辞書選択のための高速な貪欲アルゴリズム

藤井海斗（東京大学 D2）

（**相馬輔**（東京大学）との共同研究）

OPTA つくば合宿

2018/6/10

目次

辞書とは

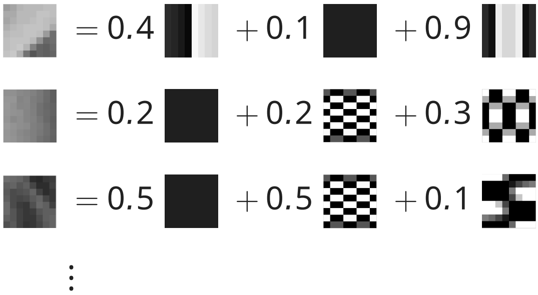
辞書とは

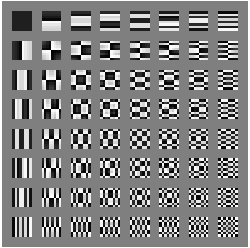
辞書

現実の信号が少数のパターンでできているなら、

“よい” 辞書によって信号のスパース表現が得られる

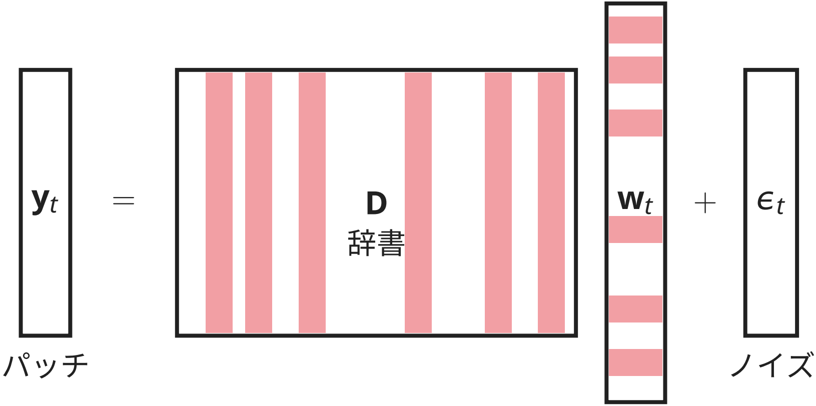
パッチ　　**辞書の元（アトム）のスパース結合**　　**辞書**





辞書

ほとんどのパッチは辞書の元のスパース結合で表現できる



応用：画像復元

画像の失われた画素を復元する問題

入力

出力

辞書作成への３つのアプローチ

既存の辞書を使う

信号処理の研究で考案された
既存の辞書を使う

欠点 データに合わせられない

辞書学習

与えられたデータから
辞書を学習する

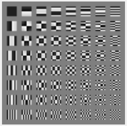
欠点 難しい非凸最適化

辞書選択 [Krause–Cevher’10]

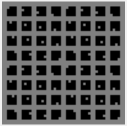
既存の辞書の和集合から
要素を選択して辞書を作る

辞書選択 [Krause–Cevher'10]

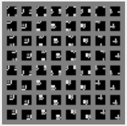
既存の辞書の和集合 V



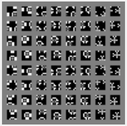
DCT 基底



Haar 基底



Db4 基底



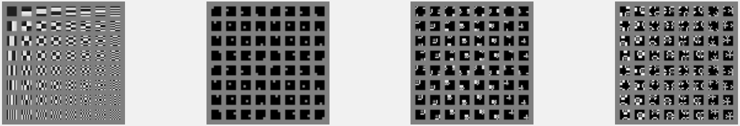
coiflet 基底

選択された辞書 $X \subseteq V$ s.t. $|X| \leq k$

パッチ $\mathbf{y}_t$ のためのアトム集合 $Z_t \subseteq X$ s.t. $|Z_t| \leq s$

辞書選択 [Krause–Cevher’10]

既存の辞書の和集合 V

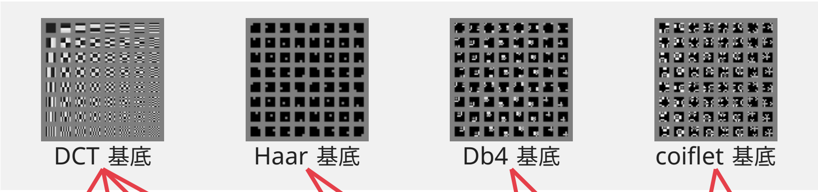


DCT 基底 Haar 基底 Db4 基底 coiflet 基底

選択された辞書 $X \subseteq V$ s.t. $|X| \leq k$

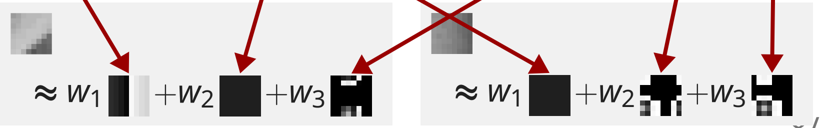
パッチ $\mathbf{y}_t$ のためのアトム集合 $Z_t \subseteq X$ s.t. $|Z_t| \leq s$

既存の辞書の和集合 V



選択された辞書 $X \subseteq V$ s.t. $|X| \leq k$

パッチ $\mathbf{y}_t$ のためのアトム集合 $Z_t \subseteq X$ s.t. $|Z_t| \leq s$



辞書選択 [Krause–Cevher'10]

アトム集合 Z_t を横に並べた行列

$$f_t(Z_t) \triangleq \|\mathbf{y}_t\|_2^2 - \min_{\mathbf{w}} \|\mathbf{y}_t - \mathbf{D}_{Z_t}\mathbf{w}\|_2^2$$

アトムの集合 Z_t によって $\mathbf{y}_t$ をどれだけうまく表現できるか

$$\text{Minimize}_{X \subseteq V} \quad \sum_{t=1}^{T} \max_{Z_t \subseteq X:\ |Z_t| \le s} f_t(Z_t)$$

$$\text{subject to} \quad |X| \le k$$

パッチ $\mathbf{y}_t$ の
スパース表現のための
アトム集合 Z_t を選択

辞書作成への３つのアプローチ

既存の辞書を使う

信号処理の研究で考案された
既存の辞書を使う

欠点 データに合わせられない

辞書学習

与えられたデータから
辞書を学習する

欠点 難しい非凸最適化

辞書選択 [Krause–Cevher’10]

既存の辞書の和集合から
要素を選択して辞書を作る

辞書選択に対する既存アルゴリズム

アルゴリズム	理論保証	実験的な 出力の質	計算時間
Modular Approximation	○	×	○
OMP evaluation	×	○	×

本研究の貢献

辞書選択のための実用的なアルゴリズムは存在しなかった

辞書学習に匹敵する良質な辞書を

大幅に高速に出力するアルゴリズムを提案

そのほかの貢献

- より複雑な疎性制約に対するアルゴリズムを提案
- 辞書選択をオンライン設定へと拡張

目次

Replacement Greedy と Replacement OMP

二段階劣モジュラ最大化のための

Replacement Greedy [Stan+'17]

Replacement Greedy と Replacement OMP

二段階劣モジュラ最大化のための

Replacement Greedy [Stan+'17]

↓

貢献 1 Replacement Greedy を辞書選択に適用

Replacement Greedy

O(s^2dknT) 時間

Replacement Greedy と Replacement OMP

二段階劣モジュラ最大化のための

Replacement Greedy [Stan+'17]

