Midterm Review

Yinyu Ye
Department of Management Science and Engineering
Stanford University
Stanford, CA 94305, U.S.A.

http://www.stanford.edu/˜yyye

(LY, Chapter 1- 6, Appendices, and Lecture Notes 1-9)

Separating hyperplane theorem

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

Theorem 1 *(Separating hyperplane theorem) Let $C \subset \mathcal{E}$, where $\mathcal{E}$ is a metric space either $\mathcal{R}^n$ or $\mathcal{M}^n$, be a* closed *convex set and let $\mathbf{y}$ be a point exterior to C. Then there is a point $\mathbf{a} \in \mathcal{E}$ such that*

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

Carathéodory's theorem

Theorem 2 *Given matrix* $A \in \mathcal{R}^{m \times n}$ *and* $\mathbf{b} \in \mathcal{R}^m$. *If*

$$\{\mathbf{x} : \ A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$$

is feasible. Then

$$\mathbf{b} = A_B \mathbf{x}_B, \ \mathbf{x}_B \geq \mathbf{0}$$

where columns in A_B *are linearly independent chosen from* $\mathbf{a}_1$,...,$\mathbf{a}_n$ *of* A.

The theorem is true whether or not A has a full row rank or not.

Farkas' Lemma

The following results are Farkas' lemma and its variants.

Theorem 3 *Let $A \in \mathcal{R}^{m \times n}$ and $\mathbf{b} \in \mathcal{R}^m$. Then, the system $\{\mathbf{x} : \ A\mathbf{x} = \mathbf{b}, \ \mathbf{x} \geq \mathbf{0}\}$ has a feasible solution $\mathbf{x}$ if and only if that system $\{\mathbf{y} : \ A^T\mathbf{y} \leq \mathbf{0}, \ \mathbf{b}^T\mathbf{y} > 0\}$ has no feasible solution $\mathbf{y}$.*

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

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

Example

Let $A = (1, 1)$ and $b = -1$. Then, $y = -1$ is an infeasibility certificate for $\{\mathbf{x} : \ A\mathbf{x} = b, \ \mathbf{x} \geq \mathbf{0}\}$.

Farkas' Lemma Variant

Theorem 4 *Let $A \in \mathcal{R}^{m \times n}$ and $\mathbf{c} \in \mathcal{R}^n$. Then, the system $\{\mathbf{y} : \ A^T\mathbf{y} \leq \mathbf{c}\}$ has a solution $\mathbf{y}$ if and only if that system $\{\mathbf{x} : \ A\mathbf{x} = \mathbf{0}, \ \mathbf{c}^T\mathbf{x} < 0, \ \mathbf{x} \geq \mathbf{0}\}$ has no feasible solution $\mathbf{x}$.*

Again, a vector $\mathbf{x} \geq \mathbf{0}$, with $A\mathbf{x} = \mathbf{0}$ and $\mathbf{c}^T\mathbf{x} < 0$, is called a infeasibility certificate for the system $\{\mathbf{y} : \ A^T\mathbf{y} \leq \mathbf{c}\}$.

example

Let $A = (1; -1)$ and $\mathbf{c} = (1; -2)$. Then, $\mathbf{x} = (1; 1)$ is an infeasibility certificate for $\{y : \ A^T y \leq \mathbf{c}\}$.

Farkas' Lemma for General Cones?

Given $\mathbf{a}_i$, $i = 1, ..., m$, and $\mathbf{b} \in \mathcal{R}^m$.

Then, the system $\{\mathbf{x} : \ \mathbf{a}_i \bullet \mathbf{x} = b_i, \ i = 1, ..., m, \ \mathbf{x} \in C\}$ has a feasible solution $\mathbf{x}$ if and only if that $-\sum_i^m y_i \mathbf{a}_i \in C^*$ and $\mathbf{b}^T \mathbf{y} > 0$ has no feasible solution $\mathbf{y}$?

It is necessary but not sufficient!

Let's write equations in a compact form:

$$\mathcal{A}\mathbf{x} = (\mathbf{a}_1 \bullet \mathbf{x}; ...; \mathbf{a}_m \bullet \mathbf{x}) \in \mathcal{R}^m$$

and

$$\mathcal{A}^T \mathbf{y} = \sum_i^m y_i \mathbf{a}_i.$$

Alternative Systems for General Cones?

Alternative System Pair I?:

$$\mathcal{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \in C,$$

and

$$-\mathcal{A}^T\mathbf{y} \in C^*, \quad \mathbf{b}^T\mathbf{y} = 1$$

Alternative System Pair II?:

$$\mathcal{A}\mathbf{x} = \mathbf{0}, \quad \mathbf{x} \in C, \quad \mathbf{c} \bullet \mathbf{x} = -1 (< 0)$$

and

$$\mathbf{c} - \mathcal{A}^T\mathbf{y} \in C^*$$

When Farkas' Lemma Holds for General Cones?

Let C be a closed convex cone in the rest of the course.

