http://maxent2010.inrialpes.fr/

30th International Workshop on Bayesian Inference and Maximun Entropy Methods in Science and Engineering

Chamonix, France, July 4-9, 2010

Program & booklet of abstracts

Thomas Bayes

Pierre-Simon Laplace

For 30 years the MaxEnt workshops have explored the use of Bayesian and Maximum Entropy methods in scientific and engineering applications. All aspects of probabilistic inference, such as techniques, applications and foundations, are of interest. The workshop includes a one-day tutorial session, state-of-the-art lectures, invited papers, contributed papers and poster presentations.

An excursion should be getting to the top of Aiguille du midi. The excursion will be on wednesday afternoon.

The banquet will be on thursday evening (July 8, 2010) and will take place at the "Grand-Hotel du Montenvers" which offers a panoramic view on the biggest French glacier : "mer de glace".

——oOo——

Program committee

General Chair: Ali Mohammad-Djafari, LSS, UMR 8506, CNRS-Supelec-Univ Paris Sud

Other chairs:
Pierre Bessière, CNRS (Technical Program/Local arrangements)
Jean-François Bercher, ESIEE/Université Paris-Est (Technical Program/Technical arrangements)

Secretary/Organization/Technical arrangements/Local arrangements
Danièle Hertzog, INRIA
Doriano-Boris Pougaza, Université Paris-Sud 11

Sponsors: we gratefully acknowledge the following sponsors, whose support have helped to the organization and success of the workshop.

- The Edwin T. Jaynes International Center for Bayesian Methods and Maximum Entropy,
- INRIA
- CNRS
- DIGITEO
- Université Paris-Sud
- ProBAYES, Inc
- ESIEE-CCIP Université Paris-Est
- INPG

Program

July 04, 2010

09:15:00-10:15:00 : **Tutorial 1 - Information Physics: The New Frontier**
by K. Knuth Page *11*

10:45:00-11:45:00 : **Tutorial 2 - Entropic Inference as the Foundation of Physics**
by A. Caticha

13:30:00-14:30:00 : **Tutorial 3 - High Up from the Ground: A Panoramic View of Inference and MaxEnt** by C. Rodriguez

14:30:00-15:30:00 : **Tutorial 4 - Foundations of Computational Inference**
by J. Skilling Page *12*

16:00:00-17:00:00 : **Tutorial 5 - Bayesian programming: principle, methodology, tools and examples** by P. Bessière

July 05, 2010

09:00:00-11:45:00 : **General methods and models**

Model Selection for Mr Studies of Stroke,
John Lee, Larry Bretthorst, Colin Derdeyn, Joshua Shimony Page *13*

Algebraic bayesian quantum theory,
Ryszard Pawel Kostecki Page *14*

Maximum Entropy Modeling For Automatic Content Scoring of Short Free-Text Responses,
Jana Sukkarieh Page *15*

Darwinian Model Building,
Do Kester, Romke Bontekoe Page *16*

13:30:00-16:45:00 : **Bayesian models of sensory motor systems**

Bayesian inference with spikes. Implication for the neural code, sensory processing and working memory,
Sophie Deneve Page *17*

Bayesian Action-Perception loop modeling: Application to trajectory generation and recognition using internal motor simulation,
Estelle Gilet, Julien Diard, Richard Palluel-Germain, Pierre Bessière Page *18*

Bayesian Modeling of an Human MMORPG Player,
Gabriel Synnaeve, Pierre Bessière Page *19*

Bayesian Visual Odometry,
Julian Center, Kevin Knuth Page *20*

Maximum entropy perception-action space: a Bayesian model of eye movement selection,
Francis Colas, Pierre Bessière, Benoît Girard Page *21*

16:45:00-19:00:00 : **Poster 1 – Foundations**

Adaptive Strategies in the Iterated Exchange Problem,
Arthur Baraov Page *22*

Finding Proper Non-informative Priors for Regression Coefficients,
Noel van Erp, Pieter van Gelder Page *23*

The Ball Is Round,
Do Kester Page *24*

Significance Indexes: Fisher vs. Bayes,
Marcio Diniz, Carlos Pereira, Adriano Polpo, Julio Stern Page *25*

Entropy Estimation For M/M/1 Queueing Systems,
Philippe Regnault Page *26*

The Full Bayesian Significance Test for Symmetry in Contingency Tables,
Gustavo Bernardo, Marcelo Lauretto, Julio Stern Page *27*

Baobabs under stress, a case of estimated pdf transfer,
Lucien Lemmens, Sebastiaan De Smedt, Roeland Samson Page *28*

The Order-Theoretic Origin of Special Relativity,
Kevin Knuth, Newshaw Bahrenyi, Philip Goyal Page *29*

Non-relativistic Gravity in Entropic Quantum Dynamics,
David Johnson, Ariel Caticha Page *30*

Mixture of skewed alpha-stable distributions,
Sayed Reza Hossaini Shojaei, Vahid Nassiri, Gholam Reza Mohammadian, Adel Mohammadpour Page *31*

Bayesian Inference for Skewed Stable Distributions,
Mona Shokripour, Vahid Nassiri, Adel Mohammadpour Page *32*

July 06, 2010

09:00:00-11:45:00 : **Algorithms for Maximum Entropy and Bayesian computation**

Advice on "Computational Methods for Bayesian Model Choice",
John Skilling Page *33*

Self-Adaptive Semantic Crossover for Genetic Programming,
Julio Stern, Rafael Inhasz Page *34*

Entropy-Based Search Algorithm for Experimental Design,
N. K. Malakar, K. H. Knuth Page *35*

Bayesian inference for MAPK signaling networks,
George Popescu, Sorina Popescu, Vasile Buzuloiu Page *36*

Nested Sampling with Constrained Hamiltonian Monte Carlo,
Michael Betancourt Page *37*

13:30:00-16:45:00 : **Information theory and statistical physics**

Information theory of quantum systems with some hydrogenic applications,
Jesus Sanchez-Dehesa, D. Manzano, P. Sánchez-Moreno, R.J. Yáñez Page *38*

On moments-based Heisenberg inequalities,
Steeve Zozor, Mariela Portesi, Pablo Sanchez-Moreno, Jesus Dehesa Page *39*

Rigorous bounds for Renyi entropies of spherically symmetric potentials,
Pablo Sanchez-Moreno, Steeve Zozor, Jesus Dehesa Page *40*

Entropic time,
Ariel Caticha Page *41*

Some topics on q-gaussians, escort distributions and Fisher information,
Jean-François Bercher Page *42*

16:45:00-19:00:00 : **Poster 2 – Applications**

Non-Extensive Radiobiology,
Oscar Sotolongo-Grau, Daniel Rodriguez-Perez, Carlos Antoranz, Oscar Sotolongo-Costa Page *43*

Online Robot Dead Reckoning Localization Using Maximum Relative Entropy Optimization With Model Constraints,
Renaldas Urniezius Page *44*

Nonlinear Source Separation: A Maximum Likelihood Approach for Quadratic Mixtures,
Chahinez Chaouchi, Yannick Deville, Shahram Hosseini Page *45*

Superresolution of compact neutron spectrometers, with applications to fusion diagnostics,
Marcel Reginatto, Andreas Zimbal Page *46*

Bayesian approach to siRNA-mediated gene silencing phenotypes quantification and off-target level estimation,
Yannis Kalaidzidis, Jochen Rink, Claudio Collinet, Akhila Chandrashaker, Martin Stötter, Eugenio Fava, Marino Zerial Page *47*

On combined estimation of expected backgrounds with statistical and systematic uncertainties in planned experiment,
Sergey Bityukov, Nikolai Krasnikov, Vera Smirnova Page *48*

Towards a Bayesian seismotectonic zoning for use in Probabilistic Seismic Hazard Assessment (PSHA),
Boris Le Goff, Delphine Fitzenz, Céline Beauval Page *49*

Application Of A Bayesian Inference Method To Reconstruct Short-Range Atmospheric Dispersion Events,
Inanc Senocak Page *50*

GNU MCSim: Bayesian Statistical Inference for System Biology Models,
Rudy Gunawan, Frederic Bois Page *51*

A Bayesian approach to Fourier Synthesis inverse problem with application in SAR imaging,
Sha Zhu, Ali Mohammad-Djafari Page *52*

Optimal Path and Sensor Selection for Target Tracking in Wireless Sensor Networks,
Majdi Mansouri, Hichem Snoussi, Moez Esseghir, Cedric Richard Page *53*

Signals and Images Foreground/Background Joint Estimation and Separation,
Boujemaa Ait-El-Fquih, Ali Mohammad-Djafari Page *54*

Maxent For Climate Impact Prediction For Mangrove Habitat In The Bay Of Bengal Rim,
Md. Saiful Islam Page *55*

Total Reliability Management by Bayesian Approach,
Ding Huang Page *56*

July 07, 2010

09:00:00-11:45:00 : **Bayesian algorithms for multi-channel image analysis**

Three-dimensional visualization and identification of objects in photon starved scenes using statistical estimation,
Bahram Javidi, Dipak Dey, M. Cho, M. Jung Page *57*

Bayesian fusion of hyperspectral astronomical images,
André Jalobeanu, Matthieu Petremand, Christophe Collet Page *58*

Unmixing hyperspectral images using Markov random fields,
Olivier Eches, Nicolas Dobigeon, Jean-Yves Tourneret, . Page *59*

An Investigation of Likelihoods and Priors for Bayesian Endmember Estimation,
Alina Zare, Paul Gader Page *60*

Multichannel SAR Image Classification by Finite Mixtures, Copula Theory and Markov Random Fields,
Vladimir A. Krylov, Gabriele Moser, Sebastiano B. Serpico, Josiane Zerubia Page *61*

July 08, 2010

09:00:00-11:45:00 : **Non parametric methods and experimental design**

The Maximum Entropy Formalism And A Generalization Of The Symmetric Dirichlet Distribution,
Dawn Holmes Page *62*

Maximum Entropies Copulas,
Doriano-Boris Pougaza, Ali Mohammad-Djafari Page *63*

Digital Particle Image Velocimetry Using Splines in Tension,
Udo von Toussaint, Silvio Gori Page *64*

Confidence Distributions In Statistical Inference,
Sergey Bityukov, Nikolai Krasnikov, Saralees Nadaraja, Vera Smirnova Page *65*

Bayesian Validation of CAIX Immunohistochemical Assays of CAIX,
Rudy Gunawan, Giuseppe Tenti, Siv Sivaloganathan Page *66*

13:30:00-16:45:00 : **Bayesian and maximum entropy inference in action: applications**

The use of bayesian programming in industrial applications,
Emmanuel Mazer Page *67*

Parameter Estimation as a Problem in Statistical Thermodynamics,
Keith Earle, David Schneider Page *68*

Reinforcement Learning by Relative Entropy Policy Search,
Jan Peters, Katharina Muelling, Yasemin Altun Page *69*

Bayesian Techniques for Parameter Inference and Knowledge Fusion in Prognostics and Health Management A Case Study,
Masoud Rabiei, Mohammad Modarres, Ali Mohammad-Djafari Page *70*

ID image characterization by entropic biometric decomposition,
Smoaca Andreea, Coltuc Daniela, Fournel Thierry Page *71*

16:45:00-19:00:00 : **Poster 3 – Methods and Algorithms**

The Posterior distribution of the Likelihood Ratio as a measure of evidence,
Isabelle Smith, André Ferrari Page *72*

A Method for Evaluating Tuning Functions of Single Neurons based on Mutual Information Maximization,
Lukas Brostek, Thomas Eggert, Seiji Ono, Michael J. Mustari, Ulrich Büttner, Stefan Glasauer Page *73*

Gaussian Process Segmentation of Co-Moving Animals,
Steven Reece, Richard Mann, Iead Rezek, Stephen Roberts Page *74*

Speed-gradient principle for description of transient dynamics in systems obeying maximum entropy principle,
Alexander Fradkov, Anton Krivtsov Page *75*

Generalized Maximum Entropy Principle, Superstatistics And Problem Of Networks Classification,
Bahruz Gadjiev Page *76*

Comparative Analysis Of Collaboration Networks,
Tatiana Progulova, Bahruz Gadjiev Page *77*

Bayesian experiments for numerical solution of differential equations and integrals,
Fábio Mendes Page *78*

A global optimization heuristic for Nonlinear UAV Task Assignment Problem - The Cross-Entropy method,
NGUYEN Duc Manh, LE THI Hoai An, PHAM Dinh Tao Page *79*

A Maximum Entropy Approach to Semi-supervised Learning,
Ayse Erkan, Yasemin Altun Page *80*

Empirical Bayes Estimation Using Ranked Set Sampling,
Mina Aminghafari, Adel Mohammadpour, Hadi Zare Page *81*

Bayesian Inference Using Maximum Entropy: Analysis And Application,
Fadi Dornaika, Fadi Chakik Page *82*

Statistical Learning Approach using Maximum Entropy: Image Texture Classification by Energy Function,
Fadi Chakik, Fadi Dornaika Page *83*

July 09, 2010

09:00:00-10:15:00 : **Foundations**

Logarithm ubiquity in statistics, information, acoustics, ...,
M. Maurin Page *84*

Modeling the World by Minimizing Relative Entropy,
Antoine van de Ven, Ben A.M. Schouten Page *85*

10:45:00-11:45:00 : **Markovian Processes**

Spectral Design in Markov Random Fields,
Jiao Wang, Jean-Baptiste Thibault, Zhou Yu, Ken Sauer, Charles Bouman Page *86*

Determination and Estimation of Generalized Entropy Rates for Countable Markov Chains,
G Ciuperca, valérie Girardin, L Lhote Page *87*

13:30:00-15:15:00 : **Information geometry and information theory**

Geometric and Topological Invariants of the Hypothesis Space,
Carlos Rodriguez Page *88*

Geometry of Covariance Matrices and Computation of Median,
Le Yang, Marc Arnaudon, Frédéric Barbaresco Page *89*

2010-07-04 - Session *Tutorial 1 - Information Physics: The New Frontier* [09:15:00-10:15:00]

Information Physics: The New Frontier

Kevin Knuth

At this point in time, two major areas of physics, statistical mechanics and quantum mechanics, rest on the foundations of probability and entropy. The last century saw several significant fundamental advances in our understanding of the process of inference, which make it clear that these are inferential theories. That is, rather than being a description of the behavior of the universe, these theories describe how observers can make optimal predictions about the universe. In such a picture, information plays a critical role. What is more is that strange little clues, such as the fact that black holes have entropy, continue to suggest that information is fundamental to physics in general.

