4.8 From 4.57 and 4.58, we have:

$$p(\mathcal{C}_1|\mathbf{x}) = \sigma\left(\ln \frac{p(\mathbf{x}|\mathcal{C}_1)p(\mathcal{C}_1)}{p(\mathbf{x}|\mathcal{C}_2)p(\mathcal{C}_2)}\right)$$

Substituting the expression for class conditional densities using 4.64, this becomes:

$$= \sigma\left(\ln\left(\frac{\frac{1}{(2\pi)^{D/2}}\frac{1}{|\mathbf{\Sigma}|^{1/2}}\exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_1)^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}_1)\right\}p(\mathcal{C}_1)}{\frac{1}{(2\pi)^{D/2}}\frac{1}{|\mathbf{\Sigma}|^{1/2}}\exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_2)^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}_2)\right\}p(\mathcal{C}_2)}\right)\right)$$

$$= \sigma\left(\ln\left(\frac{\exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_1)^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}_1)\right\}p(\mathcal{C}_1)}{\exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_2)^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}_2)\right\}p(\mathcal{C}_2)}\right)\right)$$

$$= \sigma\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_1)^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}_1) + \frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_2)^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}_2) + \ln\frac{p(\mathcal{C}_1)}{p(\mathcal{C}_2)}\right)$$

$$= \sigma\left(-\frac{1}{2}\left((\mathbf{x}-\boldsymbol{\mu}_1)^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}_1) - (\mathbf{x}-\boldsymbol{\mu}_2)^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}_2)\right) + \ln\frac{p(\mathcal{C}_1)}{p(\mathcal{C}_2)}\right)$$

$$= \sigma\left(-\frac{1}{2}\left(\mathbf{x}^T\mathbf{\Sigma}^{-1}\mathbf{x} - 2\mathbf{x}^T\mathbf{\Sigma}^{-1}\boldsymbol{\mu}_1 + \boldsymbol{\mu}_1^T\mathbf{\Sigma}^{-1}\boldsymbol{\mu}_1 - \mathbf{x}^T\mathbf{\Sigma}^{-1}\mathbf{x} + 2\mathbf{x}^T\mathbf{\Sigma}^{-1}\boldsymbol{\mu}_2 - \boldsymbol{\mu}_2^T\mathbf{\Sigma}^{-1}\boldsymbol{\mu}_2\right) + \ln\frac{p(\mathcal{C}_1)}{p(\mathcal{C}_2)}\right)$$

$$= \sigma\left(-\frac{1}{2}\left(-2\mathbf{x}^T\mathbf{\Sigma}^{-1}\boldsymbol{\mu}_1 + \boldsymbol{\mu}_1^T\mathbf{\Sigma}^{-1}\boldsymbol{\mu}_1 + 2\mathbf{x}^T\mathbf{\Sigma}^{-1}\boldsymbol{\mu}_2 - \boldsymbol{\mu}_2^T\mathbf{\Sigma}^{-1}\boldsymbol{\mu}_2\right) + \ln\frac{p(\mathcal{C}_1)}{p(\mathcal{C}_2)}\right)$$

$$= \sigma\left(\mathbf{x}^T\mathbf{\Sigma}^{-1}\boldsymbol{\mu}_1 - \frac{1}{2}\boldsymbol{\mu}_1^T\mathbf{\Sigma}^{-1}\boldsymbol{\mu}_1 - \mathbf{x}^T\mathbf{\Sigma}^{-1}\boldsymbol{\mu}_2 + \frac{1}{2}\boldsymbol{\mu}_2^T\mathbf{\Sigma}^{-1}\boldsymbol{\mu}_2 + \ln\frac{p(\mathcal{C}_1)}{p(\mathcal{C}_2)}\right)$$

$$= \sigma \left(\mathbf{x}^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) - \frac{1}{2} \boldsymbol{\mu}_1^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_1 + \frac{1}{2} \boldsymbol{\mu}_2^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_2 + \ln \frac{p(\mathcal{C}_1)}{p(\mathcal{C}_2)} \right)$$

If we applying 4.66 and 4.67, this becomes:

$$= \sigma(\mathbf{x}^T \mathbf{w} + w_0)$$

$$= \sigma(\mathbf{w}^T \mathbf{x} + w_0)$$

$$\implies p(\mathcal{C}_1 | \mathbf{x}) = \sigma(\mathbf{w}^T \mathbf{x} + w_0)$$

which verifies the results in 4.65, 4.66 and 4.67.