

CSE203B Convex Optimization

Chapter 11 Interior Point Methods

CK Cheng

Dept. of Computer Science and Engineering

University of California, San Diego

Chapter 11: Interior-Point Methods

- Formulation
 - Inequality constrained optimization
- Barrier Method
- Generalized Inequalities Problems
- Primal Dual Interior Point Methods
- Summary

Formulation: The problem

Problem: $\min f_0(x)$

Subject to $f_i \leq 0, i = 1, \dots, m$

$$Ax = b$$

Function f_i s are convex, twice continuously differentiable

We assume that $\text{rank } A = p, A \in R^{p \times n}$.

Issues:

1. m can be exponential.
2. When to put $f_i = 0$ (active)? There are 2^m combination.

Formulation: logarithmic barrier

Problem:

$$\min f_0(x) + \sum_{i=1}^m I_{f_i}(x)$$

$$s. t. \quad Ax = b$$

When $I_u = 0$ if $u \leq 0$, $I_u = \infty$. Otherwise,

$$\min f_0(x) + \frac{-1}{t} \sum_{i=1}^m \log(-f_i(x))$$

$$s. t. \quad Ax = b$$

Remark:

1. Convert inequality constraints to barrier functions.
2. Incorporate barrier functions in objective function.
3. Increase t to improve accuracy.

Formulation: logarithmic barrier

Let us set

$$\phi(x) = - \sum_{i=1}^m \log(-f_i(x)), \quad \text{dom } \phi = \{x | f_i(x) < 0\}$$

$\phi(x)$ is convex and twice differentiable

$$\nabla \phi(x) = \sum_{i=1}^m -\frac{1}{f_i(x)} \nabla f_i(x)$$

$$\nabla^2 \phi(x) = \sum_{i=1}^m \frac{1}{f_i(x)^2} \nabla f_i(x) \nabla f_i(x)^T - \frac{1}{f_i(x)} \nabla^2 f_i(x)$$

Central Path is $\{x^*(t) | t > 0\}$

$$\min \quad t f_0(x) + \phi(x)$$

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

Formulation: logarithmic barrier

Ex:

$$\begin{aligned} \text{Problem:} \quad & \min \quad c^T x \\ & \text{s.t.} \quad a_i^T x \leq b_i, \quad i = 1, \dots, m \end{aligned}$$

Log barrier formulation:

$$\min t c^T x - \sum_{i=1}^m \log(b_i - a_i^T x)$$

Hyperplane $c^T x = c^T x^*(t)$ is tangent to real curve φ through $x^*(t)$.

Solution $x^*(t)$ balance the force between $-t \nabla f_0(x)$ and $\sum_{i=1}^m -\frac{1}{f_i(x)} \nabla f_i(x)$.

Formulation: logarithmic barrier

Ex:

Problem: $\min c^T x$

s. t. $a_i^T x \leq b_i \quad i = 1, \dots, m$

$-t \nabla f_0(x) = -tc$

$$\sum_{i=1}^m -\frac{1}{f_i(x)} \nabla f_i(x) = \sum_{i=1}^m -\frac{1}{b_i - a_i^T x} a_i$$

Note that $\min \left\| \frac{1}{b_i - a_i^T x} a_i \right\|_2 = \frac{1}{\text{dist}(x, H_i)}$, $H_i = \{x \mid a_i^T x = b_i\}$

Barrier Method: Algorithm

Given strictly feasible x , $t = t^0 > 0$, $\mu > 1$, $\epsilon > 0$

Repeat (10-20)

1. **Centering step** to find solution $x^*(t)$

Problem: $\min \quad tf_0(x) + \phi(x) \quad (\text{Newton's method})$
 $s.t. \quad Ax = b$

2. Update $x = x^*(t)$

3. Stopping criterion: exit if $\frac{m}{t} < \epsilon$

4. Increase $t = \mu t$

Complexity: # Repeats (Outer iterations) = $\frac{\log(\frac{m}{\epsilon t^{(0)}})}{\log \mu}$

Plus the initial centering step $x^*(t^{(0)})$

Barrier Method: Newton's Step for **Modified** KKT

$$\begin{bmatrix} t\nabla^2 f_o(x) + \nabla^2 \phi(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ v \end{bmatrix} = - \begin{bmatrix} t\nabla f_o(x) + \nabla \phi(x) \\ 0 \end{bmatrix}$$

$$\nabla \sum_{i=1}^m (-\log(-f_i(x))) = \sum_{i=1}^m -\frac{1}{f_i(x)} \nabla f_i(x)$$

$$\begin{aligned} \nabla^2 \sum_{i=1}^m (-\log(-f_i(x))) \\ = \sum_{i=1}^m \left[-\frac{1}{f_i(x)} \nabla^2 f_i(x) + \frac{1}{f_i(x)^2} \nabla f_i(x) \nabla f_i(x)^T \right] \end{aligned}$$

