Plenary: The Optimistic Bayesian: Replica Method Analysis of Compressed Sensing

Speaker: Vivek Goyal (MIT)

Abstract

The conventional emphases in data acquisition are the density and accuracy of measurements. Compressed sensing has reminded us of the importance of modeling and has brought sparsity- and compressibility-based models to the fore. Furthermore, the variety of computations that can be performed with the same data can give a broad trade-off between computational cost and performance.

This talk will review results on sparse signal support recovery and sparse signal estimation in a non-Bayesian setting. Full support recovery in the presence of noise is difficult, and the results for estimation are quantitatively pessimistic. The replica method, a non-rigorous but successful technique from statistical physics, is used for asymptotic analysis in a Bayesian setting. This can be applied to many estimators used in compressed sensing, including basis pursuit, lasso, linear estimation with thresholding, and zero norm-regularized estimation. It reduces the analysis of compressed sensing to the study of scalar estimation problems and gives results that are more encouraging for compressed sensing.

The talk is based on joint work with Alyson Fletcher and Sundeep Rangan.

Biography

Vivek K Goyal received the B.S. degree in mathematics and the B.S.E. degree in electrical engineering from the University of Iowa and the M.S. and Ph.D. degrees in electrical engineering from the University of California, Berkeley. He was a Member of Technical Staff in the Mathematics of Communications Research Department of Bell Laboratories, Lucent Technologies and a Senior Research Engineer for Digital Fountain, Inc. He is currently Esther and Harold E. Edgerton Associate Professor of Electrical Engineering at the Massachusetts Institute of Technology.

Dr. Goyal received the Eliahu Jury Award of the University of California, Berkeley, for his thesis research, the 2002 IEEE Signal Processing Society Magazine Award, and an NSF CAREER Award. His students have been awarded three thesis or paper awards. He served a six-year term on the IEEE Signal Processing Society's Image and Multiple Dimensional Signal Processing Technical Committee, and he currently serves as permanent Co-Chair of the SPIE Wavelets conference series. Information on his forthcoming textbook coauthored with Martin Vetterli and Jelena Kovacevic is available at http://fourierandwavelets.org.

Plenary: An engineer's view of the brain: can we electronically read and write to the mind?

Speaker: Karim G. Oweiss (Michigan State University)

Abstract

Fundamental to understanding how our world is represented in our brain is the ability to observe the collective activity of ensembles of neurons acting in concert while we associate sensory stimuli with subsequent motor actions. Some recent technological advances have greatly accelerated our ability to simultaneously record and stimulate these ensembles, thereby opening up the possibility to tremendously advance our understanding of the brain and its inner workings, and to provide real time neural control of assistive devices to people with severe disabilities through sophisticated brain-machine interfaces (BMIs). These remarkable advances, however, have outstripped progress in statistical signal processing theory and algorithms specifically tailored to: 1) analyze the massive amounts of neural and behavioral data collected; 2) explain many aspects of the natural information processing mechanisms in the nervous system; and 3) perform real-time neural signal processing within the resource-constrained environment of an implantable system for clinically viable BMI applications.

In this talk, I will briefly describe some of our recent efforts towards achieving these goals. I will first discuss how some guiding principles from systems and computational neuroscience, machine learning, and compressive sensing literature may provide some useful insight into characterizing the dynamics of functional brain networks during task-specific behavior; and how changes in these dynamics, for example, during task learning or recovery from injury, can be monitored and quantified. Next, I will show how engineering a wireless, fully implantable neural interface system capable of extracting the hypothesized constituents of the neural code in a freely behaving subject can greatly help in this regard. I will conclude with a brief discussion on the potential of this framework to increase the reliability and effectiveness of brain-machine interfaces, and to develop better algorithms for statistical inference in complex networks the face of uncertainty and nonstationarity.

Biography

Karim G. Oweiss received his B.S. (1993) and M.S. (1996) degrees with honors in electrical engineering from the University of Alexandria, Egypt, and the Ph.D. degree (2002) in electrical engineering and computer science from the University of Michigan, Ann Arbor. He completed a post-doctoral training with the biomedical engineering department at the University of Michigan, Ann Arbor in 2002. In 2003, he joined the department of electrical and computer engineering and the neuroscience program at Michigan State University, where he is currently an associate professor and director of the Neural Systems Engineering Laboratory (NSEL). His research interests span the areas of statistical signal processing and information theory, neural integration and coordination in sensorimotor systems, computational neuroscience and brain-machine interfaces.

