# 优化模型与软件工具

主讲教师：董庆兴

华中师范大学 信息管理学院
qxdong@mail.ccnu.edu.cn


2017 年 10 月 10 日

# 大纲

# 范数

满足以下条件的函数 $f:\mathbb{R}^n \to \mathbb{R}, \mathbf{dom}\ f = \mathbb{R}^n$ 称为范数:

- $f$ 非负：对所有的 $\mathbf{x} \in \mathbb{R}^n$ 有 $f(\mathbf{x}) \geq 0$
- $f$ 正定：仅对 $\mathbf{x} = 0$ 有 $f(\mathbf{x}) = 0$
- $f$ 齐次：对所有有 $\mathbf{x} \in \mathbb{R}^n$ 和 $t \in \mathbb{R}$ 有 $f(t\mathbf{x}) = |t|f(\mathbf{x})$
- $f$ 满足三角不等式：$\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ 有 $f(\mathbf{x}+\mathbf{y}) \leq f(\mathbf{x}) + f(\mathbf{y})$

# 收敛

**收敛**

设有序列 $\{\mathbf{x}_k\} \subseteq \mathbb{R}^n$ 及向量 $\mathbf{x}^* \in \mathbb{R}^n$，当 $k \to \infty$ 时有 $\|\mathbf{x}_k - \mathbf{x}^*\| \to 0$ 则称 $\{\mathbf{x}_k\}$ 收敛于 $\mathbf{x}^*$，记作 $\{\mathbf{x}_k\} \to \mathbf{x}$

- $\{\mathbf{x}_k\}$ 收敛于 $\mathbf{x}^* \in \mathbb{R}^n$ 当且仅当 $\mathbf{x}_k$ 的第 $i$ 个分量收敛到 $\mathbf{x}^*$ 的第 $i$ 个分量：$\left.|\mathbf{x}_{k(i)} - \mathbf{x}^*_{(i)}| \to 0\right|_{k\to\infty}$

# 几何基本概念

**直线** 如果 $\mathbf{x}_1 \neq \mathbf{x}_2$ 为 $\mathbb{R}^n$ 中的两个点，那么 $y = \theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2, \theta \in \mathbb{R}$ 组成一条穿越 $\mathbf{x}_1$ 和 $\mathbf{x}_2$ 的直线

**线段** 在上述定义中，如果 $\theta \in [0,1]$ 则构成 $\mathbf{x}_1$ 和 $\mathbf{x}_2$ 之间的线段

**球** $\mathcal{B}(\mathbf{x}, r) = \{\mathbf{y} | \|\mathbf{y}-\mathbf{x}\| \leq r\}$ 即中心为 $\mathbf{x}$ 半径为 $r$ 的球

**单位球** $\mathcal{B} = \{\mathbf{x} \in \mathbb{R}^n | \|\mathbf{x}\| \leq 1\}$，$\mathcal{B}$ 是关于原点对称有界闭凸集

**开球** $\mathcal{B}(\mathbf{x}, r) = \{\mathbf{y} | \|\mathbf{y}-\mathbf{x}\| < r\}$ 即中心为 $\mathbf{x}$ 半径为 $r$ 的球

# 开集和内部

**内点**

对于 $\mathbf{x} \in C \subseteq \mathbb{R}^n$，如果存在 $\epsilon > 0$ 满足 $\{\mathbf{y} | \ \|\mathbf{y} - \mathbf{x}\|_2 \leq \epsilon\} \subset C$，即存在一个以 $\mathbf{x}$ 为中心的完全属于 $C$ 的球，就称其为 $C$ 的内点

**内部**

$C$ 的所有内点组成的集合称为 $C$ 的内部，用 $\textbf{int}\ C$ 表示

**开集**

如果 $\textbf{int}\ C = C$，那么 $C$ 为开集，也就是 $C$ 的每个点都是内点

- 例子：$X = \{\mathbf{x} \in \mathbb{R}^2 | x_1 \geq 0, x_2 > 0\}$ 的内部是?
- 凸集的内部也是凸集

