Distance Metric Learning for Haptic Data Classification of Braille Character Dataset

Megh Manoj Bhalerao[1]

1. *Undergraduate Student, Department of Electrical & Electronics Engineering, National Institute of Technology Karnataka, Surathkal, Mangalore 575025, India. email: 16ee234.megh@nitk.edu.in*

ABSTRACT: Machine Learning algorithms such as K-Means Clustering, K-Nearest Neighbors, Support Vector Machines (SVMs) etc. rely on a good metric to be specified over the input data. The standard Euclidean norm between data-points does not always give us a good measure of the relation between data points. Hence, it is crucial to specify a distance metric over the input data. This process of estimating a distance metric by utilizing information inherent to the data is called Distance Metric Learning (DML). Former approaches to DML make the use of class information to learn the distance metric over the dataset. Unlike the previous approaches, our approach makes the use of relative information of similarity or dissimilarity of the data-points rather than defining strict classes to separate the data-points. We obtain the information of similarity and dissimilarity of the data-points by using human perception of the same, and to learn a distance metric so as to mimic this human response. We apply DML upon a Braille Character Dataset and try to reconstruct the Human-Response Confusion Matrix using feature vectors and learned distance metric. We then project our data-points into a new vector space defined by our learned distance metric to minimize the distance between similar pairs and maximize the distance between the dissimilar pairs.

Keywords: *Convex Optimization, Haptics, Distance Metric Learning, Confusion Matrix, Braille Alphabet, Feature Vector*

INTRODUCTION & SCOPE

Distance Metric Learning is generally chosen as a preprocessing step to many machine learning algorithms. Xing et al. showed that the accuracy of K-means and constrained K-Means is brought to 100% after a distance metric is learned over the input data and the data-points are projected into a new vector space based on the learned metric. This aforementioned paper was the beginning of a solid research in DML. Here, DML is applied to the Braille Character Dataset by Jack M. Loomis. The requirements for a distance metric to be learned is a confusion matrix (also called a similarity matrix) and feature vectors corresponding to the data-points (obtained using some feature extraction techniques from the raw data).

The entire problem can be discretized into the following subsections:

1. Apply a least square regression DML algorithm for learning a distance metric on the Braille Character Dataset using different feature vectors corresponding to various Peano curves.
2. Estimate a confusion matrix corresponding to the Braille Character Dataset (using the feature vector and the learned distance metric), i.e. the pairwise distances between the feature vectors of each of the alphabets. Compare the estimated distance

matrix with the human response distance matrix.

3. Group the alphabets corresponding to the number of raised dots and restimate the confusion matrix (dimensions obviously reduce).

THE BRAILLE CHARACTER DATASET

The Confusion (Similarity) Matrix

This data is obtained from Jack Loomis' paper (1985). This confusion matrix is a 26 x 26 matrix where each of the row/ column corresponds to a letter of the English alphabet. Let this matrix be called C_b. The entries of this matrix are interpreted as follows:

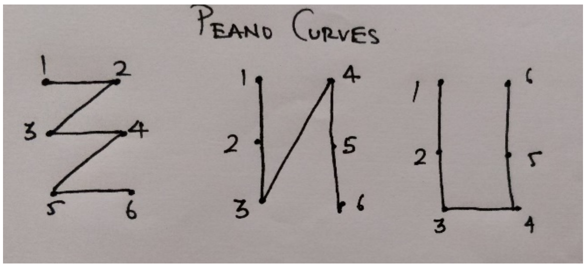
1. Consider any i^{th} row and j^{th} column of Cb i.e. (i, j) th entry of C_b.
2. C_b (i, j) indicates the fraction of people who responded with the j^{th} alphabet when the stimulus corresponding to the i^{th} alphabet was provided.
3. This tells us that the matrix need not be a symmetric matrix.
4. The matrix is converted to the standard form of a confusion matrix by symmetrizing it.

The Chosen Feature Vectors

The feature vectors are chosen by following various configurations of Peano Curves. Peano Curves are defined as continuous lines between a set of points, starting from a constant chosen point and do not cross each other. Here, Peano Curves are be drawn for the braille characters since they are nothing but a set of 6 points of which some are raised and the others aren't. The feature vectors are chosen according to the following procedure:

1. Peano Curves are drawn with the help of 6 dots of the braille characters.
2. The order in which the dots are connected decide the numbering of the dots in the braille alphabet.
3. The dots of the braille alphabet are numbered accordingly.
4. The feature vector for each alphabet is a six-dimensional vector with 1's at the places where the braille dots are raised and 0 otherwise.
5. Hence, the complete feature vector (say X_b) is a 26 x 6 vector with each row corresponding to the feature vector of one alphabet.

