A NEW ALGORITHM OF THE MONTE CARLO METHOD FOR ESTIMATING THE MAXIMUM EIGENVALUE OF AN INTEGRAL OPERATOR

MATHEMATICS

1970

SovietRxiv

Abstract

Full Text

UDC 518:517.948

MATHEMATICS

T. A. MIKHAILOV

A NEW ALGORITHM OF THE MONTE CARLO METHOD FOR ESTIMATING THE MAXIMUM EIGENVALUE OF AN INTEGRAL OPERATOR

(Presented by Academician G. I. Marchuk, 16 IX 1969)

In the present paper we consider algorithms for estimating the eigenvalue of largest modulus of an integral operator K, based on carrying out, by the Monte Carlo method, iterations of the resolvent $[I-qK]^{-1}$. In some cases these algorithms can also be applied to estimating the least (in the algebraic sense) eigenvalue.

Let $k(x', x)$ be the kernel of the operator K, where $x', x \in X$. It is assumed that X is an n-dimensional Euclidean space and $K \in [L \to L]$, where L is a certain Banach space of summable functions. We shall also assume that the eigenvalues and eigenfunctions of the operators K and K^* possess properties that ensure convergence of the iterative process

$$\lambda^{(m)} = (K^m\psi, \varphi)/(K^{m-1}\psi, \varphi), \qquad \psi \in L, \quad \varphi \in L^*, \tag{1}$$

to the eigenvalue $\lambda > 0$ of largest modulus. On the nature of these properties and of the requirements imposed on the functions ψ and φ, see, for example, ([1,2]).

Let $|\lambda q| < 1$. It is not difficult to verify that the sets of eigenfunctions of the operators $[I-qK]^{-1}$ and K coincide and that the corresponding eigenvalues $\{\mu_k\}$ and $\{\lambda_k\}$ are related by $\mu_k = (1-q\lambda_k)^{-1}$. Therefore, in view of the assumptions made above about the spectral properties of the operator K and the remark about the functions φ and ψ, we shall assume that for $q > 0$

$$\mu^{(m)} = \frac{([I-qK]^{-m}\psi, \varphi)}{([I-qK]^{-(m-1)}\psi, \varphi)} \underset{m\to\infty}{\longrightarrow} \mu = \frac{1}{1-q\lambda}. \tag{2}$$

We now consider the known Monte Carlo algorithm ([1]) for carrying out the iterative process (1). Let $\{x_n\}$ be a homogeneous Markov chain with transition

density $r(x', x)$ from x' to x, with density $r_0(x)$ of the distribution of the initial state x_0, and with probability $p(x) \geq \delta > 0$ of terminating the trajectory upon transition from the point x; N is the random number of the last state immediately preceding the termination of the trajectory. Define the auxiliary random "weight" Q_n by the formulas:

$$Q_0 = \frac{\psi(x_0)}{r_0(x)}, \qquad Q_n = Q_{n-1}\frac{k(x_{n-1}, x_n)}{r(x_{n-1}, x_n)}\frac{1}{1 - p(x_{n-1})} \quad \text{for } n \leq N,$$

$$Q_n = 0 \quad \text{for } n > N.$$

It is known that, under fairly broad assumptions ([1,3]) concerning the characteristics of the Markov chain under consideration, the following relations hold:

$$MQ_n\varphi(x_n) = (K_\psi^n, \varphi), \tag{3}$$

$$M\sum_{n=0}^{N} Q_n q^n \varphi(x_n) = \sum_{n=0}^{\infty} q^n (K_\psi^n, \varphi) = ([I - qK]^{-1}\psi, \varphi). \tag{4}$$

Relation (3) also makes it possible to realize formula (1) (see (1)) approximately by simulating chain trajectories on a computer and computing the corresponding statistical estimates.

Expression (4) defines a Monte Carlo algorithm for estimating the quantity $\mu^{(1)}$ according to formula (2).

Below it will be shown how an analogous algorithm is constructed for estimating the quantities $\mu^{(m)}$. To optimize this algorithm the following will be useful.

