ISEN 629: Engineering Optimization

Lecture 7

Sergiy Butenko

Industrial and Systems Engineering
Texas A&M University

Fall 2007

Lower complexity bounds for $\mathcal{F}_L^{\infty,1}(\mathbb{R}^n)$

We consider the following problem class (note that $\mathcal{F}_L^{\infty,1}(\mathbb{R}^n) \subset \mathcal{F}_L^{1,1}(\mathbb{R}^n)$).

Model:	$\min\limits_{x\in\mathbb{R}^n} f(x),\ f \in \mathcal{F}_L^{1,1}(\mathbb{R}^n)$.
Oracle:	First-order local black box.
Approximate solution:	$\bar{x} \in \mathbb{R}^n,\ f(\bar{x}) - f^* \leq \epsilon$.

Assumption: An iterative method $\mathcal{M}$ generates a sequence of test points $\{x_k\}$ such that

$$x_k \in x_0 + \mathrm{Lin}\{f'(x_0), \ldots, f'(x_{k-1})\},\ k \geq 1.$$

Lower complexity bounds for $\mathcal{F}_L^{\infty,1}(\mathbb{R}^n)$

We will construct a function that appears to be difficult for all schemes satisfying the above assumption.

For some constant $L > 0$, consider the family of quadratic functions

$$f_k(x) = \frac{L}{4}\left\{\frac{1}{2}\left[\left(x^{(1)}\right)^2 + \sum_{i=1}^{k-1}\left(x^{(i)} - x^{(i+1)}\right)^2 + \left(x^{(k)}\right)^2\right] - x^{(1)}\right\}$$

for $k = 1, \ldots, n$. Note that $f_k(x) = \frac{1}{2}x^T f_k''(x)x - \frac{L}{4}x^{(1)}$, so for any $x \in \mathbb{R}^n$ we have:

$$x^T f_k''(x)x = \frac{L}{4}\left[\left(x^{(1)}\right)^2 + \sum_{i=1}^{k-1}\left(x^{(i)} - x^{(i+1)}\right)^2 + \left(x^{(k)}\right)^2\right] \geq 0,$$

i.e., $f_k(x)$ is convex.

Lower complexity bounds for $\mathcal{F}_L^{\infty,1}(\mathbb{R}^n)$

In addition,

$$\begin{aligned}
x^T f_k''(x)x &= \tfrac{L}{4}\left[\left(x^{(1)}\right)^2 + \sum_{i=1}^{k-1}\left(x^{(i)} - x^{(i+1)}\right)^2 + \left(x^{(k)}\right)^2\right] \\
&\leq \tfrac{L}{4}\left[\left(x^{(1)}\right)^2 + \sum_{i=1}^{k-1} 2\left(\left(x^{(i)}\right)^2 + \left(x^{(i+1)}\right)^2\right) + \left(x^{(k)}\right)^2\right] \\
&\leq L\sum_{i=1}^{n}\left(x^{(i)}\right)^2.
\end{aligned}$$

Thus, $0 \preceq f_k''(x) \preceq LI_n$, so $f_k(x) \in \mathcal{F}_L^{\infty,1}(\mathbb{R}^n),\ 1 \leq k \leq n$.

Next, we will compute the minimum of f_k.

Lower complexity bounds for $\mathcal{F}_L^{\infty,1}(\mathbb{R}^n)$

Note that $f_k''(x) = \frac{L}{4} A_k$ with

$$A_k = \begin{pmatrix} \bar{A}_k & 0_{k,n-k} \\ 0_{n-k,k} & 0_{n-k,n-k} \end{pmatrix},$$

where $0_{i,j}$ denotes an $i \times j$ zero matrix, $\bar{A}_k$ is a $k \times k$ matrix given by

$$\bar{A}_k = \begin{pmatrix} 2 & -1 & 0 & & & & \\ -1 & 2 & -1 & & 0 & & \\ 0 & -1 & 2 & & & & \\ & \cdots & & & \cdots & & \\ & & & & 2 & -1 & 0 \\ & 0 & & & -1 & 2 & -1 \\ & & & & 0 & -1 & 2 \end{pmatrix}$$

Lower complexity bounds for $\mathcal{F}_L^{\infty,1}(\mathbb{R}^n)$

The equation

$$f_k'(x) = A_k x - e_1 = 0$$

has the following solution, which is unique with respect to the first k components:

$$\bar{x}_k^{(i)} = \begin{cases} 1 - \frac{i}{k+1}, & i = 1, \ldots, k, \\ 0, & k+1 \le i \le n. \end{cases}$$

