EE392m
Fault Diagnostics Systems
Introduction

Dimitry Gorinevsky
Consulting Professor
Information Systems Laboratory

Course Subject

- Engineering of fault diagnostics systems
- Embedded computer interacting with real world
  - Detect abnormal operation
  - Fault tolerance: more than 80% of critical control code
- Operations and maintenance
  - More than 50% of the system lifetime costs
  - Troubleshooting support
  - Condition-based maintenance – CBM

Prerequisites and Course Place

- The subject is not covered in other courses
- Prerequisites (helpful but not necessary)
  - Stat 116; EE263 or Eng 207a; EE278 or Eng 207b
- The course is about technical approaches that are actually used in fault diagnostics applications
  - Survived demands of real life
  - Used and supported by BS-level engineers in industry
  - Should be accessible to a Stanford grad student

Course Mechanics

- Class website: www.stanford.edu/class/ee392M/
- Weekly seminars
  - Follow website announcements
- Guest lecturers from diverse industries
  - Co-sponsored by NASA
    - Travel support for lecturers
  - Lecture notes will be posted as available
- Attendance
- Reference texts
  - Isermann; Chiang, Russel, & Braatz; Patton, Clark & Frank
  - Different coverage
  - Contact me if you have a specific interest

On-line (Embedded) Functions

- Embedded system, anomaly warnings
  - BIT – Built-in-Test
  - BITE – Built-in-Test Equipment
- FDIR
  - Fault Detection Identification and Recovery
- FT-RM
  - Fault Tolerance and Redundancy Management

Off-line Functions

- Reliability
  - FMECA- Failure Mode, Effects, and Criticality Analysis
  - Design time analysis – open loop
- Maintenance and Support
  - Diagnostics for maintenance
  - Troubleshooting support
  - Test equipment
  - CBM – Condition Based Maintenance
  - Pre-testing – disk drives

Fault Diagnostics in Industry

- Space systems
- Defense systems: aviation, marine, and ground
- Commercial aerospace
  - Aircraft, jet engines
- Ground vehicles
  - Locomotives, trucks, cars
- High-tech
  - Networks and IT systems
  - Disk drives
  - Server farms
- Process control
  - IC Manufacturing
  - Refineries
  - Power plants
- Oil and gas drilling

Guest Lecture Overview

#	Date	Lector	From	Diagnostics Application
1	31-Mar-09	Gorinevsky	Stanford	Introduction and overview
2	7-Apr-09	Rabover	VMTurbo (EMC)	Networks and IT systems
3	14-Apr-09	Tuv	Intel	IC Manufacturing processes
4	21-Apr-09	Felke	Honeywell	Avionics of commercial aircraft
5	28-Apr-09	Adibhatla	GE Infrastructure	Jet engines
6	5-May-09			
7	12-May-09	Urmanov	Sun	Computing systems
8	19-May-09	Bodden	Lockheed	Military aircraft systems
9	26-May-09	Kolmanovsky	Ford	Automotive powertrain
10	2-Jun-09			

