تمرین سوم تئوری درس پردازش سیگنال های EEG

نام: محمد سعید حسن نیا

شماره دانشجویی: 96108515

سوال 1

EEG اغلب توسط پتانسیل های الکترو فیزیولوژیکی مرتبط با انقباض عضلانی ناشی از جویدن، اخم کردن و ... دچار آرتیفکت می شود. فیلتر های پایین گذر معمولاً برای حذف آرتیفکت عضلانی استفاده می شوند. با این حال از آن جا که طیف فرکانس آرتیفکت های ماهیچه ای اغلب با طیف فرکانس overlap دارد فیلتر های فرکانس را اگر اعمال کنیم ممکن است اطلاعات ارزشمند سیگنال را هم از بین ببرند. روش جدید تر ICA هم کاملاً جداسازی نمی کند crosstalk و فعالیت ماهیچه ای می تواند مشاهده شود. CCA روشی جدید تر برای حذف آرتیفکت عضله از سیگنال EEG است. این الگوریتم یک الگوریتم BSS (جداسازی کور) می باشد. برای حل مسئله BSS، CCA منابع متقابل غیر همبسته را فرض می کند که حداکثر خودهمبستگی را دارند. این الگوریتم برای جداسازی آرتیفکت عضله از EEG مناسب است زیرا سیگنال EEG خودهمبستگی بالایی دارد درحالی که آرتیفکت ماهیچه خودهمبستگی کمی دارد. فرض می کنیم X و Y دو بردار تصادفی مطابق روبرو باشند و بین متغیرها همبستگی وجود داشته باشد CCA دو مجموعه از بردار های پایه را به دست می آورد (یکی برای X و یکی برای Y) به طوری که همبستگی بین Project متغیرها بر روی این بردارهای پایه حداکثر همبستگی را با یکدیگر دارند. برای مثال بر X و Y، CCA بردارهای w_x و w_y ای را پیدا می کند که correlation بین u و v را ماکزیمم کند که

$$u = w_x^T X \qquad v = w_y^T Y$$

است.

$$\max_{w_x w_y} \rho(u, v) = \frac{E[uv]}{\sqrt{E[u^2]E[v^2]}} = \frac{E\left[w_x^T X \, w_y^T Y\right]}{\sqrt{E\left[(w_x^T X)(w_x^T X)\right] E\left[(w_y^T Y)(w_y^T Y)\right]}}$$

$$= \frac{w_x^T C_{xy} w_y}{\sqrt{(w_x^T C_{xx} w_x)(w_y^T C_{yy} w_y)}}$$

حال اگر از رابطه ی صفحه ی قبل مشتق بگیریم به عبارت زیر می رسیم

$$\begin{cases} C_{xx}^{-1} C_{xy} C_{yy}^{-1} C_{yx} \hat{w}_x = \rho^2 \hat{w}_x \\ C_{yy}^{-1} C_{yx} C_{xx}^{-1} C_{xy} \hat{w}_y = \rho^2 \hat{w}_y \end{cases}$$

اگر این رابطه GEVD را حل کنیم به w_x ای میرسیم منبع اول که بیشترین خودهمبستگی را دارد می دهد در گام بعدی برای بدست آوردن بردارهای جدید از فضای عمود استفاده می کنیم تا به بیشترین ناهمبستگی برسیم.

هـ) هنگامی که BSS-CCA روی سیگنال اعمال می شود آرتیفکت عضلانی می تواند از بین رود. (دلیل آن در قسمت الف گفته شد) به وسیله ای این که ستون هایی که فعالیت آرتیفکت را نشان می دهند از بین رود یعنی

ماتریس MIX

$$X_{clean}(t) = A_{clean}\, \underline{Z(t)}$$

سورس هایی به دست آمده با استفاده از BSS

سؤال ۶) در روش PCA به دنبال ماکزیمم کردن واریانس منابع استخراج شده می باشیم

همچنین در PCA ابتدا میانگین داده ها را صفر می کنیم یعنی در PCA داریم:

$$\arg\max_W Var(Y) \xrightarrow[Var(Y)=E[Y^2]-E[Y]^2]{\text{knowing that}} \arg\max_W E[Y^2]$$

$$y(t) = W^T x \xrightarrow{\text{داریم}} \arg\max_W E[YY^T] \xrightarrow[\text{است با}]{\text{معادل}} \arg\max_W E[W^T X X^T W]$$

$$\xrightarrow[\text{است با}]{\text{معادل}} \arg\max_W W^T E[XX^T] W = W^T C_x W$$

