『多変量解析入門』　演習問題　解答　第3章

Taro Masuda
Twitter ID: @ml_taro

2022 年 2 月 12 日

はじめに

このPDFでは，小西貞則著『多変量解析入門　線形から非線形へ』(岩波書店，2010)の解答を記していきます．公式なものではなくあくまで個人として公開しているため，誤りがある可能性があります．正確性についての保証はできない旨，予めご了承ください．

なお，著作権へ配慮し，問題文については割愛させていただきます．

誤りを見つけた場合は，上記twitter ID @ml_taro までご連絡いただくか，直接PRを飛ばして頂くか，メール taro.masuda.jp あっとまーく gmail.com までご連絡ください．

第3章

問3.3

誤差の2乗和$S(\boldsymbol{w}) = \boldsymbol{\varepsilon}^\top\boldsymbol{\varepsilon}$を最小化したいので，

$$S(\boldsymbol{w}) = \boldsymbol{\varepsilon}^\top\boldsymbol{\varepsilon} = (\boldsymbol{y} - B\boldsymbol{w})^\top(\boldsymbol{y} - B\boldsymbol{w}) \tag{1}$$

を$\boldsymbol{\beta}$について微分して

$$-2B^\top\boldsymbol{y} + 2B^\top B\boldsymbol{w} = \mathbf{0}. \tag{2}$$

これを解いて，　$\hat{\boldsymbol{w}}_{\mathrm{LMS}} = (B^\top B)^{-1}B^\top\boldsymbol{y}$　を得る．

問3.4

式(3.35)の尤度関数のうち，$\boldsymbol{w}$に依存する部分は

$$\exp\left\{-\frac{1}{2\sigma^2}(\boldsymbol{y} - B\boldsymbol{w})^\top(\boldsymbol{y} - B\boldsymbol{w})\right\} \tag{3}$$

のみであるから，$(\boldsymbol{y} - B\boldsymbol{w})^\top(\boldsymbol{y} - B\boldsymbol{w})$を最小にする$\boldsymbol{w}$が最尤推定量となる．つまり，最小2乗推定量（問3.3の答え）$\hat{\boldsymbol{w}} = (B^\top B)^{-1}B^\top\boldsymbol{y}$と一致する．

問3.5

$$\frac{\partial l_\lambda(\boldsymbol{\theta})}{\partial \boldsymbol{w}} = -\frac{1}{2\sigma^2}(-2B^\top \boldsymbol{y} + 2B^\top B\boldsymbol{w}) - \lambda K\boldsymbol{w}. \tag{4}$$

これを$\mathbf{0}$とおくと

$$\left(\frac{1}{\sigma^2}B^\top B + \lambda K\right)\boldsymbol{w} = \frac{1}{\sigma^2}B^\top \boldsymbol{y} \iff \hat{\boldsymbol{w}} = \left(B^\top B + \lambda\sigma^2 K\right)^{-1} B^\top \boldsymbol{y}. \tag{5}$$

また，

$$\frac{\partial l_\lambda(\boldsymbol{\theta})}{\partial(\sigma^2)} = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}(\boldsymbol{y} - B\boldsymbol{w})^\top(\boldsymbol{y} - B\boldsymbol{w}). \tag{6}$$

これを0とおくと

$$\hat{\sigma}^2 = \frac{1}{n}(\boldsymbol{y} - B\hat{\boldsymbol{w}})^\top(\boldsymbol{y} - B\hat{\boldsymbol{w}}). \tag{7}$$

問3.6

$\boldsymbol{w}$の正則化最小2乗推定量は，ペナルティ項を含めた

$$S_\gamma(\boldsymbol{w}) = (\boldsymbol{y} - B\boldsymbol{w})^\top(\boldsymbol{y} - B\boldsymbol{w}) + \gamma\boldsymbol{w}^\top K\boldsymbol{w} \tag{8}$$

と表されるので，これを$\boldsymbol{w}$について微分して

$$\frac{\partial S_\gamma(\boldsymbol{w})}{\partial \boldsymbol{w}} = -2B^\top \boldsymbol{y} + 2B^\top B\boldsymbol{w} + 2\gamma K\boldsymbol{w} = \mathbf{0} \iff \hat{\boldsymbol{w}} = (B^\top B + \gamma K)^{-1}B^\top \boldsymbol{y}. \tag{9}$$

ここで$\gamma = \lambda\sigma^2$とおけば，正則化最尤推定量（問3.5の解）に一致する.

