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 12 - May 11th, 2011
Announcements

- On Final projects. **One** single page final project updates due next Wednesday, 5/18 at 5:00pm.

- 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.
We need to find one makeup lectures this term.

- L1 (3/30): Review
- L2 (4/1): Lovász extension
- L3 (4/6): Polymatroids
- L4 (4/8): Scratch
- L5 (4/13): Summary

- L19 (6/3):
- L20: (6/?): (need to find time/date/place).
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 $r(A) - \frac{1}{k} \mathbf{1}(A)$.
- 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.
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.

It turns out that (2) and (3) are easy to deal with, but (1) took another 16 years to solve. In fact, the problem can still be seen as unsolved, if we want a reasonable, scalable, guaranteed low-order polynomial algorithm.
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:
  - $r(A)$ 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 \( r(A) - \frac{1}{k} \mathbf{1}(A) \).

- 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 \( x \geq 0 \).
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.
Towards SFM

\[ f(A) = g(A) - \sum_{e \in E} \sum_{\text{paths } \pi \text{ that contain } e} \phi(\pi) \]

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) are easy to deal with, but (1) took another 16 years to solve. In fact, the problem can still be seen as unsolved, if we want a reasonable, scalable, guaranteed low-order polynomial algorithm.
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 minimization, we can test polyhedral membership, in at least this limited (polymatroidal polytope) sense.
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 $0 \leq f(A) - x(A), \forall A \subseteq E$. 

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

So, given a strongly polynomial time algorithm for general submodular function minimization, we can test polyhedral membership, in at least this limited (polymatroidal polytope) sense.
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 $\min f(A) - x(A) \geq 0$. 
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 minimization, we can test polyhedral membership, in at least this limited (polymatroidal polytope) sense.
We have a result very similar to that of matroids.

**Theorem 2.1**

If \( f : 2^E \rightarrow \mathbb{R}_+ \) 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 \( \max (wx : x \in P) \) is \( \forall w \) 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 $\tilde{f} : \mathbb{R}^E \to \mathbb{R}$ as

$$\tilde{f}(w) = \max(wx : x \in P_f) \quad (1)$$
An extension of $f$

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

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

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

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

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 $w(e_1) \geq w(e_2) \geq \cdots \geq w(e_m)$. 

Prof. Jeff Bilmes
EE595A/Spr 2011/Submodular Functions – Lecture 12 - May 11th, 2011
An extension of $f$

- We may consider the optimization a function $\tilde{f} : \mathbb{R}^E \rightarrow \mathbb{R}$ as
  \[
  \tilde{f}(w) = \max(wx : x \in P_f)
  \] (1)

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

- That is, we have, for submodular $f$,
  \[
  \tilde{f}(w) = \max(wx : x \in P_f)
  \] (2)
  \[
  = \sum_{i=1}^{m} w(e_i)(f(U_i) - f(U_{i-1}))
  \] (3)
  \[
  = w(e_m)f(U_m) + \sum_{i=1}^{m-1} (w(e_i) - w(e_{i+1}))f(U_i)
  \] (4)

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 $w(e_1) \geq w(e_2) \geq \cdots \geq w(e_m)$. 

Prof. Jeff Bilmes
An extension of $f$

\[ \tilde{f}(w) = \max( wx : x \in P_f ) \]  \hfill (5)

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

\[ \tilde{f}(w) = \sum_{i=1}^{m} \lambda_i f(U_i) \]  \hfill (6)

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 \( f \)

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

(5)

Therefore, if \( f \) is a submodular function, we can write

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

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, \( \tilde{f}(w) = \max(wx : x \in P_f) \) is always convex in \( w \), since it is the maximum of a set of linear functions (even when \( f \) is not submodular) or even for any \( a \)-convex set \( P_f \).
An extension of $f$

On the other hand, for any $f$ (even not submodular), we can define an extension in this way, with

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

(7)

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 $f$. 
An extension of $f$

- On the other hand, for any $f$ (even not submodular), we can define an extension in this way, with

$$\tilde{f}(w) = \sum_{i=1}^{m} \lambda_i f(U_i) \quad (7)$$

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

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

- On the other hand, for any $f$ (even not submodular), 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 $f$. 

Theorem 3.1

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

Proof.

seen in last lecture.
Who’s Extension? Who did what?

- The expression $\max( wx : x \in P_f )$ for submodular function $f$:
Who’s Extension? Who did what?

- The expression \( \max(wx : x \in P_f) \) for submodular function \( f \): Edmonds, 1970.
Who’s Extension? Who did what?

- The expression $\max(wx : x \in P_f)$ for submodular function $f$: Edmonds, 1970.
- The greedy solution to $\max(wx : x \in P_f)$ for submodular function $f$: 

Who’s Extension? Who did what?

- The expression $\max(wx : x \in P_f)$ for submodular function $f$: Edmonds, 1970.
- The greedy solution to $\max(wx : x \in P_f)$ for submodular function $f$: Edmonds, 1970.
Who’s Extension? Who did what?

- The expression $\max(wx : x \in P_f)$ for submodular function $f$: Edmonds, 1970.
- The greedy solution to $\max(wx : x \in P_f)$ for submodular function $f$: Edmonds, 1970.
- Expression of the greedy solution to $\max(wx : x \in P_f)$ for $f$ submodular as $\sum_{i=1}^{m} w(e_i)(f(U_i) - f(U_{i-1}))$: 

The convexity of $\max(wx : x \in P_f)$ when seen as a function of $w$: Obvious.

The necessity of submodular $f$ for convex $\tilde{f}(w) = \max(wx : x \in P_f)$: Lovász-1983

Simple Algebra although r.h.s. first appeared in Lovász-1983.
Who’s Extension? Who did what?

- The expression \( \max(wx : x \in P_f) \) for submodular function \( f \): Edmonds, 1970.
- The greedy solution to \( \max(wx : x \in P_f) \) for submodular function \( f \): Edmonds, 1970.
- Expression of the greedy solution to \( \max(wx : x \in P_f) \) for \( f \) submodular as \( \sum_{i=1}^{m} w(e_i)(f(U_i) - f(U_{i-1})) \): Edmonds, 1970.
Who’s Extension? Who did what?

- The expression \( \max(wx : x \in P_f) \) for submodular function \( f \): Edmonds, 1970.

