ESC

2023.1.12.

Convex sets

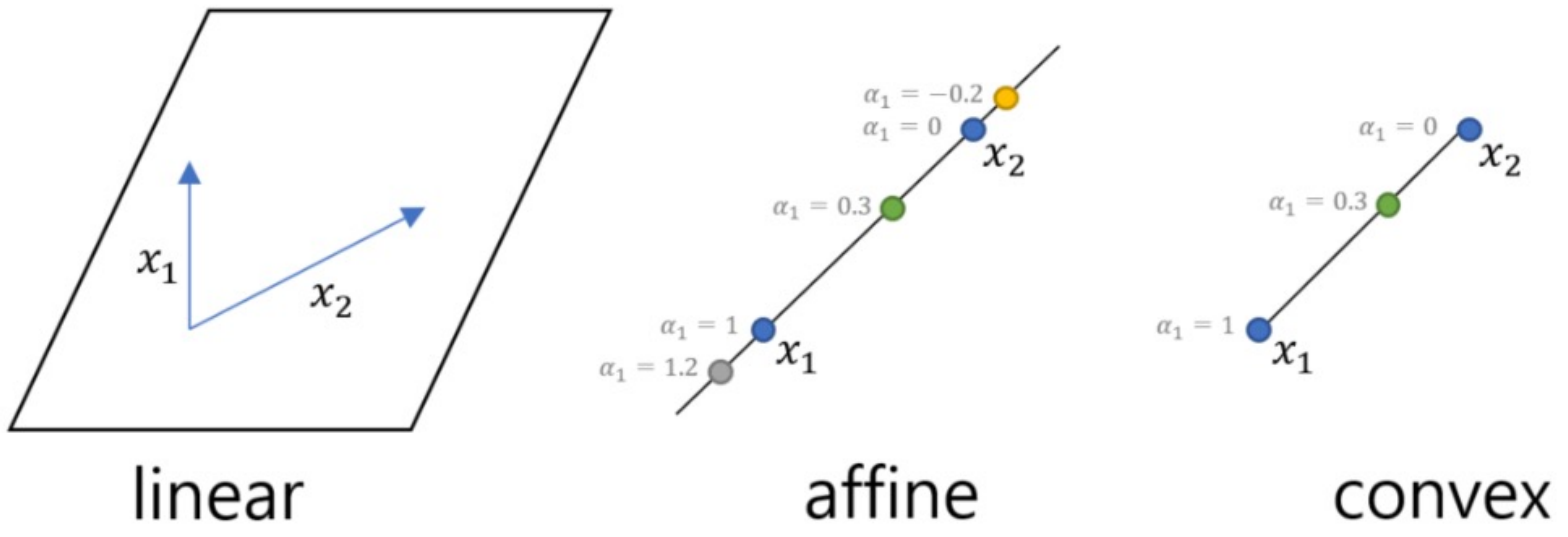
김두은

Affine sets and Convex sets

벡터 $x_1, x_2, \ldots, x_n$ 이 주어졌을 때, 이들을 결합하는 방법에는 3가지가 있다.

- Linear combination
- Affine combination
- Convex combination

$\theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_n x_n$, where [ ]



Affine sets

$\theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_n x_n$, where $\sum_{i=1}^{n} \boldsymbol{\theta_i} = \mathbf{1}$ ： affine combination

Affine set： A set which is closed under affine combination

Affine hull: $\mathbf{aff}\ \mathrm{C} = \{\theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_n x_n \mid x_1, x_2, \ldots \in C, \sum_{i=1}^{n} \theta_i = 1\}$

Affine sets: main topic

1. Affine sets are translated subspace. (example 2.1)
2. Relative interior **relint** $C = \{x \in C | B(x, r) \cap \textbf{aff } C \subseteq C\}$ (example 2.2)

Example 2.1 *Solution set of linear equations.* The solution set of a system of linear equations, $C = \{x \mid Ax = b\}$, where $A \in \mathbf{R}^{m \times n}$ and $b \in \mathbf{R}^m$, is an affine set. To show this, suppose $x_1, x_2 \in C$, *i.e.*, $Ax_1 = b$, $Ax_2 = b$. Then for any θ, we have

$$\begin{aligned} A(\theta x_1 + (1-\theta)x_2) &= \theta A x_1 + (1-\theta)Ax_2 \\ &= \theta b + (1-\theta)b \\ &= b, \end{aligned}$$

which shows that the affine combination $\theta x_1 + (1-\theta)x_2$ is also in C. The subspace associated with the affine set C is the nullspace of A.

We also have a converse: every affine set can be expressed as the solution set of a system of linear equations.

Example 2.2 Consider a square in the (x_1, x_2)-plane in $\mathbf{R}^3$, defined as

$$C = \{x \in \mathbf{R}^3 \mid -1 \leq x_1 \leq 1,\ -1 \leq x_2 \leq 1,\ x_3 = 0\}.$$

Its affine hull is the (x_1, x_2)-plane, *i.e.*, $\textbf{aff}\, C = \{x \in \mathbf{R}^3 \mid x_3 = 0\}$. The interior of C is empty, but the relative interior is

$$\textbf{relint}\, C = \{x \in \mathbf{R}^3 \mid -1 < x_1 < 1,\ -1 < x_2 < 1,\ x_3 = 0\}.$$

Its boundary (in $\mathbf{R}^3$) is itself; its relative boundary is the wire-frame outline,

$$\{x \in \mathbf{R}^3 \mid \max\{|x_1|, |x_2|\} = 1,\ x_3 = 0\}.$$

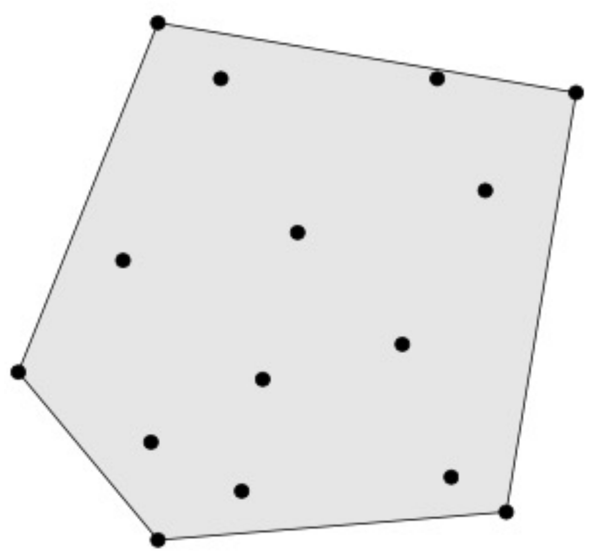
Convex sets

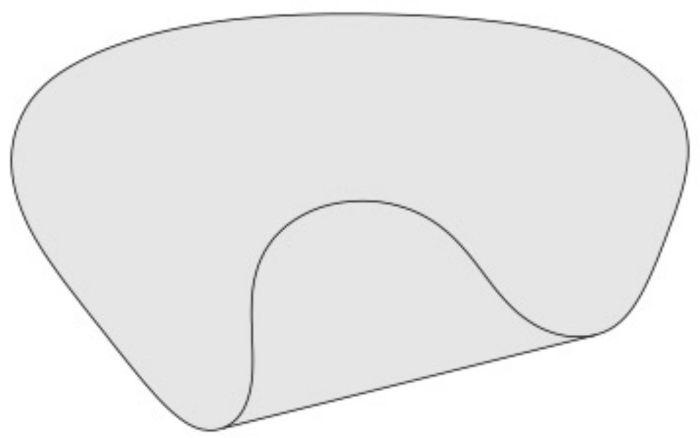
$\theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_n x_n$, where $\boldsymbol{\sum_{i=1}^{n} \theta_i = 1, \theta \geq 0}$: convex combination

Convex set: A set which is closed under convex combination

Convex hull: $\mathbf{conv}\ \mathrm{C} = \{\theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_n x_n \mid x_1, x_2, \ldots \in C, \sum_{i=1}^{n} \theta_i = 1, \theta \geq 0\}$

Suppose $\mathrm{C} \subseteq R^n$ is convex and x is a random vector with $x \in \mathrm{C}$ with probability 1. Then $Ex \in C$.



