SMOOTHEST APPROXIMATION FORMULAS

By Arthur Sard[1]
Queens College

Introduction. Consider a process of approximation which operates on a function $x = x(t)$. The error in the process may be thought of as a sum $R + \delta A$, where R is the error that would be present if x were exact and δA is the error due to errors in x. (Precise definitions are given below.) Suppose that one wishes to choose one process A from a class $\mathfrak{A}$ of processes. In some situations it is appropriate to base the choice on R alone[2]; in others it is appropriate to consider δA.

The primary purpose of the present note is to formulate a criterion of smoothest approximation: That A in $\mathfrak{A}$ is smoothest which minimizes the variance of δA. A criterion based on both R and δA is also suggested. (Sections 1 and 2.) Smoothest approximate integration formulas of one type are derived in Section 3.

Progress in the technique of estimating the covariance function of the errors in x will lead to further applications of the criterion of smoothest approximation.

1. Approximation of a functional. Suppose that X is a space of functions $x = x(t)$ each of which is continuous on $a \leqq t \leqq b$. Let $f[x]$ be a functional defined on X; that is, $f[x]$ is a real number defined for each $x \in X$. For example, X might be the space of functions with second derivatives on $[a, b]$ and $f[x]$ might be $x''(u)$, where u is a fixed number in $[a, b]$.

Suppose that $f[x]$ is to be approximated by a Stieltjes integral

$$A = \int_a^b x(t)\, d\alpha(t), \qquad x \in X, \tag{1}$$

where α is a function of bounded variation. The remainder in the approximation of $f[x]$ by A is

$$R = A - f[x].$$

If the approximation (1) operates on $x + \delta x$ instead of x, the result is $A + \delta A = \int_{t=a}^b (x + \delta x)\, d\alpha$; and the error in the approximation of $f[x]$ by $A + \delta A$ is $R + \delta A$, where

$$\delta A = \int_a^b \delta x(t)\, d\alpha(t). \tag{2}$$

Consider a class $\mathfrak{A}$ of approximations A, each of the form (1). We shall propose a criterion for characterizing the "smoothest A" in $\mathfrak{A}$, relative to the covariance function of the errors δx.

---

[1] The author gratefully acknowledges financial support received from the Office of Naval Research.

[2] "Best approximate integration formulas; best approximation formulas," *Amer. Jour. of Math.*, Vol. 71 (1949), pp. 80–91.

Assume that $\delta x = \delta x(t)$ is a stochastic process with mean zero[3] and covariance function $\sigma(t, u) = E[\delta x(t)\delta x(u)]$. Then, by (2), δA is a stochastic variable; and[4]

$$(3)\qquad \begin{aligned} E\delta A &= E\int_{t=a}^{b} \delta x\, d\alpha = \int_{t=a}^{b} 0\, d\alpha = 0, \\ E(\delta A)^2 &= v = E\left[\int_a^b \delta x(t)\, d\alpha(t) \int_a^b \delta x(u)\, d\alpha(u)\right] = \int_a^b\int_a^b \sigma(t,u)\, d\alpha(t)\, d\alpha(u). \end{aligned}$$

CRITERION. *That A (if any) in $\mathfrak{A}$ is smoothest which minimizes the variance v of δA.*

In particular cases, this criterion (least squares) has been proposed and used by Chebyshev and others. An application to approximate integration is given in section 3 below.

One may extend this discussion to cases in which the approximations A involve derivatives of x.

Remark. The criterion of *best* approximation[2] may be combined with the above criterion of *smoothest* approximation as follows: That A (if any) in $\mathfrak{A}$ is the best compromise which minimizes a specified combination of the variance of δA and the modulus of R. Here it is assumed that the remainders R satisfy the conditions for the existence of the modulus.[2]

2. Approximation of a function. One may extend the preceding discussion to the case in which $y = f[x]$ is an operation to a space of functions $y = y(u)\cdot$ $\bar{a} \leqq u \leqq \bar{b}$; and in which the approximation of $f[x]$ is

$$A = \int_a^b x(t)\, d_t\, \alpha(t, u), \qquad x \in X,$$

where, for each u, α is a function of bounded variation in t. Then, for each u, δA has a variance $v(u)$. *Criterion.* That A (if any) in a class of approximations is smoothest which minimizes $v(u)$ for all u; failing such an A, that A (if any) is smoothest which minimizes the integral of $v(u)$, or alternatively, the supremum of $v(u)$, over $\bar{a} \leqq u \leqq \bar{b}$.

3. Smoothest approximate integration formulas in a particular case.[5] Let m and n be fixed integers; $m \geqq 1$, $n \geqq 0$. *Let $\mathfrak{A} = \mathfrak{A}_{m,n}$ be the class of all approximations of*

[3] The essential point here is that $E\delta(t) = m(t)$ be known for each t; for given $m(t)$, one could and would replace $x + \delta x$ by $x + \delta x - m$.