- The greedy solution to \( \max(wx : x \in P_f) \) for submodular function \( f \): Edmonds, 1970.

- Expression of the greedy solution to \( \max(wx : x \in P_f) \) for \( f \) submodular as \( \sum_{i=1}^{m} w(e_i)(f(U_i) - f(U_{i-1})) \): Edmonds, 1970.

- Expression of the greedy solution to \( \max(wx : x \in P_f) \) for \( f \) submodular as \( w(e_m)f(U_m) + \sum_{i=1}^{m-1}(w(e_i) - w(e_{i+1}))f(U_i) = \sum_{i=1}^{m} \lambda_i f(U_i) \):
Who’s Extension? Who did what?

- The expression \( \max(wx : x \in P_f) \) for submodular function \( f \): Edmonds, 1970.
- The greedy solution to \( \max(wx : x \in P_f) \) for submodular function \( f \): Edmonds, 1970.
- Expression of the greedy solution to \( \max(wx : x \in P_f) \) for \( f \) submodular as \( \sum_{i=1}^{m} w(e_i)(f(U_i) - f(U_{i-1})) \): Edmonds, 1970.
- Expression of the greedy solution to \( \max(wx : x \in P_f) \) for \( f \) submodular as \( w(e_m)f(U_m) + \sum_{i=1}^{m-1}(w(e_i) - w(e_{i+1}))f(U_i) = \sum_{i=1}^{m} \lambda_if(U_i) \): Simple Algebra although r.h.s. first appeared in Lovász-1983.
Who’s Extension? Who did what?

- The expression $\max(wx : x \in P_f)$ for submodular function $f$: Edmonds, 1970.
- The greedy solution to $\max(wx : x \in P_f)$ for submodular function $f$: Edmonds, 1970.
- Expression of the greedy solution to $\max(wx : x \in P_f)$ for $f$ submodular as $\sum_{i=1}^{m} w(e_i)(f(U_i) - f(U_{i-1}))$: Edmonds, 1970.
- Expression of the greedy solution to $\max(wx : x \in P_f)$ for $f$ submodular as $w(e_m)f(U_m) + \sum_{i=1}^{m-1}(w(e_i) - w(e_{i+1}))f(U_i) = \sum_{i=1}^{m} \lambda_i f(U_i)$: Simple Algebra although r.h.s. first appeared in Lovász-1983
- Viewing $\tilde{f}(w) = \max(wx : x \in P_f)$ as a function of $w$: 
Who’s Extension? Who did what?

- The expression $\max(wx : x \in P_f)$ for submodular function $f$: Edmonds, 1970.
- The greedy solution to $\max(wx : x \in P_f)$ for submodular function $f$: Edmonds, 1970.
- Expression of the greedy solution to $\max(wx : x \in P_f)$ for $f$ submodular as $\sum_{i=1}^{m} w(e_i)(f(U_i) - f(U_{i-1}))$: Edmonds, 1970.
- Expression of the greedy solution to $\max(wx : x \in P_f)$ for $f$ submodular as $w(e_m)f(U_m) + \sum_{i=1}^{m-1}(w(e_i) - w(e_{i+1}))f(U_i) = \sum_{i=1}^{m} \lambda_i f(U_i)$: Simple Algebra although r.h.s. first appeared in Lovász-1983
- Viewing $\tilde{f}(w) = \max(wx : x \in P_f)$ as a function of $w$: Lovász-1983
Who’s Extension? Who did what?

- The expression $\max(wx : x \in P_f)$ for submodular function $f$: Edmonds, 1970.

- The greedy solution to $\max(wx : x \in P_f)$ for submodular function $f$: Edmonds, 1970.

- Expression of the greedy solution to $\max(wx : x \in P_f)$ for $f$ submodular as $\sum_{i=1}^{m} w(e_i)(f(U_i) - f(U_{i-1}))$: Edmonds, 1970.

- Expression of the greedy solution to $\max(wx : x \in P_f)$ for $f$ submodular as $w(e_m)f(U_m) + \sum_{i=1}^{m-1}(w(e_i) - w(e_{i+1}))f(U_i) = \sum_{i=1}^{m} \lambda_i f(U_i)$: Simple Algebra although r.h.s. first appeared in Lovász-1983

- Viewing $\tilde{f}(w) = \max(wx : x \in P_f)$ as a function of $w$: Lovász-1983

- The convexity of $\max(wx : x \in P_f)$ when seen as a function of $w$:
Who’s Extension? Who did what?

- The expression $\max(wx : x \in P_f)$ for submodular function $f$: Edmonds, 1970.

- The greedy solution to $\max(wx : x \in P_f)$ for submodular function $f$: Edmonds, 1970.

- Expression of the greedy solution to $\max(wx : x \in P_f)$ for $f$ submodular as $\sum_{i=1}^{m} w(e_i)(f(U_i) - f(U_{i-1}))$: Edmonds, 1970.

- Expression of the greedy solution to $\max(wx : x \in P_f)$ for $f$ submodular as $w(e_m)f(U_m) + \sum_{i=1}^{m-1} (w(e_i) - w(e_{i+1}))f(U_i) = \sum_{i=1}^{m} \lambda_i f(U_i)$: Simple Algebra although r.h.s. first appeared in Lovász-1983

- Viewing $\tilde{f}(w) = \max(wx : x \in P_f)$ as a function of $w$: Lovász-1983

- The convexity of $\max(wx : x \in P_f)$ when seen as a function of $w$: Obvious.
Who’s Extension? Who did what?

- The expression $\max(wx : x \in P_f)$ for submodular function $f$: Edmonds, 1970.

- The greedy solution to $\max(wx : x \in P_f)$ for submodular function $f$: Edmonds, 1970.

- Expression of the greedy solution to $\max(wx : x \in P_f)$ for $f$ submodular as $\sum_{i=1}^{m} w(e_i)(f(U_i) - f(U_{i-1}))$: Edmonds, 1970.

- Expression of the greedy solution to $\max(wx : x \in P_f)$ for $f$ submodular as $w(e_m)f(U_m) + \sum_{i=1}^{m-1}(w(e_i) - w(e_{i+1}))f(U_i) = \sum_{i=1}^{m} \lambda_i f(U_i)$: Simple Algebra although r.h.s. first appeared in Lovász-1983

- Viewing $\tilde{f}(w) = \max(wx : x \in P_f)$ as a function of $w$: Lovász-1983

- The convexity of $\max(wx : x \in P_f)$ when seen as a function of $w$: Obvious.

