近似的劣モジュラ性を用いた局所探索法の近似保証

藤井 海斗 （国立情報学研究所）

OPTA 第 14 回研究会

2021/1/8

集合関数

すべての部分集合に実数値を割り当てる関数

🍝 🍢 🥦

🍕 🍱 🥪 🍜

🍫 🍔 🍩

🍌 🍞 🥐 🍙

$f: 2^N \to \mathbb{R}$

N は要素全体の集合

（台集合と呼ぶ）

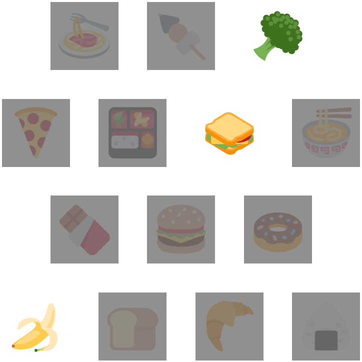
$n \triangleq |N|$

$N = \{$🍝, 🍢, 🥦, 🍕, 🍱, 🥪,
🍜, 🍫, 🍔, 🍩, 🍌, 🍞, 🥐, 🍙$\}$

集合関数

すべての部分集合に実数値を割り当てる関数

$$f(\{🥦, 🥪, 🍌\}) = 50$$



$f : 2^N \to \mathbb{R}$

N は要素全体の集合

（台集合と呼ぶ）

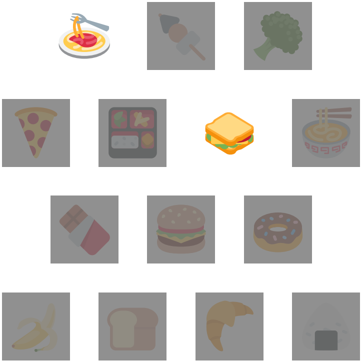
$n \triangleq |N|$

$$N = \{🍝, 🍢, 🥦, 🍕, 🍱, 🥪, 🍜, 🍫, 🍔, 🍩, 🍌, 🍞, 🥐, 🍙\}$$

集合関数

すべての部分集合に実数値を割り当てる関数

$f(\{$🍝, 🥪$\}) = 20$



$f: 2^N \to \mathbb{R}$

N は要素全体の集合

（台集合と呼ぶ）

$n \triangleq |N|$

$N = \{$🍝, 🍢, 🥦, 🍕, 🍱, 🥪,

🍜, 🍫, 🍔, 🍩, 🍌, 🍞, 🥐, 🍙$\}$

制約

実行可能な集合を集合族 $\mathcal{I} \subseteq 2^N$ で表現

例 要素数の制約

要素数 k 以下の集合はすべて実行可能

$\mathcal{I} = \{X \in 2^N \mid |X| \leq k\}$

例 予算の制約（ナップサック制約）

各要素 $v \in N$ の価格が c_v 円のとき，

合計金額が B 円以下の集合はすべて実行可能

$\mathcal{I} = \{X \in 2^N \mid \sum_{v \in X} c_v \leq B\}$

集合関数の最大化

関数値が最大となる部分集合を見つける問題

Maximize	$f(X)$	目的関数
subject to	$X \in \mathcal{I}$	制約

※ $X \subseteq N$ を入力すると $f(X)$ の値と $X \in \mathcal{I}$ かどうかを出力する
オラクルを使えると仮定

解の候補は指数的に多い

→ **全部を調べると天文学的な時間がかかる**

Q f と $\mathcal{I}$ がどんな条件を満たせば効率的に解けるか？

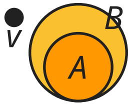
目次

劣モジュラ性

手元にある要素が増えると新しい要素による増分が減る

$f(v|A) \triangleq f(A \cup \{v\}) - f(A)$

$A \subseteq N$ に対して $v \in N$ を追加したときの**増分**



f が劣モジュラ

$\overset{\triangle}{\Longleftrightarrow}$ $A \subseteq B$ を満たす任意の $A, B \subseteq N$ と $v \in N \setminus B$ について

$f(v|A) \geq f(v|B)$ が成り立つ

$A = \{$🍌$\}, B = \{$🍌, 🍱$\}, v =$ 🍕 のとき

$f(\{$🍕, 🍌$\}) - f(\{$🍌$\}) \geq f(\{$🍕, 🍌, 🍱$\}) - f(\{$🍌, 🍱$\})$

単調性

要素を追加すると関数値が増える（減らない）

$f(v|A) \triangleq f(A \cup \{v\}) - f(A)$

$A \subseteq N$ に対して $v \in N$ を追加したときの**増分**

f が単調
$\overset{\triangle}{\Longleftrightarrow}$ 任意の $A \subseteq N$ と $v \in N$ について $f(v|A) \geq 0$ が成立

$A = \{$🍌, 🍱$\}, v =$ 🍕 のとき

$f(\{$🍕, 🍌, 🍱$\}) - f(\{$🍌, 🍱$\}) \geq 0$

劣モジュラ最大化

Maximize	$f(X)$	目的関数（劣モジュラ関数）
subject to	$X \in \mathcal{I}$	制約

各問題設定（目的関数が単調 or 非単調、さまざまな制約）に対してアルゴリズムが研究されてきた

※ 最も単純な設定

- 単調・要素数制約 $\mathcal{I} = \{X \in 2^N \mid |X| \leq k\}$
- 非単調・制約なし

でも多項式時間で最適解を求めるのは困難

近似アルゴリズム

現実には効率的に最適解を求めるのが難しい問題も多い

⟶ **近似比に保証のある多項式時間アルゴリズム**

$\alpha = 0 \longleftrightarrow \alpha = 1$

任意のアルゴリズム	最適解がつねに求まる

$\alpha \in [0, 1]$ がアルゴリズムの**近似比**（アルゴリズムは α **近似**）

$\overset{\triangle}{\Longleftrightarrow}$ 任意の問題例において $\mathbb{E}[f(X)] \geq \alpha f(X^*)$

（X アルゴリズムの解，X^* 最適解）

劣モジュラ最大化に対する主要なアプローチ

1. **貪欲法**

   劣モジュラ最大化アルゴリズムの多くが貪欲法に基づく

   連続緩和 + 丸め、部分列挙など多様な手法

2. **局所探索**

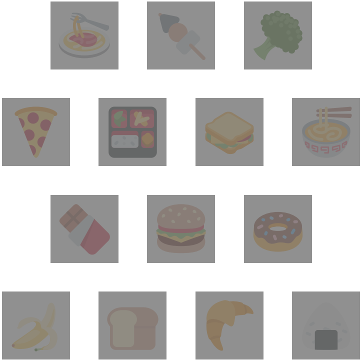
   マトロイド交叉、交換システムに対しては最良の近似比

3. その他

   ランダムサンプリングなど

単調・要素数制約に対する貪欲法 [NWF'78]

各ステップで目的関数値が最も増える要素を追加していく



$$f(\emptyset) = 0$$

貪欲法

$X_0 \leftarrow \emptyset$

for $i = 1, \cdots, k$ **do**

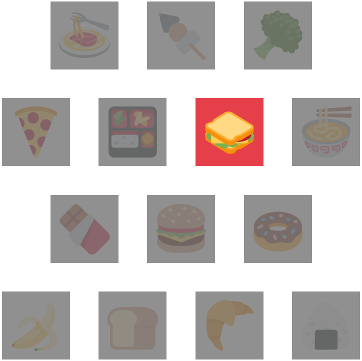
$\quad v_i \in \operatorname*{argmax}_{v \in N} f(v | X_{i-1})$

