

□問題の概要

2Aで習う最適化手法に関する問題。凸多面体や凸錐に関する理論が背景にあり、ファルカスの補題を示すための典型的な手法である。筆者はちょうど輪講でその内容をやったため比較的楽に解けたが、授業資料を証明まで覚えていないと難しいかもしれない。輪講で使った資料は末尾に載せておくので、参考にしてほしい。Chapter1が関連する。

問題数が多いので、証明の流れを素早く考え付くのが大事だろう。

(1-1)

\inf の定義より、 $\forall \varepsilon > 0$ に対して、 $\|x - z_\varepsilon\| \leq \inf_{z \in K} \|x - z\| + \varepsilon$ なる $z_\varepsilon \in K$ が存在する。よって、 $\lim_{\varepsilon \rightarrow +0} z_\varepsilon = y$ は K が閉集合なので $y \in K$ で、 $\|x - y\| \leq \inf_{z \in K} \|x - z\|$ より、 $\|x - y\| = \inf_{z \in K} \|x - z\|$ を満たす。

(1-2)

ある K 内の相異なる 2 点 y_1, y_2 がどちらも \inf を達成しているとする。すると、 $y = \lambda y_1 + (1 - \lambda)y_2$, $\lambda \in (0, 1)$ とおけば $y \in K$ であり、

$$\begin{aligned} \|x - y\| &= \|(\lambda + 1 - \lambda)x - \lambda y_1 - (1 - \lambda)y_2\| \\ &\leq \lambda \|x - y_1\| + (1 - \lambda) \|x - y_2\| \\ &= \lambda \inf_{z \in K} \|x - z\| + (1 - \lambda) \inf_{z \in K} \|x - z\| \\ &= \inf_{z \in K} \|x - z\| \end{aligned}$$

2 行目で、三角不等式 $\|x + y\| \leq \|x\| + \|y\|$ が等号で成立する条件は、 x と y が幾何ベクトルの意味で同じ向きを向いている、つまり $x = ky$, $k \in \mathbb{R}_+$ のときのみである。つまり、 $\lambda(x - y_1) = k(1 - \lambda)(x - y_2)$, $k \in \mathbb{R}_+$ となるときだけ。このとき、 $\|x - y_1\| = \|x - y_2\| = k \frac{1 - \lambda}{\lambda} \|x - y_1\|$ はある $\lambda = \lambda^*$ で成り立たない。

よってある $\lambda^* \in (0, 1)$ で $\|x - y\| < \inf_{z \in K} \|x - z\|$ となるが、これは \inf の定義に矛盾。よって、 \inf を達成する点はただ一つである。

感覚としては、凸なら相異なる最適解 2 つを直線で結んだ部分は全部最適解だが、一方で $\|x - y\|$ は x から円状に同じ値が広がるので、直線と円はいつか離れる場所があるという感じ。

(1-3)

ある超平面 c によって凸集合を分離したい。分離超平面の法線を凸集合に外向きにとって、分離超平面と凸集合が接している部分の値を d にすれば良いというイメージを持つ。

つまり、 $y = \inf_{z \in K} \|x - y\|$ とすると、 $c = x - y$ とすれば凸集合に外向きの法線ベクトルが得られる。 d は $c^T y = x^T y - \|y\|^2$ である。さて、これが実際分離超平面になっているかを示す。

$c^T x = \|x\|^2 - x^T y > d$ は容易に示せる。 $c^T z \leq d = c^T y$ ($\forall z \in K$) を示すのは少し難しい。 $c^T(z - y) = (x - y)^T(z - y) \leq 0$ を言えばよいのだが、これは凸射影であることから従う。もう少し詳しい説明は最適化の資料を参照のこと。そこまで試験で求められているかという疑問なので、この程度の説明で終える。

(2-1)

