

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
All rights reserved

2017 年 11 月 14 日

大纲

1. 分离定理
2. 广义不等式
3. 对偶锥与广义不等式
4. 最小元与极小元

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 $\min_{\mathbf{x} \in C} \|\mathbf{z} - \mathbf{x}\|_2$ 的解, \mathbf{x}^* 被称作 \mathbf{z} 在 C 上的投影并且满足

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

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 $\bar{x} \in \mathbb{R}^n$, 若 $\bar{x} \notin \text{int } C$, 则存在一个穿过 \bar{x} 的超平面使得 X 属于它的一个闭半空间, 即存在向量 $a \neq 0$ 满足

$$a^T \bar{x} \leq a^T x, \forall x \in C \quad (2)$$

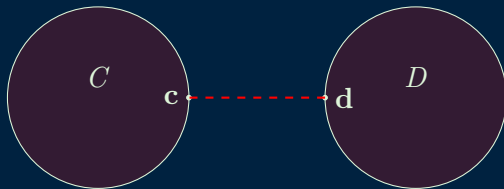
超平面分离定理

超平面分离定理

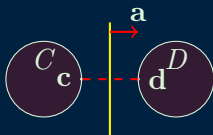
对于两个不相交的非空凸集 C 和 D , 存在 $\mathbf{a} \neq 0$ 和 \mathbf{b} 使得对于所有的 $\mathbf{x}_1 \in C$ 和 $\mathbf{x}_2 \in D$ 有 $\mathbf{a}^T \mathbf{x}_1 \leq \mathbf{b}, \mathbf{a}^T \mathbf{x}_2 \geq \mathbf{b}$, 也就是存在分离超平面 $H = \{\mathbf{a}^T \mathbf{x} = \mathbf{b}\}$

为证明上述定理, 我们首先给出一种特殊情况 (作业: 证明更普通的情况), 定义两集合之间的距离为

$\text{dist}(C, D) = \inf\{\|\mathbf{u} - \mathbf{v}\|_2 | \mathbf{u} \in C, \mathbf{v} \in D\} > 0$, 并且存在 $\mathbf{c} \in C, \mathbf{d} \in D$ 达到这个最小距离, 也就是 $\|\mathbf{c} - \mathbf{d}\|_2 = \text{dist}(C, D)$,



超平面分离定理



- 考虑上述特殊情况，则我们可以构建一个超平面 $H = \{\mathbf{a}^T \mathbf{x} = \mathbf{b}\}$ 垂直并且平分线段 \mathbf{cd}
- 已知 \mathbf{c}, \mathbf{d} 坐标，可以很方便的求出超平面 $H = \{\mathbf{a}^T \mathbf{x} = \mathbf{b}\}$ 的具体值
- 令 $\mathbf{a} = \mathbf{d} - \mathbf{c}, \mathbf{b} = \frac{\|\mathbf{d}\|_2^2 - \|\mathbf{c}\|_2^2}{2}$ ，仿射函数 $f(\mathbf{x}) = \mathbf{a}^T \mathbf{x} - \mathbf{b} = (\mathbf{d} - \mathbf{c})^T (\mathbf{x} - \frac{\mathbf{d} + \mathbf{c}}{2})$ 在 C 中非正而在 D 中非负，即超平面 $\{\mathbf{x} | \mathbf{a}^T \mathbf{x} = \mathbf{b}\}$ 分离了 C 和 D

- 首先我们用反证法证明 f 在 D 中非负（证明 f 在 C 中非正思路相同，只需交换 C, D 并考虑 $-f$ ），假设存在一点 $\mathbf{u} \in D$ ，且

$$f(\mathbf{u}) = (\mathbf{d} - \mathbf{c})^T (\mathbf{u} - \frac{\mathbf{d} + \mathbf{c}}{2}) = (\mathbf{d} - \mathbf{c})^T (\mathbf{u} - \mathbf{d} + \frac{\mathbf{d} - \mathbf{c}}{2}) = (\mathbf{d} - \mathbf{c})^T (\mathbf{u} - \mathbf{d}) + \frac{\|\mathbf{d} - \mathbf{c}\|_2^2}{2} < 0,$$
 可知

