劣モジュラ最適化

岩田　覚
(京都大学数理解析研究所)

Outline

- Submodular Functions
  - Examples
  - Discrete Convexity
- Minimizing Submodular Functions
- Symmetric Submodular Functions
- Maximizing Submodular Functions
- Approximating Submodular Functions

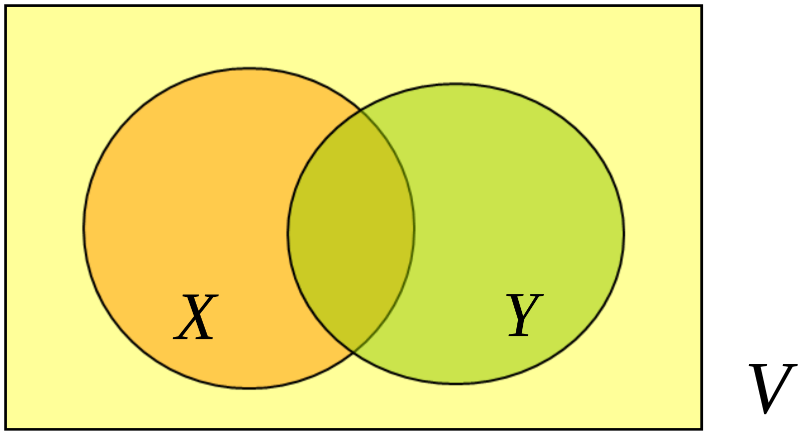
Submodular Functions

V : Finite Set

$f : 2^V \to \mathbb{R}$ $\qquad \forall X, Y \subseteq V$

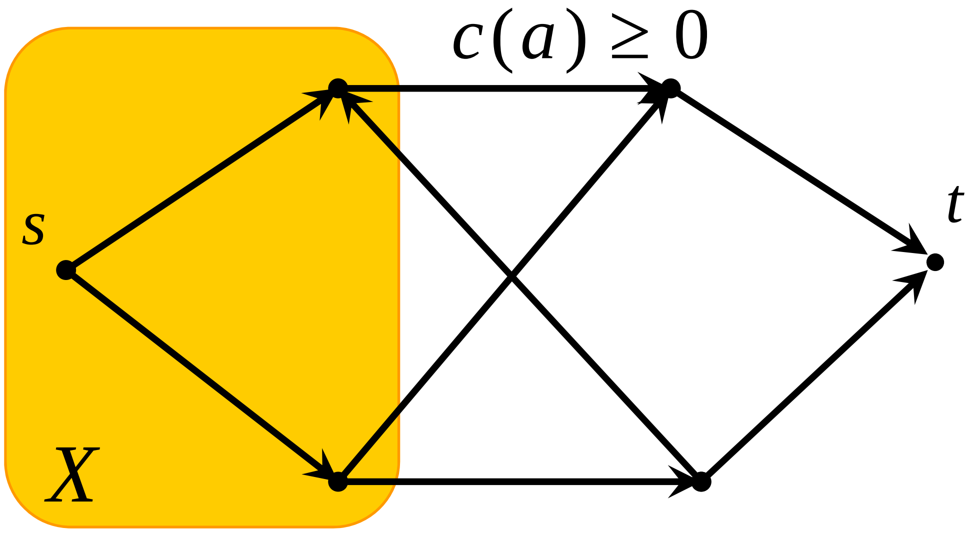
$$f(X) + f(Y) \geq f(X \cap Y) + f(X \cup Y)$$

- Cut Capacity Functions
- Matroid Rank Functions
- Entropy Functions



Cut Capacity Function

Cut Capacity $\kappa(X) = \sum \{c(a) \mid a : \text{leaving } X\}$



Max Flow Value = Min Cut Capacity

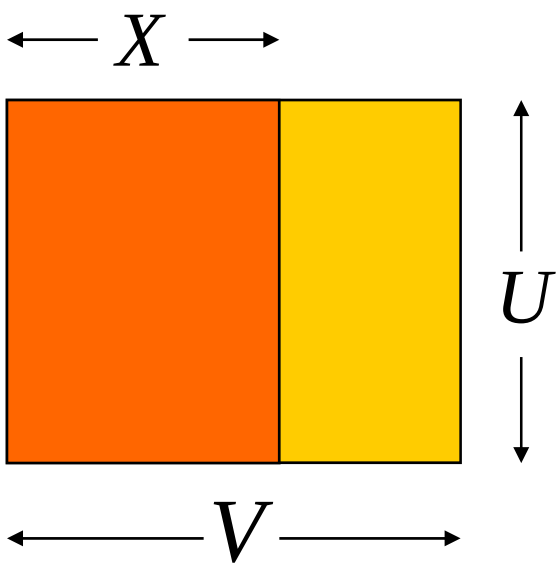
Matroid Rank Functions

Matrix Rank Function

Whitney (1935)

$$\rho(X) = \operatorname{rank} A[U, X]$$

$A =$



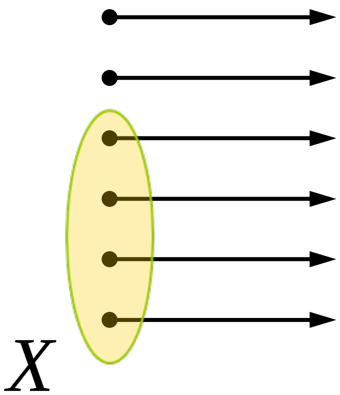
$$\forall X \subseteq V, \rho(X) \leq |X|$$

$$X \subseteq Y \Rightarrow \rho(X) \leq \rho(Y)$$

ρ : Submodular

Entropy Functions

Information Sources



$$h(\phi) = 0$$

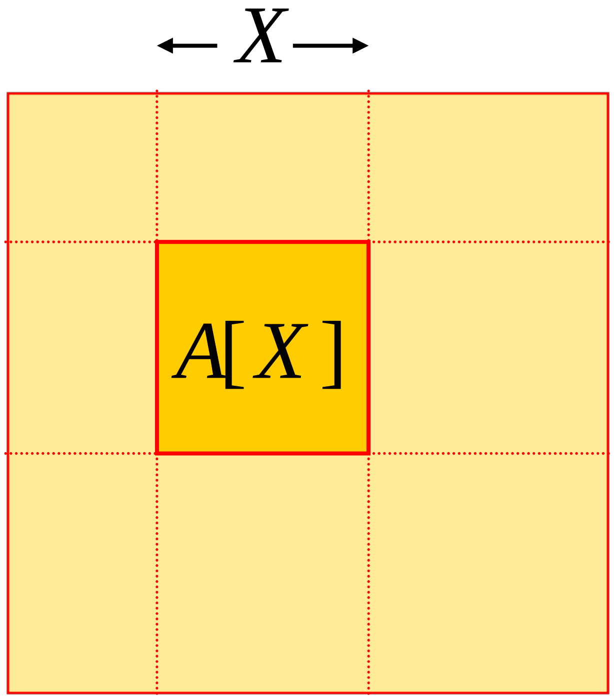
$h(X)$: Entropy of the Joint Distribution

