MATHEMATICS OF COMPUTATION, VOLUME 26, NUMBER 118, APRIL 1972

Minimum Norm Differentiation Formulas with Improved Roundoff Error Bounds

By David K. Kahaner*

Abstract. Numerical differentiation formulas of the form $\sum_{i=1}^{N} w_i f(x_i) \approx f^{(m)}(a), \alpha \leqq x_i \leqq \beta$, $\alpha \leqq a \leqq \beta$, are considered. The roundoff error of such formulas is bounded by a value proportional to $\sum_{i=1}^{N} |w_i|$. We consider formulas that have minimum norm $\sum_{i=1}^{N} w_i^2$ and converge to $f^{(m)}(a)$ as $\beta - \alpha \to 0$. The resulting roundoff error bounds can be several orders of magnitude less than corresponding bounds for high order differences.

Introduction. Many computations require the evaluation of derivatives. Often, a function $f(x)$ is known at only a discrete set of points as a result of another computation, and some derivative of f is needed. The usual situation is to approximate the derivative with a linear combination of function values,

$$f^{(m)}(0) \approx D_m f \equiv \sum_{i=1}^{N} w_i f(x_i), \qquad h = \max(x_{i+1} - x_i). \tag{1}$$

The summation is over a subset of the points at which $f(x)$ is computed.

Even if $f(x)$ were known exactly (e^x, $\sin x$), its representation in a finite word length computer involves some error. More often, the approximations to $f(x_i)$ contain substantial amounts of error. These errors tend to be magnified by the process (1), especially if the points x_i are close together. In estimating derivatives by such a procedure, one finds that answers become more accurate at first as h is decreased, and, subsequently, exhibit decreased accuracy [1]. It is therefore important to have a rigorous bound on the total error of the computation in order to select the most appropriate value of h.

The most often used approximations for (1) are obtained by taking the mth derivative of an interpolating polynomial. If this polynomial is of degree $m = N - 1$ and the x_i are uniformly spaced, $x_{i+1} - x_i = h$, its mth derivative is an "mth difference." For an mth difference, the truncation error, which can be obtained by expanding each $f(x_i)$ in a Taylor series about $x = 0$, is proportional to h, unless the points x_i are symmetrically placed with respect to 0, whence the truncation error is proportional to h^2.

To bound the computational, or roundoff error, we note that if a number α is represented on a computer, with d decimal digit word length, it can be written $\alpha = \alpha_c + \alpha_\epsilon$, with α_c the computer version of α and $|\alpha_\epsilon| \leqq \frac{1}{2} 10^{-d} |\alpha|$. The quantity $10^{-d}/2$

Received April 29, 1971.

AMS 1970 *subject classifications.* Primary 65D25, 65G05.

Key words and phrases. Numerical differentiation, minimum norm differentiation, least squares differentiation, roundoff error analysis.

* This research was supported by the U.S. Atomic Energy Commission under Contract W-7405-ENG-36 at the Los Alamos Scientific Laboratory of the University of California, Los Alamos, New Mexico 87544.

is known as the machine accuracy parameter, denoted ϵ_M. If we now consider computing (1), we may write

$$D_m f = [D_m f]_c + [D_m f]_\epsilon.$$

Doing inner products double precision and rounding afterward, we get

$$3\epsilon_M \max |f_i| \sum_{i=1}^{N} |w_i| \geqq [D_m f]_\epsilon. \tag{2}$$

If (1) represents an mth difference [1],

$$D_m f = \frac{\Delta^m f(0)}{h^m} = \sum_{k=0}^{m} (-1)^{m-k} \frac{\binom{m}{k} f(x_k)}{h^m} \tag{3}$$

and $\sum |w_i| = (2/h)^m$. Thus, the roundoff error bound is

$$3\epsilon_M \max |f_i| \, (2/h)^m. \tag{4}$$

For these formulas, it is quite apparent that the roundoff error increases rapidly with h for large m whereas the truncation error remains proportional to h or h^2. A potential improvement may be obtained by increasing the order of the approximation, i.e., increasing N relative to m. This may be done in one of two ways:

(a) Keeping h fixed and adding points outside of $[x_1, x_N]$. From this, one usually obtains greater total accuracy, because the truncation term is higher order. Nevertheless, this scheme requires the evaluation of $f(x)$ at points increasingly far from the center, which may not be convenient computationally and may in fact not always be possible because of singularities, etc.

