IOC 5275: Brain Computer Interface **(Due: 04/13/2021)**

Lab2: EEG Preprocessing and Data Cleaning

Instructor: Chun-Shu Wei *Group Number:* 1
TA: Min-Jiun Tsai *309591027, 0713223, 0860838, 0716307, 0810834, 0617028*

Submission Policy

Read all the instructions below carefully before you start working on the assignment, and before you make a submission. For this assignment, please hand in the following your report (pdf) and code (.ipynb or .m file).

- **PLAGIARISM IS STRICTLY PROHIBITED. (0 point for Plagiarism)**
- For mathematical problem(s), please show your work step by step and clarify statement of theorem you use (if any). Answering without mathematical derivations will get 0 point.
- Submission deadline: **2021.04.13 09:00:00 AM**.
- **Late submission penalty formula:**

$$original\ score \times (0.7)^{\#(days\ late)}$$

File Format

- Each group submits 1 report (.pdf and .tex file) and 1 code (.ipynb or .m).
- **Report** must contains observations, results and explanations. Please name your .pdf and .tex file as **5275_Lab2_GroupNum.pdf** and **5275_Lab2_GroupNum.tex**,respectively.
- Paper submission is not allowed. **Please use our LaTeX template to complete your report**.
- **Code** file must contains comments to explain your code. Please name your code file as **5275_Lab2_GroupNum.ipynb/.m**
- Implementation will be graded by completeness, algorithm correctness, model description, and discussion.
- **Illegal format penalty:** -5 points for violating each rule of file format.

Prerequest

To finish programming problem, you could choose Matlab or Python base on your programming preference.

Matlab 2020a+

- NYCU installation page
- NCTU installation tutorial
- EEGLab official installation page (v2020.0+ is recommended)

Python 3.7+

- MNE official installation page (0.20.7+ is recommended)

1 Mathematical problem

1.1 Find the coefficients b_n in Fourier sine series

Let us begin with the Fourier sine series

$$S(t) = \sum_{n=1}^{\infty} b_n \sin\frac{n\pi t}{T},\ \textit{in the interval } (0,T). \tag{1.1}$$

To solve b_n while $S(t)$ is given, we can use the key observation that

$$\int_0^T \sin\frac{n\pi t}{T}\sin\frac{m\pi t}{T}dt = 0\ \forall m,n\in\mathbb{Z},\ m\neq n$$

with $\sin a \sin b = \frac{1}{2}[\cos(a-b) - \cos(a+b)]$.

$$\begin{aligned}\int_0^T \sin\frac{n\pi t}{T}\sin\frac{m\pi t}{T}dt &= \frac{1}{2}\int_0^T \cos\left((n-m)\frac{\pi t}{T}\right) - \cos\left((n+m)\frac{\pi t}{T}\right)dt\\ &= \left[\frac{T}{2\pi(n-m)}\sin\frac{(n-m)\pi t}{T} - \frac{T}{2\pi(n+m)}\sin\frac{(n+m)\pi t}{T}\Big|_{t=0}^{T}\right]\\ &= 0,\ \forall m,n\in\mathbb{N},\ m\neq n\end{aligned} \tag{1.2}$$

Let's fix m, multiply (1.1) by $\sin\frac{m\pi t}{T}$, and integrate the series (1.1) term by term to get

$$\int_0^T S(t)\sin\frac{m\pi t}{T}dt = \int_0^T\left(\sum_{n=1}^{\infty} b_n\sin\frac{n\pi t}{T}\right)\sin\frac{m\pi t}{T}dt = \sum_{n=1}^{\infty} b_n\int_0^T\sin\frac{n\pi t}{T}\sin\frac{m\pi t}{T}dt = b_m\int_0^T\sin^2\frac{m\pi t}{T}dt \tag{1.3}$$

First, We can compute $\int_0^T \sin^2\frac{m\pi t}{T}dt$:

$$\begin{aligned}\int_0^T \sin^2\frac{m\pi t}{T}dt = -\frac{T}{m\pi}\int_0^T\sin\frac{m\pi t}{T}d\cos\frac{m\pi t}{T} &= -\frac{T}{m\pi}\left[\sin\frac{m\pi t}{T}\cos\frac{m\pi t}{T}\Big|_{t=0}^{T} - \frac{m\pi}{T}\int_0^T\cos^2\frac{m\pi t}{T}dt\right.\\ &= -\frac{T}{m\pi}\left[\sin\frac{m\pi t}{T}\cos\frac{m\pi t}{T}\Big|_{t=0}^{T} - \frac{m\pi}{T}\int_0^T 1-\sin^2\frac{m\pi t}{T}dt\right]\\ &= -\frac{T}{m\pi}\left[\sin\frac{m\pi t}{T}\cos\frac{m\pi t}{T}\Big|_0^T\right] + T - \int_0^T\sin^2\frac{m\pi t}{T}dt\end{aligned} \tag{1.4}$$

