Additional material

March 7, 2020

Contents

1 Example of generalization to the multivariate case

Next, as an example we show the matrix $\mathbb{S}$ and eigenfunctions and eigenvalues for a two-dimensional input vector $\boldsymbol{x} = \{\boldsymbol{x}_1, \boldsymbol{x}_2\}$ ($D = 2$) and three eigenfunctions and eigenvalues ($J = 3$) for every dimension. The number of new multidimensional eigenfunctions ϕ_j^* and eigenvalues λ_j^* is $J^D = 3^2 = 9$ ($j = \{1, \cdots, J^D\}$).

The matrix $\mathbb{S} \in \mathbb{R}^{9\times 2}$ is

$$\mathbb{S} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 2 & 1 \\ 2 & 2 \\ 2 & 3 \\ 3 & 1 \\ 3 & 2 \\ 3 & 3 \end{bmatrix}$$

and the multidimensional eigenfunctions and eigenvalues

$$\begin{aligned}
\phi_1^*(\boldsymbol{x}) &= \phi_1(\boldsymbol{x}_1)\cdot\phi_1(\boldsymbol{x}_2) & \boldsymbol{\lambda}_1^* &= \{\lambda_1(\boldsymbol{x}_1), \lambda_1(\boldsymbol{x}_2)\} \\
\phi_2^*(\boldsymbol{x}) &= \phi_1(\boldsymbol{x}_1)\cdot\phi_2(\boldsymbol{x}_2) & \boldsymbol{\lambda}_2^* &= \{\lambda_1(\boldsymbol{x}_1), \lambda_2(\boldsymbol{x}_2)\} \\
\phi_3^*(\boldsymbol{x}) &= \phi_1(\boldsymbol{x}_1)\cdot\phi_3(\boldsymbol{x}_2) & \boldsymbol{\lambda}_3^* &= \{\lambda_1(\boldsymbol{x}_1), \lambda_3(\boldsymbol{x}_2)\} \\
\phi_4^*(\boldsymbol{x}) &= \phi_2(\boldsymbol{x}_1)\cdot\phi_1(\boldsymbol{x}_2) & \boldsymbol{\lambda}_4^* &= \{\lambda_2(\boldsymbol{x}_1), \lambda_1(\boldsymbol{x}_2)\} \\
\phi_5^*(\boldsymbol{x}) &= \phi_2(\boldsymbol{x}_1)\cdot\phi_2(\boldsymbol{x}_2) & \boldsymbol{\lambda}_5^* &= \{\lambda_2(\boldsymbol{x}_1), \lambda_2(\boldsymbol{x}_2)\} \\
\phi_6^*(\boldsymbol{x}) &= \phi_2(\boldsymbol{x}_1)\cdot\phi_3(\boldsymbol{x}_2) & \boldsymbol{\lambda}_6^* &= \{\lambda_2(\boldsymbol{x}_1), \lambda_3(\boldsymbol{x}_2)\} \\
\phi_7^*(\boldsymbol{x}) &= \phi_3(\boldsymbol{x}_1)\cdot\phi_1(\boldsymbol{x}_2) & \boldsymbol{\lambda}_7^* &= \{\lambda_3(\boldsymbol{x}_1), \lambda_1(\boldsymbol{x}_2)\} \\
\phi_8^*(\boldsymbol{x}) &= \phi_3(\boldsymbol{x}_1)\cdot\phi_2(\boldsymbol{x}_2) & \boldsymbol{\lambda}_8^* &= \{\lambda_3(\boldsymbol{x}_1), \lambda_2(\boldsymbol{x}_2)\} \\
\phi_9^*(\boldsymbol{x}) &= \phi_3(\boldsymbol{x}_1)\cdot\phi_3(\boldsymbol{x}_2) & \boldsymbol{\lambda}_9^* &= \{\lambda_3(\boldsymbol{x}_1), \lambda_3(\boldsymbol{x}_2)\}
\end{aligned}$$

Now, we show another example where different number of eigenfunctions and eigenvalues are used for every dimension. We consider a three-dimensional ($D = 3$) input space, and sets of $J_1 = 2$, $J_2 = 2$ and $J_3 = 3$ eigenfunctions and eigenvalues for the first, second and third dimensions, respectively. The number of new multidimensional eigenfunctions ϕ^* and eigenvalues λ^* is $J_1 \cdot J_2 \cdot J_3 = 2 \cdot 2 \cdot 3 = 12$. The matrix $\mathbb{S} \in \mathbb{R}^{12\times 3}$ is

$$\mathbb{S} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 1 & 3 \\ 1 & 2 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \\ 2 & 1 & 1 \\ 2 & 1 & 2 \\ 2 & 1 & 3 \\ 2 & 2 & 1 \\ 2 & 2 & 2 \\ 2 & 2 & 3 \end{bmatrix}$$

