CMPUT 675: Topics in Combinatorics and Optimization **Fall 2016**

Lecture 21-23 (Oct. 26, 28, 31): The Ellipsoid Method

Lecturer: Zachary Friggstad *Scribe: Zachary Friggstad*

21.1 The Ellipsoid Method: An Overview

We discus the ellipsoid method. This was the first polynomial-time algorithm discovered for solving linear programs.

At the heart of it is a routine solving the following problem: Given a convex body $\mathcal{P}$ and radius R that ensures $\mathcal{P}$ is contained in the ball of radius R about 0, either find *some* point in $\mathcal{P}$ or determines the volume of $\mathcal{P}$ is very tiny (at most ϵ times the unit ball volume). This can be solved in time that is polynomial in $n, \log R$ and $\log \frac{1}{\epsilon}$.

To apply this high-level idea and solve linear programs, we will (roughly speaking) do the following:

- Slightly relax the constraints by an $\epsilon' > 0$ (with polynomial bit complexity) on the right hand side that ensures the following: a) if the polyhedron in question was non-empty, its volume is not significant enough that the ellipsoid method is guaranteed to find a point b) if the polyhedron was empty then it remains empty under the tiny bit of additional slack. This will use another application of Farkas' lemma. Thus, we can decide exactly if a polytope is empty or not.
- Finding a feasible primal and dual that are both optimal is a system of linear constraints (with the objective function vectors being set equal to each other). We can iteratively use the "feasibility" checking algorithm to turn some inequalities into equalities. This will allow us to find a basis of tight constraints, from which we can extract an optimum solution.

21.2 Ellipsoids

We start with a quick review of positive definite and positive semidefinite matrices. We emphasize that every time we discuss positive (semi)definiteness of a matrix that the matrix will be symmetric.

Definition 1 *A symmetric matrix $\mathbf{M} \in \mathbb{R}^{n \times n}$ is* **positive semidefinite** *(psd) if for every $\mathbf{x} \in \mathbb{R}^n$ we have $\mathbf{x}^T \cdot \mathbf{M} \cdot \mathbf{x} \geq 0$. We use $\mathbf{M} \succeq 0$ to indicate $\mathbf{M}$ is psd. We further call $\mathbf{M}$* **positive definite** *if $\mathbf{x}^T \cdot \mathbf{M} \cdot \mathbf{x} > 0$ for all nonzero $\mathbf{x} \in \mathbb{R}^n$ and use $\mathbf{M} \succ 0$ to indicate $\mathbf{M}$ is positive definite.*

Recall every symmetric matrix $\mathbf{M} \in \mathbb{R}^{n \times n}$ has precisely n real eigenvalues if you count them with appropriate multiplicity: a) using their multiplicity as roots of the characteristic polynomial of M, or b) using the dimension of their corresponding eigenspaces. The two notions of multiplicities agree for symmetric matrices over $\mathbb{R}$.

Recall the following from standard linear algebra.

Theorem 1 *The following are equivalent for a symmetric* $\mathbf{M} \in \mathbb{R}^n$.

- $\mathbf{M} \succeq 0$
- *All eigenvalues* λ_i *of* $\mathbf{M}$ *are nonnegative.*
- *We can write* $\mathbf{M} = \mathbf{U} \cdot \mathbf{D} \cdot \mathbf{U}^T$ *where* $\mathbf{U}$ *is a matrix whose columns are unit vectors and are pairwise-orthogonal and* $\mathbf{D}$ *is a diagonal matrix (i.e. all off-diagonal entries are 0) with nonnegative diagonal entries.*

In the latter case, for each $1 \leq i \leq n$ we have $D_{i,i}$ is an eigenvalue for $\mathbf{M}$ and the i'th column of $\mathbf{U}_i$ is a corresponding eigenvector. If $\mathbf{M} \succ 0$ (equivalently, all eigenvalues are strictly positive) then $\mathbf{M}^{-1} = \mathbf{U} \cdot \mathbf{D}^{-1} \cdot \mathbf{U}^T$ where $\mathbf{D}^{-1}$ is simply obtained by inverting each diagonal entry of $\mathbf{D}$. Finally, recall that the determinant of a matrix is the product of its eigenvalues. This is especially easy to see with psd matrices:

$$\det \mathbf{M} = (\det \mathbf{U}) \cdot (\det \mathbf{U^T}) \cdot (\det \mathbf{D}) = (\det(\mathbf{U} \cdot \mathbf{U^T})) \cdot (\det \mathbf{D}) = (\det \mathbf{I}) \cdot (\det \mathbf{D}) = \prod_{\mathbf{i=1}}^{\mathbf{n}} \lambda_{\mathbf{i}}.$$

