

Modern Multivariate Statistical Techniques

Regression, Classification, and Manifold Learning

Errata

“Line 5” means 5th line down from the top. “Line -5” means 5th line up from the bottom (not counting footnotes).

Corrections to First Printing

CHAPTER 1

Page 13, line 22: “thatgives” should be “that gives.”

CHAPTER 2

Page 19, line 20: “Normal21” should be “Normal22.”

Page 44, Exercise 2.4, line 3: $1x > 0$ should read $x > 0$.

Page 44, Exercise 2.10, line -3: Sentence should read “Discuss how you would estimate.” Also, 2.10 should be in boldface.

CHAPTER 3

Page 46, line -15: $(K \times J)$ -matrix.

Page 47, Equation (3.5): $-\mathbf{E}\mathbf{F}^\tau$ should be $-\mathbf{E}^{-1}\mathbf{F}^\tau$.

Page 49, line -6: \mathbf{R} should be \Re .

Page 53, line -7: “see Exercise 3.4” is not relevant here.

Page 53, Equation (3.43): Last term on rhs should be $\frac{\partial y_I}{\partial x_K}$.

Page 60, Equation (3.98): σ^{-1} should be σ^{-r} .

Page 61, line 8: (3.106) should be (3.105).

Pages 61, 62, Equations (3.109), (3.110), and (3.111): The identity matrices have different dimensions — In the top row of each matrix, the identity matrix has dimension r and in the bottom row it has dimension s .

Page 62, line 1: “ r -vector” should be “ $(r + s)$ -vector.”

Page 62, Equation (3.111): Σ_{XY} should be Σ_{YX} .

Page 64, Equation (3.127): $|\Sigma|^{-1/2n}$ should be $|\Sigma|^{-n/2}$.

Page 62, Equation (3.133): $\mathbf{I}_{(2,2)}$ should be $\mathbf{I}_{(2,2)}$.

Page 65, line 8 (2nd line of property 2): \mathcal{W}_r should be \mathcal{W}_p .

Page 65, property 4: Restate as follows.

Let $\mathcal{X} = (\mathbf{X}_1, \dots, \mathbf{X}_n)^\tau$, where $\mathbf{X}_i \sim \mathcal{N}_r(\mathbf{0}, \Sigma)$, $i = 1, 2, \dots, n$, are *independently and identically distributed (iid)*. Let \mathbf{A} be a symmetric $(n \times n)$ -matrix with $\nu = \text{rank}(\mathbf{A})$, and let \mathbf{a} be a fixed r -vector. Let $\mathbf{y} = \mathcal{X}^\tau \mathbf{a}$. Then, $\mathcal{X}^\tau \mathbf{A} \mathcal{X} \sim \mathcal{W}_r(\nu, \Sigma)$ iff $\mathbf{y}^\tau \mathbf{A} \mathbf{y} \sim \sigma_a^2 \chi_\nu^2$, where $\sigma_a^2 = \mathbf{a}^\tau \Sigma \mathbf{a}$.

Page 66, Equation (3.143): last term on rhs, $+n$ should be $-\frac{n}{2}$.

Page 67, line 3: Should read $\text{tr}(\mathbf{T}\mathbf{T}^\tau) = \sum_{i=1}^r t_{ii}^2 + \sum_{i>j} t_{ij}^2$.

Page 67, line -6: Should read “idempotent with rank $n - 1$.”

Page 67, line -3: $\mathbf{b}^\tau \mathcal{X}$ should be $\mathcal{X}^\tau \mathbf{b}$.

Page 67, Equation (3.148): n should be $n - 1$.

Page 68, lines 5, 6: $\mathbf{b}^T \mathcal{X}$ should be $\mathcal{X}^T \mathbf{b}$.
Page 72, Exercise 3.1: Should read “Show that $(\mathbf{x}^T \mathbf{y})^2 \leq (\mathbf{x}^T \mathbf{x})(\mathbf{y}^T \mathbf{y})$, where ...”
Page 72, Exercise 3.2: The second integral on the rhs should not be squared.
Page 73, Exercise 3.6: $\mathbf{A}^T \boldsymbol{\Sigma} \mathbf{A}$ should be $\mathbf{A} \boldsymbol{\Sigma} \mathbf{A}^T$.
Page 73, Exercise 3.11: \mathcal{N} should be \mathcal{N}_r .
Page 73, Exercise 3.12: $\mathbf{I}_s \otimes \boldsymbol{\Sigma}$ should be $\boldsymbol{\Sigma} \otimes \mathbf{I}_s$.
Page 74, Exercise 3.16: There is a σ^2 missing from the statement of Stein’s Lemma. It should read as follows: Show that $E\{g(X)(X - \theta)\} = \sigma^2 E\{g'(X)\}$.

