REMARKS ON THE SHOOTING METHOD FOR NONLINEAR TWO-POINT BOUNDARY-VALUE PROBLEMS

Nguyen Trung Hieu
Department of Mathematics, College of Science, VNU

Abstract. In this note, we prove a convergent theorem for the shooting method combining the explicit Euler's scheme with the Newton method for solving nonlinear two-point boundary problems (TPBVPs). Some illustrative numerical examples are also considered. A convergent result obtained before by T. Jankowski is a particular case of our result, when the boundary condition (BC) becomes linear.

1. Introduction

The shooting method for TPBVPs has been studied throughly in many works (c.f [1-10]). However, the convergence of the method did not receive adequate attention of researchers. In 1995, T. Jankowski gave an adequate proof for a convergent theorem of a shooting method.

In this paper, we will generalize this result for nonlinear ordinary differential equations (ODEs) with nonlinear boundary conditions.

Consider the problem

$$y' = f(t, y),\ t \in J = [a, b] \tag{1a}$$

$$\phi(y(a), y(b)) = 0, \tag{1b}$$

where $f : J \times \mathbb{R}^p \to \mathbb{R}^p$ is continuous in t and continuously differentiable in y, $\phi : \mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R}^p$ continuously differentiable in both variables.

If we denote $y = y(t; s)$ a solution of (1a) subject to initial condition

$$y(a) = s, \tag{1c}$$

then the problem is reduced to that of finding $s = \overline{s}$ which solves the equation

$$\phi(y(a; s), y(b; s)) = \phi(s, y(b; s)) = 0. \tag{2}$$

This equation can be solved approximately by the Newton's iteration

$$s_{j+1} = s_j - \phi'(s_j, y(b; s_j))^{-1}\phi(s_j, y(b; s_j)), \quad j \geq 0, \tag{3}$$

where $\phi'(s) = \phi_1(s, y(b; s)) + \phi_2(s, y(b; s)).y'_s(b; s)$ and ϕ_1, ϕ_2 are partial derivatives of ϕ with respect to the first and the second variables, respectively. In addition , $Y(t; s) = y'_s(t; s)$ can be found as a solution of the IVP

$$\begin{cases} Y'(t; s) &= f'_y(t, y(t; s)).Y(t; s) \\ Y(a; s) &= I. \end{cases} \tag{4}$$

Typeset by $\mathcal{A}_\mathcal{M}\mathcal{S}$-TEX

The remaining problem is to solve approximately the two IVPs (1 a,c) and (4). There are many methods for solving IVPs, but in this paper, we only use explicit Euler's method to solve them. Thus we obtain a method using Euler's scheme combined with Newton's iteration. In the following, this method will be discussed in detail. Let the integration interval be subdivided by the points

$$t_i = a + ih,\ h = \frac{b-a}{N}(i = \overline{0, N}). \tag{5}$$

First, applying the explicit Euler's scheme mentioned above, we get

$$y_h(t_0; s_{h,j}) := s_{h,j};\ :\ y_h(t_{i+1}; s_{h,j}) = y_h(t_i; s_{h,j}) + h.f(t_i, y_h(t_i; s_{h,j})); \tag{6a}$$

$$Y_h(t_0; s_{h,j}) := I;\ :\ Y_h(t_{i+1}, s_{h,j}) = [I + h.f_y'(t_i, y_h(t_i; s_{h,j}))]Y_h(t_i, s_{h,j}), \tag{6b}$$

$$i = 0, 1, ..., N-1,$$

$$s_{h0} = s_0, s_0 \in \mathbb{R}^p \text{ is an initial vector.}$$

Then we use the Newton method to improve the shooting vector

$$s_{h,j+1} = s_{h,j} - [\phi_1(s_{h,j}, y_h(b; s_{h,j}))$$

$$+\phi_2(s_{h,j}, y_h(b; s_{h,j}))Y_h(b, s_{h,j})]^{-1}\phi(s_{h,j}, y_h(b; s_{h,j})) \tag{6c}$$

$$j = 0, 1, ...$$

2. Convergence

In this note, we use the terminologies and notations of [4]. However, the definition of isolated solution should be stated as follows.

