Estimation of Covariance Matrices using Gaussian Processes

Rajbir-Singh Nirwan, Nils Bertschinger

13.03.2018

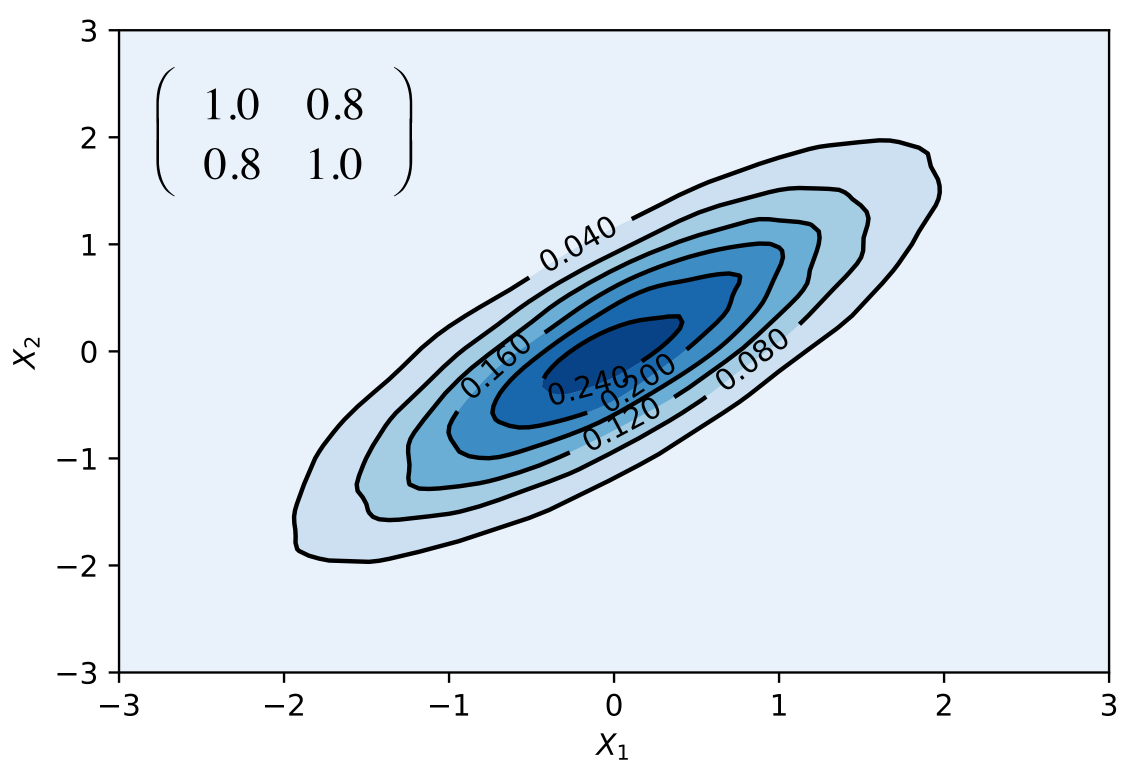
FIAS Frankfurt Institute for Advanced Studies

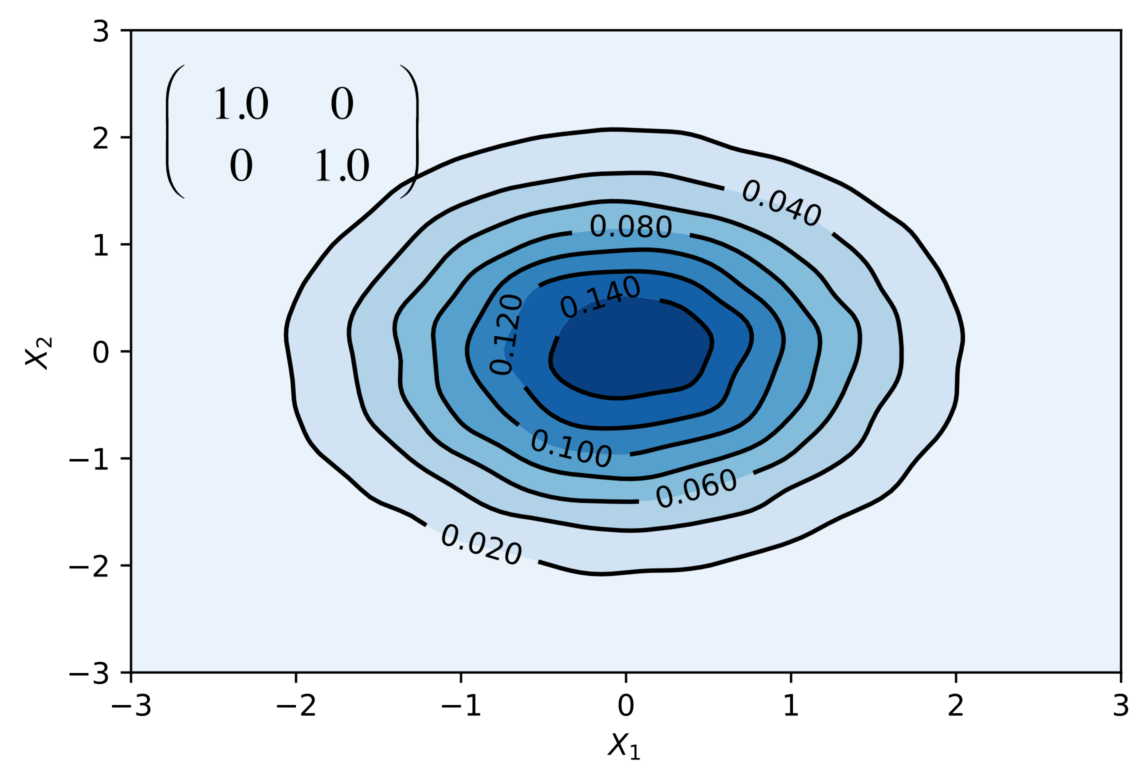
Outline

- Gaussian Processes
- Latent Variable Models
- Financial Applications

Multivariate Gaussian Distribution

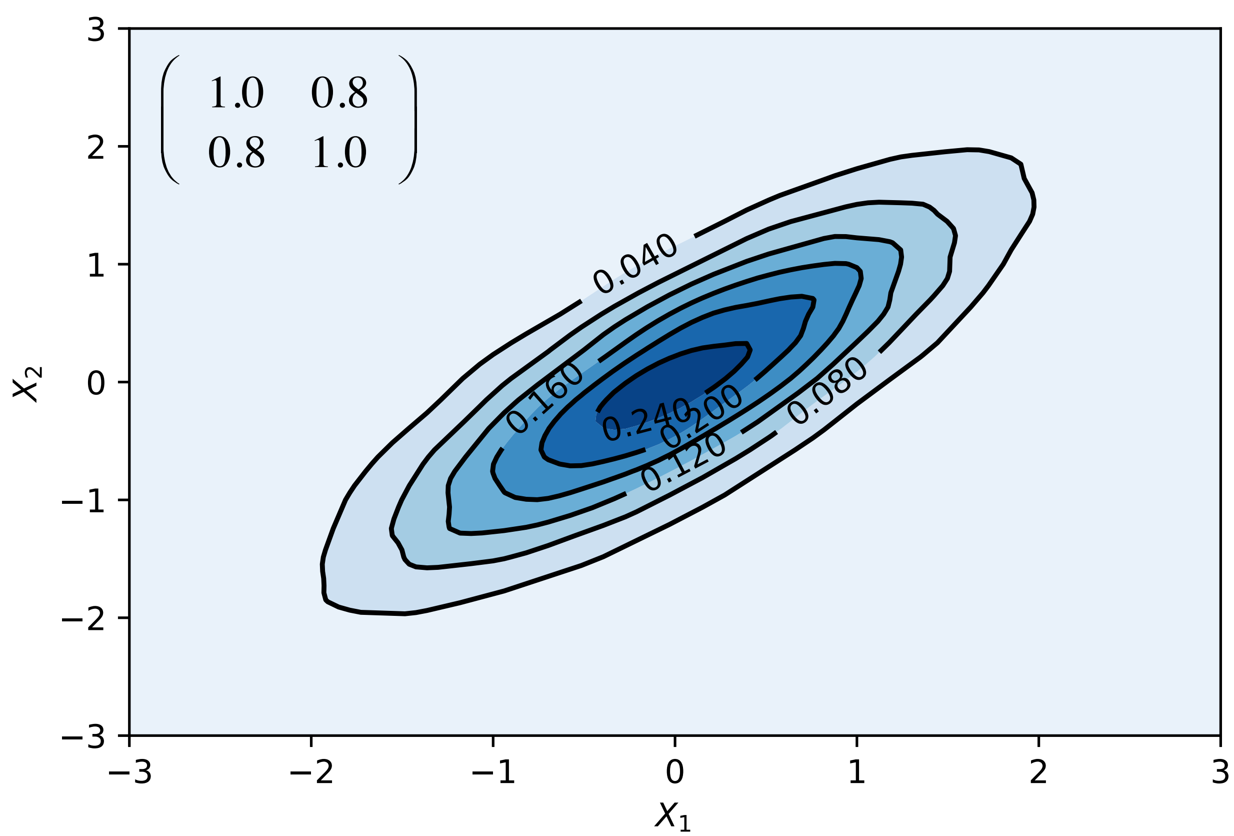
$$p(x_1,\ldots,x_D)=\frac{1}{\sqrt{(2\pi)^D\,|\Sigma|}}\,e^{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)}$$

