# 圧縮センシングアルゴリズム の構造学習における構造パラ メータの二値化

大阪大学 基礎工学部 システム科学科 知能システム学コース 飯國研究室 B4 今井智也

# 目次

•研究背景

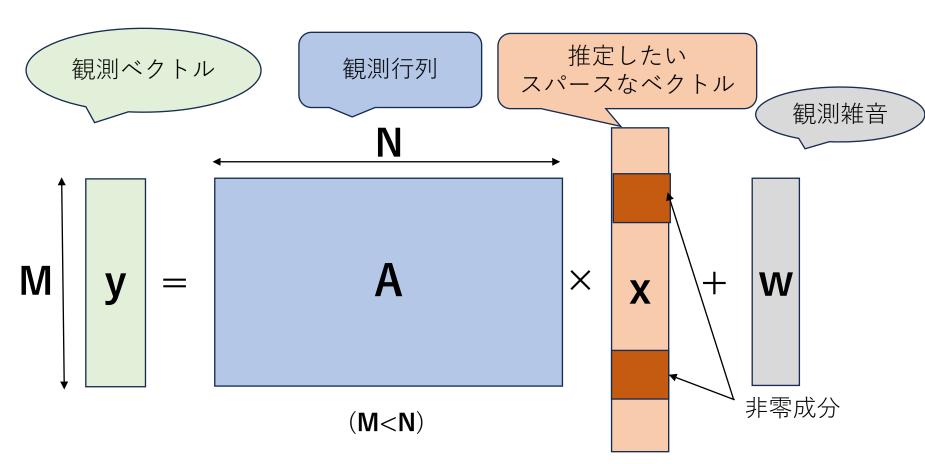
•研究目的

•実験結果

•結論

## 研究背景 | 圧縮センシング[1]

・少ない観測データ $y \in \mathbb{R}^M$ からスパースな(非零成分が少ない) $x \in \mathbb{R}^N$ を推定する



[1] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289-1306, April 2006, doi: 10.1109/TIT.2006.871582.

## 研究背景 | ISTA(iterative shrinkage thresholding algorithm) [2]

・圧縮センシングの最適化問題  $\widehat{x} = \operatorname*{argmin}_{x \in \mathbb{R}^N} \left\{ \frac{1}{2} \|y - Ax\|_2^2 + \lambda \|x\|_1 \right\}$  ( $\lambda$ :正則化係数)

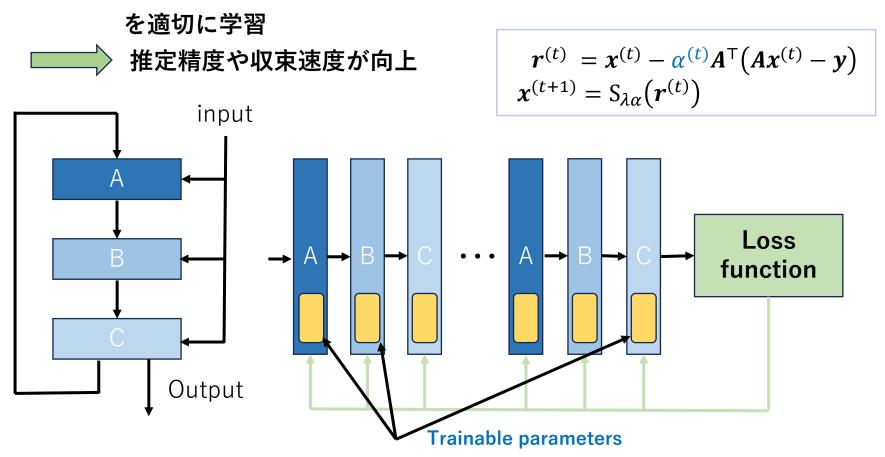
#### ·ISTA

$$m{r}^{(t)} = m{x}^{(t)} - lpha m{A}^{\mathsf{T}} m{A} m{x}^{(t)} - m{y}$$
  $\rightarrow$  勾配降下ステップ  $\mathbf{x}^{(t+1)} = S_{\lambda lpha} m{r}^{(t)}$   $\rightarrow$  縮小ステップ  $\mathbf{S}_{\lambda lpha}(x)$   $\mathbf{S}_{\lambda lpha}(x)$ :ソフト閾値関数 $(m{l}_1 m{J} m{n} m{L} m{A} m{n})$   $\mathbf{S}_{\lambda lpha}(x)$  ISTAは最適化問題の解を与える  $\mathbf{S}_{\lambda lpha}(x)$ 

