

標準関数の精度の検定

東大(理) 高橋 香 俊

東大(理) 国 井 利 泰

東大(大型計算機センター) 名 取 亮

東大(地震研) 桧 山 澄 子

§ 1 序

計算機に組み込まれている基本外部関数、および関数表の精度を検定するために、原器となるプログラムを作成したので報告する。まず原器を作る際の目標は

- (1) 多重精度をもつこと
- (2) 汎用性があること (大抵の計算機にかけられること)
- (3) 関連のある関数は、できるだけその関係を利用すること
- (4) 時間はこの際問題にしないこと
- (5) HARP(5020E)の4倍精度基本外部関数の検定を行なうこと

の5つであった。そのために次のようなことを行なった。

- (1)について；一応任意の桁まで求まるようになっているが、ここでは目標(5)のために(HARPの4倍精度は35.8桁)40桁までを正確に求めて検定を行なうことにした。そのために計算はすべて42桁で行なった。また優れた精度を得るために連分教展開式、またはニュートン近似式を用いて計算した。
- (2)について；多重精度の四則演算・入出力、4倍精度を多重精度になおすプログラムが必要になるが、これらはアセンブラでなくすべてFORTRANで用意した。
- (3)について；関数間に関連があっても、良い精度が得られるものでなければ、使わない。そのため図1に示すような計算手順になった。
- (4)について；時間はこの際問題でなかったが、 $\sqrt{2}$ 、 $\tan(\pi/4)$ の40桁までの計算は、それぞれ1620秒、12709秒であった。
- (5)について；表1の区間に対し関数の相対誤差をとり検定を行なった。結果は図2に示す。

§ 2 各関数の計算方法とアルゴリズム4

(1) $\tan(x)$

$0 \leq x \leq \frac{\pi}{4}$ の x に対し

$$\tan(x) = \frac{x}{1} - \frac{x^2}{3} + \frac{x^4}{5} - \frac{x^6}{7} + \dots \quad (1) - (i)$$

なる連分数展開を行なう。このとき第何項まで計算するかが問題になるが、§3で述べる方法で決めた項まで計算する。また連分数展開式の計算方法についても§3で述べる。次に区間の拡張を次式により行なう。

$$\tan(x) = 1 / \tan\left(\frac{\pi}{2} - x\right) \quad \left(\frac{\pi}{4} + n\pi < x < \frac{\pi}{2} + n\pi\right)$$

$$\tan(x) = -\tan(-x) \quad (x < 0)$$

サブルーチンMTAN1は(i)までの計算であり、サブルーチンMTANは区間の拡張を行なった全区間での計算である。以下の関数の計算についてもサブルーチン名に1がついたものは、限られた閉区間の x についての計算を行なうサブルーチンである。

(2) $\sin(x)$

$$0 \leq x \leq \frac{\pi}{2} \text{ のとき}$$

$$\sin(x) = 2 \tan\left(\frac{x}{2}\right) / (1 + \tan^2\left(\frac{x}{2}\right))$$

以下これを用いて

$$\frac{\pi}{2} < x \leq \pi \text{ のとき} \quad \sin(x) = \sin(\pi - x)$$

$$\pi < x \leq \frac{3}{2}\pi \text{ のとき} \quad \sin(x) = -\sin(x - \pi)$$

$$\frac{3}{2}\pi < x < 2\pi \text{ のとき} \quad \sin(x) = -\sin(2\pi - x)$$

(3) $\cos(x)$

$$0 \leq x \leq \frac{\pi}{4} \text{ のとき}$$

$$\cos(x) = (1 - \tan^2(\frac{x}{2})) / (1 + \tan^2(\frac{x}{2}))$$

以下これを用いて

$$\frac{\pi}{4} < x \leq \frac{3}{4}\pi \text{ のとき} \quad \cos(x) = \sin(\frac{\pi}{2} - x)$$

$$\frac{3}{4}\pi < x \leq \frac{5}{4}\pi \text{ のとき} \quad \cos(x) = -\cos(\pi - x)$$

$$\frac{5}{4}\pi < x \leq \frac{7}{4}\pi \text{ のとき} \quad \cos(x) = \sin(x - \frac{3}{2}\pi)$$

$$\frac{7}{4}\pi < x < 2\pi \text{ のとき} \quad \cos(x) = \cos(2\pi - x)$$

(4) $\arctan(x)$

$0 \leq x \leq \sqrt{2}-1$ のとき

$$\tan^{-1}(x) = \frac{x}{1+} - \frac{x^2}{3+} + \frac{4x^2}{5+} - \frac{k^2x^2}{(2k+1)+} \quad (4) \text{-(i)}$$

