ORF 522

Linear Programming and Convex Analysis

Ellipsoid Methods

Marco Cuturi

Reminder

- Strong Duality in LP's through Farkas Lemma
- Strong duality illustration: gravity
- Dual Simplex
- Sensitivity Analysis and scenarii for perturbation

Today

- Some reminders and formulas for ellipsoids
- Ellipsoid method for the feasibility problem
  - the bounded/full-dimensional case
  - the general case
- Ellipsoid method for optimization

Background

Background

- Simplex: US invention, Dantzig, 1947
- Klee-Minty counterexample, 1972
- People looking for polynomial pivot rules for decades.
- $'79$: Obscure "discovery" from the soviets.
- Portrayed in the paper `the mathematical sputnik of 1979`, see bb.
- The "sputnik" was the proof that LP's belonged to P.
- Proof by Khachiyan in $'79$, using earlier (unnoticed) work in convex optimization in the $'70$s.

Key geometric results

Reminder: positive definite matrices

- An important definition you all know:

  Definition 1. *A symmetric $n \times n$ matrix D is called* ***positive definite*** *(resp.* ***semidefinite*** *positive) if $\mathbf{x}^T D\mathbf{x} > 0$ (resp $\geq$) for all nonzero vectors $\mathbf{x} \in \mathbf{R}^n$.*

- In practice, run `eig` on the matrix and test positivity of eigenvalues.
- D positive definite $\Leftrightarrow D^{-1}$ positive definite.
- D positive definite, $\exists D^{\frac{1}{2}} \in \mathbf{R}^{n\times n}$ p.d. such that $D^{\frac{1}{2}}D^{\frac{1}{2}} = D$.

Reminder: ellipsoids and affine transformations

- a p.d. matrix D and a point $\mathbf{z}$ define an important kind of set

  Definition 2. *Given a p.d $n \times n$ matrix D and $\mathbf{z} \in \mathbf{R}^n$, the set*

  $$E = E(\mathbf{z}, D) = \{\mathbf{x} \in \mathbf{R}^n | (\mathbf{x} - \mathbf{z})^T D^{-1} (\mathbf{x} - \mathbf{x}) \leq 1\},$$

  is called an ***ellipsoid*** *with center $\mathbf{z}$ and axes D.*

- whenever $D = r^2 I_n$, note that $E(\mathbf{z}, r^2 I_n) = \overline{B_{\mathbf{z},r}}$.

  Definition 3. *If A is an $n \times n$ nonsingular matrix and $\mathbf{b} \in \mathbf{R}^n$, then the mapping $S : \mathbf{R}^n \mapsto \mathbf{R}^n$ defined by*

  $$S(\mathbf{x}) = D\mathbf{x} + \mathbf{b},$$

  is called an ***affine transformation***

Reminder: volumes

- An affine transformation is invertible: $S^{-1}(\mathbf{y}) = D^{-1}(\mathbf{y} - \mathbf{b})$.
- If L is any subset of $\mathbf{R}^n$, the image of S is

$$\{\mathbf{y} \in \mathbf{R}^n \mid \mathbf{y} = S(\mathbf{x}) \text{ for some } \mathbf{x} \in L\}.$$

- The volume of a set $L \subset \mathbf{R}^n$ is defined as $\mathbf{vol}(L) = \int_{\mathbf{x} \in L} d\mathbf{x}$.

Reminder: volumes

- A lemma that relates the volume of L and $S(L)$:

Lemma 1. *If* $S(\mathbf{x}) = D\mathbf{x} + \mathbf{b}$, *then*

$$\mathbf{vol}(S(L)) = |\det D|\,\mathbf{vol}(L).$$

Proof. ∘ $\mathbf{vol}(S(L)) = \int_{y \in S(L)} d\mathbf{y} = \int_{y \in S(L)} |\det J(\mathbf{x})| d\mathbf{x}$,
∘ where $J(\mathbf{x})$ is the Jacobian of the variable change $\mathbf{y} = D\mathbf{x} + \mathbf{b}$,
∘ that is $J(\mathbf{x}) = \partial S_i / \partial x_j = D$.

■

The Ellipsoid Method for the Feasibility Problem

Sequence of ellipsoids to test feasibility

- The ellipsoid method can be used to determine whether

$$P = \{\mathbf{x} \in \mathbf{R}^n \,|\, A\mathbf{x} \geq \mathbf{b}\}$$

  is **empty or not**, and in the latter case **provide a point** in it.

