Abstract

This report discusses *variational inference*, a powerful tool to approximate difficult-to-compute probability distributions. In particular, this method is useful to approximate posteriors in Bayesian inference. Two types of variational inference are introduced — *black box variational inference* and *Rényi divergence variational inference*, with the second one being a more general type of variational inference. The report also discusses stochastic optimization, a key tool in deriving variational inference algorithms, and variance reduction strategies that can help improve the efficiency of variational inference. Finally, we apply variational inference to learning in neural networks, one of the most common architectures in deep learning. All of the variational inference algorithms used in our empirical studies (in Section 4.5 and 7.4) were implemented by us with *Tensorflow* in Python. Also, we were the first to apply Gaussian variational inference with a factor covariance structure to deep learning (in Section 7.4).