Diagnostics Methods Overview

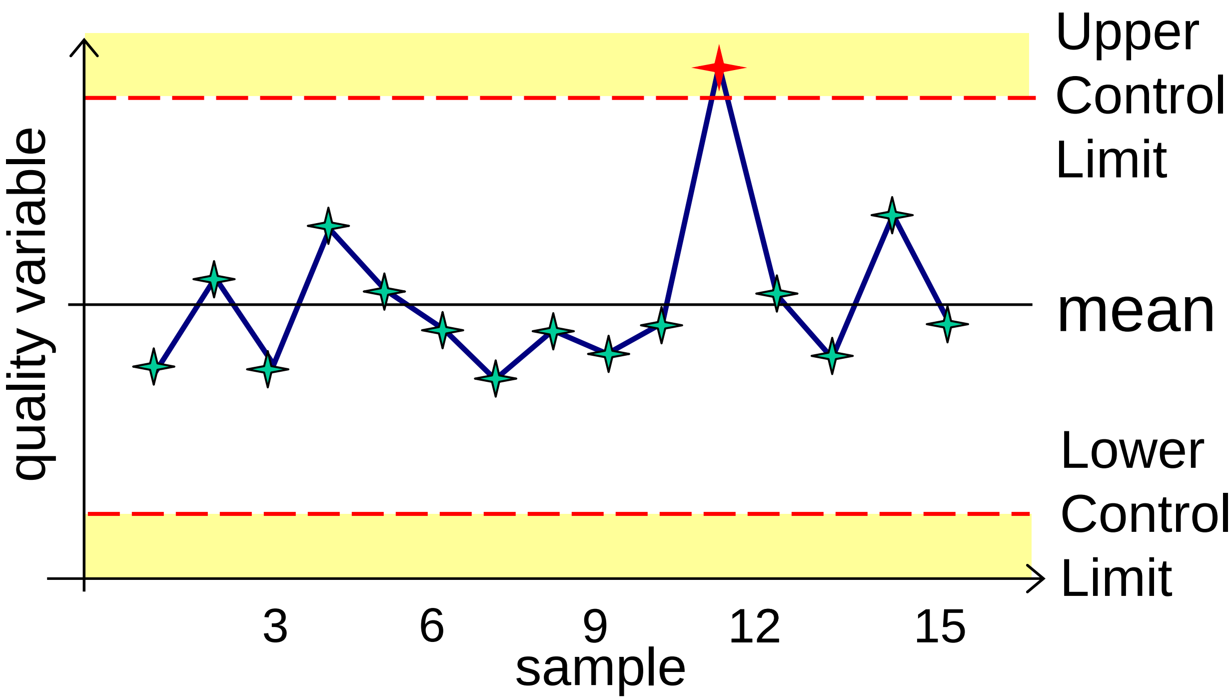
- **<u>Shewhart chart (Control chart)</u>**
- Multivariable SPC, T^2
- Model-based estimation
  - Least squares estimation
- Integrated diagnostics
  - Cascaded design

Abnormality Detection - SPC

- SPC - Statistical Process Control
  - monitoring of manufacturing processes
  - warning for off-target quality
- Main SPC method
  - Shewhart Chart (1920s)
- Also see
  - EWMA (1940s)
  - CuSum (1950s)
  - Western Electric Rules (1950s)

SPC: Shewhart Control Chart

- W.Shewhart, Bell Labs, 1924
- Statistical Process Control (SPC)
- UCL = mean + 3·σ
- LCL = mean - 3·σ



Walter Shewhart
(1891-1967)

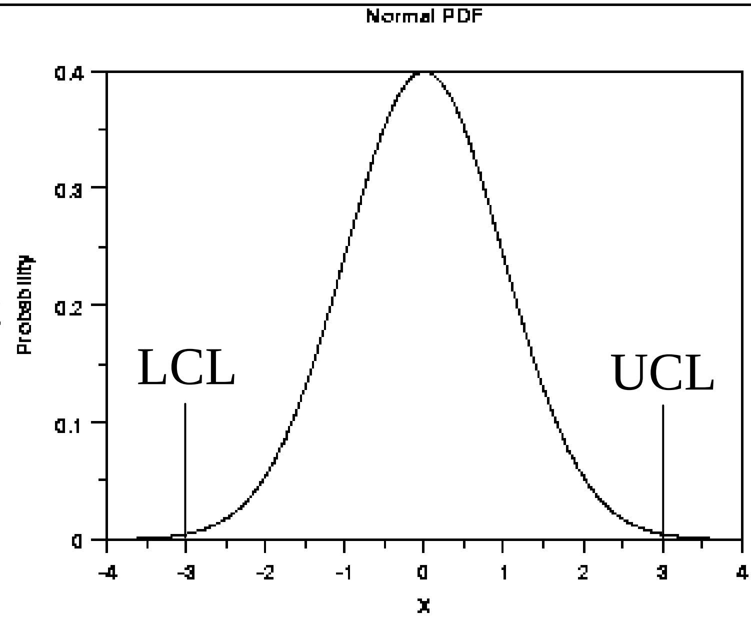
Shewhart Chart, cont'd

- Quality variable assumed randomly changing around a steady state value
- Detection: $y(t)$ > UCL=mean+3·σ
- For normal distribution, false alarm probability is less than 0.27%

