ホップフィールド型ニューラルネットによる制約条件付クラスタリング

[+]舘　俊太　　[+]武藤佳恭

教師無しクラスタリングにおいて，各クラスターの標本数やクラスター中心までの最大距離などといった条件や事前知識を反映させる手法は現実問題への応用が大きい．この分野ではファジィクラスタリングや制約条件付き K-平均法が広く研究されているが，本稿ではニューラルネットを用いて最適化組み合わせ問題として定式化する方法を述べる．提案モデルは代表点と標本点の組み合わせに対応したニューロンからなる．ニューロンの内部値に制約条件を反映する制約項を設けてこれを最小化するように各ニューロンを動作させることにより，制約条件付きクラスタリングが実現できることを示す．また実験により他のニューラルネット手法に比べても振動解や局所解への収束が少ないことを示した．

Conditional clustering model by Hopfield Neural network

[+]Shunta Tate,　[+]Yoshiyasu Takefuji

In unsupervised clustering, the technique of using prior knowledge and setting conditions such as restriction of the number of samples of each cluster and the maximum distance from each sample to a cluster center, has the large application in solving real-world problem. We used Hopfield neural network and formulated clustering problem as an optimized combination problem. The neural network consists of a neuron matrix corresponding to the combinations of the clusters and data points. Neuron's equation of evolution includes constraint terms which reflect criteria of clustering. The experiment showed the effectiveness of the model and showed that compared with other neural network techniques, there is less degree of convergence to an oscillation status or local minimum.

1 はじめに

教師無しの非階層クラスタリング手法には K-平均法[1) 2)]，ファジィ c-平均法(FCM) [3)4)]，Kohonen の自己組織化マップ[5)]などがある．データの分類基準に制約条件を設ける手法としては FCM においては可能性クラスタリング[3)]が. K-平均法には制約条件付き K-平均法などが提案されているが[1)2)]，特に後者は初期値やデータセットに最終結果が左右されやすく，局所解に収束しやすい問題がある．本論文ではホップフィールド型ニューラルネット[6)]を用いて，制約条件付きクラスタリングを解く数理モデルを提案する．ホップフィールド型ニューラルネットを用いて組み合わせ最適化問題を解く研究は古くから多数あり[6)7)8)]，クラスタリング問題も，分類される標本データとクラスターの最適組み合わせ問題として定式化してニューラルネットで解くことができる．先行手法としては MNM(Maximum Neuron Model)を用いたクラスタリングと Kohonen の自己組織化的手法が挙げられる[5)9)]．

但しこの従来手法では分類の基準として，クラスターのサイズや標本値の分布の性質といった任意の制約条件や事前知識を設定することができなかった．またクラスタリング結果が発見的に求まるので，確率モデルのような統計的な基準で結果を評価をすることができなかった．このようなニューラルネット手法でも先験知識を利用でき，また現実世界の制約条件　－例えば、資源配置の問題で代表点までの距離の上限を置きたいなど－を定式化することのできる数理モデルを提案する．

2 アルゴリズム

今，p 次元実空間 R^p 内に n 個の特徴点 $X=\{x_1,\cdots,x_n\}$ が分布しているとき，これを c 個のクラスター $W=\{w_1,\cdots,w_c\}$ に分類する問題を考える．評価関数（エネルギーコスト）としてクラスターの代表点から各特徴点への距離の総和

$$E=\sum_i \sum_j d_{ij}v_{ij}, \quad v_{ij}\in\{0,1\} \qquad (1)$$

を定める．但し d_{ij} はクラスターi の代表点 w_i から x_j

[+]慶応義塾大学大学院政策・メディア研究科
Graduate school of Media & Governance, Keio University

へのユークリッドノルム

$$d_{ij} = \sum_{p} \left\| w_i^p - x_j^p \right\|^2 \qquad (2)$$

である．この d を要素とする行列をDとする．v_{ij} を要素とする V はどの点がどのクラスターに属するかを表現する 2 値を要素とする行列である．なお，w_i は以下

$$w_i = \sum_l x_l v_{il}(t) \ / \ \sum_l v_{il}(t) \qquad (3)$$

によって計算するとする. クラスタリングは上式(1)を最小化する V を求める問題となる． ここで V をニューロンの全出力状態と見なし，静止と発火の状態を 0 と 1 の２値に対応させる. このニューロンの出力 V を決定する膜電位として実数値 u_{ij} を要素とする U を考える． そして式

$$v_{kj}(t+1) = \begin{matrix} 1 & if\ u_{kj}(t) = \max[u_{ij}(t); \forall i], \\ 0 & otherwise \end{matrix} \qquad (4)$$

