# 15.081J/6.251J Introduction to Mathematical Programming

Lecture 21: Primal Barrier
Interior Point Algorithm

# 1 Outline

SLIDE 1

1. Barrier Methods
2. The Central Path
3. Approximating the Central Path
4. The Primal Barrier Algorithm
5. Correctness and Complexity

# 2 Barrier methods

SLIDE 2

$$\begin{array}{ll} \min & f(\boldsymbol{x}) \\ \text{s.t.} & g_j(\boldsymbol{x}) \leq 0, \quad\quad j = 1, \ldots, p \\ & h_i(\boldsymbol{x}) = 0, \quad\quad i = 1, \ldots, m \end{array}$$

$$S = \{\boldsymbol{x} |\ g_j(\boldsymbol{x}) < 0,\ j = 1, \ldots, p,$$
$$h_i(\boldsymbol{x}) = 0,\ i = 1, \ldots, m\}$$

## 2.1 Strategy

SLIDE 3

- A barrier function $G(\boldsymbol{x})$ is a continous function with the property that is approaches $\infty$ as one of $g_j(\boldsymbol{x})$ approaches 0 from negative values.
- Examples:

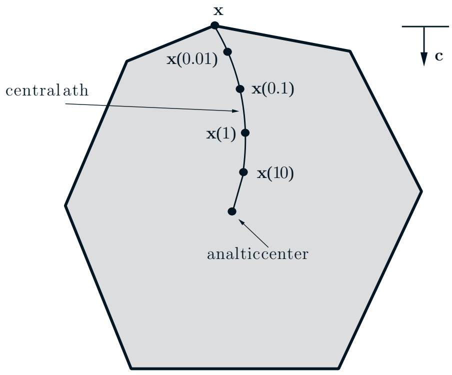
$$G(\boldsymbol{x}) = -\sum_{j=1}^{p} \log(-g_j(\boldsymbol{x})),\ G(\boldsymbol{x}) = -\sum_{j=1}^{p} \frac{1}{g_j(\boldsymbol{x})}$$

SLIDE 4

- Consider a sequence of $\mu^k$: $0 < \mu^{k+1} < \mu^k$ and $\mu^k \to 0$.
- Consider the problem

$$\boldsymbol{x}^k = \text{argmin}_{\boldsymbol{x} \in S} \left\{ f(\boldsymbol{x}) + \mu^k G(\boldsymbol{x}) \right\}$$

- Theorem Every limit point $\boldsymbol{x}^k$ generated by a barrier method is a global minimum of the original constrained problem.




## 2.2 Primal path-following IPMs for LO

SLIDE 5

$$(P)\quad \begin{aligned} \min \quad & \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0} \end{aligned} \qquad\qquad (D)\quad \begin{aligned} \max \quad & \boldsymbol{b}'\boldsymbol{p} \\ \text{s.t.} \quad & \boldsymbol{A}'\boldsymbol{p} + \boldsymbol{s} = \boldsymbol{c} \\ & \boldsymbol{s} \geq \boldsymbol{0} \end{aligned}$$

Barrier problem:

$$\begin{aligned} \min \quad & B_\mu(\boldsymbol{x}) = \boldsymbol{c}'\boldsymbol{x} - \mu \sum_{j=1}^{n} \log x_j \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \end{aligned}$$

Minimizer: $\boldsymbol{x}(\mu)$

# 3 Central Path

SLIDE 6

- As $\mu$ varies, minimizers $\boldsymbol{x}(\mu)$ form the central path
- $\lim_{\mu \to 0} \boldsymbol{x}(\mu)$ exists and is an optimal solution $\boldsymbol{x}^*$ to the initial LP
- For $\mu = \infty$, $\boldsymbol{x}(\infty)$ is called the *analytic center*

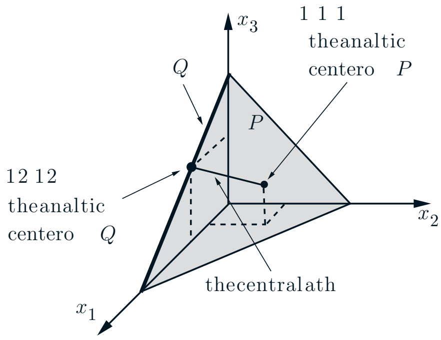
$$\begin{aligned} \min \quad & -\sum_{j=1}^{n} \log x_j \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \end{aligned}$$

SLIDE 7

## 3.1 Example

SLIDE 8

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




- $Q = \{\boldsymbol{x} \mid \boldsymbol{x} = (x_1, 0, x_3),\ x_1 + x_3 = 1,\ \boldsymbol{x} \geq \boldsymbol{0}\}$, set of optimal solutions to original LP
- The analytic center of $Q$ is $(1/2, 0, 1/2)$

$$\begin{array}{rl} \min & x_2 - \mu \log x_1 - \mu \log x_2 - \mu \log x_3 \\ \text{s.t.} & x_1 + x_2 + x_3 = 1 \end{array}$$

$$\min \quad x_2 - \mu \log x_1 - \mu \log x_2 - \mu \log(1 - x_1 - x_2).$$

$$\begin{aligned} x_1(\mu) &= \frac{1 - x_2(\mu)}{2} \\ x_2(\mu) &= \frac{1 + 3\mu - \sqrt{1 + 9\mu^2 + 2\mu}}{2} \\ x_3(\mu) &= \frac{1 - x_2(\mu)}{2} \end{aligned}$$

The analytic center: $(1/3, 1/3, 1/3)$ SLIDE 9

## 3.2 Solution of Central Path

SLIDE 10

- Barrier problem for dual:

$$\begin{array}{rl} \max & \boldsymbol{p}'\boldsymbol{b} + \mu \sum_{j=1}^{n} \log s_j \\ \text{s.t.} & \boldsymbol{p}'\boldsymbol{A} + \boldsymbol{s}' = \boldsymbol{c}' \end{array}$$

- Solution (KKT):

$$\begin{aligned} \boldsymbol{A}\boldsymbol{x}(\mu) &= \boldsymbol{b} \\ \boldsymbol{x}(\mu) &\geq \boldsymbol{0} \\ \boldsymbol{A}'\boldsymbol{p}(\mu) + \boldsymbol{s}(\mu) &= \boldsymbol{c} \\ \boldsymbol{s}(\mu) &\geq \boldsymbol{0} \\ \boldsymbol{X}(\mu)\boldsymbol{S}(\mu)\boldsymbol{e} &= \boldsymbol{e}\mu \end{aligned}$$

Slide 11

- Theorem: If $\boldsymbol{x}^*$, $\boldsymbol{p}^*$, and $\boldsymbol{s}^*$ satisfy optimality conditions, then they are optimal solutions to problems primal and dual barrier problems.
- Goal: Solve barrier problem

$$\begin{aligned} \min \quad & B_\mu(\boldsymbol{x}) = \boldsymbol{c}'\boldsymbol{x} - \mu \sum_{j=1}^{n} \log x_j \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \end{aligned}$$

# 4 Approximating the central path

Slide 12

$$\begin{aligned} \frac{\partial B_\mu(\boldsymbol{x})}{\partial x_i} &= c_i - \frac{\mu}{x_i} \\ \frac{\partial^2 B_\mu(\boldsymbol{x})}{\partial x_i^2} &= \frac{\mu}{x_i^2} \\ \frac{\partial^2 B_\mu(\boldsymbol{x})}{\partial x_i \partial x_j} &= 0, \quad i \neq j \end{aligned}$$

Given a vector $\boldsymbol{x} > \boldsymbol{0}$: Slide 13

$$\begin{aligned} B_\mu(\boldsymbol{x}+\boldsymbol{d}) &\approx B_\mu(\boldsymbol{x}) + \sum_{i=1}^{n} \frac{\partial B_\mu(\boldsymbol{x})}{\partial x_i} d_i + \\ & \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 B_\mu(\boldsymbol{x})}{\partial x_i \partial x_j} d_i d_j \\ &= B_\mu(\boldsymbol{x}) + (\boldsymbol{c}' - \mu \boldsymbol{e}'\boldsymbol{X}^{-1})\boldsymbol{d} + \frac{1}{2}\mu \boldsymbol{d}'\boldsymbol{X}^{-2}\boldsymbol{d} \end{aligned}$$

$\boldsymbol{X} = \mathrm{diag}(x_1, \ldots, x_n)$ Slide 14

Approximating problem:

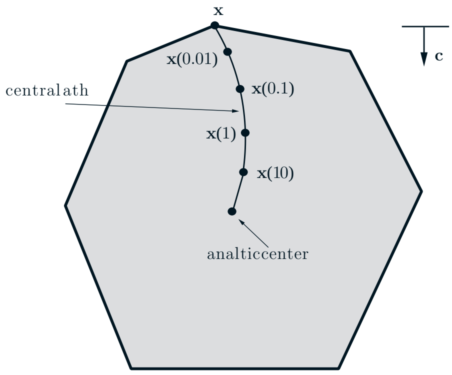
$$\begin{aligned} \min \quad & (\boldsymbol{c}' - \mu \boldsymbol{e}'\boldsymbol{X}^{-1})\boldsymbol{d} + \frac{1}{2}\mu \boldsymbol{d}'\boldsymbol{X}^{-2}\boldsymbol{d} \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{d} = \boldsymbol{0} \end{aligned}$$

Solution (from Lagrange):

$$\begin{aligned} \boldsymbol{c} - \mu \boldsymbol{X}^{-1}\boldsymbol{e} + \mu \boldsymbol{X}^{-2}\boldsymbol{d} - \boldsymbol{A}'\boldsymbol{p} &= \boldsymbol{0} \\ \boldsymbol{A}\boldsymbol{d} &= \boldsymbol{0} \end{aligned}$$

Slide 15

- System of $m+n$ linear equations, with $m+n$ unknowns ($d_j,\ j = 1, \ldots, n$, and $p_i,\ i = 1, \ldots, m$).




- Solution:

$$\begin{aligned} \boldsymbol{d}(\mu) &= \Big(\boldsymbol{I} - \boldsymbol{X}^2\boldsymbol{A}'(\boldsymbol{A}\boldsymbol{X}^2\boldsymbol{A}')^{-1}\boldsymbol{A}\Big)\left(\boldsymbol{x}\boldsymbol{e} - \frac{1}{\mu}\boldsymbol{X}^2\boldsymbol{c}\right) \\ \boldsymbol{p}(\mu) &= (\boldsymbol{A}\boldsymbol{X}^2\boldsymbol{A}')^{-1}\boldsymbol{A}(\boldsymbol{X}^2\boldsymbol{c} - \mu\boldsymbol{x}\boldsymbol{e}) \end{aligned}$$

## 4.1 The Newton connection

SLIDE 16

- $\boldsymbol{d}(\mu)$ is the *Newton direction*; process of calculating this direction is called a *Newton step*
- Starting with $\boldsymbol{x}$, the new primal solution is $\boldsymbol{x} + \boldsymbol{d}(\mu)$
- The corresponding dual solution becomes $(\boldsymbol{p}, \boldsymbol{s}) = \big(\boldsymbol{p}(\mu), \boldsymbol{c} - \boldsymbol{A}'\boldsymbol{p}(\mu)\big)$
- We then decrease $\mu$ to $\overline{\mu} = \alpha\mu$, $0 < \alpha < 1$

## 4.2 Geometric Interpretation

SLIDE 17

- Take one Newton step so that $\boldsymbol{x}$ would be close to $\boldsymbol{x}(\mu)$
- Measure of closeness

$$\left|\left|\frac{1}{\mu}\boldsymbol{X}\boldsymbol{S}\boldsymbol{e} - \boldsymbol{e}\right|\right| \leq \beta,$$

  $0 < \beta < 1$, $\boldsymbol{X} = \text{diag}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ $\boldsymbol{S} = \text{diag}(\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n)$
- As $\mu \to 0$, the complementarity slackness condition will be satisfied

SLIDE 18

## 5 The Primal Barrier Algorithm

SLIDE 19

Input

(a) $(\boldsymbol{A},\ \boldsymbol{b},\ \boldsymbol{c})$; $\boldsymbol{A}$ has full row rank;

(b) $\boldsymbol{x}^0 > \boldsymbol{0},\ \boldsymbol{s}^0 > \boldsymbol{0},\ \boldsymbol{p}^0$;

(c) optimality tolerance $\epsilon > 0$;

(d) $\mu^0$, and $\alpha$, where $0 < \alpha < 1$.

SLIDE 20

**1.** (Initialization) Start with some primal and dual feasible $\boldsymbol{x}^0 > \boldsymbol{0},\ \boldsymbol{s}^0 > \boldsymbol{0},\ \boldsymbol{p}^0$, and set $k = 0$.

**2.** (Optimality test) If $(\boldsymbol{s}^k)'\boldsymbol{x}^k < \epsilon$ stop; else go to Step 3.

**3.** Let

$$\begin{aligned} \boldsymbol{X}_k &= \text{diag}(x_1^k, \ldots, x_n^k), \\ \mu^{k+1} &= \alpha\mu^k \end{aligned}$$

SLIDE 21

**4.** (Computation of directions) Solve the linear system

$$\begin{aligned} \mu^{k+1}\boldsymbol{X}_k^{-2}\boldsymbol{d} - \boldsymbol{A}'\boldsymbol{p} &= \mu^{k+1}\boldsymbol{X}_k^{-1}\boldsymbol{e} - \boldsymbol{c} \\ \boldsymbol{A}\boldsymbol{d} &= \boldsymbol{0} \end{aligned}$$

**5.** (Update of solutions) Let

$$\begin{aligned} \boldsymbol{x}^{k+1} &= \boldsymbol{x}^k + \boldsymbol{d}, \\ \boldsymbol{p}^{k+1} &= \boldsymbol{p}, \\ \boldsymbol{s}^{k+1} &= \boldsymbol{c} - \boldsymbol{A}'\boldsymbol{p}. \end{aligned}$$

**6.** Let $k := k + 1$ and go to Step 2.

## 6 Correctness

SLIDE 22

Theorem Given $\alpha = 1 - \dfrac{\sqrt{\beta} - \beta}{\sqrt{\beta} + \sqrt{n}}$, $\beta < 1$, $(\boldsymbol{x}^0, \boldsymbol{s}^0, \boldsymbol{p}^0)$, $(\boldsymbol{x}^0 > \boldsymbol{0},\ \boldsymbol{s}^0 > \boldsymbol{0})$:

$$\left\| \frac{1}{\mu^0}\boldsymbol{X}_0\boldsymbol{S}_0\boldsymbol{e} - \boldsymbol{e} \right\| \le \beta.$$

Then, after

$$K = \left\lceil \frac{\sqrt{\beta} + \sqrt{n}}{\sqrt{\beta} - \beta} \log \frac{(\boldsymbol{s}^0)'\boldsymbol{x}^0(1 + \beta)}{\epsilon(1 - \beta)} \right\rceil$$

iterations, $(\boldsymbol{x}^K, \boldsymbol{s}^K, \boldsymbol{p}^K)$ is found:

$$(\boldsymbol{s}^K)'\boldsymbol{x}^K \le \epsilon.$$

## 6.1 Proof



- Claim (by induction): $\left|\left|\frac{1}{\mu^k}\boldsymbol{X}_k\boldsymbol{S}_k\boldsymbol{e}-\boldsymbol{e}\right|\right| \leq \beta$
- For $k=0$ we have assumed it
- Assume it holds for $k$;

$$\begin{aligned}
\left|\left|\frac{1}{\mu^{k+1}}\boldsymbol{X}_k\boldsymbol{S}_k\boldsymbol{e}-\boldsymbol{e}\right|\right| &= \left|\left|\frac{1}{\alpha\mu^k}\boldsymbol{X}_k\boldsymbol{S}_k\boldsymbol{e}-\boldsymbol{e}\right|\right| \\
&= \left|\left|\frac{1}{\alpha}\left(\frac{1}{\mu^k}\boldsymbol{X}_k\boldsymbol{S}_k\boldsymbol{e}-\boldsymbol{e}\right)+\frac{1-\alpha}{\alpha}\boldsymbol{e}\right|\right| \\
&\leq \frac{1}{\alpha}\left|\left|\frac{1}{\mu^k}\boldsymbol{X}_k\boldsymbol{S}_k\boldsymbol{e}-\boldsymbol{e}\right|\right|+\frac{1-\alpha}{\alpha}||\boldsymbol{e}|| \\
&\leq \frac{\beta}{\alpha}+\frac{1-\alpha}{\alpha}\sqrt{n} \\
&= \sqrt{\beta}
\end{aligned}$$

