Assignment 2

Songtuan Lin u6162630

March 22, 2019

Question 1

According to the definition of norm, the equation $\|\mathbf{x} - \mathbf{x_0}\|_2 \le \|\mathbf{x} - \mathbf{x_1}\|_2$ is the same as:

$$(\mathbf{x} - \mathbf{x_0})^T(\mathbf{x} - \mathbf{x_0}) \le (\mathbf{x} - \mathbf{x_1})^T(\mathbf{x} - \mathbf{x_1})$$

Which can be further expressed as:

$$\mathbf{x}^T(\mathbf{x_1} - \mathbf{x_0}) + (\mathbf{x_1} - \mathbf{x_0})^T\mathbf{x} \le \mathbf{x_1}^T\mathbf{x_1} - \mathbf{x_0}^T\mathbf{x_0} \tag{1}$$

Since $\mathbf{x}^T(\mathbf{x_1} - \mathbf{x_0}) = \left((\mathbf{x_1} - \mathbf{x_0})^T\mathbf{x}\right)^T$, which produce a constant, we have:

$$\mathbf{x}^T(\mathbf{x_1} - \mathbf{x_0}) = (\mathbf{x_1} - \mathbf{x_0})^T\mathbf{x} \tag{2}$$

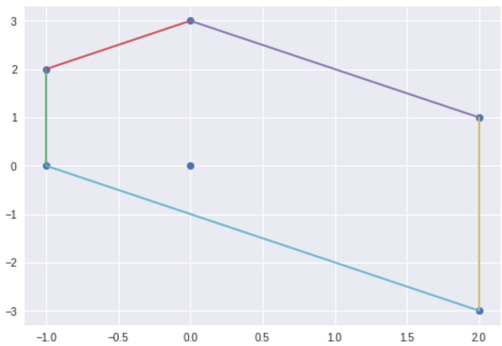
Replace equation 2 to equation 1, there is:

$$2(\mathbf{x_1} - \mathbf{x_0})^T\mathbf{x} \le \mathbf{x_1}^T\mathbf{x_1} - \mathbf{x_0}^T\mathbf{x_0}$$

Which follow the definition of half-space: $\lambda^T\mathbf{x} \le \mathbf{b}$, where $\lambda = 2(\mathbf{x_1} - \mathbf{x_0})$ and $\mathbf{b} = \mathbf{x_1}^T\mathbf{x_1} - \mathbf{x_0}^T\mathbf{x_0}$.

Question 2

The polyhedron constructed by the convex hull is the area with the 5 color line as boundary, shown in following figure:



Denote $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ as the point within $\mathcal{R}^2$, the hyper-plane defined by the 5 color line shown in above figure are:

$$\begin{bmatrix} -1 & 0 \\ 1 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \mathbf{x} = \begin{bmatrix} -1 \\ -1 \\ 2 \\ 3 \end{bmatrix}$$

As a result, the polyhedron constructed by these hyper-plane can be expressed as:

$$\mathcal{A}\mathbf{x} \preceq \mathbf{b}$$

Where $\mathcal{A} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} -1 \\ -1 \\ 2 \\ 3 \end{bmatrix}$

Question 3

(a)

Denote set $\{\mathbf{x}|\, \alpha \le \mathbf{a}^T\mathbf{x} \le \beta\}$ as $\mathcal{C}$, then, $\mathcal{C}$ is a convex set. The prove is as follow:

Assume $\mathbf{x_1}, \mathbf{x_2} \in \mathcal{C}$, their convex combination follow:

$$\mathbf{a}^T(\theta\mathbf{x_1} + (1-\theta)\mathbf{x_2}) = \theta\mathbf{a}^T\mathbf{x} + (1-\theta)\mathbf{a}^T\mathbf{x_2} \tag{3}$$

Since $\mathbf{x_1}, \mathbf{x_2} \in \mathcal{C}$, there is:

$$\begin{cases} \mathbf{a}^T\mathbf{x_1} \ge \alpha \\ \mathbf{a}^T\mathbf{x_2} \ge \alpha \end{cases} \tag{4}$$

Combine equation 3 and 4, we have:

$$\begin{aligned} \theta\mathbf{a}^T\mathbf{x} + (1-\theta)\mathbf{a}^T\mathbf{x_2} &\ge \theta\alpha + (1-\theta)\alpha \\ &= \alpha \end{aligned}$$

The equation $\theta\mathbf{a}^T\mathbf{x} + (1-\theta)\mathbf{a}^T\mathbf{x_2} \le \beta$ can be proved by using the same method. As a result, when $\mathbf{x_1}, \mathbf{x_2} \in \mathcal{C}$, $\theta\mathbf{x_1} + (1-\theta)\mathbf{x_2} \in \mathcal{C}$ hold for any $0 \le \theta \le 1$. Thus, $\mathcal{C}$ is a convex set.

