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** $\mathbf{y}$. 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** $\mathbf{y}$, and it is the hyperplane containing the origin and $\mathbf{x}$ **and** $\mathbf{y}$. When $\alpha \geq 0$ and $\beta \geq 0$, such z is called a conic combination.

锥组合

Convex Set

凸集的边界为凸曲线

- 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 a set Ω is the intersection of all convex sets containing Ω
- Any intersection of convex sets is convex
- A point in a convex set is an extreme point (极点) if and only if it cannot be represented as a convex combination of two distinct points in the set.
- A set is polyhedral if and only if it has finite number of extreme points.

多面锥

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

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

$$= \{ Ax ; x \geq 0 \}$$

可行集合为凸集！

Polyhedral Convex Cones

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

$$C = \{\mathbf{x} : \; A\mathbf{x} \leq \mathbf{0}\} \quad \textbf{or} \quad \{\mathbf{x} : \; \mathbf{x} = A\mathbf{y}, \mathbf{y} \geq \mathbf{0}\}$$

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

Carathéodory's Theorem

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

Theorem 1 *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 $\mathbf{a}_1, \ldots, \mathbf{a}_n$.

极大无关组

可以用后面αβ的选取证明

Carathéodory's Theorem

凸包

Theorem 2 (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*

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

于是有等式： $\mathrm{co}(\Omega) = \{\Omega$中至多$n+1$个点的凸组合$\}$

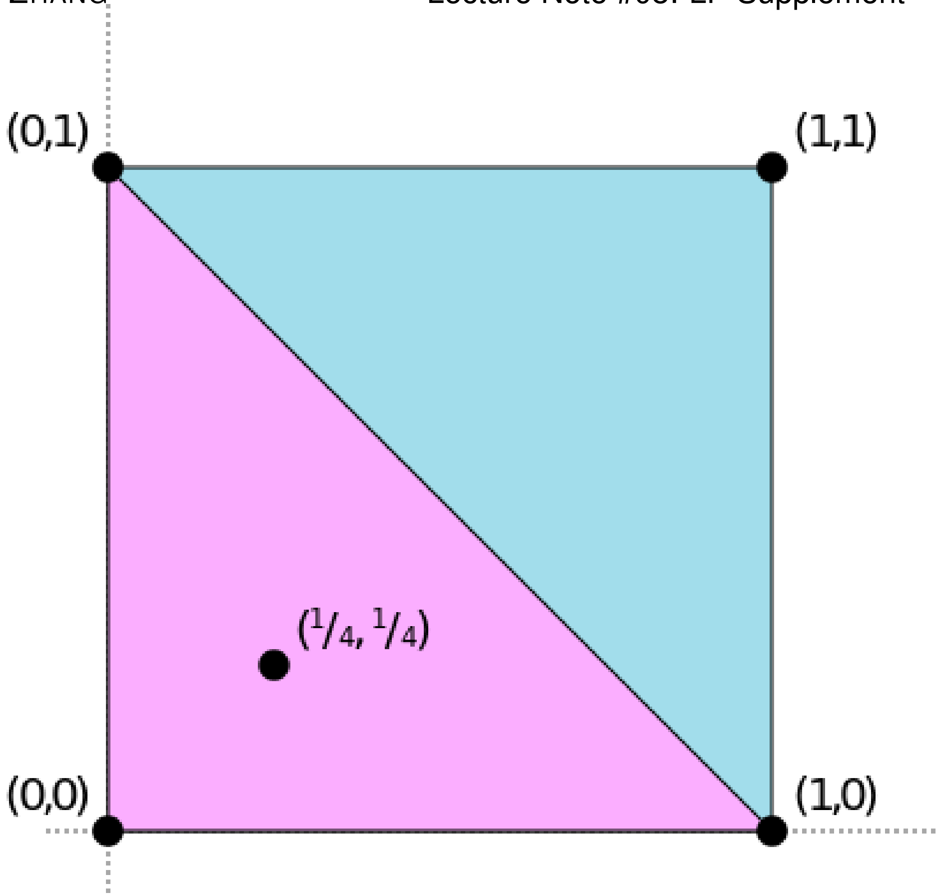


Figure 1: The convex hull of $\{(0,0),(0,1),(1,0),(1,1)\}$ is a square in $\mathcal{R}^2$

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.

