A NUMERICAL PROCEDURE FOR FILTERING AND EFFICIENT HIGH-ORDER SIGNAL DIFFERENTIATION

SALIM IBRIR*, SETTE DIOP**

* Department of Automated Production, École de Technologie Supérieure
1100, rue Notre Dame Ouest, Montreal, Québec, H3C 1K3 Canada
e-mail: s_ibrir@gpa.etsmtl.ca

** Laboratoire des Signaux et Systèmes, CNRS, Supélec, 3 rue Juliot-Curie
91190 Gif-sur-Yvette, France
e-mail: sette.diop@lss.supelec.fr

In this paper, we propose a numerical algorithm for filtering and robust signal differentiation. The numerical procedure is based on the solution of a simplified linear optimization problem. A compromise between smoothing and fidelity with respect to the measurable data is achieved by the computation of an optimal regularization parameter that minimizes the Generalized Cross Validation criterion (GCV). Simulation results are given to highlight the effectiveness of the proposed procedure.

Keywords: generalized cross validation, smoothing, differentiation, splines functions, optimization

1. Introduction

In many estimation and observation problems, estimating the unmeasured system dynamics turns on estimating the derivatives of the measured system outputs from discrete samples of measurements (Diop *et al.*, 1993; Gauthier *et al.*, 1992; Ibrir, 1999). A model of the signal dynamics may be of crucial help to achieve the desired objective. This has been magnificently demonstrated in pioneering works by R.E. Kalman (1960) and D.G. Luenberger (1971) for signals generated by known linear dynamical systems. Roughly speaking, if a signal model is known, then the resulting smooth signal can be differentiated with respect to time in order to have estimates of higher derivatives of the system output. For example, consider the problem of estimating $\nu - 1$ first derivatives, $y^{(i)},\ i = 0, 1, \ldots, \nu - 1$ of the output of a dynamic system, say, $y^{(\nu)} = f(y, \dot{y}, \ddot{y}, \ldots, y^{(\nu-1)})$, where y may be a vector, and f may contain input derivatives. But we choose not to go into technical details. If the nonlinear function f is known accurately enough, then asymptotic nonlinear observers can be designed using the results from (Ciccarella *et al.*, 1993; Gauthier *et al.*, 1992; Misawa and Hedrick, 1989; Rajamani, 1998; Tornambè, 1992; Xia and Gao, 1989). The proof of the asymptotic convergence of those observers requires various restrictive assumptions on the nonlinear function f. If f is not known accurately enough then, estimators for the derivatives of y may still be obtained via the theory of stabilization of uncertain systems, see, e.g., (Barmish and Leitmann, 1982; Chen, 1990; Chen and Leitmann, 1987; Dawson *et al.*, 1992; Leitmann, 1981). The practical convergence that is reached by the latter approach needs some matching conditions. We shall also mention the approach via sliding modes as in (Slotine *et al.*, 1987).

However, there are at least two practical situations where the available model is not of great help. First, the system model may be too poorly known. Second, it may be too complex for an extension of linear observer design theory. In those situations, and as long as practical (in lieu of asymptotic) convergence is enough for the specific application at hand, we may consider using differentiation estimators which merely ignore the nonlinear function f in their design. Differentiation estimators may be realized in both continuous time or discrete time as suggested in (Ibrir, 2001; 2003). This motivates enough the study, by observer design theorists, of more sophisticated numerical differentiation techniques for use in more involved control design problems. The numerical analysis literature is where to find the main contributions in the area, see (Anderson and Bloomfield, 1974; Craven and Wahba, 1979; De Boor, 1978; Eubank, 1988; Gasser *et al.*, 1985; Georgiev, 1984; Härdle, 1984; 1985; Ibrir, 1999; 2000; 2003; Müller, 1984; Reinsch, 1967; 1971) for more motivations and basic references. But these results have to be adapted to observer design problems since they were often envisioned so as to be used in an off-line basis.

The main difficulty that we face while designing differentiation observers, without any *a-priori* knowledge of system dynamics, is noise filtering. For this reason, robust signal differentiation can be classified as an ill-posed problem due to the conflicting goals that we aim to realize. Generally, noise filtering, precision, and the peaking phenomenon are three contradictory performances that characterize the robustness of any differentiation system.

The field of ill-posed problems has certainly been one of the fastest growing areas in signal processing and applied mathematics. This growth has largely been driven by the needs of applications both in other sciences and in industry. A problem is mathematically ill-posed if its solution does not exist, is not unique or does not depend continuously on the data. A typical example is the combined interpolation and differentiation problem of noisy data. A problem therein is that there are infinitely many ways to determine the interpolated function values if only the constraint from the data is used. Additional constraints are needed to guarantee the uniqueness of the solution to make the problem well posed. An important constraint in context is smoothness. By imposing a smoothness constraint, the analytic regularization method converts an ill-posed problem into a well-posed one. This has been used in solving numerous practical problems such as estimating higher derivatives of a signal through potentially noisy data.

As will be shown, inverse problems typically lead to mathematical models that are not well posed in Hadamard's sense, i.e., to ill-posed problems. Specifically, this means that their solutions is unstable under data perturbations. Numerical methods that can cope with this problem are the so-called regularization methods. These methods have been quite successfully used in the numerical analysis literature in approaches to the ill-posed problem of smoothing a signal from its discrete, potentially uncertain, samples (Anderson and Bloomfield, 1974; Craven and Wahba, 1979; Eubank, 1988; De Boor, 1978). One of these approaches proposed an algorithm for the computation of an optimal spline whose first derivatives are estimates of the first derivatives of the signal. These algorithms suffer from a large amount of computation they imply. One of the famous regularization criteria which have been extensively considered in numerical analysis and statistics (De Boor, 1978) is

$$\mathscr{J} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 + \lambda\int_0^t \hat{y}^{(m)}(s)\,\mathrm{d}s, \tag{1}$$

which embodies a compromise between the closeness to the measured data and smoothness of the estimate. The balance between the two distances is mastered by a particular choice of the parameter λ. It was shown that the minimum of the performance index (1) is a spline function of order $2m$, see (De Boor, 1978). Recall that spline functions are smooth piecewise functions. Since their introduction, splines have proved to be very popular in interpolation, smoothing and approximation, and in computational mathematics in general.

In this paper we present the steps of a new discrete-time algorithm which smooths signals from their uncertain discrete samples. The proposed algorithm does not require any knowledge of the statistics of the measurement uncertainties and is based on the minimization of a criterion equivalent to (1). The new discrete-time smoothing criterion is inspired by finite-difference schemes. In this algorithm the regularization parameter is obtained from the optimality condition of the Generalized Cross-Validation criterion as earlier introduced in (Craven and Wahba, 1979). We show that the smooth solution can be given as discrete samples or as a continuous-time spline function defined over the observation interval. Consequently, the regularized solution can be differentiated as many times as possible to estimate smooth higher derivatives of the measured signal.

2. Problem Statement and Solution of the Optimization Problem

Here, we consider the problem of smoothing noisy data with possibly estimating the higher derivatives $\hat{y}^{(\mu)}(t_i)$, $\mu = 0, 1, \dots, \nu - 1$ from discrete, potentially uncertain, samples $y_\ell = \bar{y}(t_\ell) + \epsilon(t_\ell)$, $\ell = i-n+1, \dots, i$, measured with an error $\epsilon(t_\ell)$ at n distinct instants, by minimizing the cost function

$$\mathscr{J} := \frac{1}{n}\sum_{\ell=i-n+1}^{i}\big[\hat{y}(t_\ell) - y(t_\ell)\big]^2 + \lambda\sum_{\ell=i-n+m}^{i-1}\Big[\hat{y}_\ell^{(m)}(\Delta t)^m\Big]^2, \quad i \in \mathbb{Z}_{\geq n},$$

where $\mathbb{Z}_{\geq n}$ is the set of positive integer numbers greater than or equal to n. For each moving window $[t_{i-n+1}, \dots, t_i]$ of length n, we minimize (2) with respect to $\hat{y}$. The first term in the criterion is the well-known least-squares criterion, and the second term represents an equivalent functional to the continuous integral

$$\int_{t_{i-n+1}}^{t_i} \hat{y}^{(m)}(t)\,\mathrm{d}t,$$

such that $\hat{y}^{(m)}(t)$ is the continuous m-th derivative of the function $\hat{y}(t)$. Here $\hat{y}_i^{(m)}$ denotes the finite-difference scheme of the m-th derivative of the continuous function $\hat{y}(t)$ at time $t = t_i$. In order to compute the m-th derivative of $\hat{y}(t)$ at time $t = t_i$ we will only use the samples

$\hat{y}_{i-m}, \hat{y}_{i-m-1}, \dots, \hat{y}_i$. Then the last cost function is written in the matrix form as

$$\mathscr{J} := \frac{1}{n}\|Y - \hat{Y}\|^2 + \lambda\|H\hat{Y}\|^2, \tag{2}$$

where

$$Y = \begin{bmatrix} y_{i-n+1} \\ y_{i-n+2} \\ \vdots \\ y_i \end{bmatrix}, \quad \hat{Y} = \begin{bmatrix} \hat{y}_{i-n+1} \\ \hat{y}_{i-n+2} \\ \vdots \\ \hat{y}_i \end{bmatrix},$$

and H is an $(n-m)\times(n)$ matrix consisting of general rows

$$(-1)^{m+j-1}C_m^{j-1}, \quad j = 1, \dots, m+1, \tag{3}$$

where C_n^k is the standard binomial coefficient. For $m = 2$, 3, and 4, the smoothness conditions are

$$\sum_{\ell=2}^{n-1}\left[\hat{y}_{\ell-1} - 2\hat{y}_\ell + \hat{y}_{\ell+1}\right]^2,$$

$$\sum_{\ell=3}^{n-1}\left[-\hat{y}_{\ell-2} + 3\hat{y}_{\ell-1} - 3\hat{y}_\ell + \hat{y}_{\ell+1}\right]^2,$$

$$\sum_{\ell=4}^{n-1}\left[\hat{y}_{\ell-3} - 4\hat{y}_{\ell-2} + 6\hat{y}_{\ell-1} - 4\hat{y}_\ell + \hat{y}_{\ell+1}\right]^2,$$

$$\sum_{\ell=5}^{n-1}\left[-\hat{y}_{\ell-4} + 5\hat{y}_{\ell-3} - 10\hat{y}_{\ell-2} + 10\hat{y}_{\ell-1} - 5\hat{y}_\ell + \hat{y}_{\ell+1}\right]^2,$$

respectively. Consequently, the corresponding matrices are

$$H_{(n-2)\times n} = \begin{bmatrix} 1 & -2 & 1 & 0 & \cdots & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & -2 & 1 \end{bmatrix},$$

$$H_{(n-3)\times n} = \begin{bmatrix} -1 & 3 & -3 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 3 & -3 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & -1 & 3 & -3 & 1 \end{bmatrix},$$

$$H_{(n-4)\times n} = \begin{bmatrix} 1 & -4 & 6 & -4 & 1 & 0 & \cdots & 0 \\ 0 & 1 & -4 & 6 & -4 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 1 & -4 & 6 & -4 & 1 \end{bmatrix}.$$

The derivative formulae (3) come from the approximation of the m-th derivative of $\hat{y}$ by the following finite-difference scheme:

$$\hat{y}_i^{(m)} = \frac{1}{(\Delta t)^m}\sum_{j=0}^{m+1}(-1)^{m+j}C_m^j\hat{y}_{i-m+j+1}. \tag{4}$$

This differentiation scheme is obtained by solving the set of the following Taylor expansions with respect to the derivatives $\hat{y}_i^{(1)}, \hat{y}_i^{(2)}, \dots, \hat{y}_i^{(m)}$:

$$\begin{aligned}
\hat{y}_{i-1} &= \hat{y}_i - \frac{\delta}{1!}\hat{y}_i^{(1)} + \frac{\delta^2}{2!}\hat{y}_i^{(2)} + \cdots + \frac{\delta^m}{m!}\hat{y}_i^{(m)},\\
\hat{y}_{i-2} &= \hat{y}_i - \frac{2\delta}{1!}\hat{y}_i^{(1)} + \frac{(2\delta)^2}{2!}\hat{y}_i^{(2)} + \cdots + \frac{(2\delta)^m}{m!}\hat{y}_i^{(m)},\\
&\vdots\\
\hat{y}_{i-m} &= \hat{y}_i - \frac{m\delta}{1!}\hat{y}_i^{(1)} + \frac{(m\delta)^2}{2!}\hat{y}_i^{(2)} + \ldots + \frac{(m\delta)^m}{m!}\hat{y}_i^{(m)},
\end{aligned}$$

where $\delta = t_i - t_{i-1}$ is the sampling period. We have selected this finite-difference scheme in order to force the matrix $H'H$ to be positive definite. The symbol $\|\cdot\|$ denotes the Euclidean norm, and λ is a smoothing parameter chosen in the interval $[0, \infty[$. We look for a solution of the last functional in the space of B-spline functions of order $k = 2m$. An interpretation of minimizing such a functional concerns the trade-off between the smoothing and the closeness to the data. If λ is set to zero, the minimization of (2) leads to a classical problem of least-squares approximation by a B-spline function of degree $2m - 1$.

We shall use splines because they often exhibit some optimal properties in interpolation and smoothing—in other words, they can often be characterized as solutions to variational problems. Roughly speaking, splines minimize some sort of "energy" functional. This variational characterization leads to a generalized notion of splines, namely, variational splines.

For each fixed measurement window, we seek the solution of (2) as

$$\hat{y}(t) := \sum_{j=i-n+1}^{i}\alpha_j b_{j,2m}(t), \quad t_{i-n+1} \le t \le t_i, \quad i \in \mathbb{Z}_{\ge n}, \tag{5}$$

where $\alpha \in \mathbb{R}^n$, and $b_{i,2m}(t)$ is the i-th B-spline basis function of order $2m$. For notational simplicity, $\hat{y}(t)$ and α are not indexed with respect to the moving window. We assume that the conditions of the optimization problem are the same for each moving window. Thus, the cost function (1) becomes

$$\mathscr{J} = \frac{1}{n}(Y - B\alpha)'(Y - B\alpha) + \lambda\alpha' B' RB\alpha \tag{6}$$

such that

$$R := H' H,$$

$$B_{i,j} := b_{j,2m}(t_\ell), \quad \ell = i-n+1,\dots,i, \quad i \in \mathbb{Z}_{\geq n}.$$

The optimum value of the control vector α is obtained via the optimality condition $\mathrm{d}\mathscr{J}/\mathrm{d}\alpha = 0$. Then we get

$$-\frac{2}{n}B'(Y - B\,\alpha) + 2\lambda B'\,R\,B\,\alpha = 0, \tag{7}$$

or

$$\begin{aligned}\alpha &= (n\lambda B'\,R\,B + B'\,B)^{-1}B'\,Y\\ &= (n\lambda RB + B)^{-1}Y.\end{aligned} \tag{8}$$

Consequently,

$$Y - B\,\alpha = n\lambda R\,B(n\lambda B'\,R\,B + B'\,B)^{-1}B'\,Y. \tag{9}$$

From (8), the continuous spline is fully determined. Hence the discrete samples of the regularized solution are computed from

$$\begin{aligned}\hat{Y} &= Y - n\lambda R\,B(n\lambda B'\,R\,B + B'\,B)^{-1}B'\,Y\\ &= \Big(I - n\lambda R\,(I + n\lambda R)^{-1}\Big)\,Y.\end{aligned} \tag{10}$$

As for the last equation, note that the discrete regularized samples are given as the output of an FIR filter where its coefficients are functions of the regularization parameter λ. The sensitivity of the solution to this parameter is quite important, so the next section will be devoted to the optimal calculation of the regularization parameter through the cross-validation criterion.