$$h(X) + h(Y) \geq h(X \cap Y) + h(X \cup Y)$$

Conditional Mutual Information ≥ 0

Positive Definite Symmetric Matrices

$A =$



$$f(\phi) = 0$$

$$f(X) = \log\det A[X]$$

Ky Fan's Inequality

$$f(X) + f(Y) \geq f(X \cap Y) + f(X \cup Y)$$

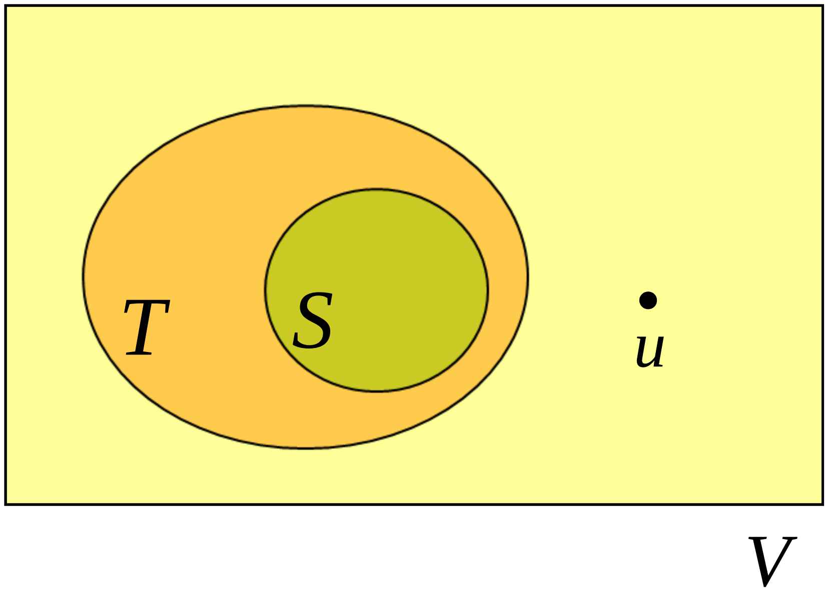
Extension of the Hadamard Inequality

$$\det A \leq \prod_{i \in V} A_{ii}$$

Discrete Concavity

$$S \subseteq T \Rightarrow$$

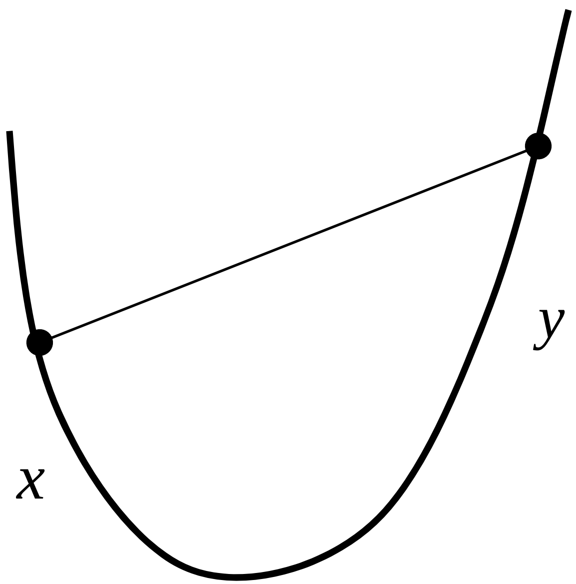
$$f(S \cup \{u\}) - f(S) \geq f(T \cup \{u\}) - f(\{u\})$$



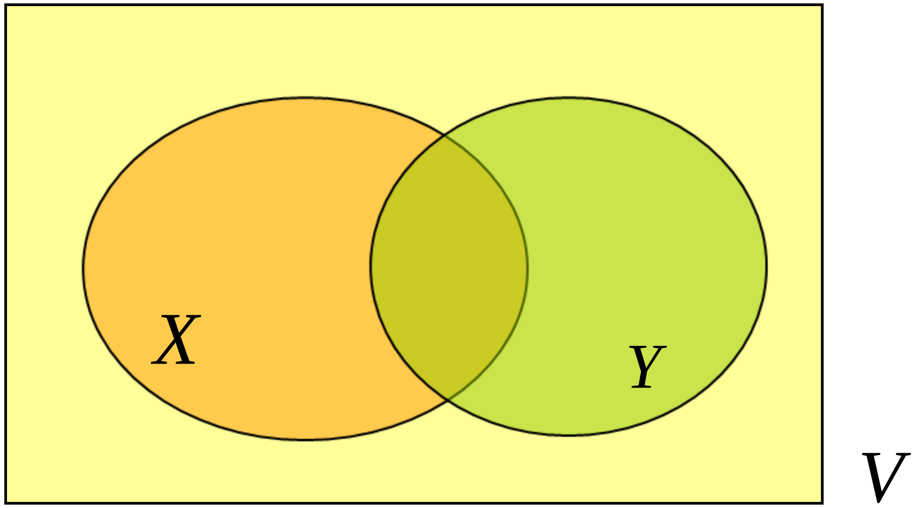
Diminishing Returns

Discrete Convexity

Convex Function

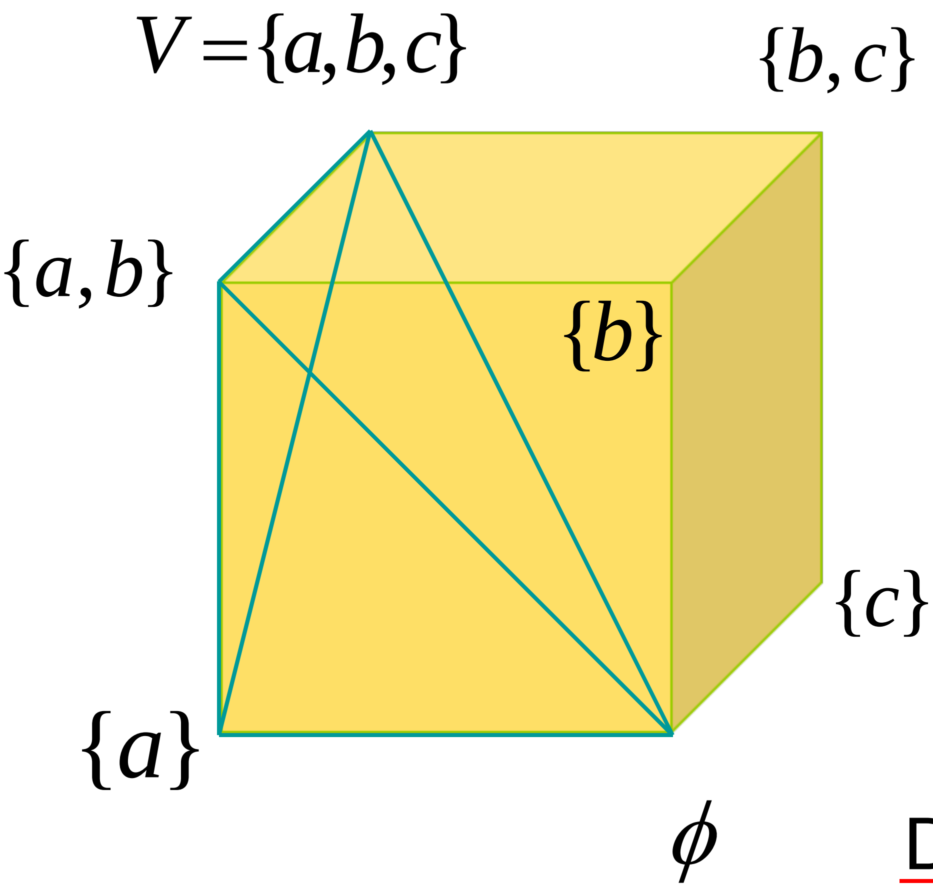


$$f(X) + f(Y) \geq f(X \cap Y) + f(X \cup Y)$$



Discrete Convexity

Lovász (1983)



