A revised look at numerical differentiation with an application to nonlinear feedback control
Mamadou Mboup, Cédric Join, Michel Fliess

To cite this version:
Mamadou Mboup, Cédric Join, Michel Fliess. A revised look at numerical differentiation with an application to nonlinear feedback control. The 15th Mediterranean Conference on Control and Automation - MED’2007, Jun 2007, Athènes, Greece. inria-00142588

HAL Id: inria-00142588
https://inria.hal.science/inria-00142588
Submitted on 19 Apr 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
A revised look at numerical differentiation with an application to nonlinear feedback control

Mamadou MBOUP†‡, Cédric JOIN*†‡, Michel FLIESS *§

* Projet ALIEN, INRIA-Futurs
† UFR Mathématiques et Informatique, Université René Descartes (Paris 5), 45 rue des Saints-Pères, 75270 Paris cedex 06, France
mboup@math-info.univ-paris5.fr
‡ CRAN (CNRS-UMR 7039), Université Henri Poincaré (Nancy I), BP 239, 54506 Vandœuvre-lès-Nancy, France
cedric.join@cran.uhp-nancy.fr
§ Équipe MAX, LIX (CNRS-UMR 7161), École polytechnique, 91128 Palaiseau, France
Michel.Fliess@polytechnique.edu

Abstract—We are presenting new and efficient methods for numerical differentiation, i.e., for estimating derivatives of a noisy time signal. They are illustrated, via convincing numerical simulations, by the analysis of an academic signal and by the feedback control of a nonlinear system.

I. INTRODUCTION

A. Numerical differentiation

Numerical differentiation, i.e., the derivatives estimation of noisy time signals, is an important but difficult ill-posed theoretical problem. It has attracted a lot of attention in many fields of engineering and applied mathematics (see, e.g., in the recent control literature [2], [3], [4], [14], [15], [16], [17], [18], [19], [23], and the references therein). Our purpose here is to improve a new approach which started in [8], [13], and in [7], [9], [10], for solving various questions in control and in signal and image processing. Let us briefly describe our differentiators which are obtained via iterated time integrals of the noisy signal.

B. Short summary of our approach

Start with the first degree polynomial time function

\[ p_1(t) = a_0 + a_1 t, \quad t \geq 0, \quad a_0, a_1 \in \mathbb{R}. \]

Rewrite thanks to classic operational calculus (cf. [25]) \( p_1 \) as \( P_1 = \frac{a_0}{s} + \frac{a_1}{s^2} \).

Multiply both sides by \( s^2 \):

\[ s^2 P_1 = a_0 s + a_1 \]

Take the derivative of both sides with respect to \( s \), which corresponds in the time domain to the multiplication by \(-t\) (cf. [25]):

\[ s^2 \frac{dP_1}{ds} + 2sP_1 = a_0 \]

The coefficients \( a_0, a_1 \) are obtained via the triangular system of equations (1)-(2). We get rid of the time derivatives, i.e., of \( sP_1, s^2 P_1 \), and \( s^2 \frac{dP_1}{ds} \), by multiplying both sides of equations (1)-(2) by \( s^{-n} \), \( n \geq 2 \). The corresponding iterated time integrals are low pass filters which attenuate the corrupting noises, which are viewed as highly fluctuating phenomena (see [6] for more details). A quite short time window is sufficient for obtaining accurate values of \( a_0, a_1 \).

The extension to polynomial functions of higher degree is straightforward. For derivatives estimates up to some finite order of a given smooth function \( f : [0, +\infty) \to \mathbb{R} \), take a suitable truncated Taylor expansion around a given time instant \( t_0 \), and apply the previous computations. Utilizing sliding time windows permit to estimate derivatives of various orders at any sampled time instant.

C. Difficulties and improvements

The above method becomes more and more ill-conditioned for higher order truncation of the Taylor expansion. This is a major impediment for obtaining good estimates for higher order derivatives in a noisy setting. Using elimination techniques and Jacobi orthogonal polynomials (see, e.g., [1], [24]), we propose here individual and independent derivatives estimators for each given order. A judicious choice of the point at which the derivatives are estimated in each sliding time window permits to take advantage of the extra-modeling capability afforded by a higher order truncation.

D. Organization of our paper

Section II discusses the mathematical foundations of our differentiators. The first illustration in Section III applies the above techniques for estimating the derivative of a noisy academic signal. The second illustration in Section IV, which is borrowed from [5], deals with the nonlinear feedback control of a DC motor joined to an inverted pendulum through a torsional spring. We provide in both cases convincing computer simulations1.

