Multi-dimensional Fault Diagnosis Using a Subspace Approach

Ricardo Dunia and S. Joe Qin
Department of Chemical Engineering
University of Texas at Austin

In *American Control Conference 1997*

Presented by Bowen Zhou

Overview

- Fault detection using principal component analysis (PCA) have been studied intensively and applied to industrial processes.
- PCA is used to define an *orthogonal* partition of the measurement space into two *orthogonal* subspaces: a principal component space (PCS) and a residual subspace (RS).
- This paper considered each fault as a subspace and addressed fundamental issues of fault detectability, reconstructability, identifiability and isolatability. (The last two issues apply to the multi-fault case only. I won't covered them in this talk due to time constraint)
- Pure theorectical paper with lots of matrix analysis.

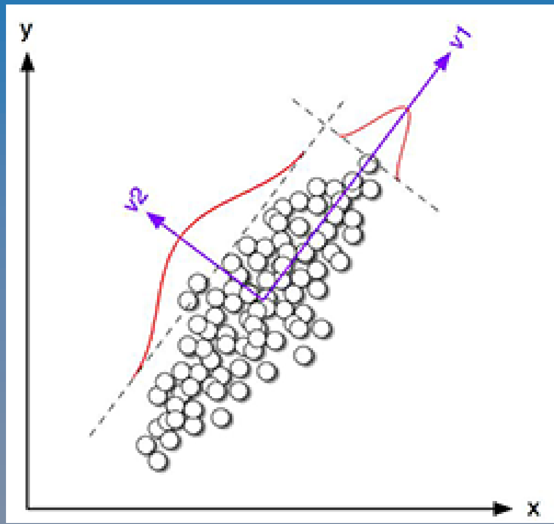
Content

- Introduction
- Detectability
- Reconstructability
- Conclusion

What is PCA?

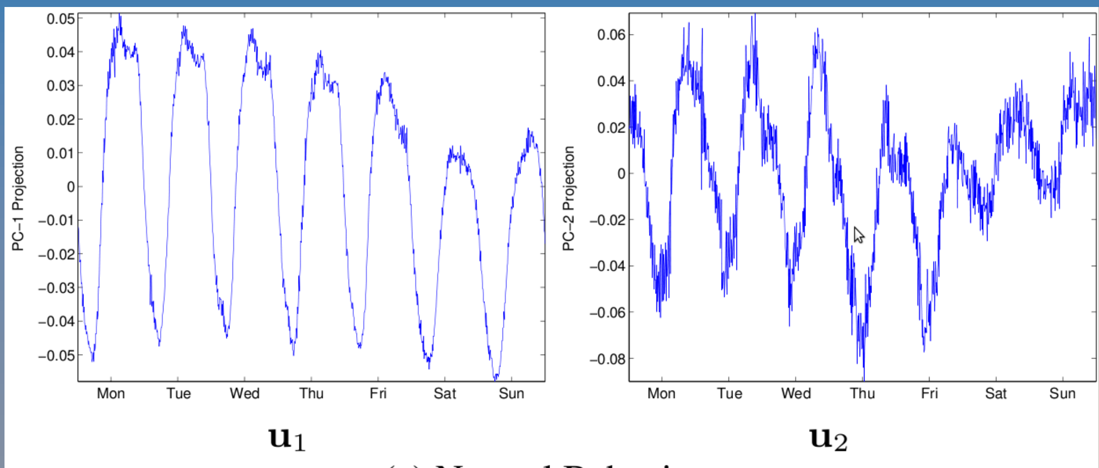
- PCA is mathematically defined as an orthogonal linear transformation that transforms the data to a new coordinate system such that the greatest variance by any projection of the data comes to lie on the first coordinate (called the first principal component), the second greatest variance on the second coordinate, and so on.
- PCA involves the calculation of the eigenvalue decomposition of a data covariance matrix or singular value decomposition of a data matrix, usually after mean centering the data for each attribute.

What is PCA?



Why use PCA to detect faults?

