HIROSHIMA MATH. J.
8 (1978), 101–121

A Posteriori Error Estimates and Iterative Methods in the Numerical Solution of Systems of Ordinary Differential Equations

Hisayoshi SHINTANI and Yuichi HAYASHI

(Received September 19, 1977)

1. Introduction

Consider the boundary value problem

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

$$f[x] = 0, \tag{1.2}$$

and let $x_0(t)$ be an approximate solution of this problem, where x and $X(x, t)$ are real n-vectors, f is an operator from $D \subset C[J]$ into R^n which is continuously Fréchet differentiable in D, and $C[J]$ is the space of all real n-vector functions continuous on $[a, b]$.

Constructing an operator equation from (1.1), (1.2) and approximating the Fréchet derivative of the operator in a neighborhood of x_0 by a linear operator independent of x, by means of an iterative method Urabe [7] proved the existence and local uniqueness of an exact solution and gave an a posteriori error estimate of x_0 in terms of $x_0(t)$ and its derivative.

The first object of this paper is to obtain the results similar to those in [7] for continuous $x_0(t)$ without assuming its differentiability. This is achieved by replacing (1.1) with an equivalent system of integral equations. Hence the results can be applied to discrete numerical solutions by means of interpolation.

The second object of this paper is to treat the case where the linear operator approximating the Fréchet derivative depends on x. This enables us to construct various iterative methods.

In Section 3 the results are applied to multipoint boundary value problems [5, 6]. In Section 4 we consider boundary value problems of the least squares type [1, 8] which arise often in system identification problems and propose some iterative methods.

2. Convergence of iterative methods and error estimates

Let R^n denote a real n-space with any norm $\|\cdot\|$ and let $C[J]$ be the space of all real n-vector functions continuous on the interval $J=[a, b]$, which is made into a Banach space with the norm $\|\cdot\|_c$ defined by

$$\|x\|_c = \sup_{t\in J}\|x(t)\| \qquad \text{for } x\in C[J].$$

For any fixed $t_0\in J$ let

$$C_0[J] = \{x\in C[J] \mid x(t_0) = 0\}.$$

Then $B=C_0[J]\times R^n$ is a Banach space with the norm $\|\cdot\|_b$,

$$\|y\|_b = \max(\|u\|_c, \|e\|) \qquad \text{for} \quad y = (u, e)\in B.$$

Let Ω' be a domain of the tx-space intercepted by two hyperplanes $t=a$ and $t=b$ such that the cross sections R_a and R_b at $t=a$ and $t=b$ make an open set in each hyperplane, and put $\Omega=R_a\cup\Omega'\cup R_b$. Let

$$D = \{x\in C[J] \mid (t, x(t))\in\Omega \quad \text{for all} \quad t\in J\}.$$

Then D is an open set in $C[J]$.

For two Banach spaces B_1 and B_2, we denote by $L(B_1, B_2)$ the set of all bounded linear operators from B_1 into B_2. For $G: D\subset B_1\to L(B_1, B_2)$ let $G(x)$ be the element of $L(B_1, B_2)$ associated with $x\in D$. When $F: D\subset B_1\to B_2$ is Fréchet differentiable at $x\in D$, we denote by $F'(x)$ the Fréchet derivative of F at x.

The identity operator and the unit matrix are denoted by the same symbol I. The product FG and the sum $F+G$ of two operators F and G are defined in the usual manner.

Let us consider the system of differential equations

$$\text{(2.1)} \qquad \frac{dx}{dt} = X(x, t), \qquad a\leqq t\leqq b,$$

with the boundary condition

$$\text{(2.2)} \qquad f[x] = 0,$$

where x and $X(x, t)$ are n-vectors, $X(x, t)$ is continuous in Ω and continuously differentiable with respect to x in Ω, and the operator $f: D\to R^n$ is continuously Fréchet differentiable in D. We assume that (2.1) has at least one solution in D.

Let $Q: D\to C_0[J]$ be defined by

$$\text{(2.3)} \qquad Qx = x(t) - x(t_0) - \int_{t_0}^{t} X(x(s), s)ds \qquad \text{for} \quad x\in D.$$

Then $Qx=0$ if and only if $x\in D$ is a solution of (2.1). Hence the boundary value problem (2.1), (2.2) is equivalent to finding a solution $x\in D$ of the equation

$$Fx = 0, \tag{2.4}$$

where $F\colon D\to B$ is defined by

$$Fx = (Qx, f[x]) \qquad \text{for} \quad x\in D. \tag{2.5}$$

Let $X_x(x, t)$ be the Jacobian matrix of $X(x, t)$ with respect to x. Then $F'(x)h$ $(x\in D)$ is given by

$$F'(x)h = (Q'(x)h, f'(x)h) \qquad \text{for} \quad h\in C[J], \tag{2.6}$$

where

$$Q'(x)h = h(t) - h(t_0) - \int_{t_0}^{t} X_x(x(s), s)h(s)ds. \tag{2.7}$$

In relation to $F'(x)$ we introduce the bounded linear operator $L(x)$ $(x\in D)$ defined by

$$L(x)h = (P(x)h,\ l(x)h) \qquad \text{for} \quad h\in C[J], \tag{2.8}$$

where

$$P(x)h = h(t) - h(t_0) - \int_{t_0}^{t} A(x(s), s)h(s)ds, \tag{2.9}$$

$A(x, t)$ is an $n\times n$ matrix continuous in Ω, and $l\colon D\to L(C[J], R^n)$ is bounded and continuous in D.

Let $\Phi_{(x)}(t)$ be the fundamental matrix of the system

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

satisfying $\Phi_{(x)}(t_0)=I$ and denote by

$$G(x) = l(x)\,[\Phi_{(x)}] \tag{2.10}$$

the matrix whose column vectors are $l(x)\Phi_i$ $(i=1, 2,\dots, n)$, where Φ_i is the i-th column vector of $\Phi_{(x)}$. Put

$$S(x) = \Phi_{(x)}G(x)^{-1}, \tag{2.11}$$

if $G(x)$ is nonsingular.

