MAEG4070 Engineering Optimization

Lecture 4 Convex Sets & Convex Functions

Yue Chen
MAE, CUHK
email: yuechen@mae.cuhk.edu.hk
Sep 19, 2022

Content of this course (tentative)

- **Optimization** — Lecture 1
  - *Theory*
    - **Linear**
      - Linear programming — Lecture 2
      - Dual Theory – Part I — Lecture 3
    - Linearization technique — Lecture 7
    - **Non-Linear** — Lecture 4
      - Unconstrained optimization — Lecture 5 & 6
      - Constrained optimization — Lecture 8 & 9
      - Dual Theory – Part II — Lecture 10
    - Distributed optimization — Lecture 11
    - Multi-objective optimization — Lecture 12
    - Robust optimization — Lecture 13
  - Programming — Lecture 14
  - Engineering examples — Lecture 15

Affine Sets

line passing through points x_1 and x_2

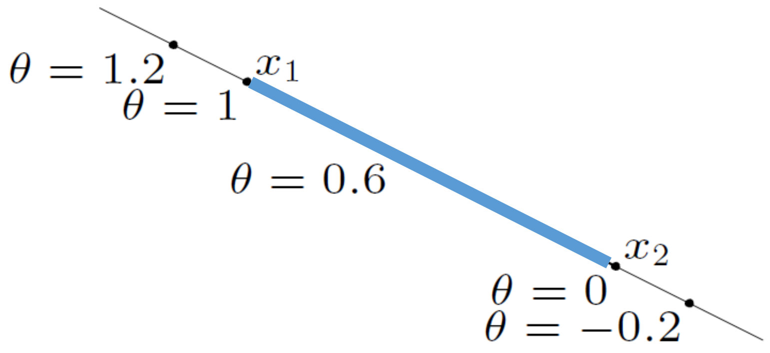
$$y = \theta x_1 + (1-\theta)x_2, \forall\theta \in \mathbb{R}$$

line segment between points x_1 and x_2

$$y = \theta x_1 + (1-\theta)x_2, \forall\theta \in [0,1]$$

Affine set: the set that contains all line through any two distinct points in the set $\mathcal{C}$

$$\forall x_1, x_2 \in \mathcal{C}, \theta \in \mathbb{R} \;\Rightarrow\; \theta x_1 + (1-\theta)x_2 \in \mathcal{C}$$



Pictures from Google image

Affine Sets

Example: solution set of linear equations $\mathcal{X} = \{x | Ax = b\}$ is an affine set.

Proof: Given any two points $x_1 \in \mathcal{X}$ and $x_2 \in \mathcal{X}$, for any $\theta \in \mathbb{R}$,
then $\theta x_1 + (1-\theta)x_2$ represents a point on the line crossing x_1 and x_2.

Since $Ax_1 = b$ and $Ax_2 = b$, we have

$$A[\theta x_1 + (1-\theta)x_2] = \theta A x_1 + (1-\theta) A x_2 = b$$

Therefore, $\theta x_1 + (1-\theta)x_2 \in \mathcal{X}$.

Affine Sets

Example: solution set of linear equations $\mathcal{X} = \{x|Ax = b\}$

Affine combination of points $x_1, \ldots, x_n$ is

$$x = \theta_1 x_1 + \cdots + \theta_n x_n,\ \ \theta_1 + \cdots + \theta_n = 1$$

Affine hull of set $\mathcal{C}$ is

- The smallest affine set that contains $\mathcal{C}$
- Set of all affine combinations of points in $\mathcal{C}$

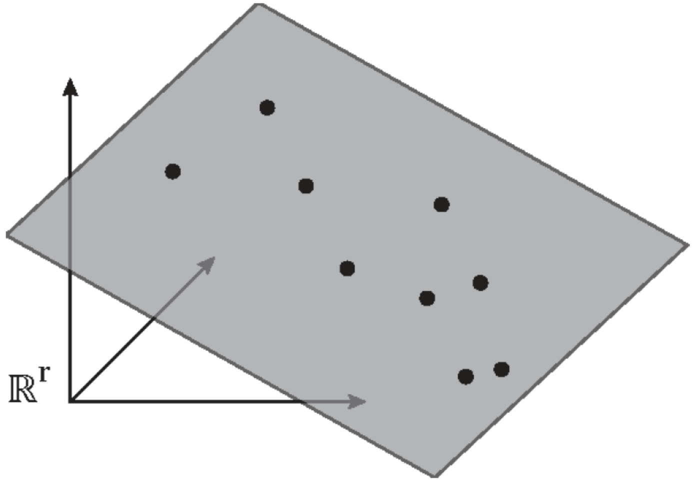
$$x = \mu_1 x_1 + \cdots + \mu_n x_n$$

$$y = \sigma_1 y_1 + \cdots + \sigma_n y_n$$

$$\forall\theta, \theta x + (1-\theta)y$$
$$= (\theta\mu_1 + (1-\theta)\sigma_1)y_1 + \cdots$$
$$+ (\theta\mu_n + (1-\theta)\sigma_n)y_n$$

⇨

$$(\theta\mu_1 + (1-\theta)\sigma_1) + \cdots$$
$$+ (\theta\mu_n + (1-\theta)\sigma_n)$$
$$= \theta\sum_{i=1}^{n}\mu_i + (1-\theta)\sum_{i=1}^{n}\sigma_i$$
$$= \theta + (1-\theta) = 1$$

