# Episode 02
# Convex Sets

**Dr. Ahmad F. Taha**

**CE 5999-02 Special Topics — Intro to Optimization**

*Email:* ahmad.taha@vanderbilt.edu

*Webpage:* http://lab.vanderbilt.edu/taha

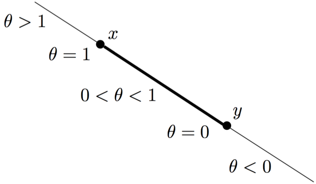
VANDERBILT UNIVERSITY

# Lines and line segments

Suppose $x, y$ are two points in $\mathbb{R}^n$

**Line** through $x$ and $y$: Points of the form $\theta x + (1-\theta)y$ with $\theta \in \mathbb{R}$

**Line segment** between $x$ and $y$: Points of the form $\theta x + (1-\theta)y$ with $\theta \in [0,1]$



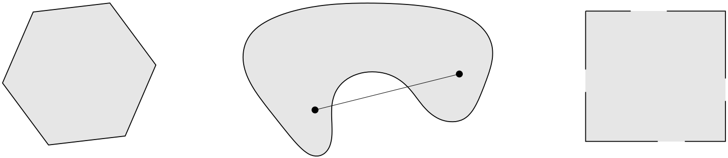
# Convex set

Set $S \subset \mathbb{R}^n$ is **convex** if

$$x, y \in S,\ \theta \in [0,1] \quad \Longrightarrow \quad \theta x + (1-\theta) y \in S$$

Geometrically: $x, y \in S \Longrightarrow$ line segment between $x, y \subset S$

Examples (one convex set and two nonconvex sets)



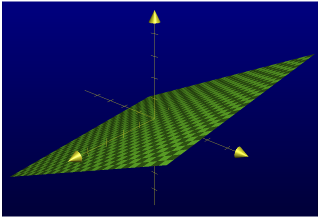
- If the border was entirely missing from the last example, we would have a convex set

# Subspaces

Set $S \subset \mathbb{R}^n$ is a **subspace** if

$$x, y \in S,\ \lambda, \mu \in \mathbb{R} \quad \Longrightarrow \quad \lambda x + \mu y \in S$$

Geometrically: $x, y \in S \Longrightarrow$ plane through $0, x, y \subset S$



Example of a subspace

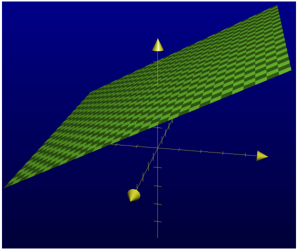
- If a set is subspace, then the origin must belong to the set
- This follows from the definition for $\lambda = \mu = 0$
- A line through the origin is also a subspace

# Affine sets

Set $S \subset \mathbb{R}^n$ is **affine** if

$$x, y \in S, \quad \lambda \in \mathbb{R} \quad \Longrightarrow \quad \lambda x + (1-\lambda)y \in S$$

Geometrically: $x, y \in S \Longrightarrow$ line through $x, y \subset S$



- An affine set is like a subspace with an offset
- A subspace is also an affine set

Subspaces and affine sets are convex.

# Convex cone

Set $S \subset \mathbb{R}^n$ is a **cone** if

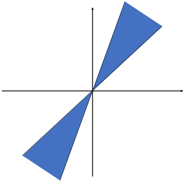
$$x \in S,\ \lambda \geq 0 \quad \Longrightarrow \quad \lambda x \in S$$

Geometrically: $x \in S \Longrightarrow$ ray through $0, x \subset S$

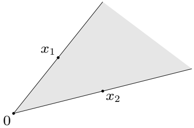
Set $S \subset \mathbb{R}^n$ is a **convex cone** if

$$x, y \in S,\ \lambda, \mu \geq 0 \quad \Longrightarrow \quad \lambda x + \mu y \in S$$

Geometrically: $x, y \in S \Longrightarrow$ "pie slice" between $0, x, y \subset S$



