Ex - 4.3

$$\tanh(a) = \frac{e^{a} - e^{-a}}{e^{a} + e^{-a}}$$

$$2\sigma(2a) - 1 = \frac{2}{1 + e^{-2a}} - 1$$

$$= \frac{2}{1 + e^{-2a}} - \frac{1 + e^{-2a}}{1 + e^{-2a}}$$

$$= \frac{1 - e^{-2a}}{1 + e^{-2a}}$$

$$\frac{1}{e^{-a}} = e^{a}$$

$$= \frac{e^{a} - e^{-a}}{e^{a} + e^{-a}} = \tanh(a)$$

Ex - 4.6

Ridge Regression

$$y(x) = w^{T} \phi(x)$$

$\phi(x) \longrightarrow$ vector of basis functions

ordinary least squares error (OLS)

$$E_D(w) = \frac{1}{2} \sum_{n=1}^{N} (t_n - w_T \phi(x_n))^2$$

we add a regularization term:

$$E_w(w) = \frac{1}{2} \sum_{j} w_j^2 = \frac{1}{2} w^T w$$

Total error function:

$$\frac{1}{2} \sum_{n=1}^{N} \{t_n - w^T \phi(x_n)\}^2 + \frac{\lambda}{2} w^T w$$

Matrix form of the data error term

$$E(w) = \frac{1}{2}(t - \phi w)^{T}(t - \phi w) + \frac{1}{2} w^T w$$

$$\nabla_w E(w) = 0$$

$$E_D = \frac{1}{2}(t - \phi w)^T (t - \phi w)$$

$$\nabla_w E_D = -\phi^T (t - \phi w) = \phi^T \phi w - \phi^T t.$$

$$\left[\nabla_w \frac{1}{2} \|Aw - b\| = A^T(Aw - b)\right]$$

$$E_w = \frac{\lambda}{2} w^T w$$

$$\nabla_w E_w = \lambda w$$

$$\therefore \nabla_w E(w) = (\phi^T \phi w - \phi^T t) + \lambda w$$

Set this equal to 0:

$$0 = -\phi^T t + \phi^T \phi w + \lambda w$$

$$\Rightarrow \quad \phi^T t = \phi^T \phi w + \lambda w$$

$$\phi^T t = (\phi^T \phi + \lambda) w$$

$$w = (\phi^T \phi + \lambda I)^{-1} \phi^T t.$$

Ex- 4.1

$$(1.1): \quad y(x, w) = w_0 + w_1 x + \cdots + w_M x^M = \sum_{j=0}^{M} w_j x^j$$

$$(1.2): \quad E(w) = \frac{1}{2} \sum_{n=1}^{M} \{y(x, w) - t_n\}^2$$

Substitute (1.1) in (1.2)

$$E(w) = \frac{1}{2} \sum_{n=1}^{N} \left[\left(\sum_{j=1}^{M} w_j x_n^j\right) - t_n\right]^2$$

$$\frac{\delta E}{\delta w_i} = 0$$

$$E_n = \frac{1}{2}\left[\left(\sum_{j=0}^{M} w_j x_n^j\right) - t_n\right]^2$$

$$\frac{\delta E_n}{\delta w_i} = \left[\left(\sum_{j=0}^{M} w_j x_n^j\right) - t_n\right] \cdot x_n^i$$

Now sum over all data-points:

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

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

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