ISEN 629: Engineering Optimization

Lecture 19

Sergiy Butenko

Industrial and Systems Engineering
Texas A&M University

Fall 2007

Introduction

We consider the following problem:

$$\min_{x \in \mathbb{R}^n} \{f_0(x) : f_j(x) \leq 0, j = 1, \ldots, m\},$$

where all functions involved are convex.

- All numerical methods we considered so far are based on the black-box concept.
- We only used some information on the functional components of the problem at some test points.
- Even though the developed numerical methods assume convexity, the only time we pay attention to the structure of the function (i.e., convexity) is <u>before</u> we apply a numerical method.
- Can we utilize the structural information in numerical schemes?

Introduction

A common problem-solving process:

1. Find a class of problems that can be solved efficiently using a particular method.
2. Describe the transformation rules for reducing a given problem to a problem from the above class.
3. Describe the class of problems for which these transformation rules are applicable.

With respect to the considered optimization problem, our goals are

- To derive a family of special convex functions, the self-concordant functions and self-concordant barriers, which can be efficiently minimized using the Newton method.
- To develop a technique for transforming our original problem into its barrier model.

Newton method revisited

Newton method: $\quad x_{k+1} = x_k - [f''(x_k)]^{-1} f'(x_k)$

How does the Newton method behave when we apply an affine transformation?
For a nondegenerate $n \times n$-matrix A, consider the function

$$\phi(y) = f(Ay).$$

Lemma (4.1.1, Affine invariance of the Newton method)
In addition to the sequence $\{x_k : k \geq 0\}$ generated by the Newton method for f, consider the sequence $\{y_k : k \geq 0\}$ generated by the Newton's method for ϕ:

$$y_{k+1} = y_k - [\phi''(y_k)]^{-1} \phi'(y_k), \; k \geq 0,$$

with $y_0 = A^{-1} x_0$. Then $y_k = A^{-1} x_k$ for all $k \geq 0$.

Newton method revisited

Theorem (1.2.5, Local convergence of Newton method)
Consider the problem $\min_{x \in \mathbb{R}^n} f(x)$ *and assume that*

1. $f \in C_M^{2,2}(\mathbb{R}^n)$;
2. *There exists a local minimum* x^* *of* f *with positive definite Hessian:* $f''(x^*) \succeq lI_n$, $l > 0$;
3. *Our starting point* x_0 *is close enough to* x^*: $\|x_0 - x^*\| < \frac{2l}{3M}$.

Then $\|x_k - x^*\| < \frac{2l}{3M}$ *for all* k *and the Newton method converges quadratically:*

$$\|x_{k+1} - x^*\| \le \frac{M\|x_k - x^*\|^2}{2(l - M\|x_k - x^*\|)}.$$

Newton method revisited

- If we choose a new basis in $\mathbb{R}^n$, then all objects in our description will change, including the region of the quadratic convergence (which is defined by the standard inner product).
- However, the Newton method is affine invariant with respect to affine transformation of variables. Thus, its real region of quadratic convergence does not depend on a particular inner product - it depends only on the local topological properties of $f(x)$.
- Can we modify our assumptions to correct this inconsistency?

Newton method revisited

Consider the Lipschitz continuity of the Hessian:

$$\|f''(x) - f''(y)\| \le M\|x - y\|, \ \forall x, y \in \mathbb{R}^n.$$

Assuming that $f \in C^3(\mathbb{R}^n)$, denote by

$$f'''(x)[u] = \lim_{\alpha \to 0} \frac{1}{\alpha}[f''(x + \alpha u) - f''(x)]$$

the 3-rd directional derivative of f along the direction u (this is an $n \times n$-matrix). Then we have

$$\|f'''(x)[u]\| \le M\|u\|,$$

therefore, for any $x \in \mathbb{R}^n$:

$$v^T f'''(x)[u]v \le M\|u\| \cdot \|v\|^2.$$

Newton method revisited

If $v = u$, then

$$v^T f'''(x)[u]v \le M\|u\| \cdot \|v\|^2.$$

becomes

$$u^T f'''(x)[u]u \le M\|u\|^3.$$

- The left-hand side is affine-invariant with respect to x (will show later);
- The right-hand side does not depend on x.
- We will replace the standard norm $\|\cdot\|$ with an affine-invariant norm defined by the Hessian $f''(x)$:

$$\|u\|_{f''(x)} = (u^T f''(x) u)^{1/2}.$$

- This leads to the definition of self-concordant functions.