In the last decade, our fundamental understanding of probability theory has led to a Bayesian revolution. In addition, we have come to recognize that the foundations go far deeper and that Cox's approach of generalizing a Boolean algebra to a probability calculus is the first specific example of the more fundamental idea of assigning valuations to partially-ordered sets. By considering this as a natural way to introduce quantification to the more fundamental notion of ordering, one obtains an entirely new way of deriving physical laws. The first result was a new theoretical foundation of measure theory, and specifically, probability theory. This new foundation both encompasses and extends the foundations proposed by Cox and Kolmogorov. Unexpectedly, in the last year, we have learned that one can derive the complex Feynman formulation of quantum mechanics directly from symmetries of observations one can make of quantum systems. At this meeting, for the first time, I will present a novel derivation of special relativity where space and time are defined to make the quantification of partially-ordered sets of events look simple.

I will introduce this new way of thinking by explaining how one can quantify partially-ordered sets and use symmetries to derive physical laws. The implication is that physical law does not reflect the order in the universe, instead it is derived from the order imposed by our description of the universe. We have now demonstrated this by deriving a surprising amount of old physics. New physics now awaits as we enter the new frontier of Information Physics.

2010-07-04 - Session *Tutorial 4 - Foundations of Computational Inference* [14:30:00-15:30:00]

Foundations of Computational Inference

John Skilling

Traditionally, computational inference has been an art-form comprising a mix of good and sometimes less-good ideas. It seems to the author that the subject is now at the cusp of becoming a principled discipline, in which various methods can be placed in the context of the solid foundations of probability and information. This talk will start at the beginning, by asking how Bayesian inference can best be accomplished. The tutorial will try to answer this question with generality and simplicity.

Model Selection For Mr Studies Of Stroke

John Lee, Larry Bretthorst, Colin Derdeyn, Joshua Shimony

For the clinical management of stroke, MR studies of cerebral perfusion have become increasing important and controversial. The controversy arises from the repeated failure of MR estimates of the ischemic penumbra to predict outcomes in several large, randomized controlled clinical trials of thrombolytic agents. While there is confounding in several study designs, the accuracy of MR measurements is clearly in question. Key measurements of cerebral perfusion involve $C(t) = FR(t) \otimes C\_A(t)$, for time-dependent tracer concentrations $C(t)$, flow F and residue function $R(t)$ convolved with an arterial input function $C\_A(t)$.

Historically and in clinical trials aforementioned, a single global input function is constructed for assessment of the entire brain [1]. Candidate input functions from regions of dynamic susceptibility near the origins of cerebral circulation are selected, averaged and often exchanged for regression results with a parametric model. Multiple local input functions have also been analyzed, uniquely estimated for each voxel, using a parametric model and Markov-chain Monte Carlo [2,3]. Tissue circulation is described by the gamma distribution with time t, time-origins $t\_0^0, t\_1^0, ..., t\_m^0$, adjustable parameters α and β and normalizing coefficients c_m: $C\_A(t) = \sum\_mc\_mG\_m(t)$, $G\_m(t) \equiv \beta^{-1-\alpha}\Gamma^{-1}(\alpha+1)(t-t\_m^0)^\alpha \exp[-(t-t\_m^0)/\beta]$. The power-law $(t-t\_m^0)^\alpha$ describes a scale-invariant dynamics that is damped by the first-order kinetics of the exponential factor. For global input functions, the residue has been modeled with an ensemble of transit times: $R(t) = \int\_t^\infty G(\alpha, \beta, u)du$, $u \equiv t - t\_0$. Contrastingly, for global input functions, the residue function has been described using a single mean intravoxel transit time, τ: $R(t) = e^{-t/\tau}$, $\tau \neq t - t\_0$.

This paper investigates the combinations of models and their model probabilities for MR perfusion data in patients with varying severity of cerebral hemodynamic impairment. Acute ischemic stroke has unique features compared to chronically evolving vaso-occlusive disease. Severe hemodynamic impairment requires increased model complexity.

References: [1] K. Mouridson et al. Neuroimage **33** 570-579 (2006). [2] J. C. Shimony et al. Max. Ent. Bay. Meth. **803**, 535542 (2005). [3] J. Lee et al. Magn. Res. Med. **In Press** (2010).

Key Words: model selection, cerebral perfusion, stroke

2010-07-05 - Session *General methods and models* [09:00:00-11:45:00]

Algebraic bayesian quantum theory

Ryszard Pawel Kostecki

We present a new approach to the foundations of quantum theory that joins maximum entropy bayesian approach to the probability and information theory with the formalism of information geometry and an algebraic approach to quantum theory. As opposed to other approaches, it can deal with conceptual and mathematical problems of quantum theory without any appeal to Hilbert space framework or frequentist or subjective interpretation.

2010-07-05 - Session *General methods and models* [09:00:00-11:45:00]

Maximum Entropy Modeling For Automatic Content Scoring of Short Free-Text Responses

Jana Sukkarieh

The education community is moving towards constructed or free-text responses instead of multiple-choice ones. Also, it is moving towards widespread computer-based assessments. At the same time, progress in natural language processing and knowledge representation has made it possible to consider free-text responses without having to fully understand the text. c-rater is a technology at Educational Testing Service for automatic content scoring of short, free-text responses. This paper intends to describe how Maximum Entropy Modeling (MaxEnt) is used to help solve the c-raters task.

Darwinian Model Building

Do Kester, Romke Bontekoe

Abstract

We present a way to generate heuristic mathematical models based on the Darwinian principles of variation and selection in a pool of individuals over many generations. Each individual has a genotype (the hereditary properties) and a phenotype (the expression of these properties in the environment). Variation is achieved by cross-over and mutation operations on the genotype which consists in the present case of a single chromosome. The genotypes 'live' in the environment of the data. Nested Sampling is used to optimize the free parameters of the models given the data, thus giving rise to the phenotypes. Selection is based on the phenotypes.

The evidences which naturally follow from the Nested Sampling Algorithm are used in a second level of Nested Sampling to find increasingly better models.

The data in this paper originate from the Leiden Cytology and Pathology Laboratory (LCPL), which screens pap smears for cervical cancer. We have data for 1750 women who on average underwent 5 tests each. The data on individual women are treated a small time series. We will try to estimate the value of the prime cancer indicator from previous tests of the same woman.

Key Words: Evolutionary Programming, Model Selection, Nested Sampling, Time Series

Bayesian inference with spikes. Implication for the neural code, sensory processing and working memory

Sophie Deneve

The Bayesian framework provides normative models that proved powerful to describe human behavior. However, the neural basis of probabilistic computations remain largely unknown. We will show that networks of integrate and fire neurons are ideally suited to infer the probability of sensory variables on-line. This has important implications for the dynamics of single neurons, single synapses, and population of neurons.

Bayesian Action-Perception loop modeling: Application to trajectory generation and recognition using internal motor simulation

Estelle Gilet, Julien Diard, Richard Palluel-Germain, Pierre Bessière

This paper is about modeling perception-action loops and, more precisely, the study of the influence of motor knowledge during perception tasks. We use the Bayesian Action-Perception (BAP) model, which deals with the sensorimotor loop involved in reading and writing cursive isolated letters. The Bayesian model includes a feedback loop from the planned trajectory back to the internal representation of letters, that implements an internal simulation of movement.

Using this probabilistic model, we simulate letter recognition, both with and without internal motor simulation. We demonstrate that, when stimuli are easy (complete letters), letter recognition with and without internal simulation yield the same performance. On the other hand, in difficult cases, like truncated letters, recognition errors can be corrected by internal simulation, which introduces motor knowledge into the perception task. Comparing the observed simulated performance yields an experimental prediction, which we detail.

Bayesian Modeling of an Human MMORPG Player

Gabriel Synnaeve, Pierre Bessière

This paper describes an application of Bayesian programming to the control of an autonomous avatar in a multiplayer role-playing game (the example is based on World of Warcraft). We model a particular task, which consist of choosing what to do and on which target in a situation where allies and foes are present. We explain the model in Bayesian programming and how we could learn the conditional probabilities from data gathered during human-played sessions. This model deals only with a sub-task of a global AI for autonomous non-playing characters. The problem that we try to solve with the presented model is: how do we choose which skill to use and on which target in a player versus environment battle? Possible targets are all our allies and foes. Possible skills are all that we know. From the input values of perception variables and the underlying conditional probabilities (that can be learned) defining our model, we will just try and get a distribution over targets and skills and pick the most probable combination that is yet possible to achieve (enough energy/mana, no cooldown). We achieved this computations through Bayesian programming with joint distributions adequate with our model. We also present results from it being applied to a simulated situation of battle with 12 skills and 6 targets. Finally, we discuss the computational complexity of such a model and how we could speed calculations up (subquestions) in cases of large battles. Indeed, our model has yet to be applied in real time along with a video game running, but we are confident in its scalability.

Bayesian Visual Odometry

Julian Center, Kevin Knuth

Visual odometry refers to tracking the motion of a body using an onboard vision system. In practice one often combines the complementary accuracy characteristics of vision and inertial measurement units. This is the approach employed by the Mars Exploration Rovers (MER), Spirit and Opportunity. While the MER navigation system is based on Bayesian methods, a number of simplifying approximations were required to deal with onboard computer limitations. Furthermore, these computational limitations severely constrain the allowable motion of the rover.

Recent advances in computer technology make it feasible to implement fully Bayesian approaches to visual odometry. Here we present an approach to odometry that combines dense stereo vision, dense optical flow, and inertial measurements. As with all true Bayesian methods, our algorithm determines error bars for all estimates. Furthermore, since this approach efficiently combines data from multiple sources, we can afford to use lower-quality sensors, such as Micro-Electro Mechanical Systems (MEMS) inertial components, which are more economical, weigh less, and consume less power than conventional inertial components.

Maximum entropy perception-action space: a Bayesian model of eye movement selection

Francis Colas, Pierre Bessière, Benoît Girard

In this article, we investigate the issue of the selection of eye movements in an eye-free Multiple Object Tracking task. We propose a Bayesian model of retinotopic maps with a complex logarithmic mapping. This model is structured in two parts: a representation of the visual scene, and a decision model based on the representation. We compare different decision models based on different features of the representation and we show that taking into account uncertainty helps predict the eye movements of subjects recorded in a psychophysics experiment. Finally, based on experimental data, we postulate that the complex logarithmic mapping has a functional relevance, as the density of objects in this space in more uniform than expected. This may indicate that the representation space and control strategies are such as the object density is of maximum entropy.

Adaptive Strategies in the Iterated Exchange Problem

Arthur Baraov

We argue for clear separation of the *exchange problem* from the *exchange paradox* to avoid confusion about the subject matter of these two distinct problems. The essence of the exchange paradox is to find the flaw behind the reasoning that leads to conclusion that expected gain of switching envelopes is always positive, irrespective of the amount of money observed in the chosen envelope. And the objective in the exchange problem is to find payoff-increasing strategies in behalf of the player.

Another important point is that the exchange problem in its current format belongs to the domain of optimal decision making $-$ it doesn't make any sense as a game of competition between the host and the player. It doesn't make much sense as a game of competition between the players either. But it takes just a tiny modification in the statement of the problem to breathe new life into it and make it practicable and meaningful game of competition, potent enough to serve as a simple yet nontrivial model for the financial market games.

Restated Exchange Problem: In a gambling house, a player is offered two sealed envelopes for sale, each holding a bank cheque; one cheque is twice the amount of the other. The player is free to open either envelope and then, after observing its content, make a decision: which envelope to buy. *The cheque amount in each envelope is the price for the other one*. The goal in this zero-sum game is the same for both adversaries $-$ the player and the house $-$ which is to maximize their net gain over indefinitely repeated trials of the game.

In this paper, we discuss adaptive strategies for the restated exchange problem in behalf of both the house and the player. The rules of the game, i.e. the boundary conditions that render the problem well-posed mathematically, are stated precisely. A few broad classes of boundary conditions are identified and treated separately.

In the continuous sample space, $x \in [0, \infty]$, we show that the objective prior pdf is of the form $\varphi(\alpha + \beta \log\_2x)/x^2$, where α and β $-$ arbitrary real numbers, and φ $-$ any nonnegative periodic function: $\varphi(z) = \varphi(z + \beta)$ and $\varphi(z) \geq 0\ \forall z \in [-\infty, \infty]$. In the discrete sample space, $x\_n + 1 = 2x\_n$, $n = 0, \pm1, \pm2, \ldots$ where x_0 is a positive seed point, the objective prior pmf is found to be $\propto 1/x\_n$.

In the case of well-behaved and stationary prior probability distribution, we offer a winning strategy in behalf of the player. With no restrictions imposed on the strategies available to the house, we build a winning strategy for the house that surprisingly does not require the unrealistic assumption of unlimited funds.

The results of computer-simulated runs of the game, demonstrating the efficiency of suggested algorithms, are presented.

Key Words: Exchange problem, Objective prior, Adaptive strategy.

Finding Proper Non-informative Priors for Regression Coefficients

Noel van Erp, Pieter van Gelder

Abstract It is a well-known fact that in problems of Bayesian model selection improper priors may lead to biased conclusions. In this paper we first give a short introduction to the procedure of Bayesian model selection. We then demonstrate for a simple model selection problem, involving two regression models, how improper uniform priors for the regression coefficients will exclude automatically the model with the most regression coefficients. Having established the problematic nature of improper priors for Bayesian model selection problems we proceed to derive an optimal parsimoneous proper uniform prior for univariate regression models. This simple univariate prior is then generalized to its multivariate equivalent. Having derived a suitable parsimonious proper noninformative uniform prior for the multivariate case, we proceed to take a closer look at the evidence values which result from using this prior. Finally, we will discuss the connection between Bayesian model comparison and classical maximum likelihood model selection.

Keywords Proper non-informative priors, Regression coefficients

The Ball Is Round

Do Kester

Abstract

The qualification matches for the European Championship of 2008 and the World Championship of 2010 for national football teams are analysed using a number of different models. Friendly matches between national teams in the same period were added to provide connectivity between the qualification pools. The model for which the Nested Sampling algorithm calculates the highest evidence, is used to predict the outcome of the championship.

It is also investigated if and how the draw into groups at the tournament influences the championship.

As the World Championship will be on-going at the time this poster is presented, some of the predictive powers of the models can be judged immediately.

