

SPMS 2014

Stochastic and Physical Monitoring Systems

Book of abstracts of the international conference

June 23 - June 28, 2014, Hotel Kavka, Malá Skála, Czech Republic

Organized by the Group of Applied Mathematics and Stochastics (GAMS),
Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering,
Czech Technical University in Prague

Supported by the SGS project SVK 33/14/F4

Foreword

Dear Colleagues,

it is our pleasure to welcome you to the 5th STOCHASTIC & PHYSICAL MONITORING SYSTEMS.

The SPMS 2014 conference is held for the fifth time with the aim to bring together students and researchers with areas of interest related to the following topics: Analysis of microscopical structure of vehicular traffic streams and traffic modeling, Monitoring and classification of acoustic signals in material defectoscopy, New statistical distances and informational divergences with applications in acoustic emission, Small area estimation of geographical characteristics gained from data sets, and Data processing in high energy particle physics, which are all together briefly called Stochastic and Physical Monitoring Systems (SPMS). The conference links both the informal character of lively student meeting and the unique platform for the research presentations and discussions of the conference participants.

The meeting is organized by the Group of Applied Mathematics and Stochastics, Department of Mathematics, Czech Technical University in Prague and this year's venue is hotel Kavka in Malá Skála Nebřich in the scenic landscape park Bohemian Paradise.

Local Organizing Committee:

Jiří Franc, Pavel Hrabák, Tomáš Hobza, Václav Kůs (Chair).

Scientific and Program Committee:

Václav Kůs, Tomáš Hobza, Milan Krbálek.

Invited Speakers:

Lucie Fajfrová, UTIA - Czech Academy of Sciences, Prague, Czech Republic,
Jiří Grim, UTIA - Czech Academy of Sciences, Prague, Czech Republic,
František Hakl, ICS - Czech Academy of Sciences, Prague, Czech Republic,
Milan Chlada, IT - Czech Academy of Sciences, Prague, Czech Republic,
Domingo Morales, Universidad Miguel Hernández de Elche, Spain,
Zdeněk Převorovský, IT - Czech Academy of Sciences, Prague, Czech Republic.

We gratefully acknowledge the institutional support provided by the funds of specific research, namely by the SGS grant SVK 33/14/F4.

The Organizers

Contents

Foreword	i
Contents	ii
In-vivo methods to determine human skin anisotropy using multi-directional ultrasonic probe	
<i>Marine Amelineau</i>	1
Parameter estimation in generalized linear models for dependent data	
<i>Petr Bohuslav</i>	1
Divergence methods in statistical separations	
<i>Petr Bouř, Václav Kůs</i>	2
Stationary distributions of interacting particle systems	
<i>Lucie Fajfrová</i>	2
The middle-ranged interactions in Dyson's particle gases	
<i>Ondřej Faltys</i>	3
Application of multivariate discrimination methods in the measurement of the inclusive top pair production cross section	
<i>Jiří Franc</i>	4
Approximating Probability Densities by Mixtures of Gaussian Dependence Trees	
<i>Jiří Grim</i>	4
Time-headways for interacting particle systems in stationary state, volume II	
<i>Pavel Hrabák</i>	5
Simulation study for consistency and robustness of Cramron Mises type estimators	
<i>Jitka Hrabáková, Václav Kůs</i>	5
Evaluation of time reversal transfer experiment with different signal lengths	
<i>Adrien Chapuis, Jan Kober</i>	6
Mapping of general dependencies among data by optimized BP-networks	
<i>Milan Chlada, Zdeněk Převorovský</i>	7
Random Matrix Revolutions	
<i>Ondřej Kollert</i>	7