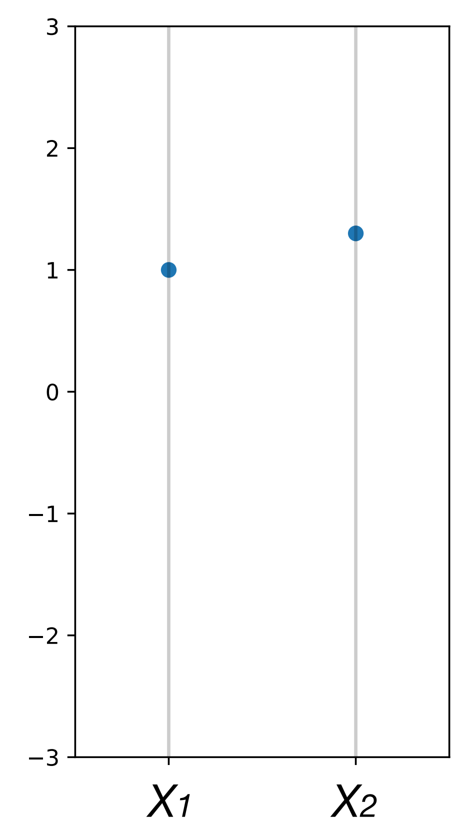




Multivariate Gaussian Distribution

$$p(x_1,\ldots,x_D)=\frac{1}{\sqrt{(2\pi)^D\,|\Sigma|}}e^{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)}$$

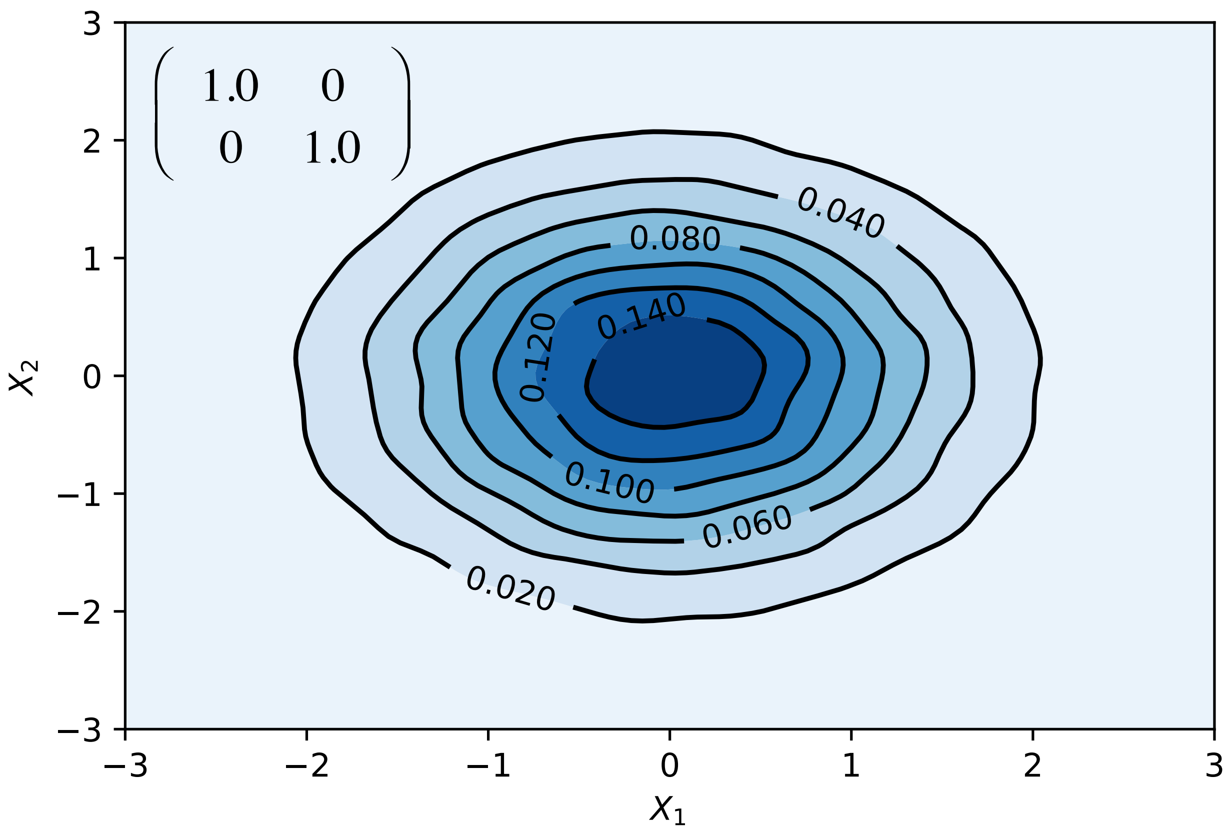


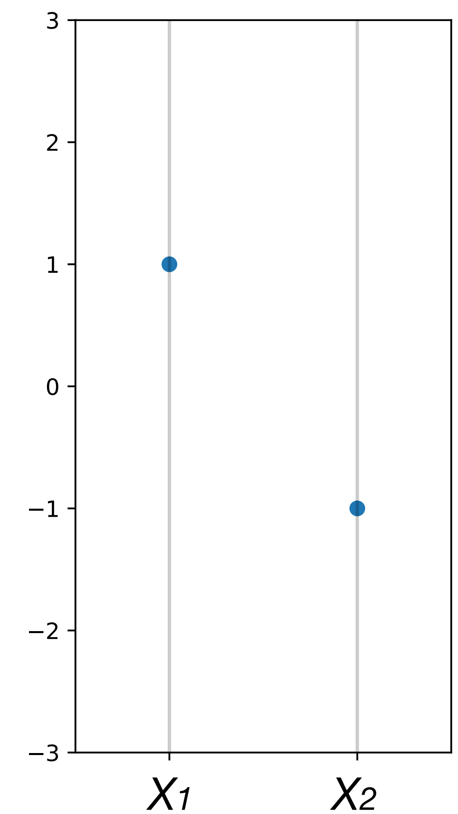


$$k(x_1,x_2)=0.8$$

Multivariate Gaussian Distribution

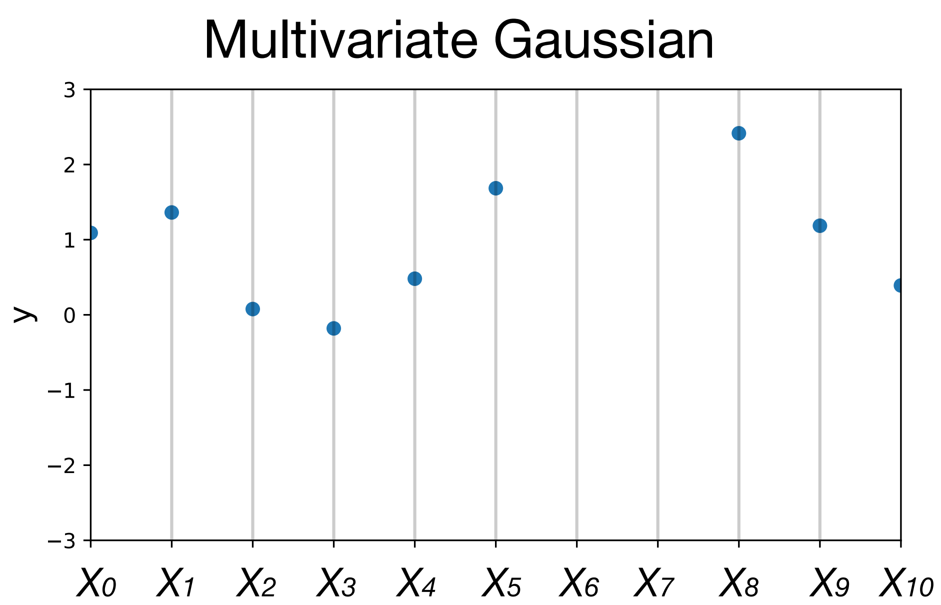
$$p(x_1,\ldots,x_D) = \frac{1}{\sqrt{(2\pi)^D \,|\Sigma|}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)}$$





