Functional Data Analysis in EEG data

Tadas Danielius

In this assignment we are going to investigate a data-driven approach and apply functional data analysis toolbox on EEG brain waves and detect of statistically reliable fluctuations. The approach is tested on EEG data which is available from "`BCI Competition III`"

Introduction

The field of computerized neurophysiology, like other fields interwined with the emergence of computer science and has grown exponentially in recent years. Observation of electrical signals from the nervous system goes back as early as 1848 when researchers such as Duboi-Reymond reported the presence of electrical signals as a marker of a peripheral nerve impulse. These early studies revealed that peripheral nerve conduction involved electricity and led Caton in 1875 to propose a similar finding for brain wave activity as noted in animal studies on monkeys and rabbits. As digital computer technology developed in the 1960s and 1970s, it became feasible to assess and quantify precisely many more EEG parameters than is possible through human visual inspection of raw EEG waveforms. However, a major challenge lies in extracting meaningful information from neuroelectrical responses typically characterized by poor signal-to-noise ratios. Especially detecting event-related potential (ERP) which is measured brain response that is the direct result of a specific sensory, cognitive, or motor event.

In this assignment we will try to apply some of the Functional Data Techniques to raw EEG

data to identify certain signal fluctuations, termed ”components”, that are statistically reliable as well as cognitively meaningful.

Origin of the Electric signals

The EEG records electrical activity from the cerebral cortex. In as much as electrocortical activity is measured in microvolts (μV), it must be amplified by a factor of 1,000000 in order to be displayed on a computer screen. Most of what we record is felt to originate from neurons, and there are a number of possible sources including action potentials, post-synaptic potentials. Action potentials induce a brief (10 ms or less) local current in the axon with a very limited potential field. Post-synaptic potentials (PSPs) are considerably longer (50-200 ms), have much greater field, and thus are more likely to be the primary generators of the EEG.

In the normal brain an action potential travels down the axon to the nerve terminal where a neurotransmitter is released. At the post-synaptic membrane the neurotransmitter produces a change in membrane conductance and trans-membrane potential. If the signal has an excitatory effect on the neuron it leads to a local reduction of the transmembrane potential (depolarization) and is called an excitatory post-synaptic potential (EPSP), typically located in the dendrites. An inhibitory post-synaptic potential (IPSP), typically located on the cell body of the neuron, The combination of EPSPs and IPSPs induces currents that flow within and around the neuron with a potential field sufficient to be recorded on the scalp. It turns out that the typical duration of a PSP, 100 ms, is similar to the duration of the average *alpha wave*. The alpha rhythm, consisting of sinusoidal or rhythmic *alpha waves*, is the basic rhythmic frequency of the normal adult brain.

Data

The recording was made with a 60 channel EEG amplifier from Neuroscan, using the left mastoid for reference and right mastoid as ground. The EEG was sampled with 250 Hz, it was filtered between 1 and 50hz with Notchfilter on. 60 EEG channels were recorded according the scheme in the figure 1 (Since the overall recording is too long for exploration I will use 1 second interval to visualize and get the basic understanding about the data. For visualization purposes I will use 4 channels.)

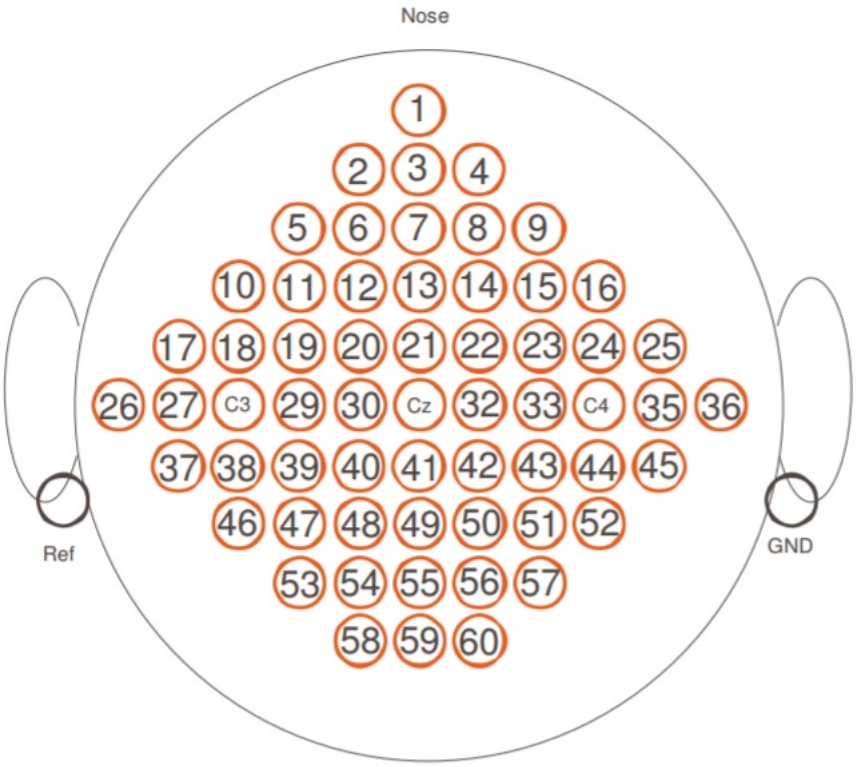


Figure 1: Position of EEG electrodes

In figure 2 we can see data from 4 channels recorded in 1 second.

