TWO-LEVEL ITERATIVE SCHEMES

MATHEMATICS

1969

Abstract

Full Text

UDC 518:517.944/.947

MATHEMATICS

Corresponding Member of the Academy of Sciences of the USSR A. A. SAMARSKII

TWO-LEVEL ITERATIVE SCHEMES

In this note we consider two-level (one-step, two-term, first-order schemes [1]) iterative schemes for solving the equation $Au = f$, where A is a linear self-adjoint operator given in a Hilbert space (h.s.). For the study of implicit schemes, the method [2] of investigating stability with respect to initial data for two-level difference schemes is used. An implicit scheme of general form is reduced to an explicit scheme (the simple-iteration scheme). The basic problem in the theory of iterative schemes is treated as the problem of estimating the norm of the resolving operator of the equivalent explicit scheme and of choosing the iteration parameters from the condition that this norm be minimal. In the case of "stationary" schemes, the problem of the minimum of the norm of the transition operator of the explicit scheme is solved. The study of the computational stability of iterative schemes is reduced to the problem of stability of schemes with respect to perturbations of all input data—the operators of the scheme, the right-hand side, and the initial approximation. In constructing schemes, methods [3] are used. A class of schemes with a factorized operator $B = (E+\omega R_1)(E+\omega R_2)$ is considered, where R_1 and R_2 are adjoint ("triangular") operators.

1. Let H be a real h.s.; $(,)$ the scalar product in H; $(H \to H)$ the set of linear operators acting from H into H. All operators considered below belong to $(H \to H)$. We use the notation [2]: $A > 0$ if $(Ax, x) > 0$ for all $x \in H$, $\|x\| \neq 0$; $A \geq B$ if $(Ax, x) \geq (Bx, x)$ for all $x \in H$, etc. Along with H we consider the energy spaces H_D, where $D = D^* > 0$, with norm $\|x\|_D = \sqrt{(Dx, x)}$, $D \in (H \to H)$, $D = A$ or B.

Let the operator equation be given

$$Au = f, \tag{1}$$

where u is the sought vector and f a given vector from H. For an approximate solution of equation (1), consider one-step iterative schemes, which we write in canonical form

$$B_k(y_{k+1} - y_k)/\tau_{k+1} + Ay_k = f, \quad k = 0, 1, \dots, \quad y_0 \in H \text{ is an arbitrary vector.} \tag{2}$$

Here k is the iteration number, y_k the k-th iterate, $\tau_{k+1} > 0$ a parameter, and B_k an arbitrary operator having an inverse B_k^{-1}. In what follows it is assumed everywhere that the initial approximation y_0 is an arbitrary vector from H. Since the solution u of equation (1) satisfies (2), for the error $z_k = y_k - u$ we have:

$$B_k(z_{k+1} - z_k)/\tau_{k+1} + Az_k = 0, \quad k = 0, 1, \dots, \quad z_0 = y_0 - u. \tag{3}$$

The iterative scheme (2) coincides in form with the two-level scheme [2] corresponding to the abstract Cauchy problem $B\,du/dt + Au = f$, $0 \le t \le t_0$, $u(0) = u_0$. It is therefore natural, instead of saying "one-step iterative scheme" (method), to say "two-level iterative scheme." If $B_k = E$, the identity operator, then (2) is called an explicit scheme; if $B_k \ne E$, then (2) is an implicit scheme. We define the family of schemes (2) by the conditions

$$A = A^* > 0, \quad B_k = B_k^* > 0. \tag{4}$$

According to (2), the implicit scheme (3) is equivalent to the explicit scheme $x_{k+1} = S_k x_k$, $k = 0, 1, \dots$, where $x_k = A^{1/2} z_k$, $S_{k+1} = E - \tau_{k+1} C_k$, $C_k = A^{1/2} B_k^{-1} A^{1/2}$, $S_{k+1}-$

transition operator (from the k-th iteration to the $(k+1)$-st iteration) for the explicit scheme. Hence it follows that

$$x_n = T_{nx} 0, \quad T_n = S_{nS_{n-1}} \dots S_1, \quad \|x_n\| \le \|T_n\| \|x_0\|, \quad \|x_k\| = \|z_k\|_A, \tag{5}$$

where T_n is the resolving operator of the explicit scheme equivalent to scheme (3). In view of (4), $T_n = T_n^*$ and $\|T_n\| = \sup_{\|x\|=1} |(T_{nx}, x)|$.

