Local Linear Embedding(LLE) and LLE Score Filter

Surupuendu and Abhijeet

DA-IICT

May 3, 2018

Aim of the Paper

Problem Statement

- To investigate the potential of LLE as a feature selection algorithm
- Introduction of LLE score to rank the features of face image dataset (ORL Dataset).

Introduction

Feature extraction and Feature selection

- Feature extraction algorithm reduces the dimensionality of data by projecting the data to lower dimensional subspace.
- Feature selection algorithm reduces the dimensionality by selecting a subset of the feature.

Shortcomings of PCA

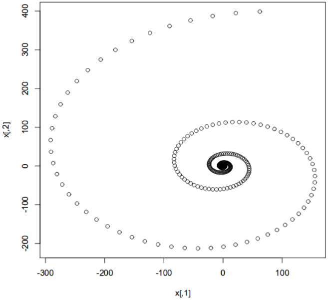

Fig:1 Spiral Data

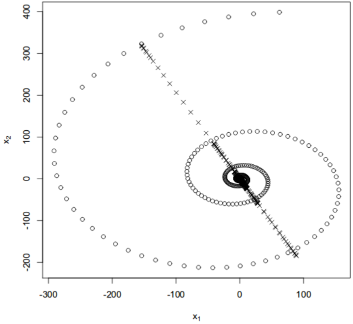

Fig:2 Data replotted with 1D PCA

LLE Algorithm

How does LLE work?

- It is a non linear manifold learning technique.
- It builds a neighbourhood of each point by searching k nearest neighbours in local region of data.
- It returns the low dimensional coordinates using the best reconstructed weights.

LLE Algorithm

1. LLE algorithm is given input as *nxp* data matrix X with rows x_i A desired number of dimensions $q < p$ An integer k for finding local neighbours , where $k >= q + 1$

$$RSS(\mathbf{w}) \equiv \sum_{i=1}^{n} \left\|\vec{x}_i - \sum_{j \neq i} w_{ij}\vec{x}_j\right\|^2 \tag{1}$$

2. The coordinates Y minimizes the reconstruction error using the weights

$$\Phi(\mathbf{Y}) \equiv \sum_{i=1}^{n} \left\|\vec{y}_i - \sum_{j \neq i} w_{ij}\vec{y}_j\right\|^2 \tag{2}$$

Computing the weights

$$G = min({w_i}^T (x_i e^T - v_i)^T (x_i e^T - v_i) {w_i}^T) \tag{3}$$

The above equation is the Gram matrix, G
where,x_i = Image point of size $nx1$
v_i = neighbour of x_i of size dxk
e = Matrix only containing ones of size $1xk$

$$w_i = \frac{eG^{-1}}{e^T G^{-1} e} \tag{4}$$

The above equation denotes the weight equation.

LLE as a Feature Selection Algorithm

In Feature selection algorithm, we recompute the weights in the lower dimension using the following equation

$$\min_{\{\hat{m}_{ij}^r, j \in \hat{J}_i\}} \| Y_i - \sum_{j \in \hat{J}_i} \hat{m}_{ij}^r Y_j \|^2 + \gamma \sum_{j \in \hat{J}_i} (\hat{m}_{ij}^r)^2 \tag{5}$$

where , γ is the coefficient of the regularizer term and is set to 10^{-5}
subject to the constraint

$$\sum_{j \in \hat{J}_i} \hat{m}_{ij}^r = 1 \tag{6}$$

LLE as a Feature Selection Algorithm

LLE Score

To reduce the features a new feature extraction method is introduced by the paper called LLE Score.

$$\mathrm{LLES}_r = \|\mathbf{M} - \hat{\mathbf{M}}^r\|_F^2. \tag{7}$$

where, $M = [w_{ij}]$
$M^r = [m_{ij}]$

MATLAB Execution

The input greyscale values of an image's pixel taken from each face dataset of 10 images.

10304x10 double

	1	2	3	4	5	6	7	8	9	10
1	48	34	60	39	63	64	43	41	44	42
2	45	35	58	44	56	60	45	44	44	41
3	45	34	68	59	52	63	57	48	41	54
4	49	34	79	54	40	53	52	44	45	48
5	46	38	49	62	43	60	45	50	48	43
6	47	36	43	69	43	41	28	46	39	40
7	45	30	52	74	29	30	33	46	45	46
8	47	27	52	61	30	28	43	44	40	45
9	48	24	56	44	30	38	57	46	33	41
10	53	27	72	47	24	55	52	50	36	41
11	52	23	68	41	23	54	64	51	31	44
12	50	20	77	35	31	49	69	47	28	44

MATLAB Execution

The output showing the reduced features onto which the image is mapped into after performing LLE.

Y | I1 | I

10x7 double

	1	2	3	4	5	6	7	8
1	-0.3805	-0.4343	0.4135	0.3931	0.1816	0.0370	0.3162	
2	0.2834	0.2787	-0.2795	-0.0843	0.4144	0.1573	0.3162	
3	0.2082	0.3128	0.2596	0.2176	-0.1908	0.3676	0.3162	
4	-0.2300	-0.0884	-0.6544	-0.0992	0.0986	-0.0339	0.3162	
5	-0.0686	-0.3310	-0.0127	-0.1806	0.2669	0.5408	0.3162	
6	0.5057	-0.5234	0.0521	-0.2011	-0.0757	-0.4936	0.3162	
7	-0.4456	0.3185	-0.0504	0.2077	0.1222	-0.5068	0.3162	
8	-0.3309	0.0635	0.0818	-0.5427	-0.6223	0.0884	0.3162	
9	0.1782	0.3717	0.4305	-0.2606	0.2568	-0.1891	0.3162	
10	0.2801	0.0319	-0.2405	0.5500	-0.4517	0.0322	0.3162	
11								
12								

Improvement on the LLE method

Improvements

- When the elements in the feature are all equal , the score comes out to be zero. So there is no discriminant information for classification.
- The Score is not scaling invariant

- "Principal Component Analysis", I.T Jolliffe
- "LLE Score: A New Filter-Based Unsupervised Feature Selection Method Based on Nonlinear Manifold Embedding and Its Application to Image Recognition" , Chao Yao , Ya- Feng Liu ,Bo Jiang, Jungong Han, and Junwei Han
- "Statistical pattern recognition " , A.K. Jain, R.P. Duin, and J.C. Mao
- "An Introduction to Local Linear Embedding" Sam T. Roweis , Lawerence K.Saul
- Nonlinear dimensionality reduction by locally linear embedding" S.T. Roweis and L.Saul