Announcements, Assignments, and Reminders

- Reminder: class web page is at our web page (http://j.ee.washington.edu/~bilmes/classes/ee596a_fall_2012/)
- Please do use our discussion board (https://catalyst.uw.edu/gopost/board/bilmes/29948/) for all questions, comments, so that all will benefit from them being answered.
Outstanding Reading

- Read chapter 1 from Fujishige book.
- Read over lecture slides, all posted on our web page (http://j.ee.washington.edu/~bilmes/classes/ee596a_fall_2012/).
Submodular Motivation Recap

- Given a set of objects $V = \{v_1, \ldots, v_n\}$ and a function $f : 2^V \rightarrow \mathbb{R}$ that returns a real value for any subset $S \subseteq V$.
- Suppose we are interested in finding the subset that either maximizes or minimizes the function, e.g., $\text{argmax}_{S \subseteq V} f(S)$, possibly subject to some constraints.
- In general, this problem has exponential time complexity.
- Example: $f$ might correspond to the value (e.g., information gain) of a set of sensor locations in an environment, and we wish to find the best set $S \subseteq V$ of sensors locations given a fixed upper limit on the number of sensors $|S|$.
- In many cases (such as above) $f$ has properties that make its optimization tractable to either exactly or approximately compute.
- One such property is submodularity.

Submodular Definitions

**Definition 3.2.4 (submodular concave)**

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

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B)$$  \hspace{1cm} (3.7)

An alternate and (as we see in lecture 3) equivalent definition is:

**Definition 3.2.5 (diminishing returns)**

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

$$f(A \cup \{v\}) - f(A) \geq f(B \cup \{v\}) - f(B)$$  \hspace{1cm} (3.8)

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

- In the last lecture, we started looking at properties of and gaining intuition about submodular functions.
- We began to see that there were many functions that were submodular, and operations on sets of submodular functions that preserved submodularity.

Some examples form last time

- Coverage functions (either via sets, or via regions in $n$-D space).
- Entropy function (as a function of sets of random variables), symmetric mutual information.
- Many functions based on graphs are either submodular or supermodular, and other functions might not be (e.g., graph strength) but involve submodularity in a critical way.
- Matrix rank - rank of a set of vectors from a set of vector indices.
- Geometric interpretation of $f(A) + f(B) \geq f(A \cup B) + f(A \cap B)$.
- Cost of manufacturing – supply side economies of scale
- Network Externalities – Demand side Economies of Scale
- Social Network Influence
- Information and Summarization - document summarization via sentence selection
The Venn and Art of Submodularity

\[
\begin{align*}
  r(A) + r(B) & \geq r(A \cup B) + r(A \cap B) \\
  &= r(A_r) + 2r(C) + r(B_r) \\
  &= r(A_r) + r(C) + r(B_r) \\
  &= r(A \cap B)
\end{align*}
\]

Operations

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

**Concave over non-negative modular**

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

$$f(A) = g(m(A))$$

then $f$ is submodular.

**Proof.**

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

$$g(m(A) + m(v)) - g(m(A)) \geq g(m(B) + m(v)) - g(m(B))$$

A form of converse is true as well.
Monotonicity

**Definition 3.3.1**

A function $f : 2^V \to \mathbb{R}$ is **monotone nondecreasing** (resp. **monotone increasing**) if for all $A \subseteq B$, we have $f(A) \leq f(B)$ (resp. $f(A) < f(B)$).

**Definition 3.3.2**

A function $f : 2^V \to \mathbb{R}$ is **monotone nonincreasing** (resp. **monotone decreasing**) if for all $A \subseteq B$, we have $f(A) \geq f(B)$ (resp. $f(A) > f(B)$).
Composition of submodular and concave

**Theorem 3.3.3**

*Given two functions, one defined on sets*

\[ f : 2^V \rightarrow \mathbb{R} \quad (3.69) \]

*and another continuous valued one:*

\[ g : \mathbb{R} \rightarrow \mathbb{R} \quad (3.70) \]

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

---

Monotone difference of two functions

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

\[ h(A) = \min(f(A), g(A)) \quad (3.71) \]

is submodular.

