

**6.21** We already know that:

$$p(t_{N+1}|\mathbf{t}, \mathbf{x}_{N+1}) = \mathcal{N}(\mathbf{k}^T \mathbf{C}_N^{-1} \mathbf{t}, c - \mathbf{k}^T \mathbf{C}_N^{-1} \mathbf{k})$$

Considering the mean:

$$\begin{aligned} & \mathbf{k}^T \mathbf{C}_N^{-1} \mathbf{t} \\ &= \mathbf{k}(\mathbf{x})^T (\beta^{-1} \mathbf{I}_N + \mathbf{K})^{-1} \mathbf{t} \end{aligned}$$

Since vector  $\mathbf{k}$  has  $n$  elements  $k(\mathbf{x}_n, \mathbf{x})$ ,  $\mathbf{k} = \alpha^{-1} \Phi \phi(\mathbf{x})$ , giving us:

$$\begin{aligned} &= (\alpha^{-1} \Phi \phi(\mathbf{x}))^T (\beta^{-1} \mathbf{I}_N + \alpha^{-1} \Phi \Phi^T)^{-1} \mathbf{t} \\ &= (\alpha^{-1} \Phi \phi(\mathbf{x}))^T (\beta) (\mathbf{I}_N + (\beta/\alpha) \Phi \Phi^T)^{-1} \mathbf{t} \\ &= \beta \alpha^{-1} \phi(\mathbf{x})^T \left( \Phi^T (\mathbf{I}_N + (\beta/\alpha) \Phi \Phi^T)^{-1} \right) \mathbf{t} \end{aligned}$$

Applying C.6, we get:

$$\begin{aligned} &= \beta \alpha^{-1} \phi(\mathbf{x})^T \left( (\mathbf{I}_M + (\beta/\alpha) \Phi^T \Phi)^{-1} \Phi^T \right) \mathbf{t} \\ &= \beta \alpha^{-1} \phi(\mathbf{x})^T \alpha (\alpha \mathbf{I}_M + \beta \Phi^T \Phi)^{-1} \Phi^T \mathbf{t} \\ &= \beta \phi(\mathbf{x})^T \mathbf{S}_N^{-1} \Phi^T \mathbf{t} \\ &= \phi(\mathbf{x})^T (\beta \mathbf{S}_N^{-1} \Phi^T \mathbf{t}) \\ &= \phi(\mathbf{x})^T \mathbf{m}_N \end{aligned}$$

Considering the variance:

$$\begin{aligned} & c - \mathbf{k}^T \mathbf{C}_N^{-1} \mathbf{k} \\ &= (k(\mathbf{x}, \mathbf{x}) + \beta^{-1}) - (\alpha^{-1} \Phi \phi(\mathbf{x}))^T \left( \beta^{-1} \mathbf{I}_N + \alpha^{-1} \Phi \Phi^T \right)^{-1} (\alpha^{-1} \Phi \phi(\mathbf{x})) \end{aligned}$$

$$\begin{aligned}
&= (\alpha^{-1}\phi(\mathbf{x})^T\phi(\mathbf{x}) + \beta^{-1}) - (\alpha^{-1}\Phi\phi(\mathbf{x}))^T \left( \beta^{-1}\mathbf{I}_N + \alpha^{-1}\Phi\Phi^T \right)^{-1} (\alpha^{-1}\Phi\phi(\mathbf{x})) \\
&= \beta^{-1} + \phi(\mathbf{x})^T \alpha^{-1}\phi(\mathbf{x}) - \phi(\mathbf{x})^T \left( \alpha^{-1}\Phi^T \left( \beta^{-1}\mathbf{I}_N + \alpha^{-1}\Phi\Phi^T \right)^{-1} \alpha^{-1}\Phi \right) \phi(\mathbf{x}) \\
&= \beta^{-1} + \phi(\mathbf{x})^T \left( \alpha^{-1}\mathbf{I}_M - \alpha^{-1}\mathbf{I}_M\Phi^T \left( \beta^{-1}\mathbf{I}_N + \Phi\alpha^{-1}\mathbf{I}_M\Phi^T \right)^{-1} \Phi\alpha^{-1}\mathbf{I}_M \right) \phi(\mathbf{x})
\end{aligned}$$

If we let

$$\begin{aligned}
\mathbf{A} &= \alpha\mathbf{I}_M \\
\mathbf{B} &= \Phi^T \\
\mathbf{C} &= \Phi \\
\mathbf{D} &= \beta^{-1}\mathbf{I}_N
\end{aligned}$$

Then, using C.7,

$$\begin{aligned}
\alpha^{-1}\mathbf{I}_M - \alpha^{-1}\mathbf{I}_M\Phi^T \left( \beta^{-1}\mathbf{I}_N + \Phi\alpha^{-1}\mathbf{I}_M\Phi^T \right)^{-1} \Phi\alpha^{-1}\mathbf{I}_M &= \left( \alpha\mathbf{I}_M + \Phi^T (\beta^{-1}\mathbf{I}_N)^{-1} \Phi \right)^{-1} \\
&= \left( \alpha\mathbf{I}_M + \beta\Phi^T\Phi \right)^{-1} \\
&= \mathbf{S}_N
\end{aligned}$$

Therefore the variance becomes:

$$= \beta^{-1} + \phi(\mathbf{x})^T \mathbf{S}_N \phi(\mathbf{x})$$

The results are identical to the ones in equation 3.58.