Definition 1. *A solution $y(t)$ of (1) is said to be isolated if the following linear TPBVP*

$$\begin{cases} z' = f_y(t, y(t))z \\ \phi_1(y(a), y(b))z(a) + \phi_2(y(a), y(b))z(b) = 0 \end{cases} \tag{7}$$

has only trivial solution $z(t) \equiv 0$.

Lemma 1. *The isolated solution is locally unique.*

Proof. Consider the space $C^1(J, \mathbb{R}^p)$ equipped with the norm $\|y\| = \max\{\|y\|_{\max}, \|y'\|_{\max}\}$, the space $C(J, \mathbb{R}^p) \times \mathbb{R}^p$ equipped with the norm $\|(\overline{x}, x_2)\| = \|\overline{x}\|_{\max} + |x_2|$, where $\|x\|_{\max} = \max_{t \in J}|x(t)|$, $|.|$ is the Euclidian norm and the mapping

$$F:\quad y \longrightarrow \begin{bmatrix} y' - f(t, y) \\ \phi(y(a), y(b)) \end{bmatrix},\ t \in J.$$

Clearly,

$$F'(y)h = \begin{bmatrix} h' - \frac{\partial f}{\partial y}(t, y)h \\ \phi_1(y(a), y(b))h(a) + \phi_2(y(a), y(b))h(b) \end{bmatrix}$$

and if y^* is an isolated solution of (1), then $\mathrm{Ker}F'(y^*) = \{0\}$. Moreover, the shooting matrix $\phi_1(y^*(a), y^*(b))U(a) + \phi_2(y^*(a), y^*(b))U(b)$ is nonsingular, where $U(t)$ is a fundamental solution

$$\begin{cases} U'(t) &= \frac{\partial f}{\partial y}(t, y^*)U(t) \\ U(a) &= I. \end{cases}$$

It implies that $\mathrm{Im}F'(y^*) = C(J, \mathbb{R}^p) \times \mathbb{R}^p$. According to the Banach inverse mapping theorem, $[F'(y^*)]^{-1}$ exists and is continuous. The inverse function theorem ensures that the equation $F(y) = 0$ has a locally unique solution. □

Theorem 1. *Let the BVP (1) have an isolated solution φ. Assume that*

i) both $f : J \times S \to \mathbb{R}^p$ and $f_y' : J \times S \to \mathbb{R}^p$ are continuous;

ii)the function $f : J \times S \to \mathbb{R}^p$ has a bounded by a constant $L > 0$ derivative with respect to the second variable, and there exists a function $\Omega : \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$\|f_y'(t, x) - f_y'(t, \overline{x})\| \leq \Omega(\|x - \overline{x}\|),$$

where the matrix norm is consistent with the vector norm;

iii) Ω is continuous, $\Omega(0) = 0$ and Ω is non-decreasing;

iv) The Euler scheme is consistent with (1a), i.e there exists a function $\epsilon : H \to \mathbb{R}^+, H = [0, h^*]$ for some $h^* > 0, \epsilon(h) \to 0$ as $h \to 0$ such that

$$\|hf(t, \varphi(t)) + \varphi(t) - \varphi(t + h)\| \leq h\epsilon(h), \ \text{for } t \in [a, b - h];$$

v) There exists a function $\delta : H \to \mathbb{R}^+$, $\delta(h) \to 0$ as $h \to 0$, such that the following condition holds

$$\|[I + hf_y'(t, \varphi(t))]Y(t, \varphi(a)) - Y(t + h; \varphi(a))\| \leq h \colon \delta(h) \ \text{for } t \in [a, b - h];$$

vi) The function ϕ in the boundary condition has Lipschitz continuous partial derivatives

$$\|\phi_1(\overline{x}, y_1) - \phi_1(x_j, y_0)\| \leq K_1(\|\overline{x} - x_j\| + \|y_1 - y_0\|);$$
$$\|\phi_2(\overline{x}, y_1) - \phi_1(x_j, y_0)\| \leq K_2(\|\overline{x} - x_j\| + \|y_1 - y_0\|).$$

Moreover

$$\|\phi_2(x, y)\| \leq K_b.$$

Then for a sufficiently small h, the shooting method (6) is convergent.

