

最优化方法

Methods of Optimization

目录

第一章 基础知识 1

1.1 数学规划 1

1.2 基本概念和符号 3

1.3 子空间、正交子空间 5

1.4 多元函数及其导数 7

1.5 凸集、凸函数和凸规划 8

1.5.1 凸集 8

1.5.2 凸函数 11

1.5.3 凸规划 11

第一章 基础知识

1.1 数学规划

规划论，又称数学规划，是运筹学的分支，研究在约束条件下通过分配资源寻求目标函数极值的数学方法，应用于经济管理、工程设计和过程控制等领域。其核心是建立约束条件与目标函数的数学模型，这种数学模型的一般形式为

$$\begin{aligned} \text{Opt. } & f(x_i, y_j, \xi_k) \\ \text{s.t. } & g_h(x_i, y_j, \xi_k) \leq (=, \geq) 0 \\ & h = 1, 2, \dots, m \end{aligned}$$

其中， x_i 是控制变量， y_j 是已知参数， ξ_k 是随机因素， f 是目标函数， g_h 是约束函数。

按照目标函数和自变量取值范围的不同，数学规划可分为：

- 线性规划： f, g 都是线性函数；
- 非线性规划： f, g 中有非线性函数；
- 多目标规划： f 是向量函数；
- 整数规划：决策变量 x_i 为整数；
- 动态规划：含多阶段决策过程；
- 随机规划：含有随机因子。

若记

$$S = \{\mathbf{x} \mid g_h(x_i, y_j, \xi_k) \leq (=, \geq) 0\} \subseteq \mathbb{R}^n$$

为约束集合，则这种数学模型可进一步简记为

$$(f \ S) \begin{cases} \min & f(\mathbf{x}) \\ \text{s.t.} & \mathbf{x} \in S \end{cases}$$

其中 $\mathbf{x} \in S$ 称模型 $(f \ S)$ 的可行解。

定义 1.1.（全局最优解、最优值）

若对于 $\mathbf{x}^* \in S$ ，满足 $f(\mathbf{x}^*) \leq f(\mathbf{x})$ ， $\forall \mathbf{x} \in S$ ，则称 \mathbf{x}^* 为 $(f \ S)$ 的全局最优解（最优解），记作 g.opt.，简记为 opt.。称此时的 $f(\mathbf{x}^*)$ 为 $(f \ S)$ 的最优值（最优目标函数值）。

定义 1.2.（局部最优解）

若对于 $\mathbf{x}^* \in S$, $\exists \mathbf{x}^*$ 的某邻域 $U(\mathbf{x}^*)$, 使得 $f(\mathbf{x}^*) \leq f(\mathbf{x})$, $\forall \mathbf{x} \in S$, 则称 \mathbf{x}^* 为 (f, S) 的局部最优解, 记作 l.opt.,

在上述定义中, 若当 $\mathbf{x} \neq \mathbf{x}^*$ 时有严格不等式成立, 则分别称 \mathbf{x}^* 为 (f, S) 的严格全局最优解和严格局部最优解.

1.2 基本概念和符号

接下来将约定一些术语的符号表示:

- \mathbb{R}^n —— n 维欧氏空间;
- $\mathbf{x} = (x_1, x_2, \dots, x_n)^\top \in \mathbb{R}^n$ ——表示 \mathbb{R}^n 中的一个点或一个向量, 其中分量 $x_i \in \mathbb{R}$;
- $\mathbf{d} = (d_1, d_2, \dots, d_n)^\top \in \mathbb{R}^n$, 其中 $\mathbf{d} \neq \mathbf{0}$ ——表示从 $\mathbf{0}$ 指向 \mathbf{d} 的方向;
- $\mathbf{x} + \lambda \mathbf{d}$ ——表示从 \mathbf{x} 点出发, 沿 \mathbf{d} 方向移动 $\lambda \mathbf{d}$ 长度得到的点.

下面是对欧氏空间中向量的运算的定义.

定义 1.3. (向量的运算)

设 $\mathbf{x} = (x_1, x_2, \dots, x_n)^\top \in \mathbb{R}^n$, $\mathbf{y} = (y_1, y_2, \dots, y_n)^\top \in \mathbb{R}^n$, 则

(1) 向量的内积运算: $\mathbf{x}^\top \mathbf{y} = \mathbf{y}^\top \mathbf{x} = \sum_{i=1}^n x_i y_i$;

(2) 向量长度: $\|\mathbf{x}\| = \sqrt{\mathbf{x}^\top \mathbf{x}} = \sqrt{\sum_{i=1}^n x_i^2}$;

(3) 两点间距离: $\|\mathbf{x} - \mathbf{y}\|$.

