□ Complexity Analysis

$$
(LP) \quad \begin{array}{ll} \text{Min} & \mathbf{c}^T\mathbf{x} \\ \text{s. t.} & \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq 0 \end{array}
$$

$$
\begin{cases} \mathbf{A}: & m \times n \\ \mathbf{b} & \in R^m \\ \mathbf{c}, \mathbf{x} & \in R^n \end{cases}
$$

$(\mathbf{A}, \mathbf{b}, \mathbf{c})$ input data with “integer-value”.
$L=$ number of binary bits needed to record all the data.
$(m, n, \mathbf{A}, \mathbf{b}, \mathbf{c})$ defines an *instance* of the LP.
(m, n, L) defines the *size* of the instance.

$$
\begin{aligned} L \sim \quad & \lceil 1 + \log_2 m + \log_2 n \\ & + \textstyle\sum_{j=1}^{n} \{1 + \log_2(1 + |c_j|)\} \\ & + \textstyle\sum_{i=1}^{m} \sum_{j=1}^{n} \{1 + \log_2(1 + |a_{ij}|)\} \\ & + \textstyle\sum_{i=1}^{m} \{1 + \log_2(1 + |b_i|)\} \rceil \end{aligned}
$$

- Complexity of an algorithm becomes a

function of the size, *i.e.* $f(m, n, L)$.

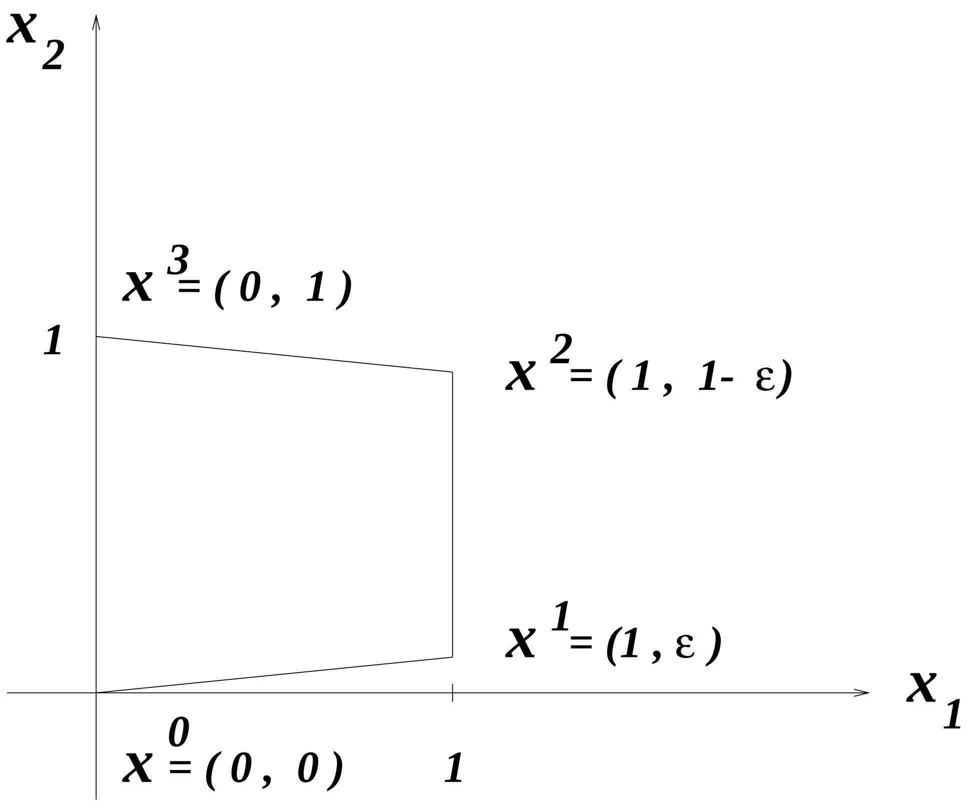
- If $\exists\ \tau > 0$ such that total # of elementary operations required by the algorithm in any instance $\leq \tau f(m, n, L)$ then, the algorithm is of order $O(f(m, n, L))$.
- If $f(m, n, L)$ is a polynomial function of m, n, L, then the algorithm is a *polynomial-time* algorithm.
- Otherwise, it is *nonpolynomial-time*.

Complexity of the Simplex Method

- Total # of elementary operations
  = (# of elementary operations at each iteration) $\times$ (# of iterations).
- # of elementary operations at each iteration of the revised simplex method $O(mn)$.
- From practical experience, the simplex method takes about (αm) iterations where $e^{\alpha} < \log_2(2 + n/m)$. Hence it is of $O(m^2 n)$.
- From the worst-case analysis, Klee and Minty [1972] showed a class of examples (in the d-dimensional space) which $2^d - 1$ iterations for the simplex method.

