دانشگاه صنعتی شریف

دانشکده مهندسی برق

پروژه درس سیگنال ها و سیستم ها

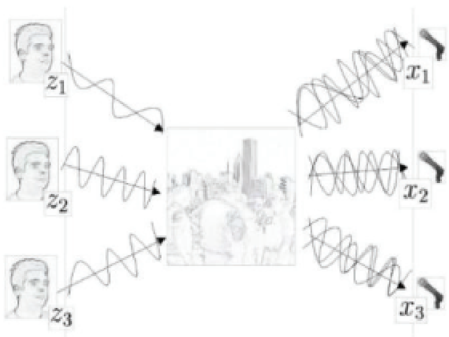
جداسازی کور منبع (منبع صوتی)

استاد درس : دکتر بهروزی

علیرضا رفیعی ساردوئی (۹۷۱۰۱۷۲۳)
مریم قربان صباغ (۹۷۱۱۰۷۱۱)

۱ مقدمه

فرض کنید شما و چند نفر دیگر در یک اتاق در حال صحبت هستید ، به تعداد افراد ضبط صوت در جاهای مختلف اتاق قرار دارد که صدا را ضبط می کند. آیا به وسیله ی این ضبط صوت ها می توان صدا های هر فرد را از بقیه جدا کرد؟ پاسخ این سوال بله است.در این پروژه می خواهیم به بررسی این موضوع بپردازیم.

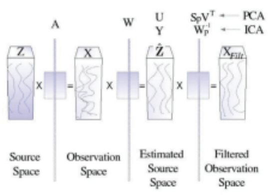


۲ خواسته ها

چند فایل صوتی در اختیار شما قرار گرفته است که صدا های ضبط شده از ۲نفر در حال صحبت هستند. در اولین قدم روی این سیگنالها تحلیل فرکانسی انجام دهید آیا با این روش می توانید که سیگنالهای منابع اصلی را از صداهای ترکیبی بازیابی کنید؟ می توانید در این بخش از تبدیل های مختلف حوزه فرکانس یا زمان-فرکانس استفاده کنید.

بله با استفاده از یک سری تبدیل ها در حوزه فرکانس می توان این کار را انجام داد البته به شرطی که این صوت ها در فضا باز تهیه شده باشند و انعکاس و بازتاب از صوت ها نداشته باشیم و در این حالت می توان با استفاده از مخلوط های کانولوتیو این کار را انجام داد . در موضوع $deep \quad learning$ با استفاده از تبدیل های $stft$ و برخی تحلیل های آماری میتوان صوت ها را از هم جدا کرد .
ناگفته نماند که این مسئله برای فضای یک اتاق هنوز حل نشده و $open$ می باشد.

حال به وسیله ی روش های زیر منابع تلاش کنید منابع صدا را از روی صدای ترکیب شده بازسازی کنید



در هرکدام از روش ها باید ماتریس تبدیلی را پیدا کنید که منابع صدا ها را از ترکیب صدا ها پیدا کنید. هر کدام از روش های زیر الگوریتم های خود را دارند آن ها را بیابید و پیاده سازی کنید

$SVD(PCA)-$

$ICA-$

شما باید درباره ی هرکدام از روش های بالا مطالعه کنید و امکان و روش جداسازی صدا را با استفاده از آنها بررسی کنید درباره ی مزایا و معایب هر روش بحث کنید

۳ روش ICA

ICA مخفف عبارت $Independet\ \ Component\ \ Analysis$ است به معنای ((**آنالیز مولفه های مستقل**)) است . به طو کلی در این روش ما به دنبال یک سری از محور های مستقل هستیم که داده ها صوت در راستای این محور ها توزیع شده اند . در این روش چند پیش فرض برای سیگنال های ترکیب شده داریم که به صورت زیر است :

$1-$ منابع تولید کننده صوت (سیگنال) باید از هم مستقل باشند .

$2-$ تعداد صوت های ضبط شده موجود از ترکیب صوت های منابع باید با تعداد منابع برابر باشد یا از آن ها بیشتر باشد.