The optimal value of f_k is given by

$$f_k^* = f_k(\bar{x}_k) = \frac{L}{4}\left[\frac{1}{2}\bar{x}_k A_k \bar{x}_k - e_1^T \bar{x}_k\right] = -\frac{L}{8} e_1^T \bar{x}_k = \frac{L}{8}\left(-1 + \frac{1}{k+1}\right).$$

Lower complexity bounds for $\mathcal{F}_L^{\infty,1}(\mathbb{R}^n)$

$$\begin{aligned} \|\bar{x}_k\|^2 &= \sum_{i=1}^{n}\left(\bar{x}_k^{(i)}\right)^2 = \sum_{i=1}^{k}\left(1 - \frac{i}{k+1}\right)^2 \\ &= k - \frac{2}{k+1}\sum_{i=1}^{k} i + \frac{1}{(k+1)^2}\sum_{i=1}^{k} i^2 \\ &\le k - \frac{2}{k+1}\frac{k(k+1)}{2} + \frac{1}{(k+1)^2}\frac{(k+1)^3}{3} = \frac{1}{3}(k+1). \end{aligned}$$

Here we used the inequality

$$\sum_{i=1}^{k} i^2 = \frac{k(k+1)(2k+1)}{6} \le \frac{(k+1)^3}{3}.$$

Lower complexity bounds for $\mathcal{F}_L^{\infty,1}(\mathbb{R}^n)$

Denote by

$$\mathbb{R}^{k,n} = \{x \in \mathbb{R}^n : x^{(i)} = 0, k+1 \le i \le n\}.$$

Note that $f_k(x) = \frac{L}{8} x^T A_k x - \frac{L}{4} e_1^T x$ with

$$A_k = \begin{pmatrix} \bar{A}_k & 0_{k,n-k} \\ 0_{n-k,k} & 0_{n-k,n-k} \end{pmatrix},$$

therefore, for any $x \in \mathbb{R}^{k,n}$ we have

$$f_p(x) = f_k(x), \;\; p = k, \ldots, n.$$

Lower complexity bounds for $\mathcal{F}_L^{\infty,1}(\mathbb{R}^n)$

Lemma (2.1.3)

Let p, $1 \le p \le n$, be fixed and let $x_0 = 0$. Then for any sequence $\{x_k\}_{k=0}^p$ such that

$$x_k \in \mathcal{L}_k = \mathrm{Lin}\{f_p'(x_0), \dots, f_p'(x_{k-1})\},$$

we have $\mathcal{L}_k \subseteq \mathbb{R}^{k,n}$.

Proof: Note that $f_p'(x) = \frac{L}{4}A_p x - \frac{L}{4}e_1$. We use induction.

1. **The induction basis.** Since $x_0 = 0$, we have $f_p'(x_0) = -\frac{L}{4}e_1 \in \mathbb{R}^{1,n}$ and $\mathcal{L}_1 = \mathbb{R}^{1,n}$.
2. **The inductive step.** Let $\mathcal{L}_k \subseteq \mathbb{R}^{k,n}$ for some $k < p$. Since A_p is three-diagonal, for any $x \in \mathbb{R}^{k,n}$ we have $f_p'(x) = \frac{L}{4}A_p x - \frac{L}{4}e_1 \in \mathbb{R}^{k+1,n}$, so $\mathcal{L}_{k+1} \subseteq \mathbb{R}^{k+1,n}$. □

Corollary (2.1.1)

For any sequence $\{x_k\}_{k=0}^p$ such that $x_0 = 0$ and $x_k \in \mathcal{L}_k$ we have $f_p(x_k) \ge f_k^*$ (since $x_k \in \mathcal{L}_k \subseteq \mathbb{R}^{k,n}$, so $f_p(x_k) = f_k(x_k) \ge f_k^*$).

Lower complexity bounds for $\mathcal{F}_L^{\infty,1}(\mathbb{R}^n)$

Theorem (2.1.7)

For any k, $1 \le k \le \frac{1}{2}(n-1)$, and any $x_0 \in \mathbb{R}^n$ there exists a function $f \in \mathcal{F}_L^{\infty,1}(\mathbb{R}^n)$ such that for any first-order method $\mathcal{M}$ we have

$$f(x_k) - f^* \ge \frac{3L\|x_0 - x^*\|^2}{32(k+1)^2},$$

$$\|x_k - x^*\|^2 \ge \frac{1}{8}\|x_0 - x^*\|^2,$$

where x^* is a minimum of $f(x)$ and $f^* = f(x^*)$.