定理 1.1. (三角不等式)

$\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, 有

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\| \quad (1.1)$$

【证明】对左式和右式分别平方后相减, 有

$$\begin{aligned} \text{左式}^2 - \text{右式}^2 &= (\mathbf{x} + \mathbf{y})^\top (\mathbf{x} + \mathbf{y}) - (\mathbf{x}^\top \mathbf{x} + \mathbf{y}^\top \mathbf{y} + 2\|\mathbf{x}\|\|\mathbf{y}\|) \\ &= (\mathbf{x}^\top \mathbf{x} + \mathbf{y}^\top \mathbf{y} + 2\mathbf{x}^\top \mathbf{y}) - (\mathbf{x}^\top \mathbf{x} + \mathbf{y}^\top \mathbf{y} + 2\|\mathbf{x}\|\|\mathbf{y}\|) \\ &= 2(\mathbf{x}^\top \mathbf{y} - \|\mathbf{x}\|\|\mathbf{y}\|) \end{aligned}$$

也即只需证明柯西不等式

$$\mathbf{x}^\top \mathbf{y} \leq \|\mathbf{x}\|\|\mathbf{y}\|$$

成立即可. 于是我们构造关于 t 的方程

$$\|\mathbf{x} + t\mathbf{y}\|^2 = \mathbf{y}^\top \mathbf{y} t^2 + 2\mathbf{x}^\top \mathbf{y} t + \mathbf{x}^\top \mathbf{x} = 0$$

显然由于向量的模长一定非负, 故该方程至多有一个解, 也即该二次方程的判别式

$$\Delta = 4(\mathbf{x}^\top \mathbf{y})^2 - 4(\mathbf{x}^\top \mathbf{x})(\mathbf{y}^\top \mathbf{y}) \leq 0$$

从而柯西不等式得证, 从而原不等式得证.

【证毕】

定义 1.4. (点列的收敛)

设 $\{\mathbf{x}^{(n)}\}$ 是定义在 \mathbb{R}^n 上的点列, 若 $\exists \mathbf{x} \in \mathbb{R}^n$, 使得 $\lim_{n \rightarrow \infty} \|\mathbf{x}^{(n)} - \mathbf{x}\| = 0$, 则称点列 $\{\mathbf{x}^{(n)}\}$ 收敛到 \mathbf{x} , 记作

$$\lim_{n \rightarrow \infty} \mathbf{x}^{(n)} = \mathbf{x}$$

定义 1.5. (向量的大小关系)

设 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, 若 $\forall i \in \{1, 2, \dots, n\}$, 都有 $x_i \leq y_i$, 则称 $\mathbf{x} \leq \mathbf{y}$. 类似可规定 $\mathbf{x} \geq \mathbf{y}$, $\mathbf{x} < \mathbf{y}$, $\mathbf{x} > \mathbf{y}$.

显然, 在如上定义中, (\mathbb{R}^n, \leq) 并不是全序集 (证明略). 特别地, $\mathbf{x} \leq \mathbf{0} \Leftrightarrow \forall i \in \{1, 2, \dots, n\}, x_i \leq 0$.

例 1.1 设 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, 若 $\forall \mathbf{y} \geq \mathbf{0}$, 都有 $\mathbf{x}^\top \mathbf{y} \leq \alpha$, 则有 $\mathbf{x} \leq \mathbf{0}$ 且 $\alpha \geq 0$.

【证明】 取 $\mathbf{y} = \mathbf{0}$, 则 $\mathbf{x}^\top \mathbf{y} = 0$, 从而 $\mathbf{x}^\top \mathbf{y}$ 的上界 $\alpha \geq 0$. 假设 $\mathbf{x} \leq \mathbf{0}$ 不成立, 则 $\exists i \in \{1, 2, \dots, n\}$, 使得 $x_i > 0$. 令 $y_i \rightarrow +\infty$, 则有 $\mathbf{x}^\top \mathbf{y} \rightarrow +\infty$, 从而与 $\mathbf{x}^\top \mathbf{y}$ 有上界 α 矛盾, 故必有 $\mathbf{x} \leq \mathbf{0}$. **【证毕】**

该命题还有以下等价表述:

- 若 $\forall \mathbf{y} \leq \mathbf{0}$, 都有 $\mathbf{x}^\top \mathbf{y} \leq \alpha$, 则有 $\mathbf{x} \geq \mathbf{0}$ 且 $\alpha \geq 0$;
- 若 $\forall \mathbf{y} \geq \mathbf{0}$, 都有 $\mathbf{x}^\top \mathbf{y} \geq \alpha$, 则有 $\mathbf{x} \geq \mathbf{0}$ 且 $\alpha \leq 0$;
- 若 $\forall \mathbf{y} \leq \mathbf{0}$, 都有 $\mathbf{x}^\top \mathbf{y} \geq \alpha$, 则有 $\mathbf{x} \leq \mathbf{0}$ 且 $\alpha \leq 0$.