$3-$ ماتریس ترکیب کنننده ، A ، که در زیر معرفی شده است باید یک ماتریس مرتبه کامل ($full\ \ rank$) باشد

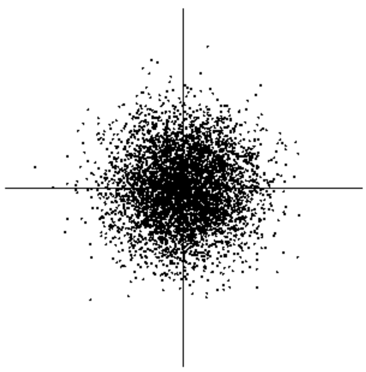
$$X = AS$$

که در آن S صوت های منبع های اصلی است و X صوت های ترکیب شده از منابع اصلی است که در واقع همان صوت هایی است که در اختیار ماست و A ماتریس ترکیب است .

$4-$ بردار های S و X باید میانگین صفر و واریانس واحد داشته باشند .

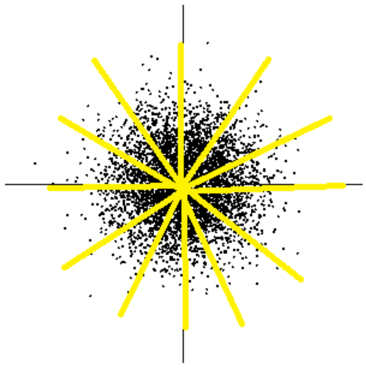
$5-$ حداکثر یکی از سیگنال های می تواند توزیع گوسی داشته باشد .

حال بررسی میکنیم چرا نمیتوانیم بیش از یک سیگنال با توزیع گوسی داشته باشیم : شکل زیر یک توزیع گوسی دو بعدی را نشان می دهد : $p(x_1, x_2) = \frac{2}{\pi}e^{-\frac{x_1^2+x_2^2}{2}}$



در آنالیز ICA ما به دنبال یک سری محور مستقل از هم می باشیم که داده ها در راستای آن ها توزیع شده

اند و با دقت در توزیع گوسی میتوان بیشمار جهت مستقل از هم یافت ، برای مثال شکل زیر را در نظر بگیرید :



همانطور که مشاهده میکنید در این شکل می توان بی شمار جهت های مستقل یافت که داده ها در راستا آن ها توزیع شده باشند . دلیل بی شمار بودن تعداد جهت ها این است که توزیع گوسی کاملاً متقارن است و به این دلیل از ستون های ماتریس A نمی توانیم اطلاعاتی کسب کنیم. پس به این دلیل ما نمیتوانیم سیگنال دارای توزیع گوسی را در این آنالیز به کار ببریم.
از جمله ابهامات در این آنالیز عدم وجود ترتیب است یعنی چون هم ماتریس A و هم S نا معلوم است می توان ترتیب هایی متفاوت به کار برد به طور مثال برای ۴ منبع صوت داریم :

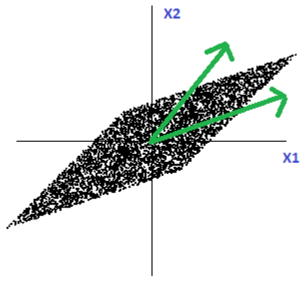
$$\begin{cases} X = a_1s_1 + a_2s_2 + a_3s_3 + a_4s_4 \\ X = a_2s_2 + a_1s_1 + a_3s_3 + a_4s_4 \\ X = a_4s_4 + a_2s_2 + a_1s_1 + a_3s_3 \\ X = a_4s_4 + a_3s_3 + a_2s_2 + a_1s_1 \\ \dots \end{cases}$$

قبل از پرداختن به جزئیات این آنالیز ، مثال هایی از کاربرد این آنالیز را بیان میکنیم :
1− جداسازی کور منبع
2− حذف نویز از تصاویر
3− پردازش سیگنال های پزشکی مثلا EEG, ECG
4− تشخیص چهره
5− فشرده سازی
6− آنالیز سری زمانی
7− $Watermarking$
حال به بررسی آنالیز ICA می پردازیم : در ابتدا باید یک سری پیش پردازش روی بردار X باید انجام دهیم تا بتوانیم از این آنالیز استفاده کنیم ، دلیل این پیش پردازش ارضا کردن شرایط اولیه آنالیز می باشد که قبلا به آن اشاره کردیم . از جمله این شرایط اولیه میانگین صفر و واریانس واحد $(UnitaryFactor)$ می باشد . شرط بعدی مستقل بودن $(Orthogonal\ \ Factor)$ مشاهدات ما از یک دیگر می باشد . با پیش پردازشی به نام سفید سازی $(Whitening)$ این شرایط را روی X ایجاد میکنیم .

