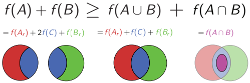
Submodular Functions, Optimization, and Applications to Machine Learning
— Fall Quarter, Lecture 3 —

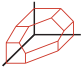
http://www.ee.washington.edu/people/faculty/bilmes/classes/ee563_spring_2018/

Prof. Jeff Bilmes

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

Oct 7th, 2020





Cumulative Outstanding Reading

- Read chapter 1 from Fujishige's book.

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, Matrix Rank, Properties
- L4(10/12):
- L5(10/14):
- L6(10/19):
- 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

Two Equivalent Submodular Definitions

Definition 3.2.1 (submodular concave)

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

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B) \tag{3.7}$$

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

Definition 3.2.2 (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:

$$f(A \cup \{v\}) - f(A) \geq f(B \cup \{v\}) - f(B) \tag{3.8}$$

- The incremental "value", "gain", or "cost" of v decreases (diminishes) as the context in which v is considered grows from A to B.
- Gain notation: Define $f(v|A) \triangleq f(A + v) - f(A)$. Then function f is submodular if $f(v|A) \geq f(v|B)$ for all $A \subseteq B \subseteq V \setminus \{v\}$, $v \in V$.

Two Equivalent Supermodular Definitions

Definition 3.2.1 (supermodular)

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

$$f(A) + f(B) \leq f(A \cup B) + f(A \cap B) \tag{3.7}$$

Definition 3.2.2 (supermodular (improving returns))

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

$$f(A \cup \{v\}) - f(A) \leq f(B \cup \{v\}) - f(B) \tag{3.8}$$

- Incremental "value", "gain", or "cost" of v increases (improves) as the context in which v is considered grows from A to B.
- A function f is submodular iff $-f$ is supermodular.
- If f both submodular and supermodular, then f is said to be **modular**, and $f(A) = c + \sum_{a \in A} \bar{f}(a)$ for some $\bar{f}$ (often $c = 0$).

Monge Matrices

- $m \times n$ matrices $C = [c_{ij}]_{ij}$ are called Monge matrices if they satisfy the Monge property, namely:

$$c_{ij} + c_{rs} \leq c_{is} + c_{rj} \tag{3.1}$$

for all $1 \leq i < r \leq m$ and $1 \leq j < s \leq n$.

Monge Matrices

- $m \times n$ matrices $C = [c_{ij}]_{ij}$ are called Monge matrices if they satisfy the Monge property, namely:

$$c_{ij} + c_{rs} \leq c_{is} + c_{rj} \tag{3.1}$$

for all $1 \leq i < r \leq m$ and $1 \leq j < s \leq n$.

- Lined up indices

$$i < r \tag{3.2}$$

$$j < s \tag{3.3}$$

Monge Matrices

- $m \times n$ matrices $C = [c_{ij}]_{ij}$ are called Monge matrices if they satisfy the Monge property, namely:

$$c_{ij} + c_{rs} \leq c_{is} + c_{rj} \tag{3.1}$$

for all $1 \leq i < r \leq m$ and $1 \leq j < s \leq n$.

- Lined up indices

$$i < r \tag{3.2}$$

$$j < s \tag{3.3}$$

- Equivalently, for all $1 \leq i, r \leq m$, $1 \leq s, j \leq n$,

$$c_{\min(i,r),\min(s,j)} + c_{\max(i,r),\max(s,j)} \leq c_{is} + c_{rj} \tag{3.4}$$

Monge Matrices Visuals

- Consider a non-negative matrix $D = (d_{i,j})$ of order $m \times n$ and form matrix $C = (c_{i,j})$ with $c_{i,j}$th entry, $1 \leq i \leq m$, $1 \leq j \leq n$:

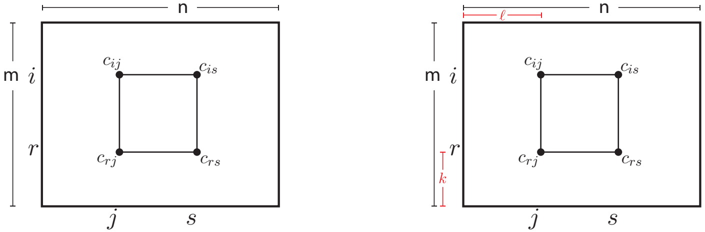
$$c_{ij} = \sum_{k=i}^{m} \sum_{\ell=1}^{j} d_{k,\ell} \tag{3.5}$$

Monge Matrices Visuals

- Consider a non-negative matrix $D = (d_{i,j})$ of order $m \times n$ and form matrix $C = (c_{i,j})$ with $c_{i,j}$th entry, $1 \leq i \leq m$, $1 \leq j \leq n$:

$$c_{ij} = \sum_{k=i}^{m} \sum_{\ell=1}^{j} d_{k,\ell} \tag{3.5}$$

- Consider four elements of the $m \times n$ matrix:

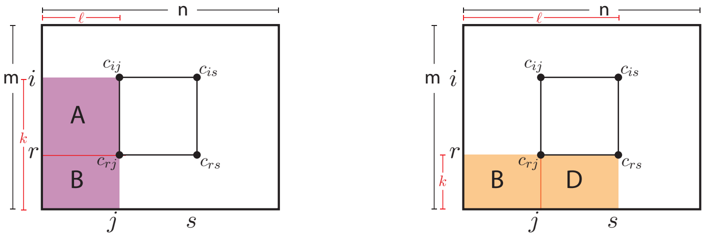


Monge Matrices Visuals

- Consider a non-negative matrix $D = (d_{i,j})$ of order $m \times n$ and form matrix $C = (c_{i,j})$ with $c_{i,j}$th entry, $1 \leq i \leq m$, $1 \leq j \leq n$:

$$c_{ij} = \sum_{k=i}^{m} \sum_{\ell=1}^{j} d_{k,\ell} \tag{3.5}$$

- Consider four elements of the $m \times n$ matrix:



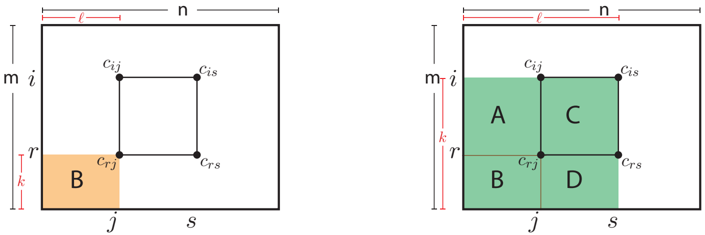
$c_{ij} = A + B$, $c_{rs} = B + D$, $c_{rj} = B$, $c_{is} = A + B + C + D$.

Monge Matrices Visuals

- Consider a non-negative matrix $D = (d_{i,j})$ of order $m \times n$ and form matrix $C = (c_{i,j})$ with $c_{i,j}$th entry, $1 \leq i \leq m$, $1 \leq j \leq n$:

$$c_{ij} = \sum_{k=i}^{m} \sum_{\ell=1}^{j} d_{k,\ell} \tag{3.5}$$

- Consider four elements of the $m \times n$ matrix:



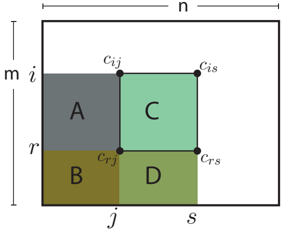
$c_{ij} = A + B$, $c_{rs} = B + D$, $c_{rj} = B$, $c_{is} = A + B + C + D$.

Monge Matrices Visuals

- Consider a non-negative matrix $D = (d_{i,j})$ of order $m \times n$ and form matrix $C = (c_{i,j})$ with $c_{i,j}$th entry, $1 \leq i \leq m$, $1 \leq j \leq n$:

$$c_{ij} = \sum_{k=i}^{m} \sum_{\ell=1}^{j} d_{k,\ell} \tag{3.5}$$

- Consider four elements of the $m \times n$ matrix:



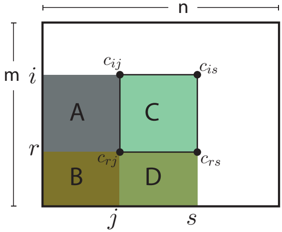
$c_{ij} = A + B$, $c_{rs} = B + D$, $c_{rj} = B$, $c_{is} = A + B + C + D$.

Monge Matrices Visuals

- Consider a non-negative matrix $D = (d_{i,j})$ of order $m \times n$ and form matrix $C = (c_{i,j})$ with $c_{i,j}$th entry, $1 \leq i \leq m$, $1 \leq j \leq n$:

$$c_{ij} = \sum_{k=i}^{m} \sum_{\ell=1}^{j} d_{k,\ell} \tag{3.5}$$

- Consider four elements of the $m \times n$ matrix:



$c_{ij} = A + B$, $c_{rs} = B + D$, $c_{rj} = B$, $c_{is} = A + B + C + D$.

- Then, $c_{ij} + c_{rs} < c_{is} + c_{rj}$.

Monge Matrices, where useful

- Useful for speeding up transportation, dynamic programming, flow, search, lot-sizing and many other problems.

Monge Matrices, where useful

- Useful for speeding up transportation, dynamic programming, flow, search, lot-sizing and many other problems.
- Example, Hitchcock transportation problem: Given $m \times n$ cost matrix $C = [c_{ij}]_{ij}$, a non-negative supply vector $a \in \mathbb{R}_+^m$, a non-negative demand vector $b \in \mathbb{R}_+^n$ with $\sum_{i=1}^m a(i) = \sum_{j=1}^n b_j$, we wish to optimally solve the following linear program:

$$\underset{X \in \mathbb{R}^{m \times n}}{\text{minimize}} \quad \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij} \tag{3.6}$$

$$\text{subject to} \quad \sum_{i=1}^{m} x_{ij} = b_j \;\; \forall j = 1, \ldots, n \tag{3.7}$$

$$\sum_{j=1}^{n} x_{ij} = a_i \;\; \forall i = 1, \ldots, m \tag{3.8}$$

$$x_{i,j} \geq 0 \;\; \forall i, j \tag{3.9}$$

Monge Matrices, Hitchcock transportation

$$
C
$$

Producers, Sources, or Supply

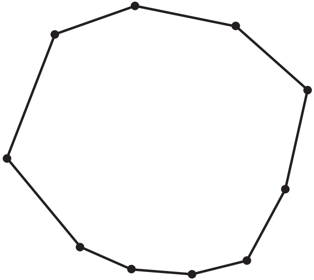
$$
\begin{array}{cc|c|c|c|c|}
\cline{3-6}
a_1 & 2 & 0 & 1 & 3 & 3 \\
\cline{3-6}
a_2 & 1 & 1 & 4 & 7 & 10 \\
\cline{3-6}
a_3 & 5 & 0 & 4 & 9 & 14 \\
\cline{3-6}
\multicolumn{2}{c}{} & \multicolumn{1}{c}{3} & \multicolumn{1}{c}{2} & \multicolumn{1}{c}{1} & \multicolumn{1}{c}{2} \\
\multicolumn{2}{c}{} & \multicolumn{1}{c}{b_1} & \multicolumn{1}{c}{b_2} & \multicolumn{1}{c}{b_3} & \multicolumn{1}{c}{b_4}
\end{array}
$$

Consumers, Sinks, or Demand

- Solving the linear program can be done easily and optimally using the "North-West Corner Rule" (a 2D greedy-like approach starting at top-left and moving down or right) in only $O(m+n)$ if the matrix C is Monge!

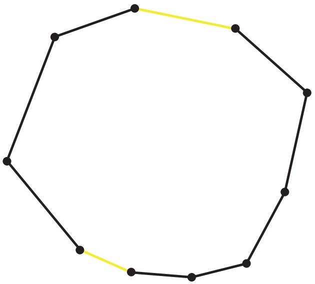
Monge Matrices and Convex Polygons

- Can generate a Monge matrix from a convex polygon - delete two segments, then separately number vertices on each chain. Distances c_{ij} satisfy Monge property (or quadrangle inequality).



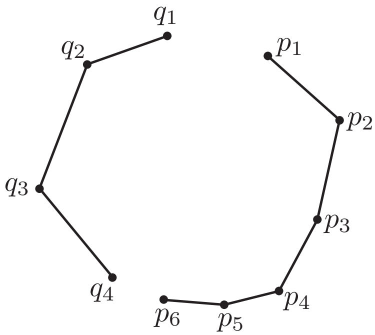
Monge Matrices and Convex Polygons

- Can generate a Monge matrix from a convex polygon - delete two segments, then separately number vertices on each chain. Distances c_{ij} satisfy Monge property (or quadrangle inequality).



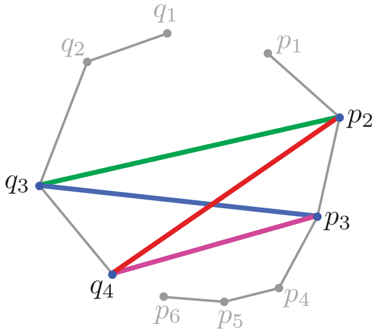
Monge Matrices and Convex Polygons

- Can generate a Monge matrix from a convex polygon - delete two segments, then separately number vertices on each chain. Distances c_{ij} satisfy Monge property (or quadrangle inequality).



Monge Matrices and Convex Polygons

- Can generate a Monge matrix from a convex polygon - delete two segments, then separately number vertices on each chain. Distances c_{ij} satisfy Monge property (or quadrangle inequality).



$$d(q_3, p_2) + d(q_4, p_3) \leq d(q_4, p_2) + d(q_3, p_3) \tag{3.10}$$

Transport unit quantities from locations q_3 and q_4 to locations p_2 and p_3; to minimize total distance traveled, routes from q_3 and q_4 must not intersect.

Monge Matrices and Submodularity

- A submodular function has the form: $f : 2^V \to \mathbb{R}$ which can be seen as $f : \{0,1\}^V \to \mathbb{R}$

Monge Matrices and Submodularity

- A submodular function has the form: $f : 2^V \to \mathbb{R}$ which can be seen as $f : \{0,1\}^V \to \mathbb{R}$
- We can generalize this to $f : \{0,1,\dots,K\}^V \to \mathbb{R}$ for some constant $K \in \mathbb{Z}_+$.

Monge Matrices and Submodularity

- A submodular function has the form: $f : 2^V \to \mathbb{R}$ which can be seen as $f : \{0,1\}^V \to \mathbb{R}$
- We can generalize this to $f : \{0,1,\ldots,K\}^V \to \mathbb{R}$ for some constant $K \in \mathbb{Z}_+$.
- We may define submodularity as: for all $x, y \in \{0,1,\ldots,K\}^V$, we have

$$f(x) + f(y) \geq f(x \vee y) + f(x \wedge y) \tag{3.11}$$

Monge Matrices and Submodularity

- A submodular function has the form: $f : 2^V \to \mathbb{R}$ which can be seen as $f : \{0,1\}^V \to \mathbb{R}$
- We can generalize this to $f : \{0,1,\ldots,K\}^V \to \mathbb{R}$ for some constant $K \in \mathbb{Z}_+$.
- We may define submodularity as: for all $x, y \in \{0,1,\ldots,K\}^V$, we have

$$f(x) + f(y) \geq f(x \vee y) + f(x \wedge y) \tag{3.11}$$

- $x \vee y$ is the (join) element-wise max of each element, that is $(x \vee y)(v) = \max(x(v), y(v))$ for $v \in V$.

