Statistique bayésienne

- Kaniav Kamary (CentraleSupélec)

  Title: Bayesian Principal Component Analysis

  Abstract: One of the main concerns of the principal component analysis (PCA) is how to choose the appropriate number of principal components to be retained. By using a probabilistic definition of the PCA, Christopher Bishop (1998) suggested a partially Bayesian approach that determines the effective reduced data dimension that relies on the informative priors for the model parameters. Since determining the exact form of the prior distribution based on the available information is not always an easy task, the Bayesian procedures may be influenced by wrong choice of the prior distributions. H. Sook OH et al. (2010) extended Christopher Bishop (1998)'s approach by using mixture and informative priors and compute posterior estimations of the parameters using MCMC algorithm. The estimation procedure, however, has a strong dependence on the choice of the prior distribution and jeopardizes the use of the approach in cases of real data analysis. This paper extends H. Sook OH et al. (2010)'s approach and proposes a noninformative prior modeling leading to a fully Bayesian estimation procedure. The approach simultaneously automates the effective dimensionality selection and estimates the principal components using a Gibbs sampler without the need of calibrating prior distribution.

- Paul Egels (Sorbonne Université)

  Title: Posterior and variational inference for deep neural networks with heavy-tailed weights

  Abstract: We consider deep neural networks in a Bayesian framework with a prior distribution sampling the network weights at random. Following a recent idea of Agapiou and Castillo (2023), who show that heavy-tailed prior distributions achieve automatic adaptation to smoothness, we introduce a simple Bayesian deep learning prior based on heavy-tailed weights and ReLU activation. We show that the corresponding posterior distribution achieves near-optimal minimax contraction rates, simultaneously adaptive to both intrinsic dimension and smoothness of the underlying function, in a variety of contexts including nonparametric regression, geometric data and Besov spaces. While most works so far need a form of model selection built-in within the prior distribution, a key aspect of our approach is that it does not require to sample hyperparameters to learn the architecture of the network. We also provide variational Bayes counterparts of the results, that show that mean-field variational approximations still benefit from near-optimal theoretical support.

- Pierre Wolinski (Université Paris-Saclay)

  Title: Une équivalence entre a priori bayésiens et pénalités en inférence variationnelle

  Abstract: En apprentissage automatique, il est courant d'optimiser les paramètres d'un modèle avec un terme de pénalité ad hoc, qui pousse les paramètres dans des directions choisies au départ (par exemple vers zéro). Le terme de pénalité apparaît naturellement en inférence variationnelle, technique permettant d'approximer l'a posteriori bayésien dans des contextes où il est trop difficile à calculer exactement. Dans ce cadre, la pénalité est proportionnelle à une divergence de Kullback–Leibler (KL) entre l'approximation de la loi a posteriori et la loi a priori. Nous caractérisons quelles pénalités peuvent prendre la forme d'une KL, et proposons une formule pour calculer la loi a priori correspondant à une pénalité donnée. Entre autres, ce point de vue permet de fournir une heuristique sur le facteur de pénalité, qui est usuellement un hyperparamètre à optimiser, dans le cas des réseaux de neurones.

- Zacharie Naulet (Université Paris-Saclay)

  Title: Bayesian nonparametric inference for « species-sampling » problems

  Abstract: Given an observed sample from a generic population of individuals belonging to species, "species-sampling" problems (SSPs) call for estimating features of the unknown species composition of additional unobservable samples from the same population. Within the broad class of SSPs, the problems of estimating coverage probabilities, the number of unseen species and coverages of prevalences have emerged in the past three decades for being the subject of numerous methodological and applied works, mostly in biological sciences but also in statistical machine learning, electrical engineering, theoretical computer science, information theory and forensic statistics. In this work, we focus on these popular SSPs, and present an overview of their Bayesian nonparametric (BNP) analysis under the Pitman–Yor process (PYP) prior. We consider the critical problem of estimating the discount parameter and the scale parameter of the PYP prior, showing a property of Bayesian consistency with respect to estimation through the hierarchical Bayes and the empirical Bayes approach, that is: the discount parameter can be always estimated consistently, whereas the scale parameter cannot be estimated consistently, thus advising caution in posterior inference. The limiting shape of posterior distributions is also investigated, showing that in the large sample regime the sequence of posterior converges to a non gaussian distribution.Joint work with Cecilia Balocchi and Stefano Favaro.