UQSay seminar on Uncertainty Quantification and related topics, organized by L2S and MSSMAT CentraleSupelec Paris-Saclay

Christian Soize (Laboratoire MSME UMR8208, Université Gustave Eiffel, chistian.soize@univ-eiffel.fr)

A probabilistic learning on manifolds as a new tool in machine learning and data science with applications to optimization, statistical inverse problems, and UQ

In Machine Learning (generally devoted to *big-data case*), the *predictive learning* (or the *supervised learning*) approach consists in identifying/learning a random mapping **F**: **w**↦**q** = **F(w). T**he parameters vector **w** (input) is modelled by a random vector **W** with known probability distribution $P_{\mathbf{W}}(d\mathbf{w})$. The vector of quantities of interest **q** (outputs) is the non-Gaussian random variable **Q** = **F**(**W**) = **f**(**W**,**U**) whose probability distribution is unknown, given an *initial dataset* (or training set) $D_N = \{(\mathbf{w}^j,\mathbf{q}^j), j=1,\ldots N\}$ of N independent realizations of random vector (**W**,**Q**)**.** The measurable mapping **f** is deterministic and **U** is a random vector whose probability distribution is known. The approach of probabilistic learning on manifold (recently introduced) will be presented, which allows for constructing a generator of an estimation of the joint probability distribution $P_{\mathbf{W},\mathbf{Q}}(d\mathbf{w},d\mathbf{q}; N)$ using only D_N, which completely characterizes random mapping **F**. In this framework, novel computational statistical tools will be presented for the *small-data challenge* for which N is relatively small and consequently, is not sufficient large for constructing converged statistical estimates. We will present two extensions of the PLoM, the first one devoted to the learning under statistical constraints defined by experiments of physical models, and the second one for the Bayesian inference in high dimension. Several applications will be presented such as, the nonconvex optimization under uncertainties of a scramjet, the model-form uncertainties in nonlinear computational dynamics, the identification of non-Gaussian random elasticity fields for random media, the ultrasonic wave propagation in biological tissues.

Joint work with R. Ghanem (USC-USA), C. Farhart and R. Tezaur (Stanford-USA), C. Safta and H. N. Najm (Sandia, Livermore-USA), C. Desceliers (Univ Eiffel), C. Thimmisetty (Lawrence Livermore NL)

Refs: hal-01325279, hal-01576263, hal-02100250, hal-02341912, hal-02341991, arXiv:1910.12717