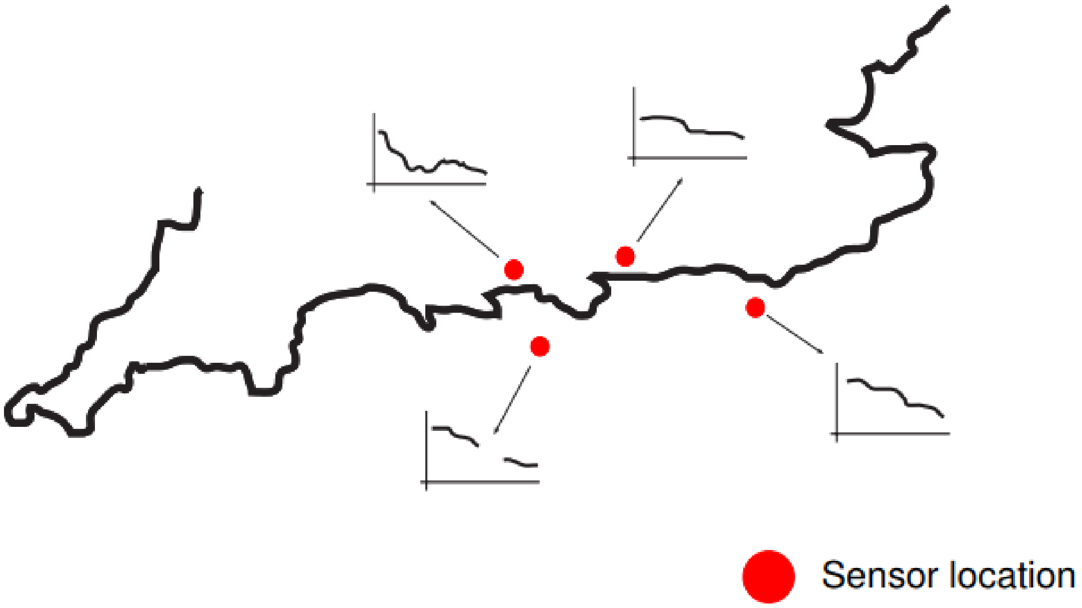
Multiple-Output Gaussian Process

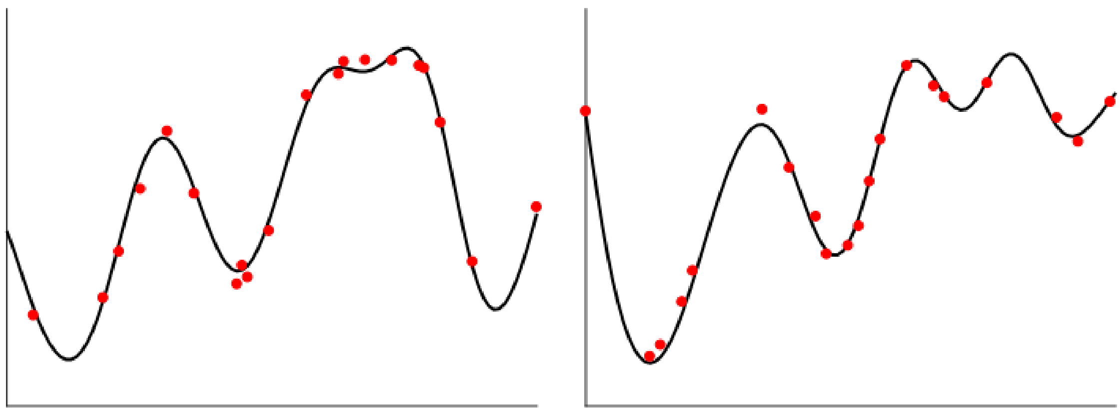
Working Situation



As the picture shows, we want to learn from the three sensors (with complete signal information) to recover the fourth one.

Dependencies between processes

Multiple-independent Output GP



$$
\begin{aligned}
f_1(\mathbf{x}) &\sim \mathcal{GP}\left(0, k_1\left(\mathbf{x}, \mathbf{x}'\right)\right) & f_2(\mathbf{x}) &\sim \mathcal{GP}\left(0, k_2\left(\mathbf{x}, \mathbf{x}'\right)\right) \\
D_1 &= \left\{\left(\mathbf{x}_{i,1}, y_1\left(\mathbf{x}_{i,2}\right)\right) \mid i = 1, \ldots, N_1\right\} & \mathcal{D}_2 &= \left\{\left(\mathbf{x}_{i,2}, y_2\left(\mathbf{x}_{i,2}\right)\right) \mid i = 1, \ldots, N_2\right\} \\
\mathbf{y}_1 &\sim \mathcal{N}\left(\mathbf{0}, \mathbf{K}_1 + \sigma_1^2\right) & \mathbf{y}_2 &\sim \mathcal{N}\left(\mathbf{0}, \mathbf{K}_2 + \sigma_2^2 \mathbf{I}\right)
\end{aligned}
$$

