連続系アルゴリズム レポート課題8

- 宿題(レポート)
  - 提出状況に応じて，期末試験の点に最大 50 点を加算します
    - まったく提出しなくても，期末試験の点で成績が出ます
    - ただし，試験を受けなければ単位は出ません
  - プログラムだけはメールの添付ファイルで
    reiji@is.s.u-tokyo.ac.jp
    に送付してください
    - 所属学科，学年，学籍番号，氏名をメール本文に明記すること
    - メールの題名(Subject)は「連続系レポート課題第8回提出者氏名」とすること
  - 手計算をするものや，計算結果と考察など，プログラム以外は A4 の紙(裏もつかってよい)にまとめてください
    - 所属学科，学年，学籍番号，氏名をレポートの最初に明記すること
    - 次回の講義の前に集めます
    - あるいは PDF で A4 サイズでもよい(プログラムと一緒にメールで送付してください)
    - ページ数制約を緩和しました

- 問題1
  - $\|A\|_2 = \sqrt{\max \lambda(A^T A)}$ を証明せよ

- 問題2
  - A を条件 $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$ を満たす適当な行列として，

    以下の方法で連立一次方程式 $Ax = b$ を解け
    - A の対角要素からなる対角行列 D とする(つまり $d_{ii} = a_{ii}$)
    - 初期ベクトルを $x_{(0)} = 0$ とする
    - 反復改良に類似の $x_{(i)} = x_{(i-1)} + D^{-1}(b - Ax_{(i-1)})$ を繰り返す
    - 残差 $r_{(i)} = b - Ax_{(i)}$ のノルムの変化をグラフ(片対数)で示せ

試験 1/23 の予定

Sherman-Morrison-Woodbury の公式

- $(A + B\,C)^{-1} = A^{-1} + A^{-1}B\,(I + C\,A^{-1}B)^{-1}\,C\,A^{-1}$
  - 証明：右辺に $A + B\,C$ を掛ける
- B が $n \times k$, C が $k \times n$ で $(A + B\,C)\,x = b$ を解く
  - A はあらかじめ LU 分解されている，また $k < n$ とする

$D = I + C\,(A^{-1}B)$ ・・・ $O(kn^2)$

D を LU 分解 ・・・ $O(k^3)$

$y = A^{-1}b$ ・・・ $O(n^2)$

$z = B\,D^{-1}\,C\,y$ ・・・ $O(kn)$

$x = y + A^{-1}z$ ・・・ $O(n^2)$

合計 $O(kn^2)$
すごく役に立つ

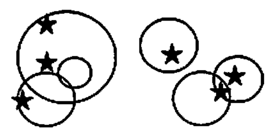
しかし，丸め誤差には弱いという

Gerschgorin の定理

- 定理：行列 A の各固有値は，中心 a_{ii} 半径

$$r_i = \sum_{j \neq i} |a_{ij}|$$

の n 個の円盤の和集合に含まれている
k 個の円盤が他の円盤と重ならないとき，これら k 個の円盤の和集合は k 個の固有値を含む

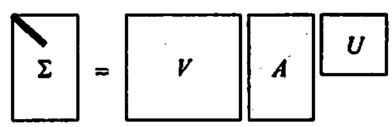


$$\Gamma_i = \{z \in \mathbf{C} \mid |z - a_{ii}| \leq r_i\}$$

注：各円盤がひとつずつ固有値を含むのではない

31

特異値分解



- A が $n \times m$ 行列のとき，ユニタリ行列 V, U で

$$\Sigma = V^H A U = \begin{pmatrix} \Sigma_r & 0 \\ 0 & 0 \end{pmatrix}, \quad \Sigma_r = diag(\sigma_1, \cdots, \sigma_r)$$

を満たすものがある（$r = \mathrm{rank}\, A$）
  - σ_i は正の実数で，A の特異値と呼ばれる
- いろんな応用がある
  - 1．低階数近似・・・行列を階数の低い行列で近似
  - 2．一般化逆行列・・・最小二乗最小ノルム解
  - 3．正規化法・・・悪条件の方程式の安定な近似解