(Nonconvex) cone




Convex cone

# Examples

- The empty set $\emptyset$, any single point (i.e., singleton) $x_0$, and the whole space $\mathbb{R}^n$ are affine (hence, convex) subsets of $\mathbb{R}^n$.
- Any line is affine. If it passes through zero, it is a subspace, hence also a convex cone.
- A line segment is convex, but not affine (unless it reduces to a point).
- A ray, which has the form $\{x_0 + \theta v | \theta \geq 0\}$, where $v \neq 0$, is convex, but not affine. It is a convex cone if its base $x_0 = 0$.
- Any subspace is affine, and a convex cone (hence convex).

# Some notations

- The notation $a \preceq b$, where $a$ and $b$ are **vectors**, means *componentwise inequality*:

$$\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \preceq \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \iff a_i \leq b_i, \ i = 1, \ldots, n$$

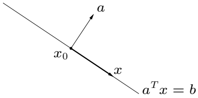
- The notation $a \prec b$ means componentwise strict inequality
- For a **matrix** $A$, the notation $A \succ 0$ means $A$ is positive definite (sometimes abbreviated as p.d.)
- For a matrix $A$, $A \succeq 0$ means $A$ is positive semidefinite
- **When we say that a matrix is positive/negative (semi-)definite or indefinite, we always imply that the matrix is symmetric**
- The notation $A \succeq B$ means that $A - B \succeq 0$, that is, $B - A$ is positive semidefinite
- Consult the modeling language documentation about how $>=$ or $<=$ is interpreted

# Hyperplanes and halfspaces

**hyperplane** is set of the form $\{x \mid a^T x = b\}$

a hyperplane is an affine set; it is a subspace if $b = 0$

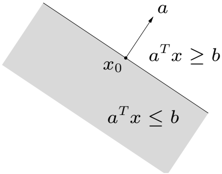
$a$ is called *normal vector*




**halfspace** is set of the form $\{x \mid a^T x \leq b\}$

a halfspace is a convex set; it is a convex cone if $b = 0$

$a$ is called *(outward) normal vector*



# Another representation of halfspaces

Consider any point $x_0$ on the hyperplane defined by $a^T x = b$ (that is, $a^T x_0 = b$).

The halfspace can be represented in two ways

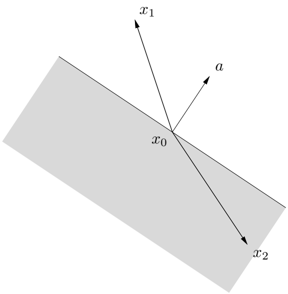
$$\{x \mid a^T x \leq b\} = \{x \mid a^T (x - x_0) \leq 0\}$$

- Vector $x_1 - x_0$ forms an acute angle with $a$, so $x_1$ is not in the halfspace

$$a^T (x_1 - x_0) > 0$$

- Vector $x_2 - x_0$ forms an obtuse angle with $a$, so $x_2$ is in the halfspace

$$a^T (x_2 - x_0) < 0$$



# Polyhedra

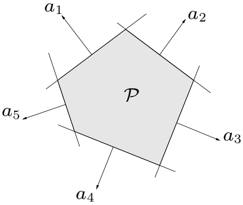
**Polyhedron** is the solution set of finitely many linear inequalities and equalities

$$\mathcal{P} = \{x \mid a_j^T x \leq b_j,\ j = 1, \ldots, m; \quad c_j^T x = d_j,\ j = 1, \ldots, p\}$$
$$= \{x \mid Ax \preceq b,\ Cx = d\}$$

where $A$ has the vectors $a_j^T$ as rows, and similarly for $C$; and

vector $b$ has entries $b_j$, and similarly for $d$

A polyhedron is a convex set. Proof?

