Numerical differentiation in noisy environment

Mamadou Mboup

Université de Reims Champagne Ardenne (URCA)
CReSTIC – UFR des Sciences Exactes et Naturelles
BP 1039, Moulin de la Housse, 51687 Reims cedex 2, France
& EPI Non-A, INRIA

Problem statement

Given $y(t) = x(t) + \varpi(t)$, $\varpi(t)$: additive noise corruption,
we want to estimate

$$x^{(n)}(t) \triangleq \frac{d^n}{dt^n} x(t) \text{ for some } n \in \mathbb{N}$$

The problem is

- important in many fields of engineering and applied mathematics: control, signal and image processing, numerical analysis, ...
- difficult and ill-posed,
- longstanding with many different approaches and a huge litterature.

UNIVERSITÉ DE REIMS CHAMPAGNE-ARDENNE

Most common approaches

- Finite difference: Not acceptable except for high SNR
- Frequency domain: transfer function $= \omega^n$
  $+$ something to cope with uncertainties
- Observer design: *e.g.* Sliding mode differentiator *à la* Levant
- Differentiate a least squares approximation/interpolation
  differentiation and approximation do not commute
- Variational approach (inverse problem): find a differentiation operator compatible with the data
  $+$ something to cope with uncertainties

something to cope with uncertainties $=$ minimize some energy, some norm
$\rightsquigarrow$ Regularisation

A revised look

Framework [Fliess & Ramirez, 2002]

- Differential algebra: already proved useful in control
- Operational calculus: for handling differential algebraic equations.

Principle: Pointwise estimation by differential elimination

- $\frac{d^n x(t)}{dt^n} \mid_{t=\tau}$ from an integral operator of $y(t)$, $t \in \mathcal{I}_\tau^T = [\tau - T,\ \tau]$
- Slide the window $\mathcal{I}_\tau^T$ to obtain the estimate $\frac{d^n x(t)}{dt^n}$ for all t.

Features: The method inherently induces

- Least-squares estimation of the unknown $x^{(n)}(t)$
- Regularisation step via a delay

Implementation and performance:

- Simple on-line implementation: discrete-time taped delay line (FIR).
- Good robustness to noise corruption

Algebraic approach: algorithm description

Let $N \geqslant n$, consider $x_N(t) = \sum_{i=0}^{N} x^{(i)}(0)\frac{t^i}{i!}$ and set $x_N(t) \circ\!\!-\!\!\bullet\ \hat{x}_N(s)$.

$$s^{N+1}\hat{x}_N(s) = s^N x(0) + s^{N-1}\dot{x}(0) + \cdots + s^{N-n}x^{(n)}(0) + \cdots + x^{(N)}(0)$$

1. Annihilators: Find $\Pi = \sum_{\text{finite}} \varrho_\ell(s)\frac{d^\ell}{ds^\ell}$, $\quad \varrho_\ell(s) \in \mathbb{C}(s)$, such that

$$\Pi\hat{x}_N = \varrho(s)x^{(n)}(0), \quad \varrho(s) \in \mathbb{C}(s)$$

2. Integral operators: Left multiply by $\sigma(s) \in \mathbb{C}(s)$ such that

$$\sigma(s)\Pi \text{ and } \sigma(s)\varrho(s) \text{ strictly proper.}$$

3. Operational estimate: Replace x_N by the actual noisy observation y

$$\rightsquigarrow \varrho(s)\tilde{x}^{(n)}(0;N) = \sum_{\text{finite}} \varrho_\ell(s)\frac{d^\ell}{ds^\ell}\hat{y} \qquad (*)$$

4. Back in time domain: If $\varrho(s) \bullet\!\!-\!\!\circ\ \rho(t)$ and $\varrho_\ell(s) \bullet\!\!-\!\!\circ\ h_\ell(t)$, then

$$(*) \bullet\!\!-\!\!\circ\ \rho(t)\tilde{x}^{(n)}(0;N) = \int_0^t \sum_{\text{finite}} \left\{h_\ell(t-\lambda)(-1)^\ell\lambda^\ell\right\} y(\lambda)d\lambda$$

where t = estimation time.

