Dual Fractional Power Approximation for Nonlinear Functions

Yasuhiro KOBAYASHI*, Masaaki OHKITA*
and Michio INOUE*

Abstract

For accurate analog simulation of continuous functions, it is advisable to use approximations of the form:

$$f(x) \simeq \alpha_0 + \alpha_1 x + \alpha_2 x^{\beta_1} + \alpha_3 x^{\beta_2}$$

in a full range of an input variable. The exponents β_1, β_2 are given by real numbers according to the nature of f(x). The accuracy is comparable to that in approximation using Chebyshev polynomials up to the fifth-degree. Especially, for odd functions, this approximation can achieve good accuracy over an extended interval.

1. Introduction

In the usual analog simulation, uniformly continuous functions have been divided into several subintervals wherein the functions are approximated by linear segments [1], quadratic arcs[2], or their combinations[3]. Unfortunately, these systems yield considerable errors near the endpoints and between them.

An approach is to use uniform(or non-breakpoint) approximation of continuous functions over the full range of an input variable. With this policy, an economical mathematical system, termed the *single fractional power approximation*, was introduced [4],[5]. To improve that accuracy or to extend that variable-range, in this paper, a fractional power is added to that system. This system, termed the dual fractional power approximation, can improve that accuracy by about a hundred times. Since a fractional power term behaves like an ordinary power series alternating in sign, one can expect better accuracy than Chebyshev approximations for odd functions. Parameter values must be given so as to satisfy a minimax error condition, although it cannot be achieved directly. We must correct their approximate values by iteration. The starting values can be determined by a method of undetermined coefficients or that of

---

This paper first appeared in Japanese in Joho-Shori (Journal of the Information Processing Society of Japan), Vol. 17, No. 8 (1976), pp. 703~710.

* Department of Electrical Engineering, Faculty of Engineering, Tottori University

interpolation. For convenience of optimization, the parameter-determination begins with interpolation of a given function at geometrically arranged points in a specified interval. Details are as follows;

2. Parameter-determination and its optimization

Consider that a single-valued continuous function is approximated by $f(x) \simeq \alpha_0 + \alpha_1 x + \alpha_2 x^{\beta_1} + \alpha_3 x^{\beta_2}$ in the interval [0, 1]. An optimal condition is made so that the approximation satisfies null error at both endpoints and minimizes the maximum magnitude of deviations within this interval. An actual criterion is given so that the error curve possesses an equal magnitude oscillatory behavior with proper alternation in this interval. Since the minimax error magnitude μ^* and the true extreme locations cannot be evaluated straightforwardly, the minimax condition in a Chebyshev sense is modified here.

For convenience of interpolation, suppose a set of distinct points X_i $(i=1,2,3,4)$ satisfying a geometrical relation $X_i = \lambda^{(i-5)}$ with $\lambda > 1$ on the x-axis within the interval (See Fig. 1). If the optimal approximation yields certain deviations ρ_i $(i=1,2,3,4)$ at these points, the following simultaneous equation is established.

$$\left.\begin{aligned} f(0) &= \alpha_0 , \\ f(X_i) + \rho_i &= \alpha_0 + \alpha_1 \lambda^{(i-5)} + \alpha_2 \lambda^{(i-5)\beta_1} + \alpha_3 \lambda^{(i-5)\beta_2} , \\ f(1) &= \alpha_0 + \alpha_1 + \alpha_2 + \alpha_3 . \end{aligned}\right\} \quad (1)$$

With the transformation: $P = \lambda^{(1-\beta_1)}$, $Q = \lambda^{(1-\beta_2)}$, (2)

one can rearrange Eq. (1) so that

$$F_i = \alpha_2 [P^{(5-i)} - 1] + \alpha_3 [Q^{(5-i)} - 1], \quad i = 1,2,3,4, \quad (3)$$

where $F_i = \lambda^{(5-i)} [f(X_i) - f(0) + \rho_i] - [f(1) - f(0)]$.

