A method for numerical integration on an automatic computer.

CLENSHAW, C.W.; CURTIS, A.R.

in: Numerische Mathematik | Numerische Mathematik - 2
9 Seite(n) (197 - 205)

Numerische Mathematik 2, 197—205 (1960)

A method for numerical integration on an automatic computer

By

C. W. CLENSHAW and A. R. CURTIS

Abstract

A new method for the numerical integration of a "well-behaved" function over a finite range of argument is described. It consists essentially of expanding the integrand in a series of Chebyshev polynomials, and integrating this series term by term. Illustrative examples are given, and the method is compared with the most commonly-used alternatives, namely SIMPSON's rule and the method of GAUSS.

1. Introduction

We consider in this paper the basic problem in numerical integration, the evaluation of the integral of a function $f(x)$ known numerically within a finite range of integration $a \leqq x \leqq b$. For the user of an automatic computer, it is desirable that a program to perform this task should give the result to a specified accuracy in a wide range of circumstances, and give an indication of the reduced accuracy achieved when these circumstances do not obtain.

Finite-difference methods are, in general, poorly suited to automatic computation because their storage requirement is large. Lagrangian formulae, which involve only function values, are usually preferred. The two most commonly used are probably SIMPSON's rule and GAUSS' formula. The former has the attraction of simplicity of form, with convenient binary weights, while the latter minimises the number of ordinates needed.

We propose here a method based on the term-by-term integration of the expansion of $f(x)$ in Chebyshev polynomials. A unique advantage of the method is that its accuracy may be checked easily *before* the integration is completed. It also shares some of the advantages of the other two methods mentioned; as with SIMPSON's rule, the number of the ordinates may be doubled if necessary without previous work being wasted, while considerable economy in the number of ordinates is also achieved. Moreover, the method will provide, with little extra complication, values of the indefinite integral $\int_a^x f(x)\, dx$ throughout the range (a, b). In contrast, SIMPSON's rule yields values of the indefinite integral at tabular points only, while GAUSS' formula is quite unsuited to indefinite integration.

Although we are primarily concerned with the integration of a non-singular function in a finite range, we may nevertheless observe that an infinite range may frequently be transformed to a finite range, or approximated by a large finite range. Further, some integrands with weak singularities are amenable to treatment with the proposed method. In general, however, special types of integral such as

$$\int_0^\infty e^{-x} f(x)\,dx \quad \text{and} \quad \int_{-\infty}^{+\infty} e^{-x^2} f(x)\,dx,$$

are probably best evaluated by special methods such as the Laguerre-Gauss and Hermite-Gauss formulae respectively (see, for example, HILDEBRAND [1] or KOPAL [2]). The second of these integrals can also be evaluated efficiently by using the simple summation formula of GOODWIN [3]. For integrals of oscillating functions over an infinite range, LONGMAN [4] has described a simple method based on the Euler transformation of series.

2. The methods of SIMPSON and GAUSS

SIMPSON's rule is given by

$$\int_0^{2h} f(x)\,dx \doteq \tfrac{1}{3}h\{f(0) + 4f(h) + f(2h)\}. \tag{1}$$

This is the three-point "Newton-Cotes" formula. The corresponding N-point formula is obtained by integrating the interpolation polynomial of degree $N-1$ which is equal to the integrand $f(x)$ at each of N equally-spaced values of the argument x. In practice, we subdivide a given range (a, b) into an even number $2m$, say, of equal segments, and use the summed form of (1), namely

$$\int_a^b f(x)\,dx \doteq \frac{1}{6m}(b-a)\,(f_0 + 4f_1 + 2f_2 + 4f_3 + 2f_4 + \cdots + 4f_{2m-1} + f_{2m}), \tag{2}$$

where

$$f_r \equiv f\left\{\frac{(2m-r)\,a + r\,b}{2m}\right\}.$$

Similar summation procedures may be adopted with the other Newton-Cotes formulae. In general, the formulae involving an *even* number of segments are preferred, since each of them has the same order of accuracy as the formula involving one more segment. Of these preferred formulae, SIMPSON's rule is the simplest, with weights which are powers of two. The higher order formulae have coefficients which must be stored, and moreover these coefficients are large and vary in sign when N is large, thereby magnifying the effect of rounding errors.

