Irregular sampling in the estimation of the response function

Vladimir Zaiats

Universitat de Vic, SPAIN

E-mails: vladimir.zaiats@uvic.es, zaiats@mat.uab.es

Some parts are based on joint works and collaborations with **F. Utzet** (Universitat Autònoma de Barcelona), **V. Buldygin** (National Technical University of Ukraine)

Talk at the SAPS 5

Le Mans, January 6, 2005

Plan of the talk

1 Linear physical systems

2 The problem

3 Random sampling: a survey

4 Results in the estimation of the response function

1. Linear physical systems

1.1. Linear systems with constant parameters

Ideal system — a system having constant parameters and such that the two basic characteristics, the **input** and the **output** are linearly related

A system has **constant parameters** if all basic properties of the system do not vary in time (time invariant)

A system is **linear** if its reaction to an input is additive and homogeneous

Additive: $f(x_1 + x_2) = f(x_1) + f(x_2)$

Homogeneous: $f(c \cdot x) = c \cdot f(x)$, for any $c = \text{const}$

1.2. Basic dynamical characteristics

$H(s)$, $s \in \mathbb{R}$ — the **impulse response function** or the **transfer function in the time domain**

Meaning: reaction of the system to a unit impulse (a pulse of unit area and infinitesimal width) enjoyed by the system s units of time before the instant where the output is measured

Then

$$y(t) = \int_{-\infty}^{\infty} H(s)x(t-s)ds = \int_{-\infty}^{\infty} x(s)H(t-s)ds, \qquad t \in \mathbb{R} \qquad (1)$$

Causal systems — those whose response at any instant does not depend on the future of the input, for any input

This means that (a necessary and sufficient condition for a system to be causal)

$$H(s) = 0, \quad s < 0, \tag{2}$$

and that the lower limit of integration in (1) equals 0 instead of $-\infty$ for causal systems

Stable systems — those in which any bounded input produces a bounded output

If

$$\int_{-\infty}^{\infty} |H(s)| ds < \infty, \tag{3}$$

then

$$|y(t)| \leq \sup_{t \in \mathbb{R}} |x(t)| \cdot \int_{-\infty}^{\infty} |H(s)| ds < \infty.$$

One can show that condition (3) is necessary and sufficient for system (1) to be stable

1.3. Stationary processes and their functional characteristics: a reminder

Assume that $X := (X(t),\ t \in \mathbb{R})$ and $Y := (Y(t),\ t \in \mathbb{R})$ are (weakly) stationary stochastic processes

One can assume that $\mathsf{E}X(t) = \mathsf{E}Y(t) \equiv 0$, $t \in \mathbb{R}$

We denote the **(auto)-correlation functions** and the **cross-correlation function** as follows:

$$K_{XX}(\tau) := \mathsf{E}[X(t)X(t+\tau)], \tag{4a}$$
$$K_{YY}(\tau) := \mathsf{E}[Y(t)Y(t+\tau)], \tag{4b}$$
$$K_{XY}(\tau) := \mathsf{E}[X(t)Y(t+\tau)], \qquad \tau \in \mathbb{R}. \tag{4c}$$

(4c) holds if X and Y are jointly stationary

Spectral densities:

$$f_{XX}(\lambda) := \int_{-\infty}^{\infty} K_{XX}(\tau) e^{-i\lambda\tau} d\tau \tag{5a}$$

$$f_{YY}(\lambda) := \int_{-\infty}^{\infty} K_{YY}(\tau) e^{-i\lambda\tau} d\tau \tag{5b}$$

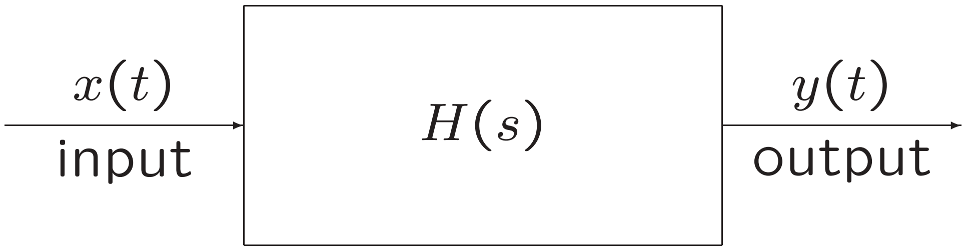
$$f_{XY}(\lambda) := \int_{-\infty}^{\infty} K_{XY}(\tau) e^{-i\lambda\tau} d\tau, \qquad \lambda \in \mathbb{R} \tag{5c}$$

f_{**} are defined if $K_{**} \in L_1(\mathbb{R})$. Fourier–Plancherel $\Rightarrow K_{**} \in L_2(\mathbb{R})$

1.4. SISO systems

SISO = single-input, single-output

MISO = multiple-input, single-output,   (MIMO, SIMO, . . .)



