Brain-Computer Interfacing
WS 2018/2019 – Lecture #06

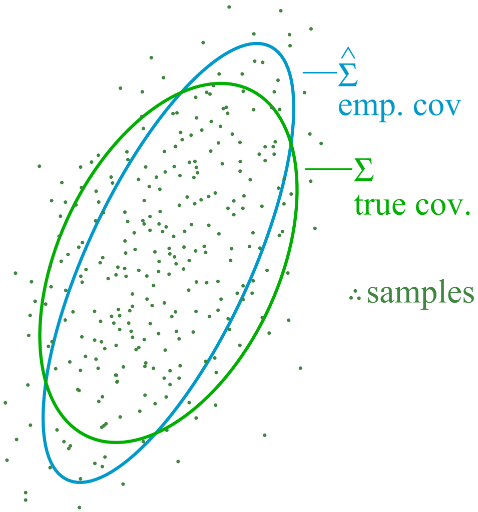
Benjamin Blankertz

Lehrstuhl für Neurotechnologie, TU Berlin

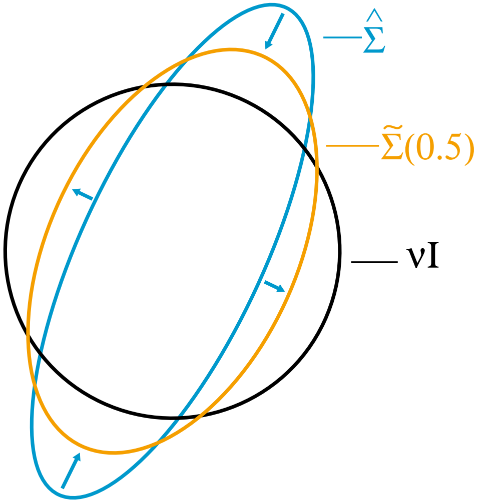
benjamin.blankertz@tu-berlin.de

21 · Nov · 2018

Cartoon in 2D:

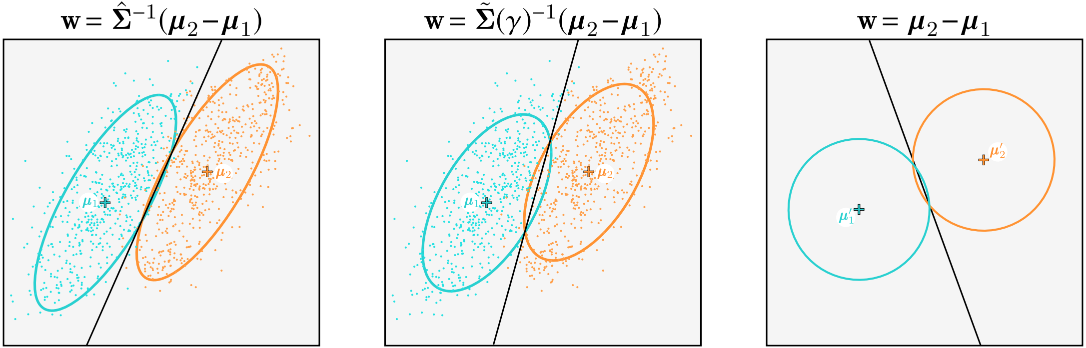


Shrinkage:



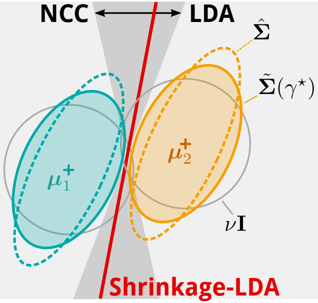
$$\tilde{\mathbf{\Sigma}}(\gamma) = (1-\gamma)\hat{\mathbf{\Sigma}} + \gamma\nu\mathbf{I}$$

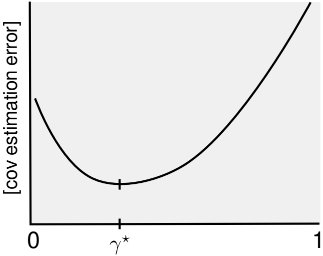
Recap: NCC, LDA and Shrinkage-LDA



same signals of interest $(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2)$ – different spatial structure of noise $(\boldsymbol{\Sigma})$
or in another view: different belief in the empirical covariance matrix.

