ERRATA

INFERENCE AND LEARNING FROM DATA

by

Ali H. Sayed

École Polytechnique Fédérale de Lausanne (EPFL), Switzerland
University of California at Los Angeles (UCLA), USA

A. H. Sayed, *Inference and Learning from Data*, 3 vols.,
Cambridge University Press, 2022.

VOL I: Foundations
VOL II: Inference
VOL III: Learning

Readers are welcome to bring to the attention of the author any typos or suggestions for improvement. Please feel free to email the author directly at ali.sayed@epfl.ch or sayed@ucla.edu.

The author is thankful for all feedback.

Last updated: March 22, 2025.

ERRATA

Inference and Learning from Data, Cambridge University Press, 2022.

Last updated March 22, 2025

1. **Notation**, p. xlvii, $\mathcal{S}(\theta)$: score function is $\nabla_\theta^{\mathsf{T}} \ln f_{\boldsymbol{x}}(x;\theta)$. Logarithm is missing.
2. **Chapter 1**, p. 50, Prob. 1.58: all norms should be squared and replaced by $\|\cdot\|_2^2$.
3. **Chapter 2**, p. 64, Table 2.1, property 7: replace A^{T} by A on RHS.
4. **Chapter 2**, p. 67, Prob. 2.11, item (c): replace A^{T} by A on the RHS.
5. **Chapter 3**, p. 117, Prob. 3.54: write $(b-a)^2$ in last bound; square is missing.
6. **Chapter 6**, p. 226, paragraph before (6.149), change to $x = r\cos\theta$.
7. **Chapter 8**, p. 270, Eq. (8.37): a multiplying factor γ^2 is missing from the second term on the right-hand side with hh^{T}."
8. **Chapter 10**, p. 332, line before (10.13), change to "upper bound on any function $g(z)$ with a convex $\mathrm{dom}(g)$."
9. **Chapter 11**, p. 342, Eq. (11.8): replace $\mathbb{I}_{C,\infty}(w)$ by $\mathbb{I}_{C,\infty}[w]$ with brackets.
10. **Chapter 11**, p. 343, rephrase sentence right before (11.15) to "function $h(w)$ is nondifferentiable at several locations, namely, whenever at least one entry of w is 0. The function $h_p(w)$ becomes"
11. **Chapter 12**, p. 380, rephrase sentence right before (12.11) and in the paragraph after it to "not differentiable when at least one entry of w is 0 for"
12. **Chapter 14**, pp. 498–499, Probs. 14.4 and 14.5: write $P(w_n^{\mathrm{best}})$ instead of $P^{\mathrm{best}}(n)$.
13. **Chapter 18**, p. 645, Eq. (18.10b): write h_ℓ instead of $h(\ell)$.
14. **Chapter 19**, p. 721, Eq. (19.167), first equality, rightmost term: brackets $\{\cdot\}$ missing between the $\lim_{n\to\infty}$ and the $=$ sign, i.e., write $\lim_{n\to\infty}\{\boldsymbol{s}(n) - \sum_{j=n}^{\infty} b(j)\}$.
15. **Chapter 23**, p. 832, Listing (23.45): replace Q by Q_u with subscript u.
16. **Chapter 28**, p. 1104, Eq. (28.65): change to $(0.5 \times 0.4238)/0.2845 \approx 0.7448$. In the following sentence, change "less" to "larger" and "virginica" to "setosa".
17. **Chapter 33**, p. 1348, in Prob. 33.1, replace $\partial g^{-1}(y_m)$ by $\partial g_m^{-1}(y)$, where $g_m^{-1}(y)$ is the mth entry of $g^{-1}(y)$.
18. **Chapter 34**, p. 1372, rephrase sentence after (34.98) as "which can be normalized into an exponential pdf using". Also, rephrase after (34.100) as "where we also added the normalization factor. Using expression (34.98) prior to normalization:"
19. **Chapter 34**, pp. 1375–1377, Section 34.4 on ADF. Replace all superscripts $(\ell-1)$ and (ℓ) by $(n-1)$ and (n) since the iteration is over n and not ℓ in this section.
20. **Chapter 34**, p. 1377. In (34.123)–(34.124) after **end** command, replace (ℓ) by (N). Initialization is $q^{(0)}_{\boldsymbol{z}|\boldsymbol{y}}(z|y) = \pi(z)$, $\mu^{(0)} = \bar{z}$, $R^{(0)} = R_z$, $\lambda^{(0)} = \mathbb{E}_{\,\pi} T(z)$.
21. **Chapter 35**, p. 1387: rephrase sentence after (35.22) as "As was explained earlier before listing (33.88), if desired ..."
22. **Chapter 35**, p. 1388, paragraph after (35.26), fifth line. Add a sentence before the words "For now" as follows: "Due to the Markovian property, it is generally sufficient to choose the rightmost factor in the form $\pi_{\boldsymbol{x}_n|\boldsymbol{x}_{n-1},\boldsymbol{y}_n}(x_n|x_{n-1},y_n)$ with conditioning only on $\boldsymbol{x}_{n-1}$ and $\boldsymbol{y}_n$. For now, ..."
23. **Chapter 35**, p. 1389, Eq. (35.29), write $\pi(x_n^j|x_{0:n-1}^j, y_{0:n-1})$ in denominator of w_n^j. Also, in last line of Eq. (35.42) on page 1352.
24. **Chapter 35**, p. 1390, Eq. (35.30), add step 5. "Normalize the weights to add up to 1 and repeat the process."
25. **Chapter 35**, p. 1393: in listing (35.41), move the line "set $\{x_{0:n}^j\} \leftarrow \{x_{0:n}^{j\star}\}$, $j = 1, 2, \ldots, J$" before the **end** command above it. Also, set initial conditions to $\boldsymbol{x}_0^i \sim \pi(x|0,y_0)$ and $w_0^j = f(x_0^j)f(y_0|x_0^j)/\pi(x_0^j|0,y_0)$.
26. **Chapter 36**, p. 1445: second line after (36.185), replace $h(y_n, z_n)$ by $T(y_n, z_n)$.
27. **Chapter 37**, p. 1485: Eqs. (37.20) and (37.21), replace N by N_d.
28. **Chapter 39**, p. 1571: Eq. (39.28), replace $Y_{1:N}$ by $Y_{1:n}$ in the expression for $\alpha(k,n)$.
29. **Chapter 40**, p. 1631: Eq. (40.108), replace $=$ by $\leftarrow$ in the 3rd equation within the **for** loop over m.

