クレジット:
Mathematics and Informatics Center 計算機実験I 2020 藤堂眞治

計算機実験 I (第 7 回)

藤堂眞治

2020/06/10

1 行列の対角化
2 疎行列に対する反復法
3 逆反復法
4 変分法

講義予定

- 2020-04-22: 第 1 回 環境整備
- 2020-05-07: 第 2 回 計算機実験の基礎
- 2020-05-13: 第 3 回 常微分方程式の解法
- 2020-05-20: 第 4 回 行列演算とライブラリ
- 2020-05-27: 第 5 回 連立一次方程式の解法
- 2020-06-03: 第 6 回 行列の対角化
- 2020-06-10: 第 7 回 疎行列に対する反復解法
- 2020-06-17: 第 8 回 特異値分解と最小二乗法

物理の問題にあらわれる行列演算

- 連立一次方程式・逆行列
  - 偏微分方程式の境界値問題
  - 非線形連立方程式に対するニュートン法
- 対角化・特異値分解
  - 固有値問題・行列関数
  - 最小二乗近似
- 計算機は大規模行列演算が得意
  - 直接法: $\sim 10^4$ 次元
  - 疎行列に対する反復解法: $\sim 10^9$ 次元
- 行列演算についてはライブラリがよく整備されている
  - それぞれの原理とその特徴を理解して正しく使うことが重要
  - 適切なライブラリを使うことで数十倍あるいはそれ以上速くなることも

行列の数値対角化

- 一般的に次元が 5 以上の行列の固有値は、あらかじめ定まる有限回の手続きでは求まらない
  - 必ず何らかの反復法 (+収束判定) が必要となる
- 密行列向きの方法
  - Jacobi 法
  - Givens 変換・Householder 法 (三重対角化) + QR 法など
- 疎行列向きの方法
  - べき乗法
  - Lanczos 法 (三重対角化) + QR 法など
- 固有ベクトル
  - QR 法で求めたものを逆変換
  - 逆反復法で精度改善

反復法

- 疎行列の場合、行列ベクトル積は高速に行える
- Givens 変換、Householder 変換などを行うと疎行列性が失われる
- 行列ベクトル積のみを用いる反復法が効果的
  - べき乗法
  - Lanczos 法

固体物理・量子統計物理の多体問題に現れる行列の例

- 強束縛近似 (tight-binding approx.) のもとでの第二量子化表示

$$H = -t \sum_{\langle i,j \rangle \sigma} (c^\dagger_{i,\sigma} c_{j,\sigma} + h.c.) + (\text{相互作用})$$

- 局在スピン模型 (ハイゼンベルグ模型)

$$H = -J \sum_{\langle i,j \rangle} S_i \cdot S_j = -J \sum_{\langle i,j \rangle} [S^z_i S^z_j + \frac{1}{2}(S^+_i S^-_j + S^-_i S^+_j)]$$

- 格子点の数を N とすると、ハミルトニアンはそれぞれ $4^N \times 4^N$、$2^N \times 2^N$ の (疎) 行列で表される。
- N が大きくなると、行列の次元は指数関数的に増加
- 量子多体系に共通する困難

べき乗法 (Power Method)

- 適当なベクトル v_1 から出発する
- v_1 が最大固有ベクトル ξ_1 と直交していないとすると

$$v_1 = c_1\xi_1 + c_2\xi_2 + c_3\xi_3 + \cdots + c_n\xi_n$$

と展開できる ($c_1 \neq 0$)。この両辺に A を次々と掛けていくと

$$\begin{aligned}
v_2 &= Av_1 = c_1\lambda_1\xi_1 + c_2\lambda_2\xi_2 + c_3\lambda_3\xi_3 + \cdots + c_n\lambda_n\xi_n \\
v_3 &= A^2v_1 = c_1\lambda_1^2\xi_1 + c_2\lambda_2^2\xi_2 + c_3\lambda_3^2\xi_3 + \cdots + c_n\lambda_n^2\xi_n \\
&\vdots \\
v_{k+1} &= A^kv_1 = c_1\lambda_1^k\xi_1 + c_2\lambda_2^k\xi_2 + c_3\lambda_3^k\xi_3 + \cdots + c_n\lambda_n^k\xi_n \\
&= c_1\lambda_1^k\Bigl[\xi_1 + \sum_{i=2}^{n}\frac{c_i}{c_1}\Bigl(\frac{\lambda_i}{\lambda_1}\Bigr)^k\xi_i\Bigr] \approx c_1\lambda_1^k\xi_1
\end{aligned}$$

べき乗法の収束

- べき乗法による固有値

$$\frac{v_{k+1}^T v_{k+1}}{v_{k+1}^T v_k} = \lambda_1 + O\Big(\big(\frac{\lambda_2}{\lambda_1}\big)^{2k}\Big)$$