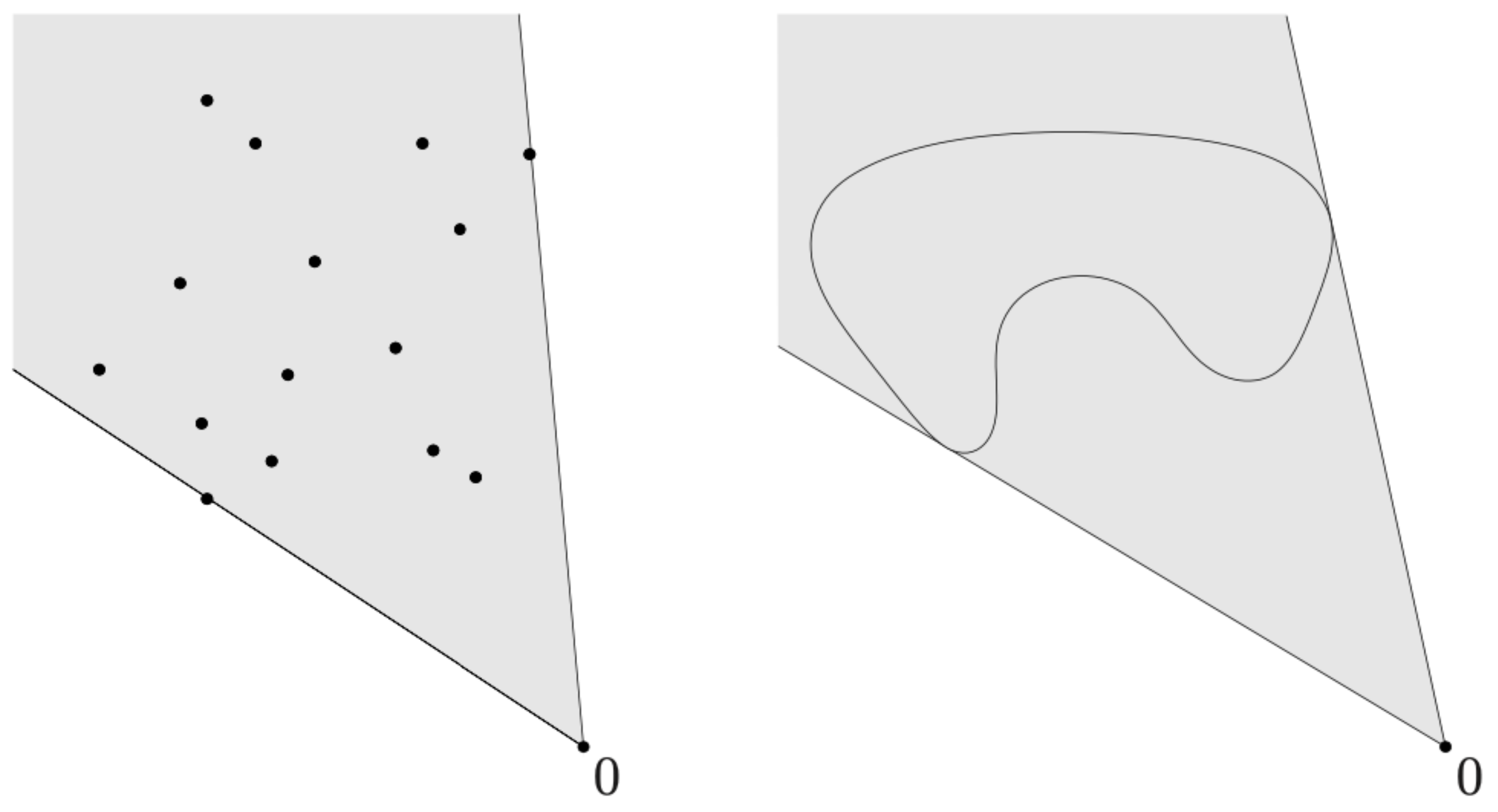


Cones

Cone: closed under positive multiplication

Convex cone: for any $x_1, x_2 \in C$ we have $\theta_1 x_1 + \theta_2 x_2 \in C$, where $\theta \geq 0$

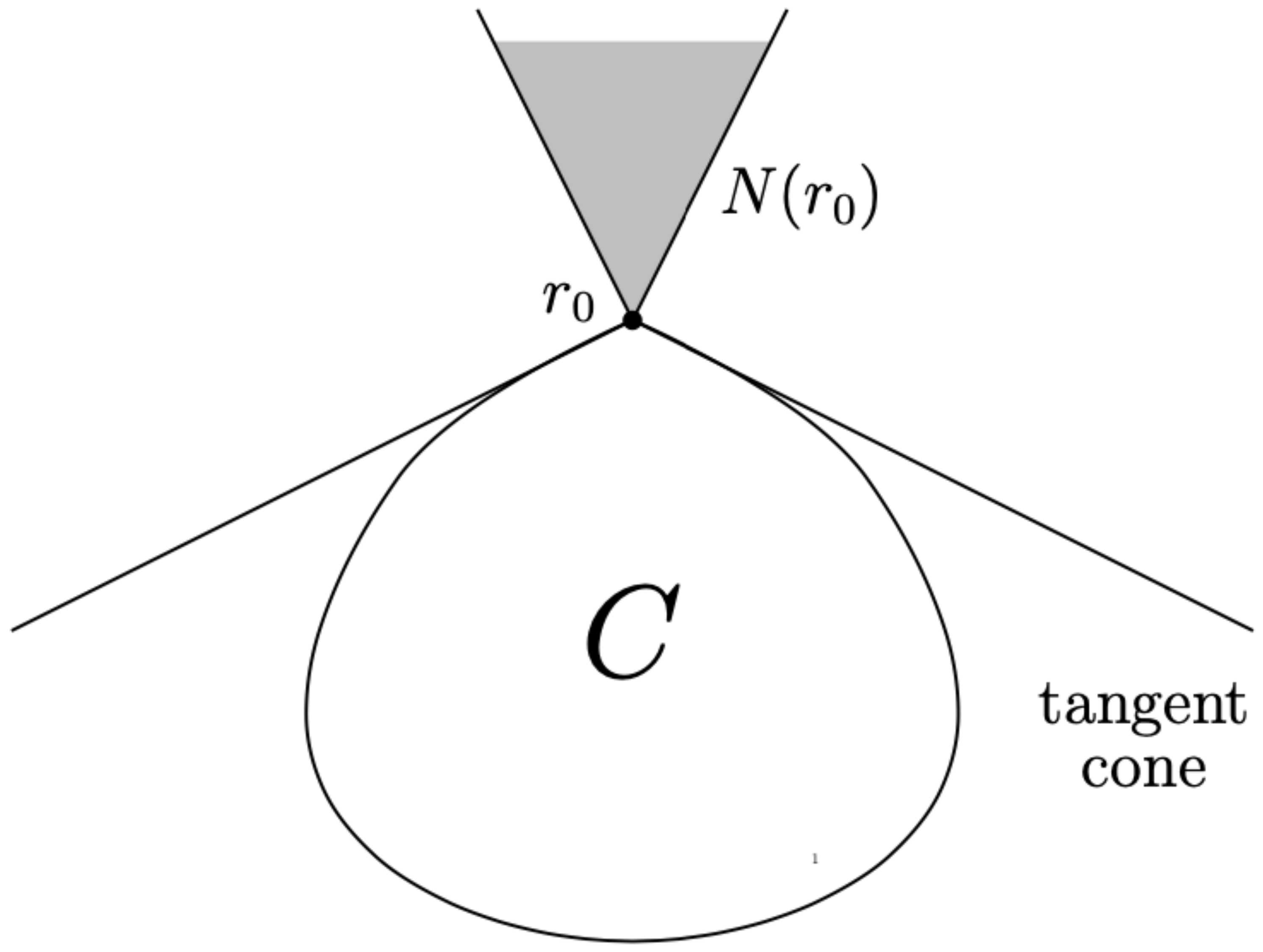
Conic hull: $\{\theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_n x_n \mid x_1, x_2, \ldots \in C, \theta \geq 0\}$



Cones

Tangent cone: x에서 아주 조금(α) 움직였을 때, $x + \alpha d \in$ C인 θd $(\theta \geq 0)$의 집합