For any $x\in D$ let $E(x)$ and $H(x)$ be the elements of $L(C[J], C[J])$ defined by

$$E(x)h = \int_{t_0}^{t} \Phi_{(x)}(t)\Phi_{(x)}^{-1}(s)h(s)ds, \tag{2.12}$$

$$H(x)h = (I - S(x)l(x))E(x)h \qquad \text{for} \quad h \in C[J], \tag{2.13}$$

and let $T(x)$: $D \to C[J]$ be the operator such that

$$T(x)\varphi = X(\varphi(t), t) - A(x)\varphi(t) \qquad \text{for} \quad \varphi \in D, \tag{2.14}$$

where

$$A(x)\varphi = A(x(t), t)\varphi(t). \tag{2.15}$$

We have the following

LEMMA 1. *For any $x \in D$ $L(x)$ has an inverse operator $L(x)^{-1}$ if and only if*

$$\det G(x) \neq 0. \tag{2.16}$$

Suppose (2.16) *is satisfied. Then for any* $y=(u, e) \in B$

$$L(x)^{-1}y = H_1(x)u + H_2(x)e, \tag{2.17}$$

so that

$$\|L(x)^{-1}\|_c \leqq \|H_1(x)\|_c + \|H_2(x)\|_c, \tag{2.18}$$

where

$$H_1(x) = I + H(x)A(x) - S(x)l(x), \tag{2.19}$$

$$H_2(x) = S(x). \tag{2.20}$$

PROOF. By (2.8), for any $y=(u, e) \in B$, the equation $L(x)h=y$ is equivalent to the system

$$P(x)h = u, \tag{2.21}$$

$$l(x)h = e. \tag{2.22}$$

The general solution of (2.21) is given by

$$h(t) = \Phi_{(x)}(t)c + u(t) + \int_{t_0}^{t} \Phi_{(x)}(t)\Phi_{(x)}^{-1}(s)A(x(s), s)u(s)ds, \tag{2.23}$$

where c is an arbitrary constant n-vector. By (2.10) and (2.12) the substitution of (2.23) into (2.22) yields

$$G(x)c + l(x)(I + E(x)A(x))u = e.$$

Hence $L(x)^{-1}$ exists and is unique if and only if c is determined uniquely for any $(u, e) \in B$, that is, $\det G(x) \neq 0$.

If (2.16) holds, then it follows that

$$c = G(x)^{-1}e - G(x)^{-1}l(x)(I + E(x)A(x))u.$$

Substituting this into (2.23), by (2.19), (2.20) and the definition of H we have (2.17) and the inequality

$$\|L(x)^{-1}y\|_c \leqq \|H_1(x)\|_c \|u\|_c + \|H_2(x)\|_c \|e\|$$
$$\leqq (\|H_1(x)\|_c + \|H_2(x)\|_c) \|y\|_b,$$

which implies (2.18).

COROLLARY. *Under the condition* (2.16) *let*

$$K(x) = I - L(x)^{-1}F \qquad \text{for} \quad x \in D. \tag{2.24}$$

Then

$$K(x)\varphi = \{H(x)T(x) + S(x)(l(x) - f)\}\varphi \qquad \text{for} \quad \varphi \in D. \tag{2.25}$$

PROOF. Substituting $u = Q\varphi$ and $e = f[\varphi]$ into (2.17), we have from (2.20)

$$H_2(x)e = S(x)f[\varphi]$$

and from (2.19) by the integration by parts

$$H_1(x)u = \{I - H(x)T(x) - S(x)l(x)\}\varphi.$$

Hence

$$L(x)^{-1}F\varphi = \{I - H(x)T(x) - S(x)(l(x) - f)\}\varphi,$$

which completes the proof.

If $L(x)$ is independent of x, so also are $K(x)$, $\Phi_{(x)}$, $G(x)$, etc.. In such a case we write these operators and matrices simply as L, K, Φ, G, etc. respectively.

By Lemma 1, its corollary and the contraction mapping theorem [4, pp. 65–66] we have the following

THEOREM 1. *Let $x_0 \in D$ be an approximate solution of* (2.4) *and suppose there exist a bounded linear operator L, an operator K, a positive number δ and nonnegative constants η, κ ($\kappa < 1$) such that*

(i) $\det G \neq 0$,

(ii) $D_\delta = \{x \in C[J] \mid \|x - x_0\|_c \leqq \delta\} \subset D$,

(iii) $\|Kx - Ky\|_c \leqq \kappa\|x - y\|_c$ *for all* $x, y \in D_\delta$,

(iv) $\|L^{-1}Fx_0\|_c \leqq \eta$,

(v) $\lambda = \eta/(1 - \kappa) \leqq \delta$.

Then the sequence $\{x_k\}$ *defined by*

$$x_{k+1} = Kx_k \qquad (k = 0, 1,\dots) \tag{2.26}$$

converges to $\hat{x} \in D_\delta$ *as* $k \to \infty$. $\hat{x}$ *is the unique solution of* (2.4) *in* D_δ, *and*

$$\|\hat{x} - x_k\|_c \leqq \kappa^k \lambda \qquad (k = 0, 1,\dots). \tag{2.27}$$

COROLLARY. *Suppose there exists a positive constant* M *such that*

$$\|X_x(x(t), t) - A(t)\|_c \leqq \kappa/M,$$

$$\|f'(x) - l\| \leqq \kappa/M \qquad \textit{for all} \quad x \in D_\delta,$$

$$\|H\|_c + \|S\|_c \leqq M.$$

Then the condition (iii) *is satisfied.*

PROOF. For any $x, y \in D_\delta$ by Corollary to Lemma 1 we have

$$Kx - Ky = H[Tx - Ty] + S(l[x - y] - f[x] + f[y]).$$

Let $h = x - y$. Then by the mean value theorem we have

$$Kx - Ky = H\left[\int_0^1 \{X_x(y(t) + \theta h(t), t) - A(t)\} h(t) d\theta\right]$$
$$+ S\int_0^1 \{l - f'(y + \theta h)\} h d\theta.$$

Since $y + \theta h \in D_\delta$, from the assumption it follows that

$$\|X_x(y(t) + \theta h(t), t) - A(t)\|_c \leqq \kappa/M,$$

$$\|l - f'(y + \theta h)\| \leqq \kappa/M \qquad \text{for all} \quad \theta \in [0, 1],$$

and so

$$\|Kx - Ky\|_c \leqq (\|H\|_c + \|S\|_c)(\kappa/M)\|h\|_c$$
$$\leqq \kappa\|x - y\|_c.$$

