数理解析研究所講究録
42巻 1968年 53-70

Numerical Solution of Multi-Point Boundary-Value Problems in Chebyshev Series-Theory of the Method

Minoru URABE

1. Introduction

In the recent paper

Urabe, M., Numerical solution of multi-point boundary value problems in Chebyshev series-Theory of the method, Numer. Math., 9(1967), 341-366,

the writer gave the mathematical certification to the method to compute solutions of boundary-value problems in Chebyshev series. Such a method had been proposed by Clenshaw and Norton in their papers:

Clenshaw, C.W. and Norton, H.J., The solution of nonlinear ordinary differential equations in Chebyshev series, Comput. J., 6(1963), 88-92,

Norton, H.J., The iterative solution of non-linear ordinary differential equations in Chebyshev series, Comput. J., 7(1964), 78-85.

By several numerical examples, they showed that such a method is of much use for practical computations. But they did not give mathematical certification to such a method. In the present note, the writer will sketch his approach and thereby he will show how such a method can be certified mathematically.

Consider the boundary value problem of the following form:

$$\frac{dx}{dt} = X(x,t) \ , \tag{1.1}$$

$$\sum_{i=0}^{N} L_i x(t_i) = \ell \ . \tag{1.2}$$

In (1.1), x and $X(x,t)$ are vectors and $X(x,t)$ is defined in the region D of the tx-space intercepted by two hyperplanes $t=-1$ and $t=1$. In (1.2),

$$-1 = t_0 < t_1 < t_2 < \cdots < t_{N-1} < t_N = 1 \ ,$$

L_i are the given square matrices and ℓ is a given vector. Our boundary condition is very general and so our boundary-value problem includes the Cauchy problem and the ordinary two-point boundary-value problem.

To seek an approximate solution of our boundary-value problem, we consider a finite Chebyshev series

$$x_m(t) = a_0 + \sqrt{2}\sum_{n=1}^{m} a_n T_n(t) \tag{1.3}$$

with undetermined coefficients $a_0, a_1, a_2, \cdots, a_m$ and we determine these unknown coefficients so that

$$\sum_{i=0}^{N} L_i x_m(t_i) = \ell \ , \tag{1.4}$$

$$\frac{dx_m(t)}{dt} = P_{m-1} X[x_m(t), t] \tag{1.5}$$

may hold. Here P_{m-1} is an operator which expresses the truncation of the Chebyshev series of the operand discarding the terms of the order higher than $m-1$. The factor $\sqrt{2}$ is introduced in an Chebyshev series merely for simplication of theoretical approach. If one can determine $x_m(t)$ so that it may satisfy (1.4) and (1.5), then it is expected that $x_m(t)$ determined in such a way will be a reasonable approximation of a solution of our boundary-value problem. In what follows, we shall call a finite Chebyshev series $x_m(t)$ satisfying (1.4) and (1.5) the Chebyshev approximation of the order m.

In our approach, we restrict the solution of (1.1) satisfying (1.2) to the one for which matrix $G=\sum_{i=0}^{N} L_i\Phi(t_i)$ is nonsingular where $\Phi(t)$ is the fundamental matrix of the first variation equation of (1.1) satisfying the initial condition $\Phi(-1)=E$ (E is a unit matrix). We shall call such a solution the isolated solution, since in a sufficiently small seighborhood of such a solution, there is no other solution satisfying (1.2).

In our approach, for vectors, we use Euclidean norms and for matrices, we use norms corresponding to Euclidean norms of vectors in a way of functional analysis. These norms will be denoted by $\|\cdots\|$. For continuous vector-functions defined on the interval $J[-1,1]$, we use two kinds of norms, namely, $\|\cdots\|_n$ and $\|\cdots\|_q$. Namely let $f(t)$ be an arbitrary continuous vector-function defined on J, then

$$\|f(t)\|_n \quad \text{means} \quad \|f(t)\|_n = \sup_{t\in J} \|f(t)\|$$

and

$$\|f(t)\|_q \quad \text{means} \quad \|f(t)\|_q = \left[\frac{1}{\pi}\int_0^{\pi} \|f(\cos\theta)\|^2 d\theta\right]^{1/2}$$

$$= [\frac{1}{\pi}\int_{-1}^{1}\frac{\|f(t)\|^2}{\sqrt{1-t^2}}\,dt]^{1/2}.$$

2. Basic propositions

Our approach is based on the following three propositions.

