

# 感受率行列の固有状態

統数研 伊庭 幸人  
東大物性研 福島 孝治

## 1 はじめに

感受率あるいは帯磁率行列の固有モード解析は、ランダム系の複雑な秩序を知る手段として、古くから調べられてきた。この解析では、秩序状態は感受率に大きな寄与があるモードと関係があると考えられるわけである。特に、秩序状態が陽に分からないような複雑な系においては、その手法は有力である。古くは「ボーズグラス」描像の解析 [1] に用いられ、また、ランダム系の Griffiths 相の性質は感受率行列の固有値分布の特異性に関連して議論されてきた [2]。最近では、系のサイズに比例する固有値をもつ (extensive) 固有モードの数と純状態の数の関係が議論され、スピングラス状態の解析に新しい視点を与えている [3]。

この解析を数値的に行なうには、モンテカルロ計算により感受率行列を求め、それを対角化することになる。このような感受率行列の数値的研究は、根本氏により2次元スピングラス模型の Griffiths 相に関して行なわれており、前回の基研研究会でもその報告がある [4]。また、Sinovaらにより、スピングラス [3] やランダム磁場問題 [5] に応用されている。しかしながら、主に2つの困難点のために、その実行は系の大きさとともに深刻になる。それらは、複雑な系において興味ある低温でモンテカルロ計算を行なうことと密な行列である感受率行列の対角化に要するメモリーと計算時間の増大である。前者は、拡張アンサンブル法によるモンテカルロ法の進展により、ある程度回避できるが、後者の問題が実行出来る系の大きさの制限となっているようである。実際に以前の計算では、系の最大スピン数  $N$  は、784 [4]、1296 [3] であり、最近の標準的なモンテカルロ計算のサイズよりも小さくとどまっている。

我々は後者の問題点を克服しうる新しい方法を考案し、幾つかの模型でテストを行なった。我々の方法は、統計学ではよく使われている主成分分析のテクニックであり、系のサイズの問題は大幅に改善している。本解説では、我々の考案した2つの方法を説明し、スピングラス系への応用例を紹介する。

## 2 従来の方法

この節では、従来の感受率行列の固有モード解析の方法を説明し、その問題点を明らかにする。具体的な模型として、 $N$  個のイジングスピン系を考え、モンテカルロ・シミュレーションにより、 $M$  個のスピン状態を持っているとする。 $\mu$  をシミュレーションによるデー

タの指標として、サイト  $i$  のスピン変数の平均値は、

$$\bar{S}_i = \frac{1}{M} \sum_{\mu} S_i^{\mu}, \quad (1)$$

となる。  $X_i^{\mu}$  をサイト  $i$  の  $\mu$  番目のデータのスピン変数の平均値からのずれとする。この変数  $X$  は  $N \times M$  の行列とみなせる (データ行列)。それを用いて、感受率行列は以下のように定義される。

$$C_{ij} = \frac{1}{M} \sum_{\mu} (S_i^{\mu} - \bar{S}_i)(S_j^{\mu} - \bar{S}_j) = \frac{1}{M} \sum_{\mu} X_i^{\mu} X_j^{\mu}. \quad (2)$$

この行列は帯磁率行列とも呼ばれるが、統計学では分散共分散行列と呼ばれるデータ解析での基本統計量である。固有モード解析は、この行列の固有値問題

$$\sum_j^N C_{ij} e_j = \lambda e_i, \quad (3)$$

を解くことである。一般に、この感受率行列は、  $N \times N$  の密行列になるために、対角化には  $N^2$  のメモリーと  $N^3$  の計算量が必要になる。例えば、Sinova らは [3]、4次元スピングラス模型に対して、  $N \leq 6^4 = 1296$  まで計算している。最近のモンテカルロ計算の平衡化が問題にならない規模は、  $N \leq 10^4$  程度が標準的であるので、少しサイズが小さいように感じられる。サイズ  $N = 1296$  の感受率行列を倍精度保存するためのメモリーは約 13MB 程度であり、現在のコンピュータのスペックからすれば大きくないので、律速になっているのは対角化の  $N^3$  の手続きだと想像される。特に、スピングラス系では、ランダム平均に数百から数千のサンプルを計算する必要があり、その手間も大変になっているであろう。

### 3 改良法

前節の問題を解決するために、2種類の方法を提案する。これまでの議論から、感受率行列を計算、保存した時点で、  $N^2$  のメモリーと  $N^3$  の計算量は避けられない。そこで、感受率行列の計算する時点でトリックを考える。

