Mar 21 (Lecture 20)

Solving Linear Programs:

→ Simplex Algorithm (1950's) (not in P)

can work with exponentially many constraints → Ellipsoid Algorithm (1970's)

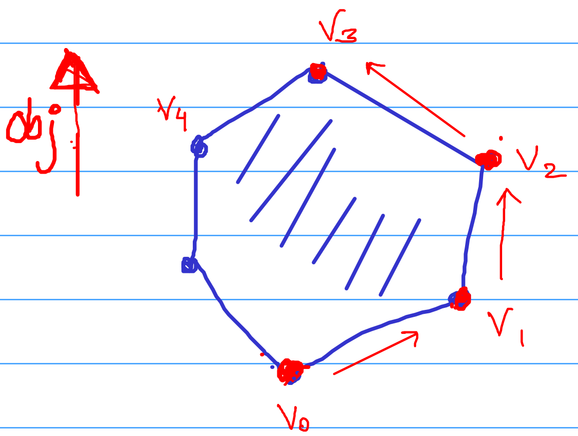
→ Interior point methods (1980's)

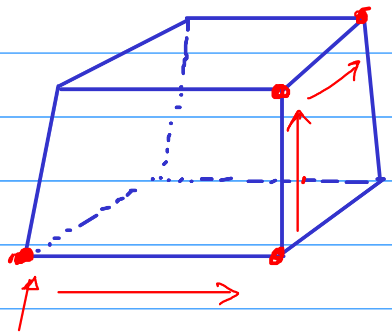
→ Gradient Descent

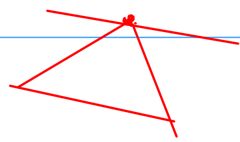
} Convex Programming

Simplex Algorithm:

- Start with a vertex of the polyhedron (also called basic feasible solution)
- At any vertex, look at the neighboring vertices (connected by an edge). Go to one of the neighbors s.t. objective value improves.
- Stop when no improvement possible







n = no. of varibles

$$\begin{cases} a_i^T x = b_i & \text{for } 1 \le i \le k \\ x \ge 0 \end{cases}$$

Vertex : intersection of n tight constraints.

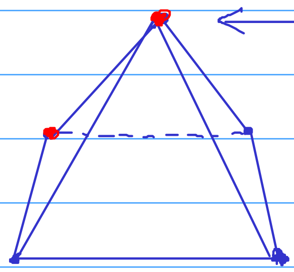
lin Ind

Edge : $n-1$ tight constraints.

Non-degeneracy assumption:

each vertex has exactly n tight constraints.

Important for bounding the running time of Simplex.



Four tight constraints

$n=3$

How to walk along an edge?

Say, we are at a vertex defined by

lin Ind

$$\left[\begin{array}{ll} \{\, a_i^T x = b_i & 1 \le i \le k \\ x_1 = 0 & \\ x_2 = 0 & \\ \vdots & \\ x_{n-k} = 0 & \end{array}\right.$$

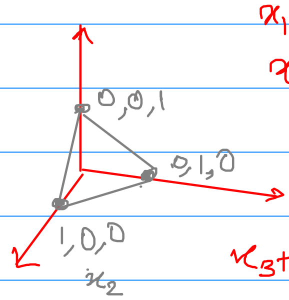
Choose to relax one tight constraint

$Ax = b$ →
$$\begin{cases} a_i^T x = b_i & 1 \le i \le k \\ x_1 \ge 0 & \\ x_2 = 0 & \\ x_3 = 0 & \\ x_{n-k} = 0 & \end{cases}$$

$x_1 \ge 0$

$x_2 \ge 0$

$x_3 \ge 0$



$x_3 + x_1 + x_2 = 1$

Find some value α for x_1 to reach the other end.

