IN PARTNERSHIP WITH:
CNRS

Université de Bordeaux

Activity Report 2015

Project-Team CQFD

Quality control and dynamic reliability

IN COLLABORATION WITH: Institut de Mathématiques de Bordeaux (IMB)

RESEARCH CENTER
Bordeaux - Sud-Ouest

THEME
Stochastic approaches

Table of contents

Project-Team CQFD

Creation of the Project-Team: 2009 January 01

Keywords:

Computer Science and Digital Science:

1.3. - Distributed Systems
3.3. - Data and knowledge analysis
6.2.3. - Probabilistic methods
6.2.4. - Statistical methods
6.2.6. - Optimization
6.4.2. - Stochastic control

Other Research Topics and Application Domains:

6.5. - Information systems

1. Members

Research Scientist

Jonatha Anselmi [Inria, Researcher]

Faculty Members

Francois Dufour [Team leader, INP Bordeaux, Professor, HdR]
Marie Chavent [Univ. Bordeaux, Associate Professor, HdR]
Alexandre Genadot [Univ. Bordeaux, Associate Professor, from Sep 2015]
Pierrick Legrand [Univ. Bordeaux, Associate Professor]
Jerome Saracco [INP Bordeaux, Professor, HdR]
Huilong Zhang [Univ. Bordeaux, Associate Professor]

PhD Students

Isabelle Charlier [Université Libre de Bruxelles, until Nov 2015]
Alizée Geeraert [Thales]
Amaury Labenne [IRSTEA, until Oct 2015]
Shuxian Li [Univ. Bordeaux, until Nov 2015]
Christophe Nivot [Inria, granted by Conseil Régional d'Aquitaine]
Jessica Sodjo [Inria]
Adrien Todeschini [Inria, granted by EVOLLIS SAS]

Post-Doctoral Fellow

Solveig Badillo [Univ. Bordeaux]

Visiting Scientists

Oswaldo Luiz Do Valle Costa [Escola Politécnica da Universidade de São Paulo, Brazil]
Luis Herrera Lezama [INP Bordeaux, from Sep 2015]
Alexei Piunovskiy [University of Liverpool, until Jun 2015]

Administrative Assistant

Sabrina Blondel-Duthil [Inria]

Other

Benoite de Saporta [Univ. Montpellier II, Professor, HdR]

2. Overall Objectives

2.1. Presentation

The core component of our scientific agenda focuses on the development of statistical and probabilistic methods for the modeling and the optimization of complex systems. These systems require mathematical representations which are in essence dynamic and stochastic with discrete and/or continuous variables. This increasing complexity poses genuine scientific challenges that can be addressed through complementary approaches and methodologies:

- Modeling: design and analysis of realistic and tractable models for such complex real-life systems and various probabilistic phenomena;
- Estimation: developing theoretical and computational procedures in order to estimate and evaluate the parameters and the performance of the system;
- Optimization: developing theoretical and numerical control tools to optimize the performance of complex systems such as computer systems and communication networks.

3. Research Program

3.1. Introduction

The scientific objectives of the team are to provide mathematical tools for modeling and optimization of complex systems. These systems require mathematical representations which are in essence dynamic, multi-model and stochastic. This increasing complexity poses genuine scientific challenges in the domain of modeling and optimization. More precisely, our research activities are focused on stochastic optimization and (parametric, semi-parametric, multidimensional) statistics which are complementary and interlinked topics. It is essential to develop simultaneously statistical methods for the estimation and control methods for the optimization of the models.

3.2. Main research topics

- Stochastic modeling: Markov chain, Piecewise Deterministic Markov Processes (PDMP), Markov Decision Processes (MDP).
  The mathematical representation of complex systems is a preliminary step to our final goal corresponding to the optimization of its performance. For example, in order to optimize the predictive maintenance of a system, it is necessary to choose the adequate model for its representation. The step of modeling is crucial before any estimation or computation of quantities related to its optimization. For this we have to represent all the different regimes of the system and the behavior of the physical variables under each of these regimes. Moreover, we must also select the dynamic variables which have a potential effect on the physical variable and the quantities of interest. The team CQFD works on the theory of Piecewise Deterministic Markov Processes (PDMP's) and on Markov Decision Processes (MDP's). These two classes of systems form general families of controlled stochastic processes suitable for the modeling of sequential decision-making problems in the continuous-time (PDMPs) and discrete-time (MDP's) context. They appear in many fields such as engineering, computer science, economics, operations research and constitute powerful class of processes for the modeling of complex system.

- Estimation methods: estimation for PDMP; estimation in non- and semi parametric regression modeling.
  To the best of our knowledge, there does not exist any general theory for the problems of estimating parameters of PDMPs although there already exist a large number of tools for sub-classes of PDMPs such as point processes and marked point processes. However, to fill the gap between these specific models and the general class of PDMPs, new theoretical and mathematical developments will be on the agenda of the whole team. In the framework of non-parametric regression or quantile regression, we focus on kernel estimators or kernel local linear estimators for complete data or censored data. New strategies for estimating semi-parametric models via recursive estimation procedures have also received an increasing interest recently. The advantage of the recursive estimation approach is to take into account the successive arrivals of the information and to refine, step after step, the implemented estimation algorithms. These recursive methods do require restarting calculation of parameter estimation from scratch when new data are added to the base. The idea is to use only the previous estimations and the new data to refresh the estimation. The gain in time could be very interesting and there are many applications of such approaches.
- Dimension reduction: dimension-reduction via SIR and related methods, dimension-reduction via multidimensional and classification methods.
  Most of the dimension reduction approaches seek for lower dimensional subspaces minimizing the loss of some statistical information. This can be achieved in modeling framework or in exploratory data analysis context.
  In modeling framework we focus our attention on semi-parametric models in order to conjugate the advantages of parametric and nonparametric modeling. On the one hand, the parametric part of the model allows a suitable interpretation for the user. On the other hand, the functional part of the model offers a lot of flexibility. In this project, we are especially interested in the semi-parametric regression model $Y = f(X'\theta) + \varepsilon$, the unknown parameter θ belongs to $\mathbb{R}^p$ for a single index model, or is such that $\theta = [\theta_1, \cdots, \theta_d]$ (where each θ_k belongs to $\mathbb{R}^p$ and $d \leq p$ for a multiple indices model), the noise ε is a random error with unknown distribution, and the link function f is an unknown real valued function. Another way to see this model is the following: the variables X and Y are independent given $X'\theta$. In our semi-parametric framework, the main objectives are to estimate the parametric part θ as well as the nonparametric part which can be the link function f, the conditional distribution function of Y given X or the conditional quantile q_α. In order to estimate the dimension reduction parameter θ we focus on the Sliced Inverse Regression (SIR) method which has been introduced by Li [44] and Duan and Li [42]
  Methods of dimension reduction are also important tools in the field of data analysis, data mining and machine learning.They provide a way to understand and visualize the structure of complex data sets.Traditional methods among others are principal component analysis for quantitative variables or multiple component analysis for qualitative variables. New techniques have also been proposed to address these challenging tasks involving many irrelevant and redundant variables and often comparably few observation units. In this context, we focus on the problem of synthetic variables construction, whose goals include increasing the predictor performance and building more compact variables subsets. Clustering of variables is used for feature construction. The idea is to replace a group of "similar" variables by a cluster centroid, which becomes a feature. The most popular algorithms include K-means and hierarchical clustering. For a review, see, e.g., the textbook of Duda [43]
- Stochastic optimal control: optimal stopping, impulse control, continuous control, linear programming.
  The first objective is to focus on the development of computational methods.

- In the continuous-time context, stochastic control theory has from the numerical point of view, been mainly concerned with Stochastic Differential Equations (SDEs in short). From the practical and theoretical point of view, the numerical developments for this class of processes are extensive and largely complete. It capitalizes on the connection between SDEs and second order partial differential equations (PDEs in short) and the fact that the properties of the latter equations are very well understood. It is, however, hard to deny that the development of computational methods for the control of PDMPs has received little attention. One of the main reasons is that the role played by the familiar PDEs in the diffusion models is here played by certain systems of integro-differential equations for which there is not (and cannot be) a unified theory such as for PDEs as emphasized by M.H.A. Davis in his book. To the best knowledge of the team, there is only one attempt to tackle this difficult problem by O.L.V. Costa and M.H.A. Davis. The originality of our project consists in studying this unexplored area. It is very important to stress the fact that these numerical developments will give rise to a lot of theoretical issues such as type of approximations, convergence results, rates of convergence,....
- Theory for MDP's has reached a rather high degree of maturity, although the classical tools such as value iteration, policy iteration and linear programming, and their various extensions, are not applicable in practice. We believe that the theoretical progress of MDP's must be in parallel with the corresponding numerical developments. Therefore, solving MDP's numerically is an awkward and important problem both from the theoretical and practical point of view. In order to meet this challenge, the fields of neural networks, neuro-dynamic programming and approximate dynamic programming became recently an active area of research. Such methods found their roots in heuristic approaches, but theoretical results for convergence results are mainly obtained in the context of finite MDP's. Hence, an ambitious challenge is to investigate such numerical problems but for models with general state and action spaces. Our motivation is to develop theoretically consistent computational approaches for approximating optimal value functions and finding optimal policies.
- An effort has been devoted to the development of efficient computational methods in the setting of communication networks. These are complex dynamical systems composed of several interacting nodes that exhibit important congestion phenomena as their level of interaction grows. The dynamics of such systems are affected by the randomness of their underlying events (e.g., arrivals of http requests to a web-server) and are described stochastically in terms of queueing network models. These are mathematical tools that allow one to predict the performance achievable by the system, to optimize the network configuration, to perform capacity-planning studies, etc. These objectives are usually difficult to achieve without a mathematical model because Internet systems are huge in size. However, because of the exponential growth of their state spaces, an exact analysis of queueing network models is generally difficult to obtain. Given this complexity, we have developed analyses in some limiting regime of practical interest (e.g., systems size grows to infinity). This approach is helpful to obtain a simpler mathematical description of the system under investigation, which leads to the direct definition of efficient, though approximate, computational methods and also allows to investigate other aspects such as Nash equilibria.

