CS 684: Approximation and Network Algorithms	10/17/01
Ellipsoid method for linear programs #8	scribe: Mikhail Kobyakov

In this lecture, we will cover the ellipsoid method to solving linear programs. One interesting point to the method, is that it needs not the linear program stated explicitly. We'll see later what we will need to know about our linear program to be able to run the algorithm.

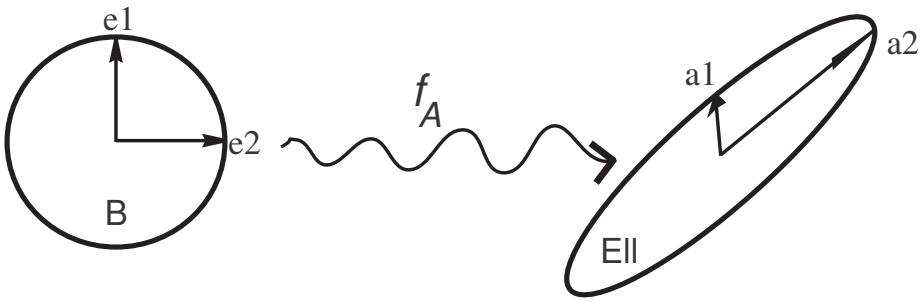
0.1 What is an ellipsoid?

Remember that a unit ball is defined as follows $B = \{ x \in \mathbb{R} \mid ||x|| \leq 1 \}$, with the ubiquitous 2-norm $||x|| = \sqrt{\sum_i x_i^2}$.

Now, let us examine a linear transformation $f_A : x \rightarrow Ax$, where A is a matrix with $\det(A) \neq 0$. We will define an ellipsoid as the image of the unit ball under such a linear transformation is an ellipsoid.

Examples:

- For a diagonal matrix
$$A = \begin{pmatrix} a & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & a \end{pmatrix}$$
$f_A(B)$ is a ball of radius a around the origin.
- For a 2x2 matrix
$$A = \begin{pmatrix} r & 0 \\ 0 & 1 \end{pmatrix}$$
$f_A(B)$ in two dimensions is an ellipse around the origin which has r horizontal radius, and 1 vertical.
- Let $A = [\ \vec{a}_1\ \vec{a}_2\ ...\ \vec{a}_n\ ]$ for a general basis $\{\ \vec{a}_i \mid i = 1, ..., n\ \}$



The picture correctly suggests that $f_A(B)$ rotates and stretches the unit ball according to the basis vectors which are the columns of A. We will take it as a sufficient proof, although it is really just a picture. However, all of these ellipsoids are centered around the origin. Our definition of a general ellipsoid will be the image of of the unit ball under linear rotation, stretch, and translation. The following is its precise form, together with two alternate and equivalent definitions.

Lemma 0.1.1 (General Ellipsoid) For an ellipsoid defined by a matrix A and a center point x_0 the following are equivalent

$$\begin{aligned} \text{Ell}(x_0, A) &= f_{A,x_0}(B) \text{ where } f_{A,x_0}(x) = x_0 + Ax & (1)\\ &= \{\ y \mid y = x_0 + Ax\ \wedge\ ||x|| \leq 1\ \} & (2)\\ &= \{\ y \mid ||A^{-1}(y - x_0)|| \leq 1\ \} & (3) \end{aligned}$$

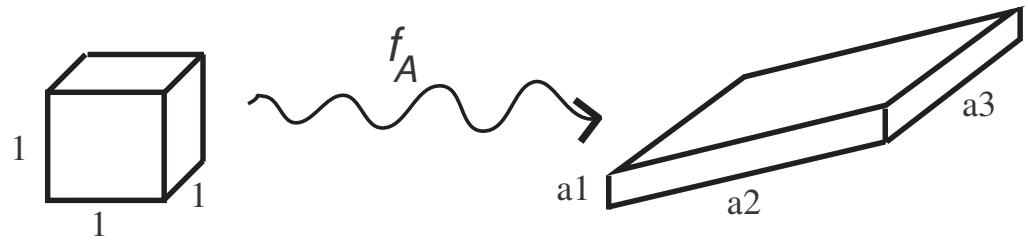
0.2 Volume of an ellipsoid

Firstly, we will need to know the volume of an n-dimensional ball, B. For simplicity that is a constant $\text{vol}_n(B) = c_n$. Taking into account the geometric meaning of the determinant of A, for some matrix the columns of which are a basis, we have the fact that

