

# Homework 1

章梦华 52214404016

2022 年 3 月 11 日

## 目录

1 凸集	1
2 凸函数	3
3 带 $l_2$ 惩罚的部分优化问题	5
4 Lipschitz 梯度与强凸性	6
5 实践：使用 CVXPY 解优化问题	8

## 1 凸集

a. 证明若函数  $f$  二次可微, 则  $f$  为凸函数的充要条件为:

- $\text{dom } f$  为凸集
- $\nabla^2 f(x) \succeq 0$ , for all  $x \in \text{dom } f$

证明. First, consider the case  $n = 1$ . We show that a twice differentiable function  $f : R \rightarrow R$  is convex if and only if  $f''(x) \geq 0$  for all  $x \in \text{dom } f$ . Assume first that  $f$  is convex and  $x, y \in \text{dom } f$ . Since  $\text{dom } f$  is convex,  $z = x + t(y - x) \in \text{dom } f$  for all  $0 < t \leq 1$ . Due to the first order convexity condition, we have

$$\begin{aligned} f(x) &\geq f(z) + f'(z)(x - z) \\ f(z) &\geq f(x) + f'(x)(z - x) \end{aligned}$$

The sum of the above two inequalities gives  $f(x) + f(z) \geq f(z) + f(x) + (x - z)(f'(z) - f'(x))$ , i.e.  $(z - x)(f'(z) - f'(x)) \geq 0$ . Divide both sides by  $(z - x)^2$  and we obtain

$$\frac{f'(z) - f'(x)}{z - x} \geq 0$$

and taking the limit as  $t \rightarrow 0$  yields  $f''(x) \geq 0$ .

To show sufficiency, assume that the function satisfies  $f''(x) \geq 0$  for all  $x \in \text{dom } f$ . By the mean value version of Taylor's theorem we have

$$f(y) = f(x) + f'(x)(y - x) + f''(z)(y - x)^2, \text{ for some } z \text{ between } y \text{ and } x.$$

Since  $f''(z) \geq 0$ , it turns out that  $f(y) \geq f(x) + f'(x)(y - x)$ . Thus,  $f$  is convex.

Now we can prove the general case, with  $f : R^n \rightarrow R$ . Let  $x, y \in R^n$  and consider  $f$  restricted to the line passing through them, i.e. the function defined by  $g(t) = f(x + t(y - x))$ , so  $g'(t) = \nabla f(x + t(y - x))^\top (y - x)$  and  $g''(t) = (y - x)^\top \nabla^2 f(x + t(y - x))(y - x)$

First assume  $f$  is convex, which implies  $g$  is convex. Due to the first order convexity condition we have

$$\begin{aligned} g(t) &\geq g(0) + g'(0)t \\ g(0) &\geq g(t) + g'(t)(-t). \end{aligned}$$

Similarly, the sum of the above two inequalities gives that  $(g'(t) - g'(0))t \geq 0$ . Divide both sides by  $t^2$  and we obtain

$$\frac{g'(t) - g'(0)}{t} \geq 0$$

and taking the limit as  $t \rightarrow 0$  yields  $g''(0) \geq 0$ . Notice that  $g''(0) = (y - x)^\top \nabla^2 f(x)(y - x) \geq 0$  holds for all  $y \in \text{dom } f$ , which implies  $\nabla^2 f(x)$  is semi positive definite. Thus,  $\nabla^2 f(x) \succeq 0$ .