Monge Matrices and Submodularity

- A submodular function has the form: $f : 2^V \to \mathbb{R}$ which can be seen as $f : \{0,1\}^V \to \mathbb{R}$
- We can generalize this to $f : \{0,1,\dots,K\}^V \to \mathbb{R}$ for some constant $K \in \mathbb{Z}_+$.
- We may define submodularity as: for all $x, y \in \{0,1,\dots,K\}^V$, we have

$$f(x) + f(y) \geq f(x \vee y) + f(x \wedge y) \tag{3.11}$$

- $x \vee y$ is the (join) element-wise max of each element, that is $(x \vee y)(v) = \max(x(v), y(v))$ for $v \in V$.
- $x \wedge y$ is the (meet) element-wise min of each element, that is, $(x \wedge y)(v) = \min(x(v), y(v))$ for $v \in V$.

Monge Matrices and Submodularity

- A submodular function has the form: $f : 2^V \to \mathbb{R}$ which can be seen as $f : \{0,1\}^V \to \mathbb{R}$
- We can generalize this to $f : \{0,1,\dots,K\}^V \to \mathbb{R}$ for some constant $K \in \mathbb{Z}_+$.
- We may define submodularity as: for all $x, y \in \{0,1,\dots,K\}^V$, we have

$$f(x) + f(y) \geq f(x \vee y) + f(x \wedge y) \tag{3.11}$$

- $x \vee y$ is the (join) element-wise max of each element, that is $(x \vee y)(v) = \max(x(v), y(v))$ for $v \in V$.
- $x \wedge y$ is the (meet) element-wise min of each element, that is, $(x \wedge y)(v) = \min(x(v), y(v))$ for $v \in V$.
- With $K = 1$, then this is the standard definition of submodularity.

Monge Matrices and Submodularity

- A submodular function has the form: $f : 2^V \to \mathbb{R}$ which can be seen as $f : \{0,1\}^V \to \mathbb{R}$
- We can generalize this to $f : \{0,1,\ldots,K\}^V \to \mathbb{R}$ for some constant $K \in \mathbb{Z}_+$.
- We may define submodularity as: for all $x, y \in \{0,1,\ldots,K\}^V$, we have

$$f(x) + f(y) \geq f(x \vee y) + f(x \wedge y) \tag{3.11}$$

- $x \vee y$ is the (join) element-wise max of each element, that is $(x \vee y)(v) = \max(x(v), y(v))$ for $v \in V$.
- $x \wedge y$ is the (meet) element-wise min of each element, that is, $(x \wedge y)(v) = \min(x(v), y(v))$ for $v \in V$.
- With $K = 1$, then this is the standard definition of submodularity.
- With $|V| = 2$, and $K + 1$ the side-dimension of the matrix, we get a Monge property (on square matrices).

Monge Matrices and Submodularity

- A submodular function has the form: $f : 2^V \to \mathbb{R}$ which can be seen as $f : \{0,1\}^V \to \mathbb{R}$
- We can generalize this to $f : \{0,1,\dots,K\}^V \to \mathbb{R}$ for some constant $K \in \mathbb{Z}_+$.
- We may define submodularity as: for all $x, y \in \{0,1,\dots,K\}^V$, we have

$$f(x) + f(y) \geq f(x \vee y) + f(x \wedge y) \tag{3.11}$$

- $x \vee y$ is the (join) element-wise max of each element, that is $(x \vee y)(v) = \max(x(v), y(v))$ for $v \in V$.
- $x \wedge y$ is the (meet) element-wise min of each element, that is, $(x \wedge y)(v) = \min(x(v), y(v))$ for $v \in V$.
- With $K = 1$, then this is the standard definition of submodularity.
- With $|V| = 2$, and $K + 1$ the side-dimension of the matrix, we get a Monge property (on square matrices).
- Non square: $f : \{0,1,\dots,K_1\} \times \{0,1,\dots,K_2\} \to \mathbb{R}$.

Two Equivalent Submodular Definitions

Definition 3.4.1 (submodular concave)

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

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B) \tag{3.7}$$

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

Definition 3.4.2 (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:

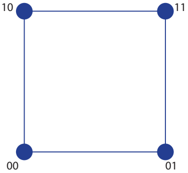
$$f(A \cup \{v\}) - f(A) \geq f(B \cup \{v\}) - f(B) \tag{3.8}$$

- The incremental "value", "gain", or "cost" of v decreases (diminishes) as the context in which v is considered grows from A to B.
- Gain notation: Define $f(v|A) \triangleq f(A + v) - f(A)$. Then function f is submodular if $f(v|A) \geq f(v|B)$ for all $A \subseteq B \subseteq V \setminus \{v\}$, $v \in V$.

The Submodular Square, and Hypercube Vertices

We can test submodularity via values on vertices of hypercube.

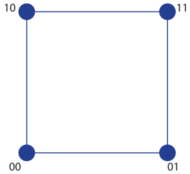
Example: with $|V| = n = 2$, this is easy:



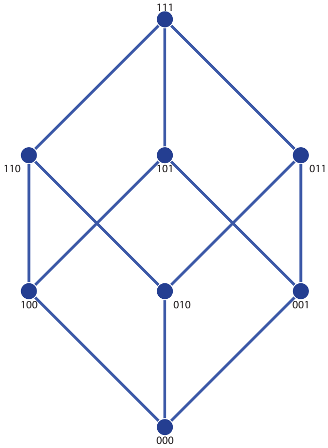
The Submodular Square, and Hypercube Vertices

We can test submodularity via values on vertices of hypercube.

Example: with $|V| = n = 2$, this is easy:



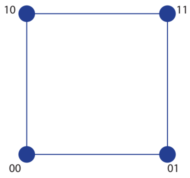
With $|V| = n = 3$, a bit harder.



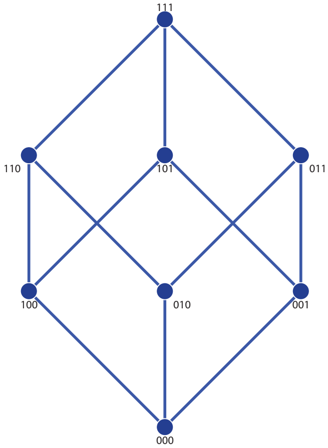
The Submodular Square, and Hypercube Vertices

We can test submodularity via values on vertices of hypercube.

Example: with $|V| = n = 2$, this is easy:



With $|V| = n = 3$, a bit harder.

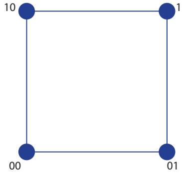


How many inequalities of form $f(A) + f(B) \geq f(A \cup B) + f(A \cap B)$?

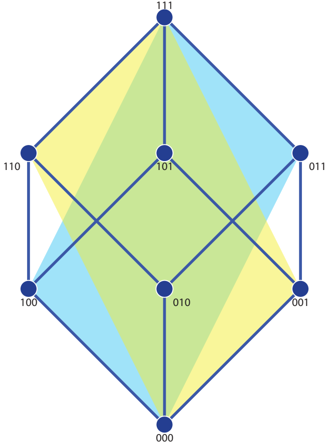
The Submodular Square, and Hypercube Vertices

We can test submodularity via values on vertices of hypercube.

Example: with $|V| = n = 2$, this is easy:



With $|V| = n = 3$, a bit harder.



How many inequalities of form $f(A) + f(B) \geq f(A \cup B) + f(A \cap B)$?

Subadditive Definitions

Definition 3.4.1 (subadditive)

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

$$f(A) + f(B) \geq f(A \cup B) \tag{3.12}$$

This means that the "whole" is less than the sum of the parts.

Two Equivalent Supermodular Definitions

Definition 3.4.1 (supermodular)

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

$$f(A) + f(B) \leq f(A \cup B) + f(A \cap B) \tag{3.7}$$

Definition 3.4.2 (supermodular (improving returns))

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

$$f(A \cup \{v\}) - f(A) \leq f(B \cup \{v\}) - f(B) \tag{3.8}$$

- Incremental "value", "gain", or "cost" of v increases (improves) as the context in which v is considered grows from A to B.
- A function f is submodular iff $-f$ is supermodular.
- If f both submodular and supermodular, then f is said to be **modular**, and $f(A) = c + \sum_{a \in A} \bar{f}(a)$ for some $\bar{f}$ (often $c = 0$).

Superadditive Definitions

Definition 3.4.2 (superadditive)

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

$$f(A) + f(B) \leq f(A \cup B) \tag{3.13}$$

- This means that the "whole" is greater than the sum of the parts.

Superadditive Definitions

Definition 3.4.2 (superadditive)

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

$$f(A) + f(B) \leq f(A \cup B) \tag{3.13}$$

- This means that the "whole" is greater than the sum of the parts.
- In general, submodular and subadditive (and supermodular and superadditive) are different properties.

Superadditive Definitions

Definition 3.4.2 (superadditive)

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

$$f(A) + f(B) \leq f(A \cup B) \tag{3.13}$$

- This means that the "whole" is greater than the sum of the parts.
- In general, submodular and subadditive (and supermodular and superadditive) are different properties.
- Ex: Let $0 < k < |V|$, and consider $f : 2^V \to \mathbb{R}_+$ where:

$$f(A) = \begin{cases} 1 & \text{if } |A| \leq k \\ 0 & \text{else} \end{cases} \tag{3.14}$$

Superadditive Definitions

Definition 3.4.2 (superadditive)

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

$$f(A) + f(B) \leq f(A \cup B) \tag{3.13}$$

- This means that the "whole" is greater than the sum of the parts.
- In general, submodular and subadditive (and supermodular and superadditive) are different properties.
- Ex: Let $0 < k < |V|$, and consider $f : 2^V \to \mathbb{R}_+$ where:

$$f(A) = \begin{cases} 1 & \text{if } |A| \leq k \\ 0 & \text{else} \end{cases} \tag{3.14}$$

- This function is subadditive but not submodular.

Modular Definitions

Definition 3.4.3 (modular)

A function that is both submodular and supermodular is called **modular**

If f is a modular function, than for any $A, B \subseteq V$, we have

$$f(A) + f(B) = f(A \cap B) + f(A \cup B) \tag{3.15}$$

In modular functions, elements do not interact (or cooperate, or compete, or influence each other), and have value based only on singleton values.

Proposition 3.4.4

If f is modular, it may be written as

$$f(A) = f(\emptyset) + \sum_{a \in A} \Big(f(\{a\}) - f(\emptyset) \Big) = c + \sum_{a \in A} f'(a) \tag{3.16}$$

which has only $|V| + 1$ parameters.

Modular Definitions

Proof.

We inductively construct the value for $A = \{a_1, a_2, \ldots, a_k\}$.
For $k = 2$,

$$f(a_1) + f(a_2) = f(a_1, a_2) + f(\emptyset) \tag{3.17}$$
$$\text{implies } f(a_1, a_2) = f(a_1) - f(\emptyset) + f(a_2) - f(\emptyset) + f(\emptyset) \tag{3.18}$$

then for $k = 3$,

$$f(a_1, a_2) + f(a_3) = f(a_1, a_2, a_3) + f(\emptyset) \tag{3.19}$$
$$\text{implies } f(a_1, a_2, a_3) = f(a_1, a_2) - f(\emptyset) + f(a_3) - f(\emptyset) + f(\emptyset) \tag{3.20}$$
$$= f(\emptyset) + \sum_{i=1}^{3} \big(f(a_i) - f(\emptyset)\big) \tag{3.21}$$

and so on . . . □

Complement function

Given a function $f : 2^V \to \mathbb{R}$, we can find a complement function $\bar{f} : 2^V \to \mathbb{R}$ as $\bar{f}(A) = f(V \setminus A)$ for any A.

Proposition 3.4.5

$\bar{f}$ is submodular iff f is submodular.

Proof.

$$\bar{f}(A) + \bar{f}(B) \geq \bar{f}(A \cup B) + \bar{f}(A \cap B) \tag{3.22}$$

follows from

$$f(V \setminus A) + f(V \setminus B) \geq f(V \setminus (A \cup B)) + f(V \setminus (A \cap B)) \tag{3.23}$$

which is true because $V \setminus (A \cup B) = (V \setminus A) \cap (V \setminus B)$ and $V \setminus (A \cap B) = (V \setminus A) \cup (V \setminus B)$ (De Morgan's laws for sets). □

Undirected Graphs

- Let $G = (V, E)$ be a graph with vertices $V = V(G)$ and edges $E = E(G) \subseteq V \times V$.

Undirected Graphs

- Let $G = (V, E)$ be a graph with vertices $V = V(G)$ and edges $E = E(G) \subseteq V \times V$.
- If G is undirected, define

$$E(X, Y) = \{\{x, y\} \in E(G) : x \in X \setminus Y, y \in Y \setminus X\} \tag{3.24}$$

as the edges strictly between X and Y.

Undirected Graphs

- Let $G = (V, E)$ be a graph with vertices $V = V(G)$ and edges $E = E(G) \subseteq V \times V$.
- If G is undirected, define

$$E(X, Y) = \{\{x, y\} \in E(G) : x \in X \setminus Y, y \in Y \setminus X\} \tag{3.24}$$

as the edges strictly between X and Y.
- Nodes define cuts. Define the cut function $\delta(X) = E(X, V \setminus X)$, set of edges with exactly one vertex in X.

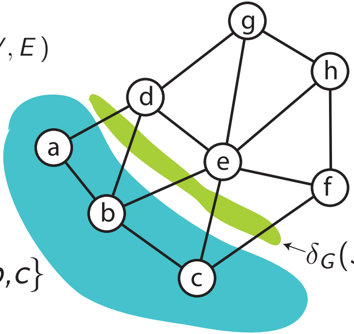
Undirected Graphs

- Let $G = (V, E)$ be a graph with vertices $V = V(G)$ and edges $E = E(G) \subseteq V \times V$.
- If G is undirected, define

$$E(X, Y) = \{\{x, y\} \in E(G) : x \in X \setminus Y, y \in Y \setminus X\} \tag{3.24}$$

as the edges strictly between X and Y.
- Nodes define cuts. Define the cut function $\delta(X) = E(X, V \setminus X)$, set of edges with exactly one vertex in X.

$G = (V, E)$



S=$\{a,b,c\}$

$\delta_G(S) = \{\{u, v\} \in E : u \in S, v \in V \setminus S\}.$
$= \{\{a,d\},\{b,d\},\{b,e\},\{c,e\},\{c,f\}\}$

Directed graphs, and cuts and flows

- If G is directed, define

$$E^+(X,Y) \triangleq \{(x,y) \in E(G) : x \in X \setminus Y, y \in Y \setminus X\} \tag{3.25}$$

as the edges directed strictly from X towards Y.

Directed graphs, and cuts and flows

- If G is directed, define

$$E^+(X,Y) \triangleq \{(x,y) \in E(G) : x \in X \setminus Y, y \in Y \setminus X\} \tag{3.25}$$

as the edges directed strictly from X towards Y.

- Nodes define cuts and flows. Define edges leaving X (out-flow) as

$$\delta^+(X) \triangleq E^+(X, V \setminus X) \tag{3.26}$$

and edges entering X (in-flow) as

$$\delta^-(X) \triangleq E^+(V \setminus X, X) \tag{3.27}$$

Directed graphs, and cuts and flows

- If G is directed, define

$$E^+(X,Y) \triangleq \{(x,y) \in E(G) : x \in X \setminus Y, y \in Y \setminus X\} \tag{3.25}$$

as the edges directed strictly from X towards Y.

