Convex Optimization

Chapter 1 引言

shijuanfeng

2013-08-19

Outline

- 优化问题的概念
- 问题分类
- 最小二乘
- 主要内容
- 软件包

优化问题的概念

$$
\begin{array}{ll}
\text{minimize} & f_0(x) \\
\text{subject to} & f_i(x) \le b_i, \quad i = 1, \ldots, m
\end{array}
$$

$$x \in \mathcal{R}^n$$

- 优化变量/Optimization Variable
- 目标函数/Objective function
- 约束函数/Constraint functions
- 解/optimal / solution

http://shijuanfeng.blogbus.com/

优化问题分类

线性规划：目标函数和约束函数都是线性函数

Dantzig的单纯形法，内点法 Chebychev逼近问题

非线性规划

目前还没有有效的解决办法 局部优化

凸优化问题：目标函数和约束函数都是凸函数

内点法

非凸优化问题

非凸问题中，凸优化的应用：

✓局部优化中利用凸优化

✓启发式算法

eg：随机算法

✓全局优化的界

eg：松弛算法

转化或者近似

最小二乘

$$\min\left|B-AX\right|^2=\min(B-AX)^T(B-AX)$$

方法：令一阶导等于0

$$A^T(B-AX)=0$$

$$\Leftrightarrow A^TAX=A^TB$$

$$\Leftrightarrow X=(A^TA)^{-1}A^TB$$

加权最小二乘

$$f(x) = \sum_{i=1}^{k} w_i (a_i^T x - b_i)^2$$

正则化最小二乘

$$f(x) = \sum_{i=1}^{k} (a_i^T x - b_i)^2 + \rho \sum_{i=1}^{n} x_i^2 \qquad \rho > 0$$

主要内容

- 理论：Ch2-5 **10.14**

  凸集，凸函数，凸优化问题，对偶

- 应用：Ch6-8 **11.25**

  拟合，统计估计，几何问题

- 算法：Ch9-11 **1.6**

  无约束优化，等式优化，内点法

- 附录：

  数学知识

  特殊的优化问题：目标函数是二次函数，且含有一个二次约束→双二次函数

  数值线性代数？ 线性代数

软件包

- LOQO：http://www.princeton.edu/~rvdb/loqo/LOQO.html
- MOSEK： http://www.mosek.com/
- AMPL： http://www.ampl.com/DOWNLOADS/
- GAMS： http://www.gams.com/

http://shijuanfeng.blogbus.com/

Summary

- 优化问题的概念
- 问题分类：2种分类方法
- 最小二乘：原型+2种变形
- 主要内容：3+1
- 软件包：4个

Convex Optimization

Chapter 2 凸集

shijuanfeng

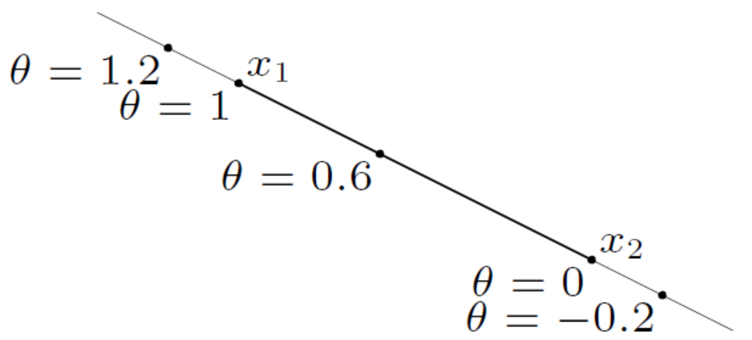
2013-08-19

Outline

- 仿射集+凸集→锥
- 重要例子
- 保凸运算
- 分离和支撑超平面
- 广义不等式
- 对偶锥与广义不等式

仿射集 Affine Set

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



- **仿射集**：过集合C内任意两点的直线均在集合C内

$$\forall x_1, x_2 \in C, \quad \theta \in R, \quad \Rightarrow \quad \theta x_1 + (1-\theta)x_2 \in C$$

