# Homework 0: Introduction to Probability & Matlab

TDA231 - Algorithms for Machine Learning & Inference

Theodor Åstrand, theast@student.chalmers.se, 931109-9114
Linnéa Otterlind, linott@student.chalmers.se, 921126-3620

Chalmers University of Technology

January, 2016

# Problem 1.1

**Problem decomposition:**

1. If you have the disease, the test will be positive with 0.99 probability.
2. If don't have the disease, the test will be negative with 0.99 probability.
3. 1 in 10 000 have the disease.

**Goal:** Find P(Y|X) : Probability that you have the disease, if you test positive.

- X : Event that you test positive
- Y : Event that you have the disease

P(X|Y) = 0.99 : Probability that you test positive, if you have the disease.

P(Y) = $\frac{1}{10000} = 0.0001$ : Probability you have the disease.

P(X|Y') = 0.01 : Probability that you test positive, if you don't have the disease.

P(Y') = $1 - 0.0001 = 0.9999$ : Probability that you don't have the disease.

**Answer:** P(Y|X) = $\frac{P(X,Y)}{P(X)} = \frac{P(X|Y) \cdot P(Y)}{P(X|Y) \cdot P(Y) + P(X|Y') \cdot P(Y')} = \frac{0.99 \cdot 0.0001}{0.99 \cdot 0.0001 + 0.01 \cdot 0.9999} = 0.0098 \approx$ 0.01

# Problem 1.2

**Goal:** Show that $Cov[X, Y] = E[XY] - E[X] \cdot E[Y] = 0$

We know that X is uniformly distributed in [-1, +1], which means that we have equal probability density to every value between -1 and 1 on the real line. This is formulated by this property:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{if a} \leq x \leq b \\ 0 & \end{cases}$$

Therefore f(x) can be seen as a constant value c.

$E[XY] = E[X^3] = \int_{-\infty}^{\infty} x^3 \cdot f_X(x)dx = c \cdot \int_{-\infty}^{\infty} x^3 dx = c \cdot (\int_{-\infty}^{-1} x^3 dx + \int_{-1}^{1} x^3 dx + \int_{1}^{\infty} x^3 dx) = c \cdot (0 + \int_{-1}^{1} x^3 dx + 0) = c \cdot \int_{-1}^{1} x^3 dx = c \cdot \left[\frac{x^4}{4}\right]_{-1}^{1} = c \cdot (\frac{1}{4} - \frac{1}{4}) = 0$

$E[X] = \int_{-\infty}^{\infty} x \cdot f_X(x)dx = c \cdot \int_{-\infty}^{\infty} xdx = c \cdot (\int_{-\infty}^{-1} xdx + \int_{-1}^{1} xdx + \int_{1}^{\infty} xdx) = c \cdot (0 + \int_{-1}^{1} xdx + 0) = c \cdot \int_{-1}^{1} xdx = c \cdot \left[\frac{x^2}{2}\right]_{-1}^{1} = c \cdot (\frac{1}{2} - \frac{1}{2}) = 0$

**Answer:** $Cov[X, Y] = E[XY] - E[X] \cdot E[Y] = 0 - 0 \cdot E[Y] = 0$

# Problem 2.1

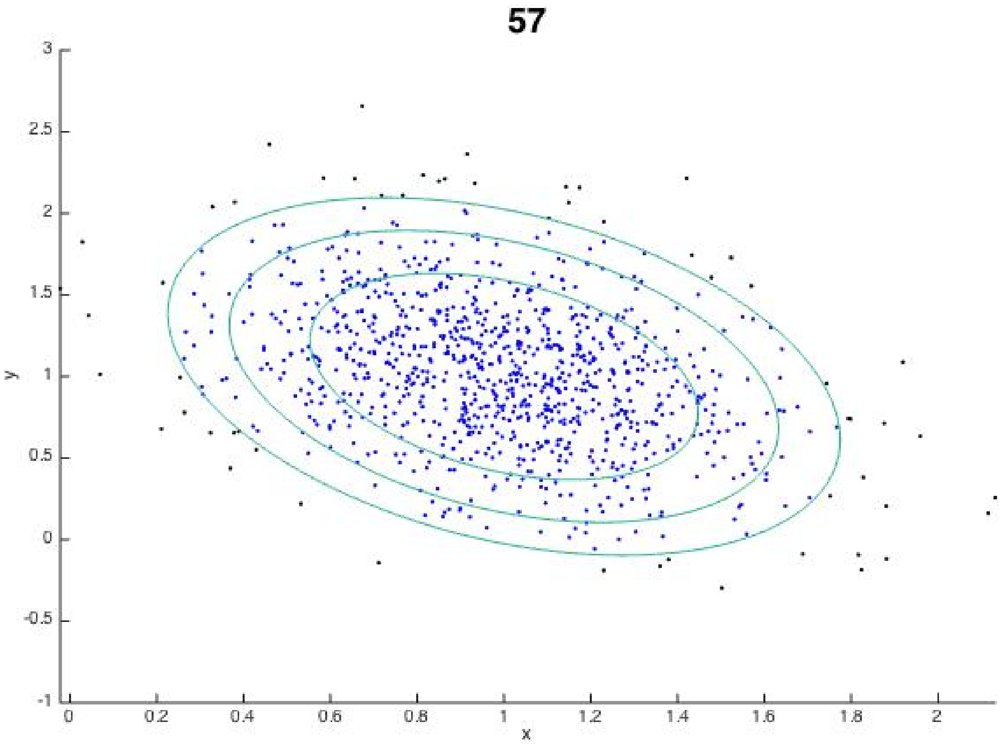


Figure 1: The final plot of problem 2.1

## Problem 2.2

From figure 2 we observe that the correlation of X and Y remains the same while the covariance of X and Y differs. This is due to the fact that each feature is scaled independently.

