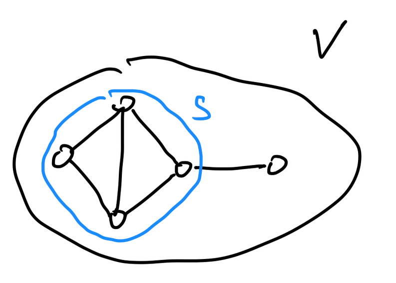
DENSEST SUBGRAPH

INPUT: $G(V,E)$, $E \subseteq \binom{V}{2}$

Q: WHAT IS THE SUBSET $S \subseteq V$ HAVING MAXIMUM DENSITY?

$$\rho(S) = \frac{|E(S)|}{|S|} \qquad E(S) = \{\{u,v\} \mid \{u,v\} \in E \wedge u,v \in S\}$$



$\rho(V) = \frac{6}{5} = 1.2$

$\rho(S) = \frac{5}{4} = 1.25$

$deg_S(v) = |\{\{u,v\} \mid \{u,v\} \in E(S)\}|$

$$\rho(S) = \frac{|E(S)|}{|S|} = \frac{1}{2} \frac{2|E(S)|}{|S|} = \frac{1}{2} \frac{\sum_{v \in S} deg_S(v)}{|S|}$$

$$= \frac{1}{2} \text{avg } deg(S)$$

K-DENSEST SUBGRAPH
FIND $S \subseteq V$, $|S| = K$, SUCH THAT $|E(S)|$ IS AS LARGE AS POSSIBLE.

$n^{\frac{1}{4}+\varepsilon}$

GOING BACK TO DENSEST SUBGRAPH

$$LP \begin{cases} \max \sum_{\{i,j\} \in E} x_{\{i,j\}} \\ x_{\{i,j\}} \le y_i & \forall \{i,j\} \in E \\ x_{\{i,j\}} \le y_j \\ \sum_{i \in V} y_i \le 1 \\ x_{\{i,j\}}, y_i \ge 0 & \forall i \in V, \forall \{i,j\} \in E \end{cases}$$

L1: FOR ANY $G(V,E)$, $\forall S \subseteq V$, $\exists$ FEASIBLE SOLUTION TO THE LP HAVING VALUE $\ge \frac{|E(S)|}{|S|} = \rho(S)$.

P: FOR $i \in S$, SET $y_i = \frac{1}{|S|}$. FOR $i \in V-S$, SET $y_i = 0$.

FOR EACH $\{i,j\} \in E$, WE SET

$$x_{\{i,j\}} = \begin{cases} 1/|S| & \text{IF } \{i,j\} \in E(S) \text{ (IF } i,j \in S), \\ 0 & \text{OTHERWISE.} \end{cases}$$

THE CONSTRAINT $\sum_{i \in V} y_i \le 1$ IS SATISFIED, GIVEN THAT

$$\sum_{i \in V} y_i = \sum_{i \in S} y_i + \sum_{i \in V-S} y_i = \sum_{i \in S} \frac{1}{|S|} + \sum_{i \in V-S} 0 = \frac{|S|}{|S|} = 1.$$

TAKE ANY $\overline{\{i,j\} \in E(S)}$. WE HAVE $x_{\{i,j\}} = \frac{1}{|S|}$.
BUT, IF $\{i,j\} \in E(S)$ THEN $i,j \in S$. THUS, $y_i = y_j = \frac{1}{|S|}$.
THEN, $\forall \{i,j\} \in E(S)$, THE CONSTRAINTS $x_{\{i,j\}} \le y_i$ AND $x_{\{i,j\}} \le y_j$ ARE BOTH SATISFIED.

IF INSTEAD $\{i,j\} \notin E(S)$, THEN $x_{\{i,j\}} = 0$. THUS, $x_{\{i,j\}} \le y_i$ AND $x_{\{i,j\}} \le y_j$ ARE SATISFIED BY $y_i, y_j \ge 0$.

THUS, OUR SOLUTION IS FEASIBLE.

THE VALUE OF THIS SOLUTION IS

$$\sum_{\{i,j\} \in E} x_{\{i,j\}} = \sum_{\{i,j\} \in E(S)} x_{\{i,j\}} = \sum_{\{i,j\} \in E(S)} \frac{1}{|S|} = \frac{|E(S)|}{|S|} = \rho(S). \quad \square$$

L2: FOR ANY FEASIBLE LP SOL. OF VALUE v, $\exists S \subseteq V$ SUCH THAT $\rho(S) \ge v$.

COR: $LP^* = OPT_{DS}$.

P: BY L1, $LP^* \ge OPT_{DS}$. BY L2, $LP^* \le OPT_{DS}$. $\square$

P OF L2:

LET $\underline{y}', \underline{x}'$ BE A FEASIBLE LP SOLUTION OF VALUE v.

LET $\underline{y}, \underline{x}$ BE THE SOLUTION TO THE LP S.T.:

- $y_i = y_i'$, $\forall i \in V$
- $x_{\{i,j\}} = \min(y_i, y_j)$, $\forall \{i,j\} \in E$.

