Dynamic regularization discriminant local preserving projection method for fault diagnosis

Yanlin He, Xiaona Yan, Qunxiong Zhu* and Yuan Xu*

College of Information Science & Technology, Beijing University of Chemical Technology, Beijing, 100029, China
*Corresponding author’s e-mail: zhuqx@mail.buct.edu.cn; xuyuan@mail.buct.edu.cn

Abstract. Due to the high dimensionality, serial correlation, and nonlinearity of industrial process data, the primary task for diagnosing fault is to extract key fault features from fault datasets. In this paper, to obtain much more inherent fault information, a dynamic regularization discriminant local preserving projection approach (DRDLPP) based on feature reduction is put forward to diagnose fault, which addresses the small sample size problem of discriminant local preserving projection (DLPP) by incorporating the regularization term into the objective function of DLPP. The enhanced performance of DRDLPP for fault diagnosis over conventional diagnostic approaches mostly benefits from two aspects: One aspect is that DRDLPP can discover local manifold fault information hidden in original sample space by preserving the local neighborhood structure of data; The other aspect is that DRDLPP has the remarkable capacity to capture dynamic information by extending the observation vector with previous observation vectors. What is more, the information criterion function is utilized to capture the optimal dimensionality reduction order and time lag of DRDLPP method. The experimental results of the Tennessee Eastman process demonstrate that the proposed DRDLPP approach provides a better visual performance and achieve lower misclassification rates in fault diagnosis.

1. Introduction
With modern industry presenting a new trend of large-scale and complex development, comprehensive industrial process monitoring technologies face new challenges. Once a fault occurs, it may cause huge economic loss and security problems. Therefore, an effective industrial fault diagnosis approach contributes to the improvement of the safety and stability of industrial production [1]. Among many fault diagnosis technologies, data-driven fault diagnosis techniques have shown high value in enhancing efficiency and production safety due to its efficiency and simplicity for many large-scale and complex industry processes.

Due to the redundancy of a large amount of information in industrial process data, it is critical to extract the data features to enhance fault diagnosis capability by using dimensionality reduction techniques. A variety of the dimensionality reduction algorithms proposed so far such as principal component analysis (PCA) [2], Fisher discriminant analysis (FDA) [3, 4] and Canonical variate analysis(CVA) [5], focus on the process data variables to diagnose faults. In multivariate statistical process analysis, the assumption imposed on the process data by researchers and pioneers for analysis simplification that an observation at a certain time is statistically uncorrelated with the “past” observations and the “future” observations. In other words, the process data are subject to independent and identical distribution. Yet this implicit assumption is usually invalid for almost any industrial production process owning to the intrinsic properties of the system. For better simulating the real
distribution characteristics of data samples, some extension methods based on traditional dimensionality reduction techniques taking serial correlation into account have been developed by using different enhancement vector techniques [6, 7].

Nevertheless, all aforementioned methods only take the global data information into account. Therefore, in order to extract local information distributed on local manifold structure of the original high dimensional sample space, some manifold learning methods based on the theory of differential geometry which have received considerable attention and been successfully used in the field of face recognition [8] and document clustering [9]. Discriminant Locality Preserving Projections (DLPP) was presented and first applied in face recognition [10]. Compared with traditional methods, DLPP not only can embed the manifold structure of original data into lower dimensional subspace, but also can discover the subspace that best separate different class information by maximizing inter-class separation and minimizing intra-class distance. Unfortunately, the DLPP algorithm has the serious Small Sample Size (SSS) problem, restricting the wide application of DLPP. Although PCA+DLPP, a two-stage approach like PCA+LPP, was used to handle SSS problem of DLPP [10], it could result in a more serious fact that it may not work as effectively in small number of fault classes.

In view of the aforementioned reasons, in order to address the SSS problem of DLPP, an approach incorporating regularization term into DLPP (RDLPP) is proposed for diagnosing faults. Additionally, in order to be closer to actual distribution characteristics of industrial process data, the enhanced RDLPP method is developed by constructing a dynamic matrix with lagged samples, which is called as dynamic RDLPP (DRDLPP). Finally, the function of information criterion combining model order and time lag is developed for capturing optimal dimensional order and time lags of DRDLPP approach.