Stochastic epidemic models	
<i>Jitka Kostková</i>	8
Robust Approach to Time Series Prediction	
<i>Vojtěch Kovalík</i>	8
Bayesian stochastic simulation oriented inference with approximate Bayesian computation	
<i>Jakub Kubiš</i>	9
Goodness of fit based variable ranking in high energy physics data	
<i>Jan Kučera</i>	9
Active learning for 3D navigation of UAVs	
<i>Jakub Malý</i>	10
Approximate Bayesian computation filtering	
<i>Filip Marada</i>	10
Clustering Method of IDM model with relation to the GNU model	
<i>Michael Matějů, Pavel Hrabák</i>	11
Divergence Decision Tree application on a data set, success of classification using SVM	
<i>Michaela Sluková, Zuzana Dvořáková</i>	11
Parallel evaluation of MVA methods without ROOT dependency	
<i>Vladimír Španihel</i>	12
Modification of gaussian mixture models for data classification in high energy physics	
<i>Michal Štěpánek</i>	13
Determination of phase transitions in two-dimensional model TASEP	
<i>Lenka Ščepánková</i>	13
Model for prediction of power consumption of a hybrid vehicle	
<i>Vít Škvára</i>	14
Annealing-based car-following model	
<i>Jiří Šleis, Milan Krbálek</i>	14
Option pricing	
<i>Viktor Šohájek</i>	15
Acoustic emission testing	
<i>Aurélie Teyssier, Zuzana Dvořáková</i>	15
Application of mixed logistic model in small area estimation	
<i>Aleš Tygl</i>	16
Some notes on SDE's in infinite dimension	
<i>Petr Veverka</i>	16
Stochastic control of tumor growth	
<i>Hynek Walner</i>	17
Program	19

In-vivo methods to determine human skin anisotropy using multi-directional ultrasonic probe

Marine Amelineau

INSA Centre Val de Loire, Blois, France

Email: marineamelineau@hotmail.com

Abstract

Human skin is a complex tissue and embodies specific mechanical behavior as inhomogeneity, anisotropy and nonlinear viscoelasticity. This paper is focused on investigation of skin anisotropy characteristics. New findings in this area are of great interest in dermatology and plastic surgery. Anisotropy characteristics of the skin are determined using two different methods and a special multi-directional flexible ultrasonic probe. The methods are based on determination of ultrasonic wave propagation velocity and attenuation. Combination of both methods and the probe enables local investigation of skin anisotropy in vivo. Methods were tested on forearm skin tissue and seems to be promising for local evaluation of the local anisotropy of inter-individual skin. Results obtained by using and comparing these two methods could help us to understand the mechanical behavior of the skin, its anisotropy properties and their changes depending on the environment.

Key words: Ultrasonic methods, Skin anisotropy.

Parameter estimation in generalized linear models for dependent data

Petr Bohuslav

Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Prague, Czech Republic

Email: peta.bohuslav@centrum.cz

Abstract

Generalized linear models represent a class of various statistical models such as linear regression, ANOVA, logistic regression and Poisson regression. These models are widely used in statistical applications and usually assume uncorrelated data. On the contrary, in longitudinal studies, for example, correlation between observations is encountered. In this case, the method called generalized estimating equations (GEE) can be employed to estimate the parameters of the model. This method gives consistent estimates of the regression parameters of the studied model under mild assumptions about

the time dependence. Let $\widehat{\beta}_G$ denote the estimate computed with GEE considering any given correlation structure and let $\widehat{\beta}_I$ be the estimate calculated with the Fisher scoring algorithm under the assumption of independence. The issue of efficiency of $\widehat{\beta}_G$ compared to $\widehat{\beta}_I$ will be discussed as well as the question of whether it is worth to use the GEE method instead of methods assuming independence regarding its computational complexity.

Key words: Generalized estimating equations, Parameter estimation, Dependent data.

Divergence methods in statistical separations

Petr Bouř, Václav Kůs

Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Prague, Czech Republic

Email: petr.bour@seznam.cz

Abstract

Using binary decision trees is a popular tool for unsupervised classification of high-dimensional data. We propose to modify the binary decision tree in order to perform supervised classification of data obtained from experiments in Fermilab. A key insight we provide is clustering driven only by a few physical parameters in which Rényi divergence obtained from quantile histograms of signal and backgrounds is maximal. Further we present our initial separation with its selected optimal parameters. Also, we show our results of several tests based on the weighted empirical distribution function and quantile histograms.