仿射集的例：直线、平面、超平面

- ✓ 仿射组合
- ✓ 维数
- ✓ 相对内部
- ✓ 相对边界

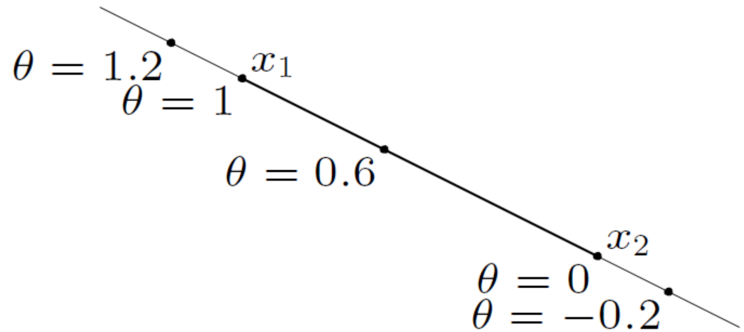
http://shijuanfeng.blogbus.com/

凸集 Convex set

仿射一定凸，凸不一定仿射

仿射∈凸

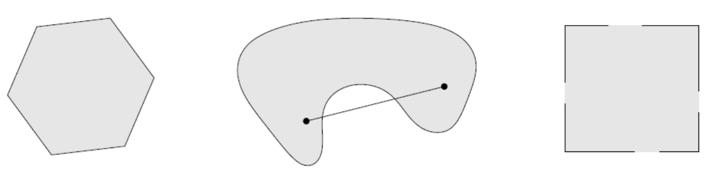
- 线段 $$x = \theta x_1 + (1-\theta)x_2$$



$$0 \le \theta \le 1$$

- **凸集：**过集合C内任意两点的线段均在集合C内

$$\forall x_1, x_2 \in C, \quad 0 \le \theta \le 1, \quad \Rightarrow \quad \theta x_1 + (1-\theta)x_2 \in C$$

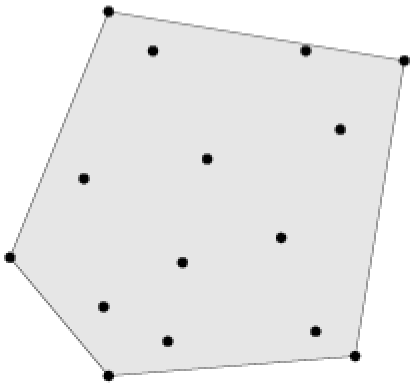


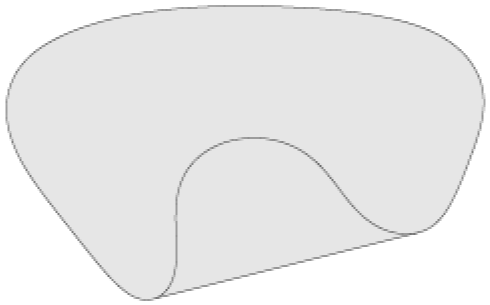
凸包

这里的“小”是如何定义的?

- 凸包：包含集合C的最小的凸集。

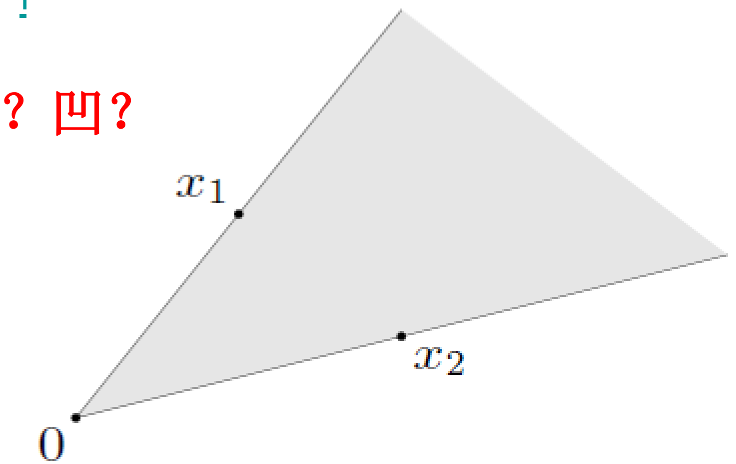
$$\operatorname{conv} C = \{\sum_{i=1}^{k} \theta_i x_i \mid x_i \in C, \theta_i \geq 0, \sum_{i=1}^{k} \theta_i = 1\}$$