によって V を発火させるとする． V を便宜的に連続値と見なし式(1)を v_{ij} で偏微分すると

$$\frac{\partial E}{\partial v_{ij}} = d_{ij} \qquad (5)$$

となるがこれに負の係数 －δをかけた値を用いて U の t についての微分方程式を作る.

$$\frac{du_{ij}}{dt} = -dd_{ij} \qquad (0 < d < 1) \qquad (6)$$

これに従って U を発展させると評価関数 E を最小化することが期待できる．なお，クラスターの中心は下式のように(0<β<1)となる係数βを用いて逐次的に W(t)の値が変化するようにして振動状態に陥るのを避ける.

$$w_i(t+1) = w_i(t) + b(w_i - w_i(t)) . \qquad (7)$$

ここで、制約条件付きの評価関数 E として新たに以下のような式を考える.

$$E = l_1 \left| \sum_i \sum_j d_{ij} v_{ij} \right| + l_2 \left| \frac{1}{2} \sum_i (n s_i - \sum_j v_{ij})^2 \right| . \qquad (8)$$

右項はそれぞれ総距離の最小化と，クラスターのサイズ(即ち，あるクラスターに含まれる点の数)の制約を表現する制約項である．但し l_m は $\sum_m l_m = 1$ となる制約項のパラメーター， s_i は $\sum_i s_i = 1$ となるクラスターのサイズを制約する係数である．この他にも制約項としては代表点への距離の分散の最小化，代表点までの最大距離の制約などが考えられるがここでは割愛する．ここで上式(8)を同様に $v_{ij}(t)$ で偏微分して $u_{ij}(t)$ の更新式として求めると，

$$\frac{\partial E}{\partial v_{ij}} = l_1 d_{ij} + l_2 (\sum_j v_{ij} - n s_i) \quad . \qquad (9)$$

u_{ij} の式はここから一次のオイラー法を用いて

$$u_{ij}(t+1) = u_{ij}(t) - t (a_1 l_1 d_{ij} + a_2 l_2 (\sum_j v_{ij}(t) - n s_i) + a_3 h u_{ij}(t)) \qquad (10)$$

となる．α_n はニューロンの内部電位と出力の間の特性を示す係数である．なお制約項のうち，距離の最小化項は最適解においても 0 にはならないので, U が発散しないように減衰項 $h u_{ij}(t)$ を加えている．ηは１以下の減衰の時定数である．またτは更新の変化分の大きさを決める時定数である．下に本モデルのアルゴリズムを示す

ステップ 1　U と W を一様乱数で初期化する
ステップ 2　V を式(4)によって計算する
ステップ 3　W を式(3),(7)によって更新する
ステップ 5　D を式(2)によって計算する
ステップ 4　U を式(10)で更新する
ステップ 6　式(8)の値が収束するまでステップ 2 に戻って反復する

パラメータセットλを調整することによって制約条件が結果に与える影響を調節することができる．なお，クラスターサイズの制約が満たされた時に式(10)右辺の α_2 の項は 0 になるが，全クラスター数と全特徴点数が

同一でない限り距離の最小化項であるα_1の項は 0 になることは無いので，標本の点の分布次第によっては振動状態を起こすことがある．また，初期の繰り返し計算の段階でα_2の項が効き過ぎると距離の最小化が進まず，総距離について大域的に最適化されない解に収束し易い．このような局所解や振動解に陥ることを少なくさせるためには，反復計算に併せてα_nを変化させる必要がある．本稿ではパラメーターα_nを下のようなシグモイド関数を用いたアニーリングスケジュールによって変化させた．

$$a_1 = \frac{1}{1+e^{-r(t-t_0)}}, \qquad (11)$$

$$a_2 = a_3 = 1 - a_1 \ . \qquad (12)$$

（$r = 0.0125,\ t_0 = 200$，(図 1)）これによって十分な反復計算の回数を取れば動作が振動せずに収束することが確かめられた．

3 実験と考察

図２のように，ガウス関数の混合分布よりなるテストデータを用意した．それぞれクラスターの数は 3,5,7 の 3 種類あり，それぞれ 3 種類について一つのクラスターが 16 個,64 個,256 個 の特徴点よりなる計９種類のテストデータである。これを用いて３つの手法の比較テストを行った．パラメータはそれぞれ $l_1 = l_2 = 1$, β =0.5, η =0.1, τ =0.2 であり，提案手法に関しては各クラスターに含まれる特徴点の数が等しい($s_1=s_2=\ldots=s_c$)とする制約条件を与えた．3 つの手法で各９データにつき 200 試行の実験を行ったが，提案手法は９つの全てのデータの全ての試行において制約条件を満たした．特徴点から代表点までの距離の総和の度数分布を図 3 に示す．3 つのグラフは左よりそれぞれクラスターの数が３，５，７の結果で 1 クラスター当りの点の数は全て 64 個である．提案手法が初期値に依存せず，最頻値への収束率が良いことがわかる．また，クラスターサイズの制約条件を課せられた本手法の解の場合，制約の無い他の二手法よりも距離の総和が少し大きくなることも判る．