をある有限項 (§3参) まで計算する。以下

$$\sqrt{2}-1 < x < \sqrt{2}+1 \text{ のとき} \quad y = \frac{\pi}{4} + \tan^{-1}(x-1)/(x+1)$$

$$\sqrt{2}+1 < x \text{ のとき} \quad y = \frac{\pi}{2} - \tan^{-1}(\frac{1}{x})$$

より (4)-(i) に帰着させる。

(5) $\arcsin(x)$

$|x| < \frac{\sqrt{2}}{2}$ のとき

$$\sin^{-1}(x) = \tan^{-1}(x/\sqrt{1-x^2})$$

$\frac{\sqrt{2}}{2} \leq |x| \leq 1$ のとき

$$\sin^{-1}(x) = \operatorname{sgn}(x) \left[\frac{\pi}{2} - \tan^{-1}\left(\frac{\sqrt{1-x^2}}{|x|}\right) \right]$$

(6) $\arccos(x)$

$$\cos^{-1}(x) = \frac{\pi}{2} - \sin^{-1}(x)$$

(7) $\tanh(x)$ $0 \leq x \leq 1$ のとき

$$\tanh(x) = \frac{x}{1+} - \frac{x^2}{3+} + \frac{x^2}{5+\dots} - \frac{x^2}{(2k+1)+\dots} \quad (7)-(i)$$

をある有限項まで計算する。

 $x > 1$ のとき二倍角

$$\tanh(2x) = 2 \tanh(x) / (1 + \tanh^2(x))$$

を用い(7)-(i)に帰着して求める。

(8) $\sinh(x)$ $|x| \leq 1$ のとき

$$\sinh(x) = 2 \tanh\left(\frac{x}{2}\right) / (1 - \tanh^2\left(\frac{x}{2}\right))$$

 $|x| > 1$ のとき二倍角

$$\sinh(2x) = 2 \sinh(x) \sqrt{1 + \sinh^2(x)}$$

を用いて計算する。

(9) $\cosh(x)$ $|x| \leq 1$ のとき

$$\cosh(x) = (1 + \tanh^2\left(\frac{x}{2}\right)) / (1 - \tanh^2\left(\frac{x}{2}\right))$$

 $|x| > 1$ のとき二倍角

$$\cosh(2x) = 2 \cosh^2(x) - 1$$

を用いて計算する。

(10) $\exp(x)$ $0 \leq x \leq 1$ のとき

$$e^x = (1 + \tanh(\frac{x}{2})) / (1 - \tanh(\frac{x}{2})) \quad (10)-(i)$$

$$x > 1 \text{ のとき } e^x = (e^{\frac{x}{2}})^2$$

$$x < 0 \text{ のとき } e^x = (e^{-\frac{x}{2}})^{-2}$$

を用いて(10)-(i)に帰着させる。

(11) $\ln(x)$

$0.5 \leq x \leq 2$ のとき、即ち $-\frac{1}{3} \leq z \leq \frac{1}{3}$ のとき

$$\ln(x) = \ln\left(\frac{1+z}{1-z}\right) = \frac{2z}{1-z} - \frac{z^2}{3-z} + \frac{4z^2}{5-\dots} - \frac{k^2 z^2}{(2k+1)-\dots} \quad (11)-(i)$$

なる連分数展開を用いて計算する。

$0 < x < 0.5$ のとき

$x \cdot 2^n > 0.5$ なる n に対して

$$y = \log(x \cdot 2^n) - n \log 2$$

$x > 2$ のとき

$\frac{x}{2^n} < 2$ なる n に対して

$$y = \log\left(\frac{x}{2^n}\right) + n \log 2$$

を計算する。

(12) \sqrt{x}

$0.01 \leq x < 1$ に対してニュートンの近似式

$$y_{n+1} = \frac{1}{2} \left(y_n + \frac{x}{y_n} \right)$$

を用いて計算する。第何項まで計算するかは、求める桁数

を N としたとき

$$|y_{n+1} - y_n| / y_n \leq 10^{-N}$$

になるまで計算する。また区間の拡張は

$x \geq 1$ のとき

$$y = 100^{\lceil \frac{k}{2} \rceil} \cdot 10^{k-2\lceil \frac{k}{2} \rceil} \sqrt{x \cdot 100^{-k}}$$

