《机器学习方法 第二版》勘误表

2025年11月15日

1. 第2章，16页

式(2.22)上面第一行：\delta^2 → \sigma^2

式(2.24) 第三项：\epsilon^2 → \sigma^2

2. 第2章，17页

上数第三行：第3 项是固有误差（irreducible error）。→ 第3 项是固有误差（irreducible error）。另有$\mathbb{E} (\epsilon^2) = \sigma^2$成立。

3. 第8章，97页

下数第六行：无向边表示概率相关关系 → 无向边表示依存关系

4. 第12章，189页

式(12.3)：P(X_u \vert X_{N(i)} → P(X_u \vert X_{N(u)}

5. 第12章，188页

上数第七行：无向图的最大图上的 → 无向图的最大团上的

6. 第12章，188页，189页，191页，207页

188页上数第七行：概率相关关系 → 依存关系

188页下数第二行：随机变量之间的相关关系 → 随机变量之间的依存关系

188页下数第一行：而变量之间的相关关系 → 而变量之间的依存关系

189页上数第七行： 概率相关关系 → 依存关系

191页定义12.1第二行： 概率相关关系 → 依存关系

207页本章概要1 第二行： 概率相关关系 → 依存关系

7. 第12章，191页

191页定义12.1第一行：由无向图 → 可以由无向图

8. 第12章，193页

图12.5下数第一行：相连的模块之间存在直接的相关性 → 相连的模块之间存在依存关系

图12.5下数第二行：而不相连的模块之间仅有间接的相关关系。注意，模型仅刻画模块之间的相关性，无法确定模块之间的因果关系。→ 而不相连的模块是条件独立的（满足成对马尔可夫性）。注意，模型仅刻画相连模块之间的依存关系，无法确定相连模块之间的因果关系。

9. 第12章，189页

上数第四行：用无向图G = (V;E) 表示概率分布P(X) → 概率分布P(X)由无向图G = (V;E)表示

10. 第12章，191页

定义12.1下数第一行：表示随机变量之间的概率依存关系。→ 表示随机变量之间的直接概率依存关系。

11. 第12章，192页

定理12.2 概率无向图模型或马尔可夫随机场的联合概率分布$P(X)$可以表示为如下形式 → 联合概率分布$P(X)$关于无向图G构成概率无向图模型或马尔可夫随机场的充分必要条件是它可以表示为如下形式

12. 第19章，320页

上数第四行：因为只需要定义马尔可夫链，而不需要定义建议分布。→ 因为只需要定义马尔可夫链，包括其中易采样的建议分布。

式(19.36)上面第一行：并选择后验概率最大的模型。→ 利用后验概率分布进行推断。

13. 第19章，321页

式(19.37)：\begin{equation} p({\bm d})=\dis\int_{\cal M} {p({\bm d}\vert {\bm m}')p({\bm m}')d{\bm m}'} \end{equation} → \begin{equation} p({\bm d})=\dis\int_{\cal M} {p({\bm d}\vert {\bm m})p({\bm m})d{\bm m}} \end{equation} 注：将'去掉

14. 第19章，314页，315页，329页

第19.3.3节：连续状态马尔可夫链 → 离散时间马尔可夫过程

第19.3.3节下第一行：连续状态马尔可夫链 → 离散时间马尔可夫过程

第19.3.4节下第一行：连续状态马尔可夫链 → 离散时间马尔可夫过程

本章概要2: 连续状态马尔可夫链 → 离散时间马尔可夫过程

15. 第19章，308页

下数第二行： 随机变量可以是离散的，也可以是连续的。→ 随机变量是离散的。

16. 第19章，309页

式(19.8)下数第一行：或马尔可夫过程（Markov process）→ 或离散时间马尔可夫链（discrete time Markov chain）

17. 第19章，309页

19.3.2节标题：离散状态马尔可夫链 → 离散时间马尔可夫链

19.3.2节第一行：离散状态马尔可夫链 → 离散时间马尔可夫链

18. 第19章，314页

314页19.3.3节第二行最后：（加上） → 离散时间马尔可夫过程是比离散时间马尔可夫链更一般的随机过程。

19. 第19章，314页，315页

314页式(19.23)上数第一行：马尔可夫链 → 马尔可夫过程

315页上数第一行：马尔可夫链 → 马尔可夫过程

20. 第19章，315页

19.3.4节第一行：以下介绍离散状态马尔可夫链的性质，可以自然推广到连续状态马尔可夫链。→ 以下介绍离散时间马尔可夫链的性质，在离散时间马尔可夫过程中有对应的推广形式。