(b)

The *rectangle* set $\mathcal{G}$ can be thought as the intersection of different sets that follow:

$$\mathcal{G} = \bigcap_{i=1}^{n} \mathcal{G}_i$$

Where $\mathcal{G}_i$ is the set: $\{\mathbf{x}|\, \alpha \le x_i \le \beta\}$. It is obvious that $\mathcal{G}_i$ is the *slab* defined in (a) with

$$\mathbf{a} = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

In which, the ith element a_i that corresponding to x_i within $\mathbf{x}$ is set to be 1 and the other elements is 0. Therefore, $\mathcal{G}_i$ is convex, which means, $\mathcal{G}$ is the intersection of convex set, which is also convex.

(c)

The set $\{\mathbf{x}|\, \mathbf{a_1}^T\mathbf{x} \le b_1, \mathbf{a_2}^T\mathbf{x} \le b_2\}$ is the same as:

$$\{\mathbf{x}|\, \mathbf{a_1}^T\mathbf{x} \le b_1\} \cap \{\mathbf{x}|\, \mathbf{a_2}^T\mathbf{x} \le b_2\}$$

For any $i \in \{0, 1\}$, the set $\{\mathbf{x}|\, \mathbf{a_i}^T\mathbf{x} \le b_i\}$ define a half-space, which is a convex set. As a result, the set $\{\mathbf{x}|\, \mathbf{a_1}^T\mathbf{x} \le b_1, \mathbf{a_2}^T\mathbf{x} \le b_2\}$ is the intersection of two convex set, which is also convex.

(d)

Assume $\mathcal{S} = \{\mathbf{s_1}, \mathbf{s_2}, \cdots, \mathbf{s_n}\}$. The original set $\{\mathbf{x}|\, \|\mathbf{x} - \mathbf{x_0}\|_2 \le \|\mathbf{x} - \mathbf{s}\|_2 \text{ for any } \mathbf{s} \in \mathcal{S}\}$ can be expressed as:

$$\bigcap_{i=1}^{n} \{\mathbf{x}|\, \|\mathbf{x} - \mathbf{x_0}\|_2 \le \|\mathbf{x} - \mathbf{s_i}\|_2\}$$

The same concept can be expanded to set $\mathcal{S}$ which has infinite element($n \to \infty$). According to Question 1, the single set $\mathbf{x}|\, \|\mathbf{x} - \mathbf{x_0}\|_2 \le \|\mathbf{x} - \mathbf{s_i}\|_2$ define a half-space and is convex. As a result, the intersection of these convex is also convex, which means, $\{\mathbf{x}|\, \|\mathbf{x} - \mathbf{x_0}\|_2 \le \|\mathbf{x} - \mathbf{s}\|_2 \text{ for any } \mathbf{s} \in \mathcal{S}\}$ is convex.

(e)

We first consider the set:

$$\mathcal{G}_i = \{\mathbf{x} | \, \|\mathbf{x} - \mathbf{s_i}\|_2 \le \|\mathbf{x} - \mathbf{s}\|_2 \; for \; all \; \mathbf{s} \in \mathcal{S}\} \tag{5}$$

Where $\mathbf{s_i} \in \mathcal{S}$ is a fixed point within $\mathcal{S}$. This set is consist of points $\mathbf{x}$ that have the minimum distance to $\mathcal{S}$ through $\mathbf{s_i}$. It is obvious that

$$\{\mathbf{x} | \, \|\mathbf{x} - \mathbf{s_i}\|_2 \le \|\mathbf{x} - \mathbf{s}\|_2\} \cap \{\mathbf{x} | \, \|\mathbf{x} - \mathbf{s_j}\|_2 \le \|\mathbf{x} - \mathbf{s}\|_2\} = \emptyset \tag{6}$$

when $\mathbf{s_i}, \mathbf{s_j} \in \mathcal{S}$ and $\mathbf{s_i} \neq \mathbf{s_j}$ as a point $\mathbf{x}$ can only achieve the minimum distance to $\mathcal{S}$ by hold $\|\mathbf{x} - \mathbf{s_i}\| < \|\mathbf{x} - \mathbf{s_j}\|$ or the inverse. Now assume $\mathcal{S} = \{\mathbf{s_1}, \mathbf{s_2}, \cdots, \mathbf{s_n}\}$ and $\mathcal{T} = \{\mathbf{t_1}, \mathbf{t_2}, \cdots, \mathbf{t_k}\}$. We define the set:

$$\mathcal{Q}_i = \{\mathbf{x} | \, \|\mathbf{x} - \mathbf{s_i}\|_2 \le \|\mathbf{x} - \mathbf{t}\|_2 \; for \; all \; \mathbf{t} \in \mathcal{T}\}$$

