Postgraduate Course
Signal Processing for Big Data
(MSc)

Instructor Information

Jesús Gustavo Cuevas del Río

E-mail: gustavo.cuevas@upm.es
Work Phone: +34 91 549 57 00 Ext: 4039

Course Information

Course Description

This course addresses the challenges of signal processing techniques when applied to large-scale data. The course presents processing algorithms suitable for large-scale data tasks involving sparse signals as the Sparse Fourier transform. Other introductory topics in the course are the extension of classical signal processing on data indexed by graphs (discrete signal processing on graphs, DSP_G). DSP_G deals with the analysis and processing of data sets in which data elements are related by dependency, similarity, physical proximity, or other properties.

Another topic is the generalization of data matrices into Tensors. Tensors are data structures indexed by more than two indices. Tensors here refer to what was previously known as multiway arrays. Tensors are becoming increasingly important in analyzing Big Data. As a result of their many indices, tensors easily become very big, causing memory and storage problems. The idea is to break the data in smaller boxes, and one of the strategies consists on compressing tensors by a low rank tensor decomposition applied to both: dense or sparse tensors. At the end of all chapters, some illustrative application scenarios are given.

The course also present some innovative topics in a chapter devoted to them. The sampling acquisition theory is reviewed in the context of new Big Data scenarios, as those based on the discovery that sparsity or compressibility can be exploited when acquiring large-scale signals (Compressive Sensing) such as the Sparse Fourier transform that only uses a subset of the input data at a time.

At the end of each chapter illustrative examples with their respective application scenarios, either PYTHON language or in MATLAB, are provided.

CONTENTS:

1. Numerical techniques for small and moderate data sets.
    1.1. Matrix storage versus tensor or graph storage.
    1.2. Principal Component Analysis (PCA).

1.3. Singular Value Decomposition (SVD).
1.4. Applications.
2. Tensors. Tensors Decompositions
2.1. Introduction to tensors.
2.1.1. Covariant, contravariant and mixed tensors. Multiway arrays and tensors.
2.1.2. Tensor Algebra: operations on tensors. Tensors product.
2.2. Tensor decompositions and applications.
2.2.1. Relation between Rank of a tensor and its decomposition.
2.2.2. Tucker Decomposition (HOPCA).
2.2.3. HOSVD Decomposition.
2.2.4. Canonical Polyadic Decomposition (CP).
2.3. Visualization of information for Big Data: Tensor Networks Diagrams.
3. Signal Processing on Graphs.
3.1. Introduction. Signal on Graphs.
3.2. Discrete Signal Processing on Graphs (DSP_G).
3.2.1. Extension of basic DSP operations to Graphs: shifting, filtering, subsampling, Z transform and Fourier Transform, etc.
3.2.2. Compression of signals on graphs.
3.2.3. Applications.
3.3. Digital Signal Processing on Graphs: Frequency Analysis.
3.3.1. Fourier transform on Graphs: Spectrum of a Graph.
3.3.2. Alternative basis.
3.3.3. Frequency response of a filter on a graph. Ordering of frequencies.
3.3.4. Filters on graphs design.
3.3.5. Applications.
3.4. Advanced Topics.
3.4.1. Strategies to fastening the Fourier transform: Sparse Fourier Transform on Graphs and Fast Fourier Transform on Graphs.

Prerequisites

Digital Signal Processing fundamentals.

Probability and Stochastic Processes for Engineers

Advanced Digital Signal Processing

In addition, a working knowledge of MATLAB and PYTHON is required.

Course Goal

To develop an understanding of the concepts, mathematical techniques (specially signal processing techniques) that underlie Big Data processing.

Summary of intended course outcomes

The students will understand fundamentals and advanced concepts in Signal Processing when the signals are attached to (defined on) graphs. They will be able to learn how to store data in form of tensors, how to process and to manipulate tensors, how to operate with tensors and how to reduce and decompose tensors.

They also will learn how to perform typical operations on signals such as shifting, filtering, subsampling, Z transform and Fourier Transform, etc., but this time on signals defined on graphs. Once they learn to do Fourier Transforms on signals defined on graphs, they will understand the concept of "Spectrum of a graph" and the related concepts of basis in a graph or filtering operations on a graph. They will learn how to do these operations on the frequency domain of a graph.

Finally they will be brought into contact with modern techniques or algorithms such as fast Fourier performing the Fourier transform or adaptation of the Fourier transform to big Sparse sets of data.

