Technical Report 67-52 September 1967

Bounds for Horner Sums

by

Manfred Reimer

Visiting Research Assistant Professor

Computer Science Center

The completion of this report was in part supported by the National Aeronautics and Space Administration under grant NsG - 398.

Abstract

The study of the effect of round-off in Horner's scheme leads to the problem of estimating the absolute values of the so-called Horner sums. In this report the problem is solved under the condition that the polynomial is either an odd or an even function and that its maximum -norm does not exceed the value one. Except for a few specific cases, the Chebyshev-polynomials then turn out to be the maximizing polynomials.

Bounds for Horner Sums[1]

by

Manfred Reimer[2]

1. Introduction

In an earlier paper, Reimer and Zeller [1] proved the following maximum property for the Chebyshev-polynomials C_n:

Consider all real polynomials

(1.) $$P(x) = a_o + a_1 x + \cdots + a_n x^n$$

satisfying

(1.2) $$\|P\| = \max_{-1\le x\le 1} |P(x)| \leqq 1$$

and

(1.3) $$P \text{ is } \begin{Bmatrix}\text{even}\\ \text{odd}\end{Bmatrix} \text{ if } n \text{ is } \begin{Bmatrix}\text{even}\\ \text{odd}\end{Bmatrix}.$$

Among these polynomials, C_n is a polynomial maximizing the absolute value of each partial sum

(1.4) $$S_i(P) = a_o + a_1 + \cdots + a_i \qquad (o \le i \le n),$$

or, equivalently, the polynomial

$$a_o + a_1 x + \cdots + a_i x^i$$

has a maximal Chebyshev-norm if $P = C_n$.

The study of the effect of round-off-errors in Horner's scheme leads to the problem of estimating the absolute values of the Horner sums

(1.5) $$H_i(P) = a_i + a_{i+1} + \cdots + a_n \quad (1 \le i \le n).$$

If P satisfies the condition (1.3), a crude partial solution for this problem can be obtained as follows (see [2]): The trivial

---

1) Computer Science Center, University of Maryland.

2) The completion of this work was in part supported by the National Aeronautics and Space Administration under grant NsG - 398.

relation

(1.6) $$S_{i-1}(P) + H_i(P) = P(1) \qquad (1 < i \leq n)$$

implies that

(1.7) $$|H_i(P)| \leq 1 + |S_{i-1}(P)| \leq 1 + |S_{i-1}(C_n)|.$$

Hence, since the $S_{i-1}(C_n)$ have alternating signs, the estimate

(1.8) $$|H_i(P)| \leq |H_i(C_n)|$$

is evidently best possible in half of all the cases. We shall prove here that there are only few exceptions for which (1.8) is not valid.

2. Lemmas

Let K be a positive integer, r one of the numbers 0 and 1, and

$$n = 2k + r.$$

If P is a real polynomial of degree n satisfying (1.3) then necessarily

$$P(x) = x^r \cdot p(x^2), \tag{2.1}$$

where p is a polynomial of degree k. In particular therefore

$$C_n(x) = x^r \cdot c(x^2). \tag{2.1.1}$$

We introduce the polynomials

$$u_\nu(x) = x^\nu (x-1)^{k-\nu} \qquad (\nu = 0,1,\dots,k) \tag{2.2}$$

as a basis for the space of all polynomials of degree k.

Then

$$c = \sum_{\nu=0}^{k} \binom{2k+r}{2\nu+r} u_\nu, \tag{2.3}$$

and as shown in [2] the condition (1.2) implies that

$$|A_\nu| \le \binom{2k+r}{2\nu+r}$$

where p is assumed to have the representation

$$p = \sum_{\nu=0}^{k} A_\nu u_\nu. \tag{2.4}$$

However, (1.2) involves one more restriction on the A_ν.

Lemma 1.

Let $A_k = 1$ and assume that the polynomial (2.4) has k zeros in the interval (0,1). Then each A_ν $(0 \le \nu \le k-1)$ is a nonnegative and strictly increasing function of each of the zeros within this interval.

