Preprocessing and Classification of ERD/ERS Signals

Florian Eichin

June 18, 2018

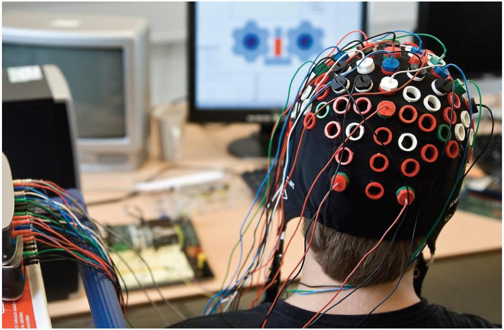
Freiburg University

Contents

How do we measure brain activity?

How do we measure brain activity?

- measure the electromagnetical effects of brain processes
- in this talk: via EEG-electrodes on the scalp



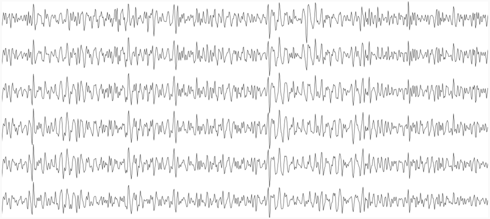
[Wolpaw, 2011]

Definition
With $x : \mathbb{N} \to \mathbb{R}^n$, we call $x(t)$ the **sample at timepoint** t.

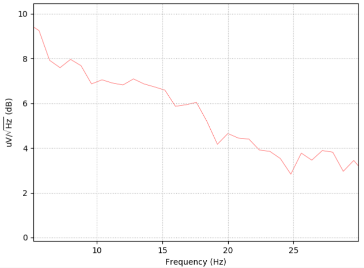
For $t, u \in \mathbb{N}$ the matrix

$$X = (x(t), x(t+1), ..., x(t+u)) \tag{1}$$

is the **epoch** at timepoint t with length u.



Frequency Spectrum

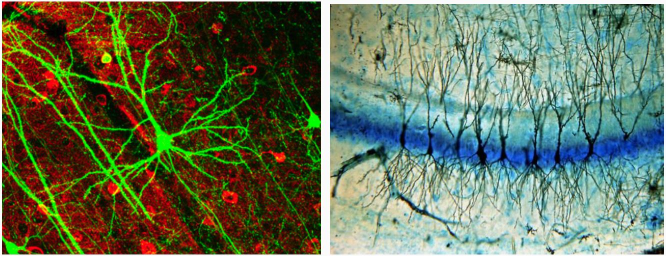


- Our signal consists of a composition electromagnetical waves
- it might be a good idea to have a look at the frequency-bands to gain information about the underlying processes

Oscillatory Analysis: Neurological Background

Pyramidal Cells

- largest contributor to the electromagnetical activity we can measure from the outside
- aligned orthogonally to the scalp on the surface of the brain



[W-CH Lee, 2006 and Aarhuus University, 2004]

Event Related (De)Synchronisations

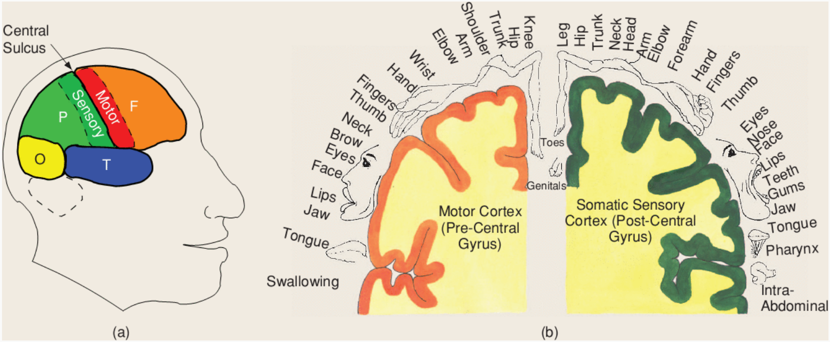
Whenever certain populations of neurons are inactive, they enter an **idle state**, synchronously oscillating at characteristic frequencies:

- α- / μ-rhythms: 8-15 Hz mostly found in the visual / motor and sensory cortex
- β-rhythm: 16-31 Hz, sensory and motor cortex

[Backyard Brains, 2014]

Event Related (De)Synchronisations

- parts of the brain are linked to certain tasks
- local ERD/ERS might help us classify between left and right movement



[Blankertz, 2008]

What are the challenges?

