Local Linear Embedding(LLE) and LLE Score Filter

Surupuendu and Abhijeet

DA-IICT

April 30, 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

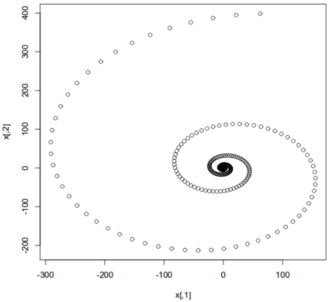

Fig:1 Spiral Data

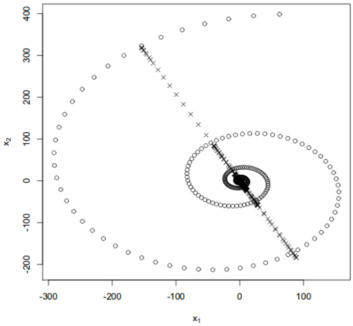

Fig:2 Data replotted with 1D PCA

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

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

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

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

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 Workspace Output

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

- Member, IEEE ,Bo Jiang, Jungong Han, and Junwei Han, Senior Member, IEEE
-