Ukrainian Mathematical Journal, Vol. 49, No. 12, 1997

EXACT ESTIMATES FOR THE RATE OF CONVERGENCE OF THE s-STEP METHOD OF STEEPEST DESCENT IN EIGENVALUE PROBLEMS

P. F. Zhuk and L. N. Bondarenko

UDC 519.6

We obtain exact (unimprovable) estimates for the rate of convergence of the s-step method of steepest descent for finding the least (greatest) eigenvalue of a linear bounded self-adjoint operator in a Hilbert space.

The investigation of the rate of convergence of an s-step method of steepest descent in the solution of linear operator equations was began by Kantorovich [1] and, for eigenvalue problems, by Birman [2]. In further works (see, e.g., [3–6]), the results obtained in [1, 2] were generalized and improved. In particular, exact (unimprovable) estimates were obtained for the rate of convergence of the s-step method of steepest descent in the solution of linear operator equations. The present work is devoted to establishing similar estimates in the problem of finding the least eigenvalue of a linear operator.

Let $A: H \to H$ be a linear bounded self-adjoint operator acting in the real Hilbert space H with the scalar product (u, v). For the spectrum of the operator A, we assume that $\mathrm{sp}(A) \subseteq \{m\} \cup [m^*, M]$, $m < m^* < M$. In this case, m is an eigenvalue of the operator A associated with a certain proper subspace $H^{(1)}$.

To find the eigenvalue m and the corresponding eigenvector, we use the s-step method of steepest descent whose successive approximations are constructed according to the rule

$$u_{k+1} = \sum_{i=0}^{s} \alpha_i^{(k)} A^i u_k, \qquad k = 0, 1, \ldots, \tag{1}$$

where u_0 is an arbitrary unit vector and the coefficients $\alpha_i^{(k)}$ are such that $\|u_{k+1}\| = 1$ and the Rayleigh ratio

$$\mu(u_{k+1}) = \frac{(Au_{k+1}, u_{k+1})}{\|u_{k+1}\|^2}$$

is minimum. We assume that $\mu_k = \mu(u_k)$, $k = 0, 1, \ldots$.

Remark 1. Consider the generalized eigenvalue problem

$$Ku = \lambda L u, \tag{2}$$

where K and L are linear self-adjoint operators, L is a positive-definite operator, and $A = L^{-1}K$ is the operator bounded in the energy space H_L. Problem (2) is reduced to the eigenvalue problem

$$Au = \lambda u \tag{3}$$

in the space H_L. Since A is a self-adjoint operator in H_L, the s-step method of steepest descent can be used for the solution of problem (3) and, hence, problem (2).

Krupskaya Kherson Pedagogic Institute, Kherson. Translated from Ukrainskii Matematicheskii Zhurnal, Vol. 49, No. 12, pp. 1694–1699, December, 1997. Original article submitted February 19, 1996.

 0041–5995/97/4912–1912 $18.00

Remark 2. If $\{(\mu_k, u_k),\ k = 0, 1, \dots\}$ is a sequence of pairs generated by the s-step method of steepest descent applied to the operator A, then $\{\chi_1 + \chi_2(\mu_k, u_k),\ k = 0, 1, \dots\}$ is the sequence of pairs of the s-step method of steepest descent applied to the operator $\chi_1 E + \chi_2 A$ (E is the unit operator and χ_1 and χ_2 are arbitrary real numbers).

It follows from Remark 2 that it is sufficient to study the s-step method of steepest descent only for an operator of the form $\chi_1 E + \chi_2 A$ $(\chi_2 \neq 0)$. As this operator (denote it by A), we choose an arbitrary operator of the indicated form with the numbers $\chi_1 + \chi_2 M = 1$, $\chi_1 + \chi_2 m > 0$, and $\chi_2 > 0$, i.e., in what follows, we assume that A is a self-adjoint operator with the boundaries $m > 0$ and $M = 1$.

We say that the s-step method of steepest descent with the initial approximation u_0 becomes stable if, for certain $k \in \{0, 1, \dots\}$ the equality $w_k = Au_k - \mu_k u_k = 0$ is valid, i.e., an eigenpair of the operator A is determined by finitely many iterations.

Denote by $\mathfrak{T}$ a subset of the unit sphere Ω of the space H consisting of elements v for which the system of vectors $v, Av, \dots, A^{s+1}v$ is linearly dependent. Assume that $\mathfrak{U} = \Omega \setminus \mathfrak{T}$.