and the multidimensional eigenfunctions and eigenvalues

$$\begin{aligned}
\phi_1^* &= \phi_1(\boldsymbol{x}_1) \cdot \phi_1(\boldsymbol{x}_2) \cdot \phi_1(\boldsymbol{x}_3) & \boldsymbol{\lambda}_1^* &= \{\lambda_1(\boldsymbol{x}_1), \lambda_1(\boldsymbol{x}_2), \lambda_1(\boldsymbol{x}_3)\} \\
\phi_2^* &= \phi_1(\boldsymbol{x}_1) \cdot \phi_1(\boldsymbol{x}_2) \cdot \phi_2(\boldsymbol{x}_3) & \boldsymbol{\lambda}_2^* &= \{\lambda_1(\boldsymbol{x}_1), \lambda_1(\boldsymbol{x}_2), \lambda_2(\boldsymbol{x}_3)\} \\
\phi_3^* &= \phi_1(\boldsymbol{x}_1) \cdot \phi_1(\boldsymbol{x}_2) \cdot \phi_3(\boldsymbol{x}_3) & \boldsymbol{\lambda}_3^* &= \{\lambda_1(\boldsymbol{x}_1), \lambda_1(\boldsymbol{x}_2), \lambda_3(\boldsymbol{x}_3)\} \\
\phi_4^* &= \phi_1(\boldsymbol{x}_1) \cdot \phi_2(\boldsymbol{x}_2) \cdot \phi_1(\boldsymbol{x}_3) & \boldsymbol{\lambda}_4^* &= \{\lambda_1(\boldsymbol{x}_1), \lambda_1(\boldsymbol{x}_2), \lambda_1(\boldsymbol{x}_3)\} \\
\phi_5^* &= \phi_1(\boldsymbol{x}_1) \cdot \phi_2(\boldsymbol{x}_2) \cdot \phi_2(\boldsymbol{x}_3) & \boldsymbol{\lambda}_5^* &= \{\lambda_1(\boldsymbol{x}_1), \lambda_1(\boldsymbol{x}_2), \lambda_2(\boldsymbol{x}_3)\} \\
\phi_6^* &= \phi_1(\boldsymbol{x}_1) \cdot \phi_2(\boldsymbol{x}_2) \cdot \phi_3(\boldsymbol{x}_3) & \boldsymbol{\lambda}_6^* &= \{\lambda_1(\boldsymbol{x}_1), \lambda_1(\boldsymbol{x}_2), \lambda_3(\boldsymbol{x}_3)\} \\
\phi_7^* &= \phi_2(\boldsymbol{x}_1) \cdot \phi_1(\boldsymbol{x}_2) \cdot \phi_1(\boldsymbol{x}_3) & \boldsymbol{\lambda}_7^* &= \{\lambda_2(\boldsymbol{x}_1), \lambda_2(\boldsymbol{x}_2), \lambda_1(\boldsymbol{x}_3)\} \\
\phi_8^* &= \phi_2(\boldsymbol{x}_1) \cdot \phi_1(\boldsymbol{x}_2) \cdot \phi_2(\boldsymbol{x}_3) & \boldsymbol{\lambda}_8^* &= \{\lambda_2(\boldsymbol{x}_1), \lambda_2(\boldsymbol{x}_2), \lambda_2(\boldsymbol{x}_3)\} \\
\phi_9^* &= \phi_2(\boldsymbol{x}_1) \cdot \phi_1(\boldsymbol{x}_2) \cdot \phi_3(\boldsymbol{x}_3) & \boldsymbol{\lambda}_9^* &= \{\lambda_2(\boldsymbol{x}_1), \lambda_2(\boldsymbol{x}_2), \lambda_3(\boldsymbol{x}_3)\} \\
\phi_{10}^* &= \phi_2(\boldsymbol{x}_1) \cdot \phi_2(\boldsymbol{x}_2) \cdot \phi_1(\boldsymbol{x}_3) & \boldsymbol{\lambda}_{10}^* &= \{\lambda_2(\boldsymbol{x}_1), \lambda_2(\boldsymbol{x}_2), \lambda_1(\boldsymbol{x}_3)\} \\
\phi_{11}^* &= \phi_2(\boldsymbol{x}_1) \cdot \phi_2(\boldsymbol{x}_2) \cdot \phi_2(\boldsymbol{x}_3) & \boldsymbol{\lambda}_{11}^* &= \{\lambda_2(\boldsymbol{x}_1), \lambda_2(\boldsymbol{x}_2), \lambda_2(\boldsymbol{x}_3)\} \\
\phi_{12}^* &= \phi_2(\boldsymbol{x}_1) \cdot \phi_2(\boldsymbol{x}_2) \cdot \phi_3(\boldsymbol{x}_3) & \boldsymbol{\lambda}_{12}^* &= \{\lambda_2(\boldsymbol{x}_1), \lambda_2(\boldsymbol{x}_2), \lambda_3(\boldsymbol{x}_3)\}
\end{aligned}$$

2 Multidimensional generalization of covariance functions and spectral densities

2.1 Square Exponential covariance function (k) and spectral density (S)

2.1.1 Using norm-L2 (Euclidean distance)

$$\boldsymbol{x}, \boldsymbol{x}' \in \mathbb{R}^{D=2};\;\; \tau_{L2} = |\boldsymbol{x}-\boldsymbol{x}'|_{L2} = \sqrt{(\boldsymbol{x}-\boldsymbol{x}')^\top(\boldsymbol{x}-\boldsymbol{x}')} = \sqrt{(x_1-x_1')^2+(x_2-x_2')^2} = \sqrt{r_1^2+r_2^2} \in \mathbb{R};\;\; k(\tau_{L2},\ell) = \exp\left(-\frac{1}{2}\frac{\tau^2}{\ell^2}\right)$$

$$\omega_{L2} = \sqrt{s_1^2+s_2^2} \in \mathbb{R};\;\; S(\omega_{L2},\ell) = \sqrt{2\pi}^D \cdot \ell^D \cdot \exp\left(-\frac{1}{2}\ell^2\omega_{L2}^2\right)$$

$\ell \in \mathbb{R}$	$\begin{aligned} k(\tau_{L2},\ell) &= k(\lvert\boldsymbol{x}-\boldsymbol{x}'\rvert_{L2},\ell) \\ &= \exp\left(-\frac{1}{2}\frac{(\boldsymbol{x}-\boldsymbol{x}')^\top(\boldsymbol{x}-\boldsymbol{x}')}{\ell^2}\right) \\ &= \exp\left(-\frac{1}{2}\frac{\sum_{i=1}^2 (x_i-x_i')^2}{\ell^2}\right) \\ &= \exp\left(-\frac{1}{2}\sum_{i=1}^D \frac{r_i^2}{\ell^2}\right) \end{aligned}$	$\begin{aligned} S(\omega_{L2},\ell) &= \sqrt{2\pi}^D \cdot \ell^D \cdot \exp\left(-\frac{1}{2}\ell^2(s_1^2+s_2^2)\right) \\ &= \sqrt{2\pi}^D \cdot \ell^D \cdot \exp\left(-\frac{1}{2}\sum_{i=1}^D \ell^2 s_i^2)\right) \end{aligned}$	-ISOTROPIC -SEPARABLE: $k(\lvert\boldsymbol{x}-\boldsymbol{x}'\rvert_{L2},\boldsymbol{\ell})$ $= k(\lvert x_1-x_1'\rvert,\ell_1)k(\lvert x_2-x_2'\rvert,\ell_2)$ $S(\omega_{L2},\boldsymbol{\ell}) = S(s_1,\ell_1)S(s_2,\ell_2)$
$\boldsymbol{\ell} \in \mathbb{R}^2$	$\begin{aligned} k(\tau_{L2},\boldsymbol{\ell}) &= k(\lvert\boldsymbol{x}-\boldsymbol{x}'\rvert_{L2},\boldsymbol{\ell}) \\ &= \exp\left(-\frac{1}{2}\sum_{i=1}^D \frac{r_i^2}{\ell_i^2}\right) \end{aligned}$	$S(\omega_{L2},\boldsymbol{\ell}) = \sqrt{2\pi}^D \cdot \prod_{i=1}^D \ell_i \cdot \exp\left(-\frac{1}{2}\sum_{i=1}^D \ell_i^2 s_i^2\right)$	
$\boldsymbol{\ell} \in \mathbb{R}^2$ Separable kernel	$k(\lvert x_1-x_1'\rvert,\ell_1)k(\lvert x_2-x_2'\rvert,\ell_2)$ $= \exp\left(-\frac{1}{2}\frac{r_1^2}{\ell_1^2}\right)\exp\left(-\frac{1}{2}\frac{r_2^2}{\ell_2^2}\right)$ $= \exp\left(-\frac{1}{2}\sum_{i=1}^D \frac{r_i^2}{\ell_i^2}\right)$	$S(s_1,\ell_1)S(s_2,\ell_2)$ $= \sqrt{2\pi}\cdot\ell_1\cdot\exp\left(-\frac{1}{2}\ell_1^2 s_1^2\right)$ $\times\sqrt{2\pi}\cdot\ell_2\cdot\exp\left(-\frac{1}{2}\ell_2^2 s_2^2\right)$ $= \sqrt{2\pi}^D \cdot \prod_{i=1}^D \ell_i \cdot \exp\left(-\frac{1}{2}\sum_{i=1}^D \ell_i^2 s_i^2\right)$	