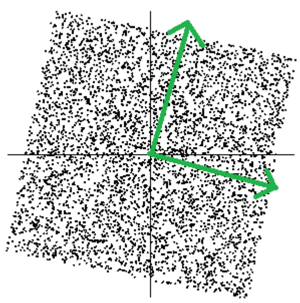
پیش پردازشِ سفید سازی : X صوت های موجود است برای اینکه واریانس آن ها را واحد کنیم باید از یک انتقال خطی استفاده کنیم ، یعنی با ضرب در ماتریس V واریانس بردار حاصل برابر با یک شود .

$$\tilde{X} = V \times X$$

به طور مثال شکل زیر را در نظر بگیرید :



تصویر نشان داده شده سیگنال دریافتی (ضبط شده) ما است. همانطور که مشاهده میکنید دو محور مشخص شده که داده ها در راستا آن ها توزیع شده اند از یک دیگر مستقل نیستند (زاویه بین آن ها قائم نیست.) حال آنکه مطلوب این آنالیز این نیست .با استفاده از ماتریسی به نام V میتوانیم این مشکل را حل کنیم اگر ماتریس مناسب V را پیدا کرده و حاصل ضرب آن در ماتریس X به شکل زیر در بیاید کار تمام است .



ماتریس V را میتوان به صورت زیر یافت :

$$V = D^{-\frac{1}{2}} E^T$$

* T نشان گر عمل ترانهاده کردن ماتریس است . $(Transpose)$
ماتریس های D, E از ماتریس کواریانس X که به صورت زیر است بدست می آیند .

$$C_X = E\{XX^T\}$$

ماتریس D در واقع یک ماتریس قطری است که عناصر روی قطر آن مقادیر ویژه ماتریس C_X می باشد و ماتریس E ماتریس بردار ویژه ماتریس C_X است که عناصر آن متناظر با عناصر ماتریس D می باشد . حال ثابت میکنیم با ضرب ماتریس V در ماتریس X واریانس آن برابر با واحد خواهد شد :

$$\tilde{X} = VX = D^{-\frac{1}{2}}E^TX = \tilde{A}S \implies E\{\tilde{X}\tilde{X}^T\} = \tilde{A}E\{SS^T\}\tilde{A}^T = \tilde{A}\tilde{A}^T = I$$

ماتریس $\tilde{A}$ در واقع همان نسخه سفید شده از ماتریس A است .
در آخر برای اینکه پیش پردازش را به اتمام برسانیم باید میانگین $\tilde{X}$ را صفر کنیم که برای اینکار کافی است میانگین را از تک تک اعضا کم کنیم .
پیش پردازش به اتمام رسید.
البته ما در اینجا از یکی روش های پیش پردازش استفاده کردیم روش های دیگری نیز وجود دارد که به صورت زیر است :

1- PCA
2- کاهش مدل
3- سفید سازی (روش استفاده شده)
4- فیلترینگ
5- ویولنت

پردازش و تخمین منبع : در صورت پروژه گفته شده بود که سیگنال های در دسترس ما یک ترکیب خطی از یکی سری منابع مستقل می باشد و ما آن را به صورت

$$X = AS$$

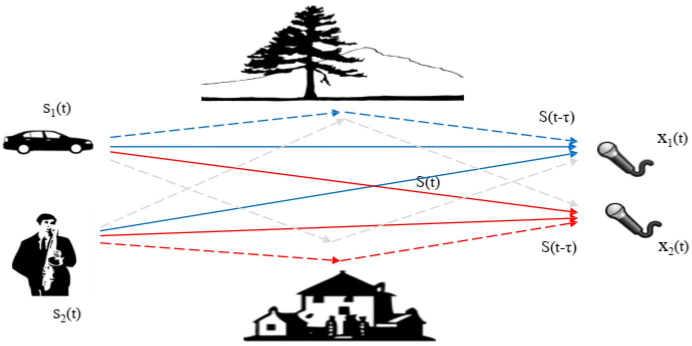
نوشتیم و با پیش پردازش سفید سازی به صورت

$$\tilde{X} = \tilde{A}S$$

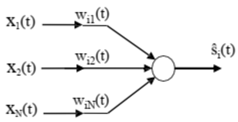
در آمد حال اگر یک ماتریس مانند W را بتوانیم پیدا کنیم که به صورت زیر باشد