$$e(t) = \frac{y(t) - \mu_0}{\sigma}$$

$P(e > 3) = 1-\Phi(3) = 0.1350 \cdot 10^{-2}$

$P(e < 3) = \Phi(-3) = 0.1350 \cdot 10^{-2}$



Shewhart Chart –Hypothesis Test

- Null hypothesis
  - given mean and covariance

$$H_0 : y(t) \sim N(\mu_0, \sigma^2) \qquad P(H_0) = p_0$$

- Fault hypothesis
  - a different mean

$$H_1 : y(t) \sim N(\mu_1 \neq \mu_0, \sigma^2) \qquad P(H_1) = p_1$$

- Hypothesis testing

given $\quad Y = y(t)$

find $\quad X \in \{H_0, H_1\}$

Bayesian Formulation

- Data: Y — Observation

  X — Underlying State

Rev. Thomas Bayes
(1702-1761)

- Bayes rule

$$P(X \mid Y) = P(Y \mid X) \cdot P(X) \cdot c$$

- Observation model: $P(Y \mid X)$
- Prior model: $P(X)$
- Maximum A posteriori Probability estimate

$$X = \arg\min \underbrace{\left(-\log P(Y \mid X) - \log P(X)\right)}_{L-\text{log-posterior index}}$$

Hypothesis Testing

- Null hypothesis

$$-\log P(Y \mid X) = \frac{1}{2\sigma^2}(y-\mu_0)^2 \qquad -\log P(X) = -\log p_0$$

- Fault hypothesis

$$-\log P(Y \mid X) = \frac{1}{2\sigma^2}(y-\mu_1)^2 = 0 \qquad -\log P(X) = -\log p_1$$

- Log-likelihood ratio

$$p_0 = 1 - p_1$$

$$\Lambda = \log \frac{P(H_1 \mid Y)}{P(H_0 \mid Y)} = L_0 - L_1$$

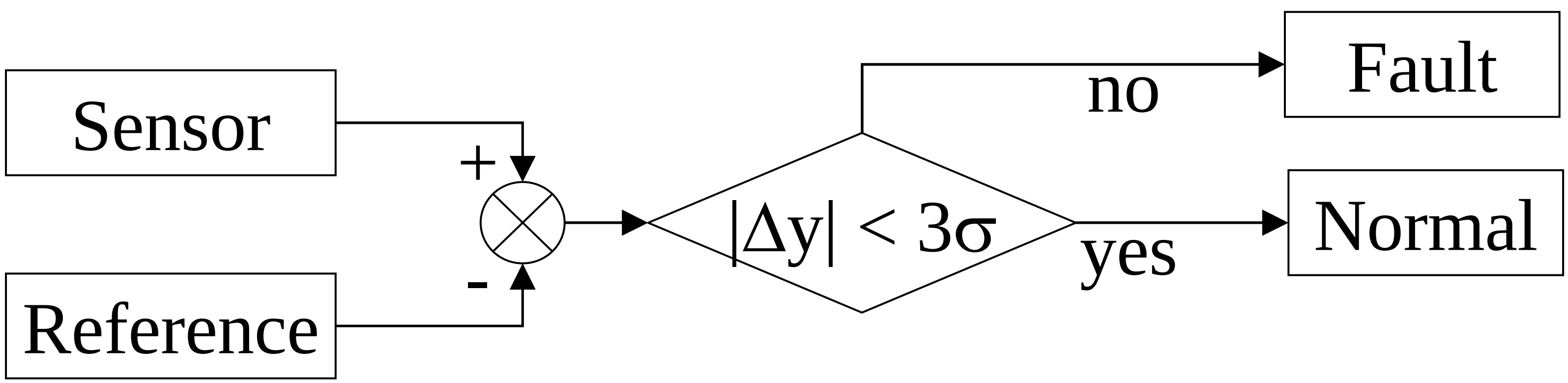
- Declare fault if

$$\Lambda = L_0 - L_1 = \frac{1}{2\sigma^2}(y-\mu_0)^2 - \log(p_0/p_1) > 0$$

$$|y-\mu_0| > \sigma\sqrt{2\log(1-p_1)/p_1} \qquad p_1 = 0.0113 \Rightarrow 3\sigma$$

Shewhart Chart: Use Examples

- SPC in manufacturing
- Fault monitoring
- Fault tolerance – sensor integrity monitoring



Diagnostics Methods Overview

- Shewhart chart (Control chart)
- **Multivariable SPC, T^2**
- Model-based estimation
  - Least squares estimation
- Integrated diagnostics
  - Cascaded design

