Health monitoring of a shaft transmission system via hybrid models of PCR and PLS

YI FANG, HYUNWOO CHO, MYONGKEE JEONG*

Department of Industrial & Information Engineering, University of Tennessee, Knoxville, TN, USA

***Abstract*:** Prediction of motor shaft misalignment is essential for the development of effective coupling and rotating equipment maintenance information systems. It can be stated as a multivariate regression problem with ill-posed data. In this paper, hybrid models of principal components regression (PCR) and partial least squares regression (PLS) have been proposed for this problem. The basic idea of hybrid models is to combine the merits of PCR and PLS to develop more accurate prediction techniques. Both the principal components defined in PCR and the latent variables in PLS are involved in a hybrid model. The experimental results show that an optimal hybrid model can outperform PCR and PLS, especially when the number of predictor variables increases. It suggests that the proposed approach may be particularly useful for complex prediction tasks that need more predictor variables. Discussions for future research are also presented.

***Keywords*:** Principal component regression, partial least square, motor shaft misalignment.

1 Introduction

A shaft transmission system is one of the most important and fundamental parts of rotary machinery. A proper shaft alignment is indispensable because it reduces excessive axial and radial forces on the most vulnerable parts of a machine system, such as the bearings, seals, and couplings (Wowk, 2000). It also minimizes the amount of shaft bending, thereby permitting full transmission of power from the driver machine to the driven machine and eliminates the possibility of shaft failure from cyclic fatigue. In addition, it minimizes the amount of wear in the coupling components, reduces mechanical seal failure, and lowers vibration levels in machine casings, bearing housings, and rotor. Therefore, monitoring and predicting condition of the shaft alignment is crucial in order to make intelligent decisions on when to perform alignment maintenance, which plays an essential role in increasing maintenance effectiveness and reducing maintenance costs. Prediction of motor shaft misalignment can be stated as a multivariate regression problem in which the number of predictors greatly exceeds the number of observations. It is also known as an ill-posed problem. Several regression models have been developed for condition monitoring with ill-posed data. Among them, the dominant two used in practice are principal-components regression (PCR) (Chatterjee and Price, 1977) and partial least squares (PLS) (Wold, 1966). An overview of PCR and PLS in condition monitoring was given by Hoskuldsson (1988) and by Joliffe (2002). The relative strengths of these two approaches are often discussed and debated, but no clear conclusion has been reached. PLS is generally regarded as being superior to PCR in prediction. However, a few cases have shown that PCR gave better prediction results than PLS (Vigneau *et al.*, 1996). Furthermore, no theoretical studies suggested that one method should predict better than the other (Wentzell *et al.*, 2003). PCR and PLS have their own unique strength and weakness although they are very similar in some regards.

The goal of this work is to combine the merits of PCR and PLS, aiming at developing more accurate prediction techniques for the health monitoring of rotating machinery. Hybrid models of PCR and PLS have been proposed in this paper. Principal components (PCs) defined in PCR as well as latent variables (LVs) in PLS can be regarded as different features of the data. Traditional PCR or PLS only extracts one kind of feature from the data: maximum variance direction in PCR and maximum correlation direction with response variables in PLS. If these two kinds of features are included in a common regression model, the prediction performance

* Corresponding author.
Address: 311 East Stadium Hall, Knoxville, TN 37996-0700, USA
Phone: 01-865-974-0234; Fax: 01-865-974-0588; E-mail: mjeong@utk.edu

may be improved. Therefore, the key of hybrid models is that the linear transformed vector may be either a principal component or a latent variable.

The remainder of this paper is organized as follows: Section 2 addresses the problem of motor shaft misalignment. In Section 3, hybrid models of PCR and PLS are introduced and the algorithm for building the optimal hybrid model is also given. A case study is then presented in Section 4. The conclusion and possible directions for future research are included in Section 5.

