

圧縮センシングの数理

Mathematics of Compressed Sensing

田中利幸 Toshiyuki TANAKA



アブストラクト 圧縮センシングの基本的な問題は、スパースなベクトルに対する線形観測に基づいて1-ノルム最適化によって元のベクトルを推定する問題において、正しい推定結果を得るために必要な観測数が元のベクトルのスパースさにどう依存するかを問うものである。Candes-Taoの結果、及びDonoho-Tannerの結果を中心に、圧縮センシングの数理について概説する。1-ノルム最適化の歴史や、低ランク行列の再構成などの関連する話題についても触れる。

キーワード 圧縮センシング, 1-ノルム最適化, 行列再構成, 積分幾何学, 情報統計力学

1. まえがき

画像、音声、動画などのデータは概してサイズが大きいため、保存の際などに圧縮されることも少なくない。これらのマルチメディアデータにはそれなりの冗長さがあるため、それを利用してサイズを圧縮することができるというわけである。しかしよく考えると、これは何だかおかしくはないだろうか。それなりの労力をかけて取得したデータの大半は、ただ圧縮に際して捨てられるだけに終わっているのである。もし、圧縮した形でデータを取得できるのであれば、その方が好ましいといえることができる。「圧縮センシング」は、そのようなことを可能にする新たな枠組みを与えるものとして近年大いに注目を集めている。

本稿では、圧縮センシングとはどういう問題なのか、またどのような研究成果が得られているのか、といったことに関して解説したい。しかしながら、圧縮センシングに関しては既に膨大な研究成果が報告されており、それらを全般的に網羅することは筆者の力量をはるかに超える。以下では、問題の数理的な側面に限定して解説を試みたい。

2. 基本的な問題設定

圧縮センシングの基本的な問題設定は、未知ベクトルを線形観測に基づいて推定する問題である。問題設定を理解するには、大学教養程度の線形代数の知識があれば十分である。未知ベクトルを N 次元の実ベクトル \mathbf{x}_0 とする。ベクトル \mathbf{x}_0 に対する線形観測は、観測ベクトル $\mathbf{a} \in \mathbb{R}^N$ を一つ選んで内積 $\mathbf{a} \cdot \mathbf{x}_0$

をとることに対応する。線形観測を M 回行うものとし、対応する観測ベクトルを $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_M\}$ とすると、 \mathbf{x}_0 に対する M 回の線形観測は、大きさが $M \times N$ の観測行列

$$A = \begin{pmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \vdots \\ \mathbf{a}_M^T \end{pmatrix} \quad (1)$$

を使って

$$\mathbf{y} = A\mathbf{x}_0 \quad (2)$$

と表現することができる。ベクトル $\mathbf{y} \in \mathbb{R}^M$ は、 M 回の線形観測の結果を成分として持つベクトルである。

観測行列 A を既知としたとき、観測結果 \mathbf{y} からベクトル \mathbf{x}_0 を推定する問題を考えてみよう。これは \mathbf{x}_0 を変数とする連立一次方程式を解く問題にほかならない。 \mathbf{x}_0 を正しく推定できるための必要十分条件は $\text{rank } A = N$ が成り立つことであるが、直観的には $M \geq N$ であれば大抵大丈夫だと思っていよい。 M 個の観測ベクトルからうまく N 個を選んで \mathbb{R}^N の基底とすることができればよいからである。一方、もし $M < N$ であれば、 \mathbf{x}_0 を正しく推定することはできない。条件式の個数 M よりも値を定めるべき変数の数 N の方が大きく、解を一意に定めることができないからである。この場合には、よく知られているように、一般逆行列を使えば最小ノルム解、すなわちユークリッドノルムが最小となる推定結果を得ることができる。

次に、ベクトル \mathbf{x}_0 が「スパース」であることを知っているものとして、先ほどと同様に観測結果 \mathbf{y} からベクトル \mathbf{x}_0 を推定する問題を考えてみる。具体的には、 \mathbf{x}_0 の N 個の成分のうち 0 でない値をとるものは高々 K 個である、ということが分かっているものとする。このとき、「ベクトル \mathbf{x}_0 は K -スパースである」と呼ぶことにする。もし、 \mathbf{x}_0 のどの成分が 0 でない値をとるかについても分かっているのであれば、話は簡単である。 $M < K$ であれば \mathbf{x}_0 を正しく推定することはできない

田中利幸 正員 京都大学大学院情報学研究所
E-mail tt@i.kyoto-u.ac.jp
Toshiyuki TANAKA, Member (Dept. Systems Science, Graduate School of Informatics, Kyoto University, Kyoto-shi, 606-8501 Japan).
電子情報通信学会 基礎・境界サイエティ
Fundamentals Review Vol.4 No.1 pp.39-47 2010 年 7 月
©電子情報通信学会 2010

し、 $M \geq K$ であれば大抵大丈夫であろう。

ここで考えたいのは、 \mathbf{x}_0 が K -スパースであることは分かっているが、どの成分が 0 でない値をとるかは分からない場合である。この場合に、観測結果 \mathbf{y} から K -スパースなベクトル \mathbf{x}_0 をどのように推定したらよいであろうか。また、正しく推定を行うためには、どのくらいの観測数 M が必要であろうか。既に述べたように、 $M \geq N$ であれば大抵は正しい推定が可能であるから、観測数 $M < N$ をどのくらい小さくできるかが問題となる。これが、圧縮センシングの基本的な問題設定である。

式 (2) は観測に全く雑音がないような状況を想定しているが、現実の問題においては観測には雑音がつきものである。そこで上記の問題設定の自然な拡張として、観測に雑音が存在する状況を考えることができる。この場合、観測結果は

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n} \quad (3)$$

という形で与えられるものとして定式化される。 \mathbf{n} は観測雑音を表すベクトルである。また、現実の問題においては、推定の対象となる未知ベクトルが厳密にスパースであることを要請するのは条件として厳しすぎる、といったこともあろう。そのような状況を想定して、未知ベクトル \mathbf{x}_0 がスパースでなくても、 \mathbf{x}_0 の成分を絶対値が大きい順に並べて大きい方から K 個を残して他を 0 としたベクトルが推定できるのであれば、推定誤差を上から押さえることが可能である。以下では扱わないが、上記のような拡張も盛んに議論されている。

3. スパースベクトルの推定法

3.1 ℓ_0 再構成

観測行列 \mathbf{A} と観測結果 \mathbf{y} とから未知ベクトル \mathbf{x}_0 を推定する一つの方法は、線形制約 $\mathbf{y} = \mathbf{A}\mathbf{x}$ を満たし、かつスパースであるような \mathbf{x} を推定結果とするやり方であろう。ベクトル $\mathbf{x} \in \mathbb{R}^N$ の「0-ノルム」 $\|\mathbf{x}\|_0$ を \mathbf{x} の非ゼロ要素の個数として定義する^(注1)。このとき、

$$\hat{\mathbf{x}}^{(0)} = \arg \min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{subj. to} \quad \mathbf{y} = \mathbf{A}\mathbf{x} \quad (4)$$

すなわち、線形制約 $\mathbf{y} = \mathbf{A}\mathbf{x}$ のもとで 0-ノルムを最小化することによって、 \mathbf{x}_0 の推定値を得るというやり方を、具体的な推定法として考えることができる。この方法は ℓ_0 再構成と呼ばれることがある。

