研究速報

エントロピー最大規範によるボルツマンマシンの解析

非会員　奥原　浩之[†]　　正　員　辻　　敏夫[††]
正　員　尾崎　俊治[††]

An Analysis of Boltzmann Machines by the Maximum Entropy Criterion

Koji OKUHARA[†], *Nonmember*, Toshio TSUJI[††] *and* Shunji OSAKI[††], *Members*

† 広島大学大学院工学研究科，東広島市
†† 広島大学工学部第二類（電気系），東広島市
Faculty of Engineering, Hiroshima University, Higashi-Hiroshima-shi, 724 Japan

あらまし　本研究では，ボルツマンマシンの定常状態に新しい解釈を与える．まず，ボルツマンマシンの定常状態の確率分布（ボルツマン分布）が，ユニット間の相関が一定のもとでエントロピーを最大にするような確率分布（正規分布）であると考える．そして，このネットワークの結合荷重と温度がラグランジュ乗数を与えることを示す．更に，定常状態では相関と結合荷重の間に成り立つ式が存在し，ネットワークの相関が与えられれば保存量が定義できることを示す．

キーワード　ボルツマンマシン，定常状態，エントロピー，相関

1. まえがき

人間の思考と結び付けて考えるには，あまりにも単純すぎるボルツマンマシンではあるが，その振舞いの示唆するところは人間の思考の振舞いの理解にとって少なからず意味があると思われる．ボルツマンマシンは，Ackley, Hinton and Senjnowski[(1)]による学習のアルゴリズムを適用することにより，外界から与えられた確率分布に対応した望ましい出力の確率分布を獲得することができる．もっとも，その学習のアルゴリズムの導出には統計物理など他の分野との間に，アナロジーが存在している．ここでは別の観点からボルツマンマシンの挙動を解析することにより，今まで見過ごされていた新たな解釈ができるようになるのではないかと考える．その足掛りとなるのが，ボルツマンマシンとエントロピーのかかわりである．

本論文では非線形処理をする n 個のユニットが相互に対称な重みで結合し，それぞれのユニットがある確率で0または1を非同期に出力するようなボルツマンマシンについて考察する．ボルツマンマシンの定常状態の確率分布が，ユニット間の相関が一定のもとでエントロピーを最大にするような確率分布であると考えるなら，ネットワークの結合荷重と温度がラグランジュ乗数を与えることが示される．更に，相関とラグランジュ乗数の間に関係式があり，相関が与えられると保存量が定義できることを示す．

2. ボルツマンマシンの定式化

各ユニットの状態を x_j，$j=1,2,\cdots,n$ とすると i 番目のユニットへの入力 X_i は，

$$X_i=\sum_{j=1}^{n} w_{ij}x_j-\theta_i \qquad (1)$$

で与えられる．ここで，w_{ij} は j 番目のユニットから i 番目のユニットへの結合荷重，θ_i は i 番目のユニットのしきい値である．ネットワークの結合荷重は対称としているので，

$$w_{ij}=w_{ji} \qquad (2)$$

である．そして，入力が X_i のとき出力 x_i がそれぞれ1および0となる確率を $P(x_i=1\,;X_i)$ および $P(x_i=0\,;X_i)$ で表すと，各ユニットは次の確率

$$P(x_i=1\,;X_i)=f\left(\frac{X_i}{T}\right) \qquad (3)$$

$$P(x_i=0\,;X_i)=1-P(x_i=1\,;X_i)=f\left(-\frac{X_i}{T}\right) \qquad (4)$$

で1または0のうちどちらかを出力する．ここで，$f(x)$ は非線形関数で通常，次のシグモイド関数

$$f(x)=\frac{1}{1+\exp(-x)} \qquad (5)$$

が用いられる．ネットワークのパターンが α である状態を $\boldsymbol{x}^\alpha\equiv(x_1{}^\alpha,x_2{}^\alpha,\cdots,x_n{}^\alpha)$ で表す．このとき，ネットワークのエネルギー関数 $E(\boldsymbol{x}^\alpha)$ が次のように，

$$E(\boldsymbol{x}^\alpha)=-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}x_i{}^\alpha x_j{}^\alpha+\sum_{i=1}^{n}\theta_i x_i{}^\alpha \qquad (6)$$

