# Contents



# 1 The Ellipsoid Method

Khachian showed in 1979 that linear programming can be solved in polynomial time. We consider LPs in inequality form:

$$\max\left\{ c^T x; Ax \leq b \right\} \tag{LP}$$

The entries of $A$, $c$, and $b$ are assumed to be integral. The ellipsoid method is about testing feasibility. Optimization is done by binary search on the objective function value, i.e., we test feasibility of the following LP

$$c^T x \geq c_0 \quad \text{and} \quad Ax \leq b \tag{1}$$

for suitably chosen constants $c_0$. More precisely, we compute an upper and a lower bound on the optimum objective value and then perform binary search on this interval.

The ellipsoid method is of limited practical value. Although its running time is polynomial, it is usually outperformed by the Simplex algorithm and also by interior point methods (another polynomial time algorithm for linear programming). The ellipsoid method is of great theoretical value as is can also be applied to implicitly defined LPs. In such LPs, the constraints are not given explicitly, but there is an algorithm (usually called, the "separation oracle") that tells whether a point $z$ satisfies all inequalities or not. In the latter case, it also returns a violated inequality.

This section is organized as follows. In Section 1.1 we introduce the Ellipsoid method and in Section 1.3 we discuss an application to an LP with exponentially many constraints. Sections 1.4 and 1.5 build intuition; we discuss the Ellipsoid method in one and two dimensions. In the one-dimensional case, the method is tantamount to binary search. In two-dimensions, the ellipsoid comes into play.

## 1.1 The Principle

We discuss how to decide feasibility of a system

$$Ax \leq b. \tag{LP}$$

The entries of $A$ and $b$ are assumed to be integral. Let $C$ be the maximum absolute value of any entry of $A$ and $b$ and let $L = n(1 + \log n + \log C)$. Then $2^L = (2nC)^n$.

**Theorem 1** *The Ellipsoid method solves linear programs in time polynomial in L.*

**Algorithm 1** The Ellipsoid Method

initialize $E$ to the ball with radius $4^{nL}$ centered at the origin. $\{S \subseteq E\}$
**while** $vol(E) \geq 2^{-(n+1)L}$ **do**
  if the center $z$ of $E$ is feasible, stop and declare the problem feasible.
  select an inequality $a_i x \leq b_i$ violated by $z$, i.e., $a_i^T z > b_i$.
  $\{$for every point $x \in S$, we have $a_i^T x \leq b_i \leq a_i^T z.\}$
  consider the "half-ellipsoid" $(1/2)E = E \cap \{x;\, a_i^T x \leq a_i^T z\}$ which is the intersection of $E$ with a half-space whose boundary passes through $z$ and replace $E$ by the smallest (in volume) ellipsoid containing $(1/2)E$. $\{S \subseteq E\}$
**end while**
stop and declare the problem infeasible.

We show how to decide whether the feasible set is empty and how to find a feasible point if there is one under the following additional assumption: *the set of feasible points inside the ball of radius* $4^{nL}$ *centered at the origin, has volume at least* $2^{-(n+1)L}$. We will justify this assumption later.

In the sequel we use $S$ to denote the set of feasible solutions inside the ball of radius $4^{nL}$ centered at the origin. Then $S$ is either empty or has volume at least $2^{-(n+1)L}$.

The Ellipsoid method is a generalization of binary search. In binary search we maintain an interval that contains the solution. In each iteration, we test whether the midpoint of the interval is a solution. If not, we proceed with one of two half-intervals. The Ellipsoid method generalizes this strategy to arbitrary dimensions. The proper generalization of intervals are ellipsoids.

An ellipsoid $E$ in $\mathbb{R}^n$ consists of all points $x \in \mathbb{R}^n$ satisfying an inequality of the form

$$(x-z)^T Q (x-z) \leq 1$$

where $z \in \mathbb{R}^n$ is the center of the ellipsoid and $Q$ is any $n \times n$ positive definite matrix[1]. For $Q = I$, we have the unit ball centered at $z$, for $Q = diag(a_1, \ldots, a_n)$ with $a_i > 0$ for all $i$, we have an ellipsoid with center $z$ and axes of length $1/\sqrt{a_i}$. An ellipsoid can also be viewed as the image of the unit ball under the affine transformation $x \mapsto R(x-z)$. A one-dimensional ellipsoid is an interval.

