

第十章 优化基础

第 29 讲 凸集

黄定江

DaSE @ ECNU

djhuang@dase.ecnu.edu.cn

- ① 29.1 凸集
- ② 29.2 凸集的保凸运算
- ③ 29.3 凸集的性质：分离超平面定理

- ① 29.1 凸集
- ② 29.2 凸集的保凸运算
- ③ 29.3 凸集的性质：分离超平面定理

29.1.1 直线与线段

定义 1

对于 R^n 中的两个点 $x_1 \neq x_2$, 形如

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

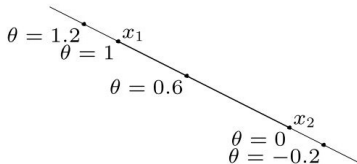
的点形成了过点 x_1 和 x_2 的直线。当 $0 \leq \theta \leq 1$ 时, 这样的点构成了连接点 x_1 和 x_2 的线段。

直线与线段

y 的表示形式

$$y = x_2 + \theta(x_1 - x_2)$$

给出了另一种解释：直线上的点 y 是基点 x_2 (对应 $\theta = 0$) 和方向 $x_1 - x_2$ (由 x_2 指向 x_1) 乘以参数 θ 的和。



当 θ 由 0 增加到 1, 点 y 相应地由 x_2 移动到 x_1 。如果 $\theta > 1$, 点 y 在超越了 x_1 的直线上。

29.1.2 仿射集

定义 2

如果通过集合 $\mathbb{C} \subseteq \mathbb{R}^n$ 中任意两点的直线仍然在集合 \mathbb{C} 中, 则称 \mathbb{C} 为仿射集。即:

$$\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{C} \Rightarrow \theta \boldsymbol{x}_1 + (1 - \theta) \boldsymbol{x}_2 \in \mathbb{C}, \forall \theta \in \mathbb{R}. \quad (1)$$

可以归纳得出:

- 一个仿射集包含其中任意点的仿射组合.

仿射集

- 如果 \mathbb{C} 是一个仿射集并且 $\boldsymbol{x}_0 \in \mathbb{C}$, 则集合

$$\mathbb{V} = \mathbb{C} - \boldsymbol{x}_0 = \{\boldsymbol{x} - \boldsymbol{x}_0 | \boldsymbol{x} \in \mathbb{C}\}$$

是一个子空间, 即关于加法和数乘是封闭的。

- 因此, 仿射集 \mathbb{C} 可以表示为

$$\mathbb{C} = \mathbb{V} + \boldsymbol{x}_0 = \{\boldsymbol{v} + \boldsymbol{x}_0 | \boldsymbol{v} \in \mathbb{V}\}$$

即一个子空间加上一个偏移。与仿射集 \mathbb{C} 相关联的子空间 \mathbb{V} 与 \boldsymbol{x}_0 的选取无关, 所以 \boldsymbol{x}_0 可以是 \mathbb{C} 中的任意一点。我们定义仿射集 \mathbb{C} 的维数为子空间 $\mathbb{V} = \mathbb{C} - \boldsymbol{x}_0$ 的维数, 其中 \boldsymbol{x}_0 是 \mathbb{C} 中的任意元素。

仿射包

定义 3

我们称由集合 $\mathbb{C} \subseteq \mathbb{R}^n$ 中的点的所有仿射组合组成的集合为 \mathbb{C} 的仿射包, 记为 $\text{aff } \mathbb{C}$:

$$\text{aff } \mathbb{C} = \{\theta_1 \mathbf{x}_1 + \cdots + \theta_k \mathbf{x}_k \mid \mathbf{x}_1, \cdots, \mathbf{x}_k \in \mathbb{C}, \theta_1 + \cdots + \theta_k = 1\}$$

仿射包

仿射包是包含 C 的最小的仿射集合，即 $\text{aff}C \subseteq S$ 。下图展示了 \mathbb{R}^3 中圆盘 S 的仿射包，为一个平面。

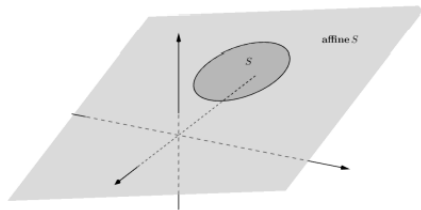


图 1

29.1.3 凸集

定义 4

如果连接集合 \mathbb{C} 中任意两点的线段都在 \mathbb{C} 内, 则称 \mathbb{C} 为凸集, 即

$$\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{C} \Rightarrow \theta \boldsymbol{x}_1 + (1 - \theta) \boldsymbol{x}_2 \in \mathbb{C}, \forall 0 \leq \theta \leq 1.$$

