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.

Class Road Map

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):
- L11 (5/6): On SFM, polymatroid member & greedy, Lovász ext.
- L12 (5/11): Lovász ext. + polymatroid props.
- 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).

Towards SFM

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

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 $r(A) - \frac{1}{k}\mathbf{1}(A)$.
- 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 $r(A) - \frac{1}{k}\mathbf{1}(A)$.
- 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.

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

$g \Rightarrow f(A) = g(A) - \sum [g(E) - g(E \setminus a)]$

$\sqrt{|M|} - |A|$

$\Rightarrow f(E) - f(E \setminus e) = 0 \ \forall e \equiv$ totally normalized

- 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:

$f(E) - f(E \setminus e)$

=smallest gain of e.

  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.

$|E|$

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 $Ax \geq b$"?

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

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 $P_f = \left\{x \in \mathbb{R}_+^E : x(A) \leq f(A), \;\; \forall A \subseteq E\right\}$.
- The membership problem, "given an $x \in \mathbb{R}_+^E$, is $x \in P_f$?", is a special case of this unsolved problem.

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 $P_f = \left\{ x \in \mathbb{R}_+^E : x(A) \leq f(A), \ \forall A \subseteq E \right\}$.
- 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 $0 \leq f(A) - x(A), \forall A \subseteq E$.

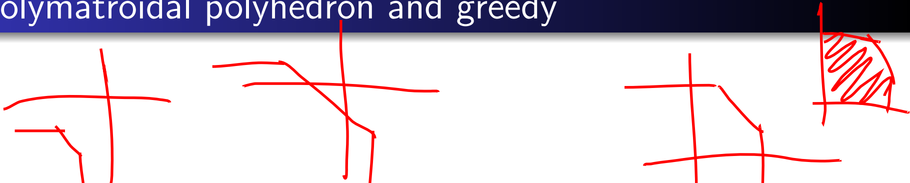
Testing membership in polymatroids

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

Polymatroidal polyhedron and greedy



- We have a result very similar to that of matroids.

Theorem 2.1

If $f : 2^E \to \mathbb{R}_+$ is given, and P is a polytope in $\mathbb{R}_+^E$ of the form $P = \{x \in \mathbb{R}_+^E : x(A) \le f(A), \forall A \subseteq E\}$, 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).

given any $f(\,)$, extended submodular polyhedron as

$$EP_f = \{ x \in \mathbb{R}^E : x(A) \le f(A) \ \forall A \subseteq E \}$$

$$x(\emptyset) = 0$$

An extension of f

- We may consider the optimization a function $\tilde{f} : \mathbb{R}^E \to \mathbb{R}$ as

$$\tilde{f}(w) = \max(wx : x \in P_f) \tag{1}$$

An extension of f

- We may consider the optimization a function $\tilde{f} : \mathbb{R}^E \to \mathbb{R}$ as

$$\tilde{f}(w) = \max(wx : x \in P_f) \tag{1}$$

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

An extension of f

- We may consider the optimization a function $\tilde{f} : \mathbb{R}^E \to \mathbb{R}$ as
$$\tilde{f}(w) = \max(wx : x \in P_f) \tag{1}$$
- Then, for any w, from the above theorem, we can compute this function using the greedy algorithm.
- That is, we have, for submodular f,
$$\begin{aligned}\tilde{f}(w) &= \max(wx : x \in P_f) &(2)\\ &= \sum_{i=1}^{m} w(e_i)(f(U_i) - f(U_{i-1})) &(3)\\ &= w(e_m)f(U_m) + \sum_{i=1}^{m-1}(w(e_i) - w(e_{i+1}))f(U_i) &(4)\end{aligned}$$
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

$$\tilde{f}(w) = \max(wx : x \in P_f) \tag{5}$$

- Therefore, if f is a submodular function, we can write
$$\tilde{f}(w) = \sum_{i=1}^{m} \lambda_i f(U_i) \tag{6}$$
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

$$\tilde{f}(w) = \max(wx : x \in P_f) \tag{5}$$

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

$$\tilde{f}(w) = \sum_{i=1}^{m} \lambda_i f(U_i) \tag{6}$$

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.