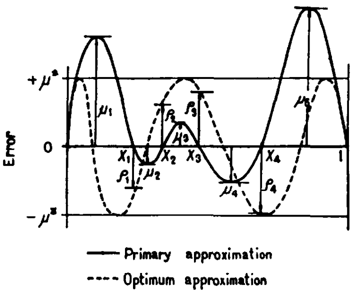

Fig. 1 Diagram of the interpolated approximations.

Moreover, upon putting

$$H_1 = F_1 - 4F_2 + 6F_3 - 4F_4 , \quad H_2 = F_2 - 3F_3 + 3F_4 , \quad H_3 = F_3 - 2F_4 , \quad H_4 = F_4 , \quad (4)$$

one can rewrite Eq. (3) as $H_i = \alpha_2 (P-1)^{(5-i)} + \alpha_3 (Q-1)^{(5-i)}$. (5)

Equation (5) is then modified so that

$$H_1 - [(P-1)+(Q-1)]H_2 + (P-1)(Q-1)H_3 = 0 , \quad H_2 - [(P-1)+(Q-1)]H_3 + (P-1)(Q-1)H_4 = 0. \quad (6)$$

Hence $$(P-1)+(Q-1)=\frac{H_2H_3-H_1H_4}{H_3^2-H_2H_4}\ , \quad (P-1)(Q-1)=\frac{H_2^2-H_1H_3}{H_3^2-H_2H_4}\ , \quad H_3^2-H_2H_4 \neq 0\ . \tag{7}$$

This means that the (P-1) value, corresponding to the fractional exponent β_1, can be evaluated as either of the two roots of a quadratic equation. Therefore,

$$P-1=(M+\sqrt{M^2-4N})/2\ , \quad Q-1=(M-\sqrt{M^2-4N})/2\ , \quad M^2-4N>0\ , \tag{8}$$

where $$M=\frac{H_2H_3-H_1H_4}{H_3^2-H_2H_4}\ , \quad N=\frac{H_2^2-H_1H_3}{H_3^2-H_2H_4}\ , \quad H_3^2-H_2H_4 \neq 0\ .$$

Then, the parameter values can be evaluated as

$$\left.\begin{aligned}
&\alpha_0 = f(0)\ , && \alpha_1 = f(1)-f(0)+(H_3-H_4M)/N\ ,\\
&\alpha_2 = \frac{2H_3-(M-\sqrt{M^2-4N})H_4}{(M+\sqrt{M^2-4N})\sqrt{M^2-4N}}\ , && \alpha_3 = \frac{-2H_3+(M+\sqrt{M^2-4N})H_4}{(M-\sqrt{M^2-4N})\sqrt{M^2-4N}}\ ,\\
&\beta_1 = -\log[(2+M+\sqrt{M^2-4N})/2\lambda]/\log\lambda\ , && \beta_2 = -\log[(2+M-\sqrt{M^2-4N})/2\lambda]/\log\lambda\ .
\end{aligned}\right\} \tag{9}$$

Note that these parameter values (except for α_0) are given as functions of ρ_i values, though difficult to evaluate analytically. For convenience, the ρ_i values are estimated in terms of neighboring extreme deviations μ_i, μ_{i+1}. At the first step of computation, all ρ_i values are given as zero; and at the second and thereafter, they are estimated so that $\rho_i = W(\mu_i + \mu_{i+i})$ with a weight W. The ρ_i values are accumulated with each step. By substituting such values into Eq.(1) iteratively, the parameter values can be modified. The modification must be repeated until all extreme deviations are leveled. Fortunately, the computation converges without oscillation and produces the optimal parameters. An experimental verification is given later.