Cone

非凸的锥？什么样子的？凹？

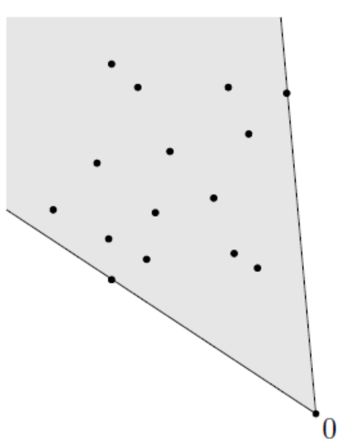


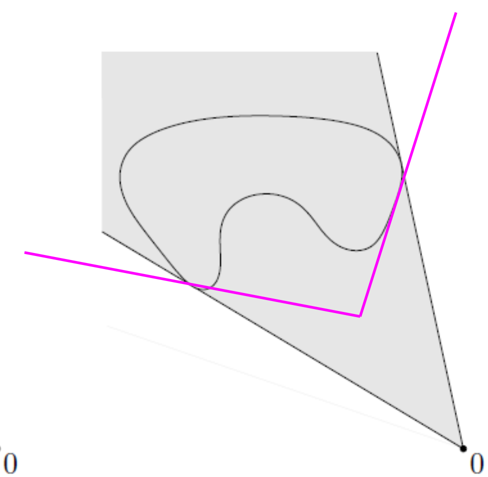
$$\forall x \in C, \theta \geq 0, \text{则有}\ \theta x \in C.$$

Convex cone

$$\forall x_1, x_2 \in C, \theta_1, \theta_2 \geq 0, \text{则有}\ \theta_1 x_1 + \theta_2 x_2 \in C.$$

