

凸集

文再文

北京大学北京国际数学研究中心

教材《最优化：建模、算法与理论》配套电子教案

<http://bicmr.pku.edu.cn/~wenzw/optbook.html>

致谢：本教案由丁思哲协助准备

- 1 范数
- 2 凸集的定义
- 3 重要的凸集举例
- 4 保凸的运算
- 5 广义不等式与对偶锥
- 6 分离超平面定理

向量范数的定义

定义

令记号 $\|\cdot\|: \mathbb{R}^n \rightarrow \mathbb{R}^+$ 是一种非负函数, 如果它满足:

- 正定性: 对于 $\forall v \in \mathbb{R}^n$, 有 $\|v\| \geq 0$, 且 $\|v\| = 0 \Leftrightarrow v = 0_{n \times 1}$;
- 齐次性: 对于 $\forall v \in \mathbb{R}^n$ 和 $\alpha \in \mathbb{R}$, 有 $\|\alpha v\| = |\alpha| \|v\|$;
- 三角不等式: 对于 $\forall v, w \in \mathbb{R}^n$, 均成立 $\|v + w\| \leq \|v\| + \|w\|$.

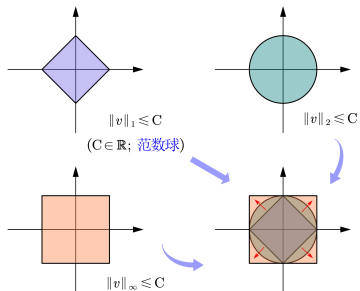
则称 $\|\cdot\|$ 是定义在向量空间 \mathbb{R}^n 上的 **向量范数**.

最常用的向量范数即我们熟知的 ℓ_p 范数(其中 $p \geq 1$):

$$\|v\|_p = \left(\sum_{i=1}^n |v_i|^p \right)^{\frac{1}{p}}; \quad \|v\|_\infty = \max_{1 \leq j \leq n} |v_j|.$$

柯西不等式: 设 $a, b \in \mathbb{R}^n$, 则 $|a^T b| \leq \|a\|_2 \|b\|_2$, 且等号成立的条件是 a 与 b 线性相关.

向量范数的定义



容易看出, $p = \infty$ 时, 有关"最大值"的定义要求向量的分量是有限的. 在一般化的空间中, 这一要求很可能不成立, 此时我们只需将"最大值"更换成"上确界"即可.

向量范数度量的是 v 与零点之间的距离. 在实际应用时, 我们通常使用 $p = 1, 2, \infty$ 的情形, 即分别使用 $\|v\|_1, \|v\|_2, \|v\|_\infty$ 度量 v 在不同意义下的距离, 这是因为它们具有鲜明的度量特征.

左图是它们各自的范数球实例, 请想一想不同范数所度量的距离分别具有怎样的特征? 这些特征分别适用于度量什么情形?

矩阵范数

矩阵范数可以由向量范数的定义推广得到。常见的矩阵范数有：

- $\|A\|_1 = \sum_{i,j} |A_{ij}|$
- $\|A\|_F = \sqrt{\sum_{i,j} A_{ij}^2} = \sqrt{\text{Tr}(AA^T)}$
- 算子范数是一类特殊的矩阵范数，它由向量范数诱导得到：

$$\|A\|_{(m,n)} = \max_{x \in \mathbb{R}^n, \|x\|_{(n)}=1} \|Ax\|_{(m)}.$$

- $p = 1$ 时, $\|A\|_{p=1} = \max_{\|x\|_1=1} \|Ax\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^m |a_{ij}|.$
- $p = 2$ 时, $\|A\|_{p=2} = \max_{\|x\|_2=1} \|Ax\|_2 = \sqrt{\lambda_{\max}(A^T A)}$, 又称为 A 的谱范数.
- $p = \infty$ 时, $\|A\|_{p=\infty} = \max_{\|x\|_\infty=1} \|Ax\|_\infty = \max_{1 \leq i \leq m} \sum_{j=1}^n |a_{ij}|.$

矩阵范数

- 核范数定义为

$$\|A\|_* = \sum_{i=1}^r \sigma_i,$$

其中 $\sigma_i (i = 1, \dots, r)$ 为 A 的所有非零奇异值, $r = \mathbf{rank}(A)$.

