Integrated Lecture
"Brain-Computer Interfacing"

Winter term 2019/20

Fachgebiet Neurotechnologie
Fakultät IV, TU Berlin
Prof. Dr. Benjamin Blankertz
benjamin.blankertz@tu-berlin.de

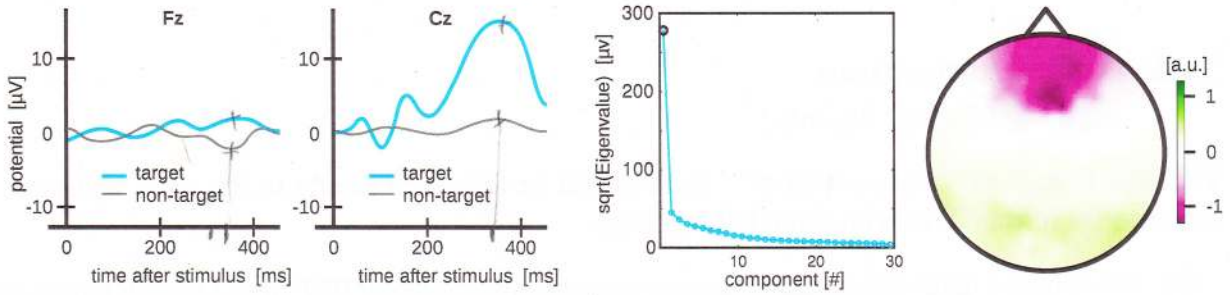
Written exam # 1 – 04.12.2019

In total there are 42 points plus 4 bonus points for the tasks. Further bonus points may be obtained for very well explained solutions. The time limit for the exam is 60 minutes.

Task 1 – Distributions of ERP features

(8 Points)

The left two graphs show the average ERPs for the classes *target* and *non-target* in the channels Fz and Cz for an ERP speller data set. We consider the distributions of the spatial features at time point $t = 340$ ms and assume that they are Gaussian distributions. The right two plots display the Eigenvalue spectrum and the map of the first Eigenvector for the *target* class and both can be assumed to look similar for the *non-target* class.



Make a sketch of the two-dimensional distributions of the potential in channels Fz and Cz at time point $t = 340$ ms for *targets* and *non-targets*.
Note: Some information about the distributions can only be estimated roughly or deduced from background knowledge about EEG. These aspects are evaluated qualitatively.

Task 2 – The Linear Model of EEG

(3+5+5* Points)

Consider an EEG data set $\mathbf{X}$ and two spatial filters $\mathbf{w}_1$ and $\mathbf{w}_2$ be given.

(a) Write down the formulas to get the components s_1 and s_2 that can be extracted from $\mathbf{X}$ with the two filters.

(b) A new synthetic EEG data set $\mathbf{Y}$ is to be generated, which consists just of the signals that are generated by the two components s_1 and s_2. Write the formulas down how this could be done.

(c) How can $\mathbf{Y}$ be generated such that s_1 has a contribution to channel Pz with a variance that is twice as high as the variance of the contribution of s_2 to that channel. (bonus task)

Task 3 – ERP Classification using Linear Discriminant Analysis (LDA)

(5+7+4 Points)

(a) Assume two classes with the Gaussian distributions $\mathcal{N}(\mu_1, \Sigma)$ and $\mathcal{N}(\mu_2, \Sigma)$ be given. Write down the formula of the two parameters $\mathbf{w}$ and b of the LDA classifier for these distributions and the LDA mapping of a test sample $\mathbf{x}$ to the class labels 1 and 2 (classification rule).

(b) We consider the following change in the means of the two classes for $\alpha \in \mathbb{R}$:

$$\hat{\mu}_1 = \mu_1 + \alpha(\mu_1 - \mu_2)$$
$$\hat{\mu}_2 = \mu_2 - \alpha(\mu_1 - \mu_2)$$

Compare the LDA classifier for this new distributions (Σ staying the same) with that of part a): How does the weight vector $\mathbf{w}$, the bias b and the classifier decision change, depending on α? A sketch (in two dimensions) might help (and a graphical solution with explanation may be awarded some points, if a mathematical solution is missing).

(c) Sketch the distributions of two classes (in two dimensions) such that the separation line rotates roughly about 40 degrees around the origin for an LDA shrinkage classifier when γ goes from 0 to 1. The distributions should be plausible for ERP features.

Task 4 – Quick questions

(2+2+2+2+2 Points)

This task consist of 'quick questions'. Each should be answered briefly to the point with one sentence (or formula), or with some bullet points.

(a) Describe the main difference between a *standard* and a *deviant* ERP in an Oddball Paradigm experiment (basis for ERP speller).

(b) Name one case in which filters and patterns (of the linear model) coincide and shortly explain the reason.

(c) What is the size of the Eigenvectors $\tilde{d}_i$ of the shrunk covariance matrix $\tilde{\Sigma}(\gamma)$, given the Eigenvalues d_i of the empirical covariance matrix Σ and parameter γ? (formula)

(d) A component is extracted from some EEG data $\mathbf{X}$ using a spatial filter $\mathbf{w}$. What does it tell you about the extracted component $\mathbf{w}^\top\mathbf{X}$, if it has a standard deviation which is larger than 100 and why?

(e) Adding 2 frontal channels (e.g. F3 and F4) drastically improves classification of an ERP based BCI setup for all participants compared to a setup which uses central and parietal channels only. What could be the reason?

Good luck!