锥包 集合C内点的所有锥组合，包含C的最小凸锥

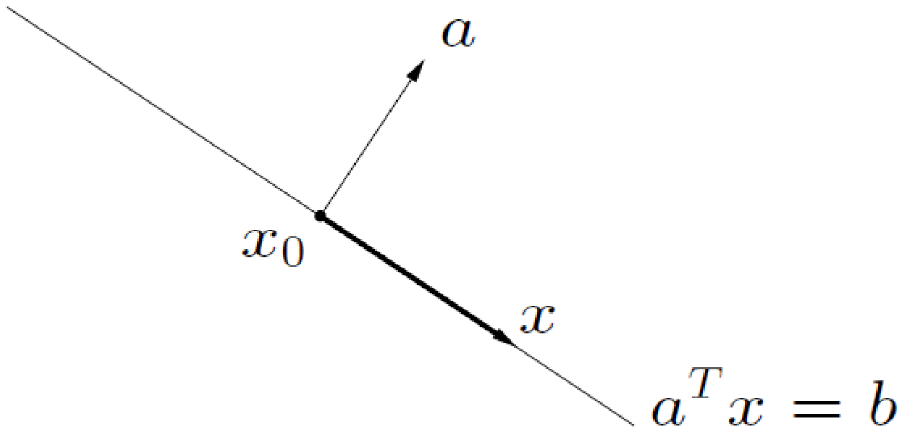


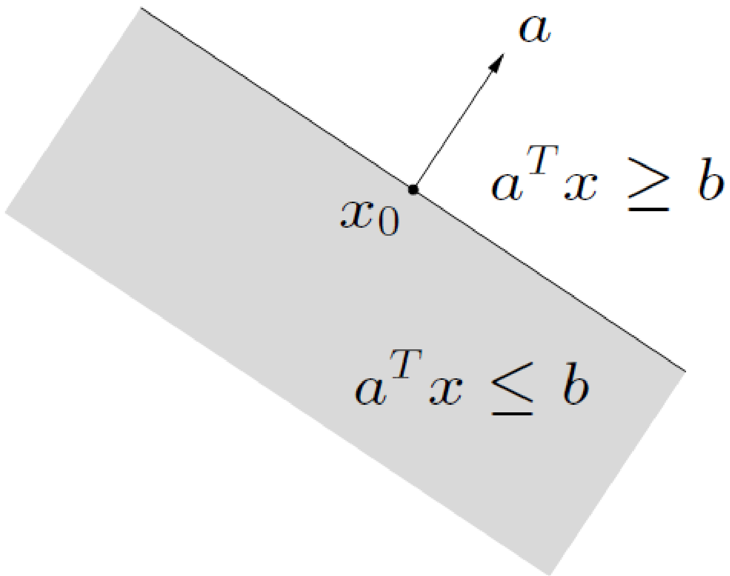


这里的“小”是如何定义的？
原点**0**是如何定义的？

重要例子——超平面&半空间

- 超平面 $\{x \mid a^T x = b\}\ (a \neq 0)$ 仿射、凸





- 半空间 $\{x \mid a^T x \leq b\}\ (a \neq 0)$ 凸

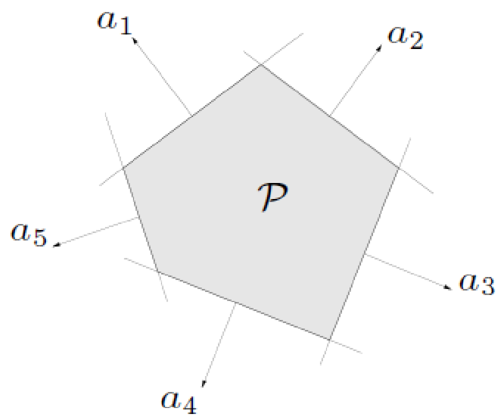
http://shijuanfeng.blogbus.com/

重要例子——多面体

凸

$$P=\{x \mid a_j^T x \le b_j, c_i^T x = d_i\}$$

有限多个半空间和超平面相交组成



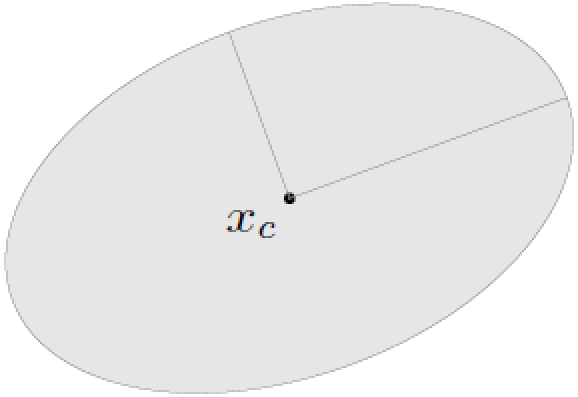
单纯形(simplex)：

?

1维单纯形：线段
2维单纯形：三角形
3维单纯形：四面体

$$\{\sum_{i=0}^{k}\theta_i v_i \mid \theta_i \ge 0, \sum_{i=0}^{k}\theta_i = 1, v_1 - v_0, \ldots, v_k - v_0 \text{线性无关}\}$$

重要例子——3类球 凸



➢Euclid球

$$B(x_c, r) = \{x \mid \|x - x_c\|_2 \le r\}$$
$$= \{x \mid (x - x_c)^T (x - x_c) \le r^2\}$$

