \\ C_{21}^*(\mathbf{h}) & C_{22}^*(\mathbf{h}) \end{pmatrix} \begin{pmatrix} \mathbf{O} & \\ \frac{\rho\sigma_1}{2\sigma_2}M(\mathbf{h}|\nu_3,a) & M(\mathbf{h}|\nu_2,a) \end{pmatrix} \\
&= \begin{pmatrix} C_{12}^*(\mathbf{h})\frac{\rho\sigma_1}{2\sigma_2}M(\mathbf{h}|\nu_3,a) & C_{11}^*(\mathbf{h})\frac{\rho\sigma_1}{2\sigma_2}M(\mathbf{h}|\nu_3,a) + C_{12}^*(\mathbf{h})M(\mathbf{h}|\nu_2,a) \\ C_{22}^*(\mathbf{h})\frac{\rho\sigma_1}{2\sigma_2}M(\mathbf{h}|\nu_3,a) & C_{21}^*(\mathbf{h})\frac{\rho\sigma_1}{2\sigma_2}M(\mathbf{h}|\nu_3,a) + C_{22}^*(\mathbf{h})M(\mathbf{h}|\nu_2,a) \end{pmatrix}
\end{aligned} \tag{35}
$$

Portanto como $C_{12}^*(\mathbf{h}) = C_{21}^{*T}(\mathbf{h})$

$$
\begin{aligned}
tr\left(\mathbf{\Sigma}_{\boldsymbol{\theta}}^{-1}(\mathbf{h})\frac{\partial \Sigma}{\partial \sigma_2^2}\right) &= tr\left(C_{22}^*(\mathbf{h})M(\mathbf{h}|\nu_2,a)\right) + 2tr\left(C_{12}^*(\mathbf{h})\frac{\rho\sigma_1}{2\sigma_2}M(\mathbf{h}|\nu_3,a)\right) \\
&= tr\left(C_{22}^*(\mathbf{h})M(\mathbf{h}|\nu_2,a)\right) + \frac{\rho\sigma_1}{\sigma_2}tr\left(C_{12}^*(\mathbf{h})M(\mathbf{h}|\nu_3,a)\right) \\
&= \frac{1}{\sigma_2^2}tr\left(C_{22}^*(\mathbf{h})C_{22}(\mathbf{h})\right) + \frac{1}{\sigma_2^2}tr\left(C_{12}^*(\mathbf{h})C_{12}(\mathbf{h})\right) \\
&= \frac{1}{\sigma_2^2}\left[tr\left(C_{22}^*(\mathbf{h})C_{22}(\mathbf{h})\right) + tr\left(C_{12}^*(\mathbf{h})C_{12}(\mathbf{h})\right)\right]
\end{aligned} \tag{36}
$$

Forma Quadrática

$$
\begin{aligned}
y^T\left[\frac{\partial \Sigma(\mathbf{h})}{\partial \theta_k}\right] y &= \begin{pmatrix} y_1^T & y_2^T \end{pmatrix} \begin{pmatrix} \mathbf{O} & \\ \rho\frac{\sigma_1}{2\sigma_2}M(\mathbf{h}|\nu_3,a) & M(\mathbf{h}|\nu_2,a) \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \\
&= \begin{pmatrix} y_1^T & y_2^T \end{pmatrix} \begin{pmatrix} \rho\frac{\sigma_1}{2\sigma_2}M(\mathbf{h}|\nu_3,a)y_2 \\ \rho\frac{\sigma_1}{2\sigma_2}M(\mathbf{h}|\nu_3,a)y_1 + M(\mathbf{h}|\nu_2,a)y_2 \end{pmatrix} \\
&= y_2^T M(\mathbf{h}|\nu_1,a)y_2 + y_1^T\rho\frac{\sigma_1}{2\sigma_2}M(\mathbf{h}|\nu_3,a)y_2 + y_2^T\rho\frac{\sigma_1}{2\sigma_2}M(\mathbf{h}|\nu_3,a)y_1 \\
&= y_2^T M(\mathbf{h}|\nu_1,a)y_2 + \rho\frac{\sigma_1}{\sigma_2}y_1^T M(\mathbf{h}|\nu_3,a)y_2 \\
&= \frac{1}{\sigma_2^2}\left[y_2^T C_{22}(\mathbf{h}|\nu_1,a) + y_1^T C_{12}(\mathbf{h}|\nu_3,a)\right]y_2
\end{aligned} \tag{37}
$$

