

圧縮センシング – 疎情報の再構成とそのアルゴリズム –

三村 和史 * (広島市立大学)

Compressed Sensing – Sparse recovery and its algorithms – Kazushi MIMURA * (Hiroshima City University)

概要

疎な信号に対するサンプリング理論である圧縮センシングについて、その概要と関連する話題について解説を行う。

1 はじめに

圧縮センシング (compressed sensing, compressive sampling) は、スパース性 (零成分が多いという性質) を持つ高次元の信号を少ない観測から復元する枠組みである。圧縮センシングは、Nyquist-Shannon の標本化定理の一般化とみなすこともできる。標本化定理によると、原信号の最大周波数成分が W 以下であるとき、 $2W$ 以上の周波数でサンプリングすれば完全再構成が可能である。圧縮センシングでのサンプリングは、原信号を等間隔で標本化するということではなく、線形変換された値をいくつか測定することと一般化して考える。そうすると、原信号がスパース表現を持つとき、そのスパース性を利用することによって、標本化定理で示されるサンプル数よりも少ないサンプルで完全再構成が可能となることが示される。

圧縮センシングについて、最近では、多くの解説記事があり、解析の概要や、応用事例、関連研究に関する話題などが詳しく紹介されている。本稿では、圧縮センシングの文献を読むために必要な用語の定義や、一部分に限られるが関連する話題について紹介したい。

*email: mimura@hiroshima-cu.ac.jp

2 問題設定

未知の N 次元ベクトル $\mathbf{x}_0 \in \mathbb{R}^N$ が, $M < N$ である $M \times N$ 行列 $A \in \mathbb{R}^{M \times N}$ を用いて, 次のように線形変換

$$\mathbf{y} = A\mathbf{x}_0 \quad (1)$$

されているものとする. このとき, より低次元な M 次元ベクトル $\mathbf{y} \in \mathbb{R}^M$ から元のベクトル $\mathbf{x}_0$ を復元する, というのが圧縮センシングの問題である. これは連立1次方程式であるが, 未知数の数が方程式の数より多いため不定となっていて, 一般には解を一意に定めることができない.

しかし, もし原信号の要素の多くが零であったとすると, 解を一意に定めることができるかもしれない. 例えば, 観測信号の次元よりも原信号の非零要素の数が小さいとき, もしも非零要素の位置を知っているならば, 非零要素を全て含むように適当に非零要素の数だけ方程式を選べば解を一意に定めることができる. 非零要素の位置が未知であれば, 問題はそれほど単純ではないのは明らかであろう.

未知のベクトル $\mathbf{x}_0 \in \mathbb{R}^N$ の非零要素が多くとも $K (\leq M)$ 個であると仮定して, 既知のベクトル $\mathbf{y} \in \mathbb{R}^M$ と既知の行列 $A \in \mathbb{R}^{M \times N}$ から (1) を満たすベクトル $\mathbf{x}_0$ を求める問題を, 圧縮センシング問題と呼ぶ. もちろん, 非零要素の位置も未知である. 未知ベクトル $\mathbf{x}_0$ の非零要素の割合を信号密度 (または疎性) $\rho := K/N$ (文献によっては K/M と定義される場合もある) といい, 行列 A の行の数と列の数の比を圧縮率 $\alpha := M/N$ という. 本稿では, 非零要素の値に符号の制約を付けない場合 (問題 $\pm$) のみを扱う.

2.1 記法

まず, 本稿で用いる記法をまとめておく. ベクトルは太文字で表記し, 特に指定しない限りは列ベクトルであるとする. 行列は大文字で表記する. $\|\cdot\|_p$ はノルムであり, $p > 0$ のときは $\|\mathbf{x}\|_p := (\sum_{i=1}^N |x_i|^p)^{1/p}$ (準ノルム), $p = 0$ のときは $\|\mathbf{x}\|_0 := |\text{supp}(\mathbf{x})|$ と定義する. ただし, $\text{supp}(\mathbf{x}) := \{i : x_i \neq 0\}$ はベクトル $\mathbf{x}$ の台である. 以降, p に依らず, 単に ℓ_p ノルムとよぶ.

行列 $A = (\mathbf{a}_1, \dots, \mathbf{a}_N) \in \mathbb{R}^{M \times N}$ の第 i 列ベクトルを $\mathbf{a}_i$ とする. 列添字の集合 $T = \{i_1, \dots, i_k\} \subset \{1, \dots, N\}$ について, すべての $i \in T$ 番目の列ベクトルの作る部分行列を $A_T = (\mathbf{a}_{i_1}, \dots, \mathbf{a}_{i_k}) \in \mathbb{R}^{M \times |T|}$ とする. また, 同様にベクトル $\mathbf{x} \in \mathbb{R}^N$ の第 i 要素を x_i , 添字の集合を $T = \{i_1, \dots, i_k\} \subset \{1, \dots, N\}$ としたとき, すべての $i \in T$ 番目の要素の作る部分ベクトルを $\mathbf{x}_T = (x_{i_1}, \dots, x_{i_k})^\top \in \mathbb{R}^{|T|}$ とする. ここで,

A^T は A の転置を表す。また、集合 T の補集合を $T^c = \{1, \dots, N\} \setminus T$ と書く。 $[i:j]$ で、 i から j までの整数の集合を表す。

ベクトルに含まれる非零要素の数が多くとも K 個であることを、 K -スパースであるという。 K -スパースなベクトルの集合を $\Sigma_K := \{x \in \mathbb{R}^N : \|x\|_0 \leq K\}$ と表記する。 $x, z \in \Sigma_K$ のとき、一般には台が一致するとは限らないので $x + z \in \Sigma_{2K}$ となる。原信号 x 自身が疎でなくても既知の基底 $\Psi \in \mathbb{R}^{N \times N}$ を用いて $x = \Psi c$ となるような K -スパースなベクトル $c \in \Sigma_K$ が存在すれば、 x 自身も K -スパースと呼ばれることがある。図1は、 $N = 3$ の場合の2-スパースなベクトルの集合 Σ_2 を表す。

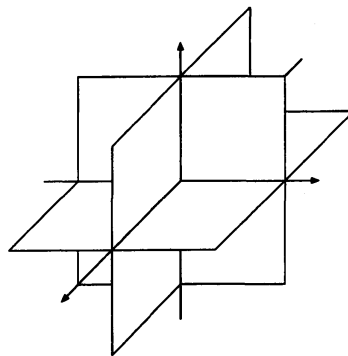


図 1: スパースなベクトルの集合。 $\Sigma_2 \subset \mathbb{R}^3$ 。

3 復元手法

3.1 概要

(1) の観測信号 y から、最もスパースな原信号 x_0 を求めるための素朴な方法は、線形制約を満たす解のうち最も非零要素数の少ない解を探す方法が考えられる。これは、全ての $|T| \leq M$ となる列添字の集合 $T \subset \{1, \dots, N\}$ について、実際に、 $y = A_T x_T$ の連立1次方程式を解くとよい。すべての T の組について繰り返すと、最もスパースな解を求めることができる。しかし、計算量は一般には N について指数的に増えてしまうため現実的な解法ではない。

復元信号 $\hat{x}$ を回帰問題 $\hat{x} = \operatorname{argmin}_x \|Ax - y\|_2^2$ によって求めること考えると、 ℓ_2 ノルムが最小となる解は一般逆行列 $A^+ = A^T(AA^T)^{-1}$ を用いて $\hat{x} = A^+ y$ となる。図2に、 $M = 100$, $N = 200$ として、 N 次元単位超球上のベクトルを一様に独立に選んだある観測行列 (辞書とも呼ばれる) A を用いて観測信号 $y = Ax_0$ を得た場合の数

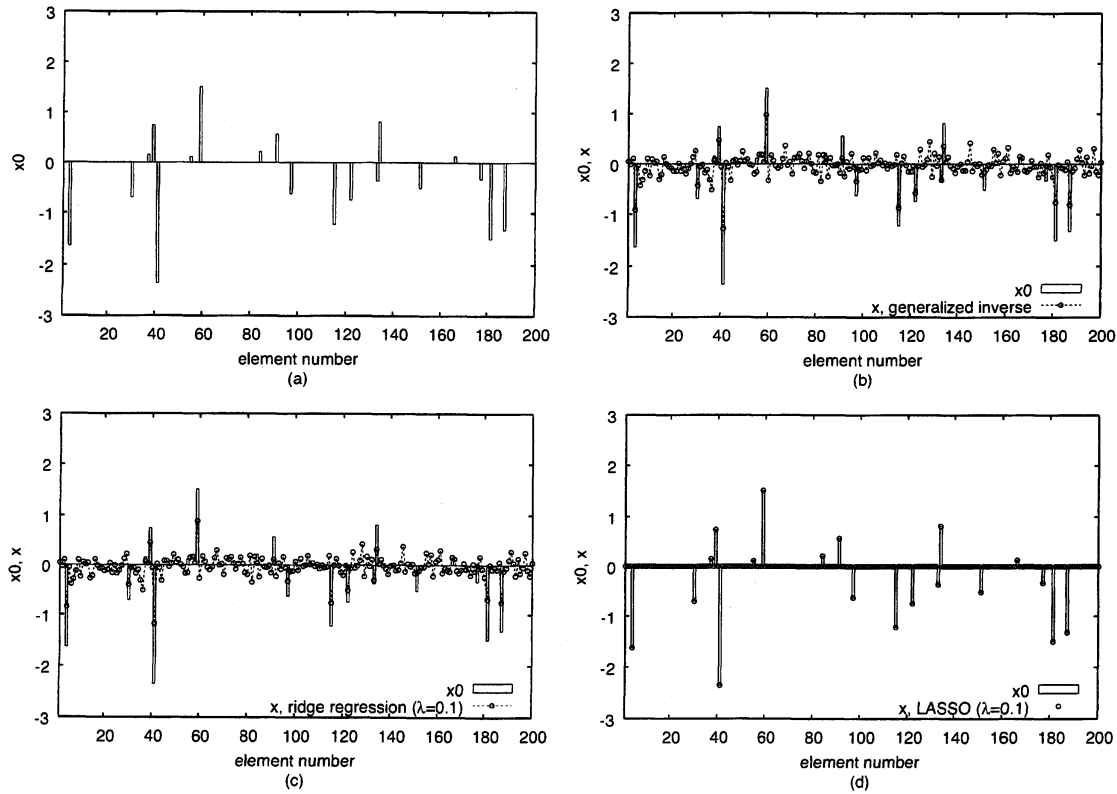


図 2: 再構成. (a) 原信号. (b) 一般逆行列. (c) ridge 回帰. (d) lasso.

値例を示す. 図 2(a) は, スパースな原信号 x_0 である. 最小 ℓ_2 ノルム解は図 2(b) のようになり, 原信号が零である要素の多くが非零になってしまう様子が確認できる.

縮小推定のひとつである Tikhonov 正則化を導入した回帰問題 (ridge 回帰) $\hat{x} = \operatorname{argmin}_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \lambda \|\mathbf{x}\|_2^2$ を考えて, より ℓ_2 ノルムが小さい解を近似的に探そうとしても状況は改善しない. パラメータ $\lambda \geq 0$ は正則化の大きさを表し, 大きいほど ℓ_2 ノルムが小さい解が得られる. このとき, 復元信号は $\hat{x} = (\mathbf{A}^\top \mathbf{A} + \lambda \mathbf{I})^{-1} \mathbf{A}^\top \mathbf{y}$ となる. パラメータが $\lambda = 0.1$ のときの結果は図 2(c) のようになり, この場合も先ほどと同様にスパースな原信号を復元できないことが例から確認できる.