Klee-Minty Example

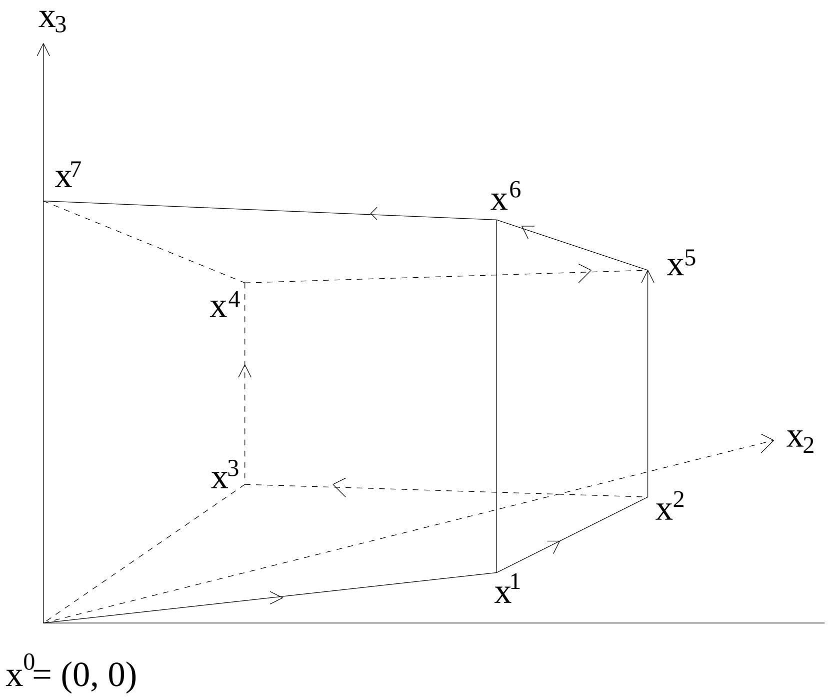
$$
\begin{array}{llll}
(2 \ \text{dim}) & \min & -x_2 & \\
& \text{s. t.} & x_1 \geq 0 & \\
& & x_1 \leq 1 & \\
& & x_2 \geq \epsilon x_1 & (0 < \epsilon < \frac{1}{2}) \\
& & x_2 \leq 1 - \epsilon x_1 & \\
& & x_1, x_2 \geq 0 &
\end{array}
$$



$\mathbf{x}^0 \to \mathbf{x}^1 \to \mathbf{x}^1 \to \mathbf{x}^3$ (optimal)

$2^2 - 1 = 3$ iterations

$$
\begin{aligned}
(3\ \text{dim}) \quad \min \quad & -x_3 \\
\text{s. t.} \quad & x_1 \geq 0 \\
& x_1 \leq 1 \\
& x_2 \geq \epsilon x_1 \\
& x_2 \leq 1 - \epsilon x_1 \\
& x_3 \geq \epsilon x_2 \\
& x_3 \leq 1 - \epsilon x_2 \\
& x_1, x_2, x_3 \geq 0
\end{aligned}
$$



