

Performance Degradation Assessment of Rolling Bearing Based on both ADMM and Sparse Combination Learning

Fei Luo^a, Yuhua Zhou^b

School of Automation Science and Engineering, South China University of Technology, Guangzhou 510640, China

^aaufeiluo@scut.edu.cn, ^bzhouyuhua0806@163.com

Keywords: rolling bearing performance degradation assessment, Convolutional Sparse Combination Learning, ADMM algorithm, IMS.

Abstract. In order to solve the problem of evaluating performance degradation of rolling bearing, this paper proposes a model containing both Alternating Direction Method of Multipliers (ADMM) and Convolutional Sparse Combination Learning. We first introduce the model during learning phase and test phase. Finally, we do an experiment on a database provided by the Center for Intelligent Maintenance Systems (IMS), University of Cincinnati. From experimental results, the proposed model is feasible.

1. Model Algorithm

1.1 Learning Phase

The learning algorithm is named as Combination Learning because Dictionary Set is a Combination of sub-dictionary. The algorithm of Combination Learning learns from Training Set to generate Dictionary Set by iterative method. During each iteration, we learn to generate a sub-dictionary that can cover as many training samples as possible, and then let residual samples which can't be covered by the existing sub-dictionary form a new Training Set for the next iteration.

Assume that the current state is in the m -th iteration. The convolutional sparse coding can be stated as a bi-convex problem

$$\min_{s_m, x_{l,n}} \frac{1}{2} \sum_{n=1}^{\tilde{N}} \left\| y_n - \sum_{l=1}^L s_{l,m} \otimes x_{l,n} \right\|_2^2 + \beta \sum_{n=1, l=1}^{\tilde{N}, L} \|x_{l,n}\|_1 \quad (1)$$

$$s.t. \quad s_m = [s_{1,m}, \dots, s_{L,m}], \quad \|s_{l,m}\|_2^2 \leq 1$$

where y_n is a sample and \tilde{N} is the No. of training samples in Training Set during the m -th iteration. $x_{l,n}$ is sparse representation and $s_{l,m}$ is corresponding convolution atom, s_m is a sub-dictionary. \otimes is convolution operator. The No. of convolution atoms L and positive scalar β are user parameters.

We will discuss in detail how to generate a sub-dictionary s_m in three parts: (A) Sparse Representations learning, (B) The m -th sub-dictionary learning, (C) Training Set updating.

(A) Sparse Representations learning

During the process of Sparse representations learning in the m -th iteration, sub-dictionary s_m is fixed, and then we start with the minimization of (1) over sparse representations x_n

$$\min_{x_n} \frac{1}{2} \sum_{n=1}^{\tilde{N}} \|y_n - S_m x_n\|_2^2 + \beta \sum_{n=1}^{\tilde{N}} \|x_n\|_1 \quad (2)$$

where S_m is an operator composed of convolutions with sub-dictionary s_m , x_n is a vector of vectorized sparse representations $x_{l,n}$. The other symbols are the same as (1).

We solve (2) by using Alternating Direction Method of Multipliers (ADMM) [2]:

$$x_n^{k+1} = \arg \min_{x_n} \sum_{n=1}^{\tilde{N}} \left(\frac{1}{2} \|y_n - S_m x_n\|_2^2 + \frac{\rho}{2} \|x_n - z_n^k + u_n^k\|_2^2 \right) \quad (3)$$

$$z_n^{k+1} = \arg \min_{z_n} \sum_{n=1}^{\tilde{N}} \left(\beta \|z_n\|_1 + \frac{\rho}{2} \|x_n^{k+1} - z_n + u_n^k\|_2^2 \right) \quad (4)$$

$$u_n^{k+1} = u_n^k + x_n^{k+1} - z_n^{k+1} \quad (5)$$

where k is the No. of iterations of ADMM.

