Midterm Review

LI-PING ZHANG

Department of Mathematical Sciences

Tsinghua University, Beijing 100084

Office: New Science Building #A302, Tel: 62798531

E-mail: lipingzhang@tsinghua.edu.cn

Affine, Line, Convex, and Conic Combinations

When $\mathbf{x}$ **and** $\mathbf{y}$ are two distinct points in R^n and α runs over R ,

$$\{\mathbf{z} : \mathbf{z} = \alpha\mathbf{x} + (1 - \alpha)\mathbf{y}\}$$

is the line determined by $\mathbf{x}$ **and** $\mathbf{y}$, and it is called the affine combination of $\mathbf{x}$ **and** $\mathbf{y}$. When $0 \leq \alpha \leq 1$, it is called the convex combination of $\mathbf{x}$ **and** $\mathbf{y}$ and it is the line segment between $\mathbf{x}$ **and** $\mathbf{y}$. Also, the set

$$\{\mathbf{z} : \mathbf{z} = \alpha\mathbf{x} + \beta\mathbf{y}\}$$

for multipliers α and β is the linear combination of $\mathbf{x}$ **and** $\mathbf{y}$. When $\alpha \geq 0$ and $\beta \geq 0$, such z is called a conic combination.

Convex Set

- Let $\Omega \subseteq R^n$. Then Ω is said to be a convex set if for every $\mathbf{x}^1, \mathbf{x}^2 \in \Omega$ and every real number $\alpha \in [0, 1]$, the point $\alpha\mathbf{x}^1 + (1-\alpha)\mathbf{x}^2 \in \Omega$.

- The convex hull of Ω is defined by

$$\textbf{co}\Omega = \left\{ \sum_{k=1}^{m} \lambda_k \mathbf{x}^k : \mathbf{x}^k \in \Omega, \lambda_k \geq 0, \sum_{k=1}^{m} \lambda_k = 1, 1 \leq m \leq n+1 \right\}.$$

- The affine hull of Ω is defined by

仿射包

$$\textbf{aff}\Omega = \left\{ \sum_{k=1}^{m} \lambda_k \mathbf{x}^k : \mathbf{x}^k \in \Omega, k = 1, \ldots, m, \sum_{k=1}^{m} \lambda_k = 1, m \geq 1 \right\}.$$

- A point in a convex set is an extreme point if and only if it cannot be represented as a convex combination of two distinct points in the set.

Carathéodory's Theorem

Theorem 1 (Carathéodory's Theorem) *Let* $\Omega \subseteq \mathcal{R}^n$ *and* $x \in \mathrm{co}(\Omega)$*. Then there exist at most* $n+1$ *points in* Ω *such that* x *can be expressed as their convex combination, that is, there exist* $x^1, \ldots, x^p \in \Omega$ *such that*

$$x = \sum_{i=1}^{p} \alpha_i x^i, \quad \sum_{i=1}^{p} \alpha_i = 1, \quad \alpha_i \geq 0 (i = 1, \ldots, p), \quad 1 \leq p \leq n+1.$$

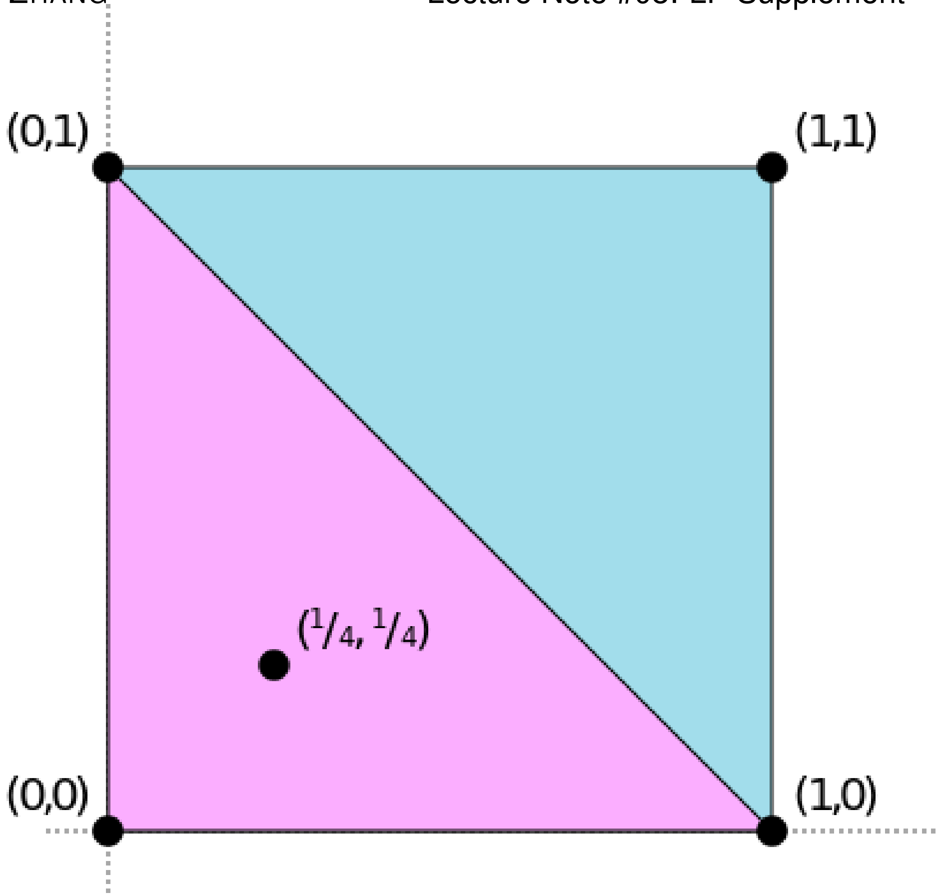


Figure 1: The convex hull of $\{(0,0),(0,1),(1,0),(1,1)\}$ is a square in $\mathcal{R}^2$

Proof of Carathéodory's Theorem

- Let $\mathbf{x}^1, \ldots, \mathbf{x}^m \in \mathcal{R}^n (m \geq n+2)$ and

$$\mathbf{x} = \sum_{i=1}^{m} \alpha_i \mathbf{x}^i, \quad \sum_{i=1}^{m} \alpha_i = 1, \alpha_i \geq 0 (i = 1, \ldots, m).$$

  Then there exist at most $n+1$ points such that $\mathbf{x}$ is their convex combination.

- $\mathrm{co}(\Omega)$ is equal to the set of all convex combinations of all finite subsets of points.

$$\mathbf{x} = \sum_{i=1}^{p} \alpha_i \mathbf{x}^i, \quad \sum_{i=1}^{p} \alpha_i = 1, \alpha_i \geq 0 (i = 1, \ldots, p), \quad p \geq n + 2.$$

$$\sum_{i=1}^{p-1} \beta_i (\mathbf{x}^i - \mathbf{x}^p) = 0.$$

Let

$$\beta_p = -\sum_{i=1}^{p-1} \beta_i, \quad \tau = \min\{\frac{\alpha_i}{\beta_i} | \beta_i > 0\}, \quad \alpha_i' = \alpha_i - \tau \beta_i.$$

Then

$$\mathbf{x} = \sum_{i=1}^{p} \alpha_i' \mathbf{x}^i, \quad \sum_{i=1}^{p} \alpha_i' = 1, \alpha_i' \geq 0 (i = 1, \ldots, p)$$

with some $\alpha_i' = 0$

Let S be the set of all convex combinations of all finite subsets of points. Then S is a convex set and $S \subseteq \mathrm{co}(\Omega)$.

