

プログラムのページ

78-05 連分数近似の最良化プログラム

浜田 穂積*

1. はじめに

関数の近似式の形として有理式を用いる場合の、最良近似式の定数を求めるには、Padé 展開の係数の補正を行うより、連分数の係数の補正を行う方が、計算時間、得られる定数の精度のいずれの点でも有利であることを既に示した¹⁾。ここでは一例として、正接関数 (tan) の定数を計算するプログラムを示す。

2. 計算手順

最良化しようとする関数は次の形の連分数で表わせるものとする。

$$F(x) = \frac{x}{a_1 + \frac{b_1 x^2}{a_2 + \frac{b_2 x^2}{a_3 + \dots + \frac{b_{n-1} x^2}{a_n}}}} \quad (1)$$

そして、次の形の近似式を求める。

$$G(x) = \frac{x}{c_1 + \frac{b_1 x^2}{c_2 + \frac{b_2 x^2}{c_3 + \dots + \frac{b_{n-1} x^2}{c_n}}} \quad (2)$$

これは $c_i = a_i + d_i$ とするとき、 a_i に対する補正量 d_i を定めて最良化する。tan x の場合、次の通りである。

$$a_i = 2i - 1, \quad b_i = -1 \quad (3)$$

さて、ここで $f(x)$, $g(x)$ を次により定義する。

$$\left. \begin{aligned} f(x) &= a_1 + \frac{b_1 x^2}{a_2 + \frac{b_2 x^2}{a_3 + \dots + \frac{b_{n-1} x^2}{a_n}}}} \\ g(x) &= c_1 + \frac{b_1 x^2}{c_2 + \frac{b_2 x^2}{c_3 + \dots + \frac{b_{n-1} x^2}{c_n}}} \end{aligned} \right\} \quad (4)$$

$f(x)$, $g(x)$ は共に偶関数である。そこで、ここでは相対誤差に関する最良化を例としてとりあげる。相対誤差関数を $E(x)$ とすると、

$$\begin{aligned} E(x) &= \frac{G(x) - F(x)}{F(x)} = \frac{x/g(x) - x/f(x)}{x/f(x)} \\ &= \frac{f(x) - g(x)}{g(x)} \end{aligned} \quad (5)$$

であり、これも偶関数である。 f_i , g_i を以下で定義する。

$$f_i = a_i + \frac{b_i x^2}{f_{i+1}} \quad (6)$$

$$g_i = a_i + d_i + \frac{b_i x^2}{g_{i+1}}, \quad g_n = a_n + d_n \quad (7)$$

このとき

$$f(x) = f_1, \quad g(x) = g_1 \quad (8)$$

である。これらを (5) に代入して次を得る。

$$\begin{aligned} E(x) &= -\frac{1}{g_1} \left(d_1 - \frac{b_1 x^2}{f_2 g_2} \left(d_2 - \dots \right. \right. \\ &\quad \left. \left. - \frac{b_{n-1} x^2}{f_n g_n} \left(d_n - \frac{b_n x^2}{f_{n+1}} \right) \dots \right) \right) \end{aligned} \quad (9)$$

これは次の漸化式で計算する。

$$\left. \begin{aligned} u_{n+1} &= -1 \\ u_i &= -\frac{1}{g_i} \left(d_i + \frac{b_i x^2 u_{i+1}}{f_{i+1}} \right) \quad (i = n, n-1, \dots, 1) \\ E(x) &= u_1 \end{aligned} \right\} \quad (10)$$

また、 d_i を得るため、 d_i に対する補正量 dd_i を計算してそれで d_i を修正する反復法を用いるという通常行われる方法を用いる。このため、偏差点の x 座標に関して、次が成立するための dd_i を求める。

$$\left. \begin{aligned} E(x_{j-1}, \dots, d_i + dd_i) + E(x_j, \dots, d_i + dd_i, \dots) &= 0 \\ (j = 1, 2, \dots, n) \end{aligned} \right\} \quad (11)$$