(3) is the quadratic problem whose optimal solution is where its derivative equals to zero:

$$x_n^{k+1} = \left(\mathbf{I} / \rho - (1/\rho)^2 S_m^T (S_m S_m^T / \rho + \mathbf{I})^{-1} S_m \right) (S_m^T y_n + \rho a) \quad (6)$$

where $a = z_n^k - u_n^k$, \mathbf{I} is unit matrix, $(S_m S_m^T / \rho + \mathbf{I})^{-1} = \left(1 + 1/\rho \sum_{l=1}^L |\hat{s}_{l,m}|^2 \right)^{-1}$, $|\hat{s}_{l,m}|$ are magnitudes of the coefficients of the Fourier transform of $\mathbf{s}_{l,m}$ [3].

We solve (4) by using the algorithm called the soft-thresholding operator [4]:

$$z_n^{k+1} = \begin{cases} 0, & \text{if } |a| \leq \varphi \\ \text{sgn}(a) \cdot (|a| - \varphi), & \text{otherwise} \end{cases} \quad (7)$$

where $a = x_n^{k+1} + u_n^k$, $\varphi = \beta/\rho$, $\text{sgn}(\cdot)$ is symbolic function.

(B) The m-th sub-dictionary learning

During the process of the m-th sub-dictionary learning, we keep sparse representations x_n fixed, and then start with the minimization of (1) over sub-dictionary \mathbf{s}_m

$$\min_{s_m} \sum_{n=1}^{\tilde{N}} \left(\frac{1}{2} \|y_n - X_n \cdot s_m\|_2^2 \right) \quad (8)$$

$$\text{s.t. } s_m = [s_{1,m}, \dots, s_{L,m}]^T, \quad \|s_{l,m}\|_2^2 \leq 1$$

where X_n is an operator composed of convolutions with sparse representations x_n , $\mathbf{s}_{l,m}$ is atom of sub-dictionary \mathbf{s}_m , L is the No. of atoms of \mathbf{s}_m . The other symbols are the same as (1).

We solve (8) by using ADMM [2]:

$$s_m^{k+1} = \arg \min_{s_m} \sum_{n=1}^{\tilde{N}} \left(\frac{1}{2} \|y_n - X_n s_m\|_2^2 + \frac{\rho}{2} \|s_m - v_m^k + u_m^k\|_2^2 \right) \quad (9)$$

$$v_{l,m}^{k+1} = \arg \min_{v_{l,m}} \left(g(v_{l,m}) + \frac{\rho}{2} \sum_{n=1}^{\tilde{N}} \|s_{l,m}^{k+1} - v_{l,m} + u_{l,m}^k\|_2^2 \right) \quad (10)$$

$$u_m^{k+1} = u_m^k + s_m^{k+1} - v_m^{k+1} \quad (11)$$

where k is the No. of iterations of ADMM, $s_m = [s_{1,m}, \dots, s_{L,m}]^T$, $v_m = [v_{1,m}, \dots, v_{L,m}]^T$, $u_m = [u_{1,m}, \dots, u_{L,m}]^T$.

The solving method of (9) is the same as (3):

$$s_m^{k+1} = \left(\mathbf{I} / \rho - (1/\rho)^2 X_n^T (X_n X_n^T / \rho + \mathbf{I})^{-1} X_n \right) (X_n^T y_n + \rho a) \quad (12)$$

where $a = v_m^k - u_m^k$, \mathbf{I} is unit matrix, $(X_n X_n^T / \rho + \mathbf{I})^{-1} = \left(1 + 1/\rho \sum_{l=1}^L |\hat{x}_{l,n}|^2 \right)^{-1}$, $|\hat{x}_{l,n}|$ are magnitudes of the coefficients of the Fourier transform of $x_{l,n}$ [3].