- 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 arbitrary 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

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

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

Lovasz: take a w, write w as $w = \sum_{i=1}^{m} \lambda_i 1_{U_i}$

$\Rightarrow$ unique values $\{\lambda_i, U_i\}_{i=1}^{m}$

from this, form $\hat{f}(w) = \sum_{i=1}^{m} \lambda_i f(U_i)$

An extension of f

- On the other hand, for any f (even not submodular), we can define an extension in this way, with
$$\tilde{f}(w) = \sum_{i=1}^{m} \lambda_i f(U_i) \tag{7}$$
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
$$\tilde{f}(w) = \sum_{i=1}^{m} \lambda_i f(U_i) \tag{7}$$
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.
- This "extension" of f, in any case, is called the Lovász extension of f.

Lovász Extension, submodularity and convexity

Theorem 3.1

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

Proof.

seen in last lecture. $\square$

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

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 $w(e_m)f(U_m) + \sum_{i=1}^{m-1}(w(e_i) - w(e_{i+1}))f(U_i) = \sum_{i=1}^{m} \lambda_i f(U_i)$: Simple Algebra although r.h.s. first appeared in Lovász-1983
  = Abel summation formula (see later in this lecture).

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.

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

Who's Extension? Who did what?

Lets have a (somewhat silly) credit assignment vote.

- Edmonds: 0
- Lovász: 0
- Edmonds-Lovász: 70

Who's Extension? Who did what?

Lets 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.

Sugeno integral
fuzzy measure
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 μ of sets. E.g., given measurable function f, we can define
$$\int_X f du = \sup I_X(s) \tag{8}$$
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 $s(x) = \sum_{i=1}^n c_i I_{X_i}(x)$ and where $I_{X_i}(x)$ 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.

Integration

- 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

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

Integration

- 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

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

$\forall b \in \{0,1\}^n$

- Consider $\mathbf{1}_e$ for $e \in E$, we have

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

/ $w_b = w(e)$ if $|b| = 1$ else $w_b = 0$.

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., $\{\mathbf{1}_e : e \in E\}$. Moreover, we are interpolating as in

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

Integration

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

$$\text{AG}(\mathbf{1}_A) = w_A \tag{12}$$

Integration

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

$$\text{AG}(\mathbf{1}_A) = w_A \tag{12}$$

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

$$\text{AG}(x) = \sum_{A \subseteq E} x(A) w_A = \underbrace{\sum_{A \subseteq E} x(A)\text{AG}(\mathbf{1}_A)}_{\text{still then some sort of interpolation.}} \tag{13}$$

Integration

- More complex aggregation functions can be constructed by defining the aggregation function on all vertices of the hypercube. I.e., for each $\mathbf{1}_A : A \subseteq E$ we might have (for all $A \subseteq E$):
$$\text{AG}(\mathbf{1}_A) = w_A \tag{12}$$
- What then might $\text{AG}(x)$ be for some $x \in \mathbb{R}^E$? Our weighted average functions might look something more like
$$\text{AG}(x) = \sum_{A \subseteq E} x(A) w_A = \sum_{A \subseteq E} x(A) \text{AG}(\mathbf{1}_A) \tag{13}$$
- Set function $f : 2^E \to \mathbb{R}$ is a game if f is normalized $f(\emptyset) = 0$.

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 $A \subseteq B$.

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 $f : \{0,1\}^m \to \mathbb{R}$.

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 $A = \{e \in E : x_e = 1\}$.

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 $f(A) = f_b(\mathbf{1}_A)$. We can also often relax f_b to any $x \in [0,1]^m$.

