

۱- سیگنال ERP_EEG (ذخیره شده در EEG_ERP.mat) شامل EEG ثبت شده از یک کانال (Pz) است. سیگنال مغزی در پاسخ به یک تحریک بینایی ثبت شده و شامل پتانسیل وابسته به رخداد P300 و EEG زمینه است. ۲۵۵۰ آزمایش انجام شده و سیگنال مغزی در پاسخ به تحریک ثبت شده است. فرکانس نمونه‌برداری ۲۴۰ هرتز بوده و در هر آزمایش از زمان تحریک تا ۱ ثانیه بعد از آن سیگنال ثبت شده است.

با فرض اینکه در تمام آزمایش‌ها تأخیر پاسخ مغزی به تحریک برابر باشد، می‌خواهیم با استفاده از متوسط‌گیری سنکرون حداقل تعداد لازم برای آزمایش‌ها را که بتوان پاسخ P300 را از آن استخراج نمود، تعیین کنیم.

الف) به‌ازای $N=100:100:2500$ ، پاسخ میانگین به دست آمده از N الگو را رسم کنید. نمودارها را در یک شکل رسم کرده و آنها را بر حسب افزایش N مقایسه نمایید.

ب) به‌ازای $N = 1:2550$ ، ماکزیمم قدرمطلق دامن سیگنال را بر حسب تعداد الگوهای میانگین‌گیری شده در یک نمودار رسم کنید.

ج) خطای root mean square بین الگوی میانگین i -ام و الگوی میانگین $i-1$ -ام بر حسب تعداد الگوی میانگین‌گیری شده $N = 1:2550$ در یک نمودار رسم کنید.

د) با استفاده از نتایج به دست آمده در بخش‌های (الف)، (ب) و (ج)، تعداد لازم برای آزمایش‌ها که بتوان پاسخ P300 را از آن استخراج نمود چند است؟ (تعداد به دست آمده را N_0 می‌نامیم)

ه) در یک نمودار پاسخ میانگین به دست آمده از بخش (د) را همراه با پاسخ‌های میانگین زیر رسم کرده و مقایسه نمایید. نتایج به دست آمده را بررسی و تحلیل کنید.

- پاسخ میانگین به ازای $N = 2550$

- پاسخ میانگین به ازای $N = \frac{N_0}{3}$

- پاسخ میانگین به ازای $N = N_0$ با انتخاب N_0 پاسخ تصادفی از بین ۲۵۵۰ پاسخ

- پاسخ میانگین به ازای $N = \frac{N_0}{3}$ با انتخاب $\frac{N_0}{3}$ پاسخ تصادفی از بین ۲۵۵۰ پاسخ

و) چند نمونه از آزمایش‌های واقعی مبتنی بر سیگنال P300 را بررسی کنید. در آزمایش‌های واقعی که از الگوهای P300 استفاده می‌شود (مانند رابط‌های مغز و رایانه مبتنی بر P300)، از چه تعداد تکرار الگوی P300 استفاده می‌شود؟ آیا این تعداد با نتایج به دست آمده در بخش‌های قبل همخوانی دارد؟ علت وجود این تفاوت را توضیح دهید.

۲- فایل SSVEP.mat شامل سیگنال SSVEP ثبت شده از یک کاربر داده شده است. این فایل شامل یک ماتریس SSVEP_Signal (سیگنال EEG ثبت شده از ۶ کانال O1، O2، P8، P7، Oz، Pz و O1 با فرکانس نمونه برداری ۲۵۰ هرتز)، یک ماتریس Events (فرکانس‌های تحریک) و یک ماتریس Event_Samples (نمونه زمانی متناظر با شروع هر تحریک) است. هدف تعیین فرکانس تحریک از روی سیگنال EEG ثبت شده در هر آزمایش است. در دو بخش زیر می‌خواهیم از دو روش "رسم محتوای فرکانسی" و "CCA" استفاده کنیم:

الف) رسم محتوای فرکانسی:

الف ۱) برای هر کانال از داده، با استفاده از یک فیلتر میان‌گذر، فرکانس‌های زیر ۱ هرتز و بالای ۴۰ هرتز را حذف کنید.

