Convex Analyses, Duality and Solution Rank

Yinyu Ye
Department of Management Science and Engineering
Stanford University

http://www.stanford.edu/~yyye

Separating hyperplane theorem

The most important theorem about the convex set is the following separating hyperplane theorem (Figure 1).

Theorem 1 *(Separating hyperplane theorem) Let C be a convex set and let b be a point exterior to the* closure *of C. Then there is a vector y such that*

$$y \bullet b > \sup_{x \in C} y \bullet x.$$

Geometric interpretation of the Theorem

The geometric interpretation of the theorem is that, given a convex set C and a point b outside of C, there exists a hyperplane strictly seperating b and C.

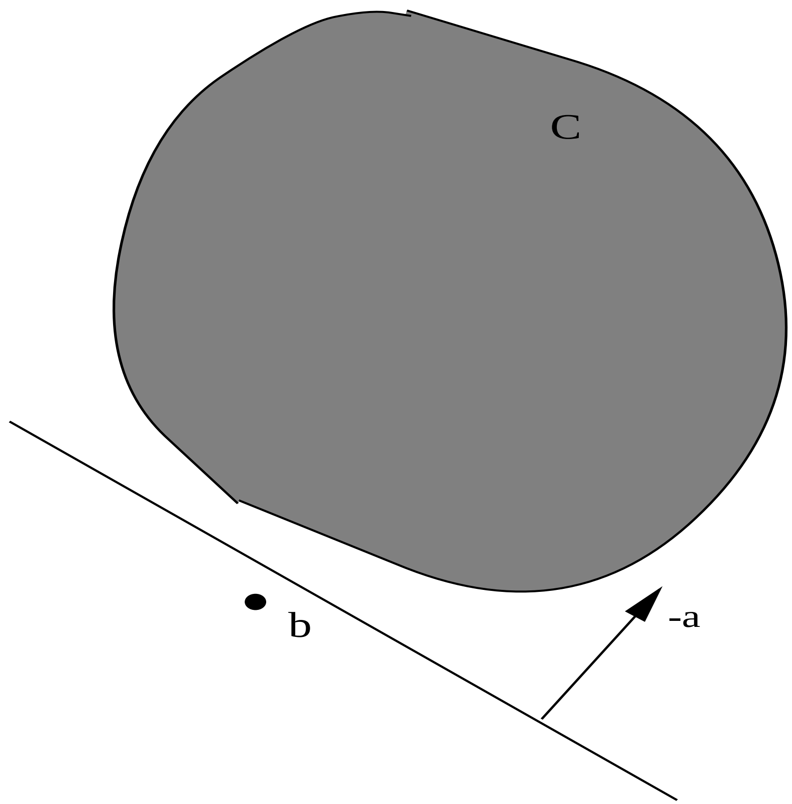


Figure 1: Illustration of the separating hyperplane theorem; an exterior point b is separated by a hyperplane from a convex set C.

System of Linear Inequalities

Solve for $\mathbf{x} \in \mathcal{R}^n$ from:

$$\begin{array}{rcl} \mathbf{a}_1\mathbf{x} & \leq & b_1 \\ \mathbf{a}_2\mathbf{x} & \leq & b_2 \\ \ldots & . & . \\ \mathbf{a}_m\mathbf{x} & \leq & b_m \end{array} \quad \Rightarrow A\mathbf{x} \leq \mathbf{b}$$