همچنین برای CCA باید کسر زیر را ماکزیمم کنیم

$$\frac{E[UV]}{\sqrt{E[U^2]E[V^2]}} = \frac{W_x^T C_{xy} W_y}{\sqrt{(W_x^T C_{xx} W_x)(W_y^T C_{yy} W_y)}}$$

براساس نکته ی مقادیر $y^T w_y$ و $x^T w_x$ به هم شبیه هستند بنا براین خواهیم داشت

	maximize	A	B	V
PCA	total scatter (variation)	C_X	I	مقادیر ویژه های C_X
CCA	correlation	$\begin{pmatrix} 0 & C_{XY} \\ C_{YX} & 0 \end{pmatrix}$	$\begin{pmatrix} C_{XX} & 0 \\ 0 & C_{YY} \end{pmatrix}$	مقادیر ویژه های $B^{-1}A$

Scanned with CamScanner

سؤال هفتم) الف) برای متغیر گاوسی $X \sim N(m, \sigma^2)$ می‌دانیم: $(Z \sim N(0,1))$

$$X = \sigma Z + m \quad , \quad f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \quad \text{for all } z \in \mathbb{R}$$

$$\Longrightarrow E\{(X-m)^4\} = \int_{-\infty}^{\infty} (x-m)^4 \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = \int_{-\infty}^{\infty} \sigma^4 z^4 \times \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$$

$$= \sigma^4 \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} z^4 e^{-\frac{z^2}{2}} dz = \sigma^4 F$$

$$F = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} z^4 e^{-\frac{z^2}{2}} dz = 2\int_0^{\infty} \frac{1}{\sqrt{2\pi}} z^4 e^{-\frac{z^2}{2}} dz$$

$$\overset{z=\sqrt{2w}}{=\!=\!=} \frac{2}{\sqrt{2\pi}} \int_0^{\infty} (2w)^2 e^{-\frac{2w}{2}} \frac{dw}{\sqrt{2w}} = \frac{2^2}{\sqrt{\pi}} \int_0^{\infty} w^{2-\frac{1}{2}} e^{-w} dw$$

$$= \frac{2^2}{\sqrt{\pi}} \Gamma\left(2+\frac{1}{2}\right) = \frac{2^2}{\sqrt{\pi}} \times \frac{4!}{(4)^2 \times 2!} \sqrt{\pi} = 3$$

$$\Longrightarrow E\{(X-m)^4\} = \sigma^4 F = 3\sigma^4 \Longrightarrow \boxed{K_4 = 3}$$

بنا براین مقدار تابع گفته شده برای متغیر گاوسی دلخواه برابر 3 می‌باشد $\Longleftarrow$ برای استفاده از آن در ICA می‌توان آن را منهای 3 کرد و مشابه آن چه در اسلایدها بیان شده است اندازه آن را برای هر منبع ماکزیمم کرد.

ب) برای این قسمت ابتدا ثابت می‌کنیم توزیع گاوسی بیشترین آنتروپی را در بین توزیع‌های با دامنه‌ی نامحدود و میانگین و واریانس محدود دارد بنا براین داریم:

we must maximize : $H(p) = \int_a^b p(x) \lg p(x)\, dx$

with respect to $p(x)$ s.t. the constraints:
$$\begin{cases} p(x) \geq 0 \\ \int_a^b p(x)\, dx = 1 \\ \int_{-\infty}^{\infty} x\, p(x)\, dx = \mu < \infty \\ \int_{-\infty}^{\infty} x^2 p(x)\, dx = \sigma^2 < \infty \end{cases}$$

با استفاده از متد لاگرانژ برای بهینه‌سازی در حضور قیدها objective func را به صورت زیر می‌نویسیم:

$$J(p) = -\int_{-\infty}^{\infty} p(x) \ln p(x)\, dx + \lambda_0 \left(\int_{-\infty}^{\infty} p(x)\, dx - 1\right) + \lambda_1 \left(\int_{-\infty}^{\infty} x\, p(x)\, dx - \mu\right) + \lambda_2 \left(\int_{-\infty}^{\infty} x^2 p(x)\, dx - \sigma^2\right)$$

