

Détection et isolation de défauts par analyse en composantes principales robuste

Y. Tharrault, G. Mouro, J. Ragot

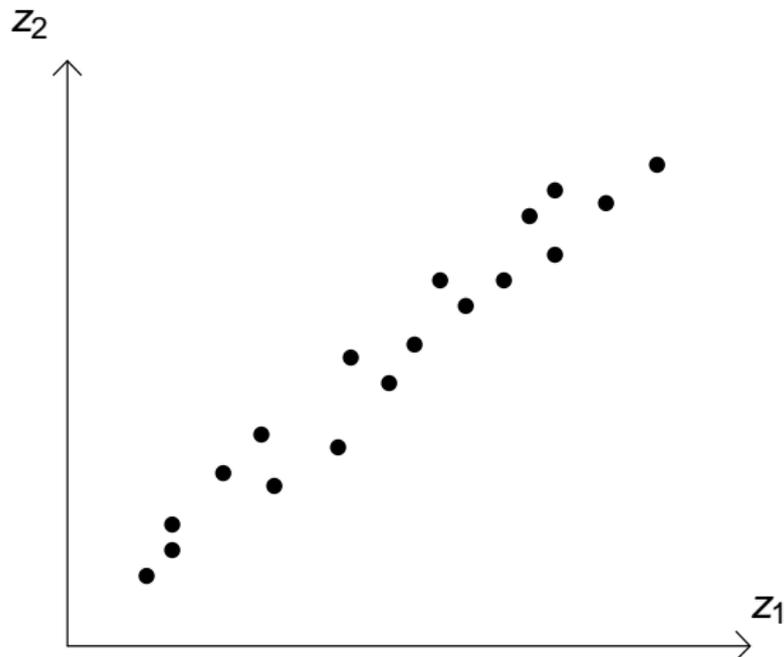
Centre de Recherche en Automatique de Nancy
Nancy-université, CNRS
2, Avenue de la forêt de Haye
54516 Vandœuvre-lès-Nancy Cedex

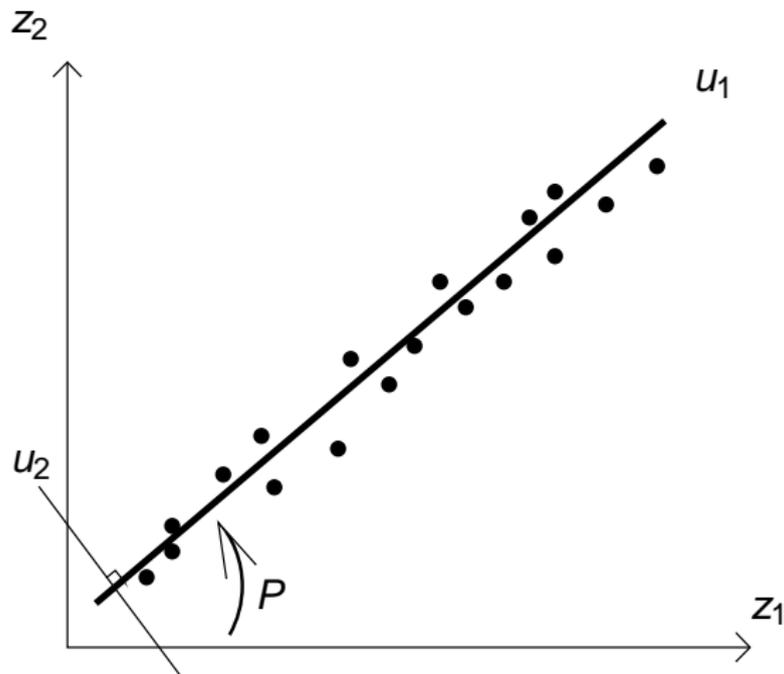
Conférence Internationale Francophone d'Automatique
3-5 septembre 2008, Bucarest, Roumanie



- 1 Principe de l'Analyse en Composantes Principales (ACP)
- 2 Limites de l'Analyse en Composantes Principales
- 3 Analyse en Composantes Principales Robuste (ACPR)
- 4 Détection et localisation de défauts multiples de capteurs
- 5 Exemple numérique : cas de défauts multiples

