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

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

Introduction

- Frisch, R., 1956. La résolution des problèmes de programme linéaire par la méthode du potentiel logarithmique. Cahiers du Seminaire D'Econometrie, pp.7-23.
- Dikin, I.I., 1967. Iterative solution of problems of linear and quadratic programming. In Doklady Akademii Nauk (Vol. 174, No. 4, pp. 747-748). Russian Academy of Sciences.
- Karmarkar, N., 1984, December. A new polynomial-time algorithm for linear programming. In Proceedings of the sixteenth annual ACM symposium on Theory of computing (pp. 302-311).
- Wright, M., 2005. The interior-point revolution in optimization: history, recent developments, and lasting consequences. Bulletin of the American mathematical society, 42(1), pp.39-56.
- Nesterov, Y. and Nemirovskii, A., 1994. Interior-point polynomial algorithms in convex programming. Society for industrial and applied mathematics.

Formulation: The problem

Softmax, $\max\{x_1, x_2 \cdots x_m\}$

$\hookrightarrow \ln \sum_i e^{\gamma x_i}$

Problem: $\min f_0(x)$

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

$Ax = b$

Function f_is are convex, twice continuously differentiable

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

Issues:

- KKT conditions on inequality constraints
  - $\lambda_i = 0$, if $f_i(x) < 0$; otherwise $\lambda_i > 0$ $(\lambda_i f_i(x) = 0)$
- m can be large.
- When to put $f_i = 0$ (active)? There are 2^{m} combination.

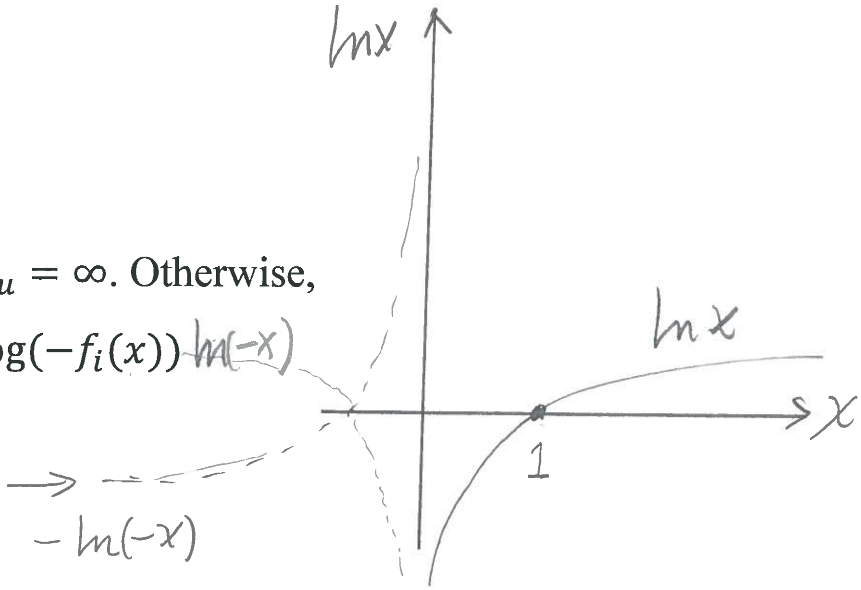
Formulation: logarithmic barrier

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

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

$\phi(x)$ 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 $\{x^*(t) | t > 0\}$

$$min \quad tf_0(x) + \phi(x)$$
$$s.t. \quad Ax = b$$

Formulation: logarithmic barrier

Ex:

$f_i(x)=a_i^Tx-b_i\le 0$

Problem: $\quad min \quad c^Tx$

$\quad s.t. \quad a_i^Tx \le b_i, \qquad i=i,\dots,m$

Log barrier formulation:

$$\min tc^Tx - \sum_{i=1}^{m} \log(b_i - a_i^Tx)$$

Hyperplane $c^Tx = c^Tx^*(t)$ *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: $min \quad c^Tx$

$s.t. \quad a_i^Tx \le b_i \qquad i=i,\dots,m$

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

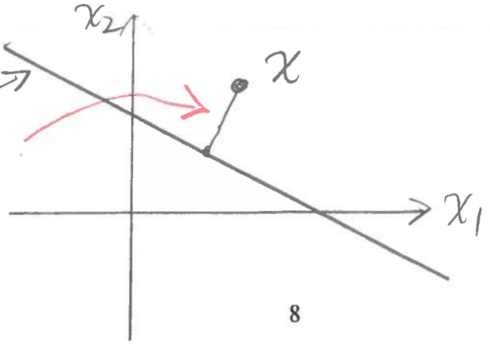
$P(x)=c^Tx-\sum_i \ln(b_i-a_i^Tx)$

$\nabla P(x) = c - \sum_i \frac{a_i}{b_i-a_i^Tx} = c - \sum_i \frac{a_i/\|a_i\|}{b_i/\|a_i\| - a_i^T/\|a\| x}$