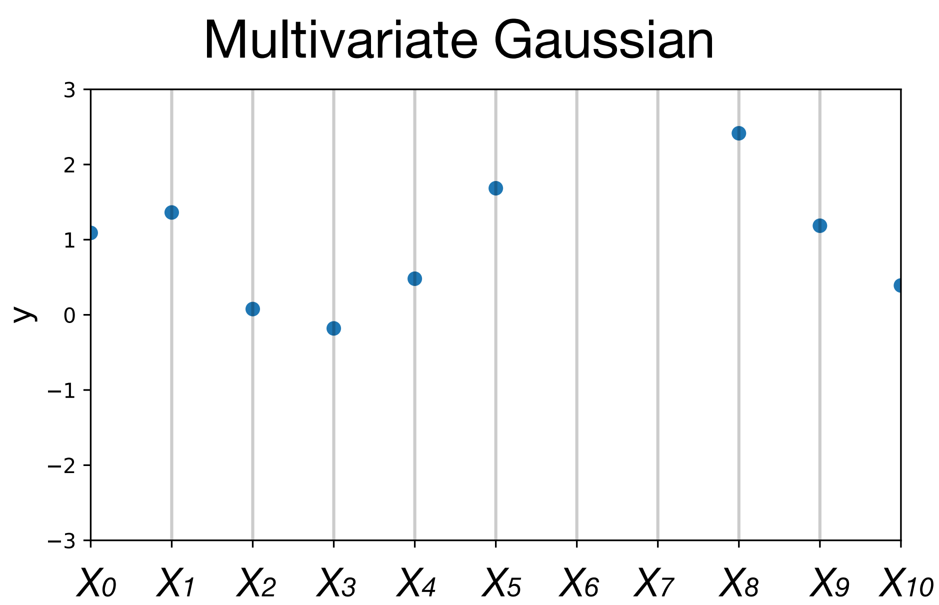
$$k(x_1, x_2) = 0$$

Multivariate Gaussian Distribution



$$Y \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} k_{00} & \cdots & k_{0N} \\ \vdots & \ddots & \vdots \\ k_{N0} & \cdots & k_{NN} \end{pmatrix}\right)$$

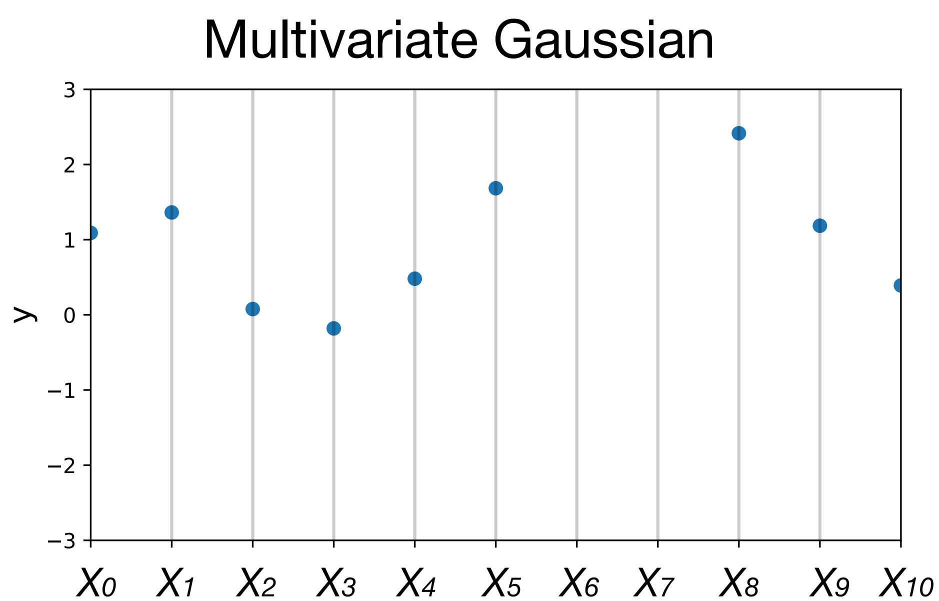
Multivariate Gaussian Distribution



$$Y \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} k_{00} & \cdots & k_{0N} \\ \vdots & \ddots & \vdots \\ k_{N0} & \cdots & k_{NN} \end{pmatrix}\right)$$

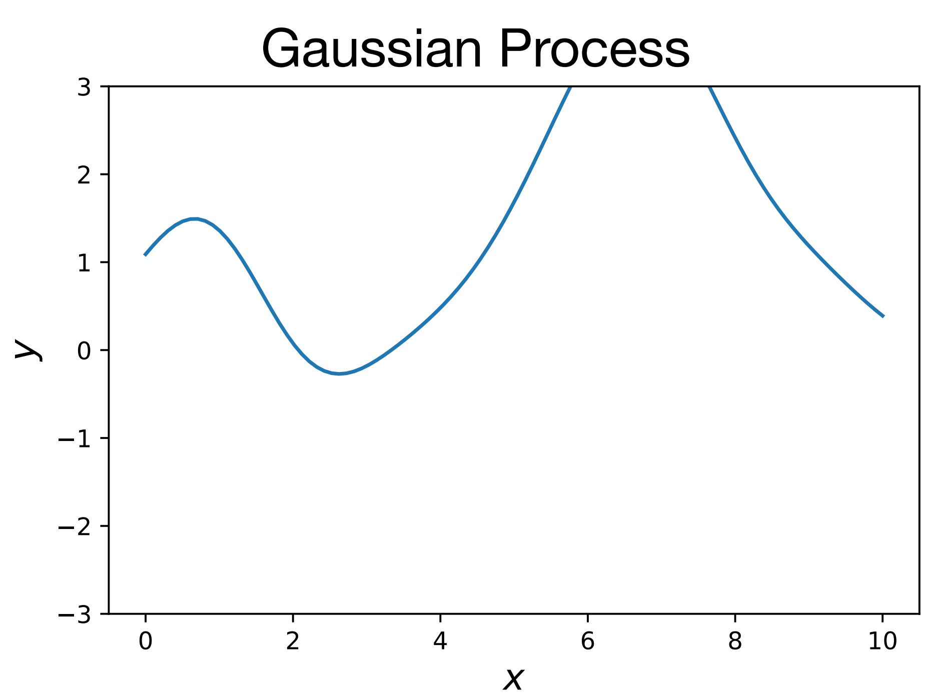
$$k(x_i, x_j) = e^{-\frac{1}{2}(x_i - x_j)^2}$$

Gaussian Processes



$$Y \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} k_{00} & \cdots & k_{0N} \\ \vdots & \ddots & \vdots \\ k_{N0} & \cdots & k_{NN} \end{pmatrix}\right)$$