$$\text{vol}_n(\text{Ell}(x_0, A)) = c_n \cdot |\det(A)|$$

The origins of the formula is in the fact that a parallelepiped is the image of a unit cube under the same linear transformation that takes B to the ellipsoid. (Note: a cube can be thought of as an n-dimensional ball under the sup norm.) It is sufficient for our purposes to take the proof of Gaussian elimination from linear algebra. The following picture is a hint to the volume of a parallelepiped P spanned by the columns of A,

$$P(\vec{a}_1, ..., \vec{a}_n) = \{\ \sum_{i=1}^{n} \vec{a}_i \cdot x_i \mid 0 \leq x_i \leq 1,\ \forall i\ \}$$



And thus, the volume of the parrallelepiped is the volume of the image of the unit cube under the same f_A as we had before.

$$|\det(A)| = \text{vol}_n(P(\vec{a}_1, ..., \vec{a}_n)) = \text{vol}_n(f_A(\{\ (x_1, ..., x_n) \mid 0 \leq x_i \leq 1, \forall i\ \}))$$

From the volume formula for cubes and parallelepipeds, we can prove it for balls and ellipsoids. The volume formula $\text{vol}_n(f_A(O)) = \text{vol}_n(O) \cdot |\det(A)|$ actually works for any object, O. If we draw a cube completely inside an object O, and then apply the linear transformation A, the cube will become a parallelepiped inside $f_A(O)$. So, by packing the original object O full with cubes, we essentially pack the image $f_A(O)$ full with parallelepipeds. From this we clearly get

$$\text{vol}_n(\text{Ell}(x_0, A)) = c_n \cdot |\det(A)|$$

1 Linear programs

1.1 What is a linear program?

So far we have seen a simple question form of a linear program "does there exist a solution for $Ax \leq b$".

$$\exists x? \text{ such that } Ax \leq b$$

A harder version of this question is "find a solution for $Ax \leq b$ to minimize $c \cdot x$". We have also seen that the hard question can be at least approximated with a binary search on a small quantity γ, and a repeated simpler question

$$\exists x? \text{ such that } Ax \leq b \ \wedge \ c \cdot x \leq \gamma$$

We will change our definition of linear program. First, we define a polytope

$$\mathcal{P} = \{ \ x \mid Ax \leq b \ \}$$

This polytope could be an unbounded or empty region in $\mathbb{R}^n$ enclosed by $n-1$-dimensional planes. For simplicity, if A is square, then there are exactly n planes. Certainly, finding such a polytope is in every way difficult, in the sense that it is difficult to find its vertices and exact boundaries. For example, a cube in n dimensions has exponentially many vertices (2^n) as it has inequalities ($2n$). On the other hand, knowing the exact boundaries allows us to immediately minimize $c \cdot x$. To make this problem a little easier, we will once again employ the technique of introducing a small quantity ϵ, and answering the simpler "is the volume of the polytope zero?"

$$\text{vol}_n(\mathcal{P}) = 0?$$

In order to establish an equivalence of the original linear program to our version, we state two facts.

