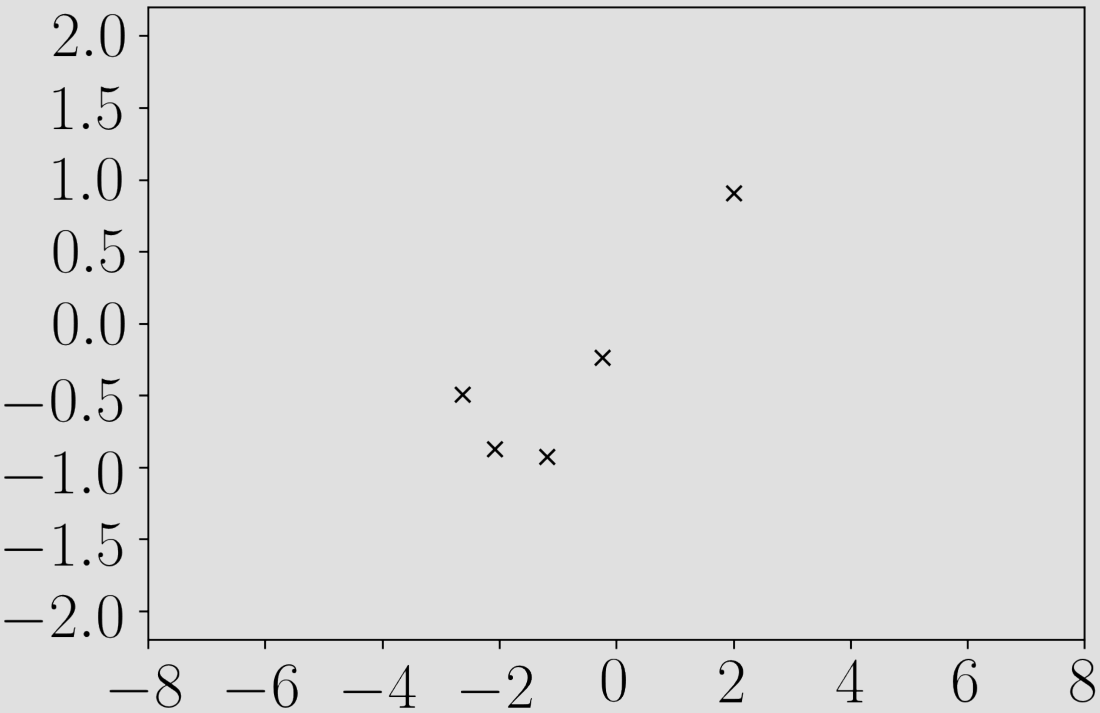
GP for machine learning

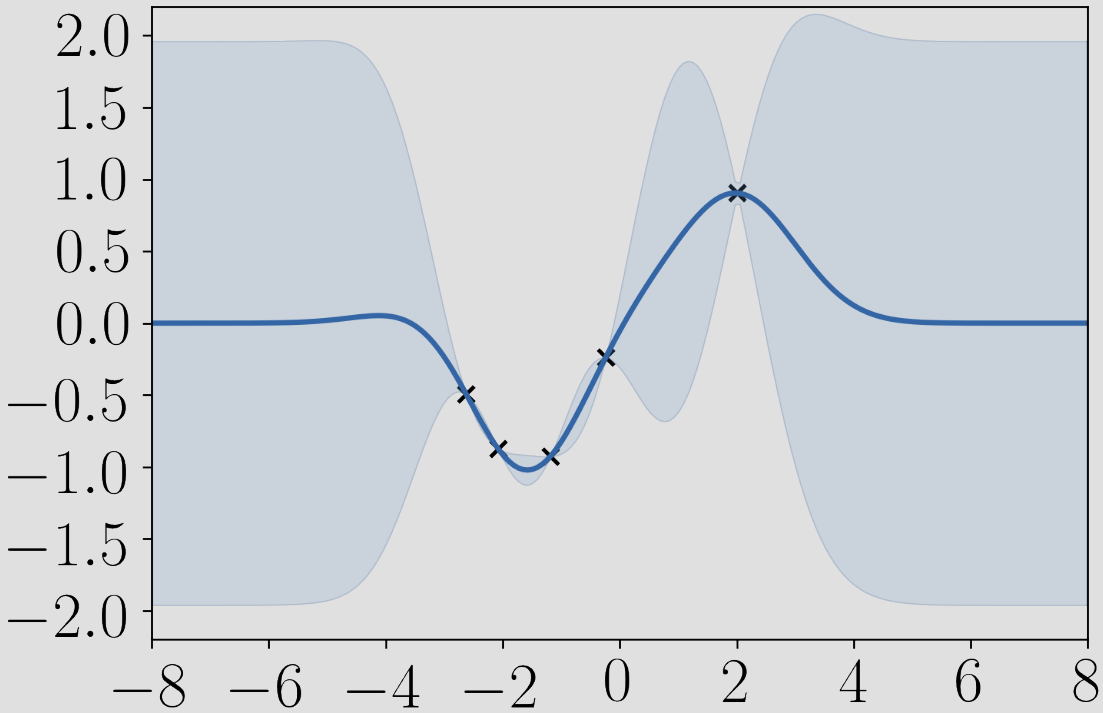
Task

$$\begin{pmatrix} f(x_1) \\ f(x_2) \\ \dots \\ f(x_n) \end{pmatrix} \Rightarrow f(x) = ?$$



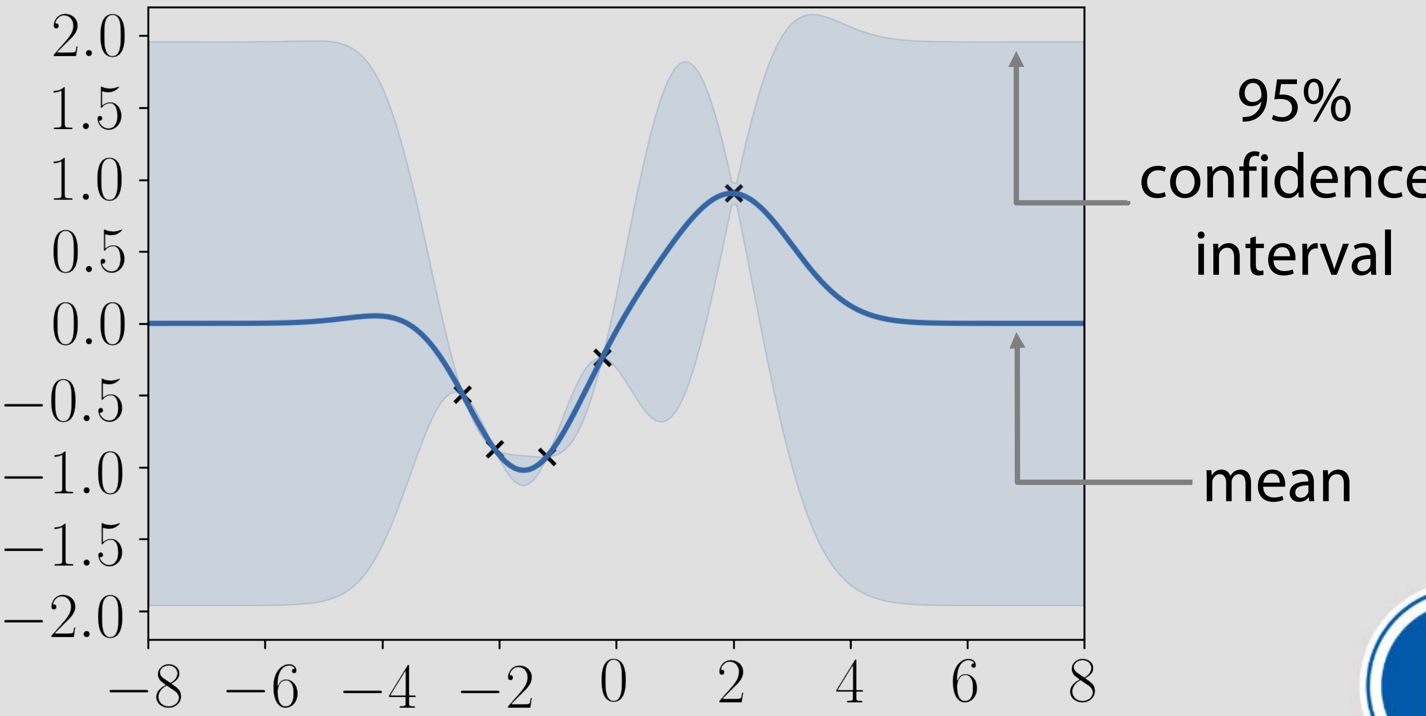
Task

$$\begin{pmatrix} f(x_1) \\ f(x_2) \\ \dots \\ f(x_n) \end{pmatrix} \Rightarrow \begin{array}{l} \cancel{f(x) = ?} \\ p(f(x)|f(x_1), \dots, f(x_n)) = ? \end{array}$$



Task

$$\begin{pmatrix} f(x_1) \\ f(x_2) \\ \dots \\ f(x_n) \end{pmatrix} \Rightarrow \begin{array}{l} \cancel{f(x) = ?} \\ p(f(x) | f(x_1), \dots, f(x_n)) = ? \end{array}$$



