Submodular Functions, Optimization, and Applications to Machine Learning

— Spring Quarter, Lecture 4 —

http://j.ee.washington.edu/~bilmes/classes/ee596b_spring_2014/

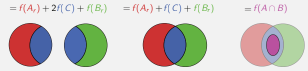
Prof. Jeff Bilmes

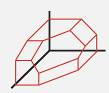
University of Washington, Seattle
Department of Electrical Engineering
http://melodi.ee.washington.edu/~bilmes

April 9th, 2014

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

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





Cumulative Outstanding Reading

- Read chapter 1 from Fujishige's book.

Announcements, Assignments, and Reminders

- our room (Mueller Hall Room 154) is changed!
- Please do use our discussion board (https://canvas.uw.edu/courses/895956/discussion_topics) for all questions, comments, so that all will benefit from them being answered.
- Weekly Office Hours: Wednesdays, 5:00-5:50, or by skype or google hangout (email me).

Class Road Map - IT-I

- L1 (3/31): Motivation, Applications, & Basic Definitions
- L2: (4/2): Applications, Basic Definitions, Properties
- L3: More examples and properties (e.g., closure properties), and examples, spanning trees
- L4: proofs of equivalent definitions, independence, start matroids
- L5:
- L6:
- L7:
- L8:
- L9:
- L10:
- L11:
- L12:
- L13:
- L14:
- L15:
- L16:
- L17:
- L18:
- L19:
- L20:

Finals Week: June 9th-13th, 2014.

Summary so far

- Summing: if $\alpha_i \geq 0$ and $f_i : 2^V \to \mathbb{R}$ is submodular, then so is $\sum_i \alpha_i f_i$.
- Restrictions: $f'(A) = f(A \cap S)$
- max: $f(A) = \max_{j \in A} c_j$ and facility location.
- Log determinant $f(A) = \log \det(\mathbf{\Sigma}_A)$

Concave over non-negative modular

Let $m \in \mathbb{R}_+^E$ be a modular function, and g a concave function over $\mathbb{R}$. Define $f : 2^E \to \mathbb{R}$ as

$$f(A) = g(m(A)) \tag{4.35}$$

then f is submodular.

Proof.

Given $A \subseteq B \subseteq E \setminus v$, we have $0 \leq a = m(A) \leq b = m(B)$, and $0 \leq c = m(v)$. For g concave, we have $g(a+c) - g(a) \geq g(b+c) - g(b)$, and thus

$$g(m(A) + m(v)) - g(m(A)) \geq g(m(B) + m(v)) - g(m(B)) \tag{4.36}$$

□

A form of converse is true as well.

Concave composed with non-negative modular

Theorem 4.2.1

Given a ground set V. The following two are equivalent:

1. *For all modular functions $m : 2^V \to \mathbb{R}_+$, then $f : 2^V \to \mathbb{R}$ defined as $f(A) = g(m(A))$ is submodular*
2. *$g : \mathbb{R}_+ \to \mathbb{R}$ is concave.*

- If g is non-decreasing concave, then f is polymatroidal.
- Sums of concave over modular functions are submodular

$$f(A) = \sum_{i=1}^{K} g_i(m_i(A)) \tag{4.35}$$

- Very large class of functions, including graph cut, bipartite neighborhoods, set cover (Stobbe & Krause).
- However, Vondrak showed that a graphic matroid rank function over K_4 (we'll define this after we define matroids) are not members.

Composition of non-decreasting submodular and non-decreasing concave

Theorem 4.2.1

Given two functions, one defined on sets

$$f : 2^V \to \mathbb{R} \tag{4.35}$$

and another continuous valued one:

$$g : \mathbb{R} \to \mathbb{R} \tag{4.36}$$

the composition formed as $h = g \circ f : 2^V \to \mathbb{R}$ (defined as $h(S) = g(f(S))$) is nondecreasing submodular, if g is non-decreasing concave and f is nondecreasing submodular.

Monotone difference of two functions

Let f and g both be submodular functions on subsets of V and let $(f - g)(\cdot)$ be either monotone increasing or monotone decreasing. Then $h : 2^V \to R$ defined by

$$h(A) = \min(f(A), g(A)) \quad (4.35)$$

is submodular.

Proof.

If $h(A)$ agrees with either f or g on both X and Y, and since

$$f(X) + f(Y) \geq f(X \cup Y) + f(X \cap Y) \quad (4.36)$$
$$g(X) + g(Y) \geq g(X \cup Y) + g(X \cap Y), \quad (4.37)$$

the result (Equation 4.35) follows since

$$\begin{matrix} f(X) + f(Y) \\ g(X) + g(Y) \end{matrix} \geq \min(f(X \cup Y), g(X \cup Y)) + \min(f(X \cap Y), g(X \cap Y)) \quad (4.38)$$

...

Saturation via the $\min(\cdot)$ function

Let $f : 2^V \to \mathbb{R}$ be an monotone increasing or decreasing submodular function and let k be a constant. Then the function $h : 2^V \to \mathbb{R}$ defined by

$$h(A) = \min(k, f(A)) \quad (4.37)$$

is submodular.

Proof.

For constant k, we have that $(f - k)$ is increasing (or decreasing) so this follows from the previous result. □

Note also, $g(a) = \min(k, a)$ for constant k is a non-decreasing concave function, so when f is monotone nondecreasing submodular, we can use the earlier result about composing a monotone concave function with a monotone submodular function to get a version of this.

Gain Notation

It will also be useful to extend this to sets.
Let A, B be any two sets. Then

$$f(A|B) \triangleq f(A \cup B) - f(B) \tag{4.41}$$

So when j is any singleton

$$f(j|B) = f(\{j\}|B) = f(\{j\} \cup B) - f(B) \tag{4.42}$$

Note that this is inspired from information theory and the notation used for conditional entropy $H(X_A|X_B) = H(X_A, X_B) - H(X_B)$.

Other properties

- Any submodular function $h : 2^V \to \mathbb{R}$ can be represented as the difference between two submodular functions, i.e., $h(A) = f(A) - g(A)$ where both f and g are submodular.
- Any submodular function f can be represented as a sum of a normalized monotone non-decreasing submodular function and a modular function, $f = \bar{f} + m$
- Any function h can be represented as the difference between two monotone non-decreasing submodular functions.

Submodular Definitions

Definition 4.3.2 (submodular concave)

A function $f : 2^V \to \mathbb{R}$ is submodular if for any $A, B \subseteq V$, we have that:

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B) \tag{4.2}$$

An alternate and (as we will soon see) equivalent definition is:

Definition 4.3.3 (diminishing returns)

A function $f : 2^V \to \mathbb{R}$ is submodular if for any $A \subseteq B \subset V$, and $v \in V \setminus B$, we have that:

$$f(A \cup \{v\}) - f(A) \geq f(B \cup \{v\}) - f(B) \tag{4.3}$$

This means that the incremental "value", "gain", or "cost" of v decreases (diminishes) as the context in which v is considered grows from A to B.

