数値計算

京都大学大学院情報学研究科
大木 健太郎
連絡先： ohki@i.kyoto-u.ac.jp

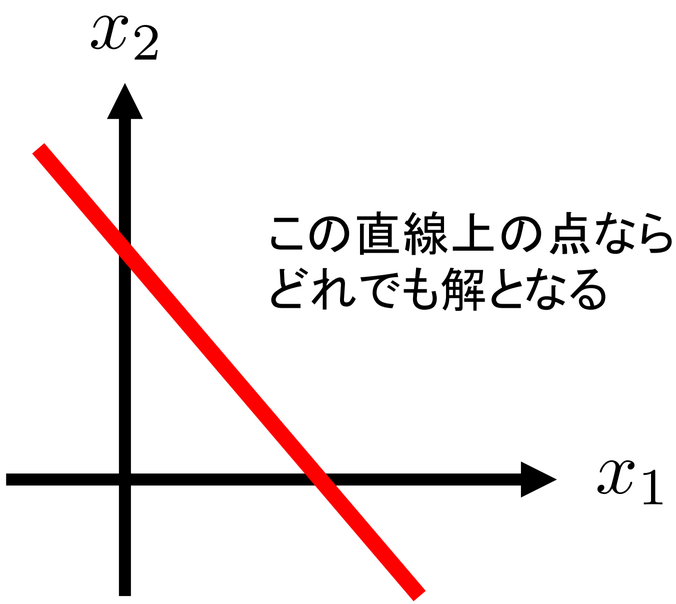
第8回
- 非線形方程式の数値解法：誤差解析
- 代数方程式の数値解法

補足：連立1次方程式（劣決定）

$$Ax = b, \quad A \in \mathbb{R}^{m\times n},\ b \in \mathbb{R}^m,\ n \neq m$$

✓ $n > m$ の場合

（例） $A = \begin{bmatrix} 1 & 1 \end{bmatrix},\ b = 1$



この直線上の点ならどれでも解となる

$\mathrm{rank}(A) = m$ の場合の一般解

任意の $c \in \mathbb{R}^n$ に対し

$$x = A^\top (AA^\top)^{-1} b + \left(I - A^\top (AA^\top)^{-1} A\right) c$$

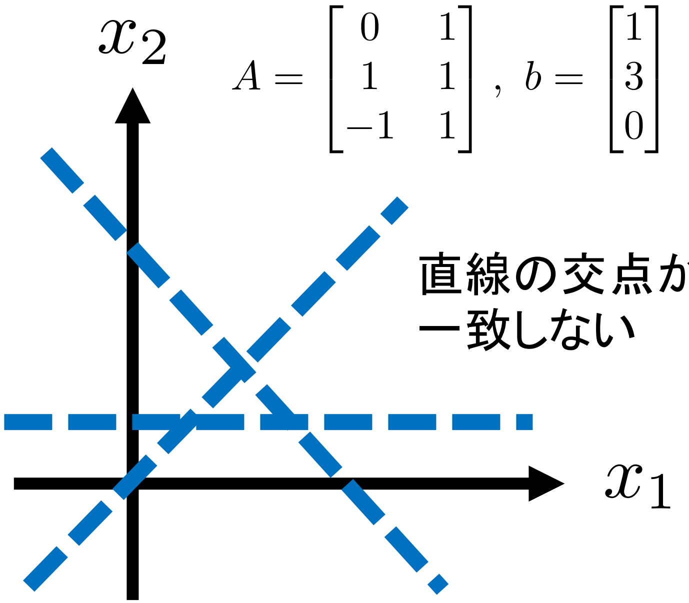
補足：連立1次方程式（優決定）

$$Ax = b, \quad A \in \mathbb{R}^{m \times n},\ b \in \mathbb{R}^m,\ n \neq m$$

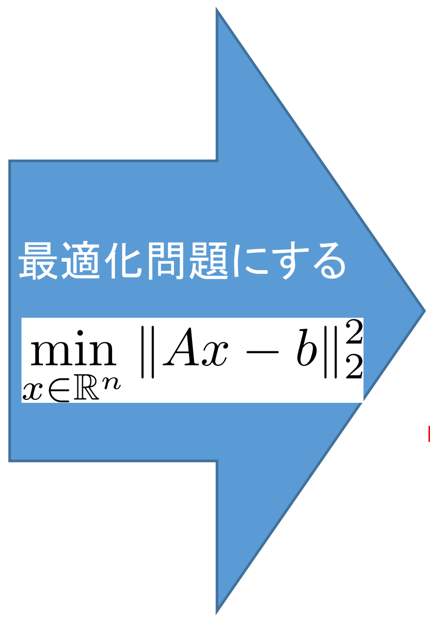
✓ $n < m$ の場合

（例）

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ -1 & 1 \end{bmatrix},\ b = \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix}$$



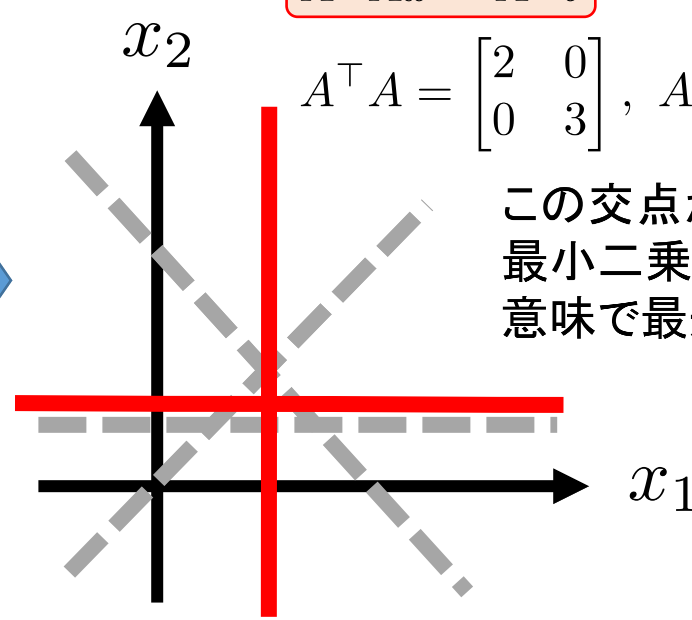
直線の交点が
一致しない



これを解けばよい

$$\boxed{A^\top A x = A^\top b}$$

