EE595A – Submodular functions, their optimization and applications – Spring 2011

Prof. Jeff Bilmes

University of Washington, Seattle
Department of Electrical Engineering
Winter Quarter, 2011
`http://ee.washington.edu/class/235/2011wtr/index.html`

Lecture 3 - April 6th, 2011

Announcements

- Weekly Office Hours: Wednesdays, 12:30-1:30pm, 10 minutes after class on Wednesdays.
- Today: Matroids

Scratch Paper

$\min\ w(T)$

s.t. T is a spanning tree

where $w \in \mathbb{R}^E$

Scratch Paper

Scratch Paper

Roots

- Submodular functions have their roots in matroid theory and in lattice theory.
- We'll spend the next few lectures reviewing key concepts from these areas.

matrices, and independence in linear space

Consider the following 3 × 8 matrix.

$$\begin{array}{cc} & \begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{pmatrix} 0 & 0 & 1 & 1 & 2 & 1 & 3 & 1 \\ 0 & 1 & 1 & 0 & 2 & 0 & 2 & 4 \\ 1 & 1 & 1 & 0 & 0 & 3 & 1 & 5 \end{pmatrix} \end{array} = \begin{array}{c} \begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array} \\ \begin{pmatrix} | & | & | & | & | & | & | & | \\ x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 \\ | & | & | & | & | & | & | & | \end{pmatrix} \end{array} \tag{1}$$

- Obviously, the rank of the matrix can't be more than 3, and considering the first three columns we see the rank is 3.
- Some vectors are linearly dependent on others. I.e., $x_4 = x_3 - x_2$, $x_5 = 2x_3 - 2x_1$, $x_6 = x_8 - 4x_2 - x_1$. $(x_4, x_3 x_2)$ are dependent
- While other vectors are linearly independent. E.g., $\{x_4, x_5, x_6\}$, $\{x_4, x_7, x_8\}$, etc.

matrices, and independence in linear space

$$\begin{array}{c|cccccccc|}
 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
1 & 0 & 0 & 1 & 1 & 2 & 1 & 3 & 1 \\
2 & 0 & 1 & 1 & 0 & 2 & 0 & 2 & 4 \\
3 & 1 & 1 & 1 & 0 & 0 & 3 & 1 & 5
\end{array}
=
\begin{array}{cccccccc}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
| & | & | & | & | & | & | & | \\
x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 \\
| & | & | & | & | & | & | & |
\end{array} \tag{2}$$

- The rank function indicates the number of linearly independent column vectors given a set of column vectors.
- Consider the column index set $V = \{1, 2, \ldots, 8\}$. Then any subset $A \subseteq V$ corresponds to a set of column vectors.
- We can define the rank function on subsets of V - the rank function of A is the number of linearly independent column vectors within A.
- E.g., $\text{rank}(\{4, 5, 6\}) = \text{rank}(\{4, 7, 8\}) = \text{rank}(\{1, 2, 3\}) = 3$ while $\text{rank}\{1, 2\} = \text{rank}(\{1, 4, 6\}) = 2$.

matrices, and independence in linear space

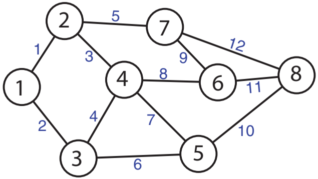
$$\begin{array}{c|cccccccc} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ \hline 1 & 0 & 0 & 1 & 1 & 2 & 1 & 3 & 1 \\ 2 & 0 & 1 & 1 & 0 & 2 & 0 & 2 & 4 \\ 3 & 1 & 1 & 1 & 0 & 0 & 3 & 1 & 5 \end{array} = \begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ \left(\begin{matrix} | \\ x_1 \\ | \end{matrix} \right. & \begin{matrix} | \\ x_2 \\ | \end{matrix} & \begin{matrix} | \\ x_3 \\ | \end{matrix} & \begin{matrix} | \\ x_4 \\ | \end{matrix} & \begin{matrix} | \\ x_5 \\ | \end{matrix} & \begin{matrix} | \\ x_6 \\ | \end{matrix} & \begin{matrix} | \\ x_7 \\ | \end{matrix} & \left. \begin{matrix} | \\ x_8 \\ | \end{matrix} \right) \end{array} \tag{3}$$

- In general, $\text{rank}(A) \leq |A|$, and vectors in A are linearly independent if and only if $\text{rank}(A) = |A|$.
- If A, B are such that $\text{rank}(A) = |A|$ and $\text{rank}(B) = |B|$, with $\underline{|A| < |B|}$, then the space spanned by B is greater, and we can find a vector in B that is linearly independent of the space spanned by vectors in A.
- In other words, $\exists$ an $b \in B$ such that $\text{rank}(A \cup \{b\}) = |A| + 1$.

Spanning trees / forests.

- We are given a graph $G = (V, E)$, and consider the edges $E = E(G)$ as an index set.
- Consider the incidence matrix of (undirected graph) G, which is the matrix $\cancel{M}A_G = (a_{v,e})_{v \in V(G), e \in E(G)}$ where

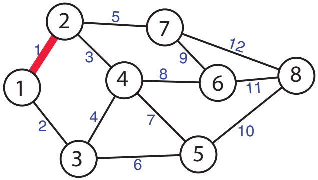
$$a_{v,e} = \begin{cases} 1 & \text{if } v \in e = (v_1, v_2) \\ 0 & \text{if } v \notin e \end{cases} \tag{4}$$



$$\begin{array}{c|cccccccccccc} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ \hline 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 3 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 4 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 5 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 7 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \\ 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{array} \tag{5}$$

Spanning trees

- We can consider edge-induced subgraphs and the corresponding matrix columns.

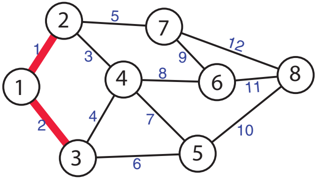


$$\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{array}\begin{array}{c} 1 \\ \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}\end{array} \qquad (6)$$

Here, $\text{rank}(\{x_1\}) = 1$.