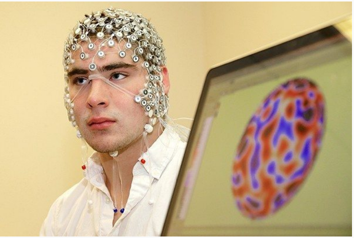
Noise and Artifacts

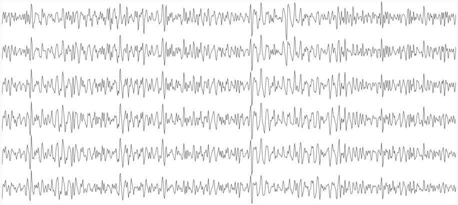
Our channels are usually not only picking up relevant neuronal activity, our signal is obscured by

- **Artifacts**:
    - environmental (power lines, electric devices, ...)
    - from BCI-hardware (amplifier, cables, electrodes, ...)
    - from the body (muscle contractions, blood flow, ...)
- **Noise** (from unrelated brain activity)

Dimensionality

- EEG-data has lots of **dimensions** (usually up to 128 channels) with lots of redundancy encoded
- the more dimensions we want to use for classification, the more data we need
- we usually only have the data from one session, that we can use for training our BCI





[University of Nebraska, 2013]

A simple approach: Laplace Filters

Filters

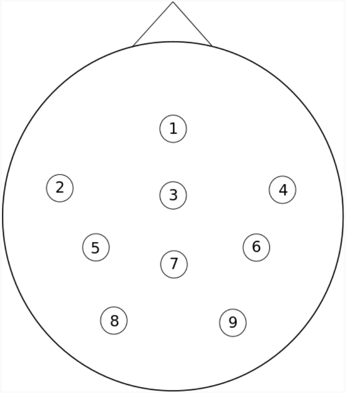
Definition
A **filter** is a linear projection $w : \mathbb{R}^n \to \mathbb{R}$ where n is the number of channels. We obtain the filtered signal by calculating:

$$x_w(t) = w \cdot x(t) \tag{2}$$

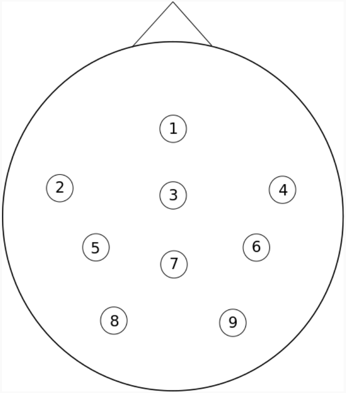
The w_i are called **weights**. Note that filters can be simultianiously applied by multiplying with a **filter matrix** W, where each row vector is a filter:

$$x_W(t) = W \cdot x(t) \tag{3}$$

Laplace Filters

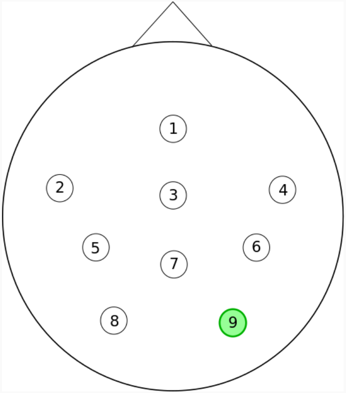


Laplace Filters



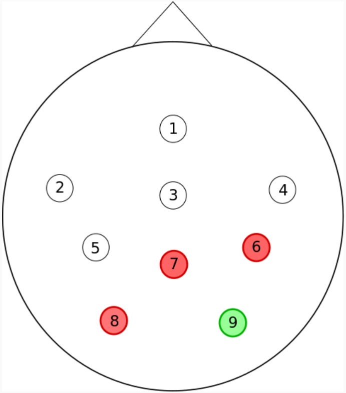
$$x(t) = (x_1(t), \ldots, x_9(t))$$

Laplace Filters



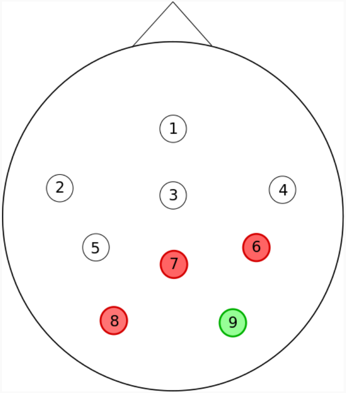
$$x(t) = (x_1(t), \ldots, x_9(t))$$

Laplace Filters



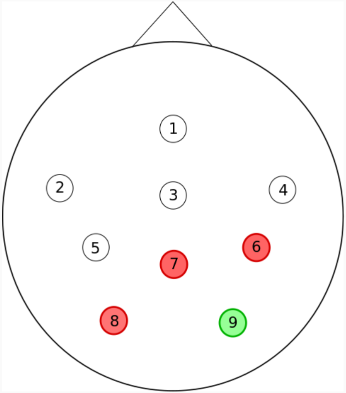
$$x(t) = (x_1(t), \ldots, x_9(t))$$

Laplace Filters



$$x(t) = (x_1(t), \ldots, x_9(t))$$
$$w = (0, \ldots, 0, -\tfrac{1}{3}, -\tfrac{1}{3}, -\tfrac{1}{3}, 1)$$

Laplace Filters

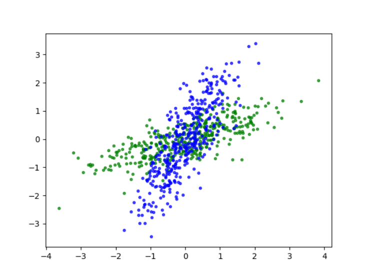


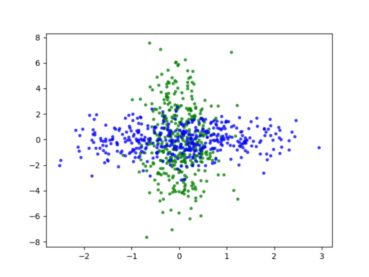
$$x(t) = (x_1(t), ..., x_9(t))$$
$$w = (0, ..., 0, -\tfrac{1}{3}, -\tfrac{1}{3}, -\tfrac{1}{3}, 1)$$
$$x_w(t) = w \cdot x(t) = -\tfrac{1}{3} \cdot (x_6 + x_7 + x_8) + x_9$$

More sophisticated: Common Spacial Patterns (CSP)

Common Spacial Patterns

- an algorithm, that learns filters from collections of data sample epochs $X_i \in \mathbb{R}^{NxC}$ for classes 1 and -1
- idea: **variance optimisation** between the two classes





Covariance Estimation

Let $X_1, ..., X_k \in \mathbb{R}^{NxC}$ be the samples for class c. Then we estimate the **covariance matrix** by

$$\Sigma_c = \frac{1}{k} \sum_{i=1}^{k} X_i^T X_i \tag{4}$$

Covariance Optimisation

- Let $\Sigma_+, \Sigma_- \in \mathbb{R}^{CxC}$ be the covariance matrix for class 1 / -1
- Subsequently find orthogonal vectors w_i that satisfy

$$w_i = argmax_{w \in \mathbb{R}^{n*}} \frac{w \Sigma_+ w^T}{w \Sigma_- w^T} \tag{5}$$

