J. Nig Ass. Math. Phys.
Vol. 4 (2000)

PROPAGATION OF ERROR IN POLYNOMIAL EVALUATION

P. B. SHOLA
DEPT OF MATHEMATICS
UNIVERSITY OF ILORIN, ILORIN, NIGERIA.

ABSTRACT

Polynomials especially those with coefficients of the same sign or strictly of alternating signs are of great importance in the approximation of many mathematical functions that occur in mathematical physics. Two standard algorithms namely the Clenshaw,and Horner's algorithms for evaluating polynomial are examined with respect to the propagation of error. Using process graph, it is established that the quantity | y'(t)/y(t)| controls the upper bound in the relative error in the value of a polynomial at t.

INTRODUCTION

Most mathematical functions are approximated using polynomials and apart from the truncation error incurred in such process (due to the replacement of the function by finite number of terms of a polynomial) the need both to quicken the evaluation process as well as to reduce the error propagation during evaluation of such polynomials has being a major concern. Three standard algorithm namely the Horner, Clenshaw and Reinsch algorithm have been devised for polynomial evaluation and have been studied by many researchers. Among such researchers are Newbery, Oliver, Razaz and Sconfelder..Newbery[1] compared Horner and Clenshaw algorithms and demonstrated that(a) the accuracy of the Horner scheme (like the Clenshaw scheme) is highly sensitive to the magnitude of the value,x, at which the polynomial is evaluated.(b)When a polynomial has coefficients of constant sign or of strictly alternating sign, a translation into chebyshev form will not bring any improvement in the accuracy of evaluation. The first hypothesis prompted Newbery to suggest the modification of Horner algorithm in such a way as to efficiently reduce the range (interval within which the polynomial is evaluated) of the argument of the polynomial to [-1/2,1/2].Oliver [3] in agreement also showed that Horner's algorithm is most accurate for small values of |x| (the absolute value of x) and that the potentially serious error magnification near the value |x|=1, which the Horner's algorithm exhibits could be avoided if Newbery's modification of Horner scheme is used instead. Since the Newbery scheme is less accurate elsewhere except near |x|=1, he suggested the use of Horner/Newbery scheme for polynomial evaluation.Oliver [4] went further to obtain an absolute error bound to be

$$\varepsilon \sum_{r=0}^{n} | p_r t^r | = \varepsilon p(| t |)$$

for monomial form (where p_r are the coefficients of the polynomial) and

$$\varepsilon \sum_{s=0}^{n} | c_s T_s(t) |$$

for chebyshev series form.He further concluded that for smaller values of |t| the inaccuracies in the coefficients may affect chebyshev form rather more except near |t|=1.

Razaz and Schonfelder [2] examined the relative error incurred during polynomial evaluation and concluded that the quantity

$$\frac{| p'(x)|_{\max}}{| p(x)|_{\min}}$$

plays an important role in the control of the relative error in the process. This paper apart from investigating the contribution of each of the three typical sources of error (namely, inexact coefficients of the polynomial, and of the value t, at which the polynomial is evaluated and the round_off error) to the total relative error on the computed value of a polynomial the paper also establish that the ratio

$$| \frac{p'(t)}{p(t)} |_{\max}$$

rather than the exaggerated value

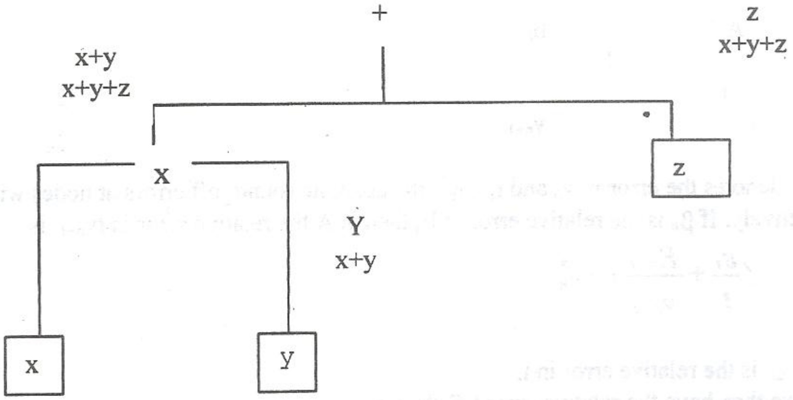
$$\frac{| p'(t)|_{\max}}{| p(t)|_{\min}}$$