#### 3.1 双対法 (Dual trick)

一つの方法は、双対法と呼ばれる主成分分析の典型的な手法で、「データ数  $M$  がスピン数  $N$  よりも少ない場合に効率の良い方法」である。双対法では、シミュレーションで得られた  $M$  個のデータ間のハミング距離行列  $D^{\mu\nu}$

$$D^{\mu\nu} = \frac{1}{N} \sum_i (S_i^{\mu} - \bar{S}_i)(S_i^{\nu} - \bar{S}_i), \quad (4)$$

の固有値問題を解く。ここで行列  $D$  は  $M \times M$  の実対称行列である。このハミング距離行列  $D$  の固有値問題と感受率行列の固有値問題は双対の関係にあることが簡単に示され

る。例えば、ハミング距離行列の固有ベクトルを  $e^\mu$ , ( $\mu = 1, \dots, M$ ) とすると、感受率行列の固有ベクトル  $e_i$  は、 $e_i = \sum_{\mu=1}^M X_i^\mu e^\mu$  と表される。すなわち、どちらかの固有値・固有ベクトルから、反対の固有値・固有ベクトルがわかるというわけである。もしもデータ数が  $O(N)$  よりも少なく済む場合は、ハミング距離行列を対角化の方が効率がよいと言える。その際に、計算量の律速は  $e^\mu$  から元の問題の固有ベクトル  $e_i$  を構成する部分で、 $O(N)$  になり、メモリー量もまたデータ行列  $X_i^\mu$  の保存のために  $O(N)$  となる。

この方法で問題となるのは、どのくらい  $M$  を小さく出来るか、あるいは、どの程度のデータ数で感受率を精度良く評価できるかである。それは調べる系の性質に依存しているので、一般的な指針はない。それぞれの系でチェックしていくしかない。必要な  $M$  が  $O(N)$  になってしまう場合は、双対法は効率が良いとは言えない。実際に比較的小さな系では固有値の  $M$  依存性が見えなくなるには  $O(N)$  のデータ数が必要である。しかし、経験的にモンテカルロ法で物理量の平均値を評価する際に  $O(N)$  のサンプル点もとらなくても、ある程度の精度は確保できることは知っているの、大きなサイズでは有効になるはずである。

### 3.2 オンライン法 (On-line trick)

もう一つの方法は、シミュレーションをしながら固有モード解析を行なう方法である。主成分分析の観点から、固有モード解析はある指定した個数  $N'$  の係数  $a^\mu$  と正規直交ベクトル  $e_i^l$ , ( $l = 1, \dots, N'$ ) について、

$$G = \sum_{\mu}^M \sum_i^N \left( X_i^\mu - \sum_l^{N'} a_l^\mu e_i^l \right) \quad (5)$$

を最小にするように求める問題とみなせる。このようなベクトル  $e_i^l$  は、未定乗数を  $\Lambda_l, \Lambda_{ll}$  として、

$$\sum_j C_{ij} e_j^l - \Lambda_l e_i^l - \sum_{l'} \Lambda_{ll'} e_i^{l'} = 0 \quad (6)$$

を満たす必要がある。これを満たし、 $G$  が最小の解を得るには、行列  $C_{ij}$  の固有値問題を解いて、最初の  $N'$  個の固有ベクトルで張られる空間内の任意の直交基底を  $e_i^l$  とすればよい。この式 (6) を逐次方程式

$$\sum_j C_{ij} e_j^l(t) - \Lambda_l e_i^l(t+1) - \sum_{l'} \Lambda_{ll'} e_i^{l'}(t+1) = 0 \quad (7)$$

とみなして、 $t$  回目の近似固有ベクトル  $e_j^l(t)$  から逐次的に  $t+1$  回目の固有ベクトル  $e_j^l(t+1)$  を解く。このままでは、 $N \times N$  の行列  $C_{ij}$  を保存する必要があるように見えるが、 $C$  の定義 (2) に戻れば、式 (7) の第一項は、

$$\sum_{j=1}^N C_{ij} e_j^l(t) = \sum_{j=1}^N \frac{1}{M} \sum_{\mu} X_i^\mu X_j^\mu e_j^l(t) = \frac{1}{M} \sum_{\mu=1}^M X_i^\mu \left( \sum_j X_j^\mu e_j^l(t) \right) (\equiv f_i^l(t)) \quad (8)$$

