Errata

Note: Some of these corrections have been incorporated into the 4th printing. Some corrections are due to errors introduced in the 4th printing.

Typos and minor changes

- p. 5, Exercise 1.4: Remove the quotation marks before and after the word "or" at the end of the first line. In other words, the statement to be proved should read "A $\Leftrightarrow$ (A and B) or (A and (not B))." [Printing 1–3]

- p. 19, Section 2.4, item 4 in list of formulas for absolute value: The left hand side should read $|\,|a| - |b|\,|$ (the symbols to the right and left should not look like "norm" ($\|$), but instead two nested symbols for absolute value, with a small space in between the two vertical bars). [Printing 1–3]

- p. 19, Section 2.4, item 4 in list of formulas for absolute value: This item should read: $||a| - |b|| \leq |a - b| \leq |a| + |b|$. The middle term is misprinted in the 4th printing. [Printing 4]

- p. 30, third equation from bottom: The subscript of the rightmost q should be $1i$, not ii. [Printing 1–current]

- p. 30, second equation from bottom: Four subscripts should be changed. The equation should read as:

  $$\alpha_{i1} q_{i-1\,1} + \alpha_{i2} q_{i-1\,2} + \cdots + \alpha_{ii} q_{i-1\,i} = 0.$$

  [Printing 1–current]

- p. 31, third equation: put a period after the equation. [Printing 1–3]

- p. 37, first displayed equation: the second component of the vector should be 1. The equation should read:

  $$\boldsymbol{x}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

  [Printing 1–3]

- p. 43, definition of "linear variety" in the middle of the page: The third (last) paragraph should be changed to:

  > A *linear variety* is a set of the form $\{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}\}$ for some matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ and vector $\boldsymbol{b} \in \mathbb{R}^n$. If $\dim \mathcal{N}(\boldsymbol{A}) = r$, we say that the linear variety has dimension r. A linear variety is a subspace if and only if $\boldsymbol{b} = \boldsymbol{0}$. If $\boldsymbol{A} = \boldsymbol{O}$, the linear variety is $\mathbb{R}^n$. If the dimension of the linear variety is less than n, then it is the intersection of a finite number of hyperplanes.

  [Printing 1–3]

- p. 43, last paragraph, 3rd line: "$\boldsymbol{A} = \mathbf{0}$" should read "$\boldsymbol{A} = \boldsymbol{O}$" (the bold-faced 0 should be changed to bold-faced O). [Printing 4]

- p. 48: Delete the last two senteces before the Exercises. [Printing 4]

- p. 53, sentence above the matrix at the bottom: Change two occurrences of Df to ∇f. [Printing 1–current]

- p. 54, Section 5.3, line 1: The symbol for the function f should be bold-faced (i.e., $\boldsymbol{f} : \Omega \to \mathbb{R}^m$). [Printing 1–current]

- p. 60, 6th line from the top: In the expression for R_m, replace $(n-1)!$ with $(m-1)!$. [Printing 1–current]

- p. 63, Exercise 5.7, part c: The symbol x_0 should be bold-faced, as in $\boldsymbol{x}_0$. [Printing 1–3]

- p. 76, Exercise 6.3, part b: The correct question should be "Find the directional derivative of f at $[1,1]^T$ with respect to a unit vector in the direction of maximal rate of increase." (The phrase "with respect to a unit vector" is missing.) [Printing 1–3]

- p. 79, 13 lines from the top: Replace the word "subsest" with "subset" (spelling error). [Printing 4]

- p. 79: In Exercise 6.13, the following assumption on the set Ω is needed: There exists $\varepsilon > 0$ such that for all $\boldsymbol{x} \in \Omega$, $\|\boldsymbol{x} - \boldsymbol{x}^*\| < \varepsilon$, there is a feasible direction $\boldsymbol{d}$ at $\boldsymbol{x}^*$ such that $\boldsymbol{x} = \boldsymbol{x}^* + \boldsymbol{d}$.