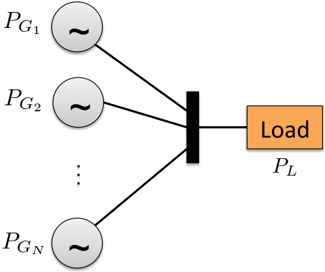



Example: The nonnegative orthant

$\mathbb{R}_+^n = \{x \in \mathbb{R}^n \mid x_i \geq 0,\ i = 1, \ldots, n\} = \{x \in \mathbb{R}^n \mid x \succeq 0\}$ is a polyhedron and a convex cone.

# Motivating example from power engineering

- Suppose $N$ power generation units serve a given load $P_L$ (e.g., a city)
- The power output of unit $i$ is $P_{G_i}$ MW
- The cost of operating unit $i$ is $C_i(P_{G_i})$ \$/h



# The economic dispatch problem

Given the cost functions $C_i(P_{G_i})$ and the load $P_L$, find the most economically generated power output.

$$\begin{aligned} \min \quad & \sum_{i=1}^{N} C_i(P_{G_i}) \\ \text{subj. to} \quad & \sum_{i=1}^{N} P_{G_i} = P_L \\ & 0 \leq P_{G_i},\ i = 1, \ldots, N \end{aligned}$$

Is the feasible set convex? Polyhedron?

What if each unit has a limit on how much it can produce, i.e., $P_{G_i} \leq P_{G_i}^{\max}$?

# Another application: Classification

Suppose we are given $N$ measurements, $x_1, \ldots, x_N$ (where $x_i \in \mathbb{R}^n$). We are also given that the first $Q$ correspond to Class 1, and the remaining $Q+1, \ldots, N$ correspond to Class 2.

If we obtain a new measurement $x$, how can we decide if it belongs to Class 1 or 2?

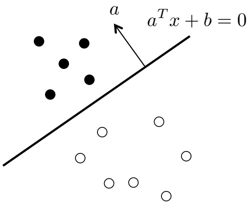
For example, the measurement vector $x_i$ may consist of height, age, weight, blood pressure, and other medical parameters. Classes 1 and 2 may correspond to 'disease' or 'no disease.'

# Linear Classifier: Definition

Suppose we are given $N$ measurements, $x_1, \dots, x_N$ (where $x_i \in \mathbb{R}^n$). We know that the first $Q$ correspond to Class 1, and the remaining $Q+1, \dots, N$ correspond to Class 2.

A **linear classifier**, or **support vector machine**, is an affine function $f(x) = a^T x + b$ that is positive for $x_1, \dots, x_Q$, and negative for $x_{Q+1}, \dots, x_N$:

$$a^T x_i + b > 0, \quad i = 1, \dots, Q; \quad a^T x_i + b < 0, \quad i = Q+1, \dots, N$$




*Prediction:* After we find the classifier, then if we obtain a new measurement $x$, we can decide if it belongs to Class 1 or Class 2 by checking if $a^T x + b > 0$ or $a^T x + b < 0$.

# Norm balls

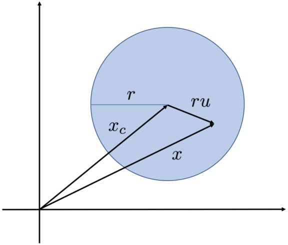
Norm ball with center $x_c \in \mathbb{R}^n$ and radius $r \geq 0$

$$B(x_c, r) = \{x \in \mathbb{R}^n \mid \|x - x_c\| \leq r\}$$

Special case: Euclidean ball

$$B(x_c, r) = \{x \in \mathbb{R}^n \mid \|x - x_c\|_2 \leq r\}$$
$$= \{x = x_c + ru \mid \|u\|_2 \leq 1\}$$

- In the figure on the side, $x - x_c = ru$
- $u$ is a vector with any direction, but length $\|u\|_2 \leq 1$



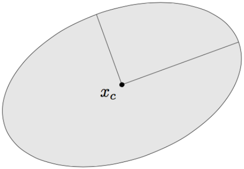

Norm balls are convex sets.

