Lecture 17: Interior Point Method, Part I – March 29

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

- Path-following Scheme
- Self-concordant Functions

Reference:

Nemirovski, *Interior Point Polynomial Time Methods in Convex Programing*, 2004, Chapter 1
Nesterov, *Introductory Lectures on Convex Optimization*, 2004, Chapter 4.1

17.1 Historical Notes

- 1947: Dantzig, Simplex method for LP
- 1973: Klee and Minty proved that Simplex Method is not a polynomial-time algorithm
- mid-1970s: Shor, Nemirovski and Yudin, Ellipsoid method for linear and convex programs
- 1979: Khachiyan proved the polynomial-time solvability of LP
- 1984: Karmarkar, polynomial algorithm (potential reduction interior point method) for LP (1967): Dikin, affine scaling algorithm (simplification of Karmarkar's algorithm) for LP
- late-1980s: Renegar, Gonzaga, path-following interior point method for LP
- 1988: Nesterov and Nemirovski, extend interior point method for convex programs

17.2 Path following Scheme

We intend to solve a general convex program

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

where f, g_i are twice continuously differentiable convex functions.

Denote $X = \{x : g_i(x) \leq 0, \forall i = 1, \dots, m\}$ as the feasible domain. Assume Slater condition holds and X is bounded, so X is a compact convex set with non-empty interior.

Barrier Method: solve a series of unconstrained problems

$$\min_x \quad tf(x) + F(x) \tag{P_t}$$

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

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

Note that for any $t > 0$, (P_t) has a unique solution in the interior of X.

Denote

$$x^*(t) = \arg\min_x \quad tf(x) + F(x)$$

The path $\{x^*(t), t > 0\}$ is called the **central path**. We have

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

To implement the above path-following scheme, need to specify:

1. the barrier function $F(x)$:
   - use self-concordant barriers, e.g. $F(x) = -\sum_{i=1}^m \log(-g_i(x))$
2. the method to solve unconstrained minimization problems (P_t):
   - use Newton method
3. the policy to update the penalty parameter t.

17.3 Self-concordant Function

Definition 17.1 *A function $f : \mathbf{R} \to \mathbf{R}$ is self-concordant if f is convex and*

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

for some constant $\kappa \geq 0$.

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

Examples:

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

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

- Linear function: $f(x) = cx$ is self-concordant with constant $\kappa = 0$:

$$f'(x) = c, f''(x) = 0, f'''(x) = 0$$

- Convex quadratic function: $f(x) = \frac{a}{2}x^2 + bx + c$ $(a > 0)$ is self-concordant with $\kappa = 0$:

$$f'(x) = ax + b, f''(x) = a, f'''(x) = 0$$

- Exponential function: $f(x) = e^x$ is <u>not</u> self-concordant:

$$f'(x) = f''(x) = f'''(x) = e^x, \frac{|f'''(x)|}{f''(x)^{3/2}} = e^{x/2} \longrightarrow +\infty \text{ as } x \to -\infty$$

- Power functions:

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

  are <u>not</u> self-concordant.

Remark: Self-concordant function is **<u>affine-invariant</u>**: If $f(x)$ is self-concordant, $\tilde{f}(y) = f(ay+b)$ is also self-concordant with the same constant.

Proof: It is easy to see that $\tilde{f}$ is convex and

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

■

When extending to a function f defined on $\mathbf{R}^n$ and $f \in \mathbf{C}^3(\mathbf{R}^n)$, we say f is self-concordant if it is self-concordant along every line, namely, $\forall x \in dom(f), h \in \mathbf{R}^n$, $\phi(t) = f(x + th)$ is self-concordant with some constant $\kappa \geq 0$.

Denote

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

as the k-th differential of f taken at x along the directions $h_1, ..., h_k$ e.g.

$$\begin{aligned} Df(x)[h] &= \phi'(0) = \langle \nabla f(x), h \rangle \\ D^2 f(x)[h, h] &= \phi''(0) = \langle \nabla^2 f(x) h, h \rangle \end{aligned}$$

Definition 17.2 *We say a function $f : \mathbf{R}^n \to \mathbf{R}$ is self-concordant if*

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

for some constant $\kappa \ge 0$

Operations Preserving Self-Concordance

Proposition 17.3

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 $\alpha f(x)$ is also self-concordant with $\kappa = \frac{\kappa}{\sqrt{\alpha}}$*

Proof: Exercise! ∎

Hence, it is straightforward to see that

Corollary 17.4 *$f(x) = -\sum_{i=1}^{m} \ln(b_i - a_i^T x)$ is standard self-concordant on int(X), where $X = \left\{x : a_i^T x \le b_i, i = 1, ..., m\right\}$*

Remark: Indeed, under regular conditions, self-concordant functions are barrier functions:

- If $dom(f)$ contains no straight line, then $\nabla^2 f(x)$ is non-degenerate (strictly convex)
- If f is closed convex, then $f(x_k) \to +\infty$ if $\{x_k\} \subseteq dom(f)$ and $x_k \to \bar{x} \in \partial(dom(f))$