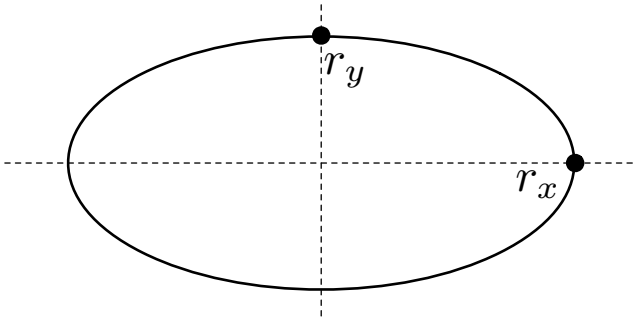
Definition 2 *Let* $\mathbf{M} \succ 0$ *and* $\mathbf{x} \in \mathbb{R}^n$. *The* **ellipsoid** *centred around* $\mathbf{x}$ *with shape* $\mathbf{M}$ *is*

$$\mathrm{E}(\mathbf{M}, \mathbf{x}) = \{\mathbf{z} \in \mathbb{R}^n : (\mathbf{z} - \mathbf{x})^T \cdot \mathbf{M}^{-1} \cdot (\mathbf{z} - \mathbf{x}) \leq 1\}.$$

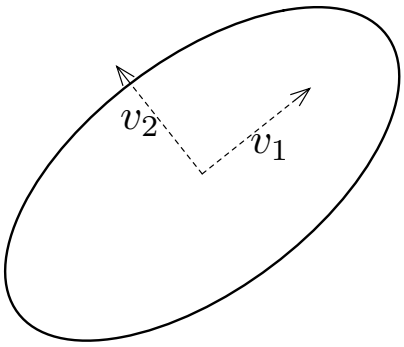
Let's get some intuition about this first. You might recall that an ellipse in the 2D-plane is the set of points (x, y) satisfying

$$\frac{x^2}{r_x^2} + \frac{y^2}{r_y^2} = 1$$

for some $r_x, r_y > 0$. This is obtained by considering the unit circle centred around the origin, and then stretching it along the x-axis by r_x and along the y-axis by r_y. The area enclosed by this ellipse is then $r_x \cdot r_y \cdot \pi$ (where π is the area of the unit circle).



To obtain a rotated ellipse, specify two orthogonal unit vectors $\mathbf{v}_1, \mathbf{v}_2$ and corresponding stretch values $r_1, r_2 > 0$. If we stretched the unit circle by r_1 along $\mathbf{v}_1$ and by r_2 along $\mathbf{v}_2$ then the set of points $\mathbf{z} = (x, y)$ lying on the boundary of this stretched circle satisfy $\frac{\langle \mathbf{z}, \mathbf{v}_1 \rangle^2}{r_1^2} + \frac{\langle \mathbf{z}, \mathbf{v}_2 \rangle^2}{r_2^2} = 1$.



In general, an n-dimensional ellipse can be given by an orthonormal basis $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ and values $r_1, r_2, \ldots, r_n > 0$: the ellipse is all points $\mathbf{z}$ satisfying $\sum_i \frac{\langle \mathbf{z}, \mathbf{v}_i \rangle^2}{r_i^2} = 1$. Let B^n be the unit ball in $\mathbb{R}^n$. The ellipsoid is just the space enclosed by this ellipse and has volume $\prod_i r_i \cdot \text{vol}(B^n)$ (i.e. scale the volume of the n-dimensional unit ball by $\prod_i r_i$). We can also centre it around any point $\mathbf{x}$ by a simple translation.

This aligns with the definition of $\text{E}(\mathbf{M}, \mathbf{x})$ above. Here, $\mathbf{M}$ is a positive definite matrix. An orthogonal basis of unit-length eigenvectors (the columns of $\mathbf{U}$) specifies the axis we will stretch the ball along to get the ellipsoid. The eigenvalues (the entries of $\mathbf{D}$) are the squares of the r_i values we use to stretch the unit ball to get the ellipsoid.

Note then that $\text{vol}(\text{E}(\mathbf{M}, \mathbf{x})) = \sqrt{\det \mathbf{M}} \cdot \text{vol}(B^n)$.

