第 9 章：复习课

课堂练习与参考解答

授课教师：曹语

课程主页：`https://yucaoyc.github.io/math3806`

练习 1（第 1 章）：资源分配建模

一家小工厂每天生产两类零件，记生产批量为 x_1, x_2（可按连续批量近似）。

零件	单位利润（千元）	每批机床工时	每批检验工时
A	3	2	1
B	2	1	2

每天最多有 8 个机床工时和 8 个检验工时，且 $x_1, x_2 \geq 0$。

练习 1:

- 写出标准形式的最小化问题；
- 找出最优解。

练习 2（第 2 章）：判断凸优化问题

给定 $\lambda > 0, r > 0$，考虑

$$
\begin{array}{ll}
\text{minimize} & \log\left(\sum_{i=1}^{m} \exp\left(\boldsymbol{a}_i^\top \boldsymbol{x} + b_i\right)\right) + \lambda \|\boldsymbol{x}\|_2^2 \\
\text{subject to} & \|\boldsymbol{B}\boldsymbol{x} - \boldsymbol{c}\|_2 \leq r, \\
& \boldsymbol{D}\boldsymbol{x} = \boldsymbol{e}.
\end{array}
$$

练习 2:

- 请判断这是凸优化问题吗？如果是，请说明目标函数和可行域分别为什么是凸的。
- 若可行域非空，$\lambda > 0$ 能否保证最优解唯一？

练习 3（第 3 章）：梯度下降与 Newton 法

考虑二次函数

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^\top \begin{bmatrix} 4 & 1 \\ 1 & 3 \end{bmatrix} \boldsymbol{x} - \begin{bmatrix} 1 \\ 2 \end{bmatrix}^\top \boldsymbol{x}.$$

从 $\boldsymbol{x}^{(0)} = \mathbf{0}$ 出发。

练习 3:

- 计算 $\nabla f(\boldsymbol{x}^{(0)})$；
- 沿负梯度方向做一次精确直线搜索，求 $\boldsymbol{x}^{(1)}$；
- 计算一次 Newton 更新，并判断是否到达最优解。

练习 4（第 4 章）：Ridge 拟合

给定

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}, \qquad \boldsymbol{y} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}.$$

考虑正则化最小二乘

$$\min_{\boldsymbol{w}\in\mathbb{R}^2} \frac{1}{2}\|\boldsymbol{A}\boldsymbol{w} - \boldsymbol{y}\|_2^2 + \frac{1}{2}\|\boldsymbol{w}\|_2^2.$$

练习 4：

- 写出一阶最优性条件；
- 简要说明正则化项在工程建模中的作用。

练习 5（第 5 章）：KKT 条件

考虑一维凸优化问题

$$\begin{aligned} \text{minimize} \quad & (x-2)^2 \\ \text{subject to} \quad & x \leq 1. \end{aligned}$$

练习 5：

- 写出 Lagrange 函数；
- 写出 KKT 条件；
- 求最优解 $x^\star$ 和最优乘子 $\lambda^\star$。

练习 6（第 6 章）：等式约束 Newton 步

考虑

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}x_1^2 + 2x_2^2 - x_1 - 2x_2 \\ \text{subject to} & x_1 + x_2 = 1. \end{array}$$

从可行点 $\boldsymbol{x}^{(0)} = (1, 0)^\top$ 出发。

练习 6：

- 写出等式约束 Newton 系统；
- 求 Newton 方向 $\Delta\boldsymbol{x}$；
- 给出更新后的点，并判断是否为最优解。

练习 7（第 7 章）：障碍子问题

考虑最简单的不等式约束问题

$$\begin{array}{ll} \text{minimize} & x \\ \text{subject to} & 1 - x \leq 0. \end{array}$$

练习 7：

- 写出参数 $t > 0$ 对应的对数障碍子问题；
- 求中心点 $x^\star(t)$；
- 计算目标函数误差 $f_0(x^\star(t)) - p^\star$。

目录

参考解答

参考解答 1：标准形式

原问题是最大化总利润（单位：千元）

$$\max_{x_1,x_2}\ 3x_1+2x_2 \quad \text{subject to} \quad 2x_1+x_2\le 8,\ x_1+2x_2\le 8,\ x_1,x_2\ge 0.$$