Dr. Oweiss is a member of the IEEE and the Society for Neuroscience. He served as a member of the board of directors of the IEEE Signal Processing Society on Brain Machine Interfaces, and continues to serve on the technical committees of the IEEE Biomedical Circuits and Systems, the IEEE Life Sciences, and the IEEE Engineering in Medicine and Biology societies. He was awarded the excellence in Neural Engineering award from the National Science Foundation in 2001. His lab is currently supported through multiple grants from the Neural Interfaces Program (NIP) as well as the Repair and Plasticity Program (RPP) at the National Institute of Neurological Disorders and Stroke.

Session Title: Sparse Inference
Organizer: Piotr Fryzlewicz (LSE)

Haeran Cho (joint work with Piotr Fryzlewicz) (LSE)
High-dimensional variable selection via tilting

We consider the variable selection problem in high-dimensional linear models. When the number of variables is large, sometimes even greater than the number of observations, it is often assumed that only a small number of variables actually contribute to the response. Then identifying those relevant variables improves estimation accuracy as well as model interpretability. We first note that marginal correlation screening can be misleading especially in high-dimensional problems due to high correlations among the variables. Therefore, a new way of measuring the association between each variable and the response is introduced, which ``tilts'' each variable such that the impact of its correlations with other variables is reduced in the ``tilted correlation''. The key ingredients of tilting procedure are (i) hard-thresholding of sample correlation matrix, (ii) projection of each variable onto a subspace of R^n, and (iii) rescaling of the projected variables. We study the conditions under which tilted correlation can successfully identify relevant variables and propose an iterative algorithm with tilted correlation screening step, which shows promising performance in extensive numerical experiments.

Clifford Lam (joint work with Qiwei Yao and Neil Bathia)
LSE, LSE and University of Melbourne
Estimation of large latent factor models for time series data

We focus on the dimension reduction for high-dimensional time series based on common factors. In particular we allow the dimension of time series p to be as large as, or even larger than, the sample size n. The estimation for the factor loading matrix and the factor process itself is carried out via an eigenanalysis for a p x p non-negative definite matrix, while the factors are not necessarily stationary. We show that when all the factors are strong in the sense that the norm of each column in the factor loading matrix is of the order p^(1/2), the estimator for the factor loading matrix, as well as the resulting estimator for the precision matrix of the original p-variant time series, are weakly consistent in L2-norm with the convergence rates independent of p. This result exhibits clearly that the `curse' is canceled out by the `blessings' in dimensionality. We also establish the asymptotic properties of the estimation when not all factors are strong. For the latter case, a two-step estimation procedure is preferred accordingly to the asymptotic theory. We further demonstrate how an eye-ball test of the number of factors can be performed, and prove an asymptotic result on this. Asymptotic forecasting results are also presented. The proposed methods together with all asymptotic properties are further illustrated in simulation studies, as well as an application to a real data set.

Patrick Wolfe (Harvard)
Title: Model Selection for Time-Varying Autoregressions

We discuss theory and applications of autoregressive models with time-varying coefficients, from the perspective of modeling human speech production.

Ming Yuan (Georgia Institute of Technology)
Title: High dimensional inverse covariance matrix estimation

More and more often in practice, one needs to estimate a high dimensional covariance matrix. In this talk, we discuss how this task is often related to the sparsity of the inverse covariance matrix. In particular, we consider estimating a (inverse) covariance matrix that can be well approximated by ``sparse'' matrices. Taking advantage of the connection between multivariate linear regression and entries of the inverse covariance matrix, we introduce an estimating procedure that can effectively exploit such ``sparsity''. The proposed method can be computed using linear programming and therefore has the potential to be used in very high dimensional problems. Oracle inequalities are established for the estimation error in terms of several operator norms, showing that the method is adaptive to different types of sparsity of the problem.

