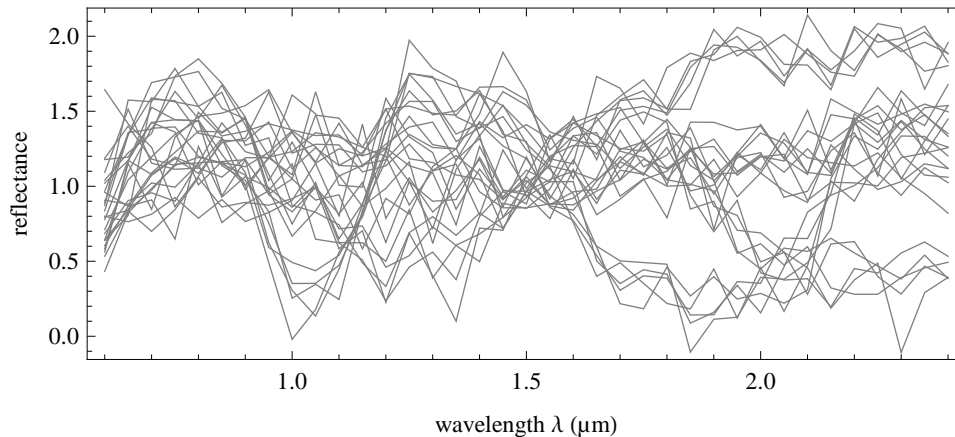**7th exercises for SIM'2021**

Ex. 1

Datafile `spectra.dat` has 26 rows — first row has the wavelengths from 0.6 μm to 2.4 μm, and the rest are reflectancies of 25 "asteroids" at those wavelengths (plotted below).

Use principle component analysis to find underlying five groups of asteroids based on their spectra. The first three principle components should be enough to distinct the groups. Plot the observations in the new PCA variable space, e.g. PCA-1 against PCA-2 and PCA-1 against PCA-3.

If you succeed, you should notice that the first 5 rows of data belong to the first group, the second 5 rows to the second group etc. The spectra of those groups correspond to common minerals in asteroids — nickel-iron (obs. 1–5), olivine (obs. 6–10), orthopyroxene (obs. 11–15), plagioclase feldspar (obs. 16–20) and spinel-bearing Allende inclusion (obs. 21–25) (from Binzel et al. (eds.) Asteroids II).

Ex. 2

Let us assume that the data in the previous exercise (`spectra.dat`), has 5 groups. The groups are ordered in the data so, that 5 first rows are nickel-iron, 5 next rows olivine, then orthopyroxene, then plagioclase feldspar, and finally spinel-bearing Allende inclusion. Use nearest-neighbor algorithm on the five new spectral observations in `new-spectra.dat`, and classify. You can use, for example, 5 nearest neighbors.

You can do PCA transformation on the data and new observations if you want, but you can also work directly on the spectral values.