Submodularity and Discrete Convexity

Satoru Fujishige

Research Institute for Mathematical Sciences
Kyoto University

4th Workshop on Discrete Optimization in Machine Learning (DISCML)

December 7, 2012
Lake Tahoe, Nevada, United States
$E$: a nonempty finite set

**A submodular function** $f : 2^E \to \mathbb{R}$

$$f(X) + f(Y) \geq f(X \cup Y) + f(X \cap Y) \quad (\forall X, Y \subseteq E)$$
Let \( E \) be a nonempty finite set.

**A submodular function** \( f : 2^E \to \mathbb{R} \)

\[
f(X) + f(Y) \geq f(X \cup Y) + f(X \cap Y) \quad (\forall X, Y \subseteq E)
\]

---

**Lemma:** \( D_{\text{min}}(f) \): the set of all minimizers of \( f \)

\[
X, Y \in D_{\text{min}}(f) \implies X \cup Y, X \cap Y \in D_{\text{min}}(f)
\]

i.e., \( D_{\text{min}}(f) \) is a distributive lattice.
$E$: a nonempty finite set

**A submodular function** $f : 2^E \rightarrow \mathbb{R}$

$$f(X) + f(Y) \geq f(X \cup Y) + f(X \cap Y) \quad (\forall X, Y \subseteq E)$$

---

**Lemma**: $\mathcal{D}_{\text{min}}(f)$: the set of all minimizers of $f$

$$X, Y \in \mathcal{D}_{\text{min}}(f) \implies X \cup Y, X \cap Y \in \mathcal{D}_{\text{min}}(f)$$
i.e., $\mathcal{D}_{\text{min}}(f)$ is a distributive lattice.

---

It leads us to a Characterization of Submodular Functions

---

**Theorem**: A set function $f : 2^E \rightarrow \mathbb{R}$ is a submodular function
\
\[\iff\]
For $\forall$ modular function $\mu : 2^E \rightarrow \mathbb{R}$, $\mathcal{D}_{\text{min}}(f - \mu)$ is a distributive lattice.


**Theorem (Birkhoff-Iri):** \( \mathcal{D} \subseteq 2^E, \emptyset, E \in \mathcal{D} \).

\( \mathcal{D} \) is a distributive lattice with respect to \( \cup \) and \( \cap \) as lattice operations.

There exists a partially ordered set (poset) \( \mathcal{P} = (\Pi(E), \preceq) \) on a partition \( \Pi(E) \) of \( E \) such that \( \mathcal{D} \) is given by

\[ \mathcal{D} = \{ X \subseteq E \mid \exists \text{ ideal } J \text{ of } \mathcal{P}: X = \bigcup_{F \in J} F \}. \]

\( (X \subseteq E \text{ is an ideal of } \mathcal{P} \text{ if } e \in X \text{ and } e' \preceq e \text{ always imply } e' \in X. \)
Theorem (Birkhoff-Iri): \( \mathcal{D} \subseteq 2^E, \ \emptyset, E \in \mathcal{D} \).

\( \mathcal{D} \) is a distributive lattice with respect to \( \cup \) and \( \cap \) as lattice operations.

There exists a partially ordered set (poset) \( \mathcal{P} = (\Pi(E), \preceq) \) on a partition \( \Pi(E) \) of \( E \) such that \( \mathcal{D} \) is given by

\[
\mathcal{D} = \{ X \subseteq E \mid \exists \text{ ideal } \mathcal{J} \text{ of } \mathcal{P}: X = \bigcup_{F \in \mathcal{J}} F \}.
\]

(\( X \subseteq E \) is an ideal of \( \mathcal{P} \) if \( e \in X \) and \( e' \preceq e \) always imply \( e' \in X \).)

---

A simple distributive lattice \( \mathcal{D} \): every component of partition \( \Pi(E) \) is a singleton, i.e., \( \mathcal{D} = 2^\mathcal{P} \) (the set of all ideals of \( \mathcal{P} = (E, \preceq) \)).
Submodular System $(\mathcal{D}, f)$ on $E$

$\mathcal{D} \subseteq 2^E$: a distributive lattice $\ (\emptyset, E \in \mathcal{D})$

$X, Y \in \mathcal{D} \implies X \cup Y, X \cap Y \in \mathcal{D}$

$f: \mathcal{D} \to \mathbb{R}$: a submodular function $\ (f(\emptyset) = 0)$

$\forall X, Y \in \mathcal{D}: f(X) + f(Y) \geq f(X \cup Y) + f(X \cap Y)$
Submodular System $(\mathcal{D}, f)$ on $E$