2.1.2 Using norm-L1

$$\boldsymbol{x}, \boldsymbol{x}' \in \mathbb{R}^{D=2};\ \ \tau_{L1} = |\boldsymbol{x} - \boldsymbol{x}'|_{L1} = |x_1 - x_1'| + |x_2 - x_2'| = r_1 + r_2 \in \mathbb{R};\ \ k(\tau_{L1}, \ell) = \exp\left(-\frac{1}{2}\frac{\tau_{L1}^2}{\ell^2}\right)$$

$$\omega_{L1} = s_1 + s_2 \in \mathbb{R};\ \ S(\omega_{L1}, \ell) = \sqrt{2\pi}^D \cdot \ell^D \cdot \exp\left(-\frac{1}{2}\ell^2\omega_{L1}^2\right)$$

$\ell \in \mathbb{R}$	$\begin{aligned} k(\tau_{L1}, \ell) &= k(\lvert\boldsymbol{x} - \boldsymbol{x}'\rvert_{L1}, \ell) \\ &= \exp\left(-\frac{1}{2}\frac{(r_1+r_2)(r_1+r_2)}{\ell^2}\right) \\ &= \exp\left(-\frac{1}{2}\frac{(r_1+r_2)}{\ell}\frac{(r_1+r_2)}{\ell}\right) \\ &= \exp\left(-\frac{1}{2}\left(\frac{r_1}{\ell}+\frac{r_2}{\ell}\right)\left(\frac{r_1}{\ell}+\frac{r_2}{\ell}\right)\right) \end{aligned}$	$\begin{aligned} S(\omega_{L1}, \ell) &= \sqrt{2\pi}^D \cdot \ell^D \cdot \exp\Big(-\frac{1}{2}\ell^2(s_1+s_2) \\ &\cdot\ (s_1+s_2)\Big) \\ &= \sqrt{2\pi}^D \cdot \ell^D \cdot \exp\Big(-\frac{1}{2}\ell(s_1+s_2) \\ &\cdot\ \ell(s_1+s_2)\Big) \\ &= \sqrt{2\pi}^D \cdot \ell^D \cdot \exp\Big(-\frac{1}{2}(\ell s_1+\ell s_2) \\ &\cdot\ (\ell s_1+\ell s_2)\Big) \end{aligned}$	-ISOTROPIC -NO SEPARABLE: $k(\lvert\boldsymbol{x} - \boldsymbol{x}'\rvert_{L1}, \boldsymbol{\ell})$ $\neq k(\lvert x_1 - x_1'\rvert, \ell_1)k(\lvert x_2 - x_2'\rvert, \ell_2)$ $S(\omega_{L1}, \boldsymbol{\ell}) \neq S(s_1, \ell_1)S(s_2, \ell_2)$
$\boldsymbol{\ell} \in \mathbb{R}^2$	$\begin{aligned} k(\tau_{L1}, \boldsymbol{\ell}) &= k(\lvert\boldsymbol{x} - \boldsymbol{x}'\rvert_{L1}, \boldsymbol{\ell}) \\ &= \exp\left(-\frac{1}{2}\left(\frac{r_1}{\ell_1}+\frac{r_2}{\ell_2}\right)\left(\frac{r_1}{\ell_1}+\frac{r_2}{\ell_2}\right)\right) \end{aligned}$	$\begin{aligned} S(\omega_{L1}, \boldsymbol{\ell}) &= \sqrt{2\pi}^D \cdot \ell_1\ell_2 \cdot \exp\Big(-\frac{1}{2}(\ell_1 s_1+\ell_2 s_2) \\ &\cdot\ (\ell_1 s_1+\ell_2 s_2)\Big) \end{aligned}$	
$\boldsymbol{\ell} \in \mathbb{R}^2$ Separable kernel	$k(\lvert x_1 - x_1'\rvert, \ell_1)k(\lvert x_2 - x_2'\rvert, \ell_2)$ $= \exp\left(-\frac{1}{2}\frac{r_1^2}{\ell_1^2}\right)\exp\left(-\frac{1}{2}\frac{r_2^2}{\ell_2^2}\right)$ $= \exp\left(-\frac{1}{2}\sum_{i=1}^{2}\frac{r_i^2}{\ell_i^2}\right)$	$S(s_1, \ell_1)S(s_2, \ell_2)$ $= \sqrt{2\pi} \cdot \ell_1 \cdot \exp\left(-\frac{1}{2}\ell_1^2 s_1^2\right)$ $\times\sqrt{2\pi} \cdot \ell_2 \cdot \exp\left(-\frac{1}{2}\ell_2^2 s_2^2\right)$ $= \sqrt{2\pi}^D \cdot \prod_{i=1}^{D}\ell_i \cdot \exp\left(-\frac{1}{2}\sum_{i=1}^{D}\ell_i^2 s_i^2\right)$	

2.1.3 Using vector difference of inputs

$$\boldsymbol{x},\boldsymbol{x}' \in \mathbb{R}^{D=2};\;\; \boldsymbol{\tau} = \boldsymbol{x}-\boldsymbol{x}' = (x_1-x_1', x_2-x_2') = (r_1,r_2) \in \mathbb{R}^2;\;\; k(\boldsymbol{\tau},\ell) = \exp\left(-\frac{1}{2}\frac{\boldsymbol{\tau}^2}{\ell^2}\right)$$

$$\boldsymbol{\omega} = (s_1,s_2) \in \mathbb{R}^2;\;\; S(\boldsymbol{\omega},\ell) = \sqrt{2\pi}^{D}\cdot \ell^D \cdot \exp\left(-\frac{1}{2}\ell^2\boldsymbol{\omega}^2\right)$$

