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

Prof. Jeff Bilmes

University of Washington, Seattle
Department of Electrical Engineering
Winter Quarter, 2011
http://ee.washington.edu/class/235/2011wtr/index.html

Lecture 2 - April 1st, 2011

# Announcements

- Weekly Office Hours: Wednesdays, 12:30-1:30pm, 10 minutes after class on Wednesdays.

# Scratch Paper

# Scratch Paper

# Scratch Paper

# Submodular Definitions

## Definition (submodular)

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

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B) \tag{1}$$

An alternate and equivalent definition is:

## Definition (diminishing returns)

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

$$f(A \cup \{v\}) - f(A) \geq f(B \cup \{v\}) - f(B) \tag{2}$$

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

# Ground set: $E$ or $V$?

Submodular functions are functions defined on subsets of some finite set, called the ground set .

- It is common in the literature to use either $E$ or $V$ as the ground set.

# Ground set: $E$ or $V$?

$A, B, C \subseteq E$

$X, Y \subseteq E$ or $V$

Submodular functions are functions defined on subsets of some finite set, called the ground set .

- It is common in the literature to use either $E$ or $V$ as the ground set.
- We will follow this inconsistency in the literature and will inconsistently use either $E$ or $V$ as our ground set (hopefully not in the same equation, if so, please point this out).

# Notation $\mathbb{R}^E$

$E = [n] = \{1, 2, \ldots, n\}$

$$\mathbb{R}^E = \{x = (x_j \in \mathbb{R} : j \in E)\} \tag{3}$$

$$\mathbb{R}_+^E = \{x = (x_j : j \in E) : x \geq 0\} \tag{4}$$

Any vector $x \in \mathbb{R}^E$ can be treated as a modular function, and vice versa. That is

$$x(A) = \sum_{a \in A} x_a \qquad A \subseteq E \tag{5}$$

$f(A)$

$\Rightarrow x(\phi) = 0$

Normalized.

# Other Notation: indicator vectors of sets

Given an $A \subseteq E$, define the vector $\mathbf{1}_A \in \mathbb{R}_+^E$ to be

$$\mathbf{1}_A(j) = \begin{cases} 1 & \text{if } j \in A; \\ 0 & \text{if } j \notin A \end{cases} \tag{6}$$

Sometimes this will be written as $\chi_A \equiv \mathbf{1}_A$.

# Other Notation: indicator vectors of sets

Given an $A \subseteq E$, define the vector $\mathbf{1}_A \in \mathbb{R}_+^E$ to be

$$\mathbf{1}_A(j) = \begin{cases} 1 & \text{if } j \in A; \\ 0 & \text{if } j \notin A \end{cases} \tag{6}$$

Sometimes this will be written as $\chi_A \equiv \mathbf{1}_A$.
Thus, given modular function $x \in \mathbb{R}^E$, we can write $x(A)$ in a variety of ways, i.e.,

$$x(A) = x \cdot \mathbf{1}_A = \sum_{i \in A} x(\overset{i}{\cancel{A}}) \tag{7}$$

# Other Notation: singletons and sets

When $A$ is a set and $k$ is a singleton (a ~~set with a~~ single item), the union is properly written as $A \cup \{k\}$, but sometimes I will write just $A + k$.

# Summing Submodular Functions

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

$$f : 2^E \to \mathbb{R} \text{ with } f(A) = f_1(A) + f_2(A) \tag{8}$$

is submodular. This follows easily since

$$\begin{aligned} f(A) + f(B) &= f_1(A) + f_2(A) + f_1(B) + f_2(B) &(9)\\ &\geq f_1(A \cup B) + f_2(A \cup B) + f_1(A \cap B) + f_2(A \cap B) &(10)\\ &= f(A \cup B) + f(A \cap B). &(11) \end{aligned}$$

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

any ~~convex~~ conic combination also works.

$f(A) = \alpha_1 f_1(A) + \alpha_2 f_2(A)$ $\alpha_1, \alpha_2 \geq 0$

# Summing Submodular and Modular Functions

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

$m' = -m$

$$f : 2^E \to \mathbb{R} \text{ with } f(A) = f_1(A) - m(A) \tag{12}$$

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

$$\begin{aligned} f(A) + f(B) &= f_1(A) - m(A) + f_1(B) - m(B) && (13)\\ &\geq f_1(A \cup B) - m(A \cup B) + f_1(A \cap B) - m(A \cap B) && (14)\\ &= f(A \cup B) + f(A \cap B). && (15) \end{aligned}$$

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

# Restricting Submodular Functions

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

$$f' : 2^E \to \mathbb{R} \text{ with } f(A) = f(A \cap S) \tag{16}$$

is submodular.