**Proof.**

If \( h(A) \) agrees with either \( f \) or \( g \) on both \( X \) and \( Y \), the result follows since

\[ f(X) + f(Y) \geq \min(f(X \cup Y), g(X \cup Y)) + \min(f(X \cap Y), g(X \cap Y)) \]

(3.72)

...
Monotone difference of two functions

...cont.

Otherwise, w.l.o.g., \( h(X) = f(X) \) and \( h(Y) = g(Y) \), giving

\[
h(X) + h(Y) = f(X) + g(Y) \geq f(X \cup Y) + f(X \cap Y) + g(Y) - f(Y)
\]

(3.73)

Assume the case where \( f - g \) is monotone increasing. Hence,

\[
f(X \cup Y) + g(Y) - f(Y) \geq g(X \cup Y)\]

giving

\[
h(X) + h(Y) \geq g(X \cup Y) + f(X \cap Y) \geq h(X \cup Y) + h(X \cap Y)
\]

(3.74)

What is an easy way to prove the case where \( f - g \) is monotone decreasing?

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

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

\[
h(A) = \min(k, f(A))
\]

(3.75)

is submodular.

**Proof.**

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

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

In general, the minimum of two submodular functions is not submodular. However, when wishing to maximize two monotone non-decreasing submodular functions, we can define function $h : 2^V \rightarrow \mathbb{R}$ as

$$h(A) = \frac{1}{2} \left( \min(k, f) + \min(k, g) \right)$$

then $h$ is submodular, and $h(A) \geq k$ if and only if both $f(A) \geq k$ and $g(A) \geq k$.

This can be useful in many applications. Moreover, this is an instance of a submodular surrogate (where we take a non-submodular problem and find a submodular one that can tell us something). We hope to revisit this again later in the quarter.

Arbitrary functions as difference between submodular funcs.

Given an arbitrary set function $f$, it can be expressed as a difference between two submodular functions: $f = g - h$ where both $g$ and $h$ are submodular.

Define

Proof.

Let $f$ be given and arbitrary.

$$\alpha \triangleq \min_{X,Y} \left( f(X) + f(Y) - f(X \cup Y) - f(X \cap Y) \right)$$  

If $\alpha \geq 0$ then $f$ is submodular, so by assumption $\alpha < 0$. Now let $h$ be an arbitrary strict submodular function and define

$$\beta \triangleq \min_{X,Y} \left( h(X) + h(Y) - h(X \cup Y) - h(X \cap Y) \right)$$

Strict means $\beta > 0$. ...
Arbitrary functions as difference between submodular funcns.

...cont.

Define \( f' : 2^V \to \mathbb{R} \) as

\[
f'(A) = f(A) + \frac{|\alpha|}{\beta} h(A)
\] (3.79)

Then \( f' \) is submodular (why?), and \( f = f'(A) - \frac{|\alpha|}{\beta} h(A) \), a difference between two submodular functions as desired.

Submodular Definitions

**Definition 3.4.4 (submodular concave)**

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

\[
f(A) + f(B) \geq f(A \cup B) + f(A \cap B)
\] (3.7)

An alternate and (as we see in lecture 3) equivalent definition is:

**Definition 3.4.5 (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:

\[
f(A \cup \{v\}) - f(A) \geq f(B \cup \{v\}) - f(B)
\] (3.8)

This means that the incremental “value”, “gain”, or “cost” of \( v \) decreases (diminishes) as the context in which \( v \) is considered grows from \( A \) to \( B \).
Submodular Definition: Group Diminishing Returns

An alternate and equivalent definition is:

Definition 3.4.1 (group diminishing returns)

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

\[
    f(A \cup C) - f(A) \geq f(B \cup C) - f(B)
\]  

(3.80)

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

Submodular Definition Basic Equivalencies

We want to show that Submodular Concave (Definition 3.4.4), Diminishing Returns (Definition 3.4.5), and Group Diminishing Returns (Definition 3.4.1) are identical. We will show that:

- Submodular Concave \( \Rightarrow \) Diminishing Returns
- Diminishing Returns \( \Rightarrow \) Group Diminishing Returns
- Group Diminishing Returns \( \Rightarrow \) Submodular Concave
Submodular Concave ⇒ Diminishing Returns

Proof.