$\ell \in \mathbb{R}$	$k(\boldsymbol{\tau},\ell) = k(\boldsymbol{x}-\boldsymbol{x}',\ell)$ $= \exp\left(-\frac{1}{2}\frac{\boldsymbol{\tau}^\top\boldsymbol{\tau}}{\ell^2}\right)$ $= \exp\left(-\frac{1}{2}\frac{(r_1,r_2)^\top(r_1,r_2)}{\ell^2}\right)$ $= \exp\left(-\frac{1}{2}\sum_{i=1}^{D}\frac{r_i^2}{\ell^2}\right)$	$S(\boldsymbol{\omega},\ell) = S(\boldsymbol{x}-\boldsymbol{x}',\ell)$ $= \sqrt{2\pi}^{D}\cdot\ell^D\cdot\exp\left(-\frac{1}{2}\ell^2\boldsymbol{\omega}^\top\boldsymbol{\omega}\right)$ $= \sqrt{2\pi}^{D}\cdot\ell^D\cdot\exp\Big(-\frac{1}{2}\ell^2(s_1,s_2)^\top$ $\cdot\;(s_1,s_2)\Big)$ $= \sqrt{2\pi}^{D}\cdot\ell^D\cdot\exp\left(-\frac{1}{2}\sum_{i=1}^{D}\ell^2 s_i^2\right)$	-ISOTROPIC -SEPARABLE: $k(\boldsymbol{x}-\boldsymbol{x}',\boldsymbol{\ell})$ $= k(\lvert x_1-x_1'\rvert,\ell_1)k(\lvert x_2-x_2'\rvert,\ell_2)$ $S(\boldsymbol{\omega},\boldsymbol{\ell}) = S(s_1,\ell_1)S(s_2,\ell_2)$
$\boldsymbol{\ell} \in \mathbb{R}^2$	$k(\boldsymbol{\tau},\boldsymbol{\ell}) = k(\boldsymbol{x}-\boldsymbol{x}',\boldsymbol{\ell})$ $= \exp\left(-\frac{1}{2}\sum_{i=1}^{D}\frac{r_i^2}{\ell_i^2}\right)$	$S(\boldsymbol{\omega},\ell) = \sqrt{2\pi}^{D}\cdot\prod_{i=1}^{D}\ell_i\cdot\exp\left(-\frac{1}{2}\sum_{i=1}^{D}\ell_i^2 s_i^2\right)$	
$\boldsymbol{\ell} \in \mathbb{R}^2$ Separable kernel	$k(\lvert x_1-x_1'\rvert,\ell_1)k(\lvert x_2-x_2'\rvert,\ell_2)$ $= \exp\left(-\frac{1}{2}\frac{r_1^2}{\ell_1^2}\right)\exp\left(-\frac{1}{2}\frac{r_2^2}{\ell_2^2}\right)$ $= \exp\left(-\frac{1}{2}\sum_{i=1}^{D}\frac{r_i^2}{\ell_i^2}\right)$	$S(s_1,\ell_1)S(s_2,\ell_2)$ $= \sqrt{2\pi}\cdot\ell_1\cdot\exp\left(-\frac{1}{2}\ell_1^2 s_1^2\right)$ $\times\sqrt{2\pi}\cdot\ell_2\cdot\exp\left(-\frac{1}{2}\ell_2^2 s_2^2\right)$ $= \sqrt{2\pi}^{D}\cdot\prod_{i=1}^{D}\ell_i\cdot\exp\left(-\frac{1}{2}\sum_{i=1}^{D}\ell_i^2 s_i^2\right)$	

2.2 Matern($\nu = 1/2$) covariance function (k) and spectral density (S)

2.2.1 Using norm-L2 (Euclidean distance)

$$\boldsymbol{x}, \boldsymbol{x}' \in \mathbb{R}^{D=2};\ \ \tau_{L2} = |\boldsymbol{x} - \boldsymbol{x}'|_{L2} = \sqrt{(\boldsymbol{x} - \boldsymbol{x}')^\top (\boldsymbol{x} - \boldsymbol{x}')} = \sqrt{(x_1 - x_1')^2 + (x_2 - x_2')^2} = \sqrt{r_1^2 + r_2^2} \in \mathbb{R};\ \ k(\tau_{L2}, \ell) = \exp\left(-\frac{\tau_{L2}}{\ell}\right)$$

$$\omega_{L2} = \sqrt{s_1^2 + s_2^2} \in \mathbb{R};\ \ S_\nu(\omega_{L2}, \ell) = \frac{2^D \pi^{D/2} \Gamma(\nu + D/2)(2\nu)^\nu}{\Gamma(\nu)\ell^{2\nu}} \left(\frac{2\nu}{\ell^2} + \omega_{L2}^2\right)^{-(\nu + D/2)};\ \ S_{1/2}(\omega_{L2}, \ell) = \frac{2^D \pi^{\frac{D}{2}} \Gamma(\frac{D+1}{2})}{\sqrt{\pi}\ell} \left(\frac{1}{\ell^2} + \omega_{L2}^2\right)^{-\frac{D+1}{2}}$$

