Based on the MSET and SPRT Liquid Nitrogen Pump Prognostics and Health Management

Qing-zhong HU\textsuperscript{1,2} and Fu-lei CHU\textsuperscript{1}

\textsuperscript{1}Department of Mechanical Engineering, Tsinghua University, Beijing 100084, P. R. China
\textsuperscript{2}Key Laboratory for the Space Launching Site Reliability Technology, Xichang, 615000, P. R. China

Keywords: Prognostics and health management, Multivariate state estimation techniques, Sequential probability ratio test.

Abstract. To improve the reliability and availability, prognostics and health management technology was applied to the liquid nitrogen pump. This paper presents a multivariate state estimation techniques and sequential probability ratio test model to predict equipment health. In the approach, correlation model among monitoring parameters in normal work condition is constructed firstly. Then, according to the similarities between the current observed feature vector and each history feature vector contained in process memory matrix, estimation of the current feature vector is calculated by using MSET, and residual signal between the current feature vector and its estimation is obtained in turn. Finally, mean test and variance test for the residual signal is executed by using SPRT, and work condition of the system is pronounced. Use the real-time data to test the model, the result showed that using this method can obtain good effect.

Introduction

As a power and testing gas resource, high pressure gas were widely applied to aerospace, aviation, shipbuilding, chemical industry and so on. Nothing can instead of them. The high pressure nitrogen system is used in new space launch site. Failure of the high pressure nitrogen system will lead to serious problem, which may cause system breakdown and huge loss. The prognostics and health monitoring (PHM) can provide the system condition information, report failure before it happens and prevent whole system from breakdown, So PHM methods of the high pressure nitrogen system are very important. Pressure and current prediction is a good solution to this problem since it can provide Cavitation erosion condition directly, this paper will present some methods related.

There are many research works on feature extraction and recognition of vibration signal, in this paper we try to use the statistical analysis prediction the cavitation erosion. Estimation of the current feature vector is calculated by using multivariate state estimation techniques (MSET), mean test and variance test for the residual signal is executed by using sequential probability ratio (SPRT).

This paper introduces a new approach related to cavitation erosion prediction based on MSET and SPRT model, and is arranged as follows: section 1 introduces MSET will be applied, section 2 introduces SPRT will be applied, section 3 describes pressure and current prediction data used in this paper, and section 4 gives conclusion.

Multivariate State Estimation Techniques

Multivariate state estimation techniques is an advanced pattern recognition technology, MSET calculation of similarity between the monitoring parameters, then the similarity judgment to estimate the state of an object, it needs the data is only part of the normal operation of the equipment. Currently MSET technology has been conducted in some equipment and equipment system in the field of application won a great success [1].

MSET used in the process of training data is equipment health history data, which is through the sensor to take to the equipment of trouble-free data, MSET to make use of the normal data after the training, it made clear the relation between each parameter to monitor each other and system health,
the relationship between the state through the calculation of weighted array will it hold, MSET weights and according to the study of the state of the model similarity or overlapping degree to early. Then you can use this MSET model of the system state identification and evaluation, the new acquisition system observation sequence of observations or to be detected data sequence input MSET model, MSET is determined by the model of the model to assess the current of the true state of these sequences. Though because of the evaluation data is not linear independent state vector, but with the system or the ongoing related physical process to a certain extent, so the health status of system described by the data obtained from monitoring of these directly or indirectly out [2]. Due to the multivariate state estimation technique such as neural networks model is built and the learning process and compared more simply and quickly, so can better meet reproduce the real-time requirements of actual use.

Multivariate state estimation process as shown in Figure 1.

![Figure 1. MSET process.](image)

The process memory matrix D is the basis of multivariate state estimation system assessment, it is also referred to as the plural form estimates of system model. Multivariate state estimation in the training process to choose enough history of training samples, the selected sample is the normal operation of the system state data, covers the system under normal working condition the record of all individual monitoring-station dynamic process.

Let \( x_i(t_j) \) is the i-th data parameter at a point in time \( t_j \) collection, all the parameters \( t_j \) point in time are combined eigenvectors of the vector is

\[
X(t) = [x_1(t), x_2(t), \ldots, x_n(t)]^T
\]  

(1)

The N characteristic parameters of each observation data storage matrix D column as a process, when there are m samples of historical data, the matrix D is

\[
D = \begin{bmatrix}
d_{1,1} & \cdots & d_{1,m} \\
\vdots & \ddots & \vdots \\
d_{n,1} & \cdots & d_{n,n}
\end{bmatrix} = \begin{bmatrix}
x_1(t_1) & \cdots & x_1(t_m) \\
\vdots & \ddots & \vdots \\
x_n(t_1) & \cdots & x_n(t_m)
\end{bmatrix}
\]  