Where $\mathbf{s_i} \in \mathcal{S}$. This set describe the point $\mathbf{x}$ that the distance between $\mathbf{x}$ and a point $\mathbf{s_i} \in \mathcal{S}$ is less than the distance between $\mathbf{x}$ and set $\mathcal{T}$. As a result, the set $\{\mathbf{x} | \, \mathbf{dist}(\mathbf{x}, \mathcal{S}) < \mathbf{dist}(\mathbf{x}, \mathcal{T})\}$ can be then expressed as:

$$\bigcup_{i=1}^{n} (\mathcal{G}_i \cap \mathcal{Q}_i)$$

According to equation 6, for any $1 \le i, j \le n$, $\mathcal{G}_i \cap \mathcal{G}_j = \emptyset$. Hence, $(\mathcal{G}_i \cap \mathcal{Q}_i) \cup (\mathcal{G}_j \cap \mathcal{Q}_j) = \emptyset$, which means, the set $\{\mathbf{x} | \, \mathbf{dist}(\mathbf{x}, \mathcal{S}) < \mathbf{dist}(\mathbf{x}, \mathcal{T})\}$ is separated and therefore can not be convex.

(f)

Assume $\mathcal{S}_1 = \{\mathbf{s_1}, \mathbf{s_2}, \cdots, \mathbf{s_n}\}$. Then, the set $\mathcal{C} = \{\mathbf{x} | \, \mathbf{x} + \mathcal{S}_1 \subset \mathcal{S}_2\}$ can be replaced as the intersection of different set as:

$$\mathcal{C} = \bigcap_{i=1}^{n} \mathcal{C}_i$$

Where $\mathcal{C}_i = \{\mathbf{x} | \, \mathbf{x} + \mathbf{s_i} \in \mathcal{S}_2\}$. It can be proved that $\mathcal{C}_i$ is convex: Assume $\mathbf{x_1}, \mathbf{x_2} \in \mathcal{C}_i$. The equation $\theta \mathbf{x_1} + (1 - \theta)\mathbf{x_2} + \mathbf{s_i}$ can be expressed as:

$$\theta \mathbf{x_1} + (1 - \theta)\mathbf{x_2} + \mathbf{s_i} = \theta(\mathbf{x_1} + \mathbf{s_i}) + (1 - \theta)(\mathbf{x_2} + \mathbf{s_i})$$

Since $\mathbf{x_1} + \mathbf{s_i}, \mathbf{x_2} + \mathbf{s_i} \in \mathcal{S}_2$ and $\mathcal{S}_2$ is convex, we have $\theta(\mathbf{x_1} + \mathbf{s_i}) + (1 - \theta)(\mathbf{x_2} + \mathbf{s_i}) \in \mathcal{S}_2$. Hence, $\mathcal{C}_i$ is convex. As a result, $\mathcal{C}$ is convex as it is the intersection of convex set.

(g)

The equation $\|\mathbf{x} - \mathbf{a}\|_2 \le \theta \|\mathbf{x} - \mathbf{b}\|_2$ is the same as $\|\mathbf{x} - \mathbf{a}\|_2^2 \le \theta^2 \|\mathbf{x} - \mathbf{b}\|_2^2$. Now, let $\theta^2 = p$, the equation mentioned before can be expanded as:

$$\mathbf{x}^T \mathbf{x} + \frac{2(p\mathbf{b} - \mathbf{a})^T}{1 - p} \mathbf{x} + \frac{\mathbf{a}^T \mathbf{a} - p \mathbf{b}^T \mathbf{b}}{1 - p} \le 0 \tag{7}$$

Which has the same form as the equation that describe a ball:

$$\|\mathbf{x} - \mathbf{x_0}\|_2^2 = \mathbf{x}^T \mathbf{x} - 2\mathbf{x_0}^T \mathbf{x} + \mathbf{x_0}^T \mathbf{x_0} \le r^2 \tag{8}$$

Compare equation 7 with 8, it is easy to figure out that:

$$\mathbf{x_0} = \frac{\mathbf{a} - p\mathbf{b}}{1 - p} \; and \; r^2 = \mathbf{x_0}^T \mathbf{x_0} - \frac{\mathbf{a}^T \mathbf{a} - p \mathbf{b}^T \mathbf{b}}{1 - p}$$