بنابراین خواهیم داشت:

$$\frac{\partial}{\partial p(x)\, dx} J(p) = -\ln p(x) - 1 + \lambda_0 + \lambda_1 x$$

$$\frac{\partial^2}{\partial p(x)^2 dx} J(p) = \frac{-1}{p(x)}$$

$$\Longrightarrow p(x) = e^{(\lambda_0 - 1) + \lambda_1 x + \lambda_2 x^2} = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-x^2}{2\sigma^2}}$$

حال اگر y یک متغیر گاوسی با ماتریس کوواریانس مشابه y باشد آن گاه $J(y) \geq 0$ است (برای توزیع گاوسی صفر می‌باشد) $\Longleftarrow$ باید wهایی را پیدا کنیم که آن را ماکزیمم کند. (یعنی از گاوسی بودن فاصله بگیریم.)

ج) mutual information یکی از مقادیری است که میزان اطلاعات یک متغیر تصادفی در مورد دیگری را اندازه‌گیری می‌کند. mutual information بالا نشان دهنده کاهش زیادی در uncertainty می‌باشد mutual information کم نشان دهنده کاهش کمی در uncertainty می‌باشد بنابراین mutual information صفر بین 2 متغیر تصادفی به این معنی است که آن دو از هم مستقل می‌باشند. در واقع به طور مثال اگر X و Y مستقل باشند آن گاه می‌دانیم دانستن X اطلاعاتی در مورد Y به ما نمی‌دهد. بنابراین در کاربرد در ICA با توجه به این که دوست داریم y_iها از هم مستقل باشند در نتیجه باید معیار mutual information را مینیمم کنیم.

سؤال 4)

می دانیم correlation ، X و Y برابر است با کواریانس نرمالایز شده.

یعنی $corr(X,Y)=\frac{cov(X,Y)}{\sigma_X \sigma_Y}$. کورلیشن 2 متغیر عددی است که بین -1 و +1 متغیر است.

$corr(X,Y)$ فقط برای یک رابطه شیب دار کامل به سمت بالا +1 است. (منظور از کامل این است که همه ی مشاهدات روی یک خط قرار دارند). همچنین فقط برای یک رابطه کامل به سمت پایین -1 است. می دانیم ICA ترتیب و دامنه ی منابع را به درستی نمی دهد. بنابراین ابتدا باید کورلیشن برای s و s_{est} را حساب کنیم. (یک ماتریس بدست می آید). سپس بزرگ ترین مقدار در آن ماتریس را محاسبه می کنیم. اندیس های i و j که برای بزرگ ترین مقدار بدست می آید نشان می دهد منبع i ام در s_{est} با منبع j ام در s متناظر است $\Leftarrow$ عدد به دست آمده در واقع میزان شباهت منبع بدست آمده به منبع واقعی را می دهد. سپس سطر i و ستون j را حذف می کنیم. 32 مرحله این کار را انجام می دهیم. در انتها هر چه جمع این 32 عدد بدست آمده بیشتر باشد نشان دهنده این است که الگوریتم ما بهتر است (یعنی منابعی که تولید می کند به منابع واقعی متناظر شباهت بیشتری دارد) بنابراین برای روند این الگوریتم داریم:

knowing that
$$\begin{cases} x(t) = \underbrace{A}_{\text{مجهول}}\,\underbrace{s(t)}_{\text{مجهول}} \\ s(t) \in R^{32\times T} \\ s_{est} \in R^{32\times T} \end{cases}$$

1) calculate C_1 # $C_1 = \frac{1}{T}\sum_{t=1}^{T} s(t)\, s_{est}^T$

2) calculate C_2 # $C_2 = \frac{1}{T}\sum_{t=1}^{T} s(t)\, s^T(t)$

3) calculate C_3 # $C_3 = \frac{1}{T}\sum_{t=1}^{T} s_{est}(t)\, s_{est}^T(t)$

4- corr_mat = zeros(32, 32)

5) for i=1:32
    for J=1:32
        $corr\_mat(i,J) = \frac{C_1(i,J)}{sqrt(C_2(i,i)\,C_3(J,J))}$
    end
end

evaluater = 0

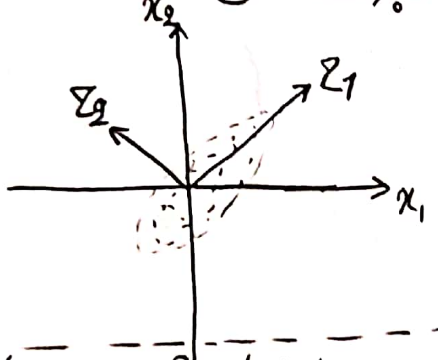
```
6) for i=1:32
    [indx, max-value] = max(corr_mat)     # find the max value in corr_mat and its indx → (i,j)
    evaluater += abs(max_value)
    omit indx(1)th row and indx(2)-th column in corr_mat
end
```

مقدار evaluater هر چقدر بیشتر باشد نشان دهنده این است الگوریتم بهتر عمل کرده است.

