SHORT COMMUNICATIONS

THE RATE OF CONVERGENCE OF THE METHOD OF STEEPEST DESCENT IN A EUCLIDEAN NORM

A.V. KNYAZEV AND A.L. SKOROKHODOV

It is proved that the error norm in the method of steepest descent decreases at the rate of a geometric progression. An estimate is given for the rate of convergence of its decrease, that is the best possible at each step of the method, for the class of ill-posed problems. The proof is based on the idea of reduction to the case of "smallest dimension" (in the given case, to two dimensions).

Consider the method of steepest descent for solving the system of linear algebraic equations $A u = f$ with a symmetric positive-definite matrix $A$ in the space $\mathbb{R}^n$ with a natural scalar product $(\cdot, \cdot)$ and norm $\| \cdot \| = (\cdot, \cdot)^{1/2}$. The formulae of this method have the form (see /1, 2/)

$$ u^{n+1} = u^n - \gamma_n (A u^n - f) \quad \gamma_n = \lambda^{-1}(A u^n - f). \quad n = 0, 1, \ldots. \quad (1) $$

where $\lambda(\cdot) = (A^*, \cdot)/(\cdot, \cdot)$ is the Rayleigh relation and $u^0$ is some initial vector.

The estimate of the rate of convergence of method (1) (that cannot be improved, see /2/) is well-known

$$ (A e^*, e^*)^m \leq \left( \frac{\delta - 1}{\delta + 1} \right)^m (A e^*, e^*)^0, $$

where $e^0 = e^i = e, \delta = \text{cond} (A) = M/m$, and the numbers $m$ and $M$ are the minimum and maximum eigenvalues of the matrix $A$ respectively. From this it is easy to obtain an estimate of the rate of convergence of method (1) in a Euclidean norm /1, 2/:

$$ \| e^i \| \leq \delta \left( \frac{\delta - 1}{\delta + 1} \right)^m \| e^0 \|. $$

**Theorem.** If $n = 0, 1, \ldots$ then the inequality

$$ e^i \leq \left( 1 - \delta^{-1} \right)^m \| e^0 \| $$

(2)

holds. Estimate (2) is the best possible at each step of method (1) for the class of ill-posed problems in the sense that for every $n = 0, 1, \ldots$ a family of errors $E^n$ can be found, such that

$$ \sup e^i \leq \left( 1 - \delta^{-1} \right)^m \| e^0 \| \leq 1. $$

(3)

**Proof.** Since $\delta$ does not depend on $n$, it is sufficient to prove estimate (2) and assertion (3) for the case $n = 0$.

We define the subspace $H^i = \text{span}(e^0, A e^0)$. If $\text{dim} H^1 = 1$ then $A e^i = \text{const} e^i$ and $e^i = 0$, i.e., method (1) converges in one iteration. Therefore, the case $\text{dim} H^1 = 2, H^i = \text{span}(e^i, e^{i-1})$ is non-trivial.

In $H^i$ we choose vectors $u_i$ and $u_i$ (Ritz vectors /3/) from the conditions

$$(u_i, u_i) = (A u_i, u_i) = 1, \quad \| u_i \| = \| u_i \| = 1.$$  

Let $\lambda_i = \lambda(u_i)$ and $\lambda_i = \lambda(u_i)$ and suppose that $\lambda_i, \lambda_i \leq \lambda_i, \lambda_i$. Then by virtue of the Courant-Fisher principal /3/, $m \leq \lambda_i, \lambda_i \leq M_i$.

We shall establish that the inequality

$$ \| e^i \| \leq \delta \left( 1 - \delta^{-1} \right)^m \| e^0 \| $$

(4)

is stronger than (2) (for $n = 0$), is correct, where $\delta = \lambda_1 / \lambda_2$, $1 \leq \delta \leq 1$.

We write the vector $e^i$ in the form of an expansion: $e^i = u_i + \mu_i$. Then $A e^i = \lambda_{i, i} u_i + \lambda_{i, 0} e_i$, and thus

$$ (A e^i, A e^i) = \lambda_{i, i} u_i^2 + \lambda_{i, 1} u_i e_i, \quad (A e^i, A e^i) = \lambda_{i, 0} u_i^2 + \lambda_{i, 0} e_i^2 $$

and thus

$$ \| e^i \| \leq \lambda_1 u_i^2 + \lambda_2 e_i^2 $$

$$ \| e^i \| \leq \| e^0 \| \left( 1 - \delta^{-1} \right)^m \| e^0 \| $$

$$ \| e^i \| \leq \left( 1 - \delta^{-1} \right)^m \| e^0 \|. $$

---

*Zh. vychisl. Mat. mat. Fiz., 28, 10, 1577-1578, 1988*
Taking into account the relation $\varepsilon^r = \varepsilon^0 - \gamma_0 A \varepsilon^0$ we note that (4) corresponds to the numerical inequality

$$\frac{(\delta - 1)^2 \alpha \beta^3 (\delta^3 \beta^3 + \alpha^3)}{(\alpha^3 + \delta^3 \beta^3) (\alpha^3 + \beta^3)} \leq (1 - \delta^{-1})^2,$$

which can be directly verified.

Note that inequality (5) is in fact an estimate of the rate of convergence of one iteration of method (1) in a Euclidean norm with matrix $A = \text{diag}(\lambda_1, \lambda_2)$ and error $\varepsilon^0 = (\alpha, \beta)^T$. Estimate (2) is proved.

We shall prove assertion (3). We chose a family of initial approximations in such a way that the family of errors $E^0$ is a linear covering of the eigenvectors of the matrix $A$ corresponding to the eigenvectors $m$ and $M$. Then $\bar{\delta} = \delta$ and, taking into account that

$$(1 - \delta^{-1})^{-1} \frac{\|e^r\|}{\|e^0\|} = \frac{\delta \alpha \beta (\delta^3 \beta^3 + \alpha^3) h}{(\alpha^3 + \delta^3 \beta^3) (\alpha^3 + \beta^3) h},$$

we have

$$\lim_{\delta \to 0} \lim_{\beta \to 0} \left[ (1 - \delta^{-1})^{-1} \frac{\|e^r\|}{\|e^0\|} \right] = 1.$$

The theorem is proved.

A similar method of proof is proposed in [4] where, in particular, analogous results are obtained for spectral problems.

REFERENCES