- Assume Submodular concave, so ∀A, B we have
  \[ f(A) + f(B) \geq f(A \cup B) + f(A \cap B). \]
- Given A, B and v ∈ V such that: A ⊆ B ⊆ V \{v\}, we have from submodular concave that:
  \[ f(A + v) + f(B) \geq f(B + v) + f(A) \quad (3.81) \]
- Rearranging, we have
  \[ f(A + v) - f(A) \geq f(B + v) - f(B) \quad (3.82) \]

Diminishing Returns ⇒ Group Diminishing Returns

Proof.

Let \( C = \{c_1, c_2, \ldots, c_k\} \). Then diminishing returns implies

\[
f(A \cup C) - f(A) = f(A \cup C) - \sum_{i=1}^{k-1} \left( f(A \cup \{c_1, \ldots, c_i\}) - f(A \cup \{c_1, \ldots, c_i\}) \right) - f(A) \quad (3.83)
\]
\[
= \sum_{i=1}^{k} f(A \cup \{c_1 \ldots c_i\}) - f(A \cup \{c_1 \ldots c_{i-1}\}) \quad (3.84)
\]
\[
\geq \sum_{i=1}^{k} f(B \cup \{c_1 \ldots c_i\}) - f(B \cup \{c_1 \ldots c_{i-1}\}) \quad (3.85)
\]
\[
= f(B \cup C) - \sum_{i=1}^{k-1} \left( f(B \cup \{c_1, \ldots, c_i\}) - f(B \cup \{c_1, \ldots, c_i\}) \right) - f(B) \quad (3.86)
\]
\[
= f(B \cup C) - f(B) \quad (3.87)
\]
Proof.
Assume group diminishing returns. Assume $A \neq B$ otherwise trivial.
Define $A' = A \cap B$, $C = A \setminus B$, and $B' = B$. Then

$$f(A' + C) - f(A') \geq f(B' + C) - f(B')$$

(3.89)

giving

$$f(A' + C) + f(B') \geq f(B' + C) + f(A')$$

(3.90)
or

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

(3.91)

which is the same as the submodular concave condition

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

(3.92)

Submodular Definition: Four Points

Definition 3.4.2 ("singleton", or "four points")

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

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

(3.93)

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

$$f(A + a) - f(A) \geq f(A + b_1 + a) - f(A + b_1) \geq f(A + b_1 + b_2 + a) - f(A + b_1 + b_2) \geq \ldots \geq f(A + b_1 + \cdots + b_k + a) - f(A + b_1 + \cdots + b_k) = f(B + a) - f(B)$$
Gain

It is often the case that we wish to express the gain of an item \( j \in V \) in some context, say \( A \), namely \( f(A \cup \{j\}) - f(A) \). This is called the gain and is used so often, that there are equally as many ways to notate this. I.e., in the literature I have seen:

\[
\begin{align*}
  f(A \cup \{j\}) - f(A) &\triangleq \rho_j(A) \\
                      &\triangleq \rho_A(j) \\
                      &\triangleq f(\{j\}|A) \\
                      &\triangleq f(j|A)
\end{align*}
\]

(3.99) (3.100) (3.101) (3.102)

We’ll use \( f(j|A) \). Note, diminishing returns can now be stated as saying that \( f(j|A) \) is a monotone non-increasing function of \( A \), since \( f(j|A) \geq f(j|B) \) whenever \( B \supseteq A \) (conditioning reduces valuation, just like entropy).

Gain Notation

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

\[
f(A|B) \triangleq f(A \cup B) - f(B)
\]

(3.103)

So when \( j \) is any singleton

\[
f(j|B) = f(\{j\}|B) = f(\{j\} \cup B) - f(B)
\]

(3.104)