If there is $\mathbf{y}$ such that $-\mathcal{A}^T\mathbf{y} \in \operatorname{int} C^*$, then Alternative System Pair I is true:

$$\mathcal{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \in C,$$

and

$$-\mathcal{A}^T\mathbf{y} \in C^*, \quad \mathbf{b}^T\mathbf{y} = 1$$

And if there is $\mathbf{x}$ such that $\mathcal{A}\mathbf{x} = \mathbf{0},\ \mathbf{x} \in \operatorname{int} C$, then Alternative System Pair II is true:

$$\mathcal{A}\mathbf{x} = \mathbf{0}, \quad \mathbf{x} \in C, \quad \mathbf{c} \bullet \mathbf{x} = -1 (< 0)$$

and

$$\mathbf{c} - \mathcal{A}^T\mathbf{y} \in C^*$$

Duality Theory

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

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

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

The dual problem can be written as:

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

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

Duality Theory

Theorem 5 *(Weak duality theorem) Let primal feasible set $\mathcal{F}_p$ and dual feasible set $\mathcal{F}_d$ be non-empty. Then,*

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

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 6 *(Strong duality theorem) Let primal feasible set $\mathcal{F}_p$ and dual feasible set $\mathcal{F}_d$ be non-empty. Then, $\mathbf{x}^* \in \mathcal{F}_p$ is optimal for (LP) and $(\mathbf{y}^*, \mathbf{s}^*) \in \mathcal{F}_d$ is optimal for (LD) if and only if*

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

Theorem 7 *(Primal-Dual relation 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 either. 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 they are both optimal, respectively.

The converse is also true; there is zero duality “gap’ if they are optimal.

Duality and Complementarity Gaps

$\mathbf{x} \in \mathcal{F}_p$ is optimal for (LP) and $(\mathbf{y}, \mathbf{s}) \in \mathcal{F}_d$ is optimal for (LD)

$$\mathbf{x}^T\mathbf{s} = \mathbf{x}^T(\mathbf{c} - A^T\mathbf{y}) = \mathbf{c}^T\mathbf{x} - \mathbf{b}^T\mathbf{y} = 0$$

implies that $x_j s_j = 0$ for all $j = 1, ..., n$, since both $\mathbf{x}$ and $\mathbf{s}$ are nonnegative. This is called the complementarity gap.

$$\begin{array}{rcl} x_j s_j & = & 0, \ \forall j \\ A\mathbf{x} & = & \mathbf{b} \\ -A^T\mathbf{y} - \mathbf{s} & = & -\mathbf{c}. \end{array}$$

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

Theorem 8 *If both (LP) and (LD) are feasible, there exists a strictly complementary solution pair $\mathbf{x}$ and $(\mathbf{y}, \mathbf{s})$ such that*

$$x_j + s_j > 0, \ \forall j.$$

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

Conic LP

$$(CLP) \quad \text{minimize} \quad \mathbf{c} \bullet \mathbf{x}$$
$$\text{subject to} \quad \mathbf{a}_i \bullet \mathbf{x} = b_i, i = 1, 2, ..., m, \ \mathbf{x} \in C,$$

where C is a convex cone.

Linear Programming (LP): $\mathbf{c}, \mathbf{a}_i, \mathbf{x} \in \mathcal{R}^n$ and $C = \mathcal{R}^n_+$

Second-Order Cone Programming (SOCP): $\mathbf{c}, \mathbf{a}_i, \mathbf{x} \in \mathcal{R}^n$ and $C = SOC$

Semidefinite Programming (SDP): $\mathbf{c}, \mathbf{a}_i, \mathbf{x} \in \mathcal{S}^n$ and $C = \mathcal{S}^n_+$

Note that cone C can be a product of many (different) convex cones.

Dual of Conic LP

The dual problem to

$$(CLP) \quad \text{minimize} \quad \mathbf{c} \bullet \mathbf{x}$$
$$\text{subject to} \quad \mathbf{a}_i \bullet \mathbf{x} = b_i, i = 1, 2, ..., m, \ \mathbf{x} \in C.$$

is

$$(CLD) \quad \text{maximize} \quad \mathbf{b}^T \mathbf{y}$$
$$\text{subject to} \quad \sum_i^m y_i \mathbf{a}_i + \mathbf{s} = \mathbf{c}, \ \mathbf{s} \in C^*,$$

where $y \in \mathcal{R}^m$ are the dual variables, $\mathbf{s}$ is called the dual slack vector/matrix, and C^* is the dual cone of C.

Theorem 9 *(Weak duality theorem)*

$$\mathbf{c} \bullet \mathbf{x} - \mathbf{b}^T \mathbf{y} = \mathbf{x} \bullet \mathbf{s} \geq 0$$

for any feasible $\mathbf{x}$ *of (CLP) and* $(\mathbf{y}, \mathbf{s})$ *of (CLD).*

CLP Duality Theories

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

Corollary 1 *Let* $\mathbf{x}^* \in \mathcal{F}_p$ *and* $(\mathbf{y}^*, \mathbf{s}^*) \in \mathcal{F}_d$*. Then,* $\mathbf{c} \bullet x^* = \mathbf{b}^T \mathbf{y}^*$ *implies that* $\mathbf{x}^*$ *is optimal for (CLP) and* $(\mathbf{y}^*, \mathbf{s}^*)$ *is optimal for (CLD).*

Is the reverse also true? That is, given $\mathbf{x}^*$ optimal for (CLP), then there is $(\mathbf{y}^*, \mathbf{s}^*)$ feasible for (CLD) and $\mathbf{c} \bullet \mathbf{x}^* = \mathbf{b}^T \mathbf{y}^*$?

This is called the Strong Duality Theorem and it is “true” for LP, but it is “False” in general cases.

When Strong Duality Theorems Holds for CLP

Theorem 10 *(Strong duality theorem) Let $\mathcal{F}_p$ and $\mathcal{F}_d$ be non-empty and at least one of them has an interior. Then, $\mathbf{x}^*$ is optimal for (CLP) and $(\mathbf{y}^*, \mathbf{s}^*)$ is optimal for (CLD) if any only if*

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

There are cases that the duality gap tends to zero but the optimal solution is not attainable.