Barrier Method: Central Path

$$\text{Min } f_0(x) + \frac{-1}{t} \sum_{i=1}^m \log(-f_i(x))$$

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

$$\text{Lagrangian: } L(x, v) = f_0(x) + \frac{-1}{t} \sum_{i=1}^m \log(-f_i(x)) + v^T (Ax - b)$$

For an optimal solution, we have $(x^*(t), \bar{v}(t))$

$$\nabla f_0(x^*) + \sum -1/(t f_i(x^*)) \nabla f_i(x^*) + A^T \bar{v} = 0$$

We can view the dual points from central path:

$$\lambda_i^*(t) = -1/(t f_i(x^*)), i = 1, \dots, m$$

Original Lagrangian:

$$L(x, \lambda, v) = f_0(x) + \sum \lambda_i f_i(x) + v^T (Ax - b)$$

Replace with $(x^*(t), \lambda^*(t), \bar{v}(t))$:

$$L(x^*, \lambda^*, \bar{v}) = f_0(x^*) + \sum \lambda_i^* f_i(x^*) + \bar{v}^T (Ax^* - b) = f_0(x^*) - \frac{m}{t}$$

Thus, we have $f_0(x^*(t)) - p^* \leq m/t$

Barrier Method: Feasible Solution Search

Search 1:

$\min s$

$s. t. f_i(x) \leq s, i = 1, \dots, m$

$Ax = b, s \in R$

Search 2:

$\min 1^T s, \quad s \in R_+^m$

$s. t. f_i(x) \leq s_i, i = 1, \dots, m$

$Ax = b$

Barrier Method: complexity analysis

#Repeats (outer iterations)

$$=\text{Ceiling}(\log(m/(\epsilon t^0))/\log\mu)$$

#Newton steps per outer iteration (self-concordance)

$$=\frac{m(\mu-1-\log\mu)}{\gamma} + \log_2 \log_2 1/\epsilon_{nt},$$

where $\gamma = \alpha\beta(1 - 2\alpha)^2/(20 - 8\alpha)$

Generalized Inequalities Problems

Problem: $\min f_0(x)$

Subject to $f_i(x) \preceq_{K_i} 0, \quad i = 1, \dots, m$, where $f_i(x) \in R^{k_i}$

$$Ax = b$$

The KKT conditions:

$$Ax^* = b$$

$$f_i(x^*) \preceq_{K_i} 0, \quad i = 1, \dots, m$$

$$\lambda_i^* \succeq_{K_i^*} 0, \quad i = 1, \dots, m$$

$$\nabla f_0(x^*) + \sum Df_i(x^*)^T \lambda_i^* + A^T v^* = 0$$

$$\lambda_i^{*T} f_i(x^*) = 0, \quad i = 1, \dots, m.$$

Note that $Df^i(x^*) \in R^{k_i \times n}$

Generalized Inequalities Problems: log barrier

Problem: $\min f_0(x)$

Subject to $f_i(x) \preceq_{K_i} 0$, $i = 1, \dots, m$, where $f_i(x) \in R^{k_i}$

$$Ax = b$$

Given a proper cone $K \subseteq R^q$, a generalized logarithm for K , $\psi: R^q \rightarrow R$ has the following two criteria:

1. Function ψ : concave, closed, twice continuously differentiable, $\text{dom } \psi = \text{int } K$, and $\nabla^2 \psi(y) \prec 0$, for $y \in \text{int } K$
2. Equality: $\psi(sy) = \psi(y) + \theta \log s$, for all $y \succ 0$, $s > 0$, where there exists a constant (**degree of ψ**) $\theta > 0$

We can derive two properties

1. If $y \succ_K 0$, then $\nabla \psi(y) \succ_{K^*} 0$ (**Proof?**)
2. $y^T \nabla \psi(y) = \theta$ (**from criterion 2**)

Generalized Inequalities Problems: log barrier

Example 1: Cone $K = R_+^n$

Function $\psi(x) = \sum_i \log x_i, x \succ 0$ is a generalized logarithm

1. Concavity: $\nabla^2 \psi(x) = \text{diag} \left(-\frac{1}{x_i^2} \right) \prec 0$
2. Log behavior: $\psi(sx) = \sum \log sx_i = \sum \log x_i + n \log s$,
where $s > 0$.

Two properties:

1. If $x \in K = R_+^n$, then

$$\nabla \psi(x) = \left(\frac{1}{x_1}, \dots, \frac{1}{x_n} \right) \succ_{K^*} 0$$

2. $x^T \nabla \psi(x) = n$.

