数値解析 28－6
（1989. 3. 14）

非線形方程式に対する存在・非存在定理と代数方程式への応用

山本 哲朗　陳 小君
（愛媛大 理）

本稿は 非線形方程式に対する成分毎存在・非存在定理を与える。次に この定理の応用として 代数方程式に対する根のGerschgorin型存在・非存在円板を決定する。我々の結果はAlefeld[1],Smith[7],Yamamoto[9-11,13]等の結果を改良することを数値例により示す。

An Existence and Nonexistence Theorem for Solutions of Nonlinear Equations and Its Application to Algebraic Polynomials

Yamamoto Tetsuro　　Chen Xiaojun

Department of Mathematics, Faculty of Science, Ehime University
Matsuyama 790, Japan

This paper gives a componentwise existence and nonexistence theorem for a solution of nonlinear equations. The theorem is applied to algebraic polynomials,and existence and nonexistence circular regions for the zeros of the polynomial are obtained. The results improve those of Alefeld [1], Smith[7] and Yamamoto[9-11,13] for distinct approximations. Finally, the results are illustrated with several examples.

1. Introduction

There are many existence theorems for solutions of nonlinear equations which are applicable to componentwice error estimates of an approximate solution obtained by some method(cf.Ortega and Rheinboldt[5],Schroder[6], Urabe[8],etc.). Among others, in [9-13], Yamamoto obtained some related results and recently Alefeld[1] generalized a result of[13].

In this paper, in §3 after preliminary section(§2), we shall give an existence and nonexistence theorem for a solution of the equation under Kantorovich type assumptions which are weaker than those of [1] and [9-11, 13].The results seems to be new and sharper.

Next, in §4, we shall apply our results to algebraic polynomials to obtain a Gerschgorin-type existence theorem of solutions under computationally verifiable conditions. The assumptions are stronger than Smith's theorem[7]. However, our result guarantees existence of a solution in each of n circular disks D_i, $i=1,2,\ldots,n$, while Smith's theorem only asserts that any connected component of the union of n circular regions Γ_i, consisting just m disks, contains exactly m zeros.

Finally, in §5, our results will be illustrated with numerical examples.

2. Preliminaries

Throughout this paper, according to Schroder[5], Urabe [8] and Yamamoto[9,13], we use the following notation and definitions.

Let

$x=(x_i), y=(y_i)\in R^n$, $A=(a_{ij}), B=(b_{ij})\in R^{n\times n}$, $H=(h_{ijk})\in R^{n\times n\times n}$

where $H=(h_{ijk})$ is a third order tensor(a bilinear operator). We define

$$\nu[x]=(|x_i|),\quad \rho(x,y)=\nu[x-y],$$

$$\nu[A]=(|a_{ij}|),\quad \rho(A,B)=\nu[A-B],\quad \nu[H]=(|h_{ijk}|).$$

We write

$$x \geq y \quad \text{or} \quad y \leq x, \quad \text{if } x_i \geq y_i, \quad i=1,2,\ldots,n$$

$$A \geq B \quad \text{or} \quad B \leq A, \quad \text{if } a_{ij} \geq b_{ij}, \quad i,j=1,2,\ldots,n$$

$$H \geq 0 \quad \text{or} \quad 0 \leq H, \quad \text{if } h_{ijk} \geq 0, \quad i,j,k=1,2,\ldots,n.$$

For a nonnegative vector $v \geq 0$, we put

$$U(x^{(0)},v)=\{x \in R^n \mid \rho(x^{(0)},x) \leq v\}.$$

In [11], Yamamoto gave the following relations for a matrix $K=(\kappa_{ij})$ and a third order tensor $H=(h_{ijk})$:

$$Ku \leq \|u\|_\infty \kappa_1, \qquad Hu^2 \leq \|u\|_\infty h_1$$

and

$$Ku \leq \|u\|_1 \kappa_\infty, \qquad Hu^2 \leq \|u\|_1 h_\infty,$$

where

$$\|u\|_\infty=\max_i |u_i|, \quad \kappa_1=(\sum_{j=1}^{n} k_{ij}), \quad h_1=(\sum_{j=1}^{n} \sum_{k=1}^{n} h_{ijk})$$

$$\|u\|_1=\sum_{i=1}^{n} |u_i|, \quad \kappa_\infty=(\max_j k_{ij}), \quad h_\infty=(\max_{j,k} h_{ijk}).$$