**Proof.**

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

$$f((A+v) \cap S) - f(A \cap S) \geq f((B+v) \cap S) - f(B \cap S) \tag{17}$$

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

$$f(A' + v) - f(A') \geq f(B' + v) - f(B') \tag{18}$$

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

# Summing Restricted Submodular Functions

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

$$f : 2^E \to \mathbb{R} \text{ with } f(A) = f_1(A \cap S_1) + f_2(A \cap S_2) \tag{19}$$

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

# Summing Restricted Submodular Functions

$f_C$

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

$$f : 2^E \to \mathbb{R} \text{ with } f(A) = \sum_{C \in \mathcal{C}} f_C(A \cap C) \tag{19}$$

is submodular. This property is critical for image processing and graphical models.

Ex: $\mathcal{C}$ = all pairs = $\{\{u,v\} : u,v \in E\}$

$\mathcal{C}$ = edges of some graph.

$p(x) \propto \exp(-f(X(x)))$  $X(x) \subseteq E$ s.t. $i \in X(x)$ if $x_i = 1$

# Max

$\mathbb{R}_+$ ~ normalization

Given $V$, let $c \in \mathbb{R}^V$ be a given fixed vector. Then

$$f : 2^V \to \mathbb{R} \text{ where } f(A) = \max_{j \in A} c_j \tag{20}$$

is submodular.

$$\boxed{\max_{j \in \emptyset} c_j = -\infty}$$


**Proof.**

Consider:

$$\max_{j \in A} c_j + \max_{j \in B} c_j \geq \max_{j \in A \cup B} c_j + \max_{j \in A \cap B} c_j \tag{21}$$

which follows since we have that

$$\max(\max_{j \in A} c_j, \max_{j \in B} c_j) = \max_{j \in A \cup B} c_j \tag{22}$$

and

$$\min(\max_{j \in A} c_j, \max_{j \in B} c_j) \geq \max_{j \in A \cap B} c_j \tag{23}$$

# Facility Location

Given $V, E$, let $c \in \mathbb{R}_{\cancel{+}}^{V \times E}$ be a given $|V| \times |E|$ ~~non-negative~~ matrix. Then

$$f : 2^E \to \mathbb{R}, \text{ where } f(A) = \sum_{i \in V} \max_{j \in A} c_{ij} \tag{24}$$

is submodular. This is ~~the~~ a facility location function.

**Proof.**

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

# Facility/Plant Location (Uncapacitated)

- Facility Location is a core problem in operations research and a strong motivation for submodular functions. Key goal is to place "facilities" to supply demand sites as efficiently as possible.
- Let $E$ be a set of possible factory/plant locations, and $V$ is a set of sites needing to be serviced (e.g., cities).
- Let $c_{ij}$ be the "benefit" (e.g., $1/c_{ij}$ is the cost) of servicing city $i$ with plant $j$.
- Let $m : 2^E \to \mathbb{R}_+^E$ be a plant construction modular function (vector). E.g., $1/m_j$ is the cost to build a plant at location $j$.
- Each city needs to be serviced by only one plant but no less than one.
- Define $f(A)$ is the ~~maximum~~ "delivery benefit" plus "construction benefit" when the plants in set $A$ are considered to be constructed.
- We can define $f(A) = m(A) + \sum_{i \in V} \max_{j \in A} c_{ij}$.
- Goal is to find a set $A$ that maximizes $f(A)$ (the benefit) placing a bound on the number of plants $A$ (e.g., $|A| \leq k$).

# Log Determinant

$\Rightarrow 2 \times 2$

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

$$f(A) = \log\det(\mathbf{\Sigma}_A) \text{ is submodular.} \tag{25}$$

(note: still submodular in the semi-def case as well)

**Proof.**

Suppose $x \in \mathbf{R}^n$ is multivariate Gaussian, that is

$$x \in p(x) = \frac{1}{\sqrt{|2\pi\mathbf{\Sigma}|}} \exp\left( -\frac{1}{2}(x-\mu)^T \mathbf{\Sigma}^{-1}(x-\mu) \right) \tag{26}$$

...

