『多変量解析入門』　演習問題　解答　第2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2章

問2.1

証明．

$$\frac{\partial S(\beta_0,\beta_1)}{\partial \beta_0} = \sum_{i=1}^{n} 2\{y_i - (\beta_0 + \beta_1 x_i)\}(-1). \tag{1}$$

これを0とおくと

$$\sum_{i=1}^{n} y_i = n\beta_0 + \beta_1 \sum_{i=1}^{n} x_i. \tag{2}$$

同様にβ_1 でも微分して0とおくと

$$\frac{\partial S(\beta_0,\beta_1)}{\partial \beta_1} = \sum_{i=1}^{n} 2\{y_i - (\beta_0 + \beta_1 x_i)\}(-x_i) = 0 \iff \sum_{i=1}^{n} x_i y_i = \beta_0 \sum_{i=1}^{n} x_i + \beta_1 \sum_{i=1}^{n} x_i^2. \tag{3}$$

□

問2.2

誤差の2乗和を最小化したいので，

$$S(\boldsymbol{\beta}) = \boldsymbol{\varepsilon}^{\mathsf{T}}\boldsymbol{\varepsilon} = (\boldsymbol{y} - X\boldsymbol{\beta})^{\mathsf{T}}(\boldsymbol{y} - X\boldsymbol{\beta}) \tag{4}$$

を$\boldsymbol{\beta}$について微分して

$$-2X^{\mathsf{T}}\boldsymbol{y}+2X^{\mathsf{T}}X\boldsymbol{\beta}=\mathbf{0}. \tag{5}$$

これを解いて，　$\hat{\boldsymbol{\beta}}_{\mathrm{LMS}}=(X^{\mathsf{T}}X)^{-1}X^{\mathsf{T}}\boldsymbol{y}$　を得る.

問2.3

尤度関数は

$$L(\boldsymbol{\beta},\sigma^2)=\frac{1}{(2\pi\sigma^2)^{n/2}}\exp\left[-\frac{1}{2\sigma^2}(\boldsymbol{y}-X\boldsymbol{\beta})^{\mathsf{T}}(\boldsymbol{y}-X\boldsymbol{\beta})\right] \tag{6}$$

となり，これの対数をとると

$$\log L(\boldsymbol{\beta},\sigma^2)=-\frac{n}{2}\log(2\pi\sigma^2)-\frac{1}{2\sigma^2}(\boldsymbol{y}-X\boldsymbol{\beta})^{\mathsf{T}}(\boldsymbol{y}-X\boldsymbol{\beta}). \tag{7}$$

これを$\boldsymbol{\beta}$について微分すると

$$-2X^{\mathsf{T}}\boldsymbol{y}+2X^{\mathsf{T}}X\boldsymbol{\beta}=\mathbf{0} \iff \hat{\boldsymbol{\beta}}_{\mathrm{ML}}=(X^{\mathsf{T}}X)^{-1}X^{\mathsf{T}}\boldsymbol{y}. \tag{8}$$

同様に，σ^2について微分すると

$$-\frac{n}{2}\frac{1}{2\pi\sigma^2}2\pi+\frac{1}{2(\sigma^2)^2}(\boldsymbol{y}-X\boldsymbol{\beta})^{\mathsf{T}}(\boldsymbol{y}-X\boldsymbol{\beta}) \iff \hat{\sigma}^2=\frac{1}{n}(\boldsymbol{y}-X\boldsymbol{\beta})^{\mathsf{T}}(\boldsymbol{y}-X\boldsymbol{\beta}). \tag{9}$$

問2.4

問2.2, 問2.3より $\hat{\boldsymbol{\beta}}_{\mathrm{LMS}}=\hat{\boldsymbol{\beta}}_{\mathrm{ML}}$.

