Problems / Subroutines

- Linear Systems $(Ax = b)$
- Linear Programming
- Semidefinite Programming (and Cholesky Decomposition)
- Eigenvalue Computation.

– Linear systems

- Gaussian Elimination can solve in time $O(n^3)$ ← row ops do not change the set of solutions.
  and bit complexity remains $\overset{\text{poly-}}{\text{bounded}}$.
- Cramer's rule says size of solutions is $\overset{\text{poly-}}{\text{bounded}}$

– LPs:

- Write LP in equational form $\max(c^T x \mid \underbrace{Ax = b, x \geq 0}_{K \subseteq \mathbb{R}^n})$ $A \in \mathbb{Q}^{m \times n}$

Then: if K is non empty and the optimal value is bounded, then

$\exists$ a basic feasible solution that achieves the optimum x^*

i.e. $\exists n$ tight LI constraints that define x^*.

so $\exists$ $\underset{\text{basis}}{B} \subseteq [m+n]$ st $\underset{\text{submatrix of } \left[\begin{smallmatrix} A \\ I \end{smallmatrix}\right] \text{ indexed by } B}{\begin{bmatrix} A \\ I \end{bmatrix}_B} x^* = \begin{bmatrix} b \\ 0 \end{bmatrix}_B$. ← subvector of $\left[\begin{smallmatrix} b \\ 0 \end{smallmatrix}\right]$ indexed by B

Again, Cramer's rule says x^* has $\overset{\text{poly}}{\text{bounded}}$ bit complexity,

being ratio of two determinants of the input. Shows that $LP \in NP$.

- ~~[illegible]~~ Eigenvalues/vectors.

These may be irrational! $\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ has eigenvalues $\pm\sqrt{2}$

So cannot write these down exactly!

- SDPs

Two kinds of problems.

- Captures eigenvalue problems — e.g. $\min t$ st $tI - A \succeq 0$

computes the λ_{max}. So solutions may be irrational.

- But even when solutions are integral, problems.

The bit-complexity of optimal solution may be exponential in inputs.

$$\min \left\{ x_n : \begin{pmatrix} 1 & 2 & & & & & \\ 2 & x_1 & & & & & \\ & & 1 & x_1 & & & \\ & & x_1 & x_2 & & & \\ & & & & 1 & x_2 & \\ & & & & x_2 & x_3 & \\ & & & & & & \ddots \end{pmatrix} \succeq 0 \right\}$$

Then want $\begin{pmatrix} 1 & 2 \\ 2 & x_1 \end{pmatrix} \succeq 0$ $\quad \begin{pmatrix} 1 & x_i \\ x_i & x_{i+1} \end{pmatrix} \succeq 0$ $\Rightarrow x_{i+1} \geq x_i^2$. $x_1 \geq 4$

$\Rightarrow x_t \geq 2^{2^t}$. ☹

In general with SDPs, this is bad news, so must also argue that the solutions we want have poly-bounded bit complexity

and that we can afford to work with approximations to them.

Bit Complexity of Gaussian Elimination

$A^{(k)} = k^{th}$ matrix during G.E., $A^{(0)} = A$

$a^{(k)}_{ij} = ij^{th}$ entry of $A^{(k)}$.

$$\begin{pmatrix} a_{11}^{(k-1)} & \cdots & & & \cdots \\ 0 & \ddots & & & \\ \vdots & & a^{(k-1)}_{k-1,k-1} & \cdots & \cdots \\ & & 0 & a^{(k-1)}_{kk} & \cdots \\ & & \vdots & \vdots & \ddots \\ 0 & & 0 & \cdots & \end{pmatrix}$$

$$a^{(k)}_{ij} = a^{(k-1)}_{ij} - a^{(k-1)}_{kj} \cdot \frac{a^{(k-1)}_{i,k}}{a^{(k-1)}_{kk}}$$

for $i \geq k+1, j \geq k$

(*)
$$\boxed{a^{(k)}_{ij} = \frac{a^{(k-1)}_{ij} a^{(k-1)}_{kk} - a^{(k-1)}_{kj} a^{k-1}_{ik}}{a^{(k-1)}_{kk}}} = \frac{\text{det of the } 2\times 2 \text{ matrix} \begin{pmatrix} a^{k-1}_{kk} & a^{k-1}_{kj} \\ a^{k-1}_{ik} & a^{k-1}_{ij} \end{pmatrix}}{a^{(k-1)}_{kk}}.$$

Claim: can be represented using few bits.

