SSM
Scuola Superiore Meridionale

Course title:
Numerical methods for Data Mining

Duration [number of hours]**: 24**

PhD Program [MERC/MPS/SPACE]**: MERC**

Name and Contact details of unit organizer(s):

Prof. Constantinos Siettos
Affiliation(s): Dipartimento di Matematica e Applicazioni "Renato Caccioppoli"
Email: constantinos.siettos@unina.it

Course Description [max 150 words]**:**

The course of the Numerical methods for Data Mining aims at giving to the students the necessary theoretical and computational background on various mathematical methods and numerical analysis methods for analysis of the dynamics of complex nonlinear large-scale models with a wide range of applications (socioeconomic, neuroscience, epidemiology, finance). Lectures will be also delivered by Dr. Dimitrios Patsatzis (MERC postdoc).

Syllabus [itemized list of course topics]**:**

A. Matrix algebra for the numerical analysis of complex systems.
- The fundamental theorem of linear algebra. Least squares estimation. The problem of regularization.
B. The problem of data embedding
- SVD, PCA, MDS, Kernel PCA, ISOMAP, Diffusion Maps
C. Modelling
- The General Linear Model and the Generalized Linear Model, Gaussian Process
- The Dynamic Mode Decomposition method and the Koopman operator theory
D. Analysis of Complex Systems
- The Equation-Free multiscale framework for the analysis of multiscale complex systems
- Matrix-free methods for the solution of large-scale problems. Numerical methods in the Krylov sub-space. GMRES and Arnoldi Methods.

Assessment [form of assessment, e.g., final written/oral exam, solutions of problems during the course, final project to be handed-in, etc.]**:**

This will be a hands-one course where the students will develop their own numerical algorithms to model the and solve large-scale complex problems. The assessment and cores will be based on the set of assignments given during the course and a final oral exam.

Suggested reading and online resources:

1. Wang, J., Geometric Structure of High-Dimensional Data and Dimensionality Reduction, Springer-Verlag, 2012
2. Kelley, C.T., Iterative Methods for Linear and Nonlinear Equations, SIAM, 1995. https://archive.siam.org/books/textbooks/fr16_book.pdf
3. Karlin, S., An introduction to Stochastic Modelling, Academic Press, 1998.
4. Dobson,A., Barnett, A., Introduction to Generalized Linear Models, Chapman and Hall/CRC2008
5. Notes and papers given to the students.