1Interested readers may obtain the corresponding computer programs from one of the authors (cedric.join@cran.uhp-nancy.fr). Reference [11], from which the second example is taken, contains many more applications to various topics in nonlinear control. Most useful discussions and comparisons may be found in [21] where an interesting concrete case-study is analyzed.
II. DERIVATIVES ESTIMATION

Let \( y(t) = x(t) + n(t) \) be a noisy observation on a finite time interval of a real-valued signal \( x(t) \), the derivatives of which we want to estimate. Assume that \( x(t) \) is analytic on this time interval. Consider without any loss of generality the convergent Taylor expansion \( x(t) = \sum_{i \geq 0} c_i t^i \) at \( t = 0 \). The truncated Taylor expansion \( x_N(t) = \sum_{i=0}^{N} c_i t^i \) satisfies the differential equation \( \frac{d^{N+1}}{dt^{N+1}} x_N(t) = 0 \). It reads in the operational domain as

\[
s^{N+1} \hat{x}_N(s) = s^N x_N(0) + s^{N-1} \hat{x}_N(0) + \ldots + x_N'(0) \tag{3}
\]

where \( \hat{x}_N(s) \) is the operational analog of \( x_N(t) \). To ease the notation, we subsequently ignore the argument \( s \).

A. Simultaneous estimation

Replace \( x_N(t) \) by the noisy observed signal \( y(t) \). Then the estimates \( \hat{x}_N^{(i)}(0) \) of the derivatives at the origin \( x_N^{(i)}(0) \) are directly obtained from the linear triangular system of equations (see [8], [13])

\[
s^{-i} \frac{d^m}{ds^m} \left\{ \hat{x}_N(0)s^N + \ldots + \hat{x}_N^{(N-1)}(0)s + \hat{x}_N^{(N)}(0) \right\} = \hat{x}_N^{(N-i)}(0), \quad m = 0, \ldots, N, \quad \nu \geq 2 \text{ ensures strict properness.}
\]

To obtain the numerical estimates, it suffices to express (4) back in the time domain, using the classical rules of operational calculus ([25]). Denote by \( T \) the estimation time. We end up with the following closed form expression

\[
\mathcal{P}_\nu(T) \begin{bmatrix} \hat{x}_N(0) \\ \hat{x}_N'(0) \\ \vdots \\ \hat{x}_N^{(N-1)}(0) \end{bmatrix} = \int_0^T Q_\nu(\tau) y(\tau) d\tau \tag{5}
\]

where the nonzero entries of the triangular matrix \( \mathcal{P}_\nu(T) \) are given, for \( i = 0, \ldots, N, \ j = 0, \ldots, N - i \), by

\[
\{ \mathcal{P}_\nu(T) \}_{ij} = \frac{(N-j)!}{(N-i-j)! (\nu - N + i + j - 1)!} T^{\nu-N+i+j-1}
\]

and

\[
\{ Q_\nu(\tau) \}_i = \sum_{\ell=0}^{i} q_{i,\ell} (T - \tau)^{\nu-N-2-\ell} \tau^{i-\ell}
\]

with

\[
q_{i,\ell} = \binom{i}{\ell} \frac{(-1)^{i-\ell}}{(N+1-\ell)!} \frac{(N+1)!}{(\nu - N - 2 - \ell)!}
\]

Finally, for each estimation time interval of length \( T \), \( I^T_n = [t, t + T] \), we obtain the derivatives estimates at time \( t \) by replacing \( y(t) \) in (5) by \( y(t+\tau) \). These estimates are however not causal. To obtain causal estimates, i.e., the estimates at time \( \tau \), based on the signal observation in \( I^T_n = [t - T, t] \), it suffices to replace \( y(t+\tau) \) by \(-y(t - \tau), \ \tau \in [0, T] \).

B. Individual estimation

The matrix \( \mathcal{P}_\nu(T) \) in (5) is in general ill-conditioned, and yields therefore poor estimates especially in a noisy setting. A solution to this problem is to obtain an independent estimator for each order of derivation \(^2\). Reconsider (3). Examine for \( 0 \leq n \leq N \) the \( n^{th} \) order derivative. Annihilate the remaining coefficients \( x_N^{(j)}(0), \ j \neq n \) by multiplying by linear differential operators of the form

\[
\Pi^N_n = \frac{d^{n+\kappa}}{ds^{n+\kappa}} \cdot \frac{1}{s} \frac{d^{N-n}}{ds^{N-n}}, \quad \kappa \geq 0
\]