The second objective of the team is to study some theoretical aspects related to MDPs such as convex analytical methods and singular perturbation. Analysis of various problems arising in MDPs leads to a large variety of interesting mathematical problems.

4. Application Domains

4.1. Dependability and safety

Our abilities in probability and statistics apply naturally to industry in particular in studies of dependability and safety.

An illustrative example which gathers several topics of team is a collaboration started in September 2013 with Airbus Defence & Space. The goal of this project is the optimization of the assembly line of the future European launcher, taking into account several kinds of economical and technical constraints. We have started with a simplified model with five components to be assembled in workshops liable to breakdowns. We have modeled the problem using the Markov Decision Processes (MDP) framework and built a simulator of the process in order to run a simulation-based optimization procedure.

A second example concerns the optimization of the maintenance of a on board system equipped with a HUMS (Health Unit Monitoring Systems) in collaboration with THALES Optronique. The physical system under consideration is modeled by a piecewise deterministic Markov process. In the context of impulse control, we propose a dynamic maintenance policy, adapted to the state of the system and taking into account both random failures and those related to the degradation phenomenon.

However the spectrum of applications of the topics of the team is larger and may concern many other fields. Indeed non parametric and semi-parametric regression methods can be used in biometry, econometrics or engineering for instance. Gene selection from microarray data and text categorization are two typical application domains of dimension reduction among others. We had for instance the opportunity via the scientific program PRIMEQUAL to work on air quality data and to use dimension reduction techniques as principal component analysis (PCA) or positive matrix factorization (PMF) for pollution sources identification and quantization.

5. Highlights of the Year

5.1. Highlights of the Year

Publication of the book: *Numerical methods for simulation and optimization of piecewise deterministic Markov processes* written by Benoîte De Saporta; Francois Dufour and Huilong Zhang in Mathematics & Statistics, Wiley, 298 pages, 2015.

6. New Software and Platforms

6.1. Package PCAmixdata

FUNCTIONAL DESCRIPTION

Mixed data type arise when observations are described by a mixture of numerical and categorical variables. The R package PCAmixdata extends standard multivariate analysis methods to incorporate this type of data. The key techniques included in the package are PCAmix (PCA of a mixture of numerical and categorical variables), PCArot (rotation in PCAmix) and MFAmix (multiple factor analysis with mixed data within a dataset). The MFAmix procedure handles a mixture of numerical and categorical variables within a group - something which was not possible in the standard MFA procedure. We also included techniques to project new observations onto the principal components of the three methods in the new version of the package.

- Participants: Marie Chavent, Amaury Labenne, Jérome Saracco
- Contact: Marie Chavent
- URL: https://cran.r-project.org/web/packages/PCAmixdata/index.html

6.2. Package divclust

FUNCTIONAL DESCRIPTION DIVCLUS-T is a divisive hierarchical clustering algorithm based on a monothetic bipartitional approach allowing the dendrogram of the hierarchy to be read as a decision tree. It is designed for numerical, categorical (ordered or not) or mixed data. Like the Ward agglomerative hierarchical clustering algorithm and the k-means partitioning algorithm, it is based on the minimization of the inertia criterion. However, it provides a simple and natural monothetic interpretation of the clusters. Indeed, each cluster is decribed by set of binary questions. The inertia criterion is calculated on all the principal components of PCAmix (and then on standardized data in the numerical case).

- Participants: Marie Chavent, Marc Fuentes
- Contact: Marie Chavent
- URL: https://github.com/chavent/divclust

6.3. Package ClustGeo

FUNCTIONAL DESCRIPTION This R package is dedicated to the clustering of objects with geographical positions. The clustering method implemented in this package allows the geographical constraints of proximity to be taken into account within the ascendant hierarchical clustering.

- Marie Chavent, Amaury Labenne, Vanessa Kuentz, Jérome Saracco
- Contact: Amaury Labenne
- URL: https://cran.r-project.org/web/packages/ClustGeo/index.html

6.4. Package QuantifQuantile

FUNCTIONAL DESCRIPTION This R package is dedicated to the estimation of conditional quantiles using optimal quantization. It allows the construction of an optimal grid of N quantizers, the estimation of conditional quantiles and the data driven selection of the size N of the grid. Graphical illustrations are available for the selection of N and of resulting estimated curves or surfaces when the dimension of the covariate is one or two.

- Isabelle Charlier, Jérôme Saracco
- Contact: Isabelle Charlier
- URL: https://cran.r-project.org/web/packages/QuantifQuantile/index.html

6.5. Biips: Software for Bayesian Inference with Interacting Particle Systems

FUNCTIONAL DESCRIPTION

Biips is a software platform for automatic Bayesian inference with interacting particle systems. Biips allows users to define their statistical model in the probabilistic programming BUGS language, as well as to add custom functions or samplers within this language. Then it runs sequential Monte Carlo based algorithms (particle filters, particle independent Metropolis-Hastings, particle marginal Metropolis-Hastings) in a black-box manner so that to approximate the posterior distribution of interest as well as the marginal likelihood. The software is developed in C++ with interfaces with the softwares R, Matlab and Octave.

- Participants: François Caron, Adrien Todeschini and Pierrick Legrand
- Contact: Adrien Todeschini
- URL: http://biips.gforge.inria.fr

6.6. VCN: Software for analysis of VCN

FUNCTIONAL DESCRIPTION

VCN is a software for the analysis of the vigilance of the patient based on the analysis of the EEG signals. The code is written in Matlab and provides an interface easy to use for someone without informatics skills.

- Participants: Pierrick Legrand, Julien Clauzel, Laurent Vezard, Charlotte Rodriguez, Borjan Geshkovski.
- Contact: Pierrick Legrand

6.7. EMGView: Software for visualisation and time-frequency analysis of bio signals

FUNCTIONAL DESCRIPTION

EMGView is a software for the visualisation and the analysis of bio-signals. The code is written in Matlab and provides an interface easy to use for someone without informatics skills.

- Participants: Luis Herrera, Eric Grivel, Pierrick Legrand, Gregory Barriere
- Contact: Pierrick Legrand

7. New Results

7.1. Control of parallel non-observable queues: asymptotic equivalence and optimality of periodic policies

The following result has been obtained by J. Anselmi (Inria CQFD), T. Nesti and B. Gaujal.

We consider a queueing system composed of a dispatcher that routes deterministically jobs to a set of non-observable queues working in parallel. In this setting, the fundamental problem is which policy should the dispatcher implement to minimize the stationary mean waiting time of the incoming jobs. We present a structural property that holds in the classic scaling of the system where the network demand (arrival rate of jobs) grows proportionally with the number of queues. Assume that each queue of type r is replicated k times and consider the set of policies that are periodic with period $k\sum_r p_r$ and such that exactly p_r jobs are sent in a period to each queue of type r. When $k \to \infty$, our main result shows that all the policies in this set are equivalent, in the sense that they yield the same mean stationary waiting time, and optimal, in the sense that no other policy having the same aggregate arrival rate to all queues of a given type can do better in minimizing the stationary mean waiting time. This property holds in a strong probabilistic sense. Furthermore, the limiting mean waiting time achieved by our policies is a convex function of the arrival rate in each queue, which facilitates the development of a further optimization aimed at solving the fundamental problem above for large systems.

7.2. Decentralized Proportional Load Balancing

The following result has been obtained by J. Anselmi (Inria CQFD), and N. Walton.

Load balancing is a powerful technique commonly used in communication and computer networks to improve system performance, robustness and fairness. In this paper, we consider a general model capturing the performance of communication and computer networks, and on top of it we propose a decentralized algorithm for balancing load among multiple network paths. The proposed algorithm is inspired by the modus operandi of the processor-sharing queue and on each network entry point operates as follows: every time a unit of load completes its service on a path, it increases by one unit the load of that path and decreases by one unit the load of a path selected at random with probability proportional to the amount of load on each of the available paths. We develop a dynamical system to argue that our load-balancer achieves a desirable network-wide utility optimization.

A paper has been accepted for publication in the SIAM Journal of Applied Mathematics.

7.3. Conditional quantile estimation through optimal quantization

The following result has been obtained by Isabelle Charlier (CQFD member), Davy Paindaveine, and Jérôme Saracco (CQFD member)

We use quantization to construct a nonparametric estimator of conditional quantiles of a scalar response Y given a d-dimensional vector of covariates X. First we focus on the population level and show how optimal quantization of X, which consists in discretizing X by projecting it on an appropriate grid of N points, allows to approximate conditional quantiles of Y given X. We show that this approximation is arbitrarily good as N goes to infinity and provide a rate of convergence for the approximation error. Then we turn to the sample case and define an estimator of conditional quantiles based on quantization ideas. We prove that this estimator is consistent for its fixed-N population counterpart. The results are illustrated on a numerical example. Dominance of our estimators over local constant/linear ones and nearest neighbor ones is demonstrated through extensive simulations in the companion paper Charlier et al. (2014).