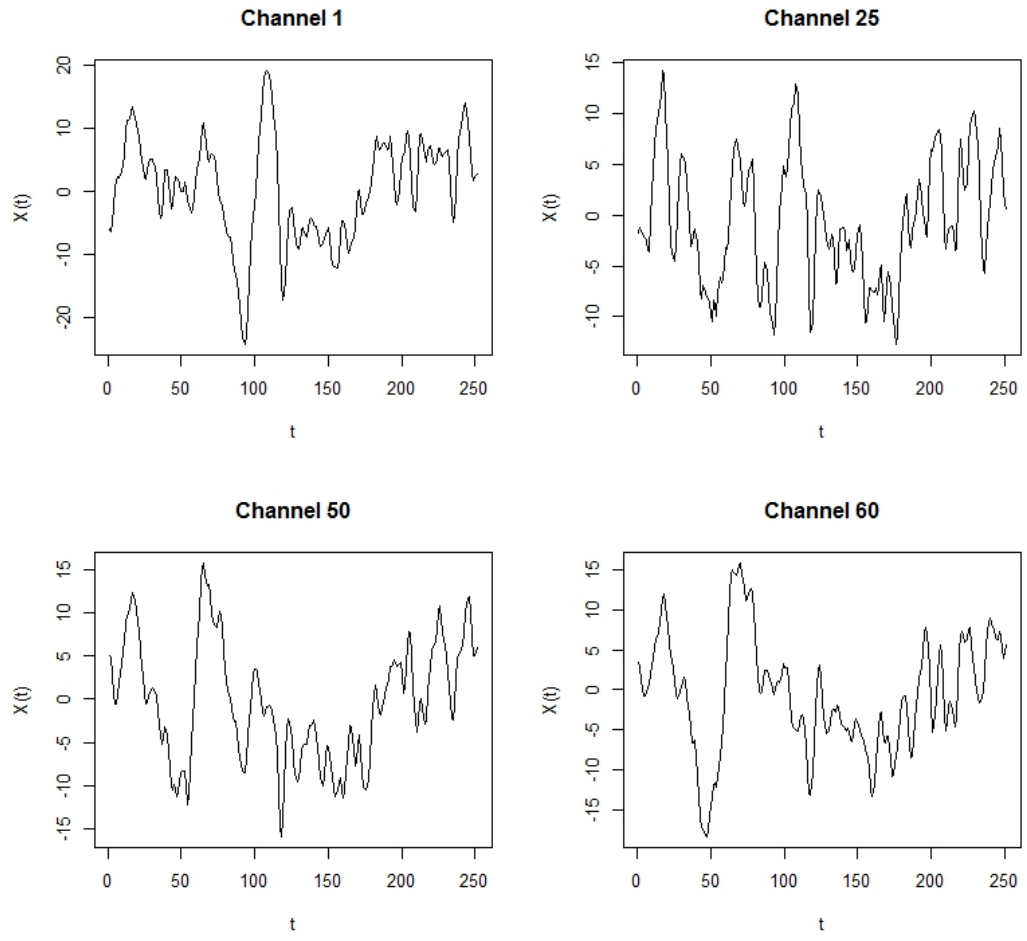


Figure 2: 4 Channels. 1sec recording

Paradigm

The subject sat in a relaxing chair with armrests. The task was to perform imagery left hand, right hand, foot or tongue movements according to a cue. The order of cues was random. The experiment consists of several runs with 40 trials each; After trial begin, the first 2 seconds were quite, at t=2s an acoustic stimulus indicated the beginning of the trial and a cross ”+” is displayed; then from t=3s an arrow to the left, right, up or down was displayed for 1 sec. At the same time the subject was asked to imagine a left hand, right hand, tongue or foot movement,

respectively, until the cross disappeared at t=7s. Each of the 4 cues was displayed 10 times within each run in a randomized order.