貢献 1 Replacement Greedy を辞書選択に適用

Replacement Greedy

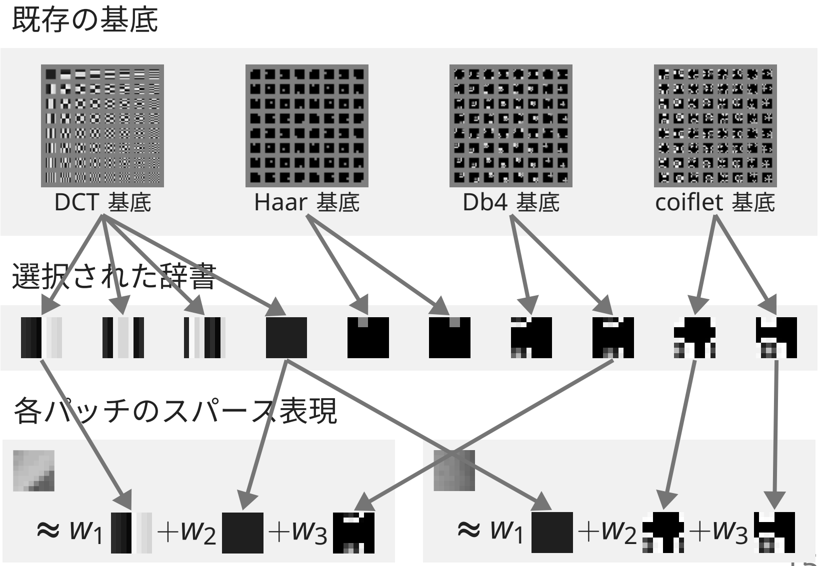
$\mathrm{O}(s^2 dknT)$ 時間

貢献 2 OMP のアイデアでおよそ $\mathrm{O}(s^2 d)$ 倍高速化

Replacement OMP

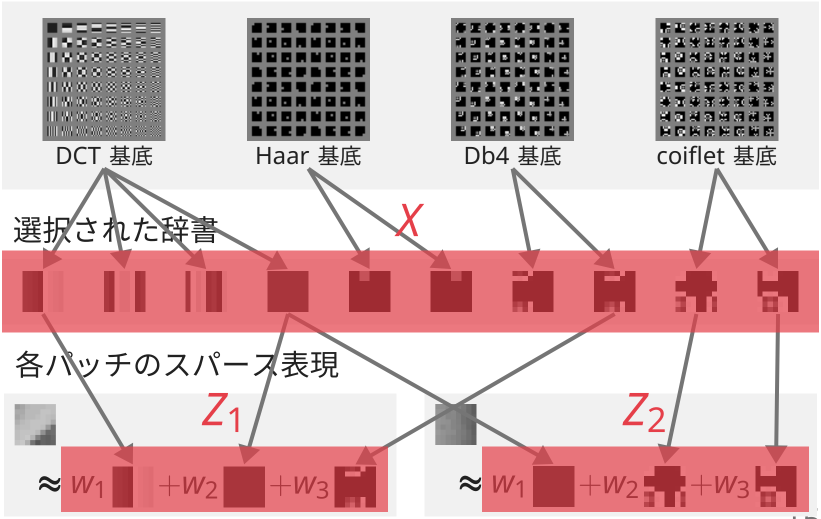
$\mathrm{O}((n + ds)kT)$ 時間

辞書選択


既存の基底
DCT 基底
Haar 基底
Db4 基底
coiflet 基底
選択された辞書
各パッチのスパース表現
≈ w_1
$+w_2$
$+w_3$
≈ w_1
$+w_2$
$+w_3$

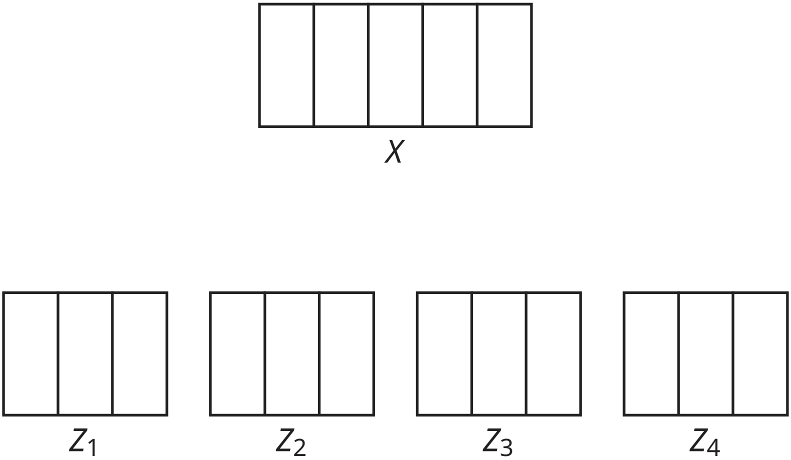
辞書選択

既存の基底


DCT 基底
Haar 基底
Db4 基底
coiflet 基底
選択された辞書
X
各パッチのスパース表現
Z1
Z2
≈ w1 +w2 +w3
≈ w1 +w2 +w3