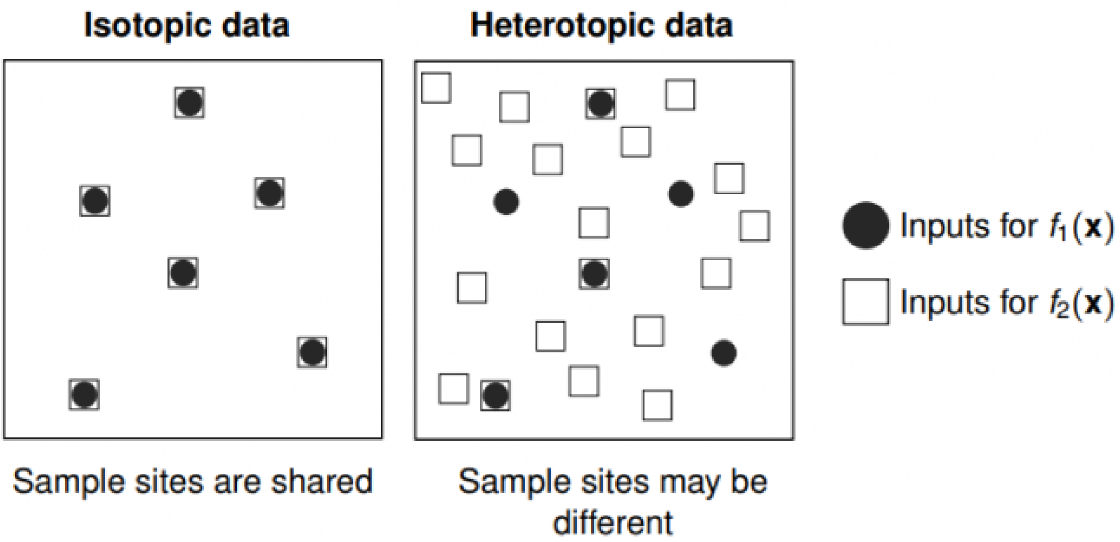
$$
\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{K}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_2 \end{bmatrix} + \begin{bmatrix} \sigma_1^2 \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \sigma_2^2 \mathbf{I} \end{bmatrix}\right) \tag{1}
$$

How to find the independences for kernel design

$$\mathbf{K}_{\mathbf{f},\mathbf{f}} = \begin{bmatrix} \mathbf{K}_1 & ? \\ ? & \mathbf{K}_2 \end{bmatrix} \tag{2}$$

Build a cross-covariance function $cov[f_1(x), f_2(x^{'})]$ such that $K_{f,f}$ is positive semi-definite.

Different input configurations of data



$$\begin{aligned} \mathcal{D}_1 &= \left\{ (\mathbf{x}_i, f_1(\mathbf{x}_i))_{i=1}^N \right\} & \mathcal{D}_1 &= \left\{ (\mathbf{x}_{i,1}, f_1(\mathbf{x}_{i,1}))_{i=1}^{N_1} \right\} \\ \mathcal{D}_2 &= \left\{ (\mathbf{x}_i, f_2(\mathbf{x}_i))_{i=1}^N \right\} & \mathcal{D}_2 &= \left\{ (\mathbf{x}_{i,2}, f_2(\mathbf{x}_{i,2}))_{i=1}^{N_2} \right\} \end{aligned} \tag{3}$$

Intrinsic Coregionalization Model

Two outputs

Sample Once

Consider two outputs $f_1(x) f_2(x)$ with $x \in \mathcal{R}^p$.

1. Sample from a GP $u(\mathbf{x}) \sim \mathcal{GP}(0, k(\mathbf{x}, \mathbf{x}'))$ to obtain $u^1(\mathbf{x})$
2. Obtain $f_1(x)$ and $f_2(x)$ by linearly transforming:

$$\begin{aligned} f_1(\mathbf{x}) &= a_1^1 u^1(\mathbf{x}) \\ f_2(\mathbf{x}) &= a_2^1 u^1(\mathbf{x}) \end{aligned}$$

For a fixed value x. we can group $f_1(x)$ and $f_2(x)$ in a vector:

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \end{bmatrix} \tag{4}$$

and this vector will be refer as a $\mathbf{vector-valued\ function}$.

The covariance for $f(x)$ is computed as:

$$\operatorname{cov}(\mathbf{f}(\mathbf{x}), \mathbf{f}(\mathbf{x}')) = \mathbb{E}\left\{ \mathbf{f}(\mathbf{x}) [\mathbf{f}(\mathbf{x}')]^\top \right\} - \mathbb{E}\{\mathbf{f}(\mathbf{x})\} [\mathbb{E}\{\mathbf{f}(\mathbf{x}')\}]^\top \tag{5}$$

