Barrier Method: Newton's Step for Modified KKT

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

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

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

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

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

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

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

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

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

$$p^* = \max_{\lambda,v} g(\lambda,v) \ge g(\lambda^*, \bar{v}) = \min_{x} L(x, \lambda^*, \bar{v}) = f_0(x^*) - \frac{m}{t}$$

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 $1^T s$, $s \in R_+^m$
$s.t. f_i(x) \leq s_i, i = 1, \dots, m$
$Ax = b$

Barrier Method: complexity analysis

#Repeats (outer iterations)
=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 $\gamma = \alpha\beta(1-2\alpha)^2/(20-8\alpha)$

Gap $f(x^*) - p^* \leq \frac{m}{t^{(\alpha)}}$ $\quad t^{(\alpha)} = \mu^\alpha t^{(0)}$

$$\frac{m}{t^{(\alpha)}} \leq \epsilon \Rightarrow \frac{m}{\mu^\alpha t^0} \leq \epsilon$$

It takes α iterations to have $\frac{m}{\mu^\alpha t^0} \leq \epsilon \Rightarrow \alpha = \frac{\log \frac{m}{\epsilon t^0}}{\log \mu}$.

$$\min f_0(x)$$
$$\text{s.t.} \quad f_i(x) \le 0 \quad i=1,\cdots m$$
$$Ax = b \qquad x \in R^n.$$

Lagrangian

$$L(x,\lambda,v) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + v^T(Ax-b)$$

$$g(\lambda,v) = \min_x L(x,\lambda,v)$$

$$\Rightarrow \nabla_x L(x,\lambda,v) =$$

x^*, λ^*, v^* $$\nabla f_0(x) + \sum_{i=1}^{m} \lambda_i \nabla f_i(x) + A^T v = 0 \quad \Longleftrightarrow$$

(KKT)

$$\begin{bmatrix} \nabla^2 f_0(x) + \sum_{i=1}^{m} \lambda_i \nabla^2 f_i(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ v \end{bmatrix}$$

$$= - \begin{bmatrix} \nabla f_0(x) + \sum_{i=1}^{m} \lambda_i \nabla f_i(x) \\ 0 \end{bmatrix}$$

Barrier Method

$$\min t f_0(x) - \sum_{i=1}^{m} \log(-f_i(x))$$
$$Ax = b.$$

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

$$g(v) = \min_x L(x,v)$$

$$\nabla_x L(x,v) = t\nabla f_0(x) - \sum_{i=1}^{m} \frac{1}{f_i(x)} \nabla f_i(x) + A^T v = 0$$

KKT condition

$$\begin{bmatrix} t\nabla^2 f_0(x) + \sum_{i=1}^{m} \frac{-1}{f_i(x)} \nabla^2 f_i(x) + \frac{1}{f_i^2(x)} \nabla f_i(x) \nabla f_i(x)^T & A^T \\ A & 0 \end{bmatrix}$$

$$= - \begin{bmatrix} t\nabla f_0(x) + \sum_{i=1}^{m} \frac{-1}{f_i(x)} \nabla f_i(x) \\ 0 \end{bmatrix}$$

A : $y^T \nabla \psi(y) = \theta$ Proof:

Condition

(1). $\psi(y)$ is a concave function in $y \in K$.

(2) $\psi(sy) = \psi(y) + \theta \log s \quad \forall\, y > 0, s > 0$.

From (2)

$$\frac{d\,\psi(sy)}{ds} = \nabla \psi(sy)^T y \quad \text{(chain rule)} \quad (3)$$

$$\frac{d\,\psi(y) + \theta \log s}{ds} = \theta / s \quad (4)$$

From (3) & (4), we have

$$\nabla \psi(sy)^T s y = \theta \Rightarrow \nabla \psi(y)^T y = \theta.$$

B : $\nabla \psi(y) >_{K^*} 0$ Proof:

Condition

(1) $\psi(y)$ is a concave function in $y \in K$ & $\nabla \psi(x) < 0$.

(2), $\nabla \psi(y)^T y = \theta$, $\psi(sy) = \psi(y) + \theta \log s$.