$$\Rightarrow \int_0^T\sin^2\frac{m\pi t}{T}dt = \frac{1}{2}\left\{-\frac{T}{m\pi}\left[\sin\frac{m\pi t}{T}\cos\frac{m\pi t}{T}\Big|_0^T\right]+T\right\} = \frac{T}{2} \Rightarrow \int_0^T S(t)\sin\frac{m\pi t}{T}dt = b_m\int_0^T\sin^2\frac{m\pi t}{T}dt = b_m\frac{T}{2}$$

$$\Rightarrow b_m = \frac{2}{T}\int_0^T S(t)\sin\frac{m\pi t}{T}dt$$

This is the famous formula for the Fourier coefficients in the series (1.1).

Problem 1. Fourier Series (10+5=15 points)

Suppose that the Fourier sine series of $g(x) = x$ *on* $(0,l)$ is given by

$$g(x) = \sum_{k=1}^{\infty} f_k(x) \tag{1.5}$$

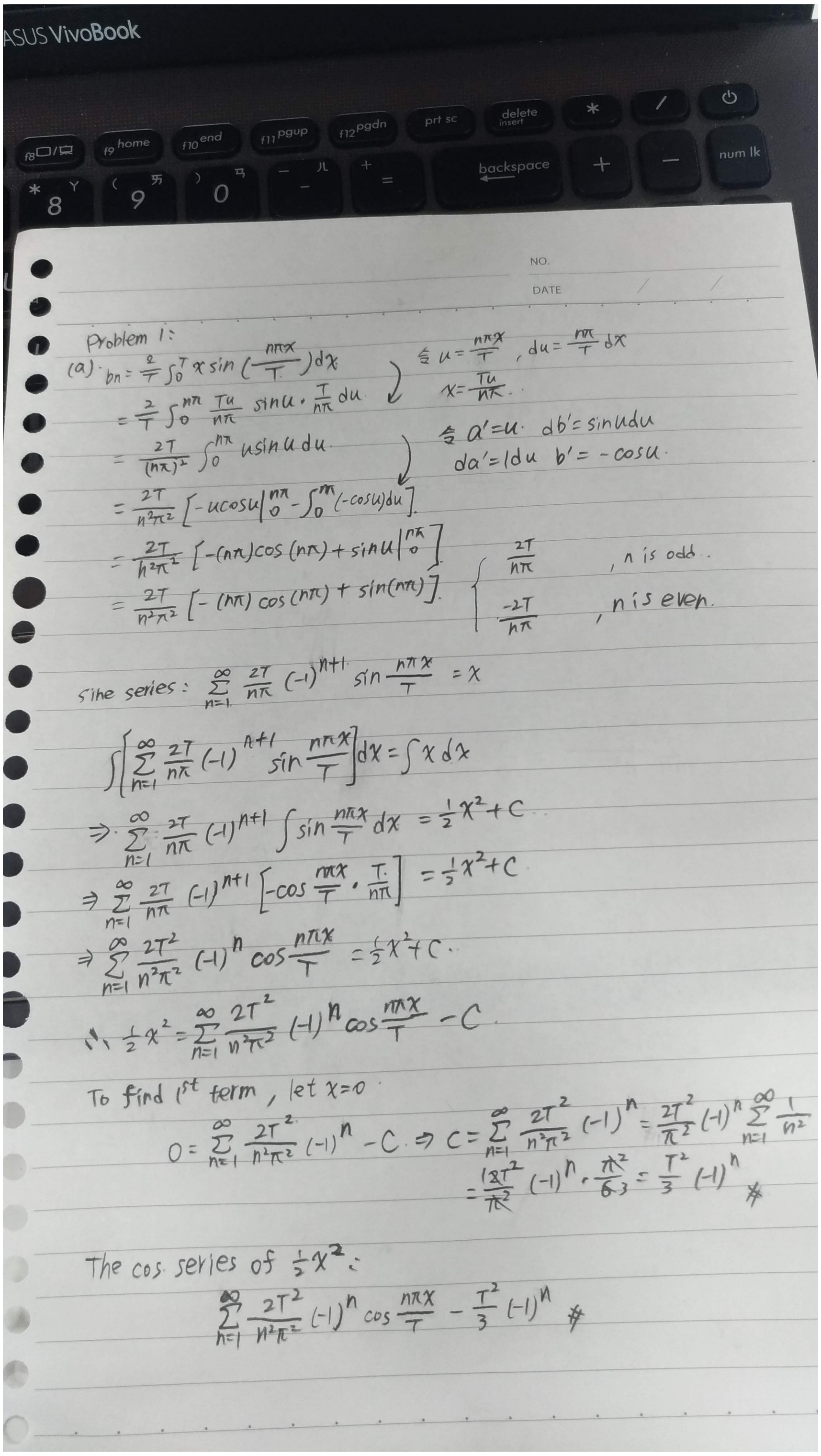
with conditions : $f_k : (0,l) \to \mathbb{R}$ is integrable and $\sum_{k=1}^{\infty} f_k(x)$ is uniformly convergent.
(a) Find the **Fourier cosine series** of the function $\frac{x^2}{2}$ and the constant of integration (the 1^{st} term of cosine series).

