

2.58 2.226 gives us the result of taking the gradient of both sides of (2.195):

$$-\nabla \ln g(\boldsymbol{\eta}) = g(\boldsymbol{\eta}) \int h(\mathbf{x}) \exp\{\boldsymbol{\eta}^T \mathbf{u}(\mathbf{x})\} \mathbf{u}(\mathbf{x}) d\mathbf{x}$$

Using definition 5.6 from MML, we get the derivative of a vector-valued function $g(\boldsymbol{\eta})\mathbf{u}(\mathbf{x})$ w.r.t the vector $\boldsymbol{\eta}$, giving us $\nabla g(\boldsymbol{\eta})\mathbf{u}(\mathbf{x})^T$ in the first term:

$$\begin{aligned} -\nabla \nabla \ln g(\boldsymbol{\eta}) &= \nabla g(\boldsymbol{\eta}) \int h(\mathbf{x}) \exp\{\boldsymbol{\eta}^T \mathbf{u}(\mathbf{x})\} \mathbf{u}(\mathbf{x})^T d\mathbf{x} \\ &\quad + g(\boldsymbol{\eta}) \int h(\mathbf{x}) \frac{\partial \exp\{\boldsymbol{\eta}^T \mathbf{u}(\mathbf{x})\}}{\partial \boldsymbol{\eta}} \mathbf{u}(\mathbf{x}) d\mathbf{x} \\ &= \nabla g(\boldsymbol{\eta}) \int h(\mathbf{x}) \exp\{\boldsymbol{\eta}^T \mathbf{u}(\mathbf{x})\} \mathbf{u}(\mathbf{x})^T d\mathbf{x} \\ &\quad + g(\boldsymbol{\eta}) \int h(\mathbf{x}) \exp\{\boldsymbol{\eta}^T \mathbf{u}(\mathbf{x})\} \frac{\partial \boldsymbol{\eta}^T \mathbf{u}(\mathbf{x})}{\partial \boldsymbol{\eta}} \mathbf{u}(\mathbf{x}) d\mathbf{x} \\ &= \nabla g(\boldsymbol{\eta}) \int h(\mathbf{x}) \exp\{\boldsymbol{\eta}^T \mathbf{u}(\mathbf{x})\} \mathbf{u}(\mathbf{x})^T d\mathbf{x} \\ &\quad + g(\boldsymbol{\eta}) \int h(\mathbf{x}) \exp\{\boldsymbol{\eta}^T \mathbf{u}(\mathbf{x})\} \frac{\partial \boldsymbol{\eta}^T \mathbf{u}(\mathbf{x}) \mathbf{u}(\mathbf{x})}{\partial \boldsymbol{\eta}} d\mathbf{x} \\ &= \nabla g(\boldsymbol{\eta}) \int h(\mathbf{x}) \exp\{\boldsymbol{\eta}^T \mathbf{u}(\mathbf{x})\} \mathbf{u}(\mathbf{x})^T d\mathbf{x} \\ &\quad + g(\boldsymbol{\eta}) \int h(\mathbf{x}) \exp\{\boldsymbol{\eta}^T \mathbf{u}(\mathbf{x})\} \frac{\partial \mathbf{u}(\mathbf{x}) \mathbf{u}(\mathbf{x})^T \boldsymbol{\eta}}{\partial \boldsymbol{\eta}} d\mathbf{x} \end{aligned}$$

Using this result, we get:

$$\begin{aligned} &= \nabla g(\boldsymbol{\eta}) \int h(\mathbf{x}) \exp\{\boldsymbol{\eta}^T \mathbf{u}(\mathbf{x})\} \mathbf{u}(\mathbf{x})^T d\mathbf{x} \\ &\quad + g(\boldsymbol{\eta}) \int h(\mathbf{x}) \exp\{\boldsymbol{\eta}^T \mathbf{u}(\mathbf{x})\} \mathbf{u}(\mathbf{x}) \mathbf{u}(\mathbf{x})^T d\mathbf{x} \end{aligned}$$

Using 2.224,

$$\begin{aligned}
&= -g(\boldsymbol{\eta})^2 \left(\int h(\mathbf{x}) \exp\{\boldsymbol{\eta}^T \mathbf{u}(\mathbf{x})\} \mathbf{u}(\mathbf{x}) d\mathbf{x} \right) \left(\int h(\mathbf{x}) \exp\{\boldsymbol{\eta}^T \mathbf{u}(\mathbf{x})\} \mathbf{u}(\mathbf{x})^T d\mathbf{x} \right) \\
&\quad + \mathbb{E}[\mathbf{u}(\mathbf{x}) \mathbf{u}(\mathbf{x})^T] \\
&= - \left(g(\boldsymbol{\eta}) \int h(\mathbf{x}) \exp\{\boldsymbol{\eta}^T \mathbf{u}(\mathbf{x})\} \mathbf{u}(\mathbf{x}) d\mathbf{x} \right) \left(g(\boldsymbol{\eta}) \int h(\mathbf{x}) \exp\{\boldsymbol{\eta}^T \mathbf{u}(\mathbf{x})\} \mathbf{u}(\mathbf{x})^T d\mathbf{x} \right) \\
&\quad + \mathbb{E}[\mathbf{u}(\mathbf{x}) \mathbf{u}(\mathbf{x})^T] \\
&= -\mathbb{E}[\mathbf{u}(\mathbf{x})] \mathbb{E}[\mathbf{u}(\mathbf{x})^T] + \mathbb{E}[\mathbf{u}(\mathbf{x}) \mathbf{u}(\mathbf{x})^T] \\
&= \mathbb{E}[\mathbf{u}(\mathbf{x}) \mathbf{u}(\mathbf{x})^T] - \mathbb{E}[\mathbf{u}(\mathbf{x})] \mathbb{E}[\mathbf{u}(\mathbf{x})^T] \\
&\implies -\nabla \nabla \ln g(\boldsymbol{\eta}) = \text{cov}[\mathbf{u}(\mathbf{x})]
\end{aligned}$$