$$(\mathbf{d} - \mathbf{c})^T (\mathbf{u} - \mathbf{d}) < 0 \quad (3)$$

- 而由于 \mathbf{d} 是 D 中离 \mathbf{c} 最近的点，那么 \mathbf{d} 是 \mathbf{c} 在 D 中的投影，从而由式 (1) 有

$$(\mathbf{c} - \mathbf{d})^T (\mathbf{u} - \mathbf{d}) \leq 0$$

- 这与式 (3) 矛盾。因此 f 在 D 中非负

仿射集与凸集分离

- 设 C 为凸集, D 为仿射集, 也就是 $D = \{F\mathbf{y} + \mathbf{g} | \mathbf{y} \in \mathbb{R}^m\}$, 其中 $F \in \mathbb{R}^{n \times m}$ 。设 C, D 不相交, 那么根据超平面分离定理, 存在 $\mathbf{a} \neq 0$ 和 b 使得 $\mathbf{a}^T \mathbf{x} \leq b, \forall \mathbf{x} \in C$; $\mathbf{a}^T \mathbf{x} \geq b, \forall \mathbf{x} \in D$
- 从而对于任意的 $\mathbf{y} \in \mathbb{R}^m$ 都有 $\mathbf{a}^T F\mathbf{y} + \mathbf{a}^T \mathbf{g} \geq b$, 也就是 $\mathbf{a}^T F\mathbf{y} \geq b - \mathbf{a}^T \mathbf{g}$
- 在 \mathbb{R}^m 上, 只有当一个线性函数为零的时候它才是有界的, 因此可知 $\mathbf{a}^T F = 0, b - \mathbf{a}^T \mathbf{g} \leq 0$
- 从而仿射集与凸集分离等价于: 存在 $\mathbf{a} \neq 0$ 使得 $\mathbf{a}^T F = 0$ 和 $\mathbf{a}^T \mathbf{x} \leq \mathbf{a}^T \mathbf{g}$ 对于所有 $\mathbf{x} \in C$ 都成立

严格分离

- 称集合 C_1 C_2 被超平面 $H = \{\mathbf{a}^T \mathbf{x} = \mathbf{b}\}$ **严格分离** 如果两集合分别位于 H 的不同的半空间内, 也就是

$$\mathbf{a}^T \mathbf{x}_1 < \mathbf{b} < \mathbf{a}^T \mathbf{x}_2, \forall \mathbf{x}_1 \in C_1, \forall \mathbf{x}_2 \in C_2$$

或

$$\mathbf{a}^T \mathbf{x}_2 < \mathbf{b} < \mathbf{a}^T \mathbf{x}_1, \forall \mathbf{x}_1 \in C_1, \forall \mathbf{x}_2 \in C_2$$

- 不相交的凸集并不一定能够被超平面严格分离 (即使集合是闭集):
例如 $C = \{\mathbf{x} \in \mathbb{R}^2 | x_1 = 0\}$ 和 $D = \{\mathbf{x} \in \mathbb{R}^2 | x_1 > 0, x_1 x_2 > 1\}$