Assume that the input is a stationary stochastic process $(X(t),\ t \in \mathbb{R})$. Then

$$Y(t) = \int_{-\infty}^{\infty} H(s)X(t-s)ds = \int_{-\infty}^{\infty} X(s)H(t-s)ds, \qquad t \in \mathbb{R} \qquad (6)$$

By taking the corresponding means, we obtain

$$K_{YY}(\tau) := \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H(u)H(v)K_{XX}(\tau + v - u)dudv \qquad (7a)$$

$$K_{XY}(\tau) := \int_{-\infty}^{\infty} H(u)K_{XX}(\tau - u)du, \qquad \tau \in \mathbb{R} \qquad (7b)$$

(for causal systems, $-\infty$ is replaced by 0)

By the Fourier transform, we have

$$f_{YY}(\lambda) = |\mathcal{H}(\lambda)|^2 f_{XX}(\lambda), \qquad (8a)$$

$$f_{XY}(\lambda) = \mathcal{H}(\lambda) f_{XX}(\lambda), \qquad \lambda \in \mathbb{R}, \qquad (8b)$$

where $\mathcal{H}(\lambda)$, $\lambda \in \mathbb{R}$, is the so-called **frequency characteristic function** of system (1):

$$\mathcal{H}(\lambda) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} H(s)e^{-i\lambda s}ds, \qquad \lambda \in \mathbb{R} \qquad (9)$$

2. The problem

The problem of estimation of the (unknown) function H is of our interest

We need a guess on how this estimation can be done

Recall formula (7b):

$$K_{XY}(\tau) = \int_{-\infty}^{\infty} H(u) K_{XX}(\tau - u) du, \qquad \tau \in \mathbb{R}$$

Imagine that X is a white noise. Then

$$K_{XX}(\tau) = A \cdot \delta_0(\tau), \qquad \tau \in \mathbb{R},$$

where δ_0 is the Dirac delta. Substitute this formula into (7b) to obtain

$$K_{XY}(\tau) = A \cdot H(\tau), \qquad \tau \in \mathbb{R}.$$

That is, the input-output cross-correlation is proportional to H given that the system is enjoyed by a white noise

Of course, white noise is not a "physical waveform" since its spectral density is flat over all frequencies ($\Rightarrow$ the average power is infinite)

However, it is a mathematical idealization of a physical waveform for which the power density spectrum is flat over a band of frequencies considerably wider than the bandwidth of the system to which it is applied as input

The estimation of H has much in common with that of the correlation function, but one should note the following crucial difference:

The correlation function is always symmetric, that is $K(\tau) = K(-\tau)$ for all $\tau \in \mathbb{R}$, whereas the impulse response function H need not be symmetric!

Moreover, if a system is causal, then $H(s) = 0$, $\tau < 0$, meaning that the function H can never (!) be symmetric in causal systems

Therefore we intuitively claim that one should expect some profound peculiarities to be revealed in the estimation of the impulse response function, with respect to that of the correlation function

An outline of the technique we are going to apply for the estimation of the function H can be summarized as follows:

- We use stationary processes as inputs; these processes are "close," in some sense, to white noise

- We estimate H by means of a discrete-time sample input-output cross-correlogram

- (New) We assume that the framework yields irregular sampling

(We also expect that the above mentioned peculiarities related to the nature of the impulse response function will somehow emerge)

Irregular sampling in the estimation of correlation functions and spectral densities

• Proceedings volume: *Time Series Analysis of Irregularly Observed Data*, Lect. Notes in Statistics, vol. 25, 1984