# 闭集

**补集**

对于 $C \subseteq \mathbb{R}^n$ 的补集就是 $\mathbb{R}^n \setminus C = \{\mathbf{x} | \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \notin C\}$

**闭集**

如果集合 $C \subseteq \mathbb{R}^n$ 的补集 $\mathbb{R}^n \setminus C = \{\mathbf{x} | \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \notin C\}$ 是开集，则 $C$ 为闭集

- 例子：$X = \{\mathbf{x} \in \mathbb{R}^2 | x_1 \geq 0, x_2 > 0\}$ 是开集还是闭集?
- $\mathbb{R}^n$ 中 $\mathbb{R}^n$ 和 $\emptyset$ 都既是开集又是闭集
- 任意个闭集的交集依然是闭集
- 有限个闭集的并集依然是闭集

# 闭包

**闭包**

集合补集的内部的补集 $\mathbf{cl}\ \ C = \mathbb{R}^n \setminus \mathbf{int}\ (\mathbb{R}^n \setminus C)$

- 如果存在一个序列 $\{\mathbf{x}_k\} \subseteq C$ 使得 $\{\mathbf{x}_k\} \to \mathbf{x}$，则称 $\mathbf{x}$ 是非空集合 $C \subseteq \mathbb{R}^n$ 的闭包点
- 那么 $C \subseteq \mathbb{R}^n$ 的闭包点的集合就是 $\mathbf{cl}\ \ C$
- 一个集合是闭集当且仅当该集合包含其闭包：$C$ 是闭集等价于 $\mathbf{cl}\ \ C \subseteq C$
- 凸集的闭包依然是凸集

# 边界

**边界**

非空集合 $C \subseteq \mathbb{R}^n$ 的边界 $\mathbf{bd}\ \ C$ 是该集合闭包和其补集的闭包的交集：$\mathbf{bd}\ \ C = \mathbf{cl}\ \ C \cap \mathbf{cl}\ (\mathbb{R}^n \setminus C)$

- 集合 $C$ 的边界也可定义为 $\mathbf{bd}\ \ C = \mathbf{cl}\ \ C \setminus \mathbf{int}\ \ C$
- 边界点 $\mathbf{x} \in \mathbf{bd}\ \ C$，对于所有 $\epsilon > 0$，存在 $\mathbf{y} \in C$ 和 $\mathbf{z} \notin C$ 满足 $\|\mathbf{y} - \mathbf{x}\|_2 \leq \epsilon, \|\mathbf{z} - \mathbf{x}\|_2 \leq \epsilon$

# 仿射

**仿射集合** 如果通过 $C \subseteq \mathbb{R}^n$ 中任意两个不同点的直线仍然在 $C$ 中，那么集合 $C$ 是仿射的。也就是，对于 $\mathbf{x}_1 \neq \mathbf{x}_2 \in C, \theta \in \mathbb{R}$ 为 $\mathbb{R}^n$ 中的两个点，那么有 $\mathbf{y} = \theta \mathbf{x}_1 + (1-\theta)\mathbf{x}_2 \in C$

**仿射组合** 如果 $\sum_1^k \theta_i = 1$，我们称具有 $y = \sum_{i=1}^k \theta_i x_i$ 形式的点为 $x_1, \cdots, x_k$ 的仿射组合

**仿射包** $C \subseteq \mathbb{R}^n$ 中的点的所有的仿射组合组成的集合为 $C$ 的仿射包，记为

$$\mathbf{aff}\ C = \left\{ \sum_{i=1}^k \theta_i \mathbf{x}_i | \mathbf{x}_1, \cdots, \mathbf{x}_k \in C, \sum_{i=1}^k \theta_i = 1 \right\}$$

- 仿射集：$\{\mathbf{x} | A\mathbf{x} = \mathbf{b}\}$
- 仿射包：$\mathbb{R}^2$ 上的单位圆环 $\{\mathbf{x} \in \mathbb{R}^2 | x_1^2 + x_2^2 = 1\}$ 的仿射包是 $\mathbb{R}^2$

