AALBORG UNIVERSITY
DENMARK

Mathematical and Statistical Analysis of Spatial Data

June 1-3, 2015

Organizers: *Morten Nielsen and Rasmus Waagepetersen*

ABSTRACTS

Bogdan Roman

"On practical compressed sensing, models and asymptotic structure"

The real world is, as usual, cruel and unforgiving when it comes to modelling its phenomena. This talk will explain why some of the fundamental conditions in classical compressed sensing (CS) are problematic or missing in a large number of practical applications, then introduce a new CS theory based on asymptotic behaviour that real-world signals possess, and show why the new concepts are a better fit and how they can be exploited to achieve substantial gains over the classical CS setup. In particular, we shall see that the success of CS depends critically on the signal structure and on its resolution. I will then pose an important question: What is then the right model? In other words, what type of signals do we recover with CS strategies that have been successfully used in practice (e.g. in MRI)? What should a future CS theory be able to describe? I will discuss a few CS models that have been proposed and a series of so called "flip tests" that can be used to test the soundness of a model, as well as endeavour to showcase examples from- or with direct applicability in practice.

Douglas Nychka

Some theory to understand the asymptotic accuracy of fixed rank Kriging

Douglas Nychka (National Center for Atmospheric Research) and William Kleiber (University of Colorado)

Kriging is a nonparametric regression method used in geostatistics for estimating curves and surfaces for spatial data. The lack of statistical large sample theory for these very useful methods is a contrast to the well developed mathematical analysis of kernel smoothers. This talk outlines an approach to understand the mathematical properties of Kriging with a focus on a recent variation known as fixed rank Kriging. One useful feature of the fixed rank representation is the use of compacted supported basis functions that allow efficient computation for large data sets. The Kriging estimate, normally derived as the best linear unbiased estimator, is also the solution of a particular penalized least squares problem. In fact Kriging estimators can also be interpreted as generalized smoothing splines where the roughness penalty is determined by the covariance function of a spatial process. We develop an equivalent kernel interpretation of geostatistical estimators that extends to fixed rank Kriging. The complete application of this theory to a multi-resolution type model is still incomplete and may be an opportunity for harmonic analysis to provide rigorous proofs.

Thomas Arildsen

"Faster Imaging of Nanoscale Surfaces"

It is important in many areas of science to be able to measure details at nanoscale or, in principle as small scale as possible. The application areas range over the medical, biological, materials science areas etc. Examples could be investigating the surface structure of some catalyst material or investigating the shape and other surface properties of biological cells. The wavelengths of visible light pose natural restrictions on how small features can be resolved in optical microscopy and therefore other techniques have been developed. Some of these techniques, such as scanning electron microscopy and atomic force microscopy, allow imaging down to individual atoms in some cases. In our work, we focus on atomic force microscopy. The speed at which an atomic force microscope can image is limited by the fact that the microscope must sense the surface of the specimen by moving a tiny mechanical probe around the surface, measuring the surface at one discrete point at a time. In order to image a surface region, this is typically done in a dense raster pattern - a relatively slow procedure depending on the size of the region of interest and the desired resolution. In order to improve the image acquisition speed, we investigate the use of advanced image processing algorithms to reconstruct an image of the surface from fewer (less dense) measurements. Here we review and discuss some of the challenges and possible approaches to image reconstruction in atomic force microscopy and demonstrate some of our recent results.

Peter Craigmile

Wavelet-based estimation for spatio-temporal processes

Wavelet transforms can be used to decompose a spatio-temporal process into an average (scaling) process, and difference of averages (wavelet) processes over successively longer scales. These processes can be used to study multi-scale behaviour and through decorrelation properties of the wavelet filter can allow for simplified modeling of the original process. In this talk we provide the statistical properties of these scaling and wavelet processes for analyzing long-range dependent spatio-temporal processes, and discuss the statistical properties of parameter estimates based on the so-called wavelet-Whittle method. We apply our methodology to a climate dataset.

This research project is joint with Wenjun Zheng, now at Fiserv.

Bubacarr Bah

Compressed sensing: structured sparsity and sparse operators

Compressed sensing seeks to exploit the simplicity (sparsity) of a signal to under sample the signal significantly. Sparsity is a first order prior information on the signal. In many applications signals exhibit an additional structure beyond sparsity. Exploiting this second order prior information about the signal enables further sub-sampling but also improves recovery. On the other hand, a lot of the sampling operators for which we are able to prove optimal recovery guarantees are dense and hence don't scale well with the dimension of the signal. Sparse operators scale better than their dense counterparts but they are more difficult to give provable guarantees on. This motivated our research on the use of such sparse sampling operators that are adjacency matrices of lossless expander graphs.

This talk will give a general overview of results on structured sparsity in compressed sensing (model-based compressed sensing). It will discuss sampling and recovery in model-based compressed sensing generally but will narrow down to give latest results our work on model-based compressed sensing with sparse sensing matrices from expanders.

