圧縮センシング

David L. Donoho, *IEEE会員。*

概要：（デジタル画像や信号）中の未知のベクトルで $\boldsymbol{R}^m$ *であると仮定し*、その一般的な線形関数を測定し、再構成することを計画している。既知の変換を用いた変換符号化により圧縮可能であることが知られており、ここで定義された非線形手順により再構成を行う場合、測定回数は.NETのサイズよりも劇的に小さくすることが可能である。従って、n m ある種の自然な画素を持つ画像は、その画素の大きさに対して m
n $O(m^{1/4}\log^{5/2}(m))$ の非適応非画素サンプル。
通常の画素サンプルとは異なり、忠実な復元が可能です。

具体的には，ある正規表現x(ウェーブレット，フーリエなど）またはタイトフレーム（カーブレット，ガボールなど）で疎な表現をしているとします．したがって，係数は，ℓ_p のボールに属します． $\leq$.その中で最も重要な係数はNの拡大により、誤差での再構成が可能です。$O(N\log(m))$*非適応*測定の設計（log( )）が可能である。
これにより、最も重要な係数を直接知っている場合と同程度の精度で再構成することができます。さらに、信号処理における線形プログラム（Basis Pursuit）*を解くこと*で、これらの重要な係数の良い近似値を測定値から抽出することができる。非適応測定は、基底/フレーム要素の「ランダムな」線形結合の特徴を持つ。我々の結果は、最適回復、-n幅、情報ベースの複雑さの概念を用いている。高次元ユークリッド空間におけるボールのGel'fand-ℓ_p -$\leq$ widthsを以下のケースで推定する。 また、Gel'f と -width に対して最適に近い部分空間を特定する基準を与える。そして、「ほとんどの」部分空間が最適に近いことを示し、凸最適化（Basis Pursuit）がこれらの最適に近い部分空間から情報を抽出する最適に近い方法であることを示す。

*索引用語-適応的*サンプリング、バナッハ空間のほぼ球形断面、基底追求、ランダム行列の固有値、ゲルファンド幅、情報に基づく複雑さ、統合されたセンシングと処理、最小ノルム分解、最適回復、部分空間の商定理、線形方程式の疎な解法。

I. イントロダクション

A 現代の技術主導型文明は、増え続けるデータを取得・活用することで、「誰もが」、取得したデータのほとんどが、ほとんど知覚的な損失なしに「捨てることができる」ことを知っています-音、画像、特殊技術データの非可逆圧縮形式が広く成功していることに照らしてみてください。ユビキタス圧縮という現象は、ごく自然な疑問を投げかける。「取得したデータのほとんどが捨てられてしまうのに、なぜわざわざ**すべての**データを取得する必要があるのか？という疑問が生じる。捨ててしまうのか？

つまり、非可逆圧縮によって「捨てられてしまう」データの部分を取得しない。さらに、このプロトコルは非適応型で並列化可能であり、信号や画像に関する知識（データが圧縮可能であるという知識以外）を事前に取得する必要はなく、アクティブセンシング戦略や適応的センシング戦略を導くために基本的なオブジェクトに関するいかなる「理解」も必要としないのである。圧縮センシングプロトコルで行われる測定は*ホログラフィック*であるため、単純なピクセルサンプルではなく、*非線形に*処理されなければならない。

この原理を応用すれば、測定時間の大幅な短縮、サンプリングレートの大幅な低下、アナログ／デジタル変換器のリソース使用量の削減が可能になるかもしれない。

A.変換圧縮の背景

我々の扱いは抽象的かつ一般的であるが，信号・画像処理の多くの場面で成立することが知られている，ある特別な仮定に依存している：*変換スパース性の*原理である．興味の対象はベクトル $x \in \boldsymbol{R}^m$ であり、それはサンプルやピクセルを持つ信号や画像である。そして、$\boldsymbol{R}^m$ のための正規直交基底 $(\psi_i : i = 1, \ldots, m)$ があるとする。これは、例えば、用途に応じて、正規ウェーブレット基底、フーリエ基底、または局所フーリエ基底であったりする。(後述するように、カーブレットやガボールフレームのようなタイトフレームへの拡張は無料である）。オブジェクトは、変換係数 $\theta_i = \langle x, \psi_i \rangle$ を持ち、これらは、 $0 < p < 2$ 、 $R > 0$ ある、という意味で、疎であることを仮定する。

$$\|\theta\|_p \equiv \left(\sum_i |\theta_i|^p\right)^{1/p} \leq R. \tag{I.1}$$

本論文では、*あたかも*直接データ取得が可能であるかのような圧縮データ取得プロトタイプを設計する。

2004年9月18日原稿受領、2005年12月15日改訂。
著者はStanford University, Stanford, CA 94305 USAのDepartment of Statisticsに所属している（e-mail: donoho@stanford.edu）。
Detection and Estimation誌のアソシエイトエディター、A. Høst-Madsen氏による通信。
デジタルオブジェクト識別子 10.1109/TIT.2006.871582

このような制約は、信号や画像の自然なクラスでは実際に従う。これが、変換符号化に基づく標準的な圧縮ツールの成功の主な理由である[1]。アイデアを定着させるために、ℓ_p 制約の2つの簡単な例を挙げる。

- *画像の境界変動モデル*。ここで画像の明るさは、単位正方形上の基礎関数 $f(x,y)$ $\leq$, $\leq$ として捉えられ、（本質的に）以下のように従う。

$1/n$ $1/n$ のピクセルを平均化した f の n^2 ピクセルのサンプルデータである。$m =$ ここではウェーブレット法を用いて、データを様々なスケールからの寄与の集合体として捉える。ここで $x^{(j)}$ はスケール j におけるデータの成分を表し、スケール j, containingelements $\left(\psi_i^j\right)$ におけるウェーブレットの正規直交基底を表すとする $3 \quad 4^j$。対応する係数は次のように $\|\theta^{(j)}\|_1 \leq 4R$ 。

$$\int_0^1 \int_0^1 |\nabla f| dxdy \leq R.$$

0018-9448/$20.00 © 2006 ieee

- *スペクトルのバンプ代数モデル*。ここで、スペクトル（質量スペクトルや磁気共鳴スペクトルなど）は、実線上の基底関数 $(f(i/n))$ のデジタルサンプルとしてモデル化されており、これは、位置、振幅、線幅が異なるいわゆるスペクトル線の重ね合わせになっている。形式的には

$$f(t) = \sum_{i=1}^{\infty} a_i g((t-t_i)/s_i).$$

ここで、パラメータ t_i は線分位置、a_i 振幅/極性、線幅 s_i で g あり、線分形状、例えばガウシアンを表しますが、他のプロファイルを考慮することも可能です。$\sum_i |a_i| \le R$ という制約を仮定する。これは、アプリケーションでは、エネルギーまたは全質量の制約を表す。再びウェーブレットの視点を取り、今回は特にスムースウェーブレットを使用する。データは様々なスケールからの分布の重ね合わせとして表すことができる。$x^{(j)}$ は、スケール j におけるスペクトルの成分を表し、要素を含むスケール 2^j におけるウェーブレットの直交基底を表すとする (ψ_i^j)。対応する係数は、再び次のようになります。$\|\theta^{(j)}\|_1 \le c \cdot R \cdot 2^{-j/2}$, [2].

この2つの例では、ℓ_1 制約が登場したが、他の ℓ_p 制約 $0 \le p < 1$ withcanも自然に登場する；下記参照。読者によっては、ℓ_p の規範を $p < 1$ は当初不思議に思われるかもしれませんが，現在では，このように小さい p ℓ_p ノルムは，スパース性の自然な数学的尺度であることがよく理解されています [3], [4]. 以下に示すように、より多くのスパース性が要求されるようになっている。また、この観点から、ℓ_p に基づく制約は、スパース性を全く必要としない。$p = 2$

なお、これらの例では、注目対象を ℓ_p の制約に従うサブバンドに分離することも許容している。実際には、以下では、関心対象は制約(I.1)に従う係数ベクトルであるという見方に固執する。これは、アプリケーションの観点から、我々の方法が、これらの例のように様々なサブバンドを別々に扱うことに対応することを意味するかもしれない。

ℓ_p 制約の重要な意味は、変換係数のスパース性である。実際、我々は、θ N θ_N が、最大の係数以外のすべてを持つベクトルに注目する場合、そのベクトルが疎であることを自明視している。0 に設定されたクライアントが

$$\|\theta - \theta_N\|_2 \le \zeta_{2,p} \cdot \|\theta\|_p \cdot (N+1)^{1/2-1/p} \qquad \text{(I.2)}$$

は、これらの分野で行われている研究の総称として「OR/IBC」を使用しており、様々な学術的貢献について百科事典的でないことは認めざるを得ません。

我々は、関心のある可能性のあるオブジェクトのクラス X を持っており、$: X \mapsto \boldsymbol{R}^n$, の情報の断片をサンプリングする情報演算子 I_n と , $\boldsymbol{R}^n$ の近似再構成を提供するアルゴリズム A_n $\boldsymbol{R}^m$ を設計することに興味があります。
ここで、*情報演算子は次のような*形式をとります。

$$I_n(x) = (\langle \xi_1, x\rangle, \ldots, \langle \xi_n, x\rangle)$$

ここで, ξ_i はサンプリングカーネルであり，必ずしもピクセルやその他の信号の単純な特徴をサンプリングしているわけではない．しかし，それらは非適応的であり，すなわち，. に依存せずに固定されている．アルゴリズム A_n は不特定多数の、場合によっては非線形再構成演算子である。x

我々は再構成の ℓ_2 誤差と、最適情報と最適アルゴリズムの挙動に興味がある。したがって、最小限の ℓ_2 誤差を比較の基準として考える。

$$E_n(X) = \inf_{A_n, I_n} \sup_{x \in X} \|x - A_n(I_n(x))\|_2.$$

つまり、ここでは、非適応的な線形サンプリングのすべての可能な方法が許容され、再構成のすべての可能な方法が許容されます。我々のアプリケーションでは，興味のあるオブジェクトのクラスは，与えられた（I.1）に対して X $x = \sum_i \theta_i \psi_i$ が従うオブジェクトの セットである．p R と。すると

$$X_{p,m}(R) = \{x : \|\theta(x)\|_p \le R\}.$$

私たちの目標は、それを $E_n(X_{p,m}(R))$ 評価し、それに近い実用的な方式を持つことです。

C.4つのサプライズ

ここで、本稿で注目する主な量的現象について説明する。

定理1: (n, m_n) を、以下のような問題サイズのシーケンスとする。$n < m_n$ $n \to \infty$ そし $m_n \sim An^\gamma$, $\gamma > 1$ $A > 0$,.では $0 < p \le 1$ とある $C_p = C_p(A, \gamma) > 0$ ように

$$E_n(X_{p,m}(R)) \le C_p \cdot R \cdot (n/\log(m_n))^{1/2-1/p}. \qquad \text{(I.3)}$$

このことは、4つの点で驚くべきことである。まず、(I.3)と(I.2)を比較してみよう。較正の下で、両者は似たような形式であることがわかる。

 言い換えると、 次

について $N \in$ 、$0, 1, 2, \ldots$ 、定数 $\zeta_{2,p}$ は $(0,2)$ $p \in$ にのみ依存する。したがって，たとえば誤差を含んだ近似を行う場合，次のようになります. $\epsilon^{(p-2)/2p}$ には、 $N \asymp$ θ の最大の用語だけを残しておく必要があります。

B.最適化回復／情報量に基づく複雑性の背景

このような問題は、*最適復元* [5] や*情報量に基づく複雑性* θ [6] の名で、（他のタイプの仮定について）議論されてきたものである。我々は x x

$n = N\log(m_n)$ のような近似の品質です。 によって提供される I_n *非適応な*情報 $N\log(m)$ を利用する X ことで，最大の変換係数を用いて得ることができる．n 驚くべきは、この近似において、どの変換係数が重要であるか事前に知ることができないことです。しかし、最適な情報演算子 I_n は非適応的で、特定のオブジェクトではなく、せいぜい $X_{p,m}(R)$ クラスに依存する程度です。ある意味で、この非適応な情報は、最適な変換係数を知っているのと N 同じくらい強力です。

このことは， 対象物に対して変換表現が最適で $x \in X_{p,m}(R)$ あり ，他のどの表現も少数の係数によって x 特徴付けることができないことに注目すると，さらに驚くべきことに思える[3], [7]. 確かに，この場合，最適な情報量を支えるサンプリング ξ_i カーネルは

演算子は、単に個々の変換係数を測定しているに過ぎないのでしょうか？実は、そうではありません。情報演算子は、ある意味で、すべての係数を大きなスープに混ぜ合わせたような、非常に複雑な「ホログラフィック」関数を測定しているのです。次の(VI.1)を見てください。$X_{p,m}(R)$ *(ホログラフィーは* ℓ_p $p \leq 1$ 3次元（3-D）イメージが干渉法によって2次元（2-D）変換を生成するプロセスである。2次元変換領域内の各値は、3次元物体全体の各部の影響を受けている。この2次元変換領域のすべて、あるいは一部から、3次元物体を干渉法で再構成することができます。2次元/3次元と干渉測定の詳細はさておき、ここでは、物体が圧縮領域に変換され、圧縮領域の各要素が元の物体のすべての部分の組み合わせであるというアナロジーを感じることができる)。

もう一つの驚きは、情報演算子の種類を増やして、適応的な演算子、例えば、ある測定が以前の測定に対応して行われる演算子を認めると、ほとんどうまくいかないということである。適応的な情報演算子の下での最小誤差を定義するE_n^{Adapt} 適応的な演算子を許すと

$$I_n^A = (\langle \xi_1, x \rangle, \langle \xi_{2,x}, x \rangle, \dots \langle \xi_{n,x}, x \rangle)$$

ここで　、各カーネル$\xi_{i,x}$ は、前のステージ $\leq j$ で $i \geq 2$ 収集された情報に依存することが許されている。$\langle \xi_{j,x}, x \rangle$ 1 $< i$
正式な設定

$$E_n^{\mathrm{Adapt}}(X) = \inf_{A_n, I_n^A} \sup_{x \in X} \left\| x - A_n \left(I_n^A(x) \right) \right\|_2$$

があります。

定理2: $0 < p \leq 1$に対して、> 0 C_p、以下のようになる。
$R > 0$

$$E_n(X_{p,m}(R)) \leq 2^{1/p} \cdot E_n^{\mathrm{Adapt}}(X_{p,m}(R)).$$

適応的な手法は、「大きな係数」を繰り返し「局在化」し、「接近化」することができるはずだというごく自然な期待にもかかわらず、適応的な情報はほとんど役に立ちません。

さらに驚きなのは、すでに興味深い事例として $p = 1$定理1と2は、OR/IBCと近似理論の既知の結果から容易に導出可能である!しかし、導出は間接的である。したがって、著者にとってかなり重要な意味を持つと思われるが、現時点では、優れた非適応型情報演算子やそれに適合する具体的なアルゴリズムについてはほとんど知られていないようである。

本論文の目的は、圧縮性の高い物体の場合 $\leq$ $0 < p$ 1をカバーする直接的な議論を行い、最適に近い情報演算子に関するディレクト情報を与え、この最適に近い情報を利用するための具体的で計算量の多いアルゴリズムについて述べることである。