Let $A = [\underbrace{A'}_{\substack{n-k \\ \text{Columns}}} \; \underbrace{A''}_{\substack{k \\ \text{Columns}}}]$

$$Ax = A' \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-k} \end{bmatrix} + A'' \begin{bmatrix} x_{n-k+1} \\ \vdots \\ x_n \end{bmatrix} = b$$

$$\begin{bmatrix} x_{n-k+1} \\ \vdots \\ x_n \end{bmatrix} = (A'')^{-1} \left(b - A' \begin{bmatrix} \alpha \\ 0 \\ 0 \\ 0 \end{bmatrix} \right)$$

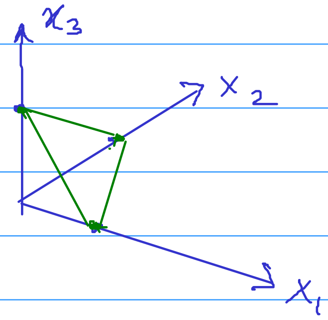
We get $x_{n-k+1}, \dots, x_n$ as a function of α.

Choose α to be the largest so that $x_{n-k+1} \geq 0$
$\vdots$
$x_n \geq 0$

n-k possible edges.
Choose one that improves objective.

Example :

$$\text{Max } 2x_1 + 3x_2 + 4x_3$$
$$x_1 \geq 0$$
$$x_2 \geq 0$$
$$x_3 \geq 0$$
$$x_1 + x_2 + x_3 = 1$$



Claim: If we are not at an optimal point then there must be an edge that strictly improves the objective. HW

$\Rightarrow$ guarantees finiteness of the algorithm.

Running time: no. of vertices. $\leq \binom{n}{k}$

In practice – good.

There is no known rule for selecting edges that guarantees polynomial time.

Open question. to find such a rule.

Good in Practice.

- Finding initial feasible solution ? (Simplex).

WLOG $b \geq 0$

$$Ax = b$$
$$x \geq 0$$

$$\min \sum z_i$$
$$\begin{cases} Ax + z = b \\ x \geq 0 \\ z \geq 0 \end{cases} \qquad \left\{ \begin{array}{l} a_i^T x + z_i = b_i \\ \text{for } 1 \leq i \leq k \end{array} \right.$$

Feasible solution. $x = 0$, $z = b$

Claim: x is a feasible solution for LHS LP iff $(x, z=0)$ is an optimal for RHS LP.

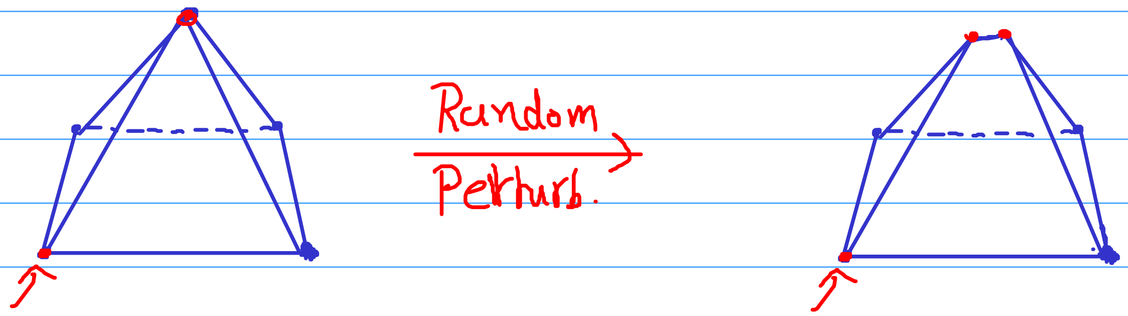
$\leftarrow$ Initial feasible for RHS LP. Find optimal using simplex.

Ensuring Non-degeneracy.

Random Perturbation

$Ax \leq b + \varepsilon$

$$\begin{cases} Ax = b \\ x \geq 0 + \varepsilon \end{cases}$$



After each constraint is randomly perturbed, each vertex will have exactly n tight constraints with high probability.

Note that small perturbation does not affect the objective value by much.

Smoothed Analysis of Simplex [2001]

(A, b) $\qquad Ax \leq b$

$(A, b+\varepsilon) \Leftarrow$ polynomial time

Spielman, Teng 2001 showed that for every input (A, b) if it is randomly perturbed to $(A, b+\varepsilon)$ then with high prob simplex will run in polynomial time on $(A, b+\varepsilon)$

This is a much stronger statement than saying that simplex will take poly time on a random input.

