ISBN 978-981-11-3671-9
Proceedings of 2017 the 7th International Workshop on Computer Science and Engineering (WCSE 2017)
Beijing, *25-27 June*, *2017*, *pp. 1124-1128*
doi: 10.18178/wcse.2017.06.196

Application of Principal Component Analysis to Aircraft Integrated Condition Monitoring and Assessment

Yanming Yang [1] +, Yingjian Liu [2] and Youbao Ding [1]

[1] Qingdao Campus, Naval Aeronautical and Astronautical University
[2] Department of Computer Science and Technology, Ocean University of China

Abstract. Principal component analysis (PCA) is a multivariate statistics technique that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. The aircraft integrated condition index is the important parameter to measure the aviation maintenance support capability. It is indispensable to scientifically analyze aircraft condition index data and to make scientific decisions on aviation maintenance to improve maintenance support capability. Aircraft integrated condition involves a large number of indexes, which is more difficult to monitor and evaluate. This paper presents a method of aircraft integrated condition monitoring and evaluation based on principal component analysis model, then gives its mathematic model and algorithm steps in detail. This method integrates multiple indexes by principal component analysis, uses the cumulative variance contribution rate to identify the principal component variables, and turns many index questions into less overall targets. To extract the most important information from the original data eliminates the information redundancy between the samples, and reduces the index dimension, so that the aircraft integrated condition monitoring and evaluation problems are simplified. The practical applications and results analysis show that the PCA-based method is feasible and effective for aircraft integrated condition monitoring and assessment.

Keywords: principal component analysis, aircraft integrated condition, monitoring and assessment, eigenvector and eigenvalue

1. Introduction

With the rapid development of aviation equipment, there are a lot of monitoring and assessment problems in the aviation maintenance support. To carry out effective monitoring and assessment of aircraft integrated condition, for the promotion of aviation maintenance work targeted, predictive, and promote scientific maintenance of aviation equipment plays a very important role. Aircraft integrated condition index data is becoming more and more complex. Therefore, there is an important significance to develop an effective approach to aircraft integrated condition monitoring and assessment. Principal component analysis (PCA) is a linear dimensionality reduction technique [1, 2]. PCA is a technique used to emphasize variation and bring out strong patterns in a dataset. It's often used to make data easy to explore and visualize [3]. PCA method can generalize lower dimensional representation of the original data, in terms of capturing the data direction that has the largest variance [4]. PCA is commonly used as one step in a series of analyses. For a comprehensive introduction to PCA, the reader is referred to [5]. PCA has been widely used to monitor the aircraft integrated condition with multiple variables and evaluate aviation maintenance support ability.

2. The Mathematical Model of Principal Component Analysis

Suppose the study object is n samples, p variable data ($n>p$). We can represent the original data as the following matrix [6]:

+ Corresponding author. Tel.: + 0532-51833335.
E-mail address: yymqd@126.com.

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} = (X_1, X_2, \cdots, X_p) \tag{1}$$

Linear transformation of $\boldsymbol{X}$ can form a new integrated variable $\boldsymbol{Y}$:

$$Y = XU \tag{2}$$

where $Y = (Y_1, Y_2, \cdots, Y_p)$, $U = (u_1, u_2, \cdots, u_p)$.

The k-th integrated variable is:

$$Y_k = Xu_k \tag{3}$$

where u_k is the coefficient of linear transformation.

According to the requirements of the PCA, the following constraints are applied to the linear transformation: (1) Y_i is not related to Y_j ($i \neq j;\ i, j = 1, 2, \cdots, p$); (2) Y_1 is the largest variance in all linear combinations of $X_1, X_2, \ldots, X_p$, Y_2 is the largest variance of all linear combinations of $X_1, X_2, \ldots, X_p$ that are not related to $Y_1, Y_2, \ldots, Y_p$ …, Y_p is the largest variance of all linear combinations of $X_1, X_2, \ldots, X_p$ that are not related to $Y_1, Y_2, \ldots, Y_p$; (3) $\sum_{i=1}^{p} u_{ik}^2 = 1$, that is $u_k' u_k = 1$.

The resulting integrated variable $Y_1, Y_2, \ldots, Y_p$ are the first, second, and p-th principal component of the original variable, respectively.

3. The Algorithm Steps of Principal Component Analysis

3.1. Calculate the Correlation Matrix

To calculate the correlation matrix of the standardized matrix $\boldsymbol{R}$:

$$R = \left(r_{ij}\right)_{p \times p} = \frac{X'X}{n-1} \tag{4}$$

where $r_{ij} = \frac{1}{n-1} \sum_{k=1}^{n} x_{ki} x_{kj},\ i, j = 1, 2, \cdots, p$.

