数理解析研究所講究録
382巻 1980年 103-111

長大桁・長大指数演算パッケージMULPACとその応用

筑波大電子情報　森　正武
京大数理研　今井敬子

§1. MULPACとその改良

MULPACシステムとは，DECUSに登録されているDEC SYSTEM-20上で稼動可能なFORTRAN用多倍長・長大指数演算のサブルーチン・パッケージのことである。MULPACで扱う数値の内部表現ならびに機能は下記の通りである。

○内部表現：連続したN WORD（Nは任意，1 WORD 36 bit）から成り，その才1 WORDは指数部，才2～才N WORDは仮数部で，それぞれ2進表示である。

○機能：四則演算サブルーチン。
大小比較サブルーチン。
上記内部表現のための長さ指定ルーチン。
上記の数値とdouble, real, integerとの間の変換ルーチン。

上記パッケージを使用するにあたっては，演算の一つ一つがサブルーチン・コールになるなど，プログラム作成あるい

はデバッグ上の不便がある。また，DECSYSTEMのSYSTEM変更により使用不可能なサブルーチンが出てくるなどの問題点もある。そこで，これらの欠点を補うため，下記のようなpre-processorを作成した。

── pre-processor ──

○入力言語：FORTRAN statement ＋

MULPAC 宣言文

例　MULPAC　X(3), Y(2, 4)

○出力言語：FORTRAN

このpre-processorは1-pass translatorであり，それ自身9K WORDから成るPascalで書かれている。現在のところ，内部表現として4 WORD（N＝4）をとってある。

入力言語の機能としては，標準のFORTRANの機能以外に下記のものが追加されている。

1．上記MULPAC宣言文で宣言された変数（以下MULPAC変数と呼ぶ）の四則，比較等の演算に対して，FORTRANと同じ記号の使用を許す。

2．FORTRANの標準のtype（integer, real, double）との混合演算を許す。

3．MULPAC変数のための関数：MPABS（絶対値），MPSIN(sin), MPCOS(cos), MPEXP(exp), MPLOG(log),

MPSQRT（平方根）。

現段階でのこのpre-processorには，内部表現が4 WORDに制限されていること以外に，MULPAC変数のままでの入出力ができないという制約がある。その主な理由は，もともとのMULPACパッケージに入出力機能が全く無いためである。それ以外に，このpre-processorを使用する際の制限事項としては以下のようなものがある。

1．pre-processorによってtranslateされた結果が2以上のstatementに分かれるようなMULPAC変数の演算を含む文は，DO文の最後の文に使用することができない。次の例は許されない例である。

```
      DO  10   I=1,100
        :
10    X=X+Y+Z      (X,Y,ZはMULPAC変数)
        ⇓
      DO  10   I=1,100
        :
10    CALL  MPADD(X,Y,WK0000)
      CALL  MPADD(WK0000,Z,X)
```

2．論理IF文のtrue statementについても同様の制限がある。

3．文関数の宣言を許していない。

4．TYPE宣言としてINTEGER*8は許されず，したがって2 WORD INTEGERの使用が許されていない。

上記制限事項は今後pre-processorをversion upするときに少しずつ取除いていきたいと考えている。

pre-processorの入力言語とその出力のプログラム例を次のページに示す。

§2．二重指数関数型数値積分公式への応用

数値計算における長大桁および長大指数のもつ有効性を，二重指数関数型数値積分公式を例にとって調べてみよう。積分の例として

$$I(\alpha)=\int_{-1}^{1}(1-x^2)^{\alpha}dx=2\int_{0}^{1}(1-x)^{\alpha}(1+x)^{\alpha}dx,\quad -1<\alpha \qquad (1)$$

を取り上げる。この積分に $x=\tanh\left(\frac{\pi}{2}\sinh t\right)$ なる変換をほどこすと

$$I(\alpha)=\int_{-\infty}^{\infty}\left(1-\tanh^2\left(\frac{\pi}{2}\sinh t\right)\right)^{\alpha}\frac{\frac{\pi}{2}\cosh t}{\cosh^2\left(\frac{\pi}{2}\sinh t\right)}dt$$

$$=\int_{-\infty}^{\infty}\frac{\frac{\pi}{2}\cosh t}{\cosh^{2(1+\alpha)}\left(\frac{\pi}{2}\sinh t\right)}dt \qquad (2)$$

