凸优化理论注释(初稿)

命题1.3.10(补) Let C be a convex subset of $\mathbb{R}^{n+m}$. For $x \in \mathbb{R}^n$, denote

$$C_x = \{y \mid (x, y) \in C\},$$

and let

$$D = \{x \mid C_x \neq \emptyset\}.$$

Then

$$\mathrm{ri}(C) = \{(x, y) \mid x \in \mathrm{ri}(D), y \in \mathrm{ri}(C_x)\}.$$

证明: 由于D是C在x轴的投影, 根据命题1.3.6,

$$\mathrm{ri}(D) = \{x \mid \text{存在} y \in \mathbb{R}^m, (x, y) \in \mathrm{ri}(C)\},$$

因此

$$\mathrm{ri}(C) = \bigcup_{x \in \mathrm{ri}(D)} (M_x \cap \mathrm{ri}(C)),$$

其中 $M_x = \{(x, y) \mid y \in \mathbb{R}^m\}$. 其中右侧包含于左侧是明显的, 对于任意的 $(x, y) \in \mathrm{ri}(C))$,

$$y \in \mathrm{ri}(C_x) = \mathrm{ri}(\mathcal{B}C) = \mathcal{B}\mathrm{ri}(C),$$

其中 $\mathcal{B}$ 是 C 到 y 轴的投影, 于是 $(x, y) \in M_x \cap \mathrm{ri}(C)$, 即得

$$\mathrm{ri}(C) = \bigcup_{x \in \mathrm{ri}(D)} (M_x \cap \mathrm{ri}(C)).$$

不妨设 C 非空, 对每个点 $x \in \mathrm{ri}(D)$, 我们有

$$M_x \cap \mathrm{ri}(C) = \mathrm{ri}(M_x) \cap \mathrm{ri}(C) = \mathrm{ri}(M_x \cap C) = \mathrm{ri}(\{x\} \times C_x) = \{x\} \times \mathrm{ri}(C_x) = \{(x, y) \mid y \in \mathrm{ri}(C_x)\}.$$

定义 1.4.1 (补) $\dfrac{x_k}{||x_k||}$ 的任意极限点都是集合 C_k 的回收方向.

不妨设 $\frac{x_k}{||x_k||} \to d$, $||d|| = 1$. 固定 k_0, 取 $x_{k_0} \in C_{k_0}$, 任取 $\alpha \geqslant 0$, 设 $d_k = \frac{x_k - x_{k_0}}{||x_k - x_{k_0}||}$, 则

$$\lim_{k\to+\infty} \frac{x_k - x_{k_0}}{||x_k - x_{k_0}||} = \lim_{k\to+\infty} \frac{x_k}{||x_k||} \frac{||x_k||}{||x_k - x_{k_0}||} = \lim_{k\to+\infty} \frac{x_k}{||x_k||} = d.$$

由 $||x_k|| \to +\infty$, 存在 $K > k_0$, 当 $k > K$ 时, $||x_k - x_{k_0}|| > \alpha$, 则由

$$x_{k_0}, x_k \in C_{k_0}, \quad k \geqslant K$$

得

$$x_{k_0} + \alpha d_k \in C_{k_0}, \quad k \geqslant K.$$

由 C_k 是闭凸集,

$$x_{k_0} + \alpha d_k \to x_{k_0} + \alpha d \in C_{k_0},$$

由 k_0 和 α 的任意性, d 为 C 的回收方向.

命题 1.4.9 (补) 半空间是收缩的.

设半空间

$$S = \{x \in \mathbb{R}^n \mid a^T x \leqslant b, a \in \mathbb{R}^n, a \neq 0, b \in \mathbb{R}\}, \quad S_0 = \{x \in \mathbb{R}^n \mid a^T x \leqslant 0, a \in \mathbb{R}^n, a \neq 0\}$$

x_k 为 S 的渐进序列, 即 $x_k \in S, x_k \neq 0, ||x_k|| \to +\infty, \frac{x_k}{||x_k||} \to d, ||d|| = 1$,

(i) $a^T d = 0$.

则

$$a^T(x_k - d) \leqslant 0, \quad x_k - d \in S, \quad S \text{ 收缩}.$$

(ii) $a^T d < 0$.

存在 x_0 满足 $a^T x_0 = b$, 于是

$$x_k - x_0 \in S_0,$$

设

$$d_k = \frac{x_k - x_0}{||x_k - x_0||},$$

则

$$d_k \in S_0, \quad d_k \to d,$$

由 $||x_k|| \to +\infty$, 存在 K, 当 $k > K$ 时,

$$x_k - x_0 > 0, \quad ||x_k - x_0|| > 2, \quad a^T d_k < \frac{2}{3} a^T d,$$

于是

$$a^T(2d_k - d) < \frac{1}{3} a^T d < 0, \quad 2d_k - d \in S_0,$$

再由

$$2d_k \in S_0, \quad S_0 \text{ 为凸锥得}$$
$$2d_k - \frac{2}{||x_k - x_0||} d \in S_0,$$

于是

$$x_k - x_0 - d \in S_0, \quad x_k - d \in S \quad (k > K).$$

即S收缩.