is essential in the control of the relative error in a polynomial evaluation. This result is obtained using a process graph (which follows the process a computer uses for its computation).A brief description of a process graph is given in the next section and the sections that follow this contain the relative error analysis of Horner's and Clenshaw's algorithms. The discussion is restricted to polynomials with all its coefficients having the same sign or of strictly alternating sign. Since a polynomial with negative coefficients can be converted to the one with positive coefficient by multiplying through by -1 (with the result of the evaluation of this polynomial negated to obtain the needed result) and that of the coefficients strictly alternating transformed , by a transformation such as t=-x, to a polynomial with polynomial of constant coefficients or of positive coefficients (depending on whether the coefficients of odd or even powers are negative) it therefore suffices to consider polynomial with positive coefficients and this we did.

PROCESS GRAPH

A process graph is a pictorial representation of the sequence in which the arithmetic operations in a calculation are carried out. it consists of a tree in which the arithmetic operations concerned are contained in the nodes.

Entering each of these nodes are lines from two other nodes that carry the operands. The lines carry the factors with which the relative error in the operands needed multiplied. For example the figure below is a process graph for the expression R=z+(y+x)



If the absolute errors in x,y and z are ε_x,ε_y, ε_z then the relative error in R is

$$(\frac{\varepsilon_x}{x}.\frac{x}{x+y}+\frac{\varepsilon y}{y}.\frac{y}{x+y}+r_*)(\frac{x+y}{x+y+z})+\frac{\varepsilon_z}{x+y+z}+r_+$$

where r_*, and r_+ are round-off error at the nodes containing * and +.

HORNER'S ALGORITHM

The Horner's algorithm for evaluating a polynomial

$$y(t)=\sum_{i=1}^{n} b_i t^i$$

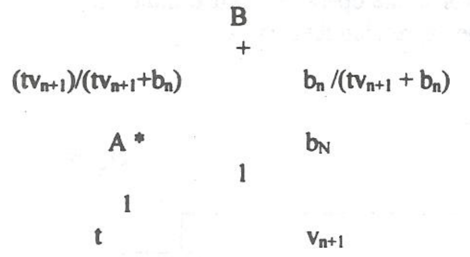
where t is assumed to lie in [-1,1] is

$$v_n = b_N$$

$$v_n = tv_{n+1} + b_n \quad n = N-1,.......1,0$$

$$y(t) = v_0$$

The graph process for this algorithm is



Let E_n denotes the error in v_n and r_n^*, r_n^+ the absolute round_off errors at nodes with * and + respectively. If β_n is the relative error in b_n then at A the relative error in tv_{n+1} is

$$(\frac{\varepsilon_t}{t}+\frac{E_{n-1}}{v_{n-1}})+\overset{*}{r_n}$$

where ε_t is the relative error in t.

At B we then have the relative error (E_n/V_n) as

$$\frac{E_n}{v_n}=\{(\frac{\varepsilon_t}{t}+\frac{E_{n-1}}{v_{n-1}})+\overset{*}{r_n}\}\frac{t\,v_{n-1}}{t\,v_{n-1}+b_n}+\frac{b_n}{t\,v_{n-1}+b_n}\beta_n+r_n$$

$$E_n=t\,E_{n-1}+(\varepsilon_t v_{n-1}+\overset{*}{r_n}t\,v_{n-1}+\beta_n b_n+r_n v_n)$$

.The solution of which is

$$E_0=\varepsilon_t y'+t\sum_{n=0}^{N-1}(\dot{r}_{n-1}+\overset{*}{r_n})t^n v_{n-1}+r_0 y+\sum_{n=0}^{N}t^n\beta_n b_n$$

where y' is the first derivative of y with respect to t and can easily be established to be

$$y'(t)=\sum_{r=0}^{N-1}v_{r-1}t^r$$

The first expression is the effect of error in the value of t .

The middle term

$$t\sum_{n=0}^{N-1}(\dot{r}_{n-1}+\overset{*}{r_n})t^n v_{n-1}+r_0 y$$

is due to the round off error and the last, from error in the coefficients.

