Kriging-based estimation of time derivatives via FIR filtering

Julien Marzat, Hélène Piet-Lahanier, Eric Walter

HAL Id: hal-00877181

https://hal.archives-ouvertes.fr/hal-00877181

Submitted on 27 Oct 2013

KRIGING-BASED ESTIMATION OF TIME DERIVATIVES VIA FIR FILTERING

Julien Marzat, Hélène Piet-Lahanier

ONERA – The French Aerospace Lab
F-91123 Palaiseau, France
julien.marzat@onera.fr

Éric Walter

Laboratoire des Signaux et Systèmes (L2S)
CNRS – SUPELEC – Univ-Paris-Sud
F-91192 Gif sur Yvette, France

ABSTRACT

Time derivatives of a noisy signal are estimated using a Kriging-based filter. Under the assumption of regularly spaced samples, the resulting differentiator turns out to be a finite impulse response filter (FIR) with constant coefficients. The method is employed to estimate the acceleration of a maneuvering target in the context of missile guidance and the results are compared with those obtained with conventional Kalman filtering.

Index Terms— FIR filtering, Kriging, numerical differentiation, noisy data, target tracking

1. INTRODUCTION

Accurate estimation of time derivatives of a noisy signal is often required in signal processing or control. This problem has received considerable attention in the literature. In particular, "model-free" control methods use estimates of time derivatives for computing the control input, thus bypassing the construction of dynamical process models [1].

The finite-difference method is usually inapplicable, since small perturbations of the signal lead to large errors in the computed derivatives [2]. Another widespread method, Savitzky-Golay filtering [3], fits polynomials on a sliding window and uses their derivatives as estimates. Other surrogate models, such as splines or wavelets, can also be used. A comparison of methods based on these approaches can be found in [4].

Techniques based on dynamical models and observers have been investigated [5], in particular with sliding modes [6]. Underlying dynamical models should be defined to make these techniques applicable. Algebraic numerical differentiation from noisy signals has also received recent attention [7, 8].

In the field of surrogate modeling, Kriging [9] (also known as Gaussian process regression [10]) is a widely-used approach that provides the best linear unbiased prediction on input-output data under some precisely stated hypotheses. Methods for evaluating derivatives based on Kriging have been proposed in [11, 12]. The method advocated in

this paper elaborates from these previous works to propose a practical scheme for estimating successive time derivatives at a low computational cost, under the form of a finite impulse response (FIR) filter with constant coefficients, in the case of regularly spaced samples.

Section 2 presents the Kriging-based method for estimating derivatives and the associated filter. Section 3 describes its application to the tracking of a maneuvering target in the context of missile guidance.

2. KRIGING-BASED EVALUATION OF DERIVATIVES

2.1. Kriging

Consider a set of n instants of time $\mathcal{T}_n = \{t_1, ..., t_n\}$ and the corresponding measurements of a signal $f(t)$, pooled in the vector $\mathbf{f}_n = [f(t_1), ..., f(t_n)]^{\mathrm{T}}$.

In its simplest form, Kriging models $f(\cdot)$ as a zero-mean Gaussian Process (GP) $F(\cdot)$ with covariance function $\mathrm{cov}(\cdot, \cdot)$ [9]. This covariance function is usually assumed stationary and written as

$$\mathrm{cov}\left(F\left(t_i\right), F\left(t_j\right)\right) = \sigma_F^2 R\left(t_i - t_j\right), \qquad (1)$$

where σ_F^2 is the GP variance and $R\left(\cdot\right)$ a parametric correlation function. The widely-used correlation function adopted here is

$$R\left(t_i - t_j\right) = \exp\left(-\frac{\left(t_i - t_j\right)^2}{\theta^2}\right), \qquad (2)$$

where θ is a scale factor to be chosen. This choice is adequate for smooth functions, yet many other correlation functions may be considered. Under these assumptions, Kriging makes it possible to predict the value of the function f at any instant of time t by

$$\widehat{f}(t) = \boldsymbol{\lambda}^{\mathrm{T}}(t)\mathbf{f}_n, \qquad (3)$$

where $\boldsymbol{\lambda}(t)$ is the solution of the linear system

$$\mathbf{R}\boldsymbol{\lambda}(t) = \mathbf{r}(t), \qquad (4)$$

in which

$$\begin{cases} \mathbf{r}\,(t) = [R\,(t - t_1)\,, \ldots, R\,(t - t_n)]^{\mathrm{T}}\,, \\ \mathbf{R}_{|ij} = R\,(t_i - t_j)\,. \end{cases} \quad (5)$$