(**)
$$\boxed{\text{Thm 1: } a^{(k)}_{ij} = \frac{\det(D^{(k)}_{ij})}{\det(D^{(k)})}. \quad \text{for } i,j \geq k+1}$$

where $D^{(k)} = A^{(0)}\big|_{\{1..k\}\times\{1..k\}}$

$D^{(k)}_{ij} = A^{(k)}\big|_{\{1..k,i\}\times\{1...k,j\}}$

Fact 2: Both these matrices have small determinants.

$\Rightarrow$ Putting the two together, get that numbers blow up only to a limited extent (i.e. bit complexity of numbers is bounded by $n \times$ orig. bit complexity).

Pf of Thm 1: $D^{(k)}_{ij} = \begin{pmatrix} a_{11}^{(k)} & & \cdots & \\ & \ddots & & \vdots \\ & & a^{(k)}_{kk} & \vdots \\ 0 & & & a^{(k)}_{ij} \end{pmatrix} \Rightarrow$ det is product of diagonals.

$D^{(k)} = \begin{pmatrix} a_{11}^{(k)} & \cdots & \cdots \\ & \ddots & \vdots \\ 0 & & a^{(k)}_{kk} \end{pmatrix} \Rightarrow$ ratio of dets is exactly $a^{(k)}_{ij}$.

Pf of Fact 2: $D_{ij}^{(k)}$ obtained from $D_{ij}^{(0)}$ by row operations

$D_{(1..k)}^{(k)}$ - . . . - $D_{(1..k)}^{(0)}$ ——

both submatrices of A $\Rightarrow$ have bounded determinants by the Hadamard bound, say

that $\det(A) \leq \|a_1\| \cdot \|a_2\| \cdot \ldots \|a_n\|$.

volume of parallelopiped — product of sides.

—— x ——

Note: $\det(D^{(k)}) = \det(D_{kk}^{(k-1)})$

define $\Delta_{ij}^{k} = \det(D_{ij}^{(k)})$

so. can use (*) & (**) on previous page to set

$$\boxed{\Delta_{ij}^{k+1} = \frac{\Delta_{ij}^{k}\,\Delta_{k+1,k+1}^{k} - \Delta_{i,k+1}^{k}\,\Delta_{k+1,j}^{k}}{\Delta_{k,k}^{k-1}}}$$

and this a determinant of some submatrix of A, so integer (!) can

use this as recurrence to work completely within the set of integers.

finally use (*) to get back $a_{ij}^{(n)}$.

—— x ——

BTW: if $A = \begin{pmatrix} 1 & & & 0 \\ -1 & 1 & & \\ -1 & -1 & 1 & \\ \vdots & & \ddots & \ddots \\ -1 & \cdots & -1 & 1 \end{pmatrix}$ then get $\begin{pmatrix} 1 & & & & 1 \\ & 1 & & & 2 \\ & & 1 & & 4 \\ & 0 & & 1 & 8 \\ & & & & 16 \end{pmatrix}$. so

numbers do get larger.

BTW 2: If not super comfortable with this, check out Trefethen-Bau for basics, the Lovasz-Gacs notes for this Edmonds proof, etc.

Aside: Can also do modular arithmetic and Chinese Remaindering, which has its own issues.

Solving LPs:

Two poly time approaches

- Ellipsoid type algorithms
- Interior Point algorithms.

Basic Subroutine: $\boxed{\text{Feasibility.}}$

Given convex compact set $K \subseteq \mathbb{R}^n$,

find a point $x \in K$ or show that $vol(K) \le \varepsilon$.

How is this set given?

"strong" separation oracle.

