Gaussian Process Regression – TD 1

Mines Saint-Étienne, Data Science, 2016 - 2017

Exercise 1

Let Z_s be a Gaussian process with mean $\mu(x) = x^2$ and kernel

$$k_s(x,y) = \frac{k(x,y) + k(-x,y) + k(x,-y) + k(-x,-y)}{4}$$

where k is a symmetric positive semi-definite function. We want to study the properties of the samples from Z_s.

1. Compute $\mathrm{E}[Z_s(x) - Z_s(-x)]$.
2. Compute $\mathrm{var}[Z_s(x) - Z_s(-x)]$.
3. What can you conclude?
4. If you want to approximate a symmetric function f given some observations $f(X) = F$, is there a difference between:
   - add extra observations $f(-X) = F$ to take the symmetry into account
   - use the kernel k_s to take the symmetry into account

Exercise 2: (2015/2016 exam)

ANOVA kernels are kernels over $\mathbb{R}^d \times \mathbb{R}^d$ of the form : $k(x,y) = \prod_{i=1}^{d} \left(1 + k_i(x_i, y_i)\right)$, where the k_i are symmetric positive semi-definite functions.

1. Using the results from the course, show that ANOVA kernels are valid covariance functions.

We now consider costly-to-evaluate function $f : [0,1]^{10} \to \mathbb{R}$, a design of experiment X based on 100 points and the set of observations F. The knowledge we have about f is that it is a smooth function that is infinitely differentiable.

2. With such settings, which kernel would you choose and what kind of Gaussian process regression model would you consider (simple Kriging, ordinary Kriging, Universal Kriging).
3. Give the expressions of the mean predictor and of the 95% confidence intervals.
4. Show that the mean predictor can be interpreted as a sum of 2^d functions with increasing interaction order.

5. Each term of this decomposition can be interpreted as a Gaussian process conditional distribution. Detail which one and deduce some confidence intervals associated to each sub-model.

6. According to an expert, the mean value of f is 6 and the interactions of order higher than 2 can be neglected. What changes can you make in the model and in the kernel expression in order to account for these informations ?

7. We now consider a particular type for the univariate kernels k_i such that $\int_0^1 k_i(s,x)\,ds = 0$ for all $x \in [0,1]$. Can you recover some of the properties of Polynomial Chaos models?

bonus Detail how to obtain a kernel k_i such that $\int_0^1 k_i(s,x)\,ds = 0$ using the conditional distribution of a Gaussian process given it has zero integral.