1.1 The error function given by 1.2 is:

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

$$\text{where } y(x, \mathbf{w}) = w_0 + w_1 x + w_2 x^2 + \ldots + w_M x^M = \sum_{j=0}^{M} w_j x^j$$

To minimize the error function w.r.t $\mathbf{w}$, we take its derivative w.r.t $\mathbf{w}$ and set it to 0.

Here, E($\mathbf{w}$) is a scalar and $\mathbf{w}$ is a vector.

The derivative of a scalar valued function with respect to a vector of variables then is a row vector. This row vector has one column for each variable we want to differentiate by.

Therefore,

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

$$= \sum_{n=1}^{N}\{y(x_n, \mathbf{w}) - t_n\}\left[\begin{array}{cccc} \frac{d\{y(x_n, \mathbf{w}) - t_n\}}{dw_0} & \frac{d\{y(x_n, \mathbf{w}) - t_n\}}{dw_1} & \ldots & \frac{d\{y(x_n, \mathbf{w}) - t_n\}}{dw_M} \end{array}\right]$$

$$= \sum_{n=1}^{N}\{y(x_n, \mathbf{w}) - t_n\}\left[\begin{array}{cccc} 1 & x_n & \ldots & (x_n)^M \end{array}\right]$$

Setting the derivative to $\mathbf{0}$ row vector, we get the following set of equations:

$$\sum_{n=1}^{N}\{y(x_n, \mathbf{w}) - t_n\}(x_n)^j = 0, \quad \forall j = 0..M$$

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

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

$$\implies \sum_{n=1}^{N}\sum_{i=1}^{M} w_i (x_n)^i (x_n)^j = \sum_{n=1}^{N} t_n (x_n)^j$$

Here we just multiplied $(x_n)^j$ to every term inside the summation

$$\implies \sum_{n=1}^{N} \sum_{i=1}^{M} w_i (x_n)^{i+j} = \sum_{n=1}^{N} t_n (x_n)^j$$
$$\implies \sum_{i=1}^{M} \sum_{n=1}^{N} w_i (x_n)^{i+j} = \sum_{n=1}^{N} t_n (x_n)^j$$
$$\implies \sum_{i=1}^{M} w_i \sum_{n=1}^{N} (x_n)^{i+j} = \sum_{n=1}^{N} t_n (x_n)^j$$
$$\implies \sum_{i=1}^{M} w_i A_{ij} = T_j \quad \text{by applying 1.123}$$

Note: i and j can be switched with no effect, as they both are indices going from 0 to M.