Research Article

Gearbox Fault Diagnosis Based on a Sparse Principal Component-Generalized Regression Neural Network

Yimo Wang and Yajun Han
Department of Automotive and Transportation Engineering, Jiangsu University of Technology, Changzhou 213001, Jiangsu, China
Correspondence should be addressed to Yimo Wang; wangym0617@163.com

Received 12 July 2022; Revised 25 August 2022; Accepted 5 September 2022; Published 20 September 2022

Academic Editor: Kai Guo

Copyright © 2022 Yimo Wang and Yajun Han. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The fault diagnosis of urban rail transit gearboxes has the characteristics of complex vibration signals and large amounts of data. The daily scheduled maintenance cannot meet the needs of gearbox maintenance, so it is necessary to predict the fault types in advance. In this paper, a method of gearbox fault prediction based on sparse principal component analysis and a generalized regression neural network is presented, and the result of fault prediction can provide a reference for making maintenance plans.

Based on principal component analysis (PCA), a sparse principal component is obtained by adding the LASSO penalty term, which reduces the risk of overfitting of PCA while obtaining a sparse solution. Then, the sparse reduced dimension principal component is input into the generalized regression neural network model for fault diagnosis. The results show that the fault diagnosis method based on the sparse principal component-generalized regression neural network model has high accuracy and is time-consuming.

1. Introduction

The gearbox is a key component of the urban rail transit walking arrangement, and its failure can threaten the normal operation of the train and bring safety problems. Fault diagnosis allows the maintainer to predict the gearbox fault types to be repaired, organize the maintenance strategy in advance, determine the maintenance time, and reduce the maintenance cost and the loss caused by the fault event.

The progress of artificial intelligence technology makes the application of neural network models in device diagnosis research common. The literature [1] applied the short-term memory (LSTM) recurrent neural network to aircraft fault prediction. Reference [2] used a CNN-GRU neural network model to predict short-term transformer faults. Reference [3] predicted the possible faults of power systems using a deep convolutional neural network model. The literature [4] used the convolutional neural network method for fault determination of characteristic turnout characteristic pattern. Reference [5] used the PCA-NARA neural network model for the degradation trend assessment and prediction of wind turbine gearboxes. Reference [6] constructed a GRU recurrent neural network-based fault prediction model for cloud data center applications. When performing data-centric fault diagnosis and prediction, data noise and redundancy can affect the accuracy of fault type determination. Based on this, this paper adopts sparse principal component analysis for sparse processing of fault vibration data on the basis of traditional principal component analysis and combines a generalized regression neural network algorithm to establish a fault diagnosis model, which can be used for gearbox fault diagnosis and provide guidance on the development of maintenance plans for urban rail transit vehicles.

2. Sparse Principal Component Analysis

Principal component analysis (PCA), which can reduce the dimensionality of multicharacteristic indicator data, is a commonly used dimensionality reduction technique [7].
This method simplifies the data structure by constructing a new set of variables and uses fewer eigenvectors to approximate the information of the original data [8]. The calculation process is as follows:

1. After normalizing the original data set with i observations and j process variables with different dimensions, the matrix $X$ is obtained, $X \in \mathbb{R}^{ni \times j}$.

2. The eigenvalue decomposition of $X$ is performed; that is, the covariance solution is performed, and the score matrix $T$ is calculated. That is,

$$
cov (X) = \frac{X^T X}{n-1} = V \Lambda \Lambda^T,
$$

$$
T = XP,
$$

where $\Lambda$ is the diagonal matrix of eigenvalues, $\Lambda \in \mathbb{R}^{j \times j}$. $V$ is the matrix [9], and $V \in \mathbb{R}^{ni \times j}$, $P$ is the load matrix, constructed from the top $n$ columns of $V$, $V \in \mathbb{R}^{j \times j}$.