$\quad X_i \leftarrow X_{i-1} \cup \{v_i\}$

return X_k

単調・要素数制約に対する貪欲法 [NWF’78]

各ステップで目的関数値が最も増える要素を追加していく



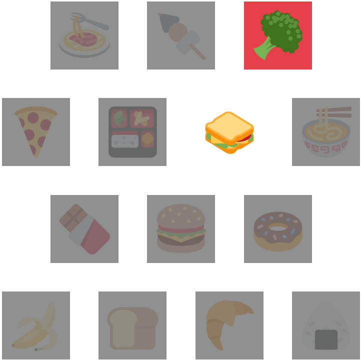
$f(\{$🥪$\}) = 50$

貪欲法

1: $X_0 \leftarrow \emptyset$
2: **for** $i = 1, \cdots, k$ **do**
3: $\quad v_i \in \operatorname{argmax}_{v \in N} f(v \mid X_{i-1})$
4: $\quad X_i \leftarrow X_{i-1} \cup \{v_i\}$
5: return X_k

単調・要素数制約に対する貪欲法 [NWF'78]

各ステップで目的関数値が最も増える要素を追加していく



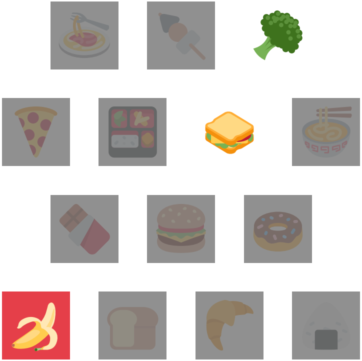
$f(\{$🥪, 🥦$\}) = 60$

貪欲法

1: $X_0 \leftarrow \emptyset$
2: **for** $i = 1, \cdots, k$ **do**
3: $\quad v_i \in \operatorname{argmax}_{v \in N} f(v | X_{i-1})$
4: $\quad X_i \leftarrow X_{i-1} \cup \{v_i\}$
5: return X_k

単調・要素数制約に対する貪欲法 [NWF'78]

各ステップで目的関数値が最も増える要素を追加していく



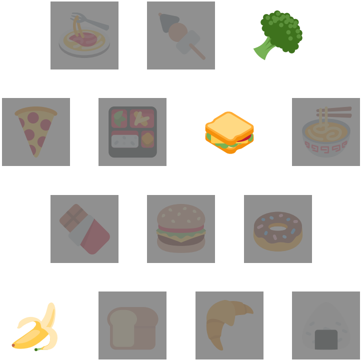
$f(\{$🥪, 🥦, 🍞$\}) = 65$

貪欲法

1: $X_0 \leftarrow \emptyset$
2: **for** $i = 1, \cdots, k$ **do**
3: $\quad v_i \in \operatorname{argmax}_{v \in N} f(v | X_{i-1})$
4: $\quad X_i \leftarrow X_{i-1} \cup \{v_i\}$
5: return X_k

単調・要素数制約に対する貪欲法 [NWF'78]

各ステップで目的関数値が最も増える要素を追加していく



$f(\{$🥪, 🥦, 🍞$\}) = 65$

貪欲法

$X_0 \leftarrow \emptyset$

for $i = 1, \cdots, k$ **do**

$\quad v_i \in \operatorname*{argmax}_{v \in N} f(v \mid X_{i-1})$

$\quad X_i \leftarrow X_{i-1} \cup \{v_i\}$

return X_k

劣モジュラ関数の性質

$f(A|B) \triangleq f(A \cup B) - f(B)$

$B \subseteq N$ に対して $A \subseteq N$ を追加したときの**増分**

補題

$f : 2^N \to \mathbb{R}$ が劣モジュラのとき

任意の集合 $A, B \subseteq N$ について $f(A|B) \leq \sum_{v \in A} f(v|B)$

証明 $A \triangleq \{a_1, \cdots, a_{|A|}\}$ 任意の順序

$$f(A|B) = \sum_{i=1}^{|A|} f(a_i | B \cup \{a_1, \cdots, a_{i-1}\}) \leq \sum_{i=1}^{|A|} f(a_i|B)$$

$$f\left(\left\{\text{🍕}, \text{🍌}\right\} \middle| \left\{\text{🍱}\right\}\right) \leq f\left(\text{🍕} \middle| \left\{\text{🍱}\right\}\right) + f\left(\text{🍌} \middle| \left\{\text{🍱}\right\}\right)$$

貪欲法の近似保証 (1/2)

定理 [Nemhauser–Wolsey–Fisher'78]

f が非負単調劣モジュラならば

$$f(X_k) \geq (1 - 1/e) \max_{X^* : |X^*| \leq k} f(X^*)$$

補題

f が劣モジュラならば　$f(v_i | X_{i-1}) \geq \frac{1}{k} \left(f(X^*) - f(X_{i-1}) \right)$

証明

$$
\begin{aligned}
f(v_i | X_{i-1}) &= \max_{v \in N} f(v | X_{i-1}) && \text{(貪欲法の性質)} \\
&\geq \frac{1}{k} \sum_{v \in X^*} f(v | X_{i-1}) \geq \frac{1}{k} f(X^* | X_{i-1}) && \text{(劣モジュラ性)} \\
&\geq \frac{1}{k} \left(f(X^*) - f(X_{i-1}) \right) && \text{(単調性)}
\end{aligned}
$$

貪欲法の近似保証 (2/2)

定理 [Nemhauser–Wolsey–Fisher'78]

f が非負単調劣モジュラならば

$$f(X_k) \geq (1 - 1/e) \max_{X^* : |X^*| \leq k} f(X^*)$$

証明 $f(X^*) - f(X_i) \leq (1 - \frac{1}{k})^i f(X^*)$ を数学的帰納法で示す

(i) $i = 0$ のとき非負性より $f(X^*) - f(\emptyset) \leq f(X^*)$

(ii) i で成り立つと仮定すると

$$f(X^*) - f(X_i) = f(X^*) - f(X_{i-1}) - f(v_i | X_{i-1})$$
$$\leq \left(1 - \frac{1}{k}\right)(f(X^*) - f(X_{i-1})) \leq \left(1 - \frac{1}{k}\right)^i f(X^*)$$

$i = k$ を代入すると $(1 - \frac{1}{k})^k \leq \frac{1}{e}$ より $f(X_k) \geq (1 - 1/e) f(X^*)$

貪欲法に関するその他の結果

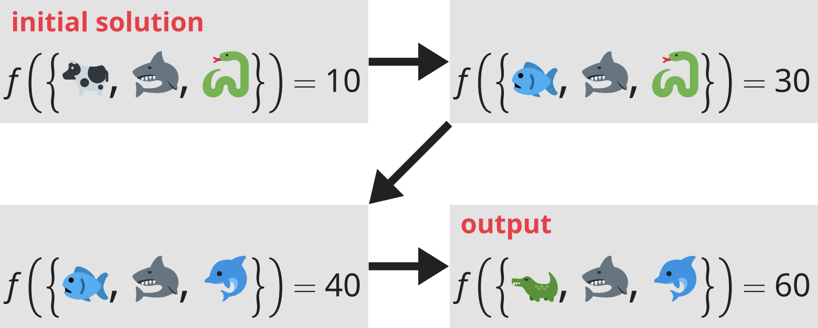
単調

- **部分列挙**によりナップサック制約に対して $(1-1/e)$ 近似 [Sviridenko'04]
- p システム制約に対して $1/(p+1)$ 近似 [Jenkyns'76, CCPV'11]
- **連続貪欲法**によりマトロイド制約に対して $(1-1/e)$ 近似 [Calinescu–Checkuri–Pál–Vondrák'11]
- …