Spanning trees

- We can consider edge-induced subgraphs and the corresponding matrix columns.

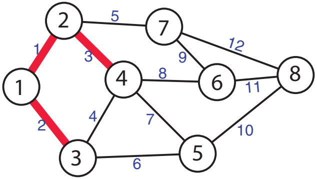


$$\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{array} \begin{array}{c} \begin{array}{cc} 1 & 2 \end{array} \\ \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \end{array} \tag{7}$$

Here, $\text{rank}(\{x_1, x_2\}) = 2$.

Spanning trees

- We can consider edge-induced subgraphs and the corresponding matrix columns.

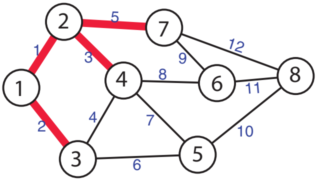


$$\begin{array}{c} \\ 1\\2\\3\\4\\5\\6\\7\\8 \end{array}\begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{array} \tag{8}$$

Here, $\text{rank}(\{x_1, x_2, x_3\}) = 3$.

Spanning trees

- We can consider edge-induced subgraphs and the corresponding matrix columns.

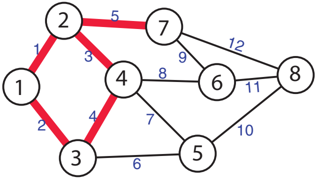


$$\begin{array}{c|cccc} & 1 & 2 & 3 & 5 \\ 1 & 1 & 1 & 0 & 0 \\ 2 & 1 & 0 & 1 & 1 \\ 3 & 0 & 1 & 0 & 0 \\ 4 & 0 & 0 & 1 & 0 \\ 5 & 0 & 0 & 0 & 0 \\ 6 & 0 & 0 & 0 & 0 \\ 7 & 0 & 0 & 0 & 1 \\ 8 & 0 & 0 & 0 & 0 \end{array} \tag{9}$$

Here, $\text{rank}(\{x_1, x_2, x_3, x_5\}) = 4$.

Spanning trees

- We can consider edge-induced subgraphs and the corresponding matrix columns.

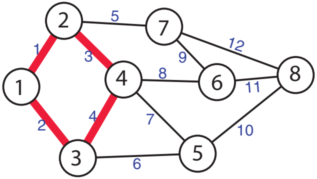


$$\begin{array}{c|ccccc} & 1 & 2 & 3 & 4 & 5 \\ \hline 1 & 1 & 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 & 1 \\ 3 & 0 & 1 & 0 & 1 & 0 \\ 4 & 0 & 0 & 1 & 1 & 0 \\ 5 & 0 & 0 & 0 & 0 & 0 \\ 6 & 0 & 0 & 0 & 0 & 0 \\ 7 & 0 & 0 & 0 & 0 & 1 \\ 8 & 0 & 0 & 0 & 0 & 0 \end{array} \tag{10}$$

Here, $\text{rank}(\{x_1, x_2, x_3, x_4, x_5\}) = 4$.

Spanning trees

- We can consider edge-induced subgraphs and the corresponding matrix columns.



$$\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{array}\begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array} \quad (11)$$

Here, $\text{rank}(\{x_1, x_2, x_3, x_4\}) = 3$ since $x_4 = x_3 + x_2 - x_1$.

Spanning trees

- In general, whenever the edges specify a cycle, there will be a linear dependence between the corresponding set of vectors in the matrix.
- This means that all forests in the graph correspond to a set of linearly independent column vectors in the matrix.
- Consider a "rank" function defined as follows: given a set of edges $A \subseteq E(G)$, the rank(A) is the size of the largest forest in the A-edge-induced subgraph of G.
- The rank of the entire graph then is then a spanning forest of the graph (spanning tree if the graph is connected).
- The rank of the graph is rank(G) $= |V| - k$ where k is the number of connected components of G.

$r(G)$

Spanning Tree Algorithms

- We are now given a positive edge-weighted connected graph $G = (V, E, w)$ where $w : E \to \mathbb{R}_+$ is a modular function the edges of the graph. The goal is to find the minimum spanning tree (MST) of the graph.
- There are several algorithms for this.

Algorithm 1: Borůvka's Algorithm

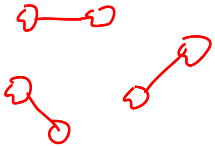
```
$F \leftarrow \emptyset$ ;
while $F$ is disconnected do
    forall the components $C_i$ of $F$ do
        $F \leftarrow F \cup \{e_i\}$ for $e_i =$ the min-weight edge out of $C_i$;
```

$\delta_G(C_i)$

T is a tree

$\text{cost}(T) = \sum_{e \in T} w(e)$)

$w \in \mathbb{R}^E$

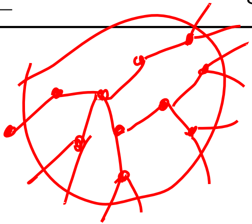


Spanning Tree Algorithms

- We are now given a positive edge-weighted graph $G = (V, E, w)$ where $w : E \to \mathbb{R}_+$ is a modular function the edges of the graph. The goal is to find the minimum spanning tree (MST) of the graph.
- There are several algorithms for this.

Algorithm 2: Jarník/Prim/Dijkstra Algorithm

```
$T \leftarrow \emptyset$ ;
while $T$ is not a spanning tree do
    $T \leftarrow T \cup \{e\}$ for $e =$
     the minimum weight edge extending the tree $T$ to a new vertex ;
```



Spanning Tree Algorithms

- We are now given a positive edge-weighted graph $G = (V, E, w)$ where $w : E \to \mathbb{R}_+$ is a modular function the edges of the graph. The goal is to find the minimum spanning tree (MST) of the graph.
- There are several algorithms for this.