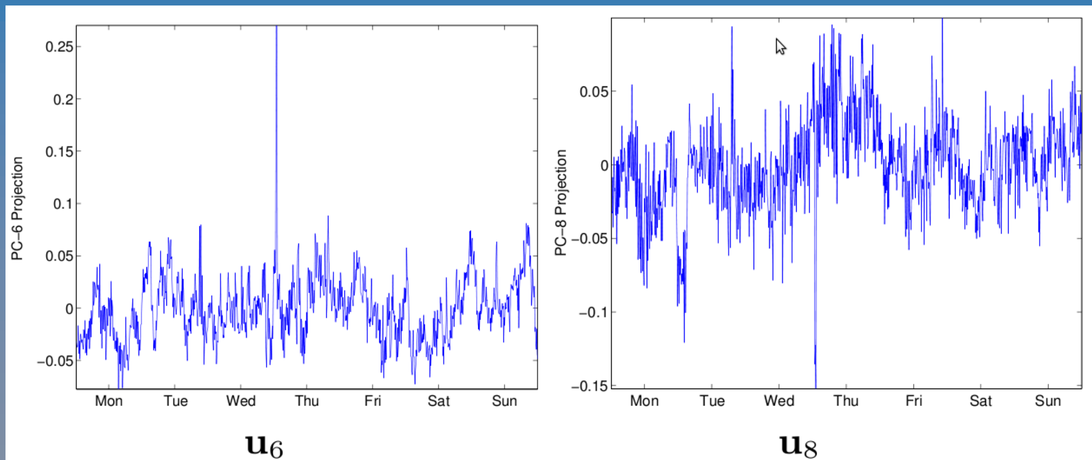
- What is NORMAL? (A plot from "Diagnosing Network-Wide Traffic Anomalies" showing NORMAL network traffic)



(a) Normal Behavior

Why use PCA to detect faults?

- What is ABNORMAL? (Another plot from "Diagnosing Network-Wide Traffic Anomalies" showing ABNORMAL network traffic)



(b) Anomalous Behavior

Why use PCA to detect faults?

- PCA is a reliable technique for capturing variable correlation.
- Violation of such correlation is assumed to be a potential anomaly.

Fault Detection

- A PCA model decomposes a normalized a sample vector into two portions

$$\mathbf{x} = \hat{\mathbf{x}} + \tilde{\mathbf{x}}$$

where $\mathbf{x} \in \mathcal{R}^{\mathbf{m}}$ is the sample vector, $\hat{\mathbf{x}}$ the modeled vector and $\tilde{\mathbf{x}}$ the residual vector.

- The modeled portion is the projection on the PCS

$$\hat{\mathbf{x}} = \mathbf{P}\mathbf{P}^{\mathbf{T}}\mathbf{x} = \mathbf{C}\mathbf{x}$$

where $\mathbf{P} \in \mathcal{R}^{\mathbf{m}\times\mathbf{l}}$ is the PCA loading matrix. The parameter l is the number of principal component (PC) retained in the PCA model. The matrix $\mathbf{C} \equiv \mathbf{P}\mathbf{P}^{\mathbf{T}}$ represents the projection on the l dimension principal component subspace, $\mathcal{S}$.

Fault Detection

- The residual $\tilde{\mathbf{x}}$ lies in the residual subspace (RS) of $\tilde{l} = m - l$ dimensions.

$$\tilde{\mathbf{x}} = (\mathbf{I} - \mathbf{C})\mathbf{x} = \tilde{\mathbf{C}}\mathbf{x}$$

where $\tilde{\mathbf{C}}$ represents the projection matrix on the residual subspace, $\tilde{\mathcal{S}}$.

- A fault may cause the sample $\mathbf{x}$ increases its projection to the residual subspace.

- As a result, the magnitude of $\tilde{\mathbf{x}}$ reaches unusual values compared to those obtained during normal conditions.

A statistic for Fault Detection

- A typical statistic for detecting anomalous behavior is the *squared prediction error* (SPE).

$$\mathbf{SPE} \equiv ||\tilde{\mathbf{x}}||^{\mathbf{2}} = \left|\left|\tilde{\mathbf{C}}\mathbf{x}\right|\right|^{\mathbf{2}}$$

and the process is considered normal if

$$\mathbf{SPE} \leq \delta^{\mathbf{2}}$$

where δ^2 denotes a confidence limit or threshold corresponding to a predefined confidence level, e.g. 90%.

Necessary Conditions for Detectability

- Denote $\mathcal{S}_\mathbf{i}$ as the subspace of the actual fault $\mathcal{F}_\mathbf{i}$, $dim(\mathcal{S}_\mathbf{i}) = \mathbf{l_i} \leq \mathbf{m}$ and $\mathcal{S}_\mathbf{i} \subseteq \mathcal{R}^\mathbf{m}$.

- A set of basis for $\mathcal{S}_\mathbf{i}$ can be chosen as columns of $\mathbf{\Xi_i} \in \mathcal{R}^{\mathbf{m}\times\mathbf{l_i}}$

- The sample vector for normal behavior is denoted by $\mathbf{x}^*$, which is unknown when a fault has occurred.

- In the presence of $\mathcal{F}_\mathbf{i}$, the sample vector $\mathbf{x}$ is represented by

$$\mathbf{x} = \mathbf{x}^* + \mathbf{\Xi_i}\mathbf{f}$$

where $\mathbf{\Xi_i}$ is orthonormal and $||\mathbf{f}||$ represents the magnitude of the fault.