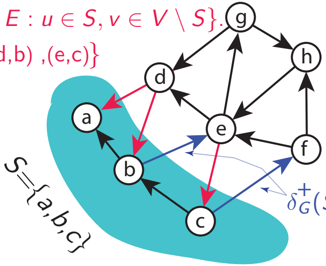
- Nodes define cuts and flows. Define edges leaving X (out-flow) as

$$\delta^+(X) \triangleq E^+(X, V \setminus X) \tag{3.26}$$

and edges entering X (in-flow) as

$$\delta^-(X) \triangleq E^+(V \setminus X, X) \tag{3.27}$$

$\delta^-_G(S) = \{(v,u) \in E : u \in S, v \in V \setminus S\}.$
$= \{$(d,a) ,(d,b) ,(e,c)$\}$



$\delta^+_G(S) = \{(u,v) \in E : u \in S, v \in V \setminus S\}.$
$= \{$(b,e) ,(c,f)$\}$

The Neighbor function in undirected graphs

- Given a set $X \subseteq V$, the neighbor function of X is defined as

$$\Gamma(X) \triangleq \{v \in V(G) \setminus X : E(X, \{v\}) \neq \emptyset\} \tag{3.28}$$

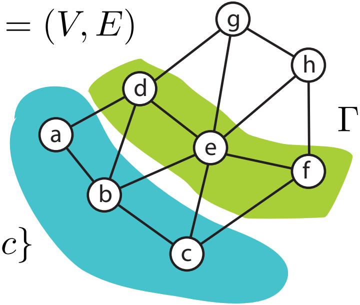
The Neighbor function in undirected graphs

- Given a set $X \subseteq V$, the neighbor function of X is defined as

$$\Gamma(X) \triangleq \{v \in V(G) \setminus X : E(X, \{v\}) \neq \emptyset\} \tag{3.28}$$

- Example:

$G = (V, E)$



$\Gamma(S) = \{d, e, f\}$

$S = \{a, b, c\}$

Directed Cut function: property

Lemma 3.5.1

For a digraph $G = (V, E)$ and any $X, Y \subseteq V$: we have

$$|\delta^+(X)| + |\delta^+(Y)| \\ = |\delta^+(X \cap Y)| + |\delta^+(X \cup Y)| + |E^+(X, Y)| + |E^+(Y, X)| \tag{3.29}$$

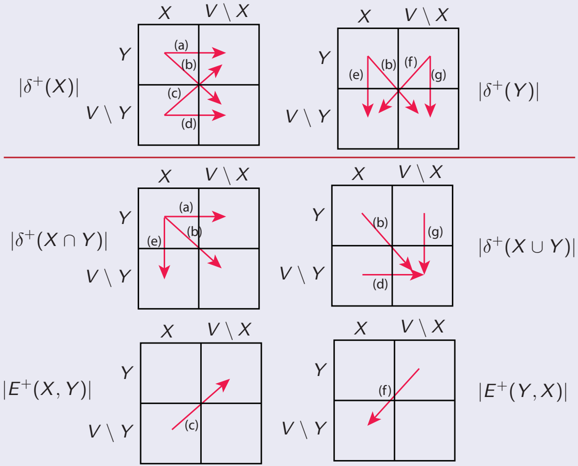
and

$$|\delta^-(X)| + |\delta^-(Y)| \\ = |\delta^-(X \cap Y)| + |\delta^-(X \cup Y)| + |E^-(X, Y)| + |E^-(Y, X)| \tag{3.30}$$

Directed Cut function: proof of property

Proof.

We can prove Eq. (3.29) using a geometric counting argument (proof for $|\delta^-(X)|$ case is similar)



Q:Why is (c) $= |E^+(X, Y)|$?

Directed cut/flow functions: submodular

Lemma 3.5.2

For a digraph $G = (V, E)$ and any $X, Y \subseteq V$: both functions $|\delta^+(X)|$ and $|\delta^-(X)|$ are submodular.

Proof.

$|E^+(X,Y)| \geq 0$ and $|E^-(X,Y)| \geq 0$. □

More generally, in the non-negative weighted edge case, both in-flow and out-flow are submodular on subsets of the vertices.

Undirected Cut/Flow & the Neighbor function: submodular

Lemma 3.5.3

For an undirected graph $G = (V, E)$ and any $X, Y \subseteq V$: we have that both the undirected cut (or flow) function $|\delta(X)|$ and the neighbor function $|\Gamma(X)|$ are submodular. I.e.,

$$|\delta(X)| + |\delta(Y)| = |\delta(X \cap Y)| + |\delta(X \cup Y)| + 2|E(X, Y)| \tag{3.31}$$

and

$$|\Gamma(X)| + |\Gamma(Y)| \geq |\Gamma(X \cap Y)| + |\Gamma(X \cup Y)| \tag{3.32}$$

Proof.

- Eq. (3.31) follows from Eq. (3.29): we replace each undirected edge $\{u, v\}$ with two oppositely-directed directed edges (u, v) and (v, u). Then we use same counting argument.

...

Undirected Cut/Flow & the Neighbor function: submodular

Lemma 3.5.3

For an undirected graph $G = (V, E)$ and any $X, Y \subseteq V$: we have that both the undirected cut (or flow) function $|\delta(X)|$ and the neighbor function $|\Gamma(X)|$ are submodular. I.e.,

$$|\delta(X)| + |\delta(Y)| = |\delta(X \cap Y)| + |\delta(X \cup Y)| + 2|E(X, Y)| \tag{3.31}$$

and

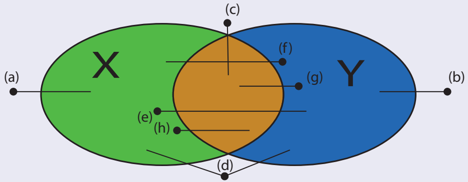
$$|\Gamma(X)| + |\Gamma(Y)| \geq |\Gamma(X \cap Y)| + |\Gamma(X \cup Y)| \tag{3.32}$$

Proof.

- Eq. (3.31) follows from Eq. (3.29): we replace each undirected edge $\{u, v\}$ with two oppositely-directed directed edges (u, v) and (v, u). Then we use same counting argument.
- Eq. (3.32) follows as shown in the following page.

...

cont.



Graphically, we can count and see that

$$\Gamma(X) = (a) + (c) + (f) + (g) + (d) \tag{3.33}$$

$$\Gamma(Y) = (b) + (c) + (e) + (h) + (d) \tag{3.34}$$

$$\Gamma(X \cup Y) = (a) + (b) + (c) + (d) \tag{3.35}$$

$$\Gamma(X \cap Y) = (c) + (g) + (h) \tag{3.36}$$

so

$$|\Gamma(X)| + |\Gamma(Y)| = (a) + (b) + 2(c) + 2(d) + (e) + (f) + (g) + (h)$$
$$\geq (a) + (b) + 2(c) + (d) + (g) + (h) = |\Gamma(X \cup Y)| + |\Gamma(X \cap Y)| \tag{3.37}$$

Undirected Neighbor functions

Therefore, the undirected cut function $|\delta(A)|$ and the neighbor function $|\Gamma(A)|$ of a graph G are both submodular.

Undirected cut/flow is submodular: alternate proof

- Another simple proof shows that $|\delta(X)|$ is submodular.

Undirected cut/flow is submodular: alternate proof

- Another simple proof shows that $|\delta(X)|$ is submodular.
- Define a graph $G_{uv} = (\{u, v\}, \{e\}, w)$ with two nodes u, v and one edge $e = \{u, v\}$ with non-negative weight $w(e) \in \mathbb{R}_+$.

Undirected cut/flow is submodular: alternate proof

- Another simple proof shows that $|\delta(X)|$ is submodular.
- Define a graph $G_{uv} = (\{u, v\}, \{e\}, w)$ with two nodes u, v and one edge $e = \{u, v\}$ with non-negative weight $w(e) \in \mathbb{R}_+$.
- Weighted cut function over those two nodes: $w(\delta_{u,v}(\cdot))$ has valuation:

$$w(\delta_{u,v}(\emptyset)) = w(\delta_{u,v}(\{u, v\})) = 0 \tag{3.38}$$

and

$$w(\delta_{u,v}(\{u\})) = w(\delta_{u,v}(\{v\})) = w \geq 0 \tag{3.39}$$

Undirected cut/flow is submodular: alternate proof

- Another simple proof shows that $|\delta(X)|$ is submodular.
- Define a graph $G_{uv} = (\{u,v\}, \{e\}, w)$ with two nodes u, v and one edge $e = \{u,v\}$ with non-negative weight $w(e) \in \mathbb{R}_+$.
- Weighted cut function over those two nodes: $w(\delta_{u,v}(\cdot))$ has valuation:

$$w(\delta_{u,v}(\emptyset)) = w(\delta_{u,v}(\{u,v\})) = 0 \tag{3.38}$$

and

$$w(\delta_{u,v}(\{u\})) = w(\delta_{u,v}(\{v\})) = w \geq 0 \tag{3.39}$$

- Thus, $w(\delta_{u,v}(\cdot))$ is submodular since $w(e) \geq 0$ and

$$w(\delta_{u,v}(\{u\})) + w(\delta_{u,v}(\{v\})) \geq w(\delta_{u,v}(\{u,v\})) + w(\delta_{u,v}(\emptyset)) \tag{3.40}$$

Undirected cut/flow is submodular: alternate proof

- Another simple proof shows that $|\delta(X)|$ is submodular.
- Define a graph $G_{uv} = (\{u, v\}, \{e\}, w)$ with two nodes u, v and one edge $e = \{u, v\}$ with non-negative weight $w(e) \in \mathbb{R}_+$.
- Weighted cut function over those two nodes: $w(\delta_{u,v}(\cdot))$ has valuation:

$$w(\delta_{u,v}(\emptyset)) = w(\delta_{u,v}(\{u, v\})) = 0 \tag{3.38}$$

and

$$w(\delta_{u,v}(\{u\})) = w(\delta_{u,v}(\{v\})) = w \geq 0 \tag{3.39}$$

- Thus, $w(\delta_{u,v}(\cdot))$ is submodular since $w(e) \geq 0$ and

$$w(\delta_{u,v}(\{u\})) + w(\delta_{u,v}(\{v\})) \geq w(\delta_{u,v}(\{u, v\})) + w(\delta_{u,v}(\emptyset)) \tag{3.40}$$

- General non-negative weighted graph $G = (V, E, w)$, define $w(\delta(\cdot))$:

$$f(X) = w(\delta(X)) = \sum_{(u,v) \in E(G)} w(\delta_{u,v}(X \cap \{u, v\})) \tag{3.41}$$

Undirected cut/flow is submodular: alternate proof

- Another simple proof shows that $|\delta(X)|$ is submodular.
- Define a graph $G_{uv} = (\{u, v\}, \{e\}, w)$ with two nodes u, v and one edge $e = \{u, v\}$ with non-negative weight $w(e) \in \mathbb{R}_+$.
- Weighted cut function over those two nodes: $w(\delta_{u,v}(\cdot))$ has valuation:

$$w(\delta_{u,v}(\emptyset)) = w(\delta_{u,v}(\{u, v\})) = 0 \tag{3.38}$$

and

$$w(\delta_{u,v}(\{u\})) = w(\delta_{u,v}(\{v\})) = w \geq 0 \tag{3.39}$$

- Thus, $w(\delta_{u,v}(\cdot))$ is submodular since $w(e) \geq 0$ and

$$w(\delta_{u,v}(\{u\})) + w(\delta_{u,v}(\{v\})) \geq w(\delta_{u,v}(\{u, v\})) + w(\delta_{u,v}(\emptyset)) \tag{3.40}$$

- General non-negative weighted graph $G = (V, E, w)$, define $w(\delta(\cdot))$:

$$f(X) = w(\delta(X)) = \sum_{(u,v) \in E(G)} w(\delta_{u,v}(X \cap \{u, v\})) \tag{3.41}$$

- This is easily shown to be submodular using properties we will soon see (namely, submodularity closed under summation and restriction).

Other graph functions that are submodular/supermodular

These come from Narayanan's book 1997. Let G be an undirected graph.

- Let $V(X)$ be the vertices adjacent to some edge in $X \subseteq E(G)$, then $|V(X)|$ (the vertex function) is submodular.

Other graph functions that are submodular/supermodular

These come from Narayanan's book 1997. Let G be an undirected graph.

- Let $V(X)$ be the vertices adjacent to some edge in $X \subseteq E(G)$, then $|V(X)|$ (the vertex function) is submodular.
- Let $E(S)$ be the edges with both vertices in $S \subseteq V(G)$. Then $|E(S)|$ (the interior edge function) is supermodular.

Other graph functions that are submodular/supermodular

These come from Narayanan's book 1997. Let G be an undirected graph.

- Let $V(X)$ be the vertices adjacent to some edge in $X \subseteq E(G)$, then $|V(X)|$ (the vertex function) is submodular.
- Let $E(S)$ be the edges with both vertices in $S \subseteq V(G)$. Then $|E(S)|$ (the interior edge function) is supermodular.
- Let $I(S)$ be the edges with at least one vertex in $S \subseteq V(G)$. Then $|I(S)|$ (the incidence function) is submodular.

Other graph functions that are submodular/supermodular

These come from Narayanan's book 1997. Let G be an undirected graph.

- Let $V(X)$ be the vertices adjacent to some edge in $X \subseteq E(G)$, then $|V(X)|$ (the vertex function) is submodular.
- Let $E(S)$ be the edges with both vertices in $S \subseteq V(G)$. Then $|E(S)|$ (the interior edge function) is supermodular.
- Let $I(S)$ be the edges with at least one vertex in $S \subseteq V(G)$. Then $|I(S)|$ (the incidence function) is submodular.
- Recall $|\delta(S)|$, is the number of edges with exactly one vertex in $S \subseteq V(G)$ is submodular (cut function). Thus, we have $I(S) = E(S) \cup \delta(S)$ and $E(S) \cap \delta(S) = \emptyset$, and thus that $|I(S)| = |E(S)| + |\delta(S)|$.

Other graph functions that are submodular/supermodular

These come from Narayanan's book 1997. Let G be an undirected graph.

- Let $V(X)$ be the vertices adjacent to some edge in $X \subseteq E(G)$, then $|V(X)|$ (the vertex function) is submodular.
- Let $E(S)$ be the edges with both vertices in $S \subseteq V(G)$. Then $|E(S)|$ (the interior edge function) is supermodular.
- Let $I(S)$ be the edges with at least one vertex in $S \subseteq V(G)$. Then $|I(S)|$ (the incidence function) is submodular.
- Recall $|\delta(S)|$, is the number of edges with exactly one vertex in $S \subseteq V(G)$ is submodular (cut function). Thus, we have $I(S) = E(S) \cup \delta(S)$ and $E(S) \cap \delta(S) = \emptyset$, and thus that $|I(S)| = |E(S)| + |\delta(S)|$. So we can get a submodular function by summing a submodular and a supermodular function.

Other graph functions that are submodular/supermodular

These come from Narayanan's book 1997. Let G be an undirected graph.