21.3 Löwner-John Ellipsoids

At the heart of the Ellipsoid Method is a method that does the following. Suppose we have an ellipsoid $\text{E}(\mathbf{M}, \mathbf{x})$ that we know contains a convex body $\mathcal{P}$. Furthermore, suppose that $\mathbf{x} \notin \mathcal{P}$ and that $\mathbf{a} \in \mathbb{R}^n, \mathbf{a} \neq \mathbf{0}$ describes a "separating hyperplane" in the sense that $\mathbf{a}^T\mathbf{z} \geq \mathbf{a}^T\mathbf{x}$ for all $\mathbf{z} \in \mathcal{P}$. Then we can find a smaller ellipsoid $\text{E}(\mathbf{M}', \mathbf{x}')$ that still contains $\mathcal{P}$. The volume of this ellipsoid is small enough (compared to the volume of $\text{E}(\mathbf{M}, \mathbf{x})$) that iterating this procedure a polynomial number of times will yield an ellipsoid with exponentially smaller volume. See Figure 21.1 for a depiction.

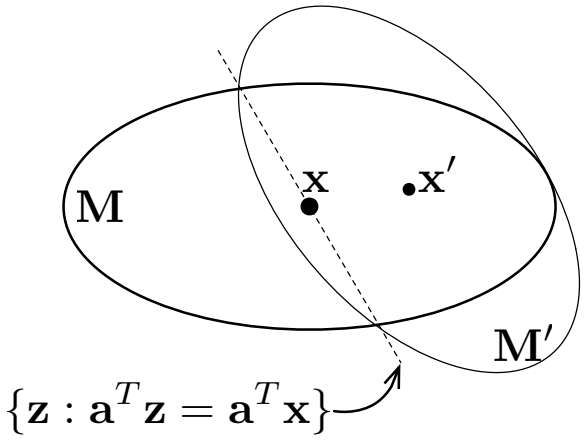


Figure 21.1: Depiction of an ellipsoid $\mathrm{E}(\mathbf{M}, \mathbf{x})$, a line through $\mathbf{x}$, and an ellipsoid $\mathrm{E}(\mathbf{M}', \mathbf{x}')$ covering the half of $\mathrm{E}(\mathbf{M}, \mathbf{x})$ contained on one side of the line.

The ellipsoid $\mathrm{E}(\mathbf{M}', \mathbf{x}')$ we will use is the following. From this point on, we assume $n \geq 2$; I think you are capable of solving linear programs with only 1 variable!

Definition 3 (Löwner-John's Ellipsoid) *Suppose $n \geq 2$. Let $\mathbf{M} \succ 0, \mathbf{x} \in \mathbb{R}^n$ and $\mathbf{a} \in \mathbb{R}^n, \mathbf{a} \neq \mathbf{0}$. Let*

$$\mathbf{b}' := \frac{1}{\sqrt{\mathbf{a}^T \mathbf{M} \mathbf{a}}} \cdot \mathbf{M}\mathbf{a}.$$

From this, define $\mathbf{M}', \mathbf{x}'$ as

$$\begin{aligned} \mathbf{M}' &:= \frac{n^2}{n^2-1}\left(\mathbf{M} - \frac{2}{n+1}\mathbf{b}'\mathbf{b}'^T\right), \\ \mathbf{x}' &:= \mathbf{x} + \frac{1}{n+1}\mathbf{b}'. \end{aligned}$$

The ellipsoid $\mathrm{E}(\mathbf{M}', \mathbf{x}')$ is the Löwner-Johns ellipsoid (for $\mathrm{E}(\mathbf{M}, \mathbf{x})$ and $\mathbf{a}$).

Of course, there are some things to prove here. For starters, we have to be convinced that $\mathbf{M}'$ is indeed positive definite and that all properties discussed above hold. You might be concerned about the square root in the calculation because we have to stay in the realm of rational numbers to implement this algorithm. It turns out it suffices to compute it within some polynomial number of bits, as long as we slightly enlarge the ellipsoid. We will briefly touch on these details later.

For the rest of this section, we fix some $\mathbf{M} \succ 0, \mathbf{x}, \mathbf{a}$ (with $\mathbf{a} \neq \mathbf{0}$) and let $\mathbf{b}', \mathbf{M}', \mathbf{x}'$ be as described in Definition 3. We first note $\mathbf{a}^T \mathbf{M} \mathbf{a} > 0$ as $\mathbf{M} \succ 0$ and $\mathbf{a} \neq \mathbf{0}$, so the construction of $\mathbf{M}'$ and $\mathbf{x}'$ is well-defined.