(2)

Put \( X_{obs} \) observation signal matrix, \( X_{est} \) matrix for the estimation of the corresponding signal, if \( X_{est} \) estimate matrix can be process storage matrix D columns linearization, there are

\[
X_{est} = DW
\]  

(3)

\( W \) in the formula (3) is a weight vector, a weight that their similarity measures, weight is in use state estimation matrix \( X_{est} \) matrix D and have clear out the process of storage, the similarity measure between the weight vector \( W \) is needed to minimize error vector \( E \) to determine, vector \( E \) for

\[
E = X_{obs} - X_{est}
\]  

(4)

In \( E \) to minimize the constraint condition, under the condition of the least square error estimation solution, for the weight vector \( W \)

\[
W = (D^T D)^{-1} D^T X_{obs}
\]  

(5)

\( D^T D \) must be reversible and type (5) can be used, to compute the weight vector \( W \).

The number of rows stored in the number of columns of matrix D is less than the memory matrix D is a necessary but not sufficient condition for DTD reversible. This case tells us that the memory
matrix $D$ process for monitoring system status monitoring parameter monitoring system should be greater than the number $n$ of the number of states $m$ (historical observations samples). However, in actual use, you must use the large number of samples to be able to signal to the system to provide statistically valid information, and ultimately to obtain the number of cases being monitored parameter is less than the number of sample data. Let DTD reversible problem is resolved, pluralistic state estimation method for nonlinear operator $\otimes$ substituted linear vector multiplication operator, so the weight vector $W$

$$W = (D^T \otimes D)^{-1}(D^T \otimes X_{obs})$$

(6)

The corresponding signal estimation matrix $X_{est}$

$$X_{est} = D(D^T \otimes D)^{-1}(D^T \otimes X_{obs})$$

(7)

The remaining training data estimation matrix $L$, denoted $List$

$$L_{est} = D(D^T \otimes D)^{-1}(D^T \otimes L)$$

(8)

Accordingly, the definition of a difference between the residual value and the actual value of the above estimates, so the actual residuals and health residuals:

$$R_X = X_{est} - X_{obs}$$

(9)

$$R_L = L_{est} - L_{obs}$$

(10)

**Sequential Probability Ratio Test**

Abraham Wald is put forward in the 1940s the Sequential Probability Ratio Test (SPRT), make the optimal choice to solve the problem of sample size [3]. It is highly suitable for the applications where testing is destructive or very expansive.

When the sample $X = [x_1, x_2, \cdots x_n]$ is normal distribution, average hypothesis test are as follows

$$\begin{align*}
H_0 & : \text{Mean } \mu_0, \text{Variance } \sigma_0^2;
H_1 & : \text{Mean } \mu_1 > \mu_0, \text{Variance } \sigma_0^2.
\end{align*}$$

Hypotheses $\mu_0$ and $\sigma_0^2$ is the normal state signals mean and variance, $\mu_1$ to mean an abnormal state signal.

In the above scenario, the likelihood function for the sample:

$$G(x) = g_1(x)g_2(x)\cdots g_n(x) = \frac{1}{(2\pi\sigma_0^2)^{n/2}} \exp\left[- \frac{1}{2\sigma_0^2} \sum_{i=1}^{n} (x_i - \mu_0)^2 \right]$$

(11)

$$F_1(x) = f_1(x)f_2(x)\cdots f_n(x) = \frac{1}{(2\pi\sigma_1^2)^{n/2}} \exp\left[- \frac{1}{2\sigma_1^2} \sum_{i=1}^{n} (x_i - \mu_1)^2 \right]$$

(12)

Likelihood ratio is calculated by the formula (11) and (12) can be obtained:

$$\Lambda(X) = \frac{F_1(x)}{G(x)} = \exp\left[\frac{(\mu_1 - \mu_0)}{\sigma_0^2} \sum_{i=1}^{n}(x_i) + \frac{n}{2\sigma_0^2} (\mu_1^2 - \mu_0^2)\right]$$

(13)

The formula (13) taking the logarithm on both sides, are:

$$\ln \Lambda(X) = \frac{(\mu_1 - \mu_0)}{\sigma_0^2} \sum_{i=1}^{n}(x_i) + \frac{n}{2\sigma_0^2} (\mu_1^2 - \mu_0^2)$$

(14)
Formula (14) does not represent the actual length of the sample, but calculated the length of the test results.