ℓ_0 再構成は優秀である。具体的には、 $M > K$ であれば、観測行列 \mathbf{A} の選び方に関してほとんどいつも ℓ_0 再構成は正しい推定結果を与えると考えてよい。このことは以下のように理解することができる。 ℓ_0 再構成における線形制約 $\mathbf{y} = \mathbf{A}\mathbf{x}$ は、 N 次元空間において \mathbf{A} の選び方に関してほとんどいつも余次元 M のアフィン部分空間 $\mathcal{L}(\mathbf{y}) = \{\mathbf{x} \in \mathbb{R}^N \mid \mathbf{y} = \mathbf{A}\mathbf{x}\}$ を定め

(注1)：これは厳密には数学的な意味での「ノルム」ではないことに注意せよ。より一般に、式 (5) で定義される「 p -ノルム」は、 $p < 1$ に対しては数学的な意味での「ノルム」ではない。また、4.3 の式 (27) で定義されるシャッテン p -ノルムについても同様である。

る。また、 K -スパースな \mathbf{x} の集合は、 K 個の基本ベクトルで張られる \mathbb{R}^N の K 次元部分空間 (${}_N C_K$ 個ある) の和集合 \mathcal{S}_K に対応する。 K -スパースなベクトル \mathbf{x}_0 は明らかに $\mathcal{L}(\mathbf{y}) \cap \mathcal{S}_K$ に属する。一方で、余次元 M のアフィン部分空間と K 次元部分空間とは、一般の位置にあれば $K < M$ のとき交点を持たない。したがって、 $K < M$ の場合には、よほど運が悪い \mathbf{A} を選んでしまった場合を除けば、 $\mathcal{L}(\mathbf{y}) \cap \mathcal{S}_K$ は空集合でなければ $\{\mathbf{x}_0\}$ に等しいと考えてよい。

ℓ_0 再構成にはしかし、重大な欠点がある。0-ノルムの最小化は一般に離散最適化問題となり、 ℓ_0 再構成は一般には NP 困難である。したがって、計算に必要な時間は N に関して指数関数的に増大すると考えられており、実用的に取り扱うことのできる問題の規模は厳しく制限されることになる。

3.2 ℓ_1 再構成

前節で見た ℓ_0 再構成の計算論的な困難を回避するために、ある種の緩和問題を考えよう。後の便宜のために、 $p \geq 0$ に対して p -ノルムを

$$\|\mathbf{x}\|_p = \begin{cases} \left(\sum_{i=1}^N |x_i|^p\right)^{1/p}, & p > 0 \\ \|\mathbf{x}\|_0, & p = 0 \end{cases} \quad (5)$$

によって定義する。そして、 ℓ_0 再構成における 0-ノルムを 1-ノルムで置き換えた 1-ノルム最適化問題

$$\hat{\mathbf{x}}^{(1)} = \arg \min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{subj. to} \quad \mathbf{y} = \mathbf{A}\mathbf{x} \quad (6)$$

を考えるのである。これは ℓ_1 再構成あるいは基底追跡 (basis pursuit) 法と呼ばれることがある。

ℓ_1 再構成には大きな利点がある。それは、 ℓ_1 再構成は線形計画問題として定式化することができ、単体 (シンプレックス) 法や内点法などのこれまでによく知られている解法を適用して効率的に解くことができることである。実際、補助変数 $\mathbf{t} = (t_1, \dots, t_N)^T$ を導入し、ベクトルに関する不等式は成分ごとに不等号が成立するものとする、 ℓ_1 再構成は

$$\min \sum_{i=1}^N t_i \quad \text{subj. to} \quad -\mathbf{t} \leq \mathbf{x} \leq \mathbf{t}, \quad \mathbf{y} = \mathbf{A}\mathbf{x} \quad (7)$$

と、線形等式/不等式制約のもとでの線形関数の最小化問題、すなわち線形計画問題として表現することができる。

もちろん、 ℓ_1 再構成が比較的容易に解けるからといって、得られた解が良い解でなければ意味がない。しかし驚くべきことに、 ℓ_1 再構成をスパースベクトルの推定に適用すると、ある条件下では $M < N$ であっても完全に正しい推定結果が得られる、すなわち $\hat{\mathbf{x}}^{(1)} = \mathbf{x}_0$ となることがあるのである。図 1 は、二次元平面において 1-スパースなベクトル (赤丸) を一次元の観測 (青の直線) に基づいて推定する問題を図示したものである。緑の正方形は、1-ノルム一定の条件を表しており、観測から定まる線形制約のもとで 1-ノルムを最小とする問題は、幾何学的には緑の正方形を青の直線と接触する条件のもとでなるべく小さくすることに対応しており、そのときの両者の接触点が ℓ_1 再

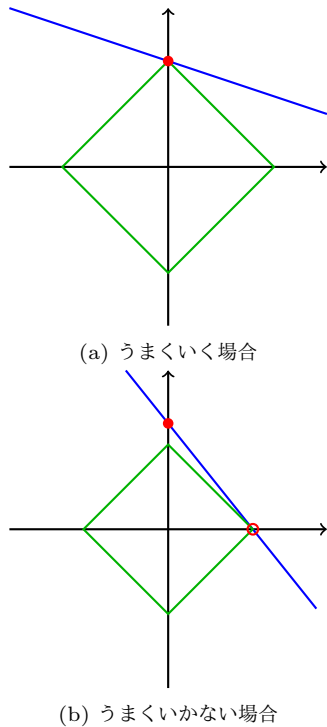


図 1 l_1 再構成 ($N = 2, K = 1, M = 1$ の場合)

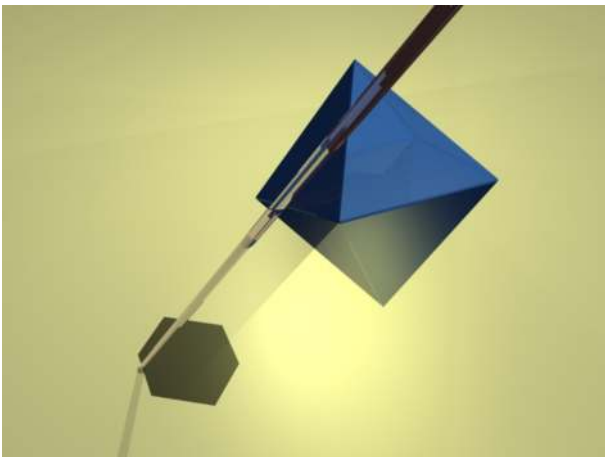


図 2 l_1 再構成 ($N = 3, K = 2, M = 2$ の場合)

構成の解に対応する。図 1 (a) は l_1 再構成が正しい推定結果を与える場合を表している。図 2 は、三次元空間における 2-スパーズなベクトルを二次元の観測に基づいて l_1 再構成によって推定している状況を示している。図の正八面体は 1-ノルム一定の条件を表しており、観測から定まる線形制約 $\mathcal{L}(\mathbf{y})$ は図中では棒として示されている。正八面体を徐々に大きくしていき、初めて棒に接触したときに、接触点が l_1 再構成の解を与える。

