Supporting Information

Journal: Industrial & Engineering Chemistry Research

Title: Process monitoring based on orthogonal locality preserving projection with maximum likelihood estimation

Authors: Jingxin Zhang; Maoyin Chen; Hao Chen; Xia Hong; Donghua Zhou*

OLPP algorithm

Algorithm of OLPP consists of following essential steps as below [1].

Step 1: Establish the adjacency graph

Let G represent a graph with N nodes. Then, the adjacency graph is established via k-nearest neighbor, which can be utilized to evaluate whether an edge should be put between two nodes [2].

Step 2: Calculate the weights $\boldsymbol{S}$

$\boldsymbol{S} \in \mathcal{R}^{N\times N}$ is a similarity matrix and obviously sparse, where S_{ij} measures the similarity between $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$.

Step 3: Calculate the orthogonal locality preserving projections

The orthogonal locality preserving projections are represented by $\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k\}$ and are calculated below.

$$\boldsymbol{A}^{(k-1)} = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k] \tag{S1}$$

$$\boldsymbol{B}^{(k-1)} = \left[\boldsymbol{A}^{(k-1)}\right]^{\mathrm{T}} \left(\boldsymbol{X}\boldsymbol{D}\boldsymbol{X}^{\mathrm{T}}\right)^{-1} \boldsymbol{A}^{(k-1)} \tag{S2}$$

The vectors $\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k\}$ can be calculated iteratively thereinafter:

• Calculate $\boldsymbol{a}_1$ as the eigenvector of $\left(\boldsymbol{X}\boldsymbol{D}\boldsymbol{X}^{\mathrm{T}}\right)^{-1}\boldsymbol{X}\boldsymbol{L}\boldsymbol{X}^{\mathrm{T}}$ corresponding to the smallest eigenvalue.

• Calculate $\boldsymbol{a}_k$ as the eigenvector of the following matrix

$$\boldsymbol{M}^{(k)} = \left\{\boldsymbol{I} - \left(\boldsymbol{X}\boldsymbol{D}\boldsymbol{X}^{\mathrm{T}}\right)^{-1} \boldsymbol{A}^{(k-1)} \left[\boldsymbol{B}^{(k-1)}\right]^{-1} \left[\boldsymbol{A}^{(k-1)}\right]^{\mathrm{T}}\right\} \cdot \left(\boldsymbol{X}\boldsymbol{D}\boldsymbol{X}^{\mathrm{T}}\right)^{-1} \boldsymbol{X}\boldsymbol{L}\boldsymbol{X}^{\mathrm{T}} \tag{S3}$$

corresponding to the smallest eigenvalue of $\boldsymbol{M}^{(k)}$.

Step 4: Orthogonal locality preserving index embedding

Let $\boldsymbol{W}_{OLPI} = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_l]$, then data $\boldsymbol{x}$ can be mapped as

$$\boldsymbol{x} \to \boldsymbol{y} = \boldsymbol{W}_{OLPI}^{\mathrm{T}} \boldsymbol{x}, \tag{S4}$$

where $\boldsymbol{y}$ is an l-dimensional expression of raw $\boldsymbol{x}$.

Relationship between OLPP and PCA [3]

According to He et al [3], $\boldsymbol{X}\boldsymbol{L}\boldsymbol{X}^{\mathrm{T}}$ can be regarded as covariance matrix if the Laplacian matrix $\boldsymbol{L} = \frac{1}{N}\boldsymbol{I} - \frac{1}{N^2}\mathbf{1}\mathbf{1}^{\mathrm{T}}$, where $\boldsymbol{I}$ is the identity matrix and $\mathbf{1}$ is a vector of all ones with proper dimension. Under the circumstances, the weight matrix $\boldsymbol{S}$ has simple format, i.e., $S_{ij} = 1/N^2$, $\forall i, j$.

$D_{ii} = \sum_i S_{ij} = 1/N$. Let $\boldsymbol{m}$ denote the sample mean, i.e., $\boldsymbol{m} = \frac{1}{N}\sum_i \boldsymbol{x}_i$. It can be proved as follows:

$$\begin{aligned}
\boldsymbol{XLX}^{\mathrm{T}} &= \frac{1}{N}\boldsymbol{X}\left(\boldsymbol{I} - \frac{1}{N}\mathbf{1}\mathbf{1}^{\mathrm{T}}\right)\boldsymbol{X}^{\mathrm{T}} \\
&= \frac{1}{N}\sum_i \boldsymbol{x}_i\boldsymbol{x}_i^{\mathrm{T}} - \frac{1}{N^2}\left(N\boldsymbol{m}\right)\left(N\boldsymbol{m}\right)^{\mathrm{T}} \\
&= \frac{1}{N}\sum_i \left(\boldsymbol{x}_i - \boldsymbol{m}\right)\left(\boldsymbol{x}_i - \boldsymbol{m}\right)^{\mathrm{T}} + \frac{1}{N}\sum_i \boldsymbol{x}_i\boldsymbol{m}^{\mathrm{T}} + \frac{1}{N}\sum_i \boldsymbol{m}\boldsymbol{x}_i^{\mathrm{T}} - \frac{1}{N}\sum_i \boldsymbol{m}\boldsymbol{m}^{\mathrm{T}} - \boldsymbol{m}\boldsymbol{m}^{\mathrm{T}} \\
&= \mathrm{E}\left[\left(\boldsymbol{x} - \boldsymbol{m}\right)\left(\boldsymbol{x} - \boldsymbol{m}\right)^{\mathrm{T}}\right]
\end{aligned} \tag{S5}$$

where $\mathrm{E}\left[\left(\boldsymbol{x} - \boldsymbol{m}\right)\left(\boldsymbol{x} - \boldsymbol{m}\right)^{\mathrm{T}}\right]$ is exactly the covariance matrix of data points [3].

It is evidently observed that $\boldsymbol{S}$ is very significant for OLPP. When global geometric structure is expected to be preserved, we just need to set $k \to \infty$ and select eigenvectors corresponding to largest eigenvalues. In this case, data points are projected along the directions of maximal variance, which implies that OLPP is equivalent to PCA in a sense [3]. When we intend to preserve local geometric information, we need to set k small enough and reserve eigenvectors associated with smallest eigenvalues. Thus, data points are projected along the directions preserving locality. The latter one is more popular and essentially the core of OLPP.

Reference

[1] Cai, D.; He, X.; Han, J.; Zhang, H.-J. Orthogonal laplacianfaces for face recognition. *IEEE Transactions on Image Processing* **2006**, *15*, 3608–3614.

[2] M. Hasanlou, F. S. Comparative study of intrinsic dimensionality estimation and dimension reduction techniques on hyperspectral images using K-NN classifier. *IEEE Geoscience and Remote Sensing Letters* **2012**, *9*, 1046–1050.

[3] He, X; Yan, S; Hu, Y; Niyogi, P; Zhang, H.J. Face recognition using laplacianfaces. *IEEE Transactions on Pattern Analysis* & Machine Intelligence **2005**, *27*, 328–340.