- 誤差の収束

$$\frac{v_{k+1}^T v_{k+1}}{v_{k+1}^T v_k} \approx \lambda_1 + e^{-2k\ln(\lambda_1/\lambda_2)}$$

- $1/\ln(\lambda_1/\lambda_2)$ 程度の反復が必要
- λ_1 と λ_2 が近い場合には、反復回数が非常に多くなる

第 2 固有値・第 3 固有値 $\cdots$

- 第 1 固有ベクトル ξ_1 の成分を行列から差し引く (減次)

$$A_1 = A - \lambda_1 \xi_1 \xi_1^T$$

  この行列は、固有値 $0, \lambda_2, \lambda_3, \cdots, \lambda_n$ を持つ
- 行列 A_1 に対してべき乗法を使うと、第 2 固有値 λ_2 と対応する固有ベクトル ξ_2 が得られる
- 第 k 固有値まで求まっている場合

$$A_k = A - \sum_{i=1}^{k} \lambda_i \xi_i \xi_i^T$$

- 実際には数値誤差のため、ベクトルの直交性は厳密ではない
- 大きい方から数個程度を求めるのが限界

Rayleigh-Ritz の方法

- $n \times n$ 行列 A について、互いに正規直交するベクトル $v_1, v_2, \cdots, v_m$ $(m < n)$ が張る部分空間の中で「最良の」固有ベクトルを求めたい
- $n \times m$ 行列

$$V = \left[v_1 v_2 \cdots v_m \right]$$

を定義すると、$V^T V = E_m$ が成り立つ (ただし $VV^T \neq E_n$)

- 部分空間内のベクトルを $w = \sum_i a_i v_i$ と表すと、$\dfrac{w^T A w}{w^T w}$ が極大値を取る (本当の固有ベクトルにできるだけ平行になる) 条件は、

$$\frac{\partial}{\partial a_i} \frac{w^T A w}{w^T w} \sim \sum_j H_{ij} a_j - \lambda a_i = 0$$

Rayleigh-Ritz の方法

- $m \times m$ 行列
$$H = V^T A V$$
に対する固有値問題 $Ha = \lambda a$
- λ: もとの行列の近似固有値 (Ritz 値)
- Va: もとの行列の近似固有ベクトル (Ritz ベクトル)
- 最大固有値に対する近似固有値が欲しい場合、最大固有ベクトルになるべく近い (しかし互いに直交する) ベクトル $v_1, v_2, \cdots, v_m$ を選べばよい

Lanczos 法

- 初期 (ランダム) ベクトル v_1 に加えて

$$Av_1, Av_1, \cdots A^{m-1}v_1$$

を正規直交化して $v_1, v_2, \cdots, v_m$ を作る (Krylov 部分空間)

- 部分空間での Ritz 値を固有値の近似値とする
- $A^k v_1$ はどんどん最大固有ベクトルに近づいていくので、$m \ll n$ でも良い近似固有値が得られると期待される

Lanczos 法

- 正規化された初期 (ランダム) ベクトル v_1 から出発する
- $v_2, v_3, \cdots$ を生成する

$$\begin{aligned} v_2 &= (Av_1 - \alpha_1 v_1)/\beta_1 \\ v_3 &= (Av_2 - \beta_1 v_1 - \alpha_2 v_2)/\beta_2 \\ &\vdots \\ v_{m+1} &= (Av_m - \beta_{m-1} v_{m-1} - \alpha_m v_m)/\beta_m \end{aligned}$$

ここで、α_i と β_i は以下のように選ぶ

$$\begin{aligned} \alpha_i &= v_i^T A v_i \\ \beta_i &= ||Av_i - \beta_{i-1} v_{i-1} - \alpha_i v_i||, \ \beta_0 = 0 \end{aligned}$$

Lanczos 法

- $v_1, v_2, v_3, \cdots, v_{m+1}$ は正規直交
- 漸化式を書き換えると

$$
\begin{aligned}
Av_1 &= \alpha_1 v_1 + \beta_1 v_2 \\
Av_2 &= \beta_1 v_1 + \alpha_2 v_2 + \beta_2 v_3 \\
Av_3 &= \beta_2 v_2 + \alpha_3 v_3 + \beta_3 v_4 \\
&\vdots \\
Av_m &= \beta_{m-1} v_{m-1} + \alpha_m v_m + \beta_m v_{m+1}
\end{aligned}
$$

Lanczos 法

- 行列で表現すると

$$
A\left[v_1 v_2 \cdots v_m\right] = \left[v_1 v_2 \cdots v_m v_{m+1}\right] \begin{bmatrix} \alpha_1 & \beta_1 & & & & \\ \beta_1 & \alpha_2 & \beta_2 & & & \\ & \beta_2 & \alpha_3 & \beta_3 & & \\ & & \beta_3 & \alpha_4 & \beta_4 & \\ & & & \ddots & \ddots & \ddots \\ & & & & \beta_{m-1} & \alpha_m \\ & & & & & \beta_m \end{bmatrix}
$$