D. ジオメトリと幅

私たちの観点からは、The-olem 1で説明した現象は、高次元の凸と非凸の "ボール"の幾何学に関係しています。この関係を見るには、このクラスが直交変換の下でなります。これが凸で原点に対して対称であり、軸に対して直交対称である　場合

はウェーブレット基底によって提供される。もし
、これが再び原点について対称で直交対称であり、凸にはならないが、それでも星型である。

この幾何学的な視点をさらに発展させるために、-widthの2つの概念を考える; [5]を参照。

定義 1.1: ℓ_2^m ノルムに関する **Gel'fand -width** は次のように定義されます。

の-次元線形部分空間上で無限大となる。

であり、 $V_n^\perp$ は標準ユークリッド内積に関する V_n の直交補語を表す。

つまり、"トラッピング
"が成立するような部分空間を探すのです。
が小さくなってしまいます。私たちの興味は、Gel'fand
-は、非適応情報の最適回復とそのような-幅の間の等価性から導き出される。

定理3 : ≤および

(I.4)

(I.5)

したがって、Gel'fand -widthは、最適な情報量と全く同じか、ほぼ同じになる。最終的には、(I.3)の不特定定数係数のおかげで、定数 $2^{1/p-1}$ の括弧書きは、私たちにとって等号と同じくらい良いものになるでしょう。私たちは通常、*最適に近い*性能、つまり定数ファクターの範囲内で E_n を得ることにしか興味がないでしょう。

Gel'fand -widthが直接研究されることは比較的まれであり[8]、Kolmogorov -widthに関する結果を見るのが一般的である。

定義1.2: $\boldsymbol{R}^m$ を有界集合とする。 **のコルモゴロフ幅は、** ℓ_2 ノルムに関して、次のように定義される。

ここで、極小値は $\boldsymbol{R}^m$ の -dimensional 線形部分空間にわたっている。

つまり、 d_n は、-次元部分空間による可能な近似の品質を測定するものである。V_n 。

この場合、Kolmogorov 幅と Gel'fand 幅の間には重要な双対関係があり、 d_n に関する公表された結果から、その特性を推測することができる。それを述べるために、 ℓ_2 ノルムではなく、近似に基づいて、明白な方法で定義される ようにしますまた、与えられた とし、 標準的な双対指標である ,.また、 $b_{p,m}$ を .の標準単位球とすると、[8]。

(I.6)

特に

$$d_n(b_{2,m};\ell_\infty) = d^n(b_{1,m},\ell_2).$$

$n < 1$

n

n X

$$d^n(X;\ell_2) = \inf_{V_n}\sup\{\|x\|_2 : x \in V_n^\perp \cap X\}$$

n

$\boldsymbol{R}^m$

$x \in X$

x

n

n

$p = 1$

$0 < p$ 1 $R > 0$

$$d^n(X_{p,m}(R)) \le E_n(X_{p,m}(R))$$
$$\le 2^{1/p-1} \cdot d^n(X_{p,m}(R)).$$

n

n

n

$X \subset$

n X

$$d_n(X;\ell_2) = \inf_{V_n}\sup_{x\in X}\inf_{y\in V_n}\|x-y\|_2$$

n

X

n

$p = 1$

d^n

$d^n(X,\ell_q)$

ℓ_q

$p \ge 1$

$q \ge 1$ p' q' $1/p' =$

$1 - 1/p$ $1/q' = 1 - 1/q$

ℓ_p^m

$$d_n(b_{p,m};\ell_q) = d^n(b_{q',m},\ell_{p'}).$$

左辺の漸近的な性質はGarnaev and Gluskin [9]によって定義されている。R ∞ これは、Kashin [10]がソボレフ空間の Kolmogorov n-width を決定する過程で、この結果の少し弱いバージョンを開発した主要な仕事に続くものである。詳しくは原著論文、あるいはPinkusの著書[8]を参照されたい。

定理4：（**カシン、ガルナエフ、グルスキン**(KGG)すべての、$m > n$で

$$d_n(b_{2,m}, \ell_\infty) \asymp (n/(1+\log(m/n)))^{-1/2}.$$

定理1は、双対性の式(I.6)と同値の式(I.4)を用いて $p = 1$ KGGを適用 $0 < p < 1$ 場合分けして導かれるようになった。定理1の $0 < p \leq 1$ 場合は双対性が使えないので、全範囲、この論文では別のアプローチをしています。

E. ミステリーズ...

KGGの結果は定理1に間接的に影響を与えるので、この方法では興味のある現象について多くを学ぶことはできない。Kashin, Garnaev, Gluskin の議論は、Kolmogorov幅のための最適に近い n-dimensional 部分空間が存在することを示す。それらは、ある性質を欠く行列の数、± 1 項目を持つ行列の総数、および、comparing によって存在が知られている、± 1 項目を持つある行列の null 空間として発生する。このアプローチの解釈可能性には限界がある。

情報演算子の暗黙性は、再構成アルゴリズムの抽象性と一致する。OR/IBC理論に基づくと、いわゆる*中心的アルゴリズム*が最適であることがわかる。この「アルゴリズム」は、与えられた情報 $y_n = I_n(x_0)$ に対して、データ y_n x

$$I_n^{-1}(y_n) = \{x : I_n(x) = y_n\}.$$

セットの中心を今定義する S

$$(S) = \inf_c \sup_{x \in S} \|x - c\|_2,$$

中心的なアルゴリズムは

$$\hat{x}_n^* = \quad \left(I_n^{-1}(y_n) \cap X_{p,m}(R)\right)$$

であり、情報I_n が最適であれば、それに従う。

$$\sup_{x_0 \in X_{p,m}(R)} \|x_0 - \hat{x}_n^*\|_2 = E_n(X_{p,m}(R));$$

は、後述の第III章をご参照ください。

この抽象的な視点は、残念ながら実用的なアプローチには結びつきません（少なくとも、この場合は $X_{p,m}(R)$。$0 < p < 1$). セット$I_n^{-1}(y_n)$は $\cap$ $X_{p,m}(R)$ の断面が $X_{p,m}(R)$

このセクションの中心を求めることは、標準的な扱いやすい計算問題にはならない。さらに、これは と がわかっていることを前提としているが、通常はそうではないだろう。

F. 結果

本論文では、主に2種類の結果を展開する。

- *近最適情報。*Gel'fと- widthsの最適に近い部分空間の問題を直接考察し、そのヌルスペースが最適に近いことを意味する-by-行列の3つの構造条件（CS1-CS3）を 導入、 $\boldsymbol{R}^m$ の-subspacesの大部分が最適に近いことを示し、ランダムサンプリングにより、圧倒的に高い確率で最適に近い情報操作子が得られることを示す。
- *最適に近いアルゴリズム我々は*、単純な非線形再構築アルゴリズムについて研究している：単純に、測定値を満足することを条件として、係数の ℓ_1 ノルムを最小化する。これは信号処理の文献でBasis Pursuitという名前で研究されており、線形計画法によって計算することができる。我々は、この方法がすべての $\leq$ に対して、ほぼ最適な結果を与えることを示す。

つまり、最適に近い情報演算子の大量供給と、線形計画法に基づく最適に近い再構成法を提供し、おそらく予想外に、非凸の 場合にも有効である 。

どのような結果が得られるかを味わうために、特定の情報とアルゴリズムの組み合わせを考えてみましょう。

- *CS情報。*列をランダムにサンプリングして生成されたanmatrixとし、異なる列は独立かつ同一に分布する（i.i.d.） $\boldsymbol{S}^{n-1}$ 上のランダムユニフォームとする。が大きい場合、圧倒的な確率で、以下のセクションII-Aで詳細に説明するCS1〜CS3の性質を持つ。このような有利な抽選を行ったとする。基底ベクトルを第１列とする 基底行列とする。CS 情報演算子はその行列である。
- ℓ_1 -- *minimization*。CS情報から再構成するために、凸最適化問題

$$(L_1) \quad \min \|\Psi^T x\|_1 \quad 会わせる \quad y_n = I_n^{\mathrm{CS}}(x).$$

つまり ℓ_1 、 $\hat{x}_1$ の係数が最小で、情報量と一致するものを探す。 y_n

情報オペレータの品質を評価するために、 I_n 、以下を設定します。

$$E_n(I_n, X) \equiv \inf_{A_n} \sup_{x \in X} \|x - A_n(I_n(x))\|_2.$$

アルゴリズムと情報の組み合わせの (A_n, I_n) 品質を評価するために 、設定します。

$$E_n(A_n, I_n, X) \equiv \sup_{x \in X} \|x - A_n(I_n(x))\|_2.$$

定理5：, を、 ,, n m_n 従う問題サイズのシーケンスとし $n < m_n \sim An^{\gamma}$ $A > 0$ $\gamma \geq 1$ 、 η_i ρ $C = C(p, (\eta_i), \rho, A, \gamma)$ I_n^{CS} を、基礎パラメータとCS行列から派生した演算子の対応するシーケンスとする（以下のセクションIIを参照）。ここで $0 < p \leq 1$.が > 0 I_n^{CS} n 存在するので $X_{p,m}(R)$ が最適に近い。 n m

$$nE_n\left(I_n^{\mathrm{CS}}, X_{p,m}(R)\right) \leq C \cdot E_n(X_{p,m}(R))$$

θ

$0 < p \quad 1$

$0 < p < 1$

Φ $n \times m$

n Φ

Ψ $m \times m$ ψ_i i I_n^{CS} $n \times m$ $\Phi\Psi^T$

$R > 0$ $n > n_0$に対して、。さらに、アルゴリズム $A_{1,n}$ の 解は (L_1) に最適である。

$$E_n\left(A_{1,n}, I_n^{\mathrm{CS}}, X_{p,m}(R)\right) \leq C \cdot E_n(X_{p,m}(R))$$

Ψ $R > 0$ $n > n_0$ために、。

このように、大きな n m Φ , に対して $\approx$ 最適に近い情報の簡単な記述と、扱いやすい最適に近い再構成アルゴリズム A がある。

Ψ

G. 潜在的なアプリケーション

その意味を理解するために、まずスペクトルのBump Al-gebraモデルを思い出してみよう。この文脈では、n $ker(\Phi)$ I_n 我々の結果は、The-olem $n = O(m^{1/3}\log(m))$ 5の情報演算子に基づく分光器では、実際には、次のように取るだけでよいということになる。の測定値を用いて、公称の測定値ではなく、そのようなスペクトルの正確な再構成を得ることができます。ただし、その場合は非線形に処理しなければならない。

画像の境界変動モデルを思い出してください。このモデルでは、定理5と似たような結果として、最適に近い情報オペレータに基づく特殊な画像処理装置では、公称の測定値ではなく、ピクセルで画像を正確に再構成するための測定 $n = O(\sqrt{m}\log^3(m))$ m m 値のみが本当に必要であることが述べられています。

これらの結果の基礎となる計算を以下に示す。また、漫画のような画像（脳スキャンのような、ある種の単純な自然画像をモデルにしていると考えられる）では、-ピクセル画像の測定回数はわずか m $O(m^{1/4}\log^{5/2}(m))$ であることを示す結果もある。

H. 内容

セクションIIでは、情報演算子の近最適化のための条件CS1〜CS3を紹介する。第III節では、抽象的な最適に近いアルゴリズムを考察し、定理1〜3を証明する。第IV節では、凸最適化問題を解くと、$\leq$のとき、(L_1) $0 < p$ 1 近最適アル ℓ_p ゴリズムが得られることを示す。第セクションVでは、弱い条件とタイトフレームへの即時の拡張を指摘する。セクションVIでは、画像、信号、配列処理における潜在的な意味を概説している。セクションVIIでは、[11]の研究を基に、条件CS1〜CS3が「ほとんどの」情報演算子で満たされることを示す。

最後に、セクションVIIIでは、2つのグループによる進行中の研究（Gilbertら [12]とCandèsら [13], [14]）に注目する。これらは n-width/OR/IBCの伝統に従って書かれていないが、（我々が説明するように）密接に関連した結果を暗示している。

通常ヌルスペース。ある演算子に対する集合の*幅を* X 定義する。

II. INFORMATION

以下のように構成された情報演算子を考える。直交行列の列を基底要素ψ_iとし、以下に示す条件を満たす特定のバイ行列を用いて、対応する情報演算子I_nを構成する。直交行列の選択によってすべてが決まるので、以下のように仮定する。というのが、このセクションを通してのアイデンティティです。

Gel'fと-widthとmin-imax誤差の関係を考慮すると、-widthで作業しても良いだろう。としておく。

条件

として

$$w(\Phi, X) \equiv \sup \|x\|_2 \quad x \in X \cap ker(\Phi). \qquad \text{(II.1)}$$

つまり、

ヌルスペースで切り取られた部分の半径である　。　一般に、Gel'fand　-widthは、以下の選択によって得られる最小の値である。

$$d^n(X) = \inf\{w(\Phi, X) : \Phi \text{ は } n \times m \text{ anmatrix}\}.$$

が存在することを　示す　。

に　依存　し、比率は $\log(m)/\log(n)$。

A.　条件CS1～CS3

以下、　$J \subset \{1,\dots,m\}$　で、　Φ_J　は、Φ の示された列だけを選択して得られる副行列を表す。. V_J は $\boldsymbol{R}^n$ の範囲とする．最後に，Φ_J;で示される $\boldsymbol{R}^n$ ベクトル上の $\ell_1(J^c)$ ℓ_1 ノルムを用いて，$\{1,\dots,m\}\backslash J$;上の*商ノルムの族を考察する.*

$$Q_{J^c}(v) = \min \|\theta\|_{\ell_1(J^c)} \quad \text{会わせる} \quad \Phi_{J^c}\theta = v.$$

これらは、.NET v Φ の列の指定された部分集合のみを用いた達成可能な最小のℓ_1 -norm 表現を記述する。

厳密には正の $n \times m$ Φ trix に課すべき3つの条件を定義する　パラメータで η_i $1 \le i \le 3$ ρ 指標化されたanma, ≪ , 。

CS1:Φ_J の最小特異値は一様に超える。$\eta_1 > 0$ $|J| < \rho n/\log(m)$ V_J

CS2：

それぞれの部分空間において、不等式

$$\|v\|_1 \ge \eta_2 \sqrt{n} \cdot \|v\|_2, \quad \forall v \in V_J$$

$|J| < \rho n/\log(m)$ V_J で統一しています。

CS3:各副空間について

$$Q_{J^c}(v) \ge \eta_3/\sqrt{\log(m/n)} \cdot \|v\|_1, \quad v \in V_J$$

で $|J| <$ 一様に$\rho n/\log(m)$。

CS1では、すべての列の小グループの間に、ある量的な線形独立性が要求されます。CS2は、少なくともℓ_1とℓ_2のノルムの比較に関する限り、小さな列のグループの線形結合は、ランダムノイズによく似たベクトルを与えると述べています。これは、幾何学的な事実によって暗示されます。すべての V_J $\{x : \Phi x = 0\}$ は、結果として生じる凸部が実際に球形に近くなるように、ℓ_1^m の球を切り裂きます。CS3　は，ある V_J のすべてのベクトルに対して，関連する商ノルムは，.NET Q_{J^c} $\boldsymbol{R}^n$ の単純な ℓ_1 ノルムより劇的に小さくなることはない，と言っています．

X $ker(\Phi)$ n Φ を適切に選ぶと、これらの条件を満たす行列は、大きな行列でも普遍的に存在することがわかる。もちろん、任意の有限および n m 、すべての規範は等価であり、ほとんどすべての任意の行列は、単に非常に小さい、および、非常に大きいを取ることによって、これらの条件を三重に満たすことができます。しかし、「非常に小さい」、「非常に大きい」の定義が Φ

$$w(\Phi, b_{p,m}) \le C \cdot d^n(b_{p,m})$$

C p

この些細な議論を成立させるために、私たちはもっと深いことを主張する。我々は、より深い真理を主張する：それは、 m An^{γ} $\leq$ のandから独立したandを選択することが可能である。

集合を考える

θ $0 = \Phi\theta$ J $k =$ $S^{n-1} \times \cdots \times S^{n-1}$

単位正規化された列を持つ $\theta =$ すべての行列の集合。 この集合に対して、自然一様尺度（各因子に一様な積の尺度 S^{n-1} ）で出現頻度を測定する。

定理6： $\leq$ ≤ 1 (n, m_n) を、以下のような問題サイズのシーケンスとする。 $n \to \infty$ $n < m_n$ $m \sim An^{\gamma}$ $A > 0$ γ 1 $\eta_i > 0$ $\rho > 0$ θ_J 、A γ $\delta > 0$ 、、≥。が $n \times m$ Φ θ_{J^c} 存在し，and ρ $1 - \delta$ にのみ依存して，パラメータ (η_i) で CS1-CS3 を満たす 全行列の割合が ，最終的に を超える.

