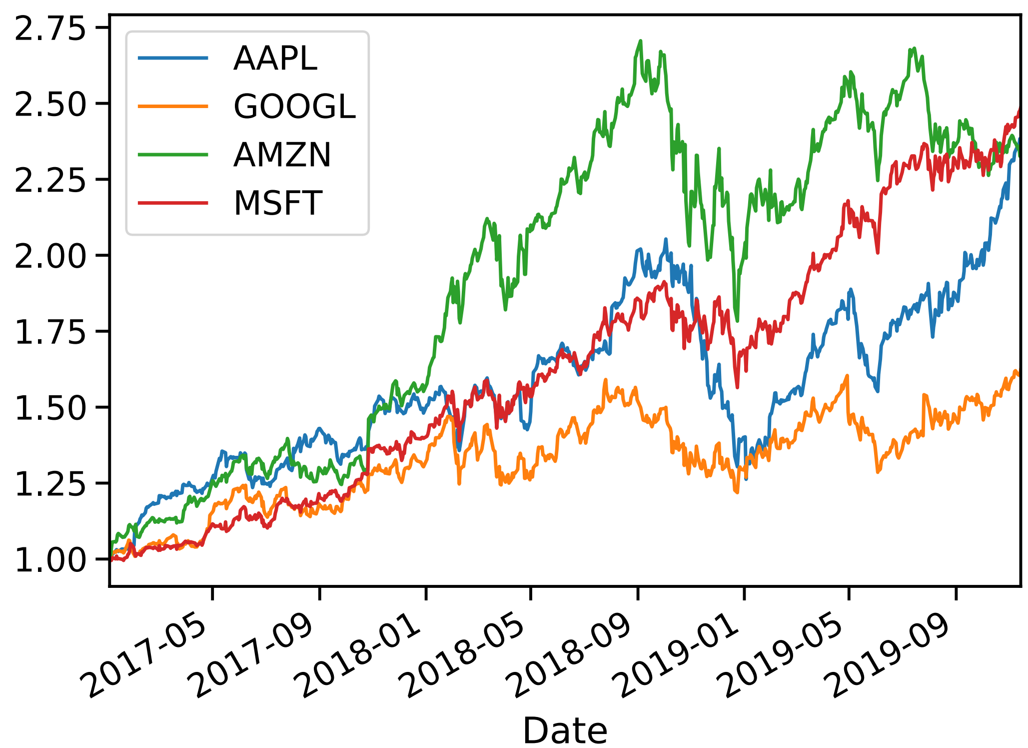
Gaussian Process Latent Variable Models in Finance

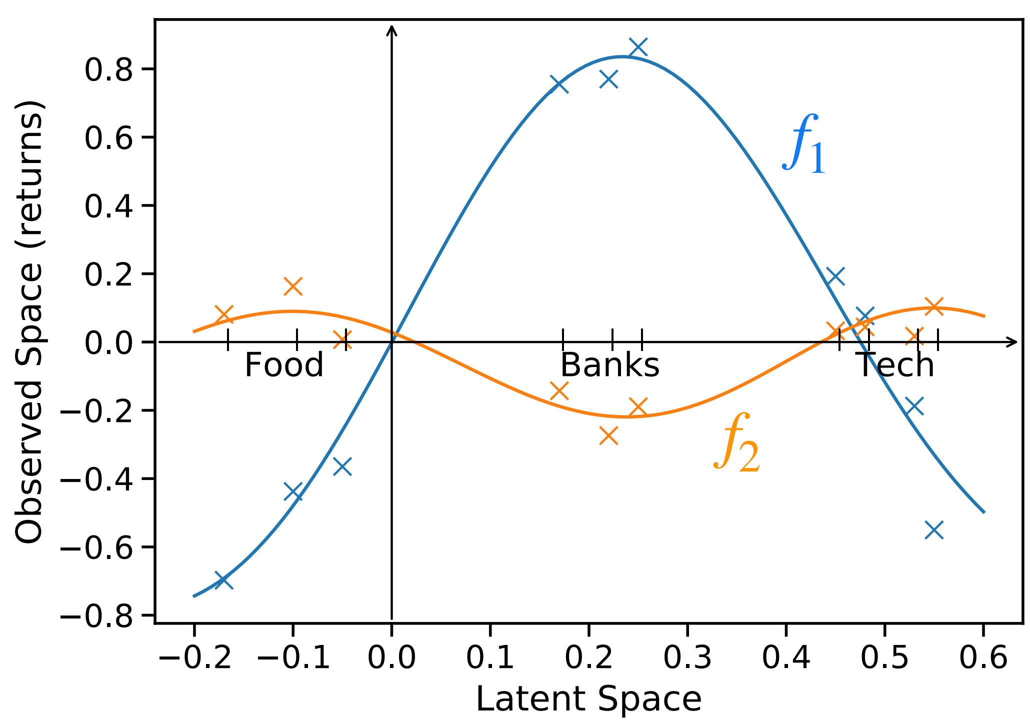
Rajbir-Singh Nirwan

November 22, 2019



RETURNS	24.10.19	25.10.19	28.10.19	29.10.19
AAPL	0.16	1.23	1.00	-2.31
GOOGL	0.11	0.41	1.95	-2.19
AMZN	1.05	-1.09	0.89	-0.80
MSFT	1.96	0.56	2.45	-0.94
FOOD				
BANK				

Generative Model



	Day1	Day2
Bank1	−0.70	0.08
Bank2	−0.44	0.16
Bank3	−0.36	0.01
Food1	0.75	−0.14
Food2	0.77	−0.27
Food3	0.86	−0.19
Tech1	0.19	0.03
Tech2	0.08	0.04
Tech3	−0.19	0.02
Tech4	−0.55	0.10

Outline

- Gaussian Processes
- Latent Variable Models
- Applications
  - Portfolio Allocation
  - Predicting missing Values
  - Structure Identification

