
Détection et isolation de défauts par analyse en composantes principales robuste

Yvon Tharrault* — Gilles Mourot* — José Ragot* — Mohamed-Faouzi Harkat**

*Centre de Recherche en Automatique de Nancy.

UMR 7039 Nancy-Université, CNRS.

2, Avenue de la forêt de Haye. F - 54 516 Vandoeuvre les Nancy Cedex.

{yvon.tharrault, gilles.mourot, jose.ragot}@ensem.inpl-nancy.fr

**Université Badji Mokhtar Annaba

Faculté des sciences de l'ingénieur, Département d'électronique

BP. 12, Sidi Amar, 23000 Annaba, Algérie

mharkat@univ-annaba.org

RÉSUMÉ. L'Analyse en Composantes Principales (ACP) est largement utilisée pour le diagnostic de fonctionnement de système. Cependant, l'un des inconvénients majeurs de l'ACP résulte de l'utilisation de techniques d'estimation par moindres carrés, techniques qui échouent souvent à s'affranchir des biais de mesure accidentels. Ce papier présente une procédure de détection et de localisation de défauts de mesure. La méthode proposée ne nécessite pas d'étude préliminaire relative à la détection et au rejet de valeurs aberrantes dans les données utilisées pour la conception du modèle ACP. Le modèle ACP obtenu étant robuste, c'est-à-dire non contaminé par les valeurs aberrantes, son utilisation pour le diagnostic est alors efficace.

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.

MOTS-CLÉS : Valeurs aberrantes, analyse en composantes principales, robustesse, signature de défauts, détection de défauts, localisation de défauts.

KEYWORDS: Outliers, principal component analysis, robustness, fault signature, fault detection, fault isolation.

1. Introduction

L'analyse en composantes principales est une technique numérique bien éprouvée dans le domaine du traitement de données pour réduire la dimension de l'espace de représentation d'un système (Jolliffe, 1986).

L'ACP est essentiellement basée sur la mise en évidence de relations linéaires entre les variables et présente un caractère d'optimalité au sens d'un critère portant sur l'erreur quadratique d'estimation en valeur moyenne (*MSE*). Il est bien connu que l'estimation basée sur l'utilisation de critères de type *MSE* est moins robuste aux valeurs aberrantes que celle issue d'autres critères comme celui de l'erreur en valeur absolue (Hubert *et al.*, 2005).

Pour tolérer la présence de valeurs aberrantes, une analyse en composantes principales robuste peut être conduite en calculant une matrice de covariance des données robuste, cette matrice étant un point clef pour la recherche du modèle *ACP*. Dans (Fekri *et al.*, 2003), les auteurs définissent une matrice de variances et covariances "locale" en ce sens que la forme proposée tend à privilégier la contribution d'observations proches au détriment d'observations éloignées dues à la présence de valeurs aberrantes.

Notre présentation est consacrée au problème de détection et de localisation de défauts dans des données qui résultent, en général, de dysfonctionnements du procédé ou du système d'acquisition de mesure. La contribution porte essentiellement sur la détection de valeurs aberrantes et leur localisation en utilisant des outils complémentaires : l'analyse en composante principale robuste, la reconstruction de données et l'analyse de résidus. La section 2 est un bref rappel, d'une part, de l'analyse en composantes principales dans le cas classique et, d'autre part, de l'analyse en composantes principales robuste aux valeurs aberrantes. Une procédure de détection et de localisation des valeurs aberrantes est ensuite proposée en section 3, puis, en section 4, est appliquée à un exemple où plusieurs variables sont simultanément en défaut.

2. Principe de l'analyse en composantes principales

Soit une matrice de données $X \in \mathbb{R}^{N \times n}$, de vecteurs lignes x_i^T , qui rassemble les N mesures effectuées sur les n variables du système.

2.1. Approche classique

Dans le cas de l'ACP classique, les données sont supposées être recueillies sur un système ayant un fonctionnement normal (absence de défauts).