By the end of the course, students should be able to:

- Deal with and analyse data stored in tensors, then reducing the range of tensor and visualize the information with the techniques explained in the course.
- Perform typical (shifting, delaying, subsampling, etc.) operations on signals defined on graphs.
- Calculate the spectrum of a signal defined on a graph.

Syllabus

Introduction: Numerical techniques for small and moderate data sets. Introduction to PYTHON.

Overview of a couple of techniques: PCA and SVD, from that many others techniques derive.

This chapter is based on report [1].
Additional material on SVD will be provided by the instructor.
Introduction to PYTHON environment. Setting of a PYTHON environment on a computer.
Assignment: development of PYTHON computer programs to simulate and study different examples of SVD and PCA techniques. Location: assigned room (or laboratory)

Tensors and tensors decomposition.
Tucker Decomposition (HOPCA).
HOSVD Decomposition.
Canonical Polyadic Decomposition (CP)

This chapter is based on references [TB-3], [2], [3], [4] and [5].
Additional material on Tensor algebra will be provided by the instructor.

Assignment: development of PYTHON or MATLAB computer programs to deal with Tensor storage of data and simulate HOPCA, HOSVD and CP decompositions. Location: assigned room (or laboratory)

Basic signal processing on graphs
Dealing with graphs. Basic operation on graphs.
Basic operations on signals defined on graphs: shifting, filtering, subsampling, Fourier transforming, Z-transforming.

This chapter is based on references [6] and [7] and [TB-1]

Additional information and software related to graphs and graphs operations will be provided by the instructor.

Assignement: development of PYTHON or MATLAB computer programs to perform operations on graphs and on signals defined on graphs. Special emphasis is set on Fourier transforming signals defined on graphs.

Signal processing on graphs: Frequency domain
Spectrum of a graphs. Obtaining the spectrum of a signal on a graph.
Filtering in the frequency domain.
Filter design in the frequency domain.

This chapter is based on references [TB-1] and [8].

Assignement: development of PYTHON or MATLAB computer programs to perform operations on graphs in the frequency domain, such as filtering. Design of a particular filter in the frequency domain.

Textbooks:

[TB-1] D. M. Cvetkovic, P. Rowlinson, and S. Simic. "An introduction to the theory of graph spectra". Cambridge University Press Cambridge, 2010.
[TB-2] Chartrand, G. "Introductory Graph Theory". Dover Books on Mathematics.
[TB-3] Kolda, T. and Bader, B. "Tensor Decompositions and Applications", SANDIA Report, SAND2007-6702.

Recommended reading material:

[RR-1] F. R. Chung. Spectral graph theory, volume 92. AMS Bookstore, 1997.

Recommended reading papers:

1. SCHLENS, Jonathon. "A Tutorial on Principal Components Analysis": Center for Neuronal Science, NYU. 2009.
2. COMON, Pierre. "Tensors: A brief Introduction", IEEE Signal Processing

Magazine, n0. 44, May 2014.

3. GORDON, C. K, "An Introduction to the Rudiments of Tensor Analysis", IEEE Transactions on Education, Vol. E-8, N0. 4, Dec. 1965.
4. KOLDA, T. G. and BADER, B. W., "Tensor Decompositions and Applications", SIAM Rev., vol. 51, n0. 3, pp. 455-500, 2009.
5. CICHOCKI, A., "Era of Big Data Processing: A New Approach via Tensor Networks and Tensor Decompositions", arXiv:1403.2048v4 [cs.ET] 24 Aug 2014.
6. SHUMAN, D. I., NARANG, S. K. FROSSARD, P., ORTEGA, A., VANDERGHEYNST, P.. "The Emerging Field of Signal Processing on Graphs", IEEE Signal Processing Magazine, Vol. 83, May 2013.
7. SANDRYHAILA, A., MOURA, J.M, "Discrete Signal Processing on Graphs", IEEE Trans. On Signal Processing, vol. 61, n0. 7, pp. 1644-1656, April 2013.
8. SANDRYHAILA, A., MOURA, J.M, "Discrete Signal Processing on Graphs: Frequency Analysis", IEEE Trans. On Signal Processing, vol. 62, n0. 12, pp. 3042-3054, June 2014.

Student Assessment Criteria

Overview of selected papers (3)	10%
Final Exam	50%
Project (computer simulations)	40%