$\mathcal{D} \subseteq 2^E$: a distributive lattice $(\emptyset, E \in \mathcal{D})$

$X, Y \in \mathcal{D} \implies X \cup Y, X \cap Y \in \mathcal{D}$

$f: \mathcal{D} \to \mathbb{R}$: a submodular function $(f(\emptyset) = 0)$

\[ \forall X, Y \in \mathcal{D}: f(X) + f(Y) \geq f(X \cup Y) + f(X \cap Y) \]

\[ P(f) = \{ x \in \mathbb{R}^E | \forall X \in \mathcal{D}: x(X) \leq f(X) \} \]

(Submodular Polyhedron)

\[ x(X) = \sum_{e \in X} x(e), \quad x(\emptyset) = 0 \]

$B(f) = \{ x \mid x \in P(f), x(E) = f(E) \}$

(Base Polyhedron)
\( \mathcal{D} \subseteq 2^E: \) a distributive lattice \( (\emptyset, E \in \mathcal{D}) \)

\[ X, Y \in \mathcal{D} \implies X \cup Y, X \cap Y \in \mathcal{D} \]

\( f: \mathcal{D} \to \mathbb{R}: \) a submodular function \( (f(\emptyset) = 0) \)

\[ \forall X, Y \in \mathcal{D}: f(X) + f(Y) \geq f(X \cup Y) + f(X \cap Y) \]

\[ \mathcal{P}(f) = \{ x \in \mathbb{R}^E | \forall X \in \mathcal{D} : x(X) \leq f(X) \} \]

(Submodular Polyhedron)

\[ x(X) = \sum_{e \in X} x(e), \quad x(\emptyset) = 0 \]

\[ \mathcal{B}(f) = \{ x \mid x \in \mathcal{P}(f), \; x(E) = f(E) \} \]

(Base Polyhedron)

Remark: Submodular system \((\mathcal{D}, f)\) \(\xrightarrow{1:1} \) Base polyhedron \( \mathcal{B}(f) \)

(Submodular polyhedron \( \mathcal{P}(f) \))
Define a supermodular system \((\mathcal{D}, g)\) and its associated supermodular polyhedron \(P(g)\) and base polyhedron \(B(g)\) in a dual manner.
Define a supermodular system \((\mathcal{D}, g)\) and its associated supermodular polyhedron \(P(g)\) and base polyhedron \(B(g)\) in a dual manner.

Duality

\[
\bar{\mathcal{D}} = \{ E \setminus X \mid X \in \mathcal{D} \} \\
f^\#(E \setminus X) = f(E) - f(X) \quad (X \in \mathcal{D})
\]

\((\bar{\mathcal{D}}, f^\#)\): the supermodular system dual to submodular system \((\mathcal{D}, f)\)

\[
B(f) = B(f^\#)
\]
Proposition: The base polyhedron of $(\mathcal{D}, f)$ has an extreme point. 
$\iff$ $\mathcal{D}$ is simple.
\((\mathcal{D}, f)\): A submodular system on \(E\)

**Proposition:** The base polyhedron of \((\mathcal{D}, f)\) has an extreme point.
\[\iff \mathcal{D} \text{ is simple.}\]

**Proposition:** The base polyhedron of \((\mathcal{D}, f)\) is **bounded**.
\[\iff \mathcal{D} = 2^E \text{ (a Boolean lattice)}\]
$(\mathcal{D}, f)$: A submodular system on $E$

**Proposition**: The base polyhedron of $(\mathcal{D}, f)$ has an extreme point.

$\iff \mathcal{D}$ is simple.

**Proposition**: The base polyhedron of $(\mathcal{D}, f)$ is bounded.

$\iff \mathcal{D} = 2^E$ (a Boolean lattice)

**Proposition**: Suppose $\mathcal{D}$ is simple. Then, all the extreme bases are nonnegative

$\iff f$ is monotone nondecreasing.
**(D, f): A submodular system on E**

**Proposition:** The base polyhedron of \((D, f)\) has an extreme point.
\[ \iff D \text{ is simple.} \]

**Proposition:** The base polyhedron of \((D, f)\) is bounded.
\[ \iff D = 2^E \quad (a \text{ Boolean lattice}) \]

**Proposition:** Suppose \(D\) is simple. Then,
\[ \iff \text{all the extreme bases are nonnegative} \]
\[ \iff f \text{ is monotone nondecreasing.} \]