More Duality Theorem for CLP

Theorem 11 *(CLP duality theorem) If one of (CLP) or (CLD) is unbounded then the other has no feasible solution.*

If (CLP) and (CLD) are both feasible, then both have bounded optimal objective values and the optimal objective values may have a duality gap.

If one of (CLP) or (CLD) has a strictly or interior feasible solution and it has an optimal solution, then the other is feasible and has an optimal solution with the same optimal value.

Optimality Conditions for SDP

$$\begin{array}{rcl} \mathbf{c} \bullet X - \mathbf{b}^T \mathbf{y} & = & 0 \\ \mathcal{A}X & = & \mathbf{b} \\ -\mathcal{A}^T \mathbf{y} - S & = & -\mathbf{c} \\ X, S & \succeq & \mathbf{0} \end{array}, \tag{1}$$

$$\begin{array}{rcl} XS & = & \mathbf{0} \\ \mathcal{A}X & = & \mathbf{b} \\ -\mathcal{A}^T \mathbf{y} - S & = & -\mathbf{c} \\ X, S & \succeq & \mathbf{0} \end{array} \tag{2}$$

Rank of SDP Solutions

At any optimal solution pair (X^*, S^*)

$$\text{rank}(X^*) + \text{rank}(S^*) \le n.$$

If the equality holds, they are a strictly complementary solution pair.

There are optimal solutions of X^* and S^* such that the rank of X^* and the rank of S^* are minimal, respectively.

There are optimal solutions of X^* and S^* such that the rank of X^* and the rank of S^* are maximal, respectively.

In certain applications, we want a solution who has the max-rank or the min-rank. Or we like to prove that any solution must have a low rank, where one way to do is to show that the dual has a high rank solution...

SNL: System of Quadratic Equations

System of nonlinear equations for $\mathbf{x}_i \in R^d$:

$$\|\mathbf{x}_i - \mathbf{x}_j\| = d_{ij}, \; \forall\,(i,j) \in N_x, \; i < j,$$
$$\|\mathbf{a}_k - \mathbf{x}_j\| = d_{kj}, \; \forall\,(k,j) \in N_a,$$

where $\mathbf{a}_k$ are possible points whose locations are known, often called anchors.

One can equivalently represent it as

$$\|\mathbf{x}_i - \mathbf{x}_j\|^2 = d_{ij}^2, \; \forall\,(i,j) \in N_x, \; i < j,$$
$$\|\mathbf{a}_k - \mathbf{x}_j\|^2 = d_{kj}^2, \; \forall\,(k,j) \in N_a,$$

which becomes a system of multi-variable-quadratic equations.

Matrix Representation of SNL

Let $X = [\mathbf{x}_1\ \mathbf{x}_2\ ...\ \mathbf{x}_n]$ be the $2 \times n$ matrix that needs to be determined. Then

$$\|\mathbf{x}_i - \mathbf{x}_j\|^2 = (\mathbf{e}_i - \mathbf{e}_j)^T X^T X(\mathbf{e}_i - \mathbf{e}_j) \text{ and}$$
$$\|\mathbf{a}_k - \mathbf{x}_j\|^2 = (\mathbf{a}_k; -\mathbf{e}_j)^T [I\ \ X]^T [I\ \ X](\mathbf{a}_k; -\mathbf{e}_j),$$

where $\mathbf{e}_i$ is the vector with 1 at the ith position and zero everywhere else.

$$\begin{aligned}
(\mathbf{e}_i - \mathbf{e}_j)^T(\mathbf{e}_i - \mathbf{e}_j) \bullet Y &= d_{ij}^2,\ \forall\, i,j \in N_x,\ i < j,\\
(\mathbf{a}_k; -\mathbf{e}_j)^T(\mathbf{a}_k; -\mathbf{e}_j) \bullet \begin{pmatrix} I & X \\ X^T & Y \end{pmatrix} &= d_{kj}^2,\ \forall\, k,j \in N_a,\\
Y - X^T X &= \mathbf{0}.
\end{aligned}$$

where Y denotes the Gram matrix $X^T X$.

SDP Relaxation

Change

$$Y - X^T X = \mathbf{0}$$

to

$$Y - X^T X \succeq \mathbf{0}.$$

This matrix inequality is equivalent to

$$Z := \begin{pmatrix} I & X \\ X^T & Y \end{pmatrix} \succeq \mathbf{0}.$$

This is the semidefinite matrix cone, and the problem becomes an SDP Feasibility problem.

SDP Feasibility Standard Form

Find a symmetric matrix $Z \in \mathbf{R}^{(2+n)\times(2+n)}$ such that

$$\begin{aligned}
Z_{1:2,1:2} &= I \\
(\mathbf{0}; \mathbf{e}_i - \mathbf{e}_j)(\mathbf{0}; \mathbf{e}_i - \mathbf{e}_j)^T \bullet Z &= d_{ij}^2, \ \forall\, i, j \in N_x,\ i < j, \\
(\mathbf{a}_k; -\mathbf{e}_j)(\mathbf{a}_k; -\mathbf{e}_j)^T \bullet Z &= d_{kj}^2, \ \forall\, k, j \in N_a, \\
Z &\succeq \mathbf{0}.
\end{aligned}$$

This is semidefinite programming feasibility system (with a null objective).

When this relaxation is exact?

One case is that the single unknown point $\mathbf{x}_1$ is connected to three anchors $\mathbf{a}_k,\ k = 1, 2, 3$.

