Logistics Review Towards Submodular Function Minimization (SFM) On Polymatroids Lovász extension Scratch Summary

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

Prof. Jeff Bilmes

University of Washington, Seattle
Department of Electrical Engineering
Spring Quarter, 2011
http://ssli.ee.washington.edu/~bilmes/ee595a_spring_2011/

Lecture 11 - May 6th, 2011
Announcements

- On Final projects. **One** single page final project proposals (revision one) are due today at 6:00pm.

- Again, all submissions must be done electronically, via our drop box. See the link https://catalyst.uw.edu/collectit/dropbox/bilmes/14888, or look at the homework on the web page.

- Email me and/or stop by office hours for ideas. The proposals next Friday are non-binding (you can change your mind later) but you should start thinking about project proposals now.

- Ideal proposal would, say, lead to a NIPS paper in June and be related to submodularity.
We need to find one makeup lectures this term.

- L1 (3/30):
- L2 (4/1):
- L3 (4/6):
- L4 (4/8):
- L5 (4/13):
- L6 (4/15):
- L7 (4/20):
- L8 (4/27):
- L9 (4/29):
- L10 (5/4):
- L12 (5/11):
- L13 (5/13):
- L14 (5/18):
- L15 (5/20):
- L16 (5/25):
- L17 (5/27):
- L18 (6/1):
- L19 (6/3):
- L20: (6/?): (need to find time/date/place).
Consider

\[ P_r = \left\{ x \in \mathbb{R}^E : x \geq 0, x(A) \leq r_M(A), \forall A \subseteq E \right\} \quad (1) \]

We saw before that \( P_r = P_{\text{ind. set}} \).

Suppose we have any \( x \in \mathbb{R}^E_+ \) such that \( x \not\in P_r \).

The most violated inequality when \( x \) is considered w.r.t. \( P_r \) corresponds to the set \( A \) that maximizes \( x(A) - r_M(A) \), i.e.,

\[ \max \{ x(A) - r_M(A) : A \subseteq E \} \]

This corresponds to \( \min \{ r_M(A) + x(E \setminus A) : A \subseteq E \} \) since \( x \) is modular and \( x(E \setminus A) = x(E) - x(A) \).

More importantly, \( \min \{ r_M(A) + x(E \setminus A) : A \subseteq E \} \) a form of submodular function minimization, namely \( \min \{ r_M(A) - x(A) : A \subseteq E \} \) for a submodular function consisting of a difference of matroid rank and modular (so no longer nec. monotone, nor positive).
Augmenting path theorem

- Thus, we consider \( x \in \mathbb{R}_+ \).
Augmenting path theorem

- Thus, we consider $x \in \mathbb{R}_+$.  

- We’ve constructed the auxiliary $s, t$ graph $G$ as previously mentioned, where for each $e \in E$, we’ve got a node in $G$, along with additional nodes (and edges) $s, t$.  

Augmenting path theorem

- Thus, we consider $x \in \mathbb{R}_+$.
- We’ve constructed the auxiliary $s, t$ graph $G$ as previously mentioned, where for each $e \in E$, we’ve got a node in $G$, along with additional nodes (and edges) $s, t$.
- We maintain $y = \sum_{i \in J} \lambda_i 1_{i} \leq x$ and thus $y \in P_{\text{ind. set}}$. 
Augmenting path theorem

- Thus, we consider $x \in \mathbb{R}_+$.  

- We’ve constructed the auxiliary $s, t$ graph $G$ as previously mentioned, where for each $e \in E$, we’ve got a node in $G$, along with additional nodes (and edges) $s, t$.

- We maintain $y = \sum_{i \in J} \lambda_i 1_{I_i} \leq x$ and thus $y \in P_{\text{ind. set}}$.

- From this, we can obtain the following theorem (most violated inequality, then, is given by $\{ e \in E : x(e) > y(e) \}$).

```latex
\begin{align*}
\text{Theorem 2.1} \\
\text{If there is a directed path from } s \text{ to } t \text{ in } G, \text{ then there exists } y' \in P \text{ with } y < y' \leq x, \text{ with } y'(E) > y(E).
\end{align*}
```
Augmenting path theorem

- Thus, we consider $x \in \mathbb{R}_+$.
- We’ve constructed the auxiliary $s, t$ graph $G$ as previously mentioned, where for each $e \in E$, we’ve got a node in $G$, along with additional nodes (and edges) $s, t$.
- We maintain $y = \sum_{i \in J} \lambda_i 1_{I_i} \leq x$ and thus $y \in P_{\text{ind. set}}$.
- From this, we can obtain the following theorem (most violated inequality, then, is given by $\{e \in E : x(e) > y(e)\}$).

**Theorem 2.1**

*If there is a directed path from $s$ to $t$ in $G$, then there exists $y' \in P$ with $y < y' \leq x$, with $y'(E) > y(E)$. If there is no such path, then there exists a set $A \subseteq E$ s.t. $y(A) = r(A)$ and $y(E \setminus A) = x(E \setminus A)$.*
Augmenting path theorem consequences

**Corollary 2.2**

For any \( x \in \mathbb{R}^E_+ \), we have

\[
\max (y(E) : y \leq x, y \in P_r) = \min (x(A) + r(E \setminus A) : A \subseteq E) \tag{2}
\]

*Note: this was not used in the theorem above, rather it is a consequence!*

**Proof.**

1. First, as we’ve seen, any \( y \in P \) with \( y \leq x \), and any \( A \subseteq E \), we have

\[
y(E) = y(A) + y(E \setminus A) \leq r(A) + x(E \setminus A) \tag{3}
\]
Augmenting path theorem consequences

**Corollary 2.2**

For any $x \in \mathbb{R}^E$, we have

$$\max (y(E) : y \leq x, y \in P_r) = \min (x(A) + r(E \setminus A) : A \subseteq E)$$  \hspace{1cm} (2)

*Note: this was not used in the theorem above, rather it is a consequence!*

**Proof.**

1. First, as we’ve seen, any $y \in P$ with $y \leq x$, and any $A \subseteq E$, we have

   $$y(E) = y(A) + y(E \setminus A) \leq r(A) + x(E \setminus A)$$  \hspace{1cm} (3)

2. So we need only find a $y$ giving equality.
Augmenting path theorem consequences

Corollary 2.2

For any \( x \in \mathbb{R}^E_+ \), we have

\[
\max (y(E) : y \leq x, y \in P_r) = \min (x(A) + r(E \setminus A) : A \subseteq E) \quad (2)
\]

Note: this was not used in the theorem above, rather it is a consequence!

Proof.

1. First, as we've seen, any \( y \in P \) with \( y \leq x \), and any \( A \subset E \), we have

\[
y(E) = y(A) + y(E \setminus A) \leq r(A) + x(E \setminus A) \quad (3)
\]

2. So we need only find a \( y \) giving equality.

3. Choose any \( y \in P \) such that \( y \leq x \) and with \( y(E) \) maximum and run algorithm.
Augmenting path theorem consequences

**Corollary 2.2**

For any $x \in \mathbb{R}_+^E$, we have

$$\max (y(E) : y \leq x, y \in P_r) = \min (x(A) + r(E \setminus A) : A \subseteq E)$$  \hspace{1cm} (2)

*Note: this was not used in the theorem above, rather it is a consequence!*

**Proof.**

1. First, as we've seen, any $y \in P$ with $y \leq x$, and any $A \subseteq E$, we have

$$y(E) = y(A) + y(E \setminus A) \leq r(A) + x(E \setminus A)$$  \hspace{1cm} (3)

2. So we need only find a $y$ giving equality.

3. Choose any $y \in P$ such that $y \leq x$ and with $y(E)$ maximum and run algorithm.

4. Then eventually exists no such $y' \in P$ s.t. $y'(E) > y(E)$, and the digraph won't have a directed path from $s$ to $t$ (by the theorem).
**Augmenting path theorem consequences**

**Corollary 2.2**

For any \( x \in \mathbb{R}_+^E \), we have

\[
\max (y(E) : y \leq x, y \in P_r) = \min (x(A) + r(E \setminus A) : A \subseteq E) \quad (2)
\]

*Note: this was not used in the theorem above, rather it is a consequence!*

**Proof.**

1. First, as we've seen, any \( y \in P \) with \( y \leq x \), and any \( A \subseteq E \), we have

\[
y(E) = y(A) + y(E \setminus A) \leq r(A) + x(E \setminus A) \quad (3)
\]

2. So we need only find a \( y \) giving equality.

3. Choose any \( y \in P \) such that \( y \leq x \) and with \( y(E) \) maximum and run algorithm.

4. Then eventually exists no such \( y' \in P \) s.t. \( y'(E) > y(E) \), and the digraph won’t have a directed path from \( s \) to \( t \) (by the theorem).