3.2. Calculate the Eigenvectors and Eigenvalues to Determine the Principal Components

Using Jacobi method to solve the characteristic equation of the correlation coefficient matrix $\boldsymbol{R}$, we obtain the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_p$ ($\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p$), and the corresponding eigenvector $e_i = (e_{i1}, e_{i2}, \cdots e_{ip})$, $i = 1, 2, \cdots, p$.

Then its eigenvalues λ_i and corresponding eigenvector e_i of coefficient matrix $\boldsymbol{R}$, can be also computed respectively as follows:

$$|R - \lambda I| = 0 \tag{5}$$

$$[\lambda_i I - R] e_i = 0 \tag{6}$$

where $\boldsymbol{R}$ is the Correlation coefficient matrix, $\boldsymbol{I}$ is the p-order unit matrix, $e_i = [e_i(1),\ e_i(2),\ \cdots, e_i(n)]'$.

3.3. Calculate the Variance Contribution Rate

The variance contribution rate of the k-th principal component Y_k :

$$\alpha_k = \lambda_k \Big/ \sum_{i=1}^{p} \lambda_i \quad , i = 1, 2, \cdots, p \tag{7}$$

The cumulative variance contribution of the former m principal components $Y_1, Y_2, \cdots, Y_m$:

$$\sum_{i=1}^{m} \alpha_i = \sum_{i=1}^{m} \lambda_i \Big/ \sum_{i=1}^{p} \lambda_i \quad , m < p \tag{8}$$

In general, the number of principal components k value is determined by the variance contribution rate $\sum_{i=1}^{m} \alpha_i \geq 85\%$.

3.4. Comprehensive Evaluation of Principal Component

First, calculate the linear weighting of each principal component:

$$\boldsymbol{F}_i = \sum_{k=1}^{p} \boldsymbol{F}_{ik} \quad , i = 1, 2, \cdots, m \tag{9}$$

where $\boldsymbol{F}_{ik} = \sum_{j=1}^{n} \boldsymbol{e}_{kj} \boldsymbol{X}'_{ji} \quad , i = 1, 2, \cdots, m; k = 1, 2, \cdots, p$.

And then k principal components of the weighted sum, that is, the final evaluation value. The weight is the variance contribution rate of each principal component.

$$\boldsymbol{F} = \sum_{i=1}^{m} \alpha_i \boldsymbol{F}_i \quad , i = 1, 2, \cdots, m \tag{10}$$

where α_i is the variance contribution rate of the i-th principal component.

4. An Application Example

4.1. Problem Description

The seven aircraft integrated condition index data of 20 aircrafts in five years: flight sorties, flight time, intact rate, all state intact rate, maintenance grounding rate, failure rate, air failure rate is shown in Table 1. Try to evaluate integrated condition for each aircraft.

Table 1: Aircraft Integrated Condition Index Data in Five Years

Aircraft ID	Flight Sorties (X1)	Flight Time/h (X2)	Intact Rate/% (X3)	All State Intact Rate/% (X4)	Grounding Rate/% (X5)	Failure Rate/% (X6)	Air Failure Rate/% (X7)
1	651	868.72	83.16	78.56	11.81	20.14	2.19
2	609	857.98	89.68	85.45	7.16	22.96	3.03
3	618	845.86	89.14	82.12	5.91	19.15	2.60
4	556	747.63	90.16	85.66	6.61	21.40	3.08
5	634	848.45	92.09	87.33	4.43	20.98	2.83
6	537	719.60	86.01	79.68	7.78	23.76	3.20
7	562	767.53	83.60	78.24	11.33	22.15	3.26
8	583	805.36	88.25	80.11	6.07	20.36	2.98
9	616	794.78	78.18	70.05	12.66	21.89	3.27
10	613	853.89	86.69	81.23	6.79	13.94	0.82
11	489	662.80	85.13	80.41	10.76	17.95	3.47
12	660	839.35	91.86	86.35	4.52	15.73	2.50
13	632	876.00	84.65	78.09	11.13	17.35	2.28
14	601	821.15	86.72	81.27	9.76	18.27	1.70
15	642	853.71	91.56	87.59	4.16	19.56	3.16
16	630	812.75	85.95	78.99	6.47	19.25	2.42
17	627	851.33	86.42	82.31	8.62	17.74	2.58
18	538	784.10	89.33	85.01	4.90	19.39	3.32
19	577	767.46	89.56	80.98	5.95	17.07	3.78
20	502	724.78	92.36	86.37	3.51	19.45	2.48