If (2) is to yield a satisfactory result, m must be made large enough to ensure that the truncation error is negligible. This may imply that many function values have to be evaluated; for functions which can be rapidly computed, however, this is usually considered worthwhile in order to permit the use of a simple and flexible formula. In such cases, however, in order to deal with functions having regions of rapidly changing behaviour, a programmer often finds it convenient to allow the interval to vary as the integration proceeds from step to step. The variation is usually controlled by comparing at a given step the approximate integrals obtained using two intervals of integration, h and $2h$, say. If these results agree, then an interval $4h$ is tried, and so on until disagreement is found, whereupon the last agreed value is accepted, and the corresponding intervals are used to start the next step. If the first two results disagree, the interval is repeatedly halved until agreement is reached.

Comparison of two results is commonly used to assess the accuracy of an integral evaluated on an automatic computer by SIMPSON's rule applied at a constant interval. However, this check is unreliable, as may be shown by a simple example. If

$$f(x) = \tfrac{23}{25}\cosh x - \cos x, \qquad a = -1, \qquad b = 1,$$

the values obtained from (2) are 0.4795546 for $m=1$, and 0.4795551 for $m=2$, whereas the true value of the integral is 0.4794282.

The possibility of check failure can be greatly reduced by comparing at least *three* successive results; although the extent of the computation is thereby increased, this seems a highly desirable safeguard for a general quadrature program.

The Gauss integration formula (HILDEBRAND [*1*] or KOPAL [*2*]) is given by

$$\int_a^b f(x)\,dx \doteq (b-a)\sum_{r=1}^{n} H_r f(x_r), \tag{3}$$

where

$$x_r = \tfrac{1}{2}(b-a)\,t_r + \tfrac{1}{2}(b+a),$$

$$H_r = \frac{1}{(1-t_r^2)\,\{P_n'(t_r)\}^2},$$

and $t_1, t_2, \ldots, t_n$ are the zeros of the Legendre polynomial $P_n(t)$. Formula (3) is exact if $f(x)$ is a polynomial of degree not exceeding $2n-1$. This considerable power is achieved at the expense of introducing a set of irrational arguments and weights, whose values need to be stored.

The checking of the Gauss formula is more difficult than that of SIMPSON's rule. Repetition with a new n requires a completely new set of abscissae x_r and weights H_r, and we cannot make use of previously computed ordinates if n is increased. The most commonly used check for an isolated integration is to compare the results of applying the Gauss formula at n and at $n+1$ points. That this is not a completely satisfactory procedure is shown by another example; the Gauss three- and four-point formulae give 1.585026 and 1.585060 respectively for the integral

$$\int_{-1}^{+1} \frac{dx}{x^4 + x^2 + 0.9},$$

whereas its correct value is 1.582233.

However, there is an important class of problems for which the use of the Gauss formula is advantageous. If it is necessary to integrate over a given range several functions whose behaviour within the range is similar, it may be possible to find an n adequate for *all* cases by a thorough investigation of a small sample.

3. Method of Chebyshev expansion

The proposed method derives its power from the economy with which a function can be represented by its expansion in Chebyshev polynomials, and the ease with which such a series can be integrated. If $f(x)$ is continuous and of bounded variation in (a, b), then it can be expanded in the form

$$f(x) \equiv F(t) = \tfrac{1}{2} a_0 + a_1 T_1(t) + a_2 T_2(t) + \cdots, \qquad (a \leqq x \leqq b), \tag{4}$$

where

$$T_r(t) = \cos(r \cos^{-1} t), \qquad t = \frac{2x - (b+a)}{b-a}, \tag{5}$$

(see N. B. S. [5]). On integration we have

$$\frac{2}{b-a} \int_a^x f(x)\, dx = \int_{-1}^t F(t)\, dt = \tfrac{1}{2} b_0 + b_1 T_1(t) + b_2 T_2(t) + \cdots, \tag{6}$$

where

$$b_r = \frac{a_{r-1} - a_{r+1}}{2r}, \qquad (r = 1, 2, \ldots) \tag{7}$$

(see CLENSHAW [6]).