- 矩阵 A, B 的内积定义为

$$\langle A, B \rangle = \text{Tr}(AB^T) = \sum_{i=1}^m \sum_{j=1}^n a_{ij} b_{ij}.$$

- 柯西不等式: 设 $A, B \in \mathbb{R}^{m \times n}$, 则

$$|\langle A, B \rangle| \leq \|A\|_F \|B\|_F,$$

等号成立当且仅当 A 和 B 线性相关.

- 1 范数
- 2 凸集的定义
- 3 重要的凸集举例
- 4 保凸的运算
- 5 广义不等式与对偶锥
- 6 分离超平面定理

凸集的几何定义

在 \mathbb{R}^n 空间中, 经过不同的两点 x_1, x_2 可以确定一条直线, 其方程为

$$y = \theta x_1 + (1 - \theta)x_2, \theta \in \mathbb{R}.$$

特别, 当 $0 \leq \theta \leq 1$ 时, 直线退化为以 x_1, x_2 为端点的线段.

定义

仿射集 如果过集合 C 中的任意两点的直线都在 C 内, 则称 C 为**仿射集**, 即

$$x_1, x_2 \in C \Rightarrow \theta x_1 + (1 - \theta)x_2 \in C, \forall \theta \in \mathbb{R}.$$

例 线性方程组 $Ax = b$ 的解集 \mathcal{X} 是仿射集, 因为 $\forall x_1, x_2 \in \mathcal{X} (x_1 \neq x_2)$ 均满足 $\theta Ax_1 + (1 - \theta)Ax_2 = b$.

反之, **任何仿射集均可表示为某一线性方程组的解集.**

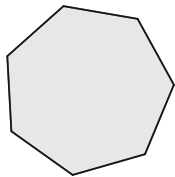
凸集的几何定义

定义

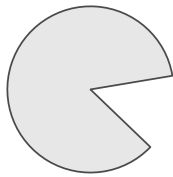
凸集 如果连接集合 C 中的任意两点的线段都在 C 内, 则称 C 为**凸集**, 即

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

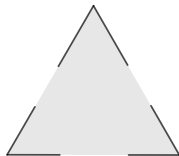
仿射集当然都是凸集.



(a)



(b)



(c)

例 在左图中我们列出了一些凸集、非凸集的例子. 其中(a)为凸集, (b)和(c)均为非凸集.

凸集的性质

定理

- 若 S 是凸集, 则 $kS = \{ks | k \in \mathbb{R}, s \in S\}$ 是凸集.
- 若 S 和 T 均是凸集, 则 $S + T = \{s + t | s \in S, t \in T\}$ 是凸集.
- 若 S 和 T 均是凸集, 则 $S \cap T$ 是凸集.
- 凸集的内部和闭包都是凸集.

上述定理前2点在凸集定义前提下是显然的. 我们简述第3点为何成立.

证明: 设 $x, y \in S \cap T$ 且 $\theta \in [0, 1]$. 由于 S 和 T 均为凸集, 则

$$\theta x + (1 - \theta)y \in S \cap T,$$

这证明 $S \cap T$ 是凸集.

实际上, **任意多凸集**的交都是凸集. 该结论在证明复杂集合是凸集时非常有用, 因为我们可以考虑将其视为任意个凸集的交.

凸组合和凸包

从凸集中可以引出凸组合和凸包的概念.

定义

凸组合 形如

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

$$\theta_1 + \cdots + \theta_k = 1, \theta_i \geq 0, i = 1, \cdots, k.$$

的点称为 x_1, \cdots, x_k 的凸组合.

定义

凸包 集合 S 的所有点的凸组合构成的点集为 S 的凸包, 记为 $\text{conv}S$.

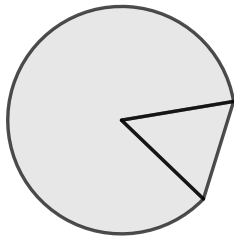
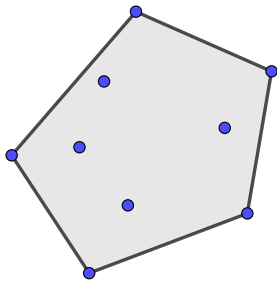
定理

凸集和凸包的关系 若 $\text{conv}S \subseteq S$, 则 S 是凸集; 反之亦然.

