GD5102 Techniques for Data Analysis

Assignment 02: Principle Component Analysis

Report submission date: 07-Oct-2020

Data, Material & Instruction

- Data: Multivariate data (file Assigment02.dat)
- Material: PPT slides, matlab scripts, some e-books
- Submission: Before 07-Oct-2020, 00:00:00

To : course.dudy@gmail.com
Subject : GD5102-TAD
File : **NIM_A02_FirstName_LastName.pdf** (*insert your matlab-script to the report*)

Tasks A

1. Remove the mean of each multivariate data and plot them at the same figure
2. Generate the data matrix X and transform it into a new set of data in PCA domain (Xnew).
3. Plot each multivariate data both in X and Xnew at the same figure. Discuss the difference among them.
4. Determine the correlation matrices of X and Xnew. Discuss the appearance of each matrix and discuss. (you may check the scatter plots of any 2-variable)
5. Transform back the Xnew into X (say Xback) and compare Xback and X. Do you see any difference between Xback & X?
6. Split X into major eigenvalues (global pattern) and the minor ones (local pattern). Plot the decomposed X.
7. What can you understand from the covariance & correlation matrices? What are the main roles of them in multivariate analysis?. Explain thoroughly.

Tasks B

1. Select your own data set and repeat the task A. How well can you understand the behavior of your data? [The selected data set should be very close to your thesis work].
2. Discuss general applications of correlation and regression in geoscience data analysis. How could they possibly be helpful in accomplishing your thesis work?

Dozent: Dr. *techn*. Dudy D. Wijaya & Dr. *rer.nat*. Wiwin Windupranata