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

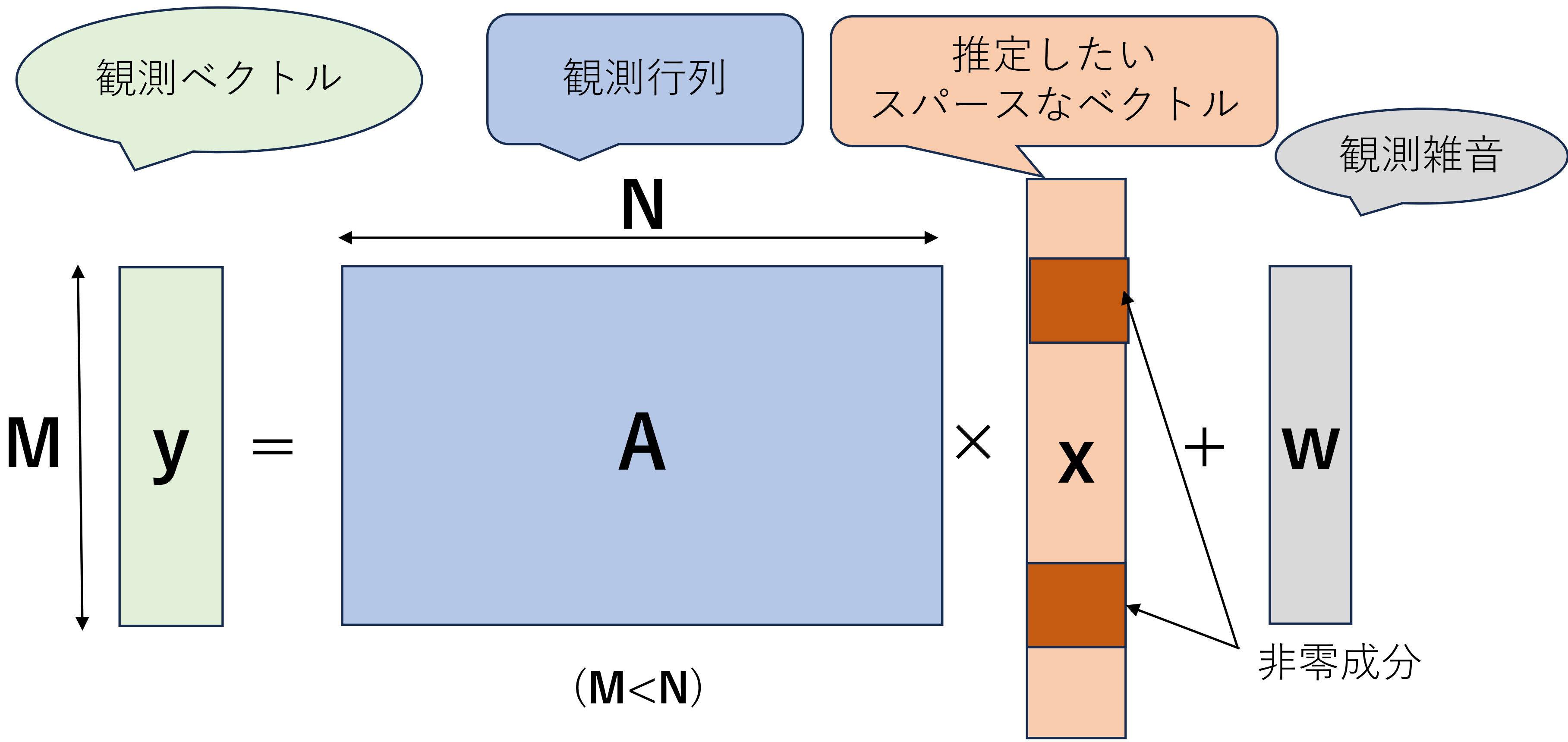
大阪大学　基礎工学部　システム科学科
知能システム学コース
飯國研究室　B4　今井智也

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$\mathbf{y} \in \mathbb{R}^{M}$からスパースな（非零成分が少ない）$\mathbf{x} \in \mathbb{R}^{N}$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