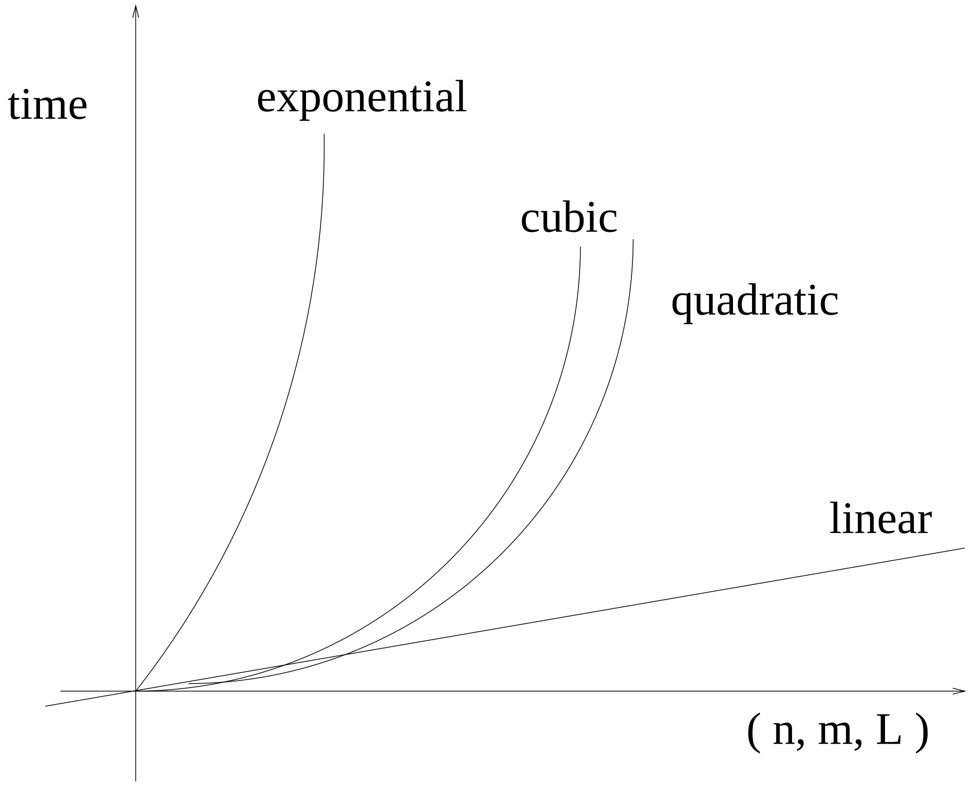
$2^3 - 1 = 7$ iterations

$$
\begin{aligned}
(d\ \text{dim}) \quad \min \quad & -x_d \\
\text{s. t.} \quad & x_1 \geq 0 \\
& x_1 \leq 1 \\
& x_2 \geq \epsilon x_1 \\
& x_2 \leq 1 - \epsilon x_1 \\
& \vdots \\
& x_d \geq \epsilon x_{d-1} \\
& x_d \leq 1 - \epsilon x_{d-1} \\
& x_i \geq 0
\end{aligned}
$$

$2^d - 1$ iterations

Hence, in theory, the simplex method is not a polynomial-time algorithm. It is an *exponential time* algorithm!

Question: What's wrong with this?



Question: Is there a polynimial-time algorithm for LP?

- Ellipsoid Method

  Levin-Shor-Khachian

  [1965] – [1970] – [1979]

  $O(n^6 L^2) \longleftarrow$ original

  $O(n^4 L^2) \longleftarrow$ improved

- Karmarkar's Algorithm [1984]

  $O(n^4 L^2) \longleftarrow$ original

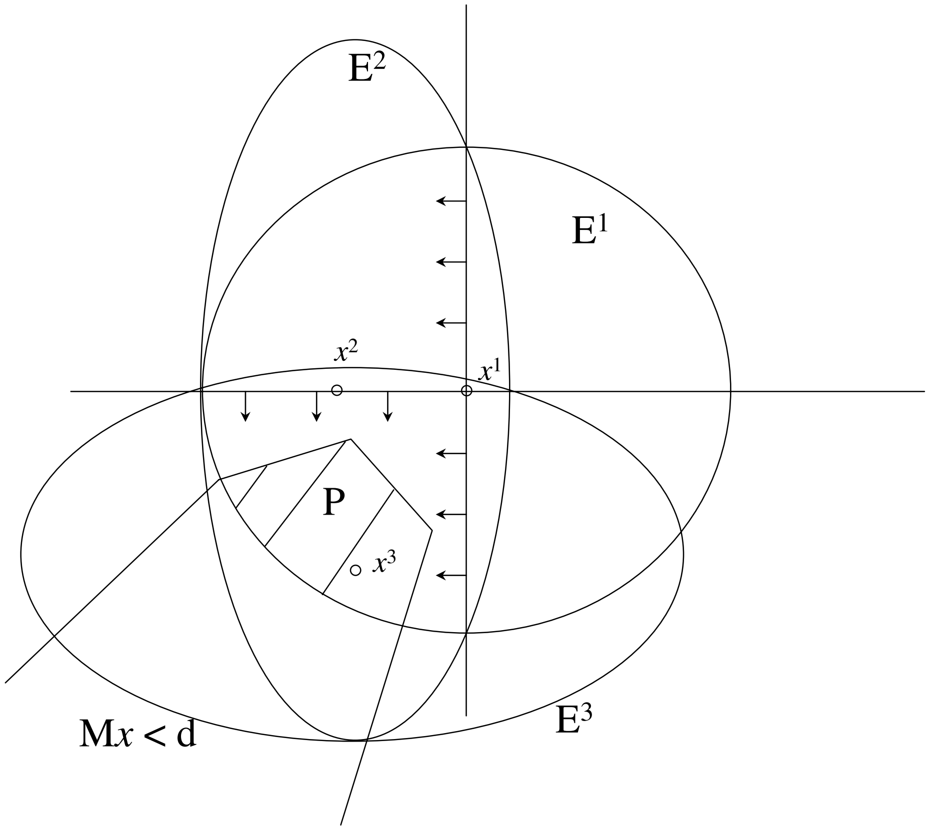
  $O(n^{3.5} L^2) \longleftarrow$ improved

- Interior Methods motivated by Karmarkar

  $O(n^3 L)$

Solving

$$\mathbf{Mx} < \mathbf{d} \qquad \begin{cases} \mathbf{M} : m \times n \\ \mathbf{x} \in R^n \\ \mathbf{d} \in R^m \end{cases}$$



$$\mathrm{Vol}(E^2) \leq e^{-1/2(n+1)}\ \mathrm{Vol}(E^1)$$
$$\mathrm{Vol}(E^3) \leq e^{-1/2(n+1)}\ \mathrm{Vol}(E^2)$$
$$\vdots$$

Observations:

(1) There is a half-ellipsoid $\frac{1}{2}E^k$ such that $P \subset \frac{1}{2}E^k$, if $\mathbf{x}^k \notin P$, for $k = 1, 2, \ldots$.

