Brief description of SLEM and FSLL manifold learning methods

Stochastic LEM Manifold Learning method

SLEM is a local stochastic ML method. In SLEM, the coefficients which construct each row ($\boldsymbol{w}_i$) of the neighborhood graph matrix between each data point and its neighbors have a PMF scheme and satisfy certain entropy as follows:

$$w_{ij} = \begin{cases} \dfrac{e^{-\frac{\|x_i - x_j\|^2}{2\sigma^2}}}{\sum_{j=1}^{k} e^{-\frac{\|x_i - x_j\|^2}{2\sigma^2}}}, & j \in N_i, S.t. H(\boldsymbol{w}_i) = H_0 \\ 0, & otherwise \end{cases} \tag{1}$$

where, w_{ij} is the coefficients between i'th data point and its j'th neighbor, $H(.)$ is the entropy operator, and the H_0 constant value can be determined by tuning step or mutual neighborhood criteria (for more details, refer to paper https://link.springer.com/article/10.1007/s10032-018-0303-4). In SLEM, the entropy value defines the locality around each data point and is related to the distribution of the data points. In fact, this condition causes the variance parameter (σ^2) for each data point is refined corresponding to the distribution of its neighbors.

After constructing the neighborhood graph matrix ($\boldsymbol{W}_{SLEM}$), the embedded data manifold are calculated by an optimization problem, same as LEM method.

(for more details, refer to paper https://link.springer.com/article/10.1007/s10032-018-0303-4).

Fusion of Stochastic LEM and LLE

In local ML methods, each method provides the embedded data in low-dimensional representation space by preserving the extracted structural information from the high-dimensional space. Therefore, the performance of these methods can be increased by extracting further structural information in the high-dimensional space. In LLE, each data point is locally presented by a linear combination of its neighbors. Also, SLEM describes each data points based on the values and distribution of its Euclidean distances to its neighbors. In this paper, a weighted fusion of the proposed SLEM and common LLE is introduced. Block diagram of the proposed FSLL method is given in figure 1. By fusion of SLEM and LLE (FSLL), we introduce the following cost function:

$$\begin{aligned} &\min_{\mathbf{Y}} \left(\sum_i \left(\lambda \left\| \boldsymbol{y}_i - \sum_{j \in N_i} w_{ij} \boldsymbol{y}_j \right\|^2 + \sum_{j \in N_i} \left\| \boldsymbol{y}_i - \boldsymbol{y}_j \right\|^2 w_{SLEM_{ij}} \right) \right) \\ &= \min_{\mathbf{Y}} \left(\lambda \underbrace{\sum_i \left(\left\| \boldsymbol{y}_i - \sum_{j \in N_i} w_{ij} \boldsymbol{y}_j \right\|^2 \right)}_{Term\ 1:\ LLE} + \underbrace{\sum_i \sum_{j \in N_i} \left\| \boldsymbol{y}_i - \boldsymbol{y}_j \right\|^2 w_{SLEM_{ij}}}_{Term\ 2:\ SLEM} \right) \end{aligned} \tag{2}$$

where λ is a regularization parameter.

