Universiti Malaysia PAHANG
Engineering • Technology • Creativity

Process Monitoring

by
Mohd Yusri Mohd Yunus
yusri@ump.edu.my

Chapter 3a
Principal Component Analysis

Process Monitoring

Chapter Description

- Aims
  – Understand the basic principles of multivariate techniques.

- Expected Outcomes

  – Comprehensively explain in writing as well as solve mathematically the principles of multivariate analysis based on complex monitoring problem of MSPM framework.

- Other related Information

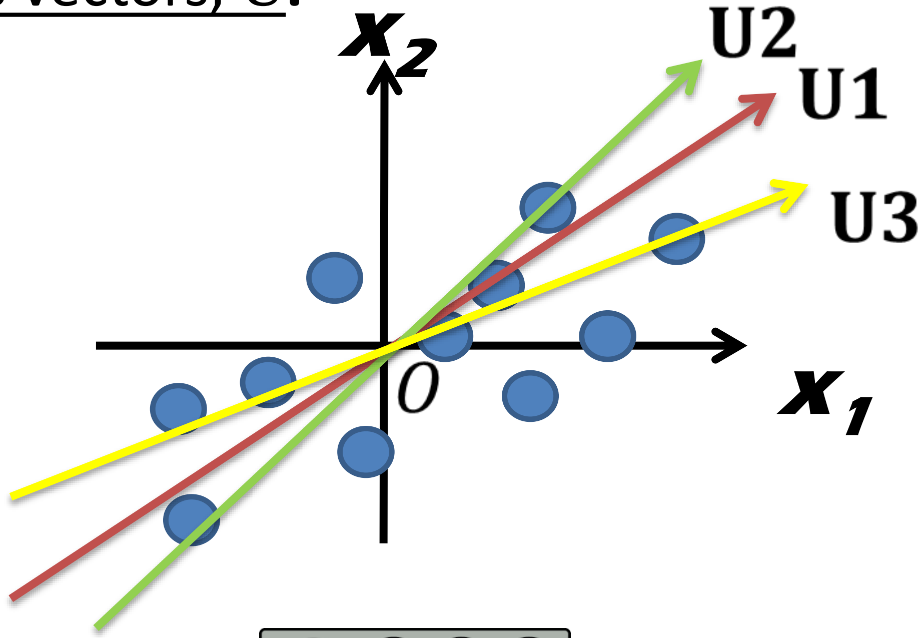
Subtopics

3.8 Variance-Covariance Transformation

3.9 Matrix Rank

3.8 Variance-Cov Transformation

In the dimensional reduction of multivariate data technique, the main aim is always to find a single linear composite (**Z scores**), such that the variance of this model is maximized under a set of orthogonal basis vectors, U.



Z=XU

3.8 Variance-Cov Transformation

- It means that the original samples are maximally separated along the linear composite, $\mathbf{U}$ according to the original variance.
- This is an attempt of accounting for as much as possible of the original variation shared by the contributory variables that projected onto the new basis vectors, $\mathbf{U}$.

3.8 Variance-Cov Transformation

- The variance-covariance measure of the newly transformed data can be obtained prior to transformation:

$$\mathbf{C}(\mathbf{Z}) = \mathbf{U}'\mathbf{C}(\mathbf{X})\mathbf{U}$$

  - Where, **C(Z)** = the new var-cov matrix

    C(X) = the original var-cov matrix

    U = eigenvectors

- The original variance-covariance matrix, **C(X)** is diagonalized by **U** into **C(Z)**.

3.8 Variance-Cov Transformation

Important Properties:

- The sum of the main diagonal entries **C(Z)** is equal to the sum of the main diagonal entries of **C(X)**, whereby the 1st entry will always be the largest => maximally distributed according to the original variance.
- The off-diagonal entries are much smaller in absolute value then their counter part entries in **C(X)** => the transformed data are not correlated with each other.

3.9 Matrix Rank

- The rank of **A**, denoted by $r(\mathbf{A})$, is defined as the maximum number of linearly independent rows (columns).
- If $r(\mathbf{A})=k$, then there exist k rows and k columns, where $k \leq \min(m, n)$.
- The number of nonzero or positive eigenvalues of **A** is equal to its rank.
- Thus, **Z**=**XA***, where **A***=selected eigenvectors (less in numbers compared to the original dimensions)

References

- Green, P.E., and Carroll, J.D., (1976). *Mathematical Tools for Applied Multivariate Analysis.* New York, USA: Academic Press.

- Jackson, J.E., (1991). *A User's Guide To Principal Components.* John Wiley and Sons. USA.

Authors Information

Credit to the authors: