15.093J Optimization Methods

Lecture 22: Barrier Interior Point Algorithms

1 Outline

SLIDE 1

1. Barrier Methods
2. The Central Path
3. Approximating the Central Path
4. The Primal Barrier Algorithm
5. The Primal-Dual Barrier Algorithm
6. Computational Aspects of IPMs

2 Barrier methods

SLIDE 2

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

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

2.1 Strategy

SLIDE 3

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

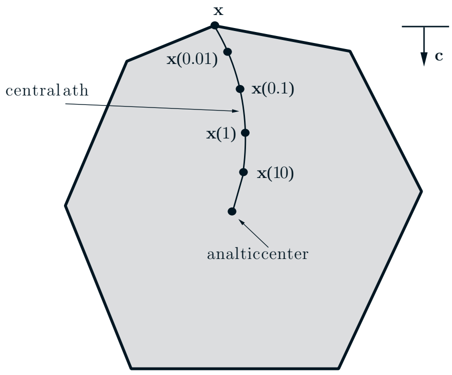
$$G(\boldsymbol{x}) = -\sum_{j=1}^{p} \log(-g_j(\boldsymbol{x})),\ G(\boldsymbol{x}) = -\sum_{j=1}^{p} \frac{1}{g_j(\boldsymbol{x})}$$

SLIDE 4

- Consider a sequence of μ^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 If 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

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

Barrier problem:

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

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

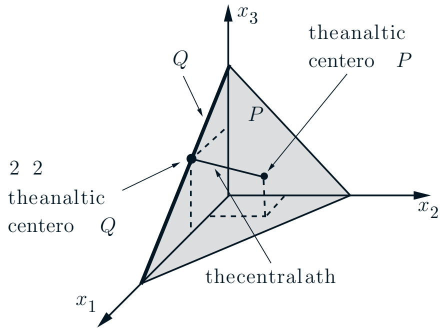
3 Central Path

SLIDE 6

- As μ 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*

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

SLIDE 7



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

3.1 Example

Slide 8

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

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

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

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

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

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

$\boldsymbol{X} = \operatorname{diag}(x_1, \ldots, x_n)$

SLIDE 14

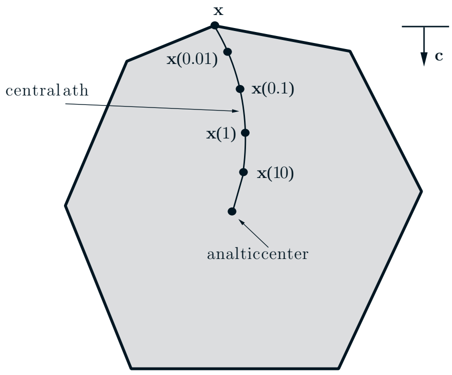
Approximating problem:

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

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:

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

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 μ 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

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

  $0<\beta<1$, $\boldsymbol{X}=\mathrm{diag}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)$ $\boldsymbol{S}=\mathrm{diag}(\boldsymbol{s}_1,\ldots,\boldsymbol{s}_n)$

- 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) μ^0, and α, 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.

5.1 Correctness

SLIDE 22

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

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

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:

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

5.2 Complexity

SLIDE 23

- 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 ϵ_0 to ϵ, with $O(n^3)$ arithmetic operations per iteration.

6 The Primal-Dual Barrier Algorithm

6.1 Optimality Conditions

SLIDE 24

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

$\boldsymbol{X}(\mu) = \text{diag}\big(x_1(\mu), \ldots, x_n(\mu)\big)$, $\boldsymbol{S}(\mu) = \text{diag}\big(s_1(\mu), \ldots, s_n(\mu)\big)$

6.2 Solving Equations

SLIDE 25

$$\boldsymbol{F}(\boldsymbol{z}) = \begin{bmatrix} \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} \\ \boldsymbol{A}'\boldsymbol{p} + \boldsymbol{s} - \boldsymbol{c} \\ \boldsymbol{X}\boldsymbol{S}\boldsymbol{e} - \mu\boldsymbol{e} \end{bmatrix}$$

$\boldsymbol{z} = (\boldsymbol{x}, \boldsymbol{p}, \boldsymbol{s})$, $r = 2n + m$

Solve

$$\boldsymbol{F}(\boldsymbol{z}^*) = \boldsymbol{0}$$

6.2.1 Newton's method

SLIDE 26

$$\boldsymbol{F}(\boldsymbol{z}^k + \boldsymbol{d}) \approx \boldsymbol{F}(\boldsymbol{z}^k) + \boldsymbol{J}(\boldsymbol{z}^k)\boldsymbol{d}$$