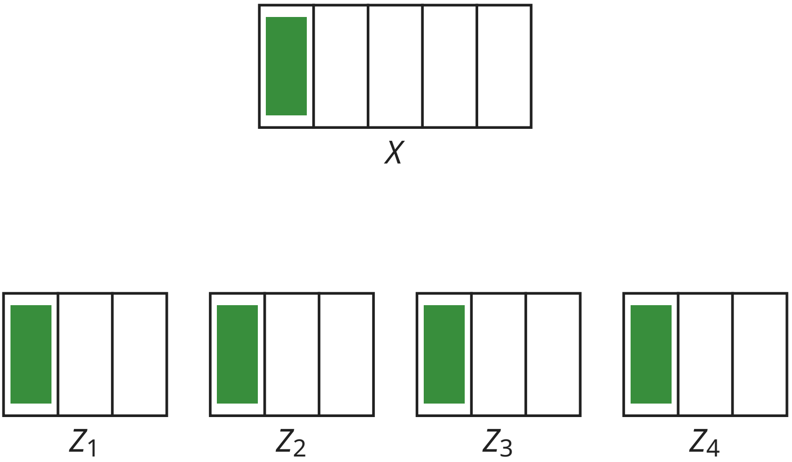
Replacement Greedy と Replacement OMP

$k = 5, s = 3, T = 4$



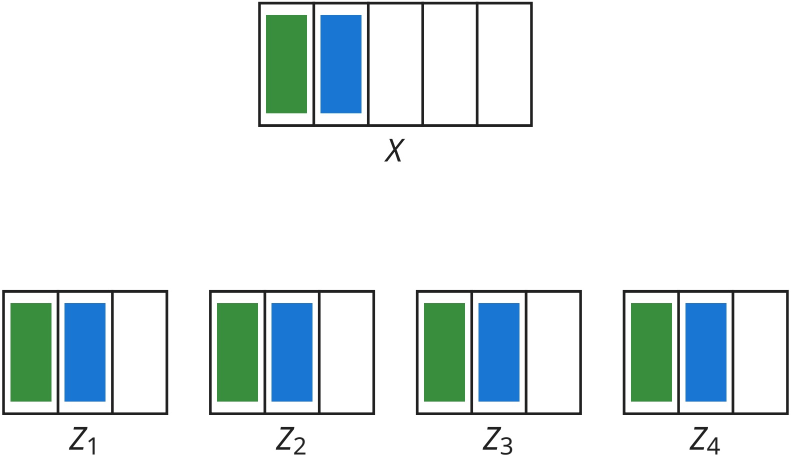
Replacement Greedy と Replacement OMP

$k = 5, s = 3, T = 4$



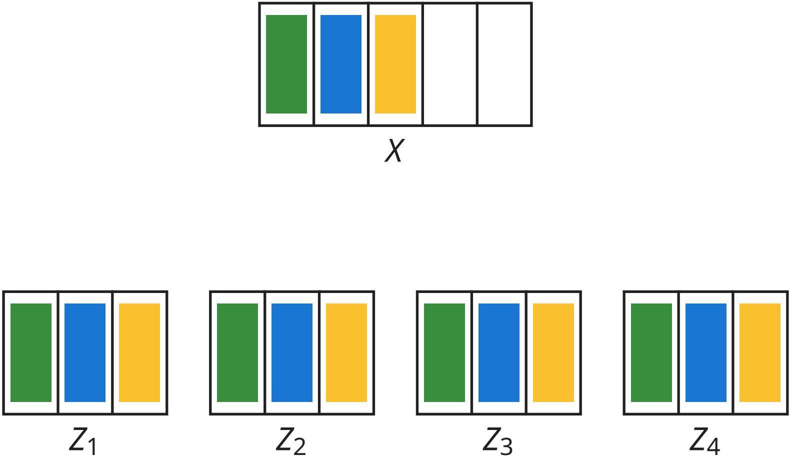
Replacement Greedy と Replacement OMP

$k = 5, s = 3, T = 4$



Replacement Greedy と Replacement OMP

$k = 5, s = 3, T = 4$



Replacement Greedy と Replacement OMP

$T = 4$

Replacement Greedy

$$\operatorname*{argmax}_{v \in V} \sum_{t=1}^{T} f_t(Z_t + v)$$

目的関数値を愚直に評価 → 遅い

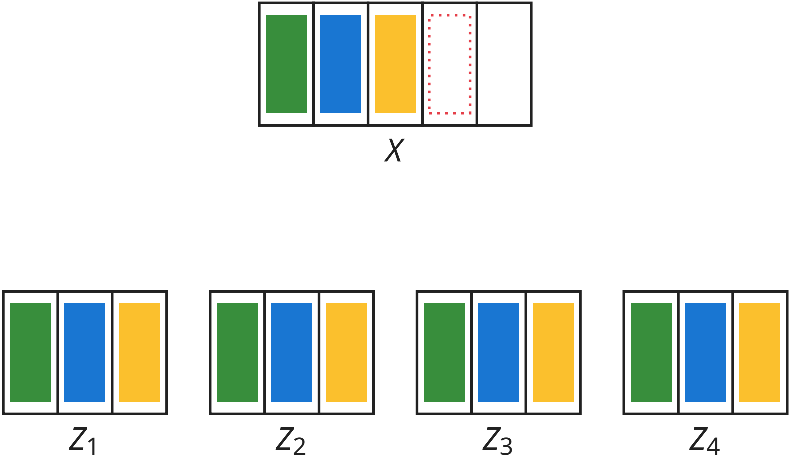
Replacement OMP

$$\operatorname*{argmax}_{v \in V} \sum_{t=1}^{T} (\mathbf{a}_v^\top \hat{\mathbf{y}}_t)^2 \quad (\hat{\mathbf{y}}_t \text{ 残差})$$

目的関数値を近似的に評価 → 速い

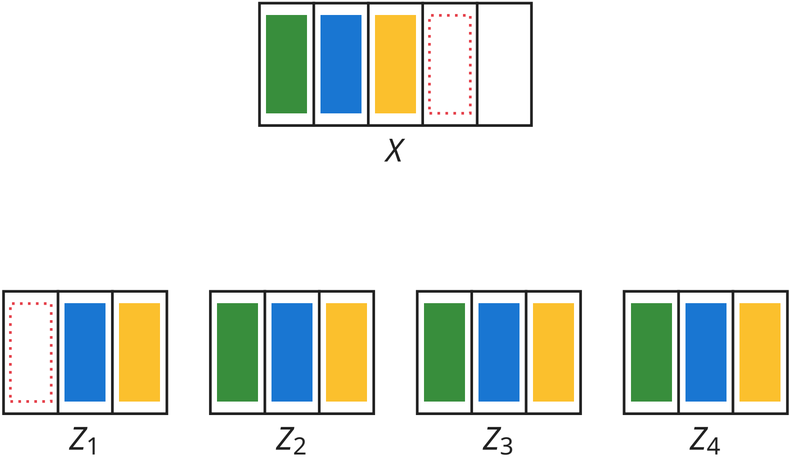
Replacement Greedy と Replacement OMP

$k = 5, s = 3, T = 4$



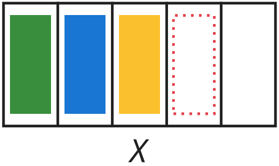
Replacement Greedy と Replacement OMP

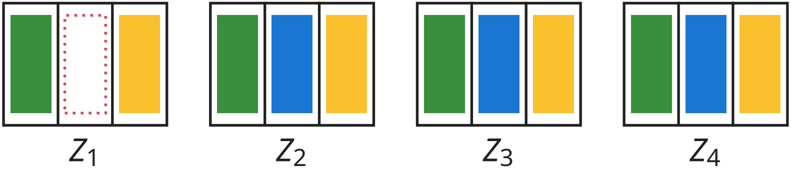
$k = 5, s = 3, T = 4$



Replacement Greedy と Replacement OMP

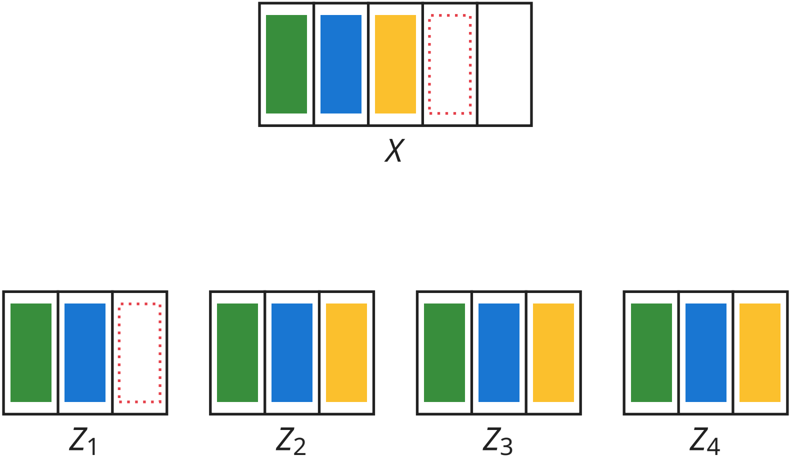
$k = 5, s = 3, T = 4$





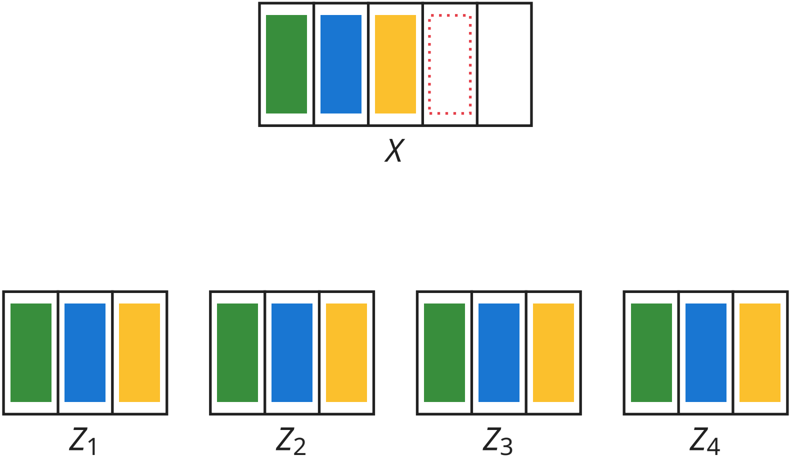
Replacement Greedy と Replacement OMP

$k = 5, s = 3, T = 4$



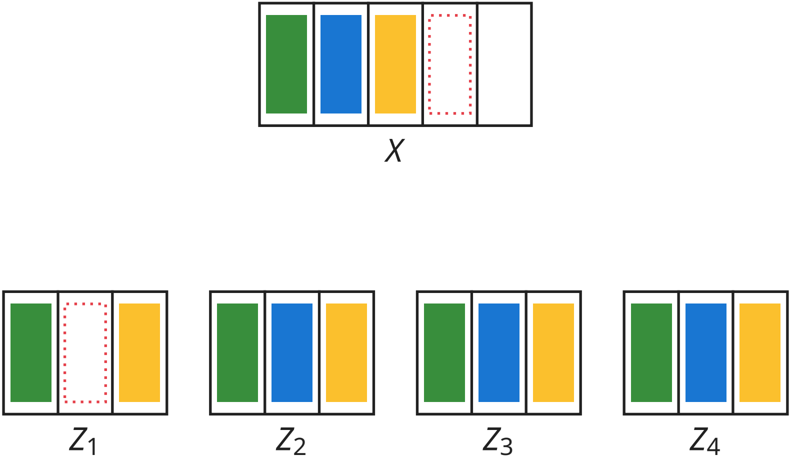
Replacement Greedy と Replacement OMP

$k = 5, s = 3, T = 4$



Replacement Greedy と Replacement OMP

$k = 5, s = 3, T = 4$



Replacement Greedy と Replacement OMP

Replacement Greedy

$$\operatorname*{argmax}_{v \in V} \sum_{t=1}^{T} \max \left\{ 0, \max_{u_t \in Z_t} f_t(Z_t - u_t + v) \right\}$$

