

数値計算の基礎

—解法と誤差—

付 録

工学博士 高倉 葉子 著

コロナ社

目 次

A. 計算機における数値表現

A.1 整数型の数の表現	1
A.2 実数型の数の表現 (浮動小数点数表示)	6
引用・参考文献	9

B. ベクトルと行列のノルム

B.1 ベクトルのノルム	10
B.1.1 ノルムの定義	10
B.1.2 エルミート行列とベクトルの内積について	12
B.1.3 新しいノルムの作り方	13
B.1.4 有限次元ベクトルノルムの同値性	14
B.2 行列のノルム	16
B.2.1 ナチュラルノルム	16
B.2.2 固有値と固有ベクトル	17
B.2.3 行列のノルムとスペクトル半径	20
引用・参考文献	23

C. 不動点反復法の収束に関する考察

C.1 1変数スカラー方程式	24
----------------------	----

ii 目 次

C.1.1 収束条件	25
C.1.2 適切な初期値の選択	27
C.1.3 議論の一般化	28
C.2 多変数の連立方程式系	28
C.2.1 収束条件：条件 ii) を満たすために	31
C.2.2 適切な初期値の選択：条件 i) を満たすために	32
引用・参考文献	33

D. 補間と誤差

D.1 多項式補間法：ラグランジュ補間とニュートン補間	34
D.2 多項式補間法の誤差	37
D.3 等間隔データ点に対する多項式補間の使用限界	40
D.4 チェビシユフ多項式による準ミニマックス近似	41
引用・参考文献	43

E. 数値微分の差分演算子表現

E.1 差分演算子と差分商	44
E.2 ニュートン補間法の差分演算子表現と数値微分	48
E.2.1 ニュートンの前進補間公式と数値微分	48
E.2.2 ニュートンの後退補間公式と数値微分	50
E.2.3 スターリングの補間公式と数値微分	51
E.3 微分演算子の差分演算子表現	54
引用・参考文献	58
付録 B の問題解答	59

A | 計算機における数値表現

計算機における数値の表現には整数型と実数型がある。数学における整数は無数個あるのに対し、計算機では有限個の“整数”しか扱えない。また数学における実数は連続で無限個の数の集合であるのに対し、計算機における“実数”は不連続（離散的）で有限個の数である。ここでは計算機における（近似的な）数の表現について述べる。

A.1 整数型の数の表現

数学では整数は無数個あるのに対し、計算機では有限の大きさで有限個の“整数”しか扱えない。10進数6桁の整数 - 123456 は

$$-123\,456 = -(1 \times 10^5 + 2 \times 10^4 + 3 \times 10^3 + 4 \times 10^2 + 5 \times 10^1 + 6 \times 10^0)$$

と書き表せるように、 β 進数 p 桁の整数は以下のように表わせる。

$$\begin{aligned} N &= \pm (d_1 d_2 \cdots d_{p-1} d_p)_{\beta\text{進数}} \\ &= \pm (d_1 \times \beta^{p-1} + d_2 \times \beta^{p-2} + \cdots + d_{p-1} \times \beta^1 + d_p \times \beta^0) \quad (\text{A.1}) \end{aligned}$$

ここに $0 \leq d_i \leq \beta - 1$ ($i = 1, \dots, p$) であるが、 $N \neq 0$ のときは $d_1 \neq 0$ 、 $N = 0$ のときは $p = 1$ 、 $d_1 = 0$ である。計算機では桁数 p に制約があるので、扱える整数の大きさの範囲は限られる。 β 進法 p 桁で表現できる整数は β^p 通りあり、符号なしの場合 0 を入れて

$$0 \leq N \leq \beta^p - 1$$

の範囲の値を表し, 符号付きの場合には負の数も入れるため, 正負それぞれ $\beta^p/2$ 通りの数が表現でき, 0 を正の側に入れて

$$-\frac{\beta^p}{2} \leq N \leq \frac{\beta^p}{2} - 1$$

の範囲の整数を表すことができる。

計算機内部で整数を表すと, 正の数の先頭部には 0 がたっている。正の数を n とすると, 計算機上では負の数 $-n$ は正の数の β の補数 $\beta^p - n$ に対応させて表現され, その先頭部は 0 ではない。ある整数 m に対応する計算機上の数をかぎ括弧を付けて $[m]$ で表すと

$$[n] = n, \quad [-n] = \beta^p - n$$

となるので

$$[-n] + [n] = \beta^p - 0 = [0]$$

という通常に加減乗除に対応する演算が可能になる。計算機内では数値は主に 2 進法で表現される。2 進数字 1 個を 1 ビット (bit) と呼び, 8 ビットをひとかたまりとして 1 バイト (byte) と呼ぶ。初めに 2 進法 4 ビット長の整数の例を示そう。表現できる整数は $2^4 = 16$ 通りあり, 符号なしの場合 0 を入れてつぎの範囲の有限な整数を表す。

$$0 \leq N \leq 2^4 - 1$$

表 A.1 に 2 進数 4 ビット長整数を 10 進, 16 進, 8 進数と対比させて示す。

つぎに 2 進数 1 バイト (8 ビット) 長の場合を示そう。表現できる整数は $2^8 = 256$ 通りであり, 符号なしの場合には 0 を入れて

$$0 \leq N \leq 2^8 - 1$$

の範囲の整数を表すことができる。符号付きの場合には負の数も入れるため, 正負それぞれ $256/2 = 128 (= 2^7)$ 通りの数が表現できるので, 0 を正の側に入れて

表 A.1 2進法 4 ビット長整数と β 進数表示
($\beta = 2, 10, 16, 8$)

10 進数	2 進数	16 進数	8 進数
0	0000	0	00
1	0001	1	01
2	0010	2	02
3	0011	3	03
4	0100	4	04
5	0101	5	05
6	0110	6	06
7	0111	7	07
8	1000	8	10
9	1001	9	11
10	1010	a	12
11	1011	b	13
12	1100	c	14
13	1101	d	15
14	1110	e	16
15	1111	f	17

$$-2^7 \leq N \leq 2^7 - 1$$

の範囲の整数を表すことができる。先ほども述べたとおり、負の数 $-n$ は、正の数 n の 2 の補数 $2^8 - n$ に対応させて表現される。これは正の数 n の 1 の補数（各ビットを反転したもの）に 1 を加えたもの

$$\underbrace{2^8 - n}_{2 \text{ の補数}} = \underbrace{(2^8 - 1)}_{\substack{\text{全ビットは 1} \\ \text{1 の補数 (ビット反転)}}} - n + 1$$

に等しい。このとき、正の数の先頭ビットは 0、負の数の先頭ビットは 1 となる。この表示方式を使えば、 $[-n] + [n] = 2^8 - 0 = [0]$ となり、通常に加減乗除に対応する演算が可能になる。

例えば $n = 11$ の場合を考えよう。

$$11 = 1 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 1 \times 2^0$$

であるので、 $n = 11$ は 2 進法 8 ビット長ではつぎのように表される。

a)

7	6	5	4	3	2	1	0
0	0	0	0	1	0	1	1

 $n = 11$ の表示

4 A. 計算機における数値表現

$-n$ を得るために、まず各ビットに 1 がたっている数 $2^8 - 1$

b)

7	6	5	4	3	2	1	0
1	1	1	1	1	1	1	1

 $2^8 - 1$ の表示

から a) の数を引く、すなわち a) の各ビットを反転させて 1 の補数を作り

c)

7	6	5	4	3	2	1	0
1	1	1	1	0	1	0	0

 $(2^8 - 1) - n$
: 1 の補数 (ビット反転)

さらに 1 を加える。これが $-n = -11$ の計算機表現 $[-n] = 2^8 - n$ である。

d)

7	6	5	4	3	2	1	0
1	1	1	1	0	1	0	1

 $(2^8 - 1) - n + 1$
 $= 2^8 - n$: 2 の補数

a) と d) , すなわち 11 と - 11 に対応する計算機表現数を加えると

e)

7	6	5	4	3	2	1	0
0	0	0	0	0	0	0	0

 0 の表示

のように 8 ビット表示では零となることが確認できる (9 桁目も有効であれば、 2^8 となるのであるが)。以上のような 1 バイト整数の内部表現を、符号なし/符号付き 10 進数と対比させて表 A.2 に示す。

同様に符号付き 2 バイト整数 (C 言語では short int 型) の内部表現を正負の数に対比させて表 A.3 に示す。2 バイトでは

$$-2^{15} \leq N \leq 2^{15} - 1$$

の範囲の整数が表現できる ($2^{15} = 32\,768$)。負の数 $-n$ は、正の数 n の 2 の補数 ($2^{16} - n = (2^{16} - 1) - n + 1$: ビットを反転させ 1 を加えた数) により表され、先頭ビットは 1 となる。通常用いられる符号付き 4 バイト整数 (最近の C 言語では int 型) では、つぎの範囲の数を扱える ($2^{31} = 2\,147\,483\,648$)。

$$-2^{31} \leq N \leq 2^{31} - 1$$

表 A.2 1 バイト整数の内部表現と符号

符号なし 10 進数	2 進数	16 進数	符号付き 10 進数	
0	00000000	00	0	↓
1	00000001	01	1	正の数
2	00000010	02	2	
3	00000011	03	3	
4	00000100	04	4	
⋮	⋮	⋮	⋮	
127	01111111	7f	127	
128	10000000	80	-128	↓
129	10000001	81	-127	負の数 $-n$ は正の数 n の 2 の補数 (ビット
130	10000010	82	-126	$2^8 - n = (2^8 - 1) - n + 1$
131	10000011	83	-125	を反転させ 1 を加えた数) として表現され,
132	10000100	84	-124	その先頭ビットは 1
⋮	⋮	⋮	⋮	
255	11111111	ff	-1	

表 A.3 符号付 2 バイト整数の内部表現

10 進数	2 進数	16 進数	2 の 補数	10 進数	2 進数	16 進数
0	00000000 00000000	0000				
1	00000000 00000001	0001	↔	-1	11111111 11111111	ffff
2	00000000 00000010	0002	↔	-2	11111111 11111110	fffe
3	00000000 00000011	0003	↔	-3	11111111 11111101	fffd
4	00000000 00000100	0004	↔	-4	11111111 11111100	fffc
5	00000000 00000101	0005	↔	-5	11111111 11111011	fffb
6	00000000 00000110	0006	↔	-6	11111111 11111010	fffa
7	00000000 00000111	0007	↔	-7	11111111 11111001	fff9
8	00000000 00001000	0008	↔	-8	11111111 11111000	fff8
9	00000000 00001001	0009	↔	-9	11111111 11110111	fff7
10	00000000 00001010	000a	↔	-10	11111111 11110110	fff6
⋮	⋮	⋮		⋮	⋮	⋮
32 766	01111111 11111110	7ffe	↔	-32 766	10000000 00000010	8002
32 767	01111111 11111111	7fff	↔	-32 767	10000000 00000001	8001
				-32 768	10000000 00000000	8000

A.2 実数型の数の表現 (浮動小数点数表示)

計算機における“実数”とは、数学的な意味での実数ではなく、有限桁からなる浮動小数点数により表される不連続 (離散的) で有限な数である。例えば 10 進数 $-1\,234.56$ と $0.000\,987\,6$ は、有効桁と 10 のべきとに分けて、それぞれ

$$-1.234\,56 \times 10^3 \qquad 9.876 \times 10^{-4}$$

と書くことができるように、一般に計算機内部における β 進 p 桁の浮動小数点数は、正規化表現

$$\begin{aligned} & \pm(d_0 . d_1 d_2 \cdots d_{p-1})_{\beta\text{進数}} \times \beta^E \\ & = \pm\left(d_0 + \frac{d_1}{\beta} + \frac{d_2}{\beta^2} + \cdots + \frac{d_{p-1}}{\beta^{p-1}}\right) \times \beta^E \end{aligned} \quad (\text{A.2})$$

で表される。ここに d_i は

$$1 \leq d_0 \leq \beta - 1 \quad (\text{A.3})$$

$$0 \leq d_i \leq \beta - 1 \quad (i = 1, \dots, p - 1) \quad (\text{A.4})$$

を満たす整数であり、特に \pm を符号部、 $d_0 . d_1 d_2 \cdots d_{p-1}$ を仮数部、 β を基底、 E を指数、 β^E を指数部と呼ぶ。通常の計算機では β は 2, 8, 10, 16 のいずれかである。桁数 p および指数 E の範囲は計算機により異なる。浮動小数点システムでは表される数が 0 でない限り、正規化表現により最初の数字 d_0 は 0 以外の数である。

パーソナルコンピュータやワークステーションで通常用いられる IEEE (Institute of Electrical and Electronics Engineers) 方式 (IEEE754 規格) では、(1) 符号部の 1 ビット (符号ビット) は 0 を正、1 を負とし、(2) $\beta = 2$ であるため正規化表現ではつねに $d_0 = 1$ となるので、 d_0 は特に内部表現中には記憶されずに用いられ、(3) 指数は 0 か正の値となるようバイアス表現されてい

る。ここでこの指数を $E0$ と表記する。IEEE SP (single precision, 単精度実数型, 図 A.1(a)) では, 符号に 1 ビット, 仮数部に 23 ビット ($d_0 = 1$ を含めると $p = 24$), バイアス表現指数 $E0$ に符号なし 8 ビット (実際の指数は $E = E0 - 127$) の計 32 ビット (4 バイト), IEEE DP (double precision, 倍精度実数型, 図 A.1(b)) では, 符号に 1 ビット, 仮数部に 52 ビット ($d_0 = 1$ を含めると $p = 53$), バイアス表現指数 $E0$ に符号なし 11 ビット (実際の指数は $E = E0 - 1023$) の計 64 ビット (8 バイト) が用いられる。

	31	30	29	...	23	22	21	0
	s	2^7	2^6	...	2^0	2^{-1}	2^{-2}	2^{-23}
符号	指数 $E0$ ($E = E0 - 127$)					仮数部 $(1.d_1d_2 \cdots d_{23})_2$ 進数			

(a) SP(単精度実数型)

	31	30	29	20	19	18	0
	s	2^{10}	2^9	2^0	2^{-1}	2^{-2}	2^{-20}
符号	指数 $E0$ ($E = E0 - 1023$)					仮数部 $(1.d_1d_2 \cdots d_{20})_2$ 進数			

	2^{-21}	2^{-52}
	仮数部 (続き) $(\cdots d_{21}d_{22} \cdots d_{52})_2$ 進数		

(b) DP(倍精度実数型)

図 A.1 不動小数点数の内部表現 (IEEE754 規格)

IEEE 方式の正規化表現における浮動小数点数は, s を符号ビット, 仮数部の 2 進数を $1.d_1d_2 \cdots$ (10 進数表現では $1 + d_1/2 + d_2/2^2 + \cdots$), バイアス表現された指数を $E0$ とすると, つぎのように書ける。

$$\text{SP: } x^{sp} = (-1)^s \times \left(1 + \frac{d_1}{2} + \frac{d_2}{2^2} + \cdots + \frac{d_{23}}{2^{23}} \right) \times (2^{E0-127}) \quad (\text{A.5})$$

$$\text{DP: } x^{dp} = (-1)^s \times \left(1 + \frac{d_1}{2} + \frac{d_2}{2^2} + \cdots + \frac{d_{53}}{2^{53}} \right) \times (2^{E0-1023}) \quad (\text{A.6})$$

ある計算機で表すことのできる浮動小数点数の絶対値の最大値をオーバーフローレベルといい F_{\max} で表し, 0 でない最小値をアンダーフローレベルといい F_{\min} で表す。 F_{\max} よりも大きい数が生ずるとオーバーフローとなり, “無限大” を

表す数で表現される。“無限大”においては、バイアス表現指数 E_0 のすべてのビットに 1 がたち、かつ仮数部のすべてのビットに 0 がたっている。符号ビットにより正負 2 種類の無限大が存在する。 E_0 の全ビットに 1 がたっているも仮数部の全ビットが 0 ではない“数”は無効演算や未決定の数値を表し、非数 NaN (Not-a-Number) として信号処理される。 F_{\min} よりも絶対値の小さい数 (0 は除く) が生ずるとアンダーフローとなるが、この場合には通常その数を 0 で置き換えて計算は続行される。なお 0 は仮数 = 0, バイアス表現指数 $E_0 = 0$ で表される。やはり正負 2 種類の 0 が存在することになる。

ここで IEEE SP 形式で表現できる実数の範囲を概算してみよう。仮数部は 24 桁の 2 進数 $d_0d_1d_2 \cdots d_{23}$ で表されるので

$$2^{24} \approx 1.7 \times 10^7$$

より有効桁数は 10 進数で約 7 桁であることがわかる。つぎに指数部に移ろう。 E_0 が符号なし 8 ビット整数 ($0 \leq E_0 \leq 255$) であることから指数 $E = E_0 - 127$ は $-127 \leq E \leq 128$ の範囲をとりうるが、実際の範囲には $-126 \leq E \leq 127$ が用いられる。これはすでに述べたように $E_0 = 255$ のときは無限大あるいは非数を表現するためであり、後で説明するように $E_0 = 0$ のときは非正規化数を許容するためである。

F_{\max} は仮数部の全ビットが 1 で指数が最大値をとるときの値であるので、

$$\begin{aligned} F_{\max}^{sp} &= (2^0 + 2^{-1} + 2^{-2} + \cdots + 2^{-23}) \times 2^{127} \\ &= (2 - 2^{-23}) \times 2^{127} \approx 2^{128} \approx 3.4 \times 10^{38} \end{aligned}$$

