Convex Optimization

Lecture 5: Convex Programming and Convex Set

Lecturer: *Dr*. Wan-Lei Zhao

Autumn Semester 2025

Contact: wlzhao@xmu.edu.cn

Outline

QP with Equation Constraints Only (1)

- There is an 8 meters long string, we want to cut it into three segments. Each segment will be used to enclose a circle. So how to cut them that minimizes the sum of three circle areas.

1. Given the lengths of three segments are x_1, x_2, and x_3
2. $x_1 + x_2 + x_3 = 8$
3. $r_1 = \frac{x_1}{2\pi}$
4. We want to minimize $\frac{x_1^2}{4\pi} + \frac{x_2^2}{4\pi} + \frac{x_3^2}{4\pi}$
5. Namely, we minimize $x_1^2 + x_2^2 + x_3^2$ under the constraint $x_1 + x_2 + x_3 = 8$

QP with Equation Constraints Only (2)

1. Minimize $x_1^2 + x_2^2 + x_3^2$ under the constraint $x_1 + x_2 + x_3 = 8$

$$\begin{aligned} \text{Max.} \quad & x_1^2 + x_2^2 + x_3^2 \\ \text{s.t.} \quad & x_1 + x_2 + x_3 = 8 \end{aligned} \tag{1}$$

How to solve

- Given $b = 8, x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, H = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$, c= $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, A=[1 1 1]
- We have

$$\begin{aligned} \text{Min.} \quad & \frac{1}{2}x^T Hx + c^T x \\ \text{sub. to} \quad & Ax = b \end{aligned} \tag{2}$$

- How to solve this problem??

Solve QP by Lagrangian Multiplier (1)

$$\begin{aligned} \text{Min. } & \frac{1}{2}x^T Hx + c^T x \\ \text{sub. to } & Ax = b \end{aligned} \tag{4}$$

- Define Lagrangian function

$$L(x, \lambda) = x^T Hx + c^T x + \lambda(Ax - b) \tag{5}$$

Solve QP by Lagrangian Multiplier (2)

- Given Lagrangian function

$$L(x, \lambda) = x^T Hx + c^T x + \lambda(Ax - b) \tag{5}$$

- Take partial derivative on x and λ

$$\nabla L_x = 0, \nabla L_\lambda = 0 \tag{6}$$

$$\Downarrow$$

$$\begin{cases} Hx + c + A\lambda = 0 \\ Ax - b = 0 \end{cases} \tag{7}$$

Solve QP by Lagrangian Multiplier (3)

$$\begin{cases} Hx + c + A\lambda = 0 \\ Ax - b = 0 \end{cases} \tag{7}$$

$$\Downarrow$$

$$\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} -c \\ b \end{bmatrix} \tag{8}$$

$$\Downarrow$$

$$\begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} H_{3\times3} & A^T_{3\times1} \\ A_{1\times3} & 0 \end{bmatrix}^{-1} \begin{bmatrix} -c_{3\times1} \\ b_{1\times1} \end{bmatrix} \tag{9}$$

Solve QP by Lagrangian Multiplier (4)

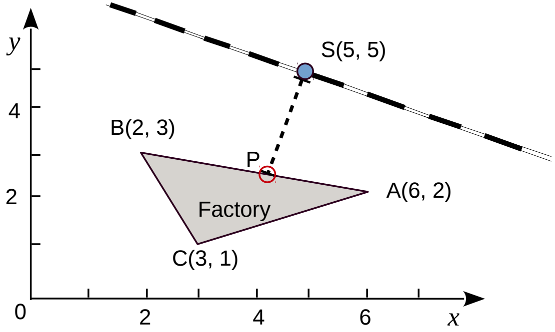
$$\begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} H_{3\times3} & A^T_{3\times1} \\ A_{1\times3} & 0 \end{bmatrix}^{-1} \begin{bmatrix} -c_{3\times1} \\ b_{1\times1} \end{bmatrix} \tag{10}$$

$$\Downarrow$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \lambda \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 & 1 \\ 0 & 2 & 0 & 1 \\ 0 & 0 & 2 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2.667 \\ 2.667 \\ 2.667 \\ -5.333 \end{bmatrix}$$