The proximal operator of (12) is equivalent to projection onto convex set \mathbb{C} [3]:

$$v_{l,m}^{k+1} = \Pi_{\mathbb{C}}(b) = \begin{cases} b/\|b\| & \text{if } \|b\|_{\mathbf{s}} > 1 \\ b \cdot \mathbf{1}_{\mathbf{s}} & \text{otherwise} \end{cases} \quad (13)$$

where $\Pi_{\mathbb{C}}(b)$ indicates that b is projected onto the convex set \mathbb{C} , $b = 1/\tilde{N} \cdot \sum_{n=1}^{\tilde{N}} (s_{l,m}^{k+1} + u_{l,m}^k)$, $\mathbf{1}_{\mathbf{s}}$ is a mask taking on 1 on support set \mathbf{s} and 0 otherwise. $\mathbb{C} = (\|v_{l,m}\|_{\mathbf{s}} \leq 1) \cap (\text{supp}(v_{l,m}) \subseteq \mathbf{s})$, \mathbf{s} is support set of $v_{l,m}$. The other symbols are the same as the above.

(C) Training Set updating

Based on sub-dictionary \mathbf{s}_m , we design the convolutional sparse reconstruction cost $\mathcal{C}_{n,m}$ of sample x_n as the following

$$C_{n,m} = \frac{1}{2} \left\| y_n - \sum_{l=1}^L s_{l,m} \otimes x_{l,n} \right\|_2^2 + \beta \sum_{l=1}^L w_{l,m} \|x_{l,n}\|_1$$

$$w_m = \{w_{1,m}, \dots, w_{L,m}\}$$
(14)

where $w_{l,m}$ is weight coefficient. $w_{l,m} = 1 - t_l / \sum_i t_i$, $t_l = \sum_n \|x_{l,n}\|_2^2$.

Remark: basic idea of designing $w_{l,m}$ is that the greater the sparse representation value is, the smaller the contribution of the corresponding atom to the cost $C_{n,m}$ is.

We summarize Training Set updating into Algorithm 1.

Algorithm 1. The update of Training Set

1. Initialize: $m = 1$, $Y_{m=1}$ is initial Training Set, Dictionary Set $S = \emptyset$, Weight Set $W = \emptyset$, λ_m is a vector of empiric values;
2. **while**
3. Obtain x_n, s_m : using parts of (A) and (B);
4. compute $C_{n,m}, w_m$: using (14);
5. Update S, W and m : add s_m into S , add w_m into W , $m++$;
6. Update Training Set Y_m : $Y_m = \begin{cases} \text{sample } y_n \text{ remains, if } C_{n,m} > \lambda_m; \\ \text{sample } y_n \text{ removes, otherwise} \end{cases}$;
7. **until** Y_m becomes Empty Set ;
8. Output: Dictionary Set S , Weight set W .

Remark: $\lambda_m = 1.1^m \cdot \lambda_0$, where λ_0 is equals to the value ranking 20% position in ascending order of $C_{n,m=1}$.

1.2 Test Phase

We can obtain Dictionary Set S and Weight Set W after learning phase of algorithm. During test phase, we will evaluate the performance degradation of rolling bearings based on S and W .

Based on Dictionary Set S obtained in learning phase, we start with the minimization of (1) over sparse representations x_t with regard to testing sample y_t

$$\min_{x_t} \frac{1}{2} \|y_t - S_m x_t\|_2^2 + \beta \|x_t\|_1$$
(15)

where S_m is an operator composed of convolutions with sub-dictionary s_m .

Since the solving process of (15) is similar to (2), we give its optimal solution directly:

$$x_t^{k+1} = \left(I / \rho - (1/\rho)^2 S_m^T (S_m S_m^T / \rho + I)^{-1} S_m \right) (S_m^T y_t + \rho a)$$
(16)

where $a = z_t^k - u_t^k$. The other symbols are the same as the above.

$$z_t^{k+1} = \begin{cases} 0, & \text{if } |a| \leq \varphi \\ \text{sgn}(a) \cdot (|a| - \varphi), & \text{otherwise} \end{cases}$$
(17)

where $a = x_t^{k+1} + u_t^k$, $\varphi = \beta / \rho$, $\text{sgn}(\cdot)$ is symbolic function.