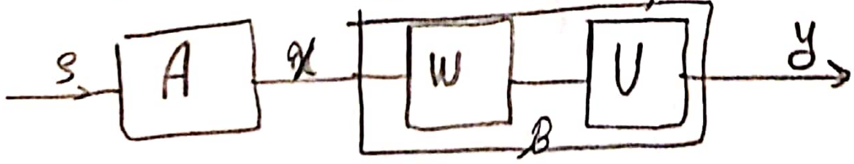
سؤال 5

الف) در این جا فضای ما 2 بعدی است. در PCA جهت بیشترین تغییرات بدست می آید. وقتی PCA اعمال می کنیم اولین بردار ویژه‌ای که به ما می دهد 1 می باشد و مقدار ویژه که با آن متناظر است متناظر است با پراکندگی در جهت 1. جهت دوم عمود بر 1 می باشد.

⟸ بهینه سازی با یک پارامتر یا 2 پارامتر مجهول



ب) اگر داشته باشیم $y = Bx$. می توانیم عناصر روی قطر اصلی ماتریس مجهول را 1 درنظر بگیریم چون ابهام در دامنه داریم. حال اگر بخواهیم به صورت مستقیم این ماتریس را به دست آوریم به طور کلی n^2-n تا مجهول خواهیم داشت درحالی که $E\{y_i y_j\}=0$ فقط $\binom{n}{2}$ تا معادله به ما می دهد. اگر داده ها را ابتدا سفید سازی کنیم U یک ماتریس دوران خواهد بود.



برای این مثال باید داشته باشیم

$$BB^T = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix} \Longrightarrow \begin{bmatrix} 1 & a \\ b & 1 \end{bmatrix}\begin{bmatrix} 1 & b \\ a & 1 \end{bmatrix} = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix} \Longrightarrow a+b=0 \to \boxed{a=-b}$$

اگر $a=-\theta$ آن گاه $b=\theta \Longrightarrow B_{new} = \begin{bmatrix} 1 & -\theta \\ \theta & 1 \end{bmatrix}$

$$x(t) \xrightarrow{PCA} z(t) \qquad B_{New} = \begin{bmatrix} 1 & -\theta \\ \theta & 1 \end{bmatrix}$$

$$y(t) = B_{New}\, z(t) \Longrightarrow \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & -\theta \\ \theta & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

$$\Longrightarrow \begin{cases} y_1 = z_1 - \theta z_2 \\ y_2 = \theta z_1 + z_2 \end{cases}$$

$$\left(\mathrm{cum}(y_1^2, y_2)\right)^2 = \left(\mathrm{cum}\left((z_1-\theta z_2), (z_1-\theta z_2), (\theta z_1 + z_2)\right)\right)^2$$

$$= E\left\{(z_1-\theta z_2)^2 (\theta z_1 + z_2)\right\}^2$$

$$= E\left\{(z_1^2+\theta^2 z_2^2 - 2\theta z_1 z_2)(\theta z_1 + z_2)\right\}^2$$

$$= E\left\{\theta z_1^3 + \theta^3 z_2^2 z_1 - 2\theta^2 z_1^2 z_2 + z_1^2 z_2 + \theta^2 z_2^3 - 2\theta z_1 z_2^2\right\}^2$$

$$= \left(\theta E\{z_1^3\} + \theta^3 E\{z_2^2 z_1\} - 2\theta^2 E\{z_1^2 z_2\} + E\{z_1^2 z_2\} + \theta^2 E\{z_2^3\} - 2\theta E\{z_1 z_2^2\}\right)^2$$

$$\left(\mathrm{cum}(y_1, y_2^2)\right)^2 = \left(\mathrm{cum}\left((z_1-\theta z_2), (\theta z_1 + z_2), (\theta z_1 + z_2)\right)\right)^2$$

