Optimal Theory and Method

最优化理论与方法

程春杰

杭州电子科技大学 自动化学院

科技馆512

Email: cjzhai@hdu.edu.cn

目录

- 线性与非线性规划
- 几个数学概念
- 凸集（考）
- 凸函数（考）

目录

- **线性与非线性规划**
- **几个数学概念**
- **凸集（考）**
- **凸函数（考）**

■ 线性与非线性规划

- 什么是**线性**规划问题和**非线性**规划问题

- 什么是**可行点**、**可行域**（可行集）

- 什么是**全局极小点**和**局部极小点**

■ 线性与非线性规划

➢线性规划问题

例1 生产计划问题

设某工厂用 4 种资源生产 3 种产品，每单位第 j 种产品需要第 i 种资源的数量为 a_{ij}，可获利润为 c_j，第 i 种资源总消耗量不能超过 b_i，由于市场限制，第 j 种产品的产量不超过 d_j，试问如何安排生产才能使总利润最大？

数学模型如下：

$$
\begin{aligned}
\max \quad & \sum_{j=1}^{3} c_j x_j \\
\text{s.t.} \quad & \sum_{j=1}^{3} a_{ij} x_j \leqslant b_i, \quad & i = 1, \cdots, 4, \\
& x_j \leqslant d_j, \quad & j = 1, 2, 3, \\
& x_j \geqslant 0, \quad & j = 1, 2, 3,
\end{aligned}
$$

目标函数和约束中自变量的幂次全为1——线性规划问题

线性与非线性规划

非线性规划问题

例2 选址问题

设有 n 个市场，第 j 个市场的位置为 (a_j, b_j)，对某种货物的需要量为 $q_j\ (j=1,\cdots,n)$。现计划建立 m 个货栈，第 i 个货栈的容量为 $c_i\ (i=1,\cdots,m)$。试确定货栈的位置，使各货栈到各市场的运输量与路程乘积之和最小。

数学模型如下：

$$\min \quad \sum_{i=1}^{m}\sum_{j=1}^{n} W_{ij}\sqrt{(x_i-a_j)^2+(y_i-b_j)^2}$$

$$\text{s.t.} \quad \sum_{j=1}^{n} W_{ij} \leqslant c_i, \quad i=1,\cdots,m,$$

$$\sum_{i=1}^{m} W_{ij} = q_j, \quad j=1,\cdots,n,$$

$$W_{ij} \geqslant 0, \qquad i=1,\cdots,m, j=1,\cdots,n.$$

目标函数和约束中自变量的幂次不全为1——非线性规划问题

线性与非线性规划

可行点和可行域

例1 生产计划问题

$$\max \quad \sum_{j=1}^{3} c_j x_j$$

$$\begin{aligned} \text{s.t.} \quad & \sum_{j=1}^{3} a_{ij} x_j \leqslant b_i, & i = 1, \cdots, 4, \\ & x_j \leqslant d_j, & j = 1, 2, 3, \\ & x_j \geqslant 0, & j = 1, 2, 3, \end{aligned}$$

例2 选址问题

$$\min \quad \sum_{i=1}^{m} \sum_{j=1}^{n} W_{ij} \sqrt{(x_i - a_j)^2 + (y_i - b_j)^2}$$

$$\begin{aligned} \text{s.t.} \quad & \sum_{j=1}^{n} W_{ij} \leqslant c_i, & i = 1, \cdots, m, \\ & \sum_{i=1}^{m} W_{ij} = q_j, & j = 1, \cdots, n, \\ & W_{ij} \geqslant 0, & i = 1, \cdots, m, j = 1, \cdots, n. \end{aligned}$$