1.3 子空间、正交子空间

定义 1.6. (子空间)

设 $\mathbf{d}^{(i)} \in \mathbb{R}^n, i = 1, 2, \dots, m$ 是线性无关的向量组, 称集合 $\{\mathbf{x} \mid \mathbf{x} = \sum_{j=1}^m \alpha_j \mathbf{d}^{(j)}, \alpha_j \in \mathbb{R}\}$ 是由向量 $\mathbf{d}^{(1)}, \mathbf{d}^{(2)}, \dots, \mathbf{d}^{(m)}$ 生成的子空间, 记作 $L(\mathbf{d}^{(1)}, \mathbf{d}^{(2)}, \dots, \mathbf{d}^{(m)})$, 简记作 L .

可以看到, $\forall L \subseteq \mathbb{R}^n$, 都有 $\mathbf{0} \in L$. 若令 $\mathbf{A} = (\mathbf{d}^{(1)}, \mathbf{d}^{(2)}, \dots, \mathbf{d}^{(m)}) \in \mathbb{R}^{n \times m}$, $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_m)^\top \in \mathbb{R}^m$, 则

$$\mathbf{x} \in L \Leftrightarrow \exists \boldsymbol{\alpha} \in \mathbb{R}^m, \mathbf{x} = \mathbf{A}\boldsymbol{\alpha}$$

并且 $\text{Rank}(\mathbf{A}) = m$, 也就是说 \mathbf{A} 是列满秩的. 我们将会在接下来的证明中用到这个性质.

定义 1.7. (正交子空间)

称集合 $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x}^\top \mathbf{y} = 0, \forall \mathbf{y} \in L\}$ 为子空间 L 的正交子空间, 记作 L^\perp .

不难看出 $\forall L \subseteq \mathbb{R}^n$, 都有 $\mathbf{0} \in L^\perp$. 且有

$$\begin{aligned} \mathbf{x} \in L^\perp &\Leftrightarrow \forall \mathbf{y} \in L, \mathbf{x}^\top \mathbf{y} = 0 \\ &\Leftrightarrow \forall \boldsymbol{\alpha} \in \mathbb{R}^m, \mathbf{x}^\top \mathbf{A}\boldsymbol{\alpha} = 0 \end{aligned}$$

定理 1.2. (子空间投影定理)

设 L 为 \mathbb{R}^n 的子空间, 则 $\forall \mathbf{z} \in \mathbb{R}^n$, 则必有以下两个命题成立:

(1) $\exists! \mathbf{x} \in L, \mathbf{y} \in L^\perp$, 使得 $\mathbf{z} = \mathbf{x} + \mathbf{y}$;

(2) 在 (1) 中的 \mathbf{x} 为问题

$$\begin{aligned} \min_{\mathbf{u}} \quad & \|\mathbf{z} - \mathbf{u}\| \\ \text{s.t.} \quad & \mathbf{u} \in L \end{aligned}$$

的唯一解, 且该问题的最优值为 (1) 中的 $\|\mathbf{y}\|$.

【证明】(1) 先证明存在性. 我们从线性方程近似解的求法得到灵感. (2) 中问题的目标实质上是求线性方程 $\mathbf{A}\boldsymbol{\alpha} = \mathbf{z}$ 在 \mathbb{R}^m 中的近似解, 为此我们将左右两边左乘 \mathbf{A}^\top , 于是得到正规方程

$$\mathbf{A}^\top \mathbf{A} \boldsymbol{\alpha} = \mathbf{A}^\top \mathbf{z}$$

由于 \mathbf{A} 是列满秩的, 所以 $\mathbf{A}^\top \mathbf{A}$ 是可逆的, 所以 $\boldsymbol{\alpha} = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{z}$, 所以取

$$\mathbf{x} = \mathbf{A}\boldsymbol{\alpha} = \mathbf{A}(\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{z}$$

接下来只需证 $\mathbf{z} - \mathbf{A}(\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{z} \in L^\perp$. 由于

$$\begin{aligned} & (\mathbf{z} - \mathbf{A}(\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{z})^\top \mathbf{A}\boldsymbol{\beta} \\ &= \mathbf{z}^\top \mathbf{A}\boldsymbol{\beta} - \mathbf{z}^\top \mathbf{A}(\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{A}\boldsymbol{\beta} \\ &= \mathbf{z}^\top \mathbf{A}\boldsymbol{\beta} - \mathbf{z}^\top \mathbf{A}\mathbf{E}\boldsymbol{\beta} \\ &= 0 \end{aligned}$$

$\forall \beta \in \mathbb{R}^m$ 都成立, 所以 $z - A(A^\top A)^{-1}A^\top z \in L^\perp$, 也即取 $y = z - A(A^\top A)^{-1}A^\top z \in L^\perp$ 即可. 存在性证毕.

