データ科学特論　課題 1

2120041 濱口和希

2020 年 5 月 21 日

1

1.1　a

$$
\begin{aligned}
\frac{\partial L(\boldsymbol{\theta}, \lambda)}{\partial \boldsymbol{\theta}} &= \sum_{i=1}^{2} \frac{\partial(-f(x_i;\theta))}{\partial \boldsymbol{\theta}}(y_i - (x_i, 1)\boldsymbol{\theta}) + \frac{\lambda}{2}\frac{\partial}{\partial \boldsymbol{\theta}}\boldsymbol{\theta}^T\boldsymbol{\theta} \\
&= -\sum_{i=1}^{2}(x_i, 1)^T(y_i - (x_i, 1)\boldsymbol{\theta}) + \lambda\boldsymbol{\theta} \\
&= -\sum_{i=1}^{2}(x_i y_i, y_i)^T + (\sum_{i=1}^{2}(x_i, 1)^T(x_i, 1) + \lambda)\boldsymbol{\theta} \\
&= \begin{pmatrix} 1+ab \\ 1+b \end{pmatrix} + \left(\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} a^2 & a \\ a & 1 \end{pmatrix} + \begin{pmatrix} \lambda & \lambda \\ \lambda & \lambda \end{pmatrix}\right)\boldsymbol{\theta} = 0
\end{aligned} \tag{1}
$$

したがって

$$
\begin{pmatrix} 1+a^2+\lambda & 1+a+\lambda \\ 1+a+\lambda & 1+\lambda \end{pmatrix}\boldsymbol{\theta} = \begin{pmatrix} 1+ab \\ 1+b \end{pmatrix}, \boldsymbol{A} = \begin{pmatrix} 1+a^2+\lambda & 1+a+\lambda \\ 1+a+\lambda & 1+\lambda \end{pmatrix}, \boldsymbol{B} = \begin{pmatrix} 1+ab \\ 1+b \end{pmatrix} \tag{2}
$$

1.2　b

$$
\begin{aligned}
\begin{vmatrix} 1+a^2+\lambda & 1+a+\lambda \\ 1+a+\lambda & 1+\lambda \end{vmatrix} &= (1+a^2+\lambda)(2+\lambda) - (1+a+\lambda)^2 \\
&= a^2(\lambda+1) - 2a(\lambda+1) + (\lambda+1) = (\lambda+1)(a-1)^2 \neq 0
\end{aligned} \tag{3}
$$

したがって $a \neq 1$ の時解が存在する

1.3 c

$argmin\boldsymbol{\theta}_\lambda = \begin{pmatrix} \theta_1 = 0 \\ \theta_2 = 0 \end{pmatrix}$ となるような λ を解けばよい

$$\lambda = \begin{pmatrix} \frac{1}{b} - 1 \\ \frac{a}{b} - 1 \end{pmatrix} \tag{4}$$

また

$$\lim_{\lambda \to \infty} \boldsymbol{\theta}_\lambda = \frac{b}{a-1} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{5}$$

2

$L(\boldsymbol{\theta}, \lambda)$ を

$$L(\boldsymbol{\theta}, \lambda) = \frac{1}{2}(\boldsymbol{Y} - \boldsymbol{X\theta})^T(\boldsymbol{Y} - \boldsymbol{X\theta}) + \frac{1}{2}\lambda\boldsymbol{\theta}^T\boldsymbol{\theta} \tag{6}$$

$$(\text{ただし} \sum_{i=1}^{n} y_i = \boldsymbol{Y}, \sum_{i=1}^{n} f(x_i; \boldsymbol{\theta}) = \boldsymbol{X\theta})$$

と書き換える．正規方程式で表したいため

$$\frac{\partial}{\partial\boldsymbol{\theta}} L(\boldsymbol{\theta}, \lambda) = 0 \text{ とすると} \tag{7}$$

$$\begin{aligned}
\frac{\partial}{\partial\boldsymbol{\theta}} L(\boldsymbol{\theta}, \lambda) = -\boldsymbol{X}^T(\boldsymbol{Y} - \boldsymbol{X\theta}) + \lambda\boldsymbol{\theta} = 0 \\
\boldsymbol{X}^T\boldsymbol{X\theta} + \lambda\boldsymbol{\theta} = \boldsymbol{X}^T\boldsymbol{Y} \\
(\boldsymbol{X}^T\boldsymbol{X} + \lambda\boldsymbol{I})\boldsymbol{\theta} = \boldsymbol{X}^T\boldsymbol{Y} \\
\boldsymbol{\theta}_\lambda = (\boldsymbol{X}^T\boldsymbol{X} + \lambda\boldsymbol{I})^{-1}\boldsymbol{X}^T\boldsymbol{Y}
\end{aligned}$$

したがってパラメータ $\boldsymbol{\theta}$ は $(\boldsymbol{X}^T\boldsymbol{X} + \lambda\boldsymbol{I})^{-1}\boldsymbol{X}^T\boldsymbol{Y}$ を満たす必要がある．