Submodular Functions, Optimization, and Applications to Machine Learning

— Fall Quarter, Lecture 6 —

http://www.ee.washington.edu/people/faculty/bilmes/classes/ee563_spring_2018/

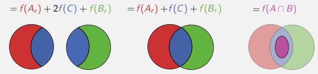
Prof. Jeff Bilmes

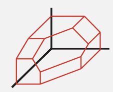
University of Washington, Seattle
Department of Electrical Engineering
http://melodi.ee.washington.edu/~bilmes

Oct 19th, 2020

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B)$$

$= f(A_r) + 2f(C) + f(B_r)$ $\quad = f(A_r) + f(C) + f(B_r)$ $\quad = f(A \cap B)$





Announcements, Assignments, and Reminders

- Homework 1 is out, due Monday 10/19/2020 at 11:59pm.
- Lecture 5 was posted to YouTube. See our announcements (https://canvas.uw.edu/courses/1397085/announcements) for the link.

Class Road Map - EE563

- L1(9/30): Motivation, Applications, Definitions, Properties
- L2(10/5): Sums concave(modular), uses (diversity/costs, feature selection), information theory
- L3(10/7): Monge, More Definitions, Graph and Combinatorial Examples,
- L4(10/12): Graph & Combinatorial Examples, Matrix Rank, Properties, Other Defs, Independence
- L5(10/14): Properties, Defs of Submodularity, Independence
- L6(10/19): Matroids, Matroid Examples, Matroid Rank, More on Partition Matroid, Laminar Matroids, System of Distinct Reps
- L7(10/21):
- L8(10/26):
- L9(10/28):
- L10(11/2):
- L11(11/4):
- L12(11/9):
- L–(11/11): Veterans Day, Holiday
- L13(11/16):
- L14(11/18):
- L15(11/23):
- L16(11/25):
- L17(11/30):
- L18(12/2):
- L19(12/7):
- L20(12/9): maximization.

Last day of instruction, Fri. Dec 11th. Finals Week: Dec 12-18, 2020

Summary: Properties so far

- Cover functions $f(A) = w(\bigcup_{a \in A} U_a)$ are submodular.
- SCCM is submodular $f(A) = \sum_{u \in U} \phi_u(m_u(A))$ where m_u is a non-negative modular and ϕ_u is concave.
- max is submodular $f(A) = \max_{j \in A} c_j$, as is facility location $f(A) = \sum_{u \in U} \max_{a \in A} s_{a,u}$.
- Matrix rank function is submodular.
- Log determinant $f(A) = \log\det(\mathbf{\Sigma}_A)$ submodular for p.d. Σ.
- Matrix rank $r(A)$, dim. of space spanned by the vector set $\{x_a\}_{a \in A}$.
- Graph cut, set cover, and incidence functions,
- quadratics with non-positive off-diagonals $f(X) = m^{\intercal}\mathbf{1}_X + \frac{1}{2}\mathbf{1}_X^{\intercal}\mathbf{M}\mathbf{1}_X$.
- Number connected components in induced graph $c(A)$, and interior edge function $E(S)$, is supermodular.
- Submodular plus modular is submodular, $f(A) = f'(A) + m(A)$.
- Complementation: $f'(A) = f(V \setminus A)$ is submodular if f is submodular (same for supermodular, modular).
- Conix mixture: $\alpha_i \geq 0$, $f_i : 2^V \to \mathbb{R}$ submodular, then so is $\sum_i \alpha_i f_i$.
- Restrictions preserve submodularity: $f'(A) = f(A \cap S)$

Summary: Other properties from last lecture

- Given non-decreasing submodular f and non-decreasing concave ϕ then $h(A) = \phi(f(A))$ is submodular.
- $h(A) = \min(f(A), g(A))$ is submodular if both f and g are, and if $f - g$ is monotone (increasing or decreasing).
- Any set function h can be represented as $h(A) = c + f(A) - g(A)$ where c is a constant, and f, g are polymatroidal.
- Gain $f(j|A)$ is like a discrete gradient $\nabla_j f(A)$.
- Any submodular g function can be represented by a sum of a totally normalized polymatroidal function $\bar{g}$ and a modular function m_g.

Many (Equivalent) Definitions of Submodularity

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B), \quad \forall A, B \subseteq V \tag{6.16}$$

$$f(j|S) \geq f(j|T), \ \forall S \subseteq T \subseteq V, \text{ with } j \in V \setminus T \tag{6.17}$$

$$f(C|S) \geq f(C|T), \forall S \subseteq T \subseteq V, \text{ with } C \subseteq V \setminus T \tag{6.18}$$

$$f(j|S) \geq f(j|S \cup \{k\}), \ \forall S \subseteq V \text{ with } j \in V \setminus (S \cup \{k\}) \tag{6.19}$$

$$f(A \cup B|A \cap B) \leq f(A|A \cap B) + f(B|A \cap B), \quad \forall A, B \subseteq V \tag{6.20}$$

$$f(T) \leq f(S) + \sum_{j \in T \setminus S} f(j|S) - \sum_{j \in S \setminus T} f(j|S \cup T - \{j\}), \ \forall S, T \subseteq V \tag{6.21}$$

$$f(T) \leq f(S) + \sum_{j \in T \setminus S} f(j|S), \ \forall S \subseteq T \subseteq V \tag{6.22}$$

$$f(T) \leq f(S) - \sum_{j \in S \setminus T} f(j|S \setminus \{j\}) + \sum_{j \in T \setminus S} f(j|S \cap T) \ \forall S, T \subseteq V \tag{6.23}$$

$$f(T) \leq f(S) - \sum_{j \in S \setminus T} f(j|S \setminus \{j\}), \ \forall T \subseteq S \subseteq V \tag{6.24}$$

