ON THE OPTIMALITY OF THE TRAPEZOIDAL METHOD WHEN INTEGRATING THE WIENER PROCESS

Peter Kopanov

Department of Mathematics, Plovdiv University
4000 Plovdiv, Bulgaria

Abstract

The paper deals with methods for approximate calculation of integrals of stochastic processes. It is shown that when integrating the Wiener process, the classical trapezoidal method is optimal if variance is taken as a criterion.

Keywords: Integral of a stochastic process; Quadrature formula; Wiener process; Trapezoidal method.

Mathematics Subject Classification (1991): 65U05, 65D32.

1 Introduction

In this paper we follow a probabilistic approach to analyze methods of approximate integration of random processes and derive useful conclusions about some classical methods used in numerical integration.

Let us describe the model. Suppose $K \subset \mathbb{R}^n$ is a compact subset of $n$-dimensional Euclidean space, $X = X(K)$ a functional space whose elements are functions defined over $K$ and $L : X \to \mathbb{R}$ is a linear functional. The standard case is to take $K = [0,1]$, the unit interval, and $X = C[0,1]$ the space of all continuous functions on $[0,1]$. 
The general goal is to find $L(x)$ for any $x \in X$. However sometimes it can be done only for special subsets of functions and in such a case we construct a sequence $\{L_n\}$ of approximators of $L$ obeying the following two properties:

(i) For any natural $n \in \mathbb{N}$ and any $x \in X$ the computation of $L_n(x)$ is relatively easy (e.g. usually $L_n$ is given by a quadrature formula).

(ii) For any $x \in X$ one holds: $L_n(x) \to L(x)$ as $n \to \infty$.

Let us note that practically we are able to compute only a finite number of the elements of the infinite sequence $\{L_n(x)\}$ and a natural requirement is to describe a rule for stopping the procedure of computing at some step (stopping rule). A typical example of a stopping rule is defined as follows: Let $\varepsilon > 0$ and the integer $k$ be fixed. We stop the computation at the first step $n, n > k$ when $|L_n(x) - L_{n-k}(x)| < \varepsilon$ and take $L_n(x)$ as an approximation of $L(x)$. We should like $L_n(x)$ to be "good" in some sense for any $\varepsilon > 0$. In particular, there are cases when we require that $|L(x) - L_n(x)| < \varepsilon |L(x)|$. It turns out that such a condition is too strong for the existence of a "good" approximation and of the corresponding stopping rule. We find useful to illustrate this situation in more details. Further on we construct a function $x \in X(K)$ such that for any fixed $M > 0, n, k \in \mathbb{N}, n > k$ we have $L(x) = M$ and the sequence of approximators $\{L_n\}$ is such that $L_{n-k}(x) = \ldots = L_n(x) = 0$. Let us describe explicitly such an example.

Take $X = C[0,1]$, $L(x) = \int_0^1 x(t)dt$, $x \in C[0,1]$ and for $m = n - k, \ldots, n$, let $\{L_m\}$ be a sequence of approximators of $L$ where

$$L_m(x) = \sum_{i=0}^{m} c_i^m x(t_i), \quad 0 = t_0^m < t_1^m < \ldots < t_m^m = 1, \quad c_i^m \geq 0, \quad \sum_{i=0}^{m} c_i^m = 1.$$  

We construct $x(t), t \in [0,1]$ as follows:

(a) $x(t) = 0$ for $t = t_j^m$, $j = 0, 1, \ldots, m$, $m = n - k, \ldots, n - 1, n$;

(b) $x(t) = qM$ for $t = (t_i^m + t_j^m)/2$ if there is no other knot of interpolation in the subinterval $[t_i^m, t_j^m]$;

(c) $x(t)$ is linear between these points.

Obviously $x \in C[0,1]$ and $L_m(x) = 0$ for $m = n - k, \ldots, n$. However $L(x) = qM/2$ and if we choose $q > 2$, then $L(x) > M$. At the same time $L(x) - L_m(x) > M$. Hence a reasonable stopping rule does not exist in this situation.
2 Main problem

In any case involving functions like that one described above we can ask the following natural question: What is the "portion" of these "bad" functions, i.e. functions \( x \in \mathbb{X} \) such that the computation of the functional \( L(x) \) can be performed with error greater than a fixed \( M, M > 0 \)? Such a question leads to the idea to follow a probabilistic approach. Note, however, that we can give a correct answer only in the framework of a well-defined probability model and the problem is reduced to finding the probability of the event \( \{|L(x) - L_n(x)| > M\} \) for fixed \( M > 0 \) and \( n \).