F_{\min} は仮数部の最初のビットが 1 で残りのビットは 0, 指数は最小値をとるときの値であるので

$$F_{\min}^{sp} = 1.0 \times 2^{-126} \approx 1.2 \times 10^{-38}$$

となる。1 章の図 1.1 における浮動小数点数のミニ模型で示したように、正規化された浮動小数点数は、その分布は 0 付近で密から急激に疎になる。正規化

表現による最小数よりさらに小さな数を表すために、IEEE 方式では $E_0 = 0$ のときはつぎのような非正規化数 ($d_0 = 0$) とする。

$$x^{sp} = (-1)^s \times \left(0 + \frac{d_1}{2} + \frac{d_2}{2^2} + \cdots + \frac{d_{23}}{2^{23}} \right) \times 2^{-126}$$

この表現による絶対値の最小値は仮数部の最下位ビットのみが立つ場合である。

$$F_{\min}^{sp} = 2^{-23} \times 2^{-126} = 2^{-149} \approx 1.4 \times 10^{-45}$$

例題 A.1 IEEE DP 形式で表される数の有効桁数，オーバーフローレベル F_{\max}^{dp} ，アンダーフローレベル F_{\min}^{dp} を概算せよ。

【解答】 仮数部は 53 桁の 2 進数で表わされるので、 $2^{53} \approx 0.9 \times 10^{16}$ より有効桁数は 10 進数で約 16 桁である。 E_0 が符号なし 11 ビット整数 ($0 \leq E_0 \leq 2^{11} - 1 = 2047$) であることから指数は $-1023 \leq E_0 - 1023 \leq 1024$ の範囲をとりうるが、実際には $-1022 \leq E \leq 1023$ が用いられる。これより正規化数では

$$\begin{aligned} F_{\max}^{dp} &= (2^0 + 2^{-1} + 2^{-2} + \cdots + 2^{-52}) \times 2^{1023} \\ &= (2 - 2^{-52}) \times 2^{1023} \approx 2^{1024} \\ &\approx 1.8 \times 10^{308} \\ F_{\min}^{dp} &= 1.0 \times 2^{-1022} \approx 2.2 \times 10^{-308} \end{aligned}$$

となる。非正規化数では F_{\min}^{dp} はつぎのように概算される。

$$F_{\min}^{dp} = 2^{-52} \times 2^{-1022} = 2^{-1074} \approx 4.9 \times 10^{-324}$$

◇

引用・参考文献

- 1) Heath, Michael T. : Scientific Computing, An Introductory Survey, McGraw-Hill (2002)
- 2) 平林雅英 : ANSI C/C++ 辞典, 共立出版 (1998)
- 3) 川崎晴久 : C&FORTRAN による数値解析の基礎, 共立出版 (1993)

B | ベクトルと行列のノルム

本来連続性をもつ関数やその微分・積分などでも，計算機では離散的な量に近似されて扱われる。これら連続性をもつ量に関する方程式は，離散化されると多くの場合連立1次方程式 $Ax = b$ に帰着される。また連立非線形方程式や連立常微分方程式を解くにあたって，解はベクトルとして扱われる。したがって数値解析においてはベクトルや行列の取扱いが不可欠である。

多元連立方程式の解法の収束性や解の誤差などを論ずるとき，ベクトルや行列の「大きさ」を表す概念，あるいは二つのベクトル間の「距離」の概念としてノルムという実数のスカラー量が必要となる。ここではベクトルと行列のノルムに関する諸概念を与える。

B.1 ベクトルのノルム

B.1.1 ノルムの定義

n 個の複素数 x_1, x_2, \dots, x_n を成分とする n 次元ベクトル

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

の集合は線形ベクトル空間になっている；すなわちベクトル $\boldsymbol{x}, \boldsymbol{y}$ とスカラー α に対して， $\boldsymbol{x} + \boldsymbol{y}$ と $\alpha\boldsymbol{x}$ はまたベクトルであり，加算と定数倍が定義できる。

線形ベクトル空間 X の要素 x の関数でつぎの条件 i) ~ iii) を満足するものをノルムといい、 $\|x\|$ と表す。

- i) すべての $x \in X$ に対して $\|x\| \geq 0$
 (等号は $x = 0$ のときに限り成立)
- ii) すべての $x \in X$ とすべての複素数 α に対して $\|\alpha x\| = |\alpha| \|x\|$
- iii) すべての $x, y \in X$ に対して $\|x + y\| \leq \|x\| + \|y\|$
- ベクトル空間のノルムの例としてはつぎのようなものがある。

$$\|x\|_1 = \sum_{i=1}^n |x_i| \quad (\text{B.1})$$

$$\|x\|_2 = \sqrt{\sum_{i=1}^n |x_i|^2} \quad (\text{B.2})$$

$$\|x\|_\infty = \max_i |x_i| \quad (\text{B.3})$$

上記の式 (B.1), (B.2), (B.3) は、 p ノルムと呼ばれる形

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p} \quad (p \geq 1) \quad (\text{B.4})$$

の特別な場合 ($p = 1, 2, \infty$) である。

問題 1 $p \rightarrow \infty$ のときの p ノルム (B.4) が式 (B.3) の形で表されることを確かめよ

p ノルムがノルムの条件 i), ii) を満たすことはすぐにわかる。条件 iii) を満たすことはミンコフスキー (Minkowski) の不等式

$$\left(\sum_{v=1}^n |a_v + b_v|^p \right)^{1/p} \leq \left(\sum_{v=1}^n |a_v|^p \right)^{1/p} + \left(\sum_{v=1}^n |b_v|^p \right)^{1/p} \quad (\text{B.5})$$

を用いて示される。ただし $p \geq 1$ である。

ノルムが定義されるとき、 n 次元線形ベクトル空間は距離づけられて n 次元ノルム空間になる。

B.1.2 エルミート行列とベクトルの内積について

行列 $A = [a_{ij}]$ の成分がすべて実数の場合, A の成分の行と列をすべて入れ換えた行列を A の転置行列といい, $A^T = [a_{ij}^T]$ で表す。成分表示では $a_{ij}^T = a_{ji}$ である。 $A = A^T$ であるとき A を対称行列と呼ぶ。 $A^T A = A A^T = I$ を満足する行列を直交行列と呼ぶ。行列 $A = [a_{ij}]$ の成分が複素数の場合, A の成分の行と列をすべて入れ換え各成分をその複素共役な値で置き換えた行列を A の転置共役行列といい, $A^* = [a_{ij}^*]$ で表す。すなわち $A^* = \bar{A}^T$, 成分表示では $a_{ij}^* = \bar{a}_{ji}$ である。 $A = A^*$ を満足する行列 A をエルミート (Hermite) 行列, $A^* A = A A^* = I$ を満足する行列 A をユニタリ (Unitary) 行列という。特に A の成分がすべて実数のとき, 転置共役行列 A^* は転置行列 A^T , エルミート行列は対称行列, ユニタリ行列は直交行列となる。

同様に列ベクトル $\boldsymbol{x} = [x_i]^T$, $\boldsymbol{y} = [y_i]^T$ についても n 行 1 列の行列とみなせば, その転置共役ベクトルは行ベクトルとなり $\boldsymbol{x}^* = [x_i^*]$, $\boldsymbol{y}^* = [y_i^*]$ で表す。 n 次元ベクトルの集合からなるベクトル空間にノルムを導入すると, 距離づけされて n 次元ノルム空間になる。同じ n 次元空間の点の集合であってもノルムの定義が異なれば別のノルム空間になる。ノルムが $\|\boldsymbol{x}\|_2$ で定義されているベクトル空間は, 成分が複素数の場合ユニタリ空間, 成分が実数の場合ユークリッド (Euclid) 空間という。

ベクトル \boldsymbol{x} , \boldsymbol{y} の内積は, 複素数ベクトルからなるユニタリ空間においては

$$(\boldsymbol{x} \boldsymbol{y}) = \boldsymbol{y}^* \boldsymbol{x} = \sum_i x_i \bar{y}_i$$

と定義され, 実数ベクトルからなるユークリッド空間においては

$$(\boldsymbol{x} \boldsymbol{y}) = \boldsymbol{y}^T \boldsymbol{x} = \boldsymbol{x}^T \boldsymbol{y} = \sum_i x_i y_i$$

のように表される。このとき次式が成立する。

$$(\boldsymbol{x} \boldsymbol{x}) = \|\boldsymbol{x}\|_2^2$$

エルミート行列 A に対し

$$A\{\mathbf{x}\} \equiv \mathbf{x}^* A \mathbf{x} = (A \mathbf{x}, \mathbf{x}) = (\mathbf{x}, A \mathbf{x}) = \sum_i \sum_j a_{ij} x_i \bar{x}_j$$

をエルミート形式という。任意の複素ベクトル \mathbf{x} に対して $A\{\mathbf{x}\}$ は実数値をとる。実際 $A\{\mathbf{x}\}$ の共役複素数は $A\{\mathbf{x}\}$ に等しいことは以下のように容易に確かめられる。

$$\bar{A}\{\mathbf{x}\} = \sum_i \sum_j \bar{a}_{ij} \bar{x}_i x_j = \sum_i \sum_j a_{ji} \bar{x}_i x_j = A\{\mathbf{x}\}$$

$A\{\mathbf{x}\}$ は、任意の \mathbf{x} に対し $A\{\mathbf{x}\} \geq 0$ なるとき非負値、さらに任意の $\mathbf{x} \neq 0$ に対し $A\{\mathbf{x}\} > 0$ なるとき正值といい、それぞれ $A \geq 0, A > 0$ で表す。任意の正方行列 P に対し $B = P^* P$ とおくと

$$B^* = (P^* P)^* = P^* P^{**} = P^* P = B$$

より $B = P^* P$ はエルミート行列となり、そのエルミート形式は

$$B\{\mathbf{x}\} = \mathbf{x}^* P^* P \mathbf{x} = (P \mathbf{x}, P \mathbf{x}) \geq 0$$

となる。 P が正則であるならば、任意の $\mathbf{x} \neq 0$ に対し $P \mathbf{x} \neq 0$ 、したがって $B\{\mathbf{x}\} > 0$ となるので、 $B = P^* P$ は正值エルミート行列となる。

B.1.3 新しいノルムの作り方

ノルム $\|\mathbf{x}\|_p$ においては、ベクトル \mathbf{x} のすべての成分は等しく重みづけられている。成分の非対称な関数であるようなノルムを創るためには、与えられたノルム $\|\mathbf{x}\|$ と正則な正方行列 T を用いて、 $\|T \mathbf{x}\|$ とすればよい。これがノルムであることの検証はやさしい。

たとえばユニタリ空間においては、内積の記法を用いるとノルム $\|\mathbf{x}\|_2$ は

$$\|\mathbf{x}\|_2 = \sqrt{(\mathbf{x}, \mathbf{x})} = \sqrt{\mathbf{x}^* \mathbf{x}} \quad (\text{B.6})$$

であり、新しいノルムを

$$\|\mathbf{x}\| = \|T \mathbf{x}\|_2 = \sqrt{\mathbf{x}^* T^* T \mathbf{x}} = \sqrt{\mathbf{x}^* B \mathbf{x}} \quad (\text{B.7})$$

で定義することができる。ここに $B = T^*T$ は正値エルミート行列 (任意の $x \neq 0$ に対し $x^*Bx > 0$) である。関係式 (B.7) はノルムの条件 i) ~ iii) を満たしている。

B.1.4 有限次元ベクトルノルムの同値性

数値解析ではベクトル列の収束を論じるとき、いかなるノルムに関する収束であるかは問題にしないことが多い。それは以下のように、有限次元ベクトル空間におけるノルムの同値性が成立するからである。

定理 B.1

任意の二つのノルム $\|x\|_\alpha$ と $\|x\|_\beta$ とは、つぎの意味で同値である。

$$m\|x\|_\alpha \leq \|x\|_\beta \leq M\|x\|_\alpha \quad (\text{B.8})$$

がすべての $x \in X$ に対して成立するような正の数 m, M が存在する。

証明 まずベクトルノルムの連続性を示す。任意のベクトル x, δ に対して

$$\|x + \delta\| \leq \|x\| + \|\delta\|$$

$$\|x\| = \|x + \delta - \delta\| \leq \|x + \delta\| + \|-\delta\| = \|x + \delta\| + \|\delta\|$$

が成立するので

$$-\|\delta\| \leq \|x + \delta\| - \|x\| \leq \|\delta\| \implies \left| \|x + \delta\| - \|x\| \right| \leq \|\delta\|$$

が得られる。 e_i を第 i 成分のみが 1 で他は 0 である単位ベクトルとすれば

$$\delta = \sum_{i=1}^n \delta_i e_i$$

と書けるので

$$\|\delta\| \leq \sum_{i=1}^n \|\delta_i e_i\| = \sum_{i=1}^n |\delta_i| \|e_i\| \leq \max_i |\delta_i| \sum_{i=1}^n \|e_i\| = M \|\delta\|_\infty$$

となる。ここで $M = \sum_{i=1}^n \|e_i\|$ 。したがって任意の ϵ が与えられたとき、 $\|\delta\|_\infty \leq \epsilon/M$ を満たす δ をとれば

$$\| \mathbf{x} + \boldsymbol{\delta} \| - \| \mathbf{x} \| \leq \epsilon$$

が成立する。これは $\| \mathbf{x} \|$ が x_1, x_2, \dots, x_n の連続関数であることを示している。

つぎに任意のノルム $\| \mathbf{x} \|_\alpha$ が $\| \mathbf{x} \|_\infty$ と同値であることを示す。 n 次元空間内の有界閉集合 $S \equiv \{ \mathbf{y} \mid \| \mathbf{y} \|_\infty = 1 \}$ の上で、 $\mathbf{y} \in S$ に対して $\| \mathbf{y} \|_\alpha$ はある点 \mathbf{y}_0 で最小、ある点 \mathbf{y}_1 で最大となる（ノルムの連続性より）。

$$\| \mathbf{y}_0 \|_\alpha \leq \| \mathbf{y} \|_\alpha \leq \| \mathbf{y}_1 \|_\alpha < \infty$$

$\mathbf{y}_0 \neq \mathbf{0}$ であるから、 $\| \mathbf{y}_0 \|_\alpha > 0$ である。これから任意の $\mathbf{x} \neq \mathbf{0}$ に対して、 $\mathbf{x} / \| \mathbf{x} \|_\infty \in S$ は

$$\| \mathbf{y}_0 \|_\alpha \leq \frac{\| \mathbf{x} \|_\alpha}{\| \mathbf{x} \|_\infty} \leq \| \mathbf{y}_1 \|_\alpha$$

を満たす。すなわち

$$m_\alpha \| \mathbf{x} \|_\infty \leq \| \mathbf{x} \|_\alpha \leq M_\alpha \| \mathbf{x} \|_\infty$$

ただし、 $m_\alpha = \| \mathbf{y}_0 \|_\alpha$ 、 $M_\alpha = \| \mathbf{y}_1 \|_\alpha$ 。これはベクトル列 $\{ \mathbf{x}^{(k)} \}$ が ∞ ノルムについて 0 に収束するならば、任意のノルムについても 0 に収束することを示している。

これより任意の 2 つのノルムに関して

$$m_\alpha \| \mathbf{x} \|_\infty \leq \| \mathbf{x} \|_\alpha \leq M_\alpha \| \mathbf{x} \|_\infty$$

$$m_\beta \| \mathbf{x} \|_\infty \leq \| \mathbf{x} \|_\beta \leq M_\beta \| \mathbf{x} \|_\infty$$

なる $m_\alpha, M_\alpha > 0$ および $m_\beta, M_\beta > 0$ が存在することとなり、 $m = m_\beta / M_\alpha$ 、 $M = M_\beta / m_\alpha$ とすれば

$$m \| \mathbf{x} \|_\alpha \leq \| \mathbf{x} \|_\beta \leq M \| \mathbf{x} \|_\alpha$$

が導出され、証明は完了する。 ♠

これはベクトル列 $\{ \mathbf{x}^{(k)} \}$ がある特定のノルムについて 0 に収束するならば、どのようなノルムについても 0 に収束することを示している。特に $\| \mathbf{x} \|_\infty$ による収束を見ると、それはすべての成分が 0 に収束することを意味する。結局、 $\mathbf{x}^{(k)} = \mathbf{z}^{(k)} - \mathbf{z}_0$ なるベクトル列 $\{ \mathbf{x}^{(k)} \}$ が 0 に収束するというときには、ノルムの定義のいかんにかかわらず $\mathbf{z}^{(k)}$ は成分ごとに \mathbf{z}_0 に収束しているわけである。このため、有限次元のベクトルのノルムはたがいに同値であるという。

B.2 行列のノルム

行列のノルムを定義する方法はいくつかある。簡単に n 次の行列は n^2 次元の一つの線形ベクトル空間になっていると考え、正方形行列 $A = (a_{ij})$ に対し以下のようなノルムが導かれる。

$$\|A\|'_1 = \sum_{i=1, j=1}^n |a_{ij}| \quad (\text{B.9})$$

$$\|A\|'_2 = \sqrt{\sum_{i=1, j=1}^n |a_{ij}|^2} \quad (\text{B.10})$$

特に式 (B.10) で定義されるノルムをユークリッド (Euclid) ノルムといい $\|A\|_E$ と表記する。しかしながらベクトルノルムがすでに定義されている場合には、以下の定義のほうが自然であろう。