$$
\begin{aligned}
\mathbb{E}\left\{\begin{bmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \end{bmatrix}\left[f_1\left(\mathbf{x}'\right) \quad f_2\left(\mathbf{x}'\right)\right]\right\} &= \begin{bmatrix} \mathbb{E}\left\{f_1(\mathbf{x}) f_1\left(\mathbf{x}'\right)\right\} & \mathbb{E}\left\{f_1(\mathbf{x}) f_2\left(\mathbf{x}'\right)\right\} \\ \mathbb{E}\left\{f_2(\mathbf{x}) f_1\left(\mathbf{x}'\right)\right\} & \mathbb{E}\left\{f_2(\mathbf{x}) f_2\left(\mathbf{x}'\right)\right\} \end{bmatrix} \\
\mathbb{E}\left\{f_1(\mathbf{x}) f_1\left(\mathbf{x}'\right)\right\} &= \mathbb{E}\left\{a_1^1 u^1(\mathbf{x}) a_1^1 u^1\left(\mathbf{x}'\right)\right\} = \left(a_1^1\right)^2 \mathbb{E}\left\{u^1(\mathbf{x}) u^1\left(\mathbf{x}'\right)\right\} \\
\mathbb{E}\left\{f_1(\mathbf{x}) f_2\left(\mathbf{x}'\right)\right\} &= \mathbb{E}\left\{a_1^1 u^1(\mathbf{x}) a_2^1\left(\mathbf{x}'\right)\right\} = a_1^1 a_2^1 \mathbb{E}\left\{u^1(\mathbf{x}) u^1\left(\mathbf{x}'\right)\right\} \\
\mathbb{E}\left\{f_2(\mathbf{x}) f_2\left(\mathbf{x}'\right)\right\} &= \mathbb{E}\left\{a_2^1 u^1(\mathbf{x}) a_2^1 u^1\left(\mathbf{x}'\right)\right\} = \left(a_2^1\right)^2 \mathbb{E}\left\{u^1(\mathbf{x}) u^1\left(\mathbf{x}'\right)\right\}
\end{aligned}
$$

So that term could be written as:

$$
\begin{aligned}
\mathbb{E}\left\{\mathbf{f}(\mathbf{x})\left[\mathbf{f}\left(\mathbf{x}'\right)\right]^{\top}\right\} &= \begin{bmatrix} \left(a_1^1\right)^2 \mathbb{E}\left\{u^1(\mathbf{x}) u^1\left(\mathbf{x}'\right)\right\} & a_1^1 a_2^1 \mathbb{E}\left\{u^1(\mathbf{x}) u^1\left(\mathbf{x}'\right)\right\} \\ a^1 a^2 \mathbb{E}\left\{u^1(\mathbf{x}) u^1\left(\mathbf{x}'\right)\right\} & \left(a_2^1\right)^2 \mathbb{E}\left\{u^1(\mathbf{x}) u^1\left(\mathbf{x}'\right)\right\} \end{bmatrix} \\
&= \begin{bmatrix} \left(a_1^1\right)^2 & a_1^1 a_2^1 \\ a_1^1 a_2^1 & \left(a_2^1\right)^2 \end{bmatrix} \mathbb{E}\left\{u^1(\mathbf{x}) u^1\left(\mathbf{x}'\right)\right\}
\end{aligned} \tag{6}
$$

The term $\mathbb{E}\{\mathbf{f}(\mathbf{x})\}$ is computed as:

$$
\mathbb{E}\left\{\begin{bmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \end{bmatrix}\right\} = \begin{bmatrix} \mathbb{E}\left\{f_1(\mathbf{x})\right\} \\ \mathbb{E}\left\{f_1(\mathbf{x})\right\} \end{bmatrix} = \begin{bmatrix} \mathbb{E}\left\{a_1^1 u^1(\mathbf{x})\right\} \\ \mathbb{E}\left\{a_2^1 u^1(\mathbf{x})\right\} \end{bmatrix}] = \begin{bmatrix} a_1^1 \\ a_2^1 \end{bmatrix} \mathbb{E}\left\{u^1(\mathbf{x})\right\} \tag{7}
$$

Putting them together, the covariance for $f(x^{'})$ follows as:

$$
\begin{bmatrix} \left(a_1^1\right)^2 & a_1^1 a_2^1 \\ a_1^1 a_2^1 & \left(a_2^1\right)^2 \end{bmatrix} \mathbb{E}\left\{u^1(\mathbf{x}) u^1\left(\mathbf{x}'\right)\right\} - \begin{bmatrix} a_1^1 \\ a_2^1 \end{bmatrix}\left[a_1^1 \quad a_2^1\right]\left\{u^1(\mathbf{x})\right\} \mathbb{E}\left\{u^1\left(\mathbf{x}'\right)\right\} \tag{8}
$$

Defining $\mathbf{a} = \left[a_1^1 \quad a_2^1\right]^{\top}$,

$$
\begin{aligned}
\operatorname{cov}\left(\mathbf{f}(\mathbf{x}), \mathbf{f}\left(\mathbf{x}'\right)\right) &= \mathbf{a}\mathbf{a}^{\top} \mathbb{E}\left\{u^1(\mathbf{x}) u^1\left(\mathbf{x}'\right)\right\} - \mathbf{a}\mathbf{a}^{\top} \mathbb{E}\left\{u^1(\mathbf{x})\right\} \mathbb{E}\left\{u^1\left(\mathbf{x}'\right)\right\} \\
&= \mathbf{a}\mathbf{a}^{\top} \underbrace{\left[\mathbb{E}\left\{u^1(\mathbf{x}) u^1\left(\mathbf{x}'\right)\right\} - \mathbb{E}\left\{u^1(\mathbf{x})\right\} \mathbb{E}\left\{u^1\left(\mathbf{x}'\right)\right\}\right]}_{k\left(\mathbf{x}, \mathbf{x}'\right)} \\
&= \mathbf{a}\mathbf{a}^{\top} k\left(\mathbf{x}, \mathbf{x}'\right)
\end{aligned}
$$

