ISEN 629: Engineering Optimization

Lecture 16

Sergiy Butenko

Industrial and Systems Engineering
Texas A&M University

Fall 2007

Cutting plane methods

Consider the problem

$$\min\{f(x) : x \in Q\},$$

where f is convex on $\mathbb{R}^n$ and Q is a compact convex set such that

$$\mathrm{int}\, Q \neq \emptyset, \ \mathrm{diam}\, Q = D < \infty.$$

We assume that we are given a separating oracle, which for any test point $\bar{x} \in \mathbb{R}^n$ returns a vector g which is

- a subgradient of f at $\bar{x}$, if $x \in Q$,
- a separator of $\bar{x}$ from Q, if $x \notin Q$.

Example

For the problem with functional constraints $\min_{x \in Q} f(x)$, where $Q = \{x \in \mathbb{R}^n : \bar{f}(x) \leq 0\}$, for $x \notin Q$ the oracle provides us with any subgradient $\bar{g} \in \partial \bar{f}(x)$, which separates x from Q, since it is supporting to the level set $\mathcal{L}_{\bar{f}}(\bar{f}(x))$.

Cutting plane methods

Consider a sequence $X = \{x_i : i \geq 0\} \subset Q$. Then the localization sets generated by this sequence are given by

$$\begin{aligned} S_0(X) &= Q, \\ S_{k+1}(X) &= \{x \in S_k(X) : g(x_k)^T(x_k - x) \geq 0\}. \end{aligned}$$

Then for all $k \geq 0$ we have $x^* \in S_k$, and, as before, we use the following notations:

$$v_i = v_f(x^*, x_i), \quad v_k^* = \min_{0 \leq i \leq k} v_i.$$

We will denote by $\mathrm{vol}_n(S)$ the n-dimensional volume of set $S \subset \mathbb{R}^n$.

Cutting plane methods

Theorem (3.2.6)

For any $k \geq 0$ we have

$$v_k^* \leq D \left[\frac{\mathrm{vol}_n(S_k(X))}{\mathrm{vol}_n(Q)}\right]^{1/n}.$$

Proof.

1. Denote by $\alpha = v_k^*/D (\leq 1)$.
2. Since $Q \subseteq B_D(x^*)$, we have
$$(1-\alpha)x^* + \alpha Q \subseteq (1-\alpha)x^* + \alpha B_D(x^*) = B_{v_k^*}(x^*).$$
3. Since Q is convex, we have
$$(1-\alpha)x^* + \alpha Q \equiv [(1-\alpha)x^* + \alpha Q] \cap Q \subseteq B_{v_k^*}(x^*) \cap Q \subseteq S_k(X).$$
4. Therefore, $\mathrm{vol}_n(S_k(X)) \geq \mathrm{vol}_n((1-\alpha)x^* + \alpha Q) = \alpha^n \mathrm{vol}_n(Q)$.

□

General cutting plane scheme

0. Choose a bounded set $E_0 \supseteq Q$.
1. k-th iteration ($k \geq 0$):
   (a) Choose $y_k \in E_k$.
   (b) If $y_k \in Q$ then compute $f(y_k), g(y_k)$. If $y_k \notin Q$ then compute $\bar{g}(y_k)$ (separates y_k from Q).
   (c) Set
   $$g_k = \begin{cases} g(y_k), & \text{if } y_k \in Q, \\ \bar{g}(y_k), & \text{if } y_k \notin Q. \end{cases}$$
   (d) Choose $E_{k+1} \supseteq \{x \in E_k : g_k^T(y_k - x) \geq 0\}$.

Note that we use supersets of Q, since Q may have a complicated structure.

General cutting plane scheme

To analyze the performance of the above scheme, we introduce the following notations:

$$Y = \{y_k : k \geq 0\}, \; X = Y \cap Q, \; i(k) = |\{y_j : 0 \leq j < k\} \cap Q|.$$

Lemma (3.2.4)

For any $k \geq 0$ we have $S_{i(k)} \subseteq E_k$.