5. Then, there is a set \( A \) such that \( y(A) = r(A) \) and \( y(E \setminus A) = x(E \setminus A) \), giving

\[
y(E) = r(A) + x(E \setminus A),
\]

thus demonstrating equality in Eq. 3, and minimality of \( r(A) + x(E \setminus A) \).
Augmenting path theorem consequences

Corollary 2.3

Given matroid $M$, we have

$$P_{\text{ind. set}} = P_r$$

(4)

We even get this a consequence!
Bounding the number of augmenting paths

Consider the algorithm implied by Theorem 4.1 as producing one augmentation, and let $G_i$ refer to the digraph at outer iteration $i$. Then we have
Bounding the number of augmenting paths

Consider the algorithm implied by Theorem 4.1 as producing one augmentation, and let $G_i$ refer to the digraph at outer iteration $i$. Then we have

**Theorem 2.4**

Let $G_0, G_1, \ldots, G_k$ be a sequence of digraphs, each having vertex set $E \cup \{s, t\}$, and correspond to such graphs each one running the algorithm implied by theorem 4.1 Assume fixed total order of $E \cup \{s\}$. Let $Q_i$ denote the CBFS path in $G_i$, for $0 \leq i < k$. If it is the case that, for $0 \leq i < k$:

1. There is an edge in $Q_i$ that is not an edge in $G_i + 1$,
2. If $(e, f)$ is an edge in $G_i + 1$ but not in $G_i$, then $e, f \in E$ and there are vertices $a, b \in Q_i$ with $a$ preceding $b$ on $Q_i$ such that: 1) either $a = f$ or $(a, f)$ is an edge in $G_i$; and 2) $b = e$ or $(e, b)$ is an edge in $G_i$,

Then we have that the number of augmentations has bound $k \leq |E|^3$. 
Bounding the number of augmenting paths

Consider the algorithm implied by Theorem 4.1 as producing one augmentation, and let $G_i$ refer to the digraph at outer iteration $i$. Then we have

**Theorem 2.4**

Let $G_0, G_1, \ldots, G_k$ be a sequence of digraphs, each having vertex set $E \cup \{s, t\}$, and correspond to such graphs each one running the algorithm implied by theorem 4.1 Assume fixed total order of $E \cup \{s\}$. Let $Q_i$ denote the CBFS path in $G_i$, for $0 \leq i < k$. If it is the case that, for $0 \leq i < k$:

1. There is an edge in $Q_i$ that is not an edge in $G_i + 1$,
2. If $(e, f)$ is an edge in $G_i + 1$ but not in $G_i$, then $e, f \in E$ and there are vertices $a, b \in Q_i$ with $a$ preceding $b$ on $Q_i$ such that: 1) either $a = f$ or $(a, f)$ is an edge in $G_i$; and 2) $b = e$ or $(e, b)$ is an edge in $G_i$,

Then we have that the number of augmentations has bound $k \leq |E|^3$. 

Prof. Jeff Bilmes
Bounding the number of augmenting paths

Consider the algorithm implied by Theorem 4.1 as producing one augmentation, and let $G_i$ refer to the digraph at outer iteration $i$. Then we have

**Theorem 2.4**

Let $G_0, G_1, \ldots, G_k$ be a sequence of digraphs, each having vertex set $E \cup \{s, t\}$, and correspond to such graphs each one running the algorithm implied by theorem 4.1 Assume fixed total order of $E \cup \{s\}$. Let $Q_i$ denote the CBFS path in $G_i$, for $0 \leq i < k$. If it is the case that, for $0 \leq i < k$:

1. There is an edge in $Q_i$ that is not an edge in $G_{i+1}$,
Bounding the number of augmenting paths

Consider the algorithm implied by Theorem 4.1 as producing one augmentation, and let $G_i$ refer to the digraph at outer iteration $i$. Then we have:

**Theorem 2.4**

Let $G_0, G_1, \ldots, G_k$ be a sequence of digraphs, each having vertex set $E \cup \{s, t\}$, and correspond to such graphs each one running the algorithm implied by theorem 4.1. Assume fixed total order of $E \cup \{s\}$. Let $Q_i$ denote the CBFS path in $G_i$, for $0 \leq i < k$. If it is the case that, for $0 \leq i < k$:

1. There is an edge in $Q_i$ that is not an edge in $G_{i+1}$,
2. If $(e, f)$ is an edge in $G_{i+1}$ but not in $G_i$, then $e, f \in E$ and there are vertices $a, b \in Q_i$ with a preceding $b$ on $Q_i$ such that: 1) either $a = f$ or $(a, f)$ is an edge in $G_i$; and 2) $b = e$ or $(e, b)$ is an edge in $G_i$,

Then we have that the number of augmentations has bound $k \leq |E|^3$. 

Prof. Jeff Bilmes EE595A/Spr 2011/Submodular Functions – Lecture 11 - May 6th, 2011
Bounding the number of augmenting paths

- Consider the algorithm implied by Theorem 4.1 as producing one augmentation, and let $G_i$ refer to the digraph at outer iteration $i$. Then we have

**Theorem 2.4**

Let $G_0, G_1, \ldots, G_k$ be a sequence of digraphs, each having vertex set $E \cup \{s, t\}$, and correspond to such graphs each one running the algorithm implied by theorem 4.1 Assume fixed total order of $E \cup \{s\}$. Let $Q_i$ denote the CBFS path in $G_i$, for $0 \leq i < k$. If it is the case that, for $0 \leq i < k$:

1. There is an edge in $Q_i$ that is not an edge in $G_{i+1}$,
2. If $(e, f)$ is an edge in $G_{i+1}$ but not in $G_i$, then $e, f \in E$ and there are vertices $a, b \in Q_i$ with a preceding $b$ on $Q_i$ such that: 1) either $a = f$ or $(a, f)$ is an edge in $G_i$; and 2) $b = e$ or $(e, b)$ is an edge in $G_i$,

Then we have that the number of augmentations has bound $k \leq |E|^3$. 
Theorem 2.5

It is possible to construct an augmentation scheme such that each augmenting path is done in accordance to Theorem 4.4. Each such augmentation is CBFS, and is called a “grand” augmentation, and is maximal in a certain way. This achieves the $O(n^3)$ time, in the number of augmentations, mentioned above.
Bounding the number of augmenting paths

Theorem 2.5

*It is possible to construct an augmentation scheme such that each augmenting path is done in accordance to Theorem 4.4. Each such augmentation is CBFS, and is called a “grand” augmentation, and is maximal in a certain way. This achieves the $O(n^3)$ time, in the number of augmentations, mentioned above.*

- Of course, the cost of each augmentation might be expensive. For matric matroids, each would be $O(r^2n^5)$ where $r$ is the number of rows of the matrix, leading to $O(r^2n^8)$ algorithm.
Bounding the number of augmenting paths

Theorem 2.5

It is possible to construct an augmentation scheme such that each augmenting path is done in accordance to Theorem 4.4. Each such augmentation is CBFS, and is called a “grand” augmentation, and is maximal in a certain way. This achieves the $O(n^3)$ time, in the number of augmentations, mentioned above.

- Of course, the cost of each augmentation might be expensive. For matric matroids, each would be $O(r^2n^5)$ where $r$ is the number of rows of the matrix, leading to $O(r^2n^8)$ algorithm.
- On the other hand, this algorithm has some intriguing properties.
Augmenting path theorem

**Theorem 3.1**

*If there is a directed path from s to t in G, then there exists $y' \in P$ with $y < y' \leq x$, with $y'(E) > y(E)$. If there is no such path, then there exists a set $A \subseteq E$ s.t. $y(A) = r(A)$ and $y(E \setminus A) = x(E \setminus A)$.*
Augmenting path theorem

**Theorem 3.1**

*If there is a directed path from s to t in G, then there exists \( y' \in P \) with \( y < y' \leq x \), with \( y'(E) > y(E) \). If there is no such path, then there exists a set \( A \subseteq E \) s.t. \( y(A) = r(A) \) and \( y(E \setminus A) = x(E \setminus A) \).*

- Recall, we are given \( x \in \mathcal{R}_+^E \). Algorithm implied by this theorem is called multiple times, setting \( y \leftarrow y' \), until no such path exists at which point we get said \( A \) and \( y \) s.t. \( y \leq x \) and \( y \) is otherwise maximal in \( P \).
Augmenting path theorem

Theorem 3.1

If there is a directed path from s to t in G, then there exists \( y' \in P \) with \( y < y' \leq x \), with \( y'(E) > y(E) \). If there is no such path, then there exists a set \( A \subseteq E \) s.t. \( y(A) = r(A) \) and \( y(E \setminus A) = x(E \setminus A) \).

- Recall, we are given \( x \in \mathcal{R}^E_+ \). Algorithm implied by this theorem is called multiple times, setting \( y \leftarrow y' \), until no such path exists at which point we get said \( A \) and \( y \) s.t. \( y \leq x \) and \( y \) is otherwise maximal in \( P \).
- This solves \( \arg\min_{A \subseteq E} (r(A) - x(A)) \) as seen a few slides back.
Augmenting path theorem

Theorem 3.1

If there is a directed path from $s$ to $t$ in $G$, then there exists $y' \in P$ with $y < y' \leq x$, with $y'(E) > y(E)$. If there is no such path, then there exists a set $A \subseteq E$ s.t. $y(A) = r(A)$ and $y(E \setminus A) = x(E \setminus A)$.