$$A^\top A = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix},\ A^\top b = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$



この交点が
最小二乗誤差の
意味で最適解

本日の内容

- 非線形方程式の数値解法
  - 加速法
- 代数方程式の数値解法
  - Sturm の方法
  - 連立法: Durand-Kerner 法
  - 連立法: Ehrlich-Aberth 法
  - 多項式の計算と微分法

収束の速度

反復法 $x^{(k+1)} = \phi(x^{(k)}),\ x^{(0)} \in \underline{K} \subset \mathbb{R}^n$

有界閉集合

仮定 ϕ は $(K, \|\bullet\|)$ 上で縮小写像

⇒ 収束先を x^* とする

定義:収束の速度

- $\displaystyle \lim_{k\to\infty} \frac{\|x^{(k+1)} - x^*\|}{\|x^{(k)} - x^*\|} = \mu$ ⇒ 線形収束, 1次収束
- $\displaystyle \lim_{k\to\infty} \frac{\|x^{(k+1)} - x^*\|}{\|x^{(k)} - x^*\|} = 0$ ⇒ 超1次収束
- $\displaystyle \lim_{k\to\infty} \frac{\|x^{(k+1)} - x^*\|}{\|x^{(k)} - x^*\|^p} > 0, \quad p > 1$ ⇒ p 次収束

Newton 法は真値の近傍で2次収束

Aitken の加速法

少ないステップで収束した方が嬉しい

⇨ **加速法**: より速い収束列を作る方法

簡単のために $n = 1$ とする

$\{x^{(k)}\}$: 収束列（収束先: x^*, 線形収束）

$$y^{(k)} = x^{(k)} - \frac{(x^{(k+2)} - x^{(k+1)})^2}{x^{(k+2)} - 2x^{(k+1)} + x^{(k)}}$$