Corollary

1. *For any k such that $i(k) > 0$ we have*
$$v^*_{i(k)}(X) \leq D \left[\frac{\mathrm{vol}_n(S_{i(k)}(X))}{\mathrm{vol}_n(Q)}\right]^{1/n} \leq D \left[\frac{\mathrm{vol}_n(E_k)}{\mathrm{vol}_n(Q)}\right]^{1/n}.$$
2. *If $\mathrm{vol}_n(E_k) < \mathrm{vol}_n(Q)$ then $i(k) > 0$.*

Proof.

Observe that $Q = S_0 = S_{i(k)} \subseteq E_k$ for all k such that $i(k) = 0$. □

General cutting plane scheme

$$v^*_{i(k)}(X) \leq D \left[\frac{\mathrm{vol}_n(E_k)}{\mathrm{vol}_n(Q)}\right]^{1/n}.$$

- To have the convergence, it suffices to ensure that $\mathrm{vol}_n(E_k) \to 0$.
- The rate of decrease of the volume defines the method's rate of convergence.
- We want to decrease the volume as fast as possible.

Method of centers of gravity

Historically first nonsmooth minimization method based on cutting planes is the method of centers of gravity. Its idea is based on the following geometrical fact.

Lemma (3.2.5)

Let S be a bounded convex set in $\mathbb{R}^n$ with nonempty interior. Define the center of gravity of this set as

$$cg(S) = \frac{1}{\mathrm{vol}_n(S)} \int_S x dx.$$

For a direction $g \in \mathbb{R}^n$ define

$$S_+ = \{x \in S : g^T(cg(S) - x) \geq 0\}.$$

Then

$$\frac{\mathrm{vol}_n(S_+)}{\mathrm{vol}_n(S)} \leq 1 - \frac{1}{e}.$$

Method of centers of gravity

0. Set $S_0 = Q$.
1. k-th iteration ($k \geq 0$):
   (a) Choose $x_k = cg(S_k)$ and compute $f(x_k), g(x_k)$.
   (b) Set $S_{k+1} = \{x \in S_k : g(x_k)^T(x_k - x) \geq 0\}$.

Theorem (3.2.7)

If f is Lipschitz continuous on $B_D(x^*)$ with a constant M then for any $k \geq 0$ we have

$$f_k^* - f^* \leq MD\left(1 - \frac{1}{e}\right)^{k/n},$$

where $f_k^* = \min_{0 \leq j \leq k} f(x_j)$.

Ellipsoid method

- Another method using approximation of localization sets is the ellipsoid method.
- It is based on the geometrical observation that a half of an ellipsoid belongs to another ellipsoid with a strictly smaller volume.
- Let H be a positive definite symmetric $n \times n$ matrix. Consider the ellipsoid

$$E(H, \bar{x}) = \{x \in \mathbb{R}^n : (x - \bar{x})^T H^{-1}(x - \bar{x}) \leq 1\}.$$

- Choose a direction $g \in \mathbb{R}^n$ and consider a half of $E(H, \bar{x})$ defined by the corresponding hyperplane

$$E_+ = \{x \in E(H, \bar{x}) : g^T(\bar{x} - x) \geq 0\}.$$

Ellipsoid method

Lemma (3.2.6)

Denote by

$$\begin{aligned} \bar{x}_+ &= \bar{x} - \tfrac{1}{n+1} \cdot \tfrac{Hg}{(g^T Hg)^{1/2}}, \\ H_+ &= \tfrac{n^2}{n^2-1}\left(H - \tfrac{2}{n+1} \cdot \tfrac{Hgg^T H}{g^T Hg}\right). \end{aligned}$$

Then

$$E_+ \subset E(H_+, \bar{x}_+)$$

and

$$\mathrm{vol}_n(E(H_+, \bar{x}_+)) \leq \left(1 - \frac{1}{(n+1)^2}\right)^{n/2} \mathrm{vol}_n(E(H, \bar{x})).$$

