https://doi.org/10.15446/recolma.v55n2.102622

Revista Colombiana de Matemáticas
Volumen 55(2021)2, páginas 167-175

New linearization method for nonlinear problems in Hilbert space

Nuevo método de linealización para problemas no lineales en espacios de Hilbert

Nada Bouazila[1], Hamza Guebbai[1,✉], Wassim Merchela[2]

[1]Université 8 Mai 1945, Guelma, Algeria

[2] Derzhavin Tambov State University, Tambov, Russia

Abstract. In this paper, we build a Newton-like sequence to approach the zero of a nonlinear Fréchet differentiable function defined in Hilbert space. This new iterative sequence uses the concept of the adjoint operator, which makes it more manageable in practice compared to the one developed by Kantorovich which requires the calculation of the inverse operator in each iteration. Because the calculation of the adjoint operator is easier compared to the calculation of the inverse operator which requires in practice solving a system of linear equations, our new method makes the calculation of the term of our new sequence easier and more convenient for numerical approximations. We provide an a priori convergence theorem of this sequence, where we use hypotheses equivalent to those constructed by Kantorovich, and we show that our new iterative sequence converges towards the solution.

Key words and phrases. Nonlinear problems, Newton-like method, Fréchet differentiability, Adjoint Operator.

2020 Mathematics Subject Classification. 49M15, 49J50.

Resumen. En este artículo, construimos una sucesión similar a la de Newton para acercarnos al cero de una función diferenciable en el sentido Fréchet no lineal definida en un espacio de Hilbert. Esta nueva sucesión utiliza el concepto del adjunto del operador, que hace que el proceso iterativo sea más manejable en la práctica en comparación al desarrollado por Kantorovich que requiere el cálculo del operador inverso en cada iteración. Dado que el cálculo del operador adjunto es más fácil en comparación con el cálculo del operador inverso que en la práctica equivale a resolver un sistema de ecuaciones, nuestra nuevo método hace que el cálculo del término de nuestra nueva sucesión sea más fácil y conveniente para la aproximación numérica. Proporcionamos un teorema de convergencia a priori de esta sucesión , donde usamos unas hipótesis equivalentes a las construidas por Kantorovich, y mostramos que nuestra nueva sucesión iterativa converge hacia la solución.

Palabras y frases clave. Problemas no lineales, método tipo Newton, diferenciabilidad de Fréchet, operador adjunto.

1. Introduction

The mathematical feat in the linearization theory of nonlinear problems in infinite dimensional spaces remains forever the metamorphosis that Kantorovich brought to the Newton's method which was basically intended for real nonlinear equations resolution (see [2, 3, 5, 7, 8]).

Kantorovich's genius appears when he gets the correspondence between $\dfrac{1}{A'(x)}$ in the real case and $(A'(x))^{-1}$ in the case of a Banach space, where the first derivative is in the classical sense and the second is a Fréchet derivative (see [1, 4, 6, 9]). For the equation

$$A(\zeta) = 0_{\mathrm{H}}, \tag{1}$$

where, $A : \Omega \longrightarrow \mathrm{H}$, Ω is an open set of the Hilbert space $(\mathrm{H}, < \cdot, \cdot >_{\mathrm{H}})$ and A is Fréchet differentiable on Ω, Kantorovich built the following sequence:

$$\begin{cases} \xi^0 \text{ chosen in } \Omega, \\ \xi^{k+1} = \xi^k - (A'(\xi^k))^{-1} A(\xi^k), \ k \geq 0, \end{cases}$$

which converges to the exact solution of (1) provided ξ^0 is well chosen. Note that $\forall k \geq 0, \ A'(\xi^k) \in \mathrm{BL(H)}$ the set of all bounded linear operators defined from H to itself.

Nevertheless, the numerical approximation of the terms $(A'(\xi^k))^{-1}$ has made this sequence very difficult to construct. In fact, to calculate ξ^{k+1} we must solve the following equation:

$$A'(\xi^k) \left(\xi^{k+1} - \xi^k \right) = -A(\xi^k),$$

which is very hard to do and usually ξ^{k+1} is numerically approximated in each iteration (see [1, 6, 9]).