- Intuitive explanation:
  - The method builds
    - ▷ a sequence E_t of ellipsoids,
    - ▷ centered on points $\mathbf{x}_t$,
    - ▷ such that $P \subset E_t$.
  - At each iteration, either
    - ▷ $\mathbf{x}_t \in P \Rightarrow$ we have proved P is nonempty.
    - ▷ $\mathbf{x}_t \notin P \Rightarrow$ a least one constraint is violated, $A_i^T \mathbf{x}_t < b_i$.
      - · Hence $P \subset E_t \cap H^-_{A_i, A_i^T \mathbf{x}_t} = Q$.
      - · We can find a **smaller** ellipsoid E_{t+1} with center $\mathbf{x}_{t+1}$ that covers Q.
      - · Loop
  - Either we stop by finding a point in P,
  - Either $\mathbf{vol}(E_t) \rightarrow 0$ and stop when $\mathbf{vol}(E_t)$ is too small to conclude $P = \emptyset$.

The main tool

Theorem 1. *Let* $E = E(\mathbf{z}, D)$ *be an ellipsoid of* $\mathbf{R}^n$ *and let* $\mathbf{a} \in \mathbf{R}^n$, $\mathbf{a} \neq \mathbf{0}$. *Consider the halfspace* $H_+ = H^+_{\mathbf{a},\mathbf{a}^T\mathbf{z}} = \{\mathbf{x} \mid \mathbf{a}^T\mathbf{x} \geq \mathbf{a}^T\mathbf{z}\}$ *and let*

$$\mathbf{z}' = \mathbf{z} + \frac{1}{n+1}\frac{D\mathbf{a}}{\sqrt{\mathbf{a}^T D\mathbf{a}}},$$

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

The matrix D *positive definite and* $E' = E(\mathbf{z}', D')$ *is an ellipsoid which satisties*