The value of b_0 is determined by the lower limit of integration; thus

$$b_0 = 2b_1 - 2b_2 + 2b_3 - \cdots. \tag{8}$$

The definite integral is given by

$$\begin{aligned} \frac{2}{b-a} \int_a^b f(x)\, dx = \int_{-1}^{+1} F(t)\, dt &= \tfrac{1}{2} b_0 + b_1 + b_2 + \cdots \\ &= 2(b_1 + b_3 + b_5 + \cdots), \end{aligned} \tag{9}$$

while the indefinite integral is given by the sum of the series (6), the evaluation of which is described later.

The coefficients in the expansion (4) may be calculated after first observing that any *polynomial* of degree N in x may be written in the form

$$\begin{aligned} f(x) \equiv F(t) &= \tfrac{1}{2} a_0 + a_1 T_1(t) + a_2 T_2(t) + \cdots + a_{N-1} T_{N-1}(t) + \tfrac{1}{2} a_N T_N(t), \\ &= \sum_{r=0}^{N}{}'' a_r T_r(t), \qquad (-1 \leqq t \leqq 1). \end{aligned} \tag{10}$$

Here $\sum''$ denotes a finite sum whose first and last terms are to be halved.

The coefficients in (10) are then given by

$$a_r = \frac{2}{N} \sum_{s=0}^{N}{}'' F_s \cos \frac{\pi r s}{N}, \tag{11}$$

where

$$F_s = F\left(\cos \frac{\pi s}{N}\right).$$

This is a consequence of the orthogonality of the cosine function with respect to the points $t_s = \cos \frac{\pi s}{N}$, expressed by the equations

$$\sum_{s=0}^{N}{}'' \cos \frac{\pi i s}{N} \cos \frac{\pi j s}{N} = \left\{ \begin{array}{lll} 0 & (i \neq j) & \\ N & (i = j = 0 & \text{or} \quad N) \\ \tfrac{1}{2} N & (i = j \neq 0 & \text{or} \quad N) \end{array} \right\}. \tag{12}$$

Finally, *any* function which satisfies the conditions necessary for its Chebyshev expansion to converge can be approximated to any required accuracy by a finite series of the form (10) the coefficients of which are given by (11).

It is important to observe that (11) can be written in the form

$$a_r = \frac{2}{N} \sum_{s=0}^{N}{}'' F_s T_s(t_r), \qquad t_r = \cos\frac{\pi r}{N}. \tag{13}$$

This form is similar to that of (10), so that the same machine subroutine can be used to evaluate the right-hand sides of these relations. Such a subroutine may be based on the recurrence method described by CLENSHAW [7], which we now outline briefly.

The series

$$F(t) = \tfrac{1}{2} a_0 + a_1 T_1(t) + \cdots + a_n T_n(t) \tag{14}$$

is evaluated by calculating successively the numbers $c_n, c_{n-1}, \ldots, c_0$ from the relation

$$c_r = 2t\, c_{r+1} - c_{r+2} + a_r, \qquad (r \leqq n), \tag{15}$$

starting with $c_{n+1} = c_{n+2} = 0$. The sum of the series is then

$$F(t) = \tfrac{1}{2}(c_0 - c_2), \tag{16}$$

This result may be verified by substituting for the a_r in (14) their expressions in terms of the c_r from (15) and using the recurrence relation for the Chebyshev polynomials