定义 1.4.2 (补) 凸紧集和多面体锥的向量和是收缩的.

设集合

$$S = C + PC,$$

其中 C 为凸紧集, $PC = \bigcap_{i=1}^{n} PC_i$ 为多面体锥, 其中

$$PC_i = \{x \in \mathbb{R}^n \mid a_i^T x \leqslant 0, a_i \in \mathbb{R}^n\}.$$

先证 $S_i = C + PC_i$ 是收缩的. 设 $z_k = x_k + y_k$ 为 S_i 的渐进序列, $z_k \in S_i$, $z_k \neq 0$, $||z_k|| \to +\infty$, $\dfrac{z_k}{||z_k||} \to d$, $||d|| = 1$, 其中 $x_k \in C$, $y_k \in PC_i$. 由 C 是紧集, 则 x_k 有界, $||y_k|| \to +\infty$, $\dfrac{y_k}{||y_k||} \to d$, $||d|| = 1$, 且存在 K, 当 $k > K$ 时, $y_k \neq 0$. 于是 y_k 是 PC_i 的渐进序列, 由命题 1.4.9 (补): 半空间是收缩的, 可知存在 K, 当 $k > K$ 时, $y_k - d \in PC_i$, 于是 $x_k + y_k - d \in S_i$ $(k > K)$, S_i 是收缩的. 由 $S = \bigcap_{i=1}^{n} S_i$, S 是收缩的 (此处应默认 PC 存在渐进序列).

定理 1.5.7 (修正版) 令 C 与 P 为 n 维欧式空间 $\mathbb{R}^n$ 的非空子集且 P 为多面体, 则能分离 C 与 P 且不包含 C 的超平面存在, 当且仅当

$$\mathrm{ri}(C) \cap P = \emptyset.$$

首先, 我们回顾上文对真分离理论的讨论, 并观察到一普遍现象：对凸集 X 和超平面 H, 如果 X 完全包含于 H 的一侧, 则有

$$X \subset H \quad \text{当且仅当} \quad \mathrm{ri}(X) \cap H \neq \emptyset \tag{(1.33)}$$

(同式(1.31). 后续的证明将反复使用这一关系.

必要性： 假设存在这样的超平面 H, 使得 C 与 P 分离, 且不包含 C. 此时不妨设 $C \subset H_+ = H \cup H_{++}, P \subset H_-$, 根据式(1.33), H 不包含任何 C 的相对内点(因为H不包含C), 即 $\mathrm{ri}(C) \cap H = \emptyset$, 于是 $\mathrm{ri}(C) \subset H_{++}$, 又由于 $P \subset H_-$, $H_- \cap H_{++} = \emptyset$ 于是 $\mathrm{ri}(C) \cap P = \emptyset$.

其中

$$H_+ = \{x \mid a^T x \geqslant b, a \in \mathbb{R}^n, a \neq 0, b \in \mathbb{R}\},$$

$$H_{++} = \{x \mid a^T x > b, a \in \mathbb{R}^n, a \neq 0, b \in \mathbb{R}\}.$$

充分性：

考虑相反方向, 假设 $\mathrm{ri}(C) \cap P = \emptyset$, 我们欲证明满足条件的超平面存在. 令

$$D = P \cap \mathrm{aff}(C).$$

如果 $D = \emptyset$, 由于 $\mathrm{aff}(C)$ 与 P 都是多面体($\mathrm{aff}(C)$是多面体是因为 : 设与 $\mathrm{aff}(C)$ 平行的子空间为 L, $\mathrm{aff}(C) - x_0 = L$, 取 $L^\perp$ 的一个基 $\{a_1, \cdots, a_m\}$, 则 $L = \{x \mid a_i^T x = 0\}$, $\mathrm{aff}(C) = \{x \mid a_i^T (x - x_0) = 0\}$, 因此 $\mathrm{aff}(C)$ 是多面体), 应用严格分离定理[见命题1.5.3的条件(3)]易证出：存在超平面 H, 使得 $\mathrm{aff}(C)$ 与 P 严格分离,因此 H 不包含 C(通过严格分离的定义即可验证). 因此我们可以假设 $D \neq \emptyset$. 现在的思路是, 先构造 C 与 D 的真分离超平面, 然后再扩展这个超平面使其分离 C 和 P. [如果 C 有内点, 则有 $\mathrm{aff}(C) = \mathbb{R}^n$ 且 $D = P$, 这时证明将大大简化]. 根据假设 $\mathrm{ri}(C) \cap P = \emptyset$, 我们可以推出

$$\mathrm{ri}(C) \cap \mathrm{ri}(D) \subset \mathrm{ri}(C) \cap (P \cap \mathrm{aff}(C)) = (\mathrm{ri}(C) \cap P) \cap \mathrm{aff}(C) = \emptyset.$$