Lemma 1 *$\mathbf{M}'$ is symmetric and $\mathbf{M}' \succ 0$*

Proof. For symmetry, simply note $\mathbf{M}$ is symmetric and $\mathbf{b}'\mathbf{b}'^T$ is also symmetric (the outer product of two vectors is always symmetric), so $\mathbf{M}'$ is also symmetric.

That $\mathbf{M}' \succ 0$ is maybe not immediately obvious as it is the difference between a positive definite matrix and a positive semidefinite matrix. But consider

$$\mathbf{M}'^{-1} = \frac{n^2-1}{n^2}\left(\mathbf{M}'^{-1} + \frac{2}{n-1} \cdot \frac{\mathbf{a}\mathbf{a}^T}{\mathbf{a}^T\mathbf{M}\mathbf{a}}\right).$$

One can verify this is indeed the inverse of $\mathbf{M}'$ by explicitly calculating $\mathbf{M}' \cdot \mathbf{M}'^{-1}$ and seeing that you get $\mathbf{I}$.

Note that $\mathbf{M}'^{-1}$ is the positive-weighted sum of a positive definite matrix $\mathbf{M}^{-1}$ and a positive semidefinite matrix $\mathbf{a}\mathbf{a}^T$ (any outer product is positive semidefinite by a quick verification: $\mathbf{z}^T(\mathbf{a}\mathbf{a}^T)\mathbf{z} = (\mathbf{z}^T\mathbf{a})^2 \geq 0$). So $\mathbf{M}'^{-1} \succ 0$, meaning $\mathbf{M}' \succ 0$ as well (the eigenvalues of $\mathbf{M}'$ are inverses of eigenvalues of $\mathbf{M}'^{-1}$). ■

We now at least know $\mathrm{E}(\mathbf{M}', \mathbf{x}')$ is well-defined as an ellipsoid we now show one of the two the main results, that it contains the half of the original ellipsoid we are interested in. The other main result is that it's volume has shrunk by a sufficient quantity.

Theorem 2 $\mathrm{E}(\mathbf{M}', \mathbf{x}') \supseteq \mathrm{E}(\mathbf{M}, \mathbf{x}) \cap \{\mathbf{z} \in \mathbb{R}^n : \mathbf{a}^T\mathbf{z} \geq \mathbf{a}^T\mathbf{x}\}$

Proof. Let $\mathbf{z} \in \mathrm{E}(\mathbf{M}, \mathbf{x}) \cap \{\mathbf{z} \in \mathbb{R}^n : \mathbf{a}^T\mathbf{z} \geq \mathbf{a}^T\mathbf{x}\}$. We show

$$(\mathbf{z} - \mathbf{x}')^T\mathbf{M}'^{-1}(\mathbf{z} - \mathbf{x}') \leq 1$$

to determine membership in $\mathrm{E}(\mathbf{M}', \mathbf{x}')$. Expanding the left side and using the expression for $\mathbf{M}'^{-1}$ presented in the proof of Lemma 1, we have

$$\begin{aligned}
& (\mathbf{z} - \mathbf{x}')^T\mathbf{M}'^{-1}(\mathbf{z} - \mathbf{x}') \\
= \;& \frac{n^2-1}{n^2} \cdot \left(\mathbf{z} - \mathbf{x} - \frac{1}{n+1}\mathbf{b}'\right)^T \left(\mathbf{M}^{-1} + \frac{2}{n-1} \cdot \frac{\mathbf{a}\mathbf{a}^T}{\mathbf{a}^T\mathbf{M}\mathbf{a}}\right)\left(\mathbf{z} - \mathbf{x} - \frac{1}{n+1}\mathbf{b}'\right) \\
= \;& \frac{n^2-1}{n^2}\left((\mathbf{z} - \mathbf{x})^T\mathbf{M}^{-1}(\mathbf{z} - \mathbf{x}) + \frac{2}{n-1}t^2 + \frac{1}{n^2-1} - \frac{2}{n-1}t\right)
\end{aligned}$$

where we set $t = \frac{\mathbf{a}^T(\mathbf{z}-\mathbf{x})}{\sqrt{\mathbf{a}^T\mathbf{M}\mathbf{a}}}$. The last equality is a straightforward (and slightly tedious) calculation. Now, $\mathbf{z} \in \mathrm{E}(\mathbf{M}, \mathbf{x})$ means the last expression is bounded by

$$f(t) := \frac{n^2-1}{n^2}\left(1 + \frac{2}{n-1}t^2 + \frac{1}{n^2-1} - \frac{2}{n-1}t\right).$$

Direct calculation verifies $f(t) \leq 1$ for $t \in [0, 1]$. Namely, $f(0) = f(1) = 1$ and $f(t)$ is a convex quadratic (i.e. it "opens upward") so $f(t) \leq \max\{f(0), f(1)\} = 1$ for every $t \in [0, 1]$.

