Parameter Estimation by Quasilinearization in Functional Differential Equations with State-Dependent Delays: a Numerical Study

Ferenc Hartung

Department of Mathematics and Computing, University of Veszprém
H-8201 Veszprém, P.O.Box 158, Hungary

Abstract

In this paper we study a parameter estimation method in functional differential equations using quasilinearization technique. We define the method and test its applicability in numerical examples. We estimate infinite dimensional parameters such as coefficient functions, delay functions and initial functions in state-dependent delay equations.

Key words: parameter estimation, state-dependent delays, quasilinearization

1 Introduction and Definition of the Scheme

Estimation of unknown parameters in various classes of differential equations, and in particular in functional differential equations (FDEs), has been investigated by many authors (see, e.g., [1], [2], [5]–[7], [13], [14], [16], [17], [19], [21]).

In this paper we consider the nonlinear state-dependent delay system

$$\dot{x}(t) = f\Big(t, x(t), x(t-\tau(t,x(t),\sigma)), \theta\Big), \qquad t \in [0,T] \tag{1}$$

with the associated initial condition

$$x(t) = \varphi(t), \qquad t \in [-r, 0]. \tag{2}$$

⋆ This research was partially supported by Hungarian National Foundation for Scientific Research Grant No. T031935 and Janos Bolyai Research Grant of the Hungarian Academy of Sciences.

Here $\theta \in \Theta$ and $\sigma \in \Sigma$ are parameters of the equation and the delay function, respectively, where Θ and Σ are normed linear spaces. In our examples the parameters will be functions, i.e., the parameter space will be infinite dimensional. We will consider the initial function φ as a parameter, too. We assume that the parameters $\gamma \equiv (\varphi, \sigma, \theta)$ are unknown, but there are measurements $X_0, X_1, \ldots, X_l$ of the solution at the points $t_0, t_1, \ldots, t_l$. Our goal is to find a parameter value which minimizes the least square cost function

$$\min J(\gamma) \equiv \sum_{i=0}^{l} (x(t_i; \gamma) - X_i)^2 \tag{3}$$

over the parameter space Γ (or over an admissible set of parameters). Denote this infinite dimensional minimization problem by P.

One standard approach used in the literature to solve this minimization problem reduces it to solving finite dimensional minimization problems:

Step 1) First take finite dimensional approximations of the parameters, γ^N, (i.e., $\gamma^N \in \Gamma^N \subset \Gamma$, $\dim \Gamma^N < \infty$, $\gamma^N \to \gamma$ as $N \to \infty$).

Step 2) Consider a sequence of approximate initial value problems ($\mathrm{IVP}_{M,N}$) corresponding to a discretization of IVP (1)-(2) for some fixed parameter $\gamma^N \in \Gamma^N$ with solutions $y^M(\cdot; \gamma^N)$ satisfying $y^M(t, \gamma^N) \to x(t, \gamma)$ as $N, M \to \infty$, uniformly on compact time intervals.

Step 3) Define the least square minimization problems ($P^{N,M}$) for each $N, M = 1, 2, \ldots$, i.e., find $\gamma^{N,M} \in \Gamma^N$, which minimizes the least squares fit-to-data criterion

$$J^{N,M}(\gamma^N) = \sum_{i=0}^{l} |y^M(t_i; \gamma^N) - X_i|^2, \qquad \gamma^N \in \Gamma^N.$$

Step 4) Assuming that the actual parameters belong to a compact subset of Γ, argue that the sequence of solutions, $\gamma^{N,M}$ ($N, M = 1, 2, \ldots$), of the finite dimensional minimization problems $P^{N,M}$ has a convergent subsequence with limit $\bar{\gamma} \in \Gamma$.

Step 5) Show that $\bar{\gamma}$ is the solution of the minimization problem P.

Note that Step 5 can be proved independently of the particular choice of the approximation schemes used in Step 1 and Step 2. This method was successfully used in [1], [2], [7] and [21] using spline-based approximation schemes in Step 2. Note that these schemes have no known extension even for the simplest classes of neutral equations. In the sequence of papers [13]–[16] and [19] we defined several versions of a numerical identification scheme and proved their theoretical convergence for a large class of FDEs including delay and neutral state-dependent FDEs. The methods were based on an approximation technique called approximation by equations with piecewise constant arguments, which was introduced for linear delay and neutral equations in [8] and was generalized for nonlinear delay and neutral state-dependent FDEs in [9] and [15], respectively.

