COMP6704 Lecture 2

Advanced Topics in Optimization

Convex Sets

Fall, 2022

Instructor: WU, Xiao-Ming

Convexity:
A Crucial Matter

- A convex function (top) is one whose graph slopes everywhere toward its minimum value.

- A nonconvex function (bottom) may have many basins, or local minima.

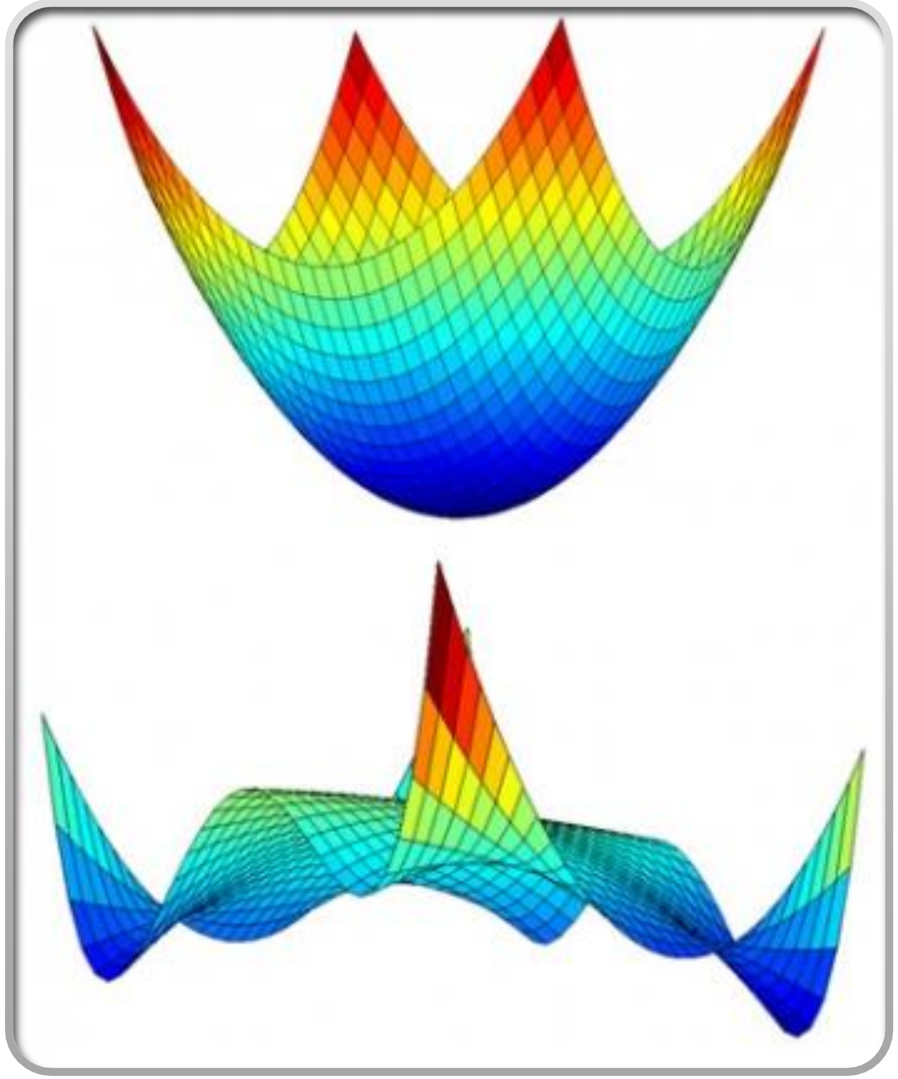


Image by Amir Ali Ahmadi

Convex Optimization: A Brief History

- The field is about 50 years old.
- Starts with the work of Von Neumann, Kuhn and Tucker, etc.
- Explodes in the 60’s with the advent of “relatively” cheap and efficient computers…
- Key to all this: fast linear algebra
- Some of the theory developed before computers even existed…

Convex Optimization: History

- **Convexity $\Longrightarrow$ Low complexity:**

"... In fact, the great watershed in optimization isn't between linearity and nonlinearity, but convexity and nonconvexity."

T. Rockafellar.

- True: Nemirovskii and Yudin [1979].
- **Very true:** Karmarkar [1984].
- **Seriously true:** Convex programming, Nesterov and Nemirovskii [1994].

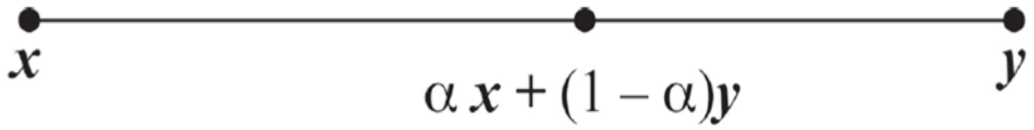
Standard Convex Complexity Analysis

- All convex minimization problems with a first order oracle (returning $f(x)$ and a sub-gradient) can be solved in polynomial time in size and number of precision digits.
- Proved using the **ellipsoid method** by Nemirovskii and Yudin [1979].
- Very slow convergence in practice.