Based on sub-dictionary s_m , weight coefficients w_m and sparse representations x_t , we can obtain the cost $C_{t,m}$ of testing sample y_t

$$C_{t,m} = \frac{1}{2} \left\| y_t - \sum_{l=1}^L s_{l,m} \otimes x_{l,t} \right\|_2^2 + \beta \sum_{l=1}^L w_{l,m} \|x_{l,t}\|_1$$
(18)

where $s_m = [s_{1,m}, \dots, s_{L,m}]$, $x_t = [x_{1,t}, \dots, x_{L,t}]$, $w_m = [w_{1,m}, \dots, w_{L,m}]$.

Finally, we can obtain the best performance assessment Φ with regard to the testing sample y_t :

$$\Phi = \min \{ (C_{t,1} - \lambda_1), \dots, (C_{t,m} - \lambda_m), \dots, (C_{t,M} - \lambda_M) \}$$
(19)

where $m \in [1, M]$, M is the No. of sub-dictionary.

Our conclusions about Φ are: (1) If $\Phi \leq 0$, the rolling bearing has good performance; (2) otherwise there is degradation in the bearing, and the greater the value of Φ is, the more severe the bearing degradation is.

2. Experimental verification

The vibration data of rolling bearings is from the Center for Intelligent Maintenance Systems (IMS), University of Cincinnati [5].

Rolling bearing chosen for our testing is Bearing 3 of Set No.1 in IMS. We assume that the bearing has good performance in the beginning period of experiment. Therefore we choose data from the first ten days as Training Set (the total No. of training samples is 368) and other residual data as Testing Set (the total No. of testing samples is 1788).

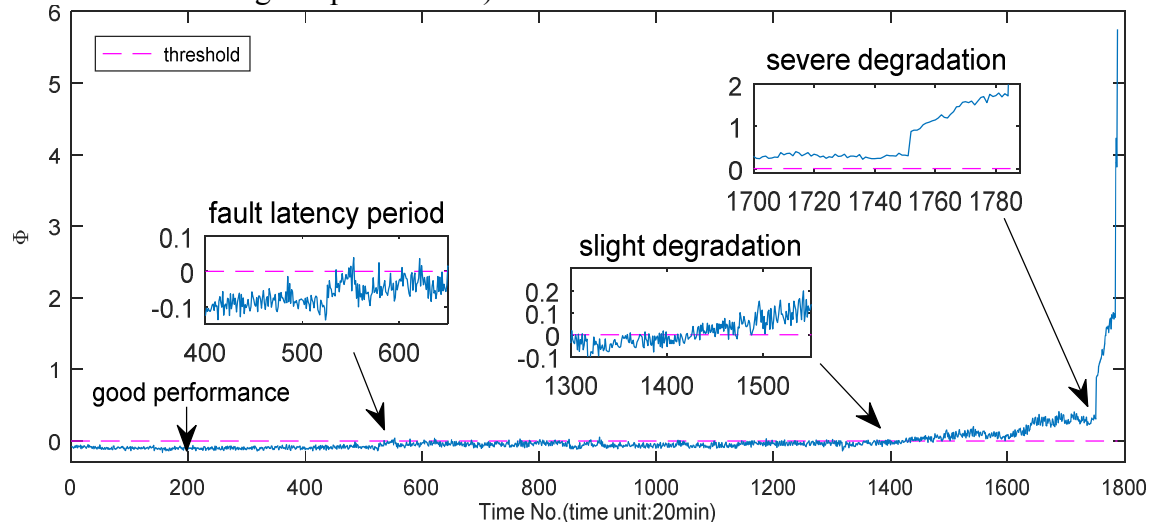


Fig. 1 The curve of ϕ -value during the entire life of bearing.