B.2.1 ナチュラルノルム

ベクトルノルムに基づいて正方形行列 A のノルムをつぎのように定義する。

$$\|A\| = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} \quad (\text{B.11})$$

ここで \sup は上限 (supremum)、すなわち最小上界を意味するが、有限次元の場合 \sup は \max 、すなわち最大値としてよい。上式で定義される行列のノルムをナチュラルノルムと呼ぶ。 A, B を行列とすると、ナチュラルノルムはつぎの性質をもつ。

- (1) $\|A\mathbf{x}\| \leq \|A\|\|\mathbf{x}\|$
- (2) $\|AB\| \leq \|A\|\|B\|$
- (3) すべての A に対して $\|A\| \geq 0$ 。ただし等号は $A = 0$ のときに限り成立。
- (4) $\|\alpha A\| = |\alpha|\|A\|$
- (5) $\|A + B\| \leq \|A\| + \|B\|$

特に性質 (3), (4), (5) は, 行列のノルムが前節で記したベクトルノルムの条件 i), ii), iii) に対応していることに注意しよう。

問題 2 行列ノルムに関する性質 (1) ~ (5) を証明せよ。

単位行列を I で表すと, $Ix = x$ であるから, 明らかに次式が成立する。

$$\|I\| = 1 \quad (\text{B.12})$$

行列 A の 1 ノルム, ∞ ノルムは以下のようにになる。2 ノルムの具体形は次項で示される。

$$\|A\|_1 = \sup_{x \neq 0} \frac{\|Ax\|_1}{\|x\|_1} = \max_j \sum_i |a_{ij}| \quad (\text{B.13})$$

$$\|A\|_\infty = \sup_{x \neq 0} \frac{\|Ax\|_\infty}{\|x\|_\infty} = \max_i \sum_j |a_{ij}| \quad (\text{B.14})$$

問題 3 ナチュラルノルムの具体形: 1 ノルム (B.13) と ∞ ノルム (B.14) を証明せよ。

B.2.2 固有値と固有ベクトル

$n \times n$ 正方行列 A を作用させてもベクトル $x \neq 0$ の方向が変わらないとき, すなわち

$$Ax = \lambda x \quad (\text{B.15})$$

が成立するとき, λ を A の固有値, x を固有値 λ に対する固有ベクトルという。上式は x に関する斉 1 次方程式

$$(A - \lambda I)x = 0$$

の形に書ける。それが $x = 0$ 以外の解をもつためには, 行列 $A - \lambda I$ は正則であってはならず, その行列式が零となること

$$|A - \lambda I| = 0$$

が必要十分である。このとき固有値 λ は特性方程式

$$p(\lambda) = |\lambda I - A| = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix} = 0 \quad (\text{B.16})$$

の根で与えられる。この特性方程式は λ の n 次多項式 (特性多項式という)

$$p(\lambda) = \lambda^n - p_1\lambda^{n-1} + p_2\lambda^{n-2} - \cdots + (-1)^n p_n \quad (\text{B.17})$$

で表されるので, その n 個の根は, 行列 A の要素がすべて実数であっても複素数を含むことがある。 λ^{n-1} の係数および定数項を比較すれば, つぎの関係

$$p_1 = \text{tr } A = \sum_{i=1}^n a_{ii} \quad (\text{B.18})$$

$$p_n = |A| \quad (\text{B.19})$$

を得る。 $\text{tr } A$ を A の対角和 (trace) という。

任意の正則行列を P とするとき, 行列 A に対する相似変換

$$\tilde{A} = P^{-1}AP \quad (\text{B.20})$$

に対して, 特性多項式は不変である。すなわち

$$\begin{aligned} |\lambda I - \tilde{A}| &= |\lambda I - P^{-1}AP| = |P^{-1}(\lambda I - A)P| \\ &= |P^{-1}||P||\lambda I - A| = |\lambda I - A| \end{aligned}$$

したがって A と \tilde{A} の固有値は等しく, また対角和も等しい。対角行列, 上三角行列 ($i > j$ に対し $a_{ij} = 0$) および下三角行列 ($i < j$ に対し $a_{ij} = 0$) の対角成分はその行列の固有値となっている。相似変換は固有値を保存するので, このような固有値を求めやすい行列に相似変換することが有効な場合がある。

行列 A の n 個の固有値を λ_i ($i = 1, \dots, n$) とする。 $|\lambda_i|$ の最大値を行列 A のスペクトル半径といい、 $\rho(A)$ と記す。

$$\rho(A) \equiv \max_i |\lambda_i| \quad (\text{B.21})$$

ナチュラルノルムの 2 ノルムは、スペクトル半径を用いて次式により表される。

$$\|A\|_2 = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \sqrt{\rho(A^*A)} \quad (\text{B.22})$$

特に A がエルミート行列ならば

$$\|A\|_2 = \rho(A) \quad (\text{B.23})$$

問題 4 ナチュラルノルムの具体形：2 ノルム (B.22) および (B.23) を証明せよ。

2 ノルム $\|A\|_2 = \sqrt{\rho(A^*A)}$ を求めるとき必要となる行列の固有値の計算は、通常かなりやっかいである。しかしながら以下の定理のように、より計算しやすいユークリッドノルム $\|A\|_E = \left(\sum_{i=1}^n \sum_{j=1}^n |a_{ij}|^2 \right)^{1/2}$ を用いた式で 2 ノルム $\|A\|_2$ を評価することはできる。

定理 B.2

すべての $n \times n$ 正方行列に対して次式が成立する。

$$\frac{1}{\sqrt{n}} \|A\|_E \leq \|A\|_2 \leq \|A\|_E \quad (\text{B.24})$$

証明 $B = A^*A$ とおくと、 B の (i, k) 成分は $b_{ik} = \sum_{j=1}^n \bar{a}_{ji} a_{jk}$ となるので

$$\text{tr } B = \sum_{i=1}^n b_{ii} = \sum_{i=1}^n \sum_{j=1}^n \bar{a}_{ji} a_{ji} = \sum_{i=1}^n \sum_{j=1}^n |a_{ji}|^2 = \|A\|_E^2 \quad (\text{B.25})$$

一方、 B はエルミート行列であるから、適当なユニタリ行列を用いた相似変換に

より, B の固有値 $\mu_i \geq 0$ を対角成分にもつ対角行列に変換できる (本付録巻末の問題 4 解答を参照のこと)。相似変換により対角和は不変であるから

$$\operatorname{tr} B = \sum_{i=1}^n \mu_i \quad (\text{B.26})$$

であり, したがって

$$\max_i \mu_i \leq \operatorname{tr} B \leq n \max_i \mu_i \quad (\text{B.27})$$

が成り立つ。これより

$$\|A\|_2^2 \leq \|A\|_F^2 \leq n \|A\|_2^2 \quad (\text{B.28})$$

を得るので, $\|A\|_2$ の評価式 (B.24) が導かれる。♠

B.2.3 行列のノルムとスペクトル半径

定理 B.3 (ナチュラルノルムの評価 I)

正方行列 A が与えられたとき, いかなるナチュラルノルムに対しても次式が成立する。

$$\|A\| \geq \rho(A) \quad (\text{B.29})$$

証明 行列 A の固有値を λ_i とし, λ_i に対する固有ベクトルを \boldsymbol{x}_i とすれば

$$A\boldsymbol{x}_i = \lambda_i \boldsymbol{x}_i \quad (\exists \boldsymbol{x}_i \neq \mathbf{0})$$

が成り立つ。上式に対してノルムをとれば

$$\|A\boldsymbol{x}_i\| = \|\lambda_i \boldsymbol{x}_i\| = |\lambda_i| \|\boldsymbol{x}_i\|$$

$$\therefore \frac{\|A\boldsymbol{x}_i\|}{\|\boldsymbol{x}_i\|} = |\lambda_i|$$

となるので, これより

$$\|A\| = \sup_{\boldsymbol{x} \neq \mathbf{0}} \frac{\|A\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \geq |\lambda_i| \quad (\text{B.30})$$

が導かれる。スペクトル半径 (B.21) を用いれば, 次式を得る。

$$\|A\| \geq \rho(A)$$



しかしながらつぎの定理が成り立つ。

定理 B.4 (ナチュラルノルムの評価 II)

正方行列 A が与えられたとき, 任意の $\epsilon > 0$ に対して

$$\|A\|_\alpha \leq \rho(A) + \epsilon \quad (\text{B.31})$$

が成り立つようなナチュラルノルム $\|A\|_\alpha$ を見出すことができる。

証明 上の条件を満たす一つのノルムを具体的に構成してみよう。行列 A が与えられたとき, これを以下に示すようなジョルダンの標準形 A_1 に相似変換するある正則行列 P が存在する。

$$A_1 = P^{-1}AP = \begin{bmatrix} \lambda_1 & \alpha_1 & & & 0 \\ & \lambda_2 & \alpha_2 & & \\ & & \ddots & \ddots & \\ & 0 & & \ddots & \alpha_{n-1} \\ & & & & \lambda_n \end{bmatrix} \quad (\text{B.32})$$

ここに λ_i は A_1 の固有値となっているが, 相似変換に関して固有値は不変であるため A の固有値でもある。 α_j は 0 または 1 である。いま δ を十分小さい正の数としてつぎのような対角行列 S を考える。

$$S = \begin{bmatrix} 1 & & & & 0 \\ & \delta & & & \\ & & \delta^2 & & \\ & 0 & & \ddots & \\ & & & & \delta^{n-1} \end{bmatrix} \quad (\text{B.33})$$

このとき変換 $S^{-1}A_1S$ を行うと

$$\begin{aligned} A_2 &= S^{-1}A_1S = (PS)^{-1}A(PS) \\ &= \begin{bmatrix} \lambda_1 & \delta\alpha_1 & & & 0 \\ & \lambda_2 & \delta\alpha_2 & & \\ & & \ddots & \ddots & \\ & 0 & & \ddots & \delta\alpha_{n-1} \\ & & & & \lambda_n \end{bmatrix} = D + G \end{aligned} \quad (\text{B.34})$$

となる。ここに

$$D = \begin{bmatrix} \lambda_1 & & & & \mathbf{0} \\ & \lambda_2 & & & \\ & & \ddots & & \\ & & & \ddots & \\ \mathbf{0} & & & & \lambda_n \end{bmatrix}, \quad G = \begin{bmatrix} 0 & \delta\alpha_1 & & & \mathbf{0} \\ & 0 & \delta\alpha_2 & & \\ & & & \ddots & \\ & & & & \ddots \\ \mathbf{0} & & & & \delta\alpha_{n-1} \\ & & & & & 0 \end{bmatrix} \quad (\text{B.35})$$

であり, δ を小さくとれば G の成分は任意に小さくすることができる。

ここで新たにベクトルのノルムを

$$\|\mathbf{x}\|_\alpha = \|(PS)^{-1}\mathbf{x}\|_2 \quad (\text{B.36})$$

により定義する (式 (B.7) 参照のこと), \mathbf{x} の代わりに

$$\mathbf{y} = (PS)^{-1}\mathbf{x} \quad (\text{B.37})$$

とおくと

$$\|\mathbf{x}\|_\alpha = \|\mathbf{y}\|_2 = (\mathbf{y}^*\mathbf{y})^{1/2}$$

であり, また $(PS)^{-1}A = A_2(PS)^{-1}$ より

$$\|A\mathbf{x}\|_\alpha = \|(PS)^{-1}A\mathbf{x}\|_2 = \|A_2\mathbf{y}\|_2 = (\mathbf{y}^*A_2^*A_2\mathbf{y})^{1/2}$$

となる。ここで式 (B.34) より

$$A_2^*A_2 = (D + G)^*(D + G) = D^*D + H$$

とおくと, H の成分は δ を十分小さくすれば任意に小さくすることができるので

$$|\mathbf{y}^*H\mathbf{y}| \leq n^2\delta'\mathbf{y}^*\mathbf{y} = \delta''\mathbf{y}^*\mathbf{y}$$

となる。ここで δ'' は任意に小さくとれるので, ϵ を任意に小さくできる数として

$$\begin{aligned} \frac{\|A\mathbf{x}\|_\alpha}{\|\mathbf{x}\|_\alpha} &= \left(\frac{\mathbf{y}^*D^*D\mathbf{y} + \mathbf{y}^*H\mathbf{y}}{\mathbf{y}^*\mathbf{y}} \right)^{1/2} \\ &\leq (\max_i |\lambda_i|^2 + \delta'')^{1/2} \leq \max_i |\lambda_i| + \epsilon \end{aligned}$$

とかけると、以上からつぎの結論を得る。

$$\sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|A\mathbf{x}\|_\alpha}{\|\mathbf{x}\|_\alpha} \leq \rho(A) + \epsilon$$

ナチュラルノルムは以下のようにとればよい。

$$\|A\|_\alpha = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|A\mathbf{x}\|_\alpha}{\|\mathbf{x}\|_\alpha} = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|(PS)^{-1}A\mathbf{x}\|_2}{\|(PS)^{-1}\mathbf{x}\|_2} \quad (\text{B.38})$$



以上よりスペクトル半径は行列ノルムの評価において重要な役割をもつことがわかる。

引用・参考文献

- 1) Fritz, John : Lectures on Advanced Numerical Analysis, Science Publishers, Inc. (1967); 藤田 宏・名取 亮 共訳 : 数値解析講義, 産業図書 (1975)
- 2) 森 正武 : 数値解析, 共立出版 (1973)
- 3) 高木貞治 : 代数学講義, 共立出版 (1980)
- 4) 佐武一郎 : 線形代数学, 裳華房 (1974)

C

不動点反復法の収束に関する考察

3章では非線形方程式の解法としてニュートン法などの反復法を扱い、4章では連立1次方程式の解法としてヤコビ法、ガウス・ザイデル法、SOR法などの反復法を紹介した。これらの反復法における収束性の議論は、不動点反復法と呼ばれる逐次解法に関連づけることができる。ここでは、まず1変数の場合の不動点反復法に関する収束性の議論を行ってから、これを多変数の連立方程式系に拡張する。

C.1 1変数スカラー方程式

関数 $\varphi(x)$ は閉区間 $K = [a, b]$ 上で定義されており、 K 上で連続で微分可能であるとする。また $x \in K$ ならば $\varphi(x) \in K$ 、すなわち φ により写像しても区間 $[a, b]$ の外に出ないものと仮定する。

$$x = \varphi(x) \tag{C.1}$$

を満足する x を不動点 (fixed point) という。 φ を作用させても x は不動であるからである。このため、式 (C.1) は不動点問題 (fixed-point problem) と呼ばれる。上式に基づく逐次アルゴリズム

$$x^{(k+1)} = \varphi(x^{(k)}) \tag{C.2}$$

を不動点反復法 (fixed-point iteration) という。ここで1変数のスカラー方程式

$$f(x) = 0 \tag{C.3}$$

の根を求めるための反復法と関連づけてみよう。その場合は関数 $\varphi(x)$ を

$$\varphi(x) = x - r(x)f(x) \quad (\text{C.4})$$

とおけばよい。 $r(x) \neq 0$ のとき方程式 $f(x) = 0$ と不動点問題 (C.1) は同値であるので、不動点反復法 (C.2) を用いることができる。 $f(x) = 0$ は通常非線形方程式を考えるが、1 次方程式であってもよい。非線形方程式の場合 $r(x)$ の特徴づけによりいろいろなスキームが考えられるが、特にニュートン法は、関数 $\varphi(x)$ を

$$\varphi(x) = x - \frac{f(x)}{f'(x)} \quad (\text{C.5})$$

と定義 (このとき $r(x) = 1/f'(x)$) したときの不動点反復法 (C.2) と同等である (3 章 参照)。

C.1.1 収束条件

漸化式 (C.2) と平均値の定理より、 $k \geq 1$ に対して

$$\begin{aligned} x^{(k+1)} - x^{(k)} &= \varphi(x^{(k)}) - \varphi(x^{(k-1)}) \\ &= \varphi'(\theta_k) (x^{(k)} - x^{(k-1)}) \end{aligned} \quad (\text{C.6})$$

なる θ_k が $x^{(k)}$ と $x^{(k-1)}$ の間に存在するので、 $K = [a, b]$ 内の $|\varphi'(x)|$ の上限を q 、すなわち

$$|\varphi'(x)| \leq q \quad (\text{C.7})$$

とすれば

$$|x^{(k+1)} - x^{(k)}| \leq q|x^{(k)} - x^{(k-1)}| \quad (\text{C.8})$$

が成り立つ。この関係を繰り返し用いると

$$|x^{(k+1)} - x^{(k)}| \leq q|x^{(k)} - x^{(k-1)}| \leq \dots \leq q^k|x^{(1)} - x^{(0)}| \quad (\text{C.9})$$

が得られるが, $\{x^{(k)}\}$ が収束するためには $\lim_{k \rightarrow \infty} (x^{(k+1)} - x^{(k)}) = 0$ でなければならない。これより収束条件が以下のように得られる。

$$|\varphi'(x)| \leq q < 1 \quad (\text{C.10})$$

と, 数学の素人である私達はこれで終わったような気になる。しかしながら数学では, (1) 数列は収束するか否か, (2) 解が考えている区間 K 上に存在するか否か, が問題となる。