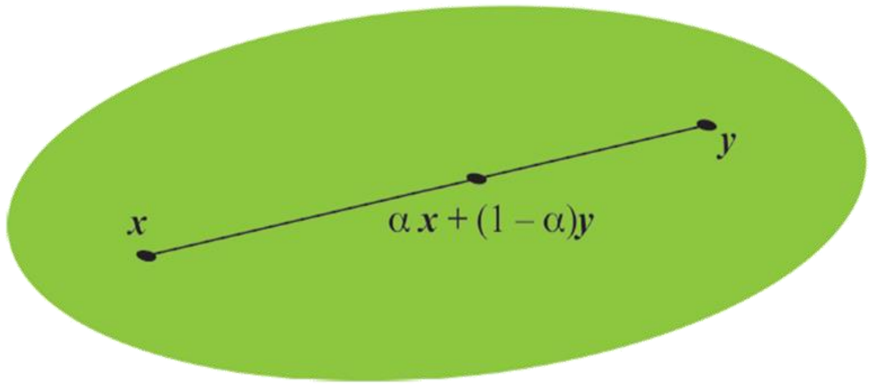
Convex Set

- A **line segment** defined by vectors x and y is the set of points of the form

$$\alpha x + (1-\alpha)y, \text{ for } \alpha \in [0,1].$$

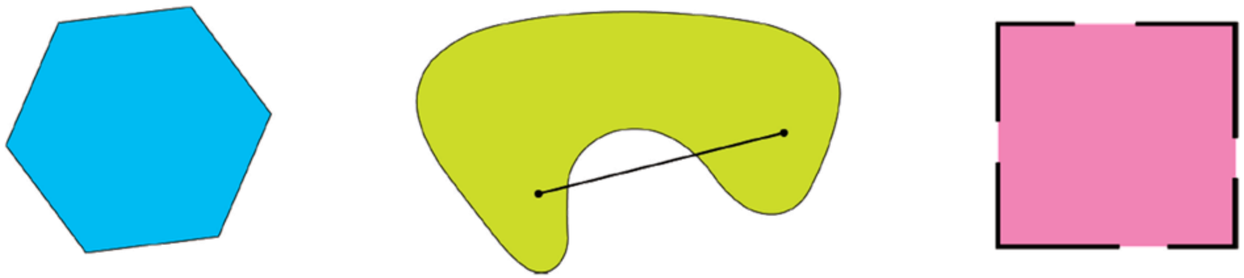


- A set $C \subset \mathbb{R}^n$ is **convex** when, with any two vectors x and y that belong to the set C, the line segment connecting x and y also belongs to C.



Examples

- Which of the following sets are convex?
  - The space $\mathbb{R}^n$
  - A **line** through two given vectors x and y
    $$l(x,y) = \{z \mid z = x + t(y-x), t \in \mathbb{R}\}$$
  - A **ray** defined by a vector x
    $$\{z \mid z = \lambda x, \lambda \geq 0\}$$
  - The **positive orthant** $\{x \in \mathbb{R}^n | x \succeq 0\}$, ($\succeq$ componentwise inequality)
  - The **set** $\{x \in \mathbb{R}^2 \mid x_1 > 0, x_2 \geq 0\}$
  - The **set** $\{x \in \mathbb{R}^2 \mid x_1 x_2 = 0\}$