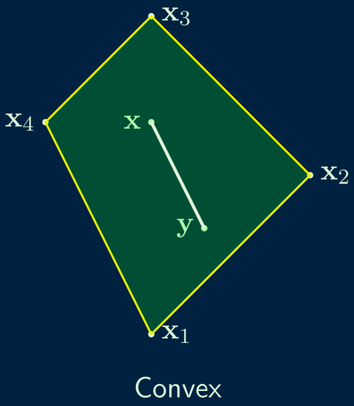
# 相对内部和相对边界

**相对内部** 定义集合 $C \subseteq \mathbb{R}^n$ 的相对内部为 $\mathbf{aff}\ C$ 的内部，即
$\mathbf{relint}\ C = \{\mathbf{x} \in C | B(\mathbf{x}, r) \cap \mathbf{aff}\ C \subseteq C, \text{ for some } r > 0\}$

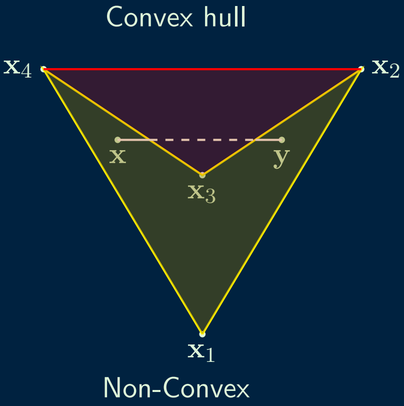
**相对边界** 同样可以定义 $\mathbf{cl}\ C \setminus \mathbf{relint}\ C$ 为 $C$ 的相对边界

- $C = \{\mathbf{x} | \theta \mathbf{x}_1 + (1-\theta)\mathbf{x}_2, \theta \in [0,1]\}$ 内部为空集，边界为自身，相对内部是连接 $\mathbf{x}_1$ 和 $\mathbf{x}_2$ 的线段（去掉端点），相对边界为 $\mathbf{x}_1$ 和 $\mathbf{x}_2$
- $C = \{\mathbf{x} \in \mathbb{R}^3 | -1 \leq x_1 \leq 1, -1 \leq x_2 \leq 1, x_3 = 0\}$。其仿射包为 $(x_1, x_2)-$ 平面，$C$ 的内部为空，边界为集合自身；但其相对内部为 $\mathbf{relint}\ C = \{\mathbf{x} \in \mathbb{R}^3 | -1 < x_1 < 1, -1 < x_2 < 1, x_3 = 0\}$，相对边界为边框 $\{\mathbf{x} \in \mathbb{R}^3 | \max\{|x_1|, |x_2|\} = 1, x_3 = 0\}$

# 凸集和凸包



$\mathbf{x}_3$
$\mathbf{x}_4$
$\mathbf{x}$
$\mathbf{x}_2$
$\mathbf{y}$
$\mathbf{x}_1$
Convex

Convex hull
$\mathbf{x}_4$
$\mathbf{x}_2$
$\mathbf{x}$
$\mathbf{y}$
$\mathbf{x}_3$
$\mathbf{x}_1$
Non-Convex

# 凸包

**凸组合** 我们称点 $\mathbf{x}=\sum_{i=1}^{k}\theta_i\mathbf{x}_i$ 为点 $\mathbf{x}_1,\cdots,\mathbf{x}_k$ 的凸组合，其中,$\sum_{i=1}^{k}\theta_i=1\ \theta_i\geq 0, i=1,\cdots,k$

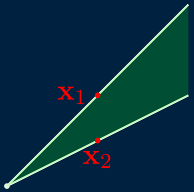
**凸包** 集合 $C$ 中所有点的凸组合的集合称为 $C$ 的凸包 $\mathbf{conv}\ \ C$