الف ۲) ۱۵ آزمایش متناظر با ۱۵ تحریک را هر یک به اندازه پنجره ۵ ثانیه‌ای جدا کنید.

الف ۳) برای هر آزمایش، محتوای فرکانسی هر شش کانال را محاسبه کرده و در یک شکل رسم کنید. می‌توانید از تابع pwelch استفاده کنید. با استفاده از legend مناسب مشخص کنید که هر نمودار مربوط به کدام کانال است.

الف ۴) آیا برای یک آزمایش، همه کانال‌ها از نظر محتوای فرکانسی یکسان هستند؟ تفاوت محتوای فرکانسی کانال‌ها به چه دلیل است؟

الف ۵) آیا برای هر آزمایش، با استفاده از محتوای فرکانسی کانال‌ها می‌توانید فرکانس غالب را تعیین کنید؟ قله‌های فرکانسی مربوط به چه فرکانس‌هایی هستند؟ هر کدام از این قله‌ها چرا ایجاد شده‌اند؟

ب) روش CCA:

ب ۱) ۱۵ آزمایش متناظر با ۱۵ تحریک را هر یک به اندازه پنجره ۵ ثانیه‌ای جدا کنید.

ب ۲) برای تعیین فرکانس‌های غالب از روش آنالیز مولفه‌های کانونی (CCA) استفاده کنید. می‌توانید از دستور canoncorr استفاده کنید. صحت نتیجه طبقه‌بندی را تعیین کنید.

توضیح: سیگنال EEG ثبت شده را به صورت ماتریس $X \in \mathbb{R}^{N \times T}$ (تعداد کانال‌ها و T : تعداد نمونه‌های زمانی) در نظر بگیرید. برای هر فرکانس تحریک (مثلاً f_1 هرتز)، یک ماتریس به صورت زیر ایجاد کنید:

$$Y_{f_1Hz} = \begin{bmatrix} \sin(2\pi f_1 t) \\ \cos(2\pi f_1 t) \\ \sin(2\pi \times 2f_1 \times t) \\ \cos(2\pi \times 2f_1 \times t) \\ \vdots \end{bmatrix}$$

هارمونیک‌ها را تا حدود ۴۰ هرتز ادامه دهید. تعداد ستون‌های ماتریس برابر با تعداد نمونه‌های زمانی T است. الگوریتم CCA را روی زوج ماتریس‌های (X, Y_{f_1Hz}) ، (X, Y_{f_2Hz}) ، (X, Y_{f_3Hz}) ، (X, Y_{f_4Hz}) و (X, Y_{f_5Hz}) اعمال کرده و بزرگ‌ترین ضریب همبستگی کانونی به دست آمده از پنج فرکانس را با هم مقایسه کنید.

ب ۳) بررسی کنید آیا می‌توان با کاهش تعداد کانال‌ها صحت طبقه‌بندی قسمت (ب ۲) را حفظ کرد؟

ب ۴) بررسی کنید آیا می‌توان با کاهش طول پنجره زمانی صحت طبقه‌بندی قسمت (ب ۳) را حفظ کرد؟

۳- می‌خواهیم یک مسئله طبقه‌بندی دو کلاسه را با استفاده از ویژگی‌های CSP حل کنیم. طبقه‌بندی بر روی داده‌های EEG مربوط به ثبت‌های تصورات ذهنی (تصور حرکت پا و انجام عمل تفريق ذهنی) است. دو سری داده در فایل CSPdata.mat در اختیار شما قرار داده شده است. داده‌های آموزش در یک آرایه سه‌بعدی TrainData به ابعاد $(30 \times 256 \times 165)$ (تعداد کانال‌ها ۳۰، تعداد نمونه‌های زمانی ۲۵۶ و تعداد آزمایش‌ها ۱۶۵ است) قرار داده شده‌اند و برچسب‌های مربوط به داده‌های آموزشی در بردار TrainLabel قرار داده شده‌اند. داده‌های آزمون در آرایه سه‌بعدی TestData به ابعاد $(30 \times 256 \times 40)$ (تعداد آزمایش‌های تست ۴۵ است) قرار داده شده‌اند.