32

条件数について

- LU 分解をしてみると，悪条件の行列では，U の対角要素に非常に小さい値が出ることが多い

だいたいこうなる

$$\begin{pmatrix} 8.5275751 & -7.1415306 & 8.5050736 & -2.9522493 & -0.1814509 \\ 0. & -7.8222027 & 3.5000545 & -6.9798529 & -5.4179199 \\ 0. & 0. & -9.1817494 & 3.8092983 & -2.7553876 \\ 0. & 0. & 0. & -3.6124519 & 2.5613292 \\ 0. & 0. & 0. & 0. & -7.9899283 \end{pmatrix} \quad \text{条件数 } 10$$

$$\begin{pmatrix} 0.6248703 & 0.4253652 & -0.2750253 & -0.1705542 & 0.1657151 \\ 0. & -0.7217696 & -0.1880898 & -0.2655097 & -0.6009392 \\ 0. & 0. & -0.7851611 & -0.6483216 & 0.3760717 \\ 0. & 0. & 0. & 0.7906930 & 0.7356841 \\ 0. & 0. & 0. & 0. & \underline{-3.571\mathrm{D}{-10}} \end{pmatrix} \quad \text{条件数 } 10^{10}$$

対称行列の摂動（摂動は非対称）

$$A = Q^{-1}\begin{pmatrix} 10 & & & & \\ & 9 & & & \\ & & 8 & & \\ & & & 7 & \\ & & & & 6 \end{pmatrix}Q$$

Q: 直行行列

$$B = A + \underbrace{E}_{\text{エラーだから}} \qquad \|E\| \approx 10^{-11}$$

計算された A の固有値
9.99999999999999993
8.99999999999999986
8.00000000000000002
6.99999999999999996
5.99999999999999996

計算された B の固有値
10.00000000000005707
8.99999999999997831
8.00000000000017819
7.00000000000003973
6.00000000000004933

対角行列の条件数 1

対角行列にでるから よい.

固有ベクトル系が悪条件のとき

$$A = X^{-1}\begin{pmatrix} 10 & & & & \\ & 9 & & & \\ & & 8 & & \\ & & & 7 & \\ & & & & 6 \end{pmatrix}X$$

$\mathrm{cond}(X) = 10^6$

$$B = A + E \qquad \|E\| \approx 10^{-11}$$

計算された A の固有値
10.000000453501436
8.999999486761467
8.000000106367200
6.999999842107712
6.000000111237143

計算された B の固有値
10.000000541771822
8.999999401220080
8.000000115333524
6.999999806133496
6.000000135559154

サイズ 5 の Jordan ブロック

$$A = Q^{-1}\begin{pmatrix} 10 & 1 & & & \\ & 10 & 1 & & \\ & & 10 & 1 & \\ & & & 10 & 1 \\ & & & & 10 \end{pmatrix}Q$$

Q: 直行行列

$$B = A + E \qquad \|E\| \approx 10^{-11}$$

計算された A の固有値
10.000851616170067
10.000851616170067
9.999674642076499
9.999674642076499
9.998947483506889

計算された B の固有値
10.004179759771153
10.001267887051625
10.001267887051625
9.996642233073947
9.996642233073947

矛盾？（1）

- 連立一次方程式 $A\,x = b$

$$\begin{pmatrix} 0.665 & 0.850 \\ 0.628 & 0.686 \end{pmatrix} x = \begin{pmatrix} 1.10 \\ 0.89 \end{pmatrix}$$

- 真の解　$x \approx \begin{pmatrix} 0.0244814 \\ 1.2749646 \end{pmatrix}$
- 近似解　$x_1 = \begin{pmatrix} 0.062 \\ 1.308 \end{pmatrix}, \quad x_2 = \begin{pmatrix} -0.356 \\ 1.598 \end{pmatrix}$
- どっちが真の解に近い？

1

矛盾？（2）

x_1のほうがちかい?

- x と x_1, x_2 の差（誤差）は約 10 倍違う

$$x_1 - x \approx \begin{pmatrix} 0.037 \\ 0.033 \end{pmatrix} \qquad x_2 - x \approx \begin{pmatrix} -0.380 \\ 0.323 \end{pmatrix}$$

小っちゃいほうがよくない

- $A\,x = b$ の解としてみると？残差の大小は逆転

$$b = \begin{pmatrix} 1.10 \\ 0.89 \end{pmatrix} \qquad Ax_1 \approx \begin{pmatrix} 1.153 \\ 0.936 \end{pmatrix} \qquad Ax_2 \approx \begin{pmatrix} 1.122 \\ 0.873 \end{pmatrix}$$

$$Ax_1 - b \approx \begin{pmatrix} 0.053 \\ 0.046 \end{pmatrix} \qquad Ax_2 - b \approx \begin{pmatrix} 0.022 \\ -0.017 \end{pmatrix}$$

2

x_2のほうがちかい?

