Interior-point methods

A supplementary note to Chapter 11 of *Convex Optimization* by S. Boyd and L. Vandenberghe

Optimization Lab.

IE department
Seoul National University

2nd December 2009

Inequality constrained minimization

Consider minimization

$$\begin{array}{ll} \min & f_0(x) \\ \text{s.t.} & f_i(x) \leq 0, \ i = 1, \ldots, m, \\ & Ax = b, \end{array} \tag{1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is convex and twice cont. diff'ble (hence domain is open) and $A \in \mathbb{R}^{p \times n}$ is of full row-rank . We also assume $\exists$ optimal solution $x^* \in \mathrm{dom} f$ such that $p^* = f(x^*)$. Furthermore, a Slater type constraint qualification holds: $\exists$ feasible x satisfying $f_i(x) < 0 \ \forall \ i = 1, \ldots, m$.
From KKT conditions, a point $x^* \in \mathcal{D}$ is optimal iff $\exists \ \lambda^*$ and ν^* such that

$$\begin{aligned} Ax^* = b, \ f_i(x^*) &\leq 0, \quad i = 1, \ldots, m \\ \lambda^* &\geq 0, \\ \nabla f_0(x^*) + \textstyle\sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) + A^T \nu* &= 0, \\ \lambda_i^* \nabla f_i(x^*) &= 0, \quad i = 1, \ldots, m. \end{aligned} \tag{2}$$

Logarithmic barrier

Minimization (1) can be rewritten as follows:

$$\begin{array}{ll} \min & f_0(x) + \sum_{i=1}^{m} I_-(f_i(x)) \\ \text{s.t.} & Ax = b, \end{array} \tag{3}$$

where I_- is indicator for the nonpositive reals:

$$I_-(u) = \begin{cases} 0, & \text{for } u \leq 0, \\ \infty, & \text{for } u > 0. \end{cases}$$

As an approximation of indicator, we can use *logarithmic barrier*,

$$\hat{I}_-(u) = -(1/t)\log(-u), \ \mathrm{dom}\hat{I}_- = -\mathbb{R}_{++}, \tag{4}$$

where $t > 0$ is a parameter that sets the accuracy of approximation: the larger, the more accurate.

Logarithmic barrier(*cont'd*)

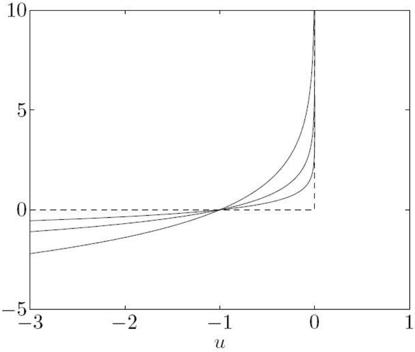

Figure 11.1 The dashed lines show the function $I_{-}(u)$, and the solid curves show $\widehat{I}_{-}(u) = -(1/t)\log(-u)$, for $t = 0.5,\ 1,\ 2$. The curve for $t = 2$ gives the best approximation.

Logarithmic barrier(*cont'd*)

Thus (3) is approximated by

$$\begin{array}{ll} \min & f_0(x) + \sum_{i=1}^m -(1/t)\log(-f_i(x)) \\ \text{s.t.} & Ax = b. \end{array} \tag{5}$$

- The term
$$\phi(x) = -\sum_{i=1}^{m} \log(-f_i(x)),$$
with $\mathbf{dom}\phi = \{x \in \mathbb{R}^n : f_i(x) < 0,\ i = 1, \ldots, m\}$, is called the *logarithmic barrier* or *log barrier* of (1).
- Notice (5) is convex since $-(1/t)\log(-u)$ is convex and increasing in u.
- Thus, with an appropriate closedness, Newton's method, for instance, can be used to solve it.

Logarithmic barrier(*cont'd*)

- Quality of (5) as approximation of (1) improves as t grows as will be seen.
- On the other hand, lager t makes minimization of $f_0 + (1/t)\phi$ via Newton's method difficult as Hessian varies rapidly near boundary of feasible region.
- This can be circumvented by solving a sequence of (5), increasing t at each iteration, starting at the solution of the previous t.

Logarithmic barrier(*cont'd*)

Gradient and hessian of log barrier

Note that

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

and

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

Logarithmic barrier(*cont'd*)

KKT conditions of log barrier approximation (5) and central path

Consider following equivalent form of (5)

$$\begin{array}{ll} \min & tf_0(x) + \phi(x) \\ \text{s.t.} & Ax = b. \end{array} \tag{6}$$

Let $x^*(t)$ be optimal solution of (6). From KKT conditions, $\exists\ \hat{\nu} \in \mathbb{R}^p$ such that

$$\begin{gathered} Ax^*(t) = b,\ f_i(x^*(t)) < 0, \\ \nabla f_0(x^*(t)) + \sum_{i=1}^m \frac{1}{-tf_i(x^*(t))} \nabla f_i(x^*(t)) + A^T(\hat{\nu}/t) = 0. \end{gathered} \tag{7}$$

The set $\{x^*(t) : t > 0\}$ is the called the *central path* associated to (1).

