Friedrich-Alexander-Universität
Technische Fakultät

Gaussian Process Regression

Richard Dirauf, M.Sc.
Machine Learning and Data Analytics (MaD) Lab
Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)
MLTS Exercise, 24.11.2022

Short Recap - Bayesian Linear Regression

MLTS Exercise 04

Remember:

BLR provides a probabilistic way to find a distribution of parameters for our models.

Advantages

- Uncertainty
- Priors

Disadvantages

- Linear with respect to the weights
- Limits expressivity
- Uncertainty independent from “observations density”

Let's take a look at stochastic processes

→ **Random Walk**

- Start somewhere at position **x**
- Measure distance **d** from **x** at point **t** in time

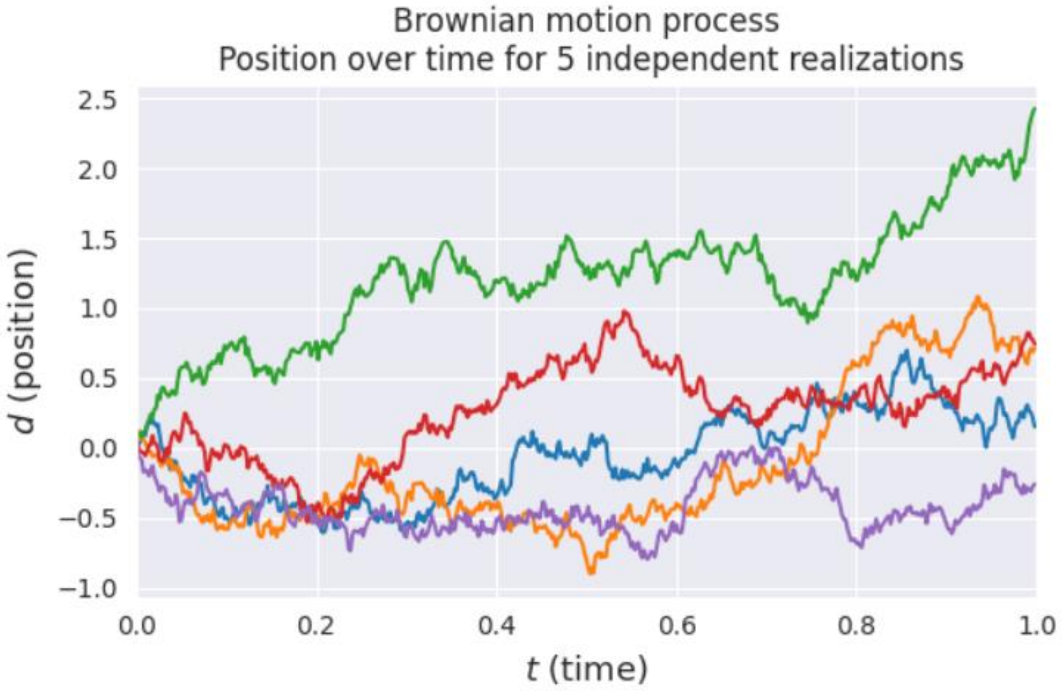
What is a Gaussian Process?

MLTS Exercise 04

Let's take a look at stochastic processes

→ Random Walk

- Start somewhere at position **x**
- Measure distance **d** from **x** at point **t** in time



What is a Gaussian Process?

MLTS Exercise 04

Gaussian Processes are distribution over functions

Defined by:

- Mean function m(x)
- Positive definite covariance function k(x, x’)

$$f(x) \sim \mathcal{GP}(m(x), k(x, x'))$$

What is a Gaussian Process?

MLTS Exercise 04

Gaussian Processes are distribution over functions

Defined by:

- Mean function m(x)
- Positive definite covariance function k(x, x')

