Exercises - Linear least squares - Probability

Linear least squares

**Exercise 1** – Give the exact solution of the linear least square problem:

\[
\text{find } \vec{x} \in \mathbb{R}^2 \text{ s.t. } \| A\vec{x} - \vec{b} \|_2 = \min_{\vec{u} \in \mathbb{R}^2} \| A\vec{u} - \vec{b} \|_2
\]

where \( A = \begin{bmatrix} -1 & 2 \\ 2 & -1 \\ 1 & 3 \end{bmatrix} \) and \( \vec{b} = \langle 3, 2, 1 \rangle \).

Probability

**Exercise 2** – There are 5 paths \( p_1, \ldots, p_5 \) that lead form the base of a mountain to its top. Each of these path also lead from the top to the base of the mountain. A round trip is a pair \( (p_i, p_j) \) where \( p_i \) is the path taken to go up and \( p_j \) the path taken to go back down. We aim at counting the number of different acceptable roundtrips under several definitions of acceptable and different.

Give a formal definition of the set of different acceptable roundtrips and give their number when:

1. **Exercise 2.1** – an acceptable roundtrip is a trip passing through any path up and any path down; two acceptable roundtrips are different if their paths up or their paths down are different;
2. **Exercise 2.2** – an acceptable roundtrip is a trip passing through two different paths up and down; two acceptable roundtrips are different if their paths up or their paths down are different;
3. **Exercise 2.3** – an acceptable roundtrip is a trip passing through two different paths up and down; two acceptable roundtrips \( r_1 \) and \( r_2 \) are different if there is at least one path in \( r_1 \) that is not in \( r_2 \).

**Exercise 3** – We roll a fair die, and consider the events:

- \( E_2 \): “we roll a multiple of 2”,
- \( E_3 \): “we roll a multiple of 3”.

**3.1** – Compute \( P(E_2) \) and \( P(E_3) \)

**3.2** – Are \( E_2 \) and \( E_3 \) independent? (Justify your answer).

**Exercise 4** –
In this exercise, we consider that for a pregnant woman, it is equiprobable to deliver a girl or a boy. A mother has two children. Compute:

4.1 – the probability that both children are boys knowing that at least one is a boy,

4.2 – the probability that both children are boys knowing that the oldest is a boy.

A mother has two children, and we pick up randomly one of them. Compute:

4.3 – the probability that both children are boys knowing that the one picked up randomly is a boy.

Exercise 5 – An urn contains 10 balls numbered from 0 to 9. Five times, we pick a ball in the urn and note its number (we do not put it back in the urn).

We obtain a tuple \( S = (s_1, s_2, s_3, s_4, s_5) \) of five numbers where \( s_i \) is the number of the \( i \)-th ball that has been picked up.

5.1 – give the probability that \( S \) contains an ascending or descending sequence of length 5, that is: either \( \forall i < 5, s_{i+1} = s_i + 1 \) or \( \forall i < 5, s_{i+1} = s_i - 1 \).

5.2 – give the probability that \( S \) contains an ascending sequence of length 4 (4 consecutive elements of the tuple \( S \) are consecutive ascending integers).

5.3 – give the probability that \( S \) contains an ascending sequence of length 5 knowing that it contains an ascending sequence of length 4.

Exercise 6 – My teachers at high-school

In my high-school in France, 60% of the teachers were women. Among all the high-school teachers in France, \( \frac{1}{3} \) of the women wear glasses and \( \frac{1}{2} \) of the men wear glasses. What is the probability for a teacher of my high-school wearing glasses to be a woman?

Exercise 7 – When I was younger, I used to forget so often the keys of my apartment (in my apartment, so that I was locked outside) that I have been able to establish the following statistics:

- knowing that I forgot my keys the day \( d \), the probability that I forgot it the day \( d + 1 \) was \( \frac{1}{10} \).
- knowing that I did not forgot my keys the day \( d \), the probability that I forgot it the day \( d + 1 \) was \( \frac{4}{10} \).

Let the day 1 be the first of January, and suppose the probability that I forgot my keys on the day 1 is \( x \), where \( 0 < x \leq 1 \). We consider \( d \leq 365 \). Note \( P_d \) the (absolute) probability that I forgot my keys the day \( d \). Note \( E_d \) the event “I forgot my keys the day \( d \)”, and \( E_d^c \) the event “I did not forgot my keys the day \( d \)”.

7.1 – Show that \( E_d \) and \( E_d^c \) form a frame of discernment.

7.2 – Give an expression of \( P_{d+1} \) in function of \( P_d \).

7.3 – Give an expression of \( P_d \) in function of \( x \) and \( d \).

7.4 – Use Matlab to fill the following table, which components are the values of \( P_d \) for different \( d \) and different initial value \( x \).
\[\begin{array}{ccccccc}
\hline 
\text{x} & \text{d = 2} & \text{d = 4} & \text{d = 6} & \text{d = 8} & \text{d = 10} & \text{d = 100} \\
\hline 
x = 0.001 & & & & & & \\
x = 0.5 & & & & & & \\
x = 0.999 & & & & & & \\
\end{array}\]

7.5 – Would you bet that I forgot my keys the 31th of December of the same year?

*Hint:* questions (d) and (e) can be addressed independently of (c).

**Exercise 8 – Email classification**

I have a list of emails that I classified (manually) in 2 categories: spam and non spam. I want to design a classification algorithm that uses machine learning for deciding if a new email arriving in my mailbox is a spam or not. This exercise is a first step trough such a classifier.

I proceed as follows: first, I choose three boolean attributes as characteristics of an email that give good evidence of whether it is a spam or not. For a given email \(e\), these attributes are \(A_1(e)\), \(A_2(e)\) and \(A_3(e)\), each taking a value in \{true, false\}. These attributes can be for instance: ”\(e\) contains my first name”, or ”\(e\) contains the words *amazing* and *sales*”. For a given email, I note:

- \(A_1\) the event: “\(A_1(e) = \text{true}\)” and \(A_1^c\) the event “\(A_1(e) = \text{false}\)”
- \(A_2\) the event: “\(A_2(e) = \text{true}\)” and \(A_2^c\) the event “\(A_2(e) = \text{false}\)”
- \(A_3\) the event: “\(A_3(e) = \text{true}\)” and \(A_3^c\) the event “\(A_3(e) = \text{false}\)”
- \(S\) the event “\(e\) is a spam” and \(S^c\) the event “\(e\) is not a spam”.

I assume that for a given email \(e\), \(A_1, A_2, A_3\) are conditionally independent knowing \(S\).

Now, I use the list of classified emails to determine \(P(S)\), \(P(A_1)\), \(P(A_2)\), \(P(A_3)\), \(P(A_1|S)\), \(P(A_2|S)\), \(P(A_3|S)\) (this is the learning step) as follows: I consider that the probability of an event is the number of emails in the classified list satisfying the event over the number of emails of the list. I obtain:

- \(P(S) = 0.4,\)
- \(P(A_1) = 0.4, P(A_2) = 0.5, P(A_3) = 0.6,\)
- \(P(A_1|S) = 0.3, P(A_2|S) = 0.6, P(A_3|S) = 0.2.\)

A new email \(e\) arrives in my mail box: it satisfies \(A_1(e) = \text{false}\), \(A_2(e) = \text{true}\) and \(A_3(e) = \text{true}\).

8.1 – Show that \(S, S^c\) form a frame of discernment.

8.2 – Use the Bayes law to give an expression of \(P(S|A_1^c \cap A_2 \cap A_3)\) involving conditional probabilities knowing \(S\) and \(S^c\).

8.3 – Give the probability that \(e\) is a spam.