matrix multiplication [Freivalds '77]

$n \times n$ matrices A, B, C

output: Yes if $AB = C$
& No if not.

Trivial alg.: Compute AB in $O(n^{2.376\ldots})$ time & compare.

Better randomized algorithm:

1. Choose a finite set S, e.g., $S = \{0, 1, \ldots, k-1\}$, where $|S| \geq 2$.
2. Pick a vector $r = (r_1, \ldots, r_n)$
   where each r_i is chosen uniformly at random from S.
3. If $(AB)r \neq Cr$ then output No
   else output Yes.

Running time: Note, $(AB)r = A(Br)$ & thus it takes $O(n^2)$ time to compute.

Lemma: $\Pr(\text{output Yes} \mid AB \neq C) \leq 1/2$

Boosting: Run t times. If all t trials output YES then output Yes, else output No.

Now, $\Pr(\text{output Yes} \mid AB \neq C) \leq 2^{-t}$

Proof of Lemma:

Let $D = AB - C$.

Assuming $AB \neq C$ we have $D \neq 0$, so there is at least one entry in D which is non-zero.

Let's assume $d_{11} \neq 0$ (if not, relabel the rows/columns)

Our goal is to show: $\Pr(Dr = 0) \leq 1/2$.

If $Dr = 0$ then $(Dr)_1 = 0$.

Note, $(Dr)_1 = \sum_{i=1}^{n} d_{1i} r_i$

if $(Dr)_1 = 0$ then $r_1 = \frac{-1}{d_{11}}(d_{12}r_2 + \cdots + d_{1n}r_n) := s^*$

When choosing r, first choose $r_2, \ldots, r_n$ & then r_1.

This is the principle of deferred decisions.

Hence, there is ≤ 1 choice of r_1 so that $(Dr)_1 = 0$,

and thus, $\Pr(Dr = 0) \leq \Pr((Dr)_1 = 0) \leq \Pr(r_1 = s^*)$

$$\leq \frac{1}{|S|} \leq \frac{1}{2}.$$

Alternatively, for all $\alpha_2, \ldots, \alpha_n \in S$, look at $\Pr\left(r_1 = s^* \,\middle|\, \begin{matrix} r_2 = \alpha_2 \\ \vdots \\ r_n = \alpha_n \end{matrix}\right)$

$$\Pr((Dr)_1 = 0) = \sum_{\alpha_2, \ldots, \alpha_n} \Pr((Dr)_1 = 0 \mid r_2 = \alpha_2, \ldots, r_n = \alpha_n) \Pr(r_2 = \alpha_2, \ldots, r_n = \alpha_n)$$

$$= \sum_{\alpha_2, \ldots, \alpha_n} \Pr(r_1 = s^*_{\alpha_2, \ldots, \alpha_n} \mid r_2 = \alpha_2, \ldots, r_n = \alpha_n) \Pr(r_2 = \alpha_2, \ldots, r_n = \alpha_n)$$

$$\leq \frac{1}{|S|} \sum_{\alpha_2, \ldots, \alpha_n} \Pr(r_2 = \alpha_2, \ldots, r_n = \alpha_n) = \frac{1}{|S|} \leq \frac{1}{2}$$

$\boxtimes$

Testing polynomial identities:

Given 2 polynomials Q & R over n variables $x_1, \ldots, x_n$
& degree $\leq d$.

Goal: test if $Q = R$?

Might have exponential # of terms/monomials so assume we have oracle access to Q & R: Given values $x_1, \ldots, x_n$ can efficiently evaluate $Q(x_1, \ldots, x_n)$ & $R(x_1, \ldots, x_n)$

As before, randomized algorithm with
small prob. of false positives
& no false negatives.

Schwartz-Zippel algorithm:

Consider $P = Q - R$. Test if $P = 0$?

uniformly at random

For a finite set S, choose $r_1, \ldots, r_n$ u.a.r. from S

Lemma: If $P \neq 0$, then $\Pr\left(P(r_1, \ldots, r_n) = 0\right) \leq \frac{d}{|S|}$

Hence, for S where $|S| \geq 2d$ we get error prob. $\leq 1/2$
& can run t times to get error prob. $\leq 2^{-t}$!

Proof of lemma: Assume $P \neq 0$.

Induct on n.

Base case: $n=1$.

Then, P is a univariate polynomial.

Since $\deg(P) \leq d$ then P has $\leq d$ roots.

Thus, $\Pr(P(r_1)=0) \leq d/|S|$.

Let k be the max degree of x_1 in P.

Then,
$$P(x_1,\dots,x_n) = M(x_2,\dots,x_n)x_1^k + N(x_1,\dots,x_n)$$
where $\deg(M) \leq d-k$ & degree of x_1 in N is $<k$.

Use principle of deferred decisions again & choose $r_2,\dots,r_n$ first & then choose r_1.

Let $\mathcal{E}$ be the event that $M(r_2,\dots,r_n)=0$

Two cases:

- $\mathcal{E}$ occurs: by induction on M (since it has $n-1$ variables) we know $\Pr(\mathcal{E}) \leq \frac{d-k}{|S|}$

- $\mathcal{E}$ does not occur: let P' be poly in x_1 after $x_2=r_2,\dots,x_n=r_n$ is plugged into $P(x_1,\dots,x_n)$.