Now assume that  $\nabla^2 f(x) \succeq 0$  for all  $x \in \text{dom } f$ . By the mean value version of Taylor's theorem we have

$$g(1) = g(0) + g'(0) + \frac{1}{2}g''(t), \text{ for some } t \text{ between } 0 \text{ and } 1.$$

Because  $\nabla^2 f(x) \succeq 0$  for all  $x \in \text{dom } f$ ,  $g''(t) = (y - x)^\top \nabla^2 f(x + t(y - x))(y - x) \geq 0$ . And then  $g(1) \geq g(0) + g'(0)$ , which is equivalent to  $f(y) \geq f(x) + \nabla f(x)^\top (y - x)$ . Hence,  $f$  is convex.  $\square$

## b. 闭集与凸集

b.1 证明多面体  $\{x : Ax \leq b\}$ ,  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$  为闭的凸集

证明. First, show that the polyhedron is closed. Assume a sequence of points  $x_n \in \{x : Ax \leq b\}$ . If this sequence has a limit point  $x$ , it is easy to see that  $Ax = \lim_{n \rightarrow \infty} Ax_n \leq b$  implying  $x$  is also in the polyhedron. Therefore, the polyhedron  $\{x : Ax \leq b\}$  is closed.

Next, show the convexity of polyhedron. If  $x, y \in \{x : Ax \leq b\}$ ,  $Ax \leq b$  and  $Ay \leq b$ . Denote  $z = \theta x + (1 - \theta)y$  where  $\theta \in [0, 1]$ .

$$Az = \theta Ax + (1 - \theta)Ay \leq \theta b + (1 - \theta)b = b.$$

Thus,  $z \in \{x : Ax \leq b\}$  and  $\{x : Ax \leq b\}$  is a convex set.

In conclusion,  $\{x : Ax \leq b\}$  is a closed convex set.  $\square$

b.2 举例:  $\mathbb{R}^2$  上的闭集的凸包不一定是闭的

Example: The convex hull of closed set

$$\left\{ (x, y) : y \geq \frac{1}{1+x^2} \right\}$$

is an open set  $\{(x, y) : y > 0\}$ .

b.3 若  $A \in \mathbb{R}^{m \times n}$ ,  $S \in \mathbb{R}^m$  为凸集, 证明集合  $S$  在  $A$  下的原像  $\{x \in \mathbb{R}^n : Ax \in S\}$  是凸集

证明. Assume  $x, y \in \{x \in \mathbb{R}^n : Ax \in S\}$ . Let  $z = \theta x + (1 - \theta)y$  where  $\theta \in [0, 1]$ . Since  $S$  is convex and  $Ax, Ay$  are both in  $S$ ,  $Az = \theta Ax + (1 - \theta)Ay \in S$ . So  $z \in \{x : Ax \in S\}$  and  $\{x \in \mathbb{R}^n : Ax \in S\}$  is convex.  $\square$

b.4 若  $A \in \mathbb{R}^{m \times n}$ ,  $S \in \mathbb{R}^n$  为凸集, 证明集合  $S$  在  $A$  下的像  $\{Ax : x \in S\}$  是凸集

证明. Assume  $x, y \in \{Ax : x \in S\}$ . There exist  $x_0, y_0 \in S$  such that  $x = Ax_0, y = Ay_0$ . Write  $z = \theta x + (1 - \theta)y, \theta \in [0, 1]$ . We can see that  $z = \theta Ax_0 + (1 - \theta)Ay_0 = A(\theta x_0 + (1 - \theta)y_0)$ . Since  $S$  is convex,  $\theta x_0 + (1 - \theta)y_0 \in S$ . As a result,  $z \in \{Ax : x \in S\}$  and  $\{Ax : x \in S\}$  is convex.  $\square$

b.5 举例: 存在  $A \in \mathbb{R}^{m \times n}$  及闭凸集  $S \subseteq \mathbb{R}^n$ , 使得  $A(S)$  不是闭集

Example: Let  $S = \{(x, y)^\top : xy \geq 1, x > 0\}$  and  $A = (1, 0)$ . Obviously,  $S$  is a closed convex set. But  $A(S) = \{x : x > 0\}$  is not closed.