$x < 0.01$ のとき

$$y = 100^{\lceil \frac{k+1}{2} \rceil} 10^{k-2\lceil \frac{k+1}{2} \rceil} \sqrt{x \cdot 100^{-k}}$$

より行なう。

§ 3 連分数展開式の計算

$\tan(x)$ 、 $\arctan(x)$ 、 $\tanh(x)$ 、 $\ln(x)$ の計算は連分数(1)-(i)、(4)-(i)、(7)-(i)、(11)-(i)、で計算された。

さて、それらの連分数を実際に計算するには、今連分数

$$V_m = \frac{x_0}{y_0 +} \frac{x_1}{y_1 +} \frac{x_2}{y_2 + \dots} \frac{x_m}{y_m}$$

において

$$P_{-1} = 0, \quad P_0 = x_0, \quad q_{-1} = 1, \quad q_0 = y_0$$

$$P_k = y_k P_{k-1} + x_k P_{k-2}$$

$$q_k = y_k q_{k-1} + x_k q_{k-2} \quad (k = 1, 2, \dots, m)$$

とすれば

$$V_m = P_m / q_m$$

となり、 $2m-1$ 回の加法、 $2m-1$ 回の乗算、1回の除法で計算される。

次に無限な連分数をどこで打ち切るかが問題になるが、40

桁まで正しく求めたとき、表2のような結果が得られた。

これより打ち切り項数の近似式 MAX は

$\tan(x)$ 、 $\tanh(x)$ については

$$\text{MAX} = \text{IFIX}(10X) + 10$$

$\arctan(x)$ については

$$\text{MAX} = \text{IFIX}(40X) + 12$$

$\ln(x)$ については

$$\text{MAX} = \text{IFIX}(-50X) + 13$$

が出るから、実際の計算では無条件に MAX まで連分教を計算し、以後

$$|V_m - V_{m+1}| / |V_m| < 10^{-N}$$

になるまで iteration を行なうことにした。

なお、プログラムのチェックは、 \sqrt{x} 以外は互に独立なプログラムを用いることで行なった。例えば $\arctan(x)$ と $\tan(x)$ のチェックでは

$$x = \arctan(\tan(x))$$

かどうかで行なった。

\sqrt{x} は

$$x = (\sqrt{x})^2$$

になるかどうかでたしかめた。

§ 4 HARP (5020 E) との比較検定

HARP (5020 E) の 4 倍精度の基本外部関数の精度の検定は、多重精度の関数値との相対誤差を取ることで行なった。

その結果 QSQRT、QLOG は 10^{-35} まで正確に求まるが、QSIN、QCOS、QATAN、QEXP は 10^{-34} までの精度しかないことがわかった。

特に、QATAN では $x \geq \frac{\pi}{2}$ で精度が落ちる。

結果は図 2 に示す。

4 倍精度の検定区間

| | | | |
|-------|-------------------|-------|---------------------|
| QSQRT | $0 \leq x \leq 1$ | QATAN | $0 \leq x \leq 2$ |
| QSIN | $0 \leq x \leq 6$ | QLOG | $0.5 \leq x \leq 2$ |
| QCOS | $0 \leq x \leq 6$ | QEXP | $-1 \leq x \leq 1$ |

表 1

| x | tanh | tan | artan | ln | z |
|-------|------|-----|-------|-----|-----------|
| 0.125 | 10 | 10 | 16 | >50 | -0.7 |
| 0.25 | 12 | 12 | 21 | 43 | -0.6 |
| 0.375 | 13 | 13 | 26 | 32 | -0.45 |
| 0.5 | 14 | 14 | 31 | 26 | -0.3 |
| 0.625 | 15 | 15 | 37 | 21 | -0.230762 |
| 0.75 | 15 | 16 | 41 | 20 | -0.142857 |
| 0.875 | 16 | 16 | 47 | 13 | -0.06 |
| 1.0 | 17 | 17 | 52 | 0 | 0.0 |