上述定理并不显然, 请尝试证明. (提示: 用数学归纳法)

凸包的例子

例 在下图中我们列出了一些离散点集和连续点集的凸包. 其中, 左子图为离散点集的凸包, 右子图为扇形连续点集的凸包.



conv S 是包含 S 的最小凸集

定理

conv S 是包含 S 的最小凸集.

证明 由凸包的定义可知, $S \in \text{conv}S$, 并且conv S 是凸集.

若再设 \mathcal{X} 是另一凸集且满足 $S \subseteq \mathcal{X} \subseteq \text{conv}S$, 下面我们需要证明只可能是 $\mathcal{X} = \text{conv}S$. 为证明此结论, 我们先证明一个重要的命题, 从而直接导出本定理的成立.

定理

对于任意向量集 S , conv S 是包含 S 的一切凸集的交集.

证明 令 \mathcal{X} 表示包含 S 的所有凸集的交集. 我们之前证明, 凸集的交是凸集, 因此 \mathcal{X} 是凸集. 因为conv S 是一个凸集且包含 S , 则 $\mathcal{X} \subseteq \text{conv}S$.

另一方面, $S \subseteq \mathcal{X}$, 因此conv $S \subseteq \text{conv}\mathcal{X}$.

再由凸集和凸包的关系得到conv $\mathcal{X} = \mathcal{X}$, 得到conv $S \subseteq \mathcal{X}$.

综上有 $\mathcal{X} = \text{conv}S$.

仿射包

仿射集和凸集的定义很像,除了 θ 的范围有所不同.受此启发,从凸组合和凸包的定义中可以自然引出仿射组合和仿射包的概念.

定义

仿射组合 形如

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

$$\theta_1 + \cdots + \theta_k = 1, \theta_i \in \mathbb{R}, i = 1, \cdots, k.$$

的点称为 x_1, \cdots, x_k 的仿射组合.

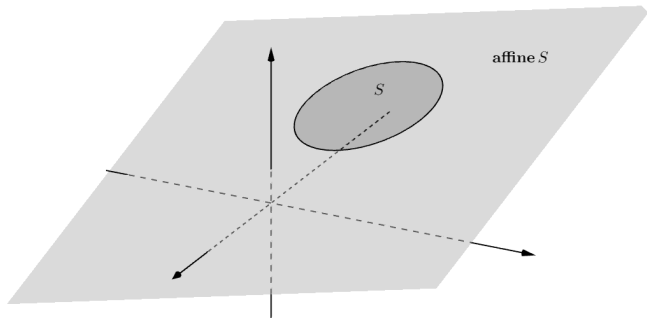
定义

仿射包 集合 S 的所有点的仿射组合构成的点集为 S 的仿射包,记为 $\text{affine}S$.

$\text{affine}S$ 是包含 S 的最小仿射集.

\mathbb{R}^3 中仿射包的例子

例 下图为 \mathbb{R}^3 中圆盘 S 的仿射包示意图,可见仿射包直接将原集合拓展为了其所在的全平面.



锥组合和凸锥

相比于凸组合和仿射组合, 锥组合不要求系数的和为1, 因此一般而言锥组合都是开放的.

定义

锥组合 形如

$$x = \theta_1 x_1 + \cdots + \theta_k x_k, \theta_i > 0 (i = 1, \cdots, k).$$

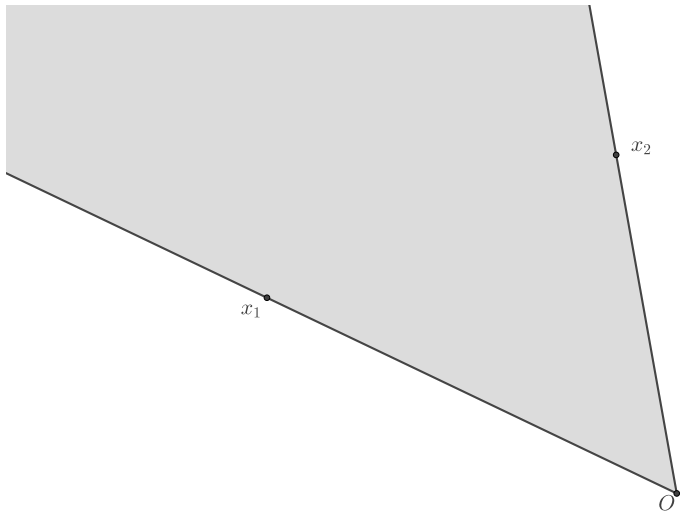
的点称为 x_1, \cdots, x_k 的锥组合.

