Fault detection and isolation with robust principal component analysis

Yvon THARRAULT, Gilles MOUROT, José RAGOT, Didier MAQUIN

Centre de Recherche en Automatique de Nancy
UMR 7039 CNRS-Nancy Université
2, Avenue de la forêt de Haye. F-54 516 Vandoeuvre-lès-Nancy Cedex
yvon.tharrault@ensem.inpl-nancy.fr

Abstract.
Principal component analysis (PCA) is a powerful fault detection technique which has been widely used in process industries. However, a main drawback of PCA is that it is based on least squares estimation techniques and hence fails to account for outliers which are common in physical processes. This paper is concerned with the fault detection and isolation problem. The proposed method does not require a data matrix without outliers for a PCA model design. Indeed, the approach directly uses the eventually corrupt database to elaborate a robust PCA model allowing fault detection. Then reconstruction principle and fault signatures analysis are used for fault isolation.

Keywords: Principal component analysis, robustness, fault signature, fault detection and isolation, outliers.

1. Introduction

Principal component analysis (PCA) is widely used as a multivariate statistical method for fault detection, isolation and diagnosis. PCA is mainly based on the description of linear relations between variables and optimises a MSE (Mean Square Error) criterion. It is well-known that the estimation based on a criterion like MSE is less robust to outliers than that resulting from other criteria like error absolute value (Hubert *et al*., 2005). Let us recall that the traditional approach of the *PCA* uses a preliminary calculation of the average of data and their covariance matrix; average and variance are sensitive to outliers, and the obtained results are often not exploitable because too biased by the influence of these outliers.

To take outliers into account, a robust covariance matrix of the data can be used to construct a robust principal component analysis. For that, Croux and Haesbroeck, (2000) built particular functions of influence and the asymptotic variances which result from them. Engelen *et al.*,(2005) proposed the robust approach *ROBPCA*, which combines revealing projections with a robust estimate of the variance matrix. This technique produces estimates which appear robust in the presence of outliers. Brown *et al.*,(2005) focused on the robust estimate of the covariance matrix for multidimensional systems. Other approaches dealing with the problem of robustness were proposed in (Böhning and Ruangroj, 2002) by using a scale-contaminated distribution law and in (Salibian-Barrera *et al*., 2006) where the authors develop an approach based on a moment calculation.

Our presentation is devoted to the problem of fault detection and isolation in data. In general, faults result from process dysfunctions or from the system of measurement acquisition. The contribution essentially deals with the detection and isolation of outliers by using complementary tools: robust principal component analysis, data reconstruction and residual analysis with the fault signatures. Section 2 is a short reminder, on one hand, of the principal component analysis in the traditional case and, on the other hand, of the robust principal component analysis. A detection and isolation procedure for outliers is proposed in section 3, then, in section 4, is applied to an example of synthesis emphasizing the generation of fault signatures.

2. PCA fault detection and isolation

Let us consider a data matrix $X \in \Re^{N \times n}$, with vector lines x_i^T, which gather N measurements collected on the n system variables.

2.1. Traditional approach

In the traditional *PCA* case, data are supposed to be collected on a system being in a normal process operation. *PCA* determines an optimal linear transformation of the data matrix X in terms of capturing the variation in the data:

$$T = XP \quad \text{et} \quad X = TP^T \tag{1}$$

with $T \in \Re^{N \times m}$ the principal component matrix and the matrix $P \in \Re^{m \times m}$ the one that contains the principal vectors which are the eigenvectors associated to the eigenvalues λ_i of the covariance matrix (or correlation matrix) Σ of X:

$$\Sigma = P\Lambda P^T \qquad \text{avec} \qquad PP^T = P^T P = I_m \tag{2}$$

where $\Lambda = diag(\lambda_1 \ldots \lambda_m)$ is a diagonal matrix with diagonal elements in decreasing magnitude order.

The relations (1) are meaningful when the dimension of the representation space is reduced. Once the component number ℓ to retain is determined, the data matrix X can be approximated. For that, the eigenvectors matrix is partitioned into the form:

$$P = \begin{pmatrix} \hat{P} & \tilde{P} \end{pmatrix} \qquad \hat{P} \in \mathcal{R}^{n \times \ell} \tag{3}$$

From the decomposition (1), $\hat{X}$ is the principal part of the data explained by the ℓ first eigenvectors and the residual part $\tilde{X}$ is explained by the remaining components:

$$\hat{X} = X\hat{P}\hat{P}^T = XC_\ell \tag{4}$$

$$E = X - \hat{X} = X(I - C_\ell) \tag{5}$$

where the matrix $C_\ell = \hat{P}\hat{P}^T$ is not equal to the identity matrix.