- Recall, we are given $x \in \mathcal{R}_+^E$. Algorithm implied by this theorem is called multiple times, setting $y \leftarrow y'$, until no such path exists at which point we get said $A$ and $y$ s.t. $y \leq x$ and $y$ is otherwise maximal in $P$.
- This solves $\text{argmin}_{A \subseteq E} (r(A) - x(A))$ as seen a few slides back.
- Suppose $x \in P$. Then, $\forall A$, $r(A) \geq x(A)$ so minimizing $r(A) - x(A)$ requires $A = \emptyset$. This then gives $y = x$ (no inequality is violated, and a certificate for $x \in P$).
Augmenting path theorem

Theorem 3.1

If there is a directed path from s to t in G, then there exists \( y' \in P \) with \( y < y' \leq x \), with \( y'(E) > y(E) \). If there is no such path, then there exists a set \( A \subseteq E \) s.t. \( y(A) = r(A) \) and \( y(E \setminus A) = x(E \setminus A) \).

- Recall, we are given \( x \in \mathcal{R}_+^E \). Algorithm implied by this theorem is called multiple times, setting \( y \leftarrow y' \), until no such path exists at which point we get said \( A \) and \( y \) s.t. \( y \leq x \) and \( y \) is otherwise maximal in \( P \).
- This solves \( \arg\min_{A \subseteq E} (r(A) - x(A)) \) as seen a few slides back.
- Suppose \( x \in P \). Then, \( \forall A, r(A) \geq x(A) \) so minimizing \( r(A) - x(A) \) requires \( A = \emptyset \). This then gives \( y = x \) (no inequality is violated, and a certificate for \( x \in P \)).
- If \( x \notin P \), minimizing \( r(A) - x(A) \) gives an \( A \) so that gives the inequality, of the form \( x(A) \leq r(A) \) that is most violated and \( E \setminus A = \{ e \in E : x(e) > y(e) \} \).
[Edmonds]

“But now, you know, this is my day in the sun.” - from A Glimpse of Heaven, 1991.
Towards SFM

- Recall the Edmonds matroid partition algorithm, was SFM for $r(A) - \frac{1}{k} 1(A)$.
Towards SFM

- Recall the Edmonds matroid partition algorithm, was SFM for \( r(A) - \frac{1}{k} 1(A) \).
- We now have an algorithm that can do SFM on \( r(A) - x(A) \) for any \( x \in \mathbb{R}^E_+ \) and any matroid rank function.
Towards SFM

- Recall the Edmonds matroid partition algorithm, was SFM for $r(A) - \frac{1}{k} \mathbf{1}(A)$.
- We now have an algorithm that can do SFM on $r(A) - x(A)$ for any $x \in \mathbb{R}^E_+$ and any matroid rank function.
- There are three limitations to this:
Towards SFM

- Recall the Edmonds matroid partition algorithm, was SFM for $r(A) - \frac{1}{k} 1(A)$.
- We now have an algorithm that can do SFM on $r(A) - x(A)$ for any $x \in \mathbb{R}^E_+$ and any matroid rank function.

- There are three limitations to this:
  1. $r(A)$ is only a matroid rank function (and thus integral) rather than a (possibly non-integral) polymatroidal function.
Towards SFM

- Recall the Edmonds matroid partition algorithm, was SFM for \( r(A) - \frac{1}{k}1(A) \).
- We now have an algorithm that can do SFM on \( r(A) - x(A) \) for any \( x \in \mathbb{R}_+^E \) and any matroid rank function.
- There are three limitations to this:
  1. \( r(A) \) is only a matroid rank function (and thus integral) rather than a (possibly non-integral) polymatroidal function.
  2. \( x \) is required to be positive \( x \geq 0 \).
Towards SFM

- Recall the Edmonds matroid partition algorithm, was SFM for $r(A) - \frac{1}{k}1(A)$.

- We now have an algorithm that can do SFM on $r(A) - x(A)$ for any $x \in \mathbb{R}_+^E$ and any matroid rank function.

- There are three limitations to this:
  1. $r(A)$ is only a matroid rank function (and thus integral) rather than a (possibly non-integral) polymatroidal function.
  2. $x$ is required to be positive $x \geq 0$.
  3. This works only for the difference between $r$ and $x$, but we’d like an algorithm that works for any arbitrary submodular function $f$, even non-monotone and/or non-non-increasing/decreasing.
Towards SFM

- Recall the Edmonds matroid partition algorithm, was SFM for $r(A) - \frac{1}{k} \mathbf{1}(A)$.
- We now have an algorithm that can do SFM on $r(A) - x(A)$ for any $x \in \mathbb{R}^E_+$ and any matroid rank function.
- There are three limitations to this:
  1. $r(A)$ is only a matroid rank function (and thus integral) rather than a (possibly non-integral) polymatroidal function.
  2. $x$ is required to be positive $x \geq 0$.
  3. This works only for the difference between $r$ and $x$, but we’d like an algorithm that works for any arbitrary submodular function $f$, even non-monotone and/or non-non-increasing/decreasing.
- It turns out that (2) and (3) is easy to deal with, but (1) took another 16 years to solve (and perhaps can still be seen as unsolved, w.r.t. wanting a scalable algorithm).
First, given any submodular function $g$, construct modular function $m : E \rightarrow \mathbb{R}$ such that $m(e) = g(E \setminus \{e\}) - g(E)$.

\[
= - \left[ g(E) - g(E \cup e) \right]
\]

\[
= - \left[ \text{gain of adding } e \text{ to } E \cup e \right]
\]

\[
= - \left[ \text{smallest possible addition rule at } e \text{ in some context} \right].
\]
First, given any submodular function $g$, construct modular function $m : E \rightarrow \mathbb{R}$ such that $m(e) = g(E \setminus \{e\}) - g(E)$.

Then construct a new function $f : \mathcal{P}(E) \rightarrow \mathbb{R}$ as

$$f(A) = g(A) + m(A) - g(\emptyset)$$

(5)
Addressing Monotonicity

• First, given any submodular function $g$, construct modular function $m : E \rightarrow \mathbb{R}$ such that $m(e) = g(E \setminus \{e\}) - g(E)$.

• Then construct a new function $f : 2^E \mathbb{R}_+$ as

\[
  f(A) = g(A) + m(A) - g(\emptyset)
\]  

(5)

• Then $f(\emptyset) = 0$, so $f$ is normalized.
Addressing Monotonicity

- First, given any submodular function $g$, construct modular function $m : E \rightarrow \mathbb{R}$ such that $m(e) = g(E \setminus \{e\}) - g(E)$.
- Then construct a new function $f : 2^E \mathbb{R}_+$ as

\[ f(A) = g(A) + m(A) - g(\emptyset) \tag{5} \]

- Then $f(\emptyset) = 0$, so $f$ is normalized.
- Also, $f$ is monotone non-decreasing (and thus non-negative) and submodular. It is submodular since sum of submodular and modular. Monotone non-decreasing follows since for $v \in B$

\[ f(B + v) - f(B) = g(B + v) - g(B) + m(v) \tag{6} \]

\[ = g(B + v) - g(B) + g(E - v) - g(E) \tag{7} \]

\[ \geq 0 \tag{8} \]

since, by submodularity, $g(B + v) - g(B) \geq g(E) - g(E - v)$.
Also, if we wish to minimize $g$, then given
\[ f(A) = g(A) + m(A) - g(\emptyset), \]
we can just minimize $f - m$ since $g(\emptyset)$ is a constant.
\[ f - m = g - g(\emptyset) \]
Addressing Monotonicity

- Also, if we wish to minimize $g$, then given $f(A) = g(A) + m(A) - g(\emptyset)$, we can just minimize $f - m$ since $g(\emptyset)$ is a constant.

- So now we have a difference of a polymatroid function $f$ and a modular function $m$. This deals with (3) above.
Addressing Monotonicity

- Also, if we wish to minimize $g$, then given $f(A) = g(A) + m(A) - g(\emptyset)$, we can just minimize $f - m$ since $g(\emptyset)$ is a constant.

- So now we have a difference of a polymatroid function $f$ and a modular function $m$. This deals with (3) above.

- Is $m \in \mathbb{R}^E_+$?
Dealing with $m \in \mathbb{R}_+^E$

- So now we reduced the problem of SFM to that of minimizing a difference between a polymatroid function $f$ and a modular function $m$ (i.e., $\min_{A \subseteq E} f(A) - m(A)$).
Dealing with $m \in \mathbb{R}_+^E$

- So now we reduced the problem of SFM to that of minimizing a difference between a polymatroid function $f$ and a modular function $m$ (i.e., $\min_{A \subseteq E} f(A) - m(A)$).

