Commentationes Mathematicae Universitatis Carolinae

Ivo Marek

On Lusterník's method of improving convergence of nonlinear iterative sequences

Commentationes Mathematicae Universitatis Carolinae, Vol. 6 (1965), No. 3, 371--380

Persistent URL: `http://dml.cz/dmlcz/105026`

Commentationes Mathematicae Universitatis Carolinae
6, 3 (1965)

ON LUSTERNIK'S METHOD OF IMPROVING CONVERGENCE OF NONLINEAR ITERATIVE SEQUENCES

Ivo MAREK, Praha

The purpose of this note is to show the possibility of using Lusternik's method of improving the convergence of iterative sequences for nonlinear problems. For linear processes Lusternik's method is well known, [2], [4].

Let $\mathfrak{X}$ be a complex Banach space and let $[\mathfrak{X}]$ denote the space of continuous linear mappings of the space $\mathfrak{X}$ into itself. The topology in $[\mathfrak{X}]$ is given, as usual, using uniform convergence on sets bounded in $\mathfrak{X}$ so that $[\mathfrak{X}]$ is also a Banach space.

We shall investigate the equation

$$Ax = f \ , \qquad (1)$$

where $f \in \mathfrak{X}$ and A is a nonlinear continuous mapping of the space $\mathfrak{X}$ into itself.

We suppose that the operator A has the following properties: Let $P \in [\mathfrak{X}]$ be a suitable operator such that $P^{-1} \in [\mathfrak{X}]$. Let the set $U \subset \mathfrak{X}$ contain the ball $S = \{x \in \mathfrak{X} \mid \| u - u_1 \| \leq \frac{1}{1-q} \| u_0 - u_1 \| \}$, where $q \in (0, 1)$, $u_1 = u_0 - PA(u_0) + Pf$.

(L) The inequality

$$\| u - v - PA(u) + PA(v) \| \leq q \| u - v \| \qquad (2)$$

holds for every pair $u, v \in U$.

(F) The operator A has a continuous Fréchet derivative A' in U.

(D) The operator $T = I - PB$, where $B = A'(u^*)$, has the dominant eigenvalue λ_1 and this value is a simple pole of the resolvent $R(\lambda, T) = (\lambda I - T)^{-1}$, where I denotes the identity operator.

From the assumption (L) there follows the existence of a unique solution u^* of the equation (1) and the convergence of the sequence $\{u_n\}$, where

$$u_{n+1} = u_n - PA(u_n) + Pf \tag{3}$$

to the solution u^*. The estimate

$$\| u^* - u_n \| \leq \frac{q^n}{1-q} \| u_1 - u_0 \|$$

for the error is known ([1] p.563).

It is easy to see that the following expression is valid for the terms of the sequence $\{h_n\}$:

$$h_{n+1} = u_{n+1} - u^* = u_n - u^* - P[A(u_n) - A(u^*)];$$

hence and from (F) we obtain that

$$h_{n+1} = Th_n + P\omega(u^*, h_n), \tag{4}$$

where $\omega(u^*, h_n) = A'(u^*)h_n - A(u^* + h_n) + A(u^*)$.

From the definition of the $A'(u^*)$ it follows that

$$\lim_{n \to \infty} \frac{\| \omega(u^*, h_n) \|}{\| h_n \|} = 0. \tag{5}$$

The expression (4) implies, easily, the equality

$$h_{n+1} = T^{n+1} h_0 + \sum_{k=0}^{n} T^{n-k} P\omega(u^*, h_k). \tag{6}$$

Now define a second sequence $\{k_n\}$ by the formula

$$(7) \qquad k_n = \frac{1}{1-\lambda_1}\,(h_{n+1} - \lambda_1 h_n).$$

We shall prove that the sequence $\{k_n\}$ converges to the zero-element more rapidly than the $\{k_n\}$; this implies that the sequence $\{v_n\}$, where

$$(8) \qquad v_n = \frac{1}{1-\lambda_1}\,(u_{n+1} - \lambda_1 u_n),$$

converges to the exact solution u^* of (1) more rapidly than the sequence $\{u_n\}$.