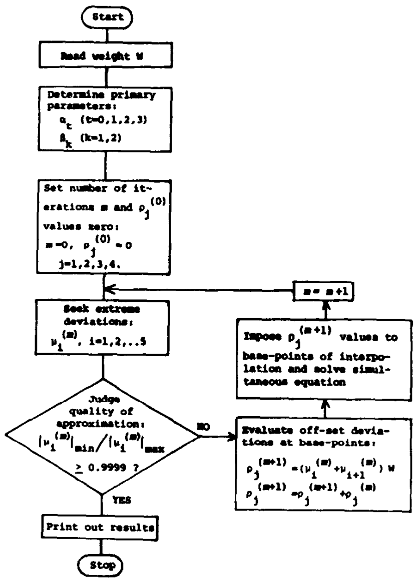

Fig. 2 Flowchart for the minimax dual fractional power approximation.

```
YJOB              OVLY,NOCC
C
C     ---------------------------------------------
C
C     MINIMAX DUAL FRACTIONAL POWER APPROXIMATION
C
C     F(X)=A0+A1*X+A2*X**P+A3*X**Q
C
C     ---------------------------------------------
C
C     N....ENTRY NUMBER OF FUNCTIONS
C          N=1....COSH(X)
C     INTERVAL (0.0 .LE. X .LE. 1.0)
C
C     ---------------------------------------------
C
      DOUBLE PRECISION FFL(4),FL(4),W
      DOUBLE PRECISION XX(5),EY(5),A(6)
      COMMON/AA/ N,AM,W,LP
      COMMON/BB/ DEB,FFL,FL
      COMMON/CC/ A
      COMMON/DD/ XX,EY
      CALL QOVFLR
      LP=0
      DO 1 I=1,4
      FL(I)=0.0D0
    1 FFL(I)=0.0D0
      CALL INPUT
    2 CALL CALCU
      CALL ERRM
      IF(LP.EQ.0) CALL OUT
      IF(DEB.GT.0.9999) GO TO 3
      LP=LP+1
      GO TO 2
    3 CALL OUT
      STOP
      END


      DOUBLE PRECISION FUNCTION FUNCA(X)
      DOUBLE PRECISION X,FUNCA,PH
      COMMON/AA/ N
      DATA PH/0.5D0/
      GO TO (1),N
    1 FUNCA=(DEXP(X)+DEXP(-X))*PH
      RETURN
      END


      DOUBLE PRECISION FUNCTION FUNCB(X)
      DOUBLE PRECISION X,FUNCB,A(6),XA,XB
      COMMON/CC/ A
      XA=X**A(5)
      XB=X**A(6)
      FUNCB=A(3)*XA+A(4)*XB+A(2)*X+A(1)
      RETURN
      END


      DOUBLE PRECISION FUNCTION ERROR(X)
      DOUBLE PRECISION X,FUNCA,FUNCB,ERROR
      ERROR=FUNCA(X)-FUNCB(X)
      RETURN
      END


      SUBROUTINE INPUT
      DOUBLE PRECISION W
      COMMON/AA/ N,AM,W
      CALL SETSEP(1H,)
      READ(10,1) N,W
      WRITE(6,3)
      CALL NAME
      WRITE(6,2) AM,W
    1 FORMAT(I3,D18.8)
    2 FORMAT(//T7,'FUNCTION=',A8,T31,'WEIGHT=',
     * D18.8)
    3 FORMAT(1H ,T7,'** MINIMAX DUAL FRACTIONAL POWER'
     * ,T40,'APPROXIMATION **')
      RETURN
      END


      SUBROUTINE NAME
      COMMON/AA/ N,AM
      DIMENSION FM(10)
      DATA FM(1)/' COSH(X)'/
      AM=FM(N)
      RETURN
      END


      SUBROUTINE CALCU
      DOUBLE PRECISION ALH,ESTD,XI(4),A(6),FL(4)
      DOUBLE PRECISION FFL(4),XM(4),XL(4),XA(2)
      DOUBLE PRECISION AOI,AIO,AII,AS,FUNCA
      DOUBLE PRECISION ZERO,ONE,TWO,THR
      DOUBLE PRECISION FOUR,SIX
      COMMON/BB/ DEB,FFL,FL
      COMMON/CC/ A
      DATA ZERO/0.0D0/,ONE/1.0D0/,TWO/2.0D0/
