POLITECNICO MILANO 1863

FAULT DETECTION

Ibrahim Ahmed, Department of Energy, Politecnico di Milano

lasar

- Part 1: Model of the Equipment Behavior in Normal Condition
 - Principal Component Analysis (PCA)
- Part 2: Statistical Test
 - 2A) Thresholds-Based
 - 2B) Sequential Probability Ratio Test (SPRT)

Overview

Context: Prognostics and Health Management

Equipment (System, Structure or Component)



In This Lecture: Fault Detection

Equipment (System, Structure or Component)



Fault Detection: What is?



Fault Detection: What is?



Fault Detection: What is not?



Fault Detection: Approaches

- Limit-based
- Model-based
- Data-driven

Limit-based fault detection: data & information

Normal operation ranges of key signals

Example:

Pressurizer of a nuclear reactor



Limit-based fault detection: the method

• Normal operation ranges of key signals



• Limit Value-Based Fault Detection

Example:

Pressurizer of a nuclear reactor



Limit-based fault detection: Limitations

- Normal operation ranges of key signals
- Limit Value-Based Fault Detection

Example:



Pressurizer of a PWR nuclear reactor

Limitations:

- No early detection
- •Not applicable to fault detection during operational transients

Control systems operations may hide small anomalies (the signal remains in the normal range although there is a process anomaly)
Considering signal individually can delay detection



Fault Detection: Approaches

- Limit-based
- Model-based
- Data-driven

Model-based & Data-driven fault detection: basic idea



Fault Detection: Approaches

- Limit-based
- Model-based
- Data-driven

Model-based fault detection: data & information

 Physics-based model of the process (used to reproduce the expected behavior of the signals in normal condition)





 Physics-based model of the process (used to reproduce the expected behavior of the signals in normal condition)

Example:



Fault Detection: Approaches

- Limit-Based
- Model Based
- Data-driven

Data-driven fault detection: data & information

Historical signal measurements in normal operation

Example:

Pressure	Liquid temperat ure	Steam temperat ure	Spray flow	Surge line flow	Heaters power	Level
150.2	321	362	539	244	0	7.2
150.4	322	363	681	304	0	7.5
150.3	323	364	690	335	1244	7.7



Water level

Data-driven fault detection: possible methods

- Statistical Approaches:
 - Principal Component Analysis (PCA)-based
 - AutoAssociative Kernel Regression (AAKR)

• ...

- Artificial Intelligence (AI)-based
 - Feedforward Neural Networks (FNNs)
 - AutoAssociative Neural Networks (AANNs)
 - AutoEncoders (AEs)
 - Self Organizing Maps

• ...

Part 1: Model of the Equipment Behavior in Normal Condition

Principal Component Analysis (PCA)

Part 2: Statistical Test

- 2A) Thresholds-Based
- 2B) Sequential Probability Ratio Test (SPRT)

In This Lecture



PART 1: Model of the Equipment Behaviour in Normal Condition

Principal Component Analysis (PCA)

Data in normal conditions



obs-nc = observation in normal condition

Training set, input and output

 Training patterns: Historical signal measurements in normal condition



- Test input: $\vec{x}^{obs} = (x_1^{obs}, ..., x_n^{obs})$ Signals measured at current time
- Test Output: $\vec{\hat{x}}^{nc} = (\hat{x}_1^{nc}, ..., \hat{x}_n^{nc})$ Signal reconstructions (expected values of the signals in normal condition)



Requirement I

Equipment is in normal condition



Requirement II

• Equipment is in abnormal condition



PCA: What is it?