From (1), we have

$$\psi(sw) \leq \psi(y) + \nabla \psi(y)(sw - y)$$
$$\forall\, sw \in K$$

(From (2)) $\qquad = \psi(y) + s \nabla \psi(y) w - \theta$

Left side $\psi(sw) = \psi(w) + \theta \log s$

Righ side $\psi(y) + s \nabla \psi(y) w - \theta$

If $w \nabla \psi(y) < 0$, the two side inequality cannot be true

$\Rightarrow \nabla \psi(y) \in K^*$

Generalized Inequalities Problems

Problem: min $f_0(x)$
Subject to $f_i(x) \preccurlyeq_{K_i} 0,\ \ i = 1, \ldots, m$, where $f_i(x) \in R^{k_i}$
$Ax = b$

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

Generalized Inequalities Problems: log barrier

Problem: min $f_0(x)$ $\Rightarrow$ $f_0(x) - \psi(-f_i(x))$
Subject to $f_i(x) \preccurlyeq_{K_i} 0,\ \ i = 1, \ldots, m$, 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: $\psi(sy) = \psi(y) + \theta \log s$, for all $y \succ_K 0, s > 0$, 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. $tr\left(X\nabla\psi(X)\right) = tr(XX^{-1}) = p.$

17

Primal-Dual Interior-Point Method

$\min f_o(x)$
$s.t. f_i(x) \leq 0, i = 1, \dots, m$
$\quad Ax = b$

Lagrangian

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

KKT Conditions

$$\nabla_x L(x, \lambda, v) = \nabla f_o(x) + \sum_{i=1}^{m} \lambda_i \nabla f_o(x) + A^T v = 0$$
$$Ax = b$$
$$f_i(x) \leq 0, i = 1, \dots, m$$
$$\lambda_i \geq 0$$
$$\lambda_i f_i(x) = 0 \ \rightarrow \ -\lambda_i f_i(x) = \frac{1}{t}, i = 1, \dots, m$$
$$(\lambda_i = -\frac{1}{tf_i(x)})$$

Primal-Dual Interior-Point Method

$$\nabla_x L = r_{dual} = \nabla f_o(x) + \sum \lambda_i \nabla f_i(x) + A^T v$$
$$r_{centrality} = -diag(\lambda)\underbrace{f(x)}_{C} - (1/t)1, (-\lambda_i f_i(x) - 1/t)$$
$$r_{primal} = Ax - b$$
$$Df(x) = \begin{pmatrix} \nabla f_1(x)^T \\ \vdots \\ \nabla f_m(x)^T \end{pmatrix}, \quad r_t = \begin{bmatrix} r_{dual} \\ r_{cent} \\ r_{pri} \end{bmatrix}, \quad y = \begin{bmatrix} \Delta x \\ \Delta\lambda \\ \Delta v \end{bmatrix}$$
$$r_t(x + \Delta x, \lambda + \Delta\lambda, v + \Delta v) = r_t(x, \lambda, v) + \nabla_y r_t^T \Delta y$$

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

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

Primal-Dual Interior-Point Method

$$r_{dual} = \nabla f_o(x) + \sum \lambda_i \nabla f_i(x) + A^T v$$
$$r_{centrality} = -diag(\lambda)f(x) - (1/t)1 \ (-\lambda_i f_i(x) - 1/t)$$
$$r_{primal} = Ax - b$$
$$\begin{array}{c} (1) \\ (2) \\ (3) \end{array} \begin{bmatrix} \nabla^2 f_o(x) + \sum_{i=1}^m \lambda_i \nabla^2 f_i(x) & Df(x)^T & A^T \\ -diag(\lambda)Df(x) & -diag(f(x)) & 0 \\ A & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta\lambda \\ \Delta v \end{bmatrix} = \begin{bmatrix} r_{dual} \\ r_{cent.} \\ r_{pri.} \end{bmatrix}$$
$$Df(x) = \begin{pmatrix} \nabla f_1(x)^T \\ \vdots \\ \nabla f_m(x)^T \end{pmatrix}, \quad r_t = \begin{bmatrix} r_{dual} \\ r_{cent} \\ r_{pri} \end{bmatrix}, \quad y = \begin{bmatrix} \Delta x \\ \Delta\lambda \\ \Delta v \end{bmatrix}$$
$$r_t(x + \Delta x, \lambda + \Delta\lambda, v + \Delta v) = r_t(x, \lambda, v) + \nabla_y r_t^T \Delta y$$

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

$$\eta(x,\lambda) = -\sum_{i=1}^{m} \lambda_i f_i(x) \approx \frac{m}{t}$$

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

Primal-dual interior-point method:

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

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

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:

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

same

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

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.