Submodular Definition: Group Diminishing Returns

An alternate and equivalent definition is:

Definition 4.3.1 (group diminishing returns)

A function $f : 2^V \to \mathbb{R}$ is submodular if for any $A \subseteq B \subset V$, and $C \subseteq V \setminus B$, we have that:

$$f(A \cup C) - f(A) \geq f(B \cup C) - f(B) \tag{4.1}$$

This means that the incremental "value" or "gain" of set C decreases as the context in which C is considered grows from A to B (diminishing returns)

Submodular Definition Basic Equivalencies

We want to show that Submodular Concave (Definition 4.3.2), Diminishing Returns (Definition 4.3.3), and Group Diminishing Returns (Definition 4.3.1) are identical. We will show that:

- Submodular Concave $\Rightarrow$ Diminishing Returns
- Diminishing Returns $\Rightarrow$ Group Diminishing Returns
- Group Diminishing Returns $\Rightarrow$ Submodular Concave

Submodular Concave $\Rightarrow$ Diminishing Returns

$f(S)+f(T) \geq f(S\cup T)+f(S\cap T) \Rightarrow f(v|A) \geq f(v|B), A \subseteq B \subseteq V \setminus v.$

- Assume Submodular concave, so $\forall S, T$ we have $f(S) + f(T) \geq f(S \cup T) + f(S \cap T)$.
- Given A, B and $v \in V$ such that: $A \subseteq B \subseteq V \setminus \{v\}$, we have from submodular concave that:

$$f(A + v) + f(B) \geq f(B + v) + f(A) \tag{4.2}$$

- Rearranging, we have

$$f(A + v) - f(A) \geq f(B + v) - f(B) \tag{4.3}$$

$\square$

Diminishing Returns $\Rightarrow$ Group Diminishing Returns

$f(v|S) \geq f(v|T), S \subseteq T \subseteq V \setminus v \Rightarrow f(C|A) \geq f(C|B), A \subseteq B \subseteq V \setminus C$.

Let $C = \{c_1, c_2, \dots, c_k\}$. Then diminishing returns implies

$$f(A \cup C) - f(A) \tag{4.4}$$

$$= f(A \cup C) - \sum_{i=1}^{k-1} \Big(f(A \cup \{c_1, \dots, c_i\}) - f(A \cup \{c_1, \dots, c_i\}) \Big) - f(A) \tag{4.5}$$

$$= \sum_{i=1}^{k} \Big(f(A \cup \{c_1 \dots c_i\}) - f(A \cup \{c_1 \dots c_{i-1}\}) \Big) \tag{4.6}$$

$$\geq \sum_{i=1}^{k} \Big(f(B \cup \{c_1 \dots c_i\}) - f(B \cup \{c_1 \dots c_{i-1}\}) \Big) \tag{4.7}$$

$$= f(B \cup C) - \sum_{i=1}^{k-1} \Big(f(B \cup \{c_1, \dots, c_i\}) - f(B \cup \{c_1, \dots, c_i\}) \Big) - f(B) \tag{4.8}$$

$$= f(B \cup C) - f(B) \tag{4.9}$$

□

Group Diminishing Returns $\Rightarrow$ Submodular Concave

$f(U|S) \geq f(U|T), S \subseteq T \subseteq V \setminus U \Rightarrow f(A) + f(B) \geq f(A \cup B) + f(A \cap B)$.

Assume group diminishing returns. Assume $A \neq B$ otherwise trivial. Define $A' = A \cap B$, $C = A \setminus B$, and $B' = B$. Then since $A' \subseteq B'$,

$$f(A' + C) - f(A') \geq f(B' + C) - f(B') \tag{4.10}$$

giving

$$f(A' + C) + f(B') \geq f(B' + C) + f(A') \tag{4.11}$$

or

$$f(A \cap B + A \setminus B) + f(B) \geq f(B + A \setminus B) + f(A \cap B) \tag{4.12}$$

which is the same as the submodular concave condition

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B) \tag{4.13}$$

Submodular Definition: Four Points

Definition 4.3.2 ("singleton", or "four points")

A function $f : 2^V \to \mathbb{R}$ is submodular iff for any $A \subset V$, and any $a, b \in V \setminus A$, we have that:

$$f(A \cup \{a\}) + f(A \cup \{b\}) \geq f(A \cup \{a,b\}) + f(A) \quad (4.14)$$

This follows immediately from diminishing returns. To achieve diminishing returns, assume $A \subset B$ with $B \setminus A = \{b_1, b_2, \ldots, b_k\}$. Then

$$\begin{aligned} f(A+a) - f(A) &\geq f(A + b_1 + a) - f(A + b_1) && (4.15)\\ &\geq f(A + b_1 + b_2 + a) - f(A + b_1 + b_2) && (4.16)\\ &\geq \ldots && (4.17)\\ &\geq f(A + b_1 + \cdots + b_k + a) - f(A + b_1 + \cdots + b_k) && (4.18)\\ &= f(B + a) - f(B) && (4.19) \end{aligned}$$

Submodular Definitions

Theorem 4.3.3

Given function $f : 2^V \to \mathbb{R}$, then

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B) \text{ for all } A, B \subseteq V \quad \text{(SC)}$$

if and only if

$$f(v|X) \geq f(v|Y) \text{ for all } X \subseteq Y \subseteq V \text{ and } v \notin B \quad \text{(DR)}$$

Proof.

(SC)⇒(DR): Set $A \leftarrow X \cup \{v\}$, $B \leftarrow Y$. Then $A \cup B = B \cup \{v\}$ and $A \cap B = X$ and $f(A) - f(A \cap B) \geq f(A \cup B) - f(B)$ implies (DR).

(DR)⇒(SC): Order $A \setminus B = \{v_1, v_2, \ldots, v_r\}$ arbitrarily. Then $f(v_i | A \cap B \cup \{v_1, v_2, \ldots, v_{i-1}\}) \geq f(v_i | B \cup \{v_1, v_2, \ldots, v_{i-1}\}),\ i \in [r-1]$

Applying telescoping summation to both sides, we get:

$$\sum_{i=0}^{r} f(v_i | A \cap B \cup \{v_1, v_2, \ldots, v_{i-1}\}) \geq \sum_{i=0}^{r} f(v_i | B \cup \{v_1, v_2, \ldots, v_{i-1}\})$$

or

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

$\square$

Use of gain: submodular bounds of a difference

- Given submodular f, and given you have $C, D \subseteq E$ with either $D \supseteq C$ or $D \subseteq C$, and have an expression of the form:

$$f(C) - f(D) \tag{4.20}$$

- If $D \supseteq C$, then for any X with $D = C \cup X$ then

$$f(C) - f(D) = f(C) - f(C \cup X) \geq f(C \cap X) - f(X) \tag{4.21}$$

or

$$f(C \cup X | C) \leq f(X | C \cap X) \tag{4.22}$$

- Alternatively, if $D \subseteq C$, given any Y such that $D = C \cap Y$ then

$$f(C) - f(D) = f(C) - f(C \cap Y) \geq f(C \cup Y) - f(Y) \tag{4.23}$$

or

$$f(C | C \cap Y) \geq f(C \cup Y | Y) \tag{4.24}$$

- Equations (4.22) and (4.24) have same form.

