Reference sol to HW 5.

1. proof: we will argue by contradiction.

suppose $\vec{x}^*$ is an optimal solution of (P) that satisfies $g(\vec{x}^*)<0$, but it is not an optimal solution of the problem.

$$\begin{aligned} &\min f(\vec{x}) \\ &\text{s.t. } \vec{x} \in X. \end{aligned}$$

them we can find $\vec{y} \in X$ and $g(\vec{y}) > 0$. s.t.

$$f(\vec{y}) < f(\vec{x}^*).$$

On the other hand, $g$ is convex over $\mathbb{R}^n$, then we obtain $g$ is continuous over $\mathbb{R}^n$. Since $g(\vec{x}^*)<0$ and $g(\vec{y})>0$, thus we can find. $\vec{z} \in (\vec{x}^*, \vec{y})$ s.t.

$$g(\vec{z}) = 0. \quad \cdots (1)$$

since $X$ is convex, $\vec{x}^*, \vec{y} \in X$, then $(\vec{x}^*, \vec{y}) \subset X$, i.e. $\vec{z} \in X$. combining with (1), we further have $\vec{z} \in X \cap \{\vec{x} : g(\vec{x}) \leq 0\}$.

Without lose of generality, we can let

$$\vec{z} = \lambda \vec{x}^* + (1-\lambda)\vec{y}, \quad \lambda \in (0,1)$$

given $\vec{z} \in (\vec{x}^*, \vec{y})$. Then

$$\begin{aligned} f(\vec{z}) = f(\lambda \vec{x}^* + (1-\lambda)\vec{y}) &\leq \lambda f(\vec{x}^*) + (1-\lambda) f(\vec{y}) \\ &< \lambda f(\vec{x}^*) + (1-\lambda) f(\vec{x}^*) \\ &= f(\vec{x}^*) \end{aligned}$$

this is a contradiction with $\vec{x}^*$ is an optimal solution of (P).

Thus, we conclude that $\vec{x}_*$ is also an optimal solution of the problem.

$$\begin{aligned} &\min f(\vec{x}) \\ &\text{s.t } \vec{x} \in X \end{aligned}$$

2. proof: argue by contradiction

suppose $A\vec{x}^* \neq A\vec{y}^*$, then let $\vec{z} = \frac{\vec{x}^* + \vec{y}^*}{2}$ we have

$$h(\vec{z}) = f\left(A\frac{\vec{x}^*+\vec{y}^*}{2}\right) + g\left(\frac{\vec{x}^*+\vec{y}^*}{2}\right)$$

$$= f\left(\tfrac{1}{2}A\vec{x}^* + \tfrac{1}{2}A\vec{y}^*\right) + g\left(\tfrac{1}{2}\vec{x}^* + \tfrac{1}{2}\vec{y}^*\right)$$

$$< \tfrac{1}{2}f(A\vec{x}^*) + \tfrac{1}{2}f(A\vec{y}^*) + \tfrac{1}{2}g(\vec{x}^*) + \tfrac{1}{2}g(\vec{y}^*)$$

$$= \tfrac{1}{2}h(\vec{x}^*) + \tfrac{1}{2}h(\vec{y}^*)$$

this is a contradiction with $\vec{x}^*$, $\vec{y}^*$ are optimal solutions. Thus we have $A\vec{x}^* = A\vec{y}^*$.

