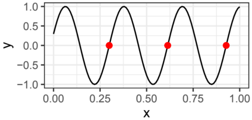
Introduction to Machine Learning

Covariance Functions for GPs - Advanced



Learning goals

- XXX
- XXX

MS-CONTINUITY AND DIFFERENTIABILITY

We wish to describe a Gaussian process in terms of its smoothness. There are several notions of continuity for random variables - one is continuity / differentiability in mean square (MS):

Definition

A Gaussian process $f(\mathbf{x})$ is said to be

- continuous in MS in $\mathbf{x}_*$, if $\mathbb{E}[|f(\mathbf{x}^{(k)}) - f(\mathbf{x}_*)|^2] \overset{k\to\infty}{\longrightarrow} 0$ for any sequence $\mathbf{x}^{(k)} \overset{k\to\infty}{\to} \mathbf{x}_*$
- MS differentiable in direction i if $\lim_{h\to 0} \mathbb{E}[|\frac{f(\mathbf{x}+h\boldsymbol{e}_i)-f(\mathbf{x})}{h}|]$ exists, where $\boldsymbol{e}_i = (0,\dots,0,1,0,\dots,0)^T$ is the unit vector in the i-th axis.

Remark: MS continuity / differentiability does not necessarily imply continuity / differentiability of the sampled function!

MS-CONTINUITY AND DIFFERENTIABILITY

MS continuity / differentiability of a Gaussian process can be derived from the smoothness properties of the kernel:

- The GP is continuous in MS if and only if the covariance function $k(\mathbf{x}, \mathbf{x}')$ is continuous
- The MS derivative of a Gaussian process exists iff the second derivative $\frac{\partial^2 k(\mathbf{x},\mathbf{x}')}{\partial \mathbf{x} \partial \mathbf{x}'}$ exists

SQUARED EXPONENTIAL COVARIANCE FUNCTION

One common used covariance function is the squared exponential covariance function:

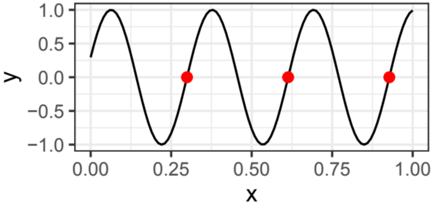
$$k(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{2\ell^2}\right)$$

Properties:

- as it depends on the distance $r = \|\mathbf{x} - \mathbf{x}'\|$ only, it is a isotropic (and thus also stationary) covariance function
- infinitely differentiable $\rightarrow$ corresponding GP is thus very smooth
- due to its strong smoothness assumptions it is often unrealistic for modeling many physical processes

UPCROSSING RATE AND CHARACTERISTIC LENGTH-SCALE

Another way to describe a Gaussian process is the expected number of up-crossings at level 0 on the unit interval, which we denote by N_0.



For an isotropic covariance function $k(r)$, it can be shown that the expected number of up-crossings can be calculated explicitly

$$\mathbb{E}[N_0] = \frac{1}{2\pi}\sqrt{\frac{-k''(0)}{k(0)}}.$$

UPCROSSING RATE AND CHARACTERISTIC LENGTH-SCALE

Example: Squared exponential

$$
\begin{aligned}
k(r) &= \exp\left(-\frac{r^2}{2\ell^2}\right) \\
k'(r) &= -k(r) \cdot \frac{r}{\ell^2} \\
k''(r) &= k(r) \cdot \frac{r^2}{\ell^4} - k(r) \cdot \frac{1}{\ell^2}
\end{aligned}
$$

The expected number of level-0 upcrossing is thus

$$
\mathbb{E}[N_0] = \frac{1}{2\pi}\sqrt{\frac{-k''(0)}{k(0)}} = \frac{1}{2\pi}\sqrt{\frac{1}{\ell^2}} = (2\pi\ell)^{-1}
$$