Pattern Recognition practical 2

Maikel Withagen (s1867733)   Steven Bosch (s1861948)

September 23, 2015

1 Covariance matrix

1.1

Using the code given in the appendix we get the following mean vector:

$$
\begin{bmatrix} 5.8000 \\ 5.0000 \\ 6.2000 \end{bmatrix}
\tag{1}
$$

And the following covariance matrix is yielded:

$$
\begin{bmatrix} 3.2000 & 0.2500 & -0.4500 \\ 0.2500 & 2.5000 & -3.7500 \\ -0.4500 & -3.7500 & 5.7000 \end{bmatrix}
\tag{2}
$$

1.2

We computed the following probability densities:

```
1  >> mvnpdf([5;5;6], mean1, cov1)
2  ans =      0.0543
3  >> mvnpdf([3;5;7], mean1, cov1)
4  ans =    6.1287e-04
5  >> mvnpdf([4;6.5;1], mean1, cov1)
6  ans =    7.0300e-29
```

2 Covariance matrix, analytically

Every element in the covariance matrix is determined by:

$$
\sigma_{ij} = \frac{1}{n-1} \sum_{n=1}^{N} (x_{in} - \mu_i)(x_{jn} - \mu_j)
\tag{3}
$$

Since $n = 2$ for all the covariance matrices calculated below, we leave that factor out (since $\frac{1}{2-1} = 1$).

2.1

Element 1,1:

$$\sigma_{1,1} = (a - \frac{a+c}{2})^2 + (c - \frac{a+c}{2})^2 \tag{4}$$

$$= (\frac{a-c}{2})^2 + (\frac{c-a}{2})^2 \tag{5}$$

$$= \frac{1}{4}(a-c)^2 + \frac{1}{4}(c-a)^2 \tag{6}$$

$$= \frac{1}{2}(a-c)^2 \tag{7}$$

Similarly for $\sigma_{2,2}$ we get $\frac{1}{2}(b-d)^2$. For the other two elements, $cov(1,2)$ and $cov(2,1)$ which are the same, we get the following:

$$\sigma(1,2) = ((a - \frac{a+c}{2})(b - \frac{b+d}{2}) + (c - \frac{a+c}{2})(d - \frac{b+d}{2})) \tag{8}$$

$$= (\frac{a-c}{2})(\frac{b-d}{2}) + (\frac{c-a}{2})(\frac{d-b}{2}) \tag{9}$$

$$= \frac{(a-c)(b-d) + (c-a)(d-b)}{4} \tag{10}$$

$$= \frac{2ab - 2ad - 2bc + 2cd}{4} \tag{11}$$

$$= \frac{1}{2}(a-c)(b-d) \tag{12}$$

This results in the following covariance matrix:

$$\begin{bmatrix} \frac{1}{2}(a-c)^2 & \frac{1}{2}(a-c)(b-d) \\ \frac{1}{2}(a-c)(b-d) & \frac{1}{2}(b-d)^2 \end{bmatrix} \tag{13}$$

2.2

Element 1,1:

$$\sigma_{1,1} = (a + k - \frac{a+c+2k}{2})^2 + (c + k - \frac{a+c+2k}{2})^2 \tag{14}$$

$$= (\frac{a-c}{2})^2 + (\frac{c-a}{2})^2 \tag{15}$$

$$= \frac{1}{4}(a-c)^2 + \frac{1}{4}(c-a)^2 \tag{16}$$

$$= \frac{1}{2}(a-c)^2 \tag{17}$$

As we can see, the ks are all removed in the simplification, due to writing sums like $a + k$ as the fraction $\frac{2a+2k}{2}$ and subtracting fractions with $2k$ in the numerator, resulting in the ks being removed. Again similarly for $\sigma_{2,2}$ we get $\frac{1}{2}(b-d)^2$. For the other two elements, $cov(1,2)$ and $cov(2,1)$ we get the following (the ks are removed in the simplification process again):

2

$$\sigma(1,2) = ((a+k - \frac{a+c+2k}{2})(b+k - \frac{b+d+2k}{2}) + (c+k - \frac{a+c+2k}{2})(d+k - \frac{b+d+2k}{2})) \quad (18)$$

$$= (\frac{a-c}{2})(\frac{b-d}{2}) + (\frac{c-a}{2})(\frac{d-b}{2}) \quad (19)$$

$$= \frac{(a-c)(b-d) + (c-a)(d-b)}{4} \quad (20)$$

$$= \frac{2ab - 2ad - 2bc + 2cd}{4} \quad (21)$$

$$= \frac{1}{2}(a-c)(b-d) \quad (22)$$

Thus in the end we get the same matrix as in the previous section:

$$\begin{bmatrix} \frac{1}{2}(a-c)^2 & \frac{1}{2}(a-c)(b-d) \\ \frac{1}{2}(a-c)(b-d) & \frac{1}{2}(b-d)^2 \end{bmatrix} \quad (23)$$

