Review: Multivariate Temporal Dictionary Learning for EEG [1]

Omar Ahmad omar.ahmad@telecom-paris.fr.com
Matias Etcheverry matias.etcheverry9@gmail.com

April 14, 2023

1 Introduction and contributions

The paper [1] presents a novel and efficient method for representing scalp electroencephalography (EEG) signals. While there are alternative technologies for studying brain activity, EEG remains popular due to its high temporal resolution and minimal latency. EEG has various applications, including detecting epileptic activity and studying event-related potentials (ERP), steady-state evoked potentials, and event-related synchronization/synchronization. In signal analysis, capturing the temporal, spatial, and spectral contents of the signal is typically the goal. Therefore, a crucial question is which dictionary can optimally represent the information recorded in the EEG.

Choosing the appropriate dictionary for representing EEG signals depends on the specific information that is being extracted, such as spectral, temporal, or spatial features. Fourier and wavelets dictionaries are suitable for spectral analysis of EEG signals but are limited in representing the various shapes of EEG patterns. The Gabor dictionary has a temporal shift-invariance property but falls short in representing evoked potentials and EEG bursts. A more recent approach is dictionary learning algorithms, which focus on data-driven representations instead of pre-existing models of the data.

Our project employed a two-fold approach, combining theoretical and experimental methods to gain a comprehensive understanding of the article and extract the primary contributions made by the authors. Firstly, we individually read the paper to gain an understanding of its content. Additionally, one of us searched for an open-source code of the paper to help us replicate the experiments and analyze the results. Fortunately, we found a code that required some adjustments to be compatible with our experiments. The modifications we made were not deeply related to the algorithm but rather focused on the initialization of the dictionary and the way inputs were handled by the functions. At the same time, the other group member searched for datasets and applications that could be used to assess the effectiveness and resilience of the techniques proposed in [1]. In fact, our experiments were slightly different from those conducted in [1]. Initially, we visualized some of the learned kernels to ensure that the algorithm was functioning properly. Then, we employed a grid search technique to determine the best hyperparameters. Next, we investigated the influence of the initialisation dictionary on the reconstructed signal. Additionally, we assessed the effectiveness of the proposed multivariate methods as a denoising approach. Finally, we utilized this method as a feature extractor for a classification task.

2 Method

In [1], an overview was given of the different time-frequency analysis tools applied for EEG. The monochannel sparse approximation is an EEG analysis where each channel is studied independently. The one channel signal is then modeled as $y = \Phi X + \epsilon$, with $y \in R^N$ of N temporal samples and $\Phi \in R^{N\times M}$ is a normalized dictionary composed of M time-frequency atoms $\{\phi_m\}_{m=1}^{M}$. In addition, another approach is the multichannel sparse approximation, where the EEG signal $y \in R^{N\times C}$ composed of several channels $c = 1..C$ are considered. The following reviewed methods link these channels spatially with a multichannel model. Another form is the shift-invariant model, where the signal is expressed as a sum of a few distinct structures called kernels. These kernels are characterized independently of their positions and are replicated at all positions throughout the signal using L shiftable kernels from a compact Ψ dictionary.

In recent times, the development of dictionary learning algorithms (DLAs) has made it possible to learn dictionary atoms in an unsupervised and data-driven manner. These algorithms use a set of iterations that alternate between sparse approximation and dictionary update to generate learned atoms that are specifically adapted to the analyzed data, rather than being generic. As a result, learned dictionaries are able to outperform generic ones in terms of processing quality.