Alefeld generalized those and obtained the following

Lemma 2.1 Let $u=(u_i) \in R^n$, $u \geq 0$, $K=(\kappa_{ij}) \in R^{n\times n}$, $K \geq 0$, $H=(h_{ijk}) \in R^{n\times n\times n}$, and $H \geq 0$. Then it hold for $p>1$, $q>1$, $q^{-1}+p^{-1}=1$, that

$$Ku \leq \|u\|_p \kappa_q, \qquad Hu^2 \leq \|u\|_p h_q, \tag{2.1}$$

where

$$\|u\|_p=(\sum_{i=1}^{n} |u_i|^p)^{\frac{1}{p}}$$

$$\kappa_q=((\sum_{j=1}^{n} \kappa_{ij}^q)^{\frac{1}{q}}) \in R^n,$$

and

$$h_q=\Big(\big(\sum_{j=1}^{n}\sum_{k=1}^{n} h_{ijk}^q\big)^{\frac{1}{q}}\Big) \in R^n.$$

The inequalities (2.1) hold for $p=\infty$ and $q=1$, or $p=1$ and $q=\infty$, where $\kappa_\infty=(\max_{1\le j\le n} k_{ij})\in R^n$ and $h_\infty=(\max_{1\le j,k\le n} h_{ijk})\in R^n$

Remark 2.1 From Lemma 2.1, we have

$$\|K\|_p=\max_{\|x\|_p=1}\|Kx\|_p\le\|\kappa_q\|_p,$$

and

$$\|H\|_p=\max_{\|x\|_p=\|y\|_p=1}\|Hxy\|_p\le\|h_q\|_p.$$

Furthermore, we have

$$\|K\|_1=\max_j\sum_{i=1}^{n}|k_{ij}|\le\|\kappa_\infty\|_1,\quad \|H\|_1=\max_{j,k}\sum_{i=1}^{n}|h_{ijk}|\le\|h_\infty\|_1,$$

$$\|K\|_\infty=\max_i\sum_{j=1}^{n}|k_{ij}|=\|\kappa_1\|_\infty,\quad \|H\|_\infty\le\max_i\sum_{j=1}^{n}\sum_{k=1}^{n}|h_{ijk}|\le\|h_1\|_\infty$$

3. An Existence and nonexistence Theorem for Solution

We consider the nonlinear equation

$$f(x)=0,\qquad x\subset D\in R^n \tag{3.1}$$

where f is Frechet differentiable. We assume that A is a nonsingular matrix which approximates $f'(x^{(0)})$, $H\ge0$ is a third order tensor, and for any $x,y\in D$

$$\rho(A^{-1}f'(x),A^{-1}f'(y))\le H\rho(x,y),\quad x,y\in D.$$

We put

$$K=\nu[A^{-1}(f'(x^{(0)})-A)],\quad \varepsilon=\nu[A^{-1}f(x^{(0)})].$$

$$m=\|K\|_p,\ \gamma=\|H\|_p,\ \eta=\|\varepsilon\|_p\ .$$

Theorem 3.1 Under the above notation and definitions, we have the following:

(i) If

$$m<1,\quad d=(1-m)^2-2\gamma\eta\ge0, \tag{3.2}$$

and $S(x^{(0)},t^*)=\{x\in R^n \mid \|x^{(0)}-x\|_p\le t^*\}\subset D$, where $t^*=2\eta/(1-m+\sqrt{d})$.

