Convex Optimization
CMU-10725

Barrier Methods

Barnabás Póczos & Ryan Tibshirani

Barrier Methods (Interior methods)

Barrier Methods

$$\min_{x \in \mathcal{S}} f(x), \qquad (P)$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is continuous, and $\mathcal{S}$ is a constraint set in $\mathbb{R}^n$

Barrier program: replace (P) with the unconstrained problem (B(c)):

$$\min f(x) + \frac{1}{c} B(x)$$

where $c > 0$

Barrier function $B : \mathbb{R}^n \to \mathbb{R}$:

(i) B is continuous,

(ii) $B(x) \geq 0$ for all $x \in int(\mathcal{S})$

(iii) $B(x) \to \infty$ as $x \to \partial\mathcal{S}$

Barrier term: we don't let the algorithm leave S [Interior method]

Inequality Constraints

$$\min_{x \in \mathcal{S}} f(x) \quad \mathcal{S} = \{x : g_i(x) \leq 0, i = 1, 2, \ldots, p\}$$

Example: $g_1(x) = x - b,\ g_2(x) = a - x$

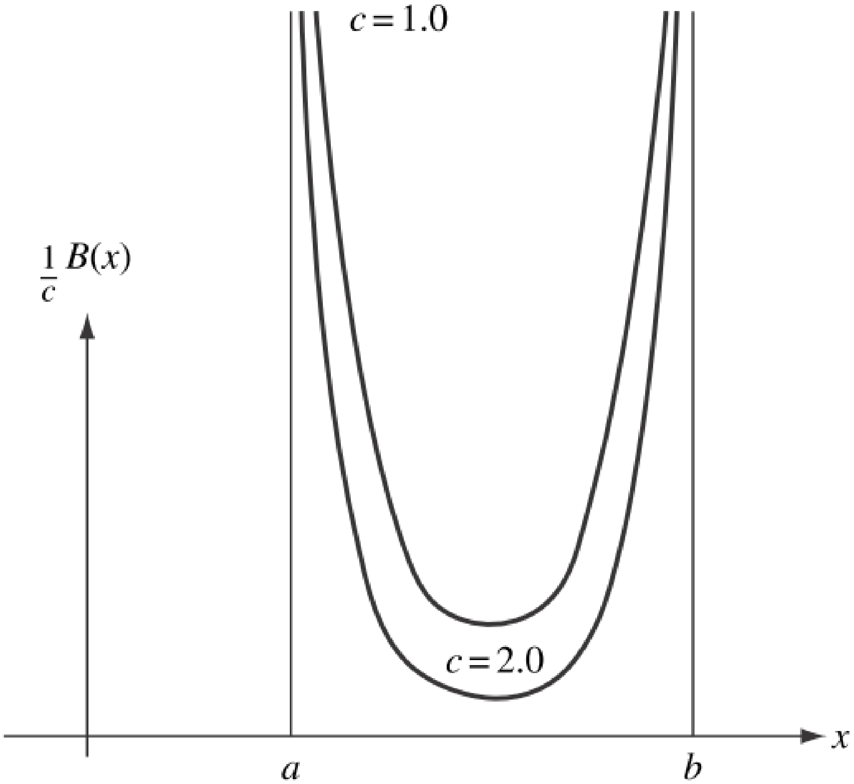
$a \leq x$
$x \leq b$

A useful barrier function in this case is:

$$B(x) = -\sum_{i=1}^{m} \frac{1}{g_i(x)}$$

Barrier program (B(c)):

$$\min f(x) + \frac{1}{c} B(x)$$



Logarithmic Barrier function

Problem: (P)

$$\min_{x \in \mathcal{S}} f(x) \qquad \mathcal{S} = \{x : g_i(x) \leq 0, i = 1, 2, \ldots, m\}$$

