

**ECE 8110: INTRODUCTION TO MACHINE LEARNING AND PATTERN
RECOGNITION**

HOMEWORK # 1

GAUSSIAN DISTRIBUTIONS AND MAXIMUM LIKELIHOOD DECODING

Vira Oleksyuk

01/31/2014

1. Denote a one-dimensional Gaussian random variable (GRV) with a mean of 1 and a variance of 2 as GRV[1,2]. Generate 10,000 data points for each of two GRVs: GRV[1,1] and GRV[-1,1]. Compute the error rate using an optimal maximum likelihood decoder with a threshold set at 0. Repeat for a threshold set at -0.5 and 0.5. Do these error rates match the predicted values?

It may seem trivial, yet it is very important to correctly generate random variables. To generate a row vector of length 10, containing Gaussian distributed numbers with mean 5 and variance 2, I can type ^[1] $R = \text{random}('norm', 5, \text{sqrt}(2), 1, 10)$ in Matlab. Function *randn()* generates samples of a Gaussian distributed random variable with mean 0 and variance 1. To obtain a mean other than zero, I had to add or subtract a constant from the generated vector. To obtain a variance other than one, multiply the generated vector by the standard deviation (square root of the variance). However, *randn* executes significantly faster than *random*, so this would be the preferred function in longer simulations.

```
tic
GRV=sqrt(2).*randn(1,10000) + 1;
Toc
Elapsed time is 0.000155 seconds.
tic
GRV=random('norm',1,sqrt(2),1,10000);
Toc
Elapsed time is 1.570847 seconds.
```

Two random variables were generated as specified.

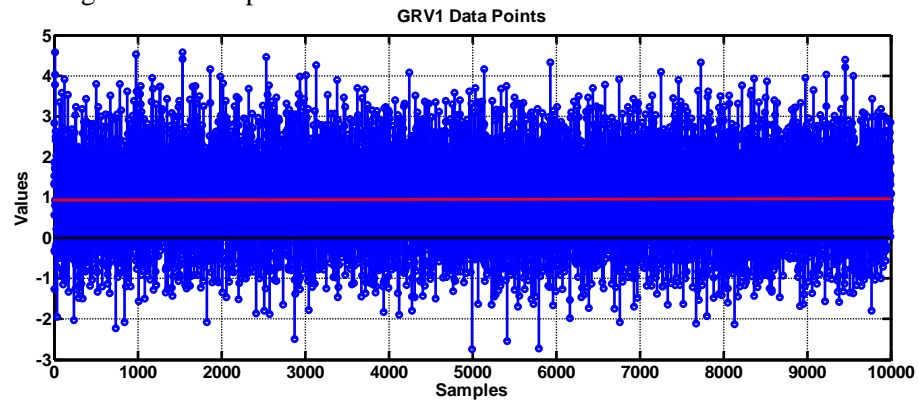


Figure 1. Plot of values for first Gaussian random variable GRV1[1,1].

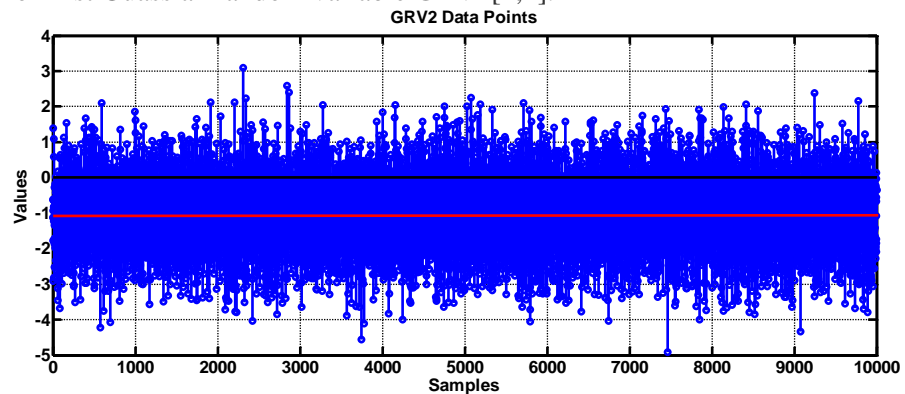


Figure 2. Plot of values for second Gaussian random variable GRV2[-1,1].

Horizontal red lines on Fig.1 and Fig.2 shows mean value.

Fig.3 and Fig.4 are histograms for each random variable with corresponding probability distribution curves in red.

