4.13 The error function is given by 4.90:

$$E(\mathbf{w}) = -\sum_{n=1}^{N}\{t_n \ln y_n + (1-t_n)\ln(1-y_n)\}$$

From 4.87, we know that $y = \sigma(\mathbf{w}^T\boldsymbol{\phi})$. Substituting for y, we get:

$$E(\mathbf{w}) = -\sum_{n=1}^{N}\{t_n \ln(\sigma(\mathbf{w}^T\boldsymbol{\phi}_n)) + (1-t_n)\ln(1-\sigma(\mathbf{w}^T\boldsymbol{\phi}_n))\}$$

Taking the gradient of the error function with respect to $\mathbf{w}$, we obtain:

$$\frac{\partial E(\mathbf{w})}{\partial \mathbf{w}} = -\frac{\partial}{\partial \mathbf{w}}\sum_{n=1}^{N}\{t_n \ln(\sigma(\mathbf{w}^T\boldsymbol{\phi}_n)) + (1-t_n)\ln(1-\sigma(\mathbf{w}^T\boldsymbol{\phi}_n))\}$$

$$= -\sum_{n=1}^{N}\left\{\frac{\partial t_n \ln(\sigma(\mathbf{w}^T\boldsymbol{\phi}_n))}{\partial \mathbf{w}} + \frac{\partial(1-t_n)\ln(1-\sigma(\mathbf{w}^T\boldsymbol{\phi}_n))}{\partial \mathbf{w}}\right\}$$

Evaluating the derivative of the first term inside the summation:

$$\frac{\partial t_n \ln(\sigma(\mathbf{w}^T\boldsymbol{\phi}_n))}{\partial \mathbf{w}}$$

$$= t_n \frac{\partial \ln(\sigma(\mathbf{w}^T\boldsymbol{\phi}_n))}{\partial \mathbf{w}}$$

$$= t_n \left(\frac{\partial \ln(\sigma(\mathbf{w}^T\boldsymbol{\phi}_n))}{\partial \sigma(\mathbf{w}^T\boldsymbol{\phi}_n)}\right)\left(\frac{\partial \sigma(\mathbf{w}^T\boldsymbol{\phi}_n)}{\partial(\mathbf{w}^T\boldsymbol{\phi}_n)}\right)\left(\frac{\partial(\mathbf{w}^T\boldsymbol{\phi}_n)}{\partial \mathbf{w}}\right)$$

$$= t_n \left(\frac{1}{\sigma(\mathbf{w}^T\boldsymbol{\phi}_n)}\right)\left(\sigma(\mathbf{w}^T\boldsymbol{\phi}_n)(1-\sigma(\mathbf{w}^T\boldsymbol{\phi}_n))\right)(\boldsymbol{\phi}_n)$$

$$= t_n\left(1-\sigma(\mathbf{w}^T\boldsymbol{\phi}_n)\right)(\boldsymbol{\phi}_n)$$

$$= t_n\boldsymbol{\phi}_n - t_n\sigma(\mathbf{w}^T\boldsymbol{\phi}_n)\boldsymbol{\phi}_n$$

Evaluating the derivative of the second term inside the summation:

$$\frac{\partial(1-t_n)\ln(1-\sigma(\mathbf{w}^T\boldsymbol{\phi}_n))}{\partial\mathbf{w}}$$

$$= (1-t_n)\frac{\partial\ln(1-\sigma(\mathbf{w}^T\boldsymbol{\phi}_n))}{\partial\mathbf{w}}$$

$$= (1-t_n)\left(\frac{\partial\ln(1-\sigma(\mathbf{w}^T\boldsymbol{\phi}_n))}{\partial(1-\sigma(\mathbf{w}^T\boldsymbol{\phi}_n))}\right)\left(\frac{\partial(1-\sigma(\mathbf{w}^T\boldsymbol{\phi}_n))}{\partial(\mathbf{w}^T\boldsymbol{\phi}_n)}\right)\left(\frac{\partial(\mathbf{w}^T\boldsymbol{\phi}_n)}{\partial\mathbf{w}}\right)$$

$$= (1-t_n)\left(\frac{1}{(1-\sigma(\mathbf{w}^T\boldsymbol{\phi}_n))}\right)\left(-\sigma(\mathbf{w}^T\boldsymbol{\phi}_n)(1-\sigma(\mathbf{w}^T\boldsymbol{\phi}_n))\right)(\boldsymbol{\phi}_n)$$

$$= (1-t_n)\left(-\sigma(\mathbf{w}^T\boldsymbol{\phi}_n)\right)(\boldsymbol{\phi}_n)$$

$$= t_n\sigma(\mathbf{w}^T\boldsymbol{\phi}_n)\boldsymbol{\phi}_n - \sigma(\mathbf{w}^T\boldsymbol{\phi}_n)\boldsymbol{\phi}_n$$

Substituting these derivatives back into the error function, we get:

$$\frac{\partial E(\mathbf{w})}{\partial\mathbf{w}} = -\sum_{n=1}^{N}\left\{t_n\boldsymbol{\phi}_n - t_n\sigma(\mathbf{w}^T\boldsymbol{\phi}_n)\boldsymbol{\phi}_n + t_n\sigma(\mathbf{w}^T\boldsymbol{\phi}_n)\boldsymbol{\phi}_n - \sigma(\mathbf{w}^T\boldsymbol{\phi}_n)\boldsymbol{\phi}_n\right\}$$

$$= -\sum_{n=1}^{N}\left\{t_n\boldsymbol{\phi}_n - \sigma(\mathbf{w}^T\boldsymbol{\phi}_n)\boldsymbol{\phi}_n\right\}$$

$$= -\sum_{n=1}^{N}\left\{t_n - \sigma(\mathbf{w}^T\boldsymbol{\phi}_n)\right\}\boldsymbol{\phi}_n$$

$$= -\sum_{n=1}^{N}\left\{t_n - y_n\right\}\boldsymbol{\phi}_n$$

$$= \sum_{n=1}^{N}\left\{y_n - t_n\right\}\boldsymbol{\phi}_n$$

which is the same as 4.91.