From Matrix Rank → Matroid

- So V is set of column vector indices of a matrix.
- Let $\mathcal{I} = \{I_1, I_2, \dots\}$ be a set of all subsets of V such that for any $I \in \mathcal{I}$, the vectors indexed by I are linearly independent.
- Given a set $B \in \mathcal{I}$ of linearly independent vectors, then any subset $A \subseteq B$ is also linearly independent. Hence, $\mathcal{I}$ is down-closed or "subclusive", under subsets. In other words,

$$A \subseteq B \text{ and } B \in \mathcal{I} \Rightarrow A \in \mathcal{I} \tag{6.1}$$

- maxInd: Inclusionwise maximal independent subsets of (i.e., the set of bases of) any set $B \subseteq V$ defined as:

$$\text{maxInd}(B) \triangleq \{A \subseteq B : A \in \mathcal{I} \text{ and } \forall v \in B \setminus A, A \cup \{v\} \notin \mathcal{I}\} \tag{6.2}$$

- Given any set $B \subset V$ of vectors, all maximal (by set inclusion) subsets of linearly independent vectors are the same size. That is, for all $B \subseteq V$,

$$\forall A_1, A_2 \in \text{maxInd}(B), \quad |A_1| = |A_2| = \text{rank}(B) \tag{6.3}$$

From Matrix Rank → Matroid

- Let $\mathcal{I} = \{I_1, I_2, \dots\}$ be the set of sets as described above.
- Thus, for all $I \in \mathcal{I}$, the matrix rank function has the property

$$r(I) = \max \{|A| : A \subseteq I \text{ and } A \in \mathcal{I}\} = |I| \tag{6.4}$$

and for any $B \notin \mathcal{I}$,

$$r(B) = \max \{|A| : A \subseteq B \text{ and } A \in \mathcal{I}\} < |B| \tag{6.5}$$

Since all maximally independent subsets of a set are the same size, the rank function is well defined.

Matroids

- 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 $\mathcal{I} = \{I_1, I_2, \ldots\}$ of E that correspond to independent elements.
- There are many definitions of matroids that are mathematically equivalent, we'll see some of them here.
- "If a theorem about graphs can be expressed in terms of edges and circuits only, it probably exemplifies a more general theorem about matroids." – Tutte

Independence System

Definition 6.3.1 (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 $(E, \mathcal{I})$.

- Set systems can be arbitrarily complex since, as stated, there is no systematic method (besides exponential-cost exhaustive search) to determine if a given set $S \subseteq E$ has $S \in \mathcal{I}$. No inherent structure.
- One useful structural property is "heredity." Namely, a set system is said to be a hereditary set system if for any $A \subset B \in \mathcal{I}$, we have that $A \in \mathcal{I}$.

Independence System

Definition 6.3.2 (independence (or hereditary) system)

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

$$\emptyset \in \mathcal{I} \quad \text{(emptyset containing)} \tag{I1}$$

and

$$\forall I \in \mathcal{I}, J \subset I \Rightarrow J \in \mathcal{I} \quad \text{(subclusive)} \tag{I2}$$

- Property (I2) called "down monotone," "down closed," or "subclusive"
- Example: $E = \{1, 2, 3, 4\}$. With $\mathcal{I} = \{\emptyset, \{1\}, \{1, 2\}, \{1, 2, 4\}\}$.
- Then $(E, \mathcal{I})$ is a set system, but not an independence system since it is not down closed (e.g., we have $\{1, 2\} \in \mathcal{I}$ but not $\{2\} \in \mathcal{I}$).
- With $\mathcal{I} = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}$, then $(E, \mathcal{I})$ is now an independence (hereditary) system.

Independence System

$$\begin{array}{c c} & \begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{pmatrix} 0 & 0 & 1 & 1 & 2 & 1 & 3 & 1 \\ 0 & 1 & 1 & 0 & 2 & 0 & 2 & 4 \\ 1 & 1 & 1 & 0 & 0 & 3 & 1 & 5 \end{pmatrix} \end{array} = \begin{array}{c} \begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array} \\ \begin{pmatrix} | & | & | & | & | & | & | & | \\ x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 \\ | & | & | & | & | & | & | & | \end{pmatrix} \end{array} \tag{6.6}$$

- 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

Independent set definition of a matroid is perhaps most natural. Note, if $J \in \mathcal{I}$, then J is said to be an independent set.

Definition 6.3.3 (Matroid)

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

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

(I2) $\forall I \in \mathcal{I}, J \subset I \Rightarrow J \in \mathcal{I}$ (down-closed or subclusive)

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

Why is (I1) is not redundant given (I2)? Because without (I1) could have a non-matroid where $\mathcal{I} = \{\}$.

On Matroid History - a brief minor digression

- Abstract properties of linear dependence (Hassler Whitney, 1935), but already then found instances of objects with those properties not based on a matrix.
- Takeo Nakasawa, 1935, also early work.
- Forgotten for 20 years until mid 1950s.
- Matroids are powerful and flexible combinatorial objects.
- The rank function of a matroid is already a very powerful submodular function (perhaps all we need for many problems).
- Understanding matroids crucial for understanding submodularity.
- Matroid independent sets (i.e., A s.t. $r(A) = |A|$) are useful constraint set, and fast algorithms for submodular optimization subject to one (or more) matroid independence constraints exist.
- Crapo & Rota preferred the term "combinatorial geometry", or more specifically a "pregeometry" and said that pregeometries are "often described by the ineffably cacaphonic term 'matroid', which we prefer to avoid in favor of the term 'pregeometry'."

Matroid

Slight modification (non unit increment) that is equivalent.

Definition 6.3.4 (Matroid-II)

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

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

(I2') $\forall I \in \mathcal{I}, J \subset I \Rightarrow J \in \mathcal{I}$ (down-closed or subclusive)

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

Note (I1)=(I1'), (I2)=(I2'), and we get (I3)$\equiv$(I3') using induction.