The point of this article is to define a Newton-like sequence that does not require the calculation of the inverse operators $(A'(\xi^k))^{-1}, \ k \geq 0$. To achieve our objective, we exploit the notion of adjoint operator offered by the Hilbert space structure, i.e, T^* the adjoint operator of $T \in \mathrm{BL(H)}$, is defined by

$$\forall x, y \in \mathrm{H} \ : < Tx, y >_{\mathrm{H}} = < x, T^* y >_{\mathrm{H}},$$

which allows us to build this new iterative sequence

$$\begin{cases} \xi^0 \text{ chosen in } \Omega, \\ \xi^{k+1} = \xi^k - \alpha_k (A'(\xi^k))^* A(\xi^k), \ k \geq 0, \end{cases}$$

where, $\{\alpha_k\}_{k\geq 0}$ are real parameters. In the next section, we give sufficient conditions on A and $\{\alpha_k\}_{k\geq 0}$ that ensure the convergence of the previous sequence to the exact solution of (1).

2. Main results

Let $(\mathrm{H}, <.,.>_{\mathrm{H}})$ be a Hilbert space, with the norm $\|.\|_{\mathrm{H}} = \sqrt{<.,.>_{\mathrm{H}}}$. We denote BL(H) the Banach space of bounded linear operators from H to H endowed with the norm:

$$\forall T \in \mathrm{BL(H)}, \ \|T\| = \sup \left\{ \|Tx\|_{\mathrm{H}} \ : \ \|x\|_{\mathrm{H}} = 1 \right\}.$$

Let $A : \Omega \subseteq \mathrm{H} \longrightarrow \mathrm{H}$ be a nonlinear Fréchet differentiable function defined on a nonempty open set Ω of H. We suppose that there is a unique solution ζ of (1) in Ω, i.e, $A(\zeta) = 0_{\mathrm{H}}$, $A'(.) : \mathrm{H} \longrightarrow \mathrm{BL(H)}$ is $\delta-$Lipschitz on Ω i.e.

$$\forall \chi, \ \xi \in \Omega, \ \|A'(\chi) - A'(\xi)\| \leq \delta \|\chi - \xi\|_{\mathrm{H}}, \quad (H1)$$

and

$$\textit{there exists } M > 0, \forall u \in \mathrm{H}, \|A'(\zeta)u\|_{\mathrm{H}} \geq M \|u\|_{\mathrm{H}}. \quad (H2)$$

This last condition is equivalent to the one required by Kantorovich, namely

$$[A'(\zeta)]^{-1} \text{ exists and is bounded.}$$

We need the following lemma to proof the main theorem of this paper.

Lemma 2.1. *For $\lambda \in]0,1[$ and $\mu \in]\lambda, 1[$, we define $\{v_k\}_{k\in\mathrm{N}}$ by: $v_0 = \mu - \lambda$, $v_{k+1} = \lambda v_k + v_k^2$, $k \geq 0$. Then, $\forall k \geq 0$, $v_k \leq \mu^k(\mu - \lambda)$ and $\lim_{k\to+\infty} v_k = 0$.*

Proof. It is clear that $g(x) = \lambda x + x^2$ is increasing over $]0, 1-\lambda[$ and has two fixed points: 0 and $1-\lambda$. Which means that $\{v_k\}_{k\in\mathrm{N}}$ is monotonic and included in $]0, 1-\lambda[$. But, we have

$$v_1 - v_0 = -\mu(1-\lambda)^2(1-\mu) < 0.$$

Then, v_k is decreasing and for $k \geq 0$,

$$v_{k+1} = (\lambda + v_k)v_k \leq (\lambda + v_0)v_k \leq \mu v_k.$$

Therefore, for all $k \geq 0$,

$$v_k \leq \mu^k(\mu - \lambda).$$

☑

Now, we state the main theorem of this paper.