接下来证明唯一性. 假设在子空间 L 中, $\exists p \neq x$, 使得 $\exists q \in L^\perp$, $z = p + q$, 则有 $q = x + y - p$, 但是

$$\begin{aligned} x^\top q &= x^\top (x + y - p) \\ &= x^\top x - x^\top p \\ &\neq 0 \end{aligned}$$

与 $\exists q \in L^\perp$ 矛盾, 所以假设不成立, 唯一性得证.

(2) 考查函数 $f(t) = \|z - x - tv\|^2$, $\forall v \in L$ 且 $v \neq 0$, 则要证明目标问题的唯一解是 x , 即证明 $f(t)$ 取最小值当且仅当 $t = 0$, 也即证明 $\forall v \in L$, $f'(0) = 0$, 且 $\forall t \neq 0$, $f'(t) \neq 0$, 而

$$\begin{aligned} f'(t) &= -2y^\top v + 2tv^\top v \\ &= 2tv^\top v \end{aligned}$$

所以上述命题成立, 从而目标问题的唯一解就是 (1) 中的 x , 代入 $u = x$ 即得到目标问题最优值为 $\|y\|$. 【证毕】

例 1.2 设 $x, y \in R^n$, 若 $\forall y \in L$, 都有 $x^\top y \leq \alpha$, 则有 $x \in L^\perp$ 且 $\alpha \geq 0$.

【证明】 取 $y = 0$, 则 $x^\top y = 0$, 从而 $x^\top y$ 的上界 $\alpha \geq 0$. 假设 $x \notin L^\perp$, 则由子空间投影定理的 (1), $\exists x^{(1)} \in L$, $x^{(2)} \in L^\perp$, 使得 $x = x^{(1)} + x^{(2)}$, 那么

$$\begin{aligned} x^\top y &= (x^{(1)} + x^{(2)})^\top y \\ &= x^{(1)\top} y + x^{(2)\top} y \\ &= x^{(1)\top} y \end{aligned}$$

令 $y = \lambda x^{(1)}$ 且 $\lambda \rightarrow +\infty$, 则 $x^{(1)\top} y \rightarrow +\infty$, 与 $x^\top y$ 有上界 α 矛盾, 故必有 $x \in L^\perp$.

【证毕】

该命题还有以下等价表述:

- 若 $\forall y \in L$, 都有 $x^\top y \geq \alpha$, 则有 $x \in L^\perp$ 且 $\alpha \leq 0$.

1.4 多元函数及其导数

多元函数 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ 可看作一个 n 维向量 \mathbf{x} 的函数 $f(\mathbf{x})$. 受篇幅所限, 我们并不给出矩阵导数的定义, 具体请参见《矩阵分析》课程相关内容. 在此我们只不加证明地给出一些常用函数的矩阵导数 (即其梯度).

定理 1.3. (一些常见函数的梯度)

- 线性函数: $f(\mathbf{x}) = \mathbf{c}^\top \mathbf{x} + \mathbf{d}$, $\nabla f(\mathbf{x}) = \mathbf{c}$, $\nabla^2 f(\mathbf{x}) = \mathbf{0}$;
- 二次函数: $f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^\top \mathbf{Q} \mathbf{x} + \mathbf{c}^\top \mathbf{x} + \mathbf{d}$, $\nabla f(\mathbf{x}) = \mathbf{Q} \mathbf{x} + \mathbf{c}$, $\nabla^2 f(\mathbf{x}) = \mathbf{Q}$, 其中 \mathbf{Q} 为对称阵.

接下来依旧不加证明地介绍多元函数的 Taylor 展开式及中值定理. 此处不加证明是因为此内容可以被读者在任何一本《数学分析》教材中找到.

定理 1.4. (Taylor 展开式)

设 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ 在 \mathbf{x}^* 的某邻域内二阶可导, 则有:

- 一阶 Taylor 展开式: $f(\mathbf{x}) = f(\mathbf{x}^*) + \nabla f^\top(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) + o(\|\mathbf{x} - \mathbf{x}^*\|)$;
- 二阶 Taylor 展开式: $f(\mathbf{x}) = f(\mathbf{x}^*) + \nabla f^\top(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) + \frac{1}{2}(\mathbf{x} - \mathbf{x}^*)^\top \nabla^2 f(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) + o(\|\mathbf{x} - \mathbf{x}^*\|^2)$.

定理 1.5. (中值定理)

设 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ 在 \mathbf{x}^* 的某邻域 $U(\mathbf{x}^*)$ 内二阶可导, 则 $\forall \mathbf{x} \in U(\mathbf{x}^*)$, 都有:

- Lagrange 中值定理: $\exists \lambda \in (0, 1)$, 记 $\mathbf{x}_\lambda = \mathbf{x}^* + \lambda(\mathbf{x} - \mathbf{x}^*)$, 则有 $f(\mathbf{x}) = f(\mathbf{x}^*) + \nabla f^\top(\mathbf{x}_\lambda)(\mathbf{x} - \mathbf{x}^*)$;
- Taylor 中值定理: $\exists \mu \in (0, 1)$, 记 $\mathbf{x}_\mu = \mathbf{x}^* + \mu(\mathbf{x} - \mathbf{x}^*)$, 则有 $f(\mathbf{x}) = f(\mathbf{x}^*) + \nabla f^\top(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) + \frac{1}{2}(\mathbf{x} - \mathbf{x}^*)^\top \nabla^2 f(\mathbf{x}_\mu)(\mathbf{x} - \mathbf{x}^*)$.