Clearly, at least one step more can be done: if we know the probability \( \mathbb{P}\{|L(x) - L_n(x)| > M\} \) we can compare different methods for numerical integration, e.g. to compare different sequences of approximators \( \{L_n\} \). Of a great importance is any case when the above probability is "small enough", i.e. if for any \( M > 0 \) it is "close" to 0. One of the goals of the present study is just as briefly mentioned above, namely, to find the probability \( \mathbb{P}\{|L(x) - L_n(x)| > M\} \) for given functional space \( \mathbb{X} \), linear functional \( L \) and approximating sequence \( \{L_n\} \).

To reach that goal we have to pass the following three steps:

1) Specify the probability model \((\mathbb{X}, \mathbb{F}, \mathbb{P})\), i.e. we have to define the probability measure \( \mathbb{P} \) on the \( \sigma \)-algebra \( \mathbb{F} \) (usually \( \mathbb{F} \) is the Borel \( \sigma \)-algebra generated by the open subsets of \( \mathbb{X} \)).

2) Define the linear operator \( L \) and the sequence of approximators \( \{L_n\} \) as random variables on the probability space \((\mathbb{X}, \mathbb{F}, \mathbb{P})\).

3) Find, when possible, the distributions of the random variables \( L, L_n \) and of the error \( D_n = L - L_n \).

If we know the distribution of \( D_n \), we can find the probability \( \mathbb{P}\{|L - L_n| > M\} \) exactly, or find its upper bound. Moreover, we can choose a probabilistic criterion for comparing different methods and then look for the best method with respect to this criterion. One possibility is to estimate the variance \( \text{Var}(D_n) \) of the random error \( D_n \). This criterion, being a natural extension of the classical Gauss least squares method, is widely used in stochastics and its applications.
3 Results

Let us focus our attention to the space $C = C[0, 1]$ of all continuous, real valued functions $x(t), \ t \in [0, 1]$ such that $x(0) = 0$. Let $F$ denote the $\sigma$-algebra of Borel subsets of $C$, and $P$ be the standard Wiener measure on $F$. The construction of the Wiener measure is well-known and can be found in the literature, see e.g. Hida (1980) or Lamperti (1977). In this case the natural process, defined on the probability space $(C, F, P)$ is usually called a standard Wiener process.

It is well-known that if $L : C[0, 1] \rightarrow \mathbb{R}$ is a linear continuous functional, then $L(x) = \int_0^1 x(t) \, d\varphi(t)$ where $\varphi \in V[0, 1]$ is a function of bounded variation. (This is exactly the Riesz theorem, see e.g. Kolmogorov and Fomin (1970).) Let now $0 = t_0 < t_1 < t_2 < \cdots < t_n = 1$, be a partition of $[0, 1]$ with a diameter $d_n = \max_{1 \leq i \leq n} (t_i - t_{i-1}) \rightarrow 0$ as $n \rightarrow \infty$ and let

$$T_n(x) = \sum_{i=1}^n x(t_i) [(\varphi(t_i) - \varphi(t_{i-1})].$$

Then $T_n(x) \rightarrow L(x)$ for any $x \in C[0, 1]$ and therefore the distribution of $L$ is a limit of the distributions of $T_n$ as $n \rightarrow \infty$. Thus the first step is to find the distribution of $T_n$. An easy transformation shows that

$$T_n = \sum_{i=1}^n x(t_i) [(\varphi(t_i) - \varphi(t_{i-1})] = \sum_{i=1}^n [\varphi(1) - \varphi(t_{i-1})] (x(t_i) - x(t_{i-1})).$$

The last quantity is a sum of independent and normally distributed random variables:

$$x(t_i) - x(t_{i-1}) \sim N(0, t_i - t_{i-1}).$$

Hence

$$T_n \sim N(0, \sigma_n^2), \ \text{where} \ \sigma_n^2 = \sum_{i=1}^n (\varphi(1) - \varphi(t_{i-1}))^2 (t_i - t_{i-1}).$$

Thus

$$L \sim N(0, \sigma^2) \ \text{with} \ \sigma^2 = \lim_{n \rightarrow \infty} \sigma_n^2 = \int_0^1 (\varphi(1) - \varphi(t))^2 \, dt.$$

Hence the following result is valid:

Proposition 1 Let $L(x) = \int_0^1 x(t) \, d\varphi(t)$. Then the random variable $L$ is normally distributed:

$$L \sim N(0, \sigma^2), \ \text{where} \ \sigma^2 = \int_0^1 (\varphi(1) - \varphi(t))^2 \, dt.$$
If \( \varphi(t) = t \), then \( L(x) = \int_0^1 x(t)dt \), and \( L \sim N(0, 1/3) \). Further on we consider in more details just this case but the idea and the techniques can be used in a more general case.