- Let $V(X)$ be the vertices adjacent to some edge in $X \subseteq E(G)$, then $|V(X)|$ (the vertex function) is submodular.
- Let $E(S)$ be the edges with both vertices in $S \subseteq V(G)$. Then $|E(S)|$ (the interior edge function) is supermodular.
- Let $I(S)$ be the edges with at least one vertex in $S \subseteq V(G)$. Then $|I(S)|$ (the incidence function) is submodular.
- Recall $|\delta(S)|$, is the number of edges with exactly one vertex in $S \subseteq V(G)$ is submodular (cut function). Thus, we have $I(S) = E(S) \cup \delta(S)$ and $E(S) \cap \delta(S) = \emptyset$, and thus that $|I(S)| = |E(S)| + |\delta(S)|$. So we can get a submodular function by summing a submodular and a supermodular function. If you had to guess, is this always the case?

Other graph functions that are submodular/supermodular

These come from Narayanan's book 1997. Let G be an undirected graph.

- Let $V(X)$ be the vertices adjacent to some edge in $X \subseteq E(G)$, then $|V(X)|$ (the vertex function) is submodular.
- Let $E(S)$ be the edges with both vertices in $S \subseteq V(G)$. Then $|E(S)|$ (the interior edge function) is supermodular.
- Let $I(S)$ be the edges with at least one vertex in $S \subseteq V(G)$. Then $|I(S)|$ (the incidence function) is submodular.
- Recall $|\delta(S)|$, is the number of edges with exactly one vertex in $S \subseteq V(G)$ is submodular (cut function). Thus, we have $I(S) = E(S) \cup \delta(S)$ and $E(S) \cap \delta(S) = \emptyset$, and thus that $|I(S)| = |E(S)| + |\delta(S)|$. So we can get a submodular function by summing a submodular and a supermodular function. If you had to guess, is this always the case?
- Consider $f(A) = |\delta^+(A)| - |\delta^+(V \setminus A)|$. Guess, submodular, supermodular, modular, or neither? Exercise: determine which one and prove it.

Number of connected components in a graph via edges

- Recall, $f : 2^V \to \mathbb{R}$ is submodular, then so is $\bar{f} : 2^V \to \mathbb{R}$ defined as $\bar{f}(S) = f(V \setminus S)$.

Number of connected components in a graph via edges

- Recall, $f : 2^V \to \mathbb{R}$ is submodular, then so is $\bar{f} : 2^V \to \mathbb{R}$ defined as $\bar{f}(S) = f(V \setminus S)$.
- Hence, if $g : 2^V \to \mathbb{R}$ is supermodular, then so is $\bar{g} : 2^V \to \mathbb{R}$ defined as $\bar{g}(S) = g(V \setminus S)$.

Number of connected components in a graph via edges

- Recall, $f : 2^V \to \mathbb{R}$ is submodular, then so is $\bar{f} : 2^V \to \mathbb{R}$ defined as $\bar{f}(S) = f(V \setminus S)$.
- Hence, if $g : 2^V \to \mathbb{R}$ is supermodular, then so is $\bar{g} : 2^V \to \mathbb{R}$ defined as $\bar{g}(S) = g(V \setminus S)$.
- Given a graph $G = (V, E)$, for each $A \subseteq E(G)$, let $c(A)$ denote the number of connected components of the (spanning) subgraph $(V(G), A)$, with $c : 2^E \to \mathbb{R}_+$. Thus, $c(\emptyset) = |V|$, and $c(E) \geq 1$.

Number of connected components in a graph via edges

- Recall, $f : 2^V \to \mathbb{R}$ is submodular, then so is $\bar{f} : 2^V \to \mathbb{R}$ defined as $\bar{f}(S) = f(V \setminus S)$.
- Hence, if $g : 2^V \to \mathbb{R}$ is supermodular, then so is $\bar{g} : 2^V \to \mathbb{R}$ defined as $\bar{g}(S) = g(V \setminus S)$.
- Given a graph $G = (V, E)$, for each $A \subseteq E(G)$, let $c(A)$ denote the number of connected components of the (spanning) subgraph $(V(G), A)$, with $c : 2^E \to \mathbb{R}_+$. Thus, $c(\emptyset) = |V|$, and $c(E) \geq 1$.
- $c(A)$ is monotone non-increasing, $c(A + a) - c(A) \leq 0$.

Number of connected components in a graph via edges

- Recall, $f : 2^V \to \mathbb{R}$ is submodular, then so is $\bar{f} : 2^V \to \mathbb{R}$ defined as $\bar{f}(S) = f(V \setminus S)$.
- Hence, if $g : 2^V \to \mathbb{R}$ is supermodular, then so is $\bar{g} : 2^V \to \mathbb{R}$ defined as $\bar{g}(S) = g(V \setminus S)$.
- Given a graph $G = (V, E)$, for each $A \subseteq E(G)$, let $c(A)$ denote the number of connected components of the (spanning) subgraph $(V(G), A)$, with $c : 2^E \to \mathbb{R}_+$. Thus, $c(\emptyset) = |V|$, and $c(E) \geq 1$.
- $c(A)$ is monotone non-increasing, $c(A + a) - c(A) \leq 0$.
- Then $c(A)$ is supermodular, i.e.,

$$c(A + a) - c(A) \leq c(B + a) - c(B) \tag{3.42}$$

with $A \subseteq B \subseteq E \setminus \{a\}$.

Number of connected components in a graph via edges

- Recall, $f : 2^V \to \mathbb{R}$ is submodular, then so is $\bar{f} : 2^V \to \mathbb{R}$ defined as $\bar{f}(S) = f(V \setminus S)$.
- Hence, if $g : 2^V \to \mathbb{R}$ is supermodular, then so is $\bar{g} : 2^V \to \mathbb{R}$ defined as $\bar{g}(S) = g(V \setminus S)$.
- Given a graph $G = (V, E)$, for each $A \subseteq E(G)$, let $c(A)$ denote the number of connected components of the (spanning) subgraph $(V(G), A)$, with $c : 2^E \to \mathbb{R}_+$. Thus, $c(\emptyset) = |V|$, and $c(E) \geq 1$.
- $c(A)$ is monotone non-increasing, $c(A + a) - c(A) \leq 0$.
- Then $c(A)$ is supermodular, i.e.,

$$c(A + a) - c(A) \leq c(B + a) - c(B) \tag{3.42}$$

  with $A \subseteq B \subseteq E \setminus \{a\}$.
- Intuition: an edge is "more" (no less) able to bridge separate components (and reduce the number of connected components) when edge is added in a smaller context than when added in a larger context.

Number of connected components in a graph via edges

- Recall, $f : 2^V \to \mathbb{R}$ is submodular, then so is $\bar{f} : 2^V \to \mathbb{R}$ defined as $\bar{f}(S) = f(V \setminus S)$.
- Hence, if $g : 2^V \to \mathbb{R}$ is supermodular, then so is $\bar{g} : 2^V \to \mathbb{R}$ defined as $\bar{g}(S) = g(V \setminus S)$.
- Given a graph $G = (V, E)$, for each $A \subseteq E(G)$, let $c(A)$ denote the number of connected components of the (spanning) subgraph $(V(G), A)$, with $c : 2^E \to \mathbb{R}_+$. Thus, $c(\emptyset) = |V|$, and $c(E) \geq 1$.
- $c(A)$ is monotone non-increasing, $c(A + a) - c(A) \leq 0$.
- Then $c(A)$ is supermodular, i.e.,

$$c(A + a) - c(A) \leq c(B + a) - c(B) \tag{3.42}$$

with $A \subseteq B \subseteq E \setminus \{a\}$.
- Intuition: an edge is "more" (no less) able to bridge separate components (and reduce the number of connected components) when edge is added in a smaller context than when added in a larger context.
- $\bar{c}(A) = c(E \setminus A)$ is number of connected components in G when we remove A; supermodular monotone non-decreasing but not normalized.

Graph Strength

- So $\bar{c}(A) = c(E \setminus A)$, the number of connected components in G when we remove A, is supermodular.

Graph Strength

- So $\bar{c}(A) = c(E \setminus A)$, the number of connected components in G when we remove A, is supermodular.
- Maximizing $\bar{c}(A)$ would be a goal for a network attacker — many connected components means that many points in the network have lost connectivity to many other points (unprotected network).

Graph Strength

- So $\bar{c}(A) = c(E \setminus A)$, the number of connected components in G when we remove A, is supermodular.
- Maximizing $\bar{c}(A)$ would be a goal for a network attacker — many connected components means that many points in the network have lost connectivity to many other points (unprotected network).
- If we can remove a small set A and shatter the graph into many connected components, then the graph is weak.

Graph Strength

- So $\bar{c}(A) = c(E \setminus A)$, the number of connected components in G when we remove A, is supermodular.
- Maximizing $\bar{c}(A)$ would be a goal for a network attacker — many connected components means that many points in the network have lost connectivity to many other points (unprotected network).
- If we can remove a small set A and shatter the graph into many connected components, then the graph is weak.
- An attacker wishes to choose a small number of edges (since it is cheap) to shatter the graph into as many components as possible.

Graph Strength

- So $\bar{c}(A) = c(E \setminus A)$, the number of connected components in G when we remove A, is supermodular.
- Maximizing $\bar{c}(A)$ would be a goal for a network attacker — many connected components means that many points in the network have lost connectivity to many other points (unprotected network).
- If we can remove a small set A and shatter the graph into many connected components, then the graph is weak.
- An attacker wishes to choose a small number of edges (since it is cheap) to shatter the graph into as many components as possible.
- Let $G = (V, E, w)$ with $w : E \to \mathbb{R}+$ be a weighted graph with non-negative weights.

Graph Strength

- So $\bar{c}(A) = c(E \setminus A)$, the number of connected components in G when we remove A, is supermodular.
- Maximizing $\bar{c}(A)$ would be a goal for a network attacker — many connected components means that many points in the network have lost connectivity to many other points (unprotected network).
- If we can remove a small set A and shatter the graph into many connected components, then the graph is weak.
- An attacker wishes to choose a small number of edges (since it is cheap) to shatter the graph into as many components as possible.
- Let $G = (V, E, w)$ with $w : E \to \mathbb{R}+$ be a weighted graph with non-negative weights.
- For $(u, v) = e \in E$, let $w(e)$ be a measure of the strength of the connection between vertices u and v (strength meaning the difficulty of cutting the edge e).

Graph Strength

- Then $w(A)$ for $A \subseteq E$ is a modular function

$$w(A) = \sum_{e \in A} w_e \tag{3.43}$$

so that $w(E(G[S]))$ is the "internal strength" of the vertex set S. *Notation: S is a set of nodes, $G[S]$ is the vertex-induced subgraph of G induced by vertices S, $E(G[S])$ are the edges contained within this induced subgraph, and $w(E(G[S]))$ is the weight of these edges. $w(E(G[S])) = \sum_{i,j \in S} w(i,j)$.*

Graph Strength

- Then $w(A)$ for $A \subseteq E$ is a modular function

$$w(A) = \sum_{e \in A} w_e \tag{3.43}$$

so that $w(E(G[S]))$ is the "internal strength" of the vertex set S.

- Suppose removing A shatters G into a graph with $\bar{c}(A) > 1$ components —

Graph Strength

- Then $w(A)$ for $A \subseteq E$ is a modular function
$$w(A) = \sum_{e \in A} w_e \tag{3.43}$$
so that $w(E(G[S]))$ is the "internal strength" of the vertex set S.
- Suppose removing A shatters G into a graph with $\bar{c}(A) > 1$ components — then $w(A)/(\bar{c}(A) - 1)$ is like the "effort per achieved/additional component" for a network attacker.

Graph Strength

- Then $w(A)$ for $A \subseteq E$ is a modular function

$$w(A) = \sum_{e \in A} w_e \tag{3.43}$$

so that $w(E(G[S]))$ is the "internal strength" of the vertex set S.

- Suppose removing A shatters G into a graph with $\bar{c}(A) > 1$ components — then $w(A)/(\bar{c}(A) - 1)$ is like the "effort per achieved/additional component" for a network attacker.
- A form of graph strength can then be defined as the following:

$$\textit{strength}(G, w) = \min_{A \subseteq E(G) : \bar{c}(A) > 1} \frac{w(A)}{\bar{c}(A) - 1} \tag{3.44}$$

Graph Strength

- Then $w(A)$ for $A \subseteq E$ is a modular function

$$w(A) = \sum_{e \in A} w_e \tag{3.43}$$

so that $w(E(G[S]))$ is the "internal strength" of the vertex set S.

- Suppose removing A shatters G into a graph with $\bar{c}(A) > 1$ components — then $w(A)/(\bar{c}(A) - 1)$ is like the "effort per achieved/additional component" for a network attacker.
- A form of graph strength can then be defined as the following:

$$\textit{strength}(G, w) = \min_{A \subseteq E(G) : \bar{c}(A) > 1} \frac{w(A)}{\bar{c}(A) - 1} \tag{3.44}$$

- Graph strength is like the minimum effort per component. An attacker would use the argument of the min to choose which edges to attack. A network designer would maximize, over G and/or w, the graph strength, $\textit{strength}(G, w)$.

Graph Strength

- Then $w(A)$ for $A \subseteq E$ is a modular function

$$w(A) = \sum_{e \in A} w_e \tag{3.43}$$

so that $w(E(G[S]))$ is the "internal strength" of the vertex set S.

- Suppose removing A shatters G into a graph with $\bar{c}(A) > 1$ components — then $w(A)/(\bar{c}(A) - 1)$ is like the "effort per achieved/additional component" for a network attacker.
- A form of graph strength can then be defined as the following:

$$\textit{strength}(G, w) = \min_{A \subseteq E(G) : \bar{c}(A) > 1} \frac{w(A)}{\bar{c}(A) - 1} \tag{3.44}$$

- Graph strength is like the minimum effort per component. An attacker would use the argument of the min to choose which edges to attack. A network designer would maximize, over G and/or w, the graph strength, *strength*(G, w).
- Since submodularity, problems have strongly-poly-time solutions.

Submodularity, Quadratic Structures, and Cuts

Lemma 3.5.4

Let $\mathbf{M} \in \mathbb{R}^{n \times n}$ be a symmetric matrix and $m \in \mathbb{R}^n$ be a vector. Then $f : 2^V \to \mathbb{R}$ defined as

$$f(X) = m^\intercal \mathbf{1}_X + \frac{1}{2}\mathbf{1}_X^\intercal \mathbf{M} \mathbf{1}_X \tag{3.45}$$

is submodular iff the off-diagonal elements of M are non-positive.

Proof.

...

Submodularity, Quadratic Structures, and Cuts

Lemma 3.5.4

Let $\mathbf{M} \in \mathbb{R}^{n \times n}$ be a symmetric matrix and $m \in \mathbb{R}^n$ be a vector. Then $f : 2^V \to \mathbb{R}$ defined as

$$f(X) = m^\intercal \mathbf{1}_X + \frac{1}{2}\mathbf{1}_X^\intercal \mathbf{M} \mathbf{1}_X \tag{3.45}$$

is submodular iff the off-diagonal elements of M are non-positive.

Proof.

- Given a complete graph $G = (V, E)$, recall that $E(X)$ is the edge set with both vertices in $X \subseteq V(G)$, and that $|E(X)|$ is supermodular.

...