3. By projecting $T$ back into the $j \times i$-dimensional space, the new matrix $\tilde{X}$ after dimensionality reduction can be obtained, and the cumulative contribution rate $v_i$ can be calculated. Usually, the top $n$ principal components corresponding to indicators with a value greater than 85% are selected [10] which is

$$
\tilde{X} = TP^T,
$$

$$
v_i = \frac{\sum_{m=1}^{i} \Lambda_m}{\sum_{m=1}^{j} \Lambda_m} \times 100%.
$$

PCA has an obvious effect on the dimensionality reduction of fault data, but it also has disadvantages; that is, it is difficult to use for nonlinear faults, and it is difficult to explain the characteristics of the principal component load of nonzero vectors [11] characteristic. Sparse dynamic principal component analysis (SDPCA) can be used to compensate for the shortcomings of the PCA method. Based on PCA, SPCA obtains the sparse principal components of data by adding LASSO penalty items [12]. This paper takes Zou [13]'s sparse method as an example to introduce the solution process of the sparse principal component.

4. Based on the principal component analysis method, the residual vector $E$ is calculated, and the residual vector is the difference between the original matrix $X$ and the new matrix after dimensionality reduction.

$$
E = X - TP^T = X - XPP^T.
$$

The principal component solution is to make the residual vector as small as possible and retain as much information as possible about the principal component in the original data, so the principal component solution can be performed using the following equation, where $Q$ is the solution of the principal component matrix $P$; that is,

$$
Q = \arg \min_{n=1}^{i} \sum_{n=1}^{i} \left\| x_n - P_n^T P_n x_n^2 \right\|.
$$

5. Adding the LASSO penalty term to sparse the obtained principal component reduces the risk of overfitting of the PCA solution while obtaining a sparse solution [13]. That is,

$$
A, B = \arg \min_{m=1}^{i} \sum_{m=1}^{i} \left\| x_n - A \beta_n, B \beta_n \right\| 
$$

$$
+ \lambda \sum_{m=1}^{k} \left\| \beta_m + \lambda_{1,m} \right\|, \quad \lambda > 0
$$

where $A$ and $B$ are the load matrix and sparse load matrix, respectively. $A$, $B$ is the solution of $A$, $B$ when the right side of the equation takes the minimum value. $\beta$ is the load vector after sparsity. $\lambda, \lambda_{1,m}$ is the coefficient of the penalty terms.

6. The solution is the cumulative contribution rate $v_s$ of the SPCA model.

$$
v_s = \frac{\sum_{m=1}^{k} \Lambda_m}{\sum_{m=1}^{j} \Lambda_m} \times 100%.
$$

$SA$ is the diagonal array of explained variance after sparsity. $P_s$ is the load vector matrix after sparsity.

3. GRNN Fault Prediction Model

A generalized regression neural network (GRNN) consists of an input layer, a pattern layer, a summation layer, and an output layer [14]. GRNN has a good fitting effect on nonlinear data when the data accuracy is poor or the amount of data is small. The structure of the GRNN network is shown in Figure 1.

The vector $X = [x_1, x_2, \ldots, x_n]^T$ is input from the input layer, and the number of original characteristic parameters is $n$. The pattern layer accepts the data output from the input layer, which is transformed into $P = [p_1, p_2, \ldots, p_n]^T$ by the neuron transfer function and passed to the summation layer. $\sigma$ in the GRNN model represents the smoothing factor, and the ideal $\sigma$ value can usually be calculated according to the distribution probability of the sample. That is,

$$
P_1 = \exp \left[ \frac{(X-X_i)^T (X-X_i)}{2\sigma^2} \right].
$$

There are two types of neurons in the summation layer, the denominator unit and numerator unit [15], and the $S_D$ and $S_N (j = 1, 2, \ldots, k$, where $k$ is the dimension of the output vector) are obtained after the output of the hidden layer is multiplied by the corresponding weight and then added and summed, and the output vector is $Y = [y_1, y_2, \ldots, y_n]^T$; that is,
Suppose $x$ and $y$ are two random variables. The joint probability density is $f(x, y)$, in which $x_0$ is the observed value of $x$, and the regression relationship between $y$ and $x$ is

$$E(y|x_0) = \int_{-\infty}^{\infty} y f(x_0, y) dy $$

(9)