Let $\mathbf{x} \in S$. There exist $\mathbf{x}^1, \ldots, \mathbf{x}^m \in \Omega$ such that

$$\mathbf{x} = \sum_{i=1}^{m} \alpha_i \mathbf{x}^i, \quad \sum_{i=1}^{m} \alpha_i = 1, \alpha_i \geq 0 (i = 1, \ldots, m).$$

Then we have $\mathbf{x} \in \mathrm{co}(\Omega)$. Clearly, it holds for $m = 1$. We now assume that it holds for $m - 1$ points. If $\alpha_m = 1$, it holds. If $\alpha_m < 1$, we have

$$\mathbf{x} = (1 - \alpha_m) \sum_{i=1}^{m-1} \frac{\alpha_i}{1 - \alpha_m} \mathbf{x}^i + \alpha_m \mathbf{x}^m \in \mathrm{co}(\Omega).$$

Proof of convex set

- All solutions to the system of linear equations and inequalities, $\{\mathbf{x} : \ A\mathbf{x} = \mathbf{b}, \ \mathbf{x} \geq \mathbf{0}\}$, form a convex set.
- Given a matrix A, let's consider the set $\mathcal{B}$ of all $\mathbf{b}$ such that the set $\{\mathbf{x} : \ A\mathbf{x} = \mathbf{b}, \ \mathbf{x} \geq \mathbf{0}\} \neq \emptyset$. Show that $\mathcal{B}$ is a convex set, where

$$\mathcal{B} = \{\mathbf{b} : \ \{\mathbf{x} : \ A\mathbf{x} = \mathbf{b}, \ \mathbf{x} \geq \mathbf{0}\} \neq \emptyset\}.$$

Polyhedral Convex Cones

- A cone C is a (convex) polyhedral if C can be represented as

两种多面锥的形式

$$C = \{\mathbf{x} : \ A\mathbf{x} \le \mathbf{0}\} \quad \text{or} \quad \{\mathbf{x} : \ \mathbf{x} = A\mathbf{y}, \mathbf{y} \ge \mathbf{0}\}$$

for some matrix A. In the latter case, C is generated by the columns of A.

- A set is polyhedral if and only if it has finite number of extreme points.

Expression of Polyhedral Cone

The following theorem states that a polyhedral cone can be generated by a set of basic directional vectors.

Theorem 2 *Given matrix* $A \in \mathcal{R}^{m\times n}$ *where* $n > m$, *take a convex polyhedral cone* $C = \{A\mathbf{x} : \ \mathbf{x} \geq \mathbf{0}\}$. *Then for any* $\mathbf{b} \in C$,

$$\mathbf{b} = \sum_{i=1}^{d} \mathbf{a}_{j_i} x_{j_i}, \quad x_{j_i} \geq 0, \ \forall i$$

for some linearly independent vectors $\mathbf{a}_{j_1}, \ldots, \mathbf{a}_{j_d}$ *chosen from* $\mathbf{a}_1, \ldots, \mathbf{a}_n$.

Proof of Theorem 2

Let $\mathbf{b} \in C$ and, without loss of generality, suppose that $\mathbf{b} = A\mathbf{x}$ and $\mathbf{x} = (x_1; \ldots; x_k; 0; \ldots; 0)$, where $x_j > 0$ for $j = 1, \ldots, k$.

If $A_{.1}, \ldots, A_{.k}$ are linearly independent, then $k \leq m$ and the conclusion holds.

Otherwise, there exist scalars $\lambda_1, \ldots, \lambda_k$ with at least one positive component such that $\sum_{j=1}^{k} \lambda_j A_{.j} = \mathbf{0}$.

Define $\alpha > 0$ as follows:

$$\alpha = \min_{1 \leq j \leq k} \left\{ \frac{x_j}{\lambda_j} : \lambda_j > 0 \right\} = \frac{x_t}{\lambda_t}.$$

Consider the point $\mathbf{x}'$ whose jth component x'_j is given by

$$x'_j = \begin{cases} x_j - \alpha\lambda_j & \text{for } j = 1, \ldots, k, \\ 0 & \text{for } j = k+1, \ldots, n. \end{cases}$$

Note that $x'_j \geq 0$ for $j = 1, \ldots, k$ and $x'_j = 0$ for $j = k+1, \ldots, n$. Moreover, $x'_t = 0$, and $A\mathbf{x}' = \mathbf{b}$.

So far, we have constructed such a new point x' with at most $k-1$ positive components. This process is continued until the positive components correspond to linearly independent columns, which results in the conclusion.

$C = \{A\mathbf{x} : \ \mathbf{x} \geq 0\}$ is a Closed Set

C is a closed set, that is, for every convergence sequence $\mathbf{c}^k \in C$, the limit of $\{\mathbf{c}^k\}$ is also in C.

The key to prove the statement is to show that $\mathbf{c}^k = A\mathbf{x}^k$ for a bounded sequence $\mathbf{x}^k \geq \mathbf{0}$. By Theorem 2, there exists a basic feasible solution $(\mathbf{x}^k_{B^k}, \mathbf{x}^k_{N^k})$ such that

$$\mathbf{c}^k = A_{B^k}\mathbf{x}^k_{B^k}, \quad \mathbf{x}^k_{B^k} \geq \mathbf{0}, \quad \mathbf{x}^k_{N^k} = \mathbf{0}.$$

Clearly, $\{\mathbf{x}^k\}$ is bounded since $\mathbf{x}^k_{B^k} = A^{-1}_{B^k}\mathbf{c}^k$ is bounded.

Remark

Note that C may not be closed if $\mathbf{x}$ is in other cones. Let

$$C = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : \begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix} \succeq \mathbf{0} \right\}.$$

Then C is not closed, since $(0; 1)$ is not in C but it is a limit point of sequence $c^k \in C$.

Separating Hyperplane Theorem

The most important theorem about the convex set is the following separating hyperplane theorem.

Theorem 3 (Separating hyperplane theorem) *Let $C \subset \mathcal{R}^n$ be a closed convex set, and let $\mathbf{b} \notin C$. Then there is a vector $\mathbf{a} \neq \mathbf{0}$ such that*

$$\mathbf{a} \bullet \mathbf{b} > \sup_{\mathbf{x} \in C} \mathbf{a} \bullet \mathbf{x}$$

Farkas' Lemma

The following results are Farkas' lemma and its variants.

Theorem 4 *Let $A \in \mathcal{R}^{m\times n}$ and $\mathbf{b} \in \mathcal{R}^m$. The system $A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}$ has a feasible solution $\mathbf{x}$ if and only if that $A^T\mathbf{y} \leq \mathbf{0}$ implies $\mathbf{b}^T\mathbf{y} \leq 0$.*

A vector $\mathbf{y}$, with $A^T\mathbf{y} \leq \mathbf{0}$ and $\mathbf{b}^T\mathbf{y} = 1$, is called a (primal) infeasibility certificate for the system $\{\mathbf{x} : \ A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$.

Geometrically, Farkas lemma means that if a vector $\mathbf{b} \in \mathcal{R}^m$ does not belong to the cone generated by $A_{.1}, \ldots, A_{.n}$, then there is a hyperplane separating $\mathbf{b}$ from $\text{cone}(A_{.1}, \ldots, A_{.n})$.

Theorem 5 *Let* $A \in \mathcal{R}^{m\times n}$ *and* $\mathbf{c} \in \mathcal{R}^n$*. The system* $A^T\mathbf{y} \le \mathbf{c}$ *has a feasible solution* $\mathbf{y}$ *if and only if that* $A\mathbf{x} = \mathbf{0}$ *and* $\mathbf{x} \ge \mathbf{0}$ *imply* $\mathbf{c}^T\mathbf{x} \ge 0$*.*

A vector $\mathbf{x} \ge \mathbf{0}$, with $A\mathbf{x} = \mathbf{0}$ and $\mathbf{c}^T\mathbf{x} = -1$, is called a (dual) infeasibility certificate for the system $A^T\mathbf{y} \le \mathbf{c}$.

Level Set and Epigraph

Let $f : \Omega \subseteq R^n \to R$ be a real-valued function.

- The level set of f is defined by

$$L(z) = \{\mathbf{x} \in \Omega : \quad f(\mathbf{x}) \leq z\}.$$

- The epigraph of f is defined by

$$\textbf{epi } f = \{(\mathbf{x}, \mu) \in R^{n+1} : \ \mathbf{x} \in \Omega, \ \ f(\mathbf{x}) \leq \mu\}.$$

Convex Function

- A function f defined on the convex set Ω is said to be convex if

$$f(\alpha \mathbf{x} + (1-\alpha)\mathbf{y}) \leq \alpha f(\mathbf{x}) + (1-\alpha) f(\mathbf{y}), \quad \forall x, y \in \Omega, 0 \leq \alpha \leq 1.$$

- f is said to be strictly convex if the above inequality holds strictly whenever x and y are distinct in Ω and and $0 < \alpha < 1$.

