Lessons learned from the theory and practice of change detection

Michèle Basseville
IRISA / CNRS, Rennes, France
michele.basseville@irisa.fr -- http://www.irisa.fr/sisthem/

Introduction

• Detection of changes
  – Stochastic models (static, dynamic) \(\leftrightarrow\) uncertainties
  – Parameterized models (physical interpretation, diagnostics)
  – Damage \(\leftrightarrow\) deviation in the parameter vector

• Many changes of interest are small
• Early detection of (small) deviations is useful
• Key issue (1): which function of the data should be handled?
• Key issue (2): how to make the decision?

Content

Changes in the mean

Changes in the spectrum

Changes in the system dynamics and vibration-based SHM

Example: flutter monitoring
Changes in the mean

Key concepts - Independent data

Likelihood

\[ p_\theta(y_i) \]

Log-likelihood ratio

\[ s_i \triangleq \ln \frac{p_{\theta_1}(y_i)}{p_{\theta_0}(y_i)} \]

\[ E_{\theta_0}(s_i) < 0 \]

\[ E_{\theta_1}(s_i) > 0 \]

Likelihood ratio

\[ \Lambda_N \triangleq \frac{p_{\theta_1}(Y^N)}{p_{\theta_0}(Y^N)} = \prod_{i=1}^{N} \frac{p_{\theta_1}(y_i)}{p_{\theta_0}(y_i)} \]

Log-likelihood ratio

\[ S_N \triangleq \ln \Lambda_N = \sum_{i=1}^{N} s_i \]

On-line change detection, unknown onset time

\[
\begin{array}{c}
\theta_0 \\
1 \\
? \\
\theta_1 \\
n
\end{array}
\]

Hypothesis testing

Hypotheses \( H_0 \quad H_1 \)

Simple \( \theta_0 \quad \theta_1 \quad \text{Known parameter values} \)

Composite \( \Theta_0 \quad \Theta_1 \quad \text{Unknown parameter values} \)

Simple hypotheses: Likelihood ratio test

If \( \Lambda_N \geq \lambda \ \text{or} \ S_N \geq h \) : decide \( H_1 \), \( H_0 \) otherwise

Composite hypotheses: Generalized likelihood ratio (GLR)

\[ \bar{\Lambda}_N = \sup_{\theta_1 \in \Theta_1} \frac{p_{\theta_1}(Y^N)}{\sup_{\theta_0 \in \Theta_0} p_{\theta_0}(Y^N)} = \frac{p_{\theta_1}(Y^N)}{p_{\theta_0}(Y^N)} \]

Maximize the likelihoods w.r.t. unknown values of \( \theta_0 \) and \( \theta_1 \)

Simple case: Known \( \theta_1 \) - CUSUM algorithm

Ratio of likelihoods under \( H_0 \) and \( H_1 \):

\[
\frac{\prod_{i=t_0}^{t_0-1} p_{\theta_0}(y_i) \cdot \prod_{i=t_0}^{k} p_{\theta_1}(y_i)}{\prod_{i=t_0}^{k} p_{\theta_0}(y_i)} = \Lambda_{t_0}^k
\]

Maximize over the unknown onset time \( t_0 \):

\[
(t_0)_k \triangleq \arg \max_{1 \leq j \leq k} \sum_{i=1}^{j-1} p_{\theta_0}(y_i) \cdot \prod_{i=j}^{k} p_{\theta_1}(y_i)
\]

\[
= \arg \max_{1 \leq j \leq k} \Lambda_j^k
\]

\[
= \arg \max_{1 \leq j \leq k} S_j^k, \quad \text{where} \quad S_j^k = \ln \Lambda_j^k
\]

\[
g_k \triangleq \max_{1 \leq j \leq k} S_j^k = \ln \Lambda_{t_0}^k
\]

Alarm time \( t_a \) : \( t_a = \min \{ k \geq 1 : g_k \geq h \} \) decision function

Estimated onset time: \( \hat{t}_0 \)
CUSUM algorithm (Contd.)

\[ g_k = \max_{1 \leq j \leq k} S_j^k \]
\[ = S_1^k - \min_{1 \leq j \leq k} S_j^i = S_1^k - m_k, \quad m_k = \min_{1 \leq j \leq k} S_j^i \]

\[ t_a = \min \{ k \geq 1 : S_1^k \geq m_k + h \} \quad \text{Adaptative threshold} \]

\[ g_k = (g_{k-1} + s_k)^+ \]
\[ g_k = (S_{k-N_k+1}^k)^+, \quad N_k \triangleq N_{k-1} \cdot I(g_{k-1}) + 1 \]

\[ (t_0)_k = t_a - N_t + 1 \quad \text{Sliding window with adaptive size} \]

CUSUM algorithm - Gaussian example (Contd.)

\[ N(\mu, \sigma^2), \quad \theta \triangleq \mu, \quad p_\theta(y) \triangleq \frac{1}{\sigma \sqrt{2\pi}} \exp \left( - \frac{(y_i - \mu)^2}{2\sigma^2} \right) \]

\[ s_i = \ln \frac{p_{\mu_1}(y_i)}{p_{\mu_0}(y_i)} \]

\[ = \frac{1}{2\sigma^2} \left( (y_i - \mu_0)^2 - (y_i - \mu_1)^2 \right) \]

\[ = \frac{\nu}{\sigma^2} \left( y_i - \mu_0 - \frac{\nu}{2} \right), \quad \nu = \mu_1 - \mu_0 \]

\[ S_1^k \text{ involves } \sum_{i=1}^k y_i : \quad \text{Integrator (with adaptive threshold)} \]

Composite case: Unknown \( \theta_1 \)

Modified CUSUM algorithms

Minimum magnitude of change

Weighted CUSUM

GLR algorithm

Double maximization

\[ g_k = \max_{1 \leq j \leq k} \sup_{\theta_1} S_j^k(\theta_1) \]

Gaussian case, additive faults: second maximization explicit.
Unknown $\theta_1$ - Gaussian example (Contd.)