Fact 1 (Non-zero volume polytope) $\exists x$ satisfying $Ax \leq b \implies \forall \epsilon > 0, \mathcal{P}_\epsilon = \{ \ x \mid Ax \leq b + \epsilon \ \}$ satisfies $\text{vol}_n(\mathcal{P}_\epsilon) \neq 0$

Fact 2 (Small enough polytope) $\forall$ small enough $\epsilon > 0$, $\nexists x$ satisfying $Ax \leq b \implies \text{vol}_n(\mathcal{P}_\epsilon) = 0$

An important note here is that if ϵ is small enough, then both facts apply, and establish an if and only if relation, that existence of a solution to $Ax \leq b$ is equivalent to $\mathcal{P}_\epsilon$ has zero volume. We can intuitively take (without proof) that $\epsilon = \lambda^{-n}$ where λ is the maximum absolute value over the entries of A, is small enough at least most of the time.

Also, using Facts 1 and 2 together with a small enough ϵ, we end up with a P_ϵ, that is either empty, or has at least some known non-zero volume, namely the volume of the ϵ-radius ball. So, if our algorithm ever discovers that the volume is VERY small, even smaller than the volume of the ϵ-radius ball, then we know the polytope must be empty.

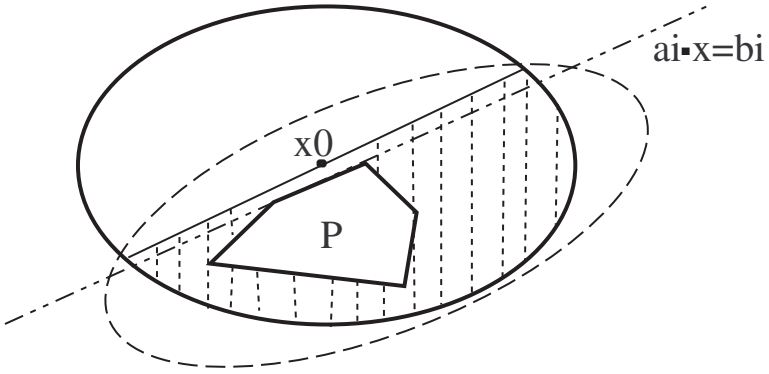
1.2 The Algorithm

This algorithm checks whether $\text{vol}_n(\mathcal{P}) = 0$. Note that we have defined $\mathcal{P} = \mathcal{P}_\epsilon$, for some small enough ϵ.

1. Start with a big ball $B(r)$ that contains our polytope, $\mathcal{P} \subset B(r)$. If $0 \leq x \leq 1$, then $r = \sqrt{n}$. In general, $L \leq x \leq U$ for some upper and lower bounds. In most practical applications, it is very straightforward to guess the bounds of x. If such bounds cannot be guessed, then the initial step is harder than simply picking the ball $B(\sqrt{max(U, -L)})$, however, we will not look at such cases.

2. General step: We know that $\mathcal{P}$ is completely inside $\text{Ell}(x_0, A)$. Test for a small neighborhood around the center, $B_\delta(x_0), \ \delta > 0$, to be entirely inside $\mathcal{P}$. This can be accomplished by checking that no inequality of the linear program is tight or violated on x_0.

   - If no inequality is tight or violated, we are done, $\text{vol}_n(\mathcal{P}) \geq \text{vol}_n(B_\delta(x_0)) > 0$.
   - Else, some row of $Ax \leq b$ has either $a_i \cdot x_0 > b_i$ or $a_i \cdot x_0 = b_i$, then the plane $a_i \cdot x = b_i$ translated to cross x_0 cuts the ellipsoid in two halves, with exactly one half containing all of $\mathcal{P}$. Then we construct a new ellipsoid, of smaller volume, containing the entire half containing $\mathcal{P}$.