根据命题1.5.6, 存在超平面 H 真分离 C 和 D. 而且 H 不包含 C, 否则 H 将也包含 $\mathrm{aff}(C)$ 从而包含 D, 这与真分离的性质相矛盾. 因此, C 只被 H 的一个闭半空间包含, 而不可能被两个闭半空间都包含. 设 $\bar{C}$ 为 $\mathrm{aff}(C)$ 与该闭半空间的交, 即 $\bar{C} = \mathrm{aff}(C) \cap H_+$, 且不妨设 $C \subset H_+$, 其中 $H = \{x \mid a^T x = b, a \in \mathbb{R}^n, a \neq 0, b \in \mathbb{R}\}$, 因为 H 不包含 C且$C \subset \bar{C} = H_+ \cap \mathrm{aff(C)}$, 于是 H 不包含 $\bar{C}$.

根据式(1.33), 我们可以推出 $H\cap \mathrm{ri}(\bar{C})=\emptyset$. 由于 $D\subset H_-$, $C\subset H_+$,

$$\mathrm{ri}(\bar{C})=\mathrm{ri}(H_+)\cap \mathrm{ri}(\mathrm{aff}(C))=H_{++}\cap \mathrm{aff}(C).$$

所以 $D\cap \mathrm{ri}(\bar{C})=\emptyset$.

于是

$$P\cap \mathrm{ri}(\bar{C})=D\cap \mathrm{aff}(C)\cap \mathrm{ri}(\bar{C})=\emptyset.$$

如果 $P\cap\bar{C}=\emptyset$, 通过严格分离定理[命题(1.5.3)条件(3)], 存在超平面使得 P 和 $\bar{C}$ 严格分离. 由于 $C\subset H_+\cap \mathrm{aff}(C)=\bar{C}$, 于是该超平面也严格分离 P 和 C, 证明完毕. 因此假设

$$P\cap\bar{C}\neq\emptyset.$$

由平移不妨设:

$$0\in P\cap\bar{C}.$$

多面体 P 可以表示成形如 $\{x\mid a_j^Tx\leqslant b_j\}$ 的半空间的交, 其中对所有 j都有 $b_j\geqslant 0$（因为 $0\in P$）且对至少一个 j 值满足 $b_j=0$（这是因为 $0\in P\cap\bar{C}=P\cap H_+\cap \mathrm{aff}(C)$, 于是 $0\in D=P\cap \mathrm{aff}(C)\subset H_-$且$0\in H_+$, 那么 $0\in H$, 若 b_j 均大于 0, 那么 0 是 P 的相对内点（由延伸引理即证）, 于是 $0\in \mathrm{ri}(P)\cap \mathrm{aff}(C)=\mathrm{ri}(D)$, 任取 $y\in \mathrm{ri(C)}\subset \mathrm{aff(C)}$, 存在 $\gamma>0$, 使得 $\gamma y\in P$, 那么 $\gamma y\in P\cap \mathrm{aff}(C)=D\subset H_-$ 于是 $\gamma a^Ty\leqslant 0$, 但因 $y\in \mathrm{ri}(C)\subset H_{++}$, 有 $a^Ty>0$ 矛盾）.

因此, 存在整数 $m\geqslant 1$ 及 $\bar{m}\geqslant m$ 和标量 $b_j>0$, 使得

$$P=\{x\mid a_j^Tx\leqslant 0, j=1,\ldots,m\}\cap\{x\mid a_j^Tx\leqslant b_j, j=m+1,\ldots,\bar{m}\}.$$

令 M 为 $\bar{C}$ 的相对边界, 即 $M=\mathrm{cl}(\bar{C})\setminus \mathrm{ri}(\bar{C})=(H_+\cap \mathrm{aff}(C))\setminus(H_{++}\cap \mathrm{aff}(C))=H\cap \mathrm{aff}(C)$.

考虑锥体:

$$K=\{x\mid a_j^Tx\leqslant 0, j=1,\ldots,m\}+M.$$

注意到 $K=\mathrm{cone}(P)+M$ (这是因为: i. 任取 $x\in P$, $\lambda\geqslant 0$, $\lambda x\in\{x\mid a_j^Tx\leqslant 0, j=1,\ldots,m\}$, 于是 $\mathrm{cone}(P)\subset\{x\mid a_j^Tx\leqslant 0, j=1,\ldots,m\}$. ii. 任取 $x\in\{x\mid a_j^Tx\leqslant 0, j=1,\ldots,m\}$, 那么存在 $\lambda\geqslant 0$, 使得 $\dfrac{x}{\lambda}\in\{x\mid a_j^Tx\leqslant b_j, j=m+1,\ldots,\bar{m}\}$, 于是 $x\in \mathrm{cone}(P)$, 即$\{x\mid a_j^Tx\leqslant 0, j=1,\ldots,m\}\subset \mathrm{cone}(P)$).

