Data Science and Machine Learning:
Mathematical and Statistical Methods

Errata

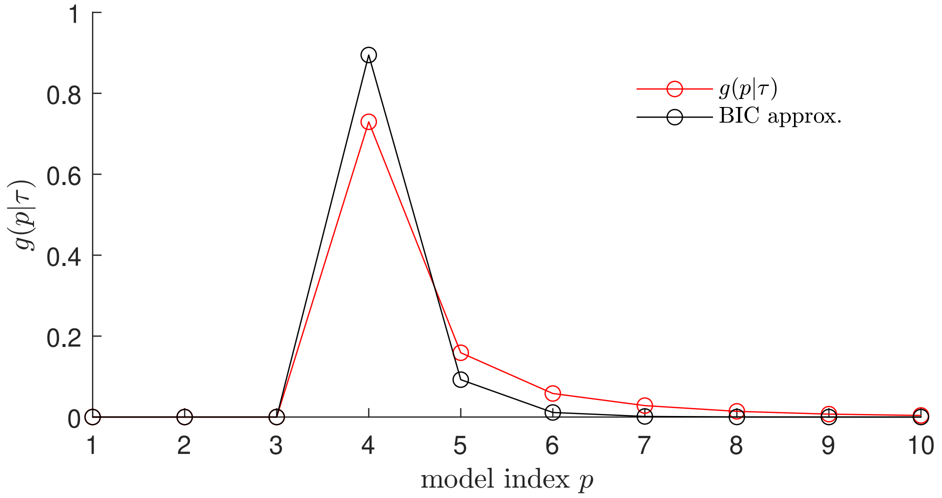
(Last Update 9th July 2021)

1. Page 37, line 3 from the top: replace $\mathbb{E}Y_i$ with $\mathbb{E}_{\mathbf{X}} Y_i$.

2. Page 38, lines 3,4 in second paragraph: replace $\ell_{\mathcal{T}_{-k}}$ symbol with ℓ_{C_k}.

3. Page 38, first line in displayed equation: replace $\ell_{\mathcal{T}_{-k}}$ symbol with $\ell_{C_k}(g_{\mathcal{T}_{-k}})$.

4. Page 57, Figure 2.16. There was a mistake in the drawing of the BIC approximation. The actual BIC approximation matches the posterior density quite well:



5. Page 72, Line −2: . . . in terms of the probability . . . (remove repeated "the").

6. Page 100, Line −8: $(1 - \alpha\boldsymbol{v})$ should be $(1 - \alpha)\boldsymbol{v}$.

7. Page 162: Line 12: $\boldsymbol{\Sigma}^{1/2}\boldsymbol{x}$ should be $\boldsymbol{\Sigma}^{-1/2}\boldsymbol{x}$.

8. Page 162: Lines 17 and 20: $\boldsymbol{\Sigma}^{1/2}(\boldsymbol{x}_i - \boldsymbol{\mu})$ should be $\boldsymbol{\Sigma}^{-1/2}(\boldsymbol{x}_i - \boldsymbol{\mu})$.

9. Page 178: fourth line below Table 5.1: replace "qualitative" with "quantitative".

10. Page 179, fourth line in Example 5.5: replace "row-wise" with "column-wise" and the vector $\boldsymbol{y}$ with $\boldsymbol{y} = [9.2988, 8.2111, 9.0688, 8.2552, 9.4978, \ldots, 8.9485]^\top$.