もちろん、 l_1 再構成が常に正しい解を与えるわけではなく、条件によっては図 1 (b) に示したように誤った結果を与えることもある。しかし後に見るように、適切な条件の下では図 1 (a) のような状況が非常に高い確率で実現されることを示すことができ、 l_1 再構成がスパースベクトルの推定に対する有用な手段を与えることになる。

3.3 1-ノルム最適化の歴史

3.3.1 概要

1-ノルムの最適化に基づく推定の方法論はガリレオやラプラスにまでさかのぼる長い歴史を持つといわれており⁽¹⁾、いろいろな分野で応用がなされたり、その性質について議論されたりしてきた。本節では、幾つかの具体例を示すことにより 1-ノルム最適化にかかわる研究の歴史を概観する。

3.3.2 Logan の現象

連続時間信号を復元する以下のような問題を考える。周波数帯域が周波数軸上で $\Omega \subset \mathbb{R}$ に制限されている信号 $f_0(t)$ に雑音 $n(t)$ が重畳して受信される状況を考え、受信信号を

$$g(t) = f_0(t) + n(t) \quad (8)$$

とおく。ただし、雑音は時間的に局在しており、時間軸上のある小さい集合 $T \subset \mathbb{R}$ においてだけ非ゼロの値をとるものとする。また、

$$\|n(t)\|_1 = \int |n(t)| dt \quad (9)$$

は有限であるものとする。

1-ノルムの最小化

$$\hat{f}(t) = \arg \min_{f \in \mathcal{B}(\Omega)} \|g(t) - f(t)\|_1 \quad (10)$$

によって、原信号 $f_0(t)$ の推定を試みるものとする。 $\mathcal{B}(\Omega)$ は、 Ω に帯域制限された絶対可積分な信号の集合である。このとき、ある条件下では原信号の完全な復元が可能であることを主張する以下の定理が成り立つ。

[定理 1] $|T||\Omega| \leq 1/4$ であれば、条件を満たす任意の雑音に対して $\hat{f}(t) = f_0(t)$ 。

証明は文献 (2) を参照されたい。この定理の驚くべき点は、条件を満たしてさえいれば、SN 比のような指標で見たときに雑音は幾ら大きくてもよいことである。定理が示す事実は、Logan が 1965 年にまとめた学位論文において $\Omega = \{\omega \leq \omega_0\}$ の場合について証明を与えたことにちなんで、Logan の現象と呼ばれている。

3.3.3 反射法地震探査

地層の境界で地震波が反射されるという性質を利用して、地表で人工的に地震波を発生させて反射波を観測することにより地層の構造などを推定することができる。このような方法は、反射法地震探査と呼ばれている。地表で発生させる地震波の時間波形を $f(t)$ 、地面で観測したときの地層のインパルス応答を $r(t)$ と置くと、観測される反射波はこれらの畳み込み $y(t) = (f * r)(t)$ によって与えられる。この関係式は、時間領域でサンプリングすることによって、行列とベクトルとを使って $\mathbf{y} = \mathbf{F}\mathbf{r}$ と表現することができる。 \mathbf{y}, \mathbf{F} から \mathbf{r} を推定することで、地層の構造についての情報を得ることができる。しかし、 $f(t)$ は帯域制限された波形であることが多く、そのような場合には行列 \mathbf{F} は悪条件となったりランク落ちしたりするため、 \mathbf{r} の推定が不安

定になりがちである。

反射法地震探査の分野では、1-ノルムに基づく推定法が有効であることが1970年代ごろから認識され、スパースな推定結果を与える頑健なアプローチとして例えば

$$\min_{\mathbf{r}} (\|\mathbf{r}\|_1 + \lambda \|\mathbf{y} - F\mathbf{r}\|_1) \quad (11)$$

という規準^{(3), (4)}や、

$$\min_{\mathbf{r}} (\|\mathbf{r}\|_1 + \lambda \|\mathbf{y} - F\mathbf{r}\|_2^2) \quad (12)$$

といった規準⁽⁵⁾などが研究されてきた。

3.3.4 Lasso

データの集合 $\{\mathbf{u}_i, v_i\} \in \mathbb{R}^N \times \mathbb{R}; i = 1, \dots, M$ から線形回帰式 $v = \mathbf{c}^T \mathbf{u}$ の係数ベクトル $\mathbf{c} \in \mathbb{R}^N$ を推定する問題を考える。回帰式の残差二乗和はベクトル $\mathbf{v} = (v_1, v_2, \dots, v_M)^T$, 及び行列

$$U = \begin{pmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \vdots \\ \mathbf{u}_M^T \end{pmatrix} \quad (13)$$

を使って、

$$\sum_{i=1}^M (v_i - \mathbf{u}_i^T \mathbf{c})^2 = \|\mathbf{v} - U\mathbf{c}\|_2^2 \quad (14)$$

と表すことができる。これを係数ベクトル \mathbf{c} について最小化することで \mathbf{c} を定めるのが、通常の最小二乗法である。

しかし、 \mathbf{u} の N 個の成分の中には、 v の値を予測するのに役に立たない成分も幾つかは混ざっているかもしれない。そのような場合には、対応する係数ベクトル \mathbf{c} の成分が0となるようにしたい。それにより、推定結果のばらつきを減らすことができるという利点があるからである。Tibshirani⁽⁶⁾は、このような性質を有する係数ベクトルの推定方法として「lasso」と呼ばれる手法を1996年に提案した。lassoは「least absolute shrinkage and selection operator」の頭文字であり、

$$\min_{\mathbf{c}} \|\mathbf{v} - U\mathbf{c}\|_2^2 \quad \text{subj. to} \quad \|\mathbf{c}\|_1 \leq a \quad (15)$$

という最適化問題の解として係数ベクトル \mathbf{c} を推定する方法である。 $a > 0$ は制御パラメータである。1-ノルムを制約条件として使うことによって、推定結果として得られる係数ベクトルがスパースになりやすくなるため、 \mathbf{u} の不要な成分の除去が実現されるというわけである。なお、式(15)における1-ノルム $\|\mathbf{c}\|_1$ を2-ノルムの二乗 $\|\mathbf{c}\|_2^2$ で置き換えたものは、縮小推定の一つとして知られているリッジ回帰と呼ばれる方法に相当する。

なお、石川真澄^{(7), (8)}は、非線形回帰モデルの一つである多層パーセプトロンを学習させる問題において、パラメータのスパース性を促すために1-ノルムを用いた「忘却つき構造学習」という学習則をlassoに先立つ1980年代末に既に提案している。

3.3.5 画像からの雑音除去

画像に雑音が含まれているような場合に、雑音成分だけを除

去して雑音のないきれいな画像を得ることができれば、画像を扱う様々な応用において有用であろう。画像にはある種の空間的構造があると期待される一方で、雑音にはそのようなものがないと考えられるので、このような両者の相違点に注目して雑音成分を除去するのが、基本的な考え方である。自然画像の場合には、撮影された物体の多くは明りようで概ね滑らかな表面を持つため、それに対応して、画像の画素ごとの輝度値、色相値などは物体境界などを除けば概ね滑らかに空間変動していると考えるのが妥当であろう。