- Is $m \in \mathbb{R}_+^E$? $m$ is given as clearly not positive.
Dealing with \( m \in \mathbb{R}^E_+ \)

- So now we reduced the problem of SFM to that of minimizing a difference between a polymatroid function \( f \) and a modular function \( m \) (i.e., \( \min_{A \subseteq E} f(A) - m(A) \)).

- Is \( m \in \mathbb{R}^E_+ \)?

No, but for any \( e \) such that \( m(e) < 0 \), \( e \) can’t be a minimizer of \( f - m \) since, assuming that \( A \) minimizes \( f(A) - m(A) \) and \( e \in A \) is such that \( m(e) < 0 \), then we have that

\[
\begin{align*}
f(A') - m(A') &< f(A) - m(A) \quad \text{where} \quad A' = A \setminus \{e\}. \\
f(A') &\leq f(A) \\
-m(A') &\leq -m(A)
\end{align*}
\]
Dealing with $m \in \mathbb{R}^E_+$

- So now we reduced the problem of SFM to that of minimizing a difference between a polymatroid function $f$ and a modular function $m$ (i.e., $\min_{A \subseteq E} f(A) - m(A)$).

- Is $m \in \mathbb{R}^E_+$?

- No, but for any $e$ such that $m(e) < 0$, $e$ can’t be a minimizer of $f - m$ since, assuming that $A$ minimizes $f(A) - m(A)$ and $e \in A$ is such that $m(e) < 0$, then we have that $f(A') - m(A') < f(A) - m(A)$ where $A' = A \setminus \{e\}$.

- This follows since $f$ is monotone non-decreasing, and $m(A) = m(A') + m(e)$, so $m(A') > m(A)$. 

---

Prof. Jeff Bilmes

EE595A/Spr 2011/Submodular Functions – Lecture 11 - May 6th, 2011
Dealing with $m \in \mathbb{R}^E_+$

- So now we reduced the problem of SFM to that of minimizing a difference between a polymatroid function $f$ and a modular function $m$ (i.e., $\min_{A \subseteq E} f(A) - m(A)$).
- Is $m \in \mathbb{R}^E_+$?
- No, but for any $e$ such that $m(e) < 0$, $e$ can’t be a minimizer of $f - m$ since, assuming that $A$ minimizes $f(A) - m(A)$ and $e \in A$ is such that $m(e) < 0$, then we have that $f(A') - m(A') < f(A) - m(A)$ where $A' = A \setminus \{e\}$.
- This follows since $f$ is monotone non-decreasing, and $m(A) = m(A') + m(e)$, so $m(A') > m(A)$.
- So we “throw away” any $e$ s.t. $m(e) < 0$. This deals with (2) above.

$E' = E \setminus M \quad M = \{e : m(e) < 0\}$

$\Rightarrow f' : 2^{E'} \rightarrow \mathbb{R}, \quad f'(A) = f(M) + f(A \cap M^c) \quad \forall A \subseteq E'$. 
Dealing with $m \in \mathbb{R}^+_E$

- So now we reduced the problem of SFM to that of minimizing a difference between a polymatroid function $f$ and a modular function $m$ (i.e., $\min_{A \subseteq E} f(A) - m(A)$).
- Is $m \in \mathbb{R}^+_E$?
- No, but for any $e$ such that $m(e) < 0$, $e$ can’t be a minimizer of $f - m$ since, assuming that $A$ minimizes $f(A) - m(A)$ and $e \in A$ is such that $m(e) < 0$, then we have that $f(A') - m(A') < f(A) - m(A)$ where $A' = A \setminus \{e\}$.
- This follows since $f$ is monotone non-decreasing, and $m(A) = m(A') + m(e)$, so $m(A') > m(A)$.
- So we “throw away” any $e$ s.t. $m(e) < 0$. This deals with (2) above.
- Therefore, SFM is as “easy” as moving from matroid rank functions to not-necessarily-integral polymatroidal functions.
Steven Smale in 2000 listed the following as one of the great unsolved problems for the next century: “Is there a polynomial time algorithm over the real numbers which decides the feasibility of the linear system of inequalities $Ax \geq b$?"
Testing membership in polymatroids

- Steven Smale in 2000 listed the following as one of the great unsolved problems for the next century: “Is there a polynomial time algorithm over the real numbers which decides the feasibility of the linear system of inequalities $Ax \geq b$?”

- Given submodular function, consider $P_f = \{x \in \mathbb{R}_+^E : x(A) \leq f(A), \ \forall A \subseteq E\}$. 
Testing membership in polymatroids

- Steven Smale in 2000 listed the following as one of the great unsolved problems for the next century: “Is there a polynomial time algorithm over the real numbers which decides the feasibility of the linear system of inequalities $Ax \geq b$?”

- Given submodular function, consider $P_f = \{ x \in \mathbb{R}_+^E : x(A) \leq f(A), \forall A \subseteq E \}$.

- The membership problem, “given an $x \in \mathbb{R}_+^E$, is $x \in P_f$?”, is a special case of this unsolved problem.
Steven Smale in 2000 listed the following as one of the great unsolved problems for the next century: “Is there a polynomial time algorithm over the real numbers which decides the feasibility of the linear system of inequalities $Ax \geq b$?"

Given submodular function, consider

$$P_f = \{x \in \mathbb{R}_+^E : x(A) \leq f(A), \forall A \subseteq E\}.$$ 

The membership problem, “given an $x \in \mathbb{R}_+^E$, is $x \in P_f$?”, is a special case of this unsolved problem.

This is true iff $0 \leq f(A) - x(A), \forall A \subseteq E.$
Steven Smale in 2000 listed the following as one of the great unsolved problems for the next century: “Is there a polynomial time algorithm over the real numbers which decides the feasibility of the linear system of inequalities $Ax \geq b$?”

Given submodular function, consider

$$P_f = \{ x \in \mathbb{R}^E_+ : x(A) \leq f(A), \forall A \subseteq E \}.$$

The membership problem, “given an $x \in \mathbb{R}^E_+$, is $x \in P_f$?”, is a special case of this unsolved problem.

This is true iff $0 \leq f(A) - x(A), \forall A \subseteq E$.

And this is true iff $\min f(A) - x(A) \geq 0$. 
Testing membership in polymatroids

- Steven Smale in 2000 listed the following as one of the great unsolved problems for the next century: “Is there a polynomial time algorithm over the real numbers which decides the feasibility of the linear system of inequalities $Ax \geq b$?”

- Given submodular function, consider $P_f = \{ x \in \mathbb{R}_+^E : x(A) \leq f(A), \ \forall A \subseteq E \}$.

- The membership problem, “given an $x \in \mathbb{R}_+^E$, is $x \in P_f$?”, is a special case of this unsolved problem.

- This is true iff $0 \leq f(A) - x(A), \ \forall A \subseteq E$.

- And this is true iff $\min f(A) - x(A) \geq 0$.

- So, given a strongly polynomial time algorithm for general submodular function, we can test polyhedral membership, in at least this limited (polymatroidal polytope) sense.
Polymatroidal polyhedron (or a “polymatroid”)

Recall from Lecture 7:

**Definition 4.1 (polymatroid)**

A **polymatroid** is a compact set $P \subseteq \mathbb{R}_+^E$ satisfying

1. $0 \in P$
2. If $y \leq x \in P$ then $y \in P$ (called **down monotone**).
3. For any $x \in \mathbb{R}_+^E$, any maximal vector $y \in P$ with $y \leq x$ (called a $P$-basis of $x$), has the same component sum $y(E)$. That is for any two maximal vectors $y^1, y^2 \in P$, we have $y^1(E) = y^2(E)$.

- A **polymatroid** is a compact set that is zero containing, down monotone, and any maximal vector $y$ in $P$, bounded by another vector $x$, has the same vector rank.
- A **matroid** a set system that is empty-set containing, down closed, and any maximal set $I$ in $\mathcal{I}$, bounded by another set $A$, has the same matroid rank.
Recall greedy algorithm (from Lec 5): Set $A = \emptyset$, and repeatedly choose $y \in E \setminus A$ such that $A \cup \{y\} \in \mathcal{I}$ with $w(y)$ as large as possible, stopping when no such $y$ exists.
Recall greedy algorithm (from Lec 5): Set $A = \emptyset$, and repeatedly choose $y \in E \setminus A$ such that $A \cup \{y\} \in \mathcal{I}$ with $w(y)$ as large as possible, stopping when no such $y$ exists.

For a matroid, we saw (Lec 5) that set system $(E, \mathcal{I})$ is a matroid iff for each weight function $w \in \mathcal{R}_+^E$, the greedy algorithm leads to a set $I \in \mathcal{I}$ of maximum weight $w(I)$.  

