RELAXATION THEORY OF JACOBI METHODS

1967

SovietRxiv

Abstract

Full Text

UDC 518:512.86

MATHEMATICS

G. D. MAISTROVSKII

RELAXATION THEORY OF JACOBI METHODS

(Presented by Academician L. V. Kantorovich on 23 XII 1966)

In the literature devoted to computational methods, Jacobi methods for reducing an arbitrary symmetric matrix A to diagonal form by successive two-dimensional rotations are widely discussed (see the bibliography in ([1])). G. Forsythe and P. Henrici ([2]) raised the question of transferring to the Jacobi method results obtained for relaxation processes. The present note contains an attempt at a systematic consideration of Jacobi methods on the basis of the theory of relaxation for a quadratic functional, developed by Yu. I. Lyubich ([3]). We establish a number of general theorems which contain, as special cases, almost all known theorems on the convergence of Jacobi processes.

Let the sequence of matrices $A_k = (a_{ij}^{(k)})_{i,j=1}^n$ be defined by the recurrence formula ($k = 0, 1, 2, \ldots;\ A_0 = A$)

$$A_{k+1} = U_k A_k U_k^{-1}, \tag{1}$$

where U_k is the rotation matrix in the coordinate plane (i_k, j_k) through some angle ψ_k. Setting*

$$\operatorname{tg} 2\alpha_k = 2a_{i_k j_k}^{(k)} \left(a_{i_k i_k}^{(k)} - a_{j_k j_k}^{(k)}\right)^{-1} \qquad (-\pi/4 < \alpha_k \le \pi/4),$$

we shall assume that $0 \le \psi_k \alpha_k^{-1} \le 2$. Then (see, for example, ([1])) the recurrence (1) defines a relaxation process for the quadratic functional

$$\Delta(X) = \sum_{i<j} x_{ij}^2$$

in the space of symmetric matrices

$$X = (x_{ij})_{i,j=1}^n.$$

Introduce the **relaxation multipliers** q_k $(0 \le q_k \le 2)$ and the **relaxation angles** θ_k $(0 \le \theta_k \le \pi/2)$, putting** (cf. [3])

$$q_k = 1 - \sin(2\alpha_k - 2\psi_k)/\sin 2\alpha_k; \qquad \cos\theta_k = a^{(k)}_{i_k j_k k} \Big/ \left[\sum_{p<s} \left(a^{(k)}_{ps}\right)^2\right]^{1/2}. \tag{2}$$

Then

$$\Delta_{k+1} - \Delta_k = q_k(2 - q_k)\cos^2\theta_k\,\Delta_k. \tag{3}$$

Theorem 1. *For convergence of the Jacobi process it is necessary and sufficient that the series*

$$\sum_{k=0}^{\infty} q_k(2 - q_k)\cos^2\theta_k \tag{4}$$

diverge.

We shall call a Jacobi process **quasi-gradient** if $\cos\theta_k \ge \gamma > 0$, and we shall call a process **strictly relaxation** if $0 < q^0 \le q_k \le 2 - q^0$. Quasi-gradient processes, in particular, are processes with the choice of the controlling sequence (i_k, j_k) in accordance with the classical Jacobi method, and processes with a two-step choice of almost the greatest–

* We take $\alpha_k = 0$ if $a^{(k)}_{i_k j k} = a^{(k)}_{i_k i k} - a^{(k)}_{j_k j k} = 0$.

** We take $q_k = 1$ when $\alpha_k = 0$.

of an extra-diagonal element (4), cyclic processes with barriers decreasing no faster than a geometric progression (5). Strictly relaxation processes include, in particular, processes with complete relaxation $(q_k \equiv 1)$, Jacobi processes with rational formulas (1), and restricted Jacobi processes (6).

Corollary 1. *For convergence of the quasigradient Jacobi process it is necessary and sufficient that the series*

$$\sum_{k=0}^{\infty} q_k(2 - q_k)$$

diverge.

Corollary 2. *For convergence of the strictly relaxation Jacobi process it is necessary and sufficient that the series*

$$\sum_{k=0}^{\infty} \cos^2 \theta_k$$

diverge.

Theorem 2. *Put*

$$\sigma = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} q_k(2-q_k)\cos^2\theta_k; \qquad \chi = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \frac{\Delta_{k+1}}{\Delta_k}.$$

For any Jacobi process the equality holds

$$\chi = 1 - \sigma. \tag{5}$$

We shall say that a Jacobi process converges normally if $\chi < 1$.

From normal convergence it obviously follows that

$$\overline{\lim_{k\to\infty}} \sqrt[k]{\Delta_k} \leqslant \chi < 1.$$

Theorem 3. *A Jacobi process converges normally if and only if $\sigma > 0$.*

Corollary 3. *For normal convergence of the quasigradient Jacobi process it is necessary and sufficient that*

$$\underline{\lim_{n\to\infty}} \frac{1}{n} \sum_{k=0}^{n-1} q_k(2-q_k) > 0.$$

Corollary 4. *For normal convergence of a strictly relaxation Jacobi process it is necessary and sufficient that*

$$\underline{\lim_{n\to\infty}} \frac{1}{n} \sum_{k=0}^{n-1} \cos^2\theta_k > 0.$$

Theorems 1–3 and their corollaries are exact analogues of the corresponding results of Yu. I. Lyubich. To compare Theorem 2 with the theorem analogous to it, Theorem 4 from (3), it should be noted that the role of the measure of conditioning for our functional $\Delta(X)$ is played by the number 1.

We shall now need the following fact (4, 7, 9).

