Mathematics for Data Science

Academic Year: (2023 / 2024) **Review date: 06-09-2023**

Department assigned to the subject: Mathematics Department

Coordinating teacher: BRANDLE CERQUEIRA, CRISTINA

Type: Compulsory ECTS Credits : 3.0

Year : 1 Semester : 1

REQUIREMENTS (SUBJECTS THAT ARE ASSUMED TO BE KNOWN)

None

OBJECTIVES

Use advanced knowledge of Linear Algebra for its application in methods of analysis of large volumes of data. Understand the foundations of the algorithms used in the analysis of large volumes of data in order to interpret the results and their meaning and validity.

Specifically:
1) Study of matrix algebra
2) Proficiency of least squares techniques
3) Proficiency of singular value decomposition techniques

DESCRIPTION OF CONTENTS: PROGRAMME

1. Matrices
(a) Matrix Operations
(b) Change of Basis Matrix
(c) Matrix of a Linear Transformation

2. Linear Systems of Equations
(a) LU factorization
(b) Cholesky factorization
(c) Applications: Iterative Methods

3. Diagonalization
(a) Diagonalization
(b) Orthogonal Diagonalization
(c) The Power Method
(d) Markov processes

4. Least Squares Problems
(a) Data fitting
(b) Orthogonal projections and Least square problems
(c) QR factorization
(d) Constrained LSP

5. Singular Value Decomposition
(a) Singular Value Decomposition
(b) The pseudoinverse
(c) Principal component analysis

LEARNING ACTIVITIES AND METHODOLOGY

The classes consist of a mixture of presentations referring to the subject and practical use through the resolution of exercises.

* Training activities

- AF1: Theoretical lesson.
- AF2: Practical lesson.
- AF5: Tutorials.
- AF7: Individual work.
- AF8: On-site evaluation tests.

* Teaching methodologies

- MD1: Class lectures by the professor with the support of computer and audiovisual media, in which the main concepts of the subject are developed and the bibliography is provided to complement the students' learning.
- MD2: Critical reading of texts recommended by the professor of the subject: press articles, reports, manuals and/or academic articles, either for later discussion in class, or to expand and consolidate the knowledge of the subject.
- MD3: Resolution of practical cases, problems, etc. posed by the teacher individually or in groups.
- MD4: Presentation and discussion in class, under the moderation of the professor of topics related to the content of the subject, as well as case studies.
- MD5: Preparation of papers and reports individually or in groups.

ASSESSMENT SYSTEM

1. Day-to-day to assesment, C.
The grade C will be obtained from Quizzes (Q), 30%, and Tasks (T), 70%:

C = 0.3*Q + 0.7*T.

2. Final exam, E, and final grade G.

-- If C >= 5 there are two options:

1. You can decide not to take the final exam. The final grade will be G=5.
2. If you want to get more than a 5 on your final grade you will have to take necessarily the final exam. In that case your final
grade will be
G = max{0.6*C + 0.4*E, 5, E}, where E is the final exam grade.

-- If C < 5 you will have to take necessarily the final exam to pass the course. The grade, in that case, will be

G = E.

% end-of-term-examination:	60
% of continuous assessment (assigments, laboratory, practicals…):	40

BASIC BIBLIOGRAPHY

- David C. Lay, Steven R. Lay, Judi J. McDonald Linear Algebra and Its Applications, Pearson; 5 edition, 2016
- Lloy N. Trefethen; David Bau, III Numerical Linear Algebra Twenty-Fifth Anniversary Edition, SIAM , 2022
- Timothy Sauer Numerical Analysis 2e, Pearson , 2012
- W. Keith Nicholson Linear Algebra with Applications, Lyryx, Open Edition Version, 2021

ADDITIONAL BIBLIOGRAPHY

- Carl D. Meyer Matrix Analysis and Applied Linear Algebra, SIAM , 2010
- Cleve Moler Numerical Methods with Matlab, SIAM , 2004
- David Watkins Fundamentals of Matrix Computations, 3rd Ed, Wiley , 2010
- James W. Demmel Applied Numerical Linear Algebra, SIAM , 1997