The method of quasilinearization for parameter estimation was introduced for ODEs in [3] and was applied to identify finite dimensional parameters in FDEs in [5] and [6]. The idea is the following: take finite dimensional approximation of the parameters (if they are infinite dimensional) $\gamma^N = (\varphi^N, \sigma^N, \theta^N)$, and consider the corresponding IVP

$$\dot{x}^N(t) = f\Big(t, x^N(t), x^N(t - \tau(t, x^N(t), \sigma^N)), \theta^N\Big), \qquad t \in [0, T] \tag{4}$$

$$x^N(t) = \varphi^N(t), \qquad t \in [-r, 0]. \tag{5}$$

Minimize the least square cost function

$$\min J^N(\gamma^N) \equiv \sum_{i=0}^{l} (x^N(t_i; \gamma^N) - X_i)^2,$$

by a gradient-based method. Note that this requires the computation of the derivative of J^N with respect to the parameter γ^N, i.e., we have to be able to compute the derivative of the solution x^N of (4)-(5) with respect to parameters. This problem was studied, e.g., in [4], [10], [11], [20] for several classes of state-independent delay equations, and in [12] and [18] for state-dependent FDEs.

The algorithm of quasilinearization can be described as follows: Take a basis $\{e_1^N, \ldots, e_N^N\}$ for the finite dimensional subspace Γ^N of Γ, and let $c = (c_1, \ldots, c_N)^T$ be the coordinates of the parameter $\gamma^N \in \Gamma^N$ with respect to this basis, i.e., $\gamma^N = \sum_{i=1}^N c_i e_i^N$. Then we identify γ^N with the column vector c, and simply write $x^N(t; c)$ instead of $x^N(t; \gamma^N)$. We approximate the parameter vector c by the fixed point iteration described by the following equations:

$$c^{(k+1)} = g(c^{(k)}), \qquad k = 0, 1, \ldots, \tag{6}$$

$$g(c) = c - (D(c))^{-1} b(c) \tag{7}$$

$$D(c) = \sum_{i=0}^{l} M^T(t_i; c) M(t_i; c) \tag{8}$$

$$b(c) = \sum_{i=0}^{l} M^T(t_i; c)(x^N(t_i; c) - X_i) \tag{9}$$

$$M(t; c) = (M_1(t; c), \ldots, M_N(t; c)) \tag{10}$$

$$M_i(t; c) = \frac{\partial x^N}{\partial \gamma}(t; c) e_i^N. \tag{11}$$

This is exactly the same scheme that was used in [5] and [6] except that there the parameter space was finite dimensional, and the set $\{e_1^N,\ldots,e_N^N\}$ was the canonical basis of $\mathbb{R}^N$. In our case $\frac{\partial x^N}{\partial\gamma}$ is a linear functional defined on a function space, e.g., on a space of continuous functions, and $\frac{\partial x^N}{\partial\gamma}(t;c)e_i^N$ denotes the value of the linear functional applied to the function e_i^N. For the derivation of this method in the finite dimensional case we refer to [3].

In the next section we present several numerical examples which will illustrate that this method works for identifying infinite dimensional parameters, as well.

2 Numerical Examples

In all of the numerical examples presented below we approximate the functions by linear spline functions. Let $\alpha=\xi_1,\xi_2,\ldots,\xi_N=\beta$ be an equidistant mesh of an interval $[\alpha,\beta]$, and $\{e_1^N,\ldots,e_N^N\}$ in (11) be the "hat" functions corresponding to the mesh $\{\xi_1,\ldots,\xi_N\}$, i.e., e_i^N is the linear spline function with the property that $e_i^N(\xi_j)=0$ if $i\neq j$, and $e_i^N(\xi_i)=1$.

Example 1 Consider the linear delay equation

$$\dot{x}(t)=\theta(t)x(t-\sigma(t)),\qquad t\in[0,2] \tag{12}$$
$$x(t)=\varphi(t),\qquad t\in[-2,0]. \tag{13}$$