また距離の総和の 200 試行回の平均値と標準偏差を表１にまとめた．距離の数値は特徴点一つ当りで割った値である．参考の為にテストデータのガウス分布の中心に各代表点を固定して算出した数値も併せて示す．表より，9 のデータのうち７つにおいて提案手法の解の収束率が他の 2 手法より勝っていることが判る．クラスターの数が７且つサイズが 64 と 256 の場合のみ標準偏差が大きくなっているが，これは特徴点が増えて特徴点の密度が上がって代表点の移動量が小さくなり，

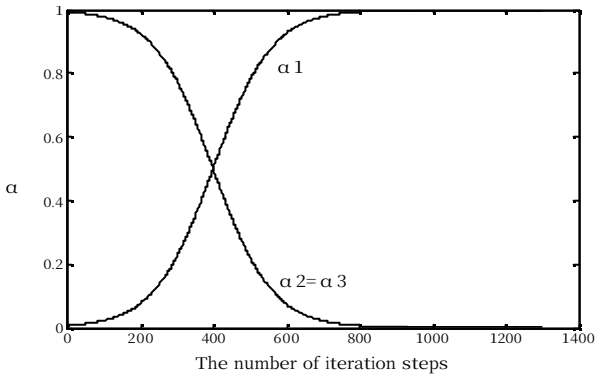

図 1.　アニーリングに用いるシグモイド関数

Fig.1 Sigmoid function which was employed for annealing.

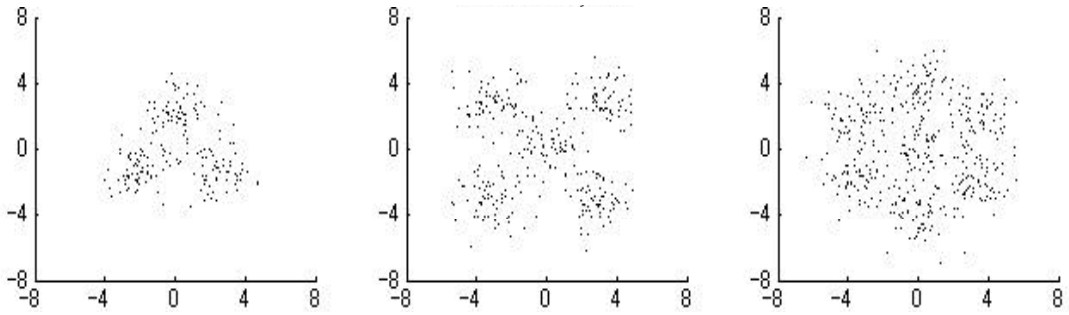

図 2.　テストデータの例．それぞれのクラスターの点の数は 64

Fig.2 Examples of data set for experiment. Each cluster is composed of 64 data point.