非単調

- **乱択貪欲法**により要素数制約に対して $1/e$ 近似 [BFNS'14]
- マトロイド制約に対して 0.385 近似 [Buchbinder–Feldman'19]
- …

組合せ最適化で広く用いられるアルゴリズム設計法



初期解から始めて、解を少しずつ変更しながら関数値を改善

単調・要素数制約に対する局所探索 [NWF'78]

各ステップで目的関数値が最も増える要素を追加していく

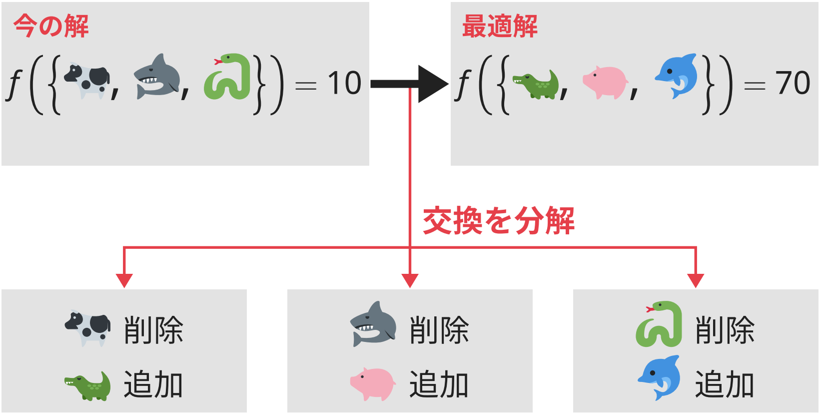
局所探索

$X \leftarrow$ 任意の要素数 k の集合
while 関数値が改善する限り **do**
  X から 1 要素だけ交換して関数値を改善
return X

※ このアルゴリズムは多項式時間で終わるとは限らない
→ 改善の幅を ϵ 以上に制限すれば多項式時間

※ 元の論文 [NWF'78] では複数要素の交換も扱われている

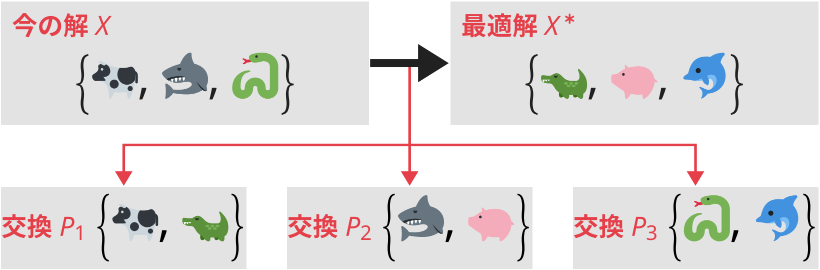
劣モジュラ関数の性質



f が非負単調劣モジュラならば、

それぞれの交換による増分の和 ≥ 最適値 − 2 × 今の値

劣モジュラ関数の性質



補題

f が非負単調劣モジュラならば、

$$\sum_{P \in \mathcal{P}} \{ f(X \triangle P) - f(X) \} \geq f(X^*) - 2f(X)$$

※ $X \triangle P \triangleq (X \setminus P) \cup (P \setminus X)$ 対称差

局所探索の近似保証 [NWF'78]

定理 [Nemhauser–Wolsey–Fisher'78]

局所探索の出力を X とすると

$$f(X) \geq 0.5 \max_{X^* : |X^*| \leq k} f(X^*)$$

証明 前頁の補題より $\sum_{P \in \mathcal{P}} \{f(X \triangle P) - f(X)\} \geq f(X^*) - 2f(X)$

X は局所探索の出力だから $f(X \triangle P) - f(X) \leq 0 \ (\forall P \in \mathcal{P})$

よって $f(X^*) - 2f(X) \leq 0$ $\quad\therefore f(X) \geq 0.5 f(X^*)$

※ 貪欲法の近似比 $1 - 1/e \approx 0.632$ より悪い

局所探索に関するその他の結果

単調

- マトロイド制約に対する $(1-1/e)$ 近似 [Filmus–Ward'14]
- p マトロイド交叉制約に対する $1/(p+\epsilon)$ 近似

  [Lee–Sviridenko–Vondrák'10]
- p 交換システム制約に対する $1/(p+\epsilon)$ 近似

  [Feldman–Naor–Schwartz–Ward'11]

非単調

- 制約なしに対する決定的 1/3 近似と乱択 2/5 近似

  [Feige–Mirrokni–Vondrák'11]
- p マトロイド交叉制約に対する $\frac{1}{p+2+1/p+\epsilon}$ 近似

  [Lee–Mirrokni–Nagarajan–Sviridenko'10]

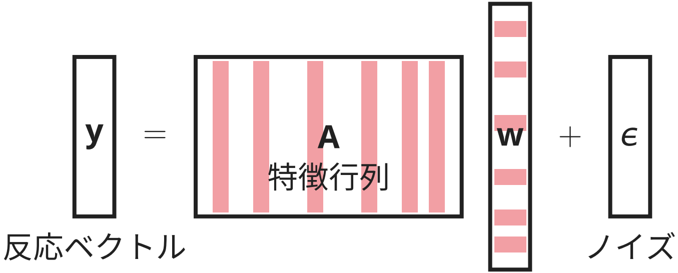
目次

線形回帰のスパースな解を求める問題

$$\text{Maximize}_{\mathbf{w}\in\mathbb{R}^n} \quad R^2 := 1 - \|\mathbf{y} - \mathbf{A}\mathbf{w}\|_2^2 \,/\, \|\mathbf{y}\|_2^2$$

$$\text{subject to} \quad \|\mathbf{w}\|_0 \leq s$$

非凸な制約



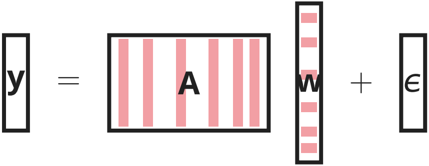
特徴選択を集合関数最大化として定式化

$\mathbf{w}$ の非ゼロ要素を選択する問題を集合関数最大化とみなす

$$\begin{aligned} &\text{Maximize}_{\mathbf{w}\in\mathbb{R}^n} && R^2 := 1 - \|\mathbf{y} - \mathbf{A}\mathbf{w}\|_2^2 \,/\, \|\mathbf{y}\|_2^2 \\ &\text{subject to} && \|\mathbf{w}\|_0 \le s \end{aligned}$$

↓ $f_{R^2}(X) \triangleq 1 - \min_{\mathbf{w}:\ \mathrm{supp}(\mathbf{w}) \subseteq X} \|\mathbf{y} - \mathbf{A}\mathbf{w}\|_2^2 \,/\, \|\mathbf{y}\|_2^2$ を導入