الف) ابتدا با استفاده از تمام داده‌های آموزشی، فیلترهای مکانی CSP را به دست آورده و بر روی داده‌های آموزشی اعمال کنید. سیگنال فیلتر شده متناظر با فیلتر اول و آخر را برای نمونه‌هایی از دو کلاس رسم کرده و با هم مقایسه کنید (مشابه صفحه ۴۳ اسلاید EEG Patterns).

ب) فیلترهای مکانی اول و آخر به دست آمده در قسمت الف) را با استفاده از تابع plottopomap.m (تمرین سری سوم) رسم کرده و با هم مقایسه کنید.

ج) با استفاده از روش 4-fold cross-validation داده‌های آموزش را به ۴ دسته تقسیم کنید. هر بار سه دسته را به عنوان داده‌های آموزشی در نظر گرفته و یک دسته را داده اعتبارسنجی قرار دهید. الگوریتم CSP را با استفاده از فقط بخش آموزشی اجرا کرده و ویژگی‌های CSP (واریانس کانال‌های جدید بعد از اعمال CSP) را از داده‌های آموزشی و داده‌های اعتبارسنجی استخراج کنید. از یک طبقه‌بند ساده (مانند kNN، SVM خطی یا LDA) استفاده کرده و آن را با استفاده از ویژگی‌های بخش آموزشی، آموزش داده و بر روی داده‌های اعتبارسنجی اعمال کنید و برچسب‌ها را تعیین کنید. این کار را چهار بار (هر بار متناظر با یک دسته از داده‌ها

به عنوان داده اعتبارسنجی) انجام داده و میانگین صحت را محاسبه نمایید. برای تعیین پارامتر مجهول تعداد فیلترهای CSP (برابر با تعداد ویژگی‌های مورد استفاده)، صحت 4-fold cross-validation را به ازای مقادیر مختلف تعداد فیلترها انجام داده و تعداد فیلتر بهینه را به دست آورید.

د) بهترین طبقه‌بند به دست آمده از قسمت (ج) را با استفاده از داده‌های آموزش، آموزش داده و بر داده‌های تست اعمال کرده و برچسب داده‌های تست را تعیین کنید و در یک بردار TestLabel ذخیره نمایید.

• توضیحاتی در مورد سیگنال‌های مورد بررسی:

- داده‌ها با استفاده از سیستم g.tec GAMMASys ثبت شده‌اند.

- از ۳۰ کانال اندازه‌گیری EEG که در موقعیت استاندارد جهانی سیستم ۲۰-۱۰ تعمیم‌یافته قرار گرفته‌اند، برای ثبت داده‌ها استفاده شده است.

- ۳۰ الکتروود ثبت:

AFz, F7, F3, Fz, F4, F8, FC3, FCz, FC4, T7, C3, Cz, C4, T8, CP3, CPz, CP4, P7, P5, P3, P1, Pz, P2, P4, P6, P8, PO3, PO4, O1, and O2.

- از فیلتر میان‌گذر بین ۰/۵ و ۱۰۰ هرتز استفاده شده و داده‌ها با فرکانس ۲۵۶ هرتز نمونه‌برداری شده‌اند. همچنین نویز برق شهر با فیلتر notch ۵۰ هرتز حذف شده است. آزمایش‌ها به بازه‌های ۱ ثانیه‌ای (۲۵۶ نمونه زمانی) تقسیم شده‌اند.

- هر آزمایش به صورت یک ماتریس (30×256) نشان داده می‌شود که ۳۰ تعداد کانال‌ها بوده و ۲۵۶ تعداد نمونه‌های زمانی است.

- تعداد کل داده‌ها: ۲۱۰ آزمایش

- داده‌های آموزشی: ۱۶۵ آزمایش

- کلاس مربوط به آنها مشخص شده است:

- ۱: تصور حرکت پا

- ۰: عمل تفریق ذهنی

- داده‌های آزمون: ۴۵ آزمایش

- هدف: تعیین برچسب داده‌های آزمون