と定義できる．ここで，$x_i{}^\alpha$ はパターン α の i 番目のニューロンの出力状態である．式(6)において改めて $w_{ii}-\theta_i$ を w_{ii} とおくと，

$$E(\boldsymbol{x}^\alpha)=-\sum_{i=1}^{n}\sum_{j\geq i}^{n} w_{ij}x_i{}^\alpha x_j{}^\alpha \qquad (7)$$

を得る．これはユニットの出力状態 $x_i{}^\alpha$ が1または0の2値であり，$x_i{}^2=x_i$ とできるからである．

式(3)または式(4)に従って状態変化を繰り返すと，各状態がネットワークの結合荷重によって一意に決まる確率で出現するような定常状態に収束することがわかっている[(2)]．この確率分布はエネルギー関数を用いて，次のようなボルツマン分布

$$P(\boldsymbol{x}^\alpha)=Q\exp\left(-\frac{E(\boldsymbol{x}^\alpha)}{T}\right) \qquad (8)$$

で表せる．ここで，Q は規格化定数である．

3. エントロピー最大規範に基づくボルツマンマシンの解析

ボルツマンマシンの i 番目と j 番目のユニットの相関 C_{ij} は,

$$C_{ij}=\sum_{\alpha} x_i^{\alpha} P(\boldsymbol{x}^{\alpha}) x_j^{\alpha} \tag{9}$$

で与えられる. もちろん規格化の条件

$$\sum_{\alpha} P(\boldsymbol{x}^{\alpha})=1 \tag{10}$$

が存在する. 定常状態を考えるのであるから相関は定数である. ここで, エントロピー S

$$S=-\sum_{\alpha} P(\boldsymbol{x}^{\alpha}) \ln P(\boldsymbol{x}^{\alpha}) \tag{11}$$

を導入して, このエントロピーを最大にするような状態 $\boldsymbol{x}^{\alpha}$ が, 確率 $P(\boldsymbol{x}^{\alpha})$ で出現しているものと考えよう.

式(9)と式(10)の拘束下で式(11)を最大とする $P(\boldsymbol{x}^{\alpha})$ は, $n(n+1)/2+1$ 個のラグランジュ乗数 λ_{ij}, $i, j=1, 2, \cdots, n, j \geq i, \lambda$ を用いて求めることができる[3]. まず, 式(9)と式(10)とラグランジュ乗数を用いて式(11)を変形し,

$$J=S-(\lambda-1)(P(\boldsymbol{x}^{\alpha})-1)-\sum_{i=1}^{n}\sum_{j \geq i}^{n} \lambda_{ij}\left(C_{ij}-\sum_{\alpha} x_i^{\alpha} P(\boldsymbol{x}^{\alpha}) x_j^{\alpha}\right) \tag{12}$$

を定義する. そして, J を $P(\boldsymbol{x}^{\alpha})$ に関して微分し得られる式を 0 とおけば,

$$-\ln P(\boldsymbol{x}^{\alpha})-\lambda-\sum_{i=1}^{n}\sum_{j \geq i}^{n} \lambda_{ij} x_i^{\alpha} x_j^{\alpha}=0 \tag{13}$$

となる. これを解いて,

$$P(\boldsymbol{x}^{\alpha})=\exp\left(-\lambda-\sum_{i=1}^{n}\sum_{j \geq i}^{n} \lambda_{ij} x_i^{\alpha} x_j^{\alpha}\right) \tag{14}$$

を得る. 更に, 式(14)を規格化の条件, 式(10)に代入すると,

$$\exp(-\lambda)\sum_{\alpha} \exp\left(-\sum_{i=1}^{n}\sum_{j \geq i}^{n} \lambda_{ij} x_i^{\alpha} x_j^{\alpha}\right)=1 \tag{15}$$

