Neurocomputing 202 (2016) 36–48

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Quantum clustering using kernel entropy component analysis

Yangyang Li [a,*], Yang Wang [a], Yuying Wang [b], Licheng Jiao [a], Yang Liu [c]

[a] Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education of China, International Research Center for Intelligent Perception and Computation, Xidian University, Xi'an 710071, PR China
[b] IBM (China) Investment Company Limited, Xi'an Branch, Global Business Solution Center, Xi'an 710000, PR China
[c] Department of Information Engineering, North China University of Water Resources and Electric Power, Zhengzhou 450045, PR China

ARTICLE INFO

Article history:
Received 26 June 2014
Received in revised form
27 October 2015
Accepted 15 March 2016
Communicated by E.W. Lang
Available online 13 April 2016

Keywords:
Quantum clustering
Kernel entropy component analysis
K-nearest neighbors

ABSTRACT

In this paper, a novel method quantum clustering using kernel entropy component analysis (KECA-QC) is proposed. This method has two phases: preprocessing and clustering stages. The main idea of preprocessing is to map the original data to a high-dimensional feature space, and to select the useful components using Renyi entropy as our similarity metric. After data preprocessing, different clusters will be distributed more or less in different places, and for high-dimensional datasets, it can achieve the purpose of dimensionality reduction at the same time. In the second phase, quantum clustering method is used, which can find clusters of any shape without knowing the number of clusters. Based on the traditional quantum clustering, we develop a new method estimating the wave function from distributions of *K*-nearest neighbors statistics, which can further reduce the running time and improve the calculation efficiency. In order to evaluate the effectiveness of this method, we compare the proposed method with *k*-means clustering (KM), the classical spectral clustering algorithm called Ng–Jordan–Weiss (NJW), the traditional QC, and kernel entropy component analysis spectral clustering algorithm (KECA-KM). The experimental results demonstrate that the proposed algorithm outperforms the compared algorithms on synthesized datasets and UCI datasets.

1. Introduction

Clustering analysis refers to the process of organizing data into meaningful homogeneous groups or clusters. Its aim is to get high similarity among the objects in the same cluster, but dissimilarity in different clusters. Clustering usually does not need a training data for learning, which is in an unsupervised classification. It has been widely used in pattern recognition, data mining and machine learning. Traditionally, the existing clustering methods can be divided into partitional, hierarchical, density-based, and model-based algorithms. Each algorithm has its own advantages and disadvantages. For example, *k*-means algorithm [1], one of the most famous clustering algorithms, can be implemented easily and efficiently, and have been successfully employed in many practical applications. However, this algorithm only works well for hyper-spherical data, or at best hyper-elliptical data. In the existing clustering algorithms, including *k*-means algorithm, most partitional approaches require a-priori knowledge about the number of clusters the data should be grouped into. However, in many cases, the number of clusters is unknowable. Moreover, the vast majority of clustering methods have many drawbacks, such as sensitivity to their initialization and noise and susceptivity to local optima.

Inspired by quantum mechanism in physics, we can solve those problems mentioned above using quantum theory. Quantum clustering (QC) [2–4], proposed by David Horn and Assaf Gottlieb, takes the clustering as a physical system. By solving the Schrödinger equation and using the gradient descent method, the resulting potential function has the minimums which correspond to the cluster centers. If the parameters of QC are fixed, it is a deterministic algorithm and belongs to one of the unsupervised clustering algorithms based on partition. Many experiments show that QC can produce satisfactory results even when some traditional clustering algorithms fail.

Nasios et al. concluded that the potential field is assimilated with the data density [5,6]. So they used the *K*-nearest neighbors statistical distribution for estimating the scale parameter, and detected the final modes by combining the local Hessian with a region growing algorithm. Li et al. improved the QC algorithm and proposed the distance-based QC [7] and parameter-estimated QC [8], to overcome the defect of the traditional QC, that is, the measured distance between two samples is relatively fixed and the kernel scale parameter is often needed to be estimated by experiments many times. Subsequently, Zhang et al. substituted the exponent measuring distance for Euclidean distance to

* Corresponding author. Tel.: +86 02988202279.
E-mail address: yyli@xidian.edu.cn (Y. Li).

http://dx.doi.org/10.1016/j.neucom.2016.03.006
0925-2312/

measure the distance between data points and the cluster centers [9], which improved the iterative procedure of QC algorithm and performed better than the Euclidean distance in data preprocessing. Gou et al. made a combination of quantum clustering and multi-elitist immune algorithm [10,11] to avoid getting stuck in local extremes the and solve the computational bottleneck. In addition, quantum clustering combination with other methods can be applied to many different areas. In Refs. [8] and [11], the improved quantum clustering algorithms are applied for the topography segmentation from SAR images of terrain and medical images segmentation respectively. Niu et al. detected the community structure in complex networks based on quantum mechanics [12]. Sun et al. applied quantum clustering to the research of fuzzy neural network model [13]. Di Buccio et al. distilled relevant documents by means of dynamic quantum clustering [14].