For $k_n = v_n - u^*$ we obtain, using (6), that

$$k_n = \frac{1}{1-\lambda_1}\{T^n(T-\lambda_1 I)h_0 + \sum_{k=0}^{n} T^{n-k} P\omega(u^*, h_k) - \lambda_1 \sum_{k=0}^{n-1} T^{n-k-1} P\omega(u^* h_k)$$

and hence

$$k_n = \frac{1}{1-\lambda_1}\{T^n(T-\lambda_1 I)h_0 + P\omega(u^*, h_n) + \sum_{k=0}^{n-1} T^{n-k-1}(T-\lambda_1 I)P\omega(u^*, h_k)\}.$$

We shall use the following notation. $C_0 = \{\lambda \mid |\lambda| = \rho\}$, where ρ is such that the $K_\sigma = \{\lambda \mid |\lambda| < \rho\}$ contains the spectrum $\sigma(T)$ of the operator T. For the circle K_1, where $K_1 = \{\lambda \mid |\lambda| < \rho_1\}$, suppose that $K_1 \cap \sigma(T) = \sigma(T) - \{\lambda_1\}$. Furthermore, let $C_1 = \{\lambda \mid |\lambda| = \rho_1\}$, $K_0 = \{\lambda \mid |\lambda - \lambda_1| < \rho_0\}$, $K_0 \cap \sigma(T) = \{\lambda_1\}$, $C_0 = \{\lambda \mid |\lambda - \lambda_1| = \rho_0\}$.

According to [5] p.305, the resolvent $R(\lambda, T)$ has a

Laurent expansion in a neighbourhood of the point λ_1,

$$R(\lambda,T)=\sum_{k=0}^{\infty} A_k(\lambda-\lambda_1)^k+\sum_{k=1}^{\infty} B_k(\lambda-\lambda_1)^{-k},$$

where $A_k \in [\mathcal{X}]$ and

$$B_1=\frac{1}{2\pi i}\int_{C_0} R(\lambda,T)d\lambda,\quad B_{k+1}=(T-\lambda_1 I)B_k .$$

In particular, in our case assumption (D) implies that $B_k=\Theta$ for $k>1$, where Θ denotes the zero-operator.

According to [5] p.290 and the Cauchy theorem, it follows that

$$h_n=\frac{1}{2\pi i}\int_{C_\sigma} R(\lambda,T)\{\lambda^n h_0 + \sum_{k=0}^{n-1}\lambda^{n-k-1}P\omega(u^*,h_k)\}d\lambda ;$$

hence

$$h_n=\lambda_1^n B_1 h_0+\sum_{k=0}^{n-1}\lambda_1^{n-k-1}B_1 P\omega(u^*,h_k)+ \tag{9}$$
$$+\frac{1}{2\pi i}\int_{C_1} R(\lambda,T)\{\lambda^n h_0+\sum_{k=0}^{n-1}\lambda^{n-k-1}P\omega(u^*,h_k)\}d\lambda$$

and therefore, for any $\eta>0$ there is a $c=c(\eta)$ such that

$$\| h_n \| \leq c\,|\tilde{\lambda}_1|^n,\quad \tilde{\lambda}_1=\lambda_1+\eta \tag{10}$$

where the constant c does not depend on n. Evidently $|\tilde{\lambda}_1|\leq q$.

Similarly we obtain the following expression for k_n:

$$k_n=\frac{1}{2\pi i}\int_{C_\sigma}\frac{\lambda^{n+1}-\lambda_1\lambda^n}{1-\lambda_1}R(\lambda,T)h_0\,d\lambda+$$
$$+\frac{1}{1-\lambda_1}P\omega(u^*,h_n)+\frac{1}{2\pi i}\int_{C_\sigma}\frac{1}{1-\lambda_1}\sum_{k=0}^{n-1}(\lambda^{n-k}-\lambda_1\lambda^{n-k-1})$$
$$R(\lambda,T)P\omega(u^*,h_k)\,d\lambda .$$

Let us put

$$W_1=\frac{1}{2\pi i}\int_{C_\sigma}\lambda^n\frac{\lambda-\lambda_1}{1-\lambda_1}R(\lambda,T)h_0\,d\lambda ,$$

$$W_2 = \frac{1}{2\pi i} \int_{C_\sigma} \frac{\lambda - \lambda_1}{1-\lambda_1} \sum_{k=0}^{n-1} \lambda^{n-k-1} R(\lambda, T) P\omega(u^*, h_k) d\lambda ,$$

$$W_3 = \frac{1}{1-\lambda_1} P\omega(u^*, h_n) .$$

We shall estimate the norms of the vectors W_1, W_2, W_3.

