第 2 章：数学基础

(1)：凸集

曹语

课程主页：https://yucaoyc.github.io/math3806

目录

背景和目标

部分优化问题相对而言稍微容易一些:

$$\min_{x \in \Omega} \ f_0(x)$$

若 Ω 为凸集，f_0 为凸函数，最优性条件以及优化算法的效率能够得到保证

目标: 在本章，我们需要学习如下问题:

- 什么是凸集? （本节课）
- 什么是凸函数?
- 如果 Ω 为凸集，f_0 为凸函数，最优解有什么性质?

目录

直线、线段和射线

倘若 $\boldsymbol{x}_1 \neq \boldsymbol{x}_2$ 为 $\mathbb{R}^n$ 空间中的两个点

两点构成的直线：

$$\boldsymbol{y} = \theta \boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2, \qquad \theta \in \mathbb{R}$$

两点构成的线段：

$$\boldsymbol{y} = \theta \boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2, \qquad \theta \in [0,1]$$

通过某个点的射线：

$$\boldsymbol{y} = \theta \boldsymbol{x}, \qquad \theta \geq 0$$

接下来，我们将基于这三个基本概念，拓展出仿射、凸集、锥等概念

仿射

定义 (仿射/Affine 的等价定义)

集合 C 是仿射的可以被（等价）定义为如下的几个方式:

(1) 如果通过集合 $C \subseteq \mathbb{R}^n$ 中任意两个不同点的直线仍然在集合 C 中。

(2) 若对任意的 $\boldsymbol{x}_1, \boldsymbol{x}_2 \in C$，$\theta \in \mathbb{R}$，皆有 $\theta \boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2 \in C$。

(3) $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k \in C$ ，且 $\theta_1 + \cdots + \theta_k = 1$ ，则

$$\sum_{i=1}^{k} \theta_i \boldsymbol{x}_i = \theta_1 \boldsymbol{x}_1 + \cdots + \theta_k \boldsymbol{x}_k \in C$$

(4) $C = V + \boldsymbol{x}_0 = \{\boldsymbol{v} + \boldsymbol{x}_0 \mid \boldsymbol{v} \in V\}$，其中 $\boldsymbol{x}_0$ 是 C 中任意的一点，V 是一个子空间（即对于加法和乘法封闭）。
并且，定义仿射集合 C 的维数为子空间 V 的维数。

例子： 线性方程组的解 $\{\boldsymbol{x} \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}\}$ 是仿射的。

即线性等式约束所涉及到的集合

仿射包/affine hull 由集合 $C \subseteq \mathbb{R}^n$ 中的点的所有仿射组合组成的集合被称为 C 的仿射包，记为 aff C：

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

性质： 仿射包是包含 C 的最小的仿射集合

仿射维数 集合 C 的仿射维数被定义为其仿射包的维数。

例子： 在 $\mathbb{R}^2$ 上的圆的仿射维数为 2

凸集

定义 (凸集/convex set)

C 是凸集具有如下的等价定义:

(1) C 中任意两点间的线段仍然在 C 中。

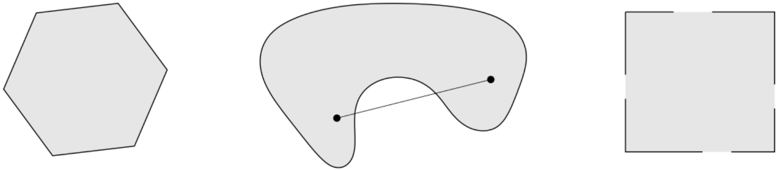
(2) 对于任意 $\boldsymbol{x}_1, \boldsymbol{x}_2 \in C$ 和 $0 \leqslant \theta \leqslant 1$ 都有

$$\theta \boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2 \in C$$

(3) 若 $\theta_i \geqslant 0, i = 1, \cdots, k$ 且 $\sum_{i=1}^{k} \theta_i = 1$，则 $\theta_1 \boldsymbol{x}_1 + \cdots + \theta_k \boldsymbol{x}_k \in C$