- The necessity of submodular $f$ for convex $\tilde{f}(w) = \max(wx : x \in P_f)$:
Who’s Extension? Who did what?

- The expression $\max(wx : x \in P_f)$ for submodular function $f$: Edmonds, 1970.
- The greedy solution to $\max(wx : x \in P_f)$ for submodular function $f$: Edmonds, 1970.
- Expression of the greedy solution to $\max(wx : x \in P_f)$ for $f$ submodular as $\sum_{i=1}^{m} w(e_i)(f(U_i) - f(U_{i-1}))$: Edmonds, 1970.
- Expression of the greedy solution to $\max(wx : x \in P_f)$ for $f$ submodular as $w(e_m)f(U_m) + \sum_{i=1}^{m-1}(w(e_i) - w(e_{i+1}))f(U_i) = \sum_{i=1}^{m} \lambda_i f(U_i)$: Simple Algebra although r.h.s. first appeared in Lovász-1983
- Viewing $\tilde{f}(w) = \max(wx : x \in P_f)$ as a function of $w$: Lovász-1983
- The convexity of $\max(wx : x \in P_f)$ when seen as a function of $w$: Obvious.
- The necessity of submodular $f$ for convex $\tilde{f}(w) = \max(wx : x \in P_f)$: Lovász-1983
Who’s Extension? Who did what?

Lets have a (somewhat silly) credit assignment vote.

- Edmonds: 0
- Lovász: 0
- Edmonds - Lovász: 0
Who’s Extension? Who did what?

Let's have a (somewhat silly) credit assignment vote.

- Edmonds:
- Lovász:
- Choquet?
Who’s Extension? Who did what?

Lets have a (somewhat silly) credit assignment vote.

- Edmonds:
- Lovász:
- Choquet?

Choquet in 1955 defined what is now known as the Choquet integral, which is a form of “non-linear” integration over discrete finite sets. This turns out to be equivalent to the Lovász extension, as we next see.

\[ \text{Seguno integral} \]
\[ \text{fuzzy measure} \]
\[ \text{non-linear measure}. \]
Integration

- Integration is just summation (e.g., the $\int$ symbol has as its origins a sum).
Integration

- Integration is just summation (e.g., the $\int$ symbol has as its origins a sum).

- Lebesgue integration allows integration w.r.t. an underlying measure $\mu$ of sets. E.g., given measurable function $f$, we can define

$$\int_X f \, du = \sup I_X(s)$$

where

$$I_X(s) = \sum_{i=1}^n c_i \mu(X \cap X_i),$$

and where we take the sup over all measurable functions $s$ such that $0 \leq s \leq f$ and

$$s(x) = \sum_{i=1}^n c_i l_{X_i}(x)$$

and where $l_{X_i}(x)$ is indicator of membership of set $X_i$, with $c_i > 0$. 
Integration

- In finite discrete spaces, Lebesgue integration is just a weighted average, and can be seen as an aggregation function.
In finite discrete spaces, Lebesgue integration is just a weighted average, and can be seen as an aggregation function.

I.e., given a weight vector $w \in [0, 1]^E$ for some finite ground set $E$, then for any $x \in \mathbb{R}^E$ we have

$$WAVG(x) = \sum_{e \in E} x(e)w(e)$$  \hspace{1cm} (9)
In finite discrete spaces, Lebesgue integration is just a weighted average, and can be seen as an aggregation function.

I.e., given a weight vector $w \in [0, 1]^E$ for some finite ground set $E$, then for any $x \in \mathbb{R}^E$ we have

$$\text{WAVG}(x) = \sum_{e \in E} x(e)w(e) \quad (9)$$

Consider $1_e$ for $e \in E$, we have

$$\text{WAVG}(1_e) = w(e) \quad (10)$$

so seen as a function on the hypercube vertices, the entire WAVG function is given based on values on a subset of the vertices of this hypercube, i.e., $\{1_e : e \in E\}$. Moreover, we are interpolating as in

$$\text{WAVG}(x) = \sum_{e \in E} x(e)w(e) = \sum_{e \in E} x(e)\text{WAVG}(1_e) \quad (11)$$
Integration

- More complex aggregation functions can be constructed by defining the aggregation function on all vertices of the hypercube. I.e., for each $1_A : A \subseteq E$ we might have (for all $A \subseteq E$):
  \[ AG(1_A) = w_A \] (12)
More complex aggregation functions can be constructed by defining the aggregation function on all vertices of the hypercube. I.e., for each $1_A : A \subseteq E$ we might have (for all $A \subseteq E$):

$$AG(1_A) = w_A \quad (12)$$

What then might $AG(x)$ be for some $x \in \mathbb{R}^E$? Our weighted average functions might look something more like

$$AG(x) = \sum_{A \subseteq E} x(A)w_A = \sum_{A \subseteq E} x(A)AG(1_A) \quad (13)$$
More complex aggregation functions can be constructed by defining the aggregation function on all vertices of the hypercube. I.e., for each $1_A : A \subseteq E$ we might have (for all $A \subseteq E$):

$$ AG(1_A) = w_A $$

(12)

What then might $AG(x)$ be for some $x \in \mathbb{R}^E$? Our weighted average functions might look something more like

$$ AG(x) = \sum_{A \subseteq E} x(A)w_A = \sum_{A \subseteq E} x(A)AG(1_A) $$

(13)

Set function $f : 2^E \rightarrow \mathbb{R}$ is a game if $f$ is normalized $f(\emptyset) = 0.$
Set function $f : 2^E \to \mathbb{R}$ is called a capacity if it is monotone non-decreasing, i.e., $f(A) \leq f(B)$ whenever $A \subseteq B$. 
Integration

- Set function $f : 2^E \rightarrow \mathbb{R}$ is called a capacity if it is monotone non-decreasing, i.e., $f(A) \leq f(B)$ whenever $A \subseteq B$.

- A Boolean function $f$ is any function $f : \{0, 1\}^m \rightarrow \{0, 1\}$ and is a pseudo-boolean function if $f : \{0, 1\}^m \rightarrow \mathbb{R}$.
Set function $f : 2^E \rightarrow \mathbb{R}$ is called a capacity if it is monotone non-decreasing, i.e., $f(A) \leq f(B)$ whenever $A \subseteq B$.

