Linear & Nonlinear Programming

Deng Chang

October 25, 2018

Exercise 2.6 (Caratheodory's theorem)

(a) $\forall y \in C$, let us consider $\Lambda = \left\{ (\lambda_1, ..., \lambda_n) \in \Re^n | \sum_{i=1}^{n} \lambda_i \mathbf{A}_i = y, \lambda_1, ..., \lambda_n \geq 0 \right\}$, and mark $A = (A_1, A_2, ..., A_n), \lambda = (\lambda_1, ..., \lambda_n)^{\mathrm{T}}$. Then $\Lambda = \{\lambda \in \Re^n \, | A\lambda = y, \lambda \geq 0\}$. $\because$ we require $y \in C$ at begin, then there exists a solution to $A\lambda = y$, from **theorem 2.5**, we can conclude : $\exists \, \tilde{A}, \tilde{y} \; s.t. \; \tilde{A}\lambda = \tilde{y}$ and rows of $\tilde{A}$ are linear independent. Now we assume $\tilde{A}$ has $\bar{m}$ row,$\bar{m} \leq m$, $\therefore$ from **theorem 2.4** one of basic feasible solution of Λ has at most $\bar{m}$ non-zero λ_i, where $\bar{m} \leq m$.

(b) $\forall y \in P$, let us consider $\Lambda = \{(\lambda_1, ..., \lambda_n) \in \Re^n | \sum_{i=1}^{n} \lambda_i A_i = y, \sum_{i=1}^{n} \lambda_i = 1, \lambda_i \geq 0, 1 \leq i \leq n\}$, use the same notation as before, and mark $B = \begin{bmatrix} A \\ \mathbf{1} \end{bmatrix}$, where $\mathbf{1}$ is a $1 \times n$ vector with all element being 1.Then $\Lambda = \left\{ \lambda \in \Re^n \; | B\lambda = \begin{bmatrix} y \\ \mathbf{1} \end{bmatrix}, \lambda \geq 0 \right\}$ $\because B\lambda = \begin{bmatrix} y \\ 1 \end{bmatrix}$ has solution, from **theorem 2.5**, $\exists \, \bar{B}.\bar{y}$ s.t $\bar{B}\lambda = \bar{y}$, and rows of $\tilde{B}$ are linear independent, and assume it has m' rows, $m' \leq m+1$, $\Lambda = \left\{ \lambda \in \Re^n \; | \bar{B}\lambda = \bar{y}, \lambda \geq 0 \right\}$, from **theorem 2.4** $\therefore$ one of basic feasible solution of Λ has at most m' non-zero λ_i where $m' \leq m+1$.

Exercise 2.10

(a) True, according to our basic linear algebra knowledge, solution to $Ax = b$ can be represented as $x = x_0 + \lambda\eta$, where $\lambda \in \mathbb{R}, \eta \in \mathbb{R}^n$, $\because n = m+1$, the freedom of η is one, it implies x is a linear ($dimision = 1$) in space (you can assume $n = 2, 3$ to help understand) , notice theorem 2.3 basic feasible solution is extreme point and vertex, and two points determine one line.
(b) False, min constant, $x \geq 0$
(c) False, use example from **(b)**
(d) True, because we consider linear optimization, objective function is linear, it is a convex problem, let x_1, x_2 be the optimal solution, then any convex combination of x_1, x_2 will be optimal.

(e) False.

$$\begin{aligned} \min \quad & x_1 \\ \text{s.t.} \quad & x_1 = 1 \\ & x_1, x_2 \geq 0 \end{aligned}$$

only (1,0) is basic feasible optimal solutions
(f) False,

$$\begin{aligned} \min \max \{x_1, 1 - x_1\} & \\ \text{s.t.} \quad & x_1 + x_2 = 1 \\ & x_1, x_2 \geq 0 \end{aligned}$$

the optimal point is (1/2,1/2) which is not extreme point

Exercise 2.12

True,this can be easily inferred by **Corollary 2.2**: every nonempty bounded polyhedron and every nonempty polyhedron in standard form has at least one basic feasible solution. we can transform (replace) every $x_i \leq 0$ to $\bar{x}_i = -x_i \geq 0$, and reformulate our problem into standard one, then use Corollary 2.2.
Or, by theorem 2.6, we need to prove The polyhedron P does not contain a line. Statement:for each x_i we have either the constraint $x_i \geq 0$ or constraint $x_i \leq 0$.we know it is impossible to contain a line.

