Session 1

Inverse bounds for learning latent structures

▪Speaker Long Nguyen (University of Michigan)

Abstract

Inverse bounds are inequalities that provide upper bound estimates of the distance of latent structures of interest in a suitable Wasserstein space in terms of the statistical distance (e.g., total variation, Hellinger distance, KL divergence) on the space of data populations. This type of bounds are useful in the deriving the rates of Bayesian posterior contraction that arise in latent structured models, including convolution models, mixture models and hierarchical models. In this talk I will present several such inverse bounds for (i) mixing measures in mixture and convolution models, (ii) the de Finetti's mixing measure in mixture of product distributions, and (iii) mixing measures for various other settings in regression and contaminated models. The derivation of such inverse bounds requires deeper exploration into the conditions of identifiability and their refinements, which shed some light about the geometry of the space of latent structured models.

Optimal transport methods for Bayesian nonparametric modeling

▪Speaker Marta Catalano (University of Warwick)

Abstract

Optimal transport (OT) methods and Wasserstein distances are flourishing in many scientific fields as an effective means for comparing and connecting different random structures. In this talk we describe the first use of an OT distance between Lévy measures with infinite mass to solve a statistical problem. Complex phenomena often yield data from different but related sources, which are ideally suited to Bayesian modeling because of its inherent borrowing of information. In a nonparametric setting, this is regulated by the dependence between random measures: we derive a general Wasserstein index for a principled quantification of the dependence gaining insight into the models' deep structure. It also allows for an informed prior elicitation and provides a fair ground for model comparison. Our analysis unravels many key properties of the OT distance between Lévy measures, whose interest goes beyond Bayesian statistics, spanning to the theory of partial differential equations and of Lévy processes.

Session 2

Compositions of discrete random structures in Bayesian nonparametrics

▪Speaker Antonio Lijoi (Bocconi University)

Abstract

Compositions of discrete random probability measures are effective tools in Bayesian nonparametrics for modeling multiple sample data. Hierarchical processes are a noteworthy example, as their infinite-dimensional layers are able to capture latent features that account for data heterogeneity and allow for borrowing of information across different samples. In this talk we consider some general families of hierarchical compositions, both in static and in dynamic settings, and will highlight their relevant distributional properties, with a special focus on the induced dependence structure and on the clustering they induce within and across different samples. We further discuss the potential of such constructions for investigating partially exchangeable multilayer stochastic block models. The presentation will be complemented by some illustrations on simulated and real data.

Masked Bayesian neural networks : Computation and optimality

▪Speaker Yongdai Kim (Seoul National University)

Abstract

As data size and computing power increase, the architectures of deep neural networks (DNNs) have been getting more and more complex and huge, and thus there is a growing need to simplify such complex and huge DNNs. In this paper, we propose a novel sparse Bayesian neural network (BNN) which searches for a good DNN with an appropriate complexity. We employ the masking variables at each node which can turn off some nodes according to the posterior distribution to yield a sparse DNN. We develop an efficient MCMC algorithm for the full Bayesian inference and prove certain theoretical optimalities of the proposed BNN (i.e. minimax optimality and adaptiveness). By extensive numerical studies, we illustrate that the proposed BNN performs well compared to other existing methods in terms of prediction accuracy and uncertainty quantification.

Session 3

A general modeling framework for partially exchangeable data: multivariate species sampling priors

▪Speaker Igor Prünster (Bocconi University)

Abstract

Partial exchangeability is the ideal probabilistic framework for analyzing data from different, though related, sources. The implications on the induced dependence structure and borrowing of information across groups are explored. These findings inspire a new general class of nonparametric priors, termed multivariate species sampling models, which is characterized by its partially exchangeable partition probability function. This class encompasses several popular dependent nonparametric priors and has the merit of highlighting their core distributional properties.

A unified construction for series representations and finite approximations of completely random measures

▪Speaker Juho Lee (KAIST)

Abstract

Infinite-activity completely random measures (CRMs) have become important building blocks of complex Bayesian nonparametric models. They have been successfully used in various applications such as clustering, density estimation, latent feature models, survival analysis, and network science. Popular infinite-activity CRMs include the (generalized) gamma process and the (stable) beta process. However, except in some specific cases, exact simulation or scalable inference with these models is challenging and finite-dimensional approximations are often considered. In this work, we propose a general and unified framework to derive both series representations and finite-dimensional approximations of CRMs. Our framework can be seen as an extension of constructions based on size-biased sampling of Poisson point process [Perman1992]. It includes as special cases several known series representations as well as novel ones. In particular, we show that one can get novel series representations for the generalized gamma process and the stable beta process. We also provide some analysis of the truncation error.