Choquet integral

Definition 3.2

1954

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

$$C_f(w) = \sum_{i=1}^{m} (w_{e_i} - w_{e_{i+1}}) f(U_i) \tag{14}$$

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:

$$C_f(w) = \sum_{i=1}^{m} w(e_i)(f(U_i) - f(U_{i-1})) \tag{15}$$

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

Choquet integral

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

$$C_f(w) = \sum_{i=1}^{m} (w_{e_i} - w_{e_{i+1}}) f(U_i) \tag{14}$$

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

- 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

$$\sum_{k=m}^{n} a_k b_k = \sum_{k=m}^{n} A_k (b_k - b_{k+1}) + A_n b_{n+1} - A_{m-1} b_m \tag{16}$$

Choquet integral

- We can thing of this as an integral over $\mathbb{R}$ of a piecewise constant function.

Choquet integral

- 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 $U_i = \{e_1, e_2, \dots, e_i\} = \{e \in E : w_e \geq w_{e_i}\}$.

Choquet integral

- 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 $U_i = \{e_1, e_2, \dots, e_i\} = \{e \in E : w_e \geq w_{e_i}\}$.
- For any $w_{e_i} > \alpha \geq w_{e_{i+1}}$ we also have $U_i = \{e_1, e_2, \dots, e_i\} = \{e \in E : w_e > \alpha\}$. for any such α,

Choquet integral

- 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 $U_i = \{e_1, e_2, \dots, e_i\} = \{e \in E : w_e \geq w_{e_i}\}$.
- For any $w_{e_i} > \alpha \geq w_{e_{i+1}}$ we also have $U_i = \{e_1, e_2, \dots, e_i\} = \{e \in E : w_e > \alpha\}$.
- Consider segmenting the real-axis at boundary points w_{e_i}, right most is w_{e_1}.

Choquet integral

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

$$F_i(\alpha) = f(\{e \in E : w_e > \alpha\}) = f(U_i) \tag{17}$$

Choquet integral

- We can generalize this to multiple segments of $\mathbb{R}$. I.e., for now, take $w \in \mathbb{R}_+^E$

$F : \mathbb{R}_+ \to \mathbb{R}$

$$F(\alpha) = \begin{cases} 0 & \text{if } \alpha > w_1 \\ f(\{e \in E : w_e > \alpha\}) & \text{if } w_{e_i} > \alpha \geq w_{e_{i+1}} \\ 0 & \text{if } w_m \geq \alpha \geq 0 \end{cases} \quad (18)$$

$$= \begin{cases} 0 & \text{if } \alpha > w_1 \text{, or } w_m \geq \alpha \geq 0 \\ f(U_i) & \text{if } w_{e_i} > \alpha \geq w_{e_{i+1}} \end{cases}$$

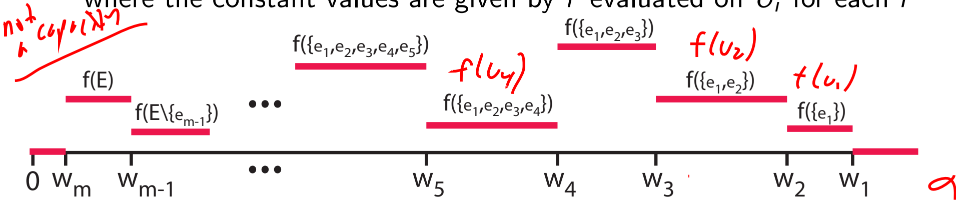
Choquet integral

- We can generalize this to multiple segments of $\mathbb{R}$. I.e.,

$$F(\alpha) = \begin{cases} 0 & \text{if } \alpha > w_1 \\ f(\{e \in E : w_e > \alpha\}) & \text{if } w_{e_i} > \alpha \geq w_{e_{i+1}} \\ 0 & \text{if } w_m \geq \alpha \geq 0 \end{cases} \tag{18}$$

not really necessary

- 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



$F(\alpha) = \begin{cases} f(\{e \in E : w_e > \alpha\}) & \alpha \geq w_m \\ 0 & 0 \leq \alpha < w_m \end{cases}$

Choquet integral

$f(\emptyset) = 0$

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

$$\tilde{f}(w) \stackrel{\text{def}}{=} \int_0^\infty F(\alpha) d\alpha \tag{19}$$