If we take

$$\sigma(t)=\begin{cases}2-t^2, & t\in[0,1],\\ 1, & t\in[1,2],\end{cases}\qquad \theta(t)=\begin{cases}-\frac{t}{t+1}, & t\in[0,1],\\ -\frac{1}{2}, & t\in[1,2],\end{cases}\qquad \varphi(t)=t^2 \tag{14}$$

as the parameters in (12)-(13), then the solution of the corresponding IVP is

$$x(t)=\begin{cases}-\frac{1}{5}t^5-\frac{1}{4}t^4+\frac{4}{3}t^3-4t+4\log(t+1), & t\in[0,1],\\ \frac{1}{60}t^6-\frac{3}{40}t^5-2t\log(t+1)-\frac{1}{24}t^4+\frac{7}{12}t^3 & \\ \qquad +\frac{83}{120}t-\frac{103}{24}+4\log(2), & t\in[1,2].\end{cases}$$

We used this function to generate measurements at the points $t_i=0.1i,\quad i=0,1,\ldots,20$. First let σ and θ be defined by (14) and consider φ as a parameter in the equation. The derivative of the solution $x(t;\varphi)$ of IVP (12)–(13) with respect to the initial function φ satisfies the variational equation

$$\dot{z}(t;\varphi,\xi) = \theta(t)z(t-\sigma(t);\varphi,\xi), \qquad t \in [0,2] \tag{15}$$

$$z(t;\varphi,\xi) = \xi(t), \qquad t \in [-2,0], \tag{16}$$

where $z(t;\varphi,\xi) = \frac{\partial x}{\partial \varphi}(t;\varphi)\xi$ denotes the derivative applied to the function ξ. This IVP was solved numerically by the approximation technique of [8] to obtain the derivative values used in (11). Then we computed one iteration of (6)–(11) starting from the constant 2 initial parameter value. The numerical results can be seen in Figure 1 using $N = 3$ and $N = 9$ dimensional spline approximations of the initial function. The solid curve represents the "true" initial function. We got the following values for the cost functions: $J^3(c^{(0)}) = J^9(c^{(0)}) = 57.574144$, $J^3(c^{(1)}) = 0.000204$ and $J^9(c^{(1)}) = 0.000001$. In this linear equation $x(t;\varphi)$ depends linearly on φ, therefore $M(t;c)$ defined by (10)-(11) is constant in c. Hence b is linear, and D is the derivative of b. Therefore iteration (6) converges in one step, since it is the Newton-iteration for finding the zero of the linear function b. We can observe that the first step gives a good approximation of the identification problem: the shape of the initial function is well approximated, and the corresponding solution fits well to the measurements.

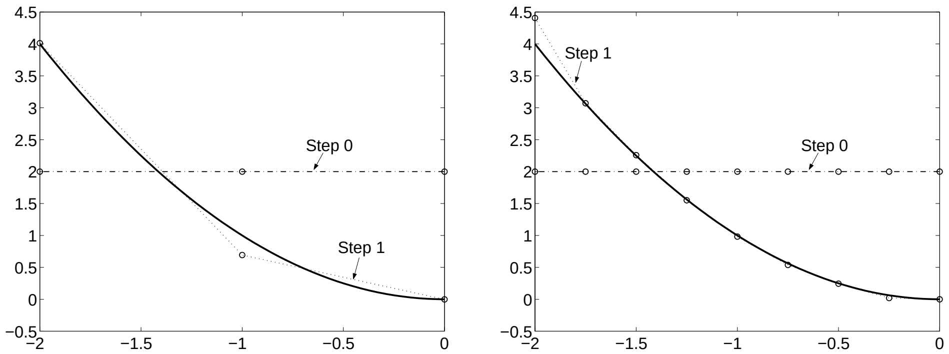

Fig. 1. Estimation of φ in IVP (12)-(13): $N = 3$ and $N = 9$.

Example 2 In this example we consider again IVP (12)-(13), but here we assume that the coefficient function θ is unknown, and the delay function σ and the initial function φ are given by (14). We used again the measurement of Example 1. The derivative of the solution with respect to θ can be computed by solving the IVP

$$\dot{z}(t;\theta,\xi) = \theta(t)z(t-\sigma(t);\theta,\xi) + \xi(t)x(t-\sigma(t);\theta), \quad t \in [0,2]$$

$$z(t;\theta,\xi) = 0, \qquad t \leq 0,$$

where $z(t;\theta,\xi) = \frac{\partial x}{\partial \theta}(t;\theta)\xi$. We applied method (6)–(11) for the constant 1 starting value. The first 3 steps of the numerical results can be seen in Figure 2. We observe fast convergence to the true parameter value. At the third step the cost function was $J^3(c^{(3)}) = 0.000170$ and $J^9(c^{(3)}) = 0.000001$, respectively.