第VII節でこのことを議論し、証明する。この証明により、条件を満たさない行列の割合が、.NETにおいて指数関数的に減少することが示される。

後の使用のために、定数と暗黙のうちに置いておき、単にCS行列、つまり、この定理で説明されるタイプのパラメータの値で与えられた条件を満たす行列、すなわち、上記の偏在性に依存せず、かつそれを許容する行列のことを言うことにする。

B. CS行列の近最適化

次に、CS条件がCS行列によって誘導される幅の最適に近い状態を意味することを示す。

定理7： (n, m_n) $n \to \infty$ $m_n \sim A \cdot n^{\gamma}$ andを持つ問題サイズの列をbeとする 。 の列を考える。

行列 Φ_{n,m_n} は CS1-CS3 の条件に従い、 η_i ρ n $p \in (0, 1]$ かつ正で.から独立した行列です。このとき、それぞれについて $C = C(p, \eta_1, \eta_2, \eta_3, \rho, A, \gamma)$ ように $n > n_0$

$$w(\Phi_{n,m_n}, b_{p,m_n}) \leq C \cdot (n/\log(m/n))^{1/2-1/p}.$$

証明する。最適化問題を考える。

$$(Q_p) \quad \sup \|\theta\|_2 \text{ subject to } \Phi\theta = 0, \|\theta\|_p \leq 1.$$

我々の目標は、以下の値を束縛する (Q_p) ことである。

$$val(Q_p) \leq C \cdot (n/\log(m/n))^{1/2-1/p}.$$

の場合、 ℓ_1 $p < 1$ 近似のための同様の議論は、以下のようになる。 (II.4)

$$\|\theta_{J^c}\|_1 \leq \zeta_{1,p} \cdot (k+1)^{1-1/p}.$$

とように選ぶ 。 における $\lfloor \rho n / \log(m) \rfloor$ の最大値の添字を表すと一般性を損なわずに，座標がエントリの中で最初に 来るように並べられ， $[\theta_J, \theta_{J^c}]$ を分割 するとする．明らかに

$$\|\theta_{J^c}\|_p \|\theta\|_p$$

(二・二)

の各エントリが少なくとも同じ大きさであるためです。

となり、(I.2)からは

今すぐしたがって、

withでは、次のようになります。

.

として

、

CS3を起動し、以下を得ることができます。

一方、V_J と$\rho n/\log(m)$を使って

、CS 2 を起動し、次のようになります。

これらと上記を組み合わせることで

$\zeta_{1,p}\rho^{1-1/p}/\eta_2\eta_3$ をとする。

$\rho n/\log(m)$を思い出し　　、CS1 を呼び出すと、次のようになります。

要するにc_1/η_1

で定理が導かれる。

III. ALGORITHMS

情報演算子I_n が与えられたとき、我々は、Gel'f と - width の推定値と品質的に互換性のある再構成を提供する再構

$$\|\theta_{J^c}\|_2 \leq \zeta_{2,p} \cdot (k+1)^{1/2-1/p}. \qquad \text{(II.3)}$$

A_n

成アルゴリズムを設計する必要がある。序文で述べたように、OR/IBCフレームワークにおける最適な方法は、いわゆる中心アルゴリズムであるが、残念ながら、我々の設定においては、一般的に効率的に計算可能ではない。ここで、定理1を証明することができる別の抽象的なアプローチについて説明する。

$$v = \Phi_J\theta_J + \Phi_{J^c}\theta_{J^c} \qquad v = \Phi_J\theta_J$$

$$= v = \Phi_{J^c}\theta_{J^c} \qquad v \in V_J \qquad |J| = k < \rho n/\log(m)$$

$$\|\theta_{J^c}\|_1 \geq Q_{J^c}(-v) \geq \eta_3/\sqrt{\log(m/n)} \cdot \|v\|_1.$$

$v \in$ $|J| = k <$

A. *実現可能点法*

OR/IBC

$$\|v\|_1 \geq \eta_2 \cdot \sqrt{n} \cdot \|v\|_2.$$

リテラルのもう一つの一般的な抽象的アルゴリズムに、

$$\|v\|_2 \leq (\eta_2\eta_3)^{-1} \cdot (\sqrt{\log(m/n)}/\sqrt{n}) \cdot \|\theta_{J^c}\|_1$$

$$\leq c_1 \cdot$$

$c_1 =$ $(n/\log(m))^{1/2-1/p}$ $|J| = k <$

いわゆる*実行可能点法がある。これは*、観測された情報と制約に適合する*任意の*再構成を見つけることを単純に目的とする。

$$\|\theta_J\|_2 \leq \|\Phi_J\theta_J\|_2/\eta_1 = \|v\|_2/\eta_1.$$

$c_2 =$ ·

中央アルゴリズムの場合と同様に、与えられた

$$+\zeta_{2,p}$$

$$\|\theta\|_2 \leq \|\theta_J\|_2 + \|\theta_{J^c}\|_2$$

$$\|y_n\|_2 = I_n(x)$$

$x \in$

$$\leq c_2 \quad (n/\log(m))^{1/2-1/p}$$

$$(\rho n/\log(m))^{1/2-1/p}$$

$$C = \left(\zeta_{1,p}\rho^{1-1/p}/(\eta_1\eta_2\eta_3) + \zeta_{2,p}\rho^{1/2-1/p}\right). \qquad \square$$

情報に対して 、

すべてのオブジェクトのコレクションを考える

$X_{p,m}(R)$ 情報の原因となり得るもの　n　y_n

$$\hat{X}_{p,R}(y_n) = \{x : y_n = I_n(x), \qquad x \in X_{p,m}(R)\}.$$

実現可能点法では、単純に $\hat{X}_{p,R}(y_n)$を、どのような手段であれ15], [6], [8]における標準的なOR/IBCの議論を適応して、次のように示すことができる。

レンマ3.1。 $I_n(x_0)$ ここで、 $y_n \equiv$ $x_0 \in X_{p,m}(R)$ I_n は最適情報演算子であり、 $\hat{x}$ beの任意の要素である。 $\hat{X}_{p,R}(y_n)$.では≤ $0 < p$ 1

$$\|x_0 - \hat{x}\|_2 \leq 2 \cdot E_n(X_{p,m}(R)). \tag{III.1}$$

つまり、どのような実現可能な点でも、最適値からファクター2の範囲内にあるのだ。

*証明する。*まず、中心アルゴリズムの最適性についての主張を正当化し、次に実行可能な点が中心アルゴリズムに近いことを示す。ここで再び $\hat{x}_n^*$ を中心アルゴリズムの結果とする。ここで

$$\text{ラジウス}(\hat{X}_{p,R}(y_n)) \equiv \inf_c \sup\left\{\|x - c\|_2 : x \in \hat{X}_{p,R}(y_n)\right\}$$
$$= \sup\left\{\|x - \hat{x}_n^*\|_2 : x \in \hat{X}_{p,R}(y_n)\right\}.$$

ここで、明らかに、 isが x_0 $X_{p,m}(R)$ $y_n = I_n(x_0)$ andにあることだけが分かっている特別な場合、測定されるとき、最小誤差は $(\hat{X}_{p,R}(y_n))$ 正半径 この誤差を実現しているため y_nような各中央アルゴリ x ズムによって、すべての誤差を最小にすることができます。この最小誤差は

$$\sup\{\text{ラジウス}(\hat{X}_{p,R}(y_n)) : y_n \in I_n(X_{p,m}(R))\} = E_n(X_{p,m}(R)).$$

今、実現可能な点は $\hat{x} \in X_{p,m}(R)$従います。

$$\|\hat{x} - \hat{x}_n^*(y_n)\|_2 \leq \text{ラジウス}(\hat{X}_{p,R}(y_n)).$$

しかし、三角形の不等式から

$$\|x_0 - \hat{x}\|_2 \leq \|x_0 - \hat{x}_n^*(y_n)\|_2 + \|\hat{x}_n^*(y_n) - \hat{x}\|_2$$

それ故に$x_0 \in \hat{X}_{p,R}(y_n)$

$$\begin{aligned}\|x_0 - \hat{x}\|_2 &\leq 2 \cdot \text{ラジウス}(\hat{X}_{p,R}(y_n)) \\ &\leq 2 \cdot \sup\{\text{ラジウス}(\hat{X}_{p,R}(y_n)) \\ &\qquad : y_n \in I_n(X_{p,m}(R))\} \\ &= 2 \cdot E_n(X_{p,m}(R)).\end{aligned}$$ □

より一般的には、情報演算子I_n が最適に近いものでしかない場合、同じ論法で次のようになる。

$$\|x_0 - \hat{x}\|_2 \leq 2 \cdot E_n(I_n, X_{p,m}(R)). \tag{III.2}$$

所与の $y_n = I_n(x_0)$ $R > 0$ ものであり $x_0 \in X_{p,m}(R)$ $0 < p \leq 1$

$$\|x_0 - \hat{x}_{p,n}\|_2 \leq 2 \cdot E_n(X_{p,m}(R)).$$

B. 定理3の証明

先に進む前に、定理3を証明するのが便利である。なお、 $\geq p$ 1はOR/IBCでよく知られているので、その $p < 1$ 場合の議論をするだけでよい $p = 1$ （ただし、我々の議論がforasでも有効であることは偶然である）。を適用することがポイントになる。

p-さんかくふとうしき

$$\|\theta + \theta'\|_p^p \leq \|\theta\|_p^p + \|\theta'\|_p^p$$

$0 < p < 1$ に対して有効である 。この不等式は Peetre と Sparr の仕事を通じて補間理論 [17] でよく知られており、直接検証することは容易である。

一般性を損なうことなく、最適な部分空間 V_n があるとする。これは固定されており、この証明で与えられる。先ほど見たように

$$E_n(X_{p,m}(R)) = \sup\{\text{ラジウス}(\hat{X}_{p,R}(y_n)) : y_n \in I_n(X_{p,m}(R))\}.$$

現在

$$d^n(X_{p,m}(R)) = \text{ラジウス}(\hat{X}_{p,R}(0))$$

であるから、明らかに$E_n \geq$ d^n。ここで，一般性を損なうことなく， x_+ andat半径境界を満たす，すなわち，以下の条件を満たすと仮定する． x_- $I_n(x_\pm) =$ y_n を満たし $c =$ かつ、(意図的に) $(\hat{X}_{p,R}(y_n))$

$$E_n(X_{p,m}(R)) = \|x_+ - c\|_2 = \|x_- - c\|_2.$$

次に、$x_-)/2^{1/p}$ を定義します。 $\theta_\pm = \Psi^T x_\pm$ セットと ξ $\Psi^T \delta$.三角形の不等式により

$$\|\theta_+ - \theta_-\|_p^p \leq \|\theta_+\|_p^p + \|\theta_-\|_p^p$$

であるからして

$$\|\xi\|_p = \|(\theta_+ - \theta_-)/2^{1/p}\|_p \leq R.$$

δ $\delta \in X_{p,m}(R)$それゆえしかし $X_{p,m}(R) \cap$ $\leq d^n(X_{p,m}(R))$

$$I_n(\delta) = I_n((x_+ - x_-)/2^{1/p}) = 0$$

の要素を取ることが、実現可能な点としてよく選ばれています。

最小ノルム、すなわち問題の解決策

$V_n^{\perp}$は，に

属するしたがって，$\|\delta\|_2$と

$$(P_p) \quad \min_x \|\theta(x)\|_p \text{会わせる} \; y_n = I_n(x)$$

ここで$\theta \in = \theta(x)\,\Psi^T x$は、変換係数のベクトル、$\ell_p^m$。この手法の優れた点は，次のようなことが必要ないことです。

は、　R　球の　$X_{p,m}(R)$　半径を知るために、最小の要素である規範は常にその内側にある。後で使うために、この解を$\hat{x}_{1,n}$前のレンマにより、この手続きはニアミニマックスである。

$$E_n(X_{p,m}(R)) = \|x_+ - x_-\|_2/2 = 2^{1/p-1}\|\delta\|_2$$
$$\leq 2^{1/p-1} \cdot d^n(X_{p,m}(R)).$$

□

C. 定理1の証明

これで、「はじめに」の定理1を証明できる状態になった。

まず、この　$p = 1$場合　、　GarnaevとGluskinの定理が含意することは、すでに「はじめに」で説明したとおりである。

を双対性で示す。$\inf_{V_n}$ ($V_n \cap X_{p,m}(R)$) $<= n d^n(X_{p,m}(R))$. $0 < p < 1$ 場合、同順位の下界と上界を示せばよい。 $E_n(X_{p,m}(R))$ d^n

下界のために、次のように定義されるエントロピー数を考える。を集合とし、そのための $e_n(X, \ell_2)$ X ϵ X netが、最大でも.cardinalityのnetを使って構築できるような最小の数を. 2^n

$c > 0$ n etとカールの定理 [18]-Pisier's book [19]の例を参照- からは、Gel'f と -width がエントロピー数を支配する定数が p あることがわかる。 $0 < p$ 1 $p < 1$

$$d^n(b_{p,m}) \geq c e_n(b_{p,m}).$$

第二に、エントロピー数は [20], [21] に従う。

$$e_n(b_{p,m}) \asymp (n/\log(m/n))^{1/2-1/p}.$$

同時に、定理7と6の組み合わせで、以下のことがわかります。

$$d^n(b_{p,m}) \leq c(n/\log(m))^{1/2-1/p}.$$

ここで、Feasible
Point法を適用すると、次のようになります。

$$E_n(X_{m,p}(1)) \leq 2d^n(b_{p,m})$$

$E_n(X_{m,p}(R))$ への即時拡張で $R > 0$
、すべての $R \neq$。
と結論づけています。

$$E_n(b_{p,m}) \asymp (n/\log(m/n))^{1/2-1/p}$$

が証明されることになった。

D. 定理2の証明

今が定理2を証明する好機である。 $\leq$ の場合、これは既知であることに注意する 1 p [22]-[25] $0 < p < 1$ の場合も論証は同じであり、単にそれを繰り返すだけである。とする。
$x = 0$ そして、どのようなアルゴリズムであれ、適応的に構築された部分空間を考える。アルゴリズムが終了すると、- x 0 dimensionalの情報ベクトルと、その情報ベクトルをすべて与えるであろうオブジェクトからなる部分空間 V_n^0 が得られる。 V_n^0 $(V_n^0 \cap X_{p,m}(R))$ のすべてのオブジェクトに対して、適応的な情報ベクトルは同じになる。ここで、その情報に関連する最小誤差はちょうどradius