Can we characterize a polymatroid in this way? That is, if we consider $\max_{x \in \mathcal{P}_f} w(x)$, where $\mathcal{P}_f$ represents the "independent vectors", is it the case that $\mathcal{P}_f$ is a polymatroid iff greedy works for this maximization? Can we even relax things so that $w \in \mathcal{R}^E$?
Polymatroidal polyhedron and greedy

- Recall greedy algorithm (from Lec 5): Set \( A = \emptyset \), and repeatedly choose \( y \in E \setminus A \) such that \( A \cup \{y\} \in \mathcal{I} \) with \( w(y) \) as large as possible, stopping when no such \( y \) exists.

- For a matroid, we saw (Lec5) that set system \((E, \mathcal{I})\) is a matroid iff for each weight function \( w \in \mathcal{R}^+_E \), the greedy algorithm leads to a set \( I \in \mathcal{I} \) of maximum weight \( w(I) \).

- Stated succinctly, considering \( \max w(I) : I \in \mathcal{I} \), then \((E, \mathcal{I})\) is a matroid iff greedy works for this maximization.
Recall greedy algorithm (from Lec 5): Set $A = \emptyset$, and repeatedly choose $y \in E \setminus A$ such that $A \cup \{y\} \in \mathcal{I}$ with $w(y)$ as large as possible, stopping when no such $y$ exists.

For a matroid, we saw (Lec5) that set system $(E, \mathcal{I})$ is a matroid iff for each weight function $w \in \mathcal{R}_E$, the greedy algorithm leads to a set $I \in \mathcal{I}$ of maximum weight $w(I)$.

Stated succinctly, considering $\max w(I) : I \in \mathcal{I}$, then $(E, \mathcal{I})$ is a matroid iff greedy works for this maximization.

Can we characterize a polymatroid in this way?
Polymatroidal polyhedron and greedy

- Recall greedy algorithm (from Lec 5): Set \( A = \emptyset \), and repeatedly choose \( y \in E \setminus A \) such that \( A \cup \{y\} \in \mathcal{I} \) with \( w(y) \) as large as possible, stopping when no such \( y \) exists.

- For a matroid, we saw (Lec5) that set system \((E, \mathcal{I})\) is a matroid iff for each weight function \( w \in \mathcal{R}^E_+ \), the greedy algorithm leads to a set \( I \in \mathcal{I} \) of maximum weight \( w(I) \).

- Stated succinctly, considering \( \max w(I) : I \in \mathcal{I} \), then \((E, \mathcal{I})\) is a matroid iff greedy works for this maximization.

- Can we characterize a polymatroid in this way?

- That is, if we consider \( \max wx : x \in P_f \), where \( P_f \) represent the "independent vectors", is it the case that \( P_f \) is a polymatroid iff greedy works for this maximization?
Recall greedy algorithm (from Lec 5): Set $A = \emptyset$, and repeatedly choose $y \in E \setminus A$ such that $A \cup \{y\} \in \mathcal{I}$ with $w(y)$ as large as possible, stopping when no such $y$ exists.

For a matroid, we saw (Lec5) that set system $(E, \mathcal{I})$ is a matroid iff for each weight function $w \in \mathbb{R}_+^E$, the greedy algorithm leads to a set $I \in \mathcal{I}$ of maximum weight $w(I)$.

Stated succinctly, considering $\max w(I) : I \in \mathcal{I}$, then $(E, \mathcal{I})$ is a matroid iff greedy works for this maximization.

Can we characterize a polymatroid in this way?

That is, if we consider $\max wx : x \in P_f$, where $P_f$ represent the “independent vectors”, is it the case that $P_f$ is a polymatroid iff greedy works for this maximization?

Can we even relax things so that $w \in \mathbb{R}^E$?
What is the greedy solution in this setting?
Polymatroidal polyhedron and greedy

- What is the greedy solution in this setting?
- Sort elements of $E$ w.r.t. $w$ so that, w.l.o.g.
  $$E = (e_1, e_2, \ldots, e_m) \text{ with } w(e_1) \geq w(e_2) \geq \cdots \geq w(e_m).$$
Polymatroidal polyhedron and greedy

- What is the greedy solution in this setting?
- Sort elements of $E$ w.r.t. $w$ so that, w.l.o.g.
  \[ E = (e_1, e_2, \ldots, e_m) \text{ with } w(e_1) \geq w(e_2) \geq \cdots \geq w(e_m). \]
- Let $k + 1$ be the first point (if any) at which we are non-positive, i.e., $w(e_k) > 0$ and $0 \geq w(e_{k+1})$.

\[ w(e_1) \geq w(e_2) \geq \cdots \geq w(e_k) > 0 \geq w(e_{k+1}) > \cdots > w(e_m) \]
What is the greedy solution in this setting?

Sort elements of $E$ w.r.t. $w$ so that, w.l.o.g. $E = (e_1, e_2, \ldots, e_m)$ with $w(e_1) \geq w(e_2) \geq \cdots \geq w(e_m)$.

Let $k + 1$ be the first point (if any) at which we are non-positive, i.e., $w(e_k) > 0$ and $0 \geq w(e_{k+1})$.

Next define partial accumulated sets $E_i$ so that for $i = 0 \ldots m$, we have w.r.t. the above sorted order:

$$ U_i \overset{\text{def}}{=} \{ e_1, e_2, \ldots e_i \} $$

(note $U_0 = \emptyset$ and $f(U_0) = 0$, and $U_i$ is always w.r.t $w$).
Polymatroidal polyhedron and greedy

- What is the greedy solution in this setting?
- Sort elements of $E$ w.r.t. $w$ so that, w.l.o.g.
  $E = (e_1, e_2, \ldots, e_m)$ with $w(e_1) \geq w(e_2) \geq \cdots \geq w(e_m)$.
- Let $k + 1$ be the first point (if any) at which we are non-positive, i.e., $w(e_k) > 0$ and $0 \geq w(e_{k+1})$.
- Next define partial accumulated sets $E_i$ so that for $i = 0 \ldots m$, we have w.r.t. the above sorted order:
  $$U_i \overset{\text{def}}{=} \{e_1, e_2, \ldots, e_i\} \quad (9)$$
  (note $U_0 = \emptyset$ and $f(U_0) = 0$, and $U_i$ is always w.r.t $w$).
- The greedy solution is the vector $x \in \mathbb{R}_+^E$ with elements defined as:
  $$x(e_1) \overset{\text{def}}{=} f(U_1) \quad (10)$$
  $$x(e_i) \overset{\text{def}}{=} f(U_i) - f(U_{i-1}) \text{ for } i = 2 \ldots k \quad (11)$$
  $$x(e_i) \overset{\text{def}}{=} 0 \text{ for } i = k + 1 \ldots m = |E| \quad (12)$$
Polymatroidal polyhedron and greedy

Theorem 4.2

The vector $x \in \mathbb{R}_+^E$ as previously defined maximizes $wx$ over $P_f$.

Proof.

...
Polymatroidal polyhedron and greedy

Theorem 4.2

The vector \( x \in \mathbb{R}_+^E \) as previously defined maximizes \( wx \) over \( P_f \).

Proof.

Consider the LP strong duality equation:

\[
\max (wx : x \in P_f) = \min \left( \sum_{A \subseteq E} y_A f(A) : y \in \mathbb{R}_+^{2^E}, \sum_{A \subseteq E} y_A 1_A \geq w \right)
\]  \hspace{1cm} (13)

Define the following vector \( y \in \mathbb{R}_+^{2^E} \) as

\[
y_U^i \triangleq w(e_i) - w(e_i + 1) \quad \text{for} \quad i = 1, \ldots, m - 1,
\]

\[
y_E \triangleq w(e_n),
\]

and otherwise (16)
Theorem 4.2

The vector \( x \in \mathbb{R}^E_+ \) as previously defined maximizes \( wx \) over \( P_f \).

Proof.

- Consider the LP strong duality equation:

\[
\max (wx : x \in P_f) = \min \left( \sum_{A \subseteq E} y_A f(A) : y \in \mathbb{R}^{2E}_+, \sum_{A \subseteq E} y_A 1_A \geq w \right)
\]  

(13)

- Define the following vector \( y \in \mathbb{R}^{2E}_+ \) as

\[
y_{U_i} \overset{\text{def}}{=} w(e_i) - w(e_{i+1}) \quad \text{for} \ i = 1 \ldots (m - 1), \quad (14)
\]

\[
y_E \overset{\text{def}}{=} w(e_m), \quad \text{and} \quad (15)
\]

\[
y_A = 0 \quad \text{otherwise} \quad (16)
\]
Polymatroidal polyhedron and greedy

**Theorem 4.2**

The vector \( x \in \mathbb{R}^E_+ \) as previously defined maximizes \( wx \) over \( P_f \).

**Proof.**

- We first see that \( x \in P_f \) (that is \( x(A) \leq f(A), \forall A \)) by induction on \( |A| \). Clearly it holds for \( A = \emptyset \).
Polymatroidal polyhedron and greedy

Theorem 4.2

*The vector* $x \in \mathbb{R}^E_+$ *as previously defined maximizes* $wx$ *over* $P_f$.