Gaussian Processes

Weight space view

$$\Phi : x \to (\phi_1(x), \phi_2(x), \ldots, \phi_D(x))$$

$$f(x) = \boldsymbol{w}^T \Phi(x)$$

Simple and easy to interpret
but limited flexibility

$$f(x) = \boldsymbol{W}_2 \sigma \left(\boldsymbol{W}_1 \Phi_1(x) \right)$$

Highly flexible
but not interpretable

Function space view

$$k : x, x' \to k(x, x')$$

Flexibility increases with
number of data points

Mercers Theorem:

$$k(x, x') = \sum_d \lambda_d \phi_d(x) \phi_d(x')$$

$$\phi(x) = x$$
$$k(x, x') = xx'$$

$$\Phi(x) = (x, x^2)$$
$$k(x, x') = xx' + x^2 x'^2$$

$$k(x, x') = (xx' + c)^d$$
$$\Phi(x) = \textit{polynomials up to order } d$$

$$k(x, x') = \exp(-0.5\ (x - x')^2 / \ell^2)$$
$$\Phi(x) = \textit{infinitly many basis functions}$$

Gaussian Processes

Any finite collection of function values at $x_1, x_2, \ldots, x_N$ is jointly Gaussian distributed

$$p\left(f\left(x_1\right), f\left(x_2\right), \ldots, f\left(x_N\right)\right) = \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{K}) = \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} k_{11} & k_{12} & \cdots & k_{1N} \\ k_{21} & k_{22} & \cdots & k_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ k_{N1} & k_{N2} & \cdots & k_{NN} \end{pmatrix}\right) \qquad k_{ij} = k(x_i, x_j)$$

Gaussian Processes

Any finite collection of function values at $x_1, x_2, \ldots, x_N$ is jointly Gaussian distributed

$$p\left(f\left(x_1\right), f\left(x_2\right), \ldots, f\left(x_N\right)\right) = \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{K}) = \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} k_{11} & k_{12} & \cdots & k_{1N} \\ k_{21} & k_{22} & \cdots & k_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ k_{N1} & k_{N2} & \cdots & k_{NN} \end{pmatrix}\right) \qquad k_{ij} = k(x_i, x_j)$$

Common Kernel Functions

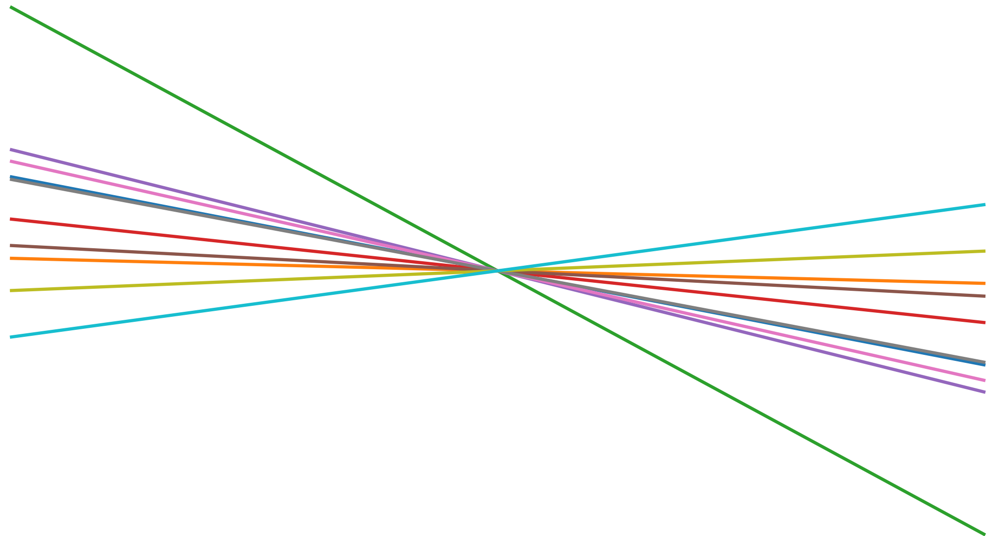
$$k_{linear}(x, x') = xx'$$

$$k_{rbf}(x, x') = \exp\left(-\frac{1}{2\ell^2}(x - x')^2\right)$$

$$k_{ou}(x, x') = \exp\left(-\frac{1}{\ell}\,|x - x'|\right)$$

$$k_{mat32}(x, x') = \left(1 + \frac{\sqrt{3}\,|x - x'|}{\ell}\right)\exp\left(-\frac{\sqrt{3}\,|x - x'|}{\ell}\right)$$

$$k_{periodic}(x, x') = \exp\left(-\frac{2}{\ell^2}\sin^2(|x - x'|)\right)$$



Gaussian Processes

Any finite collection of function values at $x_1, x_2, \ldots, x_N$ is jointly Gaussian distributed

$$p\left(f\left(x_1\right), f\left(x_2\right), \ldots, f\left(x_N\right)\right) = \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{K}) = \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} k_{11} & k_{12} & \cdots & k_{1N} \\ k_{21} & k_{22} & \cdots & k_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ k_{N1} & k_{N2} & \cdots & k_{NN} \end{pmatrix}\right) \qquad k_{ij} = k(x_i, x_j)$$

Common Kernel Functions

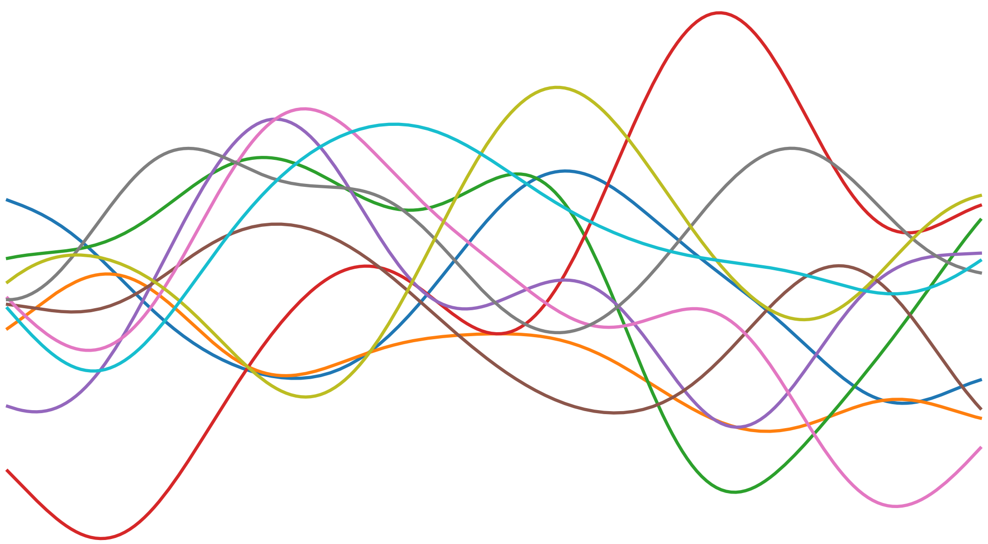
$$k_{linear}(x, x') = xx'$$

$$k_{rbf}(x, x') = \exp\left(-\frac{1}{2\ell^2}(x - x')^2\right)$$

$$k_{ou}(x, x') = \exp\left(-\frac{1}{\ell}|x - x'|\right)$$

$$k_{mat32}(x, x') = \left(1 + \frac{\sqrt{3}\,|x - x'|}{\ell}\right)\exp\left(-\frac{\sqrt{3}\,|x - x'|}{\ell}\right)$$