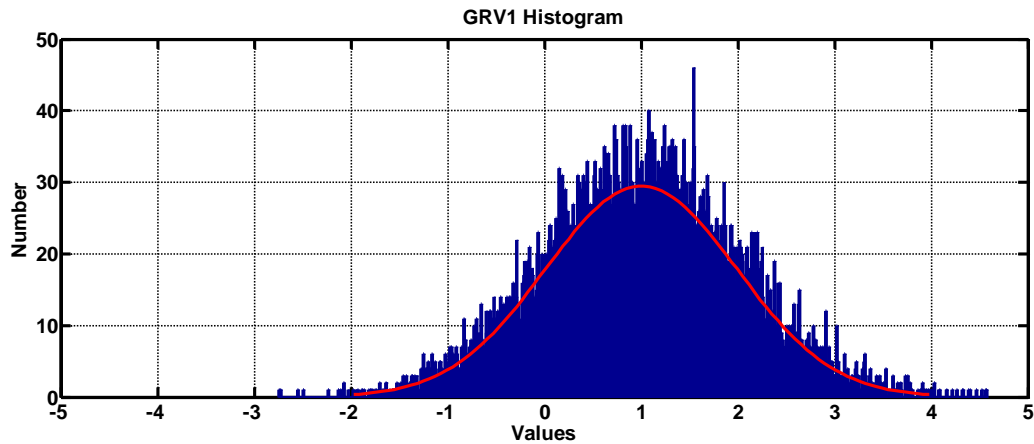


Figure 3. Histogram of GRV1

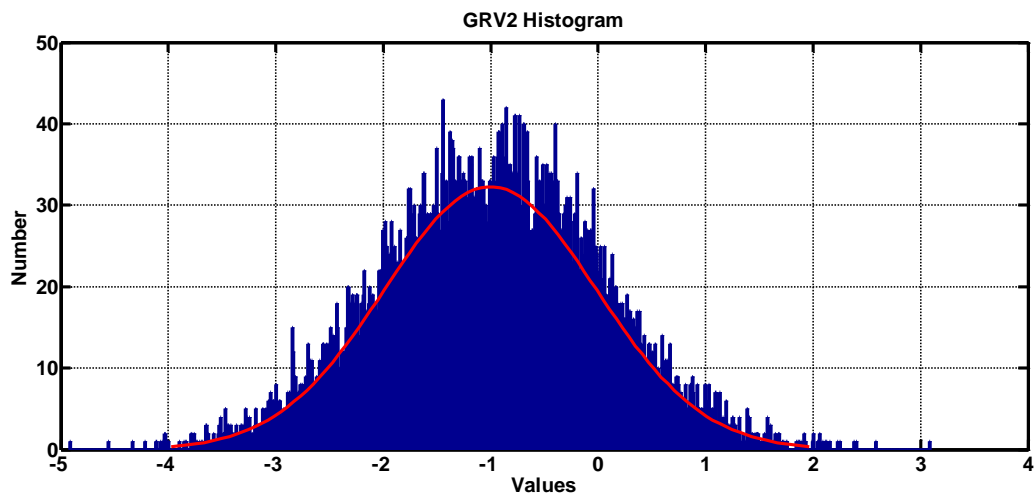


Figure 4. Histogram of GRV2

In order to calculate theoretical error rate I used *cdf* function.

```
ErrorRate_0=(1-cdf('Normal',0,-1,1) + cdf('Normal',0,1,1)) / 2; %Threshold 0
```

For optimal maximum likelihood decoder error rate calculation I did following:

```
TH1=0;
errorRate_TH0=(length(find(GRV1<TH1))+length(find(GRV2>TH1)))/(length(GRV1)+...
length(GRV2));
```

Similarly were done calculations for threshold 0.5 and -0.5.

Tab.1 contains results for error rates calculations for threshold 0, as well as for thresholds -0.5 and 0.5.

Table 1. Comparison of results from error rate calculations (one trial)

Threshold	Theoretical Error Rate	Calculated Error Rate with optimal maximum likelihood decoder
0	0.15866	0.16235
-0.5	0.18767	0.18780
0.5	0.18767	0.18855