The amount of shrinkage (γ) relates to the 'believe' in the estimation of the noise.





Shrinkage-LDA hyperplane is defined by:

$$\mathbf{w} := \tilde{\boldsymbol{\Sigma}}(\gamma^\star)^{-1}\,(\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1)$$

$$\tilde{\boldsymbol{\Sigma}}(\gamma) := (1-\gamma)\hat{\boldsymbol{\Sigma}} + \gamma\nu\mathbf{I}$$

Calculate optimal shrinkage parameter $\gamma^\star$ analytically:

[Ledoit & Wolf 2004], [Schäfer & Strimmer 2005]

$$\gamma^\star = \underset{\gamma}{\text{argmin}} \|\tilde{\boldsymbol{\Sigma}}(\gamma) - \boldsymbol{\Sigma}\|_F^2$$

$$= \frac{K}{(K-1)^2}\,\frac{\sum_{i,j=1}^d \mathrm{var}_k\,(\mathbf{Z}^k)_{ij}}{\sum_{i,j=1}^d s_{ij}^2} \quad \text{with}$$

$$\mathbf{Z}^k = (\mathbf{x}_k - \hat{\boldsymbol{\mu}})\,(\mathbf{x}_k - \hat{\boldsymbol{\mu}})^\top \quad \text{and}$$

$$s_{ij} = \left(\hat{\boldsymbol{\Sigma}} - \nu\mathbf{I}\right)_{ij}$$

Today's Topics

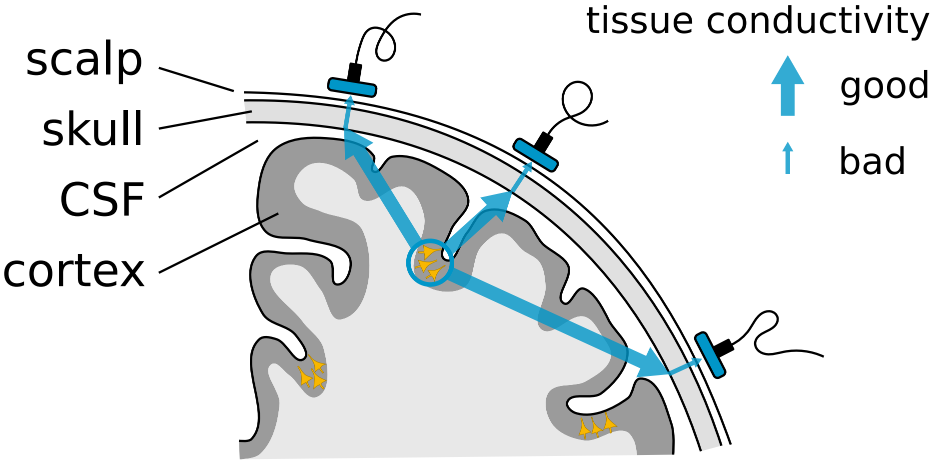
- The linear model of EEG
  - forward model, spatial patterns
  - backward model, spatial filters
  - correspondence between forward and backward model

- PCA in the framework of the linear model of EEG
  - properties of PCA factors
  - artifact removal with PCA

Linear Model of EEG

Next, we will introduce a linear model which represents the electrophysics of EEG. Although it is oversimplifying, this model it useful for understanding.

