

統計的機械学習の数理100問 with Python

鈴木 讓

2021年2月11日現在

必ずしも誤植でなく、そのように変更した方が良い情報も含まれています。プログラムの修正は、<https://bitbucket.org/prof-joe>でも反映されています。

1章

1. 頁15下9: 接点 → 切片
2. 頁21下1: $X(X^T X)X^T y \rightarrow X(X^T X)^{-1}X^T y$
3. 頁21上7: $E(\hat{\beta}_i - \beta_i)(\hat{\beta}_j - \beta_j)^T \rightarrow E(\hat{\beta}_i - \beta_i)(\hat{\beta}_j - \beta_j)$
4. 頁23図1.3: プログラムの出力が正しい
5. 頁25下4: $\beta_0 = 2 \rightarrow \beta_0 = 1$
6. 頁26上2行目、4行目、6行目: $\beta_i \rightarrow \beta_j$ (4箇所)、 $\beta \rightarrow \beta_j$ (2箇所)
7. 頁26上7: (1.17) → (1.18)
8. 頁26下1-2: 関数 $\text{pt}(x, m) \rightarrow \text{stats.t.cdf}(x, m)$ 、 $\int_x^\infty f_m(t)dt \rightarrow \int_{-\infty}^x f_m(t)dt$
9. 頁29プログラム15行目: $\text{plt.plot}(x, \text{stats.t.pdf}(x, 1)) \rightarrow \text{plt.plot}(x, \text{stats.t.pdf}(x, 98))$
10. 頁31上11: $y_* = x_*\beta + \epsilon$ 差の分布 → $y_* = x_*\beta + \epsilon$ の差の分布
11. 頁33下8: $\sqrt{\frac{RSS}{\sigma}} / (N - p - 1) \rightarrow \sqrt{\frac{RSS}{\sigma^2}} / (N - p - 1)$
12. 頁36上9: $H = X(XX^T)^{-1}X$ かつ → $H = X(XX^T)^{-1}X^T$ かつ
13. 頁37問題3: ふたつ目のプログラム2行目 $\text{size}=N$ を削除
14. 頁39問題10(b): $Evv^T = \sigma^2 I \rightarrow Evv^T = \sigma^2 \tilde{I}$
15. 頁44上2-4: y (2箇所) → y_*
16. 頁44上11-12: ペアを赤で、後者の区間のペアを青で → ペアの上側を赤で、下側を青で

第2章

1. 頁 51: $\nabla L \rightarrow \nabla l(\beta_0, \beta)$ (2箇所)、 $\nabla^2 L \rightarrow \nabla^2 l(\beta_0, \beta)$ (2箇所)、 $\frac{\partial L}{\partial \beta_j} \rightarrow \frac{\partial l(\beta_0, \beta)}{\partial \beta_j}$ 、 $\frac{\partial^2 L}{\partial \beta_j \beta_k} \rightarrow \frac{\partial^2 l(\beta_0, \beta)}{\partial \beta_j \beta_k}$
2. 頁 53 下 3: 尤度が無限大に \rightarrow パラメータが無限大に
3. 頁 54 最後のプログラムの最終行から下: `table_count(3, ans, pred)` \rightarrow `table_count(2, ans, pred)`。
正答率は $(38 + 40)/100 = 0.78$ \rightarrow 正答率は $(41 + 45)/100 = 0.86$
4. 頁 55 上 12: それが正しくない確率は $1 - \sum_{k \neq \hat{k}} P(y = k|x) = 1 - P(y = \hat{k}|x) \rightarrow$ それが正しい確率は $1 - \sum_{k \neq \hat{k}} P(y = k|x) = P(y = \hat{k}|x)$
5. 頁 55(プログラムをのぞいた) 下 6(数式中): $\mu_{-1} \Sigma^{-1} \mu_{-1} \rightarrow \mu_{-1}^T \Sigma^{-1} \mu_{-1}$
6. 頁 58 上 3: プログラム a, b, c の値に \rightarrow (削除)
7. 例 35 プログラム 17 行目: `level=1` \rightarrow `level=0` (すべて `level=0` になる)
8. 頁 69 プログラム中 11 行目: `v=np. ...` \rightarrow `v= ## 空欄 ##`
9. 頁 59 下 6: 誤り率最大 \rightarrow 誤り率最小