      DATA THR/3.0D0/,FOUR/4.0D0/,SIX/6.0D0/
      ALH=1.380277D0
      XI(4)=ONE/ALH
      XI(3)=XI(4)/ALH
      XI(2)=XI(3)/ALH
      XI(1)=XI(2)/ALH
      A(1)=FUNCA(ZERO)
      DO 1 I=1,4
    1 FL(I)=FL(I)+FFL(I)
      AS=FUNCA(ONE)-A(1)
      DO 2 I=1,4
    2 XL(I)=ALH**(5-I)*(FUNCA(XI(I))-A(1)+FL(I))-AS
      XM(1)=XL(1)-FOUR*XL(2)+SIX*XL(3)-FOUR*XL(4)
      XM(2)=XL(2)-THR*XL(3)+THR*XL(4)
      XM(3)=XL(3)-TWO*XL(4)
      XM(4)=XL(4)
      AOI=XM(2)*XM(4)-XM(3)*XM(3)
      AIO=XM(2)*XM(3)-XM(1)*XM(4)
      AII=XM(1)*XM(3)-XM(2)*XM(2)
      CALL QUAD(AOI,AIO,AII,XA,KOR)
      IF(KOR.EQ.1) GO TO 3
      IF(KOR.EQ.3) GO TO 3
      A(6)=ONE-DLOG(XA(2)+ONE)/DLOG(ALH)
      A(5)=ONE-DLOG(XA(1)+ONE)/DLOG(ALH)
      A(4)=(XM(3)-XA(1)*XM(4))/(XA(2)*(XA(2)-XA(1)))
      A(3)=(XL(4)-A(4)*XA(2))/XA(1)
      A(2)=FUNCA(ONE)-A(1)-A(3)-A(4)
      IF(KOR.EQ.2) RETURN
    3 WRITE(6,4) KOR
    4 FORMAT(1H ,T10,'KOR= ',I2)
      RETURN
      END


      SUBROUTINE QUAD(AOI,AIO,AII,XA,KOR)
C     QUADRATIC EQUATION
C     KOR=1 ..... ONE REAL ROOT
C     KOR=2 ..... TWO REAL ROOTS
C     KOR=3 ..... IMAGINARY ROOT
      DOUBLE PRECISION AOI,AIO,AII,XA(2),B,D
      DOUBLE PRECISION ZERO,FOUR
      DATA ZERO/0.0D0/,FOUR /4.0D0/
      IF(AOI.EQ.ZERO) GO TO 2
      B=AOI+AOI
      D=AIO**2-FOUR*AOI*AII
      IF(D) 1,2,3
    1 KOR=3
      RETURN
    2 KOR=1
      RETURN
    3 XA(1)=(-AIO+DSQRT(D))/B
      XA(2)=(-AIO-DSQRT(D))/B
      KOR=2
      RETURN
      END


      SUBROUTINE ERRM
      DOUBLE PRECISION X,DX,ERR1,ERR2,ERR3,ERROR
      DOUBLE PRECISION EER1,EER2,XX(5),EY(5)
      DOUBLE PRECISION EEMAX,EEMIN,A(6),EFF(4)
      DOUBLE PRECISION XAA,ZERO,PH,FFL(5),W,ESTD
      COMMON/AA/ N,AM,W,LP
      COMMON/BB/ DEB,FFL
      COMMON/CC/ A
      COMMON/DD/ XX,EY
      DATA ZERO/0.0D0/,PH/0.5D0/
      DATA ESTD/0.1D-12/
      IF(LP.GE.1) GO TO 7
      JI=0
      X=ZERO
    1 DX=0.5D-1
    2 ERR1=ERROR(X)
      ERR2=ERROR(X+DX)
      ERR3=ERROR(X+DX+DX)
      EER1=ERR2-ERR1
      EER2=ERR3-ERR2
      ERER=EER1*EER2
      IF(ERER) 4,4,3
    3 X=X+DX
      GO TO 2
    4 IF(DABS(EER2)-ESTD) 6,5,5
    5 DX=DX*PH
      GO TO 2
    6 JI=JI+1
      XX(JI)=X+DX
      XAA=XX(JI)
      EY(JI)=ERROR(XAA)
      X=XX(JI)
      IF(JI-5) 1,13,13
    7 I=1
    8 DX=0.1D-1
      X=XX(I)