A Boolean function $f$ is any function $f : \{0, 1\}^m \rightarrow \{0, 1\}$ and is a pseudo-boolean function if $f : \{0, 1\}^m \rightarrow \mathbb{R}$.

Any set function corresponds to a pseudo-boolean function. I.e., given $f : 2^E \rightarrow \mathbb{R}$, form $f_b : \{0, 1\}^m \rightarrow \mathbb{R}$ as $f_b(x) = f(A_x)$ where $A = \{e \in E : x_e = 1\}$.
Integration

- Set function $f : 2^E \to \mathbb{R}$ is called a capacity if it is monotone non-decreasing, i.e., $f(A) \leq f(B)$ whenever $A \subseteq B$.
- A Boolean function $f$ is any function $f : \{0, 1\}^m \to \{0, 1\}$ and is a pseudo-boolean function if $f : \{0, 1\}^m \to \mathbb{R}$.
- Any set function corresponds to a pseudo-boolean function. I.e., given $f : 2^E \to \mathbb{R}$, form $f_b : \{0, 1\}^m \to \mathbb{R}$ as $f_b(x) = f(A_x)$ where $A = \{ e \in E : x_e = 1 \}$.
- Also, If we have an expression for $f_b$ we can construct a set function $f$ as $f(A) = f_b(1_A)$. We can also often relax $f_b$ to any $x \in [0, 1]^m$. 
Definition 3.2

Let $f$ be any capacity on $E$ and $w \in \mathbb{R}^E_+$. The Choquet integral of $w$ w.r.t. $f$ is defined by

$$C_f(w) = \sum_{i=1}^{m} (w_{e_i} - w_{e_{i+1}}) f(U_i)$$

(14)

where in the sum, we have sorted and renamed the elements of $E$ so that $w_{e_1} \geq w_{e_2} \geq \cdots \geq w_{e_m} \geq w_{e_{m+1}} = 0$, and where $U_i = \{e_1, e_2, \ldots, e_i\}$.

We immediately see that an equivalent formula is as follows:

$$C_f(w) = \sum_{i=1}^{m} w(e_i)(f(U_i) - f(U_{i-1}))$$

(15)

where $U_0 \overset{\text{def}}{=} \emptyset$. 

Prof. Jeff Bilmes
EE595A/Spr 2011/Submodular Functions – Lecture 12 - May 11th, 2011 page 13
Definition 3.2

Let $f$ be any capacity on $E$ and $w \in \mathbb{R}^E_+$. The Choquet integral of $w$ w.r.t. $f$ is defined by

$$C_f(w) = \sum_{i=1}^{m} (w_{e_i} - w_{e_{i+1}}) f(U_i)$$

where in the sum, we have sorted and renamed the elements of $E$ so that $w_{e_1} \geq w_{e_2} \geq \cdots \geq w_{e_m} \geq w_{e_{m+1}} = 0$, and where $U_i = \{e_1, e_2, \ldots, e_i\}$.

BTW: this essentially Abel’s partial summation formula: Given two arbitrary sequences $\{a_n\}$ and $\{b_n\}$ with $A_n = \sum_{k=1}^{n} a_k$, we have

$$\sum_{k=m}^{n} a_k b_k = \sum_{k=m}^{n} A_k (b_k - b_{k+1}) + A_n b_{n+1} - A_{m-1} b_m$$
We can think of this as an integral over $\mathbb{R}$ of a piecewise constant function.
Choquet integral

- We can think of this as an integral over $\mathbb{R}$ of a piecewise constant function.
- First note, assuming $E$ is ordered according to $w$, then $U_i = \{e_1, e_2, \ldots, e_i\} = \{e \in E : w_e \geq w_{e_i}\}$. 
We can thing of this as an integral over \( \mathbb{R} \) of a piecewise constant function.

First note, assuming \( E \) is ordered according to \( w \), then
\[
U_i = \{ e_1, e_2, \ldots, e_i \} = \{ e \in E : w_e \geq w_{e_i} \}.
\]

For any \( w_{e_i} > \alpha \geq w_{e_{i+1}} \) we also have
\[
U_i = \{ e_1, e_2, \ldots, e_i \} = \{ e \in E : w_e > \alpha \}. \quad \text{for my sub \( \alpha \)}
\]
Choquet integral

- We can think of this as an integral over $\mathbb{R}$ of a piecewise constant function.
- First note, assuming $E$ is ordered according to $w$, then
  \[ U_i = \{e_1, e_2, \ldots, e_i\} = \{e \in E : w_e \geq w_{e_i}\}. \]
- For any $w_{e_i} > \alpha \geq w_{e_{i+1}}$ we also have
  \[ U_i = \{e_1, e_2, \ldots, e_i\} = \{e \in E : w_e > \alpha\}. \]
- Consider segmenting the real-axis at boundary points $w_{e_i}$, right most is $w_{e_1}$.

\[
\begin{array}{cccccccc}
\cdots & w_5 & w_4 & w_3 & w_2 & w_1 \\
\end{array}
\]
We can think of this as an integral over $\mathbb{R}$ of a piecewise constant function.

First note, assuming $E$ is ordered according to $w$, then
\[ U_i = \{e_1, e_2, \ldots, e_i\} = \{e \in E : w_e \geq w_{e_i}\}. \]

For any $w_{e_i} > \alpha \geq w_{e_{i+1}}$ we also have
\[ U_i = \{e_1, e_2, \ldots, e_i\} = \{e \in E : w_e > \alpha\}. \]

Consider segmenting the real-axis at boundary points $w_{e_i}$, right most is $w_{e_1}$.