**Proposition:** Suppose \(D\) is simple. Then,
\[ \iff \text{all the extreme bases are integral} \]
\[ \iff f \text{ is integer-valued.} \]
Polymatroid (Edmonds):

\[ D = 2^E \text{ and } f \text{ is monotone nondecreasing} \]
\[ (X \subseteq Y \subseteq E \implies f(X) \leq f(Y)). \]
\[ \iff B(f) \subseteq \mathbb{R}^E_+ \]
Polymatroid (Edmonds):

\[ \mathcal{D} = 2^E \text{ and } f \text{ is monotone nondecreasing} \]
\[(X \subseteq Y \subseteq E \implies f(X) \leq f(Y)). \]
\[ \iff B(f) \subseteq \mathbb{R}^E_+ \]

Matroid (Whitney):

Furthermore, \( f \) is integer-valued and has a unit-increase property.
\[ \forall X \in 2^E, \forall e \in E \setminus X : f(X) \leq f(X \cup \{e\}) \leq f(X) + 1 \]
(extreme bases \( \leftrightarrow \) matroid bases)
Generalized polymatroids (Frank, Hassin) and Base Polyhedra
Theorem (Tomizawa): For a bounded polyhedron $P \subseteq \mathbb{R}^E$, $P$ is a base polyhedron if all the edge vectors of $P$ are of form $(0, \ldots, 0, \pm1, 0, \ldots, 0, \mp1, 0, \ldots, 0)$.
**Theorem** (Tomizawa): For a bounded polyhedron $P \subset \mathbb{R}^E$, 

$P$ is a base polyhedron

\[ \updownarrow \]

all the edge vectors of $P$ are of form

\[(0, \cdots, 0, \pm 1, 0, \cdots, 0, \mp 1, 0, \cdots, 0)\]

**Corollary:** For a bounded polyhedron $P \subset \mathbb{R}^E$, 

$P$ is a generalized polymatroid

\[ \updownarrow \]

all the edge vectors of $P$ are of form

\[(0, \cdots, 0, \pm 1, 0, \cdots, 0, \mp 1, 0, \cdots, 0) \text{ or } (0, \cdots, 0, \pm 1, 0, \cdots, 0)\]

**Remark:** The above two are also valid for pointed polyhedra.
The Intersection Theorem and Its Equivalents

$(\mathcal{D}_i, f_i) \ (i = 1, 2)$: submodular systems on $E$

The Intersection Theorem (Edmonds):

$$\max \{ x(E) \mid x \in P(f_1) \cap P(f_2) \}$$

$$= \min \{ f_1(X) + f_2(E \setminus X) \mid X \in \mathcal{D}_1, \ E \setminus X \in \mathcal{D}_2 \}$$

(+ Integrality)
The Intersection Theorem and Its Equivalents

$(D_i, f_i) \ (i = 1, 2)$: submodular systems on $E$

---

**The Intersection Theorem** (Edmonds):

\[
\max\{x(E) \mid x \in P(f_1) \cap P(f_2)\} = \min\{f_1(X) + f_2(E \setminus X) \mid X \in D_1, \ E \setminus X \in D_2\}
\]

(+ Integrality)

---

**The Intersection Theorem’**:

\[
\max\{x \land y(E) \mid x \in B(f_1), \ y \in B(f_2)\} = \min\{f_1(X) + f_2(E \setminus X) \mid X \in D_1, \ E \setminus X \in D_2\}
\]

(+ Integrality)

---

$(x \land y)(e) = \min\{x(e), y(e)\} \ (e \in E)$.

---

Diagram:

```
x
  / \
 /  \
/    \
|    |
|    |
|    |
y
```

---

$E$
$(\mathcal{D}_1, f), (\mathcal{D}_2, g)$: a submodular system and supermodular system on $E$

**Discrete Separation Theorem** (Frank):

$$f \geq g \implies \exists z \in \mathbb{R}^E : f \geq z \geq g \quad \text{(i.e.,} \ P(f) \cap P(g) \neq \emptyset)$$

(+ Integrality)
(\mathcal{D}_1, f), (\mathcal{D}_2, g): a submodular system and supermodular system on E

\textbf{Discrete Separation Theorem} (Frank):

\[ f \geq g \implies \exists z \in \mathbb{R}^E : f \geq z \geq g \quad (\text{i.e., } P(f) \cap P(g) \neq \emptyset) \]

(+ Integrality)

\textbf{Discrete Separation Theorem'}:

\[ f \geq g \implies \exists (x \in B(f), y \in B(g)) : x \geq y \]

