26th International Conference
ENGINEERING MECHANICS 2020
Brno, Czech Republic, November 24 – 25, 2020

UNCERTAINTY QUANTIFICATION THROUGH BAYESIAN NONPARAMETRIC MODELLING

Kočková E.[*], Kučerová A.[**], Sýkora J.[***]

Abstract: *Recently there is an increasing endeavour to take into account the underlying uncertainties by stochastic modelling in order to make the numerical predictions as realistic as possible. Uncertainty quantification deals with distinct sources of nondeterminism. A lack of knowledge is expressed by epistemic uncertainties while aleatory uncertainties formulate an inherent randomness. In the case of estimating aleatory uncertainty, the task is to infer unknown but fixed probability density function and the corresponding epistemic uncertainty about this estimation. In order to avoid too strict assumptions about the unknown density function (e.g. prescription of a specific parameterised family of probability density functions), it can be modelled hierarchically by a stochastic process via the Bayesian nonparametric approach. The contribution presents application of a Dirichlet process mixture in modelling the aleatory uncertainty.*

Keywords: Uncertainty quantification, Bayesian nonparametrics, hierarchical modelling, density estimation, Markov chain Monte Carlo.

1. Introduction

For appropriate uncertainty quantification one has to distinguish between two principal types of uncertainties, specifically, they are epistemic and aleatory uncertainties (Oberkampf et al., 2002). The first uncertainty type is connected to a lack of knowledge, e.g. measurement errors or a small number of measurements. This epistemic uncertainty can be reduced by any additional information. On the other side, there is aleatory uncertainty or variability which is irreducible. The aleatory uncertainty represents natural variability or randomness of a considered quantity, which arises from neglecting some problem dimension. In other words, this variability originates from data collection, when the data are singled out e.g. from different locations or times and modelled as a random variable.

An estimation of uncertain factors influencing behaviour of an investigated system is a crucial task in predicting of future events. Inferring a probability distribution which is an infinite-dimensional object is a very complex problem. Commonly applied approaches are based on low-dimensional parameterisations of the unknown density function, traditionally they consist in prescribing some specific parameterised family of probability density functions (Sankararaman, 2013, Nagel, 2016 and Janouchová, 2018). The corresponding unknown statistical moments can be considered as uncertain random variables and inferred in the Bayesian way. This approach is based on the Bayesian parametric models whose basic feature is a fixed number of unknown parameters. The significant disadvantage of this method is the necessity of making the strong assumption about the density function structure. An inappropriate guess can lead to a totally misleading result, especially in the regions of low probability which are important e.g. in reliability analysis of building structures where the design is based on a very low failure probability.

Relaxing the density structure assumption is allowed by Bayesian nonparametric modelling which serves to model selection and adaptation according to the available data. In order to ensure consistency of the estimation, in other words to obtain undistorted inference results, some prior distribution with enough large support is necessary. In the case of density estimation, it is reasonable to use an infinite-dimensional

[*] Ing. Eliška Kočková: Faculty of Civil Engineering, Czech Technical University in Prague; Thákurova 7/2077; 166 29, Prague; CZ, eliska.kockova@fsv.cvut.cz

[**] Ing. Anna Kučerová, PhD: Faculty of Civil Engineering, Czech Technical University in Prague; Thákurova 7/2077; 166 29, Prague; CZ, Anna.Kucerova@cvut.cz

[***] Ing. Jan Sýkora, PhD: Faculty of Civil Engineering, Czech Technical University in Prague; Thákurova 7/2077; 166 29, Prague; CZ, jan.sykora.1@fsv.cvut.cz

doi: 10.21495/5896-3-274

nonparametric prior on the space of density functions, i.e. to construct a probability model for the unknown probability distribution itself (MacEachern, 2016). Commonly used nonparametric priors include stochastic processes or their mixtures, the specific setting is problem-dependent. The Gaussian processes are mostly applied in nonlinear regression problems, the mixtures of Dirichlet processes are suitable for density estimations (Gelman, 2014). Practically, despite the infinite dimensionality of the assumed prior, a finite-dimensional formulation is employed in the computations. The model complexity is determined on a basis of the available data, it means that the dimensionality of the Bayesian nonparametric model can change with a growing data set (Gershman, 2012). For a comprehensive overview of the Bayesian nonparametric methods we refer to the books Hjort et al. (2010) and Ghosal (2017).

