# IN 5520 2020
# Classification 2
## Exercises related to the lecture on 14.10.20

You will need your souce code for Exercise 1 when doing Mandatory Exercise part 2, so implement this now ☺

**Exercise 1.** Matlab exercise for classification based on a multivariate Gaussian classifier.

The images used in this exercise are under undervisningsmateriale/images on the course web page.

**Step 1: Implement a Gaussian classifier using a d-dimensional feature vector**
For the algorithm, see lecture foils. It is recommended that you use Matlab built-in functions for matrix inversion and computing the determinant, but write the remaining algorithm yourself. (You can use built-in functions mean() og cov())

**Step 2: Train the classifier**
Train the classifier on the image tm_train.png using the 4 classes defined in the mask file tm_train.png. Estimates of the class-specific mean vector and covariance matrix are found on the lecture foils. You may use Matlab functions mean() and cov() here.

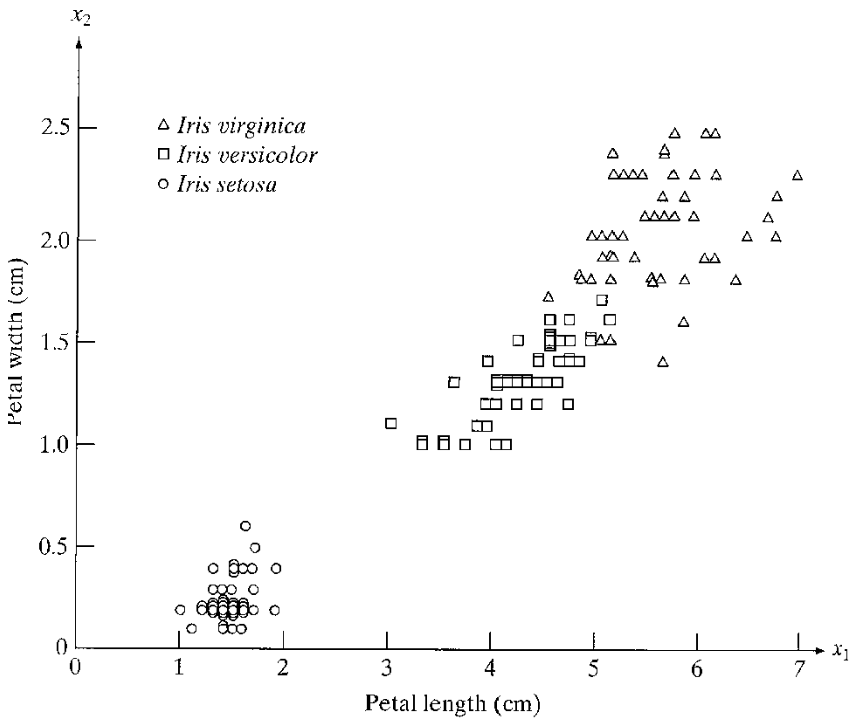
If you want to verify that your code gives the correct classification labels, check the resulting classification image you get when you classify the entire image tm.mat with the classification we produced (tm_classres.mat). In this image each pixel is assigned class labels 1-4.

**Step 3: Find the classification accuracy for classification using all features.**
Run the classification on the multivariate (6-band) input image tm.mat. Compute the percentage of correctly classified pixels when using all features, and compare it to using single features. The classification accuracy should be computed on the test image tm_test.png.

Also try the simplified covariance matrix $\Sigma=\sigma^2 I$. Which version gives the highest classification accuracy?

## 2. Finding the decision functions for a minimum distance classifier.

A classifier that uses equal diagonal covariance matrices is often called a minimum distance classifier, because a pattern is classified to class that is closest when distance is computed using Euclidean distance. See plot on next page.




a. In the above figure, find the class means just by looking at the plot.

b. If this data is classified using a minimum distance classifier, sketch the decision boundaries on the plot.

## 3. Discriminant functions

A classifier that uses Euclidean distance computes distance from pattern ***x*** to class *j* as:

$$D_j(x) = \|x - \mu_j\|$$

Show that classification with this rule is equivalent to using the discriminant function

$$d_j(x) = x^T \mu_j - \frac{1}{2}\mu_j^T \mu_j$$

# Exercise 3: Classification (From Exam 2015)

a) Consider a two-dimensional feature vector and a set of points in 2D feature space: (-3,6) (-2,4) (-1, 2) (0,0) (1, -2) (2, -4) (3, -6)
Show that the covariance matrix between the two features is :

$$\Sigma = \begin{bmatrix} 4 & -8 \\ -8 & 16 \end{bmatrix}$$

Show all your calculations.

b) Given the two features defined above, would you base you classification on 1 or 2 features? Justify your answer.

c) The discriminant functions for a multivariate Gaussian classifier are given as:

$$g_i(\mathbf{x}) = -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^T \Sigma_i^{-1}(\mathbf{x}-\boldsymbol{\mu}_i) - \frac{n}{2}\ln 2\pi - \frac{1}{2}\ln|\Sigma_i| + \ln P(\omega_i)$$

Consider two classes with equal prior probability and

$$\boldsymbol{\mu}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \boldsymbol{\mu}_2 = \begin{bmatrix} 0 \\ 2 \end{bmatrix} \quad \Sigma_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \Sigma_1^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\Sigma_2 = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \quad \Sigma_2^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \quad |\Sigma_1| = 1 \quad |\Sigma_2| = 1$$

Can the discriminant function be simplified in this case?

d) Classify the point $x = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$ by computing the value of the discriminant functions and assign it to the class corresponding to the highest probability.

e) Explain how classifier sensitivity and specitivity are computed, and discuss their importance on a medical classification problem.

f) Let us assume that we have a 2-class classification problem with a 1-dimensional feature vector f(x) which is exponentially distributed given the class-conditional parameter $\lambda_i$:

$$f(x) = \lambda_i \exp^{-\lambda_i x}$$

Find an expression for the decision boundary for this classification problem.

# Exercise 4: Classification (From Exam 2016)

Consider a two-dimensional feature vector and a set of points from 2 classes in 2D feature space:
Class 1 has points: (3,0), (5,0), (7,0), (5,2), (5, -2)
Class 2 has points: (0,5), (0,3), (0,7), (-2, 5), (2,5)
The discriminant function for Gaussian classifier is in the general form:

$$g(\mathbf{x}) = -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_\mathbf{i})^T \Sigma_i^{-1}(\mathbf{x}-\boldsymbol{\mu}_\mathbf{i}) - \frac{d}{2}\ln 2\pi - \frac{1}{2}\ln|\Sigma_i| + \ln P(\omega_i)$$

a) Compute the mean for each class

b) Show that the covariance between feature 1 and 2 is 0 for both classes.

c) Show how the discriminant function can be simplified in this case with a classifier with equal diagonal covariance matrices.

d) Find an expression for the decision boundary using this simplified discriminant function.

*e)* Compute the decision boundary if we assume equal prior probabilities.

f) Sketch the class means and the decision boundary in a plot if we assume that the two classes have equal prior probability.

g) If $P(\omega_1) = 0.75$, in which direction will the decision boundary change? Indicate this on the plot.