Key words: Binary decision trees, Quantile histograms, Rényi divergence, Supervised classification.

Stationary distributions of interacting particle systems

Lucie Fajfrová

Institute of Information Theory and Automation of the ASCR, Prague, The Czech Republic

Email: fajfrova@utia.cas.cz

Abstract

The lecture is focused on special stochastic processes which model complex phenomena related with a large number of interacting components well known from the work of Spitzer and Liggett (from the 70's) as *interacting particle systems* (IPS). We are particularly interested in so called lattice gas models where particles move among sites of a lattice, there are no creations and no annihilations possible. *The exclusion process* and *the zero range process* are the most famous such models but we will also mention more general models. We will focus on stationary distributions of these stochastic systems. The point

of this lecture is to introduce the main principles of *how to find stationary measures and how these are used* when dealing with the variants of underlying space (such as infinite lattice, finite lattice with a fixed number of particles and finite lattice with open boundaries).

Analysing properties of stationary measures enable us (among others) a mathematical insight into a phase separation and condensation phenomena which can be seen in these systems under certain conditions. This led some authors to a comparison with transport models, where particles represent cars, current of particles represents a flux of cars on the motorway and condensation phenomenon translates into appearance of a phantom traffic jam. Note here that IPS are basically described by their microscopic dynamics, i.e. local stochastic rules for interactions between individual components, and of course we are not able (it is not desirable) to deal with many particles individually when solving a traffic behaviour. We do prefer to predict behaviour on macroscopic or mesoscopic level using terms like flux or density of particles. This task has many faces from a general mathematical problem to a traffic modelling application and brings a lot of open questions.

Key words: Interacting particle system, Stationary measures, Zero range process.

The middle-ranged interactions in Dyson's particle gases

Ondřej Faltys

Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Prague, Czech Republic

Email: faltyond@fjfi.cvut.cz

Abstract

As is well known, deeper comprehension of the micro-dynamics in particle Dyson's gases seems to be essential for understanding the functioning of transport flows. The investigated model is represented by an ensemble of identical particles moving along the ring and being repulsed by middle-ranged forces. The main effort of our considerations can be found in derivation of the spacing distribution $p_\beta(r)$, where r denotes the distance-clearance between two neighboring particles. Furthermore, the parameter β corresponds to the inverse temperature of the associated thermal ensemble. Interactions in the investigated model are described by means of middle-ranged repulsive potentials. We derive the spacing distributions $p_\beta(r)$ for two various types of the middle-ranged repulsive potentials – the perturbative and logarithmic (Coulomb) potential.

Key words: Dyson's gases, Spacing distribution, Repulsive potentials.

Application of multivariate discrimination methods in the measurement of the inclusive top pair production cross section

Jiří Franc

Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Prague, Czech Republic

Email: jiri.franc@fjfi.cvut.cz

Abstract

The application of multivariate analysis techniques in experimental high energy physics have been accepted as a one of the fundamental tools in the discrimination phase, when signal is rare and background dominates. The purpose of this study is to present new approaches of the variable selection and discrimination together with the validation techniques used in High Energy Physics. We applied new MVA methods in the real data analysis and compare it with familiar methods from ROOT TMVA package. The progress of the separation by the Generalized Linear Models (GLM), Gaussian Mixture Models (GMM), Neural Networks with Switching Units (NNSU), TMVA Boosted Decision Trees, and Multi-layer Perceptron (MLP) in the measurement of the inclusive top pair production cross section on D0 Tevatron full RunII ($9.7fb^{-1}$) will be presented and possibilities of discrimination improvement will be discussed.

Key words: Multivariate methods analysis, Discrimination, Variable selection, Distribution tests, Top quark.