Here $\boldsymbol{J}(\boldsymbol{z}^k)$ is the $r \times r$ Jacobian matrix whose (i,j)th element is given by

$$\left.\frac{\partial F_i(\boldsymbol{z})}{\partial z_j}\right|_{\boldsymbol{z}=\boldsymbol{z}^k}$$

$$\boldsymbol{F}(\boldsymbol{z}^k) + \boldsymbol{J}(\boldsymbol{z}^k)\boldsymbol{d} = \boldsymbol{0}$$

Set $\boldsymbol{z}^{k+1} = \boldsymbol{z}^k + \boldsymbol{d}$ ($\boldsymbol{d}$ is the *Newton direction*) SLIDE 27

$(\boldsymbol{x}^k, \boldsymbol{p}^k, \boldsymbol{s}^k)$ current primal and dual feasible solution
Newton direction $\boldsymbol{d} = (\boldsymbol{d}_x^k, \boldsymbol{d}_p^k, \boldsymbol{d}_s^k)$

$$\begin{bmatrix} \boldsymbol{A} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{A}' & \boldsymbol{I} \\ \boldsymbol{S}_k & \boldsymbol{0} & \boldsymbol{X}_k \end{bmatrix} \begin{bmatrix} \boldsymbol{d}_x^k \\ \boldsymbol{d}_p^k \\ \boldsymbol{d}_s^k \end{bmatrix} = - \begin{bmatrix} \boldsymbol{A}\boldsymbol{x}^k - \boldsymbol{b} \\ \boldsymbol{A}'\boldsymbol{p}^k + \boldsymbol{s}^k - \boldsymbol{c} \\ \boldsymbol{X}_k \boldsymbol{S}_k \boldsymbol{e} - \mu^k \boldsymbol{e} \end{bmatrix}$$

6.2.2 Step lengths

SLIDE 28

$$\begin{aligned} \boldsymbol{x}^{k+1} &= \boldsymbol{x}^k + \beta_P^k \boldsymbol{d}_x^k \\ \boldsymbol{p}^{k+1} &= \boldsymbol{p}^k + \beta_D^k \boldsymbol{d}_p^k \\ \boldsymbol{s}^{k+1} &= \boldsymbol{s}^k + \beta_D^k \boldsymbol{d}_s^k \end{aligned}$$

To preserve nonnegativity, take

$$\begin{aligned} \beta_P^k &= \min\left\{1, \alpha \min_{\{i \mid (d_x^k)_i < 0\}} \left(-\frac{x_i^k}{(d_x^k)_i} \right) \right\}, \\ \beta_D^k &= \min\left\{1, \alpha \min_{\{i \mid (d_s^k)_i < 0\}} \left(-\frac{s_i^k}{(d_s^k)_i} \right) \right\}, \end{aligned}$$

$0 < \alpha < 1$

6.3 The Algorithm

SLIDE 29

1. (Initialization) Start with $\boldsymbol{x}^0 > \boldsymbol{0}$, $\boldsymbol{s}^0 > \boldsymbol{0}$, $\boldsymbol{p}^0$, and set $k = 0$
2. (Optimality test) If $(\boldsymbol{s}^k)'\boldsymbol{x}^k < \epsilon$ stop; else go to Step 3.
3. (Computation of Newton directions)

$$\begin{aligned} \mu^k &= \frac{(\boldsymbol{s}^k)'\boldsymbol{x}^k}{n} \\ \boldsymbol{X}_k &= \mathrm{diag}(x_1^k, \ldots, x_n^k) \\ \boldsymbol{S}_k &= \mathrm{diag}(s_1^k, \ldots, s_n^k) \end{aligned}$$

Solve linear system

$$\begin{bmatrix} \boldsymbol{A} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{A}' & \boldsymbol{I} \\ \boldsymbol{S}_k & \boldsymbol{0} & \boldsymbol{X}_k \end{bmatrix} \begin{bmatrix} \boldsymbol{d}_x^k \\ \boldsymbol{d}_p^k \\ \boldsymbol{d}_s^k \end{bmatrix} = - \begin{bmatrix} \boldsymbol{A}\boldsymbol{x}^k - \boldsymbol{b} \\ \boldsymbol{A}'\boldsymbol{p}^k + \boldsymbol{s}^k - \boldsymbol{c} \\ \boldsymbol{X}_k \boldsymbol{S}_k \boldsymbol{e} - \mu^k \boldsymbol{e} \end{bmatrix}$$