$$= E\left\{(z_1-\theta z_2)(\theta z_1 + z_2)^2\right\}^2$$

$$= E\left\{(z_1-\theta z_2)(\theta^2 z_1^2 + z_2^2 + 2\theta z_1 z_2)\right\}^2$$

$$= E\left\{\theta^2 z_1^3 + z_2^2 z_1 + 2\theta z_1^2 z_2 - \theta^3 z_1^2 z_2 - \theta z_2^3 - 2\theta^2 z_2^2 z_1\right\}^2$$

$$= \left(\theta^2 E\{z_1^3\} + E\{z_2^2 z_1\} + 2\theta E\{z_1^2 z_2\} - \theta^3 E\{z_1^2 z_2\} - \theta E\{z_2^3\} - 2\theta^2 E\{z_2^2 z_1\}\right)^2$$

$$\Longrightarrow f(\theta) = \frac{1}{\left(\mathrm{cum}(y_1^2, y_2)\right)^2 + \left(\mathrm{cum}(y_1, y_2^2)\right)^2} \qquad (5\text{-}2)$$

د) برای این قسمت می دانیم استقلال را می توانیم در کومولان های مشترک ببینیم یعنی
یعنی برای متغیرهای مستقل X و Y داریم : $\mathrm{cum}\{X^{k_1}, Y^{k_2}\} = 0$

اگر بخواهیم $f(\theta)$ ماکزیمم شود یعنی $\Leftarrow$ $\underbrace{\left(\mathrm{cum}(y_1, y_2^2)\right)^2 + \left(\mathrm{cum}(y_1^2, y_2)\right)^2}$

باید مینیمم شود و طبق صفحه بالا استقلال را می توان در کومولان های مشترک T دید بنابراین
T باید برای منابع مستقل ما باید صفر باشد $\Leftarrow$ اگر الگوریتم گرادیان افزایشی خوب عمل کند
منابع مستقل بدست می آید یعنی به مقدار اولیه الگوریتم بستگی دارد.

سؤال ۶) برای حل این سؤال با استفاده از GEVD ابتدا مفاهیم و مراحل حل با استفاده از GEVD را مروری کنیم:

مسئله ریاضی GEVD برای دو ماتریس A و B به صورت زیر تعریف می‌شود: $A v_i = \lambda_i B v_i$

برای کاربرد GEVD در حذف آرتیفکت می‌دانیم $y(t) = w^T x(t) = w^T \underbrace{(x_1^{(s)}(t) + x_2^{(s)}(t) \cdots)}_{x_s(t)} + w^T \underbrace{(x_1^{(n)} \cdots)}_{x_n(t)}$

هدف این است که w، $w^T x_n$ را تا حد ممکن کوچک کند. سپس با توجه دانسته های مان در مورد منابع برای حل با استفاده از GEVD یک معیار را در نظر می گیریم و سعی می کنیم آن را ماکزیمم کنیم. پس آن را ابتدا به یک نسبت رایلی برای ۲ بردار F و E ای تبدیل می کنیم. سپس GEVD(E,F) را حل می کنیم.
حال که نحوه‌ی مراحل و عملکرد GEVD را متوجه شدیم برای حل مسئله با استفاده از این روش داریم:

بنابر فرض سؤال داریم:

$f_s = 100$ Hz
Duration-of simulation $= 100$ $\Big\} \longrightarrow$ sample numbers $= 100 \times 100 = 10^4$ sample

observed signals $\longrightarrow x[n] \in \mathbb{R}^8$

همچنین $x(t) = x_1(t) + x_2(t) + x_3(t)$. همچنین داریم:

$s_1 \longrightarrow$ سیگنال متناوب مثلثی $\Longrightarrow$ استفاده از Periodicity Maximization $\xrightarrow{\text{مطابق اسلاید}}$ $\epsilon(w) = \frac{w^T \tilde{P}_x w}{w^T C_x w}$

$s_2 \longrightarrow$ سیگنال burst-like $\Longrightarrow$ استفاده از NON-Stationary maximization $\xrightarrow{\text{مطابق اسلاید}}$ $f(w) = \frac{w^T \tilde{C}_x w}{w^T C_x w}$