The authors in the paper Liu et al. (2019) present uncertainty quantification approach based on Bayesian nonparametric ensemble which distinguishes epistemic and aleatory uncertainties but they consider inherent stochasticity in the data generating process such as influence of an imperfect sensor. Aleatory uncertainty is estimated only on the level of output for some fixed inputs as their goal is to refine predictive property of a constructed regression model. Another application of Bayesian nonparametric modelling as a useful tool for nonlinear regression problems can be found in Müller (2013).

In this contribution, we focus on estimating probability distribution of random factors from a countable number of observations with a help of the Bayesian nonparametrics allowing to capture distribution properties such as multimodality, asymmetry or heavy-tailedness. Specifically, the unknown but fixed probability density function is expressed by a hierarchical model based on the Dirichlet process mixture, which enables to model a continuous density function (Ghosal, 2017).

2. Density estimation

The most popular nonparametric method for estimating a probability density function is a histogram, more sophisticated is a kernel density estimation widely used by frequentists (Izenman, 1991). In the Bayesian nonparametrics, the Dirichlet process is well-known tool introduced as a suitable class of prior distributions with available analytical formulations of posterior distributions given a sample of observations (Ferguson, 1973). Particularly, the Dirichlet process is a probability distribution over the set of probability distributions, i.e. every realization of the process is a probability distribution. Nevertheless, the samples of the Dirichlet process are of a discrete nature, which makes it unsuitable for the density estimation of a continuous random variable. To overcome this obstacle, a hierarchical model based on the Dirichlet process is utilized producing a mixture of Dirichlet processes also called a Dirichlet process mixture (DPM) model (Antoniak, 1974).

Assuming a set of statistically exchangeable i.i.d. samples

$$\boldsymbol{x}_1, \dots, \boldsymbol{x}_n \sim F, \tag{1}$$

where $\sim$ stands for "distributed according to" and $x_i \in \boldsymbol{R}$, the goal is to infer the unknown probability density function f as a DPM model, where

$$f(\boldsymbol{x}) = \sum_{j=1}^{\infty} w_j \, g_\theta\big(\boldsymbol{x}|\boldsymbol{\theta}_j\big), \tag{2}$$

which is an infinite weighted mixture of smooth densities from a parametric family $\boldsymbol{G} = \{g_\theta | \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$ with latent variables $\boldsymbol{\theta}$. Weights w_j represent a Dirichlet process and their sum is equal to one. Considering P_0 as a probability measure on the parameter space $\boldsymbol{\Theta}$, the DPM has the following hierarchical structure:

$$P \sim \mathrm{DP}(\alpha, P_0)$$
$$\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_n | P \sim P$$
$$\boldsymbol{x}_i | \boldsymbol{\theta}_i \sim g_\theta(\boldsymbol{x}, \boldsymbol{\theta}_i), \quad i = 1, \dots, n. \tag{3}$$

A random probability distribution P is generated by a Dirichlet process with a positive scalar α called a concentration (or precision) parameter because it defines a spread of the prior distribution P around the base (or center) distribution P_0, which is the prior expection of P. A higher value of α means a higher level of the centralization.