第3章

1. 頁 72 上 2: `cv.linear` \rightarrow `cv_linear`
2. 頁 73 上 1: 「 X_1, X_2, X_3 に依存」 削除
3. 頁 75 2 個目のセルの 6 行目: `order=np.random.choice` \rightarrow `index=np.random.choice`
4. 頁 75 3 個目のセルの 3 行目: `top_seq=list(range(0, 135, 10))` \rightarrow `top_seq=list(range(0, 150, 15))`
5. 頁 76 下 4-5: $X_S \in \mathbb{R}^{r \times p} \rightarrow X_S \in \mathbb{R}^{r \times (p+1)}$ 、 $X_{-S} \in \mathbb{R}^{(N-r) \times p} \rightarrow X_{-S} \in \mathbb{R}^{(N-r) \times (p+1)}$
6. 頁 76 下 2: $\sum_{j=1} x_j x_j^T = \sum_{j \in S} x_j x_j^T + \sum_{j \notin S} x_j x_j^T \rightarrow \sum_{j=1} x_j^T x_j = \sum_{j \in S} x_j^T x_j + \sum_{j \notin S} x_j^T x_j$
7. 頁 77 上 4: Sherman-Morrison-Woodbury \rightarrow Sherman-Morrison-Woodbury
8. 頁 77 下 16: $\hat{\beta} - (X^T X)^{-1} X_S^T y_S - (X^T X)^{-1} \rightarrow \hat{\beta} - (X^T X)^{-1} X_S^T y_S + (X^T X)^{-1}$
9. 頁 78 最初のセルの 11 行目, 頁 86 3 番目の 12 行目: `H[test, test]` \rightarrow `H[test, :][:, test]`
10. 頁 78 例 42 「の値と」 削除
11. 頁 79 上 6: 1.6 節 \rightarrow 1.7 節
12. 頁 80 と 頁 88 問題 38 のセル 8 行目: `np.std(u)` \rightarrow `:np.std(u, ddof=1)`
13. 頁 80 上 1: $\sum_{h=1}^r \{\hat{\alpha}_h - \hat{\alpha}\}^2 \rightarrow \sum_{h=1}^r \{\hat{\alpha}_h - \frac{1}{r} \sum_{j=1}^r \hat{\alpha}_j\}^2$

14. 頁 81 上 5-7 および問題 38: $N \left\{ \sum_{i=1}^N x_i \right\}^2 \rightarrow \frac{1}{N} \left\{ \sum_{i=1}^N x_i \right\}^2$, $N \left\{ \sum_{i=1}^N y_i \right\}^2 \rightarrow \frac{1}{N} \left\{ \sum_{i=1}^N y_i \right\}^2$,
 $N \left\{ \sum_{i=1}^N x_i \right\} \left\{ \sum_{i=1}^N y_i \right\} \rightarrow \frac{1}{N} \left\{ \sum_{i=1}^N y_i \right\} \left\{ \sum_{i=1}^N y_i \right\}$,
15. 頁 81 プログラム中, 頁 89 プログラムの最初のセル: $\text{var}(X), \text{var}(Y) \rightarrow \text{var}(X, \text{ddof}=1), \text{var}(Y, \text{ddof}=1)$
16. 頁 83 1m 関数の \rightarrow 線形回帰の
17. 頁 85 問題 32: 式番号 (3.7) 削除
18. 頁 85 問題 33: $X_S \in \mathbb{R}^{(N-m) \times p} \rightarrow X_S \in \mathbb{R}^{r \times (p+1)}$, $X_{-S} \in \mathbb{R}^{r \times p} \rightarrow X_S \in \mathbb{R}^{(N-r) \times (p+1)}$,
 $\hat{y} = X_S \hat{\beta}_S \rightarrow \hat{y} = X_S \hat{\beta}$
19. 頁 89 問題 39: 上から 3 個目のセルの 3 行目: $y = \text{data}[\text{index}, 0] \rightarrow y = \#\# \text{ 空欄 } \#\#$