March 24 / Lec 21

Ellipsoid Algorithm

Khachyan 1979 (first polynomial time algorithm for LPs.)

→ Works for more general convex programming (minimize a convex function over a convex set)

→ Recall that Optimization $\leq$ Feasibility.

→ Ellipsoid algorithm solves the feasibility problem.

If there is a feasible solution it will return one.

→ Does not need an explicit description of the constraints. Needs only a separation oracle.

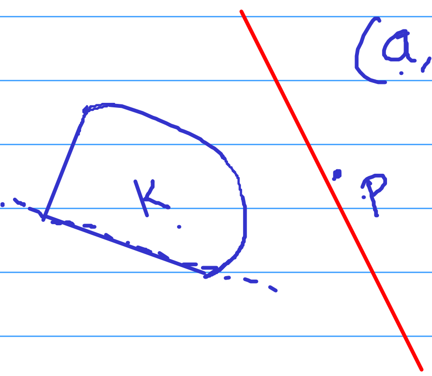
Def: Seperation oracle

Given a convex set K and a point p tells you whether $p \in K$.

If $p \notin K$ then output a separating hyperplane (a, b) s.t. $a^T p > b$

and

$a^T x \leq b$ for all $x \in K$.



→ Separation oracle obvious for a given LP
$Ax \le b$

separating hyperplane ← violated constraint for p.

→ Ellipsoid algorithm can also work for an LP with exp many constraints, if efficient separation oracle can be designed. (e.g., Steiner tree LP).

$$x \geq 0$$

$$\sum_{e \in \delta(S)} x_e \geq 1 \qquad S \text{ is a terminal separating cuts.}$$

check $\min_{S} \sum_{e \in \delta(S)} x_e \geq 1$

Separation oracle for SDPs

$$\begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -3 \end{bmatrix}$$

$4x_{11} - 2x_{12} - 2x_{21} + x_{22} < 0$

$4z_{11} - 2z_{12} - 2z_{21} + z_{22} \geq 0$ for any PSD Z

$$\left.\begin{aligned} \to\ & X \succeq 0 \\ & \sum_{ij} a_{ij} x_{ij} = b \end{aligned}\right\}$$

If X is not PSD then find some (Y_{ij}, δ) s.t. $\sum Y_{ij} X_{ij} > \delta$

$\sum Y_{ij} Z_{ij} \leq \delta$ for any PSD Z.

$$X = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$$

$Z_{22} \geq 0$ for all PSD Z

$X_{22} < 0$

Separating Hyperplane theorem:

Given a closed convex set $K \subseteq \mathbb{R}^n$, a point $p \in \mathbb{R}^n$

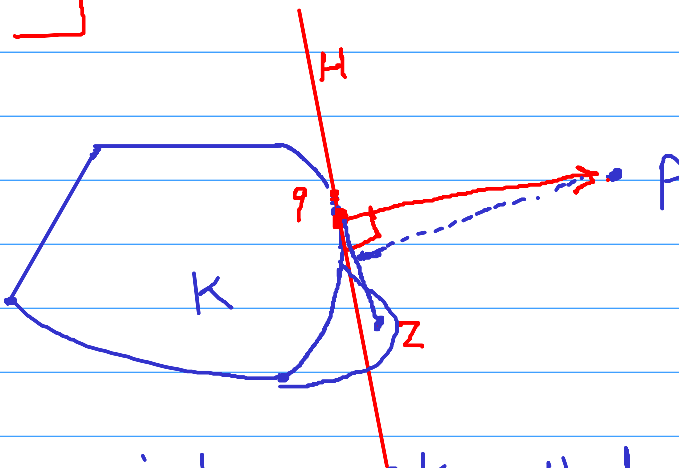
If $p \notin K$ then there exists a separating hyperplane $a \in \mathbb{R}^n$, $b \in \mathbb{R}$ s.t. $a^T p > b$

and

$a^T x \leq b$ for all $x \in K$.

$K = \{ z : 0 \leq z \leq 1 \}$
$p = (2)$

Proof



assume K is closed

Take a point $q \in K$ that is closest to p.