$$f(x) \sim \mathcal{GP}(m(x), k(x, x'))$$

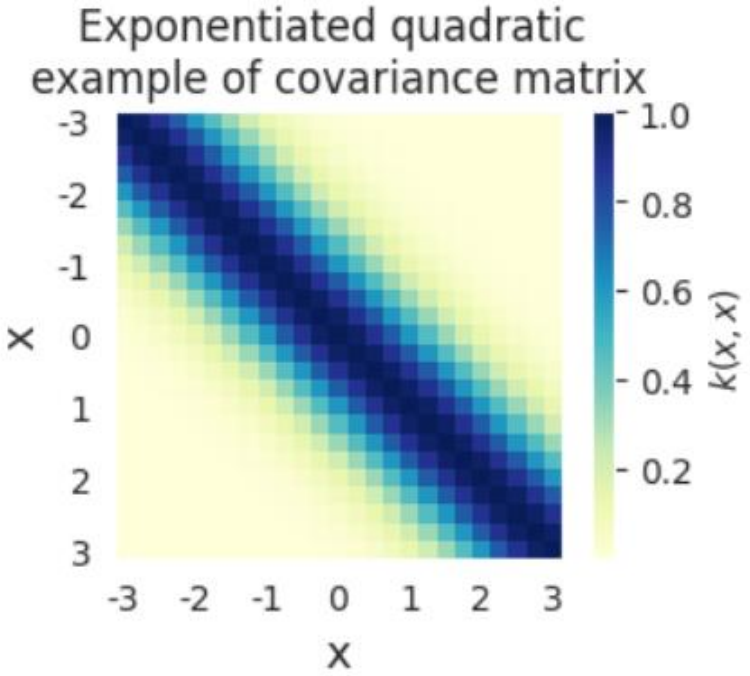
?

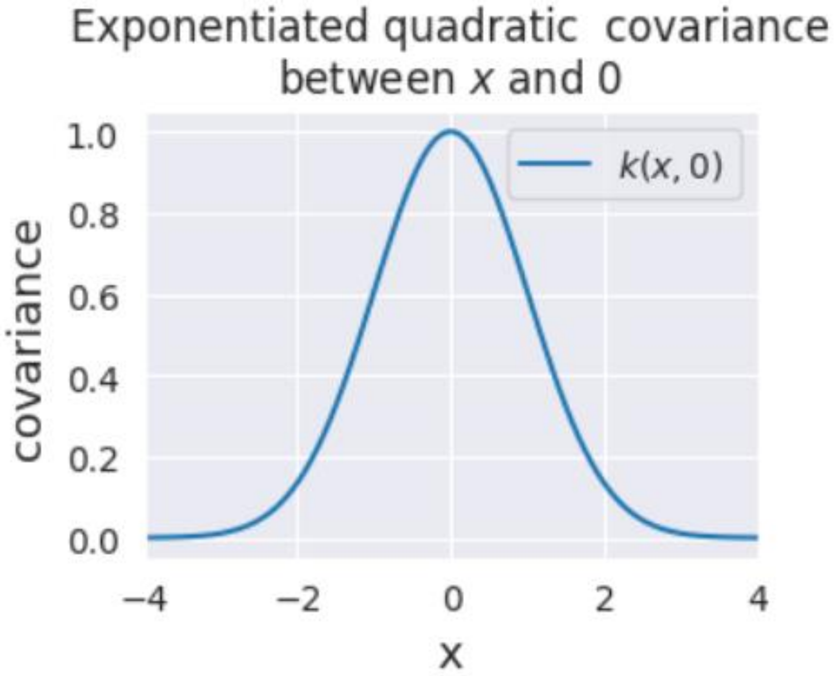
Covariance Function as a Prior

MLTS Exercise 04

Sampling from a GPR → Define m(x) and k(x, x'):