A function can be defined on a segment of $\mathbb{R}$, namely
\[ w_{e_i} > \alpha \geq w_{e_{i+1}}. \] This function $F_i : [w_{e_{i+1}}, w_{e_i}) \to \mathbb{R}$ is defined as
\[ F_i(\alpha) = f(\{e \in E : w_e > \alpha\}) = f(U_i) \quad (17) \]
We can generalize this to multiple segments of $\mathbb{R}$. I.e.,

$$F(\alpha) = \begin{cases} 
0 & \text{if } \alpha > w_1 \\
 f(\{e \in E : w_e > \alpha\}) & \text{if } w_{e_i} > \alpha \geq w_{e_{i+1}} \\
0 & \text{if } w_m \geq \alpha \geq 0 
\end{cases}$$

(18)
We can generalize this to multiple segments of \( \mathbb{R} \). I.e.,

\[
F(\alpha) = \begin{cases} 
0 & \text{if } \alpha > w_1 \\
f(\{e \in E : w_e > \alpha\}) & \text{if } w_{e_i} > \alpha \geq w_{e_{i+1}} \\
0 & \text{if } w_m \geq \alpha \geq 0 
\end{cases}
\]

Visualizing this, we see that we’ve got a piecewise constant function, where the constant values are given by \( f \) evaluated on \( U_i \) for each \( i \).
Now consider the integral, with $w \in \mathbb{R}_+^E$, and normalized $f$ so that $f(\emptyset) = 0$. Recall $w_{m+1} \overset{\text{def}}{=} 0$. 

$$\tilde{f}(w) \overset{\text{def}}{=} \int_0^\infty F(\alpha) d\alpha$$

(19)
Now consider the integral, with \( w \in \mathbb{R}_E \), and normalized \( f \) so that \( f(\emptyset) = 0 \). Recall \( w_{m+1} \stackrel{\text{def}}{=} 0 \).

\[
\tilde{f}(w) \overset{\text{def}}{=} \int_{0}^{\infty} F(\alpha) d\alpha
\]

\[
= \int_{0}^{\infty} f(\{e \in E : w_e > \alpha\}) d\alpha
\]
Now consider the integral, with $w \in \mathbb{R}^E_+$, and normalized $f$ so that $f(\emptyset) = 0$. Recall $w_{m+1} \overset{\text{def}}{=} 0$.

\begin{equation}
\tilde{f}(w) \overset{\text{def}}{=} \int_0^\infty F(\alpha) \, d\alpha
\end{equation}

\begin{align}
&= \int_0^\infty f(\{e \in E : w_e > \alpha\}) \, d\alpha \\
&= \int_{w_{m+1}}^\infty f(\{e \in E : w_e > \alpha\}) \, d\alpha
\end{align}
Now consider the integral, with $w \in \mathbb{R}^E_+$, and normalized $f$ so that $f(\emptyset) = 0$. Recall $w_{m+1} \overset{\text{def}}{=} 0$.

$$\tilde{f}(w) \overset{\text{def}}{=} \int_0^{\infty} F(\alpha) d\alpha$$

$$= \int_0^{\infty} f(\{e \in E : w_e > \alpha\}) d\alpha$$

$$= \int_{w_{m+1}}^{\infty} f(\{e \in E : w_e > \alpha\}) d\alpha$$

$$= \sum_{i=1}^{m} \int_{w_i}^{w_{i+1}} f(\{e \in E : w_e > \alpha\}) d\alpha$$
Now consider the integral, with $w \in \mathbb{R}_+^E$, and normalized $f$ so that $f(\emptyset) = 0$. Recall $w_{m+1} = 0$.

$$\tilde{f}(w) \overset{\text{def}}{=} \int_0^\infty F(\alpha) d\alpha$$

$$= \int_0^\infty f(\{e \in E : w_e > \alpha\}) d\alpha \quad \checkmark$$

$$= \int_0^\infty f(\{e \in E : w_e > \alpha\}) d\alpha$$

$$= \sum_{i=1}^m \int_{w_i+1}^{w_i} f(\{e \in E : w_e > \alpha\}) d\alpha$$

$$= \sum_{i=1}^m \int_{w_i+1}^{w_i} f(U_i) d\alpha = \sum_{i=1}^m f(U_i)(w_i - w_{i+1})$$

(19) \hspace{1cm} (20) \hspace{1cm} (21) \hspace{1cm} (22) \hspace{1cm} (23)
But we saw before that $\sum_{i=1}^{m} f(U_i)(w_i - w_{i+1})$ is just the Lovász extension of a function $f$. 

Thus, we have the following definition:

Definition 3.3

Given $w \in \mathbb{R}^{E+}$, the Lovász extension (equivalently Choquet integral) may be defined as follows:

$$\tilde{f}(w) \overset{\text{def}}{=} \int_{0}^{\infty} F(\alpha) \, d\alpha \quad (24)$$

where the function $F$ is defined as before.

Note that it is not necessary in general to require $w \in \mathbb{R}^{E+}$ (i.e., we can take $w \in \mathbb{R}^{E}$) but it is a bit more involved.
But we saw before that $\sum_{i=1}^{m} f(U_i)(w_i - w_{i+1})$ is just the Lovász extension of a function $f$.

Thus, we have the following definition:

**Definition 3.3**

Given $w \in \mathbb{R}_+^E$, the Lovász extension (equivalently Choquet integral) may be defined as follows:

$$\tilde{f}(w) \overset{\text{def}}{=} \int_0^\infty F(\alpha) d\alpha$$

(24)

where the function $F$ is defined as before.
But we saw before that $\sum_{i=1}^{m} f(U_i)(w_i - w_{i+1})$ is just the Lovász extension of a function $f$.

Thus, we have the following definition:

**Definition 3.3**

Given $w \in \mathbb{R}^E_+$, the Lovász extension (equivalently Choquet integral) may be defined as follows:

$$\tilde{f}(w) \overset{\text{def}}{=} \int_{0}^{\infty} F(\alpha) d\alpha$$

(24)

where the function $F$ is defined as before.

Note that it is not necessary in general to require $w \in \mathbb{R}_+^E$ (i.e., we can take $w \in \mathbb{R}^E$) but it is a bit more involved.
For a given $w \in [0, 1]^m$, it is easy to see that we can also define the Lovász extension as
\[
\tilde{f}(w) = \mathbb{E}[f(e \in E : w(e_i) > \alpha)]
\] (25)
where $\alpha$ is a uniform random variable in $[0, 1]$. 
Lovász extension

- For a given $w \in [0, 1]^m$, it is easy to see that we can also define the Lovász extension as
  \[ \tilde{f}(w) = \mathbb{E}[f(e \in E : w(e_i) > \alpha)] \] (25)
  where $\alpha$ is uniform random variable in $[0, 1]$.

