CSE 201 2-25-20

Matroids & the greedy algorithm:

Defn

A Matroid is a Pair $M = (S, \mathbb{I})$

where

(1) S is a finite, non-empty set and $\mathbb{I} \subseteq \mathbb{P}(S)$. The members of $\mathbb{I}$ are called independent sets, and those $\mathbb{P}(S) - \mathbb{I}$ are called Dependent sets.

(2) <u>Hereditary Property</u>:

If $B \in \mathcal{I}$ and $A \subseteq B$, then $A \in \mathcal{I}$.

(3) <u>Exchange Property</u>:

if $A \in \mathcal{I}$ & $B \in \mathcal{I}$ and $|A| < |B|$, then there exists $x \in B - A$ such that $A \cup \{x\} \in \mathcal{I}$.

<u>Note</u>: (2) implies that $\emptyset \in \mathcal{I}$ (Provided $\mathcal{I}$ is itself non-empty.)

Ex (Matric Matroids)

let P be an $n \times m$ matrix,

let $S = \{$rows of $P\}$,

~~[illegible]~~ considered as vectors in $\mathbb{R}^m$.

Then $|S| = n$. let

$$\mathcal{I} = \{A \subseteq S \mid A \text{ is linearly independent}\}$$

obviously S is finite & non-empty, and $\mathcal{I} \subseteq \mathcal{P}(S)$. Properties (2) & (3) are facts of linear algebra.

Ex. (Graphic matroids)

let $G=(V,E)$ be a graph, $E\neq\emptyset$

and $S=E$. let

$$\mathcal{I}=\{A\subseteq S \mid (V,A) \text{ is acyclic}\}$$

$$=\{\text{spanning forests in } G\}$$

(1) is obvious

(2) is also satisfied: cannot create a cycle by removing edges.

<u>Proof of (3)</u>:

let $A,B\in\mathcal{I}$ and suppose $|A|<|B|$.

so (V, A) and (V, B) are spanning forests in G.

<u>claim</u>: (V, A) contains $|V| - |A|$ trees

<u>pf.</u>

suppose (V, A) contains m trees: $T_i = (V_i, E_i)$ $(1 \le i \le m)$. Then

$$|E_i| = |V_i| - 1$$

so

$$|A| = \sum_{i=1}^{m} |E_i| = \sum_{i=1}^{m} (|V_i| - 1)$$

$$= \sum_{i=1}^{m} |V_i| - m = |V| - m$$

$\therefore \quad m = |V| - |A| \qquad \square$

likewise (V, B) contains exactly $|V| - |B|$ trees.

Now

$$|A| < |B| \Rightarrow |V| - |A| > |V| - |B|$$

$\therefore$ # trees in (V, B) < # trees in (V, A).

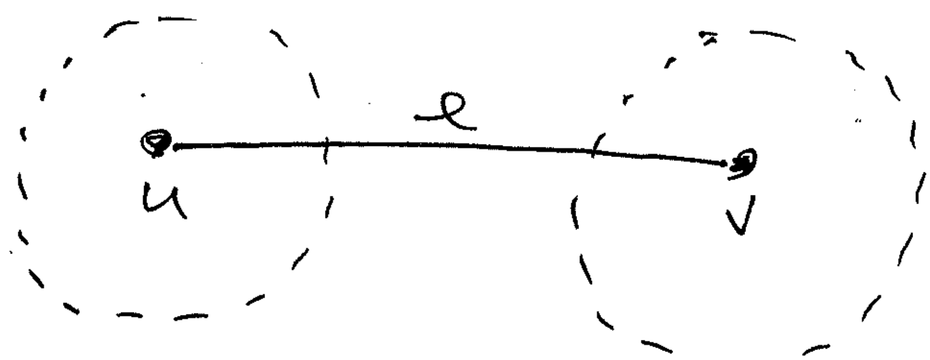
This implies that (V, B) contains a tree T which has vertices in 2 distinct trees of (V, A).

[Proof: if not, then every tree in (V, B) would be a subtree of a tree in (V, A) and every tree of (V, A) would contain at least 1 tree of (V, B). Hence

$$\#\text{trees}(V, B) \geq \#\text{trees}(V, A) \quad \times .]$$

Since T is connected, it contains an edge $e = \{u, v\}$ with ends in different trees of (V, A):

$$e \in E(T) \subseteq B$$



Distinct trees of (V, A)

By adding e to A, we do not create a cycle, i.e.

$$(V, A \cup \{e\})$$

is another sp. forest in G.

