Introduction to Convex Optimization

EE/CS/EST 135

Feb 12, 2018

Outline

- Motivation
- Recap of Linear Algebra and Real Analysis
- Convex Set
- Convex Function
- Convex Program

Outline

- **Motivation**
- Recap of Linear Algebra and Real Analysis
- Convex Set
- Convex Function
- Convex Program

Motivation

Motivation

- Many practical problems can be modeled as optimization problems:

$$\min_{x} \; f(x) \qquad x \in X$$

  - Optimal Power Flow (OPF)
  - EV Charging Scheduling

Motivation

- Many practical problems can be modeled as optimization problems:

$$\min_x \; f(x) \qquad x \in X$$

  - Optimal Power Flow (OPF)
  - EV Charging Scheduling

- Convex program
  - X is a convex set
  - f is a convex function

Motivation

- Many practical problems can be modeled as optimization problems:

$$\min_x \ f(x) \qquad x \in X$$

  - Optimal Power Flow (OPF)
  - EV Charging Scheduling

- Convex programs have good properties
  - Certificate of global optimality
  - Efficient algorithms exist
  - Powerful modeling capability

Outline

- Motivation
- **Recap of Linear Algebra and Real Analysis**
- Convex Set
- Convex Function
- Convex Program

Recap of Linear Algebra

- Euclidean space $\mathbb{R}^n$
- Vectors: $x \in \mathbb{R}^n$ Matrices: $M \in \mathbb{R}^{m \times n}$

- Transpose: $M^T \quad x^T$
- Rank: $\operatorname{rank} M$
- Trace: $\operatorname{tr} M = \sum_i M_{ii} \qquad \operatorname{tr}(AB) = \operatorname{tr}(BA)$

Recap of Linear Algebra

- Euclidean space $\mathbb{R}^n$
- Vectors: $x \in \mathbb{R}^n$ Matrices: $M \in \mathbb{R}^{m \times n}$
- Inner product: $\langle x, y \rangle$
- Norm: $\|x\| = \sqrt{\langle x, x \rangle}$
- Orthonormal basis:

$$\{u_1, \ldots, u_n\} \qquad \langle u_i, u_j \rangle = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$$

Recap of Linear Algebra

- Euclidean space $\mathbb{R}^n$
- Vectors: $x \in \mathbb{R}^n$ Matrices: $M \in \mathbb{R}^{m \times n}$

- Standard inner product: $\langle x, y \rangle = y^T x$
- Standard norm: $\|x\| = \sqrt{\langle x, x \rangle} = \sqrt{x^T x}$
- Orthonormal basis:

$$\{u_1, \ldots, u_n\} \qquad \langle u_i, u_j \rangle = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$$

Recap of Linear Algebra

- Real symmetric matrices

$$\mathbb{S}^n = \{M \in \mathbb{R}^{n \times n} : M = M^T\}$$

Recap of Linear Algebra

- Real symmetric matrices

$$\mathbb{S}^n = \{M \in \mathbb{R}^{n\times n} : M = M^T\}$$

- Eigenvalue decomposition for $M \in \mathbb{S}^n_+$

$$M = \sum_{i=1}^{n} \lambda_i u_i u_i^T$$

  - $Mu_i = \lambda_i u_i$
  - $\{u_1, \ldots, u_n\}$ forms an orthonormal basis
  - $\mathrm{rank}(M) = \#\{i : \lambda_i \neq 0\}$

Recap of Linear Algebra

- $\mathbb{S}^n$ is a real linear space with $\dim \mathbb{S}^n = \frac{1}{2}n(n+1)$
- Inner product:

$$\langle A, B \rangle = \mathrm{tr}\left(B^T A\right) = \sum_{i,j=1}^{n} A_{ij} B_{ij}$$

- Frobenius norm:

$$\begin{aligned} \|A\|_F &:= \sqrt{\langle A, A \rangle} = \sqrt{\sum_{i,j=1}^{n} A_{ij}^2} \\ &= \sqrt{\sum_{i=1}^{n} \lambda_i^2} \end{aligned}$$

Recap of Linear Algebra

- Positive semidefinite (PSD) matrices

$$M = M^T \quad \text{and} \quad x^T M x \geq 0 \quad \forall x \in \mathbb{R}^n$$

- Notation: $M \succeq 0$

Recap of Linear Algebra

- Positive semidefinite (PSD) matrices

$$M = M^T \quad \text{and} \quad x^T M x \geq 0 \quad \forall x \in \mathbb{R}^n$$

- Notation: $M \succeq 0$
- A real symmetric matrix M is PSD iff one of the following holds:
  - The eigenvalues of M are all nonnegative
  - $M = AA^T$ for some matrix A

Recap of Linear Algebra

- A real symmetric matrix M is PSD iff one of the following holds:
  - The eigenvalues of M are all nonnegative
  - $M = AA^T$ for some matrix A

Recap of Linear Algebra

- A real symmetric matrix M is PSD iff one of the following holds:
  - The eigenvalues of M are all nonnegative
  - $M = AA^T$ for some matrix A
- Corollary: A real symmetric matrix M is equal to xx^T for some $x \in \mathbb{R}^n$ iff

$$M \succeq 0 \quad \text{and} \quad \operatorname{rank} M \leq 1$$

Recap of Linear Algebra

- Complex linear space $\mathbb{C}^n$
- Complex transpose: $M^H \quad x^H$
- Hermitian matrix: $M = M^H$
- PSD matrix: $M = M^H$ and $x^H M x \geq 0 \; \forall x \in \mathbb{C}^n$

Recap of Linear Algebra

- Complex linear space $\mathbb{C}^n$
- Complex transpose: $M^H \quad x^H$
- Hermitian matrix: $M = M^H$
- PSD matrix: $M = M^H$ and $x^H M x \geq 0 \; \forall x \in \mathbb{C}^n$

- $M \in \mathbb{C}^{n \times n}$ is PSD iff

$$\begin{bmatrix} \operatorname{Re} M & \operatorname{Im} M \\ -\operatorname{Im} M & \operatorname{Re} M \end{bmatrix} \in \mathbb{S}^{2n} \text{ and is PSD}$$

Recap of Real Analysis

Recap of Real Analysis

- Open ball $B_r(x) := \{y : \|y - x\| < r\}, \quad r > 0$

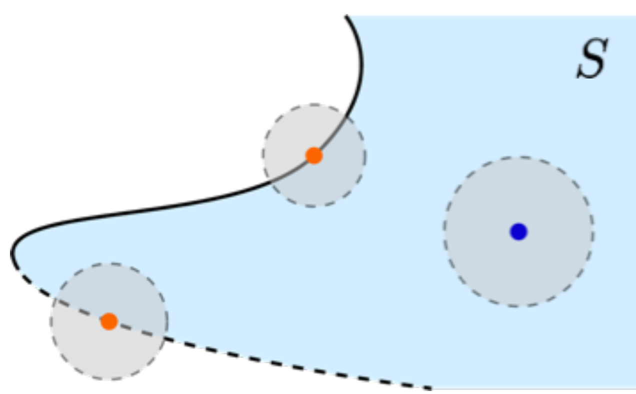
Recap of Real Analysis

- Open ball $B_r(x) := \{y : \|y - x\| < r\}, \quad r > 0$
- Interior point $x \in \operatorname{int} S$

