1.2 The regularized sum-of-squares error function given by (1.4) is;

$$\tilde{E}(\mathbf{w}) = \frac{1}{2}\sum_{n=1}^{N}\{y(x_n, \mathbf{w}) - t_n\}^2 + \frac{\lambda}{2}||\mathbf{w}||^2$$

To minimize it, we take the derivative of this error function w.r.t $\mathbf{w}$, which gives us a row vector, and set it to 0:

$$\frac{d\tilde{E}(\mathbf{w})}{d\mathbf{w}} = \frac{d}{d\mathbf{w}}\left(\frac{1}{2}\sum_{n=1}^{N}\{y(x_n, \mathbf{w}) - t_n\}^2 + \frac{\lambda}{2}||\mathbf{w}||^2\right)$$

$$= \frac{1}{2}\sum_{n=1}^{N}\frac{d}{d\mathbf{w}}\{y(x_n, \mathbf{w}) - t_n\}^2 + \frac{d}{d\mathbf{w}}\frac{\lambda}{2}||\mathbf{w}||^2$$

$$= \frac{1}{2}\sum_{n=1}^{N}\frac{d}{d\{y(x_n, \mathbf{w}) - t_n\}}\{y(x_n, \mathbf{w}) - t_n\}^2 * \frac{d(y(x_n, \mathbf{w}) - t_n)}{d\mathbf{w}} + \frac{\lambda}{2}\frac{d}{d\mathbf{w}}||\mathbf{w}||^2$$

$$= \frac{1}{2}\sum_{n=1}^{N}2\{y(x_n, \mathbf{w}) - t_n\} * \begin{bmatrix} 1 & x_n & \dots & (x_n)^M \end{bmatrix} + \frac{\lambda}{2}\frac{d}{d\mathbf{w}}\mathbf{w}^T\mathbf{w}$$

$$= \frac{1}{2}\sum_{n=1}^{N}2\{y(x_n, \mathbf{w}) - t_n\} * \begin{bmatrix} 1 & x_n & \dots & (x_n)^M \end{bmatrix} + \frac{\lambda}{2}2\mathbf{w}^T$$

$$= \sum_{n=1}^{N}\{y(x_n, \mathbf{w}) - t_n\}\begin{bmatrix} 1 & x_n & \dots & (x_n)^M \end{bmatrix} + \lambda\mathbf{w}^T$$

$$= \sum_{n=1}^{N}\left(\sum_{i=0}^{M}w_i(x_n)^i - t_n\right)\begin{bmatrix} 1 & x_n & \dots & (x_n)^M \end{bmatrix} + \lambda\mathbf{w}^T$$

If we set this derivative to $\mathbf{0}^T$, we get the following set of equations:

$$\sum_{n=1}^{N}\left(\sum_{i=0}^{M}w_i(x_n)^i - t_n\right)(x_n)^j + \lambda w_j = 0 \quad \forall j = 0...M$$

$$\implies \sum_{n=1}^{N}\left(\sum_{i=0}^{M}w_i(x_n)^i(x_n)^j - t_n(x_n)^j\right) + \lambda w_j = 0$$

$$\implies \sum_{n=1}^{N}\left(\sum_{i=0}^{M}w_i(x_n)^{i+j} - t_n(x_n)^j\right) + \lambda w_j = 0$$

$$\implies \sum_{n=1}^{N}\sum_{i=0}^{M} w_i (x_n)^{i+j} - \sum_{n=1}^{N} t_n (x_n)^j + \lambda w_j = 0$$

$$\implies \sum_{i=0}^{M}\sum_{n=1}^{N} w_i (x_n)^{i+j} - \sum_{n=1}^{N} t_n (x_n)^j + \lambda w_j = 0$$

$$\implies \sum_{i=0}^{M} w_i \sum_{n=1}^{N} (x_n)^{i+j} - \sum_{n=1}^{N} t_n (x_n)^j + \lambda w_j = 0$$

$$\implies \sum_{i=0}^{M} w_i A_{ij} - T_j + \lambda w_j = 0$$

$$\implies \sum_{i=0}^{M} w_i A_{ij} + \lambda w_j = T_j$$