3. Computing the Regularization Parameter

In this section we shall present details of a computational method for estimating the optimal regularization parameter in terms of the criterion matrices. We have seen that the spline vector α depends upon the smoothing parameter λ. In (Craven and Wahba, 1979), two ways of estimating the smoothing parameter λ were given. The first method is called the ordinary cross-validation (OCV), which consists in finding the value of λ that minimizes the OCV-criterion

$$\mathscr{R}(\lambda) := \sum_{\ell=i-n+1}^{i}\big[\hat{y}(t_\ell) - y(t_\ell)\big]^2, \quad i = n, n+1, \dots, \tag{11}$$

where $\hat{y}(t)$ is a smooth polynomial of degree $2m-1$. Reinsch (1967) suggests, roughly speaking, that if the variance of the noise σ^2 is known, then λ should be chosen so that

$$\sum_{\ell=i-n+1}^{i}\big[\hat{y}(t_\ell) - y(t_\ell)\big]^2 = n\,\sigma^2. \tag{12}$$

Let $\mathscr{A}(\lambda)$ be the $n \times n$ matrix depending on $t_{i-n+1}, t_{i-n+2}, \dots, t_i$ and λ such that

$$\begin{bmatrix}\hat{y}(t_{i-n+1})\\ \vdots\\ \hat{y}(t_i)\end{bmatrix} = \mathscr{A}(\lambda)\begin{bmatrix}y(t_{i-n+1})\\ \vdots\\ y(t_i)\end{bmatrix}. \tag{13}$$

The main result of (Craven and Wahba, 1979) shows that a good estimate of the smoothing parameter λ (also called the generalized cross-validation parameter) is the minimizer of the GCV criterion

$$\mathscr{V}(\lambda) = \frac{\frac{1}{n}\|\,(I - \mathscr{A}(\lambda))\,Y\|^2}{\big[\frac{1}{n}\mathrm{trace}(I - \mathscr{A}(\lambda))\big]^2}. \tag{14}$$

This estimate has the advantage of being free from the knowledge of the statistical properties of noise. Further, if the minimizer of $\mathscr{V}(\lambda)$ is obtained, then the estimates of higher derivatives of the function $y(t)$ could be obtained by differentiating the smooth function $\hat{y}(t)$.

Now, we outline a computational method to determine the smoothing parameter which minimizes the cross-validation criterion $\mathscr{V}(\lambda)$, where the polynomial smoothing spline $\hat{y}(t)$ is supposed to be a B-spline of degree $2m-1$. Using the definition of $\mathscr{A}(\lambda)$, we write

$$Y - \hat{Y} = Y - \mathscr{A}(\lambda)Y = \big(I - \mathscr{A}(\lambda)\big)Y. \tag{15}$$

From (7), we obtain

$$Y - \hat{Y} = n\lambda\,R\,B\,\alpha. \tag{16}$$

Substituting (8) in (16), we get

$$\begin{aligned}Y - \hat{Y} &= n\lambda\,R\,B(n\lambda\,B'\,R\,B + B'\,B)^{-1}B'Y\\ &= n\lambda R(I + n\lambda R)^{-1}Y.\end{aligned} \tag{17}$$

By comparison with (15), we deduce that

$$\big(I - \mathscr{A}(\lambda)\big) = n\lambda R(I + n\lambda R)^{-1}. \tag{18}$$

The GCV-criterion becomes

$$\mathscr{V}(\lambda) = \frac{\frac{1}{n}\|n\lambda R(I + n\lambda R)^{-1}Y\|^2}{\Big[\frac{1}{n}\mathrm{trace}\,\Big(n\lambda R(I + n\lambda R)^{-1}\Big)\Big]^2}. \tag{19}$$

We propose the classical Newton method to compute the minimum of $\mathscr{V}(\lambda)$. This yields to the following iterations:

$$\lambda_{k+1} = \lambda_k - \frac{\dot{\mathscr{V}}(\lambda_k)}{\ddot{\mathscr{V}}(\lambda_k)}, \tag{20}$$

where $\dot{\mathscr{V}}$ et $\ddot{\mathscr{V}}$ are the first and second derivatives of $\mathscr{V}$ with respect to λ, respectively.

Let

$$\begin{aligned} p &= n\lambda, \\ v &= (pR+I)^{-1}Y, \\ W &= (pR+I)^{-1}. \end{aligned}$$

Then the criterion $\mathscr{V}$ becomes

$$\mathscr{V}(p) = \frac{\frac{1}{n}\|pRv\|^2}{\left[\frac{1}{n}\text{trace}(p\,R)W\right]^2}. \tag{21}$$

Let

$$\mathscr{N} = \frac{1}{n}\|pRv\|^2 = \frac{p^2}{n}v'R'Rv, \tag{22}$$

$$\mathscr{D} = \left[\frac{1}{n}\text{trace}(pRW)\right]^2. \tag{23}$$

Differentiating the last two equations with respect to λ, we obtain

$$\frac{\mathrm{d}\mathscr{N}}{\mathrm{d}\lambda} = 2pv'R'R\left[I + p^2 R\,W\,R - p\,R\right]v, \tag{24}$$

and

$$\begin{aligned} \frac{\mathrm{d}\mathscr{D}}{\mathrm{d}\lambda} = \frac{2}{n}\text{trace}(pRW)\Big[&\text{trace}(R\,W) \\ &+ \text{trace}(p\,R^2W\,(p\,R\,W - I))\Big]. \end{aligned} \tag{25}$$

