CH 5440 Multivariate Data Analysis in Process Monitoring and Diagnosis

Assignment 5

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 (NMF requires the number of pure species to be specified by the user). For this purpose you can either use the code provided (Prof. Haesun Park of Georgia Tech) or write your own code using the ALS multiplicative update rules. Before applying NMF it is better to set all negative absorbance values in the data matrix to zero. Also an initial estimate of the non-negative matrices can be obtained by first using PCA to reduce the rank of the absorbance matrix and use absolute values of the loadings and scores matrix as initial guesses for the non-negative matrices.

(a) Use the first sample from the five replicates for each of the 26 mixtures and apply NMF to the 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 NMF to the averaged data and determine whether the pure components spectra are extracted more accurately.

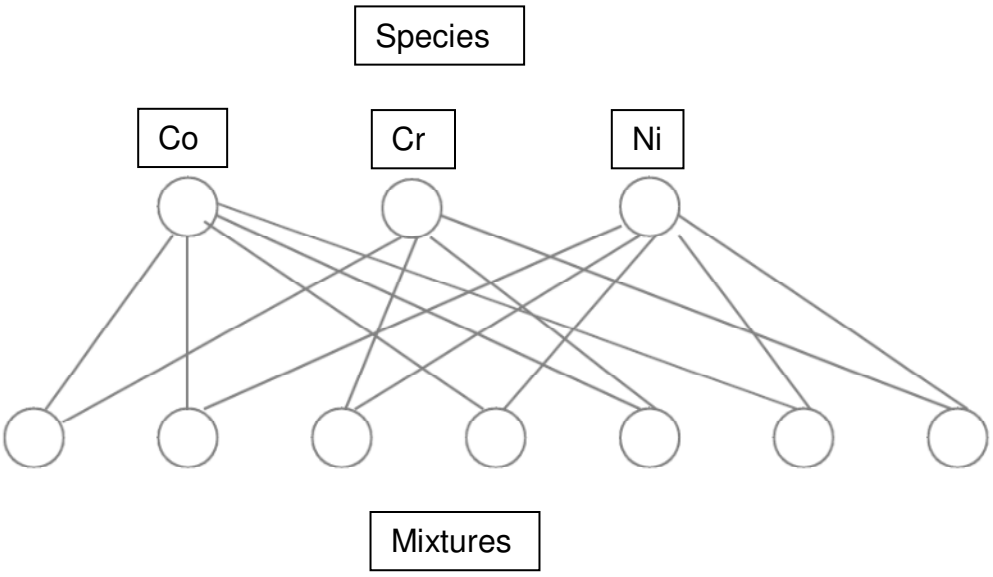
Report the correlations for the two cases in the form of a table and report your conclusions.

2. 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 in figure below is given in ncadata.mat.

(a) Show that the given network is NCA complaint.

(b) 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.

(c) 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