Theorem 2.2. *Let* $\left\{\xi^k\right\}_{k\in\mathrm{N}} \subset \mathrm{H}$ *be a Newton-like sequence defined by*

$$\begin{cases} \xi^0 \ \text{chosen} \ \ \text{in}\ \Omega, \\ \xi^{k+1} = \xi^k - \alpha_k (A'(\xi^k))^* A(\xi^k), \ k \geq 0, \end{cases}$$

with, $0 < \alpha_k \left\|A'(\xi^k)\right\|^2 \leq \alpha < 1$ *and* $\alpha_k \geq \beta > 0$ *for all* $k \geq 0$*,and under the hypothesis (H1)-(H2). Then there exists* $C > 0$ *such that,*

$$\left\|\xi^0 - \zeta\right\|_{\mathrm{H}} < C \Rightarrow \lim_{k\to+\infty} \left\|\xi^k - \zeta\right\|_{\mathrm{H}} = 0.$$

Proof. We have, for all $k \geq 0$

$$\begin{aligned} \xi^{k+1} - \zeta =& \xi^k - \zeta - \alpha_k A'(\xi^k)^* \left(A(\xi^k) - A(\zeta)\right) \\ =& \xi^k - \zeta - \alpha_k A'(\xi^k)^* \left(A'(\zeta)\left(\xi^k - \zeta\right) + o\left(\left\|\xi^k - \zeta\right\|_{\mathrm{H}}^2\right)\right) \\ =& \xi^k - \zeta - \alpha_k A'(\xi^k)^* \left(\left(A'(\xi^k) + A'(\zeta) - A'(\xi^k)\right)\left(\xi^k - \zeta\right)\right. \\ &+ o\left.\left(\left\|\xi^k - \zeta\right\|_{\mathrm{H}}^2\right)\right). \end{aligned}$$

Then, there exists $c > 0$ such that,

$$\begin{aligned} \left\|\xi^{k+1} - \zeta\right\|_{\mathrm{H}} \leq& \left\|(I - \alpha_k A'(\xi^k)^* A'(\xi^k))(\xi^k - \zeta)\right\|_{\mathrm{H}} + c\alpha_k \left\|A'(\xi^k)\right\| \left\|\xi^k - \zeta\right\|_{\mathrm{H}}^2 \\ &+ \alpha_k \left\|A'(\xi^k)^* \left(A'(\zeta) - A'(\xi^k)\right)\left(\xi^k - \zeta\right)\right\|_{\mathrm{H}}, \end{aligned}$$

where, I denotes the identity operator of BL(H). Using the fact that A' is $\delta-$Lipchitz, we obtain, for all $k \geq 0$,

$$\left\|\xi^{k+1} - \zeta\right\|_{\mathrm{H}} \leq \left\|I - \alpha_k A'(\xi^k)^* A'(\xi^k)\right\| \left\|\xi^k - \zeta\right\|_{\mathrm{H}} + \alpha_k \left(c + \delta\right) \left\|A'(\xi^k)\right\| \left\|\xi^k - \zeta\right\|_{\mathrm{H}}^2.$$

☑

For $r \in \left]0, M\delta^{-1}\right[$, we define $\Omega_r = \{\xi \in \Omega \ : \ \|\xi - \zeta\|_{\mathrm{H}} < r\}$. For all $\xi \in \Omega_r$ and for all $u \in \mathrm{H}$, we get

$$\begin{aligned} \|A'(\xi)u\|_{\mathrm{H}} &= \|(A'(\zeta) - (A'(\zeta) - A'(\xi)))\, u\|_{\mathrm{H}} \\ &\geq \|A'(\zeta)u\|_{\mathrm{H}} - \|(A'(\zeta) - A'(\xi))\, u\|_{\mathrm{H}} \\ &\geq (M - \delta \|\zeta - \xi\|_{\mathrm{H}}) \|u\|_{\mathrm{H}} \geq (M - \delta r) \|u\|_{\mathrm{H}}. \end{aligned}$$

In the same way,

$$\begin{aligned} \|A'(\xi)u\|_{\mathrm{H}} &= \|(A'(\zeta) - (A'(\zeta) - A'(\xi)))\, u\|_{\mathrm{H}} \\ &\leq \|A'(\zeta)u\|_{\mathrm{H}} + \|(A'(\zeta) - A'(\xi))\, u\|_{\mathrm{H}} \\ &\leq (\|A'(\zeta)\| + \delta r) \|u\|_{\mathrm{H}}. \end{aligned}$$