【该线性组合 $\sum_{i=1}^{k} \theta_i \boldsymbol{x}_i$ 被称为 $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k$ 的一个凸组合】

例子： 请判断如下的集合是否为凸集：



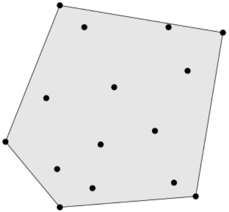
定义 (凸包/convex hull)

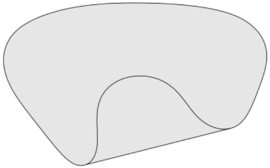
集合 C 中所有点的凸组合的集合被称为其**凸包，记为** $\operatorname{conv} C$

$$\operatorname{conv} C = \{\theta_1 \boldsymbol{x}_1 + \cdots + \theta_k \boldsymbol{x}_k \mid \boldsymbol{x}_i \in C, \theta_i \geqslant 0, i = 1, \cdots, k, \theta_1 + \cdots + \theta_k = 1\}.$$

性质： C 的凸包是包含 C 的最小的凸集。

凸包的例子：





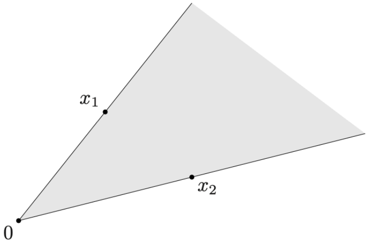
提示： 凸集的定义可以拓展至级数，以及积分形式（见课本 2.1.4）。

锥

定义 (锥/Cone)

(1) 若对于任意 $\boldsymbol{x} \in C$ 和 $\theta \geqslant 0$ 都有 $\theta \boldsymbol{x} \in C$，则称集合 C 是锥（或者非负齐次）。

(2) 若集合 C 是锥，且是凸集，则称 C 为凸锥：即对于任意 $\boldsymbol{x}_1, \boldsymbol{x}_2 \in C$ 和 $\theta_1, \theta_2 \geqslant 0$ ，都有 $\theta_1 \boldsymbol{x}_1 + \theta_2 \boldsymbol{x}_2 \in C$



提示：锥一定包含原点

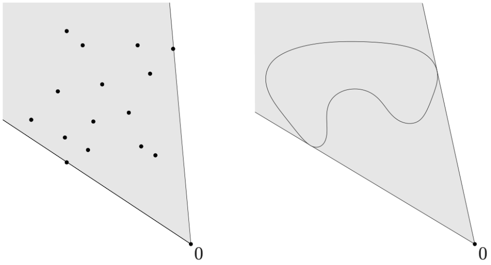
问题：能否举出一个锥，且其非凸?

定义 (锥包/conic hull)

集合 C 的锥包是 C 中元素的所有锥组合的集合，即

$$\{\theta_1 \boldsymbol{x}_1 + \cdots + \theta_k \boldsymbol{x}_k \mid \boldsymbol{x}_i \in C, \theta_i \geqslant 0, i = 1, \cdots, k\}$$

它是包含 C 的最小的凸锥



总结

两点构成的直线：

$$\theta \boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2, \qquad \theta \in \mathbb{R}$$

基于此，我们定义出仿射，和仿射包

两点构成的线段：

$$\theta \boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2, \qquad \theta \in [0,1]$$

