

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

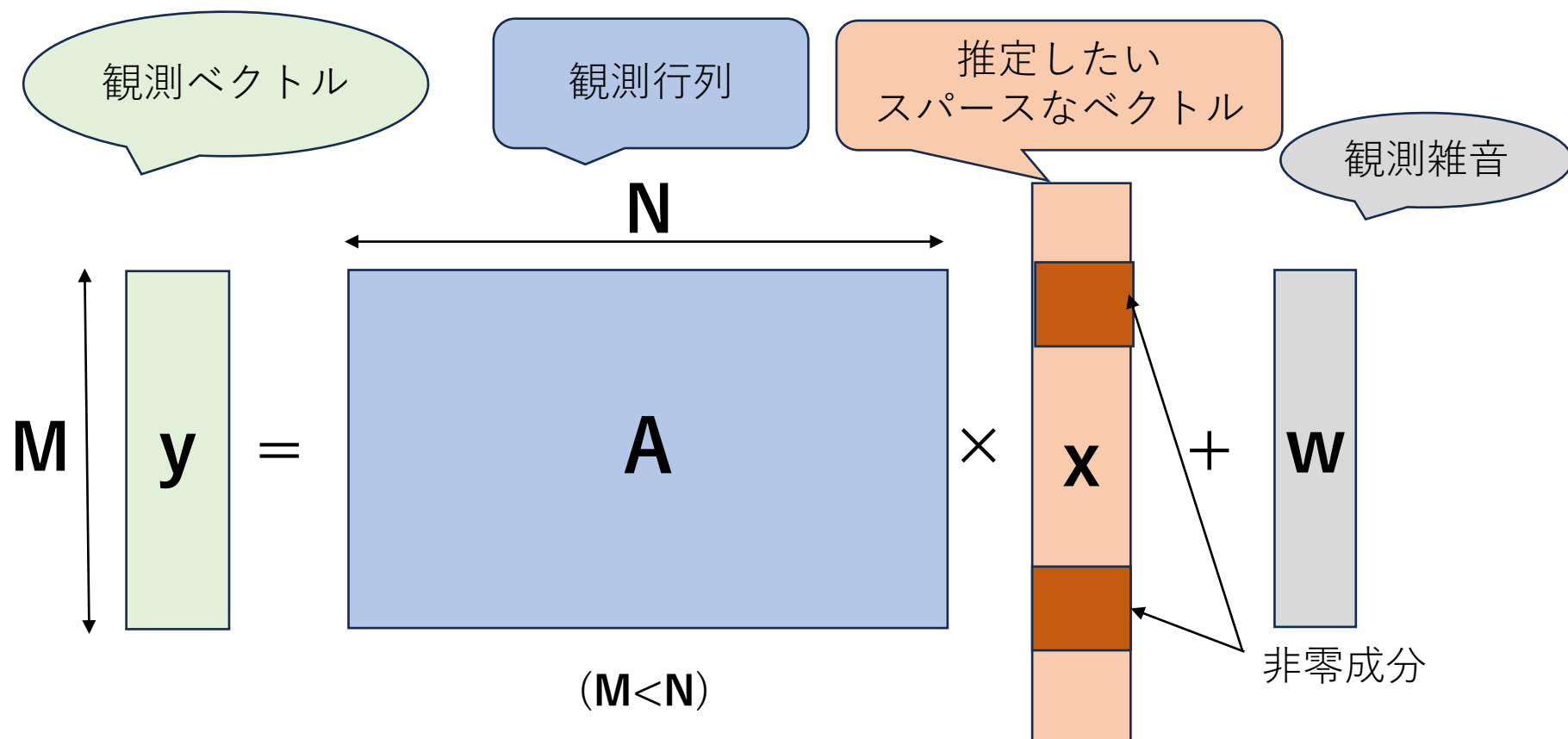
大阪大学 基礎工学部 システム科学科
知能システム学コース
飯國研究室 B4 今井智也