Definition: Logarithmic barrier function

$$B(x) = -\sum_{i=1}^{m} \mathrm{LOG}(-g_i(x))$$

Barrier method:

$$0 \leq c_1 < c_2 < \ldots < c_k < c_{k+1} < \ldots \to \infty$$

$$B(c_k) \qquad x_k = \underset{x \in \mathrm{INT}(S)}{\mathrm{ARGMIN}} \underbrace{f(x) + \frac{1}{c_k} B(x)}_{r(c_k, x)}$$

$$1/c_k = \mu_k$$

B(c) is still a constrained problem and looks more complicated than (P)

Advantage: B(c) can be solved with unconstrained optimization tools

Barrier Method

Penalty parameter:

LET $0 < c_1 < c_2 < \ldots < c_k < c_{k+1} < \ldots \to \infty$

Barrier program:

$$B(c_k): \quad x_k = \underset{x_k}{\mathrm{ARGMIN}} \; \underbrace{f(x_k) + \frac{1}{c_k} B(x_k)}_{r(c_k,\, x_k)}$$

Barrier Lemma:

1. $r(c_k, x_k) \geq r(c_{k+1}, x_{k+1})$
2. $B(x_k) \leq B(x_{k+1})$
3. $f(x_k) \geq f(x_{k+1})$
4. $f(x^*) \leq f(x_{k+1}) \leq f(x_k) \leq r(c_k, x_k)$

Proof of Barrier Lemma (1)

PROVE: $r(c_k, x_k) \geq r(c_{k+1}, x_{k+1})$

PROOF:

$$r(c_k, x_k) \doteq f(x_k) + \frac{1}{c_k} B(x_k) \underset{c_k < c_{k+1}}{\geq} f(x_k) + \frac{1}{c_{k+1}} B(x_k)$$

$$\underset{x_{k+1} \text{ is optimal with } c_{k+1}}{\geq} f(x_{k+1}) + \frac{1}{c_{k+1}} B(x_{k+1}) = r(c_{k+1}, x_{k+1})$$

Proof of Barrier Lemma (2)

PROVE $$B(x_{k}) \le B(x_{k+1})$$

PROOF:

$$(*1)\quad f(x_k) + \frac{1}{c_k} B(x_k) \le f(x_{k+1}) + \frac{1}{c_k} B(x_{k+1})$$

x_k IS OPTIMAL WITH c_k

$$(*2)\quad f(x_{k+1}) + \frac{1}{c_{k+1}} B(x_{k+1}) \le f(x_k) + \frac{1}{c_{k+1}} B(x_k)$$

c_{k+1} IS OPTIMAL WITH x_k

$(*1) + (*2) \Rightarrow$

$$\frac{1}{c_k} B(x_k) + \frac{1}{c_{k+1}} B(x_{k+1}) \le \frac{1}{c_k} B(x_{k+1}) + \frac{1}{c_{k+1}} B(x_k)$$

$$\Rightarrow \underbrace{\left(\frac{1}{c_k} - \frac{1}{c_{k+1}}\right)}_{\vee\, 0} B(x_k) \le \underbrace{\left(\frac{1}{c_k} - \frac{1}{c_{k+1}}\right)}_{\vee\, 0} B(x_{k+1})$$

$$\Rightarrow B(x_k) \le B(x_{k+1})$$

Proof of Barrier Lemma (3)

PROVE $f(x_k) \geq f(x_{k+1})$

PROOF:

$$f(x_k) + \frac{1}{c_{k+1}} B(x_k) \geq f(x_{k+1}) + \frac{1}{c_{k+1}} B(x_{k+1})$$

x_{k+1} IS OPTIMAL WITH c_{k+1}

$$\geq f(x_{k+1}) + \frac{1}{c_{k+1}} B(x_k)$$

$B(x_{k+1}) \geq B(x_k)$ FROM LEMMA(2)

$$\Rightarrow f(x_k) \geq f(x_{k+1}) \quad \square$$

Proof of Barrier Lemma (4)