Lemma 1. *Let*

$$\varepsilon = \max_{i\neq j} |a_{ij}|$$

be sufficiently small, and let the numbering of the eigenvalues be chosen so that λ_i is close to a_{ii}. Then, if $\lambda_i = \lambda_j$,

$$|a_{ij}| \leqslant C\varepsilon^2,$$

where the constant C depends only on the spectrum of the matrix A.

Let us now consider quasigradient processes with complete relaxation. From Lemma 1 it follows that in the controlling sequence (i_k, j_k) of a quasigradient process, starting from some point, there occur no pairs (i,j) such that $\lambda_i = \lambda_j$. Consequently,

$$q_k = O(\varepsilon), \qquad a^{(k+1)}_{i_k s_k} - a^{(k)}_{i_k s_k} = O(\varepsilon^2).$$

Let in the sequence

$$(i_k, j_k)_{k=0}^{N} \qquad (N = n(n-1)/2)$$

some pair (i,j) occur for $k = k_1$ and $k = k_2$ $(k_1 < k_2 \leqslant N)$. It is easy to see that

$$a^{(k_1)}_{ij} = 0, \qquad |a^{(k_2-1)}_{ij}| = O(\varepsilon^2).$$

But, by virtue of quasigradientness,

$$\left(a^{(k_2-1)}_{ij}\right)^2 \geqslant \gamma^2 \Delta_{k_2-1}.$$

Hence it follows that

Theorem 4. *A quasigradient process with complete relaxation converges quadratically:*

$$\varlimsup_{k\to\infty} \Delta_{k+N}\Delta_k^{-2} \leqslant C\gamma^{-2}$$

(where the constant C depends only on the spectrum of the matrix, $\gamma = \inf\cos\theta_k$).

For some specific quasigradient processes this result was obtained in ([4,7–9]).

Let us order the superdiagonal elements of the matrix A_k in some linear manner $\{a^{(k)}_p\}_{p=1}^N$. We shall consider strictly relaxation processes satisfying the following **condition S**: in the control sequence $\{p_k\}_{k=0}^{\infty}$, every index from $\Omega^N = \{1, 2, \dots, N\}$ occurs arbitrarily far on. A special case of such processes is the **quasicyclic** one: for some natural number l (l is the length of the quasicycle), $\Omega^N \subset \{i_k\}_0^{pl-1}$ for every natural p.

We shall say that the Jacobi process **converges locally** if convergence takes place under the assumption that the off-diagonal elements of the matrix A are sufficiently small in comparison with the quantity $d = \min|\lambda_i - \lambda_j|$ $(i \neq j)$.

Theorem 5. *Let all eigenvalues of the matrix A be simple. Then a strictly relaxation process satisfying condition S converges locally with a rate guaranteed by the estimate*

$$\varepsilon_\lambda \leq |1 - q^0|\varepsilon_0 + C\varepsilon_0^2,$$

where

$$\varepsilon_k = \max_{i \neq j} |a_{ij}^{(k)}|; \quad \lambda = \min\{l : \Omega^N \subset \{i_k\}_0^{l-1}\}; \quad 0 < q^0 \leq q_k \leq 2 - q^0$$

and the constant C depends only on the spectrum of the matrix A.

This result is analogous to the theorem on the rate of convergence of coordinate relaxation for a quadratic functional, established in ([11]).

Corollary 5. *If the spectrum of the matrix A is simple, then a quasicyclic strictly relaxation process converges at least linearly.*

For the special cyclic process ("by columns") this result was obtained in ([8]).

Corollary 6. *If the spectrum of the matrix A is simple, then a quasicyclic process with complete relaxation converges locally quadratically.*

The last result is contained in the works ([6,8,10]), but in the corresponding estimates ($\varepsilon_l \leq C(A,l)\varepsilon_0^2$) the quantity $C(A,l)_{l\to\infty} \to \infty$. In fact, as is seen from Theorem 5, the quantity C does not depend on the length l of the quasicycle.

In conclusion we note that the questions remain open of the nonlocal convergence of processes satisfying condition S, and of the local convergence of these processes when the matrix A has multiple eigenvalues.*

Physical-Technical Institute of Low Temperatures
Academy of Sciences of the Ukrainian SSR

Received
13 XII 1966

CITED LITERATURE

1. D. K. Faddeev, V. N. Faddeeva, *Computational Methods of Linear Algebra*, Moscow, 1963.

2. G. E. Forsythe, P. Henrici, Trans. Am. Math. Soc., **94**, No. 1, 1 (1960).

3. Yu. I. Lyubich, DAN, **161**, No. 6, 1274 (1965).

4. V. V. Voevodin, *Numerical Methods of Algebra*, Moscow, 1966.

5. D. A. Pope, C. Tompkins, J. Assoc. Comp. Machinery, **4**, No. 4, 459 (1957).

6. P. Henrici, J. Soc. Ind. and Appl. Math., **6**, No. 2, 144 (1958).

7. A. Schönhage, Numer. Math., **6**, 5, 440 (1964).

8. G. Schröder, Forsch. Land. Nordrhein–Westf., 1291 (1964).

9. E. D. Levinson, Zhurn. vychisl. matem. i matem. fiz., **6**, No. 3, 556 (1966).

10. A. Schönhage, Numer. Math., **3**, 374 (1961).

11. Yu. I. Lyubich, DAN, **173**, No. 1, 37 (1967).

* For one special cyclic process the latter question was answered affirmatively in ([9]), under the condition that the matrix A has only one multiple eigenvalue.

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