

ON THE QUESTION OF OPTIMAL ITERATIVE PROCESSES

MATHEMATICS

1968

SovietRxiv

View the original and related papers at <https://sovietrxiv.org/items/ru-196801.48108>

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.

Abstract

Full Text

UDC 518:512.25

MATHEMATICS

Corresponding Member of the Academy of Sciences of the USSR G. I. Marchuk,
Yu. A. Kuznetsov

ON THE QUESTION OF OPTIMAL ITERATIVE PROCESSES

We shall denote by V_n the space of n -dimensional real vectors over the field of real numbers with scalar product

$$(\varphi, \psi) = \sum_{i=1}^n \varphi_i \psi_i \quad (\varphi, \psi \in V_n).$$

Let Q be some subspace of V_n . A matrix B^* is called positive definite on Q ($B \underset{Q}{>} 0$) if $(B\varphi, \varphi) > 0$ for any $\varphi \in Q$ ($\varphi \neq \Theta$, Θ is the zero vector), and positive semidefinite on Q ($B \underset{Q}{\geq} 0$) if $(B\varphi, \varphi) \geq 0$ for all $\varphi \in Q$. The spectral radius of the matrix B will be denoted by $\rho(B)$. In the present work all matrices are assumed to be real.

Let, for the solution of the system of linear algebraic equations

$$A\varphi = f, \quad (1)$$

where A is a nonsingular matrix and $f \in V_n$, the iterative process

$$\varphi^{k+1} = \varphi^k - B \left[\sum_{i=1}^p \gamma_{ik} (AB)^{i-1} \right] (A\varphi^k - f), \quad \varphi^k \in U^0 \quad (k = 0, 1, \dots). \quad (2)$$

is proposed.

Here B is some matrix for which $B(A\varphi - f) \neq \Theta$ for any $\varphi \in U^0$ ($\varphi \neq A^{-1}f$), $\gamma_{ik} = \gamma_i(\xi^k)$ ($\xi^k = A\varphi^k - f$), and the γ_i belong to a set G of real functionals defined on any $\xi \in V_n$ ($\xi \neq \Theta$, $i = 1, \dots, p$). With respect to U^0 we shall assume that this is a closed set of the space V_n such that $U_A = \{\psi : \psi = A\varphi - f, \varphi \in U^0\}$ is a linear subspace of V_n and, for any $\varphi \in U^0$ and any real numbers α_i ($i = 1, \dots, p$), the vector

$$\psi = \varphi - B \left[\sum_{i=1}^p \alpha_i (AB)^{i-1} \right] (A\varphi - f)$$

belongs to U^0 .

By convergence of the iterative process (2) in U^0 we shall mean convergence of the sequence $\{\varphi^k\}$ to the vector $\varphi^* = A^{-1}f$, which is the unique solution of system (1), for any initial approximation

$$R_p(\gamma) = I - B \left[\sum_{i=1}^p \gamma_i (AB)^{i-1} \right] A$$

$\varphi^0 \in U^0$. The operator ($\gamma_i \in G$; $i = 1, \dots, p$, I is the identity matrix) is called the transition operator of the iterative process (2). We introduce in U_A the norm $\|\varphi\|_D = (D\varphi, \varphi)^{1/2}$, where $D = D^* \underset{U_A}{>} 0$ is some matrix, and in the space

$U = \{\psi : \psi = \varphi - \varphi^*, \varphi \in U^0\}$ the norm $\|\psi\|_{DA} = \|A\psi\|_D$.

For given matrices A , B , and D , the problem of optimizing the iterative process (2) in the space U with the norm $\|\cdot\|_{DA}$ consists in finding functionals $\gamma_i \in G$ ($i = 1, \dots, p$) such that

* Here and below, any matrix whose dimensions are not specified is assumed to be square of order n .

$$\left\| R_p(\bar{\gamma}) \right\|_{DA}^U = \inf_{\substack{\gamma_i \in G \\ 1 \leq i \leq p}} \|R_p(\gamma)\|_{DA}^U, \quad (3)$$

where

$$\|R_p(\gamma)\|_{DA}^U = \sup_{\psi \in U} (\|R_p(\gamma)\psi\|_{DA} / \|\psi\|_{DA}).$$

