

# INSA – Gaussian processes

## Continuity and differentiability of sample functions

---

Andrés F. López-Lopera

The French Aerospace Lab ONERA, France  
Information Processing and Systems Department (DTIS)  
Multidisciplinary Methods, Integrated Concepts (M2CI) Research Unit

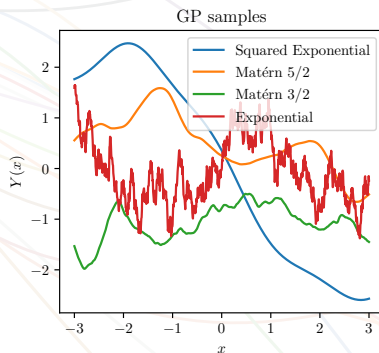
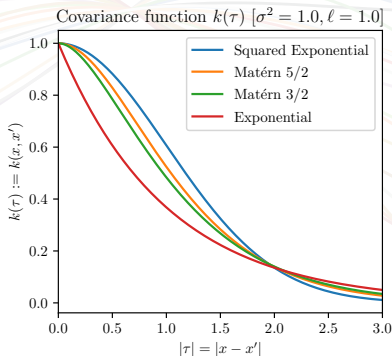
The background of the slide is a light gray with a complex pattern of thin, overlapping, wavy lines in various colors including purple, blue, green, yellow, and red. These lines create a sense of movement and depth.

# **Introduction**

---

# Introduction

- Regularity assumptions can be encoded in kernel functions:
  - periodicity, **smoothness**, stationarity, isotropy, ...



(left) Matérn covariance functions, (right) GP samples [Rasmussen and Williams, 2005]

- Let  $Y_t$  be a stochastic process (e.g. a GP):  $\{Y(t), t \in \mathbb{R}\}$
- In the following, we assume that  $Y_t$  is centred, i.e.  $\mathbb{E}\{Y_t\} = 0$ 
  - Otherwise, we consider  $Z_t = Y_t - \mathbb{E}\{Y_t\}$  instead
- We denote  $k(t, t') = \text{cov}\{Y_t, Y_{t'}\}$  the covariance function
  - We assume stationary kernels  $k(\tau) := k(t, t + \tau)$  (abuse of notation)

## Definition (Finite 2nd order moments)

A stochastic process  $Y_t$  has finite 2nd order moments if for all  $t \in \mathbb{R}$ :

$$\mathbb{E}\{|Y_t|^2\} < \infty.$$

This implies that  $\mathbb{E}\{Y_t\}$  and  $\mathbb{E}\{Y_t Y_{t'}\}$  are well defined for all  $t, t' \in \mathbb{R}$ .

- We denote  $L^2$  the set of r.v's with finite 2nd order moments
- In the following,  $L^2$  is a *Hilbert space* (3rd lecture), with inner product:

$$\langle X, Y \rangle = \mathbb{E}\{XY\}.$$

- We can also denote  $k(t, t') = \langle Y_t, Y_{t'} \rangle$  since  $Y_t$  is a centred

1. Sample functions properties in quadratic mean
2. Sample function properties

The background of the slide is white and features a series of overlapping, wavy lines in various colors including light blue, green, yellow, orange, and red. These lines create a sense of movement and depth. The text is centered on the left side of the slide.

## **Sample functions properties in quadratic mean**

---

- There is no simple relationship between the covariance function  $k$  of a stochastic process  $Y$  and the **smoothness** of its realizations
- However, one can relate  $k$  to quadratic mean properties of  $Y$ :
  - Convergence
  - Continuity
  - Differentiability



## Definition (Convergence in quadratic mean)

Let  $\{X_n\}_{n=1}^{\infty}$  be a random sequence, with r.v.'s  $X_1, X_2, \dots$  defined on the same probability space as a r.v.  $Y$ . Define  $\mathbb{E} \{|X_n|^2\} < \infty$  (finite variances).  $\{X_n\}$  converges in quadratic mean (q.m.),  $X_n \xrightarrow{q.m.} X$ , if there exists  $Y$  such that:

$$\|X_n - Y\|_{L^2} \rightarrow 0 \quad \left( \text{i.e. } \mathbb{E} \{|X_n - Y|^2\} \rightarrow 0 \right)$$

## Theorem (Loeve criterion)

$\{X_n\}$  converge in q.m. iff  $\mathbb{E} \{X_n X_m\} = \langle X_n, X_m \rangle$  converges to a finite limit  $c$  when  $n, m \rightarrow \infty$  (independently).

