

2. Convex sets

- affine and convex sets 仿射集和凸集
- some important examples
- operations that preserve convexity 保留凸度的运算，如两个集合的交集
- generalized inequalities
- separating and supporting hyperplanes 分离和支撑超平面
- dual cones and generalized inequalities 对偶锥和对偶广义不等式

证明:

$$Ax_1 = b$$

$$Ax_2 = b$$

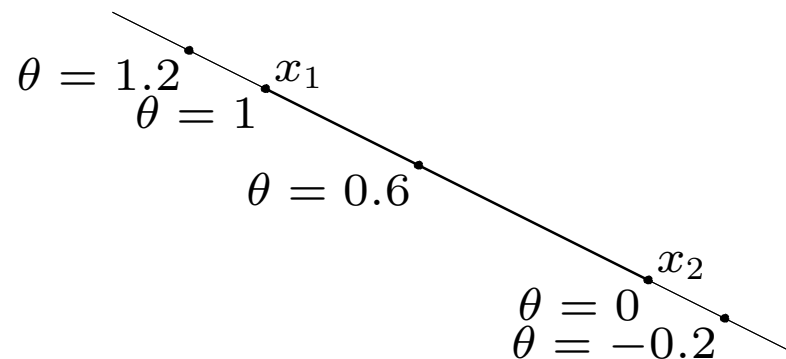
$$\begin{aligned}
Ax &= A(\theta x_1 + (1 - \theta)x_2) \\
&= \theta Ax_1 + (1 - \theta)Ax_2 \\
&= \theta b + (1 - \theta)b \\
&= b
\end{aligned}$$

Affine set

line through x_1, x_2 : all points

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

1. 线性组合
2. 两个线性系数之和为1



0 < θ < 1 时, 点处于 x_1 和 x_2 形成的线段上

仿射集可以直观理解为直线集

affine set: contains the line through any two distinct points in the set

example: solution set of linear equations $\{x \mid Ax = b\}$

(conversely, every affine set can be expressed as solution set of system of linear equations)

所以的仿射集都可以用线性方程组的解集来表示

Convex set

line segment between x_1 and x_2 : all points

$$x = \theta x_1 + (1 - \theta)x_2$$

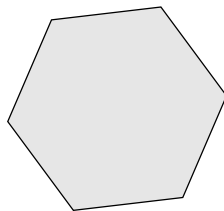
with $0 \leq \theta \leq 1$

凸集合：已知两点处于某一集合中，如果这两点的线性组合（即两点间的线段）也完全在该集合中，则该集合为凸集

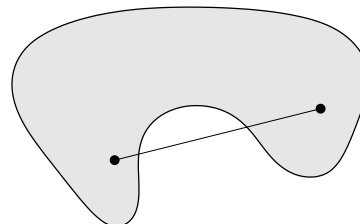
convex set: contains line segment between any two points in the set

$$x_1, x_2 \in C, \quad 0 \leq \theta \leq 1 \quad \implies \quad \theta x_1 + (1 - \theta)x_2 \in C$$

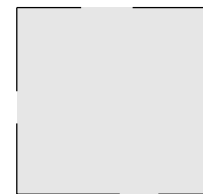
examples (one convex, two nonconvex sets)



是凸集



不是凸集



不是凸集

因为边界线有的地方不是实线，也就是不包含边界

Convex combination and convex hull

convex combination of x_1, \dots, x_k : any point x of the form

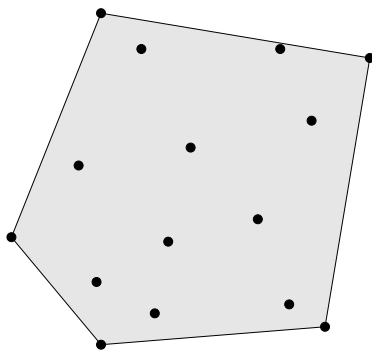
$$x = \theta_1 x_1 + \theta_2 x_2 + \dots + \theta_k x_k$$

