SovietRxiv

Soviet-era science, translated into English

RATES OF CONVERGENCE OF COORDINATE RELAXATION FOR A QUADRATIC FUNCTIONAL

1967

SovietRxiv

View the original and related papers at https://sovietrxiv.org/items/ru-196701.78502

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

SovietRxiv

Abstract

Full Text

UDC 518:517.948

MATHEMATICS

Yu. I. LYUBICH

RATES OF CONVERGENCE OF COORDINATE RELAXATION FOR A QUADRATIC FUNCTIONAL

(Presented by Academician L. V. Kantorovich on 27 IV 1966)

Consider in the real n-dimensional Euclidean space R^n the functional $\varphi(x) = (Ax, x) - 2(f, x)$ $(A^* = A > 0,\ f \in R^n)$. Let $\{e_i\}_1^n$ be an orthonormal basis, $a_{ii} = (Ae_i, e_i)$. Coordinate-relaxation processes have the following structure (see, for example, ([1]), Ch. III):

$$x_{k+1} = x_k - q_k a_{i_k i_k}^{-1} (u_k, e_{i_k}) \qquad (k = 0, 1, 2, \dots), \tag{1}$$

where $u_k = Ax_k - f$ are the residual vectors, and $q_k \in [0, 2]$ are relaxation multipliers. The "relaxationality" of the process is expressed by the fact that $\varphi(x_{k+1}) \le \varphi(x_k)$ $(k = 0, 1, 2, \dots)$. The process is called strictly relaxation, if $\varepsilon = \inf_k q_k(2 - q_k) > 0$. It follows from the results of S. Schechter ([2]) that if, in the controlling sequence $\{i_k\}_0^\infty$ of a strictly relaxation process, each index from $\Omega^n = \{1, 2, \dots, n\}$ occurs arbitrarily far out (condition S), then $\{x_k\}_0^\infty$ converges to the solution x^0 of the equation $Ax = f$ (i.e., to the minimum point of the functional $\varphi(x)$). Still earlier A. Ostrowski obtained ([3]) convergence with the rate of a geometric progression (see also ([1]), pp. 259-262) under the assumption of "quasicyclicity" : $(\exists l)(\forall N)(\Omega^n \subset \{i_k\}_N^{N+l-1})$. The quasicyclic case was also studied under weaker restrictions than strict relaxationality ([3,4]). Our aim is to investigate the rate of convergence in the general case. We do this by relying on the theory constructed in ([5]).

Notation: $\lambda(N) = \min\{l \mid \Omega^n \subset \{i_k\}_N^{N+l-1}\}$; $N_{p+1} = N_p + \lambda(N_p)$ $(p = 0, 1, 2, \dots;\ N_0 = 0)$; $\nu(k) = \min\{p \mid N_{p+1} > k\}$; $\varepsilon_p = \min q_k(2 - q_k)$ $(N_p \le k < N_{p+1})$.

Theorem. Let $\Delta(x)$ be the quadratic error measure:

$$\Delta(x) = (A(x - x^0), x - x^0)\ (= \varphi(x) + (Ax^0, x^0)).$$

Under condition S the inequality holds

SovietRxiv

$$\Delta(x_k) \leq \Delta(x_0) \prod_{p=0}^{\nu(k)-1} (1-\omega\varepsilon_p) \qquad (k=0,1,2,...),$$

where the coefficient ω ($0<\omega\leq 1$) depends only on the dimension n and on the condition number $h(A)=\|A\|\cdot\|A^{-1}\|$.

Obviously, this theorem contains the above-cited results of S. Schechter and A. Ostrowski. It also leads to the following propositions:

Corollary 1. If condition S is satisfied and

$$\sum_{p=0}^{\infty} \varepsilon_p = \infty,$$

then $\{x_k\}_0^\infty$ converges to x^0.

