ニューラルネットワークで巡回セールスマン問題を解く

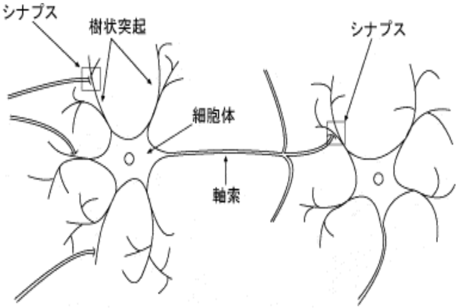
藤本勇希

名古屋大学多元数理科学研究科

2016/04/21

ニューラルネットワークとは

人間の脳の働きの一部を模倣することにより、知的な情報処理の実現を目指したものである。



各ニューロンは、他のニューロンの出力をシナプスの強さに応じて入力として受け取り、その入力に応じて発火するかどうか決定する。

ニューロンのモデル

ニューロンの発火のモデルは様々なものがあるが、出力がデジタルかアナログかで分類できる。

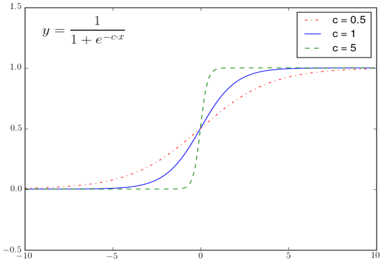


Figure: アナログ出力モデル

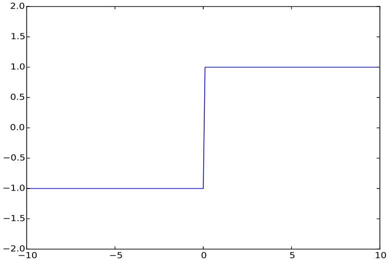


Figure: デジタル出力モデル

ゲインのスケールが大きければ大きいほど、出力の増加が急になる。
デジタル出力モデルは、ゲインのスケールの極限として解釈できる。

ニューラルネットワークの分類

大きく分けて 2 種類のネットワーク構造がある。

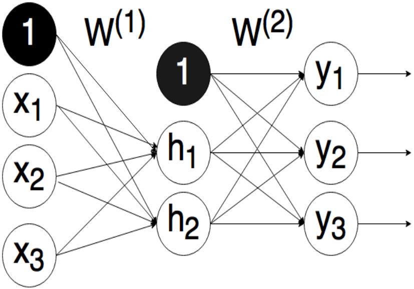


Figure: フィードフォワード型

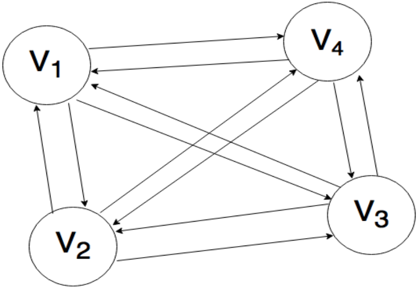


Figure: フィードバック型

パターン認識 (回帰、分類)