で生成する列 $\{y^{(k)}\}$ は元の収束列よりも速くに x^* 収束する:

$$\lim_{k \to \infty} \frac{y^{(k)} - x^*}{x^{(k+2)} - x^*} = 0$$

※ 証明は [山本哲朗: 数値解析入門, p. 74, サイエンス社(2003)] 参照

関数の微分を必要としないアルゴリズム

関数 f が微分できない場合，Newton 法は使えない

⇨ Steffensen 反復法： $x = \phi(x)$ の根を求めるアルゴリズム

$$\phi(x) = x + f(x) \quad \text{として}$$

$$z^{(k+1)} = \phi(\phi(z^{(k)})) - \frac{(\phi(\phi(z^{(k)})) - \phi(z^{(k)}))^2}{\phi(\phi(z^{(k)})) - 2\phi(z^{(k)}) + z^{(k)}}$$

- ✓ Aitken 加速と似ているが，Aitken 加速が発散しても初期値の選び方によって収束
- ✓ 収束速度はNewton法と同程度

本日の内容

- 非線形方程式の数値解法
  - 加速法
- 代数方程式の数値解法
  - Sturm の方法
  - 連立法: Durand-Kerner 法
  - 連立法: Ehrlich-Aberth 法
  - 多項式の計算と微分法

代数方程式

- 代数方程式：応用でよく現れる
  - 多項式：（例） $f(x) = \sum_{j=0}^{m} a_j x^j$
  - 代数方程式： $f(x) = 0$
- 根の存在が保証され，根の数も保証される