  There exists some open ball $B_r(x) \subseteq S$

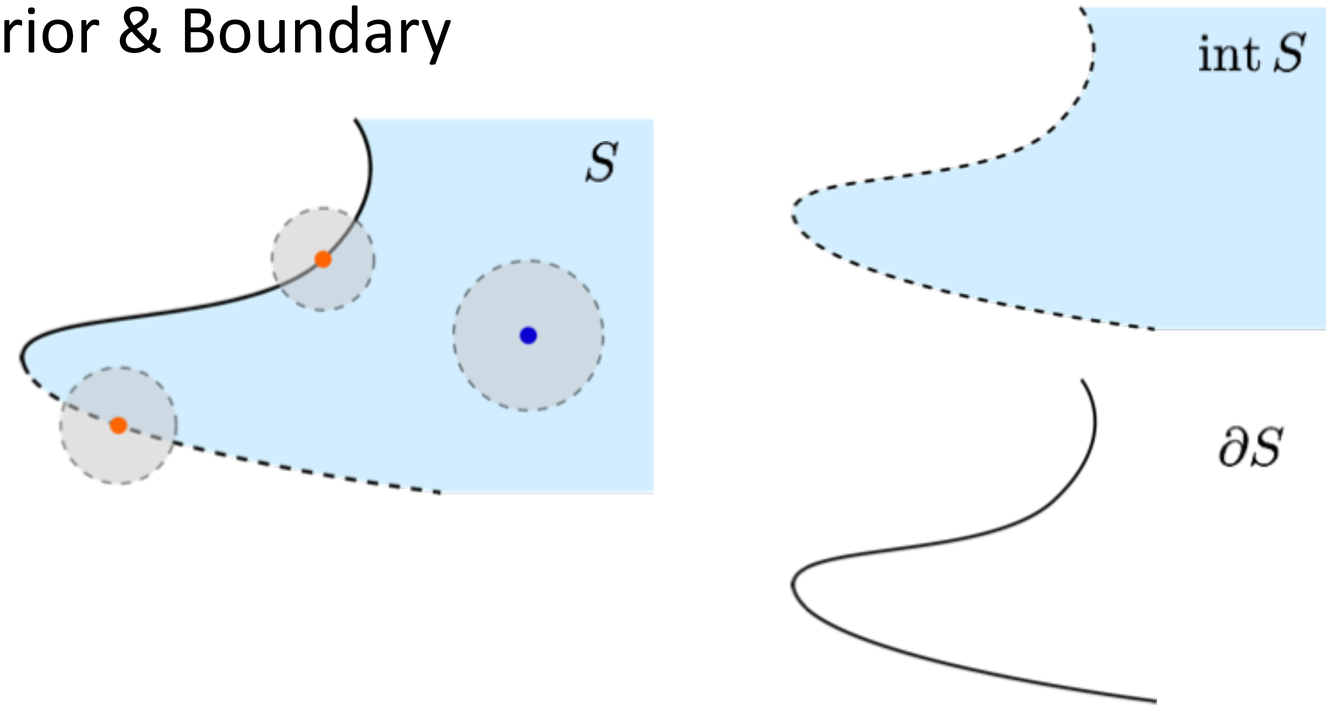
- Boundary point $x \in \partial S$

  For all $r > 0$, $B_r(x) \not\subseteq S$ and $B_r(x) \cap S \neq \varnothing$



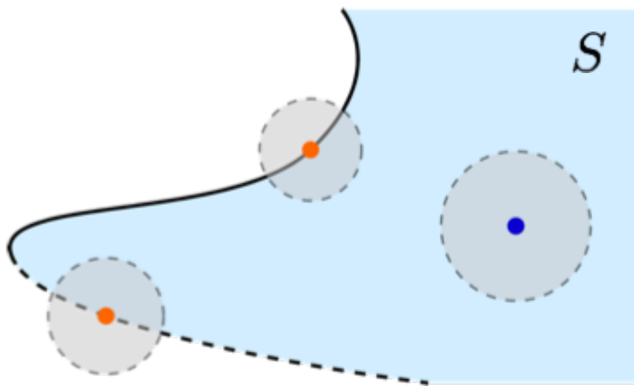
Recap of Real Analysis

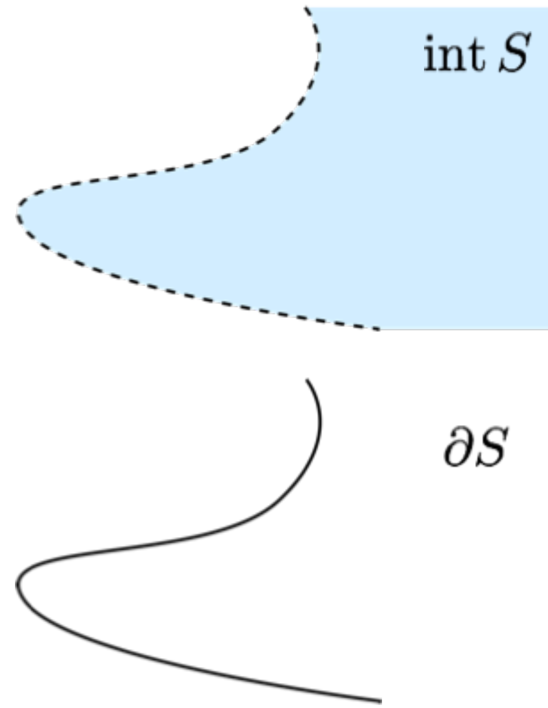
- Interior & Boundary



Recap of Real Analysis

- Interior & Boundary

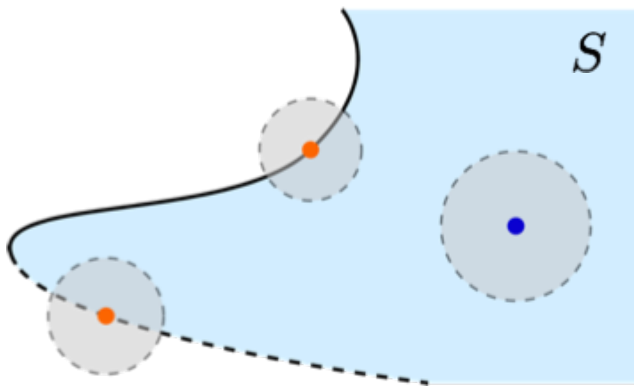


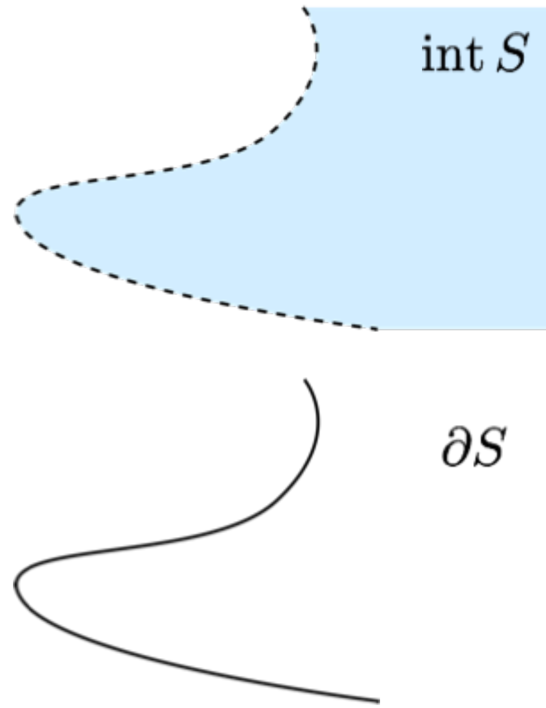


- Open set: $\text{int}\, S = S$
- Closed set: $\partial S \subseteq S$

Recap of Real Analysis

- Interior & Boundary

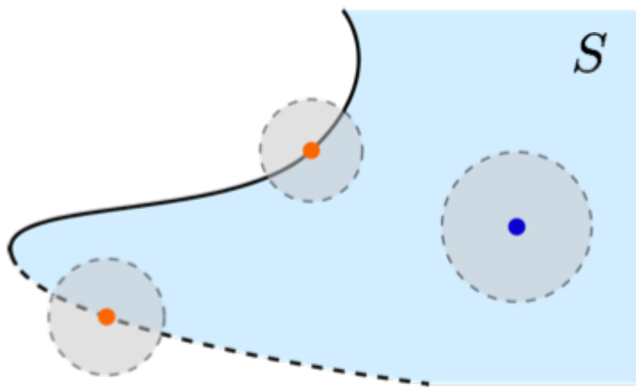


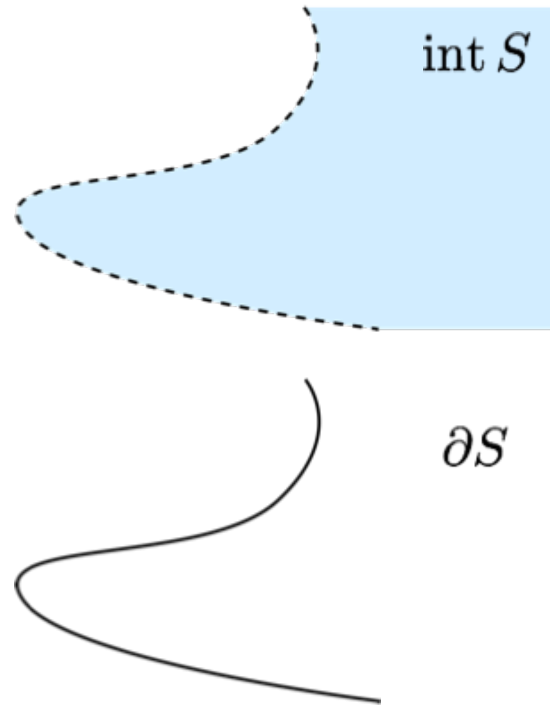


- Open set: $\mathrm{int}\, S = S$
- Closed set: $\partial S \subseteq S$
  - Complement of an open (closed) set is closed (open)
  - $\bigcup$ open sets is open, $\bigcap$ closed sets is closed

Recap of Real Analysis

- Interior & Boundary





- Open set: $\mathrm{int}\, S = S$
- Closed set: $\partial S \subseteq S$

- Bounded set: there exists some $r > 0$ s.t. $S \subseteq B_r(0)$

Recap of Real Analysis

- Compact set

Recap of Real Analysis

- Compact set
  - (Definition) Any open cover has a finite subcover

Recap of Real Analysis

- Compact set
  - (Definition) Any open cover has a finite subcover
  - (Heine - Borel) If $S \subseteq \mathbb{R}^n$, then

$$S \text{ is compact} \iff S \text{ is closed} + \text{bounded}$$

Recap of Real Analysis

- Compact set
  - (Definition) Any open cover has a finite subcover
  - (Heine - Borel) If $S \subseteq \mathbb{R}^n$, then

$$S \text{ is compact} \Longleftrightarrow S \text{ is closed} + \text{bounded}$$