表 2

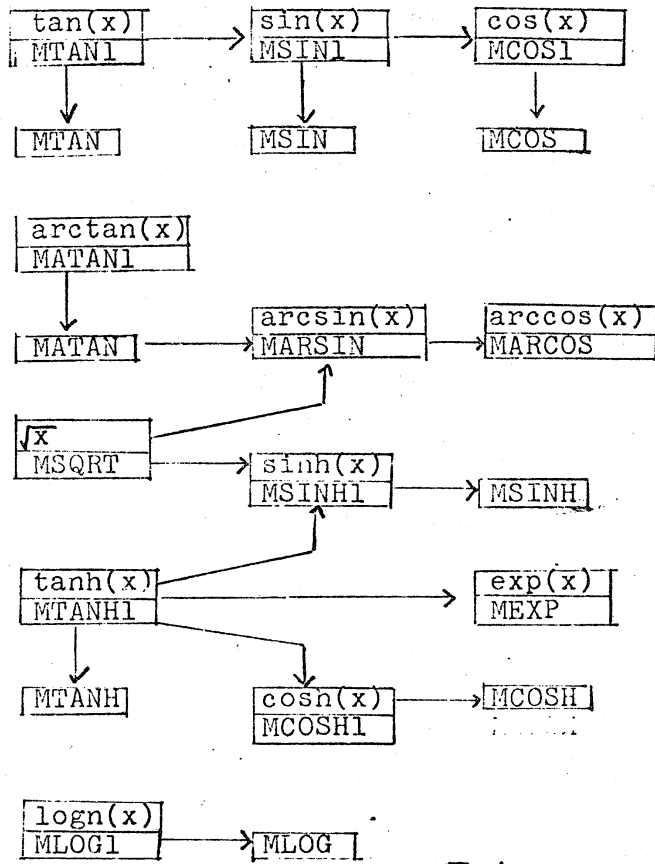
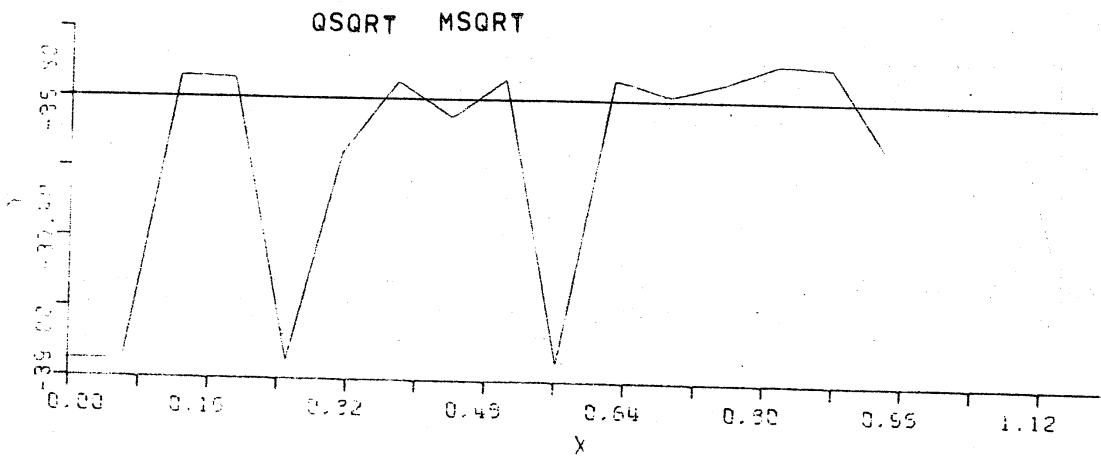