代数方程式の根に関する数値計算法

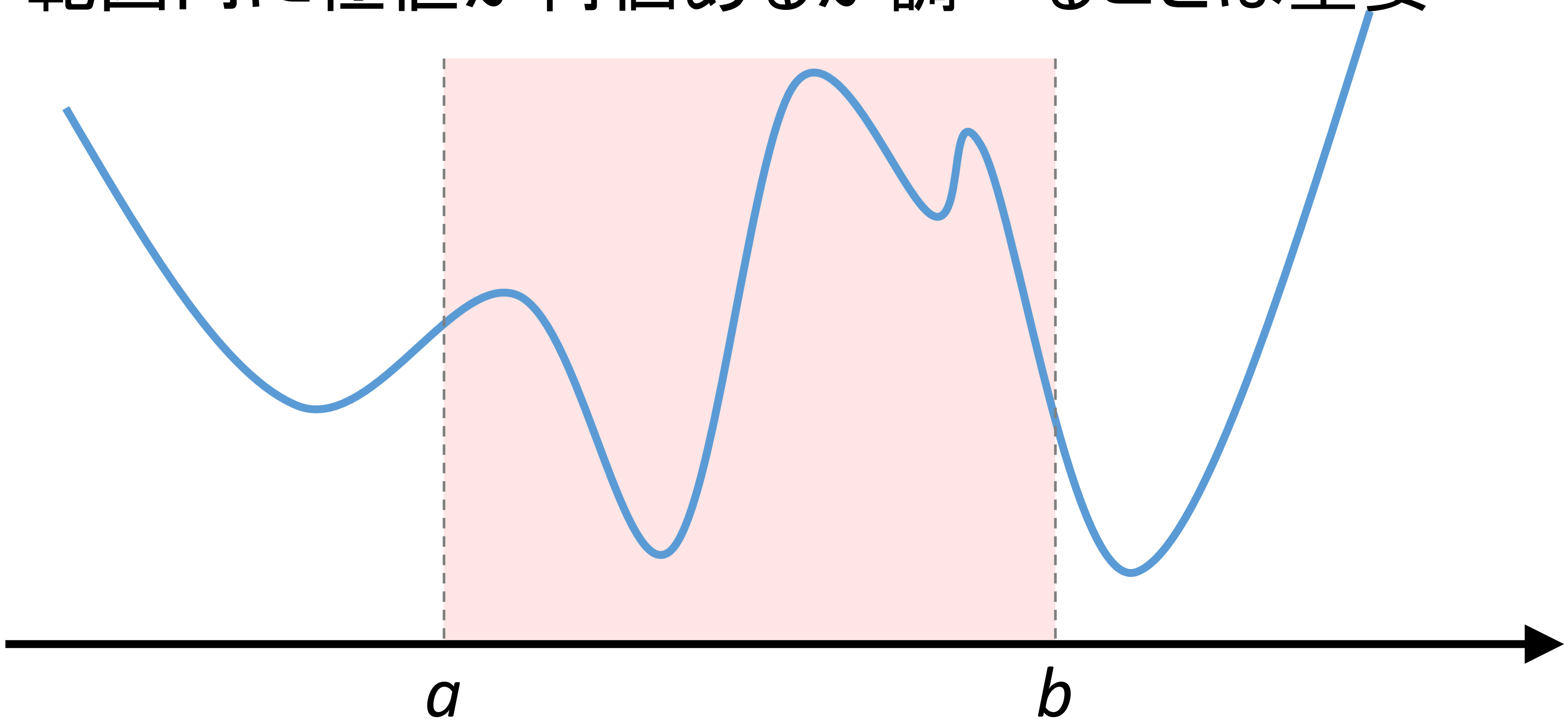
- ある指定した範囲に根がいくつ存在するか
  ⇒ Strum の方法

- 複素数の根を求める
  ⇒ Bairstow-Hitchcock 法

- 全ての根を求める
  ⇒ DKA 法 (Durand-Kerner-Aberth法)

根の数を調べること

- 最適化の問題では，パラメータを動かせる範囲が制限されることが多い
- 範囲内に極値が何個あるか調べることは重要



Sturm列

$$[a,b],\quad a,b\in\mathbb{R}, a<b$$

定義（Sturm 列）

実係数多項式の列 $f(x),\ f_1(x),\ \cdots,\ f_l(x)$ が区間[*a*, *b*]においてSturm 列であるとは，次の4つを満たすことをいう.

(1) $x\in[a,b],\ f_k(x)=0\Rightarrow f_l(x)\neq 0,\ \ k,l=0,1,\cdots,l-1,\ \ |k-l|=1$

(2) $f_k(x_0)=0\Rightarrow f_{k-1}(x_0)f_{k+1}(x_0)<0,\ \ k=1,\cdots,l-1$

(3) $(f_l(x)>0,\ \forall x\in[a,b])$ or $(f_l(x)<0,\ \forall x\in[a,b])$

(4) $f(x_0)=0\Rightarrow f(x_0)f_1(x_0)>0$

Sturm列の例

次はSturm 列の定義を満たす

$$
\begin{aligned}
f(x) &= x^2 + a_1 x + a_0, \quad a_0, a_1 \in \mathbb{R} \setminus \{0\}, a_1 \neq 2a_0 \\
f_1(x) &= 2x + a_1 \\
f_2(x) &= a_0 - \frac{1}{4} a_1^2
\end{aligned}
$$

※f は重複する根を持たないことに注意

Sturmの定理

仮定

実係数多項式の列 $f(x),\ f_1(x),\ \cdots,\ f_l(x)$ が区間[*a*, *b*]においてSturm 列であり, $f(a)f(b) \neq 0$.

$N(x)$: $x \in [a, b]$ において $[f(x), f_1(x), \cdots f_l(x)]$ を左から見たときの符号反転の数

定理2.2 (Sturmの定理)

$f(x)$ の区間 [*a*, *b*] に存在する零点の個数は,

$$n_0 = N(a) - N(b)$$

符号反転回数の数え方の注意

- 次の場合は，一回反転したものと数える

$f_{k-1}(x)$	$f_k(x)$	$f_{k+1}(x)$
$+$	0	$-$

Sturm列の生成法

1. 多項式 f は区間 $[a, b]$ で重根を持たないものとする（Sturm 列の条件(1)を満たすため）
2. まず微分し，1つ目の多項式を作る
3. f を f_1 で割り算し，剰余のマイナスを f_2 とする.
4. 以降，f_k を f_{k-1} で割り算し，剰余のマイナスを f_{k+1} とする.
5. $f_l(x)$ = 非ゼロの定数になったら終了

$$f(x) = x^2 + a_1 x + a_0, \quad a_0, a_1 \in \mathbb{R} \setminus \{0\},\ a_1 \neq 2a_0$$

$$f_1(x) = f'(x) = 2x + a_1$$

$$f_2(x) = \underbrace{\left(\frac{1}{2}x + \frac{1}{4}a_1\right)}_{g_1(x)} f_1(x) - f(x) = -a_0 + \frac{1}{4}a_1^2$$

