Exercise 1 for Matlab session 4

November 26, 2014

Biological applications of spectroscopy

There are many spectroscopic measurement techniques (such as mass spectroscopy, infra-red spectroscopy, nuclear magnetic resonance spectroscopy) having applications in biology. A *spectrum* measured with such a technique is able to provide information about the chemical content of a sample under investigation (e.g., a tissue, a fluid), but the spectrum contains a combination of responses from individual chemical components.

In Figure 1 we see an example of a spectrum with 2 spectral components (the two peaks), while in Figure 2 we can see the two individual components of the same mixture.

Note that only the position of the peaks is fixed, while the height can be different (depending on the amount in the chemical mixture). These individual components can be viewed as *basis vectors* for the *vector space* or *subspace* containing all possible spectra of chemical mixtures of these components. Indeed, all spectra will be linear combinations of the spectra of the two 'pure' components.

In this assignment you need to estimate the quantity of each component in several given mixtures. The file "spectraldata.mat" (which you should load in Matlab with "load spectraldata") contains two spectral components (named `component1` and `component2` and several mixtures (named `mixture1, mixture2, mixture3`).

Create an m-file that performs the following steps:

1. Plot `component1` (in blue) and `component2` (in red) on top of each other on the same plot (each in a different color).

2. Plot `mixture1` on top of the previous plot (in green).

3. Knowing that `mixture1` is a *linear combination* of the two components, determine a system of linear equations (SOLE) to quantify how much

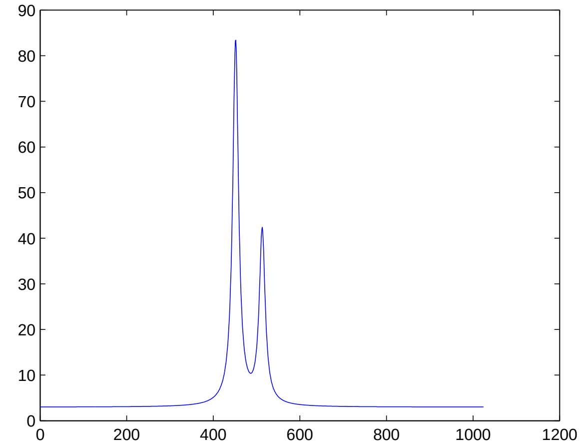

Figure 1:
A spectrum containing a mixture of 2 components.

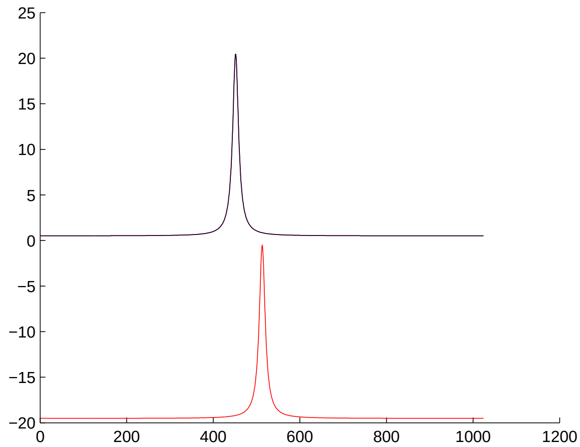

Figure 2:
Two individual spectral components. We assume that the individual components are scaled to a meaningful height corresponding to, e.g., 1 mole of substance.

of each component is present in `mixture1`. Solve the SOLE using row reduction (hint: use `rref`).

4. In a new figure, plot `mixture1` (in blue) and on top the sum of the two components scaled to the corresponding quantities found at the previous step (in red, and with dashed line). If you did it right, both plots should be perfectly on top of each other.

5. Note that real spectroscopy data is typically noisy. Make a new figure, and plot `mixture2`. You should see that `mixture2` is much noisier than `mixture1`.

6. We cannot repeat exercise 3 for `mixture2`, since the noise will make the SOLE inconsistent. Yet, we would like to check whether `mixture2` and `mixture1` are samples of the same substance. Use least-squares to (approximately) quantify the amount of `component1` and `component2` in `mixture2`. From the result, can you see whether `mixture1` and `mixture2` have a similar chemical content?

7. Repeat exercise 4 (this time for `mixture2`). Of course, this time the two plot will not perfectly match. The dashed plot can be viewed as a 'denoised' version of the blue plot.

8. Repeat exercise 6 and 7 for `mixture3`. What is (approximately) the amount of `component1` and `component2` in `mixture3`?

Hint: linear systems and least squares problems must be solved with the backslash operator "\". The traditional formula $(A^T A)^{-1} A^T \mathbf{b}$ is more sensitive to round-off errors in the computation.