2.2. Robust approach

A major difficulty of *PCA* comes from its sensitivity to outliers. In order to reduce this sensitivity, various techniques are usable and in particular that which consists in carrying out *PCA* directly on the data possibly contaminated by outliers. An alternative is to seek principal directions robust to these outliers. Fekri and Ruiz-Gazen,(2003) define a "local"

matrix of variance in the sense that the suggested form tends to emphasize the contribution of close observations in comparison with distant observations (outliers). The matrix is defined in the following way according to the observations x_i:

$$V = \frac{\displaystyle\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} w_{i,j}(x_i - x_j)(x_i - x_j)^T}{\displaystyle\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} w_{i,j}} \tag{6}$$

where the weights $w_{i,j}$ themselves are defined by:

$$w_{i,j} = \exp\left(-\frac{\beta}{2}(x_i - x_j)^T \Sigma^{-1}(x_i - x_j)\right) \tag{7}$$

β being a turning parameter to reduce the influence of the observations faraway, the authors recommend a value close to 2. Thanks to the presence of adapted weights $w_{i,j}$, *PCA* can then be carried out on this "new" matrix of covariance considered robust with respect to outliers.

3. Fault detection

3.1. Data reconstruction

PCA can be used for fault detection, these faults resulting in outliers which are highlighted by projection onto the residual space (Dunia and Qin, 1998). The *PCA* model being known, a new measurement vector x can be decomposed as below:

$$x = \hat{x} + \tilde{x} \quad , \quad \hat{x} = C_\ell\, x \quad , \quad \tilde{x} = (I - C_\ell)\, x \tag{8}$$

where $\hat{x}$ and $\tilde{x}$ are respectively the projections of x on the principal space and residual space. From (8), it is possible to estimate a particular component of the vector x, for example the Rth, where R is a subset containing the indices of the reconstructed variables. However, the presence of outliers in the observation vector x returns the estimated $\hat{x}$ sensitive to this value. It is then preferable to express this estimated $\hat{x}$ by using only the fault-free part of the observation vector x.

The reconstruction (Dunia and Qin, 1998) of process faults consists in estimating the reconstructed vector $\hat{x}_R$ by eliminating the effect of the faults. Matrix Ξ_R indicates the reconstruction directions. This matrix is orthonormal with dimension ($n \times$ number of reconstructed variables) and is constructed with 0 and 1, where 1 indicates the reconstructed variables from the other variables (with 0), for example $\Xi_R = [0\ 1\ 0\ 1\ 0]^T$ for $n = 5$ and one reconstructed variable ($R = \{2, 4\}$).

The expression for the reconstruction $\hat{x}_R$ of the variable x is given by (Dunia and Qin, 1998):

$$\hat{x}_R = [I - \Xi_R(\tilde{\Xi}_R^T\tilde{\Xi}_R)^{-1}\tilde{\Xi}_R^T]x \tag{9}$$

where $\tilde{\Xi} = (I - C_\ell)\Xi_R$.
Let us note that if $(\tilde{\Xi}_R^T\tilde{\Xi}_R)^{-1}$ exists the Rth variable is completely reconstructable or else the fault is partially reconstructable. This two cases are presented by (Dunia and Qin, 1998). In the following, only completely reconstructable faults are considering.

3.2. Residual generation

In a diagnosis objective, residuals are generated for fault detection and isolation. The reconstruction procedure is successively applied to all the components of x. The reconstructions obtained (9) are then compared with the measurements. The residuals are obtained by projecting the reconstructed variables onto the residual space. Residuals are defined by $\tilde{x}_R$, projection of $\hat{x}_R$ onto the residual space:

$$\tilde{x}_R \quad = \quad P_R^{(\ell)} x \tag{10}$$

$$P_R^{(\ell)} \quad = \quad (I - C_\ell) - \tilde{\Xi}_R(\tilde{\Xi}_R^T \tilde{\Xi}_R)^{-1}\tilde{\Xi}_R^T \tag{11}$$

Remark 1. Matrix $P_R^{(\ell)}$ has the following proprieties:

$$P_R^{(\ell)}\Xi_R = 0 \qquad \text{and} \qquad \Xi_R^T P_R^{(\ell)} = 0 \tag{12}$$

It means that the components of $\tilde{x}_R$ are not sensitive to the Rth components of x. This remark can be used to identify which component of x is disturbed by faults.