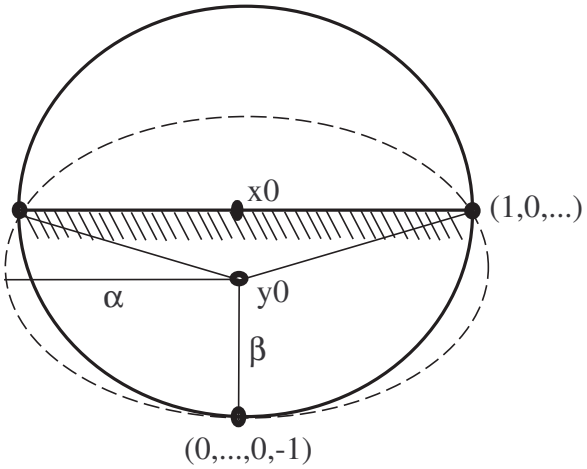
3. Iterate over smaller and smaller ellipsoids, until we find $\text{vol}_n(\mathcal{P}) \neq 0$ or $\text{vol}_n(\text{Ell})$ is very small.

Notice, that in the general step of the algorithm, we only use the fact that there is some row of A for which x_0 is tight. We do not use or even look at the whole A. This makes the algorithm use A implicitly, in the sense that we do not need the whole matrix, just a tight or violated row of it. Moreover, there may be other ways of finding tight or violated rows than looking at the matrix. For a flow problem, tightness may be expressed by a saturated edge, or a terminal node meeting its demand. We see that we may not need to explicitly define A for such cases.

1.3 Drawing a new ellipsoid

We need a general method for all ellipsoids, and inequalities that x_0 violates or is tight on. Here we only do a special case of a ball to get intuition. The rest is more just messy linear algebra. For a general instance, we use the inverse of our f_A linear transformation to get back to the ball case. Then rotate the ball to get the new inequality to the horizontal one. Then use the special case to find the new ellipsoid, and then do all these transformations back.

In the special case we have a ball $B = \{\ x \mid ||x|| \leq 1\ \}$, and the tight or violated inequality shows that $x_n \leq 0$. We will move the center of the ellipsoid from $(0, ..., 0)$ to $y_0 = (0, ..., 0, \frac{-1}{n+1})$, and we will squash the ellipsoid in the n^{th} dimension by $\beta = \frac{n}{n+1}$ and stretch the ellipsoid in every other direction by α.



Note that the new ellipsoid will be defined $\mathrm{Ell}(y_0, A')$, where

$$A' = \begin{pmatrix} \alpha & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & \alpha & 0 \\ 0 & \dots & 0 & \beta \end{pmatrix}$$

Now what is α? Vector $(1, 0, ..., 0, \frac{-1}{n+1})$ must satisfy $||A'^{-1}(y - y_0)|| \leq 1$. But, $A'^{-1}(y - y_0) = (\alpha^{-1}, 0, ..., 0, \frac{-1}{n}) \Rightarrow ||A'^{-1}(y - y_0)|| = \sqrt{\alpha^{-2} + \frac{1}{n^2}} \leq 1$. In order for the ellipsoid to exactly touch the polytope-half of the original ball at the point $(1, 0, ..., 0, \frac{-1}{n+1})$, we see the equality $\sqrt{\alpha^{-2} + \frac{1}{n^2}} = 1$. From this follows that,

$$\alpha = \sqrt{\frac{n^2}{n^2 - 1}}$$

The volume of the new ellipsoid, is of course

$$\begin{aligned} \mathrm{vol}_n(\mathrm{Ell}(y_0, A')) &= c_n \cdot |\det(A')| = c_n \cdot \alpha^{n-1} \cdot \beta \\ &= c_n \cdot (\frac{n^2}{n^2 - 1})^{\frac{n-1}{2}} \cdot \frac{n}{n+1} \\ &\approx c_n \cdot e^{\frac{1}{n^2-1}\frac{n-1}{2}} \cdot e^{\frac{-1}{n+1}} = c_n \cdot e^{\frac{-1}{2(n+1)}} \end{aligned}$$

Unfortunately, the decrease in volume is minimal, but nonetheless a multiplicative factor. Thus, over iterations of the general step, the volume will decrease exponentially, and we will decrease to a small enough volume of the ellipsoid in a polynomial number of steps.

This completes the special case drawing of the new ellipsoid, the algorithm, and the ellipsoid method.