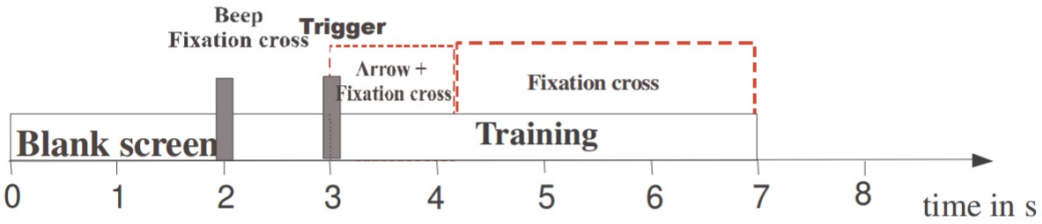
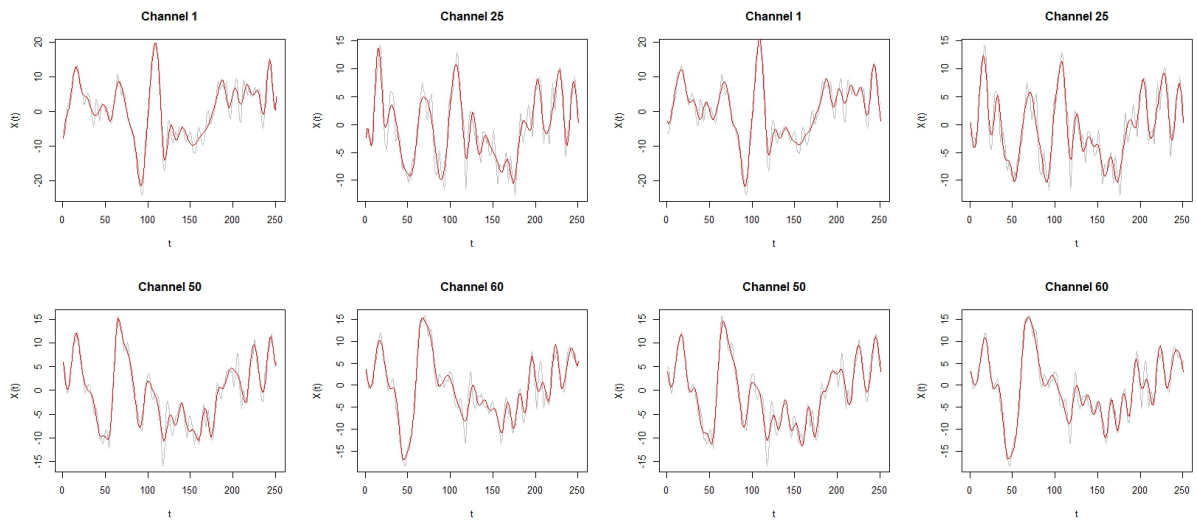


Figure 3: Timing of the paradigm

Basis system

As seen from figure 2 that raw EEG data tend to fluctuate. The goal of fitting data by temporally continuous function that is smoother and exhibits less fluctuations is to reduce noise by the function which does not over fit the raw data yet represents it's main components of variation To build functions from the observed discrete points first we need to define functional building blocks ϕ_k called *basis functions*. Then we set up a matrix of coefficients to define the function as a linear combination of these basis functions.

From raw data y_i representing signal amplitude sampled at discrete time steps $t_j, j = 1, ..n$ we need to choose smooth approximation $x(t)$ The **Fourier basis system** is the usual choice for *periodic* functions and the **spline basis system** tends to serve well for non periodic *trajectory* data. It would make more sense to use Fourier functions since EEG activity normally is rhythmic consisting in waves of approximately constant frequency. In figure 4a we can see smoothed data using **bspline** type with 39 functions and in figure 4b smoothed data using **Fourier** type with 39 functions. There is no visual difference either Fourier or bspline therefore we will stick with Fourier basis. Both bspline and Fourier captured very well dominant *Alpha waves* (7.5 and



(a) Using bspline (b) Using Fourier

Figure 4: A figure with two different smoothing basis

13 Hz) and ”lost” almost all *beta activity* which is defined as ”fast” activity. It has a frequency of **14 Hz and greater**.

Exploring data

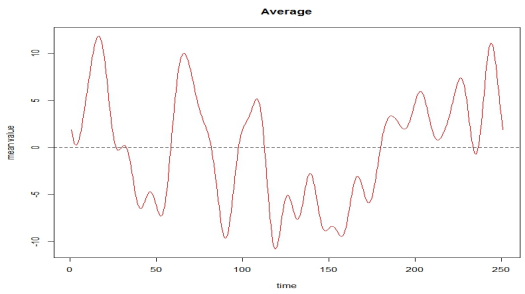


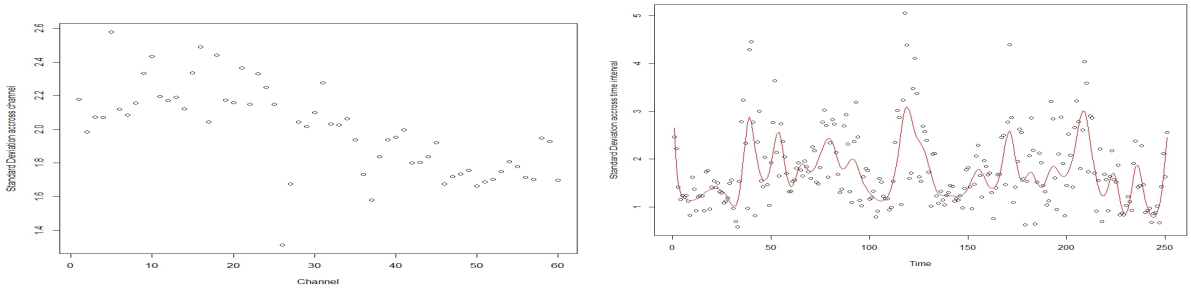
Figure 5: Mean

In figure 5 we can see the mean of all channels for 1 second interval. We can observe the dominance of *Alpha waves*. The *alpha* in fact is the `principal background` feature of the normal adult EEG. It is defined as a rhythmic frequency at **7.5 12.5 Hz**, usually of maximal amplitude in the *occipital regions*. *Alpha* is best seen when the person is in the relaxed, waking state. Indeed, as stated above, data was recorded while subject was sitting in the relaxed chair. Around t=0.5 second some visible *Beta wave* activity can be visually inspected. *Beta wave* activity is