- f is said to be strongly convex if it is convex and there exists a positive constant $c > 0$ such that for any $x, y \in \Omega$ and $0 \leq \alpha \leq 1$, 强凸的

$$f(\alpha \mathbf{x} + (1-\alpha)\mathbf{y}) \leq \alpha f(\mathbf{x}) + (1-\alpha) f(\mathbf{y}) - \frac{c}{2}\alpha(1-\alpha)\|x-y\|^2.$$

Properties of Convex Function

- The level set of a convex function f is a convex set. The converse is not true; e.g., $f(x) = x^3$.
- Let $S \subseteq R^n$ be a nonempty convex set. Then $f : S \to R$ is a convex function iff its epigraph **epi** f is a convex set.

Theorems on Convex Functions

Theorem 6 *Let $f \in C^1$. Then f is convex over a convex set Ω if and only if the gradient inequality holds, i.e.,*

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})(\mathbf{y} - \mathbf{x})$$

for all $\mathbf{x},\ \mathbf{y} \in \Omega$.

Theorem 7 *Let $f \in C^2$. Then f is convex over a open convex set Ω if and only if the Hessian matrix of f is positive semi-definite throughout Ω.*

Example Show that the Cobb-Douglas utility function $u : \mathcal{R}^2_+ \to \mathcal{R}$ defined by

$$u(x_1, x_2) = x_1^a x_2^b, \quad a, b > 0,$$

is concave iff $a + b \leq 1$.

Quasi-Concave Function

For a function f on $S \subset \mathcal{R}^n$ and a point $\mathbf{x} \in S$, the upper-contour set (lower-contour set) of f at $\mathbf{x}$ is defined by

$$U(f;\mathbf{x}) = \{\mathbf{y} \in S | f(\mathbf{y}) \geq f(\mathbf{x})\}, \quad \text{向上等值基}$$

$$L(f;\mathbf{x}) = \{\mathbf{y} \in S | f(\mathbf{y}) \leq f(\mathbf{x})\}, \quad \text{向下等值集}$$

respectively.

Definition 1 *A function f on a convex set $S \subset \mathcal{R}^n$ is quasi-concave (quasi-convex) if its upper-contour set (lower-contour set) $U(f;\mathbf{x})$ ($L(f;\mathbf{x})$) is a convex set at every $\mathbf{x} \in S$.*

Another Definition of Quasi-Concave Function

Obviously, all concave functions are quasi-concave. There is another equivalent way to define a quasi-concave function.

Theorem 8 *A function f is quasi-concave (quasi-convex) if and only if for any $\mathbf{x}, \mathbf{y} \in C$ and any $\alpha \in (0, 1)$,*

$$f(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}) \geq \min\{f(\mathbf{x}), f(\mathbf{y})\}. \tag{1}$$

$$f(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}) \leq \max\{f(\mathbf{x}), f(\mathbf{y})\}.$$

Proof: First, suppose f is quasi-concave. For any $\mathbf{x}, \mathbf{y} \in C$ and any $\alpha \in (0,1)$, we may assume, WLOG, $f(\mathbf{y}) \geq f(\mathbf{x})$. Then, by the definition, $\mathbf{x}, \mathbf{y} \in U(f;\mathbf{x})$. Since $U(f;\mathbf{x})$ is convex, $\alpha\mathbf{x} + (1-\alpha)\mathbf{y} \in U(f;\mathbf{x})$. This means

$$f(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}) \geq f(\mathbf{x}) \geq \min\{f(\mathbf{x}), f(\mathbf{y})\}.$$

On the other hand, suppose f satisfies (1) and we show $U(f;\mathbf{x})$ is convex for every $\mathbf{x}$. For any $\mathbf{y}^1, \mathbf{y}^2 \in U(f;\mathbf{x})$ and any $\alpha \in (0,1)$,

$$f(\alpha\mathbf{y}^1 + (1-\alpha)\mathbf{y}^2) \geq \min\{f(\mathbf{y}^1), f(\mathbf{y}^2)\} \geq f(\mathbf{x}).$$

Thus, by the definition, $\alpha\mathbf{y}^1 + (1-\alpha)\mathbf{y}^2 \in U(f;\mathbf{x})$. Hence, $U(f;\mathbf{x})$ is convex.

Example of Quasi-Concave Function

Theorem 9 *Let $f : \mathcal{R} \to \mathcal{R}$ be any nondecreasing function on $\mathcal{R}$. Then, f is both quasi-convex and quasi-concave on $\mathcal{R}$.*

Example: $f(x) = x^3$ is neither concave nor convex on $\mathcal{R}$, but it is both quasi-convex and quasi-concave on $\mathcal{R}$.

Linear Programming and its Dual

Consider the linear program in standard form, called the primal problem,

$$(LP) \quad \textbf{minimize} \quad \mathbf{c}^T\mathbf{x}$$
$$\textbf{subject to} \quad A\mathbf{x} = \mathbf{b}, \ \mathbf{x} \geq \mathbf{0},$$

where $\mathbf{x} \in \mathcal{R}^n$.

The dual problem can be written as:

$$(LD) \quad \textbf{maximize} \quad \mathbf{b}^T\mathbf{y}$$
$$\textbf{subject to} \quad A^T\mathbf{y} + \mathbf{s} = \mathbf{c}, \ \mathbf{s} \geq \mathbf{0},$$

where $\mathbf{y} \in \mathcal{R}^m$ and $\mathbf{s} \in \mathcal{R}^n$. The components of s are called dual slacks.

Rules to construct the dual

obj. coef. vector	right-hand-side
right-hand-side	obj. coef. vector
A	A^T
Max model	Min model
$x_j \geq 0$	jth constraint $\geq$
$x_j \leq 0$	jth constraint $\leq$
x_j free	jth constraint $=$
ith constraint $\leq$	$y_i \geq 0$
ith constraint $\geq$	$y_i \leq 0$
ith constraint $=$	y_i free

Duality Theory

Theorem 10 (Weak duality theorem) *Let feasible regions $\mathcal{F}_p$ and $\mathcal{F}_d$ be non-empty. Then,*

$$\mathbf{c}^T\mathbf{x} \geq \mathbf{b}^T\mathbf{y} \quad \textbf{where} \quad \mathbf{x} \in \mathcal{F}_p, \; (\mathbf{y}, \mathbf{s}) \in \mathcal{F}_d.$$

This theorem shows that a feasible solution to either problem yields a bound on the value of the other problem. We call $\mathbf{c}^T\mathbf{x} - \mathbf{b}^T\mathbf{y}$ the duality gap.

Theorem 11 (Strong duality theorem) *Let $\mathcal{F}_p$ and $\mathcal{F}_d$ be non-empty. Then, $\mathbf{x}^*$ is optimal for (LP) if and only if the following conditions hold:*

i) $\mathbf{x}^* \in \mathcal{F}_p$;

可行且对偶间隙为0则都最优

ii) *there is* $(\mathbf{y}^*, \mathbf{s}^*) \in \mathcal{F}_d$;

iii) $\mathbf{c}^T\mathbf{x}^* = \mathbf{b}^T\mathbf{y}^*$.

Theorem 12 (LP duality theorem) *If (LP) and (LD) both have feasible solutions then both problems have optimal solutions and the optimal objective values of the objective functions are equal.*

If one of (LP) or (LD) has no feasible solution, then the other is either unbounded or has no feasible solution. If one of (LP) or (LD) is unbounded then the other has no feasible solution.

The above theorems show that if a pair of feasible solutions can be found to the primal and dual problems with equal objective values, then these are both optimal. The converse is also true; there is no “gap”.

Optimality Conditions

For feasible $\mathbf{x}$ and $(\mathbf{y}, \mathbf{s})$, $\mathbf{x}^T\mathbf{s} = \mathbf{x}^T(\mathbf{c} - A^T\mathbf{y}) = \mathbf{c}^T\mathbf{x} - \mathbf{b}^T\mathbf{y}$ is called the complementarity gap.

Since both $\mathbf{x}$ and $\mathbf{s}$ are nonnegative, $\mathbf{x}^T\mathbf{s} = 0$ implies that $x_j s_j = 0$ for all $j = 1, \ldots, n$, where we say $\mathbf{x}$ and $\mathbf{s}$ are complementary to each other.

$$\begin{array}{rcl} X\mathbf{s} & = & \mathbf{0} \\ A\mathbf{x} & = & \mathbf{b} \\ -A^T\mathbf{y} - \mathbf{s} & = & -\mathbf{c}, \end{array} \tag{2}$$

where X is the diagonal matrix of vector $\mathbf{x}$.