定义

凸锥 若集合 S 中任意点的锥组合都在 S 中, 则称 S 为凸锥.

凸锥的例

例 下图显示了 \mathbb{R}^2 中两点 x_1, x_2 的凸锥. 可见 \mathbb{R}^2 中若两点不与原点 O 共线, 则其形成的凸锥为一个半径无穷的圆的扇形部分.



- 1 范数
- 2 凸集的定义
- 3 重要的凸集举例**
- 4 保凸的运算
- 5 广义不等式与对偶锥
- 6 分离超平面定理

超平面和半空间

定义

超平面 任取非零向量 $a \in \mathbb{R}^n$, 形如

$$\{x | a^T x = b\}$$

的集合称为超平面.

定义

半空间 任取非零向量 $a \in \mathbb{R}^n$, 形如

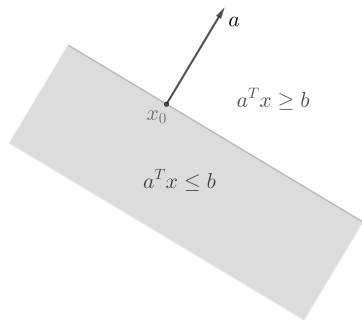
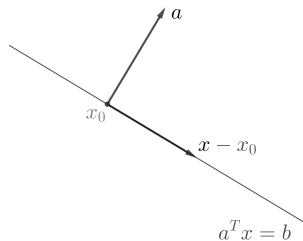
$$\{x | a^T x \leq b\}$$

的集合称为半空间.

超平面是仿射集和凸集, 半空间是凸集但不是仿射集.

超平面和半空间的例

例 下图是 \mathbb{R}^2 中超平面和半空间的例子. 其中, 左子图为超平面, 其为一
条直线; 右子图为半空间.



多面体

我们把满足线性等式和不等式组的点的集合称为多面体, 即

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

其中 $A \in \mathbb{R}^{m \times n}$, $C \in \mathbb{R}^{p \times n}$, $x \leq y$ 表示向量 x 的每个分量都小于等于 y 的对应分量.

多面体是有限个半空间和超平面的交, 因此由凸集的性质可知, 其为凸集.

范数球、椭球

如下定义的球和椭球也是常见的凸集.

定义

球 设空间中到某一定点 x_c (称为**中心**)的距离小于等于定值 r (称为**半径**)的点的集合为(范数)球, 即

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

一般而言, 我们使用 $\|\cdot\|_2$ 度量距离, 即使用2-范数球.

定义

椭球 设形如

$$\left\{x \mid (x - x_c)^T P^{-1} (x - x_c) \leq 1\right\} = \{x_c + Au \mid \|u\|_2 \leq 1\}$$

的集合为椭球, 其中 x_c 为**椭球中心**, P 对称正定, 且 A 非奇异.

范数锥

球和椭球的范围取决于 x 的范围，而锥的范围则同时取决于 x 和控制半径 t 的范围。

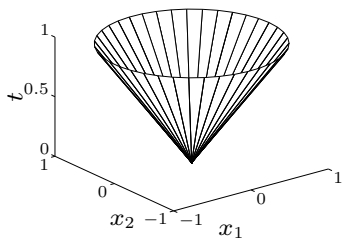
定义

范数锥 形如

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

的集合为范数锥。

锥是凸集。同时，使用 $\|\cdot\|_2$ 度量距离的锥为二次锥，也称冰淇淋锥。



特殊矩阵集合和(半)正定锥

我们介绍3类矩阵的集合.