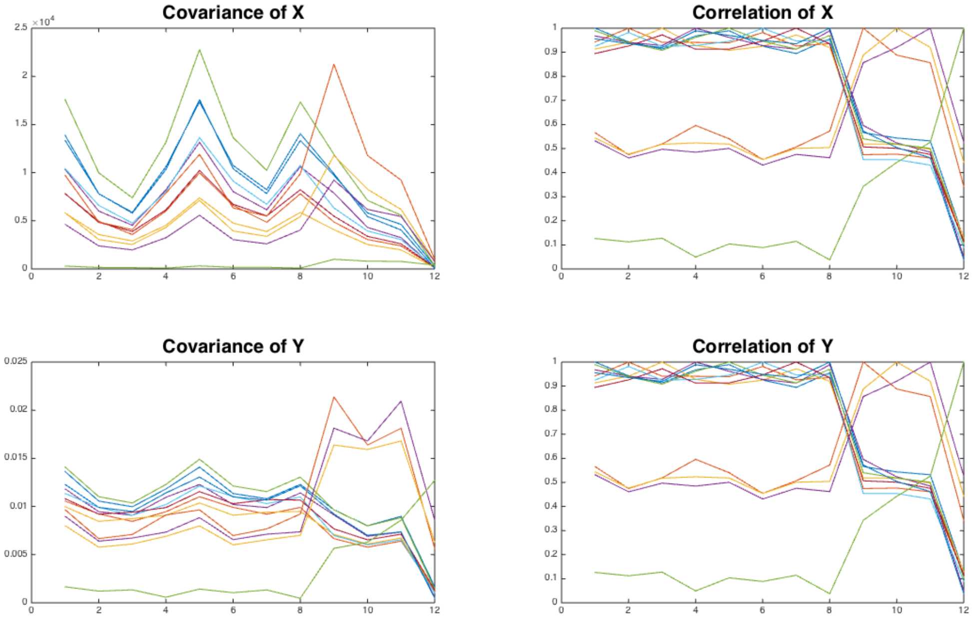


Figure 2: Plot of correlation and covariance of X and Y

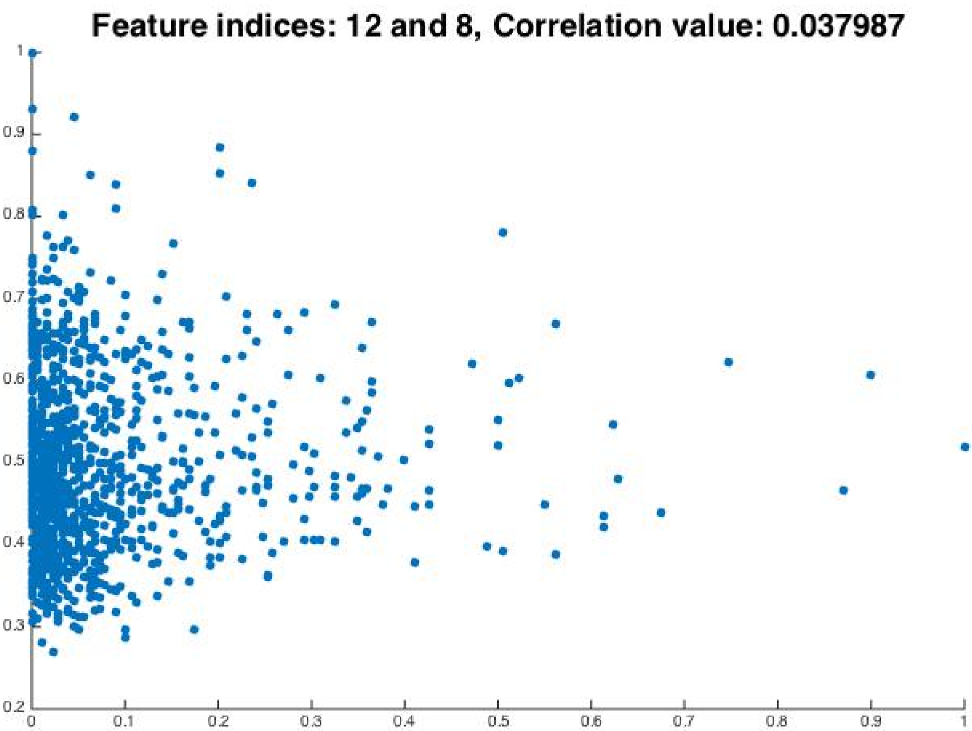


Figure 3: Scatter plot of the pair of features in Y having minimum correlation