It is not difficult to see that the problem of constructing such functionals is equivalent to finding, for any $\xi \in U_A$, a vector $\bar{\gamma} = \gamma(\xi) \in V_p$ for which

$$J_D(\bar{\gamma}, \xi) = \inf_{\gamma \in V_p} J_D(\gamma, \xi). \quad (4)$$

Here

$$J_D(\gamma, \xi) = \left\| \left(I - \sum_{i=1}^p \gamma_i S^i \right) \xi \right\|_D$$

and $S = AB$. Problem (4) is solvable for any $\xi \in U_A$ ($\xi \neq \Theta$), and the set of its solutions coincides with the set of solutions of the system

$$C^{(p)}\gamma = b^{(p)}, \quad (5)$$

where $C^{(m)}$ is an $m \times m$ matrix with elements

$$c_{ij} = (DS^i\xi, S^j\xi) \quad (1 \leq i, j \leq m)$$

and $b^{(m)} \in V_m$ is the vector with components

$$b_i = (DS^i\xi, \xi) \quad (i = 1, \dots, m; m = 1, \dots, p),$$

which is always consistent.

Since the matrix $C^{(p)}$ may be singular, problem (3) has, in general, a nonunique solution. Let Q_m be the set of vectors $\xi \in U_A$ for which the vectors $\{S^i\xi\}_{i=1}^m$ are linearly independent and

$$S^{m+1}\xi = \sum_{i=1}^m a_{iS}^i \xi \quad (m = 1, \dots, p-1)$$

and

$$Q_p = \{\xi : \xi \notin Q_m \ (m = 1, \dots, p-1), \xi \in U_A, \xi \neq \Theta\}.$$

Then the vector $\bar{\gamma}$ with components

$$\bar{\gamma}_i = \sum_{j=1}^p \tilde{c}_{ij} b_j^{(p)} \quad (i = 1, \dots, p),$$

where \tilde{c}_{ij} are the elements of the $p \times p$ matrix \tilde{C} , defined by the relation

$$\tilde{C} = \begin{vmatrix} [C^{(m)}]^{-1} & 0 \\ 0 & 0 \end{vmatrix},$$

if $\xi \in Q_m$ ($1 \leq m \leq p$), is a solution of system (5).

Thus, defining the values of the functionals γ_i^0 by the relations

$$\gamma_i^0(\xi) = \bar{\gamma}_i(\xi)$$

for all $\xi \in U_A$ ($\xi \neq \Theta$) and

$$\gamma_i^0(\xi) = 0$$

for all $\xi \notin U_A$ and $\xi = \Theta$ ($i = 1, \dots, p$), we obtain a solution of problem (3). Process (2) with functionals γ_i^0 ($i = 1, \dots, p$) constructed in this way will be called optimal.

It is not difficult to show that $J_D(\gamma, \xi)$ is a continuous functional on the set

$$Q_A = \{\psi : \psi = \varphi / \|\varphi\|, \varphi \in U_A\}$$

and that $J_D(\gamma, \xi) < 1$ ($\xi \in U_A, \xi \neq \Theta$) if and only if

$$\max_{1 \leq i \leq p} |(DS^i \xi, \xi)| > 0.$$

Then, since

$$\|R_p(\gamma)\|_{DA}^U = \inf_{\psi \in Q_A} J_D(\gamma, \psi),$$

the following is valid.

Theorem 1. Let

$$\alpha_D(s) = \inf_{Q_A} \left[\max_{1 \leq i \leq p} |(DS^i \psi, \psi)| \right].$$

Then $\alpha_D(s) < 0$ is a necessary and sufficient condition for convergence in U^0 of the optimal iterative process (2); moreover, if $\alpha_D(S) > 0$, then

$$\|R_p(\gamma)\|_{DA}^U < 1.$$

Concerning the implementation of the optimal iterative process (2), let us note that if, for some $k \geq 0$, the vector $\xi^k \in Q_p$, then the matrix $C^{(p)}$ is nonsingular and $\gamma(\xi^k) = [C^{(p)}]^{-1} b^{(p)}$, and if $\xi^k \notin Q_p$, ($\xi^k \in U_A, \xi^k \neq \Theta$) and there is some solution of system (5), then

$$\varphi^* = \varphi_k - B \left[\sum_{i=1}^p \tilde{\gamma}_i S^{i-1} \right] \xi^k,$$

and the iterative process terminates, converging to the exact solution in a finite number of steps.

