3.5

(x_n, t_n) : サンプル

r_n : 重み

重み付き二乗誤差

$$E_w(w) = \sum_{n=1}^{N} r_n (w^T x_n - t_n)^2$$

(1)

$$R = \begin{pmatrix} r_1 & 0 & \cdots & 0 \\ 0 & r_2 & & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & & r_N \end{pmatrix}$$

$\dfrac{\partial E}{\partial w} = 0$ なる w を求める.

$$\frac{\partial E}{\partial w} = \frac{\partial}{\partial w} \sum_{n=1}^{N} r_n (w^T x_n - t_n)^2$$

$$= \sum_{n=1}^{N} \frac{\partial}{\partial w} r_n (w^T x_n - t_n)^2$$

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

$$= 2 \sum_{n=1}^{N} r_n (x_n w^T x_n - t_n x_n)$$

$$= 2R \sum_{n=1}^{N} (x_n w^T x_n - t_n x_n)$$

ここで、$X = \begin{pmatrix} x_1^T \\ x_2^T \\ \vdots \\ x_N^T \end{pmatrix}$ とすると、

$$\sum_{n=1}^{N} x_n w^T x_n = (x_1 \ x_2 \ \cdots \ x_N) \begin{pmatrix} w^T x_1 \\ w^T x_2 \\ \vdots \\ w^T x_N \end{pmatrix}$$

$$= \begin{pmatrix} x_1^T \\ x_2^T \\ \vdots \\ x_N^T \end{pmatrix}^T \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N^T \end{pmatrix} \begin{pmatrix} w_0 \\ w_1 \\ \vdots \\ w_N \end{pmatrix} = X^T X w$$

と表せる.

よって.

$$\frac{\partial E}{\partial \mathbf{w}} = 2\mathbb{R}\left\{\left(\sum_{n=1}^{N} \mathbf{x}_n \mathbf{w}^T \mathbf{x}_n\right) - \left(\sum_{n=1}^{N} t_n \mathbf{x}_n\right)\right\}$$

$$= 2\mathbb{R}\left\{\left(\mathbb{X}^T\mathbb{X}\mathbf{w}\right) - \sum_{n=1}^{N} t_n \mathbf{x}_n\right\} = 0$$

これより

$$\mathbb{R}\mathbb{X}^T\mathbb{X}\mathbf{w} = \mathbb{R}\sum_{n=1}^{N} t_n \mathbf{x}_n$$

両辺に左から $(\mathbb{R}\mathbb{X}^T\mathbb{X})^{-1}$ をかけて.（$(\mathbb{R}\mathbb{X}^T\mathbb{X})^{-1}$ は正則）

$$\mathbf{w} = (\mathbb{R}\mathbb{X}^T\mathbb{X})^{-1}\mathbb{R}\sum_{n=1}^{N} t_n \mathbf{x}_n \quad \left(= \left(\begin{pmatrix} r_1\mathbf{x}_1 & \cdots & r_N\mathbf{x}_N \end{pmatrix} \begin{pmatrix} \mathbf{x}_1^T \\ \vdots \\ \mathbf{x}_N^T \end{pmatrix} \right)^{-1} \begin{pmatrix} r_1\mathbf{x}_1, & \cdots & r_N\mathbf{x}_N \end{pmatrix} \begin{pmatrix} t_1 \\ \vdots \\ t_N \end{pmatrix} \right)$$

(2) r_n を、「サンプルデータの信頼性」を表す値とする.
サンプルデータの信頼性とは、具体的には、volatile acidity などの特徴を計測するための器具の精度や、計測する人の熟練度・信頼性など指し、それらが高い場合、高品質なデータであるといえる. …(a)

重みをつけなかった場合、精度が低い計測器具のデータや、未熟であったり、作業を真面目に行わない計測者のデータを、高精度な器具による計測データ、ベテラン・真面目な計測者のデータと同等に扱ってしまうことになってしまう. そのため、このような重みをつけることで、低品質のデータの影響を小さくし、高品質のデータがより近似に影響するようにできるというメリットが期待できる. …(b)