The value computed by (3) is the best linear unbiased predictor of the value taken by $f(\cdot)$ at any t [13] (in the simple case considered here, it is also a minimum MSE linear predictor). This prediction interpolates the data, which is not appropriate in the presence of measurement noise. Zero-mean additive Gaussian white noise with variance σ_{N}^2 can be taken into account in the prediction [10] by replacing (4) by

$$\left(\mathbf{R} + \sigma_{\mathrm{N}}^2 \mathbf{I}_n\right) \boldsymbol{\lambda}(t) = \mathbf{r}(t). \quad (6)$$

In the following, $\mathbf{R} + \sigma_{\mathrm{N}}^2 \mathbf{I}_n$ is denoted by $\mathbf{R}_\sigma$. An interesting feature of GPs is their ability to estimate the variance of the prediction error, which can be used as a confidence measure in the prediction (3), as

$$\widehat{\sigma}^2(t) = \sigma_F^2 \left(1 - \boldsymbol{\lambda}^{\mathrm{T}}(t)\mathbf{r}(t)\right), \quad (7)$$

where the GP variance σ_F^2 can be estimated by maximum-likelihood,

$$\widehat{\sigma}_F^2 = \frac{1}{n}\mathbf{f}_n^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{f}_n. \quad (8)$$

As in (6), measurements corrupted by an additive zero-mean Gaussian white noise with variance σ_{N}^2 can be handled by using $\mathbf{R}_\sigma$ instead of $\mathbf{R}$ in (8).

2.2. Estimation of signal derivatives

In the predictor (3), only $\boldsymbol{\lambda}(t)$ depends on t. Therefore, an estimate of the l-th time derivative of f is

$$\widehat{f}^{(l)}(t) = \boldsymbol{\lambda}_l^{\mathrm{T}}(t)\mathbf{f}_n, \quad (9)$$

where $\boldsymbol{\lambda}_l(t)$ is the solution of the linear system

$$\mathbf{R}_\sigma \boldsymbol{\lambda}_l(t) = \mathbf{r}^{(l)}(t), \quad (10)$$

in which

$$\mathbf{r}^{(l)}\,(t) = \left[R^{(l)}\,(t - t_1)\,, \ldots, R^{(l)}\,(t - t_n)\right]^{\mathrm{T}}. \quad (11)$$

It is of course necessary for the correlation function R to be differentiable at least l times. Following [11], confidence intervals for the l-th derivative can be computed from

$$\widehat{\sigma}_l^2(t) = \widehat{\sigma}_F^2 \left((-1)^l R^{(2l)}(0) - \boldsymbol{\lambda}_l^{\mathrm{T}}(t)\mathbf{r}^{(l)}(t)\right). \quad (12)$$

This formula requires the correlation function to be differentiable $2l$ times. The correlation function (2) used in this work is appropriate, since it is infinitely differentiable.

Consider now the implementation of this filter on a sliding window of n points for estimating $f^{(l)}$, which is a common

strategy. At time t_k, the input data (time instants) and the corresponding measurements, assumed to be corrupted by an additive zero-mean Gaussian white noise with variance σ_{N}^2, are such that

$$\begin{aligned} \mathbf{t}_k &= [t_{k-n+1}, \ldots, t_k]^{\mathrm{T}}\,, \\ \mathbf{f}_k &= [f\,(t_{k-n+1})\,, \ldots, f\,(t_k)]^{\mathrm{T}}\,. \end{aligned} \quad (13)$$