Prediction

$f(x)$ is a Gaussian Process with stationary prior, $m(x) = 0$

$$p(f(x)|f(x_1), \ldots, f(x_n)) = \frac{p(f(x), f(x_1), \ldots, f(x_n))}{p(f(x_1), \ldots, f(x_n))}$$

Prediction

$f(x)$ is a Gaussian Process with stationary prior, $m(x) = 0$

$$p(f(x)|f(x_1), \dots, f(x_n)) = \frac{p(f(x), f(x_1), \dots, f(x_n))}{p(f(x_1), \dots, f(x_n))}$$

$$= \frac{\mathcal{N}(f(x), f(x_1), \dots, f(x_n)|0, \widetilde{C})}{\mathcal{N}(f(x_1), \dots, f(x_n)|0, C)}$$

Prediction

$f(x)$ is a Gaussian Process with stationary prior, $m(x) = 0$

$$p(f(x)|f(x_1),\ldots,f(x_n)) = \frac{p(f(x),f(x_1),\ldots,f(x_n))}{p(f(x_1),\ldots,f(x_n))}$$

$$= \frac{\mathcal{N}(f(x),f(x_1),\ldots,f(x_n)|0,\widetilde{C})}{\mathcal{N}(f(x_1),\ldots,f(x_n)|0,C)}$$

$$C = \begin{pmatrix} K(0) & K(x_1-x_2) & K(x_1-x_3) & \ldots & K(x_1-x_n) \\ K(x_2-x_1) & K(0) & K(x_2-x_3) & \ldots & K(x_2-x_n) \\ \ldots & \ldots & \ldots & & \ldots \\ K(x_n-x_1) & K(x_n-x_2) & K(x_n-x_3) & \ldots & K(0) \end{pmatrix}$$

Prediction

$f(x)$ is a Gaussian Process with stationary prior, $m(x) = 0$

$$p(f(x)|f(x_1),\ldots,f(x_n)) = \frac{p(f(x), f(x_1),\ldots,f(x_n))}{p(f(x_1),\ldots,f(x_n))}$$

$$= \frac{\mathcal{N}(f(x), f(x_1),\ldots,f(x_n)|0,\widetilde{C})}{\mathcal{N}(f(x_1),\ldots,f(x_n)|0,C)}$$

$$\widetilde{C} = \begin{pmatrix} K(0) & k^T \\ k & C \end{pmatrix}$$

$$k = \begin{pmatrix} K(x-x_1) \\ K(x-x_2) \\ \ldots \\ K(x-x_n) \end{pmatrix}$$

Prediction

$f(x)$ is a Gaussian Process with stationary prior, $m(x) = 0$

$$p(f(x)|f(x_1),\dots,f(x_n)) = \frac{p(f(x),f(x_1),\dots,f(x_n))}{p(f(x_1),\dots,f(x_n))}$$

$$= \frac{\mathcal{N}(f(x),f(x_1),\dots,f(x_n)|0,\widetilde{C})}{\mathcal{N}(f(x_1),\dots,f(x_n)|0,C)}$$

$$= \mathcal{N}(f(x)|\mu,\sigma^2)$$

Prediction

$f(x)$ is a Gaussian Process with stationary prior, $m(x) = 0$

$$p(f(x)|f(x_1), \dots, f(x_n)) = \frac{p(f(x), f(x_1), \dots, f(x_n))}{p(f(x_1), \dots, f(x_n))}$$

$$= \frac{\mathcal{N}(f(x), f(x_1), \dots, f(x_n)|0, \widetilde{C})}{\mathcal{N}(f(x_1), \dots, f(x_n)|0, C)}$$

$$= \mathcal{N}(f(x)|\mu, \sigma^2)$$

$$\boxed{\begin{aligned} \mu &= k^T C^{-1} f \\ \sigma^2 &= K(0) - k^T C^{-1} k \end{aligned}}$$