2. Discriminant locality preserving projection method

Consider a dataset with $m$ samples $X = \{x_1^1, x_1^2, \cdots, x_1^n, x_2^1, \cdots, x_2^n, \cdots, x_C^1, \cdots, x_C^n\}$, in which $x_i^j \in \mathbb{R}^n$ denotes the $i$th class observation and $n_1 + n_2 + \cdots + n_C = m$. The goal of DLPP aims to seek the optimal transformation matrix to transform $x_i^j$ into $y_i^j$, such that

\[ y_i^j = W^T x_i^j, \quad i = 1, 2, \ldots, n_i, \quad c = 1, 2, \ldots, C \]  

The optimization problem of DLPP is described as follows:

\[ \max_{c \in [1, C]} \frac{\sum_{i=1}^{n_c} \sum_{j=1}^{n_i} \|m_i - m_j\|^2 B_{ij}}{\sum_{i=1}^{n_i} \sum_{j=1}^{n_i} \|y_i^j - y_j^j\|^2 S_{ij}} \]  

By bringing equation (1) into equation (2), the denominator of objective function can be derived as follows:

\[ \sum_{i=1}^{n_i} \sum_{j=1}^{n_i} (y_i^j - y_j^j)^2 S_{ij} = \sum_{i=1}^{n_i} \sum_{j=1}^{n_i} W^T (x_i^j - x_j^j) S_{ij} (x_i^j - x_j^j)^T W = 2 \sum_{i=1}^{n_i} \sum_{j=1}^{n_i} W^T x_i^j D_{ij}^C (x_j^j)^T W - \sum_{i=1}^{n_i} W^T x_i^j S_{ij} (x_i^j)^T W \]

where $S_{ij}^c$ is the similarity weight to indicate the similarity between data samples within same class label and can be defined as: $S_{ij}^c = \exp(-\|x_i^j - x_j^j\|^2 / t)$ where $t$ is appropriate constant and set as overall variance of the data in this paper; $D_{ij}^C$ is a diagonal matrix whose element can be defined as $D_{ij}^C = \sum_{c=1}^{C} S_{ij}^c$ and $L = D - S$, $X = [X_1, \cdots, X_C]$, $D = \text{diag}(D_1, D_2, \ldots, D_C)$ and $S = \text{diag}(S_1, S_2, \ldots, S_C)$.

Similarly, the simple algebraic calculation that brings equation (1) into equation (2), the numerator of equation (2) can be transformed as the following form:
where $B_y$ is the similarity weight to describe the similarity between samples in different classes and can be defined as: 

$$
B_y = \exp(-\left\| f_i - f_j \right\|^2 / t); \quad f_i \text{ is the mean vector of } i\text{th class samples, namely } f_i = \frac{1}{n_i} \sum_{i=1}^{n_i} x_i; 
$$

$E$ is a diagonal matrix, $E = \sum_j B_y$, $F = [f_1, f_2, \ldots, f_c]$ and $H = E - B$. Finally, equation (2) can be rewritten as the following form:

$$
J(W) = \arg \max_w \frac{\text{tr}[W^T F H F W]}{\text{tr}[W^T X L X W]} = \frac{\text{tr}[W^T S_H W]}{\text{tr}[W^T S_L W]} 
$$

(5)

where $S_H = F H F^T$ represents the between-class scatter matrix and $S_L = X L X^T$ represents the within-class scatter matrix. The optimal solution of equation (5) can be obtained by generalized eigenvalue decomposition:

$$
S_H w = \lambda S_L w 
$$

(6)

The projection vector $w$ is the eigenvector of the corresponding eigenvalue $\lambda$.

3. Fault diagnosis approach based on dynamic regularization discriminant local preserving projection method

3.1. Dynamic regularization discriminant local preserving projection method

Through the description of DLPP in the section 2, we can find DLPP can give a more discriminant information. However, in the practical calculation and analysis process, the solution of DLPP’s objective function often stuck in the matrix decomposition difficulty because of serious SSS problem. Specifically, rank of $S_H$ is at most $c$ ($c<n$), while is a matrix, which means that $S_H$ is singular. For handling the SSS problem of DLPP, a unified approach similar to FDA is proposed to incorporate a regularization term into the inter-class scatter matrix $S_H$. Instead of maximizing equation (5), the following regularized DLPP is optimized:

$$
J(W) = \arg \max_w \frac{\text{tr}[W^T (S_H + \epsilon I) W]}{\text{tr}[W^T S_L W]} 
$$