Applying the Parzen nonparametric estimation, the probability density function $f(x_0, y)$ can be estimated from the sample data set. That is,

$$f(x_0, y) = \frac{1}{n(2\pi)^{p/2}\sigma_p} \sum_{i=1}^{n} \exp \left[ -\frac{(X-X_i)^T(X-X_i)}{2\sigma^2} \right]$$

(10)

In Equation (10), $n$ is the sample size, and $p$ is the dimensionality of the random variable $x$. Substituting this into Equation (9), exchange the order of integration and summation and simplify it to obtain:

$$y(x_0) = \frac{\sum_{i=1}^{n} \exp \left[ -\frac{(X-X_i)^T(X-X_i)}{2\sigma^2} \right]}{\sum_{i=1}^{n} \exp \left[ -\frac{(X-X_i)^T(X-X_i)}{2\sigma^2} \right]}$$

(11)

4. Gearbox Fault Prediction and Analysis

4.1. Data Preprocessing. The gearbox contains mechanical parts, such as gears, shafts, rolling bearings, and boxes, which are common power transmission devices in urban rail vehicles. Its working principle is that the rotating shaft drives the sector gears in the gearbox and transmits torque vertically to the other shaft. According to the data [16], gearbox faults occur more often in the gears and rolling bearings, 59% and 20%, respectively. This paper takes the transmission system diagnostic simulator (DDS) as the data acquisition object, the speed is 2000 r/min, and the sampling frequency is 5120 Hz. Four typical faults of the gearbox were selected for fault prediction, including bearing ball error, bearing inner and outer ring fault, gear missing teeth, and gear surface cracks. The gearbox parameters are shown in Table 1.

The collected vibration signals were taken as a sample set every 1024 points, and 30 sample sets were extracted under five working conditions: normal, bearing ball misalignment, bearing inner and outer ring fault, gear missing teeth, and gear surface cracks. A total of 29 indicators in the time domain and frequency domain were extracted from each sample set, which can comprehensively describe the faults of gears and bearings in the gearbox. However, characteristics in the frequency domain were obtained through a Fourier transform on the basis of the time domain, so there is information redundancy. In addition, some indicators are sensitive to faults, and some indicators are not sensitive to faults, so if not processed, it will cause inaccurate prediction results. This paper uses PCA and SPCA to reduce the dimensionality of 29 characteristic indicators, eliminate the indicators with low fault sensitivity, eliminate data redundancy, and improve the fault prediction accuracy.

The 150 gearbox vibration data sets under five working conditions are used as input for PCA and SPCA analysis, and the dimensionality reduction results obtained are shown in Table 2.

In Table 2, PC1~PC6 are the first six principal components after dimensionality reduction. It can be seen from the dimensionality reduction results that although the cumulative contribution rate of SPCA is 85.2% lower than that of PCA, 88.3%, there are zero elements in the principal component load extracted by the SPCA dimensionality reduction method, and the number of the corresponding zero elements are 2, 4, 8, 7, 8, and 10, which reduces the complexity of variable dimensionality reduction. The first six
The partial load matrix of the first six principal components of SPCA is shown in Table 3.

4.2. Comparison of Fault Prediction Results. The 150 groups of fault characteristic sets composed of five working conditions and 29 characteristic indicator values were transformed into a fault characteristic set composed of 6 characteristic indicators through PCA and SPCA, and 24 groups and 6 sets of fault characteristics of each working condition were randomly selected. The set consists of 120 training sets and 30 testing sets. To compare the fault prediction effects of the GRNN, PCA-GRNN, and SPCA-GRNN models, 29 time-domain and frequency-domain fault characteristic sets were used as the input of the GRNN model, and the fault characteristic sets transformed by the PCA and SPCA analysis were used as input to the PCA-GRNN and SPCA-GRNN models. The value of the smoothing factor in the GRNN model determines the prediction result. In this paper, the smoothing factor corresponding to the highest accuracy rate is determined by inputting the training samples and the changed smoothing factor value. The interval of the smoothing factor value is [0.1, 1]. When this value is 0.7, the maximum accuracy is 96%. The main parameters of the models were determined based on the input indicator characteristics of the three models and the type of faults predicted, as shown in Table 4.

