11849060 谭树杰

Class Project 4

To do the following:

(a) Design a Bayesian classifier to discriminate between two synthetic classes with the following parameters:

a. 2D case: mean vector $M_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and Covariance $\sum_1 = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$ for the first class and mean vector $M_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and Covariance $\sum_2 = \begin{bmatrix} 1 & -0.5 \\ -0.5 & 1 \end{bmatrix}$ for the second class.

b. 3D case: mean vector $M_1 = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}$ and Covariance $\sum_1 = \begin{bmatrix} 5 & 2 & 1 \\ 2 & 5 & 0 \\ 1 & 0 & 4 \end{bmatrix}$ for the first class and mean vector $M_2 = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}$ and Covariance $\sum_2 = \begin{bmatrix} 2 & 2 & 1 \\ 2 & 4 & 0 \\ 1 & 0 & 5 \end{bmatrix}$ for the second class.

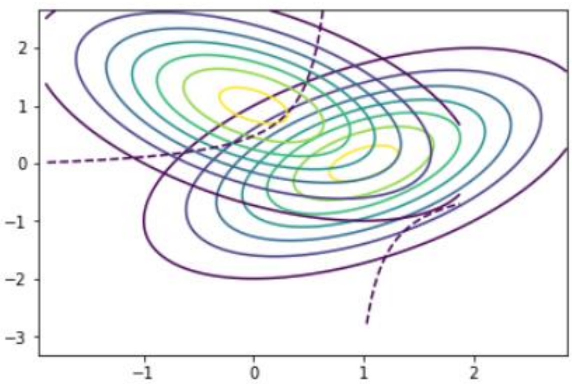
(b) Visualize the decision boundaries for (a).

(c) Compute type I and type II errors for (a).

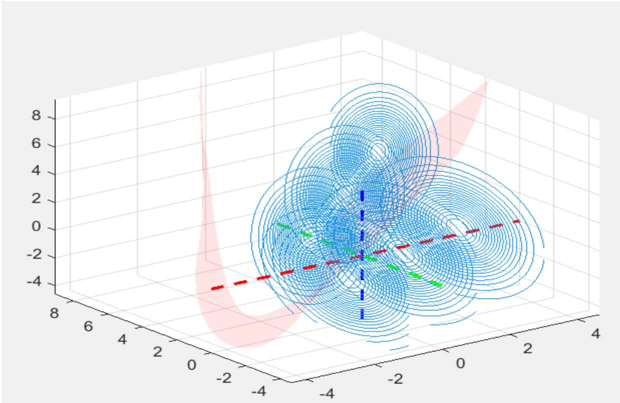
We calculate the discriminate function use following formulas:

$$g_i(\mathbf{x}) = \mathbf{x}^t \mathbf{W}_i \mathbf{x} + \mathbf{w}_i^t \mathbf{x} + w_{i0}$$
$$\mathbf{W}_i = -\tfrac{1}{2}\mathbf{\Sigma}_i^{-1}$$
$$\mathbf{w}_i = \mathbf{\Sigma}_i^{-1} \boldsymbol{\mu}_i$$
$$w_{i0} = -\frac{1}{2}\boldsymbol{\mu}_i^t \mathbf{\Sigma}_i^{-1} \boldsymbol{\mu}_i - \frac{1}{2}\ln|\mathbf{\Sigma}_i| + \ln P(\omega_i)$$

For 2D case, the decision boundary can be calculated by g1- g2, as is shown below:



For 3D case, the decision boundary can be calculated similarly, as is shown below:

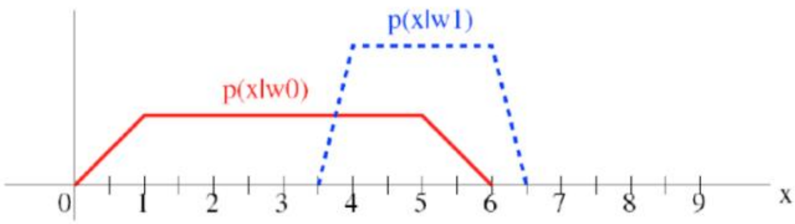


Type I and Type II errors are:

	Type I errors	Type II errors
2D case	0.654	0.345
3D case	0.933	0.767

Class Project 4-2

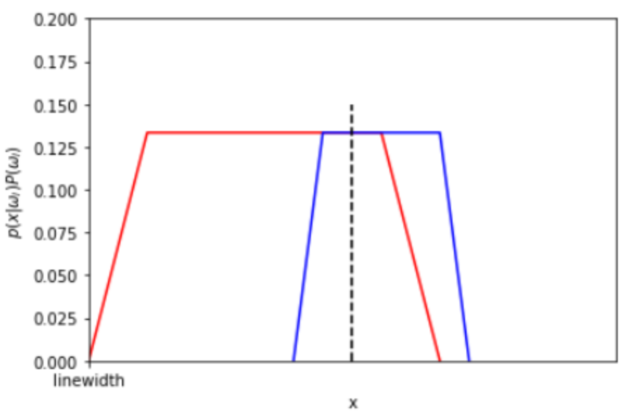
Consider a system which tries to detect the presence ω_1 or absence ω_0 of a target. The figure below illustrates the class-conditional probability distributions of observation x given the two cases:



a. Given $P(\omega_0) = 2/3$, find and sketch the decision regions R_0 and R_1 for the minimum-error detection rule and computer the associated probability of error.

b. Let $P_{fa} = P(R_1|\omega_0)$ and $P_{md} = P(R_0|\omega_1)$ be defined as the probabilities of false alarm and misdetection, respectively. Sketch an ROC curve for the decision rule: $assign\ x\ to\ \omega_1\ if\ x > \tau$. Label specific points for at least five different values of τ, including any critical values.

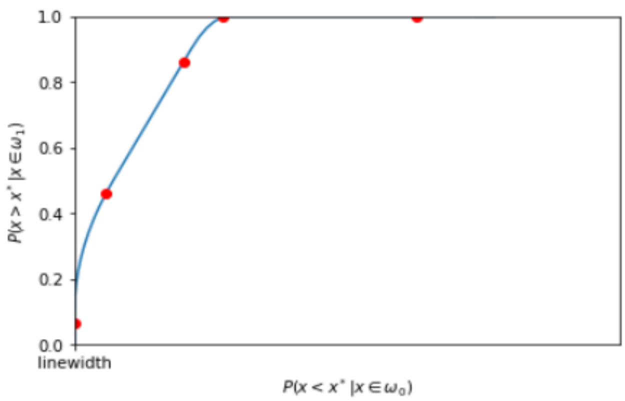
a. The decision boundary is shown below:



The probability of error is:

$$\mathrm{P}(\text{ error }) = \int_{R1} p(x|\omega_0)P(\omega_0)dx + \int_{R0} p(x|\omega_1)P(\omega_1)dx = 0.233$$

b. The ROC curve is shown below:



Class Project 4 Extra (50% over Class Project 4):

Consider a three-class classification problem where the classes ω_1, ω_2 and ω_3 are equally probable and have class-conditional probability distributions $p(\mathbf{x}|\omega_i) = N(\boldsymbol{\mu}_\mathbf{i}, \boldsymbol{\Sigma}_\mathbf{i})$ with

$$\boldsymbol{\mu}_1 = \begin{pmatrix}-2\\0\end{pmatrix}, \boldsymbol{\mu}_2 = \begin{pmatrix}0\\2\end{pmatrix}, \boldsymbol{\mu}_3 = \begin{pmatrix}0\\-2\end{pmatrix}$$

And identical covariance matrices $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}_3 = \sigma^2\mathbf{I},$ where $\sigma^2 = 0.25$.

a. Find the discriminant functions corresponding to the minimum-error decision rule.
b. Find the decision regions and boundaries associated to this rule.
c. Create 100 random patterns from each of the three class-conditional distributions and plot them in the two-dimensional feature space. Sketch the decision boundaries of part (b) on this plot. (Hint: you can use the plot and ezplot function for plotting).
d. Find the maximum likelihood estimates of $\boldsymbol{\mu}_\mathbf{i}$ and $\boldsymbol{\Sigma}_\mathbf{i}$, $i = 1,2,3$ using the sample patterns.
e. Substitute these estimates to the Bayesian decision rule and find the decision boundaries. Sketch then along with the plots in part (c).
f. Generate a new set of random patterns from each class, compute the maximum likelihood estimates of the parameters, and find the resulting decision boundaries. Prepare a new plot, like the one in part (e), that overlays the new decision boundaries on the new sample patterns. Compare these results to the previous case. Are the decision boundaries obtained using different samples different? Why/why not? Does this match your expectations?.

a. The discriminant functions are calculated as class project 4-1. We use sympy to get the result:

```
1  g1
```

```
Matrix([[-2.0*u**2 - 8.0*u - 2.0*v**2 - 10.0794415416798]])
```

```
6  g2
```

```
Matrix([[-2.0*u**2 - 2.0*v**2 + 8.0*v - 10.0794415416798]])
```

```
6  g3
```

```
Matrix([[-2.0*u**2 - 2.0*v**2 - 8.0*v - 10.0794415416798]])
```