Matroids, independent sets, and bases

- Independent sets: 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.
- A base of $U \subseteq E$: 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 $B \subset Z \subseteq U$.
- A base of a matroid: 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, or a spanning forest in a graph, or a spanning tree in a connected graph).

Matroids - important property

Proposition 6.3.5

In a matroid $M = (E, \mathcal{I})$, for any $U \subseteq E(M)$, any two bases of U have the same size.

- 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 the following is equivalent to the above.

Definition 6.3.6 (Matroid)

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

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

(I2') $\forall I \in \mathcal{I}, J \subset I \Rightarrow J \in \mathcal{I}$ (down-closed or subclusive)

(I3') $\forall X \subseteq V$, and $I_1, I_2 \in \text{maxInd}(X)$, we have $|I_1| = |I_2|$ (all maximally independent subsets of X have the same size).

Matroids - rank

- Thus, in any matroid $M = (E, \mathcal{I})$, $\forall U \subseteq E(M)$, any two bases of U have the same size.
- 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 matroid, and is the common size of all the bases of the matroid.
- We can a bit more formally define the rank function this way.

Definition 6.3.7 (matroid rank function)

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

$$r(A) = \max \{|X| : X \subseteq A, X \in \mathcal{I}\} = \max_{X \in \mathcal{I}} |A \cap X| \tag{6.7}$$

- 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 (in this case, A is a self base).

Matroids, other definitions using matroid rank $r : 2^V \to \mathbb{Z}_+$

Definition 6.3.8 (closed/flat/subspace)

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

Definition: A hyperplane is a flat of rank $r(M) - 1$.

Definition 6.3.9 (closure)

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

Therefore, a closed set A has $\text{span}(A) = A$.

Definition 6.3.10 (circuit)

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

Matroids by bases

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

Theorem 6.3.11 (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 $y \in B \setminus B'$.*

Properties 2 and 3 are called "exchange properties."
Proof here is omitted but think about this for a moment in terms of linear spaces and matrices, and (alternatively) spanning trees.

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 6.3.12 (Matroid by circuits)

Let E be a set and $\mathcal{C}$ be a collection of subsets of E that satisfy the following three properties:

1. *(C1): $\emptyset \notin \mathcal{C}$*
2. *(C2): if $C_1, C_2 \in \mathcal{C}$ and $C_1 \subseteq C_2$, then $C_1 = C_2$.*
3. *(C3): if $C_1, C_2 \in \mathcal{C}$ with $C_1 \neq C_2$, and $e \in C_1 \cap C_2$, then there exists a $C_3 \in \mathcal{C}$ such that $C_3 \subseteq (C_1 \cup C_2) \setminus \{e\}$.*

Matroids | Matroid Examples | Matroid Rank | More on Partition Matroid | Laminar Matroids | System of Distinct Reps

Matroids by circuits

Several circuit definitions for matroids.

Theorem 6.3.13 (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.

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.
- Note, if $I, J \in \mathcal{I}$, and $|I| < |J| \leq k$, and $j \in J$ such that $j \notin I$, then j is such that $|I + j| \leq k$ and so $I + j \in \mathcal{I}$.
- Rank function

$$r(A) = \begin{cases} |A| & \text{if } |A| \leq k \\ k & \text{if } |A| > k \end{cases} \tag{6.8}$$

- Note, 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

$$\text{span}(A) = \begin{cases} A & \text{if } |A| < k, \\ E & \text{if } |A| \geq k, \end{cases} \tag{6.9}$$

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

Linear (or Matric) Matroid

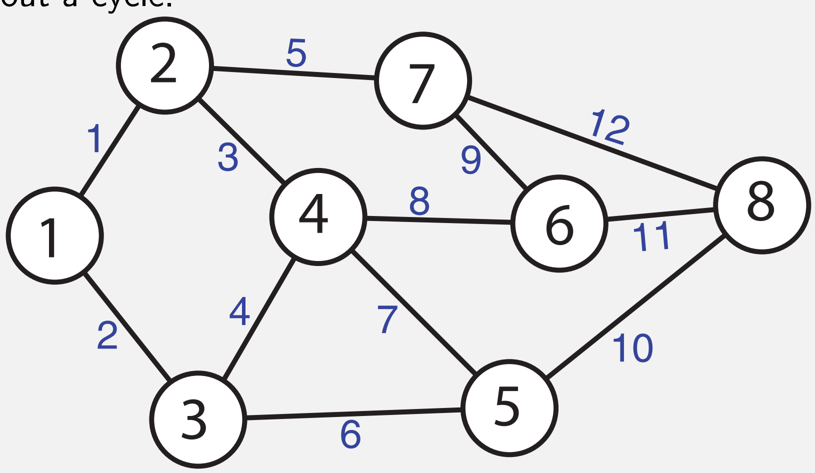
- Let $\mathbf{X}$ be an $n \times m$ matrix and $E = \{1, \ldots, m\}$
- Let $\mathcal{I}$ consists of subsets of E such that if $A \in \mathcal{I}$, and $A = \{a_1, a_2, \ldots, a_k\}$ then the vectors $x_{a_1}, x_{a_2}, \ldots, x_{a_k}$ are linearly independent.
- the rank function is just the rank 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).
- Called both linear matroids and matric matroids.

Cycle Matroid of a graph: Graphic Matroids

- 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 $G(V, A)$.
- Recall from earlier, $r(A) = |V(G)| - k_G(A)$, where for $A \subseteq E(G)$, we define $k_G(A)$ as the number of connected components of the edge-induced spanning subgraph $(V(G), A)$, and that $k_G(A)$ is supermodular, so $|V(G)| - k_G(A)$ is submodular.
- Closure function adds all edges between the vertices adjacent to any edge in A. Closure of a spanning forest is G.