Algebraic approach: algorithm description

Let $N \geqslant n$, consider $x_N(t) = \sum_{i=0}^{N} x^{(i)}(0)\frac{t^i}{i!}$ and set $x_N(t) \circ\!\!-\!\!\bullet\; \hat{x}_N(s)$.

$$s^{N+1}\hat{x}_N(s) = s^N x(0) + s^{N-1}\dot{x}(0) + \cdots + s^{N-n}x^{(n)}(0) + \cdots + x^{(N)}(0)$$

1. Annihilators: Find $\Pi = \sum_{\text{finite}} \varrho_\ell(s)\frac{d^\ell}{ds^\ell}, \quad \varrho_\ell(s) \in \mathbb{C}(s)$, such that

$$\Pi\hat{x}_N = \varrho(s)x^{(n)}(0), \quad \varrho(s) \in \mathbb{C}(s)$$

2. Integral operators: Left multiply by $\sigma(s) \in \mathbb{C}(s)$ such that

   $\sigma(s)\Pi$ and $\sigma(s)\varrho(s)$ strictly proper.

3. Operational estimate: Replace x_N by the actual noisy observation y

$$\rightsquigarrow \varrho(s)\tilde{x}^{(n)}(0;N) = \sum_{\text{finite}} \varrho_\ell(s)\frac{d^\ell}{ds^\ell}\hat{y} \qquad (*)$$

4. Back in time domain: If $\varrho(s) \bullet\!\!-\!\!\circ\; \rho(t)$ and $\varrho_\ell(s) \bullet\!\!-\!\!\circ\; h_\ell(t)$, then

$$(*) \bullet\!\!-\!\!\circ\; \rho(t)\tilde{x}^{(n)}(0;N) = \int_0^t \sum_{\text{finite}} \left\{h_\ell(t-\lambda)(-1)^\ell\lambda^\ell\right\} y(\lambda)d\lambda$$

where $t =$ estimation time.

Algebraic approach: algorithm description

Let $N \geqslant n$, consider $x_N(t) = \sum_{i=0}^{N} x^{(i)}(0)\frac{t^i}{i!}$ and set $x_N(t) \circ\!\!-\!\!\bullet\ \hat{x}_N(s)$.

$$s^{N+1}\hat{x}_N(s) = s^N x(0) + s^{N-1}\dot{x}(0) + \cdots + s^{N-n}x^{(n)}(0) + \cdots + x^{(N)}(0)$$

1. Annihilators: Find $\Pi = \sum_{\text{finite}} \varrho_\ell(s)\frac{d^\ell}{ds^\ell}$, $\quad \varrho_\ell(s) \in \mathbb{C}(s)$, such that
$$\Pi\hat{x}_N = \varrho(s)x^{(n)}(0), \quad \varrho(s) \in \mathbb{C}(s)$$
2. Integral operators: Left multiply by $\sigma(s) \in \mathbb{C}(s)$ such that
$$\sigma(s)\Pi \text{ and } \sigma(s)\varrho(s) \text{ strictly proper.}$$
3. Operational estimate: Replace x_N by the actual noisy observation y
$$\rightsquigarrow \varrho(s)\tilde{x}^{(n)}(0;N) = \sum_{\text{finite}} \varrho_\ell(s)\frac{d^\ell}{ds^\ell}\hat{y} \qquad (*)$$
4. Back in time domain: If $\varrho(s) \bullet\!\!-\!\!\circ\ \rho(t)$ and $\varrho_\ell(s) \bullet\!\!-\!\!\circ\ h_\ell(t)$, then
$$(*) \bullet\!\!-\!\!\circ\ \rho(t)\tilde{x}^{(n)}(0;N) = \int_0^t \sum_{\text{finite}} \left\{h_\ell(t-\lambda)(-1)^\ell\lambda^\ell\right\} y(\lambda)d\lambda$$
where $t =$ estimation time.