## Proof.

- The “if” part follows from

$$\begin{aligned}\mathbb{E}\{|X_n - X_m|^2\} &= \mathbb{E}\{[X_n - X_m][X_n - X_m]\} \\ &= \mathbb{E}\{X_n X_n\} + \mathbb{E}\{X_m X_m\} - 2\mathbb{E}\{X_n X_m\} \\ &= c + c - 2c = 0\end{aligned}$$

- The “only if” part follows from

$$\mathbb{E}\{X_n X_m\} \rightarrow \mathbb{E}\{XX\} = \mathbb{E}\{|X|^2\}.$$

□

## Definition (Continuity in quadratic mean)

A stochastic process  $Y_t$  is said to be continuous in q.m. at  $t = t_0$  if

$$Y_t \xrightarrow{q.m.} Y_{t_0}.$$

## Proposition

1.  $Y_t$  is continuous in q.m. at  $t = t_0$  iff  $k(u, v)$  is continuous at  $(t_0, t_0)$
2. If  $k(u, v)$  is continuous at every diagonal point  $(t, t)$ , then  $Y_t$  is continuous everywhere.

## Proof hints.

1.
  - For the “if” part, compute the expression  $\mathbb{E} \{ (Y_{t+\tau} - Y_t)^2 \}$
  - For the “iff” part, use the equality

$$\begin{aligned} k(t + \tau, t + \nu) - k(t, t) &= \langle Y_{t+\tau} - Y_t, Y_{t+\nu} - Y_t \rangle \\ &\quad + \langle Y_{t+\tau} - Y_t, Y_t \rangle + \langle Y_{t+\nu} - Y_t, Y_t \rangle \end{aligned}$$

2. Use (1) and continuity of  $\langle \cdot, \cdot \rangle$

## Definition (Differentiability in quadratic mean)

$Y_t$  is differentiable in q.m. at  $t$  if  $\frac{Y_{t+h}-Y_t}{h}$  converge in q.m.

## Proposition

1. If  $\frac{\partial^2 k}{\partial u \partial v}$  exists at  $(t, t)$ , then  $Y_t$  is differentiable in q.m. at  $t$ .
2. If  $\frac{\partial^2 k}{\partial u \partial v}$  exists for every  $(t, t)$ , then  $\frac{\partial k}{\partial u}(u, v)$  and  $\frac{\partial^2 k}{\partial u \partial v}(u, v)$  exist everywhere and we have:

$$\text{cov} \{Y'_u, Y_v\} = \frac{\partial k}{\partial u}(u, v) \quad \text{and} \quad \text{cov} \{Y'_u, Y'_v\} = \frac{\partial^2 k}{\partial u \partial v}(u, v)$$

## Proof hints.

1. Apply Loeve criterion to  $Z_n = \frac{Y_{t+h_n}-Y_t}{h_n}$  for any sequence  $h_n \rightarrow 0$
2. For the 1st derivative, use (1) and compute  $\langle \frac{Y_{u+h}-Y_u}{h}, Y_v \rangle$ . Then, develop  $\langle \frac{Y_{u+h}-Y_u}{h}, \frac{Y_{v+h}-Y_v}{h} \rangle$

## Exercise.

1. Show that if  $\frac{\partial^4 k}{\partial^2 u \partial^2 v}$  exist at  $(t, t)$ , then  $Y_t$  is twice diff. in q.m. at  $t$ .
2. In addition, if  $\frac{\partial^4 k}{\partial^2 u \partial^2 v}$  exists at every  $(t, t)$ , then all the derivatives written below exist everywhere and we have:

$$\text{cov} \{Y_u'', Y_v\} = \frac{\partial^2}{\partial u^2} k(u, v)$$

$$\text{cov} \{Y_u'', Y_v'\} = \frac{\partial^3}{\partial u^2 \partial v} k(u, v)$$

$$\text{cov} \{Y_u'', Y_v''\} = \frac{\partial^4}{\partial u^2 \partial v^2} k(u, v)$$

### Definition (2nd order stationary processes)

$Y_t$  is 2nd order stationary if for any  $t, \tau$ ,  $\mathbb{E}\{Y_t\}$  and  $\text{cov}\{Y_t, Y_{t+\tau}\}$  do not depend on  $t$ .

- If  $Y_t$  is a centred process, then  $Y_t$  is stationary if  $k(t, t')$  is a function of  $t - t'$  (see also the definition from 1st lecture).

### Proposition (Continuity and differentiability)

*Let  $Y_t$  be a stationary stochastic process.*

- 1.  $Y_t$  is continuous in q.m. at  $t = t_0$  iff  $k(\tau)$  is continuous at 0. In this case,  $Y_t$  is continuous everywhere.*
- 2. If  $k^{2p}(\tau)$  exists in an open set containing 0, then  $Y_t$  is differentiable in q.m. at order  $p$  everywhere.*

**Proof hint.** Show that the local properties of  $k(\tau)$  at 0 imply the same properties to  $k(u, v)$  at the diagonal points.