# Log Determinant

$$f(A) = \log \det(\mathbf{\Sigma}_A) \text{ is submodular.} \tag{25}$$

## ...cont.

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

$$h(X) = \log \sqrt{|2\pi e \mathbf{\Sigma}|} = \log \sqrt{(2\pi e)^n |\mathbf{\Sigma}|} \tag{26}$$

and in particular, for a variable subset $A$,

$$f(A) = h(X_A) = \log \sqrt{(2\pi e)^{|A|} |\mathbf{\Sigma}_A|} \tag{27}$$

Entropy is submodular (conditioning reduces entropy), and moreover

$$f(A) = h(X_A) = m(A) + \frac{1}{2}\log |\mathbf{\Sigma}_A| \tag{28}$$

where $m(A)$ is a modular function. □

# Concave over non-negative modular

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

**Proof.**

f(A) = g(m(A)) □

then $f$ is submodular.

**Proof.**

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

$$g(m(A) + m(v)) - g(m(A)) \geq g(m(B) + m(v)) - g(m(B)) \tag{29}$$

□

The converse is true as well. Exercise: prove this.

# Monotone difference of two functions

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

$$h(A) = \min(f(A), g(A)) \tag{30}$$

is submodular.

**Proof.**

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

$$\begin{matrix} f(X)+f(Y) \\ g(X)+g(Y) \end{matrix} \geq \min(f(X\cup Y), g(X\cup Y)) + \min(f(X\cap Y), g(X\cap Y)) \tag{31}$$

$=$ $h(X) + h(Y)$

...

# Monotone difference of two functions

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

$$h(A) = \min(f(A), g(A)) \tag{30}$$

is submodular.

**...cont.**

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

$$h(X) + h(Y) = f(X) + g(Y) \geq f(X \cup Y) + f(X \cap Y) + g(Y) - f(Y) \tag{31}$$

$g(Y) - f(Y) \geq g(X \cup Y) - f(X \cup Y)$

By monotonicity, $f(X \cup Y) + g(Y) - f(Y) \geq g(X \cup Y)$ giving

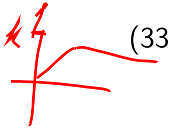
$$h(X) + h(Y) \geq g(X \cup Y) + f(X \cap Y) \geq h(X \cup Y) + h(X \cap Y) \tag{32}$$

□

# Min

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

$$h(A) = \min(k, f(A)) \tag{33}$$




is submodular.

**Proof.**

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

Note also, $g(a) = \min(k, a)$ for constant $k$ is a concave function, so we can use the earlier result about composing a concave function with a submodular function to get this result as well. □

# More on Min

Saturate trick

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

$$h(A) = \frac{1}{2}(\min(k, f) + \min(k, g)) \tag{34}$$

then $h$ is submodular, and $h(A) \geq k$ if and only if both $f(A) \geq k$ and $g(A) \geq k$.
We plan to revisit this again later in the quarter.

# Arbitrary functions as a difference between submodular functions

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

Define

**Proof.**

let f be an arbitrary function:

$$\alpha \triangleq \min_{X,Y} \left[ f(X) + f(Y) - f(X \cup Y) - f(X \cap Y) \right] \tag{35}$$

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

~~$\beta > 0$~~ $\beta \geq 0$

$$\beta \triangleq \min_{X,Y} \left[ h(X) + h(Y) - h(X \cup Y) - h(X \cap Y) \right] \tag{36}$$

If $\beta = 0$ then take it to be the smallest positive difference. $\Rightarrow$ modular doesn't work. …

# Arbitrary functions as a difference between submodular functions

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

**...cont.**

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

$$f'(A) = f(A) + \frac{|\alpha|}{\beta} h(A) \tag{35}$$

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

# Submodular Definitions

## Definition (submodular)

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

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B) \tag{1}$$

An alternate and equivalent definition is:

## Definition (diminishing returns)

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

$$f(A \cup \{v\}) - f(A) \geq f(B \cup \{v\}) - f(B) \tag{2}$$

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

# Submodular Definitions

An alternate and equivalent definition is:

**Definition (group diminishing returns)**

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

$$f(A \cup C) - f(A) \geq f(B \cup C) - f(B) \tag{36}$$

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

# Submodular Definitions

## Proposition

*group diminishing returns implies diminishing returns*

## Proof.

Obvious, set $C = \{v\}$. $\square$

# Submodular Definitions

## Proposition

*diminishing returns implies group diminishing returns*

## Proof.

Let $C = \{c_1, c_2, \dots, c_k\}$. Then diminishing returns implies yields

$$
\begin{aligned}
& f(A \cup C) - f(A) && (37)\\
&= f(A \cup C) - \sum_{i=1}^{k-1}\Big(f(A \cup \{c_1, \dots, c_i\}) - f(A \cup \{c_1, \dots, c_i\})\Big) - f(A) && (38)\\
&= \sum_{i=1}^{k} f(A \cup \{c_1 \dots c_i\}) - f(A \cup \{c_1 \dots c_{i-1}\}) = \sum_i \rho_{c_i}(A \cup \{c_1 .. c_{i-1}\}) && (39)\\
&\geq \sum_{i=1}^{k} f(B \cup \{c_1 \dots c_i\}) - f(B \cup \{c_1 \dots c_{i-1}\}) = \sum \rho_{c_i}(B \cup \{c_1 .. c_{i-1}\}) && (40)\\
&= f(B \cup C) - \sum_{i=1}^{k-1}\Big(f(B \cup \{c_1, \dots, c_i\}) - f(B \cup \{c_1, \dots, c_i\})\Big) - f(B) && (41)\\
&= f(B \cup C) - f(B) && (42)
\end{aligned}
$$

# Submodular Definitions are equivalent

**Proposition**

*The two aforementioned definitions of submodularity submodular and diminishing returns are identical.*

# Submodular Definitions are equivalent

## Proof.

Assume submodular. Assume $A \subset B$ as otherwise trivial.
Let $B \setminus A = \{v_1, v_2, \ldots, v_k\}$ and define $A^i = A \cup \{v_1 \ldots v_i\}$, so $A^0 = A$.
Then by submodular,

$A^k = B$

$$f(A^i + v) + f(A^i + v_{i+1}) \geq f(A^i + v + v_{i+1}) + f(A^i) \tag{43}$$

or

$$f(A^i + v) - f(A^i) \geq f(A^i + v_{i+1} + v) - f(A^i + v_{i+1}) \tag{44}$$

we apply this inductively, and use

$$f(A^{i+1} + v) - f(A^{i+1}) = f(\overbrace{A^i + v_{i+1}}^{A^{i+1}} + v) - f(\overbrace{A^i + v_{i+1}}^{A^{i+1}}) \tag{45}$$

and that $A^{k-1} + v_k = B$. ...

# Submodular Definitions are equivalent

## ...cont.

Assume group diminishing returns. Assume $A \neq B$ otherwise trivial. Define $A' = A \cap B$, $C = A \setminus B$, and $B' = B$. Then

$$f(A' + C) - f(A') \geq f(B' + C) - f(B') \tag{46}$$

giving

$$f(A' + C) + f(B') \geq f(B' + C) + f(A') \tag{47}$$

or

$$f(A \cap B + A \setminus B) + f(B) \geq f(B + A \setminus B) + f(A \cap B) \tag{48}$$

which is the same as

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B) \tag{49}$$

# Submodular Definitions

## Definition (singleton)

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

$$f(A \cup \{a\}) + f(A \cup \{b\}) \geq f(A \cup \{a, b\}) + f(A) \tag{50}$$

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

$$\begin{aligned} f(A + a) - f(A) &\geq f(A + b_1 + a) - f(A + b_1) && (51) \\ &\geq f(A + b_1 + b_2 + a) - f(A + b_1 + b_2) && (52) \\ &\geq \dots && (53) \\ &\geq f(A + b_1 + \dots + b_k + a) - f(A + b_1 + \dots + b_k) && (54) \\ &= f(B + a) - f(B) && (55) \end{aligned}$$

# Gain

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

$$\begin{aligned} f(A \cup \{j\}) - f(A) &\triangleq \rho_j(A) && (56)\\ &\triangleq \rho_A(j) && (57)\\ &\triangleq f(\{j\}|A) && (58)\\ &\triangleq f(j|A) && (59) \end{aligned}$$

$H(X_j | X_A)$

We'll use either $\rho_j(A)$ or $f(j|A)$. Note, diminishing returns can now be stated as saying that $\rho_j(A)$ is a monotone non-increasing function of $A$, since $\rho_j(A) \geq \rho_j(B)$ whenever $B \supseteq A$.

# Equivalent Definitions of Submodularity

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B), \;\; \forall A, B \subseteq E \tag{60}$$

$$\rho_j(S) \geq \rho_j(T), \; \forall S \subseteq T \subseteq E, \text{ with } j \in E \setminus T \tag{61}$$

$$\rho_j(S) \geq \rho_j(S \cup \{k\}), \; \forall S \subseteq E \text{ with } j \in E \setminus (S \cup \{k\}) \tag{62}$$

$$f(T) \leq f(S) + \sum_{j \in T \setminus S} \rho_j(S) - \sum_{j \in S \setminus T} \rho_j(S \cup T - \{j\}), \; \forall S, T \subseteq E \tag{63}$$


