Brain-Computer Interfacing

WS 2018/2019 – Vorlesung #11

Benjamin Blankertz

Lehrstuhl für Neurotechnologie, TU Berlin

benjamin.blankertz@tu-berlin.de

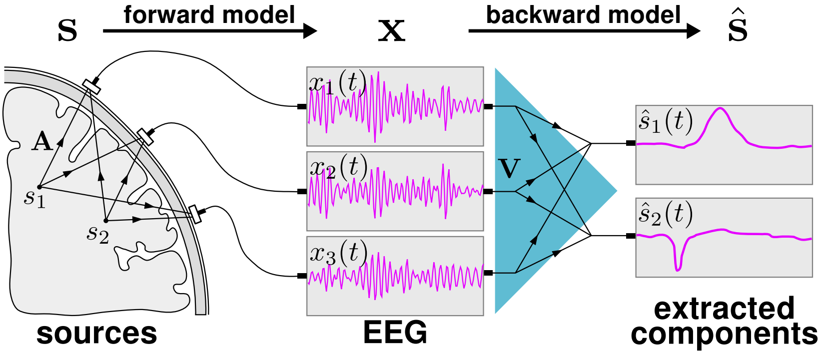
16 · Jan · 2019

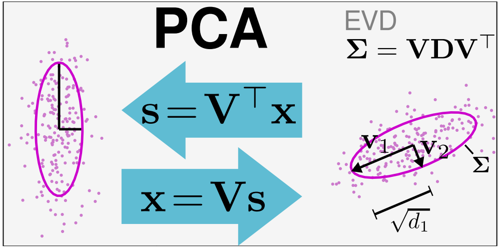
Today's Topics

Two novel ways to determine spatial filters:

- Temporal Decorrelation Seperation (TDSEP)
- Spatio-Spectral Decomposition (SSD)

Recap: Linear (Source) Model of EEG and PCA





EEG cleaning

$$\tilde{\mathbf{x}} = \mathbf{V}_0 \mathbf{V}_0^\top \mathbf{x}$$

$\mathbf{V}_0$: good EVs of $\mathbf{V}$

artifacts are discarded

Outlook

Methods that determine spatial filters are also called decomposition methods, since they decompose the mixed signals into (hopefully) meaningful components.

CSP analysis decomposes the multi-variate EEG signal into components, that are optimized wrt. a classification task. This does not imply that each component relates to a specific source in the brain.

Likewise, with PCA we cannot expect to find brain sources, since the orthogonality constraint is not neurophysiologically plausible.

Today, we will have a look at two further signal decomposition methods, that fit into the framework of linear models:

- Temporal Decorrelation Seperation (TDSEP)
- Spatio-Spectral Decomposition (SSD)

The Assumption of Independence

Independent Component Analysis (ICA) is a method used in numerous applications.

Applying ICA to EEG data often aims at the identification of sources in the brain. It relies on the following assumptions:

- the linear model of EEG holds
- sources are independent
- the number of strong sources is (less or) equal to the number of channels

ICA determines spatial filters $\mathbf{W}$ to extract the estimated source signal with the backward model.

Obstacle: Attaining independence from limited, finite signals is infisible.

In this lecture, we will only discuss one variant that is particularly simple to implement. Others rely on more complicated statistical concepts and optimization methods [Comon 1994; Bell & Sejnowski 1995; Hyvärinen et al, 2001].

Statistical Independence und Uncorrelatedness

Two random variables X and Y are independent, if knowing an observed value of X does not affect the probability distribution of Y.

In this case, for all functions $g(X)$ and $h(Y)$ (absolutely integrable wrt. X resp. Y) the following equality holds:

$$E\langle g(X)h(Y)\rangle = E\langle g(X)\rangle E\langle h(Y)\rangle$$

Choosing g and h as identity function, we obtain $E\langle XY\rangle = E\langle X\rangle E\langle Y\rangle$, which is the definition of X and Y being uncorrelated.

Independence implies uncorrelatedness, but the opposite is not true. Independence is a much stronger property.
[Proof: If $X \sim \mathcal{N}(0,1)$, $Y = X^2$ then $\mathrm{cov}(X,Y) = EX^3 = 0$, i.e., X, Y are uncorrelated, but they are obviously not independent.]

Technical Note

According to the roots of ICA, we will use a statistical notation in this lecture (treat time series $\mathbf{x}(t)$ as random variable):

$$E\langle \mathbf{x}(t) \rangle = \frac{1}{T} \sum_{t=1}^{T} \mathbf{x}(t) = \mathrm{mean}\, \langle \mathbf{x}(t) \rangle_{t=1,\ldots,T}$$

In the following we assume that the signals have zero mean.