Defining self-concordant functions

Consider a closed convex function $f(x) \in C^3(\mathrm{dom}\, f)$ with open domain. For a point $x \in \mathrm{dom}\, f$ and a direction $u \in \mathbb{R}^n$, consider the function

$$\phi(x; t) = f(x + tu)$$

as a function of $t \in \mathrm{dom}\, \phi(x; \cdot) \subseteq \mathbb{R}$. Denote by

$$\begin{aligned} Df(x)[u] &= \phi'(x;0) &&= f'(x)^T u, \\ D^2 f(x)[u,u] &= \phi''(x;0) &&= u^T f''(x) u = \|u\|^2_{f''(x)}, \\ D^3 f(x)[u,u,u] &= \phi'''(x;0) &&= u^T f'''(x)[u] u. \end{aligned}$$

(recall that $f'''(x)[u] = \lim_{\alpha \to 0} \frac{1}{\alpha}[f''(x + \alpha u) - f''(x)]$).

Definition (Self-concordant function)

We call function f self-concordant if there exists a constant $M_f \geq 0$ such that the inequality

$$D^3 f(x)[u,u,u] \leq M_f \|u\|^3_{f''(x)} \quad (\Leftrightarrow D^3 f(x)[u,u,u] \leq (D^2 f(x)[u,u])^{3/2})$$

holds for any $x \in \mathrm{dom}\, f$ and $u \in \mathbb{R}^n$.

Defining self-concordant functions

An equivalent definition of self-concordant functions (useful for establishing properties of self-concordant functions) is given by the following lemma:

Lemma (4.1.2)

A function f is self-concordant if and only if for any $x \in \mathrm{dom}\, f$ and any $u_1, u_2, u_3 \in \mathbb{R}^n$ we have

$$|D^3 f(x)[u_1, u_2, u_3]| \leq M_f \prod_{i=1}^{3} \|u_i\|_{f''(x)}.$$

Affine invariance of self-concordance

Theorem (4.1.2)

Let $\mathcal{A}(x) = Ax + b : \mathbb{R}^n \to \mathbb{R}^m$ be a linear operator. Assume that $f(y)$ is self-concordant with constant M_f. Then $\phi(x) = f(\mathcal{A}(x))$ is also self-concordant with $M_\phi = M_f$ (i.e., self-concordance is an affine-invariant property).

Proof.

1. $\phi(x)$ is closed and convex.
2. Fix $x \in \mathrm{dom}\, \phi = \{x : \mathcal{A}(x) \in \mathrm{dom}\, f\}$ and $u \in \mathbb{R}^n$, and denote by $y = \mathcal{A}(x)$, $v = Au$. Then

$$\begin{aligned} D\phi(x)[u] &= \phi'(x)^T u &&= f'(\mathcal{A}(x))^T A u = f'(y)^T v, \\ D^2\phi(x)[u,u] &= u^T \phi''(x) u &&= (Au)^T f''(\mathcal{A}(x))(Au) = v^T f''(y) v, \\ D^3\phi(x)[u,u,u] &= u^T \phi'''(x)[u] u &&= (Au)^T f'''(\mathcal{A}(x))[Au](Au) = D^3 f(y)[v,v,v]. \end{aligned}$$

and
$$\begin{aligned} |D^3\phi(x)[u,u,u]| &= |D^3 f(y)[v,v,v]| \\ &\leq M_f (v^T f''(y) v)^{3/2} = M_f (D^2\phi(x)[u,u])^{3/2}. \end{aligned}$$

□

Examples of self-concordant functions

1. Linear function.

$$f(x) = a^T x + b, \ \mathrm{dom}\, f = \mathbb{R}^n.$$

Then

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

so $f(x)$ is self-concordant with $M_f = 0$.

2. Convex quadratic function.

$$f(x) = \frac{1}{2} x^T A x + b^T x + c, \ \mathrm{dom}\, f = \mathbb{R}^n,$$

where $A = A^T \succeq 0$. Then

$$f'(x) = Ax + b, \ f''(x) = A, \ f'''(x) = 0,$$

so $f(x)$ is self-concordant with $M_f = 0$.

