第一章 —习题解答

1.2 给定函数$f: \mathscr{F} \subseteq \mathbb{R}^n \to \mathbb{R}$. 记$g(x) := -f(x)$. 假设优化问题$\min_{x\in\mathscr{F}} f(x)$有最优解. 试证: 优化问题$\max_{x\in\mathscr{F}} g(x)$也有最优解, 并且两个优化问题的最优解集相同、最优值互为相反数.

证明: 设x^*是优化问题$\min_{x\in\mathscr{F}} f(x)$的一个最优解, 则对$\forall x \in \mathscr{F}$, 都有$f(x^*) \leq f(x)$, 即$-f(x^*) \geq -f(x)$. 因为$g(x) := -f(x)$, 故对$\forall x \in \mathscr{F}$, 都有$g(x^*) \geq g(x)$. 所以$x^*$是优化问题$\max_{x\in\mathscr{F}} g(x)$的最优解. 同理可得$\max_{x\in\mathscr{F}} g(x)$的最优解也是$\min_{x\in\mathscr{F}} f(x)$的最优解. 因此, 优化问题$\max_{x\in\mathscr{F}} g(x)$有最优解, 并且两个优化问题的最优解集相同.

另外, $f(x^*)$是优化问题$\min_{x\in\mathscr{F}} f(x)$的最优值, $g(x^*)$是优化问题$\max_{x\in\mathscr{F}} g(x)$的最优值, 有$g(x^*) = -f(x^*)$, 所以两个优化问题的最优值互为相反数.

1.3 验证$x^* = (0, -3)^\top$ 是优化问题

$$\begin{aligned} \min \quad & f(x) = x_1^2 + x_2 \\ \text{s.t.} \quad & x_1^2 + x_2^2 = 9 \end{aligned}$$

的严格全局最优解.

解: 根据约束条件$x_1^2 + x_2^2 = 9$ 得$x_1^2 = 9 - x_2^2$, 带入目标函数, 原问题转化为

$$\begin{aligned} \min \quad & g(x_2) = -x_2^2 + x_2 + 9 \\ \text{s.t.} \quad & -3 \le x_2 \le 3 \end{aligned}$$

这是一个单变量二次函数求极小的问题, 易得当$x_2 = -3$ 时, 函数取得极小值，此时$x_1 = 0$, 且$\forall x_2 \in (-3, 3]$ 都有$g(x_2) > g(-3)$, 所以$x^* = (0, -3)^\top$ 是严格全局最优解.

第一章 —习题解答

1.4 验证$x^* = (1, -1)^\top$ 是优化问题

$$\min f(x) = \left(x_2 - x_1^2\right)^2 + (1 + x_1)^2$$

的严格全局最优解.

解: 由于对任意的$x \in \mathbb{R}^2$都有$\left(x_2 - x_1^2\right)^2 + (1 + x_1)^2 \geq 0$, 并且

$$\left(x_2 - x_1^2\right)^2 + (1 + x_1)^2 = 0 \quad \Longleftrightarrow \quad \begin{cases} x_2 - x_1^2)^2 = 0 \\ (1 + x_1)^2 = 0 \end{cases}$$

$$\Longleftrightarrow \quad x = (1, -1)^\top.$$

因此，命题得证.

1.5 试证明: 对任意的$x \in \mathbb{R}^n$, 有

(1) $\|x\|_\infty \le \|x\|_1 \le n\|x\|_\infty$;

(2) $\|x\|_2 \le \|x\|_1 \le \sqrt{n}\|x\|_2$;

(3) $\|x\|_\infty \le \|x\|_2 \le \sqrt{n}\|x\|_\infty$.

证明: (1) 对任意的$x \in \mathbb{R}^n$, 有

$$\max_{i\in\{1,2,\dots,n\}} |x_i| \le \sum_{i=1}^{n} |x_i| \le n \max_{i\in\{1,2,\dots,n\}} |x_i|,$$

所以, 由范数的定义有$\|x\|_\infty \le \|x\|_1 \le n\|x\|_\infty$.