Now we prove the following

THEOREM 2. *Let $x_0 \in D$ be an approximate solution of* (2.4) *and suppose there exist a bounded linear operator* $L(x)$ $(x \in D)$, $z_0 \in D_\sigma$, *positive numbers* δ, M *and nonnegative constants* η, μ, κ $(\kappa < 1)$ *such that*

(i) $D_\delta = \{x \in C[J] \mid \|x - x_0\|_c \leqq \delta\} \subset D$,

(ii) $\det G(x) \neq 0$ *for all* $x \in D_\delta$,

(iii) $\|X_x(x(t), t) - A(x_0(t), t)\|_c \leqq \kappa/M$,

$\|f'(x) - l(x_0)\| \leqq \kappa/M$ *for all* $x \in D_\delta$,

(iv) $\|H(x)\|_c + \|S(x)\|_c \leqq M$ *for all* $x \in D_\delta$,

(v) $\|L(z_0)^{-1} F x_0\|_c \leqq \eta$,

(vi) $\|A(x(t), t) - A(y(t), t)\|_c \leqq \mu\|x - y\|_c$,

$\|l(x) - l(y)\| \leqq \mu\|x - y\|_c$ *for all* $x, y \in D_\delta$,

(vii) $\beta = M\mu\lambda/(1 - \kappa) \leqq 1/4$,

(viii) $\sigma = \eta/(1 - \alpha) \leqq \delta$,

where

(2.28) $$\alpha = \{1 + \kappa - (1 - \kappa)(1 - 4\beta)^{1/2}\}/2,$$

(2.29) $$\lambda = \eta/(1 - \kappa),$$

$$D_\sigma = \{x \in C[J] \mid \|x - x_0\|_c \leqq \sigma\}.$$

Then for any sequence $\{z_k\}$ $(z_k \in D_\sigma)$ *the sequence* $\{x_k\}$ *defined by*

(2.30) $$x_{k+1} = K(z_k)x_k \qquad (k = 0, 1, \dots)$$

remains in D_σ *and converges to* $\hat{x} \in D_\sigma$ *as* $k \to \infty$. $\hat{x}$ *is the unique solution of* (2.4) *in* D_δ, *and*

(2.31) $$\|\hat{x} - x_k\|_c \leqq \alpha^k \sigma \qquad (k = 0, 1, \dots).$$

PROOF. Since $0 \leqq \kappa < 1$, by (vii) we have

(2.32) $$\kappa \leqq \alpha < 1,$$

(2.33) $$\alpha = \kappa + M\mu\sigma.$$

By (ii) and Lemma 1, $L(z)$ $(z \in D_\sigma)$ has an inverse operator $L(z)^{-1}$. Hence $K(z)$ is defined for any $z \in D_\sigma$ and, if $z_k \in D_\sigma$ and $x_k \in D$, then we have

(2.34) $$K(z_k)x_k = x_k - L(z_k)^{-1}Fx_k.$$

We show first that, if $x_k, x_{k+1} \in D_\sigma$, then

(2.35) $$\|x_{k+2} - x_{k+1}\|_c \leqq \alpha\|x_{k+1} - x_k\|_c.$$

Let $h_l = x_{l+1} - x_l$ $(l=k, k+1)$. Then

(2.36) $$\begin{aligned} h_{k+1} &= -L(z_{k+1})^{-1}Fx_{k+1} \\ &= -L(z_{k+1})^{-1}[Fx_{k+1} - Fx_k - L(z_k)h_k], \end{aligned}$$

because by (2.34)

(2.37) $$Fx_k = -L(z_k)h_k.$$

By the mean value theorem we have

(2.38) $$\begin{aligned} Fx_{k+1} - Fx_k - L(z_k)h_k &= \Big(-\int_{t_0}^{t}\{X(x_{k+1}(s), s) - X(x_k(s), s) \\ &\quad - A(z_k(s), s)h_k(s)\}ds, f[x_{k+1}] - f[x_k] - l(z_k)h_k\Big) \\ &= -(u_k, e_k), \end{aligned}$$

where

$$u_k(t) = \int_{t_0}^{t} v_k(s)ds,$$

(2.39) $$v_k(s) = \int_0^1 \{X_x(x_k(s) + \theta h_k(s), s) - A(z_k(s), s)\}h_k(s)d\theta,$$

(2.40) $$e_k = \int_0^1 \{l(z_k) - f'(x_k + \theta h_k)\}h_k d\theta.$$

From (iii) and (vi) it follows that

$$\|X_x(x_k(t) + \theta h_k(t), t) - A(x_0(t), t)\|_c \leqq \kappa/M \qquad \text{for all} \quad \theta \in [0, 1],$$

$$\|A(x_0(t), t) - A(z_k(t), t)\|_c \leqq \mu\|x_0 - z_k\|_c,$$

which yield from (2.39)

(2.41) $$\|v_k\|_c \leqq (\kappa/M + \mu\|x_0 - z_k\|_c)\,\|h_k\|_c.$$

Similarly from (2.40) we have

(2.42) $$\|e_k\| \leqq (\kappa/M + \mu\|x_0 - z_k\|_c)\,\|h_k\|_c.$$

By (2.36), (2.38) and Lemma 1

$$h_{k+1} = H_1(z_k)u_k + S(z_k)e_k \tag{2.43}$$

and by the integration by parts we have

$$H_1(z_k)u_k = H(z_k)v_k.$$

Hence (2.43) is expressed as

$$h_{k+1} = H(z_k)v_k + S(z_k)e_k. \tag{2.44}$$

By (2.44), (2.41), (2.42) and (iv) we have

$$\|h_{k+1}\|_c \leqq (\kappa + M\mu\|x_0 - z_k\|_c)\,\|h_k\|_c. \tag{2.45}$$

Since $\|x_0 - z_k\|_c \leqq \sigma$, by (2.33)

$$\kappa + M\mu\|x_0 - z_k\|_c \leqq \kappa + M\mu\sigma = \alpha,$$

and (2.35) follows from this and (2.45).