点和凸集的严格分离

点和凸集的严格分离

令集合 C 为闭凸集, $\mathbf{x}_0 \notin C$, 那么存在将 \mathbf{x}_0 与 C 严格分离的超平面

- 对于足够小的 $\varepsilon > 0$, 存在两个不相交的集合 C 和 $B(\mathbf{x}_0, \varepsilon)$ 。根据超平面分离定理, 存在 $\mathbf{a} \neq 0$ 和 b , 使得对于任意 $\mathbf{x} \in C$ 有 $\mathbf{a}^T \mathbf{x} \leq b$; 对于任意 $\mathbf{x} \in D$ 有 $\mathbf{a}^T \mathbf{x} \geq b$
- 由于 $B(\mathbf{x}_0, \varepsilon) = \{\mathbf{x}_0 + \mathbf{u} \mid \|\mathbf{u}\|_2 \leq \varepsilon\}$ 从而有
$$\mathbf{a}^T (\mathbf{x}_0 + \mathbf{u}) \geq b, \forall \|\mathbf{u}\|_2 \leq \varepsilon$$
- 由于 $\mathbf{u} = \frac{-\varepsilon \mathbf{a}}{\|\mathbf{a}\|_2} = \arg \min \mathbf{a}^T (\mathbf{x}_0 + \mathbf{u})$, 代入上式可得 $\mathbf{a}^T \mathbf{x}_0 - \varepsilon \|\mathbf{a}\|_2 \geq b$
- 所以仿射函数 $f(\mathbf{x}) = \mathbf{a}^T \mathbf{x} - b - \frac{\varepsilon \|\mathbf{a}\|_2}{2}$ 在 C 上是负的, 而在 \mathbf{x}_0 点是正的, 从而 C 和 \mathbf{x}_0 被严格分离
- 作为一个直接的结果, 我们可以得到: 一个闭凸集是包含它的所有半空间的交集
- 令 C 为闭凸集, S 为所有包含 C 的半空间, 显然 $\mathbf{x} \in C \Rightarrow \mathbf{x} \in S$ 。也就是 $C \subseteq S$
- 接下来要证明 $S \subseteq C$, 假设不成立则存在 $\mathbf{x} \in S, \mathbf{x} \notin C$, 那么根据点和凸集严格分离定理, 存在一个将 C 和 \mathbf{x} 严格分离的超平面, 也就是存在一个包含 C 但是并不包含 \mathbf{x} 的半空间。也就是 $\mathbf{x} \notin S$, 从而矛盾, 有 $S \subseteq C$

超平面分离定理的逆定理

- 分离超平面的存在表明 C 和 D 不相交是不成立的
- 例子: $C = D = \{0\} \subseteq \mathbf{R}$, 超平面 $x = 0$ 可以分离 C 和 D
- 凸性之外加条件可以成立: 例如 C, D 为凸集且 C 为开集, 如存在一个仿射函数 f 在 C 中非正而在 D 中非负, 那么 C 和 D 不相交
- f 在 C 上非正, 那么如果 f 在开集 C 上某一点为零, 那么必然会在这个点附近取正值, 这与前述矛盾, 所以 f 在 C 上为负

择一定理

严格线性不等式的择一定理

严格线性不等式

$$A\mathbf{x} \prec \mathbf{b} \quad (4)$$

无解的充要条件是凸集 $C = \{\mathbf{b} - A\mathbf{x} | \mathbf{x} \in \mathbb{R}^n\}$ 与集合 $D = \mathbb{R}_{++}^m = \{\mathbf{y} \in \mathbb{R}^m | \mathbf{y} \succ 0\}$ 不相交

- 集合 D 是开集，集合 C 是仿射集合。根据前述的结论， C 和 D 不相交的充要条件是，存在分离超平面，即存在非零的 $\lambda \in \mathbb{R}^m$ 和 $\mu \in \mathbb{R}$ 使得 C 中 $\lambda^T \mathbf{y} \leq \mu$ 而 D 中 $\lambda^T \mathbf{y} \geq \mu$
- 第一个条件意味着对于所有的 \mathbf{x} 都有 $\lambda^T (\mathbf{b} - A\mathbf{x}) \leq \mu$ ，这说明 $A^T \lambda = 0, \lambda^T \mathbf{b} \leq \mu$
- 第二个条件意味着 $\lambda^T \mathbf{y} \geq \mu$ 对于所有的 $\mathbf{y} \succ 0$ 均成立，这表明 $\mu \leq 0$ 且 $\lambda \succeq 0, \lambda \neq 0$
- 将上述结果放在一起，可知不等式组 (4) 无解的充要条件是存在 $\lambda \in \mathbb{R}^m$ 使得

$$\lambda \neq 0, \lambda \succeq 0, A^T \lambda = 0, \lambda^T \mathbf{b} \leq 0 \quad (5)$$

- 这些不等式和等式关于 $\lambda \in \mathbb{R}^m$ 也是线性的，因此我们称 (4) 和 (5) 构成一对**择一选择**：对于任意的 A 和 \mathbf{b} ，两者中仅有一组有解

