**(504)** 基礎 その他

# コヒーレンスを持つ観測によるスパース推定量の正則化項の大きさの影響

# On the Amplitude of the Regularization Term for Sparse Estimators from Measurements with Large Coherence

井原 みのり † 岩田 一貴 † 三村 和史 †
Minori Ihara † Kazunori Iwata † Kazushi Mimura †
† 広島市立大学　情報科学研究科

## 1　はじめに

未知の信号 $\boldsymbol{x}^0 \in \mathbb{R}^N$ を，行列 $A \in \mathbb{R}^{M\times N}$ を用いて

$$\boldsymbol{y} = A\boldsymbol{x}^0 + \boldsymbol{n} \tag{1}$$

と線形観測する [1, 2]．ここで，$\boldsymbol{n} \in \mathbb{R}^M$ は観測ノイズである．このとき，得られた観測信号 $\boldsymbol{y} \in \mathbb{R}^M$ と既知の観測行列 $A$ だけを用いて，原信号 $\boldsymbol{x}^0 \in \mathbb{R}^N$ を推定する．$M < N$ のとき一般的に解は不定だが，原信号の持つ非零要素の数が少なければ，解を一意に定めることができる．このように，原信号のスパース性を利用して，高次元の信号をより低次元の信号から復元する問題をスパース推定と呼ぶ．スパース推定は，情報通信・医療画像処理・天文学など多くの分野で応用があり注目されている．

本研究では，反復推定法が収束に失敗する場合のスパース推定の性能を解析的に評価することを目的とし，特に観測行列が非零の平均値をもつ場合を取り扱った．観測ノイズやコヒーレンスの大きさが推定性能に与える影響や，影響を低減するために L1 正則化項の大きさをどのように定めれば良いかを調べた．

## 2　準備

スパース推定の代表的な手法のひとつである Lasso を用いて，原信号 $\boldsymbol{x}^0$ を

$$\hat{\boldsymbol{x}} = \operatorname*{argmin}_{\boldsymbol{x}}(\|\boldsymbol{y} - A\boldsymbol{x}\|_2^2 + \lambda\|\boldsymbol{x}\|_1) \tag{2}$$

と推定する．ここで，$\lambda > 0$ は $\ell_1$ 正則化項の大きさである．

行列 $A$ に対するコヒーレンスは次のように定義される．

$$\mu_c(A) := \max_{i<j} |\boldsymbol{a}_i \cdot \boldsymbol{a}_j|. \tag{3}$$

ここで，$\boldsymbol{a}_i$ は行列 $A$ の $i$ 番目の列ベクトルを表す．行列 $A$ の各要素が (4) で示した正規分布に従うとき，コヒーレンスの期待値は，およそ，$\mathbb{E}(\boldsymbol{a}_i{\cdot}\boldsymbol{a}_{j\,(\neq i)}) = h^2/(1{+}h^2)$ となる．観測行列のコヒーレンスを調整できるようにしておき，コヒーレンスの影響も評価する．

## 3　問題設定

観測 $\boldsymbol{y} = A\boldsymbol{x}^0 + \boldsymbol{n}$ において，観測信号 $\boldsymbol{y} \in \mathbb{R}^M$ と観測行列 $A \in \mathbb{R}^{M\times N}$ だけを既知として，未知の原信号 $\boldsymbol{x}^0 \in \mathbb{R}^N$ を推定する問題を考える．このとき，観測ノイズ $\boldsymbol{n} \in \mathbb{R}^{\boldsymbol{M}}$ の各要素は $\mathcal{N}(0, \sigma_0^2)$ の独立同時分布に従うものを仮定する．観測行列の各要素が従う分布はコヒーレンスを調節するための変数 $h$ を導入して次のように設定した．

$$A_{ij} \sim \mathcal{N}\left(\frac{\mu_A}{\sqrt{M}}, \frac{\sigma_A^2}{M}\right). \tag{4}$$

ただし，$\mu_A = h/\sqrt{1+h^2}$，$\sigma_A^2 = 1/(1+h^2)$ である．$h$ が大きいほど，行列 $A$ のコヒーレンスは大きくなる．列ベクトルのノルムは 1 に正規化した．