That is,

$$\text{dist}(q, p) \leq \text{dist}(x, p) \quad \forall\ x \in K$$

Consider the hyperplane passing through q that is orthogonal to $p - q$.

H $\qquad (p-q)^T x = (p-q)^T q$

p is on one side of H

$$(p-q)^T p > (p-q)^T q$$

because $(p-q)^T (p-q) > 0$

($\because$ p and q are distinct

To show

$\forall\ x \in K \qquad (p-q)^T x \le (p-q)^T q$

For the sake of contradiction,

let's take point $z \in K$ but

$(p-q)^T z > (p-q)^T q \Rightarrow (z-q)^T (p-q) > 0$

Take a point y on the line segment joining (z, q)

$y = (1-\lambda) q + \lambda z \qquad \lambda \ge 0$

To show

dist $(y, p) <$ dist $(q, p) \leftarrow$ contradiction.

$(y-p)^T (y-p)$

$= \left[(1-\lambda) q + \lambda z - p\right]^T \left[(1-\lambda) q + \lambda z - p\right]$

$= [q-p]^T [q-p] + [-\lambda q + \lambda z]^T [-\lambda q + \lambda z] + 2 [-\lambda q + \lambda z]^T [q-p]$

$= \text{dist}(q,p) + \left[\lambda^2 \underbrace{(z-q)^T (z-q)}_{>0} + 2\lambda \underbrace{(z-q)^T (q-p)}_{<0}\right] < 0 \,?$

If you choose small enough λ then $\lambda^2 << \lambda$ and the second term will dominate.

Then the overall sum will be negative.

Corollary: Every convex set can be described by a set of linear constraints, (intersection of Halfspaces) possibly infinitely many.

Separating Hyperplane theorem can be used to prove

Farkas' Lemma

Strong LP Duality.

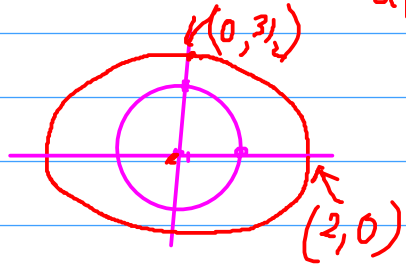
Ellipsoid Algorithm

(Ellipsoid)

Higher dimensional analogue of an ellipse.

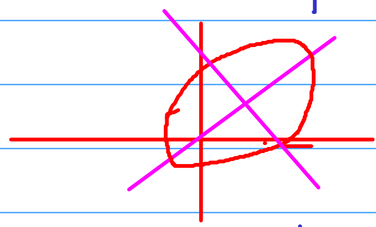
Sphere $x_1^2 + x_2^2 + \dots + x_n^2 \leq 1$ center origin

Ellipsoid $\dfrac{x_1^2}{a_1^2} + \dfrac{x_2^2}{a_2^2} \dots + \dfrac{x_n^2}{a_n^2} \leq 1$ Center origin



$$\frac{x_1^2}{2^2} + \frac{x_2^2}{(3/2)^2} \leq 1$$

axis - parallel



Any ellipsoid can be obtained from a sphere by

(1) stretching the axes

(2) rotation

(3) shifting the center.

$$S(0,1) \equiv \{ x : x^T x \leq 1 \}$$

$x^T x = \|x\|_2^2$

Ellipsoid $\rightarrow$ $\{ M^{-1}x + c : x^T x \leq 1 \}$

for some invertible matrix M and $c \in \mathbb{R}^n$

$y := M^{-1}x + c$

$x = M(y - c)$

Equivalently
$$\{ y \in \mathbb{R}^n : (y-c)^T M^T M (y-c) \leq 1 \}$$

Let $A = M^T M$ (a positive definite matrix)

Def (Ellipsoid)

$A \succ 0$

For a PD matrix A, and $c \in \mathbb{R}^n$
$$E(c, A) = \{ y \in \mathbb{R}^n : \underbrace{(y-c)^T A (y-c)} \leq 1 \}.$$

Ellipsoid Algorithm:

Input: a convex set K with a separation oracle.