Proposition 1. Let

$$F(\alpha) = 0 \tag{2.1}$$

be a given real system of equations, where α and $F(\alpha)$ are vectors of the same dimension and $F(\alpha)$ is continuously differentiable with respect to α in some region Ω of the α-space,

Assume that (2.1) has an approximate solution $\alpha = \hat{\alpha}$ for which the determinant of the Jacobian matrix $J(\alpha)$ of $F(\alpha)$ does not vanish and that there are a positive constant δ and a non-negative constant $\kappa<1$ such that

(2.2)
(i) $\Omega_\delta = \{\alpha \mid \|\alpha-\hat{\alpha}\| \leqq \delta\} \subset \Omega$,

(ii) $\|J(\alpha) - J(\hat{\alpha})\| \leqq \frac{\kappa}{M'}$ for any $\alpha \in \Omega_\delta$,

(iii) $\frac{M'r}{1-\kappa} \leqq \delta$,

where r and $M'(>0)$ are numbers such that

$$\|F(\hat{\alpha})\| \leqq r \quad \text{and} \quad \|J^{-1}(\hat{\alpha})\| \leqq M'.$$

Then system (2.1) has one and only one solution $\alpha = \bar{\alpha}$ in Ω_δ, and, for $\alpha = \bar{\alpha}$, it holds that

(2.3) $\det J(\alpha) \neq 0,$

(2.4) $\| \bar{\alpha} - \hat{\alpha} \| \leq \dfrac{M'r}{1-\kappa}$.

This proposition can be proved by means of the Newton iterative process:

$$\alpha_{n+1} = \alpha_n - J^{-1}(\hat{\alpha})F(\alpha_n) \qquad (n = 0,\ 1,\ 2,\ldots)$$

where $\alpha_0=\hat{\alpha}$. The condition (2.2) is related with the accuracy of the approximate solution $\hat{\alpha}$. By Propostion 1, if we know a certain accurate approximate solution, we can assert the existence of an exact solution and further we can get an error estimate for such an approximate solution.

Next proposition is concerned with the boundary-value problem for linear differential systems.

Proposition 2. Let

(2.5) $$\frac{dx}{dt} = A(t)x + \varphi(t)$$

be a given linear differential system, where $A(t)$ is a matrix continuous on J and $\varphi(t)$ is a vector continuous on J. Let $\Phi(t)$ be the fundemental matrix of the corresponding homogeneous system

(2.6) $$\frac{dy}{dt} = A(t)y$$

satisfying the initial condition $\Phi(-1)=E$. Then, if matrix $G=\sum_{i=0}^{N} L_i\Phi(t_i)$ is non-singular for

$$-1 = t_0 < t_1 < \cdots < t_{N-1} < t_N = 1$$

and given square matrices L_i, the given system (2.5) has one and only one solution satisfying the boundary condition

$$\sum_{i=0}^{N} L_i x(t_i) = \ell \tag{2.7}$$

for an arbitrary vector ℓ and such a solution $x=x(t)$ is given in the form

$$x(t) = \Phi(t)G^{-1}\ell + \int_{-1}^{1} H(t,s)\varphi(s)ds , \tag{2.8}$$

where $H(t,s)$ is the piece-wise continuous Green matrix corresponding to the homogeneous boundary condition

$$\sum_{i=0}^{N} L_i x(t_i) = 0 . \tag{2.9}$$

This proposition can be easily proved by means of the method of variation of constants and the explicit formula for $H(t,s)$ can be easily obtained. The matrix $H(t,s)$ is dependent only on $\Phi(t)$, in other words, it depends only on $A(t)$.

