FOUNDATIONS OF OPTIMIZATION: IE6001

Convex Sets

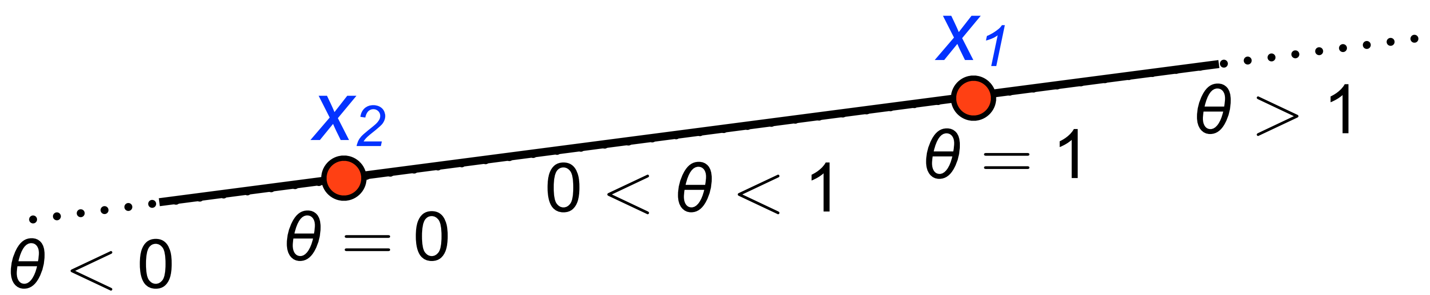
Napat Rujeerapaiboon

Semester I, AY2022/2023

Lines, Line Segments and Rays

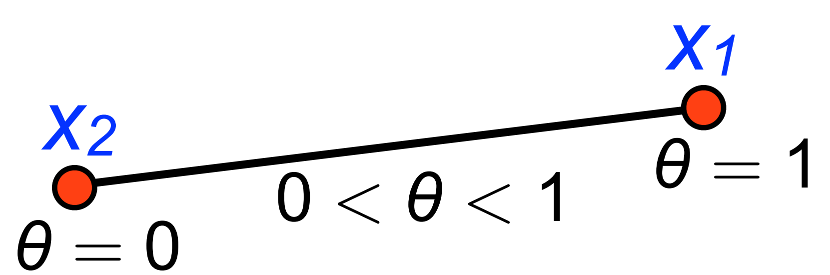
- The *line* through x_1 and x_2 is the set of all points

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



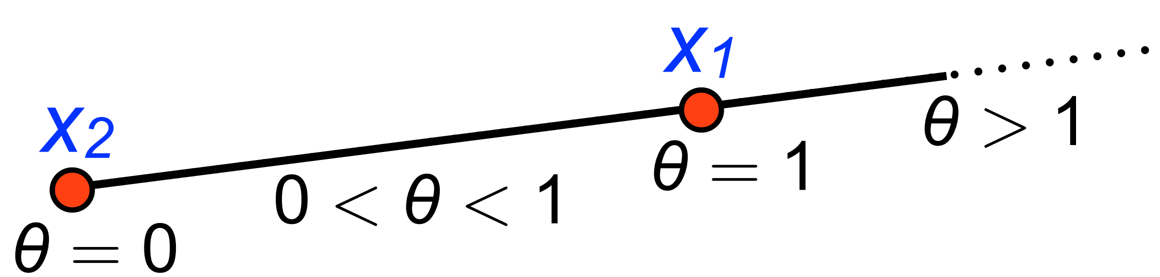
- The *line segment* between x_1 and x_2 is the set of all points

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



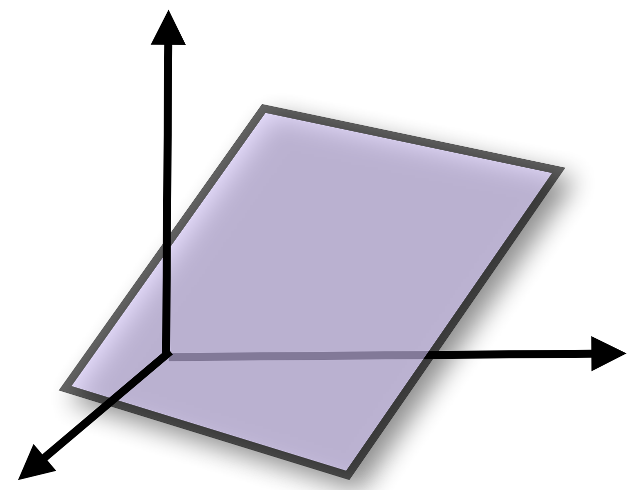
- The *ray* emanating from x_2 through x_1 is the set of points

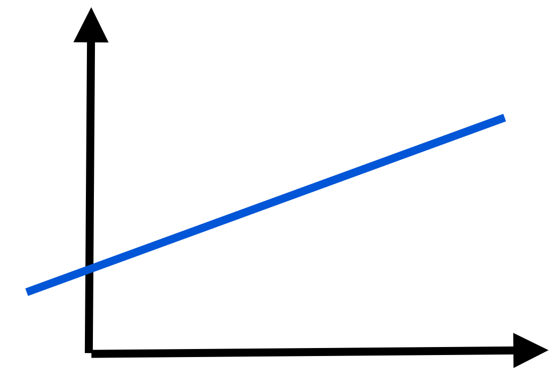
$$x = \theta x_1 + (1-\theta)x_2 \quad \forall \theta \in \mathbb{R}_+.$$

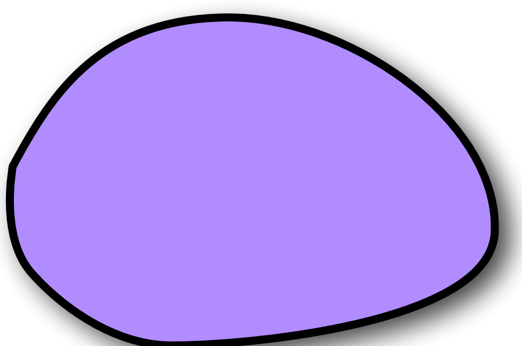


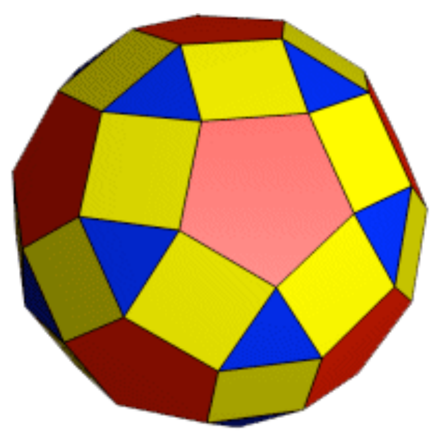
Affine, Convex and Conic Sets

A set is called *affine* if it contains all lines through any two of its points.