Generalized Inequalities Problems: log barrier

Example 2: Cone $K = \left\{ x \in R^{n+1} \mid \left(\sum_i x_i^2 \right)^{1/2} \leq x_{n+1} \right\}$

Function $\psi(x) = \log(x_{n+1}^2 - \sum_i x_i^2)$,

1. Concavity: (**exercise**)
2. Log behavior: $\psi(sx) = \psi(x) + 2 \log s$

Two properties

$$1. \frac{\partial \psi(x)}{\partial x_j} = -\frac{2x_j}{x_{n+1}^2 - \sum x_i^2}, j = 1, \dots, n$$

$$\frac{\partial \psi(x)}{\partial x_{n+1}} = \frac{2x_{n+1}}{x_{n+1}^2 - \sum x_i^2},$$

$$\nabla \psi(x) \in \text{int } K^*$$

$$2. x^T \nabla \psi(x) = 2.$$

Generalized Inequalities Problems: log barrier

Example 3: Cone $K \in S_+^p$

Function $\psi(x) = \log \det X$,

1. Concavity: (**exercise**)
2. Log behavior: $\psi(sx) = \psi(x) + p \log s$

Two properties:

1. $\log \det(sX) = \log \det(X) + p \times \log s$

$$\nabla \psi(X) = X^{-1} \succ 0$$

2. $\text{tr}(X \nabla \psi(X)) = \text{tr}(XX^{-1}) = p.$

Primal-Dual Interior-Point Method

$$\begin{aligned} \min f_o(x) \\ \text{s. t. } f_i(x) \leq 0, i = 1, \dots, m \\ Ax = b \end{aligned}$$

Lagrangian

$$L(x, \lambda, v) = f_o(x) + \sum_{i=1}^m \lambda_i f_i(x) + v^T (Ax - b)$$

KKT Conditions

$$\nabla_x L(x, \lambda, v) = \nabla f_o(x) + \sum_{i=1}^m \lambda_i \nabla f_i(x) + A^T v = 0$$

$$Ax = b$$

$$f_i(x) \leq 0, i = 1, \dots, m$$

$$\lambda_i \geq 0$$

$$\lambda_i f_i(x) = 0 \rightarrow -\lambda_i f_i(x) = \frac{1}{t}, i = 1, \dots, m$$

$$\left(\lambda_i = -\frac{1}{t f_i(x)} \right)$$

Primal-Dual Interior-Point Method

$$r_{dual} = \nabla f_o(x) + \sum \lambda_i \nabla f_i(x) + A^T v$$

$$r_{centrality} = -diag(\lambda) f(x) - (1/t) \mathbf{1} \quad (-\lambda_i f_i(x) - 1/t)$$

$$r_{primal} = Ax - b$$

$$Df(x) = \begin{pmatrix} \nabla f_1(x)^T \\ \vdots \\ \nabla f_m(x)^T \end{pmatrix}, \quad r_t = \begin{bmatrix} r_{dual} \\ r_{cent} \\ r_{pri} \end{bmatrix}, \quad y = \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta v \end{bmatrix}$$

$$r_t(x + \Delta x, \lambda + \Delta \lambda, v + \Delta v) = r_t(x, \lambda, v) + \nabla_y r_t^T \Delta y$$

$$1. r_{dual}(x + \Delta x, \lambda + \Delta \lambda, v + \Delta v) \approx r_{dual}(x, \lambda, v) + \nabla_x r_{dual}^T \Delta x + \nabla_\lambda r_{dual}^T \Delta \lambda + \nabla_v r_{dual}^T \Delta v = 0$$

$$\nabla_x r_{dual} = \nabla^2 f_o(x) + \sum_{i=1}^m \lambda_i \nabla^2 f_i(x)$$

$$\nabla_\lambda r_{dual} = Df(x)^T$$

$$\nabla_v r_{dual} = A^T$$

$$2. r_{cent.}(x + \Delta x, \lambda + \Delta \lambda, v + \Delta v) \approx r_{cent.}(x, \lambda, v) + \nabla_x r_{cent.}^T \Delta x + \nabla_\lambda r_{cent.}^T \Delta \lambda = 0$$

$$\nabla_x r_{cent.} = -diag(\lambda) Df(x)$$

$$\nabla_\lambda r_{cent.} = diag(f(x))$$

Primal-Dual Interior-Point Method

$$r_{dual} = \nabla f_o(x) + \sum \lambda_i \nabla f_i(x) + A^T v$$

$$r_{centrality} = -diag(\lambda) f(x) - (1/t) \mathbf{1} \quad (-\lambda_i f_i(x) - 1/t)$$