**Proof.**

- We first see that $x \in P_f$ (that is $x(A) \leq f(A), \forall A$) by induction on $|A|$. Clearly it holds for $A = \emptyset$.

- Assume $A \neq \emptyset$, and let $\ell$ be largest index with $e_\ell \in A$. 

...
Theorem 4.2

The vector $x \in \mathbb{R}_+^E$ as previously defined maximizes $wx$ over $P_f$.

Proof.

- We first see that $x \in P_f$ (that is $x(A) \leq f(A)$, $\forall A$) by induction on $|A|$. Clearly it holds for $A = \emptyset$.
- Assume $A \neq \emptyset$, and let $\ell$ be largest index with $e_\ell \in A$.
- Then, by induction, we have
  \[
  x(A \setminus \{e_\ell\}) \leq f(E \setminus \{e_\ell\})
  \] (17)

...
**Theorem 4.2**

*The vector* \( x \in \mathbb{R}^E_+ \) *as previously defined maximizes* \( wx \) *over* \( P_f \).

**Proof.**

- We first see that \( x \in P_f \) (that is \( x(A) \leq f(A), \forall A \)) by induction on \( |A| \). Clearly it holds for \( A = \emptyset \).

- Assume \( A \neq \emptyset \), and let \( \ell \) be largest index with \( e_\ell \in A \).

- Then, by induction, we have
  \[
  x(A \setminus \{e_\ell\}) \leq f(A \setminus \{e_\ell\}) \quad (17)
  \]

  And therefore,
  \[
  x(A) \leq f(A \setminus \{e_\ell\}) + x(e_\ell) = f(A \setminus \{e_\ell\}) + f(U_\ell) - f(U_{\ell-1}) \leq f(A) \quad (18)
  \]

  where the last inequality follows by submodularity of \( f \) (if \( \ell \leq k \)) and by monotonicity of \( f \) (if \( \ell > k \)) where \( x(e_\ell) = 0 \).

  \[
  f(\mathcal{A}) + f(U_{\ell-1}) = f(U_\ell) + f(A \setminus e_\ell) = f(A \cup U_{\ell-1} \setminus e_\ell) \quad ...
  \]
Polymatroidal polyhedron and greedy

**Theorem 4.2**

The vector $x \in \mathbb{R}^E_+$ as previously defined maximizes $wx$ over $P_f$.

**Proof.**

- We first see that $x \in P_f$ (that is $x(A) \leq f(A), \forall A$) by induction on $|A|$. Clearly it holds for $A = \emptyset$.

- Assume $A \neq \emptyset$, and let $\ell$ be largest index with $e_{\ell} \in A$.

- Then, by induction, we have

$$x(A \setminus \{e_{\ell}\}) \leq f(E \setminus \{e_{\ell}\})$$

(17)

- And therefore,

$$x(A) \leq f(A \setminus \{e_{\ell}\}) + x(e_{\ell}) = f(A \setminus \{e_{\ell}\}) + f(U_{\ell}) - f(U_{\ell - 1}) \leq f(A)$$

(18)

where the last inequality follows by submodularity of $f$ (if $\ell \leq k$) and by monotonicity of $f$ (if $\ell > k$) where $x(e_{\ell}) = 0$.

- So, therefore, we have $x \in P_f$.

...
Polymatroidal polyhedron and greedy

Theorem 4.2

The vector \( x \in \mathbb{R}_+^E \) as previously defined maximizes \( wx \) over \( P_f \).

Proof.

Now \( y \) is also feasible for the dual constraints in Eq. 13 since:

...
Polymatroidal polyhedron and greedy

Theorem 4.2

The vector \( x \in \mathbb{R}^E_+ \) as previously defined maximizes \( wx \) over \( P_f \).

Proof.

- Now \( y \) is also feasible for the dual constraints in Eq. 13 since:
- clearly, \( y \geq 0 \);
Polymatroidal polyhedron and greedy

**Theorem 4.2**

The vector $x \in \mathbb{R}_+^E$ as previously defined maximizes $wx$ over $P_f$.

**Proof.**

- Now $y$ is also feasible for the dual constraints in Eq. 13 since:
  - clearly, $y \geq 0$;
  - also, considering $y$ component wise, for any $i$, we have that
    \[
    \sum_{A : e_i \in A} y_A = \sum_{j \geq i} y_{U_j} = w(e_i). \tag{19}
    \]
Polymatroidal polyhedron and greedy

**Theorem 4.2**

The vector \( x \in \mathbb{R}^E_+ \) as previously defined maximizes \( wx \) over \( P_f \).

**Proof.**

- Now \( y \) is also feasible for the dual constraints in Eq. 13 since:
  - clearly, \( y \geq 0 \);
  - also, considering \( y \) component wise, for any \( i \), we have that
    \[
    \sum_{A: e_i \in A} y_A = \sum_{j \geq i} y_{U_j} = w(e_i). \quad (19)
    \]
- Now optimality for \( x \) and \( y \) follows from
  \[
  wx = \sum_{e \in E} w(e) x(e) = \sum_{i=1}^{m} w(e_i)(f(U_i) - f(U_{i-1})) \quad (20)
  \]
  \[
  = \sum_{i=1}^{n-1} f(U_i)(w(e_i) - w(e_{i+1})) + f(E)w(e_m) = \sum_{A \subseteq E} y_A f(A) \quad \ldots
  \]
Polymatroidal polyhedron and greedy

**Theorem 4.2**

The vector $x \in \mathbb{R}_+^E$ as previously defined maximizes $wx$ over $P_f$.

**Proof.**

The third equality (in Eq. 20) follows since

$$xw = \sum_{i=1}^m x_i w_i = \sum_{i=1}^m x_i \left( \sum_{j=1}^i w(e_j) - \sum_{i=1}^{i-1} w(e_j) \right)$$

(22)

$$= \sum_{i=1}^m x_i \left( w(U_i) - w(U_{i-1}) \right)$$

(23)

$$= \sum_{i=1}^m x_i w(U_i) - \sum_{i=1}^{m-1} x_{i+1} w(U_i)$$

(24)

$$= x_m w(U_m) + \sum_{i=1}^{m-1} (x_i - x_{i+1}) w(U_i)$$

(25)
Conversely, suppose $P$ is a polytope of form
\[ P = \left\{ x \in \mathbb{R}^E_+ : x(A) \leq f(A), \forall A \subseteq E \right\}, \]
then the greedy solution to \( \max(wx : x \in P) \) is optimum only if $f$ is submodular.

**Proof.**

- Name elements of $E$ in arbitrary order $(e_1, e_2, \ldots, e_m)$ and define $E_i = (e_1, e_2, \ldots, e_i)$. 

Theorem 4.3

Conversely, suppose $P$ is a polytope of form

$$P = \{ x \in \mathbb{R}^E_+ : x(A) \leq f(A), \forall A \subseteq E \},$$

then the greedy solution to

$$\max(wx : x \in P)$$

is optimum only if $f$ is submodular.

Proof.

- Name elements of $E$ in arbitrary order $(e_1, e_2, \ldots, e_m)$ and define $E_i = (e_1, e_2, \ldots, e_i)$.
- Define $A = \{ e_1, e_2, \ldots, e_k, e_{k+1}, \ldots, e_p \}$ and $B = \{ e_1, e_2, \ldots, e_k, e_{p+1}, \ldots, e_q \}$ for some $1 \leq p \leq q \leq m$. 

...
Polymatroidal polyhedron and greedy

**Theorem 4.3**

Conversely, suppose $P$ is a polytope of form

$$P = \{ x \in \mathbb{R}^E_+ : x(A) \leq f(A), \forall A \subseteq E \},$$

then the greedy solution to

$$\max(wx : x \in P)$$

is optimum only if $f$ is submodular.

**Proof.**

- Name elements of $E$ in arbitrary order $(e_1, e_2, \ldots, e_m)$ and define $E_i = (e_1, e_2, \ldots, e_i)$.
- Define $A = \{e_1, e_2, \ldots, e_k, e_{k+1}, \ldots, e_p\}$ and $B = \{e_1, e_2, \ldots, e_k, e_{p+1}, \ldots, e_q\}$ for some $1 \leq p \leq q \leq m$.
- Note, then $A \cap B = \{e_1, \ldots, e_k\}$. 

...
Theorem 4.3

Conversely, suppose $P$ is a polytope of form

$$P = \left\{ x \in \mathbb{R}^E_+: x(A) \leq f(A), \forall A \subseteq E \right\},$$

then the greedy solution to

$$\max(wx : x \in P)$$

is optimum only if $f$ is submodular.

Proof.

- Name elements of $E$ in arbitrary order $(e_1, e_2, \ldots, e_m)$ and define $E_i = (e_1, e_2, \ldots, e_i)$.