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

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

由于线性相关

Let

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

这里只保证新的αi不小于0，不过这样就可以有小于1

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 $\alpha_i' = 0$

于是一个方向完成

For $\forall \mathbf{x}^1, \ldots, \mathbf{x}^m \in \Omega$, let

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

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

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

写成m-1个与一个的凸组合

Let S be the set of all convex combinations of all finite subsets of points, then S is convex.

$C = \{A\mathbf{x} :\ \mathbf{x} \geq \mathbf{0}\}$ 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 Carathéodorys theorem, 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

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

半正定锥被矩阵作用后

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*

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

有一个充要条件

$\exists\ \bar{x} \in C$

s.t. $(b-\bar{x})^T(x-\bar{x}) \leq 0$

$\forall\ x \in C$

由对偶 $\begin{cases} \min & 0 \cdot x \\ \text{s.t.} & Ax = b \\ & x \geq 0 \end{cases}$ 与 $\begin{cases} \max & b^T y \\ \text{s.t.} & A^T y \leq 0 \end{cases}$

于是 $b^T y \leq 0$
（若原题可行）

Farkas' Lemma

进而可考虑对偶问题的解来看
原问题是否可行

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 $\text{cone}(A_{.1}, \ldots, A_{.n})$.

$b^T y$ 这个平面

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 $\mathbf{c}^T\mathbf{x} \underset{<0\text{即可}}{=} -1$, is called a (dual) infeasibility certificate for the system $A^T\mathbf{y} \le \mathbf{c}$.

可以用对偶的角度来说，对偶的证明依赖于Farkas引理

Concave Function

- f concave function iff for $0 \leq \alpha \leq 1$,

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

  If strict inequality holds whenever $\mathbf{x} \neq \mathbf{y}$, then f is said to be strictly concave.

- The level set of concave function f

水平集

$$L(z) = \{\mathbf{x} : \quad f(\mathbf{x}) \geq z\}$$

  is a convex set. The converse is not true; e.g., $f(x) = x^3$.

由此易见 $\{x : x^T A x + b^T x + c \leq 0\}$ 为凸集

Epigraph

上方图

- Let f be a real-valued function on $S \subset \mathcal{R}^n$. The epigraph of f is the set

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

- Let $S \subset \mathcal{R}^n$ be a nonempty convex set. Then $f : S \to \mathcal{R}$ is concave function iff **epi** f is a convex subset of $\mathcal{R}^{n+1}$.

可根据上方图的性质将凸函数分为close convex
与proper convex

Theorems on concave functions

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

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

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

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

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

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

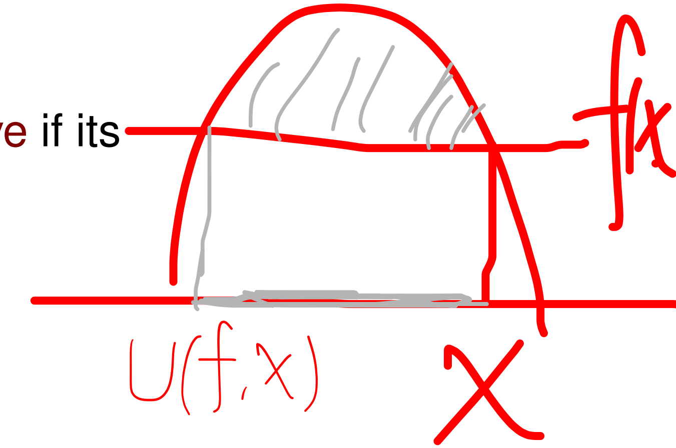
is concave iff $a + b \le 1$.

Quasi-Concave Function

For a function f on $S \subset \mathcal{R}^n$ and a point $\mathbf{x} \in S$, the upper-contour set of f at $\mathbf{x}$ consists of all points $\mathbf{y}$ with f values that are greater than or equal to $f(\mathbf{x})$:

$$U(f;\mathbf{x}) = \{\mathbf{y} \in S | f(\mathbf{y}) \geq f(\mathbf{x})\}.$$

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



易见拟凹的向上水平集都在x的左或者右侧
(不能两边的值都比x大)

进而有

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 if and only if for any $\mathbf{x}, \mathbf{y} \in C$ and any $\alpha \in (0, 1)$,*

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

In other words, the value of f at any interior of an interval is greater than or equal to the smaller one of its values at the two ends.