Lemma 1. *Let $\|T_n\| \le q_n$. Then for (2) the estimate*

$$\|y_n - u\|_A \le q_n \|y_0 - u\|_A, \qquad n = 1, 2, \dots, \tag{6}$$

is valid, and if A and B commute, then

$$\|Ay_n - f\| \le q_n \|Ay_0 - f\|. \tag{7}$$

The iterations according to scheme (2) converge in H_A if $q_n \to 0$ as $n \to \infty$. Usually it is required to find an approximate solution of problem (1) with relative error $\varepsilon > 0$, or with absolute error ε_0: $(\|y_n - u\|_A / \|y_0 - u\|_A) < \varepsilon$, or $\|y_n - u\|_A < \varepsilon_0$ under the condition $\|u\|_A \approx 1$. The first requirement is fulfilled when $\|T_n\| < \varepsilon$ or $q_n \le \varepsilon$. The number of iterations $n = n_0(\varepsilon)$ for which $\|T_n\| < \varepsilon$ depends on the choice of B_k and τ_k, which are subject to the requirement $\inf \|T_n\|$ (or $\min q_n$).

2. In what follows, schemes (2) with constant operators $B_k = B$ are considered throughout, under the assumption that

$$A = A^* > 0, \qquad B = B^* > 0, \qquad \gamma_1 B \le A \le \gamma_2 B, \qquad \gamma_2 > \gamma_1 > 0, \quad (8)$$

where γ_1 and γ_2 are prescribed constants (A and B are energetically equivalent [3], convergent [4], spectrally equivalent [5] operators with constant equivalence constants γ_1 and γ_2). For $B_k = B$, (3) is equivalent to the explicit scheme

$$x_{k+1} = S_{k+1} x_k, \qquad S_{k+1} = E - \tau_{k+1} C, \qquad k = 0, 1, \dots, \quad (9)$$

$$C = C^{(1)} = A^{1/2} B^{-1} A^{1/2} \quad \text{for } x_k = A^{1/2} z_k,$$

or

$$C = C^{(2)} = B^{-1/2} A B^{-1/2} \quad \text{for } x_k = B^{1/2} z_k.$$

Hence, and from (8), it follows that

$$C = C^* > 0, \qquad \gamma_1 E \le C \le \gamma_2 E. \quad (10)$$

Introduce the notation

$$\tau_0 = 2/(\gamma_1 + \gamma_2), \qquad \rho_0 = (\gamma_2 - \gamma_1)/(\gamma_2 + \gamma_1) = (1 - \xi)/(1 + \xi),$$

$$\rho_1 = (1 - \sqrt{\xi})/(1 + \sqrt{\xi}), \qquad \xi = \gamma_1/\gamma_2. \quad (11)$$

For scheme (9), the resolving operator $T_n = T_n(C)$ is an operator polynomial of degree n, whose norm satisfies $\|T_n(C)\| \le \max_{t \in [\gamma_1, \gamma_2]} |T_n(t)|$ (see [6]).

The parameters $\tau_1, \tau_2, \dots, \tau_n$ are found from the condition $\inf \|T_n(C)\|$, which reduces to the problem of finding $\min \max_{t \in [\gamma_1, \gamma_2]} |T_n(t)|$; its solution has the form [1]

$$\tau_k = \tau_0/(1 + \rho_0 \lambda_k), \qquad \lambda_k = \cos[(2k-1)\pi/2n], \qquad k = 1, 2, \dots, n, \quad (12)$$

and in this case

$$\|T_n\| \le q_n, \qquad q_n = 2\rho_1^n/(1 + \rho_1^{2n}).$$

Hence, and from Lemma 1 with $B_k = B$, for (2) it follows that

$$\|y_k - u\|_D \le q_n \|y_0 - u\|_D, \qquad \text{where } D = A \text{ or } D = B. \tag{13}$$

3. Let us turn to "stationary" schemes with constant $B_k = B$ and $\tau_k = \tau$:

$$B(y_{k+1} - y_k)/\tau + Ay_k = f, \qquad k = 0, 1, \dots, \qquad y_0 \in H \text{ given}. \tag{14}$$