Note that this is inspired from information theory and the notation used for conditional entropy \( H(X_A|X_B) = H(X_A, X_B) - H(X_B) \).
Many (Equivalent) Definitions of Submodularity

\[ f(A) + f(B) \geq f(A \cup B) + f(A \cap B), \ \forall A, B \subseteq E \]  
\[ f(j|S) \geq f(j|T), \ \forall S \subseteq T \subseteq E, \ \text{with} \ j \in E \setminus T \]  
\[ f(C|S) \geq f(C|T), \ \forall S \subseteq T \subseteq E, \ \text{with} \ C \subseteq E \setminus T \]  
\[ f(j|S) \geq f(j|S \cup \{k\}), \ \forall S \subseteq E \ \text{with} \ j \in E \setminus (S \cup \{k\}) \]  
\[ f(T) \leq f(S) + \sum_{j \in T \setminus S} f(j|S) - \sum_{j \in S \setminus T} f(j|S \cup T - \{j\}), \ \forall S, T \subseteq E \]  
\[ f(T) \leq f(S) + \sum_{j \in T \setminus S} f(j|S), \ \forall S \subseteq T \subseteq E \]  
\[ f(T) \leq f(S) - \sum_{j \in S \setminus T} f(j|S \setminus \{j\}) + \sum_{j \in T \setminus S} f(j|S \cap T) \ \forall S, T \subseteq E \]  
\[ f(T) \leq f(S) - \sum_{j \in S \setminus T} f(j|S \setminus \{j\}), \ \forall T \subseteq S \subseteq E \]  

We’ve already seen that Eq. 3.105 \( \equiv \) Eq. 3.106 \( \equiv \) Eq. 3.107 \( \equiv \) Eq. 3.108. We next show that Eq. 3.108 \( \Rightarrow \) Eq. 3.109 \( \Rightarrow \) Eq. 3.110 \( \Rightarrow \) Eq. 3.108.
Let $T \setminus S = \{j_1, \ldots, j_r\}$ and $S \setminus T = \{k_1, \ldots, k_q\}$.

First, we upper bound the gain of $T$ in the context of $S$:

$$f(S \cup T) - f(S) = \sum_{t=1}^{r} \left( f(S \cup \{j_1, \ldots, j_t\}) - f(S \cup \{j_1, \ldots, j_{t-1}\}) \right)$$

(3.113)

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

(3.114)

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

(3.115)

Next, lower bound $S$ in the context of $T$:

$$f(S \cup T) - f(T) = \sum_{t=1}^{q} \left[ f(T \cup \{k_1, \ldots, k_t\}) - f(T \cup \{k_1, \ldots, k_{t-1}\}) \right]$$

(3.116)

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

(3.117)

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

(3.118)
Eq. 3.108 $\Rightarrow$ Eq. 3.109

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

$$f(T \mid S) = f(S \cup T) - f(S) \leq \sum_{j \in T \setminus S} f(j \mid S) \quad (3.119)$$

and the lower bound

$$f(S \mid T) = f(S \cup T) - f(T) \geq \sum_{j \in S \setminus T} f(j \mid S \cup T \setminus \{j\}) \quad (3.120)$$

This gives upper and lower bounds of the form

expression-1 $\leq f(S \cup T) \leq$ expression-2, \quad (3.121)

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

Eq. 3.109 $\Rightarrow$ Eq. 3.110

This follows immediately since if $S \subseteq T$, then $S \setminus T = \emptyset$, and the last term of Eq. 3.109 vanishes.
Here, we set $T = S \cup \{j, k\}, j \notin S \cup \{k\}$ into Eq. 3.110 to obtain

$$f(S \cup \{j, k\}) \leq f(S) + f(j|S) + f(k|S) \quad (3.122)$$

$$= f(S) + f(S + \{j\}) - f(S) + f(S + \{k\}) - f(S) \quad (3.123)$$

$$= f(S + \{j\}) + f(S + \{k\}) - f(S) \quad (3.124)$$

$$= f(j|S) + f(S + \{k\}) \quad (3.125)$$

giving

$$f(j|S \cup \{k\}) = f(S \cup \{j, k\}) - f(S \cup \{k\}) \quad (3.126)$$

$$\leq f(j|S) \quad (3.127)$$

**Example: Rank function of a matrix**

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

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

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

- Let \( \text{rank} : 2^V \to \mathbb{Z}_+ \) be the rank function.
- In general, \( \text{rank}(A) \leq |A| \), and vectors in \( A \) are linearly independent if and only if \( \text{rank}(A) = |A| \).
- If \( A, B \) are such that \( \text{rank}(A) = |A| \) and \( \text{rank}(B) = |B| \), with \( |A| < |B| \), then the space spanned by \( B \) is greater, and we can find a vector in \( B \) that is linearly independent of the space spanned by vectors in \( A \).
- To stress a point, note that the above condition is \( |A| < |B| \), not \( A \subseteq B \) which is sufficient but not necessary.
- In other words, given \( A, B \) with \( \text{rank}(A) = |A| \) & \( \text{rank}(B) = B \), then \( |A| < |B| \iff \exists \ b \in B \) such that \( \text{rank}(A \cup \{b\}) = |A| + 1 \).

