03_Change_Detection

1. Introduction

For change detection, we mainly focus on how to detect whether the system behaving in a nominal way.

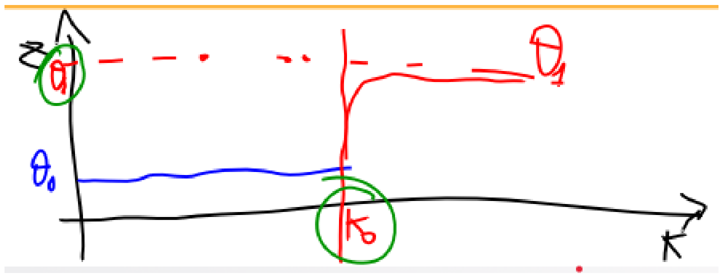
A symptoms generator + a change detection algorithm = a diagnosis method

Two Misclassification Scenarios

- False Positive: False Alarm Rate (FAR)
- False Negative: Missed Detection Rate (MDR)

Overview

Detect Change, estimate k_0 and θ_1



2. Deterministic Tests

Limit Check: Scalar Version

Limit checking simply verify the variable is inside this static, deterministic range

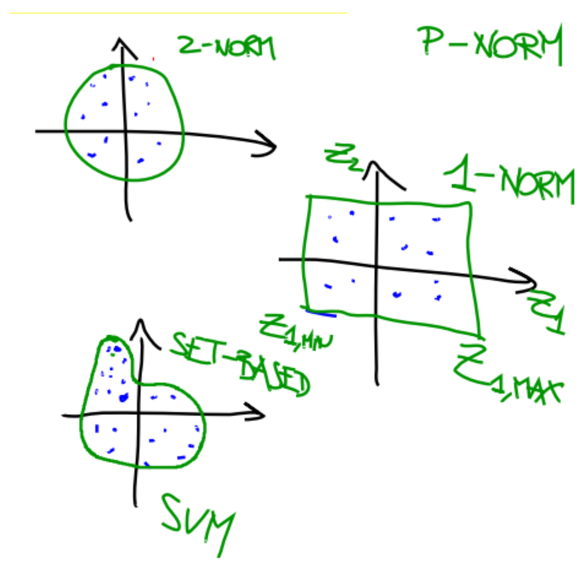
$$z_{min} \leq z \leq z_{max}$$

Problem:

z maybe a **vector** variable, or we want to check behavior over a **time window**

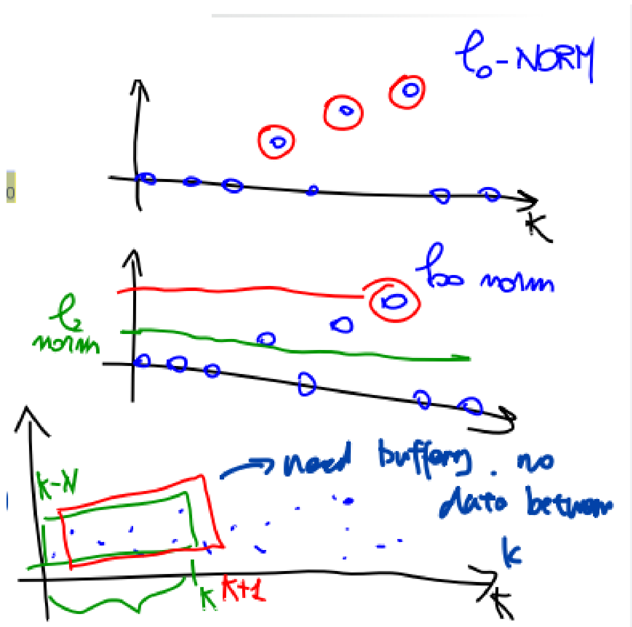
Limit Check: Vector

- norm-based limit check
- component-wise limit check
- set-based limit check: like SVM



