Statistical Machine Learning
Assignment 2

Inez Wijnands (s4149696) & Guido Zuidhof (s4160703)

Radboud University Nijmegen

04/11/2015

The entire code listing is in a separate file. The listings shown here are merely code snippets.

1 Sequential learning

1.1 Obtaining the prior

1.

$$\tilde{\boldsymbol{\Lambda}}_{a,b} = \tilde{\boldsymbol{\Sigma}}_{a,b}^{-1} \tag{1.1}$$

$$= \left(\begin{array}{cc|cc} 60 & 50 & -48 & 38 \\ 50 & 50 & -50 & 40 \\ \hline -48 & -50 & 52.4 & -41.4 \\ 38 & 40 & -41.4 & 33.4 \end{array}\right) \tag{1.2}$$

Using the precision matrix $\tilde{\boldsymbol{\Lambda}}$ we can use equations 2.69, 2.73 and 2.75 from Bishop to obtain the mean and covariance of the conditional distribution $p([x_1, x_2]^T | x_3 = x_4 = 0)$.

$$\boldsymbol{\Sigma}_p = \boldsymbol{\Lambda}_{aa}^{-1} \tag{2.73 from Bishop}$$

$$\boldsymbol{\Lambda}_{aa} = \begin{pmatrix} 60 & 50 \\ 50 & 50 \end{pmatrix} \tag{1.3}$$

$$\boldsymbol{\Sigma}_p = \begin{pmatrix} 0.1 & -0.1 \\ -0.1 & 0.12 \end{pmatrix} \tag{1.4}$$

$$\boldsymbol{\mu}_p = \boldsymbol{\mu}_{a|b} = \boldsymbol{\mu}_a - \boldsymbol{\Lambda}_{aa}^{-1}\boldsymbol{\Lambda}_{ab}(\boldsymbol{x}_b - \boldsymbol{\mu}_b) \tag{2.75 from Bishop}$$

We can fill in this equation, since $\tilde{\boldsymbol{\mu}}$ and $\boldsymbol{x}_b$ (the second partition of $\boldsymbol{x}$) are known.

$$\boldsymbol{\mu}_p = \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 60 & 50 \\ 50 & 50 \end{pmatrix}^{-1} \begin{pmatrix} -48 & 38 \\ -50 & 40 \end{pmatrix} (\begin{pmatrix} 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 2 \end{pmatrix}) \tag{1.5}$$

$$= \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 0.1 & -0.1 \\ -0.1 & 0.12 \end{pmatrix} \begin{pmatrix} -48 & 38 \\ -50 & 40 \end{pmatrix} (\begin{pmatrix} 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 2 \end{pmatrix}) \tag{1.6}$$

$$= \begin{pmatrix} 0.8 \\ 0.8 \end{pmatrix} \tag{1.7}$$

2. Using the prior $\boldsymbol{\mu}_p$ and $\boldsymbol{\Sigma}_p$, we used the numpy-equivalent in Python for the MATLAB-function `mvnrnd` to obtain the $\boldsymbol{\mu}_t$ we used for the remainder of this assignment:

```
np.random.multivariate_normal(mu_p, sigma_p, 1)
```

This resulted in:

$$\boldsymbol{\mu}_t = \begin{pmatrix} 0.28584241 \\ 1.42626702 \end{pmatrix} \tag{1.8}$$

3. The probability density is highest at the mean (as illustrated in Figure 1.1). The density decreases quickly as both x and y change, but less so when x XOR y change. In $\boldsymbol{\Sigma}_p$, the values for the x XOR y are lower, so this is consistent with our density plot.

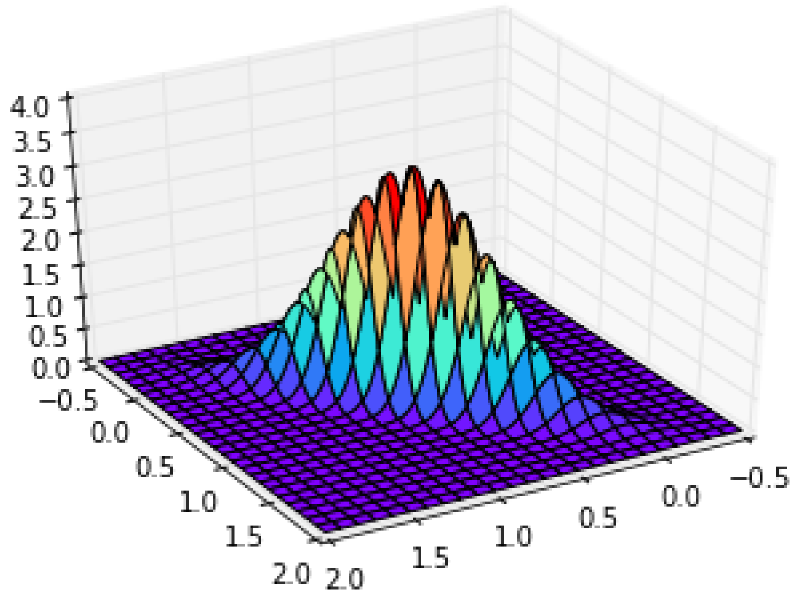

Figure 1.1: The probability density of the distribution.

1.2 Generating the data

1.

1.3 Sequential learning algorithms

1.

2.

2 The faulty lighthouse

2.1 Constructing the model

1.

2.

3.

4.

2.2 Generate the lighthouse data

1.

2.

2.3 Find the lighthouse

1.

2.

3.