$-\{ f(J \cup T) - f(J \cup T - j) \}$

$$f(T) \leq f(S) + \sum_{j \in T \setminus S} \rho_j(S), \; \forall S \subseteq T \subseteq E \tag{64}$$

$$f(T) \leq f(S) + \sum_{j \in S \setminus T} \rho_j(S \setminus \{j\}), \; \forall T \subseteq S \subseteq E \tag{65}$$

$$f(T) \leq f(S) - \sum_{j \in S \setminus T} \rho_j(S \setminus \{j\}) + \sum_{\in T \setminus S} \rho_j(S \cap T) \; \forall S, T \subseteq E \tag{66}$$

# Equivalent Definitions of Submodularity

We've already seen that Eq. 60 $\equiv$ Eq. 61 $\equiv$ Eq. 62. We next show that Eq. 62 $\Rightarrow$ Eq. 63 $\Rightarrow$ Eq. 64 $\Rightarrow$ Eq. 62.

# Eq. 62 $\Rightarrow$ Eq. 63

Let $T \setminus S = \{j_1, \ldots, j_r\}$ and $S \setminus T = \{k_1, \ldots, k_q\}$.
First upper bound T in context of S

$$f(S \cup T) - f(S) = \sum_{t=1}^{r} [f(S \cup \{j_1, \ldots, j_t\} - f(S \cup \{j_1, \ldots, j_{t-1}\})] \quad (67)$$

$$= \sum_{t=1}^{r} \rho_{j_t}(S \cup \{j_1, \ldots, j_{t-1}\}) \leq \sum_{t=1}^{r} \rho_{j_t}(S) \quad (68)$$

$$= \sum_{j \in T \setminus S} \rho_j(S) \quad (69)$$

# Eq. 62 $\Rightarrow$ Eq. 63

Let $T \setminus S = \{j_1, \ldots, j_r\}$ and $S \setminus T = \{k_1, \ldots, k_q\}$.
Next,

Lower bound S in context of T

$$f(S \cup T) - f(T) = \sum_{t=1}^{q} [f(T \cup \{k_1, \ldots, k_t\}) - f(T \cup \{k_1, \ldots, k_{t-1}\})] \tag{67}$$

$$= \sum_{t=1}^{q} \rho_{k_t}(\overset{T}{\cancel{S}} \cup \{k_1, \ldots, k_t\} \setminus \{k_t\}) \geq \sum_{t=1}^{q} \rho_{k_t}(T \cup S \setminus \{k_t\}) \tag{68}$$

$$= \sum_{j \in S \setminus T} \rho_j(S \cup T \setminus \{j\}) \tag{69}$$

# Eq. 62 $\Rightarrow$ Eq. 63

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

upper bound

$$f(S \cup T) - f(S) \leq \sum_{j \in T \setminus S} \rho_j(S) \tag{67}$$

and

Lower bound

$$f(S \cup T) - f(T) \geq \sum_{j \in S \setminus T} \rho_j(S \cup T \setminus \{j\}) \tag{68}$$

and subtracting the 2nd from the first gives the result.

(68)-(67)

$f(T) - f(S) \leq \dots$

# Eq. 63 $\Rightarrow$ Eq. 64

This follows immediately since if $S \subseteq T$, then $S \setminus T = \emptyset$, and the last term of Eq. 63 vanishes.

# Eq. 64 $\Rightarrow$ Eq. 62

Here, we set $\overbrace{T}$ $= S \cup \{j,k\}$, $j \notin S \cup \{k\}$ into Eq. 64 to obtain

$$
\begin{aligned}
\overbrace{f(S \cup \{j,k\})}^{T} &\le f(S) + \rho_j(S) + \rho_k(S) && (69)\\
&= f(S) + f(S + \{j\}) - f(S) + f(S + \{k\}) - f(S) && (70)\\
&= f(S + \{j\}) + f(S + \{k\}) - f(S) && (71)\\
&= \rho_j(S \cancel{+ \{k\}}) + f(S + \{k\}) && (72)
\end{aligned}
$$

giving

$$
\begin{aligned}
\rho_j(S \cup \{k\}) &= f(S \cup \{j,k\}) - f(S \cup \{k\}) && (73)\\
&= f(S \cup \{j,k\}) - \rho_k(S) - f(S) && (74)\\
&\le \rho_j(S) && (75)
\end{aligned}
$$

# Sources for Today's Lecture

Lovasz-1983, Nemhauser,Wilsey,Fisher-1978, Narayanan-1997, and Narasimhan,Bilmes-2005.