である。

ラジウス

と比較ことで結果が導かれる 。

IV. ベース・パーツイスト

前節の最小ノルム法には2つの欠点がある。第一に、 $\leq$ に対して

有効なアルゴリズムが必要である。第二に 、
もし数問題が非凸最適化過程を呼び起こすのであれば、それは難解であると考えられる。本節では、この2つの欠点を修正する。

A. ケース

この 場合
、凸の最適化問題である。最適化変数と等価な形で書くと、次のようになる。

$$\min_\theta \|\theta\|_1$$会わせる

これは線形計画問題として定式化することができる：
thebymatrixとする
。

線形計画問題

LP $\min_z 1^T z$ subject to(IV.1

は，例えば，R^{2m} のベクトルを解とし ，
ように分割することができる．この再構成は，.この線形プログラムは、一般に計算しやすいと考えられている。実際、この問題は信号解析の文献で Basis Pursuit [26] という名前で研究されており、その中で、非常に大規模な劣決定問題-例えば withand-が
内点最適化法を用いてうまく解決された。

性能に関しては 、
この方式がほぼ最適であることが既に分かっている
。（III.2）より、以下のようになる。

4.1:I_n が情報演算子で、ある種の

の場合、Basis Pursuit アルゴリズム $A_{1,n}(y_n)$ $\hat{x}_{1,n}$ は、すべての場合において、達成される。

特に、我々は、*任意の*
クラス、すなわち任意の、任意の、任意の、を扱うための普遍的なアルゴリズムを持っている。まず、最適に近い情報演算子を適用し、次に、Basis Pursuitによって再構築する。その結果は

に

依存する定数である
。この不等式は次のように使うことができる。
修正。未知の
物体が高度に圧縮可能であることが知られているとします。例えば、*先験的な*境界 $\|\theta\|_1 c_2 m^\alpha$， 1/2 に従うとします。$K_\epsilon (c_1 c_2/\epsilon)^2$
とする。このような物体に対しては、測定するのではなく、$K_\epsilon\, m^{2\alpha}\log(m)$
を測定するだけでよく、再構成は以下のようになります。

$$p = \|x_0 - \hat{x}_{1,n}\|_2 \le \epsilon \cdot \|x_0\|_2.$$

$p = 1$ (P_1)
$p = 1$

θ

$0 < p < 1$ このケースはすでに重要で興味深いですが、より圧縮性の高い
(P_1) データに対応し、定理1においてより印象的な性能を提供するため、指数が $1/2 - 1/p$ $\Phi\theta = y_n.$
の方がさらに強い $p = 1$ 。

n $2m$ $[\Phi \ -\Phi]$ $\hat{x}_{0,n}$ このセクションの後半 A
で、同じ解釈を、(1) (2) (3) (4) (5) (6)の性能に拡張します。
$X_{p,m}(R)$ $p < 1$を通して

( ) $Az = y_n, \quad z \ge 0$

z^*

$z^* = [u^* v^*]$ $\theta^* = u^* - v^*$ (P_1)
$\hat{x}_{1,n}$ $\Psi\theta^*$

$n = 8192$ $m = 262144$

$p = 1$

$C > 0$

$$E_n(I_n, X_{1,m}(1)) \le C \cdot E_n(X_{1,m}(1));$$

$=$

$R > 0$

$$E_n(I_n, A_{1,n}, X_{1,m}(R)) \le 2C \cdot E_n(X_{1,m}(R)).$$

$X_{1,m}(R)$ Ψ m R

$$\|x_0 - \hat{x}_{1,n}\|_2 \le c_1 \ \|\theta\|_1 \ (n/\log m)^{-1/2}$$

c_1 $\log(m)/\log(n)$

$\epsilon > 0$

x_0

$\le$ $\alpha <$

$=$ m

$n \sim$

B. ℓ_1 とℓ_0 の関係 最小化

$0 < p < 1$ 一般的なOR/IBC理論では、このようなケースを扱うには、非凸の最適化 (P_p) 問題を解く必要があり、難解に思われます。しかし、少なくとも現状では、小さな奇跡が起こる：(P_1) 解くことが再び最適に近くなるのである。このことを理解するために、まず少し回り道をして、ℓ_1 と極限状態との関係を調べます。$m \leq An$ $\rho = \rho(A, \eta)$

$\Phi\beta \neq 0$ $p \to 0$ 場合ℓ_p 空間。ここで Φ

$$(P_0) \quad \min_\theta \|\theta\|_0 \text{会わせる } \Phi\theta = y$$

ここで、$\theta + ker(\Phi)$ θ θ $\|\theta\|_0$ はちょうど.のノンゼロの数です。ここでも、Peetre and Sparr [16]の仕事以来、ℓ_0の重要性とℓ_p for $0 < p < 1$ との関係が $\beta \in R^m$ よく理解されています。より詳細については[17]を参照してください。Φ L $1_J\beta$ $\beta 1_J(i)$

通常、ℓ_0 ノルムを含むこのような問題を解くには、組合せ最適化が必要です。$\{1, \ldots, m\}$ のすべての疎な部分集合を列挙して、$\Phi\theta = y$ (P_0) (P_1)

$$\nu(\Phi, J) = \sup \left\{ \frac{\|1_J\beta\|_1}{\|\beta\|_1} : \Phi\beta = 0 \right\}.$$

解がものを探します。しかし、疎な解を持つ場合が見つかります。

定理8： CS1-CS3 を、Φ与えられた正の定数 ρ, (η_i) で満たすとする。ρ に対する解があるとき$\rho_0 > 0$その解を求めるための定数depending only on (η_i) and ρ and not on n or mが存在する。$\Phi\theta$ $\rho n/\log(m)$

$y =$ 最大で0である(P_0)とき (P_1) 両者は同じ一意解を持つ。

つまり、方程式系が大量に不足決定されているにもかかわらず、ℓ_1 最小化とスパース解がコインで結ばれ、結果が十分にスパースである場合です。

現在では、ℓ_1 と ℓ_0 の最小化の等価性に関する結果を示す広範な文献があります [27]-[34]。このテーマに関する初期の文献では、$O(n^{1/2})$ ノンゼロを許容するスパース性制約を含む条件下で等価性が見出されました。この種の結果が可能であることは意外に思われるかもしれませんが、スパース性制約 $\|\theta\|_0$ $O(n^{1/2})$ n m m は、結局のところ、期待はずれで小さいものなのです。大きなブレークスルーはCandès, Romberg, and Tao [13]の貢献で、by Fourier行列からnランダムに行を取ることによってm構築されるマトリスを研究し、オーダー$\rho = \rho(A)$ $O(n/\log(n))$ 境界を与え、以前に知られていた n m $O(n^{1/2})$ 結果よりも劇的に弱いスパース性条件件が必要であることを示しました。11]では、'ほぼ全ての'by行列について、$\leq$ nonzeros, と 等価であることが示されました。上記の結果は、「ほぼすべての」by matrices with のために、事実上、以下のことを述べています。, 同値は$O(\rho n/\log(n))$ ノンゼロまで保持される, ここで.

我々の議論は[11]と並行して行われ、ヌルスペースには条件を満たす非常に特殊な構造があることが示されました。の問題です。がスパースであるとき、与えられたアフィン 部分空間において、小さいℓ_1 ノルムを持ちうる*唯一の* 要素は自分自身で 。定理8を証明するために、まず、.のヌルスペースにおける要素の非スパース性に関するレンマが必要です。$\{1, \ldots, m\}$、与えられたベクトルに対して、letdenoteは、ミューティル化されたのベクトルを エントリと 。**濃度を**定義する

のヌルスペース中のあるベクトルに対して、ある部分集合に集中できるノルムの割合を示す。
この集中度は、小さいと大きくならない。

*レンマ4.1。が*CS1-CS3を満たすとすると、定数および.に依存する定数が存在する。

では

*証明する。*これは定理7の議論を変形したものである。 β

とする。が最も集中した部分集合であると仮定し 、その列は分割されるように 番号付けされているとする .再び CS2-CS3 を起動し、以下を得る。

$\|v\|_2$ $(\sqrt{n}/\sqrt{\log(m/n)}) \cdot \|\theta_{J^c}\|_1$

$(n/\log(m))^{1/2-1/p}$.

CS1を起動し、以下を得る。

今はもちろん、 ≤。これらをすべて組み合わせると

設定、レンマが続く 。

*定理8の証明：*andが 部分集合にsup- ported であるとする 。

の唯一の最小化器であることをまず示す。 が 、、 の解であるとする。

その後、 β .私たちは

トワイスの定義

今、 1/2 は、 i e.、を与えています 。

ここで、 レンマ4.

1の定数を思い出してDefin eso 1/2 と 。

レンマ4.1からは implies $\eta_0\rho_0^{1/2}$ $1/2$.

C. に対する基底追求の近最適性

ここで、Basis Pursuitが 全領域でほぼ最適であると主張することに戻ります 。

*定理9：*CS1-CS3 を con- stants と共に満たすものがあるとする。の問題インスタンスに対する解が 存在

$$\|\theta_0 - \hat{\theta}_{1,n}\|_2 \le C_p \cdot \|\theta_0\|_p \cdot$$

この証明には安定性のレンマが必要であり、ノルムで測定される小さな摂動に対して最小化が安定であることを示す。安定性のレンマについては、[33]-[35]を参照。

ℓ_1

Φ

$|J|$

Φ

η_i ρ η_0 η_i

$J \subset \{1, \dots, m\}$

$$|J| \le \rho_1 n/\log(m), \qquad \rho_1 \le \rho$$

$\in ker(\Phi)$ $\nu(\Phi, J) \le \eta_0 \cdot \rho_1^{1/2}$. J

$|J| < \rho n/\log(m)$

Φ $J = \{1, \dots, |J|\}$

$\beta = [\beta_J, \beta_{J^c}]$ $v = \Phi_J \beta_J$

$-v = \Phi_{J^c}\beta_{J^c}$

$\le \eta_2\eta_3 \cdot$

$\le c_1 \cdot$

$$\|\beta_J\|_2 \le \|\Phi_J\beta_J\|_2/\eta_1.$$

$\|\beta_J\|_1$ $\sqrt{|J|} \cdot \|\beta_J\|_2$

$$\|\beta_J\|_1 \le \sqrt{|J|} \cdot \eta_1^{-1}\eta_2\eta_3 \cdot \sqrt{\frac{\log(m)}{n}} \cdot \|\beta\|_1.$$

$\eta_0 = \eta_2\eta_3/\eta_1$ □

$x = \Phi\theta$ θ

$J \subset \{1, \dots, m\}$

$\nu(\Phi, J) < 1/2$ θ

(P_1) θ' (P_1)

$\theta' = \theta + \beta$ $\beta \in ker(\Phi)$

$$\|\theta'\|_1 \le \|\theta\|_1.$$

$$0 \le \|\theta\|_1 - \|\theta'\|_1 \le \|\beta_J\|_1 - \|\beta_{J^c}\|_1.$$

$\nu(\Phi, J)$

$$\|\beta_J\|_1 - \|\beta_{J^c}\|_1 \le (\nu(\Phi, J) - (1 - \nu(\Phi, J)))\|\beta\|_1.$$

$\nu(\Phi, J) <$ $2\nu(\Phi, J) - 1 < 0$

$\|\beta\|_1 \le 0$

$\theta = \theta'$

$\eta_0 > 0$ ρ_0

その$\eta_0\sqrt{\rho_0} <$ $\rho_0 \le \rho$ $|J| \le \rho_0 n/\log(m)$

$\nu(\Phi, J) \le$ $<$ □

$0 < p < 1$

$0 < p < 1$

Φ

η_i ρ $C = C(p, (\eta_i), \rho, A, \gamma)$

$\hat{\theta}_{1,n}$ (P_1) $y = \Phi\theta_0$

$(n/\log m)^{1/2-1/p}$.

ℓ_1

ℓ_1

ℓ_1 ℓ_2 ℓ_∞

Lemma 4.2: θ_0 を R^m andbe θ_0 $\theta_0(i) 1_J(i)$ ルとし、それに対応するエントリを $\nu(\Phi, J) \leq \nu_0 < 1/2$ 持つ変異したベクトルとする。ここで、$\Phi\theta_0$; (P_1) (P_1)

$$\|\theta_0 - 1_J\theta_0\|_1 \leq \epsilon$$

ところ。のインスタンスを考えてみましょう。によって $y =$ のこのインスタンスの解 $\hat{\theta}_1$ は定義され、obeys。

$$\|\theta_0 - \hat{\theta}_1\|_1 \leq \frac{2\epsilon}{1 - 2\nu_0}. \quad \text{(IV.2)}$$

Lemma 4.2 の証明: 短く $\hat{\theta}\hat{\theta}_1$ とし、以下を設定する。$\beta \hat{\theta} = \theta_0 - \in ker(\Phi)$.の定義により

$$\|\theta_0 - \hat{\theta}\|_1 = \|\beta\|_1 \leq \|\beta_{J^c}\|_1/(1 - \nu_0)$$

ながら

$$\|\beta_{J^c}\|_1 \leq \|\theta_{0,J^c}\|_1 + \|\hat{\theta}_{J^c}\|_1.$$

$\hat{\theta}$ (P_1)

$$\|\hat{\theta}_J\|_1 + \|\hat{\theta}_{J^c}\|_1 \leq \|\theta_0\|_1$$

が解決する

ように、そ

してもちろ

ん

$$\|\theta_{0,J}\|_1 - \|\hat{\theta}_J\|_1 \leq \|1_J(\theta_0 - \hat{\theta})\|_1.$$

それゆえ

$$\|\hat{\theta}_{J^c}\|_1 \leq \|1_J(\theta_0 - \hat{\theta})\|_1 + \|\theta_{0,J^c}\|_1.$$

最後に

$$\|1_J(\theta_0 - \hat{\theta})\|_1 = \|\beta_J\|_1 \leq \nu_0 \cdot \|\beta\|_1 = \nu_0 \cdot \|\theta_0 - \hat{\theta}\|_1.$$

上記を組み合わせて $\delta = \|\theta_0 - \hat{\theta}\|_1$ と $\epsilon = \|\theta_{0,J^c}\|_1$ とすると、以下のようになります。

$$\delta \leq (\nu_0\delta + 2\epsilon)/(1 - \nu_0),$$

となり、(IV.2)が成り立つ。 □

定理9の証明：定理7と同じ一般的な枠組みを用いる。$y = \Phi\theta_0$ Letwhere. の解を $\hat{\theta}$ とし、θ_0 を設定する。(P_1) $\beta = \hat{\theta} - \in ker(\Phi)$ $\|\theta_0\|_p \leq R$

