Practical session
July 3rd, 2018

Gaussian Processes metamodeling, practical session

Exercice 1: Simulation of the trajectories of a Gaussian process whose mean and covariance functions are prescribed

1.a) Let $(Z_t)_{t\in\mathbb{R}}$ be a stationary centered Gaussian process with covariance $C(t,t') = \sigma^2 R(t-t')$. Consider a grid of $[0,1]$ composed with $N = 200$ points

- Gaussian covariance: simulate some realizations of $(Z_t)_{t\in\mathbb{R}}$ on the grid with $\sigma^2 = 0.5$ and $R(h) = e^{-\left(\frac{h}{\theta}\right)^2}$. Take $\theta = 0.05$, then $\theta = 0.2$.
- Exponential covriance: simulate some realizations of $(Z_t)_{t\in\mathbb{R}}$ on the grid with $\sigma^2 = 0.5$ and $R(h) = e^{-\frac{|h|}{\theta}}$. Take $\theta = 0.05$, then $\theta = 0.2$.
- Gaussian covariance with a nugget effect: simulate some realizations of $(Z_t)_{t\in\mathbb{R}}$ on the grid with $\sigma^2 = 0.5$ and $R(h) = e^{-\left(\frac{h}{\theta}\right)^2} + \lambda\delta_0(h)$. Take $\lambda = 0.2$ and $\theta = 0.05$, then $\theta = 0.2$.

Notation: $\delta_O(h) = 0$ for $h \neq 0$, $\delta_0(0) = 1$.

1.b) Let $(Z(x))_{x\in\mathbb{R}^2}$ be a centered stationary Gaussian process with $C(x,x') = \sigma^2 R(x-x')$. Consider a grid of $[0,1]\times[0,1]$ composed with $N = 50\times 50$ points.

- Isotropic Gaussian covariance: simulate some realizations of $(Z(x))_{x\in\mathbb{R}^2}$ on the grid with $\sigma^2 = 0.5$ and $R(h) = e^{-\|\frac{h}{\theta}\|^2}$. Take $\theta = 0.1$.
- Anisotropic Gaussian covariance: simulate some realizations of $(Z(x))_{x\in\mathbb{R}^2}$ on the grid with $\sigma^2 = 0.5$ and $R(h) = e^{-\sum_{i=1}^{2}\left(\frac{h_i}{\theta_i}\right)^2}$. Take $\theta_1 = 0.1$ and $\theta_2 = 0.03$.
- Exponential anisotropic covariance: simulate some realizations of $(Z(x))_{x\in\mathbb{R}^2}$ on the grid with $\sigma^2 = 0.5$ and $R(h) = e^{-\sum_{i=1}^{2}\frac{|h_i|}{\theta_i}}$. Take $\theta_1 = 0.1$ and $\theta_2 = 0.03$.

Exercise 2: Construction of GP metamodel with a learning sample in dimension 1

We consider the analytical 1D function on $[0,1]$:

$$f(x) = \sin(30(x-0.9)^4)cos(2(x-0.9)) + \frac{x-0.9}{2}$$

Apply the following methodology:

Step 0: Representation of the function on $[0,1]$

- Evaluate f on 100 equidistributed points and plot it.

Step 1: Construction of a learning sampleand evaluation of f on this learning sample

- Vary the size of the leraning sample from $N = 10$ to 30 points.
- Vary the nature of the design: equidistributed points on $[0,1]$, random uniform sample on $[0,1]$.

Step 2: Estimation of the parameters of the GP metamodel
Let $(Z(x))_{x\in\mathbb{R}^2}$ be a Gaussian process with constant mean and Matern $5/2$ stationary covariance. That means

$$R(h) = \left(1 + \frac{\sqrt{5}h}{\theta} + \frac{5h^2}{3\theta^2}\right) \exp\left(-\frac{\sqrt{5}h}{\theta}\right).$$

Construct the metamodel based on Z conditionally to the learning sample:

- estimate the parameters of the metamodel;
- compute and plot the metamodel on the test sample;
- compute and plot the MSE on the test sample;
- compute Q^2 on the test sample.

Step 3: adaptive planification

- find the point x where the MSE is maximal and add it to the learning sample;
- update the metamodel (estimation of the hyperparameters, construction of the predictor and computation of the MSE);
- Plot the predictor and MSE as in **Step 2**.

Exercise 3: Construction of GP metamodel with a learning sample in dimension 2
We consider the schwefel2D analytical function on $[-200, 200]^2$:

$$f(x_1, x_2) = -x_1 \sin(\sqrt{[x_1|}) - x_2 \sin(\sqrt{|x_2|}).$$

Apply thr following methodology:
Step 0: Representation of the function on $[-200, 200]^2$

- evaluate f on an equidistributed grid with 70×70 points and plot f.

Step 1: Construction of a design of experiments (learning sample) and evaluation of f on this design

- vary the size of the learning sample from $N = 70$ to 100 points;
- vary the type of designs: uniform random design, LHS.

Step 2: Estimation of the parameters of the GP metamodel
Let $(Z(x)_{x\in\mathbb{R}^2})$ be a centered stationary Gaussian process with covariance Matern $3/2$.
Construct the metamodel based on Z conditionally to the points of the learning sample.

- compute and plot the predictor on the test sample;
- plot the error (absolute value) on the test sample;
- compute Q^2 on the test sample.

Exercise 4: Use of a GP metamodel for sensitivity analysis
Consider the Ishigami analytical function:

$$f(X_1, X_2, X_3) = \sin(X_1) + 7\sin(X_2)^2 + 0.1X_3^4 \sin(X_1)$$

with the X_i uniformly distributed on $[-\pi, \pi]$, for $i = 1, 2, 3$.
Step 0: Compute analytically the first-order Sobol' indices

- Show that $S_1 = 0.3139$, $S_2 = 0.4424$, $S_3 = 0$.

Step1: Construction of a DoE and evaluation of f on that design

- construct a LHS of $N = 70$ to 100 points.

Step2: Construction of the GP metamodel and control of its quality in terms of prediction
Let $(Z(x))_{x\in\mathbb{R}^2}$ be a Gaussian process with constant mean and Matern 5/2 stationary covariance. Construct the GP metamodel based on Z conditionally to the points in the learning sample.

- compute Q^2 on a test sample;
- compute Q^2 by cross validation on the learning sample.

Step 3: Estimate the Sobol' indices