Algebraic approach: algorithm description

Let $N \geqslant n$, consider $x_N(t) = \sum_{i=0}^{N} x^{(i)}(0)\frac{t^i}{i!}$ and set $x_N(t) \circ\!\!-\!\!\bullet\ \hat{x}_N(s)$.

$$s^{N+1}\hat{x}_N(s) = s^N x(0) + s^{N-1}\dot{x}(0) + \cdots + s^{N-n}x^{(n)}(0) + \cdots + x^{(N)}(0)$$

1. Annihilators: Find $\Pi = \sum_{\text{finite}} \varrho_\ell(s)\frac{d^\ell}{ds^\ell}$, $\quad \varrho_\ell(s) \in \mathbb{C}(s)$, such that

$$\Pi\hat{x}_N = \varrho(s)x^{(n)}(0), \quad \varrho(s) \in \mathbb{C}(s)$$

2. Integral operators: Left multiply by $\sigma(s) \in \mathbb{C}(s)$ such that

$$\sigma(s)\Pi \text{ and } \sigma(s)\varrho(s) \text{ strictly proper.}$$

3. Operational estimate: Replace x_N by the actual noisy observation y

$$\rightsquigarrow \varrho(s)\tilde{x}^{(n)}(0;N) = \sum_{\text{finite}} \varrho_\ell(s)\frac{d^\ell}{ds^\ell}\hat{y} \qquad (*)$$

4. Back in time domain: If $\varrho(s) \bullet\!\!-\!\!\circ\ \rho(t)$ and $\varrho_\ell(s) \bullet\!\!-\!\!\circ\ h_\ell(t)$, then

$$(*) \bullet\!\!-\!\!\circ\ \rho(t)\tilde{x}^{(n)}(0;N) = \int_0^t \sum_{\text{finite}} \left\{h_\ell(t-\lambda)(-1)^\ell\lambda^\ell\right\} y(\lambda)d\lambda$$

where $t =$ estimation time.

5. Numerical estimate: Fix the estimation time $t = T : \rho(T) \neq 0$,

$$\tilde{x}^{(n)}(0;N) = \int_0^T h(\lambda;T)y(\lambda)d\lambda$$

With $0 \to \tau$:

$$\tilde{x}^{(n)}(\tau;N) = \int_0^T h(\lambda;T)y(\tau+\lambda)d\lambda$$

With $y = x + \varpi = x_N + R_N + \varpi$,

$$\tilde{x}^{(n)}(\tau;N) = x^{(n)}(\tau) + e_{R_N}(\tau) + e_{\varpi}(\tau)$$

UNIVERSITÉ DE REIMS CHAMPAGNE-ARDENNE

Purely integral estimators

For any $\kappa, \mu \in \mathbb{N}$, the differential operator

$$\Pi_{\kappa,\mu}^{N,n} = \frac{1}{s^{N+\mu+1}} \frac{d^{n+\kappa}}{ds^{n+\kappa}} \cdot \frac{1}{s} \cdot \frac{d^{N-n}}{ds^{N-n}} \cdot s^{N+1} \tag{1}$$

is a finite-integral form annihilator for $x^{(n)}(0)$ and, it is associated with

$$\varphi_{\kappa,\mu,N}(s) = \frac{(-1)^{(n+\kappa)}(n+\kappa)!(N-n)!}{s^{\mu+\kappa+N+n+2}}. \tag{2}$$

$\rightsquigarrow$ a family of estimators $\tilde{x}^{(n)}(0;\kappa,\mu;N)$:

$$\varphi_{\kappa,\mu,N}(s)\tilde{x}^{(n)}(0;\kappa,\mu;N) = \Pi_{\kappa,\mu}^{N,n}\hat{y}$$

For $N = n$, $\tilde{x}^{(n)}(0;\kappa,\mu;N) \rightarrow \tilde{x}^{(n)}(0;\kappa,\mu)$

