Convex Optimization
CMU-10725
Barrier Methods

Barnabás Póczos & Ryan Tibshirani
Barrier Methods (Interior methods)
Barrier Methods

\[
\min_{x \in S} f(x), \quad (P)
\]

where \( f : \mathbb{R}^n \to \mathbb{R} \) is continuous, and \( S \) 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 \( x \in \text{int}(S) \)
(iii) \( B(x) \to \infty \) as \( x \to \partial S \)

**Barrier term:** we don't let the algorithm leave \( S \) [Interior method]
Inequality Constraints

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

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

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:** \((P)\)
\[
\min_{x \in S} f(x) \quad S = \{x : g_i(x) \leq 0, i = 1, 2, \ldots, m\}
\]

**Definition:** Logarithmic barrier function
\[
B(x) = -\sum_{i=1}^{m} \log(-g_i(x))
\]

**Barrier method:**
\[
0 < c_1 < c_2 < \ldots < c_\infty < c_{\infty+1} < \ldots \to \infty
\]
\[
B(c_\infty), \quad x_{\infty} = \arg\min_{x \in \text{INT}(S)} \left( f(x) + \frac{1}{c_\infty} B(x) \right)
\]
\[
\tau(c_\infty, x) \quad \lambda_{c_\infty} = \mu_{\infty}
\]

\(B(c)\) 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 \( 0 < C_1 < C_2 < \ldots < C_{k+1} \leq \ldots \rightarrow 0 \)

Barrier program:

\[
\bar{B}(C_{k+1}) : \quad x_{k+1} = \arg\min_{x_k} f(x_k) + \frac{1}{C_{k+1}} \Theta(x_k)
\]

Barrier Lemma:

1. \( \tau(C_{k+1}, x_{k+1}) \geq \tau(C_{k+1}, x_{k+1}) \)
2. \( B(x_k) \leq B(x_{k+1}) \)
3. \( f(x_k) \geq f(x_{k+1}) \)
4. \( f(x^*) \leq f(x_{k+1}) \leq f(x_k) \leq \tau(C_{k+1}, x_k) \)
Proof of Barrier Lemma (1)

\textbf{PROVE:} \quad \tau(C_{e_2}, X_{e_2}) \geq \tau(C_{e_2+1}, X_{e_2+1})

\textbf{Proof:}
\begin{align*}
\tau(C_{e_2}, X_{e_2}) &= f(X_{e_2}) + \frac{1}{C_{e_2}} B(X_{e_2}) \\
&\geq f(X_{e_2}) + \frac{1}{C_{e_2+1}} B(X_{e_2}) \\
& \quad \text{if } C_{e_2} > C_{e_2+1} \\
&= \tau(C_{e_2+1}, X_{e_2+1})
\end{align*}

\Rightarrow X_{e_2+1} \text{ is optimal with } C_{e_2+1}
Proof of Barrier Lemma (2)

**PROOF**

\[ B(x_{e_2}) \leq B(x_{e_2+1}) \]

**PROOF:**

\[(*)\quad f(x_{e_2}) + \frac{1}{C_{e_2}} B(x_{e_2}) \leq f(x_{e_2+1}) + \frac{1}{C_{e_2}} B(x_{e_2+1})\]

\[x_{e_2} \text{ is optimal with } C_{e_2}\]

\[(*)_{2}\quad f(x_{e_2+1}) + \frac{1}{C_{e_2+1}} B(x_{e_2+1}) \leq f(x_{e_2}) + \frac{1}{C_{e_2+1}} B(x_{e_2})\]

\[C_{e_2+1} \text{ is optimal with } x_{e_2}\]

\[(*)_{1} + (*)_{2} \Rightarrow \]

\[ \frac{1}{C_{e_2}} B(x_{e_2}) + \frac{1}{C_{e_2+1}} B(x_{e_2+1}) \leq \frac{1}{C_{e_2}} B(x_{e_2+1}) + \frac{1}{C_{e_2+1}} B(x_{e_2}) \]

\[ \Rightarrow \left( \frac{1}{C_{e_2}} - \frac{1}{C_{e_2+1}} \right) B(x_{e_2}) \leq \left( \frac{1}{C_{e_2}} - \frac{1}{C_{e_2+1}} \right) B(x_{e_2+1}) \]

\[ \Rightarrow B(x_{e_2}) \leq B(x_{e_2+1}) \]
Proof of Barrier Lemma (3)

\[ f(x_{e_2}) \geq f(x_{e_2+1}) \]

**Proof:**

\[
f(x_{e_2}) + \frac{1}{C_{e+1}} B(x_{e_2}) \geq f(x_{e+1}) + \frac{1}{C_{e+1}} B(x_{e+1})
\]