一方で, 正則化項を ℓ_1 ノルムにした回帰問題 (lasso) $\hat{x} = \operatorname{argmin}_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \lambda \|\mathbf{x}\|_1$ では, 疎な解が得られることはよく知られている [58]. パラメータが $\lambda = 0.1$ のときの結果は図 2(d) のようになる. この例では, 原信号が正確に再現されている. 非零要素数をコストにしなくても, ℓ_1 ノルムによる正則化項のようにスパースなとき値が小さいことが強調されるようなコスト関数であれば, スパースな原信号の復元の基準として利用可能であることが示唆される.

3.2 ℓ_p 再構成

前述のように、観測信号 $\mathbf{y}$ と観測行列 A を用いてスパースな原信号 $\mathbf{x}_0$ の再構成を行うには、線形制約 $\mathbf{y} = A\mathbf{x}$ を満たして、スパースなベクトルを探す方法が考えられる。このためには、非零要素数 (ℓ_0 ノルム) を直接小さくしなくとも、lasso のように ℓ_1 ノルムをコスト関数としてもよさそうであった。一般化すれば、 ℓ_p ノルム最小化による推定法が考えられる。

図3は、単位 ℓ_p 球である。 $0 \leq p \leq 1$ のとき、等ノルム面は軸上で「尖った」形をしている。図4は、 $K = 1, M = 1, N = 2$ の場合の ℓ_p 再構成の様子を表している。黒丸は原信号を、白丸は推定ベクトルを表す。この場合は、線形制約は1次関数となっていて、等 ℓ_p ノルム面が1次関数に接する点が、線形制約を満たす最も ℓ_p ノルムの小さい推定ベクトルとなる。図4(a), (b) は、それぞれ $p = \frac{1}{2}, p = 1$ の ℓ_p 再構成が正しい結果を与える場合を表す。等ノルム面が軸上で尖っているので、スパースなベクトルを推定できる。図4(c) は、それぞれ $p = 2$ の場合を示す。 ℓ_2 ノルム最小化では、スパースな解の部分で等ノルム面が線形制約に接することは、特殊な観測行列でなければ期待できず、推定ベクトルが正しい解であることは望めない。

もう少し詳しく様子を見るために、原信号が3次元の場合を考えよう。図5は $K = 1, M = 1, N = 3$ の場合の ℓ_p 再構成の様子を表す。図5(a), (b) は、線形制約の例である。線形制約を満たすベクトルの集合を $\mathcal{L}(\mathbf{y}) \triangleq \{\mathbf{x} \in \mathbb{R}^N : \mathbf{y} = A\mathbf{x}\}$ とすると、 $K < M$ であれば $\mathcal{L}(\mathbf{y}) \cap \Sigma_K = \{\mathbf{x}_0\}$ が期待できる。例えば、ランダムな観測行列を考える場合などは、図5(a)のように、最もスパースなベクトルが原信号と一般には一致する。図5(b)のように、 $\mathcal{L}(\mathbf{y}) \subset \Sigma_K$ の場合などは、制約条件を満たす最もスパースなベクトルが複数個あることもある。また、原信号の次元が $N = 2$ の例では、 $p \leq 1$ であれば ℓ_p 再構成に差がないように見えるが、勿論高次元では p によって様子は異なる。図5(c), (d) は、それぞれ $p = 1, p = \frac{1}{2}$ の ℓ_p 再構成の例を示す。 $p = 1$ の場合は正しい解を推定できない場合でも、 $p = \frac{1}{2}$ の場合は等 ℓ_p ノルム面が「窪んでいる」ため正しく解が推定できる場合があることがわかる。

4 解析

4.1 表現モデルと観測モデル

画像や音声といった自然にある信号は、適切な基底で展開すると簡潔な表現を持つことが多い。このような、展開して得られる係数が、多くの要素で小さい値となり、比較的少ない要素のみが大きい値となるときは、大きい値の要素だけでも、元

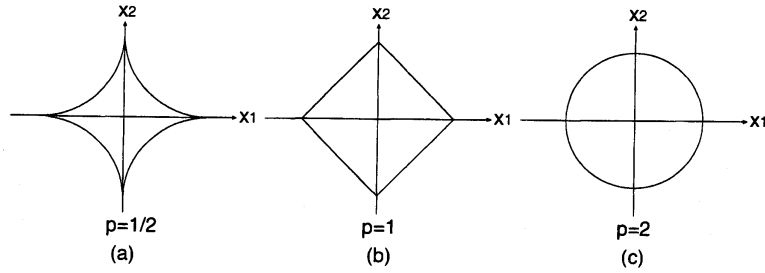


図 3: 単位 ℓ_p 球. $N = 2$, $|x_1|^p + |x_2|^p = 1$ (単位 $\ell_{1/2}$ 球は概形).

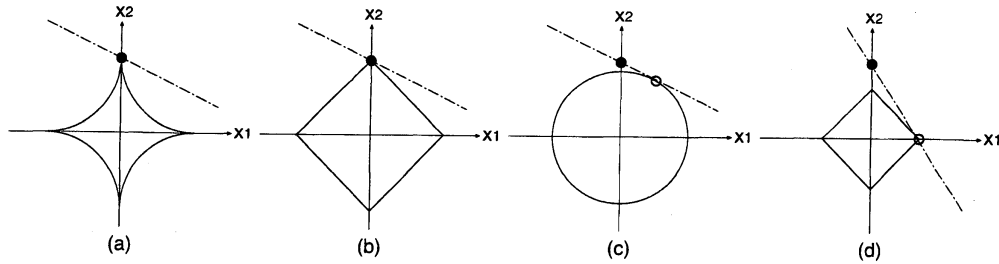


図 4: ℓ_p 再構成. $K = 1$, $M = 1$, $N = 2$ の場合.

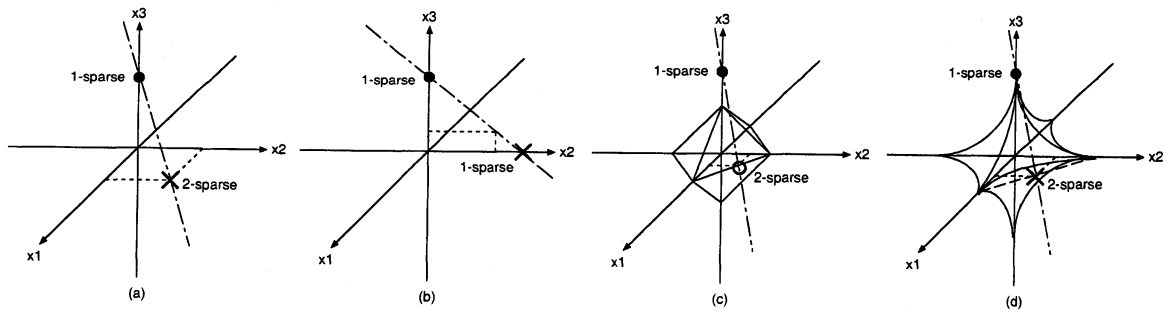


図 5: ℓ_p 再構成. $K = 1$, $M = 1$, $N = 3$ 場合.

の情報の特徴を捉えることができる. 疎な原信号は, 離散信号 $\mathbf{f} \in \mathbb{R}^N$ になんらかの変換をしたとき疎になる信号として扱うことができる. すなわち, 離散信号 $\mathbf{f}$ を既知の表現基底 (例えば DCT 基底, wavelet 基底など) $\Psi = (\psi_1, \dots, \psi_N) \in \mathbb{R}^{N \times N}$ によって, $\mathbf{f} = \sum_{n=1}^N x_{0,n} \psi_n$ のように展開して, 係数の系列 $\mathbf{x}_0 = (x_{0,1}, \dots, x_{0,N})^\top$ を原信号とする. この関係

$$\mathbf{f} = \Psi \mathbf{x}_0 \quad (2)$$

は, 表現モデルと呼ばれる.

観測信号 $\mathbf{y} = (y_1, \dots, y_M)^\top$ は, 既知の観測基底 $\Phi = (\phi_1, \dots, \phi_M) \in \mathbb{R}^{N \times M}$ に

よって、観測信号の要素を $y_m = \langle \phi_m, \mathbf{f} \rangle = \sum_{n=1}^N x_{0,n} \langle \phi_m, \psi_n \rangle$ として得るものとする。この関係

$$\mathbf{y} = \Phi^\top \mathbf{f} \quad (3)$$

は観測モデルと呼ばれる。(2) と (3) では、観測信号 $\mathbf{y}$ は、原信号 $\mathbf{x}_0$ の観測行列 $A = \Phi^\top \Psi$ による線形変換 $\mathbf{y} = A\mathbf{x}_0$ とみなすことができる。圧縮センシングの応用では、原信号 $\mathbf{x}_0$ が疎になるような表現基底 Ψ の存在を仮定することになる。

4.2 零空間特性

正しい解が再構成されることを完全再構成という。完全再構成が可能かどうかは、観測行列 A の性質によって決まる。完全再構成ができるための条件は、本質的には等価と考えられるいろいろな方法で調べられている。そのうちのひとつは零空間特性を用いた方法である。観測行列 $A \in \mathbb{R}^{M \times N}$ の零空間とは、

$$\ker(A) := \{\mathbf{x} \in \mathbb{R}^N : A\mathbf{x} = \mathbf{0}\} \quad (4)$$

をいう。完全再構成のためには、まず条件 $\mathbf{y} = A\mathbf{x}$ を満たすような疎ベクトル $\mathbf{x}$ が多くあってはいけなからう。解の一意性を示すことは、多くの方法によってなされているが、そのうちのひとつは次のスパークによって示すものである。

定義 1. (スパーク) 行列 A の一次従属な列ベクトルの数の最小値

$$\text{spark}(A) = \min_{\mathbf{z} \in \ker(A) \setminus \{\mathbf{0}\}} \|\mathbf{z}\|_0 \quad (5)$$

を行列 A の spark という [25]. □

階数とは、 $\text{spark}(A) \leq \text{rank}(A) + 1$ の関係にある。ただし、全ての列ベクトルが一次独立な行列 $A \in \mathbb{R}^{L \times N} (L \geq N)$ のとき、 $\text{spark}(A) = L + 1$ とする。また、任意の $\text{spark}(A) - 1$ 個の列ベクトルの組は、一次独立となっている。この spark を用いて、次のように解の一意性について述べるができる。

定理 1. (一意性) $\text{spark}(A) > 2K$ ならば、 $\mathbf{y} = A\mathbf{x}$ を満たす K -スパースなベクトル $\mathbf{x} \in \Sigma_K$ が多くともひとつ存在する [25]. □

証明は、付録 A.1 を参照されたい。このため、 $\text{spark}(A)$ が大きい観測行列 A ほど再構成に有利となる。この定理により、原信号が K -スパースなら一意な解が存在することになる。このように、一意性を簡潔に表現できるのは、spark を定義したこと

による恩恵である。しかし、spark を求めるための計算量は、行列 $A \in \mathbb{R}^{M \times N}$ の列ベクトルの全ての組について、その組が一次従属であるかどうかを判定する必要があるため、判定の回数が $O(2^N)$ と列の数に対して指数的に増加してしまう。この定理によって、疎な解の完全再構成は可能である、といえる。

では、 ℓ_p 最小化原理に基づいて疎な解を求めたとき、完全再構成ができる場合とはどのような場合であろうか。 ℓ_p ノルムが小さいベクトルを探す、といった疎な解の探し方に依存した完全再構成の条件を議論する方法のひとつは、次の零空間特性によるものである。

定義 2. (零空間特性) 任意の $z \in \ker(A) \setminus \{0\}$ と、 $|T| \leq K$ となる任意の添字集合 T について、

$$\|z_T\|_p^p < \|z_{T^c}\|_p^p \quad (6)$$

となるとき、行列 A は、 ℓ_p ノルムで K 次の零空間特性 (null space property: NSP) を持つという [32]. $\square$

この零空間特性を用いて、完全再構成が可能な条件を次のように述べることができる。

定理 2. (零空間特性を用いた ℓ_p 復元定理) 任意の $0 \leq p \leq 1$ について、観測行列 $A \in \mathbb{R}^{M \times N}$ が、 ℓ_p ノルムで K 次の零空間特性を持つとき、

$$\hat{x} = \underset{x \in \mathbb{R}^N}{\operatorname{argmin}} \|x\|_p \quad \text{subj. to } y = Ax \quad (7)$$

によって、任意の K -スパースなベクトルが一意的に再構成できる [32]. $\square$

証明は、付録 A.2 を参照されたい。この定理では、原信号 x_0 が疎であることを仮定した議論となっている。もし、原信号が疎でないときに、同じ再構成を行うとどうなるであろうか。原信号が疎でないときは、不定の連立 1 次方程式を解く問題と同じであり完全再構成は期待できない。このため、再構成されたベクトルと原信号との誤差の評価が重要になる。このような再構成の誤差は、次の頑健零空間特性を用いて評価されている [20].