- The convexity of the Lovász extension, the ease of minimizing convex functions, and the fact that we can recover $f$ from $\tilde{f}$ via $f(A) = \tilde{f}(\mathbf{1}_A)$ corresponds to why SFM is possible in polynomial time (which was first shown by Grötschel, Lovász, and Schrijver in 1988 as part of their Ellipsoid method.)
We want to produce some notion of generalized aggregation function. Something along the lines of this:

\[ AG(x) = \sum_{A \subseteq E} x(A)w_A = \sum_{A \subseteq E} x(A)AG(1_A) \]  

(26)

how does this correspond to Lovász extension?
We want to produce some notion of generalized aggregation function.

\[ AG(x) = \sum_{A \subseteq E} x(A)w_A = \sum_{A \subseteq E} x(A)AG(1_A) \]  

(26)

how does this correspond to Lovász extension?

Let us partition the hypercube \([0, 1]^m\) into \(q\) polytopes defined by a set of vertices \(V_1, V_2, \ldots, V_q\). This forms a “triangulation” of the hypercube.
Choquet integral and aggregation

- We want to produce some notion of generalized aggregation function.
  \[
  AG(x) = \sum_{A \subseteq E} x(A)w_A = \sum_{A \subseteq E} x(A)AG(1_A)
  \]
  (26)
  how does this correspond to Lovász extension?

- Let us partition the hypercube \([0, 1]^m\) into \(q\) polytopes defined by a set of vertices \(V_1, V_2, \ldots, V_q\). This forms a “triangulation” of the hypercube.

- For any \(x \in [0, 1]^m\) there is a \(V(x) = V_j\) for some \(j\) such that \(x \in \text{conv}(V(x))\).
Choquet integral and aggregation

For $x \in [0, 1]^m$, let us define the (unique) coefficients $\alpha_0^x(A)$ and $\alpha_i^x(A)$ so that $x$ can be represented as a weighted combination of elements in $\mathcal{V}(x)$. Note that many of these coefficients might be zero.
For $x \in [0, 1]^m$, let us define the (unique) coefficients $\alpha_0^x(A)$ and $\alpha_i^x(A)$ so that $x$ can be represented as a weighted combination of elements in $\mathcal{V}(x)$. Note that many of these coefficients might be zero.

From this, we can define an aggregation function of the form

$$AG(x) \triangleq \sum_{A: \mathbf{1}_A \in \mathcal{V}(x)} \left( \alpha_0^x(A) + \sum_{i=1}^m \alpha_i^x(A)x_i \right) AG(\mathbf{1}_A)$$

(27)

$$WAVL(x) = \bigvee_i x_i \cdot WAVL(\mathbf{1}_{e_i})$$
We can define a canonical triangulation of the hypercube in terms of permutations of the coordinates. I.e., given some permutation $\sigma$, define

$$\text{conv}(V_\sigma) = \{x \in [0,1]^n | x_{\sigma(1)} \geq x_{\sigma(2)} \geq \cdots \geq x_{\sigma(m)}\}$$

(28)

Then these $m!$ blocks of the partition are called the canonical partitions of the hypercube. In this case, we have:

![Diagram](image-url)
We can define a canonical triangulation of the hypercube in terms of permutations of the coordinates. I.e., given some permutation $\sigma$, define

$$\text{conv}(V_{\sigma}) = \{x \in [0, 1]^n | x_{\sigma(1)} \geq x_{\sigma(2)} \geq \cdots \geq x_{\sigma(m)}\}$$

(28)

Then these $m!$ blocks of the partition are called the canonical partitions of the hypercube. In this case, we have:

**Proposition 3.4**

The above linear interpolation using the canonical partition yields the Lovász extension.
Choquet integral and aggregation
Polymatroid extreme points

- The greedy algorithm does more than solve $\max(wx : x \in P_f)$. We can use it to generate vertices of polymatroidal polytopes.
Polymatroid extreme points

- The greedy algorithm does more than solve \( \max(wx : x \in P_f) \). We can use it to generate vertices of polymatroidal polytopes.
- First, consider \( P_f \) and also \( C_f \overset{\text{def}}{=} \{ x : x \in \mathbb{R}_+^E, x(e) \leq f(e) \} \)
The greedy algorithm does more than solve $\max(wx : x \in P_f)$. We can use it to generate vertices of polymatroidal polytopes.

First, consider $P_f$ and also $C_f \overset{\text{def}}{=} \{x : x \in \mathbb{R}^E_+, x(e) \leq f(e)\}$

Then $P_f \subseteq C_f$ since $f(A) = \sum_i (f(A_{i-1} + a_i) - f(A_{i-1})) \leq \sum_i f(a_i)$ for some order of elements in $A$. 
Polymatroid extreme points

The greedy algorithm does more than solve \( \max(wx : x \in P_f) \). We can use it to generate vertices of polymatroidal polytopes.

First, consider \( P_f \) and also \( C_f \) defined as \( \{ x : x \in \mathbb{R}^E_+, x(e) \leq f(e) \} \)

Then \( P_f \subseteq C_f \) since \( f(A) = \sum_i(f(A_{i-1} + a_i) - f(A_{i-1})) \leq \sum_i f(a_i) \) for some order of elements in \( A \).
Polymatroid extreme points

- Given $w \in R^E_+$, we can choose any $e \in E$ to be such that $w(e) > w(f)$ for $f \in E \setminus \{e\}$. 

Thus, intuitively, any initial vertex of the polytope can be obtained by advancing along the corresponding axis. The base polytope is defined as the extreme face of $P_f$. I.e., $B_f = P_f \cap \{x \in R^E_+ : x(E) = f(E)\}$.

Also, intuitively, we can continue advancing along the skeletal edges of the polytope to reach any other vertex, given the appropriate ordering. If we advance in all dimensions, we'll reach a vertex in $B_f$, and if we advance only in some dimensions, we'll reach a vertex in $P_f \setminus B_f$. We formalize this next:
Polymatroid extreme points

- Given $w \in \mathcal{R}_+^E$, we can choose any $e \in E$ to be such that $w(e) > w(f)$ for $f \in E \setminus \{e\}$.