$$k_{periodic}(x, x') = \exp\left(-\frac{2}{\ell^2}\sin^2(|x - x'|)\right)$$



Gaussian Processes

Any finite collection of function values at $x_1, x_2, \ldots, x_N$ is jointly Gaussian distributed

$$p\left(f\left(x_1\right), f\left(x_2\right), \ldots, f\left(x_N\right)\right) = \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{K}) = \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} k_{11} & k_{12} & \cdots & k_{1N} \\ k_{21} & k_{22} & \cdots & k_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ k_{N1} & k_{N2} & \cdots & k_{NN} \end{pmatrix}\right) \qquad k_{ij} = k(x_i, x_j)$$

Common Kernel Functions

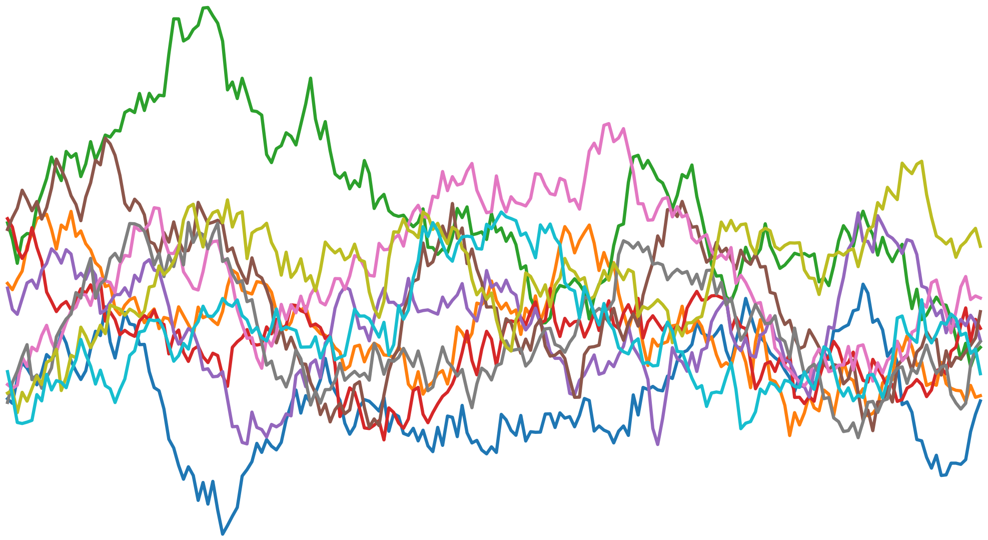
$$k_{linear}(x, x') = xx'$$

$$k_{rbf}(x, x') = \exp\left(-\frac{1}{2\ell^2}(x - x')^2\right)$$

$$k_{ou}(x, x') = \exp\left(-\frac{1}{\ell}|x - x'|\right)$$

$$k_{mat32}(x, x') = \left(1 + \frac{\sqrt{3}\,|x - x'|}{\ell}\right)\exp\left(-\frac{\sqrt{3}\,|x - x'|}{\ell}\right)$$

$$k_{periodic}(x, x') = \exp\left(-\frac{2}{\ell^2}\sin^2(|x - x'|)\right)$$



Gaussian Processes

Any finite collection of function values at $x_1, x_2, \ldots, x_N$ is jointly Gaussian distributed

$$p\left(f\left(x_1\right), f\left(x_2\right), \ldots, f\left(x_N\right)\right) = \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{K}) = \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} k_{11} & k_{12} & \cdots & k_{1N} \\ k_{21} & k_{22} & \cdots & k_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ k_{N1} & k_{N2} & \cdots & k_{NN} \end{pmatrix}\right) \qquad k_{ij} = k(x_i, x_j)$$

Common Kernel Functions

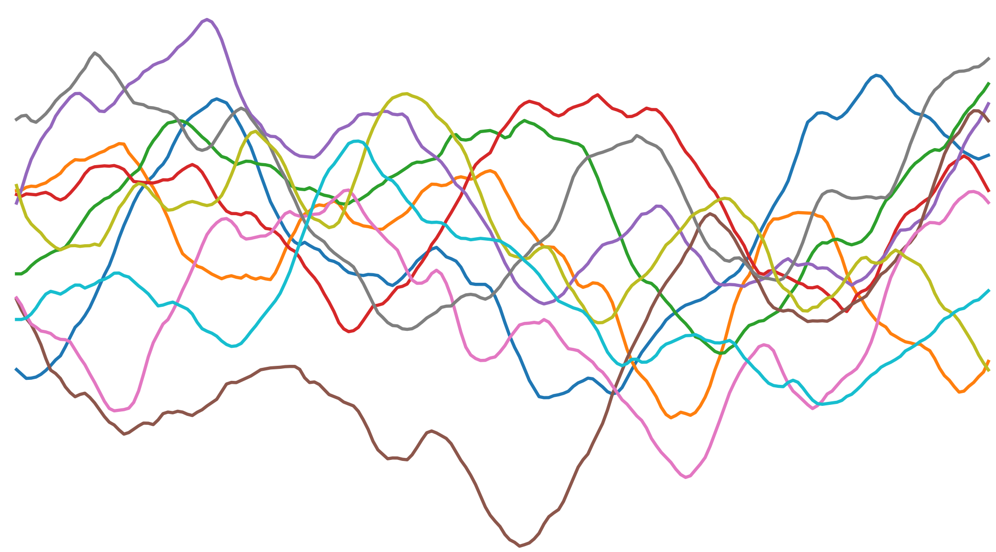
$$k_{linear}(x, x') = xx'$$

$$k_{rbf}(x, x') = \exp\left(-\frac{1}{2\ell^2}(x - x')^2\right)$$

$$k_{ou}(x, x') = \exp\left(-\frac{1}{\ell}\,|x - x'|\right)$$

$$k_{mat32}(x, x') = \left(1 + \frac{\sqrt{3}\,|x - x'|}{\ell}\right)\exp\left(-\frac{\sqrt{3}\,|x - x'|}{\ell}\right)$$