定義 3. (頑健零空間特性) 任意の $z \in \ker(A) \setminus \{0\}$ と、 $|T| \leq K$ となる任意の添字集合 T と、ある定数 $\sigma < 1$ について、

$$\|z_T\|_p^p < \sigma \|z_{T^c}\|_p^p \quad (8)$$

となるとき、行列 A は、 ℓ_p ノルムで K 次の係数 σ の頑健零空間特性 (robust NSP) を持つという [20]. $\square$

頑健零空間特性は零空間特性にさらに条件が付加されたもので、行列 A が頑健零空間特性も持つならば零空間特性を持つ。この頑健零空間特性を用いて、 ℓ_p 復元を行ってときの再構成の誤差を次のように述べることができる。

定理 3. (ℓ_p 復元の再構成誤差) 任意の $0 \leq p \leq 1$ について、 ℓ_p ノルムで K 次の係数 σ の頑健零空間特性を持つ観測行列 $A \in \mathbb{R}^{M \times N}$ を用いて、

$$\hat{x} = \underset{x \in \mathbb{R}^N}{\operatorname{argmin}} \|x\|_p \quad \text{subj. to } y = Ax \quad (9)$$

により、 K -スパースなベクトルを再構成するとき、再構成されたベクトル $\hat{x}$ の原信号 x_0 との誤差 $\|\hat{x} - x_0\|_p^p$ は、

$$\|\hat{x} - x_0\|_p^p < 2 \frac{1 + \sigma}{1 - \sigma} \|x_0 - x_0^{(K)}\|_p^p \quad (10)$$

によって上から押えられる。ただし、 $x_0^{(K)}$ は、原信号 x_0 の要素のうち絶対値の大きい K 個の要素はそのままとし、残りの要素は 0 としたベクトルである [20]。□

証明は、付録 A.3 を参照されたい。 $\|(x_0)_T\|_p^p$ は、 K -スパースなベクトルで原信号 x_0 を近似したときの誤差である。この誤差が最小になる添字の集合 T は、単に原信号 x_0 の要素のうち絶対値の大きい要素の添字を、大きいほうから K 個選んだものになる。 ℓ_p 最小化に基づく再構成を行っても、その誤差は K -スパースなベクトルでの最も小さい誤差の定数倍で押さえられている。

このように、零空間を用いて定義されている行列の spark と零空間特性は、これらを経由して、解の一意性、完全再構成が可能な条件、完全再構成ができないときの誤差などが評価できるため重要である。しかし、与えられた観測行列 A の spark を求めたり、零空間特性を持つかどうかの評価は一般には難しく、効率的な評価方法の検討や、零空間特性を満たすような観測行列の構成法の検討などが必要である。

4.3 制限等長性

完全再構成ができるための条件を調べる別の接近法は、制限等長性による議論である。原信号 $x_0 \in \Sigma_K$ を、 K -スパースなベクトルとする。ところで、 $Ax = 0$ となるような K' -スパースなベクトルがあるとすると、 $y = Ax_0 = A(x_0 + x)$ が成り立つので、 $x_0 + x$ は線形制約を満たす。この $x_0 + x$ も、 $(K + K')$ -スパースなベクトルとなっている。このような疎ベクトル x が多いと、線形制約を満たす疎ベクトルが多くなって、再構成に良くない影響を与えるであろう。もし、 x と線形変換 Ax の長さがあまり変わらない（等長性）のであれば、疎ベクトル x の線形変換が $Ax = 0$

にならずに済むので、再構成がうまくできるかもしれない [3]. 制限等長性は、このような疎ベクトルに制限した等長性を表し、次のように定義されている.

定義 4. (制限等長性) 任意の K -スパースなベクトル $\mathbf{c} \in \Sigma_K$ について不等式

$$(1 - \delta) \|\mathbf{c}\|_2^2 \leq \|A\mathbf{c}\|_2^2 \leq (1 + \delta) \|\mathbf{c}\|_2^2 \quad (11)$$

を満たす定数 $\delta \in (0, 1)$ が存在するとき、行列 $A \in \mathbb{R}^{M \times N}$ は、 K 次の制限等長性 (restricted isometry property: RIP) を持つという. また、このときの定数 δ の最小値 δ_K を K 次の RIP 定数という [11]. $\square$

RIP 定数は、制限等長定数 (restricted isometry constant: RIC) や等長性定数とも呼ばれる. RIP 定数 δ_K が小さいほど再構成に有利となる. 例えば、 $A \in \mathbb{R}^{N \times N}$ が直交行列ならば、任意の $\mathbf{x} \in \mathbb{R}^N$ に対して、 $\|A\mathbf{x}\|_2^2 = \mathbf{x}^\top A^\top A \mathbf{x} = \|\mathbf{x}\|_2^2$ となり行列 A は等長変換なので、RIP 定数は K によらず $\delta_K = 0$ である. RIP 定数 δ_K が小さいということは、任意の K 個の列ベクトルの組みが正規直交系に近い振る舞いをすることを意味する. 制限等長性を用いて、次のように完全再構成の条件 [11] が議論されるなどしている.

定理 4. (制限等長性を用いた ℓ_0 復元定理) 観測行列 A が $2K$ 次の制限等長性を持つと仮定すると、 $\mathbf{y} = A\mathbf{x}_0$ となる任意の K -スパースなベクトル $\mathbf{x}_0 \in \Sigma_K$ について、

$$\hat{\mathbf{x}} = \underset{\mathbf{x} \in \mathbb{R}^N}{\operatorname{argmin}} \|\mathbf{x}\|_0 \quad \text{subj. to } \mathbf{y} = A\mathbf{x} \quad (12)$$

は唯一の解 $\hat{\mathbf{x}}$ を持ち、 $\hat{\mathbf{x}} = \mathbf{x}_0$ となる [14]. $\square$

証明は、文献 [14] または付録 A.4 を参照されたい. ここに現れる条件は、 ℓ_0 最小化原理に基づいて疎な解を求めたとき、完全再構成ができるための十分条件となっている.

定理 5. (制限等長性を用いた ℓ_1 復元定理) 観測行列 A が $2K$ 次の制限等長性を持ち、 $\delta_{2K} < \sqrt{2} - 1$ と仮定すると、 $\mathbf{y} = A\mathbf{x}_0$ となる任意の K -スパースなベクトル $\mathbf{x}_0 \in \Sigma_K$ について、

$$\hat{\mathbf{x}} = \underset{\mathbf{x} \in \mathbb{R}^N}{\operatorname{argmin}} \|\mathbf{x}\|_1 \quad \text{subj. to } \mathbf{y} = A\mathbf{x} \quad (13)$$

は唯一の解 $\hat{\mathbf{x}}$ を持ち、 $\hat{\mathbf{x}} = \mathbf{x}_0$ となる [14]. $\square$

証明は、文献 [14] を参照されたい。また、行列 A が $2K$ 次の制限等長性を持ち、かつ $\delta_{2K} < \sqrt{2}-1$ であるとき、行列 A は ℓ_1 ノルムで K 次の係数 $\sigma = \sqrt{2}\delta_{2K}(1-\delta_{2K})^{-1}$ の頑健零空間特性を持つことが示されている [14]。このため、制限等長性は、頑健零空間特性の十分条件となっていることがわかる。

制限等長性を用いた ℓ_1 復元定理を利用するためには、与えられた行列 A の RIP 定数を評価しなければならないので、その評価方法について考えてみよう。実対称行列 $B \in \mathbb{R}^{n \times n}$ の最大固有値 $\gamma_{\max}(B)$ と最小固有値 $\gamma_{\min}(B)$ は、Rayleigh 商の最大値、最小値で表せて $\gamma_{\max}(B) = \max_{\mathbf{x} \neq 0} \mathbf{x}^\top B \mathbf{x} / (\mathbf{x}^\top \mathbf{x})$, $\gamma_{\min}(B) = \min_{\mathbf{x} \neq 0} \mathbf{x}^\top B \mathbf{x} / (\mathbf{x}^\top \mathbf{x})$ となる。これより、任意の $\mathbf{x} \in \mathbb{R}^N \setminus \{0\}$ について、 $\gamma_{\min}(B) \mathbf{x}^\top \mathbf{x} \leq \mathbf{x}^\top B \mathbf{x} \leq \gamma_{\max}(B) \mathbf{x}^\top \mathbf{x}$ を得る。また、 $A \in \mathbb{R}^{M \times N}$, $T \subset \{1, \dots, N\}$, $\mathbf{c} \in \mathbb{R}^{|T|}$ とすると、 $\|A_T \mathbf{c}\|_2^2 = \mathbf{c}^\top (A_T^\top A_T) \mathbf{c}$ である。ここで、 $A_T^\top A_T$ は実対称行列なので、さきほどの不等式が成り立って、 $\gamma_{\min}(A_T^\top A_T) \mathbf{c}^\top \mathbf{c} \leq \mathbf{c}^\top A_T^\top A_T \mathbf{c} \leq \gamma_{\max}(A_T^\top A_T) \mathbf{c}^\top \mathbf{c}$ となるため、

$$\gamma_{\min}(A_T^\top A_T) \|\mathbf{c}\|_2^2 \leq \|A_T \mathbf{c}\|_2^2 \leq \gamma_{\max}(A_T^\top A_T) \|\mathbf{c}\|_2^2 \quad (14)$$

を得る。これより、精密な K 次の RIP 定数 δ_K の評価には、全ての $|T| < K$ となる $T \subset [1:N]$ について、行列 $A_T^\top A_T$ の最大、最小固有値（または A_T の最大、最小特異値）を評価すればよいことがわかる。しかし、計算量が指数的に増えるので実際的ではない。

このため、行列のアンサンブル（確率の与えられた行列の集合）を考え、そのアンサンブルの RIP 定数が評価されている。例えば、行列 $A \in \mathbb{R}^{M \times N}$ の各要素を独立に正規分布 $\mathcal{N}(0, 1/M)$ に従って定めるガウス行列アンサンブルについて考える。ガウス行列アンサンブルでは、Marchenko-Pastur 則 [41] から $\gamma_{\min}(A_T^\top A_T)$, $\gamma_{\max}(A_T^\top A_T)$ が評価される。より詳しい固有値の評価 [38] を用いて、高い確率で K 次の制限等長性定数 δ_K の上界が、

$$\delta_K < [1 + f(\rho)]^2 - 1$$

となることが示されている [11]。ただし、 $f(\rho) := \sqrt{\rho/\delta} + \sqrt{2H(\rho)/\delta}$ とし、 $H(\rho) := -\rho \log_e \rho - (1-\rho) \log_e (1-\rho)$ を 2 値エントロピー関数とする。この上界を数値的に評価すると、上界自身が ℓ_1 再現定理の条件 ($\delta_{2K} < \sqrt{2}-1$ で K -スパース) を満たす領域があることが確認できる。

4.4 インコヒーレンス

観測行列が零空間特性を持つかどうかや、制限等長性を持つかどうかの評価は、一般には困難である。次のインコヒーレンスは、評価が簡単でかつ重要な意味を持つ。

定義 5. (インコヒーレンス) 行列 A の異なる 2 つの列ベクトル $\mathbf{a}_i, \mathbf{a}_j$ の方向余弦の絶対値の最大値

$$\mu(A) := \max_{1 \leq i < j \leq N} \frac{|\langle \mathbf{a}_i, \mathbf{a}_j \rangle|}{\|\mathbf{a}_i\|_2 \|\mathbf{a}_j\|_2} \quad (15)$$

を行列 A のインコヒーレンスという [60, 31]. □

インコヒーレンスは、単にコヒーレンスとも呼ばれる。インコヒーレンスの計算は、異なる列の全ての組合せについて方向余弦を評価すればよいので、 $M \times N$ 行列では方向余弦を $O(N^2)$ 回評価するだけでよい。このため、spark や RIP 定数などと比較して、インコヒーレンスはとても簡単に求めることができる。インコヒーレンスが小さいほど再構成に有利になることは、次から理解できる。

定理 6. (RIP 定数の上界) 全ての列ベクトルの ℓ_2 ノルムが 1 である観測行列 A について、RIP 定数 δ_K は任意の $K \in \{1, \dots, N\}$ で、

$$\delta_K \leq \mu(A)K \quad (16)$$

とインコヒーレンスを用いて上から押さえられる [60].