Remember the issue of volume conduction discussed in the first lecture:

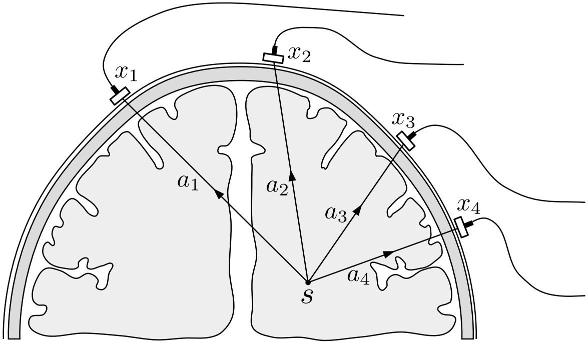


Linear Model of EEG: Propagation of Electrical Activity

- **Assumption:** The contribution of a current source $s(t)$ to the scalp potentials (at P sensors) $\mathbf{x}(t) = [x_1, \ldots, x_P]^\top$ is linear in $s(t)$:

$$\mathbf{x}(t) = [a_1 s(t), \ldots, a_P s(t)]^\top = \mathbf{a}\, s(t)$$

- The proportionality factors in vector $\mathbf{a}$ are typically unknown and depend on the spatial distribution and orientation of the current sources and the conductivity distribution of the anatomy [Parra et al, 2005].



Towards to Full Model

- Now, we consider K sources $\mathrm{s}_k(t)$ with distribution vectors $\mathbf{a}_1, \ldots, \mathbf{a}_K \in \mathbb{R}^P$.
- In the linear model, potentials are assumed to be additive. Defining the matrix $\mathbf{A}$ as being composed of the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_K$ (i.e., $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_K]$), the **forward model** is

$$\mathbf{x}(t) = \mathbf{a}_1 s_1(t) + \cdots + \mathbf{a}_K s_K(t) = \mathbf{A}\,\mathrm{s}(t)$$

- Contributions not captured by this model are considered as noise, $\mathbf{n}(t)$, typically assumed to be Gaussian distributed with mean 0.
- This gives a simple linear model representing the electrophysics of EEG:

$$\mathbf{x}(t) = \mathbf{A}\,\mathrm{s}(t) + \mathbf{n}(t)$$

Linear Model of EEG: Forward Model

The propagation of source signals $\mathbf{s}(t)$ to EEG sensors measured as $\mathbf{x}(t)$ is formalized in the **forward model**:

$$\mathbf{x}(t) = \mathbf{A}\,\mathbf{s}(t) + \mathbf{n}(t)$$

Each column of $\mathbf{A} \in \mathbb{R}^{P \times K}$ is a **spatial pattern**, which specifies how one source is propagated to the EEG sensors.

P: # of sensors, K: # of sources, $\mathbf{n}(t)$: noise

Note: The forward model does not necessarily model all sources of the brain. Sometimes it is restricted to only a few. Unmodelled sources are accounted for in the noise term.

How to Recover Sources

The linear model makes an assumption on how signals are mixed in the EEG.

For illustration, we discuss an approaches how to extract source signals from the mixture.

Signal Subtraction

Let us assume a simple example. Suppose we measure contributions from two sources s_1, s_2 at our sensors $\mathbf{x} = [x_1,\ x_2]^\top$ with a different weighting.

$$x_1(t) = \alpha s_1(t) + \beta s_2(t) + n_1(t)$$
$$x_2(t) = \gamma s_1(t) + \delta s_2(t) + n_2(t)$$

Using a suitable weighted average, we can cancel out one or the other source (ignoring the noise here):

$$\begin{aligned} \delta x_1(t) - \beta x_2(t) &= \delta\alpha s_1(t) + \cancel{\delta\beta s_2(t)} - \beta\gamma s_1(t) - \cancel{\beta\delta s_2(t)} &&= (\alpha\delta - \beta\gamma) s_1(t) \\ -\gamma x_1(t) + \alpha x_2(t) &= \cancel{-\gamma\alpha s_1(t)} - \gamma\beta s_2(t) + \cancel{\alpha\gamma s_1(t)} + \alpha\delta s_2(t) &&= (\alpha\delta - \beta\gamma) s_2(t) \end{aligned}$$

If $\frac{\alpha}{\beta} \neq \frac{\gamma}{\delta}$ (i.e. the sources are not collinear), then $\mathbf{w}_1^\top \mathbf{x}$ recovers s_1 for $\mathbf{w}_1^\top = [\delta,\ -\beta]$ and $\mathbf{w}_2^\top \mathbf{x}$ recovers s_2 for $\mathbf{w}_2^\top = [-\gamma,\ \alpha]$.

