

- p.18 Section I.3, Line 7 (just above the box) should refer to Section IV.6
- p.19 Section I.3, Statement 3 (eighth line from bottom, change sign from ' \leq ' to ' \geq '): 'Then $\text{rank}(A^T) \geq \text{rank}(A^T A) = \text{rank}(A)$.'
- p.20 Problem Set I.3, Problem 5: New second line in box (1 solution is impossible): ' $r = m < n$ $A_2 x = b$ has ∞ **solutions for every b** .'
- p.27 Section I.4, Problem 2: Delete last sentence 'Then... unique.'
- p.38 Section I.6, Last sentence of Example 3: Replace 'Increase 3 to 30 for complex λ 's.' by 'Complex eigenvalues will appear if 3 is decreased to -1 . Real matrices can easily have complex eigenvalues: for example antisymmetric matrices.'
- p.52 Section I.7, Problem 3: Last line of the problem, please replace A by S .
- p.55 Section I.7, Problem 28, Part (c): λ_1 , not λ
- p.57 Section I.8, text line 3: Replace words in the bracket by '(multiplying singular values gives | determinant of A |)'
- p.69 Section I.8, Problem 9: First line, remove $1/2$ both times. End of second line '... exactly $2Sx = 2\lambda x$.'
- p.70 Section I.8, Problem 19: The matrix with orthogonal columns of lengths 2, 3, 4 should be A and not S .
- p.73 Section I.9, Subsection 'The Frobenius Norm': In the beginning, ignore the words 'Eckart-YoungFrobenius Norm'
- p.75 Section I.9, Equation (19): Change 2nd partial derivative from $\frac{\partial E}{\partial R}$ to $\left(\frac{\partial E}{\partial R}\right)^T$
- p.75 Section I.9: See the book's website for Dan Drucker's computations!
- p.76 Section I.9: 11th line from the bottom, $A = \begin{bmatrix} 3 & -4 & 7 & 1 & -4 & -3 \\ 7 & -6 & 8 & -1 & -1 & -7 \end{bmatrix}$
- p.82 Section I.10: Last line in box, change ' $\lambda M x_1$ ' to ' $\lambda_1 M x_1$ '

- p.95 Section I.11 : Example 1, Matrix changes from $\begin{bmatrix} 5 & 15 \\ 15 & 45 \end{bmatrix}$ to $\begin{bmatrix} 10 & 20 \\ 20 & 40 \end{bmatrix}$
- p.95 Section I.11 : Example 1, Matrix changes from $\begin{bmatrix} 5 & 5 \\ 5 & 5 \end{bmatrix}$ to $\begin{bmatrix} 2 & 4 \\ 4 & 8 \end{bmatrix}$
- p.99 Section I.12 : Line 4, (columns of B).
- p.107 Section I.12 : Line 9, the word is ‘frequently’, not ‘frequency’.
- p.120 Section II.1 : Line 12 from the bottom, reference is (7) *not* (8).
- p.131 Section II.2 : Example, matrix changes from $\begin{bmatrix} 4 & x \\ 3 & x \end{bmatrix}$ to $\begin{bmatrix} 4 & * \\ 3 & * \end{bmatrix}$
- p.131 Section II.2 : Example, matrix changes from $\begin{bmatrix} 5 & x \\ 0 & x \end{bmatrix}$ to $\begin{bmatrix} 5 & * \\ 0 & * \end{bmatrix}$
- p.135 Section II.2, Problem 9 line 2 : Change $\mathbf{a}^T \mathbf{q}_1$ to $\mathbf{b}^T \mathbf{q}_1$
- p.135 Section II.2, Problem 11 line 1 : Exchange the $+$ and the T to get $QQ^T = AA^+$.
Note : Those are the same projections onto $\mathbf{C}(A)$
- p.162 Section III.1, first line : The first row of matrix M^{-1} should read 1 1 2
- p.182 Third reference : Change from ‘N’ to ‘N. Truhar’
- p.187 Two lines above equation 12 : Change bold ‘We are now maximizing...’ to ‘We are now minimizing...’
- p.187 Equation 12 : Two instances of $B\mathbf{y}$ are actually $B\mathbf{z}$
- p.188 Equation 17, $g(\mathbf{x})$ is actually $g(\mathbf{z})$
- p.189 Middle of the page, three equations labeled ADMM : Change u_{k+1} to u_k
- p.192 Equation 31 : Change $S(\mathbf{y}_{k+1})$ to $S_\lambda(\mathbf{y}_{k+1})$
- p.199 Line below Equation (11) : Change reference ‘... by using A from equation (11).’