If $r^2 \ge 0$ can be verified, then, equation 7 always describe a ball. Indeed, r^2 can be expanded as:

$$r^2 = \frac{(\mathbf{a} - p\mathbf{b})^T (\mathbf{a} - p\mathbf{b})}{(1 - p)^2} - \frac{\mathbf{a}^T \mathbf{a} - p \mathbf{b}^T \mathbf{b}}{1 - p}$$

Which is the same as:

$$r^2 = \frac{p(\mathbf{a}^T \mathbf{a} - 2\mathbf{a}^T \mathbf{b} + \mathbf{b}^T \mathbf{b})}{(1 - p)^2}$$

Since $\mathbf{a}^T \mathbf{a} - 2\mathbf{a}^T \mathbf{b} + \mathbf{b}^T \mathbf{b} = \|\mathbf{a} - \mathbf{b}\|_2^2 \ge 0$, r^2 is also greater or equal to zero. Hence, equation 7 describe a ball, which is a convex set.

Question 4

In this question, we firstly define the vector:

$$\mathbf{x} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} \tag{9}$$

Which contain all possible value that random variable x may take. After that, the vector $\mathbf{p}$ which describe the probability distribution of x:

$$\mathbf{p} = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix} \tag{10}$$

Where the element p_i within $\mathbf{p}$ corresponding to the probability that $x = a_i$. It is obvious that the set $\mathcal{P} = \{\mathbf{p} | \, \|\mathbf{p}\| = 1 \textit{ and for each } p_i \textit{ within } \mathbf{p}, p_i \geq 0\}$ is a convex set.

(a)

Follow the definition used in 9 and 10, we can define vector $\mathbf{f}$ as:

$$\mathbf{f} = \begin{bmatrix} f(a_1) \\ f(a_2) \\ \vdots \\ f(a_n) \end{bmatrix}$$

The condition $\alpha \leq \mathcal{E}f(x) \leq \beta$ define a set $\mathcal{Q}$ which is a intersection of two half-space:

$$\mathcal{Q} = \{\mathbf{p} | -\mathbf{f}^T\mathbf{p} \leq -\alpha\} \cap \{\mathbf{p} | \, \mathbf{f}^T\mathbf{p} \leq \beta\}$$

As a result, the set of $\mathbf{p}$ that satisfied the condition $\alpha \leq \mathcal{E}f(x) \leq \beta$ is the intersection of $\mathcal{P}$ and $\mathcal{Q}$, which produce a convex set.

(b)

We firstly define a indicator vector $\mathbf{i}$ as:

$$\mathbf{i} = \begin{bmatrix} \mathbb{I}(a_1 > \alpha) \\ \mathbb{I}(a_2 > \alpha) \\ \vdots \\ \mathbb{I}(a_n > \alpha) \end{bmatrix}$$

Where $\mathbb{I}(s)$ is defined as:

$$\mathbb{I}(s) = \begin{cases} 1 & \textit{if } s \textit{ is true} \\ 0 & \textit{otherwise} \end{cases}$$

Therefore, the constrain can be expressed as:

$$\mathbf{i}^T\mathbf{p} \leq \beta$$

Which obviously, define a half-space. Therefore, the set defined by this half-space: $\mathcal{Q} = \{\mathbf{p} | \, \mathbf{i}^T\mathbf{p} \leq \beta\}$ is a convex set. Consequently, the set of $\mathbf{p}$ that satisfied the constrain can be expressed as $\mathcal{P} \cap \mathcal{Q}$, which is convex.

(c)

Let $f(x)$ in (a) equal to x^2. Hence, there is:

$$\mathbf{f} = \begin{bmatrix} a_1^2 \\ a_2^2 \\ \vdots \\ a_n^2 \end{bmatrix}$$

The constrain $\mathcal{E}x^2 \geq \alpha$ can then be expressed as:

$$-\mathbf{f}^T\mathbf{p} \leq -\alpha$$

Which define a half-space. Hence, the set of $\mathbf{p}$ under this constrain is convex.

(d)

Assume $\mathbf{x} = \begin{bmatrix}1\\0\end{bmatrix}$, $\mathbf{p_1} = \begin{bmatrix}1\\0\end{bmatrix}$ and $\mathbf{p_2} = \begin{bmatrix}0\\1\end{bmatrix}$. Set $\alpha = 0.2$. Therefore, for $\mathbf{p_1}$, we have:

$$Var(\mathbf{x}) = 0 < 0.2$$

For $\mathbf{p_2}$, we have:

$$Var(\mathbf{x}) = 0 < 0.2$$

However, for $\mathbf{p_3} = 0.5\mathbf{p_1} + 0.5\mathbf{p_2}$, there is:

$$Var(\mathbf{x}) = 0.25 > 0.2$$

Hence, the set of $\mathbf{p}$ that satisfied this constrain is not convex.