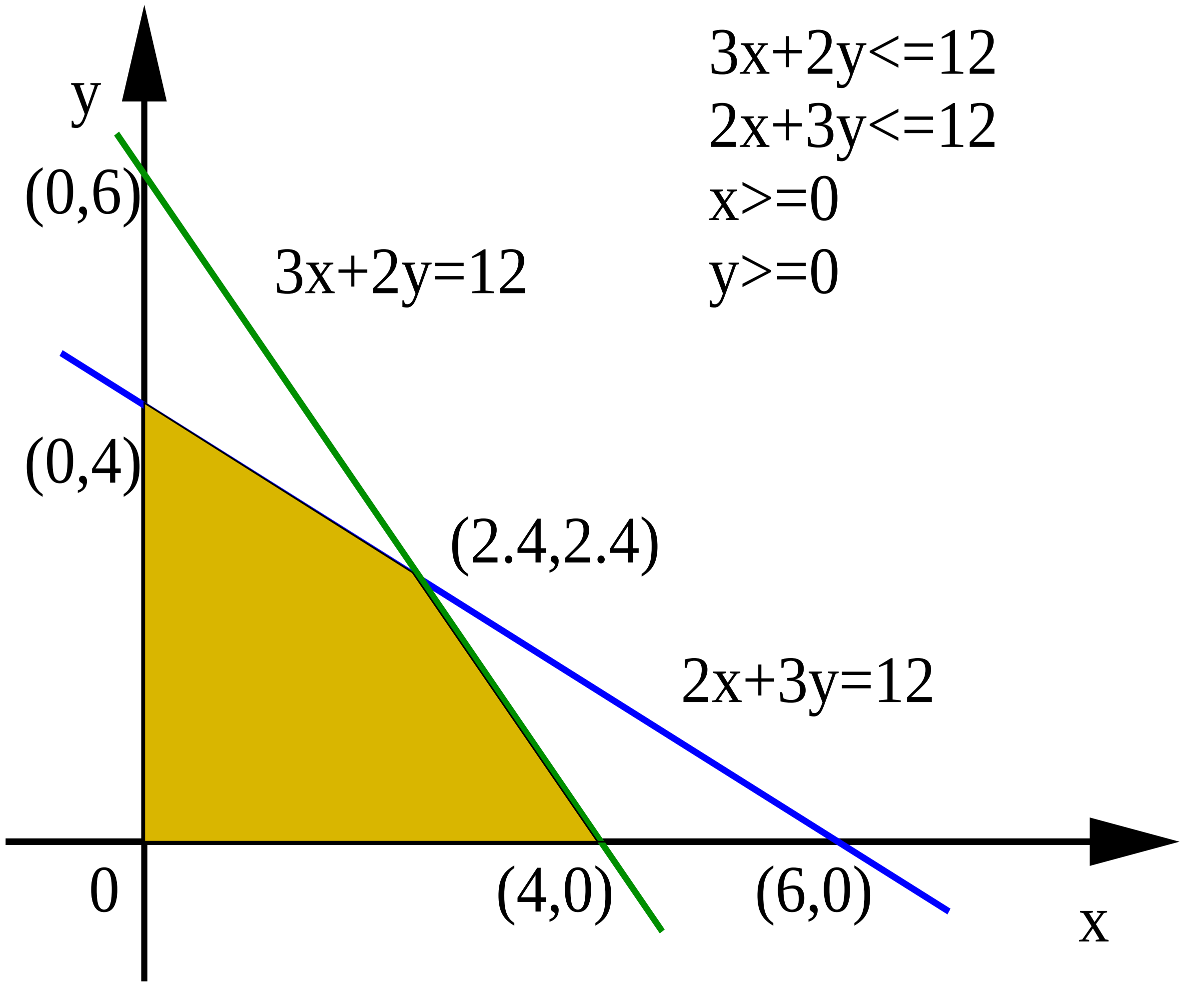


Figure 2: System of Linear Inequalities

Farkas Lemma

The following results are Farkas' lemma and its variants.

Theorem 2 *(Farkas' lemma) Let* $A \in \mathcal{R}^{m \times n}$ *and* $b \in \mathcal{R}^m$*. Then, the system* $\{x : \ Ax = b, \ x \geq 0\}$ *has a feasible solution* x *if and only if that* $A^T y \leq 0$ *and* $b^T y > 0$ *has no feasible solution.*

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

Alternative Systems

$$Ax = b, \quad x \geq 0.$$

$$A^T y \leq 0, \quad b^T y = 1 (> 0)$$

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

System of Linear Matrix Inequalities

Find $X \in \mathcal{M}^n$ such that

$$A_i \bullet X = b_i, \quad i = 1, ..., m, \quad X \succeq 0,$$

where

$$A \bullet B = \sum_{i,j} a_{ij} b_{ij} = \text{trace} A^T B,$$

and $X \succeq 0$ means that X is positive semi-definite.

Alternative Systems for SDP?

$$A_i \bullet X = b_i, \quad i = 1, ..., m, \quad X \succeq 0.$$

$$\sum_i^m y_i A_i \preceq 0, \quad b^T y = 1$$

SDP Example where Farkars' lemma failed

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, A_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

and

$$b = \begin{pmatrix} 0 \\ 2 \end{pmatrix}.$$

Farkas Lemma for SDP

Theorem 3 *(Farkas' lemma in SDP) Let* $A_i \in \mathcal{M}^n$*,* $i = 1, ..., m$*, have rank* m *(i.e.,* $\sum_i^m y_i A_i = 0$ *implies* $y = 0$*) and* $b \in \mathcal{R}^m$*. Then, there exists a symmetric matrix* $X \succ 0$ *with*

$$A_i \bullet X = b_i, \quad i = 1, ..., m,$$

if and only if $\sum_i^m y_i A_i \preceq 0$ *and* $\sum_i^m y_i A_i \neq 0$ *and* $b^T y \geq 0$ *has no feasible solution.*

Note the difference between the LP and SDP.