証明は、付録 A.5 を参照されたい。このため、インコヒーレンスが小さいときは RIP 定数 (の上界) も小さくなり、インコヒーレンスが小さいほど再構成に有利になる、と解釈できる。この議論から、例えば (ランダム行列のように) 観測行列の列ベクトルの相関が小さいと、再構成に有利になるであろうことも予想される。インコヒーレンスが RIP 定数の上界を与え、RIP 定数が $\sqrt{2}-1$ 未満のとき頑健零空間特性を持つ (このため、零空間特性も持つ)、というようにインコヒーレンス、RIP 定数、頑健零空間特性 (および零空間特性) は関連している。

5 再構成アルゴリズム

圧縮センシング問題を ℓ_1 最小化に基づいて解くときは凸計画問題となる。このため種々の凸計画問題の解法が利用できる [10]。特に、内点法などを利用して、 N 次元の原信号の再構成を $O(N^3)$ の計算量でできる。しかし、圧縮センシングは画像処理など大きな原信号へ適用される場面も多く、より計算量の少ない再構成法が重要となっている。ここでは、代表的な幾つかのアルゴリズムを紹介する。

5.1 Orthogonal Matching Pursuit (OMP)

原信号を $\mathbf{x}_0 = (x_{0,i})$, 観測行列を $A = (\mathbf{a}_1, \dots, \mathbf{a}_N)$ とするとき, 観測信号が $\mathbf{y} = A\mathbf{x}_0 = \sum_{i=1}^N x_{0,i}\mathbf{a}_i$ のように観測行列の列ベクトルの線形結合で表されることを用いる [50, 21, 60]. 残差と相関の高い列ベクトルは原信号に含まれると判定するもので, 以下のようにまとめられる.

アルゴリズム 1. (OMP)

- 入力: 観測信号 $\mathbf{y} \in \mathbb{R}^M$, 観測行列 $A \in \mathbb{R}^{M \times N}$ (非零要素数 $K \in \mathbb{N}$).
- 初期化: 推定値 $\hat{\mathbf{x}} = \mathbf{0}$, 添字集合 $T = \phi$.
 1. 終了条件 (残差のノルムが一定値以下, $|T|$ が一定値以上, 等) を満たすまで以下をくり返す.
 - (a) 残差 $\mathbf{y} - A\hat{\mathbf{x}}$ と最も相関の高い (内積の絶対値の大きい) A の列ベクトルの添字を T に追加する.
 - (b) $\hat{\mathbf{x}} = \underset{\mathbf{x}: \mathbf{x}_{T^c} = \mathbf{0}}{\operatorname{argmin}} \|\mathbf{y} - A_T \mathbf{x}_T\|_2^2$ によって $|T|$ -スパースな推定値 $\hat{\mathbf{x}}$ を求める (一般逆行列で評価する).
- 出力: 推定値 $\hat{\mathbf{x}} \in \mathbb{R}^N$. □

5.2 Iterative Hard Thresholding (IHT)

原信号 $\mathbf{x}_0$ の非零要素数 K を既知として, 次の反復法に基づくアルゴリズムが考えられている [7, 8]. 残差 $\mathbf{y} - A\hat{\mathbf{x}}^t$ によって, 現在の推定値 $\hat{\mathbf{x}}^t$ を補正する. 推定値が K -スパースであることを保つために, 補正された推定値の要素の絶対値の大きい K 個以外は 0 するもので, 次のようになっている.

アルゴリズム 2. (IHT)

- 入力: 観測信号 $\mathbf{y} \in \mathbb{R}^M$, 観測行列 $A \in \mathbb{R}^{M \times N}$, 非零要素数 $K \in \mathbb{N}$.
- 初期化: 推定値 $\hat{\mathbf{x}}^0 = \mathbf{0}$, $t = 0$.
 1. 終了条件 (残差のノルムが一定値以下, 等) を満たすまで以下をくり返す.
 - (a) $\hat{\mathbf{x}}^{t+1} = h_K(\hat{\mathbf{x}}^t + A^\top(\mathbf{y} - A\hat{\mathbf{x}}^t))$ を求める.
 - (b) $t := t + 1$ と更新する.
- 出力: 推定値 $\hat{\mathbf{x}}^{t+1} \in \mathbb{R}^N$. □

ただし, $h_K(\mathbf{x})$ は $\mathbf{x}$ の要素のうち絶対値の大きい K 個以外の要素を 0 にする関数である. アルゴリズムの導出については, 文献 [7] を参照されたい. 推定誤差の上界なども解析的に評価されている [8]. 一般には, 非零要素数は未知であるが, 既知であるとしてアルゴリズムが構成されている場合も多い. 現在では, 少ない観測によって, 非零要素数 (の上界) を評価するといった研究も進みつつある [39].

5.3 Iterative Soft Tresholding (IST)

IST[6] は, lasso (ℓ_1/ℓ_2 最適化とも呼ばれる) による再構成 $\hat{\mathbf{x}} = \operatorname{argmin}_{\mathbf{x}} \{\frac{1}{2}\|\mathbf{y} - A\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_1\}$ を反復法に基づいて行うもので, 次のようにまとめられる.

アルゴリズム 3. (IST)

- 入力: 観測信号 $\mathbf{y} \in \mathbb{R}^M$, 観測行列 $A \in \mathbb{R}^{M \times N}$, 非零要素数 $K \in \mathbb{N}$, 正則化項の大きさ $\lambda \in \mathbb{R}^+$.
- 初期化: 推定値 $\hat{\mathbf{x}}^0 = \mathbf{0}$, $t = 0$
 1. c を $A^\top A$ の最大固有値より大きい値に決める.
 2. 終了条件 (残差のノルムが一定値以下, 等) を満たすまで以下をくり返す.
 - (a) $\hat{\mathbf{x}}^{t+1} = \eta(\hat{\mathbf{x}}^t + \frac{1}{c}A^\top(\mathbf{y} - A\hat{\mathbf{x}}^t); \lambda/c)$ を求める.
 - (b) $t := t + 1$ と更新する.
- 出力: 推定値 $\hat{\mathbf{x}}^{t+1} \in \mathbb{R}^N$. □

ただし, $\eta(x; \theta) = \operatorname{sgn}(x) \max\{|x| - \theta, 0\}$ である. アルゴリズムの導出については, 文献 [6, 63] を参照されたい.

5.4 Approximate Message Passing (AMP)

Donoho らは, ℓ_1 ノルム最小化を次の同時分布

$$p(\mathbf{x}|\mathbf{y}) \propto \exp(-\beta\|\mathbf{x}\|_1)\delta(A\mathbf{x} - \mathbf{y}) \quad (17)$$

が, $\beta \rightarrow \infty$ で復号に用いる最適化問題の解を台とする一様分布となることを利用して, 反復アルゴリズムを構成した [29]. ここで, δ は Dirac のデルタ関数である. この同時分布を用いて ℓ_1 ノルム最小解 $\hat{\mathbf{x}} = \operatorname{argmin}_{\mathbf{x}} \|\mathbf{x}\|_1$ s.t. $\mathbf{y} = A\mathbf{x}$ は, $\hat{\mathbf{x}} = \operatorname{argmax}_{\mathbf{x}} p(\mathbf{x}|\mathbf{y})$ となる. 周辺分布 $p_i(x_i|\mathbf{y}) := \int d\mathbf{x}_{\setminus i} p(\mathbf{x}|\mathbf{y})$ を用いると, $\hat{x}_i =$

$\operatorname{argmax}_{x_i} \lim_{\beta \rightarrow \infty} p_i(x_i | \mathbf{y})$ と要素毎の最適化問題として表現できる。メッセージ交換アルゴリズムと呼ばれる反復法を適用して近似的に周辺分布求め、(さらにそのメッセージ交換アルゴリズムを近似して) 得られた周辺分布を用いて ℓ_1 最小化を実行するものが近似的メッセージ交換法 (approximate message passing: AMP) である。

アルゴリズム 4. (AMP)

- 入力：観測信号 $\mathbf{y} \in \mathbb{R}^M$, 観測行列 $A \in \mathbb{R}^{M \times N}$, 非零要素数 $K \in \mathbb{N}$.
- 初期化： $\mathbf{x}^0 = \mathbf{0}$, $\mathbf{z}^0 = \mathbf{y}$, $\mathbf{z}^{-1} = \mathbf{0}$, $\sigma^0 = K/N$ (K/N は例), $t = 0$.
 1. $\lambda := \alpha^{-1/2} \operatorname{argmax}_{z \geq 0} \frac{1 - (2/\alpha)[(1+z^2)\Phi(-z) - z\phi(z)]}{1 + z^2 - 2[(1+z^2)\Phi(-z) - z\phi(z)]}$ を求める.
 2. 終了条件 (残差 $\mathbf{y} - A\mathbf{x}^{t+1}$ のノルムが一定値以下, など) を満たすまで以下を繰り返す.
 - (a) $\mathbf{z}^t = \mathbf{y} - A\mathbf{x}^t + \frac{1}{\alpha}\mathbf{z}^{t-1}\frac{\|\mathbf{x}^t\|_0}{N}$ を求める.
 - (b) $\mathbf{x}^{t+1} = \eta(A^\top \mathbf{z}^t + \mathbf{x}^t; \lambda\sigma_t)$ を求める.
 - (c) (経験的) 平均 2 乗誤差 σ^{t+1} を求める.
 - (d) $t := t + 1$ と更新する.
- 出力：推定値 $\hat{\mathbf{x}} := \mathbf{x}^{t+1} \in \mathbb{R}^N$ □

ただし, $\eta(x; \theta) = \operatorname{sgn}(x) \max\{|x| - \theta, 0\}$, $\phi(z) := (2\pi)^{-1/2} \exp(-z^2/2)$, $\Phi(z) := \int_z^\infty \phi(t)dt$ である. アルゴリズムの導出については, 文献 [29] を参照されたい. 原信号と推定ベクトルの平均 2 乗誤差 (mean square error: MSE) は, $\sigma^t := N^{-1} \sum_{i=1}^N (x_{0,i} - x_i^t)^2$ であるが, 原信号 $\mathbf{x}_0$ は未知なので求めることはできない. 実際にアルゴリズムを利用する際には, 真の MSE σ^{t+1} の代わりに, その経験的 MSE (例えば推定値 $\mathbf{x}^t$ の要素の内, 絶対値の小さい方から $N - K$ 個の要素の分散や中央値) を用いる必要がある. このようにしても性能には, あまり影響を与えない. 図 6 は, 再構成が成功する場合 (推定ベクトルの MSE が 0 に近づく場合) のしきい値関数 $\eta(x; \lambda\sigma_t)$ の変化を示す. MSE が大きい場合は推定ベクトルの要素として 0 が取りやすくなり, MSE が小さくなると 0 にする効果は少なくなる.