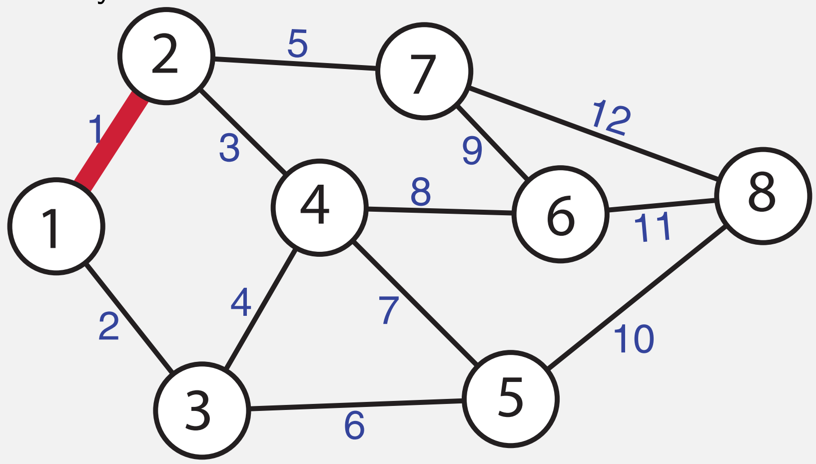
Example: graphic matroid

- A graph defines a matroid on edge sets, independent sets are those without a cycle.



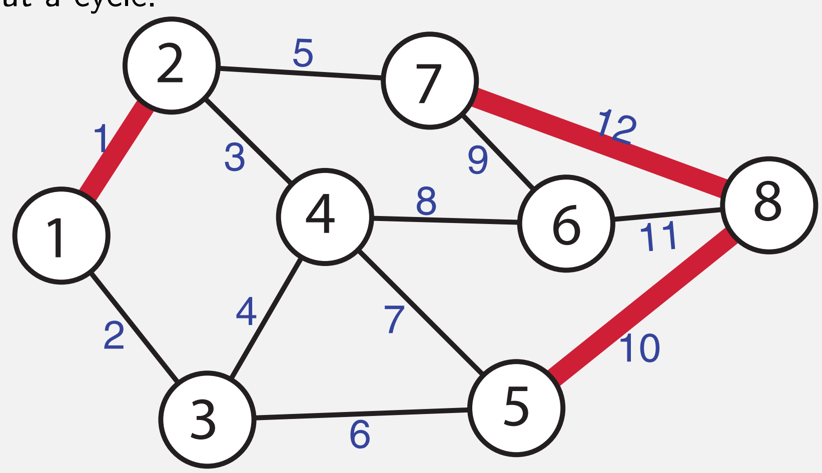
Example: graphic matroid

- A graph defines a matroid on edge sets, independent sets are those without a cycle.



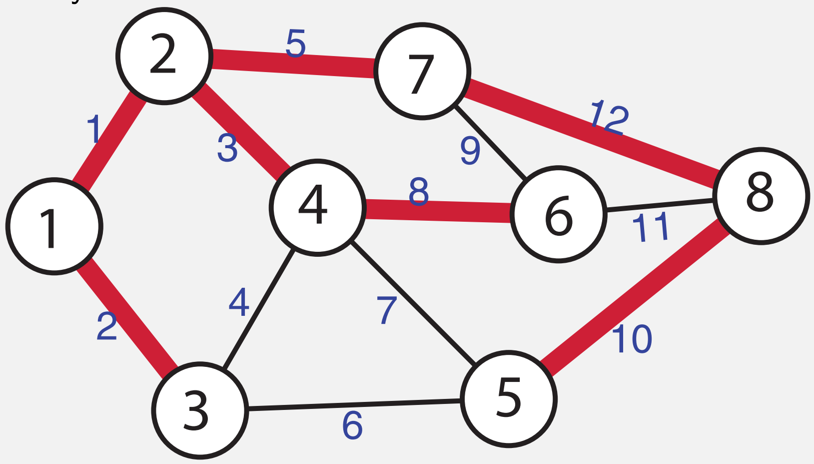
Example: graphic matroid

- A graph defines a matroid on edge sets, independent sets are those without a cycle.



Example: graphic matroid

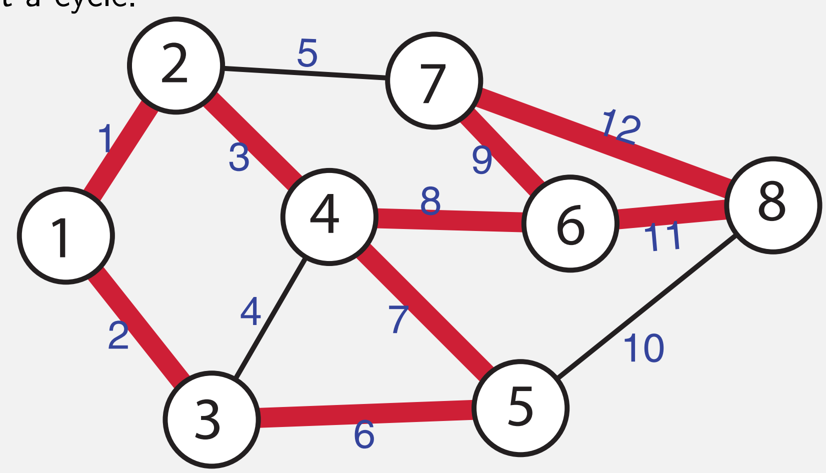
- A graph defines a matroid on edge sets, independent sets are those without a cycle.



