Block Properties for Biwm

Leonardo Uchoa

Todas as contas a seguir tem o intuito de melhorar problemas de instabilidade numérica, reduzir o custo computacional de inversas e simplesmente fazer este belíssimo código funcionar.

Contas na Log-Verossimilhança

Log do Determinante

Se

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

Então

$$det(\Sigma) = |\Sigma| = \left| C_{11}(\mathbf{h}) - C_{12}(\mathbf{h}) C_{22}^{-1}(\mathbf{h}) C_{21}(\mathbf{h}) \right| * \left| C_{22}(\mathbf{h}) \right| \tag{2}$$

Portanto, com o logarítmo fica

$$log(det(\Sigma)) = log \left| C_{11}(\mathbf{h}) - C_{12}(\mathbf{h}) C_{22}^{-1}(\mathbf{h}) C_{21}(\mathbf{h}) \right| + log \left| C_{22}(\mathbf{h}) \right| \tag{3}$$

Inversa

Aqui existem duas formas. Eu preferi usar a que irei apresentação a seguir pois ela encaixa bem com o resultado (3) do log do determinante[1].

Se

$$\mathbf{\Sigma}_{\boldsymbol{\theta}}^{-1}(\mathbf{h}) = \begin{pmatrix} C_{11}^{*}(\mathbf{h}) & C_{12}^{*}(\mathbf{h}) \\ C_{21}^{*}(\mathbf{h}) & C_{22}^{*}(\mathbf{h}) \end{pmatrix} \tag{4}$$

Então

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

$$
\begin{aligned}
C_{11}^{*}(\mathbf{h}) &= \left[C_{11}(\mathbf{h}) - C_{12}(\mathbf{h})C_{22}^{-1}(\mathbf{h})C_{21}(\mathbf{h})\right]^{-1} \\
C_{12}^{*}(\mathbf{h}) &= -\left[C_{11}(\mathbf{h}) - C_{12}(\mathbf{h})C_{22}^{-1}(\mathbf{h})C_{21}(\mathbf{h})\right]^{-1} C_{12}(\mathbf{h})C_{22}^{-1}(\mathbf{h}) \\
C_{21}^{*}(\mathbf{h}) &= -C_{22}^{-1}(\mathbf{h})C_{21}(\mathbf{h})\left[C_{11}(\mathbf{h}) - C_{12}(\mathbf{h})C_{22}^{-1}(\mathbf{h})C_{21}(\mathbf{h})\right]^{-1} \\
C_{22}^{*}(\mathbf{h}) &= C_{22}^{-1}(\mathbf{h}) + C_{22}^{-1}(\mathbf{h})C_{21}(\mathbf{h})\left[C_{11}(\mathbf{h}) - C_{12}(\mathbf{h})C_{22}^{-1}(\mathbf{h})C_{21}(\mathbf{h})\right]^{-1} C_{12}(\mathbf{h})C_{22}^{-1}(\mathbf{h})
\end{aligned}
\tag{5}
$$

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

$$
\begin{aligned}
V_1 &:= C_{22}^{-1}(\mathbf{h})C_{21}(\mathbf{h}) \\
V_2 &:= C_{11}(\mathbf{h}) - C_{12}(\mathbf{h})C_{22}^{-1}(\mathbf{h})C_{21}(\mathbf{h})
\end{aligned}
\tag{6}
$$

podemos obter uma forma simplificada para o sistema (5), já que como $C_{11}, C_{12}, C_{21}, C_{22}$ são simétricas

$$
\begin{aligned}
C_{11}^{*}(\mathbf{h}) &= V_2^{-1} \\
C_{12}^{*}(\mathbf{h}) &= -V_2^{-1}V_1^T \\
C_{21}^{*}(\mathbf{h}) &= -V_1V_2^{-1} \\
C_{22}^{*}(\mathbf{h}) &= C_{22}(\mathbf{h}) - V_1V_2^{-1}V_1^T
\end{aligned}
\tag{7}
$$

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 $C_{12}^{*}(\mathbf{h}) = C_{12}^{*T}(\mathbf{h})$.

