

概 要

講義資料の解答とかをここに書くね。中間試験・期末試験の問題はここからしか出ないよ。

Lecture 5 の回答

線形回帰モデルへの変形

$$a = -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_1)^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}_1) + \frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_2)^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}_2) + \ln \frac{p(C_1)}{p(C_2)}$$

の時、

$$\mathbf{w} = \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)$$

$$w_0 = -\frac{1}{2}\boldsymbol{\mu}_1^T \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_1 + \frac{1}{2}\boldsymbol{\mu}_2^T \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_2 + \ln \frac{p(C_1)}{p(C_2)}$$

とおくと、

$$a = \mathbf{w}^T \mathbf{x} + w_0$$

と表記できることを示す。

$$\begin{aligned} a &= -\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_1)^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}_1) + \frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_2)^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}_2) + \ln \frac{p(C_1)}{p(C_2)} \\ &= -\frac{1}{2}(\mathbf{x}^T \boldsymbol{\Sigma}^{-1} \mathbf{x} - \mathbf{x}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_1 - \boldsymbol{\mu}_1^T \boldsymbol{\Sigma}^{-1} \mathbf{x} + \boldsymbol{\mu}_1^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_1) + \\ &\quad -\frac{1}{2}(\mathbf{x}^T \boldsymbol{\Sigma}^{-1} \mathbf{x} - \mathbf{x}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_2 - \boldsymbol{\mu}_2^T \boldsymbol{\Sigma}^{-1} \mathbf{x} + \boldsymbol{\mu}_2^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_2) + \ln \frac{p(C_1)}{p(C_2)} \end{aligned}$$

ここで、 $\mathbf{x}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_1$ はスカラなので転置をかけても等しくなるので、 $(\mathbf{x}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_1)^T = \boldsymbol{\mu}_1^T \boldsymbol{\Sigma}^{-1} \mathbf{x} = \boldsymbol{\mu}_1^T \boldsymbol{\Sigma}^{-1} \mathbf{x}$ となる。ただし、 $(ABC)^T = C^T B^T A^T$ と、 $\boldsymbol{\Sigma}$ は対称行列で、対称行列の逆行列もまた対称行列となることから、 $\boldsymbol{\Sigma}^{-1T} = \boldsymbol{\Sigma}^{-1}$ を使った。これを使い、

$$\begin{aligned} a &= -\frac{1}{2}\mathbf{x}^T \boldsymbol{\Sigma}^{-1} \mathbf{x} + \boldsymbol{\mu}_1^T \boldsymbol{\Sigma}^{-1} \mathbf{x} - \frac{1}{2}\boldsymbol{\mu}_1^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_1 \\ &\quad + \frac{1}{2}\mathbf{x}^T \boldsymbol{\Sigma}^{-1} \mathbf{x} - \boldsymbol{\mu}_2^T \boldsymbol{\Sigma}^{-1} \mathbf{x} + \frac{1}{2}\boldsymbol{\mu}_2^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_2 + \ln \frac{p(C_1)}{p(C_2)} \\ &= \boldsymbol{\mu}_1^T \boldsymbol{\Sigma}^{-1} \mathbf{x} - \frac{1}{2}\boldsymbol{\mu}_1^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2^T \boldsymbol{\Sigma}^{-1} \mathbf{x} + \frac{1}{2}\boldsymbol{\mu}_2^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_2 + \ln \frac{p(C_1)}{p(C_2)} \end{aligned}$$

ここで、 $\mathbf{w}^T \mathbf{x} + w_0$ を計算すると、

$$\begin{aligned} \mathbf{w}^T \mathbf{x} + w_0 &= (\boldsymbol{\mu}_1^T - \boldsymbol{\mu}_2^T) \boldsymbol{\Sigma}^{-1} \mathbf{x} - \frac{1}{2}\boldsymbol{\mu}_1^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_1 + \frac{1}{2}\boldsymbol{\mu}_2^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_2 + \ln \frac{p(C_1)}{p(C_2)} \\ &= \boldsymbol{\mu}_1^T \boldsymbol{\Sigma}^{-1} \mathbf{x} - \boldsymbol{\mu}_2^T \boldsymbol{\Sigma}^{-1} \mathbf{x} - \frac{1}{2}\boldsymbol{\mu}_1^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_1 + \frac{1}{2}\boldsymbol{\mu}_2^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_2 + \ln \frac{p(C_1)}{p(C_2)} \end{aligned}$$