Key Words: Model selection, Nested Sampling

2010-07-05 - Poster Session – *Foundations* [16:45:00-19:00:00]

Significance Indexes: Fisher vs. Bayes

Marcio Diniz, Carlos Pereira, Adriano Polpo, Julio Stern

This paper is about the comparison of two significance indexes: **p**-value and **e**-value **p** for probability and **e** for evidence. The main objective of these indexes is to measure the consistence of the data with the sharp null hypothesis. We restrict the scope of the paper to sharp null hypotheses, **H**, against composite alternative hypotheses, **A**, as in most basic standard statistics texts.

Entropy Estimation For M/M/1 Queueing Systems

Philippe Regnault

The aim of this paper is to estimate the entropy of M/M/1 queueing systems. The marginal entropy and entropy rate of a stochastic process are well-known to measure its uncertainty. When only observations of the process are available, the need to estimate them arises. Regnault [6,7] provides estimators with good asymptotic properties of the entropy of continuous-time ergodic Markov processes.

We specialize the results of Regnault [6,7] in the context of birth-death processes used in continuous-time M/M/1 queueing systems. Links with maximum entropy arguments are used to characterize the asymptotic behavior of the estimators.

The Full Bayesian Significance Test for Symmetry in Contingency Tables

Gustavo Bernardo, Marcelo Lauretto, Julio Stern

The problem of symmetry hypothesis is fundamental in statistics analysis, where the researcher must assess the existence of a certain symmetry condition. In several applications, the state of compliance, normality or health is characterized by the existence of symmetries. In these situations, the lack of symmetry is an indicator of non-compliance, abnormality or illness. The early detection of the lack of symmetry can frequently allow the repair, maintenance or simplified treatment, thus avoiding much more expensive and complex late procedures. This kind of early detection may be helpful in avoiding severe consequences, e.g the breaking of an important part in a machine during its operation. The test for symmetry in contingency tables constitutes a broad and important subarea in Statistics, and several methods have been devised for this problem. In this work we propose the Full Bayesian Significance Test (FBST) for the problems of symmetry and point-symmetry in contingency tables. FBST is an intuitive Bayesian approach which does not assign positive probabilities to zero measure sets when testing sharp hypotheses. Numerical experiments comparing FBST performance to power-divergence statistics suggest that FBST is a good alternative for problems concerning tests for symmetry in contingency tables.

Baobabs Under Stress, A Case Of Estimated Pdf Transfer

Lucien Lemmens, Sebastiaan De Smedt, Roeland Samson

Using Bayesian techniques an ecological model for Baobabs responding to stress, is worked out. The main problem was to take the location and scale parameters for the response as well as for the covariates into account.

The Order-Theoretic Origin Of Special Relativity

Kevin Knuth, Newshaw Bahrenyi, Philip Goyal

In this paper, we present a novel derivation of special relativity and the information physics of events. We postulate that *events are fundamental*, and that some events have the potential to receive information about other events. This potential to be informed is not reciprocal, nor are all pairs of events related in such a way. This leads to the concept of a *partially-ordered set of events*, which is often called a causal set. Quantification proceeds by selecting two chains of coordinated events, each of which represents an observer, and assigning an integral valuation to each chain. Events can be projected onto each chain by identifying the earliest event on the chain that can be informed about the event. In this way, each event can be quantified by a pair of numbers, referred to a *pair*, that derives from the valuations on the chains. Pairs can be decomposed into a sum of symmetric and antisymmetric pairs, which correspond to time-like and space-like coordinates. We show that there exists a unique mapping from a real pair to a scalar, and that this gives rise to the Minkowski metric. Furthermore, we show that time is one-dimensional and space is multi-dimensional. We further assert that the scalar measure be invariant for all coordinated pairs of observers, which is as close as we get to assuming the principle of relativity. The consequences are that coordinates obey the Lorentz transformation, speed is a relevant quantity, and there is a maximum speed limit which is invariant. The result is an observer-based theory of special relativity that quantifies events with pairs of numbers. Events are fundamental and space-time is an artificial construct designed to make events look simple.

2010-07-05 - Poster Session – *Foundations* [16:45:00-19:00:00]

Non-relativistic Gravity in Entropic Quantum Dynamics

David Johnson, Ariel Caticha

Recent work has derived non-relativistic quantum theory using the concept of entropic dynamics.[1] Starting with the assumption that there is an irreducible uncertainty in the position of particles, one can introduce a statistical manifold for the probability distributions of these particles. Remarkably, one need not specify the exact form of the probability distributions. One then applies the method of Maximum Entropy and specifies a concept of time. By allowing the statistical manifold to be dynamical and by imposing energy conservation, the Schrödinger equation is recovered.

As in the standard formulation of non-relativistic quantum mechanics, entropic quantum dynamics is shown to be Galilean invariant. Furthermore, this invariance this invariance can be generalized to frames of reference in arbitrary accelerated motion (the extended Galilean transformation). Attention is given to how various quantities change in the boosted frame.

Crucial to the derivation is what is meant by a symmetry. For the physics to exhibit a symmetry, the dynamics described by two different frames of reference should be equivalent. Additionally, a notion of time for the boosted frame is required. The absolute Newtonian time of non-relativistic quantum mechanics flows at the same rate in all frames of reference. Accordingly, the very same definition of time used in the original frame is used here.

Finally, by carefully manipulating the conservation of energy condition in the original frame, one recovers the Schrödinger equation in terms of the boosted quantities. As a result, Einstein's equivalence principle in its strong form – that gravitational effects are equivalent to the fictitious effects that arise in non-inertial frames – arises naturally. The situation concerning the weak versions of the classical equivalence principle is interesting. Version A, that gravitational mass equals inertial mass, remains valid in quantum dynamics. Version B, that all bodies follow the same trajectories as they fall, is true classically but violated in quantum theory.

References: [1]Ariel Caticha, "From entropic dynamics to quantum theory," (2009), arXiv:0907.4335v3 [quant-ph]

2010-07-05 - Poster Session – *Foundations* [16:45:00-19:00:00]

Mixture of skewed alpha-stable distributions

Sayed Reza Hossaini Shojaei, Vahid Nassiri, Gholam Reza Mohammadian, Adel Mohammadpour

.

2010-07-05 - Poster Session – *Foundations* [16:45:00-19:00:00]

Bayesian Inference For Skewed Stable Distributions

Mona Shokripour, Vahid Nassiri, Adel Mohammadpour

.

2010-07-06 - Session *Algorithms for Maximum Entropy and Bayesian computation* [09:00:00-11:45:00]

Advice On "Computational Methods For Bayesian Model Choice"

John Skilling

Robert and Wraith are to be thanked for their exposé (MaxEnt 2009) of current statistical practice. After giving an overview of the principles of computational inference, this paper follows and refers to theirs, though a more critical and principled viewpoint rejects several of their suggestions as inappropriate for general use. Their mistaken account of nested sampling is corrected.

Several principles emerge from a study of inference algorithms.

- The controlling variable is the information (aka negative entropy, Kullback-Leibler).
- Inference is compressive and inherently asymmetric.
- Full Bayesian computation seeks compressive steps each acquiring no more than $O(1)$ information.
- Dimension isn't fundamental: information is.
- Topology may affect the choice of exploratory proposals, but it doesn't affect equilibrated samples.
- Point samples don't know their associated volume, and it's not generally possible to get normalisation directly from samples.
- Random Monte carlo sampling from the prior doesn't work, because the samples miss the posterior.
- Bridge sampling is diffusive, so not appropriate for large-scale compression.
- The harmonic mean algorithm is based on evaluating 0/0, and when that's fixed by weighting, its impossible to find the weights.
- The Chib approximation is based on repeatedly sampling facets of the posterior, but is blind to microscopic structure.
- Nested sampling does not have these failures, and is the only compressive algorithm that is immune to first-order phase changes.

2010-07-06 - Session *Algorithms for Maximum Entropy and Bayesian computation* [09:00:00-11:45:00]

Self-Adaptive Semantic Crossover for Genetic Programming

Julio Stern, Rafael Inhasz

Evolutionary algorithms are a class of stochastic optimization algorithms based on (inhomogeneous) Markov chains. Genetic Programming (GP) are evolutionary algorithms that work on populations, whose individuals represent possible (viable) solutions to the optimization problem. GP are meta-heuristics based on some key functions and operators inspired on evolution theories for biological species. The main goal of this paper is to define a new crossover operator named SASC or Self-Adaptive Semantic Crossover. SASC is based on meta-control techniques designed to guide the random selection of recombination points by a semantic compatibility function, a measure of functional similarity between the portions of code being swapped. It is important to realize that SASC's meta-control system is not hard-wired or pre-defined. On the contrary, it is an emerging feature, co-evolving with the population. Although the meta-control system is based on the history of each individual in the population, the required historical information is very limited. Hence, its implementation only generates a minor computational overhead.

2010-07-06 - Session *Algorithms for Maximum Entropy and Bayesian computation* [09:00:00-11:45:00]

Entropy-Based Search Algorithm for Experimental Design

N. K. Malakar, K. H. Knuth

The scientific method relies on the iterated processes of inference and inquiry. The inference phase consists of selecting the most probable models based on the available data; whereas the inquiry phase consists of using what is known about the models to select the most relevant experiment. This inquiry phase is also called the design phase of experimental design. It involves searching the parameterized space of experiments to select the experiment that promises, on average, to be maximally informative. In the case where it is important to learn about each of the model parameters, the relevance of an experiment can be measured using the Shannon entropy of the distribution of experimental outcomes predicted by a probable set of models. If the set of potential experiments is described by many parameters, we must search this high-dimensional entropy space. Brute force search methods will be slow and computationally expensive. We present an entropy-based search algorithm, called nested entropic sampling, to select the most informative experiment for efficient experimental design. This algorithm is inspired by Skillings nested sampling algorithm used in inference and borrows the concept of a rising threshold while a set of experiment samples are maintained. We demonstrate that this algorithm not only selects highly relevant experiments, but also is more efficient than brute force search. Such entropic search techniques promise to greatly benefit autonomous experimental design.

2010-07-06 - Session *Algorithms for Maximum Entropy and Bayesian computation* [09:00:00-11:45:00]

Bayesian inference for MAPK signaling networks

George Popescu, Sorina Popescu, Vasile Buzuloiu

Identification of signaling networks is a central topic in systematic analysis of cellular organization. Large scale methods based on functional protein micro-arrays offer the perspective of efficient and accurate identification of signaling networks. The prototypical MAPK signaling network has a hierarchical structure composed of three levels of nodes: the MAP3K proteins which phosphorylate MAP2K proteins which in turn, activates and phosphorylate MAPK proteins which act upon a multitude of cytoplasmic and nuclear substrates. In [1] we have described a new method for large scale identification of signaling networks that rely on detecting phosphorylation events using functional protein micro-arrays. We describe here a Bayesian inference method for MAPK signaling network identification.

We infer the MAPK network from observed protein phosphorylation signals using a Bayesian inference method. The network is built hierarchically from MAP2K to Effectors by making decisions at each level and assigning probability to phosphorylation events. Initially, all observations of MAPKK/MAPK combinations are used to predict putative phosphorylation events using a hierarchical Bayes method. Next, a small high-confidence set of putative MAPKK/MAPK phosphorylation events is selected for each of the ten MAPKs studied. We then identify the phosphorylated substrates by probing protein microarrays with all high-confidence combinations which activate Arabidopsis MAPKs. The conditional transduction of phosphorylation signal is modeled with a Bayesian network. We assign to each edge of the network probability values computed by testing on observations from replicated protein micro-array experiments. Finally, a Bayesian decision method is used to discriminate the MAPK-Effector events from background phophorylation control signal.

The structure of our reconstructed signaling network is significantly different from the signaling modules view, supporting the hypothesis of a combinatorial control of transcription through selective signal transduction in MAPK networks. Our global analysis of MAPK signaling removes some of the bias in previous studies of signaling networks guided by genotype/phenotype association methods.

2010-07-06 - Session *Algorithms for Maximum Entropy and Bayesian computation* [09:00:00-11:45:00]

Nested Sampling with Constrained Hamiltonian Monte Carlo

Michael Betancourt

Nested sampling is a powerful approach to Bayesian inference ultimately limited by the computationally demanding task of sampling from a heavily constrained probability distribution. An effective algorithm in its own right, Hamiltonian Monte Carlo is readily adapted to efficiently sample from any smooth, constrained distribution. Utilizing this constrained Hamiltonian Monte Carlo, I introduce a general implementation of the nested sampling algorithm.

Information theory of quantum systems with some hydrogenic applications

Jesus Sanchez-Dehesa, D. Manzano, P. Sánchez-Moreno, R.J. Yáñez

The information-theoretic representation of quantum systems, which complements the familiar energy description of the density-functional and wave-function-based theories, is here discussed. According to it, the internal disorder of the quantum-mechanical non-relativistic systems can be quantified by various single (Fisher information, Shannon entropy) and composite (e.g. Cramér-Rao, LMC shape and Fisher-Shannon complexity) functionals of the Schrödinger probability density $\rho(\vec{r})$. First, we examine these concepts and its application to quantum systems with central potentials. Then, we calculate these measures for hydrogenic systems, emphasizing their predictive power for various physical phenomena. Finally, some recent open problems are pointed out.

On moments-based Heisenberg inequalities

Steeve Zozor, Mariela Portesi, Pablo Sanchez-Moreno, Jesus Dehesa

In this communication, we propose to revisit the formulation of the Heisenberg uncertainty principle, stating the impossibility of refined measurement of a quantum particle position x and its momentum u, in terms of variance $\langle\|x\|^2\rangle\langle\|u\|^2\rangle \geq 1/4$. However, this inequality applies provided each variance exists. Thus, as an alternative, some authors proposed entropic versions of these inequalities, making use of Shannons entropy, Rényis entropies or Tsallis entropies. Furthermore, the primary entropic version, using Shannons entropy implies the Heisenberg relation, being in this sense more powerful than the classical Heisenberg relation. As another alternative, we revisit statistical moments-based formulations showing that inequalities involving moments of other orders than order 2 can be found. The way of making is based on the entropic versions of the Heisenberg relation together with the search for the maximal entropy under statistical moments constraints which can be summarized as follows: (1) Start with the Rényis entropy power uncertainty principle of the form $N\_\alpha(x)N\_\beta(u) \geq (\alpha, \beta)$ for each pair of positive indexes α and β [1]. (2) Search for the maximum entropy power $N\_\alpha(x)$ under the constraint $\langle\|x\|^a\rangle$ fixed, (and similarly in the momentum domain). (3) This leads to a family of uncertainty principle $\langle\|x\|^a\rangle^{2/a}\langle\|u\|^b\rangle^{2/b} \geq (a, b, \alpha, \beta)$ and the best bound of this family is $(a, b) = \max\_\alpha, \beta(a, b, \alpha, \beta)$. The couple of entropic indexes (α, β) that maximizes and the maximum are not analytically known, however the maximization can be simplified from the symmetric roles of x and u and from specific properties of (a, b, α, β). Our result improves a relation proposed very recently by Dehesa et al. [2] where the same approach was used but starting with the Shannon version of the entropic uncertainty relation. Furthermore, we can show that when $a = b$, the best bound we can find with our approach is given for $\alpha = \beta = 1$, that is the bound of [2], and in addition for $a = b = 2$, the variance-based Heisenberg relation is recovered. Note however that except for $a = b = 2$, the bound is probably not sharp since the probability distribution of the momentum corresponding to the probability distribution of the position that maximizes the entropy of the position does not coincide with the probability distribution of the momentum that maximizes the entropy of the momentum, except if both distributions are Gaussian, which occurs if and only if $a = b = 2$. In the final paper we will give an application of the uncertainty relation we found on the example of the 3-dimensional harmonic oscillator.