Caso em que $\theta_k = \rho$

Traço

$$\begin{aligned}\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1}(\mathbf{h})\frac{\partial \Sigma}{\partial \rho} &= \begin{pmatrix} C_{11}^*(\mathbf{h}) & C_{12}^*(\mathbf{h}) \\ C_{21}^*(\mathbf{h}) & C_{22}^*(\mathbf{h}) \end{pmatrix} \begin{pmatrix} \mathbf{O} & \\ \sigma_1\sigma_2 M(\mathbf{h}|\nu_3, a) & \mathbf{O} \end{pmatrix} \\ &= \begin{pmatrix} C_{12}^*(\mathbf{h})\sigma_1\sigma_2 M(\mathbf{h}|\nu_3, a) & C_{11}^*(\mathbf{h})\sigma_1\sigma_2 M(\mathbf{h}|\nu_3, a) \\ C_{22}^*(\mathbf{h})\sigma_1\sigma_2 M(\mathbf{h}|\nu_3, a) & C_{21}^*(\mathbf{h})\sigma_1\sigma_2 M(\mathbf{h}|\nu_3, a) \end{pmatrix}\end{aligned} \tag{38}$$

Portanto como $C_{12}^*(\mathbf{h}) = C_{21}^{*T}(\mathbf{h})$

$$tr\left(\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1}(\mathbf{h})\frac{\partial \Sigma}{\partial \rho}\right) = 2\sigma_1\sigma_2 tr\left(C_{12}^*(\mathbf{h})M(\mathbf{h}|\nu_3, a)\right) \tag{39}$$

Forma Quadrática

$$\begin{aligned} y^T\left[\frac{\partial \Sigma(\mathbf{h})}{\partial \theta_k}\right]y &= \begin{pmatrix} y_1^T & y_2^T \end{pmatrix} \begin{pmatrix} \mathbf{O} & \\ \sigma_1\sigma_2 M(\mathbf{h}|\nu_3, a) & \mathbf{O} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \\ &= \begin{pmatrix} y_1^T & y_2^T \end{pmatrix} \begin{pmatrix} \sigma_1\sigma_2 M(\mathbf{h}|\nu_3, a)y_2 \\ \sigma_1\sigma_2 M(\mathbf{h}|\nu_3, a)y_1 \end{pmatrix} \\ &= y_1^T\sigma_1\sigma_2 M(\mathbf{h}|\nu_3, a)y_2 + y_2^T\sigma_1\sigma_2 M(\mathbf{h}|\nu_3, a)y_1 \\ &= 2\sigma_1\sigma_2 y_1^T M(\mathbf{h}|\nu_3, a)y_2 \end{aligned} \tag{40}$$

Caso em que $\theta_k = a$

Traço

$$\begin{aligned}\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1}(\mathbf{h})\frac{\partial \Sigma}{\partial a} &= \begin{pmatrix} C_{11}^*(\mathbf{h}) & C_{12}^*(\mathbf{h}) \\ C_{21}^*(\mathbf{h}) & C_{22}^*(\mathbf{h}) \end{pmatrix} \begin{pmatrix} \sigma_1^2 M^{'}(\mathbf{h}|\nu_1, a) & \rho\sigma_1\sigma_2 M^{'}(\mathbf{h}|\nu_3, a) \\ \rho\sigma_1\sigma_2 M^{'}(\mathbf{h}|\nu_3, a) & \sigma_2^2 M^{'}(\mathbf{h}|\nu_2, a) \end{pmatrix} \\ &= \begin{pmatrix} C_{11}^*(\mathbf{h})\sigma_1^2 M^{'}(\mathbf{h}|\nu_1, a) + C_{12}^*(\mathbf{h})\rho\sigma_1\sigma_2 M^{'}(\mathbf{h}|\nu_3, a) & C_{11}^*(\mathbf{h})\rho\sigma_1\sigma_2 M^{'}(\mathbf{h}|\nu_3, a) + C_{12}^*(\mathbf{h})\sigma_2^2 M^{'}(\mathbf{h}|\nu_2, a) \\ C_{21}^*(\mathbf{h})\sigma_1^2 M^{'}(\mathbf{h}|\nu_1, a) + C_{22}^*(\mathbf{h})\rho\sigma_1\sigma_2 M^{'}(\mathbf{h}|\nu_3, a) & C_{21}^*(\mathbf{h})\rho\sigma_1\sigma_2 M^{'}(\mathbf{h}|\nu_3, a) + C_{22}^*(\mathbf{h})\sigma_2^2 M^{'}(\mathbf{h}|\nu_2, a) \end{pmatrix}\end{aligned} \tag{41}$$

