

《最优化：建模、算法与理论》勘误表

刘浩洋 户将 李勇锋 文再文

最后一次更新：2023 年 12 月 24 日

记号说明

- [2] 表示该错误于第 2 次印刷修复。不带该标记的条目表示当前版次还未修复。
- 表中的页码以**出版社版本**为准。

1 第 1 版 (2020 年 12 月)

1. [3] 第 4 页，倒数第 6 行结尾：“混合整合”改为“混合**整数**”。
2. [2] 第 10 页，1.4 节第 5 行：“简单的表示”改为“**相对简单的函数**”。
3. 第 14 页，1.5.1 节第 2 段第 2 行：“可行领域”改为“可行**邻域**”。
4. [2] 第 32 页，最后一行：“时是良定义的”改为“时是**有定义的**”。
5. 第 39 页，子标题“3. 闭函数下半连续函数”改为“3. 闭函数**与**下半连续函数”。
6. [2] 第 34 页，定义 2.5 第 1-2 行：交换“对任意方向 $V \in \mathbb{R}^{m \times n}$ ”和“存在矩阵 $G \in \mathbb{R}^{m \times n}$ ”的位置。
7. [3] 第 44 页，例 2.3：“点 (x, y, z) 构成的图形”后改为“点 (x, y, z) 构成的图形的**边界**”；图 2.9 的图注的末尾也添加“**的边界**”。
8. 第 51 页，推论 2.2 (1) 中的公式：末尾添加 $, x \neq y$ 。
9. 第 65 页，定理 2.20 证明的第一行： $v \in \mathbb{R}^m$ 改为 $v \in \mathbb{R}^n$ 。
10. 第 66 页，第 4-5 行：所有的 x 改为 x_0 ，共 4 处。

11. [3] 第 67 页, 第 1 行: “推论 2.6” 改为 “**命题 2.6**”。
12. 第 89 页, 第 12 行: “该正则项实际上是 $\|x_{\mathcal{I}_\ell}\|_2$ 的 ℓ_1 范数” 改为 “该正则项实际上是 $\{\sqrt{n_\ell}\|x_{\mathcal{I}_\ell}\|_2\}_{\ell=1}^G$ 的 ℓ_1 范数”。
13. 第 91 页, 第 3.4 节第 1 行: “应用广泛” 改为 “**广泛应用**”。
14. [3] 第 97 页, 第 11 行 (公式): $k = 1, 2, \dots, n$ 改为 $k = 1, 2, \dots, m$ 。
15. [2] 第 103 页, 3.10 K-均值聚类第 3 行: 去掉 “**的坐标**”。
16. 第 106 页, 公式 (3.11.1) 以及 (3.11.2) 下方的公式: 积分变量改为 $dx dy$ 。
17. 第 114 页, 习题 3.5: λ 改为 μ 。
18. [2] 第 123 页, 1. 线性最小二乘问题第五行 (公式) 改为:

$$\min_{x \in \mathbb{R}^n} f(x) \stackrel{\text{def}}{=} \frac{1}{2} \|Ax - b\|_2^2.$$