- Extreme Value Theorem

  Suppose X is compact and $f : X \to \mathbb{R}$ is continuous.
  Then there exist $x_{\min}, x_{\max} \in X$ such that

$$f(x_{\min}) \leq f(x) \leq f(x_{\max}) \quad \text{for all } x \in X$$

Recap of Real Analysis

Recap of Real Analysis

$X \subseteq \mathbb{R}^n$ open, $f : X \to \mathbb{R}$

Recap of Real Analysis

$X \subseteq \mathbb{R}^n$ open, $f : X \to \mathbb{R}$

- Gradient: $f(x+h) = f(x) + \langle h, \nabla f(x) \rangle + o(\|h\|)$

$$\nabla f(x) = \left[\frac{\partial f(x)}{\partial x_i} \right]_{i=1}^{n} \in \mathbb{R}^n$$

Recap of Real Analysis

$X \subseteq \mathbb{R}^n$ open, $f : X \to \mathbb{R}$

- Gradient: $f(x+h) = f(x) + \langle h, \nabla f(x) \rangle + o(\|h\|)$

$$\nabla f(x) = \left[\frac{\partial f(x)}{\partial x_i}\right]_{i=1}^n \in \mathbb{R}^n$$

- Hessian: $\nabla f(x+h) = \nabla f(x) + H_f(x)h + o(\|h\|)$

$$H_f(x) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}\right]_{i,j=1}^n \in \mathbb{S}^n$$

Recap of Real Analysis

$X \subseteq \mathbb{R}^n$ open, $f : X \to \mathbb{R}$

- Gradient: $f(x+h) = f(x) + \langle h, \nabla f(x) \rangle + o(\|h\|)$

  $$\nabla f(x) = \left[\frac{\partial f(x)}{\partial x_i} \right]_{i=1}^{n} \in \mathbb{R}^n$$

- Hessian: $\nabla f(x+h) = \nabla f(x) + H_f(x)h + o(\|h\|)$

  $$H_f(x) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j} \right]_{i,j=1}^{n} \in \mathbb{S}^n$$

  $$f(x+h) = f(x) + \langle h, \nabla f(x) \rangle + \tfrac{1}{2} h^T H_f(x) h + o(\|h\|^2)$$

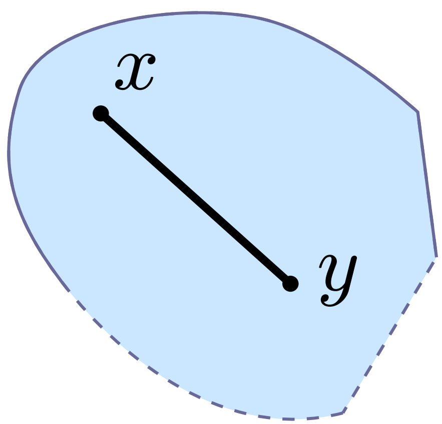
Outline

- Motivation
- Recap of Linear Algebra and Real Analysis
- **Convex Set**
- Convex Function
- Convex Program

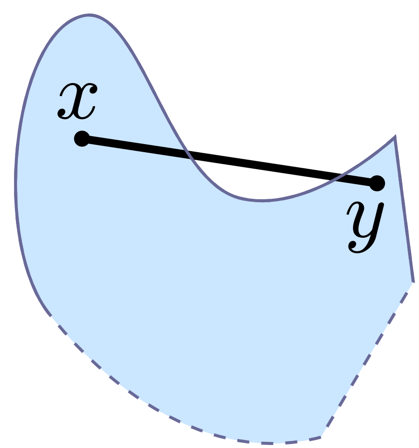
Convex Set

- Line segment: $[x, y] := \{\alpha x + (1-\alpha)y : \alpha \in [0,1]\}$

- S is called **convex** if $[x, y] \subseteq S$ for all $x, y \in S$



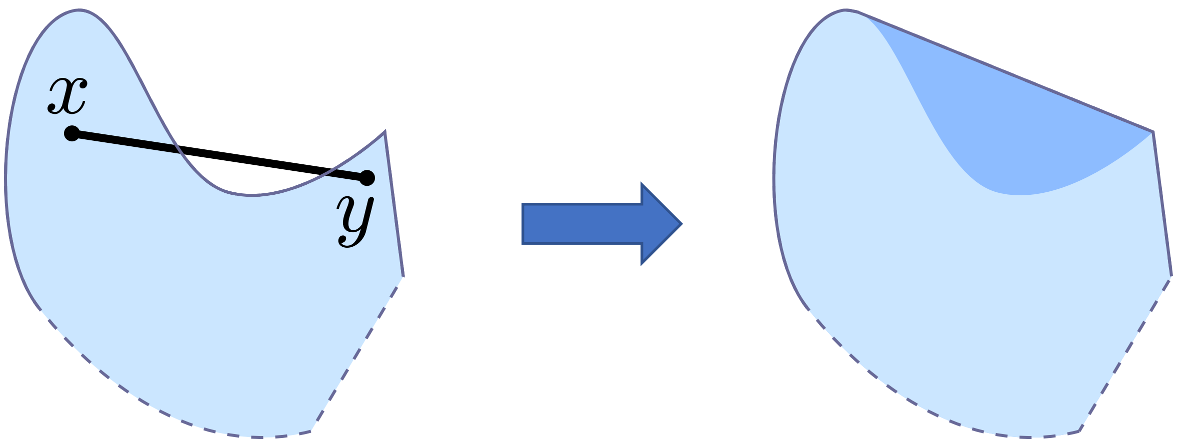
convex set



nonconvex set

Convex Set

- Convex hull $\mathrm{conv}(S)$
  - The union of all line segments $[x, y]$ for all $x, y \in S$
  - The smallest convex set containing S



nonconvex set　　　　convex hull

Examples of Convex Sets

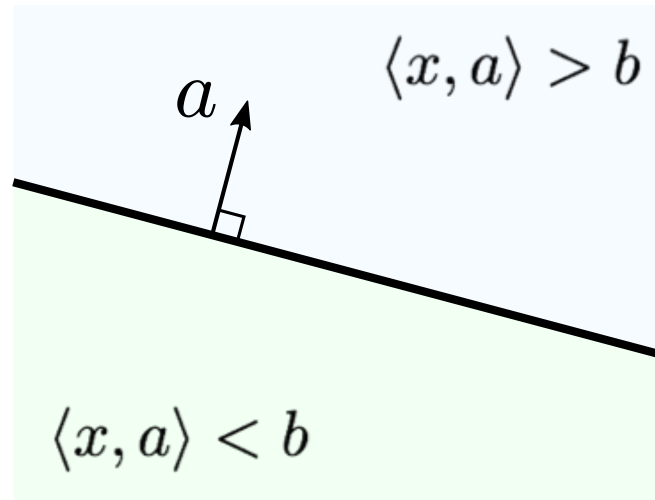
Examples of Convex Sets

- Hyperplanes

$$\{x \in \mathbb{R}^n : \langle x, a \rangle = b\} \quad a \neq 0$$

- Halfspaces

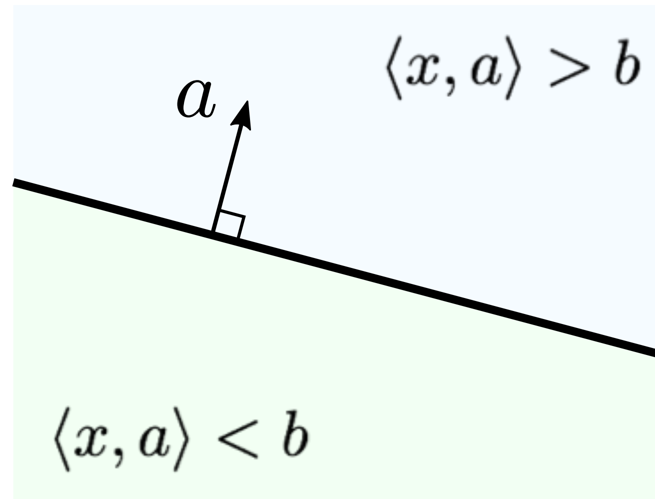
$$\{x \in \mathbb{R}^n : \langle x, a \rangle \leq b\} \quad a \neq 0$$



