NUMERICAL SOLUTION OF BOUNDARY VALUE PROBLEMS FOR SECOND ORDER FUNCTIONAL DIFFERENTIAL EQUATIONS BY THE USE OF CUBIC SPLINES

By

Manabu SAKAI

(Received Oct. 3, 1974)

1. Introduction

In the present paper we consider numerical methods for computing the solution of the boundary value problem:

$$(1) \quad \ddot{x}(t) = p(t)x(g(t)) + f(t, x(t)) \quad (0 \leq t \leq 1)$$

with boundary conditions

$$(2) \quad x(t) = u(t) \quad (\alpha \leq t \leq 0) \quad (\alpha = \min(g(t), 0))$$

$$(3) \quad x(t) = v(t) \quad (1 \leq t \leq \beta) \quad (\beta = \max(g(t), 1)).$$

It will be assumed that

(i) $p(t), g(t) \in C^2[0, 1]$,

(ii) $u(t) \in C^2[\alpha, 0]$ and $v(t) \in C^2[1, \beta]$,

(iii) $f(t, x)$ is defined and twice-continuously differentiable in a region D of (t, x) -space intercepted by two lines $t=0$ and $t=1$,

(iv) the number of changes in sign of values of $g(t)(g(t)-1)$ is finite,

(v) on each subinterval $[p/n, (p+1)/n]$ ($p=0, 1, \dots, n-1$) for $n > N$ (some integer N), $g(t)(g(t)-1) \leq 0$.

It will also be assumed that (1)–(3) has an isolated solution $\hat{x}(t)$ satisfying the internality condition

$$U = \{(t, x) \mid |x - \hat{x}(t)| \leq \delta, t \in [0, 1]\} \subset D \quad \text{for some } \delta > 0.$$

A solution $\hat{x}(t)$ is isolated if and only if the following homogeneous boundary value problem has only the trivial solution:

$$\ddot{x}(t) = \bar{p}(t)x(g(t)) + f_x(t, \hat{x}(t))x(t) \quad (0 \leq t \leq 1)$$

with boundary conditions

$$x(0) = x(1) = 0,$$

where

$$\bar{p}(t) = \begin{cases} p(t) & (0 \leq g(t) \leq 1) \\ 0 & (\text{otherwise}). \end{cases}$$

Corresponding to $\hat{x}(t)$, as is easily seen, one can determine a cubic spline $\hat{x}_h(t)$ of the form

$$\hat{x}_h(t) = \Sigma \hat{\alpha}_p Q_4(t/h - p) \quad (nh = 1)$$

so that

$$\begin{aligned} \hat{x}_h(t_p) &= \hat{x}(t_p) & (p=0, 1, \dots, n) \quad (t_p = ph) \\ \dot{\hat{x}}_h(t_p) &= \dot{\hat{x}}(t_p) & (p=0, n). \end{aligned}$$

Here we state a slight generalization of Theorem 2.3.4 in [1].

LEMMA 1. *Let $\psi(t) \in C^2[0, 1] \cap C^4[t_p, t_{p+1}]$ ($p=0, 1, \dots, n-1$) and let $\psi_h(t)$ be the cubic spline function interpolating to $\psi(t)$, then we have*

$$\left| \left(\frac{d}{dt} \right)^k [\psi_h(t) - \psi(t)] \right| = O(h^{4-k}) \quad (k=0, 1, 2) \quad (h \rightarrow 0).$$

Proof. We have only to notice that

$$\begin{aligned} & \frac{1}{h^2} (\psi_{p+1} - 2\psi_p + \psi_{p-1}) - \frac{1}{6} (\ddot{\psi}_{p+1} + 4\ddot{\psi}_p + \ddot{\psi}_{p-1}) \\ &= -\frac{h^2}{24} (\psi^{(4)}(c_p +) + \psi^{(4)}(c_p -)), \end{aligned}$$

where $t_p < c_p < t_{p+1}$ and $t_{p-1} < c_p < t_p$.

Since $\hat{x}(t) \in C^2[0, 1] \cap C^4[t_p, t_{p+1}]$ ($p=0, 1, \dots, n-1$) from the assumptions (i)–(v), it holds that

$$(4) \quad \left| \left(\frac{d}{dt} \right)^k [\hat{x}_h(t) - \hat{x}(t)] \right| = O(h^{4-k}) \quad (k=0, 1, 2)$$

uniformly on $[0, 1]$ as $h \rightarrow 0$.

