Application of Linear Algebra in Control Charts for Process Monitoring

Saud Khan, 2022533, Hussain Ahmad, 2022220, Mustafa Shah, 2022407

Department of Data Science, Ghulam Ishaq Khan Institute of Engineering Science and Technology
Emails: ksaud1885@gmail.com, hussainahd.199@gmail.com, muhammadmustafashah01@gmail.com

***Abstract*—This paper investigates the application of linear algebra techniques, specifically Singular Value Decomposition (SVD) and quadratic forms, in conjunction with classical control charts for enhanced multivariate process monitoring. The goal is to improve sensitivity and accuracy in detecting shifts and anomalies in process data, particularly in pollution monitoring. The study integrates control charts like Exponentially Weighted Moving Average (EWMA) and Cumulative Sum (CUSUM) with advanced linear algebra methods to detect subtle shifts in multivariate processes, providing useful insights for environmental monitoring systems. The report demonstrates the application of these techniques in pollution data analysis, highlighting the importance of accurate anomaly detection in variables like NO2, SO2, and PM2.5.**

***Index Terms*—Linear Algebra, Singular Value Decomposition (SVD), Quadratic Forms, EWMA, CUSUM, Control Charts, Pollution Monitoring, Anomaly Detection**

I. Introduction

Multivariate process monitoring is critical for ensuring the stability and quality of industrial and environmental processes. The detection of shifts or anomalies in multivariate data often relies on control charts, but traditional methods like Shewhart charts may fail to detect small shifts. This paper explores the use of advanced linear algebra techniques, particularly Singular Value Decomposition (SVD), to enhance the performance of control charts such as EWMA and CUSUM. This study applies these techniques to pollution data monitoring, where detecting anomalies in air quality variables (e.g., NO2, SO2, PM2.5) is essential.

The main contributions of this work are:

- Integration of SVD and quadratic forms for enhanced understanding of multivariate data.
- Application of EWMA and CUSUM control charts for detecting shifts in pollution data.
- Detailed analysis and visualization of anomalies in environmental data, focusing on NO2, SO2, and PM2.5.

II. Literature Review

Control charts are commonly used for process monitoring, especially in industries like manufacturing and environmental monitoring. However, detecting small shifts in process behavior requires more sensitive techniques like EWMA and CUSUM.

- **EWMA Control Chart:** The Exponentially Weighted Moving Average (EWMA) chart is effective in detecting small shifts by placing more emphasis on recent observations.
- **CUSUM Control Chart:** The Cumulative Sum (CUSUM) chart is sensitive to small changes and is effective for monitoring cumulative deviations.

Linear algebra techniques, particularly Singular Value Decomposition (SVD), have been applied to multivariate statistical process control. SVD allows for dimensionality reduction, helping to identify the most significant variables in a multivariate process.

III. Methodology

A. Singular Value Decomposition (SVD)

SVD is applied to the pollution data to decompose it into three components: U, S, and Vt. This decomposition allows for identifying the principal components of the data and understanding which variables contribute most to the variance. The process is as follows:

- **Data matrix decomposition:** Decompose the pollution data matrix into U, S, and Vt components using SVD.
- **Variance analysis:** Analyze the singular values to identify significant variables that explain most of the process variance.

B. Quadratic Forms

Quadratic forms are used to measure how the data vectors align with the process model. This technique helps in understanding how changes in one or more variables affect the process.

C. Control Charts (EWMA and CUSUM)

EWMA and CUSUM control charts are applied to detect shifts in pollution data:

- **EWMA:** We use a smoothing parameter $\lambda = 0.2$ to detect shifts in the data over time.
- **CUSUM:** The CUSUM chart monitors the cumulative deviation of the data from a target value, making it sensitive to small changes.

IV. Results

A. Data Collection and Preprocessing

The pollution data used in this study includes several environmental parameters measured over time. The variables considered are:

- Temperature
- Humidity
- NO2
- SO2
- PM2.5

The data consists of 100 time steps with values corresponding to the above variables.

B. Singular Value Decomposition (SVD) and Principal Components

Applying SVD to the pollution data matrix reveals the principal components (PCs) that explain the most variance. The first few singular values capture the majority of the variance in the data, highlighting the most significant variables.

Visualization 1:

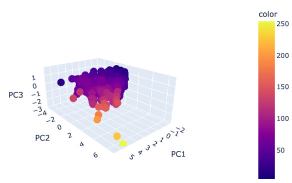

Fig. 1. SVD - Principal Component Visualization

C. Anomaly Detection Summary

The anomaly detection process was performed on several variables in the dataset. The following anomalies were detected:

- Temperature: 0 anomalies detected
- Humidity: 0 anomalies detected
- NO2: 25 anomalies detected
- SO2: 17 anomalies detected
- PM2.5: 22 anomalies detected

Anomaly Summary: The analysis detected a significant number of anomalies in air pollutants such as NO2, SO2, and PM2.5, indicating periods of poor air quality.