正常锥

称锥 $K \subseteq \mathbb{R}^n$ 为正常锥 (**Proper Cone**)，如果它满足：

- K 是凸的
- K 是闭的，包含边界
- K 是实的，即具有非空内部
- K 是尖的（有端点），即不包含直线（或者等价地， $\mathbf{x} \in K, -\mathbf{x} \in K \Rightarrow \mathbf{x} = 0$ ）

例子：平面上包含数轴的第一象限（非负象限）

广义不等式

- 正常锥 $K \subseteq \mathbb{R}^n$ 下的偏序关系为

$$\mathbf{x} \preceq_K \mathbf{y} \iff \mathbf{y} - \mathbf{x} \in K$$

- 当 $K = \mathbb{R}_+$ 时, 偏序关系 \preceq_K 就是通常意义上 \mathbb{R} 中的序 \leq
- 例子: 非负象限 $K = \mathbb{R}_+^n$ 是一个正常锥。相应的广义不等式 \preceq_K 对应于向量间的分量不等式, 及 $\mathbf{x} \preceq_K \mathbf{y}$ 等价于 $x_i \leq y_i, i = 1, \dots, n$
- 正常锥 $K \subseteq \mathbb{R}^n$ 下的严格偏序关系为

$$\mathbf{x} \prec_K \mathbf{y} \iff \mathbf{y} - \mathbf{x} \in \text{int } K$$

- 当 $K = \mathbb{R}_+$ 时, 严格偏序关系 \prec_K 就是通常意义上 \mathbb{R} 中的序 $<$
- 因此广义不等式包含了 \mathbb{R} 中的 (不严格和严格) 不等式

广义不等式性质

广义不等式 \preceq_K 有如下性质 (课后作业证明):

- 加法保序: 如果 $\mathbf{x} \preceq_K \mathbf{y}$, 且 $\mathbf{u} \preceq_K \mathbf{v}$, 那么 $\mathbf{x} + \mathbf{u} \preceq_K \mathbf{y} + \mathbf{v}$
- 传递性: 如果 $\mathbf{x} \preceq_K \mathbf{y}$, 且 $\mathbf{y} \preceq_K \mathbf{z}$, 那么 $\mathbf{x} \preceq_K \mathbf{z}$
- 非负数乘保序: 如果 $\mathbf{x} \preceq_K \mathbf{y}$, 且 $\alpha \geq 0$, 那么 $\alpha \mathbf{x} \preceq_K \alpha \mathbf{y}$
- 自反: $\mathbf{x} \preceq_K \mathbf{x}$
- 反对称: 如果如果 $\mathbf{x} \preceq_K \mathbf{y}$, 且如果 $\mathbf{y} \preceq_K \mathbf{x}$, 那么 $\mathbf{x} = \mathbf{y}$
- 极限运算保序: 如果对于 $i = 1, 2, \dots$ 均有 $\mathbf{x}_i \preceq_K \mathbf{y}_i$, 当 $i \rightarrow \infty$ 有 $\mathbf{x}_i \rightarrow \mathbf{x}$ 和 $\mathbf{y}_i \rightarrow \mathbf{y}$, 那么 $\mathbf{x} \preceq_K \mathbf{y}$

最小与极小元

- 如果对于 $\forall \mathbf{y} \in S$ 都有 $\mathbf{x} \preceq_K \mathbf{y}$, 则称 $\mathbf{x} \in S$ 是 S (关于广义不等式 \preceq_K) 的**最小元** (minimum element)
- 类似地, 我们可以定义关于广义不等式的**最大元**。并且如果一个集合有最小(或最大)元, 那么它们是唯一的
- 对于 $\mathbf{y} \in S$ 如果有 $\mathbf{y} \preceq_K \mathbf{x}$, 则有 $\mathbf{y} = \mathbf{x}$, 则称 $\mathbf{x} \in S$ 是 S (关于广义不等式 \preceq_K) 的**极小元** (minimal element)
- 同样地, 可以定义关于广义不等式的**极大元**, 一个集合可以有多个极小(大)元