2. Shaft misalignment

Shaft misalignment is one of the most common faults associated with rotating machinery, and occurs when the shaft of the driven machine and the shaft of its driver machine do not rotate on a common axis; that is, the shafts are not coaxial. Shaft misalignment is a measure of how far apart the two centerlines are from each other (Kuropatwinski, 1997). Such shift in centerlines can be in a parallel position (when the centerlines of the two shafts are parallel with each other, but not aligned along the same axis), in an angular position (when the centerlines are at an angle to each other), or in a combination of these positions as shown in Fig. 1.

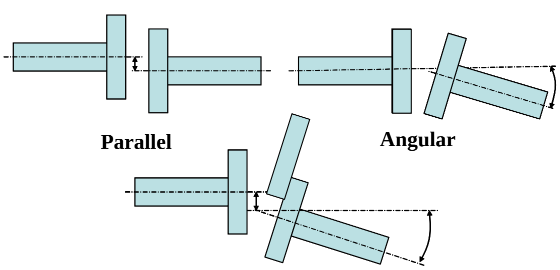

Fig.1. An illustration of parallel, angular, and combined misalignment conditions

Monitoring and prediction of motor shaft alignment is important for making intelligent decisions when performing alignment maintenance, which plays an essential role in increasing maintenance effectiveness and reducing maintenance costs. Recent studies did indicate that there is a measurable change in the input power frequency spectrum of an electric motor for different shaft alignment conditions (Hines, 1999). Thus, the purpose of this research is to develop the optimal hybrid model to predict electric motor shaft alignment conditions based on the motor's power frequency spectrum.

3 Hybrid models of PCR and PLS

3.1. *Illustration of hybrid models*

Both PCR and PLS can be formulated in a similar iterative way. The difference between them only lies in the objective functions [Bennett, 2003]. If the various objective functions are involved in a common iterative procedure, the properties of both PCR and PLS can be combined. Thus, the idea of constructing hybrid models of PCR and PLS consists of two steps. The first is to calculate PC and LV alternately in iterative steps. In this way, the orthogonal decompositions are mixed with PCs and LVs. Based on these orthogonal decompositions, the original input data is mapped into a new subspace. The second step is to make the final regression function by minimizing the least-squares error between the projected data and the response (y). The key of hybrid models lies in that the projected vector in every orthogonal decomposition could be either a principal component or a latent variable. Like PCR and PLS, when the number of components in a hybrid model reaches the number of original predictor variables, the hybrid model is equivalent to the ordinary least-square (OLS) regression technique.

As shown in Fig. 2, for the 3-dimensional data (dot), PCR sequentially calculates the first three principal components PC1, PC2 and PC3. By contrast, a hybrid model of PLS and PCR may not calculate PC2 after getting PC1, but in the space orthogonal to PC1, the first latent variable LV2 is calculated (the number 2 indicates that the computation is in the second iterative step).

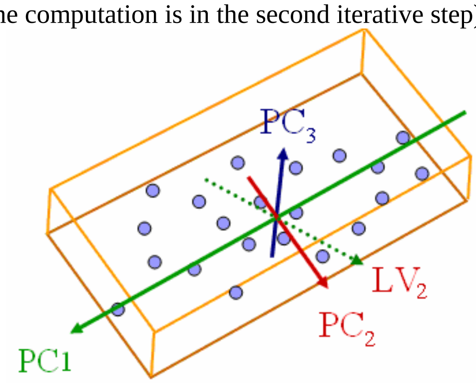

Fig.2. A possible hybrid model with PC1LV2

It has been shown that a hybrid model of PCR and PLS is generated by a combination of PCs and LVs. Different combinations create different hybrid models. Here a sequence is used to denote a hybrid model. For example, the sequence PC1LV2. The number following PC or LV means how many iterative steps (components) have already been calculated.