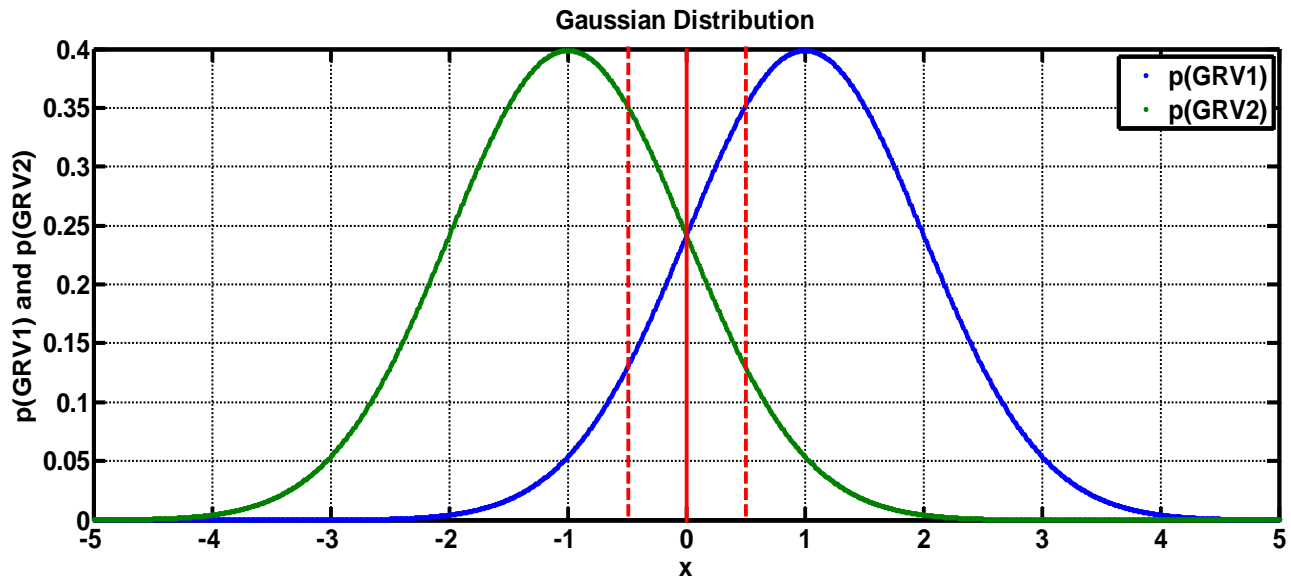


Figure 5. Probability distributions for both random variables with indicated thresholds.

As we can see from Tab.1, theoretical and corresponding calculated error rates are very similar to each other. Error rates for thresholds -0.5 and 0.5 are theoretically the same (due to symmetry of pdf's) and greater than error rates for threshold 0. This result is also supported by Fig.5. Overall area of "thresholded" parts of pdfs will increase with sliding the threshold to the left or right from zero. It is similar to giving less weight to one of the distributions. In this exercise, posteriors for both random variables were assigned to be equal = 0.5. Optimal threshold is on intersection of two distributions at zero.

As a summary, error rate will change with applications of different thresholds.

2. Generate 1,000 data points for each of two GRVs: $\text{GRV}[1, \sigma_1^2]$ and $\text{GRV}[-1, \sigma_2^2]$. Set $\sigma_2^2 = 1$ and vary σ_1^2 over the range $[0.1, 2]$. Plot the error rate as a function of σ_1^2 .

In this exercise, I built two random variables with 1000 data points. Variance was varied for GRV1 over the range $[0.1, 2]$. Distribution plot is presented on Fig. 6.

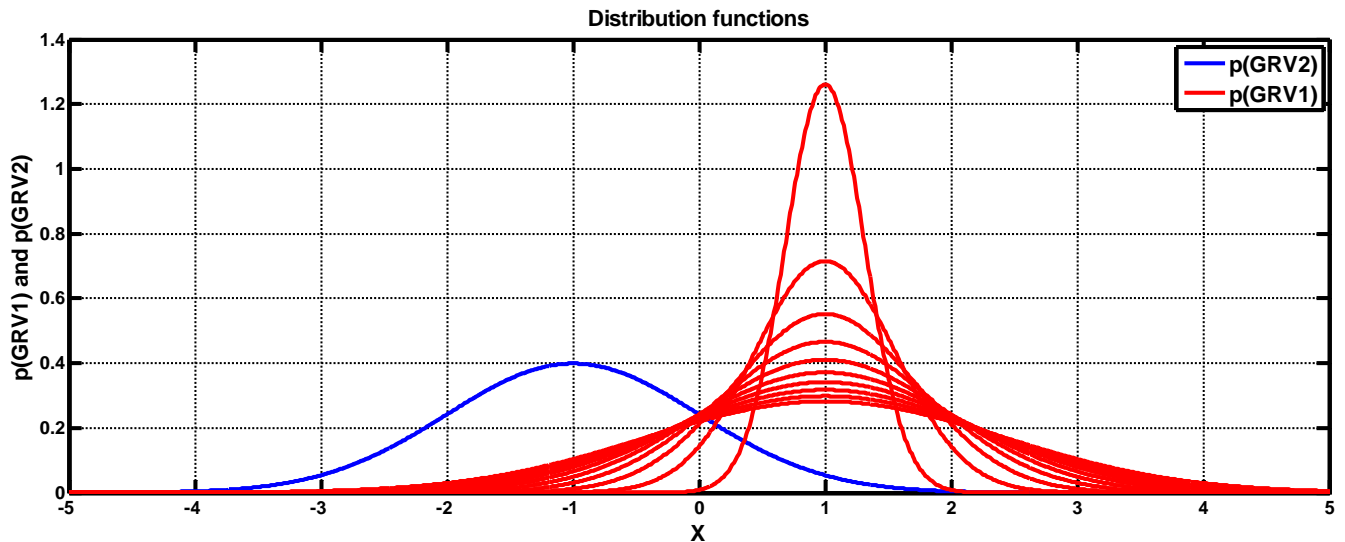


Figure 6. Distribution plots for two Gaussian random variables.