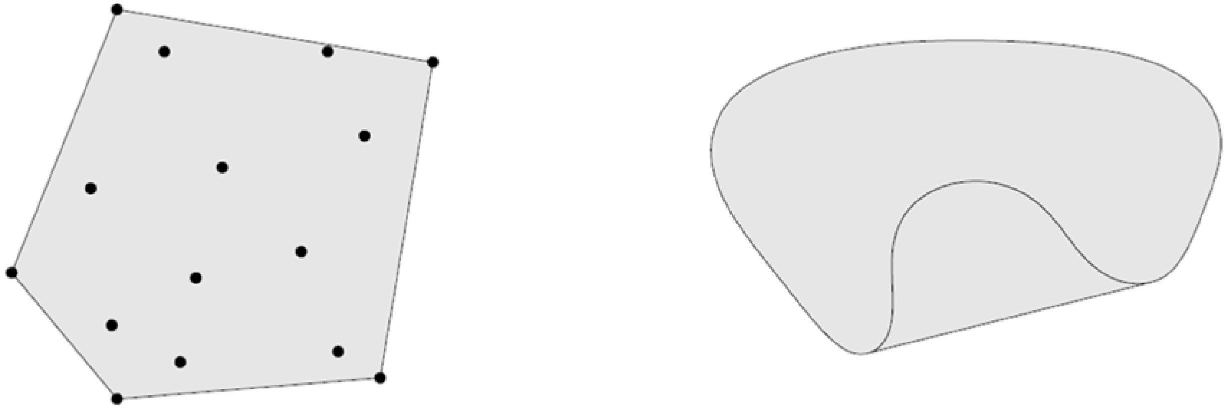
Convex Combination and Convex Hull

- **Convex Combination** of $x_1, \ldots, x_k$: any point x of the form

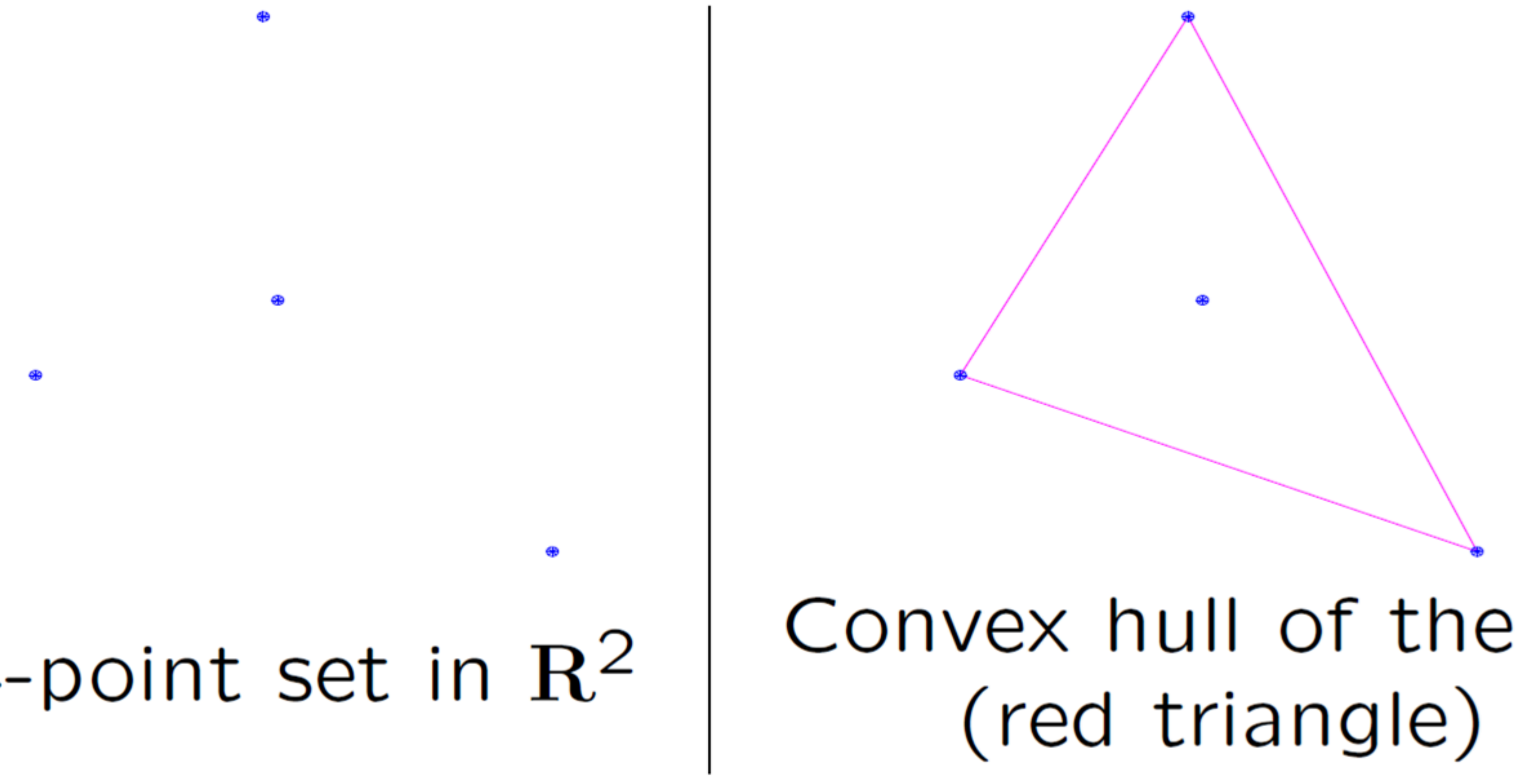
$$x = \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_k x_k$$

$$\text{with } \theta_1 + \cdots + \theta_k = 1, \theta_i \geq 0$$

- **Convex hull** CoS: set of all convex combinations of points in S



Convex Hull



4-point set in $\mathbf{R}^2$

Convex hull of the set (red triangle)

Exercise

Let $C \subseteq \mathbf{R}^n$ be a convex set, with $x_1, \ldots, x_k \in C$, and let $\theta_1, \ldots, \theta_k \in \mathbf{R}$ satisfy $\theta_i \geq 0$, $\theta_1 + \cdots + \theta_k = 1$. Show that $\theta_1 x_1 + \cdots + \theta_k x_k \in C$. (The definition of convexity is that this holds for $k = 2$; you must show it for arbitrary k.) *Hint.* Use induction on k.

Solution. This is readily shown by induction from the definition of convex set. We illustrate the idea for $k = 3$, leaving the general case to the reader. Suppose that $x_1, x_2, x_3 \in C$, and $\theta_1 + \theta_2 + \theta_3 = 1$ with $\theta_1, \theta_2, \theta_3 \geq 0$. We will show that $y = \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 \in C$. At least one of the θ_i is not equal to one; without loss of generality we can assume that $\theta_1 \neq 1$. Then we can write

$$y = \theta_1 x_1 + (1 - \theta_1)(\mu_2 x_2 + \mu_3 x_3)$$

where $\mu_2 = \theta_2/(1 - \theta_1)$ and $\mu_2 = \theta_3/(1 - \theta_1)$. Note that $\mu_2, \mu_3 \geq 0$ and

$$\boxed{\mu_1 + \mu_2} = \frac{\theta_2 + \theta_3}{1 - \theta_1} = \frac{1 - \theta_1}{1 - \theta_1} = 1.$$

Since C is convex and $x_2, x_3 \in C$, we conclude that $\mu_2 x_2 + \mu_3 x_3 \in C$. Since this point and x_1 are in C, $y \in C$.