Many (Equivalent) Definitions of Submodularity

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B), \quad \forall A, B \subseteq V \tag{4.25}$$

$$f(j|S) \geq f(j|T), \ \forall S \subseteq T \subseteq V, \text{ with } j \in V \setminus T \tag{4.26}$$

$$f(C|S) \geq f(C|T), \forall S \subseteq T \subseteq V, \text{ with } C \subseteq V \setminus T \tag{4.27}$$

$$f(j|S) \geq f(j|S \cup \{k\}), \ \forall S \subseteq V \text{ with } j \in V \setminus (S \cup \{k\}) \tag{4.28}$$

$$f(A \cup B | A \cap B) \leq f(A | A \cap B) + f(B | A \cap B), \quad \forall A, B \subseteq V \tag{4.29}$$

$$f(T) \leq f(S) + \sum_{j \in T \setminus S} f(j|S) - \sum_{j \in S \setminus T} f(j|S \cup T - \{j\}), \ \forall S, T \subseteq V \tag{4.30}$$

$$f(T) \leq f(S) + \sum_{j \in T \setminus S} f(j|S), \ \forall S \subseteq T \subseteq V \tag{4.31}$$

$$f(T) \leq f(S) - \sum_{j \in S \setminus T} f(j|S \setminus \{j\}) + \sum_{j \in T \setminus S} f(j|S \cap T) \ \forall S, T \subseteq V \tag{4.32}$$

$$f(T) \leq f(S) - \sum_{j \in S \setminus T} f(j|S \setminus \{j\}), \ \forall T \subseteq S \subseteq V \tag{4.33}$$

Equivalent Definitions of Submodularity

We've already seen that Eq. 4.25 $\equiv$ Eq. 4.26 $\equiv$ Eq. 4.27 $\equiv$ Eq. 4.28 $\equiv$ Eq. 4.29.

We next show that Eq. 4.28 $\Rightarrow$ Eq. 4.30 $\Rightarrow$ Eq. 4.31 $\Rightarrow$ Eq. 4.28.

Approach

To show these next results, we essentially first use:

$$f(S \cup T) = f(S) + f(T|S) \leq f(S) + \text{upper-bound} \tag{4.34}$$

and

$$f(T) + \text{lower-bound} \leq f(T) + f(S|T) = f(S \cup T) \tag{4.35}$$

leading to

$$f(T) + \text{lower-bound} \leq f(S) + \text{upper-bound} \tag{4.36}$$

or

$$f(T) \leq f(S) + \text{upper-bound} - \text{lower-bound} \tag{4.37}$$

Eq. 4.28 ⇒ Eq. 4.30

Let $T \setminus S = \{j_1, \dots, j_r\}$ and $S \setminus T = \{k_1, \dots, k_q\}$.
First, we upper bound the gain of T in the context of S:

$$f(S \cup T) - f(S) = \sum_{t=1}^{r} \Big(f(S \cup \{j_1, \dots, j_t\}) - f(S \cup \{j_1, \dots, j_{t-1}\}) \Big) \tag{4.38}$$

$$= \sum_{t=1}^{r} f(j_t | S \cup \{j_1, \dots, j_{t-1}\}) \leq \sum_{t=1}^{r} f(j_t|S) \tag{4.39}$$

$$= \sum_{j \in T \setminus S} f(j|S) \tag{4.40}$$

or

$$f(T|S) \leq \sum_{j \in T \setminus S} f(j|S) \tag{4.41}$$

Eq. 4.28 ⇒ Eq. 4.30

Let $T \setminus S = \{j_1, \dots, j_r\}$ and $S \setminus T = \{k_1, \dots, k_q\}$.
Next, lower bound S in the context of T:

$$f(S \cup T) - f(T) = \sum_{t=1}^{q} [f(T \cup \{k_1, \dots, k_t\}) - f(T \cup \{k_1, \dots, k_{t-1}\})] \tag{4.42}$$

$$= \sum_{t=1}^{q} f(k_t | T \cup \{k_1, \dots, k_t\} \setminus \{k_t\}) \geq \sum_{t=1}^{q} f(k_t | T \cup S \setminus \{k_t\}) \tag{4.43}$$

$$= \sum_{j \in S \setminus T} f(j | S \cup T \setminus \{j\}) \tag{4.44}$$

Eq. 4.28 ⇒ Eq. 4.30

Let $T \setminus S = \{j_1, \ldots, j_r\}$ and $S \setminus T = \{k_1, \ldots, k_q\}$.
So we have the upper bound

$$f(T|S) = f(S \cup T) - f(S) \leq \sum_{j \in T \setminus S} f(j|S) \tag{4.45}$$

and the lower bound

$$f(S|T) = f(S \cup T) - f(T) \geq \sum_{j \in S \setminus T} f(j|S \cup T \setminus \{j\}) \tag{4.46}$$

This gives upper and lower bounds of the form

$$f(T) + \text{lower bound} \leq f(S \cup T) \leq f(S) + \text{upper bound}, \tag{4.47}$$

and combining directly the left and right hand side gives the desired inequality.

Eq. 4.30 ⇒ Eq. 4.31

This follows immediately since if $S \subseteq T$, then $S \setminus T = \emptyset$, and the last term of Eq. 4.30 vanishes.

Eq. 4.31 $\Rightarrow$ Eq. 4.28

Here, we set $T = S \cup \{j,k\}$, $j \notin S \cup \{k\}$ into Eq. 4.31 to obtain

$$\begin{aligned} f(S \cup \{j,k\}) &\le f(S) + f(j|S) + f(k|S) && (4.48) \\ &= f(S) + f(S + \{j\}) - f(S) + f(S + \{k\}) - f(S) && (4.49) \\ &= f(S + \{j\}) + f(S + \{k\}) - f(S) && (4.50) \\ &= f(j|S) + f(S + \{k\}) && (4.51) \end{aligned}$$

giving

$$\begin{aligned} f(j|S \cup \{k\}) &= f(S \cup \{j,k\}) - f(S \cup \{k\}) && (4.52) \\ &\le f(j|S) && (4.53) \end{aligned}$$

Submodular Concave

- Why do we call the $f(A) + f(B) \ge f(A \cup B) + f(A \cap B)$ definition of submodularity, submodular concave?
- A continuous twice differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ is concave iff $\nabla^2 f \preceq 0$ (the Hessian matrix is nonpositive definite).
- Define a "discrete derivative" or difference operator defined on discrete functions $f : 2^V \to \mathbb{R}$ as follows:

$$(\nabla_B f)(A) \triangleq f(A \cup B) - f(A \setminus B) = f\big(B|(A \setminus B)\big) \qquad (4.54)$$

  read as: the derivative of f at A in the direction B.
- Hence, if $A \cap B = \emptyset$, then $(\nabla_B f)(A) = f(B|A)$.
- Consider a form of second derivative or 2nd difference:

$$\begin{aligned} (\nabla_C \nabla_B f)(A) &= \nabla_C[f(A \cup B) - f(A \setminus B)] \\ &= f(A \cup B \cup C) - f((A \cup C) \setminus B) \\ &\quad - f((A \setminus C) \cup B) + f((A \setminus C) \setminus B) \qquad (4.55) \end{aligned}$$

Submodular Concave

- If the second difference operator everywhere nonpositive:

$$f(A \cup B \cup C) - f((A \cup C) \setminus B) \\ - f((A \setminus C) \cup B) + f(A \setminus C \setminus B) \le 0 \qquad (4.56)$$

  then we have the equation:

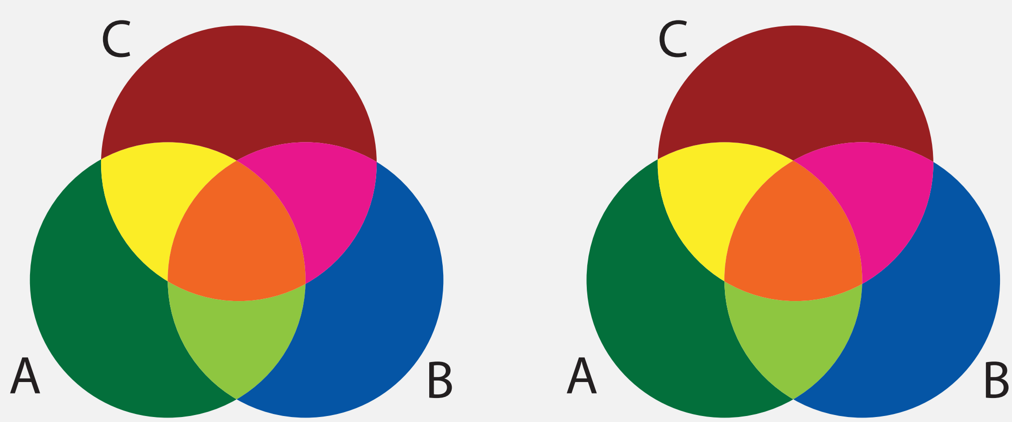
$$f((A \cup C) \setminus B) + f((A \setminus C) \cup B) \ge f(A \cup B \cup C) + f(A \setminus C \setminus B) \qquad (4.57)$$

- Define $A' = (A \cup C) \setminus B$ and $B' = (A \setminus C) \cup B$. Then the above implies:

$$f(A') + f(B') \ge f(A' \cup B') + f(A' \cap B') \qquad (4.58)$$

  and note that A' and B' so defined can be arbitrary.
- One sense in which submodular functions are like concave functions.

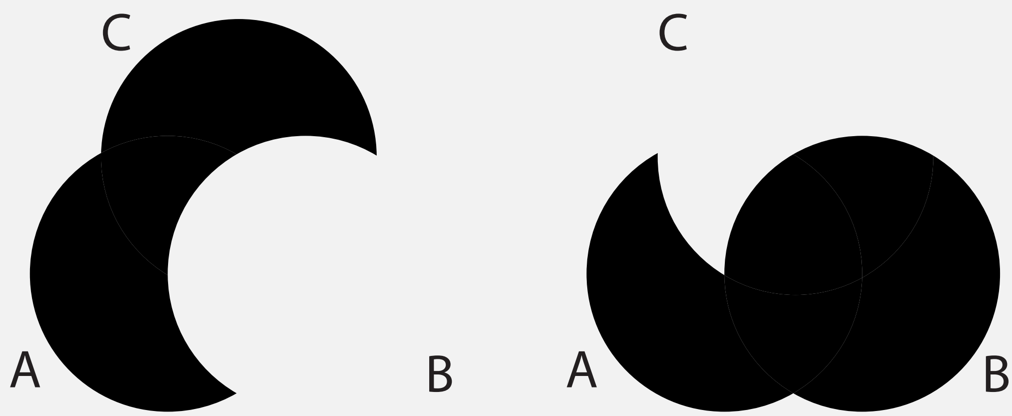
Submodular Concave



(a) $A' = (A \cup C) \setminus B$ (b) $B' = (A \setminus C) \cup B$

Figure : A figure showing $A' \cup B' = A \cup B \cup C$ and $A' \cap B' = A \setminus C \setminus B$.

Submodular Concave



(a) $A' = (A \cup C) \setminus B$ (b) $B' = (A \setminus C) \cup B$

Figure : A figure showing $A' \cup B' = A \cup B \cup C$ and $A' \cap B' = A \setminus C \setminus B$.

Submodularity and Concave

- This submodular/concave relationship is more simply done with singletons.
- Recall four points definition: A function is submodular if for all $X \subseteq V$ and $j, k \in V$

$$f(X + j) + f(X + k) \geq f(X + j + k) + f(X) \tag{4.59}$$

- This gives us a simpler notion corresponding to concavity.
- Define gain as $\nabla_j(X) = f(X + j) - f(X)$, a form of discrete gradient.
- Trivially becomes a second-order condition, akin to concave functions: A function is submodular if for all $X \subseteq V$ and $j, k \in V$, we have:

$$\nabla_j \nabla_k f(X) \leq 0 \tag{4.60}$$

Example: Rank function of a matrix

Consider the following 4×8 matrix, so $V = \{1,2,3,4,5,6,7,8\}$.

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

- Let $A = \{1,2,3\}$, $B = \{3,4,5\}$, $C = \{6,7\}$, $A_r = \{1\}$, $B_r = \{5\}$.
- Then $r(A) = 3$, $r(B) = 3$, $r(C) = 2$.
- $r(A \cup C) = 3$, $r(B \cup C) = 3$.
- $r(A \cup A_r) = 3$, $r(B \cup B_r) = 3$, $r(A \cup B_r) = 4$, $r(B \cup A_r) = 4$.
- $r(A \cup B) = 4$, $r(A \cap B) = 1 < r(C) = 2$.
- $6 = r(A) + r(B) > r(A \cup B) + r(A \cap B) = 5$

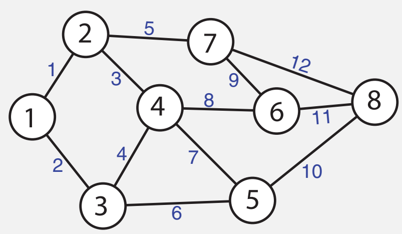
On Rank

- Let rank : $2^V \to \mathbb{Z}_+$ be the rank function.
- In general, rank$(A) \leq |A|$, and vectors in A are linearly independent if and only if rank$(A) = |A|$.
- If A, B are such that rank$(A) = |A|$ and 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.
- To stress this point, note that the above condition is $|A| < |B|$, **not** $A \subseteq B$ which is sufficient (to be able to find an independent vector) but not necessary.
- In other words, given A, B with rank$(A) = |A|$ & rank$(B) = B$, then $|A| < |B| \Leftrightarrow \exists$ an $b \in B$ such that 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 $|V| \times |E|$ incidence matrix of undirected graph G, which is the matrix $\mathbf{X}_G = (x_{v,e})_{v \in V(G), e \in E(G)}$ where

$$x_{v,e} = \begin{cases} 1 & \text{if } v \in e \\ 0 & \text{if } v \notin e \end{cases} \tag{4.61}$$



$$\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{array} \begin{array}{c} \begin{array}{cccccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \end{array} \\ \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix} \end{array} \tag{4.62}$$

Spanning trees/forests & incidence matrices

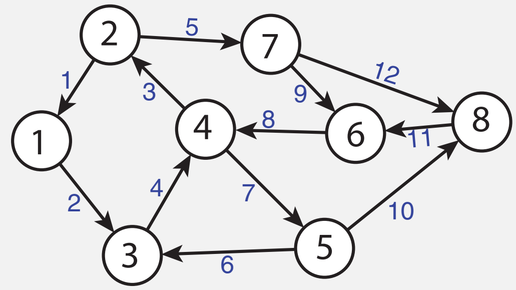
- We are given a graph $G = (V, E)$, we can arbitrarily orient the graph (make it directed) consider again the edges $E = E(G)$ as an index set.
- Consider instead the $|V| \times |E|$ incidence matrix of undirected graph G, which is the matrix $\mathbf{X}_G = (x_{v,e})_{v \in V(G), e \in E(G)}$ where

$$x_{v,e} = \begin{cases} 1 & \text{if } v \in e^+ \\ -1 & \text{if } v \in e^- \\ 0 & \text{if } v \notin e \end{cases} \tag{4.63}$$

and where e^+ is the tail and e^- is the head of (now) directed edge e.

