Minimizing Probability Currents: A General Framework for Learning

Javier R. Movellan
Institute for Neural Computation
University of California San Diego
La Jolla, CA 92093-0515

We present result on a novel theoretical framework for learning in continuous stochastic neural networks (e.g., diffusion versions of the Hopfield model). Under the framework, which we named Minimum Probability Current (MPC), the goal of learning is to minimize the probability currents created when the observable units are clamped to samples from a target distribution. The MPC framework brings together two classic neural network algorithms (backpropagation and contrastive divergence) under the same umbrella and shows new ways in which stochastic neural networks can be trained. The following theoretical results will be presented:

- In continuous time networks, Hinton's approximation to contrastive divergence (the de facto standard) becomes exact.
- In continuous time networks contrastive divergence is an MPC algorithm, i..e., it minimizes the probability velocity fields generated by stochastic diffusion.
- In continuous time systems (including partially observable ones) the probability velocity gradient, a.k.a. the contrastive divergence gradient, can be computed using a single phase without the need for stochastic relaxation.
- In continuous time networks, as the stochastic diffusion constant goes to zero, contrastive divergence converges to standard backpropagation learning with an auto-encoder architecture.
- As the stochastic diffusion term increases additional higher order terms appear that are not present in standard backpropagation. However these additional terms can be efficiently computed.

Examples will be presented on toy and medium size problems to illustrate how the MPC framework operates.

References: Movellan (Under Review). A Minimum Velocity Approach to Learning. Neural Computation.

Topic: Learning Theory. Preference: Oral/Poster