ETICS 2019

École Thématique sur les Incertitudes en Calcul Scientifique
Research School on Uncertainty in Scientific Computing

September, 22-27, Fréjus, France - https://www.caes.cnrs.fr/sejours/la-villa-clythia

Objectives

The goal of this school is to develop the skills of researchers and engineers in the domain of uncertainty management of computer codes. Some of the lectures will be followed by practical computer works. Collaborative works, round tables and poster sessions will promote exchanges between participants. The prerequisites to possess are the mathematical bases of the uncertainty quantification science.

Lecturers

Dr. Aurélien Bellet (INRIA Lille - Nord Europe, France) - Similarity and distance metric learning
Prof. Bernard Bercu (Université de Bordeaux, France) - Asymptotic behavior of stochastic algorithms with statistical applications
Prof. Jean-Michel Marin (Université de Montpellier, France) – Some computational tools for Bayesian inference
Prof. Youssef Marzouk (Massachusetts Institute of Technology, Cambridge, MA, USA)

Organization

CEA DAM

EDF R&D

CMLA
ENS Paris
Saclay

Under the scientific labeling of the GdR MASCOT-NUM

Contacts:	Guillaume Perrin	guillaume.perrin2 [at] cea.fr
	Bertrand Iooss	bertrand.iooss [at] edf.fr

Registration: https://www.azur-colloque.fr/DR04/inscription/inscription/132/fr
Registration fees 900€ including accommodation and meals

Schedule

Sunday, September, 22th: Travel (at the charge of each participant) to Fréjus

Monday, September, 23th:

9:00 - 9:30	Opening and Welcome speech	Bertrand Iooss Guillaume Perrin
9:30 – 12:30	Computational methods for Bayesian inference	Jean-Michel Marin
14:00 – 17:30	Practical session	Jean-Michel Marin

Tuesday, September, 24th:

9:00 – 12:30	Asymptotic behavior of stochastic algorithms with statistical applications	Bernard Bercu
14:00 – 17 :00	Practical session	Bernard Bercu
17 :30-19 :00	Poster session	Participants

Wednesday, September, 25th:

9:00 – 12:30	Transport methods in Bayesian computation (1/2)	Youssef Marzouk
14:00 – 20 :00	Social event in calanques rouges or free afternoon	

Thursday, September, 26th:

9:00 – 12:30	Similarity and distance metric learning	Aurélien Bellet
14:00 – 17:30	Transport methods in Bayesian computation (2/2)	Youssef Marzouk

Friday, September, 27th:

9:00 – 12:30	Practical session	Aurélien Bellet

ABSTRACTS

Jean-Michel Marin - Some computational tools for Bayesian inference

Quantifying uncertainty in statistical inference is a real issue. Bayesian techniques offer an elegant and efficient probabilistic way to resolve that question. The counterpart is that Bayesian strategies are very difficult to enforce and require the use of sophisticated simulation algorithms. After a recap on the Bayesian paradigm, we will introduce some these algorithms with an emphasize on recent techniques used in the case of intractable likelihood problems. Indeed, statistical models and data structures get increasingly complex, managing the likelihood function becomes a more and more frequent issue. We now face many realistic fully parametric situations where the likelihood function cannot be computed in a reasonable time or simply is unavailable.
As a result, while the corresponding parametric model is well-defined, standard solutions based on the likelihood function like Bayesian or maximum likelihood analyses are prohibitive to implement. To bypass this hurdle, the last decade witnessed different inferential strategies, among which composite likelihoods, indirect inference, GMMs and likelihood-free methods such as Approximate Bayesian Computation (ABC). We will focus on the latest, and consider versions of ABC that consists in using machine learning algorithms on reference tables, tables that are simulated from the Bayesian model and used as learning set.

Part 1 Recap on the Bayesian paradigm Prior distributions Bayes estimates Credible intervals Bayesian discrimination between models Difficulties with the Bayesian paradigm

Part 2 Monte Carlo and Markov chain Monte Carlo methods Standard Monte Carlo Methods Importance Sampling strategies Reminders and Complements on the Markov Chains Convergence of Markov chains Metropolis-Hastings algorithm The Gibbs sampler

Part 3 Approximate Bayesian Computation methods Methodological aspects of ABC ABC random forests

Part 4 Use of the R libraries rjags and abcrf

Youssef Marzouk - Transport methods in Bayesian computation

Bayesian inference provides a natural framework for quantifying uncertainty in parameter estimates and model predictions, and for combining heterogeneous sources of information. Characterizing the results of Bayesian inference---by simulating from the posterior distribution---often proceeds via Markov chain Monte Carlo or sequential Monte Carlo sampling, but remains computationally challenging for complex posteriors and large-scale models.
This lecture will describe a broad framework for using transport in Bayesian computation. This framework seeks deterministic couplings of the posterior measure with a tractable "reference" measure (e.g., a standard Gaussian). Such couplings are induced by transport maps, and enable direct simulation from the desired measure simply by evaluating the transport map at samples from the reference. Approximate transports can also be used to "precondition" and accelerate standard Monte Carlo schemes. Within this framework, one can describe many useful notions of low-dimensional structure associated with inference: for instance, sparse or decomposable transports underpin modeling and computation with non-Gaussian Markov random fields, and low-rank transports arise frequently in inverse problems.
We will also describe recent work specializing transport maps to the problem of nonlinear filtering in high-dimensional state-space models. The idea is to transform a forecast ensemble into samples from the current filtering distribution via a sequence of nonlinear transport maps, computed via convex optimization. Construction of the maps is regularized by leveraging potential structure in the filtering problem---e.g., decay of correlations, approximate conditional independence, and local likelihoods---thus extending notions of localization to nonlinear updates. The proposed framework can be understood as a non-Gaussian generalization of the ensemble Kalman filter.

This is joint work with Alessio Spantini, Daniele Bigoni, Ricardo Baptista, and Matthew Parno.

Aurélien Bellet - Similarity and distance metric learning

Similarity between objects plays an important role in both human cognitive processes and artificial systems for recognition and categorization. How to appropriately measure such similarities for a given task is crucial to the performance of many machine learning, pattern recognition and data mining methods. This session is devoted to metric learning, an area of machine learning which aims at automatically learning similarity and distance functions from data. In the lecture, I will review the main ideas, formulations, algorithms and statistical guarantees of metric learning. In the practical session in Python, we will apply metric learning to some synthetic and real datasets. We will use metric-learn, the reference toolbox for metric learning in Python.

Bernard Bercu – Asymptotic behavior of stochastic algorithms with statistical applications

This course is dedicated to the analysis of the asymptotic behavior of stochastic algorithms. We shall focus our attention on the well-know Robbins-Monro and Kiefer-Wolfowitz algorithms. We establish the almost sure convergence and we investigate the asymptotic normality of these two stochastic algorithms.
The proofs rely on the strong law of large numbers and the central limit theorem for martingales. Four statistical applications are also provided. The first one concerns the parametric estimation of quantiles and superquantiles. The second one is devoted to the non-parametric estimation of probability density functions. The third one deals with the semi-parametric estimation in shape invariant models. The last one is devoted to stochastic approximation of entropically regularized Wasserstein distances between two probability measures.