In general, if the rank of a feasible Z is 2, then it solves the original graph relaxation problem.

Duality Theorem for SNL

Theorem 12 *Let $\bar{Z}$ be a feasible solution for SDP and $\bar{U}$ be an optimal slack matrix of the dual. Then,*

1. *complementarity condition holds: $\bar{Z} \bullet \bar{U} = 0$ or $\bar{Z}\bar{U} = \mathbf{0}$;*
2. *$(\bar{Z}) + (\bar{U}) \leq 2 + n$;*
3. *$(\bar{Z}) \geq 2$ and $(\bar{U}) \leq n$.*

An immediate result from the theorem is the following:

Corollary 2 *If an optimal dual slack matrix has rank n, then every solution of the SDP has rank 2, that is, the SDP relaxation solves the original problem exactly.*

Basic Feasible Solution of LP

In the LP standard form when A has a full row-rank, select m linearly independent columns, denoted by the index set B, from A. Solving the m-dimension vector $\mathbf{x}_B$ from

$$A_B \mathbf{x}_B = \mathbf{b}$$

and setting the rest variables, denoted by $\mathbf{x}_N$, to zero, we obtain a solution $\mathbf{x}$ such that

$$A\mathbf{x} = \mathbf{b}.$$

Then, $\mathbf{x}$ is said to be a (primal) basic solution to (LP) with respect to the basis A_B. The entries in $\mathbf{x}_B$ are called basic variables. If a basic solution $\mathbf{x} \geq \mathbf{0}$, then $\mathbf{x}$ is called a basic feasible solution.

If one or more components in $\mathbf{x}_B$ have zero value, then the basic feasible solution $\mathbf{x}$ is said to be (primal) degenerate.

The basic feasible solution is an extreme point of the feasible region. In general (i.e., A has no full row-rank), any solution from $A_B \mathbf{x}_B = \mathbf{b},\ \mathbf{x}_B \geq \mathbf{0}$, where columns of A_B are linearly independent, is an extreme point of the feasible region.

A dual vector $\mathbf{y}$ satisfying

$$A_B^T \mathbf{y} = \mathbf{c}_B$$

is said to be the corresponding dual basic solution.

If the dual basic solution is also feasible, that is,

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

, then it is a dual Basic Feasible Solution.

If one or more slacks in $\mathbf{c}_N - A_N^T \mathbf{y}$ have zero value, that dual basic feasible solution $\mathbf{y}$ is said to be (dual) degenerate.

The dual basic feasible solution is an extreme point of the dual feasible region. In general (i.e., A has no full row-rank), any solution from $A_B^T \mathbf{y} = \mathbf{c}_B, \ \mathbf{c} - A^T \mathbf{y} \geq \mathbf{0}$, where rows of A_B are linearly independent, is an extreme point of the dual feasible region.

Theorem 13 *(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 or extreme solution ;*

ii) *if there is an optimal solution, there is an optimal basic feasible or extreme solution.*

The Simplex Algorithm

0. Initialize with a minimization problem in feasible canonical form with respect to a basic index set B. Let N denote the complementary index set.

1. Test for termination: Compute the dual solution and reduced gradient vector

$$\mathbf{y}^T = \mathbf{c}_B^T A_B^{-1} \quad \text{and} \quad \mathbf{r} = \mathbf{c} - A^T\mathbf{y}.$$

2. Check $r_e = \min_{j \in N}\{r_j\}$. If $r_e \geq 0$, stop. The solution is optimal. Otherwise determine whether the column of $\bar{A}_{.e}$ contains a positive entry. If not, the objective function is unbounded below. Terminate. Let x_e be the entering basic variable.

3. Determine the outgoing: execute the MRT to determine the outgoing variable x_o or declare the problem unbounded.

4. Update basis: update B and A_B and compute $\mathbf{x}_B = A_B^{-1}\mathbf{b}$; return to Step 1.

The Ellipsoid Method

The basic ideas of the ellipsoid method stem from research done in the nineteen sixties and seventies mainly in the Soviet Union (as it was then called) by others who preceded Khachiyan. The idea in a nutshell is to enclose the region of interest in each member of a sequence of ellipsoids whose size is decreasing, resembling the bisection method.

The significant contribution of Khachiyan was to demonstrate in two papers—published in 1979 and 1980—that under certain assumptions, the ellipsoid method constitutes a polynomially bounded algorithm for linear programming.