$$k_{periodic}(x, x') = \exp\left(-\frac{2}{\ell^2}\sin^2(|x - x'|)\right)$$



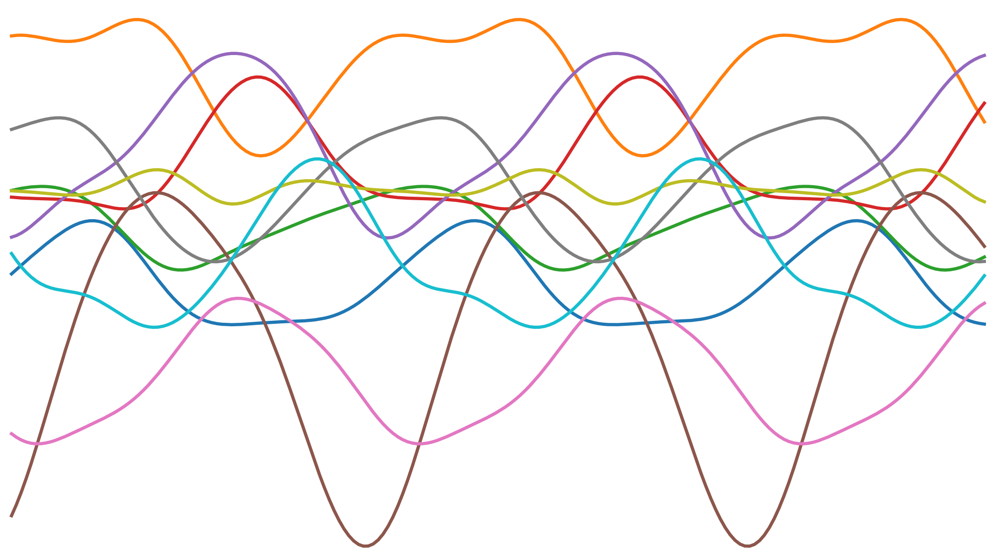
Gaussian Processes

Any finite collection of function values at $x_1, x_2, \ldots, x_N$ is jointly Gaussian distributed

$$p\left(f\left(x_1\right), f\left(x_2\right), \ldots, f\left(x_N\right)\right) = \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{K}) = \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} k_{11} & k_{12} & \cdots & k_{1N} \\ k_{21} & k_{22} & \cdots & k_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ k_{N1} & k_{N2} & \cdots & k_{NN} \end{pmatrix}\right) \qquad k_{ij} = k(x_i, x_j)$$

Common Kernel Functions

$$k_{linear}(x, x') = xx'$$



$$k_{rbf}(x, x') = \exp\left(-\frac{1}{2\ell^2}(x - x')^2\right)$$

$$k_{ou}(x, x') = \exp\left(-\frac{1}{\ell}|x - x'|\right)$$

$$k_{mat32}(x, x') = \left(1 + \frac{\sqrt{3}\,|x - x'|}{\ell}\right)\exp\left(-\frac{\sqrt{3}\,|x - x'|}{\ell}\right)$$

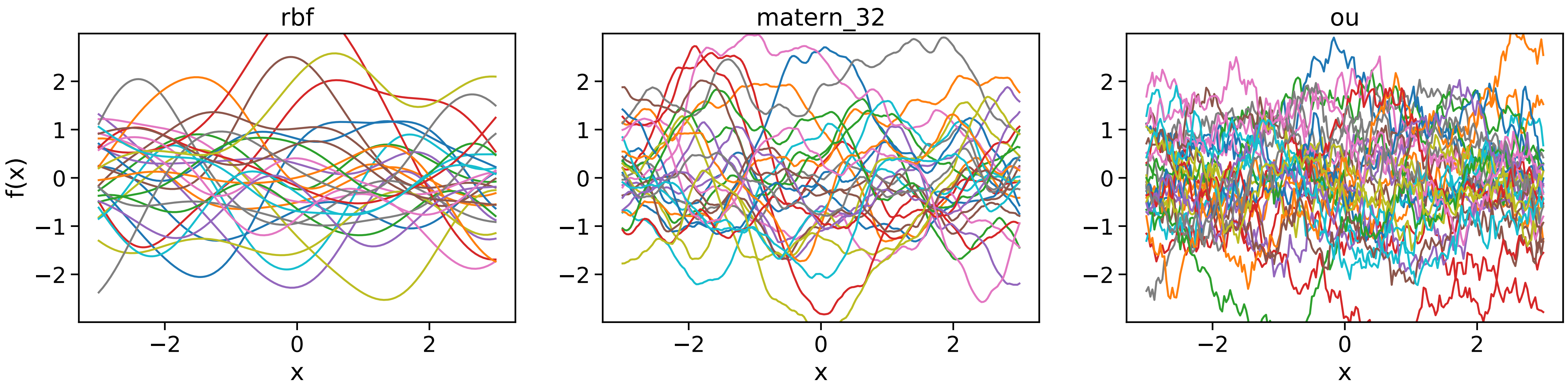
$$k_{periodic}(x, x') = \exp\left(-\frac{2}{\ell^2}\sin^2(|x - x'|)\right)$$