defined as a frequency of **14 Hz and above**. *Beta*, the second waveform is rhythmic in character and is present in the background of most subjects. Normally we should see *beta wave* all time, if it's completely absent it may represent an abnormality depending on the other features of the EEG. As mentioned above it is worth stressing out that smoothing heavily reduced information about the presence of the *beta wave* amplitude. In this work, beta waves are not much of the interest, however for medical reasons analyzing inter-hemispheric asymmetry, in particular, the side of reduced amplitude usually points to the pathological hemisphere. In that case greater care must be taken to make sure this information is not lost.

Assessing the fit to the Data

We need to do further analysis of the residuals $r_{i,j} = y_{i,j} - x_i(t_j)$ to see maybe we have missed some important features by over smoothing. Whatever the variation in the residuals conforms to the assumptions implicit in the type of the smoothing that we performed. The use of the unweighted least-squares criterion is only optimal if the residuals for all time points are normally distributed if the variance of these residuals is constant across all channels.



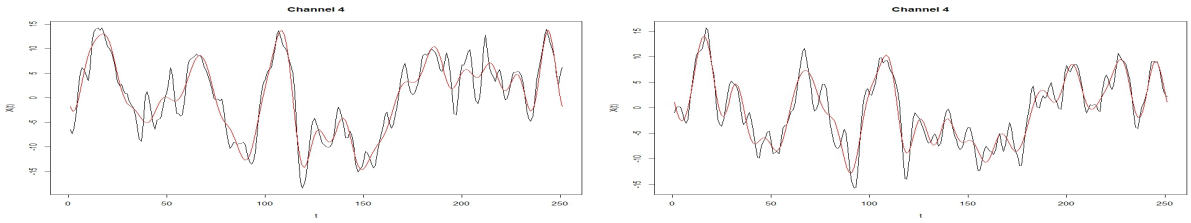
(a) Across channels

(b) The solid line is an exponentiated smooth of the log of the variance

Figure 6: Standard deviations of the residuals

Much larger variation around *frontal lobe* compared to *posterior temporal* (back of the head)

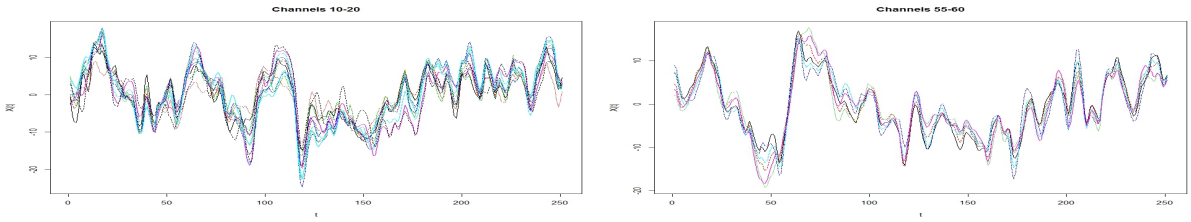
(Fig. 6a). This is interesting, there could be an indication that we might have done better to smooth different brain regions separately.



(a) Worst fit (b) Best fit

Figure 7: Two channels with best fit and worst

Figure 7a and 7b shows channel 4 (which has highest residual variation) and channel 24 with lowest variation. Indeed, we can observe that on the second part of the chart (starting from t=150) there are stronger *beta activity* which is not ”captured” well after smoothing. This does, make sense, since maximal *beta* amplitude is usually in the *frontocentral regions* whilst channels with lowest residuals variation were connected to *occipital regions* which normally has maximal *alpha* amplitude. In the figure 6b we can observe during which time period the variance of the residuals was highest. Again, comparing with data we can see, that those areas are mostly dominated by *beta waves*.



(a) Frontal lobe (b) Occipital lobe

Figure 8: Selected channels with highest and lowest residuals variation

Finally we can look at the brainwave functions for *frontal lobe* (figure 8a) with highest

residual variation and *Occipital Lobe* (figure 8b) with lowest residual variation.

Functional Principal Components Analysis

We want to see what primary modes of variation are in the data, and how many of them seem to be substantial. As in multivariate statistics, *eigenvalues* of the bivariate *variance-covariance function v(s; t)* are indicators of the importance of these principal components, and plotting *eigenvalues* is a method for determining how many principal components are required to produce a reasonable summary of the data.

In functional PCA, there is an *eigenfunction* associated with each *eigenvalue*, rather than an *eigenvector*. These *eigenfunctions* describe major variational components.

Plotting the principal component scores for pairs of harmonics can reveal interesting patterns on how curves cluster and otherwise distribute themselves within the K-dimensional subspace spanned by the eigenfunctions

Figure 11 shows some fascinating structures. The labeled points represents electrode on the head. We can see that patterns indeed exists. Inferior frontal (1-9 electrodes), mid-temporal (10-25) and posterior temporal (26-60) tend to form separate clusters.