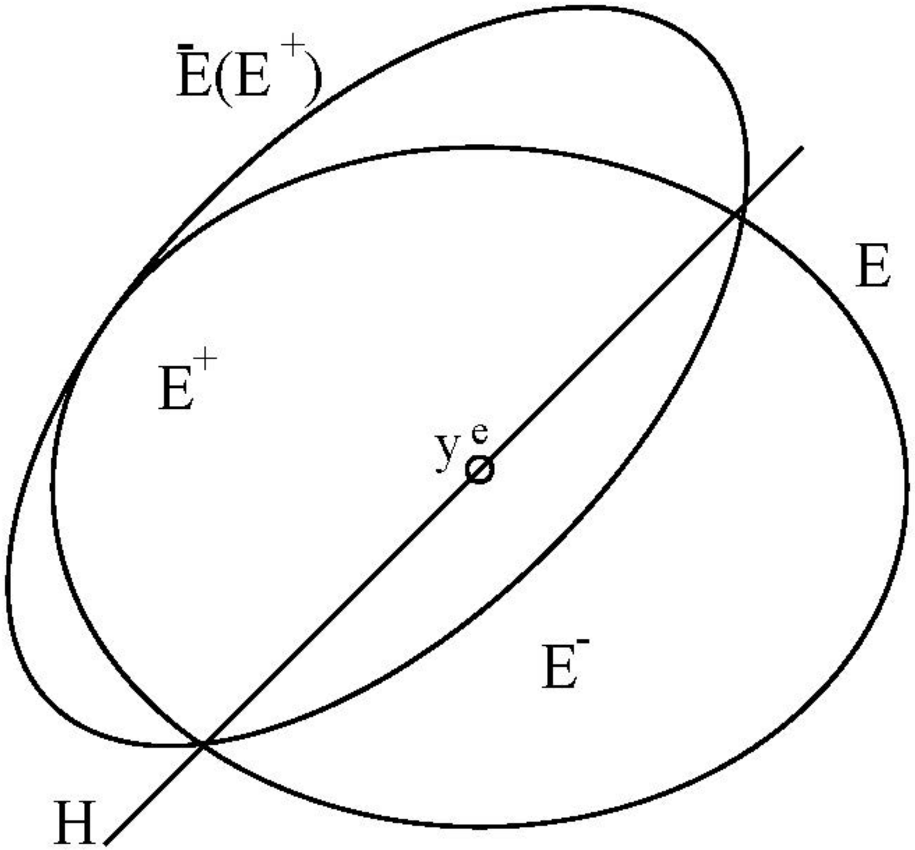

Figure 1: The least volume ellipsoid containing a half ellipsoid

Desired Theoretical Properties

- Separation Problem: Either decide the ellipsoid center $\mathbf{y}^c \in P$, where P is the target set, or find a separating hyperplane $\mathbf{a}$ such that $\mathbf{a}^T\mathbf{y} \leq \mathbf{a}^T\mathbf{y}^c$ for all $\mathbf{y} \in P$.
- Oracle to generate $\mathbf{a}$ without enumerating all hyperplanes.

Theorem 14 *If the separating (oracle) problem can be solved in polynomial time of m and $\log(R/r)$, then we can solve the standard linear programming problem whose running time is polynomial in m and $\log(R/r)$ that is independent of n, the number of inequality constraints.*

The Methodology Concept of Centers/Paths

Consider linear program

$$\begin{array}{ll} \text{maximize} & \mathbf{b}^T\mathbf{y} \\ \text{subject to} & A^T\mathbf{y} \le \mathbf{c}. \end{array}$$

Consider an objective level set

$$Y(z^0) := \{\mathbf{y} : \ A^T\mathbf{y} \le \mathbf{c},\ \mathbf{b}^T\mathbf{y} \ge z^0\},$$

and assume that it is bounded and has an interior.

Compute a “center”, $\mathbf{y}^0$, of the level set $Y(z^0)$, then move the objective hyperplane through $\mathbf{y}^0$, and now consider the smaller level set

$$Y(z^1) := \{\mathbf{y} : \ A^T\mathbf{y} \le \mathbf{c},\ \mathbf{b}^T\mathbf{y} \ge z^1 = \mathbf{b}^T\mathbf{y}^0\}$$

and repeat this process.