1.5 凸集、凸函数和凸规划

1.5.1 凸集

1. 凸集、凸组合

定义 1.8. (凸集)

设 $S \subseteq \mathbb{R}^n$, 若 $\forall \mathbf{x}^{(1)}, \mathbf{x}^{(2)} \in S, \lambda \in [0, 1]$, 必有

$$\lambda \mathbf{x}^{(1)} + (1 - \lambda) \mathbf{x}^{(2)} \in S$$

则称 S 为凸集.

规定单点集 $\{\mathbf{x}\}$ 和空集 \emptyset 都是凸集. 需要注意的是, 方程 $\mathbf{x} = \lambda \mathbf{x}^{(1)} + (1 - \lambda) \mathbf{x}^{(2)} = \mathbf{x}^{(2)} + \lambda(\mathbf{x}^{(1)} - \mathbf{x}^{(2)})$, $\lambda \in [0, 1]$ 表示连接 $\mathbf{x}^{(1)}$ 和 $\mathbf{x}^{(2)}$ 的线段.

例 1.3 设 $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$, 证明以下命题成立:

(1) 集合 $S = \{\mathbf{x} \mid \mathbf{Ax} = \mathbf{b}\}$ 是凸集;

(2) 集合 $S = \{\mathbf{x} \mid \mathbf{Ax} \leq \mathbf{b}\}$ 是凸集.

【证明】 (1) 任取 $\mathbf{x}^{(1)}, \mathbf{x}^{(2)} \in S, \lambda \in [0, 1]$, 则都有

$$\begin{aligned} & \mathbf{A}(\lambda \mathbf{x}^{(1)} + (1 - \lambda) \mathbf{x}^{(2)}) \\ &= \lambda \mathbf{Ax}^{(1)} + (1 - \lambda) \mathbf{Ax}^{(2)} \\ &= \lambda \mathbf{b} + (1 - \lambda) \mathbf{b} \\ &= \mathbf{b} \end{aligned}$$

从而 $\lambda \mathbf{x}^{(1)} + (1 - \lambda) \mathbf{x}^{(2)} \in S$, 故 S 是凸集.

(2) 任取 $\mathbf{x}^{(1)}, \mathbf{x}^{(2)} \in S, \lambda \in [0, 1]$, 则都有

$$\begin{aligned} & \mathbf{A}(\lambda \mathbf{x}^{(1)} + (1 - \lambda) \mathbf{x}^{(2)}) \\ &= \lambda \mathbf{Ax}^{(1)} + (1 - \lambda) \mathbf{Ax}^{(2)} \\ &\leq \lambda \mathbf{b} + (1 - \lambda) \mathbf{b} \\ &= \mathbf{b} \end{aligned}$$

从而 $\lambda \mathbf{x}^{(1)} + (1 - \lambda) \mathbf{x}^{(2)} \in S$, 故 S 是凸集.

【证毕】

定理 1.6. (凸集的性质)

(1) 凸集的交是凸集;

(2) 凸集的内点集^①是凸集;

(3) 凸集的闭包^②是凸集;

^① 设 $S \subseteq \mathbb{R}^n$, $\mathbf{x} \in S$, 如果存在一个开球 $B(\mathbf{x}, \varepsilon) = \{\mathbf{y} \in \mathbb{R}^n \mid \|\mathbf{y} - \mathbf{x}\| < \varepsilon\}$, 使得 $B(\mathbf{x}, \varepsilon) \subseteq S$, 则称 \mathbf{x} 是 S 的内点, S 的内点的全体称为 S 的内点集, 记作 $\text{int}(S)$. 所以 $\mathbf{x} \in \text{int}(S) \Leftrightarrow \exists \varepsilon > 0, (\|\mathbf{y} - \mathbf{x}\| < \varepsilon \Rightarrow \mathbf{y} \in S)$.

^② 设 $S \subseteq \mathbb{R}^n$, $\mathbf{x} \in \mathbb{R}^n$, 如果 $\forall \varepsilon > 0$ 都有 $B(\mathbf{x}, \varepsilon) \cap S \neq \emptyset$, 则称 \mathbf{x} 是 S 的聚点, S 的聚点的全体称为 S 的闭包, 记作 \bar{S} .

(4) 凸集边界上任意点存在支撑超平面^③;

(5) 两个互相不交的凸集之间存在分离超平面^④.

