IE 521 Convex Optimization

Lecture 16: Interior Point Method

Path Following Scheme & Self-Concordance

Niao He

16th April 2019

Outline

Path Following Scheme

Self-concordant Function

- Self-concordant in $\mathbb{R}$
- Self-concordant in $\mathbb{R}^n$
- Calculus
- Geometric Properties

Recall

- Interior-point methods play an important role in convex optimization.
- Modern LP/SOCP/SDP solvers, such as SeDuMi, SDPT3 are built on interior-point methods.

Historical Note

- **1984:** Karmarkar introduced poly-time interior point method for LP
- **late-1980s:** Renegar & Gonzaga introduced path-following interior point method for LP
- **1988:** Nesterov and Nemirovski extended interior point method for convex programs
- **after 1990s:** many solvers for convex programs

Problem Setting

$$\begin{aligned} \min_x \quad & f(x) \\ \text{s.t.} \quad & g_i(x) \leq 0, i = 1, ..., m \end{aligned} \tag{P}$$

$$X = \{x : g_i(x) \leq 0, \forall i = 1, \ldots, m\}$$

- f, g_i are twice continuously differentiable and convex
- Slater condition holds
- The feasible domain X is bounded

Remark. X is convex compact and has non-empty interior.

Path Following Scheme

Barrier Method: Solve a series of unconstrained problems

$$\min_x \quad t \cdot f(x) + F(x) \qquad (P_t)$$

where $t > 0$ is a penalty parameter and $F(x)$ is a **barrier function** that satisfies:

- $F : \text{int}(X) \to \mathbb{R}$ and $F(x) \to +\infty$ as $x \to \partial(X)$
- F is twice continuously differentiable and convex
- F is *non-degenerate*, i.e. $\nabla^2 F(x) \succ 0, \forall x \in \text{int}(X)$

Remark. For any $t > 0$, (P_t) has a unique solution in $\text{int}(X)$.

Path Following Scheme

Central Path: the path $\{x^*(t), t > 0\}$ where

$$x^*(t) = \underset{x}{\operatorname{argmin}} \{t \cdot f(x) + F(x)\}$$

Remark.

$$x^*(t) \longrightarrow x^*, \text{as } t \longrightarrow \infty$$

Question: Need to specify

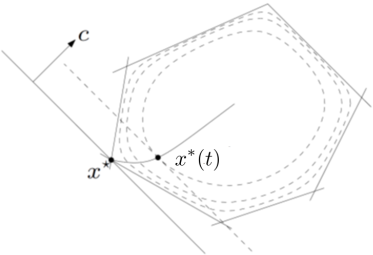
1. the barrier function $F(x)$?
2. the method to solve unconstrained problems (P_t)?
3. the policy to update the penalty parameter t?

Illustration: Linear Program

$$\begin{aligned} \min_x \quad & c^T x \\ \text{s.t.} \quad & a_i^T x \le b_i, i = 1, ..., m \end{aligned} \qquad (P)$$

Logarithmic Barrier

$$\min_x \quad c^T x - \frac{1}{t} \cdot \sum_{i=1}^m \ln(b_i - a_i^T x) \qquad (P_t)$$



Self-concordant Function in $\mathbb{R}$

Definition. $f : \mathbb{R} \to \mathbb{R}$ is self-concordant if f is convex and

$$|f'''(x)| \leq \kappa f''(x)^{3/2}, \forall x \in \mathrm{dom}(f)$$

for some constant $\kappa \geq 0$.

- When $\kappa = 2$, f is called standard self-concordant.

Example 1. Logarithmic function: $f(x) = -\ln(x), x > 0$ is standard self-concordant:

$$f'(x) = -\frac{1}{x}, f''(x) = \frac{1}{x^2}, f'''(x) = -\frac{2}{x^3}, \quad \frac{|f'''(x)|}{f''(x)^{3/2}} = 2$$

Exercise: self-concordant or not?

- Linear function: $f(x) = cx$
- Quadratic function: $f(x) = \frac{a}{2}x^2 + bx + c \ (a > 0)$
- Exponential function: $f(x) = e^x$
- Power functions:

$$f(x) = \frac{1}{x^p}(p > 0), (x > 0)$$
$$f(x) = |x|^p(p > 2)$$
$$f(x) = x^{2p}(p > 2)$$

Self-concordant Function is Affine Invariant

Proposition. If $f(x)$ is self-concordant, $\tilde{f}(y) = f(ay+b)$ is also self-concordant with the same constant.