連想記憶
組み合わせ最適化

ホップフィールド・ネットワークとは

フィードバック型のニューラルネットワークのモデルの 1 種で、次の性質を持つものである。

1. 対称的な相互作用をもつ
2. 非同期的に状態を更新する

i 番目のニューロンの状態 u_i と出力 V_i は、次の規則で更新される。

$$\frac{du_i}{dt} = -\frac{u_i}{\tau} + \sum_j T_{ij} V_j + I_i \quad (1)$$

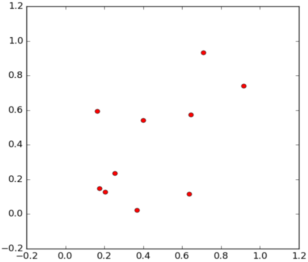
$$V_i = g(u_i) \quad (2)$$

自然な操作により、ネットワークのエネルギーが極小の状態に収束する。

巡回セールスマン問題（TSP）とは

巡回セールスマン問題（Traveling Salesman Problem）

都市の集合と各 2 都市間の移動コストが与えられたとき、全ての都市をちょうど一度ずつ巡り出発地に戻る巡回路のうち、総移動コストが最小のものを求めよ。



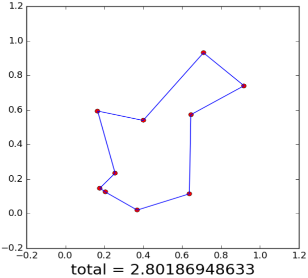
N = 10

都市の数 N に対して、異なる長さを持つ経路が $\frac{N!}{2N}$ 通り存在する。

巡回セールスマン問題（TSP）とは

巡回セールスマン問題（Traveling Salesman Problem）

都市の集合と各 2 都市間の移動コストが与えられたとき、全ての都市をちょうど一度ずつ巡り出発地に戻る巡回路のうち、総移動コストが最小のものを求めよ。



都市の数 N に対して、異なる長さを持つ経路が $\dfrac{N!}{2N}$ 通り存在する。

ニューラルネットによる解法の概要

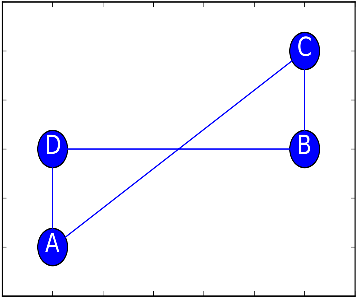
N 都市の巡回路を、N^2 個のニューロンの 0 — 1 出力で符号化する。

ホップフィールド・ネットワークの力学系では、力学系の安定点がエネルギー関数の極小点になる。

エネルギー関数の極小点と総移動コストの小さいの巡回路を対応させることができれば、収束先として良い巡回路を復号化することができる。

ニューラルネットワークによる巡回路の表現

i 番目に都市 X に訪れることを、$V_{Xi} = 1$ とする。(それ以外は 0)



$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

このとき N 都市の巡回路は、N^2 個のニューロンからなる
$N \times N$ の置換行列 $\{V_{Xi}\}$ で表現できる。

エネルギー関数への要請

力学系の安定点が表す巡回路が

1. 有効な巡回路を表す。（置換行列になっている）
2. 総移動コストが小さい。

という性質を満たすようにエネルギー関数を定義する。

エネルギー関数の性質 1

1 有効な巡回路を表す。

$$\frac{A}{2}\sum_X\sum_i\sum_{j\neq i}V_{Xi}V_{Xj}+\frac{B}{2}\sum_i\sum_X\sum_{Y\neq X}V_{Xi}V_{Yi}+\frac{C}{2}\sum_X\sum_i(V_{Xi}-n)^2 \tag{3}$$

ここで、V_{Xi} はニューロンの (X, i) 成分、 A, B, C は定数である。

各項は、次が成り立つときにのみ 0 となり、それ以外で正の値。

(a) 都市を訪れる回数は高々 1 回。

(b) 1 度に訪れる都市は高々 1 つ。

(c) 全部で n 回訪れる。

したがって、(3) 全体で巡回路の有効性を表す。

エネルギー関数の性質 2

2 総移動コストが小さい。

$$\frac{D}{2}\sum_{X}\sum_{Y\neq X}\sum_{i} d_{XY}V_{Xi}(V_{Y,i+1}+V_{Y,i-1}) \tag{4}$$

ここで、d_{XY} は 2 都市 X, Y 間の移動コスト、D は定数である。

(4) は、$\{V_{Xi}\}$ が置換行列のとき、総移動コストと一致する。

エネルギー関数の定義

エネルギー関数 E を (3) + (4) で定義する。

エネルギー関数

$$
\begin{aligned}
E := \frac{A}{2}\sum_X\sum_i\sum_{j\neq i} V_{Xi}V_{Xj} + \frac{B}{2}\sum_i\sum_X\sum_{Y\neq X} V_{Xi}V_{Yi} \\
+ \frac{C}{2}\sum_X\sum_i (V_{Xi} - n)^2 \\
+ \frac{D}{2}\sum_X\sum_{Y\neq X}\sum_i d_{XY}V_{Xi}(V_{Y,i+1} + V_{Y,i-1})
\end{aligned}
\tag{5}
$$