(+ Integrality)
(\mathcal{D}_1, f): a submodular system on \( E \)
(\mathcal{D}_2, g): a supermodular system on \( E \)

\[
f^*(x) = \max \{ x(X) - f(X) \mid X \in \mathcal{D}_1 \} \quad (x \in \mathbb{R}^E)
g^*(x) = \min \{ x(X) - g(X) \mid X \in \mathcal{D}_2 \} \quad (x \in \mathbb{R}^E)
\]

---

**Fenchel Duality Theorem (F):**

\[
\min \{ f(X) - g(X) \mid X \in \mathcal{D}_1 \cap \mathcal{D}_2 \} \\
= \max \{ g^*(x) - f^*(x) \mid x \in \mathbb{R}^E \}
\]

(+ Integrality)
(\mathcal{D}_1, f): \text{ a submodular system on } E
(\mathcal{D}_2, g): \text{ a supermodular system on } E

f^*(x) = \max\{x(X) - f(X) \mid X \in \mathcal{D}_1\} \quad (x \in \mathbb{R}^E)
g^*(x) = \min\{x(X) - g(X) \mid X \in \mathcal{D}_2\} \quad (x \in \mathbb{R}^E)

\underline{\text{Fenchel Duality Theorem (F)}:}

\min\{f(X) - g(X) \mid X \in \mathcal{D}_1 \cap \mathcal{D}_2\}
= \max\{g^*(x) - f^*(x) \mid x \in \mathbb{R}^E\}
(+ \text{ Integrality})

\underline{\text{Fenchel Duality Theorem’:}}

\min\{f(X) - g(X) \mid X \in \mathcal{D}_1 \cap \mathcal{D}_2\}
= \max\{(x - y)^-(E) \mid x \in B(f), \ y \in B(g)\}
(+ \text{ Integrality}) \quad ((x - y)^- = \min\{0, x(e) - y(e)\} \mid e \in E))
(\mathcal{D}_i, f_i) (i = 1, 2): submodular systems on \( E \)

---

**Minkowski Sum Theorem:**

\[
P(f_1) + P(f_2) = P(f_1 + f_2), \\
B(f_1) + B(f_2) = B(f_1 + f_2).
\]

Moreover, if \( f_1 \) and \( f_2 \) are integer-valued, the collections \( P_Z(\cdot) \) and \( B_Z(\cdot) \) of integer points in \( P(\cdot) \) and \( B(\cdot) \) satisfy

\[
P_Z(f_1) + P_Z(f_2) = P_Z(f_1 + f_2), \\
B_Z(f_1) + B_Z(f_2) = B_Z(f_1 + f_2).
\]
Minimum-Norm Base and SFM

$$(\mathcal{D}, f):$$ a submodular system on $E$

$w(\cdot: E \to \mathbb{R})$: a positive weight function, \(\lambda \in \mathbb{R}\).
Minimum-Norm Base and SFM

$(\mathcal{D}, f)$: a submodular system on $E$

$w(\cdot: E \rightarrow \mathbf{R})$: a positive weight function, $\lambda \in \mathbb{R}$

Parametric Vector Reduction: For any $\lambda \in \mathbb{R}$,

\[
\begin{align*}
\max \{ x(E) \mid x \in P(f), \ x \leq \lambda w \} \\
= \min \{ f(X) + \lambda w(E \setminus X) \mid X \subseteq E \}
\end{align*}
\]
Theorem (F): There exists a unique base $b^*$ such that for all $\lambda \in \mathbb{R}$ $x = b^* \land \lambda w$ attains the maximum of the following

$$\max\{x(E) \mid x \in P(f), \ x \leq \lambda w\}$$

$$= \min\{f(X) + \lambda w(E \setminus X) \mid X \subseteq E\}.$$
**Theorem (F):** There exists a **unique base** $b^*$ such that for all $\lambda \in \mathbb{R}$, $x = b^* \land \lambda w$ attains the maximum of the following

$$\max\{x(E) \mid x \in \mathcal{P}(f), \ x \leq \lambda w\}$$

$$= \min\{f(X) + \lambda w(E \setminus X) \mid X \subseteq E\}.$$ 

A base $b$ is called a **lexicographically optimal base** w.r.t. weight $w$ if it **lexicographically maximizes** the sequence

$$T(b/w) = (b(e_1)/w(e_1), \ldots, b(e_n)/w(e_n))$$

of weighted $b(e)/w(e)$ ($e \in E$) arranged in nondecreasing order of magnitude among all bases $b$. 
**Theorem (F):** There exists a unique base $b^*$ such that for all $\lambda \in \mathbb{R}$

$x = b^* \land \lambda w$ attains the maximum of the following

$$\max\{x(E) \mid x \in P(f), \ x \leq \lambda w\}$$

$$= \min\{f(X) + \lambda w(E \setminus X) \mid X \subseteq E\}.$$  

A base $\hat{b}$ is called a **lexicographically optimal base** w.r.t. weight $w$ if it **lexicographically maximizes** the sequence

$$T(b/w) = (b(e_1)/w(e_1), \ldots, b(e_n)/w(e_n))$$

of weighted $b(e)/w(e) \ (e \in E)$ arranged in nondecreasing order of magnitude among all bases $b$.

