Probabilistic Inference

Probability Theory is an integral part of machine learning as a discipline, and continues to play a major role in the development of the field. The modularity of probabilistic models has repeatedly allowed for them to be extended and developed into general frameworks for large classes of problems, like regression and density estimation. In many areas of statistics and machine learning, probabilistic formulations were not the first to the table, but still helped improve understanding of implicit assumptions of the established statistical algorithms, helped with theoretical analysis, and lit the way to algorithmic extensions. Over the years, members of the department of empirical inference have contributed substantially to the field of probabilistic learning.

Gaussian Process Regression The textbook on Gaussian processes for machine learning, written by Carl Rasmussen in collaboration with Christopher Williams[1] has played a major role in the development of the Gaussian process framework into a workhorse of machine learning. In 2009, the book was awarded the deGroot Prize for Best Text on Bayesian Inference. Nickisch and Rasmussen [4] wrote a comprehensive overview of Gaussian process classification. Rasmussen and Nickisch [5] created the GPML toolbox, by far the most comprehensive package for Gaussian processes; originally, this was conceived as the code for the book, but by now it greatly exceeds models and algorithms discussed in the textbook, and is still being actively developed. Duvenaud, Nickisch and Rasmussen [6] have provided a recent contribution on structure in Gaussian process models using classical concepts of additivity. Nickisch [3] developed a toolbox for generalized linear models, providing scalable, efficient routines for inference in sparse linear, logistic, and abstract classification models.

Approximate Inference Methods The computational complexity of probabilistic formulations demands good numeric and algebraic approximation methods. Consequently, probabilistic models and approximation methods have developed alongside each other. Among the contributions to this area from members of the department, Habeck [1, 7] introduced a framework for the estimation of free energies and densities of state from equilibrium Monte Carlo simulations, offering a more robust and flexible alternative to thermodynamic integration. Seeger and Nickisch [10] developed a highly efficient, and provably convergent algorithmic improvement to the Expectation Propagation framework. Hennig et al. [9] contributed a lightweight, and numerically stable approximation for inference in latent Dirichlet models, allowing for efficient latent nonparametric dependency modeling between discrete distributions.

Probabilistic Numerical Methods Problems of numerical mathematics, like sampling, quadrature and optimization can be interpreted as inference problems: Their goal is to infer unknown properties of a function, like the value of a definite integral or the locations of extrema, from evaluations of the function at various locations. Even if the function is analytically known and can be evaluated at any point, these quantities of interest are truly unknown, otherwise it would not be necessary to call a numerical algorithm. This view of numerics as inference is not new, but has recently attracted renewed attention, partly due to work carried out in the department. Hennig & Schuler [2] used a probabilistic formulation of the global optimization problem to develop a sample-efficient experimental design algorithm that chooses evaluation locations based on expected information gain about the optimum. More recently, Hennig & Kiefel [8] performed a probabilistic analysis of classic quasi-Newton methods, like the widely- used BFGS algorithm, to identify algebraic prior assumptions made within these algorithms, and constructed a nonparametric extension of the quasi-Newton framework. Members of the department are helping to shape the discussion in this area, for example by co-organising a NIPS workshop on probabilistic numerics, in December 2012.

[1] Rasmussen & Williams, *Gaussian Processes for Machine Learning*, MIT Press, 2006

Publications

Journal Articles

[1] M Habeck. Bayesian estimation of free energies from equilibrium simulations. *Physical Review Letters*, 109(10), 9 2012. 1

[2] P Hennig and CJ Schuler. Entropy search for information-efficient global optimization. *Journal of Machine Learning Research*, 13:1809–1837, 6 2012. 1

[3] H Nickisch. glm-ie: The generalised linear models inference and estimation toolbox. *Journal of Machine Learning Research*, 13:1699–1703, 5 2012. 1

[4] H Nickisch and CE Rasmussen. Approximations for binary Gaussian process classification. *Journal of Machine Learning Research*, 9:2035–2078, 10 2008. 1

[5] CE Rasmussen and H Nickisch. Gaussian processes for machine learning (GPML) toolbox. *Journal of Machine Learning Research*, 11:3011–3015, 11 2010. 1

Articles in Conference Proceedings

[6] D Duvenaud, H Nickisch, and CA Rasmussen. Additive Gaussian processes. In J Shawe-Taylor, RS Zemel, P Bartlett, F Pereira, and KQ Weinberger, editors, *25th Annual Conference on Neural Information Processing Systems (NIPS)*, pages 226–234, Granada, Spain, 2011. 1

[7] M Habeck. Evaluation of marginal likelihoods via the density of states. In N Lawrence and M Girolami, editors, *Proceedings of the 15th International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 22, pages 486–494, La Palma, Canary Islands, 2012. JMLR: W&CP 22. 1

[8] P Hennig and M Kiefel. Quasi-Newton methods: A new direction. In *29th International Conference on Machine Learning (ICML 2012)*, pages 1–8, Edinburgh, UK, 7 2012. 1

[9] P Hennig, D Stern, R Herbrich, and T Graepel. Kernel topic models. In *15th International Conference on Artificial Intelligence and Statistics (AISTATS)*, pages 1–9, La Palma, Canary Islands, Spain, 4 2012. 1

[10] M Seeger and H Nickisch. Fast convergent algorithms for expectation propagation approximate Bayesian inference. In G Gordon, D Dunson, and M Dudík, editors, *14th International Conference on Artificial Intelligence and Statistics (AISTATS)*, pages 652–660, Fort Lauderdale, FL, USA, 4 2011. MIT Press. 1