Logarithmic barrier(*cont'd*)

"Minimal" preliminaries on Lagrangian dual function

Consider following function, called *Lagrangian* associated to (1):

$$L(x,\lambda,\nu) = f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \nu^T Ax, \text{ for some } \lambda \geq 0,\ \nu. \tag{8}$$

- As $\lambda \geq 0$, $L(x,\lambda,\nu) \leq f(x)$ $\forall$ feasible $x \in \mathrm{dom}(f)$.
- Hence $g(\lambda,\nu) := \inf_{x\in \mathrm{dom}(f)} L(x,\lambda,\nu) \leq p^*$.
- In other words, for any $\lambda \geq 0$, $g(\lambda,\nu)$ is a lower bound on p^*.
- In particular, if (1) is convex and $\bar{x} \in \mathrm{dom} f$, $\bar{\lambda} \geq 0$, and $\bar{\nu}$ satisfy

$$\nabla f_0(\bar{x}) + \sum_{i=1}^m \bar{\lambda}_i \nabla f_i(\bar{x}) + A^T \bar{\nu} = 0, \tag{9}$$

then $L(\bar{x},\bar{\lambda},\bar{\nu}) = \inf_{x\in \mathrm{dom}(f)} L(\bar{x},\bar{\lambda},\bar{\nu}) = g(\bar{\lambda},\bar{\nu})$ is a lower bound on p^*.

Logarithmic barrier(*cont'd*)

From (7), we can observe that $x^*(t)$, $\lambda_i^*(t) = -\frac{1}{tf_i(x^*(t))} > 0$, $i = 1, \ldots, m$, and $\nu^*(t) = \hat{\nu}/t$ satisfy (9). Hence the following is a lower bound on p^*:

$$
\begin{aligned}
L(x^*(t), \lambda^*(t), \nu^*(t)) &= f_0(x^*(t)) + \sum_{i=1}^{m} \lambda_i^*(t) f_i(x^*(t)) + \nu^*(t)^T (Ax^*(t) - b) \\
&= f_0(x^*(t)) - m/t \\
&\leq p^*
\end{aligned}
$$

Thus, $f_0(x^*(t)) - p^* \leq m/t$; as t grows $x^*(t)$ gets closer to x^* as predicted.

Logarithmic barrier(*cont'd*)

Interpretation of central path via KKT conditions

From (7), $x = x^*(t)$ from central path 'almost' satisfies the KKT conditions:

$$
\begin{gathered}
Ax = b,\ f_i(x) \le 0, \\
\lambda \ge 0, \\
\nabla f_0(x) + \textstyle\sum_{i=1}^m \lambda_i \nabla f_i(x) + A^T \nu = 0, \\
-\lambda_i f_i(x) = 1/t,\ i = 1, \dots, m.
\end{gathered}
$$

Algorithm (Barrier method)

given *interior solution* x, $t := t^{(0)}$, $\mu > 1$, *tolerance* $\epsilon > 0$.

repeat

1. *Centering step: Starting at* x, *compute* $x^*(t)$ *by solving (6).*
2. *Update* $x := x^*(t)$.
3. *Quit if* $m/t < \epsilon$.
4. *Increase* t: $t := \mu t$.

- First proposed by Fiacco and McCormick in the 1960s, in name of *SUMT*.
- We assume to use Newton's method for Centering step.

- Accuracy of centering
- Choice of μ
- Choice of $t^{(0)}$

Newton step for centering from Newton step for modified KKT

The Newton step Δx_{nt} for (6) is given by

$$\begin{bmatrix} t\nabla^2 f_0(x) + \nabla^2\phi(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{\text{nt}} \\ \nu_{\text{nt}} \end{bmatrix} = - \begin{bmatrix} t\nabla f_0(x) + \nabla\phi(x) \\ 0 \end{bmatrix}. \qquad (10)$$

We will derive Δx_{nt} of (10) from the Newton step for the modified KKT conditions:

$$\begin{gathered} Ax = b, \\ \nabla f_0(x) + \textstyle\sum_{i=1}^m \lambda_i \nabla f_i(x) + A^T\nu = 0, \\ -\lambda_i f_i(x) = 1/t, \ i = 1, \ldots, m. \end{gathered} \qquad (11)$$

In doing so, we first eliminate $\lambda_i = -\frac{1}{tf_i(x)}$ to get

$$\begin{gathered} Ax = b \\ \nabla f_0(x) + \textstyle\sum_{i=1}^m \frac{1}{-tf_i(x)} \nabla f_i(x) + A^T\nu = 0. \end{gathered} \qquad (12)$$

Newton step for centering from Newton step for modified KKT(*cont'd*)

To find the Newton step for solving (12), we consider its Taylor approximation, by

$$\nabla f_0(x+v) + \sum_{i=1}^m \frac{1}{-tf_i(x+v)} \nabla f_i(x+v)$$

$$\approx \nabla f_0(x) + \sum_{i=1}^m \frac{1}{-tf_i(x)} \nabla f_i(x) + \nabla^2 f_0(x) v \tag{13}$$

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

Using this approximation in place of nonlinear terms, we get

$$Hv + A^T \nu = -g, \; Av = 0, \tag{14}$$

where,

$$H = \nabla^2 f_0(x) + \sum_{i=1}^m \frac{1}{-tf_i(x)} \nabla^2 f_i(x) + \sum_{i=1}^m \frac{1}{tf_i(x)^2} \nabla f_i(x) \nabla f_i(x)^T$$
$$g = \nabla f_0(x) + \sum_{i=1}^m \frac{1}{-tf_i(x)} \nabla f_i(x).$$

Newton step for centering from Newton step for modified KKT(*cont'd*)

But,

$$H = \nabla^2 f_0(x) + (1/t)\nabla^2\phi(x), \; g = \nabla f_0(x) + (1/t)\nabla\phi(x).$$

Hence Δx_{nt} and ν_{nt} of (10) satisfy

$$tH\Delta x_{\text{nt}} + A^T\nu_{\text{nt}} = -tg, \; A\nu_{\text{nt}} = 0.$$

Comparing this with (14), we get

$$v = \Delta x_{\text{nt}}, \; \nu = (1/t)\nu_{\text{nt}}.$$

Hence the Newton's direction for centering step is the same as the Newton's direction for solving the modified KKT conditions.

Homework

11.1, 11.2, 11.3, 11.4, 11.9, 11.10.