

# マルコフ連鎖における定常分布の学習法としての Contrastive Divergence アルゴリズム

前田 新一

京都大学 情報学研究科 システム科学専攻

高次元のボルツマンマシンのような多数の離散変数を有する離散分布の学習は、正規化定数の計算が困難であり、最尤推定やベイズ推定の適用を難しくしていた。計算困難性は、マルコフ連鎖モンテカルロ法や平均場近似などの近似計算手法によって解決が図られていたが、近年、Contrastive Divergence アルゴリズム [1] と呼ばれる学習アルゴリズムが提案され、顔画像、手書き文字の特徴抽出・再構成、トピック解析などの種々の応用に対してうまく働くことが報告されている [2]。

しかしながら、Contrastive Divergence アルゴリズムは、その直接、最適化しているコスト関数やアルゴリズムの収束性が不明であるという問題点を有していた。そのため、Contrastive Divergence アルゴリズムを理解しようとする試みがなされるとともに、得られた知見に基づいた Contrastive Divergence アルゴリズムと類似した新しいアルゴリズムも提案されている。

Sohl-Dickstein らは、マルコフ連鎖の停留点の満たすべき条件から Minimum probability flow アルゴリズムと呼ばれる収束性の保証されたアルゴリズムを提案している [3]。Minimum probability flow アルゴリズムは、Contrastive Divergence アルゴリズムとは非常に近い関係を持っているがその収束性の保証されるアルゴリズムを導出することでより学習を早めることが示されている。その一方、Lyu は、contrastive divergence だけでなく partial likelihood, pseudo-likelihood や non-local contrastive objectives といった最尤推定と異なるコスト関数をもつ複数の学習法に対してそれらの関係性を統一的に俯瞰する見方を提案している [4]。

本発表では、これら近年の研究を紹介するとともに、それらとはまた違った見方を加えて Contrastive Divergence がマルコフ連鎖における定常分布の学習法としてみなせることについて解説する。その上で、なぜ収束が保証されないのか、どのような場合に収束が保証されるかについて述べる。

## 参考文献

- [1] G. E. Hinton, "Training Products of Experts by Minimizing Contrastive Divergence", *Neural Computation*, 14(8), pp.1771–1800, (2002).
- [2] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks", *Science*, 313(5786), pp.504–507, (2006).
- [3] J. Sohl-Dickstein, P. B. Battaglino, and M. R. DeWeese, "New Method for Parameter Estimation in Probabilistic Models: Minimum Probability Flow", *Physical Review Letters*, 107(22), 22601, (2011).
- [4] S. Lyu, "Unifying Non-Maximum Likelihood Learning Objectives with Minimum KL Contraction", *Advances in Neural Information Processing Systems (NIPS)*, (2011).