Lemma η_0 $\rho_0 = \min(\rho, (4\eta_0)^{-2})$ 4.1と同様にして J θ_0 設定。

$\|\theta_0 - 1_J\theta_0\|_1 \leq \xi_{1,p} \cdot R \cdot |J|^{1-1/p} \epsilon \equiv$

$$\nu(\Phi, J) \leq \eta_0\rho_0^{1/2} \leq 1/4.$$

V. きんきゅうえんちょう

続けて述べる前に、ここまでの結果に対する2つの直接的な拡張について触れておく。これらは、以下でも他の場所でも興味深いものである。

A. タイトフレーム

これまでの我々の主な結果は、直交基底を作るという (ψ_i) 文脈で述べられてきた。実は、この結果は*タイトフレーム*でも成り立つ。タイトフレームは、ベクトルの集まりであり、これを $m \times m'$ 行列の Ψ $(m < m')$ 列として結合すると $\Psi\Psi^T = I_m.$ $\theta(x)\Psi^T x$ とすれば、以下のようにパースバル関係が成り立つ。

$$\|x\|_2 = \|\theta(x)\|_2$$

と再構成の公式 $x = \Psi\theta(x)$ がある。実は、定理7と9は、証明にパースバル関係しか必要ない。したがって、and の関係が tight x θ Φ $n \times m$ frame場合でも、同じ結果がそのまま成立する。特にCS1-CS3 を満たすanmatrix であれば、ほぼ最適な情報演算子が定義れ、最適化問題

$$(L_1) \quad \min_x \|\Psi^T x\|_1 \quad \text{subject to } I_n(x) = y_n$$

の $\rho_0 n/\log(m)$ 最大振幅のエントリをインデックスとする

から

$\|\theta_0\|_p$ となり、(II.4)となる。

となり、レンマ4.1より

レンマ4.2を適用すると

は、ほぼ最適な再構成アルゴリズム$\hat{x}_1$を定義する。

レフェリーから、ここでタイトフレームに注目する必要はない、一般的なフレームがあれば、フレーム境界を含む定数で同じ結果が得られる、という指摘がありました。これは事実であり、潜在的に非常に有用であるが、以下では使用しない。

B. 弱いℓ_pボール

これまでの主な結果は ℓ_p 空間について述べたものであるが、その証明は弱い ℓ_p 球についても同様に成り立つ 1) 。半径の弱い ℓ_p 球は、その減少する再配列 $|\theta|_{(1)}$ $|\theta|_{(2)}$ $|\theta|_{(3)}$

$$\|\beta\|_1 \leq c \cdot R \cdot |J|^{1-1/p}. \quad \text{(IV.3)}$$

の $\delta = \beta/\|\beta\|_1$ ベクトル $ker(\Phi)$ であり、$\|\delta\|_1 = 1$ 。したがって

$$\|\delta\|_2 \leq c \cdot (n/\log(m))^{-1/2}.$$

これを(IV.3)と組み合わせると

$$\|\beta\|_2 \leq c \cdot \|\beta\|_1 \ (n/\log(m))^{-1/2}.$$

、均質性によって次のように

結論づけられる。

$$\|\beta\|_2 \leq c \cdot R \cdot (n/\log(m))^{1/2-1/p}. \quad \square$$

が従うベクトルから構成されます。

$$|\theta|_{(N)} \leq R \cdot N^{-1/p}, \qquad N = 1, 2, 3, \ldots.$$

逆に，与えられた θ
、これらの等式がすべて成立する最小のものは R ノルムと定義される $\|\theta\|_{w\ell_p}$ ：。

弱い "という名称は、$\|\theta\|_{w\ell_p}$ に由来する。weak ℓ_p constraints $(p<1)$ は次のような重要な性質を持っている：最大の項目を除くすべての項目を0としたベクトルの変異版を示すと、不等式 $\geq \geq \geq \cdots$

$$\|\theta - \theta_N\|_q \leq \zeta_{q,p} \cdot R \cdot (N+1)^{1/q-1/p} \quad \text{(V.1)}$$

は $p < 1$ $q = 1, 2$and, $R = \|\theta\|_{w\ell_p}$ θ R withに対して有効である $\|\theta\|_p \geq \|\theta\|_{w\ell_p}$ $\|\theta\|_{w\ell_p} \leq R$ 実際、定理7と9は(V.1)を $Y_{p,m}(R)$ $\|\theta\|_p$ θ_N (暗黙のうちに)と共に証明に必要としただけである θ 。 N 弱規範だけで定義された空間に対する証明はそのまま適用される。

VI. 定型化されたアプリケーション

我々は、上記の抽象的な理論の3つの潜在的な応用をスケッチする。それぞれの場合において、あるクラスの関数

$$\hat{f} = \sum_k \beta_{j_0,k}\varphi_{j_0,k} + \sum_{j=j_0}^{j_1-1} \hat{\alpha}_{j,k}\psi_{j,k}$$

は、特定の n ℓ_p m または弱い ℓ_p 埋め込みに従う基底またはフレームに展開係数があり、その後、上記の抽象的な理論を適用します。

$$\|\hat{\theta} - \theta\|_2 \le c\|\theta\|_1 \cdot (n/\log(m))^{-1/2}$$

A. バンプ・アルジェブラ

$$\le c2^{-\frac{3}{2}j_0} \le c2^{-j_1/2}$$

バンプ代数 $\mathcal{F}(B)$ $\in$ [2] c $\in \mathcal{F}(B)$ に属する関数の単位区間への制限であり、バンプノルム $\mathcal{B}$ $\le$を持つ $\|f\|_{\mathcal{B}}$ B 関数 $f(t), t$

$$= \|f - P_{j_1}f\|_2 +$$

$$\|f - \hat{f}\|_2 \le \|f - P_{j_1}f\|_2 + \|P_{j_1}f - \hat{f}\|_2 \quad C$$

$$\le c2^{-j_1/2} \le C \cdot B \qquad \|\hat{\theta} - \theta\|_2$$

$[0,1]$ c $f \in \mathcal{F}(B)$ の クラスについて考えて。これは、レベル j のウェーブレット係数が $\|\theta^{(j)}\|_1$ 2^{-j} に従うことを観察した「はじめに」で述べたもので、ここで使用するウェーブレットにのみ依存するものである。以下、[36], [37], [2]と同様に、標準的なウェーブレット解析の表記を用いる。

f 古典的な線形方式では，「最短スケール」 j_1 を固定し，レジュメ係数$\beta_{j_1,k}$ を測定します. $- k$)に積分する滑らかな関数で φ $= \langle f, \varphi_{j_1,k}\rangle$. $\varphi_{j,k}$ = $2^{j/2}\varphi(2^j t$ アンチエイリアスフィルタを適用した後のスケール（ $P_{j_1}f = \sum_k \beta_{j_1,k}\varphi_{j_1,k}$

2^{-j_1}）の点サンプルとお考えください。再構成は近似誤差

$$\|f - P_{j_1}f\|_2 \le C \cdot \|f\|_{\mathcal{B}} \cdot 2^{-j_1/2}$$

は、選択されたウェーブレットにのみ C 依存するあります。 $N = 2^{j_1}$ $(\beta_{j_1,k})_k$ 単位区間に付随しているので、近似誤差は

$$\|f - P_{j_1}f\|_2 \le C \cdot \|f\|_{\mathcal{B}} \cdot N^{-1/2}.$$

圧縮センシング方式では、さらにウェーブレット $= 2^{j/2}\psi(2^j x -$ $\psi_{j,k}$ $k)$ ここで、ψ は平均ゼロの振動関数である。最も粗い 2^{j_0} スケールを選ぶ $\theta \equiv (\theta_\ell)_{\ell=1}^m$ $((\alpha_{j,k}$ 0 $2^{j_1} -$ 。我々は、レジュメ係数$\beta_{j_0,k}$ を測定する - 存在する- そして詳細なウェーブレット係数の列挙を示すとする。cients : $\le k < 2^j$): $j_0 j_1 \le j <$ 次元 m の和 E と$m =$ 2^{j_0}で $\|\theta\|_1$ 。ノルムは以下を満たす。

$$\|\theta\|_1 \le \sum_{j_0}^{j_1} \|(\alpha_{j,k})\|_1 \le \sum_{j_0}^{j_1} C \cdot \|f\|_{\mathcal{B}} \cdot 2^{-j} \le cB2^{-j_0}.$$

$n =$

これにより、我々の理論に必要な ℓ_1 ノルムに対する制約が確立される。 $c \cdot 2^{j_0}\log(2^{j_1})$ を取り、これに対する近最適情報演算子を適用し、（詳細は後述する）。 ℓ_1 の最小化の近最適アルゴリズムを適用し、誤差推定値を得る。

で、f に依存しない。全体の再構成

ハズレ

を再び独立させた。これは線形サンプリングの誤差と同じオーダーである。

圧縮センシング方式は、レジューム係数の $: 0 \le k_1, k_2 < 2^j, \sigma \in \{h, v, d\}) : j_0 \le j < j_1)$ サンプルと θ $m = 4^{j_1} - 4^{j_0}$ クライアントと。$\|\theta\|_1$ 詳細係数に m 関連 する ディメンジョンは ノルマは 2^{j_0} サンプルの

$$\|\theta\|_1 \le \sum_{j=j_0}^{j_1} \|(\theta^{(j)})\|_1 \le c(j_1 - j_0 + 1)\|f\|_{BV}.$$

$n \le c2^{j_0} \log(2^{j_1})$ $\le c \cdot j_1 \cdot 2^{j_1/3}$

合計を、情報の総量とする。これにより、従来のサンプルに基づくサンプリングと同程度の誤差を実現する。つまり、線形サンプリングのサンプル数の立方根程度で、同等の精度を得ることができる。

これによって、我々の理論を適用するために必要なノルムの制約が確立される。2^{j_1} $n = c \cdot 4^{j_0} \log^3(4^{j_1})$ このため、コアオプショナルインフォメーションオペレータを n m 適用 し Φ 、.を適用する。

このようにサンプリングを劇的に減らすためには、CS1-CS3を満足するものに基づく情報演算子が必要である。基礎となる測定カーネルは次のような形式となる。

$$\xi_i = \sum_{j=1}^{m} \Phi_{i,\ell}\phi_\ell, \qquad i = 1, \dots, n \tag{VI.1}$$

ここで、$(\phi_\ell)_{\ell=1}^m$ コレクションは単にウェーブレットの列挙であり、$\psi_{j,k}$、$j_0 \le j < j_1$ と $0 \le k < 2^j$。

B. 境界型バリエーションイメージ

ここで、「はじめに」で述べた「境界付き変動」のイメージを持つモデルについて考える。全変動が最大 $\mathcal{F}(B)$ でも [38]であり、絶対値で≤に $f(x)$ $x \in [0,1]^2$ B 束縛される 、 領域を 持つ $\|f\|_\infty$ B 関数のクラスとする。はじめに」で、レベルj のウェーブレット共効 $\le C \cdot B$ C 率は、使用するウェーブレットにのみ依存 し、$\|\theta^{(j)}\|_\infty \le C \cdot B \cdot$ $\|\theta^{(j)}\|_1$ に従うと述べました。また、2^{-j} もそうである。f で関数を近似する2つの方法を再び考えてみます。古典的な線形スキームは、すでに説明したスキームの2次元バージョンを使用します。ここでも「最も細かいスケール」j_1を固定し、レジュメ係数を $\beta_{j_1,k} = \langle f, \varphi_{j_1,k} \rangle$ 測定します。ここで、(k_1, k_2)は $k =$ 整数の組で $0 \le k_1$ $k_2 < 2^{j_1}$、インデックスを付けます。の位置に配置する。ここでは、Haarスケーリング関数

$$\varphi_{j_1,k} = 2^{j_1} \cdot 1_{\{2^{j_1}x - k \in [0,1]^2\}}.$$

近似誤差を与えて再構築 $P_{j_1} f = \sum_k \beta_{j_1,k}\varphi_{j_1,k}$ します。

$$\|f - P_{j_1} f\|_2 \le 4 \cdot B \cdot 2^{-j_1/2}.$$

$\beta_{j_1,k}$ 単位二乗に関連する係数 $N = 4^{j_1}$ が あるので、近似誤差は次のようになる。

$$\|f - P_{j_1} f\|_2 \le c \cdot B \cdot N^{-1/4}.$$

圧縮センシング方式では、Haar ウェーブレット $\psi^\sigma_{j_1,k} = 2^{j_1}\psi^\sigma(2^{j_1}x - k)$ も必要である。ここで ψ^σ は平均ゼロの振動関数で、水平方向か $(\sigma = h)$ 垂直方向(σ $(\sigma = v)$ または斜め方向 $(\sigma = d)$.最も粗い $j_0 = j_1/2$ スケール」を 選び、そのレジューム係数$\beta_{j_0,k}$ -there 4^{j_0} are ofthese を測定する。そして、詳細なウェーブレット係数の連結を "Wavelet "とする。$\theta \equiv (\theta_\ell)_{\ell=1}^m$ $((\alpha^\sigma_{j,k}$

ℓ_1

$$f = \sum_\mu \langle f, \gamma_\mu \rangle \gamma_\mu.$$

$$\|\theta^{(j)}\|_{w\ell_{2/3}} \le c(B, L), \qquad f \in \mathcal{C}^{2,2}(B, L);$$

$c \in \mathcal{F}(B)$

の最適なアルゴリズムで、得られた情報に対して、誤差の推定値を得る。

$$\|\hat{\theta} - \theta\|_2 \le c\|\theta\|_1 \quad (n/\log(m))^{-1/2} \le cB \quad 2^{-j_0}$$

$$\hat{f} = \sum_k \beta_{j_0,k}\varphi_{j_0,k} + \sum_{j=j_0} \hat{\alpha}_{j,k}\psi_{j,k}$$

$$\|f - P_{j_1}f\|_2 \le c \cdot B \quad 2^{-j_1/2}, \qquad f \in \mathcal{C}^{2,2}(B, L).$$

$$\|f - \hat{f}\|_2 \le \|f - P_{j_1}f\|_2 + \|P_{j_1}f \quad \hat{f}\|_2$$

$$\|\hat{\theta} - \theta\|_2$$

で、f に依存しない。全体の再構成

ハズレ

$$= \|f - P_{j_1}f\|_2 +$$

$$\le c2^{-j_1/2}$$

を再び独立させたもので、 $c \in \mathcal{F}(B)$ f 。これは線形サンプリングの誤差と同じオーダーである。

圧縮センシングスキームは、レジューム係数のサンプル 4^{j_0} と詳細係数に関連するサンプルの合計 $n \sim 4^{j_0} \log^3(4^{j_1})$ $4^{j_1/2}$ の測定情報を取得する。これは、 $\le c \cdot j_1^3$ 4^{j_1} のサンプルでクラスシカルサンプリングと同等の誤差を達成します。このように、Bump Algebraのケースで見たように、同等の精度を得るために必要なサンプル数は劇的に少なく、おおまかに言えば、線形サンプリングのサンプル数の平方根程度で済むのです。

