

『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 8 日

1 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 あつとまーく gmail.com までご連絡ください。

2 第2章

2.1 問2.1

証明.

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

これを0とおくと

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

同様に β_1 でも微分して0とおくと

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

□

2.2 問2.2

誤差の2乗和を最小化したいので、

$$S(\beta) = \epsilon^T \epsilon = (\mathbf{y} - X\beta)^T (\mathbf{y} - X\beta) \quad (4)$$

を β について微分して

$$-2X^T \mathbf{y} + 2X^T X \beta = \mathbf{0}. \quad (5)$$

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

2.3 問2.3

尤度関数は

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

となり, この対数をとると

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

これを β について微分すると

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

同様に, σ^2 について微分すると

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

2.4 問2.4

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

2.5 問2.5

証明.

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

これを j 番目の成分ごとに表示すると

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

$$\therefore \frac{\partial(\mathbf{c}^T \beta)}{\partial \beta} = \mathbf{c}. \quad (12)$$

□

証明.

$$\beta^T A \beta = \sum_{i=1}^{p+1} \beta_i \sum_{k=1}^{p+1} a_{ik} \beta_k \quad (13)$$

と表せるので,

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

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

これは $(A + A^T)\beta$ の第 l 成分に等しい. □

2.6 問2.6

証明.

$$E[\mathbf{Z}] = E[\mathbf{c} + A\mathbf{Y}] = E[\mathbf{c}] + E[A\mathbf{Y}] = \mathbf{c} + AE[\mathbf{Y}]. \quad (16)$$

□

証明.

$$E[(\mathbf{Z} - E[\mathbf{Z}])(\mathbf{Z} - E[\mathbf{Z}])^T] = E[(\mathbf{c} + A\mathbf{Y})(\mathbf{c} + A\mathbf{Y})^T] - (\mathbf{c} + AE[\mathbf{Y}]) (\mathbf{c} + AE[\mathbf{Y}])^T \quad (17)$$

$$= \mathbf{c}\mathbf{c}^T + 2\mathbf{c}E[\mathbf{Y}^T]A^T + AE[\mathbf{Y}\mathbf{Y}^T]A^T - \mathbf{c}\mathbf{c}^T - 2\mathbf{c}E[\mathbf{Y}^T]A^T - AE[\mathbf{Y}]E[\mathbf{Y}^T]A^T \quad (18)$$

$$= A(E[\mathbf{Y}\mathbf{Y}^T] - E[\mathbf{Y}]E[\mathbf{Y}^T])A^T = A\text{cov}(\mathbf{Y})A^T. \quad (19)$$

□

2.7 問2.7

証明.

$$P^2 = X(X^T X)^{-1} X^T X (X^T X)^{-1} X^T \quad (20)$$

$$= X(X^T X)^{-1} X^T = P. \quad (21)$$

□

証明.

$$(I_n - P)^2 = (I_n - P)(I_n - P) \quad (22)$$

$$= I_n - P - P + P^2 = I_n - 2P + P = I_n - P. \quad (23)$$

□

2.8 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)

$$= -2 \left(-\frac{n}{2} \log(2\pi\hat{\sigma}^2) - \frac{1}{2\hat{\sigma}^2} (\mathbf{y} - X\beta)^T (\mathbf{y} - X\beta) \right) \quad (24)$$

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

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

□