Examples of Convex Sets

- Hyperplanes

$$\{x \in \mathbb{R}^n : \langle x, a \rangle = b\} \quad a \neq 0$$



- Halfspaces

$$\{x \in \mathbb{R}^n : \langle x, a \rangle \leq b\} \quad a \neq 0$$

- Open and closed balls

$$B_r(x) := \{y : \|y - x\| < r\}, \quad r > 0$$

$$\overline{B}_r(x) := \{y : \|y - x\| \leq r\}, \quad r \geq 0$$

Separating Hyperplanes

Separating Hyperplanes

$S_1, S_2 \subseteq \mathbb{R}^n$ convex

Separating Hyperplanes

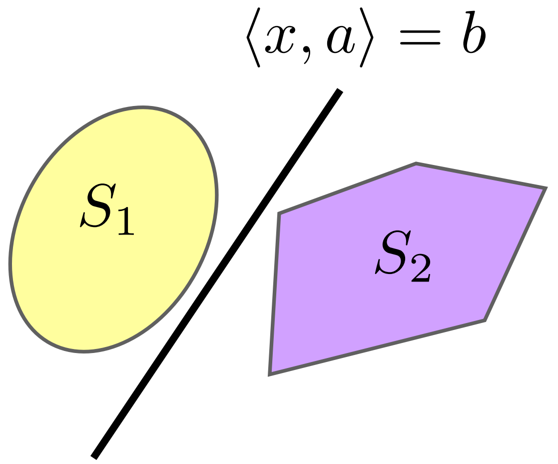
$S_1, S_2 \subseteq \mathbb{R}^n$ convex

- Separating Hyperplane Theorem

  If $\operatorname{int} S_1 \neq \varnothing$ and $S_2 \cap \operatorname{int} S_1 = \varnothing$, then there exists a nonzero $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$ such that

$$\langle x_1, a \rangle \leq b \leq \langle x_2, a \rangle$$

  for any $x_1 \in S_1$ and $x_2 \in S_2$



Separating Hyperplanes

$S_1, S_2 \subseteq \mathbb{R}^n$ convex

- Separating Hyperplane Theorem

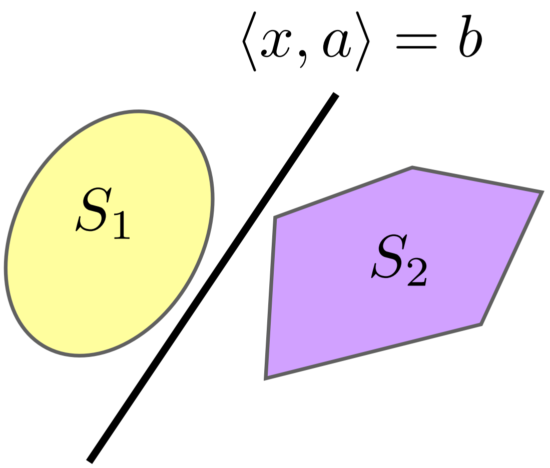
  If $\operatorname{int} S_1 \neq \varnothing$ and $S_2 \cap \operatorname{int} S_1 = \varnothing$, then there exists a nonzero $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$ such that

$$\langle x_1, a \rangle \leq b \leq \langle x_2, a \rangle$$

  for any $x_1 \in S_1$ and $x_2 \in S_2$

  In other words, $S_1 \subseteq \{x : \langle x, a \rangle \leq b\}$

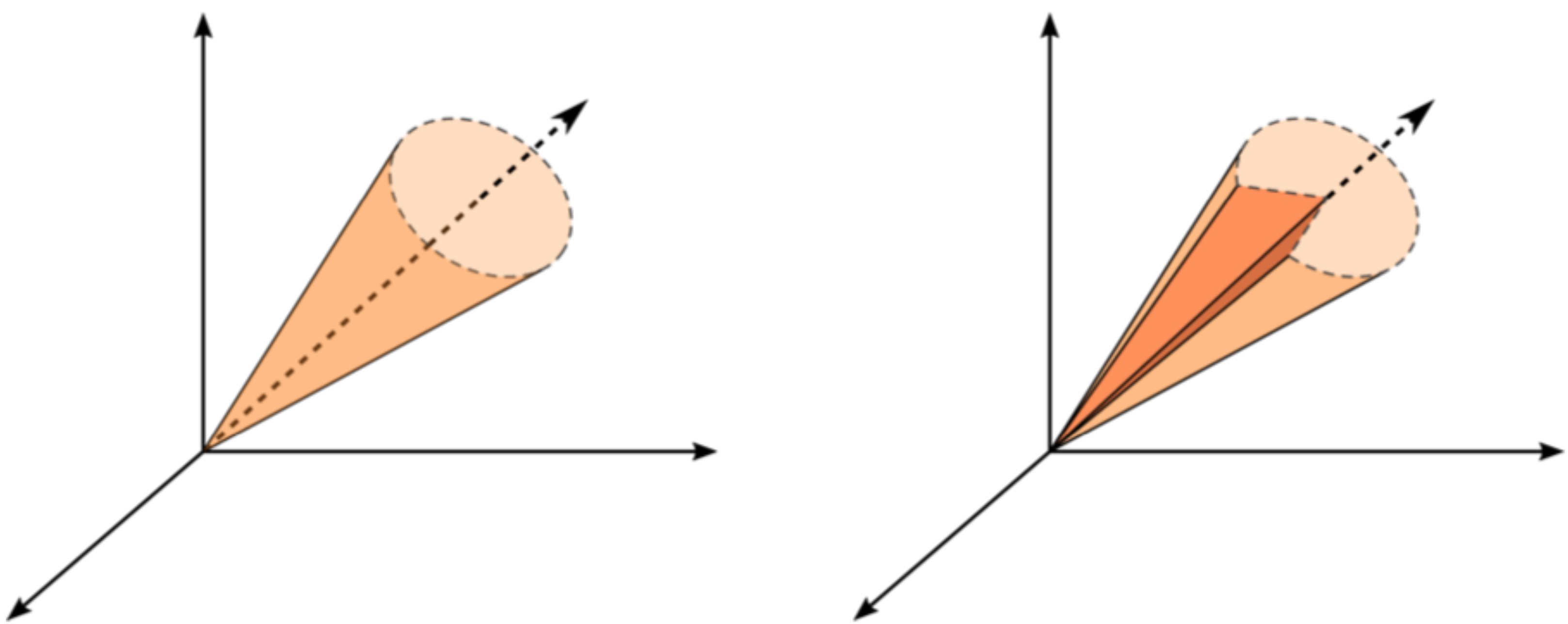
$$S_2 \subseteq \{x : \langle x, a \rangle \geq b\}$$



Convex Cones

Convex Cones

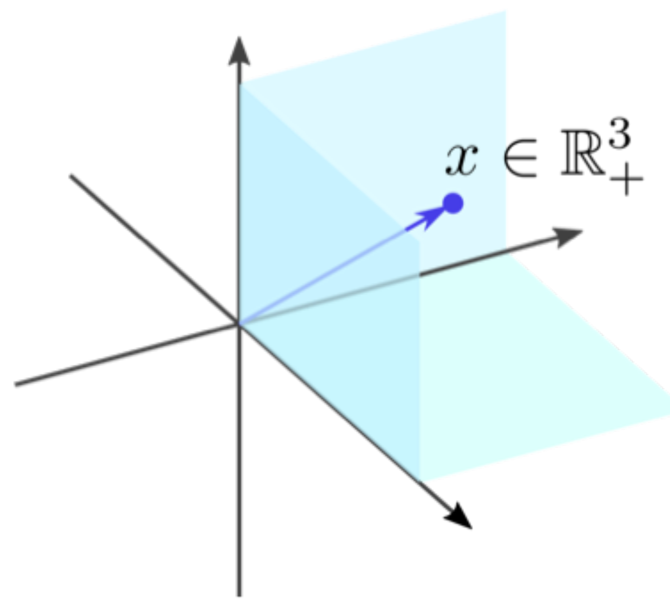
- $S \subseteq \mathbb{R}^n$ is called a **cone** if for any $x \in S$ and any positive scalar α, ones has $\alpha x \in S$



Convex Cones

- $S \subseteq \mathbb{R}^n$ is called a **cone** if for any $x \in S$ and any positive scalar α, ones has $\alpha x \in S$
- Examples of convex cones:
  - Nonnegative orthant

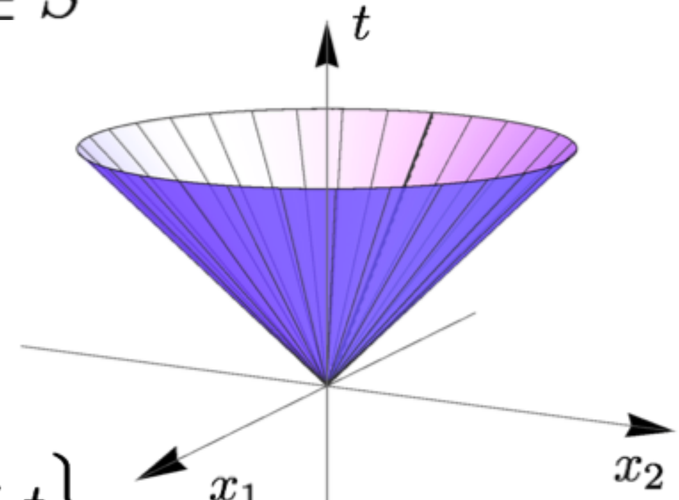
    $\mathbb{R}^n_+ := \{x \in \mathbb{R}^n : x_i \geq 0 \ \forall i\}$