$$k(x_i, x_j) = e^{-\frac{1}{2}(x_i - x_j)^2}$$

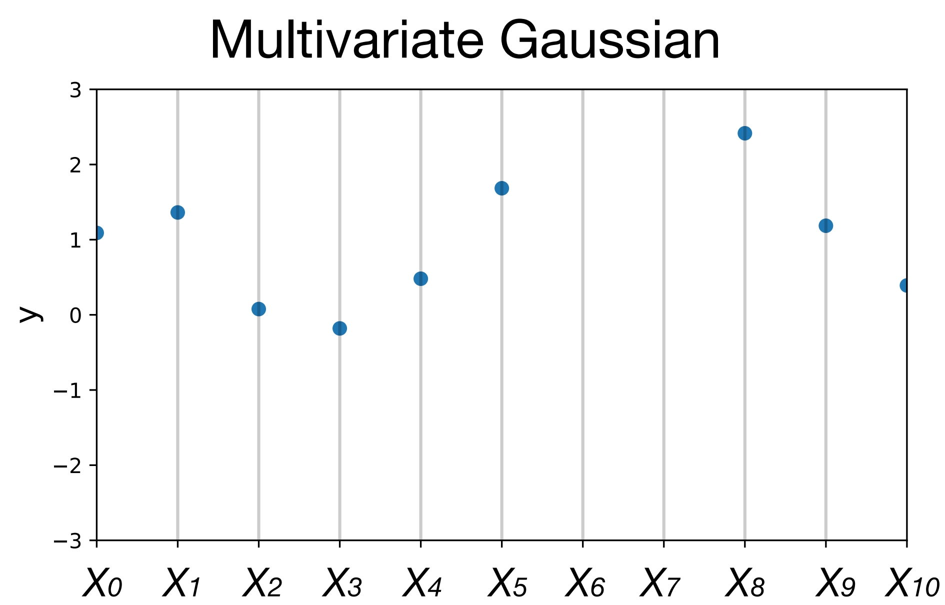


$$Y \sim GP\big(m(x), k(x, x')\big)$$

$$m(x) = 0$$

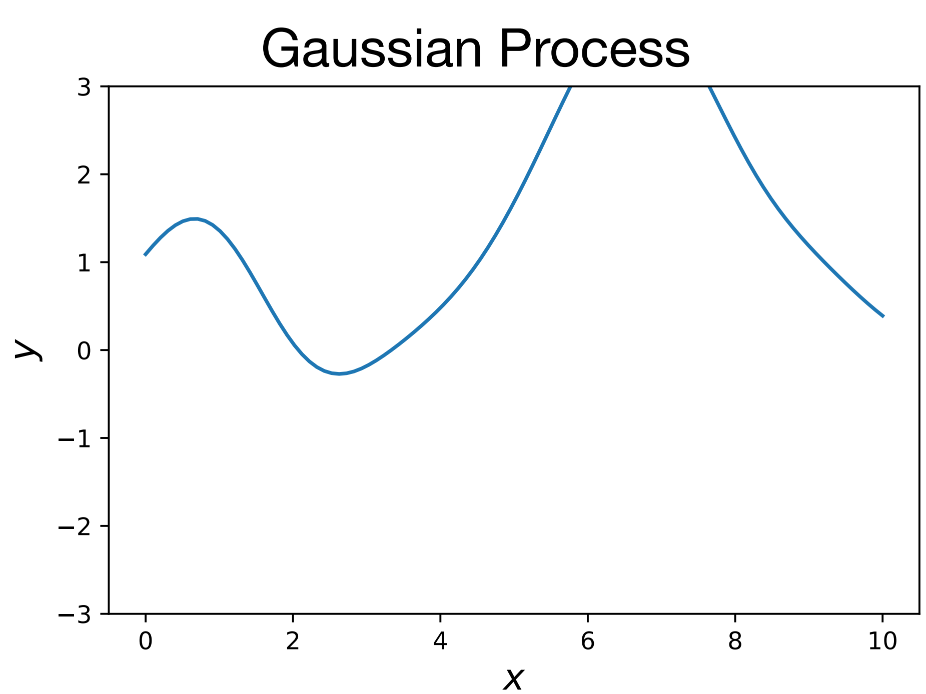
$$k(x, x') = e^{-\frac{1}{2}(x - x')^2}$$

Gaussian Processes



$$Y \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} k_{00} & \cdots & k_{0N} \\ \vdots & \ddots & \vdots \\ k_{N0} & \cdots & k_{NN} \end{pmatrix} \right)$$

$$k(x_i, x_j) = e^{-\frac{1}{2}(x_i - x_j)^2}$$



Other Kernels

Linear	$k_\theta(x_i, x_j) = \theta x_i x_j$
RBF	$k_\theta(x_i, x_j) = \theta_1 e^{-\frac{1}{2\theta_2}(x_i - x_j)^2}$
OU	$k_\theta(x_i, x_j) = \theta_1 e^{-\frac{1}{2\theta_2}\lvert x_i - x_j \rvert}$
Periodic	$k_\theta(x_i, x_j) = \theta_1 e^{-\frac{1}{2\theta_2}\left(sin^2(\theta_3(x_i - x_j))\right)}$

Linear Latent Variable Model

Problem: $Y = XW^T + \epsilon$

$$Y \in \mathbb{R}^{N \times D} \qquad N - number\ of\ data\ points$$
$$X \in \mathbb{R}^{N \times Q} \qquad D - dimension\ of\ data\ space$$
$$W \in \mathbb{R}^{Q \times D} \qquad Q - dimension\ of\ latent\ space$$

Linear Latent Variable Model

Problem: $Y = XW^T + \epsilon$

$$Y \in \mathbb{R}^{N \times D} \qquad N - \text{number of data points}$$
$$X \in \mathbb{R}^{N \times Q} \qquad D - \text{dimension of data space}$$
$$W \in \mathbb{R}^{Q \times D} \qquad Q - \text{dimension of latent space}$$

- Principle Component Analysis (PPCA)

$$y_{n,\cdot} = Wx_{n,\cdot} + \epsilon_{n,\cdot}$$

$$X \sim \mathcal{N}(0, \mathrm{I})$$
$$\epsilon \sim \mathcal{N}(0, \sigma^2 \mathrm{I})$$

$$p(Y \mid W) = \prod_n \mathcal{N}(y_{n,\cdot} \mid 0, WW^T + \sigma^2 \mathrm{I})$$

Linear Latent Variable Model

Problem: $Y = XW^T + \epsilon$

$$Y \in \mathbb{R}^{N \times D} \qquad N - \textit{number of data points}$$
$$X \in \mathbb{R}^{N \times Q} \qquad D - \textit{dimension of data space}$$
$$W \in \mathbb{R}^{Q \times D} \qquad Q - \textit{dimension of latent space}$$

- Principle Component Analysis (PPCA)

$$y_{n,\cdot} = Wx_{n,\cdot} + \epsilon_{n,\cdot}$$
$$X \sim \mathcal{N}(0, \mathrm{I})$$
$$\epsilon \sim \mathcal{N}(0, \sigma^2 \mathrm{I})$$
$$p(Y \mid W) = \prod_n \mathcal{N}(y_{n,\bullet} \mid 0, WW^T + \sigma^2 \mathrm{I})$$

- Dual Principle Component Analysis (Dual PPCA)

$$y_{\bullet,d} = Xw_{\bullet,d} + \epsilon_{\bullet,d}$$
$$W \sim \mathcal{N}(0, \mathrm{I})$$
$$\epsilon \sim \mathcal{N}(0, \sigma^2 \mathrm{I})$$
$$p(Y \mid X) = \prod_d \mathcal{N}(y_{\bullet,d} \mid 0, XX^T + \sigma^2 \mathrm{I})$$

Linear Latent Variable Model

Problem: $Y = XW^T + \epsilon$

$$Y \in \mathbb{R}^{N \times D} \qquad N - \textit{number of data points}$$
$$X \in \mathbb{R}^{N \times Q} \qquad D - \textit{dimension of data space}$$
$$W \in \mathbb{R}^{Q \times D} \qquad Q - \textit{dimension of latent space}$$

- Principle Component Analysis (PPCA)

$$y_{n,\bullet} = Wx_{n,\bullet} + \epsilon_{n,\bullet}$$

$$X \sim \mathcal{N}(0, \mathrm{I})$$
$$\epsilon \sim \mathcal{N}(0, \sigma^2 \mathrm{I})$$

$$p(Y \mid W) = \prod_n \mathcal{N}(y_{n,\bullet} \mid 0, WW^T + \sigma^2 \mathrm{I})$$

- Dual Principle Component Analysis (Dual PPCA)

$$y_{\bullet,d} = Xw_{\bullet,d} + \epsilon_{\bullet,d}$$

$$W \sim \mathcal{N}(0, \mathrm{I})$$
$$\epsilon \sim \mathcal{N}(0, \sigma^2 \mathrm{I})$$

$$p(Y \mid X) = \prod_d \mathcal{N}(y_{\bullet,d} \mid 0, XX^T + \sigma^2 \mathrm{I})$$

$$(XX^T)_{ij} = X_i^T X_j$$

Linear Latent Variable Model

Problem: $Y = XW^T + \epsilon$

$$Y \in \mathbb{R}^{N\times D} \qquad N - number\ of\ data\ points$$
$$X \in \mathbb{R}^{N\times Q} \qquad D - dimension\ of\ data\ space$$
$$W \in \mathbb{R}^{Q\times D} \qquad Q - dimension\ of\ latent\ space$$

- Principle Component Analysis (PPCA)

$$y_{n,\cdot} = Wx_{n,\cdot} + \epsilon_{n,\cdot}$$

$$X \sim \mathcal{N}(0, \mathrm{I})$$
$$\epsilon \sim \mathcal{N}(0, \sigma^2\mathrm{I})$$

$$p(Y \mid W) = \prod_n \mathcal{N}(y_{n,\bullet} \mid 0, WW^T + \sigma^2\mathrm{I})$$

- Dual Principle Component Analysis (Dual PPCA)