在线性和非线性规划中，满足约束的点——可行点

所有可行点组成的集合——可行域（可行集）

可行域是整个空间——无约束优化问题

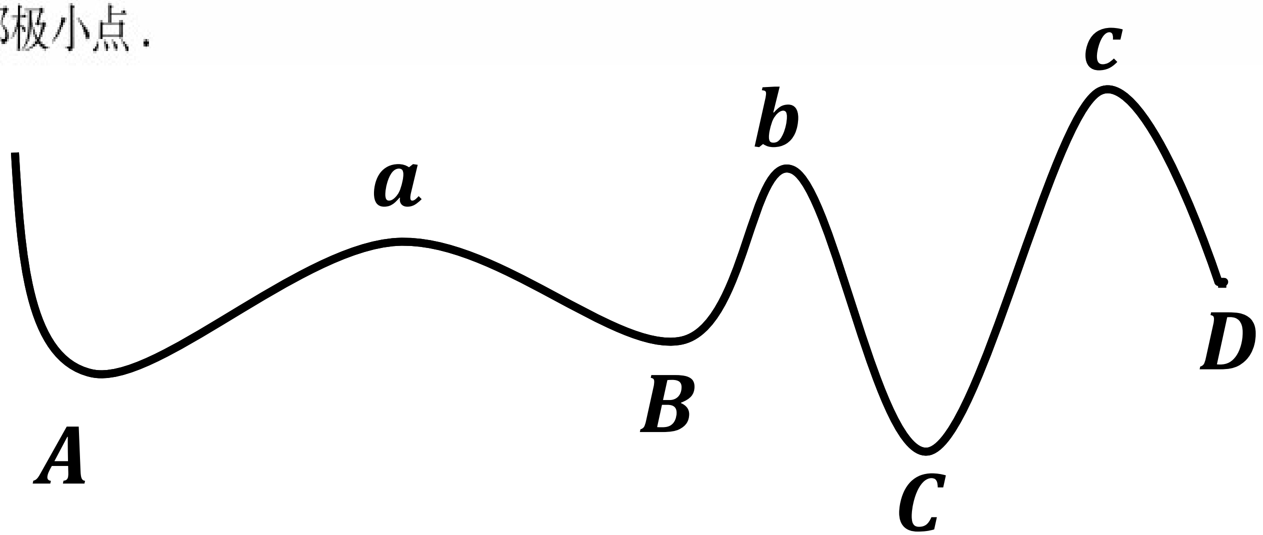
线性与非线性规划

全局极小点

定义 1.2.1 设 $f(x)$为目标函数，S 为可行域，$\bar{x}\in S$，若对每个 $x\in S$，成立 $f(x)\geq f(\bar{x})$，则称 $\bar{x}$ 为 $f(x)$在 S 上的全局极小点。

局部极小点

定义 1.2.2 设 $f(x)$为目标函数，S 为可行域，若存在 $\bar{x}\in S$ 的 $\varepsilon>0$ 邻域 $N(\bar{x},\varepsilon)=\{x \mid \|x-\bar{x}\|<\varepsilon\}$，使得对每个 $x\in S\cap N(\bar{x},\varepsilon)$成立 $f(x)\geq f(\bar{x})$，则称 $\bar{x}$ 为 $f(x)$在 S 上的一个局部极小点。



目录

- 线性与非线性规划
- 几个数学概念
- 凸集（考）
- 凸函数（考）

■ 几个数学概念

- 什么是向量范数
- 什么是序列的极限、聚点和柯西序列
- 什么是开集、闭集和紧集
- 什么是梯度、海塞矩阵和雅克比矩阵
- 什么是泰勒展开式

■ 几个数学概念

➢向量范数

定义 1.3.1 若实值函数 $\|\cdot\|: \mathbf{R}^n \to \mathbf{R}$ 满足下列条件：

(1) $\|\boldsymbol{x}\| \geqslant 0, \forall \boldsymbol{x} \in \mathbf{R}^n$；$\|\boldsymbol{x}\| = 0$ 当且仅当 $\boldsymbol{x} = \mathbf{0}$；

(2) $\|\alpha \boldsymbol{x}\| = |\alpha| \|\boldsymbol{x}\|, \forall \alpha \in \mathbf{R}, \boldsymbol{x} \in \mathbf{R}^n$；

(3) $\|\boldsymbol{x} + \boldsymbol{y}\| \leqslant \|\boldsymbol{x}\| + \|\boldsymbol{y}\|, \forall \boldsymbol{x}, \boldsymbol{y} \in \mathbf{R}^n$.

则称 $\|\cdot\|$ 为向量范数.其中 $\mathbf{R}^n$ 表示 n 维向量空间.

常用的向量范数

设 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}} \in \mathbf{R}^n$，常用的向量范数有 L_1 范数，L_2 范数和 L_∞ 范数，分别为

$$\|\boldsymbol{x}\|_1 = \sum_{j=1}^{n} |x_j|,$$

$$\|\boldsymbol{x}\|_2 = \left(\sum_{j=1}^{n} x_j^2\right)^{\frac{1}{2}},$$

$$\|\boldsymbol{x}\|_\infty = \max_j |x_j|.$$

一般地，对于 $1 \leqslant p < \infty$，L_p 范数为

$$\|\boldsymbol{x}\|_p = \left(\sum_{j=1}^{n} |x_j|^p\right)^{\frac{1}{p}}.$$

■ 几个数学概念

➢序列的极限、聚点和柯西序列

序列的极限

定义 1.3.4 设$\{x^{(k)}\}$是$\mathbf{R}^n$中一个向量序列，$\bar{x}\in\mathbf{R}^n$，如果对每个任给的$\varepsilon>0$存在正整数K_ε，使得当$k>K_\varepsilon$时就有$\|x^{(k)}-\bar{x}\|<\varepsilon$，则称序列收敛到$\bar{x}$，或称序列以$\bar{x}$为极限，记作$\lim\limits_{k\to\infty}x^{(k)}=\bar{x}$.

$$\left(\frac{1}{2}\right)^0, \left(\frac{1}{2}\right)^1, \left(\frac{1}{2}\right)^2, \left(\frac{1}{2}\right)^3 \cdots\cdots\cdots \left(\frac{1}{2}\right)^l \cdots\cdots\cdots \left(\frac{1}{2}\right)^m \cdots\cdots\cdots$$

$$(-1)^0, (-1)^1, (-1)^2, (-1)^3 \cdots\cdots (-1)^l \cdots\cdots (-1)^m \cdots\cdots$$

序列的聚点

定义 1.3.5 设$\{x^{(k)}\}$是$\mathbf{R}^n$中一个向量序列，如果存在一个子序列$\{x^{(k_j)}\}$，使$\lim\limits_{k_j\to\infty}x^{(k_j)}=\hat{x}$，则称$\hat{x}$是序列$\{x^{(k)}\}$的一个聚点.

