Theoretical Model for a Special Type of Distance Based Clustering of Feature Points Based on Distance to Complement Feature Point or Orthogonal Feature Point of Each Feature Point

Ramesh Chandra Bagadi
Fellow, Institution of Engineers (F124-1135) MIG-905, Mithilapuri Colony, VUDA Layout, Madhurawada, Visakhapatnam 530041, Andhra Pradesh State, India

Abstract- In this research investigation, the author has detailed a scheme for a special type of distance based clustering of feature points of concern based on distance to complement feature point or orthogonal feature point of each feature point.

Index terms- Clustering, Distance Based Clustering

I. INTRODUCTION

There have been many propositions regarding Clustering Models, the major among them being [1], [2], [3], [4], [5]. Also, there have been a few propositions on Overlapping Clustering Models [6], [7].

II. PROPOSED THEORETICAL MODEL

Notion of the Complement of A Given Vector

For any given Vector $A = [x_1, x_2, x_3, \ldots, x_n]$ the Complement of this Vector is given by filling it with the Complement of each element w.r.t all other elements of the vector.

That is, the complement of $x_i$, namely $x_i^c$ is

Case 1: Only Complement the Weighted Average of all other elements of this Vector except $x_i$.

Case 2: Orthogonal Complement the Weighted Average of all other elements of this Vector except $x_i$, with a Sign to be fixed as follows:

If $\sum_{i=1}^{n} x_i x_i^c$ is Positive, then $x_{(j+1)}^c$ is chosen such that $x_{(j+1)}^c x_{(j+1)}$ is Negative. And if, $\sum_{i=1}^{n} x_i x_i^c$ is Negative, then $x_{(j+1)}^c$ is chosen such that $x_{(j+1)}^c x_{(j+1)}$ is Positive. Also, $1 < j \leq (n-1)$.

That is, $A^c = [x_1^c, x_2^c, x_3^c, \ldots, x_{n-1}^c, x_n^c]$.

The weight is given by

$w_{i=p} = \frac{x_i}{\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} w_i x_i}$

Notion of the Complement of A Given n Dimensional Matrix

The Complement of any element $A(p_1, p_2, p_3, \ldots, p_{n-1}, p_n)$ of an n Dimensional Matrix $A$ with dimension sizes $l_1, l_2, l_3, \ldots, l_{n-1}, l_n$ is given as follows:

Case 1: Simple Complement

The required value is given by

$$\left\{ \sum_{l_1=1}^{l_1} \sum_{l_2=1}^{l_2} \ldots \sum_{l_{n-1}=1}^{l_{n-1}} \sum_{l_n=1}^{l_n} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \sum_{l_{i,j,k}=1}^{l_{i,j,k}} \sum_{l_{i,j,k,l_{i,j,k,l}}} a_{i,j,k,l} \right\}$$

$\left( \prod_{i=1}^{n} l_i \right) - 1$

Where the weight term is given by
\[ w_{t_1, t_2, \ldots, t_n} = \frac{\sum_{i=1}^{l_1} \sum_{j=1}^{l_2} \ldots \sum_{k=1}^{l_n} a_{t_1, t_2, \ldots, t_n}}{\left( \prod_{i=1}^{n} l_i \right) - 1} \]

**Case 2: Orthogonal Complement**

The required value is given by

\[ \sum_{i=1}^{l_1} \sum_{j=1}^{l_2} \ldots \sum_{k=1}^{l_n} \left( w_{t_1, t_2, \ldots, t_n} a_{t_1, t_2, \ldots, t_n} \right) \]

Where the weight term is given by

\[ w_{t_1, t_2, \ldots, t_n} = \frac{\sum_{i=1}^{l_1} \sum_{j=1}^{l_2} \ldots \sum_{k=1}^{l_n} a_{t_1, t_2, \ldots, t_n}}{\left( \prod_{i=1}^{n} l_i \right) - 1} \]

And the sign of the term

\[ \sum_{i=1}^{l_1} \sum_{j=1}^{l_2} \ldots \sum_{k=1}^{l_n} \left( w_{t_1, t_2, \ldots, t_n} a_{t_1, t_2, \ldots, t_n} \right) \]

is given as follows:

\[ \sum_{i=1}^{l_1} \sum_{j=1}^{l_2} \ldots \sum_{k=1}^{l_n} \left\{ a_{t_1, t_2, \ldots, t_n} b_{t_1, t_2, \ldots, t_n} \right\} \]

