Advanced Graph Algorithms Jan-Apr 2014

Lecture 13 — February, 1 2014

Lecturer: Saket Saurabh *Scribe: Sanjukta Roy*

1 Overview

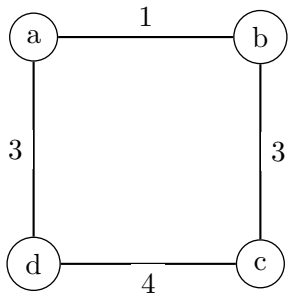
In this lecture we learn what is Matroid, the connection between greedy algorithms and matroids. We also look at some examples of Matroids e.g., Linear Matroids, Graphic Matroids etc.

2 Matroids

2.1 A Greedy Approach

Let G = (V, E) be a connected undirected graph and let $w : E \to R^{\geq 0}$ be a weight function on the edges. For MWST Kruskal's so-called greedy algorithm works.

Consider Maximum Weight Matching problem.



Application of the greedy algorithm gives (d,c) and (a,b). However, (d,c) and (a,b) do not form a matching of maximum weight.

It is obviously not true that such a greedy approach would lead to an optimal solution for any combinatorial optimization problem. It turns out that the structures for which the greedy algorithm does lead to an optimal solution, are the **matroids**.

2.2 Matroids

Definition 1. *A pair M = (E, $\mathcal{I}$), where E is a ground set and $\mathcal{I}$ is a family of subsets (called independent sets) of E, is a matroid if it satisfies the following conditions:*

(I1) $\phi \in \mathcal{I} or \mathcal{I} = \emptyset$.

(I2) If $A' \subseteq A$ and $A \in \mathcal{I}$ then $A' \in \mathcal{I}$.

(I3) If A, $B \in \mathcal{I}$ and $|A| < |B|$, then $\exists e \in (B \setminus A)$ such that $A \cup \{e\} \in \mathcal{I}$.

The axiom (I2) is also called the hereditary property and a pair M = (E, $\mathcal{I}$) satisfying (I1) and (I2) is called hereditary family or set-family.

An inclusion wise maximal set of $\mathcal{I}$ is called a basis of the matroid. Using axiom (I3) it is easy to show that all the bases of a matroid have the same size. This size is called the rank of the matroid M , and is denoted by rank(M).

2.2.1 Matroids and Greedy Algorithm

Theorem 2. *A set family M = (E, $\mathcal{I}$) is a matroid if and only if the greedy algorithm leads to a set Y in $\mathcal{I}$ of maximum weight w(Y), for each weight function $w : E \to R^{\geq 0}$.*

Proof. Our proof follows that of [2]
Forward direction : If (I1),(I2), (I3) holds then greedy algorithm finds an optimal solution, that is greedy algorithm works for matroids.

We call a Y $\in \mathcal{I}$ good if it is contained in a maximum-weight basis. We show that if Y is *good* and x $\in E \setminus$ Y gives Y $\cup$ {x}$\in \mathcal{I}$ with $w(x)$ as large as possible, then Y $\cup$ {x} is also good.

Let B be the maximum-weight basis with Y $\subset$ B. If $x \in B$ then $Y \cup \{x\}$ is clearly good.
If $x \notin$ B, $B \cup \{x\} \notin \mathcal{I}$.

case 1: If $|B| = |Y \cup \{x\}|$ then $w(Y \cup \{x\}) \geq w(B)$. Then $Y \cup \{x\} = B'$ is a basis and $w(B') = w(B)$. Since B is the max weight basis, B' is also a max weight basis containing Y $\cup\{x\}$. So $Y \cup \{x\}$ is good.

case 2: If $|B| > |Y \cup \{x\}|$. Therefore we can extend $Y \cup \{x\}$ by the elements in B to get some basis B'.
$\therefore Y \cup \{x\} \subseteq B' \subsetneq B \cup \{x\}$. By (I3), $|B'| = |B|$, so we must have

$$B' = (B \cup \{x\}) \setminus \{x'\}$$

for some $x' \in B \setminus Y$. Then x must have been chosen over that x', so $w(x) \geq w(x')$ and $w(B') \geq w(B)$. Hence B' is a maximum-weight basis containing Y $\cup\{x\}$, so Y $\cup\{x\}$ is good.