5.5 Node-Based verification (NB) Algorithm

ランダムガウス行列や, 列ベクトルを単位超球上の一様分布の実現値とした行列など, 観測行列として密な実行列を用いる場合は多い. 一方で, 疎な観測行列を用いるこ

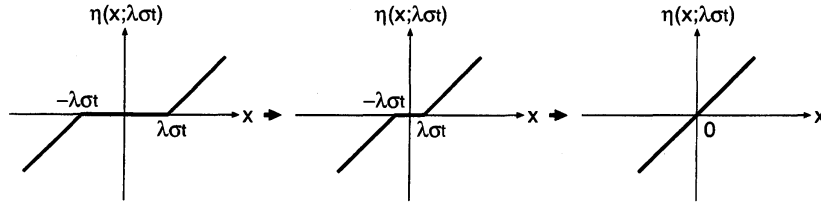


図 6: しきい値関数

とによって、計算量の低い再構成アルゴリズムの検討も行われている [61, 62, 36, 55]. ノード検証アルゴリズム (Node-Based verification Algorithm: NBA)[40, 61] は、各要素が2値であるような観測行列の場合に適用できるアルゴリズムである. 与えられた観測行列 $A = (a_{mn}) \in \{0, 1\}^{M \times N}$ で決まる2部グラフを考える. M 個の観測ノードと N 個の信号ノードを準備して, $a_{mn} = 1$ となる (m, n) についてのみ, 第 m 番目の観測ノードと第 n 番目の信号ノード間をエッジで接続する. 第 m 番目の観測ノードに隣接する信号ノードの集合を $\partial m := \{n \in [1 : N] : a_{mn} = 1\}$ とし, 第 n 番目の信号ノードに隣接する信号ノードの集合を $\partial n := \{m \in [1 : M] : a_{mn} = 1\}$ とする.

ノード検証アルゴリズムは, 有限個の原信号の成分の和が $\sum_{n \in \partial m} x_n = 0$ となるのは, それらの要素がすべて $x_n = 0, \forall n \in \partial m$ のときに限る, という仮定に基づく. 第 m 番目の観測ノードに割り当てられる値 $(d_m, \xi_m) \in \mathbb{Z} \times \mathbb{R}, \forall m \in [1 : M]$ と, 第 n 番目の信号ノードに割り当てられる値 $(s_n, w_n) \in \{0, 1\} \times \mathbb{R}, \forall n \in [1 : N]$ を, 次のように更新することによって再構成を行う.

アルゴリズム 5. (NB algorithm)

- 入力: 観測信号 $\mathbf{y} \in \mathbb{R}^M$, 観測行列 $A \in \{0, 1\}^{M \times N}$.
 - 初期化: 観測ノードの値 $(d_m, \xi_m) := (|\partial m|, y_i), \forall m$,
信号ノードの値 $(s_n, w_n) := (0, 0), \forall n$.
1. 終了条件 ($s_n = 1 \forall n$, または, 既定の反復回数に到達) を満たすまで以下を繰り返す.
 - (a) すべての信号ノード $n \in [1 : N]$ について以下を行う.
 - i. $d_m = 1$ となる観測ノード $m \in \partial n$ があれば,
 $(s_n, w_n) := (1, \xi_m)$ と更新する.
 - ii. $\xi_m = 0$ となる観測ノード $m \in \partial n$ があれば,
 $(s_n, w_n) := (1, 0)$ と更新する.

- iii. $|\{m \in \partial n : \xi_m = \xi\}| \geq 2$ (複数で ξ の値が同じ) ならば,
 $(s_n, w_n) := (1, \xi)$ と更新する.
- (b) すべての観測ノード $m \in [1 : M]$ について以下を行う.
 - i. $d_m := |\{n \in \partial m : s_n = 0\}|$ と更新する.
 - ii. $\xi_m := y_m - \sum_{n \in \partial m} s_n w_n$ と更新する.

• 出力：推定値 $\hat{\mathbf{x}} := (w_1, \dots, w_N) \in \mathbb{R}^N$ □

観測ノードの値 $d_m \in \mathbb{Z}$ は、推定値が確定していない隣接する信号ノードの数を表す。観測ノードの値 $\xi_m \in \mathbb{R}$ は、の確定していない隣接する信号ノードについての原信号の和を表す。信号ノードの値 $s_n \in \{0, 1\}$ は、推定ベクトル $\hat{\mathbf{x}}$ の第 n 番目の要素が確定していないときは 0 とし、確定したときは 1 にする。信号ノードの値 $w_n \in \mathbb{R}$ は、推定ベクトル $\hat{\mathbf{x}}$ の第 n 番目の確定した値を表す。信号ノードの処理方法を少し変更した同様なアルゴリズムもいくつか提案されている [61]。

6 典型的性能評価

これまでに、レプリカ法や経路積分法といった統計力学的な手法によっても、圧縮センシングの解析的評価が試みられている。

6.1 弱しきい値

ℓ_1 再構成の最悪評価は制限等長性を用いて、またより一般には ℓ_p 再構成の最悪評価は零空間特性を用いて詳しく議論されている。その一方で、 ℓ_p 再構成の典型的な性能は、情報統計力学の手法を用いて評価されている [35]。これには、条件付き確率

$$p(\mathbf{x}|\mathbf{y}) \propto \exp(-\beta \|\mathbf{x}\|_p^p) \delta(A\mathbf{x} - \mathbf{y}) \quad (18)$$

が、 $\beta \rightarrow \infty$ で復号に用いる ℓ_p ノルム最小化の解を台とする一様分布となることが利用される。この条件付き確率の規格化定数が、分配関数としての役割を果たし、一般の $p \geq 0$ での ℓ_p 復元の性能評価が可能になる。解析の詳細は文献 [35, 2] を参照されたい。

詳細は割愛するが、レプリカ法と呼ばれる手法によって、信号密度 $\rho = K/N$ を固定した条件下で、原信号 $\mathbf{x}_0$ と推定値 $\hat{\mathbf{x}}$ の MSE $\sigma^2 := \mathbb{E}_{A, \mathbf{x}_0}(\|\hat{\mathbf{x}} - \mathbf{x}_0\|_2^2)$ が評価すると、復元が成功する場合 ($\sigma^2 = 0$ となる場合) の最大の圧縮率 $\alpha_c(\rho)$ を得ること

ができる。特に、 $p = 1$ のときの結果は、

$$\alpha_c(\rho) = 2(1 - \rho)\Phi(\chi) + \rho \quad (19)$$

となる。ただし、 χ は、 $\chi = \alpha^{-1}[2(1 - \rho)\{(\chi + 1)\Phi(\xi^{-1/2}) - \chi^{1/2} \exp(-1/(2\chi))\}/\sqrt{2\pi}] + (1 + \chi)\rho$ の解とする。また、 $\phi(z) := (2\pi)^{-1/2} \exp(-z^2/2)$, $\Phi(z) := \int_z^\infty \phi(t)dt$ である。この結果は、Donoho-Tanner によるランダム高次元幾何学から得られる弱しきい値と一致する [26, 27]。図 7 に、 ℓ_1 再構成の弱しきい値を示す。この弱しきい値は、後述の AMP の性能評価からも得ることができる。

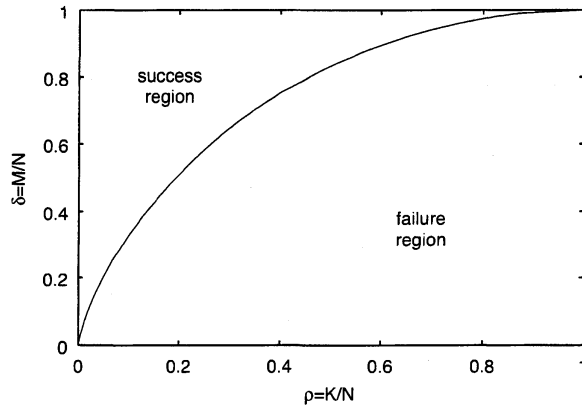


図 7: 弱しきい値.

6.2 反復復元法の典型性能

6.2.1 AMP の典型性能

Donoho らは、状態発展法 (state evolution: SE) [29] 推定されるベクトルの各要素の原信号以外の成分が正規分布であり、かつ原信号と独立ならば、アルゴリズム 4 の手順 3 で t 回反復後の推定値 $\mathbf{x}^t$ の MSE $\sigma_t^2 := \mathbb{E}_{\mathbf{x}_0}(\|\mathbf{x}^t - \mathbf{x}_0\|_2^2)$ は、

$$\sigma_{t+1}^2 = \Psi(\sigma_t^2) \quad (20)$$

となる [29]。ただし、 $\Psi(\sigma^2) := \mathbb{E}_{X_0, Z}\{[X_0 - \eta(X_0 + \frac{\sigma}{\sqrt{\alpha}}Z; \lambda\sigma)]^2\}$ である。また、 X_0 は原信号の要素と同じ確率変数であり、 $Z \sim \mathcal{N}(0, 1)$ を表す。 X_0 と Z は独立である。圧縮センシングの文脈から、再構成アルゴリズムである AMP 及びその性能解析手法である状態発展法が提案 [29] されたが、これは符号分割多元接続 (code division multiple access: CDMA) 通信方式の文脈において、Kabashima が提案したユーザ検

出アルゴリズム及びその解析法 [34] と本質的には等価であり、AMP と状態発展法は再発見と位置付けることができよう。

IST の反復式は、 $\mathbf{x}^{t+1} = \eta(\mathbf{x}_0 + [A^\top A - \mathbf{1}][\mathbf{x}_0 - \mathbf{x}^t]; \lambda\sigma^t)$ と書き直すことができる。もし、関数の中の $[A^\top A - \mathbf{1}][\mathbf{x}_0 - \mathbf{x}^t]$ が $\mathbf{x}_0$ と独立ならば解析は容易だが、 $\mathbf{x}^t$ は A に依存しているので独立ではない。ところが、AMP では残差の値を補正する項 $\frac{1}{\alpha} \mathbf{z}^{t-1} \frac{\|\mathbf{x}^t\|_0}{N}$ のため、推定するとき雑音のように働く項が原信号とは独立になる。詳細は割愛するが、少し一般化して次が成り立つことが示されている。すなわち、原信号の分布がある条件を満たすとき、殆ど至るところで微分可能な関数の任意の列 $\{\eta_t\}_{t \geq 0}$ と任意の疑似 Lipschitz 関数 $\psi: \mathbb{R}^2 \rightarrow \mathbb{R}$ について、

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \psi(x_i^{t+1}, x_{0,i}) = \mathbb{E}_{X_0, Z} \left[\psi(\eta_t(X_0 + \tau_t Z), X_0) \right] \quad (21)$$

となる [5]。ただし、 $\tau_{t+1}^2 = \frac{1}{\alpha} \mathbb{E}_{X_0, Z} \{ [\eta_t(X_0 + \tau_t Z) - X_0]^2 \}$ である。また、 $X_0 \sim p_{X_0}$, $Z \sim \mathcal{N}(0, 1)$ で、 Z は X_0 と独立であるとする。AMP の MSE の評価では、 $\psi(x, y) = (x - y)^2$, $\eta_t(x) = \eta(x; \lambda\sigma_t)$ である。