7.4. A linear programming formulation for constrained discounted continuous control for piecewise deterministic Markov processes

The following result has been obtained by Oswaldo Costa and François Dufour (CQFD member).

This work deals with the constrained discounted control of piecewise deterministic Markov process (PDMPs) in general Borel spaces. The control variable acts on the jump rate and transition measure, and the goal is to minimize the total expected discounted cost, composed of positive running and boundary costs, while satisfying some constraints also in this form. The basic idea is, by using the special features of the PDMPs, to re-write the problem via an embedded discrete-time Markov chain associated to the PDMP and re-formulate the problem as an infinite dimensional linear programming (LP) problem, via the occupation measures associated to the discrete-time process. It is important to stress however that our new discrete-time problem is not in the same framework of a general constrained discrete-time Markov Decision Process and, due to that, some conditions are required to get the equivalence between the continuous-time problem and the LP formulation. We provide in the sequel sufficient conditions for the solvability of the associated LP problem. We provide some examples to illustrate the obtained results.

7.5. Impulsive control for continuous-time Markov decision processes

The following result has been obtained by Alexey Piunovskiy and François Dufour (CQFD member).

The objective of this work is to study continuous-time Markov decision processes on a general Borel state space with both impulsive and continuous controls for the infinite-time horizon discounted cost. The continuous-time controlled process is shown to be non explosive under appropriate hypotheses. The so-called Bellman equation associated to this control problem is studied. Sufficient conditions ensuring the existence and the uniqueness of a bounded measurable solution to this optimality equation are provided. Moreover, it is shown that the value function of the optimization problem under consideration satisfies this optimality equation. Sufficient conditions are also presented to ensure on one hand the existence of an optimal control strategy and on the other hand the existence of an ε-optimal control strategy. The decomposition of the state space in two disjoint subsets is exhibited where roughly speaking, one should apply a gradual action or an impulsive action correspondingly to get an optimal or ε-optimal strategy. An interesting consequence of our previous results is as follows: the set of strategies that allow interventions at time $t = 0$ and only immediately after natural jumps is a sufficient set for the control problem under consideration.

7.6. Impulsive control for continuous-time Markov decision processes: A Linear Programming Approach

The following result has been obtained by Alexey Piunovskiy and François Dufour (CQFD member).

The objective of this work is to investigate an optimization problem for continuous-time Markov decision processes with both impulsive and continuous controls. We consider the so-called constrained problem where the objective of the controller is to minimize a total expected discounted optimality criterion associated with a cost rate function while keeping other performance criteria of the same form, but associated with different cost rate functions, below some given bounds. Our model allows multiple impulses at the same time moment. The main objective of this work is to study the associated linear program defined on a space of measures including the occupation measures of the controlled process and to provide sufficient conditions to ensure the existence of an optimal control.

7.7. Conditions for the Solvability of the Linear Programming Formulation for Constrained Discounted Markov Decision Processes

The following result has been obtained by François Dufour (CQFD member) and T. Prieto-Rumeau.

This result concerns discrete-time constrained discounted Markov decision processes (MDP) with Borel state and action spaces, compact action sets, and lower semi-continuous cost functions. We introduce a set of hypotheses related to a positive weight function which allow us to consider cost functions that might not be bounded below by a constant, and which imply the solvability of the linear programming formulation of the constrained MDP. In particular, we establish the existence of a constrained optimal stationary policy. Our results are illustrated with an application to a fishery management problem.

7.8. Comparison of Kernel Density Estimators with Assumption on Number of Modes

The following result has been obtained by Gilles Durrieu, Raphaël Coudret and Jérôme Saracco (CQFD member).

A data-driven bandwidth choice for a kernel density estimator called critical bandwidth is investigated. This procedure allows the estimation to have as many modes as assumed for the density to estimate. Both Gaussian and uniform kernels are considered. For the Gaussian kernel, asymptotic results are given. For the uniform kernel, an argument against these properties is mentioned. These theoretical results are illustrated with a simulation study that compares the kernel estimators that rely on critical bandwidth with another one that uses a plug-in method to select its bandwidth. An estimator that consists in estimates of density contour clusters and takes assumptions on number of modes into account is also considered. Finally, the methodology is illustrated using environment monitoring data.

7.9. EEG classification for the detection of mental states

The following result has been obtained by Laurent Vezard, Pierrick Legrand (CQFD member), Marie Chavent (CQFD member), Frederique Faita-Ainseba and Trujillo Leonardo

The objective of the present work is to develop a method that is able to automatically determine mental states of vigilance; i.e., a person's state of alertness. Such a task is relevant to diverse domains, where a person is expected or required to be in a particular state of mind. For instance, pilots and medical staff are expected to be in a highly alert state and the proposed method could help to detect possible deviations from this expected state. This work poses a binary classification problem where the goal is to distinguish between a "relaxed" state and a baseline state ("normal") from the study of electroencephalographic signals (EEG) collected with a small number of electrodes. The EEG of 58 subjects in the two alertness states (116 records) were collected via a cap with 58 electrodes. After a data validation step, 19 subjects were retained for further analysis. A genetic algorithm was used to select a subset of electrodes. Common spatial pattern (CSP) coupled to linear discriminant analysis (LDA) was used to build a decision rule and thus predict the alertness of the subjects. Different subset sizes were investigated and the best compromise between the number of selected electrodes and the quality of the solution was obtained by considering 9 electrodes. Even if the present approach is costly in computation time (GA search), it allows to construct a decision rule that provides an accurate and fast prediction of the alertness state of an unseen individual.

7.10. Modeling and optimization of a launcher integration process

The following result has been obtained by Christophe Nivot (CQFD member), Benoîte De Saporta, François Dufour (CQFD member), Jacques Béhar, Damien Bérard-Bergery and Charles Elegbede.

We deal with the modeling and the optimization of a launcher integration process. The subassemblies go through various types of operations which are split up into workshops. Their operating time is supposed random due to possible breakdowns or staff issues. Storage capacity of output products is limited and costly. Launches have to be performed according to a predetermined schedule, and lateness also costs money. The rate of production of the subassemblies must be decided every year. Therefore, the system can be modeled with a Markov decision process which is suitable for decision optimization and cost minimization. Indeed, one must find a balance between slow production (thus low storage levels and high probability to be late), and fast production (high storage levels but respected schedule).

We propose a model of this integration process based on Markov decision models. We present the simulation we have performed so far and discuss the difficulties of the optimization.

7.11. ClustGeo: Ascendant Hierarchical Clustering (AHC) with geographical constraints

The following result has been obtained by Marie Chavent (CQFD member), Vanessa Kuentz-Simonet, Amaury Labenne and Jerome Saracco (CQFD member).

Hierarchical Ascendant Clustering (HAC) is a well-known method of individual clustering. This method aims to bring together individuals who are similar regarding to variables which describe them. But when individuals are geographical units, the user may wish geographically close individuals to be put in same clusters and that, without too much deteriorating the quality of the partition. The proposed ClustGeo method allows geographical constraints of proximity to be taken into account within the HAC. For that purpose, a new Ward homogeneity criterion based on two different matrices of distances is proposed.

7.12. Approche bayésienne non paramétrique pour la factorisation de matrice binaire à faible rang avec loi de puissance

The following result has been obtained by Adrien Todeschini (CQFD member) and François Caron.

We introduce a low-rank Bayesian nonparametric (BNP) model for bipartite graphs. Recently, Caron (2012) proposed a BNP model where each node is given its own sociability parameter allowing to capture the power-law behavior of real world bipartite graphs. This model can be considered as a rank one nonnegative factorization of the adjacency matrix. Building on the compound random measures recently introduced by Griffin and Leisen (2014), we derive a rank p generalization of this model where each node is associated with a p-dimensional vector of sociability parameters accounting for several latent dimensions. While preserving the desired properties of interpretability, scalability and power-law behavior, our model is more flexible and provides better predictive performance as illustrated on several datasets.

7.13. Compétitions d'apprentissage automatique avec le package R rchallenge

The following result has been obtained by Adrien Todeschini (CQFD member) Robin Genuer.

In machine learning, empirical performance on real data are crucial in the success of a method. Recent years have seen the emergence of a large number of machine learning competitions. These challenges are motivated by industrial (Netflix prize) or academic (HiggsML challenge) applications and put in competition researchers and data scientists to obtain the best performance. We wanted to expose students to this reality by submitting a challenge in the context of the machine learning course. The leaderboard is displayed on an automatically updated web page allowing emulation among students. The history of the results also allows them to visualize their progress through the submissions. In addition, the challenge can continue outside of the supervised sessions promoting independence and exploration of new learning techniques and computer tools.

The system we have implemented is available as an R package for reuse by other teachers. Building on the R Markdown and Dropbox tools, it requires no network configuration and can be deployed very easily on a personal computer.

7.14. Novelty Search

The following result has been obtained by Enrique Naredo, Leonardo Trujillo and Pierrick Legrand (CQFD member).

Novelty Search (NS) is a unique approach towards search and optimization, where an explicit objective function is replaced by a measure of solution novelty. However, NS has been mostly used in evolutionary robotics while its usefulness in classic machine learning problems has been unexplored. This work presents a NS-based Genetic Programming (GP) algorithm for supervised classification. Results show that NS can solve real-world classification tasks, validated on real-world benchmarks for binary and multiclass problems. These results are made possible by using a domain-specific behavior descriptor. Two new versions of the NS algorithm are proposed, Probabilistic NS (PNS) and a variant of Minimum Criterion NS (MCNS). The former models the behavior of each solution as a random vector and eliminates all of the original NS parameters while reducing the computational overhead of the NS algorithm. The latter uses a standard objective function to constrain and bias the search towards high performance solutions. The paper also discusses the effects of NS on GP search dynamics and code growth. Results show that NS can be used as a realistic alternative for supervised classi?cation, and for binary problems the NS algorithm exhibits an implicit bloat control ability.