From this data, the l-th derivative of f at time t is evaluated by

$$\widehat{f}^{(l)}(t) = \boldsymbol{\lambda}_l^{\mathrm{T}}(t)\mathbf{f}_k, \quad (14)$$

where $\boldsymbol{\lambda}_l(t)$ is the solution of

$$\left(\mathbf{R}_k + \sigma_{\mathrm{N}}^2 \mathbf{I}_n\right) \boldsymbol{\lambda}_l(t) = \mathbf{r}_k^{(l)}\,(t)\,, \quad (15)$$

in which

$$\begin{cases} \mathbf{r}_k^{(l)}\,(t) = \left[R^{(l)}\,(t - t_{k-n+1})\,, \ldots, R^{(l)}\,(t - t_k)\right]^{\mathrm{T}}\,, \\ \mathbf{R}_{k|ij} = R\,(t_{k-n+i} - t_{k-n+j})\,. \end{cases} \quad (16)$$

Note that it is preferable to choose $t_{k-n+1} < t < t_k$, but this is not mandatory. Interesting locations to compute the derivatives could be, for example, the last time instant t_k or the middle of the sliding window. It is also possible to use Kriging for predicting the value of the derivatives at the next time instant t_{k+1}, which may be useful in predictive schemes.

The implementation of this derivative filter on a sliding window requires computing $\mathbf{r}_k$ and $\mathbf{R}_k$ (most of their components being available from time t_{k-1}), and solving (15) for $\boldsymbol{\lambda}_l(t)$ at each time step. The 0-th order derivative is equivalent to the Kriging prediction (6) and can be used to filter the data, using prior knowledge on the noise standard deviation σ_{N}.

2.3. Formulation in the regular sampling case

The method for estimating derivatives recalled in the previous section can be used when sampling is irregular. However, it is assumed in this paper that sampling is performed with a constant time step t_{s}, as is usual in signal processing. It is also assumed that the scale factor θ, involved in the correlation function (2), is kept constant.

With a constant sampling time step t_{s}, the data from (13) at time t_k can be rewritten as

$$\begin{aligned} \mathbf{t}_k &= [t_k - (n-1)t_{\mathrm{s}}, \ldots, t_k - t_{\mathrm{s}}, t_k]^{\mathrm{T}}\,, \\ \mathbf{f}_k &= [f\,(t_k - (n-1)t_{\mathrm{s}})\,, \ldots, f\,(t_k - t_{\mathrm{s}})\,, f\,(t_k)]^{\mathrm{T}}\,. \end{aligned} \quad (17)$$

Assume, for example, that the l-th derivative is to be evaluated at $t = t_k$ (other time instants can be handled by shifting indices). Equation (16) then translates into

$$\begin{cases} \mathbf{r}_k^{(l)}\,(t_k) = \left[R^{(l)}\,((n-1)t_{\mathrm{s}})\,, \ldots, R^{(l)}\,(t_{\mathrm{s}})\,, R^{(l)}\,(0)\right]^{\mathrm{T}}\,, \\ \mathbf{R}_{k|ij} = R\,((i-j)t_{\mathrm{s}})\,. \end{cases} \quad (18)$$

Since $\mathbf{r}_k^{(l)}$ and $\mathbf{R}_k$ in (18) are no longer function of k or t_k, (14) can be expressed at time t_k as the FIR filter

$$\widehat{f}^{(l)}(t_k) = \mathbf{h}_l^{\mathrm{T}} \mathbf{f}_k, \qquad (19)$$

where $\mathbf{h}_l$ is solution of

$$\left(\mathbf{R} + \sigma_{\mathrm{N}}^2 \mathbf{I}_n \right) \mathbf{h}_l = \mathbf{r}_l \qquad (20)$$

with

$$\begin{cases} \mathbf{r}_l = \left[R^{(l)} \left((n-1)t_{\mathrm{s}} \right), \dots, R^{(l)} \left(t_{\mathrm{s}} \right), R^{(l)} \left(0 \right) \right]^{\mathrm{T}}, \\ \mathbf{R}_{|ij} = R \left((i-j)t_{\mathrm{s}} \right). \end{cases}$$

$$(21)$$

The expression of $\mathbf{h}_l$ is now the same at each time step and can thus be computed once and for all, provided that the following tuning parameters have been chosen:

- n (sliding window length),

- σ_{N} (noise standard deviation),

- θ (scale factor, similar to a classical filter bandwith).

The time step t_{s} could be included in the tuning parameters if it is not considered as a given.

