6.14 For a fixed covariance, the Fisher Score is given by:

$$\mathbf{g}(\boldsymbol{\mu}, \mathbf{x}) = \nabla_{\boldsymbol{\mu}} \ln \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \mathbf{S})$$

$$= (\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \mathbf{S}))^{-1} \quad \nabla_{\boldsymbol{\mu}} \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \mathbf{S})$$

$$= (\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \mathbf{S}))^{-1} \quad \nabla_{\boldsymbol{\mu}} \left(\frac{1}{(2\pi)^{D/2}|\mathbf{S}|^{1/2}} \exp\left\{ -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \mathbf{S}^{-1}(\mathbf{x}-\boldsymbol{\mu}) \right\} \right)$$

The derivative is:

$$\nabla_{\boldsymbol{\mu}} \left(\frac{1}{(2\pi)^{D/2}|\mathbf{S}|^{1/2}} \exp\left\{ -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \mathbf{S}^{-1}(\mathbf{x}-\boldsymbol{\mu}) \right\} \right)$$

$$= \frac{1}{(2\pi)^{D/2}|\mathbf{S}|^{1/2}} \left(\frac{\partial}{\partial \boldsymbol{\mu}} \exp\left\{ -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \mathbf{S}^{-1}(\mathbf{x}-\boldsymbol{\mu}) \right\} \right)$$

$$= \frac{1}{(2\pi)^{D/2}|\mathbf{S}|^{1/2}} \exp\left\{ -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \mathbf{S}^{-1}(\mathbf{x}-\boldsymbol{\mu}) \right\} \cdot \frac{\partial}{\partial \boldsymbol{\mu}} \left\{ -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \mathbf{S}^{-1}(\mathbf{x}-\boldsymbol{\mu}) \right\}$$

$$= \frac{1}{(2\pi)^{D/2}|\mathbf{S}|^{1/2}} \exp\left\{ -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \mathbf{S}^{-1}(\mathbf{x}-\boldsymbol{\mu}) \right\} \cdot \left(-\frac{1}{2} \right) \cdot \frac{\partial}{\partial \boldsymbol{\mu}} \left\{ (\mathbf{x}-\boldsymbol{\mu})^T \mathbf{S}^{-1}(\mathbf{x}-\boldsymbol{\mu}) \right\}$$

Using result 86 of the Matrix Cookbook, this becomes:

$$= \frac{1}{(2\pi)^{D/2}|\mathbf{S}|^{1/2}} \exp\left\{ -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \mathbf{S}^{-1}(\mathbf{x}-\boldsymbol{\mu}) \right\} \cdot \left(-\frac{1}{2} \right) \cdot \left\{ -2\mathbf{S}^{-1}(\mathbf{x}-\boldsymbol{\mu}) \right\}$$

$$= \frac{1}{(2\pi)^{D/2}|\mathbf{S}|^{1/2}} \exp\left\{ -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \mathbf{S}^{-1}(\mathbf{x}-\boldsymbol{\mu}) \right\} \cdot \left\{ \mathbf{S}^{-1}(\mathbf{x}-\boldsymbol{\mu}) \right\}$$

$$= \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \mathbf{S}) \cdot \left\{ \mathbf{S}^{-1}(\mathbf{x}-\boldsymbol{\mu}) \right\}$$

Substituting these results back into the expression for $\mathbf{g}$,

$$\mathbf{g}(\boldsymbol{\mu}, \mathbf{x}) = (\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \mathbf{S}))^{-1} \cdot \left(\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \mathbf{S}) \cdot \left\{\mathbf{S}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right\}\right)$$

$$= \mathbf{S}^{-1}(\mathbf{x} - \boldsymbol{\mu})$$

Using eqn 6.34, the Fisher Information matrix is given by:

$$\mathbf{F} = \mathbb{E}_{\mathbf{x}}\left[\left(\mathbf{S}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right) \left(\mathbf{S}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right)^T\right]$$

$$= \mathbb{E}_{\mathbf{x}}\left[\mathbf{S}^{-1}(\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^T\mathbf{S}^{-T}\right]$$

$$= \mathbb{E}_{\mathbf{x}}\left[\mathbf{S}^{-1}(\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^T\mathbf{S}^{-1}\right]$$

$$= \mathbf{S}^{-1}\mathbb{E}_{\mathbf{x}}\left[(\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^T\right]\mathbf{S}^{-1}$$

$$= \mathbf{S}^{-1}\mathbf{S}\mathbf{S}^{-1}$$

$$= \mathbf{S}^{-1}$$

Finally, using eqn 6.33, the Fisher kernel is given by:

$$k(\mathbf{x}, \mathbf{x'}) = \left(\mathbf{S}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right)^T \mathbf{S} \left(\mathbf{S}^{-1}(\mathbf{x'} - \boldsymbol{\mu})\right)$$

$$= (\mathbf{x} - \boldsymbol{\mu})^T\mathbf{S}^{-T}\mathbf{S}\mathbf{S}^{-1}(\mathbf{x'} - \boldsymbol{\mu})$$

$$= (\mathbf{x} - \boldsymbol{\mu})^T\mathbf{S}^{-1}\mathbf{S}\mathbf{S}^{-1}(\mathbf{x'} - \boldsymbol{\mu})$$

$$= (\mathbf{x} - \boldsymbol{\mu})^T\mathbf{S}^{-1}(\mathbf{x'} - \boldsymbol{\mu})$$