非負結合のことを錐結合と呼ぶこともある。実際、 $C(\mathcal{A})$ は凸錐 (cone) を成す。

まず、 $C(\mathcal{A}) \cap C(\mathcal{B})$ は自明である。逆の $C(\mathcal{A}) \subset C(\mathcal{B})$ を言う。 $x \in C(\mathcal{A})$ とすると、ある m 個の a_i の組で

$$x = \sum_{i=1}^m \lambda_i a_i, \lambda \geq 0 \quad (1)$$

と書ける。ここで、そのような $\{a_i\}$ が線形独立ならば $x \in C(\mathcal{B})$ となる。線形従属ならば、ある $c_i \in \mathbb{R}$ が存在して、

$$\sum_{i=1}^m c_i a_i = 0 \quad (2)$$

である。さて、以上2つの式を組み合わせることで、 x を a_i の $m-1$ 個の非負結合で表すことができることを示していく。具体的には、 $r = \min_{c_i \neq 0} \left\{ \frac{\lambda_i}{c_i} \right\}$ として、式1から r 倍した式2を引くと、

$$x = \sum_{i=1}^m (\lambda_i - r c_i) a_i$$

が得られるが、 r の選び方より、ある i で少なくとも1つは $\lambda_i - r c_i = 0$ となる。さらに、 r の選び方より全ての i で $\lambda_i - r c_i \geq \lambda_i - \frac{\lambda_i}{c_i} c_i = 0$ となることがわかる。こうして x を $m-1$ 個の非負結合で表すことができたので、これを繰り返すことでいつか x は線形独立な a_1, \dots, a_m の部分集合の非負結合で表せるようになる。よって、 $x \in C(\mathcal{B})$ である。

この非負性を保ちつつギリギリまで減らすという操作は単体法を思い出させますね。

(2-2)

\mathcal{A} が有限集合なら、 $\{B \subset \mathcal{A} \mid B \text{ は線形独立}\}$ は有限集合になる。より詳しく、元の数 $2^{|\mathcal{A}|}$ で上から抑えられる。 $C(\mathcal{B})$ は閉集合なので、閉集合の有限和集合である $C(\mathcal{A})$ は閉集合。

一応、 \mathcal{B} が閉集合であることも示す。これは有限生成錐が多面体であり、多面体は閉集合であることから従う (Weyl-Minkowski Theorem)。とするとオーバーパワーなので、もう少し優しい証明の方針を与える。

B を生成するベクトルは線形独立なので、非負ベクトル x から Bx への線形写像は単射である。線形写像は連続で、写像の値域を $C(\mathcal{B})$ にとると、 $x \mapsto Bx$, 非負ベクトル $\rightarrow C(\mathcal{B})$ は連続な全単射写像になる。非負ベクトル全体は閉集合なので、 $C(\mathcal{B})$ は閉集合。多分これが一番簡潔な証明だと思います。

他の証明方針としては、 \mathcal{B} のサイズによる帰納法で示すものがある。 $n=1$ のとき $C(\mathcal{B})$ は直線で、閉集合。 $n=k$ で成り立つとすると、 $n=k+1$ で $C(\mathcal{B})$ 内の収束する点列をとる。収束先が $C(\mathcal{B})$ に入っていることを言いたい。 a_1, \dots, a_{k+1} で展開した係数で考えると、係数ごとに収束先を考えて言える。詳しくは [こちら](#) を参照。

(3-1)

$c^T A \lambda = c^T x > 0$ だが、 $c^T A \leq 0$ かつ $\lambda \geq 0$ なので矛盾。同時には成り立たない。

(3-2)

(P) が成り立たないとする。このとき (Q) が成り立つことを示す。

$C(A) = \{A\lambda \mid \lambda \geq 0\} \not\ni x$ である。ここで、 $C(A)$ は凸集合であることは簡単に示せる。さらに、(2-2)より、 $C(A)$ は閉集合である(とくに非空)。よって(1-3)よりある c, d が存在して、 $c^T x > d, c^T z \leq d (z \in C(A))$ を満たす。あとは、 $d = 0$ を言えばよい。まず、 $0 \in C(A)$ で、 $c^T 0 = 0$ なので、そのような d は非負である。ある $z \in C(A)$ で $c^T z > 0$ とする。 $\lambda z \in C(A) (\lambda > 0)$ なので、 $\lambda \rightarrow \infty$ とすれば、 $c^T(\lambda z) = \lambda c^T z \rightarrow \infty$ となり、上から d で抑えられることに矛盾。よって全ての $z \in C(A)$ で $c^T z \leq 0$ である。よって $d = 0$ としてよい。

参考文献

- [1] J. A. De Loera, R. Hemmecke, and M. Köppe, *Algebraic and geometric ideas in the theory of discrete optimization*. SIAM, 2012.