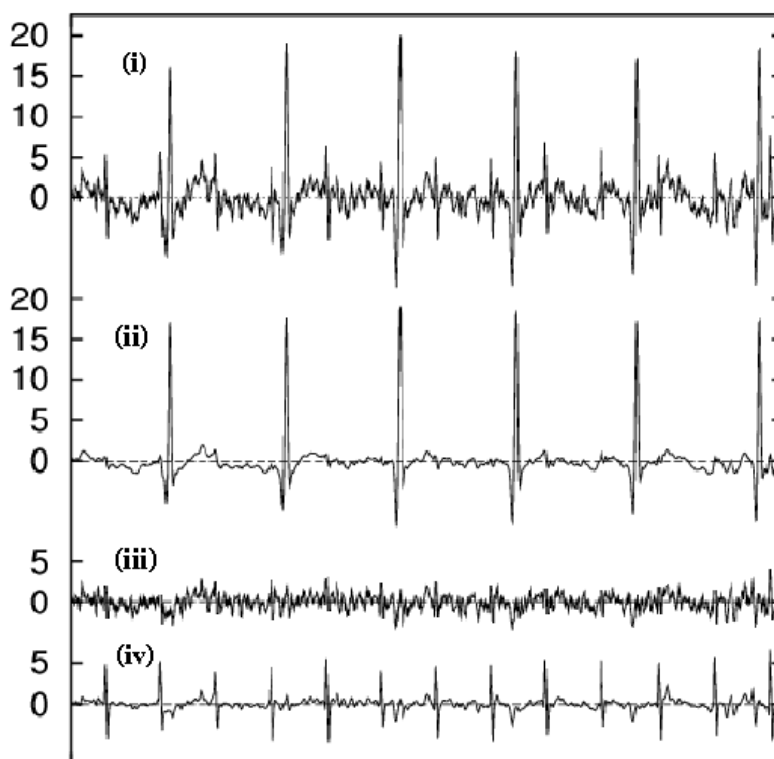


دستور کار آزمایش سوم: جداسازی سیگنال ECG مادر و جنین

در این آزمایش، شما مؤلفه‌های سیگنال قلبی (ECG) مادر (شکل ۱ (ii))، جنین (شکل ۱ (iv)) و همچنین نویز (شکل ۱ (iii)) را با استفاده از روش‌های جداسازی کور منابع (Blind Source Separation (BSS)) از سیگنال مخلوط اولیه (شکل ۱ (i)) جدا خواهید کرد. به صورت دقیق‌تر، شما دو روش Principal Component Analysis (PCA) و Analysis (using Singular Value Decomposition (SVD)) (ICA) را در این مورد به کار خواهید گرفت.



شکل ۱) از بالا: (i) مخلوط سیگنال ECG مادر و جنین، (ii) سیگنال ECG مادر، (iii) نویز و (iv) جنین. پنجره ۵ ثانیه‌ای از این سیگنال‌ها رسم شده است.

روش اول (SVD) فقط از آمارگان‌های مرتبه دوم استفاده می‌کند. در حالی که روش دوم (ICA) از آمارگان‌های مرتبه بالاتر استفاده می‌کند. در این آزمایش شما خواهید دید که در نظر گرفتن توزیع احتمالاتی مؤلفه‌ها بر روی کیفیت عملکرد این دو روش تاثیر خواهد داشت.

در کاربردهای عملی، یکی از چالش‌های جدی برای ارزیابی عملکرد روش‌های جداسازی منابع عدم شناخت ما از نحوه تولید و مشخصات آماری منابع می‌باشد. بنابراین، شما در این آزمایش، از آرتیفکت‌های ساختگی ECG و سیگنال نویز استفاده خواهید کرد و در نتیجه در انتها می‌توان کیفیت جداسازی را ارزیابی کرد.

بخش اول: توضیح ابتدایی در مورد داده‌ها

سیگنال‌های ECG مادر و جنین و همچنین سیگنال نویز به صورت `mcegl.dat`، `fecgl.dat` و `noise1.dat` به شما داده شده که می‌توانید با استفاده از دستور `load` در متلب آن‌ها را بخوانید. هر کدام از این فایل‌ها شامل یک بردار به اسم 'recorded' با یکای mV و فرکانس نمونه‌برداری 256 Hz می‌باشند.

