Fault detection and isolation with robust principal component analysis

Y. Tharrault, G. Mourot, J. Ragot

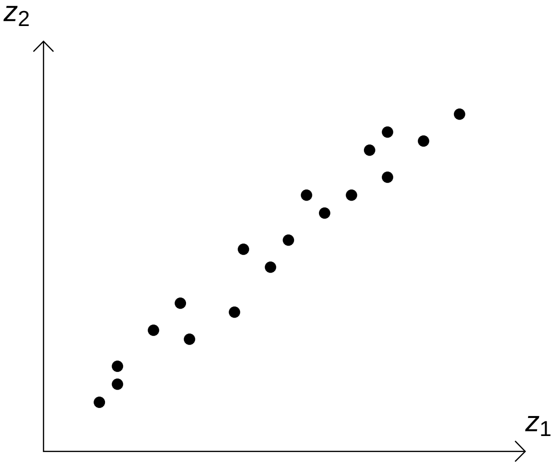
Centre de Recherche en Automatique de Nancy
Nancy-université, CNRS
2, Avenue de la forêt de Haye
54516 Vandœuvre-lès-Nancy Cedex

16th Mediterranean Conference on Control and Automation
June 25-27, 2008, Ajaccio, Corsica, France

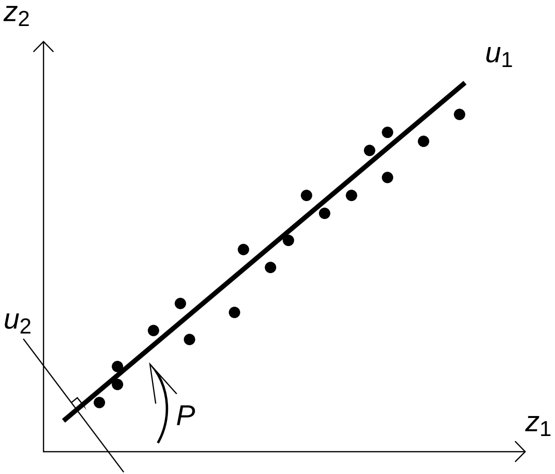
CRAN

Outline

1. Principle of the Principal component analysis
2. Robust principal component analysis
3. Multiple faults detection and isolation
4. Numerical example: multi-fault case


z_2
z_1

Principle of the Principal component analysis



Principle of the Principal component analysis

- Data matrix $X \in \Re^{N \times n}$ in a normal process operation

PCA

Maximization of the variance projections $T = X\,P$

- $T \in \Re^{N \times n}$: principal component matrix
- $P \in \Re^{n \times n}$: projection matrix

Decomposition in eigenvalues/eigenvectors of the covariance matrix

$$\Sigma = \frac{1}{N-1} X^T X = P \Lambda P^T \qquad \text{with} \qquad P P^T = P^T P = I_n$$

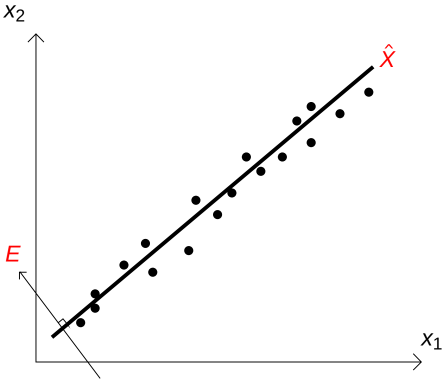
$$\Sigma = \begin{bmatrix} P_\ell & P_{n-\ell} \end{bmatrix} \begin{bmatrix} \Lambda_\ell & 0 \\ 0 & \Lambda_{n-\ell} \end{bmatrix} \begin{bmatrix} P_\ell^T \\ P_{n-\ell}^T \end{bmatrix}$$

Principle of the Principal component analysis

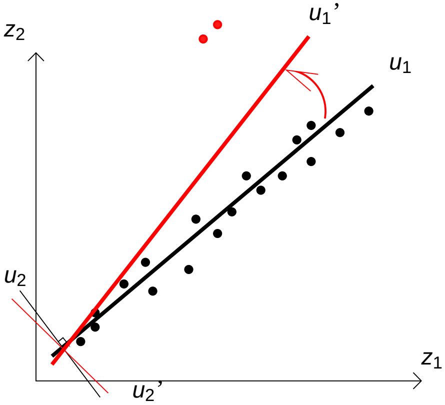
Decomposition

Principal part : $\hat{X} = X\,P_\ell P_\ell^T = X\,C_\ell$

Residual part : $E = X - \hat{X} = X\,(I - C_\ell)$



- Sensitive to outliers



Robust PCA with respect to outliers

⟶ Outliers detection and isolation

Robust approach

A scale M-estimator with a robust covariance matrix as initialization

scale M-estimator

The residual r_i

$$r_i = ||P_{n-\ell}{}^T (x_i - \mu)||^2$$

with x_i an observation
μ the mean of the data X
$P_{n-\ell}$ is the eigenvector matrix of the robust covariance matrix corresponding to its $n-\ell$ smallest eigenvalues

Robust approach

A scale M-estimator with a robust covariance matrix as initialization

scale M-estimator

The residual r_i

$$r_i = ||{P_{n-\ell}}^T (x_i - \mu)||^2$$

The general scale-M estimator minimizes the following objective function with the constraint $P^T P = I$:

$$\frac{1}{N} \sum_{i=1}^{N} \rho \left(\frac{r_i}{\hat{\sigma}} \right)$$

with $\hat{\sigma}$ the robust scale of the residual r_i and
the function ρ choose as the bisquare function ($\rho(r) = \min\{1, 1 - (1 - r)^3\}$).