(b) Please exam whether the following series converges or not. Furthermore, find the sum of the series if it exists.

$$\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n^2} = ? \tag{1.6}$$

Definition 1.1 *Uniform Convergence*
Let I be an interval on $\mathbb{R}$ and $f_k : I \to \mathbb{R}$ be a real-valued function on I. The sequence of functions $\{f_k\}_{k\in\mathbb{N}}$ is said to converge uniformly on I to the function $f : I \to \mathbb{R}$ if $\forall \epsilon > 0, \exists N = N(\epsilon) \in \mathbb{N}$ such that $\forall x \in I$ and $k > N, |f_k(x) - f(x)| < \epsilon$.

a)



NO.

DATE / /

Problem 1:

(a). $b_n = \frac{2}{T}\int_0^T x \sin(\frac{n\pi x}{T})dx$

令 $u = \frac{n\pi x}{T}$, $du = \frac{n\pi}{T}dx$

$x = \frac{Tu}{n\pi}$.

$= \frac{2}{T}\int_0^{n\pi} \frac{Tu}{n\pi} \sin u \cdot \frac{T}{n\pi} du$

$= \frac{2T}{(n\pi)^2}\int_0^{n\pi} u \sin u\, du$

令 $a' = u$. $db' = \sin u\, du$

$da' = 1du$ $b' = -\cos u$.

$= \frac{2T}{n^2\pi^2}\left[-u\cos u\Big|_0^{n\pi} - \int_0^{n\pi}(-\cos u)du\right]$.

$= \frac{2T}{n^2\pi^2}\left[-(n\pi)\cos(n\pi) + \sin u\Big|_0^{n\pi}\right]$.

$= \frac{2T}{n^2\pi^2}\left[-(n\pi)\cos(n\pi) + \sin(n\pi)\right]$.

$\begin{cases} \frac{2T}{n\pi}, & n \text{ is odd.} \\ \frac{-2T}{n\pi}, & n \text{ is even.} \end{cases}$

Sine series: $\sum_{n=1}^{\infty}\frac{2T}{n\pi}(-1)^{n+1}\sin\frac{n\pi x}{T} = x$

$\int\left[\sum_{n=1}^{\infty}\frac{2T}{n\pi}(-1)^{n+1}\sin\frac{n\pi x}{T}\right]dx = \int x\,dx$

$\Rightarrow \sum_{n=1}^{\infty}\frac{2T}{n\pi}(-1)^{n+1}\int\sin\frac{n\pi x}{T}dx = \frac{1}{2}x^2 + C$

$\Rightarrow \sum_{n=1}^{\infty}\frac{2T}{n\pi}(-1)^{n+1}\left[-\cos\frac{n\pi x}{T}\cdot\frac{T}{n\pi}\right] = \frac{1}{2}x^2 + C$

$\Rightarrow \sum_{n=1}^{\infty}\frac{2T^2}{n^2\pi^2}(-1)^n\cos\frac{n\pi x}{T} = \frac{1}{2}x^2 + C$.

$\therefore \frac{1}{2}x^2 = \sum_{n=1}^{\infty}\frac{2T^2}{n^2\pi^2}(-1)^n\cos\frac{n\pi x}{T} - C$.

To find 1st term, let $x=0$.

$0 = \sum_{n=1}^{\infty}\frac{2T^2}{n^2\pi^2}(-1)^n - C \Rightarrow C = \sum_{n=1}^{\infty}\frac{2T^2}{n^2\pi^2}(-1)^n = \frac{2T^2}{\pi^2}(-1)^n\sum_{n=1}^{\infty}\frac{1}{n^2}$

$= \frac{\cancel{2}T^2}{\cancel{\pi^2}}(-1)^n\cdot\frac{\cancel{\pi^2}}{\cancel{6}3} = \frac{T^2}{3}(-1)^n$ #

The cos series of $\frac{1}{2}x^2$:

$\sum_{n=1}^{\infty}\frac{2T^2}{n^2\pi^2}(-1)^n\cos\frac{n\pi x}{T} - \frac{T^2}{3}(-1)^n$ #

NO.

DATE / /

(b) Use Alternating Series Test

① $\lim_{n\to\infty} \frac{1}{n^2} = 0$ ② $\frac{1}{(n+1)^2} \le \frac{1}{n^2}$ $\therefore \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^2}$ converges.