Normal cone: $N_c(x) = \{g: g^T(y - x) \leq 0\ \forall y \in C\}$

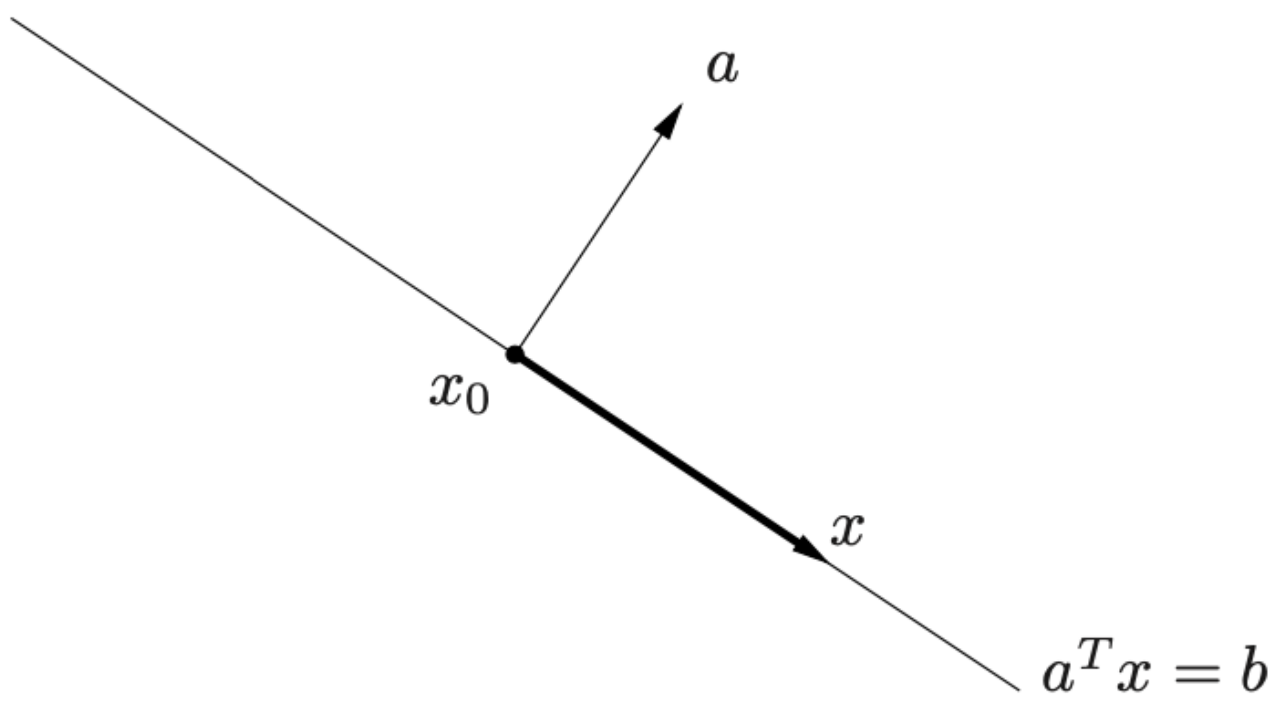


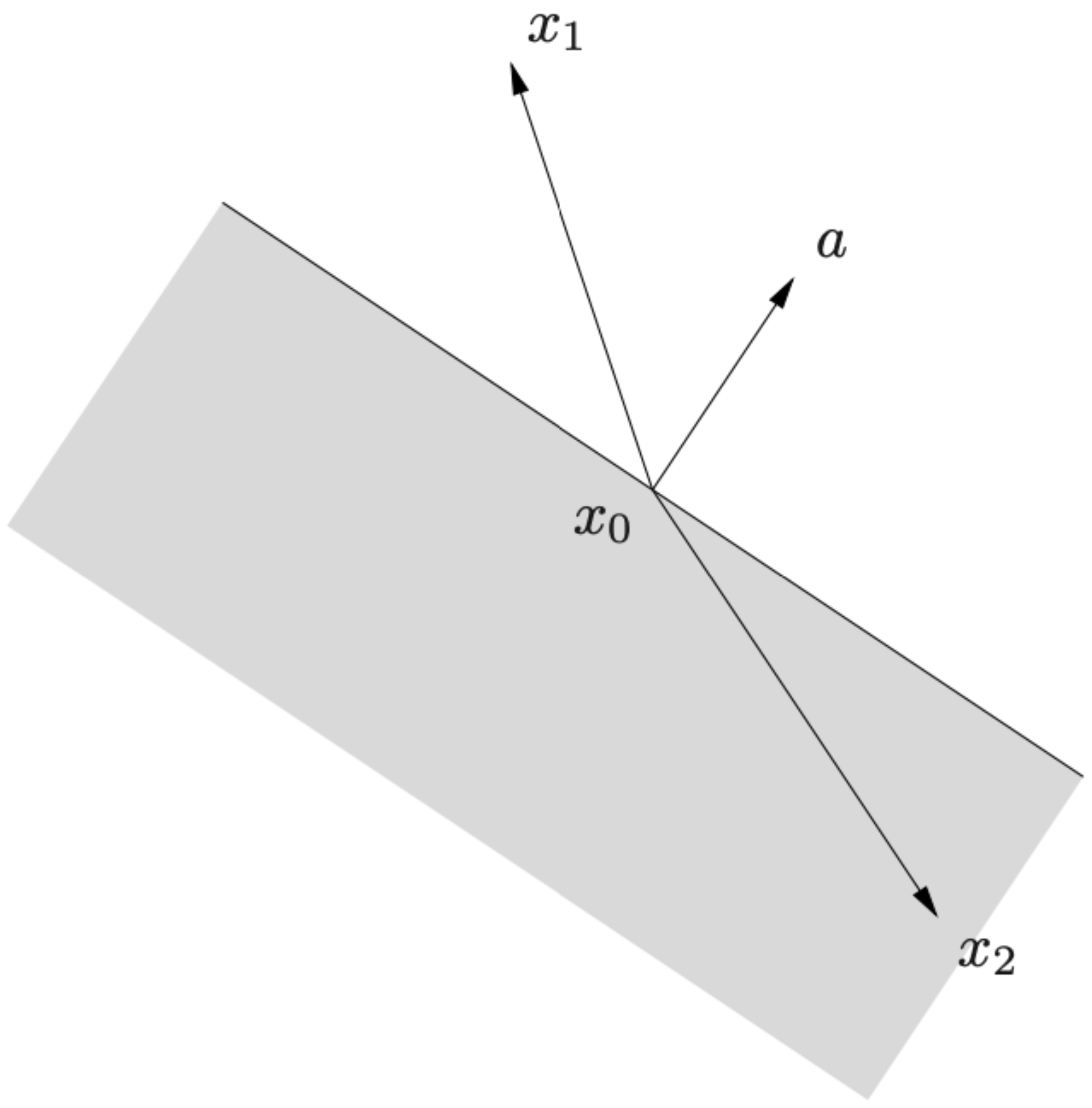
Important examples

Trivial ones: empty set, point, line

1. Hyperplanes and halfspaces

$$\{x|a^T x = b\} \rightarrow \{x|a^T(x - x_0) = 0\} \rightarrow x_0 + a^\perp$$





Important examples

2. Euclidian balls, ellipsoids, Norm balls and norm cones

$\{x_c + ru | \|u\|_2 \leq 1\}$

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

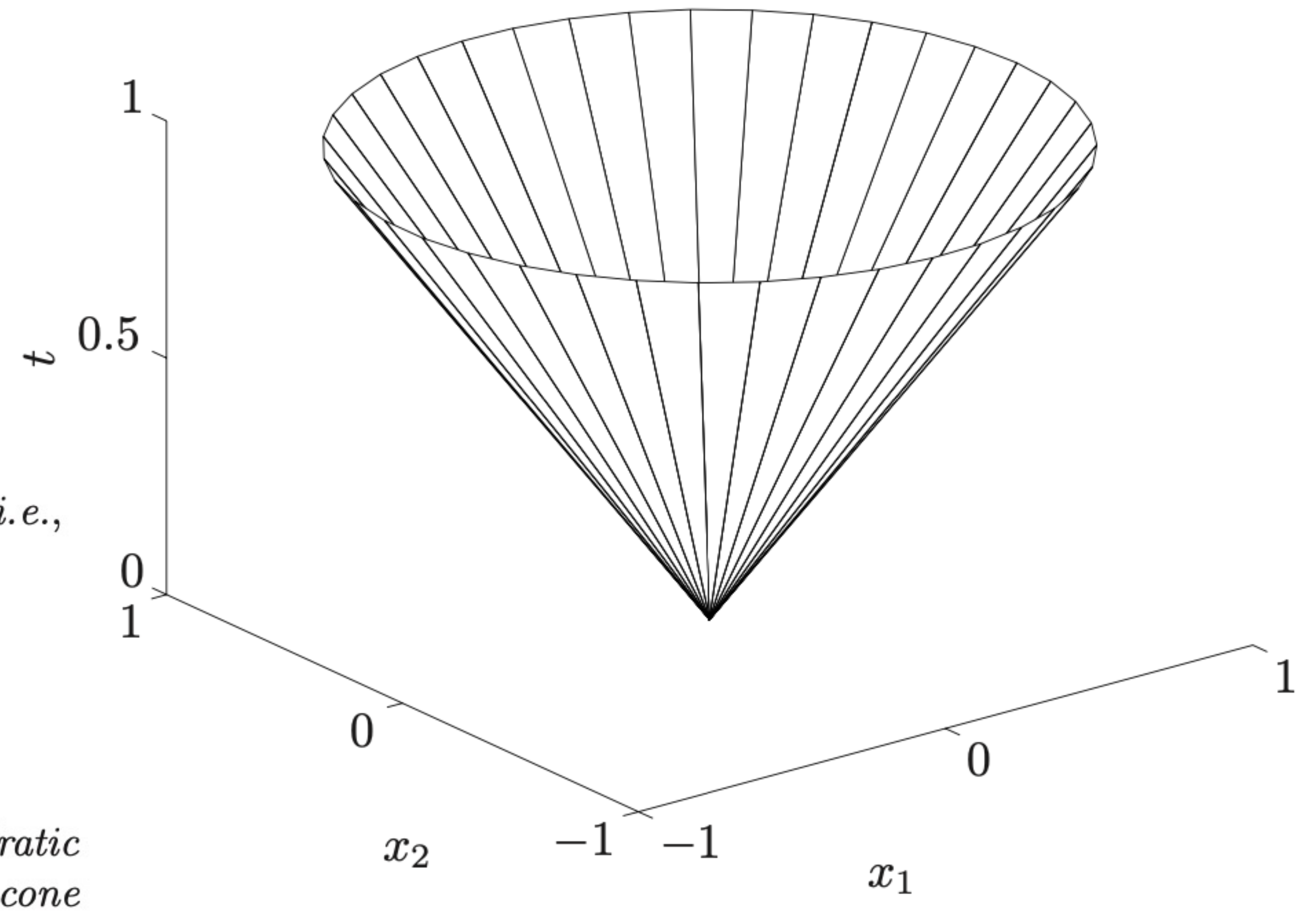
$\{x | \|x - x_c\| \leq r\}$

$\{(x, t) | \|x\| \leq t\} \subseteq R^{n-1}$

Example 2.3 *The second-order cone* is the norm cone for the Euclidean norm, *i.e.*,