定义

对称矩阵集合 记 S^n 为 $n \times n$ 对称矩阵的集合, 即

$$S^n = \{X \in \mathbb{R}^{n \times n} | X^T = X\}.$$

定义

半正定矩阵集合 记 S_+^n 为 $n \times n$ 半正定矩阵的集合, 即

$$S_+^n = \{X \in S^n | X \succeq 0\}.$$

定义

正定矩阵集合 记 S_{++}^n 为 $n \times n$ 正定矩阵的集合, 即

$$S_{++}^n = \{X \in S^n | X \succ 0\}.$$

半正定锥的例子

我们一般称 S_+^n 为半正定锥. 下图是二维半正定锥的几何形状.

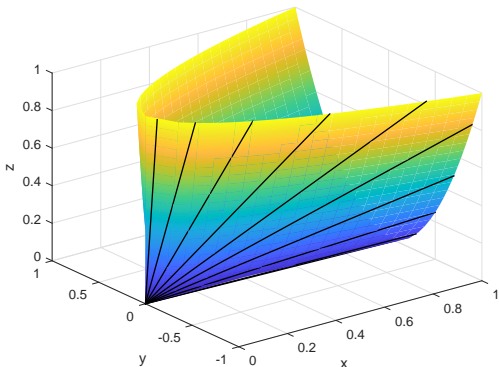
由图可知, 二维半正定锥的实际范围是

$$\{(x, y, z) \mid x \geq 0, z \geq 0, xz \geq y^2\}.$$

这实际上可以由半正定矩阵的性质直接得到:

对于矩阵 $\begin{pmatrix} x & y \\ y & z \end{pmatrix}$, 其特征值应全部大于等于0, 由此可推出

$$x \geq 0, z \geq 0, xz \geq y^2.$$



- 1 范数
- 2 凸集的定义
- 3 重要的凸集举例
- 4 保凸的运算**
- 5 广义不等式与对偶锥
- 6 分离超平面定理

仿射变换的保凸性

仿射变换(缩放、平移、投影等)也是保凸的.

定理

仿射变换的保凸性 设 $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ 是仿射变换, 即 $f(x) = Ax + b$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, 则

- 凸集在 f 下的像是凸集:

$$S \subseteq \mathbb{R}^n \text{ 是凸集} \Rightarrow f(S) = \{f(x) | x \in S\} \text{ 是凸集.}$$

- 凸集在 f 下的原像是凸集:

$$C \subseteq \mathbb{R}^m \text{ 是凸集} \Rightarrow f^{-1}(C) = \{x | f(x) \in C\} \text{ 是凸集.}$$

仿射变换的保凸性

例 线性矩阵不等式的解集

$$\{x | x_1 A_1 + \cdots + x_m A_m \preceq B\} \quad (A_i, i = 1, \cdots, m, B \in \mathcal{S}^p)$$

是凸集. 这由仿射变换可以直接得到.

例 双曲锥

$$\{x | x^T P x \leq (c^T x)^2, c^T x \geq 0, P \in \mathcal{S}_+^n\}$$

是凸集.

证明: 双曲锥可以转化为二阶锥

$$\{x | \|Ax\|_2 \leq c^T x, c^T x \geq 0, A^T A = P\},$$

而二阶锥可由二次锥 $\{(x, t) | \|x\|_2 \leq t, t \geq 0\}$ 经过仿射变换得到, 因此二阶锥、二次锥均为凸集.

透视变换和分式线性变换的保凸性

- 透视变换 $P: \mathbb{R}^{n+1} \rightarrow \mathbb{R}$:

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

透视变换下凸集的像和原像是凸集。

- 分式线性变换 $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$:

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

分式线性变换下凸集的像和原像是凸集。

- 1 范数
- 2 凸集的定义
- 3 重要的凸集举例
- 4 保凸的运算
- 5 广义不等式与对偶锥
- 6 分离超平面定理

适当锥

我们知道锥是凸集. 一个凸锥 $K \subseteq \mathbb{R}^n$ 是适当锥, 当它还满足

- K 是闭集;
- K 是实心的, 即 $\text{int}K \neq \emptyset$;
- K 是尖的, 即内部不含有直线: 若 $x \in K, -x \in K$, 则一定有 $x = 0$.

例 非负卦限 $K = \mathbb{R}_+^n = \{x \in \mathbb{R}^n | x_i \geq 0, i = 1, \dots, n\}$ 是适当锥.

例 半正定锥 $K = \mathcal{S}_+^n$ 是适当锥.