■ 几个数学概念

➢序列的极限、聚点和柯西序列

柯西序列

定义 1.3.6 设$\{x^{(k)}\}$是$\mathbf{R}^n$中一个向量序列，如果对任意给定的 $\varepsilon>0$，总存在正整数 K_ε，使得当 $m, l>K_\varepsilon$ 时，就有 $\|x^{(m)}-x^{(l)}\|<\varepsilon$，则$\{x^{(k)}\}$称为 Cauchy 序列。

$$\left(\frac{1}{2}\right)^0, \left(\frac{1}{2}\right)^1, \left(\frac{1}{2}\right)^2, \left(\frac{1}{2}\right)^3 \cdots\cdots\cdots \left(\frac{1}{2}\right)^l \cdots\cdots\cdots \left(\frac{1}{2}\right)^m \cdots\cdots\cdots$$

$$\left(-\frac{1}{2}\right)^0, \left(-\frac{1}{2}\right)^1, \left(-\frac{1}{2}\right)^2, \left(-\frac{1}{2}\right)^3 \cdots \left(-\frac{1}{2}\right)^l \cdots \left(-\frac{1}{2}\right)^m \cdots$$

柯西序列肯定有极限

■ 几个数学概念

➢开集、闭集、紧集

开集

对每一点 $\hat{x} \in S$

存在正数 ε,使得 $\hat{x}$ 的 ε 邻域 $N(\hat{x}, \varepsilon)=\{x \mid \|x-\hat{x}\|<\varepsilon\} \subset S$,则称 S 为开集

$$0 \le x < 2 \qquad 0 \le x \le 2 \qquad 0 < x < 2$$

闭集

如果 S 中每个收敛序列的极限均属于 S,则称 S 为闭集

$$0 \le x < \infty \qquad 0 \le x \le 2 \qquad 0 \le x < 2$$

紧集

S为有界闭集

■ 几个数学概念

➢梯度、海塞矩阵、雅克比矩阵

梯度

函数 f 在 $\boldsymbol{x}$ 处的梯度为 n 维列向量：

$$\Delta f(\boldsymbol{x}) = \left[\frac{\partial f(\boldsymbol{x})}{\partial x_1}, \frac{\partial f(\boldsymbol{x})}{\partial x_2}, \cdots, \frac{\partial f(\boldsymbol{x})}{\partial x_n} \right]^{\mathrm{T}}.$$

海塞矩阵

f 在 $\boldsymbol{x}$ 处的 Hesse 矩阵为 $n \times n$ 矩阵$\Delta^2 f(\boldsymbol{x})$,第 i 行第 j 列元素为

$$[\Delta^2 f(\boldsymbol{x})]_{ij} = \frac{\partial^2 f(\boldsymbol{x})}{\partial x_i \partial x_j}, \qquad 1 \leqslant i,j \leqslant n.$$

■ 几个数学概念

➢梯度、海塞矩阵、雅克比矩阵

雅克比矩阵

考虑向量值函数

$$h(\boldsymbol{x}) = \left(h_1(\boldsymbol{x}), h_2(\boldsymbol{x}), \cdots, h_m(\boldsymbol{x}) \right)^{\mathrm{T}}$$

H(x)在x处的雅克比矩阵

$$\begin{bmatrix} \frac{\partial h_1(\boldsymbol{x})}{\partial x_1} & \frac{\partial h_1(\boldsymbol{x})}{\partial x_2} & \cdots & \frac{\partial h_1(\boldsymbol{x})}{\partial x_n} \\ \frac{\partial h_2(\boldsymbol{x})}{\partial x_1} & \frac{\partial h_2(\boldsymbol{x})}{\partial x_2} & \cdots & \frac{\partial h_2(\boldsymbol{x})}{\partial x_n} \\ \cdots & \cdots & & \cdots \\ \frac{\partial h_m(\boldsymbol{x})}{\partial x_1} & \frac{\partial h_m(\boldsymbol{x})}{\partial x_2} & \cdots & \frac{\partial h_m(\boldsymbol{x})}{\partial x_n} \end{bmatrix}$$

■ 几个数学概念

➢梯度、海塞矩阵、雅克比矩阵之间的关系

$$f(\boldsymbol{x})=\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b}^{\mathrm{T}}\boldsymbol{x}+c$$

二次函数f(x)的梯度

$$\nabla f(x)=\mathbf{A}x+b$$

二次函数f(x)的海塞矩阵

$$\Delta^{2} f(\boldsymbol{x})=\boldsymbol{A}$$

二次函数的梯度和海塞矩阵一定要记住——考点

■ 几个数学概念

➢泰勒展开式

假设在开集 $S \subset \mathbf{R}^n$ 上 $f \in C^1(S)$，给定点 $\bar{\boldsymbol{x}} \in S$，则 f 在点 $\bar{\boldsymbol{x}}$ 的一阶 Taylor 展开式为

$$f(\boldsymbol{x}) = f(\bar{\boldsymbol{x}}) + \Delta f(\bar{\boldsymbol{x}})^{\mathrm{T}}(\boldsymbol{x} - \bar{\boldsymbol{x}}) + o(\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|),$$

其中 $o(\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|)$ 当 $\|\boldsymbol{x} - \bar{\boldsymbol{x}}\| \to 0$ 时，关于 $\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|$ 是高阶无穷小量.