with $\theta_1 + \dots + \theta_k = 1, \theta_i \geq 0$

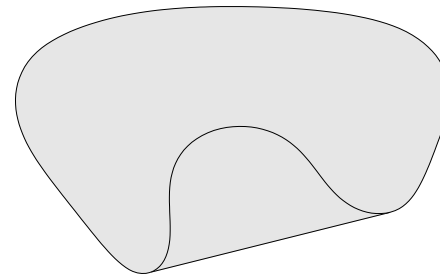
这种线性组合能够构成凸组合，当且仅当系数之和为1

convex hull $\text{conv } S$: set of all convex combinations of points in S

凸壳/凸包：是集合中所有的点的凸组合所构成的集合



图中的灰色区域就是凸包



这里划分出来的下面的区域，也是凸包

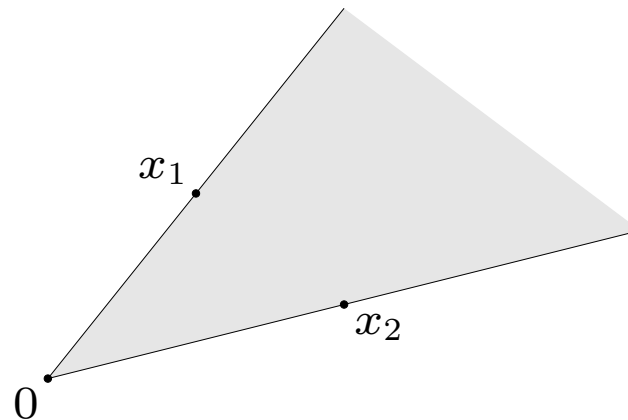
Convex cone

conic (nonnegative) combination of x_1 and x_2 : any point of the form

锥型组合：仅要求非负组合

$$x = \theta_1 x_1 + \theta_2 x_2$$

with $\theta_1 \geq 0, \theta_2 \geq 0$



锥型组合中，两个系数非负，且和为1时，属于凸锥组合

convex cone: set that contains all conic combinations of points in the set

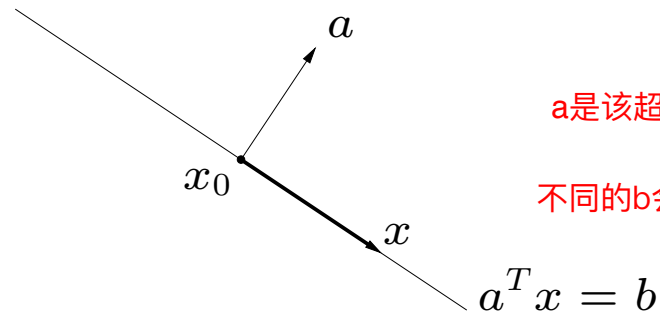
凸锥：包含集合中任意两点组成的锥型组合

Hyperplanes and halfspaces

hyperplane: set of the form $\{x \mid a^T x = b\}$ ($a \neq 0$) 在高维下， a ， x 和 b 都是向量或矩阵

超平面是单个线性方程约束的解集

在二维空间中是一条线
在三维空间中是一个平面

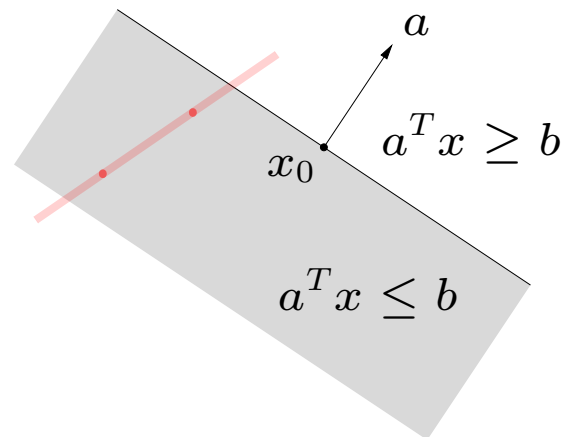


a 是该超表面的法向量

不同的 b 会得到不同且相互平行的超表面

halfspace: set of the form $\{x \mid a^T x \leq b\}$ ($a \neq 0$)