L'ACP détermine une transformation optimale (vis-à-vis d'un critère de variance) de la matrice de données X :

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

$T \in \mathbb{R}^{N \times n}$ est la matrice des composantes principales et $P = [p_1 \ p_2 \ \dots \ p_n] \in \mathbb{R}^{n \times n}$, où les vecteurs orthogonaux p_i sont les vecteurs propres correspondant aux valeurs propres λ_i de la décomposition en valeurs et vecteurs propres de la matrice de covariance (ou de corrélation) Σ de X :

$$\Sigma = P \Lambda P^T \quad \text{avec} \quad PP^T = P^T P = I_n \quad (2)$$

avec $\Lambda = \text{diag}(\lambda_1 \dots \lambda_n)$ une matrice diagonale où les termes diagonaux sont ordonnés dans l'ordre décroissant.

Il est bien connu que l'analyse des valeurs propres de la matrice Σ de covariance renseigne sur le nombre de composantes principales à retenir (Jolliffe, 1986). Une fois déterminé le nombre ℓ de composantes à retenir, la matrice X des données peut être approximée. Pour cela la matrice des vecteurs propres est partitionnée sous la forme :

$$P = (\hat{P} \quad \tilde{P}) \quad \hat{P} \in \mathbb{R}^{n \times \ell} \quad (3)$$

A partir de la décomposition (1), on peut alors expliciter la partie principale \hat{X} des données expliquées par les ℓ premiers vecteurs propres et la partie résiduelle \tilde{X} expliquée par les composantes restantes :

$$\hat{X} = X \hat{P} \hat{P}^T = X C_\ell \quad (4)$$

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

2.2. Approche robuste

Une difficulté majeure de l'ACP provient de sa sensibilité aux valeurs aberrantes. Dans (Fekri *et al.*, 2003) les auteurs définissent une matrice de covariance "locale" qui tend à privilégier la contribution d'observations proches au détriment d'observations éloignées dues à la présence de valeurs aberrantes. Cette matrice est définie par :

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

où les poids $w_{i,j}$ sont eux-mêmes définis par :

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

β étant un paramètre à régler pour obtenir effectivement une réduction de l'influence des observations éloignées, les auteurs préconisant une valeur voisine de 2. L'ACP peut alors être conduite sur cette "nouvelle" matrice de covariance réputée robuste vis-à-vis des valeurs aberrantes grâce à la présence de poids adaptés $w_{i,j}$.

3. Détection de défauts

L'ACP peut être utilisée pour la détection de défauts, les défauts se traduisant par des données aberrantes que l'on peut mettre en évidence par projection dans certains espaces. Tout d'abord, on rappelle comment reconstruire un ensemble de r variables parmi n à partir des $n - r$ variables restantes. Puis, on montre l'utilisation de cette reconstruction pour la détection de données aberrantes.

3.1. Reconstruction de données

En considérant les résultats rappelés à la section 2, on peut écrire la décomposition suivante :

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

dans laquelle \hat{x} et \tilde{x} sont respectivement la projection de x sur les espaces engendrés par les ℓ composantes principales et les $n - \ell$ composantes restantes (espace résiduel).

Soit R l'ensemble des indices des variables reconstruites, le vecteur \hat{x}_R correspondant à la reconstruction de ces r variables est donné par (Dunia *et al.*, 1998) :

$$\hat{x}_R = G_R x \quad (9a)$$

$$G_R = [I - \Xi_R (\tilde{\Xi}_R^T \tilde{\Xi}_R)^{-1} \tilde{\Xi}_R^T], \quad \tilde{\Xi}_R = (I - C_\ell) \Xi_R \quad (9b)$$

où la matrice $\Xi_R \in \mathfrak{R}^{n \times r}$ est la matrice des directions de reconstruction constituée de 0 pour indiquer une variable non reconstruite (respectivement, de 1 pour une variable reconstruite). Par exemple, prenons le cas où on souhaite reconstruire les variables $R = \{2, 4\}$ parmi 5 variables, la matrice Ξ_R est de la forme :

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

On remarque que dans le cas où $(\tilde{\Xi}_R^T \tilde{\Xi}_R)^{-1}$ existe, les r variables sont reconstruisibles sinon elles ne sont que partiellement reconstruisibles.

La structure particulière de la matrice G_R montre bien que seules les composantes du vecteur x associées à l'ensemble R sont reconstruites à partir du modèle ACP réduit à ℓ composantes principales robustes et des $n - r$ autres composantes de x .