THE $\underline{y}, \underline{x}$ SOLUTION IS FEASIBLE, INDEED:

- $\sum_{i \in V} y_i = \sum_{i \in V} y_i' \le 1$, AND
- $\forall \{i,j\} \in E$, IT HOLDS $x_{\{i,j\}} \le y_i$ AND $x_{\{i,j\}} \le y_j$.

WE HAVE TO PROVIDE A SET S HAVING DENSITY AT LEAST v.
WE ARE GOING TO PROVIDE A NUMBER OF S's, ONE OF WHICH WILL HAVE THE DESIRED PROPERTY.

$$S(r) = \{i \mid y_i \ge r\} \qquad \forall r \ge 0$$
$$E(r) = \{\{i,j\} \mid x_{\{i,j\}} \ge r\}$$

OBS: $\{i,j\} \in E(r) \iff (i \in S(r) \text{ AND } j \in S(r))$

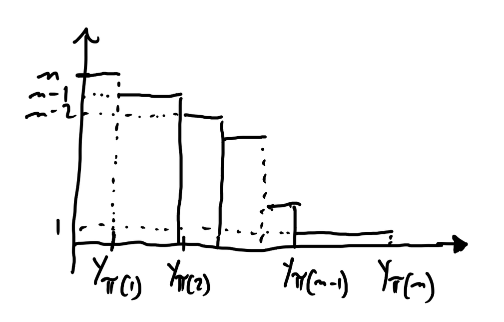
P: EXERCISE.

CL: $\exists r \ge 0$ S.T. $\rho(S(r)) = \frac{|E(r)|}{|S(r)|} \ge \sum_{\{i,j\} \in E} x_{\{i,j\}} = v$.

P: WE DEFINE A PERMUTATION π OF THE NODES, S.T.

$$0 \le y_{\pi(1)} \le y_{\pi(2)} \le y_{\pi(3)} \le \dots \le y_{\pi(n-1)} \le y_{\pi(n)} \le 1$$

$\uparrow \sum_{i \in V} y_i \le 1$ AND $y_i \ge 0 \ \forall i \in V$.



$$\int_0^1 |S(r)|\,dr = \int_0^{y_{\pi(1)}} |S(r)|\,dr + \int_{y_{\pi(1)}}^{y_{\pi(2)}} |S(r)|\,dr + \dots + \int_{y_{\pi(n)}}^{1} |S(r)|\,dr =$$

$$= \int_0^{y_{\pi(1)}} n\,dr + \int_{y_{\pi(1)}}^{y_{\pi(2)}} (n-1)\,dr + \dots + \int_{y_{\pi(n-1)}}^{y_{\pi(n)}} 1\,dr + \int_{y_{\pi(n)}}^{1} 0\,dr$$

$$= (y_{\pi(1)} - 0) \cdot n + (y_{\pi(2)} - y_{\pi(1)})(n-1) + \dots + (y_{\pi(n)} - y_{\pi(n-1)}) \cdot 1 + 0$$

$$= y_{\pi(1)} \cdot (n-(n-1)) + y_{\pi(2)}((n-1)-(n-2)) + \dots + y_{\pi(n)}(1-0)$$

$$= \sum_{i=1}^{n} y_{\pi(i)} = \sum_{i=1}^{n} y_i \le 1$$

$\uparrow$ FEASIBILITY OF THE LP SOL.

$$\int_0^1 |E(r)|\,dr = \sum_{\{i,j\} \in E} x_{\{i,j\}} = \sum_{\{i,j\} \in E} \min(y_i, y_j)$$

$$= \sum_{\{i,j\} \in E} \min(y_i', y_j')$$

$$\xrightarrow{\underline{x}', \underline{y}' \text{ IS A FEASIBLE SOL.}} \ge \sum_{\{i,j\} \in E} x'_{\{i,j\}} = v \quad \text{(VALUE OF THE ORIGINAL LP SOL).}$$

THEN, $\int_0^1 |S(r)|\,dr \le 1$ AND $\int_0^1 |E(r)|\,dr \ge v$.

BY CONTRADICTION, SUPPOSE THAT $\forall r \ge 0$ $\frac{|E(r)|}{|S(r)|} < v$.

THEN, $|E(r)| < v \cdot |S(r)|$ $\forall r \ge 0$.

THUS,

$$v \le \sum_{\{i,j\} \in E} x_{\{i,j\}} = \int_0^1 |E(r)|\,dr < \int_0^1 v\,|S(r)|\,dr$$
$$= v \int_0^1 |S(r)|\,dr$$
$$\le v \cdot 1 = v.$$

SINCE $v \not< v$, WE HAVE A CONTRADICTION.

THUS, $\exists r \ge 0$ S.T. $|E(r)| \ge v\,|S(r)| \Rightarrow \rho(S(r)) \ge v$. $\square$

THUS, WE CAN GET AN OPTIMAL (DENSEST) SUBGRAPH (ONE OF $S(y_1), S(y_2) \dots S(y_n)$ WILL DO).

I CAN SOLVE THE LP IN POLYTIME, AND WE CAN TRANSFORM THE LP SOL IN AN OPTIMAL SUBGR.

IS THIS FAST ENOUGH? NO. (THE LP HAS $\ge |E|$ VARIABLES).