Now we consider the spline function of the form

$$x_h(t) = \sum \alpha_p Q_4(t/h - p)$$

with coefficients $\alpha_{-3}, \alpha_{-2}, \dots, \alpha_{n-1}$ undetermined. The above $x_h(t)$ will be an approximate solution to the problem (1)–(3) if it satisfies

$$(i) \quad \ddot{x}_h(t_q) = \begin{cases} p(t_q)x_h(g(t_q)) + f(t_q, x_h(t_q)) & \text{if } g(t_q) \in [0, 1] \\ p(t_q)u(g(t_q)) + f(t_q, x_h(t_q)) & \text{if } g(t_q) < 0 \\ p(t_q)v(g(t_q)) + f(t_q, x_h(t_q)) & \text{if } g(t_q) > 1, \end{cases}$$

$$(ii) \quad x_h(0) = u(0) \quad \text{and} \quad x_h(1) = v(1).$$

The number of undetermined coefficients is $n+3$ and the conditions (i)–(ii) give precisely the requisite number of equations.

For the convenience of the analysis, we write (i) and (ii) in the form

$$F_q(\alpha) = \frac{1}{h^2} (\alpha_{q-3} - 2\alpha_{q-2} + \alpha_{q-1}) - f\left(t_q, \frac{1}{6} (\alpha_{q-3} + 4\alpha_{q-2} + \alpha_{q-1})\right) \\ - \begin{cases} p(t_q)\sum \alpha_p Q_4(g(t_q)/h - p) & \text{if } g(t_q) \in [0, 1] \\ p(t_q)u(g(t_q)) & \text{if } g(t_q) < 0 \\ p(t_q)v(g(t_q)) & \text{if } g(t_q) > 1, \end{cases} \\ (q=0, 1, \dots, n)$$

$$F_{-1}(\alpha) = \frac{1}{6} (\alpha_{-3} + 4\alpha_{-2} + \alpha_{-1}) - u(0),$$

$$F_{n+1}(\alpha) = \frac{1}{6} (\alpha_{n-3} + 4\alpha_{n-2} + \alpha_{n-1}) - v(1).$$

2. Existence of Spline Approximation

In what follows, for any $\varphi(t) \in C[0, 1]$, we shall denote $\max |\varphi(t)|$ by $\|\varphi\|$, and for any finite dimensional vector c , we denote its maximum norm by $\|c\|$.

In this section, we demonstrate the existence and convergence spline approximation by making use of Kantrovitch's theorem on the convergence of

$c_p(p=1, 2, \dots, 14)$ is a constant independent of h .

Newton's method (Rall [4, p. 367]).

First let us estimate $\|F(\hat{\alpha})\|$. By (4), we have

$$(5) \quad \|F(\hat{\alpha})\| \leq c_1 h^2.$$

Let $J(\alpha)$ be the Jacobian matrix of $F(\alpha)$ with respect to α_p ($p = -3, -2, \dots, n-1$). In order to investigate the properties of $J(\hat{\alpha})$, let us consider a linear system

$$(6) \quad J(\hat{\alpha})\xi = \eta,$$

where $\xi = (\xi_{-3}, \xi_{-2}, \dots, \xi_{n-1})$ and $\eta = (\eta_{-1}, \eta_0, \dots, \eta_{n+1})$.

Corresponding to ξ and η , we consider the cubic spline $\varphi(t)$ and the linear function $\lambda(t)$ defined by, respectively,

$$\varphi(t) = \sum \xi_p Q_4(t/h - p)$$

$$\lambda(t) = \sum \eta_p L_p(t),$$

where $L_p(t)$ is a piecewise linear function with the property $L_p(t_q) = \delta_{pq}$.