Matroids Matroid Examples Matroid Rank More on Partition Matroid Laminar Matroids System of Distinct Reps

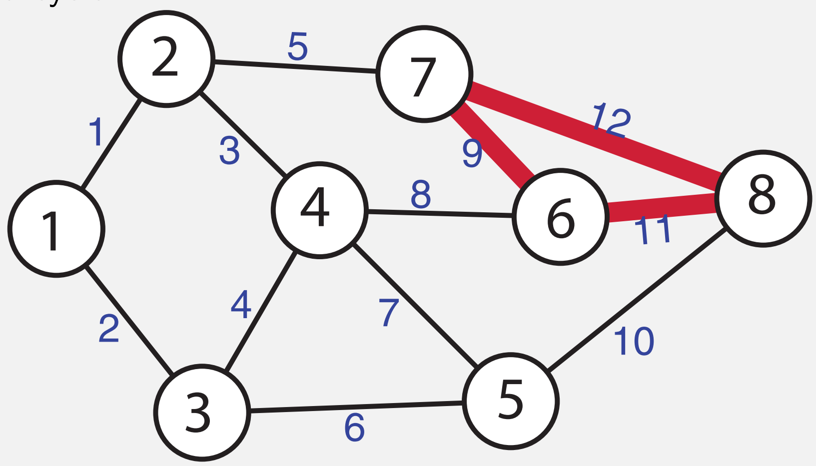
Example: graphic matroid

- A graph defines a matroid on edge sets, independent sets are those without a cycle.



Example: graphic matroid

- A graph defines a matroid on edge sets, independent sets are those without a cycle.



Partition Matroid

- Let V be our ground set.
- Let $V = V_1 \cup V_2 \cup \cdots \cup V_\ell$ be a partition of V into ℓ blocks (i.e., disjoint sets). Define a set of subsets of V as

$$\mathcal{I} = \{X \subseteq V : |X \cap V_i| \leq k_i \text{ for all } i = 1, \ldots, \ell\}. \tag{6.10}$$

  where $k_1, \ldots, k_\ell$ are fixed "limit" parameters, $k_i \geq 0$. Then $M = (V, \mathcal{I})$ is a matroid.
- Note that a k-uniform matroid is a trivial example of a partition matroid with $\ell = 1$, $V_1 = V$, and $k_1 = k$.
- Parameters associated with a partition matroid: ℓ and $k_1, k_2, \ldots, k_\ell$ although often the k_i's are all the same.
- We'll show that property (I3') in Def 6.3.4 holds. First note, for any $X \subseteq V$, $|X| = \sum_{i=1}^{\ell} |X \cap V_i|$ since $\{V_1, V_2, \ldots, V_\ell\}$ is a partition.
- If $X, Y \in \mathcal{I}$ with $|Y| > |X|$, then there must be at least one i with $|Y \cap V_i| > |X \cap V_i|$. Therefore, adding one element $e \in V_i \cap (Y \setminus X)$ to X won't break independence.

Partition Matroid

Ground set of objects, $V = \Bigg\{$

$\Bigg\}$

Partition Matroid

Partition of V into six blocks, $V_1, V_2, \ldots, V_6$

$\Bigg\}$

Partition Matroid

Limit associated with each block, $\{k_1, k_2, \dots, k_6\}$

Matroids | Matroid Examples | Matroid Rank | More on Partition Matroid | Laminar Matroids | System of Distinct Reps

Partition Matroid

Independent subset but not maximally independent.

Partition Matroid

Maximally independent subset, what is called a base.

Partition Matroid

Not independent since over limit in set six.

Partition Matroid

Is this a cycle/circuit?

Partition Matroid

No. Does it contain a cycle/circuit?

Partition Matroid

Yes.

$\}$

Matroids - rank function is submodular

Lemma 6.5.1

The rank function $r : 2^E \to \mathbb{Z}_+$ 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 $X \subseteq Y \subseteq A \cup B$.
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 (since $X \subseteq A \cap B$, $X \subseteq Y$, and $Y \subseteq A \cup B$),

$$r(A) + r(B) \geq |Y \cap A| + |Y \cap B| \tag{6.11}$$

$$= |Y \cap (A \cap B)| + |Y \cap (A \cup B)| \tag{6.12}$$

$$\geq |X| + |Y| = r(A \cap B) + r(A \cup B) \tag{6.13}$$

$\square$

A matroid is defined from its rank function

Theorem 6.5.2 (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) $r(A \cup B) + r(A \cap B) \leq r(A) + r(B)$ *for all* $A, B \subseteq E$ *(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 $r(A \cup \{v\})$, namely $r(A)$ or $r(A) + 1$.
- Hence, unit increment (if $r(A) = k$, then either $r(A \cup \{v\}) = k$ or $r(A \cup \{v\}) = k + 1$) holds.
- Thus, submodularity, normalized, monotone non-decreasing, & unit increment of rank is necessary & sufficient to define matroids.
- Can refer to matroid as (E, r), E is ground set, r is rank function.

Matroids from rank

Proof of Theorem 6.5.2 (matroid from rank).

- Given a matroid $M = (E, \mathcal{I})$, we see its rank function as defined in Eq. 6.7 satisfies (R1), (R2), and, as we saw in Lemma 6.5.1, (R3) too.
- Conversely, assume we have r satisfying (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,

$$\begin{aligned} r(X) &\geq r(Y) - r(Y \setminus X) + r(\emptyset) && (6.14) \\ &\geq |Y| - |Y \setminus X| && (6.15) \\ &= |X| && (6.16) \end{aligned}$$

implying $r(X) = |X|$, and thus $X \in \mathcal{I}$. ...

Matroids from rank

Proof of Theorem 6.5.2 (matroid from rank) cont.

