# Final Review

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

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

# Convex Cones

- A set $C$ is a cone if $\mathbf{x} \in C$ implies $\alpha\mathbf{x} \in C$ for all $\alpha > 0$
- A convex cone is cone plus convex-set.
- Dual cone:

$$C^* := \{\mathbf{y} : \ \mathbf{y} \bullet \mathbf{x} \geq 0 \quad \text{for all} \quad \mathbf{x} \in C\}$$

Linear programming is to minimize a linear objective function over linear constraints and the nonnegative cone

## 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 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 vector $\mathbf{a} \in \mathcal{E}$ such that*

$$\mathbf{a} \bullet \mathbf{y} < \inf_{\mathbf{x} \in C} \mathbf{a} \bullet \mathbf{x}.$$

## Farkas' Lemma

The following results are Farkas' lemma and its variants.

**Theorem 2** *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 $x$ 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 $\mathbf{a}_{.1}, ..., \mathbf{a}_{.n}$, then there is a hyperplane separating $\mathbf{b}$ from $\text{cone}(\mathbf{a}_{.1}, ..., \mathbf{a}_{.n})$.

**Theorem 3** *Let $A \in \mathcal{R}^{m\times n}$ and $\mathbf{c} \in \mathcal{R}^n$. Then, the system $\{\mathbf{y} : \ A^T\mathbf{y} \le \mathbf{c}\}$ has a solution $y$ if and only if that $A\mathbf{x} = \mathbf{0}$ and $\mathbf{x} \ge \mathbf{0}$ imply $\mathbf{c}^T\mathbf{x} \ge 0$.*

Again, 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 $\{\mathbf{y} : \ A^T\mathbf{y} \le \mathbf{c}\}$.

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

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 called dual slacks.

# Duality Theory

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

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

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

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

From this we have important results in the following.

**Theorem 5** *(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 they are both optimal. The converse is also true; there is no “gap.”

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 also called the complementarity gap.

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

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

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

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

# Rules to construct the dual

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

## Duality Example

Consider the combinatorial call auction market discussed in the class. This time, the market maker forms the decision problem as:

$$
\begin{array}{lll}
\max & \sum_{j=1}^{n} x_j & \\
\text{s.t.} & A\mathbf{x} - \mathbf{e} \cdot y & \leq \mathbf{0}, \ (\mathbf{p}) \\
& -\pi^T \mathbf{x} + \alpha \cdot y & \leq 0, \ (\lambda) \\
& \mathbf{x} & \leq \mathbf{q}, \ (\mu) \\
& \mathbf{x} & \geq \mathbf{0},
\end{array}
$$

where $(\pi_j, \mathbf{a}_j, q_j)$ are as defined as in our auction problem through out this course, $\mathbf{e}$ is the vector of all ones, and parameter $\alpha \geq 0$. Again, the bidder wins one dollar if the winning state is in his or her selection.

## Basic Feasible Solution

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

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

for the $m$-vector $\mathbf{x}_B$. By setting the variables, $\mathbf{x}_N$, of $\mathbf{x}$ corresponding to the remaining columns of $A$ equal 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 components of $\mathbf{x}_B$ 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 $\mathbf{x}_B$ has value zero, that basic feasible solution $\mathbf{x}$ is said to be (primal) degenerate.

A dual vector $\mathbf{y}$ satisfying

$$A_B^T \mathbf{y} = c_B$$

is said to be the corresponding dual basic solution.

If the dual basic solution is also feasible, that is

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

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

**Theorem 6** *(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.*

If there is one primal optimal basic solution that is not degenerate, then the dual optimal solution is unique.