条件数と解の精度（1）

- Vandermonde 行列
  - 非常に条件が悪い

$$A = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \end{pmatrix}$$

  - 例えば $x_i = (i-1)/(n-1)$ すなわち 0 から 1 まで等間隔とすると
    - $n = 11$ で $\kappa = 1.156\text{e}+8$　& 精度が $n=10$ まで.
    - $n = 16$ で $\kappa = 3.122\text{e}+12$
    - $n = 21$ で $\kappa = 8.482\text{e}+16$

3

条件数と解の精度（2）

$p(x) = c_0 + c_1 x + \cdots + c_{n-1}x^{n-1}$

$p(x_i) = b_i$.

- 多項式 $p(x)$ に対して $b_i = p(x_i)$ とすれば $c = (A^T)^{-1}\,b$ は $p(x)$ のべき表現を与える

べき表現ダメ.

- $p(x) = 1 + x + x^2 + \cdots + x^{n-1}$ としてやってみると

$n=11$	$n=16$		$n=21$	
1.0000000000	1.0000000011	0.0000000000	1.0000000003	0.6572977273
1.0000000000	0.9999783320	0.0000000000	0.9999962886	0.0000000000
1.0000000000	1.0007691425	3.2571071695	1.0001630724	3.8702348649
1.0000000001	0.9895508966	-0.4880069920	0.9973328298	0.0000000000
0.9999999996	1.0730547607	1.3495135338	1.0215733005	0.0000000000
1.0000000015	0.7069729957		0.9077172042	0.0000000000
0.9999999967	1.7000874333		1.1820743002	3.5959833991
1.0000000047	0.0000000000		1.0497532109	0.0000000000
0.9999999959	1.9302085130		0.0000000000	0.6372272014
1.0000000020	0.0000000000		2.8593792335	1.2212673669
0.9999999996	2.4807642139		0.0000000000	

4

all 1が正しい。　　全くダメ.

連続系アルゴリズム 12/12

ノルム.

・ベクトルのノルム

2ノルム.

$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2}$$

最大値ノルム

$$\|x\|_\infty = \max |x_i| .$$

1ノルム

$$\|x\|_1 = \sum_{i=1}^{n} |x_i| .$$

これらの関係.

$$\|x\|_2 \le \|x\|_1 \le \sqrt{n}\,\|x\|_2$$

$$\|x\|_\infty \le \|x\|_2 \le \sqrt{n}\,\|x\|_\infty .$$

$$\|x\|_\infty \le \|x\|_1 \le n\,\|x\|_\infty$$

どのノルムでも大体同じ.

・行列のノルム

Frobeniusノルム

$$\|A\|_F = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} |a_{ij}|^2}$$

これは一貫性をもつ:

$$\|Ax\|_2 \le \|A\|_F \|x\|_2 .$$

この性質をコンパチブルという.

一般に $\|AB\|_F \le \|A\|_F \|B\|_F$

cf. $\|I\|_F = \sqrt{n} \ne 1 \quad (n > 1)$

単位行列のFrobeniusノルムが1でない.

作用素ノルム

$$\|A\|_p = \sup_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_p} = \sup_{\|x\|_p = 1} \|Ax\|_p$$

今日 max でおきかえてもよい.

ベクトルのノルム.

行列のノルムが決まる.

$p = 1, 2, \infty$.

cf. $\|I\|_p = 1$

やはり一貫性がなりたつ.

$$\|Ax\|_p \leq \|A\|_p \cdot \|x\|_p$$

$$\|AB\|_p \leq \|A\|_p \cdot \|B\|_p$$

cf. 直交行列 Q に対し, $\|Q\|_2 = 1$.

具体的に示すことができる.