But factors $\alpha, \beta, \gamma, \delta$ are unknown. However, source separation algorithms (like ICA) may help to recover source activity.

Linear Model of EEG: Backward Model

The general recovering of sources is formalized in the **backward model**:

$$\hat{\mathbf{s}}(t) = \mathbf{W}^\top \mathbf{x}(t) \qquad \text{or equivalently} \qquad \hat{\mathbf{S}} = \mathbf{W}^\top \mathbf{X}$$

Each row of $\mathbf{W}^\top \in \mathbb{R}^{K \times P}$ is a **spatial filter**, that is a weight vector to be applied to the P-dimensional EEG signal.

P: # of sensors, K: # of sources

Note: The backward model is not only used for the purpose of recovering actual sources in the brain. It may serve as well for extracting components that have certain properties like good discriminability between conditions.

Mathematical Background

Generalized Inverses: Let $\mathbf{A} \in \mathbb{R}^{M\times N}$ be a matrix of full rank, i.e., M or N. Then one of the following one-sided inverses exists:

- for $M \geq N$ the **left inverse**: $\mathbf{A}_L^{-1} = (\mathbf{A}^\top\mathbf{A})^{-1}\mathbf{A}^\top$
- for $M \leq N$ the **right inverse**: $\mathbf{A}_R^{-1} = \mathbf{A}^\top(\mathbf{A}\mathbf{A}^\top)^{-1}$

The Moore-Penrose **pseudoinverse** $\mathbf{A}^+$ of $\mathbf{A}$ is defined as left or right inverse, depending on the values of M and N.

The pseudoinverse has the following properties:

$$\operatorname*{argmin}_{\mathbf{Y}} \|\mathbf{YZ} - \mathbf{X}\|^2 = \mathbf{XZ}^+ \tag{1}$$

$$\operatorname*{argmin}_{\mathbf{Z}} \|\mathbf{YZ} - \mathbf{X}\|^2 = \mathbf{Y}^+\mathbf{X} \tag{2}$$

For source signals $\mathbf{S} \in \mathbb{R}^{P\times T}$, we assume that there are more time points T than sources P. In this case, $\mathbf{S}^+ = \mathbf{S}_R^{-1} = \mathbf{S}^\top(\mathbf{S}\mathbf{S}^\top)^{-1}$ and

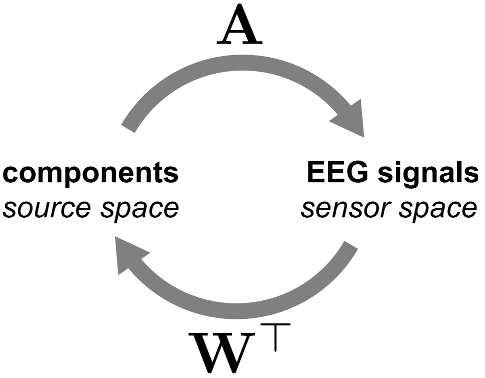
$$\operatorname*{argmin}_{\mathbf{A}} \|\mathbf{X} - \mathbf{AS}\|^2 = \mathbf{XS}^+ = \mathbf{XS}^\top(\mathbf{SS}^\top)^{-1}$$

Correspondence of Forward and Backward Model

The forward model is characterized by a matrix $\mathbf{A}$, which has for each source a projection pattern in its columns (and for each sensor a weight vector in its rows representing the mixing of the sources).

The backward model is characterized by a matrix $\mathbf{W}$, which has for each source a spatial filter in its columns.

Next, we will establish a correspondence between forward and backward model.