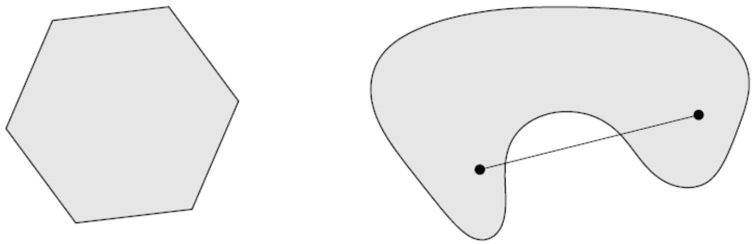


Pictures from Google image

Convex Sets

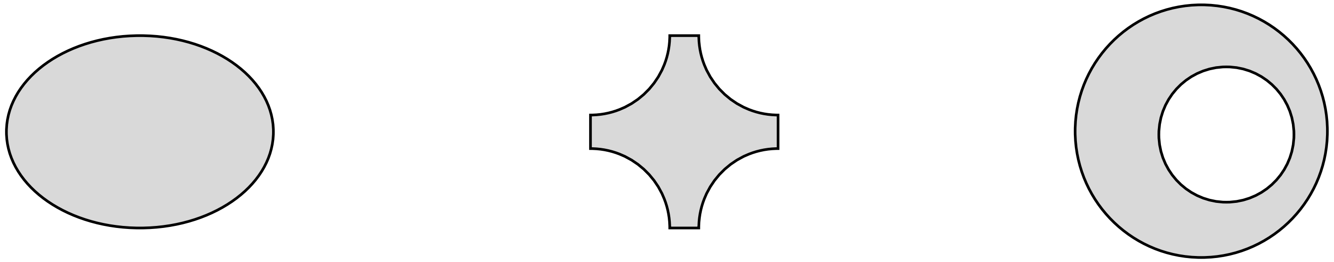
Convex set: the set that contains all line segment between any two distinct points in the set $\mathcal{C}$

$$\forall x_1, x_2 \in \mathcal{C}, \theta \in [0,1] \;\Rightarrow\; \theta x_1 + (1-\theta) x_2 \in \mathcal{C}$$



Intuitive explanation: in a convex set, you can see everywhere wherever you stand

Try it yourself: Are the following sets convex?



Pictures from Google image

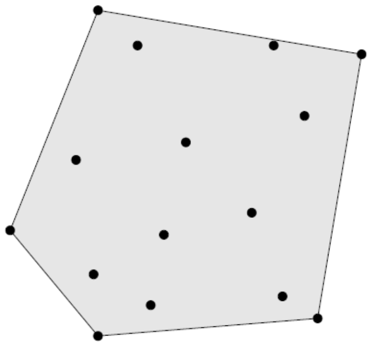
Convex Sets

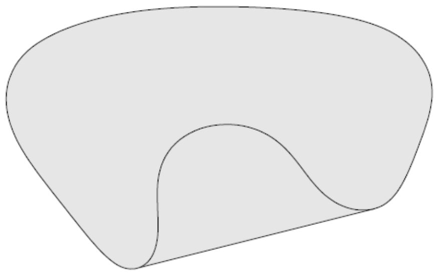
Convex combination of points $x_1, \dots, x_n$ is

$$x = \theta_1 x_1 + \cdots + \theta_n x_n,\ \theta_1 + \cdots + \theta_n = 1, \theta_k \geq 0, \forall k = 1, \dots, n$$

Convex hull of set $\mathcal{C}$ is

- The smallest convex set that contains $\mathcal{C}$
- Set of all convex combinations of points in $\mathcal{C}$





Pictures from Google image

Cones

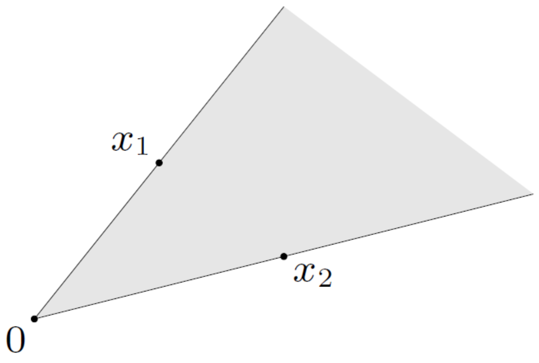
Cone: if for every $x \in \mathcal{C}$ and $\theta \geq 0$, we have $\theta x \in \mathcal{C}$

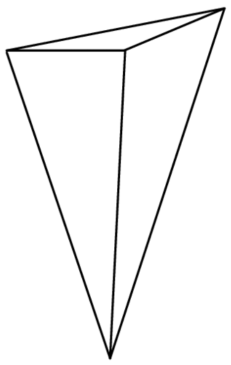
$$\forall x \in \mathcal{C}, \theta \geq 0 \;\Rightarrow\; \theta x \in \mathcal{C}$$

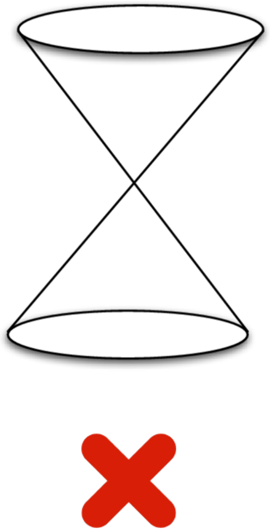
$\theta_1 x_1 \in \mathcal{C}$ (cone)
$\theta_2 x_2 \in \mathcal{C}$ (cone)
$0.5\theta_1 x_1 + 0.5\theta_2 x_2 \in \mathcal{C}$ (convex)
$2(0.5\theta_1 x_1 + 0.5\theta_2 x_2) \in \mathcal{C}$ (cone)

Convex cone: if $\mathcal{C}$ is convex and also a cone

$$\forall x_1, x_2 \in \mathcal{C}, \theta_1, \theta_2 \geq 0 \;\Rightarrow \theta_1 x_1 + \theta_2 x_2 \in \mathcal{C}$$