このとき $d_i + dd_i$ を新しい d_i と置いて (11) を反復し、 d_i が収束すれば、最良近似式が得られる。このため (11) を、 dd_i を微小量と考えて線型化し、 dd_i を解く。このとき必要な $\partial E / \partial d_i$ は次で与えられる。

$$\frac{\partial E}{\partial d_i} = (-1)^i f_1 \frac{b_1 b_2 \dots b_{i-1} x^{2i-2}}{(g_1 g_2 \dots g_i)^2} \quad (12)$$

以上の関係を用いて次の手順で計算する。

(A) d_i の初期値を求める。それには (9) の $E(x)$ が

$$x_j = \rho \cos \left(\frac{2j-1}{2n} \cdot \frac{\pi}{2} \right) \quad (j = 1, 2, \dots, n)$$

において 0 になるとして、 n 元連立方程式を解く。ただし、 g_i の計算においては d_i がまだ計算されていないので、 $d_i = 0$ すなわち $g_i = f_i$ とする。また ρ は収束半径、すなわち近似区間を $-\rho \leq x \leq \rho$ とする。

(B) $E(x)$ が極値をとる x の近似値 y_i として、

* (株)日立製作所システム開発研究所

次の値を初期値に設定する。

$$y_j = \rho \cos\left(\frac{2j}{2n} \cdot \frac{\pi}{2}\right) \quad (j=0, 1, 2, \dots, n)$$

(C) d_i の補正量 dd_i を計算して, $d_i + dd_i$ で d_i を置き換える。この手順を数回繰返す。

(D) 極値を与える x (y_i で表わす) を, その近似値の近くにある正しい値に修正する。ここでは近似値と, それをはさむ2点との3点を通る放物線の頂点の x 座標を新たな x 座標とする, という過程を反復する方法を用いている。

(E) 極値の絶対値のバラツキを調べ, 収束したと見なせるようになるまで (C), (D) を繰返す。

3. プログラム

上に述べた手順を行うプログラムの概略を図-1に, 詳細を図-2に示す。これは HITAC 8000 シリーズの EDOS/EDOS-MSO ALGOL で書かれたものである。この ALGOL は JIS 7070 相当であるが, このプログラムは JIS 4060 の機能で書かれている。

このプログラムは (1) の形の連分数の係数に対する補正量 d_i を計算して出力する。主プログラムで a_i, b_i と ρ を与える。ここでは

$$a_i = 2i - 1, \quad b_i = -1, \quad \rho = \pi/4$$

と与えて $\tan x$ の相対誤差に関する最良化を, $n=2 \sim 8$ について行う。

手続き INITDY は手順 (A) と (B) を行う。

手続き REFINE は手順 (C) と (D) を行う。手順 (C) の連立方程式の係数と右辺の符号をすべて逆にしてある。また関数 ER は (9) の $E(x)$ の値を計算するものであるが, $m=20$ とした。これらは単に便宜上の理由からだけである。(C) の繰返しは固定的