[2] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences, vol. 57, no. 11, pp. 1413–1457, 2004

## 研究背景 | 深層展開[3]

深層展開:深層学習技術を反復型アルゴリズムに適用することでパラメータ



#### LISTA(Learned ISTA):ISTAに深層展開技術を適用したアルゴリズム

[3]K. Gregor and Y. LeCun, "Learning fast approximations of sparse coding," in Proc. the 27th International Conference on International Conference on Machine Learning, Jun. 2010, pp. 399–406.

## 研究背景 | AS-ISTA (Architecture Searched-ISTA)[4]

- ・従来の深層展開ではアルゴリズムのパラメータのみを 学習して構造自体は固定していた
- ・スパース信号復元により特化した構造も学習する

#### AS-ISTA

$$m{r}^{(t)} = m{w}_{x,f}^{(t)} imes f(m{x}^{(t)}) + m{w}_{x,g}^{(t)} imes g(m{x}^{(t)})$$
 $m{x}^{(t+1)} = m{w}_{r,x}^{(t)} imes f(m{r}^{(t)}) + m{w}_{r,g}^{(t)} imes g(m{r}^{(t)})$ 
 $m{g}(m{x}) = m{x} - m{\alpha}^{(t)} m{A}^{\mathsf{T}} (m{A} m{x} - m{y})$ : 勾配降下ステップ  $m{w}_{x,f}^{(t)} = \frac{\exp(eta_1^{(t)})}{\exp(eta_1^{(t)}) + \exp(eta_2^{(t)})}, \mbox{} m{w}_{x,g}^{(t)} = \frac{\exp(eta_2^{(t)})}{\exp(eta_1^{(t)}) + \exp(eta_2^{(t)})}$ 
 $m{g}(m{x}) = m{S}_{\lambda a^{(t)}}(m{x})$ 
 $m{x}_{r,x} = \frac{\exp(eta_1^{(t)})}{\exp(eta_1^{(t)}) + \exp(eta_2^{(t)})}, \mbox{} m{w}_{r,g}^{(t)} = \frac{\exp(eta_2^{(t)})}{\exp(eta_1^{(t)}) + \exp(eta_2^{(t)})}$ 

[4]長久, 早川, 飯國, "圧縮センシングアルゴリズムの構造学習, " 第38回信号処理シンポジウム, 2023年11月.

## 研究目的IAS-ISTAの重みの二値化

#### AS-ISTA

構造パラメータ:  $\beta_1$ ,  $\beta_2$ ,  $\gamma_1$ ,  $\gamma_2$ 

**重み**部分がソフトマックス型のため重みの値は0より大きく1未満の値をとる

ISTAにくらべて2倍の計算量がかかる



重みを二値化することでどちらのステップを利用しているかを 明確にするとともに計算量を削減

### 0と1の値をとることのできる新たな重みを設定

$$\mathbf{r}^{(t)} = \mathbf{w}_{x,f}^{(t)} \times f(\mathbf{x}^{(t)}) + \mathbf{w}_{x,g}^{(t)} \times g(\mathbf{x}^{(t)})$$
$$\mathbf{x}^{(t+1)} = \mathbf{w}_{r,x}^{(t)} \times f(\mathbf{r}^{(t)}) + \mathbf{w}_{r,g}^{(t)} \times g(\mathbf{r}^{(t)})$$

#### パターン1

$$w_{x,f}^{(t)} = \beta^{(t)}, \qquad w_{x,g}^{(t)} = 1 - \beta^{(t)}$$
 $w_{r,x}^{(t)} = 1 - \gamma^{(t)}, \qquad w_{r,g}^{(t)} = \gamma^{(t)}$ 