例 $[0, 1]$ 上的有限非负多项式

$$K = \{x \in \mathbb{R}^n | x_1 + x_2 t + \dots + x_n t^{n-1} \geq 0, \forall t \in [0, 1]\}$$

是适当锥.

广义不等式

广义不等式是一种偏序(不必要保证所有对象都具有可比较性)关系, 可以使用适当锥诱导.

定义

广义不等式 对于适当锥 K , 定义偏序广义不等式为

$$x \preceq_K y \iff y - x \in K,$$

严格偏序广义不等式为

$$x \prec_K y \iff y - x \in \text{int}K.$$

例 坐标分量不等式($K = \mathbb{R}_+^n$)

$$x \preceq_{\mathbb{R}_+^n} y \iff y_i \geq x_i.$$

例 矩阵不等式($K = \mathcal{S}_+^n$)

$$X \preceq_{\mathcal{S}_+^n} Y \iff Y - X \text{ 半正定.}$$

广义不等式的性质

\preceq_K 的诸多性质在 \mathbb{R} 中与 \leq 类似.

定理

广义不等式的性质 记 \preceq_K 是定义于适当锥 K 上的广义不等式, 则

- 自反性: $x \preceq_K x$;
- 反对称性: 若 $x \preceq_K y$ 且 $y \preceq_K x$, 则 $x = y$;
- 传递性: 若 $x \preceq_K y$ 且 $y \preceq_K z$, 则 $x \preceq_K z$;
- 可加性: 若 $x \preceq_K y$ 且 $u \preceq_K v$, 则 $x + u \preceq_K y + v$;
- 非负缩放: 若 $x \preceq_K y$ 且 $\alpha \geq 0$, 则 $\alpha x \preceq_K \alpha y$.

利用偏序关系和广义不等式的定义可以轻松证明上述性质.

对偶锥

设 K 是一个锥.

定义

对偶锥 令锥 K 为全空间 Ω 的子集, 则 K 的对偶锥为

$$K^* = \{y \in \Omega \mid \langle x, y \rangle \geq 0, \forall x \in K\}.$$

对偶锥是相对于锥 K 定义的, 因此我们知道锥的同时也可以求出对偶锥. 我们将对偶锥为自身的锥称为自对偶锥

例 $K = \mathbb{R}_+^n$ 的对偶锥是它本身, 因此是自对偶锥.

例(请自证) $K = S_+^n$ 的对偶锥是它本身, 因此是自对偶锥.

例(请自证) 锥 $K = \{(x, t) \mid \|x\|_p \leq t, t > 0, p \geq 1\}$ 的对偶锥是

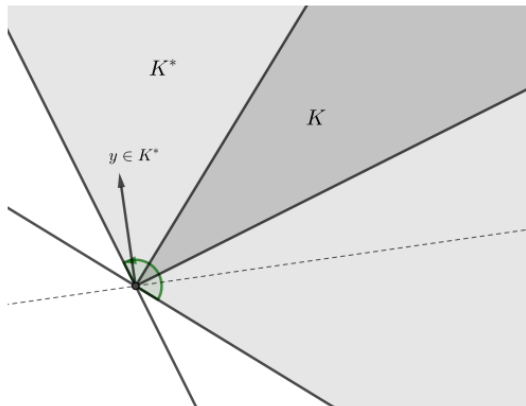
$$K^* = \{(x, t) \mid \|x\|_q \leq t, t > 0, q \geq 1, (p, q) \text{ 共轭}\}.$$

例 上例中二次锥的对偶锥是它本身, 因此是自对偶锥.

对偶锥

例 我们在下图中给出了一个 \mathbb{R}^2 平面上的一个例子. 图中深色区域表示锥 K , 根据对偶锥的定义, K^* 中的向量和 K 中所有向量夹角均为锐角或直角. 因此, 对偶锥 K^* 为图中的浅色区域.

注意, 在这个例子中, K 也为 K^* 的一部分.



对偶锥的性质

下面我们简单列举对偶锥满足的性质, 这是很重要的.