PROVE

$$f(x^*) \le f(x_{k+1}) \underset{\substack{\uparrow \\ \text{LEMMA (3)}}}{\le} f(x_k) \le r(c_k, x_k)$$

PROOF:

$$f(x_*) \underset{\substack{\uparrow \\ x^* \text{ IS OPTIMAL IN } S \\ x_k \in S}}{\le} f(x_k) \le f(x_k) + \frac{1}{c_k} \underbrace{B(x_k)}_{\ge 0} \doteq r(c_k, x_k)$$

$\square$

Primal Logarithmic Barrier Method for LP

Primal Log Barrier Method for LP

Primal problem

$$P \quad \min_{x} \; c^T x = z$$

SUBJECT TO $Ax = b$, $x \geq 0$

$x \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $m < n$

Log barrier problem

$$P(\mu) \quad \min_{x \in \mathbb{R}^n} \; f(x) = c^T x - \mu \sum_{j=1}^{n} \log(x_j)$$

S.T. $Ax = b$, $x > 0$

Derivatives: $g(x) \doteq \nabla f(x) = c - \mu D_x^{-1} e$ WHERE $D_x = \mathrm{DIAG}(x) \in \mathbb{R}^{n \times n}$, $e = (1, \ldots, 1)^T$

$G(x) = \nabla^2 f(x) = \mu D_x^{-2}$ POS DEFINITE SINCE $x > 0$

Primal Log Barrier Method for LP

Feasible solution of the log barrier problem

LET $\bar{x}$ BE A FEASIBLE SOLUTION TO $B(\mu)$

$\Rightarrow A\bar{x} = b$

WE ARE IN $\bar{x}$ FEASIBLE, NOT OPTIMAL SOLUTION.

FIND A BETTER SOLUTION $x^{+} = \bar{x} + \Delta x$

Quadratic approximation of f AT $\bar{x}$:

$$f(x) \approx f(\bar{x}) + \underbrace{\nabla f(\bar{x})^T}_{g(x)^T} \underbrace{(x-\bar{x})}_{\Delta x} + \frac{1}{2} \underbrace{(x-\bar{x})}_{\Delta x} \underbrace{\nabla^2 f(\bar{x})}_{G(\bar{x})} \underbrace{(x-\bar{x})}_{\Delta x}$$

$$= f(\bar{x}) + \underbrace{g(x)^T \Delta x + \frac{1}{2} \Delta x^T G(\bar{x}) \Delta x}_{Q(\Delta x)}$$

QUADRATIC APPROXIMATION

Primal Log Barrier Method for LP

New quadratic problem:

$$\min_{\Delta x} Q(\Delta x)$$

$$\text{s.t. } A(\underbrace{\bar{x} + \Delta x}_{x}) = b \qquad x = \bar{x} + \Delta x$$

$$\Rightarrow \min_{\Delta x} Q(\Delta x)$$

$$\text{s.t. } A\Delta x = 0 \quad \text{since } A\bar{x} = b$$

Lagrange problem:

$$L(x, \pi_x) = g(\bar{x})^T \Delta x + \frac{1}{2} \Delta x^T G(\bar{x}) \Delta x - \pi_x^T [A \Delta x]$$

Primal Log Barrier Method for LP

Lagrange problem:

$$L(x,\pi_x) = g(\bar{x})^T \Delta x + \tfrac{1}{2}\Delta x^T G(\bar{x})\Delta x - \pi_x^T [A\Delta x]$$