19. [2] 第 124 页, 第 3 行: \leq 应改为 $=$ 。
20. 第 125 页, 深度 Q 学习一节的第 2 行: “一般称 V 为价值函数” 改为 “一般称**式 (3.13.2) 中的 V** 为价值函数”。
21. [2] 第 127 页, 第 8 行: “且初值条件是已知的” 改为 “且 y_1, y_2 **的**初值条件是已知的”。
22. 第 130 页, 第 6 行: “复杂的困难” 改为 “复杂的**问题**”。
23. [2] 第 132 页, 第 8 行 (公式): $\min_{z \in \Gamma}$ 中的 \min 改为 **max**。
24. 第 132 页, 倒数第 2 行 $i = 1, 2, \dots, m$ 改为 $i = 1, 2, \dots, n$ 。
25. 第 133 页, 公式 (4.5.2) 中的 A_m 和 b_m 分别改为 A_n 和 b_n 。
26. 第 138 页, 公式 (4.6.2) 中的 α_1 改为 α 。
27. 第 140 页, 第 6 行 (公式): 最后一项改为 $e^{\text{T}} \varepsilon_{\text{xc}}(\rho)$; 第 8 行 “ V_{ion} 是离子赝势,” 之后添加 “ **e 为全 1 向量,**”。
28. 第 148 页, 习题 4.1: $b \in \mathbb{R}^n$ 改为 $b \in \mathbb{R}^m$ 。
29. [2] 第 153 页, 倒数第 6 行: $x^k \rightarrow +\infty$ 改为 $\|x^k\| \rightarrow +\infty$ 。
30. [2] 第 154 页, 第 10 行: C_{γ} 后添加句点 “.”。

31. 第 166 页, 公式 (5.4.10) 下方一句改为: “对于**强对偶原理**满足的凸问题, 不同写法的拉格朗日对偶问题是等价的”。
32. 第 170 页, 公式 (5.4.17) 中 $\min_{x \in \mathbb{R}^n}$ 改为 $\min_{x \in \mathbb{R}^n, r \in \mathbb{R}^m}$ 。
33. [2] 第 173 页, 第三个公式: $\max_{x \in \mathbb{R}^n}$ 改为 \max 。
34. 第 173 页, 倒数第 2 行: “所有点 x 处的切向量” 改为 “**点 x 处的所有切向量**”。
35. [3] 第 175 页, 图 5.5: 点 x 右侧的虚线改为**实线**, 左侧的虚线**删除**。
36. 第 178 页, 第 9 行: $c_i(z_k) \geq 0$ 改为 $c_i(z_k) \leq 0$ 。
37. 第 178 页, 倒数第 6 行的公式等号右端改为

$$d - \frac{1}{t_k} \begin{bmatrix} A(x) \\ Z^T \end{bmatrix}^{-1} \begin{bmatrix} e_k \\ 0 \end{bmatrix}.$$

38. [2] 第 181 页, 第 12 行: “或不等式约束不起作用 (即 $c_i(x^*)$ 严格大于 0)” 改为 “**或 $c_i(x^*) = 0$ (即 $i \in \mathcal{A}(x^*) \cap \mathcal{I}$)**”。
39. 第 185 页, 第 1 行: $\mathcal{D} \in \mathbb{R}^n$ 改为 $\mathcal{D} \subseteq \mathbb{R}^n$ 。
40. 第 185 页, 第 18 行: $d^* > -\infty$ 前添加 “**对偶问题最优值**”。
41. [2] 第 186 页, 图 5.7: 集合 \mathbb{A} 的边界应该是**单调下降的**。
42. 第 186 页, 公式 (5.6.3) 上方 2 行: $c_i(x)$ 改为 $c(x)$, 即去掉角标 i 。
43. [3] 第 187 页, 第 2–3 行: “则存在……满足……” 改为 “**则存在 e 使得 $\tilde{x} = x_S + e \in \mathcal{D}$, 且满足**”。
44. 第 187 页, 定理 5.13 公式 (5.6.6) 中, 去掉公式第二行的 $\forall i \in \mathcal{E}$ 。
45. [3] 第 188 页, 第 8 行: “则其对应问题的最优解” 改为 “**则其就对应 (5.6.1) 的最优解**”。
46. [3] 第 190 页, 定理 5.14 第一句: “设 Slater 条件满足” 后面添加 “; 且 $\mathcal{X} \cap \text{dom } f \neq \emptyset$ ”。
47. [3] 第 192 页, 倒数第 1 行 (公式): 等号右边第一项 $N_{\mathcal{X}_i}(x)$ 改为 $N_{\mathcal{X}_1}(x)$ 。
48. 第 197 页, 节 5.7.4 第 1 行: “第三章” 改为 “**第四章**”。
49. 第 197–198 页, 公式 (5.7.6) 和 (5.7.7) 中的 \max 改为 \min 。