Temporal Correlaton

The (functional) independence of sources $s_i(t)$ and $s_j(t)$ implies that also time shifted versions of those sources are uncorrelated, i.e.,

$$E\langle s_i(t)s_j(t+\tau)\rangle = 0 \quad \text{for all } i, j \text{ with } i \neq j$$

and any time shift τ. In matrix notation, that is

$$E\langle \mathrm{s}(t)\mathrm{s}(t+\tau)^\top\rangle \quad \text{is diagonal.}$$

*This observation gives rise to the approach of approximating independent sources by forcing **time shifted versions** of the extracted components to be **uncorrelated**.*

We define the symmetrized time-lagged covariance matrix of the multi-variate time signal $\mathrm{s}(t)$ for time shift τ as

$$\Sigma_{\mathrm{s}}(\tau) := \frac{1}{2}\left(E\langle \mathrm{s}(t)\mathrm{s}(t+\tau)^\top\rangle + E\langle \mathrm{s}(t+\tau)\mathrm{s}(t)^\top\rangle\right)$$

Temporal Decorrelaton Separation (TDSEP)

The **Temporal Decorrelaton Separation** (TDSEP) algorithm [Ziehe & Müller, 1998] aims at a demixing

$$\mathbf{y}(t) = \mathbf{W}^\top \mathbf{x}(t)$$

such that the time-lagged covariance matrices of the extracted components $\mathbf{y}$ are diagonal (thereby forcing the uncorrelatedness as claimed on the previous slide):

$$\Sigma_{\mathbf{y}}(\tau) = \mathbf{W}^\top \Sigma_{\mathbf{x}}(\tau) \mathbf{W} \quad \text{is diagonal} \tag{1}$$

It is not practical to optimize $\mathbf{W}$ wrt. *all* time lags τ. Rather, a set of time lags $\{\tau_1, \tau_2, \ldots \tau_q\}$ is chosen, and the demixing matrix $\mathbf{W}$ is determined such that eqn (1) is fulfilled jointly for all those τ as good as possible (approximate joint diagonalization).

TDSEP Technically

For only two time lags $\{\tau_1, \tau_2\}$, the diagonalization of the corresponding time-lagged covariance matrices can be accomplished exactly and also in a simple way (for more only approximately and it is more involved).

Determine the demixing matrix $\mathbf{W}$ (typically not orthogonal) such that

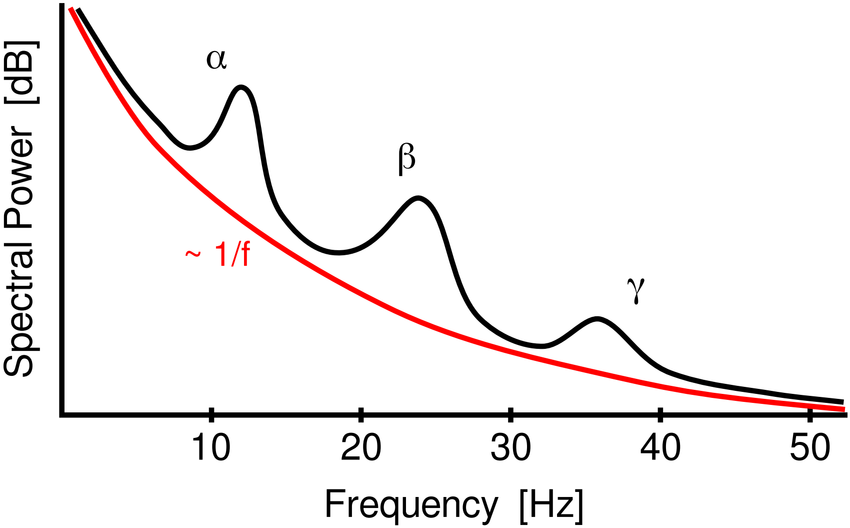
$$\mathbf{W}^\top \mathbf{\Sigma}_{\mathbf{x}}(\tau_1)\mathbf{W} = \mathbf{D} \quad \text{for a diagonal matrix } \mathbf{D} \quad \text{and} \tag{2}$$

$$\mathbf{W}^\top \mathbf{\Sigma}_{\mathbf{x}}(\tau_2)\mathbf{W} = \mathbf{I}. \tag{3}$$

As for CSP, the filter matrix $\mathbf{W}$ and $\mathbf{D}$ can be obtained by a generalized Eigenvalue decomposition.

Second Decomposition Method: SSD

Reminder: Spectrum of Macroscopic Brain Activity



A Novel Decomposition Method for Oscillatory Components

Interesting oscillatory activity in the EEG corresponds to peaks in the spectra that protrude from the $1/f$ background spectrum.

Therefore, it is of interest to extract such components for a given frequency band.