Convex Cones

- $S \subseteq \mathbb{R}^n$ is called a **cone** if for any $x \in S$ and any positive scalar α , ones has $\alpha x \in S$
- Examples of convex cones:
  - Nonnegative orthant $\mathbb{R}^n_+$
  - Second-order cone

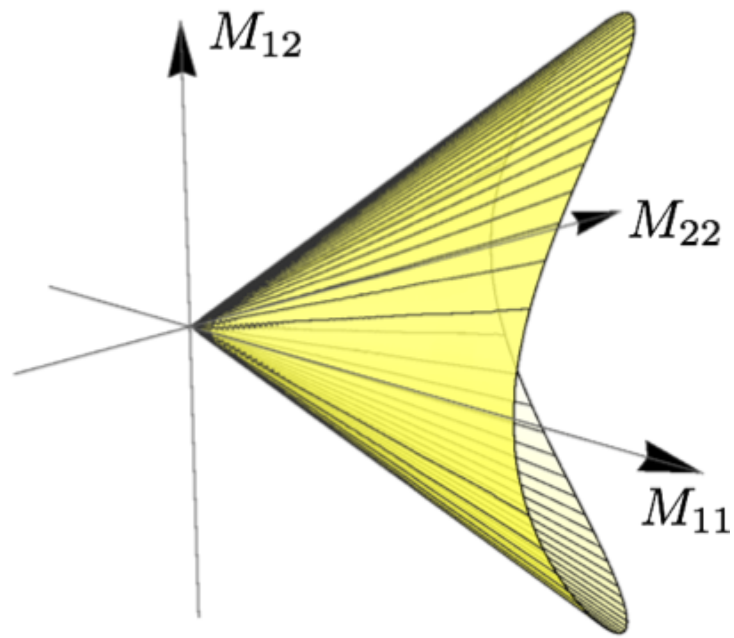
$$L^n := \left\{(x,t) \in \mathbb{R}^{n+1} : \sqrt{x^T x} \le t\right\}$$



Convex Cones

- $S \subseteq \mathbb{R}^n$ is called a **cone** if for any $x \in S$ and any positive scalar α, ones has $\alpha x \in S$
- Examples of convex cones:
  - Nonnegative orthant $\mathbb{R}^n_+$
  - Second-order cone L^n
  - PSD cone

    $\mathbb{S}^n_+ = \{M \in \mathbb{S}^n : M \succeq 0\}$



Convex Cones

- Nonnegative orthant $\mathbb{R}^n_+$
- Second-order cone L^n
- PSD cone $\mathbb{S}^n_+$

Convex Cones

- Nonnegative orthant $\mathbb{R}^n_+$
- Second-order cone L^n
- PSD cone $\mathbb{S}^n_+$

1. Closed
2. Non-empty interior
3. Pointed
   $x \in K$ and $x \neq 0 \Rightarrow -x \notin K$
4. Self-dual

Convex Cones

- Nonnegative orthant $\mathbb{R}^n_+$
- Second-order cone L^n
- PSD cone $\mathbb{S}^n_+$

1. Closed
2. Non-empty interior
3. Pointed

   $x \in K$ and $x \neq 0 \Rightarrow -x \notin K$
4. Self-dual

1+2+3 $\Rightarrow$ Possible to define a partial order:

We say $x \succeq_K y$ if $x - y \in K$

partial order
$$\begin{cases} x \succeq_K x \quad \forall x \\ x \succeq_K y \text{ and } y \succeq_K x \quad \Rightarrow \quad x = y \\ x \succeq_K y \text{ and } y \succeq_K z \quad \Rightarrow \quad x \succeq_K z \end{cases}$$

Convex Cones

- Nonnegative orthant $\mathbb{R}^n_+$
- Second-order cone L^n
- PSD cone $\mathbb{S}^n_+$

1. Closed
2. Non-empty interior
3. Pointed

   $x \in K$ and $x \neq 0 \Rightarrow -x \notin K$
4. Self-dual

1+2+3 $\Rightarrow$ Possible to define a partial order:

We say $x \succeq_K y$ if $x - y \in K$

Self-dual: $x \succeq_K 0 \Longleftrightarrow \langle x, y \rangle \geq 0 \; \forall y \in K$

Operations that Preserve Convexity

Operations that Preserve Convexity

- Intersection

$\bigcap$ convex sets is convex

Operations that Preserve Convexity

- Intersection

  $\bigcap$ convex sets is convex

- Affine transformation
  - $\{Ax + b : x \in S\}$ is convex if S is convex
  - $\{x : Ax + b \in S\}$ is convex if S is convex

Operations that Preserve Convexity

- Intersection

  $\bigcap$ convex sets is convex

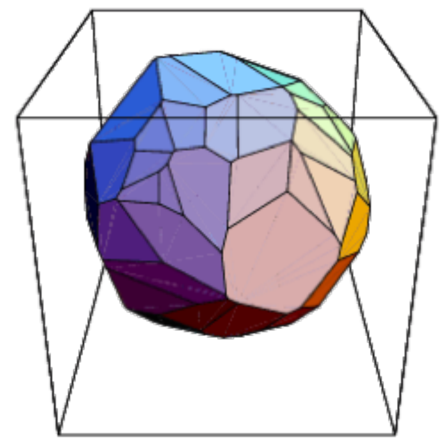
- Affine transformation
  - $\{Ax + b : x \in S\}$ is convex if S is convex
  - $\{x : Ax + b \in S\}$ is convex if S is convex
- Cartesian product, Minkowski sum, etc.

More Examples of Convex Sets

More Examples of Convex Sets

- Convex polytopes

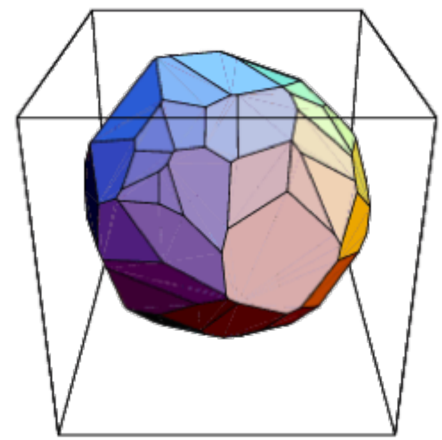
$$\{x \in \mathbb{R}^n : Ax \leq b\}$$
$$= \{x \in \mathbb{R}^n : b - Ax \in \mathbb{R}^n_+\}$$
$$= \bigcap_i \{x \in \mathbb{R}^n : a_i^T x \leq b_i\}$$



More Examples of Convex Sets

- Convex polytopes

$$\begin{aligned} &\{x \in \mathbb{R}^n : Ax \le b\} \\ = &\{x \in \mathbb{R}^n : b - Ax \in \mathbb{R}^n_+\} \\ = &\bigcap_i \{x \in \mathbb{R}^n : a_i^T x \le b_i\} \end{aligned}$$



- Solutions of linear matrix inequalities

$$\{x \in \mathbb{R}^n : x_1 A_1 + \cdots + x_n A_n + B \succeq 0\}$$

Outline

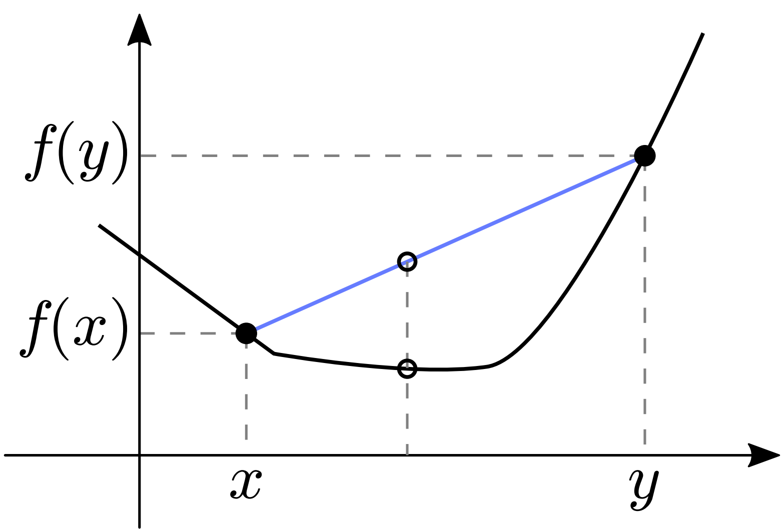
- Motivation
- Recap of Linear Algebra and Real Analysis
- Convex Set
- **Convex Function**
- Convex Program

Convex Function

- A function $f : X \to \mathbb{R}$ with a convex domain X is called **convex** if

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y)$$

for all $x, y \in X$ and $\alpha \in [0,1]$



Convex Function

- A function $f : X \to \mathbb{R}$ with a convex domain X is called **convex** if