以下では簡単のために濃淡画像に話を限定する。 (x, y) を画素の座標とし、雑音なしの画像を $f_0(x, y)$ と置こう。雑音が含まれる観測画像を

$$g(x, y) = f_0(x, y) + n(x, y) \quad (16)$$

とする。 $n(x, y)$ は画像の雑音成分を表している。雑音が含まれる画像 $g(x, y)$ から、雑音なしの画像 $f_0(x, y)$ を推定するのが、ここでの問題である。2-ノルムに基づく基本的なアプローチとしては、例えば

$$\hat{f} = \arg \min_f \left[\|g - f\|_2^2 + \lambda \int (f_{xx} + f_{yy})^2 dx dy \right] \quad (17)$$

といった形で推定画像 $\hat{f}(x, y)$ を求めるアプローチ⁽⁹⁾がある。 f の下添字は、その変数に関する偏導関数を表す。右辺の第一項目は観測画像 g に対する適合度を二乗誤差によって表しており、第二項目は推定画像の空間的な滑らかさを表している。被積分関数は平均曲率の二乗に相当する。空間的な滑らかさに対する定量的な尺度としては、二次変動

$$\int (f_{xx}^2 + 2f_{xy}^2 + f_{yy}^2) dx dy \quad (18)$$

もよく使われる。この被積分関数は、主曲率の二乗和に相当する。このように、適合度及び滑らかさを2-ノルムに基づくものとするので、対応するオイラー・ラグランジュ方程式が線形となるという利点を享受できる。

Rudinら⁽¹⁰⁾は、より自然な推定画像を与える雑音除去法として、1-ノルム的な性質を有する全変動

$$\int \sqrt{f_x^2 + f_y^2} dx dy \quad (19)$$

を滑らかさの尺度とする手法を1992年に提案した。対応するオイラー・ラグランジュ方程式は非線形となるが、2-ノルムに基づく手法と比較するとより良好な推定結果が得られるとして広く使われている。

4. 解 析

4.1 概 要

ランダム行列 A に対する ℓ_1 再構成の議論は、Candes と Tao^(注2) の研究や Donoho の研究をきっかけとして研究者の注目を集め、更に彼らの論文^{(11), (12)}が2008年にそろってIEEE

(注2) : 2006年にフィールズ賞を受賞した Terence Tao である。

情報理論サイエティの論文賞を受賞したことなども相まって、現在では理論、応用を問わず幅広い研究者が関連研究に参入し、多数の論文が継続的に発表される状況が続いている。本稿では応用については触れないこととし、圧縮センシングの数理的な側面に議論を限定するが、それでも関連文献をすべてカバーすることは筆者であれ誰であれ個人の力量の範囲を軽く超えていると思われる。本章では、 \mathbf{x}_0 を正しく推定するためには観測数 M はどの程度大きければよいのか、という圧縮センシングの基本的な問題について述べる。2. で議論したように、 \mathbf{x}_0 のスパースさを前提としないならば M は N 以上である必要があり、一方で \mathbf{x}_0 が K -スパースであり、更にどの成分が 0 でない値をとるかが分かっている場合には M は K 以上であればよい。圧縮センシングの問題はこれらの二つの問題の間に位置するものと考えられるから、必要な観測数 M は、 K よりは大きく N よりは小さいものと推測できる。以下では、この問題に関する Candes と Tao の基本的な研究成果、並びに Donoho と Tanner の研究成果について概説する。

4.2 Candes-Tao

4.2.1 制限等長性

まず、観測行列 A が与えられたとしたときに、スパースな未知ベクトルに対して ℓ_1 再構成が正しい推定結果を与えるかどうかを議論する。ある K' -スパースなベクトル \mathbf{x} に対して $A\mathbf{x} = \mathbf{0}$ になってしまうようなことがあれば、 $\mathbf{y} = A\mathbf{x}_0 = A(\mathbf{x}_0 + \mathbf{x})$ が成り立つ。 $M < N$ であるとき、 $M \times N$ 行列 A に対して $A\mathbf{x} = \mathbf{0}$ となるベクトル \mathbf{x} を必ず見つけることができるが、ベクトル \mathbf{x} が K' -スパースであれば $\mathbf{x}_0 + \mathbf{x}$ は $(K + K')$ -スパースであるから、 ℓ_1 再構成が正しい推定結果を与えることを期待するには、スパースなベクトルに限ったときに $A\mathbf{x} = \mathbf{0}$ となるベクトル \mathbf{x} がないことが望ましい。逆にもし、線形変換 A がスパースなベクトルの長さを余り変えない、という性質を有しているならば、 $A\mathbf{x} = \mathbf{0}$ となるスパースなベクトル \mathbf{x} はなく、したがって ℓ_1 再構成がうまく機能することを期待できそうである。このような性質を具体的に表現するのが、制限等長性 (restricted isometry property; RIP) と呼ばれる概念である。

$1 \leq K \leq N$ に対して、 A の等長性定数 δ_K を、不等式

$$(1 - \delta)\|\mathbf{x}\|_2^2 \leq \|A\mathbf{x}\|_2^2 \leq (1 + \delta)\|\mathbf{x}\|_2^2 \quad (20)$$

が任意の K -スパースなベクトル \mathbf{x} に対して成り立つような最小の δ の値として定義する。

A がもし $N \times N$ の直交行列であれば、等長性定数は K の値によらず 0 である。観測行列 A の列を $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N$ と置いたとき、これらの N 個のベクトルから K 個以下を任意に選び出したものがどのくらい正規直交系からずれているかを等長性定数 δ_K は表していると考えられる。

観測行列 A の制限等長性と ℓ_1 再構成の成否とを具体的に関連づけることが可能である。例えば以下が成り立つ⁽¹³⁾。

[定理 2] 観測行列 A に対して、 $\delta_{2K} < \sqrt{2} - 1$ が成り立つような $K \geq 1$ が存在するとき、任意の K -スパースな \mathbf{x}_0 に対し

て ℓ_1 再構成は正しい推定結果を与える。

4.2.2 ランダム行列

ℓ_1 再構成によって少数の観測から正しい推定結果を得ようとするのであれば、なるべく大きい K に対してなるべく小さい等長性定数 δ_K を与える観測行列 A を選ぶことができればよい。しかし、観測行列 A を適切に選ぶのは恐らく大変なので、Candes と Tao⁽¹⁴⁾ は、観測行列 A をランダム行列としたときの等長性定数がどうなるかを調べ、概略以下のような結果を得た。

[定理 3] $M \times N$ 行列 A の各要素を独立に平均 0、分散 $1/M$ の正規分布に従って定める。このとき、 $\alpha = M/N$ を有限として M, N を大きくする極限において、 $\rho = K/N$ がある値 $\rho_c(\alpha)$ より小さければ、 N に関して指数的に 1 に漸近する確率で条件 $\delta_{2K} < \sqrt{2} - 1$ が成立する。

前節の定理と合わせれば、観測行列をランダム行列 A とした場合に、 M 個の観測結果から K -スパースなベクトルを推定する問題は、 $M = \alpha N$ に応じて $K = \rho N$ が $\rho_c(\alpha)N$ より小さければ ℓ_1 再構成によって正しい推定が高い確率でなされることが結論される。この結果は、 ℓ_1 再構成の有用性を理論的に保証する結果として大きなインパクトを与えた。

