10.34 – Fall 2006
Homework #5
Due Date: Wednesday, Oct. 11th, 2006 – 9 AM

Problem 1: Singular Value Decomposition

The composition of a mixture of chemicals is usually determined by putting the mixture in a spectrometer or chromatograph. The signal coming out of an analytical-chemistry instrument for a mixture is typically a linear combination of the signals one would have obtained if the sample was pure. For example, using a spectrometer:

$$S_{mix}(\lambda_n) = \sum x_i S_i(\lambda_n)$$

Where $S(\lambda)$ is the signal measured of a mixture with mole fractions $\{x_i\}$, and $S_i(\lambda)$ is the signal that would have been measured if the mixture had been pure, i.e. if x_i=1 and all other x's had been zero. The spectrum S is a continuous function of the wavelength λ, but the spectrometer only records the spectrum at N discrete set of wavelengths λ_n.

Very often, the composition x of a mixture changes, either with time (e.g. when we are measuring batch kinetics) or as a function of some knob being turned in a chemical plant. We pull samples periodically to measure these changes, at K times $\{t_k\}$, and run each sample through the spectrometer, obtaining a big data set $S(t_k,\lambda_n)$ which we think can be deconvoluted:

$$S_{observed}(t_k, \lambda_n) = \sum x_i(t_k) S_i(\lambda_n)$$

Use SVD to write the matrix $S_{observed}$ (on the Web page) in the format (The rows in the data file corresponds to a given time and the time is measured from 0-100 seconds. The wavelength is measured from 432-482 nm, both of them are measured with linear spacing.)

$$S = \sum \sigma_i u v^T$$

Where u is a vector with K entries, v is a vector with N entries, and uv^T is a matrix formed by the outer product. From this expansion, figure out how many major (light-absorbing) components are in the mixture. Plot the normalized spectrum and time-dependence of each major light-absorbing species in the mixture.

N.B. When two spectral peaks are observed to move with the same time-dependence, this can mean (a) that two species coincidentally have about the same time dependence; (b) that a single molecule has two different absorption bands, or (c) that there are two species in rapid equilibrium with each other. It is often very difficult to distinguish between the latter two possibilities. An example of the last case is glucose, which exists as a mixture of several rapidly interconverting isomers (one with a 6-membered ring, one with a 5-membered ring, and an open-chain species.)