The subsequent explanation describes a general multivariate approach for EEG analysis that incorporates both temporal modeling based on shift-invariance and a spatial model known as multivariate. This approach uses an efficient dictionary learning technique, making it more competitive than previous models. The approach comprises two primary steps: multivariate orthogonal Matching Pursuit (M-OMP) for the sparse approximation and a multivariate dictionary learning algorithm (M-DLA) for dictionary update. The channel signals are modeled as

$$y = \Phi X + \epsilon$$

where y here is an element of $R^{N\times C}$, Φ is an element of $R^{N\times M\times C}$, and x is an element of R^M. The multiplication ΦX is now considered as an element-wise product along the M dimension. The sparse approximation is written as:

$$\min_x \|y - \Phi X\|^2$$

$$\text{s.t. } \|x\|_0 \leq K$$

There are two unknown variable Φ and X, and to solve this challenging problem, the Matching Pursuit algorithm is applied iteratively after initializing Φ. To tackle the multivariate model described earlier, we use the multivariate Orthogonal-MP extension. At each iteration the algorithm selects the atom that produces the strongest decrease (in absolute value) in the mean square error (MSE), where the error at the iteration k is $\epsilon_{k-1} = x_m\phi_m + \epsilon_k$, with $\phi_m \in R^{N\times C}$. Therefore, one can proof that :

$$m_k = \arg\max_m | < \epsilon_{k-1}, \phi_m > |$$

Now as the sparse representation is found, the multivariate dictionary is updated in the second step based on maximum likelihood criterion, on the assumption of Gaussian noise:

$$\Phi = \arg\min_\Phi \|y - \Phi X\|^2$$

$$\text{s.t.} \forall m \in \mathbb{N}_M, \|\phi_m\| = 1$$

$$\epsilon_{k-1} = \epsilon_k - \phi_{m_k} x$$

This optimization problem can be solved using a stochastic gradient descent. These 2 steps are repeated until convergence, and at each iteration the dictionary is updated before going into the sparse approximation step. In the our experiments the presented multivariate methods and the source code are used in a shift-invariant way.

3 Data

We used the dataset 2A from the Brain Computer Interface (BCI) challenge IV. This dataset was commonly used for research in brain-computer interfaces and includes electroencephalogram (EEG) recordings from 9 healthy participants who were instructed to imagine performing four motor imagery tasks involving movements of their hands and feet in two sessions of acquisition, so that a total of 18 sessions. It is worth noting that during the competition, only 9 sessions were provided to participants for evaluation, with the remaining nine being reserved for testing. Our study specifically focused on evaluating a proposed multivariate dictionary learning algorithm using this dataset. The experimental sessions were divided into six runs with brief intervals in between. Each run involved 48 trials, with 12 trials corresponding to each of the four classes under investigation. Consequently, a complete session encompassed a total of 288 trials. EEG signals are sampled at 250 Hz using C = 22 channels. At the beginning of each session, some recording was performed to estimate the EOG influence before starting the Motor imagery tasks. As illustrated in Figure 1, there is an initial period of activity preceding the motor imagery task at the start of each trial. It is crucial to consider this prior activity when examining the EEG signals in relation to the task.

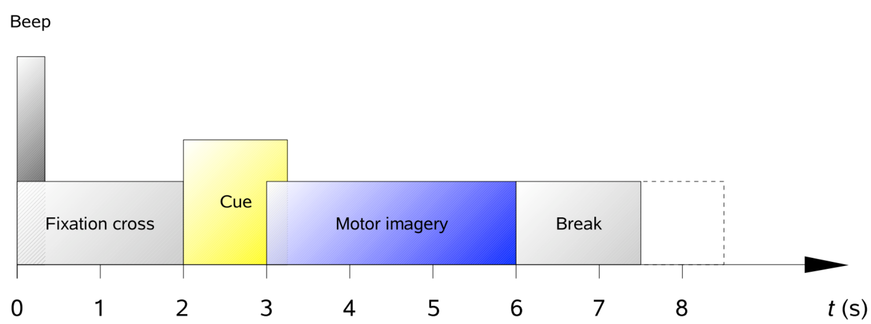

Figure 1: The trial paradigm

The limited number of healthy subjects limit the generalizability of it, and makes it misleading. Additionally, the dataset lacks diversity in terms of neurological disorders or disabilities that may benefit from BCI technology. The limited number of motor imagery tasks included in the dataset (only four) may not be enough to capture the complexity of real-world scenarios fully. Furthermore, the absence of demographic information about the subjects, such as age and gender, could make it challenging to account for individual differences in data analysis, since the brain response location could differ based on the age or the gender. Despite its usefulness in BCI research, it is essential to consider these potential limitations and problems when interpreting results obtained from this dataset.