Since for $p_2 > p_1 \geq 1$ we have

$$\|R_{p_2}(\overset{0}{\gamma})\|_{D_A}^U \leq \|R_{p_1}(\overset{0}{\gamma})\|_{D_A}^U,$$

it follows that

Lemma 1. *If the p_1 -step optimal process (2) converges in U^0 , then the p_2 -step process also converges for any $p_2 > p_1$.*

Let $U_A \perp \Phi_D$, where Φ_D is the null space of the matrix D (this, for example, always holds if D is nonsingular). Then $\xi = [D^+]^{1/2} D^{1/2} \xi$ and

$$(DS^i \xi, \xi) = (\Lambda^i \psi, \psi) \quad (i = 1, \dots, p)$$

for any $\xi \in U_A$. Here $[D^+]^{1/2}$ is a pseudoinverse matrix for the matrix $D^{1/2}$ (see (2)), $\psi = D^{1/2} \xi$, and $\Lambda = D^{1/2} A B [D^+]^{1/2}$. Since in this case

$$\omega_D(S) = \inf_{\psi \in U_D} \left[\max_{1 \leq i \leq p} |(\Lambda^i \psi, \psi)| / (\psi, \psi) \right],$$

where $U_D = \{\psi : \psi = D^{1/2} \varphi, \varphi \in U_A\}$, we have

Lemma 2. $\Lambda^{i_0} > 0$ for some $1 \leq i_0 \leq p$ is a sufficient condition for convergence in U^0 of the optimal iterative process (2).

Let us consider some special cases.

1. Let $U_A \perp \Phi_D$ and $\Lambda = \Lambda^*$. Then for any $p \geq 1$ the condition $U_D \perp \Phi_\Lambda = 0$, where Φ_Λ is the null space of the matrix Λ , is necessary for convergence in U^0 of the optimal iterative process (2). This condition is also sufficient for its convergence in U^0 for any $p \geq 2$ in the case of an arbitrary Hermitian matrix Λ , and for any $p \geq 1$ when, in addition,

$$\Lambda > 0.$$

If $\Lambda > 0$ and $U_D \perp \Phi_A$, then for the norm of the operator $R_p(\overset{0}{\gamma})$ in U the estimate

$$\|R_p(\overset{0}{\gamma})\|_{D_A}^U \leq 1/|T_p(\tau_0)|, \tag{6}$$

is valid, where

$$T_p(\tau) = [(\tau + \sqrt{\tau^2 - 1})^p + (\tau - \sqrt{\tau^2 - 1})^p] / 2$$

is the Chebyshev polynomial of order p , $\tau_0 = (M + m)/(M - m)$, $M = \rho(\Lambda)$, and $m = 1/\rho(\Lambda^+)$. For the case of an arbitrary Hermitian matrix Λ , $U_D \perp \Phi_A$, and $p \geq 2$, the estimate has the form

$$\|R_p(\gamma)\|_{D_A}^U \leq 1/|T_{\bar{p}}(\bar{\tau}_0)|.$$

Here $\bar{p} = [p/2]$ and $\bar{\tau}_0 = (\bar{M} + \bar{m})/(\bar{M} - \bar{m})$, where $\bar{M} = \rho(\Lambda^2)$ and $\bar{m} = 1/\rho([\Lambda^+]^2)$. If $A = A^* > 0$, $B = I$, and $D = A^{-1}$, then we arrive at the p -step method of steepest descent proposed by L. V. Kantorovich (7), and estimate (6) is given, for example, in monograph (1).

Under the assumptions made at the beginning of this section, for the implementation of one step of the optimal iterative process (2) one can propose the following modification of the conjugate-gradient method (see (1, 6)). Let, for some $k \geq 0$, vectors φ^k and $\xi^k = A\varphi^k - f$ be given. Then we obtain the vector φ^{k+1} by the following formulas

$$\varphi^{k+i/p} = \varphi^{k+(i-1)/p} - \gamma_i r_{i-1}^k, \quad \gamma_i = \frac{(D(A\varphi^{k+(i-1)/p} - f), Ar_{i-1}^k)}{(DAr_{i-1}^k, Ar_{i-1}^k)},$$

$$(i = 1, \dots, p),$$

where

$$r_0^k = B\xi^k, \quad r_i^k = B(A\varphi^{k+i/p} - f) - \beta_i r_{i-1}^k,$$

$$\beta_i = \frac{(DAB(A\varphi^{k+i/p} - f), Ar_{i-1}^k)}{(DAr_{i-1}^k, Ar_{i-1}^k)}, \quad (i = 1, \dots, p-1).$$

The latter is not difficult to show, since

$$(DAr_i^k, Ar_j^k) = 0 \quad \text{for } i \neq j$$

$$(Ar_i^k, Ar_j^k \in U_A, \quad 1 \leq i, j \leq p).$$

The implementation scheme proposed above ...

makes it possible to substantially increase the efficiency of the iterative processes under consideration when they are applied in practice.