Gaussian Processes

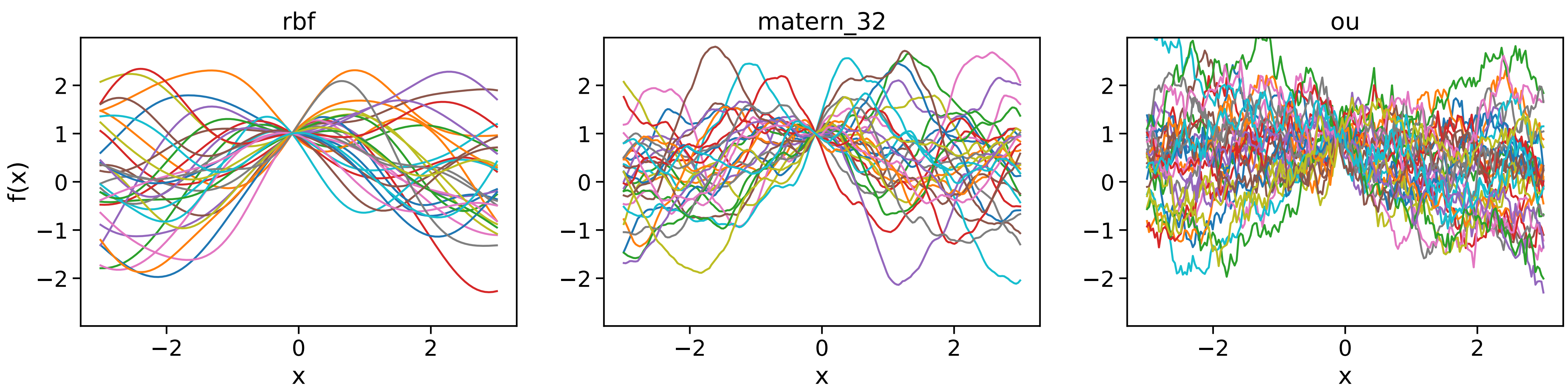
Bayes Theorem

$$p(f \mid \mathscr{D}) = \frac{p(\mathscr{D} \mid f)p(f)}{p(\mathscr{D})}$$

Prior



Posterior

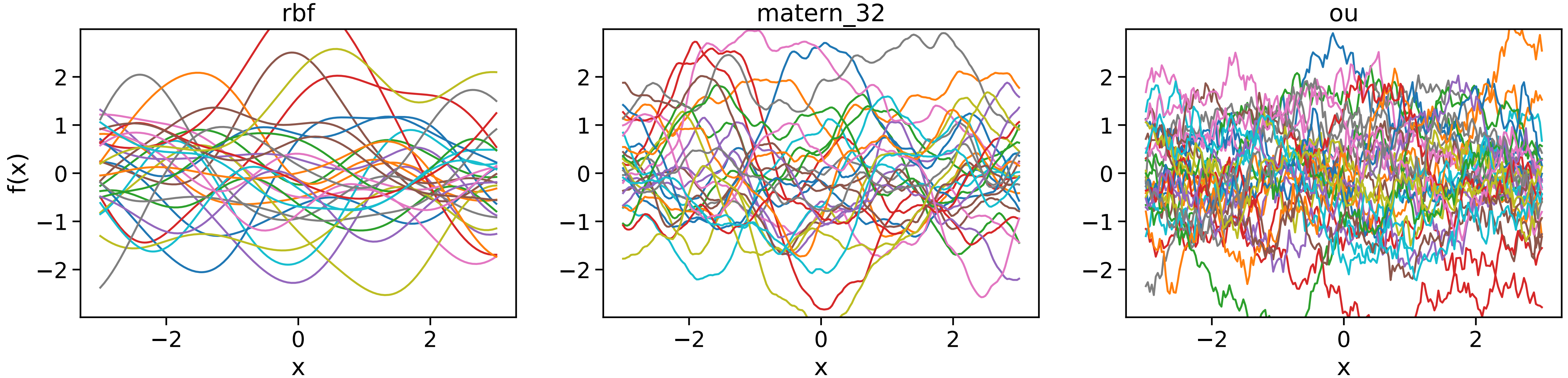


Gaussian Processes

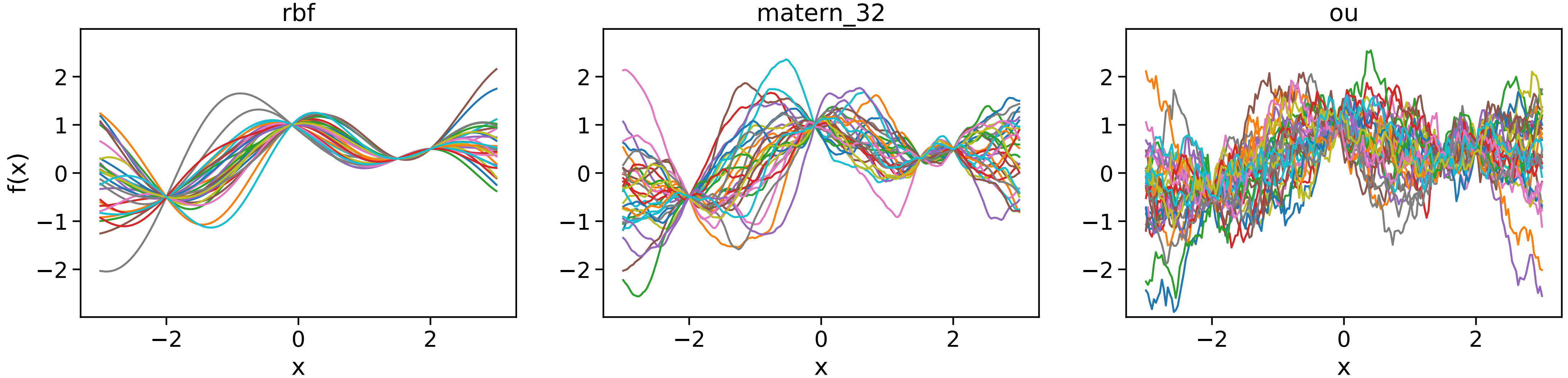
Bayes Theorem

$$p(f \mid \mathscr{D}) = \frac{p(\mathscr{D} \mid f)p(f)}{p(\mathscr{D})}$$

Prior

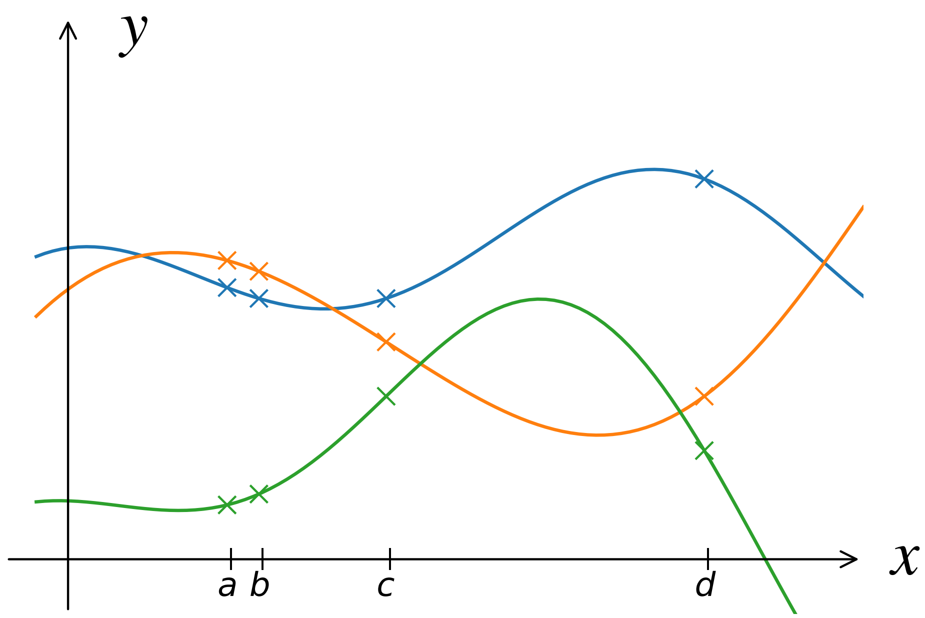


Posterior



Latent Variable Models

$$\boldsymbol{X} \in \mathbb{R}^{N \times Q} \quad \xrightarrow{f} \quad \boldsymbol{Y} \in \mathbb{R}^{N \times D}$$