Output: a point in K if K is nonempty
or say that K is empty.

Promise:

(1) bounded: $\exists\ R > 0$ s.t. $K \subseteq S(0, R)$

(2) Volume lower bound: $\exists\ r > 0$ s.t. $S(r) \subseteq K$ if K is non-empty.

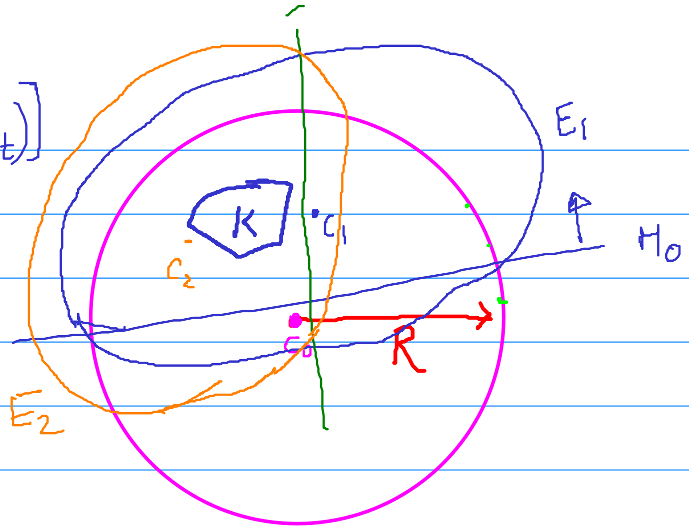
($\Rightarrow$ full dimensional)

Running time: $O\left(n^2 \log \frac{R}{r}\right)$ $\quad n \leftarrow \dim$
(no. of iterations).

$K \subseteq E_t$

K above H_t

$E_{t+1} \supseteq [E_t \cap (\text{above } H_t)]$



Algorithm outline:

$E_0 = S(0, R)$

Invariant: At round t, we will have an ellipsoid E_t with $K \subseteq E_t$

- Check if center c_t of E_t lies in K

  if yes then done.

- if No then get a hyperplane separating c_t from K.

  Say, $a^T x \leq b$ for every $x \in K$
  and $a^T c_t > b$.

Find a new Ellipsoid E_{t+1} containing the intersection of E_t and the halfspace.

$$E_{t+1} \supseteq \{ x \in E_t : a^T x \leq b \}$$

Clearly, $K \subseteq E_{t+1}$

Continue while $\mathrm{Vol}(E_t) \geq \mathrm{Vol}(S(0,r))$

If $\mathrm{Vol}(E_t)$ becomes $< \mathrm{Vol}(S(0,r))$

$\Rightarrow K$ is empty.

To bound the running time, we will show that the volume of the ellipsoid decreases by a fraction in every round.

$\mathrm{Vol}(K) \geq \mathrm{Vol}(S(r)) > \mathrm{Vol}(E_t)$

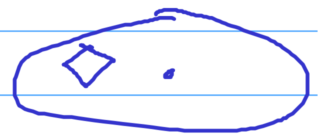
$K \subseteq E_t$

Que: Could we use a sequence of spheres instead of Ellipsoids?
or
sequence of Polytopes?

Mar 28, Lecture 22

Finding E_{t+1} containing $E_t \cap \{x : a^T x \leq b\}$

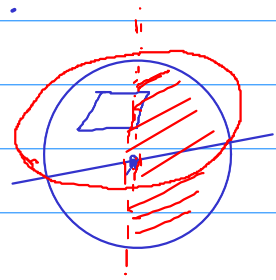
simplest case:



$$E_t = S(0,1) \qquad x_1^2 + x_2^2 + \cdots + x_n^2 \leq 1$$

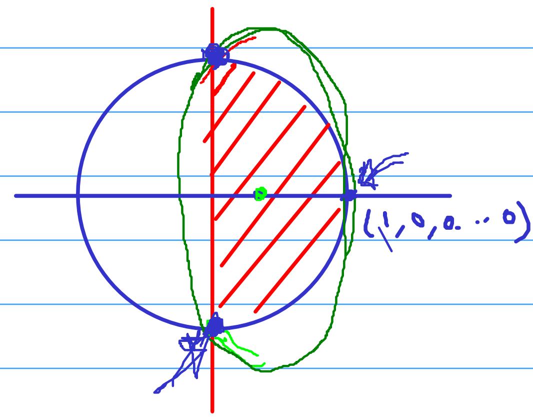
for simplicity, we will always take the separating hyperplane passing through the center.