Spanning trees/forests & incidence matrices

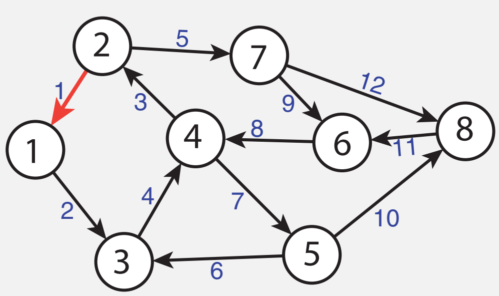
- A directed version of the graph (right) and its adjacency matrix (below).
- Orientation can be arbitrary.
- Note, rank of this matrix is 7.



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

Spanning trees

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

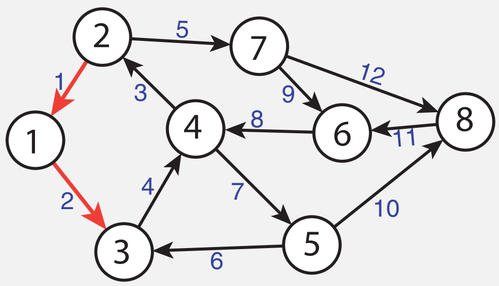


$$
\begin{array}{c}
1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8
\end{array}
\overset{\displaystyle 1}{\begin{pmatrix}
-1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0
\end{pmatrix}} \tag{4.64}
$$

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

Spanning trees

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

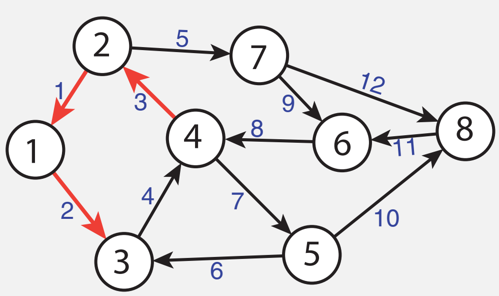


$$\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{array}\begin{array}{c} \begin{array}{cc} 1 & 2 \end{array} \\ \begin{pmatrix} -1 & 1 \\ 1 & 0 \\ 0 & -1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \end{array} \tag{4.64}$$

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

Spanning trees

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

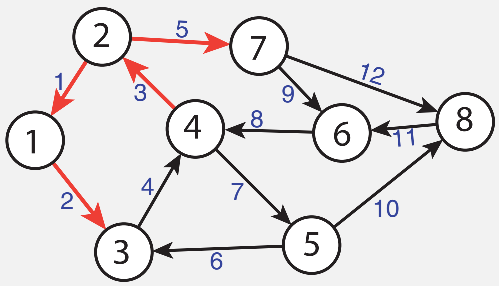


$$\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{array}\begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} -1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{array} \tag{4.64}$$

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

Spanning trees

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

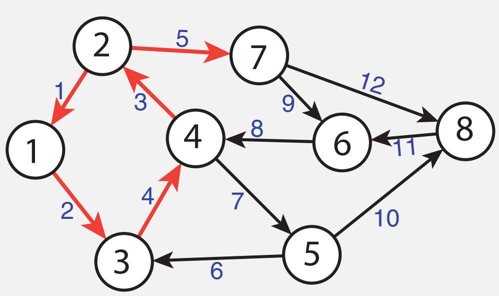


$$\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 5 \end{array} \\ \begin{pmatrix} -1 & 1 & 0 & 0 \\ 1 & 0 & -1 & 1 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array} \tag{4.64}$$

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.

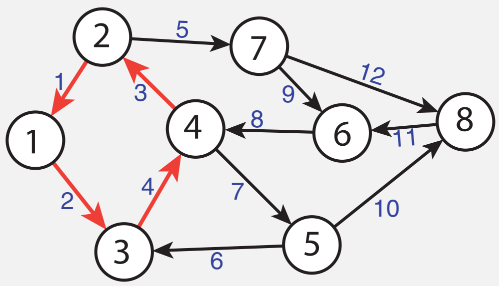


$$\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{pmatrix} -1 & 1 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 1 \\ 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array} \tag{4.64}$$

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.



$$\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{array}
\begin{array}{c} \begin{array}{cccc} \;\;1 & \;\;\;2 & \;\;\;3 & \;\;\;4 \end{array} \\
\begin{pmatrix}
-1 & 1 & 0 & 0 \\
1 & 0 & -1 & 0 \\
0 & -1 & 0 & 1 \\
0 & 0 & 1 & -1 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{pmatrix} \end{array} \tag{4.64}$$

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

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 (recall, we saw that $k_G(A)$ is a supermodular function in previous lectures).

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.
- Given a tree T, the cost of the tree is $\text{cost}(T) = \sum_{e \in T} w(e)$, the sum of the weights of the edges.
- There are several algorithms for MST:

Algorithm 1: Borůvka's Algorithm

```
F ← ∅ /* We build up the edges of a forest in F */
while G(V, F) is disconnected do
    forall the components C_i of F do
        F ← F ∪ {e_i} for e_i = the min-weight edge out of C_i;
```


Definitions of Submodularity | Independence | Matroids | Matroid Examples | Matroid Rank | Partition Matroid | System of Distinct Reps

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.
- Given a tree T, the cost of the tree is $\text{cost}(T) = \sum_{e \in T} w(e)$, the sum of the weights of the edges.
- There are several algorithms for MST:

Algorithm 2: Jarník/Prim/Dijkstra Algorithm

```
T ← ∅ ;
while T is not a spanning tree do
    T ← T ∪ {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 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.
- Given a tree T, the cost of the tree is $\text{cost}(T) = \sum_{e \in T} w(e)$, the sum of the weights of the edges.
- There are several algorithms for MST:

Algorithm 3: Kruskal's Algorithm