We define $\mathbf{B} = \mathbf{a}\mathbf{a}^{\top}$, leading to

$$
\operatorname{cov}\left(\mathbf{f}(\mathbf{x}), \mathbf{f}\left(\mathbf{x}'\right)\right) = \mathbf{B} k\left(\mathbf{x}, \mathbf{x}'\right) = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} k\left(\mathbf{x}, \mathbf{x}'\right) \tag{9}
$$

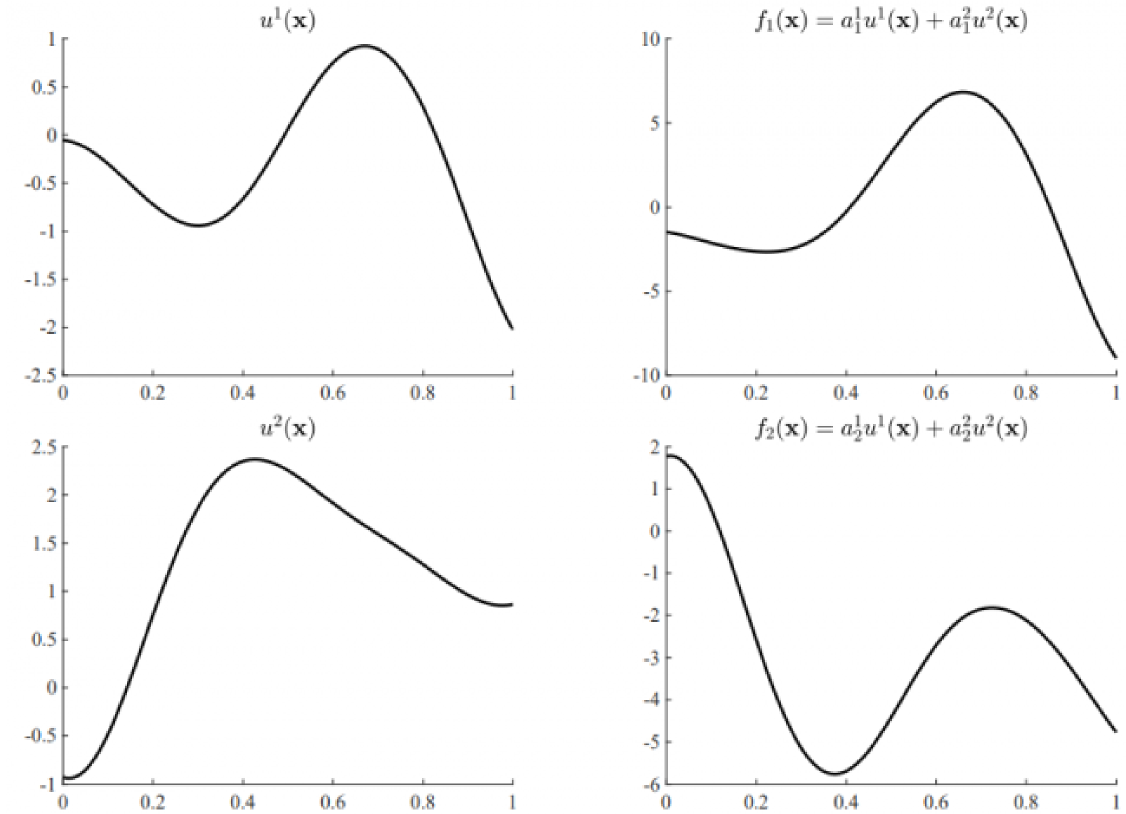
and the $\mathbf{B}$ has rank one, since it is the result of the multiplication of two column-vector.

Sample Twice

Sample **twice** from a GP $u(\mathbf{x}) \sim \mathcal{GP}\left(0, k\left(\mathbf{x}, \mathbf{x}'\right)\right)$ to obtain $u^1(\mathbf{x})$ and $u^2(\mathbf{x})$.

Adding a scaled transformation.:

$$
\begin{aligned}
f_1(\mathbf{x}) &= a_1^1 u^1(\mathbf{x}) + a_1^2 u^2(\mathbf{x}) \\
f_2(\mathbf{x}) &= a_2^1 u^1(\mathbf{x}) + a_2^2 u^2(\mathbf{x})
\end{aligned} \tag{10}
$$



**

Notice that the u_1 and u_2 are independent, although they share the same covariance k.

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} (a_1^1) & a_1^2 \\ a_2^1 & (a_2^2) \end{bmatrix} \begin{bmatrix} u^1 \\ u^2 \end{bmatrix} \tag{11}$$

The vector-valued function can be written as $f(x)$, where $\mathbf{a}^1 = \left[a_1^1 \ \ a_2^1\right]^\top$ and $\mathbf{a}^2 = \left[a_1^2 \ \ a_2^2\right]^\top$

The covariance for $f(x)$ is computed as:

$$\begin{aligned} \operatorname{cov}(\mathbf{f}(\mathbf{x}), \mathbf{f}(\mathbf{x}')) &= \mathbf{a}^1 \left(\mathbf{a}^1\right)^\top \operatorname{cov}\left(u^1(\mathbf{x}), u^1(\mathbf{x}')\right) + \mathbf{a}^2 \left(\mathbf{a}^2\right)^\top \operatorname{cov}\left(u^2(\mathbf{x}), u^2(\mathbf{x}')\right) \\ &= \mathbf{a}^1 \left(\mathbf{a}^1\right)^\top k(\mathbf{x}, \mathbf{x}') + \mathbf{a}^2 \left(\mathbf{a}^2\right)^\top k(\mathbf{x}, \mathbf{x}') \\ &= \left[\mathbf{a}^1 \left(\mathbf{a}^1\right)^\top + \mathbf{a}^2 \left(\mathbf{a}^2\right)^\top\right] k(\mathbf{x}, \mathbf{x}') \end{aligned}$$