$$\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \xrightarrow{f_i} \begin{pmatrix} a_1 & a_2 & a_3 & \\ b_1 & b_2 & b_3 & \\ c_1 & c_2 & c_3 & \cdots \\ d_1 & d_2 & d_3 & \end{pmatrix}$$

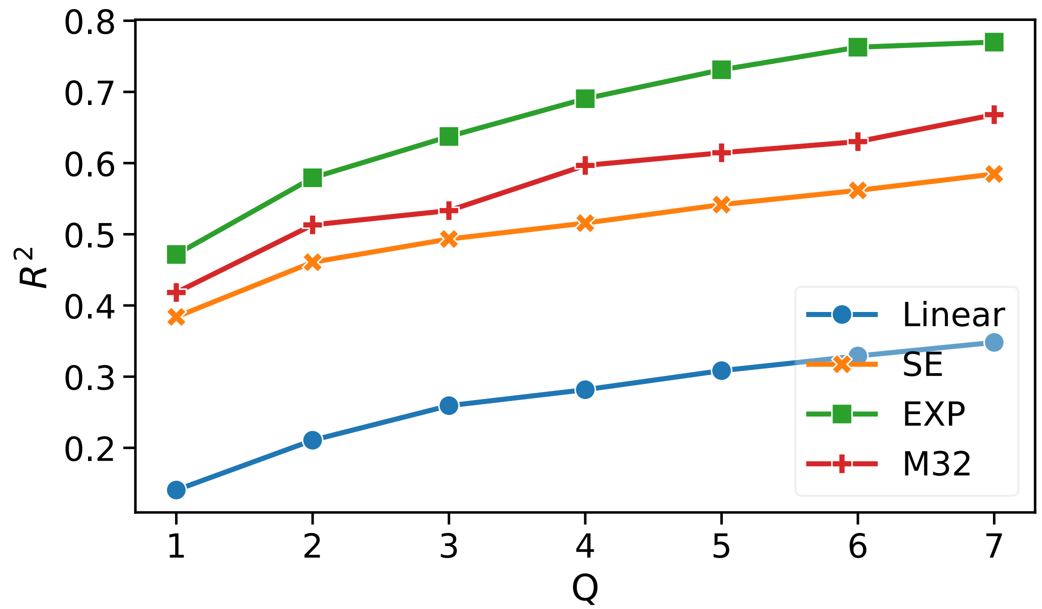
- Can we infer the hidden state $\boldsymbol{X}$ only by looking at $\boldsymbol{Y}$? Yes
- Inference using GPs also gives us the covariance $\boldsymbol{K}$ between different points

Experiments

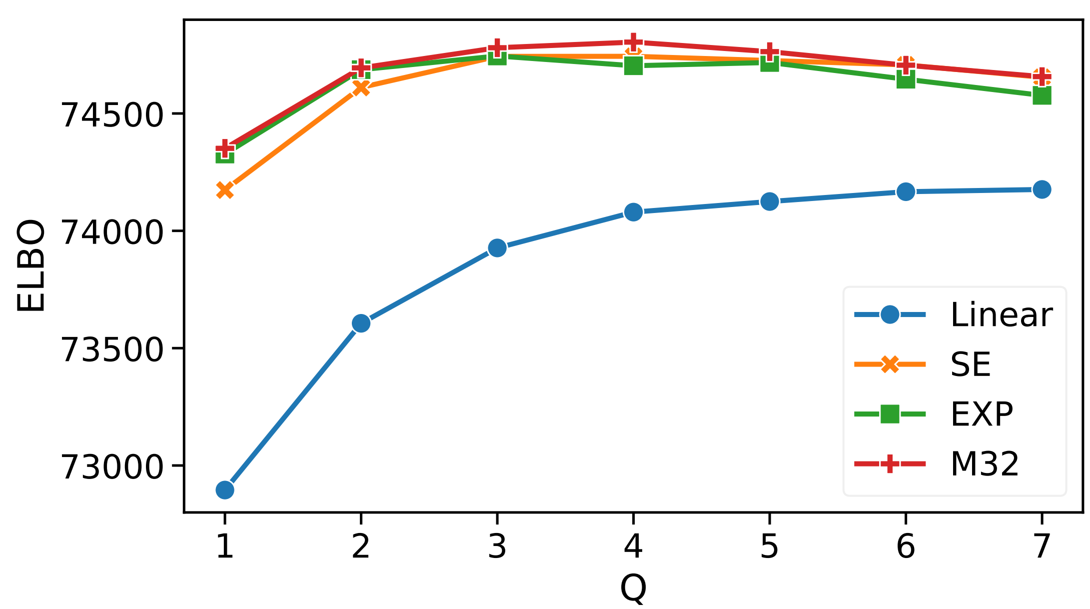
Use Variational Bayes for the inference - data $Y \in \mathbb{R}^{N \times D}$