Minimal estimator

Thm: If $q = N - n \geqslant 0$, then:

$$\tilde{x}^{(n)}(0;\kappa,\mu;N) = \sum_{\ell=0}^{q} \lambda_\ell \, \tilde{x}^{(n)}(0;\kappa_\ell,\mu_\ell), \quad \lambda_\ell \in \mathbb{Q} \tag{3}$$

where $\kappa_\ell = \kappa + q + \ell$ and $\mu_\ell = \mu + \ell$.
Moreover, if $q \leqslant n + \kappa$, then the coefficients λ_ℓ satisfy $\sum_{\ell=0}^{q} \lambda_\ell = 1$.

$\tilde{x}^{(n)}(0;\kappa,\mu)$ is termed minimal.

Université de Reims Champagne-Ardenne

Least-squares interpretation

Notations: •) $x_{LS,m}^{(n)}(\tau) = m^{th}$-order least-squares polynomial approximation of $x^{(n)}(\tau)$ for $\tau \in [0, T]$.

••) $P_i^{\{\kappa,\mu\}}(t)$: Jacobi orthogonal polynomials with weight $w_{\kappa,\mu}(t) = t^{\kappa+n}(1-t)^{\mu+n},\ t \in [0,\, 1]$

We show that $\tilde{x}^{(n)}(0;\kappa,\mu)$ is given by

$$\tilde{x}^{(n)}(0;\kappa,\mu) = x_{LS,1}^{(n)}(T\xi_1) + e_{\varpi}(0;\kappa,\mu), \tag{4}$$

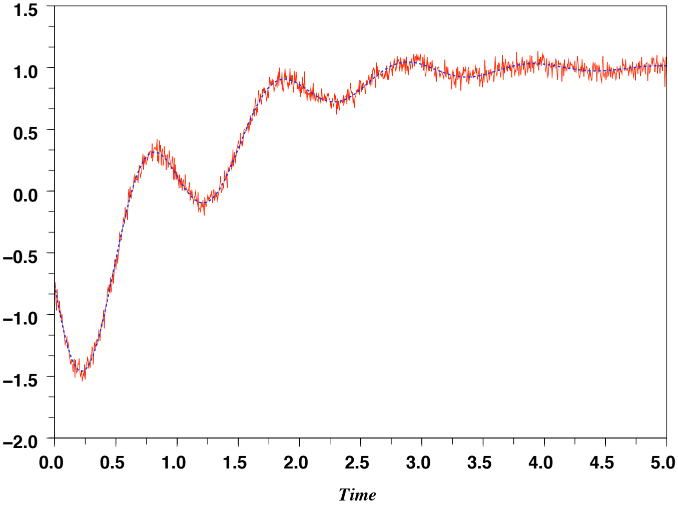
where $\xi_1 = \dfrac{\kappa+n+1}{\mu+\kappa+2(n+1)}$ = zero of $P_1^{\{\kappa,\mu\}}(t)$.

Replace $y(t)$ by $-[\mathrm{Heaviside}(t)y(\tau-t)]$, for $\tau \geqslant T$. Then

$$\tilde{x}^{(n)}(t;\kappa,\mu) \approx x^{(n)}(t-\tau_1),\, t \geqslant T$$

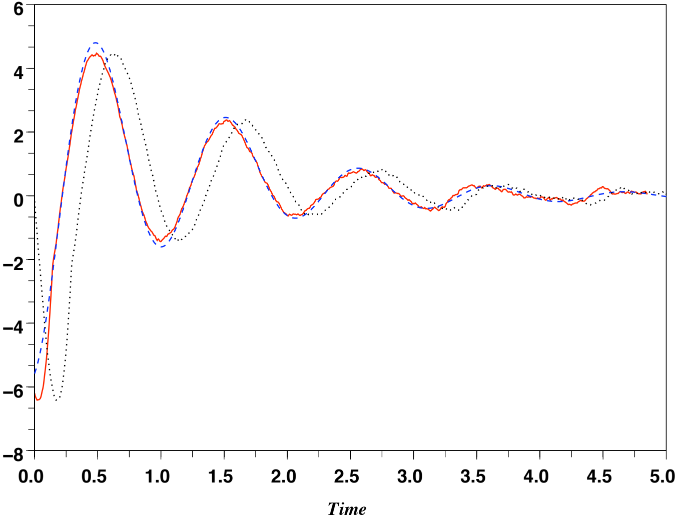
Minimal estimators introduce a delay given by $\tau_1 = T\xi_1$.