定理

对偶锥的性质 设 K 是一锥, K^* 是其对偶锥, 则满足

- K^* 是锥(哪怕 K 不是锥也成立);
- K^* 始终是闭集, 且是凸集;
- 若 $\text{int}K \neq \emptyset$, 则 K^* 是尖的, 即内部不含有直线;
- 若 K 是尖的, 则 $K^* \neq \emptyset$;
- 若 K 是适当锥, 则 K^* 是适当锥;
- **(二次对偶)** K^{**} 是 K 的凸包. 特别, 若 K 是凸且闭的, 则 $K^{**} = K$.

对偶锥诱导的广义不等式

既然适当锥的对偶锥仍是适当锥, 则可以用适当锥 K 的对偶锥 K^* 也可以诱导广义不等式. 我们在下文简称其为"对偶广义不等式".

定义

对偶广义不等式 适当锥的对偶锥 K^* 可定义广义不等式

$$x \preceq_{K^*} y \iff y - x \in K^*,$$

其满足性质:

- $x \preceq_K y \iff \lambda^T x \leq \lambda^T y, \forall \lambda \succeq_{K^*} 0$;
- $y \succeq_{K^*} 0 \iff y^T x \geq 0, \forall x \succeq_K 0$.

使用对偶广义不等式的好处是, 对偶锥始终是闭且凸的, 并可将一个偏序问题转换为满足一个偏序条件的全序问题.

- 1 范数
- 2 凸集的定义
- 3 重要的凸集举例
- 4 保凸的运算
- 5 广义不等式与对偶锥
- 6 分离超平面定理**

分离超平面定理

超平面是空间中一类特殊的凸集(仿射集),可以证明 \mathbb{R}^n 空间中的超平面恰好是 $n-1$ 维的.我们可以用超平面分离不相交的凸集.

定理

分离超平面定理 如果 C 和 D 是不相交的凸集,则存在非零向量 a 和常数 b ,使得

$$a^T x \leq b, \forall x \in C,$$

且

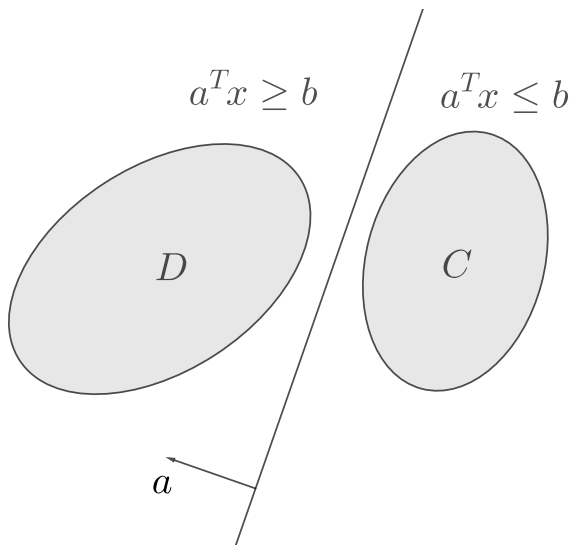
$$a^T x \geq b, \forall x \in D,$$

即超平面 $\{x | a^T x = b\}$ 分离了 C 和 D .

超平面分离定理表明,如果要**软划分** \mathbb{R}^n 中的2个凸集,则只需要求得一个适当的超平面即可.这在分类问题中属于很容易解决的问题.实际上,如果有任何一个集合不是凸集,则定理一般不成立,此时我们若要划分不同的集合,则一般需要使用更加复杂的平面.

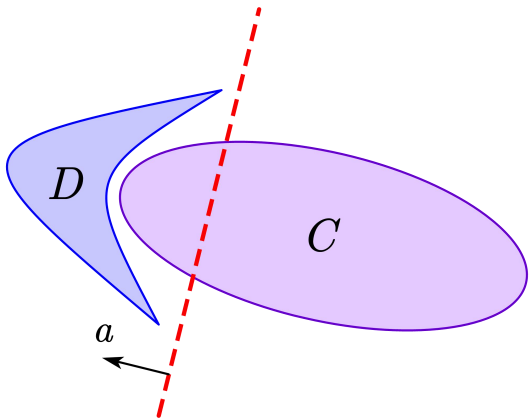
分离超平面的示意

例 下图是 \mathbb{R}^2 中的2个凸集, 我们使用超平面即可轻松划分.



分离超平面的示意

例 下图是 \mathbb{R}^2 中的两个集合, 其中一个不为凸集. 我们无法使用超平面对其划分, 而必须使用更加复杂的平面. 这就给划分问题带来了巨大的挑战.



