Détection de fautes par ACP modélisèe par des réseaux bayésiens

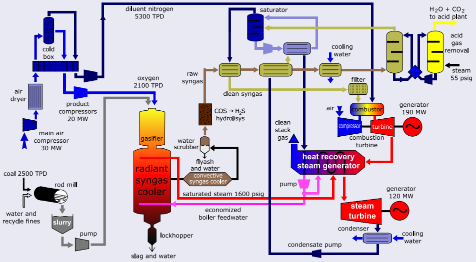
M.A.Atoui, S.Verron, A.Kobi

LARIS/ISTIA
University of Angers
L'UNAM

Réunion du GT S3
Paris

Context

A System



FDI-FDD purpose

- avoid undesirable situations/damages
- improve quality products
- cost reduction, etc

FDI-FDD methods

2 kinds of methods can be distinguished

Data-driven Methods

- Control Charts,
- PCA,
- Neural Networks,
- etc.

Model-based Methods

- Parameters Estimation,
- Observers,
- Parity Equation,
- etc.

FDI-FDD methods

2 kinds of methods can be distinguished

Data-driven Methods

- Control Charts,
- PCA,
- Neural Networks,
- etc.

Model-based Methods

- Parameters Estimation,
- Observers,
- Parity Equation,
- etc.

FDI-FDD methods

FDI-FDD methods

Data-driven methods include two steps

Fault Detection

Fault Diagnosis

FDI-FDD methods

FDI-FDD methods

Data-driven methods include two steps

Fault Detection

Fault Diagnosis

Principal Component Analysis

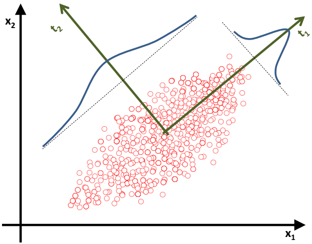
PCA :A multivariate statistical technique

projects linearly X, N normalized collected samples of m variables, onto an orthogonal space

PCA a well-known method for data reduction

Find an optimal orthogonal transformation of X :

$$\hat{T} = X\hat{P},\ \hat{T} \in \mathbb{R}^{N \times a}$$



Solution : SVD of X variance-covariance matrix

- $\Sigma_{\mathbf{x}} = P \Lambda P^T$, with
- $P \in \mathbb{R}^{m \times m}$: eigenvectors, $\Lambda = \text{diag}(\sigma_1, \ldots, \sigma_m)$: eigenvalues
- $\hat{P} \in \mathbb{R}^{m \times a}$, $\hat{\Lambda}$=diag$(\sigma_1, \ldots, \sigma_a)$

PCA to fault detection

PCA model

Seperate the original observation space X into two parts

- systematic part : $\hat{X} = \hat{T}\hat{P}^T$, noise part : $\widetilde{X} = \widetilde{T}\widetilde{P}^T$

Faut detection scheme : PCA + quadratic statistics Δ

Given a new normalized observation $x \in \mathbb{R}^m$

- monitor $\hat{t} = \hat{P}^T x$ and $\widetilde{t} = \widetilde{P}^T x$ using quadratic statistics Δ

Example : T^2 and SPE mostly used statistics

- $\mathrm{T}^2 = \hat{t}^T \hat{\Lambda} \hat{t}$, $\mathrm{SPE} = \widetilde{t}^T \widetilde{t}$
- **If** $T^2 > CL_{T^2}$ *or* $SPE > CL_{SPE}$ **Then** the system is **Out of Control** (**OC**) **Else** is **In control** (**IC**)

Bayesian framework to fault detection

Problems

- using deterministic models to deal with noisy variables
- make a decision when some variables are not observable
- represent PCA fault detection scheme through a probabilistic and graphical representation

Solution

- The use of Conditional Gaussian Networks(CGN's)

Bayesian framework : Bayesian networks

CGN a special case of BNs

Compounded of

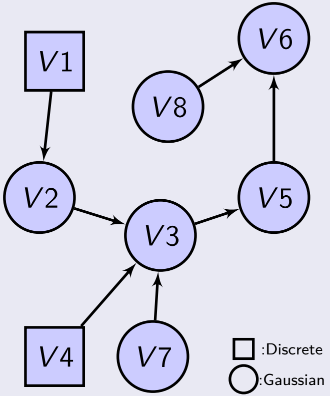
- Gaussian or discrete nodes (variables)
- Arcs (Conditional dependence/independence)