Finally, the second derivatives of $\mathscr{N}$ and $\mathscr{D}$ are respectively

$$\begin{aligned} \frac{\mathrm{d}^2\mathscr{N}}{\mathrm{d}\lambda^2} &= 2nv'\,R'\,R(I+S)v \\ &\quad + 2pn\left\{2v'R'R(I+S)\frac{\mathrm{d}v}{\mathrm{d}p} + v'R'R\frac{\mathrm{d}S}{\mathrm{d}p}v\right\}, \end{aligned} \tag{26}$$

$$\begin{aligned} \frac{\mathrm{d}^2\mathscr{D}}{\mathrm{d}\lambda^2} &= 2\Big[\text{trace}\left(RW + p\,R^2\,W(p\,R\,W - I)\right)\Big]^2 \\ &\quad + 2\,\text{trace}(p\,R\,W)\left\{\text{trace}\Big(R\frac{\mathrm{d}W}{\mathrm{d}p}\Big)\right. \\ &\quad + \text{trace}\big(R^2W(p\,R\,W - I)\big) \\ &\quad + \text{trace}\Big(p\,R^2\frac{\mathrm{d}W}{\mathrm{d}p}(p\,R\,W - I)\Big) \\ &\quad \left. + \text{trace}\left(p\,R^2\,W\Big(R\,W + p\,R\frac{\mathrm{d}W}{\mathrm{d}p}\Big)\right)\right\}, \end{aligned} \tag{27}$$

such that

$$S = p^2 R\,W\,R - p\,R, \tag{28}$$

$$\frac{\mathrm{d}S}{\mathrm{d}p} = 2p\,R\left\{W + \frac{p}{2}\frac{\mathrm{d}W}{\mathrm{d}p}\right\}R - R, \tag{29}$$

$$\frac{\mathrm{d}W}{\mathrm{d}p} = p(RW)^2 - RW, \tag{30}$$

$$\frac{\mathrm{d}v}{\mathrm{d}p} = pRWRv - Rv. \tag{31}$$

Finally, the derivatives

$$\begin{aligned} \dot{\mathscr{V}} &= \frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{\mathscr{N}}{\mathscr{D}}\right), \\ \ddot{\mathscr{V}} &= \frac{\mathrm{d}^2}{\mathrm{d}\lambda^2}\left(\frac{\mathscr{N}}{\mathscr{D}}\right) \end{aligned}$$

can be easily computed in terms of the first and second derivatives of $\mathscr{N}$ and $\mathscr{D}$.

Remark 1. It is possible to recursively use the last algorithm if we take the values of the obtained spline as noisy data for another iteration. In this case the amount of noise in the data is reduced in each step by choosing a new smoothing parameter. The user could fix *a priori* a limited number of iterations according to the specified application and the time allowed to run the algorithm.

4. Connection with Adaptive Filtering

From (10), we have

$$\hat{Y} = \mathscr{A}(\lambda)Y, \tag{32}$$

where $\mathscr{A}(\lambda) = I - n\lambda R\,(I + n\lambda R)^{-1}$. If we write $\mathscr{A}(\lambda) = (a_{i,j}(\lambda))_{1\le i,j\le n}$, then

$$\begin{aligned} \hat{y}_i &= a_{n,1}(\lambda)y_{i-n+1} + a_{n,2}(\lambda)y_{i-n+2} \\ &\quad + \cdots + a_{n,n}(\lambda)y_i. \end{aligned} \tag{33}$$

Let $\hat{y}(z)$ and $y(z)$ be the z-transforms of the discrete signals $\hat{y}_i$ and y_i, respectively. Then by taking the z-transform of (33), we obtain

$$\begin{aligned} \frac{\hat{y}(z)}{y(z)} &= a_{n,1}(\lambda)z^{-n+1} + a_{n,2}(\lambda)z^{-n+2} \\ &\quad + \cdots + a_{n,n}(\lambda). \end{aligned} \tag{34}$$

The resulting system (34) takes the form of an adaptive FIR filter, where its coefficients $(a_{n,i})_{1\le i\le n}(\lambda)$ are updated by computing a new λ in each iteration $i \in \mathbb{Z}_{\ge n}$. The updating law in our case is based on the minimization of the generalized cross-validation criterion $\mathscr{V}(\lambda)$.

If we see attentively the formulation of the generalized cross-validation criterion given by (19), we realize that this criterion is simply a weighted least-squares (LS) performance index. The LS part is given by the numerator term $\|n\lambda R(I+n\lambda R)^{-1}Y\|^2$ which is exactly the error between the smoothed discrete samples and the noisy discrete data. The weighting parameter is given by the term $(1/n)/[(1/n)\text{trace}(I-\mathscr{A}(\lambda))]^2$. Consequently, the filter (34) can be seen as a weighted least-squares (WLS) adaptive FIR filter.