定数 A, B, C が大きいとき、巡回路の有効性は保証されるが、総移動コストが小さくならないことがある。
逆に、定数 D が大きいとき、総移動コストが小さくなるが、巡回路の有効性が保証されない。

結合行列と外力

エネルギー関数 (5) を u_{Xi} で偏微分することにより、 結合行列 $T_{XiYj}(N^2 \times N^2$ 行列) と外力 I_{Xi} が得られる。

結合行列と外力

$$
\begin{aligned}
T_{XiYj} = & -A\delta_{XY}(1-\delta_{ij}) \\
& - B\delta_{ij}(1-\delta_{XY}) \\
& - C \\
& - Dd_{XY}(\delta_{j,i+1} + \delta_{j,i-1})
\end{aligned}
\tag{6}
$$

$$
I_{Xi} = +Cn \tag{7}
$$

運動方程式

任意時刻の各ニューロンの出力 V_{Xi} は、次の運動方程式によって決定する。

運動方程式

$$
\begin{aligned}
\frac{du_{Xi}}{dt} = \frac{-u_{Xi}}{\tau} - A\sum_{j\neq i} V_{Xj} - B\sum_{Y\neq X} V_{Yi} \\
- C(\sum_{X}\sum_{j} V_{Xj} - n) \\
- D\sum_{Y} d_{XY}(V_{Y,i+1} + V_{Y,i-1})
\end{aligned}
\tag{8}
$$

$$
V_{Xi} = sigmoid(u_{Xi}) = \frac{1}{2}(1 + tanh(\frac{u_{Xi}}{u_0})) \tag{9}
$$

一般性を失わずに、$\tau = 1$ として良い。
u_0^{-1} は、シグモイド関数のゲイン。

アルゴリズムの実装

ネットワークの状態 $\{V_{Xi}\}$ を次のステップによって更新する。

エネルギー最小化アルゴリズム

1. 更新するニューロンのインデックスをランダムに選択する。(非同期型更新)
2. u_{Xi} を次のように更新する。

$$u_{Xi} = -A\sum_{j\neq i} V_{Xi} - B\sum_{Y\neq X} V_{Yi} - C\sum_{Y}\sum_{j}(V_{Yj} - n) - D\sum_{Y} d_{XY}(V_{Y,i+1} + V_{Y,i-1}) \quad (10)$$

3. ニューロンの状態 V_{Xi} を更新する。

$$V_{Xi} = \frac{1}{2}(1 + \tanh(\frac{u_{Xi}}{u_0}))$$

ここで、u_0 は温度。

アルゴリズムの難しさ

このアルゴリズムには、6 つのパラメータがある。(A, B, C, D, n, u_0) 有効かつ距離の短い経路を見つけるためには、このパラメータをどのように決めるのかが重要である。

ここで、n は都市の個数でなくても良い。論文では、都市の個数 10 に対して、$n = 15$ としている。

また、u_0 は温度で、収束のスピードと探索空間の広さのトレードオフを調節する役割がある。時間が立つにつれて温度を下げる (ゲインを増加させる) ことにより、探索空間の広さと収束性の両方を確保する方法がある。(Simulated Annealing)

初期状態の決め方

n を都市の個数とする。
経路についてバイアスがない場合、

$$V_{Xi} = \frac{1}{n} \quad (\forall X, i) \tag{11}$$

$$\sum_X \sum_i V_{Xi} = n \tag{12}$$

を初期状態にするのが良さそう。

しかし、同値な経路が $2n$ 個あるので、対称的だとそのうちの一つを選ぶことができない。(完璧に垂直に立てた鉛筆は倒れない)
したがって、各要素にランダムにノイズを加える。