3.2. Génération de résidus structurés

Dans l'objectif du diagnostic, l'estimé (9a) peut être comparé aux mesures, ce qui permet de construire un résidu (Dunia *et al.*, 1998). Il est défini par la projection du vecteur des variables reconstruites dans l'espace résiduel :

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

$$P_R^{(\ell)} = (I - C_\ell) G_R \quad (11)$$

Propriété. La matrice de projection $P_R^{(\ell)}$ vérifie les deux propriétés suivantes :

$$P_R^{(\ell)} \Xi_R = 0, \quad \Xi_R^T P_R^{(\ell)} = 0 \quad (12)$$

Par conséquent, les composantes de \tilde{x}_R ne sont donc pas sensibles aux composantes de x correspondant aux indices contenus dans l'ensemble R . Cette propriété est utilisée pour localiser les composantes de x affectées par un défaut.

Pour préciser ce point fondamental, considérons le cas d'une donnée saine x^* corrompue par un bruit de valeur moyenne nulle ϵ et d le vecteur des amplitudes des défauts agissant dans les directions Ξ_F :

$$x = x^* + \epsilon + \Xi_F d \quad (13)$$

Dans cette expression x^* est la vraie valeur (et donc satisfait au modèle *ACP*) et x est une observation disponible de x^* . En vertu de (10), le résidu calculé en reconstruisant les r composantes de x s'explique :

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

dont l'espérance mathématique vaut :

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

- si la direction de reconstruction Ξ_R est celle des défauts, alors toutes les composantes du vecteur $P_R^{(\ell)} \Xi_F$ sont nulles et $\mathcal{E}(\tilde{x}_R) = 0$,
- si la direction de reconstruction Ξ_R est différente de celle des défauts, alors les composantes du vecteur $P_R^{(\ell)} \Xi_F$ ne sont pas a priori nulles.

Ainsi, en considérant toutes les reconstructions possibles, l'analyse des amplitudes des résidus \tilde{x}_R obtenus est révélatrice de la présence de défauts et permet de déterminer les composantes de la mesure affectées par des défauts.

4. Résultats numériques

Pour mettre en évidence l'aptitude de la méthode proposée à détecter des défauts de mesure, un exemple de synthèse dont les valeurs aberrantes sont parfaitement connues est utilisé. La matrice X est constituée de $N = 108$ observations d'un vecteur x à 8 composantes générées de la façon suivante :

$$\begin{aligned} x_{i,1} &= v_i^2 + \sin(0.1i), & v_i &\sim \mathcal{N}(0, 1) \\ x_{i,2} &= 2 \sin(i/6) \cos(i/4) \exp(-i/N) \\ x_{i,3} &= \log(x_{i,2}^2), & x_{i,4} &= x_{i,1} + x_{i,2} \\ x_{i,5} &= x_{i,1} - x_{i,2}, & x_{i,6} &= 2x_{i,1} + x_{i,2} \\ x_{i,7} &= x_{i,1} + x_{i,3}, & x_{i,8} &\sim \mathcal{N}(0, 1) \end{aligned} \quad (16)$$

Aux données ainsi générées ont été superposées des réalisations de variables aléatoires à distribution normale centrée et d'écart-type 0.02 ainsi que des défauts δx_1 , δx_2 , δx_3 , δx_4 représentés par des biais d'amplitude égale à 3 et définis de la façon suivante : des instants 10 à 24 pour la variable x_1 , des instants 35 à 49 pour les variables x_2 et x_3 , des instants 60 à 74 pour les variables x_3 et x_4 , des instants 85 à 99 pour la variable x_4 . Dans la suite, ces quatre intervalles sont désignés par I_1 , I_2 , I_3 , I_4 .