Pictures from Google image

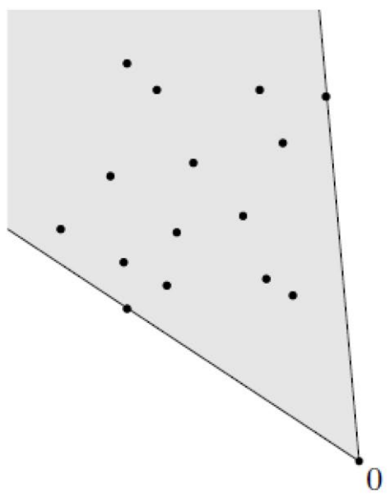
Cones

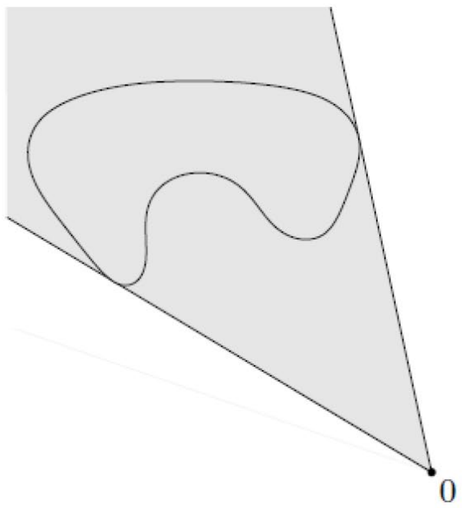
Conic combination of points $x_1, \dots, x_n$ is

$$x = \theta_1 x_1 + \cdots + \theta_n x_n, \theta_k \geq 0, \forall k = 1, \dots, n$$

Conic hull of set $\mathcal{C}$ is

- The smallest cone that contains $\mathcal{C}$
- Set of all conic combinations of points in $\mathcal{C}$





Pictures from Google image

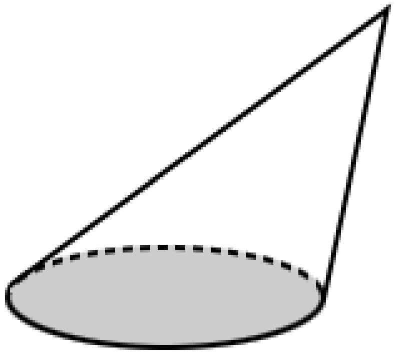
Comparison of affine set, convex set, cone

Affine set: $\forall x_1, x_2 \in \mathcal{C}, \theta \in \mathbb{R} \;\Rightarrow\; \theta x_1 + (1-\theta)x_2 \in \mathcal{C}$

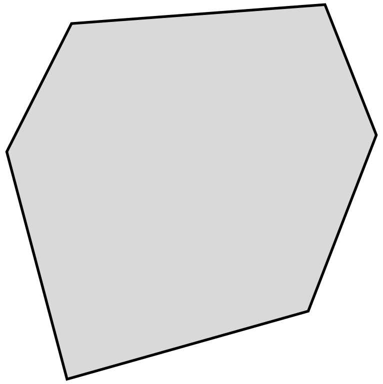
Convex set: $\forall x_1, x_2 \in \mathcal{C}, \theta \in [0,1] \;\Rightarrow\; \theta x_1 + (1-\theta)x_2 \in \mathcal{C}$

Cone: $\forall x \in \mathcal{C}, \theta \geq 0 \;\Rightarrow\; \theta x \in \mathcal{C}$

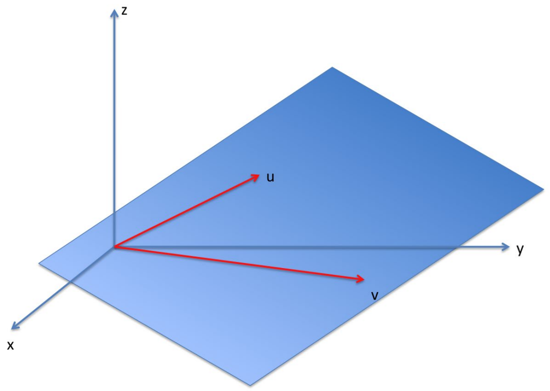
Try to identify: are these affine set, convex set, or cone?