Principe de l'Analyse en Composantes Principales





- Matrice de données $X \in \mathfrak{R}^{N \times n}$ en fonctionnement normal

ACP

Maximisation de la variance des projections $T = XP$

- $T \in \mathfrak{R}^{N \times n}$: Matrice des composantes principales
- $P \in \mathfrak{R}^{n \times n}$: Matrice de projection

Décomposition en valeurs/vecteurs propres de la matrice de covariance

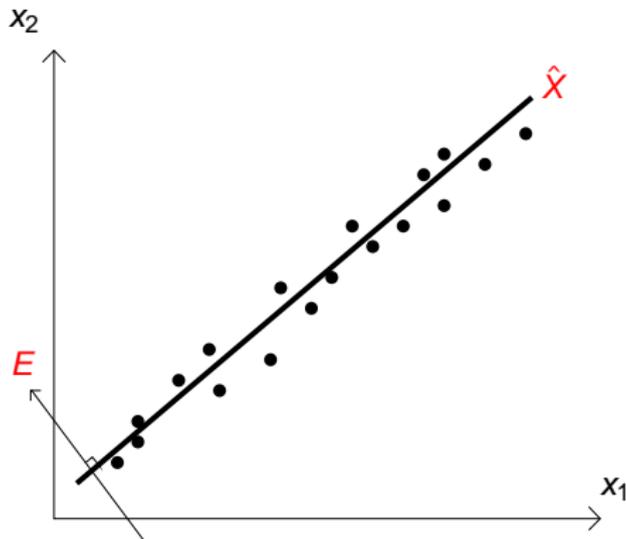
$$\Sigma = \frac{1}{N-1} X^T X = P \Lambda P^T \quad \text{avec} \quad P P^T = P^T P = I_n$$

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

Decomposition

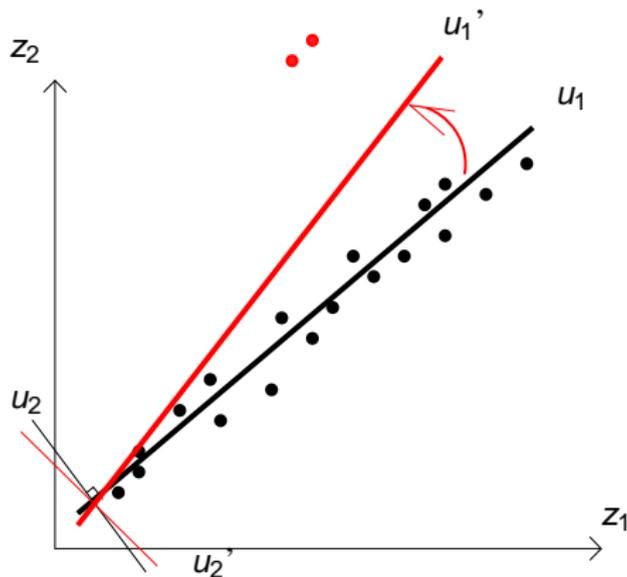
Partie principale : $\hat{X} = X P_\ell P_\ell^T = X C_\ell$

Partie résiduelle : $E = X - \hat{X} = X(I - C_\ell)$



Limites de l'Analyse en Composantes Principales

- Sensible aux valeurs aberrantes



Modèle ACP robuste aux valeurs aberrantes

- détection, localisation et suppression des valeurs aberrantes, puis modèle ACP
- estimation robuste du modèle ACP sans élimination a priori des valeurs aberrantes

Analyse en Composantes Principales Robuste (ACPR)

Approche robuste

Un MM-estimateur initialisé par une matrice de variance-covariance robuste

MM-estimateur

Les résidus $r(k)$

$$r(i) = \|P_{n-\ell}^T(x(i) - \mu)\|^2$$

avec :

- $x(i)$ une observation
- μ la moyenne robuste du jeu de données X
- $P_{n-\ell}$ la matrice des vecteurs propres de la matrice de variance-covariance robuste associés aux $n - \ell$ plus petites valeurs propres

Approche robuste

Un MM-estimateur avec pour l'initialisé une matrice de variance-covariance robuste

MM-estimateur

Les résidus $r(k)$

$$r(i) = \|P_{n-\ell}^T (x(i) - \mu)\|^2$$

Le MM-estimateur minimise le critère suivant (sous contrainte $P^T P = I$) :

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

avec :

- $\hat{\sigma}$ la dispersion robuste des résidus r
- ρ est choisie comme la fonction "Bisquare" ($\rho(r) = \min\{1, 1 - (1 - r)^2\}$).