Proof. Consider the mapping

$$z = \frac{x}{x-1}, \quad x = \frac{z}{z-1}. \tag{2.4}$$

If $x_1, x_2, \dots, x_k$ are the zeros of p in (0, 1) and $z_1, z_2, \dots, z_k$ their images under z, then

$$(z-1)^k \, p(x(z)) = \sum_{\nu=0}^{k} A_\nu z^\nu = \prod_{\nu=1}^{k} (z-z_\nu),$$

$$-\infty < z_\nu < 0 \quad (\nu = 1, 2, \dots, k).$$

Since z is on (0, 1) a strictly decreasing function of x, the statement of Lemma 1 is now evident.

Lemma 2.

Let x_ν and y_ν $(\nu = 1, \dots, k)$ denote the zeros of the non-zero polynomials

$$p = \sum_{\nu=0}^{k} A_\nu u_\nu \text{ and } q = \sum_{\nu=0}^{k} B_\nu u_\nu,$$

and suppose these zeros have been arranged as follows:

(2.5.1) $\quad 0 \leq y_k < y_{k-1} < \dots < y_1 < 1,$

(2.5.2) $\quad y_k \leq x_k \leq y_{k-1} \dots \leq x_1 < 1.$

Then

(2.6) $\quad \dfrac{A_\nu}{A_k} \geq \dfrac{B_\nu}{B_k} \geq 0 \qquad (\nu = 0, 1, \dots, k-1).$

Proof. $x_1 \neq 1$, $y_1 \neq 1$ implies that $A_k \neq 0$, $B_k \neq 0$ and therefore (2.6) follows directly from Lemma 1.

We shall now specialize the y_ν of Lemma 2 to

(2.7) $\quad y_\nu = \cos^2 \dfrac{\nu\pi}{n} \qquad (\nu = 0, 1, \dots, k);$

then (2.5.1) holds. In this case, the polynomial q of Lemma 2 can be defined as follows:

Case 1. $r = 0$. Since $y_1, \dots, y_{k-1}$ are extreme points of $c(x)$ and since $y_k = 0$, we are led to the relation

$$q(x) = \frac{1}{k} \cdot x \cdot c'(x).$$

Together with (2.2), (2.3) this results in

$$(2.8.1) \qquad q(x) = \sum_{\nu=1}^{k} \binom{2k}{2\nu-1} u_\nu(x).$$

Case 2. $r = 1$. In this case, $y_1, \ldots, y_k$ are contained in $(0, 1)$ and are extreme points of the function

$$C_n(\sqrt{x}) = \sqrt{x} \cdot c(x).$$

Thus

$$q(x) = \frac{2}{2k+1} \cdot \sqrt{x} \cdot \frac{d}{dx} (\sqrt{x} \cdot c(x)) = \frac{1}{2k+1} \{C(x) + 2x\, C'(x)\}$$

is a polynomial of degree k with $y_1, \ldots, y_k$ as its zeros. Using again (2.2), (2.3) we obtain therefore

$$(2.8.2) \qquad q(x) = \sum_{\nu=0}^{k} \binom{2k+1}{2\nu} u_\nu(x).$$

3. The Main Theorem

Let P be a polynomial of the form (1.1) satisfying the conditions (1.2) and (1.3) and let p be defined by (2.1). Obviously, we have then

$$(3.1) \qquad \begin{aligned} S_i(p) &= S_{2i+r}(P), \\ H_i(p) &= H_{2i+r}(P) \end{aligned} \qquad (i = 0, 1, \ldots, k)$$

and it can be verified easily that

$$(3.2) \qquad \begin{aligned} S_i(u_\nu) &= 0 && (0 \leq i < \nu \leq k), \\ (-1)^{k+i} S_i(u_\nu) &\geqq 1 && (0 \leq \nu \leq i < k). \end{aligned}$$

For the moment let us suppose that

$$(1.2.1) \qquad \| P \| < 1.$$

Then each of the polynomials $C_n + P$ and $C_n - P$ has a zero between each pair of successive extreme points of C_n. Passing over to c and p we see that each of the polynomials c+p and c-p satisfies the conditions placed upon p in Lemma 2 provided that q is defined by (2.8.1) and (2.8.2), respectively. This remains true even if we replace (1.2.1) by the original condition (1.2) provided we add the assumption that