Then, for all $\xi \in \Omega_r$,

$$M - \delta r \leq \|A'(\xi)\| \leq \|A'(\zeta)\| + \delta r.$$

On the other hand, we have for all $k \geq 0$, $I - \alpha_k A'(\xi^k)^* A'(\xi^k)$ is a self-adjoint operator, then

$$\begin{aligned} \left\|I - \alpha_k A'(\xi^k)^* A'(\xi^k)\right\| &= \sup_{\|u\|_{\mathrm{H}}=1} \left|< \left(I - \alpha_k A'(\xi^k)^* A'(\xi^k)\right) u, u >_{\mathrm{H}}\right| \\ &= \sup_{\|u\|_{\mathrm{H}}=1} \left|1 - \alpha_k \left\|A'(\xi^k)u\right\|_{\mathrm{H}}^2\right|. \end{aligned}$$

But, for $\|u\|_{\mathrm{H}} = 1$,

$$\alpha_k \left\|A'(\xi^k)u\right\|_{\mathrm{H}}^2 \leq \alpha < 1.$$

And, if $\xi^k \in \Omega_r$, then

$$\begin{aligned} \left\|I - \alpha_k A'(\xi^k)^* A'(\xi^k)\right\| &= \sup_{\|u\|_{\mathrm{H}}=1} \left(1 - \alpha_k \left\|A'(\xi^k)u\right\|_{\mathrm{H}}^2\right) \\ &\leq \left(1 - \beta \left(M - \delta r\right)^2\right) < 1. \end{aligned}$$

Knowing that Ω is an open set, then , there exists $\kappa > 0$

such that,

$$\Omega_\kappa = \{\xi \in X \ : \ \|\xi - \zeta\|_{\mathrm{H}} < \kappa\} \subset \Omega.$$

Set $\eta = \dfrac{c + \delta}{M - \delta r}$. Now if we take $\xi^0 \in \Omega_r$ such that

$$\left\|\xi^0 - \zeta\right\|_{\mathrm{H}} < \min(\eta^{-1} \beta \left(M - \delta r\right)^2, \kappa) = C,$$

then, for all $k \geq 1$, $\xi^k \in \Omega_r \cap \Omega_\kappa$, i.e, $\left\{\xi^k\right\}_{k \in \mathbb{N}}$ is well-defined and

$$\eta \left\|\xi^k - \zeta\right\|_{\mathrm{H}} \leq v_k,$$

where, $\{v_k\}_{k \in \mathrm{N}}$ is defined as in the previous lemma with $\lambda = \left(1 - \beta \left(M - \delta r\right)^2\right)$ and $\mu = \eta \left\|\xi^0 - \zeta\right\|_{\mathrm{H}} + \lambda$.

3. Numerical example

Let A be a function defined as

$$\begin{array}{cccc} A: & \mathbb{R}^n & \longrightarrow & \mathbb{R}^n, \\ & \xi = (\xi_1, \xi_2, .., \xi_n) & \longmapsto & y = A(\xi), \end{array}$$

such that

$$y_i = \begin{cases} 1-\xi_i^2 + \sum_{j=1}^{i-1}(1-\xi_j^2)^2 + \sum_{j=i+1}^{n}(1-\xi_j)^2 & 2 \le i \le n-1 \\ 1-\xi_1^2 + \sum_{j=2}^{n}(1-\xi_j)^2 & i=1 \\ 1-\xi_n^2 + \sum_{j=1}^{n-1}(1-\xi_j^2)^2 & i=n \end{cases}$$

It is clear that the equation $A(\xi) = 0_{\mathbb{R}^n}$ has a solution $\zeta = \underbrace{(1,1,..,1)}_{n\ times}$, A is Fréchet differentiable on $\mathbb{R}^n$, with

$$A'(\xi)(i,j) = \begin{cases} -2\xi_i & i=j \\ 4\xi_j(\xi_j^2-1) & i>j \\ 2(\xi_j-1) & i<j \end{cases}$$

so $A'(\xi)$ has the form

$$A'(\xi)(i,j) = \begin{pmatrix} -2\xi_1 & 2(\xi_2-1) & \dots & 2(\xi_n-1) \\ 4\xi_1({\xi_1}^2-1) & -2\xi_2 & \dots & 2(\xi_n-1) \\ \vdots & \vdots & \vdots & \vdots \\ 4\xi_1({\xi_1}^2-1) & 4\xi_2(\xi_2^2-1) & \dots & -2\xi_n \end{pmatrix}.$$