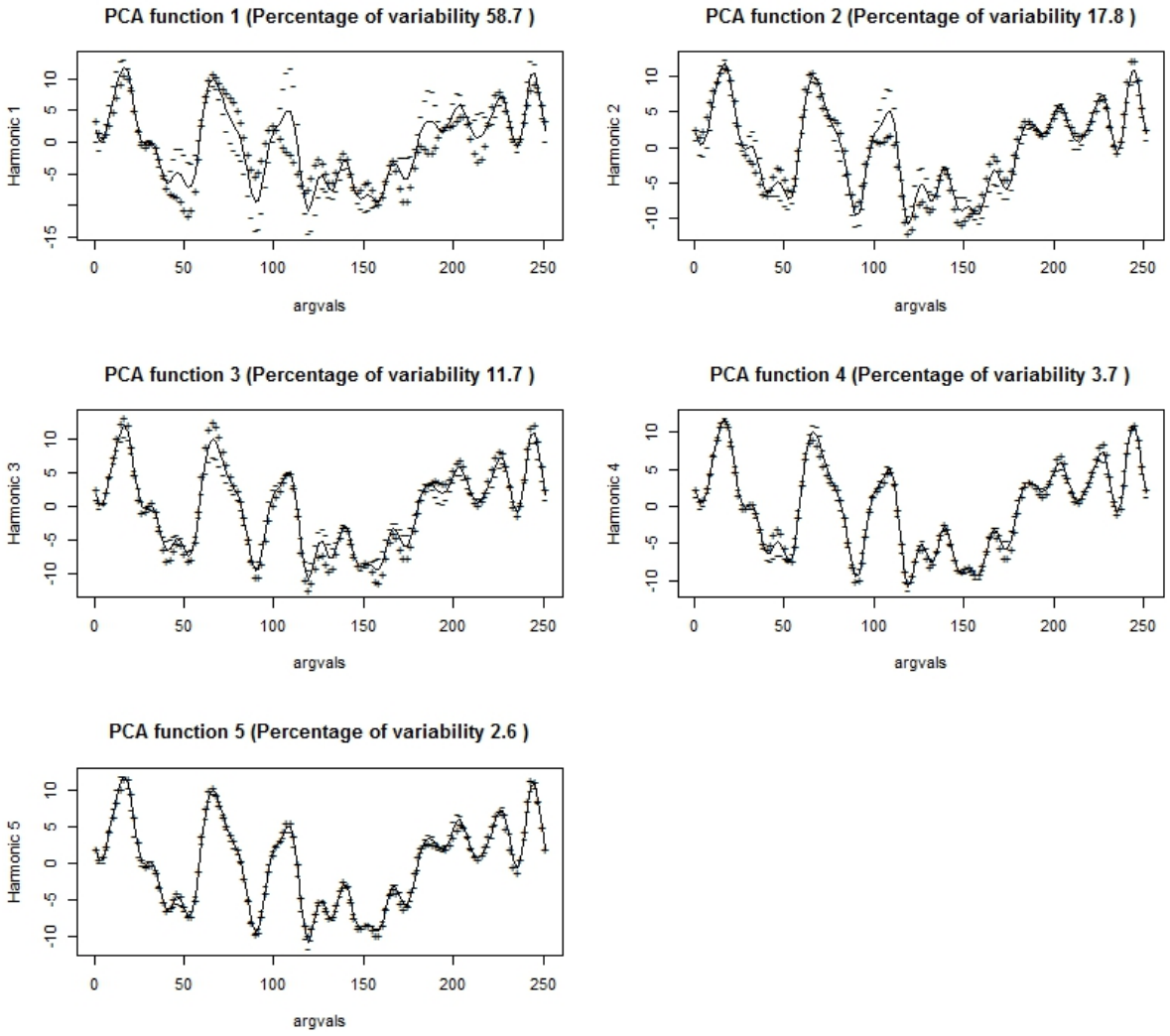


Figure 9: The 5 harmonics are shown as perturbations of the mean, which is the solid line. The +'s show what happens when small amount of a principal component is added to the mean, and the -'s show the effect of subtracting the component. Five harmonics account for 94% of the variation around the mean curve