- 集合的凸包是包含该集合的最小凸集

# 锥

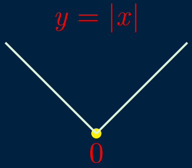
**锥** 如果对于任意 $\mathbf{x} \in C$ 和 $\theta \geq 0$ 都有 $\theta\mathbf{x} \in C$，则称 $C$ 是锥

**凸锥** 如果集合 $C$ 是锥并且是凸的，则 $C$ 为凸锥，也就是对于任意的 $\mathbf{x}_1, \mathbf{x}_2 \in C$ 和 $\theta_1, \theta_2 \geq 0$ 有 $\theta_1\mathbf{x}_1 + \theta_2\mathbf{x}_2 \geq 0$




Convex cone with origin

Convex cone without origin

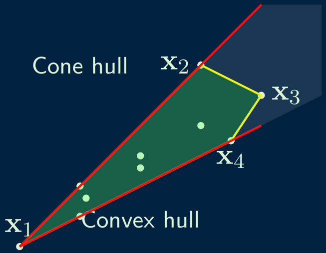



Nonconvex cone

# 锥包

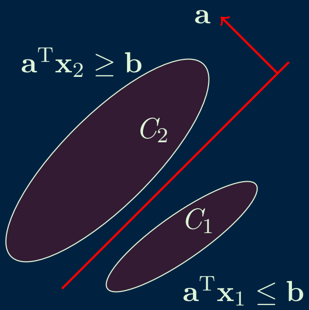
**锥组合** 我们称点 $\mathbf{x}=\sum_{i=1}^{k}\theta_i\mathbf{x}_i$ 为点 $\mathbf{x}_1,\cdots,\mathbf{x}_k$ 的锥组合（非负线性组合），其中,$\theta_i\geq 0, i=1,\cdots,k$。如果 $\mathbf{x}_i$ 均属于凸锥 $C$, 那么他们的每一个锥组合都在 $C$ 中

**锥包** 集合 $C$ 中所有点的锥组合的集合称为 $C$ 的锥包，也就是 $\{\theta_1\mathbf{x}_1+\cdots,\theta_k\mathbf{x}_k|\mathbf{x}_i\in C,\theta_i\geq 0, i=1,\cdots,k\}$

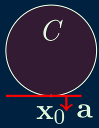


# 分离超平面

- 称集合 $C_1$ $C_2$ 被超平面 $H=\{\mathbf{a}^{\mathrm{T}}\mathbf{x}=\mathbf{b}\}$ 分离如果两集合分别位于 $H$ 的不同的半空间内，也就是 $\mathbf{a}^{\mathrm{T}}\mathbf{x}_1\le\mathbf{b}\le\mathbf{a}^{\mathrm{T}}\mathbf{x}_2,\forall\mathbf{x}_1\in C_1,\forall\mathbf{x}_2\in C_2$ 或 $\mathbf{a}^{\mathrm{T}}\mathbf{x}_2\le\mathbf{b}\le\mathbf{a}^{\mathrm{T}}\mathbf{x}_1,\forall\mathbf{x}_1\in C_1,\forall\mathbf{x}_2\in C_2$
- 超平面 $H=\{\mathbf{a}^{\mathrm{T}}\mathbf{x}=\mathbf{b}\}$ 就是集合 $C_1,C_2$ 的分离超平面