Algorithm 3: Kruskal's Algorithm

```
Sort the edges so that $w(e_1) \le w(e_2) \le \cdots \le w(e_m)$ ;
$T \leftarrow (V(G), \emptyset) = (V, E)$ ;
for $i = 1$ to $m$ do
    if $E(T) \cup \{e_i\}$ does not create a cycle in $T$ then
        $E(T) \leftarrow E(T) \cup \{e\}$ ;
```

Spanning Tree Algorithms

- We are now given a positive edge-weighted graph $G = (V, E, w)$ where $w : E \to \mathbb{R}_+$ is a modular function the edges of the graph. The goal is to find the minimum spanning tree (MST) of the graph.
- There are several algorithms for this.
- These three algorithms are all guaranteed to find the optimal minimum spanning tree in (low order) polynomial time.
- All these algorithms are related to the "greedy" algorithm. I.e., "add next whatever looks best".
- These algorithms will also always find a basis (a set of linearly independent vectors that span the underlying space) in the matrix example we saw earlier.
- The above are all examples of a matroid, which is the fundamental reason why the greedy algorithms work.

Matroid

- Matroids abstract the notion of linear independence of a set of vectors to general algebraic properties.
- In a matroid, there is an underlying ground set, say E (or V), and a collection of subsets of E that correspond to independent elements.
- There are many definitions of matroids that are mathematically equivalent, we'll see some of them here.

Independence System

$E = \{1,2,3,4\}$, $\mathcal{I} = \{\{1\}, \emptyset, \{1,2\}, \{1\,2\,4\}\}$

$\{1,2\} \in \mathcal{I}$, $\{2\} \subset \{1,2\}$

$\{2\} \notin \mathcal{I}$

Definition 4.1 (set system)

A (finite) ground set E and a set of subsets of E, $\emptyset \neq \mathcal{I} \subseteq 2^E$ is called a set system, notated $(E, \mathcal{I})$.

Definition 4.2 (independence system)

A set system $(E, \mathcal{I})$ is an independence system if

$$\emptyset \in \mathcal{I} \tag{I1}$$

$$\forall I \in \mathcal{I}, J \subset I \Rightarrow J \in \mathcal{I} \text{ (called “down monotone” or “down closed”)} \tag{I2}$$

Independence System

$$\begin{array}{c@{}c} & \begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{pmatrix} 0 & 0 & 1 & 1 & 2 & 1 & 3 & 1 \\ 0 & 1 & 1 & 0 & 2 & 0 & 2 & 4 \\ 1 & 1 & 1 & 0 & 0 & 3 & 1 & 5 \end{pmatrix} \end{array} = \begin{array}{c} \begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array} \\ \begin{pmatrix} | & | & | & | & | & | & | & | \\ x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 \\ | & | & | & | & | & | & | & | \end{pmatrix} \end{array} \tag{12}$$

- Given any set of linearly independent vectors A, any subset $B \subset A$ will also be linearly independent.
- Given any forest G_f that is an edge-induced sub-graph of a graph G, any sub-graph of G_f is also a forest.
- So these both constitute independence systems.

Matroid

if $J \in \mathcal{I}$, J is said to be an "independent set"

Independent set definition of a matroid is perhaps most natural.

Definition 4.3 (Matroid)

A set system $(E, \mathcal{I})$ is a Matroid if

(I1) $\emptyset \in \mathcal{I}$

(I2) $\forall I \in \mathcal{I}, J \subset I \Rightarrow J \in \mathcal{I}$

independence system

(I3) $\forall I, J \in \mathcal{I}$, with $|I| = |J| + 1$, then there exists $x \in \cancel{J \setminus I}$ $I \setminus J$ such that $J \cup \{x\} \in \mathcal{I}$

Matroid

Slight modification (non unit increment) that is equivalent.

Definition 4.4 (Matroid-II)

A set system $(E, \mathcal{I})$ is a Matroid if

(I1') $\emptyset \in \mathcal{I}$

(I2') $\forall I \in \mathcal{I}, J \subset I \Rightarrow J \in \mathcal{I}$

(I3') $\forall I, J \in \mathcal{I}$, with $|I| > |J|$, then there exists $x \in$ ~~$J \setminus I$~~ $I \setminus J$ such that ~~I~~ $J \cup \{x\} \in \mathcal{I}$

Note (I1)=(I1'), (I2)=(I2'), and we get (I3)≡(I3') using induction.

matrices, and independence in linear space

$$\begin{array}{c|cccccccc} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ \hline 1 & 0 & 0 & 1 & 1 & 2 & 1 & 3 & 1 \\ 2 & 0 & 1 & 1 & 0 & 2 & 0 & 2 & 4 \\ 3 & 1 & 1 & 1 & 0 & 0 & 3 & 1 & 5 \end{array} = \begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ \hline | & | & | & | & | & | & | & | \\ x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 \\ | & | & | & | & | & | & | & | \end{array} \tag{13}$$

- In general, $\text{rank}(A) \leq |A|$, and vectors in A are linearly independent if and only if $\text{rank}(A) = |A|$.
- If A, B are such that $\text{rank}(A) = |A|$ and $\text{rank}(B) = |B|$, with $|A| < |B|$, then the space spanned by B is greater, and we can find a vector in B that is linearly independent of the space spanned by vectors in A.
- In other words, $\exists$ an $b \in B$ such that $\text{rank}(A \cup \{b\}) = |A| + 1$.

Matroids

- Given a matroid $M = (E, \mathcal{I})$, a subset $A \subseteq E$ is called independent if $A \in \mathcal{I}$ and otherwise A is called dependent.
- For $U \subseteq E$, a subset $B \subseteq U$ is called a base of U if B is inclusionwise maximally independent subset of U. That is, $B \in \mathcal{I}$ and there is no $Z \in \mathcal{I}$ with $B \subset Z \subseteq U$.
- If $U = E$, then a "base of E" is just called a base of the matroid M (this corresponds to a basis in a linear space).

Matroids

Proposition 4.5 in a matroid.

For any $U \subseteq E(M)$, any two bases of U have the same size. $(\overline{I3})$

- In matrix terms, given a set of vectors U, all sets of independent vectors that span the space spanned by U have the same size.
- In fact, under (I1),(I2), this condition is equivalent to (I3). Exercise: show this
- The common size of all the bases of U is called the rank of U, denoted $r_M(U)$ or just $r(U)$ when the matroid in equation is unambiguous.
- $r(E) = r_{(E,\mathcal{I})}$ is the rank of the matrix, and is the common size of all the bases of the matroid.