$\hat{f}$: Linear Interpolation

$\hat{f}$: Convex

⇕

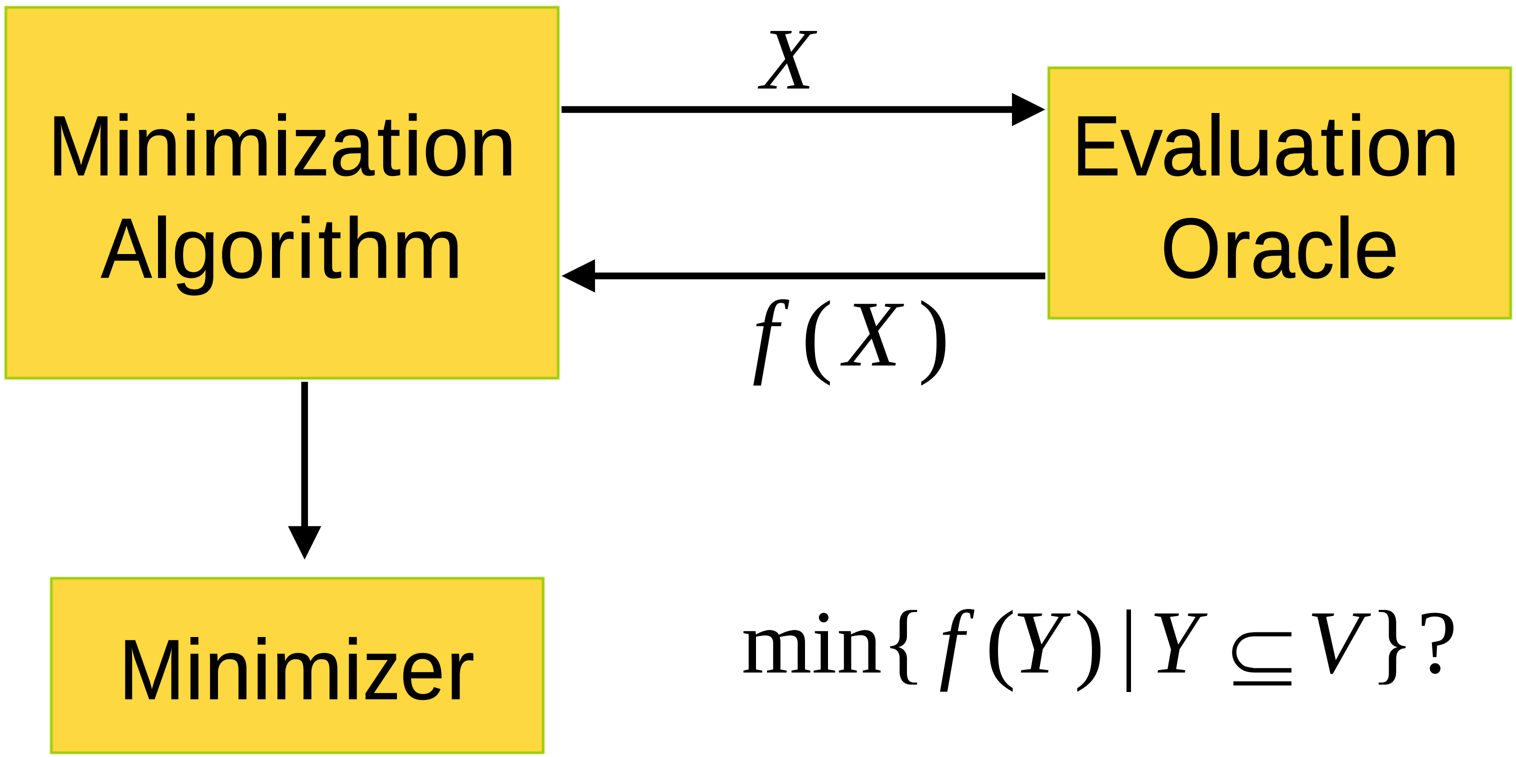
f : Submodular

Discrete Convex Analysis

Murota (2003)

Submodular Function Minimization

Assumption: $f(\phi) = 0$



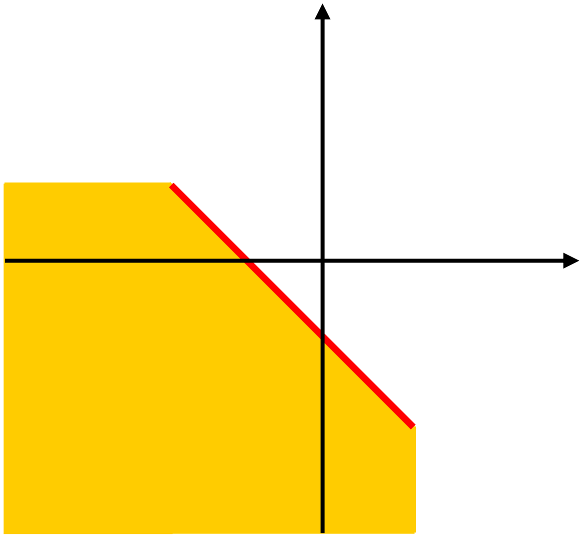
$$\min\{ f(Y) \mid Y \subseteq V \}?$$

Ellipsoid Method

Grötschel, Lovász, Schrijver (1981)

Base Polyhedra

$$\mathbf{R}^V = \{x \mid V \to \mathbf{R}\}$$



$$x(Y) = \sum_{v \in Y} x(v)$$

Submodular Polyhedron

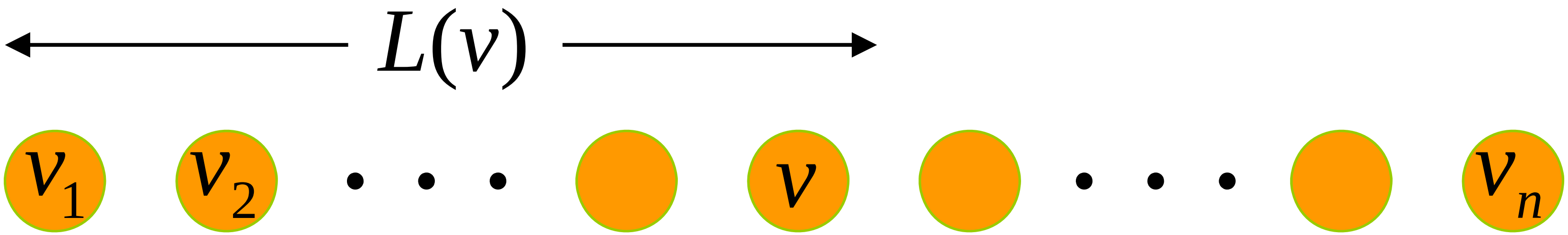
$$P(f) = \{x \mid x \in \mathbf{R}^V, \forall Y \subseteq V, x(Y) \le f(Y)\}$$

Base Polyhedron

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

Greedy Algorithm

Edmonds (1970)
Shapley (1971)



$$y(v) = f(L(v)) - f(L(v) - \{v\}) \qquad (v \in V)$$

y : Extreme Base

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} y(v_1) \\ y(v_2) \\ \vdots \\ y(v_n) \end{bmatrix} = \begin{bmatrix} f(L(v_1)) \\ f(L(v_2)) \\ \vdots \\ f(L(v_n)) \end{bmatrix}$$

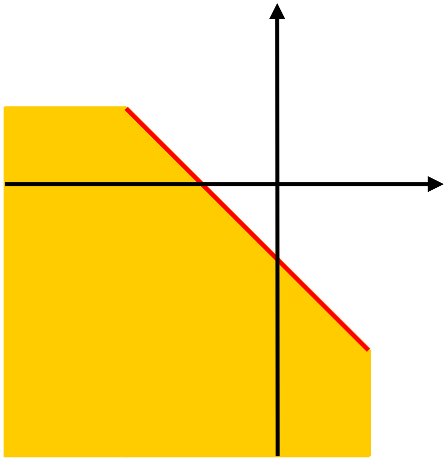
Min-Max Theorem

Theorem Edmonds (1970)