Choquet integral

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

$$\tilde{f}(w) \stackrel{\text{def}}{=} \int_0^\infty F(\alpha) d\alpha \tag{19}$$

$$= \int_0^\infty f(\{e \in E : w_e > \alpha\}) d\alpha \tag{20}$$

Choquet integral

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

$$\tilde{f}(w) \stackrel{\text{def}}{=} \int_0^\infty F(\alpha) d\alpha \tag{19}$$

$$= \int_0^\infty f(\{e \in E : w_e > \alpha\}) d\alpha \tag{20}$$

$$= \int_{w_{m+1}}^\infty f(\{e \in E : w_e > \alpha\}) d\alpha \tag{21}$$

Choquet integral

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

$$\tilde{f}(w) \stackrel{\text{def}}{=} \int_0^\infty F(\alpha) d\alpha \tag{19}$$

$$= \int_0^\infty f(\{e \in E : w_e > \alpha\}) d\alpha \tag{20}$$

$$= \int_{w_{m+1}}^\infty f(\{e \in E : w_e > \alpha\}) d\alpha \tag{21}$$

$$= \sum_{i=1}^m \int_{w_{i+1}}^{w_i} f(\{e \in E : w_e > \alpha\}) d\alpha \tag{22}$$

Choquet integral

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

$$\tilde{f}(w) \stackrel{\text{def}}{=} \int_0^\infty F(\alpha) d\alpha \tag{19}$$

$$= \int_0^\infty f(\{e \in E : w_e > \alpha\}) d\alpha \tag{20}$$

$$= \int_{w_{m+1}}^\infty f(\{e \in E : w_e > \alpha\}) d\alpha \tag{21}$$

$$= \sum_{i=1}^m \int_{w_{i+1}}^{w_i} f(\{e \in E : w_e > \alpha\}) d\alpha \tag{22}$$

$$= \sum_{i=1}^m \int_{w_{i+1}}^{w_i} f(U_i) d\alpha = \sum_{i=1}^m f(U_i)(w_i - w_{i+1}) \tag{23}$$

Choquet integral

- 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.

Choquet integral

- 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

Edmonds-

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

$$\tilde{f}(w) \stackrel{\text{def}}{=} \int_0^\infty F(\alpha) d\alpha \tag{24}$$

where the function F is defined as before.

Choquet integral

- 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:

$$\tilde{f}(w) \stackrel{\text{def}}{=} \int_0^\infty F(\alpha) d\alpha \tag{24}$$

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.

Lovász extension

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

Choquet integral and aggregation

- We want to produce some notion of generalized aggregation function. something along the lines of this

$$\mathrm{AG}(x) = \sum_{A \subseteq E} x(A) w_A = \sum_{A \subseteq E} x(A)\mathrm{AG}(\mathbf{1}_A) \tag{26}$$

how does this correspond to Lovász extension?

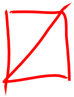
Choquet integral and aggregation

- We want to produce some notion of generalized aggregation function.

$$\mathrm{AG}(x) = \sum_{A \subseteq E} x(A) w_A = \sum_{A \subseteq E} x(A) \mathrm{AG}(\mathbf{1}_A) \tag{26}$$

how does this correspond to Lovász extension?

- Let us partition the hypercube $[0,1]^m$ into q polytopes defined by a set of vertices $\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_q$. This forms a "triangulation" of the hypercube.



Choquet integral and aggregation

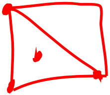
- We want to produce some notion of generalized aggregation function.

$$\mathrm{AG}(x) = \sum_{A \subseteq E} x(A) w_A = \sum_{A \subseteq E} x(A) \mathrm{AG}(\mathbf{1}_A) \tag{26}$$

  how does this correspond to Lovász extension?
- Let us partition the hypercube $[0,1]^m$ into q polytopes defined by a set of vertices $\mathcal{V}_1, \mathcal{V}_2, \dots, \mathcal{V}_q$. This forms a "triangulation" of the hypercube.
- For any $x \in [0,1]^m$ there is a $\mathcal{V}(x) = \mathcal{V}_j$ for some j such that $x \in \mathrm{conv}(\mathcal{V}(x))$.