The smoothing strategy given in this paper has a relationship with the classical LMS (Least Mean Squares) adaptive filtering discussed in the signal processing literature. Although our method of updating the filter coefficients is not quite identical to the principle of LMS adaptive filtering, the philosophy of smoothing remains the same. To highlight this fact, let us recall the principle of LMS adaptive filtering. In such a filtering strategy, the time invariance of filter coefficients is removed. This is done by allowing the filter to change coefficients according to some prescribed optimization criterion. At each instant, the desired discrete samples $\hat{y}_i$ are compared with the instantaneous filter output $\tilde{y}_i$. On the basis of this measure, the adaptive filter will change its coefficients in an attempt to reduce the error. The coefficient update relation is a function of the error signal.

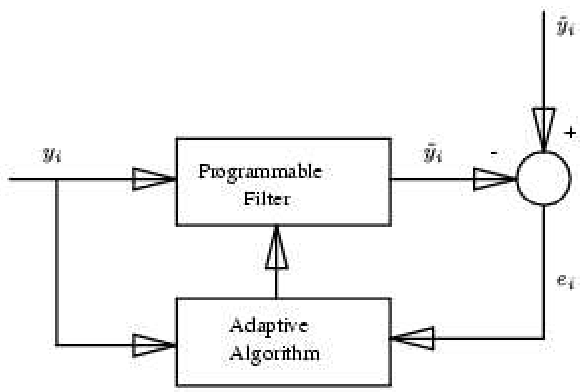

Fig. 1. Scheme of the LMS adaptive filter.

By comparison with the algorithm presented in this paper, the imposed signal $\hat{y}_i$ is not known *a priori*, but its formulation in terms of the noisy samples and the smoothing parameter λ is known. The main advantage of the GCV-based filter is that the minimum of the GCV performance index is computed independently of the knowledge of the statistical properties of noise. In addition, the information on the smoothing degree m is incorporated in the quadratic performance index (2), which makes the algorithm not only capable of filtering the discrete samples of the noisy signal but also capable of reliably reproducing the continuous higher derivatives of the signal considered.

5. Numerical Algorithm

Here, we summarize the regularization procedure in the following steps:

Step 1. Specify the desired spline of order $k = 2m$ and construct the optimal knot sequence which corresponds to the breakpoints $t_{i-n+1}, t_{i-n+2}, \dots, t_i$. See (De Boor, 1978) for more details on optimal knot computing.

Step 2. Construct B-spline basis functions that correspond to the optimal knots calculated in Step 1.

Step 3. Construct matrices H, B, R, T, and Q.

Step 4. Compute the optimal value of the smoothing parameter λ using (23)–(27).

Step 5. Compute the spline vector α.

Step 6. Compute the derivatives of the spline using the formulae

$$D^\ell\Big(\sum_i \alpha_i b_{i,k}\Big)(t) = \sum_i {\alpha_i}^{\ell+1} b_{i,k-\ell}(t),$$

where D^ℓ is the ℓ-th derivative with respect to time, and

$$ {\alpha_r}^{\ell+1} := \begin{cases} \alpha_r & \text{for } \ell = 0, \\ \dfrac{1}{k-\ell}\dfrac{{\alpha_r}^\ell - {\alpha_{r-1}}^\ell}{t_{r+k-\ell}-t_r} & \text{for } \ell > 0. \end{cases} \tag{35}$$

Step 7. In order to gradually reduce the amount of noise in the obtained smooth spline, the user has to repeat all the steps from the beginning by taking the values of the spline at $(t_{i-n+2}, \dots, t_{i+1})$ as noisy data for the next iteration.

6. Simulations

In the following simulations we suppose that we measure the noisy signal

$$y(t) = \cos(30t)\sin(t) + \epsilon(t) \tag{36}$$

for each $\delta = 0.01$ s. We assume that $\epsilon(t)$ is a norm-bounded noise of unknown variance. The exact first derivatives of the signal y are

$$\dot{y}(t) = -30\sin(30t)\sin(t) + \cos(30t)\cos(t), \tag{37}$$

$$\ddot{y}(t) = -901\cos(30t)\sin(t) - 60\sin(30t)\cos(t). \tag{38}$$

In Fig. 2 we show the noisy signal (36). In Fig. 3 we plot the exact signal (signal without noise) with the

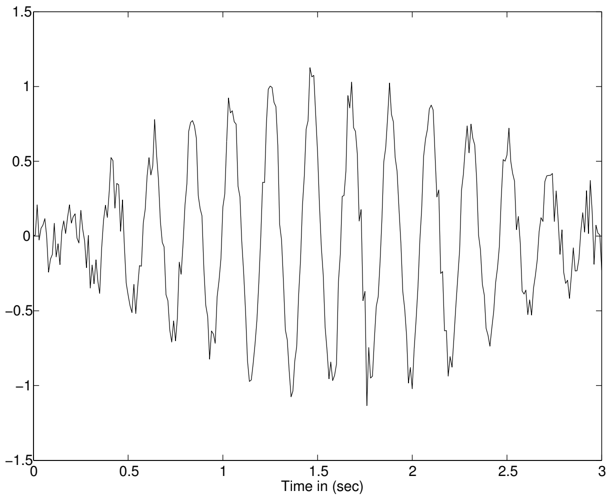

Fig. 2. Noisy signal.

