Separability of Covariance and 2-step LDA of Matrix Speller Data

Stefan Frenzel

April 12, 2012
1. Separability of Covariance Matrix

2. Learning Curves of 2-step LDA
Definitions

Separability

\( \Sigma \in \mathbb{R}^{d \times d} \) is separable if it can be written as

\[
\Sigma = V \otimes U
\]

with \( V \in \mathbb{R}^{p \times p} \), \( U \in \mathbb{R}^{q \times q} \), and \( d = pq \).
Definitions

Separability

\[ \Sigma \in \mathbb{R}^{d \times d} \text{ is separable if it can be written as} \]

\[ \Sigma = V \otimes U \]

with \( V \in \mathbb{R}^{p \times p} \), \( U \in \mathbb{R}^{q \times q} \), and \( d = pq \).

Remarks:

- \( U \) and \( V \) are not uniquely defined since \( \Sigma = \left( \frac{1}{c} U \right) \left( c V \right) \) for every nonzero \( c \)

- there are no unique ML estimators of \( U \) and \( V \) for the normal model
1. Separability of Covariance Matrix

Definitions

Matrix Normal Distribution

Let \( X \) be a \( q \times p \) random matrix. The matrix normal distribution with parameters \( \mu \), \( V \) and \( U \) (both positive definite) is defined by

\[
X \sim \mathcal{N}_{p,q} (\mu, U, V) \quad \text{if} \quad \text{vec}(X) \sim \mathcal{N}_{pq} (\text{vec}(\mu), V \otimes U).
\]

The pdf of \( X \) is given by

\[
f(x) = c^{-1} \exp \left[ -\frac{1}{2} \text{tr} \left( U^{-1}(X - \mu)V^{-1}(X - \mu)^T \right) \right]
\]

with \( c = (2\pi)^{(1/2)qp}|U|^{(1/2)p}|V|^{(1/2)q} \).
Estimation of $\mathbf{U}$ and $\mathbf{V}$

- let there be $N$ i.i.d. random samples $X_1, \ldots, X_n \in \mathbb{R}^{q \times p}$
- based on the pdf (1) the ML estimators $\hat{\mu}$, $\hat{\mathbf{V}}$ and $\hat{\mathbf{U}}$ fulfill the following equations

$$
\hat{\mu} = \frac{1}{N} \sum_{i=1}^{N} X_i \equiv \bar{X}
$$

$$
\hat{\mathbf{U}} = \frac{1}{pN} \sum_{i=1}^{N} (X_i - \bar{X}) \hat{\mathbf{V}}^{-1} (X_i - \bar{X})^T
$$

$$
\hat{\mathbf{V}} = \frac{1}{qN} \sum_{i=1}^{N} (X_i - \bar{X})^T \hat{\mathbf{U}}^{-1} (X_i - \bar{X})
$$
Estimation of $U$ and $V$

- the last two equations can be solved iteratively using a flip-flop algorithm [1]
Estimation of $U$ and $V$

- the last two equations can be solved iteratively using a flip-flop algorithm [1]
- this approach fails for our data because $U$ is singular due to the average reference
Separability of Covariance and 2-step LDA of Matrix Speller Data

1. Separability of Covariance Matrix

Estimation of $U$ and $V$

- the last two equations can be solved iteratively using a flip-flop algorithm [1]
- this approach fails for our data because $U$ is singular due to the average reference
- for estimating $U$ and $V$ based on samples from two classes (target and non-target trials) we instead use

$$\hat{U} = \frac{1}{p(N-2)} \sum_{i=1}^{N} (X_i - \bar{X}(y_i)) (X_i - \bar{X}(y_i))^T$$

$$\hat{V} = \frac{1}{q(N-2)} \sum_{i=1}^{N} (X_i - \bar{X}(y_i))^T (X_i - \bar{X}(y_i))$$

with class labels $y_i \in \{1, 2\}$ and class mean values $\bar{X}^{(1)}$ and $\bar{X}^{(2)}$
Estimation of $U$ and $V$

- data from nine subjects
- sampling frequency $f_S = 32 \text{Hz}$
- time window $0 \ldots 0.6 \text{s} \rightarrow p = 20$
- $32$ channels $\rightarrow q = 32$
- features scaled to unit variance
- mean squared error of approximation

$$
\epsilon = \frac{2}{qp(qp + 1)} \sum_{i \leq j} ((S - V \otimes U)_{ij})^2
$$

with $S$ being the pooled sample covariance matrix
Subject 1

Figure: Temporal ($V$, left) and spatial covariance matrix ($U$, right) for subject 1
Subject 1

Figure: $S$ (left) and $V \otimes U$ (right) for subject 1

$\epsilon \approx 1 \cdot 10^{-3}$
Separability of Covariance and 2-step LDA of Matrix Speller Data

1. Separability of Covariance Matrix

Subject 2

Figure: \( S \) (left) and \( V \otimes U \) (right) for subject 2

\[
\epsilon \approx 4 \cdot 10^{-4}
\]
Subject 3

Figure: \( S \) (left) and \( V \otimes U \) (right) for subject 3

\[ \epsilon \approx 4 \cdot 10^{-4} \]
Subject 4

Figure: $S$ (left) and $V \otimes U$ (right) for subject 4

\[ \epsilon \approx 5 \cdot 10^{-4} \]
1. Separability of Covariance Matrix

Subject 5

Figure: $S$ (left) and $V \otimes U$ (right) for subject 5

\[ \epsilon \approx 8 \cdot 10^{-4} \]
Subject 6

Figure: $S$ (left) and $V \otimes U$ (right) for subject 6

$\epsilon \approx 4 \cdot 10^{-4}$
Subject 7

Figure: $S$ (left) and $V \otimes U$ (right) for subject 7

$\epsilon \approx 4 \cdot 10^{-4}$
Subject 8

Figure: $S$ (left) and $V \otimes U$ (right) for subject 8

$\epsilon \approx 2 \cdot 10^{-3}$
Subject 9

Figure: $S$ (left) and $V \otimes U$ (right) for subject 9

$\epsilon \approx 3 \cdot 10^{-4}$
2-step LDA of Matrix Speller Data

- matrix speller data [2] with $pq = 640$ features
- train classification algorithms using $n$ samples from the first session
- consider both classification of fixated and counted character
- estimate of AUC based on scores of data from second session
Learning Curves (fixated character)

Figure: (a) RLDA, (b) PLDA, (c) 2-step LDA, and (d) NMC
Learning Curves (counted character)

Figure: (a) RLDA, (b) PLDA, (c) 2-step LDA, and (d) NMC
Learning Curves

Figure: Difference of AUC values of RLDA and 2-step LDA for classification of fixated (left) and counted (right) character
Separability of Covariance and 2-step LDA of Matrix Speller Data

2. Learning Curves of 2-step LDA

- Pierre Dutilleul.

- S Frenzel, E Neubert, and C Bandt.