Via a $\boxed{\text{separation oracle:}}$

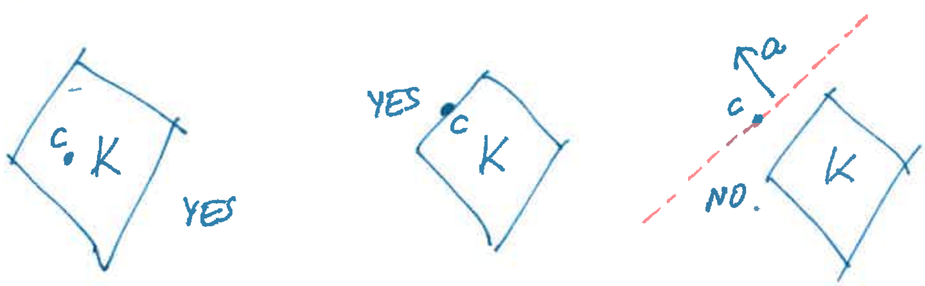
Input: a point $c \in \mathbb{R}^n$ c for "check" or "center"

Output: YES if $c \in K$.

NO if $c \notin K$. In this case output a hyperplane separating from K.

i.e. a vector a, $a \neq 0$

s.t. $K \subseteq \{x \in \mathbb{R}^n \mid a^T x < a^T c\}$



In general can handle even weak separations where tolerate errors of $\varepsilon > 0$,

but let's not go there for now.

See [GLS80] for many details.

General Approach: "Binary Search"

~~Suppose know that~~

~~(1) $K \subseteq Ball(0, R)$ for some $R > 0$~~

~~(2) point x_0, radius r, $B(x_0, r) \subseteq K$.~~

At each time, maintain region $\mathcal{E}^{(t)}$ s.t.

(a) $\mathcal{E}^{(0)} \supseteq$ ~~$Ball(0,R)$~~ K

(b) at each time test point $c^{(t)}$ "center" of $\mathcal{E}^{(t)}$ by separation oracle.

if YES, done!!

if NO, find $\mathcal{E}^{(t+1)}$

s.t. $\mathcal{E}^{(t+1)}$ contains $\underbrace{\mathcal{E}^{(t)}}_{\substack{\text{K lies here} \\ \text{by induction}}} \cap \underbrace{\{x \mid a^{(t)T} x < a^{(t)T} c^{(t)}\}}_{\substack{\text{K lies here} \\ \text{since in NO case}}}$

$\Rightarrow K \subseteq \mathcal{E}^{(t+1)}$.

(c) Argue that $vol(\mathcal{E}^{(t+1)}) \leq vol(\mathcal{E}^{(t)}) \cdot (1-\delta)$.

~~(d) Hold that $vol(K) \geq$ some ε.~~

So after T steps, $vol(\mathcal{E}^{(T)}) \leq vol(\mathcal{E}^{0}) \cdot e^{-\delta T}$

and hence setting $T = \frac{1}{\delta} \cdot \ln\left(\frac{vol(\mathcal{E}^{(0)})}{\varepsilon}\right)$, gives that $vol(\mathcal{E}^{(T)}) \leq \varepsilon$.

$\Rightarrow vol(K) \leq \varepsilon$

since $K \subseteq \mathcal{E}^{(T)}$.

This solves feasibility problem in T steps/oracle calls. ☺

in case all calls were NO.

if some call gave YES, already done!

Things to do:

Aside

(1) Show that can choose $\mathcal{E}^{(0)}$ s.t. $vol(\mathcal{E}^{(0)}) \le 2^{\overbrace{poly(\text{bit complexity}, n)}^{L}}$. done for LPs.

(2) Ensure that $\delta = 1/poly(L, n)$

for whatever choice of shape $\mathcal{E}$ and notion of center c being chosen.

(3) For feasibility problem to be useful, need

to show that small volume means something useful.

Looking ahead to LPs.

- suppose K is non empty and full dimensional, $\Rightarrow vol(K) > 0$

can show that $vol(K) \geq 2^{-poly(nL)}$.

So if set $\varepsilon = 2^{-poly(nL)}$ then will get contradiction, and hence

one call to oracle must return YES.