Let $\xi, \tilde{\xi} \in B(\zeta, R) = \{\xi \in \mathbb{R}^n : ||\zeta - \xi||_2 \le R\}$, then

$$\begin{aligned} ||A'(\xi) - A'(\tilde{\xi})|| &\le n||A'(\xi) - A'(\tilde{\xi})||_1 \\ &\le n^2 \max_{1\le i,j\le n} |A'(\xi)(i,j) - A'(\tilde{\xi})(i,j)|. \end{aligned}$$

But, $|A'(\xi)(i,i) - A'(\tilde{\xi})(i,i)| = 2|\xi - \tilde{\xi}|$, for $1 \le i \le n$. If $1 \le i < j \le n$,

$$|A'(\xi)(i,j) - A'(\tilde{\xi})(i,j)| = 2|\xi - \tilde{\xi}|,$$

and if $1 \le j < i \le n$,

$$\begin{aligned} |A'(\xi)(i,j) - A'(\tilde{\xi})(i,j)| &= |4\xi_j(\xi_j^2-1) - 4\tilde{\xi}_j(\tilde{\xi}_j^2-1)| \\ &\le 4|\xi_j^3 - \tilde{\xi}_j^3| + 4|\xi_j - \tilde{\xi}_j| \\ &\le \left(4|\xi_j^2 + \xi_j\tilde{\xi}_j + \tilde{\xi}_j^2| + 4\right)|\xi_j - \tilde{\xi}_j| \\ &\le 4\left(\frac{3}{2}|\xi_j^2 + \tilde{\xi}_j^2| + 1\right)|\xi_j - \tilde{\xi}_j|. \end{aligned}$$

On the other hand, $\xi_j^2 \leq ||\xi||_2^2 \leq (||\xi - \zeta||_2 + ||\zeta||_2)^2 \leq (R+n)^2$, for $1 \leq j \leq n$. Then,

$$\begin{aligned} ||A'(\xi) - A'(\tilde{\xi})|| &\leq 4n^2(3(R+n)^2+1)||\xi - \tilde{\xi}||_\infty \\ &\leq 4n^2(3(R+n)^2+1)||\xi - \tilde{\xi}||_2. \end{aligned}$$

Which means that $A'(.)$ is $4n^2(3(R+n)^2+1)$-Lipchitz in all ball of raduis R and center ζ . Also,

$$A'(\zeta) = \begin{pmatrix} -2 & 0 & \dots & 0 \\ 0 & -2 & & 0 \\ \vdots & & & \vdots \\ 0 & \dots & 0 & -2 \end{pmatrix},$$

which is invertible. In each iteration $k \geq 1$, we take

$$\alpha_k = 0.99\frac{1}{n^2}||A'(\xi^k)||_1^{-2} \leq 0.99||A'(\xi^k)||^{-2},$$

which gives $\alpha = 0.99$ and $\beta > 0$ in our convergence thoerem.

All the hypotheses of our theorem are verifed, we can pass now to the algorithm to get the numerical result.

Algorithm of the new Newton-like method (NL)

step 1. Give an initial point ξ^0 and $k_{\max} \geq 1$ .Set $k = 0$.
step 2. Set $n = size(\xi^0)$.
step 3. If $k \leq k_{\max}$, then stop .
step 4. $\alpha_k = 0.99\frac{1}{n^2}||A'(\xi^k)||_1^{-2}$.
step 5. $\xi^{k+1} = \xi^k - \alpha_k A'(\xi^k)^* A(\xi^k)$.
step 6. Set $k = k + 1$.
step 7. $Er(k) = \sqrt{A(\xi^k)^t A'(\xi^k)}$, then go to step 2.

$Er(k) = \sqrt{A(\xi^k)^t A'(\xi^k)}$ is the estimated error obtained in each iteration $k \geq 1$. Table 1. shows the numerical results obtained when we start with $\xi^0 = \underbrace{(2, 2, \dots, 2)}_{n\ times}$, and we vary the size of the matrix and the number of iterations.