By (v) and (viii)

$$\|x_1 - x_0\|_c = \|L(z_0)^{-1}Fx_0\|_c \leqq \eta = (1 - \alpha)\sigma \leqq \sigma,$$

so $x_1 \in D_\sigma$ and it follows from (2.35) that

$$\|x_2 - x_1\|_c \leqq \alpha\eta.$$

It can be shown by induction that

$$\|x_{k+1} - x_k\|_c \leqq \alpha^k\eta,$$

$$\|x_k - x_0\|_c \leqq (1 - \alpha^k)\sigma \qquad (k = 1, 2, \ldots),$$

so that for any integer $p \geqq 0$ we have

$$\|x_{k+p} - x_k\|_c \leqq \alpha^k\eta(1 - \alpha^p)/(1 - \alpha) \leqq \alpha^k\sigma. \tag{2.46}$$

Hence $\{x_k\}$ is a Cauchy sequence in D_σ and its limit $\hat{x}$ exists in D_σ, because D_σ is a closed set. Since by (2.37) and (vi)

$$\begin{aligned}\|Fx_k\|_b &\leqq \|L(z_k)\|_b\|x_{k+1} - x_k\|_c \\ &\leqq (\|L(z_k) - L(z_0)\|_b + \|L(z_0)\|_b)\,\|x_{k+1} - x_k\|_c,\end{aligned}$$

$$\|L(z_k) - L(z_0)\|_b \leqq 2\mu\sigma\max(b - a, 1),$$

it follows that $\|Fx_k\|_b \to 0$ as $k \to \infty$ and $F\hat{x} = 0$ by the continuity of F. Hence $\hat{x}$ is a solution of (2.4), and the estimate (2.31) follows from (2.46).

Now we consider the iterative method

$$x_{k+1} = K(z_0)x_k \qquad (k = 0, 1,\dots).$$

Since by (iii), (vi), (iv) and (2.33)

$$\|X_x(x(t), t) - A(z_0(t), t)\|_c \leqq \alpha/M,$$

$$\|f'(x) - l(z_0)\| \leqq \alpha/M \qquad \text{for all} \quad x \in D_\delta,$$

$$\|H(z_0)\|_c + \|S(z_0)\|_c \leqq M,$$

by Corollary to Theorem 1 we have

$$\|K(z_0)x - K(z_0)y\|_c \leqq \alpha\|x - y\|_c \qquad \text{for all} \quad x, y \in D_\delta$$

and by Theorem 1 (2.4) has a unique solution in D_δ. Hence $\hat{x}$ is the unique solution of (2.4) in D_δ. This completes the proof.

We note that the choice $z_k = x_0$ $(k=0, 1,\dots)$ yields the estimate

$$\|\hat{x} - x_0\|_c \leqq \lambda \tag{2.47}$$

and that the choice $z_k = x_k$ $(k=0, 1,\dots)$ is also possible.

COROLLARY 1. *Let* $c_k = x_k(t_0)$ $(k=0, 1,\dots)$. *Then under the assumptions of the theorem* x_{k+1} *and* c_{k+1} *can be written as*

$$x_{k+1} = H(z_k)T(z_k)x_k + S(z_k)(l(z_k)x_k - f[x_k]), \tag{2.48}$$

$$c_{k+1} = G(z_k)^{-1}(l(z_k)[x_k - u_k] - f[x_k]), \tag{2.49}$$

where

$$u_k = E(z_k)T(z_k)x_k.$$

PROOF. The formula (2.48) follows from Corollary to Lemma 1. Setting $t = t_0$ in (2.48), we have (2.49) because

$$H(z_k)T(z_k)x_k = u_k - S(z_k)l(z_k)u_k, \quad u_k(t_0) = 0.$$

COROLLARY 2. *Let* $\hat{c} = \hat{x}(t_0)$. *Then under the assumptions of the theorem with* $z_0 = x_0$

$$\|\hat{c} - c_0\| \leqq \min(\lambda, M_1\lambda\kappa/M + \eta_1), \tag{2.50}$$

where η_1 *and* M_1 *are nonnegative numbers such that*

$$\|Rx_0 - c_0\| \leqq \eta_1, \tag{2.51}$$

$$(2.52) \qquad \|G(x_0)^{-1}\|(1+\|l(x_0)E(x_0)\|) \leqq M_1,$$

and

$$(2.53) \qquad R = G(x_0)^{-1}\{l(x_0) - f - l(x_0)E(x_0)T(x_0)\}.$$

PROOF. Set $z_k = x_0$ $(k=0, 1,\dots)$ in (2.49). Then

$$(2.54) \qquad c_{k+1} = Rx_k \qquad (k = 0, 1,\dots).$$

Since $c_k \to \hat{c}$ as $k \to \infty$, by the continuity of f and $X(x, t)$, from (2.54) it follows that $\hat{c} = R\hat{x}$. Since $\hat{x}$, $x_0 \in D_\delta$, by (iii) and the mean value theorem we have

$$\|l(x_0)[\hat{x} - x_0] - f[\hat{x}] + f[x_0]\| \leqq (\kappa/M)\|\hat{x} - x_0\|_c,$$

$$\|T(x_0)\hat{x} - T(x_0)x_0\|_c \leqq (\kappa/M)\|\hat{x} - x_0\|_c.$$

Hence

$$\|R\hat{x} - Rx_0\| \leqq \|G(x_0)^{-1}\|(1 + \|l(x_0)E(x_0)\|)(\kappa/M)\|\hat{x} - x_0\|_c,$$

and by (2.51) and (2.52)

$$(2.55) \qquad \begin{aligned} \|\hat{c} - c_0\| &\leqq \|R\hat{x} - Rx_0\| + \|Rx_0 - c_0\| \\ &\leqq M_1(\kappa/M)\|\hat{x} - x_0\|_c + \eta_1. \end{aligned}$$

Since by (2.47)

$$(2.56) \qquad \|\hat{c} - c_0\| \leqq \|\hat{x} - x_0\|_c \leqq \lambda,$$

(2.50) follows from (2.55) and (2.56). This completes the proof.

A solution $\hat{x} \in D$ of (2.4) is said to be isolated if the Fréchet derivative $F'(\hat{x})$ of F at $\hat{x}$ has an inverse operator. With this terminology we have the following

THEOREM 3. *The solution $x = \hat{x}$ obtained in Theorem 1 is an isolated solution.*

PROOF. Let $\hat{\Phi}(t)$ be a fundamental matrix of the system

$$\frac{dy}{dt} = X_x(\hat{x}(t), t)y.$$

Then by Lemma 1 $F'(\hat{x})$ has an inverse operator if and only if $\det f'(\hat{x})[\hat{\Phi}] \neq 0$.