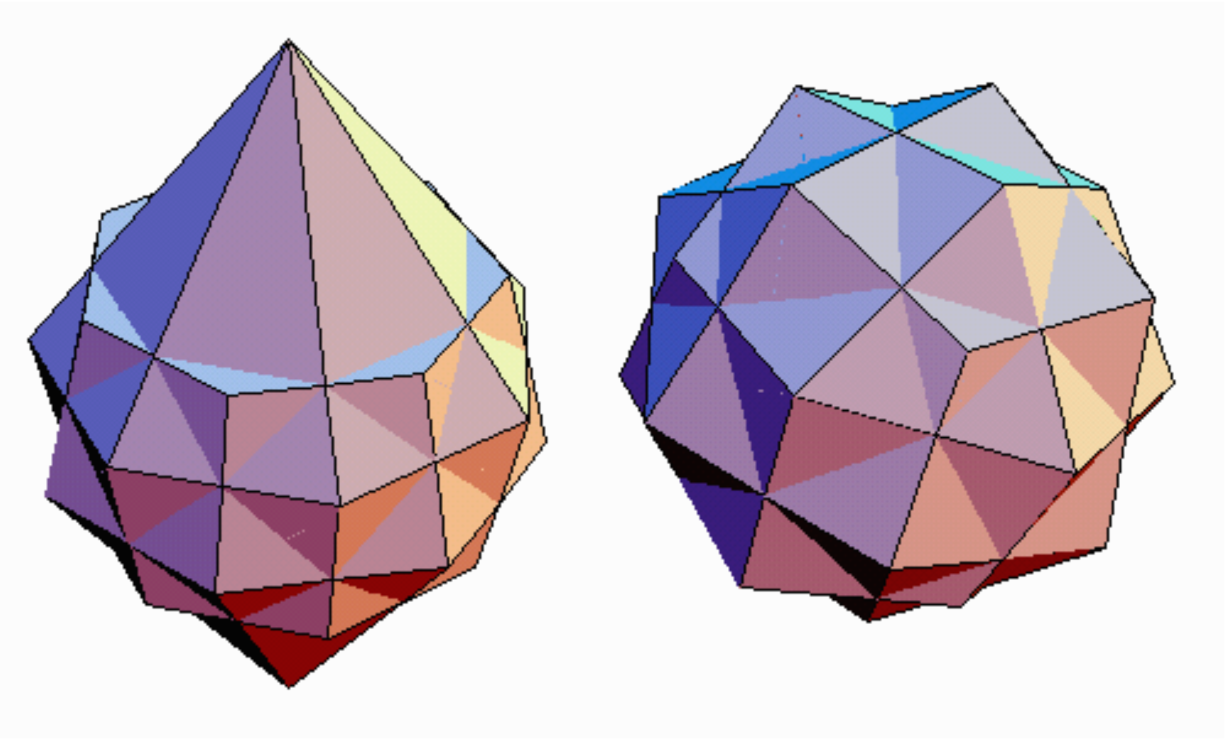
$$\begin{aligned} C &= \{(x,t) \in \mathbf{R}^{n+1} \mid \|x\|_2 \leq t\} \\ &= \left\{ \begin{bmatrix} x \\ t \end{bmatrix} \middle| \begin{bmatrix} x \\ t \end{bmatrix}^T \begin{bmatrix} I & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ t \end{bmatrix} \leq 0,\ t \geq 0 \right\}. \end{aligned}$$

The second-order cone is also known by several other names. It is called the *quadratic cone*, since it is defined by a quadratic inequality. It is also called the *Lorentz cone* or *ice-cream cone*. Figure 2.10 shows the second-order cone in $\mathbf{R}^3$.



Important examples

3. Polyhedra and simplexes



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

For affinely independent $v_0, v_1, \ldots, v_n \in R^n$, $\mathbf{conv}\{v_0, v_1, \ldots, v_n\}$

Important examples

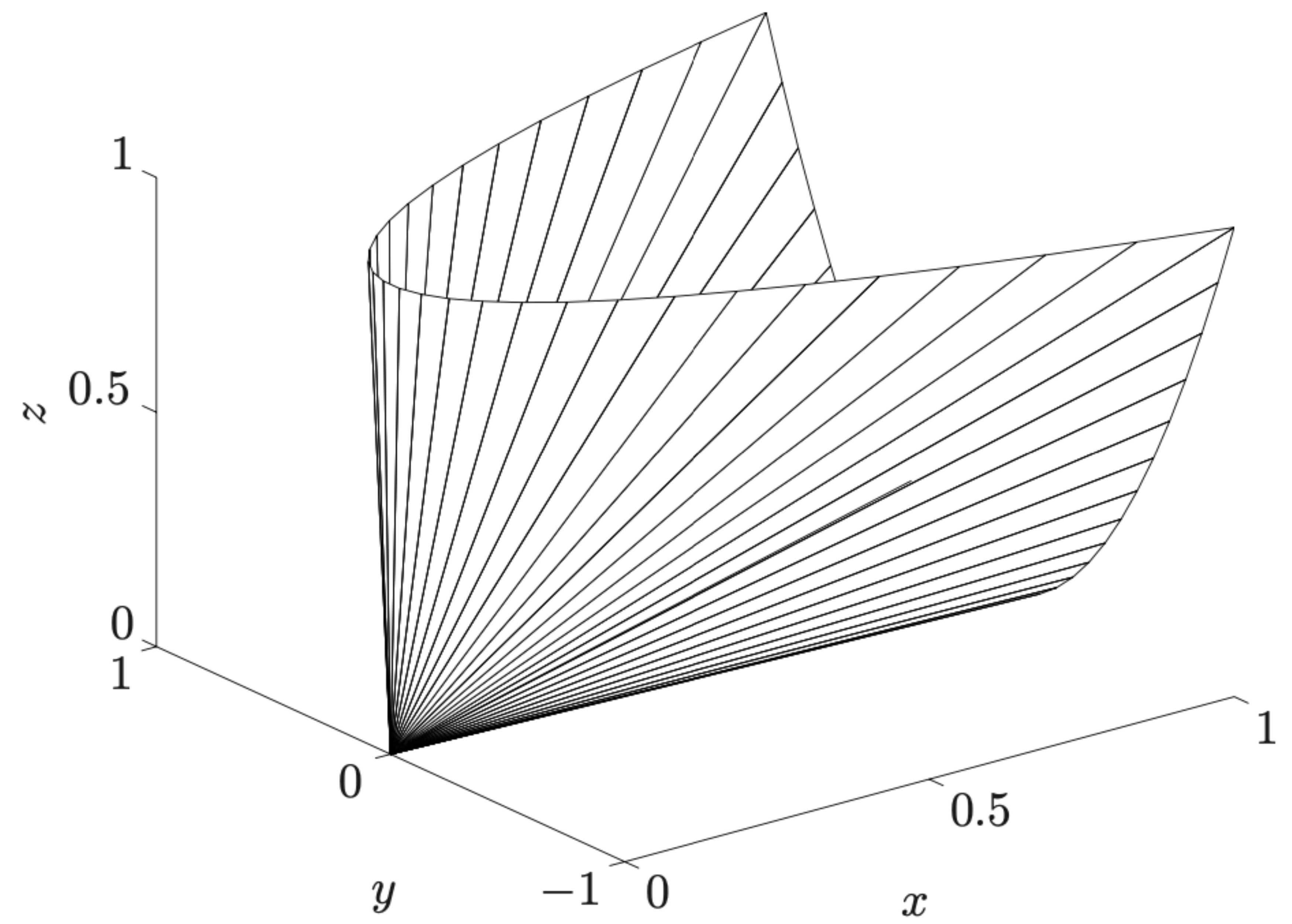
4. Positive semidefinite cones

$S_+^n \equiv$ set of symmetric positive semidefinite matrices

Example 2.6 *Positive semidefinite cone in* $\mathbf{S}^2$. We have

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

The boundary of this cone is shown in figure 2.12, plotted in $\mathbf{R}^3$ as (x, y, z).