半空间是单个线性不等式约束的解集



改变 b 只是以平行的方式上下移动其表面

对于半空间而言， a 是向外的法线

- a is the normal vector
- hyperplanes are affine and convex; halfspaces are convex

超平面既是仿射的，也是凸的

半空间只是凸的，不是仿射的

数学证明：
是凸集

• 证明:

$$a^T x_1 \leq b$$

$$a^T x_2 \leq b$$

$$\begin{aligned} a^T x &= a^T (\theta x_1 + (1 - \theta)x_2) \\ &= \theta a^T x_1 + (1 - \theta)a^T x_2 \\ &\leq \theta b + (1 - \theta)b \\ &= b \end{aligned}$$

不是仿射集，因为

证明：如图中的两点，过该两点的直线并不全在半空间区域中，因此半空间并未包含所有的仿射点

Euclidean balls and ellipsoids

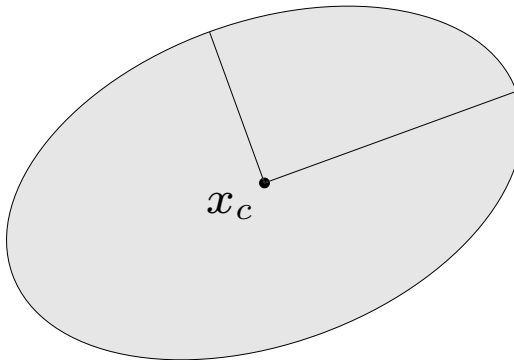
(Euclidean) ball with center x_c and radius r :

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

ellipsoid: set of the form

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

with $P \in \mathbf{S}_{++}^n$ (*i.e.*, P symmetric positive definite)



other representation: $\{x_c + Au \mid \|u\|_2 \leq 1\}$ with A square and nonsingular

Norm balls and norm cones

norm: a function $\| \cdot \|$ that satisfies

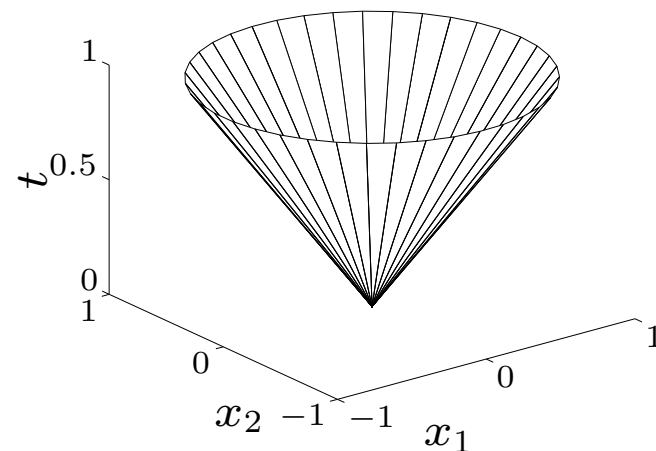
- $\|x\| \geq 0$; $\|x\| = 0$ if and only if $x = 0$
- $\|tx\| = |t| \|x\|$ for $t \in \mathbf{R}$
- $\|x + y\| \leq \|x\| + \|y\|$

notation: $\| \cdot \|$ is general (unspecified) norm; $\| \cdot \|_{\text{symb}}$ 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