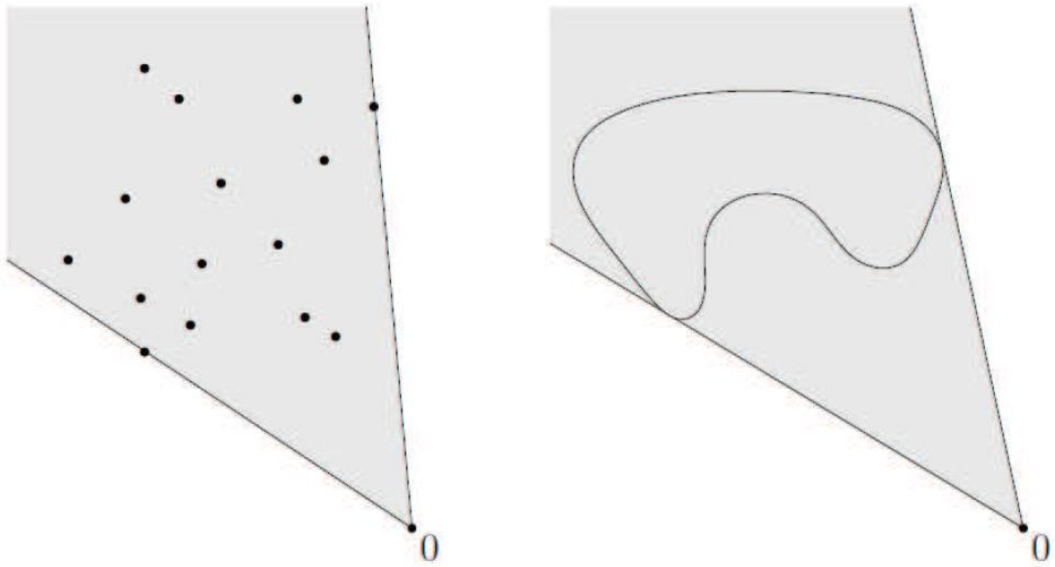
Convex cone

- A set C is called a **cone** if $x \in C \Longrightarrow \theta x \in C, \forall \theta \geq 0$.
- A set C is a **convex cone** if it is convex and a cone, i.e.,

$$x_1, x_2 \in C \Longrightarrow \theta_1 x_1 + \theta_2 x_2 \in C, \qquad \forall\ \theta_{1,}\ \theta_2 \geq 0$$

- The point $\sum_{i=1}^{k} \theta_i x_i$, where $\theta_i \geq 0, \forall i = 1, \ldots, k$, is called a **conic combination** of $x_1, \ldots, x_k$.
- The **conic hull** of a set C is the set of all conic combinations of points in C.

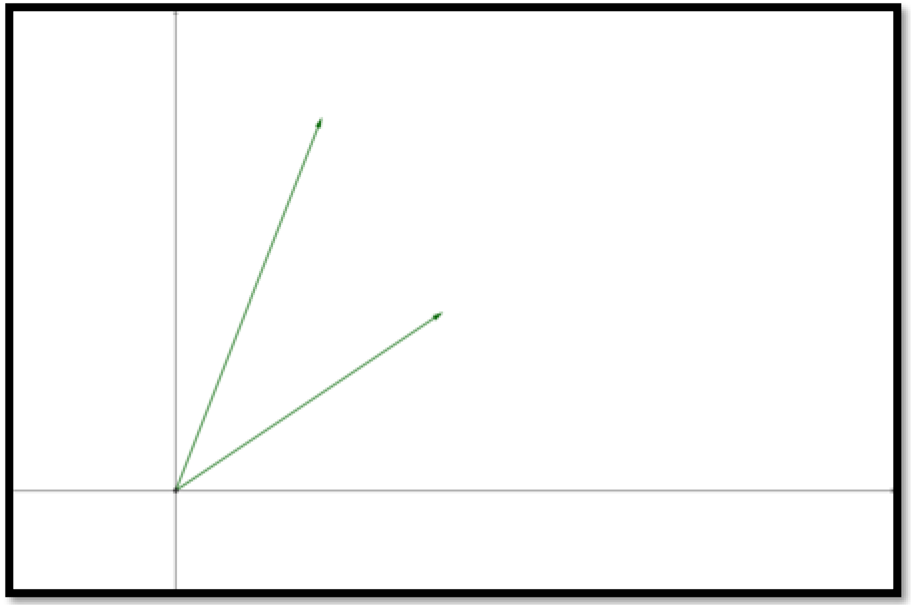
Conic Hull: Examples



Examples of conic hull

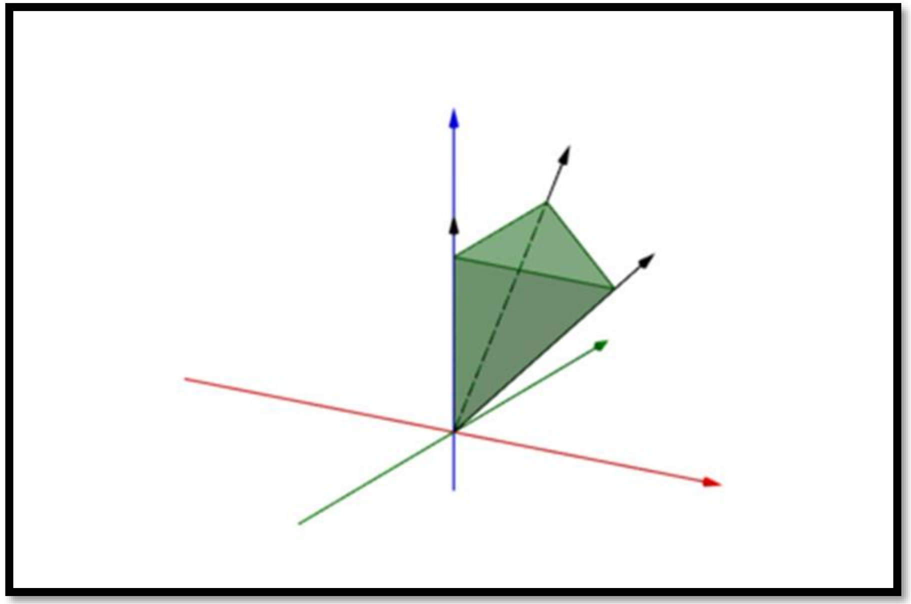
Cone

- A cone (the union of two rays) that is not a convex cone.