## 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.

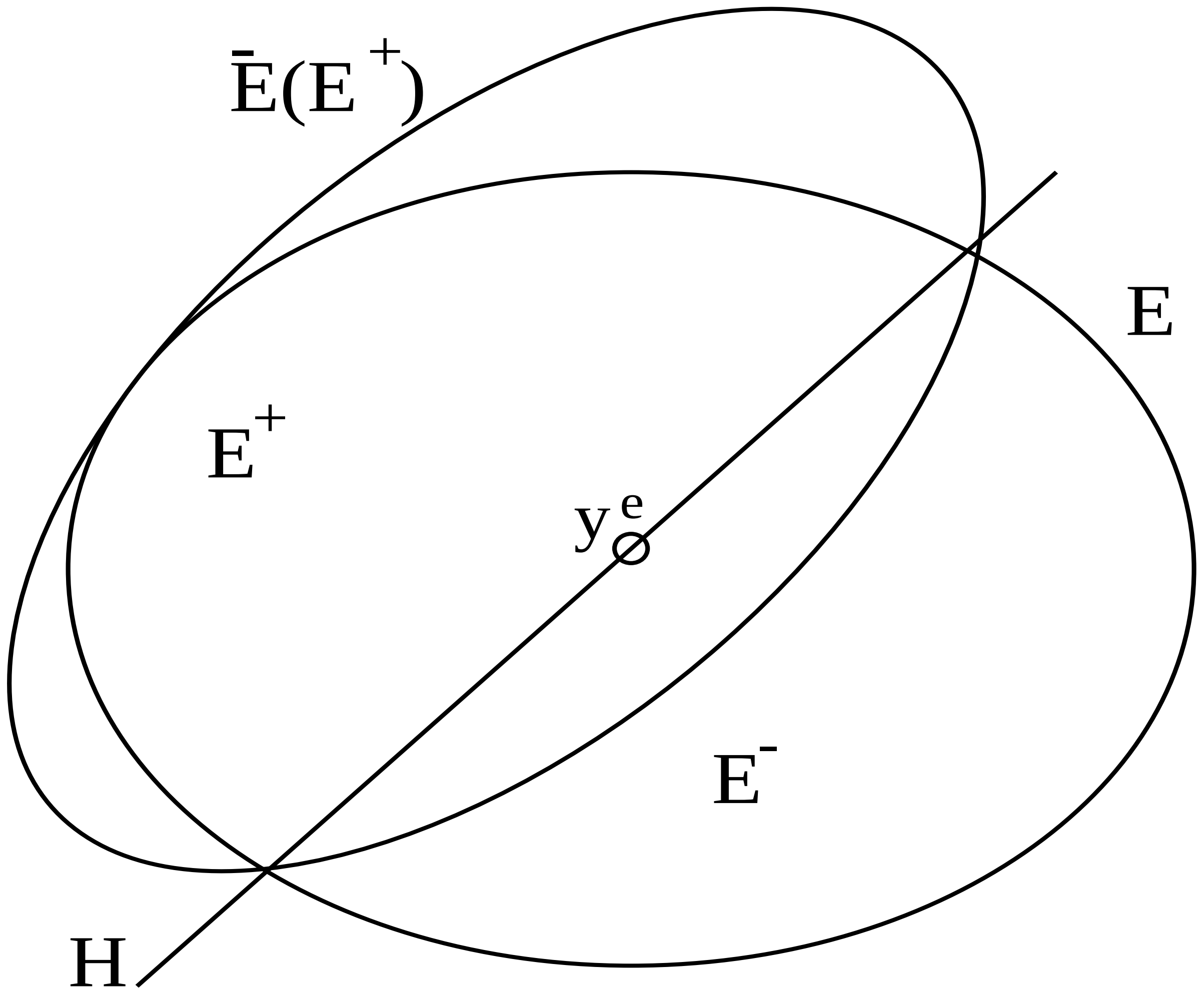


Figure 1: The least volume ellipsoid containing a half ellipsoid

## Desired Theoretical Properties

- Separation Problem: Either decide $\mathbf{x} \in P$ or find a vector $\mathbf{d}$ such that $\mathbf{d}^T\mathbf{x} \leq \mathbf{d}^T\mathbf{y}$ for all $\mathbf{y} \in P$.
- Oracle to generate $\mathbf{d}$ without enumerating all hyperplanes.

**Theorem 7** *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 Method of Centers

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.