Suppose $\hat{x}$ is not an isolated solution. Then there exists a nonzero $\hat{e} \in R^n$ such that

$$(2.57) \qquad f'(\hat{x})[\hat{\Phi}]\hat{e} = 0.$$

Put $\hat{h} = \hat{\Phi}\hat{e}$. Then by the definition of $\hat{\Phi}$ we have

$$Q'(\hat{x})\hat{h} = \hat{h}(t) - \hat{h}(t_0) - \int_{t_0}^{t} X_x(\hat{x}(s), s)\hat{h}(s)ds$$

$$= \left\{\hat{\Phi}(t) - \hat{\Phi}(t_0) - \int_{t_0}^{t} X_x(\hat{x}(s), s)\hat{\Phi}(s)ds\right\}\hat{e} = 0,$$

and by (2.57)

$$f'(\hat{x})[\hat{h}] = f'(\hat{x})[\hat{\Phi}]\hat{e} = 0.$$

Hence $F'(\hat{x})\hat{h}=0$ and we have $\hat{h}=(I-L^{-1}F'(\hat{x}))\hat{h}$.

On the other hand, by (2.24)

$$K'(\hat{x}) = I - L^{-1}F'(\hat{x}).$$

Since $\|K'(\hat{x})\|_c \leqq \kappa$ by (iii) in Theorem 1, we have $\|\hat{h}\|_c \leqq \kappa \|\hat{h}\|_c$, which implies $\hat{h}(t)=0$ because $0 \leqq \kappa < 1$. Since $\det \hat{\Phi}(t) \neq 0$ $(t \in J)$, it follows that $\hat{e}=0$, which is a contradiction and the proof is complete.

It is well known [7] that an isolated solution $\hat{x} \in D$ of the problem (2.1), (2.2) is locally unique; that is, no other solution exists in a sufficiently small neighborhood of $\hat{x}$.

In the following two sections the results in this section are applied to special boundary value problems.

3. Multipoint boundary value problems

In this section we are concerned with the multipoint boundary condition of the form:

(3.1) $$f[x] \equiv g(x(t_0), x(t_1), \ldots, x(t_N)) = 0,$$

where

(3.2) $$a = t_0 < t_1 < \cdots < t_{N-1} < t_N = b.$$

3.1. Case of nonlinear conditions

Let

(3.3) $$D_i = \{x(t_i) \mid x \in D\} \subset R_i^n \qquad (i = 0, 1, \ldots, N)$$

and let $g: D_0 \times D_1 \times \cdots \times D_N \to R^n$ be continuously Fréchet differentiable. Then

(3.4) $$f'(x)h = \sum_{i=0}^{N} B_i(x)h(t_i) \qquad \text{for} \quad h \in C[J],$$

where

(3.5) $$B_i(x) = \frac{\partial g}{\partial u_i}(x(t_0), x(t_1), \dots, x(t_N)) \qquad (i = 0, 1, \dots, N),$$

and u_i is the space variable in R_i^n.

Let us choose $l(x)=f'(x)$ in (2.8). Then

(3.6) $$G(x) = \sum_{i=0}^{N} B_i(x)\Phi_{(x)}(t_i),$$

and we have the following

LEMMA 2. *Let* $l(x)=f'(x)$. *Then*

(3.7) $$H(x)h = \int_a^b H_{(x)}(t, s)h(s)ds \qquad \textit{for} \quad h \in C[J],$$

where for $t_{k-1} \leqq s < t_k$ $(k=1, 2, \dots, N)$

(3.8) $$H_{(x)}(t, s) = \begin{cases} \Phi_{(x)}(t)(I - M_k(x))\Phi_{(x)}^{-1}(s) & \textit{if} \quad s < t, \\ -\Phi_{(x)}(t)M_k(x)\Phi_{(x)}^{-1}(s) & \textit{if} \quad s \geqq t, \end{cases}$$

(3.9) $$M_k(x) = G(x)^{-1}\sum_{i=k}^{N} B_i(x)\Phi_{(x)}(t_i).$$

PROOF. By (3.4) and (2.13) $H(x)h$ can be written as

$$H(x)h = \int_{t_0}^{t} \Phi_{(x)}(t)v(s)ds - \Phi_{(x)}(t)G(x)^{-1}\sum_{i=1}^{N} B_i(x)\int_{t_0}^{t_i} \Phi_{(x)}(t_i)v(s)ds,$$

where $v(s)=\Phi_{(x)}^{-1}(s)h(s)$. Since

$$\sum_{i=1}^{N} B_i(x)\int_{t_0}^{t_i} \Phi_{(x)}(t_i)v(s)ds = \sum_{k=1}^{N}\int_{t_{k-1}}^{t_k} \sum_{i=k}^{N} B_i(x)\Phi_{(x)}(t_i)v(s)ds,$$

we obtain (3.7).

By Lemma 1 we have the following

COROLLARY. *If* $\det G(x) \neq 0$, *then* $L(x)$ *has an inverse operator* $L(x)^{-1}$ *and for any* $y=(u, e) \in B$

(3.10) $$L(x)^{-1}y = u(t) + \int_a^b H_{(x)}(t, s)A(x(s), s)u(s)ds + S(x)(t)(e - l(x)u).$$

From this and Corollary to Lemma 1, it follows that for $\varphi \in D$

(3.11) $$K(x)\varphi = H(x)T(x)\varphi + S(x)(l(x)\varphi - f[\varphi]),$$

and Theorem 2 can be applied to the iterative method $x_{k+1}=K(x_k)x_k$ $(k=0, 1, \dots)$ to assure its convergence and to give the error estimate for x_k.

