7.3 Let's say we have 2 data points, 1 from each class. Obviously, being the only points, they will both be the "closest" to the decision surface, giving us:

$$t_1 \cdot \big(\mathbf{w}^T \boldsymbol{\phi}(\mathbf{x}_1) + b\big) = 1$$

and

$$t_2 \cdot \big(\mathbf{w}^T \boldsymbol{\phi}(\mathbf{x}_2) + b\big) = 1$$

where $t_1 = 1$ and $t_2 = -1$:

$$1 \cdot \big(\mathbf{w}^T \boldsymbol{\phi}(\mathbf{x}_1) + b\big) = 1$$

and

$$-1 \cdot \big(\mathbf{w}^T \boldsymbol{\phi}(\mathbf{x}_2) + b\big) = 1$$

If we use the results from 7.7 to 7.12, then using 7.12 we have:

$$a_1.1 + a_2.(-1) = 0 \quad \Longrightarrow \quad a_1 = a_2$$

Substituting this into 7.8, we have:

$$\mathbf{w} = a_1 t_1 \boldsymbol{\phi}(\mathbf{x}_1) + a_2 t_2 \boldsymbol{\phi}(\mathbf{x}_2) = a_1 \left(\boldsymbol{\phi}(\mathbf{x}_1) - \boldsymbol{\phi}(\mathbf{x}_2)\right)$$

Putting this back into the decision surface equations, we get:

$$a_1 \left(\boldsymbol{\phi}(\mathbf{x}_1) - \boldsymbol{\phi}(\mathbf{x}_2)\right)^T \boldsymbol{\phi}(\mathbf{x}_1) + b = 1$$

$$\Longrightarrow a_1 \big(\boldsymbol{\phi}(\mathbf{x}_1)^T \boldsymbol{\phi}(\mathbf{x}_1) - \boldsymbol{\phi}(\mathbf{x}_2)^T \boldsymbol{\phi}(\mathbf{x}_1)\big) + b = 1$$

and

$$a_1 \left(\boldsymbol{\phi}(\mathbf{x}_1) - \boldsymbol{\phi}(\mathbf{x}_2)\right)^T \boldsymbol{\phi}(\mathbf{x}_2) + b = -1$$

$$\Longrightarrow a_1 \big(\boldsymbol{\phi}(\mathbf{x}_1)^T \boldsymbol{\phi}(\mathbf{x}_2) - \boldsymbol{\phi}(\mathbf{x}_2)^T \boldsymbol{\phi}(\mathbf{x}_2)\big) + b = -1$$

Subtracting the second equation from the first,

$$\implies a_1 \left(\boldsymbol{\phi}(\mathbf{x}_1)^T \boldsymbol{\phi}(\mathbf{x}_1) - \boldsymbol{\phi}(\mathbf{x}_2)^T \boldsymbol{\phi}(\mathbf{x}_1) - \boldsymbol{\phi}(\mathbf{x}_1)^T \boldsymbol{\phi}(\mathbf{x}_2) + \boldsymbol{\phi}(\mathbf{x}_2)^T \boldsymbol{\phi}(\mathbf{x}_2) \right) = 2$$

$$\implies a_1 \left(\boldsymbol{\phi}(\mathbf{x}_1)^T \boldsymbol{\phi}(\mathbf{x}_1) - 2\boldsymbol{\phi}(\mathbf{x}_2)^T \boldsymbol{\phi}(\mathbf{x}_1) + \boldsymbol{\phi}(\mathbf{x}_2)^T \boldsymbol{\phi}(\mathbf{x}_2) \right) = 2$$

$$\implies a_1 \Big(\left(\boldsymbol{\phi}(\mathbf{x}_1) - \boldsymbol{\phi}(\mathbf{x}_2) \right)^T \left(\boldsymbol{\phi}(\mathbf{x}_1) - \boldsymbol{\phi}(\mathbf{x}_2) \right) \Big) = 2$$

$$\implies a_1 ||\boldsymbol{\phi}(\mathbf{x}_1) - \boldsymbol{\phi}(\mathbf{x}_2)||^2 = 2$$

$$\implies a_1 = \frac{2}{||\boldsymbol{\phi}(\mathbf{x}_1) - \boldsymbol{\phi}(\mathbf{x}_2)||^2}$$