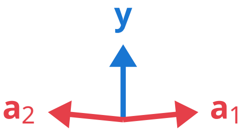
$$\text{Maximize}_{X \subseteq N \triangleq \{1,\cdots,n\}} \; f_{R^2}(X) \quad \text{subject to } |X| \le s$$



f_{R^2} は劣モジュラ関数ではない

$$f_{R^2}(X) \triangleq 1 - \min_{\mathbf{w}:\, \mathrm{supp}(\mathbf{w}) \subseteq X} \|\mathbf{y} - \mathbf{A}\mathbf{w}\|_2^2 \,/\, \|\mathbf{y}\|_2^2$$

$$\mathbf{y} := \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \mathbf{A} := \begin{bmatrix} 1 & -1 \\ 0.1 & 0.1 \end{bmatrix}$$



$$f(\emptyset) = 0$$

$$f(\{1\}) = f(\{2\}) \approx 0.0099$$

$$f(\{1, 2\}) = 1$$

$$f(1|\emptyset) \approx 0.0099$$

$$\not\geq f(1|\{2\}) \approx 0.99$$

劣モジュラ関数でない

近似的劣モジュラ性

目的関数が劣モジュラ性をもたなくても、

劣モジュラ性に近い性質（近似的劣モジュラ性）があれば

劣モジュラ最大化の技術を適用できるのではないか？

⬇

近似的劣モジュラ性を定義することで、

劣モジュラ性のない問題に対して近似アルゴリズムを設計

単調な集合関数の劣モジュラ関数への近さを表す

$\gamma = 0$ ⟷ $\gamma = 1$

劣モジュラ関数から遠い	劣モジュラ関数に近い

$\gamma_{U,k} \in [0,1]$ が単調な集合関数 f の

（集合 U と自然数 k に関する）**劣モジュラ比**

$$\overset{\triangle}{\Leftrightarrow} \gamma_{U,k}\, f(S|T) \le \sum_{v \in S} f(v|T) \quad (\forall T \subseteq U, S \subseteq N \text{ s.t. } |S| \le k)$$

劣モジュラ関数の性質（再掲）

$f(A|B) \stackrel{\triangle}{=} f(A \cup B) - f(B)$

$B \subseteq N$ に対して $A \subseteq N$ を追加したときの**増分**

補題

$f : 2^N \to \mathbb{R}$ が劣モジュラのとき
任意の集合 $A, B \subseteq N$ について $f(A|B) \leq \sum_{v \in A} f(v|B)$

証明 $A \stackrel{\triangle}{=} \{a_1, \cdots, a_{|A|}\}$ 任意の順序

$$f(A|B) = \sum_{i=1}^{|A|} f(a_i | B \cup \{a_1, \cdots, a_{i-1}\}) \leq \sum_{i=1}^{|A|} f(a_i|B)$$

$$f\left(\{🍕, 🍌\} \middle| \{🍱\}\right) \leq f\left(🍕 \middle| \{🍱\}\right) + f\left(🍌 \middle| \{🍱\}\right)$$

貪欲法の近似保証に用いた劣モジュラ関数の性質を

どれくらい緩めれば成り立つかを表す

補題

$f: 2^N \to \mathbb{R}$ が劣モジュラのとき

任意の集合 $A, B \subseteq N$ について $f(A|B) \leq \sum_{v \in A} f(v|B)$

↓ 不等式を緩める

$\gamma_{U,k} \in [0, 1]$ が単調な集合関数 f の

（集合 U と自然数 k に関する）**劣モジュラ比**

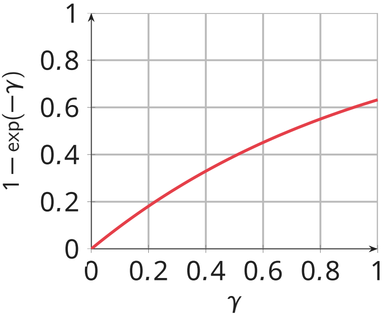
$$\overset{\triangle}{\Leftrightarrow} \gamma_{U,k} f(S|T) \leq \sum_{v \in S} f(v|T) \quad (\forall T \subseteq U, S \subseteq N \text{ s.t. } |S| \leq k)$$

劣モジュラ比による貪欲法の近似保証

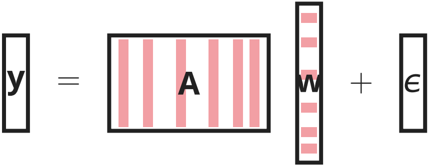
定理 [Das–Kempe'11]

単調な集合関数の劣モジュラ比が $\gamma_{U,k}$ のとき、

貪欲法の解 X は最適解の $(1-\exp(-\gamma_{X,k}))$ 近似



$$f_{R^2}(X) \triangleq 1 - \min_{\mathbf{w}\colon \mathrm{supp}(\mathbf{w}) \subseteq X} \|\mathbf{y} - \mathbf{A}\mathbf{w}\|_2^2 \,/\, \|\mathbf{y}\|_2^2$$



定理 [Das–Kempe’11]

$\mathbf{A}$ の各列が正規化されていると仮定すると

f_{R^2} の（U と k に関する）劣モジュラ比 $\gamma_{U,k}$ は次を満たす

$$\gamma_{U,k} \geq \min_{Z \subseteq [n]\colon |Z| \leq k} \lambda_{\min}(\mathbf{A}_Z^\top \mathbf{A}_Z)$$

→ 貪欲法は $1 - \exp(-\min_{Z \subseteq [n]\colon |Z| \leq k} \lambda_{\min}(\mathbf{A}_Z^\top \mathbf{A}_Z))$ 近似

R^2 を任意の制限強凹／平滑関数に一般化

$$\begin{array}{ll} \text{Maximize}_{\mathbf{w}\in\mathbb{R}^n} & u(\mathbf{w}) \quad u(\mathbf{0}) \geq 0 \\ \text{subject to} & \|\mathbf{w}\|_0 \leq s \end{array}$$

R^2 の場合

$$u(\mathbf{w}) = 1 - \|\mathbf{y} - \mathbf{A}\mathbf{w}\|_2^2 \,/\, \|\mathbf{y}\|_2^2$$

制限強凹／平滑定数が $\mathbf{A}^\top\mathbf{A}$ の部分行列の最小固有値に対応

ほかの例

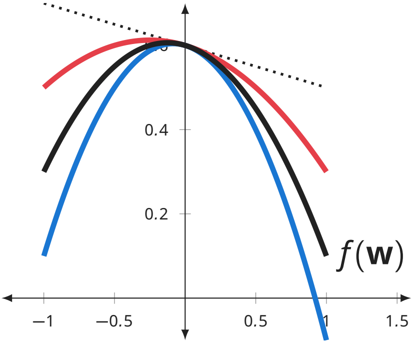
- ロジスティック回帰
- 一般化線形モデル

u は Ω において制限強凹 w.r.t. パラメータ m_Ω

$$\Leftrightarrow \forall \mathbf{x}, \mathbf{y} \in \Omega,\ u(\mathbf{y}) - u(\mathbf{x}) - \langle \nabla u(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \leq -\frac{m_\Omega}{2} \|\mathbf{y} - \mathbf{x}\|_2^2$$

u は Ω において制限平滑 w.r.t. パラメータ M_Ω

$$\Leftrightarrow \forall \mathbf{x}, \mathbf{y} \in \Omega,\ u(\mathbf{y}) - u(\mathbf{x}) - \langle \nabla u(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \geq -\frac{M_\Omega}{2} \|\mathbf{y} - \mathbf{x}\|_2^2$$



$\Omega_s = \{\mathbf{x}, \mathbf{y} \mid \|\mathbf{x}\|_0 \leq s,\ \|\mathbf{y}\|_0 \leq s, \|\mathbf{x} - \mathbf{y}\|_0 \leq s\}$ における制限強凹定数を m_s

$\Omega_{s,t} = \{\mathbf{x}, \mathbf{y} \mid \|\mathbf{x}\|_0 \leq s,\ \|\mathbf{y}\|_0 \leq s, \|\mathbf{x} - \mathbf{y}\|_0 \leq t\}$ における制限平滑定数を $M_{s,t}$ とする

解 w のサポート（非ゼロ要素の集合）を選ぶ問題とみなす

$$\text{Maximize}_{\mathbf{w}\in\mathbb{R}^n}\ u(\mathbf{w}) \quad \text{subject to } \|\mathbf{w}\|_0 \le s$$