A Novel Decomposition Method for Oscillatory Components

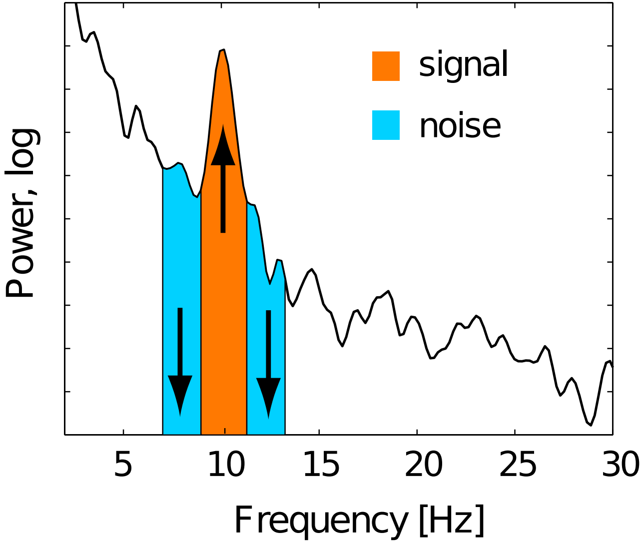
In the next part, we will discuss a decomposition technique which can be used to extract components with a predefined frequency band.

In contrast to CSP, the new method does not require two conditions in which the signal component of interest is modulated.

In contrast to general ICA algorithms (and TDSEP with more than two time shifts), this new method can be calculated more efficiently.

Idea for the Extraction Criterium

The idea is to extract a component that has strong spectral power in a predefined frequency band, but little in the flanking frequencies.



Spatio-Spectral Decomposition (SSD)

We assume any EEG component $y(t)$ to be a linear superposition of a target signal $z(t)$ which has a sharp peak at frequency f and noise $n(t)$:

$$y(t) = z(t) + n(t)$$

The goal of the **Spatio-Spectral Decomposition** (SSD, [Nikulin et al, 2011]) is to extract by spatially filtering the EEG signal $\mathbf{x}(t)$ a component $y(t)$ that has a high signal-to-noise ratio at a central frequency f:

$$\mathrm{SNR}_y(f) = \frac{P_z(f)}{P_n(f)}$$

Here, $P_z(f)$ is the spectral power of time series z at frequency f.

Spatio-Spectral Decomposition (SSD) – Part 2

We assume the target signal $z(t)$ and the noise $n(t)$ to be uncorrelated. This implies

$$P_y(f) = P_z(f) + P_n(f).$$

Furthermore, we assume that the power of signal z is negligible at frequencies $f \pm \Delta f$, and the power of the noise n at frequency f can be linearly approximated from the power of the component at flanking frequencies:

$$P_n(f) \approx \frac{P_n(f-\Delta f) + P_n(f+\Delta f)}{2} \approx \frac{P_y(f-\Delta f) + P_y(f+\Delta f)}{2}$$

Taken together we obtain

$$\begin{aligned} \frac{P_y(f)}{1/2(P_y(f-\Delta f) + P_y(f+\Delta f))} &\approx \frac{P_z(f) + P_n(f)}{P_n(f)} = \frac{P_z(f)}{P_n(f)} + 1 \\ &= \mathrm{SNR}_y(f) + 1. \end{aligned} \tag{4}$$

Spatio-Spectral Decomposition (SSD)

In order to maximize the SNR at frequency f according to eqn (4), we define the following covariance matrices:

$\boldsymbol{\Sigma}_{\mathbf{x}|f}$ calculated from signal $\mathbf{x}$, band-pass filtered around f

$\boldsymbol{\Sigma}_{\text{noise}} := {}^1\!/_2 \left(\boldsymbol{\Sigma}_{\mathbf{x}|f-\Delta f} + \boldsymbol{\Sigma}_{\mathbf{x}|f+\Delta f} \right)$

Extracting one component of SSD can be accomplished by the following maximization

$$\mathbf{w} := \underset{\mathbf{w}}{\text{argmax}} \frac{\mathbf{w}^\top \boldsymbol{\Sigma}_{\mathbf{x}|f} \mathbf{w}}{\mathbf{w}^\top \boldsymbol{\Sigma}_{\text{noise}} \mathbf{w}}$$

As for CSP, this optimization can be solved as a generalized Eigenvalue problem wrt. the matrices $\boldsymbol{\Sigma}_{\mathbf{x}|f}$ and $\boldsymbol{\Sigma}_{\text{noise}}$.