این سه سیگنال را با هم جمع کنید تا سیگنال جدیدی حاصل شود که این سیگنال با تقریب خوبی نمونه‌ای از یک سیگنال واقعی است. به سوالات زیر پاسخ دهید:

- ۱- نمودار سه سیگنال مادر، جنین، نویز و سیگنال مخلوط مشاهده شبیه‌سازی را در حوزه زمان رسم کنید.
- ۲- طیف توان (power spectrum) این سه سیگنال را با استفاده از دستور `pwelch` رسم کنید. آیا محتوای فرکانسی سیگنال مادر و جنین قابل مقایسه هستند؟ شباهت‌ها و تفاوت‌های این محتواهای فرکانسی را به اختصار ذکر کنید.
- ۳- میانگین و واریانس این سه سیگنال را محاسبه و گزارش کنید. آیا قابل مقایسه هستند؟ مقدار واریانس ارتباط خاصی با محتوای فرکانسی سیگنال‌ها دارد؟
- ۴- هیستوگرام این سیگنال‌ها را با استفاده از دستور `hist` متلب محاسبه کرده (برای تقریب PDF مؤلفه‌ها) و در گزارش خود بیاورید. ممان مرتبه ۴ این سیگنال‌ها را با استفاده از دستور `kurtosis` متلب محاسبه کرده و گزارش کنید. آیا می‌توانید با این مقادیر میزان گوسی بودن PDF آن‌ها را تشخیص دهید؟

بخش دوم: جداسازی سیگنال‌ها با استفاده از Singular Value Decomposition (SVD)

یکی از تکنیک‌های استاندارد و قدرتمند و مبتنی بر آمارگان‌های مرتبه دوم برای فیلتر کردن داده، SVD می‌باشد که در این روش، داده‌ها بر روی محورهای Orthogonal مربوط به بیشترین واریانس اطلاعات تصویر می‌شوند. این واریانس‌ها از روی داده‌ها یاد گرفته می‌شوند (Data-driven learning) و از ابتدا میزان این

واریانس‌ها قابل تنظیم نیستند. در این بخش می‌خواهیم با استفاده از SVD محتوای واریانس مؤلفه‌ها را تحلیل کنیم.

۱- داده X.dat را خوانده و با استفاده از دستور plot3ch، که فایل آن به شما داده شده است، آن را نمایش دهید. در این داده دیگر شما به مؤلفه‌های پنهان (سیگنال‌های مادر، جنین و نويز) دسترسی ندارید. در عوض شما به سیگنال مخلوط‌شده که از روی سه کانال ثبت شده دسترسی دارید.

دستور plot3ch دو شکل را نمایش می‌دهد. در اولی، این سه کانال در حوزه زمان نمایش داده می‌شود. در دومی، نمودار پراکندگی (scatter plot) تغییرات یک کانال بر اساس دو کانال دیگر رسم می‌شود. قابل ذکر است که بیشتر اطلاعات حول مبدا مختصات جمع می‌شود (که برای یک داده با میانگین صفر رخ می‌دهد). نقاطی که خارج از این خوشه مرکزی هستند بیشتر مرتبط با پیک کمپلکس‌های QRS هستند.

حالا هدف این است که با استفاده از SVD سیگنال جنین و مادر رو جدا کنیم. یادآوری می‌شود که تجزیه SVD را می‌توان به صورت $Y = USV^T$ که در آن داده اولیه $Y \in \mathbb{R}^{M \times N}$ می‌تواند به صورت سه ماتریس U ، S و V^T تجزیه شود. ماتریس $S \in \mathbb{R}^{M \times N}$ ماتریسی است که مقادیر تکین (singular values) به ترتیب نزولی بر روی قطر اصلی آن قرار داشته و بقیه المان‌های این ماتریس صفر هستند. نمودار این مقادیر تکین eigenspectrum نامیده می‌شود. ستون‌های ماتریس V ، بردارهای ویژه یک ماتریس $C = YY^T$ هستند و ماتریس U شامل تصویرهای Y بر روی بردارهای ویژه ماتریس C می‌باشند. با استفاده از دستور svd متلب، تجزیه SVD را بر روی X.dat انجام داده و ماتریس‌های U ، S و V را به دست آورید.