(2) 对任意的$x \in \mathbb{R}^n$, 一方面，由

$$\left(\sum_{i=1}^{n}|x_i|\right)^2 = \sum_{i=1}^{n}|x_i|^2 + \sum_{i\neq j}|x_i||x_j| \geq \sum_{i=1}^{n}|x_i|^2$$

可得：$\|x\|_2 = \sqrt{\sum_{i=1}^{n}|x_i|^2} \leq \sum_{i=1}^{n}|x_i| = \|x\|_1$; 另一方面，由于$|x_i||x_j| \leq \frac{1}{2}(|x_i|^2 + |x_j|^2)$, 所以$\sum_{i\neq j}|x_i||x_j| \leq (n-1)(|x_i|^2 + |x_j|^2)$. 因此，

$$\left(\sum_{i=1}^{n}|x_i|\right)^2 = \sum_{i=1}^{n}|x_i|^2 + \sum_{i\neq j}|x_i||x_j| \leq n\sum_{i=1}^{n}|x_i|^2,$$

即：$\|x\|_1 \leq \sqrt{n}\|x\|_2$.

(3) 对任意的$x \in \mathbb{R}^n$, 有

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n}|x_i|^2} \geq \sqrt{\max_{i\in\{1,2,\dots,n\}}|x_i|^2} = \|x\|_\infty,$$

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n}|x_i|^2} \leq \sqrt{n\max_{i\in\{1,2,\dots,n\}}|x_i|^2} = \sqrt{n}\|x\|_\infty.$$

1.6 设实对称矩阵$A \in \mathbb{R}^{n\times n}$的特征值为$\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$. 证明: 对任意的$x \in \mathbb{R}^n$有$\lambda_1\|x\|_2^2 \le x^\top Ax \le \lambda_n\|x\|_2^2$.

证明: 由于实对称矩阵$A \in \mathbb{R}^{n\times n}$的特征值为$\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$, 所以存在正交矩阵$Q$使得

$$A = Q^\top \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} Q := Q^\top \Lambda Q.$$

第一章 —习题解答

由于任意正交矩阵的逆矩阵也是正交矩阵，所以, 对任意的$x \in \mathbb{R}^n$, 有

$$\|Qx\|^2 = x^\top Q^\top Qx = x^\top x = \|x\|_2^2,$$

所以, 对任意的$x \in \mathbb{R}^n$, 有

$$x^\top Ax = x^\top Q^\top \Lambda Qx = \sum_{i=1}^{n} \lambda_i \left[(Qx)_i\right]^2 .$$

进而, 对任意的$x \in \mathbb{R}^n$, 有

$$x^\top Ax \geq \lambda_1 \textstyle\sum_{i=1}^{n} \left[(Qx)_i\right]^2 = \lambda_1 \|Qx\|_2^2 = \lambda_1 \|x\|_2^2,$$

$$x^\top Ax \leq \lambda_n \textstyle\sum_{i=1}^{n} \left[(Qx)_i\right]^2 = \lambda_n \|Qx\|_2^2 = \lambda_n \|x\|_2^2$$

1.7 设$A \in \mathbb{R}^{n\times n}$. 证明: $\|A\|_2 = \sqrt{\lambda_{\max}(A^\top A)}$, 其中$\lambda_{\max}(A^\top A)$表示矩阵$A^\top A$的最大特征值.