Hao Helen Zhang (North Carolina State University)
Sparse and Smooth Function Estimation by COSSO and Adaptive COSSO

We propose a new class of regularization models for sparse and smooth function estimation of nonparametric models, in the framework of smoothing spline ANOVA. The penalty functional is called the COmponent Smoothing and Selection Operator (COSSO), which is the sum of component norms and applies a soft-thresholding operator on functions to achieve estimation sparsity. The COSSO is shown to provide a unified framework for several recent proposals for model selection in linear models and smoothing spline ANOVA models. Theoretical properties of the COSSO estimator, such as existence and the rate of convergence, are studied. Quite recently, we further extend the COSSO by imposing adaptively-weighted penalties on different components according to their relative importance. We show that the adaptive COSSO is np-oracle, being able to achieving the optimal rate of convergence as the underlying true model were known. Numerous examples will be demonstrated to show performance of the COSSO and adaptive COSSO in various contexts.

Session Title: Overcomplete Representations
Organizer: Zoran Cvetkovic (KCL)

Patrick Wolfe (Harvard)

Title: Frame Theory and Algorithms for Time Series Estimation

Abstract:
We describe shrinkage estimators tailored to overcomplete representations as typically employed in speech/audio processing and other time series applications.

Matthew Fickus (Air Force Institute of Technology**)**

Title: Filter bank fusion frames

Abstract:
A fusion frame is a sequence of orthogonal projection operators whose sum can be inverted in a numerically stable way. When properly designed, fusion frames can provide redundant encodings of signals which are optimally robust against certain types of noise and erasures. However, up to this point, few implementable constructions of such frames were known; we show how to construct them using oversampled filter banks. In particular, we provide polyphase domain characterizations of filter bank fusion frames. We then use these characterizations to construct filter bank fusion frame versions of discrete wavelet and Gabor transforms, emphasizing those specific finite impulse response filters whose frequency responses are well-behaved.

Ozgur Yilmaz (University of British Columbia)

Title: Quantization of Compressed Sensing Measurements

Abstract:
It is now well established that compressed sensing can be used for dimension reduction. In this talk, we focus on another important problem: How can one quantize compressed sensing measurements in an efficient way? Quantization of compressed sensing measurements is typically justified by the robustness result for ℓ_1 recovery due to Candes, Romberg and Tao, and Donoho. In particular, if the measurement operator satisfies certain generic conditions, e.g., the restricted isometry property, this result guarantees that an approximation with accuracy of order ϵ can be obtained, via ℓ_1 recovery, if each measurement of a sparse signal is quantized within an accuracy of ϵ. This result, though critical for practicability of compressed sensing, does not suggest that the reconstruction accuracy can be improved by increasing the number of measurements while keeping the quantization resolution fixed, a standard approach in oversampled A/D conversion.

In this talk, we propose a novel approach to quantization of compressed sensing measurements that utilizes "noise-shaping". In particular, we show that frame theory techniques and sigma-delta quantization methods can be employed to significantly improve the reconstruction accuracy for sparse signals from their quantized compressed sensing measurements.

This is joint work with S. Gunturk, M. Lammers, A. Powell, and R. Saab.

Petros T Boufounos

Title: Distributed Scalar Quantization

Abstract:
Scalar quantization of overcomplete representations provides a disadvantageous trade-off between the oversampling rate and the reduction of the quantization error. More sophisticated quantization methods, such as Sigma-Delta quantization, can improve this trade-off but require feedback between the coefficients being quantized. Thus they are unsuitable to applications where each coefficient should be independently quantized, such as sensor networks. This talk revisits the fundamental reasons for the scalar quantization trade-off. We demonstrate that a simple, but very effective, modification to scalar quantization can overcome this trade-off, at the expense of more complex reconstruction. Furthermore, using a randomly generated frame and appropriate dither we provide strong guarantees on the reconstruction performance. Our results are applicable to the acquisition of both dense and sparse signals, which makes our approach immediately applicable to the quantization of compressive measurements. All in all, the proposed quantization method achieves efficient universal quantization with very small quantizer complexity.