In [7], the following condition of stabilization of the s-step method of steepest descent was proved: if $u_0 \in \mathfrak{T}$, then $w_1 = 0$; otherwise $u_k \in \mathfrak{U}$, $k = 1, 2, \dots$. Thus, the s-step method of steepest descent becomes stable if and only if $u_0 \in \mathfrak{T}$.

Since the case where $u_0 \in \mathfrak{T}$ is trivial, in what follows, we assume that $u_0 \in \mathfrak{U}$.

We see from the definition of the s-step method of steepest descent that $m \le \mu_{k+1} \le \mu_k$, $k = 0, 1, \dots$ and, hence, $\{\mu_k,\ k = 0, 1, \dots\}$ is a bounded sequence. Let $u_0^{(1)}$ be the orthogonal projection of the vector u_0 onto $H^{(1)}$. If $u_0^{(1)} = 0$, then, evidently, $\lim\limits_{k \to 0} \mu_k \ge m^*$ and, consequently, for finding m, it is necessary that $u^{(1)} \neq 0$. Under the condition $u_0^{(1)} \neq 0$, we have $\mu_k \to m$ and $k \to \infty$. Therefore, without loss of generality, we can assume in what follows that $\mu_0 < m^*$.

Denote by E_t the spectral function of the operator A and by $\sigma_k = \sigma_k(t) = (E_t u_k, u_k)$ the distribution function of the vector u_k. By definition, the function σ_k is defined and nondecreasing on the entire number axis, continuous from the left on the segment $]-\infty; 1[$, and $\sigma_k(t) = 0$ for $t \le m$ and $\sigma_k(t) = 1$ for $t \ge 1$. In addition, $\sigma_k(t) = \sigma_k(m+0)$ for $m < t \le m^*$.

Denote by Σ_k the set of points of growth of the function σ_k belonging to the segment $[m^*, 1]$. It follows from [7] that $\Sigma_{k+1} \subseteq \Sigma_k$, $k = 0, 1, \dots$ and any set Σ_k contains at least $s + 1$ points (because $u_0 \in \mathfrak{U}$). Assume that

$$\lambda_* = \min \Sigma_0, \qquad \lambda^* = \max \Sigma_0.$$

It follows from [7] that $\lambda_* = \min \Sigma_k$ and $\lambda^* = \max \Sigma_k$, $k = 0, 1, \dots$. Denote by $\pi_s(t, u_0)$ the polynomial of degree s in t with the smallest deviation from zero on the set Σ_0 normalized by the condition $\pi_s(m, u_0) = 1$. Assume that

$$\rho_s(u_0) = \max_{t \in \Sigma_0} |\pi_s(t, u_0)|.$$

We decompose the vector u_0 into orthogonal components

$$u_0 = u_0^{(1)} + u_0^{(2)}, \qquad u_0^{(1)} \in H^{(1)}, \qquad u_0^{(2)} \perp H^{(1)} \tag{4}$$

and construct the vector $\tilde{u}_0$ by the rule

$$\tilde{u}_0 = u_0^{(1)} + \rho_s(u_0) u_0^{(2)}. \tag{5}$$

Theorem 1. *The following estimates hold:*

$$\frac{\mu_1 - m}{\mu_0 - m} \le \left[\frac{\rho_s(u_0)}{\|\tilde{u}_0\|}\right]^2 \le \rho_s^2(u_0)\frac{\lambda_* - \mu_1}{\lambda_* - \mu_0}. \tag{6}$$

Proof. Assume that $\tilde{\mu}_0 = \mu(\tilde{u}_0)$, $\tilde{u} = \pi_s(A, u_0)u_0$, and $\tilde{\mu} = \mu(\tilde{u})$. It follows from decomposition (4) and equality (5) that

$$\tilde{u} = u_0^{(1)} + \tilde{u}^{(2)}, \quad \tilde{u}^{(2)} = \pi(A)u_0^{(2)}, \quad \tilde{u}^{(2)} \perp H^{(1)},$$

$$\tilde{\mu}_0 = \frac{mh + \rho^2(Au_0^{(2)}, u_0^{(2)})}{h + \rho^2\|u_0^{(2)}\|^2}, \tag{7}$$

$$\tilde{\mu} = \frac{mh + (A\tilde{u}^{(2)}, \tilde{u}^{(2)})}{h + \|\tilde{u}^{(2)}\|^2},$$

where, for brevity, $\pi(t) = \pi_s(t, u_0)$, $h = \|u_0^{(1)}\|^2$, and $\rho = \rho_s(u_0)$. In the integral form, we have

$$\tilde{\mu}_0 = \frac{mh + \rho^2\int_{\lambda_*}^{\lambda^*} t\,d\sigma_0(t)}{h + \rho^2\int_{\lambda_*}^{\lambda^*} d\sigma_0(t)},$$

$$\tilde{\mu} = \frac{mh + \int_{\lambda_*}^{\lambda^*} t\pi^2(t)\,d\sigma_0(t)}{h + \int_{\lambda_*}^{\lambda^*} \pi^2(t)\,d\sigma_0(t)}, \tag{8}$$

where the integrals are understood in the Riemann–Stieltjes sense.