Limit Check: Time Window

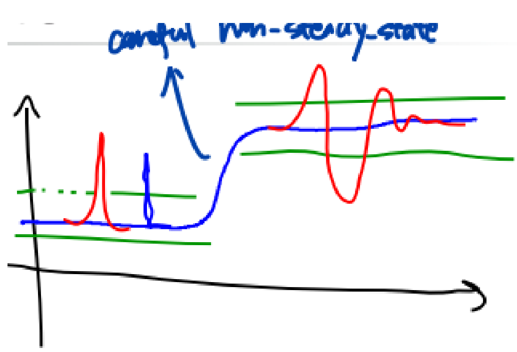
- number of non-zero samples (l_0 norm)
- peak value (l_∞ norm)
- average
- Manhattan Norm (l_1 norm)
- RMS (l_2 norm)



Limit Check: Improvements

Dynamic Check:

- relative on signal amplitude
- based on models



3. Probabilistic Test

Express nominal conditions in terms of **statistical moments, or of the Probability Density Function** (pdf) of z (always assume Gaussian)

Recursive Calculation of Mean and Variance

- **mean**: $\hat{\mu}(k) = \hat{\mu}(k-1) + \frac{1}{k}[z(k) - \hat{\mu}(k-1)]$
- **variance:** $\hat{\sigma}^2(k) = \frac{k-2}{k-1}\hat{\sigma}^2(k-1) + \frac{1}{k}[z(k) - \hat{\mu}(k-1)]^2$

Detection use t test (for test μ)

<u>**Case 1:**</u>

- known μ,
- $\hat{\mu}, \hat{\sigma}$ estimated from N samples

$$t(N-1) = \frac{\hat{\mu}(N) - \mu}{\hat{\sigma}/\sqrt{N}}$$

<u>**Case 2:**</u>

- unknown μ, unknown but constant σ;
- estimation from N_0 samples before the (hypothetical) change and N_1 after

$$t(N_0 + N_1 - 1) = \frac{\hat{\mu}_0 - \hat{\mu}_1}{\sqrt{(N_0 - 1)\,\hat{\sigma}_0^2 + (N_1 - 1)\,\hat{\sigma}_1^2}}\sqrt{\frac{N_0 N_1\,(N_0 + N_1 - 2)}{N_0 + N_1}}$$

Detection use χ test (for test σ)

<u>**Case 3:**</u>

- known σ_0,
- $\hat{\sigma}$ estimated from N samples

$$\chi^2(N-1) = \frac{(N-1)\hat{\sigma}_1^2}{\sigma_0^2}$$

Detection use F test (for test σ)

<u>**Case 4:**</u>

- σ estimated from N_0 samples before change
- From N_1 samples after change, mean is unknown and can even vary

$$F\left(N_0-1, N_1-1\right)=\frac{\hat{\sigma}_0^2}{\hat{\sigma}_1^2}$$

Pros and Cons

- To verify/falsify the null hypothesis with **good significance**, you need **a high number of samples** before and after the (hypothetical) change
- Makes **Detection Delayed**

4. Advanced Probabilistic Tests

Multivariate Case

Assumption

- Assume z is a vector
- nominal mean μ and covariance matrix C are known

Mahalanobis Distance

It is a weighted distance between observation and known mean based on the known covariance

$$D_M(\vec{x})=\sqrt{(\vec{x}-\vec{\mu})^{\top} \mathbf{C}^{-1}(\vec{x}-\vec{\mu})}.$$

Theorem: Multi-Dimensional Chebyshev Inequality

In probability theory, the **multidimensional Chebyshev's inequality** is a generalization of Chebyshev's inequality, which puts a bound on the probability of the event that a random variable differs from its expected value by more than a specified amount.

$$\operatorname{Pr}\left(\sqrt{(X-\mu)^T C^{-1}(X-\mu)}>t\right) \leq \frac{N}{t^2}$$

Change Detection using Mahalanobis Distance and Multi-Dimensional Chebyshev Inequality

By using Chebyshev Inequality, we can generate a **(conservative)** boundary based on our desired FAR α

$$\bar{d}=\frac{n}{\alpha} \quad \Rightarrow \quad \mathbb{P}\{d^2(z) \geq \bar{d}\} \leq \alpha$$

where n is the dimension of z

Log Likelihood Ratio: Neyman-Pearson's Approach

Assumption

- Assumed that no prior knowledge is available
- System to be validate:
  - $\mathcal{H}_0 \colon p(z) = p_{\theta_0}(z)$
  - $\mathcal{H}_1 \colon p(z) = p_{\theta_1}(z)$