目次

- 研究背景
- 研究目的
- 實驗結果
- 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



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

・圧縮センシングの最適化問題

観測との誤差

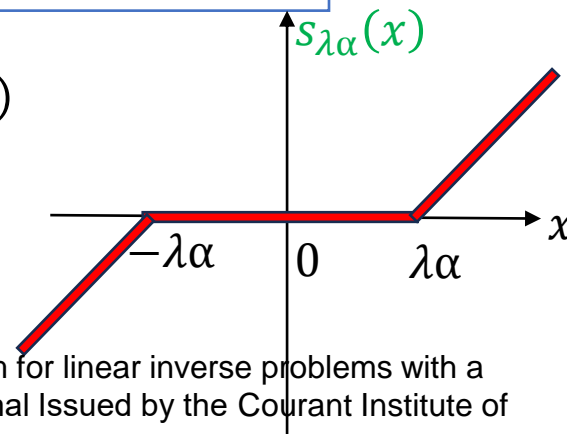
スパース性

$$\hat{\mathbf{x}} = \operatorname{argmin}_{\mathbf{x} \in \mathbb{R}^N} \left\{ \frac{1}{2} \|\mathbf{y} - \mathbf{Ax}\|_2^2 + \lambda \|\mathbf{x}\|_1 \right\} \quad (\lambda: \text{正則化係数})$$

・ISTA

$$\begin{aligned} \mathbf{r}^{(t)} &= \mathbf{x}^{(t)} - \alpha \mathbf{A}^\top (\mathbf{Ax}^{(t)} - \mathbf{y}) \rightarrow \text{勾配降下ステップ} \\ \mathbf{x}^{(t+1)} &= S_{\lambda\alpha}(\mathbf{r}^{(t)}) \rightarrow \text{縮小ステップ} \end{aligned} \quad (\alpha: \text{実定数})$$

$S_{\lambda\alpha}(\cdot)$: ソフト閾値関数 (l_1 ノルムの近接写像)



ISTAは最適化問題の解を与える

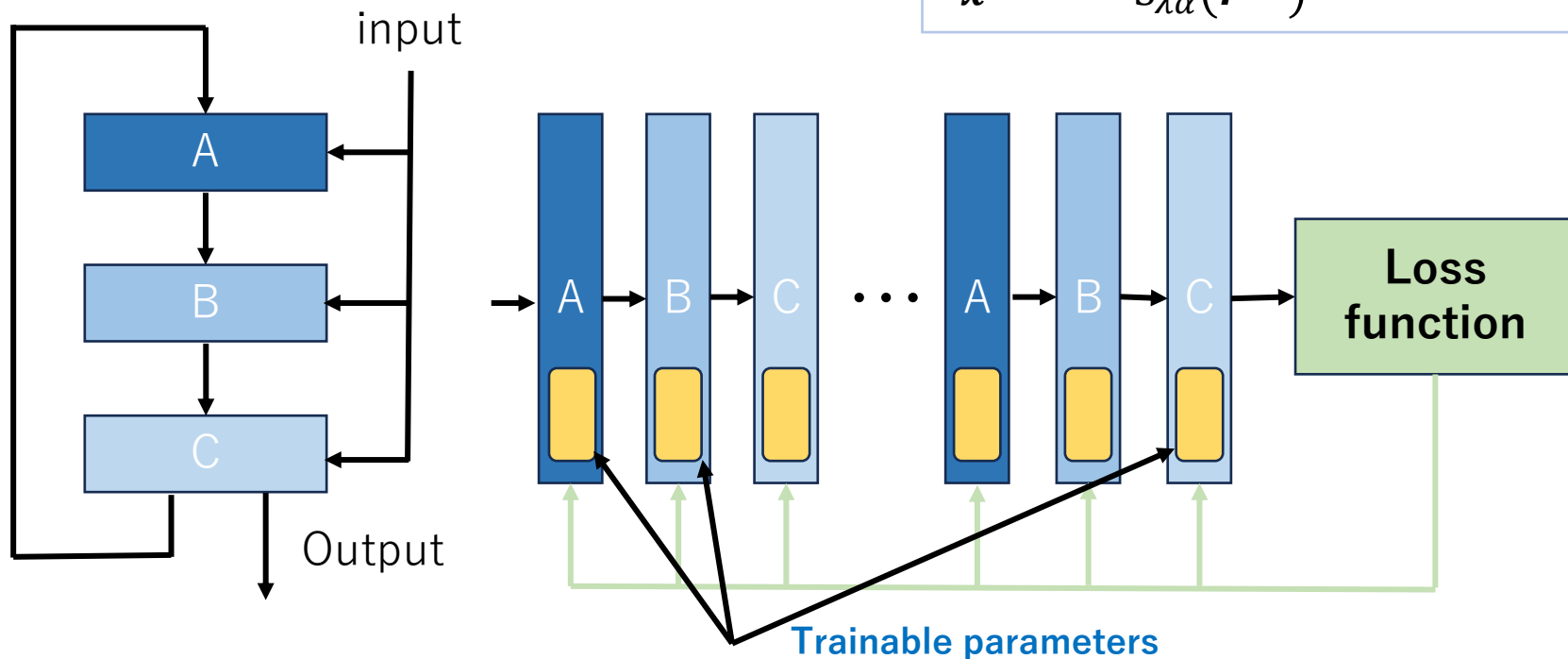
研究背景 | 深層展開[3]

