Exponential Covariance Functions

Other covariance functions produce functions that aren’t linear. One useful class of valid covariances functions for a univariate $x$ has the form

$$\text{Cov}(y_{i_1}, y_{i_2}) = \eta^2 \exp\left(-\left(\frac{|x_{i_1} - x_{i_2}|}{\rho}\right)^R\right)$$

where $0 < R \leq 2$. Usually, we let the mean function be constant (eg, zero). The process is then stationary, with variance $\eta^2$, and a length scale determined by $\rho$.

A prior using this covariance function says that nearby points are likely to have similar $y$ values, but the responses may be much different for far-away points.

When $R = 2$, the functions are infinitely differentiable; when $R < 2$, they are not differentiable. Here are 3 random functions from a GP with $R = 1.8$, $\eta = 1$, $\rho = 1$:
More Functions from Exponential Covariance Functions

These Gaussian processes use the exponential covariance function with $R = 2$ and $\eta = 1$, but one has $\rho = 1$, the other $\rho = 1/5$. Two random functions are shown for each process.

\[
\text{Cov}(y_{i_1}, y_{i_2}) = \exp(- (x_{i_1} - x_{i_2})^2) \quad \text{Cov}(y_{i_1}, y_{i_2}) = \exp(- ((x_{i_1} - x_{i_2}) / (1/5))^2)
\]
Combining Covariance Functions

If $C_1(x)$ and $C_2(x)$ are valid covariance functions, then $C_1(x) + C_2(x)$ and $C_1(x)C_2(x)$ are also valid covariance functions.

So we can build new covariance functions from old by taking sums and products. Some uses:

- The sum of a covariance function from a linear model and an exponential covariance function with small $\eta$ yields an “almost linear” model.

- Sums of covariance functions with different length scales yield functions with both “large scale” and “small scale” variation.

- Sums of univariate covariance functions, each looking at a different input variable, describe additive models.

- Products of univariate covariance functions for different inputs yield multivariate covariance functions that allow for interactions.
Univariate Functions with Combined Covariance Functions

\[
\text{Cov}(y_{i1}, y_{i2}) = 1 + x_{i1} x_{i2} \\
+ 0.1^2 \exp(-((x_{i1} - x_{i2}) / (1/3))^2) \\
\]

\[
\text{Cov}(y_{i1}, y_{i2}) = \exp(-((x_{i1} - x_{i2})^2) \\
+ 0.1^2 \exp(-((x_{i1} - x_{i2}) / (1/5))^2) \\
\]
Bivariate Functions Using an Additive Covariance Function

Here are two functions drawn from a Gaussian process whose covariance function is the sum of a function of $x_1$ and a function of $x_2$.

\[
\text{Cov}(y_{i_1}, y_{i_2}) = \exp\left( -\left(\frac{x_{i_1,1} - x_{i_2,1}}{0.88}\right)^2 \right) + \exp\left( -\left(\frac{x_{i_1,2} - x_{i_2,2}}{0.24}\right)^2 \right)
\]
Bivariate Functions Using a Product Covariance Function

Here are two functions drawn from a Gaussian process whose covariance function is the product of a function of $x_1$ and a function of $x_2$.

\[
\text{Cov}(y_{i1}, y_{i2}) = \exp \left( - \left( \frac{x_{i1,1} - x_{i2,1}}{0.88} \right)^2 - \left( \frac{x_{i1,2} - x_{i2,2}}{0.24} \right)^2 \right)
\]
How Were These Plots Generated?

To produce these plots of functions drawn from a Gaussian process, I first defined a fine grid of $x$ values, $x_1, \ldots, x_n$. For the 1D plots, the grid might have values of $-2.00, -1.99, -1.98, \ldots, 1.98, 1.99, 2.00$.

For the processes shown, the mean function is zero.

I computed the covariance matrix, $C$, of the $y_i$ values that go with each of the $x_i$. The elements are $C_{i_1, i_2} = \text{Cov}(y_{i_1}, y_{i_2})$.

I found the Cholesky decomposition of $C$, which is the lower-triangular matrix $L$ such that $C = LL^T$.

I drew a random vector $n$ of length $n$, in which the components are drawn independently from the normal distribution with mean 0 and variance 1.

I set the vector of $y_i$ values, $y = [y_1 \ldots y_n]^T$, to $Ln$.

Clearly, $E(y) = E(Ln) = L E(n) = 0$. The covariance matrix of $y$ is

$$E(yy^T) = E(Lnn^T L^T) = LE(nn^T)L^T = LL^T = C$$