Matroids -rank

We can a bit more formally define the rank function this way.

Definition 4.6

The rank of a matroid is a function $r : 2^E \rightarrow \mathbb{Z}_+$ defined by

$$r(A) = \max\{|X| : X \subseteq A, X \in \mathcal{I}\} \tag{14}$$

- From the above, we immediately see that $r(A) \leq |A|$.
- Moreover, if $r(A) = |A|$, then $A \in \mathcal{I}$, meaning A is independent.
  (full rank)

Matroids-rank

Lemma 4.7

The rank function of a matroid is submodular, that is $r(A) + r(B) \geq r(A \cup B) + r(A \cap B)$

Proof.

1. Let $X \in \mathcal{I}$ be an inclusionwise maximal set with $X \subseteq A \cap B$

☐

Matroids-rank

Lemma 4.7

The rank function of a matroid is submodular, that is $r(A) + r(B) \geq r(A \cup B) + r(A \cap B)$

Proof.

1. Let $X \in \mathcal{I}$ be an inclusionwise maximal set with $X \subseteq A \cap B$
2. Let $Y \in \mathcal{I}$ be inclusionwise maximal set with $\underline{X \subseteq Y} \subseteq A \cup B$.

$\square$

Matroids-rank

Lemma 4.7

The rank function of a matroid is submodular, that is $r(A) + r(B) \geq r(A \cup B) + r(A \cap B)$

Proof.

1. Let $X \in \mathcal{I}$ be an inclusionwise maximal set with $X \subseteq A \cap B$
2. Let $Y \in \mathcal{I}$ be inclusionwise maximal with $X \subseteq Y \subseteq A \cup B$.
3. Since M is a matroid, we know that $r(A \cap B) = r(X) = |X|$, and $r(A \cup B) = r(Y) = |Y|$. Also, for any $U \in \mathcal{I}$, $r(A) \geq |A \cap U|$.

☐

Matroids-rank

Lemma 4.7

The rank function of a matroid is submodular, that is $r(A) + r(B) \geq r(A \cup B) + r(A \cap B)$

Proof.

1. Let $X \in \mathcal{I}$ be an inclusionwise maximal set with $X \subseteq A \cap B$
2. Let $Y \in \mathcal{I}$ be inclusionwise maximal with $X \subseteq Y \subseteq A \cup B$.
3. Since M is a matroid, we know that $r(A \cap B) = r(X) = |X|$, and $r(A \cup B) = r(Y) = |Y|$. Also, for any $U \in \mathcal{I}$, $r(A) \geq |A \cap U|$.
4. Then we have

$$r(A) + r(B) \geq |Y \cap A| + |Y \cap B| \tag{15}$$

□

Matroids-rank

$f(A) = |Y \cap A|$

Lemma 4.7

The rank function of a matroid is submodular, that is $r(A) + r(B) \geq r(A \cup B) + r(A \cap B)$

Proof.

1. Let $X \in \mathcal{I}$ be an inclusionwise maximal set with $X \subseteq A \cap B$
2. Let $Y \in \mathcal{I}$ be inclusionwise maximal with $X \subseteq Y \subseteq A \cup B$.
3. Since M is a matroid, we know that $r(A \cap B) = r(X) = |X|$, and $r(A \cup B) = r(Y) = |Y|$. Also, for any $U \in \mathcal{I}$, $r(A) \geq |A \cap U|$.
4. Then we have

$$r(A) + r(B) \geq |Y \cap A| + |Y \cap B| \tag{15}$$

$$= |Y \cap (A \cap B)| + |Y \cap (A \overset{\cup}{\cancel{\cap}} B)| \tag{16}$$

□

Matroids-rank

Lemma 4.7

The rank function of a matroid is submodular, that is $r(A) + r(B) \geq r(A \cup B) + r(A \cap B)$

Proof.

1. Let $X \in \mathcal{I}$ be an inclusionwise maximal set with $X \subseteq A \cap B$
2. Let $Y \in \mathcal{I}$ be inclusionwise maximal with $X \subseteq Y \subseteq A \cup B$.
3. Since M is a matroid, we know that $r(A \cap B) = r(X) = |X|$, and $r(A \cup B) = r(Y) = |Y|$. Also, for any $U \in \mathcal{I}$, $r(A) \geq |A \cap U|$.
4. Then we have

$$\begin{aligned} r(A) + r(B) &\geq |Y \cap A| + |Y \cap B| && (15) \\ &= |Y \cap (A \cap B)| + |Y \cap (A \overset{\checkmark}{\cancel{\cap}} B)| && (16) \\ &\geq |X| + |Y| = r(A \cap B) + r(A \cup B) \quad \checkmark && (17) \end{aligned}$$

☐

Matroids

(E, r)

In fact, we can use the rank of a matroid for its definition.

Theorem 4.8 (Matroid from rank)

Let E be a set and let $r : 2^E \to \mathbb{Z}_+$ be a function. Then $r(\cdot)$ defines a matroid with r being its rank function if and only if for all $A, B \subseteq E$:

(R1) $\forall A \subseteq E \;\; 0 \leq r(A) \leq |A|$ *(non-negative cardinality bounded)*

(R2) $r(A) \leq r(B)$ *whenever* $A \subseteq B \subseteq E$ *(monotone non-decreasing)*

(R3) $r(A \cup B) + r(A \cap B) \leq r(A) + r(B)$ *forall* $A, B \subseteq E$ *(submodular)*

- So submodularity and non-negative monotone non-decreasing, and unit increase is necessary and sufficient to define the matroid.
- Given above, unit increment (if $r(A) = k$, then either $r(A \cup \{v\}) = k$ or $r(A \cup \{v\}) = k + 1$) holds.

Matroids

In fact, we can use the rank of a matroid for its definition.