(2) $\text{Vol}(E^{k+1}) \leq e^{-k/2(n+1)}\ \text{Vol}(E^1)$
This means the volume of E^k shirnks very rapidly.

How fast?

If $(E^1) = S(0, 2^{2L})$ then $\text{Vol}(E^1) = \frac{4\pi}{3}e^{6L}$
Given any $\epsilon > 0$, we want to find k such that $\text{Vol}(E^{k+1}) \leq \epsilon$.

Hence, we let

$$(e^{-k/2(n+1)})(\frac{4\pi}{3}e^{6L}) \leq \epsilon$$

i.e.

$$\frac{-k}{2(n+1)} + \ln(\frac{4\pi}{3}) + 6L \leq \ln \epsilon,$$

or

$$-k+2(n+1)\ln(\frac{4\pi}{3})+2(n+1)\,6L \leq 2(n+1)\ln\epsilon$$

Thus

$$2(n+1)[\ \ln(\frac{4\pi}{3}) + 6L - \ln\epsilon\ ] \leq k$$

As long as

$$k \geq 2(n+1)[\ \ln(\frac{4\pi}{3}) + 6L - \ln\epsilon\ ]$$

We know $\text{Vol}(E^{k+1}) \leq \epsilon$

Hence we need only $O(nL - n\ln\epsilon)$ steps!

Basic Ideas:

Step 1: Let $E^0 = S(0; 2^{2L})$ and $k = 0$ be a sufficiently large ellipsoid.

$P := E^0 \cap \{\mathbf{x} \in R^n \mid \mathbf{M}\mathbf{x} < \mathbf{d}\}$.

Step 2: Check if the center $\mathbf{x}^k$ of E^k satisfies $\mathbf{M}\mathbf{x}^k < \mathbf{d}$.

Yes, Stop!

No, remove $\frac{1}{2}E^k$ which does not contain P and replace E^k by a smaller ellipsoid E^{k+1} which contains the other $\frac{1}{2}E^k$.

Step 3: If $\text{Vol}(E^k) < 2^{-(n+1)L}$, then Stop!

The system has no solution.

Otherwise, $k \longleftarrow k + 1$ and go to Step 2.

Observations:

(1) Take $\epsilon = 2^{-(n+1)L}$, since $\ln = \log_e \sim \log_2$. Thus the algorithm takes

$$O(nL - n\log_2 2^{-(n+1)L}) \sim O(n^2 L)$$

iterations to stop.

(2) If each iteration takes $O(n^2 L)$ elamentary operations, then the total complexity of the ellipsoid method becomes

$$O(n^2 L \cdot n^2 L) = O(n^4 L^2)$$

Question:

(1) How to convert LP problems as a system of linear inequalities?

(2) How to construct E^{k+1} from the information of E^k?

MP(I)

(1) *System of Linear Inequalities:*

$$
(P) \quad \begin{array}{ll} \min & \mathbf{c}^T\mathbf{x} \\ \text{s. t.} & \mathbf{A}\mathbf{x} \le \mathbf{b} \\ & \mathbf{x} \ge 0 \end{array}
\qquad\qquad
(D) \quad \begin{array}{ll} \max & \mathbf{b}^T\mathbf{w} \\ \text{s. t.} & \mathbf{A}^T\mathbf{w} \ge \mathbf{c} \\ & \mathbf{w} \ge 0 \end{array}
$$

(i) (P) is solvable if and only if

$$
\begin{aligned}
\mathbf{A}\mathbf{x} \le \mathbf{b} &\quad \mathbf{x} \ge 0 \\
\mathbf{A}^T\mathbf{w} \ge \mathbf{c} &\quad \mathbf{w} \ge 0 \\
-\mathbf{c}^T\mathbf{x} + \mathbf{b}^T\mathbf{w} \le 0 &
\end{aligned}
$$

$$
\begin{pmatrix} A & 0 \\ -I & 0 \\ 0 & -A^T \\ 0 & -I \\ -c^T & b^T \end{pmatrix} \begin{pmatrix} x \\ w \end{pmatrix} \le \begin{pmatrix} b \\ 0 \\ -c \\ 0 \\ 0 \end{pmatrix}.
$$

has a solution $(\mathbf{x}, \mathbf{w})$.