目的関数値を愚直に評価 ➡ 遅い

Replacement OMP

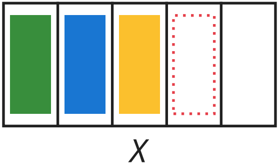
$$\operatorname*{argmax}_{v \in V} \sum_{t=1}^{T} \max \left\{ 0, \frac{1}{\Sigma_2^2} \left(\mathbf{a}_v^\top \hat{\mathbf{y}}_t \right)^2 - \min_{u_t \in Z_t} (\mathbf{w}_t^*)_{u_t}^2 \right\}$$

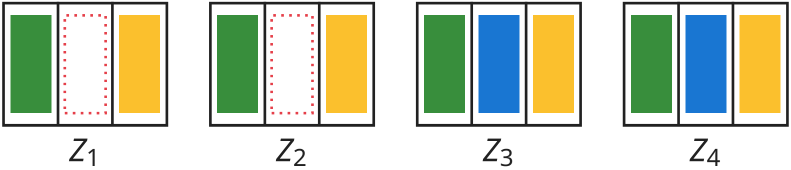
($\hat{\mathbf{y}}_t$ 残差、$\mathbf{w}_t^*$ 最適おもみベクトル)

目的関数値を近似的に評価 ➡ 速い

Replacement Greedy と Replacement OMP

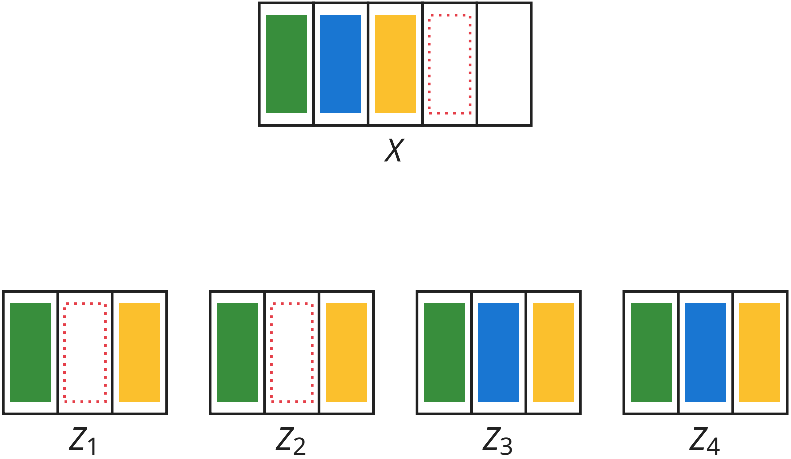
$k = 5, s = 3, T = 4$





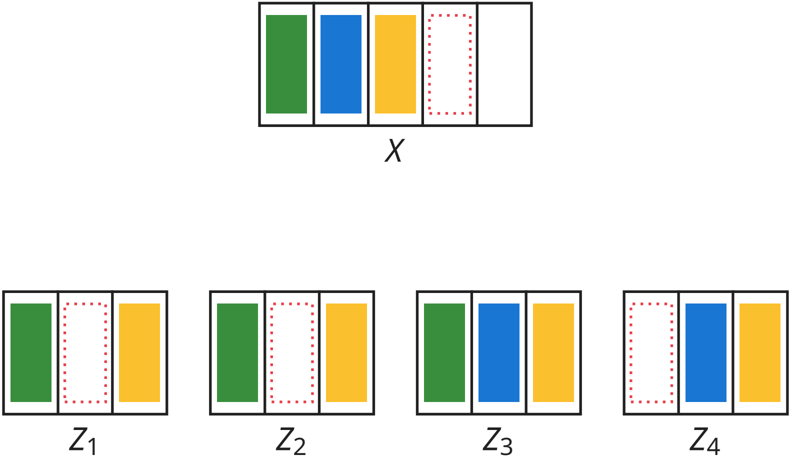
Replacement Greedy と Replacement OMP

$k = 5, s = 3, T = 4$



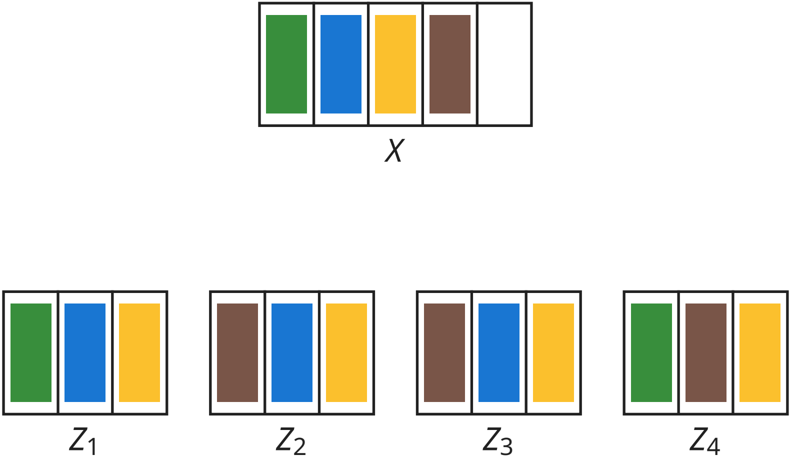
Replacement Greedy と Replacement OMP

$k = 5, s = 3, T = 4$



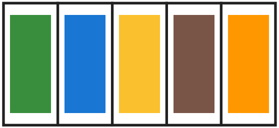
Replacement Greedy と Replacement OMP

$k = 5, s = 3, T = 4$

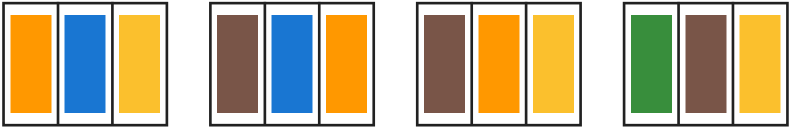


Replacement Greedy と Replacement OMP

$k = 5, s = 3, T = 4$



X



Z_1 Z_2 Z_3 Z_4

理論的な結果

定理

Replacement Greedy は $O(s^2 dknT)$ 時間、

Replacement OMP は $O((n+ds)kT)$ 時間で、

ともに $\left(\frac{\sigma_{2s}^2}{\Sigma_2^2}\right)^2 \left(1-\exp\left(-\frac{\Sigma_2^2}{\sigma_{2s}^2}\right)\right)$ 近似

$\Sigma_s \triangleq \max_{Z \subseteq V:\, |Z| \leq s} \sigma_{\max}(\mathbf{D}_Z)$ s 列の列部分行列の最大特異値

$\sigma_s \triangleq \min_{Z \subseteq V:\, |Z| \leq s} \sigma_{\min}(\mathbf{D}_Z)$ s 列の列部分行列の最小特異値