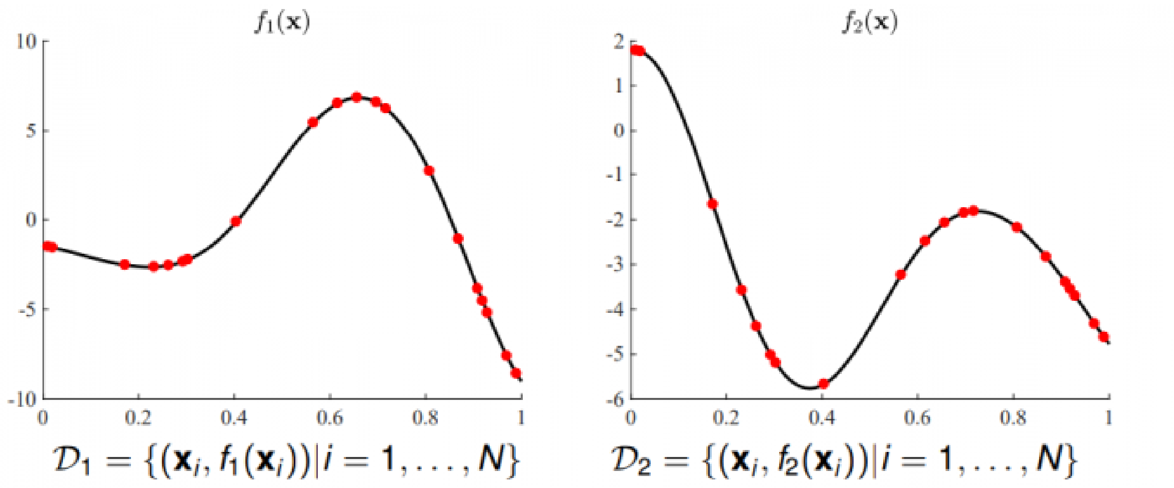
notice that u_1 and u_2 are independent, so their variance could be added directly.

we define $\mathbf{B} = \mathbf{a}^1 \left(\mathbf{a}^1\right)^\top + \mathbf{a}^2 \left(\mathbf{a}^2\right)^\top$, leading to:

$$\operatorname{cov}(\mathbf{f}(\mathbf{x}), \mathbf{f}(\mathbf{x}')) = \mathbf{B} k(\mathbf{x}, \mathbf{x}') = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} k(\mathbf{x}, \mathbf{x}') \tag{12}$$

Notice that B has rank two.

Observed Data:



$$\begin{bmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \end{bmatrix} = \begin{bmatrix} f_1(\mathbf{x}_1) \\ \vdots \\ f_1(\mathbf{x}_N) \\ f_2(\mathbf{x}_1) \\ \vdots \\ f_2(\mathbf{x}_N) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} b_{11}\mathbf{K} & b_{12}\mathbf{K} \\ b_{21}\mathbf{K} & b_{22}\mathbf{K} \end{bmatrix}\right) \tag{13}$$

The matrix $\mathbf{k} \in \mathcal{R}^{\mathbf{N*N}}$ has elements $k(x_i, x_j)$.

If we use **Kronecker product** we would get:

$$\begin{bmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \end{bmatrix} = \begin{bmatrix} f_1(\mathbf{x}_1) \\ \vdots \\ f_1(\mathbf{x}_N) \\ f_2(\mathbf{x}_1) \\ \vdots \\ f_2(\mathbf{x}_N) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \mathbf{B} \otimes \mathbf{K}\right) \tag{14}$$

General Case

Consider a set of functions $\{f_d(\mathbf{x})\}_{d=1}^{D}$.

In the ICM,

$$f_d(\mathbf{x}) = \sum_{i=1}^{R} a_d^i u^i(\mathbf{x}) \tag{15}$$

where the functions $u_i(x)$ are GPs sampled independently, and share the same covariance function $k(x, x^{'})$.

For $\mathbf{f}(\mathbf{x}) = [f_1(\mathbf{x}) \cdots f_D(\mathbf{x})]^\top$, the covariance is given as:

$$\operatorname{cov}[\mathbf{f}(\mathbf{x}), \mathbf{f}(\mathbf{x}')] = \mathbf{A}\mathbf{A}^\top k(\mathbf{x}, \mathbf{x}') = \mathbf{B}k(\mathbf{x}, \mathbf{x}') \tag{16}$$

where

$$\mathbf{A} = [\mathbf{a}^1 \mathbf{a}^2 \cdots \mathbf{a}^R] \tag{17}$$

and the Rank of B is given by R.

ICM: autokrigeability

If the outputs are considered to be noise-free, prediction using the ICM under an isotopic data case is equivalent to independent prediction over each output. This circumstance is also known as autokrigeability.