(ii) **Lemma 5.5:** Suppose

$$\mathbf{Mx} < \mathbf{d} + 2^{-L} \underline{\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}} \swarrow \text{perturbation}$$

has a solution, then $\mathbf{Mx} \leq \mathbf{d}$ has a solution

(2) *Constructing E_{k+1} from E_k:*

$$\text{In } R^n, \ \mathbf{z} \in R^n, \ r \geq 0$$

(i) Spheroid:

$$\begin{aligned} S(\mathbf{z}, r) &= \{\mathbf{x} \in R^n \mid \textstyle\sum_{j=1}^{n} (x_j - z_j)^2 \leq r^2\} \\ &= \{\mathbf{x} \in R^n \mid (\mathbf{x} - \mathbf{z})^T (\mathbf{x} - \mathbf{z}) \leq r^2\} \end{aligned}$$

$\mathbf{z}$: center, r: radius, $\text{Vol}(S)$: volume of S.

(ii) Affine transformation:

$\mathbf{A}$: nonsingular $n \times n$ matrix, $\mathbf{c} \in R^n$

$$\begin{aligned} T(\mathbf{A}, \mathbf{c}) : \quad & R^n \longrightarrow R^n \\ & \mathbf{x} \longrightarrow \mathbf{A}(\mathbf{x} - \mathbf{c}) \end{aligned}$$

$T(\mathbf{A}, \mathbf{c})$ is one-to-one and onto!

(iii) Ellipsoid:

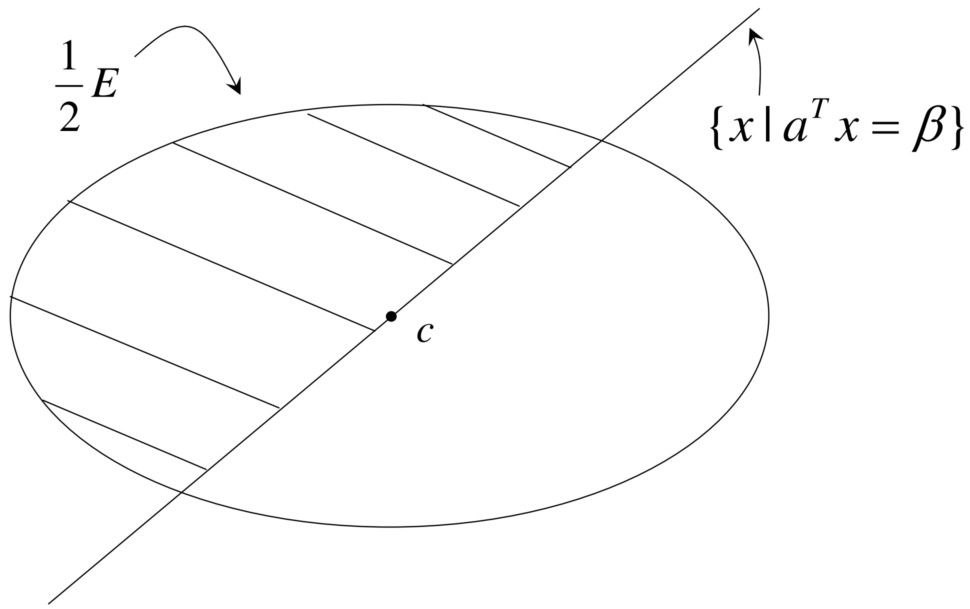
An image of $S(0,1)$ under $T(\mathbf{A},\mathbf{c})$.

$$\begin{aligned} S(0,1) &= \{\mathbf{x} \in R^n \mid \mathbf{x}^T\mathbf{x} \le 1\} \\ E &= \{\mathbf{x} \in R^n \mid [\mathbf{A}(\mathbf{x}-\mathbf{c})]^T[\mathbf{A}(\mathbf{x}-\mathbf{c})] \le 1\} \\ &= \{\mathbf{x} \in R^n \mid (\mathbf{x}-\mathbf{c})^T\mathbf{A}^T\mathbf{A}(\mathbf{x}-\mathbf{c}) \le 1\} \end{aligned}$$