CHAPTER 4

Page 76, line -7: remove comma and replace “which” by “that.”
Page 85, line 15: It should read “See Exercise 4.10.”
Page 92, line -12: It should read “See Exercise 4.11.”
Page 96, line 5: Replace (4.61) by (4.59).
Page 98, line -15: Replace (4.59) by (4.56).
Page 103, lines 6-7: Add (see Section 12.9.1) after “EM algorithm.”
Page 105, line 1: $\hat{p}_{\ell+1}$ should be defined as $N_{\ell+1}/nh_n$.

CHAPTER 5

Page 112, line -8: Exercise 5.4 should be Exercise 5.4.
Page 115, line -3: $j = 1, 2, \dots, r + 1$.
Page 116, line 6: $\boldsymbol{\beta}$ should be $\hat{\boldsymbol{\beta}}$.
Page 120, line -4: (5.65) should be (5.53).
Page 129, line -7: “the shrinkage”.
Page 129, last line of footnote: “See”.
Page 131, line -4: “represents,” not “reprents.”
Page 133, last line: Compositor’s error — Remove extra space between ‘yet’ and ‘unknown.’
Page 136, line 13: Reference to Exercise 5.11 should be removed.
Page 142, line 11: Compositor’s error — Remove “spaceskip3pt plus2pt minus2pt.”
Page 148, line -2: L_q should be ℓ_q .
Page 150, lines 8 and 9: L_2 should be ℓ_2 , and L_1 should be ℓ_1 .
Page 152, Step 3 in the Forwards-Stagewise Algorithm: remove the first $\hat{\beta}_{j_1}$.
Page 154, last line: Add reference to Draper and Smith (1998).
Page 155, line 15: (5.125) should be (5.119).
Page 155, Exercise 5.1: Change (5.12) to (5.13), and change (5.13) and (5.14) to (5.14) and (5.15), respectively.
Page 155, Exercise 5.2: The “same” matrix-inversion formula refers to that used in Exercise 5.1.

CHAPTER 6

Page 179, line -9: Should be $\boldsymbol{\Sigma}_{YY} \boldsymbol{\Gamma}$ (not just $\boldsymbol{\Sigma}_{YY}$).
Page 185, top: Replace the last vector $\hat{\boldsymbol{\mu}}^{(1)}$ by $\hat{\boldsymbol{\mu}}^{(3)}$.
Page 185, line -11: (6.86) should be (6.87).

Page 187, –16: It should read “ dE is shorthand for $\Delta \widehat{\Sigma}_{\mathcal{E}\mathcal{E}}^{(t)}$ ”.

Page 191, Exercise 6.2: Add that $\mathbf{\Gamma} = \Sigma_{YY}^{-1}$ and $\mathbf{R} = \Sigma_{YY}^{-1/2} \Sigma_{YX} \Sigma_{XX}^{-1} \Sigma_{XY} \Sigma_{YY}^{-1/2}$.

Page 192, Exercise 6.6: Add that $\widehat{\mathbf{R}} = \widehat{\Sigma}_{YY}^{-1/2} \widehat{\Sigma}_{YX} \widehat{\Sigma}_{XX}^{-1} \widehat{\Sigma}_{XY} \widehat{\Sigma}_{YY}^{-1/2}$.

Page 193, Exercise 6.11: This exercise should be in Chapter 7 because \mathbf{R} is not defined until page 220, equation (7.59).

Page 193, Exercise 6.13: It should read “Generalize Exercise 6.12”

CHAPTER 7

Page 203, line –8: Insert “is” between “ \mathbf{v}_j ” and “associated.”

Page 209, line –4: (7.33) should be (7.32).

Page 213, line –7: “ L_2 ” should be “ ℓ_2 .”

Page 214, line 12: (7.42) should be (7.44).

Page 216, line –12: (7.44) should be (7.50).

Page 216, line –11: (7.45) should be (7.52).