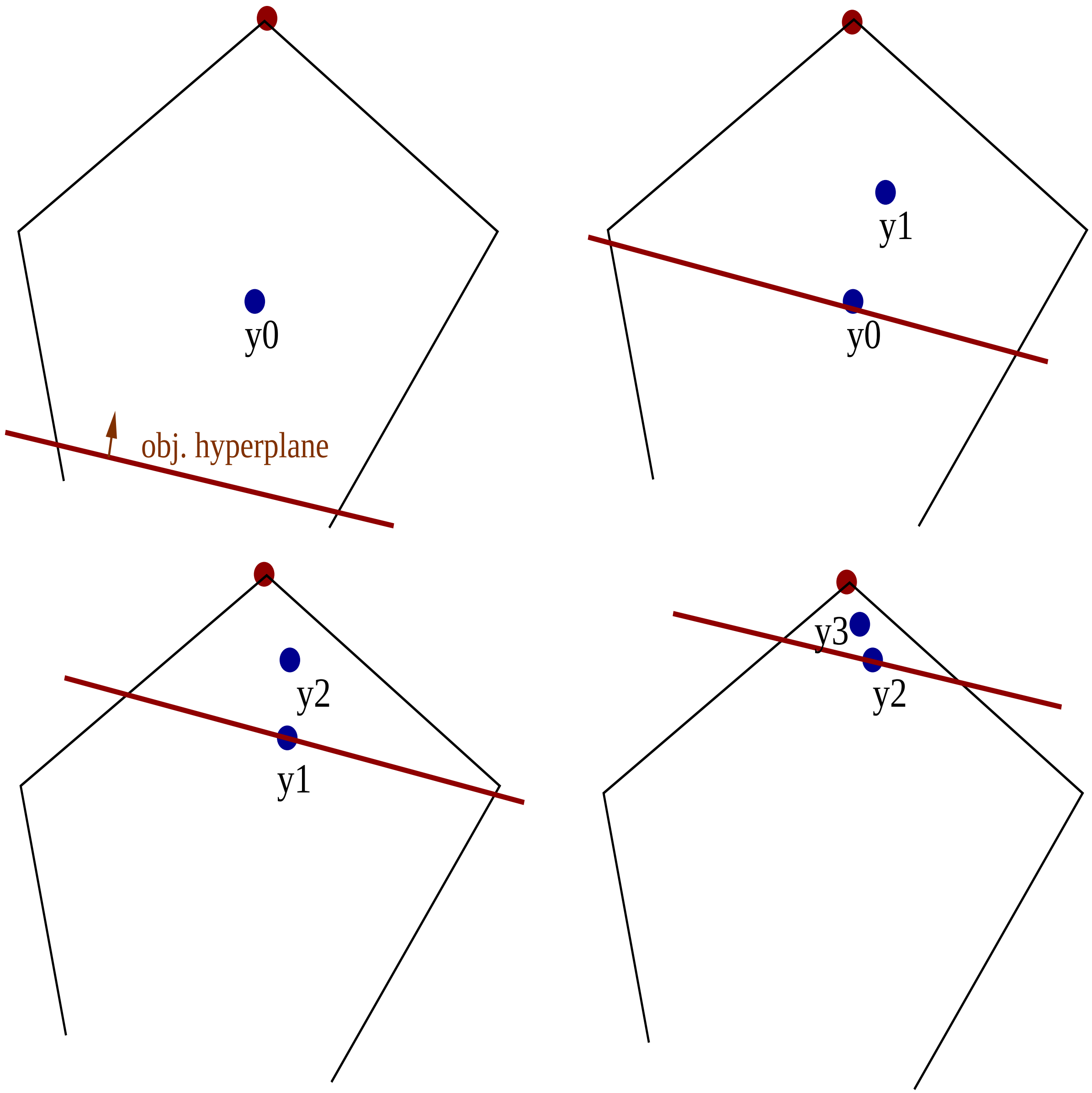


Figure 2: The analytic center-section method.