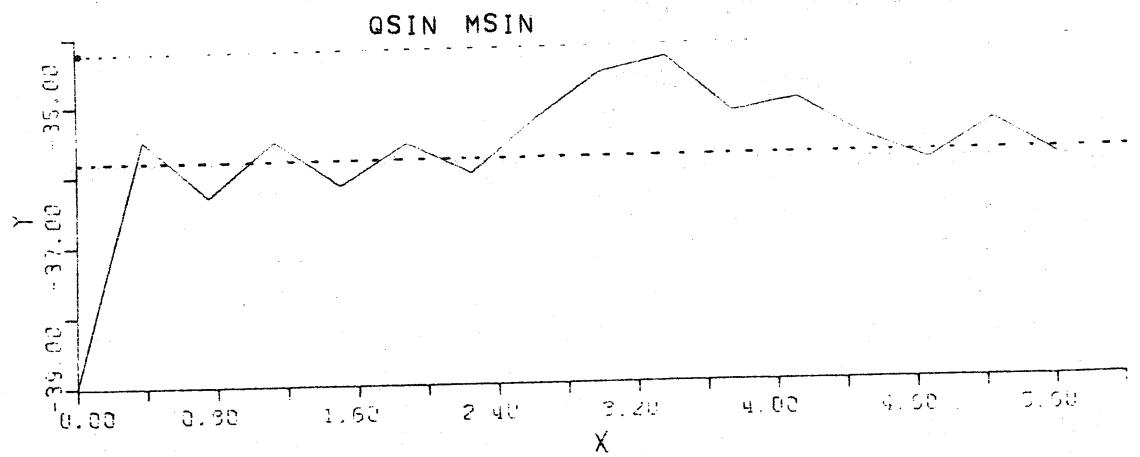
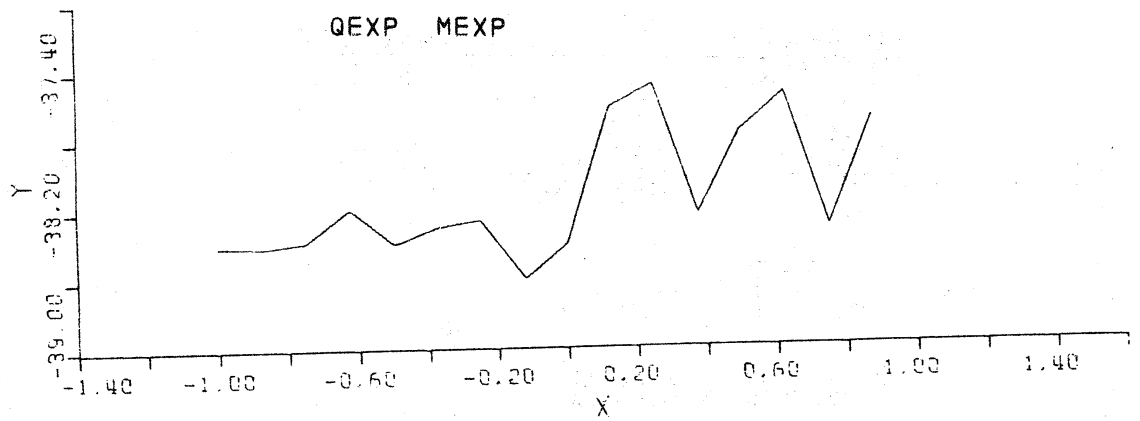
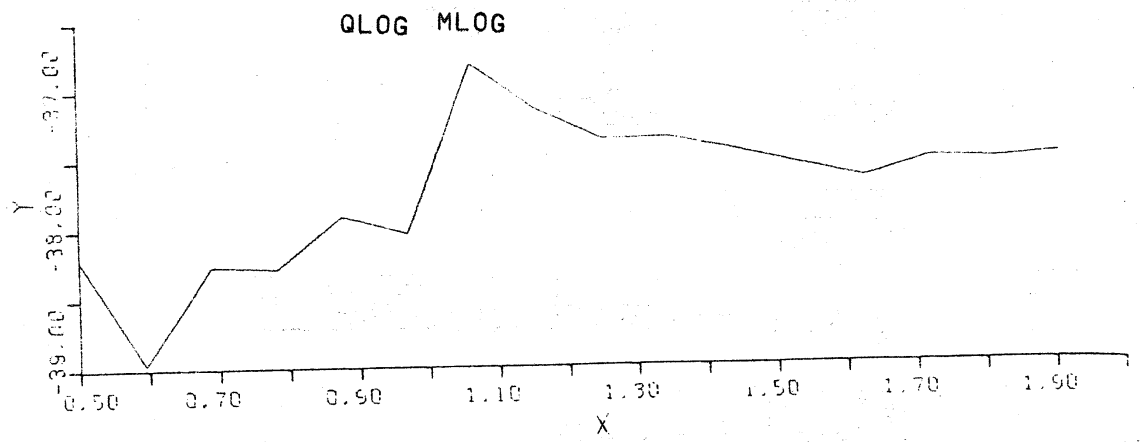