Page 218, Figure 7.7: Compositor’s error — Figure should be printed in color.

Page 227, line –12: (7.106) should be (7.109).

Page 229, line –15: Change \mathbf{x} to \mathbf{X} .

Page 231, line 1: Insert “to” between “is” and “take.”

CHAPTER 8

Page 257, equation (8.71): The last term on the rhs should read $\log_e(\pi_2/\pi_1)$.

Page 260, line –15: Exercise 8.13 should be Exercise 8.14.

Page 264, line –14: Delete “(see Exercise 8.3).”

Page 264, line –11: Change the first \mathbf{S}_{tot} to \mathbf{S}_W .

Page 266, line 6: Change “(Figure 8.2)” to “(see Figure 8.3).”

Page 269, line 7: Delete “(see Exercise 8.4).”

Page 273, Figure 8.5: Caption should refer to Table 8.9 (not 8.7).

Page 276, Figure 8.6: Compositor’s error — remove the “2pt” below three of the scatterplots (Primate Scapulae, Pendigits, and Glass).

CHAPTER 9

Page 284, footnote: Change UCI Repository address for the `cleveland.data` to `archive.ics.uci.edu/ml/datasets/Heart+Disease`.

Page 284, line 27: Change “effect” to “defect.”

Page 285, line –12: “become” should be “becomes.”

Page 288, line 8: (7.113) should be (7.115).

Page 289, equation (9.8): In the first term on the rhs, the denominator of the log should have a subscript 2+ (not 1+).

Page 300, line –6: (9.24) should be (9.34).

Page 304, Figure 9.8: Node ovals should be in bold.

Page 309, Figure 9.10: Node ovals should be in bold.

CHAPTER 10

Page 327, line 2: In the definition of B , change the maximum over $\mathbf{X}_{*i} \in \Pi_1$ to the maximum over $\mathbf{X}_i \in \Pi_1$ (i.e., remove the asterisk).

Page 330, line -2: “because” should be “became.”
Page 349, line 11: remove comma after “projections.”
Page 354, line -14: Replace \mathbf{w} by $\boldsymbol{\omega}$.
Page 368, line -2: Add “so” between “value” and “that.”

CHAPTER 11

Page 372, line -10: The support vectors in Figure 11.1 are not indentified by circles around them. Change the sentence to read: “The support vectors in Figure 11.1 are those points lying on the hyperplanes H_{-1} and H_{+1} .”
Page 373, line 8: *primal!functional* should be *primal functional*.
Page 373, equation (11.15): The zero on the rhs should be the vector $\mathbf{0}$.
Page 374, equation (11.23): the vector $\mathbf{0}$ should be the scalar 0.
Page 376, line -9: Exercise 11.2 should be Exercise 11.3.
Page 377, caption to Figure 11.2: This is a diagram of the linearly nonseparable case (not the nonlinearly separable case). Also, on line 4, remove the word “circled.”
Page 377, line -2: (11.37) should be (11.40), and (11.33) should be (11.36).
Page 378, line 1: (11.18) should be (11.21).
Page 381, line 2 (below Table 11.1): (12.55) should be (11.55).
Page 381, line -10: if $d = 2$, then $\dim(\mathcal{H}) = 33, 153$.
Page 382, line 14: “(see Section 2.2.1)” should be “(see Section 2.3.3).”
Page 383, line -6: (11.62) should be (11.59).
Page 384, line -14: “(see Exercise 11.8)” should be “(see Exercise 11.11).”
Page 385, line 1: (11.42) should be (11.47).
Page 385, line 10: “A initial grid” should be “An initial grid.”
Page 385, line -10: “amd” should be “and.”
Page 386, line -3: “computating” should be “computing.”
Page 388, equation (11.67): Delete the last equality in that display and move the definition of $(x)_+$ so that it follows the definition of hinge loss.
Page 388, line -4: (11.69) should be (11.68).
Page 390, line 15: “approaches” not “approachs.”
Page 396, lines 5, 6: The scalar 0s should be matrix $\mathbf{0}$ s.
Page 397, equation (11.110): \mathcal{Y} should be \mathcal{Y} .
Page 401, line 7: “see Exercise 11.3” should be “see Exercise 11.2.”
Page 406, Exercise 11.11: (11.40) should be (11.41).

