The homework for Xmas holidays has 3 different projects.
Do them individually and explain the design in a word document.
Upload them in the virtual campus before the 15th January 2024.

Good Luck!!

First project

Given the function

$$f(x) = \ln(1 + x^2) \ \ x\epsilon[0, e - 1]$$

You must develop an algorithm that performs the following tasks.

1. Perform the continuous regression of $\boldsymbol{f}(\boldsymbol{x})$ onto the subspace $\langle \mathbf{1}, \boldsymbol{x}, \boldsymbol{x}^2, \boldsymbol{x}^3 \rangle$, using the symbolic tools in Matlab. Plot the result (the function and the fitting).

2. Perform the discrete regression of $\boldsymbol{f}$ onto $\langle \mathbf{1}, \boldsymbol{x}, \boldsymbol{x}^2, \boldsymbol{x}^3 \rangle$. For that purpose:

    1. Introduce as regular sampler of the interval $[0, e - 1]$ using 1000 points. Call this sampler $\boldsymbol{x_s}$.
    2. Evaluate $\boldsymbol{f}(\boldsymbol{x})$ on $\boldsymbol{x_s}$, and the basis functions $\mathbf{1}, \boldsymbol{x}, \boldsymbol{x}^2$, $\boldsymbol{x}^3$ to generate the set of vectors $\{\mathbf{1}, \boldsymbol{x}, \boldsymbol{x}^2, \boldsymbol{x}^3, \boldsymbol{f}\}$.
    3. Perform the regression of $\boldsymbol{f}$ onto $\langle \mathbf{1}, \boldsymbol{x}, \boldsymbol{x}^2, \boldsymbol{x}^3 \rangle$.
    4. Plot the result (the function and the fitting).

Second project

1. Generate a random matrix M for a web page graph with N nodes.
    - First generate a random matrix A of size N. N should an integer random number between 50 and 100.
    - Second define the Matrix M:

$$T(i,j) = 1 \; if \;\; A(i,j) < 0.6$$
$$T(i,j) = 0 \; if \;\; A(i,j) \geq 0.6$$

2. Define the Google matrix for $p = 0.1$

$$G = pT+(1\text{-}p)*B,$$

where *B* is the squared matrix of size N with ones everywhere.

3. Once you have generated G show that $\lambda = 1$ is its biggest eigenvalue. For that purpose, use the command *eig(M),* finding its spectrum and plot the absolute value of the eigen values of the spectrum, showing that the curve remains under the value $\lambda = 1$.

4. **Finding** the eigenvalue corresponding to $\lambda = 1$ by power method, that is, for an initial guess $\boldsymbol{x_0} \in \mathbb{R}^{\boldsymbol{N}} = \begin{pmatrix} 1/N \\ \vdots \\ 1/N \end{pmatrix}$, perform successive iterations $G^k\boldsymbol{x_0}$, till converging.

5. Finding the page rank.

6. Compare the result with the page rank obtained calculating the popularity of a page as the total number of web pages pointing to each node and exiting from each node.

Reference
https://pi.math.cornell.edu/~mec/Winter2009/RalucaRemus/index.html

Project 3. Projecting the Ibex35 onto their components

The file Ibex35.mat contains the information for the ibex35 market for the last 60 days and the name of the companies. The last one, called **^IBEX** is the index.

We ask you:

1. To orthogonally project the index onto the companies and finding the most important companies driving the index.

   We define the importance of company j as:

$$importance\ j = 100. {|c_j|}\Big/{\sum_{j=1}^{35} |c_j|}$$

2. Perform the PCA of this data. For that purpose:

Calling X, the data matrix in ivex35.mat

1. Find the centered matrix:

$$Xc = X - \mu$$

   where μ is the average of X column wise, that is, the center of gravity of all the stocks. It has dimension 60!!

2. Find the covariance matrix:

$$C = Xc^T\, Xc.$$

3. This matrix is 60x60 and symmetric. Finding its rank and the spectrum with the command *eig.* Be careful on how the eigen values appear.

4. Finding the tree main eigenvectors (called PCAs) $\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}$ associated to the three main eigenvalues. Make sure that you select them correctly. Show that $\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}$ **is a orthonormal basis set of the PCA subspace of dimension 3.**

5. Construct $\boldsymbol{V} = [\boldsymbol{v_{1(:)}}, \boldsymbol{v_{2(:)}}, \boldsymbol{v_{3(:)}}]$ **and showing that** $\boldsymbol{V^T V} = \boldsymbol{I_3}$

6. Now project all the columns of Xc onto $\boldsymbol{V}$ and finding the 3 main pca coordinates of each stock.

7. Plot the two first coordinates in a 2D plot and the three on a 3D plot, and identify how many clusters might exist in the 2D and 3D PCA space.

8. In which cluster the Ibex35 is located?

The teachers
Oviedo 18 December 2022