The Ellipsoid Method and Its Efficiency Analysis

Yinyu Ye
Department of Management Science and Engineering
Stanford University
Stanford, CA 94305, U.S.A.

http://www.stanford.edu/˜yyye

(LY, Chapter 5.1-5.3)

Elements of Complexity Theory

The term complexity refers to the amount of resources required by a computation. Computational complexity wishes to associate to an algorithm more intrinsic measures of its time requirements

- a notion of input size,
- a set of basic operations, and
- a cost for each basic operations

The last two allow one to associate a (total) cost of a computation.

Polynomial Time Algorithms

- Bit size (bit operations) for integers and Unit size (unit cost) for real numbers.
- The former is usually referred to as the Turing model of computation, and latter is referred as the real number arithmetic model.
- An algorithm is said to be a polynomial time algorithm if its worst-case cost of computation is bounded above by a polynomial function of the input size of the problem data.

Ellipsoid Method: the first polynomial-time algorithm for LP

The basic ideas of the ellipsoid method stem from research done in the nineteen sixties and seventies mainly in the Soviet Union (as it was then called) by others who preceded Khachiyan. The idea in a nutshell is to enclose the region of interest in each member of a sequence of ellipsoids whose size is decreasing, resembling the bisection method.

The significant contribution of Khachiyan was to demonstrate in two papers—published in 1979 and 1980—that under certain assumptions, the ellipsoid method constitutes a polynomially bounded algorithm for linear programming.

Ellipsoid Representation

Ellipsoids are just sets of the form

$$E = \{\mathbf{y} \in \mathbf{R}^m : (\mathbf{y} - \bar{\mathbf{y}})^T B^{-1} (\mathbf{y} - \bar{\mathbf{y}}) \le 1\}$$

where $\bar{\mathbf{y}} \in R^m$ is a given point (called the center) and B is a symmetric positive definite matrix of dimension m. We can use the notation $\text{ell}(\bar{\mathbf{y}}, B)$ to specify the ellipsoid E defined above. Note that

$$\text{vol}(E) = (\det B)^{1/2} \text{vol}(S(\mathbf{0}, 1)).$$

where $S(\mathbf{0}, 1)$ is the unit sphere in $\mathbf{R}^m$.

Half-Ellipsoid

By a Half-Ellipsoid of E, we mean the set

$$\frac{1}{2}E_a := \{\mathbf{y} \in E : \mathbf{a}^T\mathbf{y} \leq \mathbf{a}^T\bar{\mathbf{y}}\}$$

for a given non-zero vector $\mathbf{a} \in \mathbf{R}^m$ where $\bar{\mathbf{y}}$ is the center of E.

We are interested in finding a new ellipsoid containing $\frac{1}{2}E_a$ with the least volume.

- How small could it be?
- How easy could it be constructed?

The New Containing Ellipsoid

The new ellipsoid $E^+ = \text{ell}(\bar{\mathbf{y}}^+, B^+)$ can be constructed as follows. Define

$$\tau := \frac{1}{m+1}, \qquad \delta := \frac{m^2}{m^2-1}, \qquad \sigma := 2\tau.$$

And let

$$\bar{\mathbf{y}}^+ := \bar{\mathbf{y}} - \frac{\tau}{(\mathbf{a}^T B \mathbf{a})^{1/2}} B\mathbf{a},$$

$$B^+ := \delta \left(B - \sigma \frac{B\mathbf{a}\mathbf{a}^T B}{\mathbf{a}^T B \mathbf{a}} \right).$$

Theorem 1 *Ellipsoid $E^+ = ell(\bar{\mathbf{y}}^+, B^+)$ defined as above is the ellipsoid of least volume containing $\frac{1}{2}E_a$. Moreover,*

$$\begin{aligned}
\frac{vol(E^+)}{vol(E)} &= \left(\frac{m^2}{m^2-1}\right)^{\frac{m-1}{2}} \cdot \frac{m}{m+1} = \left(1+\frac{1}{m^2-1}\right)^{\frac{m-1}{2}} \cdot \left(1-\frac{1}{m+1}\right) \\
&< \exp\left(\frac{1}{m^2-1}\cdot\frac{m-1}{2}\right)\exp\left(-\frac{1}{m+1}\right) \\
&\le \exp\left(-\frac{1}{2(m+1)}\right) \\
&< 1.
\end{aligned}$$

Here we used

$$1+a<e^a \quad \textit{and} \quad 1-a<e^{-a}$$

for $a>0$.