## c. 多面体

c.1 证明若  $P \subseteq \mathbb{R}^n$  为多面体, 则  $A(P)$  为多面体, 提示: 可使用以下事实:

$P \subseteq \mathbb{R}^{m+n}$  为多面体  $\Rightarrow \{x \in \mathbb{R}^n : (x, y) \in P \text{ for some } y \in \mathbb{R}^m\}$  是多面体

证明. Write  $P = \{x \in \mathbb{R}^n : Bx \leq b\}$ , where  $B \in \mathbb{R}^{k \times n}$ . Set  $Q = \{(x, y) \in \mathbb{R}^{m+n} : Bx \leq b, y = Ax\}$ .  $Q$  can be expressed as  $\{z \in \mathbb{R}^{m+n} : Cz \leq c\}$ , where

$$C = \begin{pmatrix} B & 0 \\ A & -1_m \\ -A & 1_m \end{pmatrix}$$

and  $c = (b, 0_m, 0_m)$ . So  $Q$  is also a polyhedron. Use the hint and we can obtain  $A(P) = \{y \in \mathbb{R}^m : (x, y) \in Q \text{ for some } x \in \mathbb{R}^n\}$  is a polyhedron.  $\square$

c.2 证明若  $Q \subseteq \mathbb{R}^m$  为多面体,  $A \in \mathbb{R}^{m \times n}$ , 则  $A^{-1}(Q)$  为多面体.

证明. Write  $Q = \{x \in \mathbb{R}^m : Bx \leq b\}$ , where  $B \in \mathbb{R}^{k \times m}$ . Set  $P = \{(x, y) \in \mathbb{R}^{m+n} : Bx \leq b, Ay = x\}$ .  $P$  can be expressed as  $\{z \in \mathbb{R}^{m+n} : Cz \leq c\}$ , where

$$C = \begin{pmatrix} B & 0 \\ -1_m & A \\ 1_m & -A \end{pmatrix}$$

and  $c = (b, 0_m, 0_m)$ . So  $P$  is also a polyhedron. Use the hint and we can obtain  $A^{-1}(Q) = \{y \in \mathbb{R}^m : (x, y) \in P \text{ for some } x \in \mathbb{R}^n\}$  is a polyhedron.  $\square$

## 2 凸函数

a. 证明熵函数:

$$f(x) = -\sum_{i=1}^n x_i \log(x_i), \text{ dom } f = \left\{x \in \mathbb{R}_{++}^n : \sum_{i=1}^n x_i = 1\right\}$$

是严格凹的.

证明. Clearly,  $\text{dom } f$  is convex. The first order derivative of  $f(x)$  is  $\nabla_i f(x) = -\log x_i - 1$  and the second order derivative is

$$\nabla_{ij}^2 f(x) = -\frac{1}{x_i} I(i=j).$$

So the Hessian matrix of  $f(x)$  is diagonal and each diagonal element is negative. Thus,  $\nabla^2 f(x)$  is negative definite and the entropy function  $f(x)$  is strictly concave.  $\square$

b. 若  $f$  为二次可微函数且  $\text{dom } f$  为凸集, 证明  $f$  为凸函数的充要条件为:

$$(\nabla f(x) - \nabla f(y))^\top (x - y) \geq 0, \forall x, y$$

这被称为梯度  $\nabla f$  的单调性.

证明. Set  $g(t) = f(x + t(y - x))$ , so  $g'(t) = \nabla f(x + t(y - x))^\top (y - x)$  and  $g''(t) = (y - x)^\top \nabla^2 f(x + t(y - x))(y - x)$ . First, show necessity. Assume  $f$  is convex, which implies  $g$  is convex. Due to the second-order convexity condition,  $g''(t) \geq 0$  for all  $t \in [0, 1]$ . So the first order derivative  $g'(t)$  is non-decreasing in its domain and we have  $g'(1) \geq g'(0)$ , i.e.

$$\begin{aligned} \nabla f(y)^\top (y - x) &\geq \nabla f(x)^\top (y - x) \\ \Rightarrow (\nabla f(x) - \nabla f(y))^\top (x - y) &\geq 0. \end{aligned}$$