Approximating Probability Densities by Mixtures of Gaussian Dependence Trees

Jiří Grim

Institute of Information Theory and Automation of the ASCR, Prague, The Czech Republic

Email: grim@utia.cas.cz

Abstract

Considering the probabilistic approach to practical problems we are increasingly confronted with the need to estimate unknown multivariate probability density functions from large high-dimensional databases produced by electronic devices. The underlying densities are usually strongly multimodal and therefore mixtures of unimodal density functions suggest themselves as a suitable approximation tool. In this respect the product mixture models are preferable because they can be efficiently estimated from data by means of EM algorithm and have some advantageous properties. However, in some cases the

simplicity of product components could appear too restrictive and a natural idea is to use a more complex mixture of dependence-tree distributions. By considering the concept of dependence tree we can explicitly describe the statistical relationships between pairs of variables at the level of individual components and therefore the approximation power of the resulting mixture may essentially increase.

Key words: Multivariate statistics, Mixtures of dependence trees, EM algorithm, Pattern recognition, Medical image analysis.

Time-headways for interacting particle systems in stationary state, volume II

Pavel Hrabák

Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Prague, Czech Republic

Email: pavel.hrabak@fjfi.cvut.cz

Abstract

Time-headway distribution is considered to be important characteristic of traffic flow. Therefore, it is very useful to investigate such characteristic in the scope of interacting particle systems used for traffic flow modeling. Focusing on the simplest (ZRP, TASEP), a method for analytical derivation of the distribution is presented for TASEP and discussed for ZRP model. Models are investigated in the stationary state which is represented by "handy" product measure.

Key words: Time-headways, ZRP, TASEP.

Simulation study for consistency and robustness of Cramrvon Mises type estimators

Jitka Hrabáková^{1,2}, Václav Kůs¹

¹ Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Prague, Czech Republic

² Department of Applied Mathematics, Faculty of Informatics, Czech Technical University, Prague, Czech Republic

Email: jitka.hrabakova@fit.cvut.cz, vaclav.kus@fjfi.cvut.cz

Abstract

The contribution focuses on the minimum distance estimators under two newly introduced modifications of Cramrvon Mises distance. The generalized power form of Cramrvon Mises distance is defined together

with the so called Kolmogorov-Cramr distance which includes both standard Kolmogorov and Cramrvon Mises distances as limiting special cases. We prove the consistency of Kolmogorov-Cramr estimators in the (expected) L1norm by direct technique employing domination relations between statistical distances. In our numerical simulation we illustrate the quality of consistency property for sample sizes of the most practical range from $n = 10$ to $n = 500$. We study dependence of consistency in L1norm on ϵ neighbourhood of the true model and further the robustness of these two newly defined estimators for normal families and contaminated samples. Numerical simulations are used to compare statistical properties of the minimum Kolmogorov-Cramr, generalized Cramrvon Mises, standard Kolmogorov, and Cramrvon Mises distance estimators of the normal family scale parameter. We deal with the corresponding order of consistency and robustness. The resulting graphs are presented and discussed for the cases of the contaminated and uncontaminated pseudo-random samples.

Key words: Minimum distance estimator, Consistency, Robustness, Cramrvon Mises distance.

Evaluation of time reversal transfer experiment with different signal lengths

Adrien Chapuis^{1,2}, Jan Kober²

¹ ENI Val de Loire, Blois, France

² Institute of Thermomechanics, Czech Academy of Sciences, Prague, The Czech Republic

Email: adrien.chapuis@insa-cvl.fr

Abstract

In this presentation, fundamentals of transferring time reversal experiment between geometrically similar objects and the dependance of the signal lengths are discussed. The time reversal experiment consists of two steps: forward propagation, when a source excites the medium and a complex wave field is created, and back propagation, resulting in focusing on the original source location and reconstruction of the source function. here we investigate the procedure of performing the first step on one sample and the second on another. in order to compare the results.

Key words: Time reversal transfer, Pentest, Signal lengths, Snr.

