

# 連続値 RBM における 最尤推定と Contrastive Divergence 推定の一致

東京大・新領域 唐木田 亮

東京大・新領域 岡田 真人

理研・脳科学総合研究センター 甘利 俊一

## 1 背景

RBM (Restricted Boltzmann Machine) は可視変数と隠れ変数から成る完全 2 部グラフ上の生成モデルで、特に Deep Learning では多層ニューラルネットワークの構成要素として使われる。RBM のモデルパラメータを最尤推定する場合、尤度計算に多数回のギブスサンプリングが必要となるため、最尤推定は計算時間の観点から実用が困難である。そこで、最尤推定の近似手法としてギブスサンプリングを有限回で打ち止める Contrastive Divergence (CD) 推定が考案され [1]、広く実用されている。隠れ変数  $\mathbf{h}$ 、可視変数  $\mathbf{v}$ 、パラメータ行列  $W$  を伴う RBM のモデル分布を  $p(\mathbf{h}, \mathbf{v}; W) \propto \exp(\mathbf{h}^T W \mathbf{v})$  とすれば、 $W$  に対する  $CD_n$  推定の更新則は、

$$\tau \frac{dW}{dt} = \langle \mathbf{h} \mathbf{v}^T \rangle_0 - \langle \mathbf{h} \mathbf{v}^T \rangle_n$$

と定義される。右辺第一項は入力データに対する平均、第二項は有限  $n$  回のギブスサンプリングにより生成したサンプル  $(\mathbf{h}^{(n)}, \mathbf{v}^{(n)})$  に対する平均である。なお、無限回のサンプリング  $n \rightarrow \infty$  が最尤推定に対応する。経験的には、 $CD_1$  推定によって最尤推定に十分近い解が得られると指摘されているが、両推定の解が一致する一般的な理論保証は知られていない。

## 2 Contrastive Divergence 推定の平衡解

本研究では、可視変数  $\mathbf{v}$  が連続実数値をとる 2 種類の RBM において、最尤推定と CD 推定に共通する解を解析的に明らかにした。具体的には、 $CD_n$  推定の平衡解を行列方程式  $dW/dt = 0$  を解析的に解き、平衡解の摂動に対する安定性 [2] を確認した。まず、隠れ変数  $\mathbf{h}$  が連続実数値の RBM では、最尤推定と  $CD_n$  推定の安定平衡解はすべて一致することを示す。また、この安定平衡解は主成分分析に基づいた入力データの次元削減に対応することを明らかにする。次に、隠れ変数  $\mathbf{h}$  が離散値  $\{0,1\}$  をとる RBM では、両推定において少なくとも解のひとつが一致しており、その解が入力データの独立成分分析に対応することを明らかにする。

我々の解析結果は、 $CD_1$  から  $CD_n$  まで一般の CD 推定で成立する。これは特に、実用上広く使われる  $CD_1$  推定で最尤推定と同じ特徴成分が抽出できることを示している。

## 参考文献

- [1] G. E. Hinton. Training products of experts by minimizing contrastive divergence. *Neural Computation*, 14(8):1771–1800, 2002.
- [2] S. Amari, T. Chen, and A. Cichocki. Stability analysis of learning algorithms for blind source separation. *Neural Networks*, 10(8):1345–1351, 1997.