The goal of this work is to show that our sequence is better in practice compared to the Newton-Kantorovich method . We have insisted on the difficulty of constructing the Newton-Kantorovich sequence in practice, and this is due to the computation of the inverse. To show the efficiency of our method, we take $n = 100$, $k = 100$ and $\xi^0 = \underbrace{(a, a, .., a)}_{n\ times}$. Table 2. shows the efficiency of our method compared to Newton-Kantorovich metohd when we vary a. NaN(Not a Number) represents in our machine a value that is not a number, like $\frac{0}{0}$ or $\frac{\text{inf}}{\text{inf}}$. In our example NaN means that the matrix $A'(\xi^k)$ is coming close to a no invertible matrix, which makes it impossible to reverse numerically.

Iterations	n=30		n=100	
	$Er(k)$	CPU time	$Er(k)$	CPU time
k=10	1.1814e-02	0.009369	5.5894e-03	0.039737
k=20	2.7653e-04	0.015863	3.2229e-04	0.084274
k=50	4.0849e-09	0.031959	6.6281e-08	0.201700
k=100	6.7793e-16	0.057180	4.7482e-14	0.291275
k=500	6.7793e-16	0.204318	1.2784e-15	1.023639

Table 1. Numerical results: Dimension vs Iteration vs Error vs Execution time.

a	NL		NK	
	$Er(k)$	CPU time	$Er(k)$	CPU time
1.01	1.2784e-15	0.311879	1.2748e-15	0.312116
1.5	6.1674e-14	0.271268	NaN	0.900919
2	4.7482e-14	0.234149	NaN	0.381291
4	1.7605e-13	0.214328	NaN	0.351824
6	8.8911e-14	0.237603	NaN	0.533779
10	3.3543e-13	0.225355	NaN	0.479043

Table 2. Starting point vs Newton-like (NL) vs Newton-Kantrovich(NK)

4. Conclusion

We have succeeded in showing that replacing the inverse by the adjoint and adding other hypothesis in the Newton-Kantorovich method, the sequence remains convergent towards the exact solution. This result makes the application of this method more convenient and easier to apply. The numerical results show the efficiency of our method especially on the plan of the starting point .

References

[1] M. Ahues, *A note on perturbed fixed slope iterations*, Appl. Math. Letters **18** (2005), no. 4, 375–380.

[2] I. K. Argyros and Á. A. Magreńán, *Extended convergence results for the newton-kantorovich iteration*, J. of Comp. and Appl. Math **286** (2015), no. 3, 54–67.

[3] F. Cianciaruso and E. De Pascale, *Estimates of majorizing sequences in the newton kantorovich method: A further improvement*, J. of Math. Anal. and Appl. **322** (2006), no. 1, 329–335.

[4] Z. K. Eshkuvatov, H. H. Hameed, and N. M. A. Nik Long, *One dimensional nonlinear integral operator with newton kantorovich method*, J. of King Saud University - Science **28** (2016), no. 2, 172–177.

[5] A. J. Ezquerro, V. Hernández, and M. Angel, *Newton's Method: an Updated Approach of Kantorovich's Theory*, Birkhäuser, 2017.

[6] L. Grammont, M. Ahues, and F. D. D'Almeida, *For nonlinear infinite dimensional equations, which to begin with: Linearization or discretization?*, J. Integral Equations Applications **26** (2014), no. 3, 413–436.

[7] L. V. Kantorovich and G. P. Akilov, *Functional Analysis*, Pergamon Press, 1982.

[8] L. V. Kantorovich and V. I. Krilov, *Approximate Methods of Higher Analysis*, Interscience Publishers Inc, 1958.

[9] A. Khellaf, W. Merchela, and S. Benarab, *New numerical process solving nonlinear infinite-dimensional equations*, Comp. Appl. Math. **39** (2020), no. 93.

(Recibido en noviembre de 2020. Aceptado en noviembre de 2021)

Laboratoire des Mathématiques Appliquées et Modélisation
Université 8 Mai 1945 Guelma, B.P. 401, Guelma 24000, Algeria
e-mail: nada.bouazila@gmail.com; bouazila.nada@univ-guelma.dz
e-mail: guebaihamza@yahoo.fr; guebbai.hamza@univ-guelma.dz

Functional Analysis Department
Derzhavin Tambov State University, 33 Internatsionalnaya St.,
Tambov 392000, Russia
e-mail: merchela.wassim@gmail.com