Proof.

- First, $\tilde{f}$ is convex;
- Second, it is easy to show that

$$\frac{|\tilde{f}'''(y)|}{\tilde{f}''(y)^{3/2}} = \frac{|a^3 f'''(ay+b)|}{\left[a^2 f''(ay+b)\right]^{3/2}} = \frac{|f'''(ay+b)|}{f''(ay+b)^{3/2}} \leq \kappa.$$

Self-concordant Function in $\mathbb{R}^n$

Definition. $f : \mathbb{R}^n \to \mathbb{R}$ is self-concordant if it is self-concordant along every line, i.e., $\forall x \in \mathrm{dom}(f), h \in \mathbb{R}^n$,

$$\phi(t) = f(x + th)$$

is self-concordant with some constant $\kappa \geq 0$.

- When $\kappa = 2$, f is called standard self-concordant.

Example 2. Logarithmic function: $f(x) = -\ln(b - a^T x)$ is standard self-concordant on its domain.

Equivalent Definition

Denote the k-th differential of f taken at $x \in \mathrm{dom}(f)$ along the directions $h_1, ..., h_k$:

$$D^k f(x)[h_1, ..., h_k] = \frac{\partial^k}{\partial t_1 ... \partial t_k}|_{t_1 = ... = t_k = 0} f(x + t_1 h_1 + ... + t_k h_k)$$

Definition. A function $f : \mathbb{R}^n \to \mathbb{R}$ is self-concordant if

$$D^3 f(x)[h, h, h] \leq \kappa (D^2 f(x)[h, h])^{3/2}, \forall x \in \mathrm{dom}(f), h \in \mathbb{R}^n$$

for some constant $\kappa \geq 0$.

Example: Logarithmic Quadratic Function

Example 3. The function below is standard self-concordant

$$f(x) = -\ln\left(-\frac{1}{2}x^T Qx + b^T x + c\right), \text{ where } Q \succeq 0.$$

Denote $q(x) = -\frac{1}{2}x^T Qx + b^T x + c$.

- Note that $f(x)$ is convex
- $Df(x)[h] = -\frac{1}{q(x)}(b^T h - x^T Qh) := \omega_1$
- $D^2 f(x)[h,h] = \frac{1}{q^2(x)}(b^T h - x^T Qh)^2 + \frac{1}{q(x)}h^T Qh := \omega_1^2 + \omega_2$
- $D^3 f(x)[h,h,h] = -\frac{2}{q^3(x)}(b^T h - x^T Qh)^3 - \frac{3}{q^2(x)}(b^T h - x^T Qh)h^T Qh = 2\omega_1^3 + 3\omega_1\omega_2$

$$\frac{|D^3 f(x)[h,h,h]|}{(D^2 f(x)[h,h])^{3/2}} = \frac{|2\omega_1^3 + 3\omega_1\omega_2|}{(\omega_1^2 + \omega_2)^{3/2}} \leq 2$$

Operations Preserving Self-Concordance

1. Affine invariant: If $f(y)$ is self-concordant with constant κ, then the function

$$\tilde{f}(x) = f(Ax + b)$$

is also self-concordant with constant κ.

2. Summation: If $f_1(x)$ and $f_2(x)$ are self-concordant with constants κ_1, κ_2, then the function

$$\tilde{f}(x) = f_1(x) + f_2(x)$$

is self-concordant with constant $\kappa = \max\{\kappa_1, \kappa_2\}$

3. Scaling: If $f(x)$ is self-concordant with constant κ, and $\alpha > 0$ then the function

$$\tilde{f}(x) = \alpha f(x)$$

is also self-concordant with $\kappa = \frac{\kappa}{\sqrt{\alpha}}$

Example

Example 4. $f(x) = -\sum_{i=1}^{m} \ln(b_i - a_i^T x)$ is standard self-concordant on $\text{int}(X)$, where

$$X = \left\{ x : a_i^T x \leq b_i, i = 1, ..., m \right\}.$$

Example 5.
$f(x_1, x_2) = -\log(x_2^2 - x_1^2) - 2\log(x_1) - 3\log(x_2)$ is self-concordant on $\text{int}(X)$, where

$$X = \{(x_1, x_2) : 0 \leq x_1 \leq x_2\}.$$

Remark. Note that $f(x)$ is also a valid barrier function on X.