Approximate the true posterior with a simple distribution

R^2 - Variance of the data captured by the model



ELBO - Lower Bound to the marginal likelihood



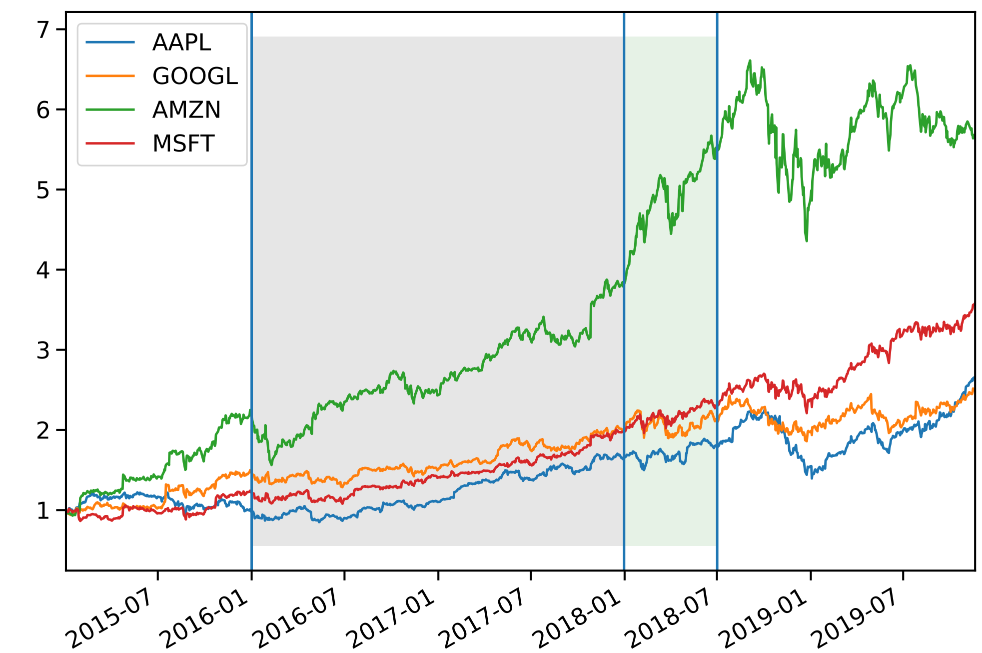
Portfolio Allocation

Given N stocks, how should I weight them to get an optimal portfolio?

Markowitz Portfolio Theory

$$\boldsymbol{w}_{opt} = \min_{\boldsymbol{w}} \left(\boldsymbol{w}^T \boldsymbol{K} \boldsymbol{w} - q \boldsymbol{w}^T \boldsymbol{\mu} \right)$$

Learn weights on previous 2 years
Hold portfolio for next 6 months



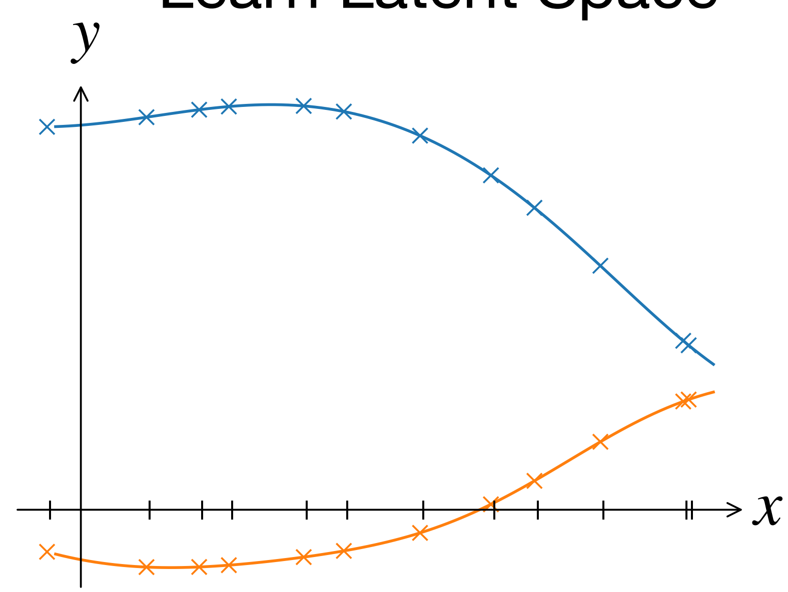
Backtesting on S&P500 from 2002 to 2018

Model	Linear	SE	EXP	M32	Sample Cov	Ledoit Wolf	Eq. Weighted
Mean	0.142	0.151	0.155	0.158	0.149	0.148	0.182
Std	0.158	0.156	0.154	0.153	0.159	0.159	0.232
Sharpe ratio	0.901	0.969	1.008	1.029	0.934	0.931	0.786

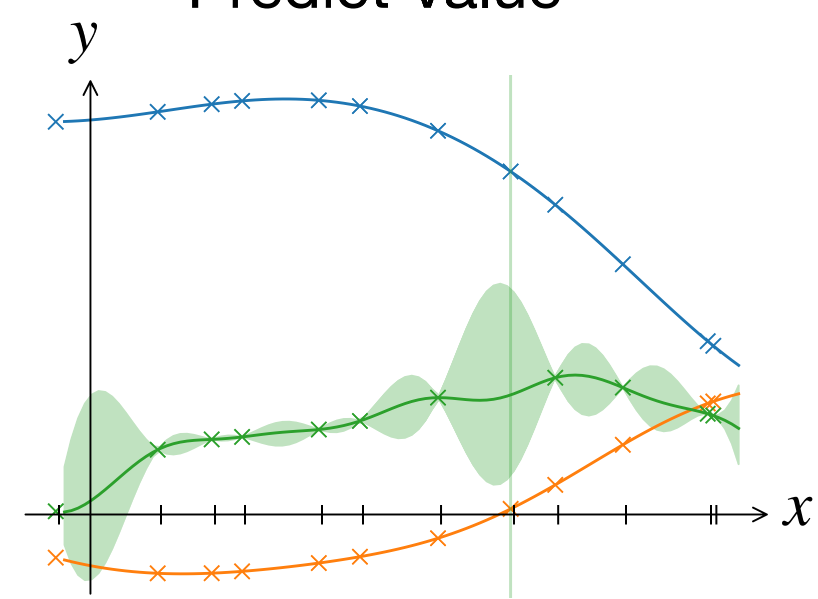
Predict Missing Values

$$\left(\begin{array}{cccc} & a_1 & a_2 & a_3 \\ & b_1 & b_2 & NaN \\ \cdots & c_1 & c_2 & c_3 \\ & d_1 & d_2 & d_3 \end{array}\right)$$