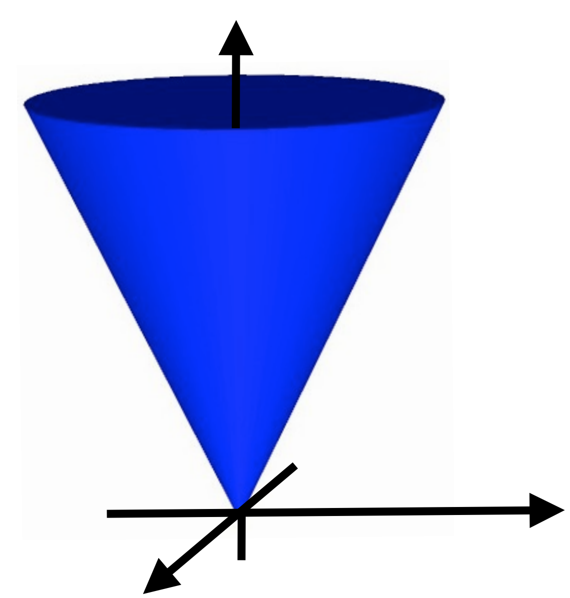






A set is called *convex* if it contains all line segments through any two of its points.

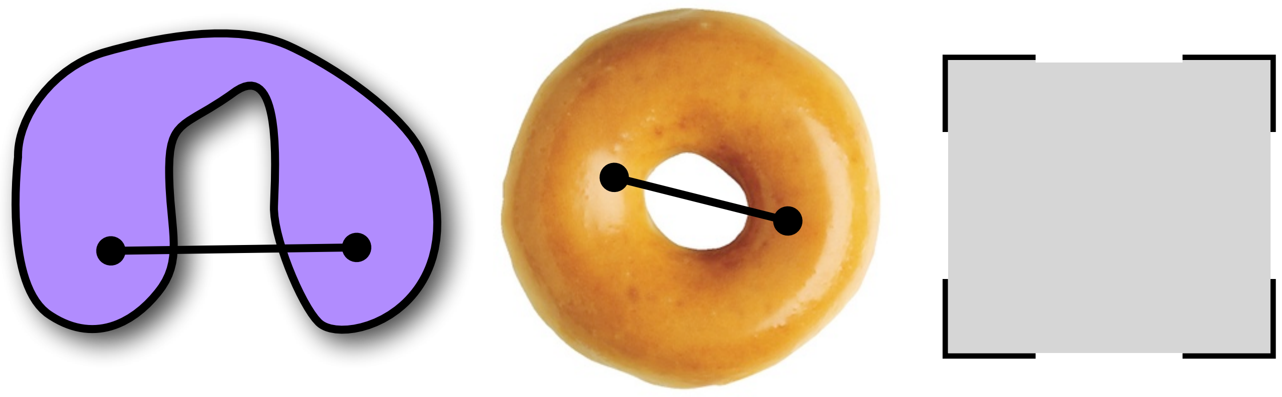
A set is called a *cone* if it contains every ray emanating from 0 through any of its points.



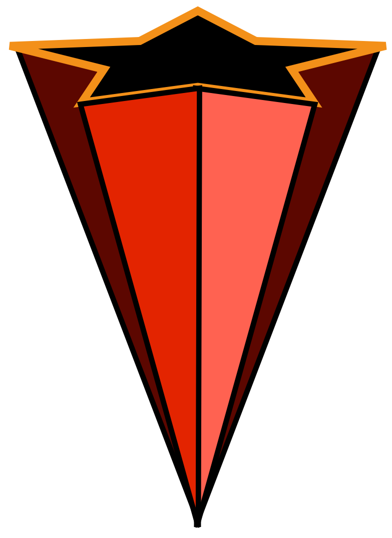
Examples

- Solution sets of *linear equations* are affine. $\{x : Ax = b\}$
- Conversely, every affine set is representable as the solution set of a system of linear equations.
- Every affine set is *convex*.

Examples of *non-convex sets*:



Example of a *non-convex cone*:



More Examples

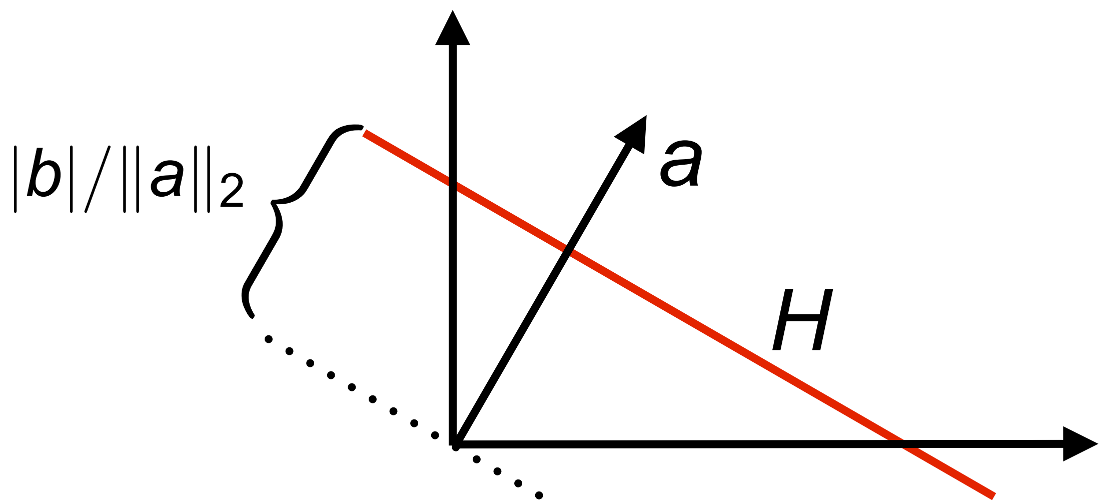
- The *empty set* $\emptyset$ is affine.
- Any *singleton* set of the form $\{x_0\}$ is affine.
- *Lines* are affine.
- *Line segments* and *rays* are convex.
- The *Euclidean Space* $\mathbb{R}^n$ is affine.
- The space $\mathbb{S}^n$ of *symmetric matrices* in $\mathbb{R}^{n\times n}$ is affine. Its dimension is $(n+1)n/2$.

Example:

$$\mathbb{S}^2 = \left\{ x_1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + x_3 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} : x \in \mathbb{R}^3 \right\}$$

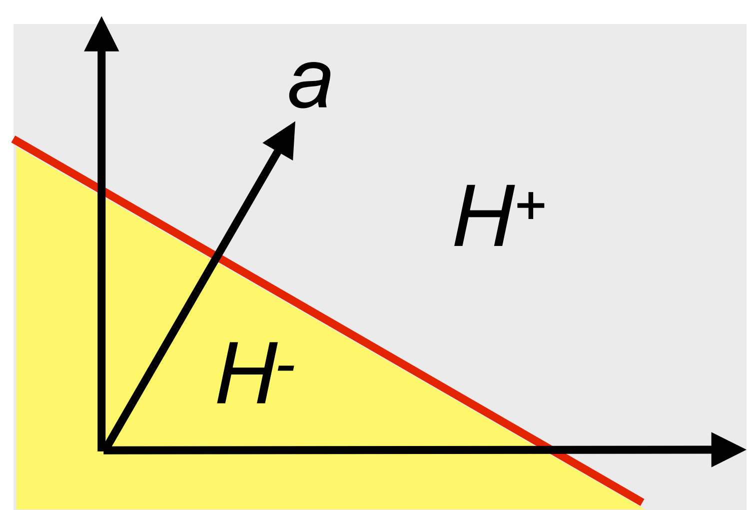
Hyperplanes and Halfspaces

- *Hyperplanes* of the form $H = \{x : a^\top x = b\}$ with normal vector $a \neq 0$ are affine.