**Theorem (F):**

1. $b^*$ is the **lexicographically optimal base** with respect to weight $w$.

2. $b^*$ is the **minimizer of** $\sum_{e \in E} b^2(e)/w(e)$ over $B(f)$. 

$\rightarrow$
Theorem (F): There exists a unique base $b^*$ such that for all $\lambda \in \mathbb{R}$, $x = b^* \land \lambda w$ attains the maximum of the following

$$\max \{ x(E) \mid x \in P(f), \ x \leq \lambda w \}$$

$$= \min \{ f(X) + \lambda w(E \setminus X) \mid X \subseteq E \}.$$ 

A base $\hat{b}$ is called a lexicographically optimal base w.r.t. weight $w$ if it lexicographically maximizes the sequence

$$T(b/w) = (b(e_1)/w(e_1), \ldots, b(e_n)/w(e_n))$$

of weighted $b(e)/w(e)$ $(e \in E)$ arranged in nondecreasing order of magnitude among all bases $b$.

Theorem (F):

(1) $b^*$ is the lexicographically optimal base with respect to weight $w$.

(2) $b^*$ is the minimizer of $\sum_{e \in E} b^2(e)/w(e)$ over $B(f)$.

($\rightarrow$ Resource Allocation Problems + submodular constraints)

$\rightarrow$
Remarks: We have
\[
\min\{f(X) + \lambda w(E \setminus X) \mid X \in \mathcal{D}\} \\
= (b^* \land \lambda w)(E) = \max\{x(E) \mid x \leq \lambda w, x \in P(f)\}
\]
Hence, for \( \lambda = 0 \)
\[
\min\{f(X) \mid X \in \mathcal{D}\} = \max\{x(E) \mid x \leq 0, x \in P(f)\},
\]
Remarks: We have
\[
\min\{f(X) + \lambda w(E \setminus X) \mid X \in \mathcal{D}\} = (b^* \wedge \lambda w)(E) = \max\{x(E) \mid x \leq \lambda w, x \in P(f)\}
\]
Hence, for \(\lambda = 0\)
\[
\min\{f(X) \mid X \in \mathcal{D}\} = \max\{x(E) \mid x \leq 0, x \in P(f)\},
\]
which implies
\[A_0 = \{e \mid b^*(e) \leq 0\}, \quad A_- = \{e \mid b^*(e) < 0\}\]
are the unique maximal and the unique minimal minimizer of \(f\).
**Remarks:** We have
\[ \min \{ f(X) + \lambda w(E \setminus X) \mid X \in \mathcal{D} \} \]
\[ = (b^* \wedge \lambda w)(E) = \max\{ x(E) \mid x \leq \lambda w, x \in P(f) \} \]
Hence, for \( \lambda = 0 \)
\[ \min \{ f(X) \mid X \in \mathcal{D} \} = \max\{ x(E) \mid x \leq 0, x \in P(f) \}, \]
which implies
\[ A_0 = \{ e \mid b^*(e) \leq 0 \}, \quad A_- = \{ e \mid b^*(e) < 0 \} \]
are the **unique maximal** and the **unique minimal minimizer** of \( f \).

\[ b^* \text{ is the minimizer of } \sum_{e \in E} b^2(e)/w(e) \]

**Minimum-norm base**
\[ \implies \text{ Submodular Function Minimization} \]

Applicability of P. Wolfe’s minimum-norm point algorithm
Submodular Function Minimization \(\leftarrow\) MNP algorithm of Wolfe

\[
A_0 = \{ e \mid b^*(e) \leq 0 \}, \quad A_- = \{ e \mid b^*(e) < 0 \}
\]

When \(f\) is integer-valued and \(w = 1\),

\[
\lambda_i = \frac{f(E^+_{\lambda_{i-1}}) - f(E^+_{\lambda_i})}{|E^+_{\lambda_i} \setminus E^+_{\lambda_{i-1}}|} \quad (i = 1, 2, \cdots, p), \quad E^+_{\lambda_0} = \emptyset
\]

\[
\lambda_1 < \cdots < \lambda_p
\]

\[
\min_i \{ \lambda_i > 0 \} \geq \frac{1}{|E|}, \quad \max_i \{ \lambda_i < 0 \} \leq -\frac{1}{|E|}
\]

\[
A_0 = \{ e \mid b^*(e) \leq \epsilon \}, \quad A_- = \{ e \mid b^*(e) < -\epsilon \} \quad (\epsilon = \frac{1}{2|E|})
\]
Maximum Weight Base Problem