アルゴリズムのまとめ

アルゴリズム	近似比	計算時間
Modular Approximation	$\frac{\sigma_s^2}{\Sigma_s^2}\left(1-\frac{1}{e}\right)$	$O((k+d)nT)$
OMP evaluation	$O(1/k)$	$O(sdk^2nT)$
Replacement Greedy	$\frac{\sigma_{2s}^4}{\Sigma_2^4}\left(1-\exp\left(-\frac{\Sigma_2^2}{\sigma_{2s}^2}\right)\right)$	$O(s^2dknT)$
Replacement OMP	$\frac{\sigma_{2s}^4}{\Sigma_2^4}\left(1-\exp\left(-\frac{\Sigma_2^2}{\sigma_{2s}^2}\right)\right)$	$O((n+ds)kT)$

目次

実験設定

データセット　（すべてのパッチは 8 × 8 pixel）

- **人工データ:**

  ランダムに決めた真の辞書からランダムに生成

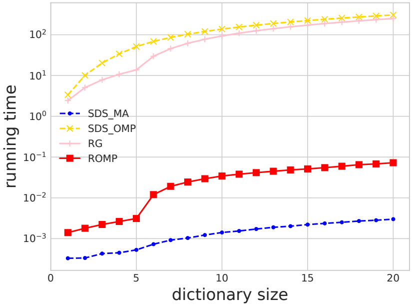
- **実データ**

  VOC2006 image dataset からランダムに抽出

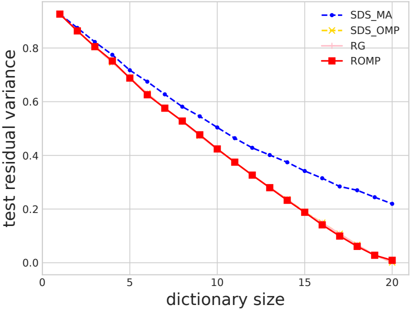
台集合　DCT、ウェーブレット（Haar, Db4, coiflet）

パラメータ　$T = 100$ or 1000、$s = 5$、試行 20 回の平均

実験結果：辞書選択の既存法と比較（人工データ）



計算時間

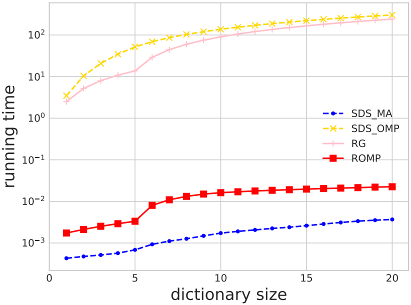


テスト残差

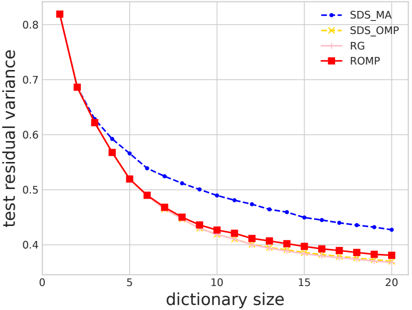
Replacement OMP は OMP Evaluation より高速

かつ Modular Approximation よりよい解を出力

実験結果：辞書選択の既存法と比較（実データ）



計算時間

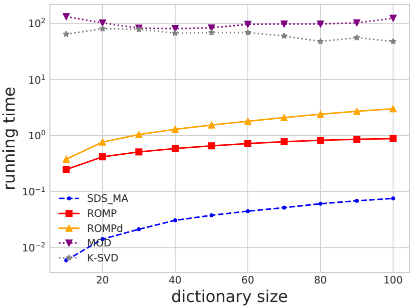


テスト残差

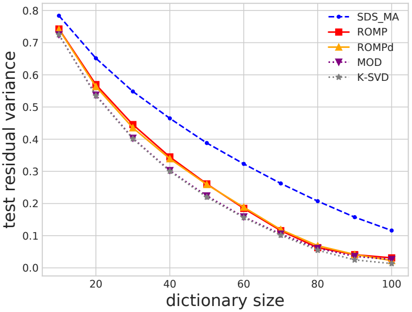
Replacement OMP は OMP Evaluation より高速

かつ Modular Approximation よりよい解を出力

実験結果：辞書学習の手法と比較（人工データ）



計算時間

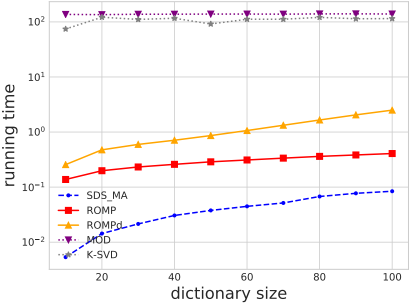


テスト残差

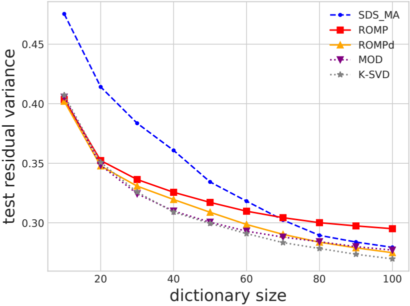
提案手法（Replacement OMP とその亜種）は

辞書学習（MOD、*k*-SVD）より高速に遜色ない解を出力

実験結果：辞書学習の手法と比較（実データ）



計算時間



テスト残差

提案手法（Replacement OMP とその亜種）は

辞書学習（MOD、*k*-SVD）より高速に遜色ない解を出力

アルゴリズムのまとめ

アルゴリズム	理論保証	実験的な 出力の質	計算時間
Modular Approximation	○	×	○
OMP evaluation	×	○	×
Replacement Greedy	○	○	×
Replacement OMP	○	○	○

本研究の貢献

辞書選択のための実用的なアルゴリズムは存在しなかった

辞書学習に匹敵する良質な辞書を

大幅に高速に出力するアルゴリズムを提案

そのほかの貢献

- より複雑な疎性制約に対するアルゴリズムを提案
- 辞書選択をオンライン設定へと拡張

目次

参考文献（1）

- E. Balkanski, B. Mirzasoleiman, A. Krause, and Y. Singer. Learning sparse combinatorial representations via two-stage submodular maximization. In ICML 2016, pp. 2207-2216.
- V. Cevher and A. Krause. Greedy dictionary selection for sparse representation. IEEE Journal of Selected Topics in Signal Processing, 5(5), pp. 979–988, 2011.
- A. Das and D. Kempe. Submodular meets spectral: Greedy algorithms for subset selection, sparse approximation and dictionary selection. In ICML 2011, pp. 1057–1064.

参考文献（2）

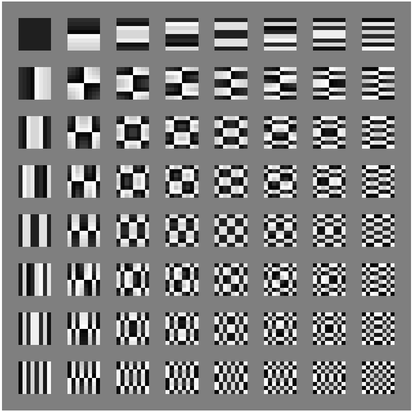
- E. R. Elenberg, R. Khanna, A. G. Dimakis, and S. Negahban. Restricted Strong Convexity Implies Weak Submodularity. In Annals of Statistics, 2018.
- A. Krause and V. Cevher. Submodular dictionary selection for sparse representation. In ICML 2010, pp. 567–574.
- B. Natarajan. Sparse approximation solutions to linear systems. SIAM Journal on Computing, 24, pp. 227–234, 1995.
- S. Stan, M. Zadimoghaddam, A. Krause and A. Karbasi. Probabilistic submodular maximization in sub-linear time. In ICML 2017, pp. 3241–3250.