From the Cauchy theorem it follows that

$$W_1 = \frac{1}{2\pi i} \int_{C_1} \lambda^n \frac{\lambda - \lambda_1}{1-\lambda_1} R(\lambda, T) h_0 \, d\lambda ;$$

therefore

$$(11) \quad \| W_1 \| \leq \frac{\rho_1^{n+1}}{1-|\lambda_1|} \max_{\lambda \in C_1} \|(\lambda - \lambda_1) R(\lambda, T)\| \, \| h_0 \| .$$

Similarly

$$W_2 = \frac{1}{2\pi i} \int_{C_1} \frac{\lambda - \lambda_1}{1-\lambda_1} R(\lambda, T) \sum_{k=0}^{n-1} \lambda^{n-k-1} P\omega(u^*, h_k) d\lambda ,$$

so that

$$\| W_2 \| \leq \frac{1}{1-|\lambda_1|} \max_{\lambda \in C_1} \|(\lambda - \lambda_1) R(\lambda, T)\| \, \|P\| \sum_{k=0}^{n-1} \rho_1^{n-k} \|\omega(u^*, h_k)\| .$$

According to (6) and (10) we obtain the estimate

$$(12) \qquad \|\omega(u^*, h_k)\| = o(|\tilde{\lambda}_1|^k) .$$

Evidently

$$\sum_{k=0}^{n-1} \rho_1^{n-k} \|\omega(u^*, h_k)\| = o\Big(\sum_{k=0}^{n-1} \rho_1^{n-k} |\tilde{\lambda}_1|^k\Big) =$$

$$= o\Big(\sum_{k=0}^{n-1} \rho_1^k |\tilde{\lambda}_1|^{n-k}\Big) = o(|\tilde{\lambda}_1|^n)$$

and thus

$$(13) \qquad \| W_2 \| = o(|\tilde{\lambda}_1|^n) .$$

The relation

(14) $$\|W_3\| = o(\|h_n\|) = o(|\tilde{\lambda}_1|^n)$$

is a direct consequence of the definition of Fréchet derivatives.

Summarizing the estimates (11),(13)and(14) we obtain the relation

$$\lim_{n\to\infty} \frac{\|k_n\|}{|\tilde{\lambda}_1|^n} = 0 .$$

This relation expresses the fact that the sequence $\{\|k_n\|\}$ converges to zero more rapidly than $\{|\tilde{\lambda}_1|^n\}$. Therefore the sequence $\{v_n\}$ converges to the solution u^* of the equation (1), in general, more rapidly than the sequence $\{u_n\}$; and this was what we set out to prove.

It is evident that the eigenvalue λ_1 of the operator $T = I - PB$, $B = A'(u^*)$, is usually unknown. But occasionally we know at least some approximations $\lambda_{(n)}$, of the value mentioned.

Let us investigate the sequence $\{\tilde{k}_n\}$ defined by the formula

$$\tilde{k}_n = \frac{1}{1-\lambda_{(n)}}(h_{n+1} - \lambda_{(n)} h_n) ,$$

assuming that

(15) $$\lim_{n\to\infty} \lambda_{(n)} = \lambda_1 ,$$

in place of the sequence $\{k_n\}$ defined by (7).

Similarly as for the k_n we obtain for the $\tilde{k}_n$ the expression

$$\tilde{k}_n = \frac{1}{1-\lambda_{(n)}}\{\lambda_1^n(\lambda_1-\lambda_{(n)})B_1 h_0 + P\omega(u^*, h_n) +$$

$$+\sum_{k=0}^{n-1} \lambda_1^{n-k-1}(\lambda_1-\lambda_{(n)})\, B_1\, P\omega(u^*, h_k) +$$

$$+\frac{1}{2\pi i}\int_{C_1} R(\lambda,T)[\lambda^n(\lambda-\lambda_{(n)})\, h_0 +$$

$$+\sum_{k=0}^{n-1} \lambda^{n-k-1}(\lambda-\lambda_{(n)})\, P\omega(u^*, h_k)]\, d\lambda \},$$

and hence the estimate

$$\|\tilde{k}_n\| \leq |\lambda_1|^n \frac{|\lambda_1-\lambda_{(n)}|}{|1-\lambda_{(n)}|}\, \|B_1 h_0\| + \frac{1}{|1-\lambda_{(n)}|}\, \|P\|\, \|\omega \cdot$$

$$(u^*, h_n)\| + \max_{\lambda\in C_1} \|(\lambda-\lambda_{(n)})R(\lambda,T)\|$$

$$[\rho_1^n \|h_0\| + \|P\| \sum_{k=0}^{n-1} \rho_1^{n-k}\, \|\omega(u^*, h_k)\|\,] \;.$$

From this result, and according to (10) and (12), it follows that

$$\|\tilde{k}_n\| \leq c\, |\lambda_1-\lambda_{(n)}|\, |\lambda_1|^n + \sigma(|\tilde{\lambda}_1|^n),$$

where c is independent of n. Consequently (15) implies the relation

$$\lim_{n\to\infty} \frac{\|\tilde{k}_n\|}{|\tilde{\lambda}_1|^n} = 0 \;.$$

Thus we have proved that the sequence $\{\tilde{k}_n\}$ converges to the zero-element, in general, more rapidly than the sequence $\{h_n\}$. In other words, the sequence $\{\tilde{v}_n\}$, where

$$(16)\cdot \quad \tilde{v}_n = \frac{1}{1-\lambda_{(n)}}\,(u_{n+1} - \lambda_{(n)}\, u_n),$$

converges to the exact solution u^* of the equation (1) more rapidly than the sequence $\{u_n\}$.