- Hyperplanes split the space into convex *halfspaces* of the form

$$H^- = \{x : a^\top x \leq b\} \text{ and } H^+ = \{x : a^\top x \geq b\}$$

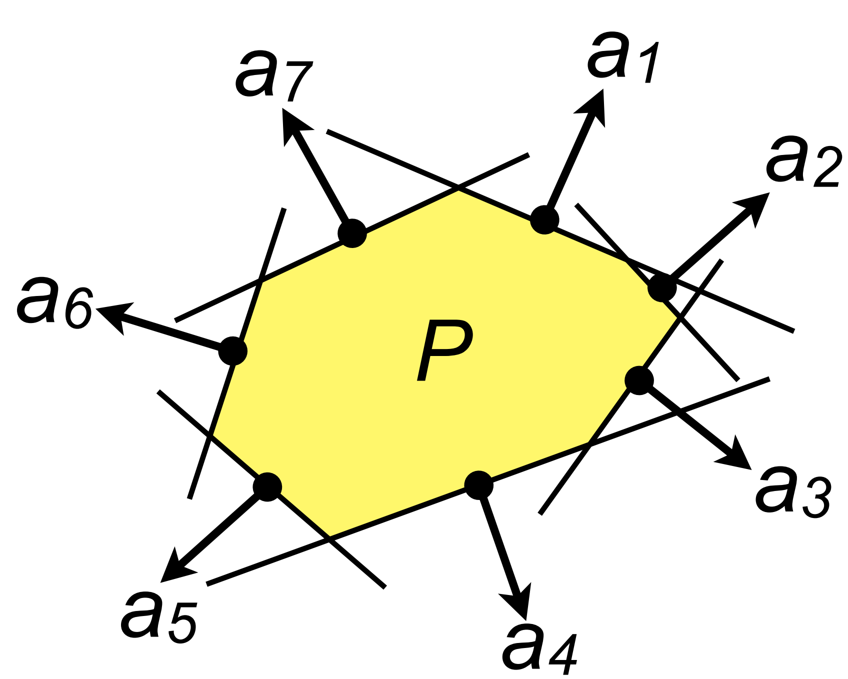


Polyhedra

A *polyhedron* is an intersection of halfspaces.

$$P = \{x : a_i^\top x \leq b_i \ \forall i = 1, \ldots, m\}$$

Equivalently, we can write $P = \{x : Ax \leq b\}$, where A is the matrix with rows $a_i^\top$, b is the vector with entries b_i, and the inequality between vectors is understood to hold componentwise.



Polyhedra are convex as they are intersections of convex sets.

Euclidean Balls

The Euclidean *ball* with radius ε around μ is representable as

$$\mathcal{B}(\mu, \varepsilon) = \{x : \|x - \mu\|_2 \leq \varepsilon\} = \{\mu + \varepsilon u : \|u\|_2 \leq 1\}.$$

It is convex by the triangle inequality of the Euclidean norm.

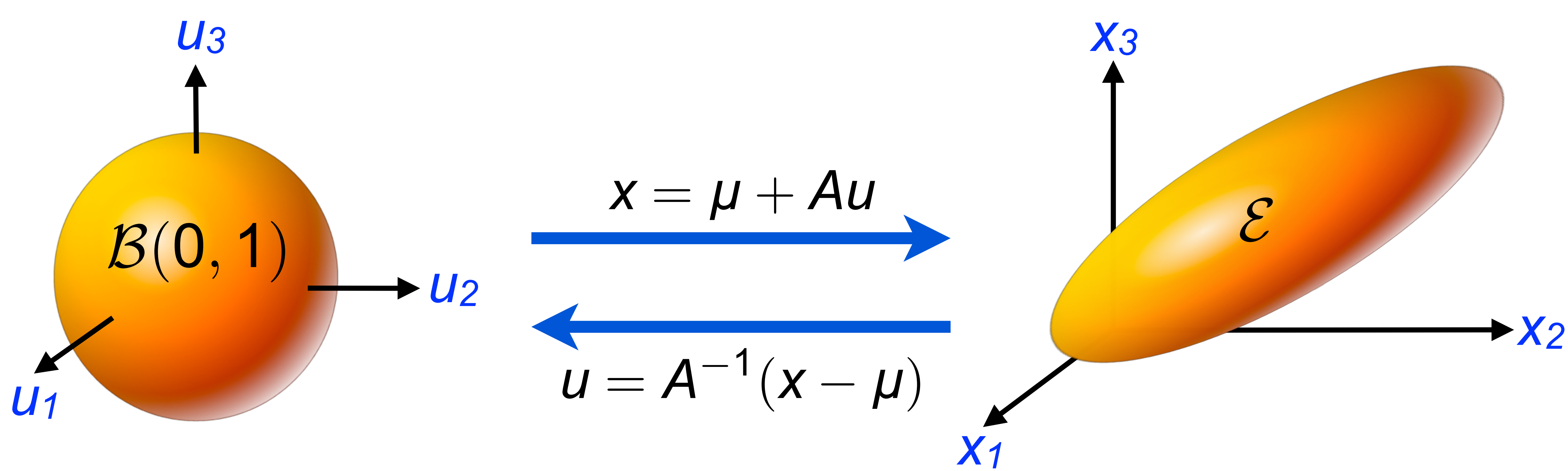
$$\begin{aligned} & x_1, x_2 \in \mathcal{B}(\mu, \varepsilon),\ \theta \in [0, 1] \\ \Longrightarrow \quad & \|\theta x_1 + (1 - \theta) x_2 - \mu\|_2 \leq \theta \|x_1 - \mu\|_2 + (1 - \theta)\|x_2 - \mu\|_2 \leq \varepsilon \\ \Longrightarrow \quad & \theta x_1 + (1 - \theta) x_2 \in \mathcal{B}(\mu, \varepsilon) \end{aligned}$$

Ellipsoids

An *ellipsoid* with center μ is a convex set of the form

$$\mathcal{E} = \{\mu + Au : \|u\|_2 \leq 1\}$$

where A is a nonsingular square matrix.



An equivalent representation is

$$\mathcal{E} = \{x : (x-\mu)^\top \Sigma^{-1}(x-\mu) \leq 1\}$$

where $\Sigma = AA^\top$ is positive definite.