THE EFFECT OF INEXACT COEFFICIENT

The effect of inexact coefficient is obtained from above by putting r_n , r_{n-1}, r_0 and ε_t equal zero to have

$$E_0 = \sum_{n=0}^{N} t^n \beta_n b_n$$

$$| E_0 | \leq \sum_{n=0}^{N} | t |^n | \beta_n || b_n |.$$

so that if the maximum attainable round off error in b_n is ε then

$$\frac{| E_0 |}{| y |} \leq \varepsilon \frac{y(| t |)}{| y(t) |}$$

It should be noted that if t>0 then this reduces simply to ε so that the bound is not greater than ε , the maximum absolute error in the coefficient initially started with. Now if t is less than zero then it is obvious that the error bound

$$\varepsilon \frac{y(-t)}{| y(t) |}$$

is greater than one since the coefficients are positive which thus warns us that the error may be magnified for this case.

THE EFFECT OF ROUND OFF ERROR

The effect of round error is

$$E_0 = \varepsilon_t y' + t \sum_{n=0}^{N-1} (r_n^{*} + r_n^{+}) t^n v_{n+1} + r_0 y$$

$$| E_0 | \leq \varepsilon\{| y' | + 2 | t | y'(| t |) + | y |\}$$

$$\leq \varepsilon\{3y'(| t |) + | y |\}$$

where ε is the machine accuracy parameters.

Consequently

$$\frac{| E_0 |}{| y(t) |} \leq \varepsilon\{1 + 3 \frac{y'(| t |)}{| y(t) |}\}$$

It can be seen that the round off error may be severe on the evaluation sometimes than that due to coefficients. This result also informs us that the error may be magnified except when

$$-\tfrac{2}{3} < \frac{y'(| t |)}{| y(t) |} < 0$$

A condition which is never satisfied since the ratio is positive.

The total bound on the relative error in the interval [-1,1] is

$$\varepsilon\{1+3\max_{t\in[-1,1]}(\frac{y'(|t|)}{|y(t)|})\}$$

ERROR DUE TO INEXACT VALUE t

Error due to inexact value of t is simply

$$E_0=\varepsilon_t y'$$

$$|\frac{E_0}{y}|\le\varepsilon\max_{t\in[-1,1]}|\frac{y'(t)}{y(t)}|$$

COMBINED EFFECT OF ALL THE ERROR

We have this to be

$$|E_0|\le\varepsilon\{|y'|+2|t|\Sigma|t^n v_{n-1}|+2|y|\}$$
$$\le\varepsilon\{|y'|+2|t|y'(|t|)+2|y|\}$$
$$\le\varepsilon\{3y'(|t|)+2|y|\}$$

so that

$$\frac{E_0}{|y(t)|}\le\varepsilon\{2+3\frac{y'(|t|)}{|y(t)|}\}$$

showing that the expression

$$\frac{y'(|t|)}{|y(t)|}$$

plays an important role in controlling the error during evaluation.

CLENSHAW ALGORITHM

With the polynomial y(t) expressed in the Chebyshev form,

$$y(t)=\sum_{s=0}^{N}{}' a_s T_s(t)$$

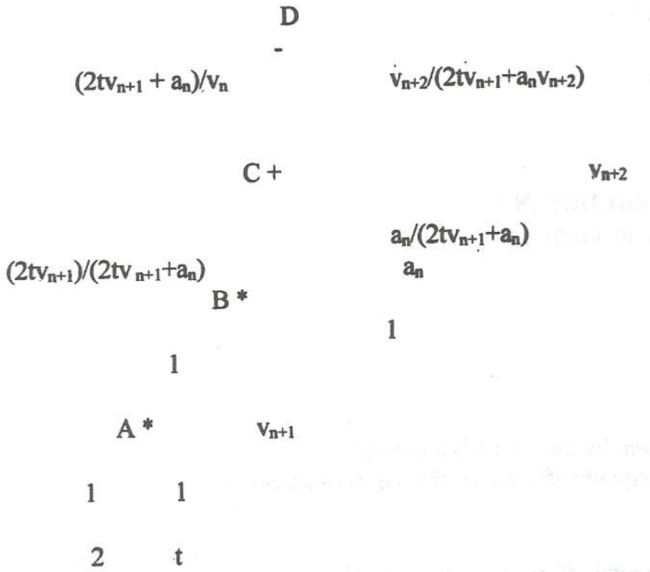
where T is a chebyshev polynomial of degree s, the Clenshaw algorithm takes the form

$$v_{N+1}=0, v_N=a_N$$
$$v_n=2tv_{n+1}+a_n-v_{n+2}$$
$$y(t)=\frac{1}{2}(v_0-v_2)$$

with n=N-1,N-2,...0.

The process graph for the algorithm is given just below.

We would assume that the integer 2 could be represented exactly without truncation of digits. this is reasonably as almost all types of computer could represent 2 exactly.