- We next show that $||\boldsymbol{X}_k^{-1}\boldsymbol{d}|| \leq \sqrt{\beta} < 1$, where $\boldsymbol{d}=\boldsymbol{x}^{k+1}-\boldsymbol{x}^k$.
- $\boldsymbol{d}$ solves

$$\begin{aligned}
\mu^{k+1}\boldsymbol{X}_k^{-2}\boldsymbol{d}-\boldsymbol{A}'\boldsymbol{p} &= \mu^{k+1}\boldsymbol{X}_k^{-1}\boldsymbol{e}-\boldsymbol{c}, \\
\boldsymbol{A}\boldsymbol{d} &= \boldsymbol{0}
\end{aligned}$$

- By left-multiplying the first equation by $\boldsymbol{d}'$

$$\mu^{k+1}\boldsymbol{d}'\boldsymbol{X}_k^{-2}\boldsymbol{d}=\boldsymbol{d}'\left(\mu^{k+1}\boldsymbol{X}_k^{-1}\boldsymbol{e}-\boldsymbol{c}\right)$$

$$\begin{aligned}
||\boldsymbol{X}_k^{-1}\boldsymbol{d}||^2 &= \boldsymbol{d}'\boldsymbol{X}_k^{-2}\boldsymbol{d} \\
&= \left(\boldsymbol{X}_k^{-1}\boldsymbol{e}-\frac{1}{\mu^{k+1}}\boldsymbol{c}\right)'\boldsymbol{d} \\
&= \left(\boldsymbol{X}_k^{-1}\boldsymbol{e}-\frac{1}{\mu^{k+1}}(\boldsymbol{s}^k+\boldsymbol{A}'\boldsymbol{p}^k)\right)'\boldsymbol{d} \\
&= \left(\boldsymbol{X}_k^{-1}\boldsymbol{e}-\frac{1}{\mu^{k+1}}\boldsymbol{s}^k\right)'\boldsymbol{d} \\
&= -\left(\frac{1}{\mu^{k+1}}\boldsymbol{X}_k\boldsymbol{S}_k\boldsymbol{e}-\boldsymbol{e}\right)'\boldsymbol{X}_k^{-1}\boldsymbol{d} \\
&\leq \left|\left|\frac{1}{\mu^{k+1}}\boldsymbol{X}_k\boldsymbol{S}_k\boldsymbol{e}-\boldsymbol{e}\right|\right|\,||\boldsymbol{X}_k^{-1}\boldsymbol{d}|| \\
&\leq \sqrt{\beta}||\boldsymbol{X}_k^{-1}\boldsymbol{d}||
\end{aligned}$$

hence, $||\boldsymbol{X}_k^{-1}\boldsymbol{d}|| \leq \sqrt{\beta} < 1$.

- We next show that $\boldsymbol{x}^{k+1}$ and $(\boldsymbol{p}^{k+1}, \boldsymbol{s}^{k+1})$ are primal and dual feasible Since $\boldsymbol{A}\boldsymbol{d} = 0$, we have

$$\boldsymbol{A}\boldsymbol{x}^{k+1} = \boldsymbol{b}$$

$$\boldsymbol{x}^{k+1} = \boldsymbol{x}^k + \boldsymbol{d} = \boldsymbol{X}_k(\boldsymbol{e} + \boldsymbol{X}_k^{-1}\boldsymbol{d}) > \boldsymbol{0},$$

because $||\boldsymbol{X}_k^{-1}\boldsymbol{d}|| < 1$

$$\boldsymbol{A}'\boldsymbol{p}^{k+1} + \boldsymbol{s}^{k+1} = \boldsymbol{c},$$

by construction and

$$\boldsymbol{s}^{k+1} = \boldsymbol{c} - \boldsymbol{A}'\boldsymbol{p}^{k+1} = \mu^{k+1}\boldsymbol{X}_k^{-1}(\boldsymbol{e} - \boldsymbol{X}_k^{-1}\boldsymbol{d}) > \boldsymbol{0},$$

because $||\boldsymbol{X}_k^{-1}\boldsymbol{d}|| < 1$

- 

$$\begin{aligned} x_j^{k+1} &= x_j^k\left(1 + \frac{d_j}{x_j^k}\right), \\ s_j^{k+1} &= \frac{\mu^{k+1}}{x_j^k}\left(1 - \frac{d_j}{x_j^k}\right). \end{aligned}$$