3.2. Case of linear conditions

Let us consider the case

$$f[x] \equiv l[x] - d = 0, \tag{3.12}$$

where

$$l[x] = \sum_{i=0}^{N} B_i x(t_i), \tag{3.13}$$

d is a constant n-vector and B_i ($i=0, 1,\ldots, N$) are constant $n \times n$ matrices. Then

$$G(x) = \sum_{i=0}^{N} B_i \Phi_{(x)}(t_i), \tag{3.14}$$

and by (3.11)

$$K(x)\varphi = H(x)T(x)\varphi + S(x)d \qquad \text{for} \quad \varphi \in D. \tag{3.15}$$

Now we are interested in the case $A(x, t)=A(t)$. The operator K is then defined by

$$K\varphi = HT\varphi + Sd \qquad \text{for} \quad \varphi \in D, \tag{3.16}$$

where

$$G = \sum_{i=0}^{N} B_i \Phi(t_i), \quad S = \Phi G^{-1}, \tag{3.17}$$

$$Hh = \int_a^b H(t, s)h(s)ds \qquad \text{for} \quad h \in C[J], \tag{3.18}$$

and for $t_{k-1} \leqq s < t_k$ ($k=1, 2,\ldots, N$)

$$H(t, s) = \begin{cases} \Phi(t)(I - M_k)\Phi^{-1}(s) & \text{if} \quad s < t, \\ -\Phi(t)M_k\Phi^{-1}(s) & \text{if} \quad s \geqq t, \end{cases} \tag{3.19}$$

$$M_k = G^{-1} \sum_{i=k}^{N} B_i \Phi(t_i). \tag{3.20}$$

In this case Theorem 1 yields the following

THEOREM 4. *Let $x_0 \in D$ be an approximate solution of the problem* (2.1), (3.12) *and suppose there exist a matrix $A(t)$ continuous on J, positive numbers δ, M and nonnegative constants η, κ ($\kappa < 1$) such that*

(i) $D_\delta = \{x \in C[J] \mid \|x - x_0\|_c \leqq \delta\} \subset D$,

(ii) $\det G \neq 0$,

(iii) $\|H\|_c \leqq M$, $\|X_x(x(t), t) - A(t)\|_c \leqq \kappa/M$ *for all* $x \in D_\delta$,

(iv) $\|r + HAr + S(l[x_0 - r] - d)\|_c \leqq \eta$,

(v) $\lambda = \eta/(1-\kappa) \leqq \delta$,

where $r = Qx_0$. *Then the sequence* $\{x_k\}$ *defined by*

$$x_{k+1} = Kx_k \qquad (k = 0, 1, \ldots) \tag{3.21}$$

remains in D_δ *and converges to* $\hat{x} \in D_\delta$ *as* $k \to \infty$. $\hat{x}$ *is the unique solution of the problem* (2.1), (3.12) *in* D_δ, *and*

$$\|\hat{x} - x_k\|_c \leqq \kappa^k \lambda \qquad (k = 0, 1, \ldots). \tag{3.22}$$

PROOF. For any $x, y \in D_\delta$ by (3.16) and the mean value theorem we have

$$Kx - Ky = H[Tx - Ty] = H\left[\int_0^1 T'(y + \theta h) h d\theta\right],$$

where $h = x - y$ and

$$T'(y + \theta h) = X_x(y(t) + \theta h(t), t) - A(t).$$

Since

$$\|T'(y + \theta h)\|_c \leqq \kappa/M \qquad \text{for all} \quad \theta \in [0, 1],$$

it follows that

$$\|Kx - Ky\|_c \leqq \|H\|_c(\kappa/M)\,\|x - y\|_c \leqq \kappa\|x - y\|_c.$$

By Corollary to Lemma 2 we have

$$L^{-1}Fx_0 = r + HAr + S(l[x_0 - r] - d)$$

and the conclusion of the theorem follows from Theorem 1.

In particular we study the cases $N=0$ and $N=1$.

In the case $N=0$, (3.12) is the initial condition $x(a)=d$, and the condition (iv) is expressed as follows:

$$\|r + HAr + \Phi(x_0(a) - d)\|_c \leqq \eta, \tag{3.23}$$

where $G = I$,

$$H(t, s) = \begin{cases} \Phi(t)\Phi^{-1}(s) & \text{if} \quad a \leqq s < t \leqq b, \\ 0 & \text{if} \quad a \leqq t \leqq s \leqq b. \end{cases} \tag{3.24}$$

The error estimate was obtained first by Fujii [2].

In the case $N=1$, (3.12) is the two point boundary condition $B_0x(a) + B_1x(b) = d$, and the condition (iv) becomes

$$\|r + HAr + S\{B_0x_0(a) + B_1x_0(b) - d - B_1r(b)\}\|_c \leqq \eta, \tag{3.25}$$

where $G=B_0+B_1\Phi(b)$, $M_1=G^{-1}B_1\Phi(b)$,

$$H(t, s) = \begin{cases} \Phi(t)(I - M_1)\Phi^{-1}(s) & \text{if} \quad a \leqq s < t \leqq b, \\ -\Phi(t)M_1\Phi^{-1}(s) & \text{if} \quad a \leqq t \leqq s \leqq b. \end{cases} \tag{3.26}$$

For the periodic boundary condition

$$x(a) = x(b), \tag{3.27}$$

the condition (iv) can be written as

$$\left\| r + HAr - S\int_a^b X(x_0(s), s)ds \right\|_c \leqq \eta, \tag{3.28}$$

where $G=I-\Phi(b)$,

$$H(t, s) = \begin{cases} S(t)\Phi^{-1}(s) & \text{if} \quad a \leqq s < t \leqq b, \\ S(t)\Phi(b)\Phi^{-1}(s) & \text{if} \quad a \leqq t \leqq s \leqq b. \end{cases} \tag{3.29}$$

In the case where $X(x, t)$ and $A(t)$ are periodic in t of period ω, suppose the problem (2.1), (3.27) has an approximate solution $x_0 \in D$ which satisfies the conditions of Theorem 4 with $a=0$ and $b=\omega$. Then $\hat{x}$ is the unique periodic solution of (2.1) with period ω in D_δ.

4. Boundary value problems of the least squares type

In this section let us assume that $X(x, t)$ is continuous in Ω and twice continuously differentiable with respect to x in Ω, and let $g: D \to R^m$ be twice continuously Fréchet differentiable in D. We consider the problem of finding a solution $x(t)$ of (2.1) which minimizes $(g[x])^*g[x]$ locally, where the symbol $*$ denotes the transpose of a matrix. Throughout this section we call this the problem (S) for simplicity.

4.1. Conditions for a local minimum

Let $x(t, c)$ be the solution of (2.1) on J such that $x(t_0, c)=c$, and let

$$\Delta = \{c \in R^n \mid x(t, c) \in D\}. \tag{4.1}$$

Then Δ is an open set in R^n.