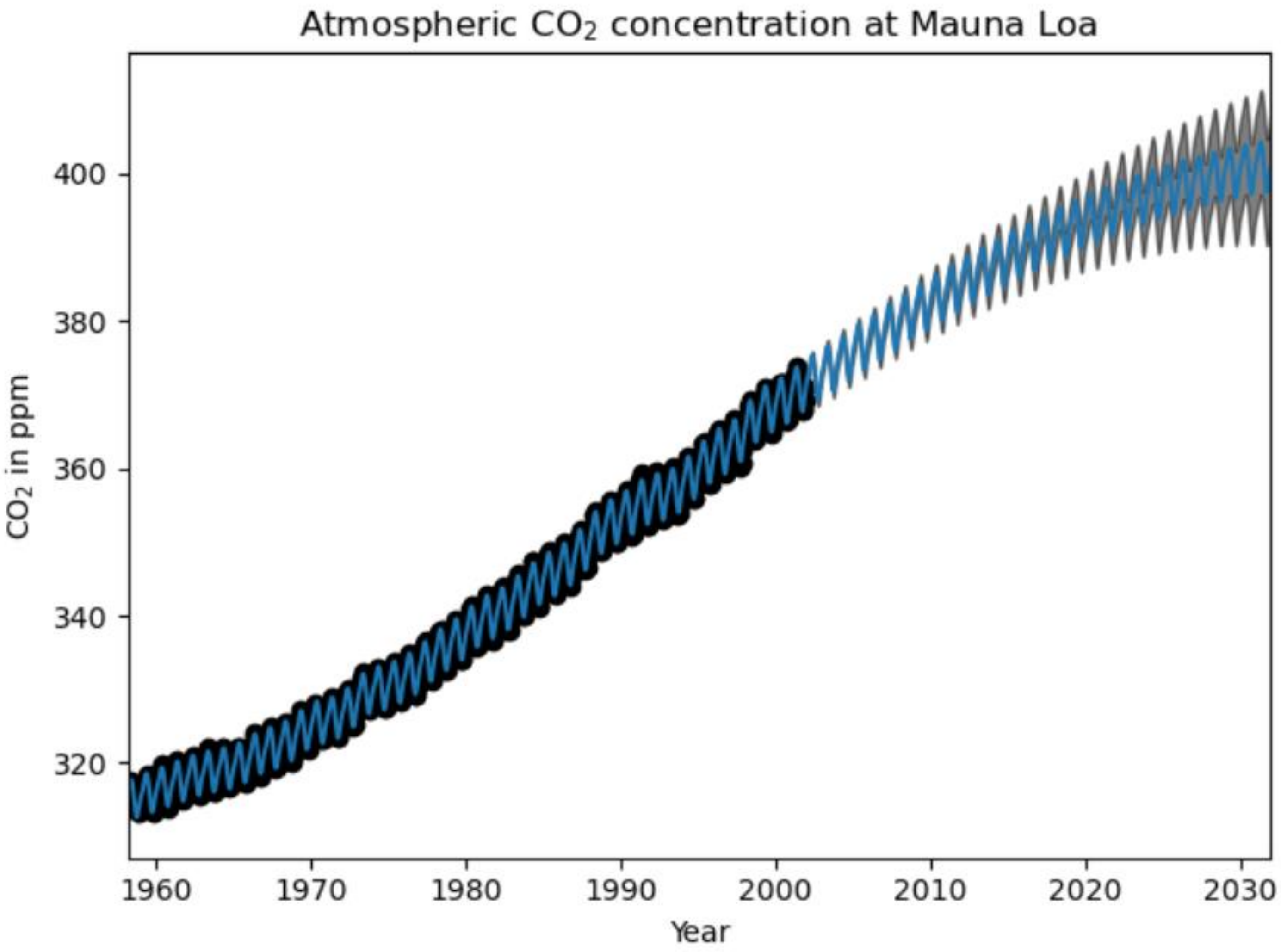
$$k(x_a, x_b) = \exp\left(-\frac{1}{2\sigma^2}\|x_a - x_b\|^2\right)$$





Suitable Kernel Functions?