3. First the gradient of the Huber function is.

$$\nabla H_\mu(\vec{x}) = \begin{cases} \frac{\vec{x}}{\mu}, & \|\vec{x}\| \le \mu. \\ \frac{\vec{x}}{\|\vec{x}\|}, & \text{else.} \end{cases}$$

Case I: $\vec{x}, \vec{y} \in \{\vec{z} : \|\vec{z}\| \le \mu\}$, then

$$\|\nabla H_\mu(\vec{x}) - \nabla H_\mu(\vec{y})\| = \left\|\frac{\vec{x}}{\mu} - \frac{\vec{y}}{\mu}\right\| = \frac{1}{\mu}\|\vec{x} - \vec{y}\|$$

Case II: $\vec{x} \in \{\vec{z} : \|\vec{z}\| \le \mu\}$, $\vec{y} \in \{\vec{z} : \|\vec{z}\| > \mu\}$, then

$$\|\nabla H_\mu(\vec{x}) - \nabla H_\mu(\vec{y})\| = \left\|\frac{\vec{x}}{\mu} - \frac{\vec{y}}{\|\vec{y}\|}\right\|$$

since $\left\|\frac{\vec{x}}{\mu} - \frac{\vec{y}}{\|\vec{y}\|}\right\|^2 - \frac{1}{\mu^2}\|\vec{x} - \vec{y}\|^2 = 1 - 2\frac{\vec{x}^T\vec{y}}{\mu\|\vec{y}\|} - \frac{1}{\mu^2}\left(\|\vec{y}\|^2 - 2\vec{x}^T\vec{y}\right)$

$$= \frac{1}{\mu^2}\left[\mu^2 - 2\mu\cos\theta\|\vec{x}\| - \|\vec{y}\|^2 + 2\|\vec{x}\|\|\vec{y}\|\cos\theta\right]$$

$$= \frac{1}{\mu^2}(\|\vec{y}\| - \mu)\left[\underbrace{-(\|\vec{y}\| + \mu)}_{<-2\mu} + 2\|\vec{x}\|\cos\theta\right]$$

$$< 0$$

Thus $\|\nabla H_\mu(\vec{x}) - \nabla H_\mu(\vec{y})\| < \frac{1}{\mu}\|\vec{x} - \vec{y}\|$

Case III: $\vec{x} \in \{\vec{z} : \|\vec{z}\| > \mu\}$. $\vec{y} \in \{\vec{z} : \|\vec{z}\| \le \mu\}$, then

$$\|\nabla H_\mu(\vec{x}) - \nabla H_\mu(\vec{y})\| = \left\|\frac{\vec{x}}{\|\vec{x}\|} - \frac{\vec{y}}{\mu}\right\|$$

Similar as Case II, we can show $\|\nabla H_\mu(\vec{x}) - \nabla H_\mu(\vec{y})\| < \frac{1}{\mu}\|\vec{x} - \vec{y}\|$.

Case IV $\vec{x} \in \{ \vec{z} : \|\vec{z}\| > u \}$, $\vec{y} \in \{ \vec{z} : \|\vec{z}\| > u \}$, then.

$$\| \nabla H_u (\vec{x}) - \nabla H_u (\vec{y}) \| = \| \frac{\vec{x}}{\|\vec{x}\|} - \frac{\vec{y}}{\|\vec{y}\|} \|$$

Since $\| \frac{\vec{x}}{\|\vec{x}\|} - \frac{\vec{y}}{\|\vec{y}\|} \|^2 - \frac{1}{u^2} \| \vec{x} - \vec{y} \|^2$

$$= 1 + 1 - 2 \frac{\vec{x}^T \vec{y}}{\|\vec{x}\| \|\vec{y}\|} - \frac{1}{u^2} ( \|\vec{x}\|^2 + \|\vec{y}\|^2 - 2 \vec{x}^T \vec{y} )$$

$$= \frac{1}{u^2} [ 2u^2 - \|\vec{x}\|^2 - \|\vec{y}\|^2 - 2 u^2 \cos\theta + 2 \|\vec{x}\| \|\vec{y}\| \cos\theta ]$$

$$\leq \frac{1}{u^2} [ 2u^2 (1 - \cos\theta) - (\|\vec{x}\|^2 + \|\vec{y}\|^2) + \cos\theta ( \|\vec{x}\|^2 + \|\vec{y}\|^2 ) ]$$

$$= \frac{1}{u^2} (1 - \cos\theta) [ 2u^2 - \|\vec{x}\|^2 - \|\vec{y}\|^2 ] < 0.$$

Thus. $\| \nabla H_u (\vec{x}) - \nabla H_u (\vec{y}) \| \leq \frac{1}{u} \| \vec{x} - \vec{y} \|$.