Submodularity, Quadratic Structures, and Cuts

Lemma 3.5.4

Let $\mathbf{M} \in \mathbb{R}^{n \times n}$ be a symmetric matrix and $m \in \mathbb{R}^n$ be a vector. Then $f : 2^V \to \mathbb{R}$ defined as

$$f(X) = m^\intercal \mathbf{1}_X + \frac{1}{2} \mathbf{1}_X^\intercal \mathbf{M} \mathbf{1}_X \tag{3.45}$$

is submodular iff the off-diagonal elements of M are non-positive.

Proof.

- Given a complete graph $G = (V, E)$, recall that $E(X)$ is the edge set with both vertices in $X \subseteq V(G)$, and that $|E(X)|$ is supermodular.
- Non-negative modular weights $w^+ : E \to \mathbb{R}_+$, $w(E(X))$ is also supermodular, so $-w(E(X))$ is submodular.

...

Submodularity, Quadratic Structures, and Cuts

Lemma 3.5.4

Let $\mathbf{M} \in \mathbb{R}^{n \times n}$ be a symmetric matrix and $m \in \mathbb{R}^n$ be a vector. Then $f : 2^V \to \mathbb{R}$ defined as

$$f(X) = m^\intercal \mathbf{1}_X + \frac{1}{2} \mathbf{1}_X^\intercal \mathbf{M} \mathbf{1}_X \tag{3.45}$$

is submodular iff the off-diagonal elements of M are non-positive.

Proof.

- Given a complete graph $G = (V, E)$, recall that $E(X)$ is the edge set with both vertices in $X \subseteq V(G)$, and that $|E(X)|$ is supermodular.
- Non-negative modular weights $w^+ : E \to \mathbb{R}_+$, $w(E(X))$ is also supermodular, so $-w(E(X))$ is submodular.
- f is a modular function $m^\intercal \mathbf{1}_A = m(A)$ added to a weighted submodular function, hence f is submodular.

...

Submodularity, Quadratic Structures, and Cuts

Proof of Lemma 3.5.4 cont.

- Conversely, suppose f is submodular.

$\square$

Submodularity, Quadratic Structures, and Cuts

Proof of Lemma 3.5.4 cont.

- Conversely, suppose f is submodular.
- Then $\forall u, v \in V$, $f(\{u\}) + f(\{v\}) \geq f(\{u,v\}) + f(\emptyset)$ and $f(\emptyset) = 0$.

$\square$

Submodularity, Quadratic Structures, and Cuts

Proof of Lemma 3.5.4 cont.

- Conversely, suppose f is submodular.
- Then $\forall u, v \in V$, $f(\{u\}) + f(\{v\}) \geq f(\{u, v\}) + f(\emptyset)$ and $f(\emptyset) = 0$.
- This requires:

$$
\begin{aligned}
0 &\leq f(\{u\}) + f(\{v\}) - f(\{u, v\}) && (3.46)\\
&= m(u) + \frac{1}{2}M_{u,u} + m(v) + \frac{1}{2}M_{v,v} && (3.47)\\
&\quad - \left(m(u) + m(v) + \frac{1}{2}M_{u,u} + M_{u,v} + \frac{1}{2}M_{v,v}\right) && (3.48)\\
&= -M_{u,v} && (3.49)
\end{aligned}
$$

So that $\forall u, v \in V$, $M_{u,v} \leq 0$. □

Set Cover and Maximum Coverage

just Special cases of Submodular Optimization

- We are given a finite set U of m elements and a set of subsets $\mathcal{U} = \{U_1, U_2, \dots, U_n\}$ of n subsets of U, so that $U_i \subseteq U$ and $\bigcup_i U_i = U$.

Set Cover and Maximum Coverage

just Special cases of Submodular Optimization

- We are given a finite set U of m elements and a set of subsets $\mathcal{U} = \{U_1, U_2, \dots, U_n\}$ of n subsets of U, so that $U_i \subseteq U$ and $\bigcup_i U_i = U$.
- The goal of minimum set cover is to choose the smallest subset $A \subseteq [n] \triangleq \{1, \dots, n\}$ such that $\bigcup_{a \in A} U_a = U$.

Set Cover and Maximum Coverage

just Special cases of Submodular Optimization

- We are given a finite set U of m elements and a set of subsets $\mathcal{U} = \{U_1, U_2, \ldots, U_n\}$ of n subsets of U, so that $U_i \subseteq U$ and $\bigcup_i U_i = U$.
- The goal of minimum set cover is to choose the smallest subset $A \subseteq [n] \triangleq \{1, \ldots, n\}$ such that $\bigcup_{a \in A} U_a = U$.
- Maximum k cover: The goal in maximum coverage is, given an integer $k \leq n$, select k subsets, say $\{a_1, a_2, \ldots, a_k\}$ with $a_i \in [n]$ such that $|\bigcup_{i=1}^{k} U_{a_i}|$ is maximized.

Set Cover and Maximum Coverage

just Special cases of Submodular Optimization

- We are given a finite set U of m elements and a set of subsets $\mathcal{U} = \{U_1, U_2, \dots, U_n\}$ of n subsets of U, so that $U_i \subseteq U$ and $\bigcup_i U_i = U$.
- The goal of minimum set cover is to choose the smallest subset $A \subseteq [n] \triangleq \{1, \dots, n\}$ such that $\bigcup_{a \in A} U_a = U$.
- Maximum k cover: The goal in maximum coverage is, given an integer $k \leq n$, select k subsets, say $\{a_1, a_2, \dots, a_k\}$ with $a_i \in [n]$ such that $|\bigcup_{i=1}^k U_{a_i}|$ is maximized.
- $f : 2^{[n]} \to \mathbb{Z}_+$ where for $A \subseteq [n]$, $f(A) = |\bigcup_{a \in A} U_a|$ is the set cover function and is submodular.

Set Cover and Maximum Coverage

just Special cases of Submodular Optimization

- We are given a finite set U of m elements and a set of subsets $\mathcal{U} = \{U_1, U_2, \dots, U_n\}$ of n subsets of U, so that $U_i \subseteq U$ and $\bigcup_i U_i = U$.
- The goal of minimum set cover is to choose the smallest subset $A \subseteq [n] \triangleq \{1, \dots, n\}$ such that $\bigcup_{a \in A} U_a = U$.
- Maximum k cover: The goal in maximum coverage is, given an integer $k \leq n$, select k subsets, say $\{a_1, a_2, \dots, a_k\}$ with $a_i \in [n]$ such that $|\bigcup_{i=1}^k U_{a_i}|$ is maximized.
- $f : 2^{[n]} \to \mathbb{Z}_+$ where for $A \subseteq [n]$, $f(A) = |\bigcup_{a \in A} U_a|$ is the set cover function and is submodular.
- Weighted set cover: $f(A) = w(\bigcup_{a \in A} U_a)$ where $w : U \to \mathbb{R}_+$.

Set Cover and Maximum Coverage

just Special cases of Submodular Optimization

- We are given a finite set U of m elements and a set of subsets $\mathcal{U} = \{U_1, U_2, \dots, U_n\}$ of n subsets of U, so that $U_i \subseteq U$ and $\bigcup_i U_i = U$.
- The goal of minimum set cover is to choose the smallest subset $A \subseteq [n] \triangleq \{1, \dots, n\}$ such that $\bigcup_{a \in A} U_a = U$.
- Maximum k cover: The goal in maximum coverage is, given an integer $k \leq n$, select k subsets, say $\{a_1, a_2, \dots, a_k\}$ with $a_i \in [n]$ such that $|\bigcup_{i=1}^k U_{a_i}|$ is maximized.
- $f : 2^{[n]} \to \mathbb{Z}_+$ where for $A \subseteq [n]$, $f(A) = |\bigcup_{a \in A} U_a|$ is the set cover function and is submodular.
- Weighted set cover: $f(A) = w(\bigcup_{a \in A} U_a)$ where $w : U \to \mathbb{R}_+$.
- Both Set cover and maximum coverage are well known to be NP-hard, but have a fast greedy approximation algorithm, and hence are instances of submodular optimization.

Vertex and Edge Covers

Also instances of submodular optimization

Definition 3.5.5 (vertex cover)

A *vertex cover* (a "vertex-based cover of edges") in graph $G = (V, E)$ is a set $S \subseteq V(G)$ of vertices such that every edge in G is incident to at least one vertex in S.

- Let $I(S)$ be the number of edges incident to vertex set S. Then we wish to find the smallest set $S \subseteq V$ subject to $I(S) = |E|$.

Definition 3.5.6 (edge cover)

A *edge cover* (an "edge-based cover of vertices") in graph $G = (V, E)$ is a set $F \subseteq E(G)$ of edges such that every vertex in G is incident to at least one edge in F.

- Let $|V|(F)$ be the number of vertices incident to edge set F. Then we wish to find the smallest set $F \subseteq E$ subject to $|V|(F) = |V|$.

Graph Cut Problems

Also submodular optimization

- Minimum cut: Given a graph $G = (V, E)$, find a set of vertices $S \subseteq V$ that minimize the cut (set of edges) between S and $V \setminus S$.

Graph Cut Problems

Also submodular optimization

- Minimum cut: Given a graph $G = (V, E)$, find a set of vertices $S \subseteq V$ that minimize the cut (set of edges) between S and $V \setminus S$.
- Maximum cut: Given a graph $G = (V, E)$, find a set of vertices $S \subseteq V$ that maximize the cut (set of edges) between S and $V \setminus S$.

Graph Cut Problems

Also submodular optimization

- Minimum cut: Given a graph $G = (V, E)$, find a set of vertices $S \subseteq V$ that minimize the cut (set of edges) between S and $V \setminus S$.
- Maximum cut: Given a graph $G = (V, E)$, find a set of vertices $S \subseteq V$ that maximize the cut (set of edges) between S and $V \setminus S$.
- Let $\delta : 2^V \to \mathbb{R}_+$ be the cut function, namely for any given set of nodes $X \subseteq V$, $|\delta(X)|$ measures the number of edges between nodes X and $V \setminus X$ — i.e., $\delta(x) = E(X, V \setminus X)$.

Graph Cut Problems

Also submodular optimization

- Minimum cut: Given a graph $G = (V, E)$, find a set of vertices $S \subseteq V$ that minimize the cut (set of edges) between S and $V \setminus S$.
- Maximum cut: Given a graph $G = (V, E)$, find a set of vertices $S \subseteq V$ that maximize the cut (set of edges) between S and $V \setminus S$.
- Let $\delta : 2^V \to \mathbb{R}_+$ be the cut function, namely for any given set of nodes $X \subseteq V$, $|\delta(X)|$ measures the number of edges between nodes X and $V \setminus X$ — i.e., $\delta(x) = E(X, V \setminus X)$.
- Weighted versions, where rather than count, we sum the (non-negative) weights of the edges of a cut, $f(X) = w(\delta(X))$.

Graph Cut Problems

Also submodular optimization

- Minimum cut: Given a graph $G = (V, E)$, find a set of vertices $S \subseteq V$ that minimize the cut (set of edges) between S and $V \setminus S$.
- Maximum cut: Given a graph $G = (V, E)$, find a set of vertices $S \subseteq V$ that maximize the cut (set of edges) between S and $V \setminus S$.
- Let $\delta : 2^V \to \mathbb{R}_+$ be the cut function, namely for any given set of nodes $X \subseteq V$, $|\delta(X)|$ measures the number of edges between nodes X and $V \setminus X$ — i.e., $\delta(x) = E(X, V \setminus X)$.
- Weighted versions, where rather than count, we sum the (non-negative) weights of the edges of a cut, $f(X) = w(\delta(X))$.
- Hence, Minimum cut and Maximum cut are also special cases of submodular optimization.

Matrix Rank functions

- Let V, with $|V| = m$ be an index set of a set of vectors in $\mathbb{R}^n$ for some n (unrelated to m). Thus, $\forall v \in V$, $\exists x_v \in \mathbb{R}^n$.

Matrix Rank functions

- Let V, with $|V| = m$ be an index set of a set of vectors in $\mathbb{R}^n$ for some n (unrelated to m). Thus, $\forall v \in V$, $\exists x_v \in \mathbb{R}^n$.
- For a given set $\{v, v_1, v_2, \dots, v_k\}$, it might or might not be possible to find $(\alpha_i)_i$ such that:

$$x_v = \sum_{i=1}^{k} \alpha_i x_{v_i} \tag{3.50}$$

If not, then x_v is linearly independent of $x_{v_1}, \dots, x_{v_k}$.

Matrix Rank functions

- Let V, with $|V| = m$ be an index set of a set of vectors in $\mathbb{R}^n$ for some n (unrelated to m). Thus, $\forall v \in V$, $\exists x_v \in \mathbb{R}^n$.
- For a given set $\{v, v_1, v_2, \ldots, v_k\}$, it might or might not be possible to find $(\alpha_i)_i$ such that:

$$x_v = \sum_{i=1}^{k} \alpha_i x_{v_i} \tag{3.50}$$

  If not, then x_v is linearly independent of $x_{v_1}, \ldots, x_{v_k}$.
- Let $r(S)$ for $S \subseteq V$ be the rank of the set of vectors S. Then $r(\cdot)$ is a submodular function, and in fact is called a matric matroid rank function.

Example: Rank function of a matrix

- Given $n \times m$ matrix $\mathbf{X} = (x_1, x_2, \ldots, x_m)$ with $x_i \in \mathbb{R}^n$ for all i. There are m length-n column vectors $\{x_i\}_i$

Skip matrix rank example

Example: Rank function of a matrix

- Given $n \times m$ matrix $\mathbf{X} = (x_1, x_2, \dots, x_m)$ with $x_i \in \mathbb{R}^n$ for all i. There are m length-n column vectors $\{x_i\}_i$
- Let $V = \{1, 2, \dots, m\}$ be the set of column vector indices.

Skip matrix rank example

Example: Rank function of a matrix

- Given $n \times m$ matrix $\mathbf{X} = (x_1, x_2, \dots, x_m)$ with $x_i \in \mathbb{R}^n$ for all i. There are m length-n column vectors $\{x_i\}_i$
- Let $V = \{1, 2, \dots, m\}$ be the set of column vector indices.
- For any $A \subseteq V$, let $r(A)$ be the rank of the column vectors indexed by A.

Skip matrix rank example

Example: Rank function of a matrix

- Given $n \times m$ matrix $\mathbf{X} = (x_1, x_2, \dots, x_m)$ with $x_i \in \mathbb{R}^n$ for all i. There are m length-n column vectors $\{x_i\}_i$
- Let $V = \{1, 2, \dots, m\}$ be the set of column vector indices.
- For any $A \subseteq V$, let $r(A)$ be the rank of the column vectors indexed by A.
- $r(A)$ is the dimensionality of the vector space spanned by the set of vectors $\{x_a\}_{a \in A}$.

Skip matrix rank example

Example: Rank function of a matrix

- Given $n \times m$ matrix $\mathbf{X} = (x_1, x_2, \dots, x_m)$ with $x_i \in \mathbb{R}^n$ for all i. There are m length-n column vectors $\{x_i\}_i$
- Let $V = \{1, 2, \dots, m\}$ be the set of column vector indices.
- For any $A \subseteq V$, let $r(A)$ be the rank of the column vectors indexed by A.
- $r(A)$ is the dimensionality of the vector space spanned by the set of vectors $\{x_a\}_{a \in A}$.
- Thus, $r(V)$ is the rank of the matrix $\mathbf{X}$.