References: [1] S. Zozor, M. Portesi and C. Vignat, Physica A 387 (18-19): 4800, (2008).

[2] J.S. Dehesa, S. Lopez-Rosa and D. Manzano, Entropy and complexity analyses of D- dimensional quantum systems. In special issue: Statistical Complexities: Application to Electronic Structure, edited by K.D. Sen. Springer, Berlin, (2010).

Key Words: Generalized Heisenberg relation, entropic uncertainty relation, maximum Rényi entropy, moments constraints

2010-07-06 - Session *Information theory and statistical physics* [13:30:00-16:45:00]

Rigorous bounds for Renyi entropies of spherically symmetric potentials

Pablo Sanchez-Moreno, Steeve Zozor, Jesus Dehesa

The Rényi and Shannon entropies are information-theoretic measures which have enabled to formulate the position-momentum uncertainty principle in a much more adequate and stringent way than the (variance-based) Heisenberg-like relation. Moreover, they are closely related to various energetic density-functionals of quantum systems. Here we find sharp upper bounds to these quantities in terms of the second order moment $\langle r^2 \rangle$ for general spherically symmetric potentials, which substantially improve previous results of this type, by means of the Rényi maximization procedure with a covariance constraint due to Costa, Hero and Vignat. The contributions to these bounds coming from the radial and angular parts of the physical wavefunctions are explicitly given.

Entropic time

Ariel Caticha

A general framework for dynamics based on the method of maximum entropy is applied to non-relativistic quantum mechanics. The basic assumption of entropic dynamics is that in addition to the particles of interest there exist hidden variables which are uncertain and to which we can associate an entropy. The evolution of the particles, which is driven by the entropy of the hidden variables, is shown to be given by the Schrodinger equation. Early in the development of quantum mechanics the magnitude of the wave function received a statistical interpretation the Born rule. An important feature of the entropic dynamics approach to quantum theory is that not just the magnitude but also the phase of the wave function receives a statistical interpretation: the phase is simply related to the entropy of the hidden variables. Our specific goal in this paper is to discuss how a dynamics driven by entropy naturally leads to an entropic notion of time. We introduce time as a convenient book-keeping device to keep track of the accumulation of change. Our task here is to develop a model that includes (a) something one might identify as an instant, (b) a sense in which these instants can be ordered, (c) a convenient concept of duration measuring the separation between instants. A welcome bonus is that our model incorporates an intrinsic directionality an evolution from past instants towards future instants. Thus, an arrow of time does not have to be externally imposed but is generated automatically. This set of concepts constitutes what we will call entropic time. I also discuss the question of the relation between entropic time, in which instants are ordered through the sequence of steps of inference, and the presumably more objective notion of physical time. We argue that for the pragmatic purpose of predicting the observable correlations among particles nothing more physical than entropic time is needed.

Some topics on q-gaussians, escort distributions and Fisher information

Jean-François Bercher

Escort distributions have been introduced in statistical physics for the characterization of chaos and multifractals. These distributions P are a simple one parameter transformation of an original distribution p according to

$$P(x) = \frac{p(x)^q}{\int p(x)^q dx}. \tag{1}$$

The parameter q behaves as a microscope for exploring different regions of the measure p: for $q > 1$, the more singular regions are amplified, while for $q < 1$ the less singular regions are accentuated.

In Tsallis extended thermostatistics, the escort-averages, defined with respect to an escort distribution, have revealed useful in order to obtain analytical results and variational equations. In particular the equilibrium distributions obtained as maxima of Rényi-Tsallis entropy subject to a 'q-variance' constraint are the so-called q-gaussian

$$P(x) \propto \left(1-(1-q)\beta x^2\right)^{\frac{1}{1-q}} \tag{2}$$

In applied fields, Tsallis distributions (q-distributions) have encountered a large success because of their remarkable agreement with experimental data.

In this contribution, we show that escort distributions (1) emerge naturally as a maximum entropy trade-off between the distribution $p(x)$ and the uniform distribution. This setting may typically describe a phase transition between two states. Then, looking for the distribution p with maximum entropy subject to a constraint on the mean energy computed with respect to the escort distribution, we arrive at the q-gaussian (2).

But escort distributions also appear in the fields of quantization and coding with interesting consequences. In the problem of quantization, we show that the reasoning above can be adopted so as to find the best distribution of quantized values, that preserves the maximum of the information content (as measured by the entropy). For the problem of coding, we recall a source coding theorem by Campbell relating a generalized measure of length to the Rényi-Tsallis entropy. We show that the associated optimal codes can easily be obtained using considerations on escort-distributions and that this provide an easy implementation procedure. We also show that these generalized lengths are bounded below by the Rényi entropy.

That q-gaussians arise from the maximization of Rényi-Tsallis entropy subject to a q-variance constraint is a known fact. We show here that the (squared) q-gaussian also appears as a minimum of Fisher information subject to the same q-variance constraint, and of course, that we recover the standard results in the limit $q = 1$. From this finding, we also obtain a generalized Cramér-Rao result. Finally, we show that analog results exist for the minimization of the escort q-Fisher information subject to a standard variance constraint.

Non-Extensive Radiobiology

Oscar Sotolongo-Grau, Daniel Rodriguez-Perez, Carlos Antoranz, Oscar Sotolongo-Costa

In order to find a theoretical expression for the survival fraction of cells under radiation a non extensive entropy approach, based on Tsallis entropy, is proposed. The obtained model is consistent with empirical models and fits better to the experimental data. The q-algebra allows to compare the damage over tumors and healthy tissues so as to choose the best treatment.

Keywords: Radiobiology, Survival fraction, Entropy

2010-07-06 - Poster session – *Applications* [16:45:00-19:00:00]

Online Robot Dead Reckoning Localization Using Maximum Relative Entropy Optimization With Model Constraints

Renaldas Urniezius

The principles of Maximum relative Entropy optimization were analyzed for dead reckoning localization of one axis differential drive robot. Model constraints were derived from the relationships between robot sensors. The real life experiment's results confirmed that online monitoring of accelerometer axiss zero bias dynamic calculation result matched the initial knowledge about its static value. Simultaneous time series updating with model constraints was derived where result is iteration-free analytical representation of entropy distribution function. A prior distribution has been calculated using Maximum Entropy principle incorporating model constraints between the sensors. Then a posteriori distribution was calculated using Maximum relative Entropy principles.

2010-07-06 - Poster session – *Applications* [16:45:00-19:00:00]

Nonlinear Source Separation: A Maximum Likelihood Approach for Quadratic Mixtures

Chahinez Chaouchi, Yannick Deville, Shahram Hosseini

This paper deals with the blind source separation problem for a specific class of nonlinear mixtures. The major part of the literature concerns the linear model while the nonlinear case is still less studied. The latter problem is more difficult to handle, so in order to simplify it several authors constrain the type of nonlinearity. In this work, the considered mixing model is the quadratic one. We use an associated separating structure and we propose a maximum likelihood approach to recover the source signals. Key words: Blind Source Separation, Quadratic Mixtures, Maximum Likelihood

Superresolution of compact neutron spectrometers, with applications to fusion diagnostics

Marcel Reginatto, Andreas Zimbal

The ability to achieve resolution that is better than the instrument resolution (i.e., superresolution) is well known in optics, where it has been extensively studied. Unfortunately, there are only a handful of theoretical studies concerning superresolution of particle spectrometers, even though experimentalists are familiar with the enhancement of resolution that is achievable when appropriate methods of data analysis are used, such as maximum entropy deconvolution and Bayesian methods. The authors are not aware of any theoretical studies that deal specifically with neutron spectrometers, despite the fact that knowledge of the superresolution factor is in many cases important. For example, in applications of neutron spectrometry to fusion diagnostics [1], the temperature of a burning plasma is an important physical parameter which may be inferred from the width of the neutron emission peak, and the ability to determine this width depends on the superresolution factor. Kosarev [2] has derived an absolute limit for resolution enhancement, known as Shannons superresolution limit for signal recovery, using arguments based on a well known theorem of Shannon. However, most calculations of superresolution factors in the literature are based on the assumption of Gaussian, translationally invariant response functions (or blurring function, to use the terminology of optics) and these results are therefore not directly applicable to neutron spectrometers which have response functions that do not satisfy these requirements. In this work, we reexamine Kosarevs analysis, derive estimates of superresolution for liquid scintillator spectrometers of a type commonly used for neutron measurements and, in particular, consider maximum entropy deconvolution and Bayesian methods and study the enhancement of resolution that is achievable with these approaches. Theoretical superresolution factors are compared to experimental results from a previous publication [1], where Bayesian and maximum entropy methods were used to analyze fusion-relevant neutron measurements carried out with a compact organic liquid scintillation spectrometer under controlled conditions.

References: [1] M. Reginatto and A. Zimbal, Rev. Sci. Instrum. 79, 023505 (2008). [2] E. L. Kosarev, Inverse Problems 6, 55-76 (1990).

Key Words: Superresolution, maximum entropy, neutron spectrometry, fusion diagnostics

Bayesian approach to siRNA-mediated gene silencing phenotypes quantification and off-target level estimation

Yannis Kalaidzidis, Jochen Rink, Claudio Collinet, Akhila Chandrashaker, Martin Stötter, Eugenio Fava, Marino Zerial

Small Interfering RNA (siRNA) and automated high-throughput high-resolution microscopy provides technological platform for systematic genome-wide survey of individual gene knockdown phenotype. Quantified knockdown phenotype can be used for gene functions elucidation and the mechanistic understanding of cellular processes in which genes participate. However, the large degree of morphological variation between cells in repetitions of biological experiment as well as variation between phenotypes of different siRNAs, which are targeting the same gene, represents a major challenge to the reliable identification of gene silencing phenotypes. We have developed a system for the high content analysis of automatically acquired high-resolution images, which describes the endosomal organelles in quantitative terms (multiparametric gene silencing profile). The stability of individual parameters of phenotypic profiles was tested by calculation the correlation between different imaging sessions and between experimental replicates. The analysis showed that different parameters reveal a wide variation of stabilities which dependent on biological variability, typical automatic imaging problems and parameter calculation details. The comparison of parameters stability between different siRNAs targeting the same gene with respective stability between different replicates provides a tool for establishing the phenotype reproducibility and mean off-target (genes silencing beside the main target) level of siRNA libraries. Bayesian approach for estimation of most probable gene profile, given (contradictive) profiles of individual siRNAs was developed. Analysis of experimental data reveals the level of off-target and gives an estimation of the minimum number of independent siRNAs which are required to infer the gene knockdown phenotype with given confidence. This approach was applied for conveying and analysis of genome-wide screen of genes involved in endocytosis [Collinet et al, Nature, 2010] where it proves it superiority in comparison with other approaches used in screening community. Beside of that, the estimation of off-target level gives an objective feedback for the new generation siRNA development and provides insight for detail understanding of siRNA-mediated gene silencing mechanism.

On combined estimation of expected backgrounds with statistical and systematic uncertainties in planned experiment

Sergey Bityukov, Nikolai Krasnikov, Vera Smirnova

The expected numbers of observed events of several backgrounds in experiment are estimated from Monte Carlo experiments. In the analysis we take into account an integrated luminosities of Monte Carlo experiments. The expected number of observed events allows to construct the distribution of probabilities of number of events which real experiment may observe (in accordance with formulae in papers [1]). The formulae allow to take into account statistical uncertainty of corresponding Monte Carlo experiment. The influence of systematics is determined by additional Monte Carlo experiments with expected number of events.

[1] S.I.Bityukov, On the Signal Significance in the Presence of Systematic and Statistical Uncertainties, JHEP **09** (2002) 060 [hep-ph/0207130]; S.I. Bityukov, N.V. Krasnikov, Signal Significance in the Presence of Systematic and Statistical Uncertainties, Nucl.Inst.
Meth. **A502**, (2003) 795.

Towards a Bayesian seismotectonic zoning for use in Probabilistic Seismic Hazard Assessment (PSHA)

Boris Le Goff, Delphine Fitzenz, Céline Beauval

The mathematical representation of seismic sources is an important part of proba- bilistic seismic hazard assessment. This process re ects the association of the seis- micity with the tectonically-active geological structures provided by seismotectonic studies. Given that most active faults are not characterized well enough, seismic sources are generally defined as area zones. This method was defined for regions of high tectonic plate velocities, leading to an important seismicity and clear surface fault expression. It is based on delimiting area zones, with

nite boundary poly- gons, within which the geological features of active tectonics and the seismicity are deemed homogeneous (e.g., focal depth, seismicity rate, and maximum magnitude). Besides the lack of data (e.g., narrow range of recorded magnitudes), the applica- tion of this method to regions where the tectonic plate velocities are lower engenders different problems: 1) a large sensitivity of resulting hazard maps on the location of zone boundaries, while these boundaries are set by expert decision; 2) the zoning can not represent the depth-variation in faulting mechanism; 3) the seismicity rate is distributed throughout the zone and we lose the location of the determinant in- formation used for their calculation. We propose an alternative procedure for area source modeling. For a given point in space, belonging to a given "zone" , from now on called seismicity model, will be expressed by a probability computed from the 2D (spatial) probability density function (pdf) for the active tectonic model used as an a priori and updated with the seismicity catalog or other new data sources (e.g., geomorphology, subsurface exploration). A point will thus be allowed to con- tribute to several models, with weights given by the value of the pdf for each model at this location. The hazard due to a source at a site will be the weighted aver- age of the hazard from the different models containing this same source. Then the results will provide the full spectrum of variability in the hazard. We will start with simple synthetic cases featuring imperfect seismological monitoring coverage, imperfect tectonic maps, and various faulting styles. We will show the alternative sets of seismicity models and we will propagate all sources of uncertainties through the ground motion prediction equations to the resulting hazard maps in order to highlight the sites where future data acquisitions are needed.