$\ell \in \mathbb{R}$	$\begin{aligned} k(\tau_{L2}, \ell) &= k(\lvert\boldsymbol{x} - \boldsymbol{x}'\rvert_{L2}, \ell) \\ &= \exp\left(-\frac{\sqrt{(\boldsymbol{x} - \boldsymbol{x}')^\top (\boldsymbol{x} - \boldsymbol{x}')}}{\ell}\right) \\ &= \exp\left(-\sqrt{\frac{(\boldsymbol{x} - \boldsymbol{x}')^\top (\boldsymbol{x} - \boldsymbol{x}')}{\ell^2}}\right) \\ &= \exp\left(-\sqrt{\frac{\sum_{i=1}^2 (x_i - x_i')^2}{\ell^2}}\right) \\ &= \exp\left(-\sqrt{\sum_{i=1}^2 \frac{r_i^2}{\ell^2}}\right) \end{aligned}$	$\begin{aligned} D &= 2 \\ S_{1/2}(\omega_{L2}, \ell) &= \frac{2\pi}{\ell}\left(\frac{1}{\ell^2} + \omega_{L2}^2\right)^{-\frac{3}{2}} \\ &= 2\pi\ell^2\left(1 + \ell^2\omega_{L2}^2\right)^{-\frac{3}{2}} \\ &= 2\pi\ell^2\left(1 + \ell^2 s_1^2 + \ell^2 s_2^2\right)^{-\frac{3}{2}} \\ D &= 3 \\ S_{1/2}(\omega_{L2}, \ell) &= \frac{8\pi}{\ell}\left(\frac{1}{\ell^2} + \omega_{L2}^2\right)^{-2} \\ &= 8\pi\ell^3\left(1 + \ell^2\omega_{L2}^2\right)^{-2} \\ &= 8\pi\ell^3\left(1 + \ell^2 s_1^2 + \ell^2 s_2^2 + \ell^2 s_3^2\right)^{-2} \end{aligned}$	-ISOTROPIC -NO SEPARABLE: $k(\lvert\boldsymbol{x} - \boldsymbol{x}'\rvert_{L2}, \boldsymbol{\ell})$ $\neq k(\lvert x_1 - x_1'\rvert, \ell_1) k(\lvert x_2 - x_2'\rvert, \ell_2)$ $S_{1/2}(\omega_{L2}, \ell) \neq S_{1/2}(s_1, \ell) S_{1/2}(s_2, \ell)$
$\boldsymbol{\ell} \in \mathbb{R}^2$	$\begin{aligned} k(\tau_{L2}, \boldsymbol{\ell}) &= k(\lvert\boldsymbol{x} - \boldsymbol{x}'\rvert_{L2}, \boldsymbol{\ell}) \\ &= \exp\left(-\sqrt{\sum_{i=1}^2 \frac{r_i^2}{\ell_i^2}}\right) \end{aligned}$	$\begin{aligned} D &= 2 \\ S_{1/2}(\omega_{L2}, \boldsymbol{\ell}) &= 2\pi\ell_1\ell_2\left(1 + \sum_{i=1}^D \ell_i^2 s_i^2\right)^{-\frac{3}{2}} \\ S_{1/2}(\omega_{L2}, \boldsymbol{\ell}) &= 2\pi\prod_{i=1}^D \ell_i\left(1 + \sum_{i=1}^D \ell_i^2 s_i^2\right)^{-\frac{3}{2}} \end{aligned}$	
$\boldsymbol{\ell} \in \mathbb{R}^2$ Separable kernel	$\begin{aligned} &k(\lvert x_1 - x_1'\rvert, \ell_1) k(\lvert x_2 - x_2'\rvert, \ell_2) \\ &= \exp\left(-\frac{r_1}{\ell_1}\right)\exp\left(-\frac{r_2}{\ell_2}\right) \\ &= \exp\left(-\sum_{i=1}^2 \frac{r_i}{\ell_i}\right) \end{aligned}$	$\begin{aligned} &S_{1/2}(s_1, \ell_1) S_{1/2}(s_2, \ell_2) \\ &= \frac{2}{\ell_1}\left(\frac{1}{\ell_1^2} + s_1^2\right)^{-1} \cdot \frac{2}{\ell_2}\left(\frac{1}{\ell_2^2} + s_2^2\right)^{-1} \\ &= 4\ell_1\ell_2\left(1 + \ell_1^2 s_1^2\right)^{-1}\left(1 + \ell_2^2 s_2^2\right)^{-1} \end{aligned}$	

2.2.2 Using norm-L1

$$\boldsymbol{x}, \boldsymbol{x}' \in \mathbb{R}^{D=2};\ \ \tau_{L1} = |\boldsymbol{x}-\boldsymbol{x}'|_{L1} = |x_1 - x_1'| + |x_2 - x_2'| = r_1 + r_2 \in \mathbb{R};\ \ k(\tau_{L1}, \ell) = \exp\left(-\frac{\tau_{L1}}{\ell}\right)$$

$$\omega_{L1} = s_1 + s_2 \in \mathbb{R};\ \ S_\nu(\omega_{L2}, \ell) = \frac{2^D \pi^{D/2} \Gamma(\nu + D/2)(2\nu)^\nu}{\Gamma(\nu)\ell^{2\nu}} \left(\frac{2\nu}{\ell^2} + \omega_{L2}^2\right)^{-(\nu+D/2)};\ \ S_{1/2}(\omega_{L2}, \ell) = \frac{2^D \pi^{\frac{D}{2}} \Gamma(\frac{D+1}{2})}{\sqrt{\pi}\ell} \left(\frac{1}{\ell^2} + \omega_{L2}^2\right)^{-\frac{D+1}{2}}$$

$\ell \in \mathbb{R}$	$\begin{aligned} k(\tau_{L1}, \ell) &= k(\lvert\boldsymbol{x}-\boldsymbol{x}'\rvert_{L1}, \ell) \\ &= \exp\left(-\frac{\tau}{\ell}\right) \\ &= \exp\left(-\frac{r_1+r_2}{\ell}\right) \\ &= \exp\left(-\sum_{i=1}^{D} \frac{r_i}{\ell}\right) \end{aligned}$	$\begin{aligned} D &= 2 \\ S_{1/2}(\omega_{L1}, \ell) &= \frac{2\pi}{\ell}\left(\frac{1}{\ell^2} + \omega_{L1}^2\right)^{-\frac{3}{2}} \\ &= 2\pi\ell^2\left(1 + \ell^2\omega_{L1}^2\right)^{-\frac{3}{2}} \\ &= 2\pi\ell^2\left(1 + \ell^2(s_1+s_2)(s_1+s_2)\right)^{-\frac{3}{2}} \\ &= 2\pi\ell^2\left(1 + (\ell s_1+\ell s_2)(\ell s_1+\ell s_2)\right)^{-\frac{3}{2}} \end{aligned}$ $\begin{aligned} D &= 3 \\ S_{1/2}(\omega_{L1}, \ell) &= \frac{8\pi}{\ell}\left(\frac{1}{\ell^2} + \omega_{L1}^2\right)^{-2} \\ &= 8\pi\ell^3\left(1 + \ell^2\omega_{L1}^2\right)^{-2} \\ &= 8\pi\ell^3\left(1 + \ell^2(s_1+s_2+s_3)(s_1+s_2+s_3)\right)^{-2} \\ &= 8\pi\ell^3\left(1 + (\ell s_1+\ell s_2+\ell s_3)(\ell s_1+\ell s_2+\ell s_3)\right)^{-2} \end{aligned}$	-ISOTROPIC -SEPARABLE: $k(\lvert\boldsymbol{x}-\boldsymbol{x}'\rvert_{L1}, \boldsymbol{\ell})$ $= k(\lvert x_1 - x_1'\rvert, \ell_1)k(\lvert x_2 - x_2'\rvert, \ell_2)$ IT SHOULD BE SEPARABLE IN THE SPECTRAL DENSITY AS WELL? $S_{1/2}(\omega_{L1}, \ell)$ should be equal to $S_{1/2}(s_1, \ell)S_{1/2}(s_2, \ell)$
$\boldsymbol{\ell} \in \mathbb{R}^2$	$k(\tau_{L1}, \boldsymbol{\ell}) = \exp\left(-\sum_{i=1}^{D} \frac{r_i}{\ell_i}\right)$	$\begin{aligned} D &= 2 \\ S_{1/2}(\omega_{L1}, \boldsymbol{\ell}) &= 2\pi\ell_1\ell_2\left(1 + (\ell_1 s_1 + \ell_2 s_2)(\ell_1 s_1 + \ell_2 s_2)\right)^{-\frac{3}{2}} \\ &= 2\pi\ell_1\ell_2\left(1 + (\ell_1 s_1 + \ell_2 s_2)(\ell_1 s_1 + \ell_2 s_2)\right)^{-\frac{3}{2}} \end{aligned}$	
$\boldsymbol{\ell} \in \mathbb{R}^2$ Separable kernel	$k(\lvert x_1 - x_1'\rvert, \ell_1)k(\lvert x_2 - x_2'\rvert, \ell_2)$ $= \exp\left(-\frac{r_1}{\ell_1}\right)\exp\left(-\frac{r_2}{\ell_2}\right)$ $= \exp\left(-\sum_{i=1}^{D} \frac{r_i}{\ell_i}\right)$	$S_{1/2}(s_1, \ell_1)S_{1/2}(s_2, \ell_2)$ $= \frac{2}{\ell_1}\left(\frac{1}{\ell_1^2} + s_1^2\right)^{-1} \cdot \frac{2}{\ell_2}\left(\frac{1}{\ell_2^2} + s_2^2\right)^{-1}$ $= 4\ell_1\ell_2\left(1 + \ell_1^2 s_1^2\right)^{-1}\left(1 + \ell_2^2 s_2^2\right)^{-1}$	