For the sum of $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^2}$; we use the result of (a):

$$\frac{1}{2}x^2 = \sum_{n=1}^{\infty} \frac{2T^2}{n^2\pi^2}(-1)^n \cos\frac{n\pi x}{T} - \frac{T^2}{3}(-1)^n$$

$$= \frac{2T^2}{\pi^2} \sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} \cos\frac{n\pi x}{T} - \frac{T^2}{3}(-1)^n.$$

$$\left[\frac{1}{2}x^2 + \frac{T^2}{3}(-1)^n\right] \cdot \frac{\pi^2}{2T^2} = \sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} \cos\frac{n\pi x}{T}$$

When $x=0$:

$$\frac{T^2}{3}(-1)^n \cdot \frac{\pi^2}{2T^2} = \sum_{n=1}^{\infty} \frac{(-1)^n}{n^2}$$

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^2} = \frac{\pi^2}{6}(-1)^{n+1} \quad \#$$

1.2 Independent Component Analysis

1.2.1 Motivation: Blind Source Separation

Model Formalization

Let n denotes number of source signal, ch denotes number of channel, and d denotes dimension of signal. The matrix $S \in \mathbb{R}^{n\times d}$ denotes source signals. We assume that recorded signals $X = AS + E \in \mathbb{R}^{ch\times d}$ are given by linear mixing system where $A \in \mathbb{R}^{ch\times n}$ is the unknown mixing matrix and $E \in \mathbb{R}^{ch\times d}$ denotes the noise. Basically, $ch \geq n$

The goal of BSS is to estimate $\hat{A}$ and $\hat{S}$ so that $\hat{S}$ provides unknown source signals as possible.

$$X = AS + E \leftarrow X = \hat{A}\hat{S}$$

Since $ch \geq n$, there are a lot of combinations (A, S) satisfy $X = AS + E$. We could apply different types of constraint to solve this system:

- PCA: Orthogonal constraint
- SCA: Sparsity constraint
- NMF: Non-negative constraint
- ICA: Statistically independent constraint

Therefore, there are many methods to solve the BSS problem depending on the constraints. What we used is depended on subject matter. In this lab, I only introduce **ICA**.

1.2.2 Model of ICA

The Cocktail Party Problem

Let X be recorded signal and S is a source signal according to above formalization. We assume that $\{s_i | i = 1, 2, ..., n\}$ is statistically independent.

$$x_i(t) = \sum_{j=1}^{n} a_{ij} s_i(t), \ \forall i \in \mathbb{Z}_n$$

Independent Component Analysis is to estimate the independent component $S(t)$ from $X(t)$.

$$X(t) = AS(t)$$

Hypothesis of ICA

- $\{s_i | i \in \mathbb{Z}_n\}$ statistically independent, that is, $P(s_1, ..., s_n) = \prod_{i=1}^{n} P(s_i)$
- $\{s_i | i \in \mathbb{Z}_n\}$ follows the Non-Gaussian distribution.
- A is regular

Therefore, we could rewrite the model as $S(t) = BX(t)$ where $B = A^{-1}$. It's only necessary to estimate B so that $\{s_i | i \in \mathbb{Z}_n\}$ is independent.

Definition 1.2 *White signal*
White signals are defined as any $z(t) \in \mathbb{R}^{d\times 1}$ which satisfying

- *Zero mean:* $E[z] = \boldsymbol{0} = m_z$
- *Unit covariance:* $C_z = E[(z - m_z)(z - m_z)^T] = E[zz^T] = I_d$

Note: If $m_z = \mathbf{0}$, then the correlation matrix $R_z = C_z + m_z m_z^T = C_z$. Recall that recorded signals are $X(t) = AS(t)$. ICA solve $S(t)$ by $S(t) = BS(t)$.

Problem 2: Whiteness property is preserved under orthogonal transformations (5 points)

Assume that an orthogonal transformation $U \in \mathbb{R}^{d\times d}$ and z is white, please prove that

$$m_{Uz} = m_z \;\&\; Cov_{Uz} = Cov_z \tag{1.7}$$

Page:

Date:

Problem2:

let $y = Uz$

y can be interpreted geomatrically as the rotation of the vector z about the origin.

$$\mu_{Uz} = \mu_y = E(y) = E(Uz) = U\,E(z) = U\cdot\mu_z = 0$$

$$\Rightarrow m_{Uz} = 0 = m_z \quad \#$$

$\because m_{Uz} = m_z = 0$

$$\therefore C_{Uz} = E[(Uz - m_{Uz})(Uz - m_{Uz})^T]$$

$$= E(Uz\cdot(Uz)^T)$$

$$= E(U\cdot z\cdot z^T\cdot U^T) \qquad z\cdot z^T = I_d$$

$$= E(U\cdot I\cdot U^T)$$

$$= E(U\cdot U^T)$$

$\because U$ is orthoganal matrix

$$\therefore E(U\cdot U^T) = I = C_{Uz} = C_z \quad \#$$

Whitening is useful for PCA and simplifies ICA problem. If we denote whitening signal as

$$z(t) = Vx(t)$$

where $V \in \mathbb{R}^{d \times d}$ is a whiten matrix of $x(t)$, then model becomes

$$s(t) = Uz(t) = UVx(t) = Bx(t)$$

where $U \in \mathbb{R}^{d \times d}$ is an orthogonal transformation matrix.