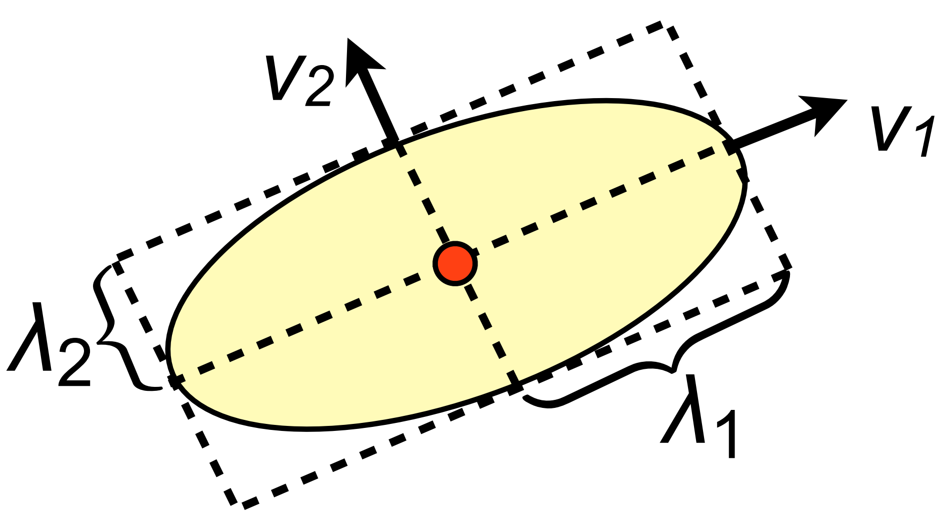
Ellipsoids

Without any loss, we can assume that A is *positive definite*.

$$\begin{aligned}
\{\mu + Au : \|u\|_2 \leq 1\} &= \{x : \|A^{-1}(x-\mu)\|_2 \leq 1\} \\
&= \{x : (x-\mu)^\top (A^{-1})^\top A^{-1}(x-\mu) \leq 1\} \\
&= \{x : (x-\mu)^\top (AA^\top)^{-1}(x-\mu) \leq 1\} \\
&= \{x : \|(AA^\top)^{-1/2}(x-\mu)\|_2 \leq 1\} \\
&= \{\mu + A'u : \|u\|_2 \leq 1\}
\end{aligned}$$

Thus, *we can set* $A \to A' = (AA^\top)^{1/2}$.

If A is positive definite, its (normalized) eigenvectors v_1, v_2,..., v_n determine the *directions* of the *semi-axes* of $\mathcal{E}$. The corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ represent the *lengths* of the *semi-axes*.



Convex Cones

Reminder:

C is a *cone* if: $x \in C,\ \theta \geq 0 \implies \theta x \in C$.

C is *convex* if: $x_1, x_2 \in C,\ \theta \in [0,1] \implies \theta x_1 + (1-\theta)x_2 \in C$.

If C is a convex cone, then: $x_1, x_2 \in C \implies x_1 + x_2 \in C$.

Examples of convex cones:

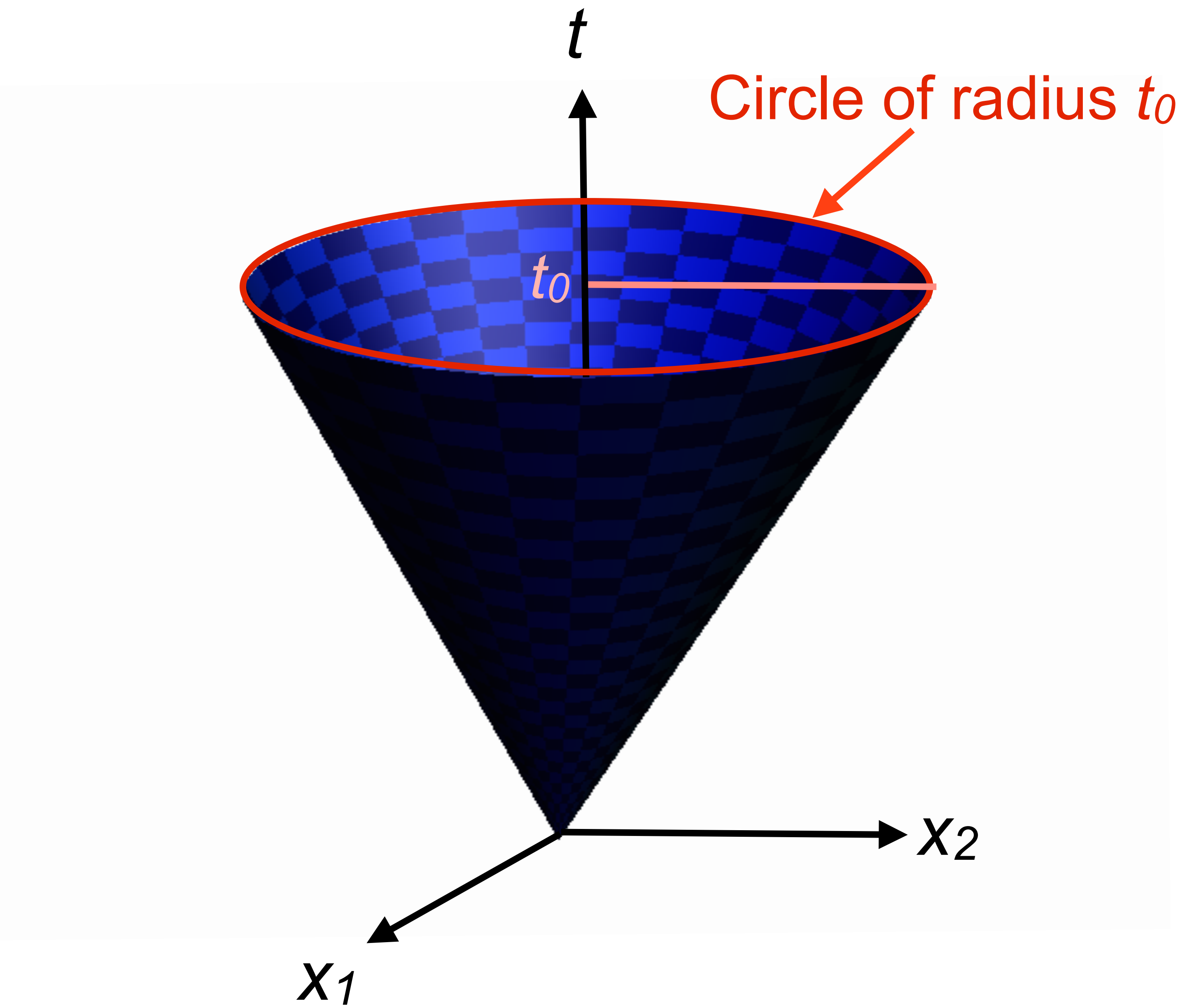
The *non-negative orthant*: $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n : x \geq 0\}$

The *second-order cone*: $\{(x,t) \in \mathbb{R}^n \times \mathbb{R} : \|x\|_2 \leq t\}$