観測との誤差

スパース性

$$\hat{\boldsymbol{x}} = \underset{\boldsymbol{x} \in \mathbb{R}^N}{\operatorname{argmin}} \left\{ \frac{1}{2} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 + \lambda \|\boldsymbol{x}\|_1 \right\}$$

(λ:正則化係数)

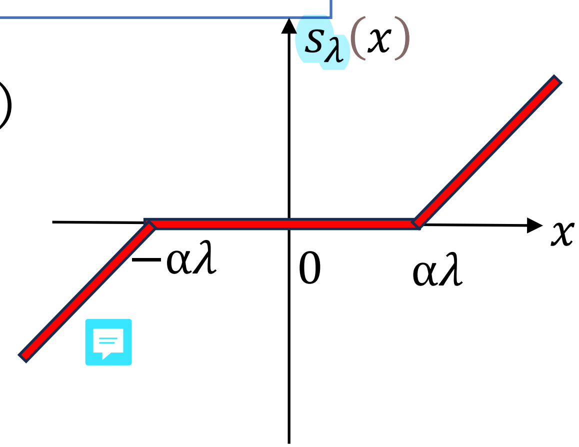
- ISTA

$$\boldsymbol{r}^{(t)} = \boldsymbol{x}^{(t)} - \alpha \boldsymbol{A}^{\top}\left(\boldsymbol{A}\boldsymbol{x}^{(t)} - \boldsymbol{y}\right)$$ →勾配降下ステップ

$$\boldsymbol{x}^{(t+1)} = \mathrm{S}_{\lambda\alpha}\left(\boldsymbol{r}^{(t)}\right)$$ →縮小ステップ

(α:実定数)

