CS470 Machine Learning
Assignment # 1*: Correlation Analysis

Uzair Akbar
(04584)

National University of Sciences and Technology (NUST), Pakistan.

1 Part a: Scatter Plots

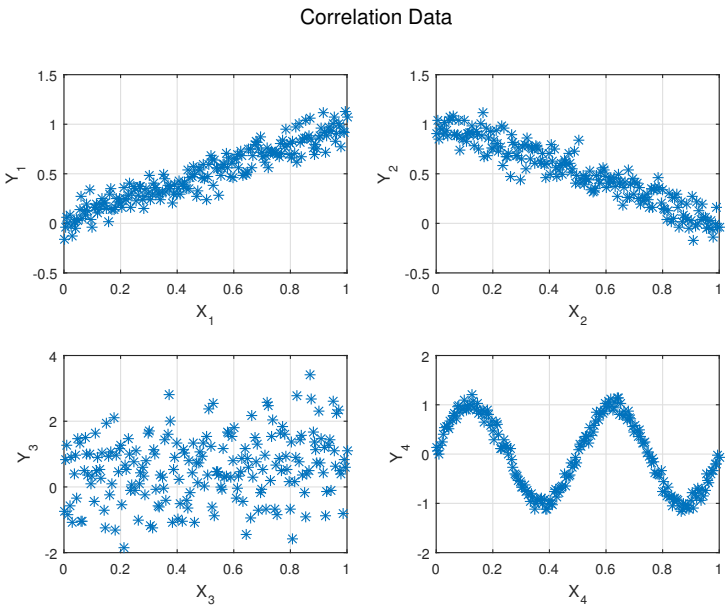

Fig. 1: Correlation Data.

* This is the implementation of problem 9 of Assignment 1. It can also be found at the following link:
https://github.com/uzairakbar/machine-learning-cs470

Data Set	Expected correlation coefficients
X_1, Y_1	+1
X_2, Y_2	-1
X_3, Y_3	0
X_4, Y_4	0

Table 1: Expected correlation coefficients.

2 Part b: Estimators

The calculated correlation coefficients for each data set were found to be the same as the matlab `corrcoef()` function.

Data Set	Correlation	Covariance	Correlation Coefficient	`corrcoef()`
X_1, Y_1	0.3315	0.0812	0.9445	0.9445
X_2, Y_2	0.1717	-0.0859	-0.9423	-0.9423
X_3, Y_3	0.3484	0.0610	0.2112	0.2112
X_4, Y_4	-0.0891	-0.0859	-0.4080	-0.4080

Table 2: Calculated estimators compared to the matlab `corrcoef()` function.

3 Appendix - Matlab Code

3.1 Functions

Correlation

```
function corr = correlation( x, y )

corr = mean(x.*y);

end
```

Covariance

```
function covar = covariance( x, y )

covar = (sum((x - mean(x)).*(y - mean(y))));
covar = covar/(length(x) - 1);

end
```

Correlation Coefficient

```
function corrcoef = correlationCoefficient( x, y )

covar = covariance( x, y );

sx = (sum((x - mean(x)).^2)/(length(x) - 1))^0.5;
sy = (sum((y - mean(y)).^2)/(length(y) - 1))^0.5;

corrcoef = covar/(sx*sy);

end
```

3.2 Main script

```
clear;
clc;

% load data into workspace
load('CorrelationData.mat');

subplot(2, 2, 1);
scatter(X1, Y1, '*');        % plot X1 and Y1 vectors
xlabel('X_1'); ylabel('Y_1');
box on; grid on;

subplot(2, 2, 2);
scatter(X2, Y2, '*');        % plot X2 and Y2 vectors
xlabel('X_2'); ylabel('Y_2');
box on; grid on;

subplot(2, 2, 3);
scatter(X3, Y3, '*');        % plot X3 and Y3 vectors
xlabel('X_3'); ylabel('Y_3');
box on; grid on;

subplot(2, 2, 4);
scatter(X4, Y4, '*');        % plot X4 and Y4 vectors
xlabel('X_4'); ylabel('Y_4');
box on; grid on;

suptitle('Correlation Data')

saveas(gcf,'scatterPlot.eps','eps2c');
```