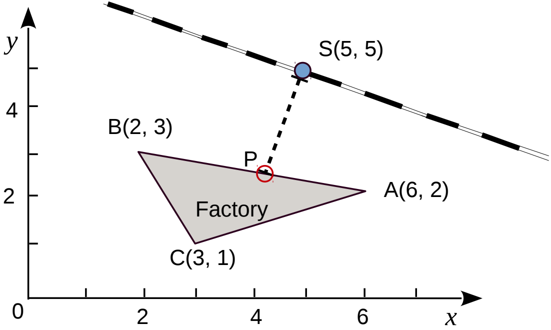
QP with Equation Constraints and Inequation Constraints

- A factory is going to build a road to train station, the locations of the station and the factory are given in Fig. 1. Please help to find the position **P** within the factory area that connects to the station. Such that the distance between **P** and the train station is the shortest.



Modeling (1)

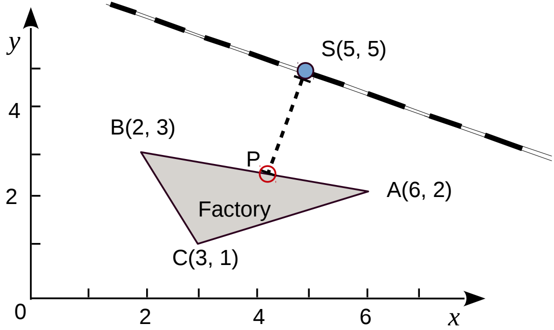
- Let's first look at the target



- Target: Min. $(x_p - 5)^2 + (y_p - 5)^2$

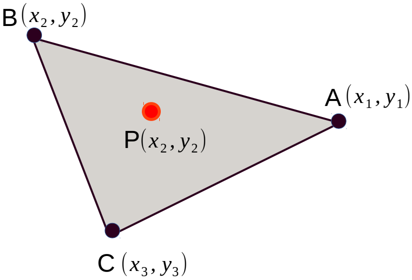
Modeling (2)

- So what is **P**



- **P** should be within the factory area

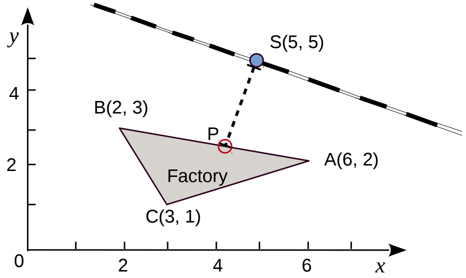
Convex Set (1)



- Area formed by A, B and C is convex
- We therefore have followig property, any point $P(x_p, y_p)$
  1. $P = \alpha_1 \cdot A + \alpha_2 \cdot B + \alpha_3 \cdot C$,
     where $\alpha_1 + \alpha_2 + \alpha_3 = 1$

Modeling (3)

- **P** could be expressed as a linear combination of A, B and C



- Namely, $P = \alpha_1{\cdot}A + \alpha_2{\cdot}B + \alpha_3{\cdot}C = \alpha_1 \begin{bmatrix} 6 \\ 2 \end{bmatrix} + \alpha_2 \begin{bmatrix} 2 \\ 3 \end{bmatrix} + \alpha_1 \begin{bmatrix} 3 \\ 1 \end{bmatrix}$
- $\alpha_1 + \alpha_2 + \alpha_3 = 1$

Modeling (4)

$$\begin{bmatrix} x_p \\ y_p \end{bmatrix} = \alpha_1 \begin{bmatrix} 6 \\ 2 \end{bmatrix} + \alpha_2 \begin{bmatrix} 2 \\ 3 \end{bmatrix} + \alpha_1 \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

- Plug the expressions for x_p and y_p into the target

$$\begin{aligned} \text{min. } & (6\alpha_1 + 2\alpha_2 + 3\alpha_3 - 5)^2 + (2\alpha_1 + 3\alpha_2 + \alpha_3 - 5)^2 \\ & \text{sub. to} \begin{cases} \alpha_1 + \alpha_2 + \alpha_3 = 1 \\ \alpha_1, \alpha_2, \alpha_3 \geq 0 \end{cases} \end{aligned} \tag{2}$$