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 coefficient might be zero.



$\alpha_0^x, \alpha_i^x(A)$ $i \in \{1, \dots, m\}$

$A \in \mathcal{V}(x)$

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 coefficient might be zero.
- From this, we can define an aggregation function of the form

$$\text{AG}(x) \stackrel{\Delta}{=} \sum_{A:\mathbf{1}_A \in \mathcal{V}(x)} \left(\alpha_0^x(A) + \sum_{i=1}^m \alpha_i^x(A) x_i \right) \text{AG}(\mathbf{1}_A) \tag{27}$$

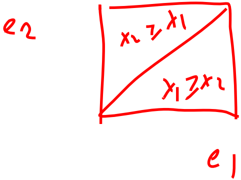
$$\text{WAVG}(x) = \sum_i x_i \,\text{WAVG}(\mathbf{1}_{e_i})$$

Choquet integral and aggregation

- We can define a canonical triangulation of the hypercube in terms of permutations of the coordinates. I.e., given some permutation σ, define

$$\text{conv}(\mathcal{V}_\sigma) = \left\{ x \in [0,1]^n \middle| x_{\sigma(1)} \geq x_{\sigma(2)} \geq \cdots \geq x_{\sigma(m)} \right\} \tag{28}$$

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



Choquet integral and aggregation

- We can define a canonical triangulation of the hypercube in terms of permutations of the coordinates. I.e., given some permutation σ, define

$$\text{conv}(\mathcal{V}_\sigma) = \left\{x \in [0,1]^n | x_{\sigma(1)} \geq x_{\sigma(2)} \geq \cdots \geq x_{\sigma(m)}\right\} \tag{28}$$

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

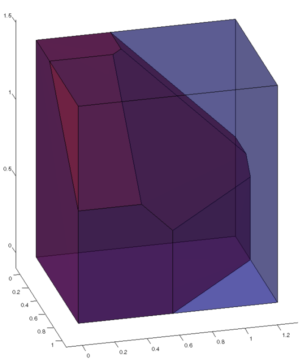
- The greedy algorithm does more than solve $\max(wx : x \in P_f)$. We can use it to generate vertices of polymatroidal polytopes.
- First, consider P_f and also $C_f \stackrel{\text{def}}{=} \{x : x \in \mathbb{R}_+^E, x(e) \leq f(e)\}$

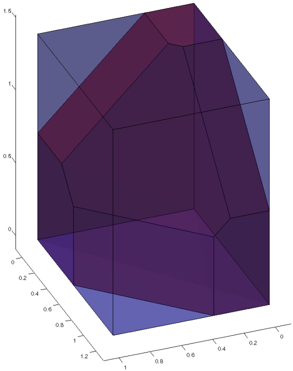
Polymatroid extreme points

- 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 $C_f \stackrel{\text{def}}{=} \{x : x \in \mathbb{R}_+^E, x(e) \leq f(e)\}$
- 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 $C_f \overset{\text{def}}{=} \{x : x \in \mathbb{R}_+^E, x(e) \leq f(e)\}$
- 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

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

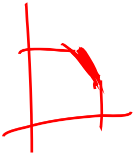
Polymatroid extreme points

- 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.

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.,

$$B_f = P_f \cap \left\{ x \in \mathbb{R}_+^E : x(E) = f(E) \right\} \tag{29}$$



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.,

$$B_f = P_f \cap \left\{x \in \mathbb{R}_+^E : x(E) = f(E)\right\} \tag{29}$$

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 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 $P_f \setminus B_f$.
- 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

$$x(e_j) = f(E_j) - f(E_{j-1}) \text{ for } \cancel{1}\,2 \le j \le i \tag{30}$$

$$x(e) = 0 \text{ for } e \in E \setminus E_i \tag{31}$$

$x(e_1) = f(E_1)$

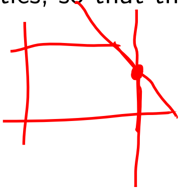
Polymatroid extreme points

- 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

$$x(e_j) = f(E_j) - f(E_{j-1}) \text{ for } 1 \leq j \leq i \tag{30}$$

$$x(e) = 0 \text{ for } e \in E \setminus E_i \tag{31}$$

- 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

Polymatroid extreme points

Theorem 4.1

For a given ordering $E = (e_1, \dots, e_m)$ 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).