第 4 章

1. 頁 93 最後のセル 7: $\text{np.log}(S_min) \rightarrow \text{np.log}(S_min/n)$
2. 頁 94 最初のセル 7: $p-1)/(TSS* \rightarrow k-1)/(TSS/$
3. 頁 94 最初のセル 11: $\text{set_q} \rightarrow \text{set_max}$
4. 頁 94 最後のセル 5,6 (2 箇所): $\text{np.log}(S) \rightarrow \text{np.log}(S/n)$
5. 頁 95 下 1, 頁 104 上 5: $f(y|x, \beta) := \frac{1}{\sqrt{(2\pi\sigma^2)^{p/2}}} \exp\{-\frac{1}{2\sigma^2} \|y - x\beta\|^2\} \rightarrow f(y_i|x_i, \beta) := \frac{1}{\sqrt{2\pi\sigma^2}} \exp\{-\frac{1}{2\sigma^2} (y_i - x_i\beta)^2\}$
6. 頁 96 上 9, 問題 40 ヒント: $\frac{\partial l^2}{\partial \sigma^2} = -\frac{N}{2\sigma^2} - \frac{\|y - X\beta\|^2}{2(\sigma^2)^2} \rightarrow \frac{\partial l}{\partial \sigma^2} = -\frac{N}{2\sigma^2} + \frac{\|y - X\beta\|^2}{2(\sigma^2)^2}$
7. 頁 97 例 47: $\nabla l = -\frac{1}{\sigma^2} \sum_{i=1}^N x_i^T (y_i - x_i\beta) \rightarrow \nabla l = \frac{1}{\sigma^2} \sum_{i=1}^N x_i^T (y_i - x_i\beta)$, $\nabla^2 l = \frac{1}{\sigma^2} \sum_{i=1}^N x_i^T x_i = \frac{1}{\sigma^2} X^T X \rightarrow \nabla^2 l = -\frac{1}{\sigma^2} \sum_{i=1}^N x_i^T x_i = -\frac{1}{\sigma^2} X^T X$, $E[\nabla l] = -\frac{1}{\sigma^2} \sum_{i=1}^N x_i^T E(y_i - x_i\beta) = 0 \rightarrow E[\nabla l] = \frac{1}{\sigma^2} \sum_{i=1}^N x_i^T E(y_i - x_i\beta) = 0$, $E[(\nabla l)^2] \rightarrow E[(\nabla l)(\nabla l)^T]$
8. 頁 98 (4.7) 式, 頁 101 下 6, 頁 102 上 5, 頁 105 下 7 (3 箇所): $(\gamma - \beta)X^T X(\gamma - \beta) \rightarrow (\gamma - \beta)^T X^T X(\gamma - \beta)$
9. P98 下 6: $\int_S f(x)dx = 0 \implies \int_S g(x)dx = 0 \rightarrow \int_S f(x)dx > 0 \implies \int_S g(x)dx > 0$
10. 頁 99 下 9: $f(z_i|x_1, \beta), \dots, f(z_N|x_N, \beta) \rightarrow f(z_1|x_1, \beta), \dots, f(z_N|x_N, \beta)$
11. 頁 103 上 1: $(-1)^j \rightarrow (-1)^{n-j}$

12. 頁 103 下 4, 頁 106 問題 45(f): $E[\log(\frac{\hat{\sigma}^2(S)}{N - k(S) - 1} / \frac{\sigma^2(S)}{N})] \rightarrow E[\log(\frac{\hat{\sigma}^2(S)}{N - k(S) - 1} / \frac{\sigma^2}{N})]$
13. 頁 104 問題 40 最小 \rightarrow 最大
14. 頁 106 問題 44(d): $\frac{N}{2} \log(2\pi\sigma^2 e) + \frac{1}{2}(k+1) \rightarrow \frac{N}{2} \log(2\pi\sigma_k^2 e) + \frac{1}{2}(k+1)$
15. 頁 106 問題 45(b): $E(\frac{U}{m}) = -\frac{1}{m} \rightarrow E(\frac{U}{m}) = \frac{2}{m}$
16. 頁 102 下 11: 右側の $\int_{-\infty}^{\infty}$ を削除