Putting this back into the expression for $\mathbf{w}$, we get:

$$\implies \mathbf{w} = \frac{2 \left(\boldsymbol{\phi}(\mathbf{x}_1) - \boldsymbol{\phi}(\mathbf{x}_2) \right)}{||\boldsymbol{\phi}(\mathbf{x}_1) - \boldsymbol{\phi}(\mathbf{x}_2)||^2}$$

Finally, solving for b using the decision surface equation for $\mathbf{x}_1$:

$$\frac{2 \left(\boldsymbol{\phi}(\mathbf{x}_1) - \boldsymbol{\phi}(\mathbf{x}_2) \right)^T \boldsymbol{\phi}(\mathbf{x}_1)}{||\boldsymbol{\phi}(\mathbf{x}_1) - \boldsymbol{\phi}(\mathbf{x}_2)||^2} + b = 1$$

$$\implies b = 1 - \frac{2 \left(\boldsymbol{\phi}(\mathbf{x}_1) - \boldsymbol{\phi}(\mathbf{x}_2) \right)^T \boldsymbol{\phi}(\mathbf{x}_1)}{||\boldsymbol{\phi}(\mathbf{x}_1) - \boldsymbol{\phi}(\mathbf{x}_2)||^2}$$

$$= \frac{\boldsymbol{\phi}(\mathbf{x}_1)^T \boldsymbol{\phi}(\mathbf{x}_1) - 2\boldsymbol{\phi}(\mathbf{x}_2)^T \boldsymbol{\phi}(\mathbf{x}_1) + \boldsymbol{\phi}(\mathbf{x}_2)^T \boldsymbol{\phi}(\mathbf{x}_2) - 2\boldsymbol{\phi}(\mathbf{x}_1)^T \boldsymbol{\phi}(\mathbf{x}_1) + 2\boldsymbol{\phi}(\mathbf{x}_2)^T \boldsymbol{\phi}(\mathbf{x}_1)}{||\boldsymbol{\phi}(\mathbf{x}_1) - \boldsymbol{\phi}(\mathbf{x}_2)||^2}$$

$$= \frac{-\boldsymbol{\phi}(\mathbf{x}_1)^T \boldsymbol{\phi}(\mathbf{x}_1) + \boldsymbol{\phi}(\mathbf{x}_2)^T \boldsymbol{\phi}(\mathbf{x}_2)}{||\boldsymbol{\phi}(\mathbf{x}_1) - \boldsymbol{\phi}(\mathbf{x}_2)||^2}$$

Verifying that the equation for $\mathbf{x}_2$ gives the same result:

$$\frac{2 \left(\boldsymbol{\phi}(\mathbf{x}_1) - \boldsymbol{\phi}(\mathbf{x}_2) \right)^T \boldsymbol{\phi}(\mathbf{x}_2)}{||\boldsymbol{\phi}(\mathbf{x}_1) - \boldsymbol{\phi}(\mathbf{x}_2)||^2} + b = -1$$

$$\implies b = -1 - \frac{2\left(\boldsymbol{\phi}(\mathbf{x}_1) - \boldsymbol{\phi}(\mathbf{x}_2)\right)^T \boldsymbol{\phi}(\mathbf{x}_2)}{||\boldsymbol{\phi}(\mathbf{x}_1) - \boldsymbol{\phi}(\mathbf{x}_2)||^2}$$

$$= \frac{-\boldsymbol{\phi}(\mathbf{x}_1)^T \boldsymbol{\phi}(\mathbf{x}_1) + 2\boldsymbol{\phi}(\mathbf{x}_2)^T \boldsymbol{\phi}(\mathbf{x}_1) - \boldsymbol{\phi}(\mathbf{x}_2)^T \boldsymbol{\phi}(\mathbf{x}_2) - 2\boldsymbol{\phi}(\mathbf{x}_1)^T \boldsymbol{\phi}(\mathbf{x}_2) + 2\boldsymbol{\phi}(\mathbf{x}_2)^T \boldsymbol{\phi}(\mathbf{x}_2)}{||\boldsymbol{\phi}(\mathbf{x}_1) - \boldsymbol{\phi}(\mathbf{x}_2)||^2}$$

$$= \frac{-\boldsymbol{\phi}(\mathbf{x}_1)^T \boldsymbol{\phi}(\mathbf{x}_1) + \boldsymbol{\phi}(\mathbf{x}_2)^T \boldsymbol{\phi}(\mathbf{x}_2)}{||\boldsymbol{\phi}(\mathbf{x}_1) - \boldsymbol{\phi}(\mathbf{x}_2)||^2}$$

Therefore, we have determined the values of both $\mathbf{w}$ and b.