Let us prove that

$$\mu_1 \le \tilde{\mu} \le \tilde{\mu}_0 \le \mu_0. \tag{9}$$

Indeed, by using Eqs. (4) and (7), we have

$$\mu_0 - \tilde{\mu}_0 = (1-\rho^2)h\Big((Au_0^{(2)}, u_0^{(2)}) - m\|u_0^{(2)}\|\Big)\|\tilde{u}_0\|^{-2} \ge 0,$$

because

$$0 \le \rho < 1, \quad \frac{(Au_0^{(2)}, u_0^{(2)})}{\|u_0^{(2)}\|^2} \ge \lambda_* > m.$$

To prove the inequality $\tilde{\mu} \le \tilde{\mu}_0$, we approximate the integrals in formulas (8) by integral sums. More exactly, we show that, for any sufficiently fine partition of the segment $[\lambda_*, \lambda^*]$, the following inequality holds:

$$\frac{mh + \sum_{i=1}^{n} t_i^{(n)} \pi^2(t_i^{(n)}) h_i^{(n)}}{h + \sum_{i=1}^{n} \pi^2(t_i^{(n)}) h_i^{(n)}} \le \frac{mh + \rho^2 \sum_{i=1}^{n} t_i^{(n)} h_i^{(n)}}{h + \rho^2 \sum_{i=1}^{n} h_i^{(n)}}, \tag{10}$$

where $h_i^{(n)}$ is the Stieltjes measure of the ith half interval of the partition and $t_i^{(n)}$ is an intermediate point of the ith half interval of the partition (if $h_i^{(n)} \ne 0$, then $t_i^{(n)} \in \Sigma_0$). We define the function

$$f(\zeta_1, \dots, \zeta_n) = \frac{mh + \sum_{i=1}^{n} t_i^{(n)} \zeta_i}{h + \sum_{i=1}^{n} \zeta_i}.$$

Assume that the considered partition of the segment $[\lambda_*, \lambda^*]$ is so fine that $f(\rho^2 h_1^{(1)}, \dots, \rho^2 h_n^{(n)}) < \lambda_*$ (this is possible because $\tilde{\mu}_0 \le \mu_0 < \lambda_*$). An elementary analysis shows that the function $f(\zeta_1, \dots, \zeta_n)$ monotonically increases on the set

$$[0, \rho^2 h_1^{(n)}] \times \dots \times [0, \rho^2 h_n^{(n)}]$$

with respect to each variable. This yields

$$f(\pi^2(t_1^{(n)}) h_1^{(n)}, \dots, \pi^2(t_n^{(n)}) h_n^{(n)}) \le f(\rho^2 h_1^{(n)}, \dots, \rho^2 h_n^{(n)}),$$

which was to be proved (note that if $h_i^{(n)} \ne 0$, then $t_i^{(n)} \in \Sigma_0$ and, hence, $\pi^2(t_i^{(n)}) \le \rho^2$).

By passing to the limit in inequality (10), we obtain $\tilde{\mu} \le \tilde{\mu}_0$.

To prove the estimate $\mu_1 \le \tilde{\mu}$, it is sufficient to note that $\tilde{u} \in \operatorname{span}(u_0, Au_0, \dots, A^s u_0)$ and, hence, $\mu(u_1) \le \mu(\tilde{u})$.

Thus, we have proved inequalities (9). Further, it follows from the estimate $\mu_1 \le \tilde{\mu}_0$ and the relation for $\tilde{\mu}_0$ in (7) that

$$\frac{\mu_1 - m}{\mu_0 - m} \le \frac{\tilde{\mu}_0 - m}{\mu_0 - m} \le \left[\frac{\rho}{\|\tilde{u}_0\|}\right]^2.$$

The left-hand side of inequality (6) is proved. To prove the right-hand side of this inequality, we set $\lambda = \mu(u_0^{(2)})$. Since $\|u_0\| = 1$, it follows from equalities (7) that

$$h + \|u_0^{(2)}\|^2 = 1, \qquad mh + \lambda\|u_0^{(2)}\|^2 = \mu_0,$$
$$\frac{mh + \rho^2\lambda\|u_0^{(2)}\|^2}{h + \rho^2\|u_0^{(2)}\|^2} = \tilde{\mu}_0. \tag{11}$$

By solving Eqs. (11) for h, $\|u_0^{(2)}\|^2$, and ρ^2, we obtain

$$\|\tilde{u}_0\|^2 = h + \rho^2\|u_0^{(2)}\|^2 = \frac{\lambda - \mu_0}{\lambda - \tilde{\mu}_0}.$$

Since $\lambda \geq \lambda_*$, it follows from estimates (9) that

$$\|\tilde{u}_0\|^2 \geq \frac{\lambda_* - \mu_0}{\lambda_* - \mu_1}.$$

Theorem 1 is proved.