This system has total $2n + m$ unknowns and $2n + m$ equations including n nonlinear equations.

LP Fundamental Theorem

Theorem 13 *Given* (LP) *where* A *has full row rank* m*,*

(i) *if there is a feasible solution, there is a basic feasible solution;*

(ii) *if there is an optimal solution, there is an optimal basic solution.*

Strict Complementarity of LP

Theorem 14 (Strict complementarity theorem) *If (LP) and (LD) both have feasible solutions then both problems have a pair of strictly complementary solutions $\mathbf{x}^* \geq \mathbf{0}$ and $\mathbf{s}^* \geq \mathbf{0}$ meaning*

$$X^*\mathbf{s}^* = 0 \quad \textbf{and} \quad \mathbf{x}^* + \mathbf{s}^* > 0.$$

Moreover, the supports

$$P^* = \{j : x_j^* > 0\} \quad \textbf{and} \quad Z^* = \{j : s_j^* > 0\}$$

are invariant for all pairs of strictly complementary solutions.

对于严格互补最优解可能很多时，其不为0的位置不变

Given (LP) or (LD), the pair of P^* and Z^* is called the (strict) complementarity partition.

原问题严格最优面

$$\{\mathbf{x}: \ A_{P^*}\mathbf{x}_{P^*} = b, \ \mathbf{x}_{P^*} > 0, \ \mathbf{x}_{Z^*} = 0\}$$

is called the primal optimal face, and

$$\{\mathbf{y}: \ c_{Z^*} - A^T_{Z^*}\mathbf{y} > 0, \ c_{P^*} - A^T_{P^*}\mathbf{y} = 0\}$$

is called the dual optimal face.

Proof of strict complementarity theorem

We only need to show that exactly one of the following holds:

either **(i) (LD) has an optimal solution with** $s_i^* > 0$

or **(ii) (LP) has an optimal solution with** $x_i^* > 0.$

Suppose now (i) is not satisfied. That is, there is no optimal solution $\mathbf{s}^*$ for (LD) with $s_i^* > 0$. Let z^* be the common value of the LP-duality equation

$$\max\{\mathbf{b}^T\mathbf{y} | A^T\mathbf{y} \le \mathbf{c}\} = \min\{\mathbf{c}^T\mathbf{x} | A\mathbf{x} = \mathbf{b},\ \mathbf{x} \ge \mathbf{0}\}. \tag{3}$$

Then,

$$A^T\mathbf{y} \le \mathbf{c}, \mathbf{b}^T\mathbf{y} \ge z^* \Rightarrow s_i \le 0,\ i.e.,\ A_{i.}^T\mathbf{y} \ge c_i.$$

z*为最优值，这个其实是等号

That is, the following system of inequalities is infeasible 由$A_i^T y \geq c_i$

$$A^T \mathbf{y} \leq \mathbf{c}, \quad -\mathbf{b}^T \mathbf{y} \leq -z^*, \quad -A_{i.}^T \mathbf{y} > -c_i.$$

By Farkas' Lemma,

$$A\mathbf{x} - \alpha \mathbf{b} = -A_{.i} \quad \textbf{and} \quad \mathbf{c}^T \mathbf{x} - \alpha z^* = -c_i$$

hold for some $\mathbf{x} \geq \mathbf{0}, \alpha \geq 0$.

$c^T x' = \alpha z^*$

Let $\mathbf{x}' = \mathbf{x} + \mathbf{e}_i$, where $\mathbf{e}_i$ is a vector with one as its i^{th} component and zero as the other. Then, $\mathbf{x}' \geq \mathbf{0}$ with $x_i' > 0$.

$Ax + A_{.i} = Ax' = \alpha b$

If $\alpha = 0$, then $A\mathbf{x} + A_{.i} = \mathbf{0}$ and $\mathbf{c}^T\mathbf{x} + c_i = 0$. Define $\bar{\mathbf{x}} = \mathbf{x}^* + \mathbf{x}'$. Then, $\bar{\mathbf{x}}$ is optimal for (LP) since $\bar{\mathbf{x}} \geq \mathbf{0}$ and

$$A\bar{\mathbf{x}} = A\mathbf{x}^* + A\mathbf{x}' = \mathbf{b} + A\mathbf{x} + A\mathbf{e}_i = \mathbf{b} + A\mathbf{x} + A_{.i} = \mathbf{b},$$

and

$$\mathbf{c}^T\bar{\mathbf{x}} = \mathbf{c}^T\mathbf{x}^* + \mathbf{c}^T\mathbf{x} + c_i = \mathbf{c}^T\mathbf{x}^*.$$

Clearly, $\bar{\mathbf{x}}_i > 0$, and hence (ii) is fulfilled.

If $\alpha > 0$, then $\mathbf{x}'/\alpha$ is optimal for (LP) as

$$A\mathbf{x}'/\alpha = \frac{1}{\alpha}(A\mathbf{x} + A_{.i}) = \mathbf{b}$$

and

$$\mathbf{c}^T\mathbf{x}'/\alpha = \frac{1}{\alpha}(\mathbf{c}^T\mathbf{x} + c_i) = z^*,$$

and $\mathbf{x}'/\alpha$ has positive i^{th} component. This shows (ii).

为了证明对于严格互补最优解可能很多时，其不为0的位置不变

只需证明

Let $(\mathbf{x}^1, \mathbf{s}^1)$ and $(\mathbf{x}^2, \mathbf{s}^2)$ be two strict complementarity solution pairs. Then, by the strong duality theorem, we have

$$0 = (\mathbf{x}^1)^T \mathbf{s}^2 = (\mathbf{x}^2)^T \mathbf{s}^1.$$

This indicates that they must have same strict complementarity partition.

An Example

Consider the primal problem:

$$
\begin{array}{llllll}
\textbf{minimize} & x_1 & +x_2 & +1.5x_3 & & \\
\textbf{subject to} & x_1 & & + & x_3 & =1 \\
 & & x_2 & + & x_3 & =1 \\
 & x_1, & x_2, & & x_3 & \geq 0;
\end{array}
$$

Its equivalent form is

$$
\begin{array}{ll}
\min & 2-0.5x_3 \\
s.t. & 0 \leq x_3 \leq 1.
\end{array}
$$

该问题是退化的

于是其对偶问题有∞多个解

Clearly, the problem has a unique optimal solution $\mathbf{x}^* = (0;0;1)$ and $P^* = \{3\}$.

The dual problem is

$$\begin{array}{llllll} \textbf{maximize} & y_1 & +y_2 & & & \\ \textbf{subject to} & y_1 & & +s_1 & = & 1 \\ & & y_2 & +s_2 & = & 1 \\ & y_1 & +y_2 & +s_3 & = & 1.5 \\ & & & \mathbf{s} & \geq & 0. \end{array}$$

Since $P^* = \{3\}$, $Z^* = \{1, 2\}$ and hence the feasible solutions on $\{y_1 + y_2 = 1.5\}$ are all strictly complementary optimal solutions.

An Application

Given a matrix $A \in \mathcal{R}^{m \times n}$, show that the system

$$Ax \geq 0,\ A^T y = 0,\ y \geq 0$$

must have a solution, denoted by $(x^*; y^*)$, such that $Ax^* + y^* > 0$.

Proof

Consider the following LP:

$$(P) \quad \textbf{min} \quad 0^T x$$
$$\textbf{s.t.} \quad Ax \geq 0,$$

and its dual:

$$(D) \quad \textbf{max} \quad 0^T y$$
$$\textbf{s.t.} \quad A^T y = 0,\ y \geq 0.$$

其严格互补解即为所求

Primal Basic Feasible Solution

In the LP standard form, select m linearly independent columns, denoted by the index set B, from A.

$$A_B x_B = b$$

for the m-vector x_B. By setting the variables, x_N, of x corresponding to the remaining columns of A equal to zero, we obtain a solution x such that

$$Ax = b.$$

Then, x is said to be a (primal) basic solution to (LP) with respect to the basis A_B. The components of x_B are called basic variables.

If a basic solution $x \geq 0$, then x is called a basic feasible solution.