Keywords: Novelty Search, Behavior-based Search, Supervised Classification, Bloat

7.15. Classification of Epileptic states

The following result has been obtained by Emigdio Z. FLores, Leonardo Trujillo and Pierrick Legrand (CQFD member).

The neurological disorder known as Epilepsy is characterized by involuntary recurrent seizures that diminish a patient's quality of life. Automatic seizure detection can help improve a patient's interaction with her/his environment, and while many approaches have been proposed the problem is still not trivially solved. In this work, we present a novel methodology for feature extraction on EEG signals that allows us to perform a highly accurate classification of epileptic states. Specifically, Hölderian regularity and Matching Pursuit are used as the main feature extraction techniques, and are combined with basic statistics to construct the final feature sets. These sets are then delivered to a Random Forests classification algorithm. Furthermore, several versions of the basic problem are tested and statistically validated producing perfect accuracy in most problems and 92% accuracy on the most difficult case. A comparison with recent results in relevant literature using a well known database reveals that our proposal achieves state-of-the-art performance.

Keywords: Epilepsy detection, Hölderian regularity, Matching Pursuit, EEG Classification

7.16. Prediction of expected performance

The following result has been obtained by Yuliana Martinez, Leonardo Trujillo and Pierrick Legrand (CQFD member).

The study of problem difficulty is an open issue in Genetic Programming (GP). The goal of this work is to generate models that predict the expected performance of a GP-based classifier when it is applied to an unseen task. Classification problems are described using domain-specific features, some of which are proposed in this work, and these features are given as input to the predictive models. These models are referred to as predictors of expected performance (PEPs). We extend this approach by using an ensemble of specialized predictors (SPEPs), dividing classification problems into specified groups and choosing the corresponding SPEP. The proposed predictors are trained using 2D synthetic classification problems with balanced datasets. The models are then used to predict the performance of the GP classifier on unseen real-world datasets that are multidimensional and imbalanced. Moreover, as we know, this work is the first to provide a performance

prediction of the GP classifier on test data, while previous works focused on predicting training performance. Accurate predictive models are generated by posing a symbolic regression task and solving it with GP. These results are achieved by using highly descriptive features and including a dimensionality reduction stage that simplifies the learning and testing process. The proposed approach could be extended to other classification algorithms and used as the basis of an expert system for algorithm selection.

7.17. Simulation of SPDEs for Excitable Media Using Finite Elements

The following result has been obtained by

M. Boulakia, A. Genadot (CQFD member) and M. Thieullen.

This result concerns the question of the discretization of Stochastic Partial Differential Equations (SPDE's) for excitable media. Working with SPDE's driven by colored noise, we consider a numerical scheme based on finite differences in time (Euler-Maruyama) and finite elements in space. Motivated by biological considerations, we study numerically the emergence of reentrant patterns in excitable systems such as the Barkley or Mitchell-Schaeffer models.

7.18. Conditional quantile estimation through optimal quantization: theoretical aspects

The following result has been obtained by J. Saracco (Inria CQFD) and I. Charlier (Inria CQFD), D. Paindaveine (ULB).

In this work, we use quantization to construct a nonparametric estimator of conditional quantiles of a scalar response Y given a d-dimensional vector of covariates X. First we focus on the population level and show how optimal quantization of X, which consists in discretizing X by projecting it on an appropriate grid of N points, allows to approximate conditional quantiles of Y given X. We show that this approximation is arbitrarily good as N goes to infinity and provide a rate of convergence for the approximation error. Then we turn to the sample case and define an estimator of conditional quantiles based on quantization ideas. We prove that this estimator is consistent for its fixed-N population counterpart. The results are illustrated on a numerical example.

7.19. Conditional quantile estimation based on optimal quantization: From theory to practice

The following result has been obtained by J. Saracco (Inria CQFD) and I. Charlier (Inria CQFD), D. Paindaveine (ULB).

In this work, small-sample properties of a nonparametric estimator of conditional quantiles based on optimal quantization, that was recently introduced(Charlier et al., JSPI, 2015),are investigated. More precisely, (i) the practical implementation of this estimator is discussed (by proposing in particular a method to properly select the corresponding smoothing parameter, namely the number of quantizers) and (ii) its finite-sample performances are compared to those of classical competitors. Monte Carlo studies reveal that the quantization-based estimator competes well in all cases and sometimes dominates its competitors, particularly when the regression function is quite complex. A real data set is also treated.While the main focus is on the case of a univariate covariate, simulations are also conducted in the bivariate case.

7.20. QuantifQuantile: An R Package for Performing Quantile Regression Through Optimal Quantization

The following result has been obtained by J. Saracco (Inria CQFD) and I. Charlier (Inria CQFD), D. Paindaveine (ULB).

In this work, we describe an R package, called QuantifQuantile, that allows to perform quantization-based quantile regression. In quantile regression, various quantiles of a response variable Y are modelled as functions of covariates (rather than its mean). An important application is the construction of reference curves/surfaces and conditional prediction intervals for Y. Recently, a nonparametric quantile regression method based on the concept of optimal quantization was proposed. This method competes very well with k-nearest neighbor, kernel, and spline methods. We describe also the various functions of the package and provide examples.

7.21. Numerical methods for simulation and optimization of piecewise deterministic Markov processes

This book is focused on theoretical and numerical aspects of simulation and optimization for piecewise deterministic Markov processes (PDMP's). PDMP's have been introduced in the literature by M. Davis as a general class of stochastic hybrid models. They form a family of Markov processes involving deterministic motion punctuated by random jumps. The motion of a PDMP includes both continuous and discrete variables. The continuous state variable represents the physical parameters of the system under consideration. The discrete mode characterizes the regimes of operation of the physical system and/or the environment. The process is defined through three local characteristics, namely the flow describing the deterministic trajectory between two consecutive jumps, the intensity function giving the jump rate and the Markov kernel specifying the post-jump location. A suitable choice of the state space and these local characteristics provides stochastic models covering a large number of problems such as engineering systems, operation research, economics, management science, biology, internet traffic, networks and reliability. The class of PDMP's is thus considered and recognized as a powerful modeling tool for complex systems.

However, surprisingly few works are devoted to the development of numerical methods for PDMP's to solve problems of practical importance such as evaluation and optimization of functionals of the process. The main objective of this book consists in presenting mathematical tools recently developed by the authors to address such problems. This book is not only focused on theoretical aspects such as proof of convergence of the approximation procedures but is also concerned with its applicability to practical problems. The approach we are proposing is general enough to be applied to several application domains. In particular, our results are illustrated by examples from the field of reliability.

Our approximation technique is based on the discretization using quantization of the underlying discrete-time Markov chain given by the post-jump locations and jump times of the PDMP. This strategy enables us to address a large class of numerical problems. In particular, in this book we focus, on the one hand, on the computation of expectation of functionals of PDMP's with applications to the evaluation of service times. On the other hand, we are interested in solving optimal control problems with applications to maintenance optimization.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

8.1.1. Airbus

Participants: Benoîte de Saporta, François Dufour, Christophe Nivot.

We are interested in the optimization of a launcher integration process. It comprises several steps from the production of the subassemblies to the final launch. The four subassemblies go through various types of operations such as preparation, integration, control and storage. These operations are split up into three workshops. Due to possible breakdowns or staff issues, the time spent in each workshop is supposed random. So is the time needed to deliver the subassemblies, for similar reasons including e.g. shipping delays. We also have to deal with constraints related to the architecture of the assembly process itself. Indeed, we have to take into account waiting policies between workshops. The workshops may work in parallel but can be blocked if their output is not transferred to the next workshop in line. Storage capacity of output products is limited.

Our goal is finding the best rates of delivery of the subassemblies, the best choice of architecture (regarding stock capacities) and the best times when to stop and restart the workshops to be able to carry out twelve launches a year according to a predetermined schedule at minimal cost. To solve this problem, we choose a mathematical model particularly suitable for optimization with randomness: Markov decision processes (MDPs).

We have implemented a numerical simulator of the process based on the MDP model. It provides the fullest information possible on the process at any time. The simulator has first been validated with deterministic histories. Random histories have then been run with exponentially distributed delivery times for the subassemblies and several families of random laws for the time spent in each workshop. Using Monte Carlo simulations, we obtain the distribution of the launch times. Preliminary optimization results allow choosing stock capacities and delivery rates that satisfy the launch schedule.

In this context, the PhD Thesis of Christophe Nivot (2013-2016) is funded by Chaire Inria-Astrium-EADS IW-Conseil régional d'Aquitaine.

8.1.2. Thales Optronique

Participants: Benoîte de Saporta, François Dufour, Alizée Geeraert.

Integrated maintenance, failure intensity, optimisation.

As part of optimizing the reliability, Thales Optronics includes systems that examine the state of their equipment. This function is performed by HUMS (Health Unit Monitoring Systems). The collaboration is the subject of the PhD of Alize Geeraert (CIFRE). The aim of this thesis is to implement in the HUMS a program based on observations that can determine the state of the system, optimize maintenance operations and evaluate the failure risk of a mission.