Mapping of general dependencies among data by optimized BP-networks

Milan Chlada, Zdeněk Převorovský

NDT Laboratory, Institute of Thermomechanics, Academy of Sciences of the Czech Republic, Prague

Email: chlada@it.cas.cz, zp@it.cas.cz

Abstract

The problem of adequate selection of parameters representing hidden information in given data can be solved by many more or less sophisticated statistical methods. In general point of view, we can say that the parameter selection consists of identifying a subset of significant attributes and discarding the remaining ones from an initially large set of potentially redundant or correlated features. In the paper, an alternative methodology for finding general non-linear dependencies is shown. It is based on the sensitivity analysis of optimized artificial neural networks (BP-networks) with suppressed over-fitting. The potentiality of this approach for parameter selection is demonstrated by processing of numerically generated model data.

Key words: Parameter selection, Artificial neural networks, Sensitivity analysis.

Random Matrix Revolutions

Ondřej Kollert

Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Prague, Czech Republic

Email: ondra.kollert@gmail.com

Abstract

Random matrix theory is quite new on the field of mathematics. The greater its importance seems to be especially in the modeling of various natural and social phenomena. Such phenomena are related to the random matrices through corresponding eigenvalues. For the description of the set of eigenvalues from the interaction dependence point of view characteristics like spacings between two eigenvalues or number of eigenvalues on a particular interval are used. For this purpose the theory of random processes is applied, particularly results from the counting processes. At the beginning of the presentation the term counting process is therefore introduced as well as some important facts about it are mentioned. Consequently fully new results are presented including for example special representation of variance and covariance of the process. However, the way of interaction between eigenvalues of random matrices

is very complicated. Thus at the end of the presentation the difference between the description of the dependance by means of counting process and the real dependence between eigenvalues is shown.

Key words: Random Matrix, Counting process, Eigenvalues, Interaction dependence.

Stochastic epidemic models

Jitka Kostková

Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Prague, Czech Republic

Email: kostkjit@gmail.com

Abstract

With expansion of traveling and communication, spread of two things has been significantly increased - infection diseases and knowledge. Both of them can be described by epidemic models. If the spread of the disease is known, it is possible to make precautions or choose optimal vaccination strategy to prevent epidemic. First, the deterministic model including birth and death process and control vaccination term is introduced. Next, three possible choices of stochastic versions are compared. These models are solved numerically using Euler method. First passage problem and optimal control problem is formulated. Finally, statistical distribution of first time of declaration of an epidemic is tested and the numerical solution of optimal vaccination strategy is found.

Key words: SIR model, Stochastic differential equations, Optimal vaccination strategy, Hitting time.

Robust Approach to Time Series Prediction

Vojtěch Kovalík

Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Prague, Czech Republic

Email: kovalik.vojtech@gmail.com

Abstract

Modeling and proper and reliable parametric estimating of self-correlated data and time series play a role in a wide range of applications. Time series analysis and statistics provide tools of model selection and parameter estimation. Unfortunately, the existence of contaminated data sets and outliers has an undesirable effect on estimation in these models. Therefore, commonly used efficient methods must be replaced by the robust ones. This work is to research linear models and robust statistical methods of parameter estimation and data prediction. Since EEG data analysis may be used in Alzheimer's disease diagnostics, selected robust methods and the corresponding algorithms are implemented and tested

on EEG data. The resulting characteristics are studied via robust and non-robust variability measures. Possible applications of the results are discussed in the conclusion.

Key words: Linear predictive model, Robust regression, Alzheimer's disease, Data analysis.

Bayesian stochastic simulation oriented inference with approximate Bayesian computation

Jakub Kubiš

Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Prague, Czech Republic

Email: jkubisg.i@gmail.com

Abstract

The approximate Bayesian computation (ABC) methods are used to generate samples from the posterior distribution in cases when the likelihood is not available. After we govern the basics of Markov chains and Bayesian inference, we address the principles of Monte Carlo (MC) methods. More complicated algorithms such as sequential Monte Carlo (SMC) or Markov chain Monte Carlo (MCMC) are described. We use these techniques to address the approximate Bayesian computation methods which are fundamentally dependent on Monte Carlo. At the end we focus on the application of several methods on the linear regression model and the regression parameters are inferred using the ABC.

Key words: Approximate Bayesian inference, Monte Carlo, Sequential Markov Chain.