Table 2 contains results for error rate calculation for threshold=0 over variance range $[0.1, 2]$.

Table 2. Results from error rate calculations

Variance GRV1	Variance GRV2	Calculated Error Rate with optimal maximum likelihood decoder
0.1000	1	0.0740
0.3111	1	0.0940
0.5222	1	0.1160
0.7333	1	0.1390
0.9444	1	0.1480
1.1556	1	0.1625
1.3667	1	0.1790
1.5778	1	0.1870
1.7889	1	0.2000
2.0000	1	0.1970

Fig.7 shows increase in error with increase of GRV1 variance.

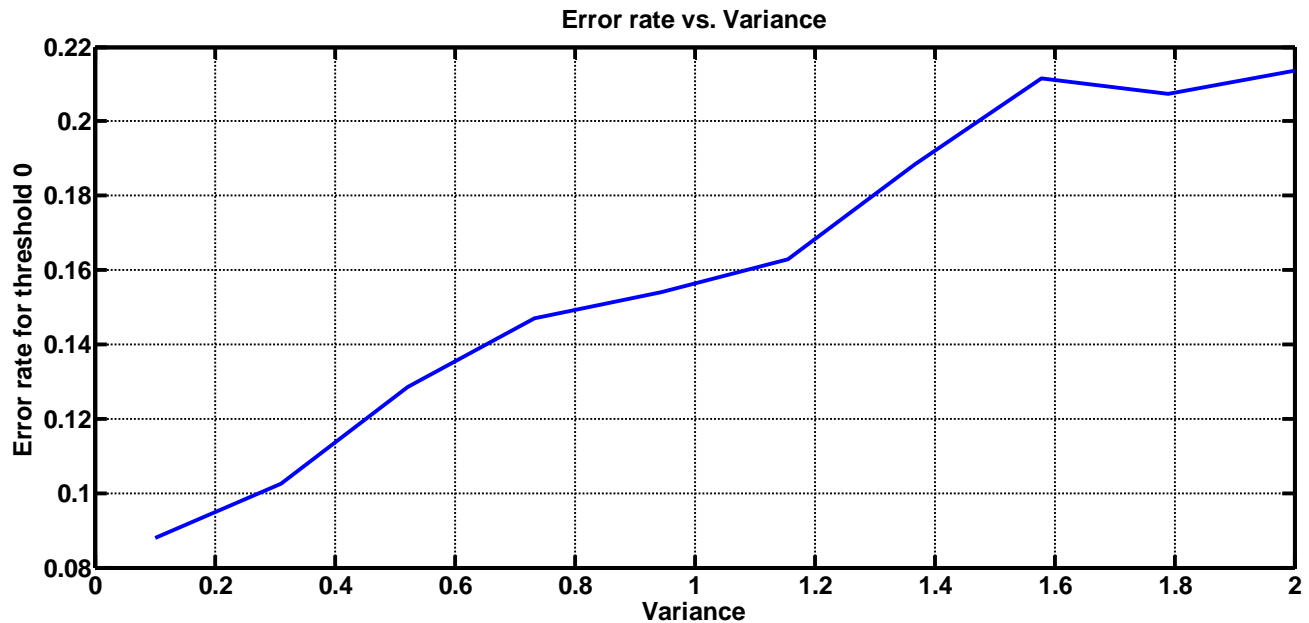


Figure 7. Error rate vs. Variance

When variance is small, separation of two pdfs is clearer than when variance increases and causes large overlaps. So later, when we will try to differentiate GRVs from each other using a threshold, this enlarged variance will cause more decision mistakes.

3. Generate a 2D GRV with a mean of $[-1,1]$ and an identity covariance matrix. Generate a second GRV with a mean of $[1,1]$ and an identity covariance matrix. Compute the error rate for the optimal maximum likelihood decoder. Does this match the theoretical prediction?

Two dimensional GRVs were generated as following:

```
GRV 2d 1 = mvnrnd(Mu 2d 1, SigmaSq1, 1000);
```

XY data for bot GRV is plotted on Fig.8 with a threshold line in red at $x=0$. We can clearly see how threshold line divides data points. From this figure, I choose threshold to be 0 or $((1+(-1))/2=0)$.