If \( b_{t_1, t_2, \ldots, t_n} \) is Positive, then the sign of

\[ a_{t_1, t_2, \ldots, t_n} \] is chosen such that

\[ a_{t_1, t_2, \ldots, t_n} = \begin{cases} 1 & \text{if } \sum_{i=1}^{l_1} \sum_{j=1}^{l_2} \ldots \sum_{k=1}^{l_n} \left( w_{t_1, t_2, \ldots, t_n} a_{t_1, t_2, \ldots, t_n} \right) > 0 \\ 0 & \text{otherwise} \end{cases} \]

is Negative.

\[ \sum_{i=1}^{l_1} \sum_{j=1}^{l_2} \ldots \sum_{k=1}^{l_n} \left\{ a_{t_1, t_2, \ldots, t_n} b_{t_1, t_2, \ldots, t_n} \right\} \]

And if \( b_{t_1, t_2, \ldots, t_n} \) is Negative, then the sign of

\[ a_{t_1, t_2, \ldots, t_n} \] is chosen such that

\[ a_{t_1, t_2, \ldots, t_n} = \begin{cases} 1 & \text{if } \sum_{i=1}^{l_1} \sum_{j=1}^{l_2} \ldots \sum_{k=1}^{l_n} \left( w_{t_1, t_2, \ldots, t_n} a_{t_1, t_2, \ldots, t_n} \right) < 0 \\ 0 & \text{otherwise} \end{cases} \]

is Positive. It should be noted that here, \( b_{t_1, t_2, \ldots, t_n} \) represents the Complement of the Matrix Element of \( A \), namely \( a_{t_1, t_2, \ldots, t_n} \) in the Complement Matrix \( B \) which is the complement of Matrix \( A \).

**Special Type of Distance Based Clustering Using Distance to Complement of the Feature Point**

Let there be \( m \) number of feature points each of \( n \) dimensions. Let them be represented by \( \bar{x}^p \), where \( p = 1 \) to \( m \). Also, let the elements of the feature points be represented by \( x^p_{pq} \), where \( p = 1 \) to \( m \) and \( q = 1 \) to \( n \). We now find the weighted average of all these feature points which is just the feature point gotten by taking the weighted averages element-wise as follows

\[ r_x_q = \frac{\sum_{p=1}^{m} w_{pq} x^p_{pq}}{\sum_{p=1}^{m} w_{pq}} \]

\[ w_{pq} = \frac{x^p_{pq}}{\sum_{p=1}^{m} x^p_{pq}} \]

with

This weighted average point is represented by \( r\bar{x}_q \) indicating that it is the most representative point for all the given feature points. Its elements are represented by \( r_x_q \) where \( q = 1 \) to \( n \).

We now find the distances between this most representative point \( r\bar{x} \) and each of all other feature points. Let these be represented by \( d(\bar{x}_p, \bar{x}) \) for \( p = 1 \) to \( m \). We now arrange these distances in increasing order. Let this order be a function \( f \)

given by a map from the Set \( \{ p \}_{p=1}^{m} \) to the same

Set \( \{ p \}_{p=1}^{m} \) but with the possibility that the map need not be necessarily congruent but such that the increasing order of distances aspect is satisfactorily met. Let these distances be denoted by

\[ d(\bar{x}, \bar{x}_1), d(\bar{x}, \bar{x}_2), \ldots, d(\bar{x}, \bar{x}_m) \]
Now, if we need K number of clusters, we find the first K Number of points that are closest to the most representative point $\bar{x}$. That is, we consider the points $\bar{x}_{p=f_1^{-1}(1)}$, $\bar{x}_{p=f_1^{-1}(2)}$, $\bar{x}_{p=f_1^{-1}(3)}$ to their respective Complement points. Let these be represented by $g_1 = d(\bar{x}_{p=f_1^{-1}(1)}, \bar{x}_{p=f_1^{-1}(2)})$, $g_2 = d(\bar{x}_{p=f_1^{-1}(2)}, \bar{x}_{p=f_1^{-1}(3)})$, ..., $g_k = d(\bar{x}_{p=f_1^{-1}(k-1)}, \bar{x}_{p=f_1^{-1}(k)})$.