5

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を適切に学習

➡ 推定精度や収束速度が向上

$$\begin{aligned} \mathbf{r}^{(t)} &= \mathbf{x}^{(t)} - \alpha^{(t)} \mathbf{A}^\top (\mathbf{A} \mathbf{x}^{(t)} - \mathbf{y}) \\ \mathbf{x}^{(t+1)} &= S_{\lambda \alpha}(\mathbf{r}^{(t)}) \end{aligned}$$



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

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

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

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

$$\begin{aligned} f(\mathbf{x}) &= \mathbf{x} - \alpha^{(t)} \mathbf{A}^\top (\mathbf{A}\mathbf{x} - \mathbf{y}) : \text{勾配降下ステップ} \\ g(\mathbf{x}) &= S_{\lambda\alpha^{(t)}}(\mathbf{x}) : \text{縮小ステップ} \end{aligned}$$

$$\mathbf{w}_{x,f}^{(t)} = \frac{\exp(\beta_1^{(t)})}{\exp(\beta_1^{(t)}) + \exp(\beta_2^{(t)})}, \mathbf{w}_{x,g}^{(t)} = \frac{\exp(\beta_2^{(t)})}{\exp(\beta_1^{(t)}) + \exp(\beta_2^{(t)})}$$

$$\mathbf{w}_{r,x}^{(t)} = \frac{\exp(\gamma_1^{(t)})}{\exp(\gamma_1^{(t)}) + \exp(\gamma_2^{(t)})}, \mathbf{w}_{r,g}^{(t)} = \frac{\exp(\gamma_2^{(t)})}{\exp(\gamma_1^{(t)}) + \exp(\gamma_2^{(t)})}$$

ソフトマックス型の重み

研究目的 | AS-ISTAの重みの二値化

7

AS-ISTA

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

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

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

$$f(\mathbf{x}) = \mathbf{x} - \alpha^{(t)} \mathbf{A}^\top (\mathbf{A}\mathbf{x} - \mathbf{y}) \quad : \text{勾配降下ステップ}$$

$$g(\mathbf{x}) = S_{\lambda\alpha^{(t)}}(\mathbf{x}) \quad : \text{縮小ステップ}$$

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

新たな重みの設定

0と1の値をとることのできる新たな重みを設定

$$\begin{aligned} r^{(t)} &= w_{x,f}^{(t)} \times f(x^{(t)}) + w_{x,g}^{(t)} \times g(x^{(t)}) \\ x^{(t+1)} &= w_{r,x}^{(t)} \times f(r^{(t)}) + w_{r,g}^{(t)} \times g(r^{(t)}) \end{aligned}$$