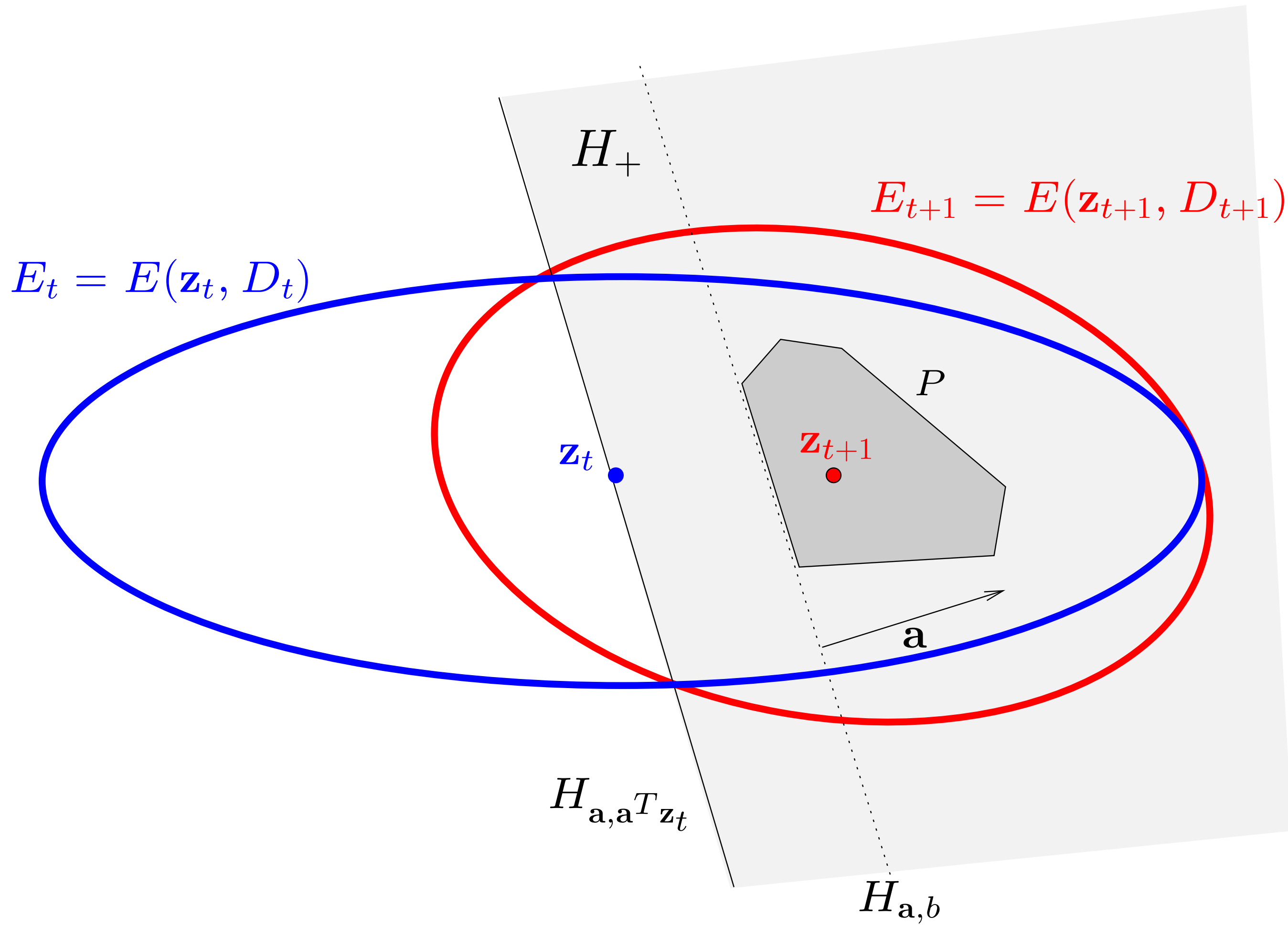
(i) $E \cap H \subset H'$

(ii) $\mathbf{vol}(E') < e^{-\frac{1}{2(n+1)}}\,\mathbf{vol}(E)$

- we are in a different mathematical world: analytical proof.

Proof

- Here is a graphical intuition of what is going on:



- We will prove this result is valid for a simple case: $\mathbf{z} = \mathbf{0}, D = I_n, \mathbf{a} = \mathbf{e}_1$.
- We will follow with a generalization to arbitrary $\mathbf{z}, D, \mathbf{a}$.

Proof of (i)

- Suppose $\mathbf{z} = \mathbf{0}, D = I_n, E_0 = E(\mathbf{0}, I_n)$ and $\mathbf{a} = \mathbf{e}_1$ which defines H_0^+.
  - In such a case,

$$E_0' = \left(\frac{\mathbf{e}_1}{n+1}, \frac{n^2}{n^2-1}\left(I_n - \frac{2}{n+1}\mathbf{e}_1\mathbf{e}_1^T\right)\right).$$

Note that the matrix is diagonal. All terms but the first equal $\frac{n^2}{n^2-1}$, the first being $(\frac{n}{n+1})^2$.

$$E_0' = \left\{\mathbf{x} \;\middle|\; \left(\frac{n+1}{n}\right)^2 \left(x_1 - \frac{1}{n+1}\right)^2 + \frac{n^2-1}{n^2}\sum_{\boldsymbol{i=2}}^{\boldsymbol{n}} x_i^2 \le 1\right\},$$

$$= \left\{\mathbf{x} \;\middle|\; \frac{n^2-1}{n^2}\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} x_i^2 + \frac{2(n+1)}{n^2}x_1^2 + \left(\frac{n+1}{n}\right)^2 \left(-\frac{2x_1}{n+1} + \frac{1}{(n+1)^2}\right) \le 1\right\}$$

$$= \left\{\mathbf{x} \;\middle|\; \frac{n^2-1}{n^2}\sum_{i=1}^{n} x_i^2 + \frac{1}{n^2} + \frac{2(n+1)}{n^2}\, x_1(x_1 - 1) \le 1\right\}.$$

Proof of (i)

- Let $\mathbf{x} \in E_0 \cap H_0^+$. Then $0 \leq x_1 \leq 1$ and therefore $x_1(x_1 - 1) \leq 0$.
- Since $\mathbf{x} \in E_0$, $\sum_{i=1}^n x_i^2 \leq 1$. Therefore,

$$\frac{n^2-1}{n^2}\sum_{i=1}^{n} x_i^2 + \frac{1}{n^2} + \frac{2(n+1)}{n^2} x_1(x_1 - 1) \leq \frac{n^2-1}{n^2} + \frac{1}{n^2} = 1$$

  meaning $\mathbf{x} \in E_0'$, hence $E_0 \cap H_0^+ \subset E_0'$.

- Consider now the general case. We build an **affine transformation** T such that

$$T(E) = E_0, \;\; T(E') = E_0' \text{ and } T(H_+) = H_0^+.$$

- The result will follow because affine transformations are such that
  - $A \subset B \Rightarrow T(A) \subset T(B)$,
  - $T(A \cap B) = T(A) \cap T(B)$,
  - $i.e.$ conserve inclusion and intersection.