Alternative Systems for SDP

$$A_i \bullet X = b_i, \quad i = 1, ..., m, \quad X \succ 0.$$

$$\sum_i^m y_i A_i \preceq 0, \; \sum_i^m y_i A_i \neq 0, \quad b^T y \geq 0$$

Linear Programming (LP)

The standard form linear programming problem is given below, which we will use throughout this book:

$$(LP) \quad \text{minimize} \quad c \bullet x$$
$$\text{subject to} \quad a_i \bullet x = b_i, \; i = 1, ..., m, \; x \geq 0,$$

where $c, \; a_i \in \mathcal{R}^n$.

With every (LP), another linear program, called the dual (LD), is the following problem:

$$(LD) \quad \text{maximize} \quad b \bullet y$$
$$\text{subject to} \quad \sum_i^n y_i a_i + s = c, \; s \geq 0,$$

where $y \in \mathcal{R}^m$ and $s \in \mathcal{R}^n$, called dual slack variable.

LP in Compact Form

The standard form linear programming problem is given below, which we will use throughout this book:

$$
\begin{array}{lll}
(LP) & \text{minimize} & c^T x \\
 & \text{subject to} & Ax = b, \; x \geq 0.
\end{array}
$$

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

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

LP Duality Theories

Theorem 4 *(Weak duality theorem) Let $\mathcal{F}_p$ and $\mathcal{F}_d$ be non-empty. Then,*

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

$$c^T x - b^T y = c^T x - (Ax)^T y = x^T(c - A^T y) = x^T s \geq 0.$$

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

From this we have important results in the following.

Theorem 5 *(Strong duality theorem) Let primal feasible region $\mathcal{F}_p$ and dual feasible region $\mathcal{F}_d$ be non-empty. Then, x^* is optimal for (LP) if and only if the following conditions hold:*

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

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

iii) $c^T x^* = b^T y^*$.

Given $\mathcal{F}_p$ and $\mathcal{F}_d$ being non-empty, we like to prove that there is $x^* \in \mathcal{F}_p$ and $(y^*, s^*) \in \mathcal{F}_d$ such that $c^T x^* \leq b^T y^*$, or to prove that

$$Ax = b,\ A^T y \leq c,\ c^T x - b^T y \leq 0,\ x \geq 0$$

is feasible.

Proof of Strong Duality Theorem

Suppose not, from Farkas' lemma, we must have an infeasibility certificate (x', τ, y') such that

$$Ax' - b\tau = 0, \; A^T y' - c\tau \leq 0, \; (x', \tau) \geq 0$$

and

$$b^T y' - c^T x' = 1$$

If $\tau > 0$, then we have

$$0 \geq (-y')^T (Ax' - b\tau) + x'^T (A^T y' - c\tau) = \tau(b^T y' - c^T x') = \tau$$

which is a contradiction.

If $\tau = 0$, then the weak duality theorem also leads to a contradiction.

Theorem 6 *(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 an equal objective value, then these are optimal for both. The converse is also true; there is no “gap.”

Optimality Conditions

$$\left\{ (x, y, s) \in (\mathcal{R}^n_+, \mathcal{R}^m, \mathcal{R}^n_+) : \begin{array}{rcl} c^T x - b^T y & = & 0 \\ Ax & = & b \\ -A^T y - s & = & -c \end{array} \right\},$$

which is a system of linear inequalities and equations. Now it is easy to verify whether or not a pair (x, y, s) is optimal.

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

If $x^T s = 0$, then we say x and s are complementary to each other.

Since both x and s are nonnegative, $x^T s = 0$ implies that $x_j s_j = 0$ for all $j = 1, \ldots, n$.

$$\begin{aligned} Xs &= 0 \\ Ax &= b \\ -A^T y - s &= -c. \end{aligned}$$

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

Theorem 7 *(Strict complementarity theorem) If (LP) and (LD) both have feasible solutions then both problems have a pair of strictly complementary solutions $x^* \geq 0$ and $s^* \geq 0$ meaning*

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

Moreover, the supports

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

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. $\{x : \; A_{P^*} x_{P^*} = b, \; x_{P^*} \geq 0, \; x_{Z^*} = 0\}$ is called the primal optimal face, and $\{y : \; c_{Z^*} - A_{Z^*}^T y \geq 0, \; c_{P^*} - A_{P^*}^T y = 0\}$ is called the dual optimal face.