Aside

— x —

Right now: focus on getting solution for Feasibility problem, with "good" δ.

- A.Y. Levin / Newman (60s) define

$\mathcal{E}^{(t)} = \mathcal{E}^{(t-1)} \cap$ feasible region, $\mathcal{E}^{(0)}$ = some ball containing K.

$c^{(t)}$ = centroid of $\mathcal{E}^{(t)}$.

get $\delta = \Theta(1)$ but difficult to compute centroids.

- Khachiyan '79 used idea of Nahum Shor '70:

$\mathcal{E}^{(t)}$ = ellipsoid that contains $\mathcal{E}^{(t-1)} \cap$ feasible region

$c^{(t)}$ = center of ellipsoid

1st poly time algo. for LPs.

get $\delta \simeq \frac{\Theta(1)}{n}$. Needs to control the bit precision of numbers involved, since perform square roots, etc.

Today: use "simplices" idea of B. Yamnitski & Leonid Levin '82

- $\mathcal{E}^{(t)}$: simplex that contains $\mathcal{E}^{(t-1)} \cap$ feasible, $\mathcal{E}^{(0)}$: simplex also.
- but $\delta = \Theta\left(\frac{1}{n^2}\right)$.

Now no longer need sq. roots, but naively the bit precision may also increase, so need to do "truncation"/"rounding". see [Bartels 2000]

— x —

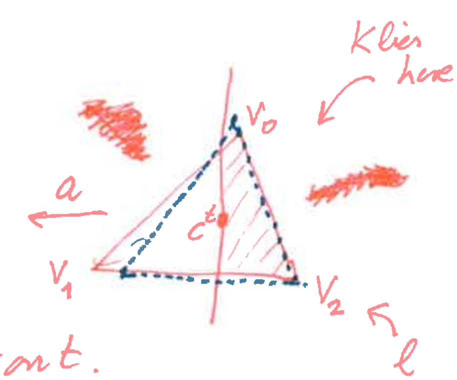
The Simplices Method

(as opposed to the Simplex Method of George Dantzig) ☺

- Start with $S^0 = \mathrm{conv}(v_0^0 \dots v_n^0)$. Set $c^{(t)} = $ centroid of $S^{(t)}$.
- Each time that we get NO,

  compute new simplex S^{t+1} by shifting some vertices of S^t.

  s.t. $\mathrm{vol}(S^{t+1}) \leq \mathrm{vol}(S^t)\left(1 - \Theta\left(\frac{1}{n^2}\right)\right)$



In more detail:

1) $S^t = \mathrm{conv}(v_0^t \dots v_n^t)$.

2) $c^t = \frac{1}{n+1} \sum_{i=0}^{n} v_i^t$

Naturally, a depends on t. not writing superscript a^t to avoid clutter.

3) if NO, get back $a \in \mathbb{R}^n$

   s.t. $K \subseteq \{x \mid a^T x < a^T c^t\}$

   a) define $e(x) := \langle a, c^t - x \rangle$. So $K \subseteq \{x \mid e(x) > 0\}$

   b) define $\ell = \arg\max_i e(v_i^t)$ ← non zero value $e(v_\ell^t)$ if $K \neq \phi$, since all v_i^t cannot be on hyperplane.

   c) $\forall i$, define $v_i^{t+1} := \left(1 - \frac{1}{d_i}\right) v_\ell^t + \frac{1}{d_i} v_i^t$ where $d_i = \left(1 - \frac{e(v_i^t)}{n^2 \cdot e(v_\ell^t)}\right)$

Observations:

- $d_i \geq 0 \ \forall i$
- Points on same side of hyperplane $\{e(x)=0\}$ as v_e $\qquad v_i' = v_e + \frac{v_i - v_e}{d_i}$

  have $e(v_i) > 0 \Rightarrow d_i < 1 \Rightarrow 1/d_i > 1$

  so these points move away from v_e

  Points on the other side of hyperplane (the "infeasible" side) have $e(v_i) \leq 0$

  so they move closer to v_e.