Finding Patterns for given Filters of a Discriminative Model

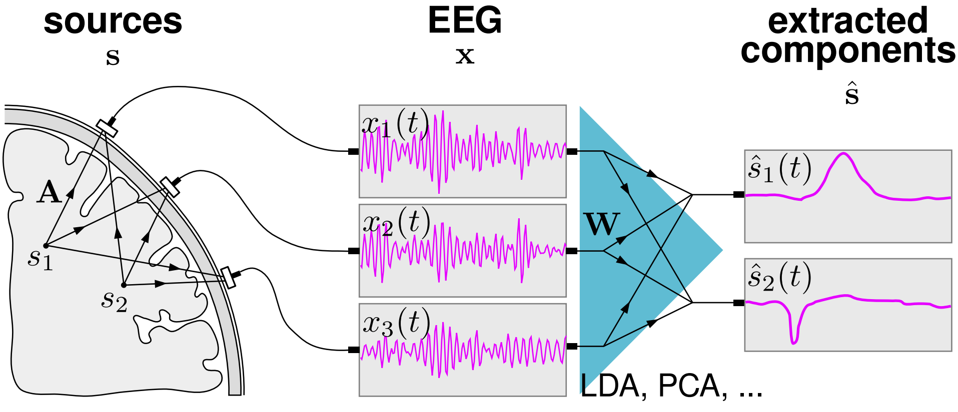
Let a filter matrix $\mathbf{W}$ be given and define $\mathbf{S} = \mathbf{W}^\top \mathbf{X}$. Assuming the noise-free case, we obtain the matrix of corresponding patterns $\hat{\mathbf{A}}$ by

$$\begin{aligned}
\hat{\mathbf{A}} &= \underset{\mathbf{A}}{\operatorname{argmin}} \|\mathbf{X} - \mathbf{A}\mathbf{S}\|^2 \\
&= \mathbf{X}\mathbf{S}^\top(\mathbf{S}\mathbf{S}^\top)^{-1} \\
&= \mathbf{X}\mathbf{X}^\top\mathbf{W}(\mathbf{S}\mathbf{S}^\top)^{-1} \\
&= \mathbf{\Sigma}_\mathbf{x}\mathbf{W}\mathbf{\Sigma}_\mathbf{s}^{-1}
\end{aligned}$$

For the general case with noise, the same result is obtained, but in a more complicated way, see [Haufe et al, 2014].

Conversely, given a pattern matrix $\mathbf{A}$, the matrix of corresponding filters $\mathbf{W}$ is determined by (provided that $\mathbf{\Sigma}_\mathbf{x}$ is invertible)

$$\mathbf{W} = \mathbf{\Sigma}_\mathbf{x}^{-1}\mathbf{A}\mathbf{\Sigma}_\mathbf{s}$$

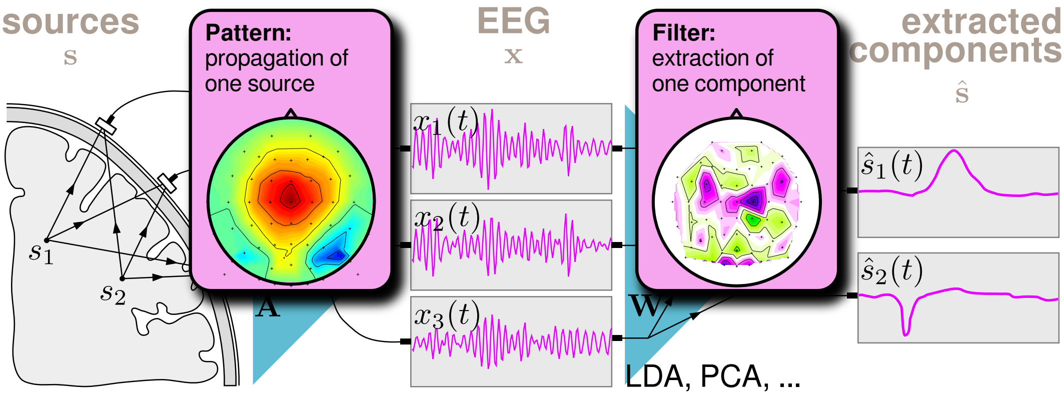


$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t)$ **forward model**

$\hat{\mathbf{s}}(t) = \mathbf{W}^\top \mathbf{x}(t)$ **backward model**

Each column of $\mathbf{A}$ is a spatial **pattern**: propagation of a source to sensors
Each row of $\mathbf{W}^\top$ is a spatial **filter**: weighting of EEG channels.