Operations that preserve convexity

- Intersection
- Affine functions
- Perspective functions
- Linear–fractional functions

Intersection, affine functions

$S = \bigcap\{\mathcal{H} \mid \mathcal{H} \text{ halfspace}, S \subseteq \mathcal{H}\}$

Affine functions: $f(x) = Ax + b, A \in R^{m\times n}, \mathrm{b} \in R^m$ → linear + translation

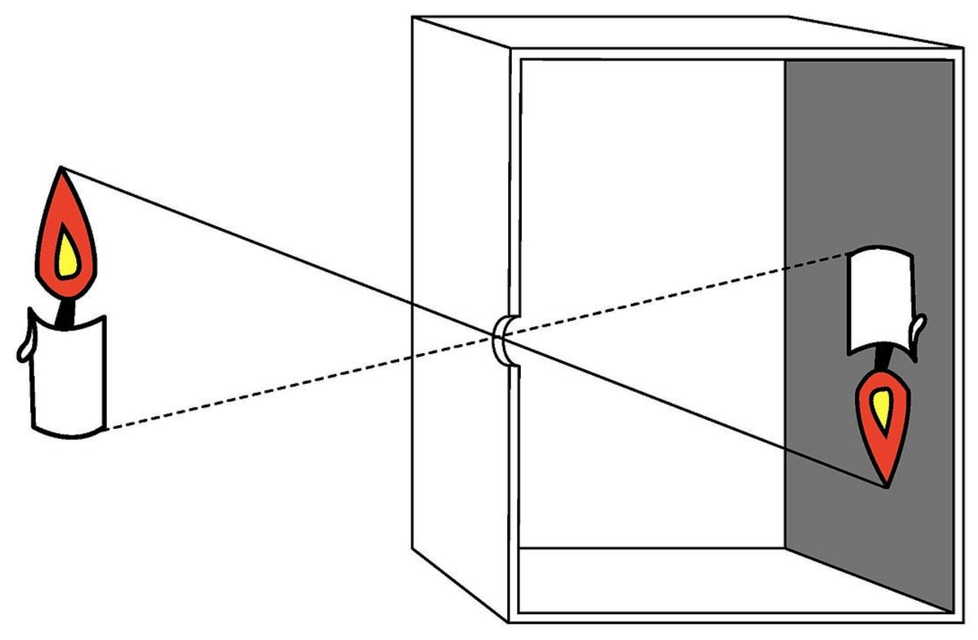
Suppose $\mathrm{C} \subseteq R^n, D \subseteq R^m$ is convex. Then affine image and affine preimage of f

$f(C) = \{f(x) | x \in C\}$, $f^{-1}(\mathrm{D}) = \{x | f(x) \in D\}$ are convex.

Ex) scaling, translation, projection, sum, partial sum

Perspective and linear−fractional functions

perspective function P: $R^n \times R_{++} \to R^n$: $P(z,t) = z/t$



Suppose g: $R^n \to R^{m+1}$ is affine:

$$g(x) = \begin{bmatrix} A \\ c^T \end{bmatrix} x + \begin{bmatrix} b \\ d \end{bmatrix}$$

Then f: $R^n \to R^m = \mathrm{P} \circ g$:

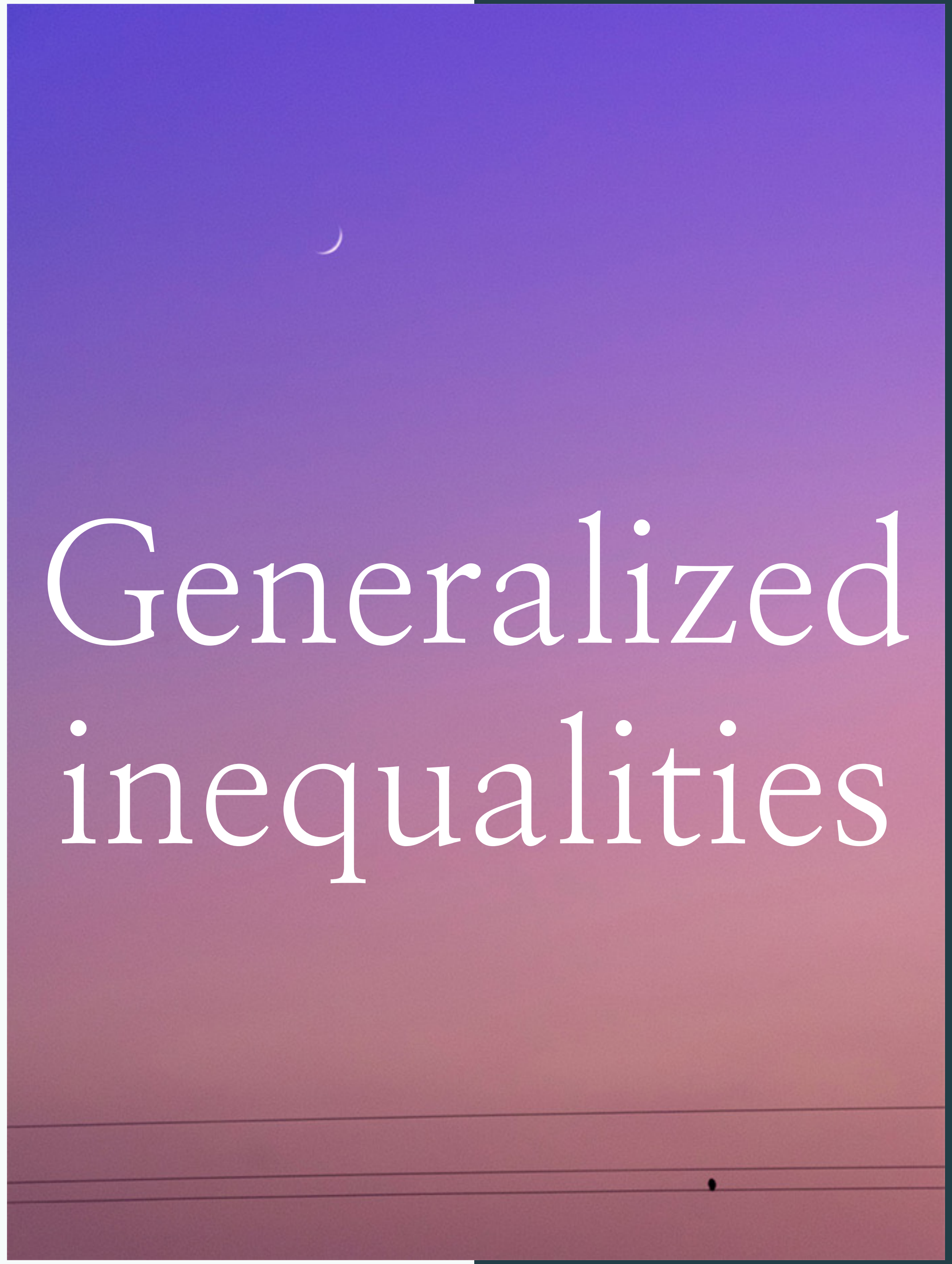
$$f(x) = \frac{Ax+b}{c^Tx+d}, \text{ where } \mathbf{dom}\, f = \{x \mid c^Tx + d > 0\}$$

Example 2.13 *Conditional probabilities.* Suppose u and v are random variables that take on values in $\{1,\dots,n\}$ and $\{1,\dots,m\}$, respectively, and let p_{ij} denote $\mathbf{prob}(u = i, v = j)$. Then the conditional probability $f_{ij} = \mathbf{prob}(u = i|v = j)$ is given by

$$f_{ij} = \frac{p_{ij}}{\sum_{k=1}^n p_{kj}}.$$

Thus f is obtained by a linear-fractional mapping from p.

It follows that if C is a convex set of joint probabilities for (u,v), then the associated set of conditional probabilities of u given v is also convex.



Generalized inequalities

Definition

Proper cone K:
cone $K \in R^n$이 다음 성질들을 만족하면 proper cone이라고 부른다.

1. k is convex.

2. k is closed. (경계를 포함하는 집합)

3. k is solid. (interior is not empty)

4. k is pointed (직선을 포함하지 않는다.)