Next, show sufficiency. Assume

$$(\nabla f(x) - \nabla f(y))^\top (x - y) \geq 0, \forall x, y.$$

Since  $f$  is twice differentiable,  $g$  is also twice differentiable. By the assumption, we have

$$(\nabla f(x + t(y - x)) - \nabla f(x))^\top (t(y - x)) \geq 0.$$

Notice that

$$t(g'(t) - g'(0)) = (\nabla f(x + t(y - x)) - \nabla f(x))^\top (t(y - x)).$$

So  $t(g'(t) - g'(0)) \geq 0$  for  $\forall t \in [0, 1]$ . Divide both sides by  $t^2$  and we obtain

$$\frac{g'(t) - g'(0)}{t} \geq 0$$

and taking the limit as  $t \rightarrow 0$  yields  $g''(0) \geq 0$ .  $g''(0) = (y - x)^\top \nabla^2 f(x)(y - x) \geq 0$  for  $\forall x, y \in \text{dom } f$  implying that  $\nabla^2 f(x) \succeq 0$  for  $\forall x \in \text{dom } f$ . Hence,  $f$  is convex.  $\square$

c. 举例: 严格凸函数并不一定能达到其最小值.

Example:  $f(x) = 1/x$ ,  $\text{dom } f = \{x \in \mathbb{R} : x > 0\}$  is strictly convex. But the minimal value is not attainable.

d. 函数  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  被称为强制的 (coercive), 如果当  $\|x\|_2 \rightarrow \infty$  时, 有  $f \rightarrow \infty$ . 强制函数的一个关键事实为其可以达到极小值. 证明一个二次可微的强凸函数是强制的 (coercive), 并因此可达到极小值.

证明. Assume  $f$  is  $m$ -strongly convex, i.e.  $g(x) = f(x) - m/2\|x\|_2^2$  is convex for some  $m > 0$ . By the first-order convexity condition,

$$\begin{aligned} g(y) &\geq g(x) + \nabla g(x)^\top (y - x) \\ f(y) - \frac{m}{2}\|y\|_2^2 &\geq f(x) - \frac{m}{2}\|x\|_2^2 + (\nabla f(x) - mx)^\top (y - x). \end{aligned}$$

Let  $\alpha = \nabla f(x) - mx$  and  $K_x = f(x) - m/2\|x\|_2^2 - \alpha^\top x$ . Then

$$f(y) \geq \frac{m}{2}\|y\|_2^2 + \alpha^\top y + K_x.$$

By Cauchy-Schwartz inequality,  $\alpha^\top y \geq -\|\alpha\|_2\|y\|_2$ . Hence,

$$f(y) \geq \frac{m}{2}\|y\|_2^2 - \|\alpha\|_2\|y\|_2 + K_x.$$

The right hand side goes to infinity when  $\|y\|_2 \rightarrow \infty$ . So  $f$  is coercive.  $\square$

e. 证明在有界多面体上的凸函数的最大值一定在其中一个顶点上. 提示: 已知一个有界的多面体可以被表示为其顶点的凸组合.

证明. Denote  $a_i$ ,  $i = 1, \dots, n$  as vertices of the bounded polyhedron  $P$  and

$$P = \left\{ x \in \mathbb{R}^m : x = \sum_{i=1}^n \lambda_i a_i, \lambda_i \geq 0 \text{ and } \sum_{i=1}^n \lambda_i = 1 \right\}.$$

Let  $f$  be the convex function defined in  $P$ . Assume that  $x_0$  is the point at which  $f$  attains its maximal value.

There exist  $\lambda_{i0} \geq 0$  such that  $\sum_{i=1}^n \lambda_{i0} = 1$  and  $x_0 = \sum_{i=1}^n \lambda_{i0} a_i$ . Since  $f$  is convex,

$$f(x_0) = f\left(\sum_{i=1}^n \lambda_{i0} a_i\right) \leq \sum_{i=1}^n \lambda_{i0} f(a_i) \leq \sum_{i=1}^n \lambda_{i0} f(x_0) = f(x_0),$$

hence the equality holds here. Because  $f(a_i) \leq f(x_0)$  for all  $1 \leq i \leq n$ ,  $f(a_i) = f(x_0)$  for those  $i$  with  $\lambda_i \neq 0$  otherwise the equality doesn't hold. Since there is at least one such  $i$ , the claim follows.  $\square$