Modeling (5)

$$\min.\ (6\alpha_1 + 2\alpha_2 + 3\alpha_3 - 5)^2 + (2\alpha_1 + 3\alpha_2 + \alpha_3 - 5)^2 \quad (2)$$
$$\text{sub. to} \begin{cases} \alpha_1 + \alpha_2 + \alpha_3 = 1 \\ \alpha_1, \alpha_2, \alpha_3 \geq 0 \end{cases}$$

- Re-organize above model, we have

$$\min.\ 40\alpha_1^2 + 13\alpha_2^2 + 10\alpha_3^2 + 36\alpha_1\alpha_2$$
$$+40\alpha_1\alpha_3 + 18\alpha_2\alpha_3 - 80\alpha_1 - 50\alpha_2 - 40\alpha_3$$
$$\text{sub. to} \begin{cases} \alpha_1 + \alpha_2 + \alpha_3 = 1 \\ \alpha_1, \alpha_2, \alpha_3 \geq 0 \end{cases}$$

Modeling (6)

- To further simplify the model

$$\begin{aligned} \text{min. } & 40\alpha_1^2 + 13\alpha_2^2 + 10\alpha_3^2 + 36\alpha_1\alpha_2 \\ & +40\alpha_1\alpha_3 + 18\alpha_2\alpha_3 - 80\alpha_1 - 50\alpha_2 - 40\alpha_3 \\ & \text{sub. to} \begin{cases} \alpha_1 + \alpha_2 + \alpha_3 = 1 \\ \alpha_1, \alpha_2, \alpha_3 \geq 0 \end{cases} \end{aligned}$$

$$\Downarrow$$

$$\text{min. } [\alpha_1 \ \alpha_2 \ \alpha_3] \begin{bmatrix} 40 & 18 & 20 \\ 18 & 13 & 9 \\ 20 & 9 & 10 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} + \begin{bmatrix} -80 \\ -50 \\ -40 \end{bmatrix}^T \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} \tag{3}$$

$$\text{sub. to} \begin{cases} \alpha_1 + \alpha_2 + \alpha_3 = 1 \\ I\alpha \succeq 0 \end{cases}$$

The model

$$\min.\ [\alpha_1\ \alpha_2\ \alpha_3]\begin{bmatrix} 40 & 18 & 20 \\ 18 & 13 & 9 \\ 20 & 9 & 10 \end{bmatrix}\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} + \begin{bmatrix} -80 \\ -50 \\ -40 \end{bmatrix}^T \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} \tag{3}$$

$$\text{sub. to} \begin{cases} \alpha_1 + \alpha_2 + \alpha_3 = 1 \\ I\alpha \succeq 0 \end{cases}$$

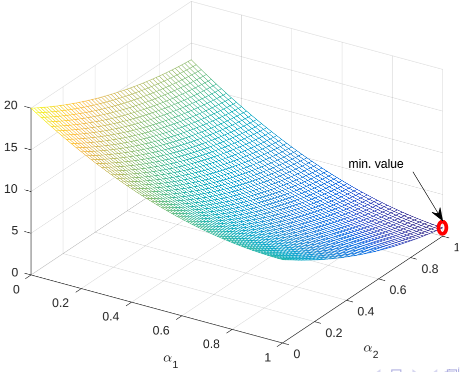
- Two observations
  1. The target function is quadratic
  2. The constraint is a linear equation[1].
- This is a typical **quadratic programming** problem

[1] In practice, there could be multiple equations

Visualize the model

$$\min. \; [\alpha_1 \; \alpha_2 \; \alpha_3] \begin{bmatrix} 40 & 18 & 20 \\ 18 & 13 & 9 \\ 20 & 9 & 10 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} + \begin{bmatrix} -80 \\ -50 \\ -40 \end{bmatrix}^T \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} \quad (3)$$

$$\text{sub. to} \begin{cases} \alpha_1 + \alpha_2 + \alpha_3 = 1 \\ I\alpha \succeq 0 \end{cases}$$



How to solve

- Given $b = 1$,$\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}$, $H = \begin{bmatrix} 40 & 18 & 20 \\ 18 & 13 & 9 \\ 20 & 9 & 10 \end{bmatrix}$, c=$\begin{bmatrix} -80 \\ -50 \\ -40 \end{bmatrix}$,
  A=[1 1 1]
- We have

$$\begin{aligned} \text{min. } & \alpha^T H\alpha + c^T\alpha \\ \text{sub. to } & A\alpha = b \\ & I\alpha \succeq 0 \end{aligned} \tag{4}$$

- How to solve this problem??