第 5 章

- 頁 112 図 5.1 のキャプション: 例 49 の実行例 \rightarrow 例 48 の実行例
- 頁 114 下 7: まず, (5.3) の第 2 項 \rightarrow (5.1) の第 2 項
- 頁 116 上 1: その場合, (5.5) の関係は成立しない。そこで, \rightarrow ここでは, (5.5) において
- 頁 116 上 5, 問題 52: $\sum_{i=1}^n x_{i,j}^2 = 1 \rightarrow \frac{1}{N} \sum_{i=1}^N x_{i,j}^2 = 1$
- 頁 118 上 2: $\hat{y}_i = \hat{\beta}_i x_{i,1} + \hat{\beta}_2 x_{i,2} \rightarrow \hat{y}_i = \hat{\beta}_1 x_{i,1} + \hat{\beta}_2 x_{i,2}$
- 頁 120 上 1,3 下 4: `cv.glmnet` \rightarrow `LassoCV`
- 頁 122 下 2 ヒント: 削除
- 頁 128 最初のセル: 削除

頁 114 下 5,8, 頁 115 上 2: $n \rightarrow N$.

$$\frac{1}{n} \sum_{i=1}^n x_{i,j} x_{i,k} = \begin{cases} 1, & j = k \\ 0, & j \neq k \end{cases} \rightarrow \frac{1}{N} \sum_{i=1}^N x_{i,j} x_{i,k} = \begin{cases} 1, & j = k \\ 0, & j \neq k \end{cases}$$

$$0 \in -\frac{1}{N} \sum_{i=1}^n x_{i,j} \left(y_i - \sum_{k=1}^p x_{i,k} \beta_k \right) \rightarrow 0 \in -\frac{1}{N} \sum_{i=1}^N x_{i,j} \left(y_i - \sum_{k=1}^p x_{i,k} \beta_k \right)$$

$$0 \in -\frac{1}{N} \sum_{i=1}^n x_{i,j} (r_{i,j} - x_{i,j} \beta_j) \rightarrow 0 \in -\frac{1}{N} \sum_{i=1}^N x_{i,j} (r_{i,j} - x_{i,j} \beta_j)$$

第 6 章

- 頁 129 下 5 (6.1) $y = \begin{cases} 1 + \epsilon, & 2m - 1 \leq |x| < 2m \\ -1 + \epsilon, & 2m - 2 \leq |x| < 2m - 1 \end{cases}, m = 1, 2, \dots$
 $\rightarrow y = \begin{cases} -1 + \epsilon, & 2m - 1 \leq |x| < 2m \\ 1 + \epsilon, & 2m - 2 \leq |x| < 2m - 1 \end{cases}, m = 1, 2, \dots$
- 頁 132 上 4,7,11,12: 添字がずれる $\beta_1, \dots, \beta_{K+4} \rightarrow \beta_0, \dots, \beta_{K+3}$

$$f(x) = \begin{cases} \beta_1 + \beta_2 x + \beta_3 x^2 + \beta_4 x^3, & \alpha_0 \leq x \leq \alpha_1 \\ \beta_1 + \beta_2 x + \beta_3 x^2 + \beta_4 x^3 + \beta_5 (x - \alpha_1)^3, & \alpha_1 \leq x \leq \alpha_2 \\ \beta_1 + \beta_2 x + \beta_3 x^2 + \beta_4 x^3 + \beta_5 (x - \alpha_1)^3 + \beta_6 (x - \alpha_2)^3, & \alpha_2 \leq x \leq \alpha_3 \\ \vdots & \vdots \\ \beta_1 + \beta_2 x + \beta_3 x^2 + \beta_4 x^3 + \beta_5 (x - \alpha_1)^3 \\ \quad + \beta_6 (x - \alpha_2)^3 + \dots + \beta_{K+4} (x - \alpha_K)^3, & \alpha_K \leq x \leq \alpha_{K+1} \end{cases}$$

