Exercise 5: Take the matrix formulation for 2 given in equation (7) and take derivatives to show that the minimum is at the matrix location given in equation (5).

Solution: Equation (7) reads

$$\chi^2 = \sum_{i=1}^{N} \frac{(y_i - f(x_i))^2}{\sigma_{y_i}^2} \equiv (\mathbf{Y} - \mathbf{AX})^{\mathbf{T}} \mathbf{C^{-1}} (\mathbf{Y} - \mathbf{AX}) \tag{1}$$

In order to take derivatives, we will first convert this matrix expression into an equivalent one involving matrix components:

$$\begin{aligned} \chi^2 &= (\mathbf{Y^T} - \mathbf{X^T A^T}) \mathbf{C^{-1}} (\mathbf{Y} - \mathbf{AX}) \\ &= (y_i - x_k A_{ki}) C_{il}^{-1} (y_l - A_{ln} x_n) \\ &= y_i C_{il}^{-1} y_l - x_k A_{ki} C_{il}^{-1} y_l - y_i C_{il}^{-1} A_{ln} x_n + x_k A_{ki} C_{il}^{-1} A_{ln} x_n \end{aligned} \tag{2}$$

where the Einstein summation convention has been utilized. Taking the derivative to find the minimum,

$$\begin{aligned} 0 = \frac{\partial \chi^2}{\partial x_m} &= -A_{mi} C_{il}^{-1} y_l - y_i C_{il}^{-1} A_{lm} + A_{mi} C_{il}^{-1} A_{ln} x_n + x_k A_{ki} C_{il}^{-1} A_{lm} \\ &= -\mathbf{A^T C^{-1} Y} - \mathbf{Y^T C^{-1} A} + \mathbf{A^T C^{-1} AX} + \mathbf{X^T A^T C^{-1} A} \end{aligned} \tag{3}$$

Notice that each of these terms is a scalar, and so must equal their transpose. For example, $(\mathbf{A^T C^{-1} y})^{\mathbf{T}} = \mathbf{y^T C^{-1} A} = \mathbf{A^T C^{-1} y}$ since $\mathbf{C^{-1}}$ is symmetric. The first two and the last two terms are equal, and so

$$\begin{aligned} 0 &= -2\mathbf{A^T C^{-1} Y} + 2\mathbf{A^T C^{-1} AX} \\ \implies \mathbf{X} &= (\mathbf{A^T C^{-1} A})^{-1} (\mathbf{A^T C^{-1} Y}) \end{aligned} \tag{4}$$

This is exactly the expression given in equation (5) of the text.