Exercise Sheet 2

Robin Schmidt
Probabilistic Inference & Learning

November 4, 2018

Gaussian Distributions

1 Structured Gaussian Models

(a) i.

The joint distribution for $\mathcal{H}_1$ is given in the lecture slides as:

$$\begin{aligned} p(x_1, x_2, x_3) &= p(x_2) \cdot p(x_1|x_2) \cdot p(x_3|x_2) & \\ x_2 &= \nu_2 & p(\nu_2) = \mathcal{N}(\nu_2; \mu_2,\, \sigma_2^2) \\ x_1 &= w_1 x_2 + \nu_1 & p(\nu_1) = \mathcal{N}(\nu_1; \mu_1,\, \sigma_1^2) \\ x_3 &= w_3 x_2 + \nu_3 & p(\nu_3) = \mathcal{N}(\nu_3; \mu_3,\, \sigma_3^2) \end{aligned}$$

This leads to the covariance matrix Σ:

$$\Sigma = \begin{bmatrix} w_1\sigma_2^2 + \sigma_1^2 & w_1\sigma_2^2 & w_1 w_3 \sigma_2^2 \\ & \sigma_2^2 & w_3\sigma_2^2 \\ & & w_3^2\sigma_2^2 + \sigma_3^2 \end{bmatrix}$$

Because there are no zeroes, matrix A and matrix B can be possible covariance matrices of $\mathcal{H}_1$.

(a) ii.

The joint distribution for $\mathcal{H}_2$ is given in the lecture slides as:

$$\begin{aligned} p(x_1, x_2, x_3) &= p(x_1) \cdot p(x_3) \cdot p(x_2|x_1, x_3) & \\ x_1 &= \nu_1 & p(\nu_1) = \mathcal{N}(\nu_1; \mu_1,\, \sigma_1^2) \\ x_3 &= \nu_3 & p(\nu_3) = \mathcal{N}(\nu_3; \mu_3,\, \sigma_3^2) \\ x_2 &= w_1 x_1 + w_3 x_3 + \nu_2 & p(\nu_2) = \mathcal{N}(\nu_2; \mu_2,\, \sigma_2^2) \end{aligned}$$

This leads to the covariance matrix Σ:

$$\Sigma = \begin{bmatrix} \sigma_1^2 & w_1\sigma_1^2 & 0 \\ & \sigma_2^2 + w_1^2\sigma_1^2 + w_3^2\sigma_3^2 & w_3\sigma_3^2 \\ & & \sigma_3^2 \end{bmatrix}$$

Because there are zeroes at $\Sigma_{1,3}$ and $\Sigma_{3,1}$, matrix C and matrix D can be possible covariance matrices of $\mathcal{H}_2$.

(b) i.

Using the joint probabilities stated before leads us to the following expressions according to the slides:

$$p(x_1, x_2, x_3) = \frac{1}{Z_1 Z_2 Z_3}\ exp\bigg(-\frac{1}{2}\bigg(\frac{x_2^2}{\sigma_2^2} + \frac{(x_1 - w_1 x_2)^2}{\sigma_1^2} + \frac{(x_3 - w_3 x_2)^2}{\sigma_3^2}\bigg)\bigg)$$

$$p(x_1, x_2, x_3) = \frac{1}{Z_1 Z_2 Z_3}\ exp\bigg(-\frac{1}{2}[x_1\ x_2\ x_3] \begin{bmatrix} \frac{1}{\sigma_1^2} & -\frac{w_1}{\sigma_1^2} & 0 \\ -\frac{w_1}{\sigma_1^2} & (\frac{1}{\sigma_2^2} + \frac{w_1^2}{\sigma_1^2} + \frac{w_3}{\sigma_3^2}) & -\frac{w_3}{\sigma_3^2} \\ 0 & -\frac{w_3}{\sigma_3^2} & \frac{1}{\sigma_3^2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\bigg)$$

The zeroes in the inverse covariance matrix Σ^{-1} at $\Sigma^{-1}_{1,3}$ and $\Sigma^{-1}_{3,1}$ imply the conditional independence of x_1 and x_3 given x_2 ($x_1 \perp\!\!\!\perp x_3 \mid x_2$). Comparing this knowledge with the four possible matrices, matrix C and matrix D seem to fulfill this condition as inverse covariance matrix Σ^{-1} for $\mathcal{H}_1$ due to the zeroes and number domains of Σ^{-1}_{ij}.

(b) ii.