It is also possible to compute confidence intervals for the value of the derivative based on the following variance estimate, obtained by combining (8) and (12),

$$\widehat{\sigma}_l^2(t_k) = \frac{1}{n} \left(\mathbf{f}_k^{\mathrm{T}} \mathbf{R}_{\sigma}^{-1} \mathbf{f}_k \right) \left((-1)^l R^{(2l)}(0) - \mathbf{h}_l^{\mathrm{T}} \mathbf{r}_l \right), \quad (22)$$

where only $\mathbf{f}_k$ changes at each time step.

For example, if the correlation function (2) is used and the first derivative is to be computed at $t = t_k$, then (21) implies that

$$\mathbf{r}_1 = \begin{bmatrix} -\frac{2(n-1)t_{\mathrm{s}}}{\theta^2} R((n-1)t_{\mathrm{s}}) \\ \vdots \\ -\frac{2t_{\mathrm{s}}}{\theta^2} R(t_{\mathrm{s}}) \\ 0 \end{bmatrix}, \qquad (23)$$

$$\mathbf{R} = \begin{bmatrix} 1 & R(t_{\mathrm{s}}) & \cdots & R((n-1)t_{\mathrm{s}}) \\ R(t_{\mathrm{s}}) & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & R(t_{\mathrm{s}}) \\ R((n-1)t_{\mathrm{s}}) & \cdots & R(t_{\mathrm{s}}) & 1 \end{bmatrix}.$$

$$(24)$$

2.4. Illustrative example

The function $f(t) = \sin(2t)$ is considered as an illustrative example, with time step $t_{\mathrm{s}} = 0.01\mathrm{s}$ and measurements corrupted by an additive zero-mean Gaussian white noise with standard deviation $\sigma_{\mathrm{N}} = 0.1$. Figure 1 shows the measured signal, its smoothing by Kriging and the first-order derivative estimated via finite difference, which is unusable. Figure 2 presents the Kriging-based estimation of the first-order

derivative with parameters $n = 100$ and $\theta = 1.5$. For comparison, the estimate obtained with a Savitzky-Golay filter using a first-order polynomial is also presented (no estimate is computed before n measurements have been acquired). Both methods achieve satisfactory filtering of the derivative, however it could be seen on this example that the polynomial fit induces a larger delay in the estimation than the Kriging-based filter. Kriging also provides 99% confidence intervals for its estimates, computed as $\widehat{f}^{(l)}(t) \pm 3\widehat{\sigma}_l(t)$ using (22).
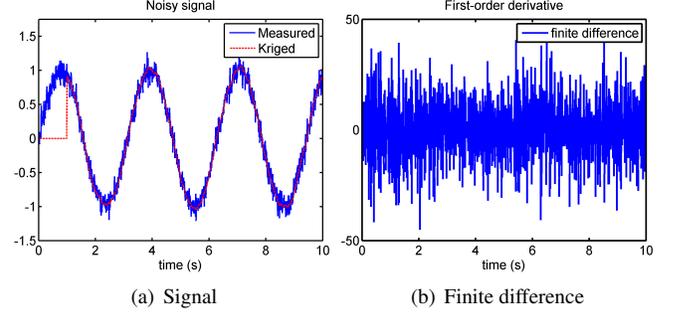

(a) Signal   (b) Finite difference

Fig. 1. Noisy signal (a) and first-order finite difference (b)

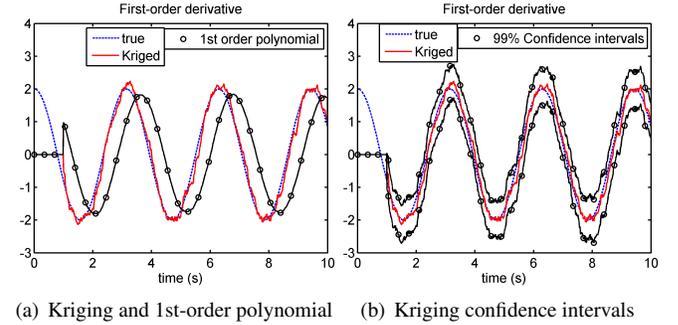(a) Kriging and 1st-order polynomial   (b) Kriging confidence intervals