MLTS Exercise 04



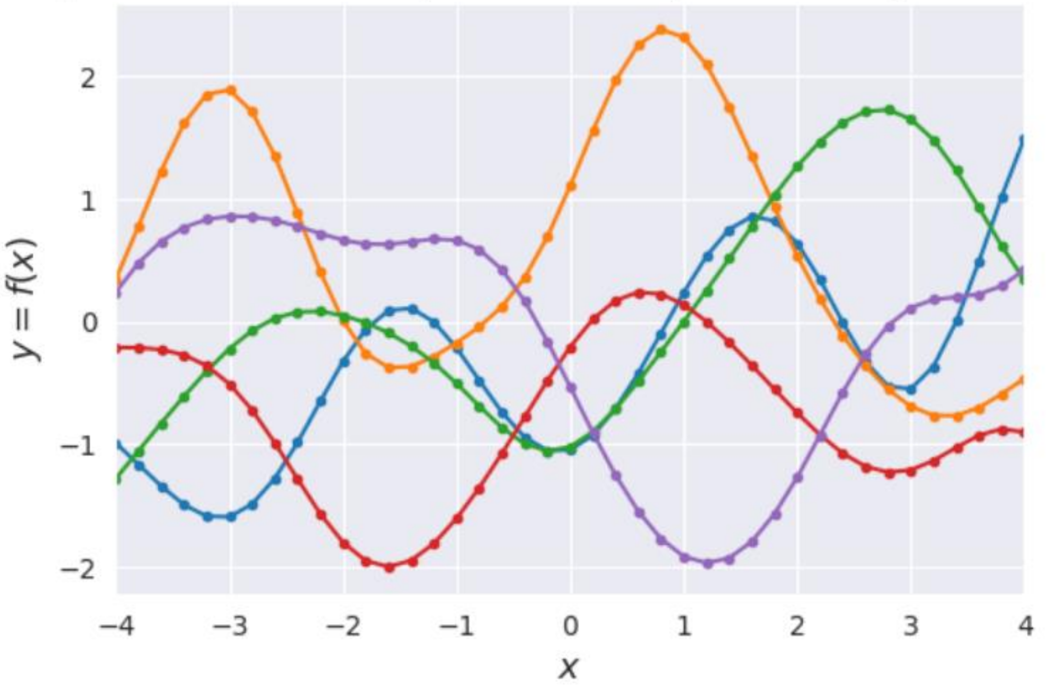
Sampling from the prior

MLTS Exercise 04

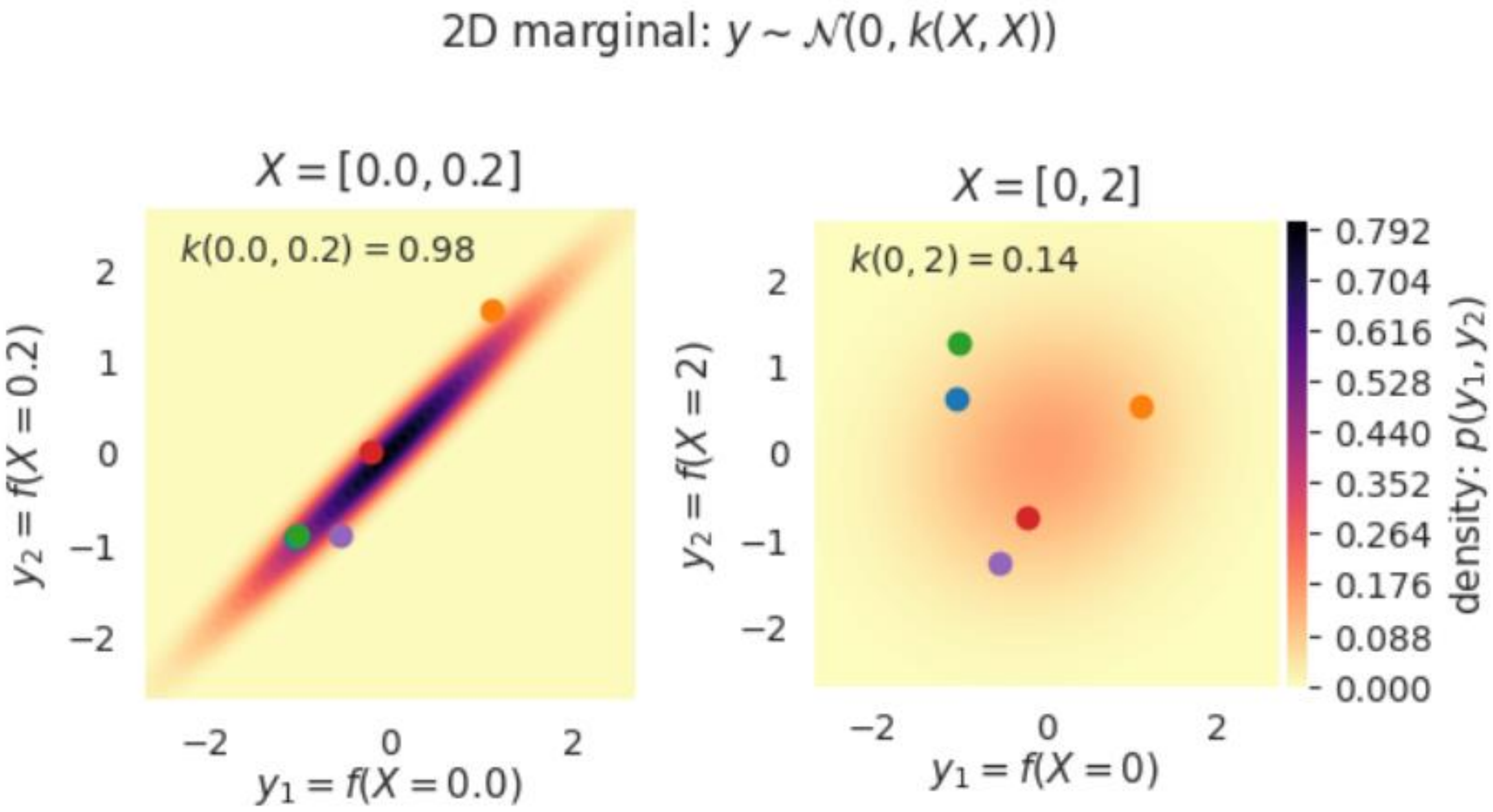
Arbitrary set of points X: $X\colon \mathbf{y} = f(X)$