(b) Keeping the total interval length fixed and adding points inside $[x_1, x_N]$ also yields a higher order truncation error and has the advantage of easily allowing extrapolation for different h's as well. In general, neither (a) not (b) will converge for $N \to \infty$, [2], although they seem to work satisfactorily for small N, m.

Minimum Norm Methods. In contrast to the above, which attempt to reduce the truncation error, we wish to consider the selection of the w_i and N to reduce the roundoff error bound. Our results will indicate possible usefulness for large m and no particular improvement for small m.**

Large roundoff error bounds occur because the weights w_i increase with $1/h$ and are not all positive. Set

$$R_\Delta = \sum_{i=1}^{N} |w_i|.$$

Then

$$R_\Delta \leqq \left(N \sum_{i=1}^{N} w_i^2 \right)^{1/2} \equiv R_B.$$

For each fixed N, we consider the selection of w_i so that $\sum_{i=1}^{N} w_i^2$ is minimized.

Fix $N > m$ and x_i, $i = 1, \cdots, N$. In analogy with the case of an mth difference,

** In [6], selection of the x_i's to minimize the L_2 norm of $f'(x)$, minus a particular approximation thereto, is considered.

we require that our selection of w's be such that (1) is exact for $f(x) = 1, x, \cdots, x^m$. This leads to the system of equations

$$\sum_{i=1}^{N} w_i(x_i)^k = m!\,\delta_{k,m}, \qquad k = 0, \cdots, m. \tag{5}$$

The solution to this system is nonunique for $N > m + 1$.

The solution of minimum L_2 norm is given in the following theorem which involves the functions $U_m(x)$.

Let $\tilde{U}_m(x)$ be a polynomial of degree m which is orthogonal on the set $\{x_1, \cdots, x_N\}$ to all other polynomials of degree less than m. Since $N > m$, this polynomial is known to exist and satisfy a three term recursion, much like orthogonal polynomials on an interval [3]. Further, if we define the polynomial $U_m(x)$ proportional to $\tilde{U}_m(x)$, but orthonormal on $\{x_1, \cdots, x_N\}$,

$$\sum_{i=1}^{N} [U_m(x_i)]^2 = 1,$$

then $U_m(x)$ can be shown to be unique.

THEOREM. *Let $U_m(x)$ be the unique mth degree polynomial, orthonormal on the set $x_1, x_2, \cdots, x_N$. Then*

$$w_i = (m!/\alpha)\,U_m(x_i), \qquad i = 1, \cdots, N,$$

is the unique minimum L_2 norm solution of (5), *where α is defined below, and $\sum w_i^2 = (m!/\alpha)^2$.*

Proof. We have immediately

$$\sum_{i=1}^{N} U_m(x_i)x_i^k = 0, \qquad k < m,$$
$$= \alpha \neq 0, \qquad k = m,$$

where $x^m = \alpha U_m(x) +$ lower order terms. Hence, w_i is a solution to (5). If $\hat{w}_i$, $i = 1, \cdots, N$, is any other solution, set $\beta_i = w_i - \hat{w}_i$. Then

$$\sum_{i=1}^{N} \hat{w}_i^2 = \sum_{i=1}^{N} [w_i^2 - 2\beta_i w_i + \beta_i^2].$$