In the optimum:

$$0 = \frac{\partial L(x,\pi_x)}{\partial \Delta x} = \underbrace{g(\bar{x})}_{c-\mu D_x^{-1}e} + \underbrace{G(\bar{x})}_{\mu D_x^{-2}}\Delta x - A^T\pi_x$$

$$\Rightarrow 0 = c - \mu D_x^{-1}e + \mu D_x^{-2}\Delta x - A^T \pi_x$$

$$\Rightarrow \underbrace{\begin{pmatrix} \mu D_x^{-2} & A^T \\ A & 0 \end{pmatrix}}_{\mathbb{R}^{(n+m)\times(n+m)}} \begin{pmatrix} -\Delta x \\ \pi_x \end{pmatrix} = \begin{pmatrix} c-\mu D_x^{-1}e \\ 0 \end{pmatrix} \Rightarrow \begin{matrix} \Delta x \leftarrow \mathbb{R}^n \\ \pi_x \leftarrow \mathbb{R}^m \end{matrix}$$

Interior-point algorithm

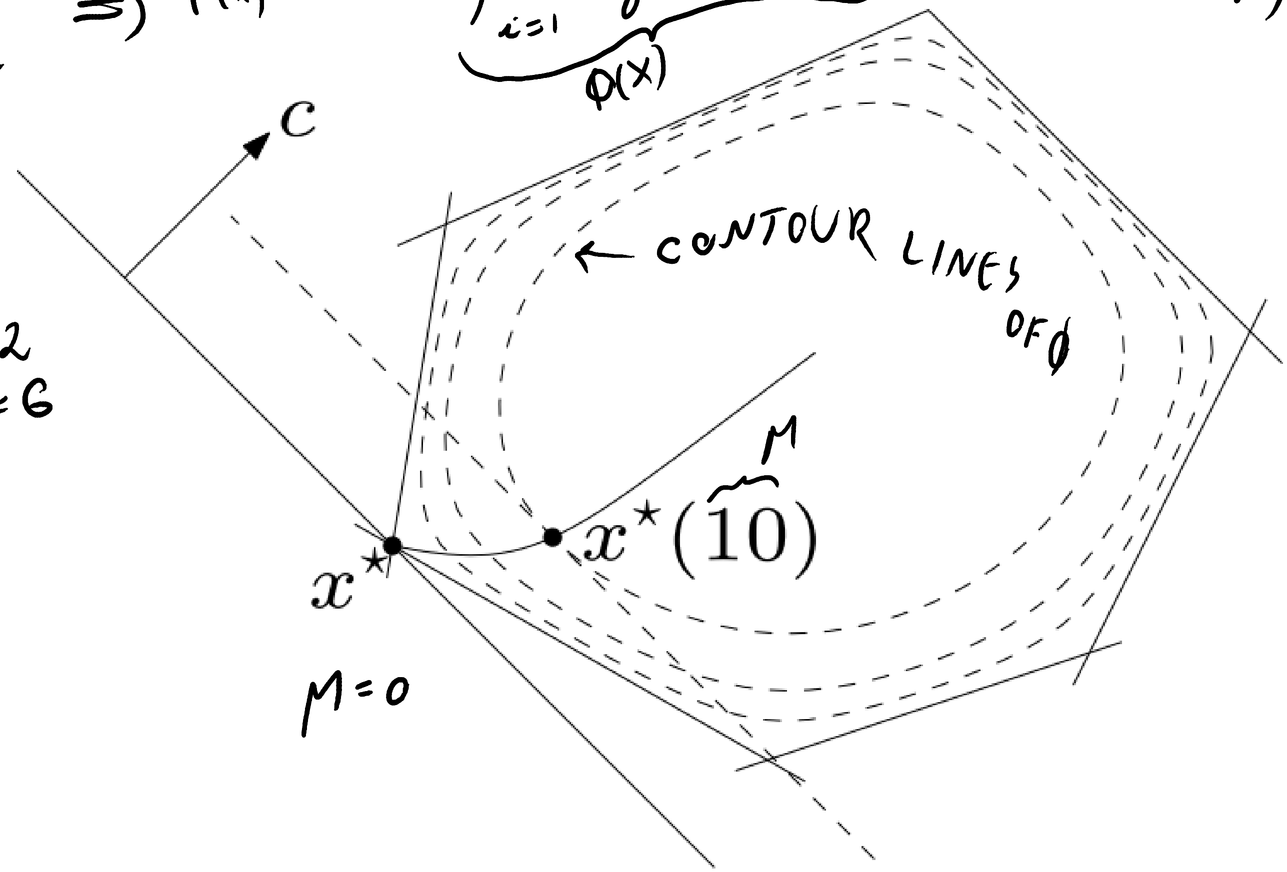
$$\min_{x \in \mathbb{R}^n} f(x) = c^T x - \mu \sum_{j=1}^{n} \log(x_j) \Rightarrow x^*(\mu)$$