4.2. Analysis Process

Calculate the Correlation Coefficient Matrix. According to the correlation between the two index data, calculate the correlation coefficient of the index data. For the same index data, the correlation coefficient is 1, with 1 as the boundary, the diagonal elements take the same value. According to the Eq. (4), we can get the correlation coefficient matrix $\boldsymbol{R}$.

$$\boldsymbol{R}=\begin{bmatrix} 1.000 & -0.913 & -0.204 & 0.057 & 0.064 & 0.147 \\ -0.913 & 1.000 & 0.184 & -0.068 & -0.092 & -0.152 \\ -0.204 & 0.184 & 1.000 & 0.517 & -0.283 & -0.280 \\ 0.057 & -0.068 & 0.517 & 1.000 & -0.294 & -0.360 \\ 0.064 & -0.092 & -0.283 & -0.294 & 1.000 & 0.920 \\ 0.147 & -0.152 & -0.280 & -0.360 & 0.920 & 1.000 \end{bmatrix}$$

Calculate the Eigenvectors and Eigenvalues. That is, the eigenvalues and eigenvectors of the correlation matrix $\boldsymbol{R}$ are calculated.

According to the Eq. (5) to calculate the eigenvalue:

$$\lambda=\left(\ 2.8288,\ 2.4312,\ 1.0293,\ 0.4230,\ 0.1951,\ 0.0684,\ 0.0242\right)'$$

According to the Eq. (6) to calculate the corresponding eigenvectors as shown in Table 2.

Table 2: Eigenvectors of the Correlation Matrix

Variables	PC1	PC2	PC3	PC4	PC5	PC6	PC7
X1	0.071	0.551	0.428	0.262	0.086	0.658	0.000
X2	0.134	0.560	0.379	-0.002	-0.114	-0.715	0.001
X3	0.574	-0.139	0.046	-0.005	-0.091	0.038	0.799
X4	0.554	-0.117	0.066	-0.175	-0.611	0.145	-0.500
X5	-0.546	0.130	-0.040	-0.063	-0.750	0.099	0.327
X6	-0.185	-0.323	0.693	-0.601	0.120	0.063	0.044
X7	-0.088	-0.478	0.430	0.732	-0.144	-0.143	-0.050

Determine the Principal Components. In order to determine the number of principal components, the cumulative variance contribution rate should be calculated first.

According to the Eq. (7) and Eq. (8) to calculate the variance contribution rate and cumulative variance contribution as shown in Table 3.

Table 3: Total Variance Explained

Component	Initial Eigenvalue			Extraction Sums of Squared Loadings		
	Total	Variance/%	Cumulative/%	Total	Variance/%	Cumulative/%
1	2.829	40.411	40.411	2.829	40.411	40.411
2	2.431	34.731	75.142	2.431	34.731	75.142
3	1.029	14.705	89.847	1.029	14.705	89.847
4	0.423	6.043	95.890			
5	0.195	2.787	98.677			
6	0.068	0.978	99.655			
7	0.024	0.345	100.000			

It can be seen from Table 3 that the first three eigenvalues are greater than 1 and the cumulative contribution rate is over 85% (89.847%), which indicates that the information reflected in the six aircraft maintenance index data can be reflected by three principal components, which explain 89.847% of total variation of variables in PCA.

After conducting PCA, we found that the first three PCs explained 89.847% of the process variation. Thus, we only used the first three PCs to obtain the variables. Therefore, the expression of the principal component can be calculated according to Eq. (9) and Table 2:

$$PC1=0.071\boldsymbol{X}_1+0.134\boldsymbol{X}_2+0.574\boldsymbol{X}_3+0.554\boldsymbol{X}_4-0.546\boldsymbol{X}_5-0.185\boldsymbol{X}_6-0.088X_7$$

$$PC2=0.551\boldsymbol{X}_1+0.560\boldsymbol{X}_2-0.139\boldsymbol{X}_3-0.117\boldsymbol{X}_4+0.130\boldsymbol{X}_5-0.323\boldsymbol{X}_6-0.478X_7$$

$$PC3=0.428\boldsymbol{X}_1+0.379\boldsymbol{X}_2+0.046\boldsymbol{X}_3+0.066\boldsymbol{X}_4-0.040\boldsymbol{X}_5+0.693\boldsymbol{X}_6+0.430X_7$$

4.3. Result Analysis

From the analysis of Fig.1 that: The first principal component includes intact rate (X3), all state intact rate (X4) and grounding rate (X5); the second principal component includes flight sorties (X1) and flight time (X2); the third principal component includes failure rate (X6) and air failure rate (X7).