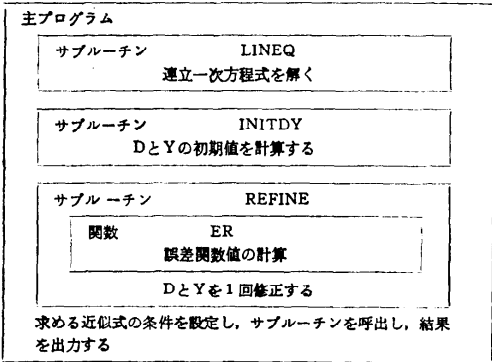


図-1 プログラムの構造

に8回としたが, これで十分であった。次に手順(D)については Muller 法の変形を用いた。すなわち, 3点を通る2次式を想定し, その頂点の x 座標を求め, 新しい3点を得る, このとき得られた点の両側に1点ずつ, 前の3点の中から選ぶ。もしこの条件が満たされなくなったら収束したと見なし, その前に得られた

```

001 BEGIN
002   INTEGER I,N;
003   REAL RADIUS,MAXERROR,CONVERGENCY;
004   REAL ARRAY A,B(1:20),C,D(1:13),E,Y(0:13);
005   PROCEDURE LINEQ(N,Z);
006     VALUE N;
007     INTEGER N;
008     REAL ARRAY Z;
009   BEGIN
010     INTEGER I,J,K;
011     FOR J:=1 STEP 1 UNTIL N DO
012       FOR I:=J+1 STEP 1 UNTIL N+1 DO
013         BEGIN
014           Z(J,I):=Z(J,I)/Z(J,J);
015           FOR K:=J+1 STEP 1 UNTIL N DO
016             Z(K,I):=Z(K,I)-Z(K,J)*Z(J,I)
017         END ELIMINATION;
018     FOR J:=N-1 STEP -1 UNTIL 1 DO
019       FOR I:=J+1 STEP 1 UNTIL N DO
020         Z(J,N+1):=Z(J,N+1)-Z(J,I)*Z(I,N+1)
021   END LINEQ;
022   PROCEDURE INITDY(R,N,A,B,D,E,Y);
023     VALUE R,N;
024     INTEGER N;
025     REAL R;
026     REAL ARRAY A,B,D,E,Y;
027   BEGIN
028     INTEGER I,J;
029     REAL HP1,S,X;
030     REAL ARRAY F(2:N+1),Z(1:N+1:N+1);
031     HP1:=1.5707963267948966;
032     FOR J:=1 STEP 1 UNTIL N DO
033       BEGIN
034         X:=R*RCOS((2*NJ-1)/(2*N)*HP1);
035         S:=A(20);
036         FOR I:=19 STEP -1 UNTIL N+2 DO
037           S:=B(I)*X*X/S+A(I);
038         FOR I:=N+1 STEP -1 UNTIL 2 DO
039           F(I):=S:=B(I)*X*X/S+A(I);
040         Z(J,1):=1;
041         FOR I:=1 STEP 1 UNTIL N DO
042           Z(J,I+1):=-Z(J,I)*B(I)*X*X/F(I+1)*N2;
043           Z(J,N+1):=-Z(J,N+1)*F(N+1)
044         END;
045       LINEQ(N,Z);
046       FOR J:=1 STEP 1 UNTIL N DO
047         D(J):=Z(J,N+1);
048       FOR J:=0 STEP 1 UNTIL N-1 DO
049         Y(J):=R*RCOS((2*NJ)/(2*N)*HP1);
050       Y(N):=0
051   END INITDY;
052   PROCEDURE REFINE(N,A,B,C,D,E,Y,ME,CONV);
053     VALUE N;
054     INTEGER N;
055     REAL ME,CONV;
056     REAL ARRAY A,B,C,D,E,Y;
057   BEGIN
058     INTEGER I,J,K;
059     REAL S,T,U,V,W,X,E1,E2,E3;
060     REAL ARRAY G(1:N),Z(0:N+1:N+1);
061     REAL PROCEDURE ER(X);
062     VALUE X;
063     REAL X;
064   BEGIN
065     INTEGER I;
066     REAL S,T,U;
067     FOR I:=20 STEP -1 UNTIL N+1 DO
068       S:=(IF I=20 THEN 0
069         ELSE B(I)*X*X/S)+A(I);
070     U:=-1;
071     FOR I:=N STEP -1 UNTIL 1 DO
072       BEGIN
073         T:=(IF I=N THEN 0
074           ELSE B(I)*X*X/T)+A(I)+D(I);
075         U:=-((U*B(I)*X*X/S+D(I))/T);
076         S:=B(I)*X*X/S+A(I)
077       END;
078     ER:=U
079   END ER;
080   COMMENT CORRECT D;
  