断言 $K\cap \mathrm{ri}(\bar{C})=\emptyset$, 下面通过反证法证明：若存在 $\bar{x}\in K\cap \mathrm{ri}(\bar{C})$, 则存在 $\alpha>0$, $\omega\in P$, $v\in M$, 使得 $\bar{x}=\alpha\omega+v$（因为 $K=\mathrm{cone}(P)+M$, 且 $0\in P$），因此 $\dfrac{\bar{x}}{\alpha}-\dfrac{v}{\alpha}\in P$. 另一方面, 因为 $\bar{x}\in \mathrm{ri}(\bar{C})=H_{++}\cap \mathrm{aff}(C)$, 且 $v\in M=H\cap \mathrm{aff}(C)$ 故而 $\dfrac{\bar{x}}{\alpha}-\dfrac{v}{\alpha}\in \mathrm{ri}(\bar{C})$. 这与 $P\cap \mathrm{ri}(C)=\emptyset$ 矛盾, 因此 $K\cap \mathrm{ri}(\bar{C})=\emptyset$.

锥体 K 是多面体（因为它是两个多面体的和（见命题2.3.4, 中间不涉及该命题及其推论的使用, 即不涉及循环论证））, 所以它是某些过原点的某些闭半空间的交（设 $K=\cap_{i=1}^{r}F_i'$, 那么 $K\subset F'=\{x\mid c^Tx\geqslant d\}$, F' 是 $F_i', i=1,\dots,r$ 中的某个, 任取 $\lambda>0, x\in K$, 由 K 是锥, 得 $\lambda c^Tx\geqslant d$, 将 λ 除到等式右侧取极限得 $c^Tx\geqslant 0$, 令 $F=\{x\mid c^Tx\geqslant 0\}$, 则 $x\in F$, 即 $K\subset F$, F 过原点, 对每个 F_i' 都如此操作得 $K\subset\cap_{i=1}^{r}F_i$）.

因为 $K=\mathrm{cone}(P)+M, 0\in\mathrm{cone}(P)$, 这些闭半空间都包含 M, 即也都包含 $\bar{C}$ 的相对边界, 进而可以推出如果有某闭半空间包含 $\mathrm{ri}(\bar{C})$ 中的向量, 则其必然包含整个 $\bar{C}$（这是因为：任取 $y\in\bar{C}$, 设 $y_r\in\mathrm{ri}(C)\subset F$, F 是 F_i' 中的某个, 由 $y\in\bar{C}=H_+\cap\mathrm{aff}(C)$, $y_r\in\mathrm{ri}(C)=H_{++}\cap\mathrm{aff}(C)$ 得 $a^Ty\geqslant 0$, $a^Ty_r>0$, i. $a^Ty=0$, 则 $y\in H$, 那么 $y\in M=H\cap\mathrm{aff}(C)\subset F$, ii. $a^Ty>0$, 则 $y\in\mathrm{ri}(\bar{C})=H_{++}\cap\mathrm{aff}(C)$, 设 $z=y+\lambda y_r, \lambda\in\mathbb{R}$, 可知存在 $\lambda<0$, 使得 $a^Tz=0$, 即 $z\in H$, 又因为$0\in\bar{C}\subset\mathrm{aff}(C)$, 则$\mathrm{aff}(C)$为子空间, $z\in\mathrm{aff}(C)$, 那么 $z\in M=H\cap\mathrm{aff}(C)$, 又因为 $y_r\in F$, F 是过原点的闭半空间, 于是 $y=z-\lambda y_r\in F(z\in M\subset F, y_r\in F, -\lambda>0$, 写出$F$的表达式即可验证), 于是 $\bar{C}\subset F$）.

因此, 由于 $K\cap\mathrm{ri}(\bar{C})=\emptyset$, 则 $F_1,\dots,F_r$ 中至少有一个集合与 $\mathrm{ri}(\bar{C})$ 不交(若均与$\mathrm{ri}(\bar{C})$相交, 则均包含$\mathrm{ri}(\bar{C})$, 而K为这些闭半空间的交, 与$K\cap\mathrm{ri}(\bar{C})=\emptyset$矛盾), 不妨假设该集合即为 F_1, 于是 K 和 $\mathrm{ri}(C)$ 在 F_1 的两侧, 所以对应于 F_1 的超平面不包含 $\bar{C}$ 的任何相对内点, 所以也不包含 C 的相对内点, 因此该超平面既不包含 C, 又能分离 K 和 C. 又由于 K 包含了 P, 于是这个超平面也能分离 P 和 C.

原版p89, 翻译版p85, 1.7小结上面最后一段

However, unless C is a cone, the support function of C will not be an indicator function, nor will the relation $(C^*)^*=C$ hold.

反例: 在 $\mathbb{R}^2$ 中, 设 $C=(\mathbb{Z},0)$, $\mathbb{Z}$ 是整数集, C 的支撑函数是示性函数, 但 C 不是锥.