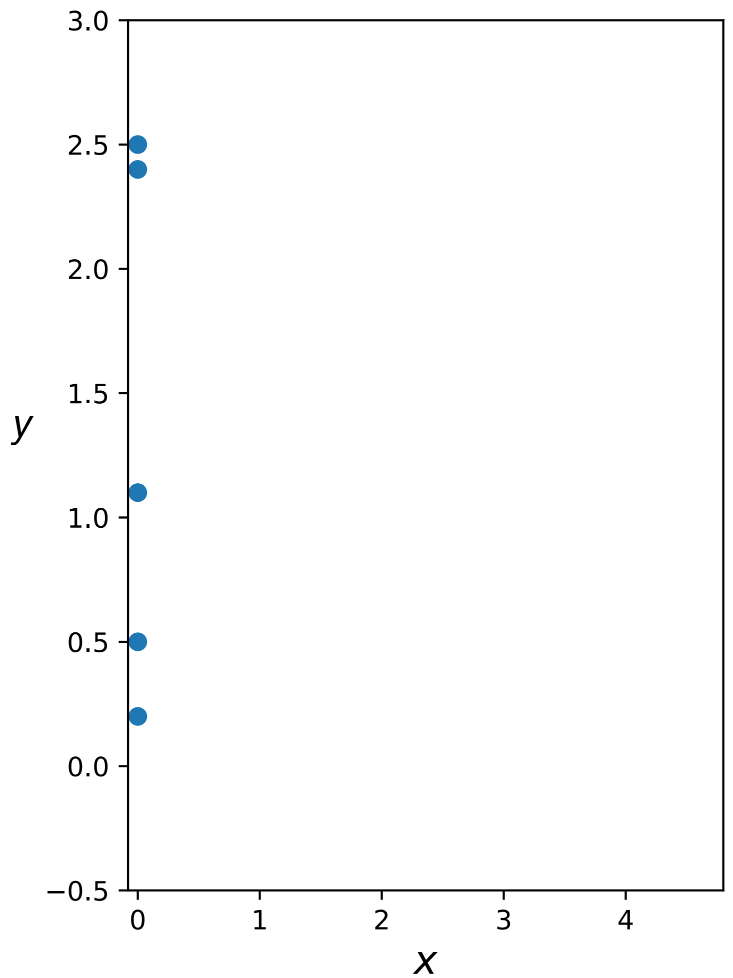
$$y_{\bullet,d} = Xw_{\bullet,d} + \epsilon_{\bullet,d}$$

$$W \sim \mathcal{N}(0, \mathrm{I})$$
$$\epsilon \sim \mathcal{N}(0, \sigma^2\mathrm{I})$$

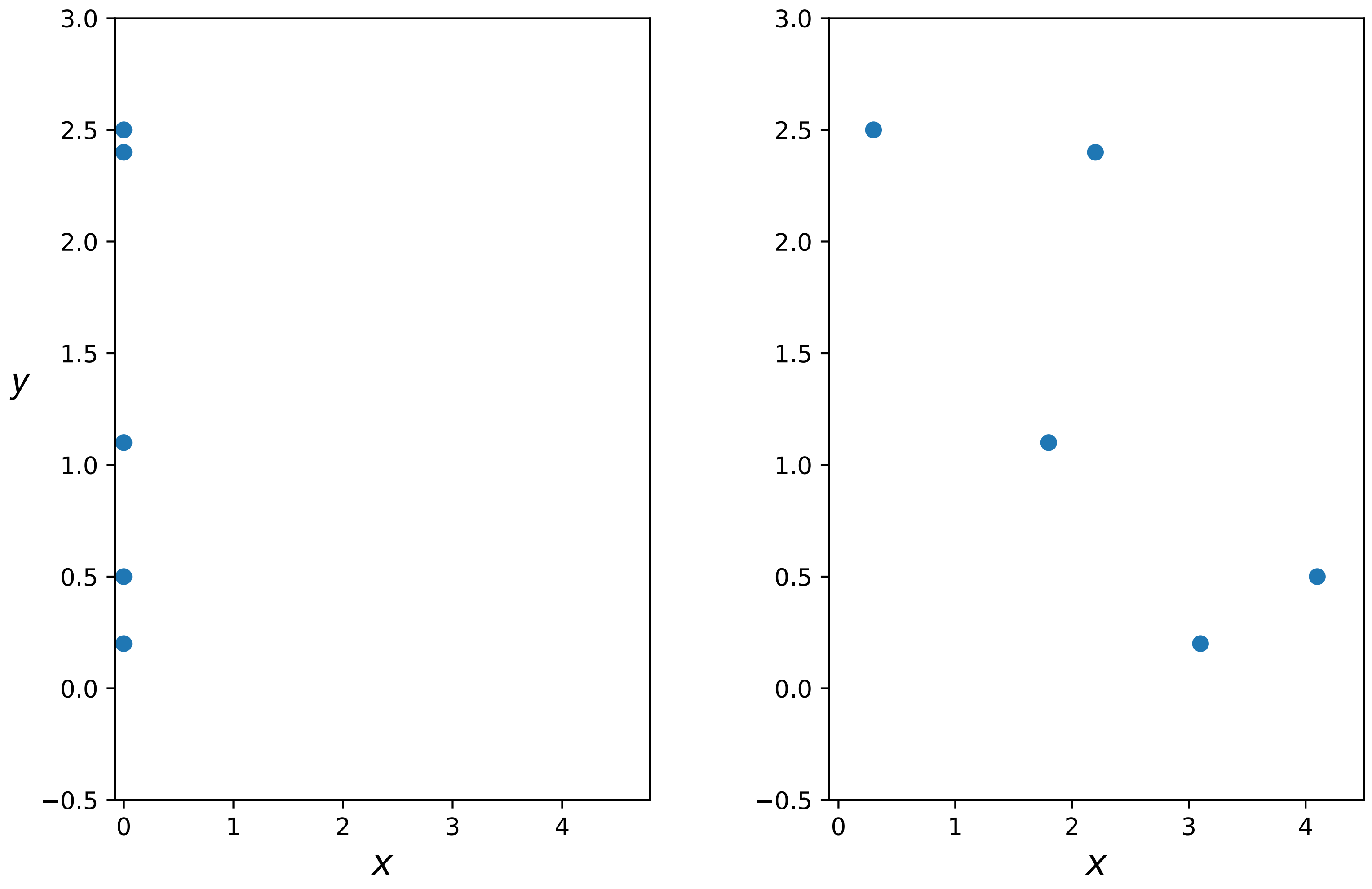
$$p(Y \mid X) = \prod_d \mathcal{N}(y_{\bullet,d} \mid 0, XX^T + \sigma^2\mathrm{I})$$