证明: 由于$A^\top A$是实对称半正定矩阵，所以存在正交矩阵Q使得

$$A^\top A = Q^\top \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} Q := Q^\top \Lambda Q.$$

一方面，由上一题的结果, 有$x^\top A^\top A x \leq \lambda_n \|x\|_2^2$, 因此,

$$\|A\|_2 = \max_{x\in\mathbb{R}^n\setminus\{0\}} \frac{\|Ax\|_2}{\|x\|_2} \leq \sqrt{\lambda_n}.$$

另一方面, 选取$x^* \in \mathbb{R}^n \setminus \{0\}$使得$Qx^* = e^n$ (其中e^n表示一个n维向量，它的第n个分量为1，其它分量为0)，注意到Q是可逆的, 这样的x^*是存在的. 因此,

$$\begin{aligned}\|Ax^*\|_2^2 &= (x^*)^\top A^\top Ax^* = (x^*)^\top Q^\top \Lambda Qx^* = \sum_{i=1}^{n} \lambda_i \left[(Qx^*)_i\right]^2 \\ &= \lambda_n \sum_{i=1}^{n} \left[(Qx^*)_i\right]^2 = \lambda_n \|Qx^*\|_2^2 = \lambda_n \|x^*\|_2^2.\end{aligned}$$

所以,

$$\|A\|_2 = \max_{x \in \mathbb{R}^n \setminus \{0\}} \frac{\|Ax\|_2}{\|x\|_2} \geq \frac{\|Ax^*\|_2}{\|x^*\|_2} = \sqrt{\lambda_n}.$$

命题得证.

第一章 —习题解答

1.8 设矩阵$G \in \mathbb{R}^{n\times n}$是对称正定矩阵, 对于向量$x \in \mathbb{R}^n$的椭球范数$\|x\|_G = \sqrt{x^\top Gx}$, 试证明: 对于任意的$x, y \in \mathbb{R}^n$，有

(1) $|x^\top y| \le \|x\|_G\|y\|_{G^{-1}}$, 当且仅当$x$和$G^{-1}y$线性相关时，等式成立;

(2) $|x^\top Gy| \le \|x\|_G\|y\|_G$, 当且仅当$x$和$y$线性相关时，等式成立.

证明: 由于G 对称正定, 故存在对称正定矩阵B, 使得$G = B^2$. 且$G^{-1} = B^{-2}, B^\top = B, B^{-\top} = B^{-1}$.

(1) $|x^\top y| = |x^\top BB^{-1}y| = |x^\top B^\top B^{-1}y| = |(Bx)^\top(B^{-1}y)|$, 根据Cauchy-Schwartz不等式, 有$|x^\top y| \le \|Bx\|\|B^{-1}y\| = \sqrt{x^\top B^\top Bx}\sqrt{y^\top B^{-\top}B^{-1}y} = \sqrt{x^\top B^2 x}\sqrt{y^\top B^{-2}y} = \sqrt{x^\top Gx}\sqrt{y^\top G^{-1}y} = \|x\|_G\|y\|_{G^{-1}}$, 其中, 等号成立当且仅当$Bx$ 和$B^{-1}y$ 线性相关, 即x 和$G^{-1}y$ 线性相关.

(2) 由(1)中的证明可得$|x^\top Gy| \le \|x\|_G\|Gy\|_{G^{-1}} = \|x\|_G\sqrt{y^\top G^\top G^{-1}Gy} = \|x\|_G\sqrt{y^\top Gy} = \|x\|_G\|y\|_G$, 其中, 等号成立当且仅当$x$ 和$G^{-1}(Gy)$ 线性相关, 即x 和y 线性相关.

第一章 —习题解答

1.12 设$\mathscr{F}$为凸集且函数$f: \mathscr{F} \subseteq \mathbb{R}^n \to \mathbb{R}$, 若对任意的点$x \in \mathscr{F}$及正数$\alpha$ 都有$f(\alpha x) = \alpha f(x)$, 则称$f$为$\mathscr{F}$上的正齐次函数. 试证明: $\mathscr{F}$上的正齐次函数f是凸函数的充要条件是对任意的$x, y \in \mathscr{F}$都有$f(x + y) \le f(x) + f(y)$.

