最优化导论第五次作业题

1. 考虑问题:

$$\begin{array}{ll} \min & f_0(x) \\ \text{s.t.} & f_i(x) \leq 0, \quad i = 1, \dots, m, \end{array}$$

其中函数 $f_i: \mathbb{R}^n \to \mathbb{R}$ 是可微且凸的。

在精确罚函数方法中，我们求解如下辅助问题:

$$\min \quad \phi(x) = f_0(x) + \alpha \max_{i=1,\dots,m} \max\{0, f_i(x)\},$$

其中参数 $\alpha > 0$。

项 $\alpha \max\{\cdot\}$ 用来惩罚 x 不可行性偏差。

如果 α 足够大，则该方法称为 精确罚函数方法（exact penalty method），因为辅助问题 (5.111) 的解也同时是原问题 (5.110) 的解。

(a) 证明 $\phi(x)$ 是凸函数。

(b) 辅助问题可写为如下形式:

$$\begin{array}{ll} \min & f_0(x) + \alpha y \\ \text{s.t.} & f_i(x) \leq y, \quad i = 1, \dots, m \\ & 0 \leq y, \end{array}$$

其中变量为 $x \in \mathbb{R}^n, y \in \mathbb{R}$。

求该问题的拉格朗日对偶问题，并用原问题 (5.110) 的对偶函数 $g(\lambda)$ 表示之。

解: (a) 第一项 $f_0(x)$ 是凸的。

第二项:

$$\max\{f_1(x), \dots, f_m(x), 0\}$$

是多个凸函数的点态最大值，因此也是凸函数。

故 $\phi(x)$ 为凸函数。

(b) 其拉格朗日函数为:

$$L(x,y,\lambda,\mu)=f_0(x)+\alpha y+\sum_{i=1}^{m}\lambda_i\left(f_i(x)-y\right)-\mu y.$$

对应的对偶函数为

$$\begin{aligned}\inf_{x,y}L(x,y,\lambda,\mu)\quad &=\inf_{x,y}\left(f_0(x)+\alpha y+\sum_{i=1}^{m}\lambda_i\left(f_i(x)-y\right)-\mu y\right)\\ &=\inf_{x}\left(f_0(x)+\sum_{i=1}^{m}\lambda_i f_i(x)\right)+\inf_{y}\left(\left(\alpha-\sum_{i=1}^{m}\lambda_i-\mu\right)y\right)\\ &=\begin{cases}g(\lambda), & \mathbf{1}^T\lambda+\mu=\alpha\\ -\infty, & \text{其他情况,}\end{cases}\end{aligned}$$

其中 $g(\lambda)$ 是原问题 (5.110) 的对偶函数。

因此对偶问题为:

$$\begin{aligned}\max\quad & g(\lambda)\\ \text{s.t.}\quad & \mathbf{1}^T\lambda+\mu=\alpha,\\ & \lambda\geq 0,\quad \mu\geq 0.\end{aligned}$$

等价地，可消去 μ，得到:

$$\begin{aligned}\max_{\lambda}\quad & g(\lambda)\\ \text{s.t.}\quad & \mathbf{1}^T\lambda\leq\alpha,\\ & \lambda\succcurlyeq 0.\end{aligned}$$

2. 推导下列问题的对偶问题:

$$\min_{x}\sum_{i=1}^{N}\| A_i x+b_i \|_2+\frac{1}{2}\| x-x_0 \|_2^2.$$

已知问题数据为 $A_i\in\mathbb{R}^{m_i\times n}$、$b_i\in\mathbb{R}^{m_i}$、$x_0\in\mathbb{R}^n$。
首先引入新变量 $y_i\in\mathbb{R}^{m_i}$，并加入等式约束

$$y_i=A_i x+b_i.$$

解: 拉格朗日函数为

$$L(x,z_1,\ldots,z_N)=\sum_{i=1}^{N}\| y_i \|_2+\frac{1}{2}\| x-x_0 \|_2^2-\sum_{i=1}^{N}z_i^T\left(y_i-A_i x-b_i\right).$$