Sampling from a Gaussian Distribution: $\mathcal{N}(0, k(X, X))$

5 different function realizations at 41 points
sampled from a Gaussian process with exponentiated quadratic kernel

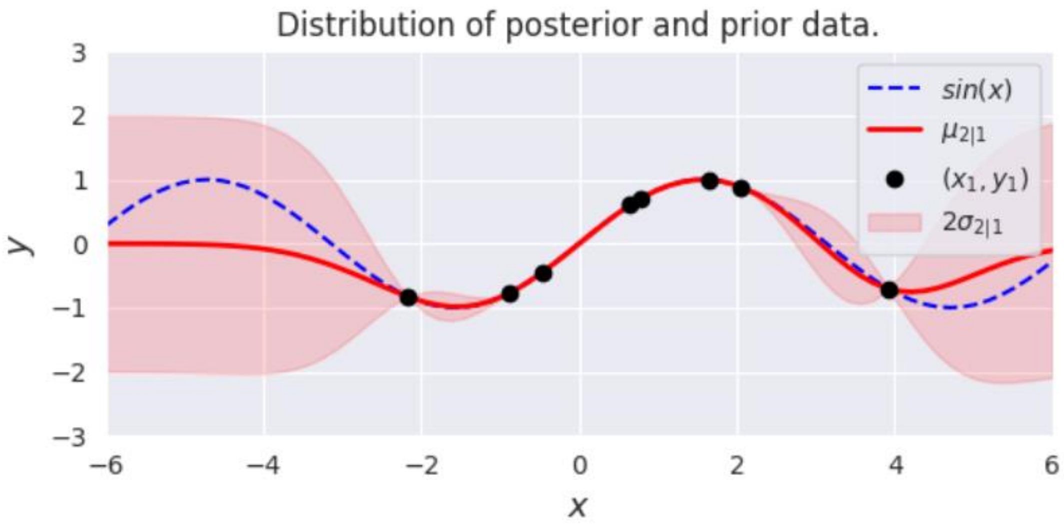


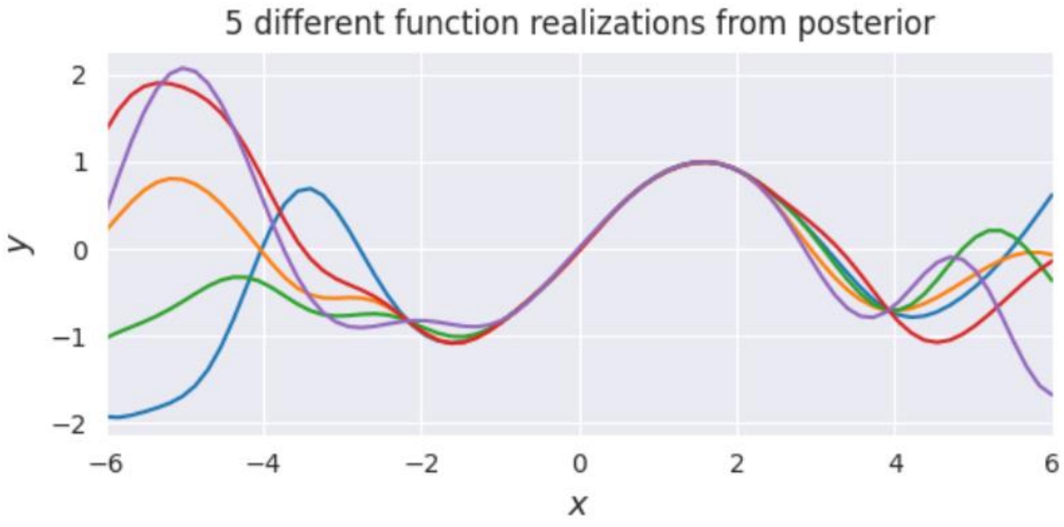
How does the covariance look like?