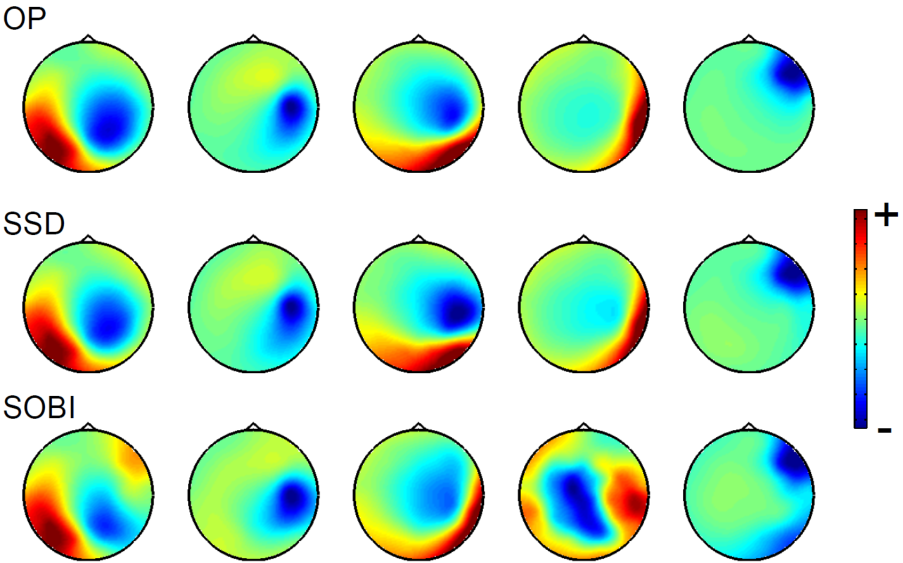
SSD: Simulated Signals for an Test Example

Simulated 64-channel EEG signals: Signals of interest were generated at a frequency range of 10–12 Hz from 5 dipoles in a realistic head model. Noise was produced with a $1/f$ type using 500 uncorrelated dipoles.

For SSD, the frequency range 10–12 Hz has been used for the target signal. As flanking frequency ranges, 8–9 and 13–14 Hz have been used.

For comparison with SSD, some ICA algorithms have been tested. Most of them did not converge, probably due to the similarity of the patterns of simulated sources. The one algorithm that did provide stable results was SOBI [Belouchrani et al, 1997] which is equivalent to TDSEP. It was used with 50 time-delayed covariance matrices with lags 5, 15, . . . , 495 ms.

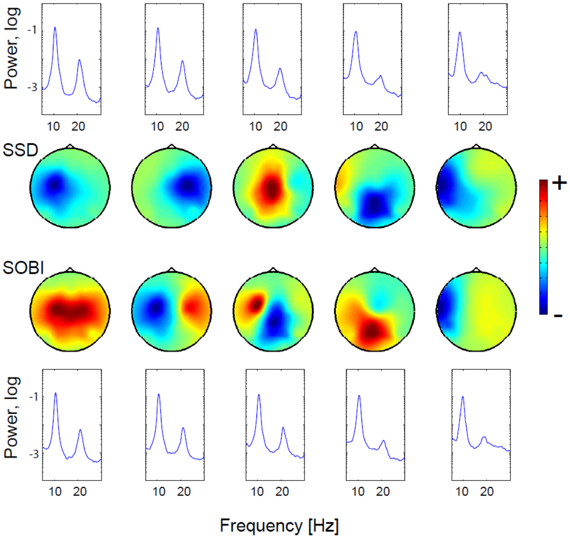
SSD: Resulting Patterns for Simulated Signals (1)



OP: Original Patterns

SSD: Results on Real Data (exemplary participant)

EEG signals have been acquired from seven participants at rest.
SSD was applied to extract components for the 8–13 Hz frequency band.



Lessons Learnt

After this lecture you should

- be familiar with the general concept of ICA,
- be capable of implementing the temporal decorrelation (TDSEP),
- know the spatio-spectral decomposition (SSD) method.

References I

- Bell, A. J. and Sejnowski, T. J. (1995).
  An information-maximization approach to blind separation and blind deconvolution.
  Neural Comput, 7(6):1129–1159.

- Belouchrani, A., Abed-Meraim, K., Cardoso, J.-F., and Moulines, E. (1997).
  A blind source separation technique using second-order statistics.
  IEEE Trans Signal Process, 45(2):434–444.

- Comon, P. (1994).
  Independent component analysis, a new concept?
  Signal Processing, 36(3):287–314.

- Hyvärinen, A., Karhunen, J., and Oja, E. (2001).
  Independent Component Analysis.
  Wiley.

- Nikulin, V. V., Nolte, G., and Curio, G. (2011).
  A novel method for reliable and fast extraction of neuronal EEG/MEG oscillations on the basis of spatio-spectral decomposition.
  NeuroImage, 55:1528–1535.

- Ziehe, A. and Müller, K.-R. (1998).
  TDSEP – an efficient algorithm for blind separation using time structure.
  In Niklasson, L., Bodén, M., and Ziemke, T., editors, *Proc. of the 8th International Conference on Artificial Neural Networks, ICANN'98*, Perspectives in Neural Computing, pages 675 – 680, Berlin. Springer Verlag.