Necessary Conditions for Detectability

- The columns of $\mathbf{\Xi_i}$ can be projected onto $\mathcal{S}$ and $\tilde{\mathcal{S}}$

$$\mathbf{\Xi_i} = \hat{\mathbf{\Xi}}_\mathbf{i} + \tilde{\mathbf{\Xi}}_\mathbf{i}$$

where $\hat{\mathbf{\Xi}}_\mathbf{i} \equiv C\mathbf{\Xi_i}$ and $\tilde{\mathbf{\Xi}}_\mathbf{i} \equiv \tilde{C}\mathbf{\Xi_i}$ belongs to $\mathcal{S}$ and $\tilde{\mathcal{S}}$, respectively.

- Project $\mathbf{x}$ onto $\tilde{\mathcal{S}}$, we get

$$\tilde{\mathbf{x}} = \tilde{\mathbf{x}}^* + \tilde{\mathbf{\Xi}}_\mathbf{i}\mathbf{f}$$

- Although $\mathbf{\Xi_i}$ has full column rank, $\tilde{\mathbf{\Xi}}_\mathbf{i}$ may not have full column rank. So we apply singular value decomposition to $\tilde{\mathbf{\Xi}}_\mathbf{i}$,

$$\tilde{\mathbf{\Xi}}_\mathbf{i} \equiv \tilde{\mathbf{\Xi}}_i^0 \tilde{D}_i \tilde{V}_i^T$$

where $\tilde{D}_i \in \mathcal{R}^{\hat{l}_i \times \hat{l}_i}$ contains non-zero singular values of $\tilde{\mathbf{\Xi}}_\mathbf{i}$.

Necessary Conditions for Detectability

- $\tilde{\mathbf{\Xi}}_i^0$ represents the non-vanishing directions of $\tilde{\mathbf{\Xi}}_\mathbf{i}$

$$\tilde{\mathbf{x}} = \tilde{\mathbf{x}}^* + \tilde{\mathbf{\Xi}}_\mathbf{i}^\mathbf{0}\tilde{\mathbf{f}}$$

where $\tilde{f} \equiv \tilde{D}_i \tilde{V}_i^T f$ represents that fault displayment projected on $\tilde{\mathcal{S}}$.

- Therefore, the SPE can be represented as follows,

$$\mathbf{SPE} = \left\| \tilde{\mathbf{x}}^* + \tilde{\mathbf{\Xi}}_\mathbf{i}^\mathbf{0}\tilde{\mathbf{f}} \right\|^\mathbf{2}$$

- If a fault happens but it does not affect the SPE, this fault cannot be detected by this equation.

Necessary Conditions for Detectability

The following three cases are the necessary conditions for detectability:

1. If $\tilde{\mathbf{\Xi}}_\mathbf{i} = 0$, the fault is not detectable no matter what $\mathbf{f}$ is.

2. If $\tilde{\mathbf{\Xi}}_\mathbf{i} \neq 0$, but $\tilde{\mathbf{\Xi}}_\mathbf{i}$ is not full ranked, the fault is not detectable if $\mathbf{f}$ is in the null space of $\tilde{\mathbf{V}}_\mathbf{i}^\mathbf{T}$,
$$\mathbf{f} \in \mathcal{N}(\tilde{\mathbf{V}}_\mathbf{i}^\mathbf{T})$$
which makes $\tilde{\mathbf{f}} = 0$.

3. If a fault is detectable, then $\tilde{\mathbf{f}} \neq 0$, which implies one the following situations:
   (a) $\tilde{\mathbf{\Xi}}_\mathbf{i}$ has full column rank, or
   (b) $\mathbf{f} \notin \mathcal{N}(\tilde{\mathbf{V}}_\mathbf{i}^\mathbf{T})$.

Sufficient Condition for Detectability

- The previous section shows that the necessary condition for a fault to be detectable is that its projection on $\tilde{\mathcal{S}}$ does not vanish. Nevertheless, $\left|\left|\tilde{\mathbf{f}}\right|\right|$ should be large enough to make the SPE exceed the confidence limit δ^2.