It yields the following estimator for \( x^{(n)}(0) \)

\[
\hat{x}_N^{(n)}(0) = \frac{(-1)^{n+\kappa}}{(n+\kappa)! (N-n)!} \frac{1}{s^{\nu+n+\kappa+1}} \Pi^N_n \left( s^{N+1} \hat{x} \right) \tag{6}
\]

which is strictly proper whenever \( \nu \) is of the form \( \nu = N + 1 + \mu, \ \mu \geq 0 \). We obtain a family of strictly proper estimators which is parametrized by \( \kappa, \mu \) and \( N \). Write \( \hat{x}_N^{(n)}(\kappa, \mu; N) \) the corresponding estimator. If \( N = n \) the differential operator \( \Pi^N_n \) reduces to \( \Pi^N_n = s^{\nu+n+1}/s^{\nu+n+1} \) we will then use the simplified notation \( \hat{x}_N^{(n)}(\kappa, \mu) \). The following result is straightforward:

**Lemma 1:** For any \( N \geq n \) and \( \mu \geq 0 \), \( \hat{x}_N^{(n)}(\kappa, \mu; N) \) in (6) belongs to the set

\[
\mathcal{F} = \text{span}_q \{ \hat{x}_0^{(n)}(\kappa_i, \mu_i), \ell = 0, \ldots, \min (n+\kappa, N-n) \} \tag{7}
\]

where \( \kappa_q = \kappa + N - n - \ell \) and \( \mu_q = \mu + \ell \).

**Proof:** Set \( q = N - n + p \). The proof follows by direct inspection, upon writing \( \Pi^N_n(s^{N+1} \hat{x}) \) in the form

\[
\Pi^N_n(s^{N+1} \hat{x}) = \sum_{i=0}^{q} \left( \frac{q}{i} \right) \frac{d^p}{ds^p} \left\{ s^{q-i} \hat{x} \right\}
\]

\[
= \sum_{i=0}^{q} \sum_{j=0}^{\min (p, q-i)} a_{i,j} s^{j-i} \Pi^{q-j} \left( s^{N+1} \hat{x} \right)
\]

where

\[
a_{i,j} = \binom{q}{i} \frac{p}{j-i} \frac{(-1)^{q-i-j}}{(q+1-i)(q-j)!}
\]

This lemma shows that an \( n^{th} \) order truncated Taylor expansion is appropriate for estimating the \( n^{th} \) order derivative.

C. Least squares interpretation\(^3\)

A common way for estimating the derivatives of a signal is to resort to a least squares polynomial fitting on an interval and then take the derivatives of the resulting polynomial function. The estimators derived here rely however on a different approach: the derivatives are estimated pointwise. This departure is furthermore apparent with the developments of the preceding subsection. Nonetheless, a least squares interpretation may be attached to our approach, as shown below.

Start with the estimation of the first order derivative.

\(^2\)The system (5) being triangular, a closed-form expression for the estimator of \( x^{(i)}(t) \) may be derived from it. The corresponding solutions would however exhibit the same sensitivity to noise perturbations.

\(^3\)The authors would like to thank A. Sedoglavic for bringing this question to their attention.
a) \( N = 1 \): With \( \nu = \mu + 2 \), equation (6) becomes
\[
\frac{1}{s^{\mu+1}+\kappa+4} \tilde{x}_0(\kappa, \mu) = \left( \frac{1}{(\kappa+2)!} \right) \left( \frac{2^{\kappa+1}}{s^{\mu+1}} + (\kappa+1) \frac{2^\kappa}{s^{\mu+2}} \right)
\]
where \( P_{1,:}^{\kappa,:}(t) = -\tau_0 = -\frac{\kappa+2}{\mu+\kappa+4} \). We therefore deduce that \( \tilde{x}_0(\kappa, \mu; 2) \) corresponds to an estimate of the first order derivative \( \dot{x}(t) \) at \( t = 0 \). The estimation of \( \dot{x}(t) \) from a second order truncation of the Taylor expansion is therefore delay free.