Prediction

$f(x)$ is a Gaussian Process with stationary prior, $m(x) = 0$

$$p(f(x)|f(x_1), \dots, f(x_n)) = \frac{p(f(x), f(x_1), \dots, f(x_n))}{p(f(x_1), \dots, f(x_n))}$$

$$= \frac{\mathcal{N}(f(x), f(x_1), \dots, f(x_n)|0, \widetilde{C})}{\mathcal{N}(f(x_1), \dots, f(x_n)|0, C)}$$

$$= \mathcal{N}(f(x)|\mu, \sigma^2)$$

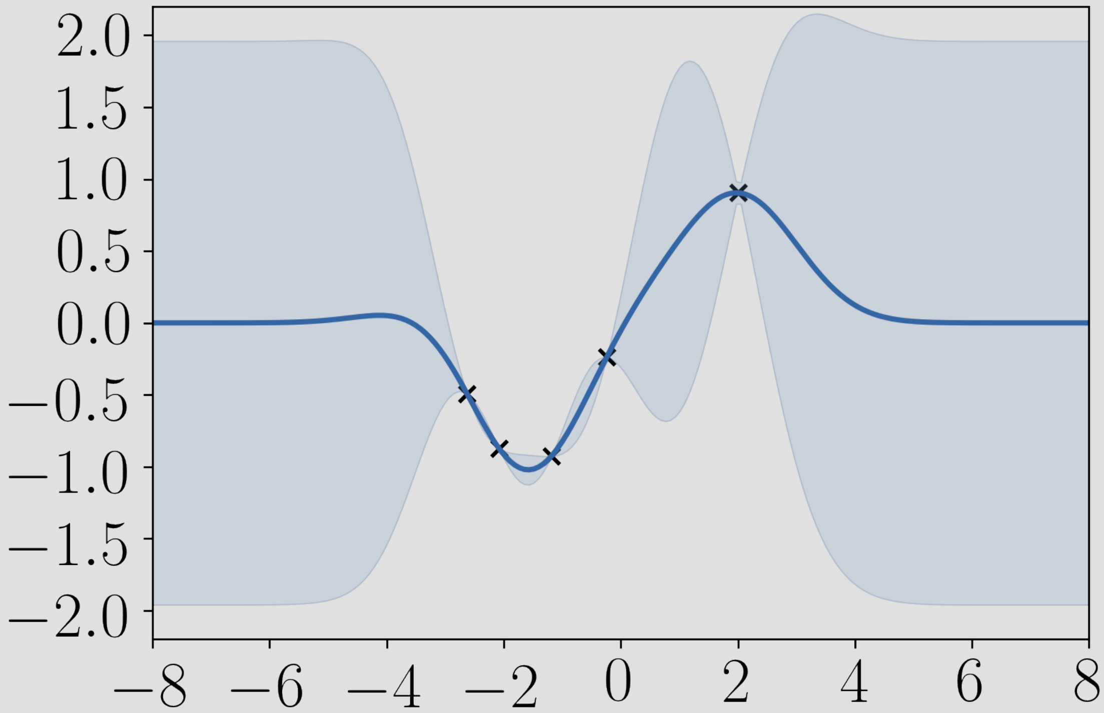
$$\boxed{\begin{aligned} \mu &= k^T C^{-1} f \\ \sigma^2 &= K(0) - k^T C^{-1} k \end{aligned}}$$

$$f = \begin{pmatrix} f(x_1) \\ f(x_2) \\ \dots \\ f(x_n) \end{pmatrix}$$

Prediction

$f(x)$ is a Gaussian Process with stationary prior, $m(x) = 0$

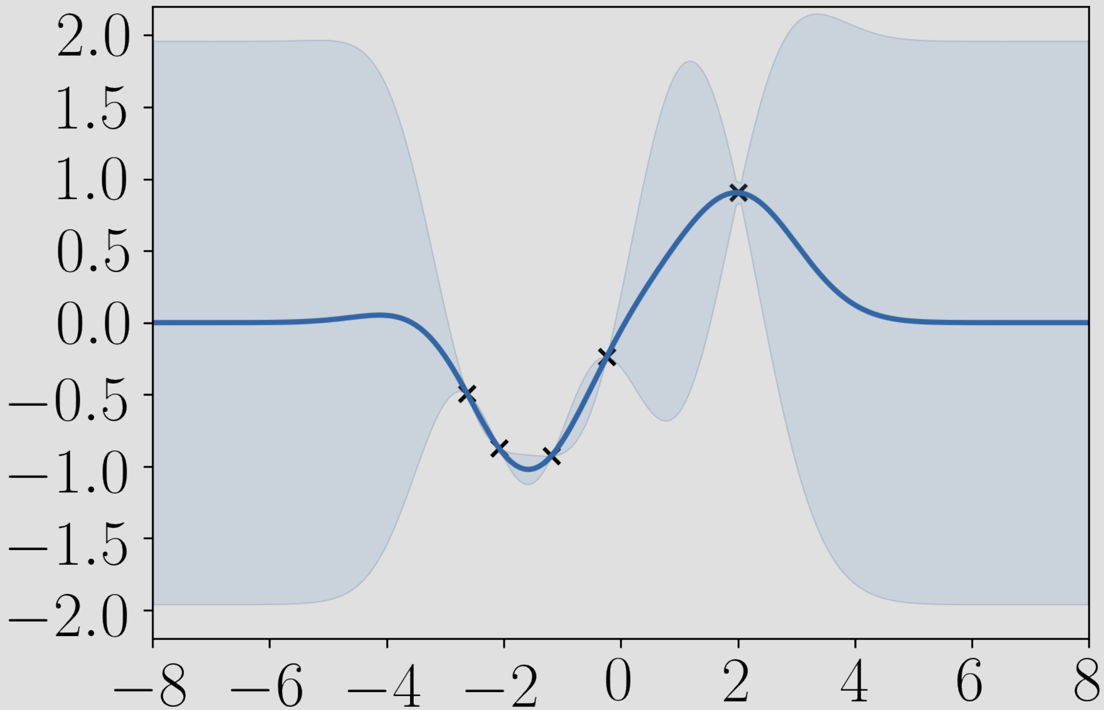
$$p(f(x)|f(x_1), \ldots, f(x_n)) = \frac{p(f(x), f(x_1), \ldots, f(x_n))}{p(f(x_1), \ldots, f(x_n))}$$