Simulation example 1



Noisy observation signal, $SNR = 25dB$.

Simulation example 1 ...



Estimation of the signal derivative: minimal $(0,0)$

Least-squares interpretation: non minimal estimators

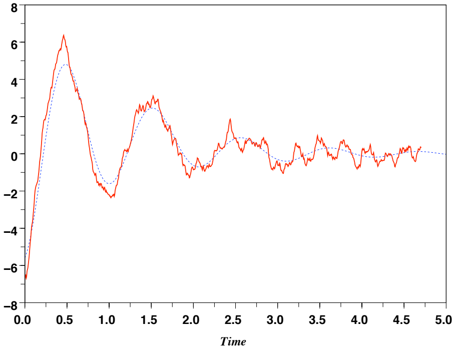
Thm: Set $q = N - n$. If $q \leqslant \kappa + n$ then

$$\tilde{x}^{(n)}(0;\kappa,\mu;N) = x^{(n)}_{LS,q}(0) + e_{\varpi}(0;\kappa,\mu;N), \tag{5}$$

i.e. $\tilde{x}^{(n)}(t;\kappa,\mu;N) \approx x^{(n)}(t)$

Non minimal estimators are delay free.

Simulation example 2



Non minimal, delay-free derivative estimation.

The abscence of delay undergoes a performance loss $\rightsquigarrow$ another paradigm:
a simple and approximate model may outperform a more precise one.

Delay and precision

Thm: For any $\xi \in [0,1]$, there exists a unique set of real coordinates $\lambda_\ell(\xi) \in \mathbb{R}, \ell = 0, \cdots, q$ such that

$$\sum_{\ell=0}^{q} \lambda_\ell(\xi)\, \tilde{x}^{(n)}(t;\kappa_\ell,\mu_\ell) = x_{LS,q}^{(n)}(t - T\xi) + e_{\varpi}^{\xi}(t;\kappa,\mu;n+q), \tag{6}$$

for some noise contribution $e_{\varpi}^{\xi}(t;\kappa,\mu;n+q)$.
Moreover, these coordinates satisfy

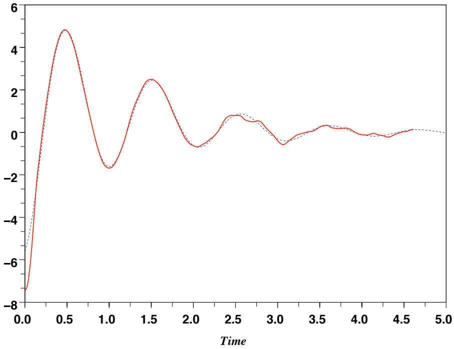
$$\sum_{\ell=0}^{q} \lambda_\ell(\xi) = 1$$

and, unless all the $\lambda_\ell(\cdot)$ are equal to zero except one, the following holds:

$$\min_\ell \lambda_\ell(\cdot) < 0.$$

Simulation example 3: $n = 1, N = 2$

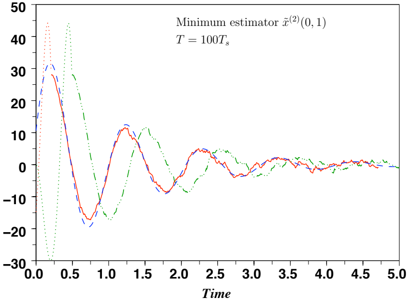
Choose ξ = zero of $P_2^{\{\kappa,\mu\}}(\cdot)$

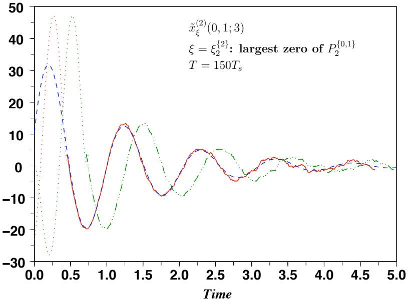


Non minimal, delayed derivative estimation.