Theorem 4.8 (Matroid from rank)

Let E be a set and let $r : 2^E \to \mathbb{Z}_+$ be a function. Then $r(\cdot)$ defines a matroid with r being its rank function if and only if for all $A, B \subseteq E$:

(R1) $\forall A \subseteq E \quad 0 \leq r(A) \leq |A|$ *(non-negative cardinality bounded)*

(R2) $r(A) \leq r(B)$ *whenever* $A \subseteq B \subseteq E$ *(monotone non-decreasing)*

(R3) $r(A \cup B) + r(A \cap B) \leq r(A) + r(B)$ *forall* $A, B \subseteq E$ *(submodular)*

- From above, $r(\emptyset) = 0$. Let $v \notin A$, then by monotonicity and submodularity, $r(A) \leq r(A \cup \{v\}) \leq r(A) + r(\{v\})$ which gives only two possible values to $r(A \cup \{v\})$.

Matroids from rank

Proof of Theorem 4.8 (matroid from rank).

- Given a matroid $M = (E, \mathcal{I})$, we see its rank function as defined in Eq. 14 satisfies (R1), (R2), and, as we saw in Lemma 4.7, (R3) too.

...

Matroids from rank

Proof of Theorem 4.8 (matroid from rank).

- Given a matroid $M = (E, \mathcal{I})$, we see its rank function as defined in Eq. 14 satisfies (R1), (R2), and, as we saw in Lemma 4.7, (R3) too.
- Next, assume we have (R1), (R2), and (R3). Define $\mathcal{I} = \{X \subseteq E : r(X) = |X|\}$. We will show that $(E, \mathcal{I})$ is a matroid.

. . .

Matroids from rank

Proof of Theorem 4.8 (matroid from rank).

- Given a matroid $M = (E, \mathcal{I})$, we see its rank function as defined in Eq. 14 satisfies (R1), (R2), and, as we saw in Lemma 4.7, (R3) too.
- Next, assume we have (R1), (R2), and (R3). Define $\mathcal{I} = \{X \subseteq E : r(X) = |X|\}$. We will show that $(E, \mathcal{I})$ is a matroid.
- First, $\emptyset \in \mathcal{I}$.

. . .

Matroids from rank

Proof of Theorem 4.8 (matroid from rank).

- Given a matroid $M = (E, \mathcal{I})$, we see its rank function as defined in Eq. 14 satisfies (R1), (R2), and, as we saw in Lemma 4.7, (R3) too.
- Next, assume we have (R1), (R2), and (R3). Define $\mathcal{I} = \{X \subseteq E : r(X) = |X|\}$. We will show that $(E, \mathcal{I})$ is a matroid.
- First, $\emptyset \in \mathcal{I}$.
- Also, if $Y \in \mathcal{I}$ and $X \subseteq Y$ then by submodularity,

...

Matroids from rank

Proof of Theorem 4.8 (matroid from rank).

- Given a matroid $M = (E, \mathcal{I})$, we see its rank function as defined in Eq. 14 satisfies (R1), (R2), and, as we saw in Lemma 4.7, (R3) too.
- Next, assume we have (R1), (R2), and (R3). Define $\mathcal{I} = \{X \subseteq E : r(X) = |X|\}$. We will show that $(E, \mathcal{I})$ is a matroid.
- First, $\emptyset \in \mathcal{I}$.
- Also, if $Y \in \mathcal{I}$ and $X \subseteq Y$ then by submodularity, $X \cap (Y \setminus X) = \emptyset$
$$r(X) \geq r(Y) - r(Y \setminus X) \tag{18}$$

. . .

Matroids from rank

Proof of Theorem 4.8 (matroid from rank).

- Given a matroid $M = (E, \mathcal{I})$, we see its rank function as defined in Eq. 14 satisfies (R1), (R2), and, as we saw in Lemma 4.7, (R3) too.
- Next, assume we have (R1), (R2), and (R3). Define $\mathcal{I} = \{X \subseteq E : r(X) = |X|\}$. We will show that $(E, \mathcal{I})$ is a matroid.
- First, $\emptyset \in \mathcal{I}$.
- Also, if $Y \in \mathcal{I}$ and $X \subseteq Y$ then by submodularity,

$$r(X) \geq r(Y) - r(Y \setminus X) \tag{18}$$

$$\geq |Y| - |Y \setminus X| \tag{19}$$

. . .

Matroids from rank

Proof of Theorem 4.8 (matroid from rank).

- Given a matroid $M = (E, \mathcal{I})$, we see its rank function as defined in Eq. 14 satisfies (R1), (R2), and, as we saw in Lemma 4.7, (R3) too.
- Next, assume we have (R1), (R2), and (R3). Define $\mathcal{I} = \{X \subseteq E : r(X) = |X|\}$. We will show that $(E, \mathcal{I})$ is a matroid.
- First, $\emptyset \in \mathcal{I}$.
- Also, if $Y \in \mathcal{I}$ and $X \subseteq Y$ then by submodularity,

$$r(X) \geq r(Y) - r(Y \setminus X) \tag{18}$$

$$\geq |Y| - |Y \setminus X| \tag{19}$$

$$= |X| \tag{20}$$

$\Rightarrow r(X) = |X|$

and thus $X \in \mathcal{I}$.

...

Matroids from rank

Proof of Theorem 4.8 (matroid from rank) cont.

- Let $A, B \in \mathcal{I}$, with $|A| < |B|$, so $r(A) = |A| < r(B) = |B|$. Let $B \setminus A = \{b_1, b_2, \dots, b_k\}$.

☐

Matroids from rank

Proof of Theorem 4.8 (matroid from rank) cont.

- Let $A, B \in \mathcal{I}$, with $|A| < |B|$, so $r(A) = |A| < r(B) = |B|$. Let $B \setminus A = \{b_1, b_2, \dots, b_k\}$.
- Suppose, to the contrary, that $\forall b \in B \setminus A$, $r(\overset{\frown}{A} + \overset{\frown}{b}) \notin \mathcal{I}$, which means for all such b, $r(A + b) = r(A) = |A|$. Then

□

Matroids from rank

Proof of Theorem 4.8 (matroid from rank) cont.