Want to show:

Fact 1: for any point x in $S^t \cap \{x \mid e(x) > 0\} \Rightarrow x \in S^{t+1}$. $\quad$ Hence $K \subseteq S^{t+1}$.

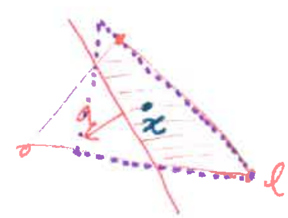
Fact 2: $vol(S^{t+1}) \leq Vol(S^t) \cdot (1 - \Omega(\frac{1}{n^2}))$.

this is the δ parameter from earlier.

Proof of Fact 1: (use $v_0 \ldots v_n$ for old points, $v_0', \ldots v_n'$ for new ones)

Let $x = \sum_i' \lambda_i v_i$, and $e(x) > 0$.

$\lambda_i \geq 0, \sum' \lambda_i = 1$.



define $\lambda_i' = \begin{cases} \lambda_i d_i & \forall i \neq \ell \\ d_\ell \lambda_\ell + \frac{e(x)}{n^2 e(v_\ell)} & \text{if } i = \ell. \end{cases}$

- $\lambda_i' \geq 0$ because all terms are positive.
- $\sum_i' \lambda_i' = \cancel{\sum_i' \lambda_i' + \lambda_e(1-d_e) + \frac{e(x)}{n^2 e(v_e)}}$

$$= \sum_i' \lambda_i d_i + \frac{e(x)}{n^2 e(v_e)} = \sum_i' \lambda_i \left(1 - \frac{e(v_i)}{n^2 e(v_e)}\right) + \frac{e(x)}{n^2 e(v_e)} = \cancel{\ldots} \rightarrow$$

$$= \sum_i' \lambda_i - \left[\frac{\sum_i' \lambda_i e(v_i) - e(x)}{n^2 e(v_\ell)} \right]$$

$$= 1.$$

but e is linear so $\sum_i' \lambda_i e(v_i) = e(\sum_i' \lambda_i v_i) = e(x)$

- Finally, claim that $x = \sum_i' \lambda_i' v_i'$.

$$\cancel{\sum_i' \lambda_i' v_i' = \sum_i \lambda_i d_i \left(v_\ell + \frac{v_i - v_\ell}{d_i}\right) + \frac{e(x)}{n^2 e(v_\ell)} \cdot v_\ell}$$

$$\cancel{v_\ell\left(\sum_i' \lambda_i d_i\right) + \sum_i' \lambda_i v_i - \left(\sum_i' \lambda_i\right) v_\ell + \frac{e(x)}{n^2 e(v_\ell)} \cdot v_\ell}$$

$$\sum_i' \lambda_i' v_i' = \sum_i \lambda_i d_i \left(\left(1 - \frac{1}{d_i}\right) v_\ell + \frac{v_i}{d_i} \right) + \frac{e(x)}{n^2 e(v_\ell)} \cdot v_\ell .$$

$$\downarrow \qquad \downarrow \qquad \downarrow$$

$$v_\ell \sum_i' \lambda_i (d_i - 1) + \sum_i' \lambda_i v_i + - \cancel{\frac{\sum_i' \lambda_i e(v_i)}{n^2 e(v_\ell)} v_\ell}$$

$$\downarrow \qquad \downarrow$$

$$-v_\ell \cdot \frac{\sum_i' \lambda_i e(v_i)}{n^2 e(v_\ell)} \qquad x$$

$$\downarrow$$

negative of last term

$$= x.$$

End of Proof of Fact 1.

——— x ———

<u>Fact 2</u>: Volume of S^{t+1} smaller than S^t by factor $\approx 1 - \frac{1}{n^2}$.