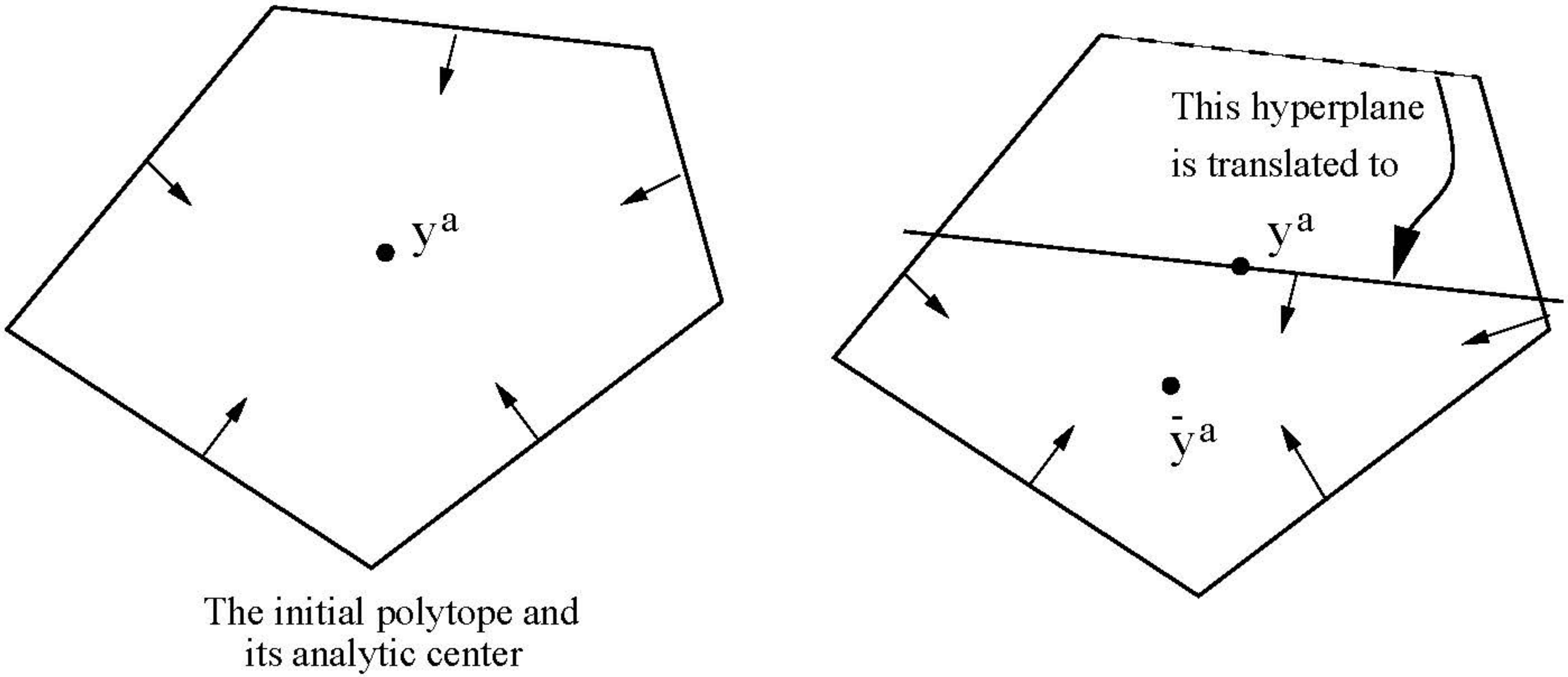

Figure 2: Cur ot translation of a hyperplane through the center.

LP with Barrier Function

Consider the LP problem with the barrier function

$$
\begin{array}{lll}
(LPB) & \text{minimize} & \mathbf{c}^T\mathbf{x} - \mu \sum_{j=1}^n \log x_j \\
 & \text{s.t.} & \mathbf{x} \in \operatorname{int} \mathcal{F}_p
\end{array}
$$

and

$$
\begin{array}{lll}
(LDB) & \text{maximize} & \mathbf{b}^T\mathbf{y} - \sum_{j=1}^n \log s_j \\
 & \text{s.t.} & (\mathbf{y}, \mathbf{s}) \in \operatorname{int} \mathcal{F}_d,
\end{array}
$$

where μ is called the barrier (weight) parameter.

They are again linearly constrained convex programs (LCCP).

Know how to derive the KKT conditions of the problem.

Common Optimality Conditions for LPB and LDB

$$
\begin{aligned}
X\mathbf{s} &= \mu\mathbf{e} \\
A\mathbf{x} &= \mathbf{b} \\
-A^T\mathbf{y} - \mathbf{s} &= -\mathbf{c};
\end{aligned}
$$

where we have

$$\mu = \frac{\mathbf{x}^T\mathbf{s}}{n} = \frac{\mathbf{c}^T\mathbf{x} - \mathbf{b}^T\mathbf{y}}{n},$$

so that it's the average of complementarity or duality gap.

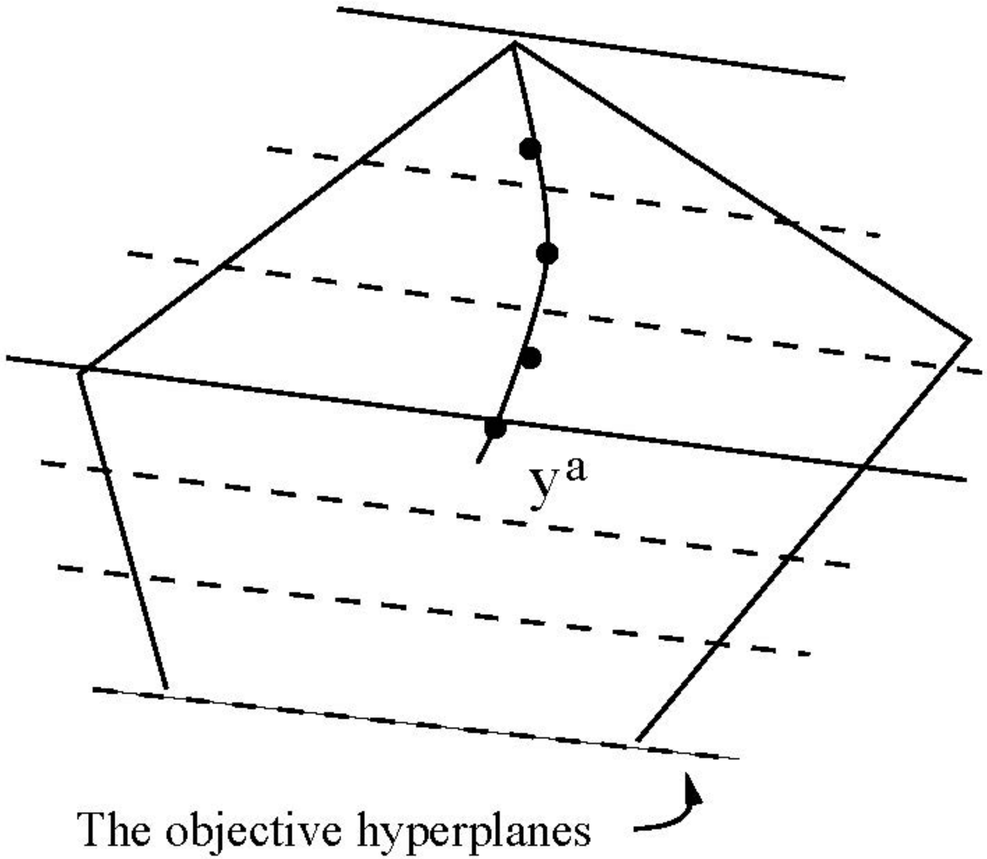

Figure 3: The central path of $\mathbf{y}(\mu)$ in a dual feasible region.