Skip matrix rank example

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 $r(A) = 3$, $r(B) = 3$, $r(C) = 2$.
- $r(A \cup C) = 3$, $r(B \cup C) = 3$.
- $r(A \cup A_r) = 3$, $r(B \cup B_r) = 3$, $r(A \cup B_r) = 4$, $r(B \cup A_r) = 4$.
- $r(A \cup B) = 4$, $r(A \cap B) = 1 < r(C) = 2$.

Example: Rank function of a matrix

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

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

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

Example: Rank function of a matrix

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

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

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

Example: Rank function of a matrix

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

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

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

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 $r(A) = 3$, $r(B) = 3$, $r(C) = 2$.
- $r(A \cup C) = 3$, $r(B \cup C) = 3$.
- $r(A \cup A_r) = 3$, $r(B \cup B_r) = 3$, $r(A \cup B_r) = 4$, $r(B \cup A_r) = 4$.
- $r(A \cup B) = 4$, $r(A \cap B) = 1 < r(C) = 2$.

Example: Rank function of a matrix

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

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

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

Example: Rank function of a matrix

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

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

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

Example: Rank function of a matrix

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

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

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

Example: Rank function of a matrix

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

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

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

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 $r(A) = 3$, $r(B) = 3$, $r(C) = 2$.
- $r(A \cup C) = 3$, $r(B \cup C) = 3$.
- $r(A \cup A_r) = 3$, $r(B \cup B_r) = 3$, $r(A \cup B_r) = 4$, $r(B \cup A_r) = 4$.
- $r(A \cup B) = 4$, $r(A \cap B) = 1 < r(C) = 2$.

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 $r(A) = 3$, $r(B) = 3$, $r(C) = 2$.
- $r(A \cup C) = 3$, $r(B \cup C) = 3$.
- $r(A \cup A_r) = 3$, $r(B \cup B_r) = 3$, $r(A \cup B_r) = 4$, $r(B \cup A_r) = 4$.
- $r(A \cup B) = 4$, $r(A \cap B) = 1 < r(C) = 2$.

Example: Rank function of a matrix

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

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

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

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 $r(A) = 3$, $r(B) = 3$, $r(C) = 2$.
- $r(A \cup C) = 3$, $r(B \cup C) = 3$.
- $r(A \cup A_r) = 3$, $r(B \cup B_r) = 3$, $r(A \cup B_r) = 4$, $r(B \cup A_r) = 4$.
- $r(A \cup B) = 4$, $r(A \cap B) = 1 < r(C) = 2$.

Example: Rank function of a matrix

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

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

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

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 $r(A) = 3$, $r(B) = 3$, $r(C) = 2$.
- $r(A \cup C) = 3$, $r(B \cup C) = 3$.
- $r(A \cup A_r) = 3$, $r(B \cup B_r) = 3$, $r(A \cup B_r) = 4$, $r(B \cup A_r) = 4$.
- $r(A \cup B) = 4$, $r(A \cap B) = 1 < r(C) = 2$.

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 $r(A) = 3$, $r(B) = 3$, $r(C) = 2$.
- $r(A \cup C) = 3$, $r(B \cup C) = 3$.
- $r(A \cup A_r) = 3$, $r(B \cup B_r) = 3$, $r(A \cup B_r) = 4$, $r(B \cup A_r) = 4$.
- $r(A \cup B) = 4$, $r(A \cap B) = 1 < r(C) = 2$.

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 $r(A) = 3$, $r(B) = 3$, $r(C) = 2$.
- $r(A \cup C) = 3$, $r(B \cup C) = 3$.
- $r(A \cup A_r) = 3$, $r(B \cup B_r) = 3$, $r(A \cup B_r) = 4$, $r(B \cup A_r) = 4$.
- $r(A \cup B) = 4$, $\boxed{r(A \cap B) = 1} < r(C) = 2$.

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 $r(A) = 3$, $r(B) = 3$, $r(C) = 2$.
- $r(A \cup C) = 3$, $r(B \cup C) = 3$.
- $r(A \cup A_r) = 3$, $r(B \cup B_r) = 3$, $r(A \cup B_r) = 4$, $r(B \cup A_r) = 4$.
- $r(A \cup B) = 4$, $r(A \cap B) = 1$ $< r(C) = 2$.

Example: Rank function of a matrix

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

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

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

Rank function of a matrix

- Let $A, B \subseteq V$ be two subsets of column indices.

Rank function of a matrix

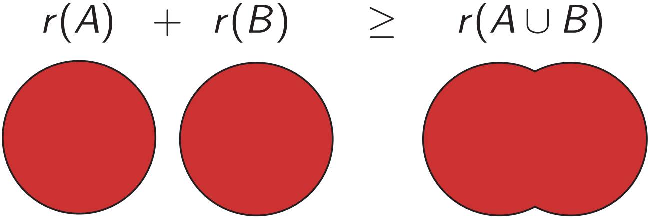
- Let $A, B \subseteq V$ be two subsets of column indices.
- The rank of the two sets unioned together $A \cup B$ is no more than the sum of the two individual ranks.

Rank function of a matrix

- Let $A, B \subseteq V$ be two subsets of column indices.
- The rank of the two sets unioned together $A \cup B$ is no more than the sum of the two individual ranks.
- In a Venn diagram, let area correspond to dimensions spanned by vectors indexed by a set. Hence, $r(A)$ can be viewed as an area.

Rank function of a matrix

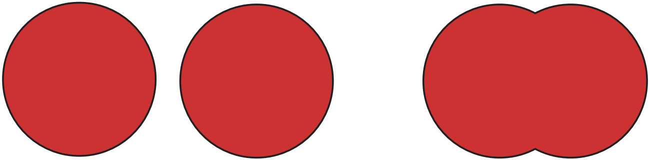
- Let $A, B \subseteq V$ be two subsets of column indices.
- The rank of the two sets unioned together $A \cup B$ is no more than the sum of the two individual ranks.
- In a Venn diagram, let area correspond to dimensions spanned by vectors indexed by a set. Hence, $r(A)$ can be viewed as an area.



Rank function of a matrix

- Let $A, B \subseteq V$ be two subsets of column indices.
- The rank of the two sets unioned together $A \cup B$ is no more than the sum of the two individual ranks.
- In a Venn diagram, let area correspond to dimensions spanned by vectors indexed by a set. Hence, $r(A)$ can be viewed as an area.

$$r(A) \quad + \quad r(B) \qquad \geq \qquad r(A \cup B)$$

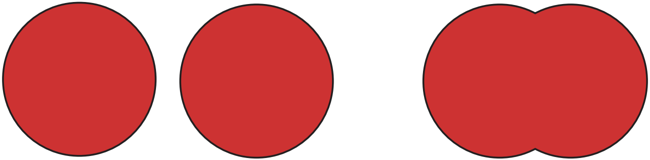


- If some of the dimensions spanned by A overlap some of the dimensions spanned by B (i.e., if $\exists$ common span), then that area is counted twice in $r(A) + r(B)$, so the inequality will be strict.

Rank function of a matrix

- Let $A, B \subseteq V$ be two subsets of column indices.
- The rank of the two sets unioned together $A \cup B$ is no more than the sum of the two individual ranks.
- In a Venn diagram, let area correspond to dimensions spanned by vectors indexed by a set. Hence, $r(A)$ can be viewed as an area.

$$r(A) \quad + \quad r(B) \qquad \geq \qquad r(A \cup B)$$



- If some of the dimensions spanned by A overlap some of the dimensions spanned by B (i.e., if $\exists$ common span), then that area is counted twice in $r(A) + r(B)$, so the inequality will be strict.
- Any function where the above inequality is true for all $A, B \subseteq V$ is called subadditive.

Rank functions of a matrix

- Vector sets A and B have a (possibly empty) common span and two (possibly empty) non-common residual spans.

Rank functions of a matrix

- Vector sets A and B have a (possibly empty) common span and two (possibly empty) non-common residual spans.
- Let C index vectors spanning all dimensions common to A and B. We call C the common span and call $A \cap B$ the common indices.

Rank functions of a matrix

- Vector sets A and B have a (possibly empty) common span and two (possibly empty) non-common residual spans.
- Let C index vectors spanning all dimensions common to A and B. We call C the common span and call $A \cap B$ the common indices.
- Let A_r index vectors spanning dimensions spanned by A but not B.

Rank functions of a matrix

- Vector sets A and B have a (possibly empty) common span and two (possibly empty) non-common residual spans.
- Let C index vectors spanning all dimensions common to A and B. We call C the common span and call $A \cap B$ the common indices.
- Let A_r index vectors spanning dimensions spanned by A but not B.
- Let B_r index vectors spanning dimensions spanned by B but not A.

Rank functions of a matrix

- Vector sets A and B have a (possibly empty) common span and two (possibly empty) non-common residual spans.
- Let C index vectors spanning all dimensions common to A and B. We call C the common span and call $A \cap B$ the common indices.
- Let A_r index vectors spanning dimensions spanned by A but not B.
- Let B_r index vectors spanning dimensions spanned by B but not A.
- Then, $r(A) = r(C) + r(A_r)$

Rank functions of a matrix

- Vector sets A and B have a (possibly empty) common span and two (possibly empty) non-common residual spans.
- Let C index vectors spanning all dimensions common to A and B. We call C the common span and call $A \cap B$ the common indices.
- Let A_r index vectors spanning dimensions spanned by A but not B.
- Let B_r index vectors spanning dimensions spanned by B but not A.
- Then, $r(A) = r(C) + r(A_r)$
- Similarly, $r(B) = r(C) + r(B_r)$.

Rank functions of a matrix

- Vector sets A and B have a (possibly empty) common span and two (possibly empty) non-common residual spans.
- Let C index vectors spanning all dimensions common to A and B. We call C the common span and call $A \cap B$ the common indices.
- Let A_r index vectors spanning dimensions spanned by A but not B.
- Let B_r index vectors spanning dimensions spanned by B but not A.
- Then, $r(A) = r(C) + r(A_r)$
- Similarly, $r(B) = r(C) + r(B_r)$.
- Then $r(A) + r(B)$ counts the dimensions spanned by C twice, i.e.,

$$r(A) + r(B) = r(A_r) + 2r(C) + r(B_r). \tag{3.51}$$

Rank functions of a matrix

- Vector sets A and B have a (possibly empty) common span and two (possibly empty) non-common residual spans.
- Let C index vectors spanning all dimensions common to A and B. We call C the common span and call $A \cap B$ the common indices.
- Let A_r index vectors spanning dimensions spanned by A but not B.
- Let B_r index vectors spanning dimensions spanned by B but not A.
- Then, $r(A) = r(C) + r(A_r)$
- Similarly, $r(B) = r(C) + r(B_r)$.
- Then $r(A) + r(B)$ counts the dimensions spanned by C twice, i.e.,

$$r(A) + r(B) = r(A_r) + 2r(C) + r(B_r). \tag{3.51}$$

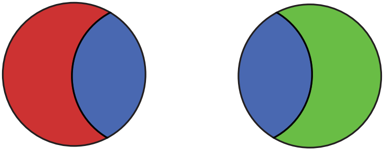
- But $r(A \cup B)$ counts the dimensions spanned by C only once.

$$r(A \cup B) = r(A_r) + r(C) + r(B_r) \tag{3.52}$$

Rank functions of a matrix

- Then $r(A) + r(B)$ counts the dimensions spanned by C twice, i.e.,

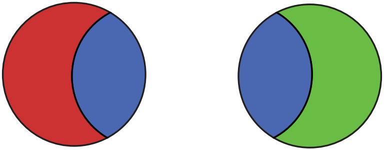
$$r(A) + r(B) = r(A_r) + 2r(C) + r(B_r)$$



Rank functions of a matrix

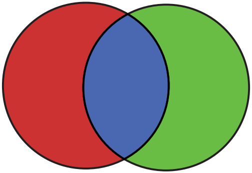
- Then $r(A) + r(B)$ counts the dimensions spanned by C twice, i.e.,

$$r(A) + r(B) = r(A_r) + 2r(C) + r(B_r)$$



- But $r(A \cup B)$ counts the dimensions spanned by C only once.

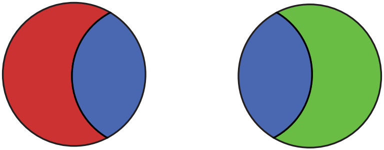
$$r(A \cup B) = r(A_r) + r(C) + r(B_r)$$



Rank functions of a matrix

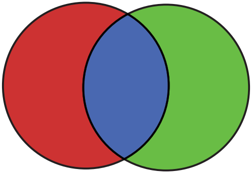
- Then $r(A) + r(B)$ counts the dimensions spanned by C twice, i.e.,

$$r(A) + r(B) = r(A_r) + 2r(C) + r(B_r)$$



- But $r(A \cup B)$ counts the dimensions spanned by C only once.

$$r(A \cup B) = r(A_r) + r(C) + r(B_r)$$

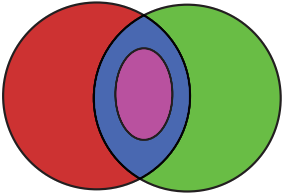


- Thus, we have subadditivity: $r(A) + r(B) \geq r(A \cup B)$. Can we add more to the r.h.s. and still have an inequality? Yes.

Rank function of a matrix

- Note, $r(A \cap B) \leq r(C)$. Why? Vectors indexed by $A \cap B$ (i.e., the common index set) span no more than the dimensions commonly spanned by A and B (namely, those spanned by the professed C).

$$r(C) \geq r(A \cap B)$$

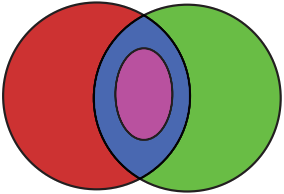


In short:

Rank function of a matrix

- Note, $r(A \cap B) \leq r(C)$. Why? Vectors indexed by $A \cap B$ (i.e., the common index set) span no more than the dimensions commonly spanned by A and B (namely, those spanned by the professed C).

$$r(C) \geq r(A \cap B)$$



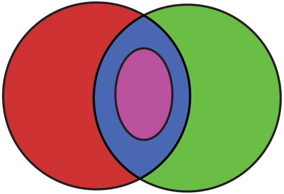
In short:

- Common span (blue) is "more" (no less) than span of common index (magenta).

Rank function of a matrix

- Note, $r(A \cap B) \leq r(C)$. Why? Vectors indexed by $A \cap B$ (i.e., the common index set) span no more than the dimensions commonly spanned by A and B (namely, those spanned by the professed C).

$$r(C) \geq r(A \cap B)$$

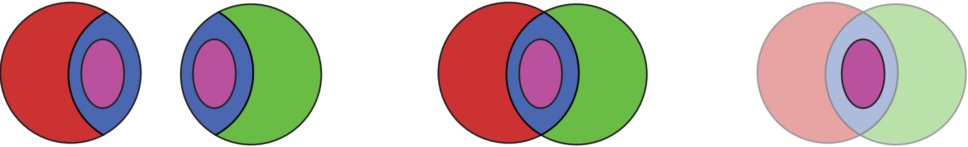


In short:

- Common span (blue) is "more" (no less) than span of common index (magenta).
- More generally, common information (blue) is "more" (no less) than information within common index (magenta).

