SSM
Scuola Superiore Meridionale

Course title:
Numerical methods for Data Mining and Machine Learning for Complex Systems

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 and Machine Learning aims at giving to the students the necessary theoretical and numerical analysis and scientific computing background on various mathematical and numerical analysis methods for dimensionality reduction, modelling of the dynamics of complex nonlinear large-scale models and the solution of the forward and inverse problem in complex systems with scientific machine learning with a wide range of applications (Fluid Dynamics, socioeconomic, neuroscience, epidemiology, finance, Crowd dynamics).

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

A. Matrix algebra for the numerical analysis of complex systems with an emphasis on matrices of large dimensionality, and the solution of matrix equations for the data mining.
- The fundamental theorem of linear algebra. Least squares estimation. The problem of regularization.
B. The problem of manifold learning/ data embedding
- SVD, PCA, MDS, Kernel PCA, ISOMAP, Diffusion Maps
- The Solution of the Pre-image Problem
C. Modelling and Numerical Methods for the approximation and analysis of stochastic Big-data
- The General Linear Model and the Gauss Markov Theorem. Uncertainty quantification
- The Dynamic Mode Decomposition method and the Koopman operator theory
D. Analysis of Complex Systems, Forward and Inverse Problems for systems modelled via Differential Equations
- Physics Informed Neural Networks
- The Equation-Free and Next Generation Equation-Free multiscale framework for the analysis of multiscale complex systems

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. Karlin, S., An introduction to Stochastic Modelling, Academic Press, 1998.

3. Notes and papers given to the students.