$$\min_{Y \subseteq V} f(Y) = \max\{x^-(V) \mid x \in B(f)\}$$

$$x^-(v) := \min\{0, x(v)\}$$

$$x^-(V) \le x(Y) \le f(Y)$$

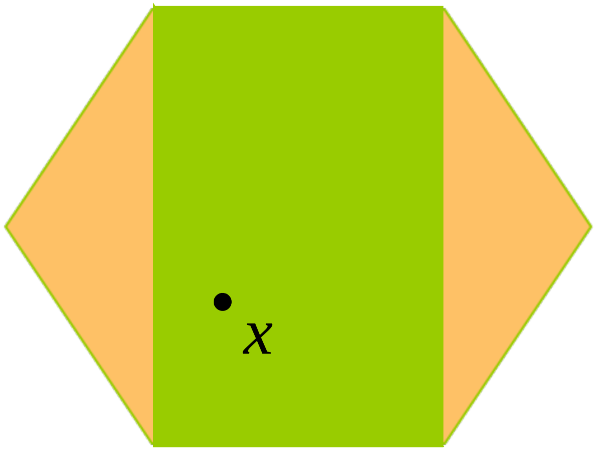


Combinatorial Approach

Extreme Base $y_L \in B(f)$

Convex Combination

$$x = \sum_{L \in \Lambda} \lambda_L y_L$$

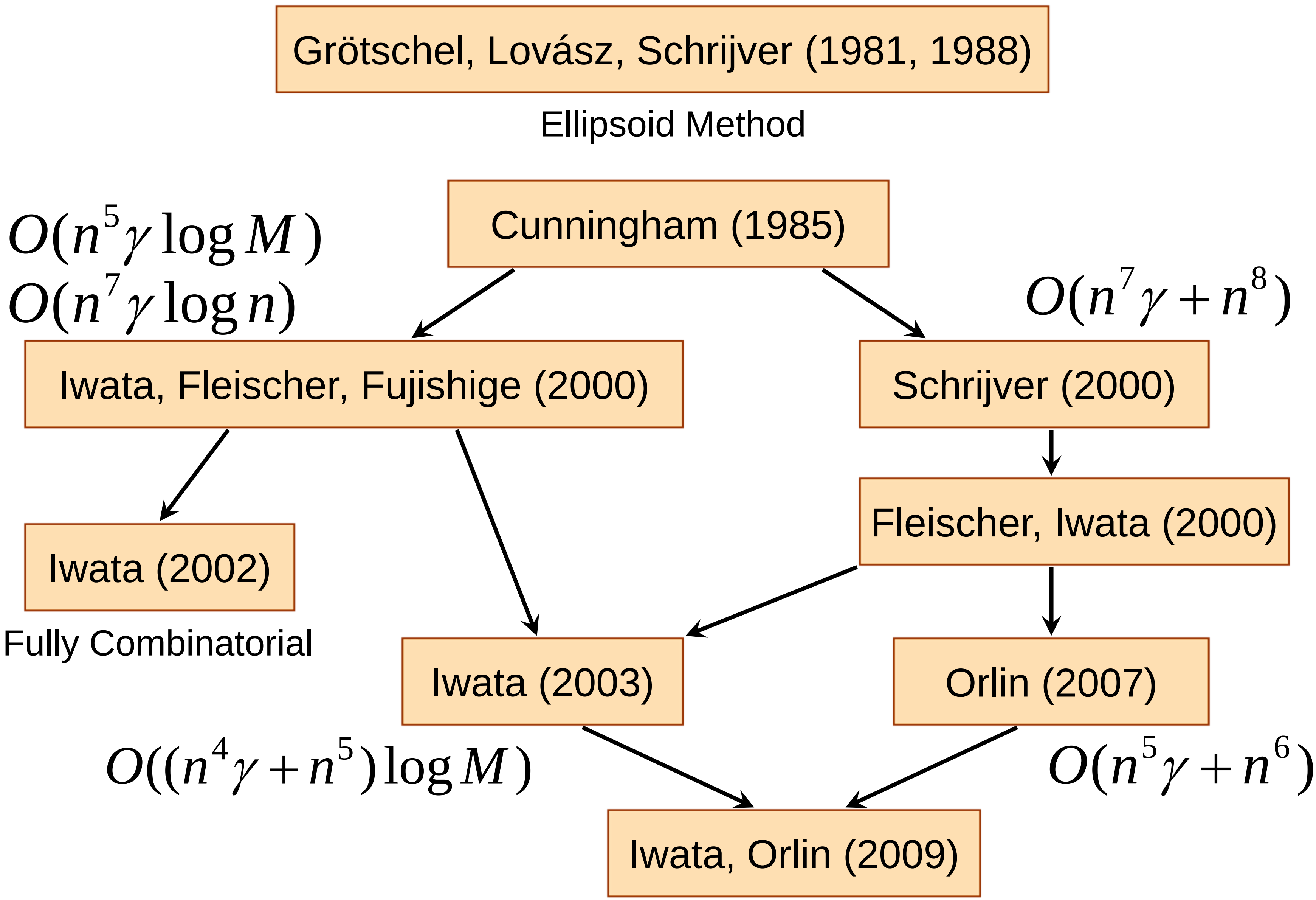


Cunningham (1985)

$$O(n^6 M\gamma \log nM)$$

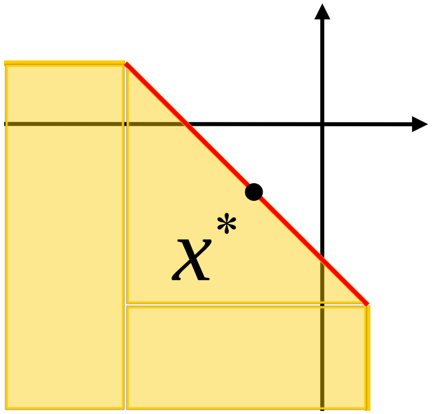
$$M = \max_{X \subseteq V} |f(X)|$$

Submodular Function Minimization



The Minimum-Norm Base

$$\text{Minimize } \|x\|^2$$
$$\text{subject to } x \in B(f)$$



Theorem   Fujishige (1984)

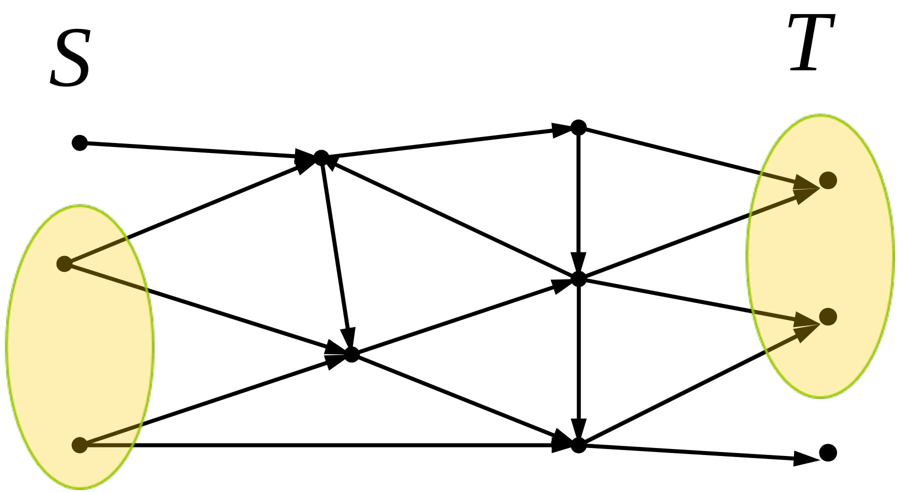
x^* : opt.sol.

$X_- := \{v \mid x^*(v) < 0\}$ ⟶ Minimal Minimizer

$X_0 := \{v \mid x^*(v) \leq 0\}$ ⟶ Maximal Minimizer

Wolfe’s Algorithm ⟶ Practically Efficient

Evacuation Problem（Dynamic Flow）



Hoppe, Tardos (2000)

$c(a)$: Capacity

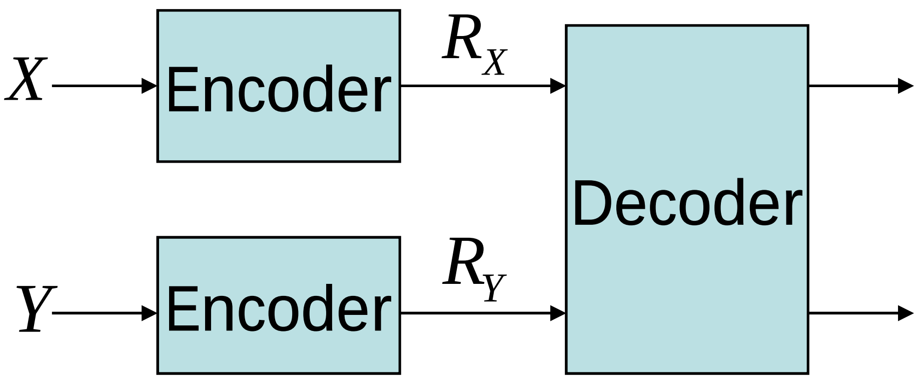
$\tau(a)$: Transit Time

$b(v)$: Supply/Demand

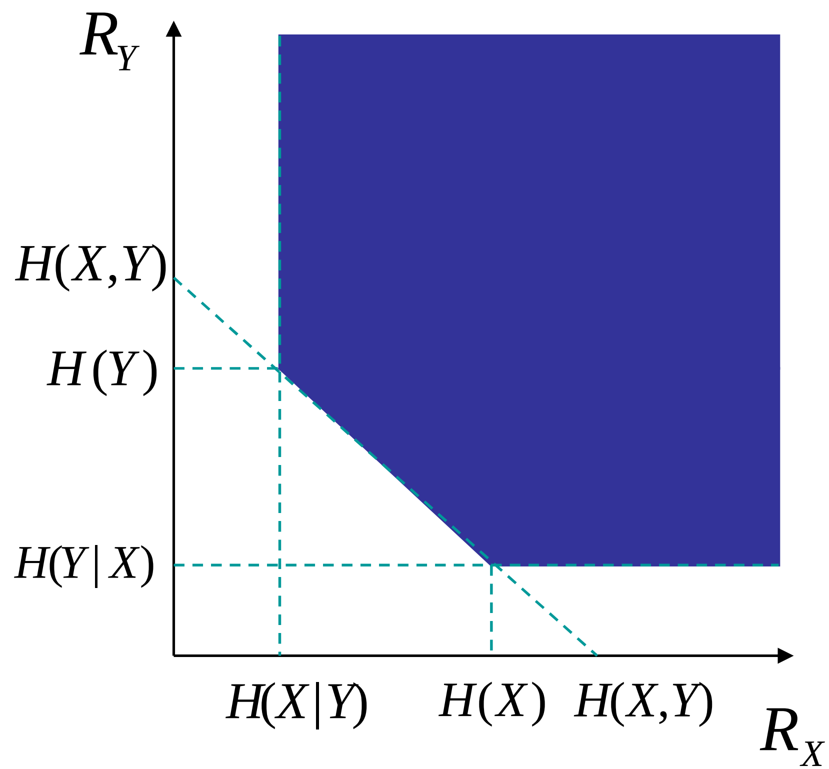
$o(X)$: Maximum Amount of Flow from $X \cap S$ to $T \setminus X$.