Prova:

Primeiramente

$$
\begin{aligned}
V_2^T &= C_{11}(\mathbf{h}) - (C_{12}(\mathbf{h})C_{22}^{-1}(\mathbf{h})C_{12}(\mathbf{h}))^T \\
&= C_{11}(\mathbf{h}) - C_{12}(\mathbf{h})(C_{12}(\mathbf{h})C_{22}^{-1}(\mathbf{h}))^T \\
&= C_{11}(\mathbf{h}) - C_{12}(\mathbf{h})C_{22}^{-1}(\mathbf{h})C_{12}(\mathbf{h}) \\
&= V_2
\end{aligned}
\tag{8}
$$

Agora

$$
\begin{aligned}
C_{12}^{*T}(\mathbf{h}) &= -(V_2^{-1}V_1^T)^T \\
&= V_1^T(V_2^{-1})^T \\
&= V_1^T(V_2^{-1}) \\
&= C_{21}^{*}(\mathbf{h})
\end{aligned}
\tag{9}
$$

[2]No artigo Σ é definida assim

Contas no Gradiente

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

Relembrar é viver: Produto de Matrizes

Se

$$\boldsymbol{\Sigma}_{\boldsymbol{\theta}}(\mathbf{h}) = \begin{pmatrix} C_{11}(\mathbf{h}) & C_{12}(\mathbf{h}) \\ C_{21}(\mathbf{h}) & C_{22}(\mathbf{h}) \end{pmatrix} \tag{10}$$

e

$$\boldsymbol{M} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \tag{11}$$

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

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

Traços: $tr\left(\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1}(\mathbf{h})\frac{\partial\Sigma}{\partial\theta_k}\right)$

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

$$\begin{aligned} \boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1}(\mathbf{h})\frac{\partial\Sigma}{\partial\sigma_1^2} &= \begin{pmatrix} C_{11}^*(\mathbf{h}) & C_{12}^*(\mathbf{h}) \\ C_{21}^*(\mathbf{h}) & C_{22}^*(\mathbf{h}) \end{pmatrix} \begin{pmatrix} M(\mathbf{h}|\nu_1, a) & \\ \frac{\rho\sigma_2}{2\sigma_1}M(\mathbf{h}|\frac{\nu_1+\nu_2}{2}, a) & \mathbf{O} \end{pmatrix} \\ &= \begin{pmatrix} C_{11}^*(\mathbf{h})M(\mathbf{h}|\nu_1, a) + C_{12}^*(\mathbf{h})\frac{\rho\sigma_2}{2\sigma_1}M(\mathbf{h}|\frac{\nu_1+\nu_2}{2}, a) & C_{11}^*(\mathbf{h})\frac{\rho\sigma_2}{2\sigma_1}M(\mathbf{h}|\frac{\nu_1+\nu_2}{2}, a) \\ C_{21}^*(\mathbf{h})M(\mathbf{h}|\nu_1, a) + C_{22}^*(\mathbf{h})\frac{\rho\sigma_2}{2\sigma_1}M(\mathbf{h}|\frac{\nu_1+\nu_2}{2}, a) & C_{11}^*(\mathbf{h})\frac{\rho\sigma_2}{2\sigma_1}M(\mathbf{h}|\frac{\nu_1+\nu_2}{2}, a) \end{pmatrix} \end{aligned} \tag{13}$$

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

$$\begin{aligned} tr\left(\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1}(\mathbf{h})\frac{\partial\Sigma}{\partial\sigma_1^2}\right) &= tr\left(C_{11}^*(\mathbf{h})M(\mathbf{h}|\nu_1, a)\right) + 2tr\left(C_{12}^*(\mathbf{h})\frac{\rho\sigma_2}{2\sigma_1}M(\mathbf{h}|\frac{\nu_1+\nu_2}{2}, a)\right) \\ &= tr\left(C_{11}^*(\mathbf{h})M(\mathbf{h}|\nu_1, a)\right) + \frac{\rho\sigma_2}{\sigma_1}tr\left(C_{12}^*(\mathbf{h})M(\mathbf{h}|\frac{\nu_1+\nu_2}{2}, a)\right) \end{aligned} \tag{14}$$