An Example

Consider the primal problem:

$$
\begin{array}{lllll}
\text{minimize} & x_1 & +x_2 & +1.5 \cdot x_3 & \\
\text{subject to} & x_1 & & + \quad x_3 & = 1 \\
& & x_2 & + \quad x_3 & = 1 \\
& x_1, & x_2, & x_3 & \geq 0;
\end{array}
$$

The dual problem is

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

$$P^* = \{3\} \quad \text{and} \quad Z^* = \{1, 2\}$$

Face and Extreme Point

Let P be a polyhedron in $\mathcal{R}^n$, F is a face of P if and only if there is a vector c for which F is the set of points attaining max $\{c^T x : \; x \in P\}$ provided this maximum is finite.

A polyhedron has only finite many faces; each face is a nonempty polyhedron.

A vector $x \in P$ is an extreme point or a vertex of P if x is not a convex combination of more than one distinct points.

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

A 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

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

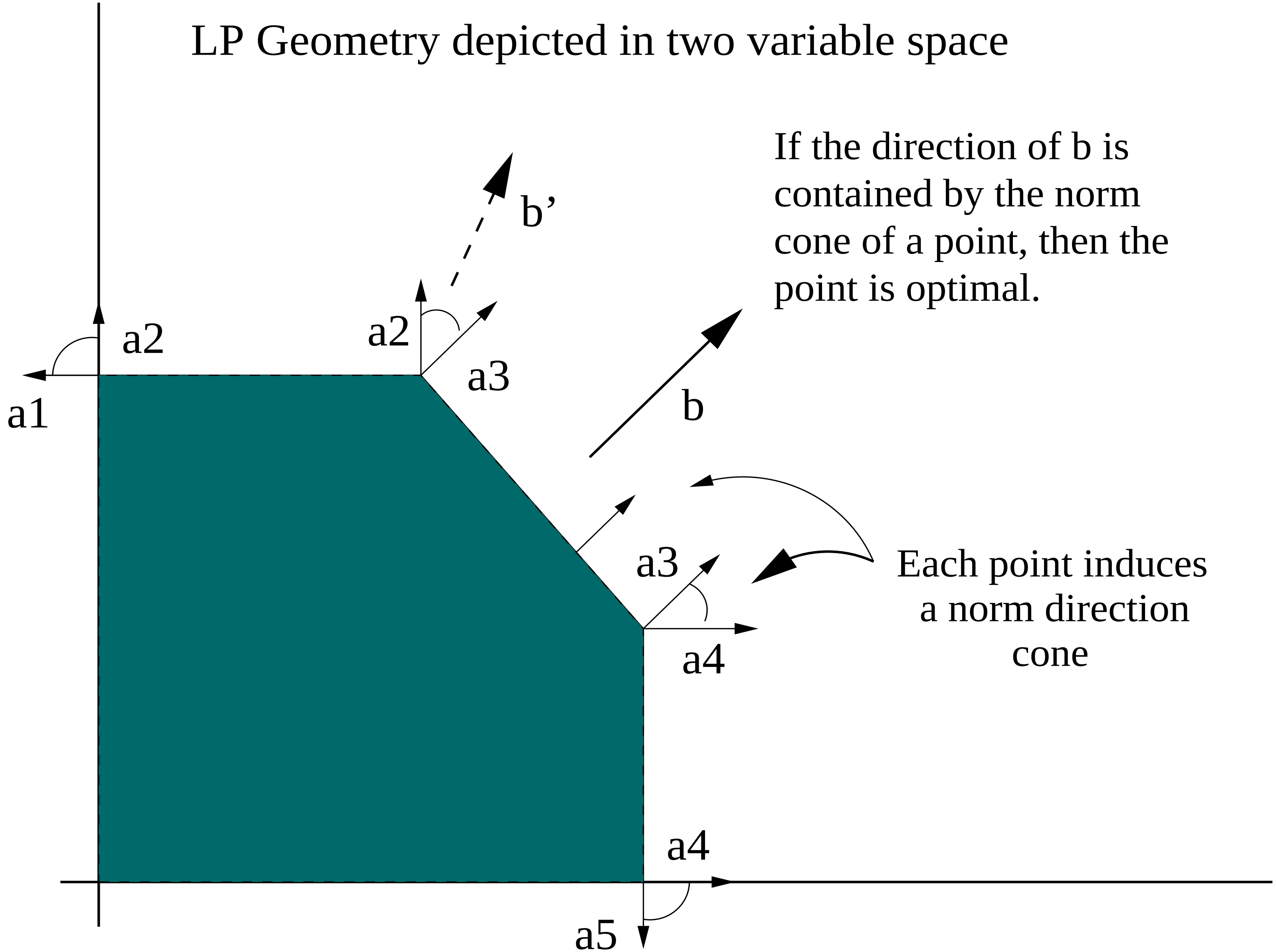
then x is called an optimal basic solution and A_B an optimal basis.