Convex Cone

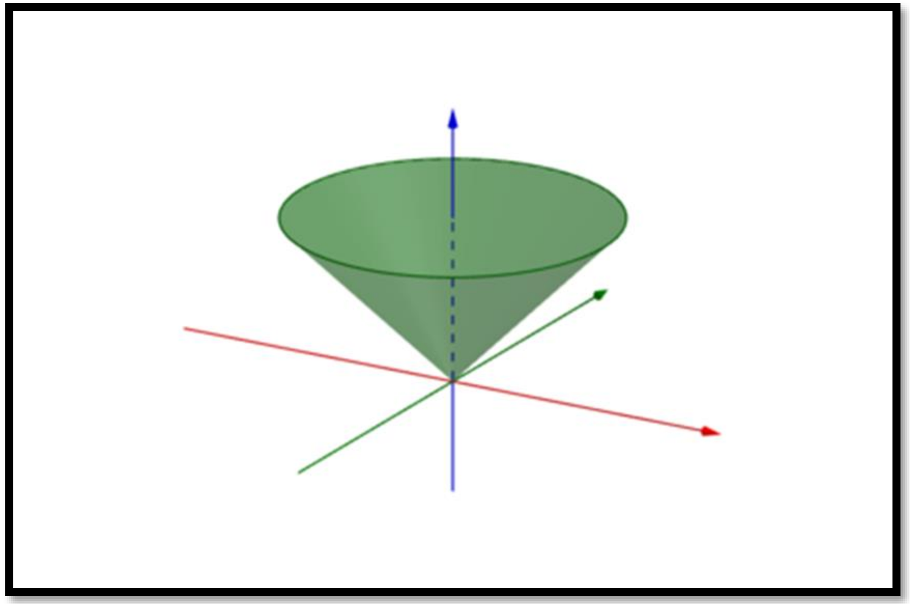
- Convex cone generated by the conic combination of three black vectors.



Pointed polyhedral cone.

Convex Cone

- Convex cone that is not a conic combination of finitely many generators.

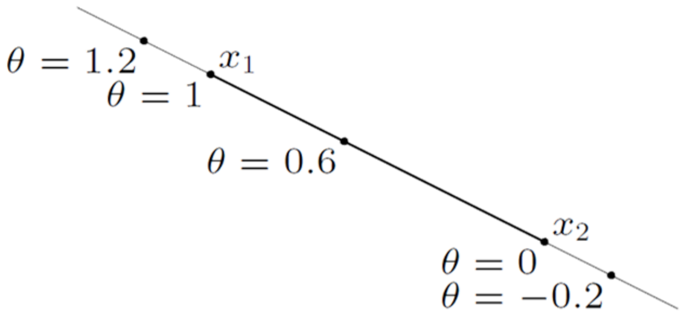


Circular pyramid, example of a cone that is not finitely generated.

Affine Set

- **line** through x_1, x_2: all points

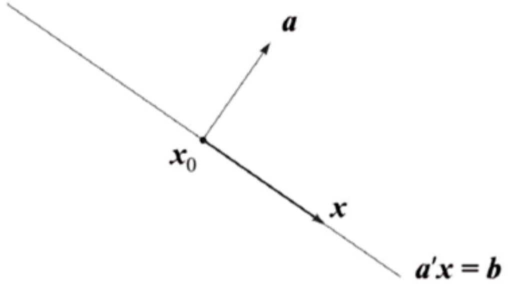
$$x = \theta x_1 + (1-\theta)x_2 \quad (\theta \in \mathbb{R})$$

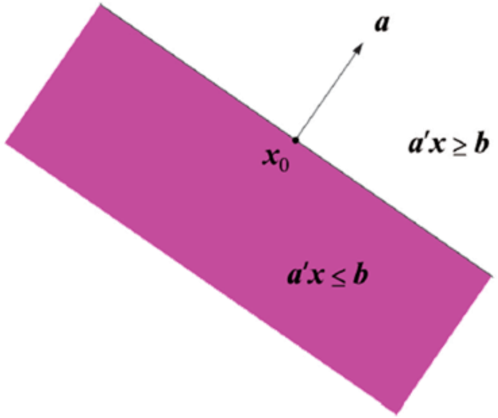


- **affine set**: contains the line through any two distinct points in the set.
- **example**: solution set of linear equations $\{x \mid Ax = b\}$

Hyperplanes and Half-spaces

- **Hyperplane** is a set of the form $\{x \mid a'x = b\}$





- **Half-space** is a set of the form $\{x \mid a'x \leq b\}$ with a nonzero vector a. The vector a is referred to as the **normal vector**.
  - A hyperplane in $\mathbb{R}^n$ divides the space into two half-spaces

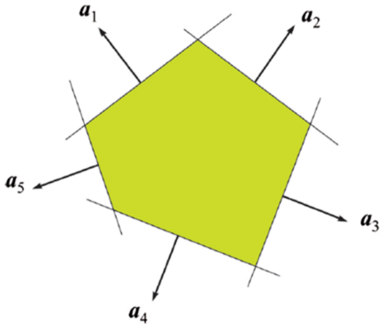
$$\{x \mid a'x \leq b\} \quad and \quad \{x \mid a'x \geq b\}$$

  - Half-spaces are convex
  - Hyperplanes are convex and affine

Polyhedral Sets

- A **polyhedral** set is given by finitely many linear inequalities

$$C = \{x \mid Ax \preceq b\} \quad \text{where } A \text{ is an } m \times n \text{ matrix}$$



  - Every polyhedral set is convex
  - **Linear Problem**

$$\begin{array}{ll} minimize & c'x \\ subject\ to & Bx \leq b, Dx = d \end{array}$$

The constraint set $\{x \mid Bx \leq b,\ Dx = d\}$ is polyhedral

Exercise

Which of the following sets are convex?