【证明】(1) 设 S_1, S_2 都是凸集, 即证明 $S_1 \cap S_2$ 是凸集, 也即证明 $\forall \mathbf{x}^{(1)}, \mathbf{x}^{(2)} \in (S_1 \cap S_2)$, $\forall \lambda \in [0, 1]$, 都有 $\lambda \mathbf{x}^{(1)} + (1 - \lambda) \mathbf{x}^{(2)} \in (S_1 \cap S_2)$. 由于 $\mathbf{x}^{(1)}, \mathbf{x}^{(2)} \in (S_1 \cap S_2)$, 则有 $\mathbf{x}^{(1)}, \mathbf{x}^{(2)} \in S_1$ 且 $\mathbf{x}^{(1)}, \mathbf{x}^{(2)} \in S_2$, 故 $\forall \lambda \in [0, 1]$, 都有 $\lambda \mathbf{x}^{(1)} + (1 - \lambda) \mathbf{x}^{(2)} \in S_1$ 且 $\lambda \mathbf{x}^{(1)} + (1 - \lambda) \mathbf{x}^{(2)} \in S_2$, 所以 $\lambda \mathbf{x}^{(1)} + (1 - \lambda) \mathbf{x}^{(2)} \in (S_1 \cap S_2)$, 原命题得证.

(2) 设 S 是一凸集, $\text{int}(S)$ 是其内点集, 任取 $\mathbf{x}^{(1)}, \mathbf{x}^{(2)} \in \text{int}(S)$, $\lambda \in [0, 1]$, 则 $\exists r > 0$, 使得

$$\|\mathbf{y}^{(1)} - \mathbf{x}^{(1)}\| \leq r \Rightarrow \mathbf{y}^{(1)} \in S,$$

$$\|\mathbf{y}^{(2)} - \mathbf{x}^{(2)}\| \leq r \Rightarrow \mathbf{y}^{(2)} \in S$$

可以证明内点定义中使用开球和使用闭球是等价的, 这里为了讨论方便, 我们使用有闭球的定义. 要证明 $\text{int}(S)$ 是凸集, 即证明 $\mathbf{x}^{(1)} + (1 - \lambda) \mathbf{x}^{(2)} \in \text{int}(S)$, 也即证明 $\exists \varepsilon > 0$, 使得

$$\|\mathbf{z} - (\mathbf{x}^{(1)} + (1 - \lambda) \mathbf{x}^{(2)})\| \leq \varepsilon \Rightarrow \mathbf{z} \in S$$

我们由三角不等式得到启发, 于是上述逻辑式的前件的左式可以作如下放缩:

$$\begin{aligned} & \|\mathbf{z} - (\mathbf{x}^{(1)} + (1 - \lambda) \mathbf{x}^{(2)})\| \\ &= \|\mathbf{z} - \lambda \mathbf{y}^{(1)} - (1 - \lambda) \mathbf{y}^{(2)} + \lambda(\mathbf{y}^{(1)} - \mathbf{x}^{(1)}) + (1 - \lambda)(\mathbf{y}^{(2)} - \mathbf{x}^{(2)})\| \\ &\leq \|\mathbf{z} - \lambda \mathbf{y}^{(1)} - (1 - \lambda) \mathbf{y}^{(2)}\| + \|\lambda(\mathbf{y}^{(1)} - \mathbf{x}^{(1)})\| + \|(1 - \lambda)(\mathbf{y}^{(2)} - \mathbf{x}^{(2)})\| \\ &\leq \|\mathbf{z} - \lambda \mathbf{y}^{(1)} - (1 - \lambda) \mathbf{y}^{(2)}\| + \lambda r + (1 - \lambda) r \\ &= \|\mathbf{z} - \lambda \mathbf{y}^{(1)} - (1 - \lambda) \mathbf{y}^{(2)}\| + r \end{aligned}$$

其中 $\mathbf{y}^{(1)}, \mathbf{y}^{(2)}$ 满足 $\|\mathbf{y}^{(i)} - \mathbf{x}^{(i)}\| \leq r, i = 1, 2$, 所以 $\mathbf{y}^{(1)}, \mathbf{y}^{(2)} \in S$, 故 $\forall \lambda \in [0, 1], \mathbf{y}^{(1)} + (1 - \lambda) \mathbf{y}^{(2)} \in S$. 我们要保证 $\mathbf{z} \in S$, 那么只需要控制 $\varepsilon = r$, 于是就有

$$0 \leq \|\mathbf{z} - \lambda \mathbf{y}^{(1)} - (1 - \lambda) \mathbf{y}^{(2)}\| \leq r - r = 0$$

从而 $\mathbf{z} = \lambda \mathbf{y}^{(1)} + (1 - \lambda) \mathbf{y}^{(2)} \in S$, 也即待证命题: $\exists \varepsilon = r > 0$, 使得

$$\|\mathbf{z} - (\mathbf{x}^{(1)} + (1 - \lambda) \mathbf{x}^{(2)})\| \leq \varepsilon \Rightarrow \mathbf{z} \in S$$

的前件可推出后件, 从而 $\mathbf{x}^{(1)} + (1 - \lambda) \mathbf{x}^{(2)} \in \text{int}(S)$, 待证命题得证.