center: $\mathbf{c}$

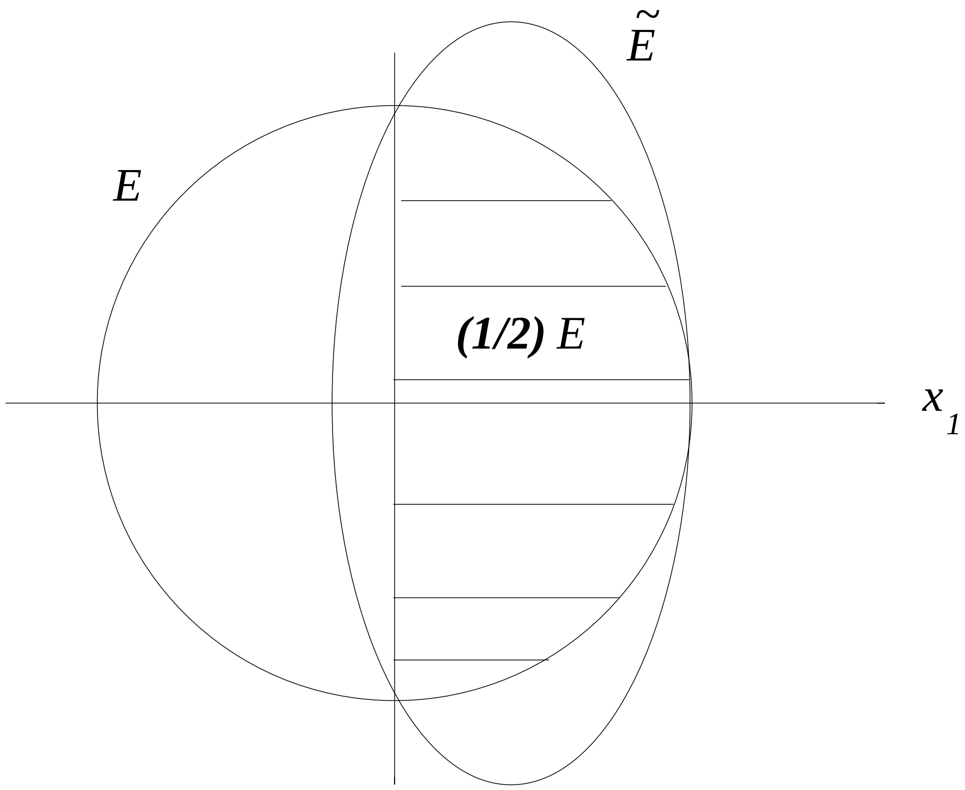
$\mathrm{Vol}(S) = \det(\mathbf{A}^{-1}) \times \mathrm{Vol}(S(0,1))$.

(iv) Half-ellipsoid:



$$\frac{1}{2}E = \{x \text{ in } E \mid a^T x \ge a^T c\}$$

(v) A small ellipsoid containing $\frac{1}{2}E$:



$E = S(0,1)$

$\frac{1}{2}E = \{\mathbf{x} \in E \mid x_1 \geq 0\}$

$\tilde{E} = \{\mathbf{x} \in R^n \mid$

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

$$\text{center: } \mathbf{c} = \begin{pmatrix} \frac{1}{n+1} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\text{scaling: } \mathbf{A} = \begin{pmatrix} \frac{n+1}{n} & & & 0 \\ & \sqrt{\frac{n^2-1}{n^2}} & & \\ & & \ddots & \\ 0 & & & \sqrt{\frac{n^2-1}{n^2}} \end{pmatrix}$$

$$\begin{aligned} \text{Vol}(\tilde{E}) &= \det(\mathbf{A}^{-1}) \times \text{Vol}(E) \\ &= (\tfrac{n}{n+1})(\tfrac{n^2-1}{n^2})^{\frac{n-1}{2}} \times \text{Vol}(E) \\ &\leq e^{-1/2(n+1)} \times \text{Vol}(E) \quad (n > 1) \end{aligned}$$

(vi) Given that

$$E_k = \{\mathbf{x} \in R^n \mid (\mathbf{x} - \mathbf{x}^k)^T \underbrace{A_k^T A_k}_{B_k^{-1}\text{ : positive definite}} (\mathbf{x} - \mathbf{x}^k) \leq 1\}$$

we test $\mathbf{M}\mathbf{x}^k < \mathbf{d}$?
Suppose that $\mathbf{a}_i^T \mathbf{x}^k \geq d_i$ for some i,
then $\mathbf{x}^k \notin P$ and

$$P \subset \frac{1}{2} E_k = \{\mathbf{x} \in E_k \mid -\mathbf{a}_i^T \mathbf{x} \geq -\mathbf{a}_i^T \mathbf{x}^k\}$$