Proof: Use the definition of convexity and the triangle inequality

# Ellipsoids

Ellipsoid $\mathcal{E} \subset \mathbb{R}^n$—three equivalent representations:

1. $\mathcal{E} = \{x \mid (x - x_c)^T A^{-1} (x - x_c) \leq 1\}$, where $A$ is a positive definite matrix
   - If $A = r^2 I$, the ellipsoid becomes the Euclidean ball with radius $r$
2. $\mathcal{E} = \{x = x_c + Bu \mid \|u\|_2 \leq 1\}$, where $B$ is a nonsingular (invertible) square matrix
   - If $B = rI$, the ellipsoid becomes the Euclidean ball with radius $r$
3. $\mathcal{E} = \{x \mid x^T Cx + 2d^T x + e \leq 0\}$, where $C$ is a p.d. matrix and $e - d^T C^{-1} d < 0$
   - The quantity $e - d^T C^{-1} d$ is called discriminant



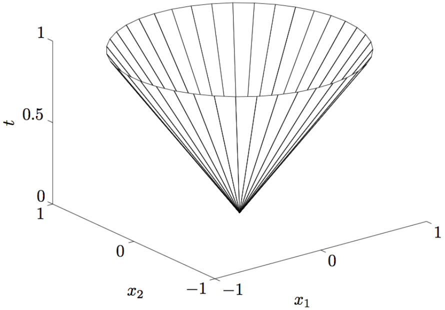

Ellispoids are convex sets.

Proof: Use the second representation and the definition of convexity

# Norm cones

Norm cone $C = \{(x,t) \in \mathbb{R}^{n+1} \mid \|x\| \leq t\}$

- Special case is when the norm $\|x\|$ is the Euclidean norm $\|x\|_2$
- The norm cone is then called second-order cone, Lorentz cone, or ice-cream cone




**Norm cones are convex cones.**

Proof: Use the definition of convexity and the triangle inequality

# The positive semidefinite cone

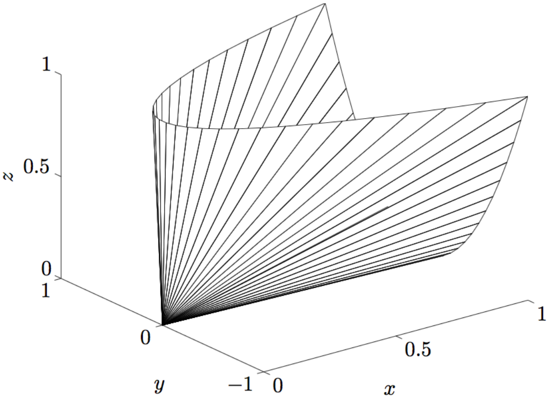
- $\mathbb{S}^n$: the set of symmetric $n \times n$ matrices
- $\mathbb{S}^n_+ = \{X \in \mathbb{S}^n \mid X \succeq 0\}$: the set of positive semidefinite matrices
- $\mathbb{S}^n_{++} = \{X \in \mathbb{S}^n \mid X \succ 0\}$: the set of positive definite matrices

The set $\mathbb{S}^n$ is a subspace; the set $\mathbb{S}^n_+$ is a convex cone; the set $\mathbb{S}^n_{++}$ is convex

Proof: Use the corresponding definition

# Positive semidefinite cone example

$$X = \begin{bmatrix} x & y \\ y & z \end{bmatrix} \in \mathbb{S}^2_+ \iff x \geq 0,\ z \geq 0,\ y^2 \leq xz$$



# Linear Matrix Inequality (LMI)

$$x_1 A_1 + \ldots + A_n x_n \preceq B$$

where $A_i \in \mathbb{S}^n$ and $B \in \mathbb{S}^n$.

The solution set of an LMI

$$S = \{x \in \mathbb{R}^n \,|\, x_1 A_1 + \ldots + A_n x_n \preceq B\}$$

is a convex set.

Proof: Use the definition of convex set.