50. 第 198 页, 倒数第 6 行: $\{c_i(Y)\}$ 改为 $\{\nabla c_i(Y)\}$ 。

51. 第 198 页, 倒数第 4 行 (公式) 改为:

$$L(Y, \lambda) = \text{Tr}(CYY^T) - \sum_{i=1}^n \lambda_i c_i(Y).$$

即公式中的 + 变为 -。

52. 第 199 页, 第 4、8、12 行公式中的 $\text{diag}(Y^T D)$ 改为 $\text{diag}(Y D^T)$ 。

53. [2] 第 217 页, 第 3 行 (公式) 改为:

$$f^* + \nabla f(x)^T(x - x^*) - \frac{\alpha}{2} \|\nabla f(x)\|^2$$

54. [3] 第 219 页, 第 9 行 (公式): 等号左边 $\|x^{k+1} - x^*\|_2^2$ 改为 $\|x^{k+1} - x^*\|^2$ 。

55. [3] 第 221 页, 第 3 行: “Q-超线性收敛速度” 改为 “R-线性收敛速度”。

56. [3] 第 234 页, 第 5 行: “对其极小化” 改为 “求其稳定点”。

57. [3] 第 244 - 245 页, 公式 (6.5.13) 改为:

$$H^{k+1} = (I - \rho_k y^k (s^k)^T)^T H^k (I - \rho_k y^k (s^k)^T) + \rho_k s^k (s^k)^T,$$

公式 (6.5.16) 改为:

$$B^{k+1} = (I - \rho_k s^k (y^k)^T)^T B^k (I - \rho_k s^k (y^k)^T) + \rho_k y^k (y^k)^T,$$

即分别交换等号右边第一项 s^k 和 y^k 的位置。

58. [3] 第 250 - 251 页, 定理 6.10、6.11: R^k 和 L^k 的表达式中 y^{i-1} 改为 y^{j-1} 。

59. [2] 第 256 页, 定理 6.12 的证明第三行 (公式) 改为:

$$L(d, \lambda) = f + g^T d + \frac{1}{2} d^T B d - \frac{\lambda}{2} (\Delta^2 - \|d\|^2).$$

60. [2] 第 257 页, 第 5 行 (公式) 改为:

$$\hat{m}(d) = f + g^T d + \frac{1}{2} d^T (B + \lambda I) d = m(d) + \frac{\lambda}{2} d^T d.$$

61. 第 259 页, 第 6 行 (公式): \min_d 改为 \min_s 。

62. [3] 第 259 页, 第 8 行: α_{k+1} 的表达式中 q^k 改为 p^k 。
63. [3] 第 259 页: 倒数第 12 行 $\|s^{k+1}\| > \Delta$ 改为 $\|s^{k+1}\| \geq \Delta$; 倒数第 11 行 “点 s^{k+1} 将处于信赖域之外” 后添加 “或边界上”; 倒数第 2 行 $\|s^k\| \leq \Delta$ 改为 $\|s^k\| < \Delta$; 倒数第 1 行 $(0, \alpha_k)$ 改为 $(0, \alpha_k]$ 。
64. [3] 第 260 页, 算法 6.9: 第 8 行 $\|s^{k+1}\| > \Delta$ 改为 $\|s^{k+1}\| \geq \Delta$ 。
65. [3] 第 261 页: 第 12 行 $\|s^{t+1}\| > \Delta$ 改为 $\|s^{t+1}\| \geq \Delta$; 第 13 行 $(0, \alpha_t)$ 改为 $(0, \alpha_t]$ 。
66. [2] 第 262 页, 倒数第 7 行: “它并没有利用” 改为 “它并没有充分利用”。
67. 第 264 页, 定理 6.16 的第 2 行: $\nabla^2 f(x) \succ 0$ 改为 $\nabla^2 f(x^*) \succ 0$ 。
68. 第 268 页, 式 (6.7.4) 后面添加 $\forall z \in \mathbb{R}^n$ 。
69. 第 268 页, 倒数第 7 行 (公式): 其中一项 $\frac{\gamma^2 \|d^k\|^2}{\beta^2 \|d^k\|}$ 改为 $\frac{\gamma^2 \|d^k\|^2}{\beta^2 \|d^k\|^2}$ 。
70. 第 269 页, 定理 6.18 的第 5 行: 矩阵 H^* 的定义改为