The Gaussianity of $X(t)$ (sums of non-gaussian random variables) must be larger than $S(t)$ (original) according to Central Limit Theorem. Let b_i be the row vector of B and the row vector of source signals $\hat{s}_i(t) = b_i^T x(t)$, we want to maximize the Non-Gaussianity of $b_i^T x(t)$. **Hence it's necessary to estimate U !**

Kurtosis is a measure of non-gaussianity

Definition 1.3 *Kurtosis*
for a random variable y,

$$kurt(y) = E[y^4] - 3(E[y^2])^2$$

That is, for white signal z,

$$kurt(z) = E[z^4] - 3(E[z^2])^2 = E[z^4] - 3$$

Which means we could solve ICA problem by

$$\hat{b} = \max_b \|kurt(b^T x(t))\| \tag{1.8}$$

We consider $z = Vx$ is a white signal given from source signal x, then we could rewrite (1.8) as:

Problem 3: Solving ICA problem by kurtosis (10 points)

$$\max \|kurt(w^T z)\| \ with \ w^T w = 1 \tag{1.9}$$

Ans:

To find the value of w that maximize the absolute value of $\|kurt(w^T z)\|$, We first calculate the partial derivation of this equation. After doing this, we can use the constraint $w^T w = 1$ to find w.

problem 3

$\max_{w} \| kurt(w^T z) \|$, $w^T w = 1$. $z = Vx$ is a white signal

$kurt(y) = E[y^4] - 3(E[y^2])^2$

$\Rightarrow \frac{\partial}{\partial w} \| kurt(w^T z) \| = sign(kurt(w^T z)) \left[\frac{\partial}{\partial w} kurt(w^T z) \right]$

$\Rightarrow \frac{\partial}{\partial w} kurt(w^T z) = \frac{\partial}{\partial w} \left(E[(w^T z)^4] - 3(E[(w^T z)^2])^2 \right)$

$= \frac{\partial}{\partial w} \left(E[(w^T z)^4] \right) - \frac{\partial}{\partial w} \left(3(E[(w^T z)^2])^2 \right)$

$\Rightarrow \frac{\partial}{\partial w} \left(E[(w^T z)^4] \right) = E\left[\left(\frac{\partial}{\partial w} w^T z \right) \cdot 4(w^T z)^3 \right] = E\left[z \cdot 4(w^T z)^3 \right]$

$\Rightarrow \frac{\partial}{\partial w} 3(E[(w^T z)^2])^2 = \frac{\partial}{\partial w} 3(E[(w^T z z^T w)])^2 = \frac{\partial}{\partial w} 3\left(E[w^T I_d w] \right)^2$

$= \frac{\partial}{\partial w} 3(E[w^T w])^2 = \frac{\partial}{\partial w} 3\left(\|w\|^2 \right)^2$

$= \left[\frac{\partial}{\partial w} \|w\|^2 \right] \cdot 6\left(\|w\|^2 \right) = 2w \cdot 6\|w\|^2 = 12 w \|w\|^2$

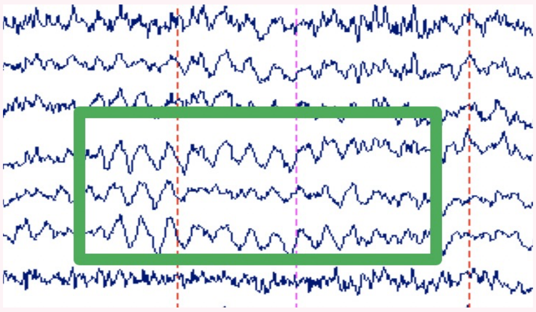
$\Rightarrow \frac{\partial}{\partial w} \left\| kurt(w^T z) \right\| = sign(kurt(w^T z)) \left(4E[z (w^T z)^3] - 12 w \|w\|^2 \right)$

2 Multiple choices

Please give a brief explanation for option(s) you choose. Answering without any description will get 0 point.