➢椭球

$$E = \{x \mid (x - x_c)^T P^{-1} (x - x_c) \le r^2\},\ P\text{为对称正定矩阵}$$

➢范数球

$$B(x_c, r) = \{x \mid \|x - x_c\| \le r\}$$

$Norm:$

$\|x\| \ge 0, \|x\| = 0$当且仅当$x = 0;$

$\|tx\| = |t|\,\|x\|, t \in \mathcal{R};$

$\|x + y\| \le \|x\| + \|y\|$

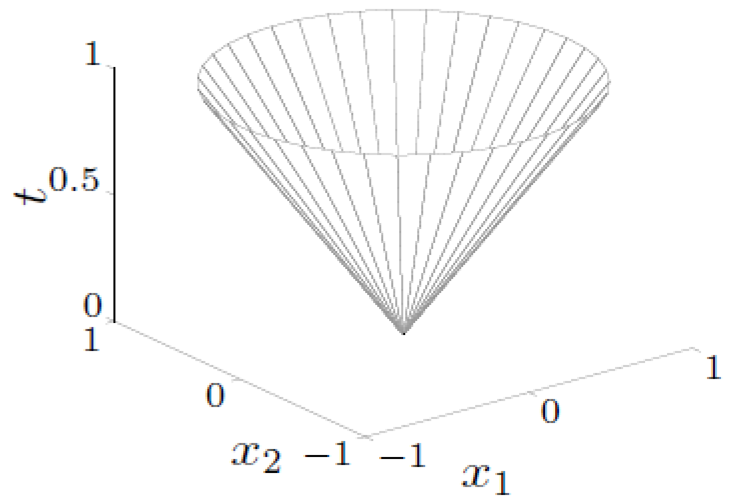
http://shijuanfeng.blogbus.com/

重要例子——2类锥

凸

➢范数锥

$$\{(x,t) \mid \|x\| \le t\}$$



➢半正定锥

n阶对称矩阵集 $S^n = \{X \in \mathcal{R}^{n\times n} \mid X = X^T\}$

n阶半正定矩阵集 $S_+^n = \{X \in S^n \mid X \succeq 0\}$

n阶对称正定矩阵集 $S_{++}^n = \{X \in S^n \mid X \succ 0\}$

➢正定

$$\forall \vec{z} \neq \vec{0}, z^T M z > 0,$$ 则M是正定矩阵

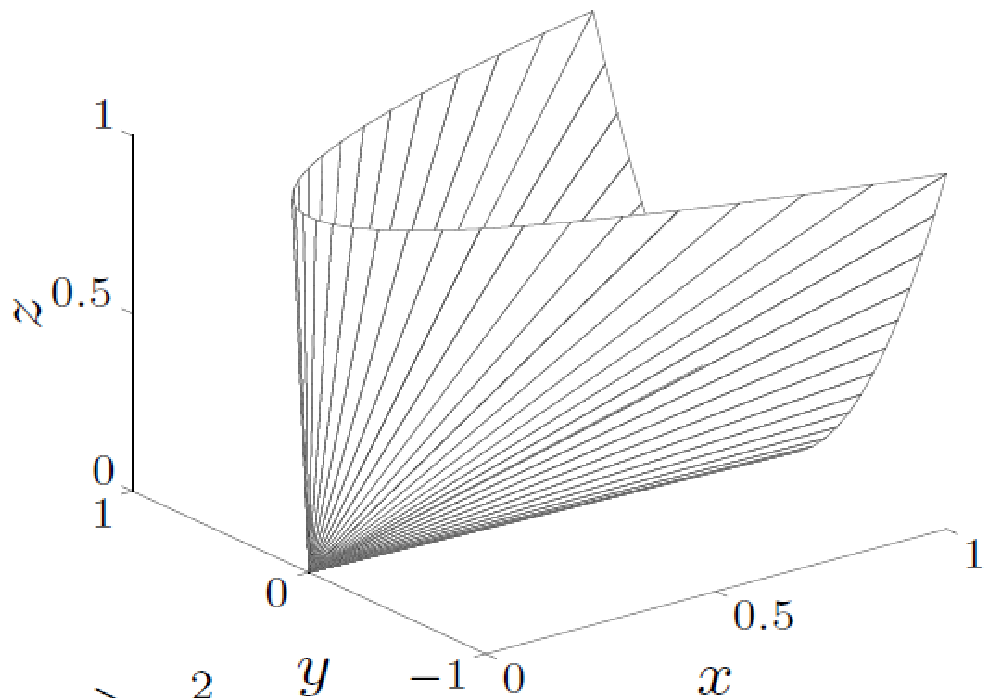
充要条件：

- M的特征值全为正
- M的各阶顺序主子式都为正
- M合同于单位阵

➢半正定

$$\forall \vec{z} \neq \vec{0}, z^T M z \geq 0,$$ 则M是半正定矩阵

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



$$\iff \quad x \geq 0, \quad z \geq 0, \quad xz \geq y^2$$

保凸运算

- 交
- 仿射函数
- 线性分式 (perspective function)