8.1.3. DCNS

Participants: Huilong Zhang, Jonatha Anselmi, François Dufour, Dann Laneuville.

This contract is with DCNS, a French industrial group specialized in naval defense and energy. In particular, DCNS designs and builds submarines and surface combatants, develops associated systems and infrastructure, and offers a full range of services to naval bases and shipyards, together with a focus into marine renewable energy. The main objective is to have robust algorithms able to build an accurate picture of the objects that are around a submarine by only using "passive sonar" information. This means that no information is transmitted by the submarine, which just listens to acustic waves coming in, to the target. We estimate the position and the velocity of moving targets through noisy observations and a Kalman-type filter. Estimates become accurate depending on the type and the number of maneuvers done by the submarine. Our goal is to combine the filter that is currently used in DCNS with a Markov decision process. This provides a systematic framework to compute the best sequence of submarine maneuvers that allows the system to determine, as soon as possible, accurate target position and velocity. The current technological transfer to DCNS stands in a stochastic optimization framework developed in Matlab that operates under the hypothesis that the target follows a uniform linear motion with constant velocity or zero acceleration. The case where targets move in a more complex manner gives concrete perspectives for further transfers to DCNS.

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. MATCHABLE project

Matchable is a startup incubated at IRA (Incubateur Régional d'Aquitaine) since Mars 2014. This startup predicts how players will behave, who is likely to spend money, who you should target with promotions/product placement, and who the developer has to pay attention to in order to prevent churners. The members of CQFD have supervised two masters internships and a postdoctoral researcher, granded by two PEPS contracts from AMIES.

9.1.2. Inter-LabEx project between CPU and TRAIL

The topic of the project is “Advanced statistical methods for analysis of multidimensional databases of human brain imaging”. The project focuses on the analysis of variability factors driving hemispheric specialization (HS) of the brain, a human specific character, for which a dedicated database has recently been built by GIN (Neurofunctional Imaging Group from L). GIN provides the database and pe./irabotrforms genotyping of fifty loci potentially affecting HS. The “Probability and Statistics” group (EPS) from the LabEx CPU works on the methodological developments of statistical tools to analyze these high dimensional data. Interactions between GIN and EPS allow to identify and to characterize the best variables, to perform additional analyses, and to suggest appropriate additional variables, especially in the case of the voxel being implemented. GIN is also involved in the interpretation of the statistical results generated throughout the project.

Dr Solveig Badillo has been hired as Postdoctoral researcher in may 2014 on this project for 20 months.

9.1.3. GIS ALBATROS, HUMO (HUman MOnitoring) project

Approche interdisciplinaire de l’évaluation de l’état cognitif de l’utilisateur. Participants: Jean-Marc André, Liliana Audin Garcia, Veronique Lespinet, Frédérique Faïta, Jérôme Sarraco, Pierrick Legrand.

Le but de ce micro-projet et de valider un protocole de recueil, traitement, et interprétation des données physiologiques pour l’évaluation de l’état de l’utilisateur. Les objectifs en sont :

- Mise en place d’un protocole de recueil de données reposant sur un des modèles théoriques classiques en psychologie cognitive (ex. modèle de la mémoire de travail, modèle attentionnel, etc.) pour servir de base à la comparaison des états cognitifs (comparaison de 2 conditions expérimentales reconnues en terme de performances cognitives).
- Mise en oeuvre d’une diversité de capteurs physiologiques conduisant au recueil de signaux variés durant les conditions expérimentales issus des protocoles de psychologie cognitive.
- Conduire une analyse statistique multivariée avec les tests existants permettant de tirer des informations quant à la structure des paramètres de monitoring.
- Discriminer les variables pertinentes : choix des signaux à retenir ; efficacité/pertinence vs diversité
- Etablir et décrire le lien éventuel de significativité entre les données physiologiques recueillies et l’état cognitif de l’utilisateur.

9.1.4. EMG analysis

Participants: Luis Herrera, Eric Grivel, Gregory Barrière, Marie chavent, Pierrick Legrand

L’analyse spectrale et temporelle des activités électromyographiques (EMGs, activités musculaires) occupe une place importante dans l’aide au diagnostic de pathologies sensorimotrices chez l’homme. Néanmoins, les outils actuellement utilisés en EMGs (transformée de Fourier notamment) demeurent limités et dans certains cas obsolètes pour le diagnostic différentiel de certaines pathologies. Tel est le cas par exemple de la maladie de Parkinson, du tremblement essentiel et de troubles cérébelleux qui s’accompagnent de tremblements similaires en termes de fréquence et d’amplitude. Notre projet, qui s’inscrit à la convergence de plusieurs disciplines (mathématiques, traitement du signal, neurobiologie et neurologie), vise à enrichir la quantité d’information pouvant être extraite des signaux EMGs à l’aide d’outils d’analyse à la pointe en matière de traitement du signal, avec pour objectif d’identifier des signatures EMG spécifiques de chaque pathologie et utilisables pour le diagnostic différentiel.

9.2. National Initiatives

9.2.1. ANR ADAPTEAU

The ANR project ADAPTEAU has been obtained for the period 2012-2016.

ADAPTEAU aims to contribute to the analysis and management of global change impacts and adaptation patterns in River-Estuarine Environments (REEs) by interpreting the scientific challenges associated with climate change in terms of: i) scale mismatches; ii) uncertainty and cognitive biases between social actors; iii) interdisciplinary dialogue on the "adaptation" concept; iv) critical insights on adaptive governance and actions, v) understanding the diversity of professional, social and economic practices vis-à-vis global change. The project aims to build an integrative and interdisciplinary framework involving biophysical and social sciences, as well as stakeholders and civil society partners. The main objective is to identify adaptive strategies able to face the stakes of global change in REEs, on the basis of what we call 'innovative adaptation options'.

We consider the adaptation of Social-Ecological Systems (SES) through the expected variations of the hydrological regimes (floods / low-flow) of the Garonne-Gironde REE—a salient issue in SW France, yet with a high potential for genericity The ADAPTEAU project will be organised as follows:

- Achieve and confront socio-economic and environmental assessments of expected CC impacts on the Garonne-Gironde river-estuarine continuum (task 1);
- Identify the emerging 'innovative adaptation options' endorsed by various social, economic, political actors of the territory (depolderisation, 'room for rivers' strategies, changes in economic activities, agricultural systems or social practices), then test their environmental, economic and social robustness through a selected subset (task 2);
- Scientists, representatives from administrators and civil society collaborate to build adaptation scenarios, and discuss them in pluralistic arenas in order to evaluate their social and economic feasibility, as well as the most appropriate governance modes (task 3).
- Disseminate the adaptation strategies to academics and managers, as well as to the broader society (task 4).

  The expected results are the definition and diffusion of new regional-scale reference frameworks for the discussion of adaptation scenarios in REE and other SESs, as well as action guidelines to better address climate change stakes.

The CQFD team work on tasks 1 and 3.

9.2.2. ANR Piece

ANR Piece (2013-2016) of the program *Jeunes chercheuses et jeunes chercheurs* of the French National Agency of Research (ANR), lead by F. Malrieu (Univ. Tours). The Piecewise Deterministic Markov Processes (PDMP) are non-diffusive stochastic processes which naturally appear in many areas of applications as communication networks, neuron activities, biological populations or reliability of complex systems. Their mathematical study has been intensively carried out in the past two decades but many challenging problems remain completely open. This project aims at federating a group of experts with different backgrounds (probability, statistics, analysis, partial derivative equations, modeling) in order to pool everyone's knowledge and create new tools to study PDMPs. The main lines of the project relate to estimation, simulation and asymptotic behaviors (long time, large populations, multi-scale problems) in the various contexts of application.

9.2.3. ANR BNPSI "Bayesian Non Parametric methods for Signal and Image Processing"

Statistical methods have become more and more popular in signal and image processing over the past decades. These methods have been able to tackle various applications such as speech recognition, object tracking, image segmentation or restoration, classification, clustering, etc. We propose here to investigate the use of Bayesian nonparametric methods in statistical signal and image processing. Similarly to Bayesian parametric methods, this set of methods is concerned with the elicitation of prior and computation of posterior distributions, but now on infinite-dimensional parameter spaces. Although these methods have become very popular in statistics and machine learning over the last 15 years, their potential is largely underexploited in signal and image processing. The aim of the overall project, which gathers researchers in applied probabilities, statistics, machine learning and signal and image processing, is to develop a new framework for the statistical signal and image processing communities. Based on results from statistics and machine learning we aim at defining new models, methods

and algorithms for statistical signal and image processing. Applications to hyperspectral image analysis, image segmentation, GPS localization, image restoration or space-time tomographic reconstruction will allow various concrete illustrations of the theoretical advances and validation on real data coming from realistic contexts.