**Xe+1 is optimal with Ce+1**

\[
\Rightarrow f(x_{e+1}) + \frac{1}{C_{e+1}} B(x_{e_2}) \geq B(x_{e+1}) \quad \text{from Lemma (2)}
\]

\[ \Rightarrow f(x_{e_2}) \geq f(x_{e_2+1}) \quad \square \]
Proof of Barrier Lemma (4)

**Prove**
\[ f(x^*) \leq f(x_{k+1}) \leq f(x_k) \leq \tau(C_k, x_k) \]

**Lemma (3)**

**Proof:**
\[ f(x^*) \leq f(x_k) \leq f(x_0) + \frac{1}{C_k} B(x_0) \leq \tau(C_k, x_k) \]

\[ x^* \text{ is optimal in } S \]
\[ x_0 \in S \]
Primal Logarithmic Barrier Method for LP
Primal Log Barrier Method for LP

Primal problem

\[
\begin{align*}
\min_{x} & \quad C^T x = z \\
\text{subject to} & \quad A x = b \\
& \quad x \geq 0
\end{align*}
\]

Log barrier problem

\[
\begin{align*}
\min_{x \in \mathbb{R}^n} & \quad f(x) = C^T x - M \sum_{j=1}^{n} \log(x_j) \\
\text{s.t.} & \quad A x = b \\
& \quad x > 0
\end{align*}
\]

Derivatives:

\[
\begin{align*}
q(x) &= \nabla f(x) = C - M D_x^{-1} e \\
G(x) &= \nabla^2 f(x) = M D_x^{-2} \quad \text{pos definite since } x > 0
\end{align*}
\]

WHERE \( D_x = \text{Diag}(x) \)
Primal Log Barrier Method for LP

Feasible solution of the log barrier problem

\[ \text{LET } x \text{ BE A FEASIBLE SOLUTION TO } B(M) \]
\[ \Rightarrow A\bar{x} = b \]

WE ARE IN \( x \) FEASIBLE, NOT OPTIMAL SOLUTION.

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

Quadratic approximation of \( f \):
\[
\begin{align*}
    f(x) & \approx f(\bar{x}) + \nabla f(\bar{x})^T (x - \bar{x}) + \frac{1}{2} (x - \bar{x})^T \nabla^2 f(\bar{x}) (x - \bar{x}) \\
    & = f(\bar{x}) + g(\bar{x})^T \Delta x + \frac{1}{2} \Delta x^T G(\bar{x}) \Delta x
\end{align*}
\]
\[ Q(\Delta x) \text{ QUADRATIC APPROXIMATION} \]
Primal Log Barrier Method for LP

New quadratic problem:

\[
\min_{\Delta x} Q(\Delta x)
\]

s.t. \( A(\overline{x} + \Delta x) = 0 \quad \overline{x} = \overline{x} + \Delta x \)

\[
\Rightarrow \min_{\Delta x} Q(\Delta x)
\]

s.t. \( A\Delta x = 0 \) \( \text{since } A\overline{x} = 0 \)

Lagrange problem:

\[
L(x, \pi x) = \varphi(\overline{x})^T \Delta x + \frac{1}{2} \Delta x^T G(\overline{x}) \Delta x - \pi_x^T \left[ A\Delta x \right]
\]
Lagrangian problem:

\[ L(x, \pi x) = \Phi(x)^T \Delta x + \frac{1}{2} \Delta x^T G(x) \Delta x - \pi x^T [A \Delta x] \]

In the optimum:

\[ 0 = \frac{\partial L(x, \pi x)}{\partial \Delta x} = \Phi(x) + G(x) \Delta x - A^T \pi_x \]

\[ 0 = c - MD_x^{-1} e + MD_x^{-2} \Delta x - A^T \pi_x \]

\[ \Rightarrow \begin{pmatrix} MD_x^{-2} & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} -\Delta x \\ \pi_x \end{pmatrix} = \begin{pmatrix} c - MD_x^{-1} e \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} \Delta x \in \mathbb{R}^n \\ \pi_x \in \mathbb{R}^{n+m} \end{pmatrix} \]
Interior-point algorithm

\[ \begin{align*}
\min_{x \in \mathbb{R}^n} & \quad f(x) = c^T x - M \sum_{j=1}^{\bar{v}} \log(x_j) \\
\text{s.t.} & \quad A x = b \\
\min_{c^T x} & \quad f(x) = c^T x - M \sum_{i=1}^{\bar{v}} \log(b_i - a_i^T x) \\
\text{s.t.} & \quad A x \leq b
\end{align*} \]

LP with \( \eta = 2 \), \( m = 6 \)

Contour lines of \( \phi \)

\( M = 0 \)

\( x^* \)

\( x^*(10) \)
Primal - Dual Logarithmic Barrier Method for LP
Symmetric Primal and Dual Forms