Feasible $\Longleftrightarrow$ $b(X) \leq o(X), \forall X \subseteq S \cup T$

Multiterminal Source Coding



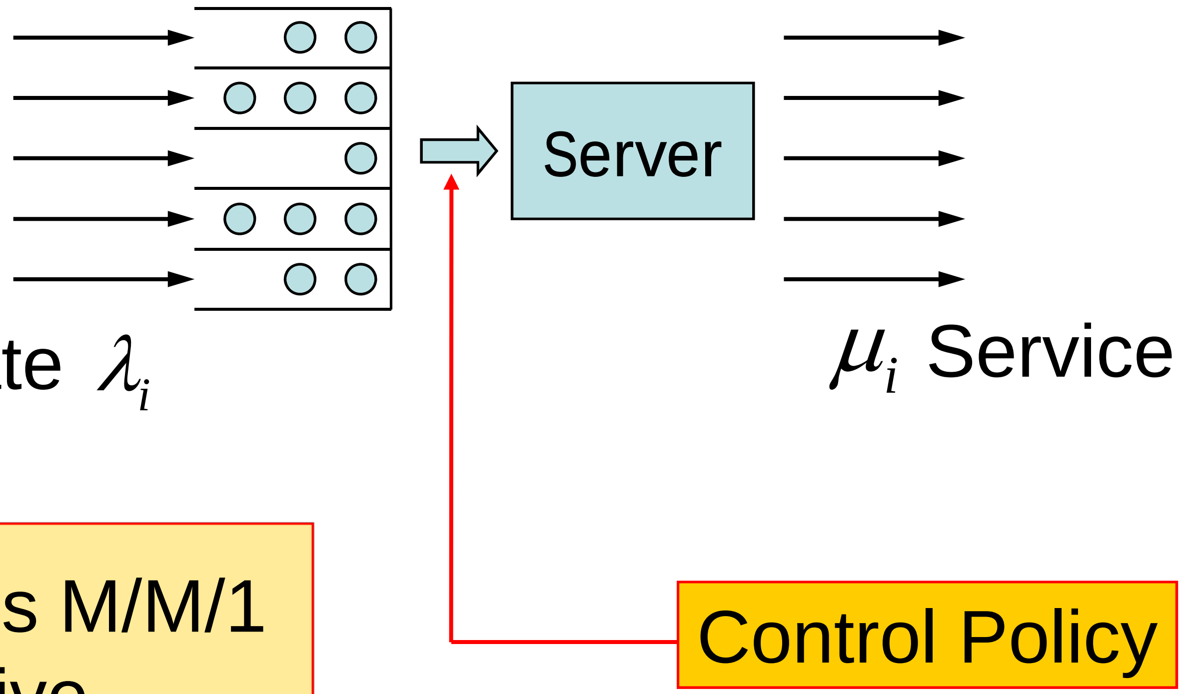
Slepian, Wolf (1973)



Multiclass Queueing Systems

Arrival Interval

Service Time



Arrival Rate λ_i

μ_i Service Rate

Multiclass M/M/1
Preemptive

Performance Region

s_j : Expected Staying Time of a Job in *j*

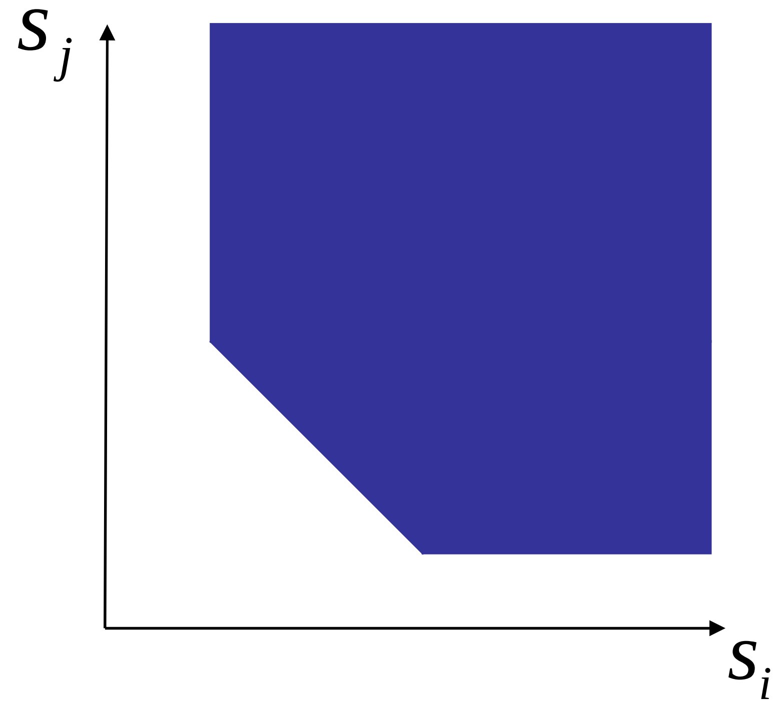
$$\rho_i := \lambda_i / \mu_i, \quad \sum_{i \in V} \rho_i < 1$$

S : Achievable

⇕

$$\sum_{i \in X} \rho_i S_i \geq \frac{\sum_{i \in X} \rho_i / \mu_i}{1 - \sum_{i \in X} \rho_i}, \forall X \subseteq V$$

Coffman, Mitrani (1980)



A Class of Submodular Functions

$x, y, z \in \mathrm{R}_+^V$

Itoko & Iwata (2005)

$h:$ Nonnegative, Nondecreasing, Convex

$$f(X) = z(X) - y(X)h(x(X)) \qquad (X \subseteq V)$$

→ Submodular

$$\sum_{i \in X} \rho_i S_i \geq \frac{\sum_{i \in X} \rho_i / \mu_i}{1 - \sum_{i \in X} \rho_i}, \forall X \subseteq V$$

$$z_i := \rho_i S_i \qquad y_i := \frac{\rho_i}{\mu_i}$$

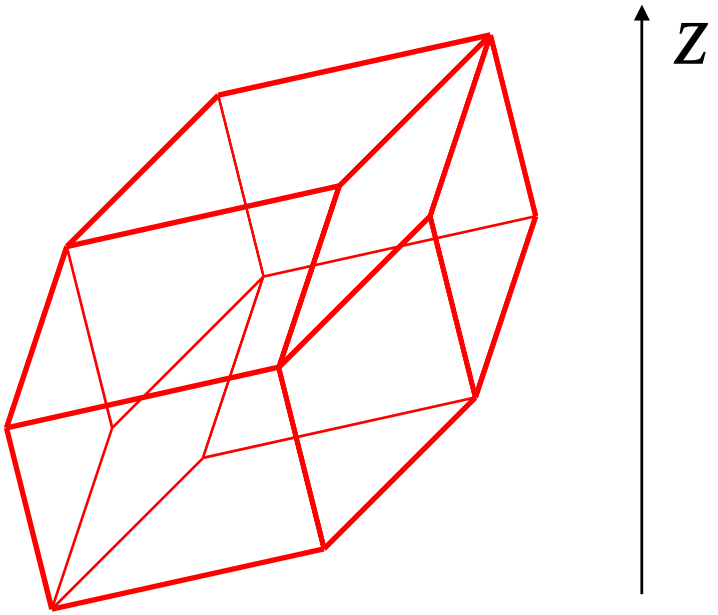
$$x_i := \rho_i \qquad h(x) := \frac{1}{1-x}$$