証明のためには、Marčenko と Pastur⁽¹⁵⁾ に始まるある種のランダム行列の固有値の漸近分布に関する結果が使われる。行列 A から L 個の列を選び出して作られる $M \times L$ 行列を \hat{A} と置こう。 \hat{A} の最大特異値 $\sigma_{\max}(\hat{A})$ 、最小特異値 $\sigma_{\min}(\hat{A})$ は $\gamma = L/M \leq 1$ を有限に保ちつつ $L, M \rightarrow \infty$ とする極限においてそれぞれ

$$\sigma_{\min}(\hat{A}) \rightarrow 1 - \sqrt{\gamma} \quad \text{a.s.}$$

$$\sigma_{\max}(\hat{A}) \rightarrow 1 + \sqrt{\gamma} \quad \text{a.s.}$$

となることはよく知られている (Marčenko-Pastur 則)。更に精密な結果として、定数 $t > 0$ に対して

$$P\left(\sigma_{\min}(\hat{A}) < 1 - \sqrt{\gamma} + o(1) - t\right) \leq e^{-Mt^2/2} \quad (21)$$

$$P\left(\sigma_{\max}(\hat{A}) > 1 + \sqrt{\gamma} + o(1) + t\right) \leq e^{-Mt^2/2} \quad (22)$$

が成り立つことも知られている⁽¹⁶⁾。 A の等長性定数は \hat{A} の特異値に

$$\delta_K = \max_{\hat{A}} \left\{ \left[\sigma_{\max}(\hat{A}) \right]^2 - 1, 1 - \left[\sigma_{\min}(\hat{A}) \right]^2 \right\} \quad (23)$$

という形で関連づけることができる。右辺は A から $L (\leq K)$ 個の列を選び出して \hat{A} を構成するすべてのやり方に関しての最大値をとる。この結果と和事象に対する確率の上界評価とを組み合わせることで、例えば、定数 $\varepsilon > 0$ 及び十分大きい K, N に対して等長性定数 δ_K に関する

$$P\left(\sqrt{1 + \delta_K} > 1 + (1 + \varepsilon)f(\rho)\right) \leq 2e^{-NH(\rho)\varepsilon/2}, \quad \rho = \frac{K}{N} \quad (24)$$

という形の確率的上界評価を得ることができる。ただし

$$f(\rho) := \sqrt{\rho/\alpha} + \sqrt{2H(\rho)/\alpha} \quad (25)$$

であり、 $H(\rho) = -\rho \log \rho - (1-\rho) \log(1-\rho)$ はエントロピー関数である。この結果を使い、 $\delta_{2K} < \sqrt{2} - 1$ が 1 に近い確率で成り立つための条件を $\rho < \rho_c(\alpha)$ という形で導くことができる。以上が、上記の定理を導く Candes-Tao の議論の概略である。

Candes らはまた、観測行列 A がランダム行列ではなく離散フーリエ変換行列からランダムに行を選び出して作られる場合についても同様の議論を行っている。

4.3 低ランク行列の再構成

関連する話題として、行列再構成の問題についても言及しておきたい。未知の $N_1 \times N_2$ 行列 X_0 に対して、線形観測を M 回行うものとし、観測結果を $\mathbf{y} = (y_1, y_2, \dots, y_M)^T$, $y_i = \langle A_i, X_0 \rangle$, $i = 1, 2, \dots, M$ とする。ただし、 $\langle X, Y \rangle = \text{tr}(X^T Y)$ である。観測行列 $\{A_1, A_2, \dots, A_m\}$ は既知であるものとし、観測を線形作用素 $\mathcal{A} : \mathbb{R}^{N_1 \times N_2} \mapsto \mathbb{R}^M$ を使って $\mathbf{y} = \mathcal{A}X_0$ と表すことにする。未知の行列 X_0 の階数 $\text{rank } X_0$ が小さいことが分かっているとき、 \mathbf{y} と \mathcal{A} とから X_0 を推定するのが、ここでの行列再構成の問題である。

ℓ_0 再構成に相当するアプローチは、

$$\hat{X}^{(0)} = \arg \min_X \text{rank } X \quad \text{subj. to } \mathbf{y} = \mathcal{A}X \quad (26)$$

のように定式化できるが、これは NP 困難であることが知られている。圧縮センシングの場合と同様に、ここでも上記の問題の緩和を考える。 X の特異値を $\sigma_1, \sigma_2, \dots$ で表すと、 X のシャッテン p -ノルムは

$$\|X\|_p = \left(\sum_{i=1}^{\min\{N_1, N_2\}} \sigma_i^p \right)^{1/p} \quad (27)$$

によって定義される。式 (26) の緩和問題として、 X の階数 $\text{rank } X$ をシャッテン 1-ノルム (トレースノルム、核型ノルムと呼ばれることもある) で置き換えた問題

$$\hat{X}^{(1)} = \arg \min_X \|X\|_1 \quad \text{subj. to } \mathbf{y} = \mathcal{A}X \quad (28)$$

を考えることができる。行列の階数は非ゼロの特異値の個数に等しいから、これをシャッテン 1-ノルムで置き換えることが ℓ_0 再構成に対する ℓ_1 再構成と同様の緩和問題になっていることが分かる。この 1-ノルム最適化問題は、 X に対称行列 Y, Z を変数として加えた半正定値計画問題

$$\begin{aligned} \min_{X, Y, Z} \text{tr} \begin{pmatrix} Y & O \\ O & Z \end{pmatrix} \\ \text{subj. to } \begin{pmatrix} Y & X \\ X^T & Z \end{pmatrix} \succeq 0, \quad \mathbf{y} = \mathcal{A}X \end{aligned} \quad (29)$$

として表現することができる (記号「 $\succeq 0$ 」は左辺の対称行列が半正定値であることを意味する)。半正定値計画問題については 1990 年代から系統的な研究が進展し、現在では既知の解法を適用することで効率的に解を得ることが可能である⁽¹⁷⁾。

圧縮センシングの場合と同様、我々が知りたいのは、比較的容易に解ける問題 (28) の解が X_0 と一致するためには、観測数 M はどのくらい必要であるか、という問題に対する答えである。Recht ら⁽¹⁸⁾はこの問題に取り組んだ。まず、線形作用素 \mathcal{A} に対する制限等長性を議論するために、 \mathcal{A} の等長性定数 δ_r を、階数が r 以下の任意の $N_1 \times N_2$ 行列 X に対して不等式

$$(1 - \delta) \|X\|_2^2 \leq \|\mathcal{A}X\|_2^2 \leq (1 + \delta) \|X\|_2^2 \quad (30)$$

が成り立つような最小の δ の値として定義する。すると、例えば以下のような結果を導くことができる。

[定理 4] 線形作用素 \mathcal{A} に対して $\delta_{5r} < 1/10$ が成り立つならば、階数が r 以下の任意の行列 X_0 に対して式 (28) は正しい推定結果を与える。

この結果に基づいて、線形作用素 \mathcal{A} をランダムに定めた場合には、観測数 M が $O(r(N_1 + N_2) \log(N_1 N_2))$ であれば高い確率で式 (28) によって階数 r の行列 X_0 が正しく推定できることなどが示されている。