Application Of A Bayesian Inference Method To Reconstruct Short-Range Atmospheric Dispersion Events

Inanc Senocak

In the event of an accidental or intentional release of chemical or biological (CB) agents into the atmosphere, first responders and decision makers need to rapidly locate and characterize the source of dispersion events using limited information from sensor networks. This inverse problem has been referred to as source term estimation, source inversion or event reconstruction. Source term estimation algorithms provide critical information for emergency response, such as how much agent was released, where, and when it was released. An estimate of the dispersion source is important to decision makers to determine the extent of contamination for emergency response, and get a more reliable account of the event for remediation efforts and forensic studies. To this end, Fusing Sensor Information from Observing Networks (FUSION) Field Trial 2007 (FFT 07) was designed to support the development of source term estimation algorithms and evaluate existing ones [1]. FTT07 data set provides detailed meteorological information and concentration measurements for short range plume dispersion experiments that were performed for a variety of release types, including single and multiple sources for continuous and puff releases.

A stochastic event reconstruction tool (SERT), which has been developed and tested for controlled pollution experiments [2], is applied to a subset of the FFT07 to estimate the source of single source short-range continuous releases. The inference in SERT is based on Bayesian inference with Markov chain Monte Carlo (MCMC) sampling. SERT adopts a probability model that takes into account both positive and zero-reading sensors. In addition to the location and strength of the dispersion event, empirical parameters in the forward model are estimated to establish a data-driven plume model. Results demonstrate the effectiveness of the Bayesian inference approach to characterize the source of a release with uncertainty quantification and its potential in data-driven modeling for air pollution modeling.

References:

[1] D.P. Storwold. West Desert Test Center, US Army Dugway Proving Ground, WDTC-TP-07-078 (2007).

[2] I. Senocak et al. Atmos. Environ. **42**, 7718 (2008).

GNU MCSim: Bayesian Statistical Inference for System Biology Models

Rudy Gunawan, Frederic Bois

Parameter estimation is a crucial component in system biology. Many complex physiologically-based pharmacokinetic (PBPK) models contain several parameters that need to be calibrated against multiple data sets before the models can be validated. Some of the parameters, such as kinetic parameters, widely vary among individuals; whereas, some may not. Prior information and uncertainty about each parameter are important consideration; this makes Bayesian inference the most feasible approach. In this context, we need a computational tool that seamlessly connects PBPK (or in general, Systems Biology Markup Language) models to numerical Bayesian inference technique, namely Markov Chain Monte Carlo. GNU MCSim was precisely developed to achieve this purpose, in a general non-linear differential context which is often found in system biology modeling. Starting with version 5.3.0, GNU MCSim reads in and simulates Systems Biology Markup Language models. Markov chain Monte Carlo simulations can be used to generate samples from the joint posterior distribution of the model parameters, given a data set and prior distributions of the parameters in the model. Hierarchical statistical models can easily be implemented.

A Bayesian approach to Fourier Synthesis inverse problem with application in SAR imaging

Sha Zhu, Ali Mohammad-Djafari

In this paper we propose a Bayesian approach to the ill-posed inverse problem of Fourier synthesis (FS) which consists in reconstructing a function from partial knowledge of its Fourier Transform (FT) with application in SAR (Synthetic Aperture Radar) imaging. The function to be estimated represents an image of the observed scene. Because this observed scene is mainly composed of point sources, we propose to use a Generalized Gaussian (GG) prior model for that. We use then the Maximum A posteriori (MAP) estimator as the desired solution.

In particular, we are interested in bistatic case of spotlight-mode SAR data from satellitates. In a first step, we consider real valued reflectivities but we account for the complex value of the measured data. The relation between the Fourier transform of the measured data and the unknown scene reflectivity is modeled by a 2D spatial FT. The inverse problem becomes then a FS and depending on the geometry of the data acquisition, only the set of locations in the Fourier space are different.

We give a detailed modelisation of the data acquistion process that we simulated. The model consisits of bistatic case based on spaceborne SAR. We then applied the proposed method on those synthetic data to measure its performances compared to some other classical methods [1].

Finally, we demonstrate the performance of the method on experimental SAR data obtained in a collaborative work by ONERA [2].

2010-07-06 - Poster session – *Applications* [16:45:00-19:00:00]

Optimal Path and Sensor Selection for Target Tracking in Wireless Sensor Networks

Majdi Mansouri, Hichem Snoussi, Moez Esseghir, Cedric Richard

Wireless Sensor Networks nodes are powered by small batteries, which are in practical simulations non rechargeable, either due to cost limitations or because they are deployed in hostile environments with high temperature, high pollution levels, or high nuclear radiation levels. These considerations enhance energy-saving and energy-efficient WSN designs. One approach to prolong battery lifetime is to use an optimal sensors selection protocol and to choose the best communication path between the candidate sensor and the cluster head. In a typical wireless sensor network, sensors are employed to achieve some specific task, e.g. tracking objects. These nodes are severely constrained in energy and in most case can not be recharged. Thus minimizing the communication costs between sensor nodes is critical to prolong the lifetime of sensor networks. Another important metric of sensor networks is the accuracy of the sensing result of the target in that several sensors in the same cluster can provide redundant data. Because of physical characteristics such as distance, modality, or noise model of individual sensors, data from different sensors can have various qualities. Therefore the accuracy depends on which sensor and which link the cluster head selects. The primary focus of this work is to study the Bayesian inference problem in distributed WSNs with particular emphasis on the trade-off between estimation precision and energy-awareness. We propose a variational approach to approximate the particle distribution to a single Gaussian distribution, while respecting the communication constraints of WSNs. The efficiency of the variational approximation relies on the fact that the online update and the compression of the filtering distribution are simultaneously performed. In addition, the variational approach has the nice property to be model-free, ensuring robustness of signal processing. We analyze the Bayesian inference issue for several specific but representative WSN applications to elaborate the variational filtering (VF) method, which can be applicable to a wider class of problems. We propose to improve the use of the VF by jointly selecting the optimal candidate sensor that participates in target tracking and its best communication path to the cluster head (CH). In the current work, firstly, we select the optimal sensor in order to provide the required data of the target and to balance the energy dissipation in the WSN. This selection is also based on the local cluster nodes density and their transmission power. Secondly, we select the best communication path that achieves the highest signal-to-noise ratio (SNR) at the CH, then we estimate the target position using the VF algorithm. The best communication path is designed to reduce the communication cost, which leads to a significant reduction of energy consumption and an accurate target tracking. The optimal sensor selection is based on Mutual Information maximization under energy constraints. This latter is computed by using the target position predictive distribution provided by the VF algorithm.

Signals and Images Foreground/Background Joint Estimation and Separation

Boujemaa Ait-El-Fquih, Ali Mohammad-Djafari

The original motivation of this work is to jointly estimate and separate a background signal $\mathbf{b}$ and a foreground one $\mathbf{f}$ from their superposition $\mathbf{y} = \mathbf{b} + \mathbf{f} + \epsilon$. The signal $\mathbf{b}$ is modeled by a homogeneous autoregressive (AR)process while $\mathbf{f}$ is non homogeneous and depending on a binary valued Markov variable $r \in \{r^b, r^f\}$ separating the foreground from the background. $\mathbf{f}|r = r^f$ and $\mathbf{f}|r = r^b$ are then modeled by two AR processes with different parameters. In other terms, this problem is modeled by a conditionally linear and Gaussian hidden Markov chain (CLGHMC), in which $(\mathbf{b}, \mathbf{f})$ is a linear and Gaussian hidden Markov chain given the Markov chain r.

So, we start by deriving a Bayesian filtering algorithm in a general conditionally hidden Markov chain framework. This algorithm propagates the filtering densities associated to the foreground $\mathbf{f}$ and the background $\mathbf{b}$ signals. We next focus on our CLGHMC model in which the filtering densities of interest are Gaussian sum distributions. We thus reduce the algorithm of the propagation of these densities to an algorithm allowing the propagation of Monte Carlo approximations of their means and covariance matrices. Finally, we perform some simulations intended to show the interest of the proposed algorithm in foreground/background signals and images separation problems.

Maxent For Climate Impact Prediction For Mangrove Habitat In The Bay Of Bengal Rim

Md. Saiful Islam

Mangroves is a restrictive habitat. Mangroves occur in coastal zones with regular saline tidal inundation which gives mangroves competitive advantage over other vegetation types as mangroves are adapted to cope with salinity. At a global scale, they flourish within tropical and sub-tropical region being restricted by cold. They also fall within rainforest regions, showing correlation with amount and pattern of annual precipitation. Some literature also suggests there is correlation of mangrove habitats with soil types. Climate change scenario will alter most of the factors of mangrove habitat with predicted sea level rises and change in temperature and precipitation. Hence MaxEnt approach is tried to combine all these different types of factors to model mangrove habitat in the Bay of Bengal rim for present and future habitat under climate change scenario. It is found that MaxEnt software for species habitat modelling can predict current mangrove extent and thereby validates the future predictions. However, some factors such as soil type is found to be statistically less contributing factor to mangroves occurrence as opposed to some papers. MaxEnt can be used to model community habitats under climate change scenario and refine understanding of factors relevant for a particular habitat.

Total Reliability Management by Bayesian Approach

Ding Huang

Bayesian statistical framework is presented for computing scheduled calibration intervals. The methodology applies the principle of maximum information entropy to construct failure rate prior, computes posterior distribution. It then derives expected reliability function given data, and the problem now becomes a decision making problem. It generates Bayesian intervals with PThe mainstream calibration interval in present practice is conventional statistics based. It fails to provide managers a tool for total reliability management. This new proposal is capable of solving, even for extreme cases, e.g., when sample size is small, no failure data is observed, or observation time is short.
References: [1] Jaynes, E. T., 1976. ‘Confidence Intervals vs Bayesian Intervals,’ in Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science, W. L. Harper and C. A. Hooker (eds.), D. Reidel, Dordrecht, p. 175; as a level 2 postscript file. [2] ISO Guide to the Expression of Uncertainty in Measurement (Geneva: ISO) 1995 [3] Establishment and Adjustment of Calibration Intervals Recommended Practice RP-1, National Conference of Standards Laboratories Calibration Interval Committee, January 2010. Key words: calibration Interval Analysis, reliability target, total reliability management, expected reliability function, decision making, risk analysis

2010-07-07 - Session *Bayesian algorithms for multi-channel image analysis* [09:00:00-11:45:00]

Three-dimensional visualization and identification of objects in photon starved scenes using statistical estimation

Bahram Javidi, Dipak Dey, M. Cho, M. Jung

In this invited talk, we present an overview of three-dimensional (3D) object recognition and reconstruction in photon starved scenes. The passive sensing is performed using multi view photon counting imaging. Both Maximum Likelihood Estimation (MLE) and Bayesian methods will be discussed to reconstruct 3D images from photon counting images. Object recognition using a variety of statistical approaches will be presented.

2010-07-07 - Session *Bayesian algorithms for multi-channel image analysis* [09:00:00-11:45:00]

Bayesian fusion of hyperspectral astronomical images

André Jalobeanu, Matthieu Petremand, Christophe Collet

We propose a new fusion method to combine hyperspectral astronomical observations acquired under varying conditions (sampling lattice, geometric distortions, PSF...) into a single model while removing most of the instrumental and acquisition-related artifacts. Our approach is based on a Bayesian framework allowing for optimal data fusion and uncertainty estimation at the same time. Thus, the results can be accurately and consistently analyzed by astronomers; for instance, astrometry and photometry could be computed and error estimation made possible in a rigorous way.

Key Words: hyperspectral, fusion, Bayesian, uncertainties, astronomy

2010-07-07 - Session *Bayesian algorithms for multi-channel image analysis* [09:00:00-11:45:00]

Unmixing hyperspectral images using Markov random fields

Olivier Eches, Nicolas Dobigeon, Jean-Yves Tourneret

This paper proposes a new spectral unmixing strategy based on the normal compositional model that exploits the spatial correlations between the image pixels. The pure materials (referred to as endmembers) contained in the image are assumed to be available (they can be obtained by using an appropriate endmember extraction algorithm), while the corresponding fractions (referred to as abundances) are estimated by the proposed algorithm. Due to physical constraints, the abundances have to satisfy positivity and sum-to-one constraints. The image is divided into homogeneous distinct regions having the same statistical properties for the abundance coefficients. The spatial dependencies within each class are modeled thanks to Potts-Markov random fields. Within a Bayesian framework, prior distributions for the abundances and the associated hyperparameters are introduced. A reparametrization of the abundance coefficients is proposed to handle the physical constraints (positivity and sum-to-one) inherent to hyperspectral imagery. The parameters (abundances), hyperparameters (abundance mean and variance for each class) and the classification map indicating the classes of all pixels in the image are inferred from the resulting joint posterior distribution. To overcome the complexity of the joint posterior distribution, Markov chain Monte Carlo methods are used to generate samples asymptotically distributed according to the joint posterior of interest. Simulations conducted on synthetic and real data are presented to illustrate the performance of the proposed algorithm.

2010-07-07 - Session *Bayesian algorithms for multi-channel image analysis* [09:00:00-11:45:00]

An Investigation of Likelihoods and Priors for Bayesian Endmember Estimation

Alina Zare, Paul Gader

A Gibbs sampler for piece-wise convex hyperspectral unmixing and endmember detection is presented. The convex geometry model is a commonly used model for hyperspectral data and assumes that each spectrum, $\mathbf{x}\_i$, is a convex combination of the endmembers (the spectral signatures of the constituent materials) in a scene. The model can be written as $\mathbf{x}\_i = \sum\_k = 1^M p\_ik\mathbf{e}\_k + \epsilon\_i$ where N is the number of pixels, M is the number of endmembers, ϵ_i is an error term, p_ik is the abundance of endmember k in pixel i, and $\mathbf{e}\_k$ is the k^{th} endmember. The abundances of this model satisfy the constraints $p\_ik \geq 0, \forall k = 1, \ldots, M$ and $\sum\_k = 1^M p\_ik = 1$. This standard model assumes that hyperspectral data reside in a single convex region. However, hyperspectral data is often non-convex. Furthermore, endmembers are generally represented as a single point in the high dimensional space. However, the spectral signature for a material varies as a function of the inherent variability of the material or environmental conditions. Therefore, it is more appropriate to represent each endmember as a full distribution to incorporate the variability and utilize this information during spectral unmixing. A Gibbs sampler that searches for several sets of endmember distributions, i.e. a piecewise convex representation, is presented. The hyperspectral data is partitioned among the sets of endmember distributions using a Dirichlet process prior that also estimates the number of needed sets. The proposed likelihood follows from a convex combination of Normal endmember distributions with a Dirichlet prior on the abundance values. A Normal distribution is also applied as a prior for the mean values of the endmember distributions.