パターン1

$$\begin{aligned} w_{x,f}^{(t)} &= \beta^{(t)}, & w_{x,g}^{(t)} &= 1 - \beta^{(t)} \\ w_{r,x}^{(t)} &= 1 - \gamma^{(t)}, & w_{r,g}^{(t)} &= \gamma^{(t)} \end{aligned}$$

パターン2

$$\begin{aligned} w_{x,f}^{(t)} &= \frac{\beta_1^{(t)}}{\beta_1^{(t)} + \beta_2^{(t)}}, & w_{x,g}^{(t)} &= \frac{\beta_2^{(t)}}{\beta_1^{(t)} + \beta_2^{(t)}} \\ w_{r,x}^{(t)} &= \frac{\gamma_1^{(t)}}{\gamma_1^{(t)} + \gamma_2^{(t)}}, & w_{r,g}^{(t)} &= \frac{\gamma_2^{(t)}}{\gamma_1^{(t)} + \gamma_2^{(t)}} \end{aligned}$$

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  $T = 1, \dots, T_{max}$  do (インクリメンタル学習)
2:   for  $i = 1, \dots, n$  do
3:      $\alpha \leftarrow Adam(\alpha, \nabla_{\alpha} L_{train}(\alpha, \theta))$  (ステップサイズの更新)
4:   end for
5:   for  $j = 1, \dots, n$  do
6:      $\theta \leftarrow Adam(\theta, \nabla L_{val}(\alpha, \theta))$  (構造パラメータの更新)
7:   end for
8:   for  $k = 1, \dots, T$  do
9:      $\theta_k \leftarrow threshold\_function(\theta_k)$  (構造パラメータの二値化)
10:  end for
11: end for

```

構造パラメータの二値化方法

パターン1

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

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

パターン2

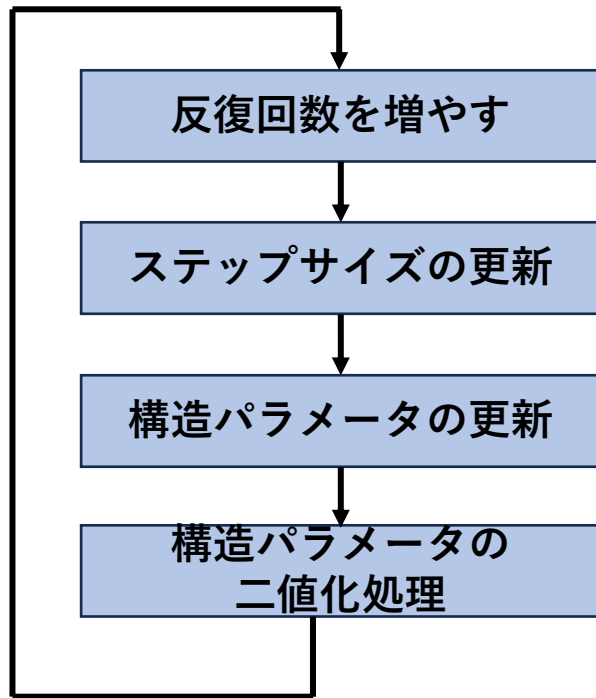
$$\beta_1^{(t)} = \begin{cases} 1 & (\beta_1^{(t)} \geq \beta_2^{(t)}) \\ 0 & (\beta_1^{(t)} < \beta_2^{(t)}) \end{cases}$$

$$\beta_2^{(t)} = \begin{cases} 1 & (\beta_2^{(t)} \geq \beta_1^{(t)}) \\ 0 & (\beta_2^{(t)} < \beta_1^{(t)}) \end{cases}$$