Fig. 2. First-order derivative estimate via Kriging

2.5. Tuning

The method proposed, as well as many other methods for estimating signal derivatives, requires the tuning of its internal parameters (here n, θ and σ_{N}), which have a strong impact on performance. In particular, for real-time control applications where an estimate should be obtained at each time step, a trade-off should be achieved between noise filtering and estimation delay. A suitable performance criterion can be defined on a set of representative test-cases and a global optimization strategy can be employed to find the best vector of tuning parameters. Such a strategy has been proposed in [14] using Kriging-based optimization, making it possible to find an adequate tuning with very few simulations of (possibly very complex) test cases. An extension has also been presented in [15] to take into account sources of uncertainty

in the test cases and determine the best possible tuning with respect to the worst case of uncertainty.

Figure 3 shows the influence of the tuning parameters θ and σ_N on the mean squared error between the estimated first-order derivative and the actual one, which can be used as a performance criterion for tuning. It can be seen that many combinations $\{\theta, \sigma_N\}$ provide an accurate estimation.
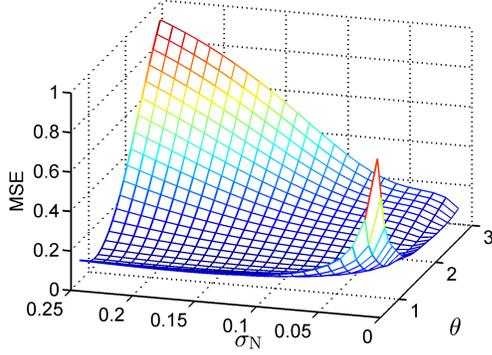
Fig. 3. Influence of the tuning parameters θ and σ_N

3. APPLICATION TO THE TRACKING OF A MANEUVERING TARGET

Endgame missile guidance aims at defining control laws that enable a missile to intercept a moving target, using radar measurements that may be embedded (seeker) or exogenous (ground radar). A maneuver (change in target acceleration) can be modeled as an unknown input acting on the joint missile-target dynamics. Classical methods, based for example on Kalman filters and simplified target modeling, have been developed to estimate the target acceleration [16]. This information is then used in a guidance law designed for ensuring collision between the missile and its target.
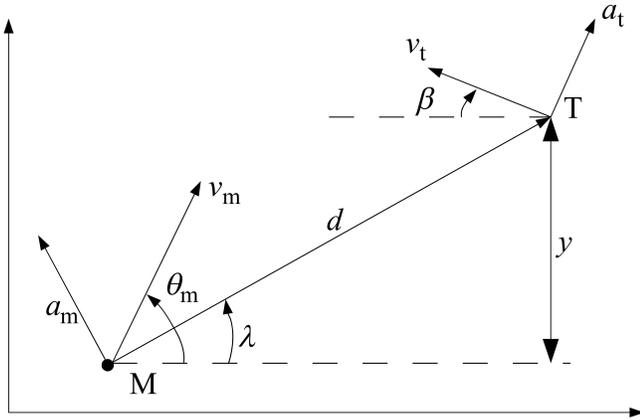
Fig. 4. Missile guidance geometry

The planar interception geometry is presented in Figure 4 [17]. The following notation is used:

- a_m is the missile acceleration (assumed to be instantaneously achieved), v_m the missile speed, θ_m the missile flight-path angle, and $\{x_m, y_m\}$ the missile position;

- a_t is the target acceleration, v_t the target speed, β the target flight-path angle, and $\{x_t, y_t\}$ the target position;

- the line of sight (LOS) is the missile-target position vector; its norm is the range d, λ is the line-of-sight angle, and v_c the closing velocity.

3.1. Kinematics

At a given time step, the target may perform a maneuver of acceleration magnitude a_t, resulting in

$$\dot{\beta} = \frac{a_t}{v_t}, \qquad v_t = \sqrt{v_{t,x}^2 + v_{t,y}^2},$$
$$\dot{x}_t = -v_t \cos(\beta), \quad \dot{y}_t = v_t \sin(\beta). \tag{25}$$

A guidange law provides the desired lateral acceleration a_m to the missile, such that

$$\dot{x}_m = v_{m,x}, \qquad \dot{y}_m = v_{m,y},$$
$$\dot{v}_{m,x} = -a_m \sin(\lambda), \quad \dot{v}_{m,y} = a_m \cos(\lambda). \tag{26}$$