Backward Direction: we show that if a set family (E, $\mathcal{I}$) does not satisfy (I2), then there will be weights w for which the greedy algorithm does not return a maximum weight element of $\mathcal{I}$. There must be an $A \notin \mathcal{I}$ and $b \notin A$ with $A \cup \{b\} \in \mathcal{I}$. Define w by

$$w(x) = \begin{cases} 1 & if x = b \\ 2 & if x \in A \\ -1 & otherwise \end{cases}$$

Then $A \cup \{x\}$ is the unique maximum weight element of $\mathcal{I}$. But the greedy algorithm can never get to $A \cup \{x\}$, since before it gets to x, it must have considered all $a \in A$, and it must have rejected

at least one of them because $A \notin \mathcal{I}$.
Finally we show that if a set family (E, $\mathcal{I}$) satisfies (I2) but not (I3), then there will be weights w for which the greedy algorithm does not return a maximum weight element of $\mathcal{I}$. There must be $Y, Z \in \mathcal{I}$ with $|Y| < |Z|$ but $Y \cup \{x\} \notin \mathcal{I}$ for any $x \in Z \setminus Y$. Define w by

$$w(x) = \begin{cases} 1 & if x \in Z \setminus Y \\ 1+\epsilon & if x \in Y, \epsilon > 0 \\ 0 & otherwise \end{cases}$$

After the first $|Y|$ iterations, the greedy algorithm will reach Y , since all its subsets are in $\mathcal{I}$ by (I2). But then no more nonzero elements can be picked, since $Y \cup \{x\} \notin \mathcal{I}$ for $x \in Z \setminus Y$. So the maximum weight found by the greedy algorithm is $|Y| \times (1+\epsilon)$. But $Z \in \mathcal{I}$ and
$w(Z) \geq |Z \cap Y| \times (1+\epsilon) + |Z \setminus Y| \times 1 \geq |Z| \geq |Y| + 1$.
So if we take $\epsilon > 0$ small enough so that $|Y|(1+\epsilon) < |Y| + 1$, then the greedy algorithm does not find the maximum weight set in $\mathcal{I}$.

□

2.3 Examples of Matroids

Uniform Matroid A pair M = (E, $\mathcal{I}$) over an n-element ground set E, is called a uniform matroid if the family of independent sets is given by

$$\mathcal{I} = \{A \subseteq E \mid |A| \leq k, \text{where k is some constant.}\}$$

This matroid is also denoted as $U_{n,k}$.
Eg: $E = \{1, 2, 3, 4, 5\}$ and $k = 2$ then $\mathcal{I}$ = {}, {1}, {2}, {3}, {4}, {5}, {1, 2}, {1, 3}, {1, 4}, {1, 5}, {2, 3}, {2, 4}, {2, 5}, {3, 4}, {3, 5}, {4, 5}

Partition Matroid A partition matroid M = (E, $\mathcal{I}$) is defined by a ground set $E = E_1 \biguplus E_2 \biguplus \cdots \biguplus E_\ell$ and by ℓ non-negative integers $k_1, \cdots, k_\ell$. A set $X \subseteq E$ is independent if and only if $|X \cap E_i| \leq k_i$ for all $i \in \{1, \cdots, \ell\}$. That is,

$$\mathcal{I} = \{X \subseteq E \mid |X \cap E_i| \leq k_i, i \in \{1, ..., \ell\}\}.$$

If X, Y $\in \mathcal{I}$ and $|Y| > |X|$, there must exist i such that $|Y \cap E_i| > |X \cap E_i|$ and this means that adding any element e in $E_i \cap (Y \setminus X)$ to X will maintain independence. M in general would not be a matroid if E_i were not disjoint. E.g., E_1 = {1, 2} and E_2 = {2, 3} and $k_1 = 1$ and $k_2 = 1$ then both Y = {1, 3} and X = {2} have at most one element of each Ei but one cant find an element of Y to add to X.