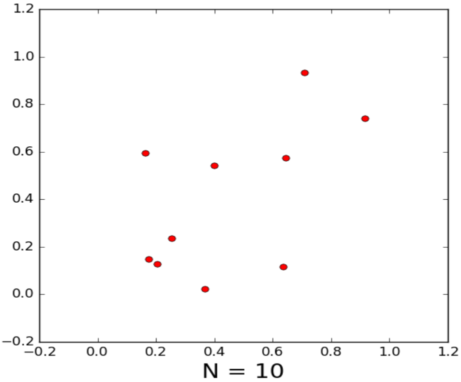
$$u_{00} = g^{-1}(\frac{1}{n}) \tag{13}$$

$$u_{Xi} = u_{00} + \delta_{Xi} \tag{14}$$

$$-0.1u_{00} < \delta_{Xi} < 0.1u_{00} \tag{15}$$

アルゴリズムの検証

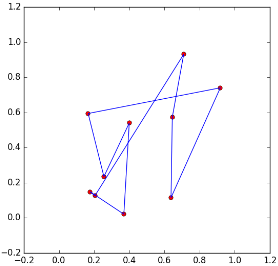
10 個の都市を $[0, 1]^2$ にランダムに配置



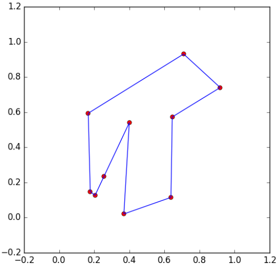
パラメータは論文と同じく
$A, B, D = 500, C = 200, n = 15, u_0 = 10$ に設定。(u_0 のみ調節を行った)
ひとつのセルに対して、10 回更新を行った。

検証の結果

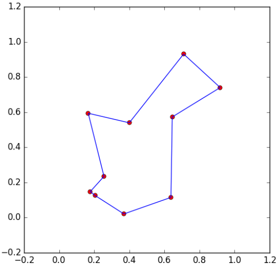
1. ランダムな巡回路
2. 得られた巡回路
3. 最小な巡回路



total = 4.71529788027



total = 3.44274777559



total = 2.80186948633

検証の結果

30 %の確率で巡回路として解釈でき、平均的にそこそこ短い巡回路が得られた。

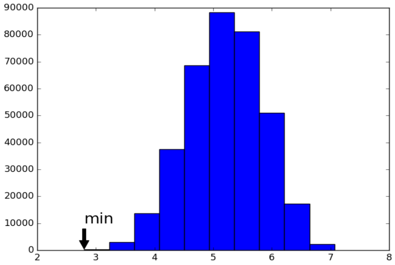


Figure: すべての巡回路の長さの頻度表

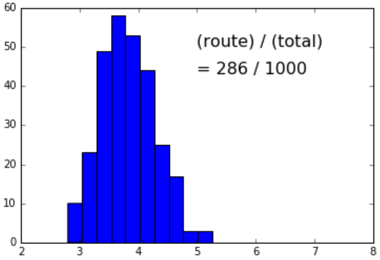


Figure: 得られた巡回路の長さの頻度表

今後の課題

1. シミュレーティド・アニーリング (Sree Aiyer[1991])
2. ボルツマンマシン (確率的に出力を決定)
3. フィードバックに遅れの効果を入れる。
4. TSP 以外の組み合わせ最適化問題への応用 (暗号の解読、　グラフ分割問題 etc)

興味あること

1. 生物の挙動をモデル化して、シミュレーションすることで、知的な振る舞いを再現・応用すること。
2. 局所的で単純なルールだけから、全体として複雑な挙動を得ること (セル・オートマトン)
3. 確率論的プログラミング (確率を用いて効率的に問題解決)

参考文献

1 J. J. Hopfield D. W. Tank, “Neural” computation of decisions in optimization problems, 1985