まず収束性から始めよう。『数列 $\{x^{(k)}\}$ が極限值 x に収束する』とは, k が限りなく増大するとき $x^{(k)}$ が一定の数 x に限りなく近づくことをいい, 記号では $\lim_{k \rightarrow \infty} x^{(k)} = x$, あるいは「 $k \rightarrow \infty$ のとき $x^{(k)} \rightarrow x$ 」と書く。数学的に記述すると, 任意の正数 ϵ が与えられたとき, それに対応して一つの番号 N が

$$k > N \quad \text{なるとき} \quad |x - x^{(k)}| < \epsilon$$

なるよう定められるのである。数列 $\{x^{(k)}\}$ が収束するための必要十分条件は, 任意の正数 $\epsilon > 0$ に対して十分大きな N をとれば

$$m, n \geq N \quad \text{なるとき} \quad |x^{(m)} - x^{(n)}| < \epsilon$$

となるようにできることである。これをコーシー (Cauchy) の収束条件という。

さて, q は式 (C.10) を満たす ($0 < q < 1$) としよう。式 (C.9) を用いると, $m > n$ に対して

$$\begin{aligned} & |x^{(m)} - x^{(n)}| \\ & \leq |x^{(m)} - x^{(m-1)}| + |x^{(m-1)} - x^{(m-2)}| + \cdots + |x^{(n+1)} - x^{(n)}| \\ & \leq (q^{m-1} + q^{m-2} + \cdots + q^n) |x^{(1)} - x^{(0)}| = \frac{q^n(1 - q^{m-n})}{1 - q} |x^{(1)} - x^{(0)}| \\ & \leq \frac{q^n}{1 - q} |x^{(1)} - x^{(0)}| \end{aligned} \quad (\text{C.11})$$

が成り立つので, 関係式 (C.11) は n を十分大きくとればいくらでも 0 に近づく。つまりコーシーの収束条件を満たすので, 数列 $\{x^{(k)}\}$ は収束し, $\lim_{k \rightarrow \infty} x^{(k)} = x$ なる極限值 x が存在する。 φ の連続性を用いると

$$\varphi(x) = \varphi\left(\lim_{k \rightarrow \infty} x^{(k)}\right) = \lim_{k \rightarrow \infty} \varphi(x^{(k)}) = \lim_{k \rightarrow \infty} x^{(k+1)} = x \quad (\text{C.12})$$

となるので, $x = \varphi(x)$ の解は存在しそれは数列 $\{x^{(k)}\}$ の収束値であることが示された。

解の一意性は以下のように示される。 $x = \varphi(x)$ の二つの解を x, y とすると, 平均値の定理と式 (C.10) より

$$\begin{aligned} |x - y| &= |\varphi(x) - \varphi(y)| = |\varphi'(\theta)| |x - y| \\ &\leq q|x - y| \end{aligned} \quad (\text{C.13})$$

となるので, $|x - y| = 0$, すなわち $x = y$ となり, 解の一意性が証明された。

なお, 収束解が考えている区間 K 上に存在するための条件については次項において述べる。

C.1.2 適切な初期値の選択

いままで, $x \in K$ ならば $\varphi(x) \in K$ と仮定してきた。ここでは $x^{(0)} \in K$ から始めて, いかにしたら反復により $\varphi(x)$ が区間 K の外に出ないかについて考察する。 $x^{(k)}$ と $x^{(0)}$ の距離を調べるために, 式 (C.11) において $m = k, n = 0$ とおくと

$$|x^{(k)} - x^{(0)}| \leq \frac{1}{1 - q} |x^{(1)} - x^{(0)}| \quad (\text{C.14})$$

を得る。すなわち $x^{(k)}$ は $x^{(0)}$ から距離 $|x^{(1)} - x^{(0)}|/(1 - q)$ の範囲の中に存在することがわかり, この範囲が区間 K の中に含まれなければならない。このためには, 初期値としては最初の反復で値があまり変わらないような, 十分真の解に近い近似値を選ぶことが大切である。

またさらに重要なことは, 数列 $\{x^{(k)}\}$ の収束値も区間 K 内になければならない。極限值が存在することはすでに示されているので, 式 (C.14) において $k \rightarrow \infty$ とすれば $x^{(k)} \rightarrow x$ より

$$|x - x^{(0)}| \leq \frac{1}{1 - q} |x^{(1)} - x^{(0)}| \quad (\text{C.15})$$

となる。したがって収束値すなわち解も区間 K 内に含まれるためには、初期値の選択として先に述べたものと同じ制約を満たせばよい。

C.1.3 議論の一般化

1 変数スカラー方程式における不動点反復法の収束性の証明において、本質的な役割を果たしたのは、 φ に平均値の定理を用いて導かれた式 (C.6) に条件 (C.10) を課したことである。これは φ が縮小写像であるための条件 (下記) に置き換えられる。

閉区間 $K \subseteq \mathbb{R}$ 上で定義された関数 φ が以下の条件を満足するとき

- i) 任意の $x \in K$ に対して $\varphi(x) \in K$
(写像しても領域の外に出ない)
- ii) 任意の $x, y \in K$ に対して

$$|\varphi(x) - \varphi(y)| \leq q|x - y|$$

が成立するような $0 < q < 1$ なる数 q が存在する。

これを K における縮小写像という。

したがって、 φ が区間 K における縮小写像であれば、 $x^{(0)}$ を初期値とする反復法 $x^{(k+1)} = \varphi(x^{(k)})$ によって生成される列 $\{x^{(k)}\}$ は、 $k \rightarrow \infty$ のとき $x = \varphi(x)$ の唯一解、すなわち不動点に収束する。このことは、収束性の議論の多変数連立方程式系への拡張 (次節) において大きな役割を果たす。

C.2 多変数の連立方程式系

多変数の連立方程式系においても、収束に関する議論の骨格は基本的にはスカラー方程式の場合とほぼ同じである。ただし連立方程式系の場合には、大小の評価尺度にノルムが用いられる。

ベクトル関数 $\varphi(x)$ は n 次元空間のある凸領域 $K \subseteq \mathbb{R}^n$ 上で定義されるとする。また $x \in K$ ならば $\varphi(x) \in K$ 、すなわち φ により写像しても領域 K の外にでないものと仮定する。スカラー方程式の場合と同様

$$\boldsymbol{x} = \boldsymbol{\varphi}(\boldsymbol{x}) \quad (\text{C.16})$$

を満足する \boldsymbol{x} を不動点 (fixed point), 式 (C.16) を不動点問題 (fixed-point problem) という。上式に基づく逐次アルゴリズム

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{\varphi}(\boldsymbol{x}^{(k)}) \quad (\text{C.17})$$

を不動点反復法 (fixed-point iteration) という。

ここで n 元連立方程式系

$$\boldsymbol{f}(\boldsymbol{x}) = \mathbf{0} \quad (\text{C.18})$$

の根を求めるための反復法と関連づけよう。その場合はベクトル関数 $\boldsymbol{\varphi}(\boldsymbol{x})$ を

$$\boldsymbol{\varphi}(\boldsymbol{x}) = \boldsymbol{x} - R(\boldsymbol{x})\boldsymbol{f}(\boldsymbol{x}) \quad (\text{C.19})$$

とおけばよい。 $R(\boldsymbol{x})$ は $n \times n$ 行列であり, もし $R(\boldsymbol{x})$ が正則であれば方程式 $\boldsymbol{f}(\boldsymbol{x}) = \mathbf{0}$ と式 (C.16) は同値であるので, 不動点反復法 (C.17) を用いることができる。 $\boldsymbol{f}(\boldsymbol{x}) = \mathbf{0}$ は連立非線形方程式であってもよいし, 連立 1 次方程式でもよい。連立非線形方程式の場合 $R(\boldsymbol{x})$ の特徴づけによりいろいろなスキームが考えられるが, 特にニュートン法は, $\boldsymbol{f}(\boldsymbol{x})$ のヤコビ行列 $J_f(\boldsymbol{x})$ を用いて

$$R(\boldsymbol{x}) = J_f(\boldsymbol{x})^{-1} \quad (\text{C.20})$$

$$J_f(\boldsymbol{x}) = [(J_f)_{ij}] = \left[\frac{\partial f_i(x_1, x_2, \dots, x_n)}{\partial x_j} \right] \quad (\text{C.21})$$

となるように R を定めたときの不動点反復法 (C.17) と同等である (3 章 参照)。連立 1 次方程式の場合 $\boldsymbol{f}(\boldsymbol{x}) = A\boldsymbol{x} - \boldsymbol{b} = \mathbf{0}$ であり, 反復解法では

$$\boldsymbol{\varphi}(\boldsymbol{x}) = M\boldsymbol{x} + N\boldsymbol{b} \quad (\text{C.22})$$

の形に書ける (4 章 参照)。

さて, 不動点反復法 (C.17) を用いると $k \geq 1$ に対して

$$\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)} = \boldsymbol{\varphi}(\boldsymbol{x}^{(k)}) - \boldsymbol{\varphi}(\boldsymbol{x}^{(k-1)}) \quad (\text{C.23})$$

が成り立つ。 $\boldsymbol{x}^{(k)}$ が解に収束するためには

$$\lim_{k \rightarrow \infty} (\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}) = \mathbf{0}$$

でなければならない。より正確には、 $\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}$ が

$$\|\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}\| \leq q \|\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(k-1)}\| \quad (\text{C.24})$$

の意味で等比数列的に 0 に収束するようにしたい (ここに $q < 1$)。 $\varphi(\boldsymbol{x})$ が 1 次式 (C.22) のときは

$$\begin{aligned} \|\varphi(\boldsymbol{x}^{(k)}) - \varphi(\boldsymbol{x}^{(k-1)})\| &= \left\| M(\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(k-1)}) \right\| \\ &\leq \|M\| \|\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(k-1)}\| \end{aligned} \quad (\text{C.25})$$

が成立するので、 M がどれかのノルムに関して $\|M\| \leq q < 1$ であればそうなる。一般には $\varphi(\boldsymbol{x})$ は非線形であることを考えると、リプシッツ条件

$$\|\varphi(\boldsymbol{y}) - \varphi(\boldsymbol{z})\| \leq q \|\boldsymbol{y} - \boldsymbol{z}\| \quad (\text{C.26})$$

を $0 < q < 1$ に対して満たせばよい。スカラー方程式の場合には C.1.3 項で述べたように、平均値の定理を用いてリプシッツ条件の形にもっていったわけである。ここで $\varphi(\boldsymbol{x})$ が縮小写像となるための条件をリプシッツ条件によって言い換える。

ある凸領域 $K \subseteq \mathbb{R}^n$ において関数 $\varphi(\boldsymbol{x})$ が以下の条件を満たすとき

- i) $\boldsymbol{x} \in K$ ならば $\varphi(\boldsymbol{x}) \in K$
- ii) 任意の $\boldsymbol{y}, \boldsymbol{z} \in K$ に対してリプシッツ条件

$$\|\varphi(\boldsymbol{y}) - \varphi(\boldsymbol{z})\| \leq q \|\boldsymbol{y} - \boldsymbol{z}\|$$

が $0 < q < 1$ なる定数について満足される。

これを領域 K における縮小写像という。

縮小写像に関しては、つぎの重要な定理が成り立つ。

定理 C.1 (縮小写像の不動点定理)

φ が凸領域 $K \subseteq \mathbb{R}^n$ における縮小写像であるとき, $\mathbf{x}^{(0)}$ を初期値とする反復法 $\mathbf{x}^{(k+1)} = \varphi(\mathbf{x}^{(k)})$ により生成される列 $\{\mathbf{x}^{(k)}\}$ は, $k \rightarrow \infty$ のとき $\mathbf{x} = \varphi(\mathbf{x})$ の唯一解, すなわち不動点に収束する。

証明 証明はスカラー方程式の場合と同様に行える。縮小写像であれば,

$$\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\| \leq q \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\| \leq \cdots \leq q^k \|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\| \quad (\text{C.27})$$

が成立し, $m > n$ に対して

$$\|\mathbf{x}^{(m)} - \mathbf{x}^{(n)}\| \leq \frac{q^n}{1-q} \|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\| \quad (\text{C.28})$$

となる(式(C.11)に対応)ので, 数列が収束するための必要十分条件たるコーシーの収束条件

『任意の正数 $\epsilon > 0$ に対して十分大きな N をとれば

$$m, n \geq N \text{ なるとき } \|\mathbf{x}^{(m)} - \mathbf{x}^{(n)}\| < \epsilon$$

とならしめる。』

が満たされ, $\lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = \mathbf{x}$ なる極限值 \mathbf{x} が存在する。 φ の連続性を用いて

$$\mathbf{x} = \varphi(\mathbf{x}) \quad (\text{C.29})$$

となることがわかる。つまり $\mathbf{x}^{(k)}$ は $\mathbf{x} = \varphi(\mathbf{x})$ の解に収束する。このとき解の一意性の確認には, 二つの解 $\mathbf{x} = \varphi(\mathbf{x})$ と $\mathbf{y} = \varphi(\mathbf{y})$ に対してスカラー方程式の場合の式(C.13)と同様のことを行う。

$$\|\mathbf{x} - \mathbf{y}\| = \|\varphi(\mathbf{x}) - \varphi(\mathbf{y})\| \leq q \|\mathbf{x} - \mathbf{y}\| \quad (\text{C.30})$$

より $\|\mathbf{x} - \mathbf{y}\| = 0$, すなわち $\mathbf{x} = \mathbf{y}$ となり, 唯一解であることが証明された。♠

それではどのような場合に $\varphi(\mathbf{x})$ が縮小写像となるかを次に述べる。

C.2.1 収束条件: 条件 ii) を満たすために

$\mathbf{y}, \mathbf{z} \in K$ の2点を結ぶ線分を

$$\mathbf{x}(t) = \mathbf{z} + t(\mathbf{y} - \mathbf{z}) \quad (0 \leq t \leq 1)$$

とする。 φ が K で連続な1階導関数をもてば, \mathbf{y}, \mathbf{z} を結ぶ線分上で

$$\begin{aligned}\varphi(\mathbf{y}) - \varphi(\mathbf{z}) &= \varphi(\mathbf{x}(t))_{t=1} - \varphi(\mathbf{x}(t))_{t=0} \\ &= \int_0^1 \frac{d}{dt} \varphi(\mathbf{x}(t)) dt\end{aligned}\quad (\text{C.31})$$

が成立する。ただし微分・積分は成分ごとに行うものとする。式 (C.31) の右辺の被積分関数を成分ごとに書くと

$$\frac{d}{dt} \varphi_i(\mathbf{x}(t)) = \sum_{j=1}^n \frac{\partial \varphi_i}{\partial x_j} \frac{dx_j}{dt} = \sum_{j=1}^n \left(\frac{\partial \varphi_i}{\partial x_j} \times (y_j - z_j) \right) \quad (\text{C.32})$$

となる。ここで φ のヤコビ行列 J_φ (サイズ: $n \times n$)

$$J_\varphi(\mathbf{x}) = [(J_\varphi)_{ij}] = \left[\frac{\partial \varphi_i}{\partial x_j} \right] \quad (\text{C.33})$$

を導入すると, 式 (C.31) は行列・ベクトル表示で結局つぎのように書ける。

$$\varphi(\mathbf{y}) - \varphi(\mathbf{z}) = \int_0^1 J_\varphi(\mathbf{x}(t)) (\mathbf{y} - \mathbf{z}) dt \quad (\text{C.34})$$

ノルムをとると

$$\begin{aligned}\|\varphi(\mathbf{y}) - \varphi(\mathbf{z})\| &\leq \int_0^1 \|J_\varphi(\mathbf{x}(t))\| \|\mathbf{y} - \mathbf{z}\| dt \\ &= \|\mathbf{y} - \mathbf{z}\| \int_0^1 \|J_\varphi(\mathbf{x}(t))\| dt\end{aligned}\quad (\text{C.35})$$

が得られる。ゆえに

$$\|J_\varphi(\mathbf{x})\| \leq q < 1 \quad (\text{C.36})$$

であれば, $0 < q < 1$ に対するリプシッツ条件

$$\|\varphi(\mathbf{y}) - \varphi(\mathbf{z})\| \leq q \|\mathbf{y} - \mathbf{z}\| \quad (\mathbf{y} \in K, \mathbf{z} \in K) \quad (\text{C.37})$$

を満たす。式 (C.36) は収束のための十分条件であり, スカラー方程式の場合の収束条件 (C.10) に対応する。

C.2.2 適切な初期値の選択: 条件 i) を満たすために

条件 i) において $x \in K$ ならば $\varphi(x) \in K$ と仮定したが, 一般にはこの仮

定が成り立つ K の範囲を決めることは難しい。ここではスカラー方程式の場合と同様に、 $x^{(0)} \in K$ から始めて、いかにしたら反復により $\varphi(x)$ が K の外に出ないかについて考察する。 $x^{(k)}$ と $x^{(0)}$ の距離を調べるために、式 (C.28) において $m = k$, $n = 0$ とおくと

$$\|x^{(k)} - x^{(0)}\| \leq \frac{1}{1-q} \|x^{(1)} - x^{(0)}\| \quad (\text{C.38})$$

を得る。すなわち $x^{(k)}$ は $x^{(0)}$ を中心とし半径 $\|x^{(1)} - x^{(0)}\|/(1-q)$ の球の中に存在することになるので、この球が領域 K の中に含まれるという制限を満たさねばならない。このためには、初期値としては最初の反復で値があまり変わらないような、十分真の解に近い近似値を選ぶことである。