(3) 我们先证明一个引理: $\mathbf{x} \in \bar{S}$ 当且仅当存在一个点列 $\{\mathbf{x}^{(n)}\} \subset S$ 满足 $\lim_{n \rightarrow \infty} \mathbf{x}^{(n)} = \mathbf{x}$. (实际上这正是聚点的另一等价定义)

^③ 设 $\boldsymbol{\alpha} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, $\beta \in \mathbb{R}$, 称集合 $H = \{\mathbf{x} \in \mathbb{R}^n \mid \boldsymbol{\alpha}^\top \mathbf{x} = \beta\}$ 是超平面, 简记作 $H: \boldsymbol{\alpha}^\top \mathbf{x} = \beta$. 设 $S \subseteq \mathbb{R}^n$, $\mathbf{x}_0 \in S$, 若超平面 $H: \boldsymbol{\alpha}^\top \mathbf{x} = \beta$ 满足:

(i) $\exists \mathbf{x}^{(0)} \in S$, 使得 $\mathbf{x}^{(0)} \in H$;

(ii) $\forall \mathbf{x} \in S, \boldsymbol{\alpha}^\top \mathbf{x} \leq \beta$

则称 H 是 S 的支撑超平面.

^④ 设 $S_1, S_2 \subseteq \mathbb{R}^n$, 且 $S_1 \cap S_2 = \emptyset$. 若存在超平面 $H: \boldsymbol{\alpha}^\top \mathbf{x} = \beta$, 使得 $\forall \mathbf{x}^{(1)} \in S_1, \boldsymbol{\alpha}^\top \mathbf{x}^{(1)} \geq \beta$, 且 $\forall \mathbf{x}^{(2)} \in S_1, \boldsymbol{\alpha}^\top \mathbf{x}^{(2)} \leq \beta$, 则称 H 是 S_1, S_2 的分离超平面.

先证 \Rightarrow . 对每一个整数 n , 取 $\varepsilon_n = \frac{1}{n}$, 则能得到一个开球列 $\{B(\mathbf{x}, \varepsilon_n)\}$. 应用选择公理, 从这个开球列中依次取一个点, 排成一个点列 $\{\mathbf{x}^{(n)}\}$, 则必然有

$$\|\mathbf{x}^{(n)} - \mathbf{x}\| = \frac{1}{n} \rightarrow 0$$

再证 \Leftarrow . 由于 $\lim_{n \rightarrow \infty} \mathbf{x}^{(n)} = \mathbf{x}$, 则 $\forall \varepsilon > 0$, 都 $\exists N > 0$, 使得 $\|\mathbf{x}^{(n)} - \mathbf{x}\| < \varepsilon$ 对于 $\forall n > N$ 成立, 从而 $\forall \varepsilon > 0$, 都一定存在一个 n , 使得 $\mathbf{x}^{(n)} \in B(\mathbf{x}, \varepsilon)$ 并且 $\mathbf{x}^{(n)} \in S$, 从而 $\mathbf{x}^{(n)} \in B(\mathbf{x}, \varepsilon) \cap S$, 故 $B(\mathbf{x}, \varepsilon) \cap S \neq \emptyset$, 所以 $\mathbf{x} \in \bar{S}$. 从而引理得证.

任取 $\mathbf{x}, \mathbf{y} \in \bar{S}$, $\lambda \in [0, 1]$, 由引理, $\exists \{\mathbf{x}^{(n)}\}, \{\mathbf{y}^{(n)}\} \subset S$, 使得 $\mathbf{x}^{(n)} \rightarrow \mathbf{x}, \mathbf{y}^{(n)} \rightarrow \mathbf{y}$. 构造点列 $\{\lambda \mathbf{x}^{(n)} + (1-\lambda)\mathbf{y}^{(n)}\}$, 则由 $\exists \{\mathbf{x}^{(n)}\}, \{\mathbf{y}^{(n)}\} \subset S$ 以及 S 是凸集, 有 $\{\lambda \mathbf{x}^{(n)} + (1-\lambda)\mathbf{y}^{(n)}\} \in S$. 由 $\mathbf{x}^{(n)} \rightarrow \mathbf{x}, \mathbf{y}^{(n)} \rightarrow \mathbf{y}$, 知 $\{\lambda \mathbf{x}^{(n)} + (1-\lambda)\mathbf{y}^{(n)}\} \rightarrow \lambda \mathbf{x} + (1-\lambda)\mathbf{y}$, 所以根据引理, $\lambda \mathbf{x} + (1-\lambda)\mathbf{y} \in \bar{S}$. 原命题得证.