The ellipsoid method maintains an ellipsoid $E$ that is known to contain $S$; $E$ is initialized to the ball with radius $4^{nL}$ centered at the origin. In each iteration, we first check the volume of $E$. If the volume of $E$ is smaller than $2^{-(n+1)L}$, $S$ is empty and we stop. Otherwise, we consider the center $z$ of $E$. If $z$ is feasible, we have found a feasible point and stop. If $z$ is infeasible, it violates at least one of the inequalities defining $S$, say $a_i^T z > b_i$. Since the points in $S$ satisfy this inequality, we have $a_i^T x \leq b_i < a_i z$ for all $x \in S$. Thus

$$S \subseteq (1/2)E := E \cap \{x;\, a_i^T x \leq a_i^T z\}.$$

[1] A matrix $Q$ is positive definite if for any nonzero vector $y$, one has $y^T Q y > 0$. Positive definite matrices have "roots", i.e., there is a matrix $R$ such that $Q = R^T R$. In fact, a matrix is positive definite if there is a non-singular matrix $R$ such that $Q = R^T R$.

$(1/2)E$ is the intersection of $E$ with a half-space whose boundary contains $z$, see Figure 1. We replace $E$ by the smallest ellipsoid containing $(1/2)E$.

What is the smallest ellipsoid, call it $E'$, containing $(1/2)E$? Let us consider a special case: $E$ is the unit ball and the half-space is $x_1 \le 0$, see Figure 2.

**Lemma 1** *Let E be the unit ball. Then*

$$E \cap \{x; x_1 \le 0\} \subseteq E' := \left\{ x;\ \left(\frac{n+1}{n}\right)^2 \left(x_1 + \frac{1}{n+1}\right)^2 + \frac{n^2-1}{n^2} \sum_{2 \le i \le n} x_i^2 \le 1 \right\}.$$

*Moreover,*

$$\frac{vol(E')}{vol(e)} \le e^{-1/(2n+2)}.$$

*Observe that $E'$ has its center at $(1/(n+1), 0, 0, \ldots, 0)$. It passes through the point $(-1, 0, \ldots, 0)$ and the points $(0, x_2, \ldots, x_n)$ with $\sum_{2 \le i \le n} x_i^2 = 1$.*

**Proof:** Consider any $x$ with $\sum_{1 \le i \le n} x_i^2 \le 1$ and $x_1 \le 0$. Then

$$\begin{aligned}
&\left(\frac{n+1}{n}\right)^2 \left(x_1 + \frac{1}{n+1}\right)^2 + \frac{n^2-1}{n^2} \sum_{2 \le i \le n} x_i^2 \\
&\qquad = \left(\frac{n+1}{n}\right)^2 \left(x_1^2 + \frac{2}{n+1} x_1 + \frac{1}{(n+1)^2}\right) + \frac{n^2-1}{n^2} \sum_{2 \le i \le n} x_i^2 \\
&\qquad = \frac{1}{n^2} \left( (2n+2)x_1^2 + (2n+2)x_1 + 1 + (n^2-1) \sum_{1 \le i \le n} x_i^2 \right) \\
&\qquad \le \frac{1}{n^2} \left( (2n+2)x_1(x_1+1) + 1 + (n^2-1) \right) \\
&\qquad \le 1,
\end{aligned}$$

where the next to last inequality follows from $\sum_i x_i^2 \le 1$ and the last inequality follows from $x_1^2 + x_1 \le 0$ for $-1 \le x_1 \le 0$. We have now shown that $E \cap \{x; x_1 \le 0\} \subseteq E'$. We will not show that $E'$ is the smallest ellipsoid with this property. In Section 1.5 you can find the argument in the two-dimensional situation.

We next turn to the volume bound. The volume of a ellipsoid is proportional to the product of the lengths of its axes. All axes of $E$ have length one and $E'$ has one axis of length $\frac{n}{n+1}$ and

$n-1$ axes of length $\sqrt{\frac{n^2}{n^2-1}}$. Thus

$$\begin{aligned}
\frac{vol(E')}{vol(E)} &= \frac{n}{n+1}\left(\frac{n^2}{n^2-1}\right)^{(n-1)/2} \\
&= \exp\left(\ln(1-\frac{1}{n+1}) + \frac{n-1}{2}\cdot\ln(1+\frac{1}{n^2-1})\right) \\
&\leq \exp\left(-\frac{1}{n+1} + \frac{n-1}{2}\frac{1}{n^2-1}\right) \\
&= \exp\left(-\frac{1}{n+1} + \frac{1}{2(n+1)}\right) \\
&= e^{-1/(2n+2)},
\end{aligned}$$

where the inequality follows from $\ln(1+s) \leq s$ for $-1 < s$. ∎

How can we deal with the case of an arbitrary ellipse $E$? We can derive the equations for $E'$ as follows:

1. translate space so that $z$ is moved into the origin.
2. rotate space such that the axes of $E$ become aligned with the coordinate axes.
3. scale the coordinates such that $E$ becomes the unit ball
4. at this point the boundary of $H$ is an arbitrary hyperplane through the origin, rotate space again so as to turn $H$ into the half-space $x_1 \leq 0$.
5. at this point we are in our special situation and hence know $E'$.
6. apply steps 1 to 5 in reverse order to $E'$.