2.2.3 Using the vector difference of inputs

$$\boldsymbol{x}, \boldsymbol{x}' \in \mathbb{R}^{D=2}; \;\; \boldsymbol{\tau} = \boldsymbol{x} - \boldsymbol{x}' = (x_1 - x_1', x_2 - x_2') = (r_1, r_2) \in \mathbb{R}^2; \;\; k(\boldsymbol{\tau}, \ell) = \exp\left(-\frac{\boldsymbol{\tau}}{\ell}\right)$$

$$\boldsymbol{\omega} = (s_1, s_2) \in \mathbb{R}^2; \;\; S_\nu(\boldsymbol{\omega}, \ell) = \frac{2^D \pi^{D/2} \Gamma(\nu + D/2)(2\nu)^\nu}{\Gamma(\nu)\ell^{2\nu}} \left(\frac{2\nu}{\ell^2} + \boldsymbol{\omega}^2\right)^{-(\nu + D/2)} \;\; ; \;\; S_{1/2}(\boldsymbol{\omega}, \ell) = \frac{2^D \pi^{\frac{D}{2}} \Gamma(\frac{D+1}{2})}{\sqrt{\pi}\ell} \left(\frac{1}{\ell^2} + \boldsymbol{\omega}^2\right)^{-\frac{D+1}{2}}$$

$\ell \in \mathbb{R}$	$k(\boldsymbol{\tau}, \ell) = k(\boldsymbol{x} - \boldsymbol{x}', \ell)$ $= \exp\left(-\frac{\boldsymbol{\tau}}{\ell}\right)$ $= \exp\left(-\frac{(r_1, r_2)}{\ell}\right)$ (using dot product?) $= \exp\left(-\sum_{i=1}^{2} \frac{r_i}{\ell}\right)$	$D = 2$ $S_{1/2}(\boldsymbol{\omega}, \ell) = \frac{2\pi}{\ell}\left(\frac{1}{\ell^2} + \boldsymbol{\omega}^2\right)^{-\frac{3}{2}}$ $= 2\pi\ell^2 \left(1 + \ell^2 \boldsymbol{\omega}^2\right)^{-\frac{3}{2}}$ $= 2\pi\ell^2 \left(1 + \ell^2 (s_1, s_2)^\top (s_1, s_2)\right)^{-\frac{3}{2}}$ $= 2\pi\ell^2 \left(1 + \ell^2 (s_1^2 + s_2^2)\right)^{-\frac{3}{2}}$ $= 2\pi\ell^2 \left(1 + \sum_{i=1}^{D} \ell^2 s_i^2\right)^{-\frac{3}{2}}$	-ISOTROPIC -SEPARABLE: $k(\boldsymbol{x} - \boldsymbol{x}', \boldsymbol{\ell})$ $= k(\|x_1 - x_1'\|, \ell_1) k(\|x_2 - x_2'\|, \ell_2)$ IT SHOULD BE SEPARABLE IN THE SPECTRAL DENSITY AS WELL? $S_{1/2}(\boldsymbol{\omega}, \ell)$ should be equal to $S_{1/2}(s_1, \ell) S_{1/2}(s_2, \ell)$
$\boldsymbol{\ell} \in \mathbb{R}^2$	$k(\boldsymbol{\tau}, \boldsymbol{\ell}) = \exp\left(-\frac{\boldsymbol{\tau}}{\boldsymbol{\ell}}\right)$ $= \exp\left(-\frac{(r_1, r_2)}{(\ell_1, \ell_2)}\right)$ (using dot product?) $= \exp\left(-\sum_{i=1}^{2} \frac{r_i}{\ell_i}\right)$	$S_{1/2}(\boldsymbol{\omega}, \boldsymbol{\ell}) = 2\pi\ell_1\ell_2 \left(1 + \sum_{i=1}^{D} \ell_i^2 s_i^2\right)^{-\frac{3}{2}}$	
$\boldsymbol{\ell} \in \mathbb{R}^2$ Separable kernel	$k(\|x_1 - x_1'\|, \ell_1) k(\|x_2 - x_2'\|, \ell_2)$ $= \exp\left(-\frac{r_1}{\ell_1}\right) \exp\left(-\frac{r_2}{\ell_2}\right)$ $= \exp\left(-\sum_{i=1}^{2} \frac{r_i}{\ell_i}\right)$	$S_{1/2}(s_1, \ell_1) S_{1/2}(s_2, \ell_2)$ $= \frac{2}{\ell_1}\left(\frac{1}{\ell_1^2} + s_1^2\right)^{-1} \cdot \frac{2}{\ell_2}\left(\frac{1}{\ell_2^2} + s_2^2\right)^{-1}$ $= 4\ell_1\ell_2 \left(1 + \ell_1^2 s_1^2\right)^{-1} \left(1 + \ell_2^2 s_2^2\right)^{-1}$	