We now arrange these distances in increasing order.

Let this order be a function $f_2$ given by a map from the Set $\{g_k\}_{k=1}^K$ to the same Set $\{g_k\}_{k=1}^K$ but with the possibility that the map need not be necessarily conjure but such that the increasing order of distances aspect is satisfactorily met.

Let these distances be represented by $g_{h=f_2^{-1}(1)}$, $g_{h=f_2^{-1}(2)}$, $g_{h=f_2^{-1}(3)}$, ..., $g_{h=f_2^{-1}(K-1)}$, $g_{h=f_2^{-1}(K)}$.

We now consider the distance $g_{h=f_2^{-1}(1)}$ and the point corresponding to it, namely $\bar{x}_{p=f_1^{-1}(h=f_2^{-1}(1))}$ and find all points that bear distance greater than or equal to the distance $g_{h=f_2^{-1}(1)}$. Now these points along with the point $\bar{x}_{p=f_1^{-1}(h=f_2^{-1}(1))}$ comprise the First Cluster.

We now consider the distance $g_{h=f_2^{-1}(2)}$ and the point corresponding to it, namely $\bar{x}_{p=f_1^{-1}(h=f_2^{-1}(2))}$ and find all points that bear distance greater than or equal to the distance $g_{h=f_2^{-1}(2)}$. Now these points along with the point $\bar{x}_{p=f_1^{-1}(h=f_2^{-1}(2))}$ comprise the Second Cluster.

We now consider the distance $g_{h=f_2^{-1}(3)}$ and the point corresponding to it, namely $\bar{x}_{p=f_1^{-1}(h=f_2^{-1}(3))}$ and find all points that bear distance greater than or equal to the distance $g_{h=f_2^{-1}(3)}$ and less than or equal to the distance $g_{h=f_2^{-1}(2)}$. Now these points along with the point $\bar{x}_{p=f_1^{-1}(h=f_2^{-1}(3))}$ comprise the Third Cluster. In this fashion, we find all K number of Clusters. It should be noted that these Clusters may be Overlapping in nature.

In this fashion, we can even find m number of Clusters for the given m number of feature points.

Special Type of Distance Based Clustering Using Distance to Orthogonal Complement of the Feature Point

Let there be m number of feature points each of n dimensions. Let them be represented by $\bar{x}_p$, where $p = 1$ to $m$. Also, let the elements of the feature points be represented by $x_{pq}$, where $p = 1$ to $m$ and $q = 1$ to $n$. We now find the weighted average of all these feature points which is just the feature point gotten by taking the weighted averages element-wise as follows:

$$r_{X_q} = \frac{\sum_{p=1}^{m} w_{pq} x_{pq}}{\sum_{p=1}^{m} w_{pq}}$$

$$w_{pq} = \frac{x_{pq}}{\sum_{p=1}^{m} x_{pq}}$$

With

$$r_{x_q}$$

indicating that it is the most representative point for all the given feature points. Its elements are represented by $r_{x_q}$ where $q = 1$ to $n$.

We now find the distances between this most representative point $r_{\bar{x}}$ and each of all other feature points. Let these be represented by $d(\bar{x}_p, r_{\bar{x}})$ for $p = 1$ to $m$. We now arrange these distances in increasing order. Let this order be a function $f_1$ given by a map from the Set $\{p\}_{p=1}^m$ to the same...
Set \( \{p\}_{p=1}^m \) but with the possibility that the map need not be necessarily congruent but such that the increasing order of distances aspect is satisfactorily met. Let these distances be denoted by 
\[
\begin{align*}
d_{\{p\}_1} & \left( x \right) \\
d_{\{p\}_2} & \left( x \right) \\
\vdots & \\
d_{\{p\}_m} & \left( x \right)
\end{align*}
\]
Now, if we need \( K \) number of clusters, we find the first \( K \) Number of points that are closest to the most representative point \( \bar{x} \). That is, we consider the points \( \bar{x}_{p_1} = \bar{x}_{f_1}^{(1)} \), \( \bar{x}_{p_2} = \bar{x}_{f_1}^{(2)} \), \( \bar{x}_{p_3} = \bar{x}_{f_1}^{(3)} \), \ldots , \( \bar{x}_{p_k} = \bar{x}_{f_1}^{(k)} \), \ldots , \( \bar{x}_{p_K} = \bar{x}_{f_1}^{(K)} \).