1. A *slab*, *i.e.*, a set of the form $\{x \in \mathbf{R}^n \mid \alpha \leq a^T x \leq \beta\}$.
2. A *rectangle*, *i.e.*, a set of the form $\{x \in \mathbf{R}^n \mid \alpha_i \leq x_i \leq \beta_i,\ i = 1, \ldots, n\}$. A rectangle is sometimes called a *hyperrectangle* when $n > 2$.
3. A *wedge*, *i.e.*, $\{x \in \mathbf{R}^n \mid a_1^T x \leq b_1,\ a_2^T x \leq b_2\}$.

Solution.

1. A slab is an intersection of two halfspaces, hence it is a convex set (and a polyhedron).
2. As in part (a), a rectangle is a convex set and a polyhedron because it is a finite intersection of halfspaces.
3. A wedge is an intersection of two halfspaces, so it is convex set. It is also a polyhedron. It is a cone if $b_1 = 0$ and $b_2 = 0$.

Semidefinite Matrices

- Given a square matrix $A \in \mathbb{R}^{n\times n}$ and a vector $x \in \mathbb{R}^n$, the scalar value $x^T Ax$ is called a **quadratic form**. Written explicitly, we see that

$$x^T Ax = \sum_{i=1}^{n} x_i (Ax)_i = \sum_{i=1}^{n} x_i \left(\sum_{j=1}^{n} A_{ij} x_j \right) = \sum_{i=1}^{n} \sum_{i=1}^{n} A_{ij} x_i x_j$$

- A symmetric matrix $A \in \mathbb{S}^n$ is ***positive definite*** (PD) if for all non-zero vectors $x \in \mathbb{R}^n$, $x^T A x > 0$. This is usually denoted $A \succ 0$ (or just $A > 0$), and often times the set of all positive definite matrices is denoted $\mathbb{S}^n_{++}$.

- A symmetric matrix $A \in \mathbb{S}^n$ is ***positive semidefinite*** (PSD) if for all vectors $x^T A x \geq 0$. This is written $A \succeq 0$ (or just $A \geq 0$), and the set of all positive semidefinite matrices is often denoted $\mathbb{S}^n_{+}$.

- Likewise, a symmetric matrix $A \in \mathbb{S}^n$ is ***negative definite*** (ND), denoted $A \prec 0$ (or just $A < 0$) if for all non-zero $x \in \mathbb{R}^n$, $x^T A x < 0$.

- Similarly, a symmetric matrix $A \in \mathbb{S}^n$ is ***negative semidefinite*** (NSD), denoted $A \preceq 0$ (or just $A \leq 0$) if for all $x \in \mathbb{R}^n$, $x^T A x \leq 0$.

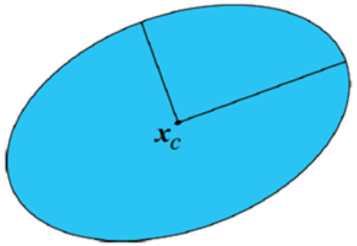
- Finally, a symmetric matrix $A \in \mathbb{S}^n$ is ***indefinite***, if it is neither positive semidefinite nor negative semidefinite — i.e., if there exists $x_1, x_2 \in \mathbb{R}^n$ such that $x_1^T A x_1 > 0$ and $x_2^T A x_2 < 0$.

Ellipsoids

- Let A be a square $(n \times n)$ matrix.
  - A is **positive semidefinite** when $x'Ax \geq 0$ for all $x \in \mathbb{R}^n$
  - A is **positive definite** when $x'Ax > 0$ for all $x \in \mathbb{R}^n, x \neq 0$
- An **ellipsoid** is a set of the form

$$\varepsilon = \{x \mid (x - x_0)' P^{-1} (x - x_0) \leq 1\}$$

Where P is symmetric and positive definite



  - x_0 is the center of the ellipsoid ε
  - A ball $\{x \mid \|x - x_0\| \leq r\}$ is a special case of the ellipsoid $(P = r^2 I)$
  - Ellipsoids are convex

Euclidean Balls and Ellipsoids

- **Euclidean ball** in $\mathbb{R}^n$ with center x_c and radius r:
- **ellipsoid** in $\mathbb{R}^n$ with center x_c:

$$\varepsilon = \{x \mid (x - x_c)^T P^{-1} (x - x_c) \leq 1\}$$

Where $P \in S_{++}^n$ (i.e., symmetric and positive definite)

  - The lengths of the semi-axes of ε are given by $\sqrt{\lambda_i}$, where λ_i are the eigenvalues of P.
  - An alternative representation of an ellipsoid: with $A = P^{1/2}$

$$\varepsilon = \{x_c + Au \mid \|u\|_2 \leq 1\}$$

- Euclidean balls and ellipsoids are convex.