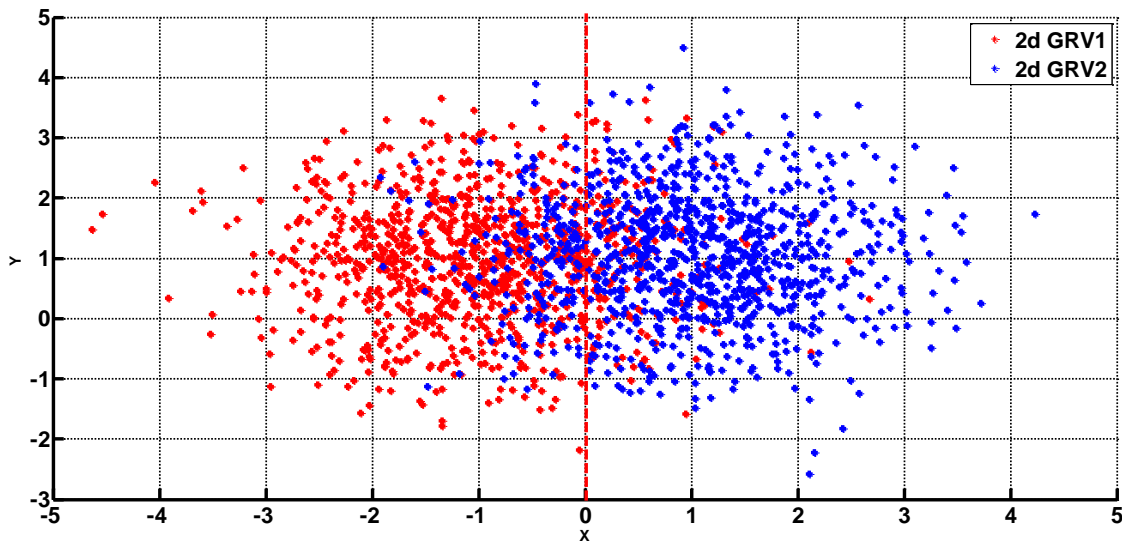


Figure 8. Data points plot for two 2d GRVs.

Next figure (Fig. 9) shows combine probability distributions and a corresponding contour plot for two 2dGRVs. They look like two overlapping peaks.

Error rates were calculated similarly to previous sections. Results are presented in Table 3.

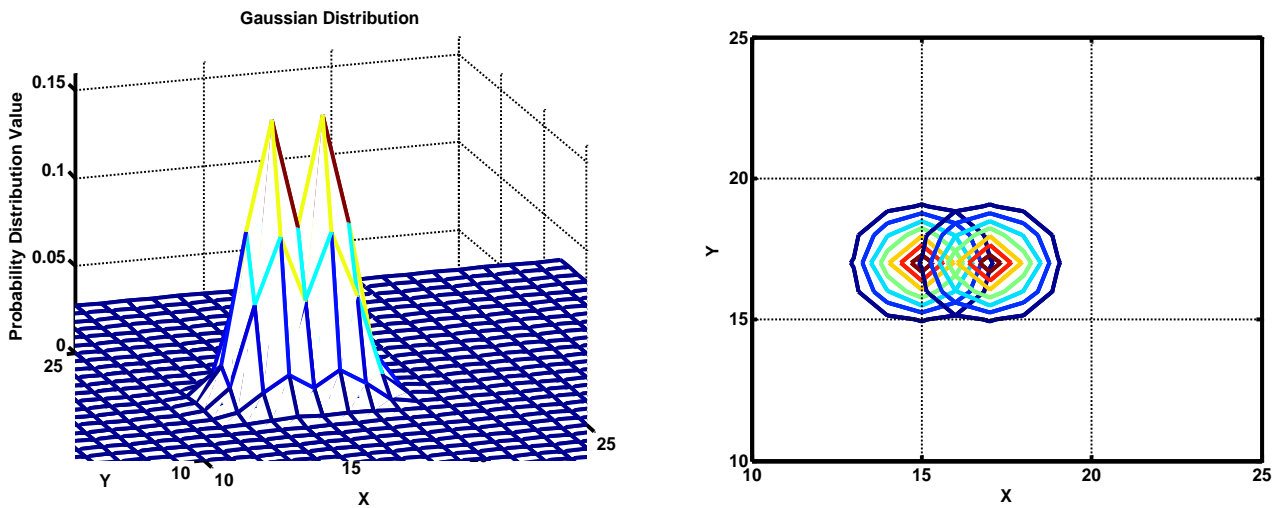


Figure 9. Gaussian distribution with a support region for two 2d GRVs

Table 3. Comparison of results from error rate calculations (one trial)

Threshold	Theoretical Error Rate	Calculated Error Rate
0	0.1587	0.1625

Calculated value of error rate using optimal maximum likelihood decoder is very close to theoretically estimated error rate.