Robust approach

Robust covariance matrix use to initialize a scale M-estimator

Robust covariance matrix

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

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

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

with β turning parameter

Reconstruction principle

Estimate r variables using the $n - r$ remaining variables and the model

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

with Ξ_R : reconstruction directions matrix

To reconstruct 2 variables ($R = 2, 4$) among 5 variables

$$\Xi_R = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}^T$$

Reconstruction condition : $(\tilde{\Xi}_R^T \tilde{\Xi}_R)^{-1}$

$$r <= n - \ell$$

If the matrix of the reconstruction directions is reorganized as follows:

$$\Xi_R = \left[\underset{(r\times r)}{I_1} \;\middle|\; \underset{((n-r)\times r)}{0} \right]^T \in \Re^{n\times r}$$

with $I_1 \in \Re^{r\times r}$ an identity matrix.

Then C_ℓ is splitted in four parts: $C_\ell = \left[\begin{array}{c|c} \underset{(r\times r)}{c_1} & \underset{(r\times(n-r))}{c_2} \\ \hline \underset{((n-r)\times r)}{c_2^T} & \underset{((n-r)\times(n-r))}{c_4} \end{array} \right] \in \Re^{n\times n}$

The reconstruction $\hat{x}_R$ of the vector x is written as follows:

$$\hat{x}_R = \left[\begin{array}{c|c} 0 & (I_1 - c_1)^{-1} c_2 \\ \hline 0 & I_2 \end{array} \right] x$$

with $I_2 \in \Re^{n-r\times n-r}$ an identity matrix.

x_R projection in the residual space

$$\tilde{x}_R = (I - C_\ell)\,\hat{x}_R$$

$$\tilde{x}_R = P_R^{(\ell)}\,x, \qquad P_R^{(\ell)} = (I - C_\ell)\,[I - \Xi_R(\tilde{\Xi}_R^T\tilde{\Xi}_R)^{-1}\tilde{\Xi}_R^T)]$$

Fault detection and isolation

A global indicator SPE_R (norm of the projection vector)

$$SPE_R = \| \tilde{x}_R \|^2$$

Fault detection

A fault is detected, if:

$$SPE > \delta_\alpha^2$$

δ_α^2 the detection threshold of SPE,
SPE correspond to the case without reconstruction *i.e.* $R = \{\phi\}$.

Fault isolation

The faulty variables of the subset $\hat{R}$ are determined as follows:

$$\hat{R} = \underset{R \in \Im}{\arg} SPE_R < \delta_\alpha^2$$

with $\Im$ is the set of all combinations of possible reconstruction directions.

Reconstruction condition : $r <= n - \ell$

The maximum reconstruction number can be calculated as follows:

$$\sum_{r=1}^{n-\ell} \mathbb{C}_n^r$$

with $\mathbb{C}_n^r$ denotes the combination of r from n.

Reduction of the number reconstruction

It can be reduced when the matrix of projected fault directions is rank-deficient or near rank-deficient.

To detect these cases, the condition number (*Rcond*) is used:

$$Rcond = \frac{\min\left(\sigma\left(\tilde{\Xi}_R\right)\right)}{\max\left(\sigma\left(\tilde{\Xi}_R\right)\right)}$$

Algorithm to determine the detectable faults and the isolable faults

1. $r = 1$ (single-fault): calculate all available directions $\tilde{\Xi}_R$.
   If $\tilde{\Xi}_R^T \tilde{\Xi}_R$ is closed to zero –> fault is not detectable
2. $r = r + 1$: calculate for all available directions $\tilde{\Xi}_R$ the values of the condition number *Rcond*.
   If *Rcond* is close to zero :
   - the r faulty variables of the subset R are not detectable.
   - all the combinations of $r - 1$ variables are only detectable.
   - all the combinations of $r - 2$ variables are isolable.