→

$$f(x) = \begin{cases} \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3, & \alpha_0 \leq x \leq \alpha_1 \\ \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 (x - \alpha_1)^3, & \alpha_1 \leq x \leq \alpha_2 \\ \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 (x - \alpha_1)^3 + \beta_5 (x - \alpha_2)^3, & \alpha_2 \leq x \leq \alpha_3 \\ \vdots & \vdots \\ \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 (x - \alpha_1)^3 \\ \quad + \beta_5 (x - \alpha_2)^3 + \cdots + \beta_{K+3} (x - \alpha_K)^3, & \alpha_K \leq x \leq \alpha_{K+1} \end{cases}$$

3. 頁 132 上 10: $x_3^3 \rightarrow x_N^3$

4. 頁 135 例 55: $d_2(x) - d_4(x) \rightarrow d_2(x) - d_3(x)$, $x_1, x_2, x_3, x_4 \rightarrow \alpha_1, \alpha_2, \alpha_3, \alpha_4$

5. 頁 138 上 14: $\lambda \int_{-\infty}^{\infty} \{g''(x)\}^2 dx \rightarrow \lambda \int_{-\infty}^{\infty} \{f''(x)\}^2 dx$

6. 頁 139 最初のセル 2: `range(2,n-1) → range(2,n)`

7. 頁 140 Python プログラム 13 行目, 問題 64: `[40, 400, 1000] → [1, 30, 80]`

8. 頁 141 上 5, 問題 65:

$$H[\lambda] := X^T (X^T X + \lambda G)^{-1} X \rightarrow H[\lambda] := X (X^T X + \lambda G)^{-1} X^T$$

9. 頁 141 セル 12, 頁 158 問題 65 セル 12: `H[test,test] → H[test,:][:,test]`

$$10. \text{ 頁 143 上 1: } \begin{bmatrix} K_\lambda(x_1, y_1) & K_\lambda(x_1, y_2) & K_\lambda(x_1, y_3) \\ K_\lambda(x_2, y_1) & K_\lambda(x_2, y_2) & K_\lambda(x_2, y_3) \\ K_\lambda(x_3, y_1) & K_\lambda(x_3, y_2) & K_\lambda(x_3, y_3) \end{bmatrix} \rightarrow \begin{bmatrix} K_\lambda(x_1, x_1) & K_\lambda(x_1, x_2) & K_\lambda(x_1, x_3) \\ K_\lambda(x_2, x_1) & K_\lambda(x_2, x_2) & K_\lambda(x_2, x_3) \\ K_\lambda(x_3, x_1) & K_\lambda(x_3, x_2) & K_\lambda(x_3, x_3) \end{bmatrix}$$

11. 頁 143 例 60: 正則性 → 正定値性

12. 頁 145 上 5: 各点 $x \in \mathbb{R} \rightarrow x \in \mathbb{R}^p$

13. 頁 145 下 5: (6.6) → (6.7)

14. 頁 145 セル 28-37, 頁 160 セル 28-37: 4 マス右へ

15. 頁 146 2 番目のセル 3 行目, 頁 148 下から 2 番目のセル 4 行目: `randn(1) → randn(n)`

16. 頁 147 例 62: $h_1(x), h_2(x), h_3(x) \rightarrow h_3(x), h_4(x), h_5(x)$

17. 頁 147 下のセル: `def poly(x,y,z=x)` を以下の 3 行に変換

```
def poly(x,y,z=None):
    if z is None:
        z=x
```

18. 頁 149 下 8: $\sum_{j=3}^{k+1} \beta_j (x - \alpha_{j-2})^3 \rightarrow \sum_{j=3}^{k+1} \beta_j (x - \alpha_{j-2})_+^3$

19. 頁 151 命題 22 の証明: 最初の項は → 第 2 項は, 第 2 項は → 最初の項は

20. 頁 159 問題 66: Narayama-Watson → Nadaraya-Watson

21. 頁 160 問題 67: $y_i - \text{beta}(x)^T [1, x_i] \rightarrow y_i - [1, x_i] \text{beta}(x)$