For $z_k = y_k - u$ we obtain the homogeneous equation (3_0) with $B_k = B$ and $\tau_{k+1} = \tau$. It is equivalent to the explicit scheme with constant transition operator $x_{k+1} = Sx_k$, $S = E - \tau C$, $k = 0, 1, 2, \dots$, where $C = C^{(1)}$ for $x_k = A^{1/2} z_k$, or $C = C^{(2)}$ for $x_k = B^{1/2} z_k$. In this case $T_n = S^n$, $\|T_n\| = \|S\|^n$, since $S = S^*$, and the problem of $\inf_\tau \|T_n\|$ reduces to the problem of finding $\inf_\tau \|S\|$. We shall use the well-known result [1] for the explicit scheme:

Theorem 1. *Let $S = E - \tau C$ and let conditions (10) be satisfied. Then $\|S\| < 1$ for $0 < \tau < 2/\gamma_2$, and $\inf_\tau \|S\|$ is attained at $\tau = \tau_0$:*

$$\inf_\tau \|S\| = \|E - \tau_0 C\| = \rho_0.$$

Corollary. For the scheme (14), (8), with $\tau = \tau_0$, the estimates

$$\|y_n - u\|_D \le \rho_0^n \|y_0 - u\|_D, \qquad D = A \text{ or } D = B. \tag{15}$$

Remark 1. In ([2]) it is shown that the necessary and sufficient condition for the estimate $\|z_n\|_D \le \rho^n \|z_0\|_D$, $\rho < 1$ ($D = A$ or $D = B$) for the scheme (14) has the form

$$\frac{1}{\tau}(1 - \rho)B \le A \le \frac{1}{\tau}(1 + \rho)B.$$

Comparing it with (8), we obtain $\tau = \tau_0$, $\rho = \rho_0$.

Remark 2. For finite-dimensional H, the estimates (15) were obtained in ([4],[7]) by the method of separation of variables. The estimates (15) are inconvenient for checking, since they contain the unknown exact solution u of equation (1).

Theorem 2. For problem (14), under (8) and with $\tau = \tau_0$, the estimates

$$\|y_{n+1} - y_n\|_D \le \rho_0^n \|y_1 - y_0\|_D, \qquad D = A \text{ or } B,$$

$$\|Ay_n - f\|_{B^{-1}} \le \rho_0^n \|Ay_0 - f\|_{B^{-1}}. \tag{16}$$

If $AB = BA$, then the estimate (7) is valid with $q_n = \rho_0^n$.

The inequality

$$\|y_{n+1} - y_n\|_D \le \varepsilon \|y_1 - y_0\|_D$$

may serve as a condition for terminating the iterations. The number of iterations in this case is $n > n_0(\varepsilon)$, where

$$n_0(\varepsilon) = [\ln(1/\varepsilon)/\ln(1/\rho_0)]$$

($[a]$ is the integer part of the number a). Without loss of generality, one may take $\|u\|_A = 1$. In this case the criterion for terminating the iterations may be the inequality

$$\|y_n - u\|_D \le \varepsilon_0.$$

Lemma 2. For the scheme (14), under (8), the inequalities

$$\frac{1}{\tau\gamma_2}\|y_{n+1} - y_n\|_D \le \|y_n - u\|_D \le \frac{1}{\tau\gamma_1}\|y_{n+1} - y_n\|_D \quad (D = A, B)$$

hold.

It follows from this that the condition $\|y_n - u\|_D \le \varepsilon_0$ will be fulfilled if

$$\|y_{n+1} - y_n\|_D \le \varepsilon_1,$$

where $\varepsilon_1 = \tau\gamma_1\varepsilon_0$.

4. Let us dwell on the question of computational stability of the scheme (14). If the computations are carried out with a finite number of digits, then because of rounding errors, in solving problem (14) we find its approximate solution $\tilde{y}_n$, which may be regarded as the exact solution of a problem with perturbed input data

$$\widetilde{B}(\tilde{y}_{k+1} - \tilde{y}_k)/\tau_0 + \widetilde{A}\tilde{y}_k = \tilde{f}_k, \qquad k = 0, 1, \dots, \qquad \tilde{y}_0 \in H. \tag{17}$$