$\square$

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 the greedy x $x \in P_f$, for each i, we see that,

$$x(E_i) = f(E_i) \qquad \forall i \in \text{greedy} \tag{34}$$

$$x(A) \leq f(A), \forall A \subseteq E \qquad \text{since } x \in P_f \tag{35}$$

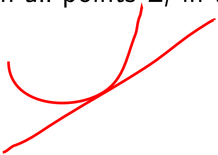
Polymatroid extreme points

- 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,

$$x(E_i) = f(E_i) \tag{34}$$

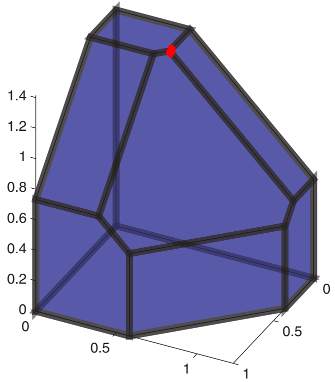
$$x(A) \leq f(A), \forall A \subseteq E \tag{35}$$

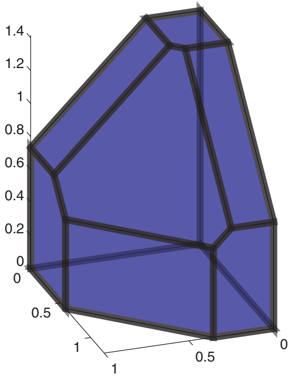
- Thus, the greedy procedure provides a modular function lower bound on f that is tight on all points E_i in the order.

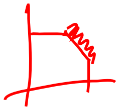


Polymatroid extreme points

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 $\{e \in E : x(e) \neq 0\} \subseteq B \subseteq \cup(A : x(A) = f(A))$, then x is generated using greedy by some ordering of A.

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 $\{e \in E : x(e) \neq 0\} \subseteq B \subseteq \cup(A : x(A) = f(A))$, then x is generated using greedy by some ordering of A.

- Note, $\text{cl}(x) = \cup(A : x(A) = f(A))$ 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).

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 $\{e \in E : x(e) \neq 0\} \subseteq B \subseteq \cup(A : x(A) = f(A))$, then x is generated using greedy by some ordering of A.

- Note, $\text{cl}(x) = \cup(A : x(A) = f(A))$ 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.

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 $\{e \in E : x(e) \neq 0\} \subseteq B \subseteq \cup(A : x(A) = f(A))$, then x is generated using greedy by some ordering of A.

- Note, $\mathrm{cl}(x) = \cup(A : x(A) = f(A))$ 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, $\mathrm{cl}(x)$ is a tight set.
- Also, $\mathrm{supp}(x) = \{e \in E : x(e) \neq 0\}$ is called the support of x.

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 $\{e \in E : x(e) \neq 0\} \subseteq B \subseteq \cup(A : x(A) = f(A))$, then x is generated using greedy by some ordering of A.

- Note, $\text{cl}(x) = \cup(A : x(A) = f(A))$ 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.

Scratch Paper

Scratch Paper

Scratch Paper

Sources for Today's Lecture

- J. Edmonds, "Submodular Functions, Matroids, and Certain Polyhedra", 1970.
- Lovász, "Submodular Functions and Convexity", 1983.
- Grabisch, Marichal, Mesiar, Pap, "Aggregation Functions", 2009.
- Bixby, Cunningham, Topkis, "The Partial Order of a Polymatroid Extreme Point", 1985.
- L. Schrijver, "Combinatorial Optimization", 2003.