第7章

1. 頁165下6: $x_{i,1} \dots \rightarrow x_{1,j}, \dots$
2. 頁166上5: 引数 \rightarrow 変数
3. 頁171 図7.5 キャプション下2: 1から12 \rightarrow 1から15
4. 頁174 3番目のセル 7行目: $\text{fr}[i]*(n-\text{fr}[i]) \rightarrow \text{fr}[i]/n*(n-\text{fr}[i])$
5. 頁175 図7.7: 中央の表の2行3列目 49 \rightarrow 1
6. 頁177下11: $\hat{f}_1(x), \dots, \hat{f}(x) \rightarrow \hat{f}_1(x), \dots, \hat{f}_B(x)$
7. 頁181上6、上9: $r_N = r_N - \hat{f}_B(x_N) \rightarrow r_N = r_N - \lambda \hat{f}_B(x_N)$
8. 頁185最後のセル下2, 頁193最後のセル: 9行目から11行目を削除して、12列目の最初を以下のように修正する。 `plt.plot(u,v, \rightarrow plt.plot(nn_seq, out_set[d-1],`
9. 頁188 問題71: すみれ \rightarrow あやめ

第8章

1. 頁196下7: $\beta^T x_i \rightarrow x_i \beta$
2. 頁196下6, 頁213 問題75: $(\beta_0, \beta) \in \mathbb{R}^p \times \mathbb{R} \rightarrow (\beta_0, \beta) \in \mathbb{R} \times \mathbb{R}^p$
3. 頁198下2: $\alpha = -1/\sqrt{2}$ より $\rightarrow \alpha = 1/\sqrt{2}$ より
4. 頁199下9: $x = x_0 \in \mathbb{R}$ において $\rightarrow x = x_0 \in \mathbb{R}^p$ において
5. 頁199下7: $(\beta - \beta^*) \leq f_0(\beta) \rightarrow (\beta - \beta^*) = f_0(\beta)$
6. 頁200 (8.18): $\sum_{i=1}^N \alpha_i y_i x_i \rightarrow \sum_{i=1}^N \alpha_i y_i x_i^T$
7. 頁202 3番目のセル, 206頁のセルの18行目-21行目, : `index=(eps<alpha[:,0])&(alpha[:,0]<C-eps)` の一行に変更
8. 頁205下2: `sum(X[:,i]*X[:,j]) \rightarrow np.dot(X[i,:],X[j,:])`
9. 頁208上3, 図8.5 キャプション, 頁210上3, 頁218 問86: ラジカルカーネル \rightarrow ラジアルカーネル
10. 頁211下1, 頁212上2, 4: $\epsilon \rightarrow \epsilon_i$
11. 頁213 問題75: $\beta_0 + x_i \beta < 0 \rightarrow y_i(\beta_0 + x_i \beta) < 0$
12. 頁217 問題83,84: `K.poly \rightarrow K_poly, K.linear \rightarrow K_linear`

第9章

1. 頁 225 下 6: L_2 ノルム \rightarrow L_2 ノルム
2. 頁 230 図 9.6: complete と centroid の図が逆
3. 頁 232 上 7: $\phi_1 \in \mathbb{R}^N, \phi_2 \in \mathbb{R}^N, \dots, \phi_p \in \mathbb{R}^N \rightarrow \phi_1 \in \mathbb{R}^p, \phi_2 \in \mathbb{R}^p, \dots, \phi_p \in \mathbb{R}^p$
4. 頁 232 下 10-11: $\phi_i \rightarrow \phi_1$ (2箇所)
5. 頁 233 上 14: 命題 9 より \rightarrow 命題 8 より
6. 頁 233 の pca 関数のセルの 5 行目: $\text{Sigma}=\text{X.T@X} \rightarrow \text{Sigma}=\text{X.T@X}/n$
7. 頁 236 の関数 f_1, f_2 (2 個目のセルも)、頁 246 問 96: f_1 が $y = T[0, 1] / T[0, 0] * x$
 f_2 が $y = T[1, 1] / T[1, 0] * x$
8. 頁 236 の 5 個目のセルの 7 行目, 頁 246 問 96: `plt.axis("equal")` \rightarrow
`plt.gca().set_aspect("equal", adjustable="box")`
9. 頁 239 下 2: 目的関数 \rightarrow 目的変数
10. 頁 244 問 91 3 行目: `single_complete` \rightarrow `dist_single`
11. 頁 247 問 99 下から 2 個目のセル: `evr=pca...` \rightarrow `evr= # 空欄 #`