分离超平面定理证明

这里仅考虑一个特殊情形。假设存在 $c \in C$ 和 $d \in D$ 使得:

$$\|c - d\|_2 = \mathbf{dist}(C, D) = \inf\{\|u - v\|_2 \mid u \in C, v \in D\} > 0.$$

定义 $a = d - c$, $b = (\|d\|_2^2 - \|c\|_2^2)/2$ 和

$$f(x) = a^T x - b = (d - c)^T(x - (d + c)/2).$$

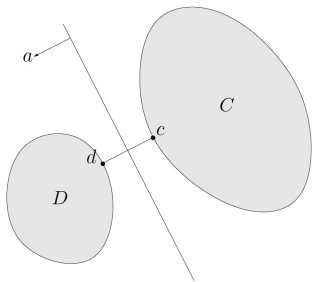
以下证: $f(x) \leq 0, \forall x \in C$ 且 $f(x) \geq 0, \forall x \in D$, 即给出了分离超平面。

- 先证 $f(x) \geq 0, \forall x \in D$ 。其它情形类似。
- 假设存在 $u \in D$, 使得

$$f(u) = (d - c)^T(u - (d + c)/2) < 0.$$

可以将 $f(u)$ 写成:

$$f(u) = (d - c)^T(u - d) + \|d - c\|_2^2/2.$$



分离超平面定理证明

- 因此有：

$$(d - c)^T(u - d) < 0.$$

- 对于 $t \in [0, 1]$ ，构造 d 与 u 的凸组合 $z(t) = d + t(u - d)$ ，因此 $z(t)$ 也在集合 D 里。由于

$$\frac{d}{dt} \|z(t) - c\|_2^2|_{t=0} = 2(d - c)^T(u - d) < 0,$$

因此存在充分小的 $t_1 \in (0, 1]$ ，使得

$$\|z(t_1) - c\|_2 < \|d - c\|_2.$$

这意味着点 $z(t_1)$ 到 c 的距离比 d 近，矛盾。

严格分离定理

我们在超平面分离时提到了**软划分**的概念,其表明若集合仅是凸集,则定理中等号可能成立,即某一凸集与超平面相交(请尝试举一个简单例子)。很多时候进一步要求超平面与任何凸集都不交,为此我们需要加强定理的条件。

定理

严格分离定理 如果 C 和 D 是不相交的凸集,且 C 是闭集, D 是紧集,则存在非零向量 a 和常数 b ,使得

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

且

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

即超平面 $\{x | a^T x = b\}$ 严格分离了 C 和 D 。

此定理的退化形式即 D 退化为单点集 $\{x_0\}$ 。此时课本中的定理成立。

支撑超平面

上述严格分离定理的退化形式要求 $x_0 \notin C$. 当点 x_0 恰好落在 C 的边界上时(此时不满足"不相交"的条件), 我们可以构造超平面.

定义

支撑超平面 给定集合 C 以及边界上的点 x_0 , 如果 $a \neq 0$ 满足 $a^T x \leq a^T x_0, \forall x \in C$, 那么称集合

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

为 C 在边界点 x_0 处的支撑超平面.

根据定义, 点 x_0 和集合 C 也被该超平面分开.

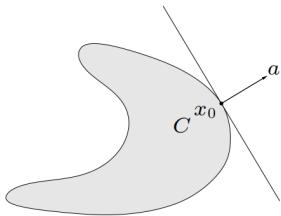
从集合上而言, 超平面 $\{x | a^T x = a^T x_0\}$ 与集合 C 在点 x_0 处相切, 并且半空间 $\{x | a^T x \leq a^T x_0\}$ 包含 C .

支撑超平面定理

注意根据凸集成立的分离超平面定理, 凸集上任何的边界点都满足支撑超平面存在的条件, 则对于凸集成立如下的定理.

定理

支撑超平面定理 若 C 是凸集, 则 C 的任意边界点处都存在支撑超平面.



支撑超平面定理有非常强的几何直观: 给定一个平面后, 可把凸集边界上的任意一点当成支撑点, 将凸集放在该平面上.

这也是凸集的特殊性质, 一般的集合甚至无法保证存在平面上的支撑点.