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

Geometry vs Algebraic

Theorem 8 *Consider the polyhedron in the standard LP form. Then, a basic feasible solution and an extreme point are equivalent; the formal is algebraic and the latter is geometric. Moreover, two neighboring extreme points are represented by two basic solutions whose bases differ by exactly one column.*

LP Geometry depicted in two variable space



Theorem 9 *(LP fundamental theorem) Given (LP) and (LD) 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.*

The above theorem reduces the task of solving a linear program to that searching over basic feasible solutions. By expanding upon this result, the simplex method, a finite search procedure, is derived. The simplex method is to proceed from one basic feasible solution (an extreme point of the feasible region) to an adjacent one, in such a way as to continuously decrease the value of the objective function until a basic minimizer is reached. In contrast, interior-point algorithms will move in the interior of the feasible region and reduce the value of the objective function, hoping to by-pass many extreme points on the boundary of the region.

Semidefinite Programming (SDP)

The standard form semidefinite programming problem is given below, which we will use throughout this book:

$$
\begin{array}{lll}
(SDP) & \text{inf} & C \bullet X \\
 & \text{subject to} & A_i \bullet X = b_i, i = 1, 2, ..., m, \;\; X \succeq 0,
\end{array}
$$

where $C, \; A_i \in \mathcal{M}^n$.

The dual problem to (SDP) can be written as:

$$
\begin{array}{lll}
(SDD) & \text{sup} & b \bullet y \\
 & \text{subject to} & \sum_i^m y_i A_i + S = C, \;\; S \succeq 0,
\end{array}
$$

where $y \in \mathcal{R}^m$ and $S \in \mathcal{M}^n$.

SDP in Compact Form

$$
\begin{array}{lll}
(SDP) & \text{inf} & C^T X \\
 & \text{subject to} & \mathcal{A}X = b, \; X \succeq 0.
\end{array}
$$

The dual problem to (SDP) can be written as:

$$
\begin{array}{lll}
(SDD) & \text{sup} & b^T y \\
 & \text{subject to} & \mathcal{A}^T y + S = C, \; S \succeq 0.
\end{array}
$$

Duality Theorems for SDP

Theorem 10 *(Weak duality theorem in SDP) Let $\mathcal{F}_p$ and $\mathcal{F}_d$, the feasible sets for the primal and dual, be non-empty. Then,*

$$C \bullet X \geq b^T y \quad \textit{where} \quad X \in \mathcal{F}_p,\ (y, S) \in \mathcal{F}_d.$$

The weak duality theorem is identical to that of (LP) and (LD).

Corollary 1 *(Strong duality theorem in SDP) Let $\mathcal{F}_p$ and $\mathcal{F}_d$ be non-empty and at least one of them has an interior. Then, X is optimal for (PS) if and only if the following conditions hold:*

i) $X \in \mathcal{F}_p$;

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

iii) $C \bullet X = b^T y$ *or* $X \bullet S = 0$.

SDP Example with a Duality Gap

$$C = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, A_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, A_2 = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

and

$$b = \begin{pmatrix} 0 \\ 10 \end{pmatrix}.$$

Optimality Conditions for SDP

$$
\begin{array}{rcl}
C \bullet X - b^T y & = & 0 \\
\mathcal{A}X & = & b \\
-\mathcal{A}^T y - S & = & -C \\
X, S & \succeq & 0
\end{array}, \qquad (1)
$$

$$
\begin{array}{rcl}
XS & = & 0 \\
\mathcal{A}X & = & b \\
-\mathcal{A}^T y - S & = & -C \\
X, S & \succeq & 0
\end{array} \qquad (2)
$$