$$(XX^T)_{ij} = X_i^T X_j = k(X_i, X_j)$$

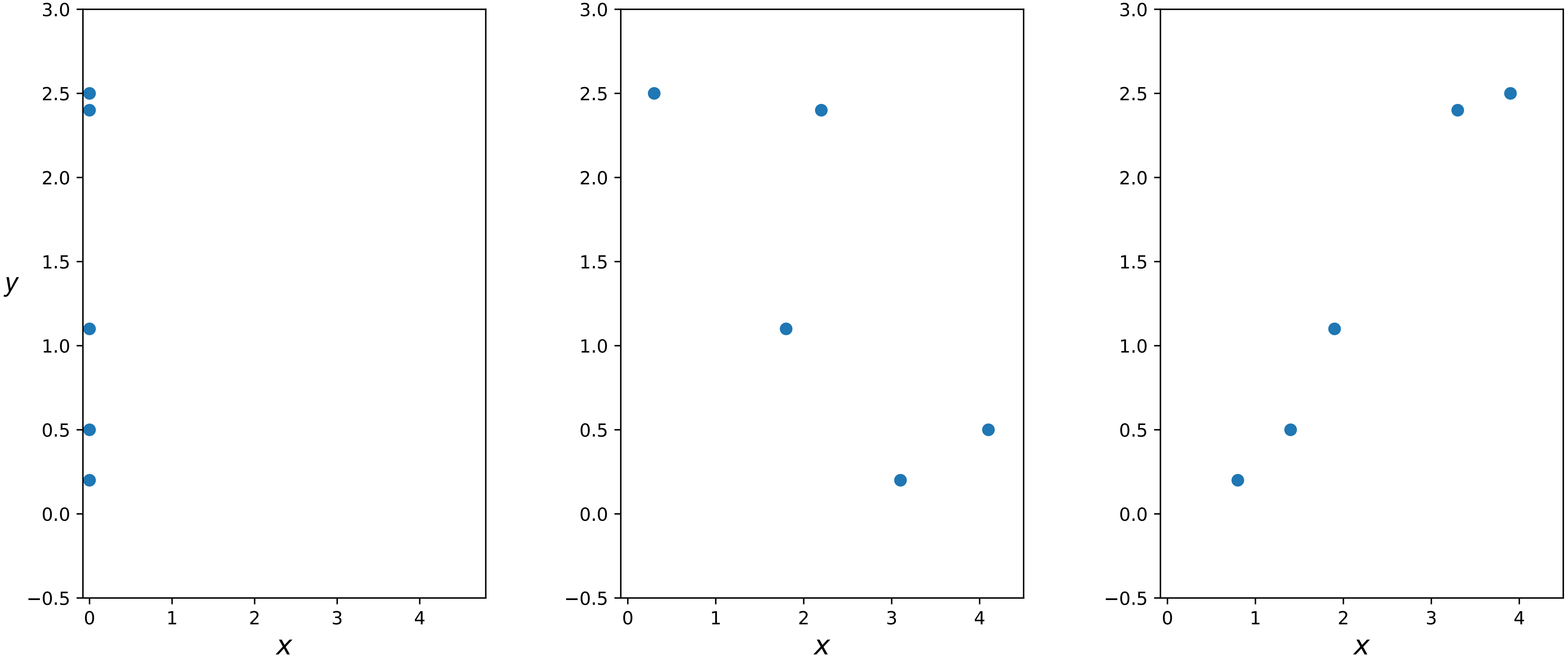
Gaussian Process Latent Variable Model



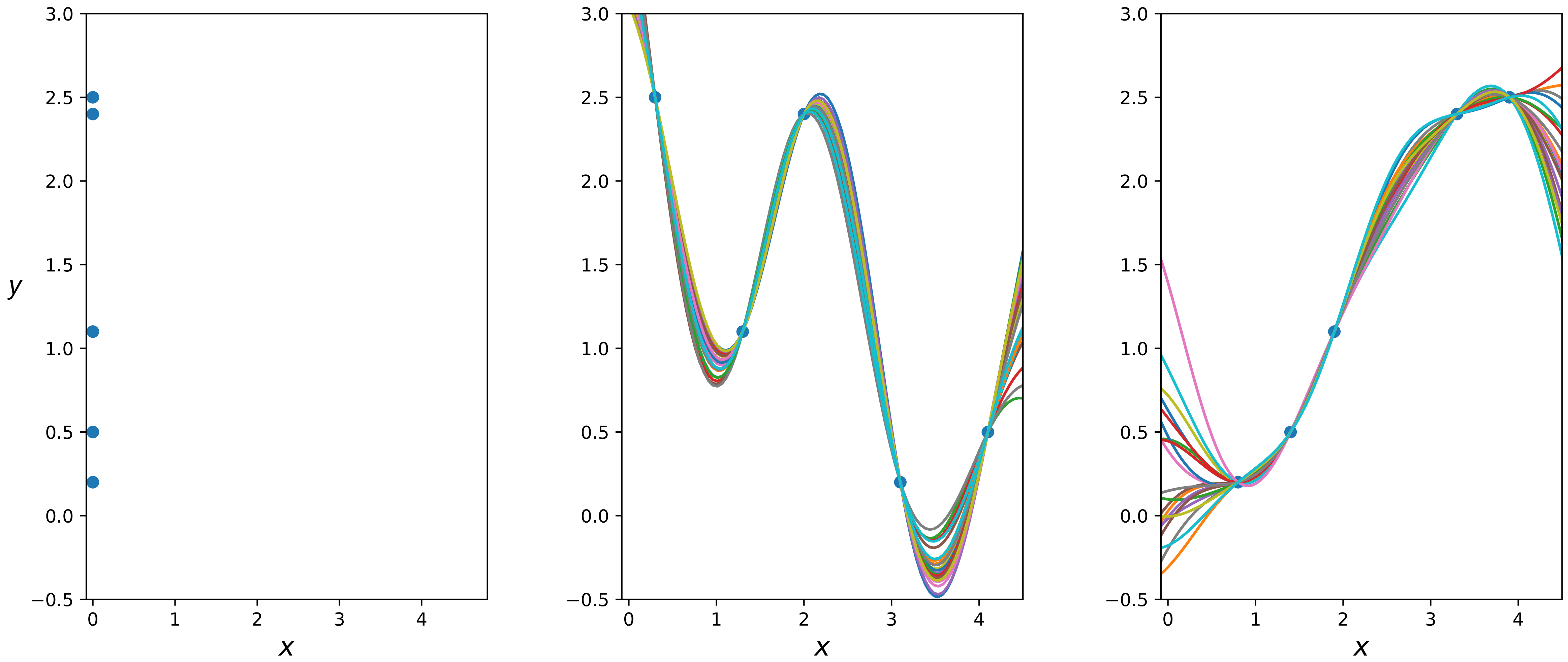
Gaussian Process Latent Variable Model



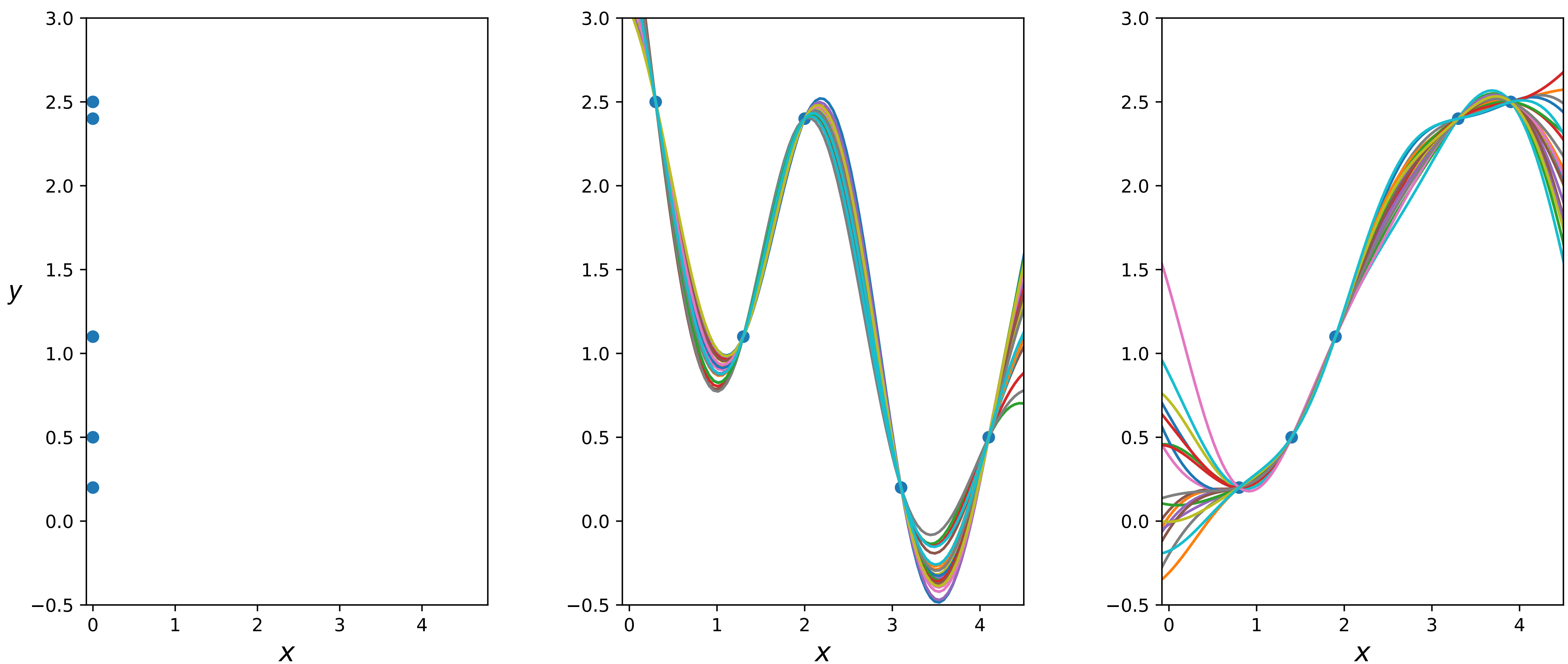
Gaussian Process Latent Variable Model



Gaussian Process Latent Variable Model



Gaussian Process Latent Variable Model



$$Y \rightarrow (K, X)$$

Applications in Finance

- CAPM

$$r_n - r_f = \beta_n (r_m - r_f) + \epsilon$$

$$\widetilde{r_n} = \beta_n \widetilde{r_m} + \epsilon$$

Applications in Finance

- CAPM $$r_n - r_f = \beta_n (r_m - r_f) + \epsilon$$

$$\widetilde{r_n} = \beta_n \widetilde{r_m} + \epsilon$$

- Given N different stocks on D days we denote the return matrix

$$R = (r_1, \ldots, r_N)^T \in \mathbb{R}^{N \times D}$$
$$R = \beta F + \epsilon$$

Applications in Finance

- CAPM $$r_n - r_f = \beta_n (r_m - r_f) + \epsilon$$

$$\widetilde{r_n} = \beta_n \widetilde{r_m} + \epsilon$$

- Given N different stocks on D days we denote the return matrix

$$R = (r_1, \ldots, r_N)^T \in \mathbb{R}^{N \times D}$$

$$R = \beta F + \epsilon \qquad p(R \mid \beta) = \prod_d \mathcal{N}(R_{\bullet,d} \mid 0, \beta\beta^T + \sigma^2 \mathrm{I})$$

Applications in Finance

- CAPM

$$r_n - r_f = \beta_n (r_m - r_f) + \epsilon$$

$$\widetilde{r_n} = \beta_n \widetilde{r_m} + \epsilon$$

- Given N different stocks on D days we denote the return matrix

$$R = (r_1, \ldots, r_N)^T \in \mathbb{R}^{N \times D}$$

$$R = \beta F + \epsilon \qquad p(R \mid \beta) = \prod_d \mathcal{N}(R_{\bullet,d} \mid 0, \beta\beta^T + \sigma^2 \mathrm{I})$$