بنابراین برای سیگنال s_1 داریم

1) calculate P_x # $P_x = E_t\{x(t)\, x^T(t+\tau)\} \xrightarrow{\tau_0 = 4s = 400\ \text{sample}} \frac{1}{10^4 - 400} \sum_{n=0}^{10^4-401} x[n]\, x[n+400]$

2) calculate $\tilde{P}_x$ # (برای تقارن) $\tilde{P}_x = \frac{1}{2}(P_x + P_x^T)$

3) calculate C_x # $C_x = E_t\{x(t)\, x^T(t)\} = \frac{1}{10^4} \sum_{n=0}^{10^4-1} x[n]\, x^T[n]$

4) $[W, \Lambda] = eig(\tilde{P}_x, C_x)$ #solve GEVD

5) find the maximum λ & corresponded $\underbrace{w_1}_{\text{eigen vector}}$ to maximize Rayleigh quotient

6) $y[n] = \begin{bmatrix} y_1[n] \\ \vdots \\ y_N[n] \end{bmatrix} = \begin{bmatrix} w_1^T \\ \vdots \\ w_N^T \end{bmatrix} x[n] = W^T x(t)$ # $y_1[n] = w_1^T x[n]$ y is a copy of s

7) $y\_new[n] = \begin{bmatrix} y_1[n] \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ # denoising (y_1 is desirable)

8) $x_1[n] = (W^T)^{-1} y_{new}[n]$

برای سیگنال s_2 نیز داریم

1) calculate $\tilde{C}_x$ # $\tilde{C}_x = E_\theta\{x(\theta)\, x^T(\theta)\} = \frac{1}{\text{sum}(T_{on})} \sum_{n \in T_{on}} x[n]\, x^T[n]$

2) calculate C_x # $C_x = \frac{1}{10^4} \sum_{n=0}^{10^4-1} x[n]\, x^T[n]$

3) $[W, \Lambda] = \text{eig}(\tilde{C}_x, C_x)$ # solve GEVD

5, 6, 7, 8) $\longrightarrow$ similar to the above for s1

برای حل مسئله با استفاده از DSS داریم:

برای این الگوریتم ابتدا باید داده‌ها را سفیدسازی کنیم در نتیجه تعداد مشاهدات برابر تعداد تعداد منابع می‌شود. ($z[n]$ را به‌عنوان مشاهدات سفیدسازی‌شده درنظر می‌گیریم)

حال الگوریتم EM را اجرا می‌کنیم:

0) initialize w_p with random value

repeat

1) step E

A) $r_1^{(i)}[n] = w_p^{(i)^T} x[n]$

B) repeated_num = $\frac{\text{period samples len}}{\text{all samples len}}$

repeated_vec = $\frac{1}{\text{repeated\_num}} \sum_{n=1}^{\text{repeated\_num}} r_1^{(i)}[400(n-1) : 400n-1]$

$r_1^{(i)+} = [\ ]$

for i = 1 : repeated_num

$r_1^{(i)+} = [r_1^{(i)+}, \text{repeated\_vec}]$

end

2) step M

$w_p^{(i)+} = \sum_{n=1}^{10^4} x[n]\, r_p^{(i)+}[n]$ # calculate new ML estimation $w_p^{(i)}$

end

حال که w_1 را به دست آوردیم داریم :

$$x_1[n] = B^{\dagger} \, w \, r_1[n]$$

ب) برای سیگنال s_2 مطابق قسمت الف عمل می‌کنیم منتهی مرحله‌ی B را به عبارت زیر باید تغییر دهیم بنابراین داریم

$$r_2^{(i)+} = r_2^{(i)}[n] \cdot T_{on}[n]$$

Scanned with CamScanner

سؤال ۷

می دانیم در الگوریتم DSS فرض بر آن است که مشاهدات ما سفیدسازی بشود و سپس الگوریتم Expectation_Maximization را روی آن اعمال کنیم. بنابراین در گام اول مشاهدات $x(t)$ را سفیدسازی می کنیم و خروجی را $z(t)$ نام گذاری می کنیم. بنابراین حالا مشاهدات برابر تعداد منابع خواهد بود. بنابراین داریم:

$$z(t) = D\,x(t) \qquad \text{و} \qquad z(t) \in \mathbb{R}^p$$

حال داده های ما آماده است تا الگوریتم EM روی آن ها اعمال شود. دقت شود که در متن صورت سؤال ذکر شده است f یک تبدیل خطی و حقیقی است. بنابراین داریم:

شروع از یک مقدار تصادفی برای بردار w_p : $(w_p \in \mathbb{R}^p)$

۱) گام E:

الف) محاسبه ی تخمین نویزی منبع:

$$s_1(t) = w_p^T z(t)$$

اطلاعات خوبی در حوزه تبدیل دارم + تبدیل خطی است

$$\Longrightarrow s_1(t) = w_p^T z(t) \xrightarrow{\text{حوزه } f} f(s_1(t)) = f(w_p^T z(t)) = w_p^T f(z(t))$$

$$\Longrightarrow f(z(t)) = \begin{cases} z^{(1)}(t^{(1)}) \\ z^{(2)}(t^{(2)}) \\ z^{(3)}(t^{(3)}) \end{cases} \Longrightarrow f(s_1(t)) = \begin{cases} w_p^T z^{(1)}(t^{(1)}) = s_1^{(1)}(t^{(1)}) & t^{(1)} = 1:50 \\ w_p^T z^{(2)}(t^{(2)}) = s_1^{(2)}(t^{(2)}) & t^{(2)} = 1:2 \\ w_p^T z^{(3)}(t^{(3)}) = s_1^{(3)}(t^{(3)}) & t^{(3)} = 1: \end{cases}$$

ب) حذف نویز براساس اطلاعات اولیه منبع:

با توجه به این که می دانیم عمده انرژی سیگنال در حوزه ی تبدیل در چه نمونه هایی قرار دارد خواهیم داشت:

$$s_1^{(1)+}(t^{(1)}) = \begin{cases} s_1^{(1)}(t^{(1)}) & t^{(1)} = t_1^{(1)} : t_2^{(1)} \\ 0 & \text{O.W} \end{cases} \qquad s_1^{(2)+} \begin{cases} s_1^{(2)}(t^{(2)}) & t^{(2)} = t_1^{(2)} : t_2^{(2)} \\ 0 & \text{O.W} \end{cases}$$

$$s_1^{(3)+}(t^{(3)}) = \begin{cases} s_1^{(3)}(t^{(3)}) & t^{(3)} = t_1^{(3)} : t_2^{(3)} \\ 0 & \text{O.W} \end{cases}$$

۲) گام M:

$$w_p^+ = \sum_{t^{(1)}=1}^{50} s_1^{(1)+}(t^{(1)})\, z^{(1)}(t^{(1)}) + \sum_{t^{(2)}=1}^{25} s_1^{(2)+}(t^{(2)})\, z^{(2)}(t^{(2)}) + \sum_{t^{(3)}=1}^{25} s_1^{(3)+}(t^{(3)})\, z^{(3)}(t^{(3)})$$

$$w_p = \frac{w_p^+}{\|w_p^+\|}$$

الگوریتم را آن قدر ادامه می دهیم (گام E و M) تا به همگرایی برسیم. در نهایت نیز خواهیم داشت:

$$s_1(t) = w^T z(t)$$

در واقع چون عملگرها خطی است... w ای که الگوریتم EM در حوزه تبدیل می دهد مانند w ای است که در حوزه زمان می گیریم.

سؤال ۸)

ابتدا برای فرضیات سؤال داریم:

$x(t) \in R^{32}$ $\begin{cases} \text{Duration time} = 100s \\ f_s = 128HZ \end{cases} \Longrightarrow n = \text{samples} = 1 : 12800$

دراین جا بخش نامطلوب سیگنال مشاهدات شبه متناوب می باشد با دوره تناوب τ_t

$$\epsilon(w) = \frac{E_t\{y(t)y(t+\tau_t)\}}{E_t\{y^2(t)\}} \xrightarrow{y(t)=w^T x(t)} \frac{E_t\{w^T x(t)\, x^T(t+\tau_t) w\}}{E\{w^T x(t)\, x^T(t)\, w\}}$$

$E\{y(t)y^T(t)\}$

$$= \frac{w^T E_t\{x(t)\, x^T(t+\tau_t)\}\, w}{w^T E_t\{x(t)\, x^T(t)\}\, w} = \frac{w^T P_x w}{w^T C_x w}$$

$$\xrightarrow[C_x \text{ متقارن}]{} \mathcal{E}(w) = \frac{\frac{1}{2}(w^T P_x w + w^T P_x^T w)}{w^T C_x w} = \frac{w^T \tilde{P}_x w}{w^T C_x w}$$

$$P_x = E_t\{x(t)\, x^T(t+\tau_t)\} = \frac{1}{12800-\tau_t} \sum_{t=1}^{12800-\tau_t} x(t)\, x^T(t-\tau_t)$$

$$\tilde{P}_x = P_x + P_x^T \qquad C_x = E_t\{x(t)\, x^T(t)\} = \frac{1}{12800} \sum_{t=1}^{12800} x(t)\, x^T(t)$$

بنابراین $\mathcal{E}(w) = \frac{w^T \tilde{P}_x w}{w^T C_x w}$ تابع هدف خواهد بود.