$$P(z,t) = z/t, z \in \mathcal{R}^n, t \in \mathcal{R}_{++}$$

- 透视函数**(**linear-fractional function**)**

=仿射+线性分式

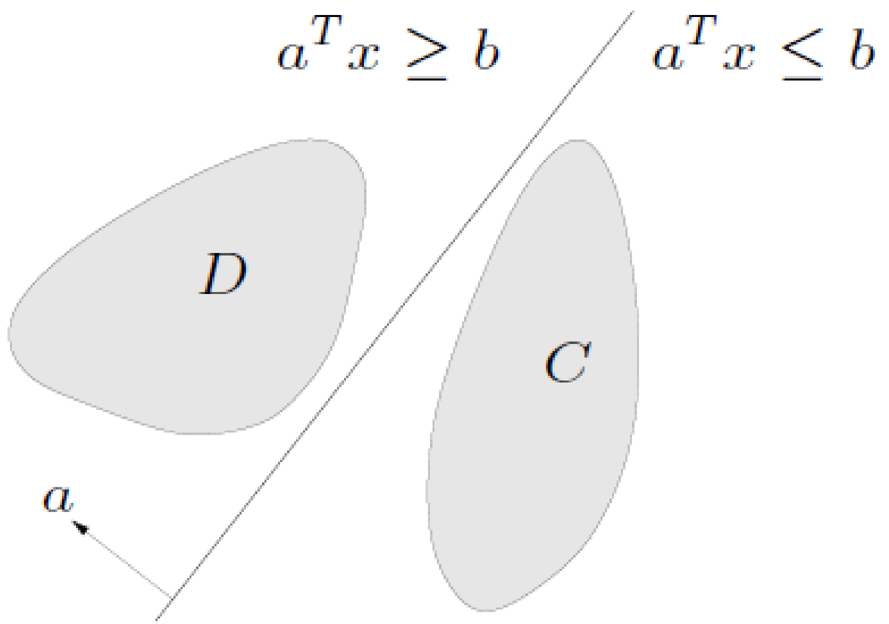
$$f(x) = (Ax+b)/(c^T x + d)$$

$$A \in \mathcal{R}^{m\times n}, b \in \mathcal{R}^m, c \in \mathcal{R}^n, d \in \mathcal{R}, c^T x + d > 0$$

分离超平面

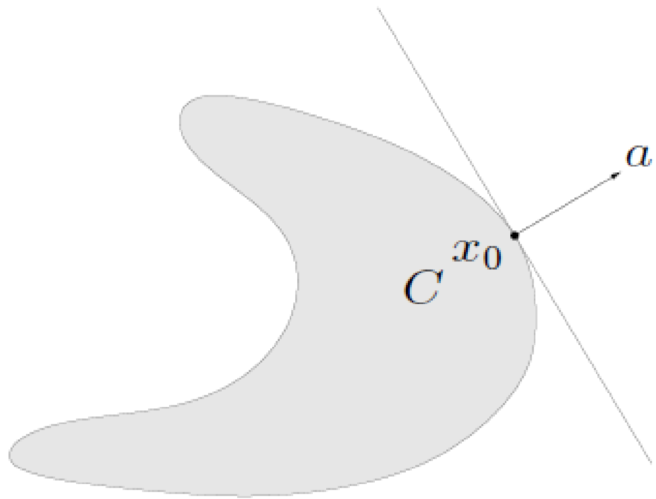
设 C 和 D 为两不相交凸集，则存在超平面将 C 和 D 分离。

$$\forall x \in C, a^T x \le b \text{且} \forall x \in D, a^T x \ge b.$$



支撑超平面

定义：设集合 C，x_0 为 C 边界上的点。若存在 $a \neq 0$，满足对任意 $x \in C$，都有 $a^T x \leq a^T x_0$ 成立，则称超平面 $\{x \mid a^T x = a^T x_0\}$ 为集合 C 在点 x_0 处的支撑超平面。