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

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

$\frac{b_i}{\|a_i\|} - \frac{a_i^T}{\|a_i\|}x$



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: $\min \quad tf_0(x) + \phi(x)$ *(Newton's method)*

   $$s.t. \quad \underset{m\times n}{A}x = \underset{m\times 1}{b}$$

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 $x^*\left(t^{(0)}\right)$

Barrier Method: Newton's Step for Modified KKT

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

n+m eq

n+m variables

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

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

$$=\sum_{i=1}^{m}[-\frac{1}{f_i(x)}\underset{n\times n}{\nabla^2 f_i(x)} + \frac{1}{f_i(x)^2}\underbrace{\underset{n\times 1 \ \times\ 1\times n}{\nabla f_i(x)\nabla f_i(x)^T}}_{n\times n}$$

Barrier Method: Central Path

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

$s.t. \quad Ax = b$

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 $(x^*(t), \bar{v}(t))$

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

We can view the dual points from central path:

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

Original Lagrangian:

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

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

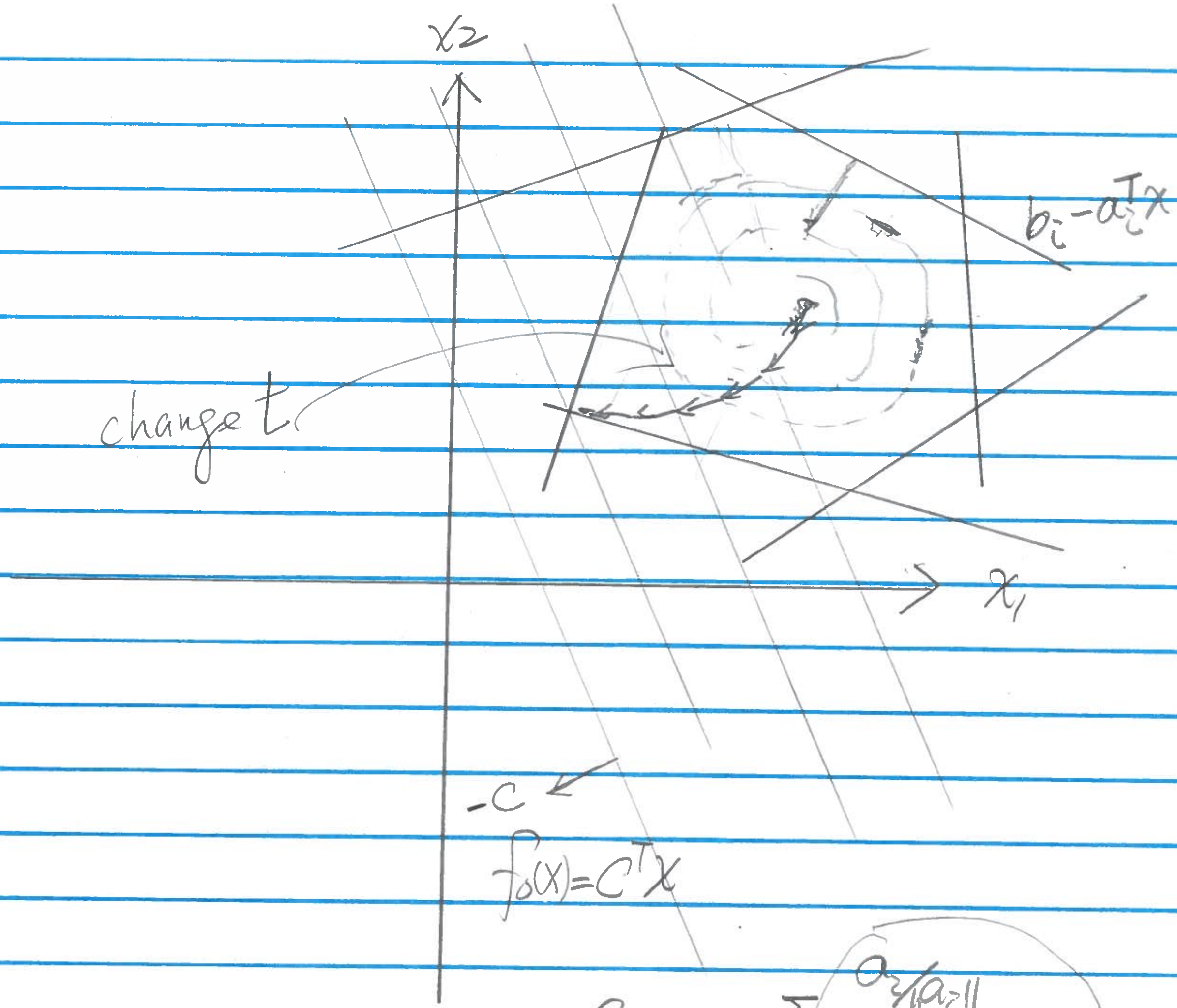
$$L(x^*, \lambda^*, \bar{v}) = f_0(x^*) + \sum\lambda_i^* f_i(x^*) + \bar{v}^T(Ax^* - b) = f_0(x^*) - \frac{m}{t} \leqslant p^*$$