Let now \( \{L_n\} \) be a sequence of approximators of \( L \) and

\[
L_n(x) = \sum_{i=1}^n c_i x(t_i)
\]

with \( t_i \) chosen as above and \( c_1 \geq 0, c_0 + c_1 + \ldots + c_n = 1 \). By the Riesz theorem

\[
L_n(x) = \int_0^1 x(t) d\varphi_n(t),
\]

where \( \varphi_n(t) \) is a step-wise function with jumps at the points \( t_i \) and size of the jumps proportional to \( c_i \). The error \( D_n(x) = L(x) - L_n(x) \) is also a linear functional and

\[
D_n(x) = \int_0^1 x(t) d(t - \varphi_n(t)).
\]

For \( D_n \) we obtain

\[
D_n \sim N(0, \delta_n^2) \quad \text{where} \quad \delta_n^2 = \int_0^1 (1 - \varphi_n(1) - t + \varphi_n(t))^2 dt.
\]

In other words, we have found the distribution of the error \( D_n \). Now we can compare the errors when using different quadrature formulas and look for the formula with the smallest variance of the error. Since

\[
\text{Var}(D_n) = \delta_n^2 = \int_0^1 (1 - \varphi_n(1) - t + \varphi_n(t))^2 dt,
\]

we want to minimize this quantity. The following transformations hold:

\[
\delta_n^2 = \int_0^1 (1 - t - (\varphi_n(1) - \varphi_n(t)))^2 dt = \\
= \int_0^1 (1 - t)^2 dt - 2 \int_0^1 (1 - t)(\varphi_n(1) - \varphi_n(t)) dt + \int_0^1 (\varphi_n(1) - \varphi_n(t))^2 dt = \\
= \frac{1}{3} - 2 \sum_{i=1}^n d_i \int_{t_{i-1}}^{t_i} (1 - t) dt + \sum_{i=1}^n d_i^2 \int_{t_{i-1}}^{t_i} dt = \\
= \frac{1}{3} + \sum_{i=1}^n [d_i^2 (t_i - t_{i-1}) + d_i (t_i - t_{i-1}) (t_i + t_{i-1} - 2)],
\]

where \( d_i = \varphi_n(1) - \varphi_n(t) \) for \( t \) in the subinterval \( (t_{i-1}, t_i) \). Thus we obtain

\[
\delta_n^2 \geq \frac{1}{3} - \frac{1}{4} \sum_{i=1}^n (t_i - t_{i-1}) (2 - t_i - t_{i-1})^2,
\]
where an equality is attained only if \( 2d_i = 1 - t_i - t_{i-1} \). The expression in the right-hand-side of the last inequality for \( \delta_n^2 \) can be minimized without difficulties. Taking into account the restriction \( 0 = t_0 < t_1 < \ldots < t_n = 1 \) we obtain for the minimal variance:

\[
t_i = \frac{i}{n}, \quad i = 0, 1, 2, \ldots, n, \quad c_0 = c_n = \frac{1}{2n}, \quad c_i = \frac{1}{n}, \quad i = 1, 2, \ldots, n - 1
\]

and for these \( t_i \) and \( c_i \) we find

\[
\delta_n^2 = \frac{1}{12n^2}.
\]

Thus we are in a position to derive an important conclusion:

The quadrature formula based on these \( t_i, c_i \) corresponds to the well-known trapezoidal method for numerical integration. Hence we have proved the following result.

**Proposition 2** Among all the methods for approximate integration, based on quadrature formulas, the trapezoidal method has the smallest variance of the error over the space \( C[0, 1] \) endowed with the Wiener measure.

It is interesting to note that the Simpson method is optimal in many cases considered in the classical numerical analysis. Perhaps a little unexpectedly, when integrating the Wiener process (i.e. the integrand is from the space \( C[0, 1] \) endowed with the standard Wiener measure) and using the variance as a criterion, we find that the trapezoidal method is optimal. In this case the Simpson method leads to a variance of the error \( \delta_n^2 = 1/(9n^2) \). Obviously, these two methods have the same order of accuracy.

**Additional remark.** The above reasoning can be extended in different directions, e.g. we could cover a case like \( L(x) = \int_K x(t)d\varphi(t) \) where \( K \subseteq \mathbb{R}^n \) is a compact subset and \( x \) is a more general Gaussian multiparameter process whose (almost all) trajectories belong to the space \( C(K) \). Some possibilities are described in Kopanov (1994).

A different approach for an approximate integration of stochastic processes and several very interesting results can be found in Cambanis and Masry (1984, 1990).

**Acknowledgment.** The author’s thanks are addressed to Dr. J. Stoyanov for proposing these topics and for his permanent encouragement and help.
REFERENCES