证明: " $\Leftarrow$ " 由f 的正齐次性可知, 对任意的点$x \in \mathscr{F}$ 及正数α, 都有$\alpha x \in \mathscr{F}$ 且$f(\alpha x) = \alpha f(x)$. 由于对$\forall x, y \in \mathscr{F}$, 都有$f(x + y) \le f(x) + f(y)$, 那么, $\forall x, y \in \mathscr{F}, \alpha \in (0, 1), f(\alpha x + (1 - \alpha)y) \le f(\alpha x) + f((1 - \alpha)y) = \alpha f(x) + (1 - \alpha)f(y)$, 对于$\alpha = 0$ 和$\alpha = 1$ 的情形，上式显然成立, 根据凸函数的定义可知f 是凸函数.

" $\Rightarrow$ " 由于f 是凸函数，故$\forall x, y \in \mathscr{F}, \alpha \in [0, 1]$, 都有$f(\alpha x + (1-\alpha)y) \le \alpha f(x)+(1-\alpha)f(y)$. 取$\alpha = \frac{1}{2}$, 得$f(\frac{1}{2}(x+y)) \le \frac{1}{2}f(x)+\frac{1}{2}f(y)$, 又因为$f$ 是正齐次函数, 所以$f(\frac{1}{2}(x + y)) = \frac{1}{2}f(x + y)$, 故$f(x + y) \le f(x) + f(y)$.

1.13 假设f为$\mathbb{R}^n$上的凸函数. 试证: 如果f在某点$x^* \in \mathbb{R}^n$处取得全局最大值, 那么f在$\mathbb{R}^n$ 上恒为常数.

证明: 假设f 在$\mathbb{R}^n$ 上不恒为常数, 由于f 在$x^* \in \mathbb{R}^n$ 处取得全局最大值, 则存在$x_1 \in \mathbb{R}^n$, 使得$f(x_1) < f(x^*)$, 同时存在$\alpha \in (0,1)$, $x_2 \in \mathbb{R}^n$, 使得$x^* = \alpha x_1 + (1-\alpha)x_2$, 且$f(x_2) \leq f(x^*)$. 由于$f$ 是凸函数, 故$f(x^*) = f(\alpha x_1 + (1-\alpha)x_2) \leq \alpha f(x_1) + (1-\alpha)f(x_2) < \alpha f(x^*) + (1-\alpha)f(x^*) = f(x^*)$, 导出矛盾, 故$f$ 在$\mathbb{R}^n$上恒为常数.

1.16 判断下列优化问题是否是凸规划问题:

(1)
$$\begin{aligned} \min \quad & f(x) = 2x_1^2 + x_2^2 + x_3^2 \\ \text{s.t.} \quad & x_1^2 + x_2^2 \le 9, \\ & 5x_1 + x_3 = 10, \\ & x_1, x_2, x_3 \ge 0. \end{aligned}$$

(2)
$$\begin{aligned} \min \quad & f(x) = x_1^4 - 4x_1x_2 + 3x_2^2 \\ \text{s.t.} \quad & x_1 + x_2 \le 5, \\ & x_1 - x_2 \le 2, \\ & x_1, x_2 \ge 0. \end{aligned}$$

解：(1) 目标函数是一个二次函数，且它的Hesse阵是一个正定矩阵，因此目标函数是凸函数. 可行域是几个凸集的交集，因此是凸集. 所以，此问题是一个凸规划.

(2) 显然，可行域是凸集. 目标函数的Hesse阵为

$$\nabla^2 f(x_1, x_2) = \begin{pmatrix} 12x_1^2 & -4 \\ -4 & 6 \end{pmatrix},$$

其行列式为$det(\nabla^2 f(x_1, x_2)) = 72x_1^2 - 16$, 所以当$x_1^2 < \frac{2}{9}$时, 其行列式的值小于零. 取点$x^* = (\frac{1}{3}, 1)$, 则$x^*$是可行点, 且目标函数的Hesse阵在此点处不是半正定的. 因此，此问题不是一个凸规划.

第一章 —习题解答