  A simple way to correct the statement of the Exercise is to put the word "convex" in front of the word "subset" in the first line of the Theorem. [Printing 1–3]

- p. 79, Exercise 6.15, first line: Change u_{x+1} to u_{k+1} (the subscript should be $k+1$). [Printing 1–current]

- p. 84, Section 7.2: The first displayed equation in the section should read $\rho_{k+1}(1-\rho_k) = 1 - 2\rho_k$ (the sign on the left side should be $-$, not $+$). [Printing 1–3]

- p. 96, last equation: Remove the negative sign in front of the quotient; i.e., $x^{(k+1)} = ...$ with no negative sign in the front of the right hand side. Same comment applies to the second equation on p. 97. [Printing 1–3]

- p. 98, 2nd line after first displayed equation: Change $\alpha_k \boldsymbol{d}^{(k)}$ to $\alpha \boldsymbol{d}^{(k)}$ (there should be no subscript k after α). [Printing 1–current]

- p. 111: In Lemma 8.1, γ_k should be defined as follows: if $\boldsymbol{g}^{(k)} = \mathbf{0}$ then $\gamma_k = 1$, and if $\boldsymbol{g}^{(k)} \neq \mathbf{0}$ then

$$\gamma_k = \alpha_k \frac{\boldsymbol{g}^{(k)T} \boldsymbol{Q} \boldsymbol{g}^{(k)}}{\boldsymbol{g}^{(k)T} \boldsymbol{Q}^{-1} \boldsymbol{g}^{(k)}} \left(2 \frac{\boldsymbol{g}^{(k)T} \boldsymbol{g}^{(k)}}{\boldsymbol{g}^{(k)T} \boldsymbol{Q} \boldsymbol{g}^{(k)}} - \alpha_k \right)$$

  (The equation given in Lemma 8.1 is ill-defined if $\boldsymbol{g}^{(k)} = \mathbf{0}$). [Printing 1–3]

- p. 112, 2nd line of Theorem 8.1: Change $\gamma_k \geq 0$ to $\gamma_k > 0$ (the inequality should be strict). [Printing 1–3]
- p. 112: In Theorem 8.1, the proof applies only to the case where $\gamma_k < 1$ for all k. If $\gamma_k = 1$ for some k, then it follows from the definition of γ_k (see item above) that $\gamma_i = 1$ for all $i \geq k$. Hence, the result of Theorem 8.1 follows trivially in this case. Therefore, to complete the proof of Theorem 8.1, we should add the following remark just after the first sentence in the proof:

  > Assume that $\gamma_k < 1$ for all k, for otherwise the result holds trivially.

  [Printing 1–3]
- p. 113, 1st and 4th lines after the end of the proof of Theorem 8.1: Change $\gamma_k \geq 0$ to $\gamma_k > 0$ (the inequality should be strict). [Printing 1–current]
- p. 114, 5th line of the proof of Theorem 8.2: Change $\gamma_k \geq 0$ to $\gamma_k > 0$ (the inequality should be strict). [Printing 1–3]
- p. 115, 7th line of the proof of Theorem 8.3: Change $\gamma_k \geq 0$ to $\gamma_k > 0$ (the inequality should be strict). [Printing 1–3]
- p. 120, Exercise 8.7: Change $\gamma_k \geq 0$ to $\gamma_k > 0$ (the inequality should be strict). [Printing 1–3]
- p. 120, last line of Exercise 8.8: add the word "to" before $\nabla f(\boldsymbol{x}^{(k+1)})$. [Printing 1–3]
- p. 128, 2nd line: Replace period after 0 with comma. [Printing 1–3]
- p. 131, 3rd line of Exercise 9.4: In $F(\boldsymbol{x}^{(k)})$, F should be bold faced ($\boldsymbol{F}(\boldsymbol{x}^{(k)})$). [Printing 1–3]
- p. 135, last equation at the bottom: The right most vector in the denominator should be

  $$\begin{bmatrix} -\frac{3}{8} \\ \frac{3}{4} \end{bmatrix}$$

  (the second entry should be $\frac{3}{4}$ instead of $\frac{3}{8}$). [Printing 1–3]
- p. 137, 7th line from the bottom: Instead of $\boldsymbol{x}^{(k)}\boldsymbol{Q}$, it should be $\boldsymbol{Q}\boldsymbol{x}^{(k)}$. [Printing 1–3]
- p. 138, line 6: For the subscript $\boldsymbol{x} \in \mathcal{V}_k$ in $\min_{\boldsymbol{x}\in\mathcal{V}_k}$, the symbol $\mathcal{V}$ should be the same as the symbol in the equation $\mathcal{V}_k = \boldsymbol{x}^{(0)} + \cdots$ that immediately follows.