ТЕХНИЧЕСКИ СЛАЙД

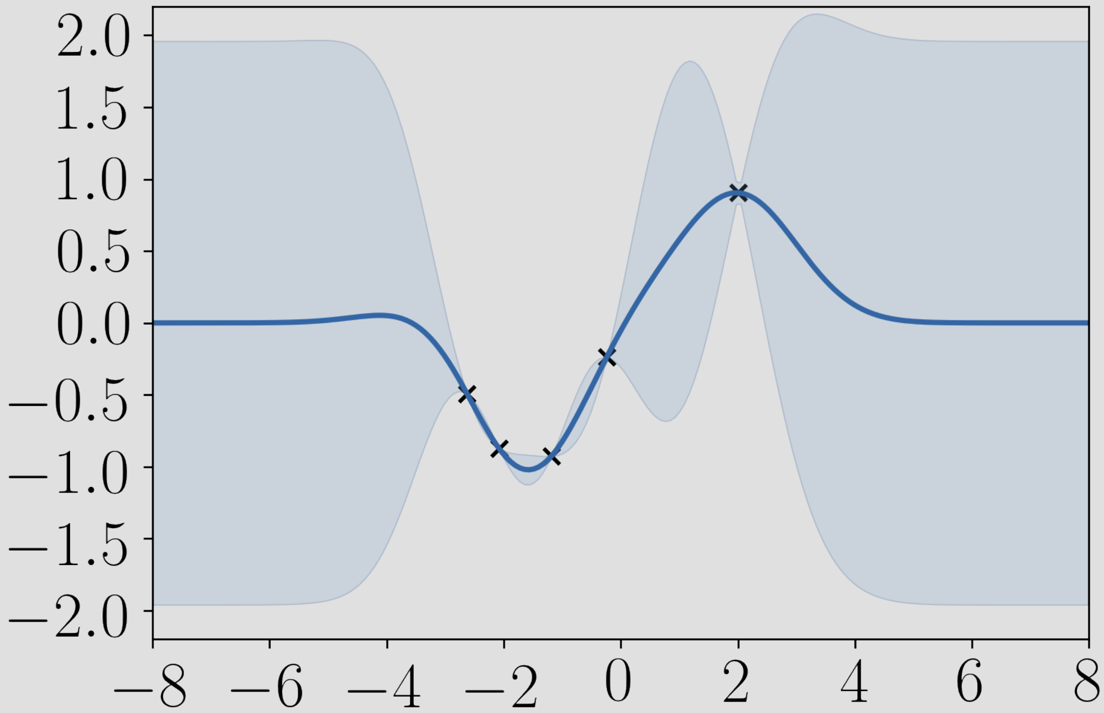
VAR = 0 @ points
GROW when moving away

$$p(f(x)|f(x_1),\dots,f(x_n)) = \frac{p(f(x),f(x_1),\dots,f(x_n))}{p(f(x_1),\dots,f(x_n))}$$



Preprocessing

- Far from data: $\mathbb{E}f(x) = 0, \ \mathrm{Var}[f(x)] = K(0)$
- Remove trend, seasonality
- Subtract mean and normalize



ТЕХНИЧЕСКИЙ СЛАЙД

POSTERIOR
IS NOT STATIONARY

REMEMBER TO INVERT

- Far from data: $\mathbb{E}f(x) = 0,\ \mathrm{Var}[f(x)] = K(0)$
- Remove trend, seasonality
- Subtract mean and normalize