## Challenges

- Continuity or differentiability in q.m. do not necessarily imply sample function continuity or differentiability.
- However, they can be easily related to stationary covariance functions



## Definition (Equivalence)

We say that  $Y_t$  and  $Z_t$  are equivalent if they have the same finite-dimensional distributions for all  $t \in \mathbb{R}$ :

$$P(\{Y_t = Z_t\}) = 1$$

## Remarks.

- This implies that two equivalent processes have the same family of finite-dimensional distributions
- Two equivalent processes do NOT have necessarily the same sample functions properties

## Example.

- Let  $Y_t$  and  $Z_t$  two stochastic processes defined over  $[0, 1]$  by:

$$Y(t) = 0 \quad \forall t$$

$$Z(t) = \begin{cases} 1, & \text{if } t = \tau \text{ for } \tau \in [0, 1] \\ 0, & \text{otherwise} \end{cases}$$

- Then  $Y_t$  and  $Z_t$  have the same finite-dimensional distributions but

$$P(\{Y_t \text{ is continuous in } [0, 1]\}) = 1$$

$$P(\{Z_t \text{ is continuous in } [0, 1]\}) = 0$$

The background of the slide is a light gray with a complex pattern of thin, overlapping, wavy lines in various colors including purple, blue, green, yellow, and orange. These lines create a sense of movement and depth.

## **Sample function properties**

---

## Theorem (Sample function continuity– Kolmogorov's theorem)

Let  $Y_t$  be a stochastic process defined over  $[0, 1]$ . Suppose that, for all  $t, t + h \in [0, 1]$ ,

$$P(\{|Y_{t+h} - Y_t| \geq g(h)\}) \leq q(h),$$

where  $g$  and  $q$  are even functions of  $h$ , non increasing as  $h \downarrow 0$ , and such that

$$\sum_{n=1}^{\infty} g(2^{-n}) < \infty \quad \text{and} \quad \sum_{n=1}^{\infty} 2^n q(2^{-n}) < \infty.$$

Then, there exists an equivalent stochastic process  $Z_t$  whose sample functions are, with probability one, continuous in  $[0, 1]$ .

**Proof.** See [Cramér and Leadbetter, 1967]

## Corollary

*If with the notation above we have*

$$\mathbb{E} \{|Y_{t+h} - Y_t|^p\} \leq c \frac{|h|}{|\log|h||^{1+r}},$$

*where  $p < r$  and  $c$  are positive constants, the conclusion of the theorem holds.*

## Proof.

· Consider  $g(h) := |\log|h||^{-b}$  with  $1 < b < r/p$  and the Markov inequality:  
 $P(|X| \geq a) \leq \mathbb{E} \{|X|^p\} / a^p$ .

· By applying the Kolmogorov's theorem, we have

$$P(\{|Y_{t+h} - Y_t| \geq g(h)\}) \leq c \frac{|h|}{|\log|h||^{1+r-bp}} = q(h)$$

· Since  $b > 1$ , then

$$\sum_{n=1}^{\infty} g(2^{-n}) = \sum_{n=1}^{\infty} \frac{1}{|\log(2^{-n})|^b} = \frac{1}{(n \log 2)^b} < \infty$$

· Since  $1 + r - bp > 1$ , then

$$\sum_{n=1}^{\infty} 2^n q(2^{-n}) = \sum_{n=1}^{\infty} \frac{c}{|\log(2^{-n})|^{1+r-bp}} = \sum_{n=1}^{\infty} \frac{c}{[n \log(2)]^{1+r-bp}} < \infty$$

□

## Theorem (Stochastic processes with finite 2nd order moments)

Let  $Y_t$  be a stochastic process defined with finite second moments. If for all  $t, t+h \in [a, b]$  the difference

$$\Delta_h^2 k(t, t) := k(t+h, t+h) - k(t+h, t) - k(t, t+h) + k(t, t)$$

satisfies the inequality  $\Delta_h^2 k(t, t) < c \frac{|h|^q}{|\log|h||^q}$ , with  $q > 3$  and  $c > 0$ , then  $Y_t$  is equivalent to a stochastic process which, with probability one, is sample continuous.

## Theorem (Stationary processes)

Let  $Y_t$  be a stationary stochastic process. If  $k''(0)$  exists, then  $Y_t$  is equivalent to a stochastic process which, with probability one, is sample continuous, i.e.  $Y_t \in \mathcal{C}$ .