The Gibbs sampler that is presented partitions the data into convex regions, determines the number of convex regions required and determines endmember distributions and abundance values for all convex regions. This method is a fully stochastic extension of a previously published methods. Results are presented on simulated and real hyperspectral data that indicate the ability of the method to effectively estimate endmember distributions and the number of sets of endmember distributions.

2010-07-07 - Session *Bayesian algorithms for multi-channel image analysis* [09:00:00-11:45:00]

Multichannel SAR Image Classification by Finite Mixtures, Copula Theory and Markov Random Fields

Vladimir A. Krylov, Gabriele Moser, Sebastiano B. Serpico, Josiane Zerubia

The last decades have witnessed an intensive development and a significant increase of interest to remote sensing, and, in particular, to synthetic aperture radar (SAR) imagery. In this paper we develop a supervised classification approach for medium and high resolution multichannel SAR amplitude images. The proposed technique combines finite mixture modeling for probability density function estimation, copulas for multivariate distribution modeling and the Markov random field approach to Bayesian image classification. The finite mixture modeling is done via a recently proposed SAR-specific dictionary-based stochastic expectation maximization approach [1] to class-conditional amplitude probability density function estimation, which is applied separately to all the SAR channels. For modeling the class-conditional joint distributions of multichannel data the statistical concept of copulas [2] is employed, and a dictionary-based copula selection method is proposed. Finally, the Markov random field approach [3] enables to take into account the contextual information and to gain robustness against the inherent noise-like phenomenon of SAR known as speckle. The designed method is an extension and a generalization to multichannel SAR of a recently developed single-channel and Dual-pol SAR image classification technique [4]. The accuracy of the developed multichannel SAR classification approach is validated on several multichannel Quad-pol RADARSAT-2 images and compared to benchmark classification techniques.

2010-07-08 - Session *Non parametric methods and experimental design* [09:00:00-11:45:00]

The Maximum Entropy Formalism And A Generalization Of The Symmetric Dirichlet Distribution

Dawn Holmes

(1) Department of Statistics and Applied Probability, South Hall,
University of California, Santa Barbara,CA 93106, USA
(email: holmes@pstat.ucsb.edu)
9 April 2010

Abstract

Bayesian networks consist of an underlying joint probability distribution and a directed acyclic graph in which a propositional variable that represents a marginal distribution is stored at each vertex in the graph. In this paper we consider two particular classes of Bayesian networks, binary trees and binary inverted trees. The method of interval constraints explored by Neapolitan is contrasted with the maxi- mum entropy method for assigning prior probabilites. Both methods are based on the principle of indifference; the distributions derived through Neapolitan's method are a generalization of the symmetric Dirichlet distribution in which prior ignorance is assumed. The principle of maximum entropy can be viewed as a generalization of the principle of indifference that allows the assumption of prior ignorance to be waived; the distributions thus derived are constrasted with those derived through the method of interval constraints. The paper draws upon and extends the authors previous work on the incorporation of interval-constraint knowledge into the prior distribution of a Bayesian network. The underlying philosophical focus, objective Bayesianism is discussed briefly.

Key Words: Bayesian networks, Maximum entropy, d-separation

Maximum Entropies Copulas

Doriano-Boris Pougaza, Ali Mohammad-Djafari

Copulas have been proved useful when dealing with the modelling of the dependence structure in a group of variables. The problem in which we are interested here is to find the bivariate distributions when we know only its marginals. This problem is an ill-posed inverse problem, in the sense that the problem has not a unique solution. One of the way to propose a unique solution to this problem is to choose an appropriate copula and then use the *Sklar's theorem* [1]. The problem then becomes the choice of copula. Two years before *Sklar's theorem* was published, Edwin Jaynes proposed, the *Principle of maximum entropy* (PME) which is the way to assign a probability law to a quantity about which we have partial information. PME has been used in many areas and originally when the partial information is knowledge of some geometric or harmonic moments [2].

Optimizing the entropy of a joint distribution subject to given marginals was studied in statistical and probabilistic litterature [3] since the 1930s. The existence of the solution has also been known. The case where the entropy considered is the Shannon entropy was also discussed more extensively by some authors, where derivation of the joint distribution with given uniform marginals on $I = [0, 1]$ and the correlation coefficient, that minimizes the mutual information is considered.

Here the partial information is only the knowledge of the marginal distributions. We propose then to study the solution we obtain when using differents entropies, Shannon, Rényi, Burg, Tsallis. Our main contribution is to consider the cases where we can obtain analytical expressions for the maximum entropy problem and so the copula families. Interestingly, when using Shannon entropy, we obtain the independent copula, which is a very known family of copula. But, when using other entropy expressions, we obtained really new families of copula, and in our knowledge, there have not been known before.

References:

[1] A. Sklar, Publ. Inst. Stat. Univ. Paris 8., Fonctions de repartition à n dimensions et leurs marges (1959).

[2] A. Mohammad-Djafari, Traitement du signal, GRETSI, Maximum d'entropie et problèmes inverses en imagerie (1994).

[3] H. Cramér, and H. Wold, J.London Math. Soc., Some theorems on distribution functions (1936).

2010-07-08 - Session *Non parametric methods and experimental design* [09:00:00-11:45:00]

Digital Particle Image Velocimetry Using Splines In Tension

Udo von Toussaint, Silvio Gori

Digital particle imaging velocimetry has become a widely used diagnostic technique for the extraction of quantitative information about flow fields. Multiple-view geometries and pulsed light sources are used to identify and track individual particles (tracers) at discrete times. The particle velocity and acceleration are then commonly derived by ill-conditioned methods based on finite-differences of the noisy measurements of the particle positions. We will present a different Bayesian approach based on an algorithm which models the particle in 3-d velocity space using splines in tension, thus automatically including the physical constraints of finite acceleration. The properties of the new algorithm will be discussed and compared with conventional approaches using mock and real world data from turbulence experiments.

2010-07-08 - Session *Non parametric methods and experimental design* [09:00:00-11:45:00]

Confidence Distributions In Statistical Inference

Sergey Bityukov, Nikolai Krasnikov, Saralees Nadaraja, Vera Smirnova

This paper reviews the new methodology for statistical inferences. Point estimators, confidence intervals and $p-$values are fundamental tools for frequentist statisticians. Confidence distributions, which can be viewed as "distribution estimators", are often convenient for constructing all of the above statistical procedures and more.

The notion of confidence distribution (CD) is in essence a Neymanian interpretation of Fisher's fiducial distribution. It contains information related to every kind of frequentist inference. The confidence distribution is a direct generalization of the confidence interval, and is a useful format of presenting statistical inference.

The following quotation from Efron (Efron, 1998) on Fisher's contribution of the fiducial distribution seems quite relevant in the context of CDs: "... here is a safe prediction for the 21st century: statisticians will be asked to solve bigger and more complicated problems. I believe there is a good chance that objective Bayes methods will be developed for such problem, and that something like fiducial inference will play an important role in this development. Maybe Fisher's biggest blunder will become a big hit in the 21st century!".

References:

Efron B., 1998. R.A. Fisher in the 21st century, Statistical Science 13, 95-122.

2010-07-08 - Session *Non parametric methods and experimental design* [09:00:00-11:45:00]

Bayesian Validation of CAIX Immunohistochemical Assays of CAIX

Rudy Gunawan, Giuseppe Tenti, Siv Sivaloganathan

Cervical cancer is a fatal disease if not detected early. According to National Cancer Institute (2009), the 5-year survival rates were 91.5% when it is detected early; however, the survival rate plummeted to 17.2% when it is detected late in its metastatic phase. Due to the site of the tumor, most diagnostic tools produce low signals and multiple samples are difficult to obtain. Hence, prescribing an effective cancer therapy is a challenging task in cervical cancers. Current treatments for cervical cancer patients include chemotherapy, radiotherapy, or a combination of both. The effectiveness of chemotherapy or radiotherapy depends crucially on the oxygen content in cervical carcinomas; thus, an accurate tumor oxygen quantification tool is needed.

The preferred tumor oxygen quantification technique is a biopsy-based immunohistochemical assay of a certain protein marker activated when tumor oxygen is low (also known as hypoxia). Since this technique is based on biopsies, there are only limited samples. Previously, we infer tumor hypoxia based on the available biopsies in the context of Bayesian framework. In the current project, we want to validate our inferences against the acclaimed gold standard method for tumor hypoxia quantification, namely the Eppendorf polarographic needle electrode; in the context of Bayesian approach, we infer whether or not the two tumor hypoxia quantification techniques correlate with each other.

In collaboration with the Princess Margaret Hospital in Toronto, Canada, we had the immunohistochemical data and needle electrode data from 24 patients with advanced cervical cancers for which we validated our inference. Successful demonstration of this analysis will provide a sound justification whether or not the indirect method (biopsy-based) of tumor hypoxia quantification can replace the direct (but more invasive) one. Finally, future directions of this research include the analysis of multi-marker data for cervical cancers to provide a more accurate description of tumor hypoxia within the tumor.

2010-07-08 - Session Bayesian and maximum entropy inference in action: applications [13:30:00-16:45:00]

The use of bayesian programming in industrial applications

Emmanuel Mazer

We will illustrate the bayesian programming approach by giving several examples related to industrial applications. We will show that bayesian programms are general enough to cover applications from banking to boiler heater.

Parameter Estimation as a Problem in Statistical Thermodynamics

Keith Earle, David Schneider

In this work, we explore the connections between parameter fitting and statistical thermodynamics using the maxent principle of Jaynes as a starting point. In particular, we show how signal averaging may be described by a suitable one particle partition function, modified for the case of a variable number of particles. These modifications lead to an entropy that is extensive in the number of measurements in the average. Systematic error may be interpreted as a departure from ideal gas behavior. In addition, we show how to combine measurements from different experiments in an unbiased way in order to maximize the entropy of simultaneous parameter fitting. We suggest how fit parameters may be interpreted as generalized coordinates and the forces conjugate to them may be derived from the system partition function. From this perspective, the parameter fitting problem may be interpreted as a process where the system (spectrum) does work against internal stresses (non-optimum model parameters) to achieve a state of minimum free energy/maximum entropy. We introduce a suitable definition of volume that allows one to define compressibilities and thus obtain further insights into the fitting process from classical thermodynamics. Finally, we show how the distribution function allows us to define a geometry on parameter space, building on previous work[1,2]. This geometry has implications for error estimation and we outline a program for incorporating these geometrical insights into an automated parameter fitting algorithm.

2010-07-08 - Session Bayesian and maximum entropy inference in action: applications [13:30:00-16:45:00]

Reinforcement Learning by Relative Entropy Policy Search

Jan Peters, Katharina Muelling, Yasemin Altun

Policy search is a successful approach to reinforcement learning. However, policy improvements often result in the loss of information. Hence, it has been marred by premature convergence and implausible solutions. As first suggested in the context of covariant policy gradients, many of these problems may be addressed by constraining the information loss. In this book chapter, we continue this path of reasoning and suggest the Relative Entropy Policy Search (REPS) method. The resulting method differs significantly from previous policy gradient approaches and yields an exact update step. It works well on typical reinforcement learning benchmark problems. We will also present a real-world applications where a robot employs REPS to learn how to return balls in a game of table tennis.

2010-07-08 - Session Bayesian and maximum entropy inference in action: applications [13:30:00-16:45:00]

Bayesian Techniques for Parameter Inference and Knowledge Fusion in Prognostics and Health Management A Case Study

Masoud Rabiei, Mohammad Modarres, Ali Mohammad-Djafari

This paper presents a case study of the application of Bayesian data analysis and parameter inference in the field of Reliability Engineering. Specifically, we address the prominent issue of fatigue crack initiation and growth during the service life of aging aircraft and the corresponding fleet management challenges. Decision making about the remaining useful life of aging airframes is very complicated and should be based upon all information available. Aerospace structures have been traditionally inspected using on-ground nondestructive techniques such as eddy current and visual inspection. New developments in recent years are heading towards continuous on-line monitoring and damage detection in the structure. All this, along with high fidelity crack growth simulation models, are used to estimate the crack size distribution at given hotspots. Such estimations always suffer from both reducible and non-reducible uncertainties. Probability and statistical techniques are used to characterize and represent such uncertainties. In the case study presented here, development of a Meta-model for crack size estimation based on fracture mechanics simulation data and Nondestructive inspection (NDI) findings is described. Bayesian regression technique is used to find the joint distribution of model parameters when dealing with uncertain data. By finding the joint distribution of parameters (rather than just the best estimates), the initial uncertainties in the input simulation and experimental data is preserved in the model that will result in more realistic final risk predictions. Updating the model parameters based on new and relevant findings is also of great interest. A major challenge is to combine the crack size estimations obtained from different sources and present them as a unified risk measure that can be used for the structural health management of the fleet. In this paper, a Hierarchical Bayesian approach is proposed to update the parameters of the Meta-model every time new information becomes available. Posterior distribution of parameters is numerically found via Markov Chain Monte Carlo (MCMC) simulation technique. MCMC simulation has been implemented in WinBUGS software package.

Key words: Knowledge Fusion, Bayesian Regression, Reliability of Airframes, Prognostics and Health Management (PHM)

2010-07-08 - Session Bayesian and maximum entropy inference in action: applications [13:30:00-16:45:00]

ID image characterization by entropic biometric decomposition

Smoaca Andreea, Coltuc Daniela, Fournel Thierry

The paper proposes a statistical-based biometric decomposition for ID image recognition, which is robust to a series of non malicious attacks generated by print/scan operations. Our goal is to label the single face expression by a signature, which is almost invariant to low filtering, noise addition and geometric attacks. Our method is based on Independent Component Analysis (ICA) in a configuration, known in literature as Architecture I. In this architecture, faces are represented as a linear combination of statistically independent basis images which are biometric characteristics as eyes, mouth, nose, eyebrows etc.