30. **Chapter 40**, p. 1632: Eq. (40.111), 3rd line below **repeat**, replace $\boldsymbol{h}'_n$ by $\mathbb{1}_M \otimes (\boldsymbol{h}'_n)^{\mathsf{T}}$.
31. **Chapter 41**, p. 1674: Eq. (41.90), remove the fifth line about pathway 8.
32. **Chapter 42**, p. 1707: Eq. (42.84), write instead: "$\{\boldsymbol{x}_k, \boldsymbol{x}_\ell\}$ are independent of each other conditioned on the other entries in $\boldsymbol{x}$." The trailing part is missing.
33. **Chapter 49**, p. 2109: Prob. 49.5, part (c). Replace $d^{\pi}(s')$ by $d^{\pi}(s)$.
34. **Chapter 50**, p. 2204, Prob. 50.9: $w^{\star}_{\text{reg}} = (I_M + \rho Q^{-1})^{-1} w^{\star}$, with Q inverted.
35. **Chapter 52**, p. 2283, Prob. 52.5: $\|h - h_a\| \leq \|h - h_n\|$, with no square.
36. **Chapter 58**, p. 2450, Eq. (58.104): replace argmax by max in the definition of the coherence measure $\mu(A)$.
37. **Chapter 59**, p. 2489, Prob. 59.4: $\sigma(z) = \ln[1/(1 + e^{-z})]$. The log is missing.
38. **Chapter 61**, p. 2544, 2nd paragraph, 4th line: $\lambda^{\star}(n_2)$. The star is missing.
39. **Chapter 63**, p. 2642: Prob. 63.12, part (e), replace $s(\|h - h'\|)$ by $s(h - h')$ in 2 locations, and the text "is a function of the distance between the feature arguments" by "is translation invariant."
40. **Chapter 64**, p. 2667, Eq. (64.28): replace $h_n^{\mathsf{T}} w \neq \gamma(n)$ by $c(w) \neq \gamma(n)$.
41. **Chapter 64**, p. 2670, Fig. 64.11: replace $\max(0, y)$ by $\max(0, -y)$.
42. **Chapter 64**, p. 2684, Prob. 64.30: replace the statement about $\boldsymbol{h}(m)$ in the first paragraph by "Let $\boldsymbol{h}$ be a scalar random feature that can assume one of M possible discrete values denoted by $\{h(m)\}$." In part (a), the probability expression on the right-hand side should become $\mathbb{P}(\boldsymbol{h} = h(m) | \boldsymbol{\gamma} = \gamma)$.
43. **Chapter 65**, p. 2756, Eq. (65.133): $D_{\text{KL}}(p \| s)$, with $\|$ instead of comma.
44. **Chapter 68**, p. 2946, Eq. (68.136): replace $z(q)$ in numerator by $e^{z(q)}$.
45. **Chapter 68**, p. 2948, line before Eq. (68.142): the reference should be to Eq. (65.61).
46. **Chapter 72**, p. 3124, Eq. (72.80), replace the superscript $^{(t)}$ by $^{(e)}$ in the rightmost term inside the summation.