假设在开集 $S \subset \mathbf{R}^n$ 上 $f \in C^2(S)$，则 f 在 $\bar{\boldsymbol{x}} \in S$ 的二阶 Taylor 展开式为

$$f(\boldsymbol{x}) = f(\bar{\boldsymbol{x}}) + \Delta f(\bar{\boldsymbol{x}})^{\mathrm{T}}(\boldsymbol{x} - \bar{\boldsymbol{x}}) + \frac{1}{2}(\boldsymbol{x} - \bar{\boldsymbol{x}})^{\mathrm{T}} \Delta^2 f(\bar{\boldsymbol{x}})(\boldsymbol{x} - \bar{\boldsymbol{x}}) + o(\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|^2),$$

其中 $o(\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|^2)$ 当 $\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|^2 \to 0$ 时，关于 $\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|^2$ 是高阶无穷小量

一阶、二阶泰勒展开式一定要会写——考点

目录

- 线性与非线性规划
- 几个数学概念
- 凸集（考）
- 凸函数（考）

■ 凸集

➢什么是凸集

➢什么是凸锥

➢什么是多面集

➢什么是极点

➢什么是方向和极方向

凸集

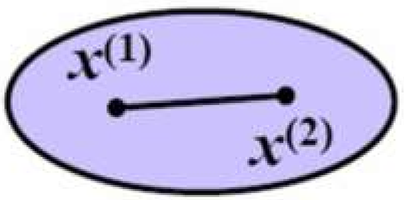
凸集

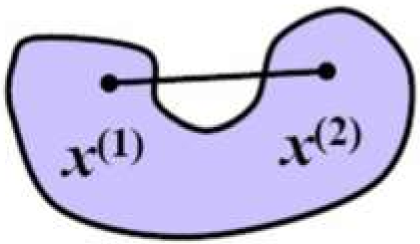
定义 1.4.1 设 S 为 n 维欧氏空间 $\mathbf{R}^n$ 中一个集合.若对 S 中任意两点,联结它们的线段仍属于 S;换言之,对 S 中任意两点 $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}$ 及每个实数 $\lambda \in [0,1]$,都有

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

则称 S 为凸集.

$$\boldsymbol{\lambda x^{(1)} + (1-\lambda)x^{(2)} = x^{(2)} + \lambda(x^{(1)} - x^{(2)})}$$





■ 凸集

➢凸集

例1

验证集合 $H^{-}=\{x \mid p^{T} x \leqslant \alpha\}$ 为凸集

解 由于对任意两点 $x^{(1)}, x^{(2)} \in H$ 及每个实数 $\lambda \in [0,1]$ 都有

$$p^{T}[\lambda x^{(1)}+(1-\lambda) x^{(2)}]=\alpha,$$

因此

$$\lambda x^{(1)}+(1-\lambda) x^{(2)} \in H .$$

集合 H 称为 $\mathbf{R}^{n}$ 中的超平面，故超平面为凸集

凸集

凸集

课堂练习题1

$$S=\{(x_1,x_2)\mid x_1^2+x_2^2\leq 10\}$$

解：对任意两点$\boldsymbol{x}^{(1)}=(\boldsymbol{x}_1^{(1)},\boldsymbol{x}_2^{(1)})$，$\boldsymbol{x}^{(2)}=\left(\boldsymbol{x}_1^{(2)},\boldsymbol{x}_2^{(2)}\right)\in \boldsymbol{S}$**及每个实数**

$\boldsymbol{\lambda}\in[\mathbf{0},\mathbf{1}]$**都有**

$$\lambda\boldsymbol{x}^{(1)}+(1-\lambda)\boldsymbol{x}^{(2)}=(\lambda\boldsymbol{x}_1^{(1)}+(1-\lambda)\boldsymbol{x}_1^{(2)},\lambda\boldsymbol{x}_2^{(1)}+(1-\lambda)\boldsymbol{x}_2^{(2)})$$

$$\left[\lambda\boldsymbol{x}_1^{(1)}+(1-\lambda)\boldsymbol{x}_1^{(2)}\right]^{\mathbf{2}}+\left[\lambda\boldsymbol{x}_2^{(1)}+(1-\lambda)\boldsymbol{x}_2^{(2)}\right]^{\mathbf{2}}$$

$$=\lambda^2\left({\boldsymbol{x}_1^{(1)}}^2+{\boldsymbol{x}_2^{(1)}}^2\right)+(1-\lambda)^2\left({\boldsymbol{x}_1^{(2)}}^2+{\boldsymbol{x}_2^{(2)}}^2\right)+2\lambda(1-\lambda)\boldsymbol{x}_1^{(1)}\boldsymbol{x}_1^{(2)}+$$