C. PiecewiseC^2 画像C^2 エッジを持つもの

ここで、 $p < 1$ 、タイトフレームへの拡張と、先に述べたweak- ℓ_p への拡張を適用する例を考える。再び画像処理の設定で、Candès and Donoho [39], [40]で議論された C^2 - C^2 $\mathcal{C}^{2,2}(B, L)$ モデルを使用します。適切な $\in [0,1]^2$ B $\le$ パラメトリザーションで1次および2次導関数 Γ $\le$ を持つ曲線 B C^2 の和である例外集合から，値，1次および2次偏導関数を持つ L 区分的に滑らかな $f(x), x$ をcol-lectionことを考えてみましょう．より色彩的に言えば、このような画像は*カートゥーン*-断片的に滑らかな曲線の境界で区切られた、一様な滑らかな挙動の*パッチ*-である。このような画像は、ある種の技術的な画像、たとえば放射線医学の分野では妥当なモデルといえるかも

$$\|f\|_2^2 = \sum_\mu |\langle f, \gamma_\mu \rangle|^2$$

しれない。

カーブレットタイトフレーム[40]は、パースバル関係を提供する滑らかなフレーム要素の集合体です

と再構成式

フレーム要素はマルチスケール構成で、スケールごとにグループ化されたフレーム係数 $\theta^{(j)}$ は、弱い ℓ_p 制約に従う。

40]と比較してください。このようなオブジェクトの場合、滑らかな2次元スケーリング関数によるスケール j_1 における古典的な線形サンプリングは、次のようになります。

これは，Bounded Variationの場合の線形サンプリングの性能と同じである．

Bounded Variationと比較した場合。

圧縮センシングスキームでは、最も粗いスケールである $j_0 j_1/4$ を選ぶ。我々は、滑らかなウェーブレット展開のレジューム係数 $\beta_{j_0,k}$ を測定し（これらの 4^{j_0} がある）、より細かいスケールのカーブレット係数 $: j_0 j_1)$ の連結を記録するの過完結のため，次元は次のようになる。

カーブレット

弱い「ノルム」は次のようになる。

$$\|\theta\|_{w\ell_{2/3}} \le \left(\sum_{j_0}^{j_1} \|(\theta^{(j)})\|_{w\ell_{2/3}}^{2/3} \right)^{3/2} \le c(j_1 - j_0 + 1)^{3/2}$$

とに依存している。これにより、我々の理論に必要な弱い ℓ_p ノルムに対する制約が確立される。我々は c' $n = c \cdot 4^{j_0} \log^{5/2}(4^{j_1})$

に対して、近最適情報演算子を適用し、and .得られた情報に対して、ℓ_1 最小化の近最適アルゴリズムを適用し、誤差推定値 n m

$$\|\hat{\theta} - \theta\|_2 \le c\|\theta\|_{2/3} \cdot (n/\log(m))^{-1} \le c' \cdot 2^{-j_1/2}$$

absoluteと。全体の再構成は

$$\hat{f} = \sum_k \beta_{j_0,k}\varphi_{j_0,k} + \sum_{j=j_0}^{j_1-1} \sum_{\mu \in M_j} \hat{\theta}_\mu^{(j)} \gamma_\mu$$

ハズレ

$$\|f - \hat{f}\|_2 \le \|f - P_{j_1} f\|_2 + \|P_{j_1} f - \hat{f}\|_2 = \|\hat{\theta} - \theta\|_2 \le c 2^{-j_1/2}$$

を再び独立させたもので、c $f \in C^{2,2}(B,L)$。これは線形サンプリングの誤差と同じオーダーである。

圧縮センシング方式は、レジューム係数と詳細係数に関連する $n \le c4^{j_0} \log^{5/2}(4^{j_1}) \le c \cdot j_1^{5/2} \cdot 4^{j_1/4}$ のサンプルの合計で、情報量が1個になる。この方式は、4^{j_1} のサンプルに基づく古典的なサンプリングに匹敵する誤差を達成する。つまり、Bump Algebraの場合よりもさらに、同等の精度を得るために必要なサンプル数は劇的に少なくなり、おおまかに言えば、線形サンプリングのサンプル数の4分の1の数で済む。

VII. NEARLY ALL MATRICES ARE CS MATRICES

定理6を次のように再定式化することができる。

定理10: , n m_n を問題サイズのシーケンスとする。$n \to \infty$ $n < m \sim An^\gamma$ と $A > 0$ $\gamma > 1$。$\Phi = \Phi_{n,m}$ とする。m S^{n-1} から独立に一様にランダムに引いた列を持つ行列とする。このとき、ある $\rho > 0$ 場合、条件CS1〜CS3は、すべての大きな n に対して圧倒的な確率で成立する。

実際、 S^{n-1} 上の列を i.i.d. 均等にサンプリングすることによって生じる $\Phi_{n,m}$ 上の確率測度は、まさに $\Phi_{n,m}$ 上の自然な均一測度であることに注意しよう。したがって、定理6は定理10から即座に導かれる。

事実上、CS条件 Φ を満たす行列はどこにでもあるため、*一様な*確率分布から*無作為に*サンプリングして生成することが合理的である。

定理10の証明は、第VII節から第VIII節にわたって行われる。

A-C; イベント Ω_n^i, $i = 1, 2, 3$, where Ω_n^i { CS1 Holds }, などを研究することによって進められる。パラメータ η_i と $\rho_i > 0$

$$P\left(\Omega_n^i\right) \to 1, \qquad n \to \infty;$$

とすると、$\min_i \rho_i$ と Ω_n を定義 $= \cap_i \Omega_n^i$ すると、次のようになります。

$$P\left(\Omega_n\right) \to 1, \qquad n \to \infty.$$

Ω_n が発生したとき、ランダムドローは CS1-CS3 に従うマトリクスをパラメータと η_i ρ > 0 して生成しているので、定理 10 が証明される。この証明は、実際には、ある β に対して、次のように示される。 $P\left(\Omega_n^c\right) \le \exp(-\beta n)$, $n > n_0$ 収束は指数関数的である。を高速化することができます。

A. 最小固有値の制御

次のレンマは、正の定数を選び ρ_1 η_1 、条件CS1が圧倒的に成立するようにするものである。 $J \subset \{1, \ldots, m\}$ Ω_n $\lambda_{min}\left(\Phi_J^T \Phi_J\right)$ $\lambda < 1$ > 0 $\rho_1 > 0$ の確率が高い。 $n > n_0$

*レンマ7.1。*を持つ (n, m_n) の列を考える。 $n \le m_n \sim An^\gamma$.イベントの定義

$$\Omega_{n,m,\rho,\lambda} = \left\{\lambda_{min}\left(\Phi_J^T \Phi_J\right) \ge \lambda,\ \forall\, |J| < \rho \cdot n/\log(m)\right\}.$$

次に、 $\lambda < 1$ それぞれについて、十分に小さい $\rho > 0$

$$P(\Omega_{n,m,\rho,\lambda}) \to 1, \qquad n \to \infty.$$

この証明には3つの考え方がある。最初の考え方は、レンマ 7.1 で注目される事象は、個々の部分集合 J

$$\Omega_{n,m,\rho,\eta} = \cap_{|J| \le \rho n/\log(m)} \Omega_{n,J}.$$

我々の計画は、すべての $\Omega_{n,J}$ の発生確率を束縛することである。第二のアイデアは、特定の部分集合に対して、最小固有値の大きなデビジョン境界を得ることである；これは次のように述べることができる。

最後の3つ目のアイデアは、個々の部分集合に対する境界が、すべての J .NETの同時動作を制御 J することができるというものである。これは次のように表現される。

*レンマ7.3：*ある $\Omega_{n,J}$ $\beta > 0$ 固定の $n > 0$ andに従ったすべてのイベントがあるとする。

$$P\left(\Omega_{n,J}^c\right) \le \exp(-n\beta),$$

それぞれについて $J \subset \{1, \ldots, m\}$ と $|J| \le \rho n$. ρ_1 をもって $\beta_1 < \beta$ ピッタウィズ $\rho_1 < \min(\beta, \rho)$ と $\beta_1 > 0$ $\beta - \rho_1$ 。次に、すべての十分大きな n

$$P\left\{\Omega_{n,J}^c \text{ あるいみで } J \subset \{1, \ldots, m\} \text{ で } |J| \le \rho_1 n/\log(m)\right\} \le \exp(-\beta_1 n).$$

本節の主要な目的であるレンマ7.1は、これら3つの考え方を組み合わせることで導かれる。

あとは、レンマ7.3を証明するだけである。とする。

$$\mathcal{J} = \{J \subset \{1, \ldots, m\} \text{with } |J| \le \rho_1 n/\log(m)\}.$$

ブールの不等式により、以下のようになることに注意。

を以下のとおりとします。

レンマ7.2 である。 に対して、 最小固有値を超える 事象を表すとする。そのとき、β_1 があり、十分に小さく、 すべての

$$P\left(\Omega_{n,J}^c\right) \le \exp(-n\beta_1)$$

$$\sqrt{\frac{\pi}{2n}}.$$

$$P\left(\cup_{\mathcal{J}}\Omega_{n,J}^c\right) \leq \sum_{J\in\mathcal{J}} P\left(\Omega_{n,J}^c\right) \leq \#\mathcal{J}\cdot\exp(-\beta n),$$

最後の不等式は、各メンバー $\mathcal{J}$ が カージナル数であることから、 $\rho_1 n/\log(m)$ 、すぐに次のようになります。 $\geq$.また、当然ながら

というわけで、≤が得られる 。β_1 を所与とすると、目的の結論が得られる。

B. 球状断面積の特性

ここで、条件 CS2 は、十分に小さいが正の値を選ぶことで、大きな値に対して圧倒的な可能性を持つことができることを示す。我々のアプローチは、バナッハ空間理論からの重要な結果である、ほぼ*球面分割現象を*適用した

$|J| \leq \rho_1 n$で統一しています 。

これは、ランダム行列の特異値の測度集中の性質を利用して、 [11] や [35] で導かれています。例えば、Szarek の仕事 [41], [42] を参照してください。

[11]に由来している。これは、バナッハ空間の単位球を適当な有限次元線形部分空間と交差させると、実質的に球形となるスライスが得られるというものである [43], [44]。我々は、この原理を の $\boldsymbol{R}^{J}$ $J\in$ ℓ_1 ノルムのために定量的に改良し、圧倒的な確率で、$\leq\rho n$ $\leq\rho m$ $m\geq$ $\rho n/\log(m)$ のための すべての演算子 が Φ_J ℓ_1^n n $3>\epsilon$ の球の球形断面を与えることを示す。我々が用いた基本

$$\log(\#\mathcal{J}) \log\binom{m}{\rho_1 n} \leq k\log(m),$$

的な議論は、Milman, Kashin, and others [44], [10], [45]の研究に由来するもので、Pisier [19] の議論を改良し、[11]のように、新しい推論を導き出した。我々は、ほぼ球形断面が存在するだけでなく、あらゆる n η_2 ρ_2

withwillで生成さほど、どこにでも存在することを結論付ける。

定義 .を

7.1: の間のレトリー 提供する「イソ kム」と言います。 Φ ϵ $\forall\,\alpha\in\boldsymbol{R}^k$ $|J|<$

$$(1-\epsilon)\cdot\|\alpha\|_2 \leq \|\Phi_J\alpha\|_1 \leq (1+\epsilon)\cdot\|\alpha\|_2. \quad \text{(VII.1)}$$

Φ_J $|J|<\rho n/\log(m)$

イ$|J| =$
$\ell_2(J)$ そして ℓ_1^n
、もし

次のレンマは、条件CS2が行列の一般的な性質であることを示す。

定理7.4：すべての $\Omega_n^2\left(\equiv \Omega_n^2(\epsilon,\rho)\right)$

Φ_Jと$\ell_1^n(J)$の間の$\rho\sqrt{\log(m)}$-isometryを提供します。 $\ell_2(J)$

それぞれについて $\epsilon > 0$、> 0 $\rho(\epsilon)$

が存在し、以下のようになる。

$$P\left(\Omega_n^2\right) \to 1, \qquad n \to \infty.$$

これを証明するためには、まず[11]で証明された個々のJ部分集合に関するレンマが必要である。

レンマ7.5： $\epsilon > 0$修正 δ。 を 選択

$$(1-3\delta)(1-\delta)^{-1} \geq (1-\epsilon)^{1/2} \tag{VII.2}$$

と

$$(1+\delta)(1-\delta)^{-1} \leq (1+\epsilon)^{1/2}. \tag{VII.3}$$

$\rho_1 = \rho_1(\epsilon)$ 選択

$\delta^2 \frac{2}{\pi}$

$v \in$

y

$i \in J^c$

ρ_1 $(1+2/\delta) <$

$\delta > 0$

とし、両者の差を表すとする。

$\rho_1 < \delta^2/128$

J $\{1,\ldots,m\}$の部分集合に対して 、

ϵ $\Omega_{n,J}^n$ は、 Φ_J が ℓ_1^nへの -isometry

を与えるという事象を表すとすると、次のようになる。

$$\max_{|J|\leq \rho_1 n} P(\Omega_{n,J}^c) \leq 2\exp(-\beta(\epsilon)n(1+o(1))).$$

ここで、レンマ7.4で注目される事象は

$$\Omega_n^2 = \cap_{|J|\leq \rho n/\log(m)} \Omega_{n,J};$$

を適用し、Lemma 7.3 の結合原理を適用する。

C. 商慣行不等式

ここで、3/4 、十分に小さい場合、次のことを示す $\eta_3 > 0$という性質があり、ほぼ全ての大きな by 行列は CS3の性質を持ちます。我々の議論は、読者が参考にすると思われる[11]から大いに借用したものである。ここでは、直観を提供したり、バナッハ空間の局所理論における密接に関連した概念（例えば、Milmanの部分空間の商定理 [19]）と比較するような試みはしていない。

$\{1,..J..,m\}$ の任意のインデックスの集まりとする。$Range(\Phi_J)$ は Σ_J $\boldsymbol{R}^n$

の線形部分空間であり，この部分空間上で，可能な符号パターンの部分集合，すなわち

$$\sigma(k) = \mathrm{sgn}\left\{\sum_J \alpha_i\phi_i(k)\right\}, \qquad 1 \leq k \leq n.$$

CS3は、 $\mathrm{sgn}(v)$ のある近似値が $Range(\Phi_I)$, $|\langle y,\phi_i\rangle| \leq 1$

for

を満たすことを示せれば、それに従う

。

と

$$\forall\, i \in J^c, \qquad |\langle \phi_i, y\rangle| \leq 1. \tag{VII.6}$$

つまり、どんな符号パターンでも、その

小さな 倍数が

$\epsilon_n\sigma$ デュアルボールの中でほとんど生きているのです。$\{x : |\langle\phi_i,x\rangle| \leq 1, i \in J^c\}$。

この結果を証明する前に、この結果がどのようにプロパティーCS3を与えるかを示す。すなわち、 $v = -\Phi_{J^c}\beta_{J^c}$

$|J| <$ $\rho_3 n/\log(m)$、および

ふくむ

$$\|\beta_{J^c}\|_1 \geq \eta_3/\sqrt{\log(m/n)} \cdot \|v\|_1. \tag{VII.7}$$

凸最適化問題を考える

$$\min \|\beta_{J^c}\|_1 \quad \text{条件として} \quad \Phi_{J^c}\beta_{J^c} = -v. \tag{VII.8}$$

これは、(IV.1)と同じような構成で、線形計画として書くことができる。線形計画法の双対定理により、原始プログラムの値は少なくとも

双対プログラムの値である。

Lemma 7.6: 一様な符号パターン埋め込み。を修正する

。次に、

for,set τ

$\max\langle v, y\rangle$ 会わせる 1,

レンマ 7.6 は、双対実行可能なベクトルの供給と、それによる双対プログラムの下界を与えてくれます。 $\mathrm{sgn}(v)$ を使って、どれが双対実行可能で、どれに従うかを見つけることができます。

を適切に選び、球面分割の定理を考慮すると、十分に大きな , が得られる。$\delta\|\sigma\|_2\|v\|_2$

; (VII.7) は 3/4 で導かれる。

1) 一様な符号パターン埋め込みの証明。 レンマ7.6の証明は、この

ケースを検討した[11]の同様の結果に密接に従ったものである 。ここで、我々の考えは、この場合の論証を

の選択の違いを反映した変更と、スパース性境界 $\rho n/\log(m)$

を用いた場合である。論証の大部分は[11]の対応する部分と同じであるため、割愛する。我々の努力の大部分は以下のレンマの生成に費やされ、*単一符号パターン*

$$\epsilon_n = (\log(m_n/(\tau n)))^{-1/2}/4. \tag{VII.4}$$

十分に $\rho_3 > 0$小さい場合は $\Omega_n^3 \equiv$

、

イベントがあります $\Omega_n^3(\rho_3, \delta)$

$P\left(\Omega_n^3\right) \to 1$with $n \to \infty$ J

$|J| < \rho_3 n/\log(m)$ as今回のイベントでは、各サブセットに対して $y(\equiv y_\sigma)$

で $\sigma \in \Sigma_J$、

各符号パターンに対して、

ベクトルwithが存在します。

$$\|y - \epsilon_n\sigma\|_2 \le \epsilon_n \quad \delta \cdot \|\sigma\|_2 \tag{VII.5}$$

の dual ball への近似的な埋め込みを証明するものである。 $|\langle\phi_i, y\rangle| <$ $i \in J^c$.