4. Generate a plot of the support region for the first GRV in (3) as you vary the covariance. Demonstrate how the shape of the support region changes depending on the nature of the covariance matrix (replicating the slides in lecture no. 2).

In this section presented Gaussian distributions with corresponding support regions. Covariance matrices were varying, and comments were made to each figure.

$$C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

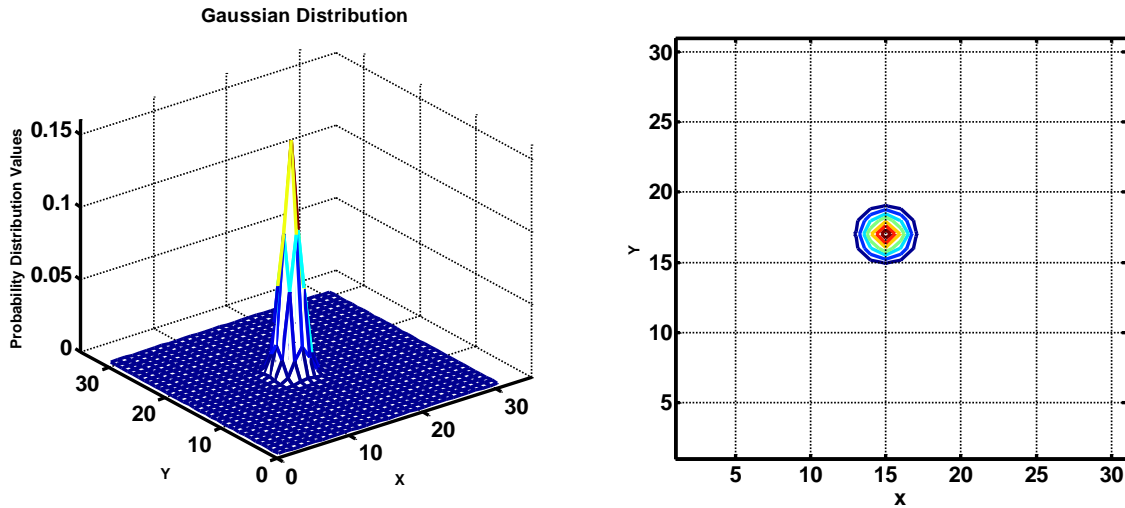


Figure 10. Gaussian distribution with a support region for two 2d GRVs (small circular region)

$$C = \begin{bmatrix} 20 & 0 \\ 0 & 20 \end{bmatrix}$$

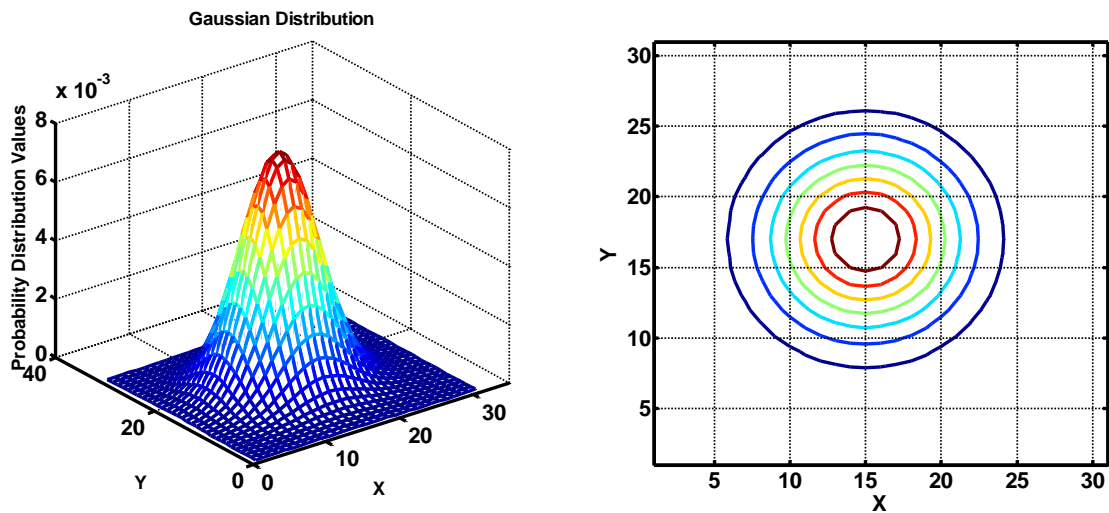


Figure 11. Gaussian distribution with a support region for two 2d GRVs (increase in size)