↓ 集合関数 $f_u(X) \triangleq \max_{\mathbf{w}:\ \mathrm{supp}(\mathbf{w})\subseteq X} u(\mathbf{w})$ を導入

$$\text{Maximize}_{X\subseteq N\triangleq[n]}\ f_u(X) \quad \text{subject to } |X| \le s$$

定理 [Das–Kempe'11]

f_u の（集合 U と自然数 k に関する）劣モジュラ比 $\gamma_{U,k}$ は

$\gamma_{U,k} \ge m_{|U|+s}/M_{|U|+1,1}$ を満たす

→ 貪欲法は $1 - \exp(-m_{2s}/M_{s+1,1})$ 近似

劣モジュラ比に関するその他の結果

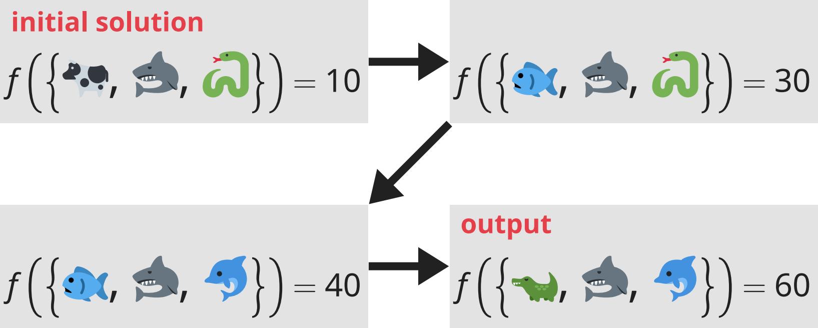
劣モジュラ最大化のアイデアを特徴選択に利用

- 分散アルゴリズム [KannaEDNG'17]
- 高速な乱択アルゴリズム [KannaEDNG'17]
- ストリーミングアルゴリズム [Elenberg–Dimakis–Feldman–Karbasi'17]
- 二段階最適化（辞書選択） [Cevher–Krause'11, Fujii–Soma'18]
- 適応的最適化 [Fujii–Sakaue'19]

目次

局所探索（再掲）

組合せ最適化で広く用いられるアルゴリズム設計法



初期解から始めて、解を少しずつ変更しながら関数値を改善

本研究の動機

貪欲法に用いられる性質を緩めることで

劣モジュラ比が定義された

⬇

局所探索に用いられる性質を緩めれば

劣モジュラ性のない問題に対して

局所探索の近似保証ができるのでは？

本研究の概要

localizability という集合関数の性質を提案

1. 集合関数が localizability を満たすなら
   局所探索がよい近似解を出力する
2. スパース最適化の目的関数は localizability を満たす
3. スパース最適化に対する局所探索は高速化できる

簡単な設定：要素数制約

要素数 s 以下の最適な部分集合を探す問題

$$\text{Maximize} \quad f(X)$$

$$\text{subject to} \quad |X| \leq s$$

簡単な設定：要素数制約

要素数 s 以下の最適な部分集合を探す問題

$$\text{Maximize} \quad f(X)$$
$$\text{subject to} \quad |X| \le s$$

$f: 2^N \to \mathbb{R}_{\ge 0}$

f の単調性を仮定

（劣モジュラとは限らない）

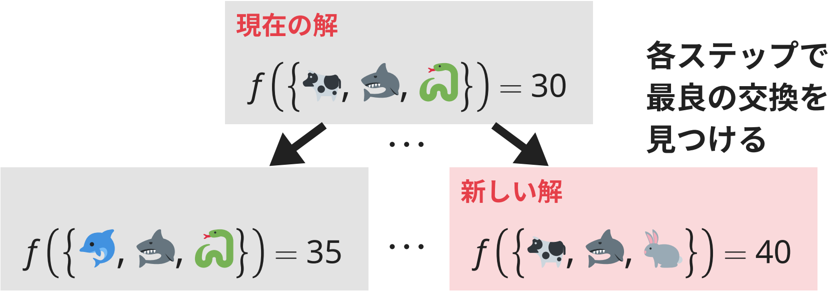
要素数制約のための局所探索

X を任意の極大な実行可能解とする

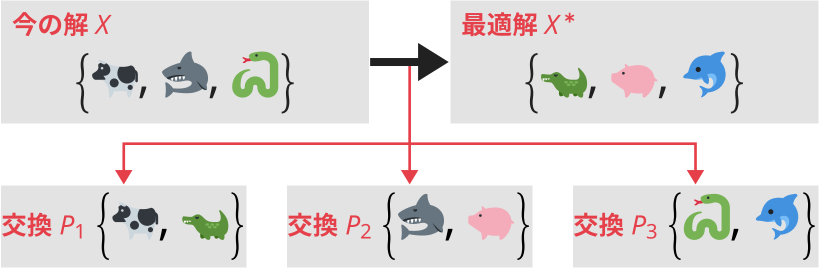
For $i = 1, \cdots, T$:

　　各 $a \in N \setminus X,\ b \in X$ について $f(X \setminus \{b\} \cup \{a\})$ を計算

　　この値を最大化する a と b で X を更新



劣モジュラ関数の性質（再掲）



補題

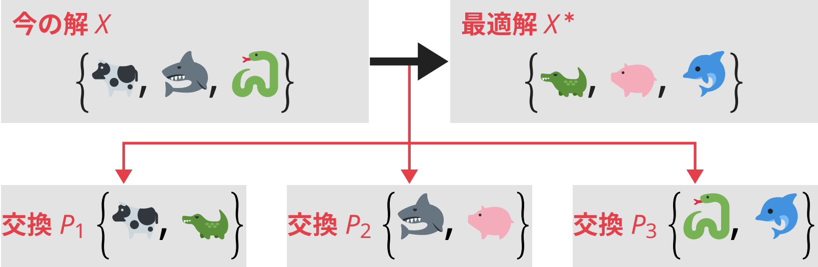
f が非負単調劣モジュラならば、

$$\sum_{P \in \mathcal{P}} \{f(X \triangle P) - f(X)\} \geq f(X^*) - 2f(X)$$

※ $X \triangle P \overset{\triangle}{=} (X \setminus P) \cup (P \setminus X)$ 対称差

Localizability (simple version)

局所探索の近似保証に用いられる性質を緩めたもの



f is (α, β)-localizable

$$\overset{\triangle}{\Longleftrightarrow} \quad \sum_{P \in \mathcal{P}} \{ f(X \triangle P) - f(X) \} \geq \alpha f(X^*) - \beta f(X)$$

$(\forall X, X^* \subseteq N$ of size at most s, $\forall \mathcal{P}$ partition of $X \triangle X^*)$

要素数制約に対する近似保証

定理

目的関数が (α, β)-localizable なら

局所探索法は $\frac{\alpha}{\beta}\left(1 - \exp\left(-\frac{\beta T}{s}\right)\right)$ 近似

T 反復回数、s 実行可能解の要素数の最大値

要素数制約に対する近似保証

定理

目的関数が (α, β)-localizable なら

局所探索法は $\frac{\alpha}{\beta}\left(1 - \exp\left(-\frac{\beta T}{s}\right)\right)$ 近似

T 反復回数、s 実行可能解の要素数の最大値

よく現れる関数に対する近似保証

	(α, β)	近似比
線形関数	$(1, 1)$	$(1 - \exp(-T/s))$
劣モジュラ関数	$(1, 2)$	$\frac{1}{2}(1 - \exp(-T/s))$