All that is left to show is $t \in [0, 1]$ for the given value of t.

1. Recall $\mathbf{z} \in \{\mathbf{z}' \in \mathbb{R}^n : \mathbf{a}^T \cdot \mathbf{z}' \geq \mathbf{a}^T \cdot \mathbf{x}\}$, this immediately shows $t \geq 0$.
2. Recall $\mathbf{z} \in \mathrm{E}(\mathbf{M}, \mathbf{x})$. This means

$$1 \geq (\mathbf{z} - \mathbf{x})^T\mathbf{M}^{-1}(\mathbf{z} - \mathbf{x}) = (\mathbf{z} - \mathbf{x} - t\mathbf{b}')^T\mathbf{M}^{-1}(\mathbf{z} - \mathbf{x} - t\mathbf{b}') + t^2 \geq t^2.$$

   The equality can be verified directly and the second inequality is because $\mathbf{M}^{-1} \succ 0$.

That is, $t \geq 0$ and $t^2 \leq 1$ so it must be $0 \leq t \leq 1$. ■

The next theorem shows the volume of $\mathrm{E}(\mathbf{M}', \mathbf{x}')$ is smaller than $\mathrm{E}(\mathbf{M}, \mathbf{x})$ by enough of a factor to ensure n iterations of the ellipsoid method decreases the volume of the ellipsoid by a constant factor.

Theorem 3 $\det \mathbf{M}' \leq e^{-1/(4n)} \cdot \det \mathbf{M}$.

Proof. This is equivalent to bounding $\det(\mathbf{M}^{-1}\mathbf{M}')$ by $e^{-n/4}$. Observe

$$\det(M^{-1}\mathbf{M}') = \left(\frac{n^2}{n^2-1}\right)^n \det\left(\mathbf{I} - \frac{2}{n+1}\cdot \mathbf{M}^{-1}\mathbf{b}'\mathbf{b}'^T\right) = \left(\frac{n^2}{n^2-1}\right)^n \det\left(\mathbf{I} - \frac{2}{n+1}\cdot\frac{1}{\mathbf{a}^T\mathbf{M}\mathbf{a}}\cdot \mathbf{a}\mathbf{a}^T\mathbf{M}\right).$$

Recall that the determinant of a matrix is the product of its eigenvalues. We claim the eigenvalues of $\mathbf{I} - \frac{2}{n+1}\cdot\frac{1}{\mathbf{a}^T\mathbf{M}\mathbf{a}}\cdot\mathbf{a}\mathbf{a}^T\mathbf{M}$ are 1 (with multiplicity $n-1$) and $1-\frac{2}{n+1}$ (with multiplicity 1). Note every vector is an eigenvector of $\mathbf{I}$ with eigenvalue 1. So it suffices to show the eigenvalues of $\mathbf{a}\mathbf{a}^T\mathbf{M}$ are 0 (with multiplicity $n-1$) and $\mathbf{a}^T\mathbf{M}\mathbf{a}$ (with multiplicity 1).

Note $\mathbf{a}\mathbf{a}^T$ is a rank-1 matrix, so then $\mathbf{a}\mathbf{a}^T\mathbf{M}$ is as well. That is, the image of the linear map $\mathbf{a}\mathbf{a}^T$ is one-dimensional as every vector $\mathbf{v}$ orthogonal to $\mathbf{a}$ has $\mathbf{a}\mathbf{a}^T\mathbf{v} = \mathbf{0}$, so the linear map $\mathbf{a}\mathbf{a}^T\mathbf{M}$ also has a one-dimensional image[1]. A rank 1 matrix always has 0 as an eigenvalue with multiplicity $n-1$ (anything orthogonal to the first row, thus any row, of the matrix is an eigenvector with eigenvalue 0). Finally, $(\mathbf{a}\mathbf{a}^T\mathbf{M})\mathbf{a} = \mathbf{a}(\mathbf{a}^T\mathbf{M}\mathbf{a})$ so $\mathbf{a}$ is an eigenvector with eigenvalue $\mathbf{a}^T\mathbf{M}\mathbf{a}$.