(4) 暂时懒得写.

(5) 暂时懒得写.

【证毕】

定义 1.9. (凸组合)

设 $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(m)} \in \mathbb{R}^n$. 若 $\forall j \in \{1, 2, \dots, m\}, \lambda_j \geq 0$ 且 $\sum_{j=1}^m \lambda_j = 1$, 则称 $\sum_{j=1}^m \lambda_j \mathbf{x}^{(j)}$ 为 $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(m)}$ 的凸组合.

定理 1.7. (凸集和凸组合的关系)

S 是凸集 $\Leftrightarrow S$ 中任意有限点的凸组合属于 S .

【证明】先证 \Rightarrow . 使用数学归纳法. 即证明 S 中任意 m 个点的凸组合属于 S . $m = 1, 2$ 时, 由于 S 是凸集, 所以显然成立. 假设 $m = k$ 时, 结论成立, 往证 $m = k+1$ 时结论也成立, 即证明当 $\sum_{j=1}^{k+1} \lambda_j = 1$ 时 $\sum_{j=1}^{k+1} \lambda_j \mathbf{x}_j \in S$. 由于 $m = k$ 时结论成立, 则有:

$$\frac{\sum_{j=1}^k \lambda_j \mathbf{x}_j}{\sum_{j=1}^k \lambda_j} = \sum_{j=1}^k \frac{\lambda_j}{\sum_{i=1}^k \lambda_i} \mathbf{x}_j \in S$$

从而

$$\sum_{j=1}^{k+1} \lambda_j \mathbf{x}^{(j)} = \sum_{j=1}^k \lambda_j \mathbf{x}^{(j)} + \lambda_{k+1} \mathbf{x}^{(k+1)} = \sum_{j=1}^k \lambda_j \cdot \frac{\sum_{j=1}^k \lambda_j \mathbf{x}^{(j)}}{\sum_{j=1}^k \lambda_j} + \lambda_{k+1} \mathbf{x}^{(k+1)} \in S$$

结论得证.

再证 \Leftarrow . 取 $m = 2$, 则若 $\lambda_1, \lambda_2 > 0$ 且 $\lambda_1 + \lambda_2 = 1$, 都有 $\lambda_1 \mathbf{x}^{(1)} + \lambda_2 \mathbf{x}^{(2)} \in S$, 这正是凸集的定义, 从而结论成立.

【证毕】

定义 1.10. (凸包)

设非空集合 $S \subseteq \mathbb{R}^n$, 由 S 中所有有限点的凸组合所构成的集合, 被称为 S 的凸包, 记作 $\text{cov}(S)$.

定理 1.8. (凸集和凸包的关系)

如果 S 是凸集, 那么 $\text{cov}(S) = S$.

【证明】 由定理 1.7 知, $\text{cov}(S) \subseteq S$, 下证 $S \subseteq \text{cov}(S)$. 一个点的凸组合就是其本身, 所以只要 $\mathbf{x} \in S$, 那么这个点本身的凸组合 $\mathbf{x} \in \text{cov}(S)$, 所以 $S \subseteq \text{cov}(S)$. **【证毕】**

定义 1.11. (多胞形)

设 $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(m)} \in \mathbb{R}^n$. 由 $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(m)}$ 的所有凸组合构成的集合称为多胞形, 记作 $H(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(m)})$.

定义 1.12. (单纯形)

设 $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(m)} \in \mathbb{R}^n$. 若多胞形 $H(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(m)})$ 满足 $\mathbf{x}^{(2)} - \mathbf{x}^{(1)}, \mathbf{x}^{(3)} - \mathbf{x}^{(1)}, \dots, \mathbf{x}^{(m)} - \mathbf{x}^{(1)}$ 线性无关, 则称该多胞形是单纯形.

2. 凸锥、半正组合**定义 1.13. (锥、凸锥)**

设 $S \subseteq \mathbb{R}^n$, 若 $\forall \mathbf{x} \in S, \lambda > 0$, 必有 $\lambda \mathbf{x} \in S$, 则称 S 为以 $\mathbf{0}$ 为顶点的锥. 若 S 还是凸集, 则称 S 为凸锥.

规定 $\{\mathbf{0}\}, \mathbb{R}^n$ 都是凸锥.

定义 1.14. (半正组合)

设 $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(m)} \in \mathbb{R}^n$. 若 $\sum_{j=1}^m \lambda_j > 0$, 则称 $\sum_{j=1}^m \lambda_j \mathbf{x}^{(j)}$ 为 $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(m)}$ 的半正组合.

定理 1.9. (凸锥和半正组合的关系)

S 是凸锥 $\Leftrightarrow S$ 中任意有限点的半正组合属于 S .

【证明】

【证毕】

1.5.2 凸函数**1. 凸函数****2. 水平集****1.5.3 凸规划**