問2.5

証明.

$$\frac{\partial(\boldsymbol{c}^{\mathsf{T}}\boldsymbol{\beta})}{\partial\boldsymbol{\beta}}=\frac{\partial}{\partial\boldsymbol{\beta}}\left(\sum_{i=1}^{p+1}c_i\beta_i\right). \tag{10}$$

これをj番目の成分ごとに表示すると

$$\frac{\partial}{\partial\beta_j}\left(\sum_{i=1}^{p+1}c_i\beta_i\right)=c_j. \tag{11}$$

$$\therefore \frac{\partial(\boldsymbol{c}^{\mathsf{T}}\boldsymbol{\beta})}{\partial\boldsymbol{\beta}}=\boldsymbol{c}. \tag{12}$$

□

証明.

$$\boldsymbol{\beta}^{\mathsf{T}}A\boldsymbol{\beta}=\sum_{i=1}^{p+1}\beta_i\sum_{k=1}^{p+1}a_{ik}\beta_k \tag{13}$$

と表せるので，

$$\frac{\partial}{\partial\beta_l}\left(\boldsymbol{\beta}^{\mathsf{T}}A\boldsymbol{\beta}\right)=\sum_{k\neq l}a_{lk}\beta_k+\sum_{i\neq l}a_{il}\beta_i+2a_{ll}\beta_l \tag{14}$$

$$= \sum_{k=1}^{p+1} (a_{lk} + a_{kl}) \beta_k \tag{15}$$

これは$(A + A^\top)\boldsymbol{\beta}$の第$l$成分に等しい. □

問2.6

証明.

$$E[\boldsymbol{Z}] = E[\boldsymbol{c} + A\boldsymbol{Y}] = E[\boldsymbol{c}] + E[A\boldsymbol{Y}] = \boldsymbol{c} + AE[\boldsymbol{Y}]. \tag{16}$$

□

証明.

$$E[(\boldsymbol{Z} - E[\boldsymbol{Z}])(\boldsymbol{Z} - E[\boldsymbol{Z}])^\top]] = E[(\boldsymbol{c} + A\boldsymbol{Y})(\boldsymbol{c} + A\boldsymbol{Y})^\top] - (\boldsymbol{c} + AE[\boldsymbol{Y}])(\boldsymbol{c} + AE[\boldsymbol{Y}])^\top \tag{17}$$

$$= \boldsymbol{c}\boldsymbol{c}^\top + 2\boldsymbol{c}E[\boldsymbol{Y}^\top]A^\top + AE[\boldsymbol{Y}\boldsymbol{Y}^\top]A^\top - \boldsymbol{c}\boldsymbol{c}^\top - 2\boldsymbol{c}E[\boldsymbol{Y}^\top]A^\top - AE[\boldsymbol{Y}]E[\boldsymbol{Y}^\top]A^\top \tag{18}$$

$$= A(E[\boldsymbol{Y}\boldsymbol{Y}^\top] - E[\boldsymbol{Y}]E[\boldsymbol{Y}^\top])A^\top = A\mathrm{cov}(\boldsymbol{Y})A^\top. \tag{19}$$

□

問2.7

証明.

$$P^2 = X(X^\top X)^{-1}X^\top X(X^\top X)^{-1}X^\top \tag{20}$$

$$= X(X^\top X)^{-1}X^\top = P. \tag{21}$$

□

証明.

$$(I_n - P)^2 = (I_n - P)(I_n - P) \tag{22}$$

$$= I_n - P - P + P^2 = I_n - 2P + P = I_n - P. \tag{23}$$

□

問2.8

証明. AICの定義は -2 (最大対数尤度) + 2 (自由パラメタの数)

$$= -2\left(-\frac{n}{2}\log(2\pi\hat{\sigma}^2) - \frac{1}{2\hat{\sigma}^2}(\boldsymbol{y} - X\boldsymbol{\beta})^\top(\boldsymbol{y} - X\boldsymbol{\beta})\right) \tag{24}$$

問2.3よりσの最尤推定量は$\hat{\sigma}^2 = \frac{1}{n}(\boldsymbol{y} - X\boldsymbol{\beta})^\top(\boldsymbol{y} - X\boldsymbol{\beta})$であることと，自由なパラメタの数は$\boldsymbol{\beta}$の次元$p+1$個と誤差分散$\hat{\sigma}^2$の合計$p+2$個なので

$$\mathrm{AIC} = n\log(2\pi\hat{\sigma}^2) + n + 2(p+2). \tag{25}$$

□