A nice fact about the transformations used in steps 1 to 5 is that, although they change volumes, they do not change the ratio of volumes and hence the bound on the ratio of the volume of $E$ and $E'$ derived above stays true.

**Lemma 2** *In every iteration, the volume of E shrinks by at least the factor*

$$f = \exp(-\frac{1}{2n+2}) < 1 .$$

The polynomiality of the method follows easily. We start with an ellipsoid whose volume is bounded by $8^{n^2L}$. Observe that the ball with radius $4^{nL}$ fits into the box with side length $2\cdot 4^{nL}$ and hence its volume is bounded by $(8^{nL})^n = 8^{n^2L}$.

In each iteration the volume shrinks by a factor $f$ and hence the volume of the ellipsoid after $k$ iterations is at most $f^k \cdot 8^{n^2L}$. We stop, if the volume goes below $2^{-(n+1)L}$. Thus if we enter iteration $k+1$, we must have $f^k 8^{n^2L} \geq 2^{-(n+1)L}$ or $k \log f + 3n^2L \geq -(n+1)L$ or

$$k \leq \frac{-(n+1)L - 3n^2L}{\log f} = O(n^3L).$$

One remark is needed at this point. The formulae for updating the ellipsoid involve additions, multiplications, divisions, and roots. We have to carry them out with finite precision. It can be shown the the claims above stay essentially true, if all calculations are carried out with $8^L$ bits of precision. This proves Theorem 1 (modulo the unproven assumptions). We turn to the assumptions.

## 1.2 Details

We sketch how to guarantee the minimum volume assumption and how to reduce optimization to feasibility testing.

**Minimum Volume:** The solution to $Ax \leq b$ might be a single point. In the Ellipsoid method one argues about the volume of the set of solutions and hence we want the situation that either there is no solution or that the solution set has a certain minimum volume. This is easy to achieve by perturbation. We need a Lemma about solutions to linear systems.

**Lemma 3** *Consider a system $Ax = b$ with non-singular $A$ and dimension $n$. Assume that all entries are integral bounded by $C$ in absolute value. Then the entries of $A^{-1}b$ are rational numbers whose numerator and denominator are bounded by $n^nC^n$.*

**Proof:** By Cramer's rule, the $j$-th coordinate of $x$ is (up to sign) equal to

$$\frac{\det A'}{\det A},$$

where $A'$ is obtained from $A$ by replacing the $j$-th column by $b$. The value of the determinant is a sum of $n!$ terms, each bounded by $C^n$. Finally, $n! \leq n^n$. ∎

**Lemma 4 (Minimum Volume and Localization)** *Let $\varepsilon = 1/(2n(nC)^n)$.*

- *$Ax \leq b$ is feasible if and only if $Ax \leq b + \varepsilon 1$ is feasible ($1$ is the vector of ones).*
- *If the latter problem is feasible, the set of solutions inside the ball with radius $4^{nL}$ centered at the origin, has volume at least $2^{-(n+1)L}$.*

**Proof:** I want to give a feeling why this might be true. Clearly, if $Ax \leq b$ is feasible, then the perturbed system is feasible.

If $Ax \leq b$ is not feasible, then there is a $y \geq 0$ such that $y^T A = 0$ and $y^T b = -1$ (this is Farkas' Lemma). In other words, the LP

$$\text{minimize } 1^T y \text{ subject to } y \geq 0,\ y^T A = 0,\ y^T b = -1$$

is feasible. Since basic solutions of LPs are solutions to linear system, Lemma 3 gives us a bound on the coordinates of a solution $y$; their absolute value is bounded by $n^n C^n$.

Next consider the perturbed system and observe that

$$y^T(b + \varepsilon 1) \leq -1 + n(nC)^n \varepsilon \leq -1/2.$$

Thus $y$ proves the infeasibility of the perturbed system. Assume to the contrary that there is an $x$ with $Ax \leq b + \varepsilon 1$. Multiplying this inequality by $y^T$ from the left yields $0 \leq y^T(b + \varepsilon 1) \leq -1/2$, a contradiction.