The prove:

Assume that we only have two outputs: f_1, f_2

the predicated mean could be written as:

$$\mu = K_{f_*,f}(K_{f,f})^{-1}f \\ K_{f,f} = B \otimes K \tag{18}$$

$$\begin{aligned} \mu &= B \otimes K_*(B \otimes K)^{-1}f \\ &= B \otimes K_*(B^{-1} \otimes K^{-1})f \\ &= BB^{-1} \otimes K_*K^{-1}f \\ &= I \otimes K_*K^{-1}f \\ &= \begin{bmatrix} K_*K^{-1} & 0 \\ 0 & K_*K^{-1} \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} \end{aligned}$$

it means, the prediction of f_1 only depends on the data set for f_1

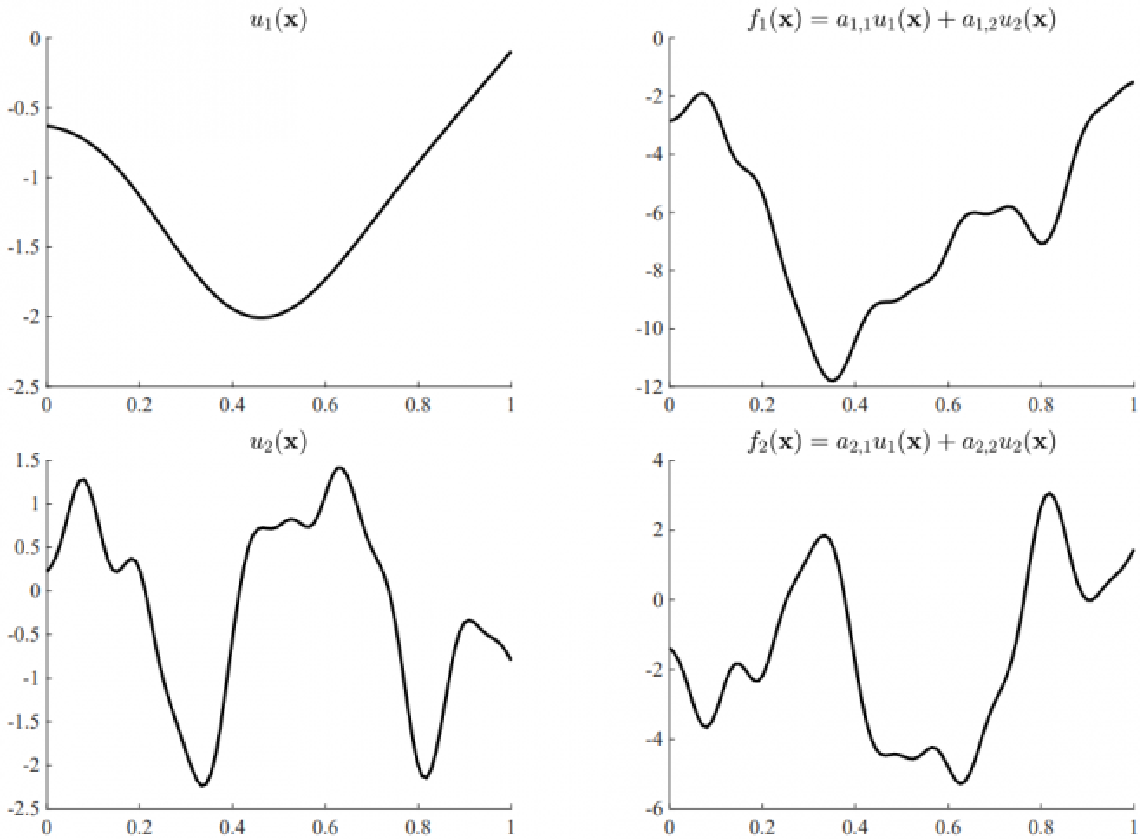
Semiparametric Latent Factor Model (SLFM)

ICM uses R samples $u^i(x)$ from $u(x)$ with the same covariance function. SLFM uses Q samples from u_q processes with different covariance functions.

Two Outputs

1. Sample from a GP $\mathcal{GP}(0, k_1(\mathbf{x}, \mathbf{x}'))$ to obtain $u_1(x)$.
2. Sample from a GP $\mathcal{GP}(0, k_2(\mathbf{x}, \mathbf{x}'))$ to obtain $u_2(x)$.
3. Adding a scaled versions:

$$\begin{aligned} f_1(\mathbf{x}) &= a_{1,1}u_1(\mathbf{x}) + a_{1,2}u_2(\mathbf{x}) \\ f_2(\mathbf{x}) &= a_{2,1}u_1(\mathbf{x}) + a_{2,2}u_2(\mathbf{x}) \end{aligned} \tag{19}$$



Similar, it can be written as:

$$\mathbf{f}(\mathbf{x}) = \mathbf{a}_1 u_1(\mathbf{x}) + \mathbf{a}_2 u_2(\mathbf{x}) \tag{20}$$

with $\mathbf{a}_1 = [a_{1,1} a_{2,1}]^\top$ and $\mathbf{a}_2 = [a_{1,2} a_{2,2}]^\top$