## Analytic Center for the Polytope

One choice of center is the one to maximize the barrier function over the level set:

$$
\begin{array}{ll}
\text{maximize} & \log s_0 + \sum_j \log s_j \\
\text{subject to} & A^T\mathbf{y} + \mathbf{s} = \mathbf{c}, \\
& \mathbf{b}^T\mathbf{y} - s_0 = z^0.
\end{array}
$$

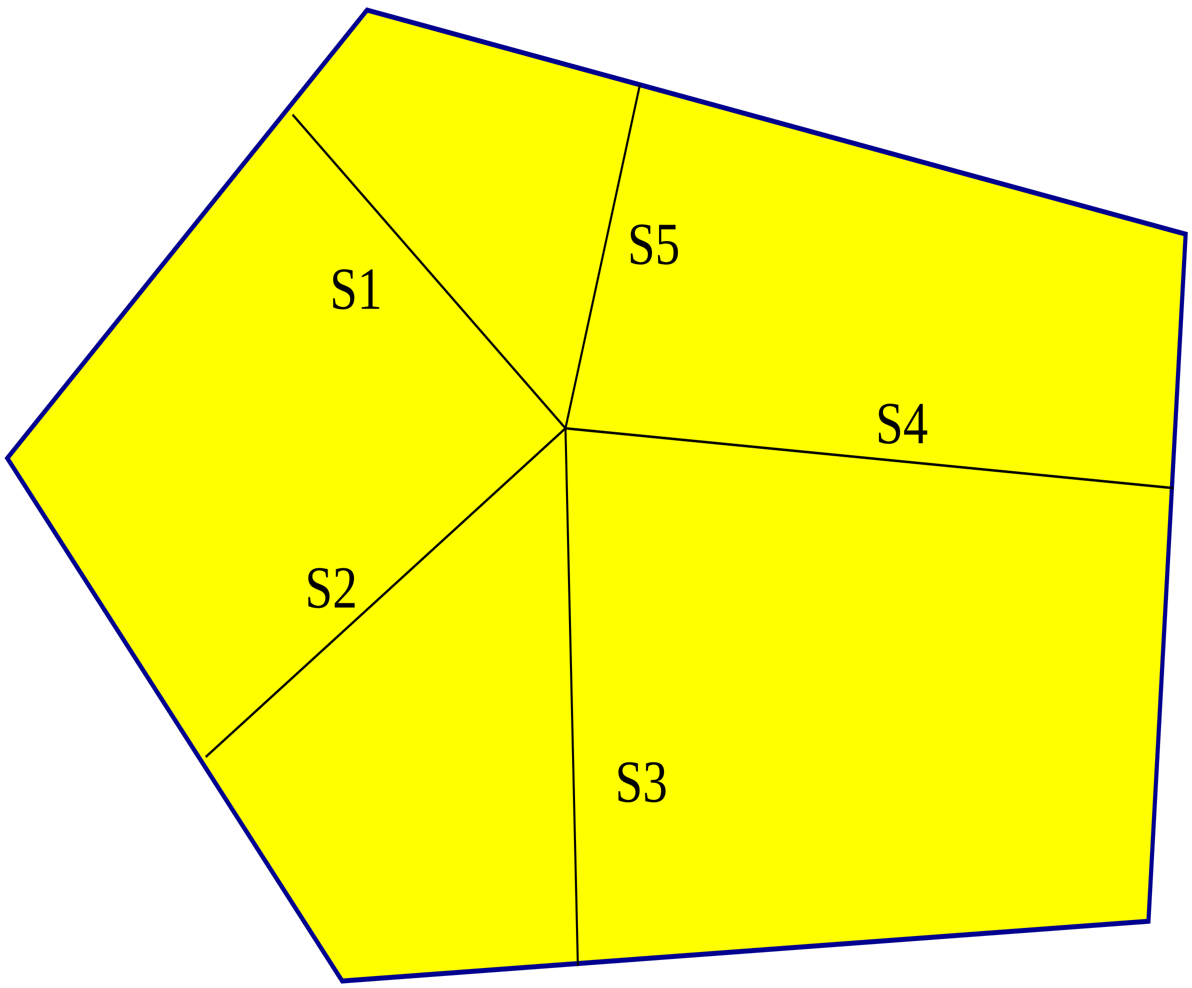


Figure 3: Analytic center maximizes the barrier function.