またさらに重要なことは、数列 $\{x^{(k)}\}$ の収束値も領域 K 内になければならない。極限値が存在することはすでに示されているので、式 (C.38) において $k \rightarrow \infty$ とすれば $x^{(k)} \rightarrow x$ より

$$\|x - x^{(0)}\| \leq \frac{1}{1-q} \|x^{(1)} - x^{(0)}\| \quad (\text{C.39})$$

となる。したがって収束値も領域 K 内に含まれるためには、初期値の選択として先に述べたものと同じ制約を満たせばよい。

引用・参考文献

- 1) Fritz, John : Lectures on Advanced Numerical Analysis, Science Publishers, Inc. (1967); 藤田 宏・名取 亮 共訳 : 数値解析講義, 産業図書 (1975)
- 2) 森 正武 : 数値解析, 共立出版 (1973)
- 3) 高木貞治 : 解析概論, 岩波書店 (1975)
- 4) 佐武一郎 : 線形代数学, 裳華房 (1974)

D | 補間と誤差

補間多項式の代表的な表式としてラグランジュ補間とニュートン補間がある。ここではニュートン補間の形で記述された補間多項式から、その誤差を導く。等間隔データ点に対する多項式補間の使用限界を示した後、その対応策としてチェビシェフ補間が誤差を最小にする傾向をもつことを示す。

D.1 多項式補間法：ラグランジュ補間とニュートン補間

独立変数に対応するたがいに異なる点 x_1, x_2, \dots, x_n とそれらの点における関数値 $f(x_1), f(x_2), \dots, f(x_n)$ が与えられているとき、 $k = 1, 2, \dots, n$ に対して $F_n(x_k) = f(x_k)$ となるようなたかだか $n - 1$ 次の補間多項式 $F_n(x)$ は、ラグランジュ (Lagrange) 補間公式の形で

$$F_n(x) = \sum_{k=1}^n f(x_k) L_k^{(n-1)}(x) \quad (\text{D.1})$$

と書ける。ここに $L_k^{(n-1)}(x)$ は

$$L_k^{(n-1)}(x_l) = \delta_{kl} \quad (\text{D.2})$$

(δ_{kl} はクロネッカーの δ) を満足する $n - 1$ 次の多項式 (ラグランジュ基底関数) であり、一意的に次式で与えられる。

$$L_k^{(n-1)}(x) = \frac{(x-x_1)(x-x_2)\cdots(x-x_{k-1})(x-x_{k+1})\cdots(x-x_n)}{(x_k-x_1)(x_k-x_2)\cdots(x_k-x_{k-1})(x_k-x_{k+1})\cdots(x_k-x_n)} \quad (\text{D.3})$$

ラグランジュ基底関数 (D.3) を用いる公式 (D.1) の計算時間を軽減するために、 n 次の多項式

$$\pi_n(x) = (x - x_1)(x - x_2) \cdots (x - x_n) \quad (\text{D.4})$$

を導入して式 (D.3) を

$$L_k^{(n-1)}(x) = \frac{\pi_n(x)}{(x - x_k)\pi_n'(x_k)} \quad (\text{D.5})$$

と表し、これを式 (D.1) に代入した式を用いることがある。

$$F_n(x) = \pi_n(x) \sum_{k=1}^n \frac{f(x_k)}{(x - x_k)\pi_n'(x_k)} \quad (\text{D.6})$$

同じ補間多項式を、データ点を零点にもつ多項式列 (ニュートン基底関数)

$$\begin{aligned} \pi_0(x) &= 1 \\ \pi_1(x) &= (x - x_1) \\ &\vdots \\ \pi_{n-1}(x) &= (x - x_1)(x - x_2) \cdots (x - x_{n-1}) \end{aligned}$$

に関して展開したニュートン (Newton) 補間公式

$$\begin{aligned} F_n(x) &= f[x_1] + (x - x_1)f[x_1, x_2] + (x - x_1)(x - x_2)f[x_1, x_2, x_3] + \cdots \\ &\quad + (x - x_1)(x - x_2) \cdots (x - x_{n-1})f[x_1, x_2, \dots, x_n] \quad (\text{D.7}) \end{aligned}$$

$$= F_{n-1}(x) + (x - x_1)(x - x_2) \cdots (x - x_{n-1})f[x_1, x_2, \dots, x_n] \quad (\text{D.8})$$

の形にも表せる。ここに、係数 $f[x_1], f[x_1, x_2], \dots$ は、以下の漸化式

$$0 \text{ 階差分商} : f[x_1] = f(x_1) \quad (\text{D.9})$$

$$1 \text{ 階差分商} : f[x_1, x_2] = \frac{f[x_2] - f[x_1]}{x_2 - x_1} \quad (\text{D.10})$$

$$2 \text{ 階差分商} : f[x_1, x_2, x_3] = \frac{f[x_2, x_3] - f[x_1, x_2]}{x_3 - x_1} \quad (\text{D.11})$$

\vdots

$$n-1 \text{ 階差分商} : f[x_1, x_2, \dots, x_n] = \frac{f[x_2, \dots, x_n] - f[x_1, \dots, x_{n-1}]}{x_n - x_1} \quad (\text{D.12})$$

により定義される $f(x)$ の差分商であり、ニュートンの補間公式 (D.7) が、与えられたデータ点を補間している、すなわち

$$F_n(x_k) = f(x_k) \quad (k = 1, 2, \dots, n) \quad (\text{D.13})$$

を満たすことは、式 (D.7) に順次 $x = x_k$ を代入していくことによって容易に確かめられる。

一般に $n - 1$ 階差分商は n 個の関数値 $f(x_1), f(x_2), \dots, f(x_n)$ の 1 次結合

$$\begin{aligned} & f[x_1, x_2, \dots, x_n] \\ &= \sum_{k=1}^n \frac{f(x_k)}{(x_k - x_1) \cdots (x_k - x_{k-1})(x_k - x_{k+1}) \cdots (x_k - x_n)} \end{aligned} \quad (\text{D.14})$$

により表される。これは数学的帰納法により証明される。まず 0 階と 1 階の差分商に対して成立しているのは明らかである。つぎに $n - 1$ 階差分商に対し式 (D.14) が成立していると仮定すると、 n 階差分商の定義に対し

$$\begin{aligned} f[x_1, x_2, \dots, x_{n+1}] &= \frac{f[x_2, x_3, \dots, x_{n+1}] - f[x_1, x_2, \dots, x_n]}{x_{n+1} - x_1} \\ &= \frac{1}{(x_{n+1} - x_1)} \left(\sum_{k=2}^n \left(\frac{1}{(x_k - x_2) \cdots (x_k - x_{n+1})} - \frac{1}{(x_k - x_1) \cdots (x_k - x_n)} \right) f(x_k) \right. \\ &\quad \left. + \frac{1}{(x_{n+1} - x_2) \cdots (x_{n+1} - x_n)} f(x_{n+1}) - \frac{1}{(x_1 - x_2) \cdots (x_1 - x_n)} f(x_1) \right) \\ &= \sum_{k=1}^{n+1} \frac{f(x_k)}{(x_k - x_1) \cdots (x_k - x_{k-1})(x_k - x_{k+1}) \cdots (x_k - x_{n+1})} \end{aligned}$$

が得られ、 n 階差分商においても式 (D.14) が成立していることが確かめられ、証明は終わる。式 (D.14) より、 $f[x_1, x_2, \dots, x_n]$ の値は x_1, x_2, \dots, x_n の順序を入れ換えても変わらないことがわかる。

差分商中の変数が一致する場合は極限により意味づけられる。例えば

$$f[x_1, x_1] = \lim_{x_2 \rightarrow x_1} f[x_1, x_2] = \lim_{x_2 \rightarrow x_1} \frac{f[x_2] - f[x_1]}{x_2 - x_1} = \frac{df(x_1)}{dx_1} \quad (\text{D.15})$$

と表され、同様にしてつぎの公式が得られる。

$$f[x_1, x_1, x_2, \dots, x_k] = \frac{d}{dx_1} f[x_1, x_2, \dots, x_k] \quad (\text{D.16})$$

x_1 について微分可能な関数 y_1, y_2, \dots, y_k を用いると,

$$\frac{d}{dx_1} f[y_1, y_2, \dots, y_k] = \sum_{i=1}^k f[y_i, y_1, y_2, \dots, y_k] \frac{dy_i}{dx_1} \quad (\text{D.17})$$

となるので, $y_1 = \dots = y_k = x_1$ とおけば,

$$\frac{d}{dx_1} f[\overbrace{x_1, \dots, x_1}^{k \text{個}}] = k f[\overbrace{x_1, x_1, \dots, x_1}^{k+1 \text{個}}] \quad (\text{D.18})$$

を得る。したがって, 式 (D.15), (D.18) より

$$\begin{aligned} \frac{d^{k-1}}{dx_1^{k-1}} f(x_1) &= \frac{d^{k-2}}{dx_1^{k-2}} \left(\frac{d}{dx_1} f(x_1) \right) = \frac{d^{k-2}}{dx_1^{k-2}} f[x_1, x_1] \\ &= \frac{d^{k-3}}{dx_1^{k-3}} \left(\frac{d}{dx_1} f[x_1, x_1] \right) = 2 \frac{d^{k-3}}{dx_1^{k-3}} f[x_1, x_1, x_1] \\ &= 2 \frac{d^{k-4}}{dx_1^{k-4}} \left(\frac{d}{dx_1} f[x_1, x_1, x_1] \right) = 3 \cdot 2 \frac{d^{k-4}}{dx_1^{k-4}} f[x_1, x_1, x_1, x_1] \\ &\quad \vdots \\ &= (k-1)! f[\overbrace{x_1, \dots, x_1}^{k \text{個}}] \end{aligned} \quad (\text{D.19})$$

が成立するので, つぎの公式が得られる。

$$f[\overbrace{x_1, \dots, x_1}^{k \text{個}}] = \frac{1}{(k-1)!} \frac{d^{k-1} f(x_1)}{dx_1^{k-1}} \quad (\text{D.20})$$

D.2 多項式補間法の誤差

n 個のデータ点を補間するたかだか $n-1$ 次の多項式は一つに定まるので, 表現方法は異なるにせよ, ラグランジュ補間法もニュートン補間法も同一の多項式を指している。

いまラグランジュ補間法 (D.6) の誤差を $\varepsilon_n(x)$ とすると

$$\varepsilon_n(x) = F_n(x) - f(x) \quad (\text{D.21})$$

である。ここで n 個のデータ点に 1 点 $(x_{n+1}, f(x_{n+1}))$ を付加して $n+1$ 個にする。この $n+1$ 点を通るニュートン補間公式 $F_{n+1}(x)$ において $x = x_{n+1}$ とおくと、式 (D.8) より

$$\begin{aligned} f(x_{n+1}) &= F_{n+1}(x_{n+1}) \\ &= F_n(x_{n+1}) + (x_{n+1} - x_1)(x_{n+1} - x_2) \cdots \\ &\quad (x_{n+1} - x_n) f[x_1, x_2, \dots, x_n, x_{n+1}] \end{aligned} \quad (\text{D.22})$$

なる関係が得られる。 x_{n+1} としては x_1, x_2, \dots, x_n 以外の任意の点を取り得るので、再びこれを x に置き換えると、式 (D.21) の誤差は次式により与えられる。

$$\varepsilon_n(x) = -(x - x_1)(x - x_2) \cdots (x - x_n) f[x_1, x_2, \dots, x_n, x] \quad (\text{D.23})$$

これは誤差の正確な表現ではあるが、 $f(x)$ の差分商をどのように見積もればよいのであろうか。

$F_n(x)$ は $f(x_1), \dots, f(x_n)$ を補間する関数であるから、 $x = x_1, \dots, x_n$ に対して

$$\varepsilon_n(x) = F_n(x) - f(x) = 0 \quad (\text{D.24})$$

である。つまり $\varepsilon_n(x) = 0$ は少なくとも n 個の零点をもつのであるから、Rolle の定理[†]より $d\varepsilon_n(x)/dx$ は少なくとも $n-1$ 個の零点をもつことになる。この論法を繰り返して $n-1$ 回 Rolle の定理を用いれば、 $d^{n-1}\varepsilon_n(x)/dx^{n-1}$ は x_1, \dots, x_n を含む最小区間の中の少なくとも 1 点 ξ において零となる、すなわち

$$\frac{d^{n-1}\varepsilon_n(\xi)}{dx^{n-1}} = \frac{d^{n-1}F_n(\xi)}{dx^{n-1}} - \frac{d^{n-1}f(\xi)}{dx^{n-1}} = 0 \quad (\text{D.25})$$

なる ξ が存在する。 $F_n(x)$ に対するニュートン補間公式 (D.7) より最高次の項は

[†] Rolle の定理： $g(x)$ は区間 $[a, b]$ で連続、 (a, b) で微分可能とする。このとき $g(a) = g(b)$ ならば、区間 (a, b) のある点において $g'(x)$ は零になる、すなわち $g'(\xi) = 0$ なる ξ が $a < \xi < b$ において存在する。

$$F_n(x) = x^{n-1}f[x_1, x_2, \dots, x_n] + \dots \quad (\text{D.26})$$

となるので、これを $n-1$ 回 x で微分すると

$$\frac{d^{n-1}F_n(\xi)}{dx^{n-1}} = (n-1)! f[x_1, x_2, \dots, x_n] \quad (\text{D.27})$$

を得る。これを式 (D.25) に代入すれば

$$\frac{d^{n-1}\varepsilon_n(\xi)}{dx^{n-1}} = (n-1)! f[x_1, x_2, \dots, x_n] - \frac{d^{n-1}f(\xi)}{dx^{n-1}} = 0 \quad (\text{D.28})$$

となるので

$$f[x_1, x_2, \dots, x_n] = \frac{1}{(n-1)!} \frac{d^{n-1}f(\xi)}{dx^{n-1}} \quad (\text{D.29})$$

を得る。したがって式 (D.23) の誤差は、

$$\varepsilon_n(x) = -\frac{(x-x_1)(x-x_2)\cdots(x-x_n)}{n!} \frac{d^n f(\xi)}{dx^n} \quad (\text{D.30})$$

と表され、誤差の見積りが可能な形となった。ここに ξ は x_1, \dots, x_n, x を含む最小区間の中のある点である。 n 階微係数 $d^n f(x)/dx^n$ の計算は一般に煩雑になることと、 ξ の値が未知であることにより、式 (D.30) を誤差の推定に用いることには限界はあるものの、誤差の振舞いを定性的にある程度は知ることができる。

なお、誤差項 (D.30) はテイラーの公式における剰余に似た形をしていることに気づく。実際、 x_1, x_2, \dots, x_n を 1 点 x_1 に一致させる極限では公式 (D.20) が成立するので、そのときニュートン補間公式 (D.7) はテイラー展開となる。

$$\begin{aligned} f(x) &= f(x_1) + \frac{(x-x_1)}{1!} \frac{df(x_1)}{dx} + \dots \\ &\quad + \frac{(x-x_1)^{n-1}}{(n-1)!} \frac{d^{n-1}f(x_1)}{dx^{n-1}} + \frac{(x-x_1)^n}{n!} \frac{d^n f(\zeta)}{dx^n} \end{aligned}$$

ζ は x_1 と x との間の点である。

$$|(x-x_1)^n| > |(x-x_1)(x-x_2)\cdots(x-x_n)|$$

となる場合、補間多項式の誤差項 $\varepsilon_n(x)$ はテイラー展開の誤差項より小さいことが多い。

D.3 等間隔データ点に対する多項式補間の使用限界

関数 $f(x)$ に対する $n-1$ 次の補間多項式は、ニュートンの補間公式により記述すると

$$F_n(x) = f[x_1] + \sum_{k=2}^n (x-x_1)(x-x_2)\cdots(x-x_{k-1})f[x_1, x_2, \dots, x_k] \quad (\text{D.31})$$

となる。無限級数

$$f[x_1] + \sum_{k=2}^{\infty} (x-x_1)(x-x_2)\cdots(x-x_{k-1})f[x_1, x_2, \dots, x_k] \quad (\text{D.32})$$

は少なくとも解析関数（領域内で微分可能な関数）に対しては $f(x)$ に収束するとも考えるかもしれないが、一般にはこれは正しくない。例えば、 $f(x) = e^x$ とし、等間隔のデータ点 $x_j = x_1 + (j-1)h$ をとってみよう。このとき式 (D.14) により

$$\begin{aligned} f[x_1, \dots, x_k] &= \sum_{j=1}^k \frac{e^{x_1+(j-1)h}}{h^{k-1}(j-1)!(k-j)!(-1)^{k-j}} \\ &= \frac{e^{x_1}}{(k-1)!h^{k-1}} \sum_{j=1}^k \frac{(k-1)!}{(j-1)!(k-j)!} (-1)^{k-j} e^{(j-1)h} \\ &= \frac{e^{x_1}}{(k-1)!} \left(\frac{e^h - 1}{h} \right)^{k-1} \end{aligned} \quad (\text{D.33})$$