Central Path for Linear Programming

The path

$$\mathcal{C} = \{(\mathbf{x}(\mu), \mathbf{y}(\mu), \mathbf{s}(\mu)) \in \operatorname{int} \mathcal{F} : \; X\mathbf{s} = \mu\mathbf{e}, \; 0 < \mu < \infty\} \, ;$$

is called the (primal and dual) central path of linear programming.

Theorem 15 *Let both (LP) and (LD) have interior feasible points for the given data set* (A, b, c)*. Then for any* $0 < \mu < \infty$*, the central path point pair* $(\mathbf{x}(\mu), \mathbf{y}(\mu), \mathbf{s}(\mu))$ *exists and is unique.*

The uniqueness proof is based on strict convexity of $-\log(.)$ function.

Potential Function for Linear Programming

For $\mathbf{x} \in \text{int}\, \mathcal{F}_p$ and $(\mathbf{y}, \mathbf{s}) \in \text{int}\, \mathcal{F}_d$, the primal-dual potential function is defined by

$$\psi_{n+\rho}(\mathbf{x}, \mathbf{s}) := (n+\rho)\log(\mathbf{x}^T\mathbf{s}) - \sum_{j=1}^{n} \log(x_j s_j),$$

where $\rho \geq 0$.

$$\psi_{n+\rho}(\mathbf{x}, \mathbf{s}) = \rho \log(\mathbf{x}^T\mathbf{s}) + \psi_n(\mathbf{x}, \mathbf{s}) \geq \rho \log(\mathbf{x}^T\mathbf{s}) + n\log n,$$

then, for $\rho > 0$, $\psi_{n+\rho}(\mathbf{x}, \mathbf{s}) \to -\infty$ implies that $\mathbf{x}^T\mathbf{s} \to 0$. More precisely, we have

$$\mathbf{x}^T\mathbf{s} \leq \exp\left(\frac{\psi_{n+\rho}(\mathbf{x}, \mathbf{s}) - n\log n}{\rho}\right).$$

Homogeneous and Self-Dual Algorithm

- It solves the linear programming problem without any regularity assumption concerning the existence of optimal, feasible, or interior feasible solutions, while it retains the currently best complexity result
- It can start at any positive primal-dual pair, feasible or infeasible, near the central ray of the positive orthant (cone), and it does not use any big M penalty parameter or lower bound.
- Each iteration solves a system of linear equations whose dimension is almost the same as that solved in the standard (primal-dual) interior-point algorithms.
- If the LP problem has a solution, the algorithm generates a sequence that approaches feasibility and optimality simultaneously; if the problem is infeasible or unbounded, the algorithm will produce an infeasibility certificate for at least one of the primal and dual problems.

Primal-Dual Alternative Systems

A pair of LP has two alternatives

(Solvable)

$$\begin{aligned} A\mathbf{x}-\mathbf{b} &= \mathbf{0} \\ -A^T\mathbf{y}+\mathbf{c} &\geq \mathbf{0}, \\ \mathbf{b}^T\mathbf{y}-\mathbf{c}^T\mathbf{x} &= 0, \\ \mathbf{y}\ \text{free},\ \mathbf{x} &\geq \mathbf{0} \end{aligned}$$

or

(Infeasible)

$$\begin{aligned} A\mathbf{x} &= \mathbf{0} \\ -A^T\mathbf{y} &\geq \mathbf{0}, \\ \mathbf{b}^T\mathbf{y}-\mathbf{c}^T\mathbf{x} &> 0, \\ \mathbf{y}\ \text{free},\ \mathbf{x} &\geq \mathbf{0} \end{aligned}$$

An Integrated Homogeneous System

The two alternative systems can be homogenized as one:

$$
\begin{array}{lrl}
(HP) & A\mathbf{x} - \mathbf{b}\tau & = \mathbf{0} \\
 & -A^T\mathbf{y} + \mathbf{c}\tau & = \mathbf{s} \geq \mathbf{0}, \\
 & \mathbf{b}^T\mathbf{y} - \mathbf{c}^T\mathbf{x} & = \kappa \geq 0, \\
 & \mathbf{y} \text{ free}, \ (\mathbf{x}; \tau) & \geq \mathbf{0}
\end{array}
$$

where the two alternatives are

$$
\text{(Solvable)}: \ (\tau > 0, \kappa = 0) \quad \text{or} \quad \text{(Infeasible)}: \ (\tau = 0, \kappa > 0)
$$

A HSD linear program

Let's try to add one more constraint to prevent the all-zero solution

$$
\begin{array}{llrrrrl}
(HSDP) & \min & & & & (n+1)\theta & \\
& \text{s.t.} & & A\mathbf{x} & -\mathbf{b}\tau & +\bar{\mathbf{b}}\theta & = \mathbf{0}, \\
& & -A^T\mathbf{y} & & +\mathbf{c}\tau & -\bar{\mathbf{c}}\theta & \geq \mathbf{0}, \\
& & \mathbf{b}^T\mathbf{y} & -\mathbf{c}^T\mathbf{x} & & +\bar{z}\theta & \geq 0, \\
& & -\bar{\mathbf{b}}^T\mathbf{y} & +\bar{\mathbf{c}}^T\mathbf{x} & -\bar{z}\tau & & = -(n+1), \\
& & \mathbf{y} \text{ free}, & \mathbf{x} \geq \mathbf{0}, & \tau \geq 0, & \theta \text{ free}. &
\end{array}
$$

Note that the constraints of (HSDP) form a skew-symmetric system and the objective coeffcient vector is the negative of the right-hand-side vector, so that it remains a self-dual linear program.