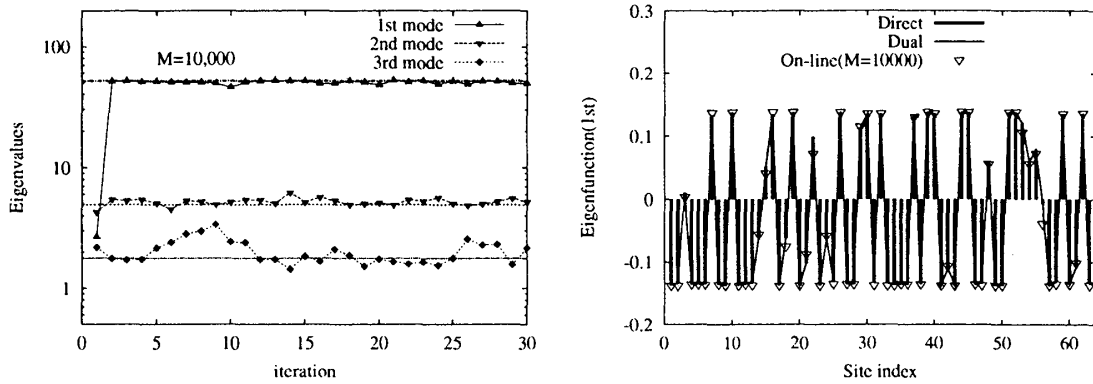


図 1: (左)3次元スピングラスへの応用. オンライン法で求めた上位3モードの固有値の繰り返し数依存性. 上から第一, 第二, 第三固有値. 直線は双対法で求めた固有値.

図 2: (右)3次元スピングラスへの応用. 直接(棒グラフ)・双対(折れ線グラフ)・オンライン法(△マーク)による第一固有ベクトルの結果の比較.

と表されるので, シミュレーション中にここまで計算してしまうと,  $N \times N'$  の行列  $f_i^l(t)$  を保存すればよいことになる. その後式 (7) を解いて, 次の近似固有ベクトルを求め, 収束するまで繰り返す. 計算する固有ベクトルの個数  $N'$  は我々の興味に合わせて設定すればよい. 少し説明が複雑になったので, 実際の手続きをまとめてみる.

1. 初期ベクトル  $e_i^l(t=0) (l=1, \dots, N')$  の設定 (例えばランダム)
2. シミュレーションで状態  $X_i^\mu (\mu=1, \dots, M)$  の生成しながら, 各近似固有ベクトルへの射影  $f_i^l$  の計算
3.  $f$  を特異値分解して, 近似固有ベクトル  $e_i^l(t+1)$  を得る
4. 2へ戻る

この方法では,  $N'$  を  $O(N^0)$  とすると, メモリーと計算量ともに  $O(N)$  である. 一方で何回もシミュレーションを繰り返す必要があることが欠点である. また, 双対法と異なり, シミュレーションデータの保存が必要ないために, 上記の2.において, シミュレーションで生成する状態数  $M$  はいくらでも大きくでき, また興味ある固有モードの個数  $N'$  を比較的自由に選択できるので効率的である.

#### 4 スピングラス系への応用

提案した方法の応用例として, 3次元イジングスピングラス模型,

$$H(\{S_i\}) = \sum_{\langle ij \rangle} J_{ij} S_i S_j \quad (9)$$

を対象とする. ここで,  $J_{ij}$  はランダム変数で,  $\pm 1$  を等確率でとるとする. スピン数が  $N = 4^3$  のこの模型に対して, 交換法を用いたモンテカルロ・シミュレーションによりスピン配位を生成し, 直接法, 双対法, オンライン法を用いて, 固有モード分解を試みた.

まず、オンライン法では、一回のシミュレーションの数を  $M = 10,000$  とし、繰り返すによる固有ベクトル等の収束を調べた (図1)。計算する固有モードの個数は  $N' = 10$  とした。図1には、同じ相互作用のモデルに対して、サンプル間隔を100ステップとし、 $M = 200$  の双対法を行なった結果を直線で示している。オンライン法で数回繰り返しを行なうとほぼ双対法の結果と一致していることがわかる。独立に計算した直接法の結果も含めて、固有ベクトルも3つの方法でよく一致していることがわかる2。

次に少し大きな系のデモンストレーションとして、2次元イジングスピングラス模型の固有モードを試みた。温度が、 $T/J = 0.4$  の最大固有モードの振幅の実空間上での様子を図3に示した。図の濃い部分がほぼ凍結しているスピンを表しており、大きなクラスターを形成していることがわかる。また、そうしたクラスターの中にも小さな薄い部分が多く紛れ込んでいる。これらは、フラストレーションの効果により、凍結できずに、フラフラ動いているスピンを表している。

ここでのサイズは以前の計算よりも数倍大きい程度であるが、秩序形成を図のような固有モードの局在非局在と関連づけて議論する際には必然的に大きなサイズの計算が求められる。