最小元和极小元的集合描述

- 元素 $\mathbf{x} \in S$ 是 S 中的一个**最小元**当且仅当 $S \subseteq \mathbf{x} + K$, 其中 $\mathbf{x} + K$ 表示可以与 \mathbf{x} 相比并且大于等于 (根据 \preceq_K) \mathbf{x} 的所有元素
- 元素 $\mathbf{x} \in S$ 是 S 中的一个**极小元**当且仅当 $\mathbf{x} - K \cap S = \{\mathbf{x}\}$, 其中 $\mathbf{x} - K$ 表示可以与 \mathbf{x} 相比并且小于等于 (根据 \preceq_K) \mathbf{x} 的所有元素

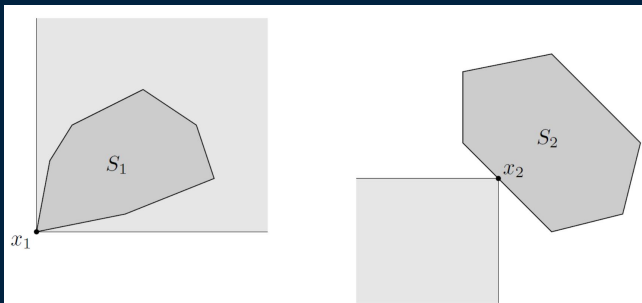
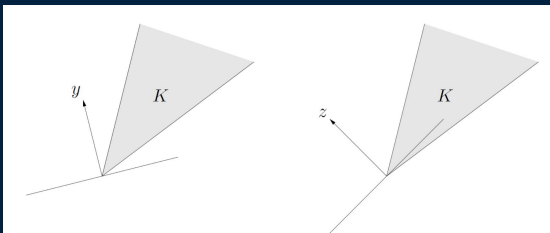


图: 集合 S_1 关于 \mathbb{R}^2 上的分量不等式有最小元 \mathbf{x}_1 。点 \mathbf{x}_2 是 S_2 的极小元

对偶锥

- 令 K 为一个锥，集合 $K^* = \{y | x^T y \geq 0, \forall x \in K\}$ 称为 K 的**对偶锥**
- 从几何上看， $y \in K^*$ 当且仅当 $-y$ 是 K 在原点上的一个支撑超平面的法线
- **非负象限**：锥 \mathbb{R}_+^n 的对偶是它本身 $y^T x \geq 0, \forall x \succeq 0 \iff y \succeq 0$



图：以 y 为内法向量的半空间包含锥 K ，因此 $y \in K^*$ 。以 z 为内法向量的半空间不包含锥 K ，因此 $z \notin K^*$

对偶锥性质

对偶锥有如下性质（课后作业证明）：

- 无论 K 是不是闭凸锥， K^* 是闭凸锥（由几何意义可见，两个支撑超平面的凸组合一定也是支撑超平面）
- $K_1 \subseteq K_2$ 可导出 $K_2^* \subseteq K_1^*$
- 如果 K 有非空内部，那么 K^* 是尖的
- 如果 K 的闭包是尖的，那么 K^* 有非空内部
- K^{**} 是 K 的对偶锥 K^* 的对偶锥，也是 K 的凸包的闭包
- 如果 K 是闭凸锥，那么 $K^* = K$
- 上述性质表明，如果 K 是一个正常锥，那么它的对偶 K^* 也是正常锥，且 $K^* = K$

广义不等式的对偶

- 正常锥 K 可以导出一个广义不等式 \preceq_K , 其对偶锥 K^* 也是正常的, 所以也能导出一个广义不等式 \preceq_{K^*} , 我们称 \preceq_{K^*} 是 \preceq_K 的对偶
- 广义不等式及其对偶有如下重要性质:
 - $\mathbf{x} \preceq_K \mathbf{y}$ 当且仅当对于任意 $\lambda \succeq_{K^*} 0$ 有 $\lambda^T \mathbf{x} \leq \lambda^T \mathbf{y}$
 - $\mathbf{x} \prec_K \mathbf{y}$ 当且仅当对于任意 $\lambda \succeq_{K^*} 0$ 和 $\lambda \neq 0$ 有 $\lambda^T \mathbf{x} < \lambda^T \mathbf{y}$
- 由于 $K^{**} = K$, 与 \preceq_{K^*} 相关的对偶广义不等式为 \preceq_K , 因此交换广义不等式及其对偶以后, 上述性质依然成立
- 例: $\lambda \preceq_{K^*} \mu$ 的充要条件是对于所有 $\mathbf{x} \preceq_K 0$ 有 $\lambda^T \mathbf{x} \leq \mu^T \mathbf{x}$