命题 2.1.4 (补) 令 $\mathbb{P}$ 为 n 维欧式空间 $\mathbb{R}^n$ 的多面体子集。

(b补) 如果 P 形如

$$P=\{x \mid Ax=b, x\geqslant 0\},$$

其中 A 是 $m\times n$ 矩阵, b 是 $\mathbb{R}^m$ 中的向量, 那么向量 $v\in P$ 是 P 的顶点当且仅当 A 对应于 v 的非零坐标的列是线性无关的（这里的对应关系是指：v 的第 i 个分量 v_i 对应 A 的第 i 个列向量 a_i）。

我们把 P 写成等价的不等式的形式

$$P=\{x \mid Ax\leqslant b, -Ax\leqslant -b, -e_ix\leqslant 0, i=1,\cdots,n\}.$$

令 S_v 为形如 $(0,\cdots,0,-1,0,\cdots,0)$ 的向量集合, 其中 -1 出现在对应 v 的某个零坐标位置上, 即

$$S_v=\{x \mid x=e_i, v_i=0\}.$$

把 (a) 部分的结果应用到上述 P 的不等式描述中, 我们看到 v 是顶点当且仅当 A 的行和 S_v 中的向量一起包含 n 个线性无关的向量。令 $\bar{A}$ 为除了把对应于 v 的零坐标的列设为 0 其他部分与 A 相同的矩阵。可知 v 是顶点当且仅当$\bar{A}$ 包含 $n-k$ 个线性无关行, 其中 k 是 S_v 中向量的个数（这是因为：将 A 的行向量和 S_v 中的向量拼接在一个矩阵中, 不妨记为

$$\begin{pmatrix} A \\ S_v \end{pmatrix},$$

该矩阵的秩是 n, 对其作初等行变换得

$$\begin{pmatrix} \bar{A} \\ S_v \end{pmatrix},$$

即得 $\bar{A}$ 包含 $n-k$ 个线性无关行, 由于该行变换是可逆的, 于是 v 是顶点当且仅当$\bar{A}$ 包含 $n-k$ 个线性无关行. 由于

$$\begin{pmatrix} \bar{A} \\ S_v \end{pmatrix}$$

的秩为 n, A 的非零列只可能是对应于 v 的非零坐标的那 $n-k$ 列, 这些列也是线性无关的, 由于这些列为 A 和 $\bar{A}$ 所共用, 结论成立。

命题 2.3.2 (Minlow-Weyl定理) (补). 一个锥体是多面体当且仅当它是有限生成的. 在 $\{a_1,\cdots,a_r\}$ 不包含 n 个线性无关向量的情形, 令 S 为由

$a_1, \cdots, a_r$ 张成的子空间. 定义有限生成锥

$$S^{\perp} + \mathrm{cone}(\{a_1, \cdots, a_r\}).$$

该集合为有限生成锥是因为：设 $\{b_1, \ldots, b_m\}$ 为 $S^{\perp}$ 的一个基, 于是 $S^{\perp} = \mathrm{cone}\{\pm b_1, \ldots, \pm b_m\}$, 那么 $S^{\perp} + \mathrm{cone}(\{a_1, \cdots, a_r\}) = \mathrm{cone}\{\pm b_1, \cdots, \pm b_m\} + \mathrm{cone}(\{a_1, \cdots, a_r\}) = \mathrm{cone}\{\pm b_1, \cdots, \pm b_m, a_1, \cdots, a_r\}$..

最后一段: 可知 $\mathrm{cone}(\{a_1, \cdots, a_r\})$ 是对应于由 $\{a_1, \cdots, a_r\}$ 中 $(n-1)$ 个线性无关向量张成的超平面子空间的交的某些半空间的交:

可反设后者存在一个向量不属于 $\mathrm{cone}(\{a_1, \cdots, a_r\})$, 对这个向量和 $\mathrm{cone}(\{a_1, \cdots, a_r\})$ 应用严格分离定理即得矛盾.

例 2.1.5(补) 这里改下一下原命题的形式: 将X中的列向量拼接称一个$\mathbb{R}^{n^2}$维的列向量x, 记

$$Q = \{x \in \mathbb{R}^{n^2} \mid \begin{pmatrix} U \\ V \end{pmatrix} x = \mathbf{1}, x \geqslant 0, \}.$$

其中U 中的行向量形如: $u = (\cdots, 0, \cdots, \alpha, \cdots, 0, \cdots)$, 其中$\alpha$为元素全为1的$n$维向量. V 中的行向量形如: $v = (e_i, \cdots, e_i)$, 其中e_i为单位向量, 这是根据原命题中P的结构构造的, 易见原命题中X为P顶点等价于x为Q顶点.

类似置换矩阵的概念, 引入针对该命题的置换向量: 设$x \in \mathbb{R}^{n^2} = R^n \times \cdots \times R^n$, 若$x$的每个$R^n$分量只有一个1其他元素为0则称为置换向量.