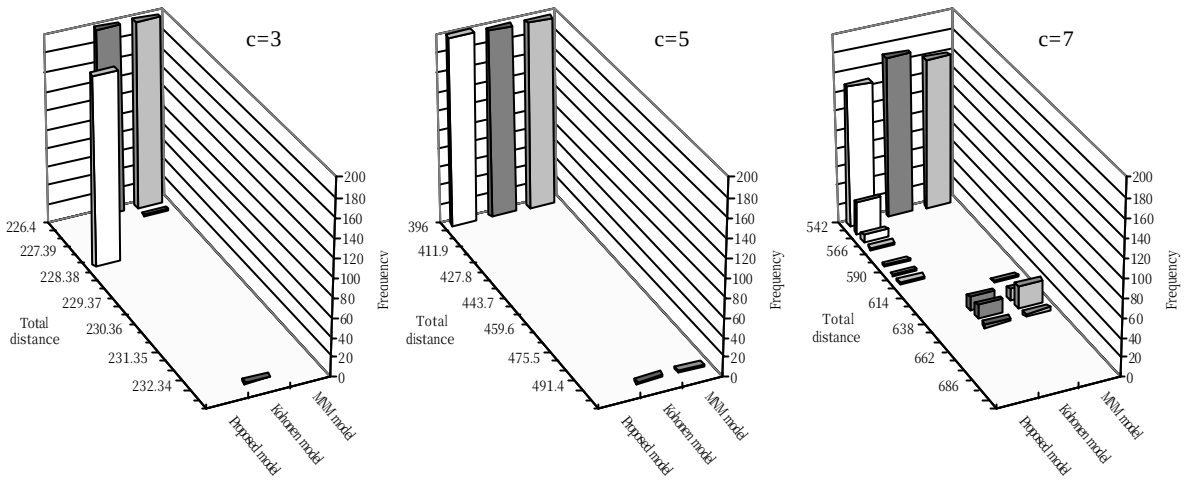

図 3 クラスタリング結果における総距離の度数分布
Fig. 3 Frequency distribution of the total distances of the clustering results.
From left to right, the number of the clusters is 3, 5, 7 respectively. Every cluster's density is 64.

The number of clusters	Method	Cluster density = 16		Cluster density = 64		Cluster density = 256	
3	The Proposed model	1.2641	(0.000)	1.1883	(0.016)	1.2394	(0.007)
	Kohonen's model	1.2658	(3.845)	1.1798	(0.457)	1.2388	(0.000)
	MNM model	1.2679	(3.723)	1.1797	(0.093)	1.2389	(0.043)
	Fixation on center	1.3357		1.1983		1.2405	
5	The Proposed model	1.0755	(0.269)	1.2388	(0.000)	1.2753	(0.073)
	Kohonen's model	1.0816	(9.348)	1.2409	(10.29)	1.2739	(0.109)
	MNM model	1.0780	(8.923)	1.2409	(10.30)	1.2739	(0.106)
	Fixation on center	1.0974		1.2456		1.2758	
7	The Proposed model	1.2483	(13.73)	1.2691	(43.98)	1.2446	(153.17)
	Kohonen's model	1.2432	(13.94)	1.2462	(34.03)	1.2279	(121.74)
	MNM model	1.2614	(14.36)	1.2574	(38.64)	1.2192	(98.20)
	Fixation on center	1.1509		1.2187		1.2060	

表 1. 200 試行での距離の総和の平均値(括弧内は標準偏差)
Table 1. The average cost of the total distances (and the standard deviation in the bracket) by 200 simulation runs.

各代表点の位置が十分最適化されずにサイズの制約条件の緩和計算が行われてしまったためと思われる．また，９のデータのうち６つの場合においてガウス分布の中心点からの数値(Fixation on center)よりも提案手法の値の方が小さい．これは与えられたデータについてはガウス分布の中心を多少外れた位置に総距離を最小化する代表点があり，提案手法がそこに収束していることを示している．

提案手法には現実の問題の様々な制約条件を応用させることができる．今後は確率分布推定との関連での評価や，またクラスター数と特徴点が増えたときに最適解への収束率が落ちることから，最適化が充分に進むための性質の良いパラメーターや境界条件について評価する必要があると考える．

1) Tung, A. K. H., Han, J., Lakshmanan, L. V. S. and Ng, R. T.: Constraint-Based Clustering in Large Databases, Proc. 2001 Int. Conf. on Database Theory (ICDT'01), London, U.K., (2001).
2) Lefkovitch, L. P.: Conditional clustering, Biometrics, 36, pp.43-58 (1980).
3) Krishnapuram, R., Keller, J. M.: A possibilistic approach to clustering, IEEE Trans. on Fuzzy syst., vol. 1, no. 2, pp. 98-110 (1993).
4) 宮本 定明，クラスター分析入門：ファジィクラスタリングの理論と応用，森北出版，東京 (1999).
5) Kohonen., T.: Self-Organizing Maps. Springer, Berlin (1997).
6) Hopfield, J. J., Tank, D. W.: Neural Computation of Decisions in Optimization Problems", Biological Cybernetics, 52, pp.141-152 (1985).
7) Takefuji, Y.: Neural Network Parallel Computing, Kluwer Publishers (1992)
8) Takefuji, Y., Oka, S.: Neural Computing for Solving Intractable Problems, The journal of the Institute of Electronics, Information, and Communication Engineers, Volume 79, Number 9 (1996).
9) Oka, S., Ogawa, T., T. Oda, and Takefuji, Y.: A New Self-Organization Classification Algorithm for Remote-Sensing Images, IEICE Trans. on Information and Systems, vol. E81-D, no. 1 (1998).