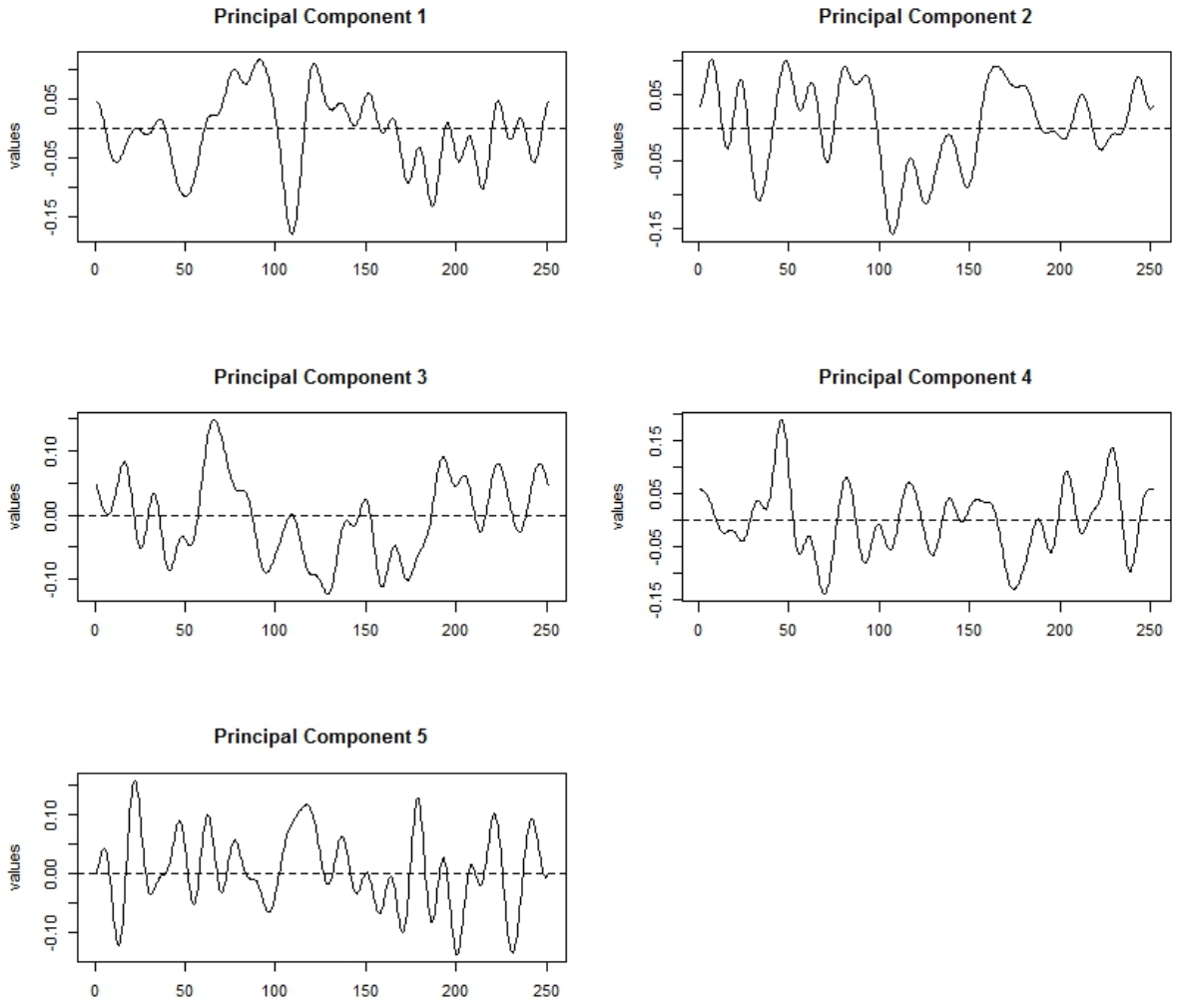


Figure 10: Principal Components

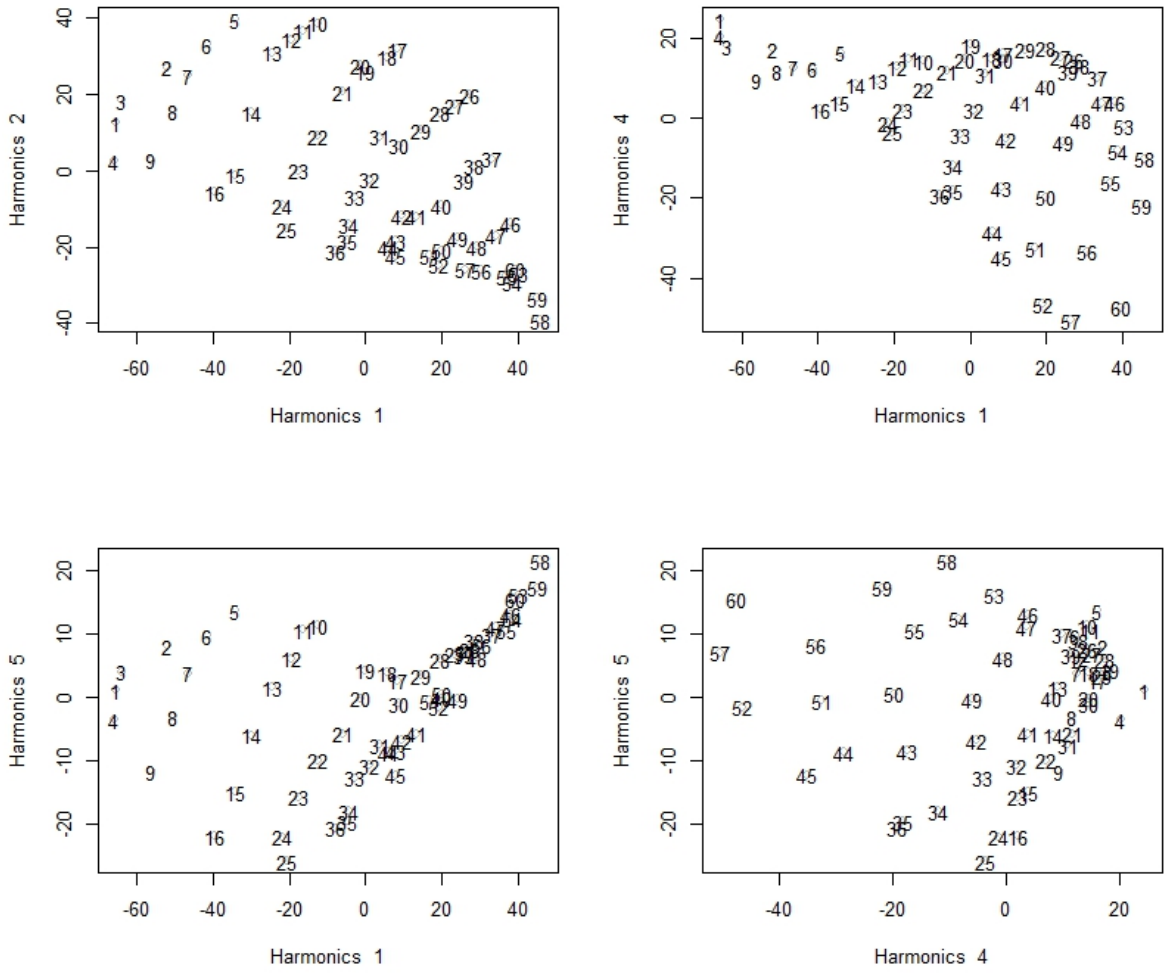


Figure 11: The labeled points represents electrode on the head. Inferior frontal (1-9 electrodes), mid-temporal (10-25) and posterior temporal (26-60) tend to form separate clusters

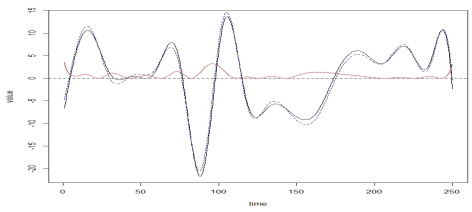


Figure 12: Reconstructed channel 1 function from principal components

Since it is enough 5 principal components to explain almost all variability we can use those components as basis instead of Fourier functions to reconstruct our functions without loosing much of the information. In the figure 12 we can observe nearly identical results. Red solid line indicates residuals.

Clustering

The aim of the cluster analysis is to build homogeneous groups (clusters) of observations representing realisations of some random variable X by working explicitly with its infinite-dimensional nature. Clustering is often used as a preliminary step for data exploration, the goal being to identify particular patterns in data that have some convenient interpretation for the user. Principal component analysis revealed that channels tend to form clusters. We expect to see clusters around different brain areas. A common problem to clustering studies is the choice of the number of clusters. The brain has specific areas that do certain types of work. There are three main areas: *Frontal Lobe*, *Parietal Lobe* and *Occipital Lobe*, thus we can start with three clusters.

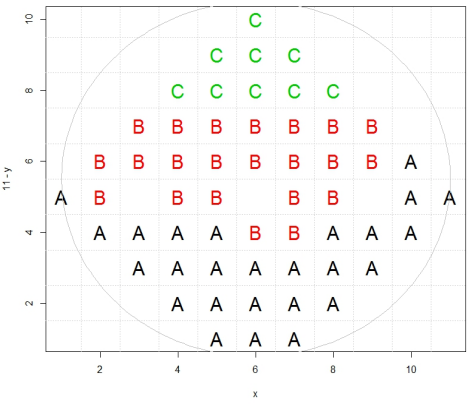