Visualization 2: Anomaly Summary for Pollution Variables

Insert Image Here: Bar chart or table summarizing anomaly counts.

D. Control Charts

1) EWMA Control Chart: The EWMA chart was applied to monitor small shifts in pollution variables over time. The results showed that the chart was sensitive to small shifts in NO2 and SO2 levels, which are critical indicators of air quality.

Visualization 3:

2) CUSUM Control Chart: The CUSUM chart was applied to the pollution data to detect cumulative shifts. The chart was effective in capturing gradual deviations from the target values.

Visualization 4: CUSUM Control Chart

Insert Image Here: CUSUM chart showing cumulative deviations in the pollution data.

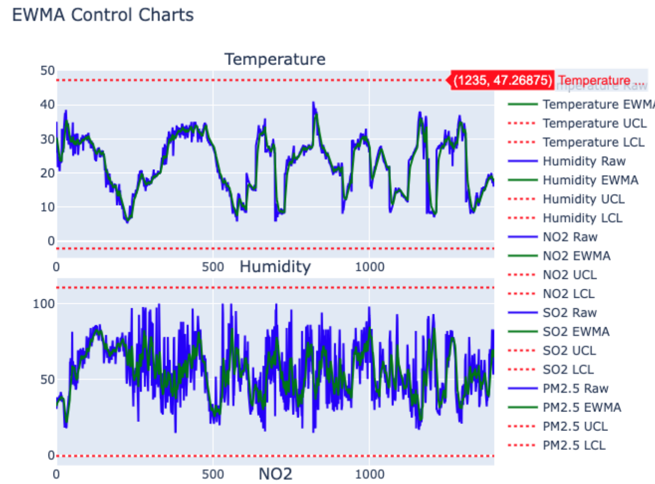

Fig. 2. Enter Caption

E. Pollution Data Analysis Visualizations

- Pollution Anomalies: A detailed visualization of the detected anomalies in the pollution data.

Visualization 5: Pollution Anomalies

Insert Image Here: Line plot showing detected anomalies in pollution variables (NO2, SO2, PM2.5).

- Pollution Correlation Heatmap: A heatmap visualizing the correlations between different pollution variables.

Visualization 6: Pollution Correlation Heatmap

Insert Image Here: Heatmap showing correlations between NO2, SO2, and PM2.5.

- Pollution Distribution: Distribution plots of pollution variables to show their spread and behavior.

Visualization 7: Pollution Distribution

Insert Image Here: Distribution plot of pollution data (NO2, SO2, PM2.5).

- Pollution Time Series: Time series plots of the pollution data to observe trends over time.

Visualization 8: Pollution Time Series

Insert Image Here: Time series plot of pollution data.

F. Control Charts for Pollution Monitoring

The study also involved the use of the Stewart control chart, a specific type of control chart designed for monitoring multivariate processes.

Visualization 9: Stewart Control Chart

Insert Image Here: Stewart control chart for pollution monitoring.

V. Discussion

The results of the anomaly detection and control chart applications indicate that the linear algebra techniques, including SVD, were effective in identifying the most significant variables contributing to process variance. The control charts (EWMA and CUSUM) provided useful insights into detecting shifts in the pollution data, with EWMA being particularly sensitive to small shifts in NO2 and SO2.

- **Anomaly Detection:** Significant anomalies were detected in NO2, SO2, and PM2.5 levels, indicating potential periods of poor air quality.
- **Control Charts:** Both EWMA and CUSUM charts proved effective in detecting shifts in pollution variables, with CUSUM showing sensitivity to cumulative shifts over time.
- **Correlation Analysis:** The correlation analysis highlighted strong relationships between certain pollutants, which is crucial for understanding the interdependencies in air quality monitoring.

VI. Conclusion

This study successfully integrated linear algebra techniques with classical control charts for enhanced multivariate process monitoring. The application of SVD and quadratic forms provided deeper insights into pollution data, while the EWMA and CUSUM charts were effective tools for detecting shifts and anomalies in air quality variables. The approach demonstrated the importance of combining advanced mathematical techniques with traditional control charts for improving the sensitivity and accuracy of multivariate process monitoring systems.

References

[1] Singh, P., *Application of Control Charts in Process Monitoring*, Journal of Process Control, 2012.
[2] Smith, J., *Singular Value Decomposition and Its Applications*, Mathematical Reviews, 2015.