#### パターン2

$$w_{x,f}^{(t)} = \frac{\beta_1^{(t)}}{\beta_1^{(t)} + \beta_2^{(t)}}, \qquad w_{x,g}^{(t)} = \frac{\beta_2^{(t)}}{\beta_1^{(t)} + \beta_2^{(t)}}$$

$$w_{r,x}^{(t)} = \frac{\gamma_1^{(t)}}{\gamma_1^{(t)} + \gamma_2^{(t)}}, \qquad w_{r,g}^{(t)} = \frac{\gamma_2^{(t)}}{\gamma_1^{(t)} + \gamma_2^{(t)}}$$

### AS-ISTAの重みの二値化

```
Algorithm
               BAS-ISTA(Binarized AS-ISTA)
Input:パラメータの初期値\alpha, \theta = (\beta_1, \beta_2, r_1, r_2)を設定
1: for T = 1, ..., T_{max} do (インクリメンタル学習)
      for i = 1, ..., n do
         \alpha \leftarrow Adam(\alpha, \nabla_{\alpha} L_{train}(\alpha, \theta)) (ステップサイズの更新)
3:
      end for
4:
5:
      for j = 1, ..., n do
         \theta \leftarrow Adam(\theta, \nabla L_{val}(\alpha, \theta)) (構造パラメータの更新)
6:
7:
      end for
8:
      for k = 1, ..., T do
       \theta_k \leftarrow threshold\_function(\theta_k) (構造パラメータの二値化)
9:
10:
      end for
```

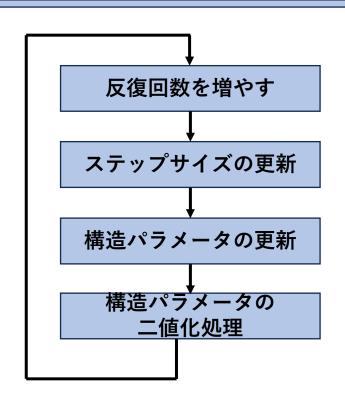
#### 構造パラメータの二値化方法

11:end for

$$\beta^{(t)} = \begin{cases} 1 & (\beta^{(t)} \ge 0.5) \\ 0 & (\beta^{(t)} < 0.5) \end{cases}$$
$$\gamma^{(t)} = \begin{cases} 1 & (\gamma^{(t)} \ge 0.5) \\ 0 & (\gamma^{(t)} < 0.5) \end{cases}$$

$$\beta_{1}^{(t)} = \begin{cases} 1 & (\beta_{1}^{(t)} \ge \beta_{2}^{(t)}) \\ 0 & (\beta_{1}^{(t)} < \beta_{2}^{(t)}) \end{cases} \qquad \gamma_{1}^{(t)} = \begin{cases} 1 & (\gamma_{1}^{(t)} \ge \gamma_{2}^{(t)}) \\ 0 & (\gamma_{1}^{(t)} < \gamma_{2}^{(t)}) \end{cases}$$
$$\beta_{2}^{(t)} = \begin{cases} 1 & (\beta_{2}^{(t)} \ge \beta_{1}^{(t)}) \\ 0 & (\beta_{2}^{(t)} < \beta_{1}^{(t)}) \end{cases} \qquad \gamma_{2}^{(t)} = \begin{cases} 1 & (\gamma_{2}^{(t)} \ge \gamma_{1}^{(t)}) \\ 0 & (\gamma_{2}^{(t)} < \gamma_{1}^{(t)}) \end{cases}$$

## AS-ISTAの重みの二値化



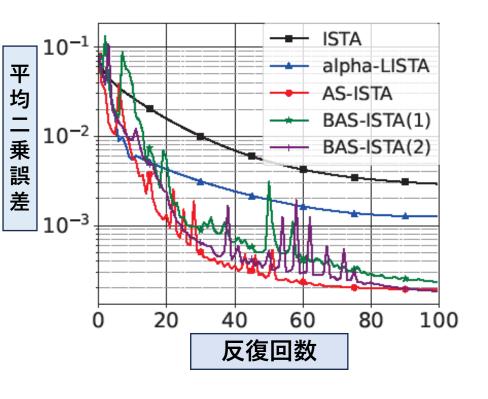
#### 構造パラメータの二値化方法

$$\beta^{(t)} = \begin{cases} 1 & (\beta^{(t)} \ge 0.5) \\ 0 & (\beta^{(t)} < 0.5) \end{cases}$$