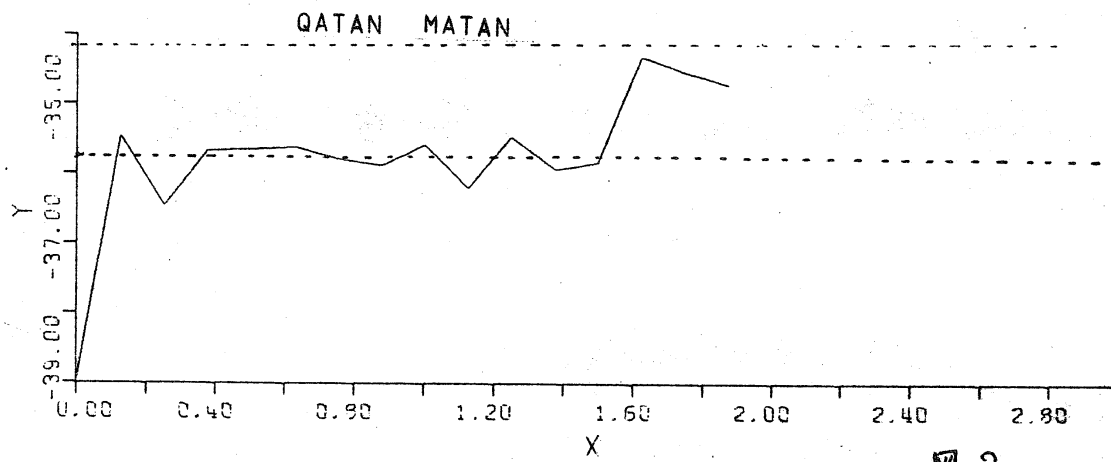
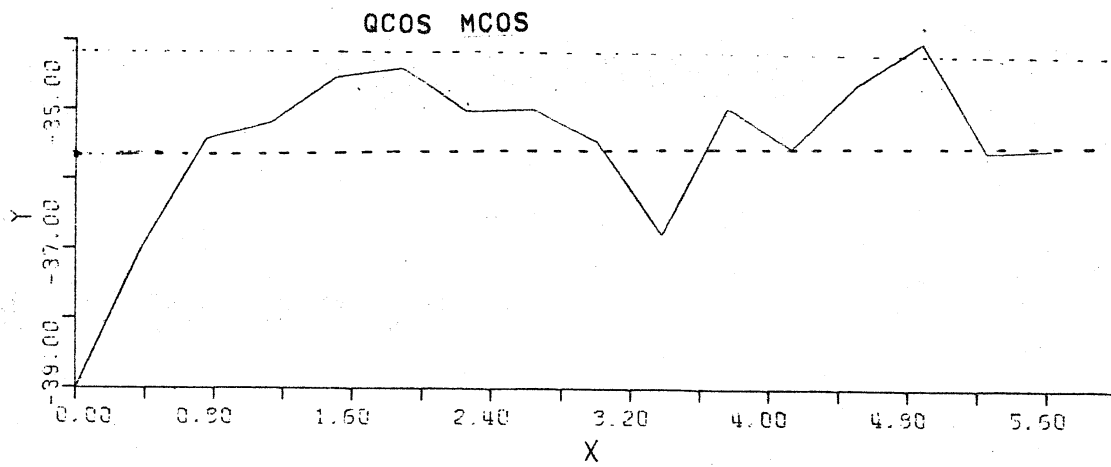
図 1

計算手順の流れ図

4倍精度の相対誤差







2

References

- Hart; Computer approximations, 1968, John Wiley
 I.D.Hill; Procedures for the basic arithmetical operations in multiple-length working, Comp.J., 1968
 PP232-235
 C.T.Fike; Computer evaluation of mathematical functions;
 1968, Prentice Hall