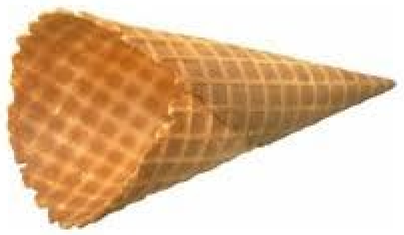
cone



convex set



affine set

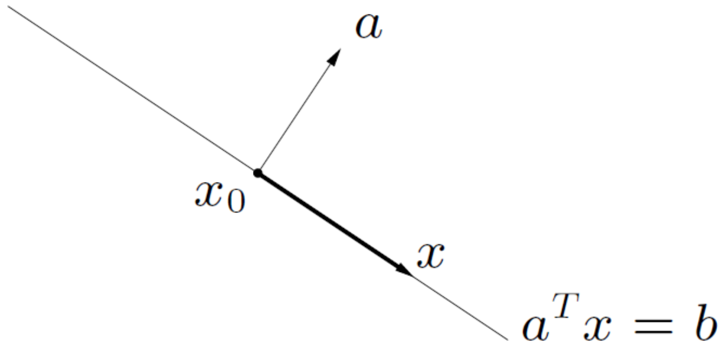


cone

Pictures from Google image

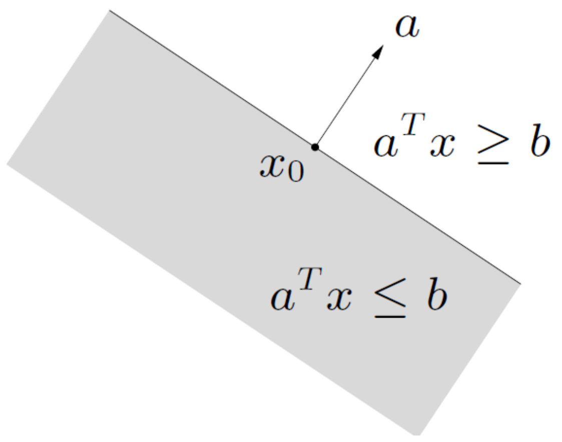
Polyhedron

Hyperplane: set of the form $\{x \mid a^T x = b, a \neq 0\}$



Hyperplane is affine and convex

Halfspace: set of the form $\{x \mid a^T x \leq b, a \neq 0\}$

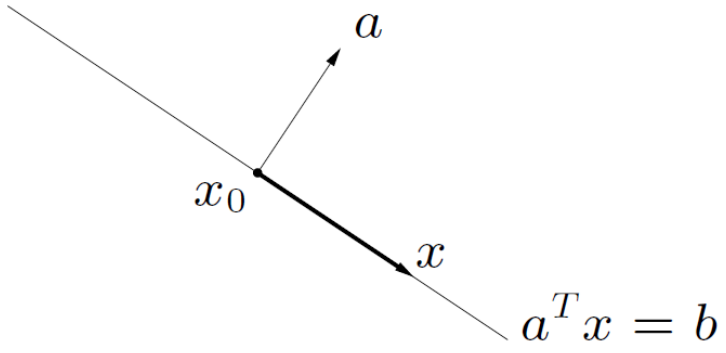


halfspace is convex

Pictures from Google image

Polyhedron

Hyperplane: set of the form $\{x \mid a^T x = b, a \neq 0\}$



Hyperplane is affine and convex

Proof: We have proved that $S = \{x \mid a^T x = b, a \neq 0\}$ is an affine set.
Next, let's prove it is convex. Suppose we have two points x_1, x_2, then for any $\theta \in [0,1]$, we have

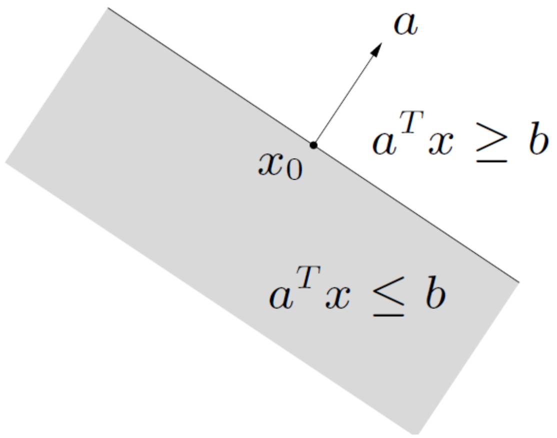
$$a(\theta x_1 + (1-\theta)x_2) = \theta a x_1 + (1-\theta)a x_2 = \theta b + (1-\theta)b = b$$

Therefore, we have $\theta x_1 + (1-\theta)x_2 \in S$.

Pictures from Google image

Polyhedron

Halfspace: set of the form $\{x \mid a^T x \leq b, a \neq 0\}$



halfspace is convex

Proof: Suppose we have two points x_1, x_2, then for any $\theta \in [0,1]$, we have

$$a(\theta x_1 + (1-\theta)x_2) = \theta a x_1 + (1-\theta) a x_2 \leq \theta b + (1-\theta) b = b$$

Therefore, we have $\theta x_1 + (1-\theta)x_2 \in S$.

not affine because we need $\theta \geq 0$

Pictures from Google image

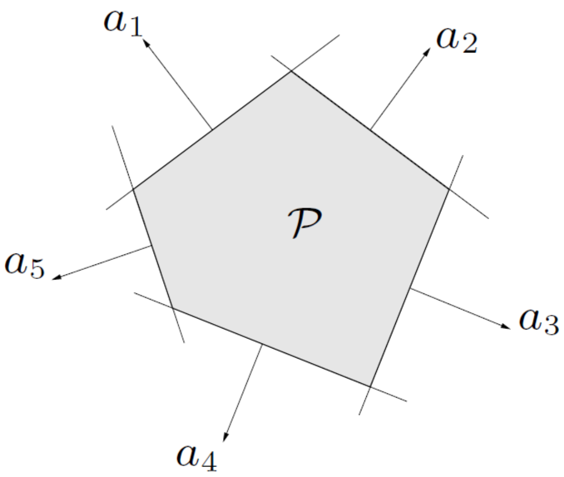
Polyhedron

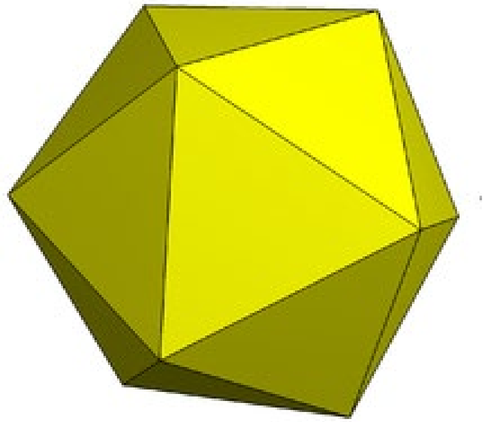
A polyhedron is defined as the solution set of linear equalities and inequalities

$$P = \{x | a_j^T x \leq b_j, j = 1, \ldots, m, c_k^T x = d_k, k = 1, \ldots, p\}$$

Or in a compact form

$$P = \{x | Ax \leq b, Cx = d\}$$





Polyhedron is intersection of finite number of halfspaces and hyperplanes

Pictures from Google image

How to prove a set $\mathcal{C}$ is convex?