先对 y_i 取下确界，对每个 i，有

$$\inf_{y_i}(\| y_i \|_2 + z_i^T y_i) = \begin{cases} 0, & \| z_i \|_2 \leq 1, \\ -\infty, & \text{否则}. \end{cases}$$

（说明：若 $\| z_i \|_2 > 1$，取 $y_i = -tz_i$ 并令 $t \to \infty$，可见该函数在下方无界；
若 $\| z_i \|_2 \leq 1$，由 Cauchy–Schwarz 不等式可知 $\| y_i \|_2 + z_i^T y_i \geq 0$，
因此最小值在 $y_i = 0$ 处取得。）

再对 x 取下确界：

对 x 最小化时，对 x 求梯度并令其为零，得到

$$x = x_0 + \sum_{i=1}^{N} A_i^T z_i.$$

代回拉格朗日函数得到对偶函数：

将上述结果代回拉格朗日函数，得到对偶函数

$$g(z_1, \dots, z_N) = \begin{cases} \sum_{i=1}^{N} (A_i x_0 + b_i)^T z_i - \frac{1}{2} \left\| \sum_{i=1}^{N} A_i^T z_i \right\|_2^2, & \| z_i \|_2 \leq 1,\ i = 1, \dots, N, \\ -\infty, & \text{否则}. \end{cases}$$

因此，对偶问题为

$$\begin{aligned} \max_{z_1, \dots, z_N} \quad & \sum_{i=1}^{N} (A_i x_0 + b_i)^T z_i - \frac{1}{2} \left\| \sum_{i=1}^{N} A_i^T z_i \right\|_2^2 \\ \text{s.t.} \quad & \| z_i \|_2 \leq 1, \quad i = 1, \dots, N. \end{aligned}$$

这就是所求的拉格朗日对偶问题。

3. 验证点：

$$(x_1, x_2) = \left(\frac{4}{5}, \frac{8}{5}\right)$$

是以下优化问题的局部/全局最优解：

$$\min\ x_1^2 + x_2^2$$

满足：

$$\begin{aligned} x_1^2 + x_2^2 \quad & \leq 5, \\ x_1 + 2x_2 \quad & = 4, \\ x_1, x_2 \quad & \geq 0. \end{aligned}$$

解： 构造拉格朗日函数:

$L(x_1, x_2, u_1, u_2, u_3, v) = x_1^2 + x_2^2 + u_1(x_1^2 + x_2^2 - 5) - u_2 x_1 - u_3 x_2 + v(x_1 + 2x_2 - 4),$

其中:

- $u_1 \geq 0$ 对应不等式 $x_1^2 + x_2^2 \leq 5$
- $u_2 \geq 0$ 对应不等式 $x_1 \geq 0$
- $u_3 \geq 0$ 对应不等式 $x_2 \geq 0$
- v 为等式约束 $x_1 + 2x_2 = 4$ 的拉格朗日乘子

(i) 原问题可行性

$$x_1^2 + x_2^2 \leq 5, \quad x_1 + 2x_2 = 4, \quad x_1, x_2 \geq 0.$$

(ii) 互补松弛

$$u_1(x_1^2 + x_2^2 - 5) = 0, \qquad u_1 \geq 0,$$

(iii) Stationarity（梯度平衡）

计算偏导:

$$\frac{\partial L}{\partial x_1} = 2x_1 + 2u_1 x_1 - u_2 + v = 0,$$

检查点 $(x_1, x_2) = \left(\frac{4}{5}, \frac{8}{5}\right)$ 是否满足 KKT

检查可行性:

$$x_1 + 2x_2 = \frac{4}{5} + \frac{16}{5} = 4$$

满足。

$$x_1^2 + x_2^2 = \frac{16}{25} + \frac{64}{25} = \frac{80}{25} = \frac{16}{5} = 3.2 < 5$$