- Let $A, B \in \mathcal{I}$, with $|A| < |B|$, so $r(A) = |A| < r(B) = |B|$. Let $B \setminus A = \{b_1, b_2, \dots, b_k\}$.
- Suppose, to the contrary, that $\forall b \in B \setminus A$, $r(A + b) \notin \mathcal{I}$, which means for all such b, $r(A + b) = r(A) = |A|$. Then

$$r(B) \leq r(A \cup B) \tag{21}$$

☐

Matroids from rank

Proof of Theorem 4.8 (matroid from rank) cont.

- Let $A, B \in \mathcal{I}$, with $|A| < |B|$, so $r(A) = |A| < r(B) = |B|$. Let $B \setminus A = \{b_1, b_2, \dots, b_k\}$.
- Suppose, to the contrary, that $\forall b \in B \setminus A$, $r(A + b) \notin \mathcal{I}$, which means for all such b, $r(A + b) = r(A) = |A|$. Then

$$r(B) \leq r(A \cup B) \tag{21}$$

$$\leq r(\underbrace{A \cup (B \setminus \{b_1\})}_{X}) + r(\underbrace{A \cup \{b_1\}}_{Y}) - r(A) \tag{22}$$

$r(X\cup Y) + r(X \cap Y) \leq r(X) + r(Y)$

$X \cap Y = A$

$X \cup Y = A \cup B$ -

$\square$

Matroids from rank

Proof of Theorem 4.8 (matroid from rank) cont.

- Let $A, B \in \mathcal{I}$, with $|A| < |B|$, so $r(A) = |A| < r(B) = |B|$. Let $B \setminus A = \{b_1, b_2, \dots, b_k\}$.
- Suppose, to the contrary, that $\forall b \in B \setminus A$, $r(A + b) \notin \mathcal{I}$, which means for all such b, $r(A + b) = r(A) = |A|$. Then

$$\begin{aligned} r(B) &\leq r(A \cup B) && (21) \\ &\leq r(A \cup (B \setminus \{b_1\})) + r(A \cup \{b_1\}) - r(A) && (22) \\ &= r(A \cup (B \setminus \{b_1\}) && (23) \end{aligned}$$

☐

Matroids from rank

Proof of Theorem 4.8 (matroid from rank) cont.

- Let $A, B \in \mathcal{I}$, with $|A| < |B|$, so $r(A) = |A| < r(B) = |B|$. Let $B \setminus A = \{b_1, b_2, \dots, b_k\}$.
- Suppose, to the contrary, that $\forall b \in B \setminus A$, $r(A + b) \notin \mathcal{I}$, which means for all such b, $r(A + b) = r(A) = |A|$. Then

$$\begin{aligned} r(B) &\leq r(A \cup B) && (21) \\ &\leq r(A \cup (B \setminus \{b_1\})) + r(A \cup \{b_1\}) - r(A) && (22) \\ &= r(A \cup (B \setminus \{b_1\}) && (23) \\ &\leq r(A \cup (B \setminus \{b_1, b_2\})) + r(A \cup \{b_2\}) - r(A) && (24) \end{aligned}$$

☐

Matroids from rank

Proof of Theorem 4.8 (matroid from rank) cont.

- Let $A, B \in \mathcal{I}$, with $|A| < |B|$, so $r(A) = |A| < r(B) = |B|$. Let $B \setminus A = \{b_1, b_2, \dots, b_k\}$.
- Suppose, to the contrary, that $\forall b \in B \setminus A$, $r(A + b) \notin \mathcal{I}$, which means for all such b, $r(A + b) = r(A) = |A|$. Then

$$\begin{aligned} r(B) &\leq r(A \cup B) && (21) \\ &\leq r(A \cup (B \setminus \{b_1\})) + r(A \cup \{b_1\}) - r(A) && (22) \\ &= r(A \cup (B \setminus \{b_1\}) && (23) \\ &\leq r(A \cup (B \setminus \{b_1, b_2\})) + r(A \cup \{b_2\}) - r(A) && (24) \\ &= r(A \cup (B \setminus \{b_1, b_2\})) && (25) \end{aligned}$$

☐

Matroids from rank

Proof of Theorem 4.8 (matroid from rank) cont.

- Let $A, B \in \mathcal{I}$, with $|A| < |B|$, so $r(A) = |A| < r(B) = |B|$. Let $B \setminus A = \{b_1, b_2, \dots, b_k\}$.
- Suppose, to the contrary, that $\overline{(}\forall b \in B \setminus A$, $r(A + b) \notin \mathcal{I}$, which means for all such b, $r(A + b) = r(A) = |A|$. Then

$$\begin{aligned}
r(B) &\leq r(A \cup B) && (21)\\
&\leq r(A \cup (B \setminus \{b_1\})) + r(A \cup \{b_1\}) - r(A) && (22)\\
&= r(A \cup (B \setminus \{b_1\}) && (23)\\
&\leq r(A \cup (B \setminus \{b_1, b_2\})) + r(A \cup \{b_2\}) - r(A) && (24)\\
&= r(A \cup (B \setminus \{b_1, b_2\})) && (25)\\
&\leq \dots \leq r(A) = |A| < |B| \qquad \Rightarrow r(B) < |B| && (26)
\end{aligned}$$

□

Matroids from rank

Proof of Theorem 4.8 (matroid from rank) cont.

- Let $A, B \in \mathcal{I}$, with $|A| < |B|$, so $r(A) = |A| < r(B) = |B|$. Let $B \setminus A = \{\underline{b_1, b_2, \ldots, b_k}\}$.
- Suppose, to the contrary, that $\forall b \in B \setminus A$, $r(A + b) \notin \mathcal{I}$, which means for all such b, $r(A + b) = r(A) = |A|$. Then

$$
\begin{aligned}
r(B) &\leq r(A \cup B) \tag{21}\\
&\leq r(A \cup (B \setminus \{b_1\})) + r(A \cup \{b_1\}) - r(A) \tag{22}\\
&= r(A \cup (B \setminus \{b_1\}) \tag{23}\\
&\leq r(A \cup (B \setminus \{b_1, b_2\})) + r(A \cup \{b_2\}) - r(A) \tag{24}\\
&= r(A \cup (B \setminus \{b_1, b_2\})) \tag{25}\\
&\leq \ldots \leq r(A) = |A| < |B| \tag{26}
\end{aligned}
$$

giving a contradiction since $B \in \mathcal{I}$.