In conclusion, we have

$$\| \nabla H_u (\vec{x}) - \nabla H_u (\vec{y}) \| \leq \frac{1}{u} \| \vec{x}, \vec{y} \|. \quad \forall \vec{x}, \vec{y} \in \mathbb{R}^n.$$

Thus.

$$H_u \in C^{1,1}_{\frac{1}{u}} ( \mathbb{R}^n )$$

4 For the proof of convexity of functions in this problem, we will use the following properties:

① addition preserves convexity

② multiplication by a nonnegative scalar preserves convexity.

③ composition of a non-descreasing convex function with a convex function. preserves convexity

④ composition of a convex function with an affine transformation. preserves convexity

Thus. if we can rewrite. the original problem into a new form. by using the above. operations, then we can conclude the problem is convex.

(i) The original problem can be rewritten as

$$\min \quad \vec{x}^T A \vec{x} + \vec{b}^T \vec{x}$$

$$\text{s.t.} \quad g_1 ( A_1 \vec{x} + \vec{b}_1 ) + g_2 ( A_2 \vec{x} + \vec{b}_2 ; A_3 \vec{x} ) - 6 \leq 0$$

$$l(x_1) - 1 \geq 0$$

$$l(x_2) - 1 \geq 0$$

$$l(x_3) - 1 \geq 0$$

where $\vec{x} = [ x_1 \ x_2 \ x_3 ]^T$ and

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$ whose eigenvalues are $\lambda_1(A) = \frac{987}{2584}$ $\lambda_2(A) = 1$. $\lambda_3(A) = \frac{2584}{987}$

$\vec{b} = [3, -4, 0]^T$.

$g_1(\vec{z}) = \|\vec{z}\|^2$ is a convex function.

$g_2(\vec{z}; y) = \frac{\|\vec{z}\|^2}{y}$, $y > 0$ is a convex function.

$l(y) = y$ is an affine function (convex and concave)

$$A_1 = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ \sqrt{\frac{3}{2}} & 0 & 0 \\ 0 & \sqrt{\frac{7}{2}} & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \vec{b}_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 2 \end{bmatrix}$$

$A_2 = [1 \ \ -1 \ \ 1]$, $\vec{b}_2 = 1$, $A_3 = [1 \ \ 1 \ \ 0]$

Thus (i) is convex based on properties ① ④.

(ii). we first show that $2x_1^2 + 4x_1x_2 + 7x_2^2 + 10x_2 + 6 \geq 0$ $\forall$. $x_1, x_2$.

since $2x_1^2 + 4x_1x_2 + 7x_2^2 + 10x_2 + 6$

$$= 2(x_1^2 + 2x_1x_2 + x_2^2 + \frac{5}{2}x_2^2 + 5x_2 + 3)$$

$$= 2\left[(x_1+x_2)^2 + \left(\sqrt{\frac{5}{2}}x_2 + \sqrt{\frac{5}{2}}\right)^2 + \frac{1}{2}\right] > 0$$

Further the original problem can be rewritten as

$$\min \ f_1(A_1\vec{x}) + f_2(\vec{x}) + \sqrt{2} f_3(A_2\vec{x} + \vec{b}_2)$$

$$\text{s.t.} \ g_1(A_3\vec{x} + \vec{b}_3; A_4\vec{x}) + \vec{x}^T A_5 \vec{x} - 7 \leq 0$$

$$l(x_1) \geq 0$$

$$l(x_2) - 1 \geq 0$$

where $\vec{x} = [x_1, x_2, x_3]^T$ and.

$f_1(y) = |y|$ is a convex function. $A_1 = [2 \ \ 3 \ \ 1]$.