   If *Rcond* is not close to zero :
   - all the combinations of $r - 1$ variables are isolable.
3. While $r \leq n - \ell$ go to step 2

Numerical example: multi-fault case

One considers:

- $X = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 \end{bmatrix}$
- $N = 108$ observations

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

	I_1 {10 : 24}	I_2 {35 : 49}	I_3 {60 : 74}	I_4 {85 : 99}
x_1	$\times$	0	0	$\times$
x_2	0	$\times$	0	0
x_3	0	$\times$	0	0
x_4	0	0	$\times$	$\times$
x_5	0	0	$\times$	0
x_6	0	0	0	$\times$

Numerical example: Useful reconstruction

1. for $r = 1$

$\tilde{\Xi}_1^T \tilde{\Xi}_1$	$\tilde{\Xi}_2^T \tilde{\Xi}_2$	$\tilde{\Xi}_3^T \tilde{\Xi}_3$	$\tilde{\Xi}_4^T \tilde{\Xi}_4$	$\tilde{\Xi}_5^T \tilde{\Xi}_5$	$\tilde{\Xi}_6^T \tilde{\Xi}_6$	$\tilde{\Xi}_7^T \tilde{\Xi}_7$	$\tilde{\Xi}_8^T \tilde{\Xi}_8$
0.84	0.72	0.46	0.71	0.41	0.40	0.46	**0.00**

–> Fault on x_8 is not detectable in the residual space

2. for $r = 2$

Rcond	1	2	3	4	5	6	7
1		0.88	0.72	0.88	0.57	0.57	0.72
2			0.79	0.73	0.42	0.68	0.80
3				0.80	0.75	0.76	**0.01**
4					0.68	0.41	0.80
5						0.79	0.75
6							0.75

–> Fault signatures on x_3 and x_7 are identical

–> Fault on x_3 or x_7 is not isolable

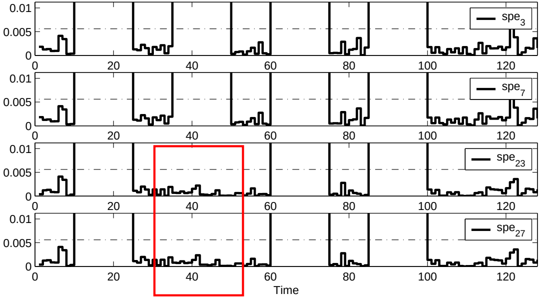
Numerical example: Useful reconstruction

1. for $r = 1$

$\tilde{\Xi}_1^T \tilde{\Xi}_1$	$\tilde{\Xi}_2^T \tilde{\Xi}_2$	$\tilde{\Xi}_3^T \tilde{\Xi}_3$	$\tilde{\Xi}_4^T \tilde{\Xi}_4$	$\tilde{\Xi}_5^T \tilde{\Xi}_5$	$\tilde{\Xi}_6^T \tilde{\Xi}_6$	$\tilde{\Xi}_7^T \tilde{\Xi}_7$	$\tilde{\Xi}_8^T \tilde{\Xi}_8$
0.84	0.72	0.46	0.71	0.41	0.40	0.46	**0.00**

–> Fault on x_8 is not detectable in the residual space

2. for $r = 2$



–> Fault signatures on x_3 and x_7 are identical

–> Fault on x_3 or x_7 is not isolable

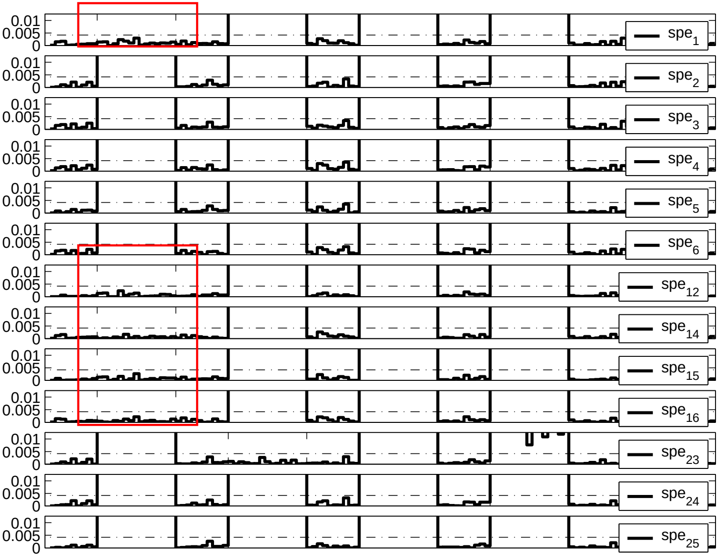
For all the directions of reconstruction *Rcond* is calculated

Rcond is close to zero with $R = \{2, 4, 5, 6\}$

Useful reconstruction

- From the 162 reconstruction possibilities
- 91 are really reconstructible
- 21 combinations are useful to isolate the faulty variables
- For the others, a set of variable is considered as faulty

Numerical example: Useful reconstruction



Numerical example: multi-fault case

	I_1	I_2	I_3	I_4
Δ_1	0	$\times$	$\times$	$\times$
Δ_{23}	$\times$	0	$\times$	$\times$
Δ_{45}	$\times$	$\times$	0	$\times$
Δ_{146}	0	$\times$	$\times$	0

Fault signatures

- in the interval I_1, x_1 is faulty
- in the interval I_2, x_2 and x_3 or/and x_7 are faulty
- in the interval I_3, x_4 and x_5 are faulty
- in the interval I_4, x_1, x_4 and x_6 are faulty

Conclusion

- Robust PCA with respect to outliers
  –> directly applicable on data containing potential faults
- use of the principle of reconstruction and projection of the reconstructed data together
  –> outliers detection and isolation
- Reduction of the computational load
  –> determination of the dectable and isolable faults