Proof of (i)

- Consider the transformation $S(x) = D^{-\frac{1}{2}}(\mathbf{x} - \mathbf{z})$.
- $S(E) = E_0$...

$$\begin{aligned} S(E) &= \{\mathbf{y} |\ \mathbf{y} = D^{-\frac{1}{2}}(\mathbf{x} - \mathbf{z}), \mathbf{x} \in E\} \\ &= \{\mathbf{y} |\ \mathbf{y} = D^{-\frac{1}{2}}(\mathbf{x} - \mathbf{z}) \text{ with } (\mathbf{x} - \mathbf{z})^T D^{-1}(\mathbf{x} - \mathbf{z}) \leq 1\} \\ &= \{\mathbf{y} |\ \mathbf{y} = \mathbf{x}', \|\mathbf{x}'\|^2 \leq 1\} \\ &= \{\mathbf{y} |\ \|\mathbf{y}\|^2 \leq 1\} = E_0 \end{aligned}$$

good start. **However $S(E') \neq E_0'$ and $S(H_+) \neq H_0^+$.**

- For any vector $\mathbf{u}$, writing $\mathbf{b} = \|u\|\mathbf{e}_1$, matrix $R = \frac{2(\mathbf{u}+\mathbf{b})(\mathbf{u}+\mathbf{b})^T}{\|\mathbf{u}+\mathbf{b}\|^2} - I_n$ is such that
  - $R^2 = I_n, R^T = R$,
  - $R\mathbf{u} = \mathbf{b}$.
- Let R be the matrix corresponding to $\mathbf{u} = D^{\frac{1}{2}}\mathbf{a}$, that is $\boldsymbol{RD^{\frac{1}{2}}\mathbf{a} = \|D^{\frac{1}{2}}\mathbf{a}\|\mathbf{e}_1}$.
- Let $T(\mathbf{x}) = R \circ S(\mathbf{x})$ and prove that it the **good affine transformation** for E, H_+ and E'.

Proof of (i)

▷ For E,

$$\begin{aligned} \mathbf{x} \in E &\Leftrightarrow (\mathbf{x}-\mathbf{z})^T D^{-1}(\mathbf{x}-\mathbf{z}) \leq 1, \\ &\Leftrightarrow (\mathbf{x}-\mathbf{z})^T D^{-\frac{1}{2}} R R D^{-\frac{1}{2}}(\mathbf{x}-\mathbf{z}) \leq 1, \\ &\Leftrightarrow R D^{-\frac{1}{2}}(\mathbf{x}-\mathbf{z}) \in E_0, \\ &\Leftrightarrow T(\mathbf{x}) \in E_0, \end{aligned}$$

hence $T(E) = E_0$

Proof of (i)

- Similarly for H_+,

$$\begin{aligned}
\mathbf{x} \in H_+ &\Leftrightarrow \mathbf{a}^T(\mathbf{x}-\mathbf{z}) \geq 0, \\
&\Leftrightarrow \boldsymbol{\mathbf{a}^T D^{\frac{1}{2}} R} R D^{-\frac{1}{2}}(\mathbf{x}-\mathbf{z}) \geq 0, \\
&\Leftrightarrow \boldsymbol{\|D^{\frac{1}{2}}\mathbf{a}\|\mathbf{e}_1^T} T(\mathbf{x}) \geq 0, \\
&\Leftrightarrow \mathbf{e}_1^T T(\mathbf{x}) \geq 0, \\
&\Leftrightarrow T(\mathbf{x}) \in H_0^+,
\end{aligned}$$

hence $T(H_+) = H_0$.

Proof of (i)

- For E': Sherman-Morrison formula: $(A + uv^T)^{-1} = A^{-1} - \frac{A^{-1}uv^TA^{-1}}{1+v^TA^{-1}u}$
- apply it to reformulate conveniently D'^{-1},