問3.8

$$z_{ij} = x_{ij} - \bar{x}_{ij}, \tag{10}$$

$$\bar{x}_j = \frac{1}{n}\sum_{i=1}^{n} x_{ij}, \tag{11}$$

$$\beta_0^* = \beta_0 + \beta_1\bar{x}_1 + \cdots + \beta_p\bar{x}_p \tag{12}$$

とおくと，

$$y_i = \beta_0 + \beta_1\bar{x}_1 + \cdots + \beta_p\bar{x}_p + \beta_1(x_{i1} - \bar{x}_1) + \cdots + \beta_p(x_{ip} - \bar{x}_p) + \varepsilon_i = \beta_0^* + \beta_1 z_{i1} + \cdots + \beta_p z_{ip} + \varepsilon_i \tag{13}$$

$$\therefore \boldsymbol{y} = \beta_0^*\mathbf{1} + Z\boldsymbol{\beta}_1 + \boldsymbol{\varepsilon}. \tag{14}$$

また，この2乗誤差と，β_1のノルム正則化誤差を最小化すれば良いので，目的関数は

$$S_\gamma(\beta_0^*, \boldsymbol{\beta}_1) = (\boldsymbol{y} - \beta_0^*\mathbf{1} - Z\boldsymbol{\beta}_1)^\top(\boldsymbol{y} - \beta_0^*\mathbf{1} - Z\boldsymbol{\beta}_1) + \gamma\boldsymbol{\beta}_1^\top\boldsymbol{\beta}_1 \tag{15}$$

$$= \boldsymbol{y}^\top\boldsymbol{y} + \beta_0^{*2}\mathbf{1}^\top\mathbf{1} + \boldsymbol{\beta}_1^\top Z^\top Z\boldsymbol{\beta}_1 - 2\beta_0^*(\boldsymbol{y} - Z\boldsymbol{\beta}_1)^\top\mathbf{1} - 2\boldsymbol{y}^\top Z\boldsymbol{\beta}_1 + \gamma\boldsymbol{\beta}_1^\top\boldsymbol{\beta}_1. \tag{16}$$

$\mathbf{1}^\top\mathbf{1} = n$に注意すると，

$$\frac{\partial S_\gamma}{\partial \beta_0^*} = 2n\beta_0^* - 2\left(\sum_{i=1}^n y_i - (Z\boldsymbol{\beta}_1)^\top \mathbf{1}\right). \tag{17}$$

これを0とおくと，$Z^\top\mathbf{1} = \mathbf{0}$ より

$$\hat{\beta}_0^* = \frac{1}{n}\left(\sum_{i=1}^n y_i - \boldsymbol{\beta}_1^\top Z^\top \mathbf{1}\right) = \bar{y}. \tag{18}$$

また$\boldsymbol{\beta}_1$について微分して$\mathbf{0}$とおくと

$$\frac{\partial S_\gamma}{\partial \boldsymbol{\beta}_1} = 2Z^\top Z\boldsymbol{\beta}_1 + 2\beta_0^* Z^\top \mathbf{1} - 2Z^\top \boldsymbol{y} + 2\gamma\boldsymbol{\beta}_1 = \mathbf{0}. \tag{19}$$

$$\therefore \hat{\boldsymbol{\beta}}_1 = (Z^\top Z + \gamma I_p)^{-1} Z^\top \boldsymbol{y}. \tag{20}$$