$$Y = W^T\tilde{X} \implies Y \simeq S$$

آنگاه سیگنال های منبع را به راحتی میتوانیم بدست آوریم اما مشکل این جاست که ما ماتریس $\tilde{A}$ را نداریم برای همین باید به دنبال راهکاری بهتر باشیم . راهکاری که در این روش استفاده میکنیم این است که ماتریس دلخواه برای ماتریس W در نظر میگیریم و سعی میکنیم که این ماتریس را بهتر و بهتر کنیم تا تقریب خوبی از سیگنال های منبع را بدست آوریم . معیاری که در اینجا برای بهبود ماتریس W استفاده میکنیم این است که ما در هر مرحله قصد داریم که استقلال سیگنال های منبع را افزایش دهیم و برای همین ماتریس W را به گونه ای تغییر میدهیم که حداکثر استقلال منابع تخمینی را بدست آوریم . البته معیاری دیگری نیز استفاده میکنیم به این صورت که میزان شباهت سیگنال را با تاخیر یافته همان سیگنال مقایسه میکنیم هرچه مقدار این شباهت بیشتر باشد سیگنال تخمینی ما بهتر است ، دلیل این کار این است که در حین ضبط کردن این سیگنال ها ، بازخوردهایی از این سیگنال ها با تاخیر τ دریافت میشود . شکل زیر میتواند در شفاف سازی این موضوع مؤثر باشد .



حال با توجه به توضیحات بالا ماتریس W مانند یک فیلتر عمل خواهد کرد که با عبور سیگنال های ضبط شده از آن سیگنال های منبع های مستقل را میتوان بدست آورد که مقدار شباهت و همبستگی بین سیگنال های منبع و تاخیر یافته های آن حداالامکان بالا باشد . تصویر زیر نشان دهنده ساده این فیلتر است .



با توجه به تعریف فیلتری بالا ، برای تاخیر زمانی داریم :

$$S_i(t\hat{-}\tau) = W_i^T X(t-\tau)$$

حال برای بررسی همبستگی بین سیگنال منبع و تاخیر یافته آن ، یک تاخیر مناسب را در نظر گرفته و همبستگی را چک میکنیم همبستگی را هم به صورت خطی و هم به صورت غیر خطی میتوان در نظر گرفت .
همبستگی خطی :

$$max \quad \psi(w_i) = E\{\tilde{S}_i(t)\tilde{S}_i(t-\tau)\} = E\{(w_i^T\tilde{X}_i(t))(w_i^T\tilde{X}_i(t-\tau))\}$$

همبستگی غیر خطی :

$$max \quad \psi(w_i) = E\{\tilde{S}_i(t)\tilde{S}_i(t-\tau)\} = E\{G(w_i^T\tilde{X}_i(t))G(w_i^T\tilde{X}_i(t-\tau))\}$$

که معمولا $G(x)$ را به صورت x^2 و یا $\log(\cosh(x))$ انتخاب میکنند و ما در کد مطلب برای راحتی از $G(x) = x^2$ استفاده میکنیم . دلیل انتخاب همبستگی غیر خطی این بود که در همبستگی غیرخطی تاثیرات نویز کمتر در نظر گرفته خواهد شد .
حال برای بهینه سازی مولفه های ماتریس W میبایست از روش های بهینه سازی استفاده کرد دو روش بهینه سازی که در زیر آمده است روش های مؤثری برای این کار می باشند :
1- روش LMS

2− روش تکرار نیوتون

روش $LMS(Least\ \ Mean\ \ Square)$

ساختار این روش به صورت زیر می باشد :

$$w_i(k+1) = w_i(k) - \mu\frac{\partial\psi}{\partial w_i}, w_i = \frac{w_i}{||w_i||}$$

که مقدار $\frac{\partial\psi}{\partial w_i}$ با توجه به رابطه ای که برای ψ بدست آوردیم به صورت زیر است :

$$\frac{\partial\psi}{\partial w_i} = E\{G'(\tilde{S}_i(t))G(\tilde{S}_i(t-\tau))\tilde{X}(t) + G(\tilde{S}_i(t))G'(\tilde{S}_i(t-\tau))\tilde{X}(t-\tau)\}$$

روش تکرار نیوتون ساختار این روش به صورت زیر می باشد :

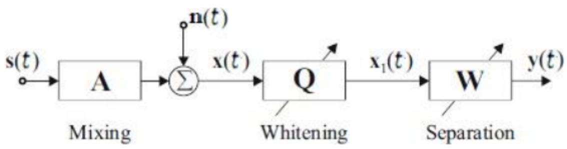
$$w_i(k+1) = w_i(k) - (\frac{\partial^2\psi}{\partial w_i^2})^{-1}(\frac{\partial\psi}{\partial w_i}), w_i = \frac{w_i}{||w_i||}$$