(Remark: sub-figure are enlarge part figures)

Fig. 1 shows the curve of performance assessment ϕ during the entire life of bearing. We can obtain the following results by observing the trend of curves in Fig 1:

(1) The ϕ value gradually increases with time, which means the bearing performance gradually degrades. This trend is in line with true situation of device degradation.

(2) The whole life of the rolling bearing is divided into four stages by the curve trend: Good Performance, Fault Latency Period, Slight Degradation, and Severe Degradation. All ϕ values in Good Performance period are less than zero. The minority of ϕ values in Fault Latency Period are more than zero. Almost all ϕ values in Slight Degradation period are more than zero and their values are very close to zero. All ϕ values in Severe Degradation period are much more than zero.

(3) Since the Fault Latency Period lasts much long time, Maintenance Engineers can arrange in advance to perform maintenance especially in the case of understaffed.

(4) We can find the best maintenance time, which is the Slight Degradation period.

(5) It lasts very short in Severe Degradation period. Once it step into the Severe Degradation period, the bearing will soon lose its ability to work. Therefore maintenance should be finished before it step into the Severe Degradation period.

ϕ can be a quantitative indicator of the bearing performance degradation according to the above analytic results, and the proposed model is feasible.

3. Conclusion

In order avoid the unnecessary economic loss and even casualties caused by the degradation of rolling bearing, we propose a method for evaluating the rolling bearing performance degradation based on Convolution Sparse Combination Learning. The main innovations are the following:

(1) In the study of bearing performance degradation assessment, there are few existing papers to use model based on sparse coding. However, the vibration signal of rolling bearing is periodic. The "convolution" operator in convolution sparse coding has shift-invariance property which is corresponding to the periodicity of the bearing vibration signal. Therefore convolution sparse coding is proper.

(2) It is difficult to obtain the prior knowledge about the characteristics in the vibration signal in practical applications. Compared with the traditional method, the proposed model, which does not need these prior knowledge, has some application value in practice.

(3) The proposed model has real-time property. In order to greatly reduce its time-complexity, ADMM algorithm and Combination Learning are applied to the model.

(4) The whole bearing life is clearly divided into four stages by the model: Good Performance, Fault Latency Period, Slight Degradation, and Severe Degradation. This result means that the proposed can be a quantitative indicator of the bearing performance degradation and that the proposed model based on Convolutional Sparse Combination Learning is feasible.

Our future works include (1) applying the proposed method to performance degradation assessment of other rotating machinery such as gears and (2) studying the method of predicting the residual life of equipment combined with sequence models.

Acknowledgments

This work was financially supported by Science and Technology Planning Project of Guangdong Province, China (Grant No. 2016B090918028), and Science and Technology Planning Project of Guangzhou City, China (Grant No. 201604010032).

Corresponding author: LUO Fei (1957-), male, Professor, Ph.D. Research direction: Mechanical equipment condition monitoring, fault diagnosis and so on. E-mail: aufeiluo@scut.edu.cn.

References

- [1]. Pan Y N. Study on Feature Extraction and Assessment Method of Rolling Element Bearing Performance Degradation [D]. Shanghai Jiao Tong University, 2011.
- [2]. Boyd S, Parikh N, Chu E, et al. Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers [J]. Foundations & Trends® in Machine Learning, 2010, 3 (1):1-122.
- [3]. Sorel M, Sroubek F. Fast convolutional sparse coding using matrix inversion lemma [J]. Digital Signal Processing, 2016, 55:44-51.
- [4]. Donoho D L. De-noising by soft-thresholding [M]. IEEE Press, 1995.
- [5]. "Bearing Data Set" was provided by the Center for Intelligent Maintenance Systems (IMS), University of Cincinnati. (<https://ti.arc.nasa.gov/tech/dash/pcoe/prognostic-data-repository/>)