$$\text{s.t. } Ax = b$$

$$\min c^T x \quad \Rightarrow f(x) = c^T x - \mu \underbrace{\sum_{i=1}^{m} \log(b_i - a_i^T x)}_{\phi(x)} \Rightarrow x^*(\mu)$$

$$\text{s.t. } Ax \le b$$

LP WITH $n=2$, $m=6$



Primal - Dual Logarithmic Barrier Method for LP

Symmetric Primal and Dual Forms

Symmetric form I [wikipedia]

$$(P)\quad \max_x\ c^T x \quad \text{s.t.}\quad Ax \le b,\ x \ge 0$$

$$(D)\quad \min_y\ y^T b \quad \text{s.t.}\quad y^T A \ge c^T,\ y \ge 0$$

Duality gap:

$$\underbrace{c^T x \le (y^T A) x}_{\ge 0} \le y^T b$$

$$c^T x \le (y^T A)\underset{\substack{\vee\!\shortmid \\ 0}}{x} \le \underset{\substack{\vee\!\shortmid \\ 0}}{y^T} b, \quad \underset{\substack{\vee\!\shortmid \\ 0}}{x}$$

Symmetric form II [Luenberger & Ye]

$$(P)\quad \min\ c^T x \quad \text{s.t.}\quad Ax \ge b,\ x \ge 0$$

$$(D)\quad \max_y\ y^T b \quad \text{s.t.}\quad y^T A \le c^T,\ y \ge 0$$

Duality gap: $y^T b \le y^T A x \le c^T x$

Asymmetric Primal and Dual Forms

Asymmetric form I [Jensen & Jonathan] $Ax + s = b \quad s \geq 0$

$$(P) \quad \max \; c^T x \quad \text{s.t} \quad Ax \leq b, \quad x \in \mathbb{R}^n$$

$$(D) \quad \min_y \; y^T b \quad \text{s.t.} \quad y^T A = c^T, \quad y \geq 0$$

Duality gap:

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

Asymmetric form II [Luenberger & Ye, R. Freund] $y^T A + s = c \quad s \geq 0$

$$(P) \quad \min \; c^T x \quad \text{s.t} \quad Ax = b, \quad x \geq 0$$

$$(D) \quad \max \; y^T b \quad \text{s.t} \quad y^T A \leq c^T, \quad y \in \mathbb{R}^m$$

In what follows, we will use this form.

Duality gap:

$$c^T x - b^T y = y^T A x + s^T x - x^T A^T y = s^T x \geq 0$$

KKT conditions

Log barrier problem [We already know this]

$$P(\mu) \qquad \min_{x \in \mathbb{R}^n} f(x) = c^T x - \mu \sum_{j=1}^{n} \log(x_j)$$

$$\text{s.t.} \quad Ax = b$$

$$x > 0$$

Derivatives: $g(x) \doteq \nabla f(x) = c - \mu D_x^{-1} e$ WHERE $D_x = \mathrm{DIAG}(x) \in \mathbb{R}^{n \times n}$, $e = (1, \dots, 1)^T$