Now, we consider each of these \( K \) points and find the distances to their respective \textit{Orthogonal Complement} points. Let these be represented by 
\[
\begin{align*}
g_1 & = d_{\{p\}_1} \left( \bar{x} \right) \\
g_2 & = d_{\{p\}_2} \left( \bar{x} \right) \\
\vdots & \\
g_K & = d_{\{p\}_K} \left( \bar{x} \right)
\end{align*}
\]
We now arrange these distances in increasing order.

Let this order be a function \( f_2 \) given by a map from the Set \( \{g_h\}_{h=1}^K \) to the same Set \( \{g_h\}_{h=1}^K \) but with the possibility that the map need not be necessarily congruent but such that the increasing order of distances aspect is satisfactorily met.

Let these be represented by 
\[
\begin{align*}
g_{h_f_2^{(1)}} & \cdot g_{h_f_2^{(2)}} \cdot g_{h_f_2^{(3)}} \cdots \cdot g_{h_f_2^{(K-1)}} \cdot g_{h_f_2^{(K)}}
\end{align*}
\]
We now consider the distance \( g_{h_f_2^{(1)}} \) and the point corresponding to it, namely \( \bar{x}_{p_1} = \bar{x}_{f_2^{(1)}} \) and find all points that bear distance less than or equal to the distance \( g_{h_f_2^{(1)}} \). Now these points along with the point \( \bar{x}_{p_1} = \bar{x}_{f_2^{(1)}} \) comprise the First Cluster.

We now consider the distance \( g_{h_f_2^{(2)}} \) and the point corresponding to it, namely \( \bar{x}_{p_2} = \bar{x}_{f_2^{(2)}} \) and find all points that bear distance greater than or equal to the distance \( g_{h_f_2^{(1)}} \) and less than or equal to the distance \( g_{h_f_2^{(2)}} \). Now these points along with the point \( \bar{x}_{p_2} = \bar{x}_{f_2^{(2)}} \) comprise the Second Cluster.

We now consider the distance \( g_{h_f_2^{(3)}} \) and the point corresponding to it, namely \( \bar{x}_{p_3} = \bar{x}_{f_2^{(3)}} \) and find all points that bear distance greater than or equal to the distance \( g_{h_f_2^{(3)}} \) and less than or equal to the distance \( g_{h_f_2^{(4)}} \). Now these points along with the point \( \bar{x}_{p_3} = \bar{x}_{f_2^{(3)}} \) comprise the Third Cluster. In this fashion, we find all \( K \) number of Clusters. It should be noted that these Clusters may be Overlapping in nature.

In this fashion, we can even find \( m \) number of Clusters for the given \( m \) number of feature points.

### III. ACKNOWLEDGMENT

The author would like to express his deepest gratitude to all the members of his Loving Family, Respectable Teachers, En-Dear-Able Friends, Inspiring Social Figures, Highly Esteemed Professors, Reverence Deserving Deities that have deeply contributed in the formation of the necessary scientific temperament and the social and personal outlook of the author that has resulted in the conception, preparation and authoring of this research manuscript document. The author pays his sincere tribute to all those dedicated and sincere folk of academia, industry and elsewhere who have sacrificed a lot of their structured leisure time and have painstakingly authored treatises on Science, Engineering, Mathematics, Art and Philosophy covering all the developments from time immemorial until then, in their supreme works. It is standing on such treasure of foundation of knowledge, aided with an iota of personal god-gifted creativity that the author bases his foray of wild excursions into the understanding of natural phenomenon and forms new premises and scientifically surmises plausible laws. The author strongly reiterates his sense of gratitude and infinite indebtedness to all such ‘Philosophical Statesmen’ that are evergreen personal librarians of Science, Art, Mathematics and Philosophy.

### REFERENCES