- We can solve this model with GP-LVM and learn the covariance between stocks

- Markowitz portfolio theory

$$w_{opt} = \min_w (w^T K w - q\mu^T w)$$

Applications in Finance

- CAPM $$r_n - r_f = \beta_n (r_m - r_f) + \epsilon$$

$$\widetilde{r_n} = \beta_n \widetilde{r_m} + \epsilon$$

- Given N different stocks on D days we denote the return matrix

$$R = (r_1, \ldots, r_N)^T \in \mathbb{R}^{N \times D}$$

$$R = \beta F + \epsilon \qquad p(R \mid \beta) = \prod_d \mathcal{N}(R_{\bullet,d} \mid 0, \beta\beta^T + \sigma^2 \mathrm{I})$$

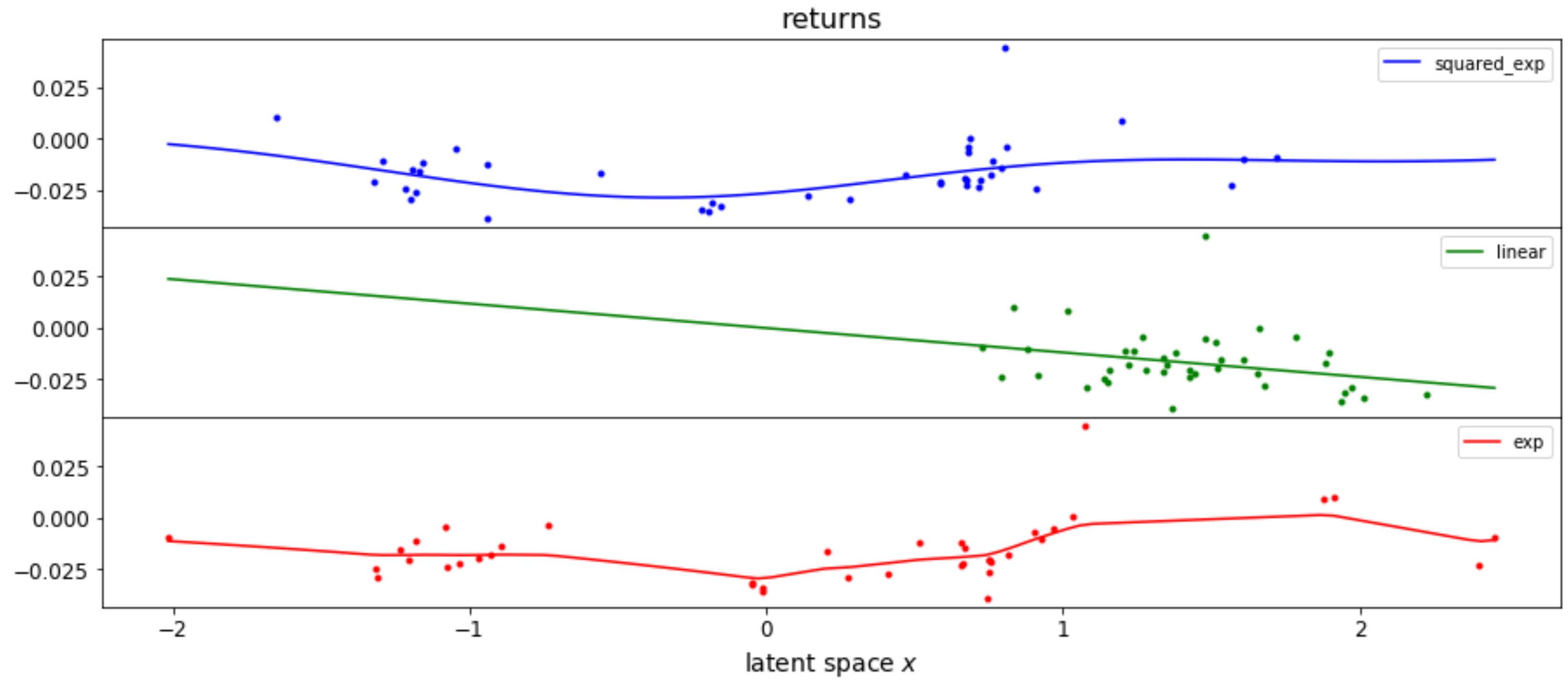
- We can solve this model with GP-LVM and learn the covariance between stocks

- Markowitz portfolio theory $$w_{opt} = \min_w (w^T K w - q\mu^T w)$$

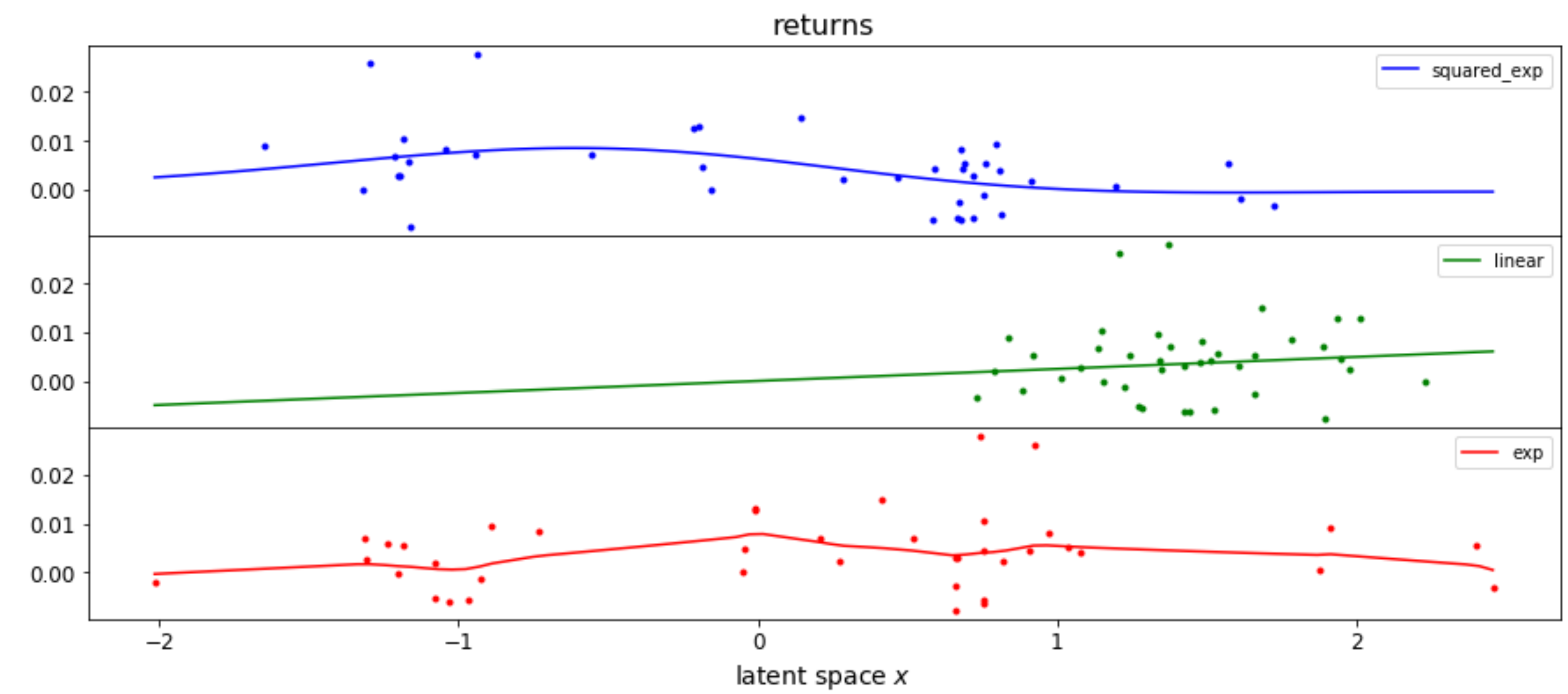
$$Y \rightarrow (K, X) \qquad\qquad R \rightarrow (K, \beta) \rightarrow w$$