Symmetric form I [wikipedia]

\[(P) \quad \max_x C^T x \quad \text{s.t.} \quad A x \leq \ell, \quad x \geq 0\]

\[(D) \quad \min_\eta \eta^T \nu \quad \text{s.t.} \quad \eta^T A \geq C^T, \quad \eta \geq 0\]

Duality gap:
\[C^T x \leq (\eta^T A)x \leq \eta^T \ell\]

Symmetric form II [Luenberger & Ye]

\[(P) \quad \min_\eta \eta^T \nu \quad \text{s.t.} \quad A x \leq \ell, \quad x \geq 0\]

\[(D) \quad \max_x C^T x \quad \text{s.t.} \quad \eta^T A \leq C^T, \quad \eta \geq 0\]

Duality gap:
\[\eta^T \nu \leq \eta^T A x \leq C^T x\]
Asymmetric Primal and Dual Forms

Asymmetric form I [Jensen & Jonathan]

\[(\rho) \quad \text{MAX} \quad C^T x \quad \text{(O)} \quad \text{MIN} \quad \eta^T \nu \]
\[\text{s.t.} \quad A x \leq b \quad \text{s.t.} \quad \eta^T A = C^T \]
\[x \in \mathbb{R}^n \quad \eta \geq 0\]

Duality gap:

\[e^T \eta - C^T x = (Ax + s)^T \eta - \eta^T A x = s^T \eta \geq 0\]

Asymmetric form II [Luenberger & Ye, R. Freund]

\[(\rho) \quad \text{MIN} \quad C^T x \quad \text{(O)} \quad \text{MAX} \quad \eta^T \nu \]
\[\text{s.t.} \quad A x = b \quad \text{s.t.} \quad \eta^T A \leq C^T \]
\[x \geq 0 \quad \eta \in \mathbb{R}^n\]

In what follows, we will use this form.

Duality gap:

\[C^T x - e^T \eta = \eta^T A x + s^T x - x^T A^T \eta = s^T x \geq 0\]
KKT conditions

**Log barrier problem** [We already know this]

\[ p(m) = \min_{x \in \mathbb{R}^n} f(x) = c^T x - m \sum_{j=1} \log (x_j) \]

s.t. \( \Delta x = 0 \)

\( x > 0 \)

**Derivatives:**

\[ \vartheta(x) = \nabla f(x) = c - m D_x^{-1} e \]

where \( D_x = \text{diag}(G(x)) \in \mathbb{R}^{n \times n} \)

\( e = (1, \ldots, 1)^T \)

**KKT stationarity condition of the log barrier problem**

\[ \nabla f(x) + \vartheta^T \nabla (\vartheta - \lambda x) = 0 \]

\[ \begin{bmatrix} c - m D_x^{-1} e \end{bmatrix}^T \Delta x = 0 \]

\[ \Rightarrow \ c - m D_x^{-1} e = A^T \gamma \]

\[ x > 0 \]
The stationarity condition can be rewritten

\[ c - M D_x^{-1} e = A^T y \]

\[ S \]

\[ m D_x^{-1} e = S \]

\[ \left( \begin{array}{c} 1 \\ D_x \\ m \end{array} \right) S = e \]

\[ \left( \begin{array}{c} 1 \\ D_x \\ m \end{array} \right) D_s e = e \]
KKT Conditions

\[
\begin{align*}
A_x &= b - x > 0 \\
A^T y + s &= c, \quad s > 0 \\
\frac{1}{M} D_x D_s e &= e
\end{align*}
\]

Lemma

If \((x, y, s)\) is a solution of \((\star \text{KKT})\),

\[\begin{align*}
\begin{cases}
\bullet \quad x \text{ is feasible for } P \\
\bullet \quad (y, s) \text{ is feasible for } D \\
\bullet \quad \text{The duality gap is } x^T s = e^T \frac{D_x D_s e}{m} = 0
\end{cases}
\end{align*}\]

We will solve \(P(m)\) and let \(m \to 0\)
Approximation of the KKT conditions

\[ \begin{align*}
\text{(a)} & \quad \Delta x &= \nu - x > 0 \\
\text{(b)} & \quad A^T \mu + s &= c \\
\text{(c)} & \quad \frac{1}{M} D_x D_s e &= e
\end{align*} \]

This is not linear in \( x \) and \( s \)!