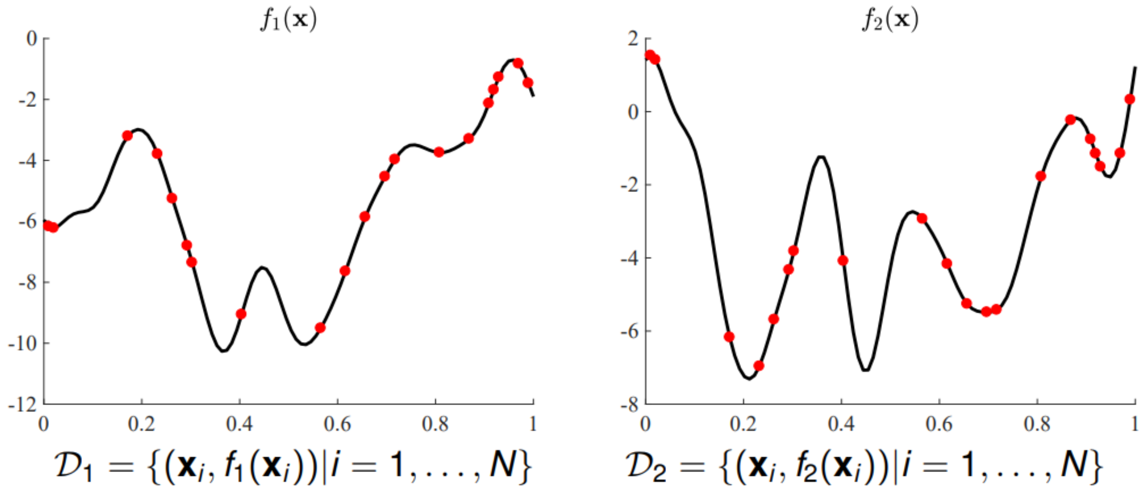
The covariance for $f(x)$ is computed as:

$$\begin{aligned} \operatorname{cov}(\mathbf{f}(\mathbf{x}), \mathbf{f}(\mathbf{x}')) &= \mathbf{a}_1(\mathbf{a}_1)^\top \operatorname{cov}(u_1(\mathbf{x}), u_1(\mathbf{x}')) + \mathbf{a}_2(\mathbf{a}_2)^\top \operatorname{cov}(u_2(\mathbf{x}), u_2(\mathbf{x}')) \\ &= \mathbf{a}_1(\mathbf{a}_1)^\top k_1(\mathbf{x}, \mathbf{x}') + \mathbf{a}_2(\mathbf{a}_2)^\top k_2(\mathbf{x}, \mathbf{x}') \end{aligned}$$

We define $\mathbf{B}_1 = \mathbf{a}_1(\mathbf{a}_1)^\top$ and $\mathbf{B}_2 = \mathbf{a}_2(\mathbf{a}_2)^\top$, leading to:

$$\operatorname{cov}(\mathbf{f}(\mathbf{x}), \mathbf{f}(\mathbf{x}')) = \mathbf{B}_1 k_1(\mathbf{x}, \mathbf{x}') + \mathbf{B}_2 k_2(\mathbf{x}, \mathbf{x}') \tag{21}$$

Notice that B_1 and B_2 have rank one.



$$\begin{bmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \end{bmatrix} = \begin{bmatrix} f_1(\mathbf{x}_1) \\ \vdots \\ f_1(\mathbf{x}_N) \\ f_2(\mathbf{x}_1) \\ \vdots \\ f_2(\mathbf{x}_N) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \mathbf{B}_1 \otimes \mathbf{K}_1 + \mathbf{B}_2 \otimes \mathbf{K}_2 \right) \tag{22}$$

General Case:

Consider a set of functions $\{f_d(\mathbf{x})\}_{d=1}^{D}$

In the SLFM,

$$f_d(\mathbf{x}) = \sum_{q=1}^{Q} a_{d,q} u_q(\mathbf{x}) \tag{23}$$

where the functions $u_q(x)$ are GPs with covariance functions $k_q(x, x^{'})$.

For $\mathbf{f}(\mathbf{x}) = [f_1(\mathbf{x}) \cdots f_D(\mathbf{x})]^{\top}$, the covariance is given as:

$$\operatorname{cov}[\mathbf{f}(\mathbf{x}), \mathbf{f}(\mathbf{x}')] = \sum_{q=1}^{Q} \mathbf{A}_q \mathbf{A}_q^{\top} k_q(\mathbf{x}, \mathbf{x}') = \sum_{q=1}^{Q} \mathbf{B}_q k_q(\mathbf{x}, \mathbf{x}') \tag{24}$$

where $A_q = a_q$.

The rank of each B_q is one.

Linear model of coregionalization (LMC)

The LMC generalizes the ICM and the SLFM allowing several independent samples from GPs with different covariances.

Consider a set of functions $\{f_d(\mathbf{x})\}_{d=1}^{D}$

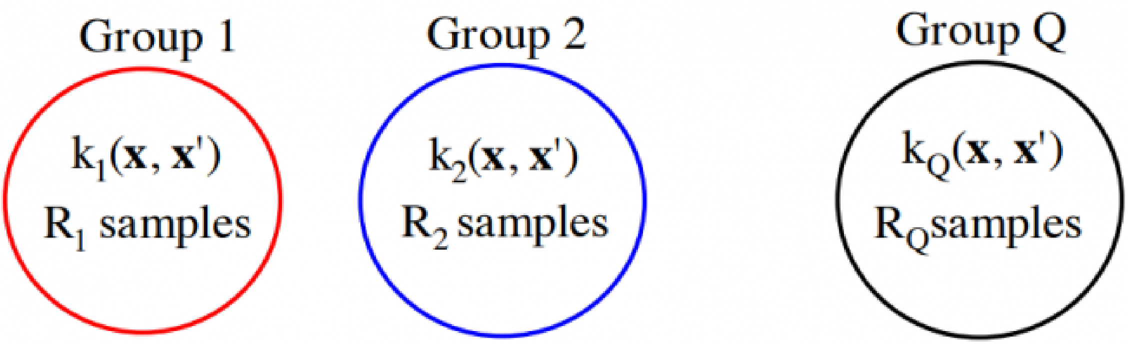
$$f_d(\mathbf{x}) = \sum_{q=1}^{Q} \sum_{i=1}^{R_q} a_{d,q}^{i} u_q^{i}(\mathbf{x}) \tag{25}$$