严格小于→约束不活跃

因此:

$$u_1 = 0$$

另外:

$$x_1 = \frac{4}{5} > 0 \Rightarrow u_2 = 0,$$

即所有不等式约束均不活跃：

$$u_1 = u_2 = u_3 = 0.$$

(ii) Stationarity 条件：

代入 $u_1 = u_2 = u_3 = 0$：

$$2x_1 + v = 0 \Rightarrow 2 \cdot \frac{4}{5} + v = 0 \Rightarrow \frac{8}{5} + v = 0 \Rightarrow v = -\frac{8}{5}$$

成立。

因此：

$$v = -\frac{8}{5}$$

我们得到了满足所有 KKT 条件的点：

$$(x_1, x_2, u_1, u_2, u_3, v) = \left(\frac{4}{5}, \frac{8}{5}, 0,0,0, -\frac{8}{5}\right).$$

由于目标函数 $x_1^2 + x_2^2$ 是凸函数，不等式约束是凸的，等式约束是仿射的（线性的），并且 Slater 条件成立，因此：KKT 条件是充分必要条件，所以该点是全局最优解。

4. 利用 KKT 条件，求集合

$$M = \{x \in \mathbb{R}^2 \colon\ x_1 + x_2 \geq 4,\ 2x_1 + x_2 \geq 5\}$$

中距离原点 (0,0) 最近的点，并判断解是否唯一。

解： 我们要求解如下优化问题：

$$\min x_1^2 + x_2^2$$

构造 Lagrangian

$$L(x_1, x_2, u_1, u_2) = x_1^2 + x_2^2 + u_1(-x_1 - x_2 + 4) + u_2(-2x_1 - x_2 + 5),$$

其中 $u_1, u_2 \geq 0$。

KKT 条件为可行性 $x_1 + x_2 \geq 4,\ 2x_1 + x_2 \geq 5$，互补松弛：

$$u_1(-x_1 - x_2 + 4) = 0, \quad u_2(-2x_1 - x_2 + 5) = 0, \quad u_1, u_2 \geq 0,$$

stationarity：

$$\frac{\partial L}{\partial x_1} = 2x_1 - u_1 - 2u_2 = 0, \quad \frac{\partial L}{\partial x_2} = 2x_2 - u_1 - u_2 = 0.$$

从互补性我们考虑以下情况：
情况 1：$u_1 = 0, u_2 = 0$，则 $x_1 = 0, x_2 = 0$，但此点不可行，舍去。
情况 2：约束 $x_1 + x_2 = 4$ 活跃，且 $u_2 = 0$，求解：

$$2x_1 - u_1 = 0, \quad 2x_2 - u_1 = 0, \quad x_1 + x_2 = 4,$$

得 $x_1 = x_2 = 2, u_1 = 4 > 0$，并且该点满足另一约束 $2x_1 + x_2 = 6 \geq 5$ 可行，得到 KKT 点 $(2,2,4,0)$。
情况 3：约束 $2x_1 + x_2 = 5$ 活跃，且 $u_1 = 0$，求解：

$$2x_1 - 2u_2 = 0, \quad 2x_2 - u_2 = 0, \quad 2x_1 + x_2 = 5,$$

得 $x_1 = 2, x_2 = 1, u_2 = 2$，但此点不满足另一约束 $x_1 + x_2 = 3 < 4$，不可行，舍去。
情况 4：两个约束均活跃 $x_1 + x_2 = 4, 2x_1 + x_2 = 5$，得 $x_1 = 1, x_2 = 3$，进一步由 stationarity：

$$u_1 + 2u_2 = 2, \quad u_1 + u_2 = 6,$$

解得 $u_1 = 10, u_2 = -4$，但 $u_2 < 0$，不合法，舍去。

因此唯一满足所有 KKT 条件的可行点为：

$$(x_1, x_2) = (2,2).$$

由于目标函数为严格凸函数且可行域为凸集合 M，投影点唯一，因此距离原点最近的点为 $(2,2)$，且解唯一。

5. 设 $n \geq 2$，考虑优化问题：

$$\min x_1$$

满足：

$$\sum_{i=1}^{n} \left(x_i - \frac{1}{n}\right)^2 \leq \frac{1}{n(n-1)}, \qquad \sum_{i=1}^{n} x_i = 1.$$