必要性： 设x为Q的顶点, 由Q的结构可知, 若x的每个分量都包含两个或两个以上非零元, 那么x零元素个数不多于$n(n-2)$ (可参考原文). 另一方面,

$$\sum_{j=1}^{n} u_j = \sum_{j=1}^{n} v_j,$$

于是$\{u_j\} \cup \{v_j\}, j = 1, \cdots, n$最多有$2n-1$个线性无关, 因此, 由命题2.1.4(b), x中最多有$2n-1$个非零元, 等价的x至少有$n(n-2)$个零元. 矛盾. 于是必存在一个分量$\bar{j}$, 使得该分量只有一个非零元. 且该元为1, 由$Vx = 1$该元在分量$\bar{j}$所在的位置(角标, 记为$\bar{i}$)对应其它分量所在相同位置($\bar{i}$)的元素也为0, 设超平面H法向量为第$\bar{j}$个分量的$\bar{i}$位置为1, 其它的元素为0, 等式右侧常数取1, 由命题2.1.1x是$Q \cap H$的顶点, 再去掉x和$Q \cap H$的第$\bar{j}$个分量和每个分量对应的第$\bar{i}$个元素, 得到新的x_{new}和$Q \cap H_{new}$. 如此继续, 通过每一步找到新包含单一非零元的分量, 可证x每个分量都具有非零元. 这就证明了x是置换向量.

命题 2.3.5 f是凸的条件可以去掉.

命题 3.3.4 令 $F : \mathbb{R}^{n+m} \mapsto (-\infty, \infty]$ 为闭真凸函数. 考虑由

$$f(x) = \inf_{z \in \mathbb{R}^m} F(x, z), \quad x \in \mathbb{R}^n$$

给定的函数 f, 假定对于某个 $\bar{x} \in \mathbb{R}^n$ 和 $\bar{\gamma} \in \mathbb{R}$, 集合

$$\{z \mid F(\bar{x}, z) \leqslant \bar{\gamma}\}$$

为非空, 且它的回收锥等于它的线形空间. 那么 f 是闭且真凸的. 进而, 对每个 $x \in \mathrm{dom}(f)$, $f(x)$ 定义中的最小点集是非空的.

由命题 3.1.1(a), f 是凸的, 由命题 1.4.5(a) f 的回收锥形如

$$R_F = \{(d_x, d_z) \mid (d_x, d_z, 0) \in R_{\mathrm{epi}(F)}\}.$$

由 F 是闭凸函数和命题 1.4.1 得 $F(\bar{x}, \cdot)$ 的非空水平集的公共回收锥形如

$$\{d_z \mid (0, d_z) \in R_F\},$$

任取 $d = (0, d_z, 0) \in R_F \cap N(P)$, 于是 $d_z \in R_{F(\bar{x},\cdot)} = L_{F(\bar{x},\cdot)}$, 由 $L_{F(\bar{x},\cdot)}$ 的定义和命题 1.4.1 得, $d = (0, d_z, 0) \in L_F$, 由命题 1.4.13, 其中取 $X = \mathbb{R}^n$ 得 $P(\mathrm{epi}(F))$ 为闭集, 由命题 3.3.1, $\mathrm{epi}(f)$ 是闭集, 即 f 是闭函数.

对任意 $x \in \mathrm{dom}(f)$, 由 F 是闭凸函数和命题 1.4.1 得, 非空水平集 $F(x, \cdot)$ 有公共回收锥, 那么也有公共线性空间, 由题设条件 $R_{F(\bar{x},\cdot)} = L_{F(\bar{x},\cdot)}$, 有 $R_{F(x,\cdot)} = L_{F(x,\cdot)}$, 再由命题 3.2.4, $F(x, \cdot)$ 最小点集非空, 进而由 F 是真函数, 任取 $x \in \mathrm{dom}(f)$, $f(x) > \infty$, 由 $\bar{x} \in \mathrm{dom}(f)$, f 真.

例 4.2.2少条件, 考虑$f = sinx$其凹闭包的计算公式明显是凹函数, 然而ϕ并不是凹的.可以加上$z'f$凹函数的条件.

命题5.1.2(b补) (线性Farkas引理). 令 A 为 $m \times n$ 矩阵而 c 为 m 维欧式空间 $\mathbb{R}^n$ 中的向量.