```
Sort the edges so that w(e_1) ≤ w(e_2) ≤ ··· ≤ w(e_m) ;
T ← (V(G), ∅) = (V, E) ;
for i = 1 to m do
    if E(T) ∪ {e_i} does not create a cycle in T then
        E(T) ← E(T) ∪ {e} ;
```


Definitions of Submodularity Independence Matroids Matroid Examples Matroid Rank Partition Matroid System of Distinct Reps

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.
- Given a tree T, the cost of the tree is $\text{cost}(T) = \sum_{e \in T} w(e)$, the sum of the weights of the edges.
- There are several algorithms for MST:
- These three algorithms are all guaranteed to find the optimal minimum spanning tree in (low order) polynomial time.
- These algorithms are all 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.

From Matrix Rank → Matroid

- So V is set of column vector indices of a matrix.
- Let $\mathcal{I}$ 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,

$$A \subseteq B \text{ and } B \in \mathcal{I} \Rightarrow A \in \mathcal{I} \tag{4.65}$$

- maxInd: Inclusionwise maximal independent subsets (or bases) of B.

$$\text{maxInd}(B) \triangleq \{A \subseteq B : A \in \mathcal{I} \text{ and } \forall v \in B \setminus A, A \cup \{v\} \notin \mathcal{I}\} \tag{4.66}$$

- 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 $B \subseteq V$,

$$\forall A_1, A_2 \in \text{maxInd}(B), \quad |A_1| = |A_2| \tag{4.67}$$

From Matrix Rank → Matroid

- Thus, for all $I \in \mathcal{I}$, the matrix rank function has the property

$$r(I) = |I| \tag{4.68}$$

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

$$r(B) = \max\{|A| : A \subseteq B \text{ and } A \in \mathcal{I}\} \leq |B| \tag{4.69}$$

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

Definition 4.5.1 (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 method to determine if a given set $S \subseteq E$ has $S \in \mathcal{I}$.
- One useful property is "heredity." Namely, a set system is a hereditary set system if for any $A \subset B \in \mathcal{I}$, we have that $A \in \mathcal{I}$.

Independence System

Definition 4.5.2 (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 is 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 (i.e., 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

$$\begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array} \\ \begin{pmatrix} 0 & 0 & 1 & 1 & 2 & 1 & 3 & 1 \\ 0 & 1 & 1 & 0 & 2 & 0 & 2 & 4 \\ 1 & 1 & 1 & 0 & 0 & 3 & 1 & 5 \end{pmatrix} \end{array} = \begin{array}{c} \begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array} \\ \begin{pmatrix} | & | & | & | & | & | & | & | \\ x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 \\ | & | & | & | & | & | & | & | \end{pmatrix} \end{array} \tag{4.70}$$

- 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 4.5.3 (Matroid)

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

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

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

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

Matroid

Slight modification (non unit increment) that is equivalent.

Definition 4.5.4 (Matroid-II)

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

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

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

(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).

Matroids - important property

Proposition 4.5.5

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 4.5.6 (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

- Recall, 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 4.5.7 (matroid rank function)

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

$$r(A) = \max \{|X| : X \subseteq A, X \in \mathcal{I}\} = \max_{X \in \mathcal{I}} |A \cap X| \tag{4.71}$$

- 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 4.5.8 (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 4.5.9 (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 4.5.10 (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 4.5.11 (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 4.5.12 (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 $C \in C_1 \cap C_2$, then there exists a $C_3 \in \mathcal{C}$ such that $C_3 \subseteq (C_1 \cup C_2) \setminus \{e\}$.*

Matroids by circuits

Several circuit definitions for matroids.

Theorem 4.5.13 (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.

Matroids by submodular functions

Theorem 4.5.14 (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:

$$\mathcal{C}(f) = \Big\{ C \subseteq E : C \text{ is non-empty, is inclusionwise-minimal, and has } f(C) < |C| \Big\} \quad (4.72)$$

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

Inclusionwise-minimal 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 $|C'| \leq f(C')$).