レンマ7.7: **個別の符号パターン埋め込み。** $\{-1,1\}^n\sigma \in$ を、y $y_0 = \epsilon_n\sigma$, ϵ_n m_n, τ δ k 0 (ϕ_i) " i $m -$ とLemma 7.6 の記述のように する $\geq$とする。

$$\langle v, y\rangle \ge \langle v, \epsilon_n\sigma\rangle - \|y - \epsilon_n\sigma\|_2\|v\|_2$$

δ

$$\ge \epsilon_n\|v\|_1 - \epsilon_n\delta\|\sigma\|_2\|v\|_2;$$

コレクションが n $\le$

$\frac{1}{4}\|v\|_1$ 与えられたとき、≤≤ n $\geq$

k)に従う

y ベクトルを出力する、以下に述べる反復アルゴリズムが存在する。$n < m < Am$ $n < m < Am$ $n < m \sim Am^\gamma$

$$|\langle\phi_i, y\rangle| \le 1, \quad i = 1, \ldots, m - k. \tag{VII.9}$$

$(\phi_i)_{i=1}^{m-k}$ を $\boldsymbol{S}^{n-1}$ i.i.d.一律とし Ω_σ δ

、以下の確率で制御される事象が存在する。

$$\mathrm{Prob}\,(\Omega_\sigma^c) \le 2n\exp\{-n\beta\} \tag{VII.10}$$

$\beta > 0$ を明示的に δ

与えることができます。この事象について

$$\|y - y_0\|_2 \le \delta \cdot \|y_0\|_2. \tag{VII.11}$$

レンマ7.7はそれ自身のセクションで証明される。ここでは、それがemma 7.6を含意することを示す。

いわゆるVapnik- y_0

$$y_1 = y_0 - \langle P_{I_0} y_0 \rangle | > 1$$
$$|\langle \phi_i, y_0 \rangle| > 1/2$$

Cervonenkis理論[46]の標準的な含意を思い起こす。

$$\#\Sigma_J \le \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{|J|} \qquad 1$$

y_2 y_3

$\rho n/\log(m)$, ならば、に注意してください。$|J| <$

$$y_{\ell+1} = y_0 - P_{J_\ell} y_0.$$

$$J_\ell = I_0 \cup \cdots \cup I_\ell$$

$$\log(\#\Sigma_J) \le \rho \cdot n + \log(n)$$

一方

$$\log \#\{J : |J| < \rho n/\log(m), J \subset \{1, \dots, m\}\} \le \rho n.$$

$$\log \#\{\sigma : \sigma \in \Sigma_J, |J| < \rho n/\log(m)\} \le n \cdot 2\rho + \log(n).$$

したがって、演算子によって生成される符号パターンの総数Φ_Jは次のようになる。

さて、レンマ7.7によって与えられたβは正なので、$\rho_3 < \beta/2$ $\rho_3 > 0$を選びます

$$I_\ell = \{i : |\langle \phi_i, y_\ell \rangle| > t_\ell\}$$

。

を定義する。

σ J

$$\Omega_n^3 = \Omega_n^3 \cap_{|J| < \rho_3 n/\log(m)} \cap_{\sigma \in \Sigma_J} \Omega_{\sigma,J}$$

$|J| < \rho_3 n/\log(m)$

ここで、$\Omega_{\sigma,J}$ は、特定の , コンビ・ネイションによって生成されるイベントのインスタンス (Lemma 7.7 のステートメントでは Ω_σ と呼ばれる) を示す。イベント上では、*任意の* Φ_J 従う関連する*すべての*符号パターンがほぼ二元的に実行可能である。今

$$P\left(\left(\Omega_n^3\right)^c\right) \le \#\{\sigma : \sigma \in \Sigma_J, |J| < \rho_3 n/\log(m)\} \times \exp(-n\beta) \le \exp\{-n(\beta - 2\rho_3) + \log(n)\}$$

は指数関数的に急速にゼロになる傾向があります。 □

D. 個別符号パターン埋め込みの証明

*1) 埋め込みアルゴリズム:*関連論文[11]では、近傍のほぼ実現可能な点から出発して、二重実現可能な点を作成するアルゴリズムを紹介した。y_0これは次のように動作する。

y

を持つインデックス $\ll$ のコレクションを I_0 とする。$|\ell|$

を設定し

P_{I_0}

$\left(\Phi_{I_0}^T \Phi_{I_0}\right)^{-1} \Phi_{I_0}^T$

最も小さい正方形プロジェクター

Φ_{I_0} .つまり、禁止されている

レベルの半分を超えているインデックスを特定し、そのインデックスを"殺す"のです。

このプロセスを続け、

ステージごとに異なる閾値 $t_\ell 1 - 2^{-\ell-1}$ を順次.Set に近づけていき、

.Set を

生成

と、 パッティング 。

が空の場合、プロセスは終了し、セットされる $y = y_\ell$。終了は I_ℓ ステージで発生しなければならない $\ell^* \le n$。終了時

$$|\langle \phi_i, y\rangle| \le 1 - 2^{-\ell^*-1}, \quad i = 1, \ldots, m.$$

したがって、デュアルフィージビリティは間違いなくあります。唯一の問題は、 y それがどの程度近いかです。 y_0

*2) 解析フレームワーク*また、[11]では、アルゴリズムの軌跡の2つの重要な記述子について、境界が開発されました。

$$\alpha_\ell = \|y_\ell - y_{\ell-1}\|_2$$

と

$$|I_\ell| = \#\left\{i : |\langle \phi_i, y_\ell\rangle| > 1 - 2^{-\ell-1}\right\}.$$

そこで展開された議論を適応する。我々は、境界 $\delta_\ell \delta_{\ell;n}$ andforandを $\nu_\ell \equiv \nu_{\ell;n}$ α_ℓ $|I_\ell|$、な形で定義する。

$$\equiv n \cdot \lambda_0 \cdot \epsilon_n^2 \quad \ell = 0, 1, 2, \ldots; \quad \equiv$$

$$= \quad \delta_{\ell;n} = \|y_0\|_2 \cdot \omega^\ell \cdot \quad \ell = 1, 2, \ldots \quad \omega_0 > 0$$

ここで、λ_0 $1/2$ と $\nu_{\ell;n}$ $\cdot \omega^{2\ell+2}/4$, $\min(1/2, \delta/2, \omega_0)$、ここで、後で定義します。サブイベント

$$E_\ell = \left\{\alpha_j \le \delta_j,\, j = 1, \ldots, \ell, \quad |I_j| \le \nu_j,\, j = 0, \ldots, \ell - 1\right\}.$$

ここで、次のように定義します。

$$\Omega_\sigma = \cap_{\ell=1}^n E_\ell;$$

というのも、 $\omega \le$ このイベントが意味するのは1/2

$$\|y - y_0\|_2 \le \left(\sum \alpha_\ell^2\right)^{1/2} \le \|y_0\|_2 \cdot \omega/(1 - \omega^2)^{1/2} \le \|y_0\|_2 \cdot \delta.$$

と組み合わせて選ばれたβ、そのことを示す。 $\tau > 0$

$$P\left(E_{\ell+1}^c | E_\ell\right) \le 2\exp\{-\beta n\}. \tag{VII.12}$$

これは、次のことを意味しています。

$$P\left(\Omega_\sigma^c\right) \le 2n\exp\{-\beta n\}$$

となり、レンマが導かれる。 □

3) ラージデビエーション事象の定義

ように

$$F_\ell = \{\alpha_\ell \le \delta_{\ell;n}\}, \quad G_\ell = \{|I_\ell| \le \nu_{\ell;n}\}$$

置く

$$E_{\ell+1} = F_{\ell+1} \cap G_\ell \cap E_\ell.$$

$$\rho_0(\tau, \omega; n) = (128)^{-1} \frac{\log(An^\gamma)}{\log(An^\gamma/\tau n)} \frac{\omega^2}{1 - \omega^2}$$

であり、これは.NETにかなり弱く依存することに注意。 n という言葉で定義されていることを思い出してください。イベントに関して E_ℓ $|J_\ell| \le \rho_0 n/\log(m)$.レンマ7.1では、暗黙のうちにquan-to-quanを定義していました。 $\lambda_1(\rho, A, \gamma)$ E_ℓ の固有値の最小値を τ 下限と。

everywhere $\Phi_J^T\Phi_J$ $\rho n/\log(m) \leq$ ピックソは $\rho_{1/2} > 0$ $\lambda_1(\rho_{1/2}, A, \gamma) > 1/2$.ピック ω_0 ソは

$$\rho_0(\tau, \omega_0; n) < \rho_{1/2}, \qquad n > n_0.$$

ω_0のこの選択で、イベントE_ℓが発生したとき。

$$\lambda_{min}\left(\Phi_{I_\ell}^T \Phi_{I_\ell}\right) > \lambda_0.$$

また、E_ℓ、$u_j = 2^{-j-1}/\alpha_j$ $2^{-j-1}/\delta_j$（≧言う）上で、$1 \leq \ell$ 11]では、確率変数i.i.d.の族 n) $(Z_i^\ell : 1 \leq i \leq m, 0 \leq \ell <$、解析の枠組みを開発した。$N\left(0, \frac{1}{n}\right)$ が紹介され、その結果

$$P\left\{G_\ell^c | E_\ell\right\} \leq P\left\{\sum_i 1_{\left\{|Z_i^\ell| > v_\ell\right\}} > \nu_\ell\right\}$$

と

$$P\left\{F_{\ell+1}^c | G_\ell, E_\ell\right\} \leq P\left\{2 \cdot \lambda_0^{-1}\left[\nu_\ell + \delta_\ell^2\left(\sum_i \left(Z_i^\ell\right)^2 1_{\left\{|Z_i^\ell| > v_\ell\right\}}\right)\right] > \delta_{\ell+1}^2\right\}$$

この論文では、2つの簡単な大偏差の境界も述べている。

Lemma 7.8: Z_i を i.i.d.$N(0,1)$ $k \geq 0$, する2

$$\frac{1}{m}\log P\left\{\sum_{i=1}^{m-k} Z_i^2 1_{\{|Z_i| > t\}} > m\Delta\right\} \leq e^{-t^2/4} - \Delta/4,$$

と

$$\frac{1}{m}\log P\left\{\sum_{i=1}^{m-k} 1_{\{|Z_i| > t\}} > m\Delta\right\} \leq e^{-t^2/2} - \Delta/4.$$

これを応用して、イベント

$$\left\{2 \cdot \lambda_0^{-1}\left[\nu_\ell + \delta_\ell^2\left(\sum_i \left(Z_i^\ell\right)^2 1_{\left\{|Z_i^\ell| > v_\ell\right\}}\right)\right] > \delta_{\ell+1}^2\right\}$$

$$\tau_\ell^2 = n \cdot v_\ell^2 = \epsilon_n^{-2}(2\omega)^{-2\ell}/4$$

$N\left(0, \frac{1}{n}\right)$ 変数で表現すると、イベントと同じになります。

$$\Delta_\ell = \frac{n}{m}\left(\lambda_0 \delta_{\ell+1}^2/2 - \nu_\ell\right)/\delta_\ell^2.$$

$n > n_0$

$$e^{-\tau_\ell^2/4} = e^{-[(16\log(m/\tau n))/16]\cdot(2\omega)^{-2\ell}} = (\tau n/m)^{(2\omega)^{-2\ell}}$$

$$\frac{1}{m}\log P\left\{F_{\ell+1}^c | G_\ell, E_\ell\right\} \leq e^{-\tau_\ell^2/4} - \Delta_\ell/4.$$

$$\left\{\sum_{i=1}^{m-k} Z_i^2 1_{\{|Z_i| > \tau_\ell\}} > m\Delta_\ell\right\}$$

$$\Delta_\ell = \frac{n}{m}\left(\omega^2/4 - \omega^2/8\right) = \frac{n}{m}\omega^2/8.$$

である

ため

、

$\omega < 1/2$ $\ell \geq 0$

最も関心のある用語はatであり、他の用語の方が常に優れている。また，実際には，$(\tau n/m)^{(2\omega)^{-2\ell}}$ に依存しない．に注目し，次のように書くことができる Δ_ℓ 。 ℓ $\ell = 0$

$$\log P\left\{F_1^c | G_0\right\} \leq m\left(\tau \cdot n/m - n/m \cdot \omega^2/32\right)$$
$$= n\left(\tau - \omega^2/32\right).$$

を思い出し、$\omega \leq \delta/2$、パッティング

$$\beta \equiv \beta(\tau; \delta) = (\delta/2)^2/32 - \tau$$

> 0 $\tau < \delta^2/128$についてはβ、取得します。

$$P\left\{F_{\ell+1}^c | G_\ell, E_\ell\right\} \leq \exp(-n\beta).$$

G_ℓ'についても同様の分析が成り立つ。 □

VIII. 結論

A. 概要

我々は、ベクトルとして表現できるオブジェクトの圧縮センシングのための抽象的なフレームワークを説明した $\boldsymbol{R}^m$ 。我々は、関心のあるオブジェクトが*先験的に*圧縮可能であると仮定し、

既知の基底またはフレームに対して $x_0 \in$ $\boldsymbol{R}$ $\|\Psi^T x_0\|_p \leq$ となるようにする 。条件CS1-CS3を満たす Ψ $p < 2$ 行列から出発し、直交基底またはタイトフレームの行列を n m Φ $n < m$ 基にして Ψ

我々は情報演算子 $\mathcal{X}_{p,m}(R)$ を定義し $I_n(x)$

$= \Phi x$ のタプルから出発し n の情報 $y_n = I_n(x_0)$

x_0 を解くことで、その近似値を復元する。

$$(L_1) \quad \min_x \|\Psi^T x\|_1 \text{会わせる } y_n = I_n(x).$$

は標準的な$N(0,1)$確率変数で記述される。

と

したがって、以下の 不等式が成り立つ 。

m

現在

と

提案する再構成ルールは凸最適化を用いており、計算量的に扱いやすい。また，CS1-CS3
を満たす行列は，.NET Φ
の列に対する一様分布からのランダムサンプリングにより構成することができる. Φ