  The same holds in the equation on line 7 of p. 139. [Printing 1–3]
- p. 138, 3rd displayed equation: The item inside the summation should be $\alpha_i \boldsymbol{d}^{(i)}$ (not $\alpha_k \boldsymbol{d}^{(k)}$). In other words, the second term on the right hand side should be

  $$\sum_{i=1}^{k} \alpha_i \boldsymbol{d}^{(i)}$$

  [Printing 1–3]

- p. 140, step 6 in the conjugate gradient algorithm: In the right hand side of the equation for β_k, the symbol $\boldsymbol{g}^{(k+1)}$ should be italicized. [Printing 1–3]

- p. 148, Proposition 11.1, last line: Change α^k to α_k (k should be a subscript). [Printing 1–current]

- p. 149, 10 lines from bottom: Change "$[\Delta\boldsymbol{x}^{(0)}, \Delta\boldsymbol{g}^{(1)}, \ldots$" to "$[\Delta\boldsymbol{x}^{(0)}, \Delta\boldsymbol{x}^{(1)}, \ldots$" (the second component should be $\Delta\boldsymbol{x}^{(1)}$). [Printing 1–current]

- p. 150, 3 lines above Theorem 11.1: replace "$alpha_k$" with "$alpha_i$" (the subscript should be i, not k). [Printing 1–current]

- p. 150: In Theorem 11.1,

  - In the second line, the range of k should be $0 \leq k < n-1$.
  - In the last line, the range of i should be $0 \leq i \leq k+1$.
  - In the last line, the superscript of the last $\boldsymbol{d}$ should be $k+1$; i.e., $\boldsymbol{d}^{(k+1)}$.

  The corrected theorem reads:

  > Consider a quasi-Newton algorithm applied to a quadratic function with Hessian $\boldsymbol{Q} = \boldsymbol{Q}^T$, such that for $0 \leq k < n-1$,
  >
  > $$\boldsymbol{H}_{k+1}\Delta\boldsymbol{g}^{(i)} = \Delta\boldsymbol{x}^{(i)}, \qquad 0 \leq i \leq k,$$
  >
  > where $\boldsymbol{H}_{k+1} = \boldsymbol{H}_{k+1}^T$. If $\boldsymbol{d}^{(i)} \neq \boldsymbol{0}$, $0 \leq i \leq k+1$, then $\boldsymbol{d}^{(0)}, \ldots, \boldsymbol{d}^{(k+1)}$ are $\boldsymbol{Q}$-conjugate.

  [Printing 1–3]

- p. 151, 9th line from the bottom: Instead of α_k, it should be a_k. [Printing 1–3]

- p. 151 of the 4th printing, 9th line from the bottom: there should be an open parenthesis "(" before $\boldsymbol{H}_k$. [Printing 4]

- p. 159, line 15: After "Lemma 10.2", add "(see also Exercise 11.1)". [Printing 1–3]

- p. 159, 4th line from bottom: Replace $(\boldsymbol{x}^T\Delta\boldsymbol{x}^{(k)})^2/\alpha_k\boldsymbol{g}^{(k)T}\boldsymbol{H}_k\boldsymbol{g}^{(k)}$ with $(\boldsymbol{x}^T\Delta\boldsymbol{x}^{(k)})^2/(\alpha_k\boldsymbol{g}^{(k)T}\boldsymbol{H}_k\boldsymbol{g}^{(k)})$ (the parentheses in the denominator are missing). [Printing 1–current]