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

$$\begin{aligned} \boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1}(\mathbf{h})\frac{\partial\Sigma}{\partial\sigma_1^2} &= \begin{pmatrix} C_{11}^*(\mathbf{h}) & C_{12}^*(\mathbf{h}) \\ C_{21}^*(\mathbf{h}) & C_{22}^*(\mathbf{h}) \end{pmatrix} \begin{pmatrix} \mathbf{O} & \\ \frac{\rho\sigma_1}{2\sigma_2}M(\mathbf{h}|\frac{\nu_1+\nu_2}{2}, a) & M(\mathbf{h}|\nu_2, a) \end{pmatrix} \\ &= \begin{pmatrix} C_{11}^*(\mathbf{h})\frac{\rho\sigma_2}{2\sigma_1}M(\mathbf{h}|\frac{\nu_1+\nu_2}{2}, a) & C_{11}^*(\mathbf{h})\frac{\rho\sigma_2}{2\sigma_1}M(\mathbf{h}|\frac{\nu_1+\nu_2}{2}, a) + C_{12}^*(\mathbf{h})M(\mathbf{h}|\nu_2, a) \\ C_{22}^*(\mathbf{h})\frac{\rho\sigma_2}{2\sigma_1}M(\mathbf{h}|\frac{\nu_1+\nu_2}{2}, a) & C_{21}^*(\mathbf{h})\frac{\rho\sigma_2}{2\sigma_1}M(\mathbf{h}|\frac{\nu_1+\nu_2}{2}, a) + C_{22}^*(\mathbf{h})M(\mathbf{h}|\nu_2, a) \end{pmatrix} \end{aligned} \tag{15}$$

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

$$
\begin{aligned}
tr\left(\mathbf{\Sigma}_{\boldsymbol{\theta}}^{-1}(\mathbf{h})\frac{\partial\Sigma}{\partial\sigma_1^2}\right) &= tr\left(C_{22}^{*}(\mathbf{h})M(\mathbf{h}|\nu_2,a)\right) + 2tr\left(C_{12}^{*}(\mathbf{h})\frac{\rho\sigma_2}{2\sigma_1}M(\mathbf{h}|\frac{\nu_1+\nu_2}{2},a)\right) \\
&= tr\left(C_{22}^{*}(\mathbf{h})M(\mathbf{h}|\nu_1,a)\right) + \frac{\rho\sigma_2}{\sigma_1}tr\left(C_{12}^{*}(\mathbf{h})M(\mathbf{h}|\frac{\nu_1+\nu_2}{2},a)\right)
\end{aligned}
\tag{16}
$$

Caso em que $\theta_k = \rho$

$$
\begin{aligned}
\mathbf{\Sigma}_{\boldsymbol{\theta}}^{-1}(\mathbf{h})\frac{\partial\Sigma}{\partial\sigma_1^2} &= \begin{pmatrix} C_{11}^{*}(\mathbf{h}) & C_{12}^{*}(\mathbf{h}) \\ C_{21}^{*}(\mathbf{h}) & C_{22}^{*}(\mathbf{h}) \end{pmatrix} \begin{pmatrix} & \mathbf{O} \\ \sigma_1\sigma_2 M(\mathbf{h}|\frac{\nu_1+\nu_2}{2},a) & \mathbf{O} \end{pmatrix} \\
&= \begin{pmatrix} C_{12}^{*}(\mathbf{h})\sigma_1\sigma_2 M(\mathbf{h}|\frac{\nu_1+\nu_2}{2},a) & C_{11}^{*}(\mathbf{h})\sigma_1\sigma_2 M(\mathbf{h}|\frac{\nu_1+\nu_2}{2},a) \\ C_{22}^{*}(\mathbf{h})\sigma_1\sigma_2 M(\mathbf{h}|\frac{\nu_1+\nu_2}{2},a) & C_{21}^{*}(\mathbf{h})\sigma_1\sigma_2 M(\mathbf{h}|\frac{\nu_1+\nu_2}{2},a) \end{pmatrix}
\end{aligned}
\tag{17}
$$

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

$$
tr\left(\mathbf{\Sigma}_{\boldsymbol{\theta}}^{-1}(\mathbf{h})\frac{\partial\Sigma}{\partial\sigma_1^2}\right) = 2\sigma_1\sigma_2 tr\left(C_{22}^{*}(\mathbf{h})M(\mathbf{h}|\frac{\nu_1+\nu_2}{2},a)\right)
\tag{18}
$$