Let the hyperplane be



$$x_1 \geq 0$$

Find an ellipsoid E_{t+1} containing the half sphere $S(0,1) \cap x_1 \geq 0$



Ellipsoid E_{t+1} center $\rightarrow$ $(\alpha, 0, 0, 0, \cdots, 0)$

$$\rightarrow \frac{(x_1-\alpha)^2}{\beta^2} + \frac{x_2^2}{\gamma^2} + \frac{x_3^2}{\gamma^2} + \cdots + \frac{x_n^2}{\gamma^2} \leq 1$$

boundary should passes through

$\longrightarrow$ $1, 0, 0, \cdots, 0$ ✓

$\longrightarrow$ $\{x : \sum x_i^2 = 1, \; x_1 = 0\}$ $(0, 1, 0, 0, 0)$ ✓

$$\frac{(1-\alpha)^2}{\beta^2} = 1 \qquad \left(\frac{-\alpha}{\beta}\right)^2 + \frac{1}{\gamma^2} = 1$$

$$\boxed{\beta^2 = (1-\alpha)^2} \qquad \boxed{\gamma^2 = \frac{1}{1 - \frac{\alpha^2}{(1-\alpha)^2}} = \frac{(1-\alpha)^2}{1-2\alpha}}$$

verify that for any such α, β, γ,

HW E_{t+1} contains the half-sphere

$$\left\{ (x_1, x_2, \cdots x_n) : x_1 \geq 0 \text{ and } \sum_{i=1}^{n} x_i^2 \leq 1 \right\}$$

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} \le 1$$

πab

$\mathrm{Vol}(S_n) \leftarrow c_n r^n$

Volume of $E_{t+1} = c_n \cdot \beta \cdot \gamma^{n-1}$

$$= c_n \cdot \frac{(1-\alpha)^n}{(1-2\alpha)^{\frac{n-1}{2}}}$$

Minimize $\frac{(1-\alpha)^n}{(1-2\alpha)^{\frac{n-1}{2}}}$

$\alpha \geq 0$

$$\Rightarrow \quad \alpha = \frac{1}{n+1}$$

Center $\left(\frac{1}{n+1}, 0, 0, 0, \ldots 0\right)$

E_{t+1} $\quad \beta = \frac{n}{n+1} \quad \gamma^2 = \frac{n^2}{(n+1)^2} \cdot \frac{n+1}{(n-1)} = \frac{n^2}{n^2-1}$

$$\frac{\left(x_1 - \frac{1}{n+1}\right)^2}{\left(\frac{n}{n+1}\right)^2} + \frac{x_2^2}{n^2/n^2-1} + \cdots + \frac{x_n^2}{n^2/n^2-1} \le 1.$$

Bounding $\quad \mathrm{Vol}(E_{t+1}) \,/\, \mathrm{Vol}(E_t)$

$$= \frac{c_n \beta \gamma^{n-1}}{c_n \cdot 1^n} = \beta \gamma^{n-1}$$

$$\frac{n}{n+1} \cdot \left(\frac{n^2}{n^2-1}\right)^{\frac{n-1}{2}} = \left(1 - \frac{1}{n+1}\right) \cdot \left(1 + \frac{1}{n^2-1}\right)^{\frac{n-1}{2}}$$

