Lecture 21: Gaussian process regression

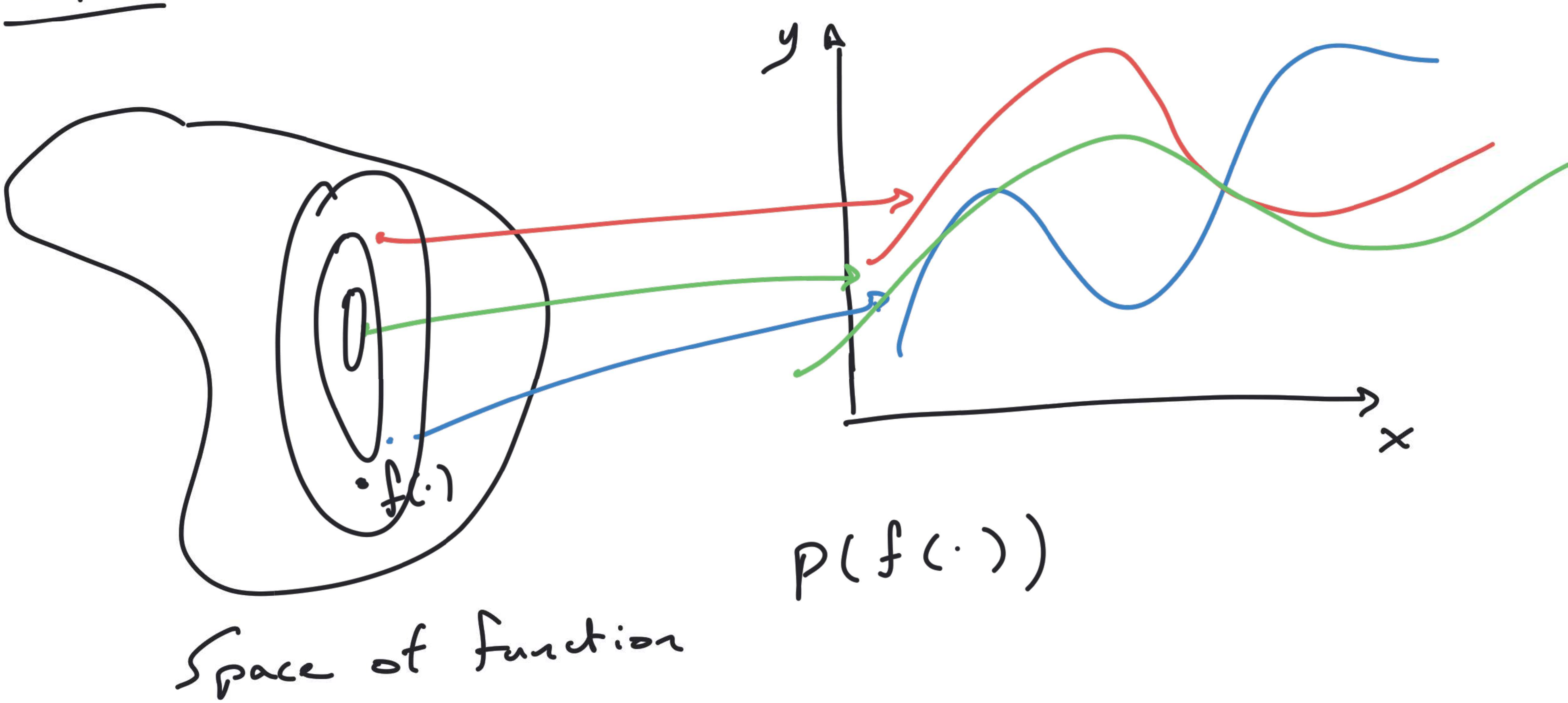
Professor Ilias Bilionis

Priors on function spaces

PREDICTIVE
SCIENCE LABORATORY

Probability measure on a function space

Inputs: $x \in \mathbb{R}$; Output: $y \in \mathbb{R}$

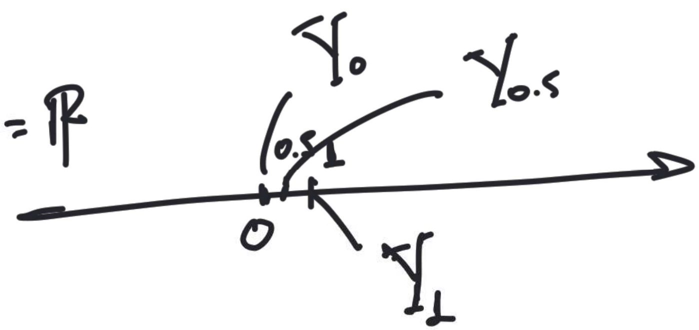


What is a stochastic process?