Therefore,

$$\begin{aligned} \frac{1}{\mu^{k+1}} x_j^{k+1} s_j^{k+1} - 1 &= \frac{1}{\mu^{k+1}} x_j^k \left(1 + \frac{d_j}{x_j^k}\right) \frac{\mu^{k+1}}{x_j^k} \left(1 - \frac{d_j}{x_j^k}\right) - 1 \\ &= -\left(\frac{d_j}{x_j^k}\right)^2. \end{aligned}$$

- $\boldsymbol{D} = \text{diag}(d_1, \ldots, d_n)$, $||\boldsymbol{u}||_1 = \sum_i |u_i|$. Note that $||\boldsymbol{u}|| \le ||\boldsymbol{u}||_1$

$$\begin{aligned} \left|\left|\frac{1}{\mu^{k+1}}\boldsymbol{X}_{k+1}\boldsymbol{S}_{k+1}\boldsymbol{e} - \boldsymbol{e}\right|\right| &= \left|\left|\boldsymbol{X}_k^{-2}\boldsymbol{D}^2\boldsymbol{e}\right|\right| \\ &\le \left|\left|\boldsymbol{X}_k^{-2}\boldsymbol{D}^2\boldsymbol{e}\right|\right|_1 \\ &= \boldsymbol{e}'\boldsymbol{X}_k^{-2}\boldsymbol{D}^2\boldsymbol{e} \\ &= \boldsymbol{e}'\boldsymbol{D}\boldsymbol{X}_k^{-2}\boldsymbol{D}\boldsymbol{e} \\ &= \boldsymbol{d}'\boldsymbol{X}_k^{-2}\boldsymbol{d} \\ &= \left|\left|\boldsymbol{X}_k^{-1}\boldsymbol{d}\right|\right|^2 \\ &\le (\sqrt{\beta})^2 \\ &= \beta, \end{aligned}$$

and hence the induction is complete.

- Since at every iteration

$$\left|\left|\frac{1}{\mu^k}\boldsymbol{X}_k\boldsymbol{S}_k\boldsymbol{e} - \boldsymbol{e}\right|\right| \le \beta$$

$$-\beta \le \frac{1}{\mu^k} x_j^k s_j^k - 1 \le \beta$$

$$n\mu^k(1 - \beta) \le (\boldsymbol{s}^k)'\boldsymbol{x}^k \le n\mu^k(1 + \beta)$$

- $$\mu^k = \alpha^k \mu^0 = \left(1 - \frac{\sqrt{\beta} - \beta}{\sqrt{\beta} + \sqrt{n}}\right)^k \mu^0 \leq e^{-k\frac{\sqrt{\beta}-\beta}{\sqrt{\beta}+\sqrt{n}}} \mu^0$$

- After

$$\left\lceil \frac{\sqrt{\beta} + \sqrt{n}}{\sqrt{\beta} - \beta} \log \frac{\mu^0 n(1+\beta)}{\epsilon} \right\rceil \leq \left\lceil \frac{\sqrt{\beta} + \sqrt{n}}{\sqrt{\beta} - \beta} \log \frac{(\boldsymbol{s}^0)'\boldsymbol{x}^0(1+\beta)}{\epsilon(1-\beta)} \right\rceil = K$$

  iterations, the primal barrier algorithm finds primal and dual solutions $\boldsymbol{x}^K$, $(\boldsymbol{p}^K, \boldsymbol{s}^K)$, that have duality gap $(\boldsymbol{s}^K)'\boldsymbol{x}^K$ less than or equal to $\epsilon$

# 7 Complexity

Slide 24

- Work per iteration involves solving a linear system with $m+n$ equations in $m+n$ unknowns. Given that $m \leq n$, the work per iteration is $O(n^3)$.

- $\epsilon_0 = (\boldsymbol{s}^0)'\boldsymbol{x}^0$: initial duality gap. Algorithm needs

$$O\left(\sqrt{n} \log \frac{\epsilon_0}{\epsilon}\right)$$

  iterations to reduce the duality gap from $\epsilon_0$ to $\epsilon$, with $O(n^3)$ arithmetic operations per iteration.

6.251J / 15.081J Introduction to Mathematical Programming
Fall 2009