VUB Leerstoel 2012-2013

Data-driven modelling: an integrative approach

Prof. Johan Suykens (KU Leuven, ESAT-SCD)

Program:

1. Advanced data-driven black-box modelling
(1 hour, inaugural lecture, Oct 3: 16:00-17:00, Promotiezaal D.2.01)

Increasingly, many application areas are critically depending on high-quality predictive models, such as in modelling energy consumption, traffic networks, process industry, environmental modelling, biomedicine, brain-machine interfaces.

How to learn powerful black-box models from given data? How to cope with diverse data characteristics across different supervised and unsupervised modelling tasks, static and dynamic systems, often in high dimensional input spaces or with massive data sets? In this talk a unifying picture is presented. Several future challenges are also outlined.

2. Support vector machines and kernel methods in systems, modelling and control
(2 hours, Oct 10: 15:00-17:00, Promotiezaal D.2.01)

First, the basics of support vector machines and least-squares support vector machines for regression are presented, with aspects of sparsity and robustness. The role of primal and dual model representations is explained.

In the next part, the modelling of dynamical systems with different model structures is discussed, together with the use of fixed-size kernel methods for tackling large scale problems, e.g. in applications of electricity load forecasting. It is shown how support vector machine approaches can also be used for solving optimal control problems and solving differential equations.

3. Data-driven modelling for biomedicine and bioinformatics
(2 hours, Oct 17: 15:00-17:00, Promotiezaal D.2.01)

Many problems in biomedicine and bioinformatics involve high dimensional data, for which support vector machines and kernel-based approaches may offer suitable solutions. We discuss application studies (including joint work with the SCD teams of biomedical data processing and bioinformatics) in ovarian cancer detection, survival analysis in breast cancer, brain-tumour recognition, microarray data and genomic data fusion.

In the second part we focus on the use of parametric models in combination with regularization schemes for sparsity, interpretability and structure detection. An interval coded scoring system is explained for interpretable clinical decision support with applications in gynecology. Towards matrix and tensor-based models the use of nuclear norm regularization is a key ingredient for structure detection, as will be shown for patient-specific seizure detection from multichannel EEG.

4. Kernel methods for exploratory data analysis and community detection
(2 hours, Oct 24: 15:00-17:00, Promotiezaal D.2.01)

Eigenvalue decompositions on given data matrices are a key ingredient in many problems of exploratory data analysis, such as in principal component analysis, spectral clustering, manifold learning, dimensionality reduction and data visualization.

We show how a least-squares support vector machine approach can be taken to characterize such problems. This enables making out-of-sample extensions, obtaining new model selection procedures and large scale methods. In this context we discuss:
- sparse and robust extensions to kernel principal component analysis
- methods of kernel spectral clustering (and incorporation of prior knowledge or additional label information with semi-supervised learning)
- kernel maps with reference point for data visualization (which convert eigenvalue problems into solving linear systems).

Applications that are given include image segmentation with highly sparse kernel models, clustering customer profiles and community detection in complex and evolving networks.

5. Complex networks, synchronization and cooperative behaviour
(2 hours, Oct 31: 15:00-17:00, Promotiezaal D.2.01)

A classic problem in synchronization theory is when given two coupled (identical or non-identical) nonlinear dynamical systems: under which conditions do they synchronize? Synchronization can take place even between chaotic systems, which are very sensitive to initial conditions. Cluster synchronization is a different type of synchronization which is known to take place in a wide variety of complex networks. In this context, synchronization techniques can also be applied to finding communities in networks.

In this talk we first focus on connections between synchronization theory and systems and control theory for the class of Lur'e systems, which includes several types of electrical circuits, recurrent neural networks and genetic oscillator models as examples. In this case global synchronization can be characterized by matrix inequalities for schemes related to absolute stability, robust Hinfinity synchronization, impulsive synchronization and time-delayed systems.

In the second part, we explore new applications of synchronization towards solving global optimization problems. Methods of coupled local minimizers are explained where a group of gradient-based local search methods are coupled through a synchronization scheme. This enables global search together with fast convergence, and computing an optimal interaction scheme for the group. A similar approach extends to coupled simulated annealing processes, which has been successfully applied e.g. to problems of tuning-parameter selection for kernel methods.

Biography:

Johan A.K. Suykens is a Professor with KU Leuven. He is author of the books "Artificial Neural Networks for Modelling and Control of Non-linear Systems" (Kluwer Academic Publishers) and "Least Squares Support Vector Machines" (World Scientific) and co-author of the book "Cellular Neural Networks, Multi-Scroll Chaos and Synchronization" (World Scientific). He is a Senior IEEE member and has served as associate editor for the

IEEE Transactions on Circuits and Systems (1997-1999 and 2004-2007) and for the IEEE Transactions on Neural Networks (1998-2009). He received an IEEE Signal Processing Society 1999 Best Paper (Senior) Award and several Best Paper Awards at International Conferences. He is a recipient of the International Neural Networks Society INNS 2000 Young Investigator Award for significant contributions in the field of neural networks. He has served as a Director and Organizer of the NATO Advanced Study Institute on Learning Theory and Practice (Leuven 2002), as a program co-chair for the International Joint Conference on Neural Networks 2004 and the International Symposium on Nonlinear Theory and its Applications 2005, as an organizer of the International Symposium on Synchronization in Complex Networks 2007 and a co-organizer of the NIPS 2010 workshop on Tensors, Kernels and Machine Learning. He has been recently awarded an ERC Advanced Grant 2011.