# 支撑超平面




- 令 $C$ 是 $\mathbb{R}^n$ 中的凸集，$C$ 的支撑半空间是包含 $C$ 的闭半空间，且有一个点在 $C$ 的边界上。$C$ 的支撑超平面就是 $C$ 支撑半空间的边界，它本身是一个超平面
- $C \subseteq \mathbb{R}^n$ 而 $\mathbf{x}_0$ 是其边界 $\mathbf{bd}\ C$ 上一点，即 $\mathbf{x}_0 \in (\mathbf{cl}\ C \setminus \mathbf{int}\ C)$，如果 $\mathbf{a} \neq 0$，并满足 $\forall \mathbf{x} \in C, \mathbf{a}^\mathrm{T}\mathbf{x} \le \mathbf{a}^\mathrm{T}\mathbf{x}_0$，则称超平面 $H = \{\mathbf{a}^\mathrm{T}\mathbf{x} = \mathbf{b}\}$ 为集合 $C$ 在点 $\mathbf{x}_0$ 处的支撑（支持）超平面
- 几何解释就是超平面 $H$ 与 $C$ 相切于点 $\mathbf{x}_0$，而且半空间 $\{\mathbf{x}|\mathbf{a}^\mathrm{T}\mathbf{x} \le \mathbf{a}^\mathrm{T}\mathbf{x}_0\}$ 包含 $C$
- 如果点 $x_0 \in \mathbf{cl}\ C$ 那么一个超平面 $H$ 分离了 $C$ 和单点集合 $\{\mathbf{x}_0\}$，则称 $H$ 在 $\mathbf{x}_0$ 处支撑了 $C$
- 换句话说，$C$ 的支撑超平面可以表示成 $H = \{\mathbf{x}|\mathbf{a}^\mathrm{T}\mathbf{x} = \mathbf{b}, \mathbf{a} \neq 0\}$，其中对于每个 $\mathbf{x} \in C, \mathbf{a}^\mathrm{T}\mathbf{x} \le \mathbf{b}$ 且至少有一个点 $\mathbf{x}_0 \in C$ 使得 $\mathbf{a}^\mathrm{T}\mathbf{x} = \mathbf{b}$
- 因此 $C$ 的支撑超平面与一个线性函数有关，该函数找到 $C$ 上的最大值。经过给定点 $\mathbf{x}_0 \in C$ 的支撑超平面对应于向量 $\mathbf{a}$，它是 $C$ 在 $\mathbf{x}_0$ 处的法向量

# 投影定理

## 投影定理

对于非空闭凸集 $C \in \mathbb{R}^n$ 和向量 $\mathbf{z} \in \mathbb{R}^n$，则存在一个唯一的向量 $\mathbf{x}^* \in C$ 为 $\min\limits_{\mathbf{x}\in C} \|\mathbf{z}-\mathbf{x}\|_2$ 的解，$\mathbf{x}$ 被称作 $\mathbf{z}$ 在 $C$ 上的投影并且满足

$$(\mathbf{z}-\mathbf{x}^*)^{\mathrm{T}}(\mathbf{x}-\mathbf{x}^*) \leq 0, \forall \mathbf{x} \in C \tag{1}$$

证明: 求解 $\min\limits_{\mathbf{x}\in C} \|\mathbf{z}-\mathbf{x}\|_2$ 等价于求解凸可微函数 $f(\mathbf{x}) = \frac{\|\mathbf{z}-\mathbf{x}\|_2^2}{2}$，可知 $\mathbf{x}^*$ 为 $f$ 在 $C$ 上的极小值当且仅当 $\nabla f(\mathbf{x}^*)^{\mathrm{T}}(\mathbf{x}-\mathbf{x}^*) \geq 0, \forall \mathbf{x} \in C$,而 $\nabla f(\mathbf{x}^*) = (\mathbf{x}^* - \mathbf{z})$, 因此可得式 (**??**)。
假设有 $\mathbf{x}_1^*$ 和 $\mathbf{x}_2^*$ 同时是 $\mathbf{z}$ 在 $C$ 上的投影，那由式 (**??**) 可得

$$(\mathbf{z}-\mathbf{x}_1^*)^{\mathrm{T}}(\mathbf{x}_2^*-\mathbf{x}_1^*) \leq 0, (\mathbf{z}-\mathbf{x}_2^*)^{\mathrm{T}}(\mathbf{x}_1^*-\mathbf{x}_2^*) \leq 0$$

不等式两端相加可得 $(\mathbf{x}_2^*-\mathbf{x}_1^*)^{\mathrm{T}}(\mathbf{x}_2^*-\mathbf{x}_1^*) = \|\mathbf{x}_2^*-\mathbf{x}_1^*\|_2^2 \leq 0$，由范数几何意义可知，$\mathbf{x}_2^* = \mathbf{x}_1^*$，从而投影向量唯一 ■