The feature vectors corresponding to Peano Curves of Row Stacking, Column Stacking and U-Shaped Stacking are taken for experimentation and shown below respectively.



CONVEX OPTIMIZATION PROBLEM FORMULATION

The metric learning algorithm for the Braille Dataset is formulated as a least squares' regression problem of the form:

$$argmin_M ||P^T MP - D_b||$$

subject to

$$M \geq 0$$

The terms are defined as follows:

1. P: The vector containing the difference between all the possible pairs of the feature vectors of the Braille Character Set.
2. M: The Distance Metric that is to be learnt.
3. D_b: The human response distance matrix.

CONVEX OPTIMIZATION ALGORITHM

As mentioned above the optimization problem here is a least squares regression algorithm and it has a closed form solution which is implemented using the algorithm given below:

Algorithm 1: Learning M using least squares regression - closed form solution

1. Input the feature vector (X_b) and the confusion matrix (Cb). Let the ith and jth entries of X_b be x_{bi} & x_{bj} respectively.
2. Define a Distance Matrix D_b = 1 - C_b. Let the (i, j) th entry of D_b be d_{bij}
3. Compute the pairwise distance between all the possible pairs of feature vector i.e. x_{bi} - x_{bj} and store it in a separate variable say P.
4. Our optimization problem is a least squares regression problem of the form

$$argmin_M ||P^T MP - D||$$

5. The problem is brought in its standard form by writing it in a slightly different manner i.e.

$$argmin_M ||Hm - D||$$

   where m is the unrolled vector of the matrix M.
6. The vector H is the vector containing the pairwise difference between all the possible pairs of the feature vectors.
7. The closed form solution of the above regression problem is given by

$$M = (H^T H)^{-1} H^T E$$

   where E is the unrolled column vector of the lower triangular part of the matrix D.

ACCURACY CHECK

The accuracy of our algorithm is measured by how relatable our estimated distance matrix is to the human response distance matrix. There is no single formula which would give us this measure of closeness/ accuracy. Rather, we have to analyze the relation using some graphical approaches. Ideally, the plot of the corresponding cell entries of the estimated distance matrix and the human response distance matrix must be a straight line of inclination of 45 degrees with the allowance for some minor deviations. But the graph we obtain is:

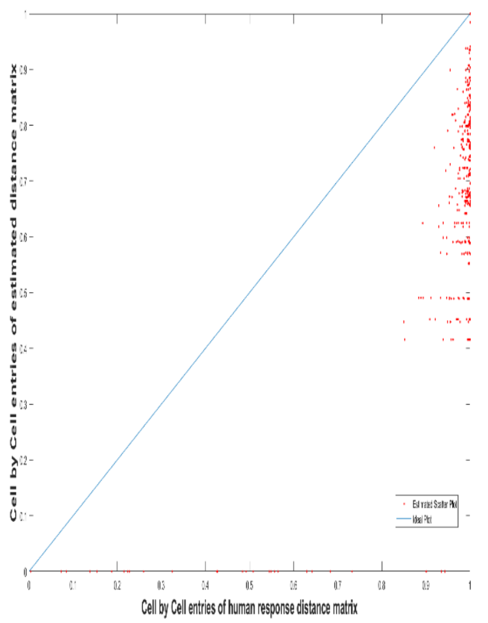


Fig 1: Y and X axes as shown above

The error of our estimate can be measured by how much our estimated data-points have deviated from the ideal plot. A way of doing this is to sum the perpendicular distances of our estimated points from the line. Mathematically it can be expressed as:

$$\in \;=\; \sum_i \frac{|x_i - y_i|}{\sqrt{2}}$$

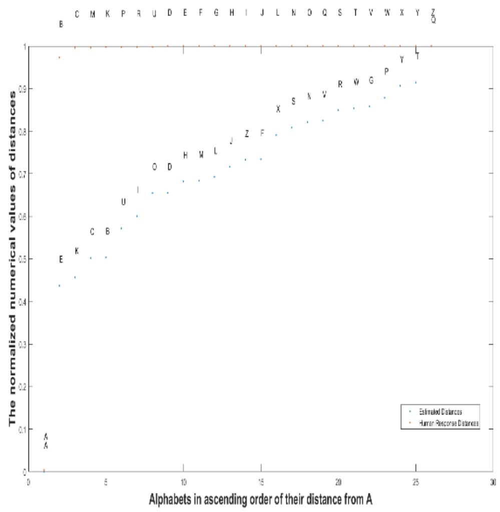
where (x_i, y_i) is the coordinate of the ith estimated datapoint.