### 3 带 $l_2$ 惩罚的部分优化问题

考虑问题

$$\min_{\beta, \sigma \geq 0} f(\beta) + \frac{\lambda}{2} \sum_{i=1}^n g(\beta_i, \sigma_i), \quad (1)$$

其中  $f$  为定义在  $\mathbb{R}^n$  上的凸函数,  $\lambda \geq 0$ , 且

$$g(x, y) = \begin{cases} x^2/y + y & \text{if } y > 0; \\ 0 & \text{if } x = 0, y = 0; \\ \infty & \text{else.} \end{cases}$$

- a. 证明  $g$  是凸函数, 即上述问题为凸优化问题. (后面我们可根据此进行部分优化, 且部分优化后的函数也是凸函数)

证明. First, the domain of  $g$  is convex. When  $y > 0$ ,  $g(x, y) = P(x^2, (0, 1)(x, y)^\top) + (0, 1)(x, y)^\top$ , where  $P$  is the perspective transformation. The affine mapping  $(0, 1)(x, y)^\top$  and  $x^2$  are both convex, so  $P(x^2, (0, 1)(x, y)^\top)$  is also convex. And the sum of two convex function is still convex implying  $g(x, y)$  is convex, i.e.

$$\min_{y \geq 0} g(x, y) = 2|x|.$$

□

- b. 证明:

$$\min_{y \geq 0} g(x, y) = 2|x|.$$

证明. Given  $x$ , the first order partial derivative of  $g(x, y)$  is

$$\frac{\partial}{\partial y} g(x, y) = 1 - \frac{x^2}{y^2}.$$

Let  $\partial g(x, y)/\partial y = 0$  and we have  $y = |x|$  since  $y \geq 0$ . In question a., we've known  $g(x, y)$  is convex, so it attains its minimum at  $y = |x|$  when  $x$  is fixed. □

- c. 证明 (1) 中对于  $\sigma \geq 0$  的优化可得  $\ell_1$  惩罚问题

$$\min_{\beta} f(\beta) + \lambda \|\beta\|_1.$$

证明. We can optimize over  $\sigma$  first and then minimize the question over  $\beta$ . The result in question b. gives that  $g(\beta_i, \sigma_i)$  attains the minimum  $2|\beta_i|$  when  $\sigma_i = |\beta_i|$ . Thus,

$$\min_{\beta, \sigma \geq 0} f(\beta) + \frac{\lambda}{2} \sum_{i=1}^n g(\beta_i, \sigma_i),$$

is equivalent to

$$\min_{\beta} f(\beta) + \frac{\lambda}{2} \sum_{i=1}^n |\beta_i| = \min_{\beta} f(\beta) + \lambda \|\beta\|_1.$$

□

## 4 Lipschitz 梯度与强凸性

令  $f$  为二次连续可微的凸函数

a. 证明以下命题等价:

- i.  $\nabla f$  为  $L$ -Lipschitz 函数
- ii. 对任意  $x, y$ ,  $(\nabla f(x) - \nabla f(y))^\top (x - y) \leq L\|x - y\|_2^2$
- iii. 对任意  $x$ ,  $\nabla^2 f(x) \preceq LI$
- iv. 对任意  $x, y$ ,  $f(y) \leq f(x) + \nabla f(x)^\top (y - x) + \frac{L}{2}\|y - x\|_2^2$

循环证明  $i \Rightarrow ii$ ,  $ii \Rightarrow iii$ ,  $iii \Rightarrow iv$ ,  $iv \Rightarrow ii$ ,  $iii \Rightarrow i$ .