Let

$$v_n^j := y_h(t_n; s_{h,j}) - \varphi(t_n); \ T_n^j := Y_h(t_n, s_{h,j})[s_{h,j} - \varphi(a)] - v_n^j;$$

$$R_n := hf(t_n, \varphi(t_n)) + \varphi(t_n) - \varphi(t_{n+1}); \ A_n^j := I + hf_y'(t_n, y_h(t_n; s_{h,j}));$$

$$q_n^j := Y_h(t_n, s_{h,j}) - Y(t_n, \varphi(a)); C_1 := e^{L(b-a)}; C_2 := (C_1 - 1)/L.$$

Introducing the norm $\|u\| = \max_{t \in J} \colon e^{-L(t-a)}\|u(t)\|, : u \in C(J, \mathbb{R}^p)$ and using the Gronwall-Bellman inequality we can show that the problem (4) has an unique solution Y is bounded by a constant C_1.

Lemma 2. *Under the conditions in the theorem 1, we have*

i) $\|v_n^j\| \le C_1\|v_0^j\| + C_2\epsilon(h)$;

ii) $\|T_n^j\| \le C_2 : \nu(h, \|v_0^j\|)$;

iii) $\|q_n^j\| \le C_2 : [C_1 : \Omega(C_1\|v_0^j\| + C_2\epsilon(h)) + \delta(h)]$,

where $\nu(h,\xi) = \Omega(C_1\xi + C_2\epsilon(h))(C_1\xi + C_2\epsilon(h)) + \epsilon(h)$.

Proof. The proof is carried out analogously to the proof of the convergence theorem in [4]. □

Lemma 3. *Let*

$$Q_h(u) = \phi_1(u, y_h(b;u)) + \phi_2(u, y_h(b;u))Y_h(b,u);$$

$$Q(u) = \phi_1(u, y(b;u)) + \phi_2(u, y(b;u))Y(b,u).$$

Then

$$Q_h(s_{h,j})v_0^{j+1} = \phi_2 T_N^j - [\phi(\overline{x}) - \phi(x_j) - \phi'(x_j)(\overline{x} - x_j)],$$

where $x_j = (s_{h,j}, y_h(b; s_{h,j})); \overline{x} = (\varphi(a), \varphi(b))$ ***and we have an estimate***

$$\|\phi(\overline{x}) - \phi(x_j) - \phi'(x_j)(\overline{x} - x_j)\| \le C_3\|v_0^j\|^2 + \epsilon_1(h)\|v_0^j\| + \epsilon_2(h),$$

where $\epsilon_i : H \to \mathbb{R}_+, : \epsilon_i(h) \to 0$ *as* $h \to 0, : i = 1,2$ *and* C_3 *is a nonnegative constant.*

Proof. The above-mentioned relation is easily obtained. Furthermore,

$$\begin{aligned}\|\phi(\overline{x}) - \phi(x_j) - \phi'(x_j)(\overline{x} - x_j)\| &= \|\int_0^1 [\phi'(x_j + t(\overline{x} - x_j) - \phi'(x_j)](\overline{x} - x_j)dt\| \\ &\le \frac{L'}{2}\|\overline{x} - x_j\|^2 \\ &\le \frac{L'}{2}(\|s_{h,j} - \varphi(a)\| + \|y(b; s_{h,j}) - \varphi(b)\|)^2 \\ &\le C_3\|v_0^j\|^2 + \epsilon_1(h)\|v_0^j\| + \epsilon_2(h).\end{aligned}$$

□

Now, we go on proving the theorem 1. It is known that, assuming that φ is an isolated solution of (1) then the matrix $Q(\varphi(a))$ is nonsingular(see, for example, [5]).