• J. Jacod: Stoch. Proc. Appl. (1993)
parametric procedures maximizing the limit of the normalized Fisher information for all values of a parameter

• K.-S. Lee & E. Masry: J. Multivar. Anal. (1992), Stoch. Proc. Appl. (1994)

• K.-S. Lee & T. H. Tsou: J. Time Ser. Anal. (1995)

• E. Masry: Adv. Appl. Probab. (1983)

• M. Rachdi: J. Nonparametr. Stat. (2004)
spectral estimation in continuous time: periodic & Poisson sampling

Correlation function (latest results)

- R. A. Davis & T. Mikosch: Ann. Statist. (1998)

consistency of the sample ACF for heavy-tailed stationary processes

- R. A. Davis & T. Mikosch: Stoch. Process. Appl. (1999)

convergence of the sample ACF in simple bilinear processes

- T. Mikosch & C. Stărică: Ann. Statist. (2000)

the limit distribution of the sample autocorrelation in GARCH $(1, 1)$, among other results related to the distribution of the maximum

- S. N. Elogne, O. Perrin & C. Thomas-Agnan: SISP (2005)

Nonparametric estimation of smooth stationary covariance functions by interpolation methods (splines)

Impulse response function

- Buldygin & Fu Li: Theory Probab. Math. Statist. (1996)
- Buldygin & Kurotschka: ROSE (1998)
- Buldygin, Utzet & Z: Qüestiió (2002), SISP (2004)

periodic discrete sampling

Deconvolution problem

- M. Ermakov: Inverse problems (1990, 2003), J. Phys. A (1992)

Volterra series

- D. Bosq & O. Lessi: Statistica (Bologna) (1995)

3. Irregular sampling: a survey

Let $X(t)$, $t \in \mathbb{R}$, be a real second-order stationary process having a univariate probability density function $\varphi(x)$, covariance function $K(t)$, and spectral density function $f(\lambda)$. Suppose that the process X is observed at instants $\{\tau_k\}_{k=1}^{n}$ which may be deterministic or random.

In the context of the covariance and spectral estimation, the question of aliasing arises naturally when the sampling points are equally-spaced (Jenkings and Watts, 1968). Thus we first characterize those sampling schemes $\{\tau_k\}$ which are "alias-free" in a sense made precise later.

It turns out that there are two distinct concepts of "alias-free" sampling leading to, correspondingly, two distinct approaches to the estimation of $K(t)$ and $f(\lambda)$.

The first concept is due to Shapiro and Silverman, *J. Soc. Indust. Appl. Math.* (1960), and it leads to estimates $\widehat{K}_n(t)$ and $\widehat{f}_n(\lambda)$ which are necessarily of the orthogonal-series type. These estimates utilize the ordered data set $\{X(\tau_k)\}_{k=1}^n$ only, i. e., they do not require the knowledge of the actual values (realization) of the sampling instants $\{\tau_k\}$ – such circumstances were reported to occur in laser anemometry (Durranti and Greated. *Laser systems in flow measurements*, Plenum, 1977).

The second concept of alias-free sampling was developed by Masry, *IEEE Trans. Information Theory* (1978), and was inspired by the work of Brillinger (Proc. 6th Berkeley Symp. Prob. Statist., 1972) on the spectral analysis of stationary interval functions. This concept leads to estimates $\widehat{K}_n(t)$ and $\widehat{f}_n(\lambda)$ based on the data set $\{X(\tau_k),\ \tau_k\}_{k=1}^n$ — here the realization of the sampling instants $\{\tau_k\}_{k=1}^n$ must be known.

Record of the sampling instants is unavailable

Rates of convergence

$$\text{MSE for } f \text{ or } K \sim \frac{1}{(\ln n)^{r-1}}, \qquad n \to \infty,$$

(or $r-2$) where r is related to the smoothness of K.