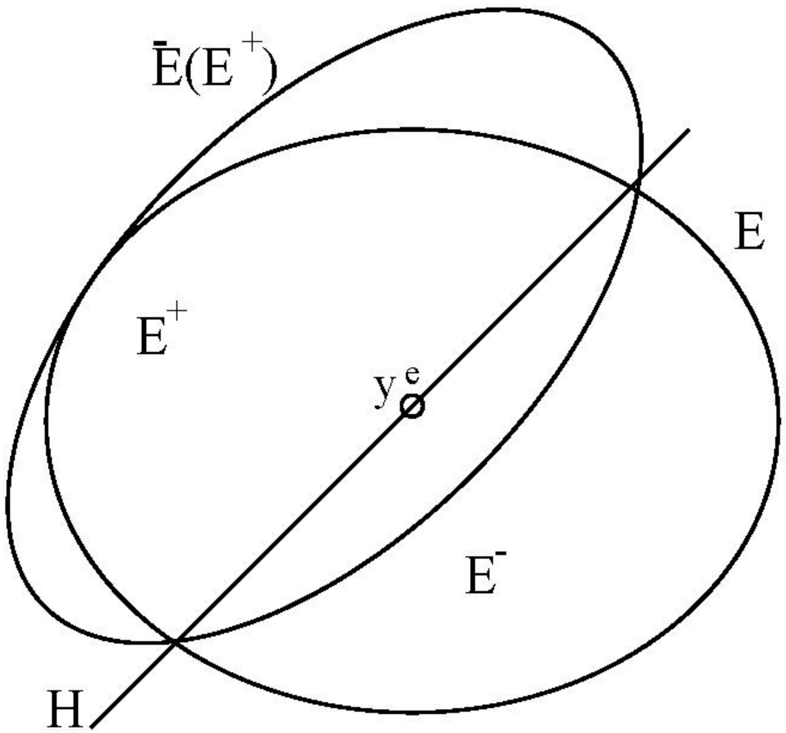

Figure 1: The least volume ellipsoid containing a half ellipsoid

Affine Transformation

Assume that $E = \text{ell}(\bar{\mathbf{y}}, B)$, where the positive definite matrix B has the factorization $B = JJ^{\mathrm{T}}$. Now consider the affine transformation $\mathbf{y} \mapsto \bar{\mathbf{y}} + J\mathbf{z}$.

Let $\mathbf{y} \in E$. Then $\mathbf{y} - \bar{\mathbf{y}} = J\mathbf{z}$ for some vector $\mathbf{z} \in R^m$. Now since $\mathbf{y} \in E$,

$$\begin{aligned} 1 &\geq (\mathbf{y} - \bar{\mathbf{y}})^T B^{-1} (\mathbf{y} - \bar{\mathbf{y}}) \\ &= (J\mathbf{z})^T (JJ^{\mathrm{T}})^{-1} (J\mathbf{z}) \\ &= \mathbf{z}^T J^{\mathrm{T}} (J^{\mathrm{T}})^{-1} J^{-1} J\mathbf{z} \\ &= \mathbf{z}^T \mathbf{z} \end{aligned}$$

so $\mathbf{z} \in S(0, 1)$, the unit sphere. Conversely, every such point maps to an element of E.

Linear Feasibility Problem

The ellipsoid method discussed here is really aimed at finding an element of a polyhedral set Y given by a system of linear inequalities.

$$Y = \{\mathbf{y} \in \mathbf{R}^m : \mathbf{a}_j^T \mathbf{y} \le c_j, \quad j = 1, \ldots n\}$$

Finding an element of Y can be thought of as being equivalent to solving a LP problem, though this requires a bit of discussion.