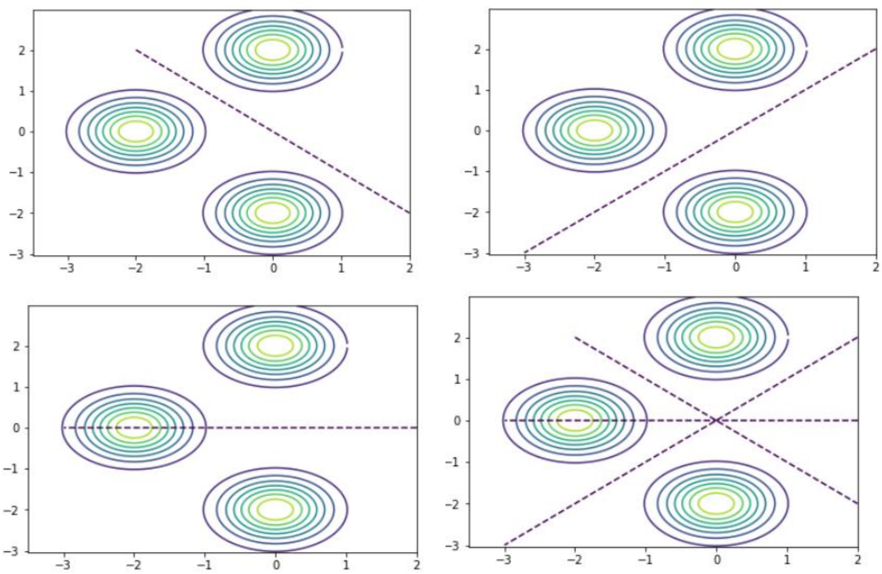
b. We calculated the boundary by g2 – g1, g3 -g1 and g3 – g1, the results are shown below:

```
In [61]:  1  g_bound1
Out[61]: Matrix([[8.0*u + 8.0*v]])

In [62]:  1  g_bound2
Out[62]: Matrix([[8.0*u - 8.0*v]])

In [63]:  1  g_bound3
Out[63]: Matrix([[-16.0*v]])
```

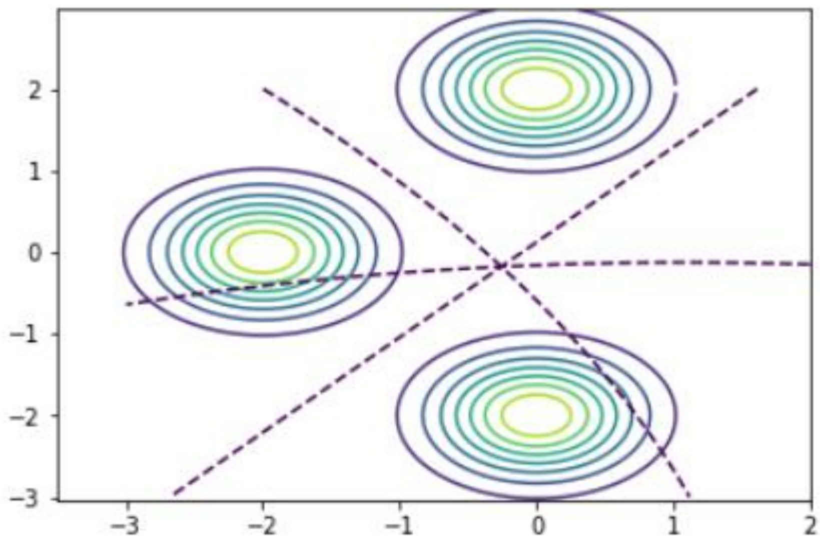
c. The decision boundary is shown below:



d. The maximum-likelihood estimate of μ_i and Σ_i are:

	μ	Σ
Class ω_1	[-1.906, -0.027]	`array([[ 0.212511 , -0.02543148], [-0.02543148, 0.24259437]])`
Class ω_2	[0.410, 1.975]	`array([[0.24875774, 0.03155827], [0.03155827, 0.3333563 ]])`
Class ω_3	[-0.032, -1.99]	`array([[0.20564183, 0.01058592], [0.01058592, 0.25161024]])`

e. The decision boundaries found using the estimated parameters



f. Compare is shown in below figure. The maximum-likelihood are very close to ground truth, thus the result match my expectation.