- p. 169, last line: Change "equaion" to "equation" (spelling error). [Printing 1–current]

- p. 178, line 14: "$\|\boldsymbol{x} - \boldsymbol{x}^*\| > 0$" should be "$\|\boldsymbol{x} - \boldsymbol{x}^*\|^2 > 0$". [Printing 1–3]

- p. 180: Change every occurrence of $R(k+1)$ in the subscripts to $R(k)+1$ (20 occurrences). [Printing 1–3]

- p. 180: Change every occurrence of $R(i+1)$ in the subscripts to $R(i)+1$ (2 occurrences). [Printing 1–current]

- p. 181, line 8: Place a close-parenthesis ")" after "[see Ref. 3, p. 70]". [Printing 1–3]
- p. 182, first 8 lines: Change every occurrence of $R(k)$ in the subscripts to $R(k)+1$ (8 occurrences). [Printing 1–current]
- p. 182, lines 4, 5, 7, and 8: Change every occurrence of $\boldsymbol{x}^{(k)}$ to $\boldsymbol{x}^{(k+1)}$ (4 occurrences). [Printing 1–current]
- p. 183, line 2, and p. 193, Exercise 12.12: Replace $\boldsymbol{a}_{R(k)}^T(\boldsymbol{x}^{(k)}-\boldsymbol{x}^*)=0$ with $\boldsymbol{a}_{R(k)+1}^T(\boldsymbol{x}^{(k+1)}-\boldsymbol{x}^*)=0$. [Printing 1–current]
- p. 187, the 4×3 matrix in the middle of the page: The element $(4,3)$ should be -72, not -73. [Printing 1–current]
- p. 187, last two equations at the bottom: The equations should be changed to:

$$\begin{aligned} \boldsymbol{B}^\dagger &= \boldsymbol{B}^T(\boldsymbol{B}\boldsymbol{B}^T)^{-1} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \\ \boldsymbol{C}^\dagger &= (\boldsymbol{C}^T\boldsymbol{C})^{-1}\boldsymbol{C}^T = [\,1/2 \quad 1/2\,]. \end{aligned}$$

  It is also useful to add the following paranthetical sentence after the above equations: (Note that the formulas for $\boldsymbol{B}^\dagger$ and $\boldsymbol{C}^\dagger$ here are different from those in Example 12.7 because of the dimensions of $\boldsymbol{B}$ and $\boldsymbol{C}$ in this example.) [Printing 1–current]
- p. 203, Figure 13.7: At the top of the figure, change "hidden neuron" to "output neuron". [Printing 1–current]
- p. 212: The chapter number ("14") is missing at the top of the page above the chapter title. [Printing 1–3]
- p. 213: Section 4.1.2, line 4: Change $P^{(k)}$ to $P(k)$. [Printing 1–current]
- p. 216, first equation in Example 14.3: The correct equation should be

$$f(x,y) = 3(1-x)^2\, e^{-x^2-(y+1)^2} - 10\left(\frac{x}{5} - x^3 - y^5\right)\, e^{-x^2-y^2} - \frac{e^{-(x+1)^2-y^2}}{3}.$$

  [Printing 1–3]
- p. 219, line 2: Replace "scheme" with "schema" (typo). [Printing 1–current]
- p. 235, line 7: $x_3 = 3$ should be $x_2 = 3$ (subscript should be "2"). [Printing 1–3]
- p. 235, 3 lines from bottom: Change "in some cases" to "In some cases" (capitalize the "i"). [Printing 1–current]
- p. 238, Figure 19.3: Remove all instances of "$h_1 =$" and "$h_2 =$" in the figure. (This is because some of the subscripts in the hs are mixed up in the figure.) [Printing 1–current]

- p. 248, last displayed equation (bottom of page): The second component of the first vector should be $-11/5$, not $-1/5$. [Printing 1–3]

- p. 249, last equation, bottom of page: There should not be a bar between n and m in $\binom{n}{m}$. [Printing 1–3]