となるので、点 $x = x_1 + h/2$ に対して、級数 (D.32) の第 k 項の絶対値はつぎの不等式で見積もられる。

$$\begin{aligned} &\left(\frac{1}{2}h\right) \left(\frac{1}{2}h\right) \left(\frac{3}{2}h\right) \left(\frac{5}{2}h\right) \cdots \left(\left(k-\frac{5}{2}\right)h\right) \frac{e^{x_1}}{(k-1)!} \left(\frac{e^h - 1}{h}\right)^{k-1} \\ &\geq \frac{1}{4} \frac{1}{(k-1)(k-2)} e^{x_1} (e^h - 1)^{k-1} \end{aligned} \quad (\text{D.34})$$

収束するためには、 $k \rightarrow \infty$ のとき、第 k 項 $\rightarrow 0$ とならなければならない。し

たがって $|e^h - 1| \leq 1$ でなければならず，データ点の間隔 h は $\log 2$ をこえてはならない。

以上はデータ点の間隔が有限な場合の例であるが，データ点を無限に増やしてデータ点間隔を限りなく零に近づける ($n \rightarrow \infty$ で $h \rightarrow 0$) ときでも，補間多項式は関数に収束するとは限らない。ルンゲはその例として，すべての実数に対して解析的な関数

$$f(x) = \frac{1}{1 + 25x^2} \quad (\text{D.35})$$

を取り上げ，区間 $[-1, 1]$ において等間隔のデータ点 x_k を無限に増やしていく (したがって次数を無限に高くしていく) とき，補間多項式が区間の端付近で発散することを証明した。この特別な例における発散の原因は虚の特異点 $x = \pm i/5$ が補間データ点 x_k の近傍に存在することによる。特異点が補間点から遠く離れているならば，解析関数 $f(x)$ に対する補間多項式は収束する (文献 1) 参照)。

D.4 チェビシユフ多項式による準ミニマックス近似

前節で等間隔データ点に対する多項式補間の使用限界が示された。それではいかなる補間を用いると，誤差を抑えることが可能であろうか。

ある区間における連続関数 $f(x)$ の近似関数を，誤差の絶対値の最大値を最小にするよう定める方法を最良近似法，あるいはミニマックス近似法という。いま， n 個のデータ点 x_1, x_2, \dots, x_n 上で $f(x)$ と一致する $n-1$ 次の近似多項式を $F_n(x)$ とし，区間 $[a, b]$ における誤差の絶対値

$$\max_{a \leq x \leq b} |f(x) - F_n(x)|$$

を最小にするように $F_n(x)$ を定める問題を考える。式 (D.21) と式 (D.30) より

$$f(x) = F_n(x) + \frac{(x - x_1)(x - x_2) \cdots (x - x_n)}{n!} \frac{d^n f(\xi)}{dx^n} \quad (\text{D.36})$$

と書ける。ここに ξ は x_1, \dots, x_n を含む最小区間の中の 1 点である。このとき最優良近似は

$$\max_{a \leq x \leq b} |f(x) - F_n(x)| = \frac{1}{n!} \max_{a \leq x \leq b} \left| (x-x_1)(x-x_2) \cdots (x-x_n) \frac{d^n f(\xi)}{dx^n} \right| \quad (\text{D.37})$$

を最小にすることを要求する。いま区間で $|d^n f(x)/dx^n| \leq M$ なる M が既知であるならば、つぎの誤差限界

$$\max_{a \leq x \leq b} |f(x) - F_n(x)| \leq \frac{M}{n!} \max_{a \leq x \leq b} |(x-x_1)(x-x_2) \cdots (x-x_n)| \quad (\text{D.38})$$

を得るので、式 (D.37) の代わりに

$$\max_{a \leq x \leq b} |(x-x_1)(x-x_2) \cdots (x-x_n)| \quad (\text{D.39})$$

を最小にするよう位置 x_1, \dots, x_n を定めることにより、ミニマックス近似に準じるような近似とすることができる。

ここでなんらかの変数変換により考える区間を $[-1, 1]$ に移したとすると、上記の条件を満足する多項式は、最高次 x^n の係数が 1 となるように 2^{n-1} で割った n 次チェビシェフ (Chebyshev) 多項式 $T_n(x)$

$$(x-x_1)(x-x_2) \cdots (x-x_n) = \frac{T_n(x)}{2^{n-1}} \quad (\text{D.40})$$

で与えられる。 n 次チェビシェフ多項式は

$$T_n(x) = \cos(n \cos^{-1} x) \quad (\text{D.41})$$

であり、その n 個の零点 x_i は

$$x_i = \cos \frac{(2i-1)\pi}{2n} \quad (i = 1, 2, \dots, n) \quad (\text{D.42})$$

により与えられる。式 (D.41) より $T_n(x) \leq 1$ であるから、このとき

$$\max_{-1 \leq x \leq 1} |(x-x_1)(x-x_2) \cdots (x-x_n)| = \frac{1}{2^{n-1}} \quad (\text{D.43})$$

となる。これが他のいかなる n 次多項式よりも最小値を与えることは、背理法により証明される。 n 次多項式 $\delta(x)$ があり、最高次の係数が 1 で区間内における $|\delta(x)|$ の最大値が $1/2^{n-1}$ より小さいとすれば

$$\phi(x) \equiv \delta(x) - \frac{T_n(x)}{2^{n-1}} \quad (\text{D.44})$$

は $T_n = 1$ をとるとき負になり, $T_n = -1$ をとるとき正になる。 T_n は区間両端を含む $n + 1$ 個の点で正負の極値 ± 1 を交互にとるので, $\phi(x)$ は n 個の零点をもつ。しかるに $\phi(x)$ は定義式 (D.44) より x^n の項がなくなり $n - 1$ 次式であるから, 矛盾する。

以上より, データ点 x_1, \dots, x_n をチェビシェフ多項式の零点 (D.42) とする多項式補間, すなわちチェビシェフ補間は, $d^n f(\xi)/dx^n$ の影響がなければ補間多項式の誤差の最大値を最小にすることが示された。

引用・参考文献

- 1) Fritz, John : Lectures on Advanced Numerical Analysis, Science Publishers, Inc. (1967); 藤田 宏・名取 亮 共訳: 数値解析講義, 産業図書 (1975)
- 2) 赤坂 隆: 数値計算, 応用数学講座第 7 巻, コロナ社 (1976)
- 3) 森 正武: 数値解析, 共立出版 (1973)

E

数値微分の差分演算子表現

9章では、テイラー級数展開からの式変形、およびラグランジュ補間多項式の微分から数値微分公式を求めた。ここではより体系的に数値微分公式を示すために、ニュートン補間法のさまざまな差分演算子表現を導き、その微分から導関数の差分演算子表現を用いた近似を示す。さらに微分演算子と差分演算子との関係からも数値微分公式を導出する。以下、幅 h の等間隔で配列している離散的な点 x

$$x_{i+k} = x_i + kh \quad (k = 0, \pm 1, \pm 2, \dots, n-1) \quad (\text{E.1})$$

にて関数 $f(x)$ の値 $f(x_{i+k})$ が与えられているとして、関数 $f(x)$ の導関数を近似的に導こう。

E.1 差分演算子と差分商

幅 h の等間隔で配列している x に対し関数値が与えられているとき、つぎのように差分演算子 (difference operator) を定義する。

$$\text{前進差分演算子: } \Delta f(x) = f(x+h) - f(x) \quad (\text{E.2})$$

$$\text{後退差分演算子: } \nabla f(x) = f(x) - f(x-h) \quad (\text{E.3})$$

$$\text{中心差分演算子: } \delta f(x) = f\left(x + \frac{h}{2}\right) - f\left(x - \frac{h}{2}\right) \quad (\text{E.4})$$

移相演算子 (shift operator) E を

$$Ef(x) = f(x+h) \quad (\text{E.5})$$

$$E^{-1}f(x) = f(x-h) \quad (\text{E.6})$$

と定義すると、正の整数 r に対して

$$E^r f(x) = f(x+rh), \quad E^{-r} f(x) = f(x-rh)$$

となり、つぎのような演算子の関係が得られる。

$$\Delta f(x) = (E-1)f(x) \quad \implies \quad \Delta = E-1 \quad (\text{E.7})$$

$$\nabla f(x) = (1-E^{-1})f(x) \quad \implies \quad \nabla = 1-E^{-1} \quad (\text{E.8})$$

$$\delta f(x) = (E^{1/2}-E^{-1/2})f(x) \quad \implies \quad \delta = E^{1/2}-E^{-1/2} \quad (\text{E.9})$$

中心差分演算子の定義には h の半整数倍が現れている。そのため平均演算子を

$$\mu f(x) = \frac{1}{2} \left(f\left(x+\frac{h}{2}\right) + f\left(x-\frac{h}{2}\right) \right) \quad (\text{E.10})$$

により定義すると、以下のように中心差分を h の（半整数倍ではなく）整数倍のみが現れるよう定義できる。

$$\begin{aligned} \mu\delta f(x) &= \mu f\left(x+\frac{h}{2}\right) - \mu f\left(x-\frac{h}{2}\right) \\ &= \frac{1}{2}(f(x+h)+f(x)) - \frac{1}{2}(f(x)+f(x-h)) \\ &= \frac{1}{2}(f(x+h)-f(x-h)) \end{aligned} \quad (\text{E.11})$$

これを演算子の関係で表すと、つぎのように書ける。

$$\begin{aligned} \mu f(x) &= \frac{1}{2}(E^{1/2}+E^{-1/2})f(x) \\ \implies \quad \mu &= \frac{1}{2}(E^{1/2}+E^{-1/2}) \end{aligned} \quad (\text{E.12})$$

$$\begin{aligned} \mu\delta f(x) &= \frac{1}{2}(E^{1/2}+E^{-1/2})(E^{1/2}-E^{-1/2})f(x) \\ \implies \quad \mu\delta &= \frac{1}{2}(E-E^{-1}) \end{aligned} \quad (\text{E.13})$$

ここで差分商と差分との関係を考えよう。前進差分に関して

$$\begin{aligned}
\Delta f(x_i) &= f(x_{i+1}) - f(x_i) = (x_{i+1} - x_i)f[x_i, x_{i+1}] \\
&= hf[x_i, x_{i+1}] \\
\Delta^2 f(x_i) &= \Delta f(x_{i+1}) - \Delta f(x_i) = h(f[x_{i+1}, x_{i+2}] - f[x_i, x_{i+1}]) \\
&= h(x_{i+2} - x_i)f[x_i, x_{i+1}, x_{i+2}] \\
&= 2! h^2 f[x_i, x_{i+1}, x_{i+2}] \\
\Delta^3 f(x_i) &= \Delta^2 f(x_{i+1}) - \Delta^2 f(x_i) \\
&= 2! h^2 (f[x_{i+1}, x_{i+2}, x_{i+3}] - f[x_i, x_{i+1}, x_{i+2}]) \\
&= 2! h^2 (x_{i+3} - x_i)f[x_i, x_{i+1}, x_{i+2}, x_{i+3}] \\
&= 3! h^3 f[x_i, x_{i+1}, x_{i+2}, x_{i+3}] \\
&\vdots
\end{aligned}$$

となり，一般に r 階前進差分 $\Delta^r f(x_i)$ は帰納的に

$$\Delta^r f(x_i) = r! h^r f[x_i, x_{i+1}, x_{i+2}, \dots, x_{i+r}] \quad (\text{E.14})$$

と表される。後退差分に関しては

$$\begin{aligned}
\nabla f(x_i) &= f(x_i) - f(x_{i-1}) = (x_i - x_{i-1})f[x_i, x_{i-1}] \\
&= hf[x_i, x_{i-1}] \\
&\vdots
\end{aligned}$$

となり，同様にして r 階後退差分 $\nabla^r f(x_i)$ は帰納的に

$$\nabla^r f(x_i) = r! h^r f[x_i, x_{i-1}, x_{i-2}, \dots, x_{i-r}] \quad (\text{E.15})$$

と表される。中心差分に関しては

$$\delta f(x_i) = f(x_{i+1/2}) - f(x_{i-1/2}) \quad (\text{E.16})$$

となって，右辺は与えられていない値を参照することになるので

$$\delta f(x_{i+1/2}) = f(x_{i+1}) - f(x_i) \quad (\text{E.17})$$

を扱う必要がある。2階の中心差分は

$$\delta^2 f(x_i) = \delta f(x_{i+1/2}) - \delta f(x_{i-1/2}) = f(x_{i+1}) - 2f(x_i) + f(x_{i-1}) \quad (\text{E.18})$$

となって、右辺は与えられた値で表される。このように、奇数階差分に対しては半整数の添字を定義しなければならないが、偶数階差分では添字は整数のみが現れる。中心差分と差分商との関係は

$$\begin{aligned}
 \delta f(x_{i+1/2}) &= f(x_{i+1}) - f(x_i) = (x_{i+1} - x_i)f[x_i, x_{i+1}] \\
 &= hf[x_i, x_{i+1}] \\
 \delta^2 f(x_i) &= \delta f(x_{i+1/2}) - \delta f(x_{i-1/2}) = h(f[x_i, x_{i+1}] - f[x_{i-1}, x_i]) \\
 &= h(x_{i+1} - x_{i-1})f[x_{i-1}, x_i, x_{i+1}] \\
 &= 2! h^2 f[x_{i-1}, x_i, x_{i+1}] \\
 \delta^3 f(x_{i+1/2}) &= \delta^2 f(x_{i+1}) - \delta^2 f(x_i) \\
 &= 2! h^2 (f[x_i, x_{i+1}, x_{i+2}] - f[x_{i-1}, x_i, x_{i+1}]) \\
 &= 2! h^2 (x_{i+2} - x_{i-1})f[x_{i-1}, x_i, x_{i+1}, x_{i+2}] \\
 &= 3! h^3 f[x_{i-1}, x_i, x_{i+1}, x_{i+2}] \\
 \delta^4 f(x_i) &= \delta^3 f(x_{i+1/2}) - \delta^3 f(x_{i-1/2}) \\
 &= 3! h^3 (f[x_{i-1}, x_i, x_{i+1}, x_{i+2}] - f[x_{i-2}, x_{i-1}, x_i, x_{i+1}]) \\
 &= 3! h^3 (x_{i+2} - x_{i-2})f[x_{i-2}, x_{i-1}, x_i, x_{i+1}, x_{i+2}] \\
 &= 4! h^4 f[x_{i-2}, x_{i-1}, x_i, x_{i+1}, x_{i+2}] \\
 &\vdots
 \end{aligned}$$

となるので、一般には奇数階と偶数階の差分に分けて帰納的につぎのように書ける。

$$\delta^{2m+1} f(x_{i+1/2}) = (2m+1)! h^{2m+1} f[x_{i-m}, \dots, x_i, \dots, x_{i+m+1}] \quad (\text{E.19})$$

$$\delta^{2m} f(x_i) = (2m)! h^{2m} f[x_{i-m}, \dots, x_i, \dots, x_{i+m}] \quad (\text{E.20})$$

最後に、前進差分 $\Delta^r f(x_i)$ と後退差分および中心差分との関係を以下に示す。

$$\begin{aligned}\Delta^r f(x_i) &= (E-1)^r f(x_i) = (1-E^{-1})^r E^r f(x_i) \\ &= \nabla^r E^r f(x_i) = \nabla^r f(x_{i+r})\end{aligned}\quad (\text{E.21})$$

$$\begin{aligned}\Delta^r f(x_i) &= (E-1)^r f(x_i) = (E^{1/2} - E^{-1/2})^r E^{r/2} f(x_i) \\ &= \delta^r E^{r/2} f(x_i) = \delta^r f(x_{i+r/2})\end{aligned}\quad (\text{E.22})$$

E.2 ニュートン補間法の差分演算子表現と数値微分

導関数の差分近似式を高精度化したり、高階の導関数の差分近似式を求めたりする場合、ニュートン補間法の差分演算子表現を用いれば系統的となる。

E.2.1 ニュートンの前進補間公式と数値微分

離散的な n 個の点 x_{i+r} , ($r = 0, 1, \dots, n-1$) において関数 $f(x)$ の値 $f(x_{i+r})$ が与えられているとき、関係式 (E.14) より差分商は

$$f[x_i, x_{i+1}, x_{i+2}, \dots, x_{i+r}] = \frac{1}{r! h^r} \Delta^r f(x_i) \quad (\text{E.23})$$

となるので、 $n-1$ 次ニュートン補間多項式 $F_n(x)$ を前進差分を用いて表すと、誤差項 $\varepsilon_n(x)$ を含めて

$$f(x) = F_n(x) - \varepsilon_n(x) \quad (\text{E.24})$$

$$\begin{aligned}F_n(x) &= f(x_i) + (x-x_i) \frac{\Delta f(x_i)}{1! h} + (x-x_i)(x-x_{i+1}) \frac{\Delta^2 f(x_i)}{2! h^2} + \dots \\ &\quad + (x-x_i) \cdots (x-x_{i+n-2}) \frac{\Delta^{n-1} f(x_i)}{(n-1)! h^{n-1}}\end{aligned}\quad (\text{E.25})$$

$$\varepsilon_n(x) = -\frac{(x-x_i)(x-x_{i+1}) \cdots (x-x_{i+n-1})}{n!} \frac{d^n f(\xi)}{dx^n} \quad (\text{E.26})$$