一般的な設定：構造的な制約

Maximize $f(X)$

subject to $X \in \mathcal{I}$

$\mathcal{I} \subseteq 2^N$ は

実行可能な部分集合の族

さまざまな制約のクラス

p 交換システム $\not\subseteq$ $\not\supseteq$ p マトロイド交叉

$\cup$

マトロイド

$\cup$

$\cup$

要素数制約

$\mathcal{I} = \{X \subseteq N \mid |X| \leq s\}$

マトロイド

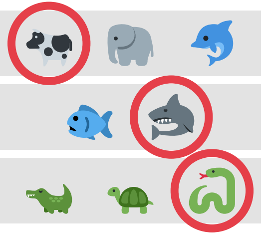
N 有限集合、$\mathcal{I} \subseteq 2^N$ s.t. $\emptyset \in \mathcal{I}$ かつ "$A \subseteq B \in \mathcal{I} \Rightarrow A \in \mathcal{I}$"

$\mathcal{M} = (N, \mathcal{I})$ は**マトロイド**

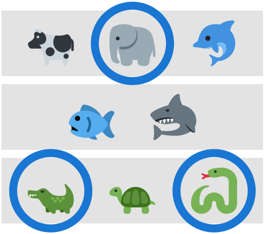
$\Updownarrow$

$\forall A, B \in \mathcal{I},\ |A| < |B| \Rightarrow \exists i \in B \setminus A,\ A \cup \{i\} \in \mathcal{I}$

例：分割マトロイド



独立



非独立

N 有限集合、$\mathcal{I} \subseteq 2^N$ s.t. $\emptyset \in \mathcal{I}$ かつ "$A \subseteq B \in \mathcal{I} \Rightarrow A \in \mathcal{I}$"

$\mathcal{M} = (N, \mathcal{I})$ は**マトロイド**

$\triangleleft\Updownarrow$

$\forall A, B \in \mathcal{I},\ |A| < |B| \Rightarrow \exists i \in B \setminus A,\ A \cup \{i\} \in \mathcal{I}$

$(N, \mathcal{I})$ は **p マトロイド交叉**

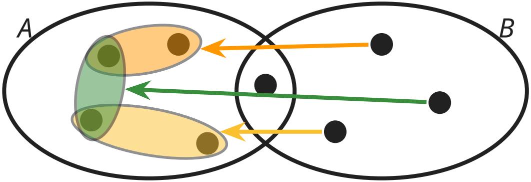
$\triangleleft\Updownarrow$

$\exists (N, \mathcal{I}_1), \cdots, (N, \mathcal{I}_p)$ マトロイド s.t. $\mathcal{I} = \bigcap_{i=1}^{p} \mathcal{I}_i$

N 有限集合、$\mathcal{I} \subseteq 2^N$ s.t. $\emptyset \in \mathcal{I}$ かつ "$A \subseteq B \in \mathcal{I} \Rightarrow A \in \mathcal{I}$"

$(N, \mathcal{I})$ は **p 交換システム**

◁⇕



$\forall A, B \in \mathcal{I}$

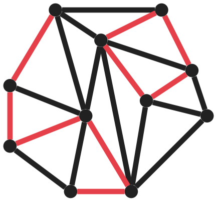
$\exists \phi\colon B \setminus A \to 2^{A \setminus B}$ s.t.

- 任意の $B' \subseteq B \setminus A$ に対して $(A \setminus (\bigcup_{b \in B'} \phi(b))) \cup B' \in \mathcal{I}$
- $|\phi(b)| \leq p$ $(\forall b \in B \setminus A)$
- 各 $a \in A \setminus B$ は $(\phi(b))_{b \in B \setminus A}$ に高々 p 回出現

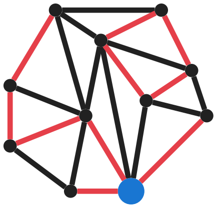
N 有限集合、$\mathcal{I} \subseteq 2^N$ s.t. $\emptyset \in \mathcal{I}$ かつ "$A \subseteq B \in \mathcal{I} \Rightarrow A \in \mathcal{I}$"

$(N, \mathcal{I})$ は **p 交換システム**

例：b マッチング



2 マッチング



2 マッチングでない

$\mathcal{I} = \{X \subseteq E \mid \forall v \in V,\ \deg_v(X) \leq b\}$ は 2 交換システム

構造的制約のための局所探索

以下の操作で得られる実行可能解全体：

q 個以下の要素の追加と
（p マトロイド交叉）$2pq$ 要素以下の削除
（p 交換システム）$pq - q + 1$ 要素以下の削除

Let $X \leftarrow \emptyset$

For $i = 1, \cdots, T$:

各 $X' \in \mathcal{F}_q(X)$について $f(X')$ を計算

そのなかで関数値を最大にする交換で X を更新

今の解

$f(\{$🐄, 🦈, 🐍$\}) = 30$

$\cdots$

$f(\{$🐬, 🦈, 🐍$\}) = 35$

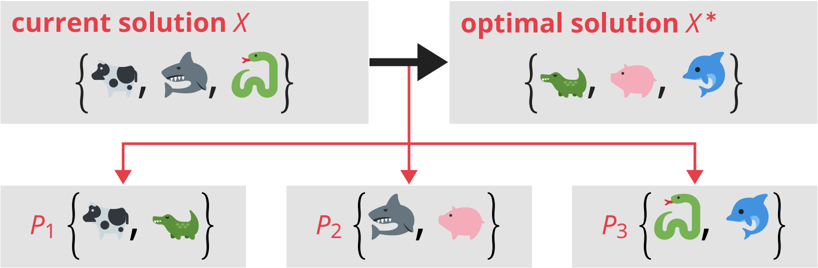
$\cdots$

次の解

$f(\{$🐄, 🦈, 🐇$\}) = 40$

各ステップで
$\mathcal{F}_q(X)$ のなかで最良の
交換を見つける

Localizability (general version)



$\mathcal{P} \subseteq 2^N$ は多重集合 s.t. $\begin{cases} \text{各 } X^* \setminus X \text{ は } k \text{ 回出現} \\ \text{各 } X \setminus X^* \text{ は } \ell \text{ 回出現} \end{cases}$

f is $(\alpha, \beta_1, \beta_2)$-localizable

$$\overset{\triangle}{\Longleftrightarrow} \quad \sum_{P \in \mathcal{P}} \{ f(X \triangle P) - f(X) \} \ \geq \ \alpha k \, f(X^*) - (\beta_1 \ell + \beta_2 k) \, f(X)$$

($\forall X, X^* \subseteq N$ of size at most s, 上記の条件を満たす任意の $\mathcal{P}$)

近似保証

定理

目的関数が $(\alpha, \beta_1, \beta_2)$-localizable なら

マトロイド 局所探索法は

$$\frac{\alpha}{\beta_1 + \beta_2}\left(1 - \exp\left(-\frac{(\beta_1 + \beta_2)T}{s}\right)\right) \text{ 近似}$$

p-MI/p-ES パラメタ $q \in \mathbb{Z}_{>0}$ の局所探索法は

$$\frac{\alpha\left(1 - \exp\left(\frac{(\beta_1(p-1+1/q)+\beta_2)T}{s}\right)\right)}{\beta_1(p - 1 + 1/q) + \beta_2} \text{ 近似}$$

q は選べるパラメタ

（各ステップで $n^{O(q)}$ 個の集合を確認しなければならない）

解のサポートに構造的な制約がある連続最適化問題

$$\begin{array}{ll} \text{Maximize}_{\mathbf{w}\in\mathbb{R}^n} & u(\mathbf{w}) \\ \text{subject to} & \mathrm{supp}(\mathbf{w}) \in \mathcal{I} \end{array}$$

