

論文メモ

文献番号	0001
日付	2021 年 9 月 21 日
名前	武川海斗

文献情報

著者	中森義輝
英文タイトル	なし
和文タイトル	ファジィクラスタリングと回帰分析
書誌情報	日本ファジィ学会誌, Vol. 8, No. 3, pp. 431-pp 439, 1996
キーワード	なし

1 論文の要約

回帰分析を用いたクラスタリング手法として, Fuzzy c - Regression Models(FCRM)[1] がある. FCRM では, ノイズ点とみなせるほど遠くにあるデータもクラスタに含めるため, 現実的でない結果になる場合がある. ノイズ点に弱い問題点を解決するために, Dave の手法 [2] を FCRM に組み込む事で解決を図った, 新しいモデルを提案した. 提案手法では, 初期クラスタ数が変化した場合にも, クラスタの形状に適応してうまくクラスタリングできることを実験で確かめた.

2 論文のトピック

本論文では, FCRM と適応型クラスタリングの考え方が重要となる.

3 ベースとなった手法

3.1 Fuzzy c -Regression Models

FCRM とはデータの分類と回帰モデル群の構築を同時に行う方法である. 与えられたデータを用いて, c 個のモデルを考える. 以下の式から導出され, この式から, U と β の交互最適化によってクラスタリングを行う手法である.

$$J(U, \beta_1, \dots, \beta_c) = \sum_{k=1}^n \sum_{i=1}^c (u_{ik})^m E_{ik}(\beta_i) \quad (1)$$

FCRM はクラスタ中心としていくつかの回帰超平面を仮定して, データとクラスタとの距離を回帰残差を用いて定義している. このために, 遠く離れた複数のデータが 1 つのクラスタに属することがある. これを解決するために, クラスタの形状を考慮し, 適切なクラスタ数を与えてあげることが肝要である.

3.2 適応型クラスタリング

適応型クラスタリングとは, クラスタの形状を動的に変更する手法である. ここで, G_{ik} がデータとクラスタのファジィ重心 v_i との 2 乗距離であるのに対し, V_{ik} はデータと線形多様体との 2 乗距離である. α_i は, ファジィ散布行列の

固有値から求める． α_i は， G_{ik} と V_{ik} の比率を定めており，各ラウンドごとに α_i を更新することで，動的にクラスタの形状を読み込んだクラスタリングを行うことができる．以下の式は適応型クラスタリングの評価規範である．

$$J_{afc} = \sum_{k=1}^n \sum_{i=1}^c (u_{ik})^m Z_{ik} \quad (2)$$

$$Z_{ik} = (1 - \alpha_i) G_{ik} + \alpha_i V_{ik} \quad (3)$$

この手法の利点として，あらかじめクラスタの形状に関する仮定をおかなくてよく，異なる形状のクラスタが存在するときでもうまくいく．

4 提案手法のコア要素

提案手法は，FCRM と適応型クラスタリングを組み合わせた手法である．

$$J(U; A_1, \dots, A_c; \bar{z}_1, \dots, \bar{z}_c) = \sum_{i=1}^c \sum_{k=1}^n (u_{ik})^m E_{ik}(A_i, \bar{z}_i) \quad (4)$$

$$E_{ik}(A_i, \bar{z}_i) = \alpha_i \times (\mathbf{y}_k - \mathbf{x}_k^o A_i)(\mathbf{y}_k - \mathbf{x}_k^o A_i)^T + (1 - \alpha_i) \times \eta \times (\mathbf{z}_k - \bar{\mathbf{z}}_i)(\mathbf{z}_k - \bar{\mathbf{z}}_i)^T \quad (5)$$

5 実験デザイン・結果と考察

実験では，クラスタ数を増やしても，クラスタの形状に適応してクラスタリングを行えることを確かめている．

6 手法の限界・今後の課題

手法の限界として，クラスタ数の決定については決定が困難なことが多い．また，説明変数の選択に関しても問題が多い．

7 特に重要な関連研究

参考文献 [1] は， c -回帰に関する論文であり，本論文の主題となる論文である．

また，参考文献 [2] は，適応型クラスタリングの手法について提唱した論文である．

次に読むべき論文のリスト

- [1] R. J. Hathaway and J. C. Bezdek, Switching Regression Models and Fuzzy Clustering, IEEE Trans on Fuzzy System Vol. 1, No. 3, pp. 195–pp. 204, 1993
- [2] R. N. Dave, An adaptive fuzzy c -elliptotype clustering algorithm, Quater Century of Fuzziness, Vol. I, pp 9–pp 12, 1990