Problem formulation

Consider a multivariate function \( f(x) : \mathbb{R}^n \rightarrow \mathbb{R}^m \) on which we put a Gaussian Process (GP) prior
\[
\mathcal{GP}(f(x)) = \mathcal{N}(\mu_f(x), K_f(x, x')).
\]
We also know that \( f(x) \) shall obey a set of linear constraints
\[
\mathcal{F}_f f(x) = 0,
\]
where \( \mathcal{F}_f \) is a matrix of linear operators, e.g. partial derivatives or integrals. We want these constraints to be included in the GP model. Hence, the problem is:

How do we select the covariance function \( K_f(x, x') \) to ensure that the constraints are fulfilled by any sample from the prior distribution of \( f(x) \)?

Solution

We model \( f(x) \) as a linear transformation of an underlying function \( g(x) \). Since GPs are closed under linear transformations, we have
\[
f(x) = \mathcal{F}_f g(x) = \mathcal{GP}(\mathcal{F}_f \mu_g(x), \mathcal{F}_f K_g(x, x') \mathcal{F}_f^T).
\]
The constraints then become
\[
\mathcal{F}_f \mathcal{F}_f g(x) = 0.
\]
These shall be fulfilled by any arbitrary function \( g(x) \), hence we require
\[
\mathcal{F}_f \mathcal{F}_f = 0.
\]
This imposes constraints on \( \mathcal{F}_f \) rather than on \( f(x) \) directly. Hence, all we need is to find the transformation \( \mathcal{F}_f \).

Toy example revisited

Let us go back to the toy example and construct \( \mathcal{F}_f \) using the algorithm above.

Step 1: Assume that \( \mathcal{F}_f \) contains the same operators as \( \mathcal{F}_0 \).
\[
g = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix}\begin{bmatrix} \partial x_1 \\ \partial x_2 \end{bmatrix} = \Gamma \xi^g.
\]

Step 2:
\[
\mathcal{F}_f \mathcal{F}_f g = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} = \gamma_{11} \partial^2 x_1 + (\gamma_{12} + \gamma_{21}) \partial x_1 \partial x_2 + \gamma_{22} \partial^2 x_2 = 0.
\]

This implies that \( \mathcal{F}_f \) is spanned by a unique vector \( \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \) which is a vector product of \( e_1 \) and \( e_2 \).

Thus, we have
\[
f(x) = \mathcal{F}_f g(x).
\]

References