For example, considering a measurement x composed with the true value x^*, a noise ϵ with null mean and one fault of amplitude d and direction Ξ_F, where F is a subset containing the indices of the reconstructed variables:

$$x = x^* + \epsilon + \Xi_F d \tag{13}$$

then the residual is:

$$\tilde{x}_R = P_R^{(\ell)}(x^* + \epsilon + \Xi_F d) = P_R^{(\ell)}(\epsilon + \Xi_F d) \tag{14}$$

and its expected value is:

$$E(\tilde{x}_R) = P_R^{(\ell)}\Xi_F d \tag{15}$$

-if the reconstruction direction Ξ_R is the same as the fault, i.e. if $R = F$,then all components of the vector $P_R^{(\ell)}\Xi_F$ are null and $E(\tilde{x}_R) = 0$

-if the reconstruction direction Ξ_R is different from the fault direction, then all components of the vector $P_R^{(\ell)}\Xi_F$ are a priori not null except the Rth components.

Then, the analysis of the residual amplitudes $\tilde{x}_R$ for all possible combinations shows the presence of faults and makes it possible to determine the components of the measurement affected by this fault.

4. Numerical example

4.1. Numerical example - mono-fault case

Data generation

A simple example based on four variables (x_1, x_2, x_3 et x_4) and two models is used. The data matrix X includes $N = 240$ measurements defined in the following way:

$$x_{i,1} = v_i^2 + 1 + \sin(0.1i), \quad v_i \sim \mathcal{N}(0,1) \tag{16}$$

$$x_{i,2} = x_{i,1}, \quad x_{i,3} = -2x_{i,1}, \quad x_{i,4} \sim \mathcal{N}(0,1)$$

Realizations of centered normal distributions with the same standard deviation equal to 0.02 are added to these four variables. The variable x_4, independent of other variables, is a perturbation for *PCA*. A constant bias of amplitude equal to 3 simulates the presence of outliers δx_1, δx_2, δx_3 affecting the variables x_1, x_2 and x_3: from 24 to 44 for the variable x_1, from 80 to 100 for the variable x_2, from 140 to 160 for the variable x_3. It's important to notice that 60 observations contained abnormal values, hence 25 percent of the data are contaminated by these values. The objective is to detect and especially isolate them.

Sensitivity analysis and theoretical fault signature

The data in the table 1 summarize the relationship between residual sensitivity $\tilde{x}_R$ (10) and outliers or faults δx_1, δx_2 and δx_3 (fault δx_4 on variable x_4 is not considered). This table was constructed while taking into account proprieties of the matrix $P_R^{(\ell)}$ (12). For example, the first four residuals $\tilde{x}_{11}$ to $\tilde{x}_{14}$ (relative to variables x_1, x_2, x_3 and x_4) were obtained by projection onto the residual space of reconstructed variables without variable x_1. As the first line and the first column of P_1^ℓ are null, according to (12), the residual $\tilde{x}_{11}$ is not sensitive to variables x_1, x_2 and x_3 and consequently to potential faults δx_1, δx_2 or δx_3 affecting these variables. Moreover, the residuals $\tilde{x}_{12}$, $\tilde{x}_{13}$ and $\tilde{x}_{14}$ are not sensitive to variable x_1 and thus to the fault δx_1 which can affect them. To summarize these different situations, the symbols $\times$ and 0 translate, or not, the fault influence on the residuals. The other parts of the table were constructed with this same principle by considering the different projection matrices $P_2^{(\ell)}$, $P_3^{(\ell)}$ and $P_4^{(\ell)}$. By analysing the dependence of the columns of the signature matrix, one can establish necessary conditions allowing the fault detection and isolation.

Let us note that only two projection matrices and two residuals are necessary for fault detection and isolation. For example, matrices $P_1^{(\ell)}$ and $P_2^{(\ell)}$ (11), allow to build the residuals $\tilde{x}_{12}$ (relative to x_2), $\tilde{x}_{21}$ (relative to x_1) which, permit to detect and isolate one of the three faults. Indeed, table (1) indicates that with these two residuals, the signature faults δx_1, δx_2 and δx_3 are respectively $(0 \quad \times)$, $(\times \quad 0)$ and $(\times \quad \times)$; these three signatures are independent and thus the faults are isolable from each other.