Zonotope in 3D

$$w(X) = (x(X), y(X), z(X))$$

$$Z = \mathrm{conv}\{w(X) \mid X \subseteq V\}$$

Zonotope

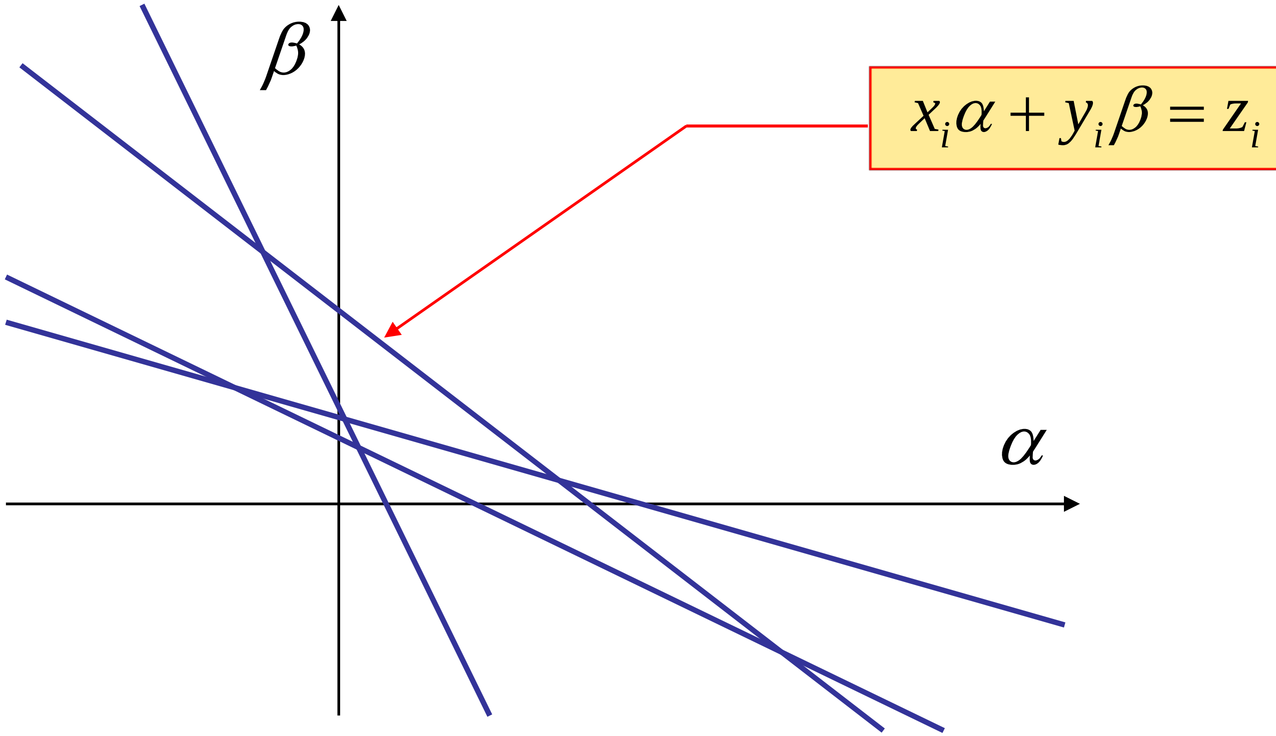


$$\tilde{f}(x, y, z) = z - yh(x)$$

$$\min\{f(X) \mid X \subseteq V\}$$
$$= \min\{\tilde{f}(x, y, z) \mid (x, y, z) : \text{Lower Extreme Point of } Z\}$$

Remark: $\tilde{f}(x, y, z)$ is NOT concave!

Line Arrangement



Enumerating All the Cells

Topological Sweeping Method
Edelsbrunner, Guibas (1989)

$\rightarrow O(n^2)$

Symmetric Submodular Functions

$$f : 2^V \to \mathbf{R}$$

Symmetric $\quad f(X) = f(V \setminus X), \quad \forall X \subseteq V.$

Crossing Submodular

$$X \cap Y \neq \phi,\ X \cup Y \neq V \Rightarrow$$
$$f(X) + f(Y) \geq f(X \cup Y) + f(X \cap Y)$$

Symmetric Submodular Function Minimization

$$\min\{ f(X) \mid \phi \neq X \subset V, X \neq V \}?$$

Maximum Adjacency Ordering

- Minimum Cut Algorithm by MA-ordering

  Nagamochi & Ibaraki (1992)
- Simpler Proofs

  Frank (1994), Stoer & Wagner (1997)
- Symmetric Submodular Functions

  Queyranne (1998)
- Alternative Proofs

  Fujishige (1998), Rizzi (2000)

Minimum Degree Ordering

Nagamochi (2007)

ISAAC'07, Sendai, Japan

Finding the family of all extreme sets for symmetric crossing submodular functions in $O(n^3\gamma)$ time.

⇓

Symmetric Submodular Function Minimization

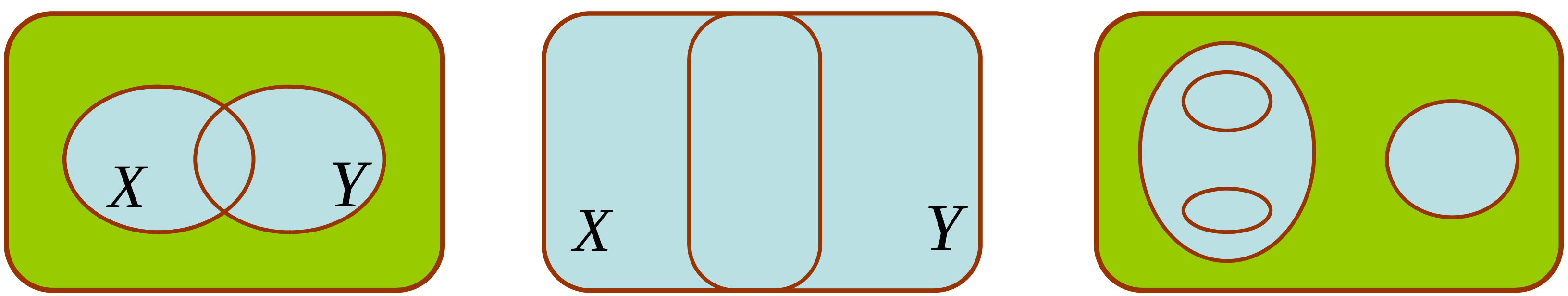
Extreme Sets

f : Symmetric Crossing Submodular Function

X : Extreme Set

$$f(Z) > f(X), \quad \forall Z \subset X : \phi \neq Z \neq X.$$

The family of all extreme sets forms a laminar.



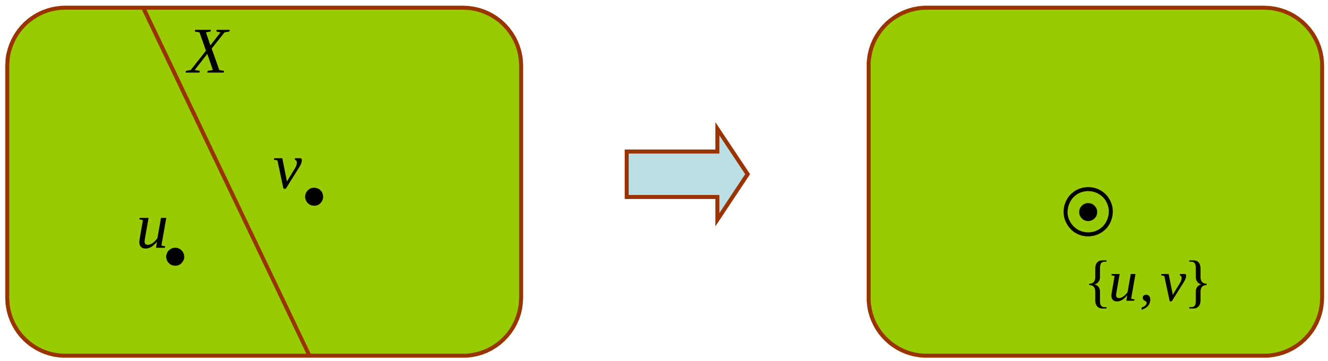
$$f(X) + f(Y) \geq f(X \setminus Y) + f(Y \setminus X)$$

Flat Pair for Symmetric Submodular Functions

Flat Pair $\{u, v\} \subseteq V \quad (u \neq v)$

$$f(X) \geq \min\{f(x) \mid x \in X\},$$

$$\forall X \subseteq V \text{ s.t.} \mid X \cap \{u, v\} \mid = 1.$$



No Extreme Sets Separate u and v.