Problem 4 (5 points)

The image here shows a **2-second** period of EEG from several different electrodes. What is the nature of the oscillating signal shown in the green box? **Select all correct options.**



(A) It is a 1 Hz oscillation

(B) It is a 5 Hz oscillation

(C) It is a 10 Hz oscillation

(D) It is a delta-band oscillation

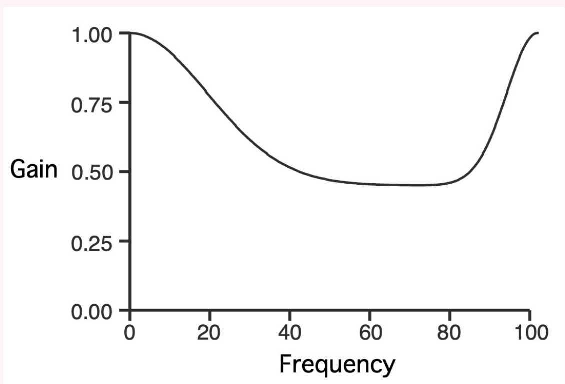
(E) It is a theta-band oscillation

(F) It is an alpha-band oscillation

Using a ruler we measured 2 second box and green box for different scales, then compute a duration inside green box we get [1.29, 1.28, 1.27] cm. After dividing number of peaks in green box by duration we get next values for oscillation frequency [10.077, 10.15, 9.9] in Hz. It gives us $\mu = 10.04$ and $\sigma = 0.1$. Also it can be *alpha-band* oscillation because its range lays from 8 to 12 Hz.

Problem 5 (5 points)

Which of the following statements are true of the frequency response function shown below?



(A) This filter would almost completely eliminate very low frequencies (because the gain is near 1.00 for low frequencies).

(B) This filter would have very little effect on very low frequencies (because the gain is near 1.00 for low frequencies).

(C) This filter would reduce frequencies of 40-80 Hz by 50% (because the gain is near 0.50 for these frequencies).

(D) This is a high-pass filter.

(E) This is a low-pass filter.

Based on the graph from the problem, the gain of the very low frequency is 1.00. This indicated that the very low frequency will have very little effect after passing through the filter, so B is correct. The gain of the frequencies of 40-80 Hz is 0.5, so the signal within these frequencies will be reduced 50 %, so C is correct.
From the graph, there is very little effect on very high frequency and very low frequency, so the filter is neither a low-pass filter, nor a high-pass filter. Therefore, D and E are not correct.

3 Programming problem

Please use the data: sXD_5678.set to answer the following problems.

3.0.1 Dataset Information

The solid black arrows represent driving trajectory. The empty circle represents deviation onset. The double circle represents response onset. The circle with a cross represents end of response. which was sufficient for subjects to experience fatigue.

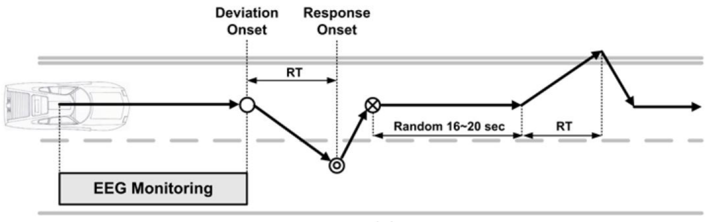


Figure: Event-related lane-departure task [Kuan-Chih Huang and Jung,2016],[Chin-Teng Lin and Jung, 2010]

Problem 6 (2+1+1+1+4+1=10 points)

Please following the following steps:

1. Plot 2D channel location map and re-reference data by $\frac{A1+A2}{2}$.
2. Down-sampling to $250Hz$.
3. Run ICA and record computational time of ICA by code.
4. Plot component map in 2D.
5. Indicate noise component(s) if it exist and explain reason why you identify this component as a noise or artifacts.
6. Plot first 10-second channel data before and after deleting noise/artifact component(s).

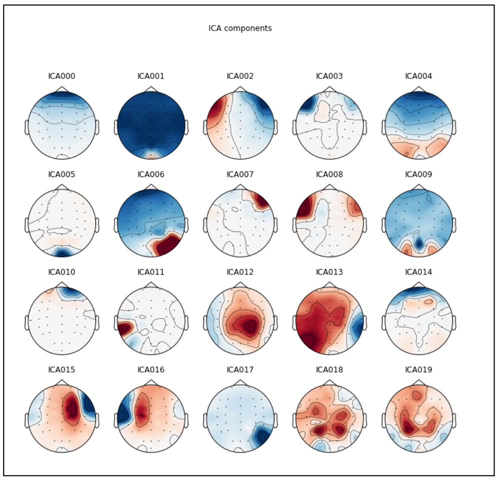
Problem 7 (0+0+0+1+1+1+4+1+8+5=20 points)

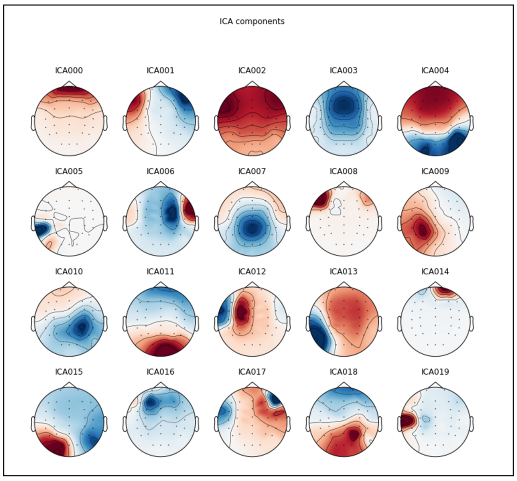
Delete vehicle position channel and then:

1. Plot 2D channel location map and re-reference data by $\frac{A1+A2}{2}$.
2. Down-sampling to $250Hz$.
3. Bandpass filtering $[1, 50]Hz$
4. Run ICA and record computational time of ICA by code.
5. Plot component map in 2D.
6. Indicate noise component(s) if it exist and explain reason why you identify this component as a noise or artifacts.
7. Plot first 10-second channel data before and after deleting noise/artifact component(s).