But

$$\sum_{i=1}^{N} \beta_i w_i = \frac{m!}{\alpha} \sum_{i=1}^{N} \beta_i U_m(x_i) = \frac{m!}{\alpha} \sum_{i=1}^{N} (w_i - \hat{w}_i)\left[\sum_{j=0}^{m} C_j x_i^j\right]$$
$$= \frac{m!}{\alpha} \sum_{j=0}^{m} \left[\sum_{i=1}^{N} w_i x_i^j - \hat{w}_i x_i^j\right] C_j = \frac{m!}{\alpha} \sum_{j=0}^{m} [m!\,\delta_{j,m} - m!\,\delta_{j,m}]C_j = 0.$$

Hence,

$$\sum_{i=1}^{N} \hat{w}_i^2 = \sum_{i=1}^{N} w_i^2 + \beta_i^2.$$

Case of Equally Spaced Points. Let the points x_i be on the interval $[a, b]$:

$$x_i = a + (i - 1)h, \qquad h = (b - a)/(N - 1).$$

We can obtain more explicit information:

THEOREM. *If w_i are the minimum norm solution to* (5) *and* $f \in C^2[a, b]$,

$$
\begin{aligned}
[D_m f] &= \sum_{i=1}^{N} w_i f(x_i) \\
&\to \frac{(2m)!}{m!} \frac{(2m+1)^{1/2}}{(b-a)^{m+1/2}} \int_a^b L_m(x) f(x)\, dx \quad \text{as } N \to \infty,
\end{aligned} \tag{6}
$$

where $L_m(x)$ is the orthonormal Legendre polynomial of degree m, on the interval $[a, b]$.

Proof. Again, $U_k(x)$ is the polynomial of degree k, orthonormal on $\{x_1, \cdots, x_N\}$. With some involved algebra, it can be shown [4] that

$$
\begin{aligned}
U_m(x) &= \frac{h^{-2m}}{(\binom{2m}{m}\binom{N+m}{2m+1})^{1/2}} h^m \frac{(2m)!}{(m!)^3} x^m + \cdots \\
&= \frac{h^{-m}(2m)!}{(\binom{2m}{m}\binom{N+m}{2m+1})^{1/2}(m!)^3} x^m + \cdots .
\end{aligned}
$$

Hence,

$$
w_i = \frac{m!\,(2m)!}{h^m (m!)^3 (\binom{2m}{m}\binom{N+m}{2m+1})^{1/2}} U_m(x_i),
$$

and

$$
\sum_{i=1}^{N} w_i^2 = \frac{h^{-2m}(2m)!}{(N+m)!\,(m!)^2} (2m+1)!\,(N+m-2m-1)!.
$$

With $h = (b-a)/(N-1)$,

$$
\sum_{i=1}^{N} w_i^2 = \frac{(N-1)^{2m}}{(N+m)\cdots(N-m)} \left[\frac{(2m)!}{(b-a)^m m!}\right]^2 (2m+1). \tag{7}
$$

From (7), we note that

$$
N \sum_{i=1}^{N} w_i^2 = \mathcal{O}(1), \qquad N \to \infty .
$$

Also, by the midpoint rule [5],

$$
\begin{aligned}
\int_{a-h/2}^{b+h/2} U_m^2(x)\, dx &= \frac{b-a}{N-1} \sum_{i=1}^{N} U_m^2(x_i) + \mathcal{O}\left(\frac{1}{N^2}\right) \\
&= \frac{b-a}{N-1} + \mathcal{O}\left(\frac{1}{N^2}\right).
\end{aligned}
$$