    9 ERR1=ERROR(X)
      ERR2=ERROR(X+DX)
      EER1=DABS(ERR2)-DABS(ERR1)
      IF(EER1) 10,11,11
   10 X=X-DX-DX
      DX=DX*PH
      X=X+DX
      GO TO 9
   11 X=X+DX
      IF(DABS(EER1)-ESTD) 12,9,9
   12 XX(I)=X+DX*PH
      EY(I)=ERROR(XX(I))
      I=I+1
      IF(I-6) 8,13,13
   13 EEMAX=DABS(EY(1))
      EEMIN=EEMAX
      DO 17 J=2,5
      IF(EEMAX-DABS(EY(J))) 14,14,15
   14 EEMAX=DABS(EY(J))
      GO TC 17
   15 IF(EEMIN-DABS(EY(J))) 17,17,16
   16 EEMIN=DABS(EY(J))
```

```
17    CONTINUE
      DEB=(EEMIN/EEMAX)
      DO 18 I=1,4
18    FFL(I)=(EY(I)+EY(I+1))*W
      RETURN
      END

      SUBROUTINE OUT
      DOUBLE PRECISION A(6),XX(5),EY(5),W
      COMMON /AA/N,AM,W,LP
      COMMON /BB/DEB
      COMMON /CC/A
      COMMON /DD/XX,EY
      WRITE(6,1) LP,DEB
      WRITE(6,2) A(1),A(2)
      WRITE(6,3) A(3),A(4)
      WRITE(6,4) A(5),A(6)
      WRITE(6,5) (JI,XX(JI),EY(JI),JI=1,5)
1     FORMAT(//T7,'ITERATIVE CORRECTION=',I3,
     * T36,'DEB=',E15.7)
2     FORMAT(1H ,5X,3HA0=,D20.11,5X,3HA1=,D20.11)
3     FORMAT(1H ,5X,3HA2=,D20.11,5X,3HA3=,D20.11)
4     FORMAT(1H ,4X,4H P =,D20.11,4X,4H Q =,D20.11)
5     FORMAT(1H ,5X,3H X(,I2,2H)=,D17.10,5X,
     * 5HEMAX=,D17.10)
      RETURN
      END
```

```
** MINIMAX DUAL FRACTIONAL POWER APPROXIMATION **

FUNCTION= COSH(X)        WEIGHT=    0.30000000E 00

ITERATIVE CORRECTION=  0      DEB=  0.2630651E-01
A0=    0.10000000000E 01      A1=    0.39815809449E-03
A2=    0.50210241904E 00      A3=    0.40580057686E-01
P =    0.20045533890E 01      Q =    0.41429576956E 01
 X( 1)= 0.6323852539E-01       EMAX=-0.8273151373E-05
 X( 2)= 0.3190979004E 00       EMAX= 0.2176377406E-06
 X( 3)= 0.4564758301E 00       EMAX=-0.2540126625E-06
 X( 4)= 0.6469665527E 00       EMAX= 0.7825217406E-06
 X( 5)= 0.9165588379E 00       EMAX=-0.6557847031E-05

ITERATIVE CORRECTION= 29      DEB=  0.9999077E 00
A0=    0.10000000000E 01      A1=    0.21056763196E-03
A2=    0.50202744141E 00      A3=    0.40842625770E-01
P =    0.20034553277E 01      Q =    0.41399901639E 01
 X( 1)= 0.3708618164E-01       EMAX=-0.2751953695E-05
 X( 2)= 0.2061242676E 00       EMAX= 0.2752206814E-05
 X( 3)= 0.4655187988E 00       EMAX=-0.2752131198E-05
 X( 4)= 0.7389050293E 00       EMAX= 0.2752054193E-05
 X( 5)= 0.9436755371E 00       EMAX=-0.2752207854E-05
```