Corollary 2. Under the conditions

$$\tilde{\nu} \equiv \lim_{k\to\infty} \frac{\nu(u)}{k} > 0, \qquad \tilde{\varepsilon} \equiv \lim_{\nu\to\infty} \frac{1}{\nu} \sum_{p=0}^{\nu-1} \varepsilon_p > 0$$

there is convergence to x^0 with the rate of a geometric progression:

$$\lim_{k\to\infty} \sqrt[k]{\Delta(x_k)} \leqslant e^{-\tilde{\omega}\tilde{\varepsilon}}.$$

Corollary 3. Under condition S and strict relaxationality, the inequality holds

$$\Delta(x_k) \leqslant \Delta(x_0)(1-\omega\varepsilon)^{\nu(k)} \qquad (k=0,1,2,...)$$

$$\varepsilon \equiv \inf_{p\geqslant 0} \varepsilon_p = \inf_{k\geqslant 0} q_k(2-q_k).$$

The theorem formulated is reduced directly to the following lemma.

Lemma. For every $n=1,2,...$ there exists a decreasing function $\omega_n(h)$ ($h\geqslant 1$), $\omega_n(1)=1$, such that for any chain of vectors $\{x_k\}_0^l$ constructed according to scheme (1) and such that $\{i_k\}_0^{l-1} \supset \Omega^n$, the inequality holds

$$\Delta(x_l) \leqslant \Delta(x_0)(1-\omega_n(h)\varepsilon_0),$$

where $h=h(A)$, $\varepsilon_0 = \min_{0\leqslant k<l} q_k(2-q_k)$.

SovietRxiv

Proof will be carried out by induction on n, on the basis of the inequality (see (5))

$$\Delta(x_{k+1}) \leqslant \Delta(x_k)(1 - h^{-1}(A)q_k(2 - q_k)\cos^2\theta_k), \tag{2}$$

where θ_k is the acute* angle between the residual vector u_k and the axis of the versor e_{i_k}. For $n = 1$ this inequality immediately leads to the required result, if one puts $\omega_1(h) = h^{-1}$. Consider the transition from $n-1$ to n $(n > 1)$.

In view of the monotonicity of the sequence $\{\Delta(x_k)\}_0^l$, one may assume that $\{i_k\}_0^{l-2} \neq \Omega^n$. For definiteness take $i_{l-1} = n$. Let

$$\psi(\xi) = \varphi(x_0 - \xi) - \varphi(x_0)$$

be a quadratic functional in the subspace R^{n-1} with basis $\{e_i\}_1^{n-1}$. It is easy to see that

$$\psi(\xi) = (PA\xi, \xi) - 2(Pu_0, \xi),$$

where P is the orthoprojector $R^n \to R^{n-1}$. The role of the function $\Delta(x)$ here will be played by

$$\delta(\xi) = (PA(\xi - \xi^0), \xi - \xi^0) = (A(\xi - \xi^0), \xi - \xi^0),$$

where $\xi^0 \in R^{n-1}$ is the solution of the equation $PA\xi = Pu_0$. Put $\tilde{x}^0 = x_0 - \xi^0$. Then $P(A\tilde{x}^0 - f) = 0$, whence it is seen that the vector $\tilde{x}^0 - x^0$ is A-orthogonal to the subspace R^{n-1}.

Consequently,

$$(A(x - \tilde{x}^0), x - \tilde{x}^0) = \Delta(x) - \Delta(\tilde{x}^0) \qquad (x - \tilde{x}^0 \in R^{n-1}).$$

Putting $x = x_0 - \xi$, we arrive at the formula

$$\delta(\xi) = \Delta(x) - \Delta(\tilde{x}^0) \qquad (\xi \in R^{n-1},\ x = x_0 - \xi). \tag{3}$$

Apply the induction hypothesis to the functional $\psi(\xi)$ and the chain $\xi_k = x_0 - x_k$ $(k = 0, 1, \dots, l-1)$. The multipliers q_k are preserved; the number ε_0 can only increase**; the condition number of the operator $PA|_{R^{n-1}}$ does not exceed $h(A)$. Therefore

$$\delta(\xi_{l-1}) \leqslant \delta(\xi_0)(1 - \eta\varepsilon_0), \tag{4}$$

SovietRxiv

* In the extreme case, a straight angle.