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,

$$\begin{array}{lll} (LP) & \textbf{minimize} & \mathbf{c}^T\mathbf{x} \\ & \textbf{subject to} & A\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0}, \end{array}$$

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

The dual problem can be written as:

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

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

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.

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

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

is called the primal optimal face, and

$$\{\mathbf{y} : \; c_{Z^*} - A_{Z^*}^T\mathbf{y} > 0, \; c_{P^*} - A_{P^*}^T\mathbf{y} = 0\}$$

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

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

Then,

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

That is, the following system of inequalities is infeasible

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

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 $\mathbf{x} \ge \mathbf{0}, \alpha \ge 0$.

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, $\mathbf{x}' \ge \mathbf{0}$ with $x_i' > 0$.

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).

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

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

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

The dual problem is

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

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:

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

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.

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.

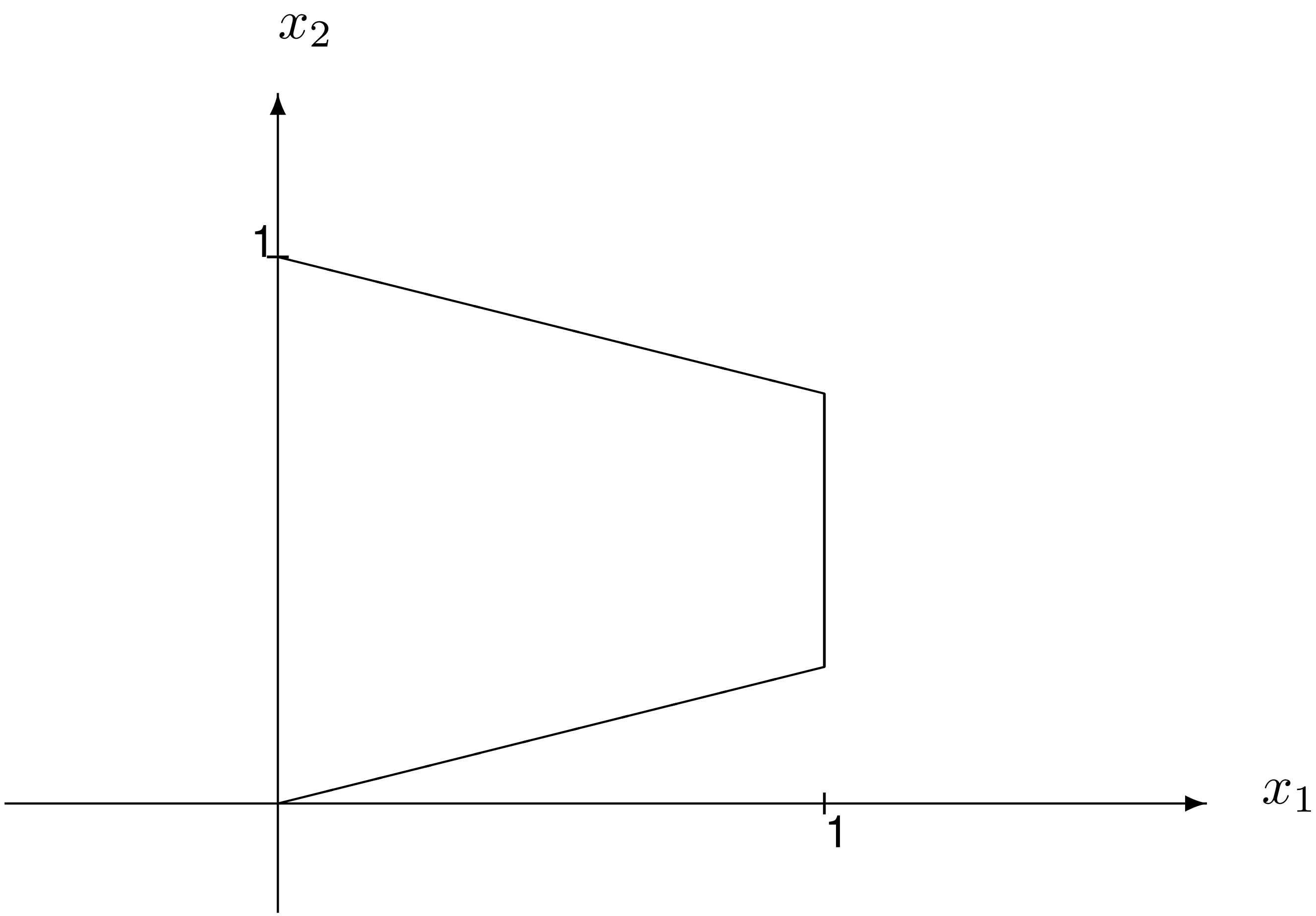
Klee and Minty Example

Consider

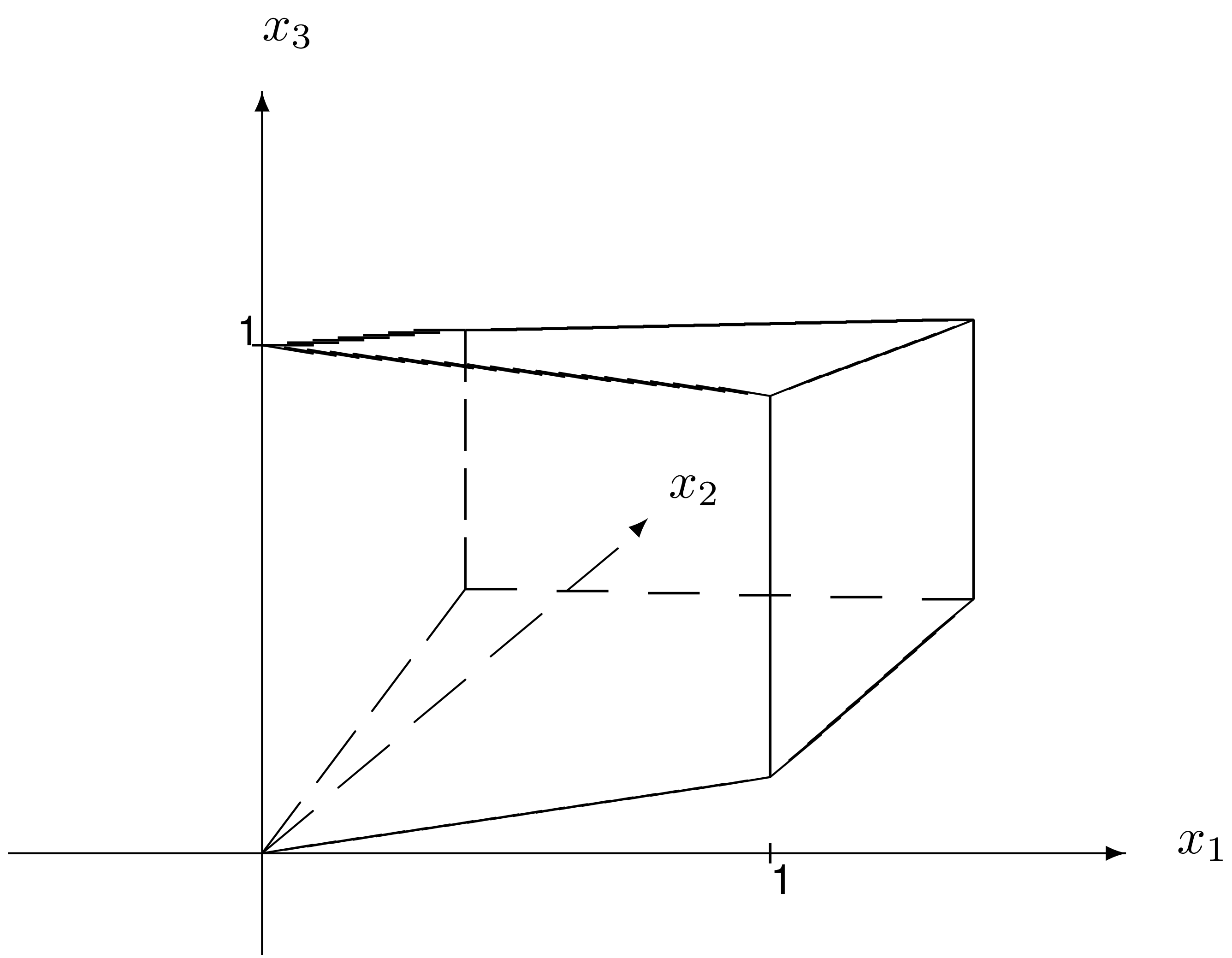
$$
\begin{array}{lll}
\max & x_n & \\
\textbf{subject to} & x_1 \geq 0 & \\
 & x_1 \leq 1 & \\
 & x_j \geq \epsilon\, x_{j-1} & j = 2, \ldots, n \\
 & x_j \leq 1 - \epsilon\, x_{j-1} & j = 2, \ldots, n
\end{array}
$$

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
T^0	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.