标准最小化形式为

$$\begin{aligned}
\text{minimize} \quad & -3x_1-2x_2 \\
\text{subject to} \quad & 2x_1+x_2-8\le 0, \\
& x_1+2x_2-8\le 0, \\
& -x_1\le 0, \quad -x_2\le 0.
\end{aligned}$$

参考解答 1（续）：顶点比较

二维线性规划的最优解在顶点处取得。顶点利润：

$$(0,0):0, \qquad (4,0):12, \qquad (0,4):8.$$

两个资源约束交点满足

$$2x_1 + x_2 = 8, \qquad x_1 + 2x_2 = 8 \quad \Longrightarrow \quad x_1 = x_2 = \frac{8}{3}.$$

利润为 $3 \cdot \frac{8}{3} + 2 \cdot \frac{8}{3} = \frac{40}{3}$，最大。

最优解为

$$\boldsymbol{x}^\star = \begin{bmatrix} 8/3 \\ 8/3 \end{bmatrix}.$$

参考解答 2：目标函数

目标函数第一项

$$\log\left(\sum_{i=1}^{m}\exp\left(\boldsymbol{a}_i^\top \boldsymbol{x} + b_i\right)\right)$$

是仿射函数的 log-sum-exp，因此是凸函数。

第二项

$$\lambda\|\boldsymbol{x}\|_2^2$$

在 $\lambda > 0$ 时是强凸函数。因此目标函数是强凸函数。

参考解答 2（续）：可行域与唯一性

约束

$$\|\boldsymbol{B}\boldsymbol{x}-\boldsymbol{c}\|_2 \le r$$

是范数球在仿射映射下的原像，是凸集；约束

$$\boldsymbol{D}\boldsymbol{x}=\boldsymbol{e}$$

是仿射集合，也是凸集。可行域是凸集交集，仍然凸。

所以该问题是凸优化问题。若可行域非空，则二次正则项使目标函数强凸并保证最优解存在且唯一。

参考解答 3：梯度和方向

记

$$\boldsymbol{P} = \begin{bmatrix} 4 & 1 \\ 1 & 3 \end{bmatrix}, \qquad \boldsymbol{q} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

则

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{P}\boldsymbol{x} - \boldsymbol{q}^\top \boldsymbol{x}, \qquad \nabla f(\boldsymbol{x}) = \boldsymbol{P}\boldsymbol{x} - \boldsymbol{q}.$$

在 $\boldsymbol{x}^{(0)} = \boldsymbol{0}$ 处，

$$\nabla f(\boldsymbol{x}^{(0)}) = -\boldsymbol{q} = \begin{bmatrix} -1 \\ -2 \end{bmatrix}.$$

负梯度方向为

$$\boldsymbol{d} = -\nabla f(\boldsymbol{x}^{(0)}) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

参考解答 3（续）：精确直线搜索

从 $\boldsymbol{x}^{(0)} = \mathbf{0}$ 出发，负梯度方向为

$$\boldsymbol{d} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

精确直线搜索步长：

$$\alpha^\star = \frac{\boldsymbol{q}^\top \boldsymbol{q}}{\boldsymbol{q}^\top \boldsymbol{P} \boldsymbol{q}} = \frac{5}{20} = \frac{1}{4}.$$

所以

$$\boldsymbol{x}_{\mathrm{gd}}^{(1)} = \frac{1}{4} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1/4 \\ 1/2 \end{bmatrix}.$$

参考解答 3（续）：Newton 更新

Newton 方向满足

$$\boldsymbol{P}\Delta\boldsymbol{x}_{\mathrm{nt}} = -\nabla f(\boldsymbol{x}^{(0)}) = \boldsymbol{q}.$$

因此

$$\Delta\boldsymbol{x}_{\mathrm{nt}} = \boldsymbol{P}^{-1}\boldsymbol{q} = \frac{1}{11}\begin{bmatrix} 3 & -1 \\ -1 & 4 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1/11 \\ 7/11 \end{bmatrix}.$$

Newton 更新为

$$\boldsymbol{x}_{\mathrm{nt}}^{(1)} = \begin{bmatrix} 1/11 \\ 7/11 \end{bmatrix}.$$

参考解答 3（续）：为什么一步到达

由于 f 是正定二次函数，最优点满足

$$\nabla f(\boldsymbol{x}^\star) = \boldsymbol{P}\boldsymbol{x}^\star - \boldsymbol{q} = \boldsymbol{0},$$