次数を下げるようにする

重根のある場合のSturm列

$$f(x) = (x+1)^2(x-5) = x^3 - 3x^2 - 9x - 5$$

のSturm 列を計算(重根の場合)

Scilabコード例

```
x = poly(0,'x');
f = x^3 – 3 x^2 – 9 x -5;
f1 = 3 x^2 – 6 x – 9;
[f2, g1] = pdiv(f, f1);
f2=-f2;
[f3, g2] = pdiv(f1, f2);
```

$$f_2(x) = \frac{1}{8}(x+1)$$ ≠ 定数

$$g_2(x) = -\frac{9}{8}x + \frac{3}{8}$$

$$f_3(x) = 0$$

重根のある場合のSturm列

$$\tilde{f}(x) = \frac{f(x)}{f_2(x)} = 8(x+1)(x-5) = 8x^2 - 32x - 40$$

のSturm 列を計算（重根の場合）

Scilabコード例

```
x = poly(0,'x');
f = 8x^2 – 32 x – 40;
f1 = 16 x – 32;
[f2, g1] = pdiv(f, f1);
f2=-f2;
```

Sturm列

$$\tilde{f}(x) = 8x^2 - 32x - 40$$

$$\tilde{f}_1(x) = 16x - 32$$

$$\tilde{f}_2(x) = 72$$

Sturmの方法

- Sturm の定理を使って，
- 区間 $[a, b]$ を与えて，根の数を調べる
- 区間 $[a, (a+b)/2]$, $[(a+b)/2, b]$ の根の数を調べる
- どんどん2分割していき，根が一つしかない区間を十分小さくして，その区間内の適当な値を根の近似値として採用する

Sturmの方法:具体例

- $f(x) = x^2 - 1x - 2 = 0$ を求める(区間[-10,10])
- Sturm 列

$$
\begin{aligned}
f(x) =& x^2 - 1x - 2 \\
f_1(x) =& 2x - 1 \\
f_2(x) =& 2 + 0.25 = 2.25
\end{aligned}
$$

- 区間[-10, 10] 内の根の数

$$[f(10),\quad f_1(10),\quad f_2(10)] = [88,\quad 19\quad 2.25] \Rightarrow N(10) = 0$$

$$[f(-10),\quad f_1(-10),\quad f_2(-10)] = [108,\quad -21\quad 2.25] \Rightarrow N(-10) = 2$$

$$\Rightarrow \quad n_{[-10,10]} = N(-10) - N(10) = 2$$

Sturmの方法：具体例

- 区間を2分割して，根の数を計算
  - 区間 [-10,0] の根の数　$n_{[-10,0]} = 1$
  - 区間 [0,10] の根の数　$n_{[0,10]} = 1$
- 再び区間を2分割して，根の数の計算を繰り返す

$$\begin{bmatrix} f(10), & f_1(10), & f_2(10) \end{bmatrix} = \begin{bmatrix} 88, & 19 & 2.25 \end{bmatrix} \Rightarrow N(10) = 0$$

$$\begin{bmatrix} f(0), & f_1(0), & f_2(0) \end{bmatrix} = \begin{bmatrix} -2, & -1 & 2.25 \end{bmatrix} \Rightarrow N(0) = 1$$

$$\begin{bmatrix} f(-10), & f_1(-10), & f_2(-10) \end{bmatrix} = \begin{bmatrix} 108, & -21 & 2.25 \end{bmatrix} \Rightarrow N(-10) = 2$$

本日の内容

- 非線形方程式の数値解法
  - 加速法
- 代数方程式の数値解法
  - Sturm の方法
  - 連立法: Durand-Kerner 法
  - 連立法: Ehrlich-Aberth 法
  - 多項式の計算と微分法