- p. 250, line 3: Replace $\begin{pmatrix} 50 \\ 4 \end{pmatrix}$ with $\begin{pmatrix} 50 \\ 5 \end{pmatrix}$. [Printing 1–current]

- p. 254, Figure 15.10: In the label for the solid line (level set), change the number "30" to "15" (the label should read $f = 3x_1 + 5x_2 = 15$). [Printing 1–current]

- p. 255, last line of Exercise 15.2: Remove the phrase, ", and write it in standard form." Add the question at the end, "What can you say about the existence of a solution to this problem?" [Printing 1–3]

- p. 261: Third line after the proof, the I in $[\boldsymbol{A}, I]$ should be boldfaced. [Printing 1–3]

- p. 270, line 5 (first equation): The term on the left hand side should be y'_{ij} (not y'_{iq}). [Printing 1–3]

- p. 270, 5 lines from bottom: "... solution on $\boldsymbol{Ax} = \boldsymbol{b}$ ..." should be "... solution to $\boldsymbol{Ax} = \boldsymbol{b}$ ..." (change "on" to "to"). [Printing 1–3]

- p. 275, 6 lines from bottom: change "tabuleau" to "tableau" (spelling error). [Printing 1–current]

- p. 275, 5 lines from bottom: add the word "function" before the word "objective" (i.e., "... the objective function ..."). [Printing 1–current]

- p. 276, 6 lines from bottom: change "tabuleau" to "tableau" (spelling error). [Printing 1–current]

- p. 277, line 7: change "tabuleau" to "tableau" (spelling error). [Printing 1–current]

- p. 277, displayed matrix, middle of page: Change the entries $\frac{8}{14}$ and $\frac{44}{14}$ to $\frac{4}{7}$ and $\frac{22}{7}$, respectively (not really a typo, just an enhancement). [Printing 1–3]

- p. 277, 6 lines from bottom: Replace "Example 16.7" with "Exercise 16.7". [Printing 1–3]

- p. 278, line 9 (first displayed equation): The space between "$\boldsymbol{A}$," and "$\boldsymbol{I}_m$" seems to be unnecessarily large. [Printing 1–3]

- p. 287, Exercise 16.5, 4th line: Insert the word "is" before $\boldsymbol{r}^T$. [Printing 1–3]

- p. 288, Exercise 16.7, 7th line: Change "=" to "≥" (i.e., it should read "$x_1, \ldots, x_7 \geq 0$"). [Printing 1–current]

- p. 293, 14 lines from bottom: Change "$\lambda_m a_{1n}+$" to "$\lambda_1 a_{1n}+$" (subscript m should be 1). [Printing 1–current]

- p. 294, line 3: Change "$\boldsymbol{c}^T\boldsymbol{x}$" to "$[\boldsymbol{c}^T, \boldsymbol{0}^T]\boldsymbol{x}$" (the "zero" components are missing). [Printing 1–current]

- p. 314, line 11, "... and a matrix $A' \in \mathbb{R}^{m\times(n+1)}$ such that ...": the matrix A' should be boldfaced; i.e., $\boldsymbol{A}'$. [Printing 1–current]

- p. 318, line 5: Change $\boldsymbol{I}_{m+1}$ to $\boldsymbol{I}_n$ (change the subscript from $m+1$ to n). [Printing 1–current]

- p. 318, line 13, "... In fact, $\boldsymbol{x}^{(0)}$ ...": Change the superscript (0) to (1). [Printing 1–current]

- p. 319, line 16: Change $\boldsymbol{I}_{m+1}$ to $\boldsymbol{I}_n$ (change the subscript from $m+1$ to n). [Printing 1–current]

- p. 320, line 6: Change $\boldsymbol{I}_{m+1}$ to $\boldsymbol{I}_n$ (change the subscript from $m+1$ to n). [Printing 1–current]

- p. 325, Definition 19.1: Replace "$\boldsymbol{g} \leq \boldsymbol{0}$" by "$\boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0}$" (the $\boldsymbol{x}$ is missing). [Printing 1–3]