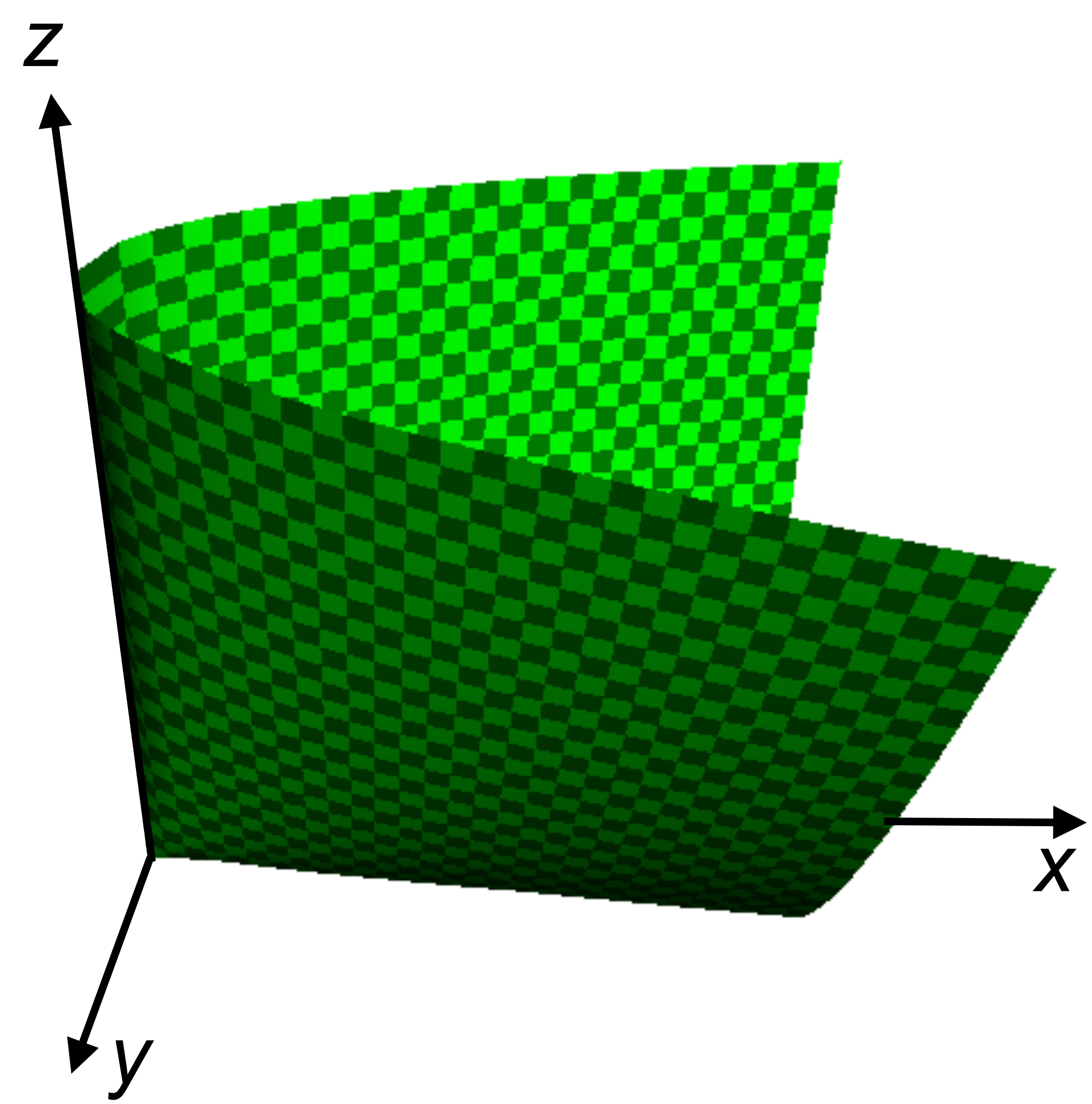
The *positive semidefinite cone*: $\mathbb{S}^n_+ = \{X \in \mathbb{S}^n : X \succeq 0\}$

$$X \in \mathbb{S}^n_+ \iff z^\top X z \geq 0\ \forall z \in \mathbb{R}^n$$

Convex Cones



Second-order cone
(also ice-cream cone or
Lorentz cone)



Positive semidefinite cone

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

Affine, Convex and Conic Combinations

- An *affine combination* of x_1, x_2, ... , x_k is any point of the form

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

for some $\theta_i \in \mathbb{R}$, $\theta_1 + \theta_2 + \cdots + \theta_k = 1$.

- A *convex combination* of x_1, x_2, ... , x_k is a point of the form

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

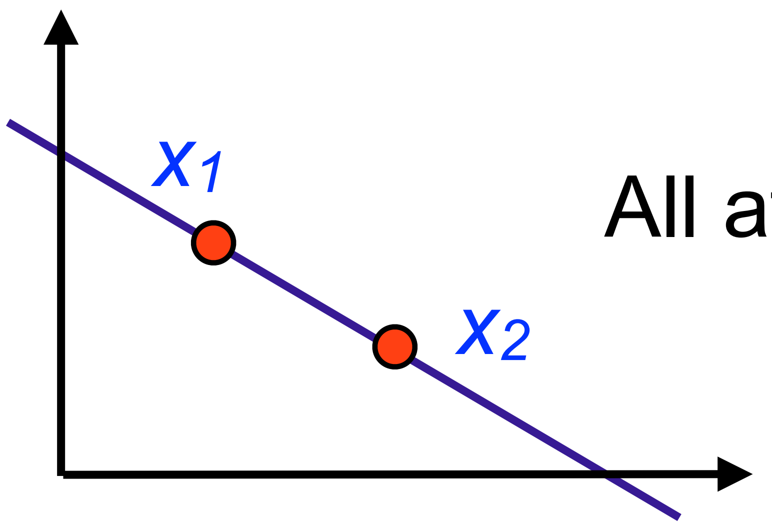
for some $\theta_i \in \mathbb{R}_+$, $\theta_1 + \theta_2 + \cdots + \theta_k = 1$.

- A *conic combination* of x_1, x_2, ... , x_k is a point of the form

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

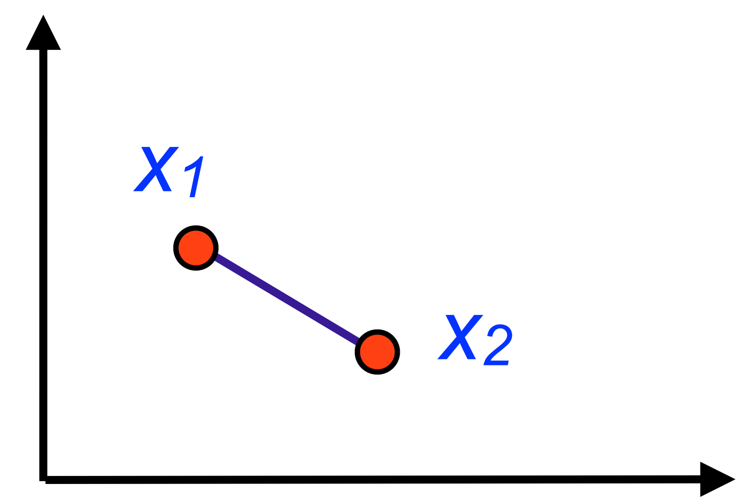
for some $\theta_i \in \mathbb{R}_+$.