- Thus, intuitively, any initial vertex of the polytope can be obtained by advancing along the corresponding axis.
Polymatroid extreme points

- Given $w \in \mathbb{R}^E_+$, we can chose any $e \in E$ to be such that $w(e) > w(f)$ for $f \in E \setminus \{e\}$.
- Thus, intuitively, any initial vertex of the polytope can be obtained by advancing along the corresponding axis.
- The base polytope is defined as the extreme face of $P_f$. I.e.,
  \[
  B_f = P_f \cap \left\{ x \in \mathbb{R}^E_+ : x(E) = f(E) \right\}
  \]  
  (29)

Also, intuitively, we can continue advancing along the skeletal edges of the polytope to reach any other vertex, given the appropriate ordering. If we advance in all dimensions, we'll reach a vertex in $B_f$, and if we advance only in some dimensions, we'll reach a vertex in $P_f \setminus B_f$. We formalize this next:
Polymatroid extreme points

- Given $w \in \mathcal{R}^E_+$, we can choose any $e \in E$ to be such that $w(e) > w(f)$ for $f \in E \setminus \{e\}$.
- Thus, intuitively, any initial vertex of the polytope can be obtained by advancing along the corresponding axis.
- The base polytope is defined as the extreme face of $P_f$. I.e.,
  \[ B_f = P_f \cap \left\{ x \in \mathbb{R}^E_+ : x(E) = f(E) \right\} \]  
  (29)
Polymatroid extreme points

- Given $w \in \mathcal{R}_+^E$, we can chose any $e \in E$ to be such that $w(e) > w(f)$ for $f \in E \setminus \{e\}$.
- Thus, intuitively, any initial vertex of the polytope can be obtained by advancing along the corresponding axis.
- The base polytope is defined as the extreme face of $P_f$. I.e.,
  \[ B_f = P_f \cap \left\{ x \in \mathbb{R}_+^E : x(E) = f(E) \right\} \tag{29} \]
- Also, intuitively, we can continue advancing along the skeletal edges of the polytope to reach any other vertex, given the appropriate ordering. If we advance in all dimensions, we’ll reach a vertex in $B_f$, and if we advance only in some dimensions, we’ll reach a vertex in $P_f \setminus B_f$. 
Polymatroid extreme points

- Given $w \in \mathcal{R}_+^E$, we can choose any $e \in E$ to be such that $w(e) > w(f)$ for $f \in E \setminus \{e\}$.

- Thus, intuitively, any initial vertex of the polytope can be obtained by advancing along the corresponding axis.

- The base polytope is defined as the extreme face of $P_f$. I.e.,

$$B_f = P_f \cap \left\{ x \in \mathbb{R}_+^E : x(E) = f(E) \right\}$$

(29)

- Also, intuitively, we can continue advancing along the skeletal edges of the polytope to reach any other vertex, given the appropriate ordering. If we advance in all dimensions, we’ll reach a vertex in $B_f$, and if we advance only in some dimensions, we’ll reach a vertex in $P_f \setminus B_f$.

- We formalize this next:
Polymatroid extreme points

- Given any arbitrary order of \( E = (e_1, e_2, \ldots, e_m) \), define \( E_i = (e_1, e_2, \ldots, e_i) \).
Polymatroid extreme points

- Given any arbitrary order of $E = (e_1, e_2, \ldots, e_m)$, define $E_i = (e_1, e_2, \ldots, e_i)$.

- A vector $x$ is generated by $E_i$ using the greedy procedure as follows
  \[ x(e_j) = f(E_j) - f(E_{j-1}) \quad \text{for} \quad 2 \leq j \leq i \]  
  \[ x(e) = 0 \quad \text{for} \quad e \in E \setminus E_i \]  
  \[ x(e_1) = f(E_1) \]
Polymatroid extreme points

- Given any arbitrary order of $E = (e_1, e_2, \ldots, e_m)$, define $E_i = (e_1, e_2, \ldots, e_i)$.

- A vector $x$ is generated by $E_i$ using the greedy procedure as follows
  \[ x(e_j) = f(E_j) - f(E_{j-1}) \text{ for } 1 \leq j \leq i \]  
  \[ x(e) = 0 \text{ for } e \in E \setminus E_i \]  
  \[ (30) \]  
  \[ (31) \]

- An extreme point of $P_f$ is a point that is not a convex combination of two other distinct points in $P_f$. Equivalently, an extreme point corresponds to setting certain inequalities in the specification of $P_f$ to be equalities, so that there is a unique single point solution.
Polymatroid extreme points

Theorem 4.1

For a given ordering $E = (e_1, \ldots, e_m)$ of $E$ and a given $E_i$ and $x$ generated by $E_i$ using the greedy procedure, then $x$ is an extreme point of $P_f$. 

Proof.

We already saw that $x \in P_f$ (in Lecture 11, proof of Theorem 4.2). To show that $x$ is an extreme point of $P_f$, note that it is the unique solution of the following system of equations:

1. $x(E_j) = f(E_j)$ for $1 \leq j \leq i$ (32)
2. $x(e) = 0$ for $e \in E \setminus E_i$ (33)
Theorem 4.1

For a given ordering $E = (e_1, \ldots, e_m)$ of $E$ and a given $E_i$ and $x$ generated by $E_i$ using the greedy procedure, then $x$ is an extreme point of $P_f$

Proof.

1. We already saw that $x \in P_f$ (in Lecture 11, proof of Theorem 4.2).
Polymatroid extreme points

Theorem 4.1

For a given ordering $E = (e_1, \ldots, e_m)$ of $E$ and a given $E_i$ and $x$ generated by $E_i$ using the greedy procedure, then $x$ is an extreme point of $P_f$

Proof.

- We already saw that $x \in P_f$ (in Lecture 11, proof of Theorem 4.2).
- To show that $x$ is an extreme point of $P_f$, note that it is the unique solution of the following system of equations

\[
x(E_j) = f(E_j) \text{ for } 1 \leq j \leq i \tag{32}
\]

\[
x(e) = 0 \text{ for } e \in E \setminus E_i \tag{33}
\]
Polymatroid extreme points

- As an example, we have $x(E_1) = x(e_1) = f(e_1)$
Polymatroid extreme points

- As an example, we have \( x(E_1) = x(e_1) = f(e_1) \)
- \( x(E_2) = x(e_1) + x(e_2) = f(e_1, e_2) \) so
  \[ x(e_2) = f(e_1, e_2) - x(e_1) = f(e_1, e_2) - f(e_1). \]
Polymatroid extreme points

As an example, we have $x(E_1) = x(e_1) = f(e_1)$

$x(E_2) = x(e_1) + x(e_2) = f(e_1, e_2)$ so

$x(e_2) = f(e_1, e_2) - x(e_1) = f(e_1, e_2) - f(e_1)$.