Therefore,

$$\det(\mathbf{M}^{-1}\mathbf{M}') = \left(\frac{n^2}{n^2-1}\right)^n \cdot \left(1 - \frac{2}{n+1}\right).$$

Now, using $\left(1+\frac{1}{x}\right)^x \leq e^1$ for all $x \geq 0$ and $\left(1-\frac{1}{x}\right)^x \leq e^{-1}$ for all $x \geq 1$ we see

$$\begin{aligned}\left(\frac{n^2}{n^2-1}\right)^n \cdot \left(1-\frac{2}{n+1}\right) &= \left(1+\frac{1}{n^2-1}\right)^{(n^2-1)\cdot\frac{n}{n^2-1}} \cdot \left(1-\frac{2}{n+1}\right)^{\frac{n+1}{2}\cdot\frac{2}{n+1}} \\ &\leq e^{\frac{n}{n^2-1}} \cdot e^{-\frac{2}{n+1}} \\ &\leq e^{-\frac{1}{4n}}\end{aligned}$$

The latter bound holds for $n \geq 3$. For $n = 2$, one directly verifies $\left(\frac{n^2}{n^2-1}\right)^n \cdot \left(1-\frac{2}{n+1}\right) = \frac{16}{27} \leq e^{-1/8}$. ■

21.4 The Ellipsoid Method

We describe the Ellipsoid method. Note we are not explicitly stating $\mathbf{x} \in \mathbb{R}^n_{\geq 0}$, only that $\mathbf{x} \in \mathbb{R}^n$ (but nonnegativity constraints may be encoded in the system $\mathbf{A}\cdot\mathbf{x} \leq \mathbf{b}$). Here, B^n denotes the unit ball. The value R in the input line of the algorithm describes a radius such that all feasible solutions are contained in a ball of radius R around $\mathbf{0}$ (which gives our initial ellipsoid).

[1]It is at most one-dimensional but $\mathbf{M}$ is invertible so it is exactly one dimensiona.

Algorithm 1 The Ellipsoid Method

Input: System of constraints $\mathbf{A} \cdot \mathbf{x} \leq \mathbf{b}$, bound $R \geq 0$ such that $\mathcal{P} = \{\mathbf{x} \in \mathbb{R}^n, \mathbf{A} \cdot \mathbf{x} \leq \mathbf{b}\} \subseteq \mathrm{E}(R^2 \cdot \mathbf{I}, \mathbf{x})$, value $\epsilon > 0$.

Output: A point $\mathbf{x} \in \mathcal{P}$ or the statement $\mathrm{vol}(\mathcal{P}) < \epsilon \cdot \mathrm{vol}(B^n)$.

```
M ← R^2 · I
x ← 0
κ ← 8n ln(R^n/ε).
for κ iterations do
  if x ∈ P then
    return x
  else
    let i be such that A_i · x > b.
    let E(M', x') be the Löwner-John ellipsoid for M, x with separating plane given by a := −A_i.
    (M, x) ← (M', x')
  end if
end for
return the statement vol(P) < ε · vol(B^n).
```

Recall we are letting B^n denote the unit ball in $\mathbb{R}^n$.

Theorem 4 *If Algorithm 1 does not find* $\mathbf{x} \in \mathcal{P}$*, then* $\mathrm{vol}(\mathcal{P}) < \epsilon \cdot \mathrm{vol}(B^n)$.

We caution this does not necessarily mean the final ellipsoid is contained in a small unit ball. It could be stretched very long and flat.

Proof. $\mathrm{E}(\mathbf{M}, \mathbf{x})$ always contains $\mathcal{P}$ by Theorem 2.

Suppose Algorithm 1 does not return a point in $\mathcal{P}$. The volume of the initial ellipsoid $\mathrm{E}(R^2 \cdot \mathbf{I}, \mathbf{0})$ is $\sqrt{\det(R^2 \cdot \mathbf{I})} \cdot \mathrm{vol}(B^n)$. Theorem 3 and the fact the final ellipsoid $\mathrm{E}(\mathbf{M}, \mathbf{x})$ contains $\mathcal{P}$ shows

$$\mathrm{vol}(\mathcal{P}) \leq \sqrt{\det \mathbf{M}} \cdot \mathrm{vol}(B^n) \leq \sqrt{e^{-\kappa/4n} \cdot \det(R^2 \cdot \mathbf{I})} \cdot \mathrm{vol}(B^n) = e^{-\kappa/8n} \cdot R^n \cdot \mathrm{vol}(B^n) = \epsilon \cdot \mathrm{vol}(B^n).$$

■

Other Considerations

All details discussed below are fleshed out in the Korte-Vygen textbook (Chapter 4.4).