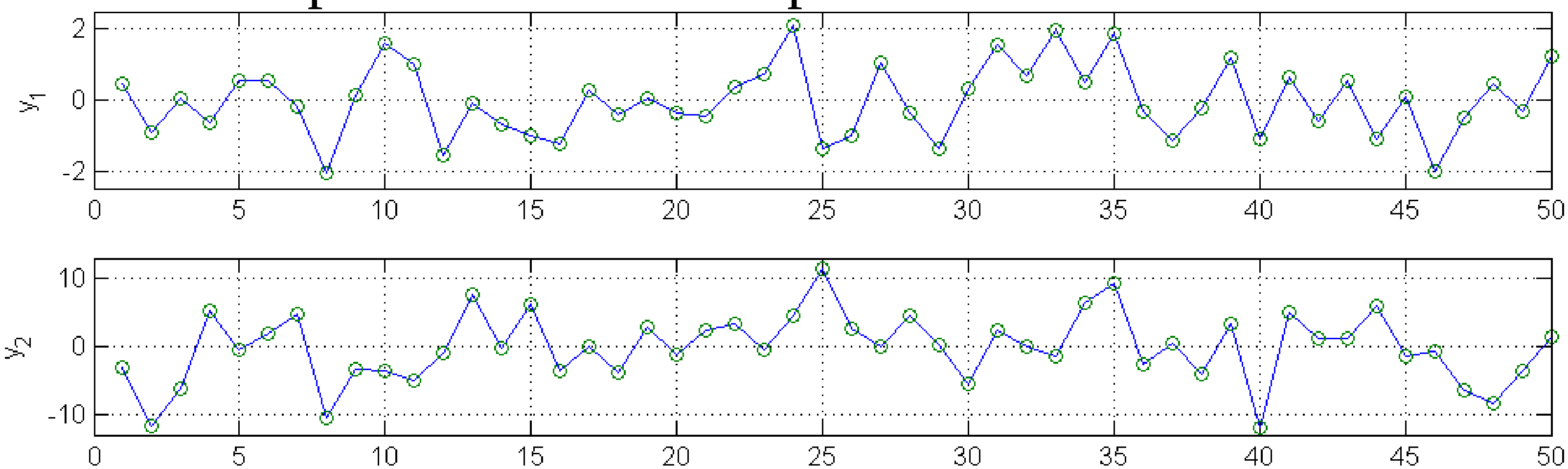
Multivariable SPC

- Univariate process

Chi-squared CDF

$$z^2 = \left(\frac{y-\mu_0}{\sigma}\right)^2 \sim \chi^2 \qquad P\left(z^2 > c^2\right) = 1 - F\left(c^2, 1\right) = \Phi(-c) + 1 - \Phi(c)$$

- Two independent univariate processes



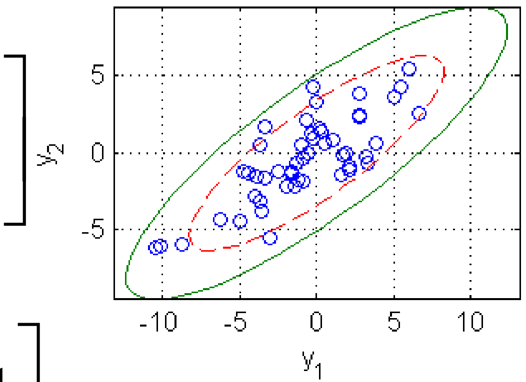
$$z^2 = \underbrace{\left(\frac{y_1-\mu_1}{\sigma_1}\right)^2}_{z_1^2} + \underbrace{\left(\frac{y_2-\mu_2}{\sigma_2}\right)^2}_{z_2^2} \sim \chi_2^2 \qquad P\left(z^2 > c^2\right) = 1 - F\left(c^2; 2\right)$$