Solve QP by Lagrangian Multiplier

$$\begin{aligned} \text{min. } & \alpha^T H \alpha + c^T \alpha \\ \text{sub. to } & A\alpha = b \\ & I\alpha \succeq 0 \end{aligned} \tag{4}$$

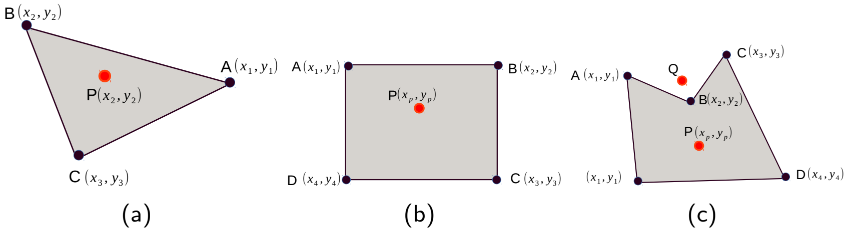
- Define Lagrangian function

$$L(\alpha, \lambda, \beta) = \alpha^T H \alpha + c^T \alpha + \lambda(A\alpha - b) + \beta^T (I\alpha - 0) \tag{5}$$

- This problem cannot be solved in the analytic way
- Before we start to solve the QP with equations and inequations
- We are going to introduce you systematically the **convex set**, **convex function**, and **convex problem**

Outline

Convex Set: the definition (1)



- $\forall x_1, x_2 \in C$, **given** $x_i = \theta x_1 + (1-\theta)x_2$, $\theta \in [0\ 1]$, **a set C is convex** *iff* $x_i \in C$

Convex Set: the definition (2)

- $\forall x_1, x_2 \in C$, **given** $x_i = \theta x_1 + (1-\theta)x_2$, $\theta \in [0\ 1]$**, a set C is convex** *iff* $x_i \in C$
- Now, given $\forall x_1, x_2, \cdots, x_k \in C$ and C is convex, $\theta_1 + \theta_2 + \cdots + \theta_k = 1$, $\theta_k \geq 0, k \geq 2$, $x_j = \theta_1 x_1 + \cdots + \theta_k x_k$ prove $x_j \in C$

1. When $k = 2$, it holds according to the definition of convex set
2. When $k \geq 2$, we prove the case when $k = 3$, and extends it to $k > 3$

Think about it in 5 minutes...

Convex Set: the definition (3)

- Given convex set C, we have $\forall x_1, x_2, x_3 \in C$. We also have $\theta_1 + \theta_2 + \theta_3 = 1, \theta_1, \theta_2, \theta_3 \in [0,\ 1]$, we are going to prove $x_i \in C$ when $x_i = \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3$
- According to the definition of convex set, we have

$$\frac{\theta_1}{\theta_1 + \theta_2} x_1 + \frac{\theta_2}{\theta_1 + \theta_2} x_2 \in C$$

- Therefore, we can build another convex combination

$$x_i = (\theta_1 + \theta_2)(\frac{\theta_1}{\theta_1 + \theta_2} x_1 + \frac{\theta_2}{\theta_1 + \theta_2} x_2) + (1 - \theta_1 - \theta_2) x_3$$

- Since $x_3 \in C$, $x_i \in C$ as well, to this end, the case is proved when $k = 3$
- In the same way, we can prove the case when $k > 3$

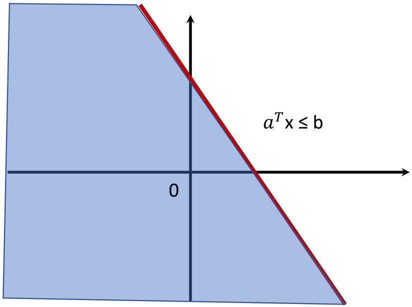
Convex Hull: the definition (1)