仮定

- $u(\mathbf{0}) \geq 0$
- $\Omega_s = \{\mathbf{x}, \mathbf{y} \mid \|\mathbf{x}\|_0 \leq s,\ \|\mathbf{y}\|_0 \leq s, \|\mathbf{x}-\mathbf{y}\|_0 \leq s\}$ において制限強凹 w.r.t. m_s
- $\Omega_{s,t} = \{\mathbf{x}, \mathbf{y} \mid \|\mathbf{x}\|_0 \leq s,\ \|\mathbf{y}\|_0 \leq s, \|\mathbf{x}-\mathbf{y}\|_0 \leq t\}$ において制限平滑 w.r.t. $M_{s,t}$

構造的制約のあるスパース最適化

解のサポートに構造的な制約がある連続最適化問題

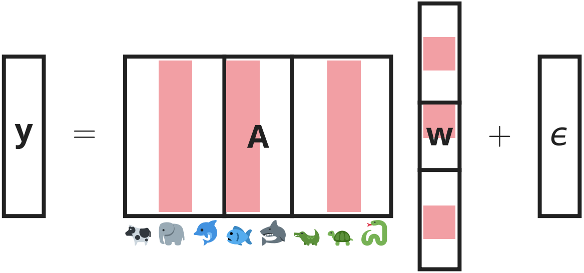
$$\begin{aligned}\text{Maximize}_{\mathbf{w}\in\mathbb{R}^n} \quad & u(\mathbf{w}) \\ \text{subject to} \quad & \mathrm{supp}(\mathbf{w}) \in \mathcal{I}\end{aligned}$$

構造的な制約

仮定

- $u(\mathbf{0}) \geq 0$
- $\Omega_s = \{\mathbf{x}, \mathbf{y} \mid \|\mathbf{x}\|_0 \leq s,\ \|\mathbf{y}\|_0 \leq s, \|\mathbf{x} - \mathbf{y}\|_0 \leq s\}$ において制限強凹 w.r.t. m_s
- $\Omega_{s,t} = \{\mathbf{x}, \mathbf{y} \mid \|\mathbf{x}\|_0 \leq s,\ \|\mathbf{y}\|_0 \leq s, \|\mathbf{x} - \mathbf{y}\|_0 \leq t\}$ において制限平滑 w.r.t. $M_{s,t}$

各分割から一つずつ特徴を選ぶ問題などを含む



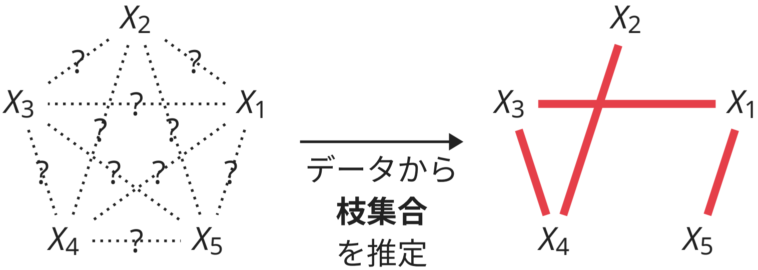
$$
\begin{aligned}
\text{Maximize} \quad & u_{R^2}(\mathbf{w}) \triangleq 1 - \frac{\|\mathbf{y} - \mathbf{A}\mathbf{w}\|_2^2}{\|\mathbf{y}\|_2^2} \\
\text{subject to} \quad & \operatorname{supp}(\mathbf{w}) \in \mathcal{I}
\end{aligned}
$$

グラフィカルモデルの構造推定

データ $\{\mathbf{x}^1, \cdots, \mathbf{x}^m\} \sim p(\mathbf{x}|\mathbf{w})$ から

$$p(\mathbf{x}|\mathbf{w}) \propto \exp\left(\sum_{(u,v)\in E} w_{uv} x_u x_v + \sum_{u\in V} w_u x_u\right)$$

スパースなイジングモデルの supp($\mathbf{w}$) を推定



各点の次数の制約は 2 交換システム制約

スパース最適化 → 集合関数最大化

supp($\mathbf{w}$) を選択する問題を集合関数最大化とみなす

$$\text{Maximize}_{\mathbf{w}\in\mathbb{R}^n} \quad u(\mathbf{w}) \quad \text{subject to} \quad \text{supp}(\mathbf{w}) \in \mathcal{I}$$

$$f_u(X) \triangleq \max_{\mathbf{w}:\, \text{supp}(\mathbf{w})\subseteq X} u(\mathbf{w})$$ を導入

$$\text{Maximize}_{X\subseteq N} f_u(X) \quad \text{subject to } X \in \mathcal{I}$$

各交換 $P \in \mathcal{P}$ が $|P| \le t$ を満たす $\mathcal{P}$

定理

f_u is $\left(\frac{m_{2s}}{M_{s,t}}, \frac{M_{s,t}}{m_{2s}}, 0\right)$-localizable with size s and exchange size t

制約	局所探索法	貪欲法
要素数制約	$\frac{m_{2s}^2}{M_{s,2}^2}(1-\epsilon_1(T))$	$1-\exp\left(-\frac{m_{2s}}{M_{s+1,1}}\right)$ †
マトロイド	$\frac{m_{2s}^2}{M_{s,2}^2}(1-\epsilon_1(T))$	$\frac{1}{(1+\frac{M_{s+1,1}}{m_s})^2}$ ‡
p-MI/p-ES	$\frac{1}{p-1+1/q}\frac{m_{2s}^2}{M_{s,t}^2}(1-\epsilon_2(T))$	N/A

$\epsilon_1(T)$ と $\epsilon_2(T)$ は $T\to\infty$ のとき 0 に収束

† [Elenberg–Khanna–Dimakis–Negahban'18]

‡ [Chen–Feldman–Karbasi'18]

高速化

$\operatorname*{argmax}_{a \in N \setminus X, b \in X} f_u(X \setminus \{a\} \cup \{b\})$ の計算には時間がかかる

➡ **f_u の値を近似的に計算することで高速化**

$\operatorname{argmax}_{a \in N \setminus X, b \in X} f_u(X \setminus \{a\} \cup \{b\})$ の計算には時間がかかる

➡ **f_u の値を近似的に計算することで高速化**

$\mathbf{w}^{(X)} \in \operatorname{argmax}_{\mathbf{w}:\, \mathrm{supp}(\mathbf{w}) \subseteq X} u(\mathbf{w})$

semi-oblivious

$\operatorname{argmax}_{a \in N \setminus X} f_u(X \setminus \{b\} \cup \{a\})$, where $b \in \operatorname{argmin}_{b \in X} (\mathbf{w}^{(X)})_b^2$

削除する要素を高速に決定

高速化

$\underset{a \in N \setminus X, b \in X}{\operatorname{argmax}} f_u(X \setminus \{a\} \cup \{b\})$ の計算には時間がかかる

➡ f_u **の値を近似的に計算することで高速化**

$$\mathbf{w}^{(X)} \in \underset{\mathbf{w}:\, \mathrm{supp}(\mathbf{w}) \subseteq X}{\operatorname{argmax}} u(\mathbf{w})$$

semi-oblivious

$$\underset{a \in N \setminus X}{\operatorname{argmax}} f_u(X \setminus \{b\} \cup \{a\}), \text{ where } b \in \underset{b \in X}{\operatorname{argmin}} (\mathbf{w}^{(X)})_b^2$$

削除する要素を高速に決定

non-oblivious

$$\underset{a \in N \setminus X, b \in X}{\operatorname{argmax}} \left\{ \frac{1}{2M_{s,2}} \left(\nabla u(\mathbf{w}^{(X)}) \right)_a^2 - \frac{M_{s,2}}{2} \left(\mathbf{w}^{(X)} \right)_b^2 \right\}$$