Bayes Rule:

$$p(\theta|\mathcal{D}) = \frac{p(\mathcal{D}|\theta)p(\theta)}{p(\mathcal{D})}$$





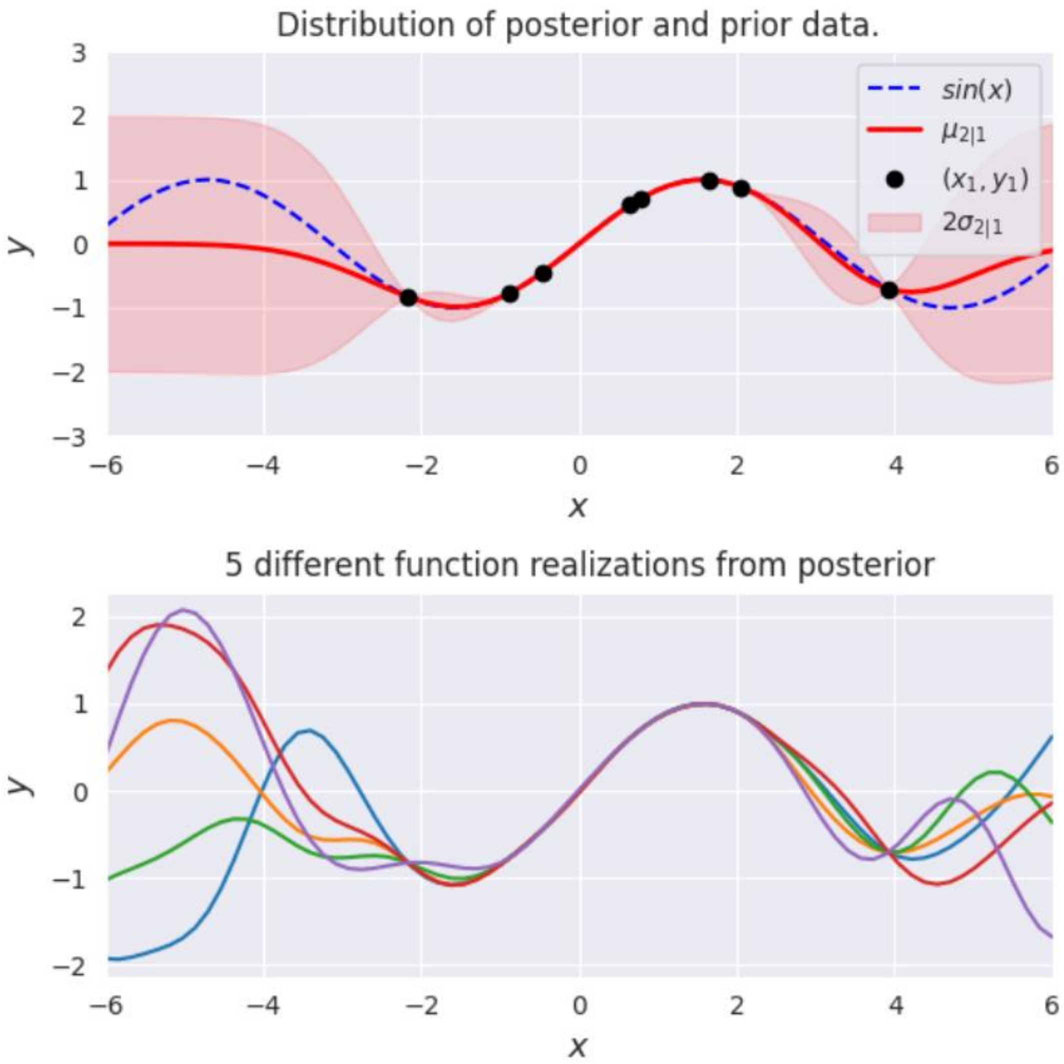
GP Noisy Observations

MLTS Exercise 04

Noisy kernel function:

$$\Sigma_{11} = k(X_1, X_1) + \sigma_\epsilon^2 I$$

?