Ellipsoid method

0. Choose $y_0 \in \mathbb{R}^n$ and $R > 0$ such that $B_R(y_0) \supseteq Q$. Set $H_0 = R^2 I_n$.
1. k-th iteration ($k \geq 0$):

$$\begin{aligned} g_k &= \begin{cases} g(y_k), & \text{if } y_k \in Q, \\ \bar{g}(y_k), & \text{if } y_k \notin Q. \end{cases} \\ y_{k+1} &= y_k - \tfrac{1}{n+1} \cdot \tfrac{H_k g_k}{(g_k^T H_k g_k)^{1/2}}, \\ H_{k+1} &= \tfrac{n^2}{n^2-1}\left(H_k - \tfrac{2}{n+1} \cdot \tfrac{H_k g_k g_k^T H_k}{g_k^T H_k g_k}\right). \end{aligned}$$

Note that this is just a special implementation of the general cutting plane scheme, where

$$E_k = \{x \in \mathbb{R}^n : (x - \bar{x})^T H^{-1}(x - \bar{x}) \leq 1\}.$$

and y_k is the center of this ellipsoid.

Ellipsoid method

As before, we use the following notations:

$$Y = \{y_k : k \geq 0\},\ X = Y \cap Q,\ i(k) = |\{y_j : 0 \leq j < k\} \cap Q|,$$

$$f_k^* = \min_{0 \leq j \leq k} f(x_j).$$

Theorem (3.2.8)

Let f be Lipschitz continuous on $B_R(x^*)$ with some constant M. Then for $i(k) > 0$ we have

$$f_{i(k)}^* - f^* \leq MR \left(1 - \frac{1}{(n+1)^2}\right)^{k/2} \cdot \left[\frac{\mathrm{vol}_n(B_R(x_0))}{\mathrm{vol}_n(Q)}\right]^{1/n}.$$

Ellipsoid method

To apply the above theorem, we need to ensure that $i(k) > 0$, i.e., $X \neq \emptyset$. Assume that there exists $\rho > 0$ and $\bar{x} \in Q$ such that

$$B_\rho(\bar{x}) \subseteq Q$$

[this can be guaranteed by assuming that all constraints are Lipschitz continuous and there is a feasible point at which all functional constraints are strictly negative (Slater condition)].
Then

$$\left[\frac{\mathrm{vol}_n(E_k)}{\mathrm{vol}_n(Q)}\right]^{1/n} \leq \left(1 - \frac{1}{(n+1)^2}\right)^{k/2} \left[\frac{\mathrm{vol}_n(B_R(x_0))}{\mathrm{vol}_n(Q)}\right]^{1/n} \leq \frac{1}{\rho} e^{-\frac{k}{2(n+1)^2}} R.$$

If $\frac{1}{\rho} e^{-\frac{k}{2(n+1)^2}} R \leq 1$, then $\mathrm{vol}_n(E_k) \leq \mathrm{vol}_n(Q) \ \Rightarrow \ i(k) > 0$ for all $k > 2(n+1)^2 \ln \frac{R}{\rho}$.
Moreover, if $i(k) > 0$ then

$$f_{i(k)}^* - f^* \leq \frac{1}{\rho} MR^2 \cdot e^{-\frac{k}{2(n+1)^2}}.$$

Ellipsoid method

- Each iteration of the ellipsoid method takes $O(n^2)$ arithmetic operations.
- In order to generate an ϵ-solution, this method needs

$$2(n+1)^2 \ln \frac{MR^2}{\rho\epsilon} = O\left(n^2 \ln \frac{1}{\epsilon}\right)$$

  calls of the oracle.
- Thus, we have a polynomial dependence on $\ln \frac{1}{\epsilon}$ and on logarithms of the class parameters M, R, ρ.
- For problem classes, whose oracle has a polynomial complexity, such algorithms are called (weakly) polynomial.