Proof: Recall that the method $\mathcal{M}$ generates a sequence of points such that

$$x_k \in x_0 + \mathrm{Lin}\{f'(x_0), \dots, f'(x_{k-1})\}, \; k \ge 1.$$

Without loss of generality, we can assume that $x_0 = 0$ (if it is not, we can consider the function $\bar{f}(x) = f(x + x_0)$).

Lower complexity bounds for $\mathcal{F}_L^{\infty,1}(\mathbb{R}^n)$

To prove the first inequality, let us fix k and apply $\mathcal{M}$ to minimizing $f(x) = f_{2k+1}(x)$. Then $x^* = \bar{x}_{2k+1}$ and $f^* = f_{2k+1}^*$. From corollary (2.1.1), we have

$$f(x_k) = f_{2k+1}(x_k) = f_k(x_k) \ge f_k^*.$$

Recall that

$$f_k^* = \frac{L}{8}\left(-1 + \frac{1}{k+1}\right)$$

and

$$\|\bar{x}_k\|^2 \le \frac{1}{3}(k+1).$$

Hence, since $x_0 = 0$ we have

$$\frac{f(x_k) - f^*}{\|x_0 - x^*\|^2} \ge \frac{\frac{L}{8}(-1 + \frac{1}{k+1} + 1 - \frac{1}{2k+2})}{\frac{1}{3}(2k+2)} = \frac{3}{8}L\frac{1}{4(k+1)^2}.$$

Lower complexity bounds for $\mathcal{F}_L^{\infty,1}(\mathbb{R}^n)$

Next, we derive the second inequality. Since $x_k \in \mathbb{R}^{k,n}$, we have

$$\begin{aligned}
\|x_k - x^*\|^2 &\ge \sum_{i=k+1}^{2k+1}\left(\bar{x}_{2k+1}^{(i)}\right)^2 = \sum_{i=k+1}^{2k+1}\left(1 - \tfrac{i}{2k+2}\right)^2 \\
&= k + 1 - \tfrac{1}{k+1}\sum_{i=k+1}^{2k+1} i + \tfrac{1}{4(k+1)^2}\sum_{i=k+1}^{2k+1} i^2 \\
(1) \quad &\ge k + 1 - \tfrac{1}{k+1}\tfrac{(3k+2)(k+1)}{2} + \tfrac{(2k+1)(7k+6)}{24(k+1)} \\
&= \tfrac{(2k+1)(7k+6)}{24(k+1)} - \tfrac{k}{2} = \tfrac{2k^2+7k+6}{24(k+1)} = \tfrac{2k^2+7k+6}{\frac{2}{3}24(k+1)^2}\left(\tfrac{1}{3}(2k+2)\right) \\
(2) \quad &\ge \tfrac{2k^2+7k+6}{16(k+1)^2}\|x_0 - \bar{x}_{2k+1}\|^2 \ge \tfrac{2(k+1)^2}{16(k+1)^2}\|x_0 - \bar{x}_{2k+1}\|^2 \\
&= \tfrac{1}{8}\|x_0 - x^*\|^2. \qquad \square
\end{aligned}$$

$$(1) \quad \sum_{i=k+1}^{2k+1} i^2 = \tfrac{1}{6}[(2k+1)(2k+2)(4k+3) - k(k+1)(2k+1)] = \tfrac{1}{6}(k+1)(2k+1)(7k+6);$$

$$(2) \quad \|\bar{x}_k\|^2 \le \tfrac{1}{3}(k+1).$$

Lower complexity bounds for $\mathcal{F}_L^{\infty,1}(\mathbb{R}^n)$

- The above theorem assumes that the number of steps of the numerical method is smaller than the space's dimension $(k \leq \frac{1}{2}(n-1))$.
- It describes the potential performance of numerical methods in the initial stage of the minimization process.
- The lower bound on the objective function value is not so bad (since the denominator is $O(k^2)$):

$$f(x_k) - f^* \geq \frac{3L\|x_0 - x^*\|^2}{32(k+1)^2}.$$

- However, the convergence to the optimal point can be extremely slow:

$$\|x_k - x^*\|^2 \geq \frac{1}{8}\|x_0 - x^*\|^2.$$

- Can we define "easier" classes of problems?