Concernant l'analyse a priori de la localisation des défauts, nous nous limitons à donner une table réduite de signatures (table 1) établies à partir des propriétés (12). Elle ne fait apparaître que quelques défauts possibles, notés δ , ceux affectant les variables 1, 2, 3, 4 et ceux affectant les couples de variables $\{1, 2\}$, $\{1, 3\}$, $\{1, 4\}$, $\{2, 3\}$, $\{2, 4\}$, $\{3, 4\}$. La première colonne est relative à la norme Δ_R des résidus obtenus en utilisant toutes les composantes de x sauf celles dont l'indice appartient à l'ensemble R :

$$\Delta_R = \|\tilde{x}_R\|^2 \quad (17)$$

	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	\times	\times	\times	\times	\times	0	0	0	0	0	\times	\times	\times	\times	\times	\times	
δ_2	\times	0	\times	\times	\times	\times	0	\times	\times	\times	0	0	0	0	\times	\times	\times	
δ_3	\times	\times	0	\times	\times	\times	\times	0	\times	\times	\times	0	\times	\times	\times	0	0	
δ_4	\times	\times	\times	0	\times	\times	\times	\times	0	\times	\times	\times	0	\times	\times	0	\times	
δ_{12}	\times	\times	\times	\times	\times	\times	0	\times										
δ_{13}	\times	0	\times															
δ_{14}	\times	\times	0	\times														
δ_{23}	\times	\times	\times	\times	0	\times	\times	\times	\times	\times	\times							
δ_{24}	\times	\times	\times	\times	\times	0	\times	\times	\times	\times	\times							
δ_{34}	\times	\times	\times	\times	\times	\times	\times	\times	0	\times	\times							

Tableau 1. Table réduite des signatures de défauts

Cette table, que le lecteur pourra étendre à l'ensemble des défauts, fournit une correspondance entre les symptômes et les défauts. Par exemple, le défaut δ_2 affecte toutes les projections sauf celles établies sans les composantes 2, $\{1, 2\}$, $\{2, 3\}$, $\{2, 4\}$, $\{2, 5\}$, $\{2, 6\}$.

A partir des données contaminées, le modèle *ACP* robuste, limité à quatre axes principaux choisis à partir de l'examen des valeurs propres de V , a été déterminé. La reconstruction est ensuite effectuée à partir de toutes les variables sauf la variable 1, puis à partir de toutes les variables sauf les variables 1 et 2, puis à partir de toutes les variables sauf les variables 1 et 3 ... la dernière reconstruction étant faite à partir de toutes les variables sauf les variables 7 et 8.

La figure 1 visualise la reconstruction des variables sans utiliser les variables $\{2, 3\}$. Elle montre la reconstruction des cinq premières variables. Les N données reconstruites ont ensuite été projetées sur l'espace résiduel. Pour chaque observation, l'indicateur de défaut Δ_R a été calculé (17).