を得る. ここで,

$$\sum_{\alpha} \exp\left(-\sum_{i=1}^{n}\sum_{j \geq i}^{n} \lambda_{ij} x_i^{\alpha} x_j^{\alpha}\right)=Z \tag{16}$$

とおく. これは分配関数[3]であって,

$$\lambda=\ln Z \tag{17}$$

である. 式(14)は,

$$P(\boldsymbol{x}^{\alpha})=Z^{-1}\exp\left(-\sum_{i=1}^{n}\sum_{j \geq i}^{n} \lambda_{ij} x_i^{\alpha} x_j^{\alpha}\right) \tag{18}$$

となる.

以上より, ラグランジュ乗数 λ_{ij} を決定できる. まず, 式(8)と式(18)より,

$$ZQ\exp\left(-\frac{E(\boldsymbol{x}^{\alpha})}{T}\right)=\exp\left(-\sum_{i=1}^{n}\sum_{j \geq i}^{n} \lambda_{ij} x_i^{\alpha} x_j^{\alpha}\right) \tag{19}$$

となる. そこで両辺の対数をとると,

$$\left(-\frac{E(\boldsymbol{x}^{\alpha})}{T}\right)^{\gamma}=-\sum_{i=1}^{n}\sum_{j \geq i}^{n} \lambda_{ij} x_i^{\alpha} x_j^{\alpha} \tag{20}$$

を得る. 但し,

$$\gamma=ZQ \tag{21}$$

である. ここで, $\gamma=1\,(Q=Z^{-1})$ とし, エネルギー関数 $E(\boldsymbol{x}^{\alpha})$ として式(7)を用い, 式(19)に代入すると,

$$\frac{\sum_{i=1}^{n}\sum_{j \geq i}^{n} w_{ij} x_i^{\alpha} x_j^{\alpha}}{T}=-\sum_{i=1}^{n}\sum_{j \geq i}^{n} \lambda_{ij} x_i^{\alpha} x_j^{\alpha} \tag{22}$$

になる. よって, λ_{ij} が

$$\lambda_{ij}=-\frac{w_{ij}}{T} \tag{23}$$

で与えられることがわかる. 式(23)は式(19)の十分条件となっている.

以上より, ユニット間の相関一定のもとでエントロピーを最大にするという観点からボルツマンマシンをとらえると, ネットワークの結合荷重 w_{ij} および温度 T はラグランジュ乗数 λ_{ij} に対応することが示された. 例えば, 学習時に行われる焼きなまし (simulated annealing) は, 制約に対するラグランジュ乗数を最初は小さくしておき, 徐々に大きくすることに対応していることがわかる.

4. 保存量の導出

次に保存量を導く. まず, 式(9)と式(10)の制約のもとで, エントロピー S を最大にするような確率分布は,

$$P(\boldsymbol{x}^{\alpha})=\frac{1}{(2\pi)^{n/2}|\boldsymbol{C}|^{1/2}}\exp\left(-\frac{\sum_{i=1}^{n}\sum_{j=1}^{n}|\boldsymbol{C}|^{ij} x_i^{\alpha} x_j^{\alpha}}{2|\boldsymbol{C}|}\right) \tag{24}$$

である[4]. 但し, $\boldsymbol{C}$ は相関 C_{ij} を要素とする相関行列, $|\boldsymbol{C}|$ は行列式, $|\boldsymbol{C}|^{ij}$ は行列 $|\boldsymbol{C}|$ の i 行 j 列の余因数を表す. 式(24)は平均 0, 共分散行列 $\boldsymbol{C}$ の正規分布を与える. 更に, このときのエントロピーの最大値は次のように $|\boldsymbol{C}|$ を用いて,

$$S=\ln\sqrt{(2\pi e)^n|\boldsymbol{C}|} \tag{25}$$

となることがわかっている[4]. ここで, 式(8)と式(24)を比較すると,

$$\frac{|\boldsymbol{C}|^{ij}}{|\boldsymbol{C}|}=\begin{cases}-\dfrac{w_{ij}}{T} & (i \neq j)\\[2ex] \dfrac{2w_{ij}}{T} & (i=j)\end{cases} \tag{26}$$