If we put

$$\psi(t) = \int_{-1}^{1} H(t,s)\varphi(s)ds, \tag{2.10}$$

then $x=\psi(t)$ satisfies the given differential system (2.5) and the homogeneous boundary condition (2.9). Since $\psi(t)$ is continuous, the formula (2.10) expresses a linear mapping in the space of continuous vector-functions. We shall call such a mapping the H-mapping corresponding to matrix $A(t)$. It is evident that for the H-mapping there are two kinds of norms corresponding to two kinds of norms of continuous vector-functions.

The last proposition is of the same character as Proposition 1. The difference is that Proposition 1 is concerned with a system of finite equations while the last proposition is concerned with a system of differential equations.

Proposition 3. In differential system (1.1), we suppose that $X(x,t)$ is continuously differentiable with respect to x in the region D of the tx-space intercepted by two hyperplanes $t=-1$ and $t=1$ and that (1.1) has an approximate solution $x=\bar{x}(t)$ lying in D and satisfying the boundary condition (1.2) approximately.

For approximate solution $x=\bar{x}(t)$, assume that

(2.11) $$\left\| \frac{d\bar{x}(t)}{dt} - X[\bar{x}(t),t] \right\| \leq r \quad \text{on} \quad J = [-1,1] \; ;$$

(2.12) $$\left\| \sum_{i=0}^{N} L_i \bar{x}(t_i) - \ell \right\| \leq \varepsilon;$$

there is a matrix $A(t)$ such that matrix $G=\sum_{i=0}^{N} L_i \Phi(t_i)$ is non-singular where $\Phi(t)$ is the fundamental matrix of the linear homogeneous system

(2.13) $$\frac{dy}{dt} = A(t)y$$

satisfying the initial condition $\Phi(-1)=E$;

there are a positive constant δ *and a non-negative constant* $\kappa < 1$ *such that*

(i) $D_\delta = \{ (t,x) \mid \|x - \bar{x}(t)\| \leqq \delta,\ t \in J \} \subset D,$

(ii) $\| \Psi(x,t) - A(t) \| \leqq \dfrac{\kappa}{M_1}$ *for any* $(t,x) \in D_\delta,$

where $\Psi(x,t)$ *is the Jacobian matrix of* $X(x,t)$ *with respect to* x *and* M_1 *is a positive constant such that*

(2.14)
$$\| H \|_n \leqq M_1$$

where H *is the H-mapping corresponding to* $A(t)$;

(iii) $$\frac{M_1 r + M_2 \varepsilon}{1 - \kappa} \leqq \delta$$

where M_2 *is a non-negative constant such that*

$$\| \Phi(t) G^{-1} \| \leqq M_2 \quad \textit{on} \quad J.$$

Then, in D_δ, *the given system* (1.1) *has one and only one solution* $x = \hat{x}(t)$ *satisfying the boundary condition* (1.2), *and this is an isolated solution and, moreover, for* $x = \hat{x}(t)$, *we have*

(2.15)
$$\| \bar{x}(t) - \hat{x}(t) \| \leqq \frac{M_1 r + M_2 \varepsilon}{1 - \kappa}.$$

This proposition can be proved by means of the iterative process:

$$(2.16)\qquad x_{n+1}(t) = \Phi(t)G^{-1}\ell + \int_{-1}^{1} H(t,s)\{X[x_n(s),s]-A(s)x_n(s)\}ds$$

$$(n = 0,\ 1,\ 2,\cdots),$$

where $x_0(t)=\bar{x}(t)$ and $H(t,s)$ is the matrix of the H-mapping H corresponding to $A(t)$. The iterative process (2.16) comes from the idea of the Newton method. the condition (2.14) is related with the accuracy of the approximate solution $x=\bar{x}(t)$ and inequality (2.15) gives an error bound for the given approximate solution.

3. Main theorems

Theorem 1. Suppose $X(x,t) \in C^2_{x,t}[D]$ and system (1.1) has an isolated solution $x = \hat{x}(t)$ satisfying (1.2) and the internality condition

(3.1) $U = \{ (t,x) \mid \| x - \hat{x}(t) \| \leqq \delta_0, \quad t \in J \} \subset D$ for some $\delta_0 > 0$.