If one or more components in x_B has value zero, the basic feasible solution x is said to be (primal) degenerate.

Dual Basic Feasible Solution

For the basis A_B, the dual vector y satisfying

$$A_B^T y = c_B$$

is said to be the corresponding dual basic solution.

If the dual basic solution is also feasible, that is

$$s = c - A^T \mathbf{y} \geq 0,$$

then $\mathbf{x}$ is called an optimal basic solution, A_B an optimal basis and $\mathbf{y}$ is said to be a dual basic feasible solution.

If one or more components in s_N has value zero, the basic feasible solution y is said to be (dual) degenerate.

Problems on the Simplex Method: Problem I

Assume that all basic feasible solutions (BFS) of a standard LP problem are non degenerate, that is, every basic variable has a positive value at every BFS. Then consider using the Simplex method to solve the problem.

Prove that, if at a pivot step there is exactly one negative reduced cost coefficient, then the corresponding entering variable will remain as a basic variable for the remaining steps of the Simplex method.

reduced cost里面只有一个负值时，他进基之后就再也不会出去

易见此时进去可以使得目标函数值严格下降

$$f_0' = f_0 + r_s \frac{\bar{b}_0}{a_{0s}}$$

Solution to Problem I

Suppose the LP is

$$\min\{f(\mathbf{x})|\ A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\},$$

where f is linear, $\mathbf{x} \in \mathcal{R}^n$ and $A \in \mathcal{R}^{m \times n}$. WLOG, assume the objective function is

$$f(x) = -x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \cdots + \alpha_{n-m} x_{n-m}$$

for some nonnegative α_j's and the current basis is

$$B = \{n - m + 1, n - m + 2, \ldots, n\}$$

In particular, the objective value is currently at 0.

Because all of the BFS's are strictly positive, the objective value decreases at each step. Let $\mathbf{x}'$ be the new BFS immediately after x_1 enters the basis. Then, $f(\mathbf{x}') < 0$.

Now, let $\hat{\mathbf{x}}$ be the BFS of an arbitrary subsequent pivot step. Then

$$0 > f(\mathbf{x}') \geq f(\hat{\mathbf{x}}) = -\hat{x}_1 + \alpha_2\hat{x}_2 + \alpha_3\hat{x}_3 + \cdots + \alpha_{n-m}\hat{x}_{n-m}.$$

Thus

$$\hat{x}_1 > \alpha_2\hat{x}_2 + \alpha_3\hat{x}_3 + \cdots + \alpha_{n-m}\hat{x}_{n-m} \geq 0.$$

In other words, x_1 is a basic variable for any subsequent pivot step.

Problem II

While solving a standard simplex form linear programming problem using the simplex method, we get the following tableau:

	x_1	x_2	x_3	x_4	x_5	
	0	0	$\bar{c}_3$	0	$\bar{c}_5$	
x_2	0	1	-1	0	β	1
x_4	0	0	2	1	γ	2
x_1	1	0	4	0	δ	3

Suppose also that the last 3 columns of the original matrix A form an identity matrix.

(a) Assume that this basis is optimal and that $\bar{c}_3 = 0$. Find an optimal basic feasible solution, other than the one described by this tableau.

(b) Suppose that $\gamma > 0$, show that there exists an optimal basic feasible solution, regardless of the values of $\bar{c}_3$ and $\bar{c}_5$.

Solution to Problem II

(a) Simply perform one iteration on the third column. We get $(x_2\ x_3\ x_4)$ is another optimal basis. The tableau is:

	x_1	x_2	x_3	x_4	x_5	
	0	0	0	0	$\bar{c_5}$	
x_2	$\frac{1}{4}$	1	0	0	$\beta + \frac{\delta}{4}$	$\frac{7}{4}$
x_4	$\frac{1}{2}$	0	0	1	$\gamma - \frac{\delta}{2}$	$\frac{1}{2}$
x_3	$\frac{1}{4}$	0	1	0	$\frac{\delta}{4}$	$\frac{3}{4}$

(b) First we see the system is feasible. If $\gamma > 0$, then consider the system corresponding to the given tableau, we have $2x_3 + x_4 + \gamma x_5 = 2$. Note that any x_i is nonnegative, from the second equation we know x_3, x_4, x_5 are bounded, then from the other 2 equations, we can prove x_1, x_2 are also bounded. Thus the object function is bounded. So there is an optimal solution over all feasible solutions. From simplex method we know the current system's optimal value only differs a constant from the original problem, so we know the original system also has an optimal solution, which means there exists an optimal basic feasible solution for the original problem.

Problem III

Given the LP problem

$$\begin{aligned} \min \quad & -2x_1 - x_2 + x_3 \\ s.t. \quad & x_1 + x_2 + 2x_3 \leq 6 \\ & x_1 + 4x_2 - x_3 \leq 4 \\ & x_1, x_2, x_3 \geq 0 \end{aligned}$$

and its optimal simplex tableau

Basic	Row	x_1	x_2	x_3	x_4	x_5	RHS
-Z	(0)	0	6	0	$\frac{1}{3}$	$\frac{5}{3}$	$\frac{26}{3}$
x_3	(1)	0	-1	1	$\frac{1}{3}$	-$\frac{1}{3}$	$\frac{2}{3}$
x_1	(2)	1	3	0	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{14}{3}$

(1) What are the optimal dual prices?

(2) Will the optimal basis change if we change $b = (6; 4)$ to $(2; 4)$? Write out the optimal tableau for the new problem via the above optimal tableau.

(3) How much can we change $c_1 = -2$ such that the optimal basis is not changed ?

Solution to Problem III

(1) In terms of the optimal simplex tableau, $r_4 = \frac{1}{3}$ and $r_5 = \frac{5}{3}$. Since

$$r = c - A^T(A_B^{-T} c_B),$$

we have

$$\begin{pmatrix} r_4 \\ r_5 \end{pmatrix} = - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} (A_B^{-T} c_B),$$

which implies the optimal dual prices

$$A_B^{-T} c_B = - \begin{pmatrix} r_4 \\ r_5 \end{pmatrix} = \begin{pmatrix} -\frac{1}{3} \\ -\frac{5}{3} \end{pmatrix}.$$

(2) From the the optimal simplex tableau,

$$A_B^{-1} = \begin{pmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}.$$

Let $b' = (2; 4)$, then

$$A_B^{-1} b' = \begin{pmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix} \begin{pmatrix} 2 \\ 4 \end{pmatrix} = \begin{pmatrix} -\frac{2}{3} \\ \frac{10}{3} \end{pmatrix},$$

and

$$c_B^T A_B^{-1} b' = \begin{pmatrix} 1 & -2 \end{pmatrix} \begin{pmatrix} -\frac{2}{3} \\ \frac{10}{3} \end{pmatrix} = -\frac{22}{3}.$$

Hence, the optimal basis is changed.

We obtain the following simplex tableau:

Basic	Row	x_1	x_2	x_3	x_4	x_5	RHS
-Z	(0)	0	6	0	$\frac{1}{3}$	$\frac{5}{3}$	$\frac{22}{3}$
x_3	(1)	0	-1	1	$\frac{1}{3}$	-$\frac{1}{3}$	-$\frac{2}{3}$
x_1	(2)	1	3	0	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{10}{3}$

We use the dual simplex method to solve the current problem. Choose x_3 as the outgoing variable and x_5 as the entering variable, and then obtain:

Basic	Row	x_1	x_2	x_3	x_4	x_5	RHS
-Z	(0)	0	1	5	2	0	4
x_5	(1)	0	3	-3	-1	1	2
x_1	(2)	1	1	2	1	0	2

This is the optimal tableau. $(2; 0; 0)$ is the optimal solution for the new problem with the optimal value -4.

(3) In this problem, we can change c_1 by $c_1' = c_1 + \Delta c_1$, so Row(0) in the final tableau will become:

$$(0, 6 - 3\Delta c_1, 0, \frac{1}{3} - \frac{1}{3}\Delta c_1, \frac{5}{3} - \frac{2}{3}\Delta c_1).$$

For these to remain nonnegative, the allowable range for Δc_1 is given by

$$6 - 3\Delta c_1 \geq 0,\ \frac{1}{3} - \frac{1}{3}\Delta c_1 \geq 0,\ \frac{5}{3} - \frac{2}{3}\Delta c_1 \geq 0 \quad \Rightarrow \quad \Delta c_1 \leq 1.$$

That is, when $c_1' \leq -1$ the optimal basis is not changed.