$$T_{r+1}(t) - 2t\, T_r(t) + T_{r-1}(t) = 0. \tag{17}$$

It only remains to consider the determination of the value of N. We propose that if no information about the behaviour of the integrand is available, the Chebyshev coefficients be evaluated using an arbitrary initial value of N in the formula (11). We then examine the coefficients of the Chebyshev polynomials of highest order in the integrated series; if these are small we may assume the chosen N to be adequate. We consider the smallness of three consecutive coefficients of a convergent series to be a reliable indication of the negligibility of *all* higher order coefficients. For instance, in the case of indefinite integration we might require b_N, $k b_{N-1}$ and $k^2 b_{N-2}$ to be negligible, where k is an arbitrary constant in the range (0, 1). The value of k could be taken to be unity; however, a smaller value will often save considerable computing time while only slightly increasing the chance of check failure. We suggest the use of $k = \frac{1}{8}$.

In the evaluation of a definite integral alternate terms vanish, and we suggest as a criterion that three successive non-zero coefficients should be small. This "safe" number of three may, of course, be increased if we desire to reduce even further the possibility, already remote, of an erroneous result.

If b_N, $k b_{N-1}$ and $k^2 b_{N-2}$ are *not* all negligible, we replace N by $2N$ and again compute the a_r and b_r of high order. The previously computed values of F_s will again be used; they are the alternate members of the new set

$$F_s = F\left(\cos\frac{\pi s}{2N}\right), \qquad (s = 0, 1, \ldots, 2N). \tag{18}$$

It is convenient to take $N=4$ first, since no smaller value can give an acceptable number of negligible coefficients, and then to proceed to $N=8$, $N=16$ and so on, to an upper limit $N=M$, dictated by the requirements of the programmer and the characteristics of the computer.

4. The case of slow convergence

The only situation remaining for consideration is that in which the desired accuracy is not attained at $N=M$. In such a case, we seek to determine the reduced accuracy achieved.

We observe that if the integrand may be expanded in an *infinite* Chebyshev series

$$F(t) = \tfrac{1}{2}A_0 + A_1 T_1(t) + A_2 T_2(t) + \cdots = \sum_{i=0}^{\infty}{}' A_i T_i(t), \tag{19}$$

then the coefficients a_r given by (11) are related to the A_r by

$$\begin{aligned} a_r &= \frac{2}{N} \sum_{s=0}^{N}{}'' \cos\frac{\pi r s}{N} \sum_{i=0}^{\infty}{}' A_i \cos\frac{\pi i s}{N} \\ &= A_r + A_{2N-r} + A_{2N+r} + A_{4N-r} + A_{4N+r} + A_{6N-r} + \cdots . \end{aligned} \tag{20}$$

The approximate value of $\int_{-1}^{+1} F(t)\,dt$, calculated from the coefficients a_r is

$$I_N = a_0 - \frac{2a_2}{1.3} - \frac{2a_4}{3.5} - \cdots - \frac{2a_{N-2}}{(N-3)(N-1)} - \frac{a_N}{(N-1)(N+1)}, \tag{21}$$

with N even, whereas the correct value of this integral is

$$I = A_0 - \frac{2A_2}{1.3} - \frac{2A_4}{3.5} - \cdots . \tag{22}$$

Assuming that the coefficients A_r may be neglected when $r \geqq 3N$, we may approximate the error $I_N - I$ by

$$\begin{aligned} E_N = 2A_{2N} &- \frac{2}{1.3}(A_{2N-2} + A_{2N+2}) - \cdots - \frac{2}{(N-3)(N-1)}(A_{N+2} + A_{3N-2}) + \\ &+ \sum_{r=1}^{N-1} \frac{2A_{N+2r}}{(N+2r-1)(N+2r+1)} . \end{aligned} \tag{23}$$