Approche robuste

Un MM-estimateur avec initialisé par une matrice de variance-covariance robuste

Matrice de variance-covariance robuste

$$V = \frac{\sum_{i=1}^{N-1} \sum_{j=i+1}^N w(i,j) (\mathbf{x}(i) - \mathbf{x}(j)) (\mathbf{x}(i) - \mathbf{x}(j))^T}{\sum_{i=1}^{N-1} \sum_{j=i+1}^N w(i,j)}$$

où les poids $w(i,j)$ sont définis par :

$$w(i,j) = \exp \left(-\frac{\beta}{2} (\mathbf{x}(i) - \mathbf{x}(j))^T \Sigma^{-1} (\mathbf{x}(i) - \mathbf{x}(j)) \right)$$

avec β un paramètre de réglage

Détection et localisation de défauts multiples de capteurs

Indicateurs de détection généralement utilisés

SPE pour détecter un défaut dans l'espace résiduel

$$SPE(k) = \|(I - P_\ell P_\ell^T)x(k)\|^2 \implies SPE(k) \leq \delta^2 ?$$

T^2 pour détecter un défaut dans l'espace principal

$$T^2(k) = x(k)^T P_\ell \Lambda_\ell^{-1} P_\ell^T x(k) \implies T_2(k) \leq \chi_l^2 ?$$

Afin d'assurer la détection de l'ensemble des défauts

Utilisation d'un indicateur combiné

$$\eta(k) = \frac{T^2(k)}{\chi_l^2} + \frac{SPE(k)}{\delta^2} = x(k)^T \Phi x(k)$$

avec $\Phi = \frac{P_\ell \Lambda_\ell^{-1} P_\ell^T}{\chi_l^2} + \frac{I - P_\ell P_\ell^T}{\delta^2}$

Reconstruction \hat{x}_R de l'état du système

Minimisant l'influence des défauts à l'indicateur de détection

$$\hat{x}_R(k) = x(k) - \Xi_R f_R$$

avec f_R : amplitude (inconnue) du défaut

Ξ_R : matrice des directions de reconstruction

Par exemple, pour reconstruire 2 variables ($R = 2, 4$) parmi 5 variables

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

Estimation de l'amplitude du défaut \hat{f}_R

$$\hat{f}_R = \underset{f_R}{\operatorname{arg\,min}} \{ \eta_R(k) \} = \underset{f_R}{\operatorname{arg\,min}} \{ \hat{x}_R(k)^T \Phi \hat{x}_R(k) \}$$

Le vecteur de reconstruction $\hat{x}_R(k)$ du vecteur $x(k)$ est donné par :

$$\hat{x}_R(k) = (I - \Xi_R(\Xi_R^T \Phi \Xi_R)^{-1} \Xi_R^T \Phi) x(k)$$

avec $\Phi = \frac{P_\ell \Lambda_\ell^{-1} P_\ell^T}{\chi_l^2} + \frac{I - P_\ell P_\ell^T}{\delta^2}$

Condition de reconstruction : $(\Xi_R^T \Phi \Xi_R)^{-1}$

Pour qu'un défaut soit reconstructible, il faut qu'il soit au minimum projeté dans l'espace principal ($r \leq \ell$) ou dans l'espace résiduel ($r \leq n - \ell$). Cette condition implique que le nombre de variables reconstruites r doit respecter l'inégalité :

$$r \leq \max(n - \ell, \ell)$$

Indice de détection η et de localisation η_R

$$\eta(k) = x(k)^T \Phi x(k) \quad \eta_R(k) = x_R(k)^T \Phi x_R(k) \quad \Phi = \frac{P_l \Lambda_l^{-1} P_l^T}{\chi_l^2} + \frac{I - P_l P_l^T}{\delta^2}$$

Détection de défauts

Un défaut est détecté, si :

$$\eta(k) > \gamma_\alpha^2$$

γ_α^2 le seuil de détection de l'indicateur η ,

Localisation de défauts

Les jeux de variables défaillantes \hat{R} sont déterminés par :

$$\hat{R} = \arg \min_{R \in \mathfrak{S}} \eta_R(k) < \gamma_\alpha^2$$

où \mathfrak{S} est l'ensemble des combinaisons des directions de reconstructions possibles.