It follows from Theorem 1 that

$$\frac{\mu_1 - m}{\lambda_* - \mu_1} \leq \rho_s^2(u_0)\frac{\mu_0 - m}{\lambda_* - \mu_0}. \tag{12}$$

Since

$$\min \Sigma_k = \lambda_*, \quad \max \Sigma_k = \lambda^*, \quad \Sigma_{k+1} \subseteq \Sigma_k, \quad k = 0, 1, \ldots,$$

by successive use of estimates (12) for the vectors $u_0, u_1, \ldots,$ we obtain

$$\frac{\mu_k - m}{\lambda_* - \mu_k} \leq \rho_s^2(u_0)\frac{\mu_{k-1} - m}{\lambda_* - \mu_{k-1}},$$

$$\frac{\mu_k - m}{\lambda_* - \mu_k} \leq \rho_s^{2k}(u_0)\frac{\mu_0 - m}{\lambda_* - \mu_0}, \quad k = 1, 2, \ldots,$$

i.e., the s-step method of steepest descent converges at least with the rate of a geometric progression with ratio $\rho_s^2(u_0)$.

Note that

$$\| u_k - e \| \leq 2\left[\frac{\mu_k - m}{\lambda_* - m}\right]^{0,5}, \quad k = 0, 1, \ldots,$$

where

$$e = \frac{u_0^{(1)}}{\|u_0^{(1)}\|} \in H^{(1)}:$$

Estimate (12) is exact (unimprovable) in the following sense: if $\rho < \sup\limits_{u_0} \rho_s(u_0)$, then

$$\frac{\mu_1 - m}{\lambda_* - \mu_1} > \rho^2 \frac{\mu_0 - m}{\lambda_* - \mu_0}$$

for a certain initial approximation u_0 with $\mu_0 < m^*$. To prove this assertion, it is sufficient to consider initial approximations in a neighborhood of the proper subspace $H^{(1)}$.

If we impose an additional restriction on the initial approximation assuming, e.g., that the value μ_1 is fixed, then estimate (12) can, generally speaking, be improved. We prove that the estimate

$$\frac{\mu_1 - m}{\mu_0 - m} \le \left[\frac{\rho_s(u_0)}{\|\tilde{u}_0\|}\right]^2,$$

which follows from inequality (6), remains unimprovable.

Theorem 2. *Let H be a finite-dimensional space. If $\dim H \ge s+2$, then, for any number μ $(m < \mu < m^*)$, there exists an initial approximation u_0 (depending on μ) such that $\mu_1 = \mu$ and*

$$\frac{\mu_1 - m}{\mu_0 - m} = \left[\frac{\rho_s(u_0)}{\|\tilde{u}_0\|}\right]^2.$$

Proof. Let $\lambda_1, \dots, \lambda_n$ $(\lambda_1 = m,\ \lambda_2 = m^*,\ \lambda_n = 1)$ be eigenvalues of the operator A. Denote by $\pi_s(t)$ the polynomial of degree s of the least deviation from zero on the set $\{\lambda_2, \dots, \lambda_n\}$ normalized by the condition $\pi_s(\lambda_1) = 1$, and denote the value of deviation by

$$\rho_s = \max_{i=2,\dots,n} |\pi_s(\lambda_i)|.$$

It follows from the definition of the polynomial $\pi_s(t)$ that there exist numbers $2 < i_1 < \dots < i_{s-1} < n$ such that

$$(-1)^j \pi_s(\lambda_{i_j}) = \pi_s(\lambda_2) = (-1)^s \pi_s(\lambda_n) = \rho_s \quad j = 1, \dots, s-1. \tag{13}$$