この線形観測における Lasso の推定性能を評価するため，原信号 $\boldsymbol{x}^0$ と Lasso による推定量 $\hat{\boldsymbol{x}}$ の平均 2 乗誤差 (MSE) $\sigma^2 := \frac{1}{N}\mathbb{E}_{A,\boldsymbol{x}^0,\boldsymbol{n}}(\|\hat{\boldsymbol{x}} - \boldsymbol{x}^0\|_2^2/N)$ を解析した．観測ノイズと観測行列のコヒーレンスの大きさが推定性能に与える影響を調べた．

## 4　解析

推定性能は原信号の各要素が従う分布に依存する．ここでは広く議論されるベルヌイガウス分布 $p_{x^0} = (1-\epsilon)\delta(x^0) + \epsilon(2\pi)^{-1/2}\exp(-{x^0}^2/2)$ を用いた．このとき，原信号の非零要素の割合 $0 \le \epsilon \le 1$ を信号密度と呼ぶ．信号密度 $\epsilon$ と圧縮率 $\delta := M/N$ を与えたときの MSE の値をレプリカ法 [10] を用いて解析的に評価した．

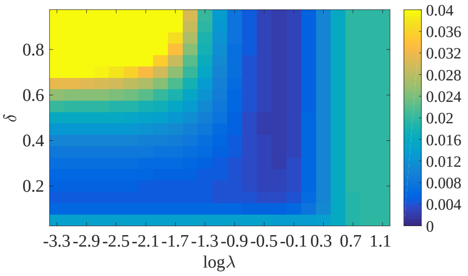
次の分布 $P(\boldsymbol{x})$

$$P(\boldsymbol{x}) = \frac{1}{Z}\exp[-\beta(\|\boldsymbol{y} - A\boldsymbol{x}\|_2^2 - \lambda\|\boldsymbol{x}\|_1)] \tag{5}$$

を考える．ただし，$Z$ は分配関数

$$Z = \int_{\mathbb{R}^N} \mathrm{d}\boldsymbol{x}\exp[-\beta(\|\boldsymbol{y} - A\boldsymbol{x}\|_2^2 - \lambda\|\boldsymbol{x}\|_1)] \tag{6}$$

である．$\beta \to \infty$ の極限において，$P(\boldsymbol{x})$ は (2) の解を台とする一様分布に収束する．レプリカ $\boldsymbol{x}^1, \ldots, \boldsymbol{x}^n$ を導入し，自由エネルギー $f = -\lim_{\beta\to\infty}\lim_{N\to\infty}\frac{1}{\beta N}\mathbb{E}_{A,\boldsymbol{x}^0,\boldsymbol{n}}[\ln Z]$ を評価した．



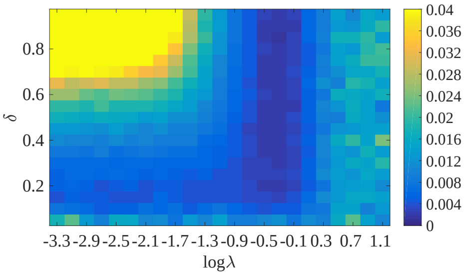




図 1: 平均 2 乗誤差．(上) 理論値，(下) 実験値．$\epsilon = 0.02$，$h = 0.02$，$\sigma_0^2 = 0.02, N = 200$ のとき．