Graphic Matroid Given a graph G, a graphic matroid is defined as M = (E, $\mathcal{I}$) where
E = E(G) edges of G are elements of the matroid

$$\mathcal{I} = \{F \subseteq E(G) \mid F \text{ is a forest in the graph G}\}$$

Claim 3. *Graphic matroids are matroid*

Proof. From definition of graphic matroids, bases of this matroids are the spanning forests.

To prove set $\mathcal{I}$ is indeed defines a matroid we need to show it satisfies (I1), (I2) and (I3). (I1) holds trivially. Since every subset of a forest is a forest, hereditary property, (I2) holds. Suppose towards contradiction (I3) does not hold. That is $|A| < |B|$ but A can not be extended by any element of $B \setminus A$. That is every edge in $B \setminus A$ creates a cycle in A.
If required, add isolated vertices to G[A] and G[B] so that they have same set of vertices, say V. Now, since $|B| > |A|$ then no. of trees in $G(B, V) \leq$ no. of of trees in $G(A, V)$

Case 1: No of trees in B = No of trees in A.That means $\exists edge\{u, v\}$ in $B \setminus A$ where either u or v is isolated vertex in $A \setminus B$. Then adding this edge $\{u, v\}$ will not create a cycle in A i.e., A can be extended.

Case 2: No of trees in B $<$ No of trees in A. Then there must be a path, p (with edges of B) in G connecting two trees of $G(A, V)$. Then $\exists$ atleast one edge, $e \in p$ and $e \notin A$. Since e will connect two disconnected components in A, it will not form a cycle in A. Therefore A can be extended by e. Hence property (I3) holds. □

Co-Graphic Matroid Given a graph G, a co-graphic matroid is defined as M = (E, $\mathcal{I}$) where and E = E(G) edges of G are elements of the matroid

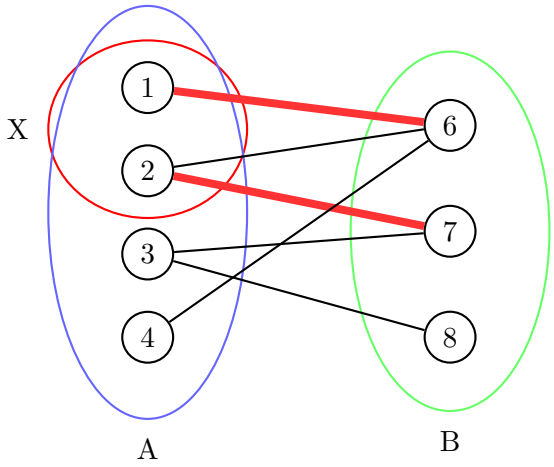
$$\mathcal{I} = \{S \subseteq E(G) \mid G \setminus S \text{ is connected}\}$$

Claim 4. *Co-Graphic Matroids are Matroid*

Proof. Property (I1) holds trivially.
If removing S does not disconnect the graph G then removing any subset of S will not disconnect G either. So hereditary property, (I2) holds.
Suppose (I3) does not holds. That is $|A| < |B|$ but A can not be extended by any element of $B \setminus A$.
$\Rightarrow$ every edge in $B \setminus A$, if removed, disconnects $G \setminus A$
$\Rightarrow$ every edge in $B \setminus A$ is a bridge in $G \setminus A$. But $G \setminus B$ is connected, since B $\in I$. Therefore no edge in B is a bridge in G.
Let e $\in B \setminus A$ disconnects $G \setminus A$ in X and Y where X and Y are two components of $G \setminus (A \cup \{e\})$. Therefore there was atleast another edge e' between X and Y which keeps $G \setminus B$ connected and $e' \in A \setminus B$.
We claim that for every edge $e \in B \setminus A$ there is a different $e' \in A \setminus B$.
Otherwise assume two different edges $e_1, e_2 \in B \setminus A$, e_1, e_2 are bridges in $G \setminus A$ and both of them are supported by the edge $e' \in A \setminus B$ so that $G \setminus B$ is not disconnected. Let, e_1 disconectes $G \setminus A$ in X_1 and Y_1, e_2 disconectes Y_1 in X_2 and Y_2 ,then e' must has endpoints in X_1 and Y_2. But then removing both e_1, e_2 disconnects X_2 from G. But $e_1, e_2 \in B \in \mathcal{I}$.
Therefore $|A| = |B|$, but $|A| < |B|$. So we reach a contradiction. Hence A can be extended. So (I3) holds. □