9.3. European Initiatives

9.3.1. FP7 & H2020 Projects

IRSES ACOBSEC

Project reference: 612689 Funded under: FP7-PEOPLE

Coordinator : Pierrick Legrand

Participants :

Université Victor Segalen Bordeaux II participation ended

Université de Bordeaux

Fundacao da Faculdade de Ciencias da Universidade de Lisboa Portugal

Universidad de Extremadura Spain

INESC ID - Instituto de Engenharia de Sistemas e Computadores, Investigacao e Desenvolvimento em Lisboa Participation ended

Over the last decade, Human-Computer Interaction (HCI) has grown and matured as a field. Gone are the days when only a mouse and keyboard could be used to interact with a computer. The most ambitious of such interfaces are Brain-Computer Interaction (BCI) systems. BCI's goal is to allow a person to interact with an artificial system using brain activity. A common approach towards BCI is to analyze, categorize and interpret Electroencephalography (EEG) signals in such a way that they alter the state of a computer. ACoBSEC's objective is to study the development of computer systems for the automatic analysis and classification of mental states of vigilance; i.e., a person's state of alertness. Such a task is relevant to diverse domains, where a person is required to be in a particular state. This problem is not a trivial one. In fact, EEG signals are known to be noisy, irregular and tend to vary from person to person, making the development of general techniques a very difficult scientific endeavor. Our aim is to develop new search and optimization strategies, based on evolutionary computation (EC) and genetic programming (GP) for the automatic induction of efficient and accurate classifiers. EC and GP are search techniques that can reach good solutions in multi-modal, non-differentiable and discontinuous spaces; and such is the case for the problem addressed here. This project combines the expertise of research partners from five converging fields: Classification, Neurosciences, Signal Processing, Evolutionary Computation and Parallel Computing in Europe (France Inria, Portugal INESC-ID, Spain UNEX, Bordeaux university, Sciences University of Lisbon) and South America (Mexico ITT, CICESE). The exchange program goals and milestones give a comprehensive strategy for the strengthening of current scientific relations amongst partners, as well as for the construction of long-lasting scientific relationships that produce high quality theoretical and applied research.

9.3.2. Collaborations in European Programs, except FP7 & H2020

Program: MTM2012-31393

Project acronym: NMMDP

Project title: Numerical methods for Markov decision processes

Duration: 01/2013 - 12/2015

Coordinator: Tomas Prieto-Rumeau

Other partners: Department of Statistics and Operations Research, UNED (Spain)

Abstract:

This project is funded by the Gobierno de Espana, Direcion General de Investigacion Cientifica y Tecnica (reference number: MTM2012-31393) for three years to support the scientific collaboration between Tomas Prieto-Rumeau, Jonatha Anselmi and François Dufour. This research project is concerned with numerical methods for Markov decision processes (MDPs). Namely, we are interested in approximating numerically the optimal value function and the optimal controls for different classes of constrained and unconstrained MDPs. Our methods are based on combining the linear programming formulation of an MDP with a discretization procedure referred to as quantization of a probability distribution, underlying the random transitions of the dynamic system. We are concerned with optimality criteria such as the total expected cost criterion (for finite horizon problems) and, on the other hand, the total expected discounted cost and the average cost optimality criteria (for infinite horizon problems).

9.4. International Initiatives

9.4.1. Inria Associate Teams not involved in an Inria International Labs

9.4.1.1. CDSS

Title: Control of Dynamic Systems Subject to Stochastic Jumps

International Partner (Institution - Laboratory - Researcher):

Universidade de São Paulo (Brazil) - Departamento de Matemática Aplicada e Estatística (ICMC) - Costa Eduardo

Start year: 2014

See also: https://team.inria.fr/cdss/fr/

The main goals of this joint team CDSS is to study the control of dynamic systems subject to stochastic jumps. Three topics will be considered throughout the next 3 years. In the first topic we study the control problem of piecewise-deterministic Markov processes (PDMP?s) considering constraints. In this case the main goal is to obtain a theoretical formulation for the equivalence between the original optimal control of PDMP?s with constrains and an infinite dimensional static linear optimization problem over a space of occupation measures of the controlled process. F. Dufour (CQFD, Inria) and O. Costa (Escola Politécnica da Universidade de São Paulo, Brazil) mainly carry out this topic. In the second topic we focus on numerical methods for solving control and filtering problems related to Markov jump linear systems (MJLS). This project will allow a first cooperation between B. de Saporta (Univ. Montpellier II) and E. Costa (Universidade de São Paulo, Brazil). The third research subject is focused on quantum control by using Lyapunov-like stochastic methods conducted by P. Rouchon (Ecole des Mines de Paris) and P. Pereira da Silva (Escola Politécnica da Universidade de São Paulo, Brazil).

9.4.2. Inria International Partners

9.4.2.1. Declared Inria International Partners

Tree-Lab, ITT. TREE-LAB is part of the Cybernetics research line within the Engineering Science graduate program offered by the Department of Electric and Electronic Engineering at Tijuana's Institute of Technology (ITT), in Tijuana Mexico. TREE-LAB is mainly focused on scientific and engineering research within the intersection of broad scientific fields, particularly Computer Science, Heuristic Optimization and Pattern Analysis. In particular, specific domains studied at TREE-LAB include Genetic Programming, Classification, Feature Based Recognition, Bio-Medical signal analysis and Behavior-Based Robotics. Currently, TREE-LAB incorporates the collaboration of several top researchers, as well as the participation of graduate (doctoral and masters) and undergraduate students, from ITT. Moreover, TREE-LAB is actively collaborating with top researchers from around the world, including Mexico, France, Spain, Portugal and USA.

9.5. International Research Visitors

9.5.1. Visits of International Scientists

Tomas Prieto-Rumeau (Department of Statistics and Operations Research, UNED, Madrid, Spain) visited the team during 2 weeks in 2015. The main subject of the collaboration is the approximation of Markov Decision Processes.

Oswaldo Costa (Escola Politécnica da Universidade de São Paulo, Brazil) collaborate with the team on the theoretical aspects of continuous control of piecewise-deterministic Markov processes. He visited the team during two weeks in 2015 supported by the Associate Team Inria: CDSS.

Alexey Piunovskiy (University of Liverpool) visited the team during 5 weeks in 2015. The main subject of the collaboration is the linear programming approach for Markov Decision Processes. This research was supported by the Clusters d'excellence CPU.

9.5.1.1. Internships

- Emigdio Z. Flores: 1 months, hosted by P. Legrand
- Luis Herrera: 3 months, hosted by P. Legrand

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific events organisation

10.1.1.1. Member of the organizing committees

- NEO 2015 - Pierrick Legrand, http://neo.cinvestav.mx/NEO2015/. The aim of the NEO is to bring together scientists from different fields and countries to discuss recent advances in numerical and evolutionary optimization. The workshop has first been held in 2013 in Tlalnepantla, Mexico, and the NEO 2015 is the third edition.
- EVOLVE 2015 - Pierrick Legrand, http://www.evolve-conference.org/. The aim of the EVOLVE international conference is to build a bridge between probability, statistics, set oriented numerics and evolutionary computing, as well as to identify new common and challenging research aspects. The event is also intended to foster a growing interest for robust and efficient new methods with a sound theoretical background and, last but not least, to unify theory-inspired methods and cutting-edge techniques aimed at reaching top-level performance. By gathering researchers with different backgrounds, ranging from computer science to mathematics, statistics and physics, to name just a few, a unified view and vocabulary can emerge where theoretical and practical advancements may echo in different domains.. The wide use and large applicability spectrum of evolutionary algorithms in real-life applications nurture the need for further developing solid theoretical grounds. Among many examples intricate mathematical objects show, as proven by acknowledged new results that evolutionary algorithms can, in some cases, act as good and fast estimators. Similarly, the handling of large quantities of data may require the use of distributed environments where the probability of failure and the stability of algorithms need to be formally addressed. Common practice confirms in many cases that theory-based results ensure performance guarantee factors for evolutionary algorithms in areas as diverse as optimization, bio-informatics or robotics. Summarizing, EVOLVE focuses on basic research and application challenges arising in theory, new paradigms and practice, thus aiming to provide a unifying view and to raise questions related to reliability, performance guarantees and modeling.

- EA 2015 - Pierrick Legrand, https://ea2015.inria.fr/. 12th Biennial International Conference on Artificial Evolution, EA 2015, held in Lyon (France). Previous EA editions took place in Bordeaux (2013), Angers (2011), Strasbourg (2009), Tours (2007), Lille (2005), Marseille (2003), Le Creusot (2001), Dunkerque (1999), Nimes (1997), Brest (1995), and Toulouse (1994). Authors had been invited to present original work relevant to Artificial Evolution, including, but not limited to: Evolutionary Computation, Evolutionary Optimization, Co-evolution, Artificial Life, Population Dynamics, Theory, Algorithmics and Modeling, Implementations, Application of Evolutionary Paradigms to the Real World (industry, biosciences, ...), other Biologically Inspired Paradigms (Swarm, Artificial Ants, Artificial Immune Systems, Cultural Algorithms...), Memetic Algorithms, Multi-Objective Optimization, Constraint Handling, Parallel Algorithms, Dynamic Optimization, Machine Learning, and hybridization with other soft computing techniques. Each submitted paper was reviewed by three members of the International Program Committee. Among the 31 submissions received, 18 papers were selected for oral presentation and 8 other papers for poster presentation. For the previous editions, a selection of the best papers which were presented at the conference and further revised were published (see LNCS volumes 1063, 1363, 1829, 2310, 2936, 3871, 4926, 5975, 7401 and 8752).

10.1.2. Scientific events selection

10.1.2.1. Member of the conference program committees

J. Anselmi has been a member of the TPC of the international conferences VALUETOOLS-2015, ScalCom-2015 and ASMTA-2015.

M. Chavent has been a member of program commitee of the SFC 2015 conference.

P. Legrand has been a member of program committee for Gecco 2015, EA 2015, NEO 2015, EVOLVE 2015.

F. Dufour has been a member of the program committee of the international SIAM conference on Control & its Application, July 2015.