KKT stationarity condition of the log barrier problem

$$\underbrace{\nabla_x f(x)}_{[c - \mu D_x^{-1} e]^T} + y^T \overbrace{\nabla_x (b - Ax)}^{-A} = 0$$

$$Ax = b$$

$$x > 0$$

$$\Rightarrow \quad c - \mu D_x^{-1} e = A^T y$$

Stationarity Condition Rewritten

The stationarity condition can be rewritten

$$c - \underbrace{\mu D_x^{-1} e}_{s} = A^T y$$

$$\Rightarrow \mu D_x^{-1} e = s$$

$$\frac{1}{\mu} D_x s = e$$

$$\Rightarrow \frac{1}{\mu} D_x D_s e = e$$

KKT Conditions

KKT Conditions

$$\left.\begin{array}{l} Ax = b,\ x > 0 \\ A^T y + s = c,\ s > 0 \\ \frac{1}{\mu} D_x D_s e = e \end{array}\right\} (*KKT)$$

Lemma

IF (x, y, s) IS A SOLUTION OF $(*KKT)$

$\Rightarrow$
- x IS FEASIBLE FOR P
- (y, s) IS FEASIBLE FOR D
- THE DUALITY GAP IS $x^T s = e^T \underbrace{D_x D_s e}_{\mu e} = \mu n$

WE WILL SOLVE $P(\mu)$ AND LET $\mu \to 0$

Approximation of the KKT conditions

$$\left.\begin{aligned} &(a)\quad Ax = b,\ x > 0 \\ &(b)\quad A^T y + s = c \\ &\qquad \frac{1}{\mu} D_x D_s e = e \end{aligned}\right\} \quad (* \text{KKT})$$

This is not linear in x and s!

β-approximation of the stationarity condition:

$$(c)\quad \left\| \frac{1}{\mu} D_x s - e \right\| \le \beta$$

THE PRIMAL-DUAL ALGORITHM SOLVES THE SYSTEM (a), (b), (c) WITH NEWTON METHOD FOR A GIVEN β AND μ

Duality Gap of beta approximation

Lemma: [duality gap after β-approximation**]**

- $Ax = b,\ x > 0$
- $A^T y + s = c,$
- $\| \frac{1}{\mu} D_x s - e \| \le \beta$

$\Bigg\}$ $(* \text{KKT} - \beta)$

WE DON'T REQUIRE $s \geq 0$!

- IF $(\bar{x}, \bar{y}, \bar{s})$ IS A β-APPROXIMATE SOLUTION OF $(*\text{-KKT}\beta)$
- & $\beta < 1$

$\Rightarrow$ DUALITY GAP CAN BE LOWER & UPPER BOUNDED:

- $n\mu(1-\beta) \le \underbrace{c^T \bar{x} - b^T \bar{y}}_{\text{DUALITY GAP}} \le n\mu(1+\beta)$
- $\bar{x}$ IS FEASIBLE FOR (P), $(\bar{y}, \bar{s})$ IS FEASIBLE FOR (D)

Duality Gap of beta approximation

Proof:

PRIMAL FEASIBILITY: $\bar{x} \geq 0$ ✓

DUAL FEASIBILITY: $\bar{y} \in \mathbb{R}^m$ ✓

WE ONLY NEED TO SHOW THAT $\bar{s} \geq 0$

FROM $\left\| \frac{1}{\mu} D_x s - e \right\| \leq \beta$

$$\Rightarrow -\beta \leq \frac{1}{\mu} x_j s_j - 1 \leq \beta$$

$$\Rightarrow 0 < (1-\beta)\mu \leq \underset{\overset{\vee|}{0}}{x_j} s_j \leq \mu(\beta+1)$$

$$\Rightarrow s_j > 0 \Rightarrow (\bar{y}, \bar{s}) \text{ IS D FEASIBLE} \qquad \square$$

DUALITY GAP:

$$n\mu(1-\beta) \leq \sum_{j=1}^{n} x_j s_j = \underbrace{x^T s}_{\text{DUALITY GAP}} \leq n\mu(\beta+1)$$

Primal-Dual Newton Step

Current iterate:

KKT:
- $A\bar{x} = b,\ \bar{x} > 0$
- $A^T\bar{y} + \bar{s} = c,\ \bar{s} > 0$
- $\frac{1}{\mu} D_{\bar{x}} D_{\bar{s}} e - e = 0$

After update:

$$A(\bar{x} + \Delta x) = b,\quad \bar{x} + \Delta x > 0$$

$$A^T(\bar{y} + \Delta y) + (\bar{s} + \Delta s) = c$$

$$\frac{1}{\mu}\left(D_{\bar{x}} + D_{\Delta x}\right)\left(D_{\bar{s}} + D_{\Delta s}\right) e - e = 0$$

Newton steps:

$$A\Delta x = 0$$

$$A^T\Delta y + \Delta s = 0$$

$$\underbrace{D_{\bar{x}} D_{\Delta s} e + D_{\Delta x} D_{\bar{s}} e}_{D_{\bar{x}}\Delta s + D_{\bar{s}}\Delta x} = \mu e - D_{\bar{x}} D_{\bar{s}} e - \underbrace{D_{\Delta x} D_{\Delta s} e}_{\text{NEGLECT THIS NONLINEAR TERM}}$$

Primal-Dual Newton System

Primal-Dual Newton System

$$\left.\begin{aligned} A\,\Delta x &= 0 \\ A^T \Delta y + \Delta s &= 0 \\ D_{\bar{s}}\,\Delta x + D_{\bar{x}}\,\Delta s &= \mu e - D_{\bar{x}} D_{\bar{s}}\, e \end{aligned}\right\} \Rightarrow \begin{matrix} \Delta x \\ \Delta y \\ \Delta s \end{matrix}$$

(* NEWTON)

The Primal Dual Algorithm

STEP 0. INITIALIZATION

(x^0, y^0, s^0) FEASIBLE SOLUTION

STEP 1. $(\bar{x}, \bar{y}, \bar{s}) = (x^k, y^k, s^k)$

STEP 2. NEWTON STEP

$$\left.\begin{array}{l} \bullet\ A\Delta x = 0 \\ \bullet\ A^T\Delta y + \Delta s = 0 \\ \bullet\ D_{\bar{s}}\Delta x + D_{\bar{x}}\Delta s = \mu e - D_{\bar{x}} D_{\bar{s}} e \end{array}\right\} \Rightarrow (\Delta x, \Delta s, \Delta y)$$

STEP 3. $x^{k+1} = \bar{x} + \Delta x$

$y^{k+1} = \bar{y} + \Delta y$

$s^{k+1} = \bar{s} + \Delta s$ $\quad \alpha < \alpha < 1$

STEP 4 $k = k+1$. $\mu_{k+1} = \alpha \mu_k$ $\left[\text{OR } \mu_{k+1} = \frac{1}{10}\frac{\bar{x}^T\bar{s}}{n}\right]$

GO BACK TO 1 $\quad$ CURRENT DUALITY GAP

Quadratic Convergence

Theorem [Local quadratic convergence of Primal-Dual Newton's method]:

- SUPPOSE THAT $(\bar{x}, \bar{y}, \bar{s})$ IS A β-APPROXIMATE SOLUTION OF $P(\mu)$ FOR SOME $0 < \beta < \frac{1}{2}$
- LET $(\Delta x, \Delta y, \Delta s)$ BE THE SOLUTION OF THE PRIMAL-DUAL NEWTON SYSTEM (* NEWTON)
- LET $(x', y', s') = (\bar{x}, \bar{y}, \bar{s}) + (\Delta x, \Delta y, \Delta s)$

$\Rightarrow (x', y', s')$ IS A $\frac{(1+\beta)}{(1-\beta)^2}\beta^2$-APPROX SOLUTION OF $P(\mu)$

Summary

- Penalty functions
- Barrier functions
- Primal log barrier method for LP
- Primal-Dual log barrier method for LP