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y)$$

for all $x, y \in X$ and $\alpha \in [0,1]$

- … is called **strictly convex** if

$$f(\alpha x + (1-\alpha)y) < \alpha f(x) + (1-\alpha) f(y)$$

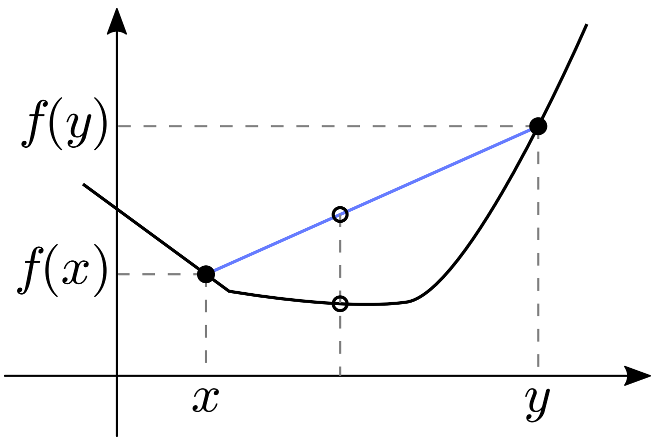
for all $x, y \in X$ with $x \neq y$ and $\alpha \in (0,1)$

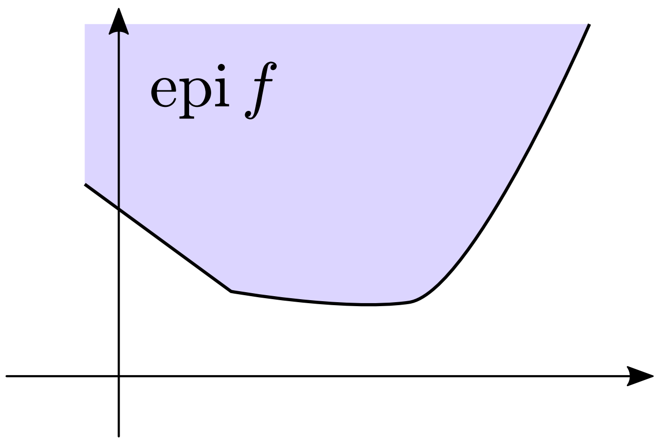
Convex Function

- A function $f : X \to \mathbb{R}$ with a convex domain X is called **convex** if

$$\operatorname{epi} f := \{(x, t) \in X \times \mathbb{R} : f(x) \leq t\}$$

is convex.





Sublevel Set

Sublevel Set

$f : X \to \mathbb{R}$

Sublevel Set

$f : X \to \mathbb{R}$

- $C_\alpha = \{x \in X : f(x) \le \alpha\}$

Sublevel Set

$f : X \to \mathbb{R}$

- $C_\alpha = \{x \in X : f(x) \leq \alpha\}$
- Sublevel sets are convex if f is convex.

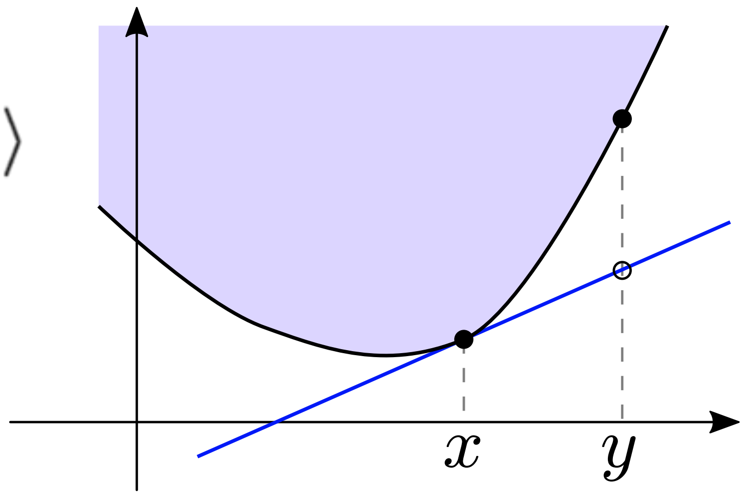
First-order Condition

First-order Condition

- A differentiable function $f : X \to \mathbb{R}$ with a convex domain X is convex iff

$$f(y) \geq f(x) + \langle y - x, \nabla f(x) \rangle$$

for all $x, y \in X$

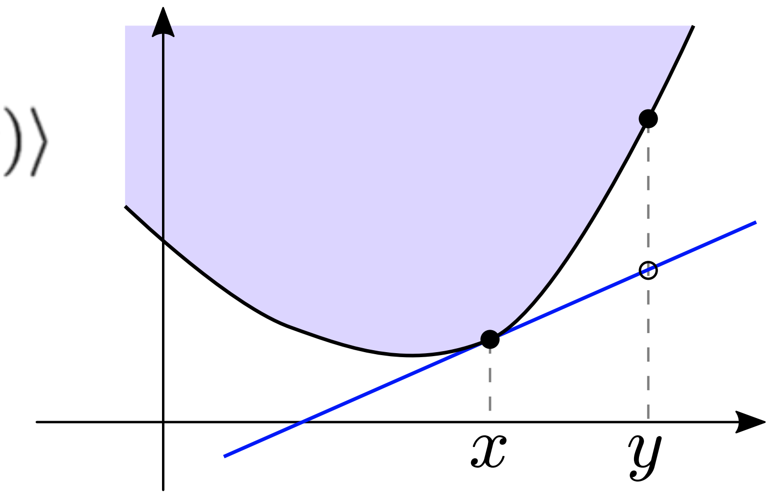


First-order Condition

- A differentiable function $f : X \to \mathbb{R}$ with a convex domain X is convex iff

$$f(y) \geq f(x) + \langle y - x, \nabla f(x) \rangle$$

for all $x, y \in X$



- The tangent plane

$$\left\{ (y, z) : \begin{bmatrix} \nabla f(x) \\ -1 \end{bmatrix}^T \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} \nabla f(x) \\ -1 \end{bmatrix}^T \begin{bmatrix} x \\ f(x) \end{bmatrix} \right\}$$

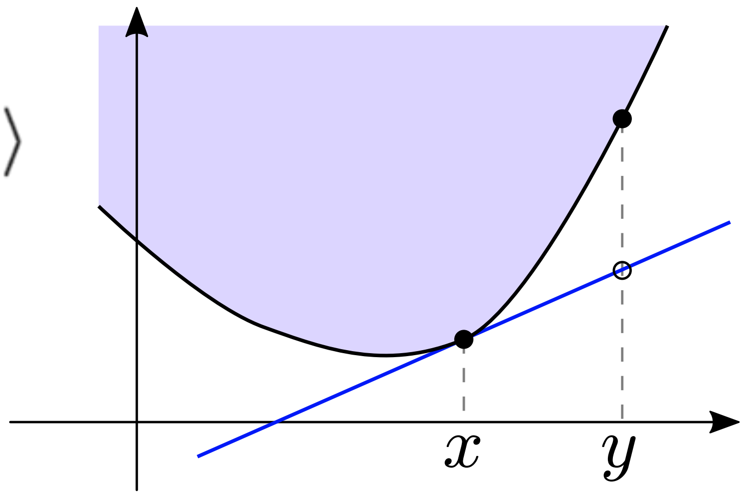
is the hyperplane separating $\mathrm{epi}\, f$ and $\{(x, f(x))\}$

First-order Condition

- A differentiable function $f : X \to \mathbb{R}$ with a convex domain X is convex iff

$$f(y) \geq f(x) + \langle y - x, \nabla f(x) \rangle$$

for all $x, y \in X$



- … is strictly convex iff

$$f(y) > f(x) + \langle y - x, \nabla f(x) \rangle$$

for all $x, y \in X$ with $x \neq y$

Second-order Condition

Second-order Condition

- A twice differentiable function $f : X \to \mathbb{R}$ with a convex domain X is convex iff

$$H_f(x) \succeq 0$$

for all $x \in X$

Operations that Preserve Convexity

- Positive weighted sum

$g(x) = \sum_k \alpha_k f_k(x)$ $\qquad$ $\alpha_k \geq 0,\ f_k$ convex $\forall k$

- Pointwise supremum of a family of convex functions

$g(x) = \max_k f_k(x)$ $\qquad$ f_k convex $\forall k$

$g(x) = \sup_y f(x, y)$ $\qquad$ $f(\cdot, y)$ convex $\forall y$

- Composition with affine functions

$g(x) = f(Ax + b)$ $\qquad$ f convex

Examples of Convex Functions

- $f(x) = Ax + b \quad x \in \mathbb{R}^n$
- $f(x) = \frac{1}{2}x^T M x + p^T x + q \quad x \in \mathbb{R}^n, \; M \succeq 0$
- $f(x) = e^x \quad x \in \mathbb{R}$
- $f(x) = -\log x \quad x > 0$

Outline

- Motivation
- Recap of Linear Algebra and Real Analysis
- Convex Set
- Convex Function
- **Convex Program**

Convex Program

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ \text{s.t.} \quad & x \in X \end{aligned}$$

- Convex program: f is convex and X is convex.