Other important variables of the engagement are

$$x = x_t - x_m, \qquad y = y_t - y_m,$$
$$d = \sqrt{x^2 + y^2}, \qquad v_c = -\dot{d} = -\frac{x\dot{x} + y\dot{y}}{d},$$
$$\lambda = \arctan\left(\frac{y}{x}\right), \quad \dot{\lambda} = \frac{x\dot{y} - y\dot{x}}{d^2}. \tag{27}$$

3.2. Linearized model

The linearized planar engagement model is often considered for estimation and control analysis. It is assumed that the velocities v_m and v_t are almost aligned with the LOS and that λ is small. The relative vertical position y, which can be used as an approximation of the miss distance at the end of the engagement, is then such that

$$\ddot{y} = a_t - a_m. \tag{28}$$

If the target were not maneuvering, (28) would simplify into $\ddot{y} = -a_m$. A potential time-varying maneuver can thus be described by the additional unknown input a_t. The engagement is then governed by

$$\begin{bmatrix} \dot{y} \\ \ddot{y} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y \\ \dot{y} \end{bmatrix} + [] + \begin{bmatrix} 0 \\ -1 \end{bmatrix} a_m + \begin{bmatrix} 0 \\ 1 \end{bmatrix} a_t, \tag{29}$$

where an equivalent noisy measurement y_{mes} of y can be obtained using an embedded seeker that would in practice provide relative angular measurements, assuming that the range and possibly the range rate can be measured by a ground radar.

3.3. Classical method for target estimation

In the nominal case with no target maneuver, proportional navigation guidance (PNG) is widely used. It computes the acceleration to be given to the missile as

$$a_{\mathrm{m}} = N v_{\mathrm{c}} \dot{\lambda}, \tag{30}$$

with N a constant gain. The objective of this guidance law is to keep the direction of the LOS constant during the engagement, which ensures interception [18]. Augmented PNG (APNG) can be used to improve interception performance when dealing with a maneuvering target. It modifies (30) into

$$a_{\mathrm{m}} = N \left(v_{\mathrm{c}} \dot{\lambda} + \frac{a_{\mathrm{t}}}{2} \right), \tag{31}$$

where the target acceleration a_{t} should be estimated on-line with as small a delay as possible.

It is usually assumed that the target performs an evasive maneuver with constant (but unknown) acceleration a_{t}. This acceleration can be estimated by, e.g., a Kalman filter using the following classical estimation model [17]

$$\begin{bmatrix} \dot{y} \\ \ddot{y} \\ \dot{a}_{\mathrm{t}} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} y \\ \dot{y} \\ a_{\mathrm{t}} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} a_{\mathrm{m}} + \begin{bmatrix} 0 \\ 0 \\ w \end{bmatrix}, \tag{32}$$

where w is a zero-mean Gaussian white noise, whose variance is a tuning parameter of the Kalman filter. The covariance and gain matrices must be updated at each time step, and the resulting performance is dependent on the underlying target model (32). As an alternative, the Kriging-based filtering method for estimating derivatives can be considered.

## 3.4. Target tracking via Kriged estimation of derivatives

Since (28) implies that $a_{\mathrm{t}} = \ddot{y} + a_{\mathrm{m}}$, an estimate of a_{t} is easy to compute if an accurate estimate of the second-order time derivative of the noisy measurement of y can be provided, for example with the Kriging-based method proposed in this paper.