Table 1. Fault signatures

	$r = 1$				$r = 2$				$r = 3$				$r = 4$			
	$\tilde{x}_{11}$	$\tilde{x}_{12}$	$\tilde{x}_{13}$	$\tilde{x}_{14}$	$\tilde{x}_{21}$	$\tilde{x}_{22}$	$\tilde{x}_{23}$	$\tilde{x}_{24}$	$\tilde{x}_{31}$	$\tilde{x}_{32}$	$\tilde{x}_{33}$	$\tilde{x}_{34}$	$\tilde{x}_{41}$	$\tilde{x}_{42}$	$\tilde{x}_{43}$	$\tilde{x}_{44}$
δx_1	0	0	0	0	$\times$	0	$\times$	$\times$	$\times$	$\times$	0	$\times$	$\times$	$\times$	$\times$	0
δx_2	0	$\times$	$\times$	$\times$	0	0	0	0	$\times$	$\times$	0	$\times$	$\times$	$\times$	$\times$	0
δx_3	0	$\times$	$\times$	$\times$	$\times$	0	$\times$	$\times$	0	0	0	0	$\times$	$\times$	$\times$	0

Fault detection

Through the use of raw data, we established a robust PCA model by applying the propositions of section 3. The analysis of the decrease of the standardized eigenvalues of the covariance matrix (85.94, 13.99, 0.04, 0.03), allows to retain two principal components ($\ell = 2$). As indicated by remark 1, in the case of the residual generated without using variable x_1, only faults affecting the variables x_2 and x_3 are detectable on the residuals $\tilde{x}_{12}$, $\tilde{x}_{13}$ and $\tilde{x}_{14}$. Graphics of the figure 1 are relative to a global indicator Δ_R (norm of the projection vector weighted by the covariance matrix $V_R = P_R^{(\ell)} \Sigma P_R^{(\ell)T}$ of these projections) computed for each observation:

$$\Delta_R = \parallel \tilde{x}_R \parallel_{V_R^{-1}}^2 \tag{17}$$

A simple jump test on the Δ_R quantity (like Page-Hinkley for example) allows to determine the index of the fault observations. Detection and isolation are realised without ambiguity and are in accordance with the theoretical results of the isolation procedure (table 1). We can notice that a traditional non-robust PCA gives no significant result for fault detection.

4.2. Numerical example - multi-fault case

Data generation

The matrix X includes $N = 108$ observations of a vector x with 8 components generated
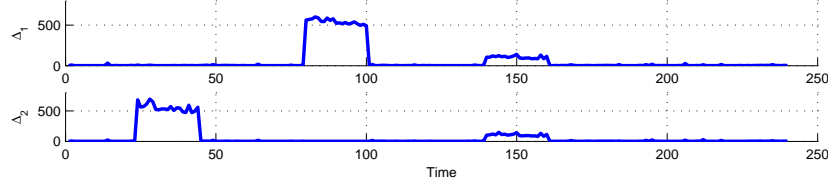
Fig. 1. Global indicator for x_1 and x_2

in the following way:

$$x_{i,1} = v_i^2 + \sin(0.1i), \quad x_{i,2} = 2\sin(i/6)\cos(i/4)\exp(-i/N), \quad v_i \sim \mathcal{N}(0,1) \quad (18)$$
$$x_{i,3} = \log(x_{i,2}^2), \quad x_{i,4} = x_{i,1} + x_{i,2}, \quad x_{i,5} = x_{i,1} - x_{i,2}$$
$$x_{i,6} = 2x_{i,1} + x_{i,2}, \quad x_{i,7} = x_{i,1} + x_{i,3}, \quad x_{i,8} \sim \mathcal{N}(0,1)$$

On the data thus generated were superimposed realizations of random variables with centered normal distribution and standard deviations equal to 0.02 as well as faults δx_1, δx_2, δx_3, δx_4 represented by a bias of amplitude equal to 3 and defined in the following way: observations from 10 to 24 for the variable x_1, observations from 35 to 49 for the variables x_2 and x_3, observations from 60 to 74 for the variables x_3 and x_4, observations from 85 to 99 for the variable x_4. In the following, these four intervals are indicated by I_1, I_2, I_3, I_4.