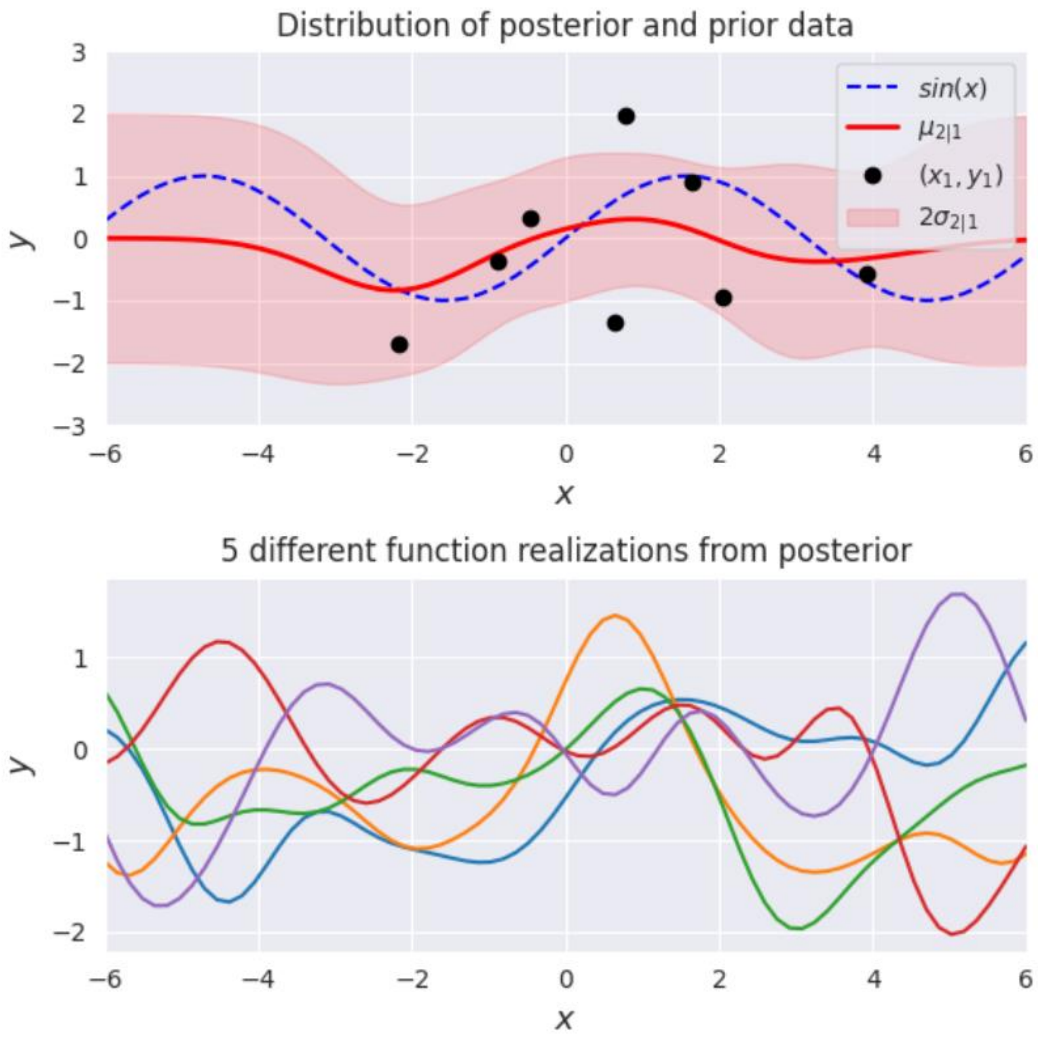


GP Noisy Observations

MLTS Exercise 04

Noisy kernel function:

$$\Sigma_{11} = k(X_1, X_1) + \sigma_\epsilon^2 I$$



Announcement

MLTS Exercise 04

Exercise 5: 08.12.2023

No Exercise next Week (01.12.2023)!!!

Any questions?

References

MLTS Exercise 04

https://katbailey.github.io/post/gaussian-processes-for-dummies/

https://peterroelants.github.io/posts/gaussian-process-tutorial/

Practice Questions

MLTS Exercise 04

- **What is a prior in the context of GPR?**
  - The probability of some event
  - The mean function
  - The covariance function
- **What are suitable kernel functions for this problem?**
  - Linear
  - Exponential
  - Sinusoidal
  - Quadratic

Practice Questions

MLTS Exercise 04

- **How can you see that there are no observations after ~2006?**
  - The black dots you plotted
  - The variance is increasing
  - 2030 is in the future and its unlikely that someone used a time-machine
  - The curve gets flatter
- **What will you expect happens if we introduce some noise to the kernel function?**
  - The uncertainty will increase for all points
  - The uncertainty will only increase where we have no observations
  - It depends on the data