By inputting the training set for model training and inputting the test set to test the performance of the model, the fault prediction results of the gearbox are obtained, as shown in Figures 2–4, and the comparison of the fault prediction results of the three models is shown in Table 5.

---

**Table 1: Gearbox parameters.**

| Component | Parameter            | Value  |
|-----------|----------------------|--------|
| Gear      | Teeth number of driven gear | 88     |
|           | Teeth number of the master gear | 44     |
|           | Spiral angle         | 9°22′  |
|           | Pressure angle       | 20°    |
|           | Modulus              | 1.5    |
|           | Engage frequency     | 959.39 Hz |
|           | Rolling body diameter | 8 mm   |
|           | Number of rolling elements | 17     |
|           | Contact angle        | 15°    |
| Rolling bearings | Pitch radius of high-speed shaft | 62 mm |
|           | Pitch radius of low-speed shaft | 72 mm |
|           | Rotational frequency of high-speed shaft | 10.902 Hz |
|           | Rotational frequency of low-speed shaft frequency | 21.804 Hz |

**Table 2: The dimensionality reduction results of PCA and SPCA.**

| Methods   | Variables          | PC1 (%) | PC2 (%) | PC3 (%) | PC4 (%) | PC5 (%) | PC6 (%) |
|-----------|--------------------|---------|---------|---------|---------|---------|---------|
| SPCA      | Contribution rate  | 57.6    | 16.3    | 4.7     | 3.3     | 2.1     | 1.2     |
|           | Cumulative contribution rate | 57.6  | 73.9    | 78.6    | 81.9    | 84      | 85.2    |
| PCA       | Contribution rate  | 54.5    | 22      | 5.5     | 3.1     | 1.8     | 1.4     |
|           | Cumulative contribution rate | 54.5  | 76.5    | 82      | 85.1    | 86.9    | 88.3    |

**Table 3: SPCA principal component partial load matrix.**

| Variables                  | PC1        | PC2        | PC3        | PC4        | PC5        | PC6        |
|----------------------------|------------|------------|------------|------------|------------|------------|
| Mean value                 | -0.199     | 0          | 0          | 0          | 0          | 0          |
| Peak value                 | -0.183     | 0.148      | 0.024      | 0          | 0          | 0          |
| Peak value                 | -0.136     | 0.041      | 0          | 0          | 0          | 0          |
| Square root amplitude      | -0.224     | 0          | 0          | 0          | 0          | 0          |
| Absolute mean value        | -0.113     | 0.119      | 0.065      | 0.140      | 0.557      | 0.005      |
| Frequency root mean square value | -0.173 | 0.124      | -0.084     | 0.118      | 0.009      | 0          |
| Frequency standard deviation | 0.159     | 0.366      | 0          | -0.120     | -0.003     | -0.106     |
| Frequency dispersion       | -0.192     | -0.075     | -0.020     | -0.065     | -0.243     | 0.268      |
| Frequency skewness         | 0.187      | 0.003      | 0          | 0          | 0          | -0.127     |
| Normalized spectral mean   | 0.163      | 0.106      | 0.231      | 0.0726     | -0.038     | 0          |
| Main frequency band position | 0         | 0          | 0.038      | 0.060      | 0.090      | 0          |
When the original data before dimension reduction are used as input for fault prediction, the accuracy rate is 43.3%, and the prediction time is 10.23 s. The fault prediction accuracy of the generalized regression neural network based on principal component analysis can reach 90%, and the prediction time is 2.63 s. Although the SPCA-GRNN model, which extracted six sparse principal elements for generalized regression neural network prediction, had the lowest cumulative contribution of the three models, the prediction accuracy reached 96.6%, and the prediction time was 82% shorter than the GRNN model. Compared with the PCA-GRNN model, it was 30% shorter and the model predicted the best results.

5. Conclusion

(1) SPCA reduces the dimension of sample data on the basis of PCA by adding the LASSO penalty item, and the six sparse principal components obtained contain more than 85% of the information of the original data, which improves the interpretability of the principal component while reducing the computational complexity.