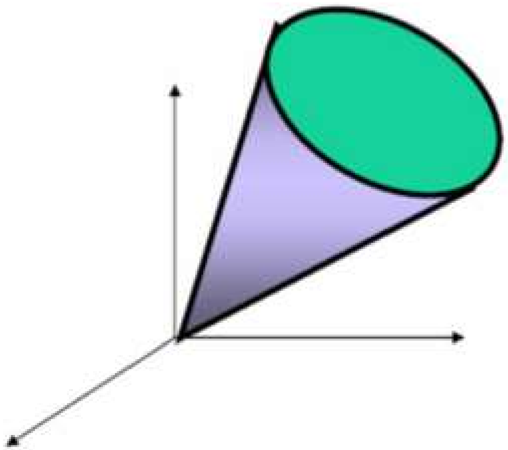
$$2\lambda(1-\lambda)\boldsymbol{x}_2^{(1)}\boldsymbol{x}_2^{(2)}\leq\mathbf{10}(\lambda^2+(1-\lambda)^2)+\mathbf{20}\lambda(1-\lambda)\leq\mathbf{10}$$

因此，$\boldsymbol{\lambda x}^{(1)}+(\mathbf{1}-\boldsymbol{\lambda})\boldsymbol{x}^{(2)}\in\boldsymbol{S}$ 集合$\boldsymbol{S}$为凸集

■ 凸集

➢凸锥

定义 1.4.2 设有集合 $C \subset \mathbf{R}^n$，若对 C 中每一点 x，当 λ 取任何非负数时，都有 $\lambda x \in C$，则称 C 为锥，又若 C 为凸集，则称 C 为**凸锥**。



■ 凸集

➢凸锥

例 1.4.4 向量集 $\alpha^{(1)},\alpha^{(2)},\cdots,\alpha^{(k)}$ 的所有非负线性组合构成的集合

$$\left\{\sum_{i=1}^{k}\lambda_i\alpha^{(i)}\,\middle|\,\lambda_i\geqslant 0,i=1,\cdots,k\right\}$$

为凸锥 .

证明：取任意一点$\boldsymbol{x}=\sum\lambda_i\alpha^{(i)}\in S,\lambda_i\geq 0$

令$\lambda\boldsymbol{x}=\sum\lambda\lambda_i\alpha^{(i)}=\sum(\lambda\lambda_i)\alpha^{(i)}$

$\lambda\lambda_i\geq 0$ **故**$\lambda\boldsymbol{x}\in S$

■ 凸集

➢多面集

定义 1.4.3 有限个半空间的交

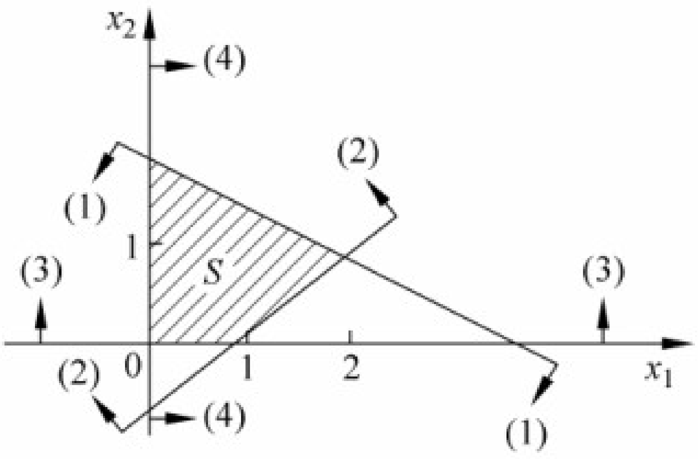
$$\{\boldsymbol{x} \mid \boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}\}$$

称为多面集，其中 $\boldsymbol{A}$ 为 $m \times n$ 矩阵，$\boldsymbol{b}$ 为 m 维向量。

例 1.4.5 集合

$$S = \{\boldsymbol{x} \mid x_1 + 2x_2 \leqslant 4, x_1 - x_2 \leqslant 1, x_1 \geqslant 0, x_2 \geqslant 0\}$$