After above preprocessing steps...

(a) Compare results (e.g. component map) and try to explain your observations.

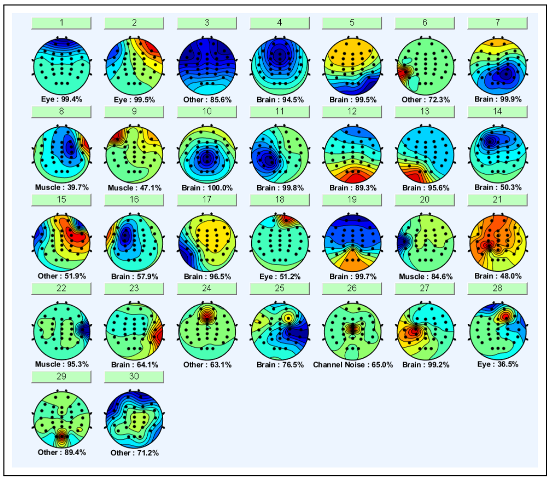


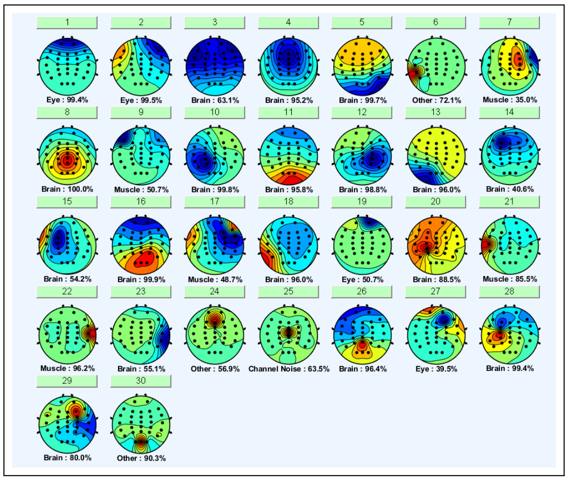


(1) criteria of selecting artifacts

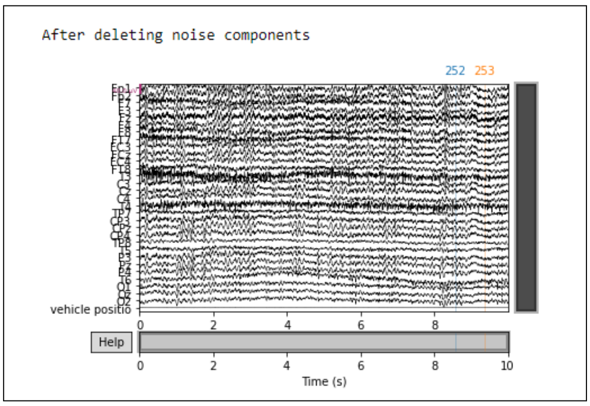
First set of components is from problem 6, where we didn't filter signals, and second set is from problem 7 with filtering. As you can see first component(ICA000) in both sets are similar and is highly possible to represent muscle activity from the face, judging from electrode locations and the fact that signal from muscles are stronger. For the second (ICA001) component in the first set, it is identified as 'vehicle positio'. For the third(ICA002),fourth (ICA003), ninth (ICA008) component in the first set we get uniform activation across all channels and we interpret that as uniform noise, when for the second set it's highly possible to represent eyes component. ICA000,ICA001, ICA002, ICA008 in the second sets are eyes for same reason as above.

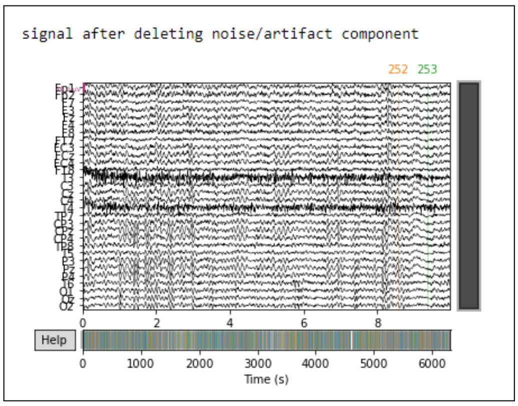
To justify our judgement upon components, we also use eeglab's iclabel feature as validation.





(2) scrolled EEG-data comparison





As we can see, problem 7(right figure) shows more cleaner signal than problem 6 (left figure) after removing artifact components.