Fig. 3 Program for the minimax dual fractional power approximation and its results.

3. Numerical experimentation

A computing procedure is outlined in Fig. 2 and detailed in Fig. 3. This program is provided for a TOSBAC 3400 model 21 digital computer. An investigation of a ratio of the minimum to the maximum absolute deviations shows how the optimization is achieved. Some of the resulting approximations are enumerated in Table 1 with each maximum absolute deviation.

Table 1 Minimax dual fractional power approximations furnishing more than 0.9999 of error-balance* in the interval[0, 1].

Table 1 Minimax dual fractional power approximations furnishing more than 0.9999 of error-balance* in the interval [0, 1]

Functions	Minimax Dual Fractional Power Approximations	Max. Abs. Errors	Full Scale Errors(%)
$\exp x$	$1.000000000 + 1.002261347\,x + 0.571024656\,x^{2.065645633} + 0.144995825\,x^{3.694510965}$	0.2193×10^{-4}	0.8068×10^{-3}
$\sinh x$	$0.000000000 + 1.000004326\,x + 0.167050480\,x^{3.001631876} + 0.008146388\,x^{5.111804903}$	0.1839×10^{-6}	0.1565×10^{-4}
$\cosh x$	$1.000000000 + 0.000210568\,x + 0.502027441\,x^{2.003455328} + 0.040842626\,x^{4.139990164}$	0.2752×10^{-5}	0.1783×10^{-3}
$\sin \frac{\pi}{2} x$	$0.000000000 + 1.570689531\,x - 0.658293028\,x^{3.012202756} + 0.087603497\,x^{4.708008110}$	0.4318×10^{-5}	0.4318×10^{-3}
$\cos \frac{\pi}{2} x$	$1.000000000 - 0.003533570\,x - 1.281127771\,x^{2.027488275} + 0.284661341\,x^{3.624865310}$	0.4250×10^{-4}	0.4250×10^{-2}
$\mathrm{erf}\, x$	$0.000000000 + 1.127857706\,x - 0.660091651\,x^{3.214527565} + 0.374934738\,x^{3.760594092}$	0.1871×10^{-4}	0.2220×10^{-2}

* Error-balance is defined by the ratio of the minimum to the maximum in absolute extreme deviations.

4. Discussion

(i) Characterization of the fractional power system

Since each fractional power term has two parameters, the use of dual fractional powers can achieve a degree of accuracy in Chebyshev approximations using the fifth-degree polynomials. If a given function behaves like an alternating series, this

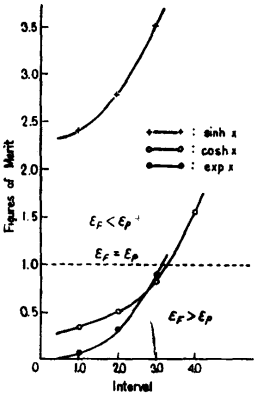

Fig. 4
Figures of Merit for both approximations using dual fractional powers and ordinary powers in the same interval.
(Figure of Merit is defined by $\varepsilon_P / \varepsilon_F$; where ε_P and ε_F show the absolute errors due to the minimax polynomial approximation of the fifth-degree and due to the dual fractional power approximation, respectively.).

mathematical system can achieve better approximation than ordinary power series. One of the advantageous features results from the capability of extended variable-range applications. This permits wide selection of variable-scaling in analog computers. The usefulness of the fractional power approximations can be found in Fig. 4.

(ii) On the local minimax approximation

The resulting approximation exhibits the equal-magnitude oscillatory behavior for its error. Then, the error satisfies six distinct roots in the closed interval [a, b], and attains an extreme deviation between two roots. The first and last extreme deviations in a Chebyshev approximation are forced out. This condition exhibits a local minimax approximation between the first and last null error points of the Chebyshev approximation. Principally, the maximum error magnitude is somewhat larger than that of the minimax approximation. In the diagram of Fig. 5, the characterization of the local minimax approximation is compared with that of the Chebyshev sense approximation.

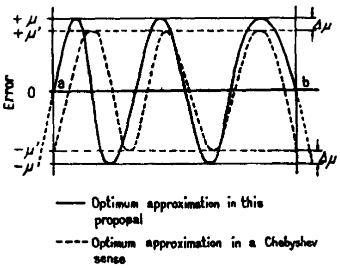

Fig. 5 Typical error-curves of both approximations in the same interval [a, b].