Goodness of fit based variable ranking in high energy physics data

Jan Kučera

Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Prague, Czech Republic

Email: kucera.ja@gmail.com

Abstract

Renyi information divergences are well-known and widely used in statistical inference due to their robustness and practical feasibility. MC simulation results for the Minimum Renyi Distance (MReD) estimates are presented and the effect of input parameter α to robustness is shown. Statistical hypothesis testing is studied with focus on analysis of the data sample "lepton plus jets" decay channels from RunII measured at the particle accelerator Tevatron in Fermilab. Results of the usage of

Renyi divergence as a possible statistic for comparing similarity or dissimilarity of two datasets is also shown.

Key words: *phi*-divergence, Top quark, Statistical Hypothesis testing.

Active learning for 3D navigation of UAVs

Jakub Malý

Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Prague, Czech Republic

Email: jakub.maly123@seznam.cz

Abstract

The contribution deals with navigation of UAVs. This task is executed in order to obtain as good estimate of unknown parameters as possible. Unknown parameters describe propagation of cloud of pollutants. Algorithm works in discrete time t where there is a state of the system x_t and measurements $y_{1:t}$. Dynamic programming is described with following formula

$$a_t^* = \arg \min_{a_t} \mathbb{E}(\mathcal{L}(x_{t+1:t+h}, a_{t:t+h-1}) | y_{1:t}).$$

Objective is to find optimal action which minimizes expected loss. In this paper expected loss denotes variance in estimate of parameters. Computation of expected loss is impossible to perform analytically due to a posterior distribution $p(x_t | y_{1:t})$, therefore it is profitable to solve this matter using Monte Carlo methods.

Key words: Sequential Monte Carlo, Particle filter, Navigation of UAVs, Active learning.

Approximate Bayesian computation filtering

Filip Marada

Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Prague, Czech Republic

Email: filipmarada@gmail.com

Abstract

Approximate Bayesian computation (ABC) methods are methods designed to facilitate the Bayesian inference of models which are given either by complex likelihood function or for example by stochastic process. The ABC filter extends the particle filtering methodology to general state-space models in which the density of the observation conditional on the state is intractable. Sequential Monte Carlo algorithm

is used to sample and fit from ABC approximation of the target probability density. The contribution presents an application of the ABC methods to a nonlinear state-space model and compares the results with the celebrated extended Kalman filter (EKF) and the unscented Kalman filter (UKF).

Key words: Approximate Bayesian computation, Filtering, State-space model, Sequential Monte Carlo.

Clustering Method of IDM model with relation to the GNU model

Michael Matějů, Pavel Hrabák

Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Prague, Czech Republic

Email: mysakbm@gmail.com

Abstract

In recent years, traffic modelling is gaining interest of mathematicians and physicists. This paper covers brief summary of known facts about ZRP processes, its stationary measures and development of GNU model. However, the main goal will be description of traffic movement in mesoscopic view by clustering of vehicles and development of corresponding quantities between GNU and clustered traffic.

Key words: Clustering, IDM, GNU model, ZRP processes.

Divergence Decision Tree application on a data set, success of classification using SVM

Michaela Sluková, Zuzana Dvořáková

Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Prague, Czech Republic

Email: slukovamisa@gmail.com, dvorakova.zuzi@gmail.com

Abstract

There exist many classification methods in classification of data, both supervised and non-supervised. This work deals with a classification method Divergence Decision Tree (DDT), where we substitute non-supervised basic classification method k-means by supervised SVM, because k-means has several disadvantage for our data set (f.e. first initialization of means of clusters). SVM is very variable method, we can use non-linear version which provides opportunity to choose appropriate kernels for each data set and so on. There is also a chance to set tolerable error ratio. Also different phi-divergence functions are compared in order to find out the dependency of the method on selected divergence. Several version

of the DDT, which differ in type of phi-divergence and input parameters of SVM, are tested on a data set coming from D? experiment from particle accelerator Tevatron Collider.

Key words: Data clustering, Decision tree, Phidivergence, SVM, Supervised learning.