Then for sufficiently large m_0, there is a Chebyshev approximation $x = \bar{x}_m(t)$ of any order $m \geqq m_0$ such that $\bar{x}_m(t)$ converges to $\hat{x}(t)$ uniformly together with its first order derivative as $m \to \infty$.

This theorem says that, for any isolated solution satisfying (3.1), we can always get its approximation as accurately as we desire by computing the Chebyshev approximation.

Theorem 2. Under the assumptions of Theorem 1, the Chebyshev approximation $x = \bar{x}_m(t)$ is determined uniquely in a sufficiently small neighborhood of $x=\hat{x}(t)$ provided the order m is sufficiently large.

This theorem says that the obtained Chebyshev approximation corresponds one-to-one to the isolated solution satisfying (3.1).

Theorem 3. In Theorem 1, the existence of the initial isolated solution $x = \hat{x}(t)$ can be assured by applying Proposition 3 to a computed Chebyshev approximation $x = \bar{x}_m(t)$ of the order sufficiently high. In the application of Proposition 3, the error bound of the computed Chebyshev approximation $x=x_m(t)$ can be also obtained.

In most practical problems, the existence of an isolated solution satisfying the given boundary condition is not known beforehand and moreover it is usually not an easy work to prove analytically the existence of an exact solution. In addition, even when we get Chebyshev approximation by computation, we are not sure about the existence of an exact solution.

Hence Proposition 3 plays an important role in practical applications, because it enables us to assure the existence of an exact solution from a computed approximate solution and furthermore it enables us to get an error bound of the computed approximate solution.

To prove our theorems, we use some properties of Chebyshev series.

Let $f(t)$ be an arbitrary vector-function continuous on J and let its Chebyshev series be

(3.2) $$f(t) \sim a_o + \sqrt{2}\sum_{n=1}^{\infty} a_n T_n(t).$$

Then it is evident that

(3.3) $$f(\cos\theta) \sim a_o + \sqrt{2}\sum_{n=1}^{\infty} a_n \cos n\theta .$$

Hence we see that

(3.4) $$a_n = \frac{\sqrt{2}}{\pi e_n}\int_0^{\pi} f(\cos\theta)\cos n\theta\, d\theta \qquad (n=0,1,2,\cdots),$$

where $e_n = \sqrt{2}$ or 1 according as $n = 0$ or $n \geqq 1$. If we apply Parseval's equality to (3.3), we readily get

(3.5) $$\| f \|_q^2 = \sum_{n=0}^{\infty} \| a_n \|^2$$

and, hence, for a finite Chebyshev series of the form

$$f_m(t) = a_o + \sqrt{2}\sum_{n=1}^{m} a_n T_n(t),$$

we have

$$(3.6) \qquad \| f_m \|_q = \|\alpha\|,$$

where

$$\alpha = \mathrm{col}(a_o, a_1, \cdots, a_m).$$

When $f(t) \in C_t'$, let the Chebyshev series of $\dot{f}(t) = \frac{d}{dt}f(t)$ be

$$(3.7) \qquad \dot{f}(t) \sim a_o' + \sqrt{2}\sum_{n=1}^{\infty} a_n' T_n(t),$$

then we can prove that

$$(3.8) \qquad a_n' = \frac{2}{e_n}\sum_{p=1}^{\infty}(n + 2p - 1)a_{n+2p-1} \qquad (n=0,1,2,\cdots)$$

and that

$$(3.9) \quad \begin{cases} \| (I-P_m)f \|_n \leqq \sigma(m) \| (I-P_{m-1})\dot{f} \|_q \leqq \sigma(m) \| \dot{f} \|_q, \\ \| (I-P_m)f \|_q \leqq \sigma_1(m) \| (I-P_{m-1})\dot{f} \|_q \leqq \sigma_1(m) \| \dot{f} \|_q \end{cases}$$

$$(m=0,1,2,\cdots;\ P_{-1} \equiv 0),$$

where

$$\sigma(m) = \sqrt{2}\left[\sum_{n=m+1}^{\infty}\frac{1}{n^2}\right]^{1/2} < \frac{\sqrt{2}}{\sqrt{m}},$$

