Math 430 Assignment 3

Arnold Jiadong Yu

September 17, 2018

Exercise 2.1

Solution:

(a) It is not a polyhedron. Let us call this set S, then

$$S = \{\mathbf{x} \in \mathbb{R}^2 | \mathbf{A}\mathbf{x} \le \mathbf{b} \text{ for } \forall\theta \in [0, \frac{\pi}{2}] \wedge \mathbf{x} \ge 0\}$$

where $\mathbf{A} = [\cos\theta \sin\theta]$ and $\mathbf{b} = 1$. Since $\theta \in [0, \frac{\pi}{2}]$, then $\mathbf{A} \ge 0$. We can write

$$[\cos\theta \sin\theta] \begin{bmatrix} x \\ y \end{bmatrix} \le 1$$

It could be intersection of infinite many halfspaces. By definition of polyhedron, it is intersection of finite many halfspaces. Hence this is not a polyhedron.

(b) It is a polyhedron. Let us call this set S, then

$$S = \{x \in \mathbb{R} | x^2 - 8x + 15 \le 0\}$$

$$\Rightarrow S = \{x \in \mathbb{R} | (x-3)(x-5) \le 0$$

$$\Rightarrow S = \{x \in [3,5]\} \Rightarrow S = [3,5]$$

In $\mathbb{R}$, S is an intersection of two halfspaces. S is a polyhedron.

(c) It is a trivial polyhedron. Since it is empty set, we can choose any $\mathbf{A}, \mathbf{b}$ to satisfy the polyhedron definition. For example,

$$P = \{\mathbf{x} \in \mathbb{R} | \mathbf{A}\mathbf{x} \ge \mathbf{b}\}$$

Let $\mathbf{A} = (-2, 2)^T \in \mathbb{R}^2$ and $\mathbf{b} = (3, -3)^T \in \mathbb{R}^2$, then $P = \emptyset$. P follows the definition of polyhedron, therefore P is polyhedron.

Exercise 2.2

Solution: Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function, then for every $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and every $\lambda \in [0, 1]$,

$$f(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) \leq \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})$$

Let $\mathbf{z}_1, \mathbf{z}_2$ be any two points in S, want to show $\lambda\mathbf{z}_1 + (1-\lambda)\mathbf{z}_2 \in S$.

$$f(\lambda\mathbf{z}_1 + (1-\lambda)\mathbf{z}_2) \leq \lambda f(\mathbf{z}_1) + (1-\lambda) f(\mathbf{z}_2)$$

$$\leq \lambda c + (1-\lambda)c = c$$

That is

$$f(\lambda\mathbf{z}_1 + (1-\lambda)\mathbf{z}_2) \leq c \Rightarrow \lambda\mathbf{z}_1 + (1-\lambda)\mathbf{z}_2 \in S$$

Therefore, S is convex by definition.

Exercise 2.5

Solution:

(a)Proof by contrapositive. Let $\mathbf{x}$ be a non-extreme point of P, then we can find two vector $\mathbf{s}, \mathbf{t} \in P$ and a scalar $\lambda \in [0, 1]$, s.t.

$$\mathbf{x} = \lambda\mathbf{s} + (1-\lambda)\mathbf{t}$$

Want to show $f(\mathbf{x})$ is not an extreme point of Q. Since both function f, g is affline, then $f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$, $g(\mathbf{y}) = \mathbf{B}\mathbf{y} + \mathbf{c}$ where $\mathbf{x} \in P, \mathbf{y} \in Q$, $\mathbf{A}, \mathbf{B}$ are matrix, $\mathbf{b}, \mathbf{c}$ are vector. Therefore,

$$f(\mathbf{x}) = f(\lambda\mathbf{s} + (1-\lambda)\mathbf{t}) = \mathbf{A}(\lambda\mathbf{s} + (1-\lambda)\mathbf{t}) + \mathbf{b}$$

$$= \lambda\mathbf{A}\mathbf{s} + (1-\lambda)\mathbf{A}\mathbf{t} + \mathbf{b} = \lambda\mathbf{A}\mathbf{s} + (1-\lambda)\mathbf{A}\mathbf{t} + \lambda\mathbf{b} + (1-\lambda)\mathbf{b}$$

$$= \lambda(\mathbf{A}\mathbf{s} + \mathbf{b}) + (1-\lambda)(\mathbf{A}\mathbf{t} + \mathbf{b}) = \lambda f(\mathbf{s}) + (1-\lambda) f(\mathbf{t})$$

That is $f(\mathbf{x})$ is not an extreme point of Q. Therefore, if $\mathbf{x}$ is not an extreme point of P, then $f(\mathbf{x})$ is not an extreme point of Q. That is if $f(\mathbf{x})$ is an extreme point of Q, then $\mathbf{x}$ is an extreme point of P.