Jesper Møller

Determinantal point processes on the sphere

Determinantal point processes (DPPs) are models for repulsiveness between points in `space', where the two most studied cases of `space' is a finite set or the d-dimensional Euclidean space. Their moment properties and density expressions are known; they can easily and quickly be simulated; rather flexible parametric models can be constructed and likelihood or moment based inference procedures apply; and they are used in mathematical physics, combinatorics, random-matrix theory, machine learning and spatial statistics – see [1] and the references therein. This talk concerns DPPs defined on the d-dimensional unit sphere $\mathbb S^d$ where the one- and two-dimensional cases are the practically most relevant cases. The connection between Mercer and Schoenberg representations of the kernel for a DPP will be established, whereby we can construct new parametric models for kernels (covariance functions) which provide tractable and flexible DPP models on $\mathbb S^d$. The talk is based on [2].

References:

[1] F. Lavancier, J. Møller and E. Rubak (2015). Determinantal point process models and statistical inference. To appear in *Journal of Royal Statistical Society: Series B (Statistical Methodology)*.

[2] J. Møller, M. Nielsen, E. Porcu and E. Rubak (2015). Determinantal point processes on the sphere. In preparation.

Remi Gribonval

Dictionary learning

Sparse modeling has become highly popular in signal processing and machine learning, where many tasks can be expressed as under-determined linear inverse problems. Together with a growing family of low-dimensional signal models, sparse models expressed with signal dictionaries have given rise to a rich set of algorithmic principles combining provably good performance with bounded complexity. In practice, from denoising to inpainting and super-resolution, applications require choosing a "good" dictionary. This key step can be empirically addressed through data-driven principles known as dictionary learning.

In this talk I will draw a panorama of dictionary learning for low-dimensional modeling. After reviewing the basic empirical principles of dictionary learning and related matrix factorizations such as PCA, K-means and NMF, we will discuss techniques to learn dictionaries with controlled computational efficiency, as well as a series of recent theoretical results establishing the statistical significance of learned dictionaries even in the presence of noise and outliers.

Finn Lindgren

Multiscale spatio-temporal modelling and different uses of sparsity

Marc Genton

Incorporating geostrophic wind information for improved space-time short-term wind speed forecasting and power system dispatch

Accurate short-term wind speed forecasting is needed for the rapid development and efficient operation of wind energy resources. This is, however, a very challenging problem. We propose to incorporate the geostrophic wind as a new predictor in an advanced space-time statistical forecasting model, the trigonometric direction diurnal (TDD) model. The geostrophic wind captures the physical relationship between wind and pressure through the observed approximate balance between the pressure gradient force and the Coriolis acceleration due to the Earth's rotation. Based on our numerical experiments with data from West Texas, our new method produces more accurate forecasts than does the TDD model using air pressure and temperature for 1- to 6-hour-ahead forecasts based on three different evaluation criteria. For example, our new method obtains between 13.9% and 22.4% overall mean absolute error improvement relative to persistence in 2-hour-ahead forecasts, and between 5.3% and 8.2% improvement relative to the best previous space-time methods in this setting. By reducing uncertainties in near-term wind power forecasts, the overall cost benefits on system dispatch can be quantified. The talk is based on joint work with Kenneth Bowman and Xinxin Zhu.

Jakob Lemvig

Reproducing Formulas for Generalized Translation Invariant Systems

Frame theory has become a central tool in analyzing and representing multivariate functions, e.g., curvelet and shearlet frames are known to provide optimally sparse approximations of signals in the class of cartoon-like images. In this talk we consider representations of functions using the so-called generalized translation invariant (GTI) frames. This is very flexible setup, covering many of the familiar representations, e.g., wavelets, shearlets, curvelets, and Gabor systems, both in their discrete and continuous variants. We give an introduction to frames and generalized translation-invariant systems. We focus on characterization of those generators of GTI systems that leads to convenient reproducing formulas. We will also mention a Walnut representation for GTI systems and its use in construction of GTI frames.

The talk is based on joint with M.S. Jakobsen (TU Denmark).

William Kleiber

Coherence for Random Fields

Multivariate spatial field data are increasingly common and whose modeling typically relies on building cross-covariance functions to describe cross-process relationships. An alternative viewpoint is to model the matrix of spectral measures. We develop the notions of coherence, phase and gain for multidimensional stationary processes. Coherence, as a function of frequency, can be seen to be a measure of linear relationship between two spatial processes at that frequency band. We use the coherence function to illustrate fundamental limitations on a number of previously proposed constructions for multivariate processes, suggesting these options are not viable for real data. We also give natural interpretations to cross-covariance parameters of the Mat\'ern class. Estimation follows from smoothed multivariate periodogram matrices. We illustrate the estimation and interpretation of these functions on two datasets, forecast and reanalysis sea level pressure and geopotential heights over the equatorial region. Examining these functions lends insight that would otherwise be difficult to detect and model using standard cross-covariance formulations.

Emilio Porcu

Nonseparable space-time covariance functions depending on the great circle distance