- Let $A, B \in \mathcal{I}$, with $|A| < |B|$, so $r(A) = |A| < r(B) = |B|$. Let $B \setminus A = \{b_1, b_2, \dots, b_k\}$ (note $1 \le k \le |B|$).
- Suppose, to the contrary, that $\forall b \in B \setminus A$, $A + b \notin \mathcal{I}$, which means for all such b, $r(A+b) = r(A) = |A| < |A| + 1$. Then

$$
\begin{aligned}
r(B) &\le r(A \cup B) && (6.17)\\
&\le r(A \cup (B \setminus \{b_1\})) + r(A \cup \{b_1\}) - r(A) && (6.18)\\
&= r(A \cup (B \setminus \{b_1\}) && (6.19)\\
&\le r(A \cup (B \setminus \{b_1, b_2\})) + r(A \cup \{b_2\}) - r(A) && (6.20)\\
&= r(A \cup (B \setminus \{b_1, b_2\})) && (6.21)\\
&\le \dots \le r(A) = |A| < |B| && (6.22)
\end{aligned}
$$

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

Matroids from rank II

Another way of using function r to define a matroid.

Theorem 6.5.3 (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 $X \subseteq E$, and $x, y \in E$:

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

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

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

Matroids by submodular functions

Theorem 6.5.4 (Matroid by submodular functions)

Let $f : 2^E \to \mathbb{Z}$ be a integer valued monotone non-decreasing submodular function. Define a set of sets as follows:

$$\mathcal{C}(f) = \Big\{ C \subseteq E : C \text{ is non-empty, is inclusionwise-minimal, and has } f(C) < |C| \Big\} \tag{6.23}$$

Then $\mathcal{C}(f)$ is the collection of circuits of a matroid on E.

Inclusionwise-minimal in this case means that if $C \in \mathcal{C}(f)$, then there exists no $C' \subset C$ with $C' \in \mathcal{C}(f)$ (i.e., $C' \subset C$ would either be empty or have $f(C') \geq |C'|$). Also, recall inclusionwise-minimal in Definition 6.3.10, the definition of a circuit.

Summarizing: Many ways to define a Matroid

Summarizing what we've so far seen, we saw that it is possible to uniquely define a matroid based on any of:

- Independence (define the independent sets).
- Base axioms (exchangeability)
- Circuit axioms
- Closure axioms (we didn't see this, but it is possible)
- Rank axioms (normalized, monotone, cardinality bounded, non-negative integral, submodular)
- Matroids by integral submodular functions.

Maximization problems for matroids

- Given a matroid $M = (E, \mathcal{I})$ and a modular value 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 covering set of sets).

Partition Matroid

- What is the partition matroid's rank function?
- A partition matroids rank function:

$$r(A) = \sum_{i=1}^{\ell} \min(|A \cap V_i|, k_i) \tag{6.24}$$

which we also immediately see is submodular using properties we spoke about last week. That is:

1. $|A \cap V_i|$ is submodular (in fact modular) in A
2. $\min(\text{submodular}(A), k_i)$ is submodular in A since $|A \cap V_i|$ is monotone.
3. sums of submodular functions are submodular.

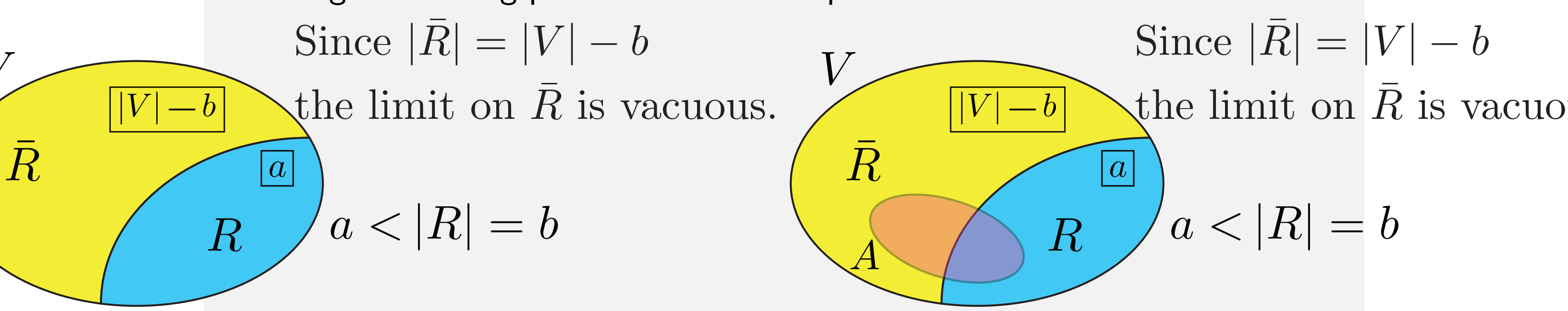
- $r(A)$ is also non-negative integral monotone non-decreasing, so it defines a matroid (the partition matroid).

From 2-partition matroid rank to truncated matroid rank

- Example: 2-partition matroid rank function: Given natural numbers $a, b \in \mathbb{Z}_+$ with $a < b$, and any set $R \subseteq V$ with $|R| = b$.
- Create two-block partition $V = (R, \bar{R})$, where $\bar{R} = V \setminus R$ so $|\bar{R}| = |V| - b$. Gives 2-partition matroid rank function as follows:

$$r(A) = \min(|A \cap R|, a) + \min(|A \cap \bar{R}|, |\bar{R}|) \tag{6.25}$$
$$= \min(|A \cap R|, a) + |A \cap \bar{R}| \tag{6.26}$$
$$= \min(|A \cap \bar{R}| + |A \cap R|, |A \cap \bar{R}| + a) \tag{6.27}$$
$$= \min(|A|, |A \cap \bar{R}| + a) \tag{6.28}$$

- Figure showing partition blocks and partition matroid limits.