目次

信号処理の知見を利用

信号処理の研究で開発された辞書を利用する

例）離散コサイン変換（DCT)、ウェーブレット、…



離散コサイン変換のサイズ 8 × 8 の基底

辞書学習

データから辞書を学習する

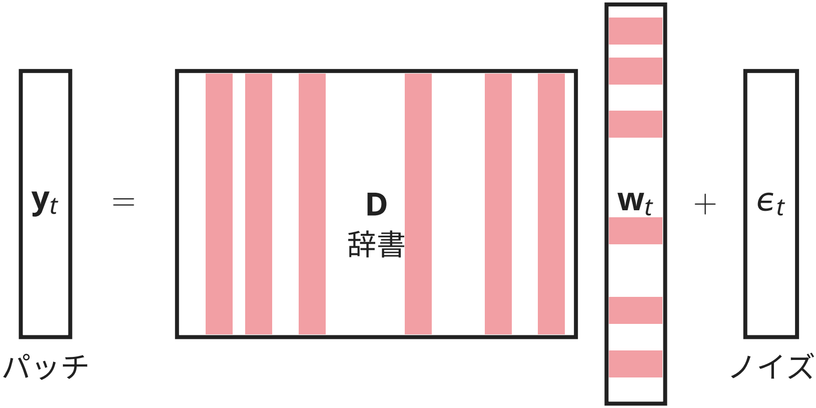
入力

- $\{\mathbf{y}_1, \cdots, \mathbf{y}_T\} \subset \mathbb{R}^d$ データ点（パッチ）の集合
- s スパース性のパラメータ
- k 辞書サイズの制約

出力 $\mathbf{D} \in \mathbb{R}^{d \times k}$ 辞書

辞書

ほとんどのパッチは辞書の元のスパース結合で表現できる



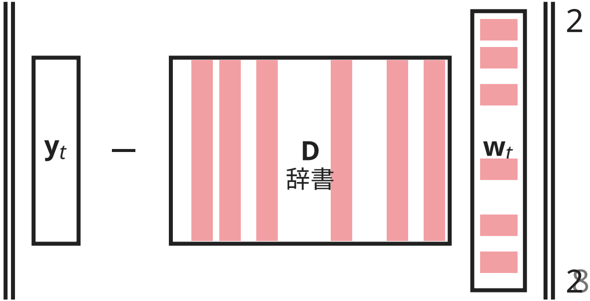
辞書学習

データから辞書を学習する問題

$$\text{Minimize}_{\mathbf{D}} \quad \sum_{t=1}^{T} \min_{\mathbf{w}_t : \|\mathbf{w}_t\|_0 \leq s} \|\mathbf{y}_t - \mathbf{D}\mathbf{w}_t\|_2^2$$

$$\text{subject to} \quad \mathbf{D} \in \mathbb{R}^{d \times k}$$

$$\sum_{t=1}^{T} \min_{\mathbf{w}_t : \|\mathbf{w}_t\|_0 \leq s}$$



辞書学習

データから辞書を学習する問題

$$\begin{aligned}&\text{Minimize}_{\mathbf{D}} && \sum_{t=1}^{T} \min_{\mathbf{w}_t : \|\mathbf{w}_t\|_0 \le s} \|\mathbf{y}_t - \mathbf{D}\mathbf{w}_t\|_2^2 \\ &\text{subject to} && \mathbf{D} \in \mathbb{R}^{d \times k}\end{aligned}$$

Remark

- 目的関数値の評価は NP 困難 [Natarajan'95]
- 最も一般的なアプローチは ℓ_0 ノルムの ℓ_1 ノルムへの緩和だが、緩和しても非凸

貪欲法

空集合から始めて、各ステップ増分最大の要素を追加

集合関数の最大化

集合関数 $f: 2^V \to \mathbb{R}_{\geq 0}$

Maximize $\quad f(S)$

subject to $\quad |S| \leq k$

要素数の制約

1: $X \leftarrow \emptyset$.

2: **for** $|X| < k$ **do**

3: $\quad v^* \in \mathrm{argmax}\{f(X+v) \mid v \in V\}$

4: $\quad X \leftarrow X + v^*$

最小化問題から最大化問題へ

$$\begin{aligned} \text{Minimize}_{X \subseteq V} \quad & \sum_{t=1}^{T} \min_{\mathbf{w}:\|\mathbf{w}\|_0 \leq s} \|\mathbf{y}_t - \mathbf{A}_X \mathbf{w}\|_2^2 \\ \text{subject to} \quad & |X| \leq k \end{aligned}$$

$$\Downarrow$$

$$\begin{aligned} \text{Maximize}_{X \subseteq V} \quad & \sum_{t=1}^{T} \left\{ \|\mathbf{y}_t\|_2^2 - \min_{\mathbf{w}:\|\mathbf{w}\|_0 \leq s} \|\mathbf{y}_t - \mathbf{A}_X \mathbf{w}\|_2^2 \right\} \\ \text{subject to} \quad & |X| \leq k \end{aligned}$$

※ 近似比の意味では、これらの問題は等価でない

二段階最適化問題としての辞書選択

$$\begin{array}{ll}\text{Maximize}_{X \subseteq V} & \displaystyle\sum_{t=1}^{T} \left\{ \|\mathbf{y}_t\|_2^2 - \min_{\mathbf{w}:\|\mathbf{w}\|_0 \leq s} \|\mathbf{y}_t - \mathbf{A}_X \mathbf{w}\|_2^2 \right\} \\ \text{subject to} & |X| \leq k \end{array}$$

$$f_t(Z_t) := \|\mathbf{y}_t\|_2^2 - \min_{\mathbf{w}} \|\mathbf{y}_t - \mathbf{A}_{Z_t} \mathbf{w}\|_2^2$$

$$\begin{array}{ll}\text{Maximize}_{X \subseteq V} & \displaystyle\sum_{t=1}^{T} \max_{Z_t \subseteq X:\ |Z_t| \leq s} f_t(Z_t) \\ \text{subject to} & |X| \leq k \end{array}$$

台集合に関する仮定

- $\mathbf{A}$ の各列の正規化を仮定（$\forall i \in V$ について $\|\mathbf{a}_i\|_2 = 1$）
- $$\Sigma_s \triangleq \max_{Z \subseteq V:\, |Z| \le s} \sigma_{\max}(\mathbf{D}_Z)$$

  s 列の列部分行列の最大特異値

  $$\sigma_s \triangleq \min_{Z \subseteq V:\, |Z| \le s} \sigma_{\min}(\mathbf{D}_Z)$$

  s 列の列部分行列の最小特異値

既存手法 1：Modular Approximation
[Krause–Cevher'10, Das–Kempe'11]

モジュラ関数（線形関数）を用いて目的関数を近似

$$\text{Maximize}_{X \subseteq V} \quad h(X) = \sum_{t=1}^{T} \max_{Z_t \subseteq X:\ |Z_t| \leq s} f_t(Z_t)$$

$$\tilde{f}_t(Z_t) = \sum_{i \in Z_t} f_t(\{i\})$$ 線形近似

$$\text{Maximize}_{X \subseteq V} \quad \tilde{h}(X) = \sum_{t=1}^{T} \max_{Z_t \subseteq X:\ |Z_t| \leq s} \tilde{f}_t(Z_t)$$

既存手法 1：Modular Approximation

[Krause–Cevher’10, Das–Kempe’11]

補題 1 $\frac{1}{\Sigma_s^2} h(X) \leq \tilde{h}(X) \leq \frac{1}{\sigma_s^2} h(X)$

補題 2 $\tilde{h}$ は単調劣モジュラ

定理 **[Das–Kempe’11]**

Modular Approximation は $\frac{\sigma_s^2}{\Sigma_s^2}\left(1 - \frac{1}{e}\right)$ 近似