In this paper, we introduce a new quantum clustering using kernel entropy component analysis [15,16]. Firstly, we use kernel entropy component analysis as a data preprocessing stage, which is to map the original data to a high-dimensional feature space and to select the useful components substituting original data. It is proved that all the information contained in the distribution of the data can be utilized by using entropy as our metric. Our approach is capable of finding clusters of any shape, without knowing the real number of clusters beforehand. The clustering results of datasets especially high-dimensional data are obviously improved after data preprocessing. In addition, quantum clustering may take a long period of time to execute. So in order to reduce the running time and improve the calculation efficiency, the distributions of K-nearest neighbors statistics are used for calculating the wave function of each quantum physics particle. Meanwhile, the introduction of local information can help improve the quantum clustering accuracy.

The rest part of this paper is organized as follows. In the next section, the quantum clustering algorithm is described. Section 3 gives the details of the proposed data preprocessing, KECA, and then the improved quantum clustering with K-nearest neighbors is described. Section 4 provides the experimental results and discussions about our algorithm. The conclusions of this study are drawn in the last part of this paper.

2. Related work – quantum clustering algorithm

Quantum mechanics describes the distribution of particles in quantum space, and clustering analysis detects the structure of samples in scale space. In this case, we can observe the similarity between data points and quantum particles. The study of quantum physicists found that the distribution of microscopic particles in the energy field is influenced by their potential energy. When the spatial distribution and evolution shrink down to a one-dimensional infinite square potential well, the particles always tend to gather together at low potential values [17]. In other words, potential function is equivalent to an abstract source, and there are more particles distributed in the potential well when the data potential function reaches minimum. Thus, quantum mechanics with this process can be used to solve clustering problems. Particles in the same cluster will eventually be placed in the same potential well but different wells for particles belonging to different cluster. There is corresponding relation between particles and data points, and we can convert the data clustering problem into a quantum clustering issues.

Quantum clustering algorithm [2–4] proposed by Horn is a novel clustering method that is based on physical intuition derived from quantum mechanics. The state of a quantum mechanics system is completely specified by the wave function [6]. According to the theory of quantum mechanics, the evolution of quantum states follows the Schrödinger equation. The Stationary Schrödinger equation can be represented as

$$H\psi = \left(-\frac{\sigma^2}{2}\nabla^2 + V(\boldsymbol{p})\right)\psi = E\psi \tag{1}$$

where $\psi(\boldsymbol{p})$ is the wave function, $V(\boldsymbol{p})$ is the potential function, H is the Hamilton operator, E is the energy eigenvalue of H, ∇ is the Laplacian operator, and σ represents the adjustable scale parameter.

From above we can conclude that particles have the same distribution state if they are at the same potential. If the wave function $\psi(\boldsymbol{p})$ is given, the potential function $V(\boldsymbol{p})$ can be worked out with the Schrödinger equation. The minima of $V(\boldsymbol{p})$ can be associated with cluster centers. The process above provides the basis for quantum clustering.

Such a wave function $\psi(\boldsymbol{p})$ can be estimated with the Gaussian kernel-based sum.

$$\psi(\boldsymbol{p}) = \sum_{i=1}^{N} e^{-\|\boldsymbol{p}-\boldsymbol{p}_i\|^2/2\sigma^2} \tag{2}$$

In total, there are N quantum particles in quantum space and $\boldsymbol{p}_i$ is one of them. Working out the Eq. (1) with the given $\psi(\boldsymbol{p})$, we can get the general expression of potential function as

$$V(\boldsymbol{p}) = E + \frac{(\sigma^2/2)\nabla^2\psi}{\psi} = E - \frac{d}{2} + \frac{1}{2\sigma^2\psi}\sum_i \|\boldsymbol{p}-\boldsymbol{p}_i\|^2 \exp\left[-\frac{\|\boldsymbol{p}-\boldsymbol{p}_i\|^2}{2\sigma^2}\right] \tag{3}$$

where d is the lowest possible eigenvalue of H, which is the dimension of data points [18]. For E is still left undefined, we require V to be nonnegative, i.e. $\min V = 0$. E can be approximated by

$$E = -\min\frac{(\sigma^2/2)\nabla^2\psi}{\psi} \tag{4}$$

Refs. [2,3] use the gradient descent algorithm to find the quantum potential minima as the cluster centers. The iterative formula is as following:

$$\boldsymbol{y}_i(t+\Delta t) = \boldsymbol{y}_i(t) - \eta(t)\nabla V(\boldsymbol{y}_i(t)) \tag{5}$$

where $\boldsymbol{y}_i(0) = \boldsymbol{p}_i$ is the initial data, $\eta(t)$ is the iteration speed and ∇V is the gradient of potential function. Eventually, data point $\boldsymbol{y}_i$ reaches a fixed value which is consistent to the cluster center. And some nearest data points should be classified into the same class.

It follows therefore that, the potential function is the cost function in QC. The location of the cluster centers in the clustering process depends on the potential structures of the data, but need not define the geometric centers or random data points as our cluster centers.

3. Quantum clustering using kernel entropy component analysis (KECA-QC)

3.1. Kernel entropy component analysis

In this subsection, the kernel entropy component analysis will be introduced. Then the performance of kernel principal component analysis (KPCA) [19,20] and singular value decomposition (SVD) [2,3], which are also widely used preprocessing methods, will be shown. And the comparison and discussion will be made with these three methods.

In 2006, from the view of information theory, Jenssen et al. combined Renyi entropy [15] with the kernel method and proposed a novel method called kernel entropy component analysis (KECA). It is able to retain the main information of data structure