Statistical Learning with Math and R

Joe Suzuki

As of 2021 Dec. 7

The modification will be made in `https://github.com/prof-joe`

Chap. 2

1. P8 Footnote: satisfies (1.2) $\mathbb{R}$ is said $\rightarrow$ satisfies (1.2) is said
2. P23 L1 below: $\sigma_{i,j} := E(\hat{\beta}_i - \beta_i)(\hat{\beta}_j - \beta_j)^T \rightarrow \sigma_{i,j} := E(\hat{\beta}_i - \beta_i)(\hat{\beta}_j - \beta_j)$
3. P26 Section Title: $\hat{\beta}_j = 0 \rightarrow \beta_j = 0$

Chap. 3

1. P52 L2 below: $\mathbb{R}^p \times \mathbb{R} \rightarrow \mathbb{R}^p \times \{-1, 1\}$
2. P54 L6 below: $1 - \sum_{k \neq \hat{k}} P(y = k|x) = 1 - P(y = \hat{k}|x) \rightarrow 1 - \sum_{k \neq \hat{k}} P(y = k|x) = P(y = \hat{k}|x)$
3. P55 Program L9: `mean(c(x.1,y.1))` $\rightarrow$ `c(mean(x.1),mean(y.1))`, `mean(c(x.2,y.2))` $\rightarrow$ `c(mean(x.2),mean(y.2))`
4. P56 L11 below: `z=array(dim=n/2)` $\rightarrow$ `z=array(dim=length(test))`
5. P59 L7-8 below (four locations) and P65 Problem 30 (two locations): `pnorm` $\rightarrow$ `dnorm`.

Chap. 4

1. P74 L1 above: Section 2.6 $\rightarrow$ Section 2.7.
2. P74 L3 below: $\hat{\sigma}^2(\hat{\alpha}) := \frac{1}{r-1}\sum_{h=1}^{r}\{\hat{\alpha}_h - \hat{\alpha}\}^2 \rightarrow \hat{\sigma}^2(\hat{\alpha}) := \frac{1}{r-1}\sum_{h=1}^{r}\{\hat{\alpha}_h - \frac{1}{r}\sum_{j=1}^{r}\hat{\alpha}_j\}^2$

Chap. 5

1. P85 L20 below: `q=T[,j]; S=sum((lm(y~X[,q])$fitted.values-y)^2)/n`
   $\rightarrow$ `q=T[,j]; S=sum((lm(y~X[,q])$fitted.values-y)^2)`
2. P87 L1 below: $f(y|x,\beta) := \frac{1}{\sqrt{(2\pi)^{p/2}}}\exp\{-\frac{1}{2\sigma^2}\|y - x\beta\|\}$
   $\rightarrow f(y_i|x_i,\beta) := \frac{1}{\sqrt{2\pi\sigma^2}}\exp\{-\frac{1}{2\sigma^2}(y_i - x_i\beta)^2\}$
3. P88 L9 above, and Prob. 40: $\frac{\partial l^2}{\partial \sigma^2} = -\frac{N}{2\sigma^2} - \frac{\|y - X\beta\|^2}{2(\sigma^2)^2} \rightarrow \frac{\partial l}{\partial \sigma^2} = -\frac{N}{2\sigma^2} + \frac{\|y - X\beta\|^2}{2(\sigma^2)^2}$
4. P89 Example 47: $\nabla l = -\frac{1}{\sigma^2}\sum_{i=1}^{N} x_i^T(y_i - x_i\beta) \rightarrow \nabla l = \frac{1}{\sigma^2}\sum_{i=1}^{N} x_i^T(y_i - x_i\beta)$、$\nabla^2 l = \frac{1}{\sigma^2}\sum_{i=1}^{N} x_i^T x_i = \frac{1}{\sigma^2}X^T X \rightarrow \nabla^2 l = -\frac{1}{\sigma^2}\sum_{i=1}^{N} x_i^T x_i = -\frac{1}{\sigma^2}X^T X$、$E[\nabla l] = -\frac{1}{\sigma^2}\sum_{i=1}^{N} x_i^T E(y_i - x_i\beta) = 0 \rightarrow E[\nabla l] = \frac{1}{\sigma^2}\sum_{i=1}^{N} x_i^T E(y_i - x_i\beta) = 0$、$E[(\nabla l)^2] \rightarrow E[(\nabla l)(\nabla l)^T]$

5. P89 Prop. 17: Suppose J is nonsingular. (at the beginning); "does not exceed" $\rightarrow$ "is not below"
6. P91 L3 above: $\int_S f(x)dx = 0 \Longrightarrow \int_S g(x)dx = 0 \rightarrow \int_S f(x)dx > 0 \Longrightarrow \int_S g(x)dx > 0$
7. P95 L5 above: Remove the right $\int_{-\infty}^{\infty}$
8. P97 L3 above: $f(y|x,\beta) := \frac{1}{\sqrt{2\pi\sigma^2}} \exp\{-\frac{1}{2\sigma^2}\|y - x\beta\|\} \rightarrow f(y_i|x_i,\beta) := \frac{1}{\sqrt{2\pi\sigma^2}} \exp\{-\frac{1}{2\sigma^2}(y_i - x_i\beta)^2\}$