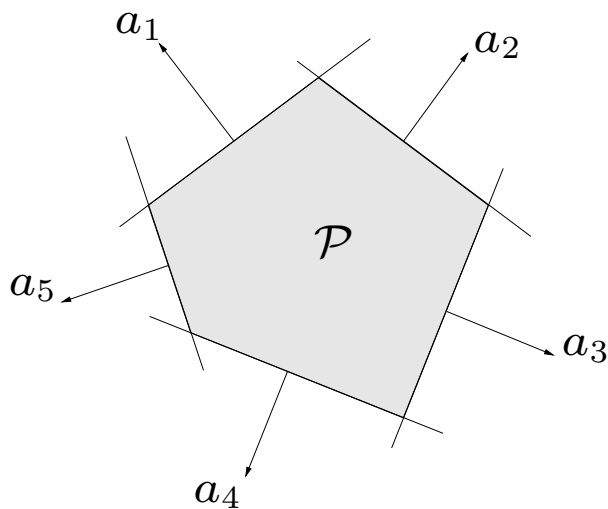
norm balls and cones are convex

Polyhedra

solution set of finitely many linear inequalities and equalities

$$Ax \preceq b, \quad Cx = d$$

($A \in \mathbf{R}^{m \times n}$, $C \in \mathbf{R}^{p \times n}$, \preceq is componentwise inequality)



polyhedron is intersection of finite number of halfspaces and hyperplanes

Positive semidefinite cone

notation:

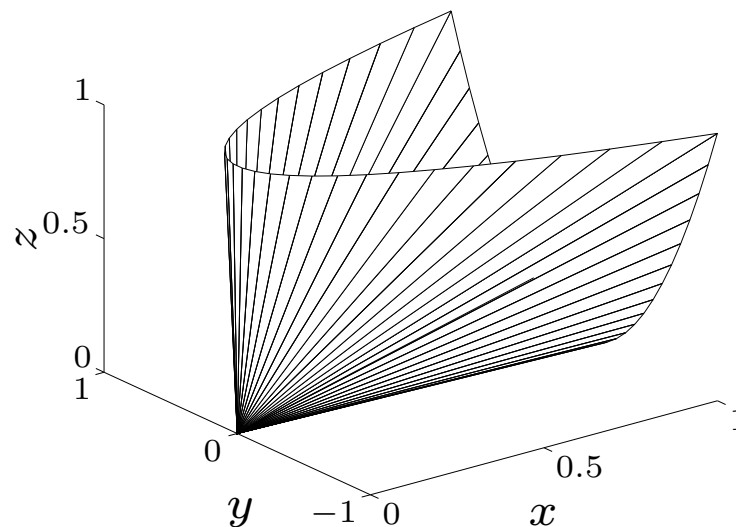
- \mathbf{S}^n is set of symmetric $n \times n$ matrices
- $\mathbf{S}_+^n = \{X \in \mathbf{S}^n \mid X \succeq 0\}$: positive semidefinite $n \times n$ matrices

$$X \in \mathbf{S}_+^n \iff z^T X z \geq 0 \text{ for all } z$$

\mathbf{S}_+^n is a convex cone

- $\mathbf{S}_{++}^n = \{X \in \mathbf{S}^n \mid X \succ 0\}$: positive definite $n \times n$ matrices

example: $\begin{bmatrix} x & y \\ y & z \end{bmatrix} \in \mathbf{S}_+^2$



Operations that preserve convexity

practical methods for establishing convexity of a set C

1. apply definition

$$x_1, x_2 \in C, \quad 0 \leq \theta \leq 1 \quad \implies \quad \theta x_1 + (1 - \theta)x_2 \in C$$

2. show that C is obtained from simple convex sets (hyperplanes, halfspaces, norm balls, . . .) by operations that preserve convexity

- intersection
- affine functions
- perspective function
- linear-fractional functions