## 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 $\mu$ is called the barrier (weight) parameter.

They are again linearly constrained convex programs (LCCP).

## 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.

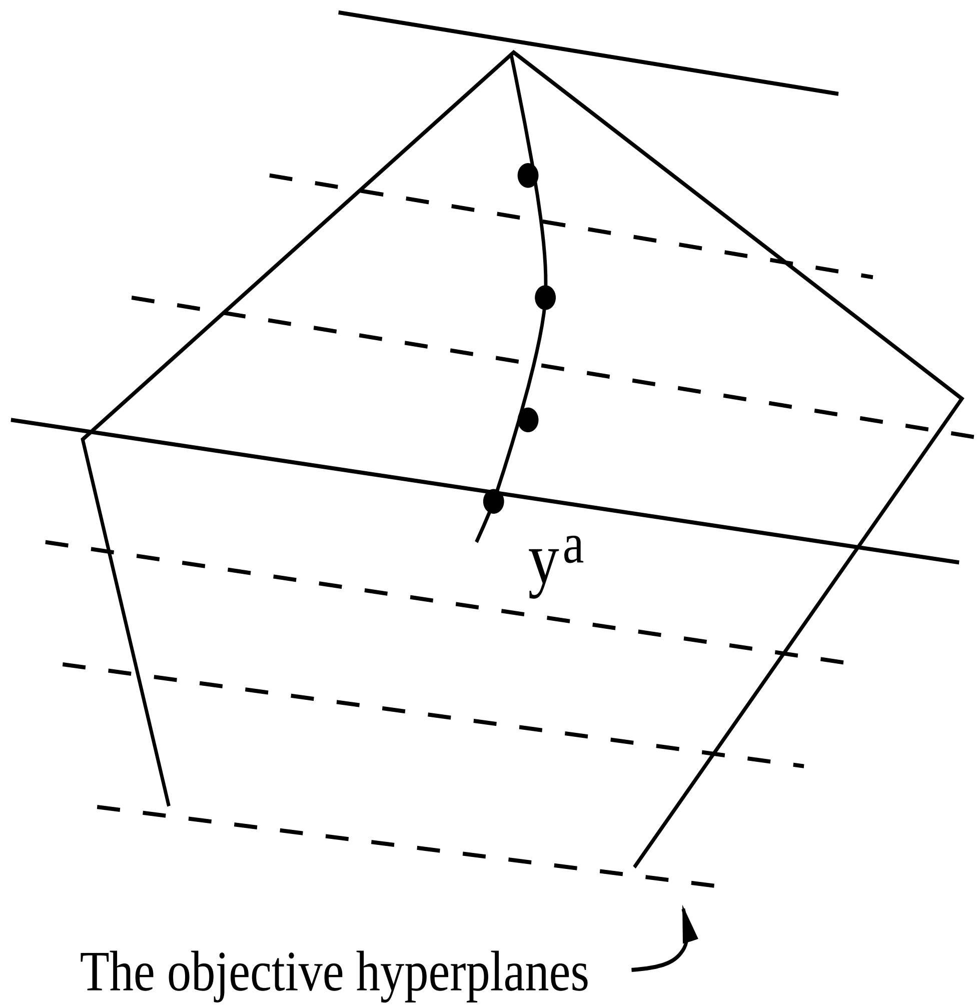


Figure 4: The central path of $\mathbf{y}(\mu)$ in a dual feasible region.