Figure 13: Channels clustered into three different clusters. Each letter represents different cluster

Figure 13 shows us performed k-means on functional data. We can clearly distinguish three clusters by brain regions. However, due to the nature of k-means algorithm which randomly

selects initial centers re-running k-means on the same data without specifying initial clusters may lead to similar but different results. But we always see forming clusters within different brain regions.

Cognitive events in EEG

Mu waves, also known as *mu rhythms* are synchronized patterns of electrical activity involving large numbers of neurons, probably of the pyramidal type, in the part of the brain that controls voluntary movement. These patterns repeat at a frequency of 7.5 - 12.5 (and primarily) Hz and are most prominent when the body is physically at rest. Unlike the alpha wave, which occurs at a similar frequency over the resting visual cortex at the back of the scalp, the mu wave is found over the motor cortex, in a band approximately from ear to ear. A person suppresses mu wave patterns when he or she performs a motor action or, with practice, when he or she visualizes performing a motor action. This suppression is called desynchronization of the wave because EEG wave forms are caused by large numbers of neurons firing in synchrony.

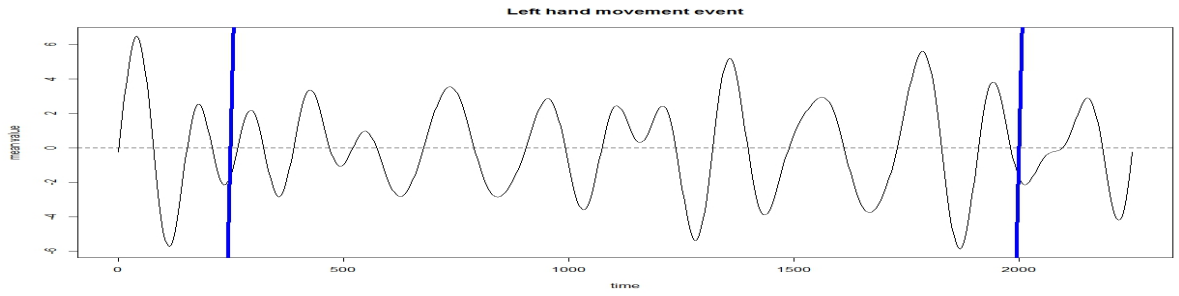


Figure 14: Left hand movement event. Smoothed average EEG data of motor cortex. Blue lines indicate begin of the trial and end.