Two Important Assumptions

(A1) There is a vector $\mathbf{y}^0 \in R^m$ and a scalar $R > 0$ such that the closed ball $S(\mathbf{y}^0, R)$ with center $\mathbf{y}^0$ and radius R

$$S(\mathbf{y}^0, R) := \{\mathbf{y} \in R^m : \|\mathbf{y} - \mathbf{y}^0\|_2 \le R\}$$

contains Y.

(A2) There is a known scalar $r > 0$ such that if Y is nonempty, then it contains a ball of the form $S(\mathbf{y}^*, r)$ with center at $\mathbf{y}^*$ and radius r.

Note that this assumption implies that if Y is nonempty then it has a nonempty interior.

Cutting Plane

At each iteration of the algorithm, we will have $Y \subset E_k$. It is then possible to check whether $\mathbf{y}^k \in Y$. If so, we have found an element of Y as required. If not, there is at least one constraint that is violated. Suppose $\mathbf{a}_j^T \mathbf{y}^k > c_j$. Then

$$Y \subset \tfrac{1}{2}E_k := \{\mathbf{y} \in E_k : \mathbf{a}_j^T \mathbf{y} \le \mathbf{a}_j^T \mathbf{y}^k\}$$

This set is a "half ellipsoid" of E_k cut through its center.

The Ellipsoid Algorithm

Input: $A \in R^{m \times n},\ \mathbf{c} \in R^n,\ \mathbf{y}^0 \in R^m$ such that Y (as defined on Slides 3-4) satisfies (A1) and (A2).

Output: $\mathbf{y} \in Y$.

Initialization: Set $B_0 = \frac{1}{R^2} I,\ K = \lceil 2m(m+1)\log(R/r) \rceil + 1.$

For $k = 0, 1, \ldots, K-1$ do

Iteration k: If $\mathbf{y}^k \in Y$, STOP: result is $\mathbf{y} = \mathbf{y}^k$. Otherwise, choose j with $\mathbf{a}_j^T \mathbf{y}^k > c_j$ and form the half ellipsoid; and update $\mathbf{y}^k$ and B_k as described earlier.

Performance of the Ellipsoid Method

Under the assumptions stated above, the ellipsoid method solves linear programs in a polynomially bounded number of iterations bounded by $O\left(m^2 \log\left(\frac{R}{r}\right)\right)$ and each iteration uses $O(m^2)$ arithmetic operations.

Computational experience shows that the number of iterations required to solve a LP problem is very close to the theoretical upper bound. This means that the method is inefficient in a practical sense.

In contrast to this, although the simplex method is known to exhibit exponential behavior on specially constructed problems such as those of Klee and Minty, it normally requires a number of iterations that is a small multiple of the number of linear equations in the standard form of the problem.

Linear Programming (LP)

$$
\text{(P)} \quad \begin{array}{ll} \max & \mathbf{c}^T\mathbf{x} \\ \text{subject to} & A\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \end{array}
$$

$$
\text{(D)} \quad \begin{array}{ll} \min & \mathbf{b}^T\mathbf{y} \\ \text{subject to} & A^T\mathbf{y} \geq \mathbf{c} \\ & \mathbf{y} \geq \mathbf{0} \end{array}
$$

By the Weak Duality Lemma[a], we have

$$
\mathbf{c}^T\mathbf{x} \leq \mathbf{b}^T\mathbf{y}.
$$

[a] LY Chapter 4.2

Integer Data

Next, we assume that the data for the problem are all integers. As a measure of the size of the problem above we let $c_j = a_{0j}$ and define

$$L = \sum_{i=0}^{m} \sum_{j=1}^{n} \lceil \log_2(\mod a_{ij} + 1) + 1 \rceil.$$

In our discussion above, we made two assumptions about Y. One of the assumptions, (A2), effectively says that if Y is nonempty, then it possesses a nonempty interior. The linear inequalities are relaxed to

$$\mathbf{a}_j^T \mathbf{y} < c_j + 2^{-L} \quad j = 1, \ldots, n. \tag{1}$$

It was shown by Gács and Lovasz (1981) that if the inequality system (1) has a solution, then so does

$$\mathbf{a}_j^T \mathbf{y} \le c_j, \quad j = 1, \ldots n.$$

Bounds from L

Therefore, we can bound

$$r \geq 2^{-L}.$$

On the other hand, we can bound

$$R \leq O(2^L).$$

Thus,

$$\log(R/r) \leq O(L),$$

which is linear (polynomial) in L.