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

Barrier Method: Feasible Solution Search

Search 1:
min s
$s.t. f_i(x) \le s, i = 1, \dots, m$
$\quad Ax = b, s \in R$

Search 2:
min $1^T s$, $\quad s \in R_+^m$
$s.t. f_i(x) \le s_i, i = 1, \dots, m$
$\quad Ax = b$



$$\underset{\sim}{c} \qquad \boxed{\sum_i \frac{a_i/\|a_i\|}{\frac{b_i}{\|a_i\|} - \frac{a_i^T}{\|a_i\|}x}}$$

$$\text{Obj}: P(x) = t\,c^T x - \sum_i \ln(b_i - a_i^T x)$$

$$\nabla P(x) \Rightarrow t\,c$$

Newton Method: 2nd ordered Taylor's exp.

Statement of Problem: $\min \; t f_0(t) + \phi(x)$

s.t. $Ax = b$

Lagrangian: $L(x, \nu) = t f_0(x) + \phi(x) + \nu^T (Ax - b)$

$$L(x+\Delta x, \nu) \approx t f_0(x) + \underline{t \nabla f_0(x)^T} \Delta x + t/2 \, \Delta x^T \underline{\nabla^2 f_0(x) \Delta x}$$

$$+ \phi(x) + \underline{\nabla \phi(x)^T} \Delta x + \frac{1}{2} \Delta x \underline{\nabla^2 \phi(x) \Delta x}$$

$$+ \nu^T (Ax - b) + \underline{\nu^T A} \Delta x$$

KKT: $\begin{bmatrix} t\nabla^2 f_0(x) + \nabla^2 \phi(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \nu \end{bmatrix} = \begin{bmatrix} t \nabla f_0(x) + \nabla \phi(x) \\ 0 \end{bmatrix}$

Barrier Method: complexity analysis

#Repeats (outer iterations)
=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 $f_i(x) \preccurlyeq_{K_i} 0,\ i=1,\dots,m,$ where $f_i(x) \in R^{k_i}$
$Ax=b$

The KKT conditions:
$Ax^*=b$
$f_i(x^*) \preccurlyeq_{K_i} 0, \qquad i=1,\dots,m$
$\lambda_i^* \succcurlyeq_{K_i^*} 0,\ i=1,\dots,m$
$\nabla f_0(x^*)+\sum Df_i(x^*)^T\lambda_i^*+A^Tv^*=0$
${\lambda_i^*}^T f_i(x^*)=0, i=1,\dots,m.$
Note that $Df^i(x^*) \in R^{k_i\times n}$

Generalized Inequalities Problems: log barrier

Problem: min $f_0(x)$

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

$Ax = b$

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

1. Function ψ: concave, closed, twice continuously differentiable, dom $\psi = int\ K$, and $\nabla^2\psi(y) \prec 0$, for $y \in int\ K$
2. Equality: $\psi(sy) = \psi(y) + \theta logs$, 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: $\nabla^2\psi(x) = diag\left(-\frac{1}{x_i^2}\right) \prec 0$
2. Log behavior: $\psi(sx) = \sum logsx_i = \sum logx_i + nlogs$, 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 $K = \left\{x \in R^{n+1} \middle| \left(\sum_i x_i^2\right)^{1/2} \leq x_{n+1}\right\}$

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

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

Two properties

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

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

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

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: $\psi(sx) = \psi(x) + plogs$

Two properties:

1. $logdet(sX) = logdet(X) + p \times logs$

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

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

Primal-Dual Interior-Point Method

$\min f_o(x)$

$s.t. f_i(x) \le 0, i = 1, \dots, m$

$Ax = b$

Lagrangian

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

KKT Conditions

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

$$Ax = b$$

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

$$\lambda_i \ge 0$$

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

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

$$\begin{bmatrix}\lambda_1 f_1 \\ \vdots \\ \lambda_m f_m\end{bmatrix} = \begin{bmatrix}-\frac{1}{t} \\ \vdots \\ -\frac{1}{t}\end{bmatrix}$$

Primal-Dual Interior-Point Method

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

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

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

$$\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_m\end{bmatrix}\begin{bmatrix}f_1 \\ \vdots \\ f_m\end{bmatrix} = \begin{bmatrix}-\lambda_1 f_1 \\ \vdots \\ -\lambda_m f_m\end{bmatrix}$$

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

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

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

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

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

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

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

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

$n+m+p$ eq.

$n+m+p$ variables

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

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

$$\lambda_i = -\frac{1}{t f_i(x)} \Rightarrow -f(x)^T \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{bmatrix} = \frac{m}{t}$$

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

Primal-dual interior-point method:

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

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

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

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

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 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.