从仿射集的定义中可以看出仿射集是凸集。

凸集

例 1

下图显示了 \mathbb{R}^2 空间中一些简单的凸和非凸集合。

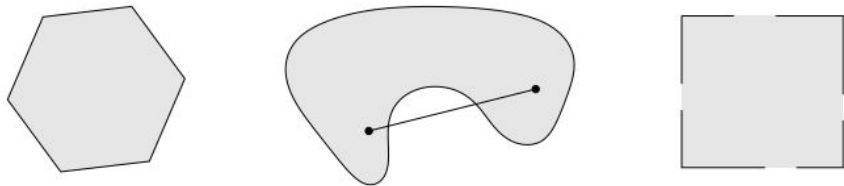


图 2: 左: 包含边界的六边形是凸的。中: 肾形集合不是凸的, 因为图中所示集合中两点间的线段不为集合所包含。右: 仅包含部分边界的正方形不是凸的。

凸组合和凸包

定义 5

形如

$$\boldsymbol{x} = \theta_1 \boldsymbol{x}_1 + \cdots + \theta_k \boldsymbol{x}_k, \quad 1 = \theta_1 + \theta_2 + \cdots + \theta_k, \theta_i \geq 0, i = 1, 2, \cdots, k$$

的点称为 $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k$ 的凸组合。集合 \mathbb{C} 中点所有可能的凸组合构成的集合称作 \mathbb{C} 的凸包，记作 $\text{conv } \mathbb{C}$ 。

凸组合和凸包

凸包是包含 \mathbb{C} 的最小的凸集。即 $\text{conv } \mathbb{C} \subseteq \mathbb{B}$ 。下图显示了凸包的定义。

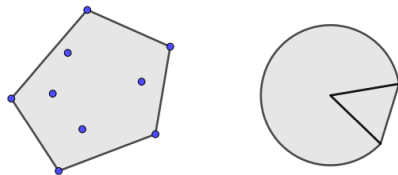


图 3

凸锥和锥组合

定义 6

如果对于任意 $x \in \mathbb{C}$ 和 $\theta \geq 0$ 都有 $\theta x \in \mathbb{C}$, 我们称集合 \mathbb{C} 是锥。

形如 $x = \theta_1 x_1 + \theta_2 x_2 \in \mathbb{C}, \theta_1 \geq 0, \theta_2 \geq 0$ 的点称为点 x_1, x_2 的锥组合。

若集合 \mathbb{C} 中任意点的锥组合都在 \mathbb{C} 中, 则称 \mathbb{C} 为凸锥。

在几何上, 具有此类形式的点构成了二维的扇形, 这个扇形以 0 为定点, 边通过 x_1 和 x_2 , 如下图所示:



29.1.4 一些简单的仿射集与凸集

首先介绍一些简单的例子。

- 空集, 任意一个点 (即单点集) $\{\mathbf{x}_0\}$ 、全空间 \mathbb{R}^n 都是 \mathbb{R}^n 的仿射 (自然也是凸的) 子集。
- 任意直线是仿射的。如果直线通过零点, 则是子空间, 因此, 也是凸锥。
- 一条线段是凸的, 但不是仿射的 (除非退化为一个点)。
- 一条射线, 即具有形式 $\{\mathbf{x}_0 + \theta\mathbf{v} | \theta \geq 0\}$, $\mathbf{v} \neq \mathbf{0}$ 的集合, 是凸的, 但不是仿射的。如果射线的基点 \mathbf{x}_0 是 $\mathbf{0}$, 则它是凸锥。
- 任意子空间是仿射的、凸锥 (自然是凸的)。

29.1.5 超平面与半空间

定义 7

任取非零向量 \mathbf{a} , 形如 $\{\mathbf{x} | \mathbf{a}^T \mathbf{x} = b\}$ 的集合称为超平面,
形如 $\{\mathbf{x} | \mathbf{a}^T \mathbf{x} \leq b\}$ 的集合称为半空间。

其中 $\mathbf{a} \in \mathbb{R}^n$, $\mathbf{a} \neq \mathbf{0}$ 是对应的超平面和半空间的法向量, 且 $b \in \mathbb{R}$ 。

- 解析地, 超平面是关于 x 的非平凡线性方程的解空间 (因此是一个仿射集合)。
- 一个超平面将 \mathbb{R}^n 分成两个半空间。
- 超平面是仿射集和凸集, 半空间是凸集但不是仿射集。