Let $\|Q^{-1}(\varphi(a))\| \le \beta_1$. From Lemma 3 and condition (vi) in theorem 1, it follows that

$$\begin{aligned}\|Q_h(s_{h,j}) - Q(\varphi(a))\| &= \|\phi_1(s_{h,j}, y_h(b; s_{h,j})) - \phi_1(\varphi(a), \varphi(b)) + \\ &+ \phi_2(s_{h,j}, y_h(b; s_{h,j}))Y_h(b, s_{h,j}) - \phi_2(s_{h,j}, y_h(b; s_{h,j}))Y(b, \varphi(a)) + \\ &+ \phi_2(s_{h,j}, y)_h(b; s_{h,j}))Y(b, \varphi(a) - \phi_2(\varphi(a), \varphi(b))Y(b, \varphi(a))\| \\ &\le K_1(\|s_{h,j} - \varphi(a)\| + \|y_h(b; s_{h,j}) - \varphi(b)\|) + \\ &+ K_b \underbrace{\|Y_h(b, s_{h,j}) - Y(b, \varphi(a))\|}_{q_N^j} + \\ &+ K_2(\|s_{h,j} - \varphi(a)\| + \|y_h(b; s_{h,j}) - \varphi(b)\|)C_1 \\ &= M_1\|v_0^j\| + M_2\epsilon(h) + K_b\|q_N^j\|.\end{aligned}$$

Therefore

$$\|Q^{-1}(\varphi(a))[Q_h(s_{h,j}) - Q(\varphi(a))]\| \le \beta_1\{M_1\|v_0^j\| + M_2\epsilon(h) + K_b\|q_N^j\|\}$$
$$\le \beta_1\{M_1\|v_0^j\| + M_2\epsilon(h) + K_bC_2[C_1\Omega(C_1\|v_0^j\| + C_2\epsilon(h)) + \delta(h)]\} =: P_j(h). \qquad (8)$$

Since Ω is continuous and $\Omega(0) = 0$, one can choose $\rho = \|v_0^0\| = \|s_0 - \varphi(a)\| << 1$, such that

$$P_0(h) \le \alpha < 1 \text{ and } M\rho + \frac{\beta_1}{1-\alpha}\epsilon_1(h) + \beta_0\Omega(C_1\rho) \le \overline{\alpha} < 1 \text{ for } h << 1, \qquad (9)$$

where $\beta_0 = \beta_1 K_b C_2 C_1/(1-\alpha)$, $M = \max\{\beta_1 M_1, M_3 = \beta_1 C_3/(1-\alpha)\}$.

Using Lemma 4.4.14 from [1], one concludes that the matrix

$$I + Q^{-1}(\varphi(a))[Q_h(s_{h,0}) - Q(\varphi(a))]$$

is nonsingular. Furthermore, $Q_h(s_0)$ is nonsingular and $\|Q_h^{-1}(s_0)\| \le \beta_1/(1-\alpha)$. Applying Lemma 3, one has

$$\|v_0^1\| \le \|Q_h(s_{h,0})^{-1}\|\,\{\|\phi_2(s_{h,0}, y_h(b; s_{h,0}))\|T_N^0\| + \|\phi(\overline{x}) - \phi(x_j) - \phi'(x_j)(\overline{x} - x_j)\|\}$$
$$\le \overline{\beta}\colon \nu(h,\rho) + M\rho^2 + \frac{\beta_1}{1-\alpha}\rho\,\epsilon_1(h) + \frac{\beta_1}{1-\alpha}\epsilon_2(h) := u_h^1\ (\overline{\beta} = \beta_1 K_b C_2/(1-\alpha)).$$

It can be proved that there exists a function $\xi, : \xi(h) \to 0$ as $h \to 0$ such that

$$\|v_0^1\| \le u_h^1 = \overline{\beta}\nu(h,\rho) + M\rho^2 + \frac{\beta_1}{1-\alpha}\rho\,\epsilon_1(h) + \frac{\beta_1}{1-\alpha}\epsilon_2(h)$$
$$\le \beta_0\rho\Omega(C_1\rho) + M\rho^2 + \frac{\beta_1}{1-\alpha}\rho\,\epsilon_1(h) + \zeta(h)$$
$$\le \overline{\alpha}\rho + \zeta(h) \le \rho := u_h^0 \text{ for } h << 1.$$

By an induction and argue as above, one sees that $Q_h^{-1}(s_{h,j})$ is nonsingular. Moreover

$$\|Q_h^{-1}(s_{h,j})\| \le \frac{\beta_1}{1-\alpha}, : \|v_0^{j+1}\| \le u_h^{j+1} \le u_h^j,$$

where $u_h^{j+1} = \overline{\beta}\nu(h, u_h^j) + M(u_h^j)^2 + \frac{\beta_1}{1-alpha}u_h^j\,\epsilon_1(h) + \frac{\beta_1}{1-\alpha}\epsilon_2(h)$.