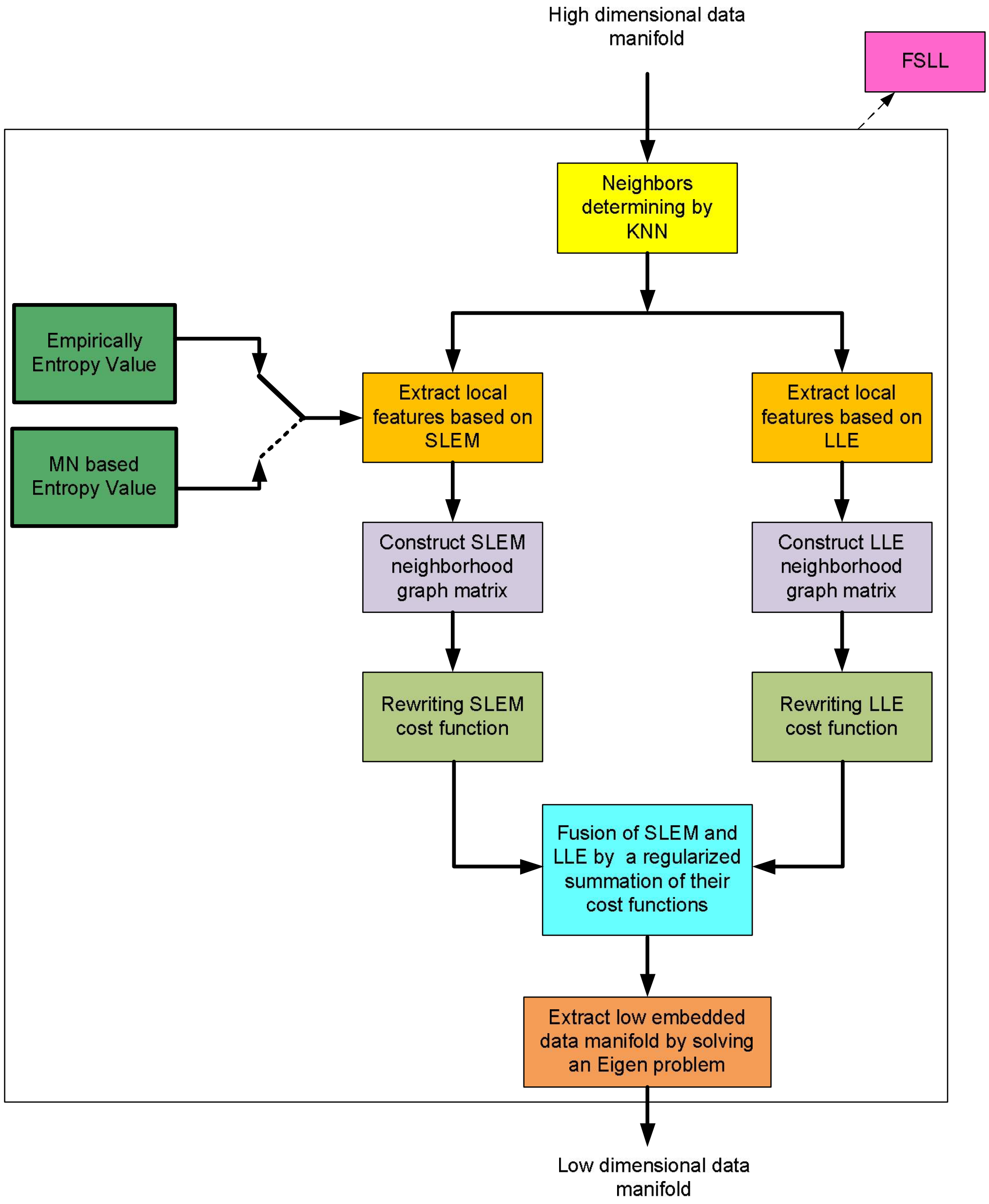

Fig 1. Block diagram of FSLL manifold learning method.

LLE part of the above cost function can be rewritten as [3]:

$$\sum_i \left(\left\| \boldsymbol{y_i} - \sum_{j \in N_i} w_{ij} \boldsymbol{y_j} \right\|^2 \right) = \boldsymbol{Y}^T (\boldsymbol{I} - \boldsymbol{W})^T (\boldsymbol{I} - \boldsymbol{W}) \boldsymbol{Y}. \tag{3}$$

Also, in the previous section, we show that SLEM can be rewritten as:

$$\sum_i \sum_{j \in N_i} \left\| \boldsymbol{y_i} - \boldsymbol{y_j} \right\|^2 w_{SLEM_{ij}} = \boldsymbol{Y}^T \boldsymbol{L_{SLEM}} \boldsymbol{Y}. \tag{4}$$

Therefore, the equation (4) can be rewritten as:

$$\sum_i \left(\lambda \left\| \boldsymbol{y_i} - \sum_{j \in N_i} w_{ij} \boldsymbol{y_j} \right\|^2 + \sum_{j \in N_i} \left\| \boldsymbol{y_i} - \boldsymbol{y_j} \right\|^2 w_{SLEM_{ij}} \right) = \lambda \boldsymbol{Y}^T (\boldsymbol{I} - \boldsymbol{W})^T (\boldsymbol{I} - \boldsymbol{W}) \boldsymbol{Y} + \boldsymbol{Y}^T \boldsymbol{L_{SLEM}} \boldsymbol{Y}$$
$$= \boldsymbol{Y}^T [\lambda (\boldsymbol{I} - \boldsymbol{W})^T (\boldsymbol{I} - \boldsymbol{W}) + \boldsymbol{L_{SLEM}}] \boldsymbol{Y} = \boldsymbol{Y}^T \boldsymbol{M_{FSLL}} \boldsymbol{Y}. \tag{5}$$

where $\boldsymbol{M_{FSLL}} = \lambda(\boldsymbol{I} - \boldsymbol{W})^T (\boldsymbol{I} - \boldsymbol{W}) + \boldsymbol{L_{SLEM}}$.

$\boldsymbol{M}_{FSLL}$ is the neighborhood graph matrix of FSLL that includes both structural features of LLE and SLEM. The proposed cost function minimization can be solved as an eigenvector problem subject to $\sum_i \boldsymbol{y_i} = 0$ and $\frac{1}{N}(\boldsymbol{Y}^T \boldsymbol{Y}) = \boldsymbol{I}$.

Generally, the proposed fusion scheme can be applied to every manifold learning method which its cost function can be rewritten as:

$$\underset{\boldsymbol{Y}}{arg\ min}\ \boldsymbol{Y}^T \boldsymbol{U} \boldsymbol{Y} \tag{6}$$

In equation (6), $\boldsymbol{U}$ is corresponding to the extracted features matrix in the high-dimensional space.

(for more details, refer to paper https://link.springer.com/article/10.1007/s10032-018-0303-4).