线性严格广义不等式的择一定理

- 设有正常锥 $K \subseteq \mathbb{R}^n$, 考虑严格广义不等式 \prec_K

$$A\mathbf{x} \prec_K \mathbf{b}, \mathbf{x} \in \mathbb{R}^n \quad (6)$$

不可行, 也就是仿射集合 $\{\mathbf{b} - A\mathbf{x} | \mathbf{x} \in \mathbb{R}^n\}$ 与开凸集 $\text{int } K$ 不相交

- 那么存在一个分离超平面, 即非零的 $\lambda \in \mathbb{R}^m$ 和 $\mu \in \mathbb{R}$ 使得对于任意 \mathbf{x} 有 $\lambda^T(\mathbf{b} - A\mathbf{x}) \leq \mu$, 对于任意的 $\mathbf{y} \in K$ 有 $\lambda^T\mathbf{y} \geq \mu$, 这种情况当且仅当 $\lambda \in K^*$ 和 $\mu \leq 0$ 才可能发生
- 综上可知当式 (6) 不可行时, 存在 λ 使得

$$\lambda \neq 0, \lambda \succeq_{K^*} 0, A^T\lambda = 0, \lambda^T\mathbf{b} \leq 0 \quad (7)$$

- 反方向证明, 假设式 (6) 与式 (7) 同时成立, 由 $\lambda \neq 0, \lambda \succeq_{K^*} 0, \mathbf{b} - A\mathbf{x} \succ_K 0$, 我们有 $\lambda^T(\mathbf{b} - A\mathbf{x}) > 0$ 。又因为 $A^T\lambda = 0$ 我们可以找到 $\lambda^T(\mathbf{b} - A\mathbf{x}) = \lambda^T\mathbf{b} \geq 0$, 这就与 $\lambda^T\mathbf{b} \leq 0$ 产生了矛盾
- 式 (4) 和式 (5) 是本择一定理在 $K = \mathbb{R}_+^m$ 上的特殊情况

最小元的对偶性质

- 我们可以用对偶广义不等式来刻画集合 $S \subseteq \mathbb{R}^m$ (可能非凸) 关于正常锥 K 导出的广义不等式的最小元和极小元
- \mathbf{x} 是 S 上关于广义不等式 \preceq_K 的最小元的充要条件是, 对于所有的 $\lambda \succ_{K^*} 0$, \mathbf{x} 是 $\mathbf{z} \in S$ 上极小化 $\lambda^T \mathbf{z}$ 的唯一最优解
- 几何上看, 这意味着对于任意的 $\lambda \succ_{K^*} 0$ 超平面 $\{\mathbf{z} | \lambda^T (\mathbf{z} - \mathbf{x}) = 0\}$ 是 \mathbf{x} 处对 S 的一个严格支撑超平面 (所谓严格就是与 S 只相交于 \mathbf{x})
- 上述结论对 S 是不是凸集无要求

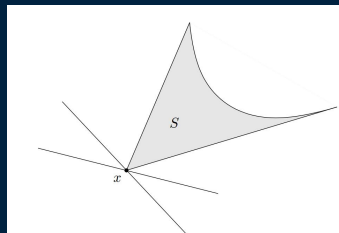


图: 最小元的对偶性质: 点 \mathbf{x} 是集合 S 中关于 \mathbb{R}_+^2 的最小元 \iff 对于任意的 $\lambda \succ_{K^*} 0$ 超平面 $\{\mathbf{z} | \lambda^T (\mathbf{z} - \mathbf{x}) = 0\}$ 在 \mathbf{x} 处对 S 的一个严格支撑, 即超平面规定的一个半空间包含了 S , 且只在 \mathbf{x} 处与 S 接触