$$T(n) \le T\left(\frac{3}{4}n\right) + 2 \times O(n)$$

$$\Rightarrow \quad O(n).$$

(more careful analysis: $\lesssim 3.4n$ comps)

$2 + 2\ln 2$

Method 2: Floyd-Rivest Alg'm ('75)

idea - random sampling

⋮

Midterm 1 Stats

mean 60.28   stdev 17.6

median 58.75

max 98

90-100	4
80-89	14
70-79	13
60-69	16
50-59	21
40-49	14
<40	14

Floyd-Rivest Alg'm

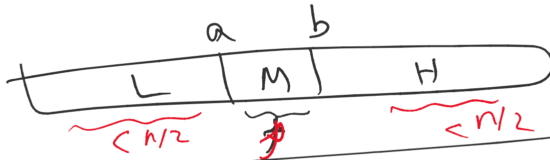
1. pick subset R of r random elements $y_1, \dots, y_r$
2. take $\left(\frac{r}{2} - c\sqrt{r}\right)$-th smallest ^{a} of R

   $\left(\frac{r}{2} + c\sqrt{r}\right)$-th smallest ^{b} of R

{ all elements $\le a$ }

$(\frac{1}{2}+c(1))$

3. Let $L = \{$ all elements $\leq a\}$
$M = \{$ all elems between a & $b\}$
$H = \{$ all elements $> b\}$



$$\boxed{\text{check: } |L| \leq \frac{n}{2},\ |H| \leq \frac{n}{2}, \text{ and } |M| \leq 4c\frac{n}{\sqrt{r}}} \quad (*)$$

(Monte Carlo)

4. return $(\frac{n}{2} - |L|)^{th}$ smallest of M

Runtime / # Comps:

line 2 $\quad O(r \log r)$

line 4 $\quad O\left(\frac{cn}{\sqrt{r}} \log \frac{cn}{\sqrt{r}}\right)$ $\leftarrow O((cn)^{2/3} \log n)$

line 3 $\quad \underbrace{n + \frac{n}{2}}_{\text{comp with } a} + \underbrace{4c\frac{n}{\sqrt{r}}}_{\text{comp with } b}$

Set r s.t. $r \approx \frac{cn}{\sqrt{r}} \Rightarrow r \approx (cn)^{2/3}$

$$\Rightarrow \# \text{ comps } \boxed{1.5n} + \underline{O((cn)^{2/3} \log n)}$$

Error Analysis:

Recall Markov's ineq:
$X > 0,\ E[X] = \mu \Rightarrow \Pr(X \geq c\mu] \leq \frac{1}{c}$.

Chebyshev's ineq:
$E[X] = \mu,\ Var(X) = \sigma^2$
$\Rightarrow \Pr(|X - \mu| \geq c\sigma) \leq \frac{1}{c^2}$.

Chernoff Bd:

Chernoff's Bd:

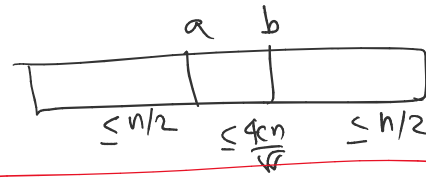
Let $X_1, \ldots, X_r$ be indep. 0-1 rand vars with $\Pr(X_i = 1) = p$.

Let $X = \sum_{i=1}^{r} X_i$. Let $\mu = pr$.

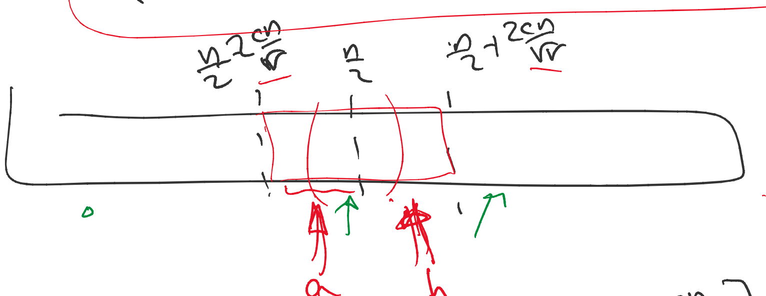
$$\Rightarrow \Pr\left[\,|X - \mu| \geq c\sqrt{\mu}\,\right] \leq e^{-\Theta(c^2)}$$

(assuming $c \leq \sqrt{\mu}$)

How to analyze Prob of (*):



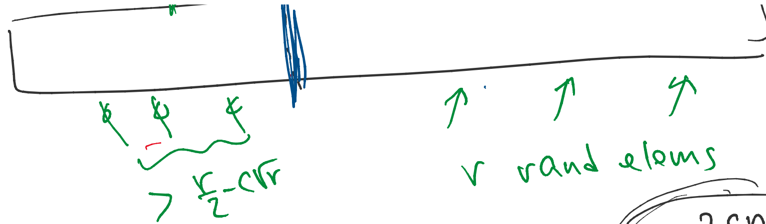
(*) is true if a has rank between $\frac{n}{2} - \frac{2cn}{\sqrt{r}}$ and $\frac{n}{2}$ & b has rank between $\frac{n}{2}$ and $\frac{n}{2} + \frac{2cn}{\sqrt{r}}$



What's Prob$\left[\text{rank of } a \text{ is } < \frac{n}{2} - \frac{2cn}{\sqrt{r}}\right]$?