$$
\begin{aligned}
D'^{-1} &= \frac{n^2-1}{n^2}\left(D^{-1} + \frac{2}{n-1}\frac{\mathbf{a}\mathbf{a}^T}{\mathbf{a}^T D\mathbf{a}}\right),\\
&= \frac{n^2-1}{n^2} D^{-1/2}\left(I + \frac{2}{n-1}\frac{D^{1/2}\mathbf{a}\mathbf{a}^T D^{1/2}}{\mathbf{a}^T D\mathbf{a}}\right) D^{-1/2},\\
&= \frac{n^2-1}{n^2} D^{-1/2}R\left(I + \frac{2}{n-1}\frac{RD^{1/2}\mathbf{a}\mathbf{a}^T D^{1/2}R}{\mathbf{a}^T D\mathbf{a}}\right) RD^{-1/2},\\
&= \frac{n^2-1}{n^2} D^{-1/2}R\left(I + \frac{2}{n-1}\frac{RD^{1/2}\mathbf{a}\mathbf{a}^T D^{1/2}R}{\mathbf{a}^T D^{1/2}RRD^{1/2}\mathbf{a}}\right) RD^{-1/2},\\
&= \frac{n^2-1}{n^2} D^{-1/2}R\left(I + \frac{2}{n-1}\mathbf{e}_1\mathbf{e}_1^T\right) RD^{-1/2},\\
&= D^{-1/2}R\left(\frac{n^2}{n^2-1}\left(I - \frac{2}{n+1}\mathbf{e}_1\mathbf{e}_1^T\right)\right)^{-1} RD^{-1/2}.
\end{aligned}
$$

Proof of (i)

- Now, for all $\mathbf{x}$ rewrite $T(\mathbf{x}-\mathbf{z'})$ using $\mathbf{z}$:

$$
\begin{aligned}
RD^{-1/2}(\mathbf{x}-\mathbf{z}') &= RD^{-1/2}(\mathbf{x}-\mathbf{z}) - \frac{1}{n+1}\frac{RD^{-1/2}D\mathbf{a}}{\sqrt{\mathbf{a}^T D\mathbf{a}}}\\
&= RD^{-1/2}(\mathbf{x}-\mathbf{z}) - \frac{1}{n+1}\frac{\boldsymbol{RD^{1/2}}\mathbf{a}}{\sqrt{\mathbf{a}^T D\mathbf{a}}}\\
&= RD^{-1/2}(\mathbf{x}-\mathbf{z}) - \frac{\mathbf{e}_1}{n+1},
\end{aligned}
$$

- Hence if $\mathbf{x}\in E' \Leftrightarrow T(\mathbf{x})\in E'_0$ and this closes the proof of (i).

Proof of (ii)

- From Lemma 1, we obtain that

$$\frac{\mathbf{vol}(E')}{\mathbf{vol}(E)} = \frac{\mathbf{vol}(T(E'))}{\mathbf{vol}(T(E))} = \frac{\mathbf{vol}(E'_0)}{\mathbf{vol}(E_0)}.$$

- Recall that

$$E'_0 = \left(\frac{\mathbf{e}_1}{n+1}, \frac{n^2}{n^2-1}\left(I_n - \frac{2}{n+1}\mathbf{e}_1\mathbf{e}_1^T\right)\right).$$

- consider the affine transformation F,

$$F(\mathbf{x}) = \left(\frac{n^2}{n^2-1}\left(I_n - \frac{2}{n+1}\mathbf{e}_1\mathbf{e}_1^T\right)\right)^{-\frac{1}{2}}\left(\mathbf{x} - \frac{\mathbf{e}_1}{n+1}\right).$$

- Note that $F(E'_0) = E_0$, or E'_0 is the image of the standard ball under transformation F^{-1}.

- Through Lemma 1, we thus have,

$$\mathbf{vol}(E_0) = \mathbf{vol}(E_0') \left| \det \left(\left(\frac{n^2}{n^2-1} \left(I_n - \frac{2}{n+1} \mathbf{e}_1 \mathbf{e}_1^T \right) \right)^{-\frac{1}{2}} \right) \right|,$$

- therefore,

$$\mathbf{vol}(E_0') = \mathbf{vol}(E_0) \sqrt{\det \left(\frac{n^2}{n^2-1} \left(I_n - \frac{2}{n+1} \mathbf{e}_1 \mathbf{e}_1^T \right) \right)},$$

and hence, using the inequality $1 + x < e^x$ valid for all $x \neq 0$ in the second line,

$$\begin{aligned} \frac{\mathbf{vol}(E_0')}{\mathbf{vol}(E_0)} &= \left(\frac{n^2}{n^2-1} \right)^{\frac{n}{2}} \sqrt{1 - \frac{2}{n+1}} = \frac{n}{n+1} \left(\frac{n^2}{n^2-1} \right)^{\frac{n-1}{2}}, \\ &= \left(1 - \frac{n}{n+1} \right) \left(1 + \frac{1}{n^2-1} \right)^{\frac{n-1}{2}} < e^{-1/(n+1)} \left(e^{1/(n^2-1)} \right)^{\frac{n-1}{2}} \\ &= e^{-\frac{1}{2(n+1)}}. \end{aligned}$$