[4] We assume here that the integrals in (3) exist and that the inversions of E and $\int d\alpha$ are valid. For this it is sufficient that δx be integrable relative to the product measure $\alpha\omega$ for all functions α corresponding to elements of $\mathfrak{A}$, where ω is the measure in the underlying probability space relative to which E is the operator $\int d\omega$. Cf. J. L. Doob, "Probability in function space," *Bull. Amer. Math. Soc.*, Vol. 53 (1947), especially pp. 26, 27.

[5] The approximate integration formulas of this section are of such a nature that one would expect them to be known. The values of J at the end are probably new.

$$\int_{-m/2}^{m/2} x(t)\,dt = f[x]$$

of the form

$$A = \sum_{i=-m/2}^{m/2} b_i x(i),$$

the $m + 1$ *constants* b_i *being such that* $A = f[x]$ *whenever* $x(t)$ *is a polynomial of degree* n. Throughout this section i is to range over the $m + 1$ values $i = -m/2, -m/2 + 1, \cdots, + m/2$. *Suppose that the errors* $\delta x(i)$ *are independent, with common variance* σ^2, *and with mean zero.* Then $\alpha(t)$ is a step function with jumps b_i at $t = i$; and

$$v = \sigma^2 \sum_i b_i^2 .$$

The smoothest approximation in $\mathfrak{A}_{m,n}$ is the one for which v is a minimum. (The $m + 1$ variables b_i in v are subject to $n + 1$ constraints due to the condition that the approximation be exact for degree n. The set $\mathfrak{A}_{m,n}$ is empty if and only if m is less than the largest even integer contained in n.)

If $n = 0$ or 1, the smoothest formula in $\mathfrak{A}_{m,n}$ is the one for which all the coefficients are equal:

$$b_i = m/(m + 1);$$

in which case

$$v = m^2\sigma^2/(m + 1).$$

If $n = 2$ or 3, the smoothest formula in $\mathfrak{A}_{m,n}$ is characterized by the following relations:

$$b_i = \lambda_0 + i^2\lambda_1 ,$$
$$\lambda_0 = m(2m^2 + 9m - 6)/2(m - 1)(m + 1)(m + 3),$$
$$\lambda_1 = -30m/(m - 1)(m + 1)(m + 2)(m + 3);$$

in which case

$$v/\sigma^2 = \lambda_0 m + \lambda_1 m^3/12.$$

Thus, the smoothest approximation in $\mathfrak{A}_{6,2}$ or in $\mathfrak{A}_{6,3}$ is the following:

$$A = \tfrac{1}{2}[x(-3) + x(3)] + \tfrac{6}{7}[x(-2) + x(2)] + \tfrac{15}{14}[x(-1) + x(1)] + \tfrac{8}{7}\, x(0).$$

By the method of Lagrange's multipliers, one may establish the following relations for the smoothest formula in $\mathfrak{A}_{m,n}$. Here i has the same range of values as before; μ and ν range over $0, 1, \cdots, [n/2]$.

$$b_i = \sum_\mu \lambda_\mu i^{2\mu},$$
$$v/\sigma^2 = \sum_\mu \lambda_\mu c_\mu ,$$

where

$$c_\mu = m^{2\mu+1}/4^\mu(2\mu + 1),$$

and λ_μ are determined by the equations

$$\sum_\mu \lambda_\mu \sum_i i^{2(\mu+\nu)} = c_\nu .$$

The class $\mathfrak{A}_{m,n}$ is such that for each $A \epsilon \mathfrak{A}_{m,n}$ there is a function $k(t)$ with the following property:[2]

$$R = A - f[x] = \int_{-m/2}^{m/2} x^{(n+1)}(t)k(t)\, dt,$$

whenever x is a function with continuous $(n + 1)$th derivative. The quantity

$$J = \int_{-m/2}^{m/2} k^2(t)\, dt$$

is useful in appraising R, since

$$R^2 \leqq J \int_{-m/2}^{m/2} x^{(n+1)}(t)^2\, dt,$$

by Schwarz's inequality.

Values of J for the smoothest formulas are as follows.

$$n = 0 : J = m^2/6(m + 1).$$

$$n = 1 : J = m^2(3m^2 + 2m + 1)/360(m + 1).$$

For $n = 2$ and 3, and $m \leqq 6$, the numerical values of J are as follows.

m	J $(n = 2)$.	J $(n = 3)$.
2	1/1,890	1/9,072
3	11/8,960	13/17,920
4	134/33,075	62,539/13,891,500
5	1,865/150,528	136,223/6,322,176
6	8/245	6,683/82,320

For the method of calculation of J, as well as the transformation of J under a linear transformation of t, the reader may consult the paper[2].