6.2.2 IST の典型性能

一見、AMP よりも単純な IST であるが、その挙動は反復過程の影響を強く受けるため、AMP よりも複雑になる。一般化した IST についての典型性能も調べてられている [49]。解析は、反復の履歴の影響を解析に取り込むことが可能な経路積分法 [22, 17, 18, 44, 45, 46, 47, 48] による。少ない反復回数での性能などの検討に役立つ可能性がある。推定値 $\mathbf{x}^{(0)}, \dots, \mathbf{x}^{(t)}$ の同時密度関数を考える。反復過程の記述に必要な量を取り出すために、パラメータ $\boldsymbol{\theta}^{(s)} = (\theta_1^{(s)}, \dots, \theta_N^{(s)})^\top$ を導入する（後で $\boldsymbol{\theta}^{(s)} = \mathbf{0} \forall s$ にする）と、

$$p[\mathbf{x}^{(0)}, \dots, \mathbf{x}^{(t)}] = \delta[\mathbf{x}^{(0)}] \prod_{s=0}^{t-1} \delta[\mathbf{x}^{(s+1)} - \eta_s(c^{-1} A^\top (\mathbf{y} - A\mathbf{x}^{(s)}) + \mathbf{x}^{(s)} + \boldsymbol{\theta}^{(s)})] \quad (22)$$

これは、更新式に関する情報をすべて含んでいる。ただし、 $\eta_t(\mathbf{x}) = \eta(\mathbf{x}; \lambda\hat{\sigma}_t/c)$, $\hat{\sigma}_0^2 = \rho$, $\hat{\sigma}_t^2 = \mathbb{E}[(x_n^t)^2 | x_{0,n} = 0]$ （零要素の MSE）とする。ここで、関数 $\eta_t(\mathbf{x})$ はベクトルに対して要素毎に作用するものとする。推定値の関数 $\mathcal{G} = \mathcal{G}(\mathbf{x}^{(0)}, \dots, \mathbf{x}^{(t)})$ の期待値は、 $\mathbb{E}_{\mathbf{x}}(\mathcal{G}) \triangleq \int_{\mathbb{R}^{(t+1)N}} [\prod_{s=0}^t d\mathbf{x}^{(s)}] p[\mathbf{x}^{(0)}, \dots, \mathbf{x}^{(t)}] \mathcal{G}$ となる。ダミー変数 $\boldsymbol{\psi}^{(s)} = (\psi_1^{(s)}, \dots, \psi_N^{(s)})^\top$ を導入して、次の母関数（特性関数）

$$Z[\boldsymbol{\psi}] \triangleq \mathbb{E}_{\mathbf{x}} \left[\exp \left(-i \sum_{s=0}^t \mathbf{x}^{(s)} \cdot \boldsymbol{\psi}^{(s)} \right) \right] \quad (23)$$

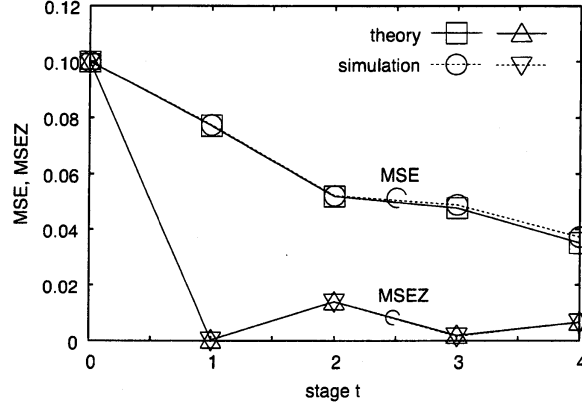


図 8: IST の反復過程.

を定義する. この母関数から, $\lim_{\psi \rightarrow 0} \partial Z[\psi] / \partial \psi_n^{(s)} = \mathbb{E}_{\mathbf{x}}(x_n^{(s)})$, $-\lim_{\psi \rightarrow 0} \partial^2 Z[\psi] / \partial \psi_n^{(s)} \partial \psi_{n'}^{(s')} = \mathbb{E}_{\mathbf{x}}(x_n^{(s)} x_{n'}^{(s')})$, $\lim_{\psi \rightarrow 0} \partial^2 Z[\psi] / \partial \psi_n^{(s)} \partial \theta_{n'}^{(s')} = \partial \mathbb{E}_{\mathbf{x}}(x_n^{(s)}) / \partial \theta_{n'}^{(s')}$ などの量を求めることが可能である. 母関数が観測行列と原信号についての期待値に集中すると仮定して, $Z[\psi]$ の代わりに, $\bar{Z}[\psi] \triangleq \mathbb{E}_{\mathbf{x}, A, \mathbf{x}_0}[\exp(-i \sum_{s=0}^t \mathbf{x}^{(s)} \cdot \boldsymbol{\psi}^{(s)})]$ を評価する.

母関数を評価すると, t 解反復後の MSE は $\sigma_t^2 = \rho - 2m^{(t)} + C^{(t,t)}$ と求められる. また, 同様に経験的 MSE も $\hat{\sigma}_t^2 = (1 - \rho)^{-1} \langle (x^t)^2 \mathbb{I}(x_0 = 0) \rangle$ と得られる. ただし, $m^{(t)} = \langle x_0 x^t \rangle$, $C^{(t,t')} = \langle x^t x^{t'} \rangle$, $G^{(t,t')} = \langle x^t (\mathbf{R}^{-1} \mathbf{v})_{t'} \rangle$ である. ここで, 詳細は割愛するが, $\mathbf{R}$ は反復過程による干渉項の共分散行列を, $\langle \cdot \rangle$ は反復過程をそれと統計的に等価な単一の要素の反復過程による期待値をそれぞれ表す. 図 8 に, $\rho = 0.1$, $\alpha = 0.5$, $\lambda = 3$, $c = 3$, $N = 2000$ での IST の反復過程の数値実験と理論値を示す. これは, 再構成が成功している場合の例である.

7 関連する話題

7.1 1 ビット圧縮センシング

観測信号の各要素が 1 ビットの情報しかないような場合の圧縮センシング問題は, 1 ビット圧縮センシング問題と呼ばれる [9]. 原信号 $\mathbf{x}_0 \in \mathbb{R}^N$ として, 観測行列 $A \in \mathbb{R}^{M \times N}$ を用いて, 2 値化された (1 ビットの) 観測信号 $\mathbf{y} = \text{sgn}(A\mathbf{x}_0) \in \{-1, 1\}^M$ が得られているとする. ただし, $\text{sgn}(x)$ は $x \geq 0$ で 1, その他で -1 をとる符号関数である. 任意の $c > 0$ に対して, $\mathbf{y} = \text{sgn}(A\mathbf{x}_0) = \text{sgn}(Ac\mathbf{x}_0)$ と観測信号が不変なので, 原信号 $\mathbf{x}_0$ の大きさは決定できない. このため, 原信号として, $\|\mathbf{x}_0\|_2^2 = 1$

と単位超球上のベクトルのみを考える。復元は、単位超球上のベクトルに限定した探索として、

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{x}\|_1 \quad \text{subj. to } \mathbf{y} = \operatorname{sgn}(A\mathbf{x}), \|\mathbf{x}\|_2^2 = 1 \quad (24)$$

のように行う。1ビット圧縮センシングでは、一般には $M > N$ と測定信号 $\mathbf{y}$ の次元の方が、原信号 $\mathbf{x}_0$ の次元よりも大きくなることに注意されたい。反復法によるアルゴリズムも提案されており、信号密度 $\rho = K/N$ を固定した条件下で、圧縮率 $\alpha = M/N$ と推定誤差の関係が評価されている。

7.2 ブラインド圧縮センシング

複数の観測信号ベクトルを観測して、観測行列と、複数の原信号ベクトルの両方を求める問題はブラインド圧縮センシング問題と呼ばれる [31]。密な原信号が疎になるような表現基底を、信号の推定と同時に求めることなどが目的である。

種々な問題設定が考えられているが [31]、最も簡潔な問題設定のひとつ [54] は以下のとおりである。観測行列を $A \in \mathbb{R}^{M \times N}$ とする。 M 次元列ベクトルを L 個観測する場合を考え、 M 次元の観測信号を列ベクトルとして観測信号を行列 $Y \in \mathbb{R}^{M \times L}$ で表し、同様に、 N 次元の原信号も列ベクトルとして原信号を行列 $X_0 \in \mathbb{R}^{N \times L}$ で表すと、観測は $Y = AX_0$ と表現できる。原信号行列に含まれる非零要素の割合を $\theta \in \mathbb{R}^+$ として、問題は、

$$(\hat{A}, \hat{X}) = \underset{A, X}{\operatorname{argmin}} \|Y - AX\|_2^2 \quad \text{subj. to } \|A\|^2 = MN, \|X\|_0 = NL\theta \quad (25)$$

となる。ただし、行列 $A = (a_{ij})$ に対して $\|A\|_2 := \sum_i \sum_j |a_{ij}|^2$ とする。また、行列 $A = (a_{ij})$ に対して $\|A\|_0$ は A の非零要素数とする。この問題は、統計力学的な手法で解析され、観測行列と原信号行列が共に正しく復元できる条件などが求められている [54]。

7.3 空間結合圧縮センシング

通信路符号化の分野で、空間結合符号 [30] が提案されている。これは、疎な帯行列の検査行列によって定義される低密度パリティ検査 (low density parity check: LDPC) 符号で、確率伝搬法 (belief Propagation: BP) 復号法によって通信路容量を達成する。

圧縮センシングでは観測行列として、ランダムガウス行列や、列ベクトルを単位超球上の一様分布の実現値とした行列など、密な実行列を用いる場合が多い。その一

方で、疎な観測行列を用いると、計算量の低い NB algorithm を適用できる [61, 62]. これに、空間結合符号のように、疎な帯行列を観測行列とした場合、ノード検証アルゴリズムの再構成の性能が向上することが示されている [37, 55]. 具体的な帯行列の構成方法については、文献 [37, 55] を参照されたい.

7.4 その他の話題

低次元信号としては、スパースなベクトルではなく低ランク行列を考えることによって、低ランク行列再構成 [51] の研究も進められている. そのほか、観測行列はランダム行列とする場合が多いが、制限等長性を持つ行列を決定論的に構成する方法 [23] についても議論されている. また、 ℓ_1 最小化のコスト関数 $|\mathbf{x}_1| + \cdots + |\mathbf{x}_N|$ によって再構成した推定ベクトルを用いて重み $c_1, \dots, c_N$ を決めて、それを用いた重み付けの和 $c_1|\mathbf{x}_1| + \cdots + c_N|\mathbf{x}_N|$ をコスト関数として、再び再構成を行うということを繰り返すという再重み付け ℓ_1 最小化 [15] の性能についても解析が行われている [57, 42].

8 まとめ

圧縮センシングは、スパース性を仮定して不良設定問題に接近する方法である. 再構成は、 ℓ_p 最小化に基づき、条件を満たす観測行列を用いることによって原信号を完全に再構成できる. このような圧縮センシングについて、その問題設定や、ほんの一部分にしかすぎないが最近の研究動向について簡単な説明を試みた. 圧縮センシングの解説は既にいくつか優れたものがあり、特に、零空間特性については文献 [1], 情報統計力学とそれに関連する学習理論などの話題との関係については文献 [2], 制限等長性についての Candes-Tao らの議論は文献 [4], Donoho のランダム高次元幾何学や低ランク行列再構成などの応用事例については文献 [3] などを併せて参照されたい. また、様々な関連研究の論文の情報を集めたサイト [64] も参照されたい. ここで触れていない数多くの興味深い議論がなされており、本稿が圧縮センシングの研究を始めるときに、ほんの少しでも役に立てば幸いである.