Condition de reconstruction : $r \leq \max(n-l, l)$

Le nombre maximum de reconstructions est le suivant :

$$\sum_{r=1}^{\max(n-l, l)} \mathbb{C}_n^r$$

où \mathbb{C}_n^r est le nombre de combinaisons possibles de choisir r variables parmi n .

Réduction du nombre de reconstruction

Détermination des directions de projections colinéaires.

Indicateur de similarité de directions de défauts

$$D(R_1, R_2) = \max\{d(R_1, R_2), \tilde{d}(R_1, R_2)\}$$

où $d(R_1, R_2)$ distance entre deux sous espaces dans l'espace principal
et $\tilde{d}(R_1, R_2)$ distance entre deux sous espaces dans l'espace résiduel :

$$d(R_1, R_2) = \|\hat{\Xi}_{R_1} (\hat{\Xi}_{R_1}^T \hat{\Xi}_{R_1})^{-1} \hat{\Xi}_{R_1}^T - \hat{\Xi}_{R_2} (\hat{\Xi}_{R_2}^T \hat{\Xi}_{R_2})^{-1} \hat{\Xi}_{R_2}^T\|_2$$

$$\tilde{d}(R_1, R_2) = \|\tilde{\Xi}_{R_1} (\tilde{\Xi}_{R_1}^T \tilde{\Xi}_{R_1})^{-1} \tilde{\Xi}_{R_1}^T - \tilde{\Xi}_{R_2} (\tilde{\Xi}_{R_2}^T \tilde{\Xi}_{R_2})^{-1} \tilde{\Xi}_{R_2}^T\|_2$$

avec $\hat{\Xi}_{R_1} = \Lambda_\ell^{-1/2} P_\ell^T \Xi_{R_1}$, $\tilde{\Xi}_{R_1} = P_{n-\ell}^T \Xi_{R_1}$ et R_1 et R_2 correspondent aux ensembles des variables de reconstruction.

- 1 $r = 1$
- 2 Calculer pour l'ensemble des directions possibles l'indicateur $D(R_1, R_2)$.
 - Si $D(R_1, R_2)$ est nul :
seul l'ensemble des variables potentiellement en défaut peut être déterminé, c'est-à-dire les variables avec les indices R_1, R_2 ou R_1 et R_2 . Il est donc nécessaire de considérer une seule direction, par exemple R_1 .
 - Si $D(R_1, R_2)$ est proche de zéro :
l'amplitude du défaut doit être importante pour assurer la localisation du défaut
 - Si $D(R_1, R_2)$ est proche de 1 :
le défaut est facilement localisable
- 3 $r = r + 1$
- 4 Tant que $r \leq \max(\ell, n - \ell)$ aller à l'étape 2

Exemple numérique : cas de défauts multiples

On considère :

- $X = [x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5 \quad x_6 \quad x_7 \quad x_8 \quad x_9]$
- $N = 128$ observations

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

$$x_{i,2} = 2 \sin(i/6) \cos(i/4) \exp(-i/N)$$

$$x_{i,3} = \log(x_{i,2}^2)$$

$$x_{i,4} = x_{i,1} + x_{i,2}$$

$$x_{i,5} = x_{i,1} - x_{i,2}$$

$$x_{i,6} = 2x_{i,1} + x_{i,2}$$

$$x_{i,7} = x_{i,1} + x_{i,3}$$

$$x_{i,8} \sim \mathcal{N}(0,1)$$

$$x_{i,9} \sim \mathcal{N}(0,1)$$

	I_1 {10 : 24}	I_2 {35 : 49}	I_3 {60 : 74}	I_4 {85 : 99}
x_1	×	×	0	0
x_2	0	×	0	0
x_3	×	0	0	0
x_4	0	0	×	0
x_8	0	0	×	×

Exemple numérique : reconstruction utile

$$\max(\ell, n - \ell) = \max(5, 4) = 5$$

Le nombre maximum de reconstructions est donc de 381 ($\mathbb{C}_9^1 + \mathbb{C}_9^2 + \dots + \mathbb{C}_9^5$).

