

情報理論に現れるランダム行列理論： 圧縮センシングとは？



竹田 晃人

茨城大学大学院理工学研究科

近年、情報科学の分野では情報の疎性を利用した情報処理技術が盛んに研究されている。それは圧縮センシングと呼ばれる手法が10年程前に考案され、情報科学の諸分野に大きいインパクトを与え、現代社会で重要視されているデータ科学の分野、具体的には医療画像技術や天体観測等で今後大きな役割を果たすことが期待されているからである。この圧縮センシングの理論は近年の研究で物理学的解析手法と関係することが分かってきた。そこで本稿では特にランダム行列理論との関連性の観点から圧縮センシングを紹介する。

圧縮センシングとは信号（又はデータ）に内在する疎性を利用し非常に少数の観測から高次元信号を復元可能とする手法である。その問題設定は極めて簡潔な数式で表現され、一言では「線型連立方程式を条件不足の下で如何に解くか」である。具体的には方程式の未知変数の数を原信号次元、方程式の本数を信号観測回数と考え、「観測回数 < 原信号次元」と仮定した上で原信号を復元する問題を考える。しかしこの設定のみでは当然解は不定で、問題は不良設定問題となる。そこでこの問題を適切に定義する為に、原信号が「疎」即ち零成分を多く持つ（疎な解）と仮定し、そのような解の求解を目標とする。

従ってこの問題では疎な解を効率良く求めることが重要だが、その為に考案されたのが l_1 ノルム最小化に基づく求解法である。この求解法は線型計画問題として表現されるため、線型計画アルゴリズムを用いれば原信号に相当する疎な解を実用上問題無い計算量で得ることが出来る。

ところでこの問題では原信号を復元する際に必要となる観測数が少ないほど応用上有利となるが、信号を完全復元するために必要な観測回数＝連立方程式の本数の下限を理論的に評価する方法が幾つか考案された。

まず「制限等長性」の概念を用いた評価法があり、 l_1 ノルム最小化で疎な解の求解が成功する十分条件はこの制限等長性を用い表現出来る。この概念とランダム行列理論（正確には最大最小固有値に関する確率不等式）を組み合わせることで、信号完全復元の為の観測数に関する条件が得られる。

それとは別に幾何学を用いた評価法がある。この解析法は l_1 ノルム最小化が線型計画問題として表現出来ることに着目し、問題を高次元の線型な幾何学の問題に焼き直して、幾何学的解釈、より正確には射影の理論から観測数条件を求めようとするものである。

これらとは独立に統計力学に基づく評価法も考案された。これは l_1 ノルム最小化を統計力学の問題、具体的にはスピングラス模型の基底状態探索問題に置き直した上で、観測数条件を相図上の相転移線から求める方法である。これは驚くべきことに幾何学による結果と全く同じ結果を導く。さらにこの解析に自由確率論に基づくランダム行列理論の解析法を応用することで観測数条件のuniversalityを議論することが出来る。

以上のように圧縮センシングとランダム行列理論は様々な解析手法を通じ密接に関係しているのである。

—Keywords—

ランダム行列理論：

ランダム行列、直観的には各行列要素が何らかの確率分布に従い生成されている行列（より正確には行列値を持つランダム変数）を扱う理論。物理学では原子核の複雑なエネルギースペクトルの理解の為にWignerにより最初に注目された。現在では物性物理学・統計力学・量子色力学・量子カオス・数論・生態学・金融工学・複雑ネットワーク等の幅広い分野に応用されている。本稿で触れる情報理論（通信工学も含んだ意味）も応用分野の一つである。

線型計画問題：

最適化問題（等式/不等式で表される制約条件の下で目的関数を最大/最小化する問題）のうち、目的関数及び全ての制約条件が線型関数で書き表される問題。この問題には単体法（シンプレックス法）・楕円体法・内点法等の効率的な解法が知られている。

制限等長性：

圧縮センシングを記述する線型連立方程式に現れる長方形行列が等長変換行列からどの程度ずれているかを定義する概念。CandèsとTaoにより導入された。「制限」とは線型連立方程式の解となるベクトルが「疎」、つまり多くの成分が0でなければならない制限を指す。制限等長性を用いることで圧縮センシングの問題で信号の完全復元が可能である十分条件を議論出来る。