Multivariable SPC

- Two correlated univariate processes $y_1(t)$ and $y_2(t)$

  $\mathrm{cov}(y_1,y_2) = Q, \quad Q^{-1} = L^{\mathrm{T}}L$

$$y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$



- Uncorrelated linear combinations

  $z(t) = L\cdot[y(t)-\mu]$

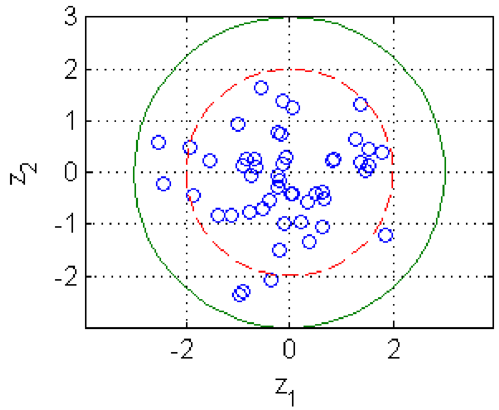
$$\|z\|^2 = (y-\mu)^T Q^{-1}(y-\mu) \sim \chi_2^2$$

$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$$

- Declare fault (anomaly) if

$$(y-\mu)^T Q^{-1}(y-\mu) > c^2$$

$$P(z^2 > c^2) = 1 - F(c^2;2)$$

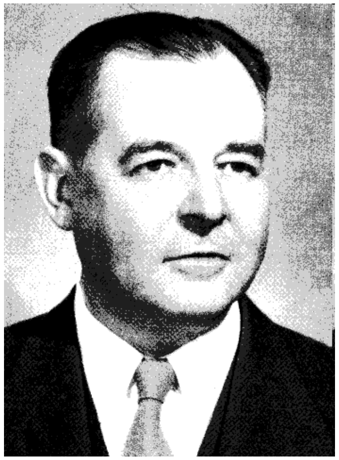


Multivariate SPC - Hotelling's T^2

- Empirical parameter estimates

$$\hat{\mu} = \frac{1}{n}\sum_{t=1}^{n} y(t) \approx E(X)$$

$$\hat{Q} = \frac{1}{n}\sum_{t=1}^{n} (y(t)-\mu)(y^T(t)-\mu^T) \approx \mathrm{cov}(y-\mu)$$



Harold Hotelling
(1895-1973)

- Hotelling's T^2 statistics is

$$T^2 = (y(t)-\mu)^T \hat{Q}^{-1} (y(t)-\mu)$$

- T^2 can be trended as a univariate SPC variable

Diagnostics Methods Overview

- Shewhart chart (Control chart)
- Multivariable SPC, T^2
- **Model-based estimation**
  - **Least squares estimation**
- Integrated diagnostics
  - Cascaded design

Least Squares Estimation

- Linear observation model:

$$Y = CX + v$$

- Fault signature model
  - Columns of C are fault signatures
  - Could be obtained from physics model
    - secant method
  - Could be identified from data:
    - regression, data mining
- Estimate
  - regularized least squares

$$\hat{X} = \left(C^T C + rI\right)^{-1} C^T Y$$

Carl Friedrich Gauss
(1777-1855)

Bayesian Estimation

- Observation model: $P(Y|X)$

$$Y = CX + v \qquad v \sim N(0, Q)$$

- Prior models: $P(X)$

$$X \sim N(0, R) \qquad -\log P(X) = \tfrac{1}{2} X^T R^{-1} X + \ldots$$

- MAP estimate:

$$\hat{X} = \arg\min \underbrace{\tfrac{1}{2}\left\| Y - CX \right\|_{Q^{-1}}^2}_{-\log P(Y|X)} + \underbrace{\tfrac{1}{2}\left\| X \right\|_{R^{-1}}^2}_{-\log P(X)}$$