Therefore the convergence rates in n are too slow

When a record of the sampling instants $\{\tau_k\}_{k=1}^{n}$ is available, in addition to the data $\{X(\tau_k)\}_{k=1}^{n}$, considerably faster quadratic-mean convergence rates are available as we are going to see now

Record of the sampling instants is available

For the estimation of the covariance function K_{XX} of a process X, Masry (1983) used the following estimate:

$$\widehat{K}_T(t) = \frac{1}{\beta^2 + c(t)} \left\{ \frac{1}{T} \sum_{i,j=1}^{N(T)} w_T(t - (\tau_i - \tau_j)) X(\tau_i) X(\tau_j) \right. \\ \left. - \frac{1}{T} \sum_{i=1}^{N(T)} w_T(t) X^2(\tau_i) \right\} \tag{10}$$

where $N(T) \stackrel{\text{def}}{=} N((0,T])$, $\beta \stackrel{\text{def}}{=} N((0,1])$, c is the density of the reduced covariance measure of $\{\tau_n\}$, and w_T is an appropriate averaging kernel.

The sample size is random; in this case the above estimate is studied as $T \to \infty$.

The averaging kernel is generated as follows. Let $w(x)$, $x \in \mathbb{R}$, be a real-valued symmetric continuous function satisfying:

(i) $\int_{-\infty}^{\infty} |w(x)| dx < \infty$

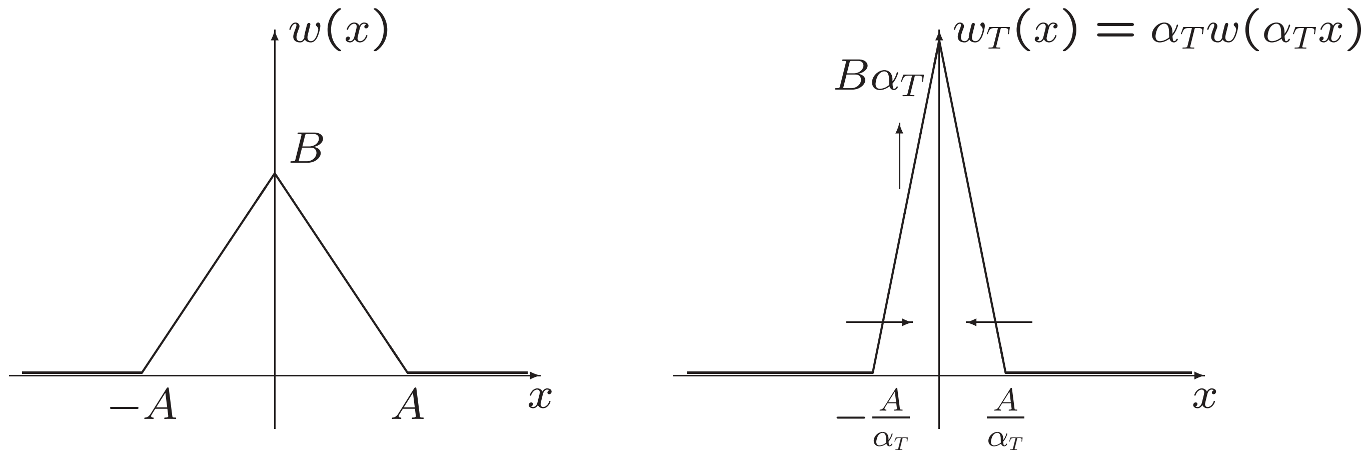
(ii) $\int_{-\infty}^{\infty} w(x) dx = 1$

(iii) $\sup_{x \in \mathbb{R}} |w(x)| < \infty$

(iv) $\lim_{x \to \infty} w(x) = 0$

Let α_T be scale factors satisfying $\alpha_T \to \infty$ and $\alpha_T / T \to 0$ as $T \to \infty$. Then we put $w_T(x) = \alpha_T w(\alpha_T x)$, $x \in \mathbb{R}$.

If the "mother" function $w(\cdot)$ has bounded support, then the picture is like this



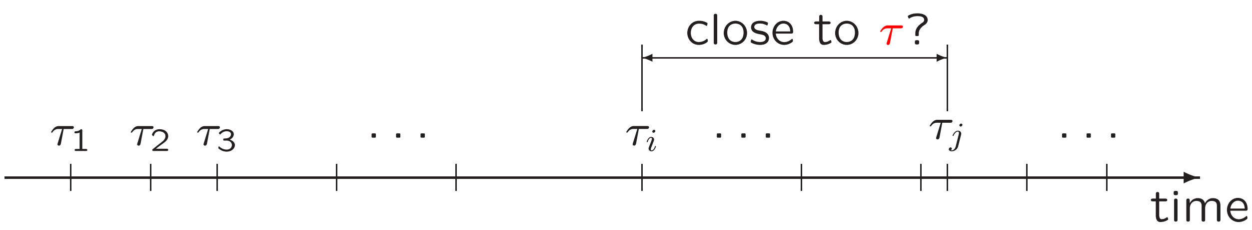
This example shows that $w_T(\cdot)$ is supposed to tend to the Dirac delta at the origin