10.1.3. Journal

10.1.3.1. Member of the editorial boards

F. Dufour is associate editor of the journal: SIAM Journal of Control and Optimization since 2009.

J. Saracco is an associate editor of the journal Case Studies in Business, Industry and Government Statistics (CSBIGS) since 2006.

10.1.3.2. Reviewer - Reviewing activities

All the members of CQFD are regular reviewers for several international journals and conferences in applied probability, statistics and operations research.

10.1.4. Invited talks

J. Anselmi gave the following invited talks;

- *Open-loop control of parallel queues: asymptotics of periodic policies*, international workshop "Modern Trends in Controlled Stochastic Processes: Theory and Applications", Liverpool, July 2015
- *Open-loop control of parallel queues: asymptotics of periodic policies*, international conference APS INFORMS, Istanbul, July 2015

M. Chavent gave invited talks at the seminar of MIAT-INRA of Toulouse in june 2015, and the seminar of the IECL of Nancy in November 2015.

A. Genadot gave a talk at the seminar of IMB, October 2015.

F. Dufour gave the invited talk *Constrained and Unconstrained Optimal Control of Piecewise Deterministic Markov Processes* international workshop "Modern Trends in Controlled Stochastic Processes: Theory and Applications", Liverpool, July 2015.

10.1.5. Research administration

M Chavent and J. Saracco are elected members of CNU 26.

B. de Saporta was an elected member of CNU 26 until sept. 2014.

J. Saracco is vice president of the french statistical society (SFdS).

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

- Licence : J. Anselmi, Probability, 12,41 hours (“équivalent TD”), L1, ENSEIRB MATMECA filiere telecom, Bordeaux, France
- Licence : J. Anselmi, Probability, 8 hours (“équivalent TD”), L1, ENSEIRB MATMECA filiere electronique, Bordeaux, France
- Licence: M. Chavent, Statistique descriptive, 36 ETD, L1, Bordeaux university, France
- License: M. Chavent, Modélisation statistique, 18 ETD, niveau L3, Bordeaux university, France
- Master : M. Chavent, Analyse des données 2, 25 ETD, niveau M2, Bordeaux university, France
- Master : M. Chavent, Apprentissage automatique, 25 ETD, niveau M2, Bordeaux university, France
- Licence : F. Dufour, Probabilités et statistiques, 16 heures, niveau L3, Institut Polytechnique de Bordeaux, école ENSEIRB-MATMECA, France.
- Master : F. Dufour, Méthodes numériques pour la fiabilité, 24 heures, niveau M1, Institut Polytechnique de Bordeaux, école ENSEIRB-MATMECA, France.
- Master : F. Dufour, Probabilités, 20 heures, niveau M1, Institut Polytechnique de Bordeaux, école ENSEIRB-MATMECA, France.
- P. Legrand, Algèbre (responsable de l’UE), Licence 1 SCIMS (108 heures)
- P. Legrand, Informatique pour les mathÃ©matiques (responsable de l’UE), Licence 1 et Licence 2 (36 heures)
- P. Legrand, Espaces Euclidiens. (responsable de l’UE), Licence 2 SCIMS (54 heures)
- P. Legrand, Formation Matlab pour le personnel CNRS (responsable de l’UE), (24 heures)
- Licence: J. Saracco, Descriptive statistics, 10.5h, L3, First year of ENSC, France
- Licence: J. Saracco, Mathematical statistics, 20h, L3, First year of ENSC, France
- Licence:J. Saracco, Data analysis (multidimensional statistics), 20h, L3, First year of ENSC, France
- Licence: J. Saracco, Mathematics (complement of linear algebra), 20h, L3, First year of ENSC, France
- Master: J. Saracco, Statistical modeling, 20h, M1, Second year of ENSC, France
- Master: J. Saracco, training project, 20h, M1, Second year of ENSC, France
- A. Genadot, Probabilités (30h), Licence MIASHS deuxième année, Université de Bordeaux.
- A. Genadot, Modélisation statistique (18h), Licence MIASHS troisième année, Université de Bordeaux.
- A. Genadot, Probabilités (30h), Master MIMSE première année, Université de Bordeaux.

10.2.2. Supervision

PhD completed : Karim Claudio, Un outil d’aide à la maîtrise des pertes dans les réseaux d’eau potable : mise en place d’un modèle de fuite multi-état en secteur hydraulique instrumenté , supervised by J. Saracco and V. Couallier.

PhD completed : Amaury Labenne, Approche Statistique du diagnostic territorial par la notion de qualité de vie, supervised by M. Chavent, J. Saracco and V. Kuentz.

PhD in progress : Adrien Todeschini, Elaboration et validation d'un système de recommandation bayésien, supervised by F. Caron and M. Chavent.

PhD completed : Isabelle Charlier, Optimal quantization applied to conditional quantile estimation, University of Bordeaux and Université Libre de Bruxelle, supervised by J. Saracco and D. Paindaveine.

PhD in progress : Christophe Nivot, Optimisation de la chaîne de montage du futur lanceur européen, September 2013, B. supervised by B. de Saporta and F. Dufour

PhD in progress : Alizé Geeraert, Contrôle optimal des processus Markoviens déterministes par morceaux et application à la maintenance, University of Bordeaux, September 2014, supervised by B. de Saporta and F. Dufour.

Nicolas Antunes: Application d'algorithmes prédictifs à l'identification de niches ecoculturelles des population du passé: approche ethnoarcheologique. Financement ERC F. D'Errico. Co-encadrement : D'Errico, Del Moral, Legrand. Cette thèse consiste à utiliser des algorithmes de type GARP pour prédire l'existence de niches écologiques à partir de données climatologiques. 2011-2014.

Emigdio Z. Flores Lopez, "Classification of mental states with genetic programming", PhD in engineering sciences. Funded by Conacyt (Consejo Nacional de Ciencia y Tecnologia), national scholarship for PNPC programs (Programa Nacional de Posgrados Calidad), Mexico. Co-encadrement: L. Trujillo (50%), P. Legrand (50%). 2013-2016.

10.2.3. Juries

J. Anselmi has been a member of the jury for the PhD defense of Henda Ben Cheikh, with thesis granted by INSA–Toulouse.

P. Legrand has been a member of the jury for the PhD defense of Nicolas Antunes, Ubx.

11. Bibliography

Major publications by the team in recent years

[1] J. ANSELMI, B. D'AURIA, N. WALTON. *Closed Queueing Networks Under Congestion: Nonbottleneck Independence and Bottleneck Convergence*, in "Math. Oper. Res.", 2013, vol. 38, n^{o} 3, pp. 469–491, http://dx.doi.org/10.1287/moor.1120.0583

[2] M. BOULAKIA, A. GENADOT, M. THIEULLEN. *Simulation of SPDEs for Excitable Media Using Finite Elements*, in "Journal of Scientific Computing", 2015, vol. 65, n^{o} 1, pp. 171-195, http://dx.doi.org/10.1007/s10915-014-9960-8

[3] I. CHARLIER, D. PAINDAVEINE, J. SARACCO. *Conditional quantile estimation through optimal quantization*, in "J. Statist. Plann. Inference", 2015, vol. 156, pp. 14–30, http://dx.doi.org/10.1016/j.jspi.2014.08.003

[4] M. CHAVENT, S. GIRARD, V. KUENTZ-SIMONET, B. LIQUET, T. M. N. NGUYEN, J. SARACCO. *A sliced inverse regression approach for data stream*, in "Comput. Statist.", 2014, vol. 29, n^{o} 5, pp. 1129–1152, http://dx.doi.org/10.1007/s00180-014-0483-4

[5] O. L. D. V. COSTA, F. DUFOUR. *Continuous average control of piecewise deterministic Markov processes*, Springer Briefs in Mathematics, Springer, New York, 2013, xii+116 p. , http://dx.doi.org/10.1007/978-1-4614-6983-4

[6] R. COUDRET, S. GIRARD, J. SARACCO. *A new sliced inverse regression method for multivariate response*, in "Comput. Statist. Data Anal.", 2014, vol. 77, pp. 285–299, http://dx.doi.org/10.1016/j.csda.2014.03.006

[7] F. DUFOUR, A. B. PIUNOVSKIY. *The expected total cost criterion for Markov decision processes under constraints*, in "Adv. in Appl. Probab.", 2013, vol. 45, n^{o} 3, pp. 837–859, http://dx.doi.org/10.1239/aap/1377868541

[8] F. DUFOUR, T. PRIETO-RUMEAU. *Approximation of average cost Markov decision processes using empirical distributions and concentration inequalities*, in "Stochastics", 2015, vol. 87, n^{o} 2, pp. 273–307, http://dx.doi.org/10.1080/17442508.2014.939979

[9] B. DE SAPORTA, F. DUFOUR, H. ZHANG, C. ELEGBEDE. *Optimal stopping for predictive maintenance of a structure subject to corrosion*, in "Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability", 2012, vol. 226, n^{o} 2, pp. 169-181, http://hal.inria.fr/hal-00554759

Publications of the year

Articles in International Peer-Reviewed Journals

[10] J. ANSELMI, B. GAUJAL, T. NESTI. *Control of parallel non-observable queues: asymptotic equivalence and optimality of periodic policies*, in "stochastic systems", December 2015, vol. 5, n^{o} 1 [*DOI :* 10.1214/14-SSY146], https://hal.archives-ouvertes.fr/hal-01102936

[11] I. CHARLIER, D. PAINDAVEINE, J. SARACCO. *Conditional Quantile Estimation based on Optimal Quantization: from Theory to Practice*, in "Computational Statistics and Data Analysis", 2015, https://hal.inria.fr/hal-01108504