代数方程式の連立法

- 多項式の根を全て求める計算法
  - Durand-Kerner 法：２次収束

$$z_i^{(k+1)} = z_i^{(k)} - \frac{p_n(z_i^{(k)})}{p_n'(z_i^{(k)})}$$

    - Durand-Kerner-Aberth 法：初期値の取り方を工夫

$$z_i^{(0)} = -\frac{a_1}{n} + r_0 \exp\left(\sqrt{-1}\left(\frac{2(i-1)\pi}{n} + \frac{\pi}{2n}\right)\right)$$

  - Ehrlich-Aberth 法：３次収束

$$z_i^{(k+1)} = z_i^{(k)} - \frac{\frac{p_n(z_i^{(k)})}{p_n'(z_i^{(k)})}}{1 - \frac{p_n(z_i^{(k)})}{p_n'(z_i^{(k)})} \sum_{j=1, j\neq i}^{n} \frac{1}{z_i^{(k)} - z_j^{(k)}}}$$

連立法の導出の準備

- 多項式

$$p_n(z) = a_0 z^n + a_1 z^{n-1} + \cdots + a_{n-1} z + a_n,$$
$$a_0 \neq 0, a_i \in \mathbb{C}, i = 0, 1, \cdots, n$$

- 書き直すと

$$\begin{aligned} p_n(z) &= a_0 z^n + a_1 z^{n-1} + \cdots + a_{n-1} z + a_n, \\ &= a_0 (z - \alpha_1) \cdots (z - \alpha_n) \end{aligned}$$

求めたい多項式の根

Durand–Kerner法の導出

- 多項式に Newton 法をそのまま適用

$$
\begin{aligned}
z_i^{(k+1)} &= z_i^{(k)} - \frac{p_n\left(z_i^{(k)}\right)}{p_n'\left(z_i^{(k)}\right)} \\
&= z_i^{(k)} - \frac{p_n\left(z_i^{(k)}\right)}{a_0 \prod_{j=1, j\neq i}^{n} \left(z_j^{(k)} - z_j^{(k)}\right)}
\end{aligned}
$$

※ 重根のある場合は，そのまま適用できないことに注意

DK法の初期値の選択

命題（[伊理正夫：数値計算，朝倉書店(1981)]，定理12.1）

$$p_n(z) = a_0 z^n + a_1 z^{n-1} + \cdots + a_{n-1} z + a_n = 0$$

の解の絶対値は，方程式

$$q_n(z) = |a_0| z^n - |a_1| z^{n-1} - \cdots - |a_{n-1}| z - |a_n| = 0$$

の唯一の正の実数解 r を超えない

$$q_n(z) = 0 \;\Rightarrow\; 1 = \frac{|a_1|}{|a_0|} z^{-1} + \frac{|a_2|}{|a_0|} z^{-2} + \cdots + \frac{|a_n|}{|a_0|} z^{-n}$$

右辺は単調減少関数なので，必ず正の実数で唯一の解をもつ

DK法の初期値の選択

仮定 $a_0 = 1$ とする

原点シフト

$$p_n(z) = p_n\left(w - \frac{a_1}{n}\right)$$
$$= w^n + c_{n-2}w^{n-2} + \cdots + c_n = 0$$

$$S(w) := w^n - |c_{n-2}|w^{n-2} - \cdots - |c_n| = 0$$

命題より，唯一の正の実数解 r が存在し，
多項式の根は全て $\Gamma := \{z \in \mathbb{C} \mid |z + a_1/n| \leq r\}$
に含まれる

DK法の初期値の選択

- $r \leq r_0$ となる r_0 を適当な方法で選ぶ（Sturm の定理を使えば，$[0, r_0]$ 内に根が存在するかどうか調べられる）

$$z_i^{(0)} = -\frac{a_1}{n} + r_0 \exp\left(\sqrt{-1}\left(\frac{2(i-1)\pi}{n} + \frac{\pi}{2n}\right)\right)$$

- 原点から半径 r_0 の円周に，等間隔で初期値を配置
- 初期値の配置を実軸対称にしないように，$\pi/2n$ 角度をずらす（初期値を複素共役対にとると，真値に収束しない場合がある）

本日の内容

- 非線形方程式の数値解法
  - 加速法
- 代数方程式の数値解法
  - Sturm の方法
  - 連立法: Durand-Kerner 法
  - 連立法: Ehrlich-Aberth 法
  - 多項式の計算と微分法