Log-Likelihood

Since Probability Distributions are often assumed to be **Gaussian,** the logarithm of this probability ratio gives very convenient calculations, given a **single observation** z

$$s(z) = \ln \frac{p_{\theta_1}(z)}{p_{\theta_0}(z)}$$

Property

The log-likelihood ratio has the following fundamental statistical property

$$E_{\theta_0}(s) = \int_{-\infty}^{\infty} s(z) p_{\theta_0}(z) \mathrm{d}z < 0$$
$$E_{\theta_1}(s) = \int_{-\infty}^{\infty} s(z) p_{\theta_1}(z) \mathrm{d}z > 0$$

E_{θ_i} denotes expectation of $s(z)$ under the distribution associated to $p_{\theta_i}(z)$

CUSUM

Assumption

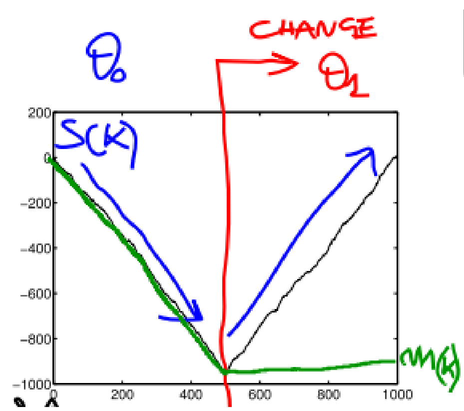
- Assume θ_0 (that is μ_0 and σ_0) is known
- We need to tune θ_1

Model

- **Cumulative Sum:**

$$S(k) = \sum_{i=1}^{k} s(z(i)) = \sum_{i=1}^{k} \ln \frac{p_{\theta_1}(z(i))}{p_{\theta_0}(z(i))}$$

S is expected to exhibit a **negative drift before** the change, and **positive after**



- **Decision Function** g

$$g(k) = S(k) - m(k)$$

  $m(k)$ is the minimum of S until time index k

- **Threshold**
  - A specific threshold is needed because **variance and noise** may lead to some positive drift even before the change
  - null hypothesis falsified for $g > h$, with h suitable

Parameter Tuning

Assume σ not change, μ_1 and h are still need to be tuned.

- for μ_1
  - replace by **minimum change** you want to detect
  - **estimate** it from data (may cause **detection delay**)
- tuning h
  - trade-off among detection time, FAR and MDR
  - know what reasonably is the **slope of** S after a change, h = desired detection time * slope

Generalized Likelihood Ratio (GLR) test

GLR aims at estimating both the **post-change parameter** θ_1 and the **change time** k_0 based on maximum likelihood thought

$$S_j^k(\theta_1) = \sum_{i=j}^{k} \ln \frac{p_{\theta_1}(z(i))}{p_{\theta_0}(z(i))}$$
$$\left(\hat{k}_0, \hat{\theta}_1\right) = \arg\left\{\max_{1 \le j \le k} \max_{\theta_1} S_j^k(\theta_1)\right\}$$
$$g(k) = \max_{1 \le j \le k} \max_{\theta_1} S_j^k(\theta_1)$$

During implementation, a maximum history buffer M can be used to reduce the complexity.

5. Summary

- change detection: A symptoms generator + a change detection algorithm = a diagnosis method
  - MAR and FAR
- Deterministic Method: Limit check
  - Scalar
  - Evaluation Function and Time window
  - Dynamic Check
- Probabilistic Method:
  - t, F, χ tests
- Advance Probabilistic Method:
  - Multivariate: MD and Chebychev Inequality
  - Log-likelihood and property: opposite sign
  - CUSUM: known θ_0, tune h and θ_1, can be easily used in multi-variable cases
    - $g(k), S(k), m(k), h$
  - GLR: known θ_0, estimate k_0 and θ_1, can be easily used in multi-variable cases
    - $g(k)$, multiple $S(k), m(k), h$