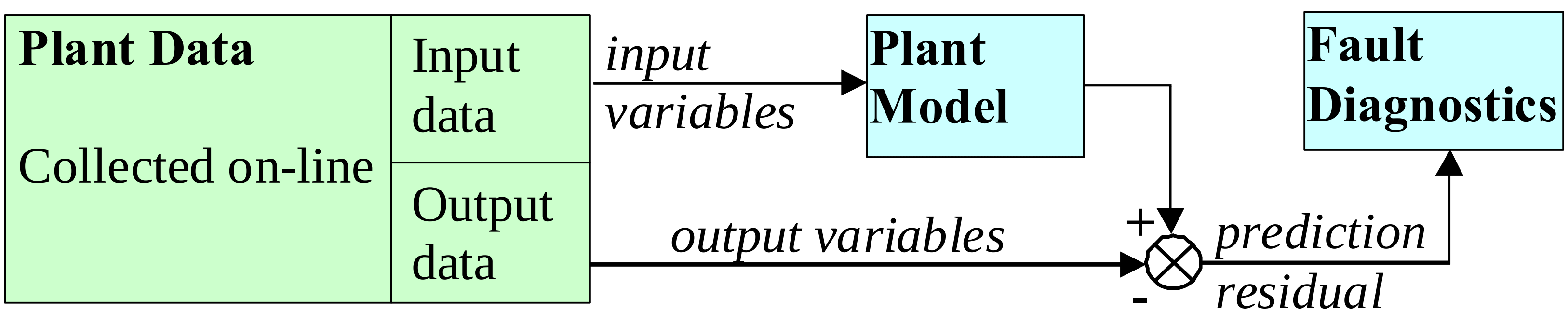
$$\hat{X} = \left(C^T Q^{-1} C + R^{-1} \right)^{-1} Q^{-1} C^T Y$$

Model-based Residuals

- Compute model-based prediction residual

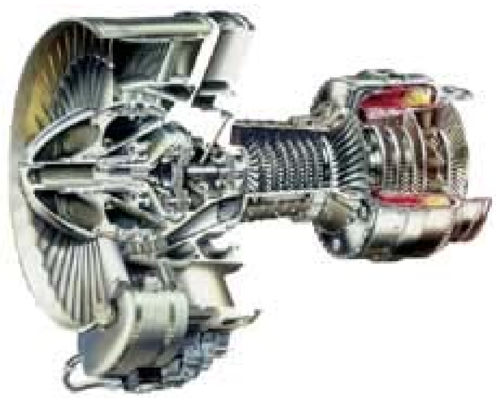
  $$Y = Y_{raw} - f(U,X)$$

- If $X = 0$ (nominal case) we should have $Y = 0$.
- Residuals Y reflect faults
  - Sensor fault model - additive output change