Chap. 6

1. P106 L2/L5 below and P197 L1 below: $n \rightarrow N$.
   $\frac{1}{n}\sum_{i=1}^{n} x_{i,j}x_{i,k} = \begin{cases} 1, & j = k \\ 0, & j \neq k \end{cases} \rightarrow \frac{1}{N}\sum_{i=1}^{N} x_{i,j}x_{i,k} = \begin{cases} 1, & j = k \\ 0, & j \neq k \end{cases}$
   $0 \in -\frac{1}{N}\sum_{i=1}^{n} x_{i,j}\left(y_i - \sum_{k=1}^{p} x_{i,k}\beta_k\right) \rightarrow 0 \in -\frac{1}{N}\sum_{i=1}^{N} x_{i,j}\left(y_i - \sum_{k=1}^{p} x_{i,k}\beta_k\right)$
   $0 \in -\frac{1}{N}\sum_{i=1}^{n} x_{i,j}(r_{i,j} - x_{i,j}\beta_j) \rightarrow 0 \in -\frac{1}{N}\sum_{i=1}^{N} x_{i,j}(r_{i,j} - x_{i,j}\beta_j)$
2. P110 Fig 6.6: $x, y \rightarrow \beta_1, \beta_2$ (The axes for the both figure).

Chap. 7

1. P119 L2 Program: `y=round(x)%%2*2-1+rnorm(n)*0.2` $\rightarrow$ `y=ceiling(x)%%2*2-1+rnorm(n)*0.2`
2. P124 Example 55: $x_1, x_2, x_3, x_4 \rightarrow \alpha_1, \alpha_2, \alpha_3, \alpha_4$
3. P127 L9 above: $\lambda \int_{-\infty}^{\infty} \{g''(x)\}^2 dx \rightarrow \lambda \int_{-\infty}^{\infty} \{f''(x)\}^2 dx$
4. P128 Example 57, and Prob. 64: `c(40, 400, 1000)` $\rightarrow$ `c(1, 30, 80)`
5. P129 L3 below, and Prob. 65: $H[\lambda] := X^T(X^TX + \lambda G)^{-1}X \rightarrow H[\lambda] := X(X^TX + \lambda G)^{-1}X^T$
6. P130 Example 58, and Prob. 65: `0*1*rnorm(n)` $\rightarrow$ `0.1*rnorm(n)`
7. P130 L1 above: $\begin{bmatrix} K_\lambda(x_1,y_1) & K_\lambda(x_1,y_2) & K_\lambda(x_1,y_3) \\ K_\lambda(x_2,y_1) & K_\lambda(x_2,y_2) & K_\lambda(x_2,y_3) \\ K_\lambda(x_3,y_1) & K_\lambda(x_3,y_2) & K_\lambda(x_3,y_3) \end{bmatrix} \rightarrow \begin{bmatrix} K_\lambda(x_1,x_1) & K_\lambda(x_1,x_2) & K_\lambda(x_1,x_3) \\ K_\lambda(x_2,x_1) & K_\lambda(x_2,x_2) & K_\lambda(x_2,x_3) \\ K_\lambda(x_3,x_1) & K_\lambda(x_3,x_2) & K_\lambda(x_3,x_3) \end{bmatrix}$
8. P133 L12 above: $x \in \mathbb{R} \rightarrow x \in \mathbb{R}^p$
9. P134 L14: (6.6) $\rightarrow$ (6.7)
10. P134 Example 62: $h_1(x), h_2(x), h_3(x) \rightarrow h_3(x), h_4(x), h_5(x)$
11. P135 Example 63, and Prob. 68: `randn(1)` $\rightarrow$ `randn(n)`
12. P139 Proof of Prop.: first term $\rightarrow$ second term, second term $\rightarrow$ first term
13. Prob. 67: $y_i - \beta(x)^T[1, x_i] \rightarrow y_i - [1, x_i]\beta(x)$

Chap. 8

1. P148 (8.3): $f(x,y) \rightarrow f_{XY}(x,y)$
2. P155 Fig7.5 caption: from 1 to 12 $\rightarrow$ from 1 to 15
3. P158 Program 7: `T=0; ... z[j])/n` $\rightarrow$ `T=0; ... z[j])`

Chap. 9

1. P174 L1 below: $\alpha' \in \mathbb{R}^m \to \alpha' \geq 0$
2. P175 L17 below: $\alpha = -1/\sqrt{2} \to \alpha = 1/\sqrt{2}$
3. P176 L4 above: $x = x_0 \in \mathbb{R} \to x = x_0 \in \mathbb{R}^p$
4. P176 L7 above: an arbitrary $\to$ each solution
5. P176 L7 above: $(\beta - \beta^*) \leq f_0(\beta) \to (\beta - \beta^*) = f_0(\beta)$
6. P177 (9.18): $\sum_{i=1}^N \alpha_i y_i x_i \to \sum_{i=1}^N \alpha_i y_i x_i^T$
7. P186 L4, 6, 9 below: $\epsilon \to \epsilon_i$

Chap. 10

1. P197 L6 below: $L_2 \to$ L2
2. P203 Example 85 Program : insert `n=100` in the first line
3. P205 pca function L4: `t(X)%*%X` $\to$ `t(X)%*%X/n`
4. P211 Problem 91: `single.complete` $\to$ `dist.single`