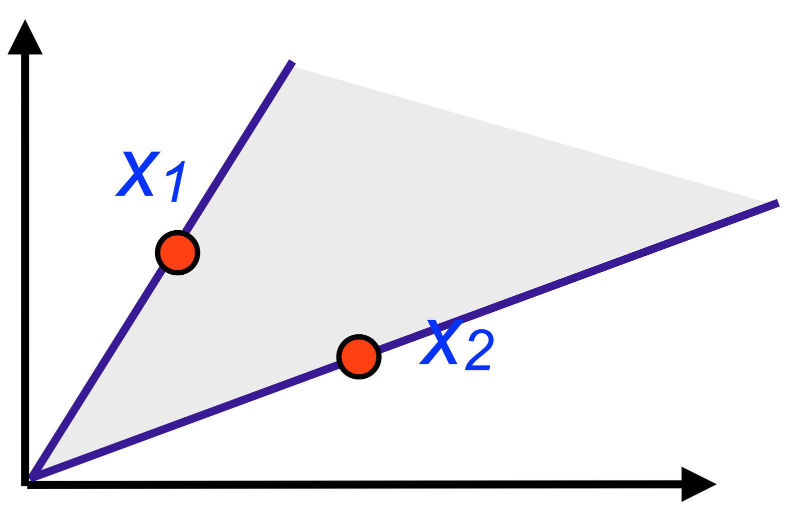
Affine, Convex and Conic Combinations



All affine combinations of x_1 and x_2

All convex combinations of x_1 and x_2

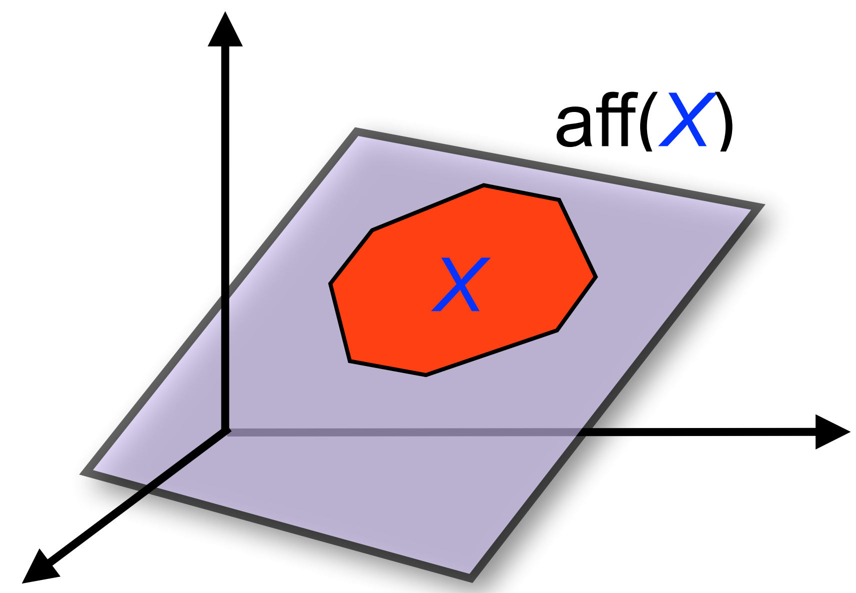


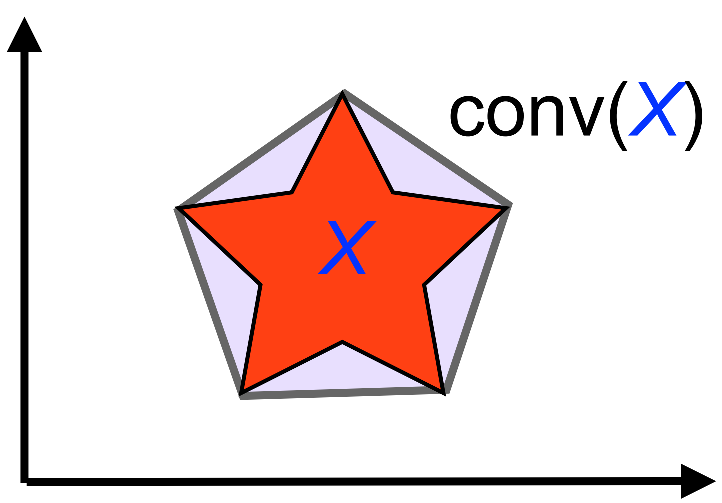


All conic combinations of x_1 and x_2

Affine, Convex and Conic Hulls

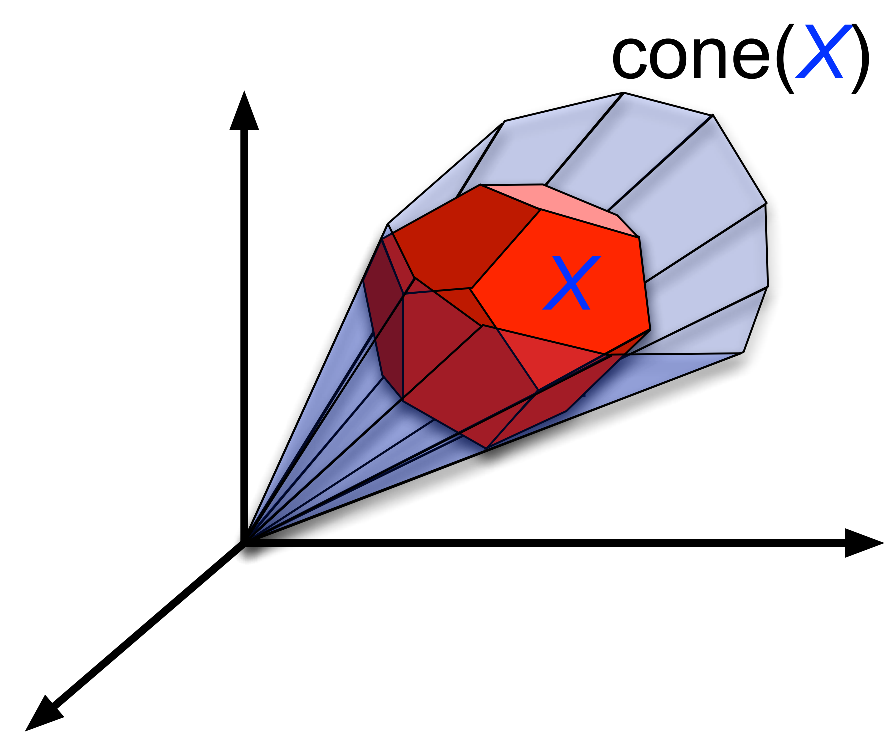
The *affine hull* of X, denoted by aff(X), is the set of all affine combinations of points in X.





The *convex hull* of X, denoted by conv(X), is the set of all convex combinations of points in X.

The *conic hull* of X, denoted by cone(X), is the set of all conic combinations of points in X.