3.2. *Model selection*

If the number of predictor variables to retain is K, there could be 2^K different hybrid models. Note that among all possible combinations of PCs and LVs, the pure PCR or PLS models are actually included, such as the combination of PC1-PC2-PC3 or LV1-LV2-LV3. One effective way to choose the optimal hybrid model is based on cross validation (CV) error (Stone, 1977). The idea behind CV is to recycle data by switching the roles of training and test samples. Based on the sample data available, different CV methods can be selected, including hold-out, k-fold, and leave-one-out methods.

In order to evaluate the predictive performance of every hybrid model, each different combination of components would be denoted by a different integral value, called the *determinant*, of a sequence. In this study, a binary numeral system is used to calculate the determinant. In a combination sequence, PC and LV are replaced by 0 and 1 respectively. So every sequence can be represented by a binary number $b_K b_{K-1} \ldots\ldots b_2 b_1$, where $b_i \in \{0,1\}$. The determinant of a sequence is computed by converting the binary number to a decimal number. For example, the sequence LV1-PC2-LV3 is denoted by $(101)_2$ and thus its determinant is 6.

3.3. The algorithm

In this section, the algorithm of constructing the optimal hybrid model of PCR and PLS is given. For convenience of presentation, fold-out cross validation is adopted in this algorithm flow. The only parameter of the algorithm is K, the number of components to retain.

Input: training input $\mathbf{X}$ and response $\mathbf{y}$, validation input $\mathbf{V}$ and response $\mathbf{z}$

Output: the optimal hybrid regression model

1. For $j = 0$ to (2^k -1) {
2. $b_K b_{K-1} b_{K-2} \cdots b_2 b_1 = (j)_2$

   /* convert j to a binary string
3. For $i=1$ to K {
4. If $b_i = 0$

   Then ($\mathbf{w_i}$, $\mathbf{t}_i$) = PCA($\mathbf{X^i}$)

   /* calculate the first PC $\mathbf{w_i}$ and scores $\mathbf{t}_i$ based on the training input residual

   $\mathbf{X^i}\mathbf{X}^{i+1} = \mathbf{X}^i - \mathbf{t}_i \mathbf{t}_i^T \mathbf{X^i}$

   /* update the training input residual

   Else ($\mathbf{w_i}$, $\mathbf{t}_i$) = PLS($\mathbf{X^i}$, $\mathbf{y}$)

   /* calculate the first LV $\mathbf{w_i}$ and scores $\mathbf{t}_i$ based on the training input residual $\mathbf{X^i}$ and response $\mathbf{y}$

   $\mathbf{X}^{i+1} = \mathbf{X}^i - \mathbf{t}_i \mathbf{t}_i^T \mathbf{X}^i$

   /* update the training input residual

   }
5. $\mathbf{S} = \left(\mathbf{T^T T}\right)^{-1} \mathbf{T^T y}$

   /* construct the least-square regression model based on the training projected data $\mathbf{T}$ and response $\mathbf{y}$
6. Calculate the test projected data $\mathbf{Q}$. Note that the test data $\mathbf{V}$ has to be projected into the subspace $\mathbf{w}$ using the same mapping as the training data did
7. $error = \|z - Qw\|$

   /* using the model calculated by step 5 to calculate the cross-validation error on the test data

   }
8. Choose the optimal regression model with the minimum cross-validation error

The proposed algorithm considers all possible combinations of PCs and LVs, and thus has exponential computation complexity. However, the K could be very small in ill-posed problems because only a few components contain most data information.. Therefore, the algorithm could be computationally efficient for ill-posed problems.

