Solutions – Assignment 5

1. **Non-negative matrix factorization** (NMF) to extract pure component spectra

   The absorbance data considered from **'Inorfull.mat'.** In Part (a), the absorbance data of the first replicate of the mixtures collectively forms the data matrix $\boldsymbol{Z_{26\times176}}$. The number of species is known a priori to be 3. The NMF code provided by *Prof. Haesun Park, GATech* is used for solving the problem. The input to the *NMF* function **('nmf.m')** are the data matrix, number of species and the initial guesses are the absolute values of the loadings and the scores matrix calculated from *svd(Z)*.

   $$\boldsymbol{B}_{26\times3} = abs\big(\boldsymbol{U}(:,1:Ns)\big)\ \ P_{3\times176} = abs(\boldsymbol{S}(1:Ns,1:Ns)\boldsymbol{V}(:,1:Ns)')$$

   In part (b), the average of the 5 replicates of every mixture is considered for the data set *Z*. The correlation matrices computed between the actual pure component spectra and the extracted ones is presented below.

-0.5629	0.2143	0.3433
0.6284	0.6342	-0.4021
0.8738	-0.0757	0.2442

Figure 1: Correlation matrix for part (a)

-0.5209	-0.1247	0.7194
0.5826	0.8541	-0.4682
0.9023	-0.0483	0.0583

Figure 2: Correlation matrix for part (b)

   Observations: Pure species spectra are unambiguously extracted using average values (since average data has less noise), the sources are not separated well using non-averaged measurements. The first extracted spectra has a high correlation with all pure species spectra. The second and third extracted spectra have maximum correlation with the same (second) pure species spectra.

2. A) $\boldsymbol{A}_{struct}$ of the given network $\begin{bmatrix} X & X & 0 \\ X & 0 & X \\ 0 & X & X \\ X & 0 & X \\ X & X & 0 \\ X & 0 & X \\ 0 & X & X \end{bmatrix}$

   For the network to be NCA compliant, the number of zeros in each column has to be $\geq m-1$. Here *m*=3, therefore the given network is NCA compliant.

B) Assuming the number of species is known to be 3, the rotation matrix $\boldsymbol{M}$ can be calculated by solving system of linear equations

$$\boldsymbol{U_1 M = A_{struct}}$$

Thus the calculated rotation matrix is given by $M = \begin{bmatrix} 1 & 2.596 & 9.325 \\ 0.27 & 1 & -11.721 \\ 1.556 & -0.211 & 1 \end{bmatrix}$

The correlation matrix is given in Figure 3

-0.9425	0.4291	0.4343
0.3662	-0.9968	-0.3815
0.4940	-0.4074	-0.9967

Figure 3: Correlation matrix for part (b)

Observations: The pure species spectra of the three species are extracted very well (correlations greater than 0.9).

C) <u>Solution using the NCA toolbox</u>

0.9425	-0.4293	-0.4342
-0.3659	0.9968	0.3815
-0.4942	0.4075	0.9967

Figure 4: Correlation matrix from NCA algorithm