11. Page 181, formula for R^2_{adjusted} at the bottom: replace $n - p - 1$ in the formula with $n - p$.

12. Page 184, formula for F_i should have the norms squared:

$$F_i = \frac{\|\boldsymbol{Y}^{(i)} - \boldsymbol{Y}^{(i-1)}\|^2 / p_i}{\|\boldsymbol{Y} - \boldsymbol{Y}^{(d)}\|^2/(n-p)} \,.$$

13. Page 211, Exercise 12 (b): $\mathbf{P}_{ii}$ should be $(1 - \mathbf{P}_{ii})$; that is 1 minus the i-th leverage.

14. Page 221, Line 8: ... one obtains the so-called ...

15. Page 219, Line −2: ... only β_1 is regularized.

16. Page 247, Algorithm 6.8.1, Line 1: $\mathbb{R}^p$ should be $\mathbb{R}^n$.

17. Page 248, Algorithm 6.8.2, Line 1: Set $\mathbf{B} \leftarrow (n\gamma \mathbf{I}_p)^{-1}$.

18. Page 235, Line 7: $\int_0^1 (g''(x))^2 \, \mathrm{d}x$ instead of $\int_0^1 (g'')^2 \, \mathrm{d}x$.

19. Page 273, 3rd line under Figure 7.9: The results are summarized in Table 7.6.

20. Page 329, line 12 from below: change y_{i-k} to y_{i-k+1}.

21. Page 331, last displayed equation:

$$\frac{\partial C}{\partial \boldsymbol{b}_l} = \frac{\partial \boldsymbol{z}_l}{\partial \boldsymbol{b}_l} \frac{\partial C}{\partial \boldsymbol{z}_l} = \boldsymbol{\delta}_l, \quad l = 1, \ldots, L.$$

22. Page 335, Algorithm 9.4.2, Line 2: ... using $\frac{\partial C}{\partial \boldsymbol{g}} = 1$ ...

23. Page 340, second displayed line:

$$[p_0, p_1, p_2, p_3] = [1, 20, 20, 1].$$

24. Page 341, Line 3: Remove the line `S = RELU`.

25. Page 351, Exercise 7(b): In the displayed formula, $\mathbf{B}$ should be replaced with $\mathbf{B}^{-1}$.

26. Page 362, First sentence in paragraph above Theorem A.4: ... the matrix $\mathbf{P}$ projects any vector in $\mathcal{V}$ onto itself.

27. Page 362, Sentence above Theorem A.4: ... where $\mathbf{U}$ is not ...

28. Page 380, third line from below: change b_{i-k} to b_{i-k+1}.

29. Page 394, line 5: ... can be computed with the aid ... (missing "the")

30. Page 404, last two lines: replace H with $\mathbf{H}$.

31. Page 414, Section B.3.4: Replace ℓ with ℓ_τ.

32. Page 433, displayed equation in the proof of Theorem C.4: replace $|\mathbf{J}_{\mathbf{g}^{-1}}(z)|$ with $|\det(\mathbf{J}_{\mathbf{g}^{-1}}(z))|$.

33. Page 439, line 4: is equal to $\Gamma(\alpha)\lambda^{-\alpha}$ times ...

34. Page 442, 4th line from the bottom: $x \geq c$ should be $x > c$.

35. Page 445, halfway on the page: $|\mathrm{e}^{\mathrm{i}x} - 1| = \left|\int_0^x \mathrm{i}\,\mathrm{e}^{\mathrm{i}\theta} \, \mathrm{d}\theta\right| \leq \left|\int_0^x |\mathrm{i}\,\mathrm{e}^{\mathrm{i}\theta}| \, \mathrm{d}\theta\right| = |x|$.

36. Page 446, displayed equation below (C.37): $O(t/n)$ should be $o(t/n)$, and in the next displayed equation, $o(1)$ should be $o(1/n)$.

37. Page 448, line 2: $O(t^3/n^{3/2})$ should be $o(t^2/n)$.

38. Page 450, first displayed equation after (C.39): The $\mathbf{\Sigma}$ in the denominator should be $\mathbf{\Sigma}_n$.

39. Page 451: Delete "ln" after "An application . . . yields"

40. Page 451, line starting with "asymptotically negligible": Replace n with $-n$ in the exponent.

41. Page 456, Sentence under (C.47): Similar to the one-dimensional case ($d = 1$), replacing the factor $1/n$ with $1/(n-1)$ gives an unbiased estimator, called the *sample covariance matrix*.

42. Page 511, line 13 from above: 'expectation of'.