超平面与半空间

超平面和半空间的几何解释如下图所示：

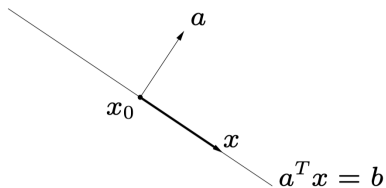


图 4: \mathbb{R}^2 中由法向量 a 和超平面上一点 x_0 确定的超平面。对于超平面上任意一点 x , $x - x_0$ (如深色箭头所示) 都垂直于 a 。

超平面与半空间

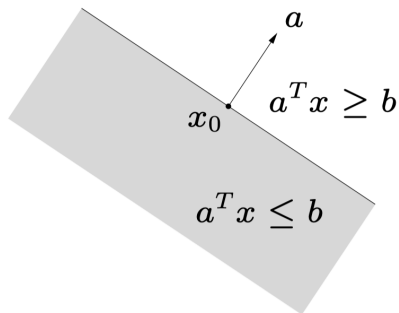


图 5: \mathbb{R}^2 上由 $a^T x = b$ 定义的超平面决定了两个半空间，由 $a^T x \geq b$ 决定的半空间（无阴影）是向 a 扩展的半空间，由 $a^T x \leq b$ 确定的半空间（阴影所示）向 $-a$ 方向扩展。向量 a 是这个半空间向外的法向量。

29.1.6 Euclid 球

定义 8

球是空间中到某个点距离（或两者差的范数）小于某个常数的点的集合，并将

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

称为中心为 \mathbf{x}_c ，半径为 r 的 **Euclid** 球。

29.1.7 椭球

定义 9

形如

$$\{\mathbf{x} | (\mathbf{x} - \mathbf{x}_c)^T \mathbf{P}^{-1} (\mathbf{x} - \mathbf{x}_c) \leq 1\}$$

的集合称为椭球，其中 $\mathbf{P} \in \mathcal{S}_{++}^n$ (即 \mathbf{P} 对称正定)。椭球的另一种表示为

$$\{\mathbf{x}_c + \mathbf{A}\mathbf{u} | \|\mathbf{u}\|_2 \leq 1\},$$

其中 \mathbf{A} 为非奇异的方阵。

29.1.8 范数球与范数锥

定义 10

设 $\|\cdot\|$ 是 \mathbb{R}^n 中的范数。称

$$\{\mathbf{x} \mid \|\mathbf{x} - \mathbf{x}_c\| \leq r\}$$

为以 r 为半径, \mathbf{x}_c 为球心的范数球。

定义 11

关于范数 $\|\cdot\|$ 的范数锥是集合

$$C = \{(\mathbf{x}, t) \mid \|\mathbf{x}\| \leq t\} \subseteq \mathbb{R}^{n+1}$$

顾名思义, 它是一个凸锥。

29.1.9 多面体

定义 12

我们将满足线性等式和不等式组的点的集合称为多面体，即

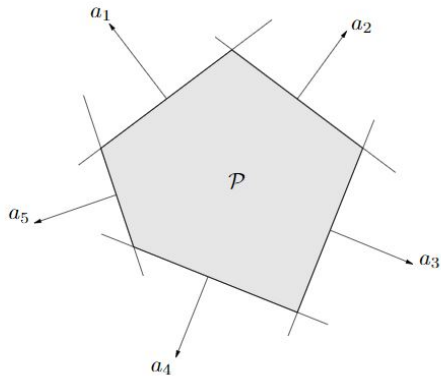
$$\{x | Ax \leq b, Cx = d\},$$

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

- 易得出多面体是有限个半空间和超平面的交集。
- 仿射集合、射线、线段和半空间都是多面体。
- 有界的多面体有时也称为多胞形。

多面体

下图显示了一个由五个半空间的交集定义的多面体。



29.1.10 半正定锥

- 我们用 S^m 表示对称 $n \times n$ 矩阵的集合, 即

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

这是一个维数为 $n(n+1)/2$ 的向量空间。

- 我们用 S_+^m 表示对称半正定矩阵的集合:

$$S_+^m = \{\mathbf{X} \in S^m | \mathbf{X} \succeq 0\}$$

- 用 S_{++}^m 表示对称正定矩阵集合:

$$S_{++}^m = \{\mathbf{X} \in S_+^m | \mathbf{X} \succ 0\}$$

半正定锥

容易证明集合 S_+^m 是一个凸锥，因此

定义 13

我们称

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

为半正定锥。

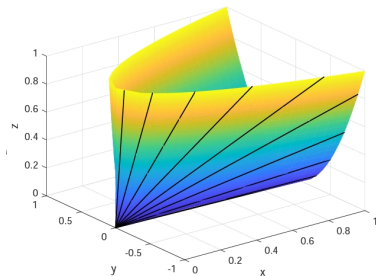


图 6: 二维半正定锥 S_+^2

- 1 29.1 凸集
- 2 29.2 凸集的保凸运算
- 3 29.3 凸集的性质：分离超平面定理

保凸运算

判定一个集合为凸集的方式:

1. 利用定义

$$\mathbf{x}_1, \mathbf{x}_2 \in C, 0 \leq \theta \leq 1 \Rightarrow \theta \mathbf{x}_1 + (1 - \theta) \mathbf{x}_2 \in C$$

2. 说明集合 C 可以由简单的凸集经过保凸运算后得到。

本节下面将介绍一些常见的保凸运算:

- 取交集
- 仿射变换
- 线性分式及透视函数

29.2.1 交集

定理 1

任意多个凸集的交为凸集，即若 $C_i, i \in \mathcal{L}$ 是凸集，则

$$\bigcap_{i \in \mathcal{L}} C_i$$

为凸集。这里 \mathcal{L} 是任意指标集（不要求可列）。

交集

例 2

- 半正定锥 S_+^m 可以表示为,

$$\cap_{z \neq 0} \{X \in S^m | z^T X z \geq 0\}.$$

对于任意 $z \neq 0$, $z^T X z$ 是关于 X 的 (不恒等于零的) 线性函数, 因此集合

$$\{X \in S^m | z^T X z \geq 0\}$$

实际上就是 S^m 的半空间。由此可见, 半正定锥是无穷个半空间的交集, 因此是凸的.

29.2.2 凸集的和

- 两个集合的和可以定义为:

$$S_1 + S_2 = \{\mathbf{x} + \mathbf{y} | \mathbf{x} \in S_1, \mathbf{y} \in S_2\}$$

如果 S_1 和 S_2 是凸集, 那么, $S_1 + S_2$ 是凸的。

29.2.3 凸集的 Cartesian 乘积

- 可以看出, 如果 S_1 和 S_2 是凸的, 那么其直积或 Cartesian 乘积

$$S_1 \times S_2 = \{(\mathbf{x}_1, \mathbf{x}_2) | \mathbf{x}_1 \in S_1, \mathbf{x}_2 \in S_2\}$$

也是凸集。

29.2.4 仿射变换

定理 2

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

(1) 凸集在 f 下的像是凸集:

$$S \subseteq \mathbb{R}^n \text{ 为凸集} \Rightarrow f(S) \stackrel{\text{def}}{=} \{f(\mathbf{x}) | \mathbf{x} \in S\} \text{ 为凸集};$$

(2) 凸集在 f 下的原像是凸集:

$$C \subseteq \mathbb{R}^m \text{ 为凸集} \Rightarrow f^{-1}(C) \stackrel{\text{def}}{=} \{\mathbf{x} \in \mathbb{R}^n | f(\mathbf{x}) \in C\} \text{ 为凸集}.$$

仿射变换

- 两个简单的例子是伸缩和平移。如果 $S \subseteq \mathbb{R}^n$ 是凸集, $\alpha \in \mathbb{R}$ 并且 $\mathbf{a} \in \mathbb{R}^n$, 那么, 集合 αS 和 $S + \mathbf{a}$ 是凸的, 其中

$$\alpha S = \{\alpha \mathbf{x} | \mathbf{x} \in S\}, \quad S + \mathbf{a} = \{\mathbf{x} + \mathbf{a} | \mathbf{x} \in S\}$$

- 一个凸集向它的某几个坐标的投影是凸的, 即: 如果 $S \subseteq \mathbb{R}^m \times \mathbb{R}^n$ 是凸集, 那么

$$T = \{\mathbf{x}_1 \in \mathbb{R}^m | (\mathbf{x}_1, \mathbf{x}_2) \in S \text{ 对于某些 } \mathbf{x}_2 \in \mathbb{R}^n\}$$

是凸集。

仿射变换

例 3

利用仿射变换保凸的性质，可以证明：

线性矩阵不等式的解集

$$\{\mathbf{x} | x_1 \mathbf{A}_1 + x_2 \mathbf{A}_2 + \cdots + x_m \mathbf{A}_m \preceq \mathbf{B}\}$$

是凸集，其中 $\mathbf{A}_i, \mathbf{B} \in S^m$ 。因为，它可以看作是一个仿射变换的原像

双曲锥

$$\{\mathbf{x} | \mathbf{x}^T \mathbf{P} \mathbf{x} \leq (\mathbf{c}^T \mathbf{x})^2, \mathbf{c}^T \mathbf{x} \geq 0\}$$