Hence, $((N-1)/(b-a))^{1/2}\, U_m(x)$ has $\mathcal{L}_2[a - h/2, b + h/2]$ norm equal to $1 + \mathcal{O}(1/N)$. Let $L_m(x)$ be the unique orthonormal Legendre polynomial of degree m on $[a, b]$:

$$
\int_a^b L_m^2(x)\, dx = 1.
$$

Then, [3, p. 290],

$$L_m(x) = \left(\frac{N-1}{b-a}\right)^{1/2} U_m(x) + o(1).$$

So

$$\begin{aligned}
\sum_{i=1}^{N} w_i f(x_i) &= \frac{m!\,(2m)!}{((b-a)/(N-1))^m (m!)^3 (\binom{2m}{m}\binom{N+m}{2m+1})^{1/2}} \sum_{i=1}^{N} U_m(x_i) f(x_i) \\
&= \frac{m!\,(2m)!}{((b-a)/(N-1))^m (m!)^3 (\binom{2m}{m}\binom{N+m}{2m+1})^{1/2}} \\
&\qquad \cdot \left[\frac{(b-a)^{1/2}}{(N-1)^{1/2}} \sum_{i=1}^{N} L_m(x_i) f(x_i) + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right)\right] \\
&= \frac{m!\,(2m)!}{((b-a)/(N-1))^m (m!)^3 (\binom{2m}{m}\binom{N+m}{2m+1})^{1/2}} \\
&\qquad \cdot \left[\frac{(N-1)^{1/2}}{(b-a)^{1/2}} \left(\frac{b-a}{N-1}\right) \sum_{i=1}^{N} L_m(x_i) f(x_i) + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right)\right] \\
&= \frac{m!\,(2m)!}{((b-a)/(N-1))^m (m!)^3 (\binom{2m}{m}\binom{N+m}{2m+1})^{1/2}} \\
&\qquad \cdot \left[\left(\frac{N-1}{b-a}\right)^{1/2} \int_{a-h/2}^{b+h/2} L_m(x) f(x)\,dx + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right)\right] \\
&= (N-1)^m \frac{(2m)!}{(b-a)^m m!} \left(\frac{2m+1}{(N+m)\cdots(N-m)}\right)^{1/2} \\
&\qquad \cdot \left[\left(\frac{N-1}{b-a}\right)^{1/2} \int_a^b L_m(x) f(x)\,dx + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right)\right].
\end{aligned}$$

Letting $N \to \infty$,

$$\sum_{i=1}^{N} w_i f(x_i) \to \frac{(2m)!}{m!} \frac{(2m+1)^{1/2}}{(b-a)^{m+1/2}} \int_a^b L_m(x) f(x)\,dx.$$

Since for each N, $\sum w_i f_i$ differentiates mth degree polynomials exactly at $x = 0$, so does (6). Formula (6) is in some sense a canonical minimum norm formula. We now investigate the truncation error in this expression.

THEOREM. *Let the points a, b, be symmetric with respect to zero (this is for convenience only) with $a \equiv -H$, $b \equiv H$. If $f^{(m+1)}(x)$ is bounded on $[-H, H]$, then*

$$\begin{aligned}
&f^{(m)}(0) - \frac{(2m)!}{m!} \frac{(2m+1)^{1/2}}{(2H)^{m+1/2}} \int_{-H}^{H} L_m(x) f(x)\,dx \\
&\qquad \leqq \frac{(2m)!}{m!\,(m+1)!} \left(\frac{2m+1}{2m+3}\right)^{1/2} \frac{H}{2^m} \max_{[-H, H]} |f^{(m+1)}|.
\end{aligned}$$

Proof. The Legendre polynomials $L_m(x)$ on $[-H, H]$ are related to those $L_m^*(x)$ on $[-1, 1]$ by $L_m(x) = (1/\sqrt{H}) L_m^*(x/H)$. So

$$\frac{(2m)!}{m!}\frac{(2m+1)^{1/2}}{(2H)^{m+1/2}}\int_{-H}^{H} L_m(x)f(x)\,dx = \frac{(2m)!}{m!}\frac{((2m+1))^{1/2}}{(2H)^{m+1/2}} \cdot \int_{-H}^{H} L_m(x)\left\{f(0)+xf'(0)+\cdots+\frac{x^m}{m!}f^{(m)}(0)+\frac{x^{m+1}f^{m+1}(\xi)}{(m+1)!}\right\}dx. \tag{8}$$