我々は、圧縮可能な物体に対して、見かけ上の非デサンプリング $(n$ $m)$ $<$
にもかかわらず、良い精度の再構成が可能であることを示す誤差境界を与え、これらの境界の最適に近い性質を Optimal Recovery と Informa- tion-Based Complexity からのアイデアを用いて説明する。さらに、この結果はデータの小さな測定誤差（ $\|y_n$ $-$ $I_n(x_0)\|_1$
small）に対しても安定であることを示す。イメージングとスペクトロスコピーに関連する潜在的なアプリケーションをスケッチする。

B. *代替処方*

CS1-CS3
の条件は、我々の結果を得るための唯一の方法ではないことを指摘する。定理9の証明は、実際には次のようなものである。

定理11： $n \times m$ Φ
anmatrixが以下の条件に従うとし、 $\rho_1 > 0$ $\eta_1 < 1/2$ 定数,
および
. $\eta_2 < \infty$

A1:最大 $n \times m$ $\nu(\Phi, J)$ n $m \times m$ 濃度(セクションIV-Bで定義 Φ は n $0 < p \leq 1$ $m_n \sim An^\gamma$ ようになります。

$$\nu(\Phi, J) \leq \eta_1, \quad |J| < \rho_1 n/\log(m_n). \quad (VIII.1)$$

$\eta_1 < n$

A2: $w(\Phi, b_{1,m_n})$ 幅（セクションIIで定義）は $\rho_1 > 0$、

$$w(\Phi, b_{1,m_n}) \leq \eta_2 \cdot (n/\log(m_n))^{-1/2}. \quad (VIII.2)$$

n に従います。

$0 < p \leq 1$ とする．いくつかの $C = C(p, (\eta_i), \rho_1)$ そしてすべての $\theta \in b_{p,m}$ に対して、(P_1) その解は推定値に $\hat{\theta}_{1,n}$ 従います Φ

$$\|\hat{\theta}_{1,n} - \theta\|_2 \leq C \cdot (n/\log(m_n))^{1/2-1/p}.$$

つまり、異なるアプローチでは、ℓ_1 のボールのみに対して良好な幅を持ち、「薄い」集合に対する集中度が低い演算子を示すかもしれません。CS1-CS3の条件が間違いなく異なるアプローチが可能であることを示すもう一つの方法は、[11]、[35]の結果を比較することです。2番目の論文は、異なる手法を使って、最初の論文と一部重複する結果を証明しています。

C. パーシャルフーリエアンサンブル

に当てはまらない最近の2つの記事について簡単に説明します。

n-幅の伝統はここで踏襲されているので、先に引用して目立つようにするのは容易ではありませんでした。

まず、我々の視点に最も近いのは、Candès, Romberg, and Taoの画期的な論文[13]である。これは、セクションIV-Bで議論した。[13]の結果は、ℓ_1 最小化を用いて、ランダムに選ばれた周波数のフーリエ変換から疎な配列を正確に回復できることを示した。配列が n $\rho^* n/\log(n)$ より少ないノンゼロを持つときはいつでも、いくつかの $\rho^* > 0$ についてである。2つ目は、Gilbert *らの論文*[12]で，異なる非線形再構成を用いて，N $m = O(\log(m)\log(M/N))$ $\boldsymbol{R}^m$ のベクトルに対する近似値を復元できることが示されています．この近似値は，周波数領域において，約ランダム M θ だが不均一なサンプルを用いて，ℓ_2 ノルムにおける最良の-項近似とほぼ同じ精度です．ここでは，．ノルムに対する上界がある（ようです）．

つまり、Gel'f と -width について最適に近い部分空間を生成し、ℓ_1 最小化によって そのような情報からすべて再構成できるようにすることです 。

さて、[13]は（実質的には）、 もし、均一測度を持つ部分フーリエアンサンブルにおいて、最大濃度条件A1 (VIII.1) が大きな（適当な定数 1/2 に対して）圧倒的な確率で成立することを証明しています。

).一方、[12]の結果は、ある非一様確率測度を用いてサンプリングした場合、条件A2 (VIII.2) が大きな , に対して圧倒的な確率で部分フーリエアンサンブルに成り立つことを示しているようである。13]と[12]は、部分フーリエ行列の異なるランダムアンサンブルに言及していますが、両者とも興味深い相対的に具体的な演算子族が、コムプレスのセンシングアプリケーション用に開発され得るという考えを補強しています。実際、CandèsはTao[47]と共に得た最近の結果について、一様にサンプリングされた部分フーリエアンサンブルに対して、ポリログ因子のモジュロでA2が成り立つことを知らせてくれた。これは非常に重要な進歩であると思われます。

校正で追加されたノート

この論文が書かれてから数ヶ月の間に、いくつかのグループがここで説明した手法と関連する手法について、合成データや実データを用いた数値実験を行っている。彼らは重要なセンサー問題への適用性を調べ、ノイズの存在下での安定性のような応用問題を研究した。EURASIPジャーナル*Applied Signal Processing*のSpecial Issue on Sparse Representationや、ICASSP 2005のスペシャルセッション、2005年5月にUniversity of Maryland Center for Scien- tific Computing and Applied Mathematicsで開かれたスパース表現に関するワークショップ、2005年11月にUniversité de RennesのSpars05で開かれたワークショップで発表された論文が参考になると思われます。

レフェリーから、Compressed Sensingはデータストリーム処理で生じる問題と似ている点があると指摘された。そこでは、ストリームを保存せずにデータストリームの基本特性（例えば、モーメント、ヒストグラム）を学習したい。要するに、比較的少ない測定回数で、比較的多くの詳細な情報を推論したいのです。大規模データベースにおける「アイスバーグクエリ」やデータストリームにおける「ヘビーヒッター」という概念は、このような文献への入り口を提供することができるかもしれません。

ACKNOWLEDGMENT

2004年春、Emmanuel Candèsは "under-sampled imaging "における部分フーリエアンサンブルの使用に関する彼のアイデアを筆者に伝え、その一部は[13]で発表された;
またプレゼンテーション[14]も参照。さらに最近、Candèsは、上記[47]の結果を筆者に知らせてくれた。これらの会話から得られたインスピレーションに感謝する次第である。また、Anna Gilbertの非適応サンプリングによるB-best Fourier係数の発見[12]について教えてくれたことに感謝し、Gilbertの仕事を明確にする会話をしたEmmanuel Candèsに感謝したい。また、Emmanuel Candèsには、Gilbertの研究を明確にするための会話を提供してもらった。Anna Gilbertは、データストリーム処理に関する文献を紹介してくれた。

参考文献

[1] D.L. Donoho, M. Vetterli, R. A. DeVore, and I. C. Daubechies, "Data compression and harmonic analysis," *IEEE Trans.Inf.Theory*, vol.44, no.6, pp.2435-2476, Oct.1998.

[2] Y.Meyer, *Wavelets and Operators*.Cambridge, U.K.: Cambridge Univ. Press, 1993.

[3] D.L. Donoho, "Unconditional bases are optimal bases for data compression and for statistical estimation," *Appl.Comput.Harmonic Anal.*1, pp.100-115, 1993.

[4] また，このような場合にも，「画像に含まれる疎な成分と最適な原子分解」，*Constructive Approx.*

[5] A.ピンクス、「近似理論における n 幅と最適回復」、*Proc.Symp.応用数学*, 第36巻，C. de Boor, Ed., Providence, RI, 1986, pp.51-66。

[6] J.F. Traub and H. Woziakowski, *A General Theory of Optimal Algo - rithms*.New York:ニューヨーク: アカデミック, 1980.

[7] D.L. Donoho, "Unconditional bases and bit-level compression," *Appl.Comput.Harmonic Anal.*3, pp.388-92, 1996.

[8] A.ピンクス、*近似理論における*n - 幅 。 New York:Springer-Verlag, 1985.

[9] A.Y. Garnaev and E. D. Gluskin, "On widths of the Euclidean ball" (in English), *Sov.Math.-Dokl.30*巻、200-203頁、1984年。

[10] B.S. Kashin, "Diameters of certain finite-dimensional sets in classes of smooth functions," *Izv.Akad.Nauk SSSR, Ser.Mat.また，"Diameters in certain finite-dimensional sets in classes smooth functions "と題する論文を発表した。*

[11] D.L. Donoho, "For most large underdetermined systems of linear equations, the minimal ℓ_1-norm solution is also the sparsest solution," *Commun.純粋応用数学*.を出版予定。

[12] A.C. Gilbert, S. Guha, P. Indyk, S. Muthukrishnan, and M. Strauss, "Near-optimal sparse fourier representations via sampling," in *Proc 34th ACM Symp.また、このような場合にも、「逐次処理」であることが望ましい。*

[13] E.J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles:このような場合、「不完備な周波数情報からの正確な信号再構成」、「*IEEE Trans.Inf.Theory*, 掲載予定。

[14] E.J. Candès, "Robust Uncertainty Principles and Signal Recovery," presented at 2nd Int.Conf.Computational Harmonic Anaysis, Nashville, TN, May 2004.

[15] C.A. Micchelli and T. J. Rivlin, "A survey of optimal recovery," in *Op- timal Estimation in Approximation Theory*, C. A. Micchelli and T. J. Rivlin, Eds:Plenum, 1977, pp.1-54.

[16] J.Peetre and G. Sparr, "Interpolation of normed abelian groups," *Ann.Math.4, vol. 92, pp.217-262, 1972*.4, vol.92, pp.217-262, 1972.

[17] J.Bergh and J. Löfström, *Interpolation Spaces. 入門*.ドイツ、ベルリン。Springer-Verlag, 1976.

[18] B.カール, "エントロピー数s数, 固有値問題", *J. Funct.Anal.*41巻、290-306頁、1981年。

[19] G. Pisier, *The Volume of Convex Bodies and Banach Space Geometry*.英国ケンブリッジ大学出版局, 1989.

[20] C.また, このような場合にも, 「曖昧さ」を解消するために, *「曖昧さ」を解消するために, 「曖昧さ」を解消するために, 「曖昧さ」を解消するために, 「曖昧さ」を解消するために, 「曖昧さ」を解消するために, 「曖昧さ」を解消するために, 「曖昧さ」を解消する必要があります。*

[21] T.を用いた、エントロピー数に対する下限推定値、*J. Approx.*

[22] S.S. Gal and C. Micchelli, "Optimal sequential and nonsequential proce- dures for evaluating a functional," *Appl.Anal.*10巻、105-120頁、1980年。

[23] A.このような場合、「曖昧なデータから線形演算子を最適に推定する」、*SIAM J. Numer.Anal.*16巻、87-105頁、1979年。

[24] M.A. Kon and E. Novak, "The adaption problem for approximating linear operators," *Bull.アメリカ数学。Soc.* , vol.23, pp.159-165, 1990.

[25] E.ノバック、「適応力について」、*J.Complexity*, vol.12, pp.199-237, 1996.

[26] S.陳、D.L.ドノホー、M.A.サンダース、「基底追求による原子分解」、*SIAM J. Sci Comp.*20巻1号、33-61頁、1999年。

[27] D.L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Trans.Inf.Theory*, vol.47, no.7, pp.2845-62, Nov. 2001.

[28] M.Elad and A. M. Bruckstein, "A generalized uncertainty principle and sparse representations in pairs of bases," *IEEE Trans.Inf.Theory*, vol.49, no.9, pp.2558-2567, Sep.2002.

[29] D.L. Donoho and M. Elad, "Optimally sparse representation from over- complete dictionaries via ℓ_1 norm minimization," *Proc.Natl. Acad.米国*、vol.100、no.5, pp.2197-2002, Mar.

[30] R.Gribonval, M. Nielsen, "Sparse representations in unions of bases," *IEEE Trans Inf Theory*, vol.49, no.12, pp.3320-3325, Dec.2003.

[31] J.J. Fuchs, "On sparse representation in arbitrary redundant bases," （任意の冗長ベースにおけるスパース表現について）。*IEEE Trans.Inf.Theory*, vol.50, no.6, pp.1341-1344, Jun.2002.

[32] J.A. Tropp, "Greed is good:スパース近似のアルゴリズム結果」、*IEEE Trans Inf.Theory*, vol.50, no.10, pp.2231-2242, Oct.2004.

[33] ——, "Just relax:ノイズの中でスパースな信号を識別するための凸計画法」*IEEE Trans Inf.Theory*, vol.52, no.3, pp.1030-1051, Mar.

[34] D.L. Donoho, M. Elad, and V. Temlyakov, "Stable recovery of sparse overcomplete representations in presence of noise," *IEEE Trans.Inf.Theory*, vol.52, no.1, pp.6-18, Jan.2006.

[35] D.L. Donoho, "For most underdetermined systems of linear equations, the minimal ℓ_1-norm near-solution approximates the sparsest near-solution," *Commun.を発表した。*を出版予定。

[36] I.C. Daubechies, *Ten Lectures on Wavelets (ウェーブレットに関する10の講義)*.Philadelphia, PA:SIAM, 1992.

[37] S.マラト、*信号処理のウェーブレットツアー*。サンディエゴ、カリフォルニア州。Aca- demic, 1998.

[38] A.Cohen, R. DeVore, P. Petrushev, and H. Xu, "Nonlinear approximation and the space ," *Amer.J. Math.*121巻、587-628頁、1999年。

[39] E.J. Candès and D. L. Donoho, "Curvelets-A surprisingly effective nonadaptive representation for objects with edges," in *Curves and Surfaces*, C. Rabut, A. Cohen, and L. L. Schumaker, Eds.Nashville, TN:Vanderbilt Univ. Press, 2000.

[40] ——また, "New tight frames of curvelets and optimal representations of ob-ject with piecewise C^2 singularities"（カーブレッツの新しいタイトフレームと区分的C特異点の最適表現）, *Comm.を用いた。そのため，このような問題点を解決するために必要な基礎的な知識を得ることができる。*

[41] S.J. Szarek, "Spaces with large distances to ℓ_∞^n and random matrices," （ランダム行列と大きな距離を持つ空間）. *アメリカJ. Math.112*巻、819-842頁、1990年。

[42] ——また, 「ランダム行列の条件数」, 「*複雑性*」, 第7巻, 131-149頁, 1991年.

[43] A.Dvoretsky, "Some results on convex bodies and banach spaces," in Japan. *シンポジウム線形空間*, エルサレム, イスラエル, 1961, pp.123-160.

[44] T.Figiel, J. Lindenstrauss, and V. D. Milman, "The dimension of almost-spherical sections of convex bodies," *Acta Math.という論文*。

[45] V.D. Milman and G. Schechtman, *Asymptotic Theory of Finite-Dimensional Normed Spaces (Leecture Notes in Mathematics)*.Berlin, Germany:Springer-Verlag, 1986, vol.1200.

[46] D.Pollard, *Empirical Processes:理論と応用*.ヘイワード、カリフォルニア州：Inst.Math .Statist., vol.2, NSF-CBMS Regional Conference Series in Probability and Statistics.

[47] E.J. Candès and T. Tao, "Near-optimal signal recovery from random project:また，このような場合にも，そのような問題を解決することができる。を用いた。Rep., 2004.