The delay improves the estimation quality

Simulation example 4: second order derivative $n = 2$





Minimal estimator, $\kappa = 0$, $\mu = 1$.

Non minimal, delayed estimator, $N = 3, \kappa = 0$, $\mu = 1$.

Delay and noise

$$\sum_{\ell=0}^{q} \lambda_\ell = 1 \Rightarrow \sum_{\ell=0}^{q} \lambda_\ell \, \tilde{x}^{(n)}(t;\kappa_\ell,\mu_\ell) = (h^{?}_{n,\kappa,\mu} \star y)(t) = \tilde{x}^{(n)}_{?}(t)$$

Affine combination $\iff$ $\tilde{x}^{(n)}_{?}(t)$: least-squares sense

Convex combination: for noise reduction

Thm: $\lambda_0 = \lambda_1 = \cdots = \lambda_q \Longrightarrow \tilde{x}^{(n)}_{?}(t)$: Minimum residual noise variance.

Let $h^{MV}_{n,\kappa,\mu}(\tau)$: Minimum noise Variance estimator filter

$h^{Min}_{n,\kappa,\mu}(\tau)$: Minimal estimator filter.

For $q = 1$, $\quad h^{MV}_{n,\kappa,\mu}(\tau) = \frac{\kappa+\mu+2n+2}{2(\mu+n+1)} h^{Min}_{n,\kappa,\mu}(\tau) + \frac{\mu-\kappa}{2(\kappa+\mu+2n+2)} h^{Min}_{n,\kappa+1,\mu}(\tau)$

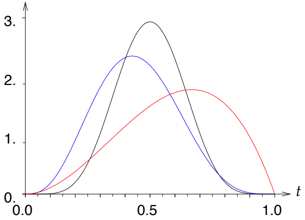
Bias and Variance are reconciled by the delay:

"Patience wears away stones"

Université de Reims Champagne-Ardenne

Tuning the parameters T, κ and μ

- If $T \nearrow \implies \begin{cases} \bullet \text{ bias } \nearrow \text{ and variance } \searrow \\ \bullet \text{ delay } \nearrow \end{cases}$
- For T fixed, $\kappa, \mu \nearrow \implies T_{effective} \searrow$



$$w_{\kappa,\mu}(t) = \alpha_{\kappa,\mu} t^{\kappa}(1-t)^{\mu}$$
$$\int_0^1 w_{\kappa,\mu}(t)dt = 1$$

$\kappa_1 < \kappa_2 < \kappa_3$
and $\mu_1 < \mu_2 < \mu_3$

Tuning the parameters

To be formal, we refer to the work of
D. Liu, W. Perruquetti, and O. Gibaru for

- An extension to κ, μ real and negative (> -1)
  $\rightsquigarrow$ smaller delay, smaller variance,
- A more complete analysis of the influence of the different error sources.

See

- Dayan, "Error analysis of Jacobi derivative estimators for noisy signals", PhD thesis, Lille, 2011.
- D. Liu, O. Gibaru, and W. Perruquetti, "Convergence Rate of the Causal Jacobi Derivative Estimator". Lecture notes in computer science, 2011.
- D. Liu, O. Gibaru, and W. Perruquetti, "Differentiation by integration with Jacobi polynomials". J. of Comput. and Applied Math., 2011.
- D. Liu, O. Gibaru, and W. Perruquetti, "Error analysis of Jacobi derivative estimators for noisy signals". Numerical Algorithms, 2011.

UNIVERSITÉ DE REIMS CHAMPAGNE-ARDENNE

Thank you for your patience