Abstract:

Recently much attention has been paid to implicit probability models, since they have been used to great success for variational inference, generation of complex data types, and more. In most all of these settings, the goal has been to find a \emph{particular member} of that model family: optimized parameters index a distribution that is close (via a divergence or classification metric) to a target distribution. Much less attention, however, has been paid to the problem of \emph{learning a model itself}. Here we define implicit probability models with specific deep network architectures and optimization procedures for learning intractable exponential family models (\emph{not} a single distribution from those models). These exponential families, which are central to some of the most fundamental problems in probabilistic inference, are learned accurately, allowing operations like posterior inference to be executed directly and generically by an input choice of natural parameters, rather than performing inference via optimization for each particular realization of a distribution within that model.

We define this two-network architecture, which we term an \emph{exponential family network} (EFN), and we specify a stochastic optimization procedure over a variant of the typical Kullback-Leibler divergence. We then demonstrate the ability of EFNs to approximately learn exponential families and the benefits of approximating distributions in such restricted model spaces. Finally we demonstrate the computational savings afforded by this approach when learning the posterior family of point-process latent intensities, given neural spike responses in primary visual cortex of macaques.