Parallel evaluation of MVA methods without ROOT dependency

Vladimír Španihel

Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague

Email: vlada.spana@gmail.com

Abstract

This paper is aimed at MVA methods evaluation on heterogeneous cluster or multi-core computer. At the beginning, there was a need for a program simultaneously performing number of learned MVA classifiers on given dataset. Each learning method of CERNs ROOT system generates two kinds of weight files. The most frequently used kind has .xml extension and may be passed as a parameter to `TMVA::Reader` while booking classification method. The second kind of weight files has extension in form `.class.C`. These files contain definitions of C++ classes (learned classifiers) which can be used for evaluation of MVA methods in arbitrary, ROOT independent, C++ code. For task distribution, OpenMpi was selected. While developing the program, many issues appeared. The main issue was to distribute learned classifiers to processes without memory wasting. Taking advantage of sending instances over MPI interface requires serialization and deserialization of objects. An easy way to send objects over MPI is adding dependency on BOOST framework but the program should be deployed with BOOST libraries. The temporary solution without adding any dependencies was very memory consuming because sending objects over MPI could not be used. So each process created vector composed of all instances of learned classifiers; however, it used just one of them. This memory issue was solved by using a function template and vector of function pointers.

Key words: TMVA, ROOT, classifier, MPI.

Modification of gaussian mixture models for data classification in high energy physics

Michal Štěpánek

Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Prague, Czech Republic

Email: michal.stepanek@fjfi.cvut.com

Abstract

In high energy physics, we deal with demanding task of signal separation from background. Since the distribution function of signal is rarely noticeably distinct from that of background, advanced classification methods have to be applied. The Model Based Clustering method including the Expectation-Maximization algorithm will be discussed together with its modifications such as weighting, missing data processing and genetic optimization. According to extensive variability of the method, data pre-processing plays a significant role for the subsequent combination of final discriminants in order to improve signal separation efficiency. Effects of variable gaussianisation, decorrelation and the number of components of distribution mixtures on the quality of data classification will be mentioned. Moreover, the results of the top quark separation from the Tevatron collider will be compared with those of standard multivariate techniques in high energy physics.

Key words: Model Based Clustering, Variable transformation, Top quark analysis.

Determination of phase transitions in two-dimensional model TASEP

Lenka Štěpánková

Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Prague, Czech Republic

Email: lencasc@gmail.com

Abstract

We will present a two-dimensional generalization of a model TASEP. The totally asymmetric simple exclusion process is a mathematical model used for traffic modelling and it is well-known as a one-dimensional model. Nowadays, it is well studied by analytical and numerical methods. One of the studied attributes of this stochastic model is phase diagram, graph that shows the relation between parameters and bulk density of the model. The study of phase transitions will be introduced. We focus on the determination of the phase transitions by observing and analyzing of density profiles for various

settings of parameters. The data are compared with analytics predictions for one dimensional version of the model. Simulations for different types of updates will be introduced. In this work are used modified updates for two-dimensional lattice. Used updates are: fully parallel, random, shuffled and ordered sequential.

Key words: TASEP, Phase diagram, Density profiles.

Model for prediction of power consumption of a hybrid vehicle

Vít Škvára

Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Prague, Czech Republic

Email: skvarvit@fjfi.cvut.cz

Abstract

Development of hybrid vehicles is the focus point of transport industry in present. While there are more sources of power onboard a hybrid vehicle with different characteristics, it is advantageous to have some sort of control over distribution of their usage. This should be done with regard to prior knowledge of route and physical model of movement. In the talk, stochastic model of fuel consumption will be presented, together with an insight into map data processing, based on data from operation of hybrid bus.

Key words: Hybrid vehicle, Vincenty's algorithm, Exponential forgetting, Kalman filtering.