2.3

Element 1,1:

$$\sigma_{1,1} = (ak - \frac{ak+ck}{2})^2 + (ck - \frac{ak+ck}{2})^2 \quad (24)$$

$$= (\frac{ak-ck}{2})^2 + (\frac{ck-ak}{2})^2 \quad (25)$$

$$= \frac{1}{4}(ak-ck)^2 + \frac{1}{4}(ck-ak)^2 \quad (26)$$

$$= \frac{1}{2}(ak-ck)^2 \quad (27)$$

$$= \frac{1}{2}k^2(a-c)^2 \quad (28)$$

Similarly for $\sigma_{2,2}$ we get $\frac{1}{2}k^2(b-d)^2$.

For the other two elements, $cov(1,2)$ and $cov(2,1)$ we get the following:

$$\sigma(1,2) = ((ak - \frac{ak+ck}{2})(bk - \frac{bk+dk}{2}) + (ck - \frac{ak+ck}{2})(dk - \frac{bk+dk}{2})) \quad (29)$$

$$= (\frac{ak-ck}{2})(\frac{bk-dk}{2}) + (\frac{ck-ak}{2})(\frac{dk-bk}{2}) \quad (30)$$

$$= \frac{k^2(a-c)(b-d) + k^2(c-a)(d-b)}{4} \quad (31)$$

$$= \frac{2k^2(ab - ad - bc + cd)}{4} \quad (32)$$

$$= \frac{1}{2}k^2(a-c)(b-d) \quad (33)$$

$$\begin{bmatrix} \frac{1}{2}k^2(a-c)^2 & \frac{1}{2}k^2(a-c)(b-d) \\ \frac{1}{2}k^2(a-c)(b-d) & \frac{1}{2}k^2(b-d)^2 \end{bmatrix} \quad (34)$$

3 2D Gaussian pdf, Mahalanobis distance

3.1

Using the code given in the appendix, we generated the following plot:
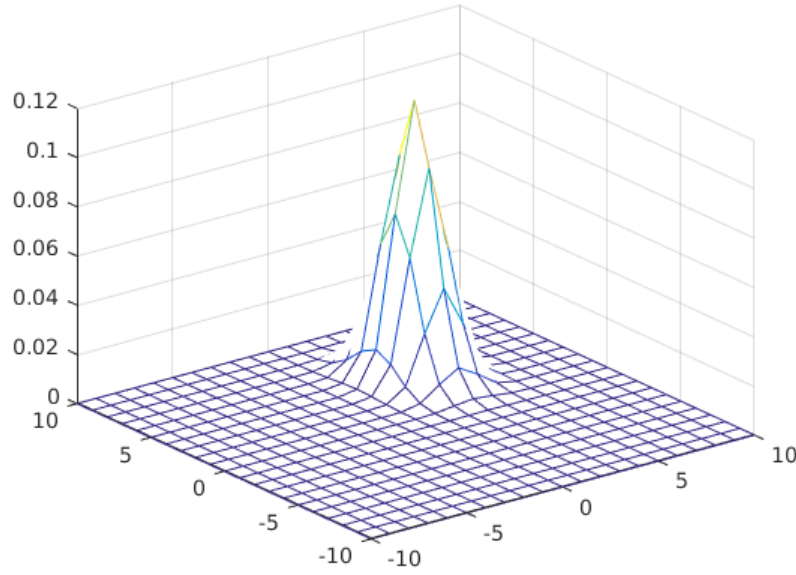
Figure 1: Two-dimensional Gaussian pdf plotted on a mesh.

3.2 2D Gaussian pdf, Mahalanobis distance

We determined the Mahalanobis distances using the following definition:

$$r^2 = (x - \mu)\Sigma^{-1}(x - \mu)' \tag{35}$$

This yielded the following distances for the given points:

```
1  >> ([10   10]-mean)*inv(cov)*([10   10]-mean).'
2  ans =      67
3  >> ([0      0]-mean)*inv(cov)*([0      0]-mean).'
4  ans =      17
5  >> ([3      4]-mean)*inv(cov)*([3      4]-mean).'
6  ans =       0
7  >> ([6      8]-mean)*inv(cov)*([6      8]-mean).'
8  ans =      17
```

3.3 Independent identically distributed random binary variables

The code in the appendix yields the following plot:

4 2D Gaussian pdf, Mahalanobis distance

4.1

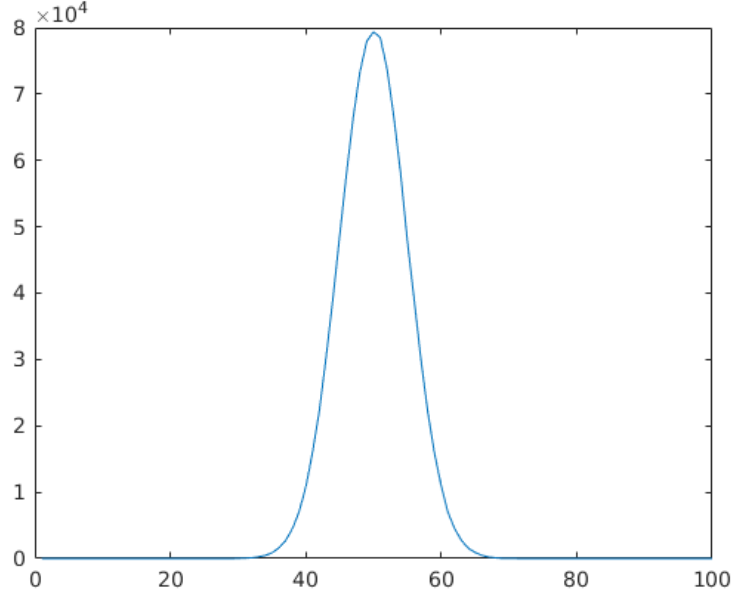Using the code given in the appendix, we generated the following plot:

4

Figure 2: Two-dimensional Gaussian pdf plotted on a mesh.

This plot resembles the binomial distribution with a sequence of 100 experiments and a probability of 0.5. Each person that tosses a coin 100 times can be viewed as a random draw from the binomial distribution with a probability P of 0.5 because of the 50/50 chances of the coin flip. The 1000000 people together form 1000000 draws from this binomial distribution, and by plotting them, an approximation of the distribution is displayed.

By tossing a coin 100 times (N) with a 50% chance to go forward (P), the theoretical mean would be N*P or 100*0.5 = 50. The variance would be N*P(1-P) = 100*0.5(1-0.5) = 25.

4.2 Multivariate normal density, discriminant functions, minimum error rate classification, unequal priors, dichotomizer

4.3 Naive Bayesian rule

Following the Naive Bayes rules as shown on the slides (so calculating the ratio):

a "We offer our dear customers a wide selection of classy watches"

$$\frac{P(spam|Customers, Watches)}{P(non-spam|Customers, Watches)} = \frac{p(customers|spam)}{p(customers|non-spam)} \frac{p(watches|spam)}{p(watches|non-spam)} \frac{p(spam)}{p(non-spam)} \tag{36}$$

$$= \frac{0.005 * 0.0003 * 0.9}{0.0001 * 0.000004 * 0.1} \tag{37}$$

$$= 33750 \tag{38}$$

So with a ratio of 33750:1 (spam:non-spam) it is very likely that the message can be classified as spam (a probability of approximately 0.99997037037).

b "Did you have fun on vacation? I sure did!"

$$\frac{P(spam|Fun,Vacation)}{P(non-spam|Fun,Vacation)} = \frac{p(fun|spam)}{p(fun|non-spam)}\frac{p(vacation|spam)}{p(vacation|non-spam)}\frac{p(spam)}{p(non-spam)} \tag{39}$$

$$= \frac{0.00015*0.00025*0.9}{0.0007*0.00014*0.1} \tag{40}$$

$$= 3.4439 \tag{41}$$

So with a ratio of 3.4439:1 (spam:non-spam) it is likely that the message can be classified as spam (with a probability of approximately 0.7096315224).

Appendix

Code for assignment 1

```
1  v1 = [4,5,6];
2  v2 = [6,3,9];
3  v3 = [8,7,3];
4  v4 = [7,4,8];
5  v5 = [4,6,5];
6
7  m1 = [v1;v2;v3;v4;v5];
8
9  mean1 = [mean(m1(:,1)); mean(m1(:,2)); mean(m1(:,3))];
10
11 cov1 = cov(m1);
12
13 mean1
14 cov1
```

Code for assignment 3

```
1  % Generate a two-dimensional Gaussian pdf with a mean [3 4] and covariance matrix
2  % 1 0
3  % 0 2
4  mean = [3 4]
5  cov = [1 0; 0 2]
6  dist = gmdistribution(mean,cov)
7
8  inputmatrix = [0 0];
9
10 counter = 1;
11 for i = -10:10
12     for j = -10:10
13         inputmatrix(counter,:) = [i j];
14
15         counter = counter + 1;
16     end
17 end
```

```
18  % 1.  Plot  this  function  on  [−10  10]  x  [−10  10] Using  the  mesh  function.
19  output = reshape(pdf(dist , inputmatrix ),[21 ,21])
20  mesh( −10:10, −10:10, output ) %
```

Code for assignment 4

```
1   output = zeros(100,1);
2   for  i=1:1000000
3       step = 0;
4       for  j = 1:100
5           if rand > 0.5
6               % Move forward
7               step = step + 1;
8           end
9       end
10      output(step) = output(step) + 1;
11  end
12
13  plot(output)
```