- p. 327, 4 lines from bottom: Put a comma between "1" and "$-2x_3$"; i.e., should be, "$[0, 1, -2x_3]^T$". [Printing 1–3]

- p. 330, Definition 19.5, 2nd line: Change "$\boldsymbol{h}(\boldsymbol{x}^*) = \boldsymbol{0}$" to "$\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}$" (remove the superscript $*$). [Printing 1–current]

- p. 330, second paragraph following Definition 19.5, 2nd line: Replace "tht" by "that" (typo). [Printing 1–3]

- p. 330, Definition 19.6, 2nd line: Change "$\boldsymbol{h}(\boldsymbol{x}^*) = \boldsymbol{0}$" to "$\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}$" (remove the superscript $*$). [Printing 1–current]

- p. 336, 2 lines above Theorem 19.2: Change "observation allow" to "observations allow" (add "s" to the end of "observation"). [Printing 1–current]

- p. 346: In the first sentence after the proof of Theorem 19.5, the word "regular" is unnecessary. [Printing 1–3]

- p. 347, line 14: Change $\{\boldsymbol{x} : 1 - \boldsymbol{x}^T\boldsymbol{P}\boldsymbol{x}\}$ to $\{\boldsymbol{x} : 1 - \boldsymbol{x}^T\boldsymbol{P}\boldsymbol{x} = 0\}$ (add "$= 0$"). [Printing 1–current]

- p. 352, 3rd line of Theorem 20.1: Remove the colon after $\boldsymbol{\lambda}^*$. It should read, "... there exist $\boldsymbol{\lambda}^* \in \mathbb{R}^m$ ..." [Printing 1–3]

- p. 353, 3rd line in the Proof of Karush-Kuhn-Tucker Theorem: Change $\{x : \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}...\}$ to $\{\boldsymbol{x} : \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}...\}$ (the symbol "$\boldsymbol{x}$" should be bold). [Printing 1–current]

- p. 354, line 13: Change $\dot{x}(t^*)$ to $\dot{\boldsymbol{x}}(t^*)$ (the symbol "$\boldsymbol{x}$" should be bold). [Printing 1–current]

- p. 359, second last line: Should be $\boldsymbol{\mu}^* \geq \mathbf{0}$ (the superscript $*$ was missing). [Printing 1–3]

- p. 360, lines 2 and 3: Change "$\{\boldsymbol{x} : \boldsymbol{h}(\boldsymbol{x}^*) = \mathbf{0}, \boldsymbol{g}(\boldsymbol{x}^*) \leq \mathbf{0}\}$" to "$\{\boldsymbol{x} : \boldsymbol{h}(\boldsymbol{x}) = \mathbf{0}, \boldsymbol{g}(\boldsymbol{x}) \leq \mathbf{0}\}$" [Printing 1–current]

- p. 360, line 3: Change "$\{\boldsymbol{x} : \boldsymbol{h}(\boldsymbol{x}^*) = \mathbf{0},\ \boldsymbol{g}(\boldsymbol{x}^*) = \mathbf{0},\ j \in J(\boldsymbol{x}^*)\}$" to "$\{\boldsymbol{x} : \boldsymbol{h}(\boldsymbol{x}) = \mathbf{0},\ \boldsymbol{g}(\boldsymbol{x}) = \mathbf{0},\ j \in J(\boldsymbol{x}^*)\}$" [Printing 1–current]

- p. 360, four lines above Theorem 20.3: Change "$\tilde{J}(\boldsymbol{x}, \boldsymbol{\mu}^*)$" to "$\tilde{J}(\boldsymbol{x}^*, \boldsymbol{\mu}^*)$" (add superscript $*$ to $\boldsymbol{x}$). [Printing 1–current]

- p. 360, second line of Theorem 20.3: Change "exists" to "exist" (remove "s"). [Printing 1–current]

- p. 360, second line of Theorem 20.3: Add "*feasible*" after the third word; i.e., the second line should read, "*there exist a feasible point ...*" [Printing 1–3]

- p. 368, first line of Definition 21.1: Change "$\Omega \to \mathbb{R}^n$" to "$\Omega \subset \mathbb{R}^n$" [Printing 1–current]