Norms

- A function $f: \mathbb{R}^n \to \mathbb{R}$ is called a **norm** , denoted $\|x\|$, if
  - non-negative: $f(x) \geq 0$, for all $x \in \mathbb{R}^n$
  - definite: $f(x) = 0$ only if $x = 0$
  - homogeneous: $f(tx) = |t|f(x)$, for all $x \in \mathbb{R}^n$ and $t \in \mathbb{R}$
  - satisfies the triangle inequality: $f(x+y) \leq f(x) + f(y)$
- Notation: $\|\cdot\|$ denotes a general norm; $\|\cdot\|_{symb}$ denotes a specific norm
- **Distance**: $dist(x, y) = \|x - y\|$ between $x, y \in \mathbb{R}^n$

Examples of Norms

- $l_p - norm$ on $\mathbb{R}^n$: $\|x\|_p = (|x_1| + \cdots + |x_n|^p)^{1/p}$
  - $l_1 - norm$ on $\mathbb{R}^n$: $\|x\|_1 = \sum_i |x_i|$
  - $l_\infty - norm$ on $\mathbb{R}^n$: $\|x\|_\infty = \max_i |x_i|$
- Quadratic norms: For $P \in S_{++}^n$, define the P-quadratic norm as

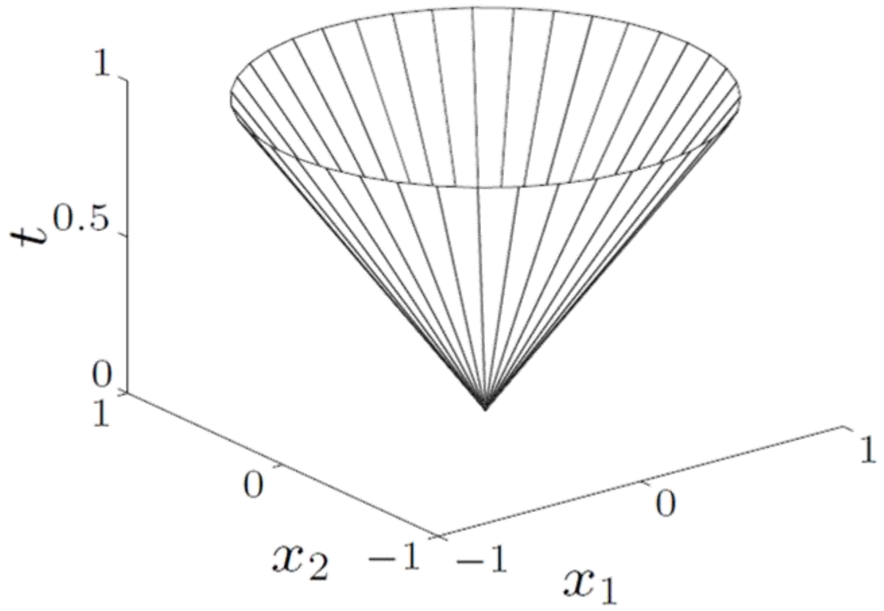
$$\|x\|_P = (x^T P x)^{1/2} = \left\|P^{\frac{1}{2}} x\right\|_2$$

Norm Balls and Norm Cones

- **norm ball** with center x_c and radius r: $\{x \mid \|x - x_c\| \le r\}$
- **norm cone**: $C = \{(x, t) \mid \|x\| \le t\} \subseteq \mathbb{R}^{n+1}$
  - The second-order cone is the norm cone for the Euclidean norm
- Norm balls and cones are convex

Norm Balls and Norm Cones

- **norm**: a function $\|\cdot\|$ that satisfies
  - $\|x\| \geq 0$; $\|x\| = 0$ if and only if $x = 0$
  - $\|tx\| = |t|\|x\|$ for $t \in \mathbb{R}$
  - $\|x + y\| \leq \|x\| + \|y\|$
- notation: $\|\cdot\|$ is general (unspecified) norm; $\|\cdot\|_{symb}$ is particular norm
- **norm ball** with center x_c and radius r: $\{x \mid \|x - x_c\| \leq r\}$
- **norm cone**: $\{(x, t) \mid \|x\| \leq t\}$
- Euclidean norm cone is called second-order cone
- Norm balls and cones are convex