The converse is trivial. That is $\mathbf{x}$ is an extreme point of P if and only if $f(\mathbf{x})$

is an extreme point of Q.
(b) Let f be a map, and g be a map such that

$$f : \mathbf{x} \mapsto (\mathbf{x}, \mathbf{z}) \text{ for } \forall \mathbf{x} \in P$$

Let $f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{c}$ where $\mathbf{A}$ is $(n+k) \times n$ matrix and $\mathbf{c}$ is a vector in $\mathbb{R}^{n+k}$

$$f(\mathbf{x}) = \begin{bmatrix} \mathbf{I}_{n\times n} \\ --- \\ \mathbf{B}_{k\times n} \end{bmatrix} \mathbf{x}_{n\times 1} + \begin{bmatrix} \mathbf{0}_{n\times 1} \\ --- \\ -\mathbf{b}_{k\times 1} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_{n\times 1} \\ --- \\ \mathbf{z}_{k\times 1} \end{bmatrix}$$

where $\mathbf{I}_{n\times n}$ is identity matrix $n \times n, \mathbf{0}_{n\times 1}$ is zero vector in $\mathbb{R}^n$,

$\mathbf{B}_{k\times n}$ is a matrix and $\mathbf{b}$ is a vector in $\mathbb{R}^k$

then $\mathbf{B}_{k\times n}\mathbf{x} - \mathbf{b} = \mathbf{z} \Rightarrow f(\mathbf{x}) \in Q$ for $\forall \mathbf{x} \in P, f$ is both affine and injective map.

$$g : (\mathbf{x}, \mathbf{z}) \mapsto \mathbf{x} \text{ for } \forall (\mathbf{x}, \mathbf{z}) \in Q$$

By assumption there are some matrix $\mathbf{C}$ and vector $\mathbf{d}$, s.t. $\mathbf{C}\mathbf{x} - \mathbf{z} = \mathbf{d}$

Let $g((\mathbf{x}, \mathbf{z})) = \mathbf{D}(\mathbf{x}, \mathbf{z}) + \mathbf{e}$, where $\mathbf{D}$ is $n \times (n+k)$ matrix and $\mathbf{e}$ is a vector in $\mathbb{R}^n$

$$g((\mathbf{x}, \mathbf{z})) = \begin{bmatrix} \mathbf{I}_{n\times n} & | & \mathbf{0}_{n\times k} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{n\times 1} \\ --- \\ \mathbf{z}_{k\times 1} \end{bmatrix} + \mathbf{0}_{n\times 1} = \mathbf{x}_{n\times 1}$$

then $g((\mathbf{x}, \mathbf{z})) \in P$ for $\forall (\mathbf{x}, \mathbf{z}) \in Q$. i.e. g is affine and injective map. By construction, both f, g are affine. Moreover, $g(f(\mathbf{x})) = \mathbf{x}$ for $\forall \mathbf{x} \in P$, $f(g(\mathbf{y})) = \mathbf{y}$ for $\forall \mathbf{y} \in Q$. These imply f, g is inverse map and P and Q are isomorphic.

Exercise 2.6

Solution:

(a) Case 1: When $n \le m$, it is trivial. It is at most m of the coefficients λ_i being nonzero.
Case 2: When $n > m$, assume there are k linear independent vectors in $\mathbf{R}^m$. WLOG, denoted them as $\mathbf{A}_1, ..., \mathbf{A}_k$, where $k \le m < n$. Let $\mathbf{A} \in C$, then

$$A = \sum_{i=1}^{n} \lambda_i \mathbf{A}_i, \text{ where } \lambda_i \ge 0 \text{ for } 1 \le i \le n$$

Consider the polyhedron,

$$\Lambda = \{(\lambda_1, ..., \lambda_n) \in \mathbb{R}^n | \sum_{i=1}^{n} \lambda_i \mathbf{A}_i = \mathbf{y}.\lambda_1, ..., \lambda_n \geq 0\}$$

By definition of basic solution and basic feasible solution, a basic solution $x \in \Lambda$ must satisfy all equality constraints and active for at least n linear independent constraints of Λ, but $\sum_{i=1}^{n} \lambda_i \mathbf{A}_i = \mathbf{y}$ are only k linear independent constraints since $\mathbf{A}_i$ is in $\mathbb{R}^m$. Therefore, we need $(n-k)$ more linear independent constraints to consider a basic solution is active. By linear algebra, in order to find a basic solution, we need to reduce the matrix to $k \times k$ square matrix. That is the other $(n-k)$ linear independent constraints must be zeros. Hence there are only k linear independent constraints are nonzero. Since $k \leq m$, then at most m of λ_i being nonzero.
(b) Consider a polyhedron,

$$\Lambda = \{(\lambda_1, ..., \lambda_n) \in \mathbb{R}^n | \sum_{i=1}^{n} \lambda_i \mathbf{A}_i = \mathbf{y}, \sum_{i=1}^{n} \lambda_i = 1, \lambda_1, ..., \lambda_n \geq 0\}$$

same argument as part a, both $\sum_{i=1}^{n} \lambda_i \mathbf{A}_i = \mathbf{y}, \sum_{i=1}^{n} \lambda_i = 1$ will provide only $k+1$ equality. In order to get a basic solution, there must be $n-k-1$ linear independent constraints are zeros. Therefore, matrix can be reduced to $(k+1) \times (k+1)$ square matrix. Since $k \leq m$, when $k = m$ there are $m+1$ linear independent constraints are nonzero. That is at most $m+1$ of the coefficients λ_i being nonzero.