Now we shall show one possible method for the construction of approximations $\lambda_{(n)}$. According to (F) the operator A possesses a continuous Fréchet derivative A' in the neighbourhood $\mathcal{U}$ of the solution u^* of equation (1); this means that $A'(w_s) \to A'(u^*)$ for $w_s \to u^*$. In particular, for the sequence $\{u_n\}$ defined by (3) we have

$$\lim_{n \to \infty} A'(u_n) = A'(u^*) .$$

The condition (15) is then fulfilled if the following assumption (D′) holds..

(D′) The operators $T_n = I - PA'(u_n)$ have dominant eigenvalues $\lambda_{(n)}$ and these values are simple poles of the resolvents $R(\lambda, T_n) = (\lambda I - T_n)^{-1}$.

We shall summarize the results obtained in the following theorem.

Theorem. Let the assumptions (F),(L),(D),(D′) be fulfilled for the mapping A of the space $\mathcal{X}$ into itself. Then the sequence $\{\tilde{v}_n\}$ defined by (16) converges to the solution u^* of the equation (1), in general, more rapidly than the sequence $\{u_n\}$ defined by (3); more precisely

$$\lim_{n \to \infty} \frac{\| \tilde{v}_n - u^* \|}{|\lambda_1|^n} = 0 .$$

If some further qualitative properties of the eigenvalue λ_1 are known, it is not necessary to calculate neither the value λ_1 nor the approximations $\lambda_{(n)}$: one can use the formula

$$w_n = \alpha_n u_{n+1} + (1 - \alpha_n) u_n ,$$

where α_n is a suitable parameter (see [3]), instead of (8) and (16). For l_n, where

$$l_n = \alpha_n h_{n+1} + (1-\alpha_n) h_n ,$$

we obtain according to (9) the following expression

$$l_n = \lambda_1^n [\alpha_n \lambda_1 + (1-\alpha_n)] B_1 h_0 + \alpha_n P\omega(u^*, h_n) + \\ + \sum_{k=0}^{n-1} \lambda_1^{n-k-1} [\alpha_n \lambda_1 + (1-\alpha_n)] B_1 P\omega(u^*, h_k) + \\ + \frac{1}{2\pi i} \int_{C_1} R(\lambda, T) [\alpha_n \lambda + (1-\alpha_n)] \{ \lambda^n h_0 + \\ + \sum_{k=0}^{n-1} \lambda^{n-k-1} P\omega(u^*, h_k) \} d\lambda .$$

From this it is evident that

$$\lim_{n\to\infty} \frac{\| l_n \|}{|\tilde{\lambda}_1|^n} = 0$$

if

$$\alpha_n \to \frac{1}{1-\lambda_1} .$$

For practical calculations one can proceed in many cases as follows. We take a constant α instead of the sequence $\{\alpha_n\}$, and form the sequence $\{\hat{w}_n\}$, where

$$\hat{w}_n = \alpha u_{n+1} + (1-\alpha) u_n .$$

For the terms of this latter one has

$$\frac{\| \hat{w}_n - u^* \|}{|\tilde{\lambda}_1|^n} \leqq c < 1 \tag{17}$$

provided that

$$|\alpha \lambda_1 + (1-\alpha)| < |\lambda_1| . \tag{18}$$

If the value λ_1 is, for example, positive, then the condition (18) is equivalent to the conditions

$$1 < \alpha < \frac{2}{1-\lambda_1} .$$

The estimate (17) may be improved by taking α near to $(1 - \lambda_1)^{-1}$.

References:

[1] L.V. KANTOROVIČ, G.P.AKILOV, Funkcional´nyj analiz v normirovannych prostranstvach.Fizmatgiz,Moskva 1959.

[2] L.A. LJUSTERNIK, Zamečanija k čislennomu rešeniju krajevych zadač uravněnija Laplasa i vyčislenijam sobstvennych značenij metodom setok. Trudy Matem.inst.AN SSSR 20 (1947),49-64.

[3] P. LUSTGARTEN, Acceleration of convergence in solving of linear equations by iterations and obtaining the latent roots of a given matrix. Rep.Venezuela Bol.Acad.Sci.,Fis.Mat.Nat.23(1963),26-33.

[4] I. MAREK, Ob odnom metodě uskorenija schodimosti iteracionnych processov.Žurnal vyč.mat.i mat.fiz.2(1962),963-971.

[5] A.E. TAYLOR, Introduction to functional analysis. J.Wiley publ.,New York 1958.

(Received September 8,1965)