Then there exists a solution x^* of (3.1) in $U(x^{(0)},u)$, where

$$u=\varepsilon+t^*\kappa_q+\frac{1}{2}t^{*2}h_q,$$

which is unique in

$$\tilde{U}=\begin{cases} U(x^{(0)},\tilde{u}), & \text{if } d=0 \\ \text{int } U(x^{(0)},\tilde{u})\cap D, & \text{if } d>0 \end{cases}$$

where

$$\tilde{u}=\varepsilon+t^{**}\kappa_q+\frac{1}{2}t^{**2}h_q.$$

and $t^{**}=(1-m+\sqrt{d})/\gamma$.

(ii) There is no solution in int $U(x^{(0)},v)$, if $x^{(0)}\ne x^*$, where

$$v=\frac{2\alpha}{1+m+\sqrt{1+m+2\gamma\alpha}}(1,1,\ldots 1)^t$$

$\alpha=\|\varepsilon\|_\infty$, $m=\|K\|_\infty$ and $\gamma=\|H\|_\infty$.

Remark 3.1 The results of [1] and [9,13] have been derived under stronger conditions

$$\|\kappa_q\|_p<1 \quad \text{and} \quad (1-\|\kappa_q\|_p)^2-2\|h_q\|_p\|\varepsilon\|_p\ge 0. \qquad (3.3)$$

than (3.2). Therefore it follows from Remark 2.1 that our existence domain improves the ones obtained in [1],[9,13].

Remark 3.2 The nonexistence domain of Theorem 3.1 improves Alefeld's one. In fact, he assumed that $\delta=\min_i \varepsilon_i >0$ and gave nonexistence domain $U(x^{(0)},\beta)$ of solution, where

$$\beta=\frac{2\delta}{1+\|\kappa_1\|_\infty+\sqrt{(1+\|\kappa_1\|_\infty)^2+2\|h_1\|_\infty\delta}}(1,1,\ldots 1)^t$$

Clealy we have $\beta\le v$.

4. An Existence and Nonexistence Theorem for Zeros of Polynomials

In this section, we consider the nth-degree polynomial

$$P(z)=z^n+a_1z^{n-1}+ \dots +a_n = \prod_{i=1}^{n} (z-z_i^*) \qquad (4.1)$$

with complex coefficients a_i. To find all zeros z_i^* $(i=1,2,\dots n)$ of the polynomial, Durand [2] and Kerner[4] considered a simultaneous iteration process

$$x_i^{(k)}=x_i^{(k-1)}-\frac{P(x_i^{(k-1)})}{\prod_{j\neq i}^{n}(x_i^{(k-1)}-x_j^{(k-1)})} \qquad \begin{matrix} i=1,2,\dots n \\ k=1,2,\dots . \end{matrix} \qquad (4.2)$$

where $x_1^{(0)},x_2^{(0)},\dots,x_n^{(0)} \in C^n$ and $x_i^{(0)}\neq x_j^{(0)}$ $(i\neq j)$.
This process is called Durand-Kerner's method or D-K method, although Weierstrass also considered this.

In [4], Kerner showed that the D-K method is equivalnt to Newton's method applied to the equation

$$f(x)=(f_1(x),\ f_2(x),\dots,f_n(x))^t=0, \qquad x\in C^n \qquad (4.3)$$

where $f_i(x)=b_i(x)-a_i$, $b_i(x)=(-1)^i \sum_{j_1<\dots<j_i} x_{j_1}\dots x_{j_i}$, $i=1,2,\dots,n$.