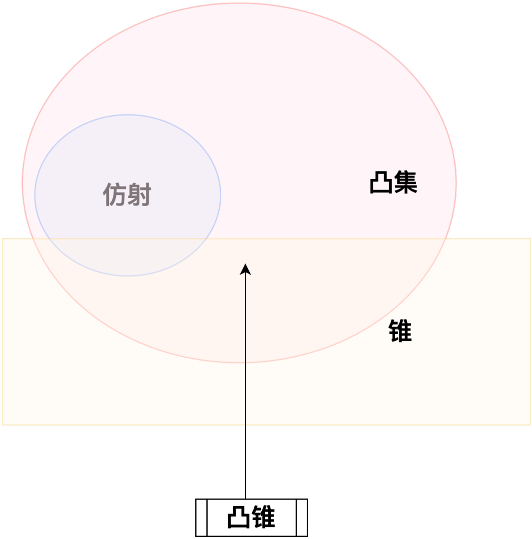
基于此，我们定义出凸集，和凸包

通过某个点的射线：

$$\theta \boldsymbol{x}, \qquad \theta \geq 0$$

基于此，我们定义出锥，和锥包

多个概念之间的关联



练习

请判断如下的集合是否属于对应的 4 个类别:

集合	仿射	凸集	锥	凸锥
空集 $\emptyset$				
单点集 $\{x_0\}$				
不穿过原点的直线				
穿过原点的直线				

前一页的答案

集合	仿射	凸集	锥	凸锥
空集 $\emptyset$	是	是	是	是
单点集 $\{x_0\}$	是	是	不一定	/
不穿过原点的直线	是	是	不是	不是
穿过原点的直线	是	是	是	是

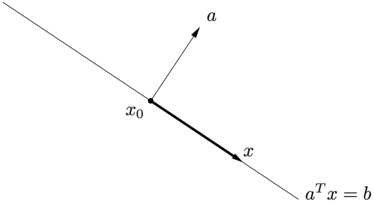
目录

(1) 超平面和半空间

超平面：对于 $\boldsymbol{a} \in \mathbb{R}^n$, $\boldsymbol{a} \neq \mathbf{0}$, $b \in \mathbb{R}$,

$$\{\boldsymbol{x} \mid \boldsymbol{a}^\top \boldsymbol{x} = b\}$$

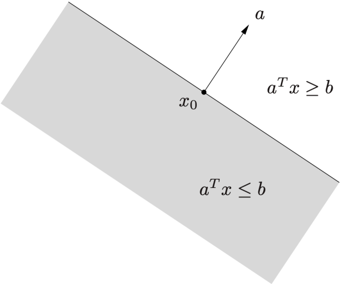
是仿射、是凸集、不是锥



（闭的）半空间：对于 $\boldsymbol{a} \in \mathbb{R}^n$, $\boldsymbol{a} \neq \mathbf{0}$, $b \in \mathbb{R}$,

$$\{\boldsymbol{x} \mid \boldsymbol{a}^\top \boldsymbol{x} \leq b\}$$

不是仿射、是凸集、不是锥



(2) Euclid 球和椭球

Euclid 球

$$
\begin{aligned}
B(\boldsymbol{x}_c, r) &= \{\boldsymbol{x} \ \mid \ \|\boldsymbol{x}-\boldsymbol{x}_c\|_2 \leqslant r\} = \left\{\boldsymbol{x} \mid (\boldsymbol{x}-\boldsymbol{x}_c)^\top (\boldsymbol{x}-\boldsymbol{x}_c) \leqslant r^2\right\} \\
&= \{\boldsymbol{x}_c + r\boldsymbol{u} \mid \|\boldsymbol{u}\|_2 \leqslant 1\}
\end{aligned}
$$

提示：它是凸集可由三角不等式证明。

椭球

$$
\begin{aligned}
\mathcal{E} &= \left\{\boldsymbol{x} \mid (\boldsymbol{x}-\boldsymbol{x}_c)^\top \boldsymbol{P}^{-1} (\boldsymbol{x}-\boldsymbol{x}_c) \leqslant 1\right\} \\
&= \left\{\boldsymbol{x}_c + \boldsymbol{P}^{1/2}\boldsymbol{u} \mid \|\boldsymbol{u}\|_2 \leq 1\right\}
\end{aligned}
$$

其中 $\boldsymbol{P} = \boldsymbol{P}^\top \succ 0$ ，即 $\boldsymbol{P}$ 是对称正定矩阵

更一般地可以定义范数球:

$$B(\boldsymbol{x}_c, r) = \{\boldsymbol{x} \mid \|\boldsymbol{x} - \boldsymbol{x}_c\| \leqslant r\}$$