For a sake of clarity, we give an example of observations from a mixture of normally distributed random variables where the observed data coincide with the random effect whose unknown probability distribution is the object of the Bayesian inference. In this case, the densities g_θ are assumed to be Gaussians with unknown mean values $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. The base distribution P_0 is assumed to be the normal-inverse-Wishart distribution which is conjugate prior distribution for $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and has its own four parameters. Multiplying this prior density by the normal likelihood gives a posterior density of the same family, which fundamentally simplifies the actual computations (Gelman, 2014). The inference is focused on the marginalized posterior distribution $p(\boldsymbol{\theta}_{1:n}|\boldsymbol{x}_{1:n})$ since the infinite-dimensional P is integrated out with a help of Polya urn representation of the Dirichlet process (Blackwell, 1973). The posterior samples can be obtained almost directly by Gibbs sampling (Spall, 2003 and MacEachern, 1998). The estimated density function in a comparison with the true density and observations is depicted in Fig. 1.

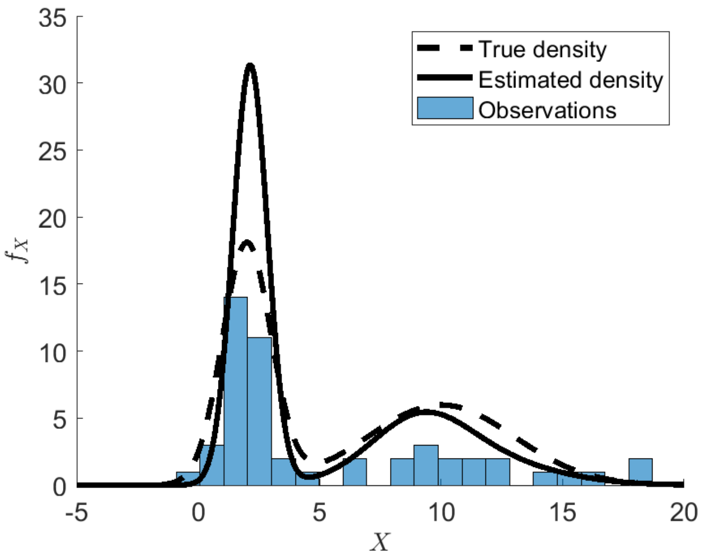

Fig. 1: Example of density estimation for mixture of two Gaussians $0.5N(2,1) + 0.5N(10,3)$*. Comparison of true and estimated probability density function based on Dirichlet process mixture of Gaussians considering set of 50 observations.*

3. Conclusions

The Bayesian nonparametric methods enable to quantify uncertainties more precisely without making restrictive assumptions about their probability distributions as it is done in the parametric approaches where the structure and a number of parameters of the estimated density function are prescribed a priori. Specifically, properties such as multimodality or asymmetry of the density function are usually omitted which can lead to unrealistic predictions and then to a wrong evaluation of risks connected to the modelled system.

Usually a limited number of observations of the uncertain effect is available and the hierarchical model based on the Dirichlet process mixture allows to share information among these samples. The nonparametric inference results in the density estimation of aleatory uncertainty formulated as a weighted finite-dimensional mixture of densities with random parameters. The number of components is determined on a basis of clustering the processed data so the parameterisation is not fixed.

This paper gives a very brief view into the world of the Bayesian nonparametrics with a simple illustrative example, however modelling density estimation especially in higher dimensions is not trivial. This topic is very actual and different effective methods have been developed in this area. Besides using the Dirichlet process mixtures, some researchers are focused on constructing hierarchical models based on the Pólya tree (Christensen, 2019). Another method is based on separating marginal and joint distribution by using copula transform (Majdara, 2019).

Finally, the important part of employing the infinite-dimensional priors for a density function space is its computational feasibility. The posterior distributions for non-conjugate priors require a special treatment to eliminate the error caused by the numerical integration and the computational demands. For example, the paper Neal (2000) presents a usage of the Metropolis-Hasting algorithm with the conditional prior equal to the proposal distribution. As an alternative to Markov chain Monte Carlo methods, Blei (2006) introduces deterministic algorithms derived from variational methods.

Acknowledgement

The financial support of the Czech Science Foundation (Project No. 18-04262S) and the Grant Agency of the Czech Technical University in Prague (Project No. SGS20/039/OHK1/1T/11) is gratefully acknowledged.