$$r_{primal} = Ax - b$$

$$\begin{matrix} (1) \\ (2) \\ (3) \end{matrix} \begin{bmatrix} \nabla^2 f_o(x) + \sum_{i=1}^m \lambda_i \nabla^2 f_i(x) & Df(x)^T & A^T \\ -diag(\lambda) Df(x) & -diag(f(x)) & 0 \\ A & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta v \end{bmatrix} = \begin{bmatrix} r_{dual} \\ r_{cent.} \\ r_{pri.} \end{bmatrix}$$

$$Df(x) = \begin{pmatrix} \nabla f_1(x)^T \\ \vdots \\ \nabla f_m(x)^T \end{pmatrix}, \quad r_t = \begin{bmatrix} r_{dual} \\ r_{cent} \\ r_{pri} \end{bmatrix}, \quad y = \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta v \end{bmatrix}$$

$$r_t(x + \Delta x, \lambda + \Delta \lambda, v + \Delta v) = r_t(x, \lambda, v) + \nabla_y r_t^T \Delta y$$

Primal Dual Interior Point Method: the surrogate duality gap

$$\eta(x, \lambda) = -f(x)^T \lambda \quad (f_i(x) < 0, \lambda \geq 0)$$

When $r_{\text{prime}} = 0$, and $r_{\text{dual}} = 0$

Primal-Dual Interior-Point Method: comparison with barrier method

Primal-dual interior-point method:

$$\begin{array}{l}
 (1) \\
 (2) \\
 (3)
 \end{array}
 \begin{bmatrix}
 \nabla^2 f_0(x) + \sum_{i=1}^m \lambda_i \nabla^2 f_i(x) & Df(x)^T & A^T \\
 -diag(\lambda) Df(x) & -diag(f(x)) & 0 \\
 A & 0 & 0
 \end{bmatrix}
 \begin{bmatrix}
 \Delta x \\
 \Delta \lambda \\
 \Delta v
 \end{bmatrix}
 =
 \begin{bmatrix}
 r_{dual} \\
 r_{cent.} \\
 r_{pri.}
 \end{bmatrix}$$

$$\begin{bmatrix}
 H_{pd} & A^T \\
 A & 0
 \end{bmatrix}
 \begin{bmatrix}
 \Delta x \\
 \Delta v + v
 \end{bmatrix}
 =
 -
 \begin{bmatrix}
 \nabla f_0(x) + \left(\frac{1}{t}\right) \sum_i \frac{-1}{f_i(x)} \nabla f_i(x) \\
 r_{pri.}
 \end{bmatrix}$$

where $H_{pd} = \nabla^2 f_0(x) + \sum \lambda_i \nabla^2 f_i(x) + \sum -(\lambda_i/f_i(x)) \nabla f_i(x) \nabla f_i(x)^T$

Barrier Method:

$$\begin{bmatrix}
 H_{bar} & A^T \\
 A & 0
 \end{bmatrix}
 \begin{bmatrix}
 \Delta x \\
 \Delta v
 \end{bmatrix}
 =
 -
 \begin{bmatrix}
 t \nabla f_0(x) + \sum_i \frac{-1}{f_i(x)} \nabla f_i(x) \\
 r_{pri.}
 \end{bmatrix}$$

where $H_{bar} = t \nabla^2 f_0(x) + \sum (-1/f_i(x)) \nabla^2 f_i(x) + \sum (1/f_i(x)^2) \nabla f_i(x) \nabla f_i(x)^T$

Primal-Dual Interior-Point Method: algorithm

Input $f_i < 0, \lambda > 0, \mu > 1, \epsilon_{feas} > 0, \epsilon > 0$

Repeat 1. Determine t , set $t := \mu m / \hat{\eta}$

2. Compute $(\Delta x, \Delta \lambda, \Delta v)$

3. Line Search and update

$$y = y + s \Delta y \quad (\Delta y = \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta v \end{bmatrix})$$

Until $\|r_{pri}\|_2 \leq \epsilon_{feas}, \|r_{dual}\|_2 \leq \epsilon_{feas},$ and $\hat{\eta} \leq \epsilon$

Remark

1. Parameter t is automatically adjusted.
2. The process proceeds even $Ax \neq b, \nabla L(x, \lambda, v) \neq 0$.
3. The search directions are similar to but not quite the same as the search directions of the barrier method.
4. The method is often more efficient than the barrier method.

Summary

- Interior point methods convert inequality constraints into costs of objective function.
- The barrier method starts with strictly feasible solution.
- The primal dual interior methods have become popular due to its efficiency and generalization.