Ehrlich–Aberth法の導出

- 多項式の対数をとる

$$\log p_n(z) = \log a_0 + \sum_{i=1}^{n} \log(z - \alpha_i)$$

- 両辺を微分し, α_i に十分近い z_i を用いて整理

$$\frac{p_n'(z)}{p_n(z)} = \frac{1}{z-\alpha_i} + \sum_{j=1, j\neq i}^{n} \left(\frac{1}{z-z_j} - \frac{z_j - \alpha_j}{(z-z_j)(z-\alpha_j)} \right)$$

$$\simeq \frac{1}{z-\alpha_i} + \sum_{j=1, j\neq i}^{n} \frac{1}{z-z_j}$$

Ehrlich-Aberth法の導出

- 両辺を整理

$$\alpha_i \simeq z - \frac{\frac{p_n(z)}{p'_n(z)}}{1 - \frac{p_n(z)}{p'_n(z)} \sum_{j=1, j\neq i}^{n} \frac{1}{z - z_j}}$$

- α_i と z を置き換える

$$z_i^{(k+1)} = z_i^{(k)} - \frac{\frac{p_n(z_i^{(k)})}{p'_n(z_i^{(k)})}}{1 - \frac{p_n(z_i^{(k)})}{p'_n(z_i^{(k)})} \sum_{j=1, j\neq i}^{n} \frac{1}{z_i^{(k)} - z_j^{(k)}}}$$

連立法の性質

- Durand-Kerner 法
  - 収束の速さ: 真値に近いと2次収束（Newton法なので）
  - 収束する初期値の範囲: 多項式の根が単根の場合，Aberthの初期値設定で収束が保証される（DKA法）
  - Aberthの初期値配置を用いると，真値から遠いところでも1次収束が保証される
- Ehrlich-Aberth 法
  - 収束の速さ:真値に近いと3次収束（証明は次ページ）

Ehrlich-Aberth法の収束の速さ

$$\delta z_i^{(k)} := z_i^{(k)} - \alpha_i$$

$\boxed{\text{仮定}}$ $|\delta z_i^{(k)}| < \epsilon,\ i = 1, 2, \cdots, n$

$$\begin{aligned}
\delta z_i^{(k+1)} =& \delta z_i^{(k)} - \left(\frac{p_n'(z_i^{(k)})}{p_n(z_i^{(k)})} - \sum_{j=1, j\neq i}^{n} \frac{1}{z_i^{(k)} - z_j^{(k)}} \right)^{-1} \\
=& \delta z_i^{(k)} - \left(\sum_{j=1}^{n} \frac{1}{z_j^{(k)} - \alpha_j} - \sum_{j=1, j\neq i}^{n} \frac{1}{z_i^{(k)} - z_j^{(k)}} \right)^{-1} \\
=& \delta z_i^{(k)} - \left(\frac{1}{z_i^{(k)} - \alpha_i} - \sum_{j=1, j\neq i}^{n} \frac{\delta z_j^{(k)}}{(z_i^{(k)} - z_j^{(k)})(z_j^{(k)} - \alpha_j)} \right)^{-1}
\end{aligned}$$

Ehrlich–Aberth法の収束の速さ(続き)

$$
\begin{aligned}
\delta z_i^{(k+1)} =& \delta z_i^{(k)} - \delta z_i^{(k)} \left(1 - \delta z_i^{(k)} \sum_{j=1, j\neq i}^{n} \frac{\delta z_j^{(k)}}{(z_i^{(k)} - z_j^{(k)})(z_j^{(k)} - \alpha_j)} \right) \\
=& \delta z_i^{(k)} \left(1 - \left(1 + \sum_{m=1}^{\infty} \left(\delta z_i^{(k)} \sum_{j=1, j\neq i}^{n} \frac{\delta z_j^{(k)}}{(z_i^{(k)} - z_j^{(k)})(z_j^{(k)} - \alpha_j)} \right)^m \right) \right) \\
=& (\delta z_i^{(k)})^2 \sum_{j=1, j\neq i}^{n} \frac{\delta z_j^{(k)}}{(z_i^{(k)} - z_j^{(k)})(z_j^{(k)} - \alpha_j)} + O(\epsilon^5)
\end{aligned}
$$

