EEG Classification: Robust vs Conventional Generalized Gaussian

Sam Nazari

May 2014

1 Abstract

In this project we apply a Robust Generalized Gaussian Detector and compare the results to a conventional Generalized Gaussian Detector. These subspace detectors are applied to both synthetic as well as data from an EEG typing system. The objective is to build a detector that is robust to artifacts without explicitly modeling artifacts.

2 Introduction

We aim to classify EEG recordings of brain activity that are subject to certain artifacts. Two common artifacts are: eye blink and jaw movement A conventional approach to artifact removal is to detect the artifact and ignore the associated section of data. This is not possible in an on-line setting such as a Brain Computer Interface. Instead, an on-line, robust detector is preferable.
In this project the artifacts are posed as interference in a learned subspace and robust subspace detection method[3] is applied to detect the target signal in the presence of artifacts.

First, the robust detection algorithm is applied in a noise-only setting to test the model fit. Then the RSVPkeyboard system data is used. RSVPkeyboard is an electroencephalography (EEG)-based brain interface typing system that uses evoked response potential (ERP) classification with the help of language models. In this system EEG data from 16 channels are recorded while subject is trying to detect target letter in a rapid serial visual presentation of symbols from alphabet. We use the introduced robust detection algorithm to detect target letters and also compare the effect of noise on this target detection algorithm performance[4].

S	known (learned)	$K \times M$
A	unknown	$K\times$ up to $K - M$
B	known (learned)	$K\times$ up to $K - M$
θ	unknown	M vector
ψ	unknown	vector
ϕ	unknown	vector
ω	unknown	scalar
v	$\mathcal{N}(0, I)$	$K \times 1$

Table 1: Variables as considered at time of detections

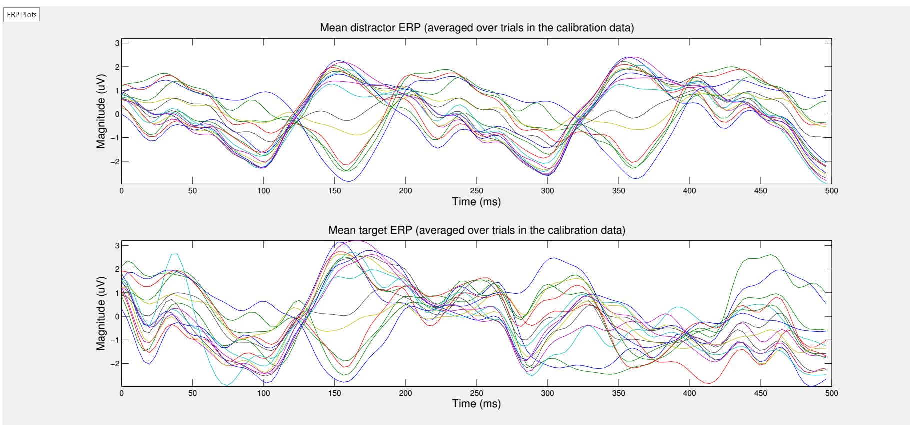

Figure 1: Averaged over trials for each channel

3 Problem formulation

A robust subspace detection method is applied to detect the target signal in noise with possible interference- artifacts. The current model assumes a Gaussian distribution for each class, target and non target in the feature space. We will consider three cases, target with unlearned interference and noise, target with learned interference and learned and target with learned and unlearned interference plus noise. The variables used throughout are defined in Table 1

4 Unknown Interference

We begin by assuming there is only the target signal and unknown interfernce

$$x = r_s + r_u + \eta \tag{1a}$$

$$= S\theta + U\psi + \eta \tag{1b}$$

we assume for each feature vector a windowed in time segment can be expressed as a sum of the target signal, r_s, an unknown interference, r_u and noise η. These

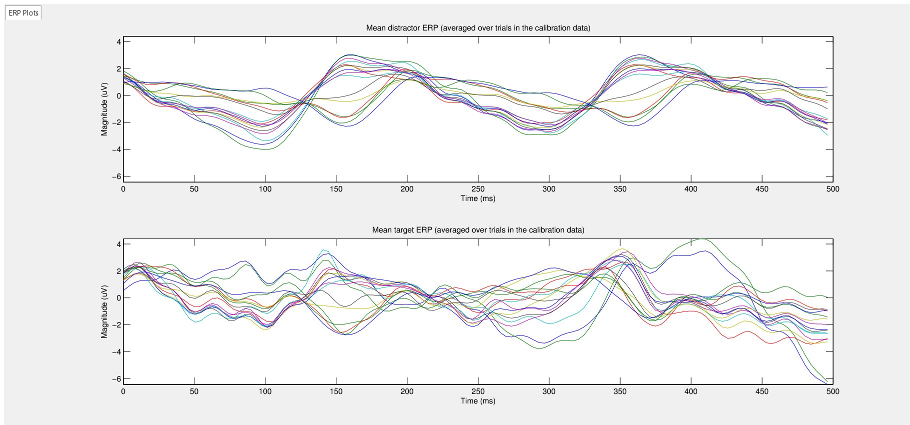

Figure 2: Averaged over trials for each channel, for eyeblink contaminated data

can each be described by the subspaces in which they exist and a magnitude. We assume that the subspace defined by the columns of S is known in detection, in practice this is estimated from a training set of data. The measurement is a $K \times 1$ vector, and r_s lies in a Mdimensional subspace defined by S.