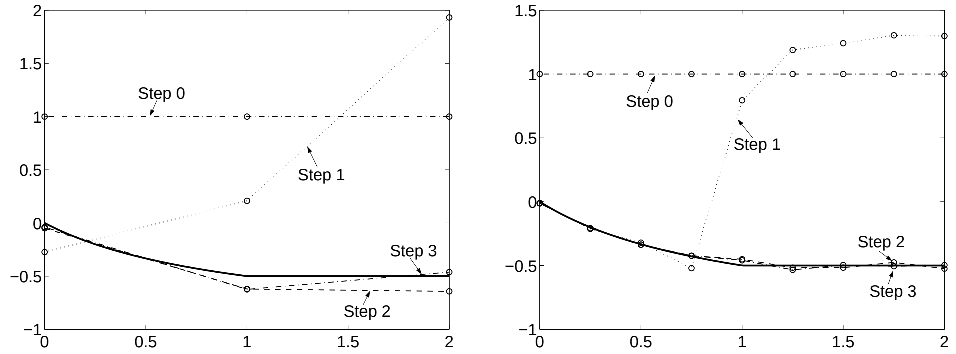

Fig. 2. Estimation of θ in IVP (12)-(13): $N = 3$ and $N = 9$.

Example 3 Consider again IVP (12)-(13). Here we assume that the delay function σ is unknown, and the coefficient function θ and the initial function φ are defined by (14). We used the same measurement as in Examples 1 and 2. We have to compute the derivative of $x(t;\sigma)$ with respect to σ. This problem was studied for the case when σ is constant in [10] and [20] and in [12] and [18] for the case when σ is a function. Consider the following variational equation

$$
\begin{aligned}
\dot{z}(t;\sigma,\xi) &= -\theta(t)\dot{x}(t-\sigma(t);\sigma)\xi(t), \qquad t \in [0,2] \\
z(t;\sigma,\xi) &= 0, \qquad t \in [-2,0]
\end{aligned}
$$

where $z(t;\sigma,\xi)$ is the candidate for $\frac{\partial x}{\partial \sigma}(t;\sigma)\xi$. It is easy to check that $x(t;\sigma)$ is continuously differentiable with respect to t for any t and σ, therefore z is well-defined, and Corollary 2 of [12] implies that $\frac{\partial x}{\partial \sigma}(t;\sigma) = z(t;\sigma,\cdot)$ for any σ and t. We did our calculations starting from a constant 2 delay function. The numerical results are given in Figure 3. The value of the cost function was $J^3(c^{(3)}) = 0.000306$ and $J^9(c^{(4)}) = 0.000003$, respectively.

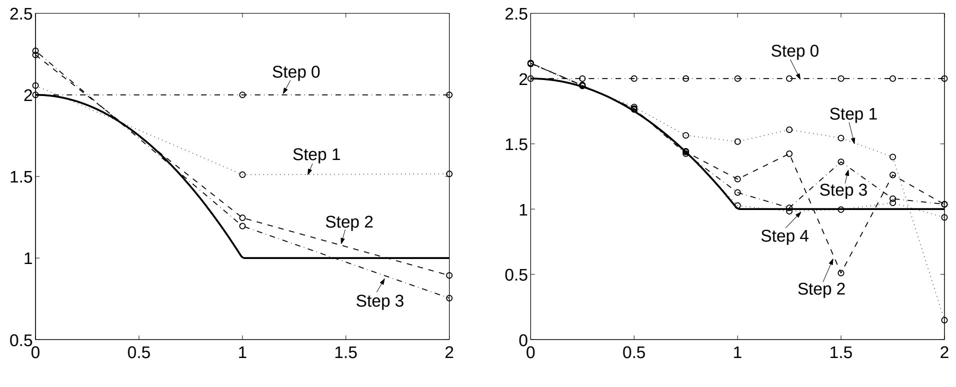

Fig. 3. Estimation of σ in IVP (12)-(13): $N = 3$ and $N = 9$

Example 4 In this example and in the following two examples we consider the state-dependent delay equation

$$\dot{x}(t) = \theta(t) x\left(t - \sigma^2(t) x^2(t)\right), \qquad t \in [0,2] \tag{17}$$

$$x(t) = \varphi(t), \qquad t \in [-1.5, 0], \tag{18}$$

where we choose

$$\theta(t) = -t, \qquad \sigma(t) = \frac{1}{t+1} \quad \text{and} \quad \varphi(t) = t^2 + 1 \tag{19}$$