Let $q: \Delta \to R^m$ and $s: \Delta \to R^1$ be defined by

$$q[c] = g[x(t, c)], \tag{4.2}$$

$$s[c] = (q[c])^*q[c]/2 \tag{4.3}$$

respectively and let $\Delta_0 \subset \Delta$ be a convex domain. Then we have the following

LEMMA 3. *For any* $c, c+e \in \Delta_0$,

$$s[c+e] = s[c] + s'(c)e + s''(c)ee/2 + U, \tag{4.4}$$

where

$$s'(c)e = (q[c])^* q'(c)e, \tag{4.5}$$

$$s''(c)ee = (q'(c)e)^* q'(c)e + (q[c])^* q''(c)ee, \tag{4.6}$$

$$|U| = o(\|e\|^2). \tag{4.7}$$

PROOF. It is easily verified that (4.5) and (4.6) hold. Since

$$U = \int_0^1 (1-\theta)\{s''(c+\theta e) - s''(c)\} ee d\theta, \tag{4.8}$$

it follows that

$$|U| \leqq \int_0^1 \|s''(c+\theta e) - s''(c)\| d\theta \|e\|^2 \tag{4.9}$$

and by the continuity of q, q' and q'' we obtain (4.7).

From this lemma we have

THEOREM 5. *Let* $\hat{c} \in \Delta_0$ *be a solution of* $s'(c)=0$ *and suppose there exists a positive constant* α *such that*

$$s''(\hat{c})ee \geqq \alpha \|e\|^2 \qquad \text{for all} \quad e \in R^n. \tag{4.10}$$

Then $s[c]$ *attains a local minimum at* $c = \hat{c}$.

COROLLARY. *Let* $\hat{c} \in \Delta_0$ *be a solution of* $s'(c)=0$ *and suppose*

$$\min_e \|q'(\hat{c})e\|_2^2 > \max_e |(q[\hat{c}])^* q''(\hat{c})ee| \tag{4.11}$$

for all $e \in R^n$ *with* $\|e\|_2 = 1$, *where* $\|\cdot\|_2$ *denotes the Euclidean norm. Then* $s[\hat{c}]$ *is a local minimum of* $s[c]$.

Since by Schwarz's inequality

$$|(q[\hat{c}])^* q''(\hat{c})ee| \leqq \|q[\hat{c}]\|_2 \|q''(\hat{c})\|_2 \|e\|_2^2, \tag{4.12}$$

(4.11) is satisfied if

$$\min_e \|q'(\hat{c})e\|_2^2 > \|q[\hat{c}]\|_2 \|q''(\hat{c})\|_2 \tag{4.13}$$

for all $e \in R^n$ with $\|e\|_2=1$. It seems that (4.13) is not an unreasonable condition if $\|q[\hat{c}]\|_2$ is small enough. We note that $\operatorname{rank} q'(\hat{c})=n$ if (4.13) is valid.

By Theorem 5 the problem (S) is reduced to finding a solution $x(t, \hat{c})$ of (2.1) which satisfies $s'(\hat{c})=0$ and (4.10).

For all $e \in R^n$ we have

$$q'(c)e = g'(x(t, c))u, \tag{4.14}$$

$$q''(c)ee = g''(x(t, c))uu + g'(x(t, c))[x_{cc}(t, c)ee], \tag{4.15}$$

where $u=x_c(t, c)e$, and x_c and x_{cc} are the first and the second Fréchet derivatives of $x(t, c)$ with respect to c respectively. From the assumption on $X(x, t)$ it follows that $x_c(t, c)$ is the fundamental matrix of the system

$$\frac{dy}{dt} = X_x(x(t, c), t)y$$

satisfying $x_c(t_0, c)=I$ and that $x_{cc}(t, c)$ is the solution of the system

$$\frac{dz}{dt}ee = X_x(x(t, c), t)zee + X_{xx}(x(t, c), t)uu$$

satisfying $x_{cc}(t_0, c)=0$, where $u=x_c(t, c)e$ and $X_{xx}(x, t)$ is the second Fréchet derivative of $X(x, t)$ with respect to x.

Substituting (4.14) into (4.5), we have

$$s'(c)e = (g[x(t, c)])^* g'(x(t, c))[x_c(t, c)]e. \tag{4.16}$$

Since $s'(c)=0$ is equivalent to $(s'(c))^*=0$, $\hat{c}$ is a solution of the equation

$$(g'(x(t, c))[x_c(t, c)])^* g[x(t, c)] = 0. \tag{4.17}$$

Hence let

$$f[x] = (g'(x)[\Theta_{(x)}])^* g[x] \qquad \text{for} \quad x \in D. \tag{4.18}$$

Then the solution $x=x(t, \hat{c})$ of (2.1) and (4.17) satisfies also the equation $f[x]=0$, where $\Theta_{(x)}(t)$ is the fundamental matrix of the system

$$\frac{dy}{dt} = X_x(x(t), t)y$$

with $\Theta_{(x)}(t_0)=I$.

Conversely a solution $\hat{x}$ of (2.1) satisfying the condition

$$f[\hat{x}] = 0 \tag{4.19}$$

is a solution of the problem (S) if it satisfies the condition of Theorem 5. Thus we are led to consider the problem of finding a solution $\hat{x} \in D$ of (2.1) satisfying the condition (4.19), which we call the problem (P) for simplicity.

4.2. Iterative methods for solving problem (P)

We propose some iterative methods for solving the problem (P).

Let f be the operator given by (4.18). Then for x, $\varphi \in D$ and $h \in C[J]$

$$f'(x)h = (g'(x)[\Theta_{(x)}])^* g'(x)h + \{g''(x)[\Theta_{(x)}, h] + g'(x)[\Psi_{(x)}h]\}^* g[x], \tag{4.20}$$

$$K(x)\varphi = u_{(x,\varphi)} + S(x)\rho_{(x,\varphi)}, \tag{4.21}$$

where

$$[\Psi_{(x)}h](t) = \int_{t_0}^{t} \Theta_{(x)}(t)\Theta_{(x)}^{-1}(s)X_{xx}(x(s), s)h(s)\Theta_{(x)}(s)ds, \tag{4.22}$$

$$G(x) = l(x)[\Phi_{(x)}], \quad S(x) = \Phi_{(x)}G(x)^{-1}, \tag{4.23}$$

$$u_{(x,\varphi)} = E(x)T(x)\varphi, \tag{4.24}$$

$$\rho_{(x,\varphi)} = l(x)[\varphi - u_{(x,\varphi)}] - f[\varphi], \tag{4.25}$$

$g''(x)[\Theta_{(x)}, h]$ denotes the matrix whose column vectors are $g''(x)[\Theta_i, h]$ $(i=1, 2,\dots, n)$, Θ_i being the i-th column vector of $\Theta_{(x)}$. It has been shown in [1, 8] that $\left(\frac{d}{dx}\Theta_{(x)}\right)h = \Psi_{(x)}h$ is given by (4.22).