We now show how this interpretation can be exploited to obtain better estimates. According to Lemma 1 and using (14), it is easy to verify that the relation
\[
\tilde{x}_0(\kappa, \mu; 2) = \lambda_0 \tilde{x}_0(\kappa, \mu + 1) + \lambda_1 \tilde{x}_0(\kappa + 1, \mu)
\]
holds for \( \lambda_0 = \kappa + 3 \) and \( \lambda_1 = -\kappa + 2 \). Let us now extend the set \( \mathcal{F} \) in (7), by allowing the coefficients of the linear combinations therein to be real, rather than rational. In particular, given a point \( \tau_1 \in [0, 1] \), one may always choose the \( \lambda_i \)'s such that (16) becomes
\[
\tilde{x}_{\tau_1}(\kappa, \mu; 2) = \left( \frac{P_{1,:}^{\kappa,:}(\tau_1), \dot{y}(T\tau_1)}{\|P_{1,:}^{\kappa,:}\|^2} \right) + \left( \frac{P_{1,:}^{\kappa,:}(\tau_1), \dot{y}(T\tau_1)}{\|P_{1,:}^{\kappa,:}\|^2} \right)
\]
In this case, \( \tilde{x}_{\tau_1}(\kappa, \mu; 2) \) will represent the estimate of \( \dot{x}(T\tau_1) \), obtained from the truncated Taylor expansion \( x_2(t) = x(t) + \dot{x}(t)(\tau - T\tau) + \frac{\ddot{x}(t)}{2} (\tau - T\tau)^2 \). A direct verification will show that the values for \( \lambda_0 \) and \( \lambda_1 \), associated to \( \tau_1 \) are given by \( \lambda_0 = \kappa + 3 - (\mu + \kappa + 5) \tau_1 \) and \( \lambda_1 = 1 - \lambda_0 \).

Let us now turn to the question pertaining to the selection of a value for \( \tau_1 \). Expansion (12) shows that a good choice for \( \tau_1 \) is given by the smallest (resp. largest) root of the polynomial \( P_{1,:}^{\kappa,:} \) when the estimation interval is \( I_{\tau_1}^T \) (resp. \( I_{\tau_1}^- \)). Indeed, choosing \( \tau_1 \) as a zero of \( P_{1,:}^{\kappa,:} \) annihilates the contribution of the orthogonal projection of the signal on \( P_{1,:}^{\kappa,:} \) in the estimation error. On the other hand, recall that \( \tau_1 \) (resp. \( 1 - \tau_1 \) for \( I_{\tau_1}^- \)) represents a delay in the estimation. Choosing the smallest (resp. largest) root thus translates to a lower delay. Note also that the delay \( \tau_1 \) is smaller than \( \tau_0 \), since the zeros of \( P_{1,:}^{\kappa,:} \) and \( P_{1,:}^{\kappa,:} \) interlace.

The above analysis may be easily generalized to the \( n \)-th order derivative estimation, \( n \geq 2 \).

III. FIRST ILLUSTRATION: DERIVATIVE OF A NOISY SIGNAL

Let \( y(t) = x(t) + n(t), 0 \leq t \leq 5 \), be a noisy measurement of the signal
\[
x(t) = \tanh(t - 1) + e^{-t/1.2} \sin(6t + \pi)
\]
The noise level, measured by the signal to noise ratio in dB, i.e., \( SNR = 10 \log_{10} \left( \sum_{i=1}^{n} \frac{\hat{y}(t_i)^2}{\hat{n}(t_i)^2} \right) \), corresponds to \( SNR = 25 \) dB (see Figure 1). In all the subsequent numerical simulations, the integrals are computed via the classical trapezoidal rule.

Begin with the first order derivative and \( N = 1 \). The estimates \( \hat{x}_0(\kappa, \mu) \), obtained from (10) with \( \kappa = \mu = 2 \), are displayed in Figure 2 (solid line). It corresponds to the estimation results in the successive intervals \( I_{\tau_1}^T \), with \( T = 60T_s \) and for \( t = iT_s, i = 0, \ldots, \left[ \frac{5}{T_s} \right] \). The exact
derivative of the noise-free signal is also displayed (dashed line) in order to gauge the estimation accuracy. For this, we have introduced a shift corresponding to the delay \( \tau_0 \) in (13). Observe how this predicted value of \( \tau_0 \) fits with the experiment. Of course with \( N = 1 \), the truncated Taylor series model is linear, resulting in poor estimates on the intervals where the signal’s dynamic is strong. For high signal-to-noise ratio, the estimates may be improved by reducing the estimation time \( T \). Alternatively, one may consider a richer signal model, e.g., with \( N = 2 \). This is the case of the next experiment. We now consider the sliding windows \( I_{T,K} \), with \( T = 110T_s \). The estimates \( \hat{x}_0(\kappa, \mu; 2) \), based on (14), (15), are plotted (solid line) in Figure 3 below, for \( \kappa = \mu = 0 \).

There is no estimation delay, as expected. However, the performance significantly degrades as compared to the preceding results although the signal model is more precise. If we now relax this delay-free constraint, it becomes possible to take advantage of the more flexible second order model for the signal.