- Define $A = \{e_1, e_2, \ldots, e_k, e_{k+1}, \ldots, e_p\}$ and $B = \{e_1, e_2, \ldots, e_k, e_{p+1}, \ldots, e_q\}$ for some $1 \leq p \leq q \leq m$.

- Note, then $A \cap B = \{e_1, \ldots, e_k\}$.

- Define $w$ as:

$$w \overset{\text{def}}{=} \sum_{i=1}^{q} 1_{e_i} = 1_{A \cup B} \quad (26)$$

...
Theorem 4.3

Conversely, suppose $P$ is a polytope of form $P = \{ x \in \mathbb{R}^E_+ : x(A) \leq f(A), \forall A \subseteq E \}$, then the greedy solution to $\max(wx : x \in P)$ is optimum only if $f$ is submodular.

Proof.

- Name elements of $E$ in arbitrary order $(e_1, e_2, \ldots, e_m)$ and define $E_i = (e_1, e_2, \ldots, e_i)$.
- Define $A = \{e_1, e_2, \ldots, e_k, e_{k+1}, \ldots, e_p\}$ and $B = \{e_1, e_2, \ldots, e_k, e_{p+1}, \ldots, e_q\}$ for some $1 \leq p \leq q \leq m$.
- Note, then $A \cap B = \{e_1, \ldots, e_k\}$.
- Define $w$ as:

$$w \overset{\text{def}}{=} \sum_{i=1}^{q} 1_{e_i} = 1_{A \cup B} \quad (26)$$

- Suppose optimum solution $x$ is given by the greedy procedure.
Polymatroidal polyhedron and greedy

**Theorem 4.3**

Conversely, suppose $P$ is a polytope of form

$$P = \{ x \in \mathbb{R}^E_+ : x(A) \leq f(A), \forall A \subseteq E \},$$

then the greedy solution to $\max(wx : x \in P)$ is optimum only if $f$ is submodular.

**Proof.**

Then

$$\sum_{i=1}^k x_i = f(U_1) + \sum_{i=2}^k (f(U_i) - f(U_{i-1})) = f(U_k) = f(A \cap B) \quad (27)$$
Theorem 4.3

Conversely, suppose $P$ is a polytope of form

$P = \{x \in \mathbb{R}_+^E : x(A) \leq f(A), \forall A \subseteq E\}$, then the greedy solution to

$max(wx : x \in P)$ is optimum only if $f$ is submodular.

Proof.

Then

\[
\sum_{i=1}^{k} x_i = f(U_1) + \sum_{i=2}^{k} (f(U_i) - f(U_{i-1})) = f(U_k) = f(A \cap B) \quad (27)
\]

\[
\sum_{i=1}^{p} x_i = f(U_1) + \sum_{i=2}^{p} (f(U_i) - f(U_{i-1})) = f(U_p) \geq f(A) \quad (28)
\]

...
Polymatroidal polyhedron and greedy

Theorem 4.3

Conversely, suppose $P$ is a polytope of form

$P = \{ x \in \mathbb{R}^E_+ : x(A) \leq f(A), \forall A \subseteq E \}$, then the greedy solution to

$\max(wx : x \in P)$ is optimum only if $f$ is submodular.

Proof.

Then

$$\sum_{i=1}^{k} x_i = f(U_1) + \sum_{i=2}^{k} (f(U_i) - f(U_{i-1})) = f(U_k) = f(A \cap B) \quad (27)$$

And

$$\sum_{i=1}^{p} x_i = f(U_1) + \sum_{i=2}^{p} (f(U_i) - f(U_{i-1})) = f(U_p)f(A) \quad (28)$$

And

$$\sum_{i=1}^{q} x_i = f(U_1) + \sum_{i=2}^{q} (f(U_i) - f(U_{i-1})) = f(U_q) = f(A \cup B) \quad (29)$$
Theorem 4.3

Conversely, suppose $P$ is a polytope of form

$$P = \{ x \in \mathbb{R}^E_+ : x(A) \leq f(A), \forall A \subseteq E \},$$

then the greedy solution to

$$\max( wx : x \in P )$$

is optimum only if $f$ is submodular.

Proof.

Thus, we have

$$\sum_{i : e_i \in B} x_i = f(A \cup B) + f(A \cap B) - f(A) \quad (30)$$
Theorem 4.3

Conversely, suppose \( P \) is a polytope of form
\[
P = \{ x \in \mathbb{R}^E_+ : x(A) \leq f(A), \forall A \subseteq E \},
\]
then the greedy solution to \( \max(wx : x \in P) \) is optimum only if \( f \) is submodular.

Proof.

- Thus, we have
  \[
  \sum_{i : e_i \in B} x_i = f(A \cup B) + f(A \cap B) - f(A)
  \]
  (30)
- But given that the greedy algorithm gives the optimal solution to \( \max(wx : x \in P) \), we have that \( x \in P \).
Polymatroidal polyhedron and greedy

**Theorem 4.3**

Conversely, suppose $P$ is a polytope of form

$$P = \{ x \in \mathbb{R}_+^E : x(A) \leq f(A), \forall A \subseteq E \},$$

then the greedy solution to

$$\max (wx : x \in P)$$

is optimum only if $f$ is submodular.

**Proof.**

- Thus, we have
  \[
  \sum_{i : e_i \in B} x_i = f(A \cup B) + f(A \cap B) - f(A) \leq f(B) \quad (30)
  \]

- But given that the greedy algorithm gives the optimal solution to
  \(\max (wx : x \in P)\), we have that $x \in P$.

- Thus,
  \[
  \chi(B) = f(A \cup B) + f(A \cap B) - f(A) = \sum_{i : e_i \in B} x_i \leq f(B) \quad (31)
  \]

ensuring the submodularity of $f$, since $A$ and $B$ are arbitrary.
Thus, summarizing this into the complete theorem, we have a result very similar to matroids.

**Theorem 4.4**

If $f : 2^E \to \mathbb{R}_+$ is given, and $P$ is a polytope in $\mathbb{R}^E_+$ of the form

$$P = \{ x \in \mathbb{R}^E_+ : x(A) \leq f(A), \forall A \subseteq E \},$$

then the greedy solution to the problem $\max (w x : x \in P)$ is optimum iff $f$ is monotone non-decreasing submodular (i.e., iff $P$ is a polymatroid).
An extension of $f$

- We may consider the optimization a function $\tilde{f} : \mathbb{R}^E \rightarrow \mathbb{R}$ as
  \[
  \tilde{f}(w) = \max(wx : x \in P_f)
  \] (32)
An extension of $f$

- We may consider the optimization a function $\tilde{f} : \mathbb{R}^E \rightarrow \mathbb{R}$ as
  $$\tilde{f}(w) = \max( wx : x \in P_f )$$  (32)

- Then, for any $w$, from the above theorem, we can compute this function using the greedy algorithm.
An extension of $f$

- We may consider the optimization a function $\tilde{f} : \mathbb{R}^E \rightarrow \mathbb{R}$ as
  \[ \tilde{f}(w) = \max(wx : x \in P_f) \]  
  (32)

- Then, for any $w$, from the above theorem, we can compute this function using the greedy algorithm.

- That is, we have
  \[ \tilde{f}(w) = \max(wx : x \in P_f) \]  
  (33)

  \[ = \sum_{i=1}^{m} w(e_i)(f(U_i) - f(U_{i-1})) \]  
  (34)

  \[ = w(e_m)f(U_m) + \sum_{i=1}^{m-1} (w(e_i) - w(e_{i+1}))f(U_i) \]  
  (35)

where $U_i = \{e_1, e_2, \ldots, e_i\}$ based on the elements of $E$ being named, w.l.o.g., in order of decreasing $w$, so that $w(e_1) \geq w(e_2) \geq \cdots \geq w(e_m)$. 
An extension of $f$

Moreover, from $\tilde{f}$ we can recover $f$. 
An extension of $f$

- Moreover, from $\tilde{f}$ we can recover $f$.
- Take $w = 1_A$ for some $A \subseteq E$. 
An extension of $f$

- Moreover, from $\tilde{f}$ we can recover $f$.
- Take $w = 1_A$ for some $A \subseteq E$.
- Then, we order $w$ so that $1_A(i) = 1$ if $i \leq |A|$, and $1_A(i) = 0$ otherwise.
An extension of \( f \)

- Moreover, from \( \tilde{f} \) we can recover \( f \).
- Take \( w = \mathbf{1}_A \) for some \( A \subseteq E \).
- Then, we order \( w \) so that \( \mathbf{1}_A(i) = 1 \) if \( i \leq |A| \), and \( \mathbf{1}_A(i) = 0 \) otherwise.
- This gives
  \[
  \tilde{f}(w) = \max(\mathbf{1}_A x : x \in P_f) \tag{36}
  \]
  \[
  = \mathbf{1}_A(m)f(U_m) + \sum_{i=1}^{m-1}(\mathbf{1}_A(i) - \mathbf{1}_A(i+1))f(U_i) \tag{37}
  \]
  \[
  = (\mathbf{1}_A(|A|) - \mathbf{1}_A(|A| + 1))f(U_i) \tag{38}
  \]
  \[
  = f(A) \tag{39}
  \]
An extension of $f$

\[
\tilde{f}(w) = \max(wx : x \in P_f) \quad (40)
\]

Therefore, if $f$ is a submodular function, we can write

\[
\tilde{f}(w) = \sum_{i=1}^{m} \lambda_i f(U_i) \quad (41)
\]

where $\lambda_m = w(e_m)$ and otherwise $\lambda_i = w(e_i) - w(e_{i+1})$, where the elements are sorted according to $w$ as before.
An extension of $f$ 

$$\tilde{f}(w) = \max(wx : x \in P_f) \quad (40)$$

- Therefore, if $f$ is a submodular function, we can write
  $$\tilde{f}(w) = \sum_{i=1}^{m} \lambda_i f(U_i) \quad (41)$$

  where $\lambda_m = w(e_m)$ and otherwise $\lambda_i = w(e_i) - w(e_{i+1})$, where the elements are sorted according to $w$ as before.