as the "true parameters". Note that the state-dependent delay term is given by $\sigma^2(t)x^2(t)$. The analytic solution of this equation is difficult to compute, therefore we obtained the measurements by numerically solving IVP (17)–(19) at the points $t_i = 0.1i$, $i = 0, 1, \ldots, 20$. First we defined θ and σ by (19), and consider the initial function φ as an unknown parameter. To compute the derivative of the solution with respect to φ we used the variational equation

$$\begin{aligned} \dot{z}(t;\varphi,\xi) = {} & t\dot{x}\left(t - \frac{x^2(t;\varphi)}{(t+1)^2};\varphi\right)\frac{2x(t;\varphi)}{(t+1)^2} z(t;\varphi,\xi) \\ & - tz\left(t - \frac{x^2(t;\varphi)}{(t+1)^2};\varphi,\xi\right), \qquad t \in [0,2], \end{aligned} \tag{20}$$

$$z(t;\varphi,\xi) = \xi(t), \qquad t \in [-1.5, 0]. \tag{21}$$

Let $x(t;\varphi)$ denote the solution of (17)-(18) corresponding to initial function φ, and let $\bar{\varphi}$ be the "true" initial function, i.e., $\bar{\varphi}(t) = t^2+1$. Note that the solution $x(t;\bar{\varphi})$ is continuously differentiable for $t \geq -1.5$. For piecewise continuously differentiable initial functions we interpret $\dot{x}(t) = \dot{\varphi}(t)$ for $t \in [-1.5, 0]$ as the right derivative $D^+\varphi(t)$. It follows from Theorem 2 of [12] that for any $t \geq 0$ the function mapping $W^{1,\infty}([-1.5,0],\mathbb{R})$ into $\mathbb{R}$, $\varphi \mapsto x(t;\varphi)$ is differentiable at $\bar{\varphi}$, and the derivative is given by $z(t;\bar{\varphi},\cdot)$. However, this theorem does not yield that $z(t;\varphi,\cdot)$ is the derivative of the solution with respect to the initial function at any other φ, and certainly not at the finite dimensional approximations φ^N generated by the method. On the other hand, Corollary 6.3 of [18] yields that the function mapping $W^{1,\infty}([-1.5,0],\mathbb{R})$ into $W^{1,p}([0,2],\mathbb{R})$, $\varphi \mapsto x(\cdot;\varphi)$, for any p satisfying $1 \leq p < \infty$ is differentiable, and the derivative is given by(20)-(21). Despite this lack of theoretical proof of differentiability in the pointwise sense, iteration (6)–(11) works well for this case too. The results can be seen in Figure 4. The cost function at the last step was $J^3(c^{(2)}) = 0.000006$ and $J^9(c^{(2)}) = 0.000005$, respectively. The graph corresponding to $N = 9$ indicates that the "true" initial interval, i.e., the portion of the initial interval which is, in fact, used to compute the solution is smaller than $[-1.5, 0]$. Using the true parameter value we can see that it is $[-1, 0]$. For more detailed discussion about the identification of the "true" initial interval we refer the reader to [17] and [19].

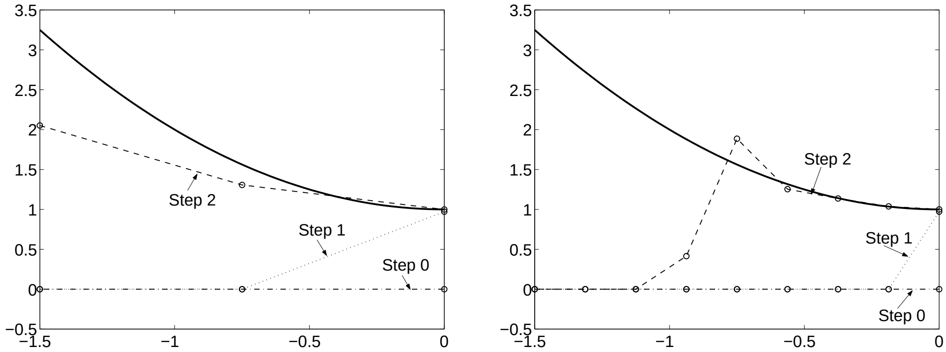

Fig. 4. Estimation of φ in IVP (17)-(18): $N = 3$ and $N = 9$.