Local Norm

Definition. The local norm of h at $x \in dom(f)$ as

$$\|h\|_x = \sqrt{h^T \nabla^2 f(x) h}.$$

Proposition. For standard self-concordant function f, it holds that

$$\left| D^3 f(x)[h_1, h_2, h_3] \right| \leq 2\|h_1\|_x \cdot \|h_2\|_x \cdot \|h_3\|_x$$

Remark. ("Lipschitz continuity") at a high level,

$$\left| \frac{d}{dt}|_{t=0} D^2 f(x + t\delta)[h, h] \right| \leq 2\|\delta\|_x D^2 f(x)[h, h]$$

The second derivative is relatively Lipschitz continuous w.r.t. the local norm defined by f.

Illustration in $\mathbb{R}$

Let f be 1-self-concordant on $\mathbb{R}$ and strictly convex

$$\frac{|f'''(x)|}{|f''(x)|^{3/2}} \leq 1 \Rightarrow \left|\frac{d}{dx}[f''(x)^{-1/2}]\right| \leq 1$$

$$\Rightarrow -y \leq \int_0^y \frac{d}{dx}[f''(x)]^{-1/2} dx \leq y$$

$$\Rightarrow -y \leq \frac{1}{\sqrt{f''(y)}} - \frac{1}{\sqrt{f''(0)}} \leq y$$

Simplifying the above terms, we arrive at $\forall 0 \leq y < (\sqrt{f''(0)})^{-1}$

$$\frac{f''(0)}{(1 + y\sqrt{f''(0)})^2} \leq f''(y) \leq \frac{f''(0)}{(1 - y\sqrt{f''(0)})^2} \qquad (\star)$$

Illustration in $\mathbb{R}$

We can further derive that $\forall 0 \le y < (\sqrt{f''(0)})^{-1}$

$$\frac{(y\sqrt{f''(0)})^2}{1+y\sqrt{f''(0)}} \le y(f'(y)-f'(0)) \le \frac{(y\sqrt{f''(0)})^2}{1-y\sqrt{f''(0)}} \qquad (\star\star)$$

$$f(y)-f(0)-f'(0)y \ge y\sqrt{f''(0)} - \ln(1+y\sqrt{f''(0)}) \qquad (\star\star\star)$$

$$f(y)-f(0)-f'(0)y \le -y\sqrt{f''(0)} - \ln(1-y\sqrt{f''(0)}) \qquad (\star\star\star)$$

Relative Lipschitz of Hessian and Gradient

Definition. **(Dikin Ellipsoid)**

$$W_r(x) = \{y : \|y - x\|_x \leq r\}$$

$$W_r^o(x) = \{y : \|y - x\|_x < r\}$$

Proposition. For $x \in \mathrm{dom}(f)$, we have $W_1^o(x) \subseteq \mathrm{dom}(f)$.

Theorem. For $x \in \mathrm{dom}(f)$, we have $\forall y \in W_1^o(x)$:

$$(1 - \|y - x\|_x)^2 \, \nabla^2 f(x) \preceq \nabla^2 f(y) \preceq (1 - \|y - x\|_x)^{-2} \, \nabla^2 f(x)$$

$$\frac{\|y - x\|_x^2}{1 + \|y - x\|_x} \leq \langle \nabla f(y) - \nabla f(x), y - x \rangle \leq \frac{\|y - x\|_x^2}{1 - \|y - x\|_x}$$

Linear Approximation

Theorem. For $x \in \text{dom}(f)$, we have

$$f(y) \geq f(x) + \langle f'(x), y - x \rangle + \omega(\| y - x \|_x), \forall y \in \text{dom}(f)$$

$$f(y) \leq f(x) + \langle f'(x), y - x \rangle + \omega^*(\| y - x \|_x), \forall y \in W_1^o(x)$$

where $\omega(t) = t - \ln(1 + t)$ and $\omega^*(t) = -t - \ln(1 - t)$.

Remark. Check Theorems 4.1.6-8 in (Nesterov, 2004).

References

- Nemirovski (2004), Interior Point Polynomial Time Methods in Convex Programming, Chapter 1
- Nesterov (2004), Introductory Lectures on Convex Optimization, Chapter 4.1