# Why bother with the positive semidefinite cone and LMIs?

- Later on, we will deal with problems where the optimization variable is a *matrix*
- These problems are formulated using LMIs
- LMIs find application in many engineering problems:
  1. Many problems in Control are formulated as optimization problems with LMIs
  2. LMIs are used to come with tractable approximations of the nonconvex problems with nonlinear equality constraints or integer variables

# Combinations and hulls

Consider $k$ points $\{x_1, \ldots, x_k\}$ in $\mathbb{R}^n$. Then $y = \theta_1 x_1 + \ldots + \theta_k x_k$ is

- linear combination if $\theta_i \in \mathbb{R}$
- affine combination if $\sum_{i=1}^{k} \theta_i = 1$
- convex combination if $\sum_{i=1}^{k} \theta_i = 1$ and $\theta_i \geq 0$
- conic combination if $\theta_i \geq 0$

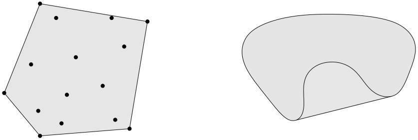
Consider an arbitrary set $S \subset \mathbb{R}$

- $\mathrm{span}(S)$: linear hull of $S$ - set of all linear combinations from $S$
- $\mathrm{aff}(S)$: affine hull of $S$ - set of all affine combinations from $S$
- $\mathrm{conv}(S)$: convex hull of $S$ - set of all convex combinations from $S$
- $\mathrm{cone}(S)$: conic hull of $S$ - set of all conic combinations from $S$

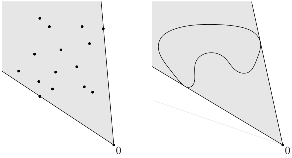
$\mathrm{span}(S)$ is a subspace, $\mathrm{aff}(S)$ is affine, $\mathrm{conv}(S)$ is convex, $\mathrm{cone}(S)$ is convex cone

# Examples

- Convex hulls



- Conic hulls




- Consider the points $e_1 = (1,0)$ and $e_2 = (0,1)$ in $\mathbb{R}^2$. Find their linear, affine, convex, conic hull.

# Images and inverse images of convex sets under affine mappings

A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is called affine if it is expressed as

$$f(x) = Ax + b$$

where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$.

1. Suppose $S \subset \mathbb{R}^n$ is a convex set. Then, the **image** of $S$ under $f$

$$f(S) = \{f(x) \,|\, x \in S\}$$

   is a convex set.

2. Suppose $S \subset \mathbb{R}^m$ is a convex set. Then, the **inverse image** of $S$ under $f$

$$f^{-1}(S) = \{x \in \mathbb{R}^n \,|\, f(x) \in S\}$$

   is a convex set.

# Set operations that preserve convexity

Let $a \in \mathbb{R}$, $b \in \mathbb{R}^n$ and suppose that sets $S \subset \mathbb{R}^n$, $S_1$, and $S_2$ are convex.

The following sets are convex.

1. *Scaling:* $aS = \{ax \,|\, x \in S\}$.
2. *Translation:* $S + b = \{x + b \,|\, x \in S\}$
3. *Sum:* $S_1 + S_2 = \{x_1 + x_2 \,|\, x_1 \in S_1, x_2 \in S_2\}$
4. *Direct or Cartesian Product:* $S_1 \times S_2 = \{(x_1, x_2) \,|\, x_1 \in S_1, x_2 \in S_2\}$
5. *Projection onto some of its coordinates:* Suppose $S \subset \mathbb{R}^m \times \mathbb{R}^n$ is convex.

   $T = \{x_1 \in \mathbb{R}^m \,|\, (x_1, x_2) \in S$ for some $x_2 \in \mathbb{R}^n\}$

   - Image of $S$ under the affine mapping $\mathbf{I}_m x_1 + \mathbf{0}_{m\times n} x_2$

# Questions And Suggestions?


**Thank You!**

Please visit

https://lab.vanderbilt.edu/taha/

**IFF** you want to know more ☺