両辺に左から $\left[v_1 v_2 \cdots v_m\right]^T$ をかけると

$$
\left[v_1 v_2 \cdots v_m\right]^T A \left[v_1 v_2 \cdots v_m\right]
$$

は 3 重対角行列となることがわかる

Lanczos 法

- 原理的には、n ステップ目で $\beta_n = 0$ となり、3 重対角化が完了する
- 実際には、数値誤差のため $v_1, v_2, v_3 \cdots$ の直交性が崩れていく
  - m を大きくしすぎると、おかしな固有値が出てくる
  - 多くの固有値・固有ベクトルが欲しい場合には Householder 法を使うべき
- Lanczos 法では、大きな固有値に対応する固有ベクトルにできるだけ近いものから部分空間を作っていく
  - 100 万次元以上の行列の場合でも $m = 100 \sim 200$ 程度で最初の数個の固有値は精度良く求まる
- 必要な操作は、行列とベクトルの積、ベクトルの内積・スケーリング・和のみ
  - 疎行列の場合、非常に効率が良い

逆反復法による固有ベクトルの精度向上

- 逆反復法
  - 近似固有値を μ とするとき、行列 $(A-\mu E)^{-1}$ を考えると、固有ベクトルは A と同じ、固有値は $(\lambda-\mu)^{-1}$。
  - μ が十分に正確であれば、$(\lambda-\mu)^{-1}$ は絶対値最大の固有値。行列 $(A-\mu E)^{-1}$ を適当な初期ベクトルにかけ続けると λ に対応する固有ベクトルに収束 (c.f. べき乗法)
  - 実際には $(A-\mu E)x' = x$ という連立方程式を繰り返し解く

変分法

- 波動関数を互いに直交する正規化された波動関数 (基底関数) の線形結合で近似する (変分波動関数、試行関数)

$$|\psi\rangle = \sum_{p=1}^{m} C_p |\phi_p\rangle \qquad (\langle\phi_p|\phi_q\rangle = \delta_{pq})$$

- エネルギーの期待値

$$E = \frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle} = \frac{\sum_{p,q} C_p^* H_{pq} C_q}{\sum_{p,q} C_p^* \delta_{pq} C_q}$$

$$H_{pq} = \langle\phi_p|H|\phi_q\rangle$$

- E ができるだけ小さくなるよう係数 C_p を最適化 (変分原理)

変分法

- $\delta E = 0$ から

$$\sum_q (H_{pq} - E\delta_{pq})C_q = 0 \qquad \text{for } {}^\forall p$$

- H_{pq}, δ_{pq} を $m \times m$ 行列と考えると、固有値問題とみなせる

$$HC = EC$$

- H はエルミート行列
- $\{\phi_p\}$ の張る部分空間での最適化 ($=$ Rayleigh-Ritz の方法)
- 変分波動関数と真の波動関数の差が ϵ 程度の時、E と真の固有エネルギーの差は ϵ^2 程度

非直交基底関数による変分法

- 重なり積分

$$S_{pq} = \langle \phi_p | \phi_q \rangle \neq \delta_{pq}$$

- 変分波動関数の正規化条件

$$\langle \psi | \psi \rangle = \sum_{p,q} C_p^* \langle \phi_p | \phi_q \rangle C_q = \sum_{p,q} C_p^* S_{pq} C_q = 1$$

- エネルギー期待値

$$E = \frac{\sum_{p,q} C_p^* H_{pq} C_q}{\sum_{p,q} C_p^* S_{pq} C_q}$$

- $\delta E = 0$ から

$$\sum_q (H_{pq} - E S_{pq}) C_q = 0 \ \Rightarrow \ HC = ESC \text{ (一般化固有値問題)}$$

一般化固有値問題

- 重なり行列 $S_{pq} = \langle \phi_p | \phi_q \rangle$
  - エルミート行列: $S_{pq} = S_{qp}^*$
  - 正定値 ($\{\phi_p\}$ が線形独立の場合):

$$x^\dagger S x = \sum_{pq} \langle \phi_p | \phi_q \rangle x_p^* x_q = || \sum_p x_p | \phi_p \rangle ||^2 > 0$$

- 一般化固有値問題 $\Rightarrow$ 2 回の固有値分解により解くことができる
  - S を固有値分解: $S = UDU^\dagger$
  - S 固有値は全て正 $\Rightarrow$ $D^{-1/2}$ を定義可
  - $HC = ESC \Rightarrow D^{-1/2} U^\dagger H U D^{-1/2} D^{1/2} U^\dagger C = E D^{1/2} U^\dagger C$
  - $H' = D^{-1/2} U^\dagger H U D^{-1/2}$、$C' D^{1/2} U^\dagger C$ とおくと

$$H'C' = EC' \qquad \text{(通常の) 固有値問題}$$

  - (1 回目の固有値分解はコレスキー分解 $A = LL^\dagger$ (L は下三角行列) を用いてもよい)