```

(つづく)

```

081  FOR K:=1 STEP 1 UNTIL 8 DO
082  BEGIN
083    FOR J:=0 STEP 1 UNTIL N DO
084    BEGIN
085      X:=Y(J);
086      G(N):=T:=A(N)+D(N);
087      FOR I:=N-1 STEP -1 UNTIL 1 DO
088        G(I):=T:=B(I)*X*X/T+A(I)+D(I);
089      S:=A(20);
090      FOR I:=19 STEP -1 UNTIL 1 DO
091        S:=B(I)*X*X/S+A(I);
092      Z(J,1):=S/T/T;
093      FOR I:=2 STEP 1 UNTIL N DO
094        Z(J,I):=-Z(J,I-1)*B(I-1)*X*X/G(I)*X*Z;
095        Z(J,N+1):=ER(X)
096      END;
097      FOR J:=N STEP -1 UNTIL 1 DO
098        FOR I:=1 STEP 1 UNTIL N+1 DO
099          Z(J,I):=Z(J,I)+Z(J-1,I);
100      LINEQ(N,Z);
101      FOR I:=1 STEP 1 UNTIL N DO
102        D(I):=D(I)+Z(I,N+1)
103      END CORRECT D;
104      COMMENT CORRECT Y;
105      E(D):=ER(Y(0));
106      FOR J:=1 STEP 1 UNTIL N-1 DO
107      BEGIN
108        U:=(Y(J-1)+Y(J))/2; E1:=ER(U);
109        V:=(Y(J)+Y(J+1))/2; E2:=ER(V);
110        W:=(Y(J)); E3:=ER(W);
111        FOR X:=(UNUR(E2-E3)+V*VR(E3-E1)+WR*W(E1-E2))/
112          /UN(E2-E3)+VM(E3-E1)+WM(E1-E2))/2
113          WHILE (X-U)*(X-V)<0 & X>W DO
114          BEGIN
115            S:=ER(X);
116            IF ABS(S)>ABS(E3) THEN
117              BEGIN
118                IF X>W THEN BEGIN V:=W; E2:=E3 END
119                ELSE BEGIN U:=W; E1:=E3 END;
120                W:=X;
121                F3:=S
122              END ELSE
123                IF X>W THEN BEGIN U:=X; E1:=S END
124                ELSE BEGIN V:=X; E2:=S END
125            END;
126            Y(J):=W;
127            E(J):=E3
128          END;
129          E(N):=ER(Y(N));
130          S:=SIGN(E(0));
131          ME:=T:=E(0)*S;
132          FOR J:=1 STEP 1 UNTIL N DO
133            BEGIN
134              S:=-S;
135              ME:=MAX2(ME,E(J)*S);
136              T:=MIN2(T,E(J)*S)
137            END;
138          CONV:=T/ME;
139          FOR I:=1 STEP 1 UNTIL N DO
140            C(I):=LOG(ABS(Z(0,I)))/ME;
141          END REFINE;
142          RADIUS:=.78539816339744831;
143          FOR I:=1 STEP 1 UNTIL 20 DO
144            BEGIN
145              A(I):=2*PI-1;
146              B(I):=-1
147            END;
148          OUTPUT1(2*' *11B" TAN(X) R="D.7D',RADIUS);
149          OUTPUT0(2*' /" N 1"11B"D(I)"13B" MRE"');
150          FOR N:=2 STEP 1 UNTIL 13 DO
151            BEGIN
152              INITDY(RADIUS,N,A,B,D,Y);
153              I:=0; CONVERGENCY:=0;
154              FOR I:=1 WHILE CONVERGENCY<.9999 & I<=15 DO
155                REFINE(N,A,B,C,D,E,Y,MAXERROR,CONVERGENCY);
156                OUTPUT3(2*' /40B-ZD.DD=-52D.48-D.6D',
157                  -LOG(1-CONVERGENCY),I-1,E(0)/MAXERROR);
158              FOR I:=1 STEP 1 UNTIL N DO
159                BEGIN
160                  IF I=1 THEN OUTPUT1(2*' /ZD',N)
161                  ELSE OUTPUT0(2*' /6B',N)
162                  OUTPUT2(2*' -ZDB=-D.16DE+DD',I,D(I));
163                  IF I=1 THEN OUTPUT1(2*' -D.30E+DDB',MAXERROR)
164                  ELSE OUTPUT0(2*' /11B',N)
165                  OUTPUT3(2*' -ZD.DD+2(-ZD.6D)',
166                    C(I),Y(I),E(I)/MAXERROR)
167                END
168              END
169            END MINIMAXIMIZATION OF CONTINUED FRACTION