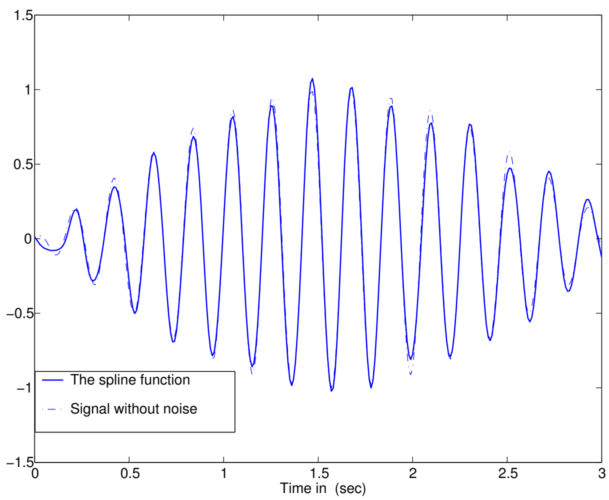

Fig. 3. Optimal spline vs. the exact signal.

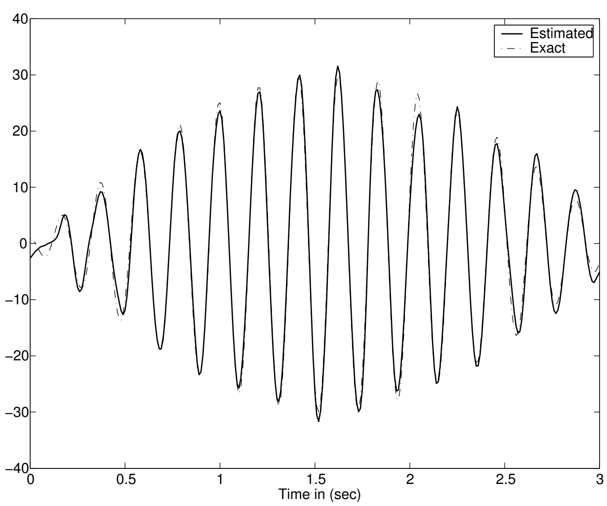

Fig. 4. First derivative of the optimal spline vs. the exact one.

continuous-time spline function, the solution to the minimization problem (2). In the whole simulation the moving window of observation is supposed to be constant of length 10. In Figs. 4 and 5 we depict the exact derivatives of the original signal with their estimated values given by the differentiation of the optimal continuous spline with respect to time. In Fig. 6, we compare the output of an LMS adaptive FIR filter of order 7 with the exact sample of the signal $y(t)$. We see clearly the superiority of the GCV-based filter in the first instants of the filtering process in comparison with the transient behaviour of the adaptive FIR filter presented in Fig. 6.

7. Conclusion

In this paper we have presented a new numerical procedure for reliable filtering and high-order signal differ-

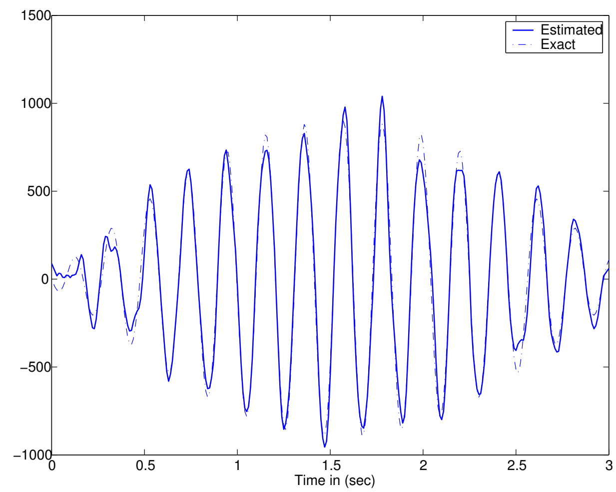

Fig. 5. Second derivative of the optimal spline and the exact second derivative of the signal.

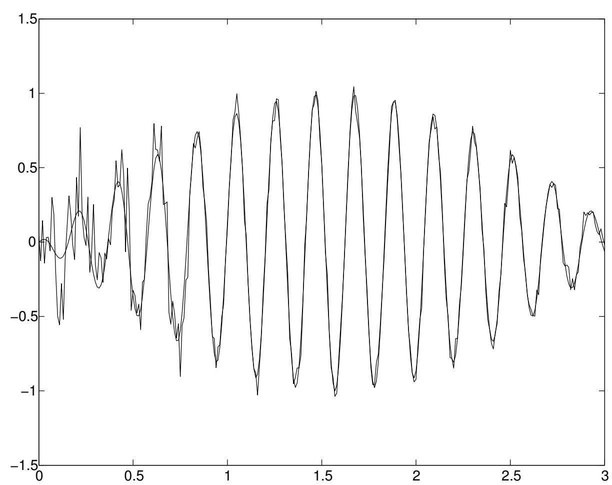

Fig. 6. Output of the adaptive FIR filter vs. the exact signal.

entiation. The design strategy consists in determining the continuous spline signal which minimizes the discrete functional being the sum of a least-squares criterion and a discrete smoothing term inspired by finite-difference schemes. The control of smoothing and the fidelity to the measurable data is ensured by the computation of one optimal regularization parameter that minimizes the generalized cross-validation criterion. The developed algorithm is able to estimate higher derivatives of a smooth signal only by differentiating its basis functions with respect to time. Satisfactory simulation results were obtained which prove the efficiency of the developed algorithm.