Why do we need this averaging kernel? A typical correlation estimate is as follows:

$$\hat{R}_T(\tau) = \frac{1}{T}\sum_{k=1}^{T} X(t_k)X(t_k+\tau), \qquad \tau \in \mathbb{R} \tag{11}$$

When irregular sampling is involved, we cannot assure that we would be able to find sampling points τ_i and τ_j satisfying $|\tau_j - \tau_i| = \tau$, even approximately

This is why the weight function w_T is introduced in order to give preference to the pairs τ_i, τ_j whose difference $|\tau_j - \tau_i|$ is equal or close to the desired lag τ in (11)



Rates of convergence

$$\text{Covariance of } \widehat{K}_T \sim \frac{\alpha_T}{T}, \quad T \to \infty$$

$$\text{CLT for } \sqrt{\frac{T}{\alpha_T}} \cdot [\widehat{K}_T(t) - K(t)]$$

Optimal quadratic-mean rate of convergence: $O(T^{-4/5})$ with $\alpha_T = O(T^{-1/5})$

Rates for the classical estimate:

$$\text{Covariance of } \widehat{K}_T \sim \frac{1}{T}, \quad T \to \infty$$

$$\text{CLT for } \sqrt{T}[\widehat{K}_T(t) - K(t)]$$

Therefore this framework for random sampling gives lower rates of convergence than the classical ones, remaining these rates still faster than those in Section "Record of the sampling instants unavailable"

On the other hand, when periodic sampling is used with the sampling rate β, consistent estimates of $K(t)$ can be obtained only(!) at the points $\tau_k = k/\beta$ on the basis of the observations $\{X(k/\beta)\}_{k=1}^{n}$. Hence fast sampling rates β will be needed for any reasonable interpolation of such estimates in contrast to estimates (10), using alias-free sampling schemes, where the average sampling rate $\beta > 0$ is completely arbitrary

4. The framework

Inputs

Let $\{X_\Delta(t), t \in \mathbb{R}\}$, $\Delta > 0$, be a family of separable stationary zero-mean Gaussian processes defined on a common probability space $(\Omega, \mathcal{F}, \mathsf{P})$. Assume that each process has spectral density f_Δ such that

$$f_\Delta(\lambda) = f(\lambda/\Delta), \quad \lambda \in \mathbb{R},$$

where the function f is as follows:

- f is symmetric; $f(0) = c/(2\pi)$, where $c > 0$ is a constant;

- $f^{*(-1)} \in L_1(\mathbb{R})$, where $f^{*(-1)}(t) \stackrel{\text{def}}{=} \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\lambda t} f(\lambda) d\lambda$.

Then $K_{XX}^{(\Delta)}(\tau) \stackrel{\text{def}}{=} K_{X_\Delta X_\Delta}(\tau) = f_\Delta^{*(-1)}(\tau) = \Delta \cdot f^{*(-1)}(\tau \cdot \Delta)$

X_Δ and Y_Δ are jointly stationary and jointly Gaussian having the spectral density

$$f_{XY}^{(\Delta)}(\lambda) = \mathcal{H}(\lambda) f_\Delta(\lambda), \qquad \lambda \in \mathbb{R} \tag{12}$$

Functional form of the estimate

Following Masry's idea, we use a similar estimate to the one he used, with auto-correlations replaced by cross-correlations. Rewriting the estimate to the integral form gives the following:

$$\hat{H}_{\Delta,T}(t) = \frac{2\pi}{\beta^2 cT} \int_0^T \int_0^T w_T(t-(t_1-t_2)) X_\Delta(t_1) Y_\Delta(t_2) N(dt_1) N(dt_2) \tag{13}$$

Recall that we have another asymptotic parameter Δ “responsable” for making the input closer to a white noise

Therefore, heuristically, we have to work with both of these asymptotics:

- $\Delta \to \infty$, in order for “$\widehat{H} \to H$”

- $T \to \infty$, in order to obtain the estimate at the desired point $\tau \in \mathbb{R}$

By the moment, we maintain these parameters independent in order to figure out the way we should make them tend to infinity.

