Equation 4.55 is:

$$\mathbf{w}^{(\tau+1)} = \mathbf{w}^{(\tau)} + \eta \boldsymbol{\phi}_n t_n$$

Assuming the algorithm makes α passes over the dataset, each $\alpha_n = \alpha$,

$$\implies \mathbf{w}^{(\tau+1)} = \mathbf{w}^{(0)} + \eta \left(\sum_n \boldsymbol{\phi}_n t_n \alpha \right)$$

$$= \mathbf{w}^{(0)} + \eta\alpha \left(\sum_n \boldsymbol{\phi}_n t_n \right)$$

We can consider the initialization $\mathbf{w}^{(0)} = 0$. This makes the learned weight vector $\mathbf{w}$ a linear combination of the vectors $t_n \boldsymbol{\phi}_n$.

$$\implies \mathbf{w}^{(\tau+1)} = \eta\alpha \sum_n \boldsymbol{\phi}_n t_n = \eta \, \alpha \, \boldsymbol{\Phi}^\top \boldsymbol{t}$$

The predictive function is given by 4.52 and 4.53:

$$y(\mathbf{x}) = f(\mathbf{w}^T \boldsymbol{\phi}(\mathbf{x})) = sign(\mathbf{w}^T \boldsymbol{\phi}(\mathbf{x}))$$

$$= sign \left(\left(\eta\alpha \sum_n \boldsymbol{\phi}_n t_n \right)^T \boldsymbol{\phi}(\mathbf{x}) \right)$$

Since η and α are positive terms, this becomes:

$$= sign \left(\sum_n t_n \left(\boldsymbol{\phi}(\mathbf{x})_n^T \boldsymbol{\phi}(\mathbf{x}) \right) \right)$$

$$= sign \left(\sum_n t_n \, k(\mathbf{x}_n, \mathbf{x}) \right)$$

So, the feature vector enters the predictive function only in the form of the kernel function.