These characteristics are first learned on a training set of ID images. Most prominent biometric features are selected by looking for the independent components with maximum local entropy. A biometric label of fixed length is associated to any ID image to be enrolled, after projection on the learned basis, uniform quantization of the obtained coefficients and Gray encoding.

Images from FERET database were used to test the performance of the method. Only individuals with neutral expression and frontal view were selected from this database. Two image sets were used, one for training (300 ID images) and another one for testing (210 ID images). The following attacks: low-pass filtering, addition of Gaussian noise and affine transformations were performed on the test set. Two parameters were tuned: the number of quantization levels and the number of biometrics. The latter one was modified, either by discarding coefficients after Principal Component Analysis in the beginning of FastICA algorithm, or by selecting the most prominent biometric features by applying an entropic criterion. The ID image binary labels were compared by using normalized Hamming distance.

The suggested method which gives a recognition rate of 100% in almost all the mentioned cases inherits the robustness of a global approach. The Hamming distance distributions for similar and dissimilar face images are clearly separated. To further enforce the discrimination power, the method can also be combined with a local approach detecting some biometric characteristic points as in photocomparison of skulls.

The Posterior distribution of the Likelihood Ratio as a measure of evidence

Isabelle Smith, André Ferrari

This paper deals with simple versus composite hypothesis testing under Bayesian and frequentist settings. The Posterior distribution of the Likelihood Ratio (PLR) concept is proposed in [1] for significance testing. The PLR is shown to be equal to 1 minus the p-value in a simple case. The PLR is used in [2] in order to calibrate p-values, Fractional Bayes Factors (BF) and others. Dempster's equivalence result is slightly extended by adding a nuisance parameter in the test.

On the other hand, in [3] the p-values and the posterior probability of the null hypothesis $\mathrm{Pr}(H\_0|x)$ (seen as a Bayesian measure of evidence against the null hypothesis) are shown to be irreconcilable. Actually, as emphasized in [4], $\mathrm{Pr}(H\_0|x)$ is a measure of accuracy of a test, not a measure of evidence in a formal sense because it does not involve the likelihood ratio.

The PLR may give such a measure of evidence and be related to a natural p-value. In this presentation, in a classical invariance framework the PLR with inner threshold 1 will be shown to be equal to 1 minus a p-value where the test statistics is the likelihood, weighted by a Jacobian to account for some volume distorsion effect.

Other analytical properties of the PLR will be proved in more general settings. The minimum of its support is equal to the Generalized Likelihood Ratio and its moments are directly related and to the (Fractional) BF for a proper prior.

Practical issues will also be considered. The PLR can be implemented using a simple Monte Carlo Markov Chain and will be applied to extrasolar planet detection using direct imaging. A hierarchical model of the images will be presented as well as an illustration of the performance of the PLR used as a detector.

2010-07-08 - Poster session – *Methods and Algorithms* [16:45:00-19:00:00]

A Method for Evaluating Tuning Functions of Single Neurons based on Mutual Information Maximization

Lukas Brostek, Thomas Eggert, Seiji Ono, Michael J. Mustari, Ulrich Büttner, Stefan Glasauer

We present a new method for evaluating tuning functions of single neurons based on mutual information maximization. This new approach allows the estimation of neuronal latencies free of model assumptions. It was used to analyze the dependence of neuronal activity in cortical area MSTd on signals related to movement of the eye and retinal image movement. The estimated latencies agree well with results based on other approaches. Compared to commonly used methods, this new approach is very robust to noise and correlations in the input variables and can be applied to every kind of stimuli design.

Gaussian Process Segmentation of Co-Moving Animals

Steven Reece, Richard Mann, Iead Rezek, Stephen Roberts

The analysis of how groups of animals move collectively and how they effectively align their movements is the focus of much current research [1]. A key research question is how individuals transfer information to retain group cohesion whilst achieving both collective and individual goals. The potential existence of distinct subgroups, the members of which are all more closely connected to each other than to other members of the collective, suggest a somewhat partitioned graph of inter-individual interactions. This would have implications for the speed and reliability of information transfer within the group and thus the effectiveness of the group's response to external factors such as the presence of predators and its stability as a cohesive unit.

Gaussian process models have previously been used to describe the flight paths of homing pigeons following habitual homeward routes from familiar release sites [2]. Such models can be similarly utilised to segment groups of co-moving animals into distinct subgroups. Based on the work of [2] we develop a Gaussian process approach to group segmentation. Our approach uses the familiar Variational Bayes (VB) algorithm [3] to cluster the animals into distinct subgroups. The VB algorithm is often used to cluster objects based on their inter-spatial distances. We extend this approach and cluster the animals according to their temporal behaviour. Our distance measure is commensurate with the correlation between individual animal paths. Simply put, animals whose behaviours are highly correlated belong to the same group. Consequently, our approach is able to discern individual groups even when these groups coincide spatially. By applying VB to a Gaussian process mixture model we are able to discern the number of groups within a collective as well as the group compositions.

We are able to exploit Gaussian process interpolation and group animals even when the data is of poor quality. Uncertainty, which is inherent in the data, is treated appropriately and our approach is able to infer groups using sparse data which arises when paths are partially observed or asynchronously observed. We demonstrate the efficacy of our approach using homing pigeon data.

References:

[1] Sumpter, Phil. Trans, Roy. Soc. B, 2006. [2] Mann et al., MaxEnt 2009. [3] Corduneanu et al., AIStats 2001.

Speed-gradient principle for description of transient dynamics in systems obeying maximum entropy principle

Alexander Fradkov, Anton Krivtsov

In [1,2] the Speed-Gradient (SG) principle for determination of nonstationary system dynamics was formulated as follows.

Speed-gradient principle: *Among all possible motions of the system only those are realized for which the input variables u change proportionally to the speed gradient $\nabla\_u\dot{Q}(x,u,t)$ of an appropriate goal functional $Q(x,t)$. If there are constraints imposed on the system motion, then the speed-gradient vector should be projected onto the set of admissible (compatible with constraints) directions.*

According to the SG-principle, to describe system dynamics one needs to introduce the goal function $Q(x,t)$. The choice of $Q(x,t)$ should reflect the tendency of natural behavior to decrease the current value $Q(x(t),t)$. It was suggested in [1,2] to apply the SG-principle for modeling transient (relaxation) dynamics for systems of N identical particles distributed over m cells and satisfying maximum entropy principle. The negative entropy $-S(x)$ was chosen as the goal function. SG-equation for transient dynamics is derived as follows: $dN(t)/dt = A\ln N(t)$, where $N(t)$ is the vector of the cell populations, A is a symmetric matrix with two zero eigenvalues corresponding to mass and energy conservation laws. Stability of the MaxEnt solution was proved analytically by Lyapunov arguments.

In this paper we compare the results obtained by SG-principle with the results of computer experiments obtained by the method of molecular dynamics for $N = 256000$. Particles are interacting via Lennard-Jones potential. According to the computation results, fast change of the distribution function for particles takes place for $t < T\_0$, where T_0 is the period of small oscillations of a particle. Then the distribution function converges slowly to the Maxwell distribution, not depending on initial conditions. Aggregating distribution into three cells (low, middle and high energies), we observed that for $t > T\_0$ the trajectory in the 3D space was moving along straight line, which is close to the prediction made by entropy SG-principle. The value of our experiments is twofold: they confirm SG-principle and show its usefulness for prediction of dynamics in the method of molecular dynamics.

References: [1] A.L. Fradkov. Cybernetical physics. Berlin-Heidelberg: Springer-Verlag,2007. [2] A.L. Fradkov. Entropy, 2008, 10 (4), 757-764 (2009).

2010-07-08 - Poster session – *Methods and Algorithms* [16:45:00-19:00:00]

Generalized Maximum Entropy Principle, Superstatistics And Problem Of Networks Classification

Bahruz Gadjiev

We present the analysis of growing networks with additional mechanisms of growth and networks with competing nodes [1, 2]. Using a generalized maximum entropy principle, we define distribution functions which characterize distributions of variables describing additional mechanisms of growth. Then, using methods of superstatistics, we define topological peculiarities of growing networks with additional mechanisms of growth and networks with competing nodes [3]. It allows us spends topological classification of the considered growing networks. In particular it is shown, that the degree distributions of these networks generally speaking is not defined by a principle networks growth. Really, the heterogenity distributions which is generated by additional mechanisms of growth cardinally changes topological structure of a networks. We present results of numerical simulations which confirm the results of analytical consideration.

Comparative Analysis Of Collaboration Networks

Tatiana Progulova, Bahruz Gadjiev

The network of collaboration represents a bipartite graph. The bipartite graph is a graph where the set of vertices can be divided into two subsets V_1 and V_2 in such a way that each edge of the graph connects vertices from different subsets. Vivid samples of the collaboration networks are movie actor networks and word networks [1]. In the given paper we carry out a comparative analysis of the word network as the collaboration network based on the novel by M. Bulgakov "Master and Margarita", the synonym network of the Russian language as well as the Russian movie actor network. We have constructed one-mode projections of these networks, defined degree distributions for them and have calculated main characteristics.

In the paper a generation algorithm of collaboration networks has been offered which allows to generate networks statistically equivalent to the studied ones. It let us reveal a structural correlation between word network, synonym network and movie actor network.

We show that degree distributions of all analyzable networks are described by the distribution of q-type [2]. The dependence of the topology of collaboration networks on restrictions, which are present with the problem of maximization of Tsallis entropy, is analyzed in detail.

We also compare the obtained results on the synonym network of the Russian language words with those on the synonym network of English and Turkish languages as well as with the results of the Hollywood movie actor network [1, 3].

References: [1] R. Albert and A.-L. Barabasi, Phys. Rev. Lett. **85**, 5234 (2000). [2] C. Tsallis, Braz. J. Phys. **29**, 1 (1999). [3] D. Strori D. et al. arXiv:0709.2476v1.

Key Words: Maximum entropy principle, Bipartite graphs, Collaboration networks, Word networks

Bayesian experiments for numerical solution of differential equations and integrals

Fábio Mendes

There is no general purpose algorithm that computes integrals or solves differential equations after a finite number of steps. If an analytic solution is not obtained, one have to use approximations. Numerical techniques often involve making an estimate of some property of a function (such as its integral or derivative) from a finite number of function evaluations. It seems natural to rephrase this task as a Bayesian experiment: given the known value of some function in a few selected points, what can be said about some other property (e.g., the integral) of this function? Which points in the domain should one select in order to obtain better estimates? In this contribution, Gaussian processes and maximum entropy methods are used to formulate the numerical solution of integrals and differential equations in one or more dimensions. The present framework is flexible, and allows one to design different methods to process different types of information. Classical quadrature rules and the method of finite differences can be recovered as limit cases. Some special-purpose algorithms of very good accuracy are also derived.

A global optimization heuristic for Nonlinear UAV Task Assignment Problem - The Cross-Entropy method

NGUYEN Duc Manh, LE THI Hoai An, PHAM Dinh Tao

The use of unmanned aerial vehicles (UAVs) for various military missions has received growing attention in the last years. Apart from the obvious advantage of not placing human life at risk, the lack of a human pilot enables significant weight savings and lower costs. In this paper, we are interested in task allocation models, in which we seek to assign a set of m UAVs to a set of n tasks in an optimal way. The optimality is quantified using target scores. We can see that this problem is NP-hard. Then, we propose an efficient approach based on the Cross-Entropy (CE) method to tackle it.

References:

[1] P. R. Chandler, M. Pachter, S. R. Rasmussen, and C. Schumacher, *Multiple Task Assignment for a UAV Team*, AIAA Guidance, Navigation, and Control Conference, Monterey, CA, Aug. (2002)

[2] J. B. Cruz, Jr. G. Chen, D. Li, and X. Wang, *Particle Swarm Optimization for Resource Allocation in UAV Cooperative Control*, AIAA Guidance, Navigation, and Control Conference, Providence, RI, Aug. (2004).

[3] J. Le Ny, and E. Feron, *An Approximation Algorithm for the Curvature Constrained Travelling Salesman Problem*, 43rd Annual Allerton Conference on Communications, Control and Computing, Monticello, IL, Sept. (2005).

[4] Rubinstein RY, Kroese D, *The cross-entropy method: a unified approach to combinatorial optimization, Monté Carlo simulation, and machine learning*, Berlin: Springer; 2004.

[5] Costa A, Dafydd O, Kroese D, *Convergence properties of the cross-entropy method for discrete optimization*, Operations Research Letters 2007;35(5):573-80.

Key Words: Task Assignment Problem, Stochastic Programming, Binary Nonlinear Programming.

A Maximum Entropy Approach to Semi-supervised Learning

Ayse Erkan, Yasemin Altun

Maximum entropy (MaxEnt) framework has been studied extensively in supervised learning. Here, the goal is to find a distribution p that maximizes an entropy function while enforcing data constraints so that the expected values of some (pre-defined) features with respect to p match their empirical counterparts approximately. Using different entropy measures, different model spaces for p and different approximation criteria for the data constraints yields a family of discriminative supervised learning methods (e.g., logistic regression, conditional random fields, least squares and boosting). This framework is known as the generalized maximum entropy framework.

Semi-supervised learning (SSL) has emerged in the last decade as a promising field that combines unlabeled data along with labeled data so as to increase the accuracy and robustness of inference algorithms. However, most SSL algorithms to date have had trade-offs, e.g., in terms of scalability or applicability to multi-categorical data. We extend the generalized MaxEnt framework to develop a family of novel SSL algorithms. Extensive empirical evaluation on benchmark data sets that are widely used in the literature demonstrates the validity and competitiveness of the proposed algorithms.

Empirical Bayes Estimation Using Ranked Set Sampling

Mina Aminghafari, Adel Mohammadpour, Hadi Zare

Using complex estimators, increase sample sizes, and changing the sampling planes are three different approaches for reducing randomize errors in the statistical models. Ranked set sampling (rss) can be considered as a third approach. A rss is an independent sample, of order statistics of sub-samples. It is useful when a sub-sample can be easily ranked, but measurements are expensive. The rss method was introduced by McIntyre [1], in the classical statistics and extended by many authors, e.g., see [2]. There are a few works on the Bayesian rss, [3,4,5,6]. It is proved that, the estimators based on the rss have smaller risk with respect to the estimators based on random sampling with the same sample size, under mild conditions.