$$H^* = \sum_{i=1}^m r_i(x^*) \nabla^2 r_i(x^*)$$

71. [2] 第 270 页, 第 12、20 行: (6.7.11) 改为 (6.7.10)。
72. [2] 第 279 页, 习题 6.4:
- (b) 问前半句修改为 “证明 $f(x)$ 在区域 $\{x \mid \|x\| \leq R \stackrel{\text{def}}{=} 1/\sqrt{K}\}$ 上是 G-利普希茨连续的”;
 - (c) 问第 2 行: “在 k ($k \leq K$) 次迭代后” 前面添加 “存在一种次梯度的取法, ”; 第三行公式改为:

$$\hat{f}^k - f^* \geq \frac{GR}{2(1 + \sqrt{K})}$$

最后一行开头 $\mathcal{O}\left(\frac{1}{\sqrt{k}}\right)$ 改为 $\mathcal{O}\left(\frac{GR}{\sqrt{K}}\right)$ 。

73. 第 280 页, 第 1 行 $\beta > 1$ 改为 $\beta > 3/2$; 第 2 行 $0 < \beta < 1$ 改为 $0 < \beta < 3/2$ 。
74. 第 288 页, 倒数第 8 行: $L(x, \lambda^*)$ 后添加 “的海瑟矩阵”。
75. 第 289 页, 定义 7.2 之后第 1 行: $h(t) = (\min\{t, 0\})^2$ 中的 \min 改为 \max 。
76. [2] 第 293 页, 倒数第四行: “随着 σ 增大” 改为 “随着 σ 减小”。

77. [3] 第 294 页, 算法 7.4 下方第 2 行: “一个常用的对数罚函数收敛准则可以是” 改为 “常用的收敛准则可以包含”; 下方公式改为

$$\left| \sigma_k \sum_{i \in \mathcal{I}} \ln(-c_i(x^{k+1})) \right| \leq \varepsilon,$$

即等号左边添加绝对值。

78. [2] 第 297 页, 删除倒数第 3–4 行。
79. [2] 第 303 页, 算法 7.6, 14 行: $\eta_{k+1}, \varepsilon_{k+1}$ 的更新方式改为:

$$\eta_{k+1} = \frac{1}{\sigma_{k+1}}, \quad \varepsilon_{k+1} = \frac{1}{\sigma_{k+1}^\alpha}.$$

80. 第 309 页, 定理 7.8 的证明做如下修改:

- 所有的上标 j_k 替换为 k ;
- 证明的最后三行中的 x^k 都替换为 x^{k+1} (共 5 处)。

81. [3] 第 311 页, 第 3 行: 公式中的 λ 改为 λ^k 。

82. 第 313 页, 倒数第 4 行: 公式中 Λ^{k+1} 的更新方式改为:

$$\Lambda^{k+1} = \Lambda^k + \sigma_k \left(\sum_{i=1}^m y_i^{k+1} A_i + S^{k+1} - C \right)$$

83. 第 314 页, 公式 (7.2.36) 中的 Λ^{k+1} 更新中的 σ 改为 σ_k 。

84. 第 321 页, 习题 7.4 的第一个公式中: 最左边的 c_k 改为 σ_k 。

85. [3] 第 333 页, 算法 8.1: 第 1 行末尾添加 “初始化 $k \leftarrow 1$.”; 第 3 行与第 4 行之间增加 $k \leftarrow k + 1$ 。