Convex Program

$$\min_{x \in \mathbb{R}^n} \quad f(x)$$
$$\text{s.t.} \quad x \in X$$

- Convex program: f is convex and X is convex.
- Any local optimum is a global optimum:

Suppose $x^* \in X$ and there exists $\delta > 0$ such that $f(x^*) \leq f(x)$ for all $x \in X$ with $\|x - x^*\| < \delta$. Then

$$f(x^*) \leq f(x) \qquad \text{for all } x \in X$$

Conic Program

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^T x \\ \text{s.t.} \quad & Ax \succeq_K b \\ & Fx = h \end{aligned}$$

Conic Program

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^T x \\ \text{s.t.} \quad & Ax \succeq_K b \\ & Fx = h \end{aligned}$$

$$\begin{aligned} & x \succeq_K y \\ \Leftrightarrow \quad & x - y \in K \end{aligned}$$

Conic Program

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^T x \\ \text{s.t.} \quad & Ax \succeq_K b \\ & Fx = h \end{aligned}$$

$$\begin{aligned} & x \succeq_K y \\ \Leftrightarrow \quad & x - y \in K \end{aligned}$$

- $K = \mathbb{R}^n_+$: Linear program (LP)
- $K = \prod_i L^{n_i}$: Second-order cone program (SOCP)
- $K = \mathbb{S}^n_+$: Semidefinite program (SDP)

Duality Theory

Duality Theory

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^T x \\ \text{s.t.} \quad & Ax \succeq_K b \\ & Fx = h \end{aligned}$$

Duality Theory

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^T x \\ \text{s.t.} \quad & Ax \succeq_K b \\ & Fx = h \end{aligned}$$

- For a conic program with K being self-dual, the Lagrangian is

$$L(x, \lambda, \mu) = c^T x - \langle \lambda, Ax - b \rangle - \langle \mu, Fx - h \rangle$$

where $\lambda \in K$

Duality Theory

$$\min_{x \in \mathbb{R}^n} \quad c^T x$$
$$\text{s.t.} \quad Ax \succeq_K b$$
$$Fx = h$$

- For a conic program with K being self-dual, the Lagrangian is

$$L(x, \lambda, \mu) = c^T x - \langle \lambda, Ax - b \rangle - \langle \mu, Fx - h \rangle$$

  where $\lambda \in K$

- Because K is self-dual, we have

$$L(x, \lambda, \mu) \leq c^T x \qquad \forall x \in X, \forall \lambda \in K, \forall \mu$$

Duality Theory

$$\min_{x \in \mathbb{R}^n} \quad c^T x$$
$$\text{s.t.} \quad Ax \succeq_K b$$
$$Fx = h$$

- For a conic program with K being self-dual, the Lagrangian is

$$L(x, \lambda, \mu) = c^T x - \langle \lambda, Ax - b \rangle - \langle \mu, Fx - h \rangle$$

where $\lambda \in K$

- Because K is self-dual, we have

$$L(x, \lambda, \mu) \leq c^T x \qquad \forall x \in X, \forall \lambda \in K, \forall \mu$$

$$\inf_{x \in X} L(x, \lambda, \mu) \leq \inf_{x \in X} c^T x \qquad \forall \lambda \in K, \forall \mu$$

Duality Theory

$$\min_{x \in \mathbb{R}^n} \quad c^T x$$
$$\text{s.t.} \quad Ax \succeq_K b$$
$$Fx = h$$

- For a conic program with K being self-dual, the Lagrangian is

$$L(x, \lambda, \mu) = c^T x - \langle \lambda, Ax - b \rangle - \langle \mu, Fx - h \rangle$$

where $\lambda \in K$

- Because K is self-dual, we have

$$L(x, \lambda, \mu) \leq c^T x \qquad \forall x \in X, \forall \lambda \in K, \forall \mu$$

$$\inf_{x \in X} L(x, \lambda, \mu) \leq \inf_{x \in X} c^T x \qquad \forall \lambda \in K, \forall \mu$$

$$\inf_{x \in \mathbb{R}^n} L(x, \lambda, \mu) \leq \inf_{x \in X} L(x, \lambda, \mu) \qquad \forall \lambda \in K, \forall \mu$$

Duality Theory

$$\min_{x\in\mathbb{R}^n} \quad c^T x$$
$$\text{s.t.} \quad Ax \succeq_K b$$
$$Fx = h$$

- For a conic program with K being self-dual, the Lagrangian is

$$L(x, \lambda, \mu) = c^T x - \langle \lambda, Ax - b \rangle - \langle \mu, Fx - h \rangle$$

where $\lambda \in K$

- Because K is self-dual, we have

$$\inf_{x\in\mathbb{R}^n} L(x, \lambda, \mu) \leq \inf_{x \in X} c^T x \qquad \forall \lambda \in K, \forall \mu$$

Duality Theory

$$\min_{x \in \mathbb{R}^n} \quad c^T x$$
$$\text{s.t.} \quad Ax \succeq_K b$$
$$Fx = h$$

- For a conic program with K being self-dual, the Lagrangian is

$$L(x, \lambda, \mu) = c^T x - \langle \lambda, Ax - b \rangle - \langle \mu, Fx - h \rangle$$

where $\lambda \in K$

- Because K is self-dual, we have

$$\inf_{x \in \mathbb{R}^n} L(x, \lambda, \mu) \leq \inf_{x \in X} c^T x \qquad \forall \lambda \in K, \forall \mu$$

$$\sup_{\lambda \in K, \mu} \inf_{x \in \mathbb{R}^n} L(x, \lambda, \mu) \leq \inf_{x \in X} c^T x$$

Duality Theory

$$\min_{x \in \mathbb{R}^n} \quad c^T x$$
$$\text{s.t.} \quad Ax \succeq_K b$$
$$Fx = h$$

- For a conic program with K being self-dual, the Lagrangian is

$$L(x, \lambda, \mu) = c^T x - \langle \lambda, Ax - b \rangle - \langle \mu, Fx - h \rangle$$

where $\lambda \in K$

- Weak duality:

$$\begin{aligned} \max_{\lambda, \mu} \quad & \inf_{x \in \mathbb{R}^n} L(x, \lambda, \mu) \\ \text{s.t.} \quad & \lambda \succeq_K 0 \end{aligned} \quad \leq \quad \begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^T x \\ \text{s.t.} \quad & Ax \succeq_K b \\ & Fx = h \end{aligned}$$

Duality Theory

- Weak duality:

$$\begin{aligned} \max_{\lambda,\mu} \quad & \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu) \\ \text{s.t.} \quad & \lambda \succeq_K 0 \end{aligned} \quad \leq \quad \begin{aligned} \min_{x\in\mathbb{R}^n} \quad & c^T x \\ \text{s.t.} \quad & Ax \succeq_K b \\ & Fx = h \end{aligned}$$

Duality Theory

- Strong duality:

$$\begin{aligned} \max_{\lambda,\mu} \quad & \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu) \\ \text{s.t.} \quad & \lambda \succeq_K 0 \end{aligned} \quad = \quad \begin{aligned} \min_{x\in\mathbb{R}^n} \quad & c^T x \\ \text{s.t.} \quad & Ax \succeq_K b \\ & Fx = h \end{aligned}$$

Duality Theory

- Strong duality:

$$\begin{array}{ll} \max_{\lambda,\mu} & \inf_{x\in\mathbb{R}^n} L(x,\lambda,\mu) \\ \text{s.t.} & \lambda \succeq_K 0 \end{array} \quad = \quad \begin{array}{ll} \min_{x\in\mathbb{R}^n} & c^T x \\ \text{s.t.} & Ax \succeq_K b \\ & Fx = h \end{array}$$

- For a convex conic program with K being self-dual, strong duality holds under Slater's condition

$$\exists x_0 \text{ s.t. } Ax_0 - b \in \operatorname{int} K \text{ and } Fx_0 = h$$

Duality Theory

- Strong duality:

$$\begin{aligned} \max_{\lambda,\mu} \quad & \langle \lambda, b \rangle + \langle \mu, h \rangle \\ \text{s.t.} \quad & c - A^T\lambda - F^T\mu = 0 \\ & \lambda \succeq_K 0 \end{aligned} \quad = \quad \begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^T x \\ \text{s.t.} \quad & Ax \succeq_K b \\ & Fx = h \end{aligned}$$

- For a convex conic program with K being self-dual, strong duality holds under Slater's condition

$$\exists x_0 \text{ s.t. } Ax_0 - b \in \operatorname{int} K \text{ and } Fx_0 = h$$

Duality Theory

- Strong duality:

$$\begin{aligned} \max_{\lambda,\mu} \quad & \langle \lambda, b \rangle + \langle \mu, h \rangle \\ \text{s.t.} \quad & c - A^T\lambda - F^T\mu = 0 \\ & \lambda \succeq_K 0 \end{aligned} \quad = \quad \begin{aligned} \min_{x\in\mathbb{R}^n} \quad & c^T x \\ \text{s.t.} \quad & Ax \succeq_K b \\ & Fx = h \end{aligned}$$

- Primal feasible points produce upper bounds of optimal value.
  Dual feasible points produce lower bounds of optimal value.