```

図-2 連分数の最良化プログラム

新しい点のx座標の値を極値をとるxの値とする。

手順(E)は主プログラムで行う。

その前にあらかじめ手続き REFIN E の中で次の値を計算しておく。

$$s = \text{sgn}(E(y_0))$$

$$e_{\max} = \max_j \{(-1)^j s E(y_j)\}$$

$$e_{\min} = \min_j \{(-1)^j s E(y_j)\}$$

$$k = e_{\min} / e_{\max}$$

そして $k > 0.9999$ となったら収束したと見なす。ただし繰返しは多くても15回までとする。図-3に出力の一部を示す。このうち(f)の最後の値は、 e_{\max} だけ関数値を変化させる d_i の値を示すものであり、 d_i の桁数を定める目安となる¹⁾。

4. おわりに

正接関数 (tan) の $|x| \leq \pi/4$ にお

TAN(X) R=0.7853982						
N	I	D(I)	MRE	(a)	(b)	(d)
2	1	1.3037929752988941E-03	1.302E-03	2.9990	0.575938	1.000000
	2	-1.2934359229009223E-01		1.66	0.000000	-1.000000
3	1	-5.7731383860934924E-06	5.773E-06	5.45	0.690019	0.999996
	2	1.3694983146858293E-03		4.22	0.609994	-1.000000
	3	-1.3667814820018976E-01		2.63	0.000000	0.999996
4	1	1.4163813770077187E-08	1.416E-08	4.95	2	-0.999989
	2	-6.0756854900322005E-06		7.95	0.730582	0.999997
	3	1.1660949257125745E-03		6.83	0.568360	-0.999989
	4	-1.40657160e2610192E-01		5.23	0.312588	1.000000
13	1	-4.1194531795307344E-38	4.120E-38	4.33	2	-0.999953
	2	1.9695814420733019E-34		37.49	0.779863	0.999964
	3	-4.349098308086438E-31		36.36	0.763312	-0.999989
	4	7.4204756818177160E-28		34.77	0.735909	0.999993
	5	-1.0904144276730556E-24		32.88	0.697941	-0.999983
	6	1.427673327703718E-21		30.77	0.649826	0.999982
	7	-1.6788108774384405E-18		26.48	0.592132	-0.999953
	8	1.7637126661460532E-15		26.04	0.525581	0.999964
	9	-1.6308953813025866E-12		23.48	0.451058	-0.999986
	10	1.2945807170060710E-09		20.81	0.369605	1.000000
	11	-8.4631619766865312E-07		18.05	0.282410	-0.999997
	12	4.3063751824866296E-04		15.19	0.190788	0.999980
	13	-1.4975671554311330E-01		12.26	0.096156	-0.999961
				9.26	0.000000	0.999953

MRE: Maximum Relative Error

(a) $-\log_{10}(1-k)$

(b) 収束計算の繰返し回数

(c) $\log_{10} \left(\max_j \left| \frac{\partial E(x_j)}{\partial d_i} \right| / e_{\max} \right)$

ただし e_{\max} は MRE の値

(d) 極値をとる x の値と、そのときの誤差 $E(x_i)/e_{\max}$

図-3 出力例

ける最良近似連分数を, $n=2\sim 13$ について計算したプログラムを示した. いずれの場合も, 収束計算の外側のループ2回で, バラツキ (k) が 0.9999 に収まった. 計算条件は特によすぎるとはいうわけではないので, この結果は計算法のよさを示していると考えられる.

プログラムは ALGOL で記述したので, 不自然な制約も少なく, 読みやすいものになった. また, いわゆる構造的プログラミングの一つの条件である **go to**

文なしのプログラムができたのも, ALGOL を用いたメリットであろう.

参 考 文 献

- 1) 浜田穂積: 有理式近似および連分数近似の最良化について, 情報処理, Vol. 19, No. 11, pp. 1065~1071 (1978).

(昭和52年4月27日受付)

(昭和53年1月17日再受付)