86. [2] 第 335 页, 3. 小波模型求解: “下面用近似点梯度算法求解小波模型” 改为 “下面考虑小波分解模型”; 公式 (8.1.7) 前一行结尾: “我们有等价模型” 改为 “可以使用近似点梯度法求解对应的合成模型”。

87. [2] 第 337 页, 第 14 行末尾: “的最小值” 改为 “的最小值点”。

88. 第 346 页, 算法 8.6 的第 6 行: $(x^k - x^{k-1})$ 改为 $(u - x^{k-1})$ 。

89. [3] 第 347 页, 算法 8.8: 第 1 行的 x^0 和 y^0 交换位置, $k \leftarrow 0$ 改为 $k \leftarrow 1$ 。

90. [3] 第 348 页, 算法 8.9: 第 1 行末尾添加 “初始化 $k \leftarrow 1$.”; 第 5 行与第 6 行之间增加 $k \leftarrow k + 1$.
91. [2] 第 348 页, 8.2.3 应用举例, 第 13 行 (公式) 中的 y^k 改为 y^{k-1} 。类似修改还包含: 第 349 页倒数第 3 行、第 350 页第 8 行。
92. [2] 第 349 页, 2. 小波模型求解后一行: “针对小波模型求解” 改为 “针对合成小波模型求解”。
93. [3] 第 350 页, 倒数第 2 行: 删除 “的凸性、”。
94. [3] 第 361 页, 第 9 行 (公式): $\Gamma_{\lambda t_k}^2$ 改为 $\mathbf{I}_{\lambda t_k}^2$ (改为斜体)。
95. 第 361 页, 倒数第 1 行 (公式): $\Gamma_{\mu t_k}^2(U)$ 改为 $\mathbf{I}_{\lambda t_k}^2(U)$
96. 第 362 页, 第 1 行和第 3 行 (公式) 中的 $\Gamma_{\mu t_k}^2(U)$ 改为 $\nabla_U \mathbf{I}_{\lambda t_k}^2(U)$
97. 第 362 页, 第 5 行: 公式中的 \min 改为 \max 。
98. [3] 第 363 页最后一行公式添加编号 (8.3.14), 并去掉原 364 页 (8.3.14) 的编号; 第 364 页第 5 行 (8.3.4) 改为 (8.3.14)。
99. 第 368 页, 倒数第 2 行: 公式中的 $r_2(X)$ 改为 $r_2(\mathbf{Y})$ 。
100. [2] 第 358 页, 第 4 行: “增广拉格朗日方法” 改为 “增广拉格朗日函数法”。
101. [3] 第 386 页, 定理 8.11: “且假设 8.2 满足,” 的后面添加 “ $\{z^k\}$ 是有界序列”。
102. [3] 第 389 页, 总结第 (3) 条的内容改为 “假设 Ψ 是一个 KL 函数且迭代序列 $\{z^k\}$ 有界, 证明 $\{z^k\}$ 是一个柯西列”。
103. 第 391 页, 第 6 – 7 行的公式改为:
- $$f(x_2) - \frac{\mu}{2} \|x_2\|^2 \geq f(x_1) - \frac{\mu}{2} \|x_1\|^2 + (y_1 - \mu x_1)^\top (x_2 - x_1),$$
- $$f(x_1) - \frac{\mu}{2} \|x_1\|^2 \geq f(x_2) - \frac{\mu}{2} \|x_2\|^2 + (y_2 - \mu x_2)^\top (x_1 - x_2),$$
104. [2] 第 391 页, 引理 8.5 证明的倒数第二行: “定理 6.1” 改为 “引理 6.1”。
105. [2] 第 392 页, 公式 (8.5.8) 第三行: 去掉 $h(y)$ 左边的 “(”。
106. [3] 第 404 页, 倒数第 5 行 (公式): $h^*(\bar{z})$ 改为 $h^*(z)$ 。
107. [2] 第 408 页, 公式 (8.6.3): $L_\rho(x_1^k, x_2^k, y^k)$ 改为 $L_\rho(x_1, x_2, y^k)$ 。