(3.3) $$P(1) \neq \pm 1.$$

Let

$$w = \sum_{\nu=0}^{k} A_\nu u_\nu$$

be one of the polynomials c+p and c-p; then

$$A_k = w(1) = 1 \pm p(1) > 0.$$

From (1.6) and (3.2) it follows that

(3.4) $$(-1)^{k+i} H_i(w) = (-1)^{k+i} A_k + \sum_{\nu=0}^{i-1} A_\nu \left| S_{i-1}(u_\nu) \right| \qquad (1 \leq i \leq k).$$

Using $A_k > 0$, (3.2) and (2.8.1) or (2.8.2), whatever the case may be, together with Lemma 2 we obtain from (3.4) the estimate

(3.5) $$(-1)^{k+i} H_i(w) \geqq A_k \left\{ (-1)^{k+i} + \sum_{\nu=0}^{i-1} \frac{1}{2k+r} \binom{2k+r}{2\nu+r-1} \right\} \qquad (1 \leq i \leq k).$$

Suppose now that one of the following conditions holds:

(3.6.1) $\quad 1 \leq i \leq k, \; i \equiv k \bmod 2;$

(3.6.2) $\quad 2 \leq i \leq k, \; i \not\equiv k \bmod 2.$

Then, for both of the two possible choices of w, the right-hand-side of (3.5) is nonnegative and it follows that

$$\left[H_i(c)\right]^2 - \left[H_i(p)\right]^2 = H_i(c+p) \,.\, H_i(c-p) \overset{(>)}{=} 0$$

where the equality sign occurs at best when

(3.7) $\quad i = 2, \; k \equiv 1 \bmod 2, \; r = 0.$

Because of (3.1) this finally leads to the estimate

(3.8) $$\left| H_{2i+r}(P) \right| \overset{(<)}{=} \left| H_{2i+r}(C_n) \right|.$$

Let us now drop the conditon (3.3), i.e. let us assume that

$$P(1) = \pm 1$$

(for the following we select a fixed sign). By continuity (3.6) then remains valid and more precisely

$$(3.9) \qquad (-1)^{k+i} H_i(c \pm p) \ (\overset{>}{=}) \, 0$$

holds under the same conditons as above. However, if $w = c \mp p$ then $A_k = 0$. Assume that

$$P \neq C_n, \ P \neq -C_n;$$

then it is a well-known fact (Markoff's inequality) that

$$|P'(1)| < C_n'(1).$$

This implies that $x = 1$ is a simple root of $C_n \mp P$ and likewise of $c \mp p$. However, we have

$$A_{k-1} = w'(1) > 0$$

and w satisfied the condition (2.5.2) for p in Lemma 2 if equality is permitted also in the rightmost inequality. Consequently zero is at best a simple root. Using again mapping (2.4) and applying Descartes' rule to $A_o + A_1 z + .. + A_{k-1} z^{k-1}$ we find that

$$A_o \geq 0, \ A_\nu > 0 \qquad (\nu = 1, 2, \ldots, k-1)$$

and hence (3.4) implies that

$$(3.10) \qquad (-1)^{k+i} H_i(c \mp p) > 0 \qquad (2 \leq i \leq k).$$

Therefore (3.8) is obtained from (3.9), (3.10) and again without the equality sign in the case

$$(3.11) \qquad 2 \leq i \leq k; \ i \neq 2 \text{ if } r = 0 \text{ and } k \equiv 1 \bmod 2; \ P \neq \pm C_n.$$

Finally we observe the self-evident fact that (3.8) is valid for

$$(3.6.3) \qquad i = 0, \ k \geq 1.$$

Moreover, the cases covered by (3.6.1), (3.6.2) and (3.6.3) are obviously exactly those excluded by the condition

E: $i = 1$; $r = 0$ or 1; $k = 2, 4, 6, ---$

Altogether we have therefore obtained the following result:

Theorem.

Let $P(x) = a_n x^n + a_{n-2} x^{n-2} + a_{n-4} x^{n-4} + ---$

be a real polynomial satisfying

$$\|P\| = \max_{-1 \le x \le 1} |P(x)| \le 1.$$

Let $C_n(x) = \alpha_n x^n \quad \alpha_{n-2} x^{n-2} + \alpha_{n-4} x^{n-4} + ---$

be the Chebyshev-polynomial of degree n. Then

(3.12) $$| a_\nu + a_{\nu+2} + - + a_n | \overset{\le}{(=)} | \alpha_\nu + \alpha_{\nu+2} + --- + \alpha_n |$$

is valid for

$$0 \le \nu \le n, \quad \nu \equiv n \bmod 2$$

except in the following cases

E_o : $\nu = 2$; $n = 4, 8, 12, ---$,

E_1 : $\nu = 3$; $n = 5, 9, 13, ---$.