(7)

where $\epsilon \in (0,1)$ and $I$ is an identity matrix. By adjusting the appropriate $\epsilon$ in the inter-class scatter matrix $S_H$, DLPP can be solved while not add any numerical problems, then optimal projection matrix $W$ can be obtained. Compared with PCA+DLPP method in which dimensionality reduction order is at most $c$, it do not subject to $c$ restriction and can obtain a larger dimensionality reduction order than $c$. In reference [11], regularized discriminant analysis is investigated by Friedman with regard to the SSS problem. In fact, equation (7) is a special case of his regularized discriminant analysis. Consequently, the optimal transformation matrix $W$ maximizes the objective of equation (7) can be computed by solving the generalized eigenvalue decomposition problem:

$$
(S_H + \epsilon I) w = \lambda S_L w 
$$

(8)

Let $w_0, w_1, \ldots, w_c$ be the solution of equation (12), ordered according to their eigenvalues $\lambda_0 \geq \lambda_1 \geq \cdots \geq \lambda_{c-1}$. Thus, the transformation matrix $W$ is as follows:

$$
x \rightarrow y = W^T x 
W = [w_0, w_1, \ldots, w_c] 
$$

(9)

In addition, there is usually an assumption imposed on process data by researchers and pioneers for analysis simplification that an observation at a certain time is statistically uncorrelated with the “past” observations and the “future” observations. Nevertheless, this implicit assumption is usually invalid for most industrial processes due to the intrinsic properties of the system. Therefore, it is more practical and
effective to develop a fault diagnosis method considering the potential serial correlations of process data. For this, a dynamic data matrix is constructed in equation (10), which contains more dynamic information by combining each observation vector with the previous \( l \) samples,

\[
X = \begin{bmatrix}
X_k^T & X_{k-1}^T & \cdots & X_{k-l}^T \\
X_{k-1}^T & X_{k-2}^T & \cdots & X_{k-l-1}^T \\
\vdots & \vdots & \ddots & \vdots \\
X_{k+i-n}^T & X_{k+i-n-1}^T & \cdots & X_{k-n}^T
\end{bmatrix}
\]  

(10)

where \( X_k^T \in R^n \) is observation at time \( k \).

The RDLP model formulated from equation (1) to (9) can be utilized to incorporate dynamic information of data for faulty classification, then we call this method Dynamic Regularization Discriminant Local Preserving Projection Method (DRDLPP).

3.2. Information criterion

In this paper, the information criterion function includes two parameters, namely dimension reduction order \( a \) and time lag \( l \). A feasible method on basis of Information Criterion (IC) presented by Akaike [12] to determine the optimal value for both parameters. The formula is as follows:

\[
IC = f(a,l) + \frac{a}{n}
\]  

(11)

where \( f(a,l) \) indicates the error classification rate under the order \( a \) and time lag \( l \), \( n \) is the average number of samples in each fault dataset. Although the misclassification rate of the training set is significantly reduced with the increase of \( a \), misclassification rate of the training set is usually first reduced and then increased after the certain order \( a \). Meanwhile, computational complexity will become larger and larger. The second term added in equation (11) measures the degree of model complexity to punish unlimited increase dimension just considering the lower misclassification rate while not computational complexity.

3.3. The steps of the DRDLPP method for fault diagnosis

The overall steps of our proposed DRDLPP-based fault diagnosis method can be roughly divided into two parts, namely training step and testing step. The detailed steps are as follows:

Training step:

Step 1: Stack observation vector in the way as equation (10) and determine the order of dimensional reduction \( a \) and the time lag \( l \) according to the information criterion function in equation (11).

Step 2: Calculate the two similarity matrices, \( S_i \) and \( S_{ii} \), and solve the generalized eigenvalue decomposition of equation (8).

Step 3: Obtain the transformation matrix \( W \) consisting of the first \( a \) projection vectors according to equation (9).

Testing step:

Step 4: Construct new data vectors in the form as equation (12), project each dataset according to equation (11), and utilize discriminant function based on Bayesian framework which can be described as:

\[
g_j(x) = -\frac{1}{2}(x - \bar{x}_j)^TW\left(\frac{1}{n_j-1}W^T S_j W\right)^{-\frac{1}{2}}W^T(x - \bar{x}_j) - \frac{1}{2}\ln\left[\det\left(\frac{1}{n_j-1}W^T S_j W\right)\right]
\]  

(12)

And an new observation \( x \) is classified into class \( j \), \( g_j(x) > g_i(x) \) if \( \forall i \neq j \).