The analysis shows that the first principal component is expressed as the maintenance support capability, the second principal component is expressed as the strength of the aircraft, and the third principal component is expressed as the inherent reliability of the aircraft.

In order to visually display the assessment result, the distribution state of the principal component score may be represented by a score plot. The main component of the eigenvalue is the coordinate axis, which shows the main component score state, as shown in Fig.2. It can be seen from the figure, for the first principal component, the more the point on the right side, indicating the greater the aircraft maintenance support capacity; for the second main component, the upper point, indicating the greater use of aircraft strength.

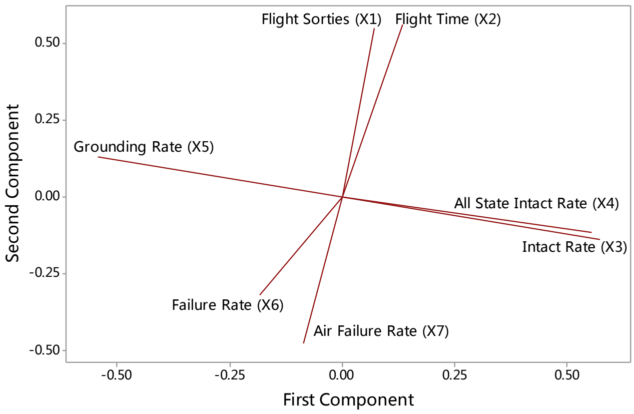

Fig. 1: Loading plot of aircraft integrated condition.

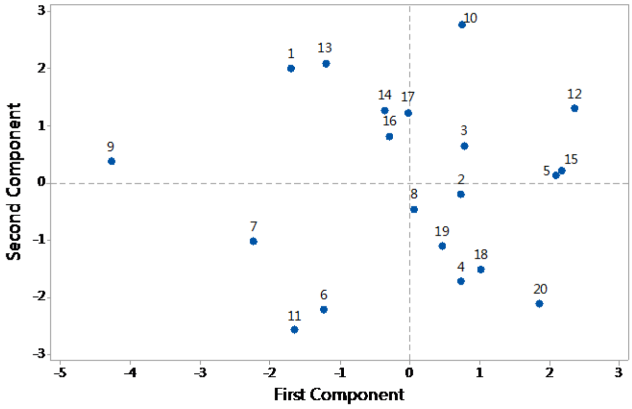

Fig. 2: Score plot aircraft integrated condition.

5. Conclusion

Aiming at the aircraft integrated condition monitoring and assessment, an effective approach based on PCA is put forward. Based on the principle of principal component analysis, the mathematical model and algorithm steps are given in detail. In order to improve the operability of the method, taking the aircraft integrated condition index data as an example, this paper discusses how to use the PCA-based method to monitor and evaluate the aircraft integrated condition. And the visual analysis results are presented graphically. The results show that the method proposed in this paper is an effective monitoring and assessment approach to aircraft integrated condition. The method can also be applied to other multivariate evaluation issues in the field of aeronautical equipment maintenance support.

6. References

[1] Jolliffe, I. T., and J. Cadima. "Principal component analysis: a review and recent developments. " Philosophical Transactions of the Royal Society A Mathematical Physical & Engineering Sciences 374.2065(2016):20150202.

[2] Shen, Dan, H. Shen, and J. S. Marron. "A general framework for consistency of principal component analysis." Journal of Machine Learning Research 17.1(2016):5218-5251.

[3] Behdad, Mohammad, et al. "On principal component analysis for high-dimensional XCSR." Evolutionary Intelligence 5.2(2012):129-138.

[4] Wang, Rixin, et al. "Fault detection of flywheel system based on clustering and principal component analysis." Chinese Journal of Aeronautics 28.6(2015):1676-1688.

[5] Abdi. H., & Williams, L.J. (2010). "Principal component analysis". Wiley Interdisciplinary Reviews: Computational Statistics. 2 (4): 433-459.

[6] Zhengshun, and Kangling. "Online process monitoring for complex systems with dynamic weighted principal component analysis." Chinese Journal of Chemical Engineering 24.6(2016):775-786.