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



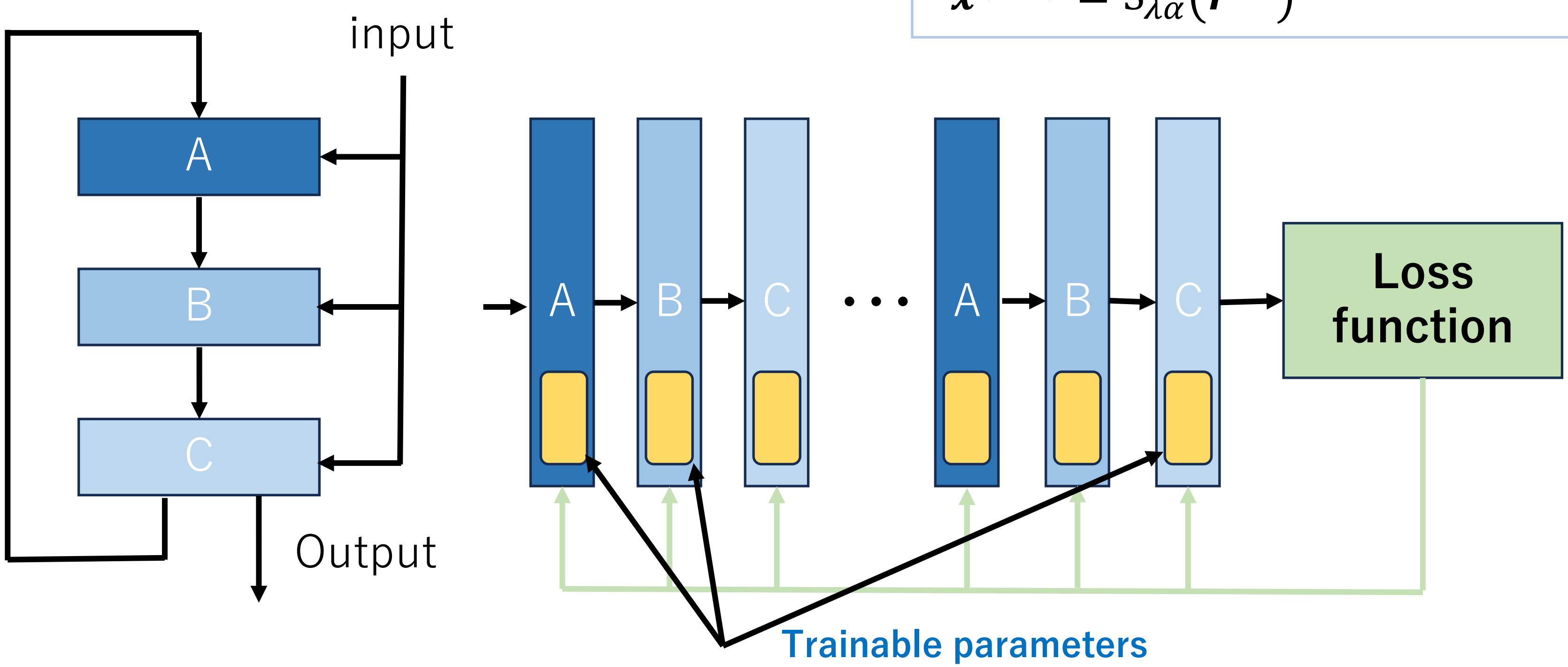
ISTAは最適化問題の解を与える

研究背景 | 深層展開[2]

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

$$\boldsymbol{r}^{(t)} = \boldsymbol{x}^{(t)} - \alpha^{(t)} \boldsymbol{A}^{\top}\left(\boldsymbol{A}\boldsymbol{x}^{(t)} - \boldsymbol{y}\right)$$
$$\boldsymbol{x}^{(t+1)} = S_{\lambda\alpha}\left(\boldsymbol{r}^{(t)}\right)$$



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

[2]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)[3]

- 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

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

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

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

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

$g(\boldsymbol{x}) = \mathrm{S}_{\lambda\alpha^{(t)}}(\boldsymbol{x})$:縮小ステップ

$$w_{x,f}^{(t)} = \frac{\exp\left(\beta_1^{(t)}\right)}{\exp\left(\beta_1^{(t)}\right)+\exp\left(\beta_2^{(t)}\right)}, \quad w_{x,g}^{(t)} = \frac{\exp\left(\beta_2^{(t)}\right)}{\exp\left(\beta_1^{(t)}\right)+\exp\left(\beta_2^{(t)}\right)}$$

$$w_{r,x}^{(t)} = \frac{\exp\left(\gamma_1^{(t)}\right)}{\exp\left(\gamma_1^{(t)}\right)+\exp\left(\gamma_2^{(t)}\right)}, \quad w_{r,g}^{(t)} = \frac{\exp\left(\gamma_2^{(t)}\right)}{\exp\left(\gamma_1^{(t)}\right)+\exp\left(\gamma_2^{(t)}\right)}$$

ソフトマックス型の重み

[3]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｜ASーISTAの重みの二値化

ASーISTA

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

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

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

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

$g(\boldsymbol{x}) = \mathrm{S}_{\lambda\alpha^{(t)}}(\boldsymbol{x})$ ：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~~各式で~~どちらのステップを利用しているかを明確にする~~ことを目指し、~~計算量を削減~~する~~

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

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

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

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

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

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

重みの２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

for $T = 1, \dots, T_{max}$ do　(インクリメンタル学習)

　for $i = 1, \dots, n$ do

　　$\alpha \leftarrow Adam(\alpha, \nabla_\alpha L_{train}(\alpha, \theta))$ (ステップサイズの更新)

　end for

　for $j = 1, \dots, n$ do

　　$\theta \leftarrow Adam(\theta, \nabla L_{val}(\alpha, \theta))$　(構造パラメータの更新)

　end for

　for $k = 1, \dots, T$ do

　　$\theta_k \leftarrow threshold\_function(\theta_k)$　(構造パラメータの二値化)