階数 r の $N_1 \times N_2$ 行列の自由度は $r(N_1 + N_2 - r)$ である。このことは行列の特異値分解を考えれば直ちに理解できる。上述の Recht らの結果は、正しく行列再構成を行うためには因子 $\log(N_1 N_2)$ に相当する分だけ観測数を上乘せれば十分であることを示していると解釈することができる。更に進んだ解析⁽¹⁹⁾によれば、この因子は除くことができ、結果として階数 r の行列の自由度に比例する程度の観測数があれば、高い確率で式 (28) によって階数 r の行列 X_0 が正しく推定できることも分かっている。

低ランク行列の再構成は、共起行列の特異値分解に基づくデータ解析手法である潜在意味解析 (latent semantic analysis; LSA)⁽²⁰⁾ や、その応用である協調フィルタリングとの関連においても注目を集めている。このような観点からは、再構成の対象を行列から更に高階のテンソルとする一般化も重要であり、多くの研究がなされている。

4.4 Donoho-Tanner

4.4.1 問題 ±, 問題 +

圧縮センシングの問題に戻り、本節では Donoho と Tanner による結果について、その概略を述べる。

未知ベクトル \mathbf{x}_0 が K -スパースであるだけでなく、非ゼロ要素は必ず正の値をとる、という制約を更に付加した問題設定を考えることができる。この問題を「問題 +」と呼び、ベクトル \mathbf{x}_0 の要素に対する非負値制約のないもとの問題を「問題 ±」と呼ぶことにしよう。問題 ± と同様に、問題 + に対しても ℓ_1 再構成

$$\hat{\mathbf{x}}^{(1)} = \arg \min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{subj. to } \mathbf{y} = A\mathbf{x}, \quad \mathbf{x} \geq \mathbf{0} \quad (31)$$

を考えることができる。これもやはり線形計画問題

$$\min \sum_{i=1}^N t_i \quad \text{subj. to } \mathbf{0} \leq \mathbf{x} \leq \mathbf{t}, \quad \mathbf{y} = A\mathbf{x} \quad (32)$$

として表現することができるため、比較的容易に解くことがで

きる。

4.4.2 ランダム高次元幾何学

ℓ_1 再構成がうまくいくかどうかは、幾何学的に特徴づけることが可能である。例えば図 2 では、線形制約 $\mathcal{L}(\mathbf{y})$ を表す棒の向きと平行に光を当てており、この光によって正八面体が六角形の影を作っている。棒と正八面体とは左手前の辺で互いに接触しており、それに対応して、棒の影と六角形の影とが接触している。棒と正八面体との接触点が \mathbf{x}_0 に対応していれば、図の光の向きに対応する観測行列に基づいて ℓ_1 再構成を行うことにより正しい解が得られる。言い換えれば、 \mathbf{x}_0 が 2-スパースであるとき、 \mathbf{x}_0 のスパースさの構造に対応する正八面体の辺が「影」においても辺であれば、 ℓ_1 再構成は正しい推定結果を与えるのである。

より一般に ℓ_1 再構成がうまくいくかどうかは、ある種の N 次元多胞体^(注3) Q を線形変換 A によって M 次元空間に射影したとき、 Q の K 次元面が射影先で面をなすかどうか、という幾何学的性質と対応している。Donoho と Tanner はこの対応に着目し、後者の幾何学の問題を考察することで ℓ_1 再構成がうまくいくために必要な観測数に関する検討を行った。

N 次元多胞体 Q として具体的に何を考えなければならないかは、問題が \pm であるか $+$ であるかによって違ってくる。問題 \pm に対しては

$$Q = C^N := \{\mathbf{x} \in \mathbb{R}^N; \|\mathbf{x}\|_1 \leq 1\} \quad (33)$$

をとる。例えば、図 2 に見られるように、 C^3 は正八面体である。一方、問題 $+$ に対しては、 T^{N-1} を \mathbb{R}^N の N 個の基本ベクトル $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_N\}$ の凸包として定義される単体としたとき、 Q は原点と T^{N-1} との凸包として定義される単体 T_0^N とする。その上で、ランダム行列 A を使って多胞体 Q を \mathbb{R}^M へ線形射影して得られる多胞体 AQ の性質を議論する。

例えば、 F を多胞体 Q の K -次元面としたとき、 AF は多胞体 AQ の面となるかどうか、ということを知ることができる。直観的には、射影先の空間の次元 M が余り小さいと、 Q の面は射影によって AQ の内側に「取り込まれ」てしまうし、更に次元 K が大きい面の方が射影によって「取り込まれ」やすいと考えられる。 Q として $(N-1)$ 次元単体 T^{N-1} をとった最も基本的な問題設定においても、 Q の面が射影によって内側に「取り込まれ」るかどうかは、問題の特徴づける次元 N, M, K だけでなく、 F として Q のどの面を考えるか、及びランダム行列 A の実現値がどうであるかといったことに依存する。しかし、 N, K が共に十分大きければ、 $\rho = K/N$ に対して $\alpha = M/N$ がおよそ

$$\frac{\rho}{-2 \log \rho} \quad (34)$$

よりも大きければ、 A のランダムさに関して非常に高い確率で答は「イエス」であり、逆に小さければ高い確率で「ノー」であることが分かっている。

また、少し要求を厳しくして、多胞体 Q のすべての K -次元面 F に対して、 AF は多胞体 AQ の面となるかどうか、ということを知ることができる。 $Q = T^{N-1}$ に対しては、 N, K が共に十分大きければ、 $\alpha = M/N$ がおよそ

$$\frac{\rho}{-2e \log(2\sqrt{\pi\rho})} \quad (35)$$

よりも大きければ、 A のランダムさに関して非常に高い確率で答は「イエス」であり、逆に小さければ高い確率で「ノー」であることが分かっている。

多胞体 Q の K -次元面の個数を $f_K(Q)$ と表記する。また、 A のランダムさに関する期待値を $\mathbb{E}_A(\dots)$ と表す。 ℓ_1 再構成に直接関係するのは、多胞体 Q の K -次元面の個数が M 次元空間へのランダムな射影によって減るのか、減らないのか、という問題である。具体的には、以下の二つの定理が重要である。射影による面の個数の減少の程度を、第 1 の定理は射影前の個数 $f_K(Q)$ との比において相対的に、第 2 の定理は実際の個数においてそれぞれ議論しているのが、両者の相違点である。

[定理 5] $*$ $\in \{\pm, +\}$ とする。 $\alpha = M/N, \rho = K/N$ を有限に保ちつつ N, M, K を無限大とする極限において、問題 $*$ に対して「弱しきい値」 $\alpha_W^*(\rho)$ が存在し、 $\alpha > \alpha_W^*(\rho)$ であれば $\mathbb{E}_A[f_K(AQ)]/f_K(Q) = 1 - o(1)$ が成り立ち、逆に $\alpha < \alpha_W^*(\rho)$ であればある定数 $\varepsilon > 0$ に対して $\mathbb{E}_A[f_K(AQ)]/f_K(Q) < 1 - \varepsilon$ が成り立つ。