Examples of self-concordant functions

3. Logarithmic barrier for a ray.

$$f(x) = -\ln(x),\ \mathrm{dom}\, f = \{x \in \mathbb{R} : x > 0\}.$$

Then

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

Hence, $f(x)$ is self-concordant with $M_f = 2$.

Examples of self-concordant functions

4. Logarithmic barrier for a region defined by a quadratic. Let $A = A^T \succeq 0$. Consider the concave function

$$\phi(x) = -\frac{1}{2}x^T Ax + b^T x + c.$$

Let $f(x) = -\ln \phi(x),\ \mathrm{dom}\, f = \{x \in \mathbb{R}^n : \phi(x) > 0\}$. Then

$$\begin{aligned} Df(x)[u] &= -\tfrac{1}{\phi(x)}[b^T u - x^T Au], \\ D^2 f(x)[u,u] &= \tfrac{1}{\phi^2(x)}[b^T u - x^T Au]^2 + \tfrac{1}{\phi(x)}u^T Au, \\ D^3 f(x)[u,u,u] &= -\tfrac{2}{\phi^3(x)}[b^T u - x^T Au]^3 - \tfrac{3}{\phi^2(x)}[b^T u - x^T Au]u^T Au \end{aligned}$$

If we denote by $\omega_1 = Df(x)[u]$ and $\omega_2 = \frac{1}{\phi(x)}u^T Au \geq 0$, then

$$\begin{aligned} D^2 f(x)[u,u] &= \omega_1^2 + \omega_2 \geq 0 \text{ (since } D^2 f(x)[u,u] = \|u\|^2_{f''(x)}), \\ |D^3 f(x)[u,u,u]| &= |2\omega_1^3 + 3\omega_1\omega_2|. \end{aligned}$$

Assume that $\omega_1 \neq 0$ (the only nontrivial case). Denote by $\alpha = \omega_2/\omega_1^2$, then

$$\frac{|D^3 f(x)[u,u,u]|}{(D^2 f(x)[u,u])^{3/2}} \leq \frac{2|\omega_1|^3 + 3|\omega_1|\omega_2}{(\omega_1^2 + \omega_2)^{3/2}} = \frac{2(1 + \frac{3}{2}\alpha)}{(1+\alpha)^{3/2}} \leq 2.$$

Thus, $f(x)$ is self-concordant with $M_f = 2$.

Properties of self-concordant functions

Theorem (4.1.1)

Let functions f_1 and f_2 be self-concordant with constants M_1 and M_2, respectively, and let $\alpha, \beta > 0$. Then the function $f(x) = \alpha f_1(x) + \beta f_2(x)$ is self-concordant with constant

$$M_f = \max\left\{\frac{1}{\sqrt{\alpha}}M_1, \frac{1}{\sqrt{\beta}}M_2\right\}$$

and $\mathrm{dom}\, f = \mathrm{dom}\, f_1 \cap \mathrm{dom}\, f_2$.

Properties of self-concordant functions

Corollary (4.1.1)

If f is self-concordant with some constant M_f and $A = A^T \succeq 0$, then the function

$$\phi(x) = f(x) + \frac{1}{2}x^T Ax + b^T x + c$$

is also self-concordant with the constant $M_\phi = M_f$.

Corollary (4.1.2)

Let function f be self-concordant with some constant M_f and $\alpha > 0$. Then the function $\phi(x) = \alpha f(x)$ is also self-concordant with the constant $M_\phi = \frac{1}{\sqrt{\alpha}}M_f$.

Properties of self-concordant functions

Theorem (4.1.3)

Let f be self-concordant. If $\mathrm{dom}\, f$ contains no straight line, then the Hessian $f''(x)$ is nondegenerate at any x from $\mathrm{dom}\, f$.

Theorem (4.1.4)

Let f be self-concordant. Then for any $\bar{x}$ in the boundary of $\mathrm{dom}\, f$ and any sequence

$$\{x_k\} \subset \mathrm{dom}\, f : x_k \to \bar{x}$$

we have $f(x_k) \to +\infty$.

Thus, $f(x)$ is a barrier function for $\mathrm{cl}(\mathrm{dom}\, f)$.