$$C = \begin{bmatrix} 40 & 0 \\ 0 & 5 \end{bmatrix}$$

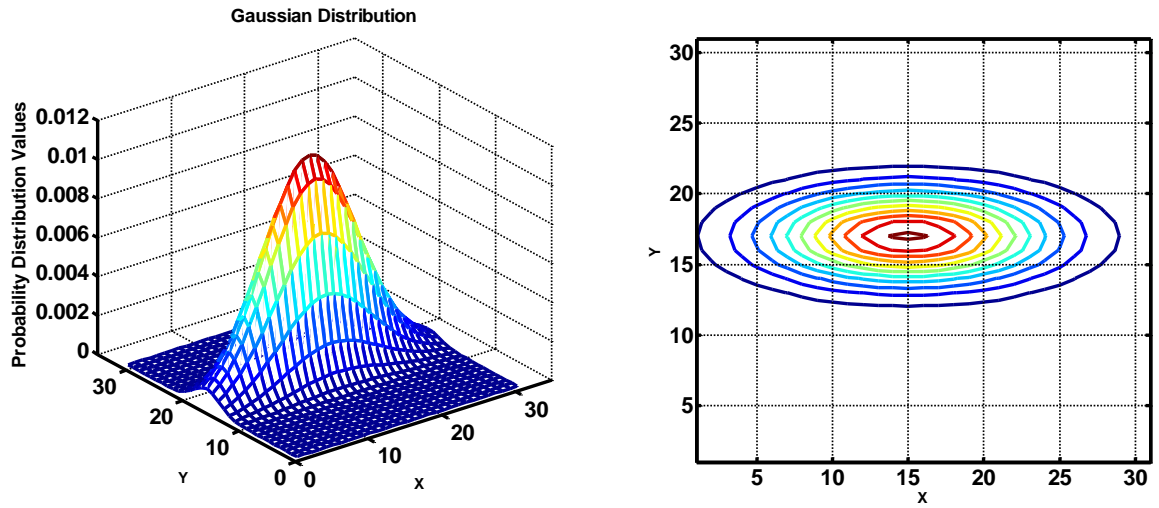


Figure 12. Gaussian distribution with a support region for two 2d GRVs (extension in X-direction))

$$C = \begin{bmatrix} 5 & 0 \\ 0 & 40 \end{bmatrix}$$

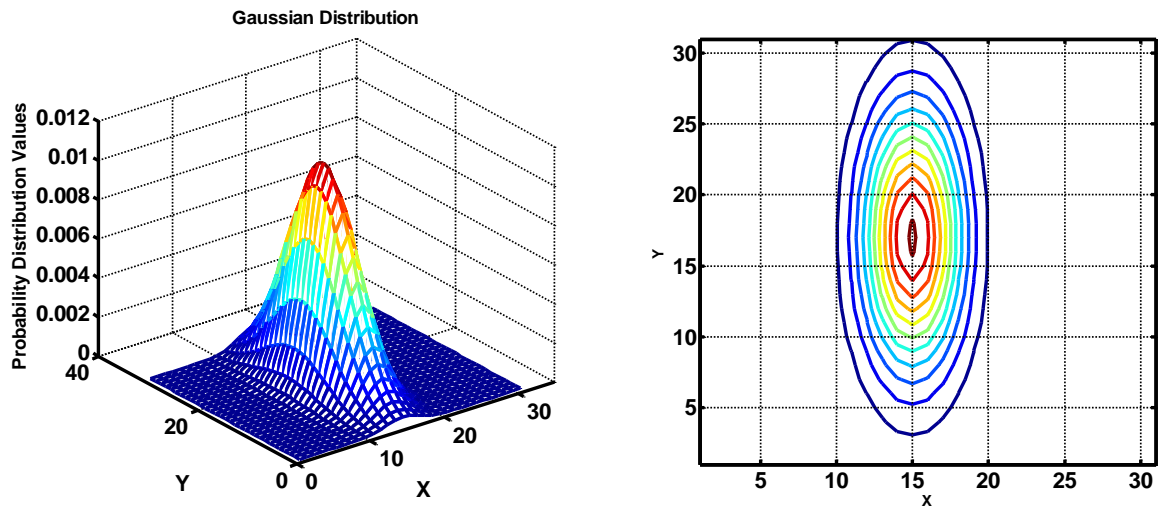


Figure 13. Gaussian distribution with a support region for two 2d GRVs (extension in Y-direction))

