Emergence of compositional representations in restricted Boltzmann machines

R. Monasson, J. Tubiana

CNRS & Ecole Normale Supérieure, Paris, France

Machine learning has undergone spectacular achievements in the recent years, thanks in particular to the development of the so-called deep learning methods. However, unsupervised learning, that is, the automatic extraction of the complex set of features composing real high-dimensional data still faces formidable challenges. In this talk I will review one of the simplest approach for unsupervised learning, the so-called Restricted Boltzmann machines (RBM), proposed about 30 years ago. RBM are empirically known to be efficient for modelling data distribution, and to be able to generate distributed and graded representations of the data, a property sometimes referred to as compositional phase. RBMs may also be seen as the elementary building bricks of more complex, deeper architectures.

Despite empirical evidence, our understanding of how and why RBMs function is still not satisfactory. Based on analytical calculations using the tools and concepts of the statistical physics of disordered systems, I will discuss the structural conditions (sparsity of the weights, low effective temperature, nonlinearities in the activation functions of hidden units, and adaptation of fields maintaining the activity in the visible layer) allowing RBM to operate in such a compositional phase. In addition, I will show that RBM exhibit remarkable energy landscape properties, which allow them to equilibrate very quickly despite the presence of multiple low-lying configurations. This fast mixing dynamics is at odds with the usual dynamics taking place in other models, such as the Hopfield model, in which equilibrations requires activated processes overcoming large barriers and is very slow. Evidence will be provided by the replica analysis of an adequate statistical ensemble of random RBMs and by RBM trained on the handwritten digits data set MNIST. Finally I will discuss the dynamics of learning of RBM, and, in particular, how they learn distributions with invariances. This study suggests practical way of improving the existing learning algorithms.

[l] J. Tubiana, R. Monasson, Phys. Rev. Lett. **118**, 138301 (2017).