Detailed Canonical Tableau for Production

If the original LP is the production problem:

$$
\begin{array}{ll}
\text{maximize} & \mathbf{c}^T\mathbf{x} \\
\text{subject to} & A\mathbf{x} \le \mathbf{b}(> \mathbf{0}), \\
 & \mathbf{x} \ge \mathbf{0}.
\end{array}
$$

The initial canonical tableau for minimization would be

B	$-\mathbf{c}^T$	$\mathbf{0}$	0
Basis Indices	A	I	$\mathbf{b}$

The intermediate canonical tableau would be

B	$\mathbf{r}^T$	$-\mathbf{y}^T$	$\mathbf{c}_B^T\bar{\mathbf{b}}$
Basis Indices	$\bar{A}$	A_B^{-1}	$\bar{\mathbf{b}}$

How Good is the Simplex Method

Very good on average, but the worse case ...?

When the simplex method is used to solve a linear program the number of iterations to solve the problem starting from a basic feasible solution is typically a small multiple of m, e.g., between $2m$ and $3m$.

At one time researchers believed—and attempted to prove—that the simplex algorithm (or some variant thereof) always requires a number of iterations that is bounded by a polynomial expression in the problem size.

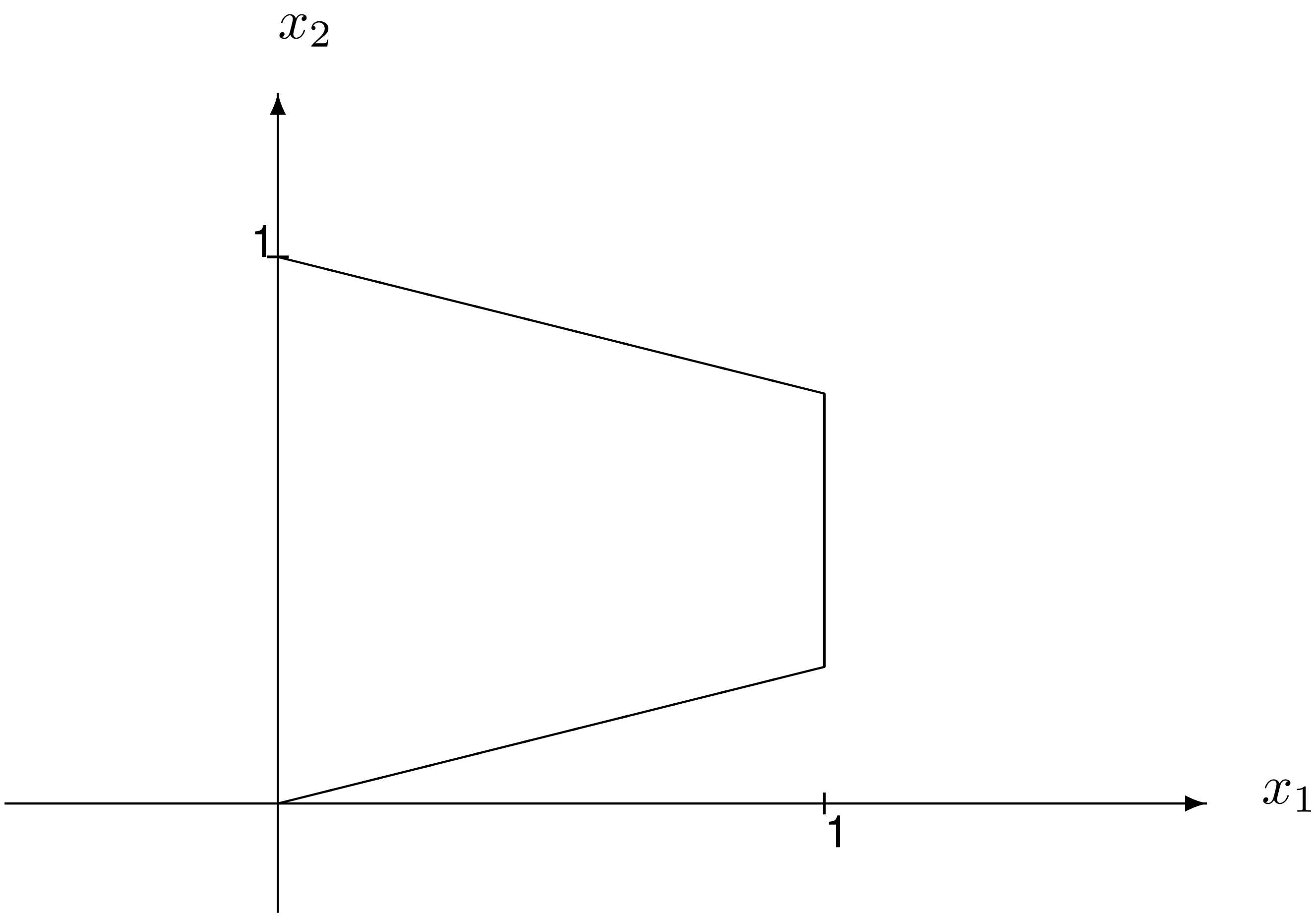
Klee and Minty Example

Consider

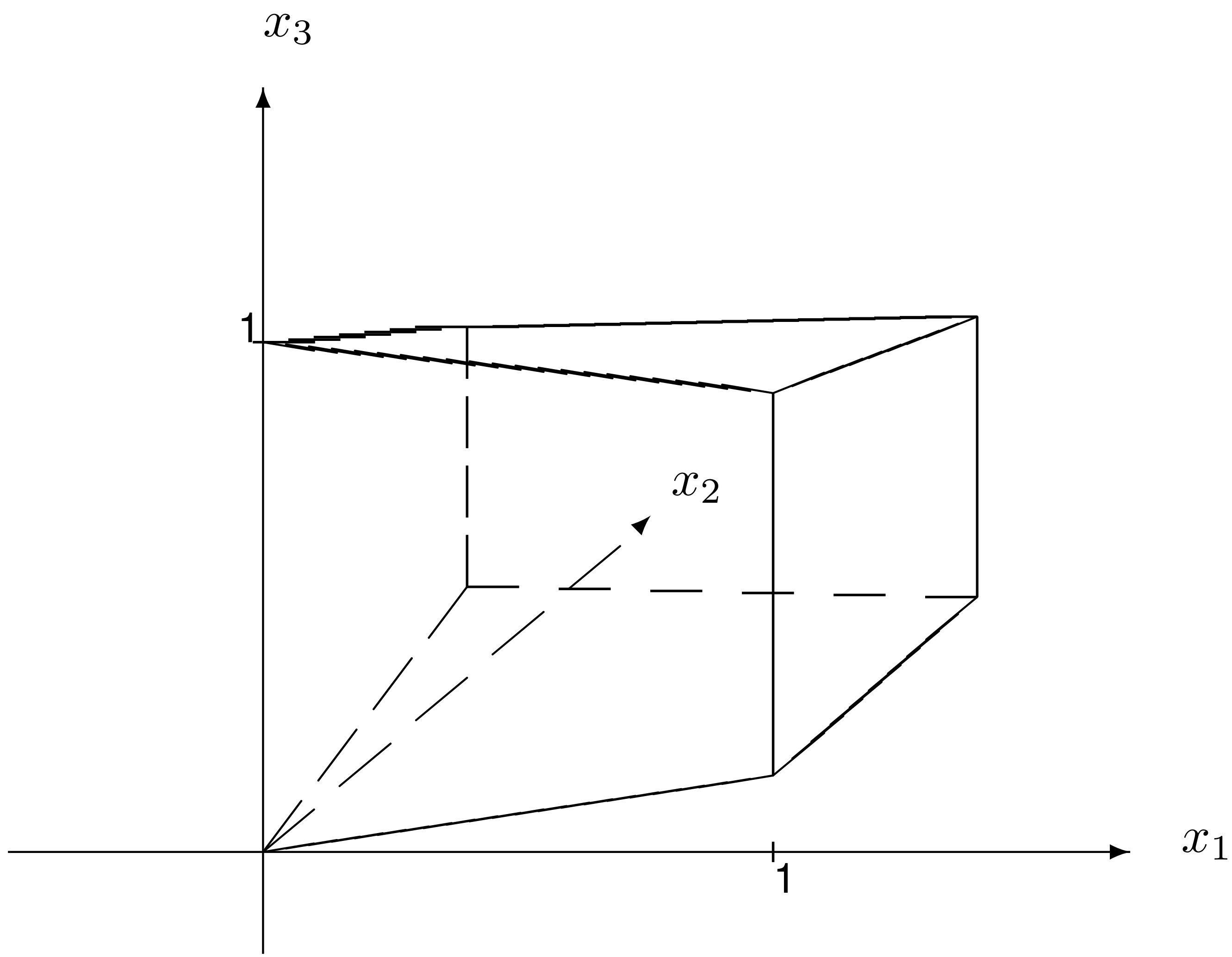
$$\begin{array}{lll} \max & x_n & \\ \textbf{subject to} & x_1 \geq 0 & \\ & x_1 \leq 1 & \\ & x_j \geq \epsilon\, x_{j-1} & j = 2, \ldots, n \\ & x_j \leq 1 - \epsilon\, x_{j-1} & j = 2, \ldots, n \end{array}$$