  Note, $P' \neq 0$ & $\deg(P') \leq k$ & it has 1 variable.

  Thus $\Pr(P'(r_1)=0 \mid \overline{\mathcal{E}}) \leq k/|S|$

because have term αx_1^k where $\alpha = M(r_2,\dots,r_n)$

$$
\begin{aligned}
\Pr\big(P(r_1,\ldots,r_n)=0\big)
&= \Pr\big(P(r_1,\ldots,r_n)=0 \mid \mathcal{E}\big)\times\Pr(\mathcal{E}) \\
&\quad + \Pr\big(P(r_1,\ldots,r_n)=0 \mid \overline{\mathcal{E}}\big)\times\Pr(\overline{\mathcal{E}}) \\
&\le \Pr(\mathcal{E}) + \Pr\big(P(r_1,\ldots,r_n)=0 \mid \overline{\mathcal{E}}\big) \\
&\le (d-k)/|S| + \Pr\big(P'(r_1)=0 \mid \overline{\mathcal{E}}\big) \\
&\le \frac{d-k}{|S|} + \frac{k}{|S|} = \frac{d}{|S|}. \qquad \blacksquare
\end{aligned}
$$

Bipartite perfect matching via polynomial identity testing

Given a bipartite graph $G=(V_1, V_2, E)$ where $|V_1|=|V_2|=n$,
Does G contain a perfect matching?

Tutte matrix: For graph G, Define $n \times n$ matrix A_G where:

$$a_{ij} = \begin{cases} x_{ij} & \text{if } (i,j)\in E \\ 0 & \text{o/w} \end{cases}$$

where the x_{ij}'s are variables.

Lemma: $\det(A_G) \neq 0$ iff G contains a perfect matching.

Proof: Note, $\det(A_G) = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)}$

where S_n is the set of $n!$ permutations of $\{1,\dots,n\}$
and $\mathrm{sgn}(\sigma) = (-1)^{N(\sigma)}$ where $N(\sigma) = \#$ of inversions in σ

$= (-1)^{\#\{\text{even cycles in } \sigma\}}$

$= (-1)^{n-k}$ where $k = \#$cycles in σ.

If G contains a perfect matching σ then $\prod_i a_{i,\sigma(i)} = \prod_i x_{i,\sigma(i)}$
& every other perfect matching has ≥ 2 diff. monomials
So no cancellations.

And if $\det(A_G)$ has a non-zero term then that corresponds to a perf. matching $\boxminus$

Can we use our poly. id. testing alg. to test if
G has a perfect matching?

Yes by checking if $\det(A_G)$ (which is just a poly.)
is 0 or not.

How to find a perfect matching?

Try an edge e, Check if smaller graph
contains a p. m. or not.

<u>Input:</u> $G=(V,E)$

Set $M=\emptyset$

While $M \neq$ perfect matching:

- Choose an edge $e=(i,j)$ of G
- Test if $G' = G \setminus \{i,j\}$ (remove vertices i & j) has a perfect matching (using Schwartz-Zippel alg.)
  - if YES then let $M = M \cup e$ & $G = G'$
  - if No then let ~~$G = G \setminus \{(i,j)\}$~~ (drop edge e) $G = G \setminus e$

Output M.

This takes $O(m)$ rounds.

(Need to do $t = O(\log m)$ trials of Schwarz-Zippel in each round to get each one to have error prob. $\leq \frac{1}{\text{Poly}(m)}$)

Can we do it in parallel: test each edge simultaneously?

[Mulmuley, Vazirani, Vazirani 87]:

<u>Isolation Lemma:</u> Let $S_1, \ldots, S_k$ be subsets of a set S, $|S| = m$.

Let each $x \in S$ have weight w_x which is chosen independently & u.a.r. from $\{1, \ldots, \ell\}$.

Then, $\Pr(\exists$ <u>unique</u> set S_i of min weight$) \geq \cancel{\left(1 - \frac{1}{\ell}\right)^m}$

$$\geq 1 - \frac{m}{\ell}.$$

Note, for a set S_i, $w(S_i) = \sum_{x \in S_i} w_x$

Proof of isolation lemma:

For $x \in S$, say x is tied if $\min_{S_j \ni x} w(S_j) = \min_{S_j \not\ni x} w(S_j)$

Note, $\exists$ tied element $y \in S$ iff min weight subset is not unique.

Fix $x \in S$ & let's prove $\Pr(x \text{ is tied}) \leq 1/\ell$.

Use principle of deferred decisions:

Fix w_y for all $y \in S, y \neq x$.

Let $W^+ = \min_{S_j \ni x} w(S_j) - w_x$

& $W^- = \min_{S_j \not\ni x} w(S_j)$

So W^+ & W^- are functions of w_y $\forall y \neq x$.

Note, x is tied iff $w_x = W^- - W^+$.

$$\Pr(x \text{ is tied} \mid w_y \text{ for all } y \neq x) \leq 1/\ell$$

$$\Pr(\exists \text{ tied } y) \leq \sum_{y \in S} \Pr(y \text{ is tied}) \leq \frac{m}{\ell}$$

$$\Pr\left(\exists \text{ unique subset } S_i \text{ of min weight}\right) \geq 1 - \frac{m}{\ell}$$

$\boxtimes$