In the present work, we focus on the role of rss to improve the empirical Bayes estimators. Robbins [7], for the fist time, relax the assumptions of the Bayesian framework to calculate a Bayes estimator when the prior distribution is unknown (non-parametric empirical Bayes) or its parameters are unknown (parametric empirical Bayes), e.g., see [8]. The main difference between structures of the empirical Bayes with the Bayesian framework is the independence of marginal observations. On the other hand, rss sample is a set of independent order statistics which can be generated from an empirical Bayes data framework. The main contribution in this paper, is to show some advantages of the rss sample with respect to the random sample for the empirical Bayes estimators, which are not considered in the former works [3,4,5,6] in the Bayesian rss. The advantages are presented through the analytic results and simulation studies. A few applications are also given.

References:

[1] G. A. McIntyre. A Method for Unbiased Selective Sampling Using Ranked Sets, Australian Journal of Agricultural Research, 3:385-390, (1952).
[2] Z. Chen, et al., Ranked Set Sampling: Theory and Application, Lecture Notes In Statistics, No. 176, Springer (2003).
[3] P. H. Kvam, R. C. Tiwari, Bayes Estimation of a Distribution Function Using Ranked Set Samples, Environmental and Ecological Statistics, 6(1):11-22, (1999).
[4] M. F. Al-Saleh, et al., Bayesian Estimation Using Ranked Set Sampling, Biometrical Journal, 42(4):489–500, (2000).
[5] M. Lavin, The 'Bayesics' of Ranked Set Sampling, Environmental and Ecological Statistics, 6(1):47-57, (1999).
[6] N. Gemayel, Bayesian Nonparametric Models for Ranked Set Sampling, Seminars of Department of Statistics, The Ohio State Univ., (2010).
[7] H. Robbins, An Empirical Bayes Approach to Statistics, Proc. of the 3rd Berkeley Symp. on Math. Stat. and Prob., 1:157-163, (1956).
[8] T. Lwin, J. S. Maritz, Empirical Bayes Methods, Chapman
Hall (1989).

Bayesian Inference Using Maximum Entropy: Analysis And Application

Fadi Dornaika, Fadi Chakik

Recently, Support Vector Machines (SVMs) have been proposed as a powerful tool for solving classification problems. While Maximum Entropy Principle (MEP) aims at determining the probability distribution that explains examples, SVMs aim at determining the best separating surface for the purpose of classification. The main contributions of the paper are as follows. First, we provide a MEP formulation for the task of binary classification and probability distribution estimation based on observables. Second, we establish a link between MEP and SVMs leading to a very interesting property by which the joint probability distribution and the binary classification based on the Maximum Entropy Principle can be solved using the powerful SVM-based optimization tools. Third, we provide an application of the MEP algorithm to breast cancer diagnosis.

Statistical Learning Approach using Maximum Entropy: Image Texture Classification by Energy Function

Fadi Chakik, Fadi Dornaika

The Maximum Entropy Principle (MEP) is based on the maximization of the missing information, given the training data. The information given by examples is used by means of quantities called observables, which are defined on the examples. The observables are chosen a priori, hopefully in a clever way. Although the MEP determines the complete probability distribution, the latter can be used to obtain the most probable output to a given input, for either classification or interpolation tasks. In other words, the MEP can serve as a classification tool. In this paper, we compare the performance of Maximum Entropy and Neural Networks approaches for texture classification in satellite images. The texture classification is data driven, that is, the classifiers are built using a set of labeled examples. Dimensionality reduction is achieved through the use of Energy functions. To this end, Gabor filters are used. We also combine several models via model mixtures to construct a classifier with a low test error.

Logarithm ubiquity in statistics, information, acoustics, ...

M. Maurin

Abstract. Here the central figure is the logarithm function because its numerous properties in mathematics and its interventions in information theory, statistics, acoustics, functional equations, Among all the possibilities of the logaritm and also the exponential, here we introduce a new statistic named the expo-dispersion well suited for logarithmic data as the levels of magnitudes. It is a statistic which characterizes the disperson of data as do variance and entropy, and they have many common behaviours and properties.

Modeling the World by Minimizing Relative Entropy

Antoine van de Ven, Ben A.M. Schouten

Relative entropy in the form of the Kullback-Leibler divergence $KLD\,(w||m)$ can be used as a measure to indicate the difference between the world, represented by w and the internal model of the agent, given by m.

m gives the probability distributions for the predictions, expectations and beliefs by the internal model, and can be generated by a cognitive architecture, like a brain, that can model the world and find causes in the world. It depends on many variables like internal states, memory and more.

w represents the probability distributions for the data coming from the world as sensed and observed by the agent. For example if w represents a virtual world then the function for w would need to include the computation of the virtual world to calculate what would be the sensory input for the agent. It depends on many variables, like location, orientation, time, rules and objects in the world and more.

If something new is experienced and learned it can be seen as an additional constraint for the internal model. Our hypothesis is that there will only be minimal belief updating, and that this happens by minimizing the relative entropy [1,2] between the new model m_t and the old model m_{t-1}. This relative entropy can be interpreted as the surprise and is equal to Bayesian Surprise. In this paper we derive the following general formula for the change of the internal model through time:

$$KLD\,(w||m_t) = KLD\,(w||m_0) - \sum_{i=1}^{t} KLD\,(m_i||m_{i-1})$$

where KLD stands for the Kullback-Leibler divergence. With this formula and the theorem that any KLD is always positive (or zero) it is proven that the internal model m converges to w, and thus builds a model of the world.

References:

[1] Antoine van de Ven. A minimum relative entropy principle for the brain. In Proceedings of the Ninth International Conference on Epigenetic Robotics. Lund University Cognitive Studies, 145, 2009.

[2] Antoine van de Ven and Ben A.M. Schouten, A minimum relative entropy principle for AGI, Proceedings of the Third Conference on Artificial General Intelligence, Advances in Intelligent Systems Research, March 2010, doi:10.2991/agi.2010.26

Key Words: Relative Entropy, Kullback-Leibler divergence, Bayesian Surprise

Spectral Design in Markov Random Fields

Jiao Wang, Jean-Baptiste Thibault, Zhou Yu, Ken Sauer, Charles Bouman

Design of Markov random fields (MRF) as image models has largely been focused on the scalar function penalizing local pixel differences, often with the goal of better preserving discontinuities than quadratic regularization. In this paper we study the spatial design of weighting coefficients in higher-order MRFs, with more focus on preservation of specific frequency components. The design may be taken directly from spectral models of the signal to be recovered, or with the forward and inverse operators included, for the system response of the entire process. We apply the resulting designs with edge-preserving penalties and find that the spectral methods help alleviate inherent bias of the maximum a posteriori probability estimator.

Determination and Estimation of Generalized Entropy Rates for Countable Markov Chains

G. Ciuperca, Valérie Girardin, L. Lhote

We study entropy rates of random sequences for general entropy functionals including the classical Shannon and Rényi entropies and the more recent Tsallis and Sharma-Mittal ones.

We obtain an explicit formula for the entropy rate for a large class of entropy functionals, as soon as the process satisfies a regularity property. Countable Markov chains are proven to satisfy this property, under simple explicite conditions on the transition probabilities. Then, we focus on the estimation of the generalized marginal entropy and entropy rate for parametric Markov chains. Estimators with good asymptotic properties are built through a plug-in procedure using a maximum likelihood estimation of the parameter.

Geometric and Topological Invariants of the Hypothesis Space

Carlos Rodriguez

The form and the shape of a hypothesis space imposes natural objective constraints to any inferential process. This contribution summarizes what is currently known and the mathematics that are thought to be needed for new developments in this area. For example, it is well known that the quality of best possible estimators deteriorates with increasing volume, dimension and curvature of the hypothesis space. It is also known that regular statistical parametric models are finite dimensional Riemannian manifolds admitting a family of dual affine connections. Fisher information is the metric induced on the hypothesis space by the Hellinger distance. Nonparametric models are infinite dimensional manifolds. Global negative curvature implies asymptotic inadmissibility of uniform priors. When there is uncertainty about the model and the prior, entropic methods are more robust than standard Bayesian inference. The presence of some types of singularities allow the existence of faster than normal estimators . . ., etc.

The large number of fundamental statistical concepts with geometric and topological content suggest to try to look at Riemannian Geometry, Algebraic Geometry, K-theory, Algebraic Topology, Knot-theory and other branches of current mathematics, not as empty esoteric abstractions but as allies for statistical inference.

Geometry of Covariance Matrices and Computation of Median

Le Yang, Marc Arnaudon, Frédéric Barbaresco

In this paper, we consider the manifold of covariance matrices of order n parametrized by reflection coefficients which are derived from Levinson's recursion of autoregressive model. The explicit expression of the reparametrization and its inverse are obtained. With the Riemannian metric given by the Hessian of a Kähler potential, we show that the manifold is in fact a Cartan-Hadamard manifold with lower sectional curvature bound -4. The explicit expressions of geodesics are also obtained. After that we introduce the notion of Riemannian median of points lying on a Riemannian manifold and give a simple algorithm to compute it. Finally, some simulation examples are given to illustrate the applications of the median method to radar signal processing.

Index of authors

Ait-El-Fquih, Boujemaa, 54
Altun, Yasemin, 69, 80
Aminghafari, Mina, 81
Andreea, Smoaca, 71
Antoranz, Carlos, 43
Arnaudon, Marc, 89

Bahrenyi, Newshaw, 29
Baraov, Arthur, 22
Barbaresco, Frédéric, 89
Beauval, Céline, 49
Bercher, Jean-François, 42
Bernardo, Gustavo, 27
Bessière, Pierre, 18, 19, 21
Betancourt, Michael, 37
Bityukov, Sergey, 48, 65
Bois, Frederic, 51
Bontekoe, Romke, 16
Bouman, Charles, 86
Bretthorst, Larry, 13
Brostek, Lukas, 73
Buzuloiu, Vasile, 36
Büttner, Ulrich, 73

Caticha, Ariel, 30, 41
Center, Julian, 20
Chakik, Fadi, 82, 83
Chandrashaker, Akhila, 47
Chaouchi, Chahinez, 45
Cho, M., 57
Ciuperca, G, 87
Colas, Francis, 21
Collet, Christophe, 58
Collinet, Claudio, 47

Daniela, Coltuc, 71
De Smedt, Sebastiaan, 28
Dehesa, Jesus, 39, 40
Deneve, Sophie, 17
Derdeyn, Colin, 13
Deville, Yannick, 45
Dey, Dipak, 57
Diard, Julien, 18
Diniz, Marcio, 25
Dobigeon, Nicolas, 59
Dornaika, Fadi, 82, 83

Earle, Keith, 68
Eches, Olivier, 59
Eggert, Thomas, 73
Erkan, Ayse, 80
Esseghir, Moez, 53

Fava, Eugenio, 47
Ferrari, André, 72
Fitzenz, Delphine, 49
Fradkov, Alexander, 75

Gader, Paul, 60
Gadjiev, Bahruz, 76, 77
Gilet, Estelle, 18
Girard, Benoît, 21
Girardin, Valérie, 87
Glasauer, Stefan, 73
Gori, Silvio, 64
Goyal, Philip, 29
Gunawan, Rudy, 51, 66

Holmes, Dawn, 62
Hossaini Shojaei, Sayed Reza, 31
Hosseini, Shahram, 45
Huang, Ding, 56

Inhasz, Rafael, 34
Islam, Md. Saiful, 55

Jalobeanu, André, 58
Javidi, Bahram, 57
Johnson, David, 30
Jung, M., 57

Kalaidzidis, Yannis, 47
Kester, Do, 16, 24
Knuth, K. H., 35
Knuth, Kevin, 11, 20, 29
Kostecki, Ryszard Pawel, 14
Krasnikov, Nikolai, 48, 65
Krivtsov, Anton, 75
Krylov, Vladimir A., 61

Lauretto, Marcelo, 27
Le Goff, Boris, 49
Le Thi, Hoai An, 79
Lee, John, 13

Lemmens, Lucien, 28
Lhote, L, 87

Malakar, N. K., 35
Mann, Richard, 74
Mansouri, Majdi, 53
Manzano, D., 38
Maurin, M., 84
Mazer, Emmanuel, 67
Mendes, Fábio, 78
Modarres, Mohammad, 70
Mohammad-Djafari, Ali, 52, 54, 63, 70
Mohammadian, Gholam Reza, 31
Mohammadpour, Adel, 31, 32, 81
Moser, Gabriele, 61
Muelling, Katharina, 69
Mustari, Michael J., 73

Nadaraja, Saralees, 65
Nassiri, Vahid, 31, 32
Nguyen, Duc Manh, 79

Ono, Seiji, 73

Palluel-Germain, Richard, 18
Pereira, Carlos, 25
Peters, Jan, 69
Petremand, Matthieu, 58
Pham, Dinh Tao, 79
Polpo, Adriano, 25
Popescu, George, 36
Popescu, Sorina, 36
Portesi, Mariela, 39
Pougaza, Doriano-Boris, 63
Progulova, Tatiana, 77

Rabiei, Masoud, 70
Reece, Steven, 74
Reginatto, Marcel, 46
Regnault, Philippe, 26
Rezek, Iead, 74
Richard, Cedric, 53
Rink, Jochen, 47
Roberts, Stephen, 74
Rodriguez, Carlos, 88
Rodriguez-Perez, Daniel, 43

Samson, Roeland, 28
Sanchez-Dehesa, Jesus, 38
Sanchez-Moreno, Pablo, 39, 40
Sauer, Ken, 86
Schneider, David, 68
Schouten, Ben A.M., 85
Senocak, Inanc, 50
Serpico, Sebastiano B., 61
Shimony, Joshua, 13
Shokripour, Mona, 32
Sivaloganathan, Siv, 66
Skilling, John, 12, 33
Smirnova, Vera, 48, 65
Smith, Isabelle, 72
Snoussi, Hichem, 53
Sotolongo-Costa, Oscar, 43
Sotolongo-Grau, Oscar, 43
Stern, Julio, 25, 27, 34
Stötter, Martin, 47
Sukkarieh, Jana, 15
Synnaeve, Gabriel, 19
Sánchez-Moreno, P., 38

Tenti, Giuseppe, 66
Thibault, Jean-Baptiste, 86
Thierry, Fournel, 71
Tourneret, Jean-Yves, 59

Urniezius, Renaldas, 44

van de Ven, Antoine, 85
van Erp, Noel, 23
van Gelder, Pieter, 23
von Toussaint, Udo, 64

Wang, Jiao, 86

Yang, Le, 89
Yu, Zhou, 86
Yáñez, R.J., 38

Zare, Alina, 60
Zare, Hadi, 81
Zerial, Marino, 47
Zerubia, Josiane, 61
Zhu, Sha, 52
Zimbal, Andreas, 46
Zozor, Steeve, 39, 40