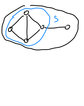
DENSEST SUBGRAPH

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

OUTPUT: A SUBGRAPH $S \subseteq V$ OF MAXIMUM DENSITY $p(S) = \frac{|E(S)|}{|S|}$



$$E(S) = \{ e \mid e \in E \wedge e \subseteq S \}$$

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

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

$$2\,p(S) = \max_{v \in S} d_{G_S}(v)$$

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

THE LP HAS UP TO $\Theta(n^2)$ VARIABLES.

IF n IS VERY LARGE ($10^6 \ldots 10^9$ NODES), n^2 IS TOO SLOW.

CHARIKAR

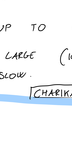
GREEDY($G(V,E)$):
$S_0 \leftarrow V$
FOR $i = 1, 2, \ldots, n$
LET v_i BE A NODE HAVING MINIMUM DEGREE IN $G[S_{i-1}]$
$S_i \leftarrow S_{i-1} - \{v_i\}$
RETURN A SET S_i HAVING LARGEST DENSITY $p(S_i)$.
(THE OUTPUT IS THE "BEST" OF $S_0, S_1, \ldots, S_{n-1}$)

$G[S]$, FOR $S \subseteq V$, IS THE GRAPH HAVING S AS ITS SET OF NODES, AND $E(S)$ AS ITS SET OF EDGES

EX: PROVE THAT GREEDY DOES NOT ALWAYS RETURN AN OPTIMAL SOL.

DEF: AN ORIENTATION Φ OF THE EDGES OF AN UNDIRECTED GRAPH $G(V,E)$ IS A FUNCTION THAT ASSIGNS TO EACH $e \in E$ ONE OF ITS ENDPOINTS.

$\Phi : E \to V$
$\Phi(e) \in e$



$2^{|E|}$ = # OF ORIENTATIONS OF $G(V,E)$

DEF: GIVEN AN ORIENTATION Φ OF $G(V,E)$, LET $d_\Phi(v)$ BE THE IN-DEGREE OF $v \in V$, ACCORDING TO Φ

OBS.: IF Φ IS AN ORIENTATION OF $G(V,E)$, THEN $\sum_{v \in V} d_\Phi(v) = |E|$.

DEF.: THE MAX-DEGREE OF Φ, IS $\Delta_\Phi = \max_{v \in V} d_\Phi(v)$.

L: $\max_{\emptyset \subsetneq S \subseteq V} p(S) = \max_{\emptyset \subsetneq S \subseteq V} \frac{|E(S)|}{|S|} \le \Delta_\Phi$, $\forall$ ORIENTATION Φ.

P: LET $\emptyset \subsetneq S \subseteq V$ BE GIVEN. RECALL THAT

$$E(S) = \{ \{w,v\} \mid \{w,v\} \in E \wedge \{w,v\} \subseteq S \}$$

EACH $\{w,v\} \in E(S)$ IS GOING TO BE ORIENTED TOWARDS ONE OF ITS ENDPOINTS, THAT IS, TOWARDS A NODE OF S.

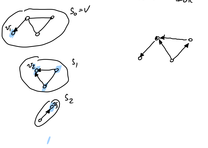
$$|E(S)| \le \sum_{v \in S} d_\Phi(v) \le \sum_{v \in S} \Delta_\Phi = |S| \cdot \Delta_\Phi .$$

THUS, $\frac{|E(S)|}{|S|} \le \Delta_\Phi$. ☐

WHILE RUNNING THE GREEDY ALGORITHM WE BUILD AN ORIENTATION:

- WHENEVER NODE v_i (THE iTH NODE TO BE REMOVED) IS REMOVED FROM THE GRAPH (WHICH, AT THAT POINT HAS NODE SET S_{i-1}) WE ORIENT EACH EDGE OF $G[S_{i-1}]$ THAT IS INCIDENT ON v_i TOWARDS v_i.

THE RESULTING ORIENTATION IS CALLED Φ_{GR}.



OBSERVE THAT, WITH THE ORIENTATION Φ_{GR}, $d_{\Phi_{GR}}(v)$ EQUALS THE DEGREE OF v IN THE GRAPH, WHEN v IS ABOUT TO BE REMOVED.

L: LET M BE THE MAXIMUM $p(S_i)$ VALUE, $M = \max_{i=0,\ldots,n-1} p(S_i)$. (THEN, M IS THE VALUE OF THE GREEDY SOLUTION.)

THEN, $\Delta_{\Phi_{GR}} \le 2M$.