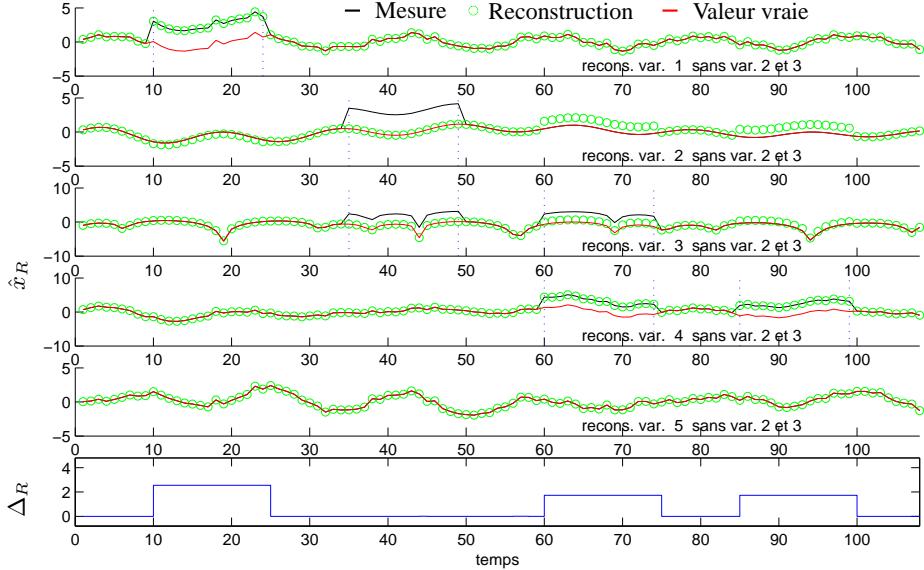


Figure 1. Reconstruction des variables sans utiliser les variables $\{2, 3\}$

Analysons la figure 1. Les variables $\{2, 3\}$, biaisées pour les observations des intervalles I_2 et I_3 , ne sont pas utilisées pour la reconstruction, la variable x_4 étant également biaisée sur l'intervalle I_3 . Pour les observations de l'intervalle I_2 , les reconstructions sont donc correctes pour toutes les variables. En effet, on observe la superposition des variables reconstruites avec les variables vraies c'est-à-dire non biaisées (en pratique ces dernières sont inconnues, mais dans cette phase où les données sont générées, la comparaison est possible). On a également indiqué la mesure de la variable afin de la comparer à la reconstruction. Ce résultat est confirmé par le dernier graphique de la figure 1 où l'on a tracé la norme du vecteur des résidus (17). Pour les observations de l'intervalle I_2 cette norme est proche de la valeur 0 témoignant ainsi de l'absence de valeurs aberrantes dans les variables utilisées pour la reconstruction et la projection, c'est-à-dire toutes les variables sauf x_2 et x_3 . On note aussi que les trois autres groupes d'observations (I_1, I_3, I_4) sont affectés de défauts, sans savoir exactement quelle(s) composante(s) du vecteur de mesure en sont la cause. Finalement, compte tenu de la présence de défauts dans les quatre intervalles, l'examen de la figure 1 conclut à :

- dans l'intervalle I_2 , x_2 et x_3 sont en défaut,
- dans chaque intervalle I_1, I_3, I_4 , d'autres variables que x_2 et x_3 sont en défaut.

Les autres projections (non présentées ici) se construisent et s'interprètent de façon analogue. La table 2 résume les conclusions issues de l'analyse des projections. La ligne Δ_1 est relative à la projection des résidus avec reconstruction sans utiliser la première variable ... Le symbole 0 témoigne de l'absence de défaut dans l'intervalle

considéré. Le diagnostic est donc :

- dans I_1 , x_1 est en défaut,
- dans I_2 , x_2 et x_3 sont en défaut,
- dans I_3 , x_3 et x_4 sont en défaut,
- dans I_4 , x_4 est en défaut.

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

Tableau 2. Signatures de défauts

5. Conclusion

Lorsque des valeurs aberrantes corrompent les données, l'*ACP* classique s'avère inefficace, alors que sa version robuste donne des résultats tout à fait satisfaisants. Sur les exemples traités, la présence d'environ 25 pour-cent de données aberrantes autorise une estimation correcte des directions principales c'est-à-dire peu sensibles à ces valeurs. On peut donc construire un modèle *ACP* directement à partir des données disponibles contenant d'éventuels défauts.

Le résultat le plus important concerne le diagnostic de fonctionnement de système, appliqué ici à la détection et localisation de données aberrantes. Pour cela, on a montré comment construire des indicateurs de défaut. L'utilisation conjointe des principes de reconstruction et de projection des données a permis de détecter et d'isoler de façon efficace les données aberrantes.

La procédure proposée ici n'est pas, en principe, limitée par le nombre de variables. Cependant, le volume de calcul risque de devenir incompatible avec un traitement en ligne des données et une réduction du nombre de reconstructions et projections est envisageable ; ce point méritera une attention particulière dans la suite de nos travaux.

Remerciement

Ce travail a été réalisé avec le soutien du Programme TASSILI n°07 MDU 714.

6. Bibliographie

- Dunia R., Qin S. J., A subspace approach to multidimensional fault identification and reconstruction. *AIChE Journal*, vol.44, n°8, 1998, p.1813-1831.
- Fekri M., Ruiz-Gazen A., Robust weighted orthogonal regression in the errors-in-variables model. *Journal of Multivariate Analysis*, vol.88, 2003, p.89-108.
- Hubert M., Rousseeuw P.J., Van den Branden K., ROBPCA : a New Approach to Robust Principal Component Analysis. *Technometrics*, vol.47, n°1, 2005, p. 64-79.
- Jolliffe I., *Principal Component Analysis*. Springer-Verlag, New York, 1986.