A weight function $w : E \to [0, 1]$

Maximize \( \sum_{e \in E} w(e)x(e) \)

subject to \( x \in B(f) \)
Let $D = 2^E$

### Maximum Weight Base Problem

A weight function $w : E \rightarrow [0, 1]$

Maximize $\sum_{e \in E} w(e)x(e)$

subject to $x \in B(f)$

For a permutation $\sigma = (e_1, e_2, \ldots, e_n)$ of $E$, define

$\Delta(\sigma): \quad 1 \geq x(e_1) \geq x(e_2) \geq \cdots \geq x(e_n) \geq 0$

$S_i = \{e_1, \ldots, e_i\} \quad (i = 1, \ldots, n)$

$S_0 = \emptyset \subset S_1 \subset \cdots \subset S_n = E$

and

$b^\sigma(e_i) = f(S_i) - f(S_{i-1}) \quad (i = 1, \ldots, n)$

Then, $b^\sigma$ is a **maximum weight base** for $w \in \Delta(\sigma)$.

(← Greedy Algorithm) (Edmonds)
\( \hat{f}(w) = \text{the value of a maximum weight base} \)

(the support function of \( B(f) \) restricted on \( [0, 1]^E \))

(*) \( \hat{f} \) is a linear function on each cell \( \Delta(\sigma) \).
The set of cells for \( n! \) permutations \( \sigma \) defines a simplicial division of unit hypercube \([0, 1]^E\) (Freudenthal simplicial division).

For any set function \( f : 2^E \to \mathbb{R} \), the piecewise-linear function \( \hat{f} \) obtained by linear interpolation on every cell \( \Delta(\sigma) \) is called the Lovász extension (or the Choquet integral) of \( f \).
Theorem (Lovász): For any set function \( f : 2^E \to \mathbb{R} \), \( f \) is a submodular function if and only if its Lovász extension \( \hat{f} \) is convex.
Theorem (Lovász): For any set function $f : 2^E \rightarrow \mathbb{R}$, $f$ is a submodular function if and only if its Lovász extension $\hat{f}$ is convex.

In other words,

**Submodular functions**

$\iff$ **Convex extensible w.r.t. the Freudenthal simplicial division**
Subgradients and Subdifferentials of Submodular Functions

The subdifferential of $f$ at $X \in \mathcal{D}$

$$\partial f(X) = \{x \in \mathbb{R}^E \mid \forall Y \in \mathcal{D} : x(Y) - x(X) \leq f(Y) - f(X)\}$$

Note: $\langle x, \chi_Y - \chi_X \rangle \leq f(\chi_Y) - f(\chi_X), \quad \partial f(X) = \partial \hat{f}(\chi_X)$
Submodular function $f : 2^E \rightarrow \mathbb{R}$

\[\hat{f}\]

The Lovász extension $\hat{f}$ is convex (and linear on every cell $\Delta(\sigma)$)

\[\hat{f}\]

Base polyhedron $B(f)$ (edge vectors $(0, \cdots, 0, \pm 1, 0, \cdots, 0, \mp 1, 0, \cdots, 0)$)

\[\hat{f}\]

Greedy algorithm works
Submodular function $f : 2^E \to \mathbb{R}$

Base polyhedron $B(f)$

$\Downarrow$

Lovász extension $\hat{f}$ is convex

Greedy algorithm

$\Downarrow$

Submodular integrally convex function

Valuated matroid

(Favati-Tardella (1990))

(Dress-Wenzel (1990, 1992))

$\Downarrow$

(Convex conjugate)

$\Downarrow$

$L^{-}-L^{-}\$-convex function $\iff$ $M^{-}/M^{-}\$-convex function

(Murota (1998), F-Murota (2000))

(Murota (1996), Murota-Shioura (1999))

Discrete Convex Analysis (Kazuo Murota)
A simplicial division of the plane (triangulation)

The Freudenthal simplicial division
A simplicial division of the plane (triangulation)

The Freudenthal simplicial division

Consider a function $f$ defined on the integer lattice $\mathbb{Z}^n$. 
Discrete convex functions with respect to Freudenthal simplicial division $= L^\natural$-convex functions defined on $\mathbb{Z}^n$ (due to Murota)

This is equivalent to the

Submodular integrally convex function
due to Favati and Tardella (1990)

($L^\natural$-concave functions are defined similarly.)

