

ボルツマンマシンと情報統計力学

安田 宗樹

東北大学大学院情報科学研究科 応用情報科学専攻

近年のインターネットや携帯電話の普及により、様々な種類の情報データが個人ユーザーレベルでやりとりされるようになり、それにともない情報科学も多種多様な情報データに対して適応的な処理を行わざるを得なくなってきた。情報通信媒体に日々触れているならば、一口に情報データといっても静止画・動画や音楽、単なる文書に至るまで実に様々であることは容易に想像できよう。

現状の多くの研究がそうであるように、個々の情報データに対して個別に方法論を展開し、開発につなげていくことも有意義で実践的な方針であるが、もう一方として情報データをより一般的に扱える理論を進化させていくというのも今後のデータサイエンスを発展させていくなかで重要になってきている。統計的機械学習理論 (Statistical Machine Learning Theory) はまさにそのような理論枠組みのひとつである。機械学習理論とは扱いたい情報データの性質を受け取った情報データから自動的に抽出するための理論であり、統計的機械学習理論はベイズの理論を基礎とした確率的情報処理のための機械学習理論である。

本講演ではパラメトリックな統計的機械学習理論のモデルのひとつであるボルツマンマシンとよばれる確率モデルを扱う。ボルツマンマシンは図1に示されるようなグラフィカルモデル (Graphical Model) $G(V, E)$ 上に定義される確率モデルである。各ノード

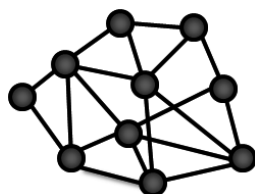


図 1: 10 ノードのボルツマンマシン。

$i \in V$ は確率変数 x_i をあらわしており、リンク $(i, j) \in E$ は確率変数同士の直接の相互作用を表現している。ボルツマンマシンは次の確率分布として定義される。

$$P(\mathbf{x} | \mathbf{h}, \mathbf{J}) := \frac{1}{Z(\mathbf{h}, \mathbf{J})} \exp \left(\sum_{i \in V} h_i x_i + \sum_{(i, j) \in E} J_{ij} x_i x_j \right) \quad (1)$$

\mathbf{h} は各確率変数に対するバイアスであり、 \mathbf{J} は隣接した確率変数同士の相互作用のパラメータである。扱いたい情報データの統計性 (確率的な規則性) にもっとも近くなるようにこれらのパラメータを調整することにより、情報データの統計性をこのモデル内で再現することがボルツマンマシンを用いた統計的機械学習理論の第一の課題である。学習 (Learning) とよばれるこの操作は最尤推定 (Maximum Likelihood Estimation) を用いることで達成される。

ボルツマンマシン上での推論

モデル(1)の学習の過程が無事に終わったとして、次に必要になるのが推論 (Inference) である。推論とは学習を終えたモデル上でのローカルな統計量 (例えばとあるノードの平均値やノード対の共分散) を計算することに相当する。しかしながら、一般に厳密なこの操作は NP-hard であり、ノードの数が多くなっていくことでこの操作は計算量的に困難になる。そこで何からの代替手段を考えねばならなくなるのだが、よく見てみるとモデル(1)は統計力学における磁性体の基本モデルであるイジングスピンモデルと等価であることがすぐにわかる。したがってボルツマンマシン上では物理学で発展してきた伝統的な手法が特別な変形なしにそのまま利用できる。

ひとつまず思い浮かぶのがモンテカルロ法などの確率的近似アルゴリズムを用いる方法であろう。しかし、それとは別にビリーフプロパゲーション (Belief Propagation) などの平均場近似に代表される決定論的近似アルゴリズムによる推論アルゴリズムの設計も可能な方針のひとつである。本講演では決定論的近似アルゴリズムによる推論の方にスポットライトを当て、新しい推論アルゴリズムを提案する。

本講演での提案アルゴリズムはスパースな系の近似に強いビリーフプロパゲーションの特性と、密な系の近似に強い適応 TAP 近似の両方の特性を併せ持つ強力な近似法になっており、より頑健な推論アルゴリズムを設計するために重要な知見を与えうる手法となっている (図2)。

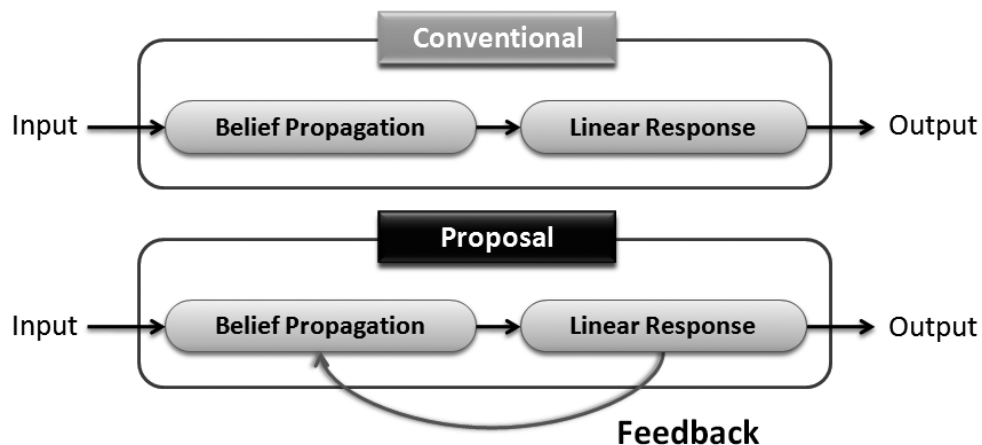


図 2: 提案アルゴリズムのスキーム。従来法 (感受率伝搬法; Susceptibility Propagation) ではビリーフプロパゲーションの結果に対して線形応答を適用し出力を得るため、線形応答による近似の補正結果はビリーフプロパゲーションに寄与することがない。対して提案法では線形応答の結果をビリーフプロパゲーションにフィードバックさせる構造になっているため、線形応答による補正をビリーフプロパゲーションに活かすことが可能となっている。