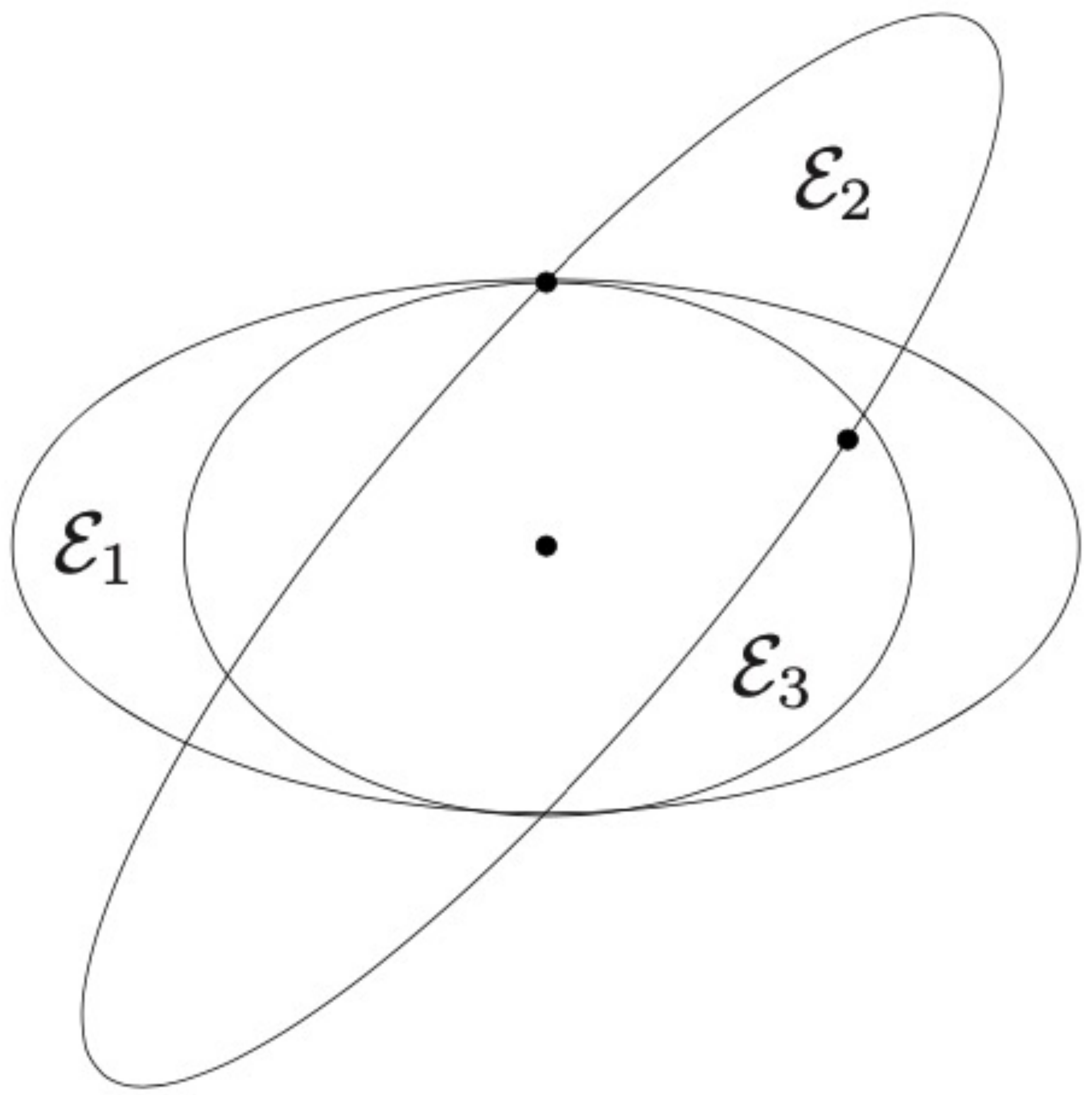
이 proper cone을 이용해 정의된 generalized inequality는 다음과 같다.

$$x \preceq_K y \iff y - x \in K, \; x \prec_K y \iff y - x \in \mathbf{int}\, K$$

Property, minimum vs minimal

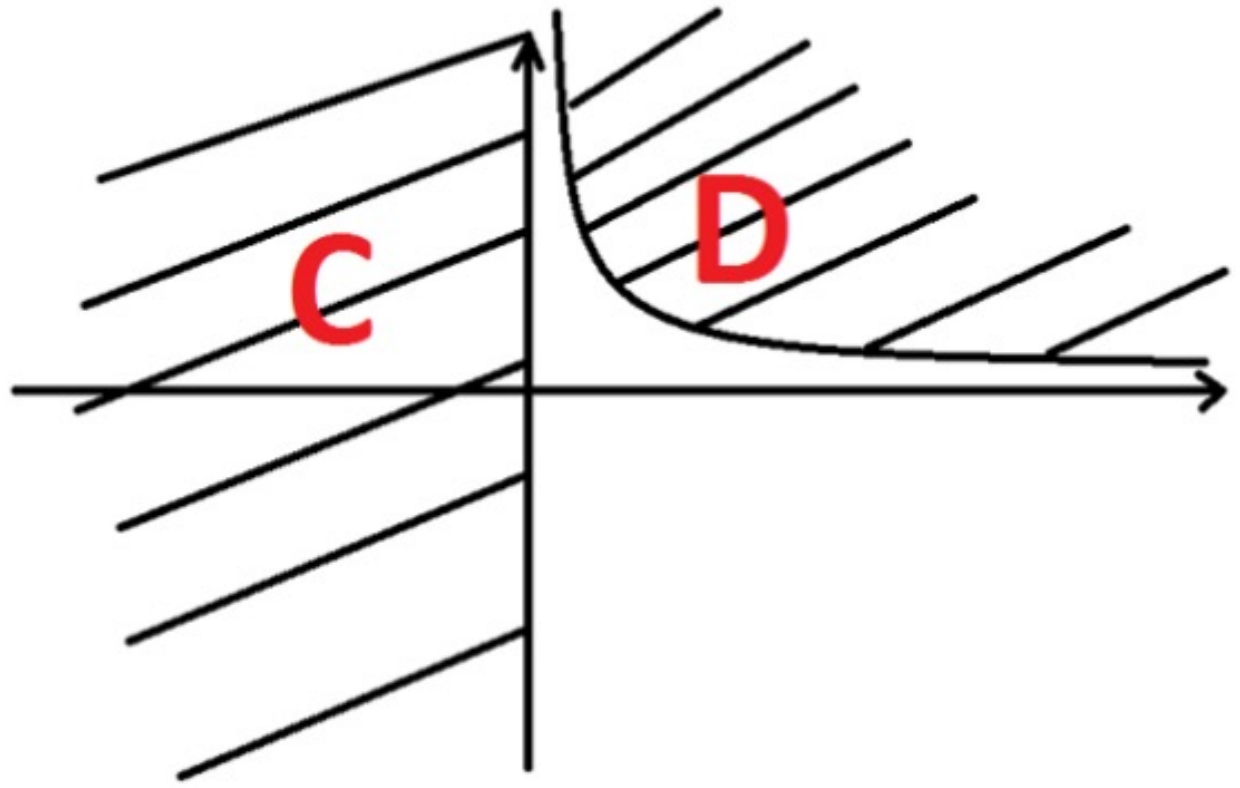
$$S \subseteq x + K \ vs. (x - K) \cap S = \{x\}$$

- preserved under addition: if $x \preceq_K y$ and $u \preceq_K v$, then $x + u \preceq_K y + v$.
- transitive: if $x \preceq_K y$ and $y \preceq_K z$ then $x \preceq_K z$.
- preserved under nonnegative scaling: if $x \preceq_K y$ and $\alpha \geq 0$ then $\alpha x \preceq_K \alpha y$.
- reflexive: $x \preceq_K x$.
- antisymmetric: if $x \preceq_K y$ and $y \preceq_K x$, then $x = y$.
- preserved under limits: if $x_i \preceq_K y_i$ for i = 1,2,..., $x_i \to x$ and $y_i \to y$ as $i \to \infty$, then $x \preceq_K y$.



Separating hyperplane theorem

어떤 두 conex set C와 D가 disjoint일 때, 그 두 집합은 어떤 hyperplane $a^T x = b$로 구분될 수 있다.
i.e. 하나는 한쪽 halfspace, 다른 것은 반대쪽에 속한다. (역은 성립하지 않을 수 있다.)



이때 등호가 없는 separation이 성립할 수도 있는데, (i.e. $C \in a^T x > b$, $D \in a^T x < b$)
이 그림처럼 항상 가능하지는 않다.

Proof of separating hyperplane theorem

Let C and D be closed convex sets in R^n with at least one of them bounded, and assume that $C \cap D = \emptyset$. Then $\exists a \in R^n$, $a \neq 0$, $b \in R$ $s.t.$

$a^T x > b, \ \forall x \in D$ and $a^T x < b, \ \forall x \in C$

Step 1. $dist(C, D) = \inf ||u - v||$ s.t. $u \in C, v \in D$

$$a = d - c, b = \frac{||d||^2 - ||c||^2}{2}.$$

Step 2. claim: $f(x) > 0, \ \forall x \in D$ and $f(x) < 0, \ \forall x \in C.$ $\rightarrow$ $f\left(\frac{c+d}{2}\right) = (d-c)^T\left(\frac{c+d}{2}\right) - \frac{||d||^2 - ||c||^2}{2} = 0.$

Step 3. For the sake of contradiction, assume that $(d-c)^T \bar{d} - \frac{||d||^2 - ||c||^2}{2} \leq 0.$

Proof of separating hyperplane theorem

Step 4.

Define $g(x) = ||x - c||^2$. We claim that $\bar{d} - d$ is a descent direction for g at d. Indeed,

$$\begin{aligned}
\nabla g^T(d)(\bar{d} - d) &= (2d - 2c)^T(\bar{d} - d) \\
&= 2(-||d||^2 + d^T\bar{d} - c^T\bar{d} + c^T d) \\
&= 2(-||d||^2 + (d - c)^T\bar{d} + c^T d) \\
&\leq 2\left(-||d||^2 + \frac{||d||^2 - ||c||^2}{2} + c^T d\right) \\
&= -||d||^2 - ||c||^2 + 2c^T d \\
&= -||d - c||^2 < 0
\end{aligned}$$

Step 5.

