The error function we wish to minimize is:

$$E(\boldsymbol{w}) = \frac{1}{2}\sum_{n=1}^{N}\{t_n - \boldsymbol{w}^T\boldsymbol{\phi}(\boldsymbol{x}_n)\}^2 \tag{1}$$

We take its derivatives and set them to zero:

$$\nabla E(\boldsymbol{w}) = -\sum_{n=1}^{N}\{t_n - \boldsymbol{w}^T\boldsymbol{\phi}(\boldsymbol{x}_n)\}\boldsymbol{\phi}(\boldsymbol{x}_n)^T \tag{2}$$

$$\nabla E(\boldsymbol{w}) \;=\; 0^T \tag{3}$$

$$\Leftrightarrow \sum_{n=1}^{N} t_n\boldsymbol{\phi}(\boldsymbol{x}_n)^T \;=\; \sum_{n=1}^{N}\boldsymbol{w}^T\boldsymbol{\phi}(\boldsymbol{x}_n)\boldsymbol{\phi}(\boldsymbol{x}_n)^T \tag{4}$$

In order to write these sums using matrix notation, define the matrix $\boldsymbol{\Phi}$:

$$\boldsymbol{\Phi} = \begin{pmatrix} \phi_0(\boldsymbol{x}_1) & \phi_1(\boldsymbol{x}_1) & \dots & \phi_{M-1}(\boldsymbol{x}_1) \\ \phi_0(\boldsymbol{x}_2) & \phi_1(\boldsymbol{x}_2) & \dots & \phi_{M-1}(\boldsymbol{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_0(\boldsymbol{x}_N) & \phi_1(\boldsymbol{x}_N) & \dots & \phi_{M-1}(\boldsymbol{x}_N) \end{pmatrix} \tag{5}$$

Now,

$$\nabla E(\boldsymbol{w}) = 0^T \Leftrightarrow \boldsymbol{t}^T\boldsymbol{\Phi} \;=\; \boldsymbol{w}^T\boldsymbol{\Phi}^T\boldsymbol{\Phi} \tag{6}$$

Taking transpose of both sides:

$$\boldsymbol{\Phi}^T\boldsymbol{t} \;=\; \boldsymbol{\Phi}^T\boldsymbol{\Phi}\boldsymbol{w} \tag{7}$$

Assuming the inverse of $\boldsymbol{\Phi}^T\boldsymbol{\Phi}$ exists (e.g. basis functions are not linearly dependent):

$$\boldsymbol{w} \;=\; \left(\boldsymbol{\Phi}^T\boldsymbol{\Phi}\right)^{-1}\boldsymbol{\Phi}^T\boldsymbol{t} \tag{8}$$