The performance of Kriging-based FIR and Kalman filtering have been compared on two typical trajectories: an evading maneuver of magnitude 20 m/s^2 (Figure 5) and a weaving maneuver of magnitude 70 m/s^2 and frequency 0.64 Hz (Figure 6). Figure 7 displays the first and second-order derivatives as estimated by FIR filtering, together with the actual values and estimated 99% confidence intervals. The engagement parameters, used for all the results presented, are $v_{\mathrm{m}} = 500$ m/s, $v_{\mathrm{t}} = 200$ m/s, $d = 3000$ m, $\beta = 5°$, equivalent noise standard deviation of 0.25 m on the measurement of y and time step equal to 0.01 s.
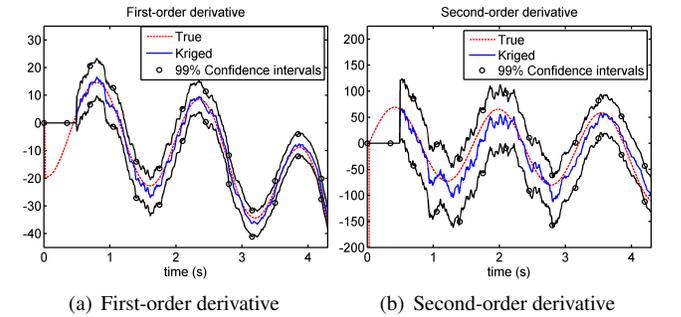
The Kriging-based scheme achieves a performance similar to that of the Kalman filter The Kriging-based method was tuned with $n = 50$, $\theta = 1.5$ and $\sigma_{\mathrm{N}} = 10^{-2}$ for the two maneuvers. Many good trade-offs between noise filtering and

(a) Kriging filter   (b) Kalman filter
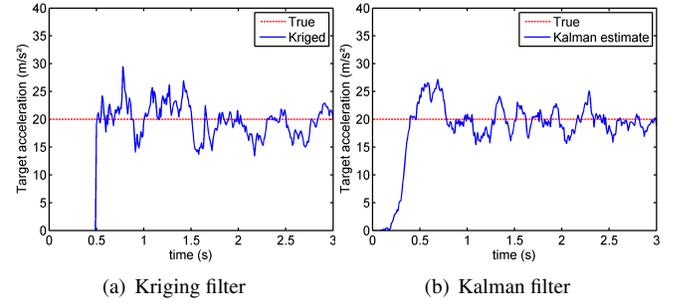
Fig. 5. Estimation of evasive target maneuver with Kriging (a) and Kalman (b) filters


(a) Kriging filter   (b) Kalman filter
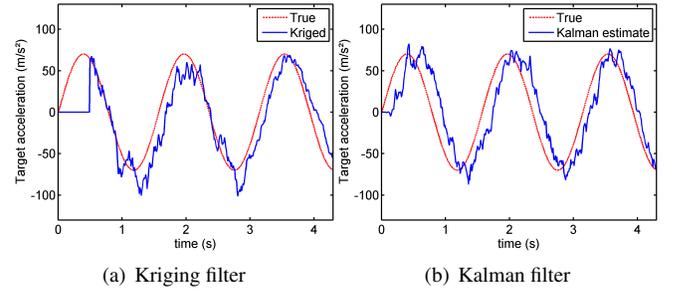
Fig. 6. Estimation of weaving target maneuver with Kriging (a) and Kalman (b) filters

(a) First-order derivative   (b) Second-order derivative

Fig. 7. First-order (a) and second-order (b) derivatives provided by the Kriging filter in the weaving case

estimation delay can be achieved by modifying the values of θ and σ_{N}, as seen in Figure 3. Note that increasing θ or σ_{N} smooths the prediction, so the situation slightly differs from the one of the Kalman filter where the process and noise variances have opposite effects.

The estimated acceleration can be used in the guidance loop with the APNG law (31). Figure 8 presents the comparison of PNG (without estimation) and APNG (with the Kriged estimate) for intercepting the evading target maneuvering at 20 m/s^2. It can be seen that the estimated acceleration is accurate enough to anticipate the target movement and facilitate interception.
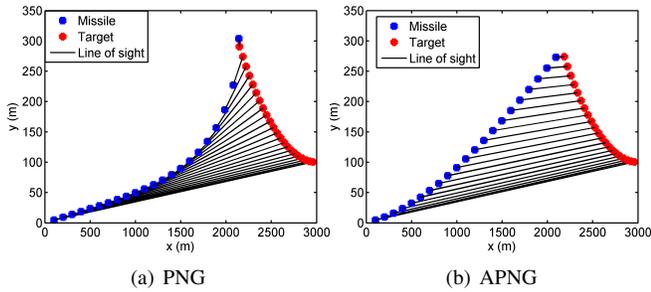
Fig. 8. Target interception with PNG (a) and APNG (b)

4. CONCLUSIONS AND PERSPECTIVES

A filter based on Gaussian processes has been proposed to estimate the time derivatives of a noisy signal. Under the assumption that the measurements are regularly sampled with a known time step, Kriging-based derivative estimation on a sliding window is carried out by an FIR filter. This simple method is thus an attractive alternative, which deserves to be further investigated. Its applicability has been illustrated for estimating the motion of a maneuvering target in the context of missile guidance with promising results. The optimal tuning of the filter parameters on more complex test cases can be performed by considering the methodology proposed in [14, 15].