$$\Rightarrow \quad \delta z_i^{(k+1)} = O(\epsilon^3)$$

3次の収束

補足

- 多項式に重根がある場合は，例えば
  伊理正夫: 数値計算，朝倉書店 (1981)
  を参照すること
- Matlab や Scilab には，係数ベクトルから多項式を生成するコマンド（poly）と，多項式の根を計算するコマンド（roots）がある．そこではJenkins-Traub法（コンパニオン行列の固有値を求める方法）が用いられている
  ※　固有値の計算法については次回以降の講義で扱う

本日の内容

- 非線形方程式の数値解法
  - 加速法
- 代数方程式の数値解法
  - Sturm の方法
  - 連立法: Durand-Kerner 法
  - 連立法: Ehrlich-Aberth 法
  - 多項式の計算と微分法

多項式の効率的な計算法

- 多項式の効率的な計算法 : Horner 法
- 通常の多項式の計算 : $n(n+1)/2$ 回の乗算
- Horner 法 : n 回の乗算

$$p_n(z) = a_0 z^n + a_1 z^{n-1} + \cdots + a_n$$

$$\begin{aligned} p_n(\zeta) =& a_0\zeta^n + a_1\zeta^{n-1} + \cdots + a_n \\ =& (\zeta - z)\left(b_0^{(1)}\zeta^{n-1} + b_1^{(1)}\zeta^{n-2} + \cdots + b_{n-1}^{(1)}\right) + b_n^{(1)} \end{aligned}$$

$$p_n(z) = b_n^{(1)}$$

Horner 法

形式的に整理する

$$
\begin{aligned}
p_n(\zeta) =& a_0\zeta^n + a_1\zeta^{n-1} + \cdots + a_n \\
=& (\zeta - z)\left(b_0^{(1)}\zeta^{n-1} + b_1^{(1)}\zeta^{n-2} + \cdots + b_{n-1}^{(1)}\right) + b_n^{(1)}
\end{aligned}
$$

$$\Rightarrow \quad p_n(z) = b_n^{(1)}$$

$b_l^{(1)}, l = 0, 1, \cdots, n$ を計算すればよい

$$
\left\{
\begin{aligned}
b_0^{(1)} =& a_0 \\
b_l^{(1)} =& z b_{l-1}^{(1)} + a_l
\end{aligned}
\right.
\quad \Rightarrow \quad p_n(z) = b_n^{(1)}
$$

多項式の微分法

- 多項式を Taylor 展開

$$
\begin{aligned}
p_n(\zeta) =& p_n(z) + (\zeta - z)p_n'(z) + (\zeta - z)^2 \frac{1}{2!} p_n''(z) + \cdots + (\zeta - z)^n \frac{1}{n!} p_n^{(n)}(z) \\
=& p_n(z) + (\zeta - z)\underline{q_1(\zeta, z)}
\end{aligned}
$$

$$
\begin{aligned}
q_1(\zeta, z) =& p_n'(z) + (\zeta - z)\frac{1}{2!} p_n''(z) + \cdots + (\zeta - z)^{n-1} \frac{1}{n!} p_n^{(n)}(z) \\
=& b_0^{(1)} \zeta^{n-1} + b_1^{(1)} \zeta^{n-2} + \cdots + b_{n-1}^{(1)}
\end{aligned}
$$

$$
p_n'(z) = b_0^{(1)} z^{n-1} + b_1^{(1)} z^{n-2} + \cdots + b_{n-1}^{(1)}
$$

Horner法で計算可能

多項式の高次微分

- *m* 回微分の場合

$$\begin{cases} b_0^{(m+1)} = a_0 \\ b_l^{(m+1)} = z b_{l-1}^{(m+1)} + b_l^{(m)}, \quad l = 1, 2, \cdots, n-m \end{cases}$$

$$\frac{d^m p_n(z)}{dz^m} = m! b_{n-m}^{(m+1)}$$

※ *z* を固定して，Horner 法を繰り返し適用

次回の講義予定

- 代数方程式(今回終わらなかった分)
- 行列の固有値問題(教科書第3章)