Truncated Matroid Rank Function

- Define truncated matroid rank function. Start with 2-partition matroid rank $r(A) = \min(|A \cap R|, a) + \min(|A \cap \bar{R}|, |\bar{R}|)$, $a < b$. Define:

$$f_R(A) = \min\left\{ \boxed{r(A)}, b \right\} \tag{6.29}$$

$$= \min\left\{ \boxed{\min(|A|, |A \cap \bar{R}| + a)}, b \right\} \tag{6.30}$$

$$= \min\left\{ |A|, a + |A \cap \bar{R}|, b \right\} \tag{6.31}$$

- Defines a matroid $M = (V, f_R) = (V, \mathcal{I})$ (Goemans et. al.) with

$$\mathcal{I} = \{I \subseteq V : |I| \leq b \text{ and } |I \cap R| \leq a\}, \tag{6.32}$$

Useful for showing hardness of constrained submodular minimization. Consider sets $B \subseteq V$ with $|B| = b$. Recall R fixed, and $|R| = b$.

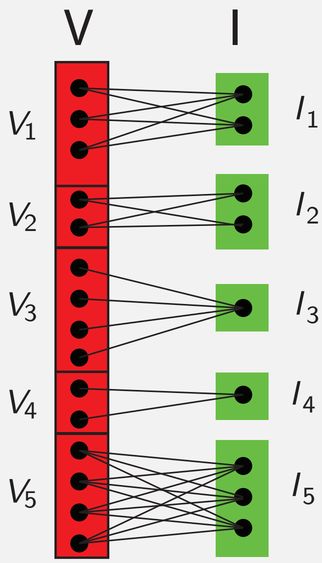
- For R, we have $f_R(R) = \min(b, a, b) = a < b$.
- For any B with $|B \cap R| \leq a$, $f_R(B) = b$.
- For any B with $|B \cap R| = \ell$, with $a \leq \ell \leq b$, $f_R(B) = a + b - \ell$.
- R, the set with minimum valuation amongst size-b sets, is hidden within an exponentially larger set of size-b sets with larger valuation.

Partition Matroid, rank as matching

- A partition matroid can be viewed using a bipartite graph.
- Letting V denote the ground set, and $V_1, V_2, \ldots$ the partition, the bipartite graph is $G = (V, I, E)$ where V is the ground set, I is a set of "indices", and E is the set of edges.
- $I = (I_1, I_2, \ldots, I_\ell)$ is a set of $k = \sum_{i=1}^{\ell} k_i$ nodes, grouped into ℓ clusters, where there are k_i nodes in the i^{th} group I_i, and $|I_i| = k_i$.
- $(v, i) \in E(G)$ iff $v \in V_j$ and $i \in I_j$.

Partition Matroid, rank as matching

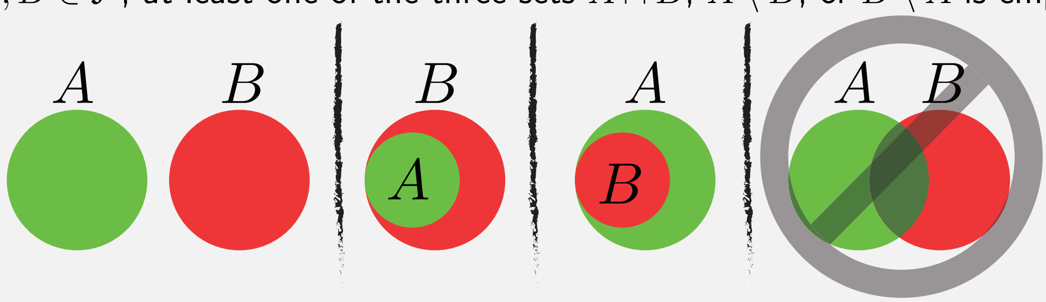
- Example where $\ell = 5$, $(k_1, k_2, k_3, k_4, k_5) = (2, 2, 1, 1, 3)$.



- Recall, $\Gamma : 2^V \to \mathbb{R}$ as the neighbor function in a bipartite graph, the neighbors of X is defined as $\Gamma(X) = \{v \in V(G) \setminus X : E(X, \{v\}) \neq \emptyset\}$, and recall that $|\Gamma(X)|$ is submodular.
- Here, for $X \subseteq V$, we have $\Gamma(X) = \{i \in I : (v, i) \in E(G) \text{ and } v \in X\}$.
- For such a constructed bipartite graph, the rank function of a partition matroid is $r(X) = \sum_{i=1}^{\ell} \min(|X \cap V_i|, k_i) =$ the maximum matching involving X.

Laminar Family and Laminar Matroid

- We can define a matroid with structures richer than just partitions.
- A set system $(V, \mathcal{F})$ is called a laminar family if for any two sets $A, B \in \mathcal{F}$, at least one of the three sets $A \cap B$, $A \setminus B$, or $B \setminus A$ is empty.



- Family is laminar $\exists$ no two properly intersecting members: $\forall A, B \in \mathcal{F}$, either A, B disjoint ($A \cap B = \emptyset$) or comparable ($A \subseteq B$ or $B \subseteq A$).
- Suppose we have a laminar family $\mathcal{F}$ of subsets of V and an integer k_A for every set $A \in \mathcal{F}$. Then $(V, \mathcal{I})$ defines a matroid where

$$\mathcal{I} = \{I \subseteq E : |I \cap A| \leq k_A \text{ for all } A \in \mathcal{F}\} \tag{6.33}$$

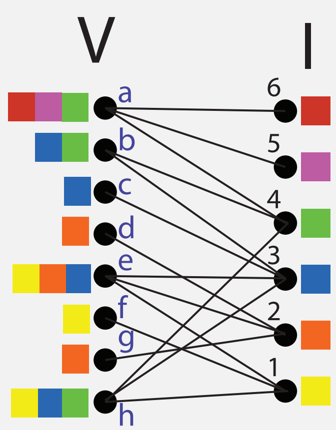
- Exercise: what is the rank function here?