- Given a set **S**, the minimal convex set that include set S is called the convex hull of S, it is denoted as **conv(S)**
- Namely, $\forall x_1, x_2 \in S, \theta \in [0, 1]$, we have

$$\theta x_1 + (1 - \theta) x_2 \in Conv(S)$$

Representative Convex Sets – Half-space

- Given plane $\{x|a^Tx \leq b, x, a \in R^n\}$, the half-space is a convex set



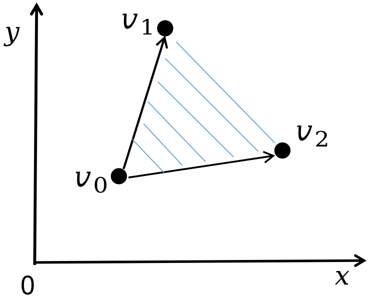
Representative Convex Sets – Polyhedral

- Given $a_i, c_j, x \in R^n$, the polyhedral is defined as

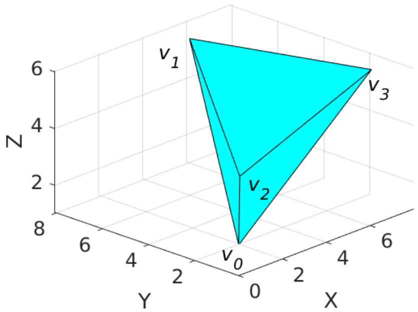
$$P = \left\{ x \middle| \begin{array}{c} a_i^T x \le b_i, i = 1, \cdots, m \\ c_j^T x = d_j, j = 1, \cdots, p \end{array} \right\}. \tag{6}$$

- One $a_i^T x \le b_i$ defines a half-space, which is convex
- One equation $c_j^T x = d_j$ defines a plane, which is convex as well
- The intersection of convex sets are convex as well

Representative Convex Sets – Simplex (1)



(d) 2D Simplex



(e) 3D Simplex

- Given $v_0, v_1, \cdots, v_k \in R^n$ and $k \leq n$, $v_1 - v_0, \cdots, v_k - v_0$ are linearly independent
- The simplex is the convex hull of $v_0, v_1, \cdots, v_k$

Representative Convex Sets – Simplex (2)

- Given $v_0, v_1, \cdots, v_k \in R^n$ and $k \leq n$, $v_1 - v_0, \cdots, v_k - v_0$ are linearly independent
- The simplex is the convex hull ($\mathbf{H}$) of $v_0, v_1, \cdots, v_k$
- $\forall x = \theta_0 v_0 + \cdots + \theta_i v_i + \cdots + \theta_k v_k$, $\sum_{i=0}^{k} \theta_i = 1$, $\theta_i \geq 0$
- $\Rightarrow x \in \mathbf{H}$

In reverse,

- $\forall x \in H$, it can be expressed as

$$x = \theta_0 v_0 + \cdots + \theta_i v_i + \cdots + \theta_k v_k, \tag{7}$$
$$\text{where} \quad \sum_{i=0}^{k} \theta_i = 1, \theta_i \geq 0$$

Representative Convex Sets – Convex Cone (1)

- Given the set $x_1, x_2, \cdots, x_k \in C$ and $\theta_1, \theta_2, \cdots, \theta_k \geq 0$
- $\theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_k x_k \in C$
- Please show set **C** is convex

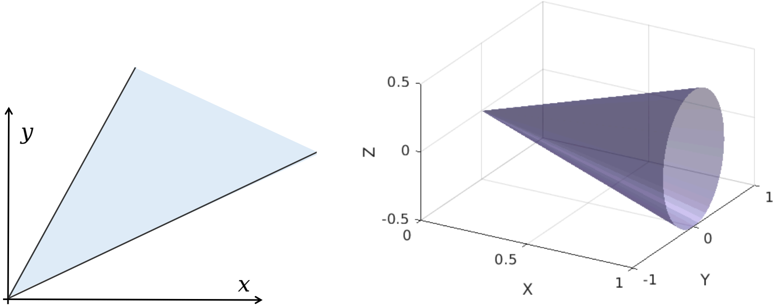


Figure: Convex Cone in 2D and 3D.