SLIDE 30

4. (Find step lengths)

$$\beta_P^k = \min\left\{1, \alpha \min_{\{i|(d_x^k)_i<0\}} \left(-\frac{x_i^k}{(d_x^k)_i}\right)\right\}$$
$$\beta_D^k = \min\left\{1, \alpha \min_{\{i|(d_s^k)_i<0\}} \left(-\frac{s_i^k}{(d_s^k)_i}\right)\right\}$$

5. (Solution update)

$$\begin{aligned} \boldsymbol{x}^{k+1} &= \boldsymbol{x}^k + \beta_P^k \boldsymbol{d}_x^k \\ \boldsymbol{p}^{k+1} &= \boldsymbol{p}^k + \beta_D^k \boldsymbol{d}_p^k \\ \boldsymbol{s}^{k+1} &= \boldsymbol{s}^k + \beta_D^k \boldsymbol{d}_s^k \end{aligned}$$

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

6.4 Insight on behavior

Slide 31

- Affine Scaling

$$\boldsymbol{d}_{\text{affine}} = -\boldsymbol{X}^2\Big(\boldsymbol{I} - \boldsymbol{A}'(\boldsymbol{A}\boldsymbol{X}^2\boldsymbol{A}')^{-1}\boldsymbol{A}\boldsymbol{X}^2\Big)\boldsymbol{c}$$

- Primal barrier

$$\boldsymbol{d}_{\text{primal}-\text{barrier}} = \Big(\boldsymbol{I} - \boldsymbol{X}^2\boldsymbol{A}'(\boldsymbol{A}\boldsymbol{X}^2\boldsymbol{A}')^{-1}\boldsymbol{A}\Big)\Big(\boldsymbol{X}\boldsymbol{e} - \frac{1}{\mu}\boldsymbol{X}^2\boldsymbol{c}\Big)$$

- For $\mu = \infty$

$$\boldsymbol{d}_{\text{centering}} = \Big(\boldsymbol{I} - \boldsymbol{X}^2\boldsymbol{A}'(\boldsymbol{A}\boldsymbol{X}^2\boldsymbol{A}')^{-1}\boldsymbol{A}\Big)\boldsymbol{X}\boldsymbol{e}$$

- Note that

$$\boldsymbol{d}_{\text{primal}-\text{barrier}} = \boldsymbol{d}_{\text{centering}} + \frac{1}{\mu}\boldsymbol{d}_{\text{affine}}$$

- When μ is large, then the centering direction dominates, i.e., in the beginning, the barrier algorithm takes steps towards the analytic center
- When μ is small, then the affine scaling direction dominates, i.e., towards the end, the barrier algorithm behaves like the affine scaling algorithm

7 Computational aspects of IPMs

Slide 32

Simplex vs. Interior point methods (IPMs)

- Simplex method tends to perform poorly on large, massively degenerate problems, whereas IP methods are much less affected.
- Key step in IPMs

$$\big(\boldsymbol{A}\boldsymbol{X}_k^2\boldsymbol{A}'\big)\boldsymbol{d} = \boldsymbol{f}$$

- In implementations of IPMs $\boldsymbol{A}\boldsymbol{X}_k^2\boldsymbol{A}'$ is usually written as

$$\boldsymbol{A}\boldsymbol{X}_k^2\boldsymbol{A}' = \boldsymbol{L}\boldsymbol{L}',$$

  where $\boldsymbol{L}$ is a square lower triangular matrix called the *Cholesky factor*

- Solve system

$$\big(\boldsymbol{A}\boldsymbol{X}_k^2\boldsymbol{A}'\big)\boldsymbol{d} = \boldsymbol{f}$$

  by solving the triangular systems

$$\boldsymbol{L}\boldsymbol{y} = \boldsymbol{f}, \qquad \boldsymbol{L}'\boldsymbol{d} = \boldsymbol{y}$$

- The construction of $\boldsymbol{L}$ requires $O(n^3)$ operations; but the actual computational effort is highly dependent on the sparsity (number of nonzero entries) of $\boldsymbol{L}$

- Large scale implementations employ heuristics (reorder rows and columns of $\boldsymbol{A}$) to improve sparsity of $\boldsymbol{L}$. If $\boldsymbol{L}$ is sparse, IPMs are stronger.

8 Conclusions

SLIDE 33

- IPMs represent the present and future of Optimization.
- Very successful in solving very large problems.
- Extend to general convex problems

15.093J / 6.255J Optimization Methods
Fall 2009