(b) Explain why it takes less time this time?

There are two main possible factors. First, in problem 7, the non-EEG channel 'vehicle position' has been removed. In problem 6, vehicle position is taken into account as a independent input signal with other EEG input signals. This makes ICA algorithm harder to find independent components based on given input signals. Secondly, in problem 7, the low frequency drifts and high frequency noises would possibly be filtered out, and ICA algorithm could compute more efficiently.

Problem 8 (10+10+10=30 points)

Design your own EEG preprocessing strategy:
(a) Describe your design idea (e.g. Apply CleanLine function in EEGLab to eliminate environmental artifacts and apply lowpass filtering to remove drift... ...)
(b) Compare the performance (computational time) and results with proplem 8.
(c) Explain potential reason(s) why performance of your preprocessing strategy is superior/inferior to performance of problem 8?

a) Design idea

   1) Re-referencing using mean(A1, A2)
   2) Determine power line frequency using DFFT.
   3) Filter power line noise by using notch filter 60 Hz, 10 Hz notch width. (From the dataset source paper, we could also find that the data was collected in Taiwan, and the power frequency is indeed 60 Hz.)
   4) Apply custom Band-pass filter with parameters: with edge 1 - 30 Hz, with transition bandwidth 0.7 Hz at low frequency and 70 Hz at higher frequency.[10] The brain signal is usually lower than 30 Hz [7], so we would like to filter out the signals has frequency larger than that. The lower band-stop frequency is 0.3, so the drifts at the low frequency can be removed. In addition, having a long transition bandwidth can increase the precision in time domain. [8]
   5) Down sampling 250Hz
   6) Run ICA and plot the component in 2D
   7) Remove artifact components

b) In problem 7, we only filter data between 1-50 Hz unlike our problem 8 implementation,in which more frequencies we intepreted as not EEG frequency are filtered out. Under the assumption that the data is cleaner, since ICA is to find data source, the algorithm could converge even faster than in problem 7.[9] Computational time for ICA: Problem 6 - 1min 48s; Problem 7 - 1min 41s; Problem 8 - 1min 12s.

c) 1) The notch filter exclude the noise of power-line.
   2) Our band-pass filter has a lower upper edge comparing with the filter of problem 7. This indicates that we filter out more noise. Furthermore, the longer transition bandwidth could also provide a more precise result.
   3) After the two steps, we have a cleaner data that speed up the process of ICA.

References

[1] Kuan-Chih Huang, Chin-Teng Lin, and Tzyy-Ping Jung. *Tonic and phasic eeg and behavioral changes induced by arousing feedback.* Neuroimage, 52(2):633–642, 2010.

[2] Mike X Cohen. *Analyzing neural time series data : theory and practice.* Cambridge, Massachusetts :The MIT Press, 2014.

[3] Chin-Teng Lin Kuan-Chih Huang and Tzyy-Ping Jung. *An eeg-based fatigue detection and mitigation system.* International Journal of Neural Systems, 26(4), 2016.

[4] Ganesh R. Naik and Dinesh K Kumar.*An Overview of Independent Component Analysis and Its Applications*, Informatica, 35:63–81, 2011.

[5] Donald L. Schomer and Fernando H. Lopes da Silva. *Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*,Lippincott William & Wilkins, 2011. **ISBN** 9780781789424.

[6] Tatsuya Yokota. *Independent Component Analysis for Blind Source Separation*, Remote Sensing, 3(6):1104–1138, 2012, Molecular Diversity Preservation International.

[7] ICLabel Tutorial: EEG Independent Component Labeling. URL: https://labeling.ucsd.edu/tutorial/labels

[8] FIR filter recommended transition bandwidth. URL: https://eeglab.org/others/Firfilt_FAQ.html#q-what-is-the-recommended-transition-bandwidth-for-a-windowed-sinc-fir-high-pass-filter

[9] Aapo Hyvärinen and Erkki Oja. *Independent Component Analysis: Algorithms and Applications* Neural Networks, 13(4-5):411-430, 2000

[10] Peng W. *(2019) EEG Preprocessing and Denoising. In: Hu L., Zhang Z. (eds) EEG Signal Processing and Feature Extraction. Springer, Singapore.*

4 Feedback for Lab 2

4.1 Work Division

Student ID	Name	Be responsible for...
0617028	Ya-Lin Huang	problem 6, 7, 8
0810834	Yi-Ching Chiu	problem 1
0716307	Chia-Ying Hsieh	problem 7
0860838	Ivan Kakhaev	problem 8 (6, 7 manor changes)
0713223	Jian-Xue Huang	problem 3 (problem 4 help to solve)
309591027	Chang-Ling Tsai	problem 2 and 8. Helped solve problem 4 and 5

4.2 Suggestions and Comments

Video could be better, next time give us more time please!

4.2.1 For instructor

4.2.2 For teaching assistant(s)

For Min-Jiun

For Eric