[12] I. CHARLIER, D. PAINDAVEINE, J. SARACCO. *Conditional quantile estimation through optimal quantization*, in "Journal of Statistical Planning and Inference", 2015, vol. 156, pp. 14 - 30 [*DOI :* 10.1016/J.JSPI.2014.08.003], https://hal.inria.fr/hal-01108482

[13] I. CHARLIER, D. PAINDAVEINE, J. SARACCO. *QuantifQuantile : an R package for performing quantile regression through optimal quantization*, in "The R Journal", 2015, https://hal.inria.fr/hal-01108505

[14] O. COSTA, F. DUFOUR. *A linear programming formulation for constrained discounted continuous control for piecewise deterministic Markov processes*, in "Journal of Mathematical Analysis and applications", April 2015, vol. 424, n^{o} 2 [*DOI :* 10.1016/J.JMAA.2014.11.036], https://hal.archives-ouvertes.fr/hal-01246215

[15] B. DE SAPORTA, E. COSTA. *Approximate Kalman-Bucy filter for continuous-time semi-Markov jump linear systems*, in "IEEE Transactions on Automatic Control", 2016, vol. 61, n^{o} 8 [*DOI :* 10.1109/TAC.2015.2495578], https://hal.archives-ouvertes.fr/hal-01062618

[16] F. DUFOUR, A. B. PIUNOVSKIY. *Impulsive control for continuous-time Markov decision processes*, in "Advances in Applied Probability", March 2015, vol. 47, n^{o} 1 [*DOI :* 10.1239/AAP/1427814583], https://hal.archives-ouvertes.fr/hal-01246222

[17] F. DUFOUR, A. B. PIUNOVSKIY. *Impulsive Control for Continuous-Time Markov Decision Processes: A Linear Programming Approach*, in "Applied Mathematics and Optimization", 2015 [*DOI :* 10.1007/s00245-015-9310-8], https://hal.archives-ouvertes.fr/hal-01246229

[18] F. DUFOUR, T. PRIETO-RUMEAU. *Conditions for the Solvability of the Linear Programming Formulation for Constrained Discounted Markov Decision Processes*, in "Applied Mathematics and Optimization", 2015 [*DOI :* 10.1007/s00245-015-9307-3], https://hal.archives-ouvertes.fr/hal-01246228

[19] G. DURRIEU, R. COUDRET, J. SARACCO. *Comparison of Kernel Density Estimators with Assumption on Number of Modes*, in "Communications in Statistics - Simulation and Computation", January 2015, vol. 44, n° 1, pp. 196-216 [*DOI :* 10.1080/03610918.2013.770530], https://hal.archives-ouvertes.fr/hal-01074437

[20] Y. MARTINEZ, L. TRUJILLO, P. LEGRAND, E. GALVAN-LOPEZ. *Prediction of Expected Performance for a Genetic Programming Classifier*, in "Genetic Programming and Evolvable Machines", 2016, https://hal.inria.fr/hal-01252141

[21] L. VÉZARD, P. LEGRAND, M. CHAVENT, F. FAÏTA-AÏNSEBA, L. TRUJILLO. *EEG classification for the detection of mental states*, in "Applied Soft Computing", 2015, vol. 32, pp. 113-131 [*DOI :* 10.1016/J.ASOC.2015.03.028], https://hal.inria.fr/hal-01207506

Invited Conferences

[22] B. DE SAPORTA. *Asymétrie et mémoire dans la division cellulaire*, in "5e Rencontres Scientifiques Sherbrooke-Montpellier", Sherbrooke, Canada, 2015, https://hal.archives-ouvertes.fr/hal-01193163

International Conferences with Proceedings

[23] M. CASTELLI, L. TRUJILLO, L. VANNESCHI, S. SILVA, E. Z-FLORES, P. LEGRAND. *Geometric Semantic Genetic Programming with Local Search*, in "Proceedings of the Conference on Genetic and Evolutionary Computation - GECCO '15", Madrid, Spain, 2015 [*DOI :* 10.1145/2739480.2754795], https://hal.inria.fr/hal-01207505

[24] C. NIVOT, B. DE SAPORTA, F. DUFOUR, J. BÉHAR, D. BÉRARD-BERGERY, C. ELEGBEDE. *Modeling and optimization of a launcher integration process*, in "ESREL 2015", Zurich, Switzerland, 2015, https://hal.archives-ouvertes.fr/hal-01202585

[25] E. Z-FLORES, L. TRUJILLO, O. SCHÜTZE, P. LEGRAND. *A Local Search Approach to Genetic Programming for Binary Classification*, in "Proceedings of the Conference on Genetic and Evolutionary Computation - GECCO '15", Madrid, Spain, 2015 [*DOI :* 10.1145/2739480.2754797], https://hal.inria.fr/hal-01207504

National Conferences with Proceedings

[26] M. CHAVENT, V. KUENTZ-SIMONET, A. LABENNE, J. SARACCO. *ClustGeo: Ascendant Hierarchical Clustering (AHC) with geographical constraints*, in "47èmes Journées de Statistique de la SFdS", Lille, France, June 2015, https://hal.inria.fr/hal-01246856

[27] A. TODESCHINI, F. CARON. *Approche bayésienne non paramétrique pour la factorisation de matrice binaire à faible rang avec loi de puissance*, in "47e Journées de Statistique", Lille, France, Société Française de Statistique, June 2015, https://hal.inria.fr/hal-01157151

[28] A. TODESCHINI, F. CARON. *Approche bayésienne non paramétrique pour la factorisation de matrice binaire à faible rang avec loi de puissance*, in "47èmes Journées de Statistique de la SFdS", Lille, France, Société Française de Statistique, June 2015, https://hal.inria.fr/hal-01256860

[29] A. TODESCHINI, R. GENUER. *Compétitions d'apprentissage automatique avec le package R rchallenge*, in "47èmes Journées de Statistique de la SFdS", Lille, France, Société Française de Statistique, June 2015, https://hal.inria.fr/hal-01157147

Conferences without Proceedings

[30] M. CHAVENT, M. FUENTES. *Monothetic divisive clustering. The divclust R package*, in "4èmes Rencontres R", Grenoble, France, June 2015, https://hal.inria.fr/hal-01246857

[31] M. CHAVENT, V. KUENTZ, A. LABENNE, B. LIQUET, B. LIQUET, J. SARACCO. *Multivariate analysis of mixed data: The PCAmixdata R package*, in "The useR! Conference 2015", Aalborg, Denmark, June 2015, https://hal.inria.fr/hal-01246858

[32] E. NAREDO, L. TRUJILLO, F. F. D. VEGA, S. SILVA, P. LEGRAND. *Diseñando Problemas Sintéticos de Clasificación con Superficie de Aptitud Deceptiva*, in "MAEB 2015", Merida, Spain, 2015, https://hal.inria.fr/hal-01207507

Scientific Books (or Scientific Book chapters)

[33] B. DE SAPORTA, F. DUFOUR, H. ZHANG. *Numerical Methods for Simulation and Optimization of Piecewise Deterministic Markov Processes*, Mathematics and statistics series, Wiley-ISTE, 2015, https://hal.archives-ouvertes.fr/hal-01249897

Other Publications

[34] J. ANSELMI, N. WALTON. *Decentralized Proportional Load Balancing*, January 2015, working paper or preprint, https://hal.inria.fr/hal-01103205

[35] L. CHAARI, S. BADILLO, T. VINCENT, G. DEHAENE-LAMBERTZ, F. FORBES, P. CIUCIU. *Hemodynamic-Informed Parcellation of fMRI Data in a Joint Detection Estimation Framework*, November 2015, Submitted to IEEE Transactions on Medical Imaging [*DOI :* 10.1007/978-3-642-33454-2_23], https://hal.archives-ouvertes.fr/hal-01228007

[36] L. CHAARI, S. BADILLO, T. VINCENT, G. DEHAENE-LAMBERTZ, F. FORBES, P. CIUCIU. *Subject-level Joint Parcellation-Detection-Estimation in fMRI*, January 2016, working paper or preprint, https://hal.inria.fr/hal-01255465

[37] E. COSTA, B. DE SAPORTA. *Linear minimum mean square filters for Markov jump linear systems*, 2015, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01251334

[38] B. DELYON, B. DE SAPORTA, N. KRELL, L. ROBERT. *Investigation of asymmetry in E. coli growth rate*, 2015, working paper or preprint, https://hal.inria.fr/hal-01201923

[39] S. GIRARD, J. SARACCO. *Supervised and unsupervised classification using mixture models*, December 2015, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01241818

[40] C. NIVOT, B. DE SAPORTA, F. DUFOUR, D. BÉRARD-BERGERY, C. ELEGBEDE. *Optimization of a launcher integration process: a Markov decision process approach* *, 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01269541

[41] A. TODESCHINI, F. CARON. *Exchangeable Random Measures for Sparse and Modular Graphs with Overlapping Communities*, February 2016, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01270854

References in notes

[42] N. DUAN, K.-C. LI. *Slicing regression: a link-free regression method*, in "Ann. Statist.", 1991, vol. 19, n^{o} 2, pp. 505–530, http://dx.doi.org/10.1214/aos/1176348109

[43] R. DUDA, P. HART, D. STORK. *Pattern Classification*, John Wiley, 2001

[44] K.-C. LI. *Sliced inverse regression for dimension reduction*, in "J. Amer. Statist. Assoc.", 1991, vol. 86, n^{o} 414, pp. 316–342, With discussion and a rejoinder by the author