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

$$r(A) = \begin{cases} |A| & \text{if } |A| \leq k \\ k & \text{if } |A| > k \end{cases} \tag{4.73}$$

- 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

$$\text{span}(A) = \begin{cases} A & \text{if } |A| < k, \\ E & \text{if } |A| \geq k, \end{cases} \tag{4.74}$$

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

Linear (or Matric) Matroid

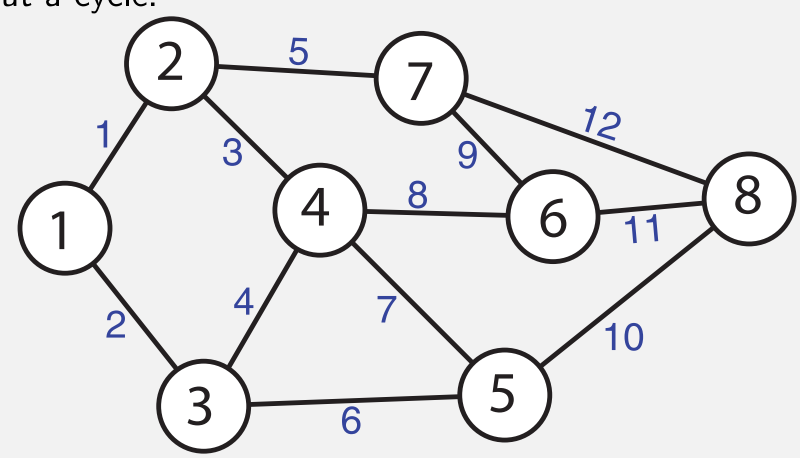
- Let $\mathbf{X}$ be an $n \times m$ matrix and $E = \{1, \ldots, m\}$
- Let $\mathcal{I}$ consists of subsets of E such that if $A \in \mathcal{I}$, and $A = \{a_1, a_2, \ldots, a_k\}$ then the vectors $x_{a_1}, x_{a_2}, \ldots, x_{a_k}$ are linearly independent.
- the rank function is just the rank 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 $G(V, A)$.
- Closure function adds all edges between the vertices adjacent to any edge in A. Closure of a spanning forest is G.

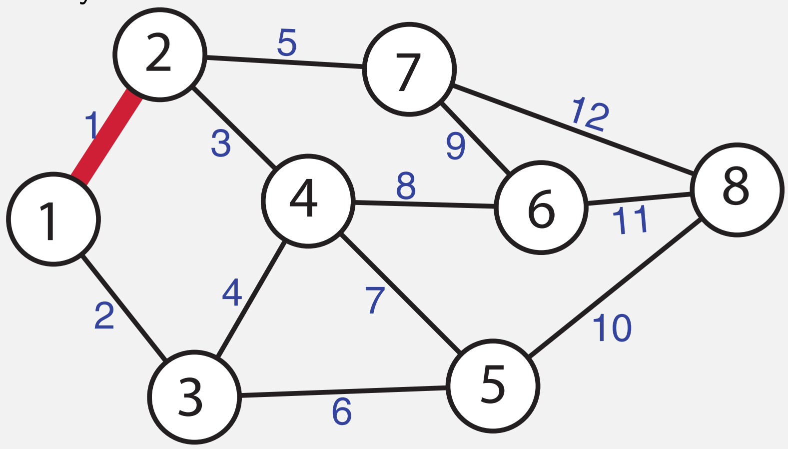
Example: graphic matroid

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



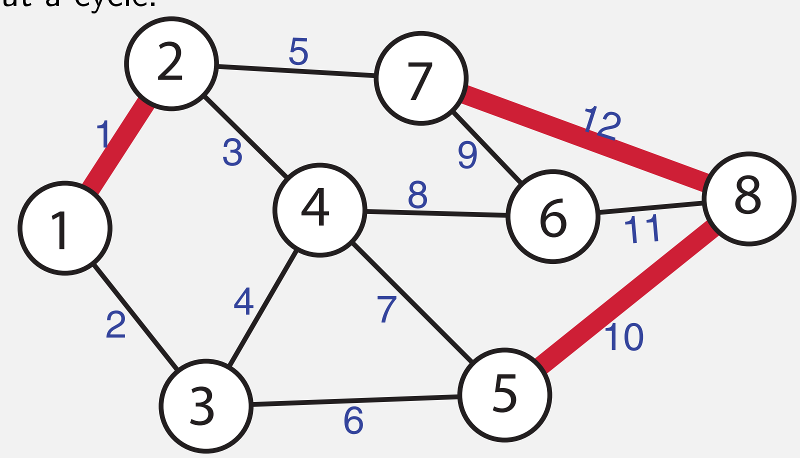
Example: graphic matroid

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



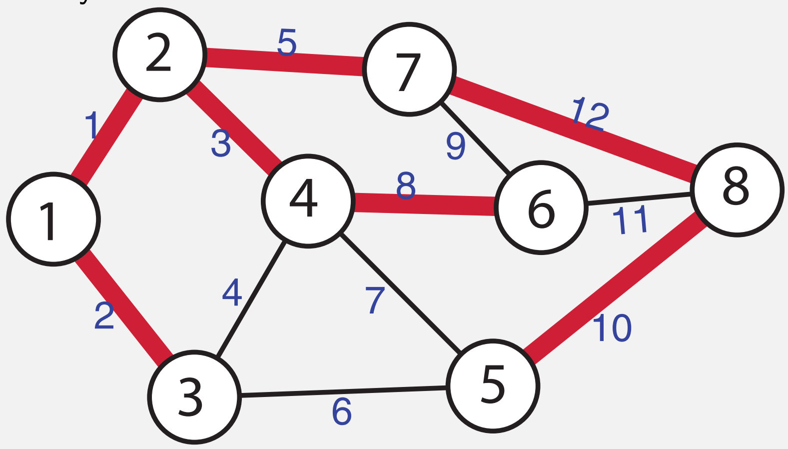
Example: graphic matroid

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



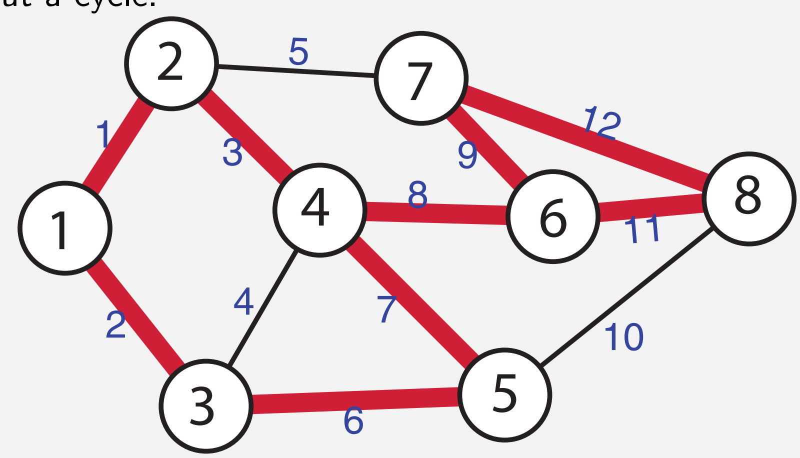
Example: graphic matroid

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



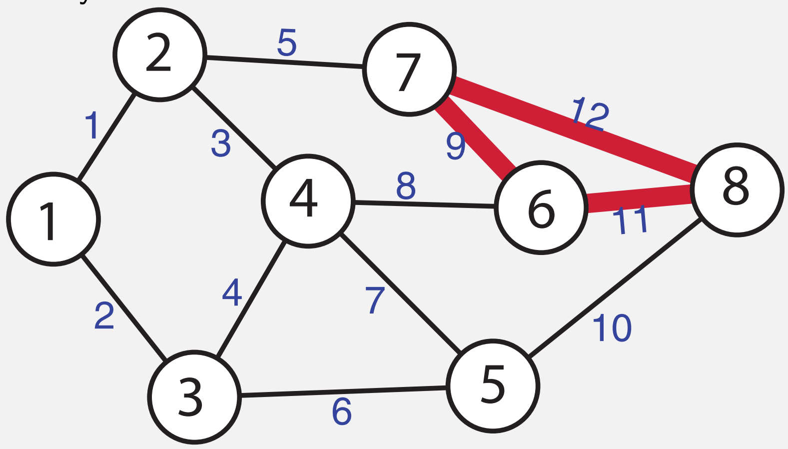
Example: graphic matroid

- 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 blocks or disjoint sets (disjoint union). Define a set of subsets of V as

$$\mathcal{I} = \{X \subseteq V : |X \cap V_i| \leq k_i \text{ for all } i = 1, \ldots, \ell\}. \tag{4.75}$$

  where $k_1, \ldots, k_\ell$ are fixed 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 $k_1 = k$.
- We'll show that property (I3') in Def 4.5.6 holds. 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, $V_1, V_2, \dots, V_6$

$\Bigg\}$

Partition Matroid

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

}

Definitions of Submodularity Independence Matroids Matroid Examples Matroid Rank Partition Matroid 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.

Matroids - rank

Lemma 4.7.1

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

$$r(A) + r(B) \geq |Y \cap A| + |Y \cap B| \quad (4.76)$$
$$= |Y \cap (A \cap B)| + |Y \cap (A \cup B)| \quad (4.77)$$
$$\geq |X| + |Y| = r(A \cap B) + r(A \cup B) \quad (4.78)$$

□

Definitions of Submodularity Independence Matroids Matroid Examples Matroid Rank Partition Matroid System of Distinct Reps

Matroids

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

Theorem 4.7.2 (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)*

- 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.
- A matroid is sometimes given as (E, r) where E is ground set and r is rank function.

Matroids

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

Theorem 4.7.2 (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 $r(A \cup \{v\})$.

Matroids from rank

Proof of Theorem 4.7.2 (matroid from rank).

- Given a matroid $M = (E, \mathcal{I})$, we see its rank function as defined in Eq. 4.71 satisfies (R1), (R2), and, as we saw in Lemma 4.7.1, (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,

$$r(X) \geq r(Y) - r(Y \setminus X) - r(\emptyset) \tag{4.79}$$
$$\geq |Y| - |Y \setminus X| \tag{4.80}$$
$$= |X| \tag{4.81}$$

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

...

Matroids from rank

Proof of Theorem 4.7.2 (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 $k \le |B|$).
- 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

$$\begin{aligned}
r(B) &\le r(A \cup B) && (4.82)\\
&\le r(A \cup (B \setminus \{b_1\})) + r(A \cup \{b_1\}) - r(A) && (4.83)\\
&= r(A \cup (B \setminus \{b_1\}) && (4.84)\\
&\le r(A \cup (B \setminus \{b_1, b_2\})) + r(A \cup \{b_2\}) - r(A) && (4.85)\\
&= r(A \cup (B \setminus \{b_1, b_2\})) && (4.86)\\
&\le \dots \le r(A) = |A| < |B| && (4.87)
\end{aligned}$$

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

Matroids from rank II

Another way of using function r to define a matroid.

Theorem 4.7.3 (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 $A \subseteq E$, 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* $r(X \cup \{x, y\}) = r(X)$.

Matroid and Rank

- Thus, we can define a matroid as $M = (V, r)$ where r satisfies matroid rank axioms.
- Example: 2-partition matroid rank function: Given natural numbers $a, b \in \mathbb{Z}_+$ with $a > b$, and any set $R \subseteq V$ with $|R| = a$, two-block partition $V = (R, \bar{R})$, define:

$$r(A) = \min(|A \cap R|, b) + \min(|A \cap \bar{R}|, |\bar{R}|) \quad (4.88)$$

$$= \min(|A \cap R|, b) + |A \cap \bar{R}| \quad (4.89)$$