Corresponding to (6), we have

$$(7) \quad \ddot{\varphi}(t) = T[f_x(t, \hat{x}_h(t))\varphi(t) + \bar{p}(t)\varphi(g(t))] + \lambda(t) \quad (0 \leq t \leq 1)$$

with boundary conditions

$$\varphi(0) = \eta_{-1} \quad \text{and} \quad \varphi(1) = \eta_{n+1},$$

where T is an operator defined by

$$(T\psi)(t) = \sum \psi(t_p) L_p(t).$$

Equation (7) can be rewritten as follows:

$$(8) \quad -\ddot{\varphi}(t) + f_x(t, \hat{x}_h(t))\varphi(t) + \bar{p}(t)\varphi(g(t)) = R(t) + \lambda(t),$$

where $R(t) = (I - T)[f_x(t, \hat{x}_h(t))\varphi(t) + \bar{p}(t)\varphi(g(t))]$.

Now we consider the solution $\varphi_1(t)$ satisfying the homogeneous boundary conditions of (8):

$$\varphi_1(0) = \varphi_1(1) = 0.$$

Here we consider an operator K defined by

$$(K\psi)(t) = \int_0^1 G(t, s) [f_x(s, \hat{x}(s))\psi(s) + \bar{p}(s)\psi(g(s))] ds,$$

where

$$G(t, s) = \begin{cases} s(t-1) & (0 \leq s \leq t) \\ t(s-1) & (t \leq s \leq 1) \end{cases}$$

From the assumption (iv), K is a compact operator from $C[0, 1]$ to $C[0, 1]$.

Since $\hat{x}(t)$ is isolated, by making use of Fredholm's alternating theorem, we have

$$(9) \quad \|(I - K)^{-1}\| \leq c_2.$$

From (4) and (8), we have the inequality of the form

$$(10) \quad \|\varphi_1\| \leq c_3 (\|R\| + \|\eta\|) \quad \text{for sufficiently small } h.$$

Next, we consider the solution $\varphi_2(t)$ satisfying

$$-\ddot{\varphi}_2(t) + f_x(t, \hat{x}_h(t))\varphi_2(t) + \bar{p}(t)\varphi_2(g(t)) = 0 \quad (0 \leq t \leq 1)$$

with boundary conditions

$$\varphi_2(0) = \eta_{-1} \quad \text{and} \quad \varphi_2(1) = \eta_{n+1}.$$

By (9), we have the inequality of the form

$$(11) \quad \|\varphi_2\| \leq c_4 \|(\eta_{-1}, \eta_{n+1})\|.$$

Hence by (10) and (11), we have

$$(12) \quad \|\varphi\| \leq c_3 (\|R\| + \|\lambda\|) + c_4 \|(\eta_{-1}, \eta_{n+1})\|.$$

In order to estimate $\|R\|$, we need the following lemma.

LEMMA 2. Let $\psi(t) \in C[0, 1]$, then we have

$$\|(I - T)\psi\| \leq \frac{h}{2} \|\dot{\psi}\|.$$

By this lemma, we have

$$(13) \quad \|R\| \leq \frac{h}{2} \| [f_{xt}(t, \hat{x}_h(t)) + f_{xx}(t, \hat{x}_h(t))\dot{\hat{x}}_h(t)] \varphi(t) \\ + f_x(t, \hat{x}_h(t))\dot{\varphi}(t) + \bar{p}(t)\varphi(g(t)) + \bar{p}(t)\dot{\varphi}(g(t))\dot{g}(t) \|.$$

On the other hand, if h is sufficiently small, by (4), there exists $c_5 > 0$ such that

$$|f_{xt}(t, \hat{x}_h(t))|, |f_{xx}(t, \hat{x}_h(t))|, |f_x(t, \hat{x}_h(t))| \leq c_5.$$

Therefore we have

$$(14) \quad \|R\| \leq \frac{h}{2} [c_5(1+c_6)\|\varphi\| + c_5\|\dot{\varphi}\| + c_7\|\varphi\| + c_8\|\dot{\varphi}\|],$$

where $\|\dot{\hat{x}}_h\| \leq c_6, \|\dot{\bar{p}}\| \leq c_7$ and $\|\bar{p}\dot{g}\| \leq c_8$.

Here we need the well-known inequality:

Let $\psi(t) \in C^2[0, 1]$, then there exists a positive constant $C(\varepsilon)$ such that $\|\dot{\varphi}\| \leq \varepsilon\|\ddot{\varphi}\| + C(\varepsilon)\|\varphi\|$ for any $\varepsilon > 0$.