  - Actuator fault model - additive input change



Example: Jet Engine Model

- Nonlinear jet engine model
  - static map



- Residuals

$$Y = Y_{raw} - f(U, X)$$

- Linearized model

$$Y = CX + v$$

$$C = -\frac{\partial f(U, X)}{\partial X}$$

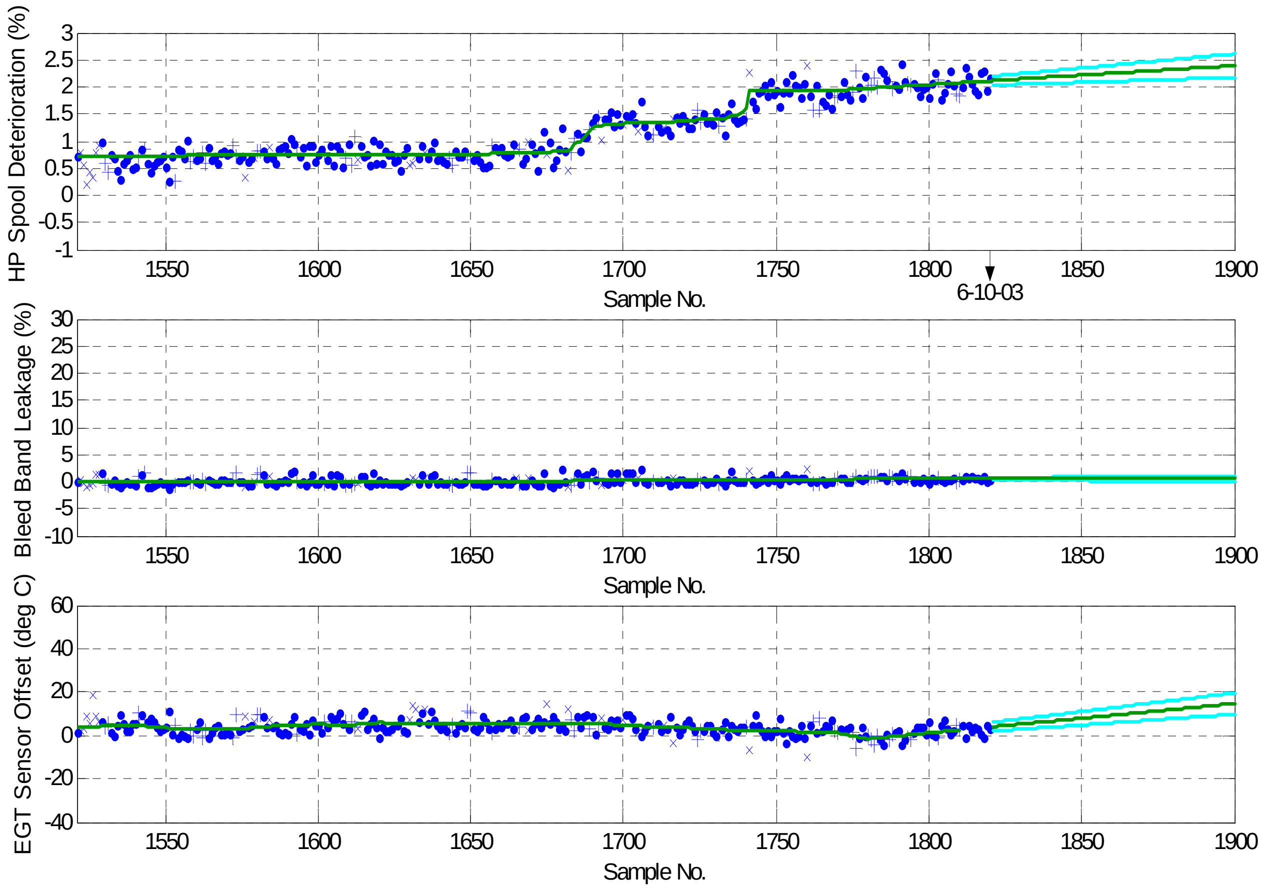
$$X = \begin{bmatrix} \text{Turbine deterioration} \\ \text{Bleed band leak} \\ \text{EGT sensor drift} \end{bmatrix}$$