The Freudenthal simplicial division of $\mathbb{R}^2$. 

$\rightarrow$
Characterization by **mid-point convexity** due to Favati-Tardella

\[ f(x) + f(y) \geq f\left(\left\lfloor \frac{1}{2}(x + y) \right\rfloor \right) + f\left(\left\lceil \frac{1}{2}(x + y) \right\rceil \right) \quad (\forall x, y \in \mathbb{Z}^n). \]

**Remark:**

\[ x + y = \left\lfloor \frac{1}{2}(x + y) \right\rfloor + \left\lceil \frac{1}{2}(x + y) \right\rceil, \]

\[ \hat{f}\left(\frac{1}{2}(x + y)\right) = \frac{1}{2} \left\{ f\left(\left\lfloor \frac{1}{2}(x + y) \right\rfloor \right) + f\left(\left\lceil \frac{1}{2}(x + y) \right\rceil \right) \right\} \]

\[ \implies \frac{1}{2} \{\hat{f}(x) + \hat{f}(y)\} \geq \hat{f}\left(\frac{1}{2}(x + y)\right) \quad (\forall x, y \in \mathbb{Z}^n). \]
\( f: \textbf{\textit{L}}^n\text{-convex function} \) on integer lattice \( \mathbb{Z}^n \)

\( \hat{f}: \text{the convex extension of } f \text{ on the Freudenthal simplicial division} \)

---

**Convex conjugate \( \hat{f}^* \) of \( \hat{f} \) (or \( f \)) (Legendre-Fenchel transform)**

\[
\hat{f}^*(p) = \sup \{ \langle p, x \rangle - \hat{f}(x) \mid x \in \mathbb{R}^n \} \\
= \sup \{ \langle p, x \rangle - f(x) \mid x \in \mathbb{Z}^n \} \quad (p \in (\mathbb{R}^n)^*)
\]

where \( \langle p, x \rangle = \sum_{i=1}^{n} p(i)x(i) \).

\( \hat{f}^*(p) \) is an \( \textbf{M}^\#\text{-convex function} \) (Murota-Shioura)

---

\( f^*: \text{restriction of convex conjugate } \hat{f}^* \text{ on } (\mathbb{Z}^n)^* \)

\( f \text{ is integer-valued} \implies \hat{f}^* \text{ is the convex extension of } f^* \)

\( f^* \text{ for integer-valued } f \text{ is exactly an integer-valued } \textbf{M}^\#\text{-convex function} \text{ on } (\mathbb{Z}^n)^* \) (Murota-Shioura)

---
\[ \hat{f}^*(p) = \sup\{\langle p, x \rangle - \hat{f}(x) \mid x \in \mathbb{R}^n \} \]
\[ = \sup\{\langle p, x \rangle - f(x) \mid x \in \mathbb{Z}^n \} \quad (p \in (\mathbb{R}^n)^*) \]

For \( x_0 \in \mathbb{Z}^n \) the set of all \( p \in (\mathbb{R}^n)^* \) satisfying
\[ \hat{f}^*(p) = \langle p, x_0 \rangle - f(x_0), \quad \text{i.e.} \]
\[ f(x) \geq f(x_0) + \langle p, x - x_0 \rangle \quad (\forall x \in \mathbb{Z}^n) \]

is the subdifferential \( \partial \hat{f}(x_0)(= \partial f(x_0)) \) at \( x_0 \).
**Corollary:** For a pointed polyhedron $P \subset \mathbb{R}^E$, $P$ is a generalized polymatroid

\[ \uparrow \]

all the edge vectors of $P$ are of form

$(0, \cdots, 0, \pm 1, 0, \cdots, 0, \mp 1, 0, \cdots, 0)$ or $(0, \cdots, 0, \pm 1, 0, \cdots, 0)$
(Recall)

**Corollary:** For a pointed polyhedron $P \subset \mathbb{R}^E$,

- $P$ is a generalized polymatroid

  \[
  \uparrow
  \]

- all the edge vectors of $P$ are of form $(0, \cdots, 0, \pm 1, 0, \cdots, 0, \mp 1, 0, \cdots, 0)$ or $(0, \cdots, 0, \pm 1, 0, \cdots, 0)$

---

**Hence,** the subdifferential $\partial f(z)$ of an $L^\natural$-convex function $f$ at an integral point $z$ is a **generalized polymatroid**.

