7.4 The value ρ of the margin for the maximum-margin hyperplane is given by the perpendicular distance of the decision boundary from support vector $\mathbf{x}_n$, given by equation 7.2:

$$\rho = \frac{t_n \left(\mathbf{w}^T \boldsymbol{\phi}(\mathbf{x}_n) + b\right)}{||\mathbf{w}||}$$

Since we scale $\mathbf{w}$ such that $t_n \left(\mathbf{w}^T \boldsymbol{\phi}(\mathbf{x}_n) + b\right) = 1$, we get:

$$\rho = \frac{1}{||\mathbf{w}||}$$

$$\implies \frac{1}{\rho^2} = ||\mathbf{w}||^2$$

which is the same as the result in 7.125.

Since the second norm is a convex function and the constraints are all linear, strong duality holds and we can say that:

$$\min\ L(\mathbf{w}) = \max\ \tilde{L}(\mathbf{a})$$

$$\implies \frac{1}{2}||\mathbf{w}||^2 = \sum_{n=1}^{N} a_n - \frac{1}{2}\sum_{n=1}^{N}\sum_{m=1}^{N} a_n a_m t_n t_m \boldsymbol{\phi}(\mathbf{x}_n)^T \boldsymbol{\phi}(\mathbf{x}_m)$$

$$\implies \frac{1}{2}||\mathbf{w}||^2 = \sum_{n=1}^{N} a_n - \frac{1}{2}||\mathbf{w}||^2$$

$$\implies ||\mathbf{w}||^2 = \frac{1}{\rho^2} = \sum_{n=1}^{N} a_n$$

where from 7.8, we know that:

$$||\mathbf{w}||^2 = \left(\sum_{n=1}^{N} a_n t_n \boldsymbol{\phi}(\mathbf{x}_n)\right)^T \left(\sum_{m=1}^{N} a_m t_m \boldsymbol{\phi}(\mathbf{x}_m)\right)$$

$$= \sum_{n=1}^{N}\sum_{m=1}^{N} a_n a_m t_n t_m \boldsymbol{\phi}(\mathbf{x}_n)^T \boldsymbol{\phi}(\mathbf{x}_m)$$