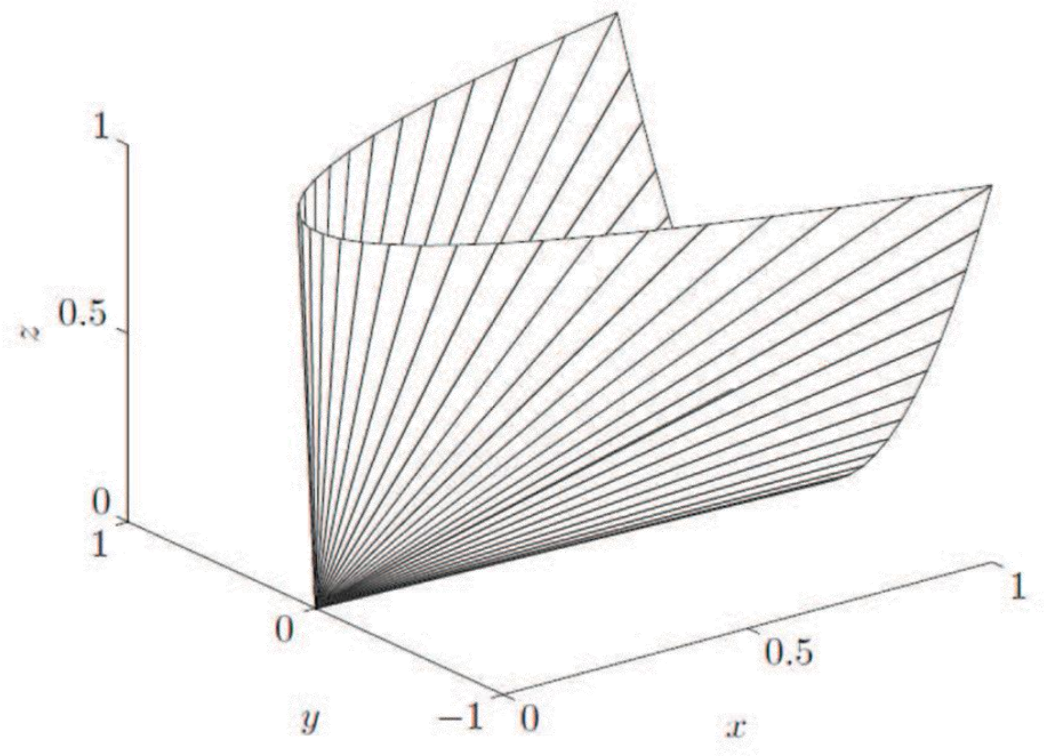


Positive Semidefinite Cone

- Notation:
  - S^n: the set of symmetric $n \times n$ matrices
  - $S_+^n = \{X \in S^n \mid X \succeq 0\}$: symmetric positive semidefinite matrices
  - $S_{++}^n = \{X \in S^n \mid X \succ 0\}$ symmetric positive definite matrices
- S_+^n is a convex cone, called positive semi-definite cone. S_{++}^n comprise the cone interior; all singular positive semi-definite matrices reside on the cone boundary.

Positive Semi-definite Cone: Example

$$X = \begin{bmatrix} x & y \\ y & z \end{bmatrix} \in S_+^2 \Leftrightarrow x \geq 0, z \geq 0, xz \geq y^2$$



Positive Semi-definite cone: S_+^2

Practical Methods for Establishing Convexity of a Set

- Establish the convexity of a given set X
  - The set is one of the "recognizable" (simple) convex sets such as polyhedral, simplex, norm cone, etc.
  - Prove the convexity by directly applying the definition for every $x, y \in X$ and $\alpha \in (0,1)$, show that $\alpha x + (1-\alpha)y$ is also in X.
  - Show that the set is obtained from one of the simple convex sets through an operation that preserves convexity.

Operations Preserving Convexity

- Let $C \subseteq \mathbb{R}^n, C_1 \subseteq \mathbb{R}^n, C_2 \subseteq \mathbb{R}^n$, and $K \subseteq \mathbb{R}^m$ be convex sets. Then, the following sets are also convex:
  - The **intersection** $C_1 \cap C_2 = \{x \mid x \in C_1 \, and \; x \in C_2\}$
  - The **sum** $C_1 + C_2$ of two convex sets
    - The **translated** set $\mathrm{C} + \alpha$
  - The **scaled** set $tC = \{tx \mid x \in C\}$ for any $t \in \mathbb{R}$
  - The **Cartesian product** $C_1 \times C_2 = \{(x_1, x_2) \mid x_1 \in C_1, \; x_2 \in C_2\}$
  - The **coordinate projection** $\{x_1 \mid (x_1, x_2) \in C \; for \; some \; x_2\}$
  - The **image** AC under a linear transformation $A: \mathbb{R}^n \mapsto \mathbb{R}^m$:
  $$AC = \{y \in \mathbb{R}^m \mid y = Ax \; for \; some \; x \in C\}$$
  - The **inverse image** $A^{-1}K$ under a linear transformation $A: \mathbb{R}^n \mapsto \mathbb{R}^m$:
  $$A^{-1}K = \{x \in \mathbb{R}^n \mid Ax \in K\}$$

Exercise

Exercise 1 *Show that the intersection S of any numbers of convex sets S_i is a convex set.*

Solution: Let $C_1, \cdots, C_k$ be convex sets, and define $C := \cap C_k$. Assume $x, y \in C$ and $\lambda \in [0, 1]$. By definition of C we have $x, y \in C_i$, $i \in \{1, \cdots, k\}$, and by convexity we have $\lambda x + (1 - \lambda) y \in C_i$, $i \in \{1, \cdots, k\}$. This means $\lambda x + (1 - \lambda) y \in C$, so C is convex. (A similar argument holds for an infinite number of sets.)

This is not true for the union of convex sets. Let $C_1 := \{0\}$ and $C_2 := \{1\}$ as subsets of $\mathbb{R}$. Both sets are clearly convex, but the union of them is not convex.

Intersection: Example 1

- Show that the positive semi-definite cone S_+^n is convex.

Proof

S_+^n can be expressed as

$$S_+^n = \bigcap_{z \neq 0} \{X \in S^n \mid z^T X z \geq 0\}.$$

Since the set

$$\{X \in S^n \mid z^T X z \geq 0\}$$

Is a half-space in S^n, it is convex. S_+^n is the intersection of an infinite number of half-spaces, so it is convex.

Affine Function: Theorem

Suppose

$f: \mathbb{R}^n \to \mathbb{R}^m$ is an affine function (i.e., $f(x) = Ax + b$).

Then

- the image of a convex set under f is convex

$S \subseteq \mathbb{R}^n$ is convex $\Longrightarrow f(S) = \{f(x) \mid x \in S\}$ is convex

- the inverse image of a convex set under f is convex

$B \subseteq \mathbb{R}^m$ is convex $\Longrightarrow f^{-1}(B) = \{x \mid f(x) \in B\}$ is convex

Affine Function : Example 1

Show that the ellipsoid

$$\varepsilon = \{x \mid (x - x_c)^T P^{-1}(x - x_c) \leq 1\}$$

where $P \in S_{++}^n$ is convex.

Proof

Let

$$S = \{u \in \mathbb{R}^n \mid \|u\|_2 \leq 1\}$$

denote the unit ball in $\mathbb{R}^n$. Define an affine function

$$f(u) = P^{1/2}u + x_c$$

ε is the image of S under f, so is convex.