Duality Theory

- Strong duality:

$$\begin{array}{lllll} \max\limits_{\lambda,\mu} & \langle \lambda, b \rangle + \langle \mu, h \rangle & & \min\limits_{x\in\mathbb{R}^n} & c^T x \\ \text{s.t.} & c - A^T\lambda - F^T\mu = 0 & = & \text{s.t.} & Ax \succeq_K b \\ & \lambda \succeq_K 0 & & & Fx = h \end{array}$$

- Certificate of optimality: If $\hat{x}$ is primal feasible and $(\hat{\lambda}, \hat{\mu})$ is dual feasible, and

$$c^T\hat{x} = \langle \hat{\lambda}, b \rangle + \langle \hat{\mu}, h \rangle$$

then $\hat{x}$ is an optimal solution.

KKT Conditions

KKT Conditions

- $$Ax \succeq_K b$$
  $$Fx = h$$
  primal feasibility

  $$c - A^T\lambda - F^T\mu = 0$$
  $$\lambda \succeq_K 0$$
  dual feasibility

  $$\langle \lambda, Ax - b \rangle = 0$$
  complementary slackness

KKT Conditions

- $$Ax \succeq_K b$$
  $$Fx = h$$
  primal feasibility

  $$c - A^T\lambda - F^T\mu = 0$$
  $$\lambda \succeq_K 0$$
  dual feasibility

  $$\langle \lambda, Ax - b \rangle = 0$$
  complementary slackness

- Necessary and sufficient conditions for optimality (under Slater's condition)

Linear Program

Linear Program

- $K = \mathbb{R}^m_+$

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^T x \\ \text{s.t.} \quad & Ax \geq b \\ & Fx = h \end{aligned} \qquad\qquad \begin{aligned} \max_{\lambda, \mu} \quad & b^T \lambda + h^T \mu \\ \text{s.t.} \quad & c - A^T \lambda - F^T \mu = 0 \\ & \lambda \geq 0 \end{aligned}$$

Linear Program

- $K = \mathbb{R}^m_+$

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^T x \\ \text{s.t.} \quad & Ax \geq b \\ & Fx = h \end{aligned} \qquad \begin{aligned} \max_{\lambda,\mu} \quad & b^T \lambda + h^T \mu \\ \text{s.t.} \quad & c - A^T \lambda - F^T \mu = 0 \\ & \lambda \geq 0 \end{aligned}$$

- Scheduling for EV charging

Semidefinite Program

Semidefinite Program

- $K = \mathbb{S}^m_+$

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & c^T x \\ \text{s.t.} \quad & \textstyle\sum_{i=1}^n x_i A_i \succeq B \\ & Fx = h \end{aligned}$$

$$\begin{aligned} \max_{\Lambda \in \mathbb{S}^m, \mu} \quad & \mathrm{tr}(B\Lambda) + h^T \mu \\ \text{s.t.} \quad & \mathrm{tr}(A_i \Lambda) = (c - F^T \mu)_i \quad \forall i \\ & \Lambda \succeq 0 \end{aligned}$$

QCQP and SDP Relaxation

QCQP and SDP Relaxation

- QCQP: quadratically constrained quadratic program

$$
\begin{aligned}
\min_{x \in \mathbb{R}^n} \quad & x^T P_0 x \\
\text{s.t.} \quad & x^T P_i x \leq 0, \quad i = 1, \ldots, m
\end{aligned}
$$

QCQP and SDP Relaxation

- QCQP: quadratically constrained quadratic program

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & x^T P_0 x \\ \text{s.t.} \quad & x^T P_i x \leq 0, \quad i = 1, \ldots, m \end{aligned}$$

- Could be non-convex if some P_i is not PSD.
- Generally NP-hard.

QCQP and SDP Relaxation

- QCQP: quadratically constrained quadratic program

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & x^T P_0 x \\ \text{s.t.} \quad & x^T P_i x \leq 0, \quad i = 1, \ldots, m \end{aligned}$$

$$x^T P x = \operatorname{tr}(x^T P x) = \operatorname{tr}(P x x^T)$$

QCQP and SDP Relaxation

- QCQP: quadratically constrained quadratic program

$$
\begin{aligned}
\min_{W \in \mathbb{S}^n, x \in \mathbb{R}^n} \quad & \operatorname{tr}(P_0 W) \\
\text{s.t.} \quad & \operatorname{tr}(P_i W) \leq 0, \quad i = 1, \ldots, m \\
& W = xx^T
\end{aligned}
$$

$$x^T P x = \operatorname{tr}(x^T P x) = \operatorname{tr}(P x x^T)$$

QCQP and SDP Relaxation

- QCQP: quadratically constrained quadratic program

$$
\begin{aligned}
\min_{W \in \mathbb{S}^n} \quad & \operatorname{tr}(P_0 W) \\
\text{s.t.} \quad & \operatorname{tr}(P_i W) \leq 0, \quad i = 1, \ldots, m \\
& W \succeq 0 \\
& \operatorname{rank} W \leq 1
\end{aligned}
$$

QCQP and SDP Relaxation

- QCQP: quadratically constrained quadratic program

$$
\begin{aligned}
\min_{W \in \mathbb{S}^n} \quad & \operatorname{tr}(P_0 W) \\
\text{s.t.} \quad & \operatorname{tr}(P_i W) \leq 0, \quad i = 1, \ldots, m \\
& W \succeq 0
\end{aligned}
$$

- Semidefinite relaxation of QCQP

Algorithms

Algorithms

- Unconstrained optimization ($X = \mathbb{R}^n$)
  - Gradient descent & its variants
  - Newton & quasi-Newton method

Algorithms

- Unconstrained optimization ($X = \mathbb{R}^n$)
  - Gradient descent & its variants
  - Newton & quasi-Newton method
- Constrained optimization
  - Projected gradient descent & its variants
  - Dual ascent & its variants
  - Simplex method for LP
  - Interior point method
  - Distributed algorithms

Software

- Solvers
  - SDPT3, Sedumi (LP+SOCP+SDP, MATLAB), CVXOPT (Python)
  - IPOPT (nonlinear opt, local solution)
  - Gurobi (LP+SOCP+…), Mosek (LP+SOCP+SDP+…)
- Interfaces and modelling tools
  - CVX, YALMIP (MATLAB)
  - CVXPY (Python)

References

- Linear algebra:
  - G. Strang. *Linear Algebra and Its Applications*.
  - P. D. Lax. *Linear Algebra and Its Applications*.
- Real analysis:
  - W. Rudin. *Principles of Mathematical Analysis*.
  - N. L. Carothers. *Real Analysis*.
- Convex Optimization:
  - S. Boyd and L. Vandenberghe. *Convex Optimization*.
  - A. Ben-Tal and A. Nemirovski. *Lectures on Modern Convex Optimization*.
- Numerical Methods:
  - J. Nocedal and S. J. Wright. *Numerical Optimization*.

Backup Slides

Projection onto Closed Convex Sets

$S \subseteq \mathbb{R}^n$ closed

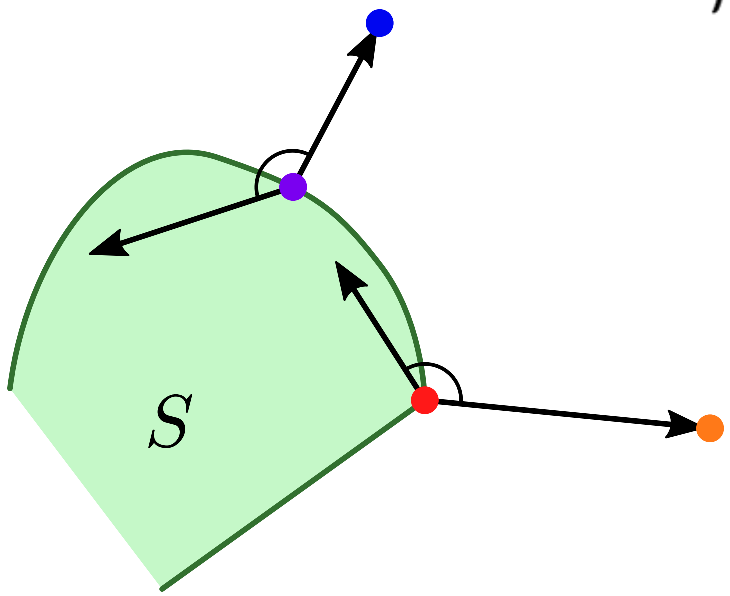
- Projection onto a closed set $\mathcal{P}_S(x) = \arg\min_{y \in S} \|y - x\|$
  - Points in S that are closest to x
- Projection onto a closed convex set
  - (Motzkin) S is convex $\iff \mathcal{P}_S(x)$ is unique for all $x \in \mathbb{R}^n$
  - If S is convex, then

$$y = \mathcal{P}_S(x) \iff \langle x - y, z - y \rangle \leq 0 \text{ for all } z \in S$$

Projection onto Closed Convex Sets

$S \subseteq \mathbb{R}^n$ closed and convex

- Projection onto a closed convex set



$$\mathcal{P}_S(x) = \arg\min_{y \in S} \|y - x\|$$

$$y = \mathcal{P}_S(x)$$
$$\text{iff } \langle x - y, z - y \rangle \leq 0 \quad \forall z \in S$$