This is illustrated in the following simulation (see Figure 4), where we keep the same settings for \( T, \kappa \) and \( \mu \). The solid line curve in Figure 4 represents the estimates obtained from (16)-(17) where \( \tau_1 \) is the largest root of \( P_{3}^{\kappa,\mu} \). Using the same idea with \( N = 3 \), we obtain the estimates \( \hat{x}_0(\kappa, \mu; 3) \), plotted with dashed line. The associated estimator reads as

\[
\hat{x}_0(\kappa, \mu; 3) = \sum_{i=0}^{2} \lambda_i \hat{x}_0(\kappa + 2 - i, \mu + i)
\]

where the real coefficients \( \lambda_i, i = 0, 1, 2 \) are chosen so that \( \hat{x}_0(\kappa, \mu; 3) \) corresponds to \( \hat{x}(\tau_2 T) \), with \( \tau_2 \) being a root of \( P_{3}^{\kappa,\mu} \). If \( \kappa = \mu \), which is the case here, then \( \tau_2 = 0.5 \) is a common root of all Jacobi polynomials of odd degree. The dashed line curve was obtained with this choice of \( \tau_2 \).

IV. SECOND ILLUSTRATION: NONLINEAR FEEDBACK CONTROL

A. System description

Consider with [5] the mechanical system, which consists of a DC-motor joined to an inverted pendulum through a torsional spring:

\[
\begin{align*}
J_m \ddot{\theta}_m &= \kappa(\theta_l - \theta_m) - B \dot{\theta}_m + K \tau \\
J_l \dot{\theta}_l &= -\kappa(\theta_l - \theta_m) - mgh \sin(\theta_l) \\
y &= \theta_l
\end{align*}
\]
θ_m and θ_l represent respectively the angular deviation of the motor shaft and the angular position of the inverted pendulum. J_m, J_l, h, m, κ, B, K_r and g are physical parameters which are assumed to be constant and known. This system, which is linearizable by static state feedback, is flat (cf. [12], [22]); y = θ_l is a flat output, which is measured.

B. Control design

Asymptotic tracking of a given smooth reference trajectory \( y^*(t) = \theta_l^*(t) \) is achieved by the feedback controller

\[
u = \frac{1}{\kappa} \left( \frac{J_m}{\kappa} (J_l v + \kappa \dot{y}_e + mgh (\dot{y}_e \cos y_e - (\dot{y}_e)^2 \sin y_e)) + J_l \dot{y}_e + mgh \sin y_e \frac{p}{\kappa} (J_l y_e^{(2)} + \kappa \dot{y}_e + mgh \dot{y}_e \cos y_e) \right)
\]

where

\[
v = [y^*(t)]^{(4)} - \gamma_4 (y_e^{(3)} - [y^*(t)]^{(3)}) - \gamma_2 (y_e - \dot{y}^*(t)) - \gamma_1 (y_e - y^*(t))
\]

The subscript “e” denotes the estimated value. The design parameters γ_1, ..., γ_4 are chosen so that the resulting characteristic polynomial is Hurwitz. The state θ_m is estimated via

\[\hat{\theta}_m = \frac{1}{\kappa} (J_l \dot{y}_e + mgh \sin y_e) + y_e\]

C. Numerical simulations

The physical parameters have the same numerical values as in [5]: J_m = 3.7 \times 10^{-3} \text{ kgm}^2, J_l = 9.3 \times 10^{-3} \text{ kgm}^2, h = 1.5 \times 10^{-1} \text{ m}, m = 0.21 \text{ kg}, B = 4.6 \times 10^{-2} \text{ m}, K_r = 8 \times 10^{-2} \text{ NmV}^{-1}. The numerical simulations (see also [11]), which are much better than in [13], where less efficient differentiators were employed, are presented in Figures 5. Robustness has been tested with an additive white Gaussian noise \( \mathcal{N}(0; 0.01) \) on the output y (this noise level is quite relevant in such applications). Note that the off-line estimations of \( \ddot{y} \) and \( \hat{\theta}_m \), where a “small” delay is allowed, are more accurate than the on-line estimation of \( \ddot{y} \).

V. CONCLUSION

The basic elements of our differentiators are essentially integrators, the advantage of which is twofold: easy implementation and good robustness with respect to noise corruption. Our estimators may be given a least squares interpretation (see [20] for more details), although our approach is based on quite different mathematical ingredients, which are mainly of algebraic flavor. This interpretation is a key point leading to improved numerics.

REFERENCES

Fig. 5. Simulation results for DC-motor