(b) 不等式组 $Ay \geqslant c$ 有解当且仅当

$$A'x = 0, \quad x \geqslant 0 \qquad \Rightarrow \qquad c'x \leqslant 0.$$

充分性: 设集合 $P = \{(A'x, -c'x) \mid x \geqslant 0\}$, $S = \{(0, s) \mid s \leqslant 0\}$, P 可写作:

$$A'\mathbb{R}^m_+ \times (-c)'\mathbb{R}^m_+,$$

因此是多面体, 由假设可知 $\mathrm{ri}(S) \cap P = \emptyset$, 对 P 和 S 应用命题 1.5.7(多面体的真分离定理), 可知存在不包含 S 的超平面 $H : (y, \omega)'(u, v) = b, y, u \in$

$\mathbb{R}^n, \omega, v \in \mathbb{R}$ 使得:

$$(y,\omega)'(A'x,-c'x) \geqslant b \geqslant (y,\omega)'(0,s), \qquad x \geqslant 0, \quad s \leqslant 0,$$

由 $S \cap P = 0$得$b=0$, 又由于 H 不包含 S, 那么 $\omega \neq 0$, 取 $x=0$, 得到 $0 \geqslant \omega s$, 由于 $s \leqslant 0$, 则 $\omega > 0$, 于是不妨设 $\omega = 1$ 有

$$y'A'x - c'x \geqslant s, \qquad x \geqslant 0, \quad s \leqslant 0,$$

令 $s \to 0$ 得 $(y'A'-c')x \geqslant 0, x \geqslant 0$, 于是 $y'A' \geqslant c'$, 等价于 $Ay \geqslant c$.

命题 5.2.1 (a补) 如果 q^* 有限, 仍根据命题 1.4.12, 对偶问题有最优解 μ^*. 设矩阵 $A=(a_1,\ldots,a_r)'$. 于是对于任意满足 $A'y=0, y \geqslant 0$ 的 y, 都有 $b'y \leqslant 0$, 否则与 $b'\mu$ 取最大矛盾. 由命题 5.1.2 (b), 不等式 $A'x \geqslant c$ 有解, 于是 f^* 有限. 前面已证明 f^* 有限的情况 $f^* = q^*$.

p 169 中间段 与命题 5.3.7有关的一个结果是...

其中的: ($L(x,\mu^*) \leqslant F(x,0)$ 对所有 $x \in X$ 成立, 这意味着 $F(\cdot,0)$ 具有紧的水平集), 这是因为: F 是闭凸真函数, 若 F 存在无界水平集, 由于 F 是闭凸真函数, 根据命题 1.4.5, F 回收锥 R_F 存在非零回收方向 d, 根据 $L(x,\mu^*) \leqslant F(x,0)$, d 也是 $L(\cdot,\mu^*)$ 回收锥 $R_{L(\cdot,\mu^*)}$ 的回收方向, 这与 x^* 是唯一最小解矛盾.

翻译版 p171 下面的讨论

(1) 如果 q^* 有限, 并且 $\mathrm{ri}(\mathrm{dom}(f_1^\star)) \cap (A' \cdot \mathrm{ri}(-\mathrm{dom}(f_2^\star))) \neq \emptyset$, 那么 $f^* = g^*$ 并且对偶问题至少有一个最优解.

证明: 由于 $q(\lambda) = \inf\limits_{x_1 \in \mathbb{R}^n}\{f_1(x_1) - \lambda' A x_1\} + \inf\limits_{x_2 \in \mathbb{R}^n}\{f_2(x_2) + \lambda' x_2\} = -\sup\limits_{x_1 \in \mathbb{R}^n}\{\lambda' A x_1 - f_1(x_1)\} - \sup\limits_{x_2 \in \mathbb{R}^n}\{-\lambda' x_2 - f_2(x_2)\} = -f_1^\star(A'\lambda) - f_2^\star(-\lambda)$, 那么 $f^{\star\star} = \inf\limits_{\lambda \in \mathbb{R}^n}\{f_1^\star(A'\lambda) + f_2^\star(-\lambda)\} = \inf\limits_{\lambda \in \mathbb{R}^n}(-q(\lambda)) = \sup\limits_{\lambda \in \mathbb{R}^n} q(\lambda) = -q^*$, 于是根据命题 5.3.8, $f^{\star\star} = q^{\star\star}$, 并且可以计算出 $q^\star(\mu) = -\inf\limits_{u \in \mathbb{R}^n}(f_1^{\star\star}(Au) + f_2^{\star\star}(Au) = -f^*$.

(2) 更一般地, 如果 f_1 的形式为

$$f_1(x) = \begin{cases} \tilde{f}(x), & x \in X \\ \infty, & x \notin X \end{cases}$$

其中 X 是多面体集而 $\tilde{f}: \mathbb{R}^n \mapsto \mathbb{R}$ 是凸集 C 上的函数并且满足 $X \subset \mathrm{ri}(C)$, 如果 $(A \cdot \mathrm{ri}(\mathrm{dom}(f_1))) \cap \mathrm{dom}(f_2) \neq \emptyset$,

证明: 考虑优化问题:

$$\begin{aligned} \text{minimize} \quad & f_1(x) + f_2(Ax), \\ \text{subject to} \quad & x \in \mathrm{dom}(f_1) \cap A^{-}\mathrm{dom}(f_2) \end{aligned}$$

应用命题 5.3.6 即证.