(iii) On the optimization approach

The usual techniques have required repetition of tedious numerical calculations. This is because increments of the parameter values have been evaluated as a set of solutions of a linear problem between the shift of the extremum points and the deviation of the extremum magnitude. Therefore, all matrix-elements are necessarily refreshed in every cycle of correction. For economical calculation, an improved method is proposed, which repeats interpolation at the base-points with off-set deviations. Proper deviations are given in terms of the weighted algebraic sum of neighboring extreme deviations. With the pertinent weight(a number 1/3 ∿ 1/20 is available), we can achieve rapid convergence of the computation. Figure 6 shows how the approximation to the function cosh x ($0 \le x \le 1$) is optimized with such weights. In this process, the off-set deviations ρ_i at the base-points converge to the proper values as shown in Fig. 7. Obviously, the convergence of the quality of approximation corresponds to that of the ρ_i values. This fact shows that there is a practical reason in our assumption that $\rho_i = W(\mu_i + \mu_{i+1})$. In some applications, this algorithm fails sometimes in the optimization when there is no true extreme deviation between two preliminary arranged base-points. Arrangement of the base-points with equal spaces will succeed in such a case, although the parameter values cannot be evaluated explicitly.

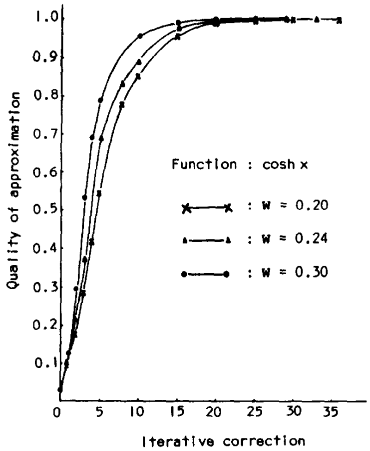

Fig. 6 Convergence of the approximation to the function cosh x ($0 \le x \le 1$) for the weights 0.20, 0.24 and 0.30.

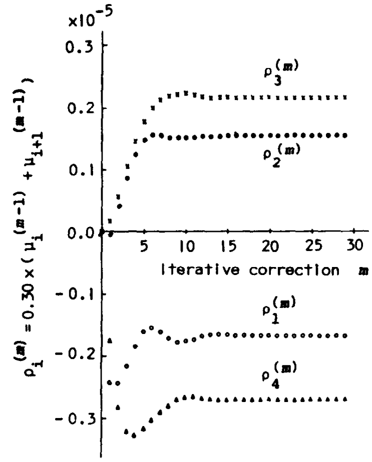

Fig. 7 Convergence of the off-set deviations with the weight 0.30 for the function cosh x ($0 \le x \le 1$).

5. Conclusions

The usefulness of the dual fractional power approximation has been clarified. Since each of the fractional power terms contains two parameters, one can economize on two ordinary power terms without diminishing the accuracy. Therefore, the degree of accuracy becomes comparable to that in approximation using the Chebyshev-polynomials up to the fifth-degree. For odd functions, especially, this approximation can achieve good accuracy over an extended interval. An algorithm of the optimal parameter-determination can be completed, which achieves a minimax condition with null error terminals.

This approximation system will serve as an economical and accurate simulation in analog computer applications.

References

[1] G. A. Korn and T. M. Korn, Electronic Analog and Hybrid Computers, New York, McGraw-Hill, 1964, pp. 233 ∿ 240.

[2] Y. Kobayashi, M. Ohkita and T. Chiba, "Arbitrary Function Generator," Japan. Pat. no. 9537, 1975.

[3] A. Nathan, "Linear and Nonlinear Interpolators," IEEE Trans. Electron. Comput., vol. EC-12, pp. 526 ∿ 532, 1963.

[4] Y. Kobayashi, M. Ohkita and M. Inoue, "Fractional Power Approximation and Its Generation," Trans. IMACS, vol. 18, no. 12, pp. 115 ∿ 122, 1976.

[5] Y. Kobayashi, M. Ohkita and M. Inoue, "Generation of Functions Approximated by a Fractional Power," Trans. IEE, Japan, vol. 96, no. 8, pp. 173 ∿ 178, 1976.