The Sliding Objective Hyperplane Method

Consider

$$
\begin{array}{lll}
 & \min & \mathbf{b}^T\mathbf{y} \\
(\mathrm{D}) & \text{subject to} & A^T\mathbf{y} \geq \mathbf{c} \\
 & & \mathbf{y} \geq \mathbf{0}
\end{array}
$$

At the center $\mathbf{y}^k$ of the ellipsoid, if a constraint is violated then add the corresponding constraint hyperplane as the cut; otherwise, add objective hyperplane

$$\mathbf{b}^T\mathbf{y} \geq \mathbf{b}^T\mathbf{y}^k$$

as the cut.

Desired Theoretical Properties

- Separation Problem: either decide $\mathbf{x} \in P$ or find a vector $\mathbf{d}$ such that $\mathbf{d}^T\mathbf{x} \leq \mathbf{d}^T\mathbf{y}$ for all $\mathbf{y} \in P$.
- Oracle to generate $\mathbf{d}$ without enumerating all hyperplanes.

Theorem 2 *If the separating (oracle) problem can be solved in polynomial time of m and $\log(R/r)$, then we can solve the standard linear programming problem whose running time is polynomial in m and $\log(R/r)$ that is independent of n, the number of inequality constraints.*

LP with an Exponentially Large Number of Inequalities: TSP

Travelling Salesman Problem (TSP): given an undirected graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ where $\mathcal{N}$ is the set of n nodes and length c_e for every edge $e \in \mathcal{E}$, the goal is to find a tour (a cycle that visits all nodes) of minimal length.

To model the problem, we define for every edge e a variable x_e, which is 1 if e is in the tour and 0 otherwise. Let $\delta(i)$ be the set of edges incident to node i, then

$$\sum_{e \in \delta(i)} x_e = 2, \; \forall i \in \mathcal{N}.$$

Let $S \subset \mathcal{N}$ and

$$\delta(S) = \{e : \; e = (i,j), \; i \in S, \; j \notin S\}.$$

Then,

$$\sum_{e \in \delta(S)} x_e \geq 2, \quad \forall S \subset \mathcal{N}, \; S \neq \emptyset, \mathcal{N}.$$

LP Relaxatrion of TSP

$$
(\text{TSP}) \quad \begin{array}{ll}
\min & \sum_{e \in \mathcal{E}} c_e x_e \\
\text{subject to} & \sum_{e \in \delta(i)} x_e = 2, \ \forall i \in \mathcal{N}, \\
& \sum_{e \in \delta(S)} x_e \geq 2, \ \forall S \subset \mathcal{N}, \ S \neq \emptyset, \mathcal{N}, \\
& 0 \leq x_e \leq 1 \ \forall e \in \mathcal{E}.
\end{array}
$$

This problem has an exponential number of inequalities since there are $2^n - 2$ of proper subsets of S

Oracle to Check the Separation

Given x_e^*, we would like to check if

$$\sum_{e \in \delta(S)} x_e^* \geq 2, \ \forall S \subset \mathcal{N}, \ S \neq \emptyset, \mathcal{N}.$$

Assign x_e^* as the capacity for every edge $e \in \mathcal{E}$, then the problem is to check if the min-cut of the graph is greater than or equal to 2.

This problem can be formulated as Maximum Flow problems (how?) and can be solved as a small LP.

Convex Feasibility Problem

The ellipsoid method can be used to find an element of a convex set Y given by a system of convex inequalities.

$$Y = \{\mathbf{y} \in \mathbf{R}^m : f_j(\mathbf{y}) \le 0, \quad j = 1, \ldots n\}$$

where each $f_\cdot(\mathbf{y})$ is a continuous convex function.

Finding an element of Y can be thought of as being equivalent to solving a convex optimization problem when its sub-gradient vector is computable.

How to generate a separation hyperplane?

The tangent or gradient hyperplane of the convex function of the violated inequality.