The Venn and Art of Submodularity

$$\underbrace{r(A) + r(B)}_{= r(A_r) + 2r(C) + r(B_r)} \geq \underbrace{r(A \cup B)}_{= r(A_r) + r(C) + r(B_r)} + \underbrace{r(A \cap B)}_{= r(A \cap B)}$$



Polymatroid rank function

- Let S be a set of subspaces of a linear space (i.e., each $s \in S$ is a subspace of dimension ≥ 1).

Polymatroid rank function

- Let S be a set of subspaces of a linear space (i.e., each $s \in S$ is a subspace of dimension ≥ 1).
- For each $X \subseteq S$, let $f(X)$ denote the dimensionality of the linear subspace spanned by the subspaces in X.

Polymatroid rank function

- Let S be a set of subspaces of a linear space (i.e., each $s \in S$ is a subspace of dimension ≥ 1).
- For each $X \subseteq S$, let $f(X)$ denote the dimensionality of the linear subspace spanned by the subspaces in X.
- We can think of S as a set of sets of vectors from the matrix rank example, and for each $s \in S$, let X_s being a set of vector indices.

Polymatroid rank function

- Let S be a set of subspaces of a linear space (i.e., each $s \in S$ is a subspace of dimension ≥ 1).
- For each $X \subseteq S$, let $f(X)$ denote the dimensionality of the linear subspace spanned by the subspaces in X.
- We can think of S as a set of sets of vectors from the matrix rank example, and for each $s \in S$, let X_s being a set of vector indices.
- Then, defining $f : 2^S \to \mathbb{R}_+$ as follows,

$$f(X) = r(\cup_{s \in X} X_s) \tag{3.53}$$

we have that f is submodular, and is known to be a polymatroid rank function.

Polymatroid rank function

- Let S be a set of subspaces of a linear space (i.e., each $s \in S$ is a subspace of dimension ≥ 1).
- For each $X \subseteq S$, let $f(X)$ denote the dimensionality of the linear subspace spanned by the subspaces in X.
- We can think of S as a set of sets of vectors from the matrix rank example, and for each $s \in S$, let X_s being a set of vector indices.
- Then, defining $f : 2^S \to \mathbb{R}_+$ as follows,

$$f(X) = r(\cup_{s \in X} X_s) \tag{3.53}$$

  we have that f is submodular, and is known to be a polymatroid rank function.
- In general (as we will see) polymatroid rank functions are submodular, normalized $f(\emptyset) = 0$, and monotone non-decreasing ($f(A) \leq f(B)$ whenever $A \subseteq B$).

Polymatroid rank function

- Let S be a set of subspaces of a linear space (i.e., each $s \in S$ is a subspace of dimension ≥ 1).
- For each $X \subseteq S$, let $f(X)$ denote the dimensionality of the linear subspace spanned by the subspaces in X.
- We can think of S as a set of sets of vectors from the matrix rank example, and for each $s \in S$, let X_s being a set of vector indices.
- Then, defining $f : 2^S \to \mathbb{R}_+$ as follows,

$$f(X) = r(\cup_{s \in X} X_s) \tag{3.53}$$

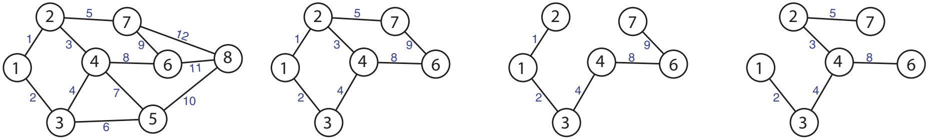
  we have that f is submodular, and is known to be a polymatroid rank function.
- In general (as we will see) polymatroid rank functions are submodular, normalized $f(\emptyset) = 0$, and monotone non-decreasing ($f(A) \leq f(B)$ whenever $A \subseteq B$).
- We use the term non-decreasing rather than increasing, the latter of which is strict (also so that a constant function isn't "increasing").

Spanning trees

- Let E be a set of edges of some graph $G = (V, E)$, and let $r(S)$ for $S \subseteq E$ be the maximum size (in terms of number of **edges**) spanning forest in the graph induced by edges S.

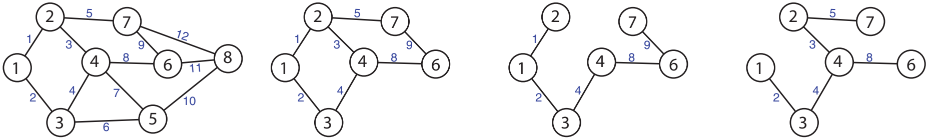
Spanning trees

- Let E be a set of edges of some graph $G = (V, E)$, and let $r(S)$ for $S \subseteq E$ be the maximum size (in terms of number of **edges**) spanning forest in the graph induced by edges S.
- Example: Given $G = (V, E)$, $V = \{1, 2, 3, 4, 5, 6, 7, 8\}$, $E = \{1, 2, \ldots, 12\}$. $S = \{1, 2, 3, 4, 5, 8, 9\} \subset E$. Two spanning trees have the same edge count (the rank of S).



Spanning trees

- Let E be a set of edges of some graph $G = (V, E)$, and let $r(S)$ for $S \subseteq E$ be the maximum size (in terms of number of **edges**) spanning forest in the graph induced by edges S.
- Example: Given $G = (V, E)$, $V = \{1, 2, 3, 4, 5, 6, 7, 8\}$, $E = \{1, 2, \dots, 12\}$. $S = \{1, 2, 3, 4, 5, 8, 9\} \subset E$. Two spanning trees have the same edge count (the rank of S).



- Then $r(S)$ is submodular, and is another matrix rank function corresponding to the incidence matrix of the graph.

Summing Submodular Functions

Given E, let $f_1, f_2 : 2^E \to \mathbb{R}$ be two submodular functions. Then

$$f : 2^E \to \mathbb{R} \text{ with } f(A) = f_1(A) + f_2(A) \tag{3.58}$$

is submodular.

Summing Submodular Functions

Given E, let $f_1, f_2 : 2^E \to \mathbb{R}$ be two submodular functions. Then

$$f : 2^E \to \mathbb{R} \text{ with } f(A) = f_1(A) + f_2(A) \tag{3.58}$$

is submodular. This follows easily since

$$\begin{aligned} f(A) + f(B) &= f_1(A) + f_2(A) + f_1(B) + f_2(B) && (3.59) \\ &\geq f_1(A \cup B) + f_2(A \cup B) + f_1(A \cap B) + f_2(A \cap B) && (3.60) \\ &= f(A \cup B) + f(A \cap B). && (3.61) \end{aligned}$$

I.e., it holds for each component of f in each term in the inequality.

Summing Submodular Functions

Given E, let $f_1, f_2 : 2^E \to \mathbb{R}$ be two submodular functions. Then

$$f : 2^E \to \mathbb{R} \text{ with } f(A) = f_1(A) + f_2(A) \tag{3.58}$$

is submodular. This follows easily since

$$\begin{aligned} f(A) + f(B) &= f_1(A) + f_2(A) + f_1(B) + f_2(B) && (3.59) \\ &\geq f_1(A \cup B) + f_2(A \cup B) + f_1(A \cap B) + f_2(A \cap B) && (3.60) \\ &= f(A \cup B) + f(A \cap B). && (3.61) \end{aligned}$$

I.e., it holds for each component of f in each term in the inequality. In fact, any conic combination (i.e., non-negative linear combination) of submodular functions is submodular, as in $f(A) = \alpha_1 f_1(A) + \alpha_2 f_2(A)$ for $\alpha_1, \alpha_2 \geq 0$.

Summing Submodular and Modular Functions

Given E, let $f_1, m : 2^E \to \mathbb{R}$ be a submodular and a modular function.

Summing Submodular and Modular Functions

Given E, let $f_1, m : 2^E \to \mathbb{R}$ be a submodular and a modular function. Then

$$f : 2^E \to \mathbb{R} \text{ with } f(A) = f_1(A) - m(A) \tag{3.62}$$

is submodular (as is $f(A) = f_1(A) + m(A)$).

Summing Submodular and Modular Functions

Given E, let $f_1, m : 2^E \to \mathbb{R}$ be a submodular and a modular function. Then

$$f : 2^E \to \mathbb{R} \text{ with } f(A) = f_1(A) - m(A) \tag{3.62}$$

is submodular (as is $f(A) = f_1(A) + m(A)$). This follows easily since

$$\begin{aligned} f(A) + f(B) &= f_1(A) - m(A) + f_1(B) - m(B) && (3.63)\\ &\geq f_1(A \cup B) - m(A \cup B) + f_1(A \cap B) - m(A \cap B) && (3.64)\\ &= f(A \cup B) + f(A \cap B). && (3.65) \end{aligned}$$

Summing Submodular and Modular Functions

Given E, let $f_1, m : 2^E \to \mathbb{R}$ be a submodular and a modular function. Then

$$f : 2^E \to \mathbb{R} \text{ with } f(A) = f_1(A) - m(A) \tag{3.62}$$

is submodular (as is $f(A) = f_1(A) + m(A)$). This follows easily since

$$\begin{aligned} f(A) + f(B) &= f_1(A) - m(A) + f_1(B) - m(B) &(3.63)\\ &\geq f_1(A \cup B) - m(A \cup B) + f_1(A \cap B) - m(A \cap B) &(3.64)\\ &= f(A \cup B) + f(A \cap B). &(3.65) \end{aligned}$$

That is, the modular component with $m(A) + m(B) = m(A \cup B) + m(A \cap B)$ never destroys the inequality. Note of course that if m is modular than so is $-m$.

Restricting Submodular functions

Given E, let $f : 2^E \to \mathbb{R}$ be a submodular functions. And let $S \subseteq E$ be an arbitrary fixed set. Then

$$f' : 2^E \to \mathbb{R} \text{ with } f'(A) \triangleq f(A \cap S) \tag{3.66}$$

is submodular.

Restricting Submodular functions

Given E, let $f : 2^E \to \mathbb{R}$ be a submodular functions. And let $S \subseteq E$ be an arbitrary fixed set. Then

$$f' : 2^E \to \mathbb{R} \text{ with } f'(A) \triangleq f(A \cap S) \tag{3.66}$$

is submodular.

Proof.

$\square$

Restricting Submodular functions

Given E, let $f : 2^E \to \mathbb{R}$ be a submodular functions. And let $S \subseteq E$ be an arbitrary fixed set. Then

$$f' : 2^E \to \mathbb{R} \text{ with } f'(A) \triangleq f(A \cap S) \tag{3.66}$$

is submodular.

Proof.

Given $A \subseteq B \subseteq E \setminus v$, consider

$$f((A + v) \cap S) - f(A \cap S) \geq f((B + v) \cap S) - f(B \cap S) \tag{3.67}$$

□

Restricting Submodular functions

Given E, let $f : 2^E \to \mathbb{R}$ be a submodular functions. And let $S \subseteq E$ be an arbitrary fixed set. Then

$$f' : 2^E \to \mathbb{R} \text{ with } f'(A) \triangleq f(A \cap S) \tag{3.66}$$

is submodular.

Proof.

Given $A \subseteq B \subseteq E \setminus v$, consider

$$f((A+v) \cap S) - f(A \cap S) \geq f((B+v) \cap S) - f(B \cap S) \tag{3.67}$$

If $v \notin S$, then both differences on each size are zero.

□

Restricting Submodular functions

Given E, let $f : 2^E \to \mathbb{R}$ be a submodular functions. And let $S \subseteq E$ be an arbitrary fixed set. Then

$$f' : 2^E \to \mathbb{R} \text{ with } f'(A) \triangleq f(A \cap S) \tag{3.66}$$

is submodular.

Proof.

Given $A \subseteq B \subseteq E \setminus v$, consider

$$f((A + v) \cap S) - f(A \cap S) \geq f((B + v) \cap S) - f(B \cap S) \tag{3.67}$$

If $v \notin S$, then both differences on each size are zero. If $v \in S$, then we can consider this

$$f(A' + v) - f(A') \geq f(B' + v) - f(B') \tag{3.68}$$

with $A' = A \cap S$ and $B' = B \cap S$. Since $A' \subseteq B'$, this holds due to submodularity of f. □

Summing Restricted Submodular Functions

Given V, let $f_1, f_2 : 2^V \to \mathbb{R}$ be two submodular functions and let $S_1, S_2 \subseteq V$ be two arbitrary fixed sets. Then

$$f : 2^V \to \mathbb{R} \text{ with } f(A) = f_1(A \cap S_1) + f_2(A \cap S_2) \tag{3.69}$$

is submodular. This follows easily from the preceding two results.

Summing Restricted Submodular Functions

Given V, let $f_1, f_2 : 2^V \to \mathbb{R}$ be two submodular functions and let $S_1, S_2 \subseteq V$ be two arbitrary fixed sets. Then

$$f : 2^V \to \mathbb{R} \text{ with } f(A) = f_1(A \cap S_1) + f_2(A \cap S_2) \tag{3.69}$$

is submodular. This follows easily from the preceding two results.

Given V, let $\mathcal{C} = \{C_1, C_2, \ldots, C_k\}$ be a set of subsets of V, and for each $C \in \mathcal{C}$, let $f_C : 2^V \to \mathbb{R}$ be a submodular function. Then

$$f : 2^V \to \mathbb{R} \text{ with } f(A) = \sum_{C \in \mathcal{C}} f_C(A \cap C) \tag{3.70}$$

is submodular.

Summing Restricted Submodular Functions

Given V, let $f_1, f_2 : 2^V \to \mathbb{R}$ be two submodular functions and let $S_1, S_2 \subseteq V$ be two arbitrary fixed sets. Then

$$f : 2^V \to \mathbb{R} \text{ with } f(A) = f_1(A \cap S_1) + f_2(A \cap S_2) \tag{3.69}$$

is submodular. This follows easily from the preceding two results.
Given V, let $\mathcal{C} = \{C_1, C_2, \dots, C_k\}$ be a set of subsets of V, and for each $C \in \mathcal{C}$, let $f_C : 2^V \to \mathbb{R}$ be a submodular function. Then

$$f : 2^V \to \mathbb{R} \text{ with } f(A) = \sum_{C \in \mathcal{C}} f_C(A \cap C) \tag{3.70}$$

is submodular. This property is critical for image processing and graphical models. For example, let $\mathcal{C}$ be all pairs of the form $\{\{u, v\} : u, v \in V\}$, or let it be all pairs corresponding to the edges of some undirected graphical model.

Max - normalized

Given V, let $c \in \mathbb{R}_+^V$ be a given fixed vector. Then $f : 2^V \to \mathbb{R}_+$, where

$$f(A) = \max_{j \in A} c_j \tag{3.71}$$

is submodular and normalized (we take $f(\emptyset) = 0$).

Proof.