**Proof hint.** Apply Corollary with  $p = 2$ .

## Theorem (Sample function differentiability)

Let  $Y_t$  be a stochastic process defined over  $[0, 1]$ . Suppose that the hypothesis of Kolmogorov's theorem hold, and that, for all  $t - h, t, t + h \in [0, 1]$ ,

$$P(\{|Y_{t+h} + Y_{t-h} - 2Y_t| \geq g_1(h)\}) \leq q_1(h),$$

where  $g_1$  and  $q_1$  are even functions of  $h$ , non increasing as  $h \downarrow 0$ , and such that

$$\sum_{n=1}^{\infty} 2^n g_1(2^{-n}) < \infty \quad \text{and} \quad \sum_{n=1}^{\infty} 2^n q_1(2^{-n}) < \infty.$$

Then,  $Y_t$  is equivalent to a process which, with probability one, has continuous sample function derivatives in  $[0, 1]$ .

**Proof.** See [Cramér and Leadbetter, 1967].



## Corollary

*If the conditions of the corollary of the Kolmogorov's theorem are satisfied, and if*

$$\mathbb{E} \{|Y_{t+h} + Y_{t-h} - 2Y_t|^p\} \leq c \frac{|h|^{1+p}}{|\log|h||^{1+r}},$$

*where  $p < r$  and  $c$  are positive constants, the conclusion of the theorem holds.*

**Proof hint.** Apply the Markov inequality.

## Theorem (Stochastic processes with finite 2nd order moments)

*Let  $Y_t$  be a stochastic process defined with finite second moments. If for all  $t, t + h$ , the 4th difference  $\Delta_h^4 k(t, t)$  satisfies the inequality  $\Delta_h^4 k(t, t) < c \frac{|h|^3}{|\log|h||^q}$ , with  $q > 3$  and  $c > 0$ , then  $Y_t$  is equivalent to a stochastic process which, with probability one, has continuous sample function derivatives.*

## Theorem (Stationary processes)

*Let  $Y_t$  be a stationary stochastic process. If  $k^{(4)}(0)$  exists, then  $Y_t$  is equivalent to a stochastic process which, with probability one, has  $C^1$  sample functions.*

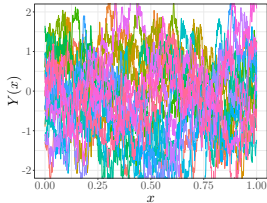
**Proof hint.** Apply Corollary with  $p = 2$ .

## Theorem (Differentiability in high orders)

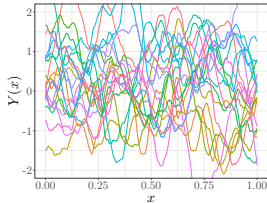
*There are analogous results. In particular, if  $Y_t$  is a stationary stochastic process and if  $k^{(2k+2)}(0)$  exists, then  $Y_t$  is equivalent to a process which, with probability one, has  $C^k$  sample functions.*

# Effects of 1D stationary kernels in GP samples

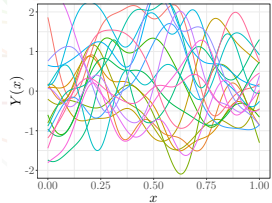
1D stationary kernel	$k_{\sigma^2, \ell}(\tau)$	Class
Exponential	$\sigma^2 \exp \left\{ -\frac{ \tau }{\ell} \right\}$	$\mathcal{C}$
Matérn 3/2	$\sigma^2 \left( 1 + \sqrt{3} \frac{ \tau }{\ell} \right) \exp \left\{ -\sqrt{3} \frac{ \tau }{\ell} \right\}$	$\mathcal{C}^1$
Matérn 5/2	$\sigma^2 \left( 1 + \sqrt{5} \frac{ \tau }{\ell} + \frac{5}{3} \frac{\tau^2}{\ell^2} \right) \exp \left\{ -\sqrt{5} \frac{ \tau }{\ell} \right\}$	$\mathcal{C}^2$
Squared Exponential (SE)	$\sigma^2 \exp \left\{ -\frac{1}{2} \frac{\tau^2}{\ell^2} \right\}$	$\mathcal{C}^\infty$



(a) Exponential



(b) Matérn 3/2



(c) Squared Exponential

■ ■ ■ ■ GP samples

The background of the slide is a light gray with a complex pattern of thin, overlapping, wavy lines in various colors including blue, green, yellow, orange, and red. These lines create a sense of movement and depth.

## Conclusions

---

- Continuity and differentiability in quadratic mean have been studied
  - They do not imply sample function continuity or differentiability
  - They can be related to stationary covariance functions
- Sample function continuity/differentiability can be shown but at the cost of technicality

- Harald Cramér and M. Ross Leadbetter. *Stationary and Related Stochastic Processes - Sample Function Properties and Their Applications*. Wiley, 1967.
- Marc G. Genton. Classes of kernels for machine learning: A statistics perspective. *Journal of Machine Learning Research*, 2001.
- Carl E. Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. MIT Press, 2005.
- Michael L. Stein. *Interpolation of Spatial Data: Some Theory for Kriging*. Springer, 1999.
- A. M. Yaglom. *Correlation Theory of Stationary and Related Random Functions*. Springer, 1987.