　end for

end for

構造パラメータの二値化方法

パターン１

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

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

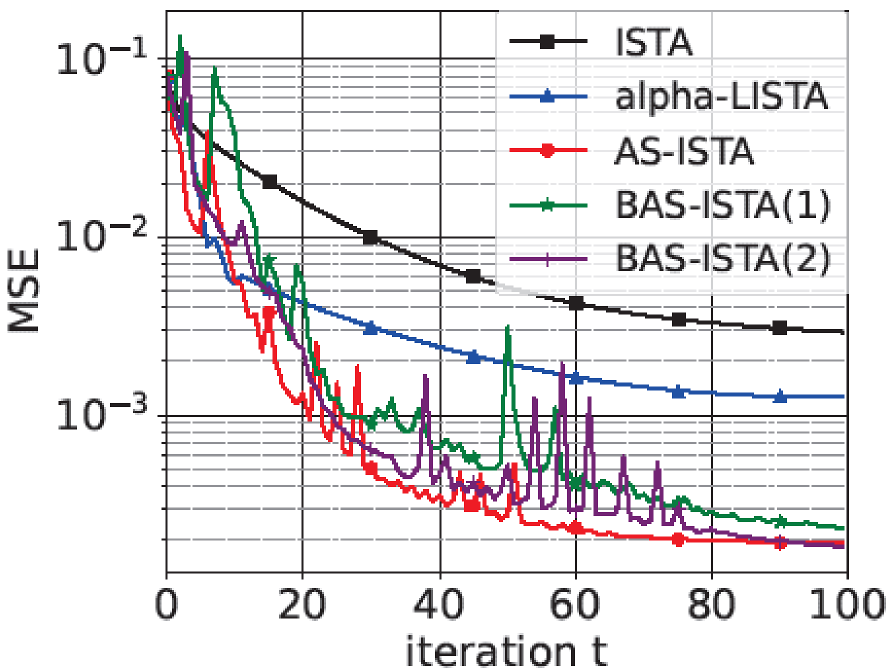
パターン2

$$\beta_1^{(t)} = \begin{cases} 1 & (\beta_1^{(t)} \geq \beta_2^{(t)}) \\ 0 & (\beta_1^{(t)} < \beta_2^{(t)}) \end{cases} \qquad \gamma_1^{(t)} = \begin{cases} 1 & (\gamma_1^{(t)} \geq \gamma_2^{(t)}) \\ 0 & (\gamma_1^{(t)} < \gamma_2^{(t)}) \end{cases}$$

$$\beta_2^{(t)} = \begin{cases} 1 & (\beta_2^{(t)} \geq \beta_1^{(t)}) \\ 0 & (\beta_2^{(t)} < \beta_1^{(t)}) \end{cases} \qquad \gamma_2^{(t)} = \begin{cases} 1 & (\gamma_2^{(t)} \geq \gamma_1^{(t)}) \\ 0 & (\gamma_2^{(t)} < \gamma_1^{(t)}) \end{cases}$$

実験の結果

AS-ISTAとBAS-ISTAのグラフの比較



alpha-LISTA

$$\boldsymbol{r}^{(t)} = \boldsymbol{x}^{(t)} - \alpha^{(t)} \boldsymbol{A}^{\top}\left(\boldsymbol{A}\boldsymbol{x}^{(t)} - \boldsymbol{y}\right)$$

$$\boldsymbol{x}^{(t+1)} = S_{\lambda\alpha}\left(\boldsymbol{r}^{(t)}\right)$$

BAS-ISTA(1)：パターン1

$$\boldsymbol{r}^{(t)} = \beta^{(t)} \times f\left(\boldsymbol{x}^{(t)}\right) + \left(1 - \beta^{(t)}\right) \times g\left(\boldsymbol{x}^{(t)}\right)$$

$$\boldsymbol{x}^{(t+1)} = \left(1 - \gamma^{(t)}\right) \times f\left(\boldsymbol{r}^{(t)}\right) + \gamma^{(t)} \times g\left(\boldsymbol{r}^{(t)}\right)$$

BAS-ISTA(2)：パターン2

$$\boldsymbol{r}^{(t)} = \frac{\beta_1^{(t)}}{\beta_1^{(t)} + \beta_2^{(t)}} \times f\left(\boldsymbol{x}^{(t)}\right) + \frac{\beta_2^{(t)}}{\beta_1^{(t)} + \beta_2^{(t)}} \times g\left(\boldsymbol{x}^{(t)}\right)$$

$$\boldsymbol{x}^{(t+1)} = \frac{\gamma_1^{(t)}}{\gamma_1^{(t)} + \gamma_2^{(t)}} \times f\left(\boldsymbol{r}^{(t)}\right) + \frac{\gamma_2^{(t)}}{\gamma_1^{(t)} + \gamma_2^{(t)}} \times \boldsymbol{g}\left(\boldsymbol{r}^{(t)}\right)$$

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 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

結論

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．ま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