To prove a set $\mathcal{C}$ is convex, we can

- Apply definition

$$\forall x_1, x_2 \in \mathcal{C}, \theta \in [0,1] \quad \Rightarrow \quad \theta x_1 + (1-\theta) x_2 \in \mathcal{C}$$

- Show that $\mathcal{C}$ is obtained from simple convex sets by operations that preserve convexity, e.g.
  - ✓ Intersection
  - ✓ Affine mapping
  - ✓ Perspective mapping
  - ✓ Linear-fractional mapping
  - ✓ epigraph

Example - 1

Prove that the set $S = \{(x_1, x_2) | x_1 + x_2 \leq 6, -x_1 + 2x_2 \leq 1\}$ is convex.

Proof: Given any $x = (x_1, x_2) \in S, y = (y_1, y_2) \in S$. For any $\theta \in [0,1]$, we have

$$\theta x + (1-\theta)y = (\theta x_1 + (1-\theta)y_1, \theta x_2 + (1-\theta)y_2)$$

Then we check $\theta x + (1-\theta)y$ is in S

$$\theta x_1 + (1-\theta)y_1 + \theta x_2 + (1-\theta)y_2 = \theta(x_1 + x_2) + (1-\theta)(y_1 + y_2) \leq 6$$

$$-\theta x_1 - (1-\theta)y_1 + 2\theta x_2 + 2(1-\theta)y_2 = \theta(-x_1 + 2x_2) + (1-\theta)(-y_1 + 2y_2) \leq 1$$

Therefore,$\theta x + (1-\theta)y \in S$.

Example -2 Positive semidefinite cone

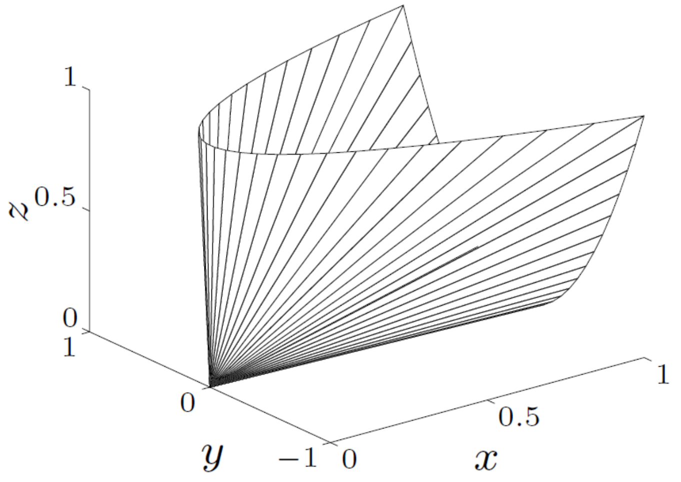
Let $\boldsymbol{S}^n$ be a set of symmetric $n \times n$ matrices, define

$$\mathbf{S}^n_+ = \{X \in \mathbf{S}^n | X \succeq 0\}$$
$$\mathbf{S}^n_{++} = \{X \in \mathbf{S}^n | X \succ 0\}$$

Convex cone (try to prove it)

Then we have

$$X \in \mathbf{S}^n_+ \;\Rightarrow\; z^T X z \geq 0 \text{ for all } z \in \mathbb{R}^n$$



Pictures from Google image

Example -2 Positive semidefinite cone

$$\mathbf{S}_+^n = \{X \in \mathbf{S}^n | X \succeq 0\}$$

Proof: 1) ***Convex***

$$\forall X_1, X_2 \in X \in \mathbf{S}_+^n \text{ and } \theta \in [0,1], \text{ we have for all z}$$

$$\begin{aligned} & z^T(\theta X_1 + (1-\theta)X_2)z \\ = \ & \theta z^T X_1 z + (1-\theta) z^T X_2 z \\ \geq \ & 0 \end{aligned}$$

$$\text{Therefore, } (\theta X_1 + (1-\theta)X_2) \in \mathbf{S}_+^n.$$

Example -2 Positive semidefinite cone

$$\mathbf{S}_+^n = \{X \in \mathbf{S}^n | X \succeq 0\}$$

Proof: 2) ***Cone***

$\forall X \in \mathbf{S}_+^n$ and $\theta \geq 0$, we have for all z

$$z^T \theta X z = \theta z^T X z \geq 0$$

Therefore, $\theta X \in \mathbf{S}_+^n$

Example – 3 Intersection of two sets

Suppose we have two convex sets $\mathcal{C}$ and $\mathcal{S}$. Then, let's prove $\mathcal{C} \cap \mathcal{S}$ is a convex set.

Proof:

Given any $x_1, x_2 \in \mathcal{C} \cap \mathcal{S}$, then we have $x_1, x_2 \in \mathcal{C}$ and $x_1, x_2 \in \mathcal{S}$

Since $\mathcal{C}$ and $\mathcal{S}$ are all convex sets, then given any $\theta \in [0,1]$

We have

$$\theta x_1 + (1-\theta)x_2 \in \mathcal{C}$$
$$\theta x_1 + (1-\theta)x_2 \in \mathcal{S}$$

