6.1

ノートのタイトル 2016/01/26

最小二乗法のアルゴリズム

(6.2) $J(w) = \frac{1}{2}\sum_{n=1}^{N}\{w^T\phi(x_n) - t_n\}^2 + \frac{\lambda}{2}w^Tw$ … ゴサ関数

(6.3) $w = -\frac{1}{\lambda}\sum_{n=1}^{N}\{w^T\phi(x_n) - t_n\}\phi(x_n)$ … $\nabla_w J = 0$ の解

(6.4) $y(x) = w^T\phi(x)$ … y の予測値

双対表現

(6.7) $J(a) = \frac{1}{2}a^TKKa - a^TKt + \frac{1}{2}t^Tt + \frac{\lambda}{2}a^TKa$ … ゴサ関ね

(6.8) $a = (K + \lambda I_N)^{-1}t$ … $\nabla_a J = 0$ の解

(6.9) $y(x) = a^T\Phi\phi(x) = a^Tk(x)$ … y の予測値

(a_n が $\phi(x_n)$ の要素の線形結合で書けることを示す。)

つまり $a_n = \phi_1(x_n)u_1 + \phi_2(x_n)u_2 = \phi^T(x_n)u$ という風に書けることを示す.

(6.8) より

$$(K + \lambda I_N)a = t$$

$N=3$ として成分で書くと

$$\left\{\begin{pmatrix}\phi(x_1)^T\phi(x_1) & \phi(x_1)^T\phi(x_2) & \phi(x_1)^T\phi(x_3)\\ \phi(x_2)^T\phi(x_1) & \phi(x_2)^T\phi(x_2) & \phi(x_2)^T\phi(x_3)\\ \phi(x_3)^T\phi(x_1) & \phi(x_3)^T\phi(x_2) & \phi(x_3)^T\phi(x_3)\end{pmatrix} + \begin{pmatrix}\lambda & 0 & 0\\ 0 & \lambda & 0\\ 0 & 0 & \lambda\end{pmatrix}\right\}\begin{pmatrix}a_1\\a_2\\a_3\end{pmatrix} = \begin{pmatrix}t_1\\t_2\\t_3\end{pmatrix}$$

これより

$$\phi(x_1)^T\underbrace{\{\phi(x_1)a_1 + \phi(x_2)a_2 + \phi(x_3)a_3\}}_{=\Phi^Ta} = a_1\left(-\lambda + \frac{t_1}{a_1}\right)$$

$$\therefore\ a_1 = \phi(x_1)^T\Phi^Ta\,l_1,\quad l_1 = \left(-\lambda + \frac{t_1}{a_1}\right)^{-1}$$

$$a_2 = \phi(x_2)^T\Phi^Ta\,l_2,\quad l_2 = \left(-\lambda + \frac{t_2}{a_2}\right)^{-1}$$

$$a_3 = \phi(x_3)^T\Phi^Ta\,l_3,\quad l_3 = \left(-\lambda + \frac{t_3}{a_3}\right)^{-1}$$

N が4以上でも同様に

$$a_i = \phi(x_i)^T\Phi^Ta\,l_i,\quad l_i = \left(-\lambda + \frac{t_i}{a_i}\right)^{-1} \quad \cdots ①$$

(双対表現から元の最小二乗法の式に戻す)

①より
$$\begin{pmatrix}a_1\\a_2\\a_3\end{pmatrix} = \begin{pmatrix}\phi(x_1)^T\Phi^Ta\,l_1\\ \phi(x_2)^T\Phi^Ta\,l_2\\ \phi(x_3)^T\Phi^Ta\,l_3\end{pmatrix} = \begin{pmatrix}l_1 & 0 & 0\\ 0 & l_2 & 0\\ 0 & 0 & l_3\end{pmatrix}\begin{pmatrix}\phi(x_1)^T\Phi^Ta\\ \phi(x_2)^T\Phi^Ta\\ \phi(x_3)^T\Phi^Ta\end{pmatrix}$$

よって $a = L\Phi\Phi^Ta,\quad L = \begin{pmatrix}l_1 & 0 & 0\\ 0 & l_2 & 0\\ 0 & 0 & l_3\end{pmatrix}$ … ②

②より $L\Phi\Phi^T = I_N$ … ③

③より $LK = I_N$

③より $\Phi^TL\Phi\Phi^T = \Phi^T \quad \therefore\ \Phi^TL\Phi = I_W$

② を $a = L\Phi w$, $w = \Phi^T a$ とおく
(6.1), (6.8), (6.9) から a を消去して w のみで表わす。その結果が元の最小二乗法の式に戻ってることを確認する。

(6.7) に入れると　← L は対角なので $L^T = L$

$$
\begin{aligned}
J(a) &= \frac{1}{2} w^T \Phi^T L K K L \Phi w - w^T \Phi^T L K t + \frac{1}{2} t^T t + \frac{\lambda}{2} w^T \Phi^T L K L \Phi w \\
&= \frac{1}{2} w^T \Phi^T \Phi w - w^T \Phi^T t + \frac{1}{2} t^T t + \frac{\lambda}{2} w^T \Phi^T L \Phi w \\
&= \frac{1}{2} (\Phi w - t)^T (\Phi w - t) + \frac{\lambda}{2} w^T w
\end{aligned}
$$

となり (6.2) を得る

(6.9) に入れると

$$
y(x) = a^T \Phi \phi(x) = w^T \Phi^T L \Phi \phi(x) = w^T \phi(x)
$$

となり元の最小二乗法の予測値を得る

(6.8) より

$$
\Phi \Phi^T a + \lambda a = t
$$

両辺に Φ^T をかけて

$$
\Phi^T \Phi \Phi^T a + \lambda \Phi^T a = \Phi^T t
$$

$w = \Phi^T a$ を入れると

$$
\Phi^T \Phi w + \lambda w = \Phi^T t
$$

$$
\begin{aligned}
\therefore\ w &= -\frac{1}{\lambda} \Phi^T (\Phi w - t) \\
&= -\frac{1}{\lambda} \begin{pmatrix} \phi(x_1) & \phi(x_2) & \phi(x_3) \end{pmatrix} \begin{pmatrix} \phi(x_1)^T w - t_1 \\ \phi(x_2)^T w - t_2 \\ \phi(x_3)^T w - t_3 \end{pmatrix}
\end{aligned}
$$

← $\phi(x_1)^T w = \underset{\text{スカラーなので}}{(\phi(x_1)^T w)^T} = w^T \phi(x_1)$

$$
\begin{aligned}
&= -\frac{1}{\lambda} \begin{pmatrix} \phi(x_1) & \phi(x_2) & \phi(x_3) \end{pmatrix} \begin{pmatrix} w^T \phi(x_1) - t_1 \\ w^T \phi(x_2) - t_2 \\ w^T \phi(x_3) - t_3 \end{pmatrix} \\
&= -\frac{1}{\lambda} \sum_{n=1}^{N} \{ w^T \phi(x_n) - t_n \} \phi(x_n)
\end{aligned}
$$

となり (6.3) を得る。