We set $q(t) = (t-\mu)\pi_s(t)$, $J = \{1, 2, i_1, \dots, i_{s-1}, n\}$. Since $m < \mu < m^*$, the sequence $q(\lambda_1), q(\lambda_2), q(\lambda_{i_1}), \dots, q(\lambda_n)$ has exactly $s+1$ changes in sign. Therefore, the system of equations linear with respect to ζ_j^2, $j \in J$,

$$\sum_{j\in J} \zeta_j^2 = 1, \qquad \sum_{j\in J} \lambda_j^i q(\lambda_j)\zeta_j^2 = 0, \qquad i = 0, 1, \dots, s,$$

has a real solution $\zeta_j^* \ne 0$, $j \in J$. Indeed, otherwise, by virtue of the Stiemke theorem [8], there exists a polynomial $l(t) = \tau_0 + \tau_1 t + \dots + \tau_s t^s \not\equiv 0$ with real coefficients such that $q(\lambda_j) l(\lambda_j) \ge 0$, $j \in J$. But in this case, the polynomial $l(t)$ has $s+1$ real roots (with regard to their multiplicity) and, hence, $l(t) \equiv 0$. Thus, we arrive at a contradiction.

Let e_j be an arbitrary unit eigenvector of the operator A corresponding to the eigenvalue λ_j. Let us show that the vector

$$u_0 = \sum_{j\in J} \zeta_j^* e_j$$

satisfies the condition of the theorem.

Consider the polynomial

$$q_1(t) = (t-\mu_1) p_0(t), \qquad p_0(t) = \sum_{i=0}^{s} \alpha_i^{(0)} t^i.$$

Since $(Au_1 - \mu_1 u_1, A^i u_0) = 0$, $i = 0, 1, \dots, s$, we have

$$\sum_{j \in J} \lambda_j^i q_1(\lambda_j)(\zeta_j^*)^2 = 0, \quad i = 0, 1, \dots, s,$$

and, hence, for an arbitrary number α, we obtain

$$\sum_{j \in J} \lambda_j^i [q_1(\lambda_j) - \alpha q(\lambda_j)](\zeta_j^*)^2 = 0, \quad i = 0, 1, \dots, s. \tag{14}$$

Let α^* be such that the degree of the polynomial $q_1(t) - \alpha^* q(t)$ is at most s. It follows from Eqs. (14) that $q_1(t) = \alpha^* q(t)$. This yields

$$\mu_1 = \mu, \quad p_0(t) = \alpha^* \pi_s(t), \quad u_1 = \frac{\tilde{u}}{\|\tilde{u}\|} \quad (\tilde{u} = \pi_s(A) u_0).$$

Since

$$\Sigma_0 = \{\lambda_2, \lambda_{i_1}, \dots, \lambda_{i_{s-1}}, \lambda_n\},$$

we conclude that $\pi_s(t, u_0) = \pi_s(t)$ and, by virtue of (13), $\rho_s(u_0) = \rho_s$. Thus, we get

$$\mu_1 - m = \mu(\tilde{u}) - m = \left[\frac{\rho_s(u_0)}{\|\tilde{u}\|}\right]^2 (\mu_0 - m).$$

It remains to note that $\|\tilde{u}\| = \|\pi_s(A) u_0\| = \|\tilde{u}_0\|$. The theorem is proved.

REFERENCES

1. L. V. Kantorovich, "Functional analysis and applied mathematics," *Usp. Mat. Nauk*, **3**, No. 6, 89–184 (1948).
2. M. Sh. Birman, "Calculation of eigenvalues by the method of steepest descent," *Zap. Leningrad. Gorn. Inst.*, **27**, No. 1, 209–215 (1952).
3. A. B. Kovrigin, "Estimation of the rate of convergence of the k-step gradient method," *Vestn. Leningrad. Univ.*, Issue 13, 34–36 (1970).
4. V. G. Prikazchikov, "Strict estimates of the rate of convergence of the iteration method for the calculation of eigenvalues," *Zh. Vychisl. Mat. Mat. Fiz.*, **15**, No. 5, 1330–1333 (1975).
5. A. V. Knyazev, *Calculation of Eigenvalues and Eigenvectors in Network Problems: Algorithms and Estimates of Error* [in Russian], Department of Computational Mathematics, Academy of Sciences of the USSR, Moscow (1986).
6. A. F. Zabolotskaya, "On a method of steepest descent," in: *Communications on Applied Mathematics* [in Russian], Computer Center, Academy of Sciences of the USSR, Moscow (1988), p. 26.
7. P. F. Zhuk, "Asymptotic behavior of the s-step method of steepest descent in eigenvalue problems in a Hilbert space," *Mat. Sb.*, **184**, No. 12, 87–122 (1993).
8. V. S. Charin, *Linear Transformations and Convex Sets* [in Russian], Vyshcha Shkola, Kiev (1978).