※ MA はアトムのあいだの相関を無視しているため、
実用的な性能はあまりよくない

Orthogonal Matching Pursuit

スパース線形回帰に対する貪欲法の一種

$$
\begin{aligned}
\text{Maximize}_{Z \subseteq X} \quad & f(Z) = \|\mathbf{y}\|_2^2 - \min_{\mathbf{w}} \|\mathbf{y} - \mathbf{A}_{Z_t}\mathbf{w}\|_2^2 \\
\text{subject to} \quad & |Z| \leq s
\end{aligned}
$$

Orthogonal Matching Pursuit

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

$$\mathbf{w}_Z^* \in \operatorname*{argmin}_{\mathbf{w}} \|\mathbf{y} - \mathbf{A}_Z\mathbf{w}\|_2^2$$

$$Z \leftarrow Z + v \text{ where } v \in \operatorname*{argmax}_{v \in V} \left| \mathbf{a}_v^\top \big(\underbrace{\mathbf{y} - \mathbf{D}_Z\mathbf{w}_Z^*}_{\text{残差}} \big) \right|$$

既存手法 2：OMP evaluation
[Krause–Cevher'10, Das–Kempe'11]

目的関数値を OMP で近似的に計算

$$\text{Maximize}_{X \subseteq V} \quad h(X) = \sum_{t=1}^{T} \max_{Z_t \subseteq X:\ |Z_t| \leq s} f_t(Z_t)$$

$Z_t^{X,\mathrm{OMP}}$ f_t を目的関数、X を台集合としたときの OMP の解

$$\text{Maximize}_{X \subseteq V} \quad \tilde{h}(X) = \sum_{t=1}^{T} f_t(Z_t^{X,\mathrm{OMP}})$$

既存手法 2：OMP evaluation
[Krause–Cevher'10, Das–Kempe'11]

目的関数値を OMP で近似的に計算

定理 [Das–Kempe'11]

OMP evaluation は $\dfrac{\sigma_s^2}{\Sigma_s^2} \cdot \dfrac{1 - \exp(-p\sigma_s^2)}{k - kp\sigma_s^2 + 1}$ 近似

ただし、$p = (1 - \exp(-\sigma_{2s}^4))/\Sigma_s^2$

※ この近似比は Modular Approximation より大幅に悪い

目次

二段階劣モジュラ最大化

$$\begin{aligned} \text{Maximize}_{X \subseteq V} \quad & \sum_{t=1}^{T} \max_{Z_t \subseteq X:\, |Z_t| \le s} f_t(Z_t) \\ \text{subject to} \quad & |X| \le k \end{aligned}$$

- $f_t(Z) = \|\mathbf{y}_t\|_2^2 - \min_{\mathbf{w}} \|\mathbf{y}_t - \mathbf{A}_Z \mathbf{w}\|_2^2$ 劣モジュラでない

  ➡ **辞書選択**

- f_t が単調劣モジュラ

  ➡ **二段階劣モジュラ最大化** [Balkanski+'16]

Replacement Greedy [Stan+'17] は二段階劣モジュラ最大化のためのアルゴリズムとして提案された

Replacement Greedy

$$\text{Maximize}_{X \subseteq V} \sum_{t=1}^{T} \max_{Z_t \subseteq X:\, |Z_t| \le s} f_t(Z_t) \text{ subj. to } |X| \le k$$

0. 初期化 $X := \emptyset, Z_t := \emptyset \ (\forall t \in [T])$

1. 最初の s ステップ

   $\operatorname{argmax}_{v \in V} \sum_{t=1}^{T} f_t(Z_t + v)$ を X と $Z_t (\forall t \in [T])$ に追加

2. $s+1$ ステップ目から k ステップ目

   $\operatorname{argmax}_{v \in V} \sum_{t=1}^{T} \max \left\{ 0, \max_{u_t \in Z_t} f_t(Z_t - u_t + v) \right\}$ を X に追加し、各 Z_t に置換を適用

理論的な結果

定理

Replacement Greedy は

$O(s^2dknT)$ 時間で $\frac{\sigma_{2s}^2}{\Sigma_2^2}\left(1-\exp\left(-\frac{\Sigma_2}{\sigma_{2s}}\right)\right)$ 近似解を出力

※ Modular Approximation の近似比 $\frac{\sigma_s^2}{\Sigma_s^2}\left(1-\frac{1}{e}\right)$ との比較は困難

Replacement Greedy + OMP = ?

OMP のアイデアを用いて Replacement Greedy を高速化

スパース線形回帰

Forward Regression $O(s^2dn)$ → OMP $O((s^2+n)sd)$

辞書選択

Replacement Greedy $O(s^2dknT)$ → Replacement OMP $O((n+ds)kT)$

Forward Regression vs. OMP

$$\mathrm{Max.}_{Z \subseteq X}\, f(Z) = \|\mathbf{y}\|_2^2 - \min_{\mathbf{w}} \|\mathbf{y} - \mathbf{A}_{Z_t}\mathbf{w}\|_2^2 \text{ s.t. } |Z| \le s$$

FR 各ステップで $\operatorname{argmax}_{v \in V} f(Z + v)$ を Z に追加

OMP 各ステップで $\operatorname{argmax}_{v \in V} |\mathbf{a}_v^\top \hat{\mathbf{y}}|$ を Z に追加

ただし、$\underbrace{\hat{\mathbf{y}} \triangleq \mathbf{y} - \mathbf{D}_Z \mathbf{w}^*}_{\text{残差}}$ および $\mathbf{w}^* \in \underbrace{\operatorname{argmin}_{\mathbf{w}} \|\mathbf{y} - \mathbf{D}_Z \mathbf{w}\|_2^2}_{Z \text{ に対して最適な } \mathbf{w}}$

定理 [Das–Kempe'11, Elenberg+'18]

これらのアルゴリズムはともに $(1 - \exp(-\sigma_s^2))$ 近似

Replacement OMP（最初の s ステップ）

Replacement Greedy

$$\underset{v \in V}{\mathrm{argmax}} \sum_{t=1}^{T} f_t(Z_t + v) \text{ を } X \text{ と } Z_t(\forall t) \text{ に追加}$$

Replacement OMP

$$\underset{v \in V}{\mathrm{argmax}} \sum_{t=1}^{T} (\mathbf{a}_v^\top \hat{\mathbf{y}}_t)^2 \text{ を } X \text{ と } Z_t(\forall t) \text{ に追加}$$

$$\left(\underbrace{\hat{\mathbf{y}}_t \triangleq \mathbf{y}_t - \mathbf{A}_{Z_t}\mathbf{w}_t^*}_{\text{残差}} \text{、} \underbrace{\mathbf{w}_t^* \in \underset{\mathbf{w}}{\mathrm{argmin}} \|\mathbf{y}_t - \mathbf{A}_{Z_t}\mathbf{w}\|_2^2}_{Z \text{ に対して最適な } \mathbf{w}} \right)$$

Replacement OMP（$s+1$ ステップ目以降）

Replacement Greedy

$$\operatorname*{argmax}_{v \in V} \sum_{t=1}^{T} \max\left\{0, \max_{u_t \in Z_t} f_t(Z_t - u_t + v)\right\}$$ を

X に追加し、各 Z_t に置換を適用

Replacement OMP

$$\operatorname*{argmax}_{v \in V} \sum_{t=1}^{T} \max\left\{0, \frac{1}{\Sigma_2^2}\left(\mathbf{a}_v^\top \hat{\mathbf{y}}_t\right)^2 - \min_{u_t \in Z_t}(\mathbf{w}_t^*)_{u_t}^2\right\}$$ を

X に追加し、各 Z_t に置換を適用

$$\left(\hat{\mathbf{y}}_t \triangleq \mathbf{y}_t - \mathbf{A}_{Z_t}\mathbf{w}_t^*、\mathbf{w}_t^* \in \operatorname*{argmin}_{\mathbf{w}} \|\mathbf{y}_t - \mathbf{A}_{Z_t}\mathbf{w}\|_2^2\right)$$