( $i \Rightarrow ii$ ). Since  $\nabla f$  is  $L$ -Lipschitz function,

$$\|\nabla f(x) - \nabla f(y)\|_2 \leq L\|x - y\|_2 \text{ for all } x, y.$$

Multiply both sides by  $\|x - y\|_2$  and we obtain

$$\|\nabla f(x) - \nabla f(y)\|_2 \|x - y\|_2 \leq L\|x - y\|_2^2.$$

Cauchy-Schwarz inequality guarantees

$$\|\nabla f(x) - \nabla f(y)\|_2 \|x - y\|_2 \geq (\nabla f(x) - \nabla f(y))^\top (x - y).$$

Thus, ii. holds.

( $ii \Rightarrow iii$ ). Let  $x = y + t\alpha$ , where  $t \in \mathbb{R}$  and  $\alpha \in \mathbb{R}^n$ ,  $\|\alpha\|_2 = 1$ . Apply Taylor's expansion and we have

$$\nabla f(x) = \nabla f(y) + \nabla^2 f(y)(x - y) + o(\|x - y\|).$$

According to result in (ii),

$$\begin{aligned} (\nabla f(x) - \nabla f(y))^\top (x - y) &= (\nabla^2 f(y)(x - y))^\top (x - y) + o(\|x - y\|_2^2) \leq L\|x - y\|_2^2 \\ &\Rightarrow t^2 \alpha^\top \nabla^2 f(y) \alpha + o(t^2 \|\alpha\|_2^2) \leq L(t^2 \|\alpha\|_1^2) \\ &\Rightarrow \alpha^\top \nabla^2 f(y) \alpha + o(1) \leq L \quad (\text{since } \|\alpha\|_2 = 1). \end{aligned}$$

Let  $\alpha$  be the eigenvector of the maximal eigenvalue of  $\nabla^2 f(y)$  and taking  $t \rightarrow 0$  gives  $\lambda_{\max} \nabla^2 f(y) \leq L$ . Hence,  $\nabla^2 f(x) \preceq LI$ .

( $iii \Rightarrow iv$ ). By mean value version of Taylor's theorem,

$$f(y) = f(x) + \nabla f(x)^\top (y - x) + \frac{1}{2}(y - x)^\top \nabla^2 f(\xi)(y - x), \text{ where } \xi \text{ between } x \text{ and } y.$$

Since  $\nabla^2 f(\xi) \preceq LI$  confirmed by (iii),

$$f(y) \leq f(x) + \nabla f(x)^\top (y - x) + \frac{L}{2}\|y - x\|_2^2.$$

(iv  $\Rightarrow$  ii). From (iv), we have

$$f(y) \leq f(x) + \nabla f(x)^\top (y - x) + \frac{L}{2} \|y - x\|_2^2,$$

$$f(x) \leq f(y) + \nabla f(y)^\top (x - y) + \frac{L}{2} \|x - y\|_2^2.$$

The sum of the two inequalities gives

$$\begin{aligned} f(y) + f(x) &\leq f(x) + f(y) + (\nabla f(y) - \nabla f(x))^\top (x - y) + L\|x - y\|_2^2 \\ \Rightarrow (\nabla f(x) - \nabla f(y))^\top (x - y) &\leq L\|x - y\|_2^2. \end{aligned}$$

(iii  $\Rightarrow$  i). Apply the mean value version of Taylor's theorem and then

$$\nabla f(x) - \nabla f(y) = \nabla^2 f(\xi)(x - y), \quad \text{where } \xi \text{ between } x \text{ and } y.$$

Taking the norm of both sides gives

$$\|\nabla f(x) - \nabla f(y)\|_2 = \|\nabla^2 f(\xi)(x - y)\|_2.$$

Since  $\nabla^2 f(x) \preceq LI$  for  $\forall x$ ,  $\nabla^2 f(\xi) \preceq LI$ . By Cauchy-Schwarz inequality, we have

$$\|\nabla f(x) - \nabla f(y)\|_2 = \|\nabla^2 f(\xi)(x - y)\|_2 \leq \|\nabla^2 f(\xi)\|_2 \|x - y\|_2 \leq L\|x - y\|_2.$$

b. 证明以下命题等价:

i.  $f$  为  $m$ -强凸函数

ii. 对任意  $x, y$ ,  $(\nabla f(x) - \nabla f(y))^\top (x - y) \geq m\|x - y\|_2^2$

iii. 对任意  $x$ ,  $\nabla^2 f(x) \succeq mI$

iv. 对任意  $x, y$ ,  $f(y) \geq f(x) + \nabla f(x)^\top (y - x) + \frac{m}{2} \|y - x\|_2^2$