# 支撑超平面定理

**支撑超平面定理**

设 $C$ 是 $\mathbb{R}^n$ 中的非空凸集，向量 $\bar{\boldsymbol{x}} \in \mathbb{R}^n$，若 $\bar{\boldsymbol{x}} \notin \textbf{int}\ C$，则存在一个穿过 $\bar{\boldsymbol{x}}$ 的超平面使得 $X$ 属于它的一个闭半空间，即存在向量 $\boldsymbol{a} \neq \mathbf{0}$ 满足

$$\mathbf{a}^{\mathrm{T}}\bar{\mathbf{x}} \leq \mathbf{a}^{\mathrm{T}}\mathbf{x},\ \forall \mathbf{x} \in C \tag{2}$$

证明：考虑 $C$ 的闭包 $\textbf{cl}\ C$，可知 $\textbf{cl}\ C$ 为凸集（作业：证明这一结论）。可知一定存在 $\{\mathbf{x}_k\}$ 为一个 $\textbf{cl}\ C$ 外收敛于 $\bar{\mathbf{x}}$ 的点列即 $\{\mathbf{x}_k\} \to \bar{\mathbf{x}}$ 且 $\forall k, \mathbf{x}_k \notin C$（思考：为什么？）。令 $\mathbf{y}_k$ 为 $\mathbf{x}_k$ 在 $\textbf{cl}\ C$ 上的投影，则根据投影定理有

$$(\mathbf{y}_k - \mathbf{x}_k)^{\mathrm{T}}(\mathbf{x} - \mathbf{y}_k) \geq 0,\ \forall \mathbf{x} \in \textbf{cl}\ C$$

因此对于所有 $\mathbf{x} \in \textbf{cl}\ C$ 和所有 $k$ 有

$$\begin{aligned}
(\mathbf{y}_k - \mathbf{x}_k)^{\mathrm{T}}\mathbf{x} &\geq (\mathbf{y}_k - \mathbf{x}_k)^{\mathrm{T}}\mathbf{y}_k \\
&= (\mathbf{y}_k - \mathbf{x}_k)^{\mathrm{T}}(\mathbf{y}_k - \mathbf{x}_k) + (\mathbf{y}_k - \mathbf{x}_k)^{\mathrm{T}}\mathbf{x}_k \\
&\geq (\mathbf{y}_k - \mathbf{x}_k)^{\mathrm{T}}\mathbf{x}_k
\end{aligned}$$

# 支撑超平面定理：证明

所以我们可以将上述不等式重写为

$$\mathbf{a}_k^{\mathrm{T}}\mathbf{x} \geq \mathbf{a}_k^{\mathrm{T}}\mathbf{x}_k,\ \forall \mathbf{x} \in \mathbf{cl}\ \ C,\ \forall k \tag{3}$$

其中

$$\mathbf{a}_k = \frac{\mathbf{y}_k - \mathbf{x}_k}{\|\mathbf{y}_k - \mathbf{x}_k\|}$$

可知 $\|\mathbf{a}_k\| = 1, \forall k$，因此 $\{\mathbf{a}_k\}$ 有界。从而 $\{\mathbf{a}_k\}$ 存在一个收敛子序列 $\{\mathbf{a}_k\}_{\mathcal{K}}$（Bolzano-Weierstrass 定理）收敛到 $a \neq 0$，由 $\mathbf{x}_k \to \bar{\mathbf{x}}, \mathbf{a}_k \to \mathbf{a}$，考虑式 (**??**) 中 $\mathbf{a}_k \in \{\mathbf{a}_k\}_{\mathcal{K}}$，令 $k \to \infty$ 可得

$$\mathbf{a}^{\mathrm{T}}\mathbf{x} \geq \mathbf{a}^{\mathrm{T}}\bar{\mathbf{x}},\ \forall \mathbf{x} \in C$$

命题得证。■