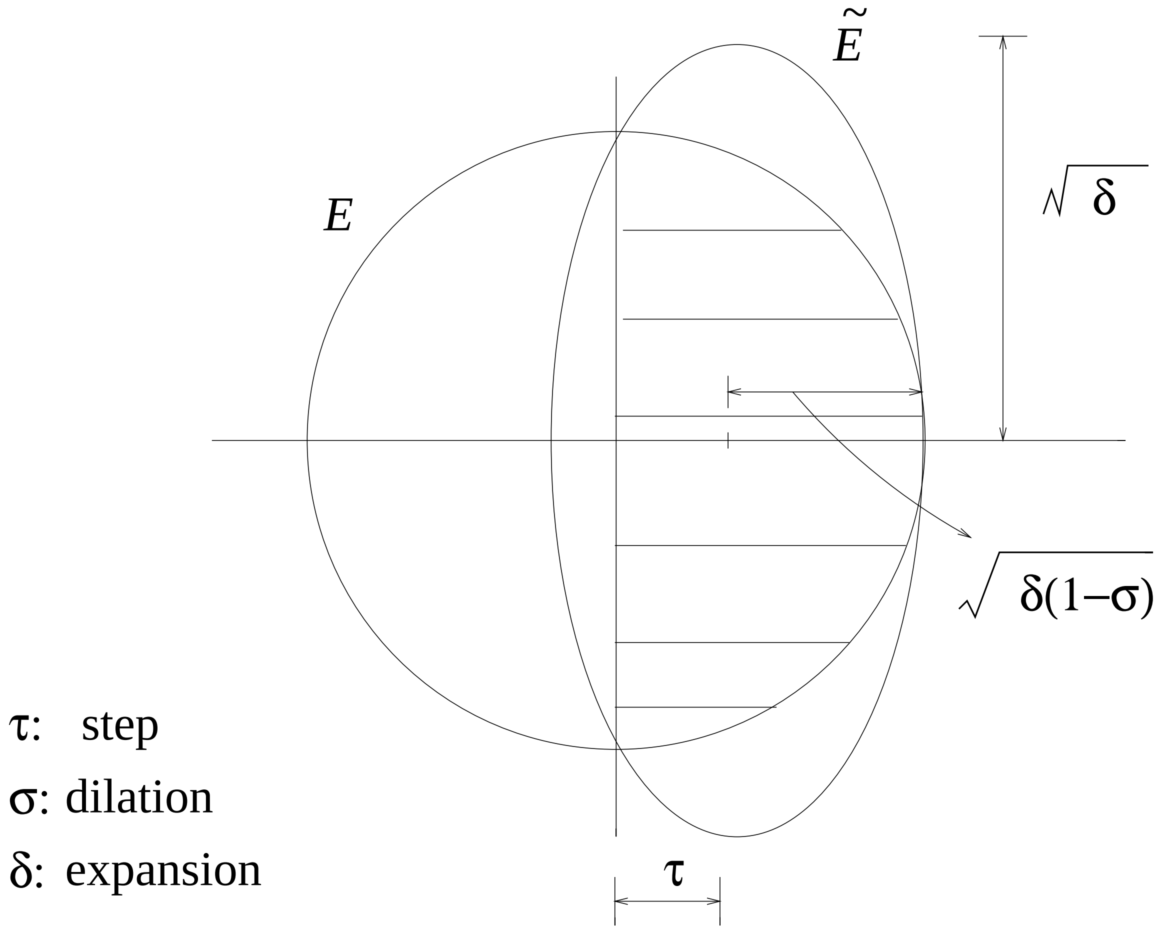
Mapping E_k back to $S(0,1)$, we can derive

$$E_{k+1} = \tilde{E} =$$
$$\{\mathbf{x} \in R^n \mid (\mathbf{x} - \mathbf{x}^{k+1})^T \mathbf{B}_{k+1}^{-1} (\mathbf{x} - \mathbf{x}^{k+1}) \leq 1\}$$

where $\mathbf{x}^{k+1} = \mathbf{x}^k - \tau(\mathbf{B}_k \mathbf{a}_i / \sqrt{\mathbf{a}_i^T \mathbf{B}_k \mathbf{a}_i})$,
$\tau = 1/(n+1)$ and

$$\mathbf{B}_{k+1} = \delta(\mathbf{B}_k - \sigma[\mathbf{B}_k \mathbf{a}_i (\mathbf{B}_k \mathbf{a}_i)^T / (\mathbf{a}_i^T \mathbf{B}_k \mathbf{a}_i)])$$

where $\delta = n/\sqrt{n^2 - 1}$ and $\sigma = 2/(n+1)$.



$\tilde{E}$
E
$\sqrt{\delta}$
$\sqrt{\delta(1-\sigma)}$
τ: step
σ: dilation
δ: expansion
τ

Ellipsoid Method for LP:

Step 1: $k \longleftarrow 0$

$E^k = S(0, 2^{2L})$

$\mathbf{B}_k = 2^{2L} I$

$\mathbf{x}^k = 0$

Step 2: If $\mathbf{M}\mathbf{x}^k < \mathbf{d}$ then STOP!

$\mathbf{x}^k$ is a solution.