$\boxed{\text{Use } \quad 1+y \le e^y \quad \forall y \in \mathbb{R}}$

$$\le e^{-1/n+1} \cdot \left[e^{1/n^2-1}\right]^{\frac{n-1}{2}}$$

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

≤ 1

$$e^y = 1 + y + \underbrace{\frac{y^2}{2!} + \frac{y^3}{3!}}_{\geq 0}$$

If in every round we have volume decrease by a factor of $e^{-1/2n+2}$

then in K rounds

$$e^{-K/2n+2} \cdot \mathrm{vol}(S(R)) = \mathrm{vol}(S(r))$$

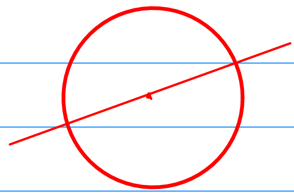
$$e^{-K/2n+2} \cdot c_n R^n = c_n r^n$$

number of rounds $= (2n+2)n \log\left(\frac{R}{r}\right)$

General case:

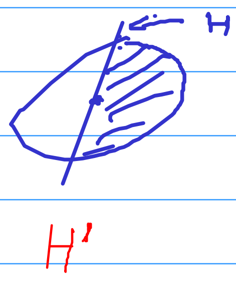
Arbitrary Ellipsoid and hyperplane.

(1) $S(0,1)$ and $\underline{a}^T x \geq 0$



Just need to rotate the above computed ellipsoid appropriately.

(2) Arbitrary ellipsoid



$$\begin{array}{ccc} E_t, H & \xrightarrow{T} & S(0,1), H' \\ & & \downarrow \\ E_{t+1} & \xleftarrow{T^{-1}} & E'_{t+1} \supseteq (S(0,1) \cap H'_2) \end{array}$$

Claim Containment, and ratio of volumes invariant under Linear Transformation

Examples

$$\frac{x_1^2}{4} + \frac{x_2^2}{9} \leq 1$$

$$\begin{bmatrix} 1/2 & 0 \\ 0 & 1/3 \end{bmatrix} \longrightarrow$$

$y_1 = x_1/2$

$y_2 = x_2/3$

$y_1^2 + y_2^2 \leq 1$

$$(x_1+x_2)^2 + (x_1-x_2)^2 \leq 1$$

$y_1 = x_1 + x_2$

$y_2 = x_1 - x_2$

$x_1 = \frac{y_1+y_2}{2}$

$x_2 = \frac{y_1-y_2}{2}$

$$(x_1-\alpha)^2 + (x_2-\beta)^2 \leq 1$$

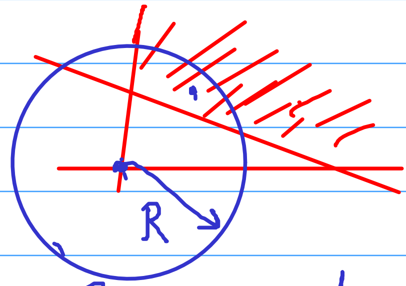
$\longrightarrow y_1 = x_1 - \alpha \qquad y_2 = x_2 - \beta$

Solving an LP via Ellipsoid.

Que: is $Ax \le b$ feasible.

Bounding R and r for the convex set

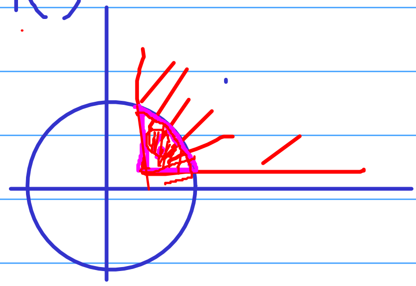
$Ax \le b$



An LP can be unbounded. So cannot bound R.

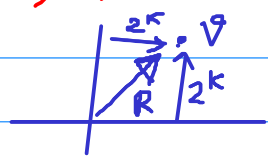
It is good enough if we can guarantee that if $Ax \le b$ is feasibe then there is at least one point in

$$\{x : Ax \le b\} \cap S(0,R)$$



$n \leftarrow$ no. of var
$k \leftarrow$ no. of const
$\ell \leftarrow$ bit size of coeff

Just take $K = \{x : Ax \le b\} \cap S(0,R)$



say, A is $k \times n$.

