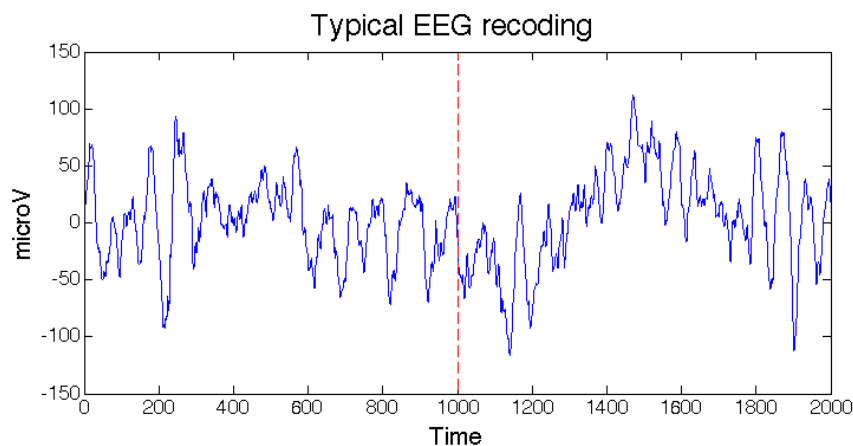# Introduction to Computational Neuroscience
Practice II: Data Analysis - Continuous Data

Ilya Kuzovkin, Raul Vicente

March 9, 2015

This is how usual EEG recording of one channel looks like:

On the x axis there is time, on the y axis strength of the recorded signal in μV. The dashed line at the time point $x = 1000$ is the moment when the *stimulus* (picture) was shown to the test subject.

Exercise 1: Event Related Potential (ERP) (1pt)
By looking at the plot above one cannot say that showing the stimulus had any considerable effect on the test subject. It seems that there is an increase starting at the time point 1200 and ending at the time point 1500, but that can be a random event, which is not related to the stimulus.

In such studies the way to go is to conduct the same experiment several times and then average the results. In this way the response, if it is there, will reveal itself much more clearly.

In the `data` folder you have file `erptrials.csv`. One row is one trial recorded for 2 seconds with sampling rate of 1000 Hz. The stimulus was shown at the time point 1000 (1 second). There are 79 trials in this file. Your task is to plot an average of all 79 trials and see whether there is a clear ERP response or not.

Exercise 2: Frequency Analysis (2pt)
The most popular operation that you can do with continuous brain data is converting it from the *time domain* to the *frequency domain*. Due to the fact that any function can be represented as a sum of sinusoids we can decompose our signal into such sinusoids and observe from which frequency components it consists. In terms of the brain data such transformation makes particular sense, because of the *brain rhythms* – different frequencies of the firings of the neurons

are related to the different kinds of mental activity[1].

In this exercise we will plot *power spectrum* and see how *alpha wave* emerges when the test subject's eyes are closed. In the `data` folder you can find two files: `eyes_open.csv` and `eyes_closed.csv`. The data is recorded from the channel `Pz`. One row contains 4 seconds of the signal, sampling rate is 512. Each file has 15 recordings.

Your task is to perform Fourier analysis on both datasets, plot power spectra and compare the results. Do it as follows:

1. Plot some of the signals just to see how they look like.

2. For each recorded signal (2048 data points):

   (a) Use `fft(signal)` to compute power spectrum of this signal. You will get a vector of complex numbers.

   (b) Use `abs(result_of_fft)` to obtain the magnitude[2].

3. Sum together 15 power spectrum distributions, that you got from the previous step.

4. And divide the resulting vector by 15 to obtain the average.

5. Plot it. You will see that the right part of the graph is mirror image of the left part. Discard the right part.

6. Your x axis goes from 1 to 1024, which does not correspond to actual frequencies. Compute correct x axis as follows:

   ```
   dt = 1/512;              % time step
   df = 1/4;                % frequency step
   fNQ = 1/dt/2;            % fNQ is the maximal frequency, in our case it is 256
                            %   if you have discarded part of the data in step 5
                            %   or 512 if you did not
   xaxis = (0:df:fNQ-df);   % points for your X axis, should be of the same length
                            %   as the vector of frequencies
   ```

7. And plot it again

   ```
   plot(xaxis, powers);
   ```

   Now your axis goes from 0 to 256 (or 512) with step of 4, which corresponds to the frequency range `pow` function produced.

After making the plot more beautiful and focusing on the range from 0 to 30 Hz you should obtain the result that looks something like this:

---

[1]`http://en.wikipedia.org/wiki/Electroencephalography#Wave_patterns`
[2]`http://www.mathworks.se/help/matlab/ref/abs.html`
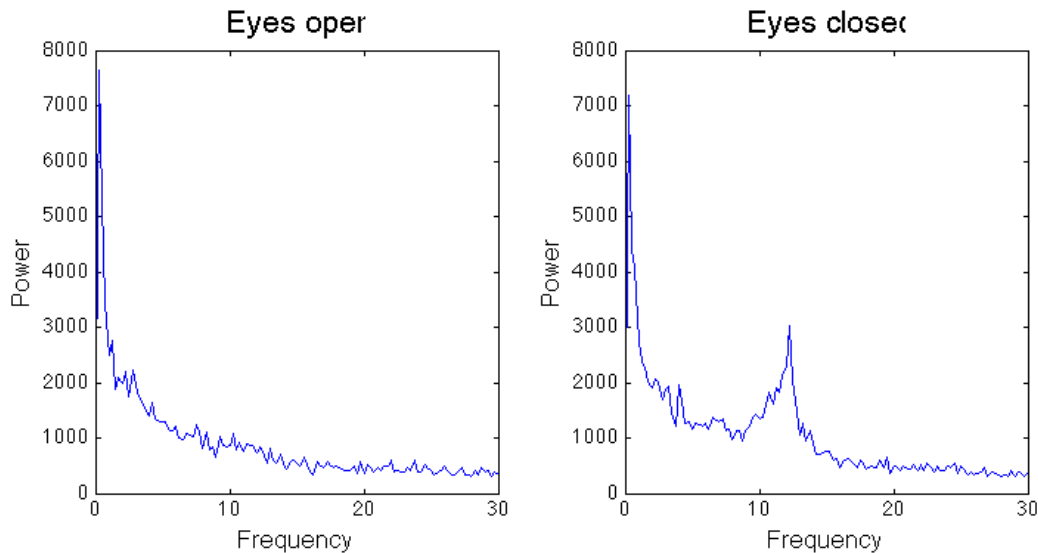
## Exercise 3: Epilepsy (2pt)

In the folder `data/epilepsy` you have two data files: `sz4_pre.dat` and `sz4_ict.dat`. This data was recorded using ECoG[3] array. The file with the suffix `_pre` is recorded before the seizure and the `_ict` file is recorded during the epilepsy seizure. Read the `README.txt` file for the description of the data. Your task is to perform Fourier analysis on both files and see how brain activity changes during the epilepsy seizure.

Hint. Reuse the code from the previous exercise.

Hint. The data has readings from the 76 electrodes. First make Fourier transform on each of them separately, then compute the power spectra (each one separately) and then compute the average and plot that.

Exercise 4: Nyquist theorem (1pt)

What does Nyquist theorem say? What does it mean? What are the implications?
What does this picture illustrate?
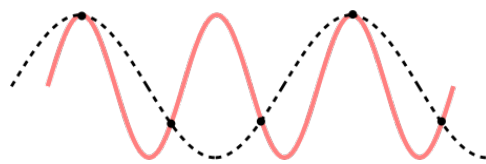
Prepare to present and explain the theorem to your fellow students.

Exercise 5*: Discrete Fourier Transform (DFT) (bonus 2pt)

Now let us have a look on what happens behind the scenes when you call the `fft`[4] function. Fourier transform is possible because, as we mentioned before, any function can be represented as a sum of sinusoidal functions of different frequencies and amplitudes. If you have some variable changing over time you can transform this data from the *time domain* (time on the

---

[3]http://en.wikipedia.org/wiki/Electrocorticography

[4]Actually Fast Fourier Transform (FFT) is a specific algorithm. In this exercise we do not analyse that, but rather look at the equations, which FFT solves in his own mysterious way.

x axis and the value of the variable on the y axis) to the *frequency domain* (frequency on the x axis and it's *power* on the y axis) to see if there is any periodicity in the changes of that variable.

Transformation is defined both for continuous and discrete data. On practice we mostly work with a discrete signal and therefore we use Discrete Fourier Transform (DFT). The equation for the discrete transform is

$$X_f = \sum_{t=0}^{N-1} x_t e^{-i2\pi f \frac{t}{N}}$$

where

- f is the frequency of the component, contribution of which you are trying to estimate.
- X_f is a complex number, that encodes both amplitude and phase of the sinusoidal component.
- x_t is a value of your signal at the time point t.
- N is the total number of time points.

Let us say that you would like to know contribution of sinusoidal components of frequencies from 1 to 100 Hz. Then you should apply this equation 100 times. As a result you will get 100 complex numbers. To extract amplitude from these numbers use the following equation on each of them:

$$A_f = 2\frac{\sqrt{\mathrm{Re}(X_f)^2 + \mathrm{Im}(X_f)^2}}{N}$$

Now you have all the contributions and you have completed Discrete Fourier Transform of your signal. For example if 7th number is "25" then it means that the amplitude of the sinusoidal component with frequency $7 \times$ df Hz is 25.

Here you see a curve, that is a sum of three sinusoids. Your task is to apply DFT and find out frequencies and amplitudes of those components. You can find data points for this curve in the file `curve.csv`.
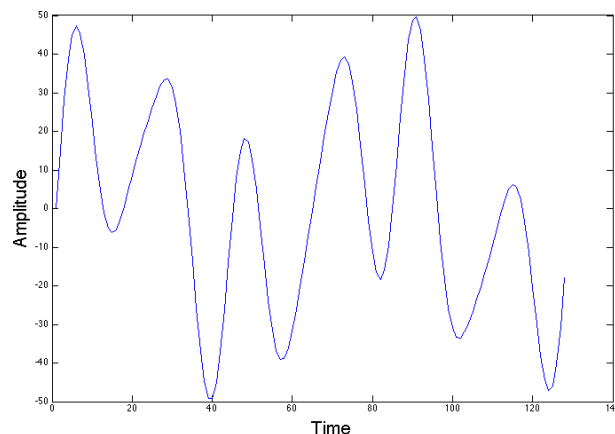
Please submit a `pdf` report with the answers to the questions, plots and comments about your solutions. Also submit the code for the programming exercise(s). Pack those into the `zip` archive and upload to the course web page.