Question 5

(a)

As the function $f(\mathbf{x}, \mathbf{z})$ is convex over $\mathbf{x}$, it's Hessian Matrix over $\mathbf{x}$ is:

$$\mathcal{H}_{\mathbf{x}} = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 x_n} \\ \frac{\partial^2 f}{\partial x_2 x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n x_1} & \frac{\partial^2 f}{\partial x_n x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

Which satisfied $\mathcal{H}_{\mathbf{x}} \succeq 0$. The Hessian Matrix of f over $\mathbf{z}$: $\mathcal{H}_{\mathbf{z}}$ has the similar form as $\mathcal{H}_{\mathbf{x}}$ and satisfied: $\mathcal{H}_{\mathbf{z}} \preceq 0$. As a result, the Hessian Matrix over $\mathbf{x}$ and $\mathbf{z}$ should be:

$$\nabla^2 f(\mathbf{x}, \mathbf{z}) = \begin{bmatrix} \mathcal{H}_{\mathbf{x}} & \mathcal{H}_{\mathbf{z}} \\ \mathcal{H}_{\mathbf{z}} & \mathcal{H}_{\mathbf{x}} \end{bmatrix}$$

(b)

Firstly, we fix $\mathbf{x} = \tilde{\mathbf{x}}$. Since f is concave over $\mathbf{z}$ and $\nabla f(\tilde{\mathbf{x}}, \tilde{\mathbf{z}}) = 0$, according to the first-order condition of concave function, we have, for any $\mathbf{z}$:

$$f(\tilde{\mathbf{x}}, \mathbf{z}) < f(\tilde{\mathbf{x}}, \tilde{\mathbf{z}}) \tag{11}$$

In order to prove $f(\tilde{\mathbf{x}}, \tilde{\mathbf{z}}) < f(\mathbf{x}, \tilde{\mathbf{z}})$. We fix $\mathbf{z} = \tilde{\mathbf{z}}$. Since f is convex over $\mathbf{x}$, according to the first-order condition of convex function, there is:

$$f(\tilde{\mathbf{x}}, \tilde{\mathbf{z}}) + \nabla f(\tilde{\mathbf{x}}, \tilde{\mathbf{z}})(\mathbf{x} - \tilde{\mathbf{x}}) < f(\mathbf{x}, \tilde{\mathbf{z}})$$

Since $\nabla f(\tilde{\mathbf{x}}, \tilde{\mathbf{z}}) = 0$, we can prove that:

$$f(\tilde{\mathbf{x}}, \tilde{\mathbf{z}}) < f(\mathbf{x}, \tilde{\mathbf{z}})$$

In order to prove that $\sup_{\mathbf{z}} \inf_{\mathbf{x}} f(\mathbf{x}, \mathbf{z}) = f(\tilde{\mathbf{x}}, \tilde{\mathbf{z}})$. We first prove $\inf_{\mathbf{x}} f(\mathbf{x}, \mathbf{z}) = f(\tilde{\mathbf{x}}, \mathbf{z})$: if this equation do not hold, which means, there exist a $\mathbf{x}^* \neq \tilde{\mathbf{x}}$ that minimize $f(\mathbf{x}, \mathbf{z})$ over $\mathbf{z}$, which means, $f(\mathbf{x}^*, \tilde{\mathbf{z}}) \leq f(\tilde{\mathbf{x}}, \tilde{\mathbf{z}})$. However, this conflict with $f(\tilde{\mathbf{x}}, \tilde{\mathbf{z}}) < f(\mathbf{x}, \tilde{\mathbf{z}})$. As a result, $\inf_{\mathbf{x}} f(\mathbf{x}, \mathbf{z}) = f(\tilde{\mathbf{x}}, \mathbf{z})$. Furthermore, according to equation 11, there is:

$$\sup_{\mathbf{z}} f(\tilde{\mathbf{x}}, \mathbf{z}) = f(\tilde{\mathbf{x}}, \tilde{\mathbf{z}})$$

As a result, $\sup_{\mathbf{z}} \inf_{\mathbf{x}} f(\mathbf{x}, \mathbf{z}) = f(\tilde{\mathbf{x}}, \tilde{\mathbf{z}})$. The similar method can be used to prove $\inf_{\mathbf{x}} \sup_{\mathbf{z}} f(\mathbf{x}, \mathbf{z}) = f(\tilde{\mathbf{x}}, \tilde{\mathbf{z}})$

(c)