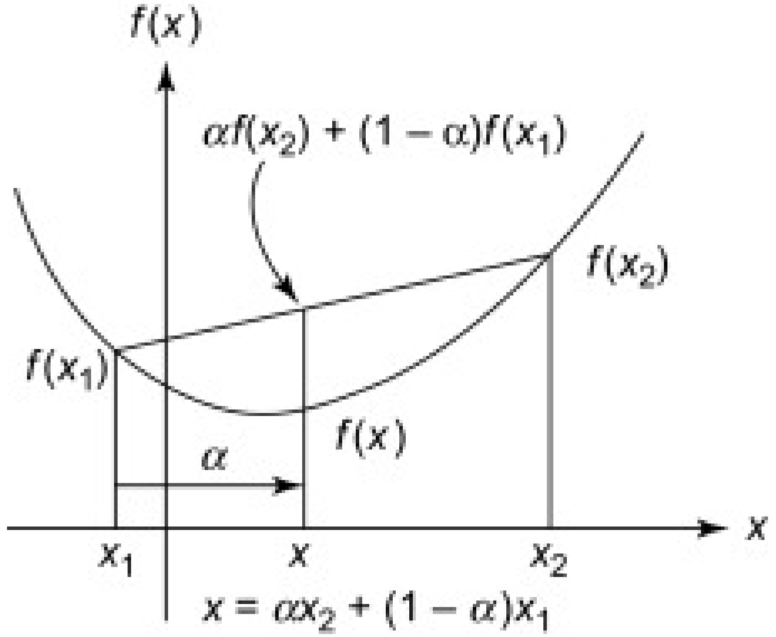
Therefore,

$$\theta x_1 + (1-\theta)x_2 \in \mathcal{C} \cap \mathcal{S}$$

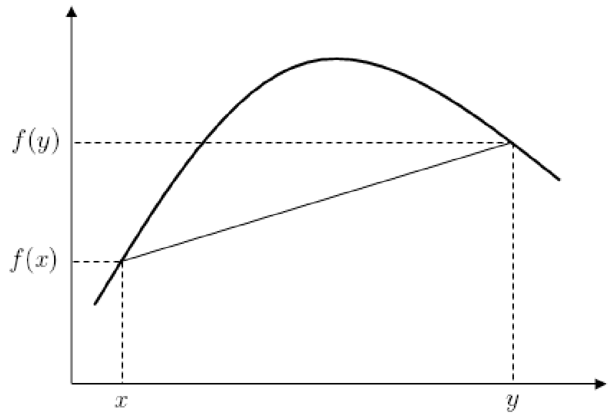
Convex function

Function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ is **convex** if dom(f) is a convex set, and the following inequality holds

$$f(\theta x_1 + (1-\theta)x_2) \leq \theta f(x_1) + (1-\theta) f(x_2), \forall \theta \in [0,1], \forall x_1, x_2 \in dom(f)$$



If we change $\leq$ into $\geq$, then it is **concave**



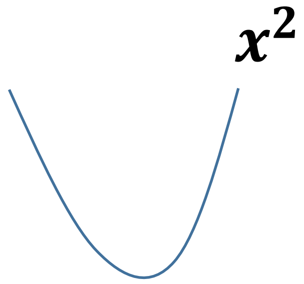
Pictures from Google image

Convex function

Function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is **strictly convex** if dom(f) is a convex set, and the following inequality holds

$$f(\theta x_1 + (1-\theta)x_2) < \theta f(x_1) + (1-\theta) f(x_2), \forall \theta \in (0,1), \forall x_1 \neq x_2 \in dom(f)$$

Function f is strongly convex if $\exists \alpha > 0$: $f(x) - \alpha \|x\|_2^2$ is convex



f is (strictly, strongly) concave if $-f$ is (strictly, strongly) convex

stronger

Gradient

The *gradient* of a scalar-valued differentiable function f of several variables is the vector field (or vector-valued function) ∇f, whose value at a point x is the vector, whose components are the partial derivatives of f at x. i.e.

$$\nabla f(x) = \left[\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_N}\right]^T$$

For example, $f(x) = 3x_1^2 + 5x_2 + x_1 x_2$, then

$$\frac{\partial f}{\partial x_1} = 6x_1 + x_2, \frac{\partial f}{\partial x_2} = 5 + x_1$$

Therefore

$$\nabla f(x) = [6x_1 + x_2, 5 + x_1]^T$$

Review of mathematics

Hessian

the Hessian matrix or Hessian is a square matrix of second-order partial derivatives of a scalar-valued function, or scalar field

$$\mathbf{H}(\mathbf{x}^*) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_N} \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_N \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_N^2} \end{bmatrix}$$

$$\frac{\partial f}{\partial x_1} = 6x_1 + x_2, \frac{\partial f}{\partial x_2} = 5 + x_1$$

For example, $f(x) = 3x_1^2 + 5x_2 + x_1 x_2$, we have

$$\frac{\partial^2 f}{\partial x_1^2} = 6, \frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1} = 1, \frac{\partial^2 f}{\partial x_2^2} = 0$$

Therefore

$$H(x) = \begin{bmatrix} 6 & 1 \\ 1 & 0 \end{bmatrix}$$

For twice continuously differentiable functions, Hessian is always symmetric. (which can be used to double check the calculation of cross term)