We consider the iterative method

$$x_{k+1} = K(x_k)x_k \qquad (k = 0, 1,\dots). \tag{4.26}$$

Various methods are obtained by the choice of $A(x, t)$ and $l(x)$. Some typical examples are given below.

Case 1. $A(x, t) = X_x(x, t), \qquad l(x) = f'(x).$

The method (4.26) is nothing but the Newton method.

Case 2. $A(x, t) = X_x(x, t), \qquad l(x) = (g'(x)[\Phi_{(x)}])^* g'(x).$

The method in this case is the so-called Gauss-Newton method.

Case 3. $A(x, t) = A(t), \qquad l(x) = (g'(x)[\Phi])^* g'(x).$

When g is a linear operator, S is computed once for all.

Case 4. $A(x, t) = A(t), \qquad l(x) = (g_0[\Phi])^* g_0,$

where g_0 is an operator approximating $g'(x)$. In practical applications it will be convenient usually to choose $A(t)=X_x(x_0(t), t)$ and $g_0=g'(x_0)$. The advantage of the method in this case lies in computing S only once.

Theorem 2 can be applied to the methods in Cases 1, 2 and 3 to assure their convergence and to give the error estimates; Theorem 1 can also be applied to that in Case 4.

In particular we are concerned with the case

$$g[x] = g(x(t_0), x(t_1),\dots, x(t_N)), \tag{4.27}$$

where $g: D_0\times D_1\times\cdots\times D_N\to R^m$ is twice continuously Fréchet differentiable, and t_i and D_i $(i=0, 1,\dots, N)$ are given by (3.2) and (3.3) respectively. All summations are assumed to be taken from 0 to N. Let

$$C_{ij}(x) = \frac{\partial^2 g}{\partial u_i \partial u_j}(x(t_0), x(t_1),\dots, x(t_N)) \qquad (i, j=0, 1,\dots, N), \tag{4.28}$$

$$R(x) = \textstyle\sum_i B_i(x)\Phi_{(x)}(t_i), \quad V(x) = \sum_i B_i(x)\Theta_{(x)}(t_i),$$

where $B_i(x)$ $(i=0, 1,\dots, N)$ are given by (3.5). Then for $x\in D$ and $h\in C[J]$ we have

$$f[x] = V(x)^* g[x], \tag{4.29}$$

$$\begin{aligned} f'(x)h = {} & V(x)^*\textstyle\sum_i B_i(x)h(t_i) \\ & + \{\textstyle\sum_{i,j} h(t_i)^* C_{ij}(x)\Theta_{(x)}(t_j) + \sum_i B_i(x)[\Psi_{(x)}h](t_i)\}^* g[x]. \end{aligned} \tag{4.30}$$

In Case 2 we have

$$G(x) = R(x)^* R(x), \tag{4.31}$$

$$\rho_{(x,\varphi)} = R(x)^*\textstyle\sum_i B_i(x)\{\varphi(t_i) - u_{(x,\varphi)}(t_i)\} - f[\varphi]. \tag{4.32}$$

In Case 3 $G(x)$, $u_{(x,\varphi)}$ and $\rho_{(x,\varphi)}$ are obtained from (4.31), (4.24) and (4.32) respectively with $\Phi_{(x)}$ replaced by Φ.

Finally let us consider the case

$$g[x] = \textstyle\sum_i B_i x(t_i) - d, \tag{4.33}$$

where d is a constant m-vector and B_i $(i=0, 1,\dots, N)$ are constant $m\times n$ matrices. Then for $x\in D$ and $h\in C[J]$ we have

$$f[x] = (\textstyle\sum_i B_i\Theta_{(x)}(t_i))^* g[x], \tag{4.34}$$

$$f'(x)h = (\textstyle\sum_i B_i\Theta_{(x)}(t_i))^*\sum_i B_i h(t_i) + (\sum_i B_i[\Psi_{(x)}h](t_i))^* g[x]. \tag{4.35}$$

The iterative method in Case 2 has been given by Banks and Groome [1] and those in Case 1 and Case 3 have been obtained by Urabe [8].

References

[1] H. T. Banks and G. M. Groome, Jr., *Convergence theorems for parameter estimation by quasilinearization,* J. Math. Anal. Appl., **42** (1973), 91–109.

[2] M. Fujii, *An a posteriori error estimation of the numerical solution by step-by-step methods for systems of ordinary differential equations,* Bull. Fukuoka Univ. Ed. III, **23** (1973), 35–44.

[3] J. M. Ortega and W. C. Rheinboldt, *Iterative solution of nonlinear equations in several variables,* Academic Press, New York, 1970.

[4] L. B. Rall, *Computational solution of nonlinear operator equations,* Wiley, New York, 1969.

[5] M. Urabe, *Galerkin's procedure for nonlinear periodic systems,* Arch. Rational Mech. Anal., **20** (1965), 120–152.

[6] M. Urabe, *An existence theorem for multi-point boundary value problems,* Funkcial. Ekvac., **9** (1966), 43–60.

[7] M. Urabe, *The Newton method and its application to boundary value problems with nonlinear boundary conditions,* Proc. US-Japan Seminar on Differential and Functional Equations, pp. 383–410, Benjamin, New York, 1967.

[8] M. Urabe, *On the Newton method to solve problems of the least squares type for ordinary differential equations,* Mem. Fac. Sci. Kyushu Univ. Ser. A, **29** (1975), 173–183.

Department of Mathematics,
Faculty of Science,
Hiroshima University
and
Department of Mathematics,
Faculty of Science,
Toyama University