We shall assume that $\widetilde{A} = \widetilde{A}^* > 0$, $\widetilde{B} = \widetilde{B}^* > 0$, i.e. (17) belongs to the original family of schemes (14). As a measure of the perturbation of the operators A and B, it is natural to take the relative change in the energies of these operators:

$$|((\widetilde{A}-A)x, x)| \le \alpha(Ax, x), \qquad |((\widetilde{B}-B)x, x)| \le \alpha(Bx, x) \qquad (0 \le \alpha < 1) \tag{18}$$

for all $x \in H$, which is equivalent to the conditions

$$\|\widetilde{A}^{-1/2}(\widetilde{A} - A)\widetilde{A}^{-1/2}\| \le \alpha, \qquad \|\widetilde{B}^{1/2}(\widetilde{B} - B)\widetilde{B}^{-1/2}\| \le \alpha.$$

Theorem 3. Let $\tilde{y}_n$ be the solution of problem (17), and let conditions (8), (18) be fulfilled. Then, for

$$\alpha \le \xi/(4 + \xi), \qquad \xi = \gamma_1/\gamma_2,$$

the estimates are valid

$$\|\tilde{y}_n - u\|_D \le \tilde{\rho}^n\|y_0 - u\|_D + \alpha\sqrt{\frac{1+\alpha}{1-\alpha}}\,(1+\tilde{\rho}^n)\|f\|_{D^{-1}} + \frac{2(1-\tilde{\rho}^n)}{\gamma_1}\max_{0\le k\le n-1}\|\tilde{f}_k - f\|_{D^{-1}},$$

where

$$\tilde{\rho} = \rho_0 + \frac{2\alpha}{1-\alpha}(1+\rho_0) \le \frac{1+\rho_0}{2} < 1 \quad \text{for} \quad \alpha < \frac{\xi}{4+\xi}, \qquad D = A \text{ or } B.$$

5. Not only the number of actions Q for computing one iteration, but also the number of iterations $n_0(\varepsilon)$, depends on the choice of the operator B in the scheme (14). Therefore it is natural to choose B in some admissible family of operators so that: 1) the ratio $\xi = \gamma_1/\gamma_2$ is maximal (ρ_0 minimal), 2) B is an economical operator (Q minimal in some sense, for example, in order relative to ξ as $\xi \to 0$). In constructing B, one usually proceeds from some operator $R = R^* > 0$ (see ([1–7])), energo-

-tically equivalent to A and B:

$$c_1 R \le A \le c_2 R, \qquad c_2 \ge c_1 > 0; \tag{19}$$

$$\mathring{\gamma}_1 B \le R \le \mathring{\gamma}_2 B. \tag{20}$$

Then inequalities (8) hold with constants $\gamma_1 = c_1\mathring{\gamma}_1,\ \gamma_2 = c_2\mathring{\gamma}_2$.

Let us investigate the case of a factorized operator (f.o.)

$$B = (E + \omega R_1)(E + \omega R_2), \qquad R = R_1 + R_2, \tag{21}$$

where R_1 and R_2 are adjoint operators, $R_2 = R_1^*$, so that $B = B^* > 0$, and $\omega > 0$ is a parameter. The numbers $\mathring{\gamma}_1$ and $\mathring{\gamma}_2$ in this case depend on the parameter ω, which should be chosen so that the ratio $\mathring{\gamma}_1/\mathring{\gamma}_2 = f(\omega)$ is maximal ([8]).

Theorem 4. Let $R_2 = R_1^*$, $R = R_1 + R_2$, and $R \ge \delta E$, $\|R_2 x\|^2 \le \frac{1}{4}\Delta(Rx, x)$, $\Delta \ge \delta > 0$. Then the constants $\mathring{\gamma}_1$ and $\mathring{\gamma}_2$ in (20) are given by the formulas

$$\mathring{\gamma}_1 = \delta/2(1+\sqrt{\eta}), \qquad \mathring{\gamma}_2 = \delta/4\sqrt{\eta}, \qquad \mathring{\gamma}_1/\mathring{\gamma}_2 = 2\sqrt{\eta}/(1+\sqrt{\eta}),$$

$$\eta = \delta/\Delta \quad \text{for} \quad \omega = 2/\sqrt{\delta\Delta}\,^*.$$

Remark 1. Knowing $\mathring{\gamma}_1, \mathring{\gamma}_2$, we find $\gamma_1 = c_1\mathring{\gamma}_1$, $\gamma_2 = c_2\mathring{\gamma}_2$ and $\tau_0 = 2/(\gamma_1+\gamma_2)$. For the number of iterations $n_0(\varepsilon)$, as $\eta \to 0$ the estimate

$$n_0(\varepsilon) = O\left(\frac{1}{\sqrt{\eta}} \ln \frac{1}{\varepsilon}\right)$$

is valid.