2.3 Matern($\nu = 3/2$) covariance function (k) and spectral density (S)

2.3.1 Using norm-L2 (Euclidean distance)

$$\boldsymbol{x}, \boldsymbol{x}' \in \mathbb{R}^{D=2};\ \ \tau_{L2} = |\boldsymbol{x}-\boldsymbol{x}'|_{L2} = \sqrt{(\boldsymbol{x}-\boldsymbol{x}')(\boldsymbol{x}-\boldsymbol{x}')} = \sqrt{(x_1-x_1')^2+(x_2-x_2')^2} = \sqrt{r_1^2+r_2^2} \in \mathbb{R};\ \ k(\tau_{L2},\ell) = \left(1+\frac{\sqrt{3}\tau}{\ell}\right)\exp\left(-\frac{\sqrt{3}\tau}{\ell}\right)$$

$$\omega_{L2} = \sqrt{s_1^2+s_2^2} \in \mathbb{R};\ \ S_\nu(\omega_{L2}) = \frac{2^D\pi^{D/2}\Gamma(\nu+D/2)(2\nu)^\nu}{\Gamma(\nu)\ell^{2\nu}}\left(\frac{2\nu}{\ell^2}+\omega_{L2}^2\right)^{-(\nu+D/2)}\ ;\ \ S_{3/2}(\omega_{L2}) = \frac{2^D\pi^{D/2}\Gamma(\frac{D+3}{2})\sqrt{3}^3}{\frac{1}{2}\sqrt{\pi}\ell^3}\left(\frac{3}{\ell^2}+\omega_{L2}^2\right)^{-\frac{D+3}{2}}$$

$\ell \in \mathbb{R}$	$\begin{aligned} k(\tau_{L2},\ell) &= k(\lvert\boldsymbol{x}-\boldsymbol{x}'\rvert_{L2},\ell) \\ &= \left(1+\frac{\sqrt{3}\sqrt{\sum_{i=1}^2 r_i^2}}{\ell}\right)\exp\left(-\frac{\sqrt{3}\sqrt{\sum_{i=1}^2 r_i^2}}{\ell}\right) \\ &= \left(1+\sqrt{\frac{\sum_{i=1}^2 3r_i^2}{\ell^2}}\right)\exp\left(-\sqrt{\frac{\sum_{i=1}^2 3r_i^2}{\ell^2}}\right) \end{aligned}$	$\begin{aligned} D &= 2 \\ S_{3/2}(\omega_{L2},\ell) &= \frac{6\pi\sqrt{3}^3}{\ell^3}\left(\frac{3}{\ell^2}+\omega^2\right)^{-\frac{5}{2}} \\ &= 6\pi\sqrt{3}^3\ell^2\left(3+\ell^2\omega^2\right)^{-\frac{5}{2}} \\ &= 6\pi\sqrt{3}^3\ell^2\left(3+\ell^2 s_1^2+\ell^2 s_2^2\right)^{-\frac{5}{2}} \\ D &= 3 \\ S_{3/2}(\omega_{L2},\ell) &= 32\pi\sqrt{3}^3\ell^3\left(3+\ell^2\omega^2\right)^{-3} \end{aligned}$	-ISOTROPIC -NO SEPARABLE: $k(\lvert\boldsymbol{x}-\boldsymbol{x}'\rvert_{L2},\boldsymbol{\ell})$ $\neq k(\lvert x_1-x_1'\rvert,\ell_1)k(\lvert x_2-x_2'\rvert,\ell_2)$ $S_{3/2}(\omega_{L2},\ell) \neq S_{3/2}(s_1,\ell)S_{1/2}(s_2,\ell)$
$\boldsymbol{\ell} \in \mathbb{R}^2$	$\begin{aligned} k(\tau_{L2},\boldsymbol{\ell}) &= k(\lvert\boldsymbol{x}-\boldsymbol{x}'\rvert_{L2},\boldsymbol{\ell}) \\ &= \left(1+\sqrt{\frac{\sum_{i=1}^2 3r_i^2}{\ell_i^2}}\right)\exp\left(-\sqrt{\frac{\sum_{i=1}^2 3r_i^2}{\ell_i^2}}\right) \end{aligned}$	$\begin{aligned} S_{3/2}(\omega_{L2},\boldsymbol{\ell}) &= 6\pi\sqrt{3}^3\ell_1\ell_2\left(3+\ell_1^2 s_1^2+\ell_2^2 s_2^2\right)^{-\frac{5}{2}} \\ &= 6\pi\sqrt{3}^3\prod_{i=1}^D \ell_i\left(3+\sum_{i=1}^D \ell_i^2 s_i^2\right)^{-\frac{5}{2}} \end{aligned}$	
$\boldsymbol{\ell} \in \mathbb{R}^2$ Separable kernel	$\begin{aligned} & k(\lvert x_1-x_1'\rvert,\ell_1)k(\lvert x_2-x_2'\rvert,\ell_2) \\ &= \left(1+\sqrt{3}\frac{r_1}{\ell_1}\right)\exp\left(-\sqrt{3}\frac{r_1}{\ell_1}\right) \\ &\cdot\left(1+\sqrt{3}\frac{r_2}{\ell_2}\right)\exp\left(-\sqrt{3}\frac{r_2}{\ell_2}\right) \\ &= \left(1+\sqrt{\frac{3r_1^2}{\ell_1^2}}\right)\left(1+\sqrt{\frac{3r_2^2}{\ell_2^2}}\right)\exp\left(-\sum_{i=1}^2\sqrt{\frac{3r_i^2}{\ell_i^2}}\right) \\ &= \left(1+\sum_{i=1}^2\sqrt{\frac{3r_i^2}{\ell_i^2}}+\frac{3r_1^2 r_2^2}{\ell_1^2\ell_2^2}\right)\exp\left(-\sum_{i=1}^2\sqrt{\frac{3r_i^2}{\ell_i^2}}\right) \end{aligned}$	$\begin{aligned} & S_{3/2}(s_1,\ell_1)S_{3/2}(s_2,\ell_2) \\ &= \frac{4\sqrt{3}^3}{\ell_1^3}\left(\frac{3}{\ell_1^2}+s_1^2\right)^{-3}\cdot\frac{4\sqrt{3}^3}{\ell_2^3}\left(\frac{3}{\ell_2^2}+s_2^2\right)^{-3} \\ &= 4^2\sqrt{3}^6\boldsymbol{\ell}^\top\boldsymbol{\ell}\left(3+\ell_1^2 s_1^2\right)^{-2}\left(3+\ell_2^2 s_2^2\right)^{-2} \end{aligned}$	

3 Brief details about computational demands and model inference

This paper focus on the basis function approximation via Laplace eigenfunctions for stationary covariance functions proposed by **?**. This method has an attractive computational cost as this basically turns the regular GP model into a linear model.