By the formula (14) corresponding to the sequential probability ratio test for the decision:

\[
\ln \frac{\beta}{1 - \alpha} < \left( \frac{\mu_1 - \mu_0}{\sigma_0^2} \right)^2 \sum_{i=1}^{n} (x_i) + \frac{n}{2\sigma_0^2} (\mu_1^2 - \mu_0^2) < \ln \frac{1 - \beta}{\alpha}
\] (15)

For signal samples \( X = [x_i, x_{i+1}, \ldots, x_{i+k}] \), \( i=1, \ldots, n, k=0 \). When \( i = 1 \) when the sample \( X \) is substituted into the formula (15) is calculated to obtain the signal samples likelihood ratio for value, if the value is less than the lower limit of \( A \) or greater than the upper limit \( B \) of lower values, the working end, taking \( i = i + 1, k = 0 \), the next one begins the test sample; if the logarithm likelihood ratio \( a \) and the lower limit value \( B \) between the command, the \( i = i + 1, k = k + 1 \), the above steps to continue, repeated cycles until all determined signal samples \( X = [x_i, x_{i+1}, \ldots, x_{i+k}] \) the likelihood ratio of the value of all state.

In the likelihood ratio after completion of numerical samples each state, we can calculate the sequential probability ratio test signal sample index, as follows:

Step 1: First to assignment \( s(i) \), each state is equal to the logarithm of the likelihood ratio;

Step 2: Make the data, and if \( s(i) \geq \ln B \), then let \( s_1(i) = \ln B \); if \( s(i) \leq \ln B \), then let \( s_1(i) = \ln A \); if \( \ln A < s(i) < \ln B \), then let \( s_1(i) = s(i) \).

Step 3: If \( s_1(i) > 0 \), let \( s_2(i) = \frac{s_1(i)}{\ln B} \); otherwise, let \( s_2(i) = \frac{s_1(i)}{\ln A} \).

Step 4: obtained by the above method, \( s_2(i) \in [-1, 1] \), where \( s_2(i) = 1 \) in the sample corresponding to the abnormal state. \( s_2(i) = -1 \) corresponds to the state of health of a sample, it is unable to determine between \((-1, 1)\) between the critical state.

**MSET and SPRT Applications**

**Data Standardization**

For the realization of the state on pump performance of fault prediction and health assessment, selection of pump actual operation data. Which collect data consists of electrical current, export pressure, liquid temperature, gas temperature, speeds, which showed in Figure 2.

In order to facilitate input of the network, improve the eigenvector of the clustering, normalized the feature vector with \( x* = (x - \text{mean}(x))/\text{std}(x) \), after Standardized, the characteristic vector for each element number between 0-1. which shown in Figure 3.

![Figure 2. Observation data of the nitrogen system.](image)

![Figure 3. The standardization data figure.](image)

**Liquid Nitrogen Pump MSET Memory Matrix Structure**

Memory matrix D is the key to the multivariate state estimation, is also the main factors influencing the estimation results.
According to the basic principle of memory matrix $D$ state choice, choose low temperature flow of liquid nitrogen pump in evaluation period of extreme value as a sequence of data in memory matrix $D$. Then according to the training matrix $T$ and memory matrix $D$, $L$ the relation between the residual matrix $T = DUL$ determined the remaining training data matrix $L$.

Memory matrix and the remaining matrix selected, the calculation will need to estimate the matrix, the matrix $X_{\text{obs}}$ observation data into the equation (7) is obtained by the estimated data matrix $X_{\text{est}}$, then equation (8) obtained by the remaining training data matrix calculation of $L$ estimation matrix $L_{\text{est}}$. According to formula (9) and (10) to obtain actual pump liquid nitrogen residual health residuals. Figure 4 shows the actual residual $R_{\text{a}}$, Fig. 5 shows a healthy residual $R_{\text{h}}$.

By the residual plots apparent liquid nitrogen cryogenic pump significant fault has occurred in the period from 800 to 1000, it has been a marked liquid nitrogen pump cavitation phenomenon.

**Liquid Nitrogen Pump SPRT**

In order to further test the effectiveness of system for sequential probability ratio test of existing data. After calculating the likelihood ratio value is shown in Figure 6, from it can also be in the low temperature liquid nitrogen pump in the work time of 800 ~ 1000 this time clearly happened fluctuations, and the residual value, but also from the picture to see other place also obviously have change, say abnormal data at this time have been to. Further converted to get SPRTIndex in Figure 7, more obvious exceptions can be found from the figure.

**Conclusion**

This paper application of MSET and SPRT model for high pressure nitrogen system, which can solve numerous historical fault data under the condition of liquid nitrogen pump health state estimation problem. The results show that it can effectively and practically.

**References**