$$\gamma_1^{(t)} = \begin{cases} 1 & (\gamma_1^{(t)} \geq \gamma_2^{(t)}) \\ 0 & (\gamma_1^{(t)} < \gamma_2^{(t)}) \end{cases}$$

$$\gamma_2^{(t)} = \begin{cases} 1 & (\gamma_2^{(t)} \geq \gamma_1^{(t)}) \\ 0 & (\gamma_2^{(t)} < \gamma_1^{(t)}) \end{cases}$$

AS-ISTAの重みの二値化



構造パラメータの二値化方法

パターン1

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

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

パターン2

$$\beta_1^{(t)} = \begin{cases} 1 & (\beta_1^{(t)} \geq \beta_2^{(t)}) \\ 0 & (\beta_1^{(t)} < \beta_2^{(t)}) \end{cases}$$

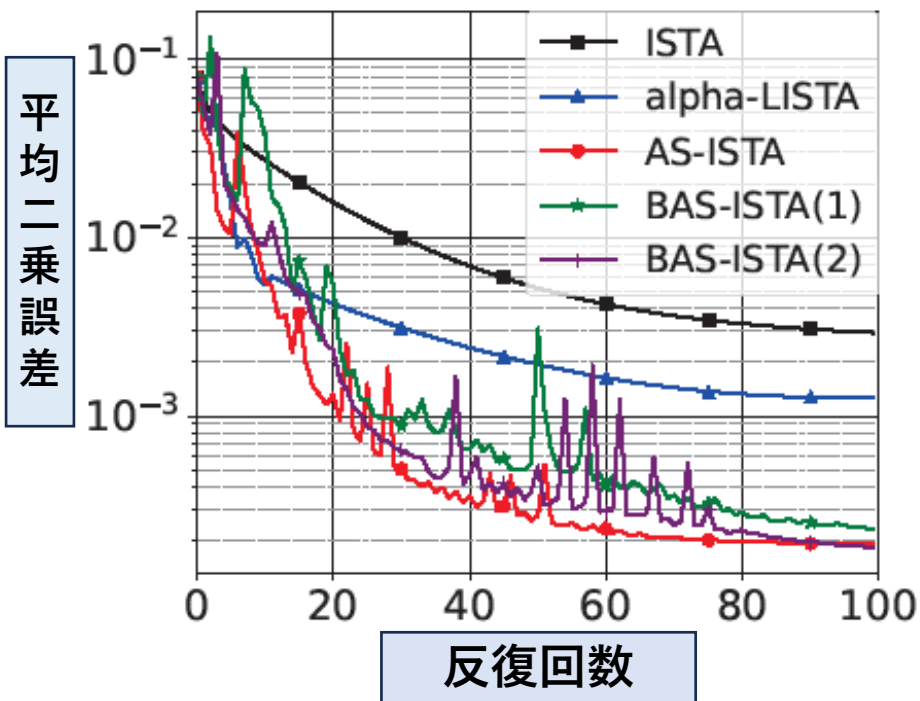
$$\beta_2^{(t)} = \begin{cases} 1 & (\beta_2^{(t)} \geq \beta_1^{(t)}) \\ 0 & (\beta_2^{(t)} < \beta_1^{(t)}) \end{cases}$$

$$\gamma_1^{(t)} = \begin{cases} 1 & (\gamma_1^{(t)} \geq \gamma_2^{(t)}) \\ 0 & (\gamma_1^{(t)} < \gamma_2^{(t)}) \end{cases}$$

$$\gamma_2^{(t)} = \begin{cases} 1 & (\gamma_2^{(t)} \geq \gamma_1^{(t)}) \\ 0 & (\gamma_2^{(t)} < \gamma_1^{(t)}) \end{cases}$$

実験の結果

AS-ISTAとBAS-ISTAのグラフの比較



alpha-LISTA

$$\begin{aligned} \mathbf{r}^{(t)} &= \mathbf{x}^{(t)} - \alpha^{(t)} \mathbf{A}^\top (\mathbf{A} \mathbf{x}^{(t)} - \mathbf{y}) \\ \mathbf{x}^{(t+1)} &= S_{\lambda \alpha}(\mathbf{r}^{(t)}) \end{aligned}$$

BAS-ISTA(1) : パターン 1

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

BAS-ISTA(2) : パターン 2

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

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	1.72×10^{-2}	50.7
AS-ISTA	3.39×10^{-2}	—
BAS-ISTA(1)	1.84×10^{-2}	54.3
BAS-ISTA(2)	1.83×10^{-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