For the second part, assume feasibility of the perturbed system. Then the original system is feasible and has a solution $x$ well inside the ball of radius $4^{nL}$ (there is a feasible point whose coordinates are bounded by $(nC)^n$ and $(nC)^n \ll 4^{nL}$). Consider any $x'$ whose coordinates differ from the coordinates of $x$ by at most $\delta$. Then

$$Ax' = Ax + A(x' - x) \leq b + \delta nC1 \leq b + \varepsilon 1$$

where $\delta$ is such that $nC\delta = \varepsilon$. Thus the feasible region of the perturbed system contains a cube of side length $2\delta$ and the volume bound follows (since $(2\delta)^n \geq 2^{-(n+1)L}$. ∎

**Feasibility and Optimization:** We argue that if one can decide feasibility of LPs in polynomial time then one can compute optimal solutions in polynomial time. We first show how to deal with unbounded LPs and then how to deal with bounded LPs.

An LP is unbounded if and only if it and the "companion LP" $\max\left\{0;\, Ax \leq 0, c^T x \geq 1\right\}$ are feasible. Thus we can decide unboundedness if we can decide feasibility. Observe first that if both problems are feasible, say with feasible solutions $x_0^*$ and $x_1^*$, then $x_0^* + tx_1^*$ is feasible for the original problem for any $t \geq 0$. The objective value grows without bounds. Conversely, if the LP is unbounded, the simplex algorithm yields $x_0^*$ and $x_1^*$ such that $x_0^* + tx_1^*$ is feasible for any $t \geq 0$ and the objective functions grows as a function of $t$. For $t = 0$, we conclude that $x_0^*$ is feasible for LP. The fact that $x_0^* + tx_1^*$ is feasible for every positive $t$ implies $Ax_1^* \leq 0$ and the fact that the objective function grows without bounds implies $Ax_1^* > 0$.

For a bounded LP, Lemma 3 tells us that the coordinates of the optimal solution are bounded by $(nC)^n$. Thus the objective value is bounded in absolute value by $M := nC(nC)^n$. We use binary search on the interval $[-M, +M]$ to determine the optimal objective value. Any search step is a feasibility test of an LP of the form

$$c^T x \leq c_0 \quad Ax \leq b.$$

When can we stop the search. Let $x$ and $x'$ be two distinct vertices of the feasible set and assume $c^T x \neq c^T x'$. Then

$$c^T(x - x') \geq \frac{1}{(nC)^{2n}}$$

since $x - x'$ is a vector of rationals with denominator at least $(nC)^{2n}$ and the entries of $c$ are integral.

Thus we can stop the search after

$$\log(2M \cdot (nC)^{2n}) = O(L^{O(1)}$$

iterations.

**Summary:** The ellipsoid method is not a practical method. However, its theoretical interest is immense.

- First, it shows that linear programming is a polynomial time problem.
- Second, it shows that it suffices to have a "separation oracle" for solving LPs. A separation oracle takes a point $z$ and tells whether $z$ is feasible. If $z$ is infeasible it also provides a violated inequality. Observe that a separation oracle is all that is needed in step 4.

## 1.3 An Application: The Subtour Elimination LP

Let $G = (V,E)$ be an undirected graph and let $c : E \to \mathbb{R}_{\geq 0}$ be a non-negative weight function on the edges of $G$. The following integer linear program solves the Traveling Salesman Problem on $G$. We have a variable $x_e$ associated with each edge $e$ of $G$. The variables are constrained to have values 0 or 1. The intended meaning is that the edges $e$ with $x_e = 1$ form the Traveling Salesman Tour. Every tour must use exactly two of the edges incident to any vertex, i.e.,

$$\sum_{e \in \delta(v)} x_e = 2 \quad \text{for each } v \in V \tag{2}$$

where $\delta(v)$ is the set of edges incident to $v$. For a set $S$ of vertices, let $\delta(S)$ be the set of edges having exactly one endpoint in $S$. Then

$$\sum_{e \in \delta(S)} x_e \geq 2 \quad \text{for each } S \subseteq V \text{ with } \emptyset \neq S \neq V. \tag{3}$$

**Lemma 5** *The zero-one solutions of (2) and (3) are precisely the Traveling Salesman tours.*

**Proof:** A tour uses exactly two edges incident to every vertex. Moreover, for every set $S$ of vertices with $\emptyset \neq S \neq V$ at least two edges of the tour have exactly one endpoint in $S$.