$f_2(\vec{x}) = \|\vec{x}\|^2$ is a convex function.

$f_3(\vec{z}) = \|\vec{z}\|$ is a convex function $A_2 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & \sqrt{\frac{5}{2}} & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $\vec{b}_2 = \begin{bmatrix} 0 \\ \sqrt{\frac{5}{2}} \\ \sqrt{\frac{1}{2}} \end{bmatrix}$

$g_1(\vec{z}; y) = \frac{\|\vec{z}\|^2}{y}$, $y>0$ is a convex function. $A_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. $A_4 = [1\ 0\ 0]$

$A_5 = \begin{bmatrix} 2 & 2 & 1 \\ 2 & 5 & 1 \\ 1 & 1 & 10 \end{bmatrix}$ whose eigenvalues are $\lambda_1(A_5) = \frac{1439}{1473}$, $\lambda_2(A_5) = \frac{2641}{472}$, $\lambda_3(A_5) = \frac{1804}{173}$

Thus the original problem is convex. based on properties ① ② ④.

(iii). The problem can be rewritten as

$$\min \ h_1(f_1(A_1\vec{x}; A_2\vec{x})) + f_2(A_3\vec{x} + \vec{b}_3)$$

$$\text{s.t.} \ \vec{x}^T A_4 \vec{x} - 100 \le 0$$

$$l(A_5\vec{x}) - 2 = 0$$

$$l(A_6\vec{x}) - 1 \ge 0$$

where $\vec{x} = [x_1\ x_2\ x_3]^T$ and

$h_1(y) = y^2$, $y>0$ is a non-descreasing convex function

$f_1(\vec{z}; y) = \frac{\|\vec{z}\|^2}{y}$, $y>0$ is a convex function, $A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, $A_2 = [1\ 1\ 0]$,

$f_2(\vec{z}) = \|\vec{z}\|$ is a convex function, $A_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$. $\vec{b}_3 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

$A_4 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}$ whose eigenvalues are $\lambda_1(A_4) = 0$, $\lambda_2(A_4) = \frac{577}{985}$, $\lambda_3(A_4) = \frac{1393}{408}$

$l(y) = y$ is an affine function. $A_5 = [1\ 1\ 1]$, $A_6 = [1\ 1\ 0]$.

Thus the original problem is a convex problem. based on properties ① ③ ④.

(iv) The problem can be rewritten as

$$\min \ h_1(f_1(A_1\vec{x}; A_2\vec{x}) + f_1(A_2\vec{x}; A_1\vec{x})) + f_2(A_5\vec{x} + \vec{b}_1)$$

$$\text{s.t.} \ h_2(h_2(g_1(A_6\vec{x} + \vec{b}_2)) + 1) + g_2(A_7\vec{x}) + g_2(A_8\vec{x}) + g_2(A_9\vec{x}) - 200 \le 0$$

$$h_3(g(\vec{x})) - 40 \le 0$$

$$l(A_{11}\vec{x}) - 1 \ge 0$$

$$l(A_{12}\vec{x}) - 1 \ge 0.$$

where $\vec{x} = [\,x_1 \quad x_2 \quad x_3\,]^T$ and

$h_1(y) = y^2$, $y>0$ is a non-descreasing convex function.

$f_1(\vec{z}; y) = \frac{\|\vec{z}\|^2}{y}$, $y>0$ is a convex function, $A_1 = [1, 0, 0]$, $A_2 = [0 \; 1 \; 0]$

$f_2(\vec{z}) = \|\vec{z}\|_1$ is a convex function, $A_5 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $\vec{b}_1 = \begin{bmatrix} 5 \\ 5 \\ 5 \end{bmatrix}$

$h_2(y) = y^2$, $y>0$ is a nondescreasing convex function.