$x(E_3) = x(e_1) + x(e_2) + x(e_3) = f(e_1, e_2, e_3)$ so

$x(e_3) = f(e_1, e_2, e_3) - x(e_2) - x(e_1) = f(e_1, e_2, e_3) - f(e_1, e_2)$
Polymatroid extreme points

- As an example, we have \( x(E_1) = x(e_1) = f(e_1) \)
- \( x(E_2) = x(e_1) + x(e_2) = f(e_1, e_2) \) so
  \( x(e_2) = f(e_1, e_2) - x(e_1) = f(e_1, e_2) - f(e_1). \)
- \( x(E_3) = x(e_1) + x(e_2) + x(e_3) = f(e_1, e_2, e_3) \) so
  \( x(e_3) = f(e_1, e_2, e_3) - x(e_2) - x(e_1) = f(e_1, e_2, e_3) - f(e_1, e_2) \)
- And so on ...
Polymatroid extreme points

- As an example, we have $x(E_1) = x(e_1) = f(e_1)$
- $x(E_2) = x(e_1) + x(e_2) = f(e_1, e_2)$ so
  $x(e_2) = f(e_1, e_2) - x(e_1) = f(e_1, e_2) - f(e_1)$.
- $x(E_3) = x(e_1) + x(e_2) + x(e_3) = f(e_1, e_2, e_3)$ so
  $x(e_3) = f(e_1, e_2, e_3) - x(e_2) - x(e_1) = f(e_1, e_2, e_3) - f(e_1, e_2)$
- And so on ...

Also, since $x \in P_f$, for each $i$, we see that,

$$x(E_i) = f(E_i) \quad \forall i \in \text{index}$$  \hspace{1cm} (34)

$$x(A) \leq f(A), \forall A \subseteq E \quad \text{s.t.} \quad x \in P_+$$  \hspace{1cm} (35)
Polymatroid extreme points

- As an example, we have $x(E_1) = x(e_1) = f(e_1)$
- $x(E_2) = x(e_1) + x(e_2) = f(e_1, e_2)$ so $x(e_2) = f(e_1, e_2) - x(e_1) = f(e_1, e_2) - f(e_1)$.
- $x(E_3) = x(e_1) + x(e_2) + x(e_3) = f(e_1, e_2, e_3)$ so $x(e_3) = f(e_1, e_2, e_3) - x(e_2) - x(e_1) = f(e_1, e_2, e_3) - f(e_1, e_2)$.
- And so on ...

- Also, since $x \in P_f$, for each $i$, we see that,
  \[ x(E_i) = f(E_i) \quad (34) \]
  \[ x(A) \leq f(A), \forall A \subseteq E \quad (35) \]

- Thus, the greedy procedure provides a modular function lower bound on $f$ that is tight on all points $E_i$ in the order.
Polymatroid extreme points

some examples
Polymatroid extreme points

Moreover, we have

**Corollary 4.2**

*If* $x$ *is an extreme point of* $P_f$ *and* $A \subseteq E$ *is given such that*

$$\{ e \in E : x(e) \neq 0 \} \subseteq B \subseteq \bigcup (A : x(A) = f(A)),$$

*then* $x$ *is generated using greedy by some ordering of* $A$. 
Moreover, we have

**Corollary 4.2**

*If* \( x \) *is an extreme point of* \( P_f \) *and* \( A \subseteq E \) *is given such that*  
\[ \{ e \in E : x(e) \neq 0 \} \subseteq B \subseteq \bigcup (A : x(A) = f(A)), \]
*then* \( x \) *is generated using greedy by some ordering of* \( A \).

- Note, \( \text{cl}(x) = \bigcup (A : x(A) = f(A)) \) *is the closure of* \( x \) *(recall that sets* \( A \) *such that* \( x(A) = f(A) \) *are called tight, and such sets are closed under union and intersection, see Lecture 7, in proof of Theorem 4.3, starting Eq. 50).*
Moreover, we have

**Corollary 4.2**

*If* \( x \) *is an extreme point of* \( P_f \) *and* \( A \subseteq E \) *is given such that*

\[ \{ e \in E : x(e) \neq 0 \} \subseteq B \subseteq \bigcup (A : x(A) = f(A)) \],

*then* \( x \) *is generated using greedy by some ordering of* \( A \).

- Note, \( \text{cl}(x) = \bigcup (A : x(A) = f(A)) \) *is the closure of* \( x \) *(recall that sets* \( A \) *such that* \( x(A) = f(A) \) *are called tight, and such sets are closed under union and intersection, see Lecture 7, in proof of Theorem 4.3, starting Eq. 50).*

- Thus, \( \text{cl}(x) \) *is a tight set.*
Moreover, we have

**Corollary 4.2**

*If $x$ is an extreme point of $P_f$ and $A \subseteq E$ is given such that $\{e \in E : x(e) \neq 0\} \subseteq B \subseteq \bigcup (A : x(A) = f(A))$, then $x$ is generated using greedy by some ordering of $A$.***

- Note, $\text{cl}(x) = \bigcup (A : x(A) = f(A))$ is the closure of $x$ (recall that sets $A$ such that $x(A) = f(A)$ are called tight, and such sets are closed under union and intersection, see Lecture 7, in proof of Theorem 4.3, starting Eq. 50).
- Thus, $\text{cl}(x)$ is a tight set.
- Also, $\text{supp}(x) = \{e \in E : x(e) \neq 0\}$ is called the support of $x$. 
Moreover, we have

**Corollary 4.2**

*If* $x$ *is an extreme point of* $P_f$ *and* $A \subseteq E$ *is given such that*
\[
\{ e \in E : x(e) \neq 0 \} \subseteq B \subseteq \bigcup (A : x(A) = f(A)),
\]
*then* $x$ *is generated using greedy by some ordering of* $A$.

- Note, $\text{cl}(x) = \bigcup (A : x(A) = f(A))$ *is the closure of* $x$ *recall that sets* $A$ *such that* $x(A) = f(A)$ *are called tight, and such sets are closed under union and intersection, see Lecture 7, in proof of Theorem 4.3, starting Eq. 50)*.
- Thus, $\text{cl}(x)$ *is a tight set.*
- Also, $\text{supp}(x) = \{ e \in E : x(e) \neq 0 \}$ *is called the support of* $x$.
- For arbitrary $x$, $\text{supp}(x)$ *is not tight, but for an extreme point, $\text{supp}(x)$ is.*
Sources for Today’s Lecture