実験設定：特徴選択

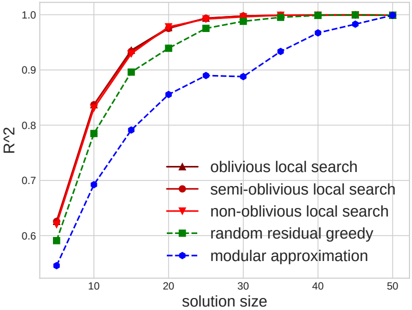
データセット

- $\mathbf{A}_{ij} \sim \mathrm{Unif}([0,1])$ の各列を正規化
- 特徴全体をランダムに s グループに分割し
  各グループから 1 つずつランダムに選んで S^* とする
- $\mathbf{y} = \mathbf{A}_{S^*}\mathbf{w} + \epsilon$（ただし $\mathbf{w}_i \sim \mathcal{N}(0,1)$, $\epsilon_i \sim \mathcal{N}(0, 0.2)$）

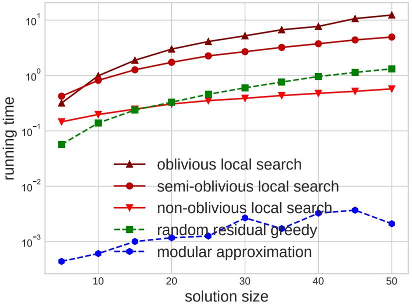
比較手法

- **Residual random greedy** [Chen–Feldman–Karbasi'18]
- **Modular approximation**

実験結果：特徴選択



R^2（近似比）



計算時間

- 提案手法は比較手法よりよい R^2 を達成
- 高速化により計算時間を 1/10 以下に短縮

実験設定：グラフィカルモデル

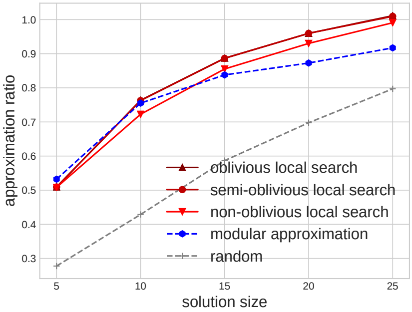
データセット

- 頂点数 10、次数 5 のランダムレギュラーグラフを真のグラフ構造とする
- イジングモデルの各枝 uv のパラメタ w_{uv} は $\{\pm 0.5\}$ からランダムに選ぶ
- Gibbs サンプリングで生成した 100 点を入力

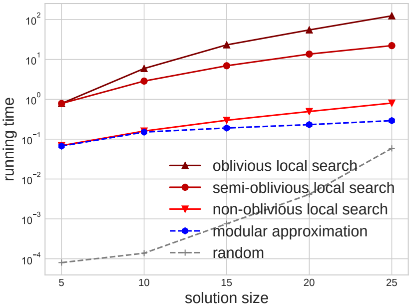
比較手法 **Modular approximation**

目的関数を線形近似して最大おもみマッチング

実験結果：グラフィカルモデル



近似比



計算時間

- 提案手法は比較手法よりよい近似比を達成
- 高速化により計算時間を 1/10 以下に短縮

本研究のまとめ

localizability という集合関数の性質を提案

1. 集合関数が localizability を満たすなら
   局所探索がよい近似解を出力する
2. スパース最適化の目的関数は localizability を満たす
3. スパース最適化に対する局所探索は高速化できる

参考文献 (1)

- Niv Buchbinder, Moran Feldman: Constrained Submodular Maximization via a Nonsymmetric Technique. *Math. Oper. Res.* 44(3): 988-1005 (2019).
- Niv Buchbinder, Moran Feldman, Joseph Naor, Roy Schwartz: Submodular Maximization with Cardinality Constraints. SODA 2014: 1433-1452.
- Gruia Calinescu, Chandra Chekuri, Martin Pál, Jan Vondrák: Maximizing a Monotone Submodular Function Subject to a Matroid Constraint. *SIAM J. Comput.* 40(6): 1740-1766 (2011).
- Volkan Cevher, Andreas Krause: Greedy Dictionary Selection for Sparse Representation. *IEEE J. Sel. Top. Signal Process.* 5(5): 979-988 (2011).
- Lin Chen, Moran Feldman, Amin Karbasi: Weakly Submodular Maximization Beyond Cardinality Constraints: Does Randomization Help Greedy? ICML 2018: 803-812.

参考文献 (2)

- Abhimanyu Das, David Kempe: Approximate Submodularity and its Applications: Subset Selection, Sparse Approximation and Dictionary Selection. *J. Mach. Learn. Res.* 19: 3:1-3:34 (2018).
- Ethan R. Elenberg, Alexandros G. Dimakis, Moran Feldman, Amin Karbasi: Streaming Weak Submodularity: Interpreting Neural Networks on the Fly. NIPS 2017: 4044-4054.
- Ethan R. Elenberg, Rajiv Khanna, Alexandros G. Dimakis, and Sahand Negahban: Restricted strong convexity implies weak submodularity. *Ann. Statist.* 46(6B): 3539-3568 (2018).
- Uriel Feige, Vahab S. Mirrokni, Jan Vondrák: Maximizing Non-monotone Submodular Functions. *SIAM J. Comput.* 40(4): 1133-1153 (2011).
- Moran Feldman, Joseph Naor, Roy Schwartz, Justin Ward: Improved Approximations for k-Exchange Systems - (Extended Abstract). ESA 2011: 784-798.

参考文献 (3)

- Kaito Fujii, Shinsaku Sakaue: Beyond Adaptive Submodularity: Approximation Guarantees of Greedy Policy with Adaptive Submodularity Ratio. ICML 2019: 2042-2051
- Kaito Fujii, Tasuku Soma: Fast greedy algorithms for dictionary selection with generalized sparsity constraints. NeurIPS 2018: 4749-4758.
- T. A. Jenkyns, The efficacy of the “greedy” algorithm, in *Proceedings of the 7th Southeastern Conference on Combinatorics, Graph Theory and Computing*: 341 – 350 (1976).
- Rajiv Khanna, Ethan R. Elenberg, Alexandros G. Dimakis, Sahand N. Negahban, Joydeep Ghosh: Scalable Greedy Feature Selection via Weak Submodularity. AISTATS 2017: 1560-1568
- Jon Lee, Vahab S. Mirrokni, Viswanath Nagarajan, Maxim Sviridenko: Maximizing Nonmonotone Submodular Functions under Matroid or Knapsack Constraints. *SIAM J. Discret. Math.* 23(4): 2053-2078 (2010).

参考文献 (4)

- Jon Lee, Maxim Sviridenko, Jan Vondrák: Submodular Maximization over Multiple Matroids via Generalized Exchange Properties. Math. Oper. Res. 35(4): 795-806 (2010).
- Sahand N. Negahban, Pradeep Ravikumar, Martin J. Wainwright, and Bin Yu: A Unified Framework for High-Dimensional Analysis of *M*-Estimators with Decomposable Regularizers. *Statist. Sci.* 27(4): 538-557 (2012).
- George L. Nemhauser, Laurence A. Wolsey, Marshall L. Fisher: An analysis of approximations for maximizing submodular set functions - I. *Math. Program.* 14(1): 265-294 (1978).
- Maxim Sviridenko: A note on maximizing a submodular set function subject to a knapsack constraint. *Oper. Res. Lett.* 32(1): 41-43 (2004).

-