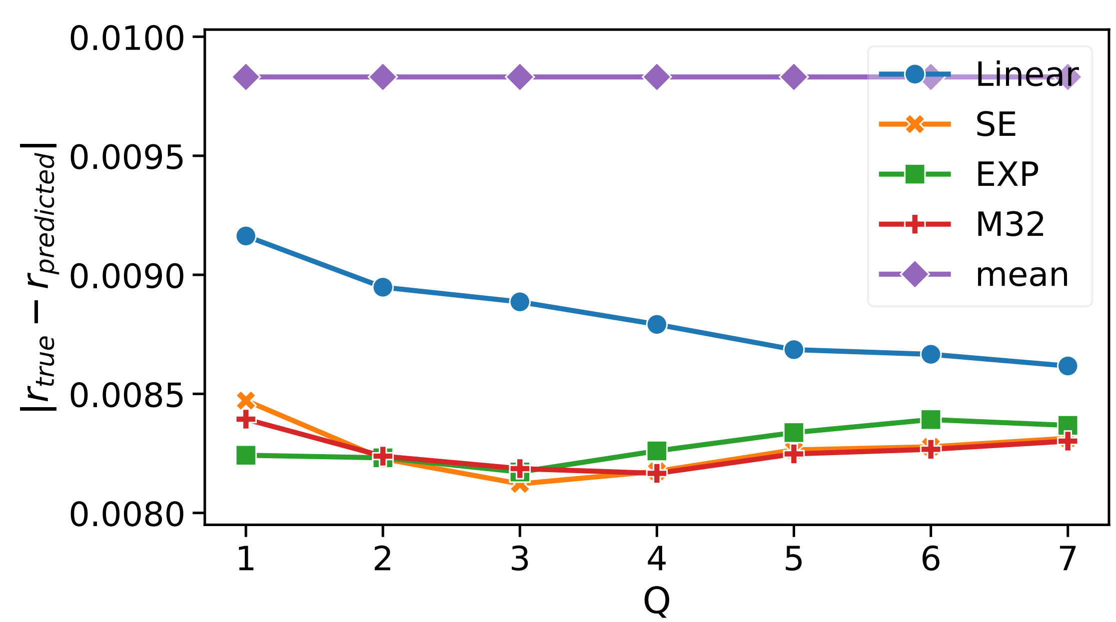
Learn Latent Space



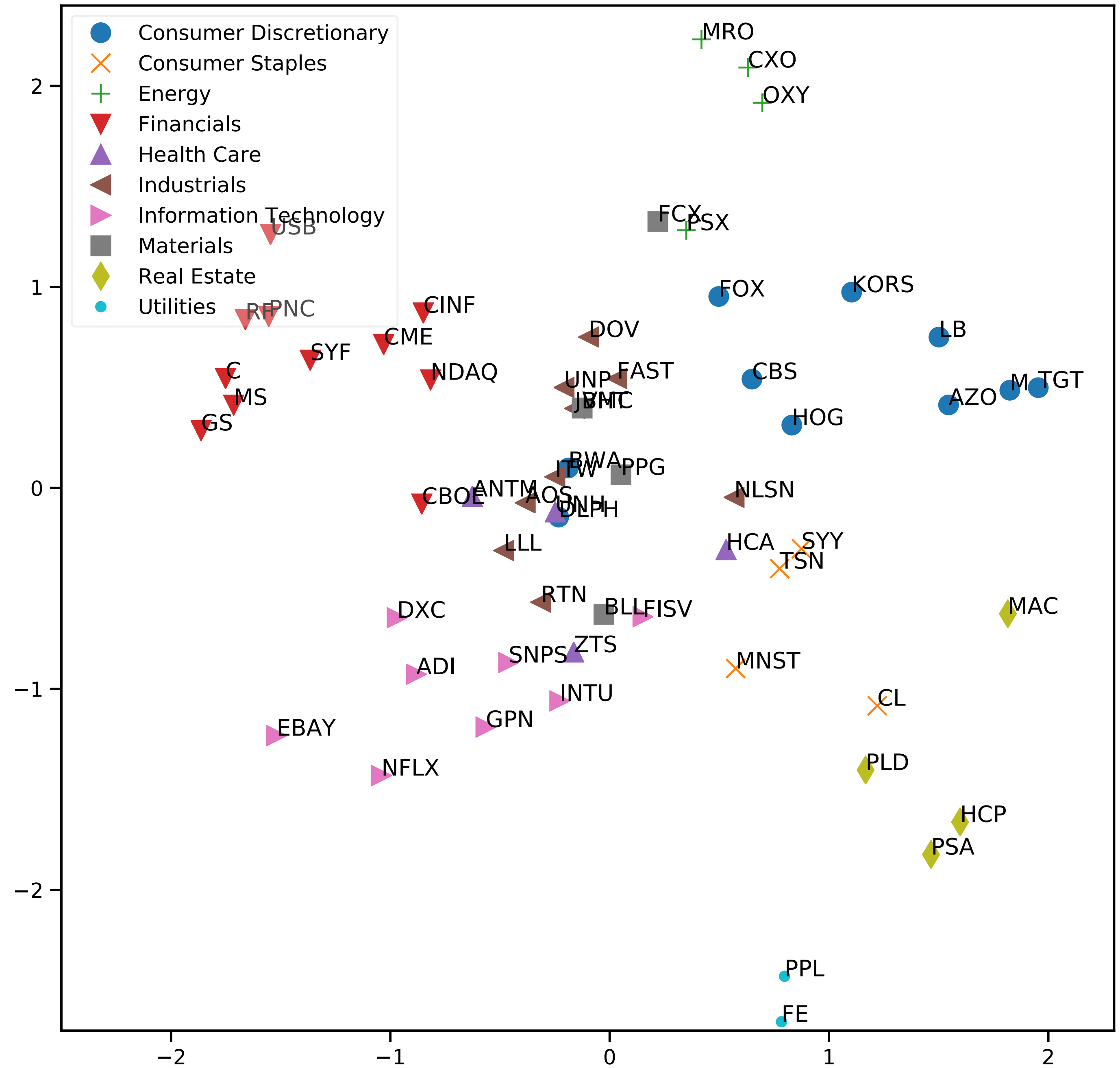
Predict Value



Prediction of missing values for held out dataset



Visualization of Latent Space



Summary

- Use of Gaussian Processes in Finance
- Build Portfolio based on the Covariance between Stocks
- Better Predictor for Missing values
- Latent Space Structure Identification

Applications of Gaussian Process Latent Variable Models in Finance
RS Nirwan, N Bertschinger - Proceedings of SAI Intelligent Systems Conference, 2019