From (8), by this inequality, we have

$$(15) \quad \begin{aligned} \|\dot{\varphi}\| &\leq \varepsilon\|\ddot{\varphi}\| + C(\varepsilon)\|\varphi\| \\ &\leq \varepsilon[\|R\| + \|\lambda\| + c_5\|\varphi\| + c_7\|\varphi\|] + C(\varepsilon)\|\varphi\|. \end{aligned}$$

By (14) and (15) ($\varepsilon=1$), we have the estimate of $\|R\|$ of the form

$$(16) \quad \|R\| \leq c_9 h [\|\varphi\| + \|\lambda\|] \quad \text{for sufficiently small } h.$$

By (12) and (16), we have the inequality of the form

$$\|\varphi\| \leq c_{10} [\|\lambda\| + \|(\eta_{-1}, \eta_{n+1})\|] \quad \text{for sufficiently small } h,$$

from which follows

$$(17) \quad \|\xi\| \leq c_{11} \|\eta\|.$$

The inequality (17) implies the non-singularity of $J(\hat{\alpha})$ and the inequality

$$(18) \quad \|J^{-1}(\hat{\alpha})\| \leq c_{12}.$$

Let $\alpha = (\alpha_{-3}, \alpha_{-2}, \dots, \alpha_{n-1})$ and $\beta = (\beta_{-3}, \beta_{-2}, \dots, \beta_{n-1})$ be arbitrary vectors such that

$$\left| \frac{1}{6} (\alpha_{p-3} + 4\alpha_{p-2} + \alpha_{p-1}) - \hat{x}(t_p) \right| \leq \delta,$$

and

$$\left| \frac{1}{6} (\beta_{p-3} + 4\beta_{p-2} + \beta_{p-1}) - \hat{x}(t_p) \right| \leq \delta \quad (p=0, 1, \dots, n).$$

By means of the mean-value theorem we have

$$(19) \quad \|J(\alpha) - J(\beta)\| = \max_p \left(\sum_q |J_{pq}(\alpha) - J_{pq}(\beta)| \right) \leq K \|\alpha - \beta\|,$$

where c_{13} is a positive constant such that

$$|f_{xx}(t, x)| \leq c_{13} (=K) \quad \text{for all } (t, x) \in U.$$

Now, let us note that there is a positive constant c_{14} such that

$$\|\hat{x}_h - \hat{x}\| \leq c_{14} h^4.$$

From (19), we have

$$(20) \quad \|J(\alpha) - J(\hat{\alpha})\| \leq K \|\alpha - \hat{\alpha}\| \quad \text{for } \|\alpha - \hat{\alpha}\| \leq r (= \delta - c_{14} h^4)$$

for $h \leq h_1$ provided h_1 sufficiently small.

From (5) and (18), we have

$$(21) \quad \|J^{-1}(\hat{\alpha})\| \leq B_0 (=c_{12}),$$

$$(22) \quad \|J^{-1}(\hat{\alpha})F(\hat{\alpha})\| \leq \eta_0 (=c_1 c_{12} h^2)$$

for $h \leq h_2$, provided that h_2 is sufficiently small.

Hence, by applying Kantorovich's theorem to the determining equation $F(\alpha) = 0$, we have

THEOREM. *In a sufficiently small neighbourhood of the isolated solution $\hat{x}(t)$ of the problem (1)–(3), there is a cubic spline function of the form*

$$\bar{x}_h(t) = \Sigma \bar{\alpha}_p Q_4(t/h - p)$$

such that (i) the coefficient $\bar{\alpha} = (\bar{\alpha}_{-3}, \bar{\alpha}_{-2}, \dots, \bar{\alpha}_{n-1})$ satisfies the determining equation $F(\alpha) = 0$, and

$$(ii) \quad \|x_h - x\| = O(h^2) \quad (h \rightarrow 0).$$

COROLLARY.

$$\left\| \left(\frac{d}{dt} \right)^k [\bar{x}_h - \hat{x}] \right\| = O(h^2) \quad (k=0, 1, 2) \quad (h \rightarrow 0).$$

Proof. We have only to notice that the piecewise linear function $\bar{x}_h(t)$ satisfies the following equations:

$$\bar{x}_h(t_q) = p(t_q) \bar{x}_h(g(t_q)) + f(t_q, \bar{x}_h(t_q)) \quad (q=0, 1, \dots, n).$$

3. Numerical Illustration

In the following we present two numerical examples to illustrate the application of the method of cubic spline.