## Central Path for Linear Programming

The path

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

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

**Theorem 8** *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.*

## 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

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

## Primal-Dual Potential Reduction Algorithm for LP

Once we have a pair $(\mathbf{x}, \mathbf{y}, \mathbf{s}) \in \operatorname{int} \mathcal{F}$ with $\mu = \mathbf{x}^T\mathbf{s}/n$, we can generate a new iterate $\mathbf{x}^+$ and $(\mathbf{y}^+, \mathbf{s}^+)$ by solving for $\mathbf{d}_x$, $\mathbf{d}_y$ and $\mathbf{d}_s$ from the system of linear equations:

$$\begin{aligned} S\mathbf{d}_x + X\mathbf{d}_s &= \mathbf{r} := \tfrac{\mathbf{x}^T\mathbf{s}}{n+\rho}\mathbf{e} - X\mathbf{s}, \\ A\mathbf{d}_x &= \mathbf{0}, \\ -A^T\mathbf{d}_y - \mathbf{d}_s &= \mathbf{0}. \end{aligned} \tag{1}$$

Let $\mathbf{d} := (\mathbf{d}_x, \mathbf{d}_y, \mathbf{d}_s)$. To show the dependence of $\mathbf{d}$ on the current pair $(\mathbf{x}, \mathbf{s})$ and the parameter $\gamma$, we write $\mathbf{d} = \mathbf{d}(\mathbf{x}, \mathbf{s}, \gamma)$. Note that $\mathbf{d}_x^T\mathbf{d}_s = -\mathbf{d}_x^T A^T\mathbf{d}_y = 0$ here. Th results still hold even if $\mathbf{d}_x^T\mathbf{d}_s \geq 0$.

**Lemma 1** *Let the direction* $\mathbf{d} = (\mathbf{d}_x, \mathbf{d}_y, \mathbf{d}_s)$ *be generated by equation (1), and let*

$$\theta = \frac{\alpha\sqrt{\min(X\mathbf{s})}}{\|(XS)^{-1/2}(\frac{\mathbf{x}^T\mathbf{s}}{(n+\rho)}\mathbf{e} - X\mathbf{s})\|}\ , \tag{2}$$

*where* $\alpha$ *is a positive constant less than 1. Let*

$$\mathbf{x}^+ = \mathbf{x} + \theta\mathbf{d}_x, \quad \mathbf{y}^+ = \mathbf{y} + \theta\mathbf{d}_y, \quad \textit{and} \quad \mathbf{s}^+ = \mathbf{s} + \theta\mathbf{d}_s.$$

*Then, we have* $(\mathbf{x}^+, \mathbf{y}^+, \mathbf{s}^+) \in \operatorname{int}\mathcal{F}$ *and*

$$\psi_{n+\rho}(\mathbf{x}^+, \mathbf{s}^+) - \psi_{n+\rho}(\mathbf{x}, \mathbf{s})$$

$$\leq -\alpha\sqrt{\min(X\mathbf{s})}\|(XS)^{-1/2}(\mathbf{e} - \frac{(n+\rho)}{\mathbf{x}^T\mathbf{s}}X\mathbf{s})\| + \frac{\alpha^2}{2(1-\alpha)}\ .$$

## 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)

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

or

(Infeasible)

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

## An Integrated Homogeneous System

The two alternative systems can be homogenized as one:

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

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

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

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).

## The Alternating Direction Method with Multipliers

We consider

$$\min \quad f_1(\mathbf{x}_1) + f_2(\mathbf{x}_2) \quad \text{s.t.} \quad A_1\mathbf{x}_1 + A_2\mathbf{x}_2 = \mathbf{b},\ \mathbf{x}_1 \in X_1,\ \mathbf{x}_2 \in X_2;$$

where $X_1$ and $X_2$ are (simple) convex sets.

Define its Augmented Lagrangian

$$L(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}) = f_1(\mathbf{x}_1) + f_2(\mathbf{x}_2) - \mathbf{y}^T(A_1\mathbf{x}_1 + A_2\mathbf{x}_2 - \mathbf{b}) + \frac{\beta}{2}\|A_1\mathbf{x}_1 + A_2\mathbf{x}_2 - \mathbf{b}\|^2.$$

Then, for any given $(\mathbf{x}_1^k, \mathbf{x}_2^k, \mathbf{y}^k)$, we compute a new iterate pair

$$\mathbf{x}_1^{k+1} = \arg\min_{\mathbf{x}_1 \in X_1} L(\mathbf{x}_1, \mathbf{x}_2^k, \mathbf{y}^k),\ \mathbf{x}_2^{k+1} = \arg\min_{\mathbf{x}_2 \in X_2} L(\mathbf{x}_1^{k+1}, \mathbf{x}_2, \mathbf{y}^k)$$

and

$$\mathbf{y}^{k+1} = \mathbf{y}^k - \beta(A_1\mathbf{x}_1^{k+1} + A_2\mathbf{x}_2^{k+1} - \mathbf{b}).$$

Again, the iterates converge for any $\beta > 0$ with the same speed as the SDM.