$$\gamma^{(t)} = \begin{cases} 1 & (\gamma^{(t)} \ge 0.5) \\ 0 & (\gamma^{(t)} < 0.5) \end{cases}$$

$$\beta_{1}^{(t)} = \begin{cases} 1 & (\beta_{1}^{(t)} \ge \beta_{2}^{(t)}) \\ 0 & (\beta_{1}^{(t)} < \beta_{2}^{(t)}) \end{cases} \qquad \gamma_{1}^{(t)} = \begin{cases} 1 & (\gamma_{1}^{(t)} \ge \gamma_{2}^{(t)}) \\ 0 & (\gamma_{1}^{(t)} < \gamma_{2}^{(t)}) \end{cases}$$
$$\beta_{2}^{(t)} = \begin{cases} 1 & (\beta_{2}^{(t)} \ge \beta_{1}^{(t)}) \\ 0 & (\beta_{2}^{(t)} < \beta_{1}^{(t)}) \end{cases} \qquad \gamma_{2}^{(t)} = \begin{cases} 1 & (\gamma_{2}^{(t)} \ge \gamma_{1}^{(t)}) \\ 0 & (\gamma_{2}^{(t)} < \gamma_{1}^{(t)}) \end{cases}$$

## AS-ISTAとBAS-ISTAのグラフの比較



#### alpha-LISTA

$$r^{(t)} = x^{(t)} - \alpha^{(t)} A^{\mathsf{T}} (Ax^{(t)} - y)$$
  
$$x^{(t+1)} = S_{\lambda\alpha}(r^{(t)})$$

BAS-ISTA(1): パターン 1

$$\mathbf{r}^{(t)} = \beta^{(t)} \times f(\mathbf{x}^{(t)}) + (1 - \beta^{(t)}) \times g(\mathbf{x}^{(t)})$$
$$\mathbf{x}^{(t+1)} = (1 - \gamma^{(t)}) \times f(\mathbf{r}^{(t)}) + \gamma^{(t)} \times g(\mathbf{r}^{(t)})$$

BAS-ISTA(2): パターン2

$$\mathbf{r}^{(t)} = \frac{\beta_1^{(t)}}{\beta_1^{(t)} + \beta_2^{(t)}} \times f(\mathbf{x}^{(t)}) + \frac{\beta_2^{(t)}}{\beta_1^{(t)} + \beta_2^{(t)}} \times g(\mathbf{x}^{(t)}) 
\mathbf{x}^{(t+1)} = \frac{\gamma_1^{(t)}}{\gamma_1^{(t)} + \gamma_2^{(t)}} \times f(\mathbf{r}^{(t)}) + \frac{\gamma_2^{(t)}}{\gamma_1^{(t)} + \gamma_2^{(t)}} \times g(\mathbf{r}^{(t)})$$

重みを二値化したにもかかわらず、最終的にBAS-ISTAはAS-ISTAと 同様の精度を得ることができた

## 実験の結果

## 100反復時のアルゴリズムの計算時間の比較

表1 各アルゴリズムの100 反復にかかる計算時間とAS-ISTAとの 比較

アルゴリズム	計算時間[s]	AS-ISTA比(%)
ISTA	$1.72 \times 10^{-2}$	50.7
AS-ISTA	$3.39 \times 10^{-2}$	—
BAS-ISTA(1)	$1.84 \times 10^{-2}$	54.3
BAS-ISTA(2)	$1.83 \times 10^{-2}$	54.0

BAS-ISTAはAS-ISTAに比べて計算時間を削減することができた

### まとめ

・圧縮センシングのためのアルゴリズムであるISTAの精度を向上 させるため、パラメータと構造を学習させたAS-ISTAの重みの 二値化を行った

•

- ・二値化を行った際でも精度は変わらない結果を得た.
- ・反復時間の削減を行うことができた

## 今後の課題

重みを二値化した場合でも反復途中の精度の落ちないような学習 方法を検討する

### まとめ

- ・圧縮センシングのためのアルゴリズムであるISTAの精度を 向上させるため、パラメータと構造を学習させたAS-ISTAの 重みの二値化を行った
- ・二値化を行った際でも精度は変わらない結果を得た
- ・反復時間の削減を行うことができた

## 今後の課題

重みを二値化した場合でも反復途中の精度の落ちないような学習 方法を検討する