1.25 We are given:

$$E[L(\mathbf{t},\mathbf{y}(\mathbf{x}))] = \int\int ||\mathbf{y}(\mathbf{x})-\mathbf{t}||^2\, p(\mathbf{x},\mathbf{t})\, d\mathbf{x}\, d\mathbf{t}$$

Using D.5, if $E[L(\mathbf{t},\mathbf{y}(\mathbf{x}))]$ can be treated as a functional that is defined by an integral over a function $G(\mathbf{y},\mathbf{y}',\mathbf{x})$, such that:

$$F(\mathbf{y}) = E[L(\mathbf{t},\mathbf{y}(\mathbf{x}))]$$

$$\implies \int\int ||\mathbf{y}(\mathbf{x})-\mathbf{t}||^2\, p(\mathbf{x},\mathbf{t})\, d\mathbf{x}\, d\mathbf{t} = \int G(\mathbf{y},\mathbf{y}',\mathbf{x})\, d\mathbf{x}$$

$$\implies \int ||\mathbf{y}(\mathbf{x})-\mathbf{t}||^2\, p(\mathbf{x},\mathbf{t})\, d\mathbf{t} = G(\mathbf{y},\mathbf{y}',\mathbf{x})$$

Here, the functional actually does not depend on $\mathbf{y}'$, and is of the form $G(\mathbf{y},\mathbf{x})$.

Therefore, as per page 705 in Bishop, in this case, stationarity simply requires that $\dfrac{\partial G}{\partial \mathbf{y}(\mathbf{x})} = 0$ for all values of $\mathbf{x}$.

Taking the partial derivative, we get:

$$\frac{\partial G}{\partial \mathbf{y}(\mathbf{x})} = \frac{\partial}{\partial \mathbf{y}(\mathbf{x})}\left(\int ||\mathbf{y}(\mathbf{x})-\mathbf{t}||^2\, p(\mathbf{x},\mathbf{t})\, d\mathbf{t}\right) = 0$$

$$\implies \frac{\partial}{\partial \mathbf{y}(\mathbf{x})}\left(\int (\mathbf{y}(\mathbf{x})-\mathbf{t})^T(\mathbf{y}(\mathbf{x})-\mathbf{t})\, p(\mathbf{x},\mathbf{t})\, d\mathbf{t}\right) = 0$$

$$\implies \frac{\partial}{\partial \mathbf{y}(\mathbf{x})}\left(\int (\mathbf{y}(\mathbf{x})^T\mathbf{y}(\mathbf{x}) - 2\mathbf{y}(\mathbf{x})^T\mathbf{t} + \mathbf{t}^T\mathbf{t})\, p(\mathbf{x},\mathbf{t})\, d\mathbf{t}\right) = 0$$

The derivative can be taken inside the integral as per Leibniz Integral Rule.

$$\implies \int \left(\frac{\partial(\mathbf{y}(\mathbf{x})^T\mathbf{y}(\mathbf{x}))}{\partial \mathbf{y}(\mathbf{x})} - \frac{\partial(2\mathbf{y}(\mathbf{x})^T\mathbf{t})}{\partial \mathbf{y}(\mathbf{x})} + \frac{\partial(\mathbf{t}^T\mathbf{t})}{\partial \mathbf{y}(\mathbf{x})}\right) p(\mathbf{x},\mathbf{t})\, d\mathbf{t} = 0$$

$$\implies \int \left(2\mathbf{y}(\mathbf{x})^T - 2\mathbf{t}^T + \mathbf{0}^T\right)\, p(\mathbf{x},\mathbf{t})\, d\mathbf{t} = 0$$

$$\implies \int 2\mathbf{y}(\mathbf{x})^T \, p(\mathbf{x},\mathbf{t}) \, d\mathbf{t} = \int 2\mathbf{t}^T \, p(\mathbf{x},\mathbf{t}) \, d\mathbf{t}$$

$$\implies \int \mathbf{y}(\mathbf{x})^T \, p(\mathbf{t}|\mathbf{x})p(\mathbf{x}) \, d\mathbf{t} = \int \mathbf{t}^T \, p(\mathbf{t}|\mathbf{x})p(\mathbf{x}) \, d\mathbf{t}$$

$$\implies \mathbf{y}(\mathbf{x})^T \, p(\mathbf{x}) \, \int p(\mathbf{t}|\mathbf{x}) \, d\mathbf{t} = p(\mathbf{x}) \, \int \mathbf{t}^T \, p(\mathbf{t}|\mathbf{x}) \, d\mathbf{t}$$

$$\implies \mathbf{y}(\mathbf{x})^T \, \int p(\mathbf{t}|\mathbf{x}) \, d\mathbf{t} = \int \mathbf{t}^T \, p(\mathbf{t}|\mathbf{x}) \, d\mathbf{t}$$

By using definition of conditional expectation from 1.37, we get:

$$\implies \mathbf{y}(\mathbf{x})^T \, (1) = E_{\mathbf{t}}^T[\mathbf{t}|\mathbf{x}]$$

$$\implies \mathbf{y}(\mathbf{x}) = E_{\mathbf{t}}[\mathbf{t}|\mathbf{x}]$$

For a single target variable $\mathbf{t}$, this becomes:

$$y(\mathbf{x}) = E_t[t|\mathbf{x}]$$

which is the same as 1.89.