$g_1(\vec{z}) = \|\vec{z}\|^2$, is a convex function. $A_6 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, $\vec{b}_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$

$g_2(y) = y^4$ is a convex function. $A_7 = [1 \; 0 \; 0]$, $A_8 = [0 \; 1 \; 0]$, $A_9 = [0 \; 0 \; 1]$.

$h_3(\vec{z}) = \max\{\vec{z}\}$ is a non-descreasing convex function.

$g(\vec{x}) = [\; \vec{x}^T A_{10} \vec{x}; \;\; A_{11}\vec{x}, \;\; A_{12}\vec{x} \;]$

$A_{10} = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 9 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ whose eigenvalues are $\lambda_1(A_{10}) =$

$A_{11} = [1, 0, 0]$, $A_{12} = [0, 1, 0]$.

$l(y) = y$ is an affine function.

Thus, the original problem is convex based on properties ① ③ ④.

solutions (see code)

> Note: for simplicity of clarification, we always use $\vec{x} = [x_1, x_2, x_3]^T$ in the above analysis, but we can dynamicly change $\vec{x}$ based on the form of the formula, see code implementation

5. proof: since $\vec{x}^*$ is a stationary point iff $\nabla f(\vec{x}^*)^T(\vec{x} - \vec{x}^*) \geq 0$, $\forall \vec{x} \in C$,

$C = \{\vec{x} : \vec{a}^T\vec{x} = 1, \; \vec{a} \in \mathbb{R}^n_{++}\}$

Thus we only need to show (Given $\vec{a}^T\vec{x}^* = 1$)

$$\nabla f(\vec{x}^*)^T(\vec{x} - \vec{x}^*) \geq 0, \; \forall \vec{x} \in C \iff \frac{\frac{\partial f}{\partial x_1}(\vec{x}^*)}{a_1} = \frac{\frac{\partial f}{\partial x_2}(\vec{x}^*)}{a_2} = \cdots = \frac{\frac{\partial f}{\partial x_n}(\vec{x}^*)}{a_n}$$

"$\Leftarrow$" assume $\vec{a}^T\vec{x}^* = 1$ and $\frac{\frac{\partial f}{\partial x_1}(\vec{x}^*)}{a_1} = \frac{\frac{\partial f}{\partial x_2}(\vec{x}^*)}{a_2} = \cdots = \frac{\frac{\partial f}{\partial x_n}(\vec{x}^*)}{a_n}$

Then $\forall \vec{x} \in C$, we have

$$\nabla f(\vec{x}^*)^T (\vec{x} - \vec{x}^*) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\vec{x}^*)(x_i - x_i^*)$$

$$= \sum_{i=1}^{n} a_i \frac{\frac{\partial f}{\partial x_i}(\vec{x}^*)}{a_i}(x_i - x_i^*)$$

$$= \frac{\frac{\partial f}{\partial x_1}(\vec{x}^*)}{a_1} \sum_{i=1}^{n} a_i (x_i - x_i^*)$$

$$= \frac{\frac{\partial f}{\partial x_1}(\vec{x}^*)}{a_1} (\vec{a}^T\vec{x} - \vec{a}^T\vec{x}^*) = \frac{\frac{\partial f}{\partial x_1}(\vec{x}^*)}{a_1}(1-1) = 0 \geq 0.$$

"$\Rightarrow$" assume $\vec{a}^T\vec{x}^* = 1$ and $\nabla f(\vec{x}^*)^T(\vec{x} - \vec{x}^*) \geq 0$, $\forall \vec{x} \in C$.

we will use arguement by contradiction. suppose there exist two different indices $i \neq j$, s.t. $\frac{\frac{\partial f}{\partial x_i}(\vec{x}^*)}{a_i} > \frac{\frac{\partial f}{\partial x_j}(\vec{x}^*)}{a_j}$.