[定理 6] $*$ $\in \{\pm, +\}$ とする。 $\alpha = M/N, \rho = K/N$ を有限に保ちつつ N, M, K を無限大とする極限において、問題 $*$ に対して「強しきい値」 $\alpha_S^*(\rho)$ が存在し、 $\alpha > \alpha_S^*(\rho)$ であれば $f_K(Q) - \mathbb{E}_A[f_K(AQ)] = o(1)$ が成り立ち、逆に $\alpha < \alpha_S^*(\rho)$ であれば $f_K(Q) - \mathbb{E}_A[f_K(AQ)] \rightarrow \infty$ が成り立つ。

問題 $+$ に対する弱しきい値の存在は、Vershik と Sporyshev⁽²¹⁾ によって証明された。上記の二つの定理、及び強弱それぞれのしきい値の具体的な形を導いたのは Donoho⁽²²⁾ 及び Donoho と Tanner⁽²³⁾ である。これらの解析に際して必要となる数学的な道具立てに関しては、例えば積分幾何学の教科書⁽²⁴⁾ を参照して頂きたい。

Donoho と Tanner は、強弱それぞれのしきい値が ℓ_1 再構成がうまくいくために必要な観測数の下限を与えることを見いだした。具体的には、以下が成り立つ。

[定理 7] $*$ $\in \{\pm, +\}$ とする。 $\alpha = M/N, \rho = K/N$ を有限に保ちつつ $N, M, K \rightarrow \infty$ とする極限において、ほとんどすべての K -スパースなベクトル \mathbf{x}_0 に対してランダムな観測行列 A に基づく ℓ_1 再構成が正しい推定結果を与える確率が 1 に近づくための条件は、 $\alpha > \alpha_W^*(\rho)$ である。

[定理 8] $*$ $\in \{\pm, +\}$ とする。 $\alpha = M/N, \rho = K/N$ を有限に保ちつつ $N, M, K \rightarrow \infty$ とする極限において、任意の K -スパースなベクトル \mathbf{x}_0 に対してランダムな観測行列 A に基づく ℓ_1 再構成が正しい推定結果を与える確率が 1 に近づくための条件は、 $\alpha > \alpha_S^*(\rho)$ である。

弱しきい値 $\alpha_W^*(\rho)$ については、媒介変数を用いた具体的な表式が分かっている。媒介変数を $t \in [0, \infty)$ とすると、問題 \pm に対しては

(注3)：多胞体は、三次元空間における多面体の概念を一般の次元に拡張したものである。

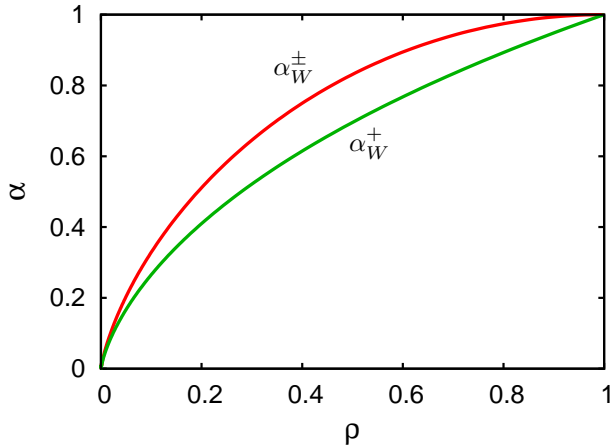


図3 弱しきい値

$$\alpha^{-1} = 1 + \sqrt{\frac{\pi}{2}} t e^{t^2/2} [1 - 2Q(t)]$$

$$\frac{\rho}{1-\rho} = 2 \left(\frac{e^{-t^2/2}}{t\sqrt{2\pi}} - Q(t) \right) \quad (36)$$

である。また、問題+に対しては

$$\alpha^{-1} = 1 + \sqrt{\frac{\pi}{2}} t e^{t^2/2} [2 - 2Q(t)]$$

$$\frac{\rho}{1-\rho} = \frac{e^{-t^2/2}}{t\sqrt{2\pi}} - Q(t) \quad (37)$$

である。ただし、

$$Q(t) = \int_t^\infty \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx \quad (38)$$

である。それぞれを図示すると、図3のようになる。

興味深いことに、上記の弱しきい値は、情報統計力学に基づく ℓ_1 再構成の解析によって得られるしきい値^{(25), (26)}と全く同一である。この事実は、情報統計力学とランダム高次元幾何学との間の深い関連を示唆するものではないかと筆者は考えている。

5. 構造的知識の利用

研究の展開の一例として、筆者らの最近の研究成果について概略を御紹介したい。前章までの話では、未知ベクトル \mathbf{x}_0 には特に構造を仮定せず、いわば一様にスパースであるようなベクトルを対象としていた。一方で、現実の問題は何らかの構造を持っている場合も少なくないと考えられる。ここで具体的に考えたいのは、例えば未知ベクトル \mathbf{x}_0 がブロック構造を持ち、ブロックごとにスパースさの程度が異なるような状況である。

未知ベクトル \mathbf{x}_0 の推定に際してこのような構造的知識が入手可能であるものとして、それを活用して推定の性能を高めるにはどうすればよいだろうか。素朴な発想として、 ℓ_1 再構成における 1-ノルムに重みを導入し、スパースなブロックに属する成分には大きい重みを付け、逆にスパースでないブロックに属する成分に対しては重みを小さくすることによって、 ℓ_1 再構成によって得られる推定結果が所望の構造を持つよう促すというやり方が考えられよう。このような発想それ自体は決して新し

いものではない⁽²⁷⁾。筆者らは、重み付き ℓ_1 再構成においてどのように重みを定めればよいかという問題を検討した⁽²⁸⁾。

入手可能な構造的知識としては、未知ベクトル \mathbf{x}_0 の各成分 x_{0i} , $i = 1, \dots, N$ が非ゼロの値をとる確率 $\rho_i = P(x_{0i} \neq 0)$ がすべての i について既知であることを仮定する。また、重みベクトル $\mathbf{w} = (w_1, w_2, \dots, w_N)$, $w_i \geq 0$ を使って重み付き 1-ノルムを

$$\sum_{i=1}^N w_i |x_i| \quad (39)$$

と定義する。非ゼロ確率のベクトル $\boldsymbol{\rho} = (\rho_1, \rho_2, \dots, \rho_N)$ 及び重みベクトル \mathbf{w} を指定すれば、情報統計力学のアプローチによってしきい値 $\alpha_c(\boldsymbol{\rho}, \mathbf{w})$ を $\boldsymbol{\rho}, \mathbf{w}$ の関数として得ることができる。最適な重み付き ℓ_1 再構成を行うには、このしきい値 $\alpha_c(\boldsymbol{\rho}, \mathbf{w})$ がなるべく小さくなるように重みベクトル \mathbf{w} を定めればよい。

得られる結果は大変興味深いものである。問題土についての結果の概略を述べよう。まず、最適な重みは以下のようにして求めればよい。各成分に対して、 ρ_i から w_i を以下の非線形方程式によって定める。

$$\frac{\rho_i}{1-\rho_i} = 2 \left(\frac{e^{-w_i^2/2}}{w_i\sqrt{2\pi}} - Q(w_i) \right) \quad (40)$$

