# Exercise Set 0: Prerequisite Knowledge

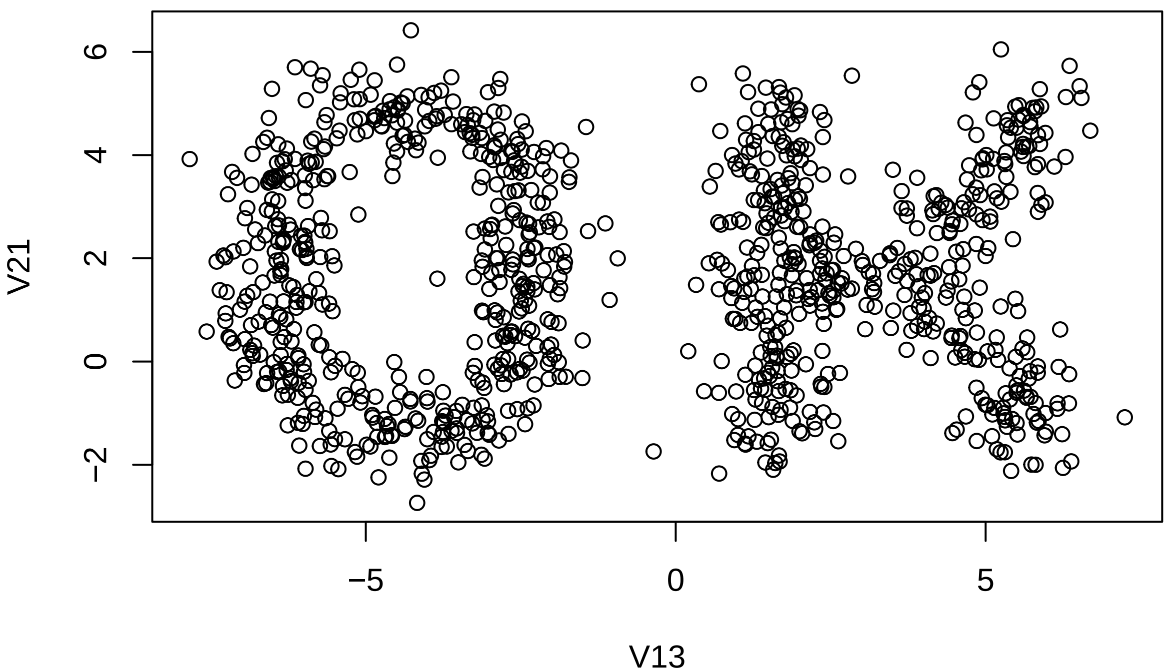
Enrico Buratto

4/10/2021

## Problem 1

### Task a

```
x <- read.csv("./x.csv")
variances <- sapply(x, var)
variances <- variances[order(variances, decreasing = TRUE)]
first <- x[names(variances[1])]
second <- x[names(variances[2])]
to_plot <- data.frame(first, second)
plot(to_plot)
```




## Problem 2

### Task a

In order to prove that $\lambda_i$ and $x_i$ are, respectively, eigenvalues and eigenvectors of $\mathbf{B}$ as well, we must prove that

$$B_i x_i = \lambda_i x_i \ \forall i = 1..n$$

We have that

$$(1.)\ B_i x_i = \lambda_i \begin{pmatrix} x_{1i}x_{1i} & \dots & x_{1i}x_{ni} \\ \vdots & \vdots & \vdots \\ x_{ni}x_{1i} & \dots & \dots x_{ni}x_{ni} \end{pmatrix} \begin{pmatrix} x_{1i} \\ \vdots \\ x_{ni} \end{pmatrix}$$

$$(2.)\ \lambda_i x_i = \lambda_i \begin{pmatrix} x_{1i} \\ \vdots \\ x_{ni} \end{pmatrix}$$

Developing (1.), we have that

$$B_i x_i = \lambda_i \begin{pmatrix} x_{1i} \sum_{j=1}^n x_{ji}^2 \\ \vdots \\ x_{ni} \sum_{j=1}^n x_{ji}^2 \end{pmatrix}$$

Since the summations of the last vector are equal to 1 due to orthonormality, the vector is $x_i$. Thus

$$B_i x_i = \lambda_i x_i \ \forall i = 1, ..., n$$

And then $x_i$ and $\lambda_i$ are eigenvectors and eigenvalues of B.

### Task b

The eigenvalues for the given matrix are

$$\lambda_1 = 2 + \sqrt{5}$$
$$\lambda_2 = 2 - \sqrt{5}$$

The associated eigenvectors are

$$v_1 = \begin{pmatrix} \frac{-1+\sqrt{5}}{2} \\ 1 \end{pmatrix}$$

$$v_2 = \begin{pmatrix} \frac{-1-\sqrt{5}}{2} \\ 1 \end{pmatrix}$$

The equation can be shown to be satisfied solving this formula (passages omitted due to brevity)

$$A = (2+\sqrt{5}) \begin{pmatrix} \frac{-1+\sqrt{5}}{2} \\ 1 \end{pmatrix} \begin{pmatrix} \frac{-1+\sqrt{5}}{2} & 1 \end{pmatrix} + (2-\sqrt{5}) \begin{pmatrix} \frac{-1-\sqrt{5}}{2} \\ 1 \end{pmatrix} \begin{pmatrix} \frac{-1-\sqrt{5}}{2} & 1 \end{pmatrix}$$

With further (trivial) calculations can be proved that

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$$

## Problem 3

### Task a

In order to show that $E$ is a linear operator, two properties have to be demonstrated:

1. $E[X+Y] = E[X] + E[Y]$
2. $E[tX] = tE[X]$

**Proof of (1.)** Since $X$ and $Y$ are defined on the same sample space, their sum is $X + Y$. Then

$$E[X + Y] = \sum_{\omega \in \Omega} (X + Y)(\omega)P(\omega)$$

For summation property

$$\sum_{\omega \in \Omega} (X + Y)(\omega)P(\omega) = \sum_{\omega \in \Omega} X(\omega)P(\omega) + \sum_{\omega \in \Omega} Y(\omega)P(\omega)$$

And then finally

$$E[X + Y] = E[X] + E[Y]$$

**Proof of (2.)** This proof is even simpler:

$$E[tX] = \sum_{\omega \in \Omega} tX(\omega)P(\omega) = t \sum_{\omega \in \Omega} X(\omega)P(\omega)$$

And then

$$E[tX] = tE[X]$$

### Task b

The proof is simple

$$Var[X] = E[(X - m)^2] = E[(X - E(X))^2] = E[X^2 + E[X]^2 - 2XE[X]]$$

Since the expectation is a linear operator, we can then write

$$Var[X] = E[X^2] + E[X]^2 - 2E[X]E[X] = E[X^2] + E[X]^2 - 2E[X]^2 = E[X^2] - E[X]^2$$

## Problem 4

### Task a

We know that

$$P(X \mid Y) = \frac{P(X \wedge Y)}{P(Y)}$$

We can use this formula in the other way as well, so

$$P(Y \mid X) = \frac{P(Y \wedge X)}{P(X)}$$

Starting from these premises, it's easy to see that

$$P(Y \wedge X) = P(X \mid Y) \cdot P(Y)$$

and also

$$P(Y \wedge X) = P(Y \mid X) \cdot P(X)$$

since $P(X \wedge Y) = P(Y \wedge X)$. We can then set $P(X \mid Y) \cdot P(Y)$ equal to $P(Y \mid X) \cdot P(X)$

$$P(X \mid Y) \cdot P(Y) = P(Y \mid X) \cdot P(X)$$

Isolating $P(X \mid Y)$ it's then trivial that

$$P(X \mid Y) = \frac{P(Y \mid X) \cdot P(X)}{P(Y)}$$

### Task b

First of all, we define two boolean variables:

- $T = \begin{cases} 0 \; if \; test \; is \; negative \\ 1 \; if \; test \; is \; positive \end{cases}$
- $A = \begin{cases} 0 \; if \; person \; is \; not \; allergic \\ 1 \; if \; person \; is \; allergic \end{cases}$

We then have to find $P(A = 1 \mid T = 1)$. Using the Bayes theorem

$$P(A = 1 \mid T = 1) = \frac{P(T = 1 \mid A = 1) \cdot P(A = 1)}{P(T = 1)}$$

The single components of the formula are:

- $P(T = 1 \mid A = 1)$, i.e. the probability that the test is positive if the person is allergic, that is equal to 0.85
- $P(A = 1)$ can be calculated using the marginal probabilty formula

$$P(A = 1) = P(A = 1 \wedge T = 1) + P(A = 1 \wedge T = 0) = 0.2 \cdot 0.85 + 0.2 \cdot 0.15 = 0.2$$

- Similarly, $P(T = 1)$ can be calculated with the marginal probabilty formula

$$P(T1) = P(T = 1 \wedge A = 1) + P(A = 0 \wedge T = 1) = 0.2 \cdot 0.85 + 0.8 \cdot 0.23 = 0.354$$

Putting all together,

$$P(A = 1 \mid T = 1) = \frac{0.85 \cdot 0.2}{0.354} = 0.48$$

Then, the probabilty that a person is really allergic to pollen if the test result is positive 48%.

## Problem 5

### Task a

The value $x \in \mathbb{R}$ that minimizes the value of $f(x)$ can be found calculating the first derivative and setting it to zero, so $f'(x) = 0$. The first derivative is

$$f'(x) = 4ax^3 + b$$

and it's easy to find that it has only one solution (or, better, three coincident solutions), that is

$$x_0 = \sqrt[3]{\frac{-b}{4a}}$$

This point is a critical point; in order to be a minimum, it must be the point for which the second derivative on that point is greater than 0, so

$$f"(x) = 12ax^2 > 0$$

on

$$x_0 = \sqrt[3]{\frac{-b}{4a}} \Rightarrow 12a(\sqrt[3]{\frac{-b}{4a}})^2 > 0$$

that is true for every $a > 0$. The value that minimized $f(x)$ is then

$$\sqrt[3]{\frac{-b}{4a}}$$

with $a > 0$.

### Task b

Since the first derivative has three coincident solutions, it means that there is only one critical point; furthermore, in order to be that critical point a minimum point, the second derivative in that point must be greater than 0. We can then conclude that the condition for the function to have a finite and unique point is to have $a > 0$, since with this condition $(\sqrt[3]{\frac{-b}{4a}})^2$ can't be undefined ($a \neq 0$) and the equation is true.

## Problem 6

### Task a

```
1. Fibonacci(n)
2.     a <- 1
3.     b <- 1
4.     print a
5.     print b
6.     i <- 2
7.     while i<n
8.         c = a+b
9.         print c
10.        a = b, b = c
11.        i += 1
```

### Task b

The time complexity of the algorithm is $O(n)$; let's analyze the algorithm in-depth:

- Instruction 1. is constant ($O(1)$), since it only receives a number $n$ in input
- Instructions $2. - 6.$ and $8. - 11.$ can be executed in $O(1)$ time, since they are only variable assignments and prints
- Instructions $8. - 11.$ are repeated $O(n)$ times

The algorithm time complexity is then a sum of constants, except for instructions $8. - 11.$ that are repeated exactly $n - 2$ times ($O(n)$).