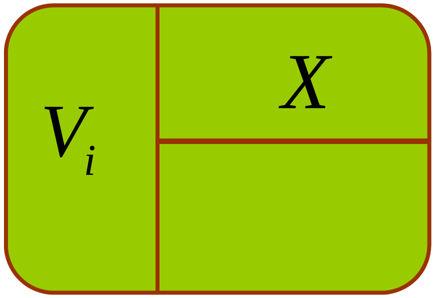
Shrink $\{u, v\}$ into a single vertex.

MD-Ordering for Symmetric Submodular Functions

MD-ordering

$$V_i := \{v_1, \ldots, v_i\}$$

$$v_1, v_2, \ldots, v_{n-1}, v_n \in V$$



Each v_j has minimum value of $f_{j-1}(v)$ among $v \in V \setminus V_{j-1}$.

$$f_i(X) := f(X) + f(V_i \cup X) \qquad (X \subseteq V \setminus V_i)$$

Symmetric, Crossing Submodular

MD-Ordering for Symmetric Submodular Functions

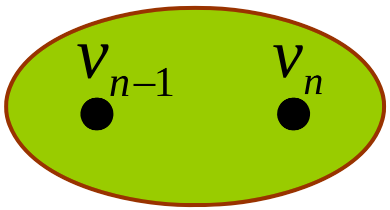
The last two vertices v_{n-1}, v_n of an MD-ordering form a flat pair.

Proof by Induction:

$\{v_{n-1}, v_n\}$: Flat Pair for f_i on $V \setminus V_i$

$$i = n-2, \ldots, 1, 0.$$

$$i = n-2$$



Time Complexity

- Finding an MD-ordering in $O(n^2\gamma)$ time.

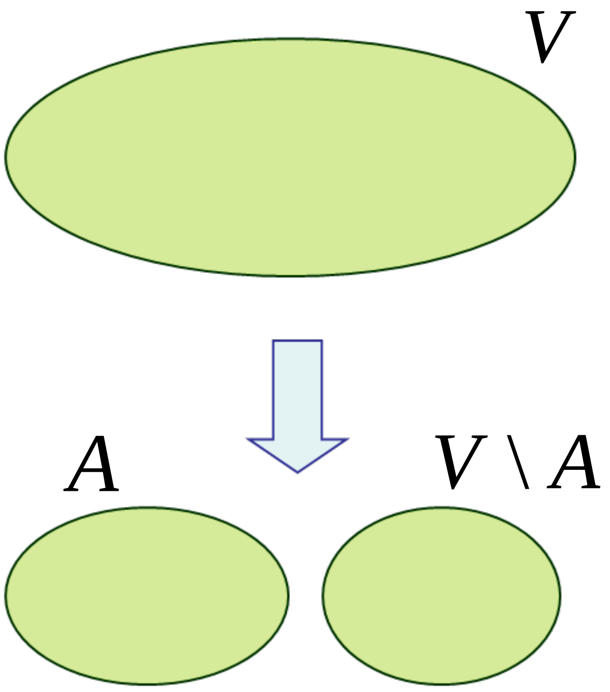
- Finding all the extreme sets in $O(n^3\gamma)$ time.

- Minimizing symmetric submodular functions in $O(n^3\gamma)$ time.

Application to Clustering

Random variables $X_1, \ldots, X_n$

Partition V into A and $V \setminus A$ as independent as possible



Minimize $I(X_A; X_{V \setminus A})$

subject to $\phi \neq A \neq V$

$$I(X_A; X_B) = H(X_B) - H(X_B \mid X_A)$$
$$= H(X_A) + H(X_B) - H(X_A, X_B)$$

Application to Clustering

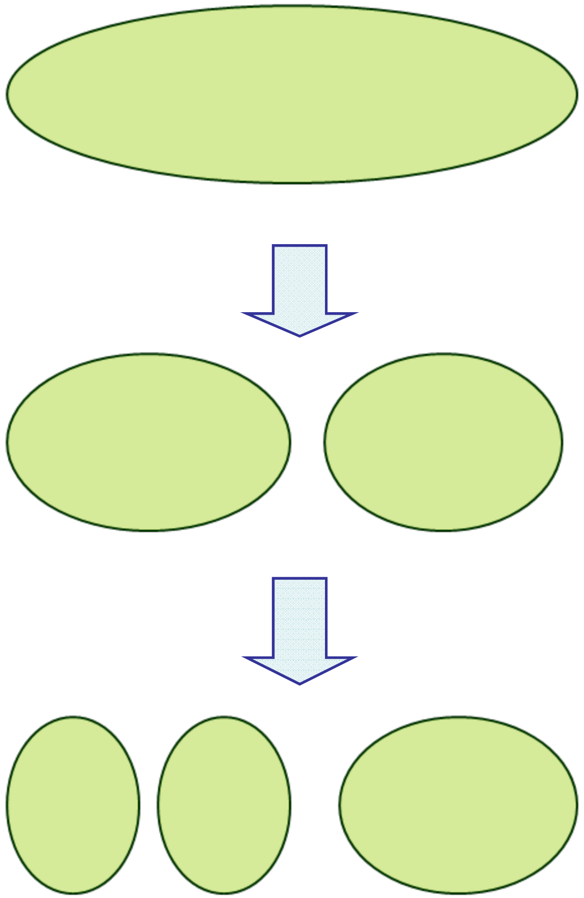
Minimize $$\sum_{i=1}^{k} f(V_i)$$

subject to $$V = V_1 \cup \cdots \cup V_k$$

$$V_i \cap V_j = \phi \quad (i \neq j)$$

$(2-1/k)$-Approximation

Greedy Split



Submodurlar Function Maximization

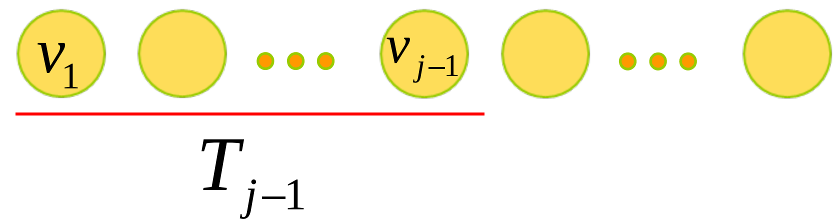
Approximation Algorithms

- Nemhauser, Wolsey, Fisher (1978)
  - Monotone SF / Cardinality Constraint
  - (1−1/e)-Approximation
- Feige, Mirrokni, Vondrák (FOCS 2007)
  - Nonnegative SF
  - 2/5-Approximation
- Vondrák (STOC 2008)
  - Monotone SF / Matroid Constraint
  - (1−1/e)-Approximation