Minimum magnitude of change $\nu_m$

Lesson 2

**Decreasing** mean

$T_k^1 \triangleq \sum_{i=1}^k (y_i - \mu_0 + \frac{\nu_m}{2})$

**Increasing** mean

$U_k^1 \triangleq \sum_{i=1}^k (y_i - \mu_0 - \frac{\nu_m}{2})$

$M_k \triangleq \max_{1 \leq j \leq k} T_j^1$

$m_k \triangleq \min_{1 \leq j \leq k} U_j^1$

$t_a = \min \{ k \geq 1 : M_k - T_k^1 \geq h \}$

$t_a = \min \{ k \geq 1 : U_k^1 - m_k \geq h \}$

Key concepts - Dependent data (Contd.)

**Which residuals?**

Lesson 3

- **Likelihood ratio** computationally complex

- **Efficient score**: likelihood sensitivity w.r.t. parameter vector

- Other estimating functions

Innovation not sufficient for monitoring the dynamics.

Key concepts - Dependent data (Contd.)

**Changes in the spectrum**

**Conditional likelihood** $p_{\theta}(y_i | Y_1^{i-1})$

**Log-likelihood ratio**

$s_i \triangleq \ln \frac{p_{\theta_1}(y_i | Y_1^{i-1})}{p_{\theta_0}(y_i | Y_1^{i-1})}$

$E_{\theta_0}(s_i) < 0$

$E_{\theta_1}(s_i) > 0$

**Likelihood ratio**

$\Lambda_N \triangleq \frac{p_{\theta_1}(Y_N^1)}{p_{\theta_0}(Y_N^1)} = \prod_{i} \frac{p_{\theta_1}(y_i | Y_1^{i-1})}{p_{\theta_0}(y_i | Y_1^{i-1})}$

**Log-likelihood ratio**

$S_N \triangleq \ln \Lambda_N = \sum_{i=1}^N s_i$

Key concepts - Dependent data (Contd.)

$\theta_0$: reference parameter, known (or identified)

$Y_k$: $N$-size sample of new measurements

**Build a residual $\zeta$ significantly non zero** when damage

Residual $\leftrightarrow$ **Estimating function** $\zeta_N(\theta, Y_1^N)$

Characterized by: $E_{\theta_0} \zeta_N(\theta, Y_1^N) = 0 \iff \theta = \theta_0$

Mean sensitivity $J(\theta_0)$ and covariance $\Sigma(\theta_0)$ of $\zeta_N(\theta_0)$
The residual is asymptotically Gaussian

\[
\zeta_N(\theta_0) \to \begin{cases} 
\mathcal{N}(0, \Sigma(\theta_0)) & \text{if } P_{\theta_0} \\
\mathcal{N}(\mathcal{J}(\theta_0) \delta\theta, \Sigma(\theta_0)) & \text{if } P_{\theta_0 + \delta\theta \sqrt{N}} \text{ small deviations}
\end{cases}
\]

(On-board) $\chi^2$-test in the residual

\[
\zeta_N^T \Sigma^{-1} \mathcal{J} (\mathcal{J}^T \Sigma^{-1} \mathcal{J})^{-1} \mathcal{J}^T \Sigma^{-1} \zeta_N \geq h
\]

Noises and uncertainty on $\theta_0$ taken into account.

Changes in the dynamics and vibration-based SHM

**Structural monitoring: Eigenstructure monitoring**

\[
\begin{align*}
X_{k+1} &= F X_k + V_k \\
Y_k &= H X_k
\end{align*}
\]

Canonical parameter: $\theta \triangleq \begin{pmatrix} \Lambda \\ \text{vec} \Phi \end{pmatrix}$ modes mode shapes

Example: flutter monitoring

Monitoring a damping coefficient

- Test $\rho \geq \rho_c$ against $\rho < \rho_c$
- Write the subspace-based residual $\zeta$ as a cumulative sum
- Introduce a minimum change magnitude (actual change magnitude unknown)
- Run two CUSUM tests in parallel (actual change direction unknown)
Flutter monitoring (Contd)

Application to real datasets

Ariane booster launcher during a launch scenario on the ground (not during a real flight test)

Running the test with 2 critical values

With Laurent Mével, Maurice Goursat, Albert Benveniste

Free COSMAD Toolbox to used with Scilab

http://www.irisa.fr/sisthem/cosmad/

Automatic identification. Each symbol: processing 5 sec. data.

Conclusion

Advanced statistical signal processing mandatory for SHM

A statistical framework enlightens the meaning and increases the power of a number of familiar operations integration, averaging, sensitivity, adaptive thresholds & windows

Change detection useful for (vibration-based) SHM

Current investigations

- flutter monitoring (with Dassault, Airbus),
- handling the temperature effect on civil structures (with LCPC)