Lemma 1. Let the functions $k(x', x)$ and $\varphi(x)$ be nonnegative,

$$\xi_x = \varphi(x) + \sum_{n=1}^{\infty} q^n Q_n \varphi(x_n), \qquad f^* = [I - qK^*]^{-1}\varphi.$$

If

$$r_0(x) = \frac{\psi(x) f^*(x)}{(\psi, f^*)}, \qquad r(x', x) = \frac{k(x', x) f^*(x)}{[K^* f^*](x')}, \qquad p(x) \equiv 0, \tag{5}$$

then $\xi_x = f^*(x)$ with probability 1 and, consequently, $D\xi_x = 0$ for a.e. x.

The assertion of the lemma follows directly from the results of [4].

We note that in the general case the relation $M\xi_x = f^*(x)$ holds (see, for example, (3)). We now pass to the description of the required algorithm.

Lemma 2. Let λ_1 be the eigenvalue of maximum modulus of the integral operator K_1, whose kernel is equal to the absolute value of the kernel of the operator K. If (3) is satisfied for every n and $|\lambda_1 q| < 1$, then

$$([I - qK]_{\psi}^{-m}, \varphi) = M \sum_{n=0}^{\infty} q^n a_n^{(m)} Q_n \varphi(x_n), \tag{6}$$

where $a_n^{(m)} = C_{m+n-1}^n$.

Proof. The relation

$$[I - qK]^{-m} = \sum_{n=0}^{\infty} q^n a_n^{(m)} K^n \tag{7}$$

is the expansion of an "operator binomial" with negative exponent. It is valid by virtue of the properties of the binomial series and the corresponding proposition of the spectral theory of functions of operators (see [5], Ch. 7, § 3, 10). Moreover, the series in (7) converges in the uniform topology of operators; therefore for each function $\psi \in L$ it converges in the corresponding norm.

Consequently,

$$([I - qK]_{\psi}^{-m}, \varphi) = \sum_{n=0}^{\infty} q^n a_n^{(m)} (K^n \psi, \varphi). \tag{8}$$

In order to pass from (6) to (8), it is necessary to average termwise (i.e., to integrate with respect to the probability measure in the space of trajectories) the series in (6). Such averaging will be admissible if, in (6), the functions $k(x', x)$, $\varphi(x)$, $\psi(x)$, and the quantity q are replaced by their absolute values. The sum of the series will then be finite in view of the condition $|q\lambda_1| < 1$. Thus, for the partial sums of the series (6) there exists an integrable majorant, and it can be averaged termwise by Lebesgue' s theorem (see [5], Ch. 3, § 6, 16). The lemma is proved.

The coefficients $a_n^{(m)}$ satisfy the relations:

$$a_n^{(m+1)} = a_n^{(m)} + a_{n-1}^{(m+1)}, \qquad a_n^{(m+1)} = \sum_{i=0}^{n} a_i^{(m)}. \tag{9}$$

Using (2), (6), and (9), for $q > 0$ we obtain the following estimate for λ:

$$\lambda \approx \frac{1}{q}\left(1 - \frac{1}{\mu^{(m)}}\right) = \frac{(K[I - qK]_{\psi}^{-m}, \varphi)}{([I - qK]_{\psi}^{-m}, \varphi)} = \frac{M \sum_{n=1}^{\infty} q^{n-1} a_{n-1}^{(m)} Q_n \varphi(x_n)}{M \sum_{n=0}^{\infty} q^n a_n^{(m)} Q_n \varphi(x_n)}.$$

The last equality in fact describes the required Monte Carlo algorithm. The coefficients $a_n^{(m)}$ in the course of the computations are easily calculated with the aid of the first of relations (9).

Let us denote by f_1 and f_1^* the first eigenfunctions of the operators K and K^*, respectively. It is not difficult to see that if $\psi = f_1$ or $\varphi = f_1^*$, then already $\mu^{(1)} = \mu$. But the statistical error of such an estimate, of course, will be different from zero. A way of reducing this error is indicated by

Lemma 3. Let the functions $\psi(x)$, $\varphi(x)$, and $k(x', x)$ be nonnegative, $q > 0$, $\varphi = f_1^*$, and

$$\xi^{(m)} = \sum_{n=0}^{\infty} a_n^{(m)} q^n \varphi(x_n) Q_n.$$

Then, if (5) is satisfied for $f^* = f_1^*$, the equality

$$\xi^{(m)} = \frac{1}{1-q\lambda}\xi^{(m-1)} = \frac{1}{(1-q\lambda)^m},$$

holds with probability 1, and consequently $D\xi^{(m)} = 0$.