$\therefore$ $A \cup \{e\} \in \mathcal{I}$. ∎

<u>Defn</u>

Given $M = (S, \mathcal{I})$ a matroid and $A \in \mathcal{I}$, we call $x \in S$ an <u>extension</u> of A iff

$$A \cup \{x\} \in \mathcal{I}.$$

<u>Defn</u>

$A \in \mathbb{I}$ is called <u>maximal</u> iff it has no extensions.

(equivalently: no superset of A is independent.) i.e.

$$\forall x \in S - A : \; A \cup \{x\} \notin \mathbb{I}$$

A maximal independent set is also called a <u>Base</u> of the matroid.

Ex. in a matric matroid

A Base is a basis for the row space of P $(n \times m)$

Ex. in a graphic matroid

A base is the edge set of a sp. tree in G.

Recall: if $D \subseteq S$ and $D \notin \mathbb{I}$ we call D dependent.

Defn $C \subseteq S$ is called a cycle iff it is a minimal dependent set i.e. C is dependent, but it has no dependent subsets.

Ex. a cycle in a graphic matroid is the edge set of a cycle in G.

Theorem

Any two bases in a matroid have same cardinality.

Def: this cardinality is called the rank of the matroid: $\text{rank}(M)$.

<u>Proof.</u>

Let $A, B \in \mathbb{I}$ be bases in M, and suppose, to get a X, that $|A| < |B|$. By (3), there exists $x \in B - A$ such that $A \cup \{x\} \in \mathbb{I}$, contradicting that A is maximal. Hence $|A| \geq |B|$. Likewise $|A| \leq |B|$ by a similar argument. So $|A| = |B|$. ∎

Defn

A <u>weighted matroid</u> $M = (S, \mathcal{I})$ is a matroid with a weight function on S:

$$w : S \longrightarrow \mathbb{R}^+ = (0, \infty)$$

extend to subsets $A \subseteq S$:

$$w(A) = \sum_{x \in A} w(x)$$

a set $A \in \mathcal{I}$ is <u>optimal</u> iff it has maximum weight of all independent sets.

<u>note:</u>

A optimal $\Longrightarrow$ A maximal

why? if A optimal but not maximal, then $\exists\, x \in S - A$ s.t. $A \cup \{x\} \in \mathcal{I}$, and

$$w(A \cup \{x\}) = w(A) + w(x) > w(A) \quad \text{\Large ⨳}.$$

(using $w(x) > 0$).

Problem

Given a weighted matroid $M = (S, \mathcal{I})$ & $w : S \to \mathbb{R}^+$, find an optimal subset of S.

i.e. find $A \subseteq S$ s.t.

- $A \in \mathcal{I}$
- $w(A) \geq w(B)$ for all $B \in \mathcal{I}$.

Fact.
Many optimization problems are 'equivalent' to this on.

Any problem equiv. to this can be solved by a greedy algorithm.

<u>Ex.</u> MWST Problem in a weighted, connected graph $G=(V,E,w)$

How?

Consider the matric matroid

$$M_G = (E(G), \{\text{sp. forests in } G\})$$

let $w': E(G) \to \mathbb{R}^+$ be defined by

$$w'(e) = w_0 - w(e)$$

where w_0 is a fixed number greater than weight of all edges in G.

So $w'(e) > 0$ for all $e \in E(G)$.

Recall a maximal ind. set (Base) in M_G, $A \subseteq E(G)$ corresponds to a sp. tree in G.

Note

$$w'(A) = \sum_{e \in A} w'(e)$$

$$= \sum_{e \in A} (w_0 - w(e))$$

$$= |A| \cdot w_0 - \sum_{e \in A} w(e)$$

$$= |A| \cdot w_0 - w(A)$$

$$= (|V|-1)v_0 - w(A)$$

so $w'(A) + w(A) = (|V|-1)w_0$

Thus any sp. tree that maximizes $w'(\cdot)$ necessarily minimizes $w(\cdot)$, and conversely

"The greedy algorithm"

$(S, \mathcal{I})$
$\parallel$

Greedy (M, w)

1. $A = \phi$
2. sort S by descending weights
3. for each $x \in S$ // in order
4. if $A \cup \{x\} \in \mathcal{I}$
5. $A = A \cup \{x\}$
6. return A

Theorem

Greedy (M, w) returns an independent set $A \subseteq S$ of maximum weight.

Proof see Brassard & Bratley.

Problem

A unit time task is a job requiring 1 unit of time.

Given $S = \{a_1, a_2, \ldots, a_n\}$ of n unit-time tasks, a schedule for S is a permutation of S giving execution order.

Assume any schedule begins at time $t = 0$

Ex. $S = \{a_1, a_2, a_3\}$