[*Hint*: The above mentioned order of things suggests that we should first make $T \to \infty$ and then $\Delta \to \infty$]

Some asymptotic results

Lemma 1 *Suppose that* $H \in L_2(\mathbb{R})$*. Then* $\lim_{\Delta\to\infty} K_{XY}^{(\Delta)}(t) = \frac{c}{2\pi}H(t)$ *for any* $t \in S_{\text{con}}(H)$ *(the set of continuity of the function* H*).*

Lemma 2 *We have*

$$\mathsf{E}[\hat{H}_{\Delta,T}(t)] = \frac{2\pi}{c}\int_{-\infty}^{\infty} w_T(t-u)K_{XY}^{(\Delta)}(u)A_T(u)du, \qquad t \in \mathbb{R}, \qquad (14)$$

where

$$A_T(u) = \begin{cases} 1 - \dfrac{|u|}{T}, & \textit{if } |u| \le T, \\ 0, & \textit{if } |u| > T. \end{cases}$$

Put

$$\varphi_{\Delta,T}(t,s) = \frac{2\pi}{c}\int_{-\infty}^{\infty} w_T(t-u)A_T(u)f_{\Delta}^{*(-1)}(u-s)du, \qquad s,t \in \mathbb{R} \tag{15}$$

This family (w.r.t. Δ and T) is an approximate identity as $\Delta \to \infty$ and $T \to \infty$

Theorem 1 *If $t \in S_{con}(H)$, then $\widehat{H}_{\Delta,T}(t) \to H(t)$ as $\Delta \to \infty$ and $T \to \infty$.*

Remark: The bias of estimate (13) does not depend on the sampling rate β.

Theorem 2 *Under the assumption $H \in L_1(\mathbb{R})$ and given that the input processes are Gaussian, we have*

$$\text{Cov}\left\{\widehat{H}_{\Delta,T}(t), \widehat{H}_{\Delta,T}(s)\right\} \sim C \cdot \frac{\alpha_T \cdot \Delta}{T} \tag{16}$$

where C is a constant whose value can be calculated explicitly.

Remark: The rate $\frac{\alpha_T \cdot \Delta}{T}$ is of course slower than $\frac{\alpha_T}{T}$ since $\Delta \to \infty$, though Δ may tend to infinity at an arbitrarily slow rate.

On the other hand, the rate at which $\Delta \to \infty$ cannot be faster than that of $T/\alpha_T \to \infty$.

Theorem 3 *Under the hypothesis of Theorem 2 the standartized variates*

$$[T/(\alpha_T\Delta)]^{1/2}\{\hat{H}_{\Delta,T}(t_i)-\mathsf{E}[H(t_i)]\},\quad i=1,\ldots,m \tag{17}$$

are jointly asymptotically normal, as $\Delta\to\infty$ *and* $T\to\infty$*, with zero means and a covariance structure given in Theorem 2.*

Remark: The covariance structure of the limiting process is markedly richer than that observed under the regular discrete sampling.

Another interesting feature is the form of the limit covariance function for the standartized variates:

$$C_\infty(t_1,t_2) = \int_{-\infty}^{\infty}\Big[H(u+\frac{t_1-t_2}{2})H(u+\frac{t_2-t_1}{2})+ +H(u-\frac{t_1+t_2}{2})H(-u-\frac{t_1+t_2}{2})\Big]\,du$$

It is clear that the second term equals zero if the system is causal, i.e. if $H(s)=0$ for $s<0$. [If a system is known to be causal in advance, there is no sense in considering negative t_1 and t_2 !]

In this case, the remaining first term depends on the difference t_1-t_2 only, meaning that the process whose correlation function is C_∞ will be stationary, in contrast to the non-causal case

In the estimation of the correlation function, both terms are present in the limit correlation function C_∞, always making the limit stochastic process to be non-stationary

In the framework of the response function estimation, C_∞ generates a stationary or non-stationary limit process depending on whether the original system is causal or not