In [9],Yamamoto gave n circular regions,each of which contains at least one zero of the polynomial. In this section, as an application of Theorem 3.1, we shall give new existence circular regions under a Kantorovich-type condition and nonexistence circular regions with nocondition.The radius of the former regions are in general smaller than those of Smith and Yamamoto.
We put

$$\varepsilon_i=|x_i^{(1)}-x_i^{(0)}|, \qquad \eta=\sum_{i=1}^{n}\varepsilon_i$$

$$\beta_{ij}=|x_i^{(0)}-x_j^{(0)}|^{-1}, \qquad \beta=\max_{i,j}\beta_{ij}$$

and

$$a=\eta\beta .$$

Let $H=(h_{ijk})$ be defined as follows

$$h_{ijk}=\begin{cases}\beta(1+\frac{2\beta\eta}{n-2})^{n-2}, & \text{if } j\neq k, i=j \text{ or } j\neq k,\ i=k\\ \beta^2(1+\frac{2\beta\eta}{n-2})^{n-2}, & \text{if } j\neq k, i\neq k, i\neq k\\ 0 & \text{if } j=k.\end{cases}$$

Furthermore, we put

$$h=\begin{cases}2a^2(1+2a/(n-2))^{n-3}+2a(1+2a/(n-2)^{n-2}, & n\geq 3\\ 2a, \quad n=2\end{cases}$$

and

$$\gamma=\|H\|_1=\max_{j,k}\sum_{i=1}^{n}h_{ijk}=\begin{cases}2\beta(1+\frac{2\beta\eta}{n-2})^{n-2}+2\eta\beta^2(1+\frac{2\beta\eta}{n-2})^{n-3} & , n\geq 3\\ 2\beta \quad , & n=2.\end{cases}$$

Theorem 4.1 (i)If $h\leq 1/2$, then we have the following:

(1) In each closed disk

$$\Gamma_i\ :\ |z-z_i^{(0)}|\leq \gamma_i=\varepsilon_i+\frac{1}{2}t^{*2}(h_\infty)_i$$

there exists at least one zero $z_i^{(0)}$ of $P(z)$.

(2) Any connected component of the union of the disk

$$\tilde{\Gamma}_i:\ |z-z^{(0)}|=\tilde{\gamma}_i=\varepsilon_i+\frac{1}{2}t^{**2}(h_\infty)_i$$

where $t^{**}=\dfrac{1+\sqrt{1-2h}}{\gamma}$ and $h_\infty=(\max_{i,j} h_{ijk})\in R^n$, consisting of just m disks contains exactly m zeros of $P(z)$.

Furthermore, if $h<1/2$, then all the zeros of $P(z)$ are simple.

(ii) In each open disk

$$D_i:\ |z-z_i^{(0)}|<\frac{2\alpha}{1+\sqrt{1+2(n-1)\gamma\alpha}}$$

there is no zeros of $P(z)$, where $\alpha=\max_i \varepsilon_i$.

5.Numerical Examplex

In this section, we shall illustrate our results with some examples.

Examplex 1([11][1]). We first consider the algebraic eigenvalue problem

$$Ty=\lambda y \tag{5.1}$$

where

$$T=\begin{pmatrix} 1 & 1 & 0.5 \\ 1 & 1 & 0.25 \\ 0.5 & 0.25 & 2 \end{pmatrix}$$

and λ and y are eigenvalue and eigenvector of T, respectively. We assume that $y=(y_i)\in R^n$ satisfies $\|y\|_2^2=\sum_{i=1}^{n}|y_i|^2=1$. Let $x=(y_1,...,y_n,\lambda)$. Then (5.1) can be written in the form

$$f(x)=\begin{pmatrix} (T-\lambda I)y \\ (1-\|y\|_2^2)/2 \end{pmatrix}=0$$

Furthermore, we have

$$f'(x)=\begin{pmatrix} T-\lambda I & -y \\ -y^t & 0 \end{pmatrix},$$

$$f''(x)=\left(\begin{matrix} 0,..,0 & -1 \\ & 0 \\ 0 & \\ -1,0,..,0 & \end{matrix}\right| \quad \cdots \quad \left|\begin{matrix} 0,..,0 & 0 \\ & 0 \\ 0,..,-1 & -1\ 0 \end{matrix}\right| \left.\begin{matrix} -1 & & \\ & \ddots & \\ & & -1 \\ & & & 0 \end{matrix}\right).$$

We choose an initial approximation $x^{(0)}=(-0.7,0.7,0.15,0.0)$ and put $A=f'(x^{(0)})$. Then we have

$K=0,\quad H=\nu[f'(x^{(0)})^{-1}f''(x^{(0)})],\quad \varepsilon=\nu[f'(x^{(0)})^{-1}f(x^{(0)})]$.

