

演習問題 1.02

正則化された二乗和誤差関数 (1.4) を最小にする係数 w_i が満たす、式 (1.122) に類似した線形方程式系を書き下せ。

$$\sum_{j=0}^M A_{ij} w_j = T_i$$

… (1.122)

ただし、

$$A_{ij} = \sum_{n=1}^N (x_n)^{i+j}, \quad T_i = \sum_{n=1}^N (x_n)^i t_n$$

… (1.123)

ここで、下付き添え字の i や j は成分を表し、 $(x)^i$ は x の j 乗を表す。

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

… (1.1)

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

… (1.4)

[解]

式 (1.4) に式 (1.1) を代入することにより、式 (1.4) は以下のように変形できる。

$$\tilde{E}(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^N \left\{ \sum_{j=0}^M w_j x_n^j - t_n \right\}^2 + \frac{\lambda}{2} \sum_{j=0}^M w_j^2$$

ここで、 $\tilde{E}(\mathbf{w})$ を最小化するために、 $\tilde{E}(\mathbf{w})$ を w_i について微分すると、

$$\frac{d\tilde{E}(\mathbf{w})}{dw_i} = \frac{1}{2} \cdot 2 \sum_{n=1}^N \left(\sum_{j=0}^M w_j x_n^j - t_n \right) \left(\sum_{j=0}^M w_j x_n^j - t_n \right)' + 2 \cdot \frac{\lambda}{2} w_i$$

$\mathbf{w} = w_j = \{w_0, w_1, \cdots, w_M\}$ であり、 w_i は w_j の中に含まれる 1 個の要素である。これは、すなわち

$$y(x_n, \mathbf{w}) = \sum_{j=0}^M w_j x_n^j = w_0 x_n^0 + w_1 x_n^1 + w_2 x_n^2 + \cdots + w_i x_n^i + \cdots + w_M x_n^M$$

を満たす。ここで、 i の範囲は、 $0 \leq i \leq M$ となる。 $y(x_n, \mathbf{w})$ を w_i で微分すると、

$$\frac{d y(x_n, \mathbf{w})}{d w_i} = x_n^i$$

となる。

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

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

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

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

上記の式は、式 (1.123) を用いることにより、

$$\sum_{j=0}^M A_{ij} w_j = T_i - \lambda w_i$$

と書き表すことができる。以上より、正則化された二乗和誤差関数 (1.4) を最小にする係数 w_i が満たす、式 (1.122) に類似した線形方程式系で書き下すことができた。