where the functions u_q^i are GPs with zero means and covariance functions:

$$\operatorname{cov}\left[u_q^i(\mathbf{x}), u_{q'}^{i'}(\mathbf{x}')\right] = k_q(\mathbf{x}, \mathbf{x}') \tag{26}$$

if $i = i^{'}$ and $q = q^{'}$

There are Q groups of samples. For each group, there are R_q samples obtained independently from the same GP with covariance $k_q(x, x^{'})$.



The LMC corresponds to the sum of Q ICMs.

Suppose we have D = 2, Q = 2, and R_q=2. According to LMC:

$$\begin{aligned} f_1(\mathbf{x}) &= a_{1,1}^1 u_1^1(\mathbf{x}) + a_{1,1}^2 u_1^2(\mathbf{x}) + a_{1,2}^1 u_2^1(\mathbf{x}) + a_{1,2}^2 u_2^2(\mathbf{x}) \\ f_2(\mathbf{x}) &= a_{2,1}^1 u_1^1(\mathbf{x}) + a_{2,1}^2 u_1^2(\mathbf{x}) + a_{2,2}^1 u_2^1(\mathbf{x}) + a_{2,2}^2 u_2^2(\mathbf{x}) \end{aligned} \tag{27}$$

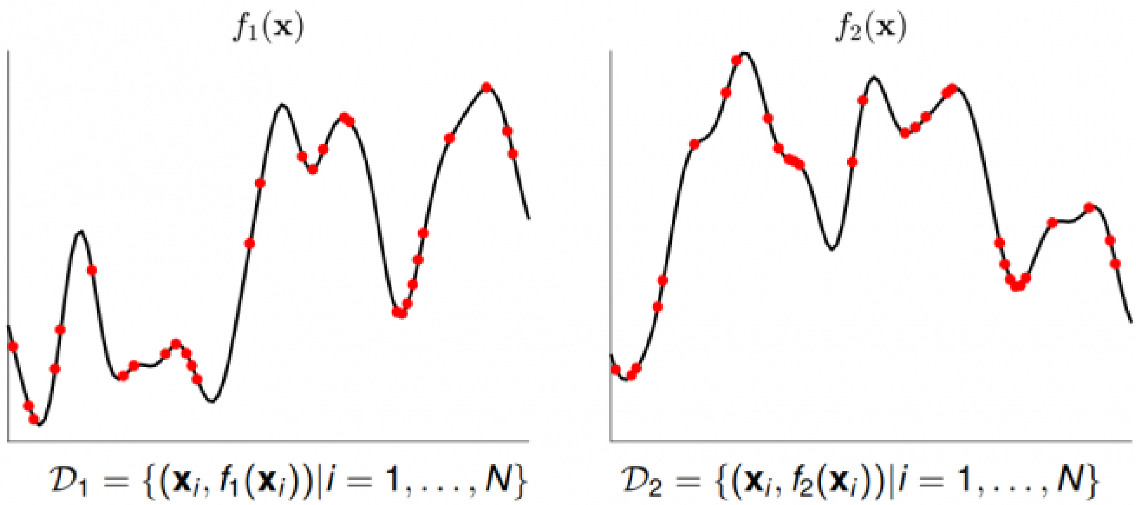
For $\mathbf{f}(\mathbf{x}) = [f_1(\mathbf{x}) \cdots f_D(\mathbf{x})]^\top$, the covariance $\operatorname{cov}[\mathbf{f}(\mathbf{x}), \mathbf{f}(\mathbf{x}')]$ is given as:

$$\operatorname{cov}[\mathbf{f}(\mathbf{x}), \mathbf{f}(\mathbf{x}')] = \sum_{q=1}^{Q} \mathbf{A}_q \mathbf{A}_q^\top k_q(\mathbf{x}, \mathbf{x}') = \sum_{q=1}^{Q} \mathbf{B}_q k_q(\mathbf{x}, \mathbf{x}') \tag{28}$$

where $\mathbf{A}_q = \left[\mathbf{a}_q^1 \mathbf{a}_q^2 \cdots \mathbf{a}_q^{R_q}\right]$.

The rank of each B_q is R_q.

The matrices B_q are known as the coregionalization matrices.



$$\begin{bmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \end{bmatrix} = \begin{bmatrix} f_1(\mathbf{x}_1) \\ \vdots \\ f_1(\mathbf{x}_N) \\ f_2(\mathbf{x}_1) \\ \vdots \\ f_2(\mathbf{x}_N) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \sum_{q=1}^{Q} \mathbf{B}_q \otimes \mathbf{K}_q \right) \tag{29}$$