Spanning trees/forests

- We are given a graph \( G = (V, E) \), and consider the edges \( E = \mathcal{E}(G) \) as an index set.
- Consider the \( |V| \times |E| \) incidence matrix of undirected graph \( G \), which is the matrix \( X_G = (x_{v,e})_{v \in \mathcal{V}(G), e \in \mathcal{E}(G)} \) where
  \[
  x_{v,e} = \begin{cases} 
  1 & \text{if } v \in e \\
  0 & \text{if } v \notin e
  \end{cases}
  \]
Spanning trees/forests & incidence matrices

- We are given a graph $G = (V, E)$, we can orient the graph (make it directed) consider again the edges $E = E(G)$ as an index set.

- Consider instead the $|V| \times |E|$ incidence matrix of undirected graph $G$, which is the matrix $X_G = (x_{v,e})_{v \in V(G), e \in E(G)}$ where

$$x_{v,e} = \begin{cases} 
1 & \text{if } v \in e^+ \\
-1 & \text{if } v \in e^- \\
0 & \text{if } v \notin e 
\end{cases} \quad (3.130)$$

and where $e^+$ is the tail and $e^-$ is the head of edge $e$. 

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

(3.131)
Spanning trees

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

\[
\begin{pmatrix}
1 \\
2 \\
3 \\
4 \\
5 \\
6 \\
7 \\
8 \\
\end{pmatrix}
\begin{pmatrix}
-1 \\
1 \\
0 \\
0 \\
0 \\
0 \\
0 \\
0 \\
\end{pmatrix}
\]

(3.132)

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

Spanning trees

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

\[
\begin{pmatrix}
1 \\
2 \\
3 \\
4 \\
5 \\
6 \\
7 \\
8 \\
\end{pmatrix}
\begin{pmatrix}
1 & 2 \\
1 & 1 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 0 \\
\end{pmatrix}
\]

(3.132)

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

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

\[
\begin{bmatrix}
1 & 2 & 3 \\
1 & -1 & 1 & 0 \\
2 & 1 & 0 & -1 \\
3 & 0 & -1 & 0 \\
4 & 0 & 0 & 1 \\
5 & 0 & 0 & 0 \\
6 & 0 & 0 & 0 \\
7 & 0 & 0 & 0 \\
8 & 0 & 0 & 0 \\
\end{bmatrix}
\]  
(3.132)

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

Spanning trees

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

\[
\begin{bmatrix}
1 & 2 & 3 & 5 \\
1 & -1 & 1 & 0 & 0 \\
2 & 1 & 0 & -1 & 1 \\
3 & 0 & -1 & 0 & 0 \\
4 & 0 & 0 & 1 & 0 \\
5 & 0 & 0 & 0 & 0 \\
6 & 0 & 0 & 0 & 0 \\
7 & 0 & 0 & 0 & -1 \\
8 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\]  
(3.132)

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

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

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

Spanning trees

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

Here, \( \text{rank}\left(\{x_1, x_2, x_3, x_4\}\right) = 3 \) since \( x_4 = -x_1 - x_2 - x_3 \).
Spanning trees

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

Spanning Tree Algorithms

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

Algorithm 1: Borůvka’s Algorithm

1. $F \leftarrow \emptyset$ /* We build up the edges of a forest in $F$ */
2. while $G(V, F)$ is disconnected do
3.  forall the components $C_i$ of $F$ do
4.  $F \leftarrow F \cup \{e_i\}$ for $e_i$ = the min-weight edge out of $C_i$;
Spanning Tree Algorithms

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

- Given a tree \( T \), the cost of the tree is \( \text{cost}(T) = \sum_{e \in T} w(e) \), the sum of the weights of the edges.

- There are several algorithms for this.