Example. Let $A = \widetilde{B} - C$, $A \succ 0$, and $U_A = \{\psi : \psi = C\varphi, \varphi \in V_n\}$, where

$$\widetilde{B} = \begin{vmatrix} F_1 & 0 \\ -F_2 & F_1^* \end{vmatrix}_{V_n} > 0, \quad C = \begin{vmatrix} L & L \\ L & L \end{vmatrix} = C^* \succeq 0,$$

and $F_1, F_2 = F_2^*$ and $L = L^*$ are $n/2 \times n/2$ matrices. (Systems of equations with such matrices arise in the numerical solution of kinetic equations [4].) Choose for process (2) $B = \widetilde{B}^{-1}$ and $D = C^+$. Since $U_A \perp \Phi_D$ and $\rho(BC) = \|BC\|_{D_A}^U < 1$,

it is easy to see that $\Lambda = D^{1/2}AB[D^+]^{1/2} = \Lambda^* \underset{U_n}{>} 0$, and, consequently, the optimal iterative process (2) converges in

$$U^0 = \{\varphi : \varphi = A^{-1}\psi + J, \psi \in U_A\}$$

for any $p \geq 1$. Note that for any $f \in V_n$ the vector $Bf \in U^0$, since $ABf - f = -CBf \in U_A$. Using (6), it is not difficult to show that

$$\lim_{\rho(BC) \rightarrow 1} \ln \rho(BC) / \ln \|R_p(\gamma^0)\|_{DA}^U \leq 1/2p^2$$

for any fixed $p \geq 1$. In other words, in this case asymptotically one step of the optimal iterative process (2) is at least equivalent to $2p^2$ steps of the simple iteration method

$$\widetilde{B}\varphi^{k+1} = C\varphi^k + f, \quad \varphi^k \in V_n \quad (k = 0, 1, \dots).$$

Let us note that, to implement the iterative process (2), knowledge of the matrix C^+ is not necessary if the vector Bf is chosen as the initial one.

2. Let $B_\tau = \prod_{i=1}^l (I + \tau B_i)$ be a nonsingular matrix, τ a real parameter, $B = B_\tau^{-1}$, and $D = I$. Iterative processes with matrices B of this form arise when solving system (1) by the splitting method (see [4, 5]).

Lemma 3. If $A \underset{V_n}{>} 0$, then there exists $\tau_0 > 0$ such that for any $\tau \in [0, \tau_0]$ the matrix B_τ is nonsingular and the optimal iterative process (2) converges in V_n for all $p \geq 1$.

Suppose that $l = 2$, $A = B_1 + B_2$, $B_1 \underset{V_n}{>} 0$; $B_2 \underset{V_n}{>} 0$, and choose

$$D = [(I + \tau B_1)^{-1}]^*(I + \tau B_1)^{-1}.$$

Lemma 4. Under the assumptions made above, the optimal iterative process (2) converges in V_n for any $\tau > 0$ and $p \geq 1$.

In conclusion, the authors express their gratitude to S. G. Godunov and V. I. Lebedev for valuable comments.

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

Received 27 V 1968

REFERENCES

1. D. K. Faddeev, V. N. Faddeeva, *Computational Methods of Linear Algebra*, Moscow, 1963.
2. F. R. Gantmacher, *The Theory of Matrices*, Moscow, 1966.
3. G. I. Marchuk, N. N. Yanenko, Proceedings of the Symposium on Applied and Computational Mathematics, Novosibirsk, 1965.
4. G. I. Marchuk, *Uspekhi Mat. Nauk*, 161, No. 1, 66 (1965).
5. A. A. Samarskii, *Zh. Vychisl. Mat. i Mat. Fiz.*, 6, No. 6, 665 (1966).
6. J. M. Daniel, *SIAM J. Num. Anal.*, 4, No. 1, 10 (1967).
7. L. V. Kantorovich, *Uspekhi Mat. Nauk*, 3, No. 2, 89 (1948).

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

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.