۲- ستون‌های ماتریس V را با استفاده از دستور plot3dv، که فایل آن به شما داده شده است، رسم کنید. به نکات زیر توجه فرمایید:

الف) تابع plot3dv ماتریس V و مقادیر تکین مربوطه را به صورت ورودی دریافت می‌کند که با استفاده از آن‌ها نرمالیزاسیون را انجام دهد.

ب) تصویر خروجی با فرمت fig. و همچنین ماتریس‌های خروجی از روش SVD را ذخیره کنید.

یادآوری می‌شود که در ماتریس $U \in \mathbb{R}^{M \times M}$ ، که شامل تصویرهای داده اولیه بر روی بردارهای ویژه ماتریس C می‌باشد، تنها N تصویر اول معنادار می‌باشد. زیرا ابعاد تصاویر معنادار (M) نمی‌تواند از تعداد مقادیر تکین (N) بیشتر باشد.

۳- سه ستون اول ماتریس U را رسم کنید. برای هر کدام از این تصاویر، ویژگی‌های غالب (مادر، جنین، نويز یا مخلوط) را تشخیص دهید. کدام یک از این تصاویر برای تشخیص جنین بیشتر کمک‌کننده است؟ همچنین eigenspectrum را با استفاده از دستور stem رسم کرده و در گزارش خود بیاورید.

قابل ذکر است که این نوع projection دیگر از نظر کلینیکی معنادار نیست. تشخیص‌های سنتی از روی سیگنال‌های ECG به وسیله چک کردن کران‌های بالا و پایین آستانه‌های تمایز در تصاویر داده انجام می‌شود. اما با استفاده از تجزیه SVD، ما داده را به فضای غیر کلینیکی می‌بریم که در آن فضا مشخصاً آستانه‌های تشخیصی با فضای کانال‌های ECG تفاوت دارند. بنابراین، اگر ما بخواهیم اطلاعات کلینیکی را در تجزیه SVD بازیابی کنیم باید به صورتی داده اولیه را با استفاده از معکوس تجزیه SVD، پس از انجام اصلاحات مطلوب، به فضای سنسورهای ECG بازگردانیم. به این منظور مقدار تکین مربوط به مؤلفه جنینی را نگه داشته و بقیه مقادیر تکین را صفر کرده و به این صورت ماتریس S را اصلاح کرده و در رابطه SVD گذاشته (با به کار گیری همان ماتریس U و V) و به فضای سنسور جنینی برگردید.

۴- سه کانال سنسور جنینی بازیابی‌شده را نمایش دهید و تاثیر تجزیه SVD و اصلاحات انجام‌شده را توصیف کنید. آیا بازیابی سیگنال جنینی موفقیت آمیز بوده‌است؟

بخش سوم: جداسازی منابع با استفاده از ICA

در این بخش، ما از یک تکنیک مبتنی بر آمارگان‌های مرتبه بالاتر به نام Independent Component Analysis (ICA) برای جداسازی سیگنال مادر، جنین و نویز استفاده خواهیم کرد. می‌توان الگوی مخلوط را به صورت "منابع \times ماتریس مخلوط = مشاهدات" و یا به صورت دقیق‌تر ریاضی به صورت زیر نوشت:

$$\mathbf{X}^T = \mathbf{A}\mathbf{Z}^T$$

قابل ذکر است که برای نمایش ماتریس مشاهدات به صورت channels \times timesamples طرفین رابطه ترانهاد شده است که در آن N تعداد کانال‌ها و M تعداد نمونه‌های زمانی می‌باشد ($\mathbf{X} \in \mathbb{R}^{M \times N}$).

بنابراین، تخمین منابع \mathbf{Z} را می‌توان به صورت زیر نوشت:

$$\hat{\mathbf{Z}}^T = \mathbf{W}\mathbf{X}^T$$

که در آن ماتریس \mathbf{W} ماتریس جداساز بوده و ماتریس \mathbf{W}^{-1} تخمینی از ماتریس ترکیب \mathbf{A} می‌باشد.