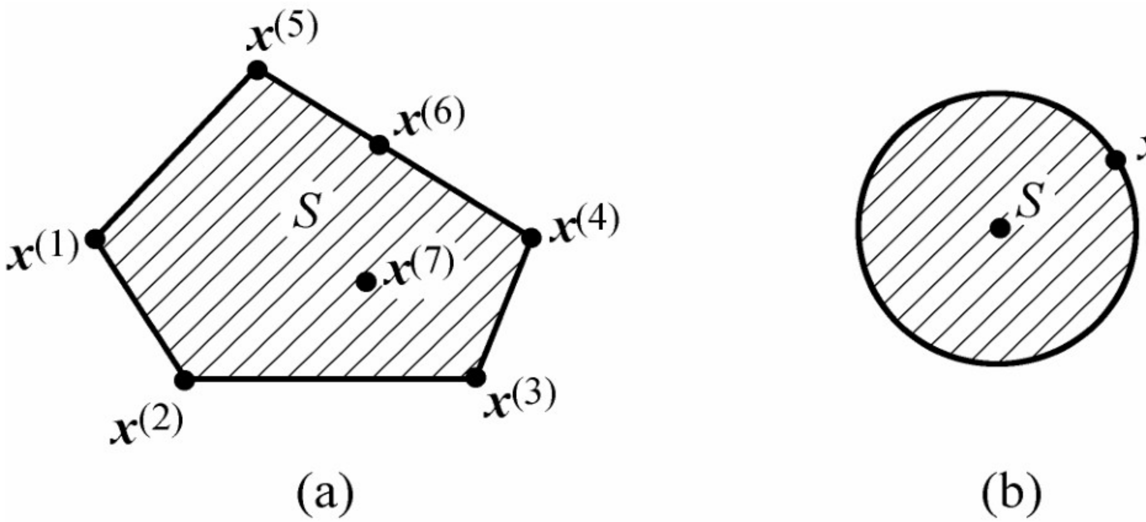
为多面集．其几何表示如图



■ 凸集

➢极点

定义 1.4.4 设 S 为非空凸集，$\boldsymbol{x}\in S$，若 $\boldsymbol{x}$ 不能表示成 S 中两个不同点的凸组合；换言之，若假设 $\boldsymbol{x}=\lambda\boldsymbol{x}^{(1)}+(1-\lambda)\boldsymbol{x}^{(2)}$ $(\lambda\in(0,1))$，$\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}\in S$，必推得 $\boldsymbol{x}=\boldsymbol{x}^{(1)}=\boldsymbol{x}^{(2)}$，则称 $\boldsymbol{x}$ 是凸集 S 的**极点**.



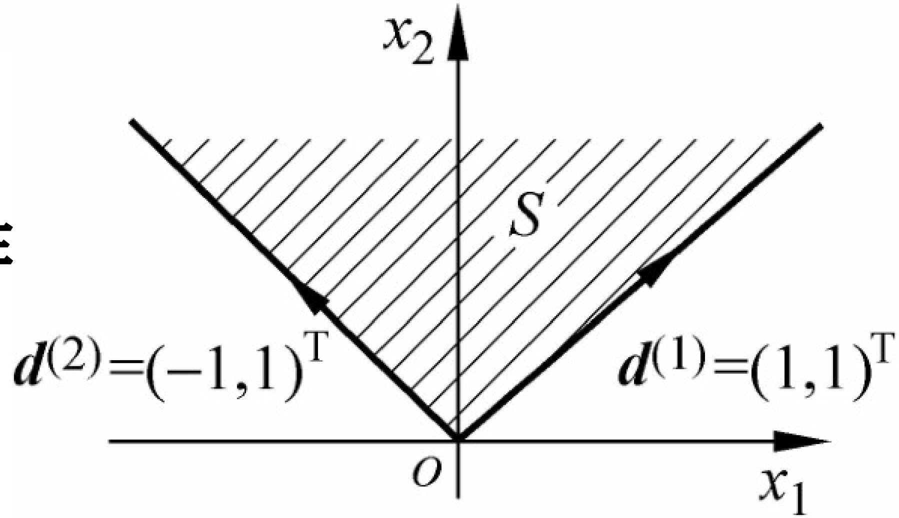
■ 凸集

➢方向和极方向

定义 1.4.5 设 S 为 $\mathbb{R}^n$ 中的闭凸集，$\boldsymbol{d}$ 为非零向量，如果对 S 中的每一个 $\boldsymbol{x}$，都有射线

$$\{\boldsymbol{x}+\lambda\boldsymbol{d} \mid \lambda\geqslant 0\}\subset S,$$

则称向量 $\boldsymbol{d}$ 为 S 的**方向**．又设 $\boldsymbol{d}^{(1)}$ 和 $\boldsymbol{d}^{(2)}$ 是 S 的两个方向，若对任何正数 λ，有 $\boldsymbol{d}^{(1)}\neq\lambda\boldsymbol{d}^{(2)}$，则称 $\boldsymbol{d}^{(1)}$ 和 $\boldsymbol{d}^{(2)}$ 是两个不同的方向．若 S 的方向 $\boldsymbol{d}$ 不能表示成该集合的两个不同方向的正的线性组合，则称 $\boldsymbol{d}$ 为 S 的**极方向**．



有界集合不存在方向，更不存在极方向

目录

- 线性与非线性规划
- 几个数学概念
- 凸集（考）
- 凸函数（考）

■ 凸函数

- 什么是凸函数

- 凸函数有哪些性质

- 凸函数的判别

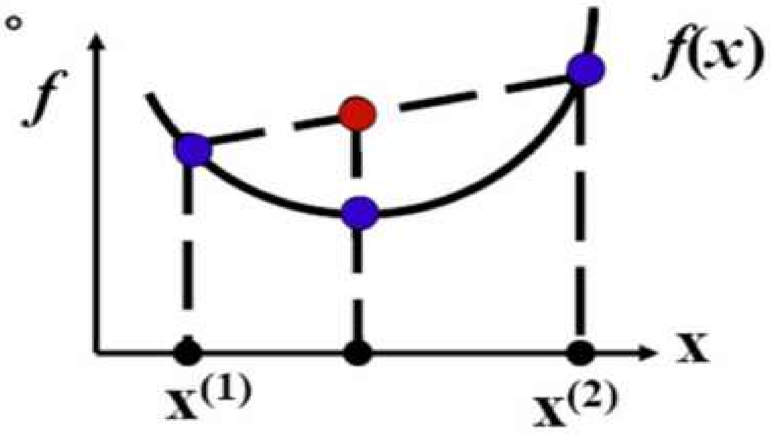
■ 凸函数

➢凸函数

◆定义2．2．1 设S为E^n中的非空凸集，f是定义在S上的实函数。如果对任意的$x^{(1)},x^{(2)}\in S$及每个数$\lambda\in(0，1)$，都有