Linear Model of EEG



$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t)$$

forward model

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

backward model

Each column of $\mathbf{A}$ is a spatial **pattern**: propagation of a source to sensors
Each row of $\mathbf{W}^\top$ is a spatial **filter**: weighting of EEG channels.

Notes on the Linear Model

$$\mathbf{x}(t) = \mathbf{A}\,\mathbf{s}(t) + \mathbf{n}(t)$$

- The EEG signals 'live' in the sensor space. Their unit is μV.
- The components (source signals) 'live' in the source space. They have no unit (label, e.g., with [a.u.] for arbitrary unit).
- **The order of sources is undetermined.** Permuting the sources $\mathbf{s}$ and columns (patterns) of $\mathbf{A}$ in the same way amounts to the same model.
- **The scaling of the sources is undetermined.** Scaling a source and scaling inversely the corresponding column of $\mathbf{A}$ amounts to the same model.

PCA as a Tool for Investigating EEG

Next, we will see how Principal Component Analysis (PCA) can be used in the framework of the linear model for EEG analysis.

Principal Component Analysis (PCA)

Let $\mathbf{x}(t)$ be a multivariate time series, e.g., EEG signals. Applying PCA to $\mathbf{x}(t)$ means

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

where $\mathbf{V} \in \mathcal{O}(n)$ is the matrix of Eigenvectors $[\mathbf{v}_1, \ldots \mathbf{v}_P]$ that is obtained from the Eigenvalue decomposition of the covariance matrix $\hat{\mathbf{\Sigma}}$ of $\mathbf{x}(t)$

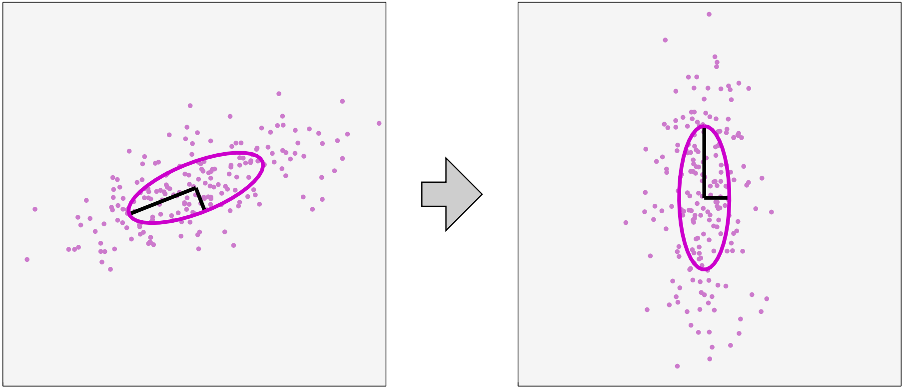
$$\hat{\mathbf{\Sigma}} = \mathbf{V}\mathbf{D}\mathbf{V}^\top \qquad \text{with diagonal matrix } \mathbf{D} \tag{3}$$

In particular, we get one component $y_i(t)$ (spatially filtered time series) for each Eigenvector (spatial filter) $\mathbf{v}_i$:

$$y_i(t) = \mathbf{v}_i^\top \mathbf{x}(t)$$

PCA in Feature Space

As seen in a previous lecture, a multiplication of a vector with the matrix $\mathbf{V}$ of Eigenvectors (from the left) corresponds to a rotation (plus possibly mirroring) which maps the Eigenvectors to the coordinate axes:



Properties of PCA Factors

- For simplicity, we assume $\mathbf{x}(t)$ to have zero mean.
- Then $\boldsymbol{\Sigma}_{\mathbf{X}} = \frac{1}{T-1}\mathbf{X}\mathbf{X}^\top$.