Otherwise, find i such that $\mathbf{a}_i^T \mathbf{x}^k \geq d_i$

Calculate

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \tau(\mathbf{B}_k \mathbf{a}_i / \sqrt{\mathbf{a}_i^T \mathbf{B}_k \mathbf{a}_i}),$$

$$\mathbf{B}_{k+1} = \delta \left(\mathbf{B}_k - \sigma [\mathbf{B}_k \mathbf{a}_i (\mathbf{B}_k \mathbf{a}_i)^T / (\mathbf{a}_i^T \mathbf{B}_k \mathbf{a}_i)] \right)$$

Step 3: If $\mathrm{Vol}(E^k) < 2^{-(n+1)L}$, STOP! $P = \emptyset$

Otherwise, $k \longleftarrow k + 1$

Go To Step 2!

$$\delta = n/\sqrt{n^2 - 1},\ \sigma = 2/(n+1),\ \tau = 1/(n+1)$$

Notes:

1. Lemma 5.3 and lemma 5.4
   $\Rightarrow \text{Vol}(P) \geq 2^{-(n+1)L}$ where
   $P := \{\mathbf{x} \mid \mathbf{M}\mathbf{x} < \mathbf{d}, \mid x_i \mid \leq 2^L, i = 1, \ldots, n\}$
2. Khachian showed that the ellipsoid method for LP terminates within $O(n^2 L)$ iterations with an exact solution.
3. Each iteration takes $O(n^2 L)$ elementary operations for $\mathbf{x}^{k+1}$ and $\mathbf{B}_{k+1}$.
   Hence, the total complexity is $O(n^4 L^2)$.
4. Since the optimal solutions $(\mathbf{x}, \mathbf{w}) \in R^{n+m}$ falls in a lower dimensional hyperplane $\mathbf{c}^T\mathbf{x} - \mathbf{b}^T\mathbf{w} = 0$, the perturbed solution set has a very small volume (≈ 0)
   Hence, the convergence of the ellipsoid method is extremely slow.
5. Since

   $$\mathbf{B}_{k+1} = \delta\left(\mathbf{B}_k - \sigma[\mathbf{B}_k\mathbf{a}_i(\mathbf{B}_k\mathbf{a}_i)^T/(\mathbf{a}_i^T\mathbf{B}_k\mathbf{a}_i)]\right)$$

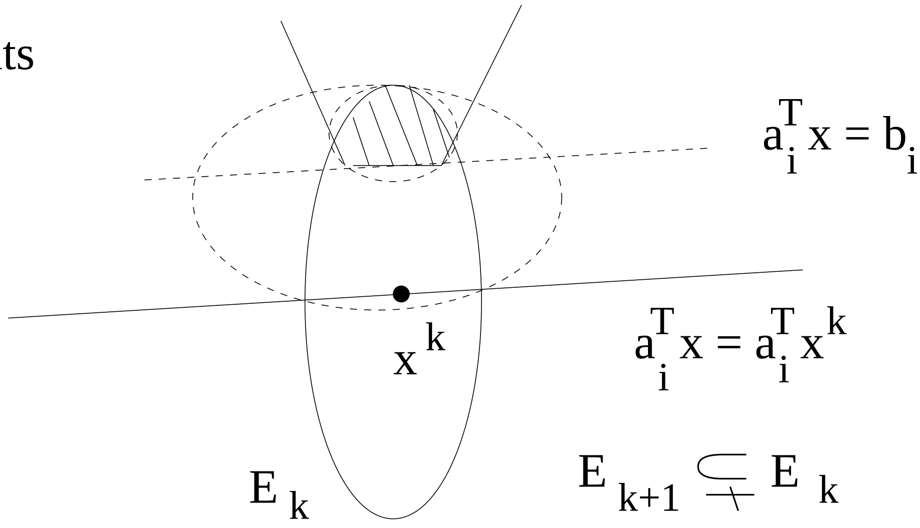
   Even $\mathbf{B}_0$ is sparse, $\mathbf{B}_k$ can be very dense.

Therefore for large scale problems, "sparsity" becomes an issue!

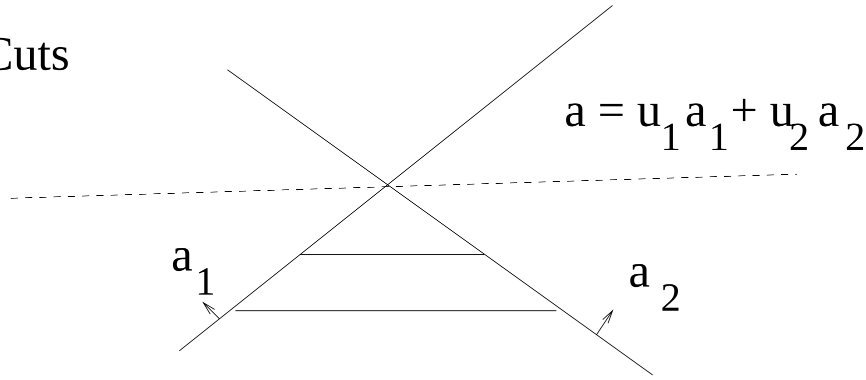
6. After all, Ellipsoid method is
   Good in theory, but bad in pratice.

7. Modification

(a) Deep Cuts



(b) Surrogate Cuts



(c) Parallel Cuts