Approximate Maximization

Maximize $f(S)$ subject to $|S| \le k$

Greedy Algorithm



$$T_0 := \phi$$

$$\left.\begin{aligned} v_j &:= \arg\max f(T_{j-1} \cup \{v\}) \\ T_j &:= T_{j-1} \cup \{v_j\} \\ \rho_j &:= f(T_j) - f(T_{j-1}) \end{aligned}\right\} \quad j = 1, \ldots, k$$

$$f(T_k) \ge \left(1 - 1/\mathrm{e}\right) f(S), \quad \forall S : |S| \le k$$

Approximate Maximization

S^* : Optimal Solution $\eta := f(S^*)$

$$\eta \le k\rho_j + \sum_{i=1}^{j-1} \rho_i \quad (j = 1, \ldots, k)$$

$$f(T_k) = \sum_{i=1}^{k} \rho_i$$

$$\because) \; f(S^*) \le f(S^* \cup T_{j-1})$$

$$\le f(T_{j-1}) + \sum_{u \in S^* \setminus T_{j-1}} \left[f(T_{j-1} \cup \{u\}) - f(T_{j-1}) \right]$$

$$\le \underline{f(T_{j-1})} + k\rho_j$$

$$\to \sum_{i=1}^{j-1} \rho_i$$

Approximate Maximization

Minimize $\sum_{i=1}^{k} \rho_i$

subject to
$$\begin{bmatrix} k & 0 & \cdots & 0 \\ 1 & k & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 1 & \cdots & 1 & k \end{bmatrix} \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_k \end{bmatrix} \geq \begin{bmatrix} \eta \\ \eta \\ \vdots \\ \eta \end{bmatrix}$$

$$\rho_j \geq 0 \qquad (j = 1, \ldots, k)$$

$$\sum_{i=1}^{k} \rho_i = f(T_k)$$

$$\eta := f(S^*)$$

$$\hat{\rho}_j := \frac{\eta}{k}\left(1 - \frac{1}{k}\right)^{j-1}$$

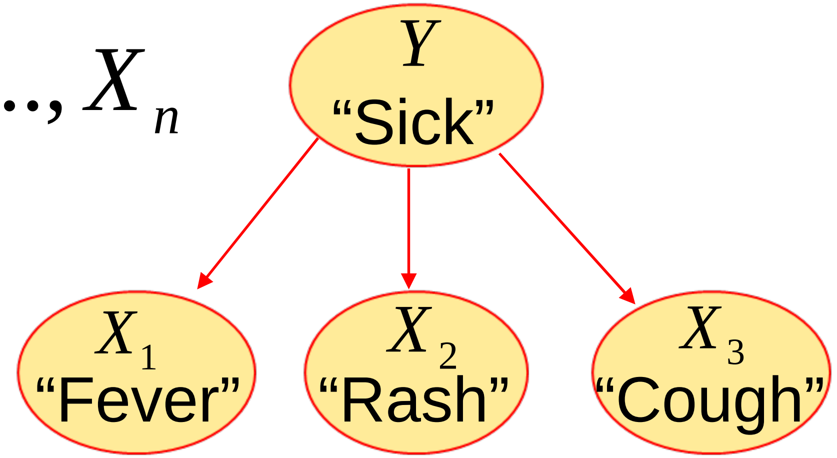
$$(j = 1, \ldots, k)$$

$$\sum_{i=1}^{k} \hat{\rho}_i = \eta\left[1 - \left(1 - \frac{1}{k}\right)^k\right] \geq \eta\left(1 - \frac{1}{\mathrm{e}}\right)$$

Feature Selection

Random Variables $Y, X_1, \ldots, X_n$

Predict Y from subset X_A



Maximize $I(X_A; Y)$

subject to $|A| \leq k$

X : Conditionally Independent

$$I(X_A; Y) := H(Y) - H(Y \mid X_A)$$

$$= H(X_A) - H(X_A \mid Y)$$

Submodular

Krause & Guestrin (2005)

Submodular Welfare Problem

Utility Functions $f_1, \ldots, f_k$

(Monotone, Submodular)

Maximize $$\sum_{i=1}^{k} f(V_i)$$

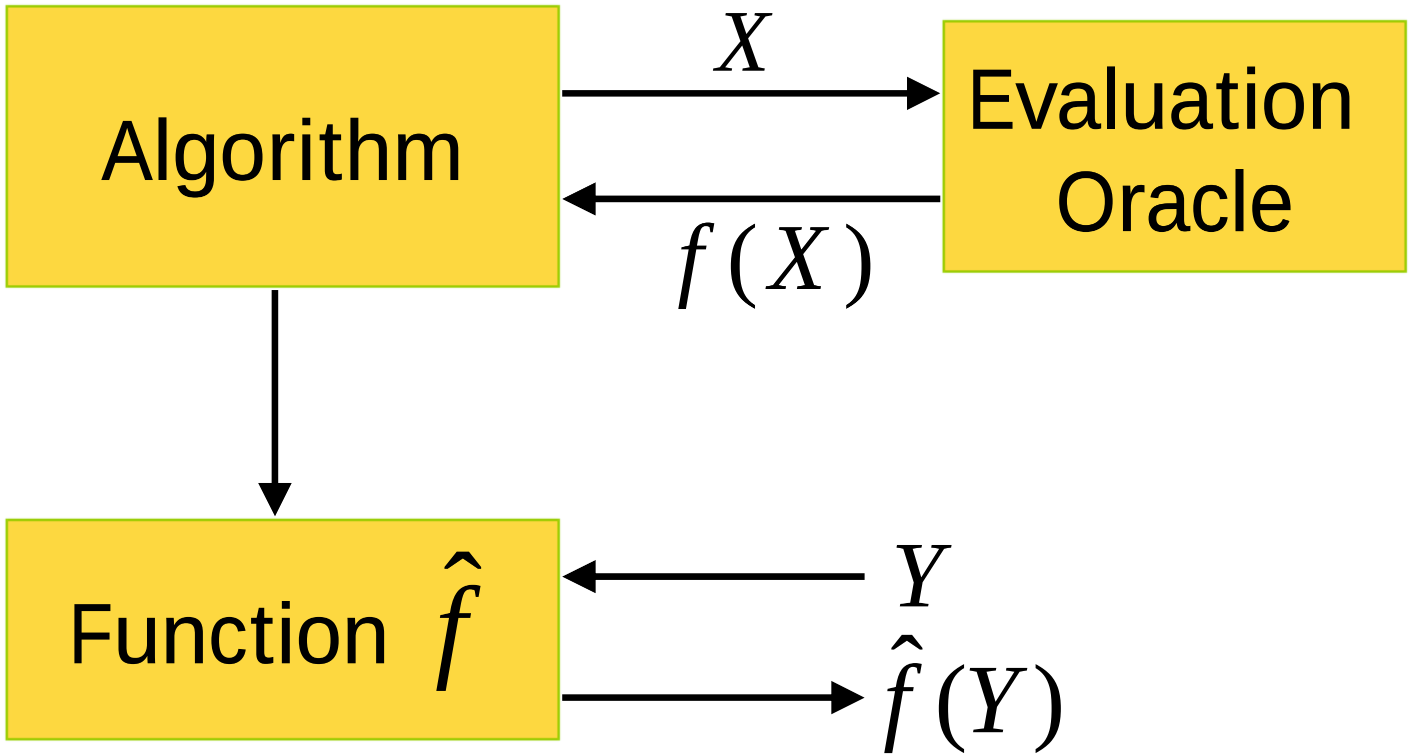
subject to $$V = V_1 \cup \cdots \cup V_k$$

$$V_i \cap V_j = \phi \ (i \neq j)$$

$(1 - 1/\mathrm{e})$ -Approximation

Vondrák (2008)

Approximating Submodular Functions



Approximating Submodular Functions

Assumption $f(\phi) = 0, \quad f(X) \geq 0, \forall X \subseteq V.$

Problem

Construct a set function $\hat{f}$ such that

$$\hat{f}(X) \leq f(X) \leq \alpha(n)\hat{f}(X), \quad \forall X \subseteq V.$$

For what function α is this possible?

Remarks

$\alpha(n) = 1$ for cut capacity functions

$\alpha(n) = n$ for general monotone submodular functions

Approximating Submodular Functions

Goemans, Harvey, Iwata, Mirrokni (2009)

- Algorithm with $\alpha(n) = \sqrt{n+1}$ for matroid rank functions.
- Algorithm with $\alpha(n) = O(\sqrt{n}\log n)$ for monotone submodular functions.
- No polynomial algorithm can achieve a factor better than $\alpha(n) = \Omega(\sqrt{n}/\log n)$ even for matroid rank functions.

Submodular Load Balancing

Svitkina & Fleischer (2008)

$f_1, \ldots, f_m$: Monotone Submodular Functions

$$\min_{\{V_1, \ldots, V_m\}} \max_j f_j(V_j)?$$

$$f_j(X) := \sum_{v \in X} p_{jv}$$ → Scheduling

2-Approximation Algorithm

Lenstra, Shmoys, Tardos (1990)

$O(\sqrt{n} \log n)$-Approximation Algorithm

Pointers

- S. Fujishige: *Submodular Functions and Optimization*, Elsevier, 2005.
- ICML 2008 Tutorial Session http://www.submodularity.org
- 永野清仁，河原吉伸，岡本吉央：離散凸最適化手法による機械学習の諸問題へのアプローチ，回路とシステム軽井沢ワークショップ，2009.