Production Problem I

$$\max\ \mathbf{p}^T\mathbf{x} \quad \text{s.t.} \quad A\mathbf{x} \le \mathbf{r}, \quad \mathbf{x} \ge \mathbf{0}$$

where

- $\mathbf{p}$: profit margin vector
- A: resources consumption rate matrix
- $\mathbf{r}$: available resource vector
- $\mathbf{x}$: production level decision vector

Production Problem II: Liquidation Pricing

- $\mathbf{y}$: the fair price vector
- $A^T\mathbf{y} \geq \mathbf{p}$: competitiveness
- $\mathbf{y} \geq 0$: positivity
- $\min\ \mathbf{r}^T\mathbf{y}$: minimize the total liquidation cost

$$
Primal: \quad \begin{array}{lllll}
\text{maximize} & x_1 & +2x_2 & & \\
\text{subject to} & x_1 & & \leq 1 \\
& & x_2 & \leq 1 \\
& x_1 & +x_2 & \leq 1.5 \\
& x_1, & x_2 & \geq 0.
\end{array}
$$

$$
Dual: \quad \begin{array}{llllll}
\text{minimize} & y_1 & +y_2 & +1.5y_3 & \\
\text{subject to} & y_1 & & +y_3 & \geq 1 \\
& & y_2 & +y_3 & \geq 2 \\
& y_1, & y_2, & y_3 & \geq 0.
\end{array}
$$

Optimal Value Function I

For fixed matrix A and right-hand-side vector $\mathbf{b}$, the optimal value is a function of objective coefficient vector $\mathbf{c}$:

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

Homework 6: Show that $f_c(\mathbf{c})$ is a concave function in $\mathbf{c}$.

The Max-Flow problem

Given a directed graph with nodes $1, ..., n$ and edges A, where node 1 is called source and node n is called the sink, and each edge (i, j) has a flow rate capacity u_{ij}. The Max-Flow problem is to find the largest possible flow rate from source to sink.

Let x_{ij} be the flow rate from node i to node j. Then the problem can be formulated as

$$
\begin{array}{ll}
\text{maximize} & x_{n1} \\
\text{subject to} & \sum_{j:(i,j)\in A} x_{ij} - \sum_{j:(j,i)\in A} x_{ji} - x_{n1} = 0,\ \forall i = 1, \\
& \sum_{j:(i,j)\in A} x_{ij} - \sum_{j:(j,i)\in A} x_{ji} = 0, \forall i = 2, ..., n-1, \\
& \sum_{j:(i,j)\in A} x_{ij} - \sum_{j:(j,i)\in A} x_{ji} + x_{n1} = 0, \forall i = n, \\
& 0 \le x_{ij} \le u_{ij},\ \forall (i,j) \in A.
\end{array}
$$

The dual of the Max-Flow problem: Min-Cut

$$
\begin{array}{ll}
\text{minimize} & \sum_{(i,j)\in A} u_{ij} z_{ij} \\
\text{subject to} & y_i - y_j + z_{ij} \geq 0, \ \forall (i,j) \in A, \\
& -y_1 + y_n = 1, \\
& z_{ij} \geq 0, \ \forall (i,j) \in A.
\end{array}
$$

This problem is called the Min-Cut problem.

Two-Person Zero-Sum Game

Let P be the payoff matrix of a two-person, ”column” and ”row”, zero-sum game.

$$P = \begin{pmatrix} +3 & -1 & -4 \\ -3 & +1 & +4 \end{pmatrix}$$

Players usually use randomized strategies in such a game. A randomized strategy is a vector of probabilities, each associated with a particular decision.

Nash Equilibrium

In a Nash Equilibrium, if your (column) strategy is a pure strategy (one where you always play a single action), the expected payout for the (dominating) action that you are playing should be greater than or equal to the expected payout for any other action. If you are playing a randomized strategy, the expected payout for each action included in your strategy should be the same (if one were lower, you won't want to ever choose that action) and these payouts should be greater than or equal to the actions that aren't part of your strategy.

LP formulation of Nash Equilibrium

”Column” strategy:

$$\begin{array}{ll} \max & v \\ s.t. & ve \le Px \\ & e^T x = 1 \\ & x \ge 0. \end{array}$$

”Row” strategy:

$$\begin{array}{ll} \min & u \\ s.t. & ue \ge P^T y \\ & e^T y = 1 \\ & y \ge 0. \end{array}$$

They are dual to each other.