Example 5 For IVP (17)-(18) consider θ as the unknown parameter, and let σ and φ be defined by (19). We used the same measurements as in Example 4. The derivative of the solution $x(t;\theta)$ with respect to θ is computed by

$$\begin{aligned}
\dot z(t;\theta,\xi) &= -\theta(t)\dot x\left(t-\frac{x^2(t;\theta)}{(t+1)^2};\theta\right)\frac{2x(t;\theta)}{(t+1)^2}z(t;\theta,\xi)\\
&\quad + \theta(t)z\left(t-\frac{x^2(t;\theta)}{(t+1)^2};\theta,\xi\right)+\xi(t)x\left(t-\frac{x^2(t;\theta)}{(t+1)^2};\theta\right),\ t\in[0,2]\\
z(t;\theta,\xi) &= 0, \qquad t\in[-1.5,0].
\end{aligned}$$

where we use $D^+x(0;\theta)$ instead of $\dot x(0;\theta)$ when x is not differentiable at 0. We have the same problem with this derivative as in Example 4, but, again, here we can also observe good convergence of our scheme to the true parameter (see Figure 5). We have $J^3(c^{(2)}) = 0.001721$ and $J^9(c^{(2)}) = 0.000725$.

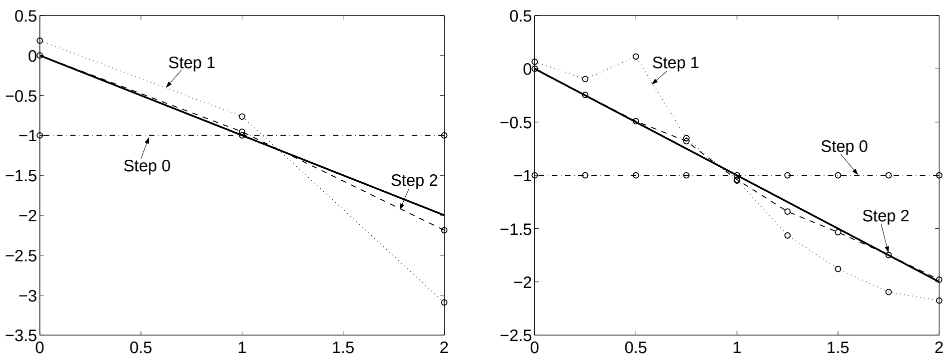

Fig. 5. Estimation of θ in IVP (17)-(18): $N = 3$ and $N = 9$.

Example 6 Finally, consider IVP (17)-(18) with σ as unknown, and let θ and φ be defined by (19). We used $z(t;\sigma,\cdot)$ defined by

$$\dot z(t;\sigma,\xi) = \theta(t)\Big\{-\dot x\big(t-\sigma^2(t)x^2(t;\sigma);\sigma\big)2\sigma^2(t)x(t;\sigma)z(t;\sigma,\xi)$$

$$
\begin{aligned}
&- \dot{x}\Big(t - \sigma^2(t)x^2(t;\sigma);\sigma\Big)2\sigma(t)\xi(t)x^2(t;\sigma) \\
&+ z\Big(t - \sigma^2(t)x^2(t;\sigma);\sigma,\xi\Big)\Big\}, \qquad t \in [0,2], \qquad (22)
\end{aligned}
$$

$$
z(t;\sigma,\xi) = 0, \qquad t \leq 0, \qquad (23)
$$

to generate the derivative of the solution with respect to σ. In this case a much weaker "pointwise differentiability" result can be proved than that of the previous two cases (see Theorem 3 in [12]), but still, the "derivative" generated by IVP (22)-(23) is good enough to produce nice approximations of the function σ (see Figure 6). We had $J^3(c^{(2)}) = 0.000008$ and $J^9(c^{(2)}) = 0.000018$, respectively, at the last step.

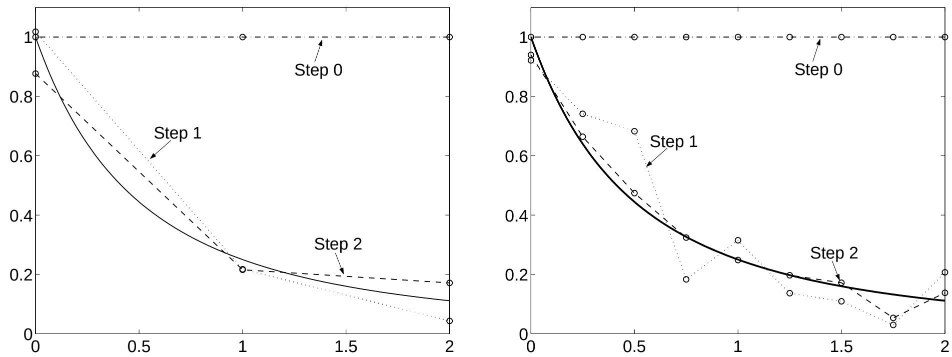

Fig. 6. Estimation of σ in IVP (17)-(18): $N = 3$ and $N = 9$.