Exercise 2.16

1. This can **not** be feasible set in standard form in $\mathbb{R}^n$
2. This can be feasible set in standard form in higher dimension. Because we can introduce a slack variable, x_{n+1}

and require $x_n + x_{n+1} = 1, x_{n+1} \geq 0$ then

$$\left\{ x \in \mathbb{R}^{n+1} | \begin{bmatrix} I_{n-1} & & \\ & 1 & 1 \end{bmatrix} x = [0 \ ... \ 0 \ \ 1]^{\mathrm{T}}, x \geq 0 \right\}$$

Separating Hyperplane Theorem

Theorem 1 Suppose C and D are nonempty dis-joint convex sets, *i.e.*, $C \cap D = \emptyset$. Then there exist $a \neq 0$ and b such that $a^T x \leq b$ for all $x \in C$ and $a^T x \geq b$ for all $x \in D$. In other words, the affine function $a^T x - b$ is non-positive on C and non-negative on D. The hyperplane $\{x | a^T x = b\}$ is called a separating hyperplane for the sets C and D, or is said to separate the sets C and D.(This theorem is copying from convex optimization Stephen Boyd)

Theorem 2 Let $C \subset \mathcal{E}$, where $\mathcal{E}$ is either $\mathcal{R}_n$ or $\mathcal{M}_n$, be a closed convex set and let b be a point exterior to C. Then there is a vector $a \subset \mathcal{E}$ such that

$$a^T b > \sup_{x \in C} a^T x, \qquad x \subset C$$

where a is the norm direction of the hyperplane.

Proof

In this part, firstly, we prove **theorem 1** and, use it to prove **theorem 2**.

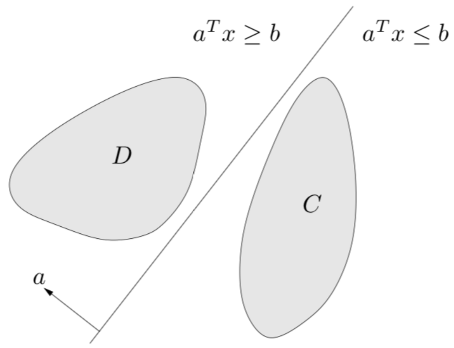

Figure 1: The hyperplane $\left\{x|a^T x = b\right\}$ separates the disjoint convex sets C and D. The affine function $a^T x - b$ is nonpositive on C and nonnegative on D

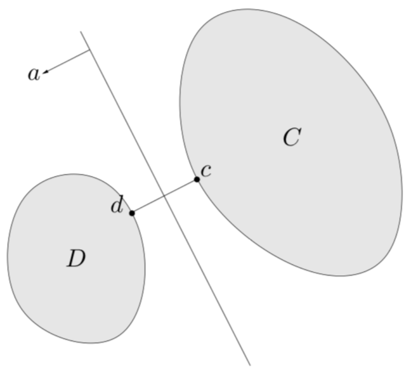

Figure 2: Construction of a separating hyperplane between two convex sets. The points $c \in C$ and $d \in D$ are the pair of points in the two sets that are closest to each other. The separating hyperplane is orthogonal to, and bisects, the line segment between c and d.

Here we assume the distance between C and D, defined as

$$\mathbf{dist}(C, D) = \inf \left\{ \|u - v\|_2 \, | u \in C, v \in D \right\},$$

is positive, and that there exist points $c \in C$ and $d \in D$ that achieve the minimum distance, *i.e.*, $\|c-d\|_2 = \mathbf{dist}(C, D)$.(These conditions are satisfied, for example, when C and D are closed and one set is bounded.)

define

$$a = d - c, \qquad b = \frac{\|d\|_2^2 - \|c\|_2^2}{2}$$

We will show that the affine function

$$f(x) = a^T x - b = (d-c)^T(x - (1/2)(d+c))$$

is nonpositive on C and nonnegative on D, i.e., that the hyperplane $\{x | a^T x = b\}$ separates C and D. This hyperplane is perpendicular to the line segment between c and d, and passes through its midpoint, as shown in figure 2.

we first show that f is nonnegative on D, The proof that f is nonpositive on C is similar (or follows by swapping C and D and considering $-f$). Suppose there were a point $u \in D$ for which

$$f(u) = (d-c)^T(u - (1/2)(d+c)) < 0$$

We can express $f(u)$ as

$$f(u) = (d-c)^T(u - d + (1/2)(d-c)) = (d-c)^T(u-d) + 1/2\,\|\text{-}c\|_2^2$$

We see that implies $(d-c)^T(u-d) < 0$. Now we observe that

$$\frac{d}{dt}\|d + t(u-d) - c\|_2^2 \bigg|_{t=0} = 2(d-c)^T(u-d) < 0,$$

so for some small $t > 0$, with $t \le 1$, we have

$$\|d + t(u-d) - c\|_2 < \|d - c\|_2$$

i.e., the point $d + t(u-d)$ is closer to c than d is. Since D is convex and contain d and u, we have $d + t(u-d) \in D$. But t this is impossible, since d is assumed to be the point in D that is closest to C.

Until now, Theorem 1 is proven. Now let us focus on theorem 2. Because point b is exterior to C, we can find a small ball with radius ϵ(small enough), centering at b that isn't intersect with C, let us denote this small ball as $B(b, \epsilon)$.

$$B(b, \epsilon) \cap C = \emptyset$$

By theorem 1, we can find a separating hyperplane $\{x \ | a^T x = d\}$ for C and $B(b, \epsilon)$, we can assume $\forall x \in B(b, \epsilon), a^T x - d > 0$ and $\forall x \in C, a^T x - d < 0$, (otherwise, we can use $-a$ to replace a)
$\therefore \sup\limits_{x \in C} a^T x < d$, $\ \because$ point $b \in B(b, \epsilon), \therefore a^T b > d > \sup\limits_{x \in C} a^T x$