Multi-Firm LP Alliance I

Consider a finite set I of firms each of whom has operations that have representations as linear programs. Suppose the linear program representing the operations of firm i in I entails choosing an n-column vector $\mathbf{x} \geq \mathbf{0}$ of activity levels that maximize the firm's profit

$$\mathbf{c}^T\mathbf{x}$$

subject to the constraint that its consumption $A\mathbf{x}$ of resources minorizes its available resource vector $\mathbf{b}^i$, that is,

$$A\mathbf{x} \leq \mathbf{b}^i.$$

Multi-Firm LP Alliance II

An alliance is a subset of the firms. If an alliance S pools its resource vectors, the linear program that S faces is that of choosing an n-column vector $\mathbf{x} \geq \mathbf{0}$ that maximizes the profit $\mathbf{c}^T\mathbf{x}$ that S earns subject to its resource constraint

$$A\mathbf{x} \leq \mathbf{b}^S = \sum_{i \in S} \mathbf{b}^i.$$

Let V^S be the resulting maximum profit of S. The grand alliance is the set I of all firms.

$$\begin{array}{rlll} V^S := & \max & \mathbf{c}^T\mathbf{x} & \\ & \text{s.t.} & A\mathbf{x} & \leq \sum_{i \in S} \mathbf{b}^i, \\ & & \mathbf{x} & \geq \mathbf{0}, \end{array}$$

Multi-Firm LP Alliance III: Core

Core is the set of payment vector $\mathbf{z} = (z_1,\ ...,\ z_{|I|})$ to each company such that

$$\sum_{i \in I} z_i = V^I$$

and

$$\sum_{i \in S} z_i \geq V^S,\ \forall S \subset I.$$

Theorem 11 *For each optimal dual price vector for the linear program of the grand alliance, allocating each firm the value of its resource vector at those prices yields a profit allocation vector in the core.*

Combinatorial Auction Pricing: The dual of the model

$$
\begin{array}{lrl}
\min & \mathbf{q}^T\mathbf{y} & \\
\text{s.t.} & A^T\mathbf{p} + \mathbf{y} & \geq \pi, \\
 & \mathbf{e}^T\mathbf{p} & = w, \\
 & (\mathbf{p}, \mathbf{y}) & \geq 0,
\end{array}
$$

where $\mathbf{p}$ represent state prices.

Homework 7: Prove the following conditions hold:

$x_j > 0$	$\mathbf{a}_j^T\mathbf{p} + y_j = \pi_j$ so that $\mathbf{a}_j^T\mathbf{p} \leq \pi_j$
$0 < x_j < q_j$	$y_j = 0$ so that $\mathbf{a}_j^T\mathbf{p} = \pi_j$
$x_j = 0$	$y_j = 0$ so that $\mathbf{a}_j^T\mathbf{p} \geq \pi_j$

Carathéodory's theorem

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

Theorem 12 *Let convex polyhedral cone*

$$C = \textit{cone}(a_1, ..., a_n) = \{Ax = \sum_j a_j x_j : x \geq 0\}.$$

and $b \in C$*. Then,* $b \in \textit{cone}(a_{i_1}, ..., a_{i_m})$ *for some linearly independent vectors* a_{i_1},...,a_{i_m} *chosen from* a_1,...,a_n*.*

If $Ax = b,\ x \geq 0$ has a feasible solution, it has a feasible solution where it has at most m nonzero entries.

Rounding to a low-rank solution for LP

1. Start at any feasible solution x^0 and, without loss of generality, assume $x^0 > 0$, and let $k = 0$ and $A^0 = A$.
2. Find any $A^k d = 0, \; d \neq 0$, and let $x^{k+1} = x^k + \alpha d$ where α is chosen such as $x^{k+1} \geq 0$ and at least one of x^{k+1} equals 0.
3. Eliminate the the variable(s) in x^{k+1} and column(s) in A^k corresponding to $x_j^{k+1} = 0$, and let the new matrix be A^{k+1}.
4. Return to step 2.

The Eigenvalue Decomposition

Let X be a positive semidefinite matrix of rank r, Then, the eigenvalue decomposition of X

$$X = \sum_{j=1}^{r} \lambda_j v_j v_j^T,$$

such that for all j,

$$v_j^T v_i = 0 \quad \|v_j\|^2 = 1, \quad j \neq i, j = 1, ...r.$$

Bound on Rank of SDP Solutions

Theorem 13 *Let* X^* *be an minimizer of (SDP). By solving a linear program (in strongly polinomial time) we can compute another minimizer of (SDP) whose rank* r *satisfying*

$$\frac{r(r+1)}{2} \leq m.$$

Barvinok (2001), Pataki, Alfakih/Wolkowicz (1999).