The sequence $\{u_h^j\}_k$ is nonincreasing and nonnegative, so $u_h = \lim_{j\to\infty} u_h^j$ exits and u_h satisfies the equation

$$u_h = \overline{\beta}\nu(h, u_h) + M(u_h)^2 + \frac{\beta_1}{1-\alpha}u_h\,\epsilon_1(h) + \frac{\beta_1}{1-\alpha}\epsilon_2(h).$$

Moreover, if $u_h \to u$ as $h \to 0$, then u is a non-negative solution of

$$u = \overline{\beta}\Omega(C_1 u)(C_1 u) + Mu^2. \qquad (10)$$

Since $0 \le u_h \le u_h^0 = \rho$, it implies $u \le \rho$. The estimate $\beta_0\Omega(C_1 u) + Mu \le \overline{\alpha} < 1$ ensures that $u = 0$ is an unique solution of (10). This and (i) in Lemma 2 yield the assertion, which was to be proved.

3. Numerical Experiments

In this section, we consider three examples. In each example, step sizes and an initial vector are given. The convergence of the method in each example is expressed by $\max|y_h^j(t_i) - y_h^{j-1}(t_i)|$, where $y_h^j(t_i)$ denotes $y_h(t_i, s_{hj})$.

Example 1. Consider the problem

$$y' = -4ty^{3/2},\ t \in [0,1],$$

$$\sin(y(0)) + \cos(y(1)) - \sin\frac{1}{(1+\sqrt{2})^2} - \cos\frac{1}{(2+\sqrt{2})^2} = 0.$$

This problem has been studied in [4] with the linear boundary condition $y(0) - 2y(1) = 0$. In both cases, it has an exact solution $y(t) = (t^2+1+\sqrt(2))^{-2}$ and $y(0) = 0.17157287525381$. Furthermore, $f_y'(t,y) = -6ty^{1/2}$ is not bounded. However, it is bounded in a bounded neighbourhood of the isolated solution. We can put $\Omega(\nu) = 6\nu^{1/2}$. Thus Ω and the function in the boundary condition satisfy the conditions in the theorem 1.

N	s_j	Time	Accuracy
10	0.17173830227481	$0.11s$	10^{-4}
40	0.17161116830044	$0.27s$	10^{-5}
150	0.17158289850619	$0.88s$	10^{-6}

Table 1: Numerical result of example 1 with initial vector $s_0 = 0.5$.

Starting from the initial vector $s_0 = 0.5$ we get the improved shooting points given in Table 1. The convergent speed of the method ($N = 10$) is given in Table 2.

j	$\max\|y_h^j(t_i) - y_h^{j-1}(t_i)\|$
1	5.00×10^{-1}
3	2.38×10^{-2}
4	6.07×10^{-5}
5	4.07×10^{-10}
6	1.11×10^{-16}

Table 2: The convergence of the method in example 1 with $N = 10$.

Example 2. Consider the problem

$$y' = z, : z' = y,\ t \in [0,1],$$

$$e^{-0.01y^2(0)} + \cos y(1) - e^{-0.01} - \cos(e^{-1}) = 0,$$

This problem has an exact solution $y = e^{-t}, : z = -e^{-t}$ and $y(0) = 1, : z(0) = -1$. It is easy to show that the function in the boundary condition satisfies all the conditions in the theorem 1.

N	$s_j(1)$	$s_j(2)$	Time	Accuracy
30	1	-0.99462269847685	$1.81s$	10^{-2}
250	1	-0.99937163278951	$2.58s$	10^{-4}
2000	1	-0.99992170530796	$10.93s$	10^{-5}

Table 3: Numerical result of example 2 with initial vector $s_0 = [1.5; -1.5]$.

j	$\max\lvert y_h^j(t_i) - y_h^{j-1}(t_i)\rvert$	$\max\lvert z_h^j(t_i) - z_h^{j-1}(t_i)\rvert$
1	1.45×10^0	1.45×10^0
5	6.39×10^{-2}	9.82×10^{-2}
7	8.44×10^{-6}	1.26×10^{-5}
8	5.82×10^{-11}	8.62×10^{-11}
9	7.22×10^{-16}	7.22×10^{-16}

Table 4: The convergence of the method in example 2 with $N = 30$.