$$f(\lambda x^{(1)}+(1-\lambda)x^{(2)})\leq \lambda f(x^{(1)})+(1-\lambda)f(x^{(2)})$$

则称f为S上的凸函数。



■ 凸函数

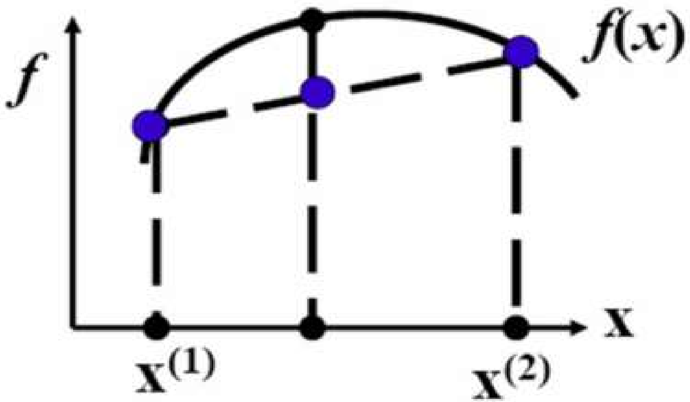
➢凸函数

如果对任意互不相同的$x^{(1)}, x^{(2)} \in S$，及每一个数$\lambda \in (0,\ 1)$，都有

$$f(\lambda x^{(1)}+(1-\lambda)x^{(2)}) < \lambda f(x^{(1)})+(1-\lambda)f(x^{(2)})$$

则称f为S上的严格凸函数。

如果$-f$为S上的凸函数，则称f为S上的凹函数



■ 凸函数

➢凸函数性质

◆定理2.2.1 设f是定义在凸集S上的凸函数，实数 $\lambda \geqslant 0$，则 λf 也是定义在S上的凸函数

◆定理2.2.2 设f_1和f_2是定义在凸集S上的凸函数，则f_1+f_2也是定义在S上的凸函数

■ 凸函数

➢凸函数

例1 一元函数 $f(x)=|x|$ 是 $\mathbf{R}^1$ 上的凸函数.

解 对任意的 $x^{(1)}, x^{(2)} \in \mathbf{R}^1$ 及每个数 $\lambda \in (0,1)$,均有

$$
\begin{aligned}
f(\lambda x^{(1)}+(1-\lambda)x^{(2)}) &= |\lambda x^{(1)}+(1-\lambda)x^{(2)}| \\
&\leqslant \lambda |x^{(1)}| + (1-\lambda)|x^{(2)}| \\
&= \lambda f(x^{(1)}) + (1-\lambda) f(x^{(2)}),
\end{aligned}
$$

因此,由定义 1.4.7 知,$f(x)=|x|$ 为凸函数.

■ 凸函数

➢凸函数

练习题2

$$f(x_1, x_2) = x_1^2 - 2x_1x_2 + x_2^2 + x_1 + x_2;$$

解：对任意两点$\boldsymbol{x}^{(1)} = (x_1^{(1)}, x_2^{(1)})$，$\boldsymbol{x}^{(2)} = \left(x_1^{(2)}, x_2^{(2)}\right) \in S$**及每个实数**

$\lambda \in (0,1)$**都有**

$$\lambda \boldsymbol{x}^{(1)} + (1-\lambda)\boldsymbol{x}^{(2)} = \left(\lambda x_1^{(1)} + (1-\lambda)x_1^{(2)}, \lambda x_2^{(1)} + (1-\lambda)x_2^{(2)}\right)$$

$$f\left(\lambda \boldsymbol{x}^{(1)} + (1-\lambda)\boldsymbol{x}^{(2)}\right) \le \lambda f\left(\boldsymbol{x}^{(1)}\right) + (1-\lambda) f(\boldsymbol{x}^{(2)})$$

因此，$f(x)$为凸函数

从定义出发进行凸函数的判断，计算量往往比较大

■ 凸函数

➢一阶可微凸函数的判别

◆定理2.2.7 设S是E^n中非空开凸集，$f(x)$是定义在S上的可微函数，则$f(x)$为凸函数的充要条件是对任意两点$x^{(1)},x^{(2)}\in S$,都有