Annealing-based car-following model

Jiří Šleis, Milan Krbálek

Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Prague, Czech Republic

Email: sleisjir@fjfi.cvut.cz

Abstract

The intention of this article is to examine whether the arrangement of vehicles in the vicinity of an intersection is a consequence of traffic rules, complicated evaluation-procedures, and sophisticated decision-making procedures inside a drivers brain or, on contrary, it is a consequence of general stochastic nature of queuing systems. For solving this dilemma a stochastic alternative based on principles of the so-called simulated annealing for both above-discussed models is created. Thus, an unimodal scheme simulating a time-evolution of vehicular ensembles without any division into modes and without a concept of the safe values for some quantities will be introduced.

Key words: Time clearances, Signalized inter- section, Car-following model, Simulated annealing.

Option pricing

Viktor Šohájek

Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Prague, Czech Republic

Email: viki_sohajek@yahoo.com

Abstract

Option is a financial instrument, that gives the owner the right to buy or sell a certain amount of asset for a fixed price in the future. The value of underlying asset is changing stochastically through time. We will show, how stochastic Ito calculus can be used to derive a price, for which the option should be traded and what conditions should be taken into account.

Key words: Option, Financial mathematics, Black-Schols approach.

Acoustic emission testing

Aurélie Teyssier^{1,2}, Zuzana Dvořáková^{2,3}

¹ ENI Val de Loire, Blois, France

² Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Prague, Czech Republic

³ Institute of Thermomechanics, Czech Academy of Sciences, Prague, The Czech Republic

Email: aurelie.teyssier@insa-cvl.fr, dvorakova.zuzi@gmail.com

Abstract

Acoustic Emission (AE) is a non-destructive method in that aims to do a wide group of analysis to examine the behaviour and evaluate the properties of a material deforming under stress without causing any damage. The defectoscopy principle is used to characterize and identify the damage developments. Reliable identification and classification of pre-localized AE sources is one of the most important and also most difficult problems in AE monitoring. In order to classify different AE sources, suitable signal parameters and a classification method are necessary. In this work, the signals are distinguished by means of newly suggested parameters obtained directly from signals and from signal spectra. The parameters are calculated from a continuous version of formerly suggested successful parameters. As the classification method k-means and the Gaussian Mixture Model (GMM) are used. All suggested parameters along with the classification method are tested on the laboratory measured different types of AE signals.

Key words: Acoustic Emission, Classification method.

Application of mixed logistic model in small area estimation

Aleš Tygl

Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Prague, Czech Republic

Email: ales.tygl@gmail.com

Abstract

Small area estimation (SAE) is in need of reliable methods for statistical inference. Direct estimation from particular small areas might not be feasible and models that borrow strength from other areas are used. For binary data, a logistic model is one of the most used in practice. In this work, the logistic mixed model is introduced. Then, parameter estimation techniques based on Laplace approximation of likelihood are outlined. The pseudo empirical best linear unbiased predictors (pseudo-EBLUP) of area parameters are derived. The random area effects logistic model is compared to a logistic model with area effects treated as fixed in a set of Monte Carlo simulations. Performance of both models is discussed. Finally, the results are compared to the corresponding results of a Fay-Herriot model.

Key words: Generalized linear mixed models, Small area estimation, Logistic regression.

Some notes on SDE's in infinite dimension

Petr Veverka

Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Prague, Czech Republic

Email: petr.veverka@fjfi.cvut.cz

Abstract

In the talk, the concept of probability measures in infinite dimensional real separable Hilbert space H will be introduced in short and the differences compared to the finite dimension case will be discussed. A construction of a Wiener process in H will be mentioned together with a SDE driven by such a process. Finally, the connection to stochastic PDE's will be provided.

Key words: Gaussian measures in Hilbert spaces, Wiener process, SDE.

Stochastic control of tumor growth

Hynek Walner

Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Prague, Czech Republic

Email: hynekw@gmail.com

Abstract

Pontryagin's maximum principle, with its origin in theoretical mechanics, is well known tool for optimizing controlled differential equations. One can find stochastic analogy of this principle when solving stochastic control problems. We will demonstrate several possible forms of stochastic maximum principle by applying to controlled stochastic differential equation describing tumor growth and present limitations of this approach.

Key words: Stochastic control, Maximum principle, BSDE, Cancer modelling.