Choose the initial vector $s_0 \equiv [1.5; -1.5]$ we have Table 3. The convergent speed of the method ($N = 30$) is given in Table 4.

Example 3. Consider the problem

$$y'' = \frac{3}{2}y^2, \ t \in [0,1],$$
$$e^{-0.001y^2(0)} - e^{-0.016} = 0,$$
$$e^{-0.001y^2(1)} - e^{-.0001} = 0.$$

Put $z = y'$, the equation becomes

$$\begin{bmatrix} y \\ z \end{bmatrix}' = \begin{bmatrix} z \\ 3/2y^2 \end{bmatrix}, \ t \in [0,1].$$

The partial derivative f_y' is not bounded in the whole space but it is bounded in a neighborhood of an isolated solution. The remaining conditions of the theorem 1 can be easily checked.

N	$s_j(1)$	$s_j(2)$	Time	Accuracy
40	4	-8.02978329590422	$1.70s$	10^{-2}
400	4	-8.00297581304037	$4.56s$	10^{-4}
4000	4	-8.00029758256503	$30.3s$	10^{-5}

Table 5: Numerical result of example 3 with initial vector $s_0 = [4.5; -8.5]$.

This problem has an exact solution $y = 4/(1+t)^2$ and $y(0) = 4, : z(0) = -8$. Choose the initial vector $s_0 = [4.5; -8.5]$ we have Table 5. The convergent speed of the method ($N = 40$) is shown in Table 6.

j	$\max\lvert y_h^j(t_i) - y_h^{j-1}(t_i)\rvert$	$\max\lvert z_h^j(t_i) - z_h^{j-1}(t_i)\rvert$
1	3.99×10^0	7.54×10^0
5	7.81×10^{-2}	1.60×10^{-1}
7	7.07×10^{-6}	1.44×10^{-5}
8	1.73×10^{-10}	3.50×10^{-10}
9	5.60×10^{-14}	1.09×10^{-13}

Table 6: The convergence of the method in example 3 with $N = 40$.

References

[1] Ascher, U., Mattheu, R., Russel, R., *Numerical Solution for Boundary Value Problems for Ordinary Differential Equations*, Prentice Hall, Englewood Cliffs, New Jersey, 1998.

[2] Bernfeld, Stephen R., Lakshmikantham, V., *An Introduction to Nonlinear Boundary Values Problems*, Academic Press, Inc, 1974.

[3] Chow, K. L., Enright, W. H., Distributed Parallel Shooting for BVPODEs, *Technical Report of Department of Computer Science*, University of Toronto, 1995.

[4] Jankowski, T., Approximate Solutions of Boundary-Value Problems for Systems of Ordinary Differential Equations, *Comp. Maths Math. Phys.*, Vol 35, **No. 7**(1995), pp 837-843.

[5] Keller, H. B., Nummerical solution of two-point boundary value problems. *Soc. Industr. and Appl. Math.*, **24**(1976), 1-61.

[6] Pašic, H., Multipoint Boundary-Value Solution of Two-Point Boundary-Value Problems, *Journal of Optimization Theory and Applications*, Vol. 100(1999), pp 397-416.

[7] Shampine, L. F., Reichelt., M. W., The MATLAB ODE Suite, *SIAM Journal on Scientific Computing*, 18-1, 1997.

[8] Roberts, S., Shipman, J., *Two-Point Boundary-Value Problems: Shooting Methods*, Elsevier, 1972.

[9] Stoer, J., and Bulirsch, R., *Introduction to Numerical Analysis*, Springer-Verlag NewYork, 1993.

[10] Vasiliev, N.I., Klokov, Yu. A., *Fundamentals of theory of BVPs for ODEs*, Riga, Zinatne, 1978.