Intersection

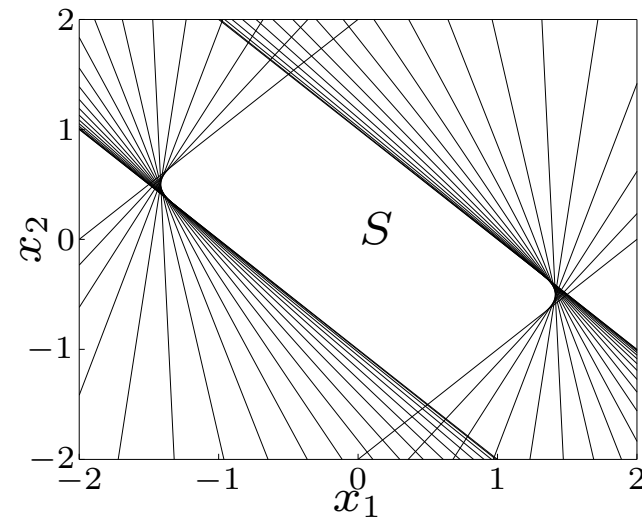
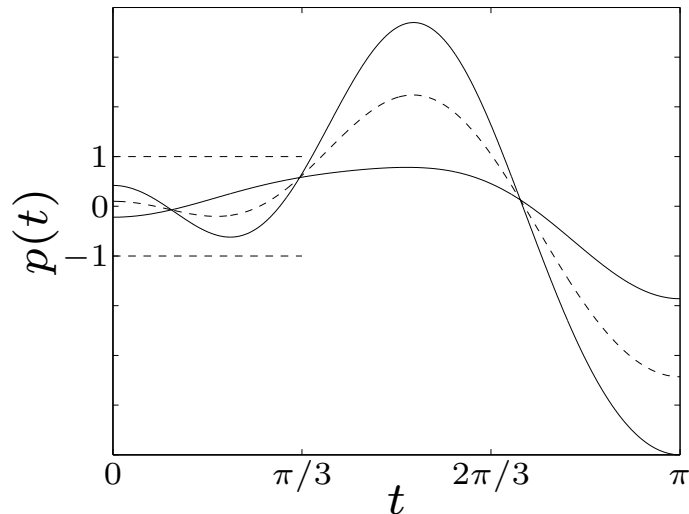
the intersection of (any number of) convex sets is convex

example:

$$S = \{x \in \mathbf{R}^m \mid |p(t)| \leq 1 \text{ for } |t| \leq \pi/3\}$$

where $p(t) = x_1 \cos t + x_2 \cos 2t + \cdots + x_m \cos mt$

for $m = 2$:



Affine function

suppose $f : \mathbf{R}^n \rightarrow \mathbf{R}^m$ is affine ($f(x) = Ax + b$ with $A \in \mathbf{R}^{m \times n}$, $b \in \mathbf{R}^m$)

- the image of a convex set under f is convex

$$S \subseteq \mathbf{R}^n \text{ convex} \implies f(S) = \{f(x) \mid x \in S\} \text{ convex}$$

- the inverse image $f^{-1}(C)$ of a convex set under f is convex

$$C \subseteq \mathbf{R}^m \text{ convex} \implies f^{-1}(C) = \{x \in \mathbf{R}^n \mid f(x) \in C\} \text{ convex}$$

examples

- scaling, translation, projection
- solution set of linear matrix inequality $\{x \mid x_1 A_1 + \cdots + x_m A_m \preceq B\}$
(with $A_i, B \in \mathbf{S}^p$)
- hyperbolic cone $\{x \mid x^T P x \leq (c^T x)^2, c^T x \geq 0\}$ (with $P \in \mathbf{S}_+^n$)

Perspective and linear-fractional function

perspective function $P : \mathbf{R}^{n+1} \rightarrow \mathbf{R}^n$:

$$P(x, t) = x/t, \quad \text{dom } P = \{(x, t) \mid t > 0\}$$

images and inverse images of convex sets under perspective are convex

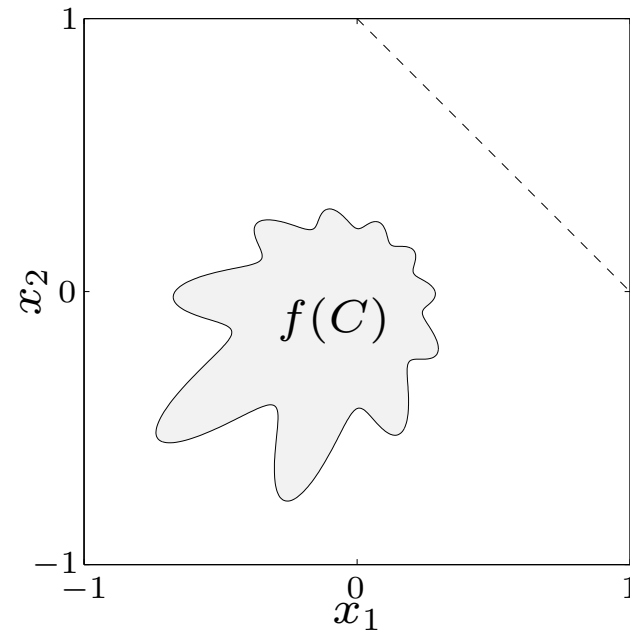
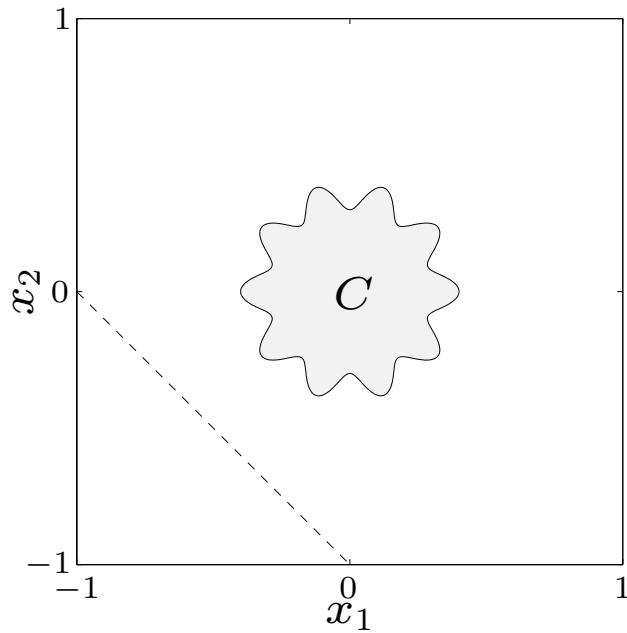
linear-fractional function $f : \mathbf{R}^n \rightarrow \mathbf{R}^m$:

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

images and inverse images of convex sets under linear-fractional functions are convex

example of a linear-fractional function

$$f(x) = \frac{1}{x_1 + x_2 + 1}x$$



Generalized inequalities

a convex cone $K \subseteq \mathbf{R}^n$ is a **proper cone** if

- K is closed (contains its boundary)
- K is solid (has nonempty interior)
- K is pointed (contains no line)

examples

- nonnegative orthant $K = \mathbf{R}_+^n = \{x \in \mathbf{R}^n \mid x_i \geq 0, i = 1, \dots, n\}$
- positive semidefinite cone $K = \mathbf{S}_+^n$
- nonnegative polynomials on $[0, 1]$:

$$K = \{x \in \mathbf{R}^n \mid x_1 + x_2 t + x_3 t^2 + \dots + x_n t^{n-1} \geq 0 \text{ for } t \in [0, 1]\}$$

generalized inequality defined by a proper cone K :

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

examples

- componentwise inequality ($K = \mathbf{R}_+^n$)

$$x \preceq_{\mathbf{R}_+^n} y \iff x_i \leq y_i, \quad i = 1, \dots, n$$

- matrix inequality ($K = \mathbf{S}_+^n$)

$$X \preceq_{\mathbf{S}_+^n} Y \iff Y - X \text{ positive semidefinite}$$

these two types are so common that we drop the subscript in \preceq_K

properties: many properties of \preceq_K are similar to \leq on \mathbf{R} , *e.g.*,

$$x \preceq_K y, \quad u \preceq_K v \implies x + u \preceq_K y + v$$

Minimum and minimal elements

\preceq_K is not in general a *linear ordering*: we can have $x \not\preceq_K y$ and $y \not\preceq_K x$