Sensitivity analysis

Concerning the a priori analysis of fault isolation, we limit ourselves to giving a reduced table of signatures (table 2) established from the properties (12). It reveals only some possible faults, noted δ in the first line, those affecting variables 1, 2, 3, 4 and those affecting the couples of variables $\{1,2\}, \{1,3\}, \{1,4\}, \{2,3\}, \{2,4\}, \{3,4\}$. The first column relates to the norm Δ_R of the residual vectors obtained by reconstruction-projection of the variables by using all the components of x except that with the indices R. The residuals are defined by (10). This table, which the reader will be able to extend , provides a correspondence be-

Table 2. Reduced table of faults signature

	Δ_1	Δ_2	Δ_3	Δ_4	Δ_5	Δ_6	Δ_{12}	Δ_{13}	Δ_{14}	Δ_{15}	Δ_{16}	Δ_{23}	Δ_{24}	Δ_{25}	Δ_{26}	Δ_{34}	Δ_{35}	Δ_{36}
δ_1	0	×	×	×	×	×	0	0	0	0	0	×	×	×	×	×	×	×
δ_2	×	0	×	×	×	×	0	×	×	×	×	0	0	0	0	×	×	×
δ_3	×	×	0	×	×	×	×	0	×	×	×	0	×	×	×	0	0	0
δ_4	×	×	×	0	×	×	×	×	0	×	×	×	0	×	×	0	×	×
δ_{12}	×	×	×	×	×	×	0	×	×	×	×	×	×	×	×	×	×	×
δ_{13}	×	×	×	×	×	×	×	0	×	×	×	×	×	×	×	×	×	×
δ_{14}	×	×	×	×	×	×	×	×	0	×	×	×	×	×	×	×	×	×
δ_{23}	×	×	×	×	×	×	×	×	×	×	×	0	×	×	×	×	×	×
δ_{24}	×	×	×	×	×	×	×	×	×	×	×	×	0	×	×	×	×	×
δ_{34}	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	0	×	×

tween the symptoms Δ_R and the faults δ_R. For example, the defect δ_2 affects all projections except those established without components 2, $\{1,2\}, \{2,3\}, \{2,4\}, \{2,5\}, \{2,6\}$.

Fault detection

From the contaminated data, the robust *PCA* model, with four principal axes, was chosen. Without carrying out the reconstruction, the observations were projected onto the residual space. The analysis of the residual norm thus generated by using all the variables reveals the presence of defects in the four intervals I_1, I_2, I_3, I_4, without being able to call into question a particular variable. This phase of detection is now supplemented by a stage of fault isolation.

The reconstruction is then carried out from all the variables except the variable 1, then starting from all the variables except the variables 1 and 2,... the last reconstruction being made from all the variables except the variables 7 and 8.

The figures 2 visualize the reconstructions of variables without using the variable 1. This figure shows the reconstruction of the first seven variables which are to be associated with the column Δ_1 of table 2 specifying the isolable faults.

The N reconstructed data were then projected onto the residual space. For each observation fault indicator Δ_R (17) were calculated.
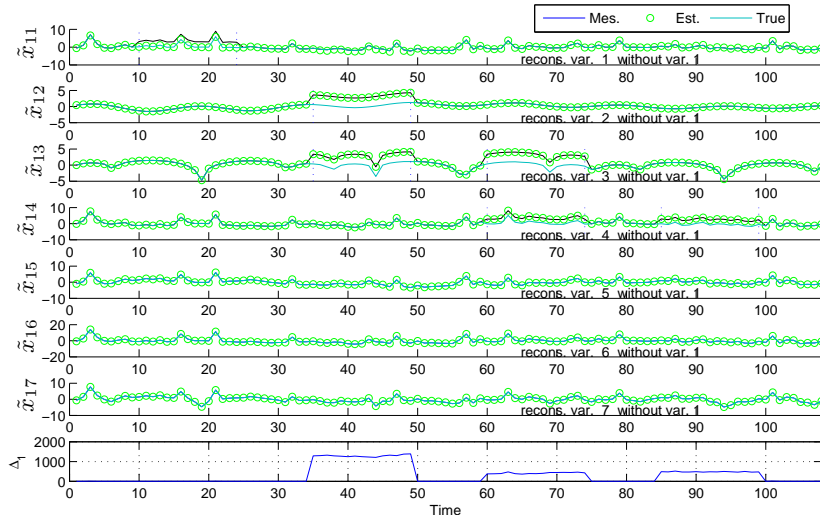