Conversely, consider a zero-one solution and let $X$ be the set of edges with $x_e = 1$. Since $X$ contains exactly two edges incident to every vertex, the edges in $X$ form a set of disjoint cycles.

Assume that there is more than one cycle. Let $S$ be the vertex set of one of the cycles. Then $\sum_{e \in \delta(S)} x_e = 0$, a contradiction to (3). ∎

The constraints in (2) are called *degree constraints* and the constraints in (3) are called *subtour-elimination constraints*. Of course, the goal is to minimize the total cost of the edges in the tour, i.e.,

$$\text{minimize} \quad \sum_e c_e x_e.$$

We next relax the condition $x_e \in \{0,1\}$ to $0 \le x_e \le 1$. We obtain an LP. There are $2n$ degree constraints and $2^n - 2$ subtour elimination constraints. We show how to solve this LP with the Ellipsoid method in polynomial time.

Assume that the edge costs $c_e$ are integers in the range from 0 to $C$. Then the objective value of the LP lies between 0 and $nL$; the objective value of the LP is not necessarily integral. We want to find the smallest integer $c_0$ such that the LP

$$\sum_e c_e x_e \le c_0 \text{ and } 0 \le x_e \le 1 \text{ and } (2) \text{ and } (3)$$

is feasible. Then $c_0$ is a lower bound for the ILP. We use the Ellipsoid method.

We can start with a ball of radius $m$, since we have $m$ variables bounded by one.[2] Let $z$ be the center of the current ellipsoid. We check the constraints

$$\sum_e c_e x_e \le c_0 \text{ and } 0 \le x_e \le 1 \text{ and } (2)$$

by substituting $z$ for $x$. We check the subtour-elimination constraints algorithmically.

We set up an auxiliary graph $G'$. $G'$ is isomorphic to $G$; the weight of edge $e$ is equal to $z_e$ (= the entry of the vector $z$ indexed by $e$). We compute a minimum edge cut[3] in $G'$. If the value of this cut is less than two, it gives us a violated subtour-elimination constraint. If the value of this cut is equal to two (why can it be no larger than two?), there is no violated subtour-elimination constraint.

The result of this section is remarkable. We have an LP with $m$ variables, one for each edge. Hence the optimal solution is defined by $m$ of the constraints. Which constraints are relevant is determined by the edge weights. There are exponentially many subtour-elimination constraints. The ellipsoid method finds the relevant constraints in polynomial time without inspecting all of the constraints.

## 1.4 Binary Search

We start with a one-dimensional problem. Consider the following situation.

---

[2] Argue that one can start with a ball of radius $n$.

[3] The pedestrian way of solving a min-cut problem is to iterate over all pairs $(a,b)$ of distinct vertices of $G'$. For each pair one computes the minimum $(a,b)$-cut by a max-flow computation with source $a$ and sink $b$. There are more efficient algorithms known [SW97, KS96]. For an implementation see [MN99].

- The goal is to find an $x \in \mathbb{R}$ having a certain property $P$ or to tell that no $x \in R$ has property $P$.

  In the application to linear programming: $x$ has property $P$ if it satisfies (1).

- We know that the set $S$ of $x$ having property $P$ is either empty or an interval of length at least $\ell$.

  In the application to linear programming: the set of feasible $x$ is either empty or is a convex set of volume at least $\ell$. A convex set in one dimension is an interval.

- We know that the absolute value of any $x$ with property $P$ is bounded by $U$.

  In the application to linear programming: the feasible set is contained in a ball of radius $U$.

- Given an $x \in \mathbb{R}$ we can test whether $x$ has property $P$. If $x$ does not have property $P$, either $S \subseteq [-\infty \mathinner{..} x]$ or $S \subseteq [x \mathinner{..} +\infty]$. This follows from the fact that $S$ is an interval. We can tell which of the two cases applies[4].

  In the application to linear programming: either $x$ satisfies all inequalities or we can find an inequality violated by $x$.

The problem just described can be solved by binary search.

1. Initialize an interval $I$ to $[-U \mathinner{..} U]$. Then $S \subseteq I$. In the course of the algorithm we will shrink $I$ and maintain the invariant $S \subseteq I$.

2. If the length of $I$ is less than $\ell$, we stop and declare that there is no $x$ with property $P$. This is correct, since $S \subseteq I$ by our invariant and since a nonempty $S$ has a length of at least $\ell$. Thus if the length of $I$ is less than $\ell$, $S$ must be empty.

3. So assume that the length of $I$ is larger than $\ell$. Let $z$ be the midpoint of $I$. If $z$ has property $P$, we stop.