$x \in S$ is **the minimum element** of S with respect to \preceq_K if

$$y \in S \implies x \preceq_K y$$

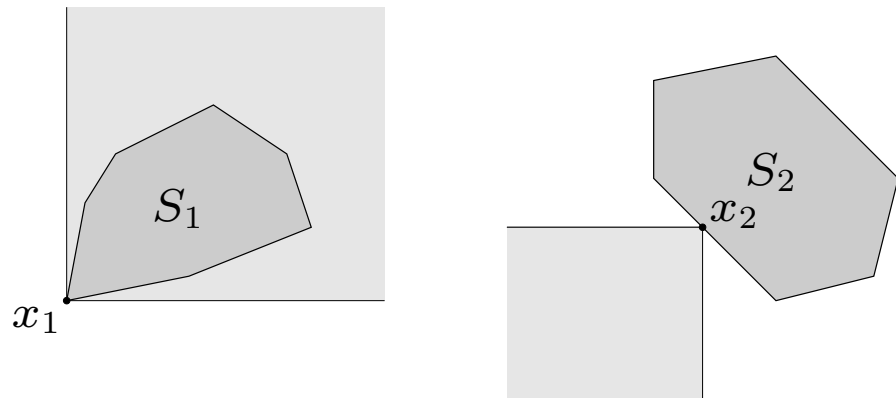
$x \in S$ is **a minimal element** of S with respect to \preceq_K if

$$y \in S, \quad y \preceq_K x \implies y = x$$

example ($K = \mathbf{R}_+^2$)

x_1 is the minimum element of S_1

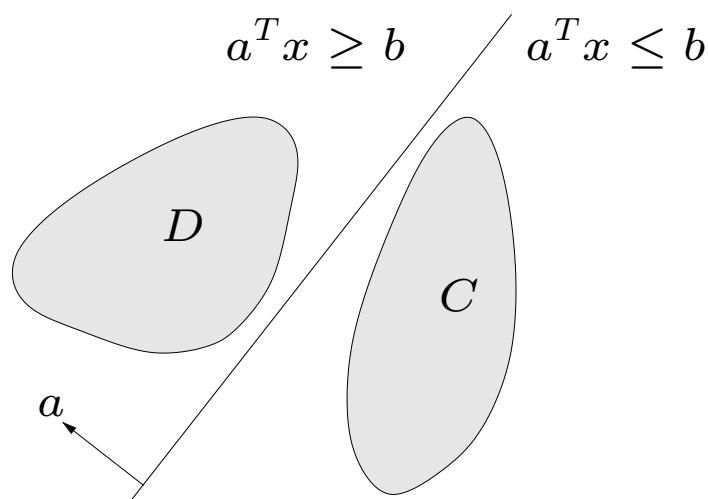
x_2 is a minimal element of S_2



Separating hyperplane theorem

if C and D are nonempty disjoint convex sets, there exist $a \neq 0$, b s.t.

$$a^T x \leq b \text{ for } x \in C, \quad a^T x \geq b \text{ for } x \in D$$



the hyperplane $\{x \mid a^T x = b\}$ separates C and D

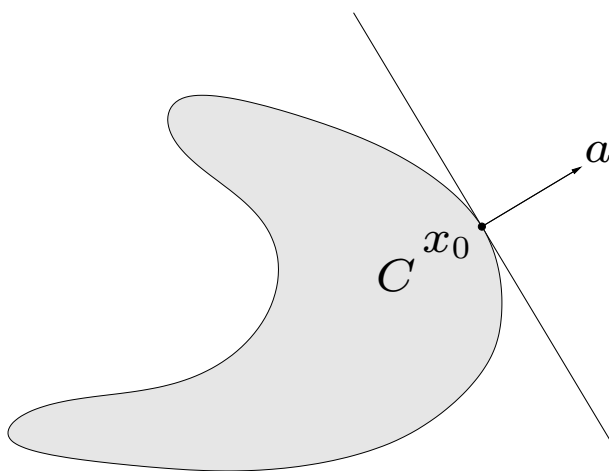
strict separation requires additional assumptions (*e.g.*, C is closed, D is a singleton)