CHAPTER 12

Page 409, line -6 (excluding footnote): “incorrectly” should be “incorrect.”
Page 413, line -11: s_2 should be s_j .
Page 413, line -11: Summation for \bar{x}_ℓ should be over k (not ℓ).
Page 413, lines -15 to -7: The “1-correlation” criterion is used not for clustering variables, but for clustering observations.
Page 424, line -6: Replace “Kmed” by “med” in the subscript of ESS.
Page 451, line 4: (12.14) should be (12.31).
Page 451, line 9: (12.20) should be (12.36).
Page 451, line 10: (12.21) should be (12.37).

Page 457, line -10: (13.49) should be (12.49) and (13.50) should be (12.50).
Page 457, line -2: Reference to Section 9.5.1 should be to Section 8.5.1.
Page 458, line 14: (13.56) should be (12.56).
Page 458, line 16: (13.54) should be (12.54).
Page 458, line -3: reference to Exercise 12.9 should be to Exercise 12.8.
Page 460, line 11: reference to Figures 13.11 and 13.12 should be to Figures 12.11 and 12.12, respectively.
Page 460, line 18: The website should read:
`www-stat.stanford.edu/~owen/clickwrap/plaid.html`
Page 461, Exercise 12.1: Comparisons to Table 12.10 should be to Table 12.6.

CHAPTER 13

Page 470, line -8: The reference to Buja, Swayne, Littman, and Hofmann (2002) should be to Buja, Swayne, Littman, Dean, Hofmann, and Chen (2008). For the complete reference, see below.
Page 475, line -3: There should be no right parenthesis (“}”) at the end of the line. Also, the last three letters (L, K, Y) in the middle row of the display should each be placed one position to the right.
Page 480, line 7: Reference to Section 7.2.2 should be to Section 7.2.3.
Page 498, line 6: (13.34) should be (13.35).
Page 502, line -14: Reference to Section 16.7.3 should be to Section 16.6.3.
Page 503, Exercise 13.1: The color-stimuli experiment is outlined in Section 13.2.1, not 14.2.1. Also, comparison of the solution should be made with Figure 13.3.
Page 504, Exercise 13.5: Compare to the monotonicity property in (13.28), not (13.27).

CHAPTER 14

Page 515, footnote: The website should read:
`www.stat.lsa.umich.edu/~culpmp/math/ada/img.html`
Page 532, Table 14.3: In step 3, last bullet, and in step 4, β_t should be ρ_t . In step 4, the summation should be $\sum_{t=0}^T$.
Page 534, lines 14, 18, 21: L_1 should be ℓ_1 .
Page 534, line -13: L_2 should be ℓ_2 .

CHAPTER 15

Page 507, line -18: Reference should be made to Section 5.4.3, not to 5.5.3.
Page 561, bottom line: Should read “maximize negentropy $\mathcal{J}(Y)$.”
Page 571, Figure 15.3: Compositor’s error — Figure should be printed in color.

CHAPTER 16

Page 599–600: Figure 16.1 is not divided into “noiseless” (upper panel) and “noisy” (lower panel) cases. The noisy case is not displayed.
Page 616, line -17: Should be 20,000 points (not 2,000).
Page 624, line 16: (10.60) should be (16.43).
Page 624, line -16: (16.56) should be (16.44).

Page 630, line -21: The website should read:
`lear.inrialpes.fr/~verbeek/software.`

CHAPTER 17

REFERENCES

Page 672, Reference [77]: The reference should read:

Buja, A. and Swayne, D.F. (2002). Visualization methodology for multidimensional scaling, *Journal of Classification*, **19**, 7–43.

Page 672, Reference [78]: The reference should be updated to read:

Buja, A., Swayne, D.F., Littman, M.L., Dean, N., Hofmann, H., and Chen, L. (2008). Data visualization with multidimensional scaling, *Journal of Computational and Graphical Statistics*, **17**, 444–472.

Page 674, Reference [99]: Page numbers should be 979–985.

Page 691, Reference [342]: *Hour*nal should be *Journal*.

Page 698, Reference [441]: Should be *Journal of Machine Learning Research*.

Acknowledgements

The author is grateful to Donna Ankerst, Habib Esmaeili, Russell Ivory, Stephen B. Vardeman and Jue (Andy) Wang for pointing out many of these errata.

Last revised: 9th November 2009