- Example: Truncated matroid rank function.

$$f_R(A) = \min\{r(A), a\} \quad (4.90)$$

$$= \min\left\{|A|, b + |A \cap \bar{R}|, a\right\} \quad (4.91)$$

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

$$\mathcal{I} = \{I \subseteq V : |I| \leq a \text{ and } |I \cap R| \leq b\}, \quad (4.92)$$

useful for showing hardness of constrained submodular minimization.

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, submodular)

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 covering set of sets).

Partition Matroid

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

$$r(A) = \sum_{i=1}^{\ell} \min(|A \cap V_i|, k_i) \tag{4.93}$$

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

1. $|A \cap V_i|$ is submodular (even 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.

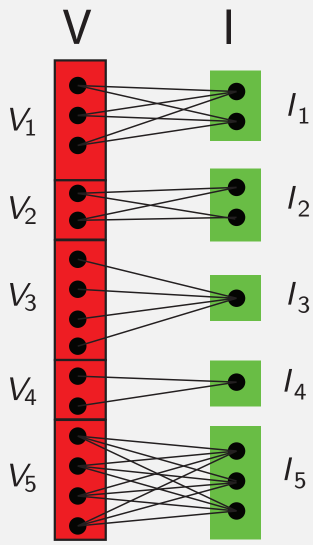
- $r(A)$ is also non-negative integral monotone non-decreasing, so it defines a matroid (the partition matroid).

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 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.
- $(v, i) \in E(G)$ iff $v \in V_j$ and $i \in I_j$.

Partition Matroid, rank as matching

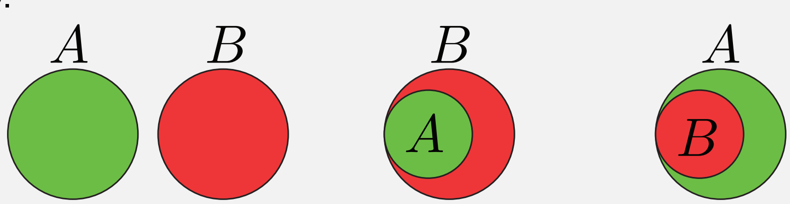
- Example where $\ell = 5$, $(k_1, k_2, k_3, k_4, k_5) = (2, 2, 1, 1, 3)$.



- 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 $r(X) = \sum_{i=1}^{\ell} \min(|X \cap V_i|, k_i) =$ maximum matching involving X.

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 if it has no two "properly intersecting" members: i.e., intersecting $A \cap B \neq \emptyset$ and not comparable (one is not contained in the other).
- 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

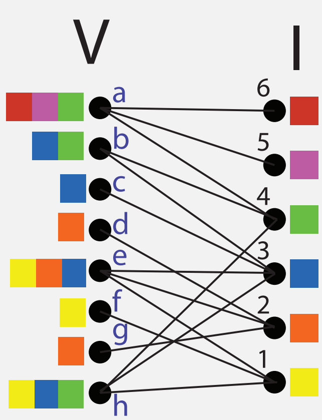
$$\mathcal{I} = \{I \subseteq E : |X \cap A| \leq k(A) \text{ for all } A \in \mathcal{F}\} \tag{4.94}$$

System of Representatives

- Let $(V, \mathcal{V})$ be a set system (i.e., $\mathcal{V} = (V_k : i \in I)$ where $\emptyset \subset V_i \subseteq V$ for all i).
- A family $(v_i : i \in I)$ with $v_i \in V$ for index set I 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 $\pi(i)$, meaning the i^{th} representative is meant to represent set $V_{\pi(i)}$. Consider the house of representatives, v_i = "John Smith", while i = King County.*
- In a system of representatives, there is no requirement for the representatives to be distinct. I.e., we could have $v_1 \in T$, where v_1 represents both V_1 and V_2.
- We can view this as a bipartite graph.

System of Representatives

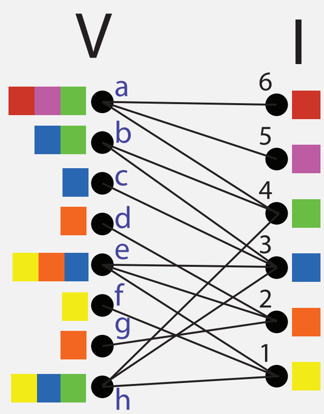
- We can view this as a bipartite graph. The groups of V are marked by color tags on the left, and also via right neighbors in the graph.
- Here, $\ell = 6$, and $\mathcal{V} = (V_1, V_2, \ldots, V_6)$ $= (\{e, f, h\}, \{d, e, g\}, \{b, c, e, h\}, \{a, b, h\}, \{a\}, \{a\})$.



- A system of representatives would make sure that there is a representative for each color group. For example,
- The representatives are shown as colors on the left.
- Here, the set of representatives is not distinct. 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, $\ell = 6$, and $\mathcal{V} = (V_1, V_2, \ldots, V_6)$ $= (\{e, f, h\}, \{d, e, g\}, \{b, c, e, h\}, \{a, b, h\}, \{a\}, \{a\})$.

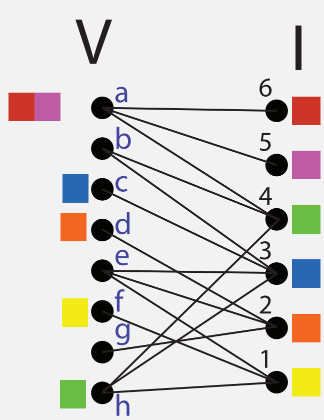


- A system of representatives would make sure that there is a representative for each color group. For example,
- The representatives are shown as colors on the left.
- Here, the set of representatives is not distinct. In fact, due to the red and pink group, a distinct group of representatives is impossible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

Definitions of Submodularity Independence Matroids Matroid Examples Matroid Rank Partition Matroid System of Distinct Reps

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, $\ell = 6$, and $\mathcal{V} = (V_1, V_2, \ldots, V_6)$ $= (\{e, f, h\}, \{d, e, g\}, \{b, c, e, h\}, \{a, b, h\}, \{a\}, \{a\})$.



- A system of representatives would make sure that there is a representative for each color group. For example,
- The representatives are shown as colors on the left.
- Here, the set of representatives is not distinct. 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., $\mathcal{V} = (V_k : i \in I)$ where $V_i \subseteq V$ for all i). Hence, $|I| = |\mathcal{V}|$.
- A family $(v_i : i \in I)$ with $v_i \in V$ for index set I 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. Lets re-state (and rename) this as a:

Definition 4.9.1 (transversal)

Given a set system $(V, \mathcal{V})$ 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

$$x \in V_{\pi(x)} \text{ for all } x \in T \tag{4.95}$$

- Note that due to it being a bijection, all of I and T are "covered" (so this makes things distinct).