- Clearly, $\tilde{f}(w)$ is always convex in $w$, since it is the maximum of a set of linear functions.
An extension of $f$

- Recall, for any such $w \in \mathbb{R}^E$, we have

$$
\begin{bmatrix}
w_1 \\
w_2 \\
\vdots \\
w_n
\end{bmatrix} = (w_1 - w_2) \begin{bmatrix} 1 \\
0 \\
\vdots \\
0
\end{bmatrix} + (w_2 - w_3) \begin{bmatrix} 1 \\
1 \\
0 \\
\vdots \\
0
\end{bmatrix} + \cdots + (w_{n-1} - w_n) \begin{bmatrix} 1 \\
1 \\
0 \\
\vdots \\
1
\end{bmatrix} + (w_n) \begin{bmatrix} 1 \\
1 \\
0 \\
\vdots \\
1
\end{bmatrix}
$$

(42)
An extension of $f$

- Recall, for any such $w \in \mathbb{R}^E$, we have

$$
\begin{pmatrix}
  w_1 \\
  w_2 \\
  \vdots \\
  w_n
\end{pmatrix} = (w_1 - w_2) \begin{pmatrix}
  1 \\
  0 \\
  \vdots \\
  0
\end{pmatrix} + (w_2 - w_3) \begin{pmatrix}
  1 \\
  1 \\
  \vdots \\
  0
\end{pmatrix} + \\
\vdots + (w_{n-1} - w_n) \begin{pmatrix}
  1 \\
  1 \\
  \vdots \\
  0
\end{pmatrix} + (w_n) \begin{pmatrix}
  \vdots \\
  1 \\
  1
\end{pmatrix}
$$

(42)

- If we take $w$ in decreasing order, then each coefficient of the vectors is non-negative (except possibly the last one, $w_n$).
An extension of $f$

Define sets $U_i$ based on this decreasing order as follows, for $i = 0, \ldots, n$

$$U_i \overset{\text{def}}{=} \{ e_1, e_2, \ldots, e_i \}$$ (43)
An extension of $f$

- Define sets $U_i$ based on this decreasing order as follows, for $i = 0, \ldots, n$

$$U_i \overset{\text{def}}{=} \{e_1, e_2, \ldots, e_i\}$$ (43)

- Note that

\[
\mathbf{1}_{U_0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{1}_{U_1} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad \mathbf{1}_{U_\ell} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{\ell \times (n - \ell)}
\] (44)
An extension of $f$

Thus, for any $f$, we can define an extension in this way, with

$$\tilde{f}(w) = \sum_{i=1}^{m} \lambda_i f(U_i)$$  \hspace{1cm} (45)$$

with the $U_i$'s and sorted order of $w$ defined as above, so that $w = \sum_{i=1}^{m} \lambda_i \mathbf{1}_{U_i}$
An extension of $f$

Thus, for any $f$, we can define an extension in this way, with

$$\tilde{f}(w) = \sum_{i=1}^{m} \lambda_i f(U_i) \quad (45)$$

with the $U_i$'s and sorted order of $w$ defined as above, so that $w = \sum_{i=1}^{m} \lambda_i 1_{U_i}$

Lovász showed that if a function $\tilde{f}(w)$, so defined is convex, then the underlying $f$ must be submodular.
An extension of $f$

Thus, for any $f$, we can define an extension in this way, with

$$\tilde{f}(w) = \sum_{i=1}^{m} \lambda_i f(U_i)$$

with the $U_i$'s and sorted order of $w$ defined as above, so that $w = \sum_{i=1}^{m} \lambda_i 1_{U_i}$

Lovász showed that if a function $\tilde{f}(w)$, so defined is convex, then the underlying $f$ must be submodular.

This “extension” of $f$, in any case, is called the Lovász extension of $f$. 

"Lovász extension\"
Polymatroidal polyhedron and greedy

**Theorem 5.1**

A function $f : 2^E \to \mathbb{R}$ is submodular iff its Lovász extension $\tilde{f}$ of $f$ is convex.

**Proof.**
Polymatroidal polyhedron and greedy

**Theorem 5.1**

A function \( f : 2^E \to \mathbb{R} \) is submodular iff its Lovász extension \( \tilde{f} \) of \( f \) is convex.

**Proof.**

- We’ve already shown that if \( f \) is submodular, its extension can be written this way, and thus is convex.
Polymatroidal polyhedron and greedy

Theorem 5.1

A function $f : 2^E \rightarrow \mathbb{R}$ is submodular iff its Lovász extension $\tilde{f}$ of $f$ is convex.

Proof.

- We’ve already shown that if $f$ is submodular, its extension can be written this way, and thus is convex.
- Conversely, suppose the Lovász extension $\tilde{f}$ of $f$ is a convex function.
Polymatroidal polyhedron and greedy

Theorem 5.1

A function \( f : \mathcal{P}(E) \to \mathbb{R} \) is submodular iff its Lovász extension \( \tilde{f} \) of \( f \) is convex.

Proof.

- We’ve already shown that if \( f \) is submodular, its extension can be written this way, and thus is convex.
- Conversely, suppose the Lovász extension \( \tilde{f} \) of \( f \) is a convex function.
- We note that, based on the extension definition, \( \tilde{f}(\alpha w) = \alpha \tilde{f}(w) \) for any \( \alpha \in \mathbb{R}^+ \). I.e., \( f \) is a positively homogeneous convex function.
Polymatroidal polyhedron and greedy

**Theorem 5.1**

A function $f : 2^E \to \mathbb{R}$ is submodular iff its Lovász extension $\tilde{f}$ of $f$ is convex.

**Proof.**

- We’ve already shown that if $f$ is submodular, its extension can be written this way, and thus is convex.

- Conversely, suppose the Lovász extension $\tilde{f}$ of $f$ is a convex function.

- We note that, based on the extension definition, $\tilde{f}(\alpha w) = \alpha \tilde{f}(w)$ for any $\alpha \in \mathbb{R}_+$. I.e., $f$ is a positively homogeneous convex function.

- Given $A, B \subseteq E$, we have that
  \[
  \tilde{f}(1_A + 1_B) = \tilde{f}(1_{A \cup B} + 1_{A \cap B}) \tag{46}
  \]
  \[
  = f(A \cup B) + f(A \cap B). \tag{47}
  \]

  Exercise: show this.

...
Polymatroidal polyhedron and greedy

Theorem 5.1

A function \( f : 2^E \rightarrow \mathbb{R} \) is submodular iff its Lovász extension \( \tilde{f} \) of \( f \) is convex.

Proof.

Also, since \( \tilde{f} \) is convex, we have

\[
\tilde{f}(0.5\mathbf{1}_A + 0.5\mathbf{1}_B) \leq 0.5\tilde{f}(\mathbf{1}_A) + 0.5\tilde{f}(\mathbf{1}_B)
\]

\[
= 0.5(f(A) + f(B))
\]
Polymatroidal polyhedron and greedy

**Theorem 5.1**

A function $f : 2^E \rightarrow \mathbb{R}$ is submodular iff its Lovász extension $\tilde{f}$ of $f$ is convex.

**Proof.**

- Also, since $\tilde{f}$ is convex, we have
  
  \[
  \tilde{f}(0.5\mathbf{1}_A + 0.5\mathbf{1}_B) \leq 0.5\tilde{f}(\mathbf{1}_A) + 0.5\tilde{f}(\mathbf{1}_B) \tag{48}
  \]
  
  \[
  = 0.5(f(A) + f(B)) \tag{49}
  \]

- Thus,

  \[
  f(A \cup B) + f(A \cap B) \leq f(A) + f(B) \tag{50}
  \]

  as required..
1. What did Edmond, Punth/Nelly, Lovász show Add slide.
Sources for Today’s Lecture