4 Experiments

4.1. *Experimental data*

Data was obtained from Oak Ridge National Laboratory Advanced Motor Testing Facility using a three-phase 50-horsepower AC motor attached through three different flexible couplings to a 150-horsepower dynamometer at controlled parallel and angular offset conditions. The data was acquired by digitizing two of the three phase currents and voltages off of the digit tape at a sampling rate of 6000 Hertz. The third of the three phase currents and voltages were calculated using a variation of Ohm's law. The total power for each event was then calculated by summing the three individual powers. There are some differences in the power time domain for each alignment condition, such as different frequency components, or even some very small differences in magnitude. Advanced signal processing techniques were then used to transform the raw voltage and current data into the power time waveform. Therefore, for this analysis the input data is the power frequency spectrum and the response data is the misalignment condition. The response data ranges from 0 to 50 mils (1mil= 2.540×10^{-5} m) for the parallel offset and from 0 to 15 mils for the angular offset. The sizes of the entire input set and response set are 50×3000 (50 time series were collected, each with 3000 points) and 50×2 respectively. The condition number of the input data is 6.1246e+093, so it is a typical ill-posed problem. The objective then is to use the optimal hybrid model of PCR and PLS to predict the misalignment condition based on the power frequency spectrum. Although the response data has two variables, parallel misalignment and angular misalignment, they are predicted independently in this experiment

4.2. *Choosing the number of predictors in hybrid models*

PCR and PLS algorithms were used to determine the parameter *K*, the number of predictors. Because only 50 observations were available, leave-one-out cross validation (LOOCV) has been used for accurately evaluating the prediction performance of regression models. Both PCR and PLS achieve the minimum LOOCV MSEs with the number of predictors at 9 on this dataset. Furthermore, the first 9 PCs can explain the 99.64% variance of the input data. For the angular response, similar results were produced. Therefore, the number of predictor variables to retain in hybrid models was specified as 9. Although the original input data has a large number of variables (i.e. 3000), most data information is actually contained in much fewer components (i.e. 9). Thus, the proposed algorithm with exponential complexity appears to be computationally efficient.

4.3. *Experimental results*

As *K* is specified as 9, the optimal model chosen by the proposed algorithm for the parallel condition is that with a determinant of 75 equivalent to binary 001001011. Therefore the corresponding combination sequence of PCs and LVs is PC1-PC2-LV3-PC4-PC5-LV6-PC7-LV8-LV9. The optimal prediction model for angular condition is that with a determinant of 99 whose corresponding combination sequence is PC1-PC2-LV3-LV4-PC5-PC6-PC7-LV8-LV9 (001100011). Note that, the model with determinant of 0 is a PCR model and that with determinant of 511 is a PLS model. The results show that the optimal hybrid models outperform PCR and PLS in both parallel and angular misalignment prediction conditions.

Table 1 contains all experimental results when *K* covers from 2 to 10. These experimental results validate that the optimal hybrid models achieve the best prediction results when *K* is 9. The Table 1 also shows that the optimal hybrid model predicts more accurately than PCR and PLS when *K* is greater than 3. It suggests that the proposed approach may be particularly useful for complex prediction tasks that need more predictors. In addition, the MSEs for angular offset are much smaller than the MSEs for parallel offset, which implies that modeling the parallel offset is slightly difficult, at least for the given training data.

5 Conclusion and future work

In this paper, a new multivariate regression tool for modeling shaft misalignment has been proposed. It aims to develop more accurate prediction models by combining the merits of PCR and PLS. The experimental results presented have shown great potential use of such work in prediction tasks. Future research includes creating nonlinear hybrid models since the relationship between machine misalignment and machine vibration is nonlinear (Hines, 1999). A possible solution is based on Kernel Trick (Scholkopf *et al.*, 1998), which has been proven an efficient approach to deal with nonlinear problems. Kernel PCR and Kernel PLS have been proposed recently and achieve good prediction results (Rosipal, 2001). Thus, it is expected that the Kernel hybrid models of PCR and PLS could work well on some nonlinear cases. Furthermore, model selection method in this paper has an exponential algorithm complexity. Although in many ill-posed cases the number of significant components is as small as 10 or fewer, it may be necessary in some instances to develop more efficient model selection methods by constraining the search for candidate hybrid models.