Figure 14 shows smoothed and averaged EEG recording of single trial of left hand movement. Only channels (17-45) from motor cortex of the cerebral cortex which is involved in the

planning, control, and execution of voluntary movements. However, the pattern of Mu waves aren't clear enough.

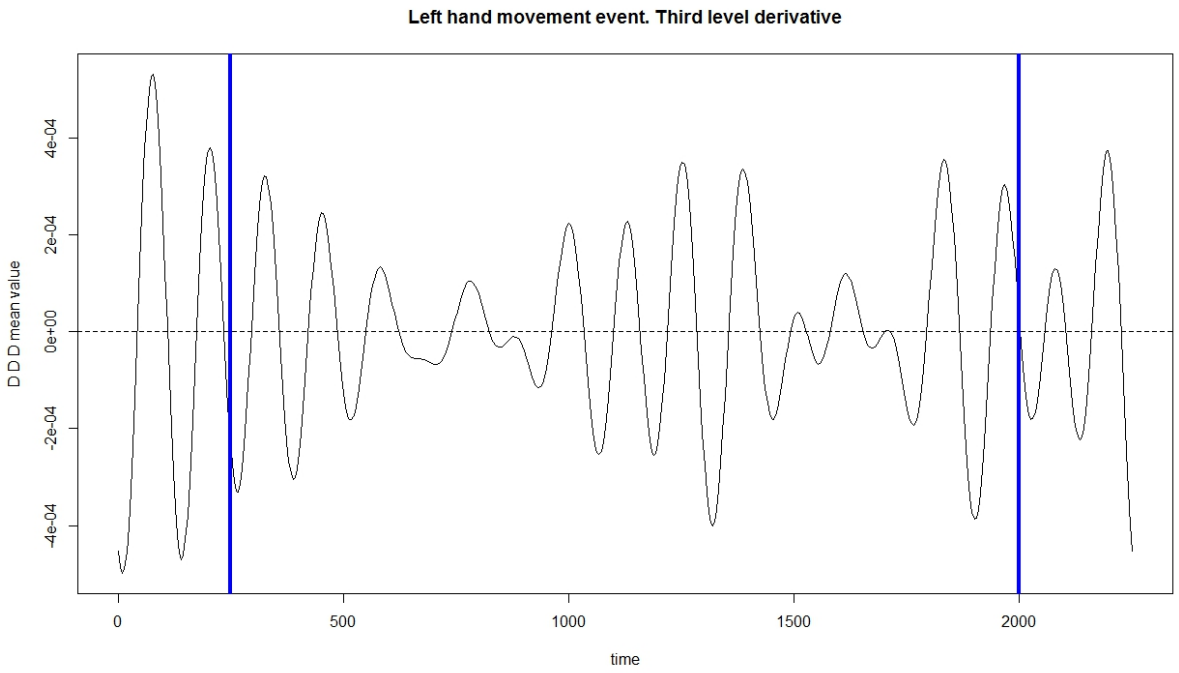


Figure 15: Left hand movement event. Third level derivative EEG data of motor cortex. Blue lines indicate begin of the trial and end.

Figure 15 shows third level derivative of this function is used to detect changes. We can observe amplitude changes from **t=500** which is actually "beep" signal indicating that the subject is starting to plan the movement. From **t=1000** the cross is showed on the screen which has no effect on Mu waves and starting from **t=1500** actual movement happens which we also can observe. These components can be roughly associated with cognitive processes known to occur around these times.

Classifying cognitive events

Since our data is labelled we can try to attempt to predict which cognitive event user has performed using classification via functional regression based and functional discriminant analysis. I choose two cognitive events left hand movement and right hand movement, split data into training set and validation set. Figure 16 shows decision boundary between classes.

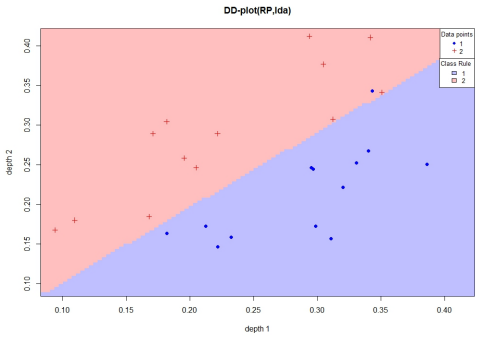


Figure 16: Left hand movement and right hand movement events decision boundary

After model has been built I have applied it on testing set. Table 1 shows confusion matrix.

-	Class 1	Class 2
Class 1	10	5
Class 2	4	9

Table 1: Confusion matrix