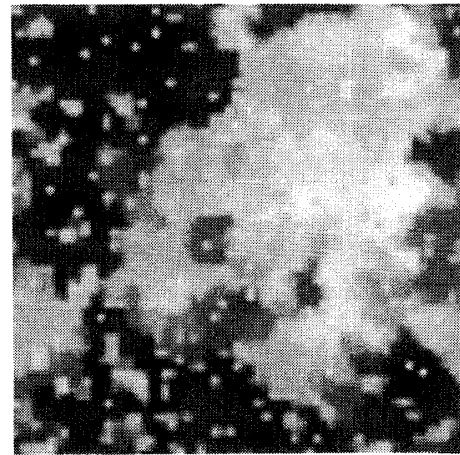


図3: 2次元スピングラスの例。  $N = 64 \times 64$  の系の第一固有モードの成分の大きさの実空間上への濃淡プロット。

我々の提案する方法はそのような研究への発展に有用であると考えている。また、ランダム系の秩序の様子を視覚的に捉えることは、理論の発展にも役立つであろう。

## 5 まとめとこれから

我々は感受率行列の固有モード解析に対する2種類の効率の良い方法を提案した。また、それらをスピングラス系へ応用し、実際に機能することを示した。この固有モード解析は統計学の主成分分析と呼ばれるものと同じであり、我々の方法は統計学での知見を用いたり、発展させたものである。我々はスピン系への応用しか試みていないが、方法は非常に一般的であり、様々な応用が期待される。特にオンライン法は適用範囲が広いと考えられるが、ここでは単に使えることを示すだけで、実用上はもう少し検討が必要な部分もある。例えば、初期ベクトルの設定や計算モード数  $N'$  のとり方や収束判定条件等である。

本解説では、感受率行列の数値解析上の困難の解消という観点で論じてきた。実際に、ここでの双対法を用いて、短距離型スピングラス模型の大規模数値計算を行ない、これま

での調べられたスピンの約 10 倍大きなサイズを調べることが可能になり、新しい結果を得ている [6].

一方で、主成分分析の観点からは、この解析はシミュレーションによって得られたデータの数値解析と見ることが出来る。このような主成分分析を用いたシミュレーション結果の解析は、タンパク質の分野ではすでに研究されているが、最近ではスピングラス分野でもクラスター解析や主成分分析等の研究が始まりつつある。

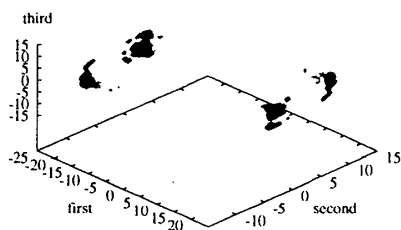


図 4: ある相互作用配置の 3 次元イジングスピングラス模型のシミュレーションデータの主軸への射影プロット。データが幾つかの「かたまり」へ分類される様子がわかる。

例えば、一つの試みは、図 4 に示すように、シミュレーションから得られたデータを主成分分析し、各データの主軸への射影プロットを見ることである。データであるスピン配位を直接目で見て、そこから秩序の様子を理解することはランダム系では難しいが、射影プロットにより、それらのデータは幾つかのグループ分けが可能のように見える。温度依存性等の観察から、第一軸は時間反転対称性を表していることがわかる。第一軸で大きく 2 つに分かれたデータは、さらに第二軸についても 2 つにわかれているように見える。これは階層的な連続相転移を連想させるが、これはある一つのサンプルでの事象にすぎない。“ランダム平均”をこのような解析にどのように取り込むかはスピングラス研究には重要な課題である。

また、グループ分けされたデータ (スピン配位) は励起状態の様子を反映していると想像され、実空間上で見ることでその形状についての情報が得られるであろう。今後は、いわゆるドロップレット励起の直接観察につながるものと考えられ、さらなる発展が望まれる。

## 参考文献

- [1] J. A. Hertz, L. Fleishman and P. W. Anderson, Phys. Rev. Lett. 43, (1979) 942.
- [2] A. J. Bray and M. A. Moore, J. Phys. C15 (1982) L765., A. J. Bray, Phys. Rev. Lett. 59 (1987) 586.
- [3] J. Sinova et al, Phys. Rev. Lett. 85, (2000) 2609, J. Sinova et al, Phys. Rev. B85, (2000) 104427.
- [4] 根本 幸児, 山田 定徳, 物性研究 74-2 (2000) 122.
- [5] J. Sinova and G. Canright cond-mat/0103071.
- [6] K. Hukushima and Y. Iba, in preparation.