5. REFERENCES

[1] M. Fliess and C. Join, "Model-free control and intelligent PID controllers: Towards a possible trivialization of nonlinear control ?," in *Proceedings of the 15th IFAC Symposium on System Identificaiton, Saint-Malo, France*, 2009, pp. 1531–1550.

[2] A. G. Ramm and A. B. Smirnova, "On stable numerical differentiation," *Mathematics of computation*, vol. 70, no. 235, pp. 1131–1153, 2001.

[3] A. Savitzky and M.J.E. Golay, "Smoothing and differentiation of data by simplified least squares procedures.," *Analytical chemistry*, vol. 36, no. 8, pp. 1627–1639, 1964.

[4] K. Ahnert and M. Abel, "Numerical differentiation of experimental data: local versus global methods," *Computer Physics Communications*, vol. 177, no. 10, pp. 764–774, 2007.

[5] B. Carlsson, A. Ahlén, and M. Sternad, "Optimal differentiation based on stochastic signal models," *IEEE Transactions on Signal Processing*, vol. 39, no. 2, pp. 341–353, 1991.

[6] G. Bartolini, A. Pisano, and E. Usai, "First and second derivative estimation by sliding mode technique," *Journal of Signal Processing*, vol. 4, no. 2, pp. 167–176, 2000.

[7] M. Mboup, C. Join, and M. Fliess, "Numerical differentiation with annihilators in noisy environment," *Numerical Algorithms*, vol. 50, pp. 439–467, 2009.

[8] J. Reger and J. Jouffroy, "On algebraic time-derivative estimation and deadbeat state reconstruction," in *Proceedings of the 48th IEEE Conference on Decision and Control, Shanghai, China*, 2009, pp. 1740–1745.

[9] G. Matheron, "Principles of geostatistics," *Economic Geology*, vol. 58, no. 8, pp. 1246–1266, 1963.

[10] C.E. Rasmussen and C.K.I. Williams, *Gaussian Processes for Machine Learning*, Springer-Verlag, New York, 2006.

[11] E. Vazquez and E. Walter, "Estimating derivatives and integrals with kriging," in *Proceedings of the 44th IEEE Conference on Decision and Control, Seville, Spain*, 2005, pp. 8156–8161.

[12] T. Holsclaw, B. Sanso, H.K.H. Lee, K. Heitmann, S. Habib, D. Higdon, and U. Alam, "Gaussian process modeling of derivative curves," *Technometrics*, vol. 55, no. 1, pp. 57–67, 2013.

[13] T. J. Santner, B. J. Williams, and W. Notz, *The Design and Analysis of Computer Experiments*, Springer-Verlag, Berlin-Heidelberg, 2003.

[14] J. Marzat, E. Walter, H. Piet-Lahanier, and F. Damongeot, "Automatic tuning via Kriging-based optimization of methods for fault detection and isolation," in *Proceedings of the IEEE Conference on Control and Fault-Tolerant Systems, Nice, France*, 2010, pp. 505–510.

[15] J. Marzat, E. Walter, and H. Piet-Lahanier, "Min-max hyperparameter tuning with application to fault detection," in *Proceedings of the 18th IFAC World Congress, Milan, Italy*, 2011, pp. 12904–12909.

[16] R.A. Singer and K.W. Behnke, "Real-time tracking filter evaluation and selection for tactical applications," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 7, no. 1, pp. 100–110, 1971.

[17] P. Zarchan, *Tactical and Strategic Missile Guidance*, AIAA Progress in Astronautics and Aeronautics, Washington DC, 1990.

[18] N. A. Shneydor, *Missile Guidance and Pursuit: Kinematics, Dynamics & Control*, Horwood Publishing Limited, Chichester, 1998.