PCA:

- Space transformation
- From an *n*-dimensional space to a *l*-dimensional space (l < n)
- Retaining most of the information (loosing the least information)



IDEA OF PCA

• Two signals are highly correlated or dependent \rightarrow One is enough! $x_2 \uparrow u$



Key underlying phenomena
 → Areas of variance in data
 → Focus on directions along which X₂ the observations have largest variance



 X_1

PCA: Training set, input and output = Slide 25

 Training patterns: Historical signal measurements in normal condition



- Test input: $\vec{x}^{obs} = (x_1^{obs}, ..., x_n^{obs})$ Signals measured at current time
- Test Output: $\vec{\hat{x}}^{nc} = (\hat{x}_1^{nc}, ..., \hat{x}_n^{nc})$ Signal reconstructions (expected values of the signals in normal condition)



PCA for fault detection: operational steps (1)

Step 1: find Principal Components (PCs) in the training set X^{obs-nc} .

 PC1 → is the direction of maximum variance

 PC2 -> is orthogonal to PC1 and describes the maximum residual variance



3) PC3 ->

is **orthogonal** to PC1 and PC2 and describes the maximum residual variance

Step 1: Mathematical details (1A)

Objective: find principal components

Procedure:

• Compute $V = \text{covariance matrix of } X^{obs-nc}$



Step 1: Mathematical Details (1B)

Objective: find principal components

Procedure:

- Compute $V = \text{covariance matrix of } X^{obs-nc}$
- Find the *n* eigenvectors $\vec{p}_1, \vec{p}_2, ..., \vec{p}_n$ of *V* and the corresponding eigenvalues $\lambda_1 \ge \lambda_2 \ge \lambda_3 \ge \cdots \ge \lambda_n$



33

Step 1: Properties of the PCs (I)

$$P = \begin{bmatrix} \vec{p}_1, \vec{p}_2 \end{bmatrix} = \begin{bmatrix} 0.28 & -0.96 \\ 0.96 & 0.28 \end{bmatrix}$$

> *P* is an orthonormal basis:



Step 1: Properties of the PCs (II)

- \succ *P* is an orthonormal basis:
 - Data can be transformed from the original to the transformed bases and viceversa without any loss of information (multiplication for P and P^T)
 - \vec{u} = the projection of \vec{x} on the new basis is given by: $\vec{u} = \vec{x} \cdot P$



Step 1: Properties of the PCs (III)

- - P is an orthonormal basis:
 - Data can be transformed from the original to the transformed bases and viceversa without any loss of information (multiplication for P and P^T)
 - \vec{u} = the projection of \vec{x} on the new basis is given by: $\vec{u} = \vec{x} \cdot P$
 - \vec{x} can be obtained from \vec{u} by: $\vec{x} = \vec{u} \cdot P^T$


Step 1: Properties of the PCs (III)

- - P is an orthonormal basis:
 - Data can be transformed from the original to the transformed bases and viceversa without any loss of information (multiplication for P and P^T)
 - \vec{u} = the projection of \vec{x} on the new basis is given by: $\vec{u} = \vec{x} \cdot P$
 - \vec{x} can be obtained from \vec{u} by: $\vec{x} = \vec{u} \cdot P^T$
 - The percentage of variance retained by the *i*-th principal component is:

$$%Var(PC_i) = \frac{\lambda_i}{\sum_{i=1,\dots,n} \lambda_i}$$

PCA for fault detection: operational steps (1)

Step 2 [PCA approximation]: ignore the PCs of lower significance.



- Lost small information
- Reduce the number of dimensions from *n*=10 to = 4

PCA for fault detection: operational steps (2)

- - Step 2 [PCA approximation]: ignore the PCs of lower significance.

map the observation \vec{x}^{obs} in a subspace $\Re^l \subset \Re^n$ identified by the first | < n eigenvectors $\vec{p}_1, ..., \vec{p}_l$: $\vec{x}^{obs} P_l$ with $P_l = [\vec{p}_1, ..., \vec{p}_l]$



PCA for fault detection: operational steps (3)

- Step 2 [PCA approximation]: ignore the PCs of lower significance.

map the observation \vec{x}^{obs} in a subspace $\Re^l \subset \Re^n$ identified by the first l < n eigenvectors $\vec{p}_1, \dots, \vec{p}_l$:

 $\vec{x}^{obs}P_l$ with $P_l = [\vec{p}_1, ..., \vec{p}_l]$ • Step 3: [Antitransformation]: signal reconstructions $\vec{\hat{x}}^{nc} = \vec{x}^{obs}P_l P_l^T$



PCA for fault detection: Summary

Historical data $X^{obs-nc} =$

$$\begin{bmatrix} x_{11}^{obs-nc} & \dots & x_{1j} & \dots & x_{1n}^{obs-nc} \\ \dots & \dots & \dots & \dots \\ x_{k1} & \dots & x_{kj} & \dots & x_{kn} \\ \dots & \dots & \dots & \dots \\ x_{N1}^{obs-nc} & \dots & x_{Nj} & \dots & x_{Nn}^{obs-nc} \end{bmatrix}$$

abs-nc

Find P_l from $X^{\mathit{obs-nc}}$

PCA for fault detection: Summary

Historical data $X^{obs-nc} = \begin{bmatrix} x_{11}^{obs-nc} & \dots & x_{1j} & \dots & x_{1n}^{obs-nc} \\ \dots & \dots & \dots & \dots & \dots \\ x_{k1} & \dots & x_{kj} & \dots & x_{kn} \\ \dots & \dots & \dots & \dots & \dots \\ x_{N1}^{obs-nc} & \dots & x_{Nj} & \dots & x_{Nn}^{obs-nc} \end{bmatrix}$

Find P_l from X^{obs-nc}

- Measured signals at present time: $\vec{x}^{obs} = (x_1^{obs}, ..., x_n^{obs})$
 - Transform and project $\vec{x}^{obs}P_l$

I'm looking at the measurements considering only the directions that are most meaningful in normal condition (directions of maximum variance)

• Antitrnansform
$$\hat{\vec{x}}^{nc} = \vec{x}^{obs} P_l P_l^T$$

Signal reconstructions



 $\vec{x}^{nc} \cong \vec{x}^{obs} \rightarrow$ normal condition

I loose only the irrelevant noise

POLITECNICO MILANO 1863

 $\hat{x}^{nc} \neq \hat{x}^{obs} \rightarrow$ abnormal condition <u>The process is changed</u>

Exercise 1

•Measured signals at present time: $\vec{x}^{obs} = (x_1^{obs}, x_2^{obs})$ •

- •Signal reconstructions?
- •Normal or abnormal condition?



•available historical signal measurements in normal plant condition

Exercise 1: Solution

•Measured signals at present time: $\vec{x}^{obs} = (x_1^{obs}, x_2^{obs})$ •Step 1: find principal components: \vec{p}_1 , \vec{p}_2



•available historical signal measurements in normal plant condition

Exercise 1: Solution

•Measured signals at present time: $\vec{x}^{obs} = (x_1^{obs}, x_2^{obs})$ •Step 1: find principal components \vec{p}_1, \vec{p}_2



Exercise 1: Solution







Exercise 2

- •Measured signals at present time:
- •Signal reconstructions?
- •Normal or abnormal condition?



$$\vec{x}^{obs} = (x_1^{obs}, x_2^{obs}) \quad \bullet$$

•available historical signal measurements in normal plant condition

Exercise 2: Solution



Exercise 2: Solution



Exercise 2: Solution



Computational time:

- Training time = computational time necessary to find the Principal Components is proportional to the number of measured signals n
- Execution time: very short (only 2 matrix multiplications)
 → OK for online applications

Performance:

Unsatisfactory for dataset characterized by highly nonlinear relationships



Part 2: Statistical Test



54

- Part 1: Model of the Equipment Behavior in Normal Condition
 - Principal Component Analysis (PCA)
- Part 2: Statistical Test
 - 2A) Thresholds-Based
 - 2B) Sequential Probability Ratio Test (SPRT)

PART 2: Statistical Test

- Thresholds-based
- Sequential Probability Ratio Test (SPRT)

Abnormal condition detection: decision

Basics of the decision: residual analysis





- Methods
 - Thresholds-based approach
 - Stochastic approaches:
 - Q Statistics
 - Sequential Probability Ratio Test (SPRT)

PART 2 A: Statistical Test

- Thresholds-based
- Sequential Probability Ratio Test (SPRT)

Thresholds-based



Thresholds-Based: Remarks

- Easy to apply
- Thresholds setting is difficult and error-prone



PART 2 B: Statistical Test

- General Idea
- Sequential Probability Ratio Test (SPRT)

Stochastic approaches

- Residual (r)= random variable described by a probability law
- The probability law is different in case of normal/abnormal condition





• $R_T = \{r^{(1)}, \dots, r^{(T)}\}$ sequence of residuals at time $t = 1, \dots, T$, where $r^{(t)} = x^{obs}(t) - \hat{x}^{nc}(t)$

- Binary hypothesis test:
 - Null hypothesis $(H_0) \equiv$ Normal condition



• Alternative hypothesis $(H_1) \equiv$ Abnormal condition $r^{(t)} \sim \mathcal{N}(\mu_1, \sigma), \forall t$



SPRT: the decision



SPRT Theorem



SPRT for the positive mean test

- Null hypothesis $(H_0) \equiv$ Normal condition $r^{(t)} \sim \mathcal{N}(0, \sigma)$
- Alternative hypothesis $(H_1) \equiv$ Abnormal condition $r^{(t)} \sim \mathcal{N}(\mu_1, \sigma)$

$$L_T = \frac{P(r^{(1)}, \dots, r^{(T)} | H_1)}{P(r^{(1)}, \dots, r^{(T)} | H_0)} = e^{-\frac{1}{2\sigma^2} \sum_{t=1}^T \mu_1 (\mu_1 - 2r^{(t)})} = e^{\frac{\mu_1}{\sigma^2} \sum_{t=1}^T \left(r^{(t)} - \frac{\mu_1}{2} \right)}$$



$$\ln(L_T) = \frac{\mu_1}{\sigma^2} \sum_{t=1}^T \left(r^{(t)} - \frac{\mu_1}{2} \right) = \frac{\mu_1}{\sigma^2} \sum_{t=1}^{T-1} \left(r^{(k)} - \frac{\mu_1}{2} \right) + \frac{\mu_1}{\sigma^2} \left(r^T - \frac{\mu_1}{2} \right)$$
Sequential
$$= \ln(L_{T-1}) + \frac{\mu_1}{\sigma^2} \left(r^{(T)} - \frac{\mu_1}{2} \right)$$
Formula!

SPRT: Example



- the residual variance in normal condition (σ^2)
- the expected offset amplitude (μ_1)
- the maximum acceptable false alarm rate (α)
- the maximum acceptable missing alarm rate (β)

Example

Time interval	Simulated
	Offset
[0-200]	No
[201-400]	Yes (amplitude =
	0.11)
[401-600]	Yes
	(amplitude =
	0.23)
[601-800]	Yes
	(amplitude =
	0.34)
[801-1000]	Yes
	(amplitude =
	0.46)



Example: residuals



Example: SPRT



Example of Applications
Example of Application 1



COMPONENT TO GAS TURBINE TO BE MONITORED



Temperature location 6 (°C)

POLITECNICO MILANO 1863

Example of Application 2





- Main Bearing+ Planetary Gear box + Gearbox + Generator
- Monitoring system: 6 accelerometers and 1 sensor measuring the rotating speed

Bedplate

Example of Application 3*

COMPONENT TO BE MONITORED

Reactor Coolant Pump of PWR Nuclear Power Plant



Measured signals	48 (Temperatures, pressures, flows,)
Available data	Historical signal measurements in normal plant condition [1 year, frequency=1/30 Hz]

* Work developed with EDF-R&D