Hence $\exists \bar{a} > 0$ s.t. $\forall \alpha \in (0, \bar{\alpha})$

$$g(d + \alpha(d - \bar{d})) < g(d)$$

i.e.,

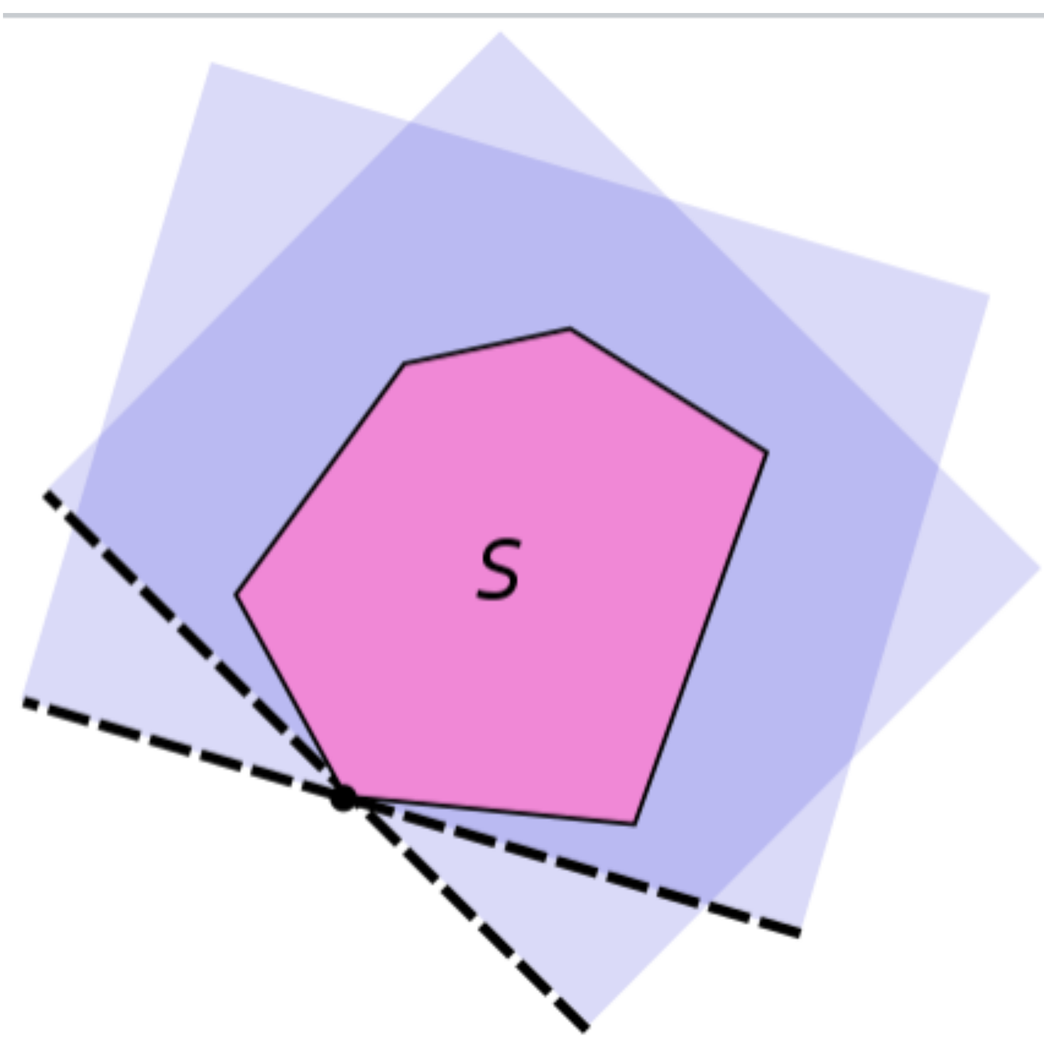
$$||d + \alpha(d - \bar{d}) - c||^2 < ||d - c||^2.$$

But this contradicts that d was the closest point to c. $\square$

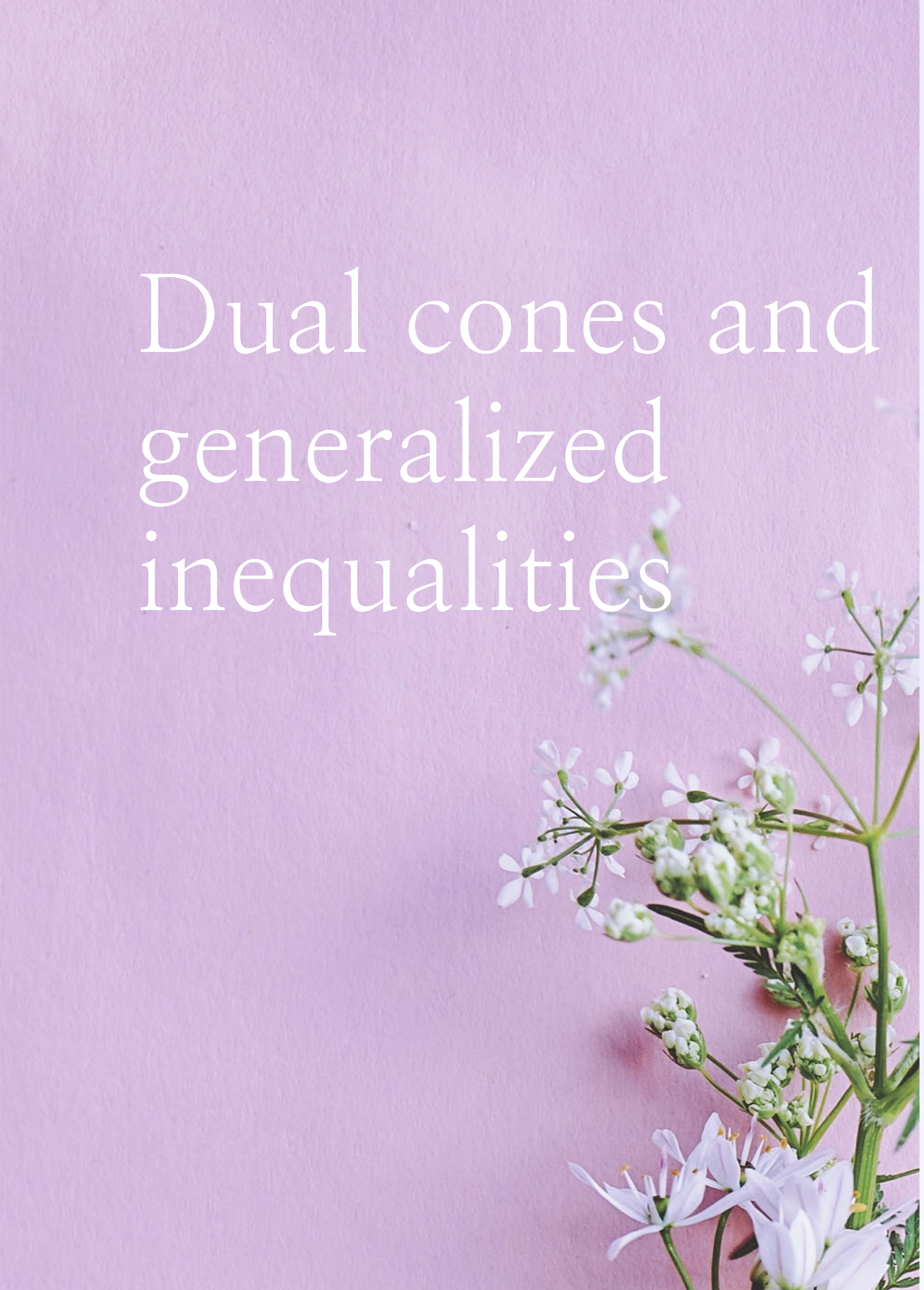
Supporting hyperplane theorem

어떤 집합이 supporting hyperplane을 갖는다는 것은 그 집합이 hyperplane을 통해 분리된 halfspace 중 하나에 온전히 속하며, hyperplane과 접하는 경계점이 적어도 하나 존재한다는 것이다.

convex set C의 경우 경계면의 점 x_0와 접하는 supporting hyperplane이 적어도 하나 존재한다.



반대로, 만약 C가 interior를 가지는 closed nonempty set이고 모든 경계면의 점이 그 점을 포함하는 supporting hyperplane을 가질 경우, C는 convex set이며 모든 supporting hyperplane의 교집합 역시 convex set이다.