と書ける。ただし $\xi \in (x_i, x_{i+n-1})$ 。この式は、無次元変数 α

$$x = x_i + \alpha h \quad (0 \leq \alpha \leq n-1) \quad (\text{E.27})$$

を導入すると $x - x_{i+r} = h(\alpha - r)$ となり、つぎのように簡単な形になる。

$$f(x_i + \alpha h) = F_n(x) - \varepsilon_n(x) \quad (\text{E.28})$$

$$\begin{aligned} F_n(x) &= f(x_i) + \alpha \Delta f(x_i) + \frac{\alpha(\alpha-1)}{2!} \Delta^2 f(x_i) \\ &\quad + \frac{\alpha(\alpha-1)(\alpha-2)}{3!} \Delta^3 f(x_i) + \cdots \\ &\quad + \frac{\alpha(\alpha-1) \cdots (\alpha-n+2)}{(n-1)!} \Delta^{n-1} f(x_i) \end{aligned} \quad (\text{E.29})$$

$$\varepsilon_n(x) = -\frac{h^n \alpha(\alpha-1) \cdots (\alpha-n+1)}{n!} \frac{d^n f(\xi)}{dx^n} \quad (\text{E.30})$$

式 (E.29) はニュートンの前進補間公式と呼ばれる。誤差項を含めた式 (E.28) を x で微分すると, $d\alpha/dx = 1/h$ より

$$\begin{aligned} f'(x_i + \alpha h) &= \frac{1}{h} \Delta f(x_i) + \frac{2\alpha-1}{2h} \Delta^2 f(x_i) + \frac{3\alpha^2-6\alpha+2}{6h} \Delta^3 f(x_i) + \cdots \\ &\quad + \frac{1}{h(n-1)!} \frac{d}{d\alpha} (\alpha(\alpha-1) \cdots (\alpha-n+2)) \Delta^{n-1} f(x_i) \\ &\quad + \frac{h^{n-1}}{n!} \frac{d}{d\alpha} (\alpha(\alpha-1) \cdots (\alpha-n+1)) \frac{d^n f(\xi)}{dx^n} \end{aligned} \quad (\text{E.31})$$

を得る。 $\alpha = 0, 1, 2, \dots$ ととると $x_i + \alpha h = x_{i+\alpha}$ はあらかじめ与えられた点となるので, 上式より 1 階導関数に対するさまざまな差分公式が得られる。例えば $n = 2$ として右辺第 1 項までとると誤差項は $(h/2)(2\alpha-1)(d^2 f(\xi)/dx^2)$ となり, $\alpha = 0$ とおくと 1 次精度前進差分公式, $\alpha = 1$ とおくと 1 次精度後退差分公式が得られる。 $n = 3$ として右辺第 2 項までとると誤差項は $(h^2/6)(3\alpha^2-6\alpha+2)(d^3 f(\xi)/dx^3)$ となり, $\alpha = 0$ とおくと 2 次精度前進差分公式, $\alpha = 1$ とおくと 2 次精度中心差分公式, $\alpha = 2$ とおくと 2 次精度後退差分公式が得られる。

さらに式 (E.31) を微分すると

$$\begin{aligned} f''(x_i + \alpha h) &= \frac{1}{h^2} \Delta^2 f(x_i) + \frac{6\alpha-6}{6h^2} \Delta^3 f(x_i) + \cdots \\ &\quad + \frac{1}{h^2(n-1)!} \frac{d^2}{d\alpha^2} (\alpha(\alpha-1) \cdots (\alpha-n+2)) \Delta^{n-1} f(x_i) \\ &\quad + \frac{h^{n-2}}{n!} \frac{d^2}{d\alpha^2} (\alpha(\alpha-1) \cdots (\alpha-n+1)) \frac{d^n f(\xi)}{dx^n} \end{aligned} \quad (\text{E.32})$$

を得る。上式より 2 階導関数に対する差分公式が得られる。 $n = 3$ として右辺第 1 項までとって $\alpha = 0$ とおくと 1 次精度前進差分公式, $\alpha = 1$ とおくと 2 次精度中心差分公式, $\alpha = 2$ とおくと 1 次精度後退差分公式が得られる。このようにしてさらに高精度, 高階の導関数の数値微分公式が得られていく。

E.2.2 ニュートンの後退補間公式と数値微分

同様にして, ニュートンの後退補間公式を導くことができる。離散的な n 個の点 x_{i-r} , ($r = 0, 1, \dots, n-1$) において関数 $f(x)$ の値 $f(x_{i-r})$ が与えられているとき, 関係式 (E.15) より差分商は

$$f[x_i, x_{i-1}, x_{i-2}, \dots, x_{i-r}] = \frac{1}{r! h^r} \nabla^r f(x_i) \quad (\text{E.33})$$

となるので, $n-1$ 次ニュートン補間多項式 $F_n(x)$ を後退差分を用いて表し, 無次元変数 α

$$x = x_i + \alpha h \quad (0 \geq \alpha \geq -(n-1)) \quad (\text{E.34})$$

を導入すると, 誤差項 $\varepsilon_n(x)$ を含めて以下のように簡単な形に書ける。

$$f(x_i + \alpha h) = F_n(x) - \varepsilon_n(x) \quad (\text{E.35})$$

$$\begin{aligned} F_n(x) = & f(x_i) + \alpha \nabla f(x_i) + \frac{\alpha(\alpha+1)}{2!} \nabla^2 f(x_i) \\ & + \frac{\alpha(\alpha+1)(\alpha+2)}{3!} \nabla^3 f(x_i) + \dots \\ & + \frac{\alpha(\alpha+1) \cdots (\alpha+n-2)}{(n-1)!} \nabla^{n-1} f(x_i) \end{aligned} \quad (\text{E.36})$$

$$\varepsilon_n(x) = - \frac{h^n \alpha(\alpha+1) \cdots (\alpha+n-1)}{n!} \frac{d^n f(\xi)}{dx^n} \quad (\text{E.37})$$

式 (E.36) はニュートンの後退補間公式と呼ばれる。誤差項を含めた式 (E.35) を x で微分すると, $d\alpha/dx = 1/h$ より

$$\begin{aligned}
f'(x_i + \alpha h) &= \frac{1}{h} \nabla f(x_i) + \frac{2\alpha + 1}{2h} \nabla^2 f(x_i) + \frac{3\alpha^2 + 6\alpha + 2}{6h} \nabla^3 f(x_i) + \cdots \\
&+ \frac{1}{h(n-1)!} \frac{d}{d\alpha} (\alpha(\alpha+1)\cdots(\alpha+n-2)) \nabla^{n-1} f(x_i) \\
&+ \frac{h^{n-1}}{n!} \frac{d}{d\alpha} (\alpha(\alpha+1)\cdots(\alpha+n-1)) \frac{d^n f(\xi)}{dx^n} \quad (\text{E.38})
\end{aligned}$$

を得る。 $\alpha = 0, -1, -2, \dots$ ととると $x_i + \alpha h = x_{i+\alpha}$ はあらかじめ与えられた点となるので、上式より1階導関数に対するさまざまな差分公式が得られる。例えば $n = 2$ として右辺第1項までとると誤差項は $(h/2)(2\alpha + 1)(d^2 f(\xi)/dx^2)$ となり、 $\alpha = 0$ とおくと1次精度後退差分公式、 $\alpha = -1$ とおくと1次精度前進差分公式が得られる。 $n = 3$ として右辺第2項までとると誤差項は $(h^2/6)(3\alpha^2 + 6\alpha + 2)(d^3 f(\xi)/dx^3)$ となり、 $\alpha = 0$ とおくと2次精度後退差分公式、 $\alpha = -1$ とおくと2次精度中心差分公式、 $\alpha = -2$ とおくと2次精度前進差分公式が得られる。

E.2.3 スターリングの補間公式と数値微分

さらに差分商と中心差分との関係からニュートン補間の中心差分表現を導くことができる。補間点を $x_i, x_{i+1}, x_{i-1}, x_{i+2}, x_{i-2}, \dots$ の順序にとっていくと、ニュートンの補間多項式は

$$\begin{aligned}
f(x) &= f[x_i] + (x - x_i)f[x_i, x_{i+1}] + (x - x_i)(x - x_{i+1})f[x_i, x_{i+1}, x_{i-1}] \\
&+ (x - x_i)(x - x_{i+1})(x - x_{i-1})f[x_i, x_{i+1}, x_{i-1}, x_{i+2}] + \cdots
\end{aligned} \quad (\text{E.39})$$

と書ける。差分商の値はかぎ括弧 [] の中の値の順序によらないので、関係式 (E.19), (E.20) より

$$\begin{aligned}
f[x_i, x_{i+1}] &= \frac{1}{h} \delta f(x_{i+1/2}) \\
f[x_i, x_{i+1}, x_{i-1}] &= \frac{1}{2! h^2} \delta^2 f(x_i) \\
f[x_i, x_{i+1}, x_{i-1}, x_{i+2}] &= \frac{1}{3! h^3} \delta^3 f(x_{i+1/2}) \\
&\vdots
\end{aligned}$$

となる。これらを先のニュートン補間多項式に代入し，無次元変数 α

$$x = x_i + \alpha h \quad (\text{E.40})$$

を導入すると，以下のように簡単な形になる。補間点数が，(1) 偶数個か，(2) 奇数個か，により補間式の最終項と誤差項の形が異なってくることに注意しよう。

$$\begin{aligned}
&f(x_i + \alpha h) \\
&= f(x_i) + \alpha \delta f(x_{i+1/2}) + \frac{\alpha(\alpha-1)}{2!} \delta^2 f(x_i) \\
&\quad + \frac{\alpha(\alpha^2-1)}{3!} \delta^3 f(x_{i+1/2}) + \frac{\alpha(\alpha^2-1)(\alpha-2)}{4!} \delta^4 f(x_i) + \dots \\
&+ (\text{補間の最終項}) \left\{ \begin{array}{l} (1) \frac{\alpha(\alpha^2-1) \dots (\alpha^2 - (m-1)^2)}{(2m-1)!} \delta^{2m-1} f(x_{i+1/2}) \\ (2) \frac{\alpha(\alpha^2-1) \dots (\alpha^2 - (m-1)^2)(\alpha-m)}{(2m)!} \delta^{2m} f(x_i) \end{array} \right\} \\
&+ (\text{誤差項}) \left\{ \begin{array}{l} (1) \frac{h^{2m} \alpha(\alpha^2-1) \dots (\alpha^2 - (m-1)^2)(\alpha-m)}{(2m)!} \frac{d^{2m} f(\xi)}{dx^{2m}} \\ (2) \frac{h^{2m+1} \alpha(\alpha^2-1) \dots (\alpha^2 - m^2)}{(2m+1)!} \frac{d^{2m+1} f(\xi)}{dx^{2m+1}} \end{array} \right\} \\
&\hspace{15em} (\text{E.41})
\end{aligned}$$

式 (E.41) は実質的にはニュートン補間公式の一変形に過ぎないが，特にガウス (Gauss) の前進補間公式と呼ばれる。今度は補間点を $x_i, x_{i-1}, x_{i+1}, x_{i-2}, x_{i+2}, \dots$ の順序にとっていくとき，ニュートン補間は

$$\begin{aligned}
& f(x_i + \alpha h) \\
&= f(x_i) + \alpha \delta f(x_{i-1/2}) + \frac{\alpha(\alpha+1)}{2!} \delta^2 f(x_i) \\
&\quad + \frac{\alpha(\alpha^2-1)}{3!} \delta^3 f(x_{i-1/2}) + \frac{\alpha(\alpha^2-1)(\alpha+2)}{4!} \delta^4 f(x_i) + \cdots \\
&+ (\text{補間の最終項}) \left\{ \begin{array}{l} (1) \frac{\alpha(\alpha^2-1) \cdots (\alpha^2 - (m-1)^2)}{(2m-1)!} \delta^{2m-1} f(x_{i-1/2}) \\ (2) \frac{\alpha(\alpha^2-1) \cdots (\alpha^2 - (m-1)^2)(\alpha+m)}{(2m)!} \delta^{2m} f(x_i) \end{array} \right\} \\
&+ (\text{誤差項}) \left\{ \begin{array}{l} (1) \frac{h^{2m} \alpha(\alpha^2-1) \cdots (\alpha^2 - (m-1)^2)(\alpha+m)}{(2m)!} \frac{d^{2m} f(\xi)}{dx^{2m}} \\ (2) \frac{h^{2m+1} \alpha(\alpha^2-1) \cdots (\alpha^2 - m^2)}{(2m+1)!} \frac{d^{2m+1} f(\xi)}{dx^{2m+1}} \end{array} \right\}
\end{aligned} \tag{E.42}$$

となり，これはガウスの後退補間公式と呼ばれる。

ガウスの公式 (E.41)，(E.42) の平均をとり奇数階の差分を平均演算子を用いて表すと，以下に示すスターリング (Stirling) の補間公式が得られる。スターリングの公式では，(1) 補間公式が奇数階の差分の項までで終わっている場合と，(2) 偶数階の差分の項までの場合，とで誤差項は異なる。

$$f(x_i + \alpha h) = F(x) - \varepsilon(x) \tag{E.43}$$

(1)の場合

$$\begin{aligned}
& F(x) \\
&= f(x_i) + \alpha \mu \delta f(x_i) + \frac{\alpha^2}{2!} \delta^2 f(x_i) + \frac{\alpha(\alpha^2-1)}{3!} \mu \delta^3 f(x_i) \\
&+ \frac{\alpha^2(\alpha^2-1)}{4!} \delta^4 f(x_i) + \cdots + \frac{\alpha(\alpha^2-1) \cdots (\alpha^2 - (m-1)^2)}{(2m-1)!} \mu \delta^{2m-1} f(x_i)
\end{aligned} \tag{E.44}$$

$$\begin{aligned}
\varepsilon(x) &= -\frac{h^{2m} \alpha(\alpha^2-1) \cdots (\alpha^2 - (m-1)^2)}{2(2m)!} \\
&\quad \times \left((\alpha-m) \frac{d^{2m} f(\xi_1)}{dx^{2m}} + (\alpha+m) \frac{d^{2m} f(\xi_2)}{dx^{2m}} \right)
\end{aligned} \tag{E.45}$$

(2)の場合

$$\begin{aligned}
F(x) &= f(x_i) + \alpha\mu\delta f(x_i) + \frac{\alpha^2}{2!}\delta^2 f(x_i) + \frac{\alpha(\alpha^2-1)}{3!}\mu\delta^3 f(x_i) \\
&+ \frac{\alpha^2(\alpha^2-1)}{4!}\delta^4 f(x_i) + \cdots + \frac{\alpha^2(\alpha^2-1)\cdots(\alpha^2-(m-1)^2)}{(2m)!}\delta^{2m} f(x_i)
\end{aligned} \tag{E.46}$$

$$\varepsilon(x) = -\frac{h^{2m+1}\alpha(\alpha^2-1)\cdots(\alpha^2-m^2)}{(2m+1)!} \frac{d^{2m+1}f(\xi)}{dx^{2m+1}} \tag{E.47}$$

ここに

$$\begin{aligned}
\mu\delta^{2k-1}f(x_i) &= \frac{1}{2}(\delta^{2k-1}f(x_{i+1/2}) + \delta^{2k-1}f(x_{i-1/2})) \\
&(k = 1, 2, \dots, m)
\end{aligned} \tag{E.48}$$

であり, ξ_1, ξ_2, ξ はそれぞれの補間点 $x_i, x_{i\pm 1}, x_{i\pm 2}, \dots$ および x を含む最小区間中にある値である。誤差項を含めた式 (E.43) を x で微分すると, 数値微分公式が得られる。例えば (2) の場合, $d\alpha/dx = 1/h$ より

$$\begin{aligned}
f'(x_i + \alpha h) &= \frac{1}{h}\mu\delta f(x_i) + \frac{\alpha}{h}\delta^2 f(x_i) + \frac{3\alpha^2-1}{6h}\mu\delta^3 f(x_i) + \frac{2\alpha^3-\alpha}{12h}\delta^4 f(x_i) + \cdots \\
&+ \frac{1}{h(2m)!} \frac{d}{d\alpha} (\alpha^2(\alpha^2-1)\cdots(\alpha^2-(m-1)^2)) \delta^{2m} f(x_i) \\
&+ \frac{h^{2m}}{(2m+1)!} \frac{d}{d\alpha} (\alpha(\alpha^2-1)\cdots(\alpha^2-m^2)) \frac{d^{2m+1}f(\xi)}{dx^{2m+1}}
\end{aligned} \tag{E.49}$$

となる。例えば $\alpha = 0$ とすると 1 階導関数の中心差分近似公式が得られる。

E.3 微分演算子の差分演算子表現

一般的な数値微分の公式を導くために, 微分演算子 D

$$Df(x) \equiv \frac{df(x)}{dx} \tag{E.50}$$

を用いる。テイラー級数展開

$$f(x+h) = f(x) + \frac{h}{1!}f'(x) + \frac{h^2}{2!}f''(x) + \frac{h^3}{3!}f'''(x) + \frac{h^4}{4!}f^{(4)}(x) + \dots \quad (\text{E.51})$$

は、この微分演算子 D と移相演算子 E を用いて

$$Ef(x) = \left(1 + \frac{hD}{1!} + \frac{h^2D^2}{2!} + \dots\right) f(x) = e^{hD}f(x) \quad (\text{E.52})$$