References

Antoniak, C. E. (1974) Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems. The annals of statistics, 2, 6, pp. 1152-1174.

Blackwell, D. and MacQueen, J. B. (1973) Ferguson Distributions Via Polya Urn Schemes. The Annals of Statistics, 1, 2, pp. 353-355.

Blei, D. M. and Jordan, M. I. (2006) Variational inference for Dirichlet process mixtures. Bayesian analysis, 1, 1, pp. 121-143.

Ferguson, T. S. (1973) A Bayesian analysis of some nonparametric problems. The Annals of Statistics, 1, 2, pp. 209-230.

Christensen, J. and Ma, L. (2019) A Bayesian hierarchical model for related densities by using Pólya trees. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 82, 1, pp. 127-153.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A. and Rubin, D. B. (2014) Bayesian data analysis. Third edition. Boca Raton: CRC Press. ISBN 978-143-9840-955.

Gershman, S. J. and Blei, D. M. (2012) A tutorial on Bayesian nonparametric models. Journal of Mathematical Psychology, 56, 1, pp. 1-12.

Ghosal, S. and Van der Vaart, A. (2017) Fundamentals of nonparametric Bayesian inference. Cambridge University Press. Cambridge series in statistical and probabilistic mathematics, 28. ISBN 978-0-521-87826-5.

Hjort, N. L., Holmes, C., Müller, P. and Walker, S. G. (2010) Bayesian nonparametrics. New York: Cambridge University Press. Cambridge series in statistical and probabilistic mathematics, 28. ISBN 978-0-521-51346-3.

Izenman, A. J. (1991) Recent developments in nonparametric density estimation. Journal of the American Statistical Association, 86, 413, pp. 205-224.

Janouchová, E. and Kučerová, A. (2018) Bayesian inference of heterogeneous viscoplastic material parameters, in: NMM 2018 - Nano & Macro Mechanics 2018. Praha: Czech Technical University in Prague, 2018. pp. 41-45. Acta Polytechnica CTU Proceedings. vol. 15. ISSN 2336-5382. ISBN 978-80-01-06457-3.

Liu, J., Paisley, J., Kioumourtzoglou, M.-A. and Coull, B. (2019) Accurate uncertainty estimation and decomposition in ensemble learning, in: Proc. 33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, pp. 8950-8961.

MacEachern, S. N. and Müller, P. (1998) Estimating mixture of Dirichlet process models. Journal of Computational and Graphical Statistics, 7, 2, pp. 223-238.

MacEachern, S. N. (2016) Nonparametric Bayesian methods: a gentle introduction and overview. Communications for Statistical Applications and Methods, 23, 6, pp. 445-466.

Majdara, A. and Nooshabadi, S. (2019) Nonparametric Density Estimation Using Copula Transform, Bayesian Sequential Partitioning and Diffusion-Based Kernel Estimator. IEEE Transactions on Knowledge and Data Engineering.

Müller, P. and Mitra, R. (2013) Bayesian nonparametric inference – why and how. Bayesian Analysis, 8, 2, pp. 269-302.

Nagel, J. B. and Sudret, B. (2016) A unified framework for multilevel uncertainty quantification in Bayesian inverse problems. Probabilistic Engineering Mechanics, 43, pp. 68-84.

Neal, R. M. (2000) Markov chain sampling methods for Dirichlet process mixture models. Journal of computational and graphical statistics, 9, 2, pp. 249-265.

Oberkampf, W. L., DeLand, S. M., Rutherford, B. M., Diegert, K. V. and Alvin, K. F. (2002) Error and uncertainty in modeling and simulation. Reliabity Engineering & System Safety, 75, 3, pp. 333-357.

Sankararaman, S. and Mahadevan, S. (2013) Separating the contributions of variability and parameter uncertainty in probability distributions. Reliability Engineering & System Safety, 112, pp. 187-199.

Spall, J. C. (2003) Estimation via Markov Chain Monte Carlo. IEEE Control Systems Magazine, 23, 2, pp. 34-45.