where $0 < \epsilon < 0.5$. This presentation of the problem emphasizes the idea (see the figures below) that the feasible region of the problem is a perturbation of the n-cube.

In the case of $n = 2$ and $\epsilon = 1/4$, the feasible region of the linear program above looks like



For the case where $n = 3$, the feasible region of the problem above looks like



The formulation above does not immediately reveal the standard form representation of the problem. Instead, we consider a different one, namely

$$\begin{array}{llll} \max & \displaystyle\sum_{j=1}^{n} 10^{n-j} x_j & & \\ \textbf{subject to} & \displaystyle 2\sum_{j=1}^{i-1} 10^{i-j} x_j + x_i & \leq 100^{i-1} & i = 1, \ldots, n \\ & x_j & \geq 0 & j = 1, \ldots, n \end{array}$$

The problem above[a] also be used is easily cast as a linear program in standard form. Unfortunately, it is less apparent how to exhibit the relationship between its feasible region and a perturbation of the unit cube.

[a] It should be noted that there is no need to express this problem in terms of powers of 10. Using any constant $C > 1$ would yield the same effect (an exponential number of pivot steps).

Example

$$\begin{array}{lrcrcrcr}
\max & 100x_1 & + & 10x_2 & + & x_3 & & \\
\textbf{subject to} & x_1 & & & & & \leq & 1 \\
& 20x_1 & + & x_2 & & & \leq & 100 \\
& 200x_1 & + & 20x_2 & + & x_3 & \leq & 10,000
\end{array}$$

In this case, we have three constraints and three variables (along with their nonnegativity constraints). After adding slack variables, we get a problem in standard form. The system has $m = 3$ equations and $n = 6$ nonnegative variables. In tableau form, the problem is

T^0

$-z$	x_1	x_2	x_3	x_4	x_5	x_6	1
1	100	10	1	0	0	0	0
0	1	0	0	1	0	0	1
0	20	1	0	0	1	0	100
0	200	20	1	0	0	1	10,000
				●	●	●	

The bullets below the tableau indicate the columns that are basic.

T^1

$-z$	x_1	x_2	x_3	x_4	x_5	x_6	1
1	0	10	1	–100	0	0	–100
0	1	0	0	1	0	0	1
0	0	1	0	–20	1	0	80
0	0	20	1	–200	0	1	9,800
	●				●	●	

T^2

$-z$	x_1	x_2	x_3	x_4	x_5	x_6	1
1	0	0	1	100	–10	0	–900
0	1	0	0	1	0	0	1
0	0	1	0	–20	1	0	80
0	0	0	1	200	–20	1	8,200
	●	●				●	

T^3

$-z$	x_1	x_2	x_3	x_4	x_5	x_6	1
1	–100	0	1	0	–10	0	–1,000
0	1	0	0	1	0	0	1
0	20	1	0	0	1	0	100
0	–200	0	1	0	–20	1	8,000
		●		●		●	

T^4

$-z$	x_1	x_2	x_3	x_4	x_5	x_6	1
1	100	0	0	0	10	–1	–9,000
0	1	0	0	1	0	0	1
0	20	1	0	0	1	0	100
0	–200	0	1	0	–20	1	8,000
		●	●	●			

T^5

$-z$	x_1	x_2	x_3	x_4	x_5	x_6	1
1	0	0	0	–100	10	–1	–9,100
0	1	0	0	1	0	0	1
0	0	1	0	–20	1	0	80
0	0	0	1	200	–20	1	8,200
	●	●	●				

T^6

$-z$	x_1	x_2	x_3	x_4	x_5	x_6	1
1	0	–10	0	100	0	–1	–9,900
0	1	0	0	1	0	0	1
0	0	1	0	–20	1	0	80
0	0	20	1	–200	0	1	9,800

T^7

$-z$	x_1	x_2	x_3	x_4	x_5	x_6	1
1	–100	–10	0	0	0	–1	–10,000
0	1	0	0	1	0	0	1
0	20	1	0	0	1	0	100
0	200	20	1	0	0	1	10,000
			●	●	●		

$$(x_1, x_2, x_3, x_4, x_5, x_6) = (0, 0, 10^4, 1, 10^2, 0)$$

is optimal and that the objective function value is $10,000$.

Along the way, we made $2^3 - 1 = 7$ pivot steps. The objective function made a strict increase with each change of basis.

Remark. The instance of the linear program (1) in which $n = 3$ leads to $2^3 - 1$ pivot steps when the greedy rule is used to select the pivot column. The general problem of the class (1) takes $2^n - 1$ pivot steps. To get an idea of how bad this can be, consider the case where $n = 50$. Now $2^{50} - 1 \approx 10^{15}$. In a year with 365 days, there are approximately 3×10^7 seconds. If a computer were running continuously and performing T iterations of the Simplex Algorithm per second, it would take approximately

$$\frac{10^{15}}{3T \times 10^7} = \frac{1}{3T} \times 10^8 \quad \textbf{years}$$

to solve the problem using the Simplex Algorithm with the greedy pivot selection rule.

An interesting connection

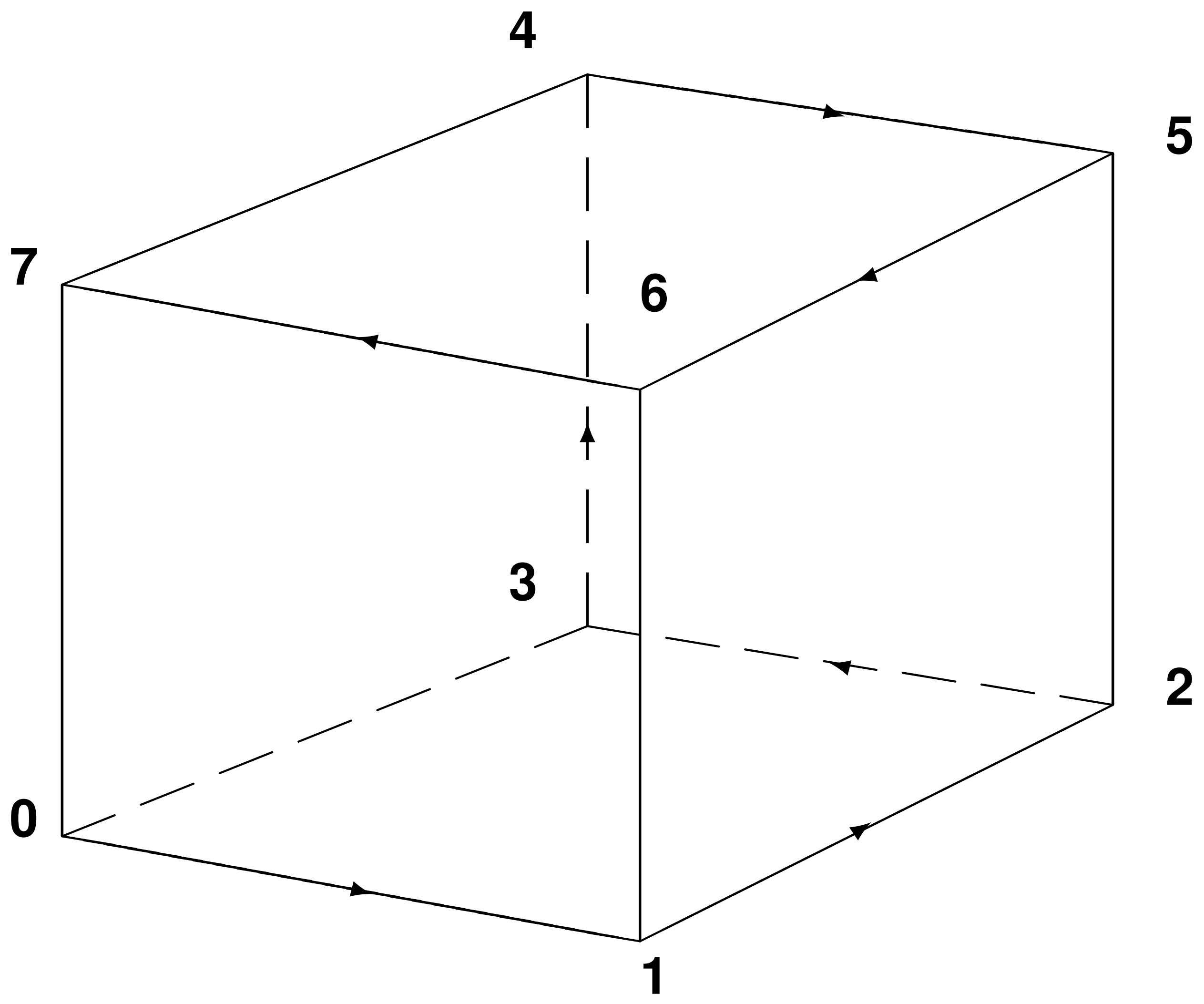
Consider the eight vectors $v^k = (v_1^k, v_2^k, v_3^k)$ where $k = 0, 1, \ldots, 7$ and