RESULTS – LEAST SQUARES REGRESSION ALGORITHM

The estimated distance metric obtained using the least squares algorithm is the following: The closeness of the estimated distance matrix and the original human response distance matrix is given by a comparative study of the plots of the distance versus the letter.

The plots are shown below:

Fig 2: For Row Stacking
Fig 3: For Column Stacking
Fig 4: For U-Shaped Stacking



Explanation of the plots shown:

1. The X-axis depicts each letter of the English language spaced at equal intervals.
2. The order of letters is not alphabetical but rather the ascending order of the distances from the reference alphabet.
3. The reference alphabet in this case is the letter A.
4. The Y-axis depicts the numerical value of the distance of the reference letter with respect to the letter written above each of the data-point.
5. The Y-axis distance is normalized by dividing each of the entry in the estimated distance by the sum of entries of that row in the distance matrix to make all entries between 0 and 1.
6. Normalization is done since our human response distance matrix's entries are also between 0 and 1, as it depicts the fraction of people.

A Different Confusion Matrix
An observation of the above plots (of the distance of each character with respect to others) suggests that the similarity perception of the braille alphabet could depend upon different aspects of the alphabets such as:

1. The number of elevated dots
2. The orientation of the dots i.e. vertical or horizontal alignment.

Initially, we convert our 26 x 26 confusion matrix into a 5 x 5 confusion matrix by grouping the rows and columns suitably according to the groups. After that we sum all the entries of the corresponding group as well as one group with the other and place it in the corresponding columns of our new confusion matrix. After this we normalize each of the row of this new matrix so as to make all the entries sum up to 1. We symmetrize this matrix. The new 5 x 5 confusion matrix (D_f = 1 - C_f) is given below:

# of dots raised	1	2	3	4	5
1	0.0040	0.9919	0.9993	0.9996	1.0000
2	0.9919	0.0953	0.9021	0.9789	0.9899
3	0.9993	0.9021	0.2893	0.7901	0.9156
4	0.9996	0.9789	0.7901	0.3416	0.5676
5	1.0000	0.9899	0.9156	0.5676	1.0000

# of dots raised	1
1	0
2	0.
3	0.
4	1.
5	1.

After constructing D_f, we learn our distance metric M based on D_f and a set of new feature vectors (X_g) which consists of the pairwise distances of one alphabet group with itself and also one alphabet group with the other.

The convex optimization problem is:

$$argmin_M \; ||X_{sum} M_u - E||$$

subject to

$$M \geq 0$$

The terms are defined as follows:

1. X_{sum}: The vector containing the pairwise difference between all the possible pairs of the feature vectors (from different groups as

mentioned above) of the Braille Character Set.

2. M_u: The unrolled version of the Distance Metric that is to be learnt.
3. E: The elements of the reduced human response distance matrix arranges such that they correspond to the calculated distance (5 x 5).

The re-estimated distance matrix is the following:

The accuracy of such a methodology can be estimated as follows:

1. Generate all possible triplets (i, j, k) from the grouped-up alphabets such

   that (i, j) are in the same group and k is from a different group. Let the total number of triplets be N.
2. From all the possible triplets' corresponding feature vectors (x_i, x_i ,x_k) from the alphabet groups count the ones such that $d(x_i, x_j) < d(x_i, x_k)$ where d is the distance between the feature vectors calculated using the learned distance metric M. Let this count obtained be C.
3. The accuracy percentage can be calculated as:

$$Acc = \frac{C}{N} \times 100$$

4. The accuracy obtained by such a method for L2 norm minimization comes out to be around 46.46%

REFERENCES

[1] Xing, E P et al (2004), Distance Metric Learning with application to Clustering and Side Information, NIPS, 2003, pp. 521-528.

[2] Loomis J M, Tactile Recognition of Raised Character: A parametric study, Bulletin of the Psychonomic Society, 1985, Vol 31, No.2, pp. 18 - 20