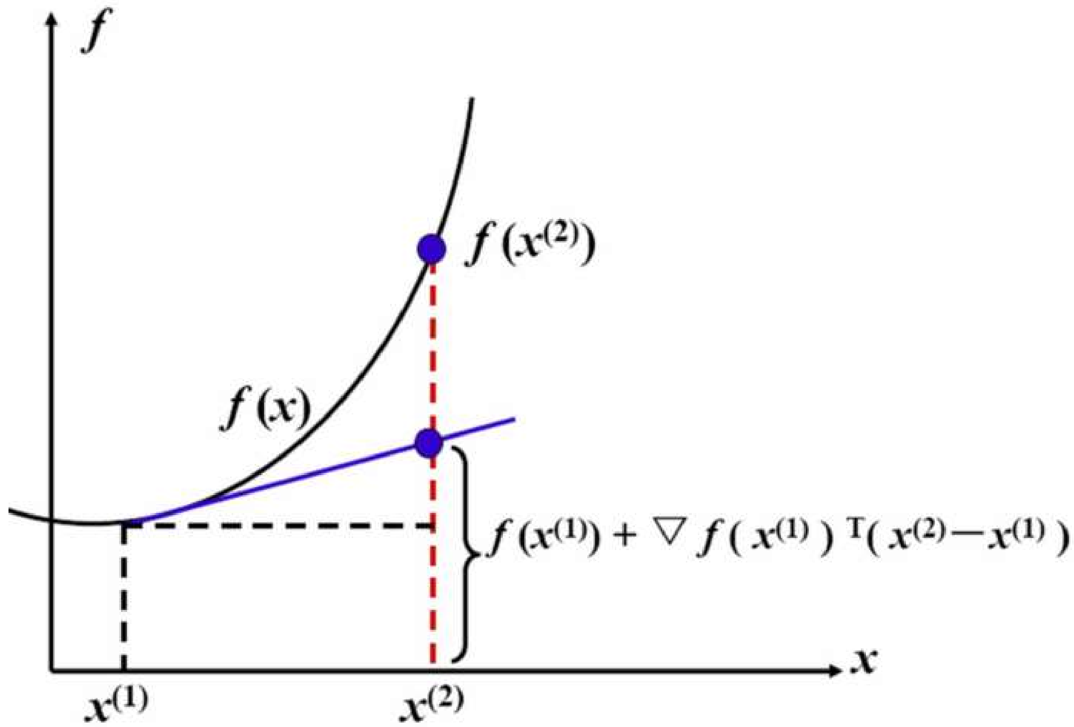
$$f(x^{(2)})\geqslant f(x^{(1)})+\nabla f(x^{(1)})^T(x^{(2)}-x^{(1)})$$

而$f(x)$为严格凸函数的充要条件是对任意的互不相同的$x^{(1)},x^{(2)}\in S$，成立

$$f(x^{(2)})> f(x^{(1)})+\nabla f(x^{(1)})^T(x^{(2)}-x^{(1)})$$

■ 凸函数

➢一阶可微凸函数的判别



■ 凸函数

➢二阶可微凸函数的判别

◆定理2.2.8 设S是E^n中非空开凸集，$f(x)$是定义在S上的二次可微函数，则$f(x)$为凸函数的充要条件是在每一点$x\in S$处，Hessian矩阵半正定

◆定理2.2.9 设S是E^n中非空开凸集，$f(x)$是定义在S上的二次可微函数，如在每一点$x\in S$，Hessian矩阵正定，则$f(x)$为严格凸函数

■ 凸函数

➢二阶可微凸函数的判别

如何判定一个矩阵为半正定矩阵?

判断方法1：若所有特征值均不小于零，则称为半正定。

判断方法2：所有的主子式非负。

如何判定一个矩阵为正定矩阵?

判断方法1：若所有特征值均大于零，则称为正定。

判断方法2：所有的顺序主子式均大于0。

■ 凸函数

➢二阶可微凸函数的判别

$$\begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}$$

$$\begin{bmatrix} 2 & -4 \\ -4 & 2 \end{bmatrix}$$

$$(2 + e^{x_1+x_2}) \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

$$e^{-(x_1+x_2)} \begin{bmatrix} x_1 - 2 & x_1 - 1 \\ x_1 - 1 & x_1 \end{bmatrix}$$

■ 凸函数

➢二阶可微凸函数的判别

$$\begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}$$

半正定矩阵

$$\begin{bmatrix} 2 & -4 \\ -4 & 2 \end{bmatrix}$$

不定矩阵

$$(2+e^{x_1+x_2})\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

半正定矩阵

$$e^{-(x_1+x_2)}\begin{bmatrix} x_1-2 & x_1-1 \\ x_1-1 & x_1 \end{bmatrix}$$

不定矩阵

■ 凸函数

➢凸函数的判别

$$f(x_1, x_2) = x_1^2 - 2x_1x_2 + x_2^2 + x_1 + x_2;$$

解： 1、判断f(x)是一阶可微还是二阶可微

二阶可微

2、写出f(x)的海塞矩阵

$$f(x) = \frac{1}{2}(x_1, x_2)\begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}(x_1, x_2)^T + x_1 + x_2$$

海塞矩阵A= $\begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}$

3、判断海塞矩阵是否半正定

■ 作业

P23~24：

第**1**题

第**9**题的偶数项