循环证明  $i \Rightarrow ii$ ,  $ii \Rightarrow iii$ ,  $iii \Rightarrow iv$ ,  $iv \Rightarrow i$ .

(i  $\Rightarrow$  ii). Since  $f$  is  $m$ -strongly convex,  $g(x) = f(x) - m/2\|x\|_2^2$  is convex for some  $m > 0$ .  $\nabla g(x) = \nabla f(x) - mx$ .

It follows from the monotone gradient condition for convexity of  $g(x)$ , i.e.

$$\begin{aligned} (\nabla g(x) - \nabla g(y))^\top (x - y) &\geq 0 \\ \Rightarrow (\nabla f(x) - \nabla f(y))^\top (x - y) &\geq m\|x - y\|_2^2. \end{aligned}$$

(ii  $\Rightarrow$  iii). Let  $x = t\alpha + y$ , where  $t \in \mathbb{R}$  and  $\alpha \in \mathbb{R}^n$ ,  $\|\alpha\| = 1$ . Use Taylor's theorem and we have

$$\nabla f(x) - \nabla f(y) = \nabla^2 f(y)(x - y) + o(\|x - y\|).$$

Then multiplying both sides by  $x - y$  gives

$$\begin{aligned} (\nabla f(x) - \nabla f(y))^\top (x - y) &= (x - y)^\top \nabla^2 f(y)(x - y) + o(\|x - y\|_2^2) \geq m\|x - y\|_2^2 \\ \Rightarrow t^2 \alpha^\top \nabla^2 f(y) \alpha + o(t^2 \|\alpha\|_2^2) &\geq mt^2 \|\alpha\|_2^2 \\ \Rightarrow \alpha^\top \nabla^2 f(y) \alpha + o(1) &\geq m \end{aligned}$$

Let  $\alpha$  be the eigenvector of the minimal eigenvalue of  $\nabla^2 f(y)$  and taking  $t \rightarrow 0$  gives  $\lambda_{\min} \nabla^2 f(y) \geq m$ .

Hence,  $\nabla^2 f(x) \succeq mI$ .



(iii  $\Rightarrow$  iv). Let  $g(x) = f(x) - \frac{m}{2}\|x\|_2^2$ . Since  $\nabla^2 g(x) = \nabla^2 f(x) - mI \succeq 0$ ,  $g(x)$  is convex. The convexity of  $g(x)$  gives

$$\begin{aligned} g(y) &\geq g(x) + \nabla g(x)^\top (y - x) \\ &\Leftrightarrow f(y) - \frac{m}{2}\|y\|_2^2 \geq f(x) - \frac{m}{2}\|x\|_2^2 + (\nabla f(x) - mx)^\top (y - x) \\ &\Leftrightarrow f(y) \geq f(x) + \nabla f(x)^\top (y - x) + \frac{m}{2}\|x - y\|_2^2. \end{aligned}$$

(iv  $\Rightarrow$  i). Let  $g(x) = f(x) - \frac{m}{2}\|x\|_2^2$ .  $f(y) \geq f(x) + \nabla f(x)^\top (y - x) + \frac{m}{2}\|x - y\|_2^2$  implies that  $g(y) \geq g(x) + \nabla g(x)^\top (y - x)$ , i.e.  $g(x)$  is convex. Thus,  $f(x)$  is  $m$ -strongly convex.

## 5 实践：使用 CVXPY 解优化问题

见 “CVXPY 实践结果.pdf” 文件