Computer experiments, exercises

INSA Toulouse 2021 - O. Roustant

1 Kriging equations (Gaussian process approach)

We recall the formula for Gaussian conditioning: if (Z_1, Z_2) is a Gaussian vector $\mathcal{N}(\mu, \Sigma)$ with $\Sigma = \begin{pmatrix} \Sigma_{1,1} & \Sigma_{1,2} \\ \Sigma_{2,1} & \Sigma_{2,2} \end{pmatrix}$ then $Z_2 | Z_1 = z_1$ is a Gaussian vector with:

$$\begin{aligned} \mathbb{E}[Z_2|Z_1 = z_1] &= \mu_2 + \Sigma_{2,1}\Sigma_{1,1}^{-1}(z_1 - \mu_1) \quad (1) \\ \mathrm{Cov}[Z_2|Z_1 = z_1] &= \Sigma_{2,2} - \Sigma_{2,1}\Sigma_{1,1}^{-1}\Sigma_{1,2} \quad (2) \end{aligned}$$

Let $(Y(x))_{x \in T}$ be a Gaussian process on some abstract set T, with mean μ and kernel k. Denote:

- $x^{(1)}, \dots, x^{(n)}$: vectors of length d corresponding to design points.
- $y = (y_1, \dots, y_n)^\top$: column vector of observations at design points.
- $X = \{x^{(1)}, \dots, x^{(n)}\}$: set of design points (not necessarily a matrix here)
- $k(X, X)$: matrix of size $n \times n$ whose element (i, j) is equal to $k(x^{(i)}, x^{(j)})$.
- $k(X, x)$: column vector of length n whose element i is equal to $k(x^{(i)}, x)$.
- $k(x, X)$: row vector, equal to the transpose of $k(X, x)$.
- $\mathbb{1}$: the column vector of ones (length n)

1.1 Noise-free observations and interpolation

1. Prove that the conditional process $\left(Y(x) | \{Y(x^{(1)}) = y_1, \dots, Y(x^{(n)}) = y_n\}\right)_{x \in T}$ is Gaussian, with mean and covariance function given by:

$$\begin{aligned} m_k(x) &= \mu + k(x, X)k(X, X)^{-1}(y - \mu\mathbb{1}) \\ c_k(x, x') &= k(x, x') - k(x, X)k(X, X)^{-1}k(X, x'). \end{aligned}$$

2. What can you say if x is a design point?
3. Some properties directly inherit from Gaussian vectors: which ones?

1.2 Noisy observations and filtering

Let us know consider the situation where $Y(x^{(i)})$ is observed with noise, i.e. $Y(x^{(i)}) + \epsilon_i$ is really observed. We assume that $\epsilon_1, \ldots, \epsilon_n$ are independent $\mathcal{N}(0, \tau_i^2)$, and independent of Y.

1. By adapting the previous exercise, explain why the process Y conditional on $Y(x^i)+\epsilon_i = y_i, i = 1, \ldots, n$ is Gaussian, and give its mean and covariance function.

2. What's happening at the design points?

3. Explain the differences and similarities between noisy and noise-free observations.

2 Kernels and models

1. Show that the following operations on (covariance) kernels also give valid kernels:

   - Multiplication by a positive scalar: λk
   - Sum or tensor sum: $k_1 + k_2$, $k_1 \oplus k_2$
   - Product or tensor product: $k_1 \times k_2$, $k_1 \otimes k_2$
   - Mapping the inputs with a function f (warping): $k(f(x), f(x'))$

2. Conversely, build a covariance kernel suited to model functions with the following properties:

   - 2D additives functions
   - 1D even functions, odd functions, functions symmetric with respect to $x = 1/2$
   - 1D periodic functions of period 2π

   Explain why, intuitively, the sample paths of the conditional GP (see Ex. 1) should share the same properties. Check your answers by simulating GP sample paths with these covariance functions.

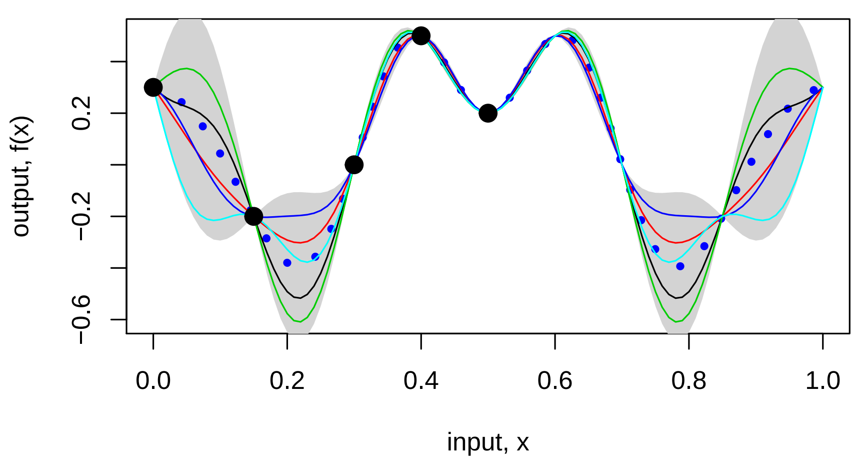

Figure 1: Gaussian process with a symmetric kernel (from R package `DiceKriging`).