所以 Newton 法一步到达全局最优点。

参考解答 4：Ridge 拟合

目标函数为

$$\frac{1}{2}\|\boldsymbol{A}\boldsymbol{w}-\boldsymbol{y}\|_2^2+\frac{1}{2}\|\boldsymbol{w}\|_2^2.$$

一阶最优性条件给出：

$$(\boldsymbol{A}^\top\boldsymbol{A}+\boldsymbol{I})\boldsymbol{w}=\boldsymbol{A}^\top\boldsymbol{y}.$$

参考解答 4（续）：正则化含义

正则化项 $\frac{1}{2}\|\boldsymbol{w}\|_2^2$ 会惩罚过大的参数，使拟合更稳定，降低对噪声数据的敏感性。

工程上可以理解为：不只要求训练误差小，还希望参数不要过度放大噪声。

参考解答 5：KKT 条件

把约束写成

$$f_1(x) = x - 1 \leq 0.$$

Lagrange 函数为

$$L(x, \lambda) = (x-2)^2 + \lambda(x-1), \qquad \lambda \geq 0.$$

KKT 条件：

$$\begin{aligned} &x^\star - 1 \leq 0, \\ &\lambda^\star \geq 0, \\ &\lambda^\star(x^\star - 1) = 0, \\ &2(x^\star - 2) + \lambda^\star = 0. \end{aligned}$$

参考解答 5（续）：求解

由于无约束最优点 $x=2$ 不可行，约束应当紧：

$$x^\star = 1.$$

代入平稳性：

$$2(1-2)+\lambda^\star = 0 \quad \Longrightarrow \quad \lambda^\star = 2.$$

参考解答 6：梯度与 Hessian

目标函数为

$$f(x_1, x_2) = \frac{1}{2}x_1^2 + 2x_2^2 - x_1 - 2x_2.$$

梯度与 Hessian 为

$$\nabla f(\boldsymbol{x}) = \begin{bmatrix} x_1 - 1 \\ 4x_2 - 2 \end{bmatrix}, \qquad \nabla^2 f(\boldsymbol{x}) = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}.$$

在 $\boldsymbol{x}^{(0)} = (1, 0)^\top$ 处，

$$\nabla f(\boldsymbol{x}^{(0)}) = \begin{bmatrix} 0 \\ -2 \end{bmatrix}.$$

约束矩阵为 $\boldsymbol{A} = \begin{bmatrix} 1 & 1 \end{bmatrix}$。

参考解答 6（续）：Newton 系统

Newton 系统：

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 4 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \\ w \end{bmatrix} = - \begin{bmatrix} 0 \\ -2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}.$$

由 Newton 系统得到

$$\Delta x_1 + w = 0, \qquad 4\Delta x_2 + w = 2, \qquad \Delta x_1 + \Delta x_2 = 0.$$

解得

$$\Delta x_1 = -\frac{2}{5}, \qquad \Delta x_2 = \frac{2}{5}, \qquad w = \frac{2}{5}.$$

参考解答 6（续）：更新和验证

因此

$$\boldsymbol{x}^{(1)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} -2/5 \\ 2/5 \end{bmatrix} = \begin{bmatrix} 3/5 \\ 2/5 \end{bmatrix}.$$

验证 KKT 条件：

$$x_1^{(1)} + x_2^{(1)} = 1,$$

且

$$\nabla f(\boldsymbol{x}^{(1)}) = \begin{bmatrix} -2/5 \\ -2/5 \end{bmatrix}, \qquad \nabla f(\boldsymbol{x}^{(1)}) + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \frac{2}{5} = \boldsymbol{0}.$$

所以该点是全局最优解。

参考解答 7：障碍子问题

约束 $1 - x \leq 0$ 等价于 $x > 1$ 上的严格可行域。

参数 $t > 0$ 对应的对数障碍子问题为

$$\min_{x>1} \; tx - \log(x-1).$$

一阶条件：

$$t - \frac{1}{x-1} = 0 \quad \Longrightarrow \quad x^\star(t) = 1 + \frac{1}{t}.$$

参考解答 7（续）：目标误差

原问题最优解为 $x^\star = 1$，最优值为

$$p^\star = 1.$$

因此中心点的目标误差为

$$f_0(x^\star(t)) - p^\star = 1 + \frac{1}{t} - 1 = \frac{1}{t}.$$