۱- تکنیک ICA را بر روی ماتریس مشاهدات (\mathbf{X}^T) با استفاده از تابع ica، که فایل آن به شما داده شده‌است، پیاده کنید و ماتریس‌های \mathbf{W} و $\hat{\mathbf{Z}}$ را به دست آورید و ذخیره کنید. سپس ماتریس \mathbf{W}^{-1} که تخمینی از ماتریس مخلوط \mathbf{A} است را نیز محاسبه و ذخیره کنید.

۲- نمودار پراکندگی (scatter plot) داده اولیه را با استفاده از تابع plot3ch و سه منبع (مؤلفه) را با استفاده از plot3dv رسم کنید. قابل ذکر است که تابع plot3dv ستون‌های ماتریس \mathbf{W}^{-1} را به عنوان ورودی می‌گیرد و می‌تواند چندین بار فراخوانی شود. شکل به دست آمده را با فرمت fig. ذخیره کنید.

۳- سه ستون ماتریس ($\hat{\mathbf{Z}}$)، به طور معادل سه سطر ماتریس $\hat{\mathbf{Z}}^T$ را رسم کنید. ستونی از ماتریس \mathbf{W}^{-1} که به نظر شما بیشتر از بقیه ستون‌ها معرف محتوای مؤلفه جنینی می‌باشد را نگه داشته و بقیه ستون‌ها را صفر کنید. از این ماتریس اصلاح‌شده \mathbf{W}^{-1} و ماتریس $\hat{\mathbf{Z}}$ استفاده کرده و مؤلفه جنینی را به فضای سنسورهای ECG باز گردانید.

۴- داده بازیابی‌شده با ICA را رسم کرده و ذخیره کنید. به نظر شما مؤلفه جنینی با موفقیت بازیابی شده‌است؟

بخش چهارم: مقایسه‌ها

۱- نمودارهای زیر را در یک شکل در فضای سه‌بعدی (متناظر با سه کانال) رسم کنید:

- نمودار پراکندگی ماتریس مشاهدات \mathbf{X} .

- نمودار پراکندگی ماتریس مشاهدات بازیابی‌شده در دو بخش دوم و سوم.

- راستای ستون‌های ماتریس \mathbf{V} .

- راستای ستون‌های ماتریس \mathbf{W}^{-1} .

نمودارها را با هم مقایسه کنید و نتایج به دست آمده از دو روش را تحلیل کنید.

زاویه بین راستاهای مختلف را با دستور dot محاسبه کنید. نرم‌محورهای مختلف را نیز محاسبه و مقایسه کنید.

سوال اساسی این است که این دو روش چقدر در جداسازی موفق بوده‌اند؟ تنها راه پاسخگویی در این موارد این است که دانش ما درباره منابع و نویز کامل باشد. بنابراین، با توجه به اینکه در اینجا از داده‌های شبیه‌سازی استفاده شده‌است، دانش ما درباره سیگنال مطلوب (fECG) و نامطلوب (mECG+noise) کامل است. درباره ماتریس داده اولیه \mathbf{X} ، سیگنال مطلوب fECG در fecg2.dat، که نسخه scale شده داده fecg1.dat می‌باشد، به شما داده شده‌است.

۲- با رسم سیگنال مطلوب بازسازی‌شده به دست آمده از روش‌های SVD و ICA (یکی از کانال‌های سیگنال بازسازی‌شده با این دو روش) و سیگنال ایده‌آل مطلوب fecg2، عملکرد نسبی آن‌ها را مقایسه کرده و روش برتر را مشخص کنید.

۳- سپس مقدار ضریب همبستگی (correlation coefficient) بین سیگنال ایده‌آل fecg2 و سیگنال بازسازی‌شده توسط این دو روش (یکی از کانال‌های سیگنال بازسازی‌شده با این دو روش) را با دستور corrcoef متلب محاسبه کرده و مقایسه را انجام دهید.

۴- مقایسه نهایی بین دو روش SVD و ICA را انجام داده و مزیت‌ها و معایب این دو روش را به صورت کلی بیان نمایید.

۵- مهم‌ترین نکته‌ای که از این آزمایش آموخته‌اید را بیان نمایید.