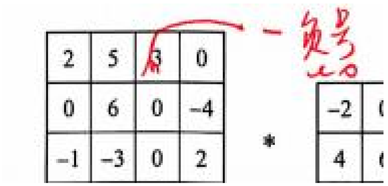
《深度学习：基础与概念》勘误表（1 版 3 印）

序号	章	页	行	误	正
1	1	12	图 1.10 图题	分别对应于 lnλ=-∞和 lnλ=0	分别对应于 lnλ=-18 和 lnλ=0
2	2	45	倒 1	如果一个函数仅在λ=0 和λ=1 时 KL(p‖q)=0 成立	如果一个函数仅在λ=0 和λ=1 时等号成立
3	2	46	倒 3	从 $p(\boldsymbol{x})$中产生的 $q(\boldsymbol{x} \mid \boldsymbol{\theta})$	从 $p(\boldsymbol{x})$中产生的 $\boldsymbol{x}_n$
4	2	47	倒 7	而将 $\boldsymbol{x}$ 视为观测到新数据 $\boldsymbol{y}$ 后的后验分布	而将 $p(\boldsymbol{x} \mid \boldsymbol{y})$视为观测到新数据 $\boldsymbol{y}$ 后的后验分布
5	3	67	倒 11	条件分布 $p(\boldsymbol{x}_a，\boldsymbol{x}_b)$也是高斯分布	条件分布 $p(\boldsymbol{x}_a \mid \boldsymbol{x}_b)$也是高斯分布
6	3	84	倒 5	最大似然估计量的解仅通过 $\sum_n \boldsymbol{\eta}_{\mathrm{ML}}$	最大似然估计量的解仅通过 $\boldsymbol{u}(\boldsymbol{x}_n)$
7	3	92	习题 3.9	多元高斯分布 $\mathcal{N}(\boldsymbol{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma})\partial$	多元高斯分布 $\mathcal{N}(\boldsymbol{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma})$

8	4	102	7	最小二乘的几何表示	最小二乘解的几何表示
9	4	102	8	考虑一个 n 维空间	考虑一个 N 维空间
10	4	103	11	这称为最小二乘（least-mean-square)法或 LMS 算法	这称为最小均方（Least-Mean-Square，LMS)算法
11	4	105	倒 10、11	$f(t \mid \boldsymbol{x})$	$f(\boldsymbol{x})$
12	5	131	倒 1	逻辑斯蒂 sigmoid 函数$\sigma(a)$的倒数为	逻辑斯蒂 sigmoid 函数$\sigma(a)$的逆为
13	5	138	倒 4	请注意，非线性变换$\Phi_0(\boldsymbol{x})$并不能消除这种重叠	请注意，非线性变换$\Phi(\boldsymbol{x})$并不能消除这种重叠
14	5	142	图 5.17	由两个高斯混合分布及其累积分布函 $f(a)$	由两个高斯混合分布及其累积分布函数 $f(a)$
15	5	144	(5.95) 上一行	误差函数的梯度减小为	误差函数的梯度归约为
16	6	174	15	网络输出的总数由 $L+2K$ 给出	网络输出的总数由$(L+2)K$ 给出
17	7	195	1	7.4 正则化	7.4 归一化
18	8	218	倒 5	伴随变量 $\overline{v}_1 \sim \overline{v}_2$	伴随变量 $\overline{v}_7 \sim \overline{v}_1$

19	10	275	(10.18)		
20	11	284	图 11.5	但所有条件分布 $p(x_i \mid x_i-1)$都共享一组参数	但所有条件分布 $p(x_i \mid x_{i-1})$都共享一组参数
21	12	305	3	强大归纳偏差	强大归纳偏置
22	12	316	2	受限于输入向量所跨的子空间	受限于输入向量所张成的子空间
23	12	331	倒 11	用一个特殊 token（用<rnask>表示）	用一个特殊 token（用<mask>表示）
24	12	343	倒 2	方差为单位方差的高期分布	方差为单位方差的高斯分布
25	15	398	3	边缘分布的形式是: $p(\boldsymbol{x})=\Sigma_z p(\boldsymbol{x},z)$	边缘分布的形式是 $p(\boldsymbol{x})=\sum_z p(\boldsymbol{x}，z)$
26	15	405	4	后验分布 $p(\boldsymbol{X}，\boldsymbol{Z} \mid \boldsymbol{\theta})$	后验分布 $p(\boldsymbol{Z} \mid \boldsymbol{X}，\boldsymbol{\theta})$