The calculation of $\mathbf{x}'$ requires a square root. It turns out it is sufficient to compute it up to a polynomial number of bits after the decimal place. This does require us to slightly enlargen the coefficient $\frac{n^2}{n^2-1}$ in front of the definition of $\mathbf{M}'$ to ensure this truncated representation of $\mathbf{x}'$ still contains the space we want. A very mild increase ensures this and still guarantees geometric decrease in the volume every n steps.

Also, while the number of iterations is bounded by a polynomial in the bit complexity of the input, we still do not know for sure that the values themselves representing $\mathbf{M}$ and $\mathbf{x}$ will stay bounded throughout the algorithm. This is indeed the case, the Korte-Vygen textbook shows that they have polynomial bit complexity.

These are the only details we did not provide that allow us to see how to run the ellipsoid method in polynomial time.

21.5 Feasibility Testing

Here we show how to determine whether $\mathcal{P} = \emptyset$ or not. The algorithm correctly decides whether or not $\mathcal{P} = \emptyset$, but in the case $\mathcal{P} \neq \emptyset$ it finds a solution $\mathbf{x}$ that is only guaranteed to be "very close" to being feasible (i.e. it might violate constraints by a tiny amount). Later we will fix this problem.

We assume every entry of $\mathbf{A}$ and $\mathbf{b}$ is an integer. We also assume that if nonnegativity is desired, it is encoded in the constraints. This is without loss of generality (up to a factor of n in the input size), as discussed in an earlier lecture and we can simply add the constraint $-\mathbf{x}_i \leq 0$.

Let $\mathcal{P} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A} \cdot \mathbf{x} \leq \mathbf{b}\}$. Let $\Gamma = 2 \cdot n^{n/2} \cdot 2^{n\Delta}$ where 2^Δ is an upper bound on the absolute value of any entry of $\mathbf{A}, \mathbf{b}$ (so Δ is the maximum bit complexity of an input value). By earlier bounds (via Hadamard's bound) each extreme point of $\mathcal{P}$ has each component being at most $\Gamma/2$.

Consider some $\epsilon' > 0$ (to be determined momentarily). Let

$$\mathcal{Q}_{\epsilon'} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A} \cdot \mathbf{x} \leq \mathbf{b} + \epsilon' \cdot \mathbf{1}\} \cap \{\mathbf{x} \in \mathbb{R}^n : -\Gamma \cdot \mathbf{1} \leq \mathbf{x} \leq \Gamma \cdot \mathbf{1}\}$$

where $\mathbf{1}$ is the all-1 vector.

Finally, let $\Gamma' = (m+1)^{(m+1)/2} \cdot 2^{(m+1)\cdot\Delta}$.

Lemma 2 *If $\mathcal{P} \neq \emptyset$, then $\mathrm{vol}(\mathcal{Q}_{\epsilon'}) \geq \left(\frac{\epsilon'}{n\Gamma}\right)^n \cdot \mathrm{vol}(B^n)$. If $\mathcal{P} = \emptyset$, then $\mathcal{Q}_{\epsilon'} = \emptyset$ for $\epsilon' \leq \frac{1}{2m\Gamma'}$.*

Proof. Suppose $\mathcal{P} \neq \emptyset$ and let $\mathbf{x} \in \mathcal{P}$. Let $\mathbf{y} \in \mathbb{R}^n$ be such that $\max_i |\mathbf{y}_i - \mathbf{x}_i| \leq \frac{\epsilon'}{n\Gamma}$. Then for any constraint $\mathbf{A}_i, \mathbf{b}_i$ we have

$$\overline{\mathbf{A}}_i \cdot \mathbf{y} = \overline{\mathbf{A}}_i \cdot \mathbf{x} - \overline{\mathbf{A}}_i \cdot (\mathbf{x} - \mathbf{y}) \leq \overline{\mathbf{b}}_i + \Gamma \sum_i \frac{\epsilon'}{n\Gamma} \leq \mathbf{b}_i + \epsilon'.$$

Also, $-\Gamma \leq \mathbf{y}_i \leq \Gamma$ for each i because $0 \leq \mathbf{x}_i \leq \Gamma/2$. Note that the square with side lengths ℓ contains a ball with radius ℓ. So $\mathrm{vol}(\mathcal{Q}_{\epsilon'}) \geq \left(\frac{\epsilon'}{n\Gamma}\right)^n \cdot \mathrm{vol}(B^n)$.