Remark 2. If one uses scheme (2) with $B_k = B$, where B is the f.o. (21), and τ_{k+1} are determined according to (12), then

$$n_0(\varepsilon) = O\left(\frac{1}{\sqrt[4]{\eta}} \ln \frac{1}{\varepsilon}\right).$$

The question of the computational stability of such a scheme remains open.

Let us consider one more f.o.

$$B = (E + \omega_1 R_1)(E + \omega_2 R_2), \qquad \text{where } R_1 R_2 = R_2 R_1,\ R_\alpha = R_\alpha^* > 0,$$

$$\alpha = 1, 2, \quad R_1 + R_2 = R \tag{22}$$

Theorem 5. Let the conditions $\delta_\alpha E \le R_\alpha \le \Delta_\alpha E;\ \Delta_\alpha > 0;\ \alpha = 1, 2;\ \delta_2 > 0;\ \delta_1[1 + \delta_2(1/\Delta_1 + 1/\Delta_2)] + \delta_2 > 0$ be satisfied. Then, for the values of the parameters ω_1 and ω_2 equal to

$$\omega_1 = \omega_0/(1 - \omega_0 c_0) > 0, \qquad \omega_2 = \omega_0/(1 + \omega_0 c_0) > 0, \qquad \omega_0 = 1/\sqrt{\delta_1' \Delta_1'} =$$

$$= 1/\sqrt{\delta_2' \Delta_2'}, \quad \delta_1' = \delta_1 + c_0, \quad \delta_2' = \delta_2 - c_0, \quad \Delta_1' = \Delta_1 + c_0, \quad \Delta_2' = \Delta_2 - c_0,$$

$$c_0 = (\delta_2 \Delta_2 - \delta_1 \Delta_1)/(\delta_1 + \delta_2 + \Delta_1 + \Delta_2),$$

the equivalence constants of $R = R_1 + R_2$ and the f.o. (22) are equal to

$$\mathring{\gamma}_1 = \frac{1 - \bar{\rho}}{\omega_1 + \omega_2}, \qquad \mathring{\gamma}_2 = \frac{1 + \bar{\rho}}{\omega_1 + \omega_2}, \qquad \frac{\mathring{\gamma}_1}{\mathring{\gamma}_2} = \frac{1 - \bar{\rho}}{1 + \bar{\rho}}, \qquad \bar{\rho} = \frac{1 - \sqrt{\eta_1'}}{1 + \sqrt{\eta_1'}} \frac{1 - \sqrt{\eta_2'}}{1 + \sqrt{\eta_2'}},$$

$$\eta_\alpha' = \frac{\delta_\alpha'}{\Delta_\alpha'}, \qquad \alpha = 1, 2.$$

Schemes with an f.o. of the form (22) occur in solving difference elliptic equations by the method of alternating directions (see, for example, ([8–10])).

Received
2 X 1968

CITED LITERATURE

[1] D. K. Faddeev, V. N. Faddeeva, *Computational Methods of Linear Algebra*, Moscow, 1960.
[2] A. A. Samarskii, DAN, **181**, No. 4, 808 (1968).
[3] A. A. Samarskii, Zhurn. vychisl. matem. i matem. fiz., **7**, No. 1, 62 (1967); **7**, No. 5, 1093 (1967).
[4] E. G. D' yakonov, Zhurn. vychisl. matem. i matem. fiz., **6**, No. 1, 12 (1966); DAN, **163**, No. 6, 1314 (1965).
[5] S. G. Mikhlin, *Numerical Implementation of Variational Methods*, "Nauka," Moscow, 1966.
[6] L. V. Kantorovich, G. P. Akilov, *Functional Analysis in Normed Spaces*, Moscow, 1959.
[7] J. E. Gunn, J. Soc. Industr. and Appl. Math., Numer. Anal., Sec. B, **2**, No. 1, 24 (1965).
[8] A. A. Samarskii, Zhurn. vychisl. matem. i matem. fiz., **6**, No. 4, 665 (1966).
[9] E. L. Wachspress, J. Soc. Industr. and Appl. Math., **11**, No. 3, 994 (1963).
[10] R. S. Varga, *Information Processing*, Paris, 1960, p. 85–90.

* In the proof of Theorem 4 the inequality $B \geq 2\omega R$, indicated by E. Nikolaev, was used.

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