Fig. 2. Variables reconstruction without using variable 1 and global indicator

Let us analyse the figure 2. The variable 1, biased for the observations of the interval I_1, is not used for the reconstruction and the other variables which are used for the reconstruction do not present any bias. For these observations, the reconstructions are thus correct, emphasizing the first graph (starting from the top of the figure) which shows the superposition of the reconstructed variables (symbol 'o') with the true variables (in practice the latter are unknown, but at this stage where the data are generated, the comparison is possible). The measurement of the variable is also indicative (continuous line) in order to compare it with the reconstruction.

This result is confirmed by the last graph of the figure 2 where the norm of the vector projection (17) was traced. For the observations of the interval I_1 this norm is close to the value 0 thus testifying to the absence of outliers in the variables used for the reconstruction and projection, i.e. all the variables except x_1. Let us note that the three other groups of observations (I_2, I_3, I_4) are affected by faults, without knowing exactly which components of the measurement vector are faulty. Finally, by taking into account the fault presence in the four intervals, the examination of the figure 2 concludes that:

- in each interval I_2, I_3, I_4, a variable other than x_1 is faulty

Other projections (not presented here) are built and are interpreted in a similar way. The table 3 summarizes the conclusions resulting from the projection analysis. The line Δ_1 relates to the reconstructed residuals without using the first variable, the symbol 0 attests the fault absence in the considered interval. The diagnosis is then:

- in the interval I_1, x_1 is faulty
- in the interval I_2, x_2 and x_3 are faulty
- in the interval I_3, x_3 and x_4 are fautly
- in the interval I_4, x_4 is faulty

Table 3. Fault signatures

	I_1	I_2	I_3	I_4
Δ_1	0	×	×	×
Δ_{23}	×	0	×	×
Δ_{24}	×	×	×	0
Δ_{34}	×	×	0	0

5. Conclusion

Simulation results confirm that for data not contaminated by errors, classical *PCA* and robust *PCA* give similar results. In the other situations where outliers corrupt the data, traditional *PCA* proves to be ineffective, whereas its robust version gives completely satisfactory results. On the treated examples, the presence of approximately 25 percent of outliers authorizes a correct estimation of the principal directions, then the estimation is not very sensitive to these values. A *PCA* model can thus be built directly from the available data containing potential faults.

The most important result concerns the diagnosis of the systems, applied here to the detection and isolation of outliers. For that, we showed how to build fault indicators. The use of the principle of reconstruction and projection of the reconstructed data together made it possible to detect and isolate outliers in an effective way.

The procedure suggested here, is not limited in theory by the number of variables. However, the computational load is likely to become incompatible with an on-line treatment of the data and a reduction of the reconstruction and projection number is possible; this point will require a detailed attention in the follow-up of our work.

References

Böhning, D. and Ruangroj, R. (2002). A note on the maximum deviation of the scale-contaminated normal to the best distribution, *Metrika* **55**: 177–182.

Brown, C., Brcich, R. and Zoubir, A. (2005). Adaptive m-estimators for robust covariance estimation, *Applied Stochastic Models and Data Analysis*.

Croux, C. and Haesbroeck, G. (2000). Principal component analysis based on robust estimators of the covariance or correlation matrix: Influence functions and efficiencies, *Biometrika* **87**: 603–618.

Dunia, R. and Qin, S. (1998). A subspace approach to multidimensional fault identification and reconstruction, *AIChE Journal* **44 (8)**: 1813–1831.

Engelen, S., Hubert, M. and Vanden Branden, K. (2005). A comparison of three procedures for robust pca in high dimensions, *Austrian Journal of Statistics* **34 (2)**: 117–126.

Fekri, M. and Ruiz-Gazen, A. (2003). Robust weighted orthogonal regression in the errors-in-variables model, *Journal of Multivariate Analysis* **88**: 89–108.

Hubert, M., Rousseeuw, P. and Van den Branden, K. (2005). Robpca: a new approach to robust principal component analysis, *Technometrics* **47 (1)**: 64–79.

Salibian-Barrera, M., Van Aelst, S. and Willems, G. (2006). Pca based on multivariate mm-estimators with fast and robust bootstrap, *Journal of the American Statistical Association* **101**: 1198 – 1211.