$$\sigma_1(m) = \frac{1}{m+1}.$$

Formula (3.8) is frequently convenient for theoretical approach more than well-known formula

$$e_{n-1}a'_{n-1} - a'_{n+1} = 2na_n \qquad (n=1,2,\cdots).$$

From (3.9), we can easily see that

(3.10)

$$\|(I-P_m)f\|_n \leqq \frac{\sigma(m)}{m(m-1)} \|\dddot{f}\|_q,$$

$$\|(I-P_m)f\|_q \leqq \frac{1}{(m+1)m(m-1)} \|\dddot{f}\|_q \qquad (m \geqq 2)$$

if $f(t) \in C_t^3\,[J]$.

For Chebyshev series, the operation P_m and $\frac{d}{dt}$ are not commutative since $\frac{d}{dt} P_m f(t)$ is in general a polynomial of degree $m-1$ while $P_m \frac{d}{dt} f(t)$ is in general a polynomial of degree m. However, if we put

$$\frac{d}{dt} P_m f = \dot{f}_m ,$$

then, after some manipulations, we can prove that

$$\|\dot{f}-\dot{f}_m\|_q \leqq \frac{\sqrt{m+2}}{\sqrt{2}} \|(I-P_{m-1})\dot{f}\|_q + \|(I-P_{m+1})\dot{f}\|_q$$

and, if the Chebyshev series of $\dot{f}(t)$ is uniformly convergent,

$$\|\dot{f}-\dot{f}_m\|_n \leqq (m+2) \|(I-P_{m-1})\dot{f}\|_q + \|(I-P_{m+1})\dot{f}\|_n .$$

When $f(t) \in C_t^3\,[J]$, the Chebyshev series of $\dot{f}(t)$ is uniformly convergent, hence applying (3.9) twice to the right-hand side of the above inequalities, we have

$$
(3.11)\qquad
\begin{aligned}
&\| \dot{f}-\dot{f}_m \|_n \leqq \left[\frac{m+2}{m(m-1)} + \frac{\sigma(m+1)}{m+1}\right] \| \dddot{f} \|_q , \\
&\| \dot{f}-\dot{f}_m \|_q \leqq \left[\frac{\sqrt{m+2}}{\sqrt{2}m(m-1)} + \frac{2}{(m+2)(m+1)}\right] \| \dddot{f} \|_q \\
&\hfill (m \geqq 2).
\end{aligned}
$$

Now we shall sketch the proof of our theorems.

By (3.4), the equations (1.4) and (1.5) are equivalent to the following system of equations

$$
(3.12)\qquad
\begin{aligned}
&F_0(\alpha) \overset{\text{def}}{=\!=} \sum_{i=0}^{N} L_i x_m(t_i) - \ell = 0, \\
&F_n(\alpha) \overset{\text{def}}{=\!=} \frac{\sqrt{2}}{\pi e_{n-1}} X[x_m(\cos\theta), \cos\theta] \cos(n-1)\theta \, d\theta \\
&\qquad - a'_{n-1}(\alpha) = 0 \\
&\hfill (n=1, 2, \cdots, m),
\end{aligned}
$$

where $\alpha = \mathrm{col}(a_0, a_1, \cdots, a_m)$ and

$$
(3.13)\qquad a'_{n-1}(\alpha) = \frac{2}{e_{n-1}} [n a_n + (n+2) a_{n+2} + \cdots] \qquad (n=1, 2, \cdots, m)
$$

We shall call the system of equations (3.12) the determining equation of Chebyshev approximations and we shall write this briefly in a vector form as follows:

$$
(3.14)\qquad F^{(m)}(\alpha) = 0 .
$$

Now, in Theorem 1, the existence of an isolated solution $x = \hat{x}(t)$ satisfying the boundary condition (1.2) and the internality condition (3.1) is assumed. Hence, from this isolated solution $\hat{x}(t)$, we can make a finite Chebyshev series

$$P_m\hat{x}(t) = \hat{x}_m(t) = \hat{a}_0 + \sqrt{2}\sum_{n=1}^{m}\hat{a}_n T_n(t) .$$