**Algorithm 2: Jarník/Prim/Dijkstra Algorithm**

1. \( T \leftarrow \emptyset \);
2. while \( T \) is not a spanning tree do
3. \( T \leftarrow T \cup \{e\} \) for \( e = \) the minimum weight edge extending the tree \( T \) to a new vertex ;

**Algorithm 3: Kruskal’s Algorithm**

1. Sort the edges so that \( w(e_1) \leq w(e_2) \leq \cdots \leq w(e_m) \);
2. \( T \leftarrow (V(G), \emptyset) = (V, E) \);
3. for \( i = 1 \) to \( m \) do
4. if \( E(T) \cup \{e_i\} \) does not create a cycle in \( T \) then
5. \( E(T) \leftarrow E(T) \cup \{e\} \);
Spanning Tree Algorithms

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

Matroid

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

**Definition 3.6.1 (set system)**

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

- Set systems can be arbitrarily complex since, as stated, there is no method to determine if a given set $S \subseteq E$ has $S \in \mathcal{I}$.
- One useful property is “heredity.” Namely, a set system is a hereditary set system if for any $A \subset B \in \mathcal{I}$, we have that $A \in \mathcal{I}$.

**Definition 3.6.2 (independence system)**

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

\[
\emptyset \in \mathcal{I} \quad (I1)
\]

and

\[
\forall I \in \mathcal{I}, J \subset I \Rightarrow J \in \mathcal{I} \quad (I2)
\]

- Property I2 is called “down monotone” or “down closed”.
- Example: $E = \{1, 2, 3, 4\}$. With $\mathcal{I} = \{\emptyset, \{1\}, \{1, 2\}, \{1, 2, 4\}\}$.
- Then $(E, \mathcal{I})$ is a set system, but not an independence system since it is not down closed (i.e., we have $\{1, 2\} \in \mathcal{I}$ but not $\{2\} \in \mathcal{I}$).
- With $\mathcal{I} = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}$, then $(E, \mathcal{I})$ is now an independence (hereditary) system.
Given any set of linearly independent vectors $A$, any subset $B \subset A$ will also be linearly independent.

Given any forest $G_f$ that is an edge-induced sub-graph of a graph $G$, any sub-graph of $G_f$ is also a forest.

So these both constitute independence systems.

### Matroid

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

**Definition 3.6.3 (Matroid)**

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

1. $\emptyset \in \mathcal{I}$
2. $\forall I \in \mathcal{I}, J \subset I \Rightarrow J \in \mathcal{I}$
3. $\forall I, J \in \mathcal{I}$, with $|I| = |J| + 1$, then there exists $x \in I \setminus J$ such that $J \cup \{x\} \in \mathcal{I}$. 
Matroid

Slight modification (non unit increment) that is equivalent.

**Definition 3.6.4 (Matroid-II)**

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

1. \(\emptyset \in \mathcal{I}\)  \((I1')\)
2. \(\forall I \in \mathcal{I}, J \subset I \Rightarrow J \in \mathcal{I}\) (or “down-closed”)  \((I2')\)
3. \(\forall I, J \in \mathcal{I}, \text{with } |I| > |J|, \text{ then there exists } x \in I \setminus J \text{ such that } J \cup \{x\} \in \mathcal{I}\)  \((I3')\)

Note \((I1)\equiv(I1'), (I2)\equiv(I2'), \text{ and we get } (I3)\equiv(I3') \text{ using induction.}\n
Matroids, independent sets, and bases

- **Independent sets**: Given a matroid \(M = (E, \mathcal{I})\), a subset \(A \subseteq E\) is called **independent** if \(A \in \mathcal{I}\) and otherwise \(A\) is called **dependent**.

- **A base of \(U \subseteq E\)**: For \(U \subseteq E\), a subset \(B \subseteq U\) is called a **base** of \(U\) if \(B\) is inclusionwise maximally independent subset of \(U\). That is, \(B \in \mathcal{I}\) and there is no \(Z \in \mathcal{I}\) with \(B \subset Z \subseteq U\).

- **A base of a matroid**: If \(U = E\), then a “base of \(E\)” is just called a base of the matroid \(M\) (this corresponds to a basis in a linear space).
Sources for Today’s Lecture


End