Proof of the theorem

(This proof is written by Mehdi Mohseni and myself based on our discussions in the class.)

Consider SDP in the standard form:

$$
\begin{aligned}
z^* := \quad & \text{Minimize} && C \bullet X \\
& \text{Subject to} && A_i \bullet X = b_i, \ i = 1, ..., m \\
& && X \succeq 0.
\end{aligned}
\tag{3}
$$

Theorem 14 *Let X^* be a minimizer of (3). Then we can compute a minimizer of (3) whose rank r satisfies $r(r+1)/2 \le m$ in polinomial time.*

Proof

If the rank of X^*, r, satisfies the inequality, then we need do nothing. Thus, we assume $r(r+1)/2 > m$, and let

$$VV^T = X^*, \quad V \in R^{n\times r}.$$

Then consider

$$\begin{array}{ll} \text{Minimize} & V^TCV \bullet U \\ \text{Subject to} & V^TA_iV \bullet U = b_i, \; i = 1, ..., m \\ & U \succeq 0. \end{array} \tag{4}$$

Note that V^TCV, V^TA_iVs and U are $r \times r$ symmetric matrices and

$$V^TCV \bullet I = C \bullet VV^T = C \bullet X^* = z^*.$$

Moreover, for any feasible solution of (4) one can construct a feasible solution for (3) using

$$X(U) = VUV^T \quad \text{and} \quad C \bullet X(U) = V^T CV \bullet U. \tag{5}$$

Thus, the minimal value of (4) is also z^*, and $U = I$ is a minimizer of (4).

Now we show that any feasible solution U to (4) is a minimizer for (4); thereby $X(U)$ of (5) is a minimizer for (3). Consider the dual of (4)

$$\begin{array}{rll} z^* := & \text{Maximize} & b^T y = \sum_{i=1}^m b_i y_i \\ & & \\ & \text{Subject to} & V^T CV \succeq \sum_{i=1}^m y_i V^T A_i V^T. \end{array} \tag{6}$$

Let y^* be a dual maximizer. Since $U = I$ is an interior optimizer for the primal, the strong duality condition holds, i.e.,

$$I \bullet (V^T CV - \sum_{i=1}^m y_i^* V^T A_i V^T) = 0$$

so that we have

$$V^TCV - \sum_{i=1}^{m} y_i^* V^T A_i V^T = 0.$$

Then, any feasible solution of (4) satisfies the strong duality condition so that it must be also optimal.

Consider the system of homogeneous linear equations

$$V^T A_i V \bullet W = 0, \; i = 1, ..., m$$

where W is a $r \times r$ symmetric matrices (does not need to be definite). This system has $r(r+1)/2$ real number variables and m equations. Thus, as long as $r(r+1)/2 > m$, we must be able to find a symmetric matrix $W \neq 0$ to satisfy all m equations. Without loss of generality, let W be either indefinite or negative semidefinite (if it is positive semidefinite, we take $-W$ as W), that is, W has at least one negative eigenvalue, and consider

$$U(\alpha) = I + \alpha W.$$

Choosing $\alpha^* = 1/|\bar{\lambda}|$ where $\bar{\lambda}$ is the least eigenvalue of W, we have

$$U(\alpha^*) \succeq 0$$

and it has at least one 0 eigenvalue or $\text{rank}(U(\alpha^*)) < r$, and

$$V^T A_i V \bullet U(\alpha^*) = V^T A_i V \bullet (I + \alpha^* W) = V^T A_i V \bullet I = b_i, \; i = 1, ..., m.$$

That is, $U(\alpha^*)$ is a feasible and so it is an optimal solution for (4). Then,

$$X(U(\alpha^*)) = VU(\alpha^*)V^T$$

is a new minimizer for (3), and $\text{rank}(X(U(\alpha^*))) < r$.

This process can be repeated till the system of homogemeous linear equations has only all zero solution, which is necessarily given by $r(r+1)/2 \leq m$. Thus, we must be able to find an SDP solution whose rank satisfies the inequality. The total number of such reduction steps is bounded by $n-1$ and each step uses no more than $O(m^2n)$ arithmetic operations. Therefore, the total number of arithmetic operations is a polynomial in m and n, i.e., in (strongly) polynomial time given the least eigenvalue of W.