

Homework #5

Due April 15, 2024, 11:59pm

Problem 5.1

Calculate the frequency spectrum of the following signals, sampled at 44.1kHz:

- Guitar sounds are given in the two files `nylon16.wav` and `strat16.wav`. Determine the frequency of the fundamental for both sounds.
- Calculate the relative energy in each of the first 3 harmonics (exclude the fundamental)
- Estimate the decay time of the third harmonic (in sec)

Hint:

- The energy in a signal is the square of the absolute value of the Fourier transform.*
- The frequency spectrum is the energy as a function of frequency*

Problem 5.2

In the class data directory on the SciServer you will find a file `covid-19-us.csv`. This contains cumulative data on the COVID-19 cases for the whole US from the day the first infection was detected. The data is from the New York Times data collection on GITHUB (see <https://github.com/nytimes/covid-19-data>).

The problems presented here will serve as a warm-up exercise in learning the basics of Python and the basic data input/output and graphing techniques (*hint: read the csv file into a pandas DataFrame*).

- Read the file and create a plot of the cumulative cases as a function of time
- On a separate plot, create a count of the cumulative deaths as a function of time
- Repeat both plots when you show the numbers of cases on a logarithmic scale
- On the cumulative charts visually find the segments when there was an exponential growth in the number of cases, indicated by a linear increase of the $\log(\text{counts})$. Estimate the slope visually, and overplot a straight line with approximately matching the trend of the data. Interpret the slope in terms of the number of days that the cases double.

In order to be able to use the shared data, you should have the following imports at the top of your iPython notebook:

```
import pandas
import os
```

followed by

```
path = '../../../AS_171_205_2021/AS_171_205_2021-Course/data/'
files = [f for f in os.listdir(path) if f.endswith(".csv")]
files
```

You should see the following output:

```
['covid-19-us.csv', 'covid-19-us-states.csv',
 'covid-19-us-counties.csv']
```

If you do not see this directory, you need to go to the SciServer Compute page with your Containers, and you need to create a new container. When you do that, make sure that you select the User Volume

Finally, you can read the first data file with

```
df =pandas.read_csv(path+files[0])
```

There is an example notebook for database access below list of the homeworks.

Problem 5.3

Consider the data set describing wine quality vs different attributes, described at:

<http://archive.ics.uci.edu/ml/datasets/Wine+Quality>

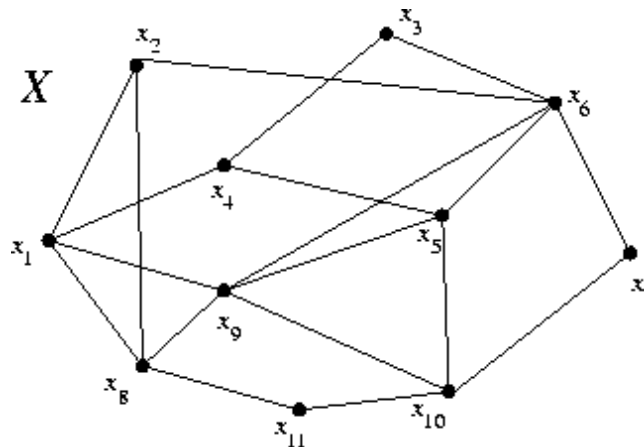
The data itself is at `winequality.zip`, in the data directory [1].

- Use Principal Component Analysis to find a projection in which the excellent and poor wines can be best separated. Perform the analysis for both the red and white wines. If possible, try to determine the three most important factors influencing the decision.
- Repeat the analysis using t-SNE and UMAP

Problem 5.4

Consider the graph X with 11 nodes shown on the figure below.

- Create the adjacency matrix A for the graph X .
- Using the adjacency matrix, calculate the number of closed triangles on X .
- Calculate the Laplacian matrix of X and calculate its eigenvalues and eigenvectors.
- Use the eigenvectors corresponding to the two lowest eigenvalues (different from 0) as the x, y coordinates to re-plot the graph.
- Use an algorithmic method to determine the shortest path from node 1 to node 7.
- Use an algorithmic solution to find the pair of nodes with the maximal shortest path (*hint: try Dijkstra's algorithm*)



Problem 7.3.

The file DS.wav contains a recording of the mysterious 'cricket' sound hurting US diplomats in Cuba. Determine the frequency spectrum of the signal. Do you notice anything unusual? Break the sound into 1 sec long overlapping segments, stepping in 0.5 sec intervals and calculate the power spectrum and plot the frequency spectrum as a function of time. What is the accuracy of the frequency measurement for these spectra?