Since the operator (8) differentiates polynomials of degree m exactly, we get

$$= f^{(m)}(0) + \frac{(2m)!}{m!}\frac{((2m+1))^{1/2}}{(2H)^{m+1/2}}\int_{-H}^{H} L_m(x)\frac{x^{m+1}f^{(m+1)}(\xi(x))}{(m+1)!}\,dx$$

$$= f^{(m)}(0) + \frac{(2m)!}{m!}\frac{((2m+1))^{1/2}}{(2H)^{m+1/2}}\frac{1}{\sqrt{H}}\int_{-H}^{H}\frac{L_m^*\left(\frac{x}{H}\right)x^{m+1}f^{(m+1)}(\xi(x))}{(m+1)!}\,dx$$

$$= f^{(m)}(0) + \frac{(2m)!}{m!}\frac{(2m+1)^{1/2}}{(2H)^{m+1/2}}\frac{1}{\sqrt{H}}H^{m+2}\int_{-1}^{+1}\frac{L_m^*(t)t^{m+1}f^{(m+1)}(\xi(tH))}{(m+1)!}\,dt$$

$$= f^{(m)}(0) + \frac{(2m)!}{m!}\frac{(2m+1)^{1/2}}{2^{m+1/2}}H\int_{-1}^{1}\frac{L_m^*(t)t^{m+1}f^{(m+1)}(\xi(tH))}{(m+1)!}\,dt.$$

Thus, if $f^{(m+1)}$ is bounded on the interval $[-H, H]$, this latter term goes to zero with H. This corresponds to the truncation error for mth differences. There, the error goes to zero with h, the mesh spacing, here, with the total interval length. A bound on the truncation error is

$$\frac{(2m)!}{m!\,(m+1)!}\frac{(2m+1)^{1/2}}{2^{m+1/2}}H\max|f^{(m+1)}|\int_{-1}^{1}|L_m^*(t)t^{m+1}|\,dt$$

$$\leqq \frac{(2m)!}{m!\,(m+1)!}\frac{(2m+1)^{1/2}}{2^{m+1/2}}H\max|f^{(m+1)}|\left(\int L_m^{*2}\,dt\int t^{2m+2}\right)^{1/2}dt$$

$$= \frac{(2m)!}{m!\,(m+1)!}\left(\frac{2m+1}{2m+3}\right)^{1/2}\frac{H}{2^m}\max|f^{(m+1)}|.$$

COROLLARY. *If $f^{(m+2)}$ is bounded on $[-H, H]$, the truncation error is $O(H^2)$. Now, we require that the interval is symmetric about zero.*

Proof. This is immediate if we use one more term in the Taylor series expansion of $f(x)$ and note that $L_m(x)x^{m+1}$ is an odd function on $[-H, H]$, hence integrates to zero. This corollary is analogous to the result for central differences.

To summarize this section, the operators (1), with w_i selected as the minimum norm solution to (5) and (6), provide approximations to $f^{(m)}(0)$ which are exact for mth degree polynomials. The roundoff error in (1) is bounded by

$$3\epsilon_M \max|f_i|\,R_\Delta \leqq 3\epsilon_M \max|f_i|\,R_B$$

which for sufficiently large N is, by (7),

$$3\epsilon_M \max|f_i|\,R_\Delta \leqq 3\epsilon_M \max_{[a,b]}|f(x)|\frac{(2m)!\,(2m+1)^{1/2}}{(2H)^m m!}.$$

If $x = 0$ is the center of the interval, the truncation error in (6) is bounded by

$$\frac{(2m)!}{m!\,(m+2)!}\left(\frac{2m+1}{2m+5}\right)^{1/2}\frac{H^2}{2^m}\max|f^{(m+2)}|.$$

The roundoff error bound for mth central differences is $3\epsilon_M \max|f_i|\,(2m/2H)^m$.