$$v_j^k = \begin{cases} 1 & \textbf{if } x_j \textbf{ is basic in tableau } k \\ 0 & \textbf{otherwise} \end{cases}$$

Looking at the eight tableaus $\mathrm{T}^0, \mathrm{T}^1, \ldots, \mathrm{T}^7$, we see that

$$\begin{aligned} v^0 &= (0,0,0) & v^4 &= (0,1,1) \\ v^1 &= (1,0,0) & v^5 &= (1,1,1) \\ v^2 &= (1,1,0) & v^6 &= (1,0,1) \\ v^3 &= (0,1,0) & v^7 &= (0,0,1) \end{aligned}$$

Now suppose we regard these vectors as the coordinates of the vertices of the 3-cube $[\,0, 1]$.


4
5
7
6
3
2
0
1

The figure above illustrates the fact that the sequence of vectors v^k corresponds to a path on the edges of the 3-cube. The path visits each vertex of the cube once and only once. Such a path is said to be Hamiltonian.

Queations on Homeworks

Problem 1. Show that the following problem is unbounded.

$$\begin{aligned} \max \quad & x_1 + x_2 \\ s.t. \quad & x_1 - x_2 - x_3 = 1 \\ & -x_1 + x_2 + 2x_3 \geq 1 \\ & x_1, x_2, x_3 \geq 0 \end{aligned}$$

Solution: It is clear that the primal problem is feasible since $x = (3; 0; 2)$ is a feasible solution. Its dual problem is

$$\begin{aligned} \min \quad & y_1 + y_2 \\ s.t. \quad & y_1 - y_2 \geq 1 \\ & -y_1 + y_2 \geq 1 \\ & -y_1 + 2y_2 \geq 0 \\ & y_1 \text{ free}, y_2 \leq 0. \end{aligned}$$

Clearly, the dual problem is infeasible. By the LP-duality theorem, the primal problem is unbounded.

Problem 2. Let $\Omega \subseteq R^n$ be a nonempty convex set and $\bar{\Omega}$ be its closure. Show that

$$\textbf{int}\Omega = \textbf{int}\bar{\Omega} \quad \textbf{and} \quad \partial\Omega = \partial\bar{\Omega}.$$

Solution: Clearly, $\textbf{int}\Omega \subseteq \textbf{int}\bar{\Omega}$. We will prove that $\textbf{int}\bar{\Omega} \subseteq \textbf{int}\Omega$.

Let $x \in \textbf{int}\bar{\Omega}$. If $\textbf{int}\Omega \neq \emptyset$, take $y \in \textbf{int}\Omega$, then there exists $\delta > 0$ such that

$$z = x + \delta(x - y) \in \bar{\Omega} \quad \Rightarrow \quad x = \frac{1}{1+\delta} z + \frac{\delta}{1+\delta y} \in \textbf{int}\Omega.$$

If $\textbf{int}\Omega = \emptyset$, then $\bar{\Omega} = \partial\Omega$ and $\textbf{int}\bar{\Omega} = \emptyset$. If not, there exists $\delta > 0$ such that $B_\delta(x) \subset \partial\Omega$. Take $y \in B_\delta(x)$, there is a sequence $\{y^k\} \subset \Omega$ with $y^k \to y$. Since Ω is convex set, there exists $\varepsilon > 0$ such that $B_\varepsilon(y^k) \subset \Omega$ for a fixed large k. This is a contradiction.

$$\bar{\Omega} = \textbf{int}\Omega \cup \partial\Omega = \textbf{int}\bar{\Omega} \cup \partial\bar{\Omega} \quad \Rightarrow \quad \partial\Omega = \partial\bar{\Omega}.$$

提 $y \in \text{int}\,\Omega$

How to Linearize the Abs Function I

$$\begin{array}{ll} \min & \sum_i \left|\sum_j a_{ij}x_j - b_i\right| \\ \textbf{s.t.} & 0 \leq x_j \leq 1, \forall j. \end{array}$$

最小一乘法

For each of the industry codes, the model will determine a probability which indicates the likelihood that a transaction was personal.

- Let x_j be such a probability that a transaction is personal for industry code j.
- Let a_{ij} be the transaction amount for account i and industry code j.
- Let b_i be the amount paid by personal remit for account i.
- $\sum_j a_{ij}x_j$ is the expected personal expenses for account i.
- Wed like to choose x_j such that $\sum_j a_{ij}x_j$ matches b_i for all i.

Model Example

The model will determine the probability that a transaction from each industry code is personal in such a manner which will minimize the sum of the absolute errors between predicted personal remittances and actual personal remittances.

Each Column represents an Industry Code

Personal Remittances

				...					
Account	1	2	3	...	n			Actual	
1	\$156	\$0	\$87		\$25			\$200	
2	\$200	\$25	\$0		\$0			\$195	
...	\$0	\$134	\$35		\$60			\$210	

Value of transactions in period

To deal with the abs function, we introduce auxiliary variables y_i. Let

$$|z_i| = y_i, \quad i = 1, \ldots, m.$$

引入辅助变量

Relax it to linear inequalities:

$$-y_i \leq z_i \leq y_i, \quad i = 1, \ldots, m.$$

If the sum of y_is is minimized, the equality must hold.

$$\begin{array}{ll} \min & \sum_i y_i \\ \textbf{s.t.} & -y_i \leq \sum_j a_{ij} x_j - b_i \leq y_i \quad \forall i, \\ & 0 \leq x_j \leq 1, \forall j. \end{array}$$

This is an LP problem:

$$\begin{array}{ll} \min & e^T y \\ \textbf{s.t.} & -y \leq Ax - b \leq y, \\ & 0 \leq x \leq e. \end{array}$$

还可以考虑A变化的情况

或者考虑x的0范数或者1范数有一定的限制

How to Linearize the Abs Function II

Introduce auxiliary variables u_i and v_i:

$$z_i = u_i - v_i, \quad u_i \geq 0, \quad \mathbf{v}_i \geq 0.$$

Relax it to linear inequalities:

$$\min |z_i| \Leftrightarrow \min u_i + v_i.$$

If the sum is minimized, the equality must hold.

$$\begin{array}{ll} \min & e^T(u+v) \\ \textbf{s.t.} & Ax - b = u - v, \\ & 0 \leq x \leq e, \quad u \geq 0, v \geq 0. \end{array}$$

得到等式约束

This is an LP problem.