IE 521: Convex Optimization **Spring 2017, UIUC**

Lecture 21: Interior Point Method - Part V – April 12

Instructor: Niao He *Scribe: Shuanglong Wang*

Courtesy warning: These notes do not necessarily cover everything discussed in the class. Please email TA (swang157@illinois.edu) if you find any typos or mistakes.

In this lecture, we cover the following topics

- Revisit barrier method
- Self-concordant barrier
- Restate path-following scheme

Reference: Nesterov, Introductory Lectures on Convex Optimization, 2004, Chapter 4.2

21.1 Revisit Barrier Method

Recall the barrier method solves the general convex problem

$$\min_{x \in X} f(x), \quad \text{where } X = \{x : g_i(x) \leq 0, i = 1, ..., m\}$$

by solving a sequence of unconstrained minimization: $\min_x \ tf(x) + F(x)$, where $F(x)$ is barrier function defined on $\text{int}(X)$, and $t > 0$ is penalty parameter.

Without loss of generality, let us consider problem of the form,

$$\min_{x \in X} \quad c^T x \tag{P}$$

where X is a closed bounded convex set with non-empty interior. Let F be standard self-concordant with $cl(dom(F)) = X$. We want to solve (P) by tracing the central path

$$x^*(t) = \arg\min_x \underbrace{tc^T x + F(x)}_{F_t(x)}$$

By optimality condition: $\forall t > 0$,

$$tc + \nabla F(x^*(t)) = 0$$

Note that $F_t(x)$ is standard self-concordant and (damped) Newton method achieves a local convergence when $\lambda_{F_t}(x) \leq \frac{1}{4}$.

When increasing $t \to t'$, we would want to preserve $\lambda_{F_{t'}}(x^*(t)) \leq \frac{1}{4}$ and make t' as large as possible. We have

$$\begin{aligned}\lambda_{F_{t'}}(x^*(t)) &= \| \nabla F_{t'}(x^*(t)) \|_{x^*(t),*} \\ &= \| (t'-t)c + tc + \nabla F(x^*(t)) \|_{x^*(t),*} \\ &= \| (t'-t)c \|_{x^*(t),*} \\ &= \| (\frac{t'}{t} - 1)\nabla F(x^*(t)) \|_{x^*(t),*}\end{aligned}$$

Hence,

$$\frac{t'}{t} = 1 + \frac{1}{4\lambda_F(x^*(t))}$$

To ensure $t \to +\infty$, need $\lambda_F(x)$ to be uniformly bounded from above, namely,

$$\lambda_F^2(x) = \nabla F(x)^T[\nabla^2 F(x)]^{-1}\nabla F(x) \leq \nu$$

for some ν. This leads to the definition of self-concordant barriers.

21.2 Self-concordant Barriers

21.2.1 Definition and Examples

Definition 21.1 *Let $v \geq 0$. We call F a ν-self-concordant barrier (v-s.c.b.) for set $X = cl(dom(F))$ if F is standard self-concordant and satisfies*

$$|DF(x)[h]| \leq v^{1/2}\sqrt{D^2F(x)[h,h]} \quad \forall x \in dom(F), h \in \mathbf{R}^n \tag{$\star$}$$

Remark

1. The inequality $(\star)$ implies that $|\nabla F(x)^T h| \leq \nu \| h \|_x^2$, i.e. F is Lipschitz continuous, w.r.t. the local norm defined by F.
2. When F is non-degenerate, $(\star)$ is equivalent to
$$\lambda_F^2(x) = \nabla F(x)^T[\nabla^2 F(x)]^{-1}\nabla F(x) \leq \nu$$
3. The following are equivalent:
$$(\star) \iff \nabla^2 F(x) \succcurlyeq \frac{1}{\nu}\nabla F(x)[\nabla F(x)]^T$$

Examples

- Constant function: $f(x) = c$ is 0-s.c.b.

- Linear function: $f(x) = a^T x + c (a \neq 0)$, is not s.c.b.
- Quadratic function: $f(x) = \frac{1}{2} x^T Q x + q^T x + c \; Q \succ 0$ is not s.c.b.
- Logarithmic function: $f(x) = -\ln x \; (x > 0)$ is 1-s.c.b.

$$\frac{(f'(x))^2}{f''(x)} = (-\frac{1}{x})^2 x^2 = 1$$

21.2.2 Self-concordance Preserving Operators

Proposition 21.2 *The following are true:*

1. *If $F(x)$ is a ν-self-concordant barrier, then $\tilde{F}(y) = F(Ay + b)$ is a ν-self-concordant barrier.*
2. *If $F_i(x)$ is ν_i-self-concordant barrier, $i = 1, 2$, then $F_1(x) + F_2(x)$ is $(\nu_1 + \nu_2)$-self-concordant barrier.*
3. *If $F(x)$ is a ν-self-concordant barrier, then $\beta F(x)$ with $\beta \geq 1$ is a $(\beta\nu)$-self-concordant barrier.*

Example: The function

$$F(x) = -\sum_{i=1}^{m} \ln(b_i - a_i^T x)$$

is m-self-concordant barrier for the set $\{x : Ax \leq b\}$

Remark: Indeed, for any closed convex set $X \subseteq \mathbf{R}^n$ with non-empty interior, there exists a (βn)-self-concordant barrier for X.