$(\mathbf{y} = \mathbf{0},\ \mathbf{x} = \mathbf{e},\ \tau = 1,\ \theta = 1)$ is a strictly feasible point for (HSDP).

Sample Problem 1

Let $A_1 \in R^{m\times n}$, $A_2 \in R^{m\times p}$ be two given matrices, and let $\mathbf{c}_1 \in R^n$, $\mathbf{c}_2 \in R^p$ be two given *non-negative vectors*. Consider the problem

$$\begin{array}{ll} \min & \mathbf{c}_1^T\mathbf{x}_1 + \mathbf{c}_2^T\mathbf{x}_2 \\ \text{s.t.} & A_1\mathbf{x}_1 + A_2\mathbf{x}_2 = \mathbf{b} \\ & \mathbf{x}_1, \quad \mathbf{x}_2 \geq \mathbf{0}, \end{array}$$

and assume the problem is *feasible*.

(a) Sow that the problem has an optimal solution $(\mathbf{x}_1^*, \mathbf{x}_2^*)$.

(b) Let $(\mathbf{x}_1, \mathbf{x}_2)$ be a feasible solution to the problem and its objective value equals $\mathbf{b}^T y$ where $\mathbf{y}$ satisfies

$$A_1^T\mathbf{y} \leq \alpha_1\mathbf{c}_1$$

$$A_2^T\mathbf{y} \leq \alpha_2\mathbf{c}_2,$$

where α_1 and α_2 are two scalars greater than or equal to 1, then

$$\mathbf{c}_1^T\mathbf{x}_1 + \mathbf{c}_2^T\mathbf{x}_2 = \mathbf{b}^T\mathbf{y} \le \alpha_1 \cdot \mathbf{c}_1^T\mathbf{x}_1^* + \alpha_2 \cdot \mathbf{c}_2^T\mathbf{x}_2^*.$$

(α_1, α_2) is usually called the bi-factor approximation ratio and used in approximating algorithms for combinatorial optimization.

(a) Consider the dual problem:

$$\begin{array}{ll} \max & \mathbf{b}^T\mathbf{y} \\ \text{subject to} & A_1^T\mathbf{y} \le \mathbf{c}_1, \\ & A_2^T\mathbf{y} \le \mathbf{c}_2. \end{array}$$

Since $\mathbf{c}_1 \ge \mathbf{0}$ and $\mathbf{c}_2 \ge \mathbf{0}$, $\mathbf{y} = \mathbf{0}$ is a feasible point for the dual. By LP duality, since both the primal and dual problems are feasible, both must have optimal solutions.

(b) Let $(\mathbf{x}_1^*, \mathbf{x}_2^*)$ be a primal optimal solution, and let $(\mathbf{x}_1, \mathbf{x}_2)$ be a primal feasible solution with value $c\_1^T \mathbf{x}_1 + \mathbf{c}_2^T \mathbf{x}_2 = \mathbf{b}^T \mathbf{y}$, where $\mathbf{y}$ satisfies

$$A_1^T \mathbf{y} \le \alpha_1 \mathbf{c}_1, \quad A_2^T \mathbf{y} \le \alpha_2 \mathbf{c}_2.$$

Since $(\mathbf{x}_1^*, \mathbf{x}_2^*)$ is primal feasible, $A_1 \mathbf{x}_1^* + A_2 \mathbf{x}_2^* = \mathbf{b}$, and $(\mathbf{x}_1^*, \mathbf{x}_2^*) \ge \mathbf{0}$. Then,

$$(A_1^T \mathbf{y})^T \mathbf{x}_1^* \le \alpha_1 \mathbf{c}_1^T \mathbf{x}_1^*,$$

$$(A_2^T \mathbf{y})^T \mathbf{x}_2^* \le \alpha_2 \mathbf{c}_2^T \mathbf{x}_2^*.$$

Sum up all these inequalities, we have

$$\mathbf{b}^T \mathbf{y} = (A_1 \mathbf{x}_1^* + A_2 \mathbf{x}_2^*)^T \mathbf{y} = (A_1^T \mathbf{y})^T \mathbf{x}_1^* + (A_2^T \mathbf{y})^T \mathbf{x}_2^* \le \alpha_1 \mathbf{c}_1^T \mathbf{x}_1^* + \alpha_2 \mathbf{c}_2^T \mathbf{x}_2^*.$$

Sample Problem 2

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.

Sample Problem 3

Consider the standard LP primal and dual pair. Prove that, if both of them have interior feasible solutions, the optimal solution set is bounded.

Sample Problem 4

Consider the RL/MDP fixed-point computation by linear program:

$$
\begin{array}{lrcl}
\text{maximize}_{\mathbf{y}} & \sum_{i=1}^{m} y_i & & \\
\text{subject to} & y_1 - \gamma \mathbf{p}_j^T \mathbf{y} & \leq & c_j,\ j \in \mathcal{A}_1 \\
& & \vdots & \\
& y_i - \gamma \mathbf{p}_j^T \mathbf{y} & \leq & c_j,\ j \in \mathcal{A}_i \\
& & \vdots & \\
& y_m - \gamma \mathbf{p}_j^T \mathbf{y} & \leq & c_j,\ j \in \mathcal{A}_m.
\end{array}
$$

Prove at the optimal solution $\mathbf{y}^*$:

$$
y_i^* = \min_{j \in \mathcal{A}_i} \{c_j + \gamma \mathbf{p}_j^T \mathbf{y}^*\}\ \forall i
$$

What is the dual of the problem? What is the sum of a feasible solution of the dual?