Affine, Convex and Conic Hulls

Elementary Results:

1) The *affine hull* of X is the smallest affine set containing X (i.e., the intersection of all affine sets containing X).

2) The *convex hull* of X is the smallest convex set containing X. (i.e., the intersection of all convex sets containing X).

$$\mathrm{conv}(X) = \left\{ \sum_{i=1}^{k} \theta_i x_i : x_i \in X,\ \theta_i \geq 0\ \forall i,\ \sum_{i=1}^{k} \theta_i = 1 \right\}.$$

3) The *conic hull* of X is the smallest **convex** cone containing X (i.e., the intersection of all **convex** cones containing X).

Checking Convexity

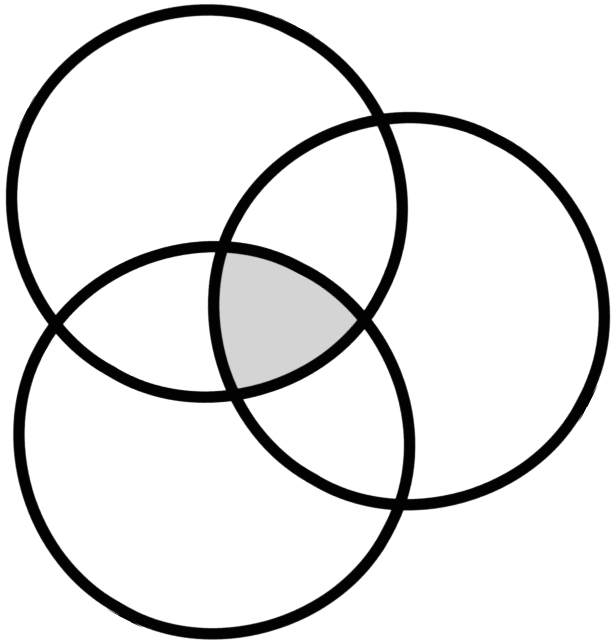
To check whether C is convex one can

- verify the *basic condition*

$$x_1, x_2 \in C,\ \theta \in [0, 1] \implies \theta x_1 + (1 - \theta) x_2 \in C$$

- or show that C is obtained from simple convex sets (balls, ellipsoids, halfspaces etc.) via *transformations that preserve convexity* such as
  - intersection
  - Minkowski sum
  - affine transformation
  - linear-fractional transformation

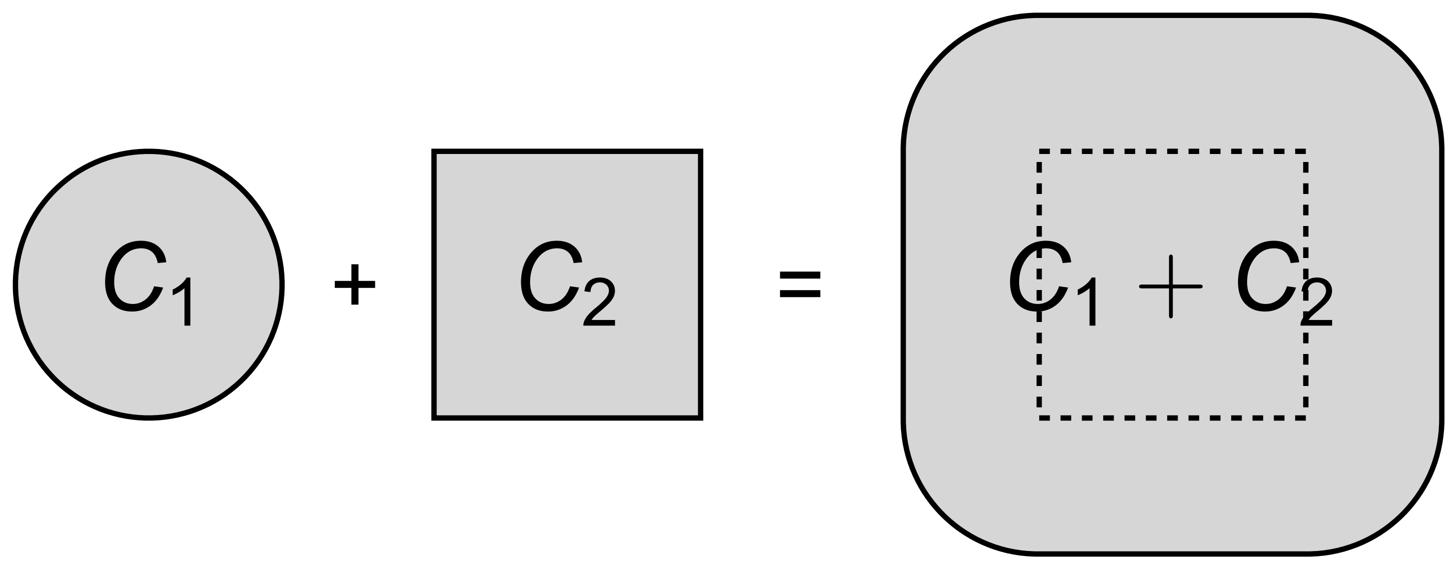
Intersection & Minkowski Sum



$C = \bigcap_{i \in \mathcal{I}} C_i$ is convex if C_i is convex $\forall i \in \mathcal{I}$

$$x_1, x_2 \in C_i \; \forall i \in \mathcal{I},\; \theta \in [0,1] \quad \Longrightarrow \quad \theta x_1 + (1-\theta) x_2 \in C_i \; \forall i \in \mathcal{I}$$
$$\Longrightarrow \quad \theta x_1 + (1-\theta) x_2 \in C$$

Minkowski sum:



C_1, C_2 convex $\quad \Longrightarrow \quad C_1 + C_2 = \{x_1 + x_2 : x_1 \in C_1, x_2 \in C_2\}$ convex

Example

The *positive semidefinite cone* is convex as it is an intersection of (infinitely many) convex halfspaces in $\mathbb{S}^n$.

$$\begin{aligned} \mathbb{S}^n_+ &= \{X \in \mathbb{S}^n : z^\top X z \geq 0 \ \forall z \in \mathbb{R}^n\} \\ &= \bigcap_{z \in \mathbb{R}^n} \{X \in \mathbb{S}^n : z^\top X z \geq 0\} \end{aligned}$$

Convexity-Preserving Transformations

Given transformation $f : \mathbb{R}^n \to \mathbb{R}^m$

Definitions:

The *image* of $C \subseteq \mathbb{R}^n$ is the set $f(C) = \{f(x) : x \in C\}$

The *preimage* of $C \subseteq \mathbb{R}^m$ is the set $f^{-1}(C) = \{x : f(x) \in C\}$

The image/preimage of a convex set C is convex if f is

- *affine* (e.g., rotation, translation, scaling, projection)

$$f(x) = Ax + b, \; A \in \mathbb{R}^{m \times n}, \; b \in \mathbb{R}^m$$

- *linear fractional*

$$f(x) = \frac{Ax + b}{c^\top x + d}, \; A \in \mathbb{R}^{m \times n}, \; b \in \mathbb{R}^m, \; c \in \mathbb{R}^n, \; d \in \mathbb{R}$$