が成り立つことがわかる．式(23)と式(26)から，

$$\lambda_{ij}=\begin{cases}\dfrac{|\boldsymbol{C}|^{ij}}{|\boldsymbol{C}|} & (i \neq j)\\[2ex] -\dfrac{2|\boldsymbol{C}|^{ij}}{|\boldsymbol{C}|} & (i=j)\end{cases} \tag{27}$$

となる．

一方，式(14)を式(11)に代入すると，

$$S=\lambda\sum_{\alpha}P(\boldsymbol{x}^{\alpha})+\sum_{\alpha}P(\boldsymbol{x}^{\alpha})\sum_{i=1}^{n}\sum_{j\geq i}^{n}\lambda_{ij}x_i^{\alpha}x_j^{\alpha} \tag{28}$$

となる．ここで，式(9)と式(10)を用いると，

$$S=\lambda+\sum_{i=1}^{n}\sum_{j\geq i}^{n}\lambda_{ij}C_{ij} \tag{29}$$

と書ける．ボルツマンマシンの定常状態では式(27)により λ_{ij} の値は与えられるので，これと式(17)から，

$$\ln\sqrt{(2\pi e)^n|\boldsymbol{C}|}=\ln Z+\frac{1}{|\boldsymbol{C}|}\sum_{i=1}^{n}\left(\sum_{j\geq i}^{n}|\boldsymbol{C}|^{ij}-2\sum_{j=i}^{n}|\boldsymbol{C}|^{ij}\right)C_{ij} \tag{30}$$

となる．これを整理して，

$$\sum_{i=1}^{n}\left(\sum_{j\geq i}^{n}|\boldsymbol{C}|^{ij}-2\sum_{j=i}^{n}|\boldsymbol{C}|^{ij}\right)C_{ij}-|\boldsymbol{C}|\ln\sqrt{(2\pi e)^n|\boldsymbol{C}|}=-|\boldsymbol{C}|\ln Z \tag{31}$$

を得る．熱力学における保存量の表記法[5]

$$U-TS=F \tag{32}$$

と比較することにより，式(31)の左辺第1項は内部エネルギー U を，右辺は自由エネルギー F を，また相関行列の行列式 $|\boldsymbol{C}|$ は温度 T をそれぞれ表していると考えることができる．このようにしてユニット間の相関が与えられると保存量が定義できることが示された．

5. むすび

本論文では，ボルツマンマシンの定常状態をエントロピー最大という観点から解析し，ボルツマンマシンを制約付きの最適化問題としてとらえると共に，保存量を導出できることを示した．ボルツマンマシンの学習は定常状態を用いてなされるので，本論文で示した結果は学習過程にも適用可能である．それによると，ボルツマンマシンの学習における焼きなましや結合荷重を変化させることは，実はラグランジュ乗数を変化させていることに対応する．今後，更にボルツマンマシンの最適化問題への応用，ならびに学習について研究を進めていく予定である．

文 献

(1) Ackley D. H., Hinton G. E. and Sejnowski T. J. : "A learning algorithm for Boltzmann Machines", Cognitive Sci., **9**, pp. 147-169 (1985).
(2) 甘利俊一，太原育夫："認知情報処理", pp. 74-76,オーム社 (1991).
(3) Herman H. 著，牧島邦夫，小森尚志共訳："協同現象の数理", pp. 61-64,東海大学出版会 (1986).
(4) 瀧 保夫："情報論 I", p. 174, 岩波全書 (1978).
(5) 小出昭一郎："熱学", p. 66, 東京大学出版会 (1984).

(平成4年7月1日受付，9月14日再受付)