定理： 凸集边界上任意一点均存在支撑超平面。

定理： 若一个闭的非中空集合，在边界上的任意一点存在支撑超平面，则该集合为凸集。

广义不等式

- Proper Cone/真锥

1. K为凸集；
2. K为闭集；
3. K非中空；
4. K有端点。

偏序关系：

$$x \preceq_K y \Leftrightarrow y - x \in K$$

严格偏序关系：

$$x \prec_K y \Leftrightarrow y - x \in \operatorname{int} K$$

http://shijuanfeng.blogbus.com/

广义不等式的性质

$1. x \preceq_K x;$

$2. x \preceq_K y, y \preceq_K x \Rightarrow x = y;$

$3. x \preceq_K y, y \preceq_K z \Rightarrow x \preceq_K z;$

$4. x \preceq_K y, u \preceq_K v \Rightarrow x + u \preceq_K y + v;$

$5. x \preceq_K y, \alpha > 0 \Rightarrow \alpha x \preceq_K \alpha y;$

$6. x_i \preceq_K y_i, \lim x_i = x, \lim y_i = y \Rightarrow x \preceq_K y.$

广义不等式的性质

$1. x \prec_K y \Rightarrow x \preceq_K y;$

$2. x \nprec_K x;$

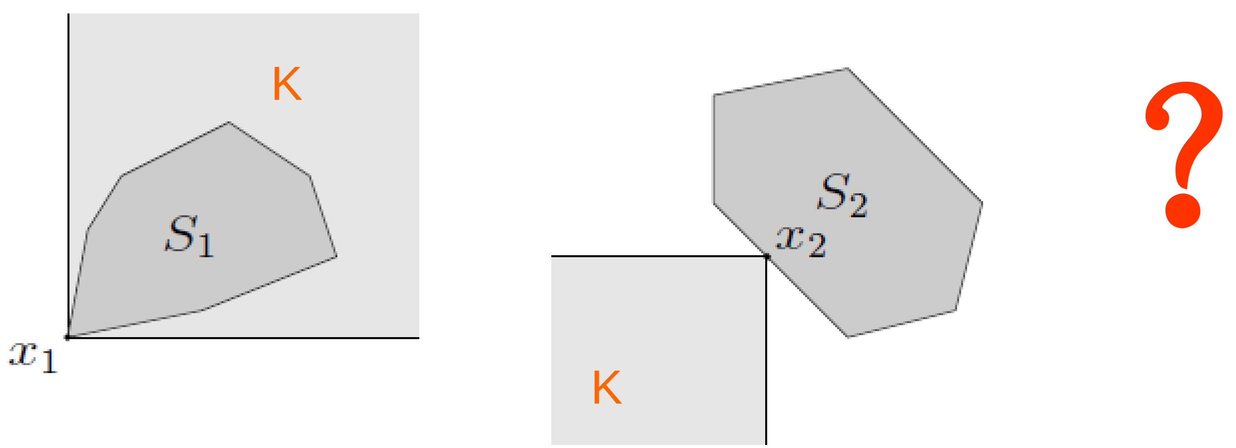
$3. x \prec_K y, u \preceq_K v \Rightarrow x + u \prec_K y + v;$

$4. x \prec_K y, \alpha > 0 \Rightarrow \alpha x \prec_K \alpha y$

$5. x \prec_K y, u$足够小$\Rightarrow x + u \prec_K y.$

最值和极值

- 最小元 $$y \in S \implies x \preceq_K y$$
- 极小元 $$y \in S, \quad y \preceq_K x \implies y = x$$



对偶锥

$$K^* = \{y \mid x^T y \geq 0, \forall x \in K\}$$

examples

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

广义不等式的对偶

- 广义不等式与对偶等价性质

$$x \preceq_K y \Leftrightarrow \lambda^T x \le \lambda^T y, \text{for all } \lambda \succeq_{K^*} 0;$$

$$x \prec_K y \Leftrightarrow \lambda^T x \le \lambda^T y, \text{for all } \lambda \succeq_{K^*} 0, \lambda \ne 0.$$

- 最小元的对偶性

x为集合S中关于K偏序的最小元

$\Leftrightarrow$ 对所有$\lambda \succ_{K^*} 0$, x为使$\lambda^T z (z \in S)$最小的值.

- 极小元的对偶性

$\lambda \succ_{K^*} 0$, x为使$\lambda^T z (z \in S)$最小的值 $\Rightarrow$ x为极小元.

Summary

- 仿射集+凸集→锥
- 重要例子：超平面&半空间；多面体； 3类球； 2类锥
- 保凸运算：交，仿射函数，线性分式&透视函数
- 分离和支撑超平面：定义+定理
- 广义不等式：定义、性质、最小与极小元
- 对偶锥与广义不等式：对偶锥、广义不等式的对偶、最小元、极小元的对偶性

Thanks!