Transversal Matroids Let G be a bipartite graph with the vertex set V (G) being partitioned as A and B. The transversal matroid $M = (E, \mathcal{I})$ of G has E = A as its ground set,

$$\mathcal{I} = \{X \mid X \subseteq A, \text{there is a matching that saturates } X\}$$



In the next lecture we will give a proof that transversal matroids are indeed matroids using representability of this matroid.

2.4 Matroid Representation

Definition 5. *A matroid $M = (E, \mathcal{I})$ is representable over a field $\mathbb{F}$ if there exist vectors in $\mathbb{F}^\ell$ that correspond to the elements such that the linearly independent sets of vectors precisely correspond to independent sets of the matroid.*

Representation is required because we need compact representation to for the family of independent sets.Also to be able to test easily whether a set belongs to the family of independent sets.

Linear Matroid Let A be a matrix over an arbitrary field $\mathbb{F}$ and let E be the set of columns of A. Given A we define the matroid M = (E, $\mathcal{I}$) as follows.
A set X $\subseteq$ E is independent (that is X $\in \mathcal{I}$) if the corresponding columns are linearly independent over $\mathbb{F}$.

$$A = \begin{pmatrix} * & * & * & \cdots & * \\ * & * & * & \cdots & * \\ * & * & * & \cdots & * \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ * & * & * & \cdots & * \end{pmatrix} \quad * \text{ are elements of field } \mathbb{F}$$

Linear Matroids are Matroids

Proof. Property (I1) holds trivially. If $Y \subseteq E$ (columns of A) contains all linearly independent columns then all subsets of Y also contains linearly independent columns. So hereditary property holds.
If X is another set of linearly independent columns such that $|X| < |Y|$ and $\forall c \in Y \setminus X, X \cup \{c\} \notin \mathcal{I}$

Then c = $\sum_{x_i \in X} \alpha_i x_i$ then $\forall y \in Y$, y = $\sum_{x_i \in X} \beta_i x_i$. But we assumed $Y \in \mathcal{I}$. That gives a contradiction. Hence it satisfies (I3). □

2.4.1 Linear Matroids and Representable Matroids

If a matroid can be defined by a matrix A over a field $\mathbb{F}$, then we say that the matroid is representable over $\mathbb{F}$.
Let E = $\{e_1, \cdots, e_m\}$ and ℓ be a positive integer.

$$\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ \vdots \\ \ell \end{array} \begin{array}{c} \begin{array}{ccccc} e_1 & e_2 & e_3 & \cdots & e_\ell \end{array} \\ \begin{pmatrix} * & * & * & \cdots & * \\ * & * & * & \cdots & * \\ * & * & * & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ * & * & * & \cdots & * \end{pmatrix}_{\ell \times m} \end{array}$$

A matroid M = (E, $\mathcal{I}$) is called representable or linear if it is representable over some field $\mathbb{F}$.
Rank of the matroid is the rank of this matrix.
Let M = (E, $\mathcal{I}$) be linear matroid and Let E = $\{e_1, \cdots, e_m\}$ and d=rank(M). We can always assume (using Gaussian Elimination) that the matrix has following form:

$$\left[\ I_{d\times d} \mid D\ \right]_{d\times m}$$

Here $I_{d\times d}$ is a $d \times d$ identity matrix.

References

[1] Wiley-Interscience Series, *Combinatorial Optimization*, William J. Cook, William H. Cunningham, William R. Pueyblank, Alexander Schrijver

[2] *Matroid Theory*, James G. Oxley