Example 1

Consider the linear boundary value problem:

$$\ddot{x}(t) = -x(t-0.1) + 10 \sin(10t-1) - 1000 \sin 10t \quad (0 \leq t \leq 1)$$

with boundary conditions

$$x(t) = 10 \sin 10t \quad (t \leq 0) \quad \text{and} \quad x(1) = 10 \sin 10.$$

Its exact solution is $10 \sin 10t$. A selection of numerical results is presented in Table 1.

Table 1

	$ \hat{x}(t) - \bar{x}_h(t) $		$ \hat{x}(t) - x_d(t) $
	$h=0.05$	$h=0.025$	
0.2	2.1×10^{-1}	5.5×10^{-2}	0.2×10^{-3}
0.4	1.0	2.7	1.0
0.6	0.2	0.5	1.4
0.8	2.9	8.2	1.2

By $x_d(t)$, we mean the improved approximation by deferred correction.

Example 2 (Nevers-Schmitt [3])

Next, we consider the nonlinear case:

$$\ddot{x}(t) = -\frac{1}{16} \sin x(t) + t - (t+1)x(t-1) \quad (0 < t < 2)$$

with boundary conditions

$$x(t) = t - \frac{1}{2} \quad (t \leq 0) \quad \text{and} \quad x(2) = -\frac{1}{2}.$$

We compute the initial approximation $x_1(t)$ of Newton's method as follows: by collocation conditions, we have

$$\ddot{x}_1(0) = -\frac{1}{16} \sin x_1(0) - x_1(-1) = -\sin \frac{1}{2} + \frac{3}{2}$$

$$\ddot{x}_1(1) = -\frac{1}{16} \sin x_1(1) + 1 - 2x_1(0) = -\sin x_1(1) + 2$$

$$\ddot{x}_1(2) = -\frac{1}{16} \sin x_1(2) + 2 - 3x_1(1) = -\sin \frac{1}{2} - 3x_1(1) + 2,$$

by consistency condition of cubic spline with mesh size $h=1$

$$x_1(2) - 2x_1(1) + x_1(0) = \frac{1}{6} \{ \ddot{x}_1(2) + 4\ddot{x}_1(1) + \ddot{x}_1(0) \},$$

from which follows the single equation of $x_1(1)$:

$$72x_1(1) - 2 \sin x_1(1) = -140 - \sin \frac{1}{2}.$$

The approximate values at $t=0.5, 1.0$ and 1.5 are give in Table 2.

For comparison, Table 2 contains the values computed by Cryer [2] and Nevers-Schmitt [3].

Table 2

$n+1$	$t=0.5$		$t=1.0$		$t=1.5$	
21	-1.543	107	-2.082	050	-1.962	812
41	-1.543	458	-2.082	681	-1.963	245
Cryer 64	-1.543	319	-2.082	247	-1.963	052
128	-1.543	542	-2.082	725	-1.963	348
Nevers-Schmitt	-1.543	053	-2.081	821	-1.962	343

References

- [1] J. AHLBERG, E. NILSON and J. WALSH: *The theory of splines and its applications*. Academic Press, New York, 1967.
- [2] C. CRYER: *The numerical solution of boundary value problems for second order functional differential equations by finite differences*. Numer. Math. **20** (1973), 288-299.

- [3] K. NEVERS and K. SCHMITT: *An application of the shooting method to boundary value problems for second order delay equations.* J. Math. Anal. Appl. **36** (1969), 588–597.
- [4] L. RALL: *Computational solution of nonlinear operator equation.* Wiley, New York, 1969.
- [5] M. SAKAI: *Piecewise cubic interpolation and two-point boundary value problems.* Publ. Res. Inst. Inst. Math. Sci., Kyoto Univ. **7** (1971), 345–362.
- [6] M. SAKAI: *Cubic spline function and difference method.* Mem. Fac. Sci., Kyushu Univ. **28** (1974), 43–58.