謝辞

本稿は、2011 年 10 月に開催された京都大学数理解析研究所共同利用研究集会「時間周波数解析の理論とその理工学的応用」での著者の講演内容に基づいている. 同

研究集会での講演の機会を与えていただいた芦野隆一氏（大阪教育大学）に深く感謝する。本稿に含まれる結果の一部は、文部科学省科学研究費補助金基盤研究 (C)（課題番号 2250136）の援助を受けて行われた。

A 証明

取り上げた定理のうち、初等的な議論で示されるいくつかの基本的な定理の証明をまとめておく。

A.1 定理 1 の証明

文献 [25] に従う。観測行列 A が $\text{spark}(A) > 2K$ を満たすとする。ある観測信号 $\mathbf{y}$ に対して、 $\mathbf{y} = A\mathbf{x} = A\mathbf{x}'$ となるような K -スパースなベクトル $\mathbf{x}, \mathbf{x}' \in \Sigma_K$ が存在すると仮定する。このとき、 $A(\mathbf{x} - \mathbf{x}') = \mathbf{0}$ となっている。 ℓ_0 ノルムは、 $\|\mathbf{x} - \mathbf{x}'\|_0 \leq \|\mathbf{x}\|_0 + \|\mathbf{x}'\|_0$ を満たすので、 $\mathbf{x} - \mathbf{x}'$ は $2K$ -スパースなベクトルである。 $\text{spark}(A) > 2K$ なので、観測行列 A の全ての $2K$ 個の列ベクトルの集合は一次独立である。このことから、 $\mathbf{x} - \mathbf{x}' = \mathbf{0}$ であることがわかる。すなわち、 $\mathbf{y} = A\mathbf{x}$ を満たす K -スパースなベクトル $\mathbf{x} \in \Sigma_K$ が存在するとき、それは一意に定まる。（より詳しく $\text{spark}(A) \leq 2K$ では一意に定まらないことも示されている。）

A.2 定理 2 の証明

文献 [32, 1] に従う。任意の $0 < p \leq 1$ と任意の正数 $u > 0, v > 0$ について、 $(u + v)^p \leq u^p + v^p$ となる。これより、任意の $0 < p \leq 1$ と任意の $u, v \in \mathbb{R}$ について、 $|u + v|^p \leq (|u| + |v|)^p \leq |u|^p + |v|^p$ となる。 $u = a + b, v = -b$ とおくと、任意の $0 < p \leq 1$ と任意の $a, b \in \mathbb{R}$ について、

$$|a + b|^p \geq |a|^p - |b|^p \quad (\text{A.1})$$

を得る。証明に、この不等式を用いる。

観測行列 $A \in \mathbb{R}^{M \times N}$ が ℓ_p ノルムで K 次の零空間特性を持つ、すなわち (6) を満たすと仮定する。また、原信号を任意の K -スパースなベクトル $\mathbf{x}_0 = (x_{0,1}, \dots, x_{0,N}) \in \Sigma_K$ とし、その台を $T = \text{supp}(\mathbf{x}_0)$ とする。任意の零空間の非零元 $\mathbf{z} = (z_1, \dots, z_N) \in$

$\ker(A) \setminus \{0\}$ に対して,

$$\begin{aligned}
\|\mathbf{x}_0 + \mathbf{z}\|_p^p - \|\mathbf{x}_0\|_p^p &= \sum_{i \in \{1, \dots, N\}} |x_{0,i} + z_i|^p - \sum_{i \in \{1, \dots, N\}} |x_{0,i}|^p \\
&= \left(\sum_{i \in T} |x_{0,i} + z_i|^p + \sum_{i \in T^c} |0 + z_i|^p \right) - \left(\sum_{i \in T} |x_{0,i}|^p + \sum_{i \in T^c} |0|^p \right) \\
&\geq_{(a)} \sum_{i \in T} (|x_{0,i}|^p - |z_i|^p) + \sum_{i \in T^c} |z_i|^p - \sum_{i \in T} |x_{0,i}|^p \\
&= - \sum_{i \in T} |z_i|^p + \sum_{i \in T^c} |z_i|^p \\
&= -\|\mathbf{z}_T\|_p^p + \|\mathbf{z}_{T^c}\|_p^p >_{(b)} 0
\end{aligned} \tag{A.2}$$

となる。ただし、不等号 $\geq_{(a)}$ で (A.1) を、不等号 $>_{(b)}$ で (6) をそれぞれ用いた。これより、 $\|\mathbf{x}_0 + \mathbf{z}\|_p^p \geq \|\mathbf{x}_0\|_p^p$ を得る。すなわち、 $\mathbf{y} = A\mathbf{x}$ を満たす $\mathbf{x}$ のうち最も $\|\mathbf{x}\|_p$ の小さいものが $\mathbf{x} = \mathbf{x}_0$ となる。また、 $0^0 = 0$ と定義すると、 $p = 0$ の場合についても同様に示される。

A.3 定理3の証明

文献 [20, 1] に従う。観測行列 $A \in \mathbb{R}^{M \times N}$ が、 ℓ_p ノルムで K 次の係数 σ の頑健零空間特性を持つ、すなわち (8) を満たすと仮定する。再構成されたベクトル $\hat{\mathbf{x}}$ は ℓ_p ノルムが最小なので、 $\|\hat{\mathbf{x}}\|_p^p < \|\mathbf{x}_0\|_p^p$ となっている。また、 $A\hat{\mathbf{x}} = A\mathbf{x}_0$ なので、 $\hat{\mathbf{x}} = \mathbf{x}_0 + \mathbf{z}$ となるような $\mathbf{z} \in \ker(A)$ が存在する。任意の大きさが $|T| \leq K$ であるような添字の集合 $T \subset \{1, \dots, N\}$ について、

$$\begin{aligned}
\|\mathbf{x}_0\|_p^p &\geq \|\hat{\mathbf{x}}\|_p^p = \|\mathbf{x}_0 + \mathbf{z}\|_p^p = \sum_{i \in \{1, \dots, N\}} |x_{0,i} + z_i|^p = \sum_{i \in T} |x_{0,i} + z_i|^p + \sum_{i \in T^c} |x_{0,i} + z_i|^p \\
&\geq_{(a)} \sum_{i \in T} (|x_{0,i}|^p - |z_i|^p) + \sum_{i \in T^c} (|z_i|^p - |x_{0,i}|^p) \\
&= \|(\mathbf{x}_0)_T\|_p^p - \|\mathbf{z}_T\|_p^p + \|\mathbf{z}_{T^c}\|_p^p - \|(\mathbf{x}_0)_{T^c}\|_p^p
\end{aligned} \tag{A.3}$$

が成り立つ。ここで、不等号 $\geq_{(a)}$ で (A.1) を用いた。また、 $\|\mathbf{x}_0\|_p^p = \|(\mathbf{x}_0)_T\|_p^p + \|(\mathbf{x}_0)_{T^c}\|_p^p$ なので、 $\|\mathbf{z}_{T^c}\|_p^p \leq 2\|(\mathbf{x}_0)_{T^c}\|_p^p + \|\mathbf{z}_T\|_p^p$ となる。これに、さらに (8) を用いると、

$$\|\mathbf{z}_{T^c}\|_p^p < \frac{2}{1 - \sigma} \|(\mathbf{x}_0)_{T^c}\|_p^p \tag{A.4}$$

を得る。この関係を用いて、誤差を評価すると、

$$\|\hat{\mathbf{x}} - \mathbf{x}_0\|_p^p = \|\mathbf{z}\|_p^p = \|\mathbf{z}_T\|_p^p + \|\mathbf{z}_{T^c}\|_p^p <_{(b)} (1 + \sigma) \|\mathbf{z}_{T^c}\|_p^p <_{(c)} 2 \frac{1 + \sigma}{1 - \sigma} \|(\mathbf{x}_0)_{T^c}\|_p^p \quad (\text{A.5})$$

となる。ただし、不等号 $<_{(b)}$ で再び (8) を、不等号 $<_{(c)}$ で (A.4) をそれぞれ用いた。集合 $T \subset \{1, \dots, N\}$ は、大きさが $|T| \leq K$ (K 次の頑健零空間特性を用いたので) であれば任意でよい。最も $\|(\mathbf{x}_0)_{T^c}\|_p^p$ を小さくする T を選んでも、この不等式は成り立つ。この誤差が最小になる添字の集合 T は、単に原信号 $\mathbf{x}_0$ の要素のうち絶対値の大きい要素の添字を、大きいほうから K 個選んだものになる。

A.4 定理 4 の証明

$2K$ 次の制限等長性より、任意の非零の $2K$ -スパースなベクトル $\mathbf{x} \in \Sigma_{2K} \setminus \{\mathbf{0}\}$ に対して、

$$\|A\mathbf{x}\|_2^2 \geq (1 - \delta_{2K}) \|\mathbf{x}\|_2^2 > 0 \quad (\text{A.6})$$

となる。このため、行列 A の任意の $2K$ 個以下の列ベクトルの組は一次独立である。よって、 $\text{spark}(A) > 2K$ となるので、定理 1 より、多くともひとつの $\mathbf{y} = A\mathbf{x}$ を満たす K -スパースなベクトル $\mathbf{x} \in \Sigma_K$ が存在する。原信号は K -スパースなので、解は存在して $\hat{\mathbf{x}} = \mathbf{x}_0$ となる。

A.5 定理 6 の証明

任意の $a, b \in \mathbb{R}$ について、 $(|a| - |b|)^2 = |a|^2 - 2|ab| + |b|^2 \geq 0$ より、 $\frac{a^2 + b^2}{2} \geq |ab|$ が成り立つことを用いる。観測行列 $A = (\mathbf{a}_1, \dots, \mathbf{a}_N) \in \mathbb{R}^{M \times N}$ の列ベクトルの ℓ_2 ノルムが 1 であることを仮定しているので、インコヒーレンスは $\mu(A) = \max_{1 \leq i < j \leq N} |\langle \mathbf{a}_i, \mathbf{a}_j \rangle|$ である。大きさが $1 \leq |T| \leq K$ の添字の集合を $T \subset \{1, \dots, N\}$

とすると, 任意の $\mathbf{x} \in \mathbb{R}^{|T|}$ について線形変換後のノルムが,

$$\begin{aligned}
\|A_T \mathbf{x}\|_2^2 &= (A_T \mathbf{x})^\top (A_T \mathbf{x}) = \left(\sum_{i \in T} x_i \mathbf{a}_i \right)^\top \left(\sum_{j \in T} x_j \mathbf{a}_j \right) = \sum_{i \in T} \sum_{j \in T} x_i x_j \mathbf{a}_i^\top \mathbf{a}_j \\
&= \sum_{i \in T} x_i^2 + \sum_{i \in T} \sum_{j \in T \setminus \{i\}} x_i x_j \mathbf{a}_i^\top \mathbf{a}_j \\
&\leq \sum_{i \in T} x_i^2 + \sum_{i \in T} \sum_{j \in T \setminus \{i\}} |x_i x_j| \|\mathbf{a}_i^\top \mathbf{a}_j\| \leq \sum_{i \in T} x_i^2 + \mu(A) \sum_{i \in T} \sum_{j \in T \setminus \{i\}} |x_i x_j| \\
&\leq \sum_{i \in T} x_i^2 + \mu(A) \sum_{i \in T} \sum_{j \in T} \frac{x_i^2 + x_j^2}{2} \leq \sum_{i \in T} x_i^2 + \mu(A) |T| \sum_{i \in T} x_i^2 \\
&\leq (1 + \mu(A)K) \|\mathbf{x}\|_2^2
\end{aligned} \tag{A.7}$$

と上から押えられる. 同様に, $\|A_T \mathbf{x}\|_2^2 \geq \sum_{i \in T} x_i^2 - \sum_{i \in T} \sum_{j \in T \setminus \{i\}} |x_i x_j| \|\mathbf{a}_i^\top \mathbf{a}_j\| \geq (1 - \mu(A)K) \|\mathbf{x}\|_2^2$ と下からも抑えられる. よって, 制限等長性の定義 (11) より $\delta_K \leq \mu(A)K$ を得る.