21.2.3 Properties of Self-concordant Barriers

Lemma 21.3 *(Boundedness) Let f be a ν-self-concordant barrier for X. Then for any $x \in int(X), y \in X$, we have $\langle \nabla F(x), y - x \rangle \leq \nu$*

Proof is omitted and can be found in Theorem 4.2.4. in (Nesterov, 2004).

Theorem 21.4 *For any $t > 0$, we have*

$$c^T x^*(t) - \min_{x \in X} c^T x \leq \frac{\nu}{t}$$

Proof:

$$c^T x^*(t) - c^T y = -t^{-1} \nabla F(x^*(t))^T (x^*(t) - y) \leq \frac{\nu}{t}$$

■

21.3 Restate Path-following Scheme

We would like to address the convex problem

$$\min_{x \in X} c^T x$$

by tracing the central path

$$x^*(t) = \arg\min_x \left\{ F_t(x) := tc^T x + F(x) \right\}$$

where $F(x)$ is a ν-self-concordant barrier for $X = cl(dom(F))$.

Theorem 21.5 *For any $t > 0$, we have*

$$c^T x^*(t) - \min_{x \in X} c^T x \leq \frac{\nu}{t}$$

Proof: By optimality condition: $tc + \nabla F(x^*(t)) = 0$, i.e. $c = -t^{-1} \nabla F(x^*(t))$,

$$\begin{aligned} \forall y \in X : \quad c^T x^*(t) - c^T y &= -t^{-1} \nabla F(x^*(t))^T (x^*(t) - y) \\ &= t^{-1} \nabla F(x^*(t))^T (y - x^*(t)) \\ &\leq \frac{\nu}{t} \end{aligned}$$

■

Now consider an approximate solution x that is close to $x^*(t)$:

$$\lambda_{F_t}(x) \leq \beta$$

where β is small enough.

Theorem 21.6 *If $\lambda_{F_t}(x) \leq \beta$,*

$$c^T x - \min_{x \in X} c^T x \leq \frac{1}{t}(\nu + \frac{\sqrt{\nu}\beta}{1 - \beta})$$

Proof: First of all,

$$\begin{aligned} c^T x - c^T x^*(t) &\leq \| c \|_{x^*(t),*} \cdot \| x - x^*(t) \|_{x^*(t)} \\ &= t^{-1} \| \nabla F(x^*(t)) \|_{x^*(t),*} \cdot \| x - x^*(t) \|_{x^*(t)} \end{aligned}$$

Recall from last lecture that for $x^* = \arg\min_x f(x)$ with standard self concordant f:

$$\| x - x^* \|_{x^*} \leq \frac{\lambda_f(x)}{1 - \lambda_f(x)}$$

Hence,

$$\| x - x^*(t) \|_{x^*(t)} \leq \frac{\lambda_{F_t}(x)}{1 - \lambda_{F_t}(x)} \leq \frac{\beta}{1-\beta}$$

Thus, $c^T x - c^T x^*(t) \leq \frac{\sqrt{\nu}}{t} \frac{\beta}{1-\beta}$

$$\begin{aligned} c^T x - \min_{x \in X} c^T x &= c^T x - c^T x^*(t) + c^T x^*(t) - \min_{x \in X} c^T x \\ &\leq \frac{\sqrt{\nu}}{t} \frac{\beta}{1-\beta} + \frac{\nu}{t} \end{aligned}$$

■

Now we can formally describe the path-following scheme.

Path-following Scheme

- Initialize (x_0, t_0) with $t_0 > 0$ and $\lambda_{F_{t_0}}(x_0) \leq \beta$ ($\beta \in (0, \frac{1}{4})$)
- For $k \geq 0$, do

$$\begin{aligned} t_{k+1} &= t_k(1 + \frac{\gamma}{\sqrt{\nu}}) \\ x_{k+1} &= x_k - [\nabla^2 F(x_k)]^{-1}[t_{k+1} c + \nabla F(x_k)] \end{aligned}$$

Theorem 21.7 (Rate of convergence) *In the above scheme, one has*

$$c^T x_k - \min_{x \in X} c^T x \leq O(1) \frac{\nu}{t_0} \exp\left\{ -O(1) \frac{k}{\sqrt{\nu}} \right\}$$

where the constant factor $O(1)$ depends solely on β and γ.

Remark [Complexity]

- The total complexity of Newton steps does not exceed

$$N(\epsilon) \leq O\left(\sqrt{\nu} \log \frac{\nu}{t_0 \epsilon} \right)$$

- The total arithmetic cost of finding an ϵ-solution by the above scheme does not exceed

$$O\left(M \sqrt{\nu} \log \frac{\nu}{t_0 \epsilon} \right)$$

  where M is the arithmetic cost for computing $\nabla F(x)$, $\nabla^2 F(x)$ and solving a Newton system.

Remark [Initialization] In order to obtain a fair (x_0, t_0), s.t.

$$\lambda_{F_{t_0}}(x_0) \leq \beta$$

one can apply the following trick. Suppose $\hat{x} \in dom(F)$ is given. Consider the auxiliary path

$$y^*(t) = \arg\min_x [-t \nabla F(\hat{x})^T x + F(x)]$$

When $t = 1$, $y^*(1) = \hat{x}$. When $t \to 0$, $y^*(t) = x_F := \arg\min_x F(x)$.

We can trace $y^*(t)$ as t decreases from 1 to 0, until we approach to a point (y_0, t_0) such that $\lambda_{F_{t_0}}(x_0) \leq \beta$.