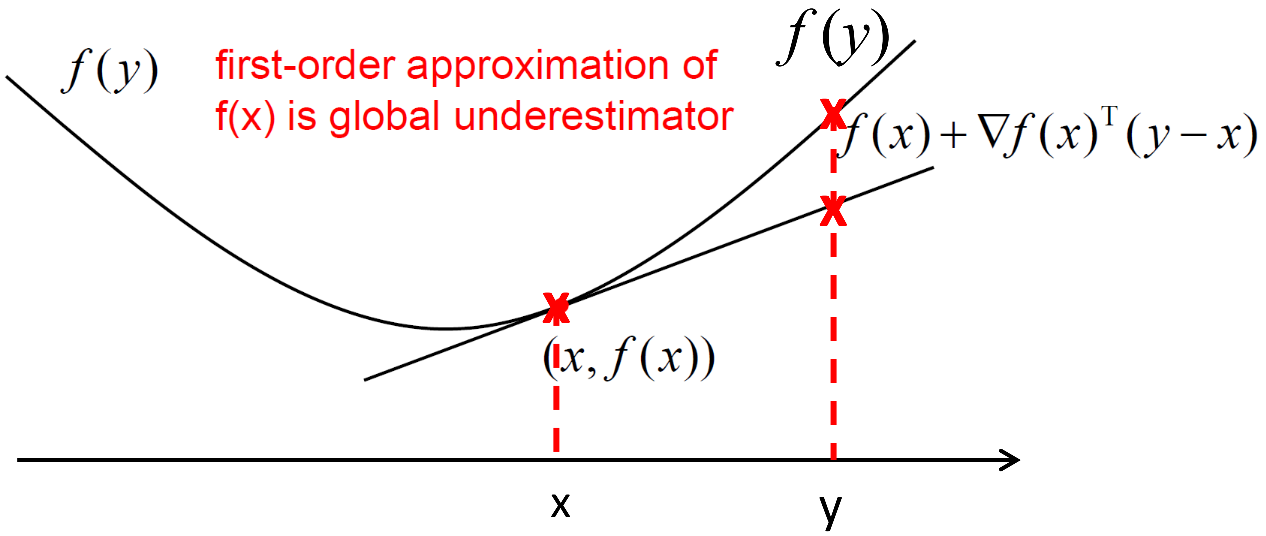
Convex function

Apart from proving the convexity by definition, in the following, we provide two conditions, i.e. first-order condition & second-order condition

Suppose f is differentiable and $\nabla f(x)$ exists at each $x \in dom(f)$

First-order condition f with convex domain is convex iff

$$f(y) \geq f(x) + \nabla f(x)^T (y - x), \forall x, y \in dom(y)$$



Pictures from Google image

Convex function

Simple proof:

$ty + (1-t)x$

Necessity: consider $x + t(y - x) \in dom(f), \forall t \in [0, 1]$. Since $f(x)$ is convex, then we have

$$f(x + t(y - x)) \leq (1 - t)f(x) + tf(y)$$

which is equivalent to

$$f(y) \geq f(x) + \frac{f(x + t(y - x)) - f(x)}{t}$$

Let $t \to 0_+$ we can get the conclusion.

Convex function

Sufficiency: for any $x \neq y$, let $z = tx + (1-t)y, \forall t \in [0,1]$

$$f(x) \geq f(z) + f^{'}(z)^T(x-z) \qquad \times t$$

$$f(y) \geq f(z) + f^{'}(z)^T(y-z) \qquad \times (1-t)$$

$$tf(x) + (1-t)f(y) \geq f(z) + (tx + (1-t)y)f^{'}(z) - zf^{'}(z)$$

$$tf(x) + (1-t)f(y) \geq f(z)$$

This shows the convexity of $f(x)$.

Convex function

Suppose f is twice differentiable and the Hessian $H(x)$ exists at every $x \in dom(f)$.

Second-order condition function f with convex domain is

- convex iff

positive semidefinite

$$H(x) \succeq 0, \forall x \in dom(f)$$

$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is positive semidefinite iff $a \geq 0$ and $ad - bc \geq 0$

- Strictly convex iff

positive definite

$$H(x) \succ 0, \forall x \in dom(f)$$

- Strongly convex iff

$$H(x) - \alpha I \succeq 0, \forall x \in dom(f)$$

Example

Convex functions

- Affine: $ax + b$ on $\mathbb{R}$ for any *a* and *b*
- Quadratic function: $ax^2 + bx + c$ on $\mathbb{R}$ for any $a \geq 0$
- Exponential: e^{ax} on $\mathbb{R}$ for any *a*
- Negative entropy: $xlog(x)$ on $\mathbb{R}_{++}$

Try to prove $f(x) = ax^2 + bx + c$ on $\mathbb{R}$ for any $a \geq 0$ is convex.

Proof:

$$f'(x) = 2ax + b, f''(x) = 2a \geq 0$$

According to the second-order condition, it is convex.

Epigraph

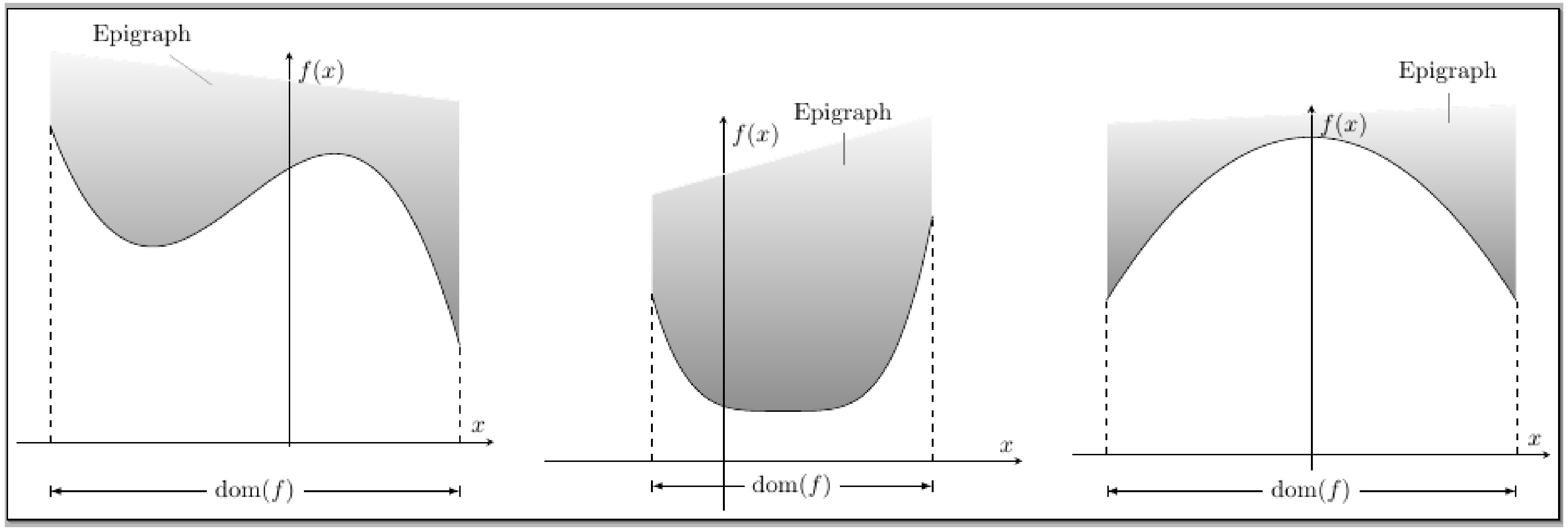
The **graph** of a function $f: \mathbb{R}^n \to \mathbb{R}$ is defined as