System of Representatives

- Let $(V, \mathcal{V})$ be a set system (i.e., $\mathcal{V} = (V_i : i \in I)$ where $\emptyset \subset V_i \subseteq V$ for all i), and I is an index set. Hence, $|I| = |\mathcal{V}|$.
- Here, the sets $V_i \in \mathcal{V}$ are like "groups" and any $v \in V$ with $v \in V_i$ is a member of group i. Groups need not be disjoint (e.g., interest groups of individuals).
- A family $(v_i : i \in I)$ with $v_i \in V$ is said to be a system of representatives of $\mathcal{V}$ if $\exists$ a bijection $\pi : I \to I$ such that $v_i \in V_{\pi(i)}$.
- v_i is the representative of set (or group) $V_{\pi(i)}$, meaning the i^{th} representative is meant to represent set $V_{\pi(i)}$.
- Example: Consider the house of representatives, $v_i =$ "Pramila Jayapal", while $i =$ "King County, WA-7".
- In a system of representatives, there is no requirement for the representatives to be distinct. I.e., we could have some $v_1 \in V_1 \cap V_2$, where v_1 represents both V_1 and V_2.
- We can view this as a bipartite graph.

System of Representatives

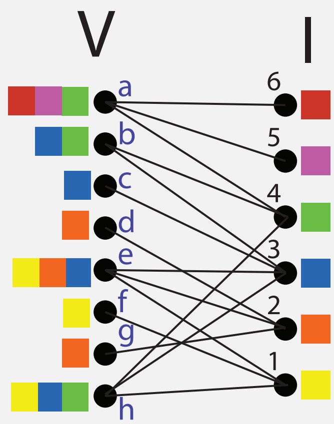
- We can view this as a bipartite graph. The groups of V are marked by color tags on the left, and also via right neighbors in the graph.
- Here, $\ell = 6$ groups, with $\mathcal{V} = (V_1, V_2, \ldots, V_6)$ $= \big(\{e, f, h\}, \{d, e, g\}, \{b, c, e, h\}, \{a, b, h\}, \{a\}, \{a\}\big)$.



- A system of representatives would make sure that there is a representative for each color group. For example,
- The representatives $(\{a, c, d, f, h\})$ are shown as colors on the left.
- Here, the set of representatives is not distinct. Why? In fact, due to the red and pink group, a distinct group of representatives is impossible (since there is only one common choice to represent both color groups).

System of Representatives

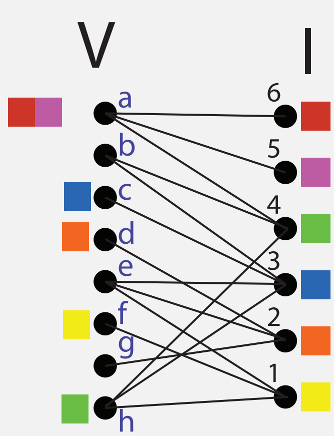
- We can view this as a bipartite graph. The groups of V are marked by color tags on the left, and also via right neighbors in the graph.
- Here, $\ell = 6$ groups, with $\mathcal{V} = (V_1, V_2, \dots, V_6)$ $= \left(\boxed{\{e, f, h\}}, \boxed{\{d, e, g\}}, \boxed{\{b, c, e, h\}}, \boxed{\{a, b, h\}}, \boxed{\{a\}}, \boxed{\{a\}} \right)$.



- A system of representatives would make sure that there is a representative for each color group. For example,
- The representatives $(\{a, c, d, f, h\})$ are shown as colors on the left.
- Here, the set of representatives is not distinct. Why? In fact, due to the red and pink group, a distinct group of representatives is impossible (since there is only one common choice to represent both color groups).

System of Representatives

- We can view this as a bipartite graph. The groups of V are marked by color tags on the left, and also via right neighbors in the graph.
- Here, $\ell = 6$ groups, with $\mathcal{V} = (V_1, V_2, \dots, V_6)$ $= \left(\boxed{\{e, f, h\}}, \boxed{\{d, e, g\}}, \boxed{\{b, c, e, h\}}, \boxed{\{a, b, h\}}, \boxed{\{a\}}, \boxed{\{a\}} \right)$.



- A system of representatives would make sure that there is a representative for each color group. For example,
- The representatives $(\{a, c, d, f, h\})$ are shown as colors on the left.
- Here, the set of representatives is not distinct. Why? In fact, due to the red and pink group, a distinct group of representatives is impossible (since there is only one common choice to represent both color groups).

System of Distinct Representatives

- Let $(V, \mathcal{V})$ be a set system (i.e., $\mathcal{V} = (V_i : i \in I)$ where $V_i \subseteq V$ for all i), and I is an index set. Hence, $|I| = |\mathcal{V}|$.
- A family $(v_i : i \in I)$ with $v_i \in V$ is said to be a system of distinct representatives of $\mathcal{V}$ if $\exists$ a bijection $\pi : I \leftrightarrow I$ such that $v_i \in V_{\pi(i)}$ and $v_i \neq v_j$ for all $i \neq j$.
- In a system of distinct representatives, there is a requirement for the representatives to be distinct. We can re-state (and rename) this as a:

Definition 6.8.1 (transversal)

Given a set system $(V, \mathcal{V})$ and index set I for $\mathcal{V}$ as defined above, a set $T \subseteq V$ is a transversal of $\mathcal{V}$ if there is a bijection $\pi : T \leftrightarrow I$ such that

$$x \in V_{\pi(x)} \text{ for all } x \in T \tag{6.34}$$

- Note that due to $\pi : T \leftrightarrow I$ being a bijection, all of I and T are "covered" (so this makes things distinct automatically).