- the obtained W will project our data into a space, where the first (last) coordinate refers to the feature that has the highest variance for class 1 (-1)

Analytical Solution

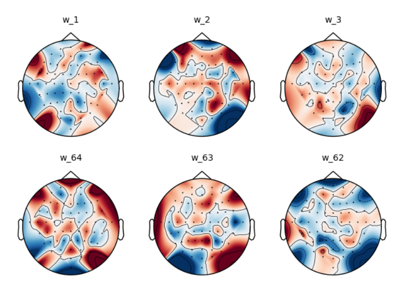
- A solution for W can also be obtained by solving

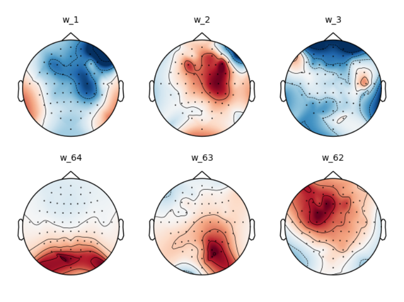
$$w\Sigma_{+} = \lambda \cdot w\Sigma_{-} \tag{6}$$

- the eigenvecors $w_1, ..., w_n$ with eigenvalues $\lambda_1 > \lambda_2 > ... > \lambda_n$ yield $W = (w_1, ..., w_n)^T$
- λ_i corresponds to the amount of variance in coordinate i in our surrogate space

Forward and Backward Model

Definition
W is also called **backward model**, $A = W^{-1}$ the **forward model**.





Relation to Principal Component Analysis

CSP is a **supervised generalization** of PCA for two classes.
If class -1 is is uncorrelated, CSP yields the same filter matrix as PCA for class 1.

Proof.
Let Σ_+ and Σ_- be the covariance matrices for class 1 and -1.
Then $\Sigma_- = I$ and we have

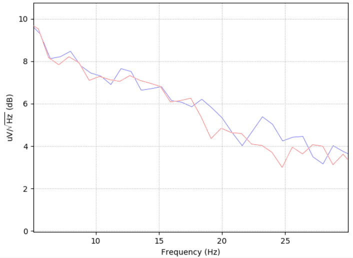
$$w_i = argmax_{w \in \mathbb{R}^{n*}} \frac{w\Sigma_+ w^T}{w\Sigma_- w^T} = argmax_{w \in \mathbb{R}^{n*}} \frac{w\Sigma_+ w^T}{ww^T} \tag{7}$$

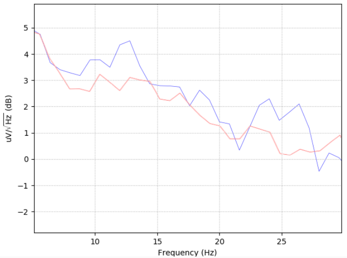
which is the optimization criterion for PCA. $\square$

How good is it?

Noise Reduction and Dimensionality

- CSP implements noise reduction and lower dimensionality
- α- and β-bands are more prominent in the target epochs after applying filters yielding an opportunity to discriminate from the non-target epochs





Application

Widely used among other methods such as PCA and ICA

- Easy to use and implement
- Low runtime, just a linear mapping

On the other hand:

- Hyperparameters (length, timepoint of sample epochs)
- supervised: need for labelled data
- need for training data beforehand

Sources

- Wolpaw, Brain-Computer Interfaces: principles and practice. Oxford Univ. Press, 2011.
- Blankertz B, Tomioka R, Lemm S, Kawanabe M, MÃ¼ller KR, Optimizing Spatial Filters for Robust EEG Single-Trial Analysis. IEEE Signal Process Mag, 25(1):41-56, 2008.
- mnetools.github.io/dev/auto_examples/decoding/plot_decoding_csp_eeg.html
- Schalk, G McFarland, DJ Hinterberger T, Birbaumer N, Wolpaw JR, BCI2000: A General-Purpose Brain-Computer Interface (BCI) System. IEEE TBME 51(6):1034-1043, 2004