□

Matroids from rank II

Another way of using function r to define a matroid.

Theorem 4.9 (Matroid from rank II)

Let E be a finite set and let $r : 2^E \to \mathbb{Z}_+$ be a function. Then $r(\cdot)$ defines a matroid with r being its rank function if and only if for all $A \subseteq E$, and $x, y \in E$:

(R1') $r(\emptyset) = 0$;

(R2') $r(X) \leq r(X \cup \{y\}) \leq r(X) + 1$;

(R3') *If* $r(X \cup \{x\}) = r(X \cup \{y\}) = r(X)$, *then* $r(X \cup \{x, y\}) = r(X)$.

Matroids, more defs

Definition 4.10

closed/flat/subspace A subset $A \subseteq E$ is closed or a flat or a subspace of matroid M if for all $x \in \cancel{A \subseteq} E$, $r(A \cup \{x\}) = r(A) + 1$.

$E \setminus A$,

Definition 4.11

closure Given $A \subseteq E$, the closure of A, is defined by $\sigma(A) = \{b \in E : r(A \cup \{b\}) = r(A)\}$.

Therefore, a closed set A has $\sigma(A) = A$.

Definition 4.12

circuit A subset $A \subseteq E$ is circuit or a cycle if it is an inclusionwise minimally dependent set (i.e., if for any $a \in A$, $r(A \setminus \{a\}) = |A| - 1$).

$+ \; r(A) < |A|$

In general, besides independent sets and rank functions, there are other equivalent ways to characterize matroids.

Matroids by bases

Theorem 4.13

Matroid (by bases) Let E be a set and $\mathcal{B}$ be a nonempty collection of subsets of E. Then the following are equivalent.

1. *$\mathcal{B}$ is the collection of bases of a matroid;*
2. *if $B, B' \in \mathcal{B}$, and $x \in B' \setminus B$, then $B' - x + y \in \mathcal{B}$ for some $y \in B \setminus B'$.*
3. *If $B, B' \in \mathcal{B}$, and $x \in B' \setminus B$, then $B - y + x \in \mathcal{B}$ for some $y \in B \setminus B'$.*

Matroids by bases

Theorem 4.13

Matroid (by bases) Let E be a set and $\mathcal{B}$ be a nonempty collection of subsets of E. Then the following are equivalent.

1. *$\mathcal{B}$ is the collection of bases of a matroid;*
2. *if $B, B' \in \mathcal{B}$, and $x \in B' \setminus B$, then $B' - x + y \in \mathcal{B}$ for some $y \in B \setminus B'$.*
3. *If $B, B' \in \mathcal{B}$, and $x \in B' \setminus B$, then $B - y + x \in \mathcal{B}$ for some $y \in B \setminus B'$.*

exchange ~~[illegible]~~ properties.

Proof here is omitted but think about this for a moment in terms of linear spaces and matrices, and (alternatively) spanning trees.

Scratch Paper

Matroids by circuits

A set is independent if and only if it contains no circuit. Therefore, it is not surprising that circuits can also characterize a matroid.

Theorem 4.14

Matroid (by circuits) Let E be a set and $\mathcal{C}$ be a collection of nonempty subsets of E, such that no two sets in $\mathcal{C}$ are contained in each other. Then the following are equivalent.

1. *$\mathcal{C}$ is the collection of circuits of a matroid;*
2. *if $C, C' \in \mathcal{C}$, and $x \in C \cap C'$, then $(C \cup C') \setminus \{x\}$ contains a set in $\mathcal{C}$;*
3. *if $C, C' \in \mathcal{C}$, and $x \in C \cap C'$, and $y \in C \setminus C'$, then $(C \cup C') \setminus \{x\}$ contains a set in $\mathcal{C}$ containing y;*

Matroids by circuits

A set is independent if and only if it contains no circuit. Therefore, it is not surprising that circuits can also characterize a matroid.

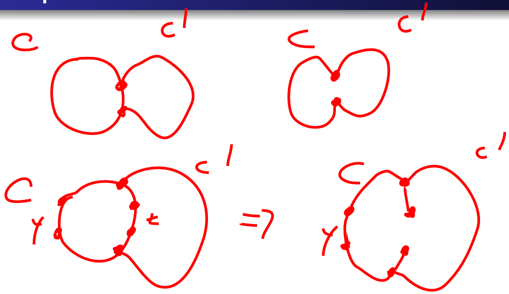
Theorem 4.14

Matroid (by circuits) Let E be a set and $\mathcal{C}$ be a collection of nonempty subsets of E, such that no two sets in $\mathcal{C}$ are contained in each other. Then the following are equivalent.

1. *$\mathcal{C}$ is the collection of circuits of a matroid;*
2. *if $C, C' \in \mathcal{C}$, and $x \in C \cap C'$, then $(C \cup C') \setminus \{x\}$ contains a set in $\mathcal{C}$;*
3. *if $C, C' \in \mathcal{C}$, and $x \in C \cap C'$, and $y \in C \setminus C'$, then $(C \cup C') \setminus \{x\}$ contains a set in $\mathcal{C}$ containing y;*

Again, think about this for a moment in terms of linear spaces and matrices, and spanning trees.

Scratch Paper



contain a cycle.

Maximization problems for matroids

- Given a matroid $M = (E, \mathcal{I})$ and a modular cost function $c : E \to \mathbb{R}$, the task is to find an $X \in \mathcal{I}$ such that $c(X) = \sum_{x \in X} c(x)$ is maximum.
- This seems remarkably similar to the max spanning tree problem.

