Chapter 3 Linear Models For Regression

Yue Yu

3.1 Linear Basis Function Models

(3.8)

Recall $t = y(\mathbf{x}, \mathbf{w}) + \epsilon$, where $\epsilon \sim \mathcal{N}(0, \beta^{-1})$. This is equivalent to

$$
\begin{aligned}
p(\epsilon; \beta) &= \frac{\beta}{\sqrt{2\pi}} \exp\left\{ -\frac{\beta}{2} \epsilon^2 \right\} \\
&= \frac{\beta}{\sqrt{2\pi}} \exp\left\{ -\frac{\beta}{2} (t - y(\mathbf{x}, \mathbf{w}))^2 \right\},
\end{aligned}
$$

which implies that

$$
p(t|\mathbf{x}, \mathbf{w}, \beta) = \mathcal{N}(t|y(\mathbf{x}, \mathbf{w}), \beta^{-1}).
$$

(3.13)

This equation should be

$$
\nabla_{\mathbf{w}} \ln p(\mathbf{t}|\mathbf{w}, \beta) = \beta \sum_{n=1}^{N} \left\{ t_n - \mathbf{w}^{\mathrm{T}} \boldsymbol{\phi}(\mathbf{x}_n) \right\} \boldsymbol{\phi}(\mathbf{x}_n), \tag{*}
$$

because

$$
\nabla_{\mathbf{w}} (\mathbf{w}^{\mathrm{T}} \boldsymbol{\phi}(\mathbf{x}_n)) = \boldsymbol{\phi}(\mathbf{x}_n).
$$

(3.14)

According to (*), this equation should be

$$
\mathbf{0} = \sum_{n=1}^{N} t_n \boldsymbol{\phi}(\mathbf{x_n}) - \left(\sum_{n=1}^{N} \boldsymbol{\phi}(\mathbf{x_n}) \boldsymbol{\phi}(\mathbf{x_n})^{\mathrm{T}} \right) \mathbf{w}. \tag{**}
$$

(3.15)

By defining a design matrix $\boldsymbol{\Phi}$ in the form of (3.16), (**) can be reduced to

$$
\mathbf{0} = \boldsymbol{\Phi} \mathbf{t} - \boldsymbol{\Phi}^{\mathrm{T}} \boldsymbol{\Phi} \mathbf{w}.
$$

Solving for $\mathbf{w}$, we obtain

$$
\mathbf{w}_{\mathrm{ML}} = (\boldsymbol{\Phi}^{\mathrm{T}} \boldsymbol{\Phi})^{-1} \boldsymbol{\Phi} \mathbf{t}.
$$

(3.19)

$$
\begin{aligned}
\frac{\partial}{\partial w_0} E_D(\mathbf{w}) &= -\sum_{n=1}^{N} \left\{ t_n - w_0 - \sum_{j=1}^{M-1} w_j \phi_j(\mathbf{x}_n) \right\} \\
&= 0.
\end{aligned}
$$

Solving for w_0, we obtain

$$w_0 = \frac{1}{N} \sum_{n=1}^{N} t_n - \frac{1}{N} \sum_{j=1}^{M-1} w_j \sum_{n=1}^{N} \phi_j(\mathbf{x}_n)$$

$$= \bar{t} - \sum_{j=1}^{M-1} w_j \overline{\phi_j},$$

where

$$\bar{t} = \frac{1}{N} \sum_{n=1}^{N} t_n, \qquad \overline{\phi_j} = \frac{1}{N} \sum_{n=1}^{N} \phi_j(\mathbf{x}_n).$$