108. 第 410 页, 公式 (8.6.10): A_2 前添加 ρ ;
第 411 页, 第 1 行的公式改为 $\rho A_1^T A_2 (x_2^{k-1} - x_2^k)$ 。
109. [2] 第 411 页, 8.6.2 节第 2 行: 删除“组合”。
110. [3] 第 413 页, 倒数第 6 行: 删除“针”。
111. [3] 第 413 页, 定理 8.15: 最前面添加“如果 $w^1 = -tA_2x_2^0$, 那么”。
112. [2] 第 414 页, 第 6 行 (公式) 改为:
- $$x_1^k = \arg \min_{x_1} \left\{ f_1(x_1) + (x^k)^T (A_1 x_1 - b) + \frac{t}{2} \|A_1 x_1 - b - \frac{w^k}{t}\|_2^2 \right\}.$$
113. [3] 第 414 页, 倒数第 8 行: $\|A_1 x_1 + A_2 x_2^{k-1} - b\|$ 改为 $\|A_1 x_1 + A_2 x_2^{k-1} - b\|^2$ 。
114. [3] 第 414 页, 倒数第 7 行: $\|A_1 x_1^k + A_2 x_2 - b\|$ 改为 $\|A_1 x_1^k + A_2 x_2 - b\|^2$ 。
115. [2] 第 415 页, 第 12、15、倒数第 1 行的公式: 所有的 x^k 改为 x_1^k 。
116. 第 415 页, 第 15 行公式: $f(x_1)$ 改为 $f_1(x_1)$ 。
117. [2] 第 415 页, 2. 缓存分解: $A_1^T A$ 改为 $A_1^T A_1$ 。
118. [2] 第 416 页, 5. 超松弛: 公式中的 $+(1 - \alpha_k)$ 改为 $-(1 - \alpha_k)$ 。
119. [2] 第 418 页, 第三个公式中关于 z^{k+1} 的更新: $\|x\|_1$ 改为 $\|z\|_1$ 。
120. 第 428 页, 第 5 行 (公式): 对调 \tilde{A} 和 \hat{A} 的值。
121. [3] 第 442 页, 图 8.12: 标题改为“使用不同类型的随机梯度法求解逻辑回归问题”。
122. [3] 第 447 页, 第 1 行: “定理 8.8” 改为“引理 8.8”。
123. 第 450 页, 倒数第 10 行: $\nabla f_{s_k}(x_k)$ 改为 $\nabla f_{s_k}(x^k)$ 。
124. [3] 第 453 页, 倒数第 2 行: “相对于普通梯度算法” 改为“相对于普通的随机梯度算法”。
125. [3] 第 457 页, 定理 8.25 第一句: “设每个 $f_i(x)$ 是可微的” 改为“设每个 $f_i(x)$ 是可微凸函数”。
126. [2] 第 462 页, 习题 8.5(a): “当 $t_k = \gamma_k \lambda_k$ 时” 改为“当 $t_k = \gamma_k \lambda_k$ 且 $h(x) = 0$ 时”。
127. [2] 第 462 页, 习题 8.11: “基于格式 (8.4.3)” 改为“基于格式 (8.4.4)”。

128. 第 464 页, 习题 8.20: 证明的公式改为

$$\mathbb{E}[\|\nabla f_{s_k}(x^k)\|^2] \leq L^2 \mathbb{E}[\|x^k - x^*\|^2] + \mathbb{E}[\|\nabla f_{s_k}(x^k) - \nabla f(x^k)\|^2],$$

同时删除 “ α_k 为第 k 步的步长”。