Minimization problems for matroids

- Given a matroid $M = (E, \mathcal{I})$ and a modular cost function $c : E \to \mathbb{R}$, the task is to find a basis $B \in \mathcal{B}$ such that $c(B)$ is minimized.
- This sounds like a set cover problem (find the minimum cost set ~~of covering sets~~).

  covering

Uniform Matroid

- Given E, consider $\mathcal{I}$ to be all subsets of E that are at most size k. That is $\mathcal{I} = \{A \subseteq E : |A| \leq k\}$.
- Then $(E, \mathcal{I})$ is a matroid called a k-uniform matroid.

$I, J \in \mathcal{I}$

$|I| < |J| \leq k$

an $j \in J$ s.t. $j \notin I$

is such that

$|I + j| \leq k$

$\therefore \; I + j \in \mathcal{I}$

Uniform Matroid

- Given E, consider $\mathcal{I}$ to be all subsets of E that are at most size k. That is $\mathcal{I} = \{A \subseteq E : |A| \leq k\}$.
- Then $(E, \mathcal{I})$ is a matroid called a k-uniform matroid.
- Rank function

$$r(A) = \begin{cases} |A| & \text{if } |A| \leq k \\ k & \text{if } |A| > k \end{cases} \tag{27}$$

Uniform Matroid

- Given E, consider $\mathcal{I}$ to be all subsets of E that are at most size k. That is $\mathcal{I} = \{A \subseteq E : |A| \leq k\}$.
- Then $(E, \mathcal{I})$ is a matroid called a k-uniform matroid.
- Rank function

$$r(A) = \begin{cases} |A| & \text{if } |A| \leq k \\ k & \text{if } |A| > k \end{cases} \tag{27}$$

- Therefore, this function is submodular. Not surprising since $r(A) = \min(|A|, k)$ which is a non-decreasing concave function applied to a modular function.

Uniform Matroid

- Given E, consider $\mathcal{I}$ to be all subsets of E that are at most size k. That is $\mathcal{I} = \{A \subseteq E : |A| \leq k\}$.
- Then $(E, \mathcal{I})$ is a matroid called a k-uniform matroid.
- Rank function

$$r(A) = \begin{cases} |A| & \text{if } |A| \leq k \\ k & \text{if } |A| > k \end{cases} \tag{27}$$

- Therefore, this function is submodular. Not surprising since $r(A) = \min(|A|, k)$ which is a non-decreasing concave function applied to a modular function.
- Closure function

$$\sigma(A) = \begin{cases} A & \text{if } |A| < k, \\ \cancel{S}\ E & \text{if } |A| \geq k, \end{cases} \tag{28}$$

- A "free" matroid sets $k = |E|$, so everything is independent.

Linear Matroid

- Let $\mathbf{X}$ be an $n \times m$ matrix and $E = \{1, \dots, m\}$
- Let $\mathcal{I}$ consists of subsets of E such that if $A \in \mathcal{I}$, and $A = \{a_1, a_2, \dots, a_k\}$ then the vectors $x_{a_1}, x_{a_2}, \dots, x_{a_k}$ are linearly independent.
- the rank function is just the rank of the space spanned by the corresponding set of vectors.
- rank is submodular, it is intuitive that it satisfies the diminishing returns property (a given vector can only become linearly dependent in a greater context, thereby no longer contributing to rank).

Cycle Matroid of a graph

- Let $G = (V, E)$ be a graph. Consider $(E, \mathcal{I})$ where the edges of the graph E are the ground set and $A \in \mathcal{I}$ if the edge-induced graph $G(V, A)$ by A does not contain any cycle.
- Then $M = (E, \mathcal{I})$ is a matroid.
- $\mathcal{I}$ contains all forests.
- Bases are spanning forests (spanning trees if G is connected).

Cycle Matroid of a graph

- Let $G = (V, E)$ be a graph. Consider $(E, \mathcal{I})$ where the edges of the graph E are the ground set and $A \in \mathcal{I}$ if the edge-induced graph $G(V, A)$ by A does not contain any cycle.
- Then $M = (E, \mathcal{I})$ is a matroid.
- $\mathcal{I}$ contains all forests.
- Bases are spanning forests (spanning trees if G is connected).
- Rank function $r(A)$ is the size of the largest spanning forest contained in $G(V, A)$.
- Closure function adds all edges between the vertices adjacent to any edge in A. Closure of a spanning forest is G.

Partition Matroid

- Let $E = E_1 \cup E_2 \cup \dots \cup E_\ell$ be a partitioning of E into disjoint sets (disjoint union). Define a set of subsets of E as
$$\mathcal{I} = \{X \subseteq E : |X \cap E_i| \leq k_i \text{ for all } i = 1, \dots, \ell\}. \tag{29}$$
where $k_1, \dots, k_\ell$ are fixed parameters. Then $M = (E, \mathcal{I})$ is a matroid.
- Note that a uniform matroid is a trivial example of a partition matroid with $\ell = 1$.
- We'll show that property (I3') in Def 4.4 holds. If $X, Y \in \mathcal{I}$ with $|Y| > |X|$, then there must be at least one i with $|Y \cap E_i| > |X \cap E_i|$. Therefore, adding one element $e \in E_i \cap (Y \setminus X)$ to X won't break independence.

Matroid and the greedy algorithm

- Let $\mathcal{I}$ be a set of subsets of E that is down-closed. Consider a modular weight function $w : E \to \mathbb{R}$, and we want to find the $A \in \mathcal{I}$ that maximizes $w(A)$.
- Greedy algorithm: Set $A = \emptyset$, and repeatedly choose $y \in E \setminus A$ such that $A \cup \{y\} \in \mathcal{I}$ with $w(y)$ as large as possible, stopping when no such y exists.

Theorem 5.1

Let $\mathcal{I}$ be a non-empty collection of subsets of a set E, down-closed. Then the pair $(E, \mathcal{I})$ is a matroid if and only if for each weight function $w \in \mathcal{R}^E$, the greedy algorithm leads to a set $I \in \mathcal{I}$ of maximum weight $w(I)$.

Sources for Today's Lecture

Korte,Vygen-2005, Vondrak-2010, Schrivjer-2003, Oxley-1992.