参考文献

- [1] 平林 晃, “Compressed Sensing –基本原理と最新研究動向–,” 信学技報告, SIP2009-, CAS2009-, VLD2009-, 1 (2009).
- [2] 樺島 祥介, “圧縮センシングへの統計力学的アプローチ,” 日本神経回路学会誌, vol. 17, no. 2, 70 (2010).
- [3] 田中 利幸, “圧縮センシングの数理,” *IEICE Fundamentals Review*, vol. 4, no. 1, 39 (2010).
- [4] 和田山 正, “圧縮センシングにおける完全再現十分条件について,” 日本神経回路学会誌, vol. 17, no. 2, 63 (2010).
- [5] M. Bayati and A. Montanari, “The dynamics of message passing on dense graphs, with applications to compressed sensing,” *Proc. of the 2010 IEEE Int’l Sympo. on Info. Theory (ISIT)*, 1528 (2010). [the longer version will appear in *IEEE Transactions on Information Theory*]
- [6] A. Beck and M. Teboulle, “A Fast Shrinkage-Thresholding Algorithms for Linear Inverse Problems,” *SIAM J. Imaging Sciences*, vol. 2, no. 1, 182 (2009).

- [7] T. Blumensath and M. Davis, "Iterative Thresholding for Sparse Approximations," *Journal of Fourier Analysis and Applications*, vol. 14, 629 (2008).
- [8] T. Blumensath and M. Davis, "Iterative hard thresholding for compressed sensing," *Appl. Comput. Harmon. Anal.*, vol. 27, no. 3, 265 (2009).
- [9] P. T. Boufounos and R. G. Baraniuk, "1-Bit compressive sensing," *Proc. of the 42nd Annual Conf. on Info. Sci. & Syst. (CISS2008)*, 16 (2008).
- [10] S. Boyd, L. Vandenberghe, "Convex Optimization," Cambridge University Press (2004).
- [11] E. J. Candés and T. Tao, "Decoding by linear programming," *IEEE Trans. Info. Theory*, vol. 51, no. 12, 4203 (2005).
- [12] E. J. Candés, J. Romberg and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Info. Theory*, vol. 52, no. , 489 (2006).
- [13] E. J. Candés and T. Tao, "Near-Optimal Signal Recovery From Random Projections: Universal Encoding Strategies?" *IEEE Trans. Info. Theory*, vol. 52, no. 12, 5406 (2006).
- [14] E. J. Candés, "The restricted isometry property and its implications for compressed sensing," *Compte Rendus de l'Academie des Sciences, Paris, Serie I*, 589 (2008).
- [15] E. J. Candés, M. B. Wakin and S. P. Boyd, "Enhanceing sparsity by reweighted ℓ_1 minimization," *Journal of Fourier Analysis and Applications*, vol. 14, nos. 5-6, 877 (2008).
- [16] J. F. Claerbout and F. Muir, "Robust modeling with erratic data," *Geophysics*, 38, 826 (1973).
- [17] A. C. C. Coolen, "Statistical mechanics of recurrent neural networks II. Dynamics," *Preprint arXiv cont-mat/0006011* (2000).
- [18] A. C. C. Coolen, *The Mathematical Theory of Minority Games*, Oxford Univ. Press (2005).

- [19] E. T. Copson, *Asymptotic Expansions*, Cambridge Univ. Press (1965).
- [20] M. Davies and R. Gribonval, "On L_p minimisation, instance optimality, and restricted isometry constants for sparse approximation," *Proc. Sampling Theory and Applications (SampTA)*, 4 pages (2009).
- [21] G. Davis, S. Mallat, and M. Avellaneda, "Adaptive greedy approximation," *J. Constr. Approx.*, vol. 13, 57 (1997).
- [22] C. De Dominicis, "Dynamics as a substitute for replicas in systems with quenched random impurities," *Phys. Rev. B*, vol. 18, 4913 (1978).
- [23] R. A. DeVore, "Deterministic constructions of compressed sensing matrices," *Journal of Complexity*, vol. 23, 918 (2007).
- [24] D. L. Donoho, "Uncertainty Principles and Signal Recovery," *SIAM J. Appl. Math.*, vol. 49, no. 3, 906 (1989).
- [25] D. L. Donoho, "Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ^1 minimization," *Proc. Natl. Acad. Sci.*, vol. 100, no. 5, 2197 (2003).
- [26] D. L. Donoho and J. Tanner, "Neighborliness of randomly projected simplices in high dimensions," *Proc. of the National Academy of Sciences, USA*, vol. 102, no. 27, 9452 (2005).
- [27] D. L. Donoho, "High-dimensional centrally-symmetric polytopes with neighborliness proportional to dimension," *Discrete & Computational Geometry*, vol. 35, no. 4, 617 (2006).
- [28] D. L. Donoho, "Compressed Sensing," *IEEE Trans. Info. Theory*, vol. 52, no. 4, 1289 (2006).
- [29] D. L. Donoho, A. Maleki and A. Montanari, "Message-passing algorithms for compressed sensing," *Proc. of the National Academy of Sciences (PNAS)*, vol. 106, no. 45, 18914 (2009).
- [30] A. J. Felström and K.S. Zigangirov, "Time-varying periodic convolutional codes with low-density parity-check matrix," *IEEE Trans. Info. Theory*, vol. 45, no. 6, 2181 (1999).

- [31] S. Gleichman and Y. C. Eldar, "Blind Compressed Sensing," *IEEE Trans. Info. Theory*, vol. 57, no. 10, 6958 (2011).
- [32] R. Gribonval and M. Nielsen, "Sparse representation in unions of bases," *IEEE Trans. Info. Theory*, vol. 49, no. 12, 3320 (2003).
- [33] J. A. F. Heindel, "Dynamics of learning by neurons and agents," *Doctoral Thesis*, King's College London (2001).
- [34] Y. Kabashima, "A CDMA multiuser detection algorithm on the basis of belief propagation," *J. Phys. A: Math. Gen.*, vol. 36, no. 43, 11111 (2003).
- [35] Y. Kabashima, T. Wadayama, T. Tanaka, "A typical reconstruction limit for compressed sensing," *Journal of Statistical Mechanics: Theory and Experiments*, 109003 (2009).
- [36] Y. Kabashima and T. Wadayama, "A signal recovery algorithm for sparse matrix based compressed sensing," *arXiv1102.3220* (2011).
- [37] S. Kudekar and H. Pfister, "The effect of spatial coupling on compressive sensing," *Proc. 48th Annual Allerton Conf. on Commun., Control and Computing*, 347 (2010).
- [38] M. Ledoux, *The Concentration of Measure Phenomenon*, Mathematical Surveys and Monographs, no. 89, American Mathematical Society (2001).
- [39] Miles E. Lopes, "Estimating Unknown Sparsity in Compressed Sensing," *arXiv:1204.4227* (2012).
- [40] M. Luby and M. Mitzenmacher, "Verification-based decoding for packet-based low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol.51, no.1, 120 (2005).
- [41] V. A. Marcenko and L. A. Pastur, "Distribution of eigenvalues for some sets of random matrices," *Mathematics for the USSR-Sbornik*, vol. 1, no. 4, 457 (1967).
- [42] R. Matsushita and T. Tanaka, "Critical compression ratio of iterative reweighted l1 minimization for compressed sensing," *Proc. of the 2011 Information Theory Workshop (ITW)*, 568 (2011).

- [43] N. Merhav, "Statistical Physics and Information Theory," now publishing (2009).
- [44] K. Mimura K and M. Okada, "Generating functional analysis of CDMA detection dynamics," *J. Phys. A: Math. Gen.*, vol. 38, no. 46, 9917 (2005).
- [45] K. Mimura and M. Okada, "Generating functional analysis of CDMA multiuser detection dynamics," *Proc. of the 2006 IEEE Int'l Sympo. on Info. Theory (ISIT)*, 2373 (2006).
- [46] K. Mimura and M. Okada, "Generating functional analysis of CDMA soft-decision multiuser detector," *Proc. of the 2007 IEEE Int'l Sympo. on Info. Theory (ISIT)*, 2076 (2007).
- [47] K. Mimura and A. C. C. Coolen, "Generating functional analysis of LDGM channel coding with many short loops," *Proc. of the 2009 IEEE Int'l Sympo. on Info. Theory (ISIT)*, 1829 (2009).
- [48] K. Mimura and M. Okada, "Generating Functional Analysis for Iterative CDMA Multiuser Detectors," [arXiv:1103.1439](https://arxiv.org/abs/1103.1439) (2011).
- [49] K. Mimura, "Generating functional analysis of iterative reconstruction algorithms for compressed sensing," *Proc. of the 2011 IEEE Int'l Sympo. on Info. Theory (ISIT)*, 1432 (2011).
- [50] Y. C. Pati, R. Rezaifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," *Proc. 27th Annu. Asilomar Conf. Signals Syst. Comput.*, vol. 1, 40 (1993).
- [51] B. Recht, M. Fazel and P. A. Parrilo, "Guaranteed Minimum-Rank Solutions of Linear Matrix Equations via Nuclear Norm Minimization," *SIAM Review*, vol. 52, no. 3, 471 (2010).
- [52] T. J. Richardson and R. L. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Trans. Info. Theory*, vol. 47, no. 2, 599 (2001).
- [53] F. Santosa and W. W. Symes, "Linear Inversion of Band-Limited Reflection Seismograms," *SIAM J. Sci. and Stat. Comput.*, vol. 7, no. 4, 1307 (1986).

- [54] A. Sakata, Y. Kabashima “Statistical Mechanics of Dictionary Learning,” *arXiv:1203.6178* (2012).
- [55] Y. Sato, K. Kasai and K. Sakaniwa, “Compressed sensing via node-based verification recovery with band sparse matrices,” *Proc. of the 34th Sympo. on Info. Theory and its App. (SITA2011)*, 526 (in Japanese) (2011).
- [56] T. Tanaka and M. Okada, “Approximate Belief Propagation, Density Evolution, and Statistical Neurodynamics for CDMA Multiuser Detection,” *IEEE Trans. Info. Theory*, vol. 51, no. 2, 700 (2005).
- [57] T. Tanaka and J. Raymond, “Optimal incorporation of sparsity information by weighted L_1 optimization,” *Proc. of the 2010 IEEE Int’l Sympo. on Info. Theory (ISIT)*, 1598 (2010).
- [58] R. Tibshirani, “Regression shrinkage and selection via the lasso,” *Journal of the Royal Statistical Society B*, vol. 58, no. 1, 267 (1996).
- [59] J. Tropp and A. Gilbert, “Signal recovery from partial information via orthogonal matching pursuit,” *IEEE Trans. Info. Theory*, vol. 53, no. 12, 4655 (2007).
- [60] J. Tropp and S. J. Wright, “Computational methods for sparse solution of linear inverse problems,” *Proc. of IEEE Appl. & Comput. Math.*, vol. 98, no. 6, 948 (2010).
- [61] F. Zhang and H.D. Pfister, “List-message passing achieves capacity on the q-ary symmetric channel for large q,” *arXiv:0806.3243* (2008).
- [62] F. Zhang and H. D. Pfister, “On the iterative decoding of high-rate LDPC codes with applications in compressed sensing,” *arXiv0903.2232* (2009).
- [63] M. Zibulevsky and M. Elad, “L1 and L2 Optimization in Signal and Image Processing,” *IEEE Signal Process. Mag.*, vol. 27, no. 3, 76 (2010).
- [64] “Compressive Sensing Resources,” <http://dsp.rice.edu/cs>