\( \beta \)-approximation of the stationarity condition:

\[ \begin{align*}
\text{(c)} & \quad \|
\frac{1}{M} D_x s - e \| &\leq \rho
\end{align*} \]

The primal-dual algorithm solves the system \((a), (b), (c)\) with Newton method for a given \( \rho \) and \( M \)
Duality Gap of beta approximation

Lemma: [duality gap after $\beta$-approximation]

- $\Delta x = \bar{b}, \ x > 0$
- $A^T \bar{y} + s = \bar{c},$
- $\| \frac{1}{M} D_X s - e \| \leq \beta$

\begin{equation}
\begin{aligned}
\text{(KKT-1)}
\end{aligned}
\end{equation}

We don't require $s \geq 0!$

If $(\bar{x}, \bar{y}, \bar{s})$ is a $\beta$-approximate solution of $(\text{KKT-1})$

- $\beta M < 1$

$\Rightarrow$ Duality gap can be lower & upper bounded:

\begin{equation}
\begin{aligned}
&\eta M (1-\beta) \leq C^T \bar{x} - b^T \bar{y} \leq \eta M (1+\beta) \\
&\bar{x} \text{ is feasible for (P), (}\bar{y}, \bar{s}) \text{ is feasible for (D)}
\end{aligned}
\end{equation}
Duality Gap of beta approximation

Proof:

**PRIMAL FEASIBILITY:** $\bar{x} \geq 0 \checkmark$

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

We only need to show that $\bar{z} \geq 0$

From $\| \frac{1}{M} D_x s - e \| \leq \beta$

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

$\Rightarrow 0 < (1 - \beta) M \leq x_j s_j \leq M(\beta + 1)$

$\Rightarrow s_j > 0 \Rightarrow (\bar{x}_0, \bar{y})$ is D FEASIBLE

**DUALITY GAP:**

$nM(1 - \beta) \leq \sum x_j s_j = x^T s \leq nM(\beta + 1)$

$\checkmark$ Duality Gap
Primal-Dual Newton Step

Current iterate:

\[ KKT: \begin{align*}
\Delta \vec{x} &= \vec{b}, \quad \bar{x} > 0 \\
\bar{z}^T \bar{x} + \bar{s} &< c, \quad \bar{s} > 0 \\
\frac{1}{m} D_{\bar{x}} D_{\bar{s}} e - e &= 0
\end{align*} \]

After update:

\[ A(\bar{x} + \Delta x) = \vec{b}, \quad \bar{x} + \Delta x > 0 \\
A^T(\bar{y} + \Delta y) + (\bar{s} + \Delta s) = c \\
\frac{1}{m} \left( D_{\bar{x}} + D_{\Delta x} \right)(D_{\bar{s}} + D_{\Delta s}) e - e &= 0
\]

Newton steps:

\[ A \Delta x = 0 \\
A^T \Delta \bar{y} + \Delta s = 0 \\
\frac{D_{\Delta x}}{D_{\bar{x}}} D_{\Delta s} e + \frac{D_{\bar{x}}}{D_{\bar{s}}} D_{\Delta s} e \\
\frac{D_{\bar{x}}}{D_{\bar{s}}} \Delta s + \frac{D_{\bar{s}}}{D_{\bar{s}}} \Delta x \\
\text{NEGLECT THIS NONLINEAR TERM}
\]
Primal-Dual Newton System

\[ \begin{align*}
\Delta A \Delta x &= 0 \\
A^T \Delta y + \Delta s &= 0 \\
O_{\tilde{\xi}} \Delta x + D_{\tilde{x_0}} \Delta s &= m e - D_{\tilde{x_0}} D_{\tilde{\xi}} e \\
\end{align*} \]

(\text{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**

- \(A \Delta x = 0\)
- \(A^T \Delta y + \Delta s = 0\)
- \(D_s^x \Delta x + D_y^x \Delta s = m e - D_s^x D_s^y e\)

**STEP 3.**

- \(x^{k+1} = \bar{x} + \Delta x\)
- \(y^{k+1} = \bar{y} + \Delta y\)
- \(s^{k+1} = \bar{s} + \Delta s\) \(\sigma < 1\)

**STEP 4.** \(k = k + 1, M^{k+1} = 2M_k\)

Go back to 1

\[\text{OR } M^{k+1} = \frac{1}{10} x^T s\]

Current dual, iter gap
Theorem [Local quadratic convergence of Primal-Dual Newton’s method]:

- Suppose that \((\tilde{x}, \tilde{y}, \tilde{s})\) is a \(\rho\)-approximate solution of \(P(M)\) for some \(0 < \rho < \frac{1}{2}\).
- Let \((\Delta x, \Delta y, \Delta s)\) be the solution of the Primal-Dual Newton system \((**\text{NEWTON})\).
- Let \((x', y', s') = (\tilde{x}, \tilde{y}, \tilde{s}) + (\Delta x, \Delta y, \Delta s)\).

\[\Rightarrow (x', y', s') \text{ is a } \frac{(1+\rho)}{(1-\rho)^2} \rho^2\text{-approximate solution of } P(M).\]
Summary

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