となるので、 $a = \mathbf{w}^T \mathbf{x} + w_0$ が示せた。

シグモイド関数の微分

$\sigma(x) = \frac{1}{1+\exp(-x)}$ の時、 $\frac{d\sigma(x)}{dx} = \sigma(x)(1 - \sigma(x))$ を示そう。

$$\begin{aligned} \frac{d\sigma(x)}{dx} &= \frac{d(1 + e^{-x})^{-1}}{dx} \\ &= -(1 + e^{-x})^{-2}(-1)e^{-x} \quad (\because \frac{de^{-x}}{dx} = -e^{-x}) \\ &= (1 + e^{-x})^{-2}e^{-x} \end{aligned}$$

ここで、右辺を計算すると、

$$\begin{aligned}\sigma(x)(1 - \sigma(x)) &= (1 + e^{-x})^{-1} \frac{1 + e^{-x} - 1}{1 + e^{-x}} \\ &= (1 + e^{-x})^{-2} e^{-x}\end{aligned}$$

よって、 $\frac{d\sigma(x)}{dx} = \sigma(x)(1 - \sigma(x))$ を示した。

交差エントロピー誤差関数の微分

$\frac{\partial E(\mathbf{w})}{\partial \mathbf{w}} = \sum_{n=1}^N (t_n - y_n) \phi_n$ を示そう。

$$\begin{aligned}\frac{\partial E(\mathbf{w})}{\partial \mathbf{w}} &= \sum_{n=1}^N \frac{\partial}{\partial \mathbf{w}} (t_n \ln y_n + (1 - t_n) \ln(1 - y_n)) \\ &= \sum_{n=1}^N (t_n \frac{\partial}{\partial \mathbf{w}} \ln y_n + (1 - t_n) \frac{\partial}{\partial \mathbf{w}} \ln(1 - y_n))\end{aligned}$$

まずは $\frac{\partial}{\partial \mathbf{w}} \ln y_n$ を計算しよう。

$$\begin{aligned}\frac{\partial}{\partial \mathbf{w}} \ln y_n &= \frac{1}{y_n} \frac{\partial}{\partial \mathbf{w}} y_n \quad (\because \frac{d \ln x}{dx} = \frac{1}{x}) \\ &= \frac{1}{\sigma(\mathbf{w}^T \phi_n)} \frac{\partial}{\partial \mathbf{w}} \sigma(\mathbf{w}^T \phi_n) \\ &= \frac{1}{\sigma(\mathbf{w}^T \phi_n)} \sigma(\mathbf{w}^T \phi_n) (1 - \sigma(\mathbf{w}^T \phi_n)) \phi_n \quad (\because \frac{d\sigma(x)}{dx} = \sigma(x)(1 - \sigma(x))) \\ &= (1 - y_n) \phi_n\end{aligned}$$

次に、 $\frac{\partial}{\partial \mathbf{w}} \ln(1 - y_n)$ を計算しよう。

$$\begin{aligned}\frac{\partial}{\partial \mathbf{w}} \ln(1 - y_n) &= \frac{1}{1 - y_n} \frac{\partial}{\partial \mathbf{w}} (1 - y_n) \\ &= \frac{1}{1 - y_n} \cdot -\frac{\partial}{\partial \mathbf{w}} y_n \\ &= \frac{1}{1 - y_n} \cdot -y_n(1 - y_n) \phi_n \\ &= -y_n \phi_n\end{aligned}$$

これにより、

$$\begin{aligned}\frac{\partial E(\mathbf{w})}{\partial \mathbf{w}} &= \sum_{n=1}^N (t_n(1 - t_n) \phi_n + (1 - t_n)(-y_n \phi_n)) \\ &= \sum_{n=1}^N (t_n - y_n) \phi_n\end{aligned}$$

と示せた。

意地悪な問題

$\frac{\partial E(\mathbf{w})}{\partial \mathbf{w}} = \sum_{n=1}^N (t_n - y_n) \phi_n = 0$ となるような \mathbf{w} を解析的に導出してみよう。

$$\begin{aligned}\sum_{n=1}^N t_n \phi_n &= \sum_{n=1}^N y_n \phi_n \\ \sum_{n=1}^N t_n \phi_n &= \sum_{n=1}^N \sigma(\mathbf{w}^T \phi_n) \phi_n\end{aligned}$$

さて、線形回帰の時には、

$$\sum_{n=1}^N t_n \phi_n^\top = \mathbf{w}^\top \sum_{n=1}^N \phi_n \phi_n^\top \quad (1)$$

となり、 $\Phi = \begin{pmatrix} \phi(x_1) \\ \phi(x_2) \\ \dots \\ \phi(x_N) \end{pmatrix}$ とおくことによって、 $\Phi^\top \mathbf{t} = \mathbf{w}^\top \Phi^\top \Phi$ と変形できた。しかし、

今回は非線形の関数 σ があるために \mathbf{w} を Σ の外側に移動できない！(非線形関数は行列では表現できないことに注意しよう)