CH 5440 Multivariate Data Analysis in Process Monitoring and Diagnosis

Assignment 4

1. Apply Non-negative matrix factorization (NMF) to extract the pure component spectra from the mixture UV absorbance data set used in assignment 3. Assume that the number of pure components in the mixture is known to be three. Use PCA to first obtain the denoised absorbance matrix of rank three and apply NMF to the denoised data to extract three sources.

(a) Draw one random sample from every five replicate measurements for each of the 26 mixtures and apply NMF denoised data. Compare the extracted pure component spectra to the measured pure component spectra using correlation and identify which of the pure component spectra are being extracted well (Note that the order in which NMF extracts the pure components cannot be ascertained, that is, there is a permutation ambiguity). Therefore you have to compare each extracted spectra with every pure spectra to determine the permutation order.

(b) Determine the average of the five replicates for each mixture. Apply PCA and NMF to the averaged data and determine whether the pure components spectra are extracted more accurately.

Report the correlations for the two cases in the form of a table and report your conclusions. For applying NMF you can use the code developed by Prof. Haesun Park of Georgia Tech (which is uploaded) or you can write your own code.

2. The data set contains UV absorbances for seven 3-component mixtures obtained by preparing mixtures consisting of Co, Cr, and Ni salts in nitric acid according to the experimental design shown by figure below. Show that the given network is NCA complaint.

(a) Apply PCA to the data set and obtain denoised spectra of mixtures assuming that the number of pure species is known. Determine the rotation matrix using the experimental design information and apply it to the denoised spectra and also estimate the pure component spectra using this rotation matrix. Report the rotation matrix as well as correlation between the estimated and true pure component spectra.

(b) Apply NCA to estimate the pure component spectra and compare with the true spectra given as part of the data set (use correlation coefficients). For applying NCA use NCA toolbox which has been downloaded from Prof James Liao's site (UCLA).

Species

Co   Cr   Ni

Mixtures