Since the series (19) converges for $-1 \leqq t \leqq 1$, the coefficients A_r satisfy an inequality of the form

$$|A_r| \leqq \frac{K_N}{r}, \quad (r \geqq N), \tag{24}$$

where K_N is independent of r, and is assumed to be as small as (24) will allow. Then we have

$$\begin{aligned} |A_{2N-2s} + A_{2N+2s}| &\leqq \tfrac{1}{2} K_N \left(\frac{1}{N-s} + \frac{1}{N+s}\right) \\ &= K_N \frac{N}{N^2 - s^2} < \frac{K_N}{N}\left(1 + \frac{4s^2}{3N^2}\right), \quad (s < \tfrac{1}{2}N) . \end{aligned} \tag{25}$$

It is clear that the coefficient of $A_{2N\pm 2s}$ in (23) does not exceed $\frac{2}{(2s-1)(2s+1)}$ in absolute value for $s=1, 2, \ldots, \frac{1}{2}N-1$. Therefore

$$\begin{aligned} |E_N| &< K_N\left\{\frac{1}{N} + \sum_{s=1}^{\frac{1}{2}N-1} \frac{2}{4s^2-1}\,\frac{1}{N}\left(1+\frac{4s^2}{3N^2}\right)\right\} \\ &= K_N\left\{\frac{2}{N} - \frac{1}{N(N-1)} + \frac{2}{3N^3}\sum_{s=1}^{\frac{1}{2}N-1}\frac{4s^2}{4s^2-1}\right\} < \frac{2}{N}\,K_N. \end{aligned} \tag{26}$$

The largest of $2\,|a_{N-4}|$, $2\,|a_{N-2}|$ and $|a_N|$ may conveniently be used as an approximation to $2K_N/N$. Thus, if the upper limit $N=M$ is reached before the desired accuracy is attained, we may assess the magnitude of the error to be that of the largest of these quantities, or, more conveniently and almost equivalently, the largest of $4N\,|b_{N+1}+b_{N-1}+b_{N-3}|$, $4N\,|b_{N+1}+b_{N-1}|$ and $4N\,|b_{N+1}|$ (see (7)).

In the case of an indefinite integral similar considerations apply, but here *all* the coefficients b_r are required, and the estimate of $2K_N/N$ may therefore be based on the values of b_{N+1}, b_N and b_{N-1} instead of b_{N+1}, b_{N-1} and b_{N-3}. This procedure is illustrated in the second of the examples which follow.

5. Examples

As our first example we consider again the definite integral,

$$\int_{-1}^{+1} \frac{dx}{x^4+x^2+0.9}\,,$$

and suppose we require the result to six-decimal accuracy. Taking $N=4$, we evaluate the integrand at $x_s=\cos\frac{1}{4}\pi s$ $(s=0, 1, 2, 3, 4)$ and, applying formula (11) with $r=4$, obtain a_4 and hence $b_5=+0.00609544$, which is not negligible to six decimal places. We therefore take $N=8$, evaluate the integrand at the intermediate points and apply (11) with $r=8$ to obtain a_8, and hence $b_9=-0.00007329$, which is still not negligible. Again doubling N we find a new a_{16}, and $b_{17}=+0.00000002$ which can be neglected. We therefore calculate a_{14}, keeping $N=16$, and thence $\frac{1}{8}b_{15}$, and, finding this to be only $+0.00000004$ we proceed to a_{12} and $\frac{1}{64}b_{13}=-0.00000007$. This also is sufficiently small, so we accept $N=16$ as being large enough, and proceed to compute the terms b_r (r odd), shown in Table 1.

Finally, the integral is estimated by the present method to be

$$2(b_1+b_3+b_5+\cdots)=1.58223296,$$

which is correct to eight decimals.

As an example of indefinite integration which also illustrates the problem of slow convergence we consider

$$\int_{-1}^{x} |x+\tfrac{1}{2}|^{\frac{1}{2}}\,dx.$$

The coefficients b_r produced with $N=4$, 8 and 16 are unacceptably large, and a computer program would continue to double N until sufficient accuracy was obtained, or until it reached its limit M. For a program working to about 30 binary figures, $M=64$ might be chosen, but for the purpose of illustration we take $M=16$. We find for the coefficients b_r in the integrated series the values given in Table 2.