(3) 如果 f_1 和 f_2 都是多面体的, 那么相对内点的条件是不必要的. 这种情况下, 问题 (5.17) 等价于一个线性规划 (此时只需将目标函数中的 $f_1(x_1)$ 和 $f_2(x_2)$ 分别换元即可).

图 5.4.1 "相应的MC/MC最大相交点问题的解集恰为次微分$\partial f(x)$"应为$-\partial f(x)$

命题 5.4.6 (补) (该命题的证明暂不确定是否存在问题) 存在满足$1 \geqslant \bar{m} \geqslant m$的$\bar{m}$, 使得每个$i = 1, \cdots, \bar{m}$对应的函数$f_i$是多面体函数且

$$(\cap_{i=1}^{\bar{m}} \mathrm{dom}(f_i)) \cap (\cap_{i=\bar{m}+1}^{m} \mathrm{ri}(\mathrm{dom}(f_i))) \neq \emptyset.$$

则有

$$\partial F(x) = \partial f_1(x) + \cdots + \partial f_m(x), \qquad \forall x \in (\cap_{i=1}^{\bar{m}} \mathrm{dom}(f_i)) \cap (\cap_{i=\bar{m}+1}^{m} \mathrm{ri}(\mathrm{dom}(f_i))).$$

注: 原文中的$x \in \mathbb{R}^n$是错的.考虑$f_1(x) = 0, f_2(x) = 1 - \sqrt{x}$.

由多面体函数定义, 可证其有效域内次梯度非空, 设$f(x_1, \cdots, x_m) = f_1(x_1) + \cdots + f_n(x_m)$, 注意到$\mathrm{dom} f(x_1, \cdots, x_m) = \mathrm{dom} f_1(x_1) \times \cdots \times \mathrm{dom} f_m(x_m)$, $\forall (x_1, \cdots, x_m) \in (\cap_{i=1}^{\bar{m}} \mathrm{dom}(f_i)) \cap (\cap_{i=\bar{m}+1}^{m} \mathrm{ri}(\mathrm{dom}(f_i)))$, 设$d \in \partial f(x_1, \cdots, x_m)$, $d = (d_1, \cdots, d_n) \in \mathbb{R}^{mn}$, 于是

$$f(z_1, \cdots, z_m) \geqslant f(x_1, \cdots, x_m) + \sum_{i=1}^{m} d^i(z_i - x_i), \qquad \forall z_i \in \mathrm{dom}(f_i)$$

也即

$$\sum_{i=1}^{m} (f_i(z_i) - z_i' d_i) \geqslant \sum_{i=1}^{m} (f_i(x_i) - x_i' d_i), \qquad z_i \in \mathrm{dom}(f_i).$$

因此$(x_1, \cdots, x_m)$是如下优化问题的最优解:

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^{m} (f_i(y_i) - z_i' d_i), \\ \text{subject to} \quad & y_i \in \mathrm{dom}(f_i), \quad z_i = y_i, i = 1, \cdots, m. \end{aligned}$$

根据命题 5.3.6, 强对偶关系成立且对偶最优解存在, 且存在对偶最优解(此处将(y_i,z_i)的笛卡尔积看作变量, 并将定义域写作笛卡尔积的形式)$(\lambda_1,\cdots,\lambda_m)$使得

$$(x_1,x_1,x_2,x_2,\cdots,x_m,x_m)\in \arg\min_{y_i\in\mathbb{R}^n,z_i\in\mathbb{R}^n}\sum_{i=1}^{m}(f(y_i)-{z_i}'d_i+{\lambda_i}'(z_i-y_i)),$$

由于上式对变量z_i无约束, 则有$d_i=\lambda_i$, 于是

$$(x_1,\cdots,x_m)\in arg\min_{y_i\in\mathbb{R}^n}\sum_{i=1}m(f(y_i)-{\lambda_i}'y_i),$$

或者

$$f_i(z_i)\geqslant f_i(x_i)+{\lambda_i}'(z_i-x_i),\qquad \forall z_i\in\mathbb{R}^n.$$

因此$\lambda_i\in\partial f(x_i)$, 所以

$$\partial f(x_1,\cdots,x_m)=\partial f_1(x_1)+\cdots+\partial f_m(x_m),\qquad \forall x_i\in \mathrm{ri}(\mathrm{dom}(f_i))),$$

取$x_i=x$即得结论.

命题 5.6.2(补) 翻译版p203下面:

$$q(\mu)=\inf_{Ax-b\geqslant u,Bx\geqslant c}\mu' u=\sup_{A'u+B'\nu=0,\nu\geqslant 0}(-b'\mu-c'\nu).$$

这是因为: 当 $A'\mu+B'\nu=0,\quad \nu\geqslant 0$ 时, $-b'\mu-c'\nu$ 是定值.

附录前最后一段 没懂, 感觉第一句不对.

翻译问题 命题1.1.8. 命题1.3.10. 命题1.4.11. 命题1.5.4. 命题1.6.1(d). 命题2.1.4(b)的证明. 命题3.2.2.