Consider

$$\max_{j \in A} c_j + \max_{j \in B} c_j \geq \max_{j \in A \cup B} c_j + \max_{j \in A \cap B} c_j \tag{3.72}$$

which follows since we have that

$$\max(\max_{j \in A} c_j, \max_{j \in B} c_j) = \max_{j \in A \cup B} c_j \tag{3.73}$$

and

$$\min(\max_{j \in A} c_j, \max_{j \in B} c_j) \geq \max_{j \in A \cap B} c_j \tag{3.74}$$

□

Max

Given V, let $c \in \mathbb{R}^V$ be a given fixed vector (not necessarily non-negative). Then $f : 2^V \to \mathbb{R}$, where

$$f(A) = \max_{j \in A} c_j \tag{3.75}$$

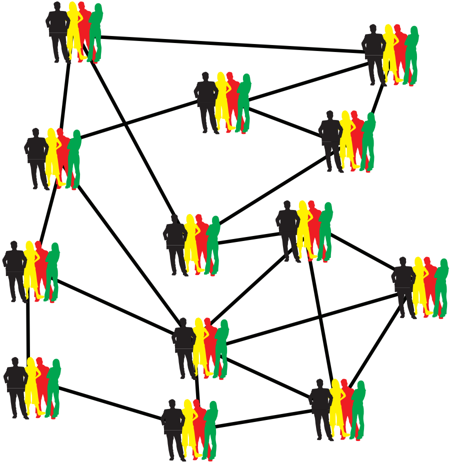
is submodular, where we take $f(\emptyset) \leq \min_j c_j$ (so the function need not be normalized).

Proof.

The proof is identical to the normalized case. □

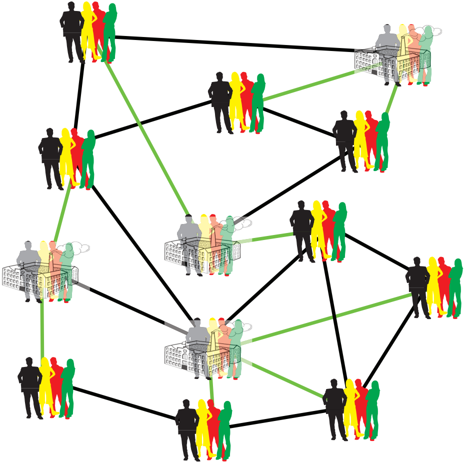
Facility/Plant Location (uncapacitated)

- Core problem in operations research, early motivation for submodularity.
- Goal: as efficiently as possible, place "facilities" (factories) at certain locations to satisfy sites (at all locations) having various demands.



Facility/Plant Location (uncapacitated)

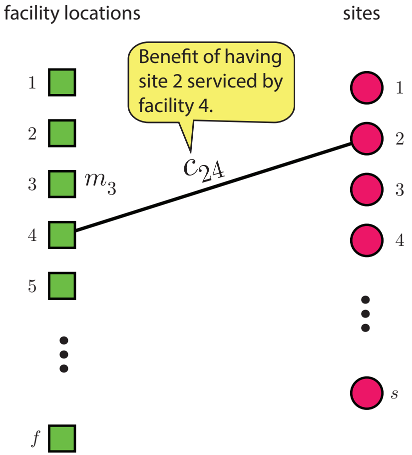
- Core problem in operations research, early motivation for submodularity.
- Goal: as efficiently as possible, place "facilities" (factories) at certain locations to satisfy sites (at all locations) having various demands.



Facility/Plant Location (uncapacitated)

- Core problem in operations research, early motivation for submodularity.
- Goal: as efficiently as possible, place "facilities" (factories) at certain locations to satisfy sites (at all locations) having various demands.
- We can model this with a weighted bipartite graph $G = (F, S, E, c)$ where F is set of possible factory/plant locations, S is set of sites needing service, E are edges indicating (factory,site) service possibility pairs, and $c : E \to \mathbb{R}_+$ is the benefit of a given pair.
- Facility location function has form:

$$f(A) = \sum_{i \in S} \max_{j \in A} c_{ij}. \qquad (3.76)$$



Facility/Plant Location (uncapacitated) w. plant benefits

- Let $F = \{1, \ldots, f\}$ be a set of possible factory/plant locations for facilities to be built.

Facility/Plant Location (uncapacitated) w. plant benefits

- Let $F = \{1, \ldots, f\}$ be a set of possible factory/plant locations for facilities to be built.
- $S = \{1, \ldots, s\}$ is a set of sites (e.g., cities, clients) needing service.

Facility/Plant Location (uncapacitated) w. plant benefits

- Let $F = \{1, \dots, f\}$ be a set of possible factory/plant locations for facilities to be built.
- $S = \{1, \dots, s\}$ is a set of sites (e.g., cities, clients) needing service.
- Let c_{ij} be the "benefit" or "value" (e.g., $1/c_{ij}$ is the cost) of servicing site i with facility location j.

Facility/Plant Location (uncapacitated) w. plant benefits

- Let $F = \{1, \ldots, f\}$ be a set of possible factory/plant locations for facilities to be built.
- $S = \{1, \ldots, s\}$ is a set of sites (e.g., cities, clients) needing service.
- Let c_{ij} be the "benefit" or "value" (e.g., $1/c_{ij}$ is the cost) of servicing site i with facility location j.
- Let m_j be the benefit (e.g., either $1/m_j$ is the cost or $-m_j$ is the cost) to build a plant at location j.

Facility/Plant Location (uncapacitated) w. plant benefits

- Let $F = \{1, \ldots, f\}$ be a set of possible factory/plant locations for facilities to be built.
- $S = \{1, \ldots, s\}$ is a set of sites (e.g., cities, clients) needing service.
- Let c_{ij} be the "benefit" or "value" (e.g., $1/c_{ij}$ is the cost) of servicing site i with facility location j.
- Let m_j be the benefit (e.g., either $1/m_j$ is the cost or $-m_j$ is the cost) to build a plant at location j.
- Each site should be serviced by only one plant but no less than one.

Facility/Plant Location (uncapacitated) w. plant benefits

- Let $F = \{1, \ldots, f\}$ be a set of possible factory/plant locations for facilities to be built.
- $S = \{1, \ldots, s\}$ is a set of sites (e.g., cities, clients) needing service.
- Let c_{ij} be the "benefit" or "value" (e.g., $1/c_{ij}$ is the cost) of servicing site i with facility location j.
- Let m_j be the benefit (e.g., either $1/m_j$ is the cost or $-m_j$ is the cost) to build a plant at location j.
- Each site should be serviced by only one plant but no less than one.
- Define $f(A)$ as the "delivery benefit" plus "construction benefit" when the locations $A \subseteq F$ are to be constructed.

Facility/Plant Location (uncapacitated) w. plant benefits

- Let $F = \{1, \dots, f\}$ be a set of possible factory/plant locations for facilities to be built.
- $S = \{1, \dots, s\}$ is a set of sites (e.g., cities, clients) needing service.
- Let c_{ij} be the "benefit" or "value" (e.g., $1/c_{ij}$ is the cost) of servicing site i with facility location j.
- Let m_j be the benefit (e.g., either $1/m_j$ is the cost or $-m_j$ is the cost) to build a plant at location j.
- Each site should be serviced by only one plant but no less than one.
- Define $f(A)$ as the "delivery benefit" plus "construction benefit" when the locations $A \subseteq F$ are to be constructed.
- We can define the (uncapacitated) facility location function

$$f(A) = \sum_{j \in A} m_j + \sum_{i \in S} \max_{j \in A} c_{ij}. \tag{3.77}$$

Facility/Plant Location (uncapacitated) w. plant benefits

- Let $F = \{1, \dots, f\}$ be a set of possible factory/plant locations for facilities to be built.
- $S = \{1, \dots, s\}$ is a set of sites (e.g., cities, clients) needing service.
- Let c_{ij} be the "benefit" or "value" (e.g., $1/c_{ij}$ is the cost) of servicing site i with facility location j.
- Let m_j be the benefit (e.g., either $1/m_j$ is the cost or $-m_j$ is the cost) to build a plant at location j.
- Each site should be serviced by only one plant but no less than one.
- Define $f(A)$ as the "delivery benefit" plus "construction benefit" when the locations $A \subseteq F$ are to be constructed.
- We can define the (uncapacitated) facility location function

$$f(A) = \sum_{j \in A} m_j + \sum_{i \in S} \max_{j \in A} c_{ij}. \tag{3.77}$$

- Goal is to find a set A that maximizes $f(A)$ (the benefit) placing a bound on the number of plants A (e.g., $|A| \leq k$).

Facility Location

Given V, E, let $c \in \mathbb{R}^{V \times E}$ be a given $|V| \times |E|$ matrix. Then

$$f : 2^E \to \mathbb{R}, \text{ where } f(A) = \sum_{i \in V} \max_{j \in A} c_{ij} \tag{3.78}$$

is submodular.

Proof.

We can write $f(A)$ as $f(A) = \sum_{i \in V} f_i(A)$ where $f_i(A) = \max_{j \in A} c_{ij}$ is submodular (max of a i^{th} row vector), so f can be written as a sum of submodular functions. □

Thus, the facility location function (which only adds a modular function to the above) is submodular.

Log Determinant

- Let $\mathbf{\Sigma}$ be an $n \times n$ positive definite matrix. Let $V = \{1, 2, \ldots, n\} \equiv [n]$ be an index set, and for $A \subseteq V$, let $\mathbf{\Sigma}_A$ be the (square) submatrix of $\mathbf{\Sigma}$ obtained by including only entries in the rows/columns given by A.

Log Determinant

- Let $\mathbf{\Sigma}$ be an $n \times n$ positive definite matrix. Let $V = \{1, 2, \dots, n\} \equiv [n]$ be an index set, and for $A \subseteq V$, let $\mathbf{\Sigma}_A$ be the (square) submatrix of $\mathbf{\Sigma}$ obtained by including only entries in the rows/columns given by A.
- We have that:

$$f(A) = \log\det(\mathbf{\Sigma}_A) \text{ is submodular.} \tag{3.79}$$

Log Determinant

- Let $\mathbf{\Sigma}$ be an $n \times n$ positive definite matrix. Let $V = \{1, 2, \ldots, n\} \equiv [n]$ be an index set, and for $A \subseteq V$, let $\mathbf{\Sigma}_A$ be the (square) submatrix of $\mathbf{\Sigma}$ obtained by including only entries in the rows/columns given by A.
- We have that:

$$f(A) = \log\det(\mathbf{\Sigma}_A) \text{ is submodular.} \tag{3.79}$$

- The submodularity of the log determinant is crucial for determinantal point processes (DPPs) (defined later in the class).

Log Determinant

- Let $\mathbf{\Sigma}$ be an $n \times n$ positive definite matrix. Let $V = \{1, 2, \ldots, n\} \equiv [n]$ be an index set, and for $A \subseteq V$, let $\mathbf{\Sigma}_A$ be the (square) submatrix of $\mathbf{\Sigma}$ obtained by including only entries in the rows/columns given by A.
- We have that:

$$f(A) = \log\det(\mathbf{\Sigma}_A) \text{ is submodular.} \tag{3.79}$$

- The submodularity of the log determinant is crucial for determinantal point processes (DPPs) (defined later in the class).

Proof of submodularity of the logdet function.

Suppose $X \in \mathbf{R}^n$ is multivariate Gaussian random variable, that is

$$x \in p(x) = \frac{1}{\sqrt{|2\pi\mathbf{\Sigma}|}} \exp\left(-\frac{1}{2}(x-\mu)^T\mathbf{\Sigma}^{-1}(x-\mu)\right) \tag{3.80}$$

...

Log Determinant

...cont.

Then the (differential) entropy of the r.v. X is given by

$$h(X) = \log \sqrt{|2\pi e \mathbf{\Sigma}|} = \log \sqrt{(2\pi e)^n |\mathbf{\Sigma}|} \tag{3.81}$$

and in particular, for a variable subset A,

$$f(A) = h(X_A) = \log \sqrt{(2\pi e)^{|A|} |\mathbf{\Sigma}_A|} \tag{3.82}$$

Entropy is submodular (further conditioning reduces entropy), and moreover

$$f(A) = h(X_A) = m(A) + \frac{1}{2} \log |\mathbf{\Sigma}_A| \tag{3.83}$$

where $m(A)$ is a modular function. □

Note: still submodular in the semi-definite case as well.

Summary: Properties so far

- 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.

Summary: Properties so far

- 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 $f(A) = \sum_{u \in U} \max_{a \in A} s_{a,u}$.

Summary: Properties so far

- 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 $f(A) = \sum_{u \in U} \max_{a \in A} s_{a,u}$.
- Log determinant $f(A) = \log\det(\mathbf{\Sigma}_A)$ submodular for p.d. Σ.

Summary: Properties so far

- 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 $f(A) = \sum_{u \in U} \max_{a \in A} s_{a,u}$.
- Log determinant $f(A) = \log \det(\mathbf{\Sigma}_A)$ submodular for p.d. Σ.
- Matrix rank function is submodular.

Summary: Properties so far

- 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 $f(A) = \sum_{u \in U} \max_{a \in A} s_{a,u}$.
- Log determinant $f(A) = \log\det(\mathbf{\Sigma}_A)$ submodular for p.d. Σ.
- Matrix rank function is submodular.
- Graph cut, set cover, and incidence functions, and quadratics with non-positive off-diagonals, are all submodular.

Summary: Properties so far

- 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 $f(A) = \sum_{u \in U} \max_{a \in A} s_{a,u}$.
- Log determinant $f(A) = \log\det(\mathbf{\Sigma}_A)$ submodular for p.d. Σ.
- Matrix rank function is submodular.
- Graph cut, set cover, and incidence functions, and quadratics with non-positive off-diagonals, are all submodular.
- Number of connected components in induced graph, and interior edge function, is supermodular.

Summary: Properties so far

- 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 $f(A) = \sum_{u \in U} \max_{a \in A} s_{a,u}$.
- Log determinant $f(A) = \log\det(\mathbf{\Sigma}_A)$ submodular for p.d. Σ.
- Matrix rank function is submodular.
- Graph cut, set cover, and incidence functions, and quadratics with non-positive off-diagonals, are all submodular.
- Number of connected components in induced graph, and interior edge function, is supermodular.
- Submodular plus modular is submodular, $f(A) = f'(A) + m(A)$.

Summary: Properties so far

- 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 $f(A) = \sum_{u \in U} \max_{a \in A} s_{a,u}$.
- Log determinant $f(A) = \log\det(\mathbf{\Sigma}_A)$ submodular for p.d. Σ.
- Matrix rank function is submodular.
- Graph cut, set cover, and incidence functions, and quadratics with non-positive off-diagonals, are all submodular.
- Number of connected components in induced graph, and interior edge function, 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 and m is modular. (supermodular) if f is submodular (supermodular).

Summary: Properties so far

- 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 $f(A) = \sum_{u \in U} \max_{a \in A} s_{a,u}$.
- Log determinant $f(A) = \log\det(\mathbf{\Sigma}_A)$ submodular for p.d. Σ.
- Matrix rank function is submodular.
- Graph cut, set cover, and incidence functions, and quadratics with non-positive off-diagonals, are all submodular.
- Number of connected components in induced graph, and interior edge function, 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 and m is modular. (supermodular) if f is submodular (supermodular).
- Summing: if $\alpha_i \geq 0$ and $f_i : 2^V \to \mathbb{R}$ is submodular, then so is $\sum_i \alpha_i f_i$.

Summary: Properties so far

- 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 $f(A) = \sum_{u \in U} \max_{a \in A} s_{a,u}$.
- Log determinant $f(A) = \log\det(\mathbf{\Sigma}_A)$ submodular for p.d. Σ.
- Matrix rank function is submodular.
- Graph cut, set cover, and incidence functions, and quadratics with non-positive off-diagonals, are all submodular.
- Number of connected components in induced graph, and interior edge function, 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 and m is modular. (supermodular) if f is submodular (supermodular).
- 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 preserve submodularity: $f'(A) = f(A \cap S)$