Kjersti Engan (University of Stavanger)

Title: Learning overcomplete dictionaries for sparse representation of images

Abstract:
Sparse representation of signals or images using an overcomplete dictionary assumes an ability to describe the signals as linear combinations of a few atoms from the pre-specified dictionary. Thus the choice of such a dictionary, or frame, is crucial for the success of this model. This talk will focus on learning dictionaries for sparse representations by using a training set. We will emphasize on our algorithms starting with Method of optimal Directions (MOD) from 1998, developed further to the family of Iterative Least Squares Dictionary Learning Algorithms (ILS-DLA), and with our newest online dictionary learning algorithm, Recursive Least Squares Dictionary Learning Algorithm (RLS-DLA) from 2010.

We will demonstrated some of the many possible applications using learned dictionaries on sparse representation of images, such as image compression, texture classification and denoising of images.

Anna Scaglione

Title: Signal recovery based on sparse representations: ideas and solution for communication systems

Abstract:
Overcomplete representations have recently been at the center of renewed attention because of the advances in the field of compressed sensing and of sub-Nyquist sampling for signals lying on union of subspaces. In applying these ideas in communication theory the ultimate objective is to understand how many resolvable degreesof freedom in the signal are available at the transmitter to encode information.

Formulating this theory requires a clear framework to understand what features of the transmit signal allow or prevent the received signal reconstruction, and how one can fold in the equation a constraint in the sampling complexity at receiver end. In this talk we show how multipath channel distortion and other forms of synchronization error affect the parametrization of the received signal and discuss results pertinent to compressed channel sensing and signal recovery that are pertinent to communications. We will then move to a more general framework to generalize the problem and give identifiability conditions on the input. We will conclude by discussing what are the implications of the theory in the design of communication front ends.

This is joint work with Y. Eldar, Simon Li and Matthew Sharp

Alex Powell

Title: Coarse quantization for random interleaved oversampling of bandlimited signals.

Abstract:
This work investigates the compatibility of unsynchronized interleaved sampling with Sigma-Delta analog-to-digital conversion. We consider the setting where a bandlimited signal is sampled on a union of N randomly shifted sampling grids (each of which may individually correspond to sub-Nyquist sampling). We prove that if the time-ordered samples are quantized with a first order Sigma-Delta modulator then with high probability the overall quantization error is of order N^{-1} log(N). This bound is within a logarithmic factor of related error bounds for uniform sampling; the analysis makes use of tools from random matrix theory and order statistics.

This is joint work with Jared Tanner, Yang Wang, and Ozgur Yilmaz.

Tutorial: Methods of analysis in Compressed Sensing
Instructor: Dr Jared Tanner (Edinburgh)

Organizer: Dr Simon Schultz (Imperial)
Session Title: Signal Processing and Information Estimation Problems in Neuroscience

Marcelo Montemurro

Accurate assessment of timing information in neural signals

Physical and chemical signals from the external world are processed by sensory organs and converted into patterns of neural activity. A key question in Neuroscience is to understand what aspects of the neural response are most relevant for the encoding of information about an external stimulus. An ideal tool to probe different coding hypotheses is provided by information theory (IT). Within this rigorous framework it is possible to quantify the knowledge about the stimulus that can be obtained from specific features of the neural response. Thus, IT can be used to assess the performance of different candidate neural codes. However, a major practical problem with the application of information theory to neural signals is that direct estimates of Shannon information from spike or Local Field Potentials data are seriously affected by systematic errors, or bias, due to limited sampling of the stimulus response statistics. These systematic errors may falsely lead to conclude that the brain is encoding information about a particular stimulus when in fact it is not. Here I will review recently developed methods that require significantly less amounts of data to obtain tight bounds to the information conveyed by the neural response to repeated presentations of the same stimulus set.

Analysis of neural population coding using two-photon calcium imaging

Simon R Schultz

Studies of the population code used to represent and process sensory information in mammalian cerebral and cerebellar cortices have historically been based upon experimental data acquired via single-unit recording techniques, and more recently data from micro-electrode arrays (MEAs). These techniques introduce a number of systematic biases: single-unit recording is biased towards selecting cells with high spontaneous activity; MEA recording, while addressing this issue, still grossly undersamples the computational circuitry.