Similar to the calculation wrt whitening, we get:

$$\begin{aligned}
\boldsymbol{\Sigma}_{\mathbf{Y}} &= \boldsymbol{\Sigma}_{\mathbf{V}^\top\mathbf{X}} \\
&= \mathbf{V}^\top\boldsymbol{\Sigma}_{\mathbf{X}}\mathbf{V} \\
&= \mathbf{V}^\top(\mathbf{V}\mathbf{D}\mathbf{V}^\top)\mathbf{V} \\
&= \mathbf{D}
\end{aligned} \tag{4}$$

- The extracted components are pairwise **uncorrelated** (non-diagonal elements of (4) are zero).
- The variance of the i-th extracted component (factor) is equal to the i-th Eigenvalue (diagonal elements of (4)).

PCA and the Linear Model of EEG

In the linear model, $\mathbf{V}$ is the forward and $\mathbf{V}^\top$ is the backward model, i.e., spatial patterns and spatial filters coincide:

Using the EVD of $\mathbf{\Sigma_x} = \mathbf{V}\mathbf{D}\mathbf{V}^\top$, we obtain for the covariance matrix of the components $\mathbf{\Sigma_s} = \mathbf{D}$ and therefore the given filter matrix $\mathbf{W}^\top = \mathbf{V}^\top$:

$$
\begin{aligned}
\hat{\mathbf{A}} &= \mathbf{\Sigma_x}\mathbf{V}\mathbf{\Sigma_s}^{-1} \\
&= \mathbf{V}\mathbf{D}\mathbf{V}^\top\mathbf{V}\mathbf{D}^{-1} \\
&= \mathbf{V}\mathbf{D}\mathbf{D}^{-1} \\
&= \mathbf{V}
\end{aligned}
$$

PCA and the Linear Model of EEG

Often a submatrix $\mathbf{V}_0 = [\mathbf{v}_{p(1)}, \ldots, \mathbf{v}_{p(k)}]^\top$ with $p(i) \in \{1, \ldots, P\}$ is used in a backward–forward projection:

$$\mathbf{x}(t) \mapsto \mathbf{V}_0 \mathbf{V}_0^\top \mathbf{x}(t)$$

Some examples of possible usage:

- $\mathbf{V}_0$ may consist of the Eigenvectors corresponding to the largest Eigenvalues (compression): low variance directions (assumed to be irrelevant) are discarded.
- In $\mathbf{V}_0$ the Eigenvectors corresponding to some very high Eigenvalues (say $\sqrt{d_i} > 100$; assumed to be non-EEG artifacts) are discarded.
- $\mathbf{V}_0$ consists of few Eigenvectors that are assumed to correspond to eye movements (according to the scalp maps of the spatial patterns). In this case one obtains an estimation of the influence of the eye movements in the EEG.

PCA Matlab Demo

Lessons Learnt

After this lecture you should

- be capable of describing the linear model
  - including the role of spatial patterns and spatial filters,
- be able to go back and forth between sensor and EEG space with the forward and the backward model,
  - in particular with PCA.

References I

- Haufe, S., Meinecke, F., Görgen, K., Dähne, S., Haynes, J.-D., Blankertz, B., and Bießmann, F. (2014).
  On the interpretation of weight vectors of linear models in multivariate neuroimaging.
  NeuroImage, 87:96–110. Neuroimage single best paper of 2014 Award.
- Ledoit, O. and Wolf, M. (2004).
  A well-conditioned estimator for large-dimensional covariance matrices.
  J Multivar Anal, 88:365–411.
- Parra, L. C., Spence, C. D., Gerson, A. D., and Sajda, P. (2005).
  Recipes for the linear analysis of EEG.
  NeuroImage, 28(2):326–341.
- Schäfer, J. and Strimmer, K. (2005).
  A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics.
  Stat Appl Genet Mol Biol, 4:Article32.