- p.214 First line of ‘Practice’: Change ‘ $(0, 1, 0) \otimes (d_0, d_1, d_2) = (d_1, d_2, d_0)$ ’ to ‘ $(0, 1, 0) \otimes (d_0, d_1, d_2) = (d_2, d_0, d_1)$ ’
- p.216 Section IV.2, Equation 11: Correct matrix is
$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & i^2 & i^3 \\ 1 & i^2 & i^4 & i^6 \\ 1 & i^3 & i^6 & i^9 \end{bmatrix}$$
- p.238 Section IV.5, Problem 5: ‘The symbol 5’ (not ‘The symbol S ’)
- p.241 Third line of first paragraph: Change ‘ $n - m + 1$ ’ to ‘ $m - n + 1$ ’
- p.245 Section IV.7, last printed line: Replace by ‘For 5 nodes and 2 clusters, C has only two columns (the centroids of the clusters) in $A \approx CR$.’
- p.265 Line after Equation (4): Change $(x_N - m^2)$ to $(x_N - m)^2$
- p.270 Last line in second box: Change ‘ $(A_N - \frac{1}{2}) / 2\sqrt{N}$ ’ to ‘ $(A_N - \frac{1}{2}) \times 2\sqrt{N}$ ’
- p.278 Equation 12: Change ‘ $e^{-\lambda/\beta}$ ’ to ‘ $e^{-x/\beta}$ ’
- p.280 Equation 18: Change $e^{-(\mathbf{x}-\boldsymbol{\mu})^T C^{-1}(\mathbf{x}-\boldsymbol{\mu})}$ to $e^{-(\mathbf{x}-\boldsymbol{\mu})^T C^{-1}(\mathbf{x}-\boldsymbol{\mu})/2}$
- p.286 Line 15 earlier ended with ‘ μ_1, μ_5, \dots ’. For better understanding, ‘ μ_1, μ_3, \dots ’
- p.291 Correct ‘Joel Tropp’ reference is ‘arXiv:1501.01571’
- p.297 Section V.4, Equation (10): Modified,

$$\mathbf{V} = \mathbf{E} \left[(\mathbf{X} - \overline{\mathbf{X}}) (\mathbf{X} - \overline{\mathbf{X}})^T \right] \quad \mathbf{V}_{ij} = \sum \mathbf{p}_{ij} (\mathbf{X}_i - \mathbf{m}_i) (\mathbf{X}_j - \mathbf{m}_j)$$
- p.300 Last paragraph, first line: Replace ρ_{xy} by σ_{xy} ; second line: Replace σ_{xy} by ρ_{xy}
- p.346 Example 3, second line: Change ‘ $\mathbf{F}_{\min} = -\frac{1}{2} \mathbf{a}^T \mathbf{S} \mathbf{a}$ ’ to ‘ $\mathbf{F}_{\min} = -\frac{1}{2} \mathbf{a}^T \mathbf{S}^{-1} \mathbf{a}$ ’
- p.357 Last line of figure: add absolute value sign, ‘ $S(x_i) = \text{sign}(x_i) \cdot \max(|x_i| - c, 0)$ ’
- p. 395 Section VII.2, line 14: Replace ‘It is hard to predict whether deep ConvNets will be replaced by ResNets.’ with ‘By sending information far forward, features that are learned early don’t get lost before the output. Residual networks have become highly successful deep networks.’
- p. 417 Line 6 from the bottom: ‘The diagonalization (3) breaks down...’