Inference

Update the network

- Given evidences
- Calculate posterior probabilities

Example



Conditional Gaussian networks

Gaussian linear node

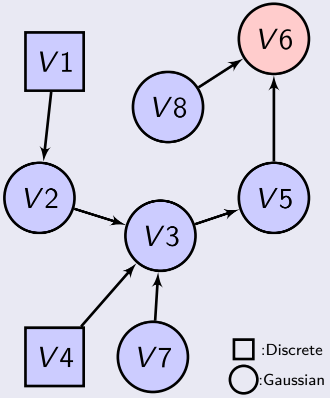
- Gaussian node
- with only Gaussian parents

Its conditional distribution

$$p(\mathbf{V6}|Pa(\mathbf{V6}) = v) = \mathcal{N}(\mu_{\mathbf{V6}} + Wv; \Sigma_{\mathbf{V6}})$$

$\mu_{\mathbf{v6}}$: is a parameter governing the mean
$\Sigma_{\mathbf{V6}}$: is the covariance matrix
W : is the regression coefficient.

Example



Probabilistic representation of PCA

PPCA

Define $\mathbf{x}$ as the linear combination of $a < m$ mutually uncorrelated latent variables plus an additive noise

$$\mathbf{x} = \hat{P}\widehat{\mathbf{t}} + \epsilon$$

$$\mathbf{x}|\widehat{\mathbf{t}} \sim \mathcal{N}(\hat{P}\widehat{\mathbf{t}}, \Psi), \ \widehat{\mathbf{t}} \sim \mathcal{N}(0, I), \epsilon \sim \mathcal{N}(0, \Psi)$$

PPCA vs PCA

Let $\Psi = vI$, $v \sim 0$, $\epsilon \sim \mathcal{N}(0, vI)$, then

$$p(\widehat{\mathbf{t}}|\mathbf{x}) = \mathcal{N}(M\hat{P}^T\mathbf{x}, \ vM), \ M \approx (\hat{P}^T\hat{P})^{-1}$$

Given an observation $x_n^T \in \mathbb{R}^m$ of X :

$$\hat{t}_n = \mathbb{E}(\widehat{\mathbf{t}}|\mathbf{x} = x_n) = M\hat{P}^T x_n \approx \hat{P}^T x_n$$

PCA model in a CGN

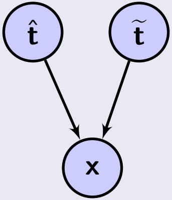
PCA model vs CGN

Let the probabilistic input-output model

$$\mathbf{x} = \hat{P}\hat{\mathbf{t}} + \tilde{P}\tilde{\mathbf{t}} + \epsilon$$

$$\mathbf{x}|\mathbf{t} \sim \mathcal{N}(P\mathbf{t}, \Psi), \ \mathbf{t} \sim \mathcal{N}(0, I), \epsilon \sim \mathcal{N}(0, \Psi), \Psi = vI$$

CGN



Inference : Bayesian calculations

$$p(\mathbf{t}|\mathbf{x}) = \mathcal{N}(MP^T\mathbf{x}, \ vM), \ M \approx (P^TP)^{-1}$$

$$\mathbf{t} = [\hat{\mathbf{t}} \ \tilde{\mathbf{t}}]^T, \tilde{\mathbf{t}} \in \mathbb{R}^{m-a}, PP^T = I$$

Conditional Gaussian networks

Gaussian node with only discrete parent

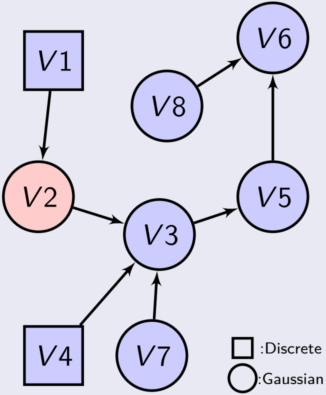
its conditional distribution for each value $i_{Pa(\mathbf{V3})}$ of its parents $Pa(\mathbf{V3})$:

$$p(\mathbf{V3}|Pa(\mathbf{V3}) = i_{Pa(\mathbf{V3})}) = \mathcal{N}(\mu_{i_{Pa(\mathbf{V3})}}; \Sigma_{i_{Pa(\mathbf{V3})}}),\ i_{Pa(\mathbf{V3})} \in I_{Pa(\mathbf{V3})}$$

CGNs for data classification

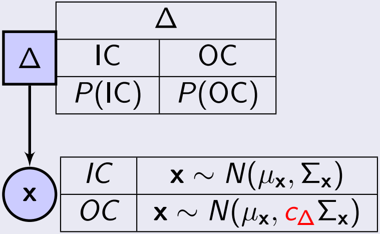
- GGN Classifiers (CGNC) equivalent to DA
- include a discrete node indexing the different classes
- affects an observation to the class with the higher posterior probability

Example



Generalization of the quadratic statistic

A CGN to fault detection



Settings : CGNC vs quadratic statistics

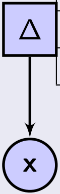
- equivalent to the Quadratic discriminant analysis
- discriminate between IC and OC
- an upper variation is assigned to OC
- the decision rule and the network parameters are defined using only the free fault data

How to match the decisions made by Δ ?