4. Otherwise, we replace $I$ by either the left half or the right half of $I$ and continue with step 2. We exclude the half which is known to contain no point in $S$ and hence the invariant $S \subseteq I$ is maintained.

We have already argued correctness of our method. Let us next bound the number of iterations. After $k$ executions of step four the length of $I$ is $2^{-k}2U$, since we start with an interval of length $2U$ and since every execution of step four halves the length of the interval.

Assume that the loop body is executed $k+2$ times, i.e., in the $k+2$-nd iteration we stop in either step 2 or step 3. Since we did not stop in iteration $k+1$, the length of $I$ at the beginning of the $k+1$-st iteration (=end of $k$-th iteration) is at least $\ell$. Thus $2^{-k}(2U) \geq \ell$ or $k \leq \log((2U)/\ell)$.

[4] This item is usually called a *separation oracle*: the word oracle emphasizes that, at least for the purposes of the current discussion, it is irrelevant how the decision is made, separation indicates that the decision separates $z$ from $S$.

**Theorem 2** *Binary search solves the search problem described above with at most*

$$2 + \log \frac{2U}{\ell}$$

*iterations.*

**Proof:** by the preceding discussion. ∎

Before we generalize to higher dimensions, we observe that it is not really important that the length of $I$ is halved in every iteration. Any reduction by a constant factor $f < 1$ would do. Assume the length of $I$ is reduced by a factor $f < 1$ in every iteration. Then the last paragraph has to be changed to:

Assume that the loop body is executed $k+2$ times, i.e., in the $k+2$-nd iteration we stop in either step 2 or step 3. Since we did not stop in iteration $k+1$, the length of $I$ at the beginning of the $k+1$-st iteration (=end of $k$-th iteration) is at least $\ell$. Thus $f^k(2U) \geq \ell$ or $k \leq \log_{1/f}((2U)/\ell) = \frac{\log((2U)/\ell)}{\log(1/f)}$. Thus only the basis of the log-function changes.

## 1.5 The Ellipsoid Method in Two Dimension

The ellipsoid method generalizes binary search to higher dimension. We discuss the generalization to two dimensions. Let us first generalize our search problem. Consider the following situation.

- The goal is to find an $x \in \mathbb{R}^2$ having a certain property $P$ or to tell that no $x \in R^2$ has property $P$.

  In the application to linear programming: $x$ has property $P$ if it satisfies (1).

- (Minimum Area) We know that the set $S$ of $x$ having property $P$ is either empty or a convex set of area at least $\ell$.

  In the application to linear programming: the set of feasible $x$ is either empty or has volume at least $\ell$.

- (Localization) We know that any $x$ with property $P$ is contained in a disk with radius $U$ centered at the origin.

  In the application to linear programming: the feasible set is contained in a ball of radius $U$.

- (Separation Oracle) Given an $x \in \mathbb{R}^2$ we can test whether $x$ has property $P$. If $x$ does not have property $P$, there is a line $l$ passing through $x$ such one of the (open) half-spaces defined by $l$ contains no point in $S$. This follows from the fact that $S$ is convex. We assume that we can determine such a line and the empty open half-space (equivalently, the closed half-space $H$ with boundary $l$ and containing $S$)

  In the application to linear programming: either $x$ satisfies all inequalities or we can find one which is violated by $x$.

How can we generalize binary search? What is the proper generalization of the interval $I$ which is known to contain $S$? Khachian showed that ellipses work. This leads to the following algorithm.

1. Initialize $E$ to the disk with radius $U$ centered at the origin. Then $S \subseteq E$. In the course of the algorithm we will shrink $E$ and maintain the invariants that

   - $E$ is an ellipse
   - $E$ contains $S$.

2. If the area of $E$ is less than $\ell$, we stop and declare that there is no $x$ with property $P$. This is correct, since $S \subseteq I$ by our invariant and a non-empty $S$ has an area of at least $\ell$. Thus if the area of $E$ is less than $\ell$, $S$ must be empty.

3. So assume that the area of $E$ is larger than $\ell$. Let $z$ be the center of $E$. If $z$ has property $P$, we stop.

4. Otherwise, by our assumption, we can determine a closed half-plane $H$ having $z$ in its boundary and containing $S$. Define $(1/2)E$ as the intersection of $E$ and $H$. Then $S \subseteq (1/2)E$ and the area of $(1/2)E$ is one-half the area of $E$.