On the other hand, suppose $\mathcal{P} = \emptyset$. By Farkas' Lemma there is some $\mathbf{y} \in \mathbb{R}^m_{\geq 0}$ such that $\mathbf{A}^T \cdot \mathbf{y} \geq 0$ and $\mathbf{b}^T \cdot \mathbf{y} = -1$. We may take $\mathbf{y}$ to be an extreme point of $\{\mathbf{z} \in \mathbb{R}^m_{\geq 0} : \mathbf{A}^T \cdot \mathbf{z} = 0, \mathbf{b}^T \cdot \mathbf{z} = -1.\}$. Then $0 \leq \mathbf{y}_i \leq \Gamma'$ for each $1 \leq i \leq m$.

Using this $\mathbf{y}$, we then see $(\mathbf{b} + \epsilon' \cdot \mathbf{1})\mathbf{y} = -1 + \epsilon' \sum_i \mathbf{y}_i \leq -1 + \epsilon' \cdot m\Gamma \leq -\frac{1}{2}$. A quick inspection of our earlier proof of Farkas' lemma shows this still certifies $\mathcal{Q}_{\epsilon'} = \emptyset$. ■

The volume of $\mathcal{Q}_{\epsilon'}$ can be seen to be even larger in the case $\mathcal{P} \neq \emptyset$ with a more careful analysis, but this will suffice.

Now we just invoke the ellipsoid method to test if $\mathcal{Q}_{\epsilon'}$ has a small volume or not. Use $\epsilon' = \frac{1}{2m\Gamma'}$ and invoke it with the target volume parameter ϵ being

$$\epsilon = \left(\frac{\epsilon'}{n\Gamma}\right)^n.$$

Let the bounding radius R be $n\Gamma$: this contains the cube centred at $\mathbf{0}$ with side lengths 2Γ, which contains all points in $\mathcal{Q}_{\epsilon'}$. As the bit complexity of ϵ', ϵ is polynomial in the input size, this takes polynomial time.

21.6 From Feasibility to Optimization

We now know how to decide if a given collection of linear constraints admits a feasible solution. Unfortunately the way we applied the ellipsoid method to solving this problem may produce a solution that slightly violates constraints. However, it still correctly decides emptiness of $\mathcal{P}$.

We first run the check on $\mathcal{P}$ itself to see if there is any solution. The exercise has you decide how to decide if the linear program is unbounded. So we now assume the primal has an optimal solution. By strong duality, this means the dual has an optimal solution as well and that the respective optimal solutions have equal value.

That is, $\mathcal{Q} = \{\mathbf{x} \in \mathbb{R}^n_{\geq 0}, \mathbf{y} \in \mathbb{R}^m_{\geq 0} : \mathbf{A} \cdot \mathbf{x} \leq \mathbf{b}, \mathbf{A}^T \cdot \mathbf{y} \geq \mathbf{c}, \mathbf{c}^T \cdot \mathbf{x} = \mathbf{b}^T \cdot \mathbf{y}\}$ has a feasible solution and any feasible solution corresponds to both an optimal primal and dual solution.

To actually find an optimal solution, try replacing inequality constraints in $\mathcal{Q}$ with equality constraints one at a time. If such a replacement yields a nonempty set, keep it. Otherwise revert it to the inequality constraint.

For any extreme point optimum $\mathbf{x}, \mathbf{y}$ of the primal and dual, respectively, we see a full-rank subsystem of tight constraints in the primal and dual. So when we have reached a point when no more primal or dual constraints can be converted to equality constraints, we know we have full-rank subsystems of tight constraints so simply solving the corresponding system of linear equations (which can be done in polynomial time using Gaussian elimination with a careful choice of steps, see the textbook) yields the optimal primal and dual solutions.

21.7 Separation Oracles

In the Ellipsoid method, it suffices to generate any violated constraint. Of course, we can search through the constraints to find one. But this gives us more flexibility: we can perhaps have exponentially many constraints not explicitly listed and simply have an algorithm that produces a violated constraint for a proposed point $\mathbf{x}$ or determine $\mathbf{x} \in \mathcal{P}$.

So, can one optimize when we are given this "oracle" access to exponentially many constraints? Yes, but the approach has to be slightly revised. For example, we can't go through the process of guessing tight constraints as with $\mathcal{Q}$ in the previous section.

The approach will be touched on next lecture.