	$-z$	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
T^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

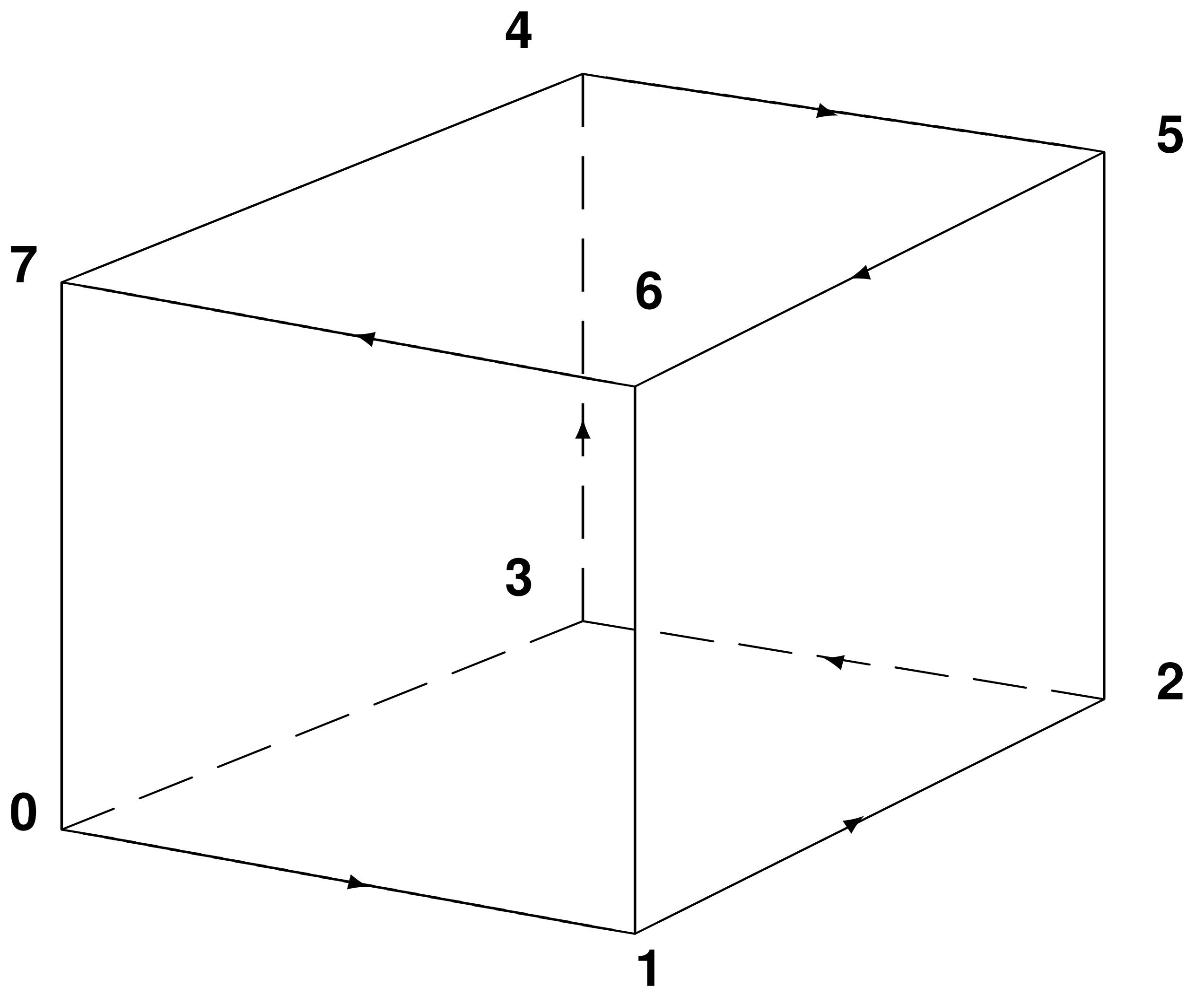
Consider the eight vectors $v^k = (v_1^k, v_2^k, v_3^k)$ where $k = 0, 1, \ldots, 7$ and

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

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

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

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.