Dual cones and generalized inequalities

1. Dual cones

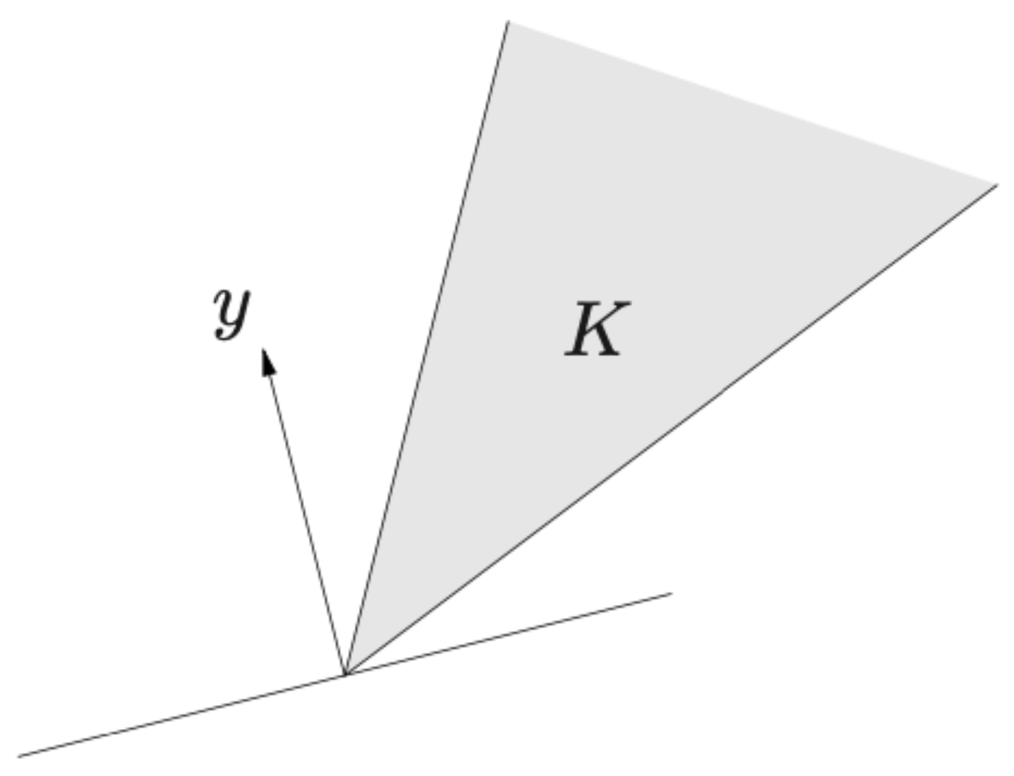
2. Dual generalized inequalities

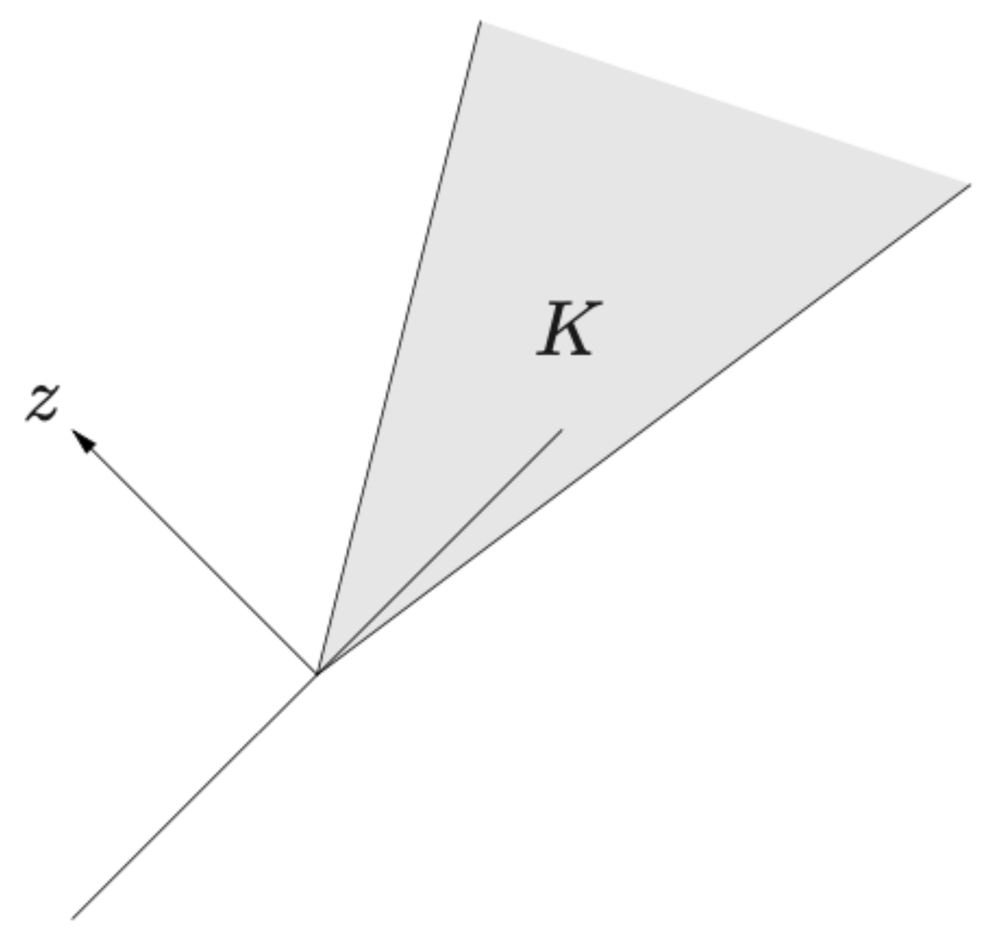
3. minimum and minimal elements via dual inequalities

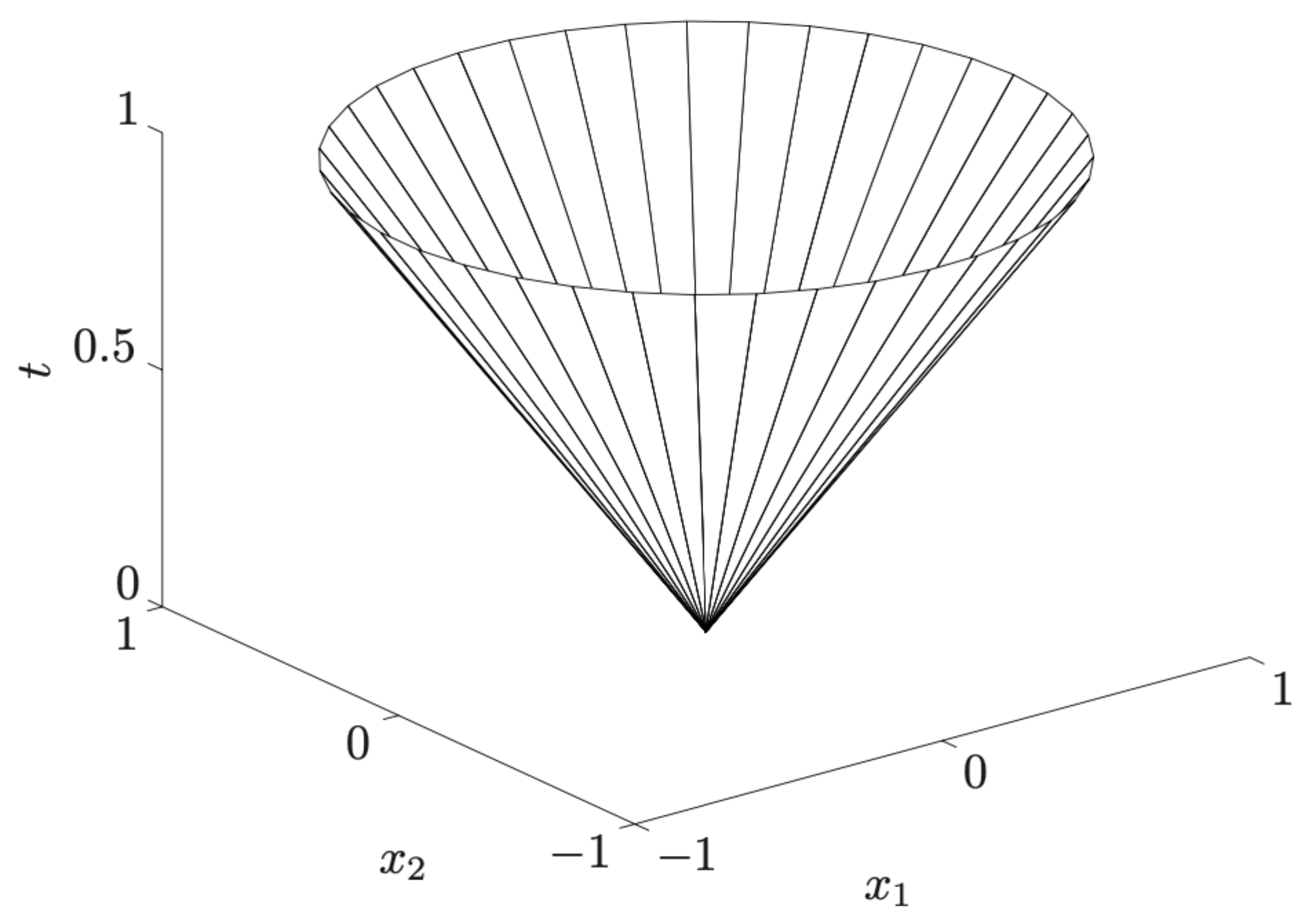
Dual cones

$K^* = \{y | y^T x \geq 0 \; for \; all \; x \in K\}$

Geometrically, $y \in K^* \; iff \; -y$ is the normal of a hyperplane that supports K at the origin







Dual generalized inequalities

$y \succcurlyeq_{K^*} 0 \leftrightarrow y^T x \geq 0 \; for \; all \; x \succcurlyeq_K 0$

- $x \preceq_K y$ if and only if $\lambda^T x \leq \lambda^T y$ for all $\lambda \succeq_{K^*} 0$.
- $x \prec_K y$ if and only if $\lambda^T x < \lambda^T y$ for all $\lambda \succeq_{K^*} 0$, $\lambda \neq 0$.

$K^{**} = K$

$\lambda \preceq_{K^*} \mu$ if and only if $\lambda^T x \leq \mu^T x$ for all $x \succeq_K 0$.