Supporting hyperplane theorem

supporting hyperplane to set C at boundary point x_0 :

$$\{x \mid a^T x = a^T x_0\}$$

where $a \neq 0$ and $a^T x \leq a^T x_0$ for all $x \in C$



supporting hyperplane theorem: if C is convex, then there exists a supporting hyperplane at every boundary point of C

Dual cones and generalized inequalities

dual cone of a cone K :

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

examples

- $K = \mathbf{R}_+^n$: $K^* = \mathbf{R}_+^n$
- $K = \mathbf{S}_+^n$: $K^* = \mathbf{S}_+^n$
- $K = \{(x, t) \mid \|x\|_2 \leq t\}$: $K^* = \{(x, t) \mid \|x\|_2 \leq t\}$
- $K = \{(x, t) \mid \|x\|_1 \leq t\}$: $K^* = \{(x, t) \mid \|x\|_\infty \leq t\}$

first three examples are **self-dual** cones

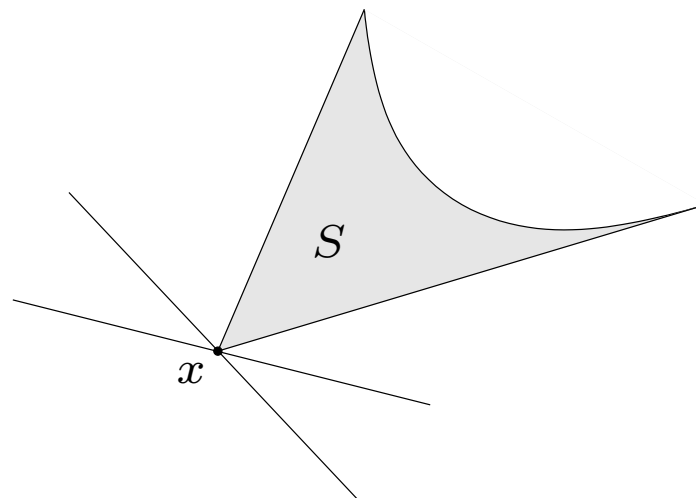
dual cones of proper cones are proper, hence define generalized inequalities:

$$y \succeq_{K^*} 0 \iff y^T x \geq 0 \text{ for all } x \succeq_K 0$$

Minimum and minimal elements via dual inequalities

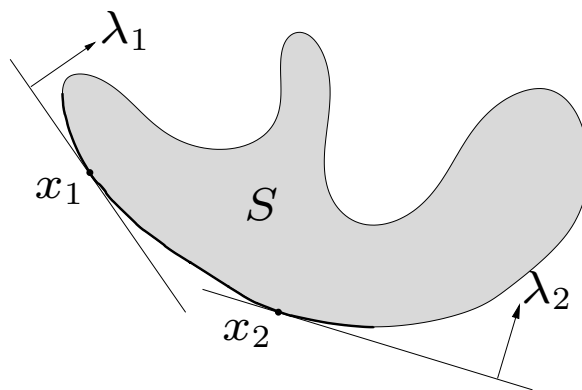
minimum element w.r.t. \preceq_K

x is minimum element of S iff for all $\lambda \succ_{K^*} 0$, x is the unique minimizer of $\lambda^T z$ over S



minimal element w.r.t. \preceq_K

- if x minimizes $\lambda^T z$ over S for some $\lambda \succ_{K^*} 0$, then x is minimal



- if x is a minimal element of a *convex* set S , then there exists a nonzero $\lambda \succeq_{K^*} 0$ such that x minimizes $\lambda^T z$ over S

optimal production frontier

- different production methods use different amounts of resources $x \in \mathbf{R}^n$
- production set P : resource vectors x for all possible production methods
- efficient (Pareto optimal) methods correspond to resource vectors x that are minimal w.r.t. \mathbf{R}_+^n

example ($n = 2$)

x_1, x_2, x_3 are efficient; x_4, x_5 are not