** The multiplier q_{l-1} drops out of the system of multipliers.

where $\eta = \omega_{n-1}(h(A))$. Taking formula (3) and the equality $\xi_0 = 0$ into account, from (4) we obtain

$$\Delta(x_{l-1}) \le (1 - \eta\varepsilon_0)\Delta(x_0) + \eta\varepsilon_0\Delta(\tilde{x}^0). \qquad (5)$$

Now, in connection with the forthcoming passage from (5) to an estimate of $\Delta(x_l)$, let us estimate from below $\cos^2\theta_{l-1}$ (see (2)). We have ($\tilde{u}^0 = A\tilde{x}^0 - f$)

$$\|u_{l-1} - \tilde{u}^0\|^2 \le \|A\|(A(x_{l-1} - \tilde{x}^0),\, x_{l-1} - \tilde{x}^0) = \|A\|\delta(\xi_{l-1}),$$

whence, by (4) and (3),

$$\|u_{l-1} - \tilde{x}^0\| \le \sqrt{\|A\|(1 - \eta\varepsilon_0)[\Delta(x_0) - \Delta(\tilde{x}^0)]}. \qquad (6)$$

But from elementary geometric considerations

$$\cos\theta_{l-1} \ge 1 - 2\|u_{l-1} - \tilde{u}^0\|\,\|\tilde{u}^0\|^{-1} \qquad (7)$$

(the vector $\tilde{u}^0$ is collinear with the unit vector e_n). Since

$$\|\tilde{u}^0\| \ge \sqrt{\|A^{-1}\|^{-1}\Delta(\tilde{x}^0)},$$

it follows, by (6) and (7), that

$$\cos\theta_{l-1} \ge 1 - 2\sqrt{h(A)(1 - \eta\varepsilon_0)(\rho^{-1} - 1)};$$

where $\rho = \Delta(\tilde{x}^0) : \Delta(x_0) \le 1$.

We shall regard ρ, ε_0, h as variables ($0 \le \rho \le 1,\ 0 \le \varepsilon_0 \le 1,\ h \ge 1$) and set*

$$r_n(\rho, \varepsilon_0, h) = \max\left(1 - 2\sqrt{h(1 - \eta\varepsilon_0)(\rho^{-1} - 1)},\, 0\right) \quad (\eta = \omega_{n-1}(h)). \qquad (8)$$

By virtue of (2), (5), (8),

$$\Delta(x_l) \le \Delta(x_0)[1 - \eta\varepsilon_0(1 - \rho)][1 - h^{-1}\varepsilon_0 r_n^2(\rho, \varepsilon_0, h)] = \Delta(x_0)[1 - \gamma_n(\rho, \varepsilon_0, h)\varepsilon_0],$$

SovietRxiv

where $h = h(A)$, and the function $\gamma_n(\rho, \varepsilon_0, h)$ is defined in an obvious way. This function is everywhere positive, continuous in (ρ, ε_0), decreases in h, and does not exceed 1. It remains to take

$$\omega_n(h) = \min_{(\rho,\varepsilon_0)} \gamma_n(\rho, \varepsilon_0, h),$$

and the lemma is proved.**

Kharkov State University
named after A. M. Gorky

Received
12 IV 1966

CITED LITERATURE

[1] D. K. Faddeev, V. N. Faddeeva, *Computational Methods of Linear Algebra*, Moscow-Leningrad, 1963.
[2] S. Schechter, Comm. Pure and Appl. Math., **12**, No. 2, 313 (1959).
[3] A. M. Ostrowski, Rend. Math. e Appl. (5), **14**, 140 (1954).
[4] A. S. Kelbasinskii, Vestn. Mosk. Univ., (I), No. 5, 40 (1960).
[5] Yu. I. Lyubich, DAN, **161**, No. 6, 1274 (1965).

* In this case $r_n(0, \varepsilon_0, h) = 0$ by definition.
** We note that $\omega_n(h) \geq h^{-1}$ $(n = 1, 2, ...)$, for $\gamma_n(1, \varepsilon_0, h) \equiv h^{-1}$.

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

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