1ノルム（列和ノルム）

$$\|A\|_1 = \max_j \left\{ \sum_{i=1}^{n} |a_{ij}| \right\}$$

$\downarrow$

∞ノルム（行和ノルム）

$$\|A\|_\infty = \max_i \left\{ \sum_{j=1}^{n} |a_{ij}| \right\}$$

$\longrightarrow$

2ノルム（スペクトルノルム）

$$\|A\|_2 = \sqrt{\max \lambda(A^T A)}$$

$\lambda(B)$ は B の固有値の集合.

<u>方針</u>:

実際に最大値をとる x を上手く作ってやる.

また, それを超えることはない.

誤差と残差.

$A\mathbf{x} = \mathbf{b}$ の解 $\mathbf{x}$, 近似解 $\tilde{\mathbf{x}}$

誤差 $\mathbf{e} = \tilde{\mathbf{x}} - \mathbf{x}$

残差 $\mathbf{r} = A\tilde{\mathbf{x}} - \mathbf{b}$.

丸め誤差レベル.

$$\mathbf{e} = A^{-1}\mathbf{r}.$$

$$\|\mathbf{e}\| \le \underline{\|A^{-1}\|\,\|\mathbf{r}\|}$$

どのノルムでもよい.

誤差の大きさ　残差の大きさ.

A^{-1}の大きさぐらい拡大されうる.

条件数.

係数行列に小さい誤差をもち, $\hat{A}$ であるとする.

$\hat{A}\tilde{\mathbf{x}} = \mathbf{b}$ とする.

$$\begin{aligned}
\tilde{\mathbf{x}} - \mathbf{x} &= \hat{A}^{-1}\mathbf{b} - A^{-1}\mathbf{b} \\
&= (\hat{A}^{-1} - A^{-1})\mathbf{b} \\
&= A^{-1}(A - \hat{A})\hat{A}^{-1}\mathbf{b}.
\end{aligned}$$

"相対誤差"　　一貫性.

$$\begin{aligned}
\frac{\|\tilde{\mathbf{x}} - \mathbf{x}\|}{\|\tilde{\mathbf{x}}\|} &= \frac{\|A^{-1}(A - \hat{A})\hat{A}^{-1}\mathbf{b}\|}{\|\hat{A}^{-1}\mathbf{b}\|} \\
&\le \frac{\|A^{-1}\|\,\|A - \hat{A}\|\,\|\hat{A}^{-1}\mathbf{b}\|}{\|\hat{A}^{-1}\mathbf{b}\|} \\
&= \|A^{-1}\|\,\|A\|\,\frac{\|\hat{A} - A\|}{\|A\|} \\
&= K(A)\boxed{\frac{\|\hat{A} - A\|}{\|A\|}}
\end{aligned}$$

計算する解するために $\|\mathbf{x}\|$ でなく $\|\tilde{\mathbf{x}}\|$ とする.

挿入

Aの相対誤差.

$K(A) = \|A\|\,\|A^{-1}\|$　条件数という.

Aの相対誤差の条件数倍

$\gtrsim$ xの相対誤差.

* 直交行列 Q

$$\|Q\|_2 = \|Q^{-1}\|_2 = 1$$

$$\kappa(Q) = 1$$

* 対角スケーリング

$$Ax = b$$

対角行列 D.

$$DAx = Db$$

AとDAでは条件数がちがう

→ Dの選び方によって
精度を高めることができる

DAの各行のノルムが1になるようにする.

誤差を含んだ $b \to \hat{x}$
を $\hat{A}^{-1}$と書くことにする.

真の解 $x = \hat{x} + A^{-1}(b - A\hat{x})$.

↓

$\hat{A}^{-1}$におきかえたら,
何がおこる?

cf. $\hat{A}^{-1}$はつくらない.
過去に分解の結果
で解く.

反復改良.

$$x^{(k+1)} = x^{(k)} + \underline{\hat{A}^{-1}(b - Ax^{(k)})}$$

1回代入してみる.

$$\begin{aligned} b - Ax^{(k+1)} &= b - A\left(x^{(k)} + \hat{A}^{-1}(b - Ax^{(k)})\right) \\ &= b - Ax^{(k)} - A\hat{A}^{-1}(b - Ax^{(k)}) \\ &= (I - A\hat{A}^{-1})(b - Ax^{(k)}) \end{aligned}$$

$I - A\hat{A}^{-1}$の固有値の絶対値が1未満 → 残差は減る.