Think about it in 5 minutes...

Representative Convex Sets – Convex Cone (2)

- Given the set $x_1, x_2, \cdots, x_k \in C$ and $\theta_1, \theta_2, \cdots, \theta_k \geq 0$
- $\theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_k x_k \in C$
- Please show set **C** is convex
- **Proof:**
- Given $x, y \in C$, $\lambda \in [0, 1]$, $z = \lambda x + (1 - \lambda) y$
- According to the definition of convex cone, $z \in C$ as well
- So **C** is convex

— proved —

Representative Convex Sets – Symmetric Matrix (1)

- Given the set $S^n = \{X \in R^{n \times n} | X = X^T\}$, S is convex
- Prove it based on the definition of convex set

Think about it in 5 minutes...

Representative Convex Sets – Symmetric Matrix (2)

- Given the set $S^n = \{X \in R^{n \times n} | X = X^T\}$, S^n is convex
- **Proof:**
- Given $A, B \in S^n$,$\theta \in [0,\ 1]$
$$X = \theta A + (1-\theta)B$$
- Obviously, X is symmetric and $X \in R^{n \times n}$

Representative Convex Sets – Semi-def. Sym. Matrix

- Given the set $S_+^n = \{X \in R^{n\times n} | X = X^T, X \succeq 0\}$, S_+^n is convex
- **Proof:**
- Given $A, B \in S_+^n$,$\theta \in [0,\ 1]$
- $\forall x \in R^n$, We have $x^TAx \geq 0$, $x^TBx \geq 0$

$$\begin{aligned} x^TAx \geq 0 \\ x^TBx \geq 0 \end{aligned} \tag{8}$$

- So we have

$$\begin{aligned} \theta x^TAx + (1-\theta)x^TBx \geq 0 \\ x^T(\theta A + (1-\theta)B)x \geq 0 \end{aligned} \tag{9}$$

— **proved** —

Representative Convex Sets – Definite Sym. Matrix

- Given the set $S^n_{++} = \{X \in R^{n\times n} | X = X^T, X \succ 0\}$, S^n_{++} is convex
- The proof is quite similar
- We leave it as a homework

Summary over Convex Sets

Name	Expression
Plane	$a^T x + b = 0$
Half-space	$a^T x \leq b$
Polyhedral	$P = \left\{ x \middle\vert \begin{array}{c} a_i^T x \leq b_i, i = 1, \cdots, m \\ c_j^T x = d_j, j = 1, \cdots, p \end{array} \right\}.$
Simplex	$conv\{v_0, \cdots, v_k\}$, $rank([v_1 - v0; \cdots ; v_k - v_0]) = k$
Symmetric Mat.	$S^n = \{X \in R^{n\times n} \vert X = X^T\}$
S.-D. Symm. Matrix	$S_+^n = \{X \in R^{n\times n} \vert X = X^T, X \succeq 0\}$
Def. Symm. Matrix	$S_{++}^n = \{X \in R^{n\times n} \vert X = X^T, X \succ 0\}$

Outline

Intersection of Two Convex Sets (1)

- Given $A, B \in R^n$ are convex, then $A \cap B$ is convex
- How to prove it holds?

Think about it in 5 minutes...

Intersection of Two Convex Sets (2)

- Given sets $A, B \in R^n$ are convex, then $A \cap B$ is convex
- **Proof:**
- Given $x_1, x_2 \in A \cap B$, $\theta \in [0,\ 1]$
- $x_i = \theta x_1 + (1 - \theta) x_2$

1. Since $x_1, x_2 \in A$, A is convex, so $x_i \in A$
2. Since $x_1, x_2 \in B$, B is convex, so $x_i \in B$
3. As x_i is shared by A and B, it is true that $x_i \in A \cap B$

— proved —

Affine Transformation (1)

- Given $f : R^n \rightarrow R^m$ is an affine tranform, namely $f(x) = Ax + b$
- $A \in R^{m \times n}, b \in R^m$
- If $S \in R^n$ is convex, then $Q = \{f(x) | x \in S\}$ is convex as well
- How to prove it??