Example: Fault Estimates

- Maintenance decision support tool

Honeywell LF507 Engine Fleet

Estimates of (fault) performance parameter deterioration



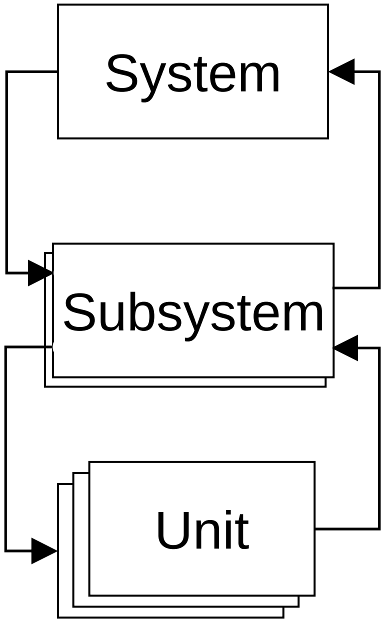
Ganguli, Deo, & Gorinevsky, IEEE CCA'04

Diagnostics Methods Overview

- Shewhart chart (Control chart)
- Multivariable SPC, T^2
- Model-based estimation
  - Least squares estimation
- **Integrated diagnostics**
  - **Cascaded design**

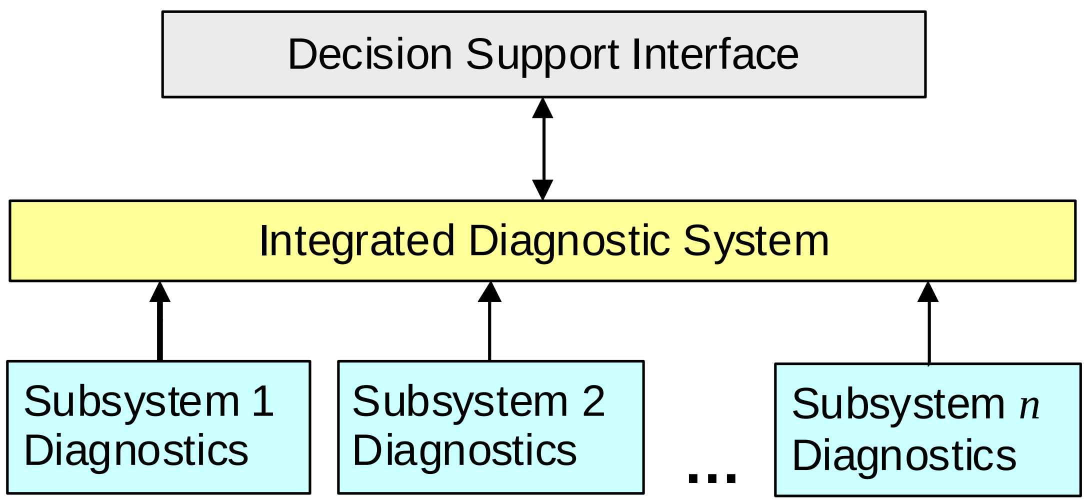
Cascaded Design

- Increasing complexity and integration of system ↑
- Slower time scale ↑
- Simple inner loop models
- Examples
  - Control systems
  - Estimation and data fusion
  - Fault diagnostics systems



Integrated System Diagnostics

- Complex integrated systems
- Examples
  - Aerospace vehicle, e.g. B777
  - Large scale computer network
  - Medical equipment



Discrete Fault Signatures

Model of root cause fault k : $Y^k = B^k$

Root Cause → ↓ Symptom Code	#0 Null	#1	#2	#3	#4	#5	#6
#1	0	0	1	1	0	0	1
#2	0	0	1	1	0	1	0
#3	0	1	0	1	0	0	1
#4	0	1	0	0	0	0	0
#5	0	1	0	0	0	1	0
#6	0	0	0	0	1	1	0
#7	0	1	1	0	1	0	0

Estimation Algorithm

- Diagnosis problem:

  Given data Y, diagnose root cause k

- Solution:

$$k = \arg\min \left\| Y - B^k \right\|_1$$

- Minimal Hamming distance
- Justifications
  - Case-based reasoning (table of fault cases)
  - Model-based reasoning (fault signature model)
  - Bayesian

Bayesian Justification

- Data

$$Y \qquad X = \left\{ H_0 : 0, \quad H_1 : B^1, \quad \ldots, \quad H_n : B^n \right\}$$

- Observation model

$$P(Y \mid X) \qquad P(y_j = b_j^k \mid H_k) = 1 - p_1 \qquad y_\mathrm{j} \text{ follows the model}$$

$$P(y_j \neq b_j^k \mid H_k) = p_1 \qquad \text{deviates from the model}$$

$$-\log P(Y \mid H_k) = \sum_{y_j = b_j^k} -\log(1 - p_1) + \sum_{y_j \neq b_j^k} -\log p_1 =$$

$$-n \cdot \log(1 - p_1) - \sum_{j=0}^{n} | y_j - b_j^k | \cdot \underbrace{\left(\log p_1 - \log(1 - p_1)\right)}_{>0,\ \text{for } p_1 < 1/2}$$

$$-\log P(Y \mid H_k) = c + w \left\| y - b^k \right\|_1$$

Bayesian Justification

- Prior model

$$P(X) \qquad P(H_k) = 1/(n+1)$$

$$-\log P(H_k) = d$$

- MAP Estimate

$$k = \arg\min\left(-\log P\left(Y \mid H_k\right) - \log P\left(H_k\right)\right)$$

$$k = \arg\min_k\left(c + w\left\| y - b^k \right\|_1 + d\right)$$

$$k = \arg\min_k \left\| y - b^k \right\|_1$$

Conclusions

- Basic diagnostics estimation methods
  - Are known for long time
  - Used in on-line systems for less time
  - Can be explained in several ways, e.g., Bayesian
- Engineering of fault diagnostics systems
  - Is new and current
  - Will be discussed in guest lectures
  - Not just diagnostics algorithms

Guest Lectures: Approaches

#	Date	Lector	From	Application	**Approach**
2	7-Apr-09	Rabover	EMC	Network	**Integrated diagnostics**
3	14-Apr-09	Tuv	Intel	IC Manufacture	**Fault signature ID**
4	21-Apr-09	Felke	Honeywell	Aircraft	**Integrated diagnostics**
5	28-Apr-09	Adibhatla	GE Infra	Jet Engines	**Multivariate estimation**
7	12-May-09	Urmanov	Sun	Computing	**Multivariate SPC, ML**
8	19-May-09	Bodden	Lockheed	Aircraft	**Fault tolerance**
9	26-May-09	Kolmanovsky	Ford	Automotive	**Model-based**