Portanto como $C_{12}^*(\mathbf{h}) = C_{21}^{*T}(\mathbf{h})$

$$\begin{aligned} tr\left(\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1}(\mathbf{h})\frac{\partial \Sigma}{\partial a}\right) = \ & \sigma_1^2 tr\left(C_{11}^*(\mathbf{h})M^{'}(\mathbf{h}|\nu_1, a)\right) + \\ & \sigma_2^2 tr\left(C_{22}^*(\mathbf{h})M^{'}(\mathbf{h}|\nu_2, a)\right) + \\ & 2\rho\sigma_2\sigma_2 tr\left(C_{12}^*(\mathbf{h})M^{'}(\mathbf{h}|\nu_3, a)\right) \end{aligned} \tag{42}$$

Por outro lado, temos que

$$M^{'}(\mathbf{h}|\nu, a) = \frac{\partial M(\mathbf{h}|\nu, a)}{\partial a} = \frac{(\frac{2}{a})^{1-\nu}d^\nu}{\Gamma(\nu)}\left[2\nu K_\nu(ad) - adK_{\nu+1}(ad)\right] \tag{43}$$

em que $d := ||\mathbf{h}||$.

Forma Quadrática

$$
\begin{aligned}
y^T\left[\frac{\partial\Sigma(\mathbf{h})}{\partial\theta_k}\right]y &= \begin{pmatrix} y_1^T & y_2^T \end{pmatrix} \begin{pmatrix} \sigma_1^2 M^{'}(\mathbf{h}|\nu_1,a) & \\ \rho\sigma_1\sigma_2 M^{'}(\mathbf{h}|\nu_3,a) & \sigma_2^2 M^{'}(\mathbf{h}|\nu_2,a) \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \\
&= \begin{pmatrix} y_1^T & y_2^T \end{pmatrix} \begin{pmatrix} \sigma_1^2 M^{'}(\mathbf{h}|\nu_1,a)y_1 + \rho\sigma_1\sigma_2 M^{'}(\mathbf{h}|\nu_3,a)y_2 \\ \rho\sigma_1\sigma_2 M^{'}(\mathbf{h}|\nu_3,a)y_1 + \sigma_2^2 M^{'}(\mathbf{h}|\nu_2,a)y_2 \end{pmatrix} \\
&= y_1^T\sigma_1^2 M^{'}(\mathbf{h}|\nu_1,a)y_1 + y_1^T\rho\sigma_1\sigma_2 M^{'}(\mathbf{h}|\nu_3,a)y_2 + y_2^T\rho\sigma_1\sigma_2 M^{'}(\mathbf{h}|\nu_3,a)y_1 + y_2^T\sigma_2^2 M^{'}(\mathbf{h}|\nu_2,a)y_2 \\
&= \sigma_1^2 y_1^T M^{'}(\mathbf{h}|\nu_1,a)y_1 + \sigma_2^2 y_2^T M^{'}(\mathbf{h}|\nu_2,a)y_2 + 2\rho\sigma_1\sigma_2 y_1^T M^{'}(\mathbf{h}|\nu_3,a)y_2
\end{aligned} \tag{44}
$$