- The sufficient condition for detectability can be obtained using the following triangular inequality

$$||\tilde{\mathbf{x}}|| = \left|\left|\tilde{\mathbf{x}}^* + \tilde{\mathbf{\Xi}}_\mathbf{i}^\mathbf{0}\tilde{\mathbf{f}}\right|\right| \geq \left|\left|\tilde{\mathbf{f}}\right|\right| - ||\tilde{\mathbf{x}}^*||$$

and the substitution of $||\tilde{\mathbf{x}}^*|| \leq \delta$ gives

$$||\tilde{\mathbf{x}}|| \geq \left|\left|\tilde{\mathbf{f}}\right|\right| - \delta$$

- To make the fault sufficiently detectable, the condition $\mathbf{SPE} = ||\tilde{\mathbf{x}}||^{\mathbf{2}} > \delta^{\mathbf{2}}$ should be satisfied, which requires

$$\left|\left|\tilde{\mathbf{f}}\right|\right| > \mathbf{2}\delta$$

- Therefore, a fault $\mathcal{F}_i$ is guaranteed detectable with a confident limit δ^2 if $\left|\left|\tilde{\mathbf{f}}\right|\right| > \mathbf{2}\delta$

Fault Reconstruction

- When a fault is detected, it is desirable to determine the necessary adjustments to bring the system back to normal conditions.

- In this section, it is assumed that the real fault has been identified as $\mathcal{F}_i$, and the normal portion $\mathbf{x}^*$ will be reconstructed from the corrupted vector $\mathbf{x}$.

- The reconstruction consists of estimating $\mathbf{x}^*$ by eliminating the effect of $\mathcal{F}_i$ from the sample vector $\mathbf{x}$,

$$\mathbf{x_i} = \mathbf{x} - \mathbf{\Xi_i f_i}$$

where $\mathbf{f}_i$ is an estimate of the fault magnitude $\mathbf{f}$.

Fault Reconstruction

- The best estimate of $\mathbf{x}^*$ is given by minimizing the distance from $\mathbf{x}_i$ to the PCS

$$\mathbf{f_i} = \mathbf{argmin} \, ||\tilde{\mathbf{x}}_\mathbf{i}||^2$$

where $\tilde{\mathbf{x}}_\mathbf{i} = \tilde{\mathbf{x}} - \tilde{\boldsymbol{\Xi}}_\mathbf{i}\mathbf{f_i}$ is the projected vector of $\mathbf{x_i}$ to RS.

- Let $\tilde{\mathbf{x}}_\mathbf{i} = \mathbf{0}$, solve the equation to get,

$$\mathbf{f_i} = (\tilde{\boldsymbol{\Xi}}_\mathbf{i}{}^\mathbf{T}\tilde{\boldsymbol{\Xi}}_\mathbf{i})^{-1}\tilde{\boldsymbol{\Xi}}_\mathbf{i}{}^\mathbf{T}\tilde{\mathbf{x}}$$

- This result shows that $\mathbf{f}_i$ can be uniquely calculated if and only if $(\tilde{\boldsymbol{\Xi}}_\mathbf{i}{}^T\tilde{\boldsymbol{\Xi}}_\mathbf{i})^{-1}$ exists, i.e. $dim\{\tilde{\mathcal{S}}_\mathbf{i}\} = \mathbf{l_i}$
- This case is referred as the *complete reconstruction*. As a consequence, the necessary and sufficient condition for *complete reconstructability* is that $\tilde{\boldsymbol{\Xi}}_\mathbf{i}$ has full column rank.

Fault Reconstruction

- What if $\tilde{\boldsymbol{\Xi}}_\mathbf{i}$ does NOT have full column rank?
- There is no unique solution for $\mathbf{f_i}$ because $\tilde{\boldsymbol{\Xi}}_\mathbf{i}{}^T\tilde{\boldsymbol{\Xi}}_\mathbf{i}$ does not have an inverse.
- Such a fault can only be *partially* eliminated along the non-vanishing directions $\tilde{\boldsymbol{\Xi}}_i^0$

$$\mathbf{f_i} = \tilde{\mathbf{V}}_\mathbf{i}\tilde{\mathbf{D}}_\mathbf{i}^{-1}(\tilde{\boldsymbol{\Xi}}_\mathbf{i}^\mathbf{0})^\mathbf{T}\tilde{\mathbf{x}}$$

- The above equation shows that $\mathbf{f_i}$ can be calculated as long as $\tilde{\mathbf{D}}_\mathbf{i} \neq \mathbf{0}$, which implies $\tilde{\boldsymbol{\Xi}}_\mathbf{i} \neq 0$.
- Therefore, a fault is *partially reconstructable* if and only if $\tilde{\boldsymbol{\Xi}}_\mathbf{i} \neq 0$

Conclusions

- Fault detection, reconstruction, identification and isolation are proposed using a multidimensional approach which describes faults using the fault subspace.

- The sufficient and necessary conditions for fault detectability and reconstructability are analyzed using this approach.

- Identifiability and Isolatability concern about the separation of multiple faults in the same system.

Questions?