$$\{(x, f(x)) \mid x \in dom(f)\} \subseteq \mathbb{R}^{n+1}$$

The **epigraph** of a function $f: \mathbb{R}^n \to \mathbb{R}$ is defined as

$$epi\ f = \{(x, t) \mid x \in dom(f), f(x) \leq t\} \subseteq \mathbb{R}^{n+1}$$

$f(x)$ is convex if and only if $epi\ f$ is a convex set.



Pictures from Google image

How to prove a function $f(x)$ is convex?

To prove a function $f(x)$ is convex, we can

- Verify definition
- For twice differentiable functions, apply *second-order condition*
- Show that $f(x)$ is obtained from simple convex functions by operations that preserve convexity, e.g.
  - ✓ Nonnegative weighted sum
  - ✓ Composition with affine function
  - ✓ Pointwise maximum
  - ✓ Composition
  - ✓ Minimization
  - ✓ Perspective

Example - 1

Prove that $f(x_1, x_2) = x_1^2 - 2x_1x_2 + 4x_2^2 + 3x_1$ is convex.

Proof: The gradient of $f(x_1, x_2) = x_1^2 - 2x_1x_2 + 4x_2^2 + 3x_1$ is

$$\frac{\partial f}{\partial x_1} = 2x_1 - 2x_2 + 3, \qquad \frac{\partial f}{\partial x_2} = -2x_1 + 8x_2$$

The Hessian is

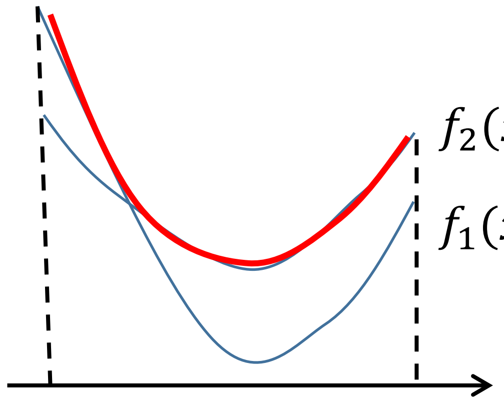
$$H(x) = \begin{bmatrix} 2 & -2 \\ -2 & 8 \end{bmatrix}$$

$2 \geq 0$

$2 \times 8 - (-2) \times (-2) = 16 - 4 = 12 \geq 0$

$H(x)$ is positive semi-definite. Therefore, $f(x_1, x_2)$ is a convex function.

Example – 2 Pointwise maximum of convex functions



If $f_1, \ldots, f_m$ are convex, then $f(x) = \max_i\{f_i(x)\}$ is convex

$$
\begin{aligned}
f(tx + (1-t)y) &= \max_i\{f_i(tx + (1-t)y)\} \\
&\le \max_i\{t f_i(x) + (1-t) f_i(y)\} \\
&\le t \max_i\{f_i(x)\} + (1-t) \max_i\{f_i(y)\} \\
&= t f(x) + (1-t) f(y)
\end{aligned}
$$

Suppose the maximum is the i^* term

$$\max_i\{f_i(x)\} \ge f_{i^*}(x)$$
$$\max_i\{f_i(y)\} \ge f_{i^*}(y)$$
$$t \max_i\{f_i(x)\} + (1-t) \max_i\{f_i(y)\}$$
$$\ge t f_{i^*}(x) + (1-t) f_{i^*}(y)$$
$$= \max_i\{t f_i(x) + (1-t) f_i(y)\}$$

Example: piecewise-linear functions

$$f(x) = \max\{a_1^T x + b_1, \ldots, a_m^T x + b_m\}$$

Example – 3 Minimization

$\mathcal{C}$ is a convex set, $f(x,y)$ is convex in the (x,y) space, then the function

$$g(x) = \min_{y \in \mathcal{C}} f(x,y)$$

is convex in x

Proof: For any x_1 and x_2, the minimizers are y_1 and y_2

$$g(x_1) = f(x_1, y_1),\ g(x_2) = f(x_2, y_2)$$

$$\begin{aligned} g(tx_1 + (1-t)x_2) &= \min_{y \in \mathcal{C}} f(tx_1 + (1-t)x_2, y) \\ &\leq f(tx_1 + (1-t)x_2, ty_1 + (1-t)y_2) \\ &\leq tf(x_1, y_1) + (1-t)f(x_2, y_2) \\ &= tg(x_1) + (1-t)g(x_2) \end{aligned}$$

Thanks！