If in (3.12) equality holds and if one of the following conditions is satisfied

$$\nu \ge 6;\ n = 6, 10, 14, ---,$$

(3.13) $$\nu \ge 4;\ n = 4, 8, 12, ---,$$

$$\nu \ge 5;\ n = 5, 7, 9, ---$$

then

$$P = \pm C_n.$$

Proof. The statements of the theorem are self-evident in the case $n = 0$ and $n = 1$. If $n \ge 2$ set $\nu = 2i + r$ and recall the meaning of $H_\nu(P)$. The exceptions E_o and E_1 correspond to E for $r = 0$ and $r = 1$, respectively, and (3.12) is identical to (3.8) Finally, (3.13) is a decomposed version of (3.11). Thus the theorem has been proved in its entirety.

Note that the result does not apply in the exceptional cases

E_o and E_1. In fact, the example

$$P(x) = C_6(x) - K \cdot x \cdot C_6'(x)$$

shows that when (3.13) is violated the equality in (3.12) does not imply that $P = \pm C_n$.

In this case the assumptions about P made in the theorem are satisified for some interval

$$0 \leq K < K_o.$$

Yet, because of

$$x \cdot C_6'(x) = 192\, x^6 - 192\, x^4 + 36\, x^2$$

we have

$$H_4(P) = H_4(C_6)$$

for any choice of K.

4. Exceptional Cases.

We shall now discuss the situation when one of the conditions E_o and E_1 applies.

E_o. Because of

$$H_2(P) = P(1) - P(0)$$

(1.2) implies that

(4.1) $$|H_2(P)| \leq 2 \qquad (n = 4, 8, 12, ---).$$

The example $P = C_2$ then demonstrates that the bound in (4.1) is best possible. However, since $C_n(1) = C_n(0)$, this bound is not attained for $P = C_n$.

E_1. Assume for the moment that

(4.2) $$\|P\| < 1.$$

Since $C_n(\lambda)$ attains each value between -1 and +1 within the interval

(4.3) $$\cos\frac{\pi}{n} < \lambda < 1,$$

we can choose an s in this interval such that

$$P(1) = C_n(s).$$

Then $C_n(sx) - P(x)$ is an odd polynomial with exactly k positive roots between 1 and the smallest positive extreme point of $C_n(sx)$. This implies that

$$P'(0) < \left[\frac{d}{dx} C_n(sx)\right]_{x=0} = s\, C_n'(0),$$

because otherwise an additional zero of $C_n(sx) - P(x)$ could be found in (0, 1). Thus

$$P'(0) - P(1) < s\, C_n'(0) - C_n(s).$$

Since $-P$ satisfies the same conditions as P there is another number t in the interval (4.3) with

$$-P'(0) + P(1) < t\, C_n'(0) - C_n(t).$$

Now

$$P(1) - P'(0) = a_3 + a_5 + \cdots + a_n = H_3(p),$$

and thus

$$|H_3(p)| \leq \max_{\cos\frac{\pi}{n} \leq \lambda \leq 1} |\lambda\, C_n'(0) - C_n(\lambda)|$$

holds, even if we admit (1.2) instead of the condition (4.2). The maximum on the right can be determined by elementary means; it is assumed only at $\lambda = \cos\frac{\pi}{n}$. Therefore

$$|H_3(P)| \leq 1 + n \cos\frac{\pi}{n} \qquad (n = 5, 9, 13, \cdots), \tag{4.4}$$

and the bound is attained by

$$P(x) = \pm\, C_n\left(x \cdot \cos\frac{\pi}{n}\right).$$

5. References

[1] Reimer, M. and Zeller, K.: Abschätzung der Teilsummen reeller Polynome. Math. Zeitschr. 99, 101-104 (1967).

[2] Reimer, M.: Normenschranken für die Horner-Summen. To appear soon in Z. angew. Math. Mech.