(2) Aiming at the gearbox fault prediction problem, this paper adopts the prediction method combining sparse principal component analysis and a generalized regression neural network. After comparing the fault prediction results of the three models of GRNN, PCA-GRNN, and SPCA-GRNN, it is concluded that the prediction ability is SPCA-GRNN > PCA-GRNN > GRNN.
(3) In this paper, sparse principal component analysis is introduced into the research of fault vibration data processing, and combined with GRNN, it is used for gearbox fault prediction. Future research will consider the feasibility of integrating kernel principal component analysis with other machine learning and fault prediction effects.

Data Availability

The labeled dataset used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The author declares no conflicts of interest.

Acknowledgments

This research was supported by the Jiangsu University Basic Science (Natural Science) Research Project—Research on Coordination Optimization of Driverless Vehicle Tracking System in Narrow Area (No. 22KJB120002).

References

[1] X. Wang, J. Wu, C. Liu, H. Yang, Y. Du, and W. Niu, “Fault time series prediction based on LSTM recurrent neural network,” Journal of Beihang University, vol. 44, no. 4, pp. 772–784, 2018.

[2] W. Yang, C. X. Pu, K. Yang, A. A. Zhang, and G. L. Qu, “Transformer short-term fault prediction method based on CNN-GRU combined neural network,” Power System Protection and Control, vol. 50, no. 6, pp. 107–116, 2022.

[3] H. Ren, Z. J. Hou, B. Vyakaranam, H. Wang, and P. Etingov, “Power system event classification and localization using a convolutional neural network,” Frontiers in Energy Research, vol. 8, no. 1, p. 607826, 2020.

[4] M. Sun, Research on Fault Diagnosis Method of Switch Machine Based on Neural Network, University of Electronic Science and Technology of China, Chengdu, China, 2020.

[5] Y. Xing, Evaluation and Prediction of Wind Turbine Gearbox Deterioration Trend Based on PCA-NAR Neural Network, Southwest Jiaotong University, Chengdu, China, 2019.

[6] X. Hu, “Application fault prediction method of cloud data center based on GRU recurrent neural network,” Railway Computer Application, vol. 31, no. 2, pp. 7–11, 2022.

[7] C. Hui and Y. Zhou, “Research on equipment failure prediction based on PCA-BP neural network,” Logistics Engineering and Management, vol. 43, no. 6, pp. 150–153, 2021.

[8] Y. Wang and R. Liu, “Vehicle pantograph fault diagnosis based on principal component analysis-probabilistic neural network,” Urban Rail Transit Research, vol. 24, no. 1, pp. 88–92, 2021.

[9] Y. Duan, P. Wu, and J. Gao, “Fault detection method based on sparse dynamic principal component analysis,” Computer Measurement & Control, vol. 27, no. 4, pp. 46–50, 2019.

[10] K. Zhang, K. Zhang, and L. Kun, “Principal component analysis-neural network rockburst grade prediction model,” Chinese Journal of Safety Science, vol. 31, no. 3, pp. 96–104, 2021.

[11] A. K. Seghouane, N. Shokouhi, and I. Koch, “Sparse principal component analysis with preserved sparsity pattern,” IEEE Transactions on Image Processing, vol. 28, no. 7, pp. 3274–3285, 2019.

[12] Y. Duan, Research on Fault Detection Based on Sparse Principal Component Analysis [D], Zhejiang Sci-Tech University, Hangzhou, China, 2019.

[13] T. Hastie, H. Zou, and R. Tibshirani, “Sparse principal component analysis,” Journal of Computational & Graphical Statistics, vol. 15, no. 2, pp. 265–286, 2006.

[14] L. Gu, Research and Application of Greenhouse Tomato Fertilization Regulation Based on Generalized Regression Neural Network, Heilongjiang University, Harbin, China, 2021.

[15] G. Cheng, Rolling Force Prediction Based on Generalized Regression Neural Network, Metallurgical Automation Research and Design Institute, Beijing, China, 2021.

[16] B. Liang, Research on State Modelling and Remaining Life Prediction of Wind Turbine Gearbox, North China Electric Power University, Beijing, China, 2019.