$M^\natural$-convex function $\hat{f}^*$ is an **affine function** on every such **generalized polymatroid**.

**Remark:** If $f$ is an **integer-valued** function, $\partial f(z)$ is an **integral** generalized polymatroid.
$M^2$-concave function $g$
Simultaneous Exchange Axiom for $M^\bullet$-convex functions

$f : \mathbb{Z}^n \to \mathbb{R} \cup \{+\infty\}$ (Murota-Shioura)

\[ \text{dom } f = \{ x \mid f(x) < +\infty \} \]
\[ \text{supp}^+(x) = \{ i \mid x(i) > 0 \}, \quad \text{supp}^-(x) = \{ i \mid x(i) < 0 \} \]

\[ \text{(M}^\bullet\text{-EXC)} \quad \text{For } x, y \in \text{dom } f \text{ and } i \in \text{supp}^+(x - y), \]
\[ f(x) + f(y) \geq \min \left[ f(x - \chi_i) + f(y + \chi_i), \right. \]
\[ \left. \min_{j \in \text{supp}^-(x - y)} \{ f(x - \chi_i + \chi_j) + f(y + \chi_i - \chi_j) \} \right]. \]

\[ \text{(\rightarrow Simultaneous Exchange Axiom for Generalized Polymatroids)} \]
Discrete Separation Theorem (L^5)

\( f \): an integer-valued \( L^5 \)-convex function on \( \mathbb{Z}^n \)

\( g \): an integer-valued \( L^5 \)-concave function on \( \mathbb{Z}^n \)

Suppose \( f \geq g \).
Suppose \( x^* \) minimizes \( f(x) - g(x) \) over \( \mathbb{Z}^n \).
Common sub- and supergradient $p \in \partial f(x^*) \cap \partial g(x^*)$
Intersections of subdifferentials and superdifferentials
(integral generalized polymatroids)
There exists a common integral sub- and supergradient 
\[ p \in \partial f(x^*) \cap \partial g(x^*) \]
due to the following integrality:

**Intersection Theorem for Generalized Polymatroids** (Edmonds, Frank):
The intersection of any two integral generalized polymatroids is integral if it is nonempty.
There exists a common integral sub- and supergradient
\[ p \in \partial f(x^*) \cap \partial g(x^*) \]
due to the following integrality:

**Intersection Theorem for Generalized Polymatroids** (Edmonds, Frank):
The intersection of any two integral generalized polymatroids is integral if it is nonempty.

Hence,
\[ g(x^*) - \langle p, x^* \rangle \leq \beta \leq f(x^*) - \langle p, x^* \rangle \]
\( (p, x^*): \text{integral vectors} \implies \text{There exists an integer } \beta \).
There exists a common integral sub- and supergradient
\[ p \in \partial f(x^*) \cap \partial g(x^*) \]
due to the following integrality:

**Intersection Theorem for Generalized Polymatroids** (Edmonds, Frank):
The intersection of any two integral generalized polymatroids is integral if it is nonempty.

Hence,
\[ g(x^*) - \langle p, x^* \rangle \leq \beta \leq f(x^*) - \langle p, x^* \rangle \]
\((p, x^*): \text{integral vectors} \implies \text{There exists an integer } \beta.\)

**Discrete Separation Theorem** (Murota):
\[ \forall z \in \mathbb{Z}^n : f(z) \geq g(z) \]
\[ \implies \exists (p \in (\mathbb{Z}^n)^*, \beta \in \mathbb{Z}) : f(z) \geq \langle p, z \rangle + \beta \geq g(z) \ (\forall z \in \mathbb{Z}^n) \]
**Discrete Fenchel Duality Theorem** (Murota):

For any integer-valued $L^\natural$-convex function $f : \mathbb{Z}^n \to \mathbb{Z}$ and $L^\natural$-concave function $g : \mathbb{Z}^n \to \mathbb{Z}$,

$$\inf \{ f(x) - g(x) \mid x \in \mathbb{Z}^n \} = \sup \{ g^\circ(p) - f^\bullet(p) \mid p \in (\mathbb{Z}^n)^* \}.$$
Discrete Fenchel Duality Theorem (Murota):

For any integer-valued $L^\natural$-convex function $f : \mathbb{Z}^n \to \mathbb{Z}$ and $L^\natural$-concave function $g : \mathbb{Z}^n \to \mathbb{Z}$,

$$\inf\{f(x) - g(x) \mid x \in \mathbb{Z}^n\} = \sup\{g^\circ(p) - f^\bullet(p) \mid p \in (\mathbb{Z}^n)^*\}.$$
For more information see the following monographs.