P: FOR $i = 1, 2, \ldots, n$, v_i IS A NODE OF MINIMUM DEGREE IN $G[S_{i-1}]$.

IN GENERAL, THE MINIMUM DEGREE IN A GRAPH IS NOT LARGER THAN THE AVERAGE DEGREE OF THE SAME GRAPH. THUS,

$$d_{\Phi_{GR}}(v_i) = \min_{v \in S_{i-1}} d_{S_{i-1}}(v) \le \frac{\sum_{v \in S_{i-1}} d_{S_{i-1}}(v)}{|S_{i-1}|} = \frac{2|E(S_{i-1})|}{|S_{i-1}|} = 2\,p(S_{i-1})$$

THEN,

$$\Delta_{\Phi_{GR}} = \max_{v \in V} d_{\Phi_{GR}}(v) = \max_{i=1,\ldots,n} d_{\Phi_{GR}}(v_i) \le \max_{i=1,\ldots,n} 2\,p(S_{i-1}) = 2M . \ ☐$$

COR.: GREDY RETURNS A 2-APPRX TO DENSEST SUBGRAPH.

P: $M \ge \frac{\Delta_{\Phi_{GR}}}{2} \ge \frac{\max_{\emptyset \subsetneq S \subseteq V} p(S)}{2}$. THUS, THE GREDY SOL. IS NO WORSE THAN A 2-APPRX. ☐

A GENERAL LP

PRIMAL

$$\begin{cases} \max\ c_1 x_1 + c_2 x_2 + \cdots + c_n x_n \\ y_1:\ a_{11} x_1 + a_{12} x_2 + \cdots + a_{1n} x_n \le b_1 \\ y_2:\ a_{21} x_1 + a_{22} x_2 + \cdots + a_{2n} x_n \le b_2 \\ \vdots \\ y_m:\ a_{m1} x_1 + a_{m2} x_2 + \cdots + a_{mn} x_n \le b_m \\ x_1, \ldots, x_n \ge 0 \end{cases}$$

$x_i = x_i^+ - x_i^-$
$x_i^+ \ge 0$
$x_i^- \ge 0$

$$\begin{cases} x_1 + 2x_2 \ge 10 \\ x_1 + 2x_2 \le 10 \\ -x_1 - 2x_2 \le -10 \iff x_1 + 2x_2 \ge 10 \end{cases}$$

THE DUAL OF THE ABOVE LP IS

$$\begin{cases} \min\ b_1 y_1 + b_2 y_2 + \cdots + b_m y_m \\ x_1:\ a_{11} y_1 + a_{21} y_2 + \cdots + a_{m1} y_m \ge c_1 \\ x_2:\ a_{12} y_1 + a_{22} y_2 + \cdots + a_{m2} y_m \ge c_2 \\ \vdots \\ x_n:\ a_{1n} y_1 + a_{2n} y_2 + \cdots + a_{mn} y_m \ge c_n \\ y_1, \ldots, y_m \ge 0 \end{cases}$$

IF $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$, $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$, $c = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$

AND $\underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $\underline{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$, THEN

PRIMAL: $\begin{cases} \max\ c^T \underline{x} \\ A\underline{x} \le b \\ \underline{x} \ge 0 \end{cases}$ AND DUAL: $\begin{cases} \min\ b^T \underline{y} \\ A^T \underline{y} \ge c \\ \underline{y} \ge 0 \end{cases}$

RECALL THAT $(N\,H)^T = H^T N^T$. ✓✓

WEAK DUALITY THEOREM:
IF $\underline{x}$ IS A FEASIBLE SOLUTION TO THE PRIMAL, AND
IF $\underline{y}$ " " " " " " DUAL, THEN

$$\text{PRIMAL}(\underline{x}) = c^T \underline{x} \le b^T \underline{y} = \text{DUAL}(\underline{y}) .$$

FEASIBILITY OF $\underline{y}$ FOR DUAL

P: $c^T \underline{x} = \underline{x}^T c \le \underline{x}^T A^T \underline{y} = (A\underline{x})^T \underline{y} \le b^T \underline{y}$. ☐

FEASIBILITY OF $\underline{x}$ FOR PRIMAL

STRONG DUALITY THEOREM: IF $\underline{x}$ IS OPTIMAL FOR THE PRIMAL, AND $\underline{y}$ IS OPTIMAL FOR THE DUAL, (BOTH THE PRIMAL AND THE DUAL HAVE OPTIMAL FINITE SOLUTIONS),

$$c^T \underline{x} = b^T \underline{y} .$$