Recently developed multiphoton calcium imaging technology has the potential to address these problems by allowing almost every cell in an area to be recorded. However, limitations upon the SNR and temporal resolution of both microscopes and fluorophores mean that there are significant signal processing and data analysis problems to be solved in order to take maximum advantage of this technique. In this talk, I will discuss our approaches to these problems with reference to in vivo multiphoton imaging of cerebellar Purkinje cell complex spike calcium signals.

Data Analysis and visualization for multi-neuron recordings

Ken Harris (Imperial)

The brain is made of billions of neurons, which together form the world's most powerful information-processing machine. Despite decades of research, the fundamental principle by which these cells work together is still unknown. In the last few years has it become possible to record simultaneously from large enough numbers of neurons to address this question experimentally. However methods for dealing with this kind of data - highly multivariate point processes - are in their infancy. Here we discuss some methods for analysis and visualization of neuronal population recordings, and some conclusions obtained using them on the function of the cortex.

Gaussian Process methods for estimating orientation preference maps

Jakob Macke (UCL)

Gaussian process methods for estimating orientation preference maps

A striking feature of cortical organization is that the encoding of many stimulus features, for example orientation or direction selectivity, is arranged into topographic maps. Functional imaging methods such as optical imaging of intrinsic signals, voltage sensitive dye imaging or functional magnetic resonance imaging are important tools for studying the structure of cortical maps. As functional imaging measurements are usually noisy, statistical processing of the data is necessary to extract maps from the imaging data. I will present a probabilistic model of functional imaging data based on Gaussian processes. In comparison to conventional approaches, the model yields superior estimates of cortical maps from smaller amounts of data. In addition, it allows us to obtain quantitative uncertainty estimates, i.e. error bars on properties of the estimated map. I will show how the probabilistic model can be used to study the coding properties of the map and the role of noise correlations by decoding the stimulus from single trials of an imaging experiment.

Session Title: Estimation and Optimization in Machine Learning
Organizer: Massi Pontil

Alessandro Verri (DISI - Universita' degli Studi di Genova)

Title: Regularized Nonlinear Variable Selection

Abstract: A regularization method for nonlinear variable selection is proposed. Selection is obtained through a data dependent term which penalizes large values of partial derivatives of the regression function. In the talk, we first motivate the problem and nature of the penalty term. Then, we describe an algorithm for solving the underlying optimization problem and discuss its convergence. Finally, we report numerical experiments and some preliminary results on real data which show the merit of the proposed method.

Alain Rakotomamonjy (Université de Rouen)

Title: learning large-margin discriminative wavelet dictionaries

Abstract: This paper addresses the problem of optimal feature extraction from wavelet representations. Our work aims at building features by selecting wavelet coefficients resulting from signal or image decomposition on a adapted wavelet basis. For this purpose, we jointly learn in a large-margin context the wavelet shapes as well as the appropriate scale and translation of the wavelets. This problem is posed as a multiple kernel learning problem where the number of kernels can be very large. For solving such a problem, we introduce a novel multiple kernel learning algorithm based on active constraints methods. We furthermore propose some variants of this algorithm that can produce approximate solutions more efficiently. Empirical analysis shows that our active constraint MKL algorithm achieves state-of-the art efficiency. When applied to wavelet kernel learning, our experimental results show that our proposed approaches is competitive with respect to the state of the art wavelet dictionary learning methods.

Raphael Hauser (Computing Laboratory, University of Oxford)

Title: A Parallel SVD Algorithm for Use in Matrix-Completion Problems

Abstract: A popular approach to the low-rank matrix completion problem is to relax it by a nuclear-norm minimization problem. Algorithms to solve this relaxation typically make it necessary to solve a large number of thin SVD computations on large matrices. The SVD computation represents an important bottleneck in the overall algorithm. Our new algorithm allows to split these computations and distribute them across a network of loosely coupled computational resources that may run at different speeds.

Mark Herbster (Dept of Computer Science, UCL)

Title: Resistive geometry for graph-based transduction.