   Unfortunately, $(1/2)E$ is not an ellipsoid. Here is where the ellipsoid method goes beyond binary search. In binary search $(1/2)I$ is an interval and hence we immediately proceed to the next iteration. In the ellipsoid method we need one further step.

   Set $E$ to the smallest (in area) ellipse containing $(1/2)E$, see Figure 1.

Clearly, if the method terminates, it terminates with the correct answer. The key for the termination proof and the running time analysis is to show that the area of $E$ is multiplied by a factor $f$ less than one in every iteration.

How can we determine the smallest ellipsoid containing $(1/2)E$? Let us start with a particularly simple situation, see Figure 2: $E$ is a unit disk centered at the origin and $H$ is the left half-plane.

The figure suggests to use an ellipse $E'$ that has its center on the negative $x$-axis, passes through points $(-1,0)$, $(0,1)$ and $(0,-1)$, and has axes parallel to the coordinate axes. Thus

$$E' = \{\, (x,y) \in R^2;\ ((x-c)/a)^2 + (y/b)^2 \leq 1 \,\}$$

for appropriate constants $a$, $b$, and $c$.

**Lemma 6** *For $c = -1/3$, $a = 2/3$, and $b = 2/\sqrt{3}$, the ellipsoid $E'$ defined above contains $E \cap H$. Moreover,*

$$\frac{area(E')}{area(E)} = \frac{(2/3)^2}{\sqrt{1/3}} = \frac{4\sqrt{3}}{9} \leq 0.8$$

*and $E'$ is the smallest ellipse (in area) containing $E \cap H$.*

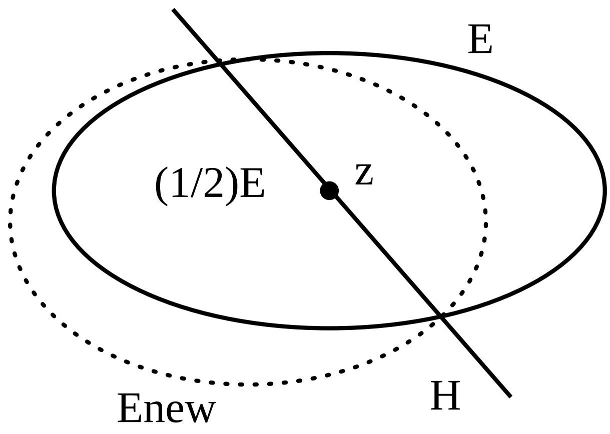


Figure 1: The figure shows an ellipse $E$, and a halfspace $H$ having $z$, the center of $E$, on its boundary; $(1/2)E$ is the intersection of $E$ and $H$. *Enew* is an ellipsoid containing $(1/2)E$. It is not the smallest such ellipsoid (my mastering of xfig did not suffice for this purpose).

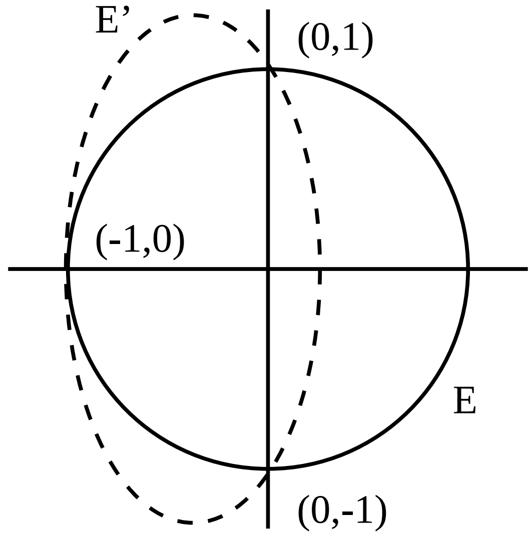


Figure 2: $E$ is the unit disk and $H$ is the left half-plane. The figure suggests that the smallest ellipse $E'$ containing $E \cap H$ has its center on the negative $x$-axis, passes through points $(-1,0)$, $(0,1)$ and $(0,-1)$, and has axes parallel to the coordinate axes.

**Proof:**

We have $-1 \leq c \leq 0$ and $a$ and $b$ are the length of the axes of our ellipse, $a > 0$ and $b > 0$. The area of $E'$ is $\pi ab$ and we are going to choose $a$ and $b$ such that the area is minimal. Since $(-1,0)$ lies on the boundary of $E'$, the length of the horizontal axis is $1+c$. Thus $a = 1+c$. We must clearly have $2a \geq 1$. Thus $-1/2 \leq c \leq 0$. Since $(0,1)$ lies on the boundary of $E'$ we have $(c/(1+c))^2 + (1/b)^2 = 1$. Thus $(1/b)^2 = 1-(c/(1+c))^2 = ((1+c)^2 - c^2)/(1+c)^2$ and hence $b = (1+c)/\sqrt{1+2c}$. At this point, we are left with a single parameter $c$. For any $c$ with $-1/2 < c \le 0$, the formulae above determine $a$ and $b$.