证明：

$$\left(0, \frac{1}{n-1}, \dots, \frac{1}{n-1}\right)$$

是该问题的最优解。

解：

$$x^\star = \left(0, \frac{1}{n-1}, \dots, \frac{1}{n-1}\right)$$

满足可行性约束：

$$\sum_{i=1}^{n} x_i = 0 + \sum_{i=2}^{n} \frac{1}{n-1} = 1,$$

并且

$$\left(0 - \frac{1}{n}\right)^2 + (n-1)\left(\frac{1}{n-1} - \frac{1}{n}\right)^2 = \frac{1}{n(n-1)},$$

说明不等式约束为活跃。

构造 Lagrangian：

$$L = x_1 + u\left(\sum_{i=1}^{n}\left(x_i - \frac{1}{n}\right)^2 - \frac{1}{n(n-1)}\right) + v\left(\sum_{i=1}^{n} x_i - 1\right),$$

KKT 条件包括 stationarity：

$$\frac{\partial L}{\partial x_1} = 1 + 2u\left(x_1 - \frac{1}{n}\right) + v = 0,$$

在最优点：

- $x_1 = 0$
- $x_i = \frac{1}{n-1},\ i \geq 2$

代入得到方程系统：

$$1 - \frac{2u}{n} + v = 0,$$

解得：

$$u = \frac{n-1}{2} \geq 0, \qquad v = -\frac{1}{n}.$$

因此该点满足可行性、stationarity 和互补松弛，构成 KKT 点。

由于目标函数为线性的（凸），不等式约束为凸，等式约束为线性，因此 KKT 条件也是充分条件，所以：

$$\left(0, \frac{1}{n-1}, \ldots, \frac{1}{n-1}\right)$$

是全局最优解。

6. 设 $f:\mathbb{R}\to\mathbb{R}$ 为凸函数，且 $\mathbb{R}_+ \subseteq \mathrm{dom}f$。定义其“滑动平均”函数

$$F(x)=\frac{1}{x}\int_0^x f\,(t)\,dt,\qquad \mathrm{dom}F\,=\mathbb{R}_{++}.$$

证明 F 为凸函数。（f 可微。）

证明:

由积分求导法则，

$$F'(x)=-\frac{1}{x^2}\int_0^x f\,(t)\,dt+\frac{1}{x}f(x).$$

再求一次导数得

$$\begin{aligned}F''(x)\quad&=\frac{2}{x^3}\int_0^x f\,(t)\,dt-\frac{2}{x^2}f(x)+\frac{1}{x}f'(x)\\&=\frac{2}{x^3}\int_0^x \big(f(t)-f(x)\big)\,dt+\frac{1}{x}f'(x)\\&=\frac{2}{x^3}\int_0^x \big(f(t)-f(x)-f'(x)(t-x)\big)\,dt.\end{aligned}$$

由于 f 凸且可微，对任意 $x,t\in\mathrm{dom}f$ 有不等式

$$f(t)\ \geq\ f(x)+f'(x)(t-x).$$

故被积函数

$$f(t)-f(x)-f'(x)(t-x)\ \geq\ 0,$$

从而

$$F''(x)\ =\ \frac{2}{x^3}\int_0^x \underbrace{\big(f(t)-f(x)-f'(x)(t-x)\big)}_{\geq 0}\ dt\ \geq\ 0.$$

因此 F 在 $\mathbb{R}_{++}$ 上二阶导非负，故 F 为凸函数。证毕。