Entries are rational numbers with ℓ bits

To show: $R \le 2^{poly(k,n,\ell)}$ $\qquad \log R = poly(k,n,\ell)$

Claim: Any vertex has coordinates with at most $poly(k,n,\ell)$ bits. $\qquad R \le 2^k \sqrt{n}$

$\boxed{\begin{array}{l} A'x = b \\ x = A'^{-1} b \end{array}}$

Thus, we can bound $R \le 2^{poly(k,n,\ell)}$

$Ax \le b$
$x_1 = 1$

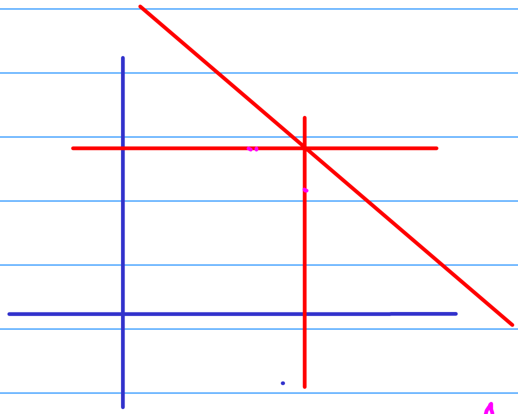
Lower bounding the volume.

① If not Full dimensional (non zero volume)
→ perturb constraints by small amount

$$a_j^T x \le b_j + \varepsilon$$

<u>Example</u>

$x \le 1$
$y \le 1$
$x + y \ge 2$



$x \le 1$
$y \le 1$
$x + y \ge 2$

If $Ax \le b$ non-feasible $\Rightarrow$ $Ax \le b + \varepsilon$ non-feasible.

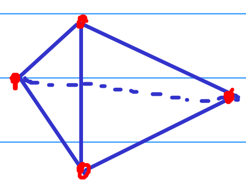
② volume lower bound

 $\ge 2^{-poly(n, |\ell|)}$

co-ordinates of <u>vertices</u> are like P/q
where $q \le \underbrace{2^{poly(m,n)}}$

Let us lower bound the volume
when there are exactly <u>$n+1$</u> vertices.

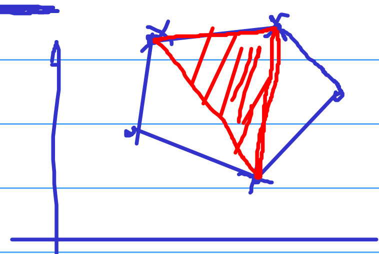
$2x_1 \le x_n$
$2x_2 \le x_n$ $\quad 0 \le x_n \le 1$
$2x_{n-1} \le x_n$



2-simplex

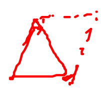
3-simplex

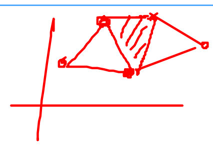
n-dim



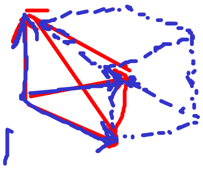
<u>n-simplex</u>

$\frac{1}{2} \det \left[v_1 - v_0 , v_2 - v_0 \right]$





Let vertices be $v_0, v_1, v_2, v_3, \dots, v_n$

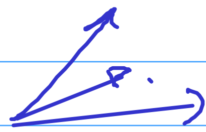


$\frac{1}{6}$

$$\text{volume}\left(\text{parallelopiped}\left(v_1 - v_0,\ v_2 - v_0,\ v_3 - v_0,\ \dots,\ v_{n+1} - v_0\right)\right)$$

$$= (n!)\ \text{volume}\left(\text{simplex}\left(v_0, v_1, v_2, \dots, v_{n+1}\right)\right)$$

HW



vol (parllelopiped)

$$= \det \begin{bmatrix} v_1 - v_0 & v_2 - v_0 & \cdots & v_n - v_0 \end{bmatrix} \geq q^{-n}$$

HW

$\geq 2^{-poly}$

If the largest denominator of any entry is q the $\det \geq q^{-n}$

$$\text{Volume (simplex)} \geq 2^{-poly(m,n,b)}$$

$r = \frac{1}{2^{+poly}}$

No. of iterations in Ellipsoid

$$= O\left(n \log \frac{R^n}{\text{vol}\cdot LB}\right)$$

$$= O\left(poly(m, n, b)\right)$$

Unfortunately the poly( ) is too large to be practical.