References

Anderson R.S. and Bloomfield P. (1974): *A time series approach to numerical differentiation*. — Technom., Vol. 16, No. 1, pp. 69–75.

Barmish B.R. and Leitmann G. (1982): *On ultimate boundness control of uncertain systems in the absence of matching assumptions*. — IEEE Trans. Automat. Contr., Vol. AC-27, No. 1, pp. 153–158.

Chen Y. H. (1990): *State estimation for non-linear uncertain systems: A design based on properties related to the uncertainty bound*. — Int. J. Contr., Vol. 52, No. 5, pp. 1131–1146.

Chen Y. H. and Leitmann G. (1987): *Robustness of uncertain systems in the absence of matching assumptions*. — Int. J. Contr., Vol. 45, No. 5, pp. 1527–1542.

Ciccarella G., Mora M.D. and Germani A. (1993): *A Luenberger-like observer for nonlinear systems*. — Int. J. Contr., Vol. 57, No. 3, pp. 537–556.

Craven P. and Wahba G. (1979): *Smoothing noisy data with spline functions: Estimation the correct degree of smoothing by the method of generalized cross-validation*. — Numer. Math., Vol. 31, No.4, pp. 377–403.

Dawson D.M., Qu Z. and Caroll J.C. (1992): *On the state observation and output feedback problems for nonlinear uncertain dynamic systems*. — Syst. Contr. Lett., Vol. 18, No.3, pp. 217–222.

De Boor C., (1978): *A Practical Guide to Splines*. — New York: Springer.

Diop S., Grizzle J.W., Morral P.E. and Stefanoupoulou A.G. (1993): *Interpolation and numerical differentiation for observer design*. — Proc. Amer. Contr. Conf., Evanston, IL, pp. 1329–1333.

Eubank R.L. (1988): *Spline Smoothing and Nonparametric Regression*. — New York: Marcel Dekker.

Gasser T., Müller H.G. and Mammitzsch V. (1985): *Kernels for nonparametric curve estimation*. — J. Roy. Statist. Soc., Vol. B47, pp. 238–252.

Gauthier J.P., Hammouri H. and Othman S. (1992): *A simple observer for nonlinear systems: Application to bioreactors*. — IEEE Trans. Automat. Contr., Vol. 37, No. 6, pp. 875–880.

Georgiev A.A. (1984): *Kernel estimates of functions and their derivatives with applications*. — Statist. Prob. Lett., Vol. 2, pp. 45–50.

Härdle W. (1984): *Robust regression function estimation*. — Multivar. Anal., Vol. 14, pp. 169–180.

Härdle W. (1985): *On robust kernel estimation of derivatives of regression functions*. — Scand. J. Statist., Vol. 12, pp. 233–240.

Ibrir S. (1999): *Numerical algorithm for filtering and state observation*. — Int. J. Appl. Math. Comp. Sci., Vol. 9, No.4, pp. 855–869.

Ibrir S. (2000): *Méthodes numriques pour la commande et l'observation des systèmes non linéaires*. — Ph.D. thesis, Laboratoire des Signaux et Systèmes, Univ. Paris-Sud.

Ibrir S. (2001): *New differentiators for control and observation applications*. — Proc. Amer. Contr. Conf., Arlington, pp. 2522–2527.

Ibrir S. (2003): *Algebraic riccati equation based differentiation trackers*. — AIAA J. Guid. Contr. Dynam., Vol. 26, No. 3, pp. 502–505.

Kalman R.E. (1960): *A new approach to linear filtering and prediction problems*. — Trans. ASME J. Basic Eng., Vol. 82, No. D, pp. 35–45.

Leitmann G. (1981): *On the efficacy of nonlinear control in uncertain linear systems*. — J. Dynam. Syst. Meas. Contr., Vol. 102, No.2, pp. 95–102.

Luenberger D.G. (1971): *An introduction to observers*. — IEEE Trans. Automat. Contr., Vol. 16, No.6, pp. 596–602.

Misawa E.A. and Hedrick J.K. (1989): *Nonlinear observers. A state of the art survey*. — J. Dyn. Syst. Meas. Contr., Vol.111, No.3, pp. 344–351.

Müller H.G. (1984): *Smooth optimum kernel estimators of densities, regression curves and modes*. — Ann. Statist., Vol. 12, pp. 766–774.

Rajamani R. (1998): *Observers for Lipschitz nonlinear systems*. — IEEE Trans. Automat. Contr., Vol. 43, No. 3, pp. 397–400.

Reinsch C.H. (1967): *Smoothing by spline functions*. — Numer. Math., Vol. 10, pp. 177–183.

Reinsch C.H. (1971): *Smoothing by spline functions ii*. — Numer. Math., Vol. 16, No.5, pp. 451–454.

Slotine J.J.E., Hedrick J.K. and Misawa E.A. (1987): *On sliding observers for nonlinear systems*. — J. Dynam. Syst. Meas. Contr., Vol. 109, No.3, pp. 245–252.

Tornambè A. (1992): *High-gain observers for nonlinear systems*. — Int. J. Syst. Sci., Vol. 23, No.9, pp. 1475–1489.

Xia X.-H. and Gao W.-B. (1989): *Nonlinear observer design by observer error linearization*. — SIAM J. Contr. Optim., Vol. 27, No. 1, pp. 199–216.

Received: 26 January 2004
Revised: 28 May 2004