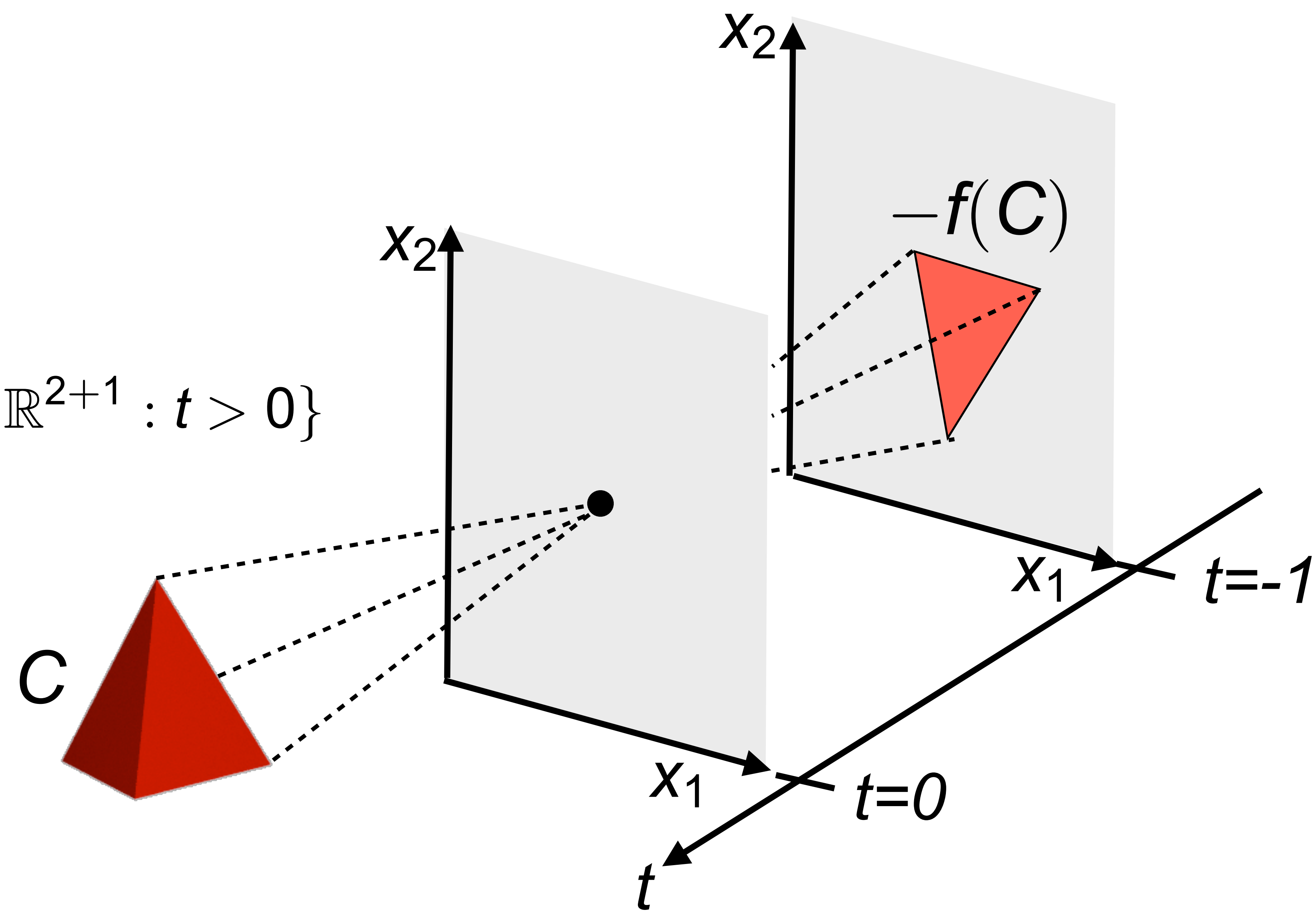
Here we require $c^\top x + d > 0 \; \forall x \in C$

Perspective Function

$$f : \mathbb{R}^{n+1} \to \mathbb{R}^n,\ f(x, t) = \frac{x}{t}$$

(special case of a linear fractional transformation)

$$C \subseteq \{(x, t) \in \mathbb{R}^{2+1} : t > 0\}$$



Proper Cones

Definition: A convex cone $K \subseteq \mathbb{R}^n$ is called *proper* if it is

- *closed* (i.e., it contains its boundary)
- *solid* (i.e., it contains a ball with positive radius)
- *pointed* (i.e., it contains **no** entire line)

Examples:

The *non-negative orthant*: $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n : x \geq 0\}$

The *second-order cone*: $\{(x, t) \in \mathbb{R}^n \times \mathbb{R} : \|x\|_2 \leq t\}$

The *positive semidefinite cone*: $\mathbb{S}^n_+ = \{X \in \mathbb{S}^n : X \succeq 0\}$

Counterexamples:

$\mathbb{R}^n_{++} \cup \{0\}$ does not contain its boundary

$\{tv : t \in \mathbb{R}_+\}$ for $v \in \mathbb{R}^n,\ v \neq 0,$ does not contain a ball

$\mathbb{R}^n$ is a convex cone that contains all lines through zero

Conic Inequalities

If $K \subseteq \mathbb{R}^n$ is a *proper convex cone*, then we can define generalized inequalities "$\preceq_K$", "$\succeq_K$", "$\prec_K$" and "$\succ_K$" on $\mathbb{R}^n$.

By definition, we have

$$x \preceq_K y \iff y - x \in K, \; x \prec_K y \iff y - x \in \mathrm{int}(K),$$
$$x \succeq_K y \iff x - y \in K, \; x \succ_K y \iff x - y \in \mathrm{int}(K).$$

Examples:

$K = \mathbb{R}^n_+$: $x \preceq_K y \iff x_i \le y_i \; \forall i = 1, \ldots, n$

$K = \{(x,t) : \|x\|_2 \le t\}$: $(x,t) \preceq_K (y,s) \iff \|y - x\|_2 \le s - t$

$K = \mathbb{S}^n_+$: $X \preceq_K Y \iff Y - X$ is positive semidefinite

Shorthand notation for linear and semidefinite inequalities:

$$x \preceq_{\mathbb{R}^n_+} y \rightarrow x \le y, \quad X \preceq_{\mathbb{S}^n_+} Y \rightarrow X \preceq Y$$

Main Take-Away Points

- **Definitions:** affine, convex and conic sets; affine, convex and conic combinations; affine, convex and conic hulls
- **Elementary convex sets:** hyperplanes, halfspaces, polyhedra, balls, ellipsoids
- **Elementary convex cones:** non-negative orthant, second-order cone, positive semidefinite cone
- **Convexity-preserving operations:** intersection; Minkowski-sum; images and pre-images of affine transformations, linear-fractional transformations (e.g., the perspective function)
- **Definitions:** proper cones; conic inequalities