**命題 1** レプリカ対称性仮定と，$N^{-1/2}\sum_{j=1}^{N}\hat{x}_j \approx 0$ とする仮定のもとで，平均 *2* 乗誤差 $\sigma^2$ は，

$$\sigma^2 = \epsilon - 2m + Q \tag{7}$$

となる．ただし，$m$，$Q$ は次の連立方程式の解である．

$$\begin{aligned}\operatorname*{extr}_{\hat{Q},\hat{\chi},\hat{m},Q,\chi,m}\Bigg\{&\delta\frac{\sigma_0^2+(Q-2m+\epsilon)\sigma_A^2\delta^{-1}}{1+2\chi\sigma_A^2\delta^{-1}}\\&+\hat{m}m-\frac{\hat{Q}Q}{2}+\frac{\hat{\chi}\chi}{2}\\&+(1-\epsilon)\int_{\mathbb{R}}\mathrm{D}z\lambda\phi_1\left(\frac{z\sqrt{\hat{\chi}}}{\lambda};\frac{\hat{Q}}{\lambda}\right)\\&+\epsilon\int_{\mathbb{R}}\mathrm{D}z\lambda\phi_1\left(\frac{z\sqrt{\hat{\chi}+\hat{m}^2}}{\lambda};\frac{\hat{Q}}{\lambda}\right)\Bigg\}.\end{aligned} \tag{8}$$

ただし，$\phi_1(h;k) := \mathrm{argmin}_x((k/2)x^2 - hx + |x|)$，$\mathrm{D}z := \mathrm{d}z(2\pi)^{-1/2}\mathrm{e}^{-z^2/2}$ とする． □

## 5 結果

図 1 に，L1 正則化項の大きさ $\lambda$ と圧縮率 $\delta$ を変化させた場合の MSE の値を示す．理論値は，計算機実験より得られた実験値をよく説明していることが確認できる．

L1 ノルム正則化項の大きさ $\lambda$ が適切な値のときは，圧縮率 $\delta$ が大きいほど MSE が小さく，すなわち推定精度がよくなる．しかし，$\lambda$ の値が適切でない場合は，MSE の値は圧縮率に対して非単調に変化するなど，観測数が多くなるほど MSE が大きくなってしまうことがあることがわかる．観測にノイズがない場合は，観測数が多くて正則化項が小さいほどよい．しかし，観測にノイズが含まれる場合においては，観測数が少ない（圧縮率 $\delta$ が小さい）ほうがよい場合があることがわかる．さらに，観測数によらずほぼ同じ値の正則化項の大きさが最適となる領域があることがわかる．

応用上は，交差検証法によって正則化項の大きさが決められるが，観測数を変更したならば，交差検証法をやりなおして決める必要があることを意味する．また，交差検証法のデータの分割の方法によっては，交差検証法によって選ばれる正則化項の大きさに強い影響があることも意味している．

## 6 まとめ

観測ノイズを含む線形観測による Lasso 推定量の平均 2 乗誤差を解析的に評価した．観測にノイズが含まれる場合においては，観測数が多くて正則化項が小さいほどよいという直感的な予想に反して，観測数が少ない（圧縮率 $\delta$ が小さい）ほうがよい場合があることを示した．また，正則化項の大きさ $\lambda$ が大きいほうが良い場合があることを理論的に示した．加えて，そのような性質が，コヒーレンスの大きさによってどのように変化するかを示した．

## 参考文献

[1] D. L. Donoho, *IEEE Trans. Info. Theory*, vol.52, no.4, pp.1289–1306, Apr.2006.

[2] E. J. Candés, J. Romberg, and T. Tao, *IEEE Trans. Info. Theory*, vol.52, no.2, pp.489–509, Feb.2006.

[3] D. L. Donoho, A. Maleki and A. Montanari, *Proc. of the National Academy of Sciences (PNAS)*, vol.106, no.45, pp.18914–18919, Sep.2009.

[4] S. Rangan, P. Schniter, and A. Fletcher, *Proc. of ISIT2014*, pp.236–240, Jul.2014.

[5] S. Rangan, *IEEE Trans. Info. Theory*, vol.62, no.12, pp.7464–7474, Dec.2016.

[6] F. Caltagirone, L. Zdeborová, and F.Krzakala, *Proc. of ISIT2014*, pp.1812-1816, Jun.2014.

[7] J. Ma and L. Ping, *IEEE Access*, vol.5, pp.2020–2033, Jan.2017.

[8] S. Rangan, P. Schniter, and A. K. Fletcher, *Proc. of ISIT2014*, pp.1588–1592, Jul.2017.

[9] K. Takeuchi, *Proc. of ISIT2014*, pp.501–505, Jul.2017.

[10] Y. Kabashima, T. Wadayama, and T. Tanaka, *Journal of Statistical Mechanics: Theory and Experiments*, L09003, Sep.2009.