4. Application to Tennessee Eastman process

This section investigates the performance of the proposed DRDLPP approach in terms of visual illustration and classification accuracy, in comparison with commonly FDA and RDLP methods.
without dynamic information, for fault diagnosis based on case study from Tennessee Eastman process (TEP). The process flowsheet of TEP with closed-loop control structure is shown in figure 1 and more detailed information about TEP can be obtained in reference [13].

Two data sets were produced for each fault, namely training data set and testing data set. The training data set and test data set for each fault contain 480 and 800 observations with every 3 min sample time, and a total of 52 process variables are recorded in each observation, including all manipulated and measurement variables except the reactor stirrer’s agitation speed. The test data containing fewer observations is used for establishing fault diagnosis model, and training test sets containing more observations are used to verify the performance of the established diagnostic model.

In this case study, IDV 2, 5, 8, 12, 13, and 14 from TEP are used to verify the performance of DRDLPP, RDLPP and FDA. Detailed information on the six faults is listed in table 1.

Table 1. Detail description of the selected six faults

| ID     | Fault description                  | Type         |
|--------|------------------------------------|--------------|
| IDV 2  | B composition, A/C ratio constant  | Step change  |
| IDV 5  | Condenser cooling water inlet temperature | Step change |
| IDV 8  | A, B, C feed composition           | Random variation |
| IDV12  | Condenser cooling water inlet temperature | Random variation |
| IDV 13 | Reactor kinetics                   | Slow drift   |
| IDV 14 | Reactor cooling water valve        | Sticking     |

Table 2. Classification results for DRDLPP, RDLPP and FDA for IDV 2, 5, 8, 12, 13 and 14

| ID     | Misclassification rate (%) | FDA    | RDLPP | DRDLPP |
|--------|----------------------------|--------|-------|--------|
| IDV 2  | 22.50                      | 23.50  | 3.76  |
| IDV 5  | 2.50                       | 2.13   | 2.01  |
| IDV 8  | 43.38                      | 28.50  | 34.09 |
| IDV12  | 34.88                      | 21.88  | 15.91 |
| IDV 13 | 74.13                      | 42.25  | 40.98 |
| IDV 14 | 11.00                      | 3.63   | 3.51  |
| Overall misclassification rate (%) | 31.40 | 20.32 | 16.71 |

4.1. Results on the visual performance

In this section, the visual performance of the DRDLPP, RDLPP, and FDA approaches are analysed and discussed. The first three main features for IDV 2, 5, 8, 12, 13, and 14 are extracted by three methods and 3-D visual scatter plots and histograms are employed to show the projection results in figure 3. The histograms are drawn from three groups of inter-class distance points generated by the above three methods in ascending order from left to right and the distance points are normalized to be [0, 1], where the two black double lines are thresholds at the upper 25% and 75% quantiles. As shown in figure 3(a),
figure 3(b) and figure 3(c), the scatter plot generated by DRDLPP has better visual performance than two other methods (RDLPP and FDA) in terms of clustering the same class and separating different classes. In particular, it is worth mentioning that the visual performance of DRDLPP proposed are better than RDLPP without dynamic information, which further demonstrates that the dynamic matrix considering high serial correlations help improve visual performance. Additionally, the histograms (see figure 3(d), figure 3(e) and figure 3(f)) show clearly that a set of distance data produced by DRDLPP has a more even distribution and a larger proportion of large inter-class distances. All in all, this enables the faulty datasets with small between-class distances to be better separate by the proposed DRDLPP method than other two methods.

![Figure 3](image)

Figure 3. Visual results and between-class distances for IDV 2, 5, 8, 12, 13, and 14 produced by DRDLPP, RDLPP and FDA: (a) - (c) Visual result for DRDLPP, RDLPP and FDA, respectively. (d) - (f) between-class distances for DRDLPP, RDLPP and FDA, respectively.