Put $\hat{\alpha} = \mathrm{col}(\hat{a}_0, \hat{a}_1, \cdots, \hat{a}_m)$. Then, using (3.9), (3.10) and (3.11), we can prove that

(3.15) $$\| F^{(m)}(\hat{\alpha}) \| \leqq Km^{-3/2}$$

for any $m \geqq m_1$ and some K independent of m provided m_1 is sufficiently large. This shows that $\alpha = \hat{\alpha}$ is an approximate solution of (3.14) provided m is large. Hence we apply our Proposition 1 to (3.14) in order to prove the existence of an exact solution of (3.14), namely, the existence of a Chebyshev approximation.

However, to apply Proposition 1 to (3.14), we have to know some properties of the Jacobian matrix $J_m(\alpha)$ of $F_m(\alpha)$. To derive the necessary properties of $J_m(\alpha)$, we consider the linear system

(3.16) $$J_m(\alpha)\xi + \gamma = 0.$$

Corresponding to $\alpha = \mathrm{col}(a_0, a_1, \cdots, a_m)$, $\xi = \mathrm{col}(u_0, u_1, \cdots, u_m)$ and $\gamma = \mathrm{col}(c_0, c_1, \cdots, c_m)$, put

$$x_m(t) = a_0 + \sqrt{2}\sum_{n=1}^{m} a_n T_n(t),$$

$$y(t) = u_0 + \sqrt{2}\sum_{n=1}^{m} u_n T_n(t)$$

$$\varphi(t) = c_1 + \sqrt{2}\sum_{n=1}^{m-1} c_{n+1} T_n(t).$$

Then we readily see that (3.16) yields the following system of equations:

$$(3.17)\begin{cases} \sum_{i=0}^{N} L_i y(t_i) + c_o = 0 , \\ \frac{dy(t)}{dt} = P_{m-1} \left\{ \Psi[x_m(t),t]y(t) \right\} + \varphi(t). \end{cases}$$

If we rewrite the second of (3.17) in the form

$$\frac{dy(t)}{dt} = \Psi[x_m(t),t]y(t) + \left\{ \varphi(t) - (I-P_{m-1}) \Psi[x_m(t),t]y(t) \right\},$$

then this can be regarded as the linear equation with respect to $y(t)$. The first of (3.17) then gives a boundary condition for $y(t)$. Hence we apply Proposition 2 to (3.17). Then, using the corespondence between (3.16) and (3.17), we can get the necessary properties of $J_m(\alpha)$. Then we check the conditions of Proposition 1 for (3.14) and we see that all these conditions are satisfied by $x = \hat{x}_m(t)$ provided m is sufficiently large. From this, we see the existence of Chebyshev approximations.

The convergence of Chebyshev approximations together with their first order derivatives can be proved without difficulty by using the error estimation (2.4) of Proposition 1.

Theorem 2 can be easily proved if we use the properties of the solutions of the equation of the form (3.17).

Since the Chebyshev approximation $\bar{x}_m(t)$ converges to $\hat{x}(t)$ uniformly, we can prove Theorem 3 without difficulty if we use the isolatedness of $x = \hat{x}(t)$.

4. Remarks about the numerical solution of the determining equation

As well known, the Chebyshev coefficients of the known function can be easily evaluated if we use the techniques of Fourier analysis. Hence, if we apply the Newton method to the determining equation, then we can easily solve the determining equation numerically without calculating the explicit form of the determining equation. Since the convergence of Chebyshev series is usually very rapid, we can find the first approximation necessary for Newton method by solving the determining equation with very small number of unknown Chebyshev coefficients. One or two unknown coefficients are frequently sufficient. Numerical examples are shown in Norton's paper. When $X(x,t)$ is linear in x, the determining equation is linear in α and the proof of our theorem 1 implies the determinant of coefficients of unknown α does not vanish. In such a case, of course, the first approximation is not necessary for computation of Chebyshev approximations. Numerical examples based on our theory will be published elsewhere.