21. 第19章，321页

19.6.1节下数第三行：构造的马尔可夫链 → 实际构造更一般的离散时间马尔可夫过程

22. 第19章，329页

本章概要2，第三行：有离散状态马尔可夫链和连续状态马尔可夫链 → 有离散时间马尔可夫链和更一般的离散时间马尔可夫过程

本章概要2，第四行：称为马尔可夫链的平稳分布 → 称为马尔可夫链或马尔可夫过程的平稳分布

23. 第1-4册，目录

连续状态马尔可夫链 → 离散时间马尔可夫过程

24. 第24章，388页，390页

388页下数第十二行、下数十三行：词符 → 词元

390页下数第七行：词符 → 词元

25. 第24章，388页，390页，397页

388页下数第十行：音符 → 听觉词元

390页下数第七行：音符 → 听觉词元

397页上数第一行：音符 → 听觉词元

26. 第24章，388页，390页，397页

388页下数第十一行：图符 → 视觉词元

390页下数第七行：图符 → 视觉词元

397页上数第一行：图符 → 视觉词元

27. 第24章，388页

388页下数第十行：audio → auditory

28. 第32章，561页

第三段第四行：从而 → 最后

29. 第32章，562页

第一段第二行：起初漂浮在水面 → 起初从水面向水中渗透

30. 第32章，562页，563页

562页下数第一段第一行：马尔可夫链 → 马尔可夫过程

562页下数第一段第二行：马尔可夫链 → 马尔可夫过程

563页上数第一段第一行：马尔可夫链 → 马尔可夫过程

563页上数第一段第二行：马尔可夫链 → 马尔可夫过程

31. 第32章，565页

定理32.1上数第一段：以下定理说明反向过程的状态转移分布也是高斯分布。→ （删除）

定理32.1：→ （删除）

32. 第32章，567页

上数第一段第一行：马尔可夫链 → 马尔可夫过程

上数第一段第二行：在此基础上可以得出，反向过程的转移概率分布是高斯分布，并且能够求出 → 在此基础上能够求出

上数第三段第一行：由定理32.4知，这个转移概率分布 → 假设这个转移概率分布

33. 第32章，570页

式(32.20)：\epsilon_1 → \epsilon

34. 第32章，571页

下数第二段第一行：两者存在一一对应关系 → 两者存在归一化条件下的一一对应关系

35. 第32章，573页

第四段第一行：有相似之处，对原始数据随机添加少量噪声，使用加噪数据学习分数匹配模型。这样可以学到对噪声更强健的模型，同时也可以解决分数匹配方法计算效率不高的问题。 → 有相似之处。对原始数据随机添加少量噪声，学习加噪数据的分数函数，用以估计原始数据的分数函数，同时解决分数匹配方法计算效率不高的问题。

36. 第32章，573页

式(32.28)右侧第二项：\frac{\delta}{2} → \delta

式(32.28)右侧第三项：\sqrt{\delta} → \sqrt{2 \delta}

37. 第32章，574页

第二段：如果第三项中... 是一种随机梯度方法。→（删除整个段落）

38. 第32章，576页

算法32.4下数第四行右侧第二项：\frac{\delta}{2} → \delta

算法32.4下数第四行右侧第三项：\sqrt{\delta} → \sqrt{2 \delta}

39. 第32章，577页

引理32.3下数第一行：假设均值$\sqrt{\bar{\alpha}}_t {\bm x}_0$就等于后验期望 → 均值$\sqrt{\bar{\alpha}}_t {\bm x}_0$近似等于后验期望

40. 第32章，578页

32.3.2节第一段第二行：可以用基于连续时间的随机微分方程 → 可以用随机微分方程

41. 第32章，579页

第二段第一行：如果将时间离散化，就得到对应的随机差分方程，本章介绍的扩散模型的迭代公式实际都是随机差分方程。随机差分方程可以让我们更方便地计算SDE。→ 如果将时间离散化，进行迭代计算，就得到其数值计算方法，最基本的方法有欧拉-丸山法。本章介绍的扩散模型的迭代公式实际都是SDE的数值计算方法。

42. 第32章，579页，585页

579页式(32.45)： \frac{\beta_t}{2} → \beta_t

585页本章概要9第二个公式： \frac{\beta_t}{2} → \beta_t

43. 第32章，579页

式(32.42)下数第一行：可以将它表示为一个随机差分方程。→（删除）

式(32.42)下面第一行：时间差分 → 时间步长

式(32.42)下数第一行：随机差分方程就变成。→ 迭代公式就变成。

44. 第32章，579页，580页

定理32.2： 定理32.2 → 定理32.1

式(32.44)上数第一行：定理32.2 → 定理32.1

式(32.52)上数第一行：定理32.2 → 定理32.1

45. 第32章，579

式(32.45)上数第一行：对应的随机差分方程如下，→ 对应的迭代公式如下，

46. 第32章，580页

第二段第一行：迭代公式(32.47)可以表示为一个随机差分方程。当$\Delta t \rightarrow 0$时，随机差分方程就变成SDE。→ 迭代公式(32.47)，当$\Delta t \rightarrow 0$时，就变成SDE。

式(32.53)上数第一行：对应的随机差分方程如下，→ 对应的迭代公式如下，

47. 第32章，585页

本章概要7：\frac{\delta}{2} → \delta

本章概要7： \sqrt{\delta} → \sqrt{2 \delta}

48. 第39章，685页

上数第二行：
策略梯度定理有以下简洁的证明。→ 策略梯度定理有以下简洁的证明\footnote{该证明源自 OpenAI Spinning Up in Deep RL, Introduction to RL, Part 3。}。