Denote the vector $\vec{x} \in C$ as.

$$x_k = \begin{cases} x_k^*, & k \neq i, j \\ x_i^* - \frac{1}{a_i}, & k = i. \\ x_j^* + \frac{1}{a_j}, & k = j. \end{cases}$$

Then $\nabla f(\vec{x}^*)^T(\vec{x} - \vec{x}^*)$

$$= a_i \frac{\frac{\partial f}{\partial x_i}(\vec{x}^*)}{a_i}(x_i - x_i^*) + a_j \frac{\frac{\partial f}{\partial x_j}(\vec{x}^*)}{a_j}(x_j - x_j^*)$$

$$= \frac{\frac{\partial f}{\partial x_i}(\vec{x}^*)}{a_i}(a_i x_i - a_i x_i^*) + \frac{\frac{\partial f}{\partial x_j}(\vec{x}^*)}{a_j}(a_j x_j - a_j x_j^*)$$

$$= \frac{\frac{\partial f}{\partial x_i}(\vec{x}^*)}{a_i}(-1) + \frac{\frac{\partial f}{\partial x_j}(\vec{x}^*)}{a_j}(+1) < 0$$

This is a contradition with $\nabla f(\vec{x}^*)^T(\vec{x} - \vec{x}^*) \geq 0$.

Thus we can conclude the statement in the problem. #

6. (i) First we can rewrite the problem as.

$$(Q) \quad \begin{aligned} \min \quad & \vec{x}^T A \vec{x} + \vec{b}^T \vec{x} \\ \text{s.t} \quad & x_1, x_2, x_3 \geq 0. \end{aligned}$$

where. $A = \begin{bmatrix} 2 & 1 & -1 \\ 1 & 3 & 0 \\ -1 & 0 & 4 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} -8 \\ -4 \\ -2 \end{bmatrix}$

Since the eigenvalues of A are. $\lambda_1(A) = \frac{511}{456}$, $\lambda_2(A) = \frac{2477}{740}$, $\lambda_3(A) = \frac{2189}{483}$.

Thus the original problem is convex. Note that. $\forall \vec{x} \in C = \{ \vec{x} : \vec{x} \succeq \vec{0} \}$.

let $\vec{x}^* = (\frac{17}{7}, 0, \frac{6}{7})^T$, then.

$$\begin{aligned} \nabla f(\vec{x}^*)^T (\vec{x} - \vec{x}^*) &= (2A\vec{x}^* + \vec{b})^T (\vec{x} - \vec{x}^*) \\ &= \left( 2 \begin{bmatrix} 2 & 1 & -1 \\ 1 & 3 & 0 \\ -1 & 0 & 4 \end{bmatrix} \begin{bmatrix} \frac{16}{7} \\ 0 \\ \frac{6}{7} \end{bmatrix} + \begin{bmatrix} -8 \\ -4 \\ -2 \end{bmatrix} \right)^T \begin{pmatrix} x_1 - \frac{17}{7} \\ x_2 \\ x_3 - \frac{6}{7} \end{pmatrix} \\ &= \begin{bmatrix} 0 & \frac{6}{7} & 0 \end{bmatrix} \begin{bmatrix} x_1 - \frac{17}{7} \\ x_2 \\ x_3 - \frac{6}{7} \end{bmatrix} = \frac{6}{7} x_2 \geq 0. \end{aligned}$$

Thus $\vec{x}^*$ is a stationary point. Since the optimization is a convex optimization problem, then $\vec{x}^*$ is an optimal solution of $(Q)$.

(ii) Since

$$\begin{aligned} & \| 2A\vec{x} + \vec{b} - 2A\vec{y} - \vec{b} \| \\ &= 2 \| A\vec{x} - A\vec{y} \| \\ &\leq 2 \|A\| \, \| \vec{x} - \vec{y} \| \end{aligned}$$

Then $L = 2\|A\| = \frac{1695}{187}$

on the other hand, from the class we know. that.

$$P_C(\vec{x}) = [\vec{x}]_+ , \quad \text{where} \quad [x_i]_+ = \begin{cases} x_i, & x_i \geq 0 \\ 0. & x_i < 0. \end{cases}$$