- p. 375, Theorem 21.4: Remove the items in brackets ("(strictly convex)" and "(positive definite)"). Also remove the last sentence in the proof. The condition $\boldsymbol{F}(\boldsymbol{x}) > 0$ is only sufficient but not necessary for strict convexity. [Printing 1–current]

- p. 377, second line of Theorem 21.5: Add "a" before "local" (missing article). [Printing 1–current]

- p. 381, 11 lines from bottom: Change "$\boldsymbol{\mu}^{*T} \geq \mathbf{0}$" to "$\boldsymbol{\mu}^{*T} \geq \mathbf{0}^T$" (missing transpose). [Printing 1–current]

- p. 382, Exercise 21.5, second displayed equation: $\boldsymbol{\mu}_j^*$ should not be boldfaced. [Printing 1–current]

- p. 395, Exercise 22.4, part b: $\boldsymbol{S}$ should have square-roots of eigenvalue terms. In other words,

$$\boldsymbol{S} = \operatorname{diag}\left(\sqrt{\lambda_1(\boldsymbol{A}\boldsymbol{A}^T)}, \ldots, \sqrt{\lambda_m(\boldsymbol{A}\boldsymbol{A}^T)}\right)$$

  is a diagonal matrix with diagonal elements that are the *square-roots* of the eigenvalues of $\boldsymbol{A}\boldsymbol{A}^T$. [Printing 1–3]

- p. 397, reference [47]. Replace "Nautral" with "Natural" [Printing 1–current]

- p. 398, reference [59]. Replace "Uhr" with "Uhl" [Printing 1–current]

- p. 408, index entry for "Step size": Replace "10" with "102" [Printing 1–current]

Other Changes (all printings)

- p. 103: The notation arg min, while standard, may be unfamiliar to some. Therefore, the following paragraph should be added in Section 1.2 (p. 5), just before the last paragraph:

  > Given a real valued function f, the notation arg min denotes the argument that minimizes the function (a point in the domain of f), assuming such a point is unique. For example, if $f : \mathbb{R} \to \mathbb{R}$ is given by $f(x) = (x+1)^2 + 3$, then $\arg\min f(x) = -1$. If we write $\arg\min_{x\in\Omega}$, then we treat Ω as the domain of f. For example, for the function f above, $\arg\min_{x\geq 0} f(x) = 0$. In general, we can think of $\arg\min_{\boldsymbol{x}\in\Omega}$ as the solution (assuming it exists and is unique) of the optimization problem
  >
  > $$\begin{array}{ll} \text{minimize} & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{x} \in \Omega. \end{array}$$

  An entry for "arg min" should also be added to the index.

- p. 112: Just after the proof of Lemma 8.1, add the following remarks:

  > Note that $\gamma_k \leq 1$ for all k, since $\gamma_k = 1 - V(\boldsymbol{x}^{(k+1)})/V(\boldsymbol{x}^{(k)})$ and V is a nonnegative function. If $\gamma_k = 1$ for some k, then $V(\boldsymbol{x}^{(k+1)}) = 0$, which implies that $\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^*$. In this case, we also have that $\boldsymbol{x}^{(i)} = \boldsymbol{x}^*$ and $\gamma_i = 1$ for all $i \geq k+1$.

- p. 373, Theorem 21.3: The assumption that Ω be open is not necessary, as long as $f \in \mathcal{C}^1$ on some open set that contains Ω (e.g., $f \in \mathcal{C}^1$ on $\mathbb{R}^n$).

- p. 375, Theorem 21.4: The theorem can be strengthened to include nonopen sets by modifying the condition to be $(\boldsymbol{y}-\boldsymbol{x})^T \boldsymbol{F}(\boldsymbol{x})(\boldsymbol{y}-\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x}, \boldsymbol{y} \in \Omega$ (and assuming $f \in \mathcal{C}^2$ on some open set that contains Ω; e.g., $f \in \mathcal{C}^2$ on $\mathbb{R}^n$).

May 1, 2001