Think about it in 5 minutes...

Affine Transformation (2)

- Given $f : R^n \rightarrow R^m$ is an affine tranform, namely $f(x) = Ax + b$
- $A \in R^{m \times n}, b \in R^m$
- If $S \in R^n$ is convex, then $Q = \{f(x) | x \in S\}$ is convex as well
- **Proof:**
- Given $x, y \in S^n$, after $f(\cdot)$ transformation, we have
- $Ax + b$ and $Ay + b$, they are inside Q, we should check whether $z \in Q$
- $z = \theta(Ax + b) + (1 - \theta)(Ay + b)$, $\theta \in [0,\ 1]$

$$z = \theta(Ax + b) + (1 - \theta)(Ay + b)$$
$$z = A(\theta x + (1 - y)\theta) + b$$

- Since $\theta x + (1 - y)\theta \in S$, $z \in Q$ as well

— proved –

Perspective Transformation (1)

- The perspective transformation is defined as $P: \ R^{n+1} \rightarrow R^n$, $(\widetilde{x}, x_{n+1}), x_{n+1} > 0$,
- Namely, $P(\widetilde{x}, x_{n+1}) = \frac{\widetilde{x}}{x_{n+1}}$, $\widetilde{x} \in R^n$
- Now let's prove this transformation maintains the convexity of a convex set

Think about it in 5 minutes...

Perspective Transformation (2)

- Given $x = (\widetilde{x}, x_{n+1})$, and $y = (\widetilde{y}, y_{n+1})$
- A segment connecting x and y is $\forall \theta \in [0, 1]$
- We should prove that $\theta x + (1 - \theta) y$ is a segment after the projection

Think about it in 5 minutes...

Perspective Transformation (3)

- Given $x = (\widetilde{x}, x_{n+1})$, and $y = (\widetilde{y}, y_{n+1})$
- A segment connecting x and y is $\forall \theta \in [0, 1]$
- We should prove that $\theta x + (1-\theta)y$ is a segment after the projection
- **Proof:**

$$\theta x + (1-\theta)y \rightarrow P(\theta x + (1-\theta)y)$$

$$\begin{aligned} P(\theta x + (1-\theta)y) &= \frac{\theta \widetilde{x} + (1-\theta)\widetilde{y}}{\theta x_{n+1} + (1-\theta)y_{n+1}} \\ &= \frac{\theta x_{n+1}}{\theta x_{n+1} + (1-\theta)y_{n+1}} \cdot \frac{\widetilde{x}}{x_{n+1}} + \frac{(1-\theta)y_{n+1}}{\theta x_{n+1} + (1-\theta)y_{n+1}} \cdot \frac{\widetilde{y}}{y_{n+1}} \end{aligned}$$

Perspective Transformation (4)

- We should prove that $\theta x + (1-\theta)y$ is a segment after the projection
- **Proof:**

$$\theta x + (1-\theta)y \rightarrow P(\theta x + (1-\theta)y)$$

$$P(\theta x + (1-\theta)y) = \frac{\theta \widetilde{x} + (1-\theta)\widetilde{y}}{\theta x_{n+1} + (1-\theta)y_{n+1}}$$

$$= \frac{\theta x_{n+1}}{\theta x_{n+1} + (1-\theta)y_{n+1}} \cdot \frac{\widetilde{x}}{x_{n+1}} + \frac{(1-\theta)y_{n+1}}{\theta x_{n+1} + (1-\theta)y_{n+1}} \cdot \frac{\widetilde{y}}{y_{n+1}}$$

- Let $\mu = \frac{\theta x_{n+1}}{\theta x_{n+1}+(1-\theta)y_{n+1}}$. Note that $0 \leq \mu \leq 1$, we have

$$= \mu \cdot \frac{\widetilde{x}}{x_{n+1}} + (1-\mu) \cdot \frac{\widetilde{y}}{y_{n+1}}$$

$$P(\theta x + (1-\theta)y) = \mu \cdot P(x) + (1-\mu) \cdot P(y)$$