<u>Pf</u>: Volume of a simplex $= \frac{1}{n!} \det \begin{pmatrix} v_1 - v_0 \\ v_2 - v_0 \\ v_n - v_0 \end{pmatrix}$.
$\operatorname{conv}(v_0, v_1, \dots v_n)$

For simplicity say that $v_\ell = v_0$. ~~and say v_ℓ = the zero vector.~~

then recall $v_i' = v_0 + \dfrac{v_i - v_0}{d_i}$ $\qquad v_0' = v_0$.

$$\mathrm{vol}(S') = \frac{1}{n!}\cdot\det\begin{pmatrix} v_1' - v_0' \\ \vdots \\ v_n' - v_0' \end{pmatrix} = \frac{1}{n!}\det\begin{pmatrix} \frac{v_1 - v_0}{d_1} \\ \vdots \\ \frac{v_n - v_0}{d_n} \end{pmatrix} = \frac{\mathrm{vol}(S)}{\prod_{i\neq 0} d_i}$$

How small can $\prod_{i\neq 0} d_i$ be?

- By construction: $d_i = 1 - \dfrac{e(v_i)}{n^2 e(v_0)} \geq 1 - \dfrac{1}{n^2}$ $\qquad$ since $e(v_0) \geq e(v_i)$

- Moreover $\displaystyle\sum_{i\neq 0} d_i = \sum_{i\neq 0}\left(1 - \frac{e(v_i)}{n^2 e(v_0)}\right)$

$$= n - \underbrace{\sum_{i=0}^{n} \frac{e(v_i)}{n^2 e(v_0)}}_{\frac{e(\text{centroid})}{n^2 e(v_0)} = 0} + \underbrace{\frac{e(v_0)}{n^2 e(v_0)}}_{1/n^2}$$

by defn of $e(x) = a^T(c - x)$.

$$= n + 1/n^2.$$

So to minimize $\prod_{i\neq 0} d_i$ subject to $d_i \geq 1 - 1/n^2$ and $\sum_{i\neq 0} d_i = n + 1/n^2$

have all equal to $1 - 1/n^2$ and last one be $1 + 1/n$.

which gives $\geq \left(1 - \frac{1}{n^2}\right)^{n-1}\cdot\left(1 + \frac{1}{n}\right) = \dfrac{\left(1 - \frac{1}{n^2}\right)^n}{\left(1 - \frac{1}{n}\right)}$.

~~Taylor series says this $\geq$~~

Taking logs, $\log\left(\frac{(1-\frac{1}{n^2})^n}{(1-\frac{1}{n})}\right) = -n.\sum_{i=1}^{\infty}\frac{1}{in^{2i}} + \sum_{i=1}^{\infty}\frac{1}{in^i}$

$$= \left(-\frac{n}{n^2}-\frac{n}{2n^4}-\frac{n}{3n^6}\cdots\right)+\left(\frac{1}{n}+\frac{1}{2n^2}+\frac{1}{3n^3}+\cdots\right)$$

$$\geq \frac{1}{2n^2}-\frac{1}{6n^3}$$

$\Rightarrow \prod d_i \geq e^{1/2n^2 - 1/6n^3}$

$\Rightarrow$ reduction in volume $\leq e^{-1/2n^2+1/6n^3} \overset{n}{=} \left(1-\theta\left(\frac{1}{n^2}\right)\right)$ when n large.

———— x ————

Same kind of ideas hold for the Ellipsoid also, but now deal with psd matrices defining those ellipsoids.

One concern, btw: the numbers in the description of V_i may blow up. So Sven Bartels gives a rounding technique to "blow up" the simplex a bit and get vertices defined using bounded precision. See paper for details.

———— x ————

Solves feasibility problem. To use to solve LP feasibility, use the fact that

for full dimensional polytopes ~~equations~~

① ~~original polytope can be~~

there is an optimal solution

that has bounded bit precision (so putting a bounding box of the form $|x_i| \leq 2^{poly(L)}$ is OK)

discuss soon...

Hence volume of $S^0 \leq 2^{poly(L)}$

② if $vol(K) > 0 \Rightarrow vol(K) > 2^{-poly(nL)}$ $\Rightarrow$ setting ε to this value is OK

$\Rightarrow$ runtime $= \frac{1}{\delta}\cdot\log\left(\frac{vol(S^0)}{\varepsilon}\right) = n^2 \cdot poly(nL)$. ☺

Next: Separation oracles —