Furthermore, we choose $p=1, q=\infty$. Then we obtain $\|H\|_1\|\varepsilon\|_1=0.4074514<1/2$. We can thus conclude that there exists a solution x^* in $U(x^{(0)},u)$, where

$u=(\ 0.0475545,\ 0.0293976,\ 0.1691796,\ 0.0271919)$

which is unique in $\tilde{U}(x^{(0)},\tilde{u})$, where

$\tilde{u}=(\ 0.0934822,\ 0.0491608,\ 0.2014925,\ 0.0676589$

It is interseting to remark that $\|h_1\|_\infty\|\varepsilon\|_\infty=0.7518750$ and $\|h_\infty\|_1\|\varepsilon\|_1=0.5163171$, so that the conditions of Corollary

3.2 in [1] and Theorem 1 in [13] are not satisifed. Hence, existence of a solution can not be guaranteed by results of [1][13].

Example 2([9]).Consider the polynomial

$$P(z)=z^5-10z^4+43z^3-104z^2+150z-100=(z-2)(z^2-2z+5)(z^2-6z+10)$$

We choose approximate roots of P(z)

z_1=(3.00010,1.00010), z_2=(3.00000,-0.99990), z_3=(0.99990, 2.00000), z_4=(0.99990,-2.00000), z_5=(2.00010,0.00000)

Smith's radii of circular regions containing all the zeros of P(z) are

r_1	r_2	r_3	r_4	r_5
0.0018920	0.0013537	0.0013401	0.0013398	0.0011174

Since h=0.0019982<1/2, we can use Theorem 4.1 to give improved radii

d_1	d_2	d_3	d_4	d_5
0.0003791	0.0002714	0.0002687	0.0002687	0.0002242

References

[1]G.Alefeld, Componentwise inclusion and exclusion sets for solutions of quadratic equations in finite dimensional spaces, Numer.Math.48(1986),391-416.

[2]E.Durand, Solutions Numeriques des Equations Algebriques. Paris:Masson 1960.

[3]W.B.Gragg and R.A.Tapia, Optimal error bounds for the Newton-Kantorovich theorem. SIAM J.Numer.Anal.11(1974) 10-13.

[4]I.Kerner, Ein Gesamtschrittverfahren zur berechnung der nullstellen von polynomen,Num.Math.8(1966),290-294.

[5]J.M.Ortega and W.C.Rheinboldt,Iterative Solution of Nonlinear Equation in Several Variables. Academic Press, New York 1970.

{6]J.Schroder,Nichtlineare majoranten beim verfahren der schrittweisen naherung, Arch.Math. 7(1956),471-484.

[7]B.T.Smith, Error bounds for zeros of a polynomial based

upon Gerschgorin's theorem,JA.C.M. 17(1970),661-674.
[8]M.Urabe,A posteriori componentwise error estimation of approximate solution to nonlinear equations, Lect.Notes in Comp.Sci.29,Springer(1975)
[9]T.Yamamoto, An existence theorem for solution of nonlinear systems and its application to algebraic equation.3rd USA-JAPAN computer conference proceedings, AFIPS and IPSJ, 1978.
[10]T.Yamamoto, Componentwise error estimates for approximate solutions of nonlinear equations,J.Inf. Processing2(1979),121-126.
[11]T.Yamamoto, Error bounds for computed eigenvalues and eigenvectors, Numer.Math.34(1980),189-199.
[12]T.Yamamoto,Error bounds for computed eigenvalues and eigenvectors. ,Numer.Math.40(1982),201-206.
[13]T.Yamamoto, Error bounds for approximate solutions of systems of equations, Japan J. of Appl. Math. 1(1984) 157-171.