- The design matrix of the proposed linear model, which is composed of a basis of Laplace eigenfunctions, can be computed analytically and does not depend on the hyperparameters of the model, then it has to be computed only once with $O(n+m)$ computational demands.
- In learning the covariance function hyperparameters the proposed approximate GP model has a computational complexity of $O(nm+m)$ in every step of the optimizer. This demand includes:
  - The linear model is computed with complexity $O(nm)$, computed in every step of the optimizer.
  - The weights associated to the basis functions in this linear model is a m-dimensional vector (m is the number of basis functions) and their computation is an operation with $O(m)$ computational demands. The weights depend on the hyperparameters, then they have to be computed in every step of the optimizer.
- The computation of the automatic differentiation to compute the gradients in this linear model scales $O(n)$? which is related to the computational complexity, an operation that must be computed in every step of the optimizer.
- Using maximizing marginal likelihood methods, the proposed model has a overall complexity of $O(nm^2)$. After this, evaluating the marginal likelihood and marginal likelihood gradients is an $O(m^3)$ operation in every step of the optimizer.
- The parameter posterior distribution in this approximate GP model is m-dimensional ($m << n$) which helps the use of GP priors as latent functions, especially when sampling methods for inference are used. GP prior as latent functions is needed in generalized models.
- In regular GPs and other approximate GP models and Splines models these features do not have so nice properties:
  - In regular GPs, the automatic differentiation to compute the gradients of the covariance function scales $O(n^2)$, the dimension of the covariance matrix, and the full inversion of the covariance matrix scales $O(n^3)$. These two operations have to be computed at every step of the HMC or optimizer.
  - In regular GPs, the parameter posterior distributions is N-dimensional. It is known that when N is of medium or large size there is high correlation between the N-dimensional latent function and the hyperparameters of the GP prior.
  - In conventional sparse GP approximations, based on inducing points, although the rank of the GP is reduced considerably to the number of inducing points, this still needs to do the autodiff and covariance matrix inversion.

- The Splines models are also a sort of basis functions expansion model, then the computational demands are similar to that in this approach. However in Splines models the lengthscale hyperparameter tend to be fixed and then the fit is covered by the magnitude parameter. In that sense, Splines models tend to loose the useful interpretation of the lengthscale parameter.

4 Brief details about computational demands and model inference

This study focus on the basis function approximation via Laplace eigenfunctions for stationary covariance functions proposed by **?**. This method has an attractive computational cost as this basically turns the regular GP model into a linear model.

- The design matrix of the proposed linear model, which is composed of a basis of Laplace eigenfunctions, can be computed analytically and does not depend on the hyperparameters of the model, then it has to be computed only once with $O(n \cdot m)$ computational demands.
- In learning the covariance function hyperparameters the proposed approximate GP model has a computational complexity of $O(nm + m)$ in every step of the optimizer. This demand includes:
  - The linear model is computed with complexity $O(nm)$, computed in every step of the optimizer.
  - The weights associated to the basis functions in this linear model is a m-dimensional vector (m is the number of basis functions) and their computation is an operation with $O(m)$ computational demands. The weights depend on the hyperparameters, then they have to be computed in every step of the optimizer.
- The computation of the automatic differentiation to compute the gradients in this linear model scales ¿$O(n)$? which is related to the computational complexity, an operation that must be computed in every step of the optimizer.
- Memory requirements of automatic differentation are reduced because this rather scales with the computational complexity instead of with the usual memory requirements for the posterior density computation.
- Using maximizing marginal likelihood methods, the proposed model has a overall complexity of $O(nm^2)$. After this, evaluating the marginal likelihood and marginal likelihood gradients is an $O(m^3)$ operation in every step of the optimizer.
- The parameter posterior distribution in this approximate GP model is m-dimensional ($m \ll n$) which helps the use of GP priors as latent functions, especially when sampling methods for inference are used. GP prior as latent functions is needed in generalized models.
- In regular GPs and other approximate GP models and spline models these features do not have so nice properties:
  - In regular GPs, the automatic differentiation to compute the gradients of the covariance function scales $O(n^2)$, the dimension of the covariance matrix, and the full inversion of the covariance matrix scales $O(n^3)$. These two operations have to be computed at every step of the HMC or optimizer.

- In regular GPs, the parameter posterior distributions is N-dimensional. It is known that when N is of medium or large size there is high correlation between the N-dimensional latent function and the hyperparameters of the GP prior.
- In conventional sparse GP approximations, based on inducing points, although the rank of the GP is reduced considerably to the number of inducing points, this still needs to do the autodiff and covariance matrix inversion.
- While Sparse Spectrum GP is based on a sparse spectrum, the reduced-rank method proposed in this study aims to make the spectrum as 'full' as possible at a given rank.
- The spline models are also a sort of basis functions expansion model, then the computational demands are similar to that in this approach. However in spline models the lengthscale hyperparameter tend to be fixed and then the fit is covered by the magnitude parameter. In that sense, spline models tend to loose the useful interpretation of the lengthscale parameter.
- Recent spline models can reproduce the Matern family of covariance functions (see, e.g., **?**), however our approach can reproduce basically all of the stationary covariance functions.

5 Contributions of the work

As said above the proposed method was already developed by **?** where they fully develop, describe and generalize the methodology. Though, they do not put much effort in describing and analyzing the relation among the key factors of the box size (or boundary condition), the number of basis functions, and the smoothness or roughness of the function. The performance and accuracy of the method are directly related with the number of basis functions and the box size. At the same time, successful values for these two factors depend on the smoothness or roughness of the process to be modeled. The time of computation is mainly dependent on the number of basis functions. Our main contributions to this recently developed methodology for low-rank GP model by **?** goes around these aspects.

• Firstly, clear summarized formulae of the method for the univariate and multivariate cases is presented.

• We investigate the relations going on among these factors, the number of basis functions, the box size, and the lengthscale of the functions.

• We make recommendations for the values of these factors based on the recognized relations among them. We provide useful graphs of these relations that will help the users to improve performance and save time of computation.

• We also diagnose if the chosen values for the number of basis functions and the box size are adequate to fit to the actual data.

• We describe the generalization of the method to the multidimensional case.

• We implement the approach in a fully probabilistic framework and for the Stan programming probabilistic software.

• We show several illustrative examples, simulate and real datasets, of the performance of the model, and accompanied by their Stan codes.