是凸集，其中 $\mathbf{P} \in S_+^m$ 。因为，它可以看作是 $\mathbf{x} \rightarrow (\mathbf{P}^{1/2} \mathbf{x}, \mathbf{c}^T \mathbf{x})$ 变换下的原像，而值域是凸锥。

29.2.5 透视函数

定义 14

我们定义 $P: \text{dom } P \rightarrow \mathbb{R}^n$, $P(z, t) = z/t$ 为透视函数, 其定义域为 $\text{dom } P = \mathbb{R}^n \times \mathbb{R}_+$ 。透视函数对向量进行伸缩, 或称为规范化, 使得最后一维分量为 1 并舍弃之。

透视函数

- 如果 $C \subseteq \text{dom } P$ 是凸集，那么它的像

$$P(C) = \{P(\mathbf{x}) | \mathbf{x} \in C\}$$

也是凸集。

- 一个凸集在透视函数下的原象也是凸的: 如果 $C \subseteq \mathbb{R}^n$ 为凸集，那么

$$P^{-1}(C) = \{(\mathbf{x}, t) \in \mathbb{R}^{n+1} | \mathbf{x}/t \in C, t > 0\}$$

是凸集。

29.2.6 线性分式函数

- 线性分式函数由透视函数和仿射函数复合而成。设 $g: \mathbb{R}^n \rightarrow \mathbb{R}^{m+1}$ 是仿射的，即

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

其中 $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{c} \in \mathbb{R}^n$ 并且 $d \in \mathbb{R}$ 。则由 $f = P \circ g$ 给出的函数 $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$

$$f(\mathbf{x}) = (\mathbf{Ax} + \mathbf{b})/(\mathbf{c}^T \mathbf{x} + d), \text{dom } f = \{\mathbf{x} | \mathbf{c}^T \mathbf{x} + d > 0\}$$

称为线性分式 (或投射) 函数。如果 $\mathbf{c} = \mathbf{0}$, $d > 0$ 。则 f 的定义域为 \mathbb{R}^n ，并且 f 是仿射函数。因此，我们可以将仿射和透视函数视为特殊的线性分式函数。

线性分式函数

- 类似于透视函数，线性分式函数也是保凸的。

如果 C 是凸集并且在 f 的定义域中 (即任意 $\mathbf{x} \in C$ 满足 $\mathbf{c}^T \mathbf{x} + d > 0$) , 那么 C 的象 $f(C)$ 也是凸集。

根据前述的结果可以直接得到这个结论: C 在仿射映射下的象是凸的, 并且在透视函数 P 下的映射 (即 $f(C)$) 是凸的。

- 类似地, 如果 $C \subseteq \mathbb{R}^m$ 是凸集, 那么其原象 $f^{-1}(C)$ 也是凸的。

- 1 29.1 凸集
- 2 29.2 凸集的保凸运算
- 3 29.3 凸集的性质：分离超平面定理

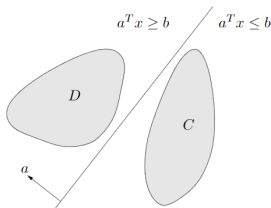
29.3.1 分离与支撑超平面

定理 3

超平面分离定理: 假设 C 和 D 是两个不相交的凸集, 则存在非零向量 a 和常数 b , 使得

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

即超平面 $\{x | a^T x = b\}$ 分离了 C 和 D 。如下图所示:



29.3.2 分离超平面定理

严格分离（即上式成立严格不等号）需要更强的假设。例如当 C 是闭凸集， D 是单点集时，我们有如下严格分离定理：

定理 4

严格分离定理：设 C 是闭凸集，点 $x_0 \notin C$ ，则存在非零向量 a 和常数 b ，使得

$$a^T x < b, \forall x \in C \text{ 且 } a^T x_0 > b.$$

29.3.3 支撑超平面

当点 x_0 恰好在凸集 C 的边界上时，可以构造支撑超平面。

定义 15

给定集合 C 及其边界上一点 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。

从几何上来说，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

根据凸集的分离超平面定理，我们有如下支撑超平面定理：

定理 5

如果 C 是凸集，则在 C 的任意边界点处都存在支撑超平面。

支撑超平面定理从几何上看，可以理解为：

给定一个平面后，可以把凸集边界上的任意一点当成支撑点将凸集放置在该平面上。

本讲小结

凸集

- 仿射集与仿射包
- 凸集与凸包
- 常见的凸集与仿射集

凸集的保凸运算

- 保凸运算：交集、和、Cartesian 乘积、仿射变换、透视函数、线性分式函数
- 凸集的性质：分离超平面定理

凸集是凸优化中很重要的概念。通常很难从定义判定一个集合是否是凸集，可以借助保凸运算进行判断。