## The ADMM for LP with Inequalities

We consider

$$\min \quad \mathbf{c}^T\mathbf{x} \quad \text{s.t.} \quad A\mathbf{x}+\mathbf{s}=\mathbf{b},\ \mathbf{s}\geq \mathbf{0}.$$

$$L(\mathbf{x},\mathbf{s},\mathbf{y}) = \mathbf{c}^T\mathbf{x} - \mathbf{y}^T(A\mathbf{x}+\mathbf{s}-\mathbf{b}) + \frac{\beta}{2}\|A\mathbf{x}+\mathbf{s}-\mathbf{b}\|^2.$$

Then, for any given $(\mathbf{x}^k, \mathbf{s}^k \geq \mathbf{0}, \mathbf{y}^k)$, we compute a new iterate pair

$$\mathbf{x}^{k+1} = \arg\min_{\mathbf{x}} L(\mathbf{x}, \mathbf{s}^k, \mathbf{y}^k)$$

$$\mathbf{s}^{k+1} = \arg\min_{\mathbf{s}\geq\mathbf{0}} L(\mathbf{x}^{k+1}, \mathbf{s}, \mathbf{y}^k)$$

and

$$\mathbf{y}^{k+1} = \mathbf{y}^k - \beta(A\mathbf{x}^{k+1}+\mathbf{s}^{k+1}-\mathbf{b}).$$

Note that the solution of $\mathbf{s}$ can be computed in a close form!

## The ADMM for LP in Standard Form

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

We consider an equivalent problem:

$$(LP) \quad \text{minimize} \quad \mathbf{c} \bullet \mathbf{x}_1$$
$$\text{subject to} \quad A\mathbf{x}_1 = \mathbf{b},\ \mathbf{x}_1 - \mathbf{x}_2 = \mathbf{0},\ \mathbf{x}_2 \geq \mathbf{0},$$

## The ADMM for LP

Consider its Augmented Lagrangian

$$L(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}, \mathbf{s}) = \mathbf{c}^T\mathbf{x}_1 - \mathbf{y}^T(A\mathbf{x}_1 - \mathbf{b}) - \mathbf{s}^T(\mathbf{x}_1 - \mathbf{x}_2) + \frac{\beta}{2}\|A\mathbf{x}_1 - \mathbf{b}\|^2 + \frac{\beta}{2}\|\mathbf{x}_1 - \mathbf{x}_2\|^2.$$

Then, for any given $(\mathbf{x}_1^k, \mathbf{x}_2^k, \mathbf{y}^k, \mathbf{s}^k)$, we compute a new iterate pair

$$\mathbf{x}_1^{k+1} = \arg\min_{\mathbf{x}_1} L(\mathbf{x}_1, \mathbf{x}_2^k, \mathbf{y}^k, \mathbf{s}^k)$$

$$\mathbf{x}_2^{k+1} = \arg\min_{\mathbf{x}_2 \geq \mathbf{0}} L(\mathbf{x}_1^{k+1}, \mathbf{x}_2, \mathbf{y}^k, \mathbf{s}^k)$$

and

$$\mathbf{y}^{k+1} = \mathbf{y}^k - \beta(A\mathbf{x}_1^{k+1} - \mathbf{b}) \quad \text{and} \quad \mathbf{s}^{k+1} = \mathbf{s}^k - \beta(\mathbf{x}_1^{k+1} - \mathbf{x}_2^{k+1}).$$

The minimization over $\mathbf{x}_1$ is a unconstrained optimization, and again the minimization over $\mathbf{x}_2$ can be computed in a close form!