$= \text{Prob}\left[\underbrace{\left(\begin{array}{l}\text{\# elems in } R \\ \text{with rank } < \frac{n}{2} - \frac{2cn}{\sqrt{r}}\end{array}\right)}_{X} > \frac{r}{2} - c\sqrt{r}\right]$

(recall $a = \left(\frac{r}{2} - c\sqrt{r}\right)^{th}$ smallest in R.)



Let $X_i = \begin{cases} 1 & \text{if } \operatorname{rank}(y_i) \leq \frac{n}{2} - \frac{2cn}{\sqrt{r}} \\ 0 & \text{else} \end{cases}$

$$p = \frac{\frac{n}{2} - \frac{2cn}{\sqrt{r}}}{n} = \frac{1}{2} - \frac{2c}{\sqrt{r}}$$

$$X = \sum_{i=1}^{r} X_i \qquad \mu = rp = \underbrace{\frac{r}{2} - 2c\sqrt{r}}$$

$p = \Pr(X_i = 1)$

$$\Pr\left(X > \underbrace{\frac{r}{2} - c\sqrt{r}}\right)$$

$$\leq \Pr\left(X > \mu + \underbrace{\cancel{c\sqrt{r}}}_{\Theta(c\sqrt{\mu})}\right)$$

$$\leq \underline{e^{-\Theta(c^2)}}.$$

Similar for $\Pr\left[\text{rank of } a \text{ is } > \frac{n}{2}\right]$ ...

e.g. set $c = n^{0.1}$ $\Rightarrow$ # comps $1.5n + o(n)$

but err prob. $e^{-\Theta(n^{0.1})}$

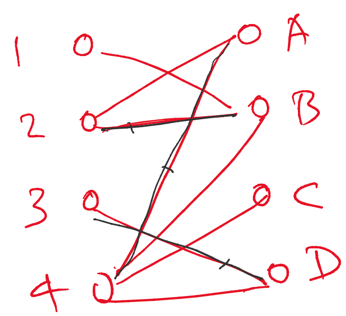
can convert to Las Vegas.

OPTIMIZATION: Matching, Flows, & LP

OPTIMIZATION: Matching, Flows, &

Problem Given unweighted bipartite graph $G=(V,E)$
$|V|=n$, $|E|=m$, $n \lesssim m \leq n^2$

find matching with max size
subgraph with max deg 1



$\{2B, 3D, 4A\}$

$\{1B, 2A, 3D, 4C\}$

brute force: $\Omega(n!)$
greedy not work
DP not work
⋮

"Hungarian" Method (König-Egerváry '55 / Edmonds '65)

idea. local improvement
(start with feasible sol'n,
modify & improve ...)

Def Let M be a matching.
An augmenting path P for M is
a path $v_1 v_2 v_3 \ldots v_k$ in G
s.t. $v_1 v_2 \notin M$, $v_2 v_3 \in M$, $v_3 v_4 \notin M$, ...
... $v_{k-1} v_k \notin M$

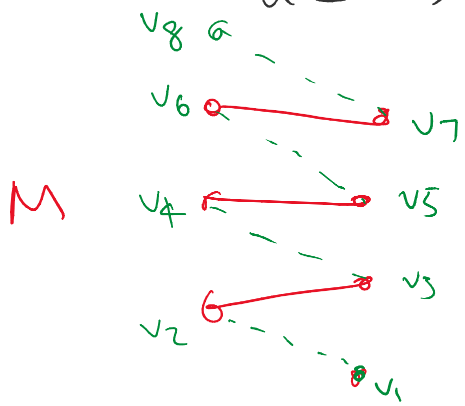
alternating path

... by M

s.t. $v_1 v_2 \notin$, , ,

... $v_{k-1} v_k \notin M$

and v_1, v_k are not covered by M.



Notation:

$S \oplus T$

$= S \cup T - (S \cap T)$

$\Rightarrow$ $M \oplus P$ is a matching

with $|M \oplus P| \geq |M| + 1$.

Algm:

$M = \emptyset$

repeat {

find augmenting path P for M

if not exist, return M

$M = M \oplus P$

}

To prove: matching M is optimal

$\Longleftrightarrow$ it has no augmenting path.

each iteration: $O(m)$ time

iterations: $\leq n$

$\Rightarrow$ $\boxed{O(mn)}$

more details to come ...