Using the joint probabilities stated before leads us to the following expressions according to the slides:

$$p(x_1, x_2, x_3) = \frac{1}{Z_1 Z_2 Z_3}\ exp\bigg(-\frac{1}{2}\bigg(\frac{x_1}{\sigma_1^2} + \frac{x_3}{\sigma_3^2} + \frac{x_2 - w_1 x_1 - w_3 x_3}{\sigma_2^2}\bigg)\bigg)$$

$$p(x_1, x_2, x_3) = \frac{1}{Z_1 Z_2 Z_3}\ exp\bigg(-\frac{1}{2}[x_1\ x_2\ x_3] \begin{bmatrix} (\frac{1}{2\sigma_1^2} + \frac{w_1^2}{\sigma_2^2}) & -\frac{w_2}{\sigma_2^2} & \frac{w_1 w_3}{\sigma_2^2} \\ -\frac{w_2}{\sigma_2^2} & \frac{1}{\sigma_2^2} & -\frac{w_3}{\sigma_2^2} \\ \frac{w_1 w_3}{\sigma_2^2} & -\frac{w_3}{\sigma_2^2} & (\frac{1}{2\sigma_3^2} + \frac{w_3^2}{\sigma_2^2}) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\bigg)$$

Because there are no zeroes matrix A and matrix B could be the inverse covariance matrix Σ^{-1} for $\mathcal{H}_2$.

2 Python Exercise

(a)

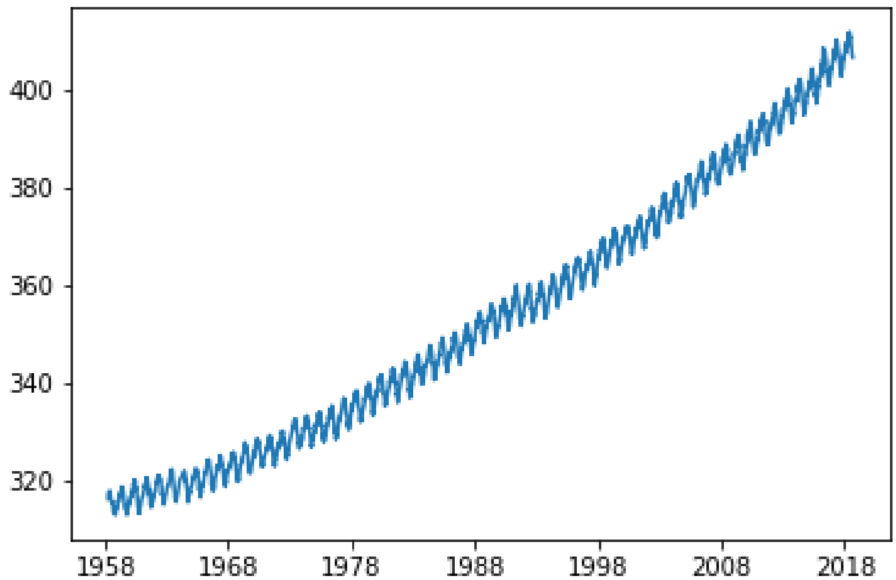

Figure 1: Plot of MaunaLoa.csv data

(b)

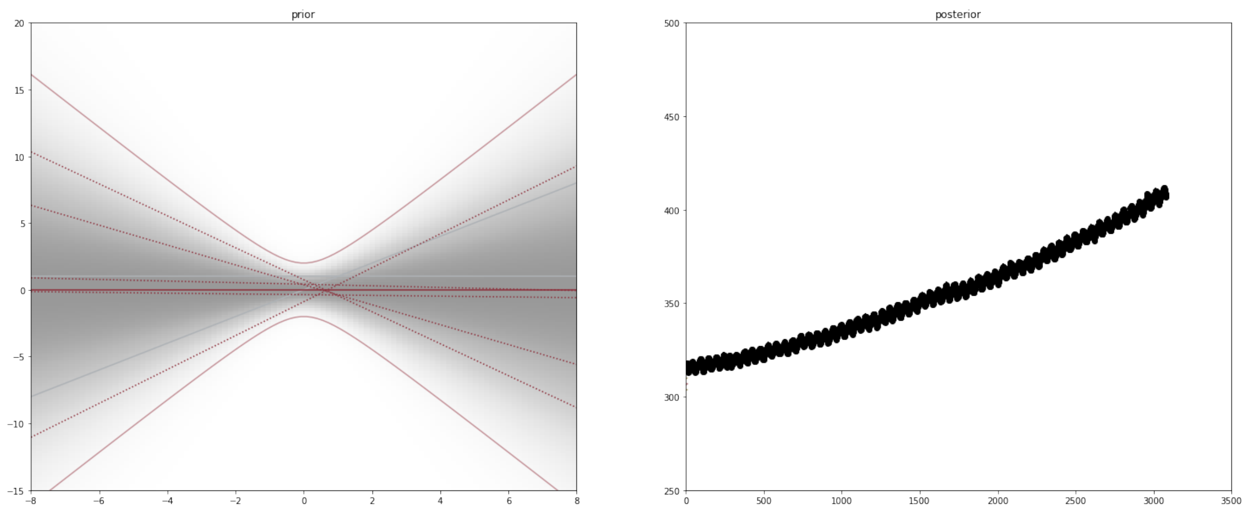

Figure 2: Posterior Gaussian prediction