Replacement OMP のための事前計算

Replacement OMP の実行時に Σ_2 の値が必要

解決策 1 <u>Σ_2 の値を事前に計算しておく</u>

$$\Sigma_2 = \max_{Z \subseteq V:\ |Z| \leq 2} \sigma_{\max}(\mathbf{D}_Z) = 1 + \max_{i \neq j} \left| \mathbf{a}_i^\top \mathbf{a}_j \right|$$

は $O(n^2 d)$ 時間で計算可能

解決策 2 <u>Σ_2 の代わりに上界 2 を使う</u>

$\Sigma_2 \leq 2$ は各列が正規化された任意の $\mathbf{A}$ について成立

近似比は $\dfrac{\sigma_{2s}^2}{4}\left(1 - \exp\left(-\dfrac{2}{\sigma_{2s}}\right)\right)$

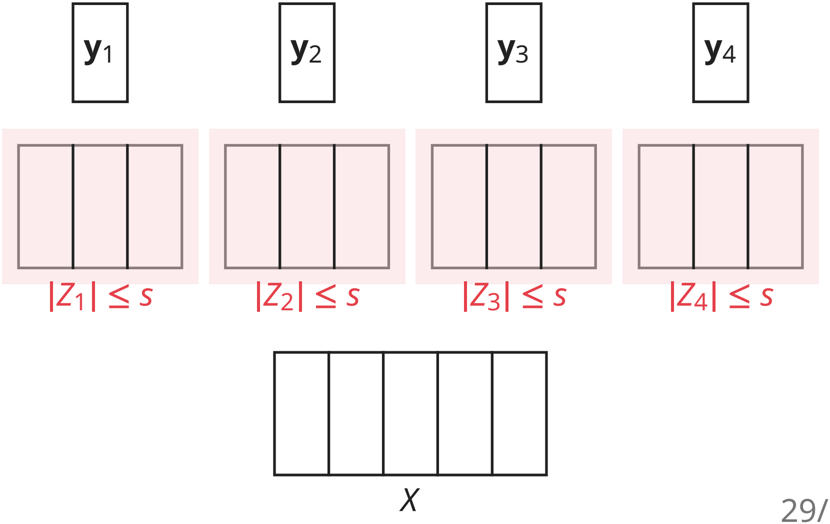
アルゴリズムのまとめ

アルゴリズム	近似比	計算時間
Modular Approximation	$\frac{\sigma_s^2}{\Sigma_s^2}\left(1-\frac{1}{\mathrm{e}}\right)$	$\mathrm{O}((k+d)nT)$
OMP evaluation	複雑	$\mathrm{O}(sdk^2nT)$
Replacement Greedy	$\frac{\sigma_{2s}^2}{\Sigma_2^2}\left(1-\exp\left(-\frac{\Sigma_2}{\sigma_{2s}}\right)\right)$	$\mathrm{O}(s^2dknT)$
Replacement OMP	$\frac{\sigma_{2s}^2}{\Sigma_2^2}\left(1-\exp\left(-\frac{\Sigma_2}{\sigma_{2s}}\right)\right)$	$\mathrm{O}((n+ds)kT)$

目次

いままでの疎性制約

各パッチに対して s 個まで選べる



動機

パッチによって線形表現に必要なアトム数が異なるのでは？

単純

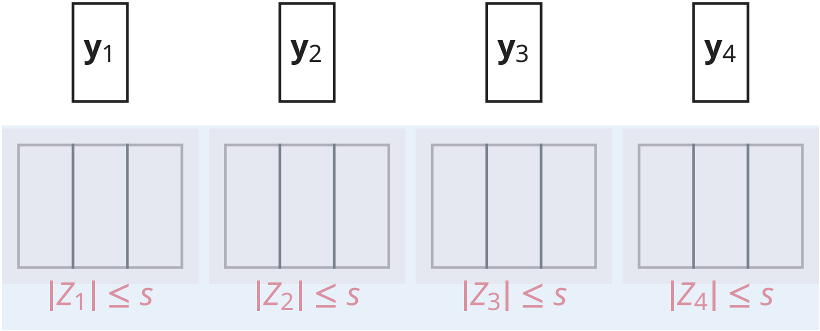
→ 必要なアトムは少ない

複雑

→ 必要なアトムは多い

平均疎性制約 [Cevher–Krause'11]

元の制約に加えて、選べるアトム数の合計を s' 個までに制限

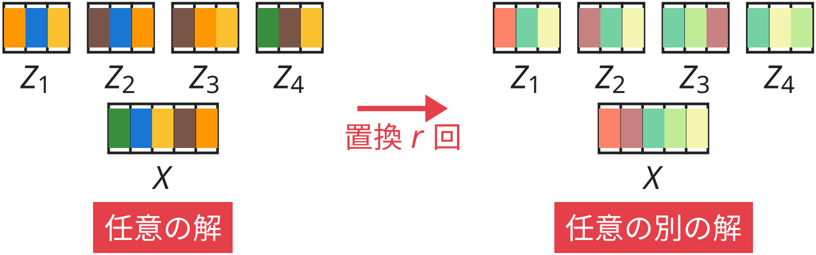


$$\sum_{t=1}^{T} |Z_t| \leq s'$$

一般的な疎性制約のクラス

制約の複雑さを表すパラメータ r を導入

ある解から別の解へ、何回置換すれば移せるか



一般化 Replacement OMP

一般的な疎性制約のために Replacement OMP を一般化

定理

パラメータ r の疎性制約に対して、

Replacement OMP は $\frac{\sigma_{2s}^2}{\Sigma_2^2}\left(1-\exp\left(-\frac{r}{k}\frac{\Sigma_2}{\sigma_{2s}}\right)\right)$ 近似

さまざまな疎性制約

疎性制約	r の上界
元の制約	$r \leq k$
マトロイドの直和	$r \leq k$
平均疎性制約（全体のみ）	$r \leq 2k - 1$
平均疎性制約	$r \leq 3k - 1$

目次

オンライン学習

さまざまなオンライン意思決定問題を含む枠組み

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

1. 学習者が行動 $X_t \in \mathcal{F}$ を決める
2. 敵対者が目的関数 g_t を明らかにする
3. 学習者が利得 $g_t(X_t)$ を受け取る

学習者  敵対者

α リグレット

学習者が得た利得と、最適な固定戦略の利得の α 倍の差

$$\mathrm{regret}_{\alpha}(T) \triangleq \alpha \underbrace{\max_{X^* \in \mathcal{F}} \sum_{t=1}^{T} g_t(X^*)}_{\text{最適な固定戦略の利得}} - \underbrace{\sum_{t=1}^{T} g_t(X_t)}_{\text{学習者の利得}}$$

※ α がオフライン設定の近似比に対応

オンライン辞書選択

データ点の列 $\mathbf{y}_1, \cdots, \mathbf{y}_T$ から辞書を学習

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

1. 学習者が辞書 $\mathbf{A}_{X_t}$ を決める
2. 敵対者がデータ点 $\mathbf{y}_t$ を明らかにする
3. 学習者が利得 $g_t(X_t)$ を受け取る

ただし、$g_t(X_t) = \|\mathbf{y}_t\|_2^2 - \min_{\mathbf{w}:\|\mathbf{w}\|_0 \le s} \|\mathbf{y}_t - \mathbf{A}_{X_t}\mathbf{w}\|_2^2$

理論的な結果

アルゴリズム	α
Online Modular Approximation	$\frac{\sigma_s^2}{\Sigma_s^2}\left(1-\frac{1}{e}\right)$
Online Replacement Greedy Online Replacement OMP	$\frac{\sigma_{2s}^4}{\Sigma_2^4}\left(1-\exp\left(-\frac{\Sigma_2^2}{\sigma_{2s}^2}\right)\right)$

定理

これらのアルゴリズムはそれぞれの α について

$\mathrm{regret}_\alpha(T) \leq k\sqrt{2T\ln n}$ を満たす