4 Results

4.1 Learned kernels

Figure 2 illustrates four learned kernels, each of which is of size $\in R^{N\times C}$. These learned kernels exhibit smooth and continuous patterns and appear to capture relevant information from the original signal. Additionally, the efficiency of this method is demonstrated by the plots in Figure 2. Specifically, the second graph clearly highlights the spectral variability across channels, indicating that this method accounts for spatial relationships among different channels.

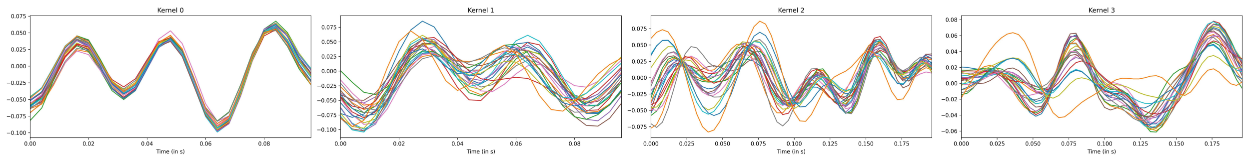

Figure 2: Learned kernels

4.2 Initialisation impact

During our in-depth investigation of this algorithm, a question arose regarding the need for an initial dictionary for multivariate dictionary learning. Therefore, we conducted an experiment to explore the impact of the initial dictionary on the final results. As shown in Figure 3, the reconstructed signals exhibit notable differences depending on the choice of the initial dictionary. Notably, starting with a dictionary extracted from the training signals appears to yield the best performance, significantly outperforming Gabor and random initializations.

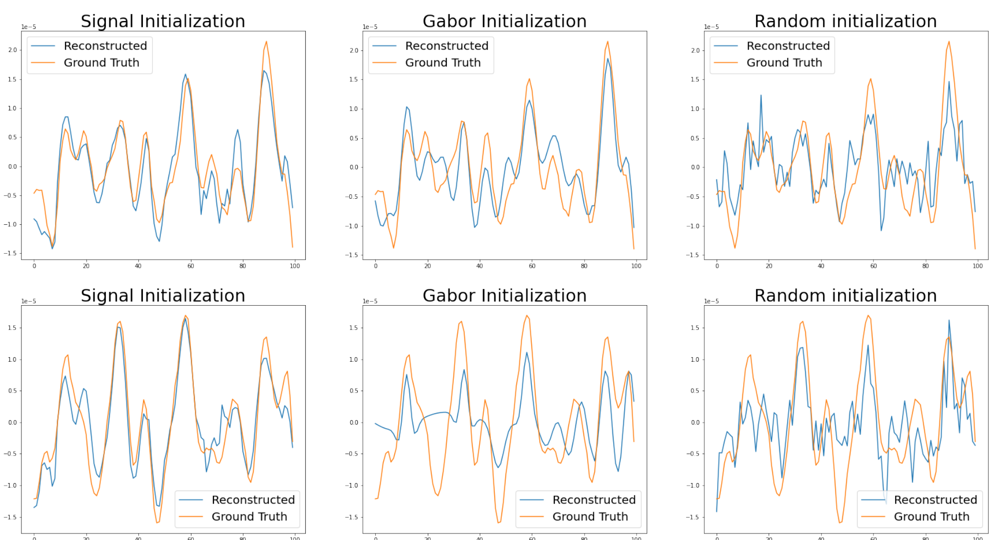

Figure 3: Learned kernels

4.3 Classification of the motory imagery

Exposing a stimulus to an individual elicits a neurological response that triggers a brief mental representation of a specific body part. The objective of the dataset analysis is to classify these signals exclusively from EEG recordings. The dictionary learning approach was employed to achieve this goal. However, due to computational limitations, we restricted our investigation to the EEG data of two participants and employed a simplified framework that assumes a priori knowledge of the exact time t of the stimulus presentation.