If we denote as r_{nb} r_{nc} r_{nd} the round_off error at B, C, D and E_n the error in v_n then, using the process graph above, the relative error E_n in v_n satisfies the recurrence relation

$$E_n = 2t\,E_{n+1} - E_{n+2} + 2t(\varepsilon + r_{na} + r_{n\beta} + r_{nc} + r_{nd})\,v_{n+1}$$

$$(\alpha_n + r_{nc} + r_{nd})\,a_n - v_{n+2}\,r_{nd}$$

If e_y is the relative error in y then it is easily shown that

$$e_y = \{\,\varepsilon_t\,y' + (r_- + r_*)y + 2t\sum_{i=0}^{N}{}'(r_{ia} + r_{i\beta} + r_{ic} + r_{nd})\,v_{i+1}T_i$$

$$+\sum_{i=0}^{N}{}'(\alpha_i + r_{ic} + r_{nd})\,a_iT_i + \sum_{i=0}^{N}{}'\,r_{id}\,v_{i+2}T_i\,\}\,y$$

The first term again represents the effect of round-off error in t.

Apart from

$$\sum_{i=0}^{N}{}'\,\alpha_i\,a_iT_i$$

which shows the effects of inexact coefficients the other terms represent the effect of round off error during calculation.

EFFECT OF TRUNCATION OF COEFFICIENT

The effect of truncation of coefficients is

$$e_y = (\sum \alpha_i a_i T_i)/y$$

so that

$$| e_y | \leq \frac{\varepsilon \sum | a_i T_i |}{| y |}$$

EFFECT OF INEXACT VALUES IN t

The effect of inexact values in t is given by

$$| e_y | = \frac{| \varepsilon y' |}{| y(t) |}$$

$$\leq \varepsilon \frac{y'(| t |)}{| y(t) |}$$

which is the same as that given by the Horner's algorithm.
This indicates that both are equally affected by the inexact values of t.

COMBINED EFFECT

The relative error due to the total effect of all the types of error is

$$e_y = \{ \varepsilon_t y' + (r_{\cdot} + r_{\bullet}) y + 2t \sum_{i=0}^{N}{}' (r_{ia} + r_{i\beta} + r_{ic} + r_{id}) v_{i-1} T_i$$

$$\sum_{i=0}^{N}{}' (\alpha_i + r_{ic} + r_{id}) a_i T_i + \sum_{i=0}^{N}{}' r_{id} v_{i-2} T_i \}/y$$

$$| e_y | \leq \varepsilon \{ | y' | + 2 | y | + 8 \sum_{i=0}^{N}{}' | v_{i-1} | | T_i | + 3 \sum_{i=0}^{N}{}' | a_i T_i |$$

$$\sum_{i=0}^{N}{}' | v_{i-2} T_i | \} / | y |$$

Though we may not be able to simplify this further, the expression shows that the error bound is dependent on the ratio y'(|t|)/|y(t)|. This bound is obviously greater than the produced by Horner's algorithm.

CONCLUSION

The analysis of relative error propagation of Horner's and clenshaw's algorithms for evaluating polynomials was undertaken using process graph. Three sources of error namely the error due to inexact coefficient, round off error and inexact value, of t, at which the polynomial is evaluated, are identified and the contribution of each of these to the overall relative error is determined.

It is also shown that the important quantity that controls the error bound in the polynomial evaluation is max (y'(|t|)/|y(t)|) rather than the exaggerated value

$$\frac{|\, y'(t)\,|_{\max}}{|\, y(t)\,|_{\min}}$$

given by Schonfelder and Razaz[2].

The magnitude of the effect of error due to the inexact coefficient is dependent on

$$\varepsilon \max \frac{y(|\,t\,|)}{|\, y(t)\,|}$$

which simplifies to ε for t>0.

It therefore follows that the error due to inexact coefficient when t>0 is not greater than the maximum error in the coefficient started with.

REFERENCES

[1] NEWBERY A.C.R (1974). Error analysis for polynomial evaluation. maths comput. 28 pp 789.

[2] RAZAZ M. & SCHONFELDER J.L. (1980). Error control with polynomial approximations IMA J.Num Analysis 1, 105.

[3] OLIVER J.(1979) Rounding error propagation in polynomial evaluation, J. of Comput & App. Maths vol. 5 No. 2.

[4] OLIVER J. (1982) The accurate evaluation of polynomial approximations to Library function, IMA J. of Num. Analysis No. 2 pp 63-72.