と書けるので、つぎの演算子の関係が得られる。

$$E = e^{hD} \quad (\text{E.53})$$

式 (E.53) に、前進差分演算子 Δ との関係 (E.7) を用いると

$$D = \frac{1}{h} \log E = \frac{1}{h} \log (1 + \Delta) \quad (\text{E.54})$$

となるので、1 階導関数 $f'(x_i)$ は、演算子により記号的につぎのように展開される。

$$\begin{aligned} f'(x_i) &= Df(x_i) = \frac{1}{h} (\log (1 + \Delta)) f(x_i) \\ &= \frac{1}{h} \left(\Delta - \frac{1}{2} \Delta^2 + \frac{1}{3} \Delta^3 - \frac{1}{4} \Delta^4 + \dots \right) f(x_i) \end{aligned} \quad (\text{E.55})$$

これが 1 階導関数の前進差分近似の一般形である。これは、ニュートンの前進補間公式に対する 1 階導関数 (E.31) において $\alpha=0$ としたものと同じである。

r 階導関数 $f^{(r)}(x_i)$ の前進差分近似の一般形は

$$\begin{aligned} f^{(r)}(x_i) &= D^r f(x_i) = \frac{1}{h^r} (\log (1 + \Delta))^r f(x_i) \\ &= \frac{1}{h^r} \left(\Delta^r - \frac{r}{2} \Delta^{r+1} + \frac{r(3r+5)}{24} \Delta^{r+2} - \dots \right) f(x_i) \end{aligned} \quad (\text{E.56})$$

と表せる。

後退差分演算子 ∇ については関係 (E.8) を用いると、式 (E.53) より

$$D = -\frac{1}{h} \log E^{-1} = -\frac{1}{h} \log (1 - \nabla) \quad (\text{E.57})$$

となるので, 同様にして 1 階導関数と r 階導関数は以下のように表せる。

$$\begin{aligned} f'(x_i) &= Df(x_i) = -\frac{1}{h} (\log (1 - \nabla)) f(x_i) \\ &= \frac{1}{h} \left(\nabla + \frac{1}{2} \nabla^2 + \frac{1}{3} \nabla^3 + \frac{1}{4} \nabla^4 + \dots \right) f(x_i) \end{aligned} \quad (\text{E.58})$$

$$\begin{aligned} f^{(r)}(x_i) &= D^r f(x_i) = \left(-\frac{1}{h} \log (1 - \nabla) \right)^r f(x_i) \\ &= \frac{1}{h^r} \left(\nabla^r + \frac{r}{2} \nabla^{r+1} + \frac{r(3r+5)}{24} \nabla^{r+2} + \dots \right) f(x_i) \end{aligned} \quad (\text{E.59})$$

この $f'(x)$ に関する公式は, ニュートンの後退補間公式に対する 1 階導関数 (E.38) において $\alpha=0$ としたものと同じである。

中心差分演算子 δ については関係 (E.9) を用いると, 式 (E.53) より

$$\delta = E^{1/2} - E^{-1/2} = e^{hD/2} - e^{-hD/2} = 2 \sinh \frac{hD}{2} \quad (\text{E.60})$$

となるので

$$D = \frac{2}{h} \sinh^{-1} \frac{\delta}{2} \quad (\text{E.61})$$

を得る。これより微分演算子 D は δ を用いて記号的に

$$D = \frac{2}{h} \sinh^{-1} \frac{\delta}{2} = \frac{1}{h} \left(\delta - \frac{1^2}{2^2 \cdot 3!} \delta^3 + \frac{1^2 \cdot 3^2}{2^4 \cdot 5!} \delta^5 - \dots \right) \quad (\text{E.62})$$

と展開されるが, δ の奇数次の差分は半整数の添字をもつ x の関数値でしか表現されないため, 奇数階の導関数に対しては適切ではない。偶数階の導関数は

$$f''(x_i) = D^2 f(x_i) = \frac{1}{h^2} \left(\delta^2 - \frac{1}{12} \delta^4 + \frac{1}{90} \delta^6 - \frac{1}{560} \delta^8 + \dots \right) f(x_i) \quad (\text{E.63})$$

$$\begin{aligned} f^{(4)}(x_i) &= D^4 f(x_i) = \frac{1}{h^4} \left(\delta^4 - \frac{1}{6} \delta^6 + \frac{7}{240} \delta^8 - \dots \right) f(x_i) \quad (\text{E.64}) \\ &\vdots \end{aligned}$$

と表せる。奇数階の導関数に対しては、平均演算子 (E.12) は演算子 δ を用いて

$$\mu = \frac{1}{2}(E^{1/2} + E^{-1/2}) = \left(1 + \frac{1}{4}\delta^2\right)^{1/2} \quad (\text{E.65})$$

と変形されることから、演算子 (E.61) を以下のように書き換えて用いる。

$$D_\mu = \frac{2\mu}{h} \frac{\sinh^{-1} \frac{\delta}{2}}{\sqrt{1 + \frac{1}{4}\delta^2}} = \frac{\mu}{h} \left(\delta - \frac{1^2}{3!}\delta^3 + \frac{1^2 \cdot 2^2}{5!}\delta^5 - \dots \right) \quad (\text{E.66})$$

D の上記表現と偶数階の導関数 (E.63), (E.64), \dots を考慮すれば、奇数階の導関数はつぎのように表せる。

$$\begin{aligned} f'(x_i) &= Df(x_i) = D_\mu f(x_i) \\ &= \frac{\mu}{h} \left(\delta - \frac{1}{6}\delta^3 + \frac{1}{30}\delta^5 - \frac{1}{140}\delta^7 + \dots \right) f(x_i) \end{aligned} \quad (\text{E.67})$$

$$\begin{aligned} f'''(x_i) &= D^3 f(x_i) = D_\mu D^2 f(x_i) \\ &= \frac{\mu}{h^3} \left(\delta^3 - \frac{1}{4}\delta^5 + \frac{7}{120}\delta^7 - \dots \right) f(x_i) \end{aligned} \quad (\text{E.68})$$

$$\begin{aligned} f^{(5)}(x_i) &= D^5 f(x_i) = D_\mu D^4 f(x_i) \\ &= \frac{\mu}{h^5} \left(\delta^5 - \frac{1}{3}\delta^7 + \dots \right) f(x_i) \end{aligned} \quad (\text{E.69})$$

\vdots

この $f'(x)$ に関する公式は、スターリングの補間公式に対する 1 階導関数 (E.49) において $\alpha=0$ としたものと同一である。ここで δ の奇数次の項は

$$\mu\delta f(x_i) = \frac{1}{2}(E^1 - E^{-1})f(x_i) = \frac{1}{2}(f(x_{i+1}) - f(x_{i-1})) \quad (\text{E.70})$$

$$\begin{aligned} \mu\delta^3 f(x_i) &= \delta^2(\mu\delta f(x_i)) = (E^1 - 2 + E^{-1})\frac{1}{2}(f(x_{i+1}) - f(x_{i-1})) \\ &= \frac{1}{2}(f(x_{i+2}) - 2f(x_{i+1}) + 2f(x_{i-1}) - f(x_{i-2})) \end{aligned} \quad (\text{E.71})$$

$$\begin{aligned} \mu\delta^5 f(x_i) &= \delta^2(\mu\delta^3 f(x_i)) \\ &= (E^1 - 2 + E^{-1})\frac{1}{2}(f(x_{i+2}) - 2f(x_{i+1}) + 2f(x_{i-1}) - f(x_{i-2})) \\ &= \frac{1}{2}(f_{i+3} - 4f(x_{i+2}) + 5f(x_{i+1}) - 5f(x_{i-1}) + 4f(x_{i-2}) - f_{i-3}) \end{aligned} \quad (\text{E.72})$$

など, 平均演算子 μ により整数番目の x の関数値で表されていることに留意。

引用・参考文献

- 1) 一松 信: 数値解析, 新数学講座 13, 朝倉書店 (1982)
- 2) 赤坂 隆: 数値計算, 応用数学講座第 7 巻, コロナ社 (1976)

付録 B の問題解答

問題 1 (ベクトルノルムの具体表現)

∞ ノルム: $\|\mathbf{x}\|_\infty = \max_i |x_i|$ の証明

$\gamma = \max_i |x_i|$ とおけば

$$\|\mathbf{x}\|_\infty = \lim_{p \rightarrow \infty} \left(\sum_{i=1}^n |x_i|^p \right)^{1/p} = \lim_{p \rightarrow \infty} \gamma \left(\sum_{i=1}^n \left| \frac{x_i}{\gamma} \right|^p \right)^{1/p}$$

となるが, ここに $|x_i/\gamma| \leq 1$ である。

$$\begin{aligned} \left| \frac{x_i}{\gamma} \right| < 1 \text{ のとき } \quad \lim_{p \rightarrow \infty} \left| \frac{x_i}{\gamma} \right|^p &= 0 \\ \left| \frac{x_i}{\gamma} \right| = 1 \text{ のとき } \quad \lim_{p \rightarrow \infty} \left| \frac{x_i}{\gamma} \right|^p &= 1 \end{aligned}$$

$|x_i/\gamma| = 1$ となる項の個数を k 個 ($1 \leq k \leq n$) とおけば,

$$\|\mathbf{x}\|_\infty = \gamma \lim_{p \rightarrow \infty} (k)^{1/p} = \gamma = \max_i |x_i|$$

問題 2 (行列ノルムの性質)

(1) $\|A\mathbf{x}\| \leq \|A\|\|\mathbf{x}\|$ の証明

$$\|A\| = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} \quad \text{より} \quad \|A\| \geq \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} \quad (\forall \mathbf{x} \neq \mathbf{0})$$

したがって

$$\|A\|\|\mathbf{x}\| \geq \|A\mathbf{x}\|$$

(2) $\|AB\| \leq \|A\|\|B\|$ の証明

(1) を繰り返し用いると

$$\|(AB)\mathbf{x}\| = \|A(B\mathbf{x})\| \leq \|A\|\|B\mathbf{x}\| \leq \|A\|\|B\|\|\mathbf{x}\|$$

したがって

$$\frac{\|(AB)\mathbf{x}\|}{\|\mathbf{x}\|} \leq \|A\|\|B\| \quad (\forall \mathbf{x} \neq \mathbf{0})$$

$$\therefore \|AB\| = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|(AB)\mathbf{x}\|}{\|\mathbf{x}\|} \leq \|A\|\|B\|$$

(3) すべての A に対して $\|A\| \geq 0$ (等号は $A = 0$ のときに限り成立) の証明

$\|A\| = \sup_{\mathbf{x} \neq \mathbf{0}} (\|A\mathbf{x}\|/\|\mathbf{x}\|)$ は非負の数の上限だから $\|A\| \geq 0$ であることは明らか。

等号は $A = 0$ のときに限り成立することは、以下のように示される。

$$\begin{aligned} A = 0 &\iff A\mathbf{x} = \mathbf{0} \iff \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} = 0 \quad (\forall \mathbf{x} \neq \mathbf{0}) \\ &\iff \|A\| = 0 \end{aligned}$$

(4) $\|\alpha A\| = |\alpha|\|A\|$ の証明

ベクトルノルムの性質 ii) より, $\|\alpha A\mathbf{x}\| = |\alpha|\|A\mathbf{x}\|$ であるから

$$\begin{aligned} \|\alpha A\| &= \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|\alpha A\mathbf{x}\|}{\|\mathbf{x}\|} = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{|\alpha|\|A\mathbf{x}\|}{\|\mathbf{x}\|} \\ &= |\alpha| \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} = |\alpha|\|A\| \end{aligned}$$

(5) $\|A + B\| \leq \|A\| + \|B\|$ の証明

$$\begin{aligned} \|(A + B)\mathbf{x}\| &= \|A\mathbf{x} + B\mathbf{x}\| \\ &\leq \|A\mathbf{x}\| + \|B\mathbf{x}\| \quad \leftarrow \text{ベクトルノルムの性質 iii) より} \\ &\leq \|A\|\|\mathbf{x}\| + \|B\|\|\mathbf{x}\| \quad \leftarrow \text{行列ノルムの性質 (1) より} \end{aligned}$$

したがって、次式を得る。

$$\frac{\|(A + B)\mathbf{x}\|}{\|\mathbf{x}\|} \leq \|A\| + \|B\| \quad (\forall \mathbf{x} \neq \mathbf{0})$$

$$\therefore \|A + B\| = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|(A + B)\mathbf{x}\|}{\|\mathbf{x}\|} \leq \|A\| + \|B\|$$

問題 3 (ナチュラルノルムの具体表現 1)

(1) 1 ノルム: $\|A\|_1 = \max_k \sum_i |a_{ik}|$ の証明

$$\begin{aligned} \|A\mathbf{x}\|_1 &= \sum_{i=1}^n \left| \sum_{j=1}^n a_{ij}x_j \right| \leq \sum_{i=1}^n \sum_{j=1}^n |a_{ij}||x_j| = \sum_{j=1}^n \left(|x_j| \sum_{i=1}^n |a_{ij}| \right) \\ &\leq \left(\sum_{j=1}^n |x_j| \right) \left(\max_k \sum_{i=1}^n |a_{ik}| \right) = \|\mathbf{x}\|_1 \max_k \sum_{i=1}^n |a_{ik}| \end{aligned}$$

上式の等号は、 \mathbf{x} として単位ベクトル e_k をとれば成立する。ただし k に $\sum_{i=1}^n |a_{ik}|$ が最大となる k の値を用いる。これより

$$\|A\|_1 = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|A\mathbf{x}\|_1}{\|\mathbf{x}\|_1} = \max_k \sum_i |a_{ik}|$$

(2) ∞ ノルム : $\|A\|_\infty = \max_i \sum_k |a_{ik}|$ の証明

$$\begin{aligned} \|A\mathbf{x}\|_\infty &= \max_i \left| \sum_{k=1}^n a_{ik}x_k \right| \leq \max_i \left(\sum_{k=1}^n |a_{ik}| |x_k| \right) \\ &\leq \|\mathbf{x}\|_\infty \max_i \sum_{k=1}^n |a_{ik}| \end{aligned}$$

上式の等号は、 \mathbf{x} として $x_k = |a_{ik}|/a_{ik}$ なる成分をもつベクトルをとれば成立する。ただし i には $\sum_{k=1}^n |a_{ik}|$ が最大となる i の値を用いる。これより

$$\|A\|_\infty = \max_i \sum_k |a_{ik}|$$

問題 4 (ナチュラルノルムの具体表現 2)

2 ノルム : $\|A\|_2 = \sqrt{\rho(A^*A)}$ の証明 (特に A がエルミート行列のときは $\|A\|_2 = \rho(A)$)

$B = A^*A$ はエルミート行列 ($B^* = B$) であるから、適当なユニタリ行列 T ($T^*T = I$) を用いた相似変換により、以下のように対角化できる。

$$T^*BT = D$$

相似変換に関して固有値は不変であるから、 D は B の固有値を対角成分にもつ対角行列である。 B の固有値を μ_i とすると、 D は以下のように書ける。

$$D = \begin{bmatrix} \mu_1 & & & 0 \\ & \mu_2 & & \\ & & \ddots & \\ 0 & & & \mu_n \end{bmatrix}$$

ここで B の固有値 μ_i に属する固有ベクトルを \mathbf{x}_i とすると、 $\mathbf{x}_i^*B\mathbf{x}_i = \mu_i\mathbf{x}_i^*\mathbf{x}_i$ となるので

$$\mu_i = \frac{\mathbf{x}_i^*B\mathbf{x}_i}{\mathbf{x}_i^*\mathbf{x}_i} = \frac{\mathbf{x}_i^*A^*A\mathbf{x}_i}{\mathbf{x}_i^*\mathbf{x}_i} = \frac{\|A\mathbf{x}_i\|^2}{\|\mathbf{x}_i\|^2}$$

より $\mu_i \geq 0$ である。さて、任意のベクトル $\mathbf{x} \neq \mathbf{0}$ に対して

$$\frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \frac{\sqrt{\mathbf{x}^*A^*A\mathbf{x}}}{\sqrt{\mathbf{x}^*\mathbf{x}}} = \frac{\sqrt{\mathbf{x}^*B\mathbf{x}}}{\sqrt{\mathbf{x}^*\mathbf{x}}}$$

となるが, $\mathbf{y} = T^* \mathbf{x}$ とおけば \mathbf{y} も任意のベクトルとなり, $\mathbf{x} = T\mathbf{y}$ を用いて

$$\frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \frac{\sqrt{\mathbf{y}^* T^* B T \mathbf{y}}}{\sqrt{\mathbf{y}^* T^* T \mathbf{y}}} = \frac{\sqrt{\mathbf{y}^* D \mathbf{y}}}{\sqrt{\mathbf{y}^* \mathbf{y}}} \leq \frac{\sqrt{\sum_i \mu_i |y_i|^2}}{\sqrt{\sum_i |y_i|^2}} = \sqrt{\max_i \mu_i}$$

を得る。もし μ_k が最大固有値ならば, \mathbf{y} として成分 y_i が δ_{ik} であるような単位ベクトル e_k をとれば, 等号は成立する。したがって

$$\|A\|_2 = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \sqrt{\max_i \mu_i} = \sqrt{\rho(A^* A)}$$

もし A がエルミート行列であれば, A の固有値 λ_i に対する固有ベクトルを \mathbf{x}_i とすると

$$A^* A \mathbf{x}_i = A A \mathbf{x}_i = A(\lambda_i \mathbf{x}_i) = \lambda_i^2 \mathbf{x}_i$$

より $A^* A$ の固有値は λ_i^2 となるので, 次式が導かれる。

$$\|A\|_2 = \max_i |\lambda_i| = \rho(A)$$