$\mathcal{X}$: set of inputs

Stochastic process on $\mathcal{X}$ is a collection of random variables Y_x, $x \in \mathcal{X}$

$\mathcal{X} = \mathbb{R}$

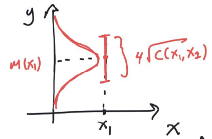


The Gaussian process prior

$f(\cdot) \sim GP(m(\cdot), c(\cdot,\cdot))$

(random function)

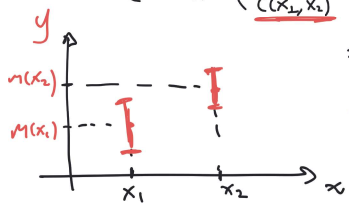
$\underline{\underline{n=1}}$ x_1, $f(\cdot)$ r.f. $\Rightarrow$ $f(x_1)$ r.v. $\xRightarrow[\text{of GP}]{\text{By def.}}$ $f(x_1) \sim N(m(x_1), c(x_1,x_1))$

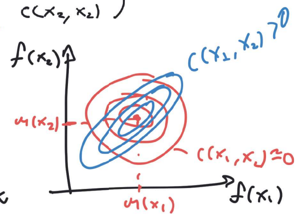


$\underline{\underline{n=2}}$ x_1, x_2, $f(\cdot)$ r.f. $\Rightarrow$ $f_{1:2} = (f(x_1), f(x_2))$ random vector

$\xRightarrow[\text{of GP}]{\text{By def.}}$ $f_{1:2} \sim N(m_{1:2}, C_2)$

$m_{1:2} = \begin{pmatrix} m(x_1) \\ m(x_2) \end{pmatrix}$, $C_2 = \begin{pmatrix} c(x_1,x_1) & \underline{c(x_1,x_2)} \\ \underline{c(x_2,x_1)} & c(x_2,x_2) \end{pmatrix}$





$\underline{n}$

$x_{1:n} = (x_1, x_2, \dots, x_n)$

$f_{1:n} = (f(x_1), f(x_2), \dots, f(x_n))$ random vector

$$f_{1:n} \sim N(m_{1:n}, C_n)$$

$$m_{1:n} = \begin{pmatrix} m(x_1) \\ \vdots \\ m(x_n) \end{pmatrix}, \quad C_n = \begin{pmatrix} c(x_1,x_1) & \dots & c(x_1,x_n) \\ & \ddots & \vdots \\ & & c(x_n,x_n) \end{pmatrix} = \left(c(x_i,x_j)\right)_{i,j=1}^{n}$$

Stochastic Process > Kolmogorov Extension Theorem.

The mean function

$$f(\cdot) \sim GP(\boxed{m(\cdot)}, c(\cdot, \cdot))$$

$$f(x) \sim N(m(x), c(x,x))$$

$$\mathbb{E}[f(x)] = m(x)$$

what you expect $f(x)$ to be before you see any data.

PREDICTIVE
SCIENCE LABORATORY

The covariance function

$$f(\cdot) \sim GP(m(\cdot), \boxed{c(\cdot,\cdot)})$$

$$f(x) \sim N(m(x), c(x,x))$$

$$\downarrow$$

$$V[f(x)] = c(x,x)$$

$$(f(x_1), f(x_2)) \sim N\left(\begin{pmatrix} m(x_1) \\ m(x_2) \end{pmatrix}, \begin{pmatrix} c(x_1,x_1) & c(x_1,x_2) \\ c(x_1,x_2) & c(x_2,x_2) \end{pmatrix}\right)$$

$$C[f(x_1), f(x_2)] = c(x_1, x_2)$$

How correlated the function values at x_1 and x_2 are?

$c(x_1,x_2) \uparrow$ $f(x_1)$ and $f(x_2)$ are more correlated.

$c(x_1,x_2) \downarrow$ $f(x_1)$ and $f(x_2)$ are less correlated.

PREDICTIVE SCIENCE LABORATORY