Table 1

r	b_r
1	+0.85844113
3	−0.07354558
5	+0.00645162
7	−0.00015279
9	−0.00010230
11	+0.00002844
13	−0.00000436
15	+0.00000030
17	+0.00000002

Table 2

r	b_r	r	b_r
1	+0.707670	9	+0.000062
2	+0.127592	10	−0.001338
3	+0.020533	11	+0.001061
4	−0.022044	12	−0.000180
5	+0.008786	13	−0.000427
6	+0.001172	14	+0.000516
7	−0.004192	15	−0.000161
8	+0.002548	16	−0.000178
		17	+0.000118

From (8) we obtain $b_0=1.250724$, and from Table 2 we find that the values of $64\,|b_{17}+b_{16}+b_{15}|$, $64\,|b_{17}+b_{16}|$ and $64\,|b_{17}|$ are 0.014144, 0.003840 and 0.007552 respectively. Hence we conclude that the indefinite integral may contain errors up to 0.0141. As a check, the approximate value of the definite integral is

$$\int_{-1}^{+1} |x+\tfrac{1}{2}|^{\frac{1}{2}}\,dx=\tfrac{1}{2}b_0+b_1+\cdots=1.466900,$$

compared with the correct value 1.460447 ,and is thus in error by about 0.0065.

It may be of interest to note that a program based on the proposed method has been prepared for the DEUCE at the National Physical Laboratory. The upper limit M of the number of terms is 64, and the program evaluated the integral $\int_{-1}^{+1}|x+\frac{1}{2}|^{\frac{1}{2}}\,dx$ with an error of 0.00078. This may be compared with the error of 0.00317 committed by the 32-point Gauss formula, and of 0.00036 by the 64-point Gauss formula. We see that the Chebyshev formula, which is much more convenient than the Gauss, may sometimes nevertheless be of comparable accuracy.

Conclusion

The principal advantages of the Simpson and Gauss methods are to some extent shared with the present method of Chebyshev expansion.

SIMPSON'S rule has simple and convenient weights, and its accuracy is easy to check because the interval may be halved without wasting the previous computations. Its main disadvantage is that many ordinates may be required.

The method of Gauss requires the least possible number of ordinates. Its most serious drawback is that this economy may be lost in checking the accuracy of a result, because the weights and abscissae of the n-point formulae are completely different for each value of n.

The present method effects a compromise. It is as easy to check as SIMPSON'S rule, yet achieves considerable economy in the number of ordinates required, and will readily provide values of an indefinite, as well as a definite, integral. This combination of economy and reliability make it a very suitable basis for general subroutines for numerical integration on an automatic computer.

The work described above has been carried out as part of the research programme of the National Physical Laboratory and is published by permission of the Director of the Laboratory.

References

[1] HILDEBRAND, F. B.: Introduction to numerical analysis. New York: McGraw-Hill 1956.

[2] KOPAL, Z.: Numerical analysis. London: Chapman & Hall 1955.

[3] GOODWIN, E. T.: Evaluation of integrals of the form $\int_{-\infty}^{+\infty} f(x) e^{-x^2} dx$. Proc. Cambridge Phil. Soc. **45**, 241—245 (1949).

[4] LONGMAN, I. M.: Note on a method for computing infinite integrals of oscillatory functions. Proc. Cambridge Phil. Soc. **52**, 764—768 (1956).

[5] National Bureau of Standards Appl. Math. Series No. 9. Tables of Chebyshev Polynomials. Washington: Government Printing Office 1952.

[6] CLENSHAW, C. W.: The numerical solution of linear differential equations in Chebyshev series. Proc. Cambridge Phil. Soc. **53**, 134—149 (1957).

[7] CLENSHAW, C. W.: A note on the summation of Chebyshev series. Math. Tab., Wash. **9**, 118 (1955).

Mathematics Division,
National Physical Laboratory,
Teddington, Middlesex

(Received December 21, 1959)