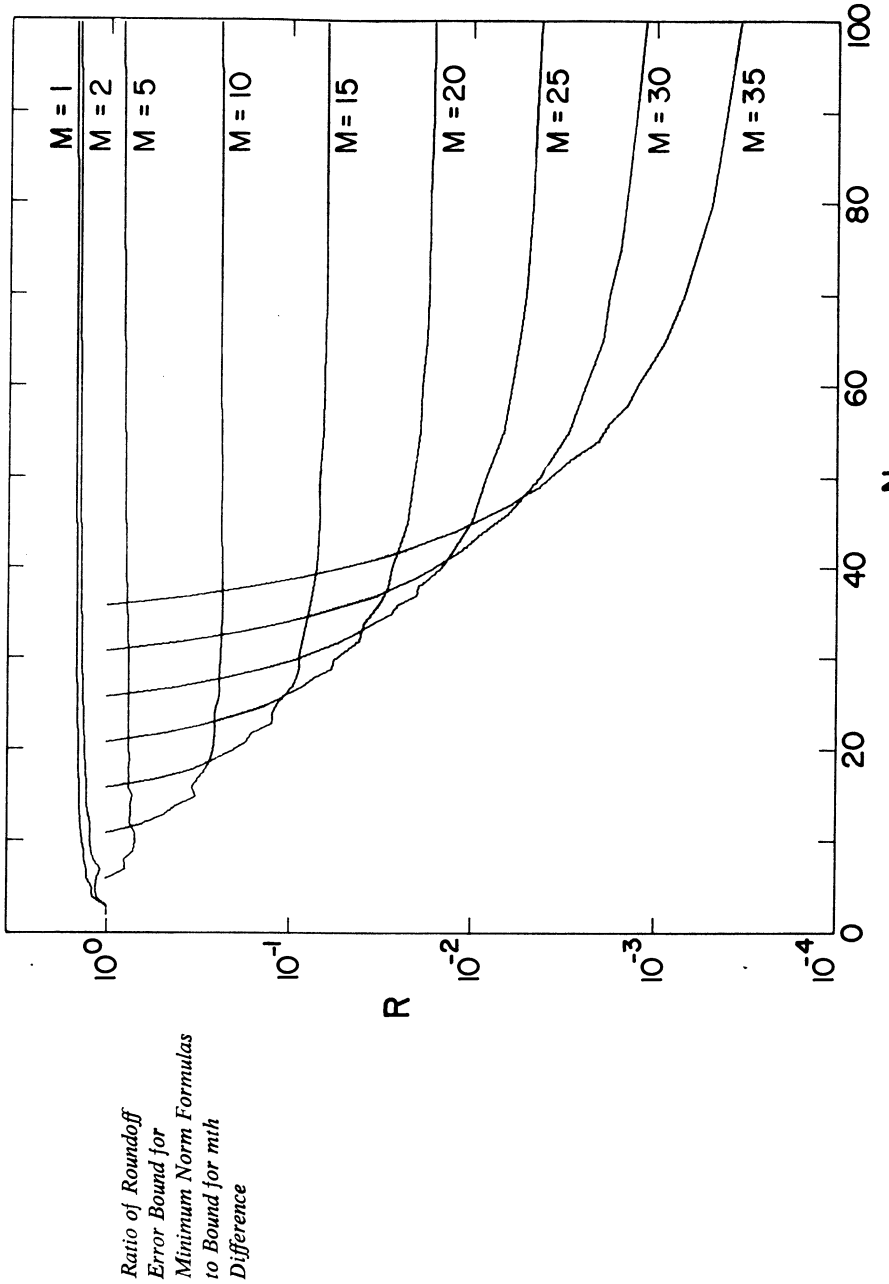

FIGURE I. *Roundoff Error Bounds for Minimum Norm Differentiation Formulas Divided by Roundoff Bounds for mth Difference*