### 4.2. Results on the fault diagnosis accuracy

We evaluate the diagnosis performance of the DRDLPP, RDLPP and FDA on the above six faults. From figure 2, the dimensional reduction order $a$ and the time lag $l$ for the DRDLPP method are optimized as $a=10$ and $l=1$ according to the IC equation (11). In order to make a fair comparison, the dimension reduction order of the other two methods is also selected as $a=10$. The diagnostic results for testing data are shown in table 2. From overall view, the proposed DRDLPP method performs better than FDA and RDLPP. The overall misclassification rate for DRDLPP is the lowest and is 16.71%, especially compared to 31.40% for FDA, which is a factor of about 2.0 enhanced diagnostic performance for DRDLPP. After giving special attention to IDV 2, surprising result can be found that the misclassification of DRDLPP is 3.76%. Compared to 22.50% to FDA and 23.50% to RDLPP, it is a great improvement for diagnostic performance. The main reasons for performance improvement are concerned with the following two aspects: (1) The one is that the proposed DRDLPP is a manifold learning technique that can extract local neighborhood structure the original data and map it into subspace, while FDA only consider the global information of the data, thereby enabling FDA to be easily affected by ‘outlier’ fault categories. (2) The other major aspect is that the dynamic information from the process data is well utilized by the proposed DRDLPP approach based on the incorporation of the time-lag technique.
5. Conclusion
Generally, a dynamic regularization discriminant local preserving projection approach (DRDLPP) is presented for diagnosing process faults, which solves the small sample size problem of DLPP by integrating a regularization term into the objective function of DLPP. Compared with traditional fault techniques, DRDLPP not only can discover the geometrical structure of the latent manifold but also maximizes the scatter between different classes and minimizes the scatter within same class simultaneously. In addition, DRDLPP has the superior ability to capture dynamic information from process data by adding previous sample vectors. The experiment results on Tennessee Eastman process demonstrate that proposed DRDLPP method obtains a better visual performance and achieves higher diagnostic accuracy.

Acknowledgments
This research is supported by the National Natural Science Foundation of China under Grant Nos. 61573051 and 61473026 and Fundamental Research Funds for the Central Universities under Grant Nos. JD1914 and XK1802-4.

References
[1] Yin, S., Ding, S.X., Xie, X., Luo, H. (2014) A review on basic data-driven approaches for industrial process monitoring. IEEE Transactions on Industrial Electronics, 61(11): 6418-6428.
[2] Papaioannou, A., Zafeiriou, S. (2014) Principal component analysis with complex kernel: the widely linear model. IEEE Transactions on Neural Networks & Learning Systems, 25(9): 1719-1726.
[3] Chiang, L.H., Kotanchek, M.E., Kordon, A.K. (2004) Fault diagnosis based on fisher discriminant analysis and support vector machines. Computers & Chemical Engineering, 28(8): 1389-1401.
[4] Jin, X., Zhao, M., Chow, T.W.S., Pecht, M. (2014) Motor bearing fault diagnosis using trace ratio linear discriminant analysis. IEEE Transactions on Industrial Electronics, 61(5): 2441-2451.
[5] Xiaoang, D., Xuemin, T. (2006) Nonlinear process fault diagnosis based on kernel canonical variate analysis. Control & Decision, 21(10): 1109-1113.
[6] Jiang, B., Huang, D., Zhu, X., Yang, F., Braatz, R.D. (2015) Canonical variate analysis-based contributions for fault identification. Journal of Process Control, 26: 17-25.
[7] Qin, S.J., Zhou, D., Li, G. (2014). A new method of dynamic latent-variable modeling for process monitoring. IEEE Transactions on Industrial Electronics, 61(11): 6438-6445.
[8] Weiwei, Y. (2009) Two-dimensional discriminant locality preserving projections for face recognition. Pattern Recognition Letters, 30(15): 1378-1383.
[9] He, X., Cai, D., Liu, H., Ma, W.Y. (2004) Locality preserving indexing for document representation. In: International Conference on Research and Development in Information Retrieval. United Kingdom. pp. 96–103.
[10] Yu, W., Teng, X., Liu, C. (2006) Face recognition using discriminant locality preserving projections. Image and Vision Computing, 24(3): 239-248.
[11] Friedman, J.H. (1989). Regularized discriminant analysis. Journal of the American Statistical Association, 84(405): 165-175.
[12] Akaike, H. (2011). Akaike’s Information Criterion. Springer-Verlag, Berlin Heidelberg.
[13] Chiang, L.H., Russell, E.L., Braatz, R.D. (2001) Fault detection and diagnosis in industrial systems. Springer-Verlag, London, U.K.