Initially, EEG signals associated with motor imagery are extracted by selecting a time interval of 2 seconds from 0.6 seconds after the presentation of a cue until 2.6 seconds post-stimulus.The extracted EEG records consist of 574 trials having 4 different classes, which are partitioned into training and testing sets. A multivariate temporal dictionary is then fitted on the training data, and each record is represented by a code that indicates the activation, index, and offset of the kernel. It's worth noting that a kernel may be activated multiple times in the same trial. As feature extraction, we compute the minimum, maximum, and total activation of each kernel.

A Support Vector Classifier process yields an accuracy of 46.09% on the test set, where the chance level is 25%. The confusion matrix in Figure 4 reveals that a classification confusion occurs between the right and left hand, potentially stemming from participant confusion or underlying neural mechanisms. To investigate the relationship between reconstruction quality and classification accuracy, we plot the reconstruction rate against the classification test results in Figure 5. A statistical analysis suggests that higher reconstruction quality is associated with lower classification accuracy, likely indicating that the algorithm struggles to handle excessive noise in the data. To improve the classification performance, it may be necessary to reduce the number of used kernels.

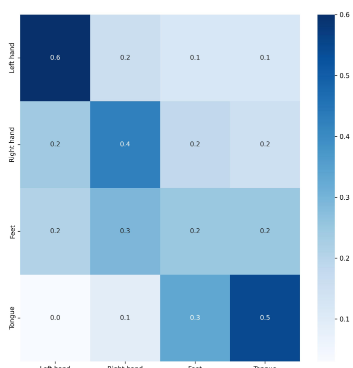

Figure 4: Confusion matrix

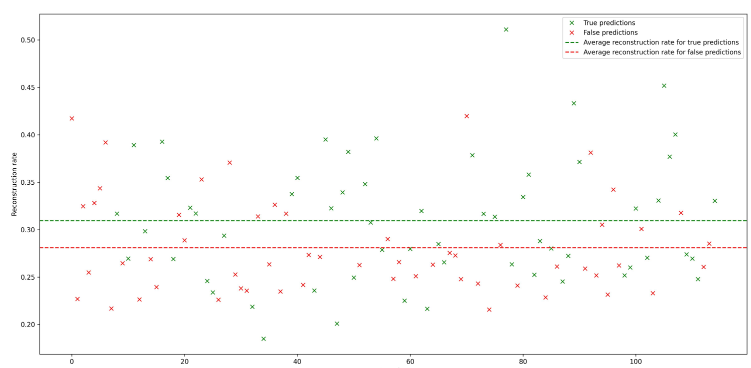

Figure 5: Reconstruction rate with classification errors

5 Conclusion

In conclusion, several key findings emerged from our analysis. First, we observed that the initialization has a significant effect on the reconstruction performance. Second, computational time proved to be a limiting factor in our investigation, with longer computational time required to obtain higher quality results. Third, while the reconstruction rate alone may not be sufficient to evaluate the quality of the denoiser, our results indicate that our approach is not a very effective denoiser. Fourth, finding the optimal hyperparameters was a challenging task. Fifth, we found that the kernels employed in our model do not necessarily have a physical interpretation, which may complicate the interpretation of our results.

In recent years, there has been growing interest in applying deep learning techniques to extract features from time-series EEG data. Convolutional neural networks (CNNs), recurrent neural networks (RNNs), and their variations, such as long short-term memory (LSTM) networks, have been used for feature extraction and classification of EEG signals. These methods have shown promising results in a variety of EEG-based applications, including motor imagery.

References

[1] Quentin Barthélemy et al. "Multivariate Temporal Dictionary Learning for EEG". In: (2013). DOI: 10.48550/ARXIV.1303.0742. URL: https://arxiv.org/abs/1303.0742.