

MA324 : TP2

You may use TP2.py using your favorite Python IDE (e.g. spyder). Some parts are the same as in TP1, so you may gain time if you saved your work.

Exercise 1 (*Fitting curves*)

1. Using the numpy/linspace create an 1D array x that contains 100 samples uniformly distributed on $[0, 1]$

2. Simulate and plot the model

$$y = 10 + 2x + 8x^2 + \varepsilon$$

where ε are centered normal iid with variance 2

3. Form the matrix X , the vector Y and apply the formula seen during the lecture 2 to estimate the model parameters. You can compute the matrix inverse with `np.linalg.inv`
4. Plot the obtained model and the data
5. Do the same as in questions 1-4 for the model

$$y = 10 + 2x + 8x^2 + 10 \sin(\pi x) + \varepsilon$$

.

Exercise 2 (*Orbit parameter*)

1. We observe the following data We assume that the model is $f(\theta) = \frac{p}{1-e \cos(\theta)}$, as in EX1 of

Angles in degrees	43	45	52	93	108	126
Observed values	4.7126	4.5542	4.0419	2.2187	1.8910	1.7599

TD2. Estimate the parameters p and e from this data

Exercise 3 (*Dealing with outliers : (simplified) ransac algorithm*)

1. Using the numpy/linspace create an 1D array x that contains 100 samples uniformly distributed on $[0, 1]$ and build the model $y = 10 + 2x$.
2. Modify at random (uniformly) say 10 values of y and replace these values by realizations of iid centered Gaussian random variable of unit variance and plot the result
3. Write a function that given two models decide¹ if a model has more inliers than the other. We say that (x, y) is an inlier if for model parameter β_1, β_2 we have $|\beta_1 x + \beta_2 - y| < Threshold$. Where *Threshold* is a given value. Hint : compare the number of inliers for both models.
4. Write a function that pick 2 samples at random (uniformly) and compute the associated model using these 2 samples and return the computed parameters
5. Set *Threshold* = 1, and iterate say 10 times : computing a model at random ; compare it with new one and keep the new one if the new model has more inliers
6. Compare both strategies : the naive one done in TP1 where we performed the estimation keeping the outliers and the model estimated via ransac.

1. that is return a boolean that indicates if a model is more suitable than the other