A first simplification

- The **first version** of the ellipsoid method we study assumes that polyhedra are "regular", no pathological cases.
- **full-dimensional** is one such criterions:

  Definition 4. *A polyhedron P is **full-dimensional** if it has positive volume.*

- In practice this means that the **dimension** of P is n,
- That is the smallest vector subspace of $\mathbf{R}^n$ that contains P is $\mathbf{R}^n$.
- In the **first version we study**, we assume that P is either

(a) empty,
(b) **bounded** and **full-dimensional**, namely
  - $P \subset E(\mathbf{x}_0, r^2 I)$ whose volume is V,
  - and $\mathbf{vol}(P) > v$ for $v > 0$.

- we **assume** we are given $\mathbf{x}_0, r$ and v (lower bound), and V (upper bound).

The ellipsoid algorithm

- **Input**: $P = \{A, \mathbf{b}, \mathbf{c}\}, \mathbf{x}_0, r, v, V$
- **Output**: a feasible point $\mathbf{x}^*$ in P or the statement that P is empty.
- **Algorithm**:

1. *initialization* Let $t^\star = \lceil 2(n+1)\log(V/v) \rceil, D_0 = r^2 I, E_0 = E(\mathbf{x}_0, D_0), t = 0$
2. *main loop*
   (a) if $t = t^\star$ stop, P is empty.
   (b) if $\mathbf{x}_t \in P$ stop, P is nonempty.
   (c) if $\mathbf{x}_t \notin P$, find a violated constraint $A_i^T \mathbf{x} < b_i$ and set $\mathbf{a} = A_i$.
   (d) Let

$$\mathbf{x}_{t+1} = \mathbf{x}_t + \frac{1}{n+1} \frac{D_t \mathbf{a}}{\sqrt{\mathbf{a}^T D_t \mathbf{a}}},$$

$$D_{t+1} = \frac{n^2}{n^2 - 1} \left(D_t - \frac{2}{n+1} \frac{D_t \mathbf{a} \mathbf{a}^T D_t}{\mathbf{a}^T D_t \mathbf{a}} \right).$$

   (e) $t \leftarrow t + 1$.

The ellipsoid algorithm

Theorem 2. *Let P be a bounded polyhedron that is* **either** ***empty*** *or* ***full-dimensional*** *and for which the prior information $\mathbf{x}_0, r, v, V$ is available. Then the ellipsoid method decides correctly whether P is empty or gives a point $\mathbf{x}$ in P.*

- **Proof** If $\mathbf{x}_t \in P$ for $t < t^\star$ then the algorithm correctly decides that P is nonempty.
- Let us assume $\mathbf{x}_0, \cdots, \mathbf{x}_{t^\star - 1} \notin P$. We show that P is empty.
  - $P \subset E_k, k = 1, \cdots, t^\star$ because E_k is constructed at each step to contain P.
  - We also have that $\frac{\mathbf{vol}(E_{t+1})}{\mathbf{vol}(E_t)} < e^{-\frac{1}{2(n+1)}}$, thus

$$\frac{\mathbf{vol}(E_{t^\star})}{\mathbf{vol}(E_0)} < e^{-\frac{t^\star}{2(n+1)}},$$

    since $t^\star = \lceil 2(n+1)\log(V/v) \rceil$, $\mathbf{vol}(E_{t^\star}) < V e^{-\log(\frac{V}{v})} = v$.
  - The ellipsoid method has not terminated $\Rightarrow$ $\mathbf{vol}(P) \leq v \Rightarrow P$ is empty.

without boundedness and full-dimensionality assumptions

- Through **alternative assumptions on A and $\mathbf{b}$**, we can get rid of the boundedness and full-dimensionality assumptions which were crucial.
- The discussion is rather **technical** but interesting to follow.
- Complete proofs are omitted, only sketch given.
- Details are well explained in Bertsimas-Tsitsiklis's book.