Abstract: In graph-based transduction we are given a fixed set of objects, some of which are labeled and some of which are unlabeled, and we wish to predict the unlabeled objects. A graph is then defined where an edge between objects indicates similarity

between objects. If the graph is weighted then the weights indicate the degree of similarity. These include the min-cut method and the harmonic energy minimisation. We interpret these methods as specific instances of the minimization of a p-energy. When p=2 the analogy is that the graph is an electrical network; the edges are now resistors whose resistance is reciprocal of the similarity. The fixed labels from {-1,1} now correspond to potential (voltage) constraints and the algorithm for labeling the graph is then to find the set of consistent voltages which minimize the energy dissipation and then to predict with the ``sign'' of the voltages. We extend the analogy for general p, which leads to natural analogues of Kirchoff's laws, Ohm's law, the conservation of energy principle, and the "rules" of resistors in series and parallel.

Jean Morales (Dept of Computr Science, UCL)

Title: Structured sparsity with convex penalty functions.

Abstract: A linear regression can benefit from the knowledge that the underlying vector is sparse. The combinatorial problem of selecting the nonzero components of that vector can be convexified by regularising the least squares error with a penalty function like the Lasso. However in real data there is often more information about the vector: this is not only known to be sparse, but possibly some prior knowledge about the sparsity pattern can be used to improve predictive performance. Here, a convex penalty function is presented which encodes such a prior knowledge as a set of linear constraints on the absolute values of the components of the vector. This penalty function reduces to the Lasso as a special case, but it is flexible enough to allow a large set of constraints to be taken into account, thus modelling many different patterns of sparsity.

Andres Argyriou (Toyota Technology Institute, Chicago)

Title: Spectral Regularization and Multi-task Learning

Abstract: Multi-task learning and transfer learning extend the standard paradigm of supervised learning. In multi-task learning, samples for multiple related tasks are given and the goal is to learn a function for each task while generalizing well (transferring learned knowledge) on new tasks. Many interesting applications, as well as formulations like matrix completion or matrix factorization can be expressed by this framework. I will present some formulations of such problems as convex programs with matrix variables. Of special interest are some recent approaches based on regularization with Schatten Lp norms, such as the trace (nuclear) norm. I will also briefly discuss possible algorithms for these problems.

Curzio Basso (DISI - Universita' degli Studi di Genova)

Title: Learning Frames for Sparse Coding

Abstract: Recently, considerable research efforts have been devoted to the design of methods to learn from data overcomplete dictionaries for sparse coding. However, coding new data with a generic dictionary requires the solution of a possibly expensive optimization problem. We suggest to overcome this drawback by learning, together with the dictionary, a matrix directly mapping the data to their optimal encodings.

We propose two strategies to accomplish this, both based on the theory of frames in finite dimensions and solved leveraging on proximal methods. The algorithm jointly minimizes the reconstruction error of the dictionary, a penalty on the dictionary depending on the scheme adopted, and an ℓ_1-based, sparsity-inducing penalty on the encodings. The results obtained on synthetic data and real images show that the algorithms are capable of recovering dictionaries with the expected characteristics.

Contributed Talks – Sparsity and Estimation
Chair: Pier Luigi Dragotti (Imperial)

Speaker: Bubacarr Bah (with Jared Tanner)
Title: Improved bounds on restricted isometry constants for Gaussian matrices

The Restricted Isometry Constants (RIC) of a matrix A measures how close to an isometry is the action of A on vectors with few nonzero entries, measured in the l^2 norm. Specifically, the upper and lower RIC of a matrix A of size n N is the maximum and the minimum deviation from unity (one) of the largest and smallest, respectively, square of singular values of all $\binom{N}{k}$ matrices formed by taking k columns from A. Calculation of the RIC is intractable for most matrices due to its combinatorial nature; however, many random matrices typically have bounded RIC in some range of problem sizes (k,n,N). We provide the best known bound on the RIC for Gaussian matrices, which is also the smallest known bound on the RIC for any large rectangular matrix. Our results are built on the prior bounds of Blanchard, Cartis, and Tanner in *Compressed Sensing: How sharp is the Restricted Isometry Property?* with improvements achieved by grouping submatrices that share a substantial number of columns.

Spearker: Yingsong Zhang
Title: Fast and accurate L0-based Sparse Signal Recovery

Speaker: Aris Gretsistas
Title: Stagewise Conjugate Gradient Polytope Faces Pursuit for large-scale sparse recovery problems

Speaker: Kezhi Li
Title: Deterministic Compressed-Sensing Matrices: where Toeplitz meets Golay