となるが，これにきざみ幅一定の台形則を適用すれば，二重

```
C
      MULPAC X,Y
      DOUBLE PRECISION A,B,C
      READ(5,1000) A,B
 1000 FORMAT(2F10.5)
      X=A
      IF(X+1.0.GT.X*B) GO TO 1   ①
      Y=X*X+2.0*MPSIN(B)  ②
      GO TO 2
    1 Y=X+MPABS(X)-B  ③
    2 C=Y
      WRITE(6,2000) A,B,C
 2000 FORMAT(' A= ',D25.18/' B= ',D25.18/' C= ',D25.18)
      STOP
      END
C



C
      DOUBLE PRECISION A,B,C
        DIMENSION X(4) , Y(4)
        DIMENSION WZ0000(4),WZ0001(4),WZ0002(4),WZ0003(4),WZ0004(4),
     -  WZ0005(4),WZ0006(4),WZ0007(4),WZ0008(4),WZ0009(4),WKZERO(4)
        INTEGER WKBL01,WKBL02,WKBL03,WKBL04,WKBL05
        DATA WKZERO/0,0,0,0/
        CALL ERRSET(0)
        CALL MPINI
      READ(5,1000) A,B
 1000 FORMAT(2F10.5)
        CALL   MPDCON(A,X)
!     X=A
        CALL   MPCONS(1.0,WZ0000)
        CALL   MPADD(X,WZ0000,WZ0001)
        CALL   MPDCON(B,WZ0002)
        CALL   MPMUL(X,WZ0002,WZ0003)
        CALL   MPTEST(WZ0001,WZ0003,WKBL01)
      IF(WKBL01.GT.0) GOTO 1
        CALL   MPMUL(X,X,WZ0000)
        CALL   MPDCON(B,WZ0001)
        CALL   MPSIN(WZ0001,WZ0002)
        CALL   MPCONS(2.0,WZ0003)
        CALL   MPMUL(WZ0003,WZ0002,WZ0004)
        CALL   MPADD(WZ0000,WZ0004,Y)
!     Y=X*X+2.0*MPSIN(B)
      GO TO 2
1       CALL   MPABS(X,WZ0000)
        CALL   MPADD(X,WZ0000,WZ0001)
        CALL   MPDCON(B,WZ0002)
        CALL   MPSUB(WZ0001,WZ0002,Y)
!   1 Y=X+MPABS(X)-B
2       CALL   MPDPV(Y,C)
!   2 C=Y
      WRITE(6,2000) A,B,C
 2000 FORMAT(' A= ',D25.18/' B= ',D25.18/' C= ',D25.18)
      STOP
      END
C
```

① ② ③

指数関数型公式による結果が得られる。

αが-1に近いとき，この計算で二つの問題が生ずる。

(i) $1-x^2=(1-x)(1+x)$ の計算における，$x=\pm1$に近いところでの桁落ち。

(ii) 重み $\frac{\pi}{2}\cosh t/\cosh^2(\frac{\pi}{2}\sinh t)$ の$t\to\pm\infty$におけるアンダー・フロー。

このうち，(i) の問題は長大桁の採用により，また (ii) の問題は長大指数の採用により解決することができる。これらを簡単に考察してみよう。

(i) 桁落ちの問題

εをいわゆる計算機イプシロンとすると，$1-x$の計算ではつねにε程度の誤差が生ずると考えられる。xが1に近づいて，その差が計算機イプシロン以下になると，桁落ちによって$1-x$の計算結果は0になり，積分は通常はそこで打切らざるを得ない。そのxの上限をx_{max}，すなわち

$$\varepsilon = 1 - x_{max} \tag{3}$$

と置く。すると，$1-x$の計算で生ずる誤差の積分への寄与ΔI_εはほぼ次のようになる。

$$\Delta I_\varepsilon = 2\int_0^{x_{max}} \{(1-x+\varepsilon)^\alpha-(1-x)^\alpha\}(1+x)^\alpha dx + 2\int_{x_{max}}^1 (1-x)^\alpha(1+x)^\alpha dx \tag{4}$$