Table 1. The LOOCV error rates with PCR, PLS and the optimal hybrid model

K	**PCR**		**PLS**		**Optimal Hybrid Model**	
	Parallel	Angular	Parallel	Angular	Parallel	Angular
10	6.3259e-2	1.2564e-3	8.2317e-2	5.6314e-4	3.6479e-3	8.7461e-5
9	**4.5123e-3**	**3.4315e-4**	**1.4657e-3**	**7.4153e-5**	**2.5559e-4**	**5.5223e-6**
8	1.2684	3.5749e-2	1.0361e-2	1.0129e-4	9.3247e-4	1.2479e-5
7	4.3695	1.9654	9.8621e-2	9.1476e-4	5.2947e-3	9.5514e-5
6	91.237	2.2143	0.1579	6.3471e-3	1.2568e-2	3.2694e-4
5	98.865	9.9176	0.3214	1.0874e-2	0.0974	6.3247e-3
4	96.364	27.3695	1.8694	6.3727e-2	1.6987	0.8591e-2
3	227.68	25.9873	1.5697	0.1458	20.317	2.3697
2	235.41	25.416	3.1843	1.5147	41.585	8.1211

References:

Bennett, K. P. and Embrechts, M. J. (2003). An optimization perspective on partial least squares, In *Proceedings of the NATO Advanced Study Institute on Learning Theory and Practice*, IOS Press Amsterdam, 190, 227-250.

Casada, D. A. (1996). Using motor data as a diagnostic tool: A collection of scattered thoughts. Presentation given at the University of Tennessee, Knoxville.

Chatterjee, S., and Price, B. (1977). *Regression Analysis by Examples*. Wiley and Sons, New York and Toronto.

Chawla, J.S. (1990) A note on ridge regression. *Statistics & Probability Letters*, 9: 343-345.

Eisenmann, R. C., Sr. (1998). *Machinery Malfunction Diagnosis and Correction*. New Jersey: Prentice Hall PTR.

Geladi, P., & Kowlaski B. (1986). Partial least square regression: A tutorial. Analytica Chemica Acta, 35, 1–17.

Golub, G.H., Heath, M., and Wahba, G., (1979). Generalized cross-validation as a method for choosing a good ridge parameter, *Technometrics*, 21:2, 215-223.

Hines, J., Wesley, J., Edmondson, A., and Nower, D. (1999). Motor shaft misalignment versus bearing load analysis. *Maintenance Technology*, 5, 11-17.

Hoskuldsson, A. (1988). PLS regression methods. *Journal of Chemometrics*, 2: 211-228.

Joliffe, I.T. (2002). *Principal Component Analysis* (2nd ed.). New York: Springer-Verlag.

Kuropatwinski, J. J., Jesse, S., Hines, J. W., Edmondson, A., and Carley, J. (1997) Prediction of motor misalignment using neural networks, *Proceedings of Maintenance and Reliability Conference (MARCON 97).*

Rosipal R., and Trejo, L.J. (2001) "Kernel Partial Least Squares Regression in Reproducing Kernel Hilbert Space." *Journal of Machine Learning Research.* 2:97-123.

Scholkopf B., Smola, A., and Muller, K. R. (1998) Nonlinear component analysis as a kernel eigenvalue problem. *Neural Computation*, 10:1299-1319.

Stone, M. (1977) Asymptotics for and against cross-validation, *Biometrika*, 64, 29-35.

Vigneau, E., Bertrand, D., and Qannari, M. (1996) Application of latent root regression for calibration in near-infrared spectroscopy. Comparison with principal component regression and partial least squares, *Chemometrics and Intelligent Laboratory Systems*, 35, 231-238.

Wentzell, Peter D., and Vega Montoto, Lorenzo. (2003) "Comparison of principal components regression and partial least squares regression through generic simulations of complex mixtures." *Chemometrics and Intelligent Laboratory Systems*, 65, 257– 279.

Piotrowski, J. (1995). *Shaft Alignment Handbook*. New York: Marcel Dekker, Inc.

Wold, H. (1962) Estimation of principal components and related models by iterative least squares. *Multivariate Analysis*. New York: Academic Press. 391-420.

Wowk, Victor (2000). *Machine Vibration: Alignment*. New York: McGraw Hill.