$$C = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$$

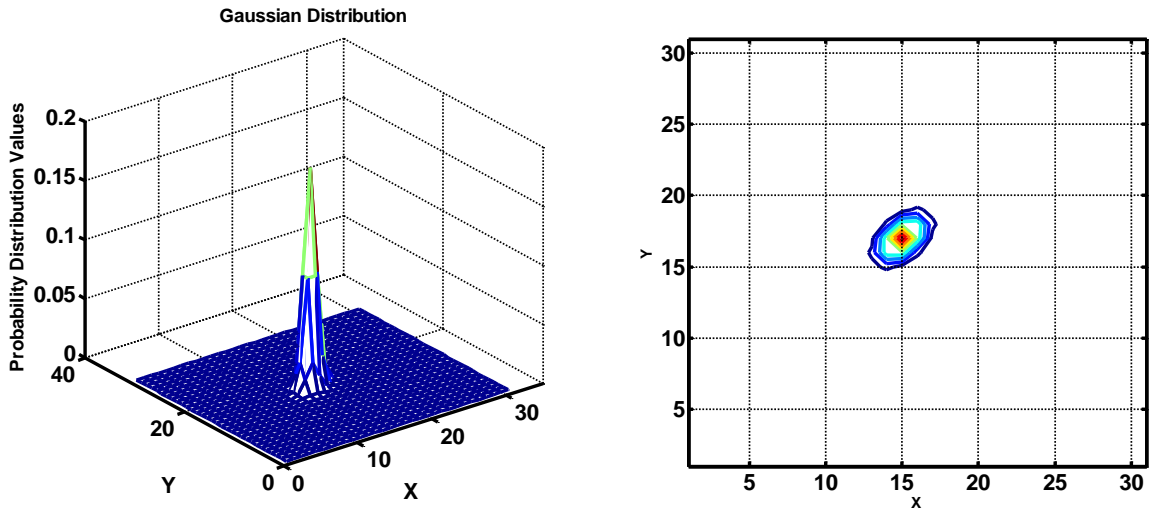


Figure 14. Gaussian distribution with a support region for two 2d GRVs (45 degree rotation)

$$C = \begin{bmatrix} 5 & 0.5 \\ 0.5 & 2 \end{bmatrix}$$

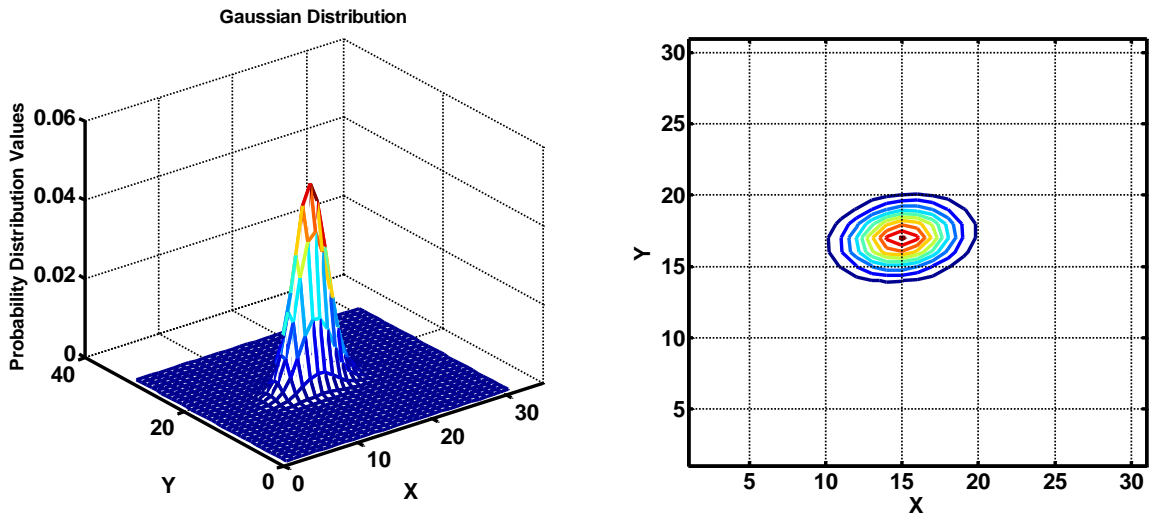


Figure 15. Gaussian distribution with a support region for two 2d GRVs (81 and 9 degree rotation)

Discussion

In this HW the goal was to learn how to generate and work with Gaussian distributions (one and multidimensional). We also learn method of maximum likelihood decoding for error rate calculation.

We observed that choice of threshold value, as well as variance of data distributions, causes change in error rate.

We also learned how covariance matrix influences on shape of probability distribution and on shape of its support region.

References

- [1] Retrieved from: <https://engineering.purdue.edu/VISE/ee438L/matlab/help/pdf/random.pdf>