We have

$$\left(\frac{2m}{2H}\right)^m \Big/ \frac{(2m)!\ (2m+1)^{1/2}}{(2H)^m m!} = \frac{(2m)^m m!}{(2m)!\ (2m+1)^{1/2}}$$

$$\sim \frac{(2m)^m m^{m+1/2} e^{-m}}{(2m)^{2m+1/2} e^{-2m} (2m+1)^{1/2}} = \frac{m^{m+1/2} e^m}{2^{m+1/2} m^{m+1/2} (2m+1)^{1/2}}$$

$$= \left(\frac{e}{2}\right)^m \frac{1}{\sqrt{2}(2m+1)^{1/2}} \to \infty \quad \text{as } m \to \infty .$$

While both roundoff error bounds are $\mathcal{O}(1/H^m)$, the difference bound is substantially greater for large m. As far as the truncation error is concerned, the situation is reversed. Both (6) and the mth difference have $O(H^2)$ bounds, with the coefficients of the latter being smaller for large m than (6).

If we use a higher order interpolating polynomial rather than an mth difference, the roundoff bound will increase and that comparison will be more favorable to (6), whereas the truncation error bound will decrease.

A calculation shows that for equally spaced points the minimum norm weights are given by

$$w_i = \frac{i!}{m!\ \binom{N+m}{2m+1}} \left(\frac{N-1}{2}\right)^m$$

$$\cdot \sum_{\nu=0}^{\min\left[m, i\right]} \frac{(m+\nu)!}{(m-\nu)!\ (\nu!)^2} (-1)^{m-\nu} \frac{(N-\nu-1)\cdots(N-m+1)(N-m)}{(i-\nu)!},$$

with $R_\Delta = \sum_{i=0}^{N} |w_i|$. We have computed R_Δ for various values of m and N. Graphs of some of these calculations are included in Fig. I. Each curve has been scaled so that its value at $N = m + 1$ is 1, i.e., we have divided the ordinate values of each curve by the roundoff error bound for the corresponding mth difference. When m is small, $m \leqq 3$, R_Δ assumes its minimum near $N = m + 1$. For larger m, $m \geqq 10$, R_Δ makes a substantial decrease, about one order of magnitude, with about $m/10$ additional function evaluations. Further increasing N reduces R_Δ more slowly to a broad minimum, although not monotonely, and the limiting value of R_Δ tends to be slightly above its minimum but substantially less than the mth difference value. Thus, for $m = 35$, R_Δ is reduced about four orders of magnitude over the corresponding bound for a 35th difference.

Conclusion. We have defined minimum norm differentiation formulas and shown they exist by exhibiting them. Whether these will become useful remains to be seen. If they turn out to have application, it will clearly be for higher derivatives. In any case, the results indicate an interesting alternative way of looking at rounding problems.

University of California
Los Alamos Scientific Laboratory
Los Alamos, New Mexico 87544

1. A. RALSTON, *A First Course in Numerical Analysis*, McGraw-Hill, New York, 1965. MR **32** #8479.
2. P. J. DAVIS, *Interpolation and Approximation*, Blaisdell, Waltham, Mass., 1963. MR **28** #393.
3. F. B. HILDEBRAND, *Introduction to Numerical Analysis*, McGraw-Hill, New York, 1956. MR **17,** 788.
4. C. JORDAN, *Calculus of Finite Differences*, Chelsea, New York, 1950.
5. P. J. DAVIS & P. RABINOWITZ, *Numerical Integration*, Blaisdell, Waltham, Mass., 1967. MR **35** #2482.
6. A. SCHÖNHAGE, "Optimale Punkte für Differentiation und Integration," *Numer. Math.*, v. 5, 1963, pp. 303–331. MR **29** #1481.