右辺オ1項は積分される関数値のもつ誤差から生ずる積分誤差の上限で，オ2項は x_{max} で打切ったことにより生ずる誤差である。ここで，α は -1 に近い場合を考え，積分は $x=-1$ の近くでの寄与が大きいとして，$(1+x)^{\alpha} \fallingdotseq 2^{\alpha} \fallingdotseq 2^{-1}$，$(1-x_{max})^{\alpha} \gg 1$ とすると，ΔI_{ε} は次のようになる。

$$\Delta I_{\varepsilon} \fallingdotseq 2^{1+\alpha}\varepsilon\alpha\int_{0}^{x_{max}}(1-x)^{\alpha-1}dx + 2^{1+\alpha}\int_{x_{max}}^{1}(1-x)^{\alpha}dx$$

$$= \varepsilon\left[(1-x_{max})^{\alpha}-1\right] + \frac{1}{\alpha+1}(1-x_{max})^{\alpha+1}$$

$$= \frac{\alpha+2}{\alpha+1}\varepsilon^{\alpha+1} \qquad (5)$$

DECSYSTEM-20 の FORTRAN の2倍精度における計算機イプシロンは

$$\varepsilon = 2^{-62} = 2.2\times10^{-19}$$

程度であるから，例えば $\alpha=-0.9$ とすると，$I(-1)\fallingdotseq 10.7177\cdots$ に対し $\Delta I_{\varepsilon} \fallingdotseq 1.5\times10^{-1}$ となり，2倍精度では十分な精度は期待できない。しかし，仮数部に3WORD を使う現在の MULPAC を利用しても

$$\varepsilon = 2^{-36\times3} = 2^{-108} = 3.1\times10^{-33}$$

であるから，$\alpha=-0.9$ のときには $\Delta I_{\varepsilon} \fallingdotseq 6.2\times10^{-3}$ 程度であってあまりその効果はない。もし，$\alpha=-0.9$ に対して $\Delta I_{\varepsilon} \fallingdotseq 10^{-15}$

程度の値を得ようとすると，$\varepsilon = \left(\frac{0.1}{1.1}\times 10^{-15}\right)^{\frac{1}{0.1}} \fallingdotseq 2^{-5.33} = (2^{-36})^{14.8}$ より，15 WORD 程度の仮数部が必要となる。したがって，α が -1 に近いときには，多倍長演算はあまり効率的ではない。仮数部 3 WORD の現在の MULPAC で 10^{-15} 程度の精度が得られるのは，$\alpha = -0.53$ あたりまでである。

この例のように，具体的に関数形が陽に定義できる被積分関数の場合には，

$$1 \mp \tanh x = \frac{2e^{\mp x}}{e^{x}+e^{-x}}$$

の関係を利用してうまくプログラムを行うことにより，桁落ちによる上記の誤差 ΔI_ε は避けることができる [1]。

(ii) アンダー・フローの問題

$|t|\to\infty$ のとき，(2) の重み $\frac{\pi}{2}\cosh t / \cosh^2\left(\frac{\pi}{2}\sinh t\right)$ の分母が著しく大きくなってアンダー・フローを起こすので，汎用サブルーチンではこれを避けるために

$$\cosh^{-2}\left(\frac{\pi}{2}\sinh t_{max}\right) \fallingdotseq 4\exp\left(-\frac{\pi}{2}\exp t_{max}\right) = \delta \qquad (6)$$

を満たす t_{max}，すなわち

$$t_{max} = \log\left(\frac{2}{\pi}\log\frac{4}{\delta}\right) \qquad (7)$$

において積分の上下限を打切るようにしなければならない。

ただし，δは浮動小数点数の最小の絶対値で，DECSYSTEM-20では約10^{-38}であり，対応するt_{max}は約4.03である。この打切により，積分には

$$\Delta I_\delta \fallingdotseq 2\int_{t_{max}}^{\infty} \frac{1}{\cosh^{2(1+\alpha)}\left(\frac{\pi}{2}\sinh t\right)}dt \fallingdotseq \exp\left(-\frac{\pi}{2}(1+\alpha)\exp t_{max}\right) \qquad (8)$$

程度の誤差が生ずる。

DECSYSTEM-20の倍精度演算とMULPACを使用して得た誤差を次に示す。MULPACでは積分値は$t_{max}=6.25$で打切ってある。このとき，(6)より$\delta \fallingdotseq 10^{-354}$である。

α	$I(\alpha)$	MULPAC	2倍精度	ΔI_δ
-0.8	5.74349	-0.4×10^{-15}	-0.4×10^{-7}	0.6×10^{-7}
-0.9	10.7177	-0.7×10^{-15}	0.5×10^{-3}	0.2×10^{-3}
-0.99	100.695	0.8×10^{-2}	$0.4\times10^{+2}$	0.4×10^{0}

最後のΔI_δは，(8)によって計算した2倍精度に対する値である。この場合には，$\alpha=-0.9$程度でもMULPACで十分な効果が得られることがわかる。なお，この計算は，(i)の桁落ちによる誤差は生じないようにプログラムを組んで実行させた。

参考文献

[1] 森正武「曲線と曲面 ― 計算機による作図と追跡」教育出版 1974。