- The issue we face is handling **unbounded** and **not fully-dimensional** polyhedra.
- P for which $\mathbf{x}_0, r, V$ and v are not known.
- Three successive lemmas to solve these issues and replace the assumptions by a bound on A and $\mathbf{b}$'s elements.

without boundedness and full-dimensionality assumptions

Lemma 2. *Let A be an $m \times n$ integer matrix, $\mathbf{b}$ a vector in $\mathbf{R}^m$ and U* ***an upper bound*** *on the* ***absolute values*** *of* ***all entries of A and $\mathbf{b}$***. *Then,*

(a) every ***extreme point*** *of the polyhedron $P = \{\mathbf{x} \in \mathbf{R}^n \mid A\mathbf{x} \geq \mathbf{b}\}$ satisfies*

$$-(nU)^n \leq x_j \leq (nU)^n, \; j = 1, \cdots, n$$

(b) every ***extreme point*** *of the standard form polyhedron $P = \{\mathbf{x} \in \mathbf{R}^n \mid A\mathbf{x} = \mathbf{b}\}$ satisfies*

$$-(mU)^m \leq x_j \leq (mU)^m, \; j = 1, \cdots, n$$

- **Proof idea** use Cramer rule and determinants of minors.
- **Remark** the extreme points of P are in $P_B = \{\mathbf{x} \in P \mid |x_j| \leq (nU)^{\{n,m\}}\}$; $P_B \subset E_B(\mathbf{0}, n(nU)^{2n} I)$ and $\mathbf{vol}(E_B) \leq (2n(nU)^n)^n = (2n)^n (nU)^{n^2}$.

without boundedness and full-dimensionality assumptions

Lemma 3. *Let* $P = \{\mathbf{x} \in \mathbf{R}^n \mid A\mathbf{x} \geq \mathbf{b}\}$ *and assume all entries of* A *and* $\mathbf{b}$ *have integer entries bounded by* U *in absolute value. Let*

$$\epsilon = \frac{1}{2(n+1)}((n+1)U)^{-(n+1)}$$

and

$$P_\epsilon = \{\mathbf{x} \in \mathbf{R}^n \mid A\mathbf{x} \geq \mathbf{b} - \epsilon\mathbf{1}\}.$$

Then we have that

(a) if P *is empty,* P_ϵ *is empty.*

(b) if P *is nonempty, then* P_ϵ *is full-dimensional.*

- **Proof idea**

  (a) use duality: if P is empty, any primal problem involving P is infeasible, and a dual formulation of a problem involving P must be feasible with unbounded objective. Modifying that dual, recover the dual of a problem involving P_ϵ and show it is also unbounded, implying P_ϵ is empty.

  (b) show we can inscribe a small ball centered on a feasible point of P in P_ϵ.

without of boundedness and full-dimensionality assumptions

Lemma 4. *Let* $P = \{\mathbf{x} \in \mathbf{R}^n \mid A\mathbf{x} \geq \mathbf{b}\}$ *be a full-dimensional bounded polyhedron and assume all entries of* A *and* $\mathbf{b}$ *have integer entries bounded by* U *in absolute value. Then*

$$\mathbf{vol}(P) > n^{-n}(nU)^{-n^2(n+1)}$$

- **Proof idea** lower bound the volume of P by the volume of the convex combination of $n+1$ arbitrary extreme points of P,
- consider such points.
- the volume of their convex combination is a determinant of a matrix using the coordinates of these points.
- such a determinant value can be lower bounded by the rhs.

Ready for an application to arbitrary polyhedra

- Consider now that $P = \{\mathbf{x} \in \mathbf{R}^n \mid A\mathbf{x} \geq \mathbf{b}\}$ where all entries of A and $\mathbf{b}$ are integers bounded by U in absolute value and assume the rows of A span $\mathbf{R}^n$.
- if P is bounded, either empty or full-dimensional,
  - Choose $v = n^{-n}(nU)^{-n^2(n+1)}$ and $V = (2n)^n(nU)^{n^2}$.
  - Get an upper bound $\lceil 2(n+1)\log(V/v)\rceil$ of the order of $O(n^4\log(nU))$.
- if P is arbitrary,
  - check whether P_B is empty which is equivalent.
  - studying $P_{B,\epsilon}$ which is bounded **and** fully-dimensional is also equivalent.
  - use the technique above,
  - the upper bound becomes $O(n^6\log(nU))$
- **Conclusion**: The **linear programming feasibility problem** with integer data can be solved in **polynomial time**.