References

[1] H. T. Banks, J. A. Burns and E. M. Cliff, Parameter estimation and identification for systems with delays, SIAM J. Control and Opt., 19:6 (1981) 791–828.

[2] H. T. Banks and P. K. Daniel Lamm, Estimation of delays and other parameters in nonlinear functional differential equations, SIAM J. Control and Opt., 21:6 (1983) 895–915.

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

[4] D. W. Brewer, The differentiability with respect to a parameter of the solution of a linear abstract Cauchy problem, SIAM. J. Math. Anal. Appl. 13:4 (1982) 6-7–620.

[5] D. W. Brewer, Quasi-Newton methods for parameter estimation in functional differential equations, Proc. 27th IEEE Conf. on Decision and Control, Austin, TX, (1988) 806–809.

[6] D. W. Brewer, J. A. Burns, E. M. Cliff, Parameter identification for an abstract Cauchy problem by quasilinearization, Quart. Appl. Math. 51:1 (1993) 1–22.

[7] J. A. Burns and P. D. Hirsch, A difference equation approach to parameter estimation for differential-delay equations, Appl. Math. Comp. 7 (1980) 281–311.

[8] I. Győri, On approximation of the solutions of delay differential equations by using piecewise constant arguments, Internat. J. of Math. & Math. Sci., 14:1 (1991) 111–126.

[9] I. Győri, F. Hartung and J. Turi, On numerical approximations for a class of differential equations with time- and state-dependent delays, Appl. Math. Letters, 8:6 (1995) 19–24.

[10] J. K. Hale, L. A. C. Ladeira, Differentiability with respect to delays, J. Diff. Eqns., 92 (1991) 14-26.

[11] J. K. Hale and S. M. Verduyn Lunel, Introduction to Functional Differential Equations, Spingler-Verlag, New York, 1993.

[12] F. Hartung, On differentiability of solutions with respect to parameters in a class of functional differential equations, Func. Diff. Eqns., 4:1-2 (1997) 65–79.

[13] F. Hartung, T. L. Herdman and J. Turi, Identifications of parameters in hereditary systems, Proceedings of ASME Fifteenth Biennial Conference on Mechanical Vibration and Noise, Boston, Massachusetts, September 1995, DE-Vol 84-3, Vol.3, Part C, 1061–1066.

[14] F. Hartung, T. L. Herdman and J. Turi, Identifications of parameters in hereditary systems: a numerical study, Proceedings of the 3rd IEEE Mediterranean Symposium on New Directions in Control and Automation, Cyprus, July 1995, 291–298.

[15] F. Hartung, T. L. Herdman, and J. Turi, On existence, uniqueness and numerical approximation for neutral equations with state-dependent delays, Appl. Numer. Math., 24 (1997) 393–409.

[16] F. Hartung, T. L. Herdman, and J. Turi, Parameter identification in classes of hereditary systems of neutral type, Appl. Math. and Comp., 89 (1998) 147–160.

[17] F. Hartung, T. L. Herdman, and J. Turi, Parameter identification in neutral functional differential equations with state-dependent delays, Nonlin. Anal., 39 (2000) 305–325.

[18] F. Hartung and J. Turi, On differentiability of solutions with respect to parameters in state-dependent delay equations, J. Diff. Eqns. 135:2 (1997) 192–237.

[19] F. Hartung and J. Turi, Identification of Parameters in Delay Equations with State-Dependent Delays, J. Nonlinear Analysis: Theory, Methods and Applications, 29:11 (1997) 1303–1318.

[20] V.-M. Hokkanen and G. Morosanu, Differentiability with respect to delay, Differential and Integral Equations, 11:4 (1998) 589–603.

[21] K. A. Murphy, Estimation of time- and state-dependent delays and other parameters in functional differential equations, SIAM J. Appl. Math., 50:4 (1990) 972–1000.