Using Interpoint Distances to Develop a New Multivariate Control Chart Based on Change-Point Detection

Claudio Giovanni Borroni[1], Manuela Cazzaro[1], Paola Maddalena Chiodini[1]
[1] University of Milano - Bicocca (Italy)

ABSTRACT

The change-point paradigm has been successfully applied to statistical process control by building many parametric and nonparametric charts for sequential monitoring. We concentrate here on a novel implementation of the change-point paradigm, based on dynamic windows, which forces comparisons only between samples of the same size. Despite that implementation has been successfully applied to univariate control, some applications often need to consider more than a single variable at a time. Indeed, not only different aspects of quality must be considered, but their dependence structure can also provide relevant information. Thus, in this talk, we investigate the application of the dynamic-window approach to multivariate control variables. Specifically, we propose a preliminary reduction of the dimension of such variables, before the change-point methodology is applied. We evaluate some techniques of reduction based on inter-point distances. A comparison study is conducted to: i) identify techniques which are stable with respect to the actual dimension and shape of the underlying distribution; ii) identify techniques which can provide a fast signal when the underlying distribution undergoes shifts in location or scale. We focus on multivariate charts which perform averagely well across a wide number of cases, more than on those which are extremely powerful just in isolated cases.

Keywords: Multivariate statistical process control, Change-point detection, interpoint distances.

Back to Sessions List

Some Change-Point Design-Based Distribution-free Approaches for Monitoring High-Dimensional Data

Amitava Mukherjee

Production, Operations and Decision Sciences Area, XLRI – Xavier School of Management, India.

ABSTRACT

This paper introduces some purely distribution-free schemes for monitoring high-dimensional data streams using a change point design setup, which combines the advantages of exponentially weighted moving averages. Proposed procedures are based on some distance measures and ranks. Our proposed procedures require collecting only a few reference samples from the in-control process at the outset of Phase II monitoring, and not many Phase I observations, unlike traditional Phase II distribution-free schemes for high-dimensional data. A regression-based approach is considered for determining the control limits. An industrial application in monitoring a process involving semiconductor quality is discussed. We also discuss comparative detection performance using Monte Carlo methods. The paper concludes with some remarks and recommendations for future research.

Keywords: Change Point Design; Distribution-free; High-dimensional; Process Monitoring

Large-Scale Decentralized Fault Diagnosis for Multi-Group Data with Auxiliary Information via Distributed Multiple Testing

Zhihan Zhang[1], Wendong Li[1], Fugee Tsung[2], Dongdong Xiang[1,*]

[1]*KLATASDS-MOE, School of Statistics, East China Normal University*

[2]*Department of Industrial Engineering and Decision Analytics, Hong Kong University of Science and Technology*

ABSTRACT

In the era of big data, efficiently diagnosing faults in high-dimensional data streams (HDS) is critical for numerous industrial applications. This paper addresses a novel decentralized fault diagnosis problem involving multi-group HDS with multi-sequence auxiliary information (MAI). Traditional diagnostic methods, which are designed for a single group of single-sequence HDS, struggle with the complexity and volume of such data, often leading to suboptimal diagnostic performance. To overcome this challenge, we propose a distributed fault diagnosis framework that leverages advanced multiple testing techniques and data fusion strategies to analyze multi-group HDS with MAI. Under this framework, we introduce a generalized multi-sequence local index of significance for the data streams in each group, based on a Cartesian hidden Markov model, to effectively fuse information from auxiliary sequences. This is then integrated into a distributed multiple testing procedure for group-wise diagnosis of the target data sequence. The proposed procedure minimizes the group-wise expected number of false positives in the target sequence while controlling the overall group-wise missed discovery rate at a specified level. Numerical studies demonstrate that the proposed method outperforms state-of-the-art diagnostic techniques, providing more reliable and effective fault diagnostics.

Keywords: Auxiliary Information; Distributed Computing; Fault Diagnosis; Multi-Group Data; Multiple Testing

Control Chart for High-Dimensional Dynamic Process Monitoring

Peihua Qiu

University of Florida

ABSTRACT

In air pollution surveillance, additive manufacturing, and other fields, monitoring high-dimensional data streams presents unique challenges. The in-control (IC) distributions often change over time due to seasonality and other factors, making it difficult to apply standard statistical process control (SPC) charts. Traditional SPC methods assume that IC process observations at different times are independent and identically distributed—an assumption that is often invalid in these settings. In this talk, we present a novel process monitoring method that integrates principal component analysis with sequential learning. This approach effectively handles high dimensionality, time-varying IC distributions, serial data correlation, and nonparametric data distributions. It has proven to be a reliable analytic tool for the online monitoring of high-dimensional dynamic processes. This is joint research with Dr. Xiulin Xie.

Keywords: Control charts; Correlation; Dynamic processes; Online monitoring