The ellipsoid method for linear programming

The primal-dual ellipsoid method for linear programming

- Consider the dual pair of problems

$$\begin{array}{ll} \text{minimize} & \mathbf{c}^T\mathbf{x} \\ \text{subject to} & A\mathbf{x} \geq \mathbf{b} \end{array} \quad , \quad \begin{array}{ll} \text{maximize} & \mathbf{b}^T\mu \\ \text{subject to} & A^T\mu = \mathbf{c} \\ & \mu \geq 0 \end{array}$$

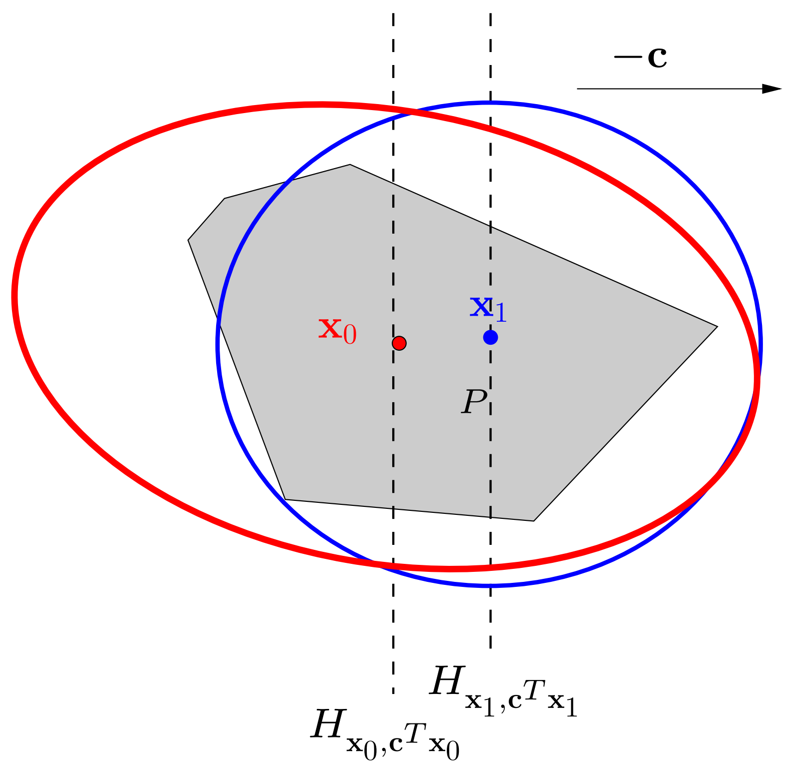
- By **strong duality**, both the primal and dual optimization problems have optimal solutions iff the following set of linear inequalities is feasible:

$$\mathbf{b}^T\mu = \mathbf{c}^T\mathbf{x}, \quad A\mathbf{x} \geq \mathbf{b},$$
$$A^T\mu = \mathbf{c}, \quad \mu \geq 0.$$

- Just test the **existence** of a feasible point $(\mathbf{x}, \mu)$ using the ellipsoid method.
- This is enough to obtain an optimum to the problem by **weak duality**.
- **Conclusion**: The **linear programming problem** with integer data can be solved in **polynomial time**.

Alternative implementation: Sliding objective ellipsoid

- Start with a problem encoded by P and $\mathbf{c}$. Assume the problem minimizes $\mathbf{c}^T\mathbf{x}$.
- Through the ellipsoid method, find a point $\mathbf{x}$ in P, write $\mathbf{x}_0 = \mathbf{x}$ and $P_0 = P$, t.
- loop as long as P_t is not empty,
  - Add a constraint to P_t: $P_{t+1} = P \cap \{\mathbf{x} \,|\mathbf{c}^T\mathbf{x} < \mathbf{c}^T\mathbf{x}_t\}$.
  - Rerun the ellipsoid method on that set and find $\mathbf{x}_t \in P_t$
- The solution is then $\mathbf{x}_t$.



Practical considerations

- in theory,
  - The **ellipsoid method** guarantees a **polynomial** upperbound on convergence to the solution of the order $O(n^6 \log(nU))$.
  - For the **simplex**, such an upperbound is **exponential**.
- in practice?
  - simplex's convergence time is usually **linear** in the number of constraints.
  - the ellipsoid method converges steadily, but **very slowly**. even with improvements that select better cuts.
- The **merit** of the ellipsoid method is that it confirmed what people were thinking, but were hoping to prove through the simplex appraoch (at least in the US).
- Spurred further research in interior point methods.
- Also useful in **general convex programming**, next course.

Next time

- An overview of interior point methods,
  - Affine scaling algorithm,
  - Potential reduction algorithm,
  - Path following algorithm.