Given the model above we desire a hypothesis test for the presence of the signal, that should work, even if ψ is nonzero and U is unknown.

$$\mathcal{H}_0 : U\psi + \eta \tag{2a}$$

$$\mathcal{H}_1 : S\theta + U\psi + \eta \tag{2b}$$

We apply a likelihood ratio test, which after taking the log yields the test statistics in Equation 3 for the no interference (conventional, Equation 3a) and interference (Equation 3b) cases.

$$\lambda(x) = (\frac{1}{w_1}\|x - s\hat{\theta}\|_2)^2 + (\frac{1}{w_0}\|x\|_2)^2 \tag{3a}$$

$$\lambda(x) = (\frac{1}{w_1}\|x - S\hat{\theta}\|_2)^2 + (\frac{1}{w_0}\|x - N\hat{\phi}\|_2)^2 \tag{3b}$$

A geometric interpretation of the problem is useful for the next step in the derivation. Consider the feature vector x to be in a space with components in the direction of and orthogonal to the signal subspace, S.
We assume use of a maximum likelihood estimate $\hat{\theta}$ for the unknown gain; this distance $\|x - s\hat{\theta}\|_2$ is the projection of x onto the subspace S.

$$P_S = S^T(S^TS)^{-1}S^T$$

With no interference, in the case where the noise variance is known, we can set a threshold for the magnitude of the projection of x into S for detection. This

Variance	Conventional Detector	Robust Detector
$\omega_0 = \omega_1 = \omega$, known	$\lambda(x) = \frac{x^T P_S x}{2\omega^2}$	$\lambda(x) = \frac{x^T (P_S - P_N) x}{2\omega^2}$
ω_0, ω_1 unknown	$\lambda(x) = \frac{x^T x}{x^T P_N x}$	$\lambda(x) = \frac{x^T P_S x}{x^T P_N x}$

Table 2: Various Detector statistics from [3]

results in the χ^2 test. In the case where the noise variance is unknown, we can consider the ratio of the portion of x in the subspace of S and in the orthogonal space $N = S_\perp$. This is the F statistic[3].

Given this intuition, we present four detector statistics in Table 2 where P_N is the projection matrix onto the part of the measurement space orthogonal to the signal.

4.1 Known Interference

To model the non-target, ongoing, or background EEG activity, we add a known (or learned) interference term, r_b. This lies in a subspace B which is not necessarily orthogonal to the signal space.

$$x = r_s + r_b + r_a + \eta \tag{4}$$

$$= S\theta + B\phi + A\psi + \omega v \tag{5}$$

One test for this model is given by the Adaptive Matched Subspace detector[1]. In this method, we define Z as the space spanned by the full measurement, $Z = [SB]$. To denote an orthogonal subspace, we use a superscript, $\perp$, i.e. $P_B^\perp = I - P_B$.
This method assumes no unknown interference, only the learned interference. Here, λ is derived via likelihood ratio test method, but $\lambda - 1$ is used as the test statistic to make numerator and denominator independent (with 1 in the form of $\frac{P_Z^\perp}{P_Z^\perp}$). The test statistic is given in Equation 6

$$\lambda(x) = \frac{x^T (P_B^\perp) x}{x^T P_Z^\perp x} \tag{6a}$$

$$D_{AMSD} = \frac{x^T (P_B^\perp - P_Z^\perp) x}{x^T P_Z^\perp x} \tag{6b}$$

A related, but slightly different method, Orthogonal subspace projection method focuses on a signal-to-noise (SNR) related statistic[2]. This method aims to maximize the SNR using projections into orthogonal subspace and matched filer(template).

$$SNR = \lambda(x) = \frac{x^T (P_B^\perp s)^2}{x^T P_B^\perp x} \tag{7}$$

Finally, we consider the case from the extensions section of [3], where there is both learned and unlearned interference. Following the same methodology as in section 4, the final detectors for this most realistic case are in Equation 8.

$$\lambda(x) = \frac{x^T(P_S - P_B)x}{x^T(P_N - P_B)x} \tag{8}$$

5 MATLAB Scripts

In order to run the MATLAB scripts just run mainRbst.m.

References

[1] Joshua Broadwater. A hybrid algorithm for subpixel detection in hyperspectral imagery. *Geoscience and Remote ...*, 1(1):1601–1604, 2004.

[2] CI Chang. Orthogonal subspace projection (OSP) revisited: a comprehensive study and analysis. *IEEE Transactions on Geoscience and Remote Sensing,*, 43(3):502–518, 2005.

[3] M.N. Desai and R.S. Mangoubi. Robust gaussian and non-gaussian matched subspace detection. *IEEE Transactions on Signal Processing*, 51(12):3115–3127, December 2003.

[4] Umut Orhan, Deniz Erdogmus, Brian Roark, Barry Oken, and Melanie Fried-Oken. Offline analysis of context contribution to ERP-based typing BCI performance. *Journal of neural engineering*, 10(6):066003, December 2013.