范数锥:

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

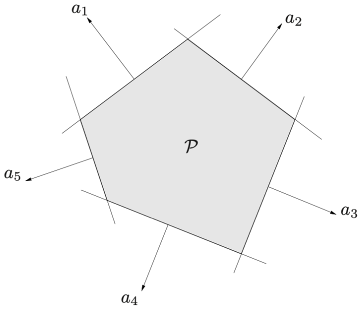
练习: 证明范数锥是凸锥

(3) 多面体/polyhedron

多面体是有限个半空间和超平面的交集：

$$\mathcal{P} = \left\{ \boldsymbol{x} \;\middle|\; \begin{array}{l} \boldsymbol{a}_j^\top \boldsymbol{x} \leqslant b_j,\ j = 1, \cdots, m, \\ \boldsymbol{c}_j^\top \boldsymbol{x} = d_j,\ j = 1, \cdots, p \end{array} \right\} \overset{\text{简写为}}{=} \left\{ \boldsymbol{x} | \boldsymbol{A}\boldsymbol{x} \preceq \boldsymbol{b}, \boldsymbol{C}\boldsymbol{x} = \boldsymbol{d} \right\}$$

即一般的线性规划问题所涉及到的集合



$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a}_1^\top \\ \boldsymbol{a}_2^\top \\ \vdots \\ \boldsymbol{a}_m^\top \end{bmatrix}, \qquad \boldsymbol{C} = \begin{bmatrix} \boldsymbol{c}_1^\top \\ \boldsymbol{c}_2^\top \\ \vdots \\ \boldsymbol{c}_p^\top \end{bmatrix}$$

由定义可知，超平面、半空间都是特殊的多面体

练习： 证明多面体是凸集

单纯形/Simplex（特殊的多面体）

设 $k+1$ 个点 $\boldsymbol{v}_0, \cdots, \boldsymbol{v}_k \in \mathbb{R}^n$ 仿射独立，即 $\boldsymbol{v}_1 - \boldsymbol{v}_0, \cdots, \boldsymbol{v}_k - \boldsymbol{v}_0$ 线性独立，则这些点决定了一个单纯形：

$$C = \operatorname{conv}\{\boldsymbol{v}_0, \cdots, \boldsymbol{v}_k\} = \left\{\theta_0 \boldsymbol{v}_0 + \cdots + \theta_k \boldsymbol{v}_k \mid \boldsymbol{\theta} \succeq 0,\ \mathbf{1}^\top \boldsymbol{\theta} = 1\right\}$$

该单纯形的仿射维数为 k

- $k=1$，线段
- $k=2$，三角形
- $k=3$，四面体

例子：证明如上定义的单纯形是一个有界的多面体。

提示：该集合是紧集，所以在单纯形中优化连续函数一定存在极值。

(4) 对称矩阵

如下是几个和矩阵空间相关的凸集:

- n 阶对称矩阵的集合:

$$\mathbf{S}^n = \{\boldsymbol{X} \in \mathbb{R}^{n\times n} \mid \boldsymbol{X} = \boldsymbol{X}^\top\}$$

- n 阶对称半正定矩阵:

$$\mathbf{S}_+^n = \{\boldsymbol{X} \in \mathbb{R}^{n\times n} \mid \boldsymbol{X} = \boldsymbol{X}^\top, \boldsymbol{X} \succeq 0\}$$

（它是一个**凸锥**）

- n 阶对称半正定矩阵:

$$\mathbf{S}_{++}^n = \{\boldsymbol{X} \in \mathbb{R}^{n\times n} \mid \boldsymbol{X} = \boldsymbol{X}^\top, \boldsymbol{X} \succ 0\}$$

目录

总结

- 知道仿射、凸集、锥等基本概念的定义
- 掌握介绍过的 4 类重要例子；能判断其类别

阅读作业 & 参考资料：

- 课本第 2.1，2.2 章