Applications in Finance

Day 1

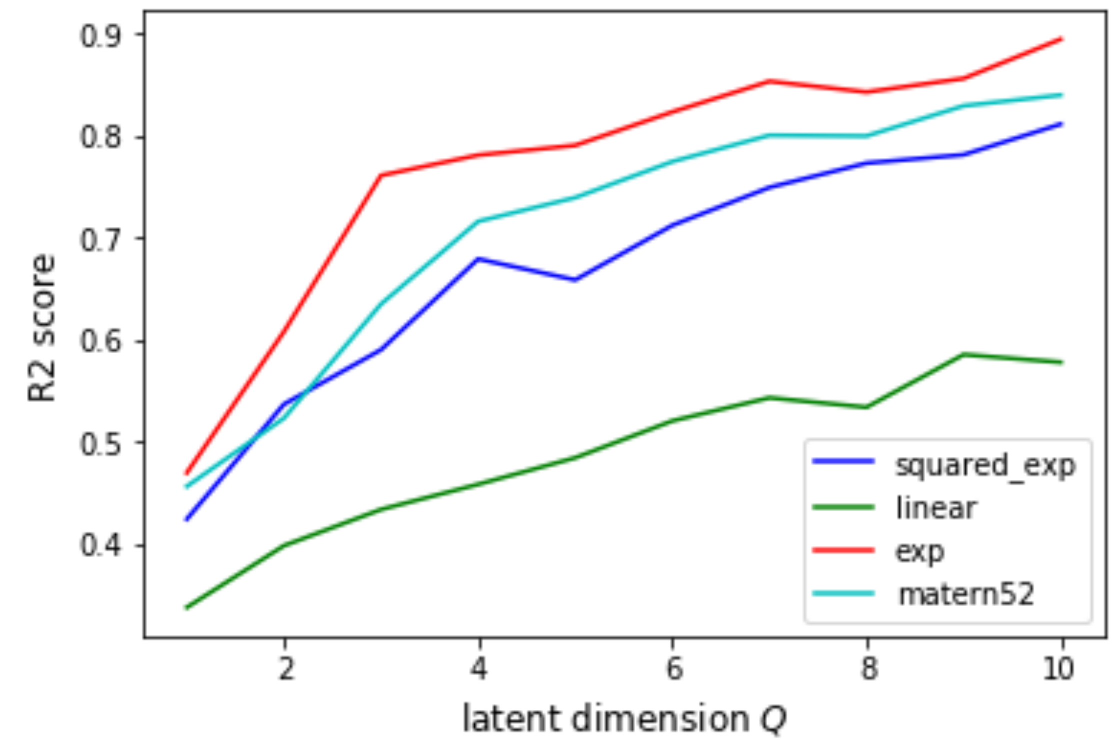


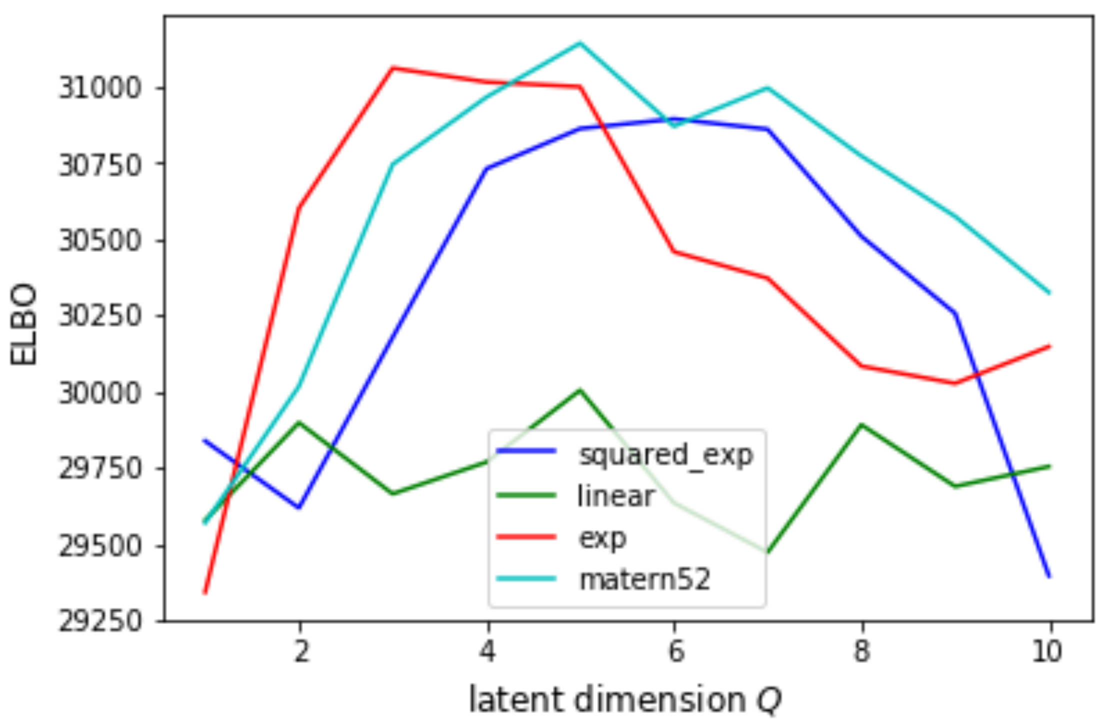
Day 2



Applications in Finance

R squared:



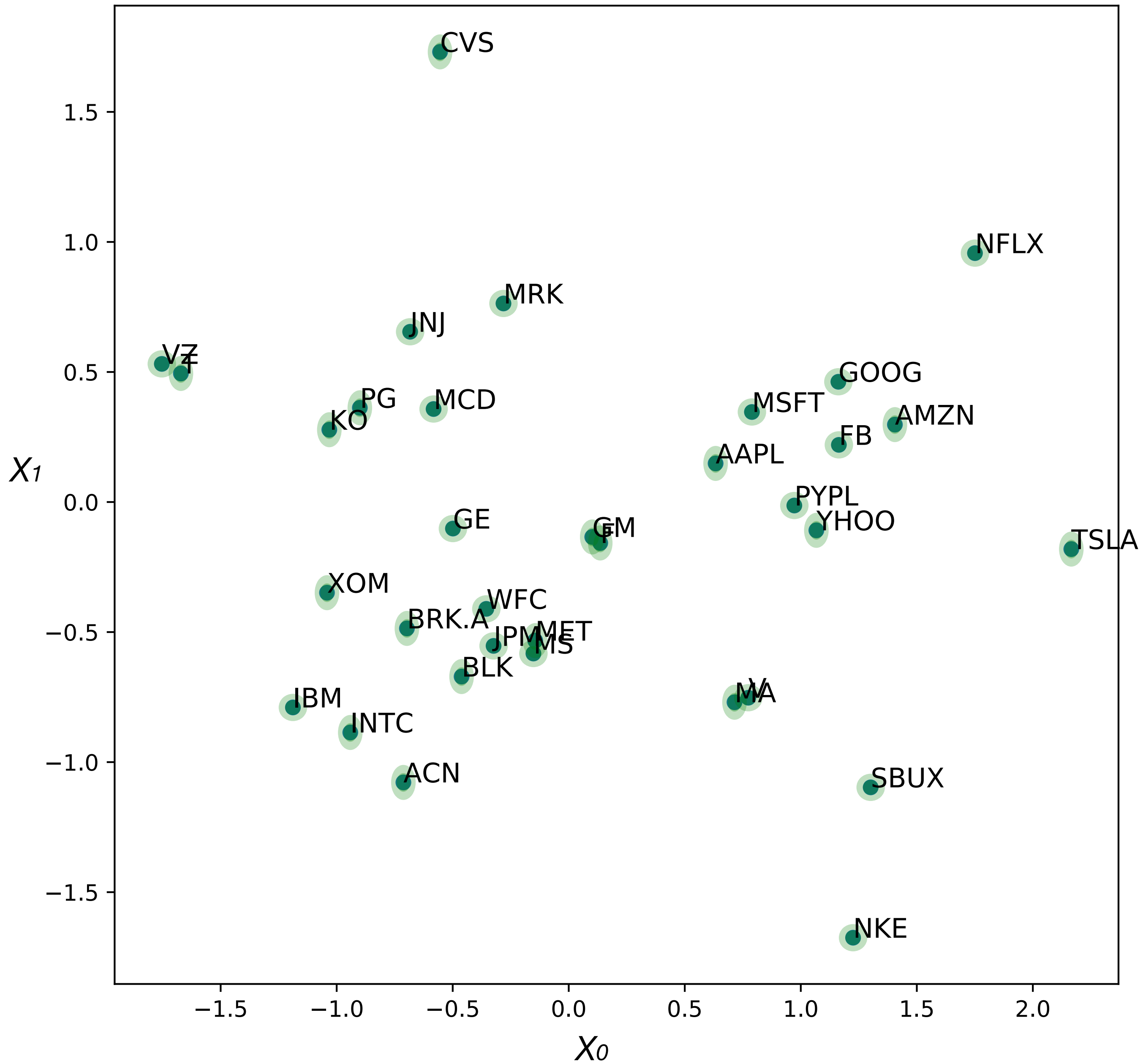


ELBO: Evidence Lower Bound

Applications in Finance

Returns for learning period of 1 year and
reweighting period of 6 months

Applications in Finance



Thanks for your attention!

Supervisor: Prof. Dr. Nils Bertschinger
Funder: Dr. h. c. Helmut O. Maucher

Thanks for your attention!

Supervisor: Prof. Dr. Nils Bertschinger
Funder: Dr. h. c. Helmut O. Maucher

Thanks for your attention!

Supervisor: Prof. Dr. Nils Bertschinger
Funder: Dr. h. c. Helmut O. Maucher