Ph 21 - Assignment 6 - Principal Componenent Analysis

Yovan Badal

06/09/2018

1 2D Linear dataset

We simulate a simple 2D array of linearly dependent datapoints, including errors (simulated by multiplying the samples by samples from a normal distribution with mean 1). Testing our PCA analysis code gives the following output:

```
Linear PCA = [[  0.17336095  -0.09487871   0.11116776   0.78036476   0.16470093
    0.53517108   0.28634902  -0.46264184   1.13014248  -1.10907052]
 [ -0.83481595  -4.4613282  -10.37175513 -18.51247361 -20.46597215
  -27.16817782 -30.88819511 -32.1996602  -44.78789363 -38.92322583]]
Variance over axes = [0.3547789280245619, 191.33097622323984]
```

We can easily observe that our PCA analysis correctly identifies a single degree of freedom in our data, since only one of the components contributes significantly to the covariance.

2 4D Linear dataset

We simulate a simple 4D array of linearly dependent datapoints by a simple extension of the code for the 2D linear dataset, including error. Testing our PCA analysis code gives the following output:

```
PCA = [[  0.17336095  -0.09487871   0.11116776   0.78036476   0.16470093
    0.53517108   0.28634902  -0.46264184   1.13014248  -1.10907052]
 [ -0.83481595  -4.4613282  -10.37175513 -18.51247361 -20.46597215
  -27.16817782 -30.88819511 -32.1996602  -44.78789363 -38.92322583]]
Variance over axes = [197.48303438984914, 0.04637431453798641, 2.3041924450580025, 1.255255715040029]
```

Again we observe that our PCA analysis correctly identifies a single degree of freedom in our data, since only one of the components contributes significantly to the covariance.