$r = 1$

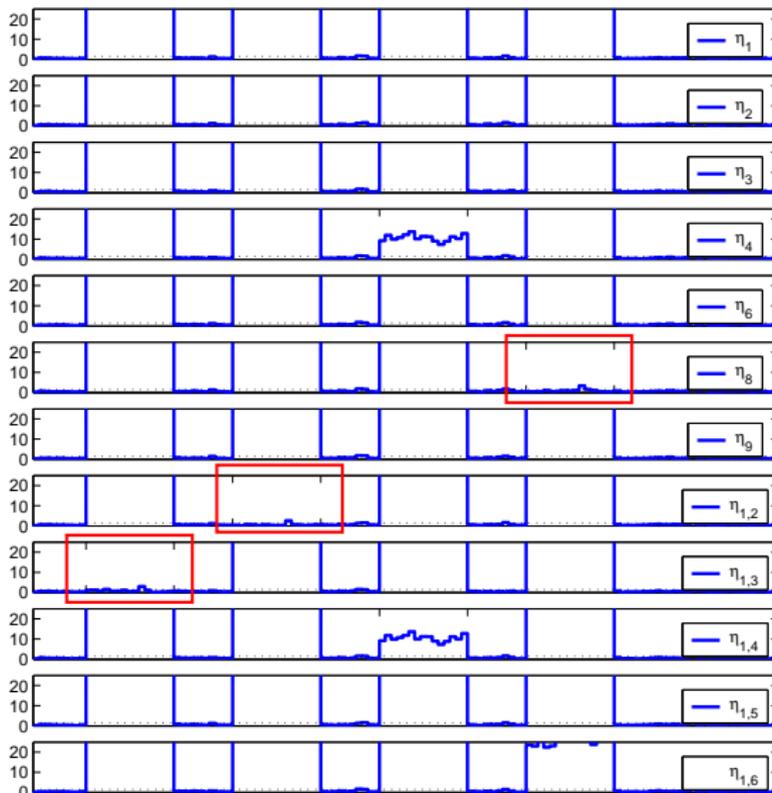
k	R_1								
	1	2	3	4	5	6	7	8	9
1	0	1	0.99	1.00	0.92	0.92	1.00	1.00	0.99
2		0	1.00	0.96	0.73	0.97	1.00	1.00	1.00
3			0	1.00	0.96	1.00	0.40	1.00	1.00
R_2 4				0	0.97	0.73	1.00	0.99	1.00
5					0	0.97	0.99	1.00	1.00
6						0	1.00	0.99	1.00
7							0	1.00	0.99
8								0	1.00

$r = 2$

Les directions $\{1, 3\}$ et $\{1, 7\}$ sont colinéaires

Le nombre de reconstructions nécessaires est de 150 au lieu de 381

Exemple numérique : reconstruction utile



Défaut x_1 sur l_1 et l_2

Défaut x_2 sur l_2

Défaut x_3 sur l_1

Défaut x_4 sur l_3

Défaut x_8 sur l_3 et l_4

Sur l_1 , défauts x_1 et x_3

Sur l_2 , défauts x_1 et x_2

Sur l_3 , défauts x_4 et x_8

Sur l_4 , défauts x_8

	l_1	l_2	l_3	l_4
$\eta_{1,3}$	0	×	×	×
$\eta_{1,2}$	×	0	×	×
$\eta_{4,8}$	×	×	0	0
η_8	×	×	×	0

Signature des défauts

- dans l'intervalle l_1 , x_1 et x_3 ou / x_1 et x_7 sont en défaut
- dans l'intervalle l_2 , x_1 et x_2 sont en défaut
- dans l'intervalle l_3 , x_4 et x_8 sont en défaut
- dans l'intervalle l_4 , x_8 est en défaut

Conclusion et perspectives

Conclusion et perspectives

- ACP robuste aux valeurs aberrantes
- Utilisation d'un indice combiné pour détecter des défauts sur l'ensemble des variables
- Utilisation du principe de reconstruction afin de localiser les défauts
- Détermination des défauts isolables
→ évite l'explosion combinatoire des scénarios de défauts

Perspectives

- Systèmes dynamiques
- Prise en compte de la persistance des défauts