ب) دراین قسمت قید نرم بودن را هم اضافه می کنیم. می دانیم فرکانس نمونه برداری برابر f_s می شود یعنی فاصله بین هر ۲ نمونه ی نمونه برداری شده برابر $\frac{1}{f_s}$ است. دراین صورت با استفاده از مشتق اولیه داریم:

$$y'(t) = \frac{(y(t) - y(t-1))}{\Delta x} = \frac{(y(t) - y(t-1))}{\frac{1}{f_s}} = f_s\,(y(t) - y(t-1))$$

می دانیم $y(t) = w^T x(t)$ بنابراین داریم:

$$y'(t) = w^T (x(t) - x(t-1)) f_s$$

می خواهیم y هموار باشد بنابراین مینیمم $E_t\{y'^2_{(t)}\}$ را می خواهیم بنابراین داریم:

$$E_t\{y'^2_{(t)}\} = E_t\{y'_{(t)} y'^T_{(t)}\} = E_t\{w^T (x(t) - x(t-1)) f_s (x(t) - x(t-1))^T w f_s\}$$

$$= f_s^2 w^T E_t\{(x(t) - x(t-1))(x(t) - x(t-1))^T\} w$$

$$= f_s^2 w^T \times \Big(\underbrace{\frac{1}{f_s} \sum_{t=2}^{12800} (x(t) - x(t-1))(x(t) - x(t-1))^T}_{R_x} \Big) w$$

$$= w^T f_s L w = w^T \tilde{R}_x w$$

بنابراین ما می خواهیم $E_t\{y'(t)^2\} = w^T \tilde{R}_x w$ مینیمم شود. از طرفی در تابع قسمت قبل می خواستیم آن را ماکزیمم کنیم بنابراین برای تابع این قسمت خواهیم داشت:

$$\mathcal{E}_{new}(w) = \frac{w^T \tilde{P}_x w}{w^T C_x w + \lambda w^T \tilde{R}_x w} = \frac{w^T \tilde{P}_x w}{w^T (C_x + \lambda \tilde{R}_x) w}$$

د) تابع بالا به فرم یک تابع رایلی است بنابراین مطابق اسلایدها به حل مسئله GEVD می پردازیم:

$$GEVD(\tilde{P}_x, C_x + \lambda \tilde{R}_x) \longrightarrow w_1, \dots, w_{32} \ , \ \lambda_1, \dots, \lambda_{32} \quad \lambda_1 \geq \dots \geq \lambda_n$$

چون λ_1 بزرگ ترین مقدار ویژه است داریم:

$$S_1 = w_1^T x(t) \qquad W = [w_1, \dots, w_{32}]$$

هـ) S_1 منبع مطلوب ما نیست بنابراین داریم:

$$S_N(t) = \begin{bmatrix} S_1(t) \\ 0 \\ \vdots \\ 0 \end{bmatrix} \Longrightarrow \boxed{x_d(t) = x(t) - (W^T)^{-1} S_N(t)}$$

سؤال ۹

کافی است از power method استفاده کنیم. این الگوریتم یک روش برای یافتن مقدار ویژه ماکزیمم می‌باشد به این صورت که یک بردار تصادفی را ضرب در ماتریس می‌کنیم و تقسیم بر نرم تا فقط مؤلفه‌ی مقدار ویژه ماکزیمم باقی بماند و بقیه تا حد امکان تضعیف شوند و این عمل را تکرار می‌کنیم. ۵. DSS هم کاملاً قابل پیاده سازی با این متد است

⟸ DSS هم به مسئله‌ی قطری سازی همزمان دو ماتریس covariance مشابه GEVD قابل بیان است زیرا براساس این متد DSS به بردار ویژه و مقدار ویژه تعمیم یافته می‌رسد