こうして得られる w_i を第 i 成分として持つベクトル \mathbf{w} が最適な重みベクトルとなる。すなわち、重みの最適化は成分ごとに独立に行うことができるのである。また、最適な重みを使った場合のしきい値は、

$$\alpha_c^{\text{opt}} = \frac{1}{N} \sum_{i=1}^N \alpha_W^\pm(\rho_i) \quad (41)$$

という形となる。右辺に現れる $\alpha_W^\pm(\rho)$ は、 \mathbf{x}_0 の各要素が非ゼロとなる確率が ρ である一様な系の弱しきい値 (36) である。図3から見て取れるように、弱しきい値 α_W^\pm は ρ の関数として上に凸であるから、イエンセンの不等式を想起すれば、重みを最適化すれば必ずしきい値を小さくできることが式 (41) から直ちに分かる。また、 $\alpha_W^\pm(\rho)$ という量は、確率 ρ で非ゼロ値をとる実数値確率変数に付随する、ある種の情報の量を表しているのではないかと考えることもできよう。更に、以上の解析結果をランダム高次元幾何学の文脈において再解釈することも可能である。

6. むすび

以上のように、圧縮センシングの数理は、最適化、ランダム行列、積分幾何学、情報統計力学などの多様な分野にまたがるという点で刺激的であり、数理科学の研究対象として非常に興味深いものである。本稿では触れなかったが、 ℓ_1 再構成を効率的に解くためのアルゴリズムの研究や、 ℓ_1 再構成よりも高い性能を有する定式化並びに解法の探求などの問題に関しても、数多くの研究がなされている。後者に関して、特に $p < 1$ に対する ℓ_p 再構成の理論的性能に関する議論を主に扱った日本語の解説⁽²⁹⁾があるので、併せて御参照頂きたい。また、線形観測に基づく

推定という問題の構造は、幅広い分野への応用を可能にするものと考えられ、実際に非常に多くの研究がなされているが、それらの応用についても本稿では全く触れることができなかった。

圧縮センシングに関する研究が国際的に非常に活発化している一方で、研究動向を概観できるような資料を見つけることはかえって難しくなっているのではないかと感じている。本稿が、圧縮センシングの数理に関心のある研究者が関連する研究に踏み出す足掛かりとして幾らかでも役立つことを願っている。

文 献

- (1) R. Gonin and A. H. Money, *Nonlinear L_p -norm Estimation*, Marcel Dekker, 1989.
- (2) D. L. Donoho and P. B. Stark, "Uncertainty principles and signal recovery," *SIAM J. Appl. Math.*, vol. 49, no. 3, pp. 906–931, June. 1989.
- (3) H. L. Taylor, S. C. Banks, and J. F. McCoy, "Deconvolution with the l_1 norm," *Geophysics*, vol. 44, no. 1, pp. 39–52, Jan. 1979.
- (4) J. F. Claerbout and F. Muir, "Robust modeling with erratic data," *Geophysics*, vol. 38, no. 5, pp. 826–844, Oct. 1973.
- (5) F. Santosa and W. W. Symes, "Linear inversion of band-limited reflection seismograms," *SIAM J. Sci. Stat. Comput.*, vol. 7, no. 4, pp. 1307–1330, Oct. 1986.
- (6) R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. R. Statist. Soc. B*, vol. 58, no. 1, pp. 267–288, 1996.
- (7) M. Ishikawa, "Learning of modular structured networks," *Artif. Intell.*, vol. 75, no. 1, pp. 51–62, May 1995.
- (8) M. Ishikawa, "Structural learning with forgetting," *Neural Netw.*, vol. 9, no. 3, pp. 509–521, April 1996.
- (9) B. R. Hunt, "The application of constrained least squares estimation to image restoration by digital computer," *IEEE Trans. Comput.*, vol. C-22, no. 9, pp. 805–812, Sept. 1973.
- (10) L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, nos. 1–4, pp. 259–268, Nov. 1992.
- (11) E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.
- (12) D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, April. 2006.
- (13) E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *C. R. Acad. Sci. I, Math.*, vol. 346, nos. 9–10, pp. 589–592, May 2008.
- (14) E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.
- (15) V. A. Marčenko and L. A. Pastur, "Distribution of eigenvalues for some sets of random matrices," *Mathematics of the USSR-Sbornik*, vol. 1, no. 4, pp. 457–483, April 1967.
- (16) M. Ledoux, *The Concentration of Measure Phenomenon*, Mathematical Surveys and Monographs, no. 89, American Mathematical Society, 2001.
- (17) S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.
- (18) B. Recht, M. Fazel, and P. A. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," [online] arXiv:0706.4138v1 [math.OA], 2007.
- (19) B. Recht, W. Xu, and B. Hassibi, "Null space conditions and thresholds for rank minimization," [online] arXiv:0809.1260v1 [math.OA], 2008.
- (20) S. Deerwester, S. T. Dumais, G. W. Fumas, T. K. Landauer, and R. Harshman, "Indexing by latent semantic analysis," *J. Am. Soc. Inf. Sci.*, vol. 41, no. 6, pp. 391–407, Sept. 1990.
- (21) A. M. Vershik and P. V. Sporyshev, "Asymptotic behavior of the number of faces of random polyhedra and the neighborliness problem," *Sel. Math. Sov.*, vol. 11, no. 2, pp. 181–201, 1992.
- (22) D. L. Donoho, "High-dimensional centrally-symmetric polytopes with neighborliness proportional to dimension," *Discrete & Comput. Geom.*, vol. 35, no. 4, pp. 617–652, 2006.
- (23) D. L. Donoho and J. Tanner, "Neighborliness of randomly-projected simplices in high dimensions," *Proc. Nat. Acad. Sci., USA*, vol. 102, no. 27, pp. 9452–9457, 2005.
- (24) R. Schneider and W. Weil, *Stochastic and Integral Geometry*, Springer, 2008.
- (25) Y. Kabashima, T. Wadayama, and T. Tanaka, "A typical reconstruction limit for compressed sensing based on L_p -norm minimization," *J. Stat. Mech.: Theory Exp.*, L09003, 2009.
- (26) D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proc. Nat. Acad. Sci., USA*, vol. 106, no. 45, pp. 18914–18919, 2009.
- (27) E. J. Candès, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted l_1 minimization," *J. Fourier Anal. Appl.*, vol. 14, nos. 5–6, pp. 877–905, 2008.
- (28) T. Tanaka and J. Raymond, "Optimal incorporation of sparsity information by weighted L_1 optimization," *Proc. 2010 IEEE Int. Symp. Inf. Theory*, Austin, Texas, U.S.A., June 2010, pp. 1598–1602; [online] arXiv:1001.1873v1 [cs.IT], 2010.
- (29) 平林 晃, "Compressed Sensing – 基本原理と最新研究動向 –, " 信学技報, CAS2009-11, VLD2009-16, SIP2009-28, pp.55-60, July 2009.

(平成 22 年 5 月 11 日受付)



田中利幸 (正員)

1988 東大・工・電子卒, 1990 同大学院工学系研究科電子工学専攻修士課程了, 1993 同大学院工学系研究科電子工学専攻博士課程了, 博士(工学). 都立大・工・電子・情報・助手, 講師, 助教授, 首都大東京・システムデザイン・助教授を経て, 現在, 京大大学院情報学研究所システム科学専攻教授, 情報理論等の研究に従事.