Proof. Using (9) and rearranging the terms of the series, we obtain

$$\xi^{(m)} = \sum_{n=0}^{\infty} q^n a_n^{(m-1)} Q_n \left(\sum_{i=n}^{\infty} q^{i-n} \frac{Q_i}{Q_n} f_1^*(x_i) \right). \tag{10}$$

Since from $\lambda f_1^* = K f_1^*$ it follows that

$$\frac{1}{1-q\lambda} f_1^* = [I - qK^*]^{-1} f_1^*,$$

then, on the basis of Lemma 1, the last series in expression (10) is equal to $f_1^*(x_n)/(1-q\lambda)$ with probability 1. This, obviously, proves Lemma 3.

Let us note several positive properties of the proposed algorithm. Expression (6) uses the simulation results more fully than (3), in particular if the eigenvalue is calculated simultaneously with the calculation of the functional (4). In this case one should also take into account the circumstance that, with a corresponding choice of the value of the parameter q, the iterative process (2) may converge considerably faster than process (1). Further, a priori estimates of the function f_1^* on the basis of Lemma 3 can be used to reduce the variance of the statistical estimates.

One of the most important applications of the algorithms under consideration is the problem of determining the critical parameters of nuclear reactors (see [2]).

By the Monte Carlo method, such problems are usually solved by estimating the effective multiplication factor over generations of fissions. The method considered here makes it possible conveniently to calculate the multiplication factor over generations of particle collisions, which in some cases can substantially refine the estimate of the critical parameters.

Let us also note that expression (6), for $q < 0$, makes it possible to estimate by the Monte Carlo method the smallest (in the algebraic sense) eigenvalue, provided that the corresponding iterative process converges.

The algorithms and conclusions obtained automatically extend to the case when K is a linear algebraic operator.

To compare the efficiency of two Monte Carlo algorithms, the maximal eigenvalue λ of the integral operator of the Peierls equation (see (2)), describing isotropic scattering of particles in a plane layer of matter $0 \leq \tau \leq 3$, was computed. It was assumed that: $\psi(\tau) = 1$ for $0 \leq \tau \leq 1$, $\varphi \equiv 1$, $\tau(\tau', \tau) = \text{const} \cdot k(\tau', \tau)$, $q = 1$, and the probability of termination of a trajectory is equal to the probability of a particle leaving the layer, i.e., in fact the physical process of particle transport was simulated. It turned out that an estimate of the value $\lambda = 0.860\ldots$ (see (2), Appendix 1) with an error of 0.01 is achieved in process (1) for $m = 5$, and in process (2) for $m = 2$. The root-mean-square statistical errors σ of the corresponding "Monte Carlo" estimates are approximately in the ratio $7 : 2$.

In conclusion, we note that λ can be estimated analogously on the basis of iterations of certain other functions of the operator, for example $\exp(qK)$. However, an assertion of the type of Lemma 3 holds only in the case where the resolvent is used.

Computing Center
of the Siberian Branch of the USSR Academy of Sciences
Novosibirsk

Received
9 IX 1969

CITED LITERATURE

[1] V. S. Vladimirov, *Probability Theory and Its Applications*, **1**, No. 1, 113 (1956).
[2] G. I. Marchuk, *Numerical Methods for the Calculation of Nuclear Reactors*, Moscow, 1958.
[3] D. Kertiss, *UMN*, **12**, No. 5, 149 (1957).
[4] A. G. Mikhailov, *Journal of Computational Mathematics and Mathematical Physics*, **8**, No. 4, 1085 (1968).
[5] N. Dunford and J. T. Schwartz, *Linear Operators (General Theory)*, IL, 1962.

Note: Figure translations are in progress. See original paper for figures.