The area of $E'$ is $\pi ab = \pi(1+c)^2/\sqrt{1+2c}$. For $c = 0$, $E'$ is equal to $E$ and the area is equal to $\pi$, the area of the unit disk. For $c = -0.1$, the area of $E'$ is $\pi 0.9^2/\sqrt{0.8} \approx 0.9\pi$ and hence the area of $E'$ is only 90% of the area of $E$. We can do even better, but there is actually no need to do so. Already at this point, we know that the number of iterations is logarithmic in $U/\ell$.

The area is minimized[5] for $c = -1/3$. Then

$$\frac{area(E')}{area(E)} = \frac{(2/3)^2}{\sqrt{1/3}} = \frac{4\sqrt{3}}{9} \le 0.8 \ .$$

The center of $E'$ is at $(-1/3, 0)$ and the axes have length $2/3$ and $2/\sqrt{3}$ respectively. Thus $1/a^2 = 9/4$ and $1/b^2 = 3/4$.

We still need to verify that $E'$ contains $E \cap H$. So let $(x,y)$ be arbitrary with $x^2 + y^2 \le 1$ and $x \le 0$. Then

$$\frac{(x-c)^2}{a^2} + \frac{y^2}{b^2} = \frac{9}{4}(x+\frac{1}{3})^2 + \frac{3}{4}y^2 = \frac{6}{4}x^2 + \frac{3}{2}x + \frac{1}{4} + \frac{3}{4}(x^2+y^2) \le \frac{3}{2}x^2 + \frac{3}{2}x + 1 \le 1 \ .$$

The next to last inequality follows from $x^2 + y^2 \le 1$ and the last inequality follows from $x^2 + x \le 0$ for $-1 \le x \le 0$. ∎

How can we deal with the case of an arbitrary ellipse $E$? We can derive the equations for $E'$ as follows:

1. translate space so that $z$ is moved into the origin.
2. rotate the plane such that the axes of $E$ become aligned with the coordinate axes.
3. scale $x$ and $y$-coordinates such that $E$ becomes the unit disk
4. at this point the boundary of $H$ is an arbitrary line through the origin, rotate space again so as to turn $H$ into the left half-plane.
5. at this point we are in our special situation and hence know $E'$.
6. apply steps 1 to 5 in reverse order to $E'$.

A nice fact about the transformations used in steps 1 to 5 is that, although they change volumes they do not change the ratio of volumes and hence the bound on the ratio of the volume of $E$ and $E'$ derived above stays true.

[5]The derivative of $(1+c)^2/\sqrt{(1+2c)}$ with respect to $c$ is $2(1+c)(1+c)^{-1/2} - (1+c)^2(1+2c)^{-3/2} = (2(1+c)(1+2c) - (1+c)^2)(1+2c)^{-3/2} = (1+c)(2(1+2c) - (1+c))(1+2c)^{-3/2}$. We have maxima for $c = -1$ (an illegal value) and $2+4c-1-c = 0$ or $3c = -1$ or $c = -1/3$.

**Theorem 3** *The ellipsoid method solves the two-dimensional search problem with at most*

$$2 + \log_{1/f} \frac{4U^2}{\ell} = 2 + \frac{\log(4U^2/\ell)}{\log(9/(4\sqrt{3}))}$$

*iterations;* $f = 4\sqrt{3}/9$.

**Proof:** We start with the disk of radius $U$ centered at the origin. Its area is $\pi U^2 \leq 4U^2$. We terminate when the area is smaller than $\ell$ and we reduce the area by a factor $f = 4\sqrt{3}/9$ in every iteration. Therefore the number of iterations is at most what is stated in the theorem. ∎

# References


[Chv93] V. Chvatal. *Linear Programming*. Freeman, 93.

[KS96] D.R. Karger and C. Stein. A new approach to the minimum cut problem. *Journal of the ACM*, 43(4):601–640, 1996.

[MN99] K. Mehlhorn and S. Näher. *The LEDA Platform for Combinatorial and Geometric Computing*. Cambridge University Press, 1999.

[SW97] M. Stoer and F. Wagner. A simple min-cut algorithm. *Journal of the ACM*, 44(4):585–591, July 1997.