If $\Delta > CL_{\Delta}$ **Then** the system is **Out of Control** (**OC**) **Else** is **In control** (**IC**)

Generalization of the quadratic statistic

A CGN to fault detection



Δ	
IC	OC
P(IC)	P(OC)

IC	$\mathbf{x} \sim N(\mu_{\mathbf{x}}, \Sigma_{\mathbf{x}})$
OC	$\mathbf{x} \sim N(\mu_{\mathbf{x}}, c_{\Delta}\Sigma_{\mathbf{x}})$

rules : CGNC vs quadratic statistics

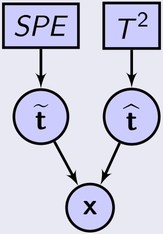
- We have for a given x^* : $\Delta = CL_{\Delta}$
- Similarly we can write
  $\zeta_{\Delta}^{ic} = P(IC_{\Delta}|x^*)$
- $\zeta_{\Delta}^{ic} = \dfrac{P(IC_{\Delta}|x^*)}{P(IC_{\Delta}|x^*) + P(IC_{\Delta}|x^*)}$
- $\zeta_{\Delta}^{ic} = \dfrac{\gamma}{1+\gamma}, \zeta_{\Delta}^{oc} = 1 - \zeta_{\Delta}^{ic}$ with
  $\gamma = \omega \dfrac{e^{\frac{(c_{\Delta}-1)}{2c} CL_{\Delta}}}{c_{\Delta}^{\frac{m}{2}}}, \omega = \dfrac{P(OC)}{P(IC)}$

Probabilistic control limit

If $P(OC_{\Delta}|x) > \zeta_{\Delta}^{oc}$ **Then** the system is declared OC **else** IC

PCA fault detection scheme in a CGN

CGN_1 : A multivariate form



T^2, SPE $\in \{IC, OC\}$

Cpt's

- CPT of $\widehat{\mathbf{t}}$:

IC_{T^2}	$\widehat{\mathbf{t}} \sim \mathcal{N}(\mu_{\widehat{\mathbf{t}}}; \hat{\Lambda})$
OC_{T^2}	$\widehat{\mathbf{t}} \sim \mathcal{N}(\mu_{\widehat{\mathbf{t}}}; \mathbf{c}_{T^2}\hat{\Lambda})$

- CPT of $\widetilde{\mathbf{t}}$:

IC_{SPE}	$\widetilde{\mathbf{t}} \sim \mathcal{N}(\mu_{\widetilde{\mathbf{t}}}; I)$
OC_{SPE}	$\widetilde{\mathbf{t}} \sim \mathcal{N}(\mu_{\widetilde{\mathbf{t}}}; \mathbf{c}_{SPE} I)$

- CPT of $\mathbf{x}$: $\mathbf{x} \sim \mathcal{N}(P\mathbf{t}; vI)$

Configuration

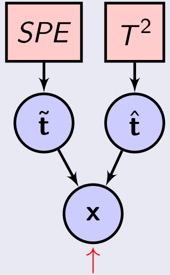
- evidence : $x = [x_1, \ldots, x_m]^T$
- **If** $P(OC_{T^2}|x) > \zeta^{oc}_{T^2}$ *or* $P(OC_{SPE}|x) > \zeta^{OC}_{SPE}$ **Then** OC **else** IC

PCA fault detection scheme in a CGN

Bayesian formulation

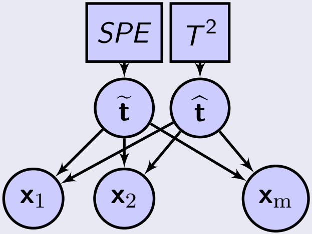
- $\mathbf{x} = P\mathbf{t} + \epsilon, \ \epsilon \sim \mathcal{N}(0, vI)$
- $P(OC_{T^2}|x) = \int\int P(x|\widehat{\mathbf{t}}, \widetilde{\mathbf{t}})P(\widehat{\mathbf{t}}|OC_{T^2})d\widetilde{\mathbf{t}}d\widehat{\mathbf{t}}$
- $P(OC_{SPE}|x) = \int\int P(x|\widehat{\mathbf{t}}, \widetilde{\mathbf{t}})P(\widetilde{\mathbf{t}}|OC_{SPE})d\widetilde{\mathbf{t}}d\widehat{\mathbf{t}}$
- $P(IC_{T^2}|x) = 1 - P(OC_{T^2}|x)$
- $P(IC_{SPE}|x) = 1 - P(OC_{SPE}|x)$

CGN_1 : a univariate form



PCA fault detection scheme in a CGN

CGN_2 : a univariate form



T^2, SPE $\in \{IC, OC\}$

Cpt's

- CPT of $\hat{\mathbf{t}}$:

IC_{T^2}	$\hat{\mathbf{t}} \sim \mathcal{N}(\mu_{\hat{\mathbf{t}}}; \hat{\Lambda})$
OC_{T^2}	$\hat{\mathbf{t}} \sim \mathcal{N}(\mu_{\hat{\mathbf{t}}}; \mathbf{c}_{T^2}\hat{\Lambda})$

- CPT of $\tilde{\mathbf{t}}$:

IC_{SPE}	$\tilde{\mathbf{t}} \sim \mathcal{N}(\mu_{\tilde{\mathbf{t}}}; I)$
OC_{SPE}	$\tilde{\mathbf{t}} \sim \mathcal{N}(\mu_{\tilde{\mathbf{t}}}; \mathbf{c}_{SPE} I)$

- CPT of $\mathbf{x}$: $\mathbf{x} \sim \mathcal{N}(B_i \mathbf{t}; v)$

Configuration

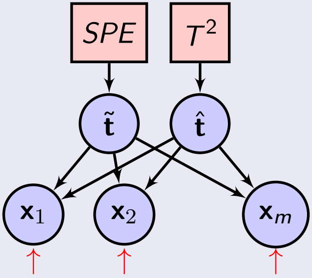
- evidence : $\mathbf{x}_1, \ldots, \mathbf{x}_m$
- **If** $P(OC_{T^2}|x_1, \ldots, x_m) > \zeta^{oc}_{T^2}$ *or* $P(OC_{SPE}|x_1, \ldots, x_m) > \zeta^{OC}_{SPE}$ **Then** OC **else** IC

PCA fault detection scheme in a CGN

Bayesian formulation

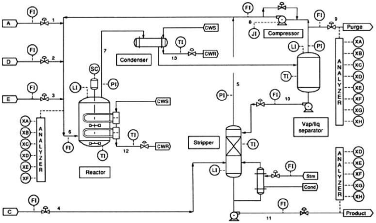
- $\mathbf{x}_\mathrm{i} = B_i\mathbf{t} + \epsilon, \ \epsilon \sim \mathcal{N}(0, v)$
- $P^T = [B_1^T \dots B_m^T], B_i^T \in \mathbb{R}^m$
- $P(OC_{T^2}|x_1, \dots, x_m) =$
  $\int\int P(x_1, \dots, x_m|\widehat{\mathbf{t}}, \widetilde{\mathbf{t}})P(\widehat{\mathbf{t}}|OC_{T^2})d\widetilde{\mathbf{t}}d\widehat{\mathbf{t}}$
- $P(OC_{SPE}|x_1, \dots, x_m) =$
  $\int\int P(x_1, \dots, x_m|\widehat{\mathbf{t}}, \widetilde{\mathbf{t}})P(\widetilde{\mathbf{t}}|OC_{SPE})d\widehat{\mathbf{t}}td\widetilde{\mathbf{t}}$
- $P(IC_{T^2}|x) = 1 - P(OC_{T^2}|x)$
- $P(IC_{SPE}|x) = 1 - P(OC_{SPE}|x)$

CGN_2 : a univariate form



1. Context
2. Problems
3. Propositions
4. Application
5. Conclusions & Outlooks

Case study : TEP, a well-known benchmark



Contains 52 variables

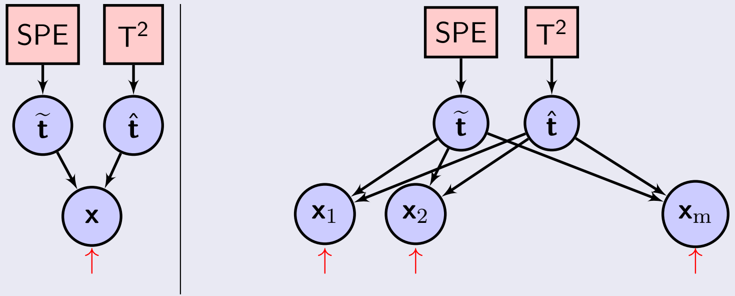
- 41 observed process variables
- 11 manipulated variables.

Settings

- consider 22 observed and 11 manipulated (m=33, a =9)
- compare the proposed method and PCA for fault detection
- use 21 test data sets (1 for normal operating conditions and others for 20 different faults).

Application

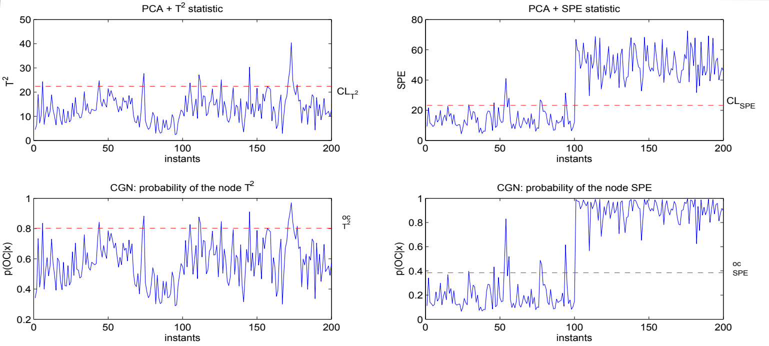
CGNs vs PCA



Test : $x(k) = [x_1(k), \ldots, x_m(k)]^T$

If $P(OC_{T^2}|x(k)) > \zeta_{T^2}^{oc}$ *or* $P(OC_{SPE}|x(k)) > \zeta_{SPE}^{OC}$ **Then** OC **else** IC

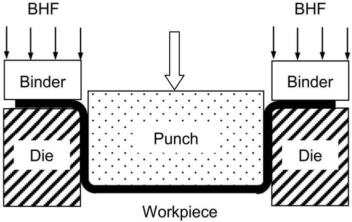
Results comparison



Test : PCA vs CGN, with $c_{SPE} = c_{T^2} = 1.005$

- consider 200 observations of respectively : the fault free and the Fault 4 data sets.
- PCA and CGNs give the same results at any instant

Case study : The hot forming process



Contains 5 variables :

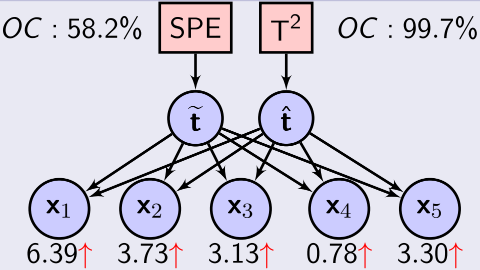
- four process variables : $\mathbf{x}_1$: temperature ; $\mathbf{x}_2$ material flow stress ; $\mathbf{x}_3$: tension in work-piece ; and $\mathbf{x}_4$: blank holding force.
- one quality variable : $\mathbf{x}_5$ final dimension of work-piece.

Simulations

- simulate 100 fault-free observations of each variable
- generate 31 faulty scenarios (introducing a mean shift with an amplitude (ms))
- compare the proposed method and PCA for fault detection

Application

Example

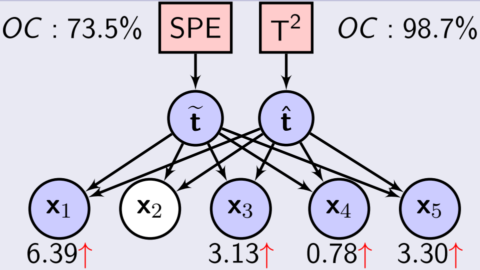


Test : all the variables are observable

- An upper violation of $\zeta_{T^2}^{oc} = 88.5\%$ and $\zeta_{SPE}^{oc} = 47.9\% \rightarrow$ OC
- Gives the same MD and FA rates obtained by PCA fault detection scheme

Application

Example



Test : some variables are not observed

- An upper violation of $\zeta_{T^2}^{oc} = 88.5\%$ and $\zeta_{SPE}^{oc} = 47.9\% \rightarrow$ OC
- The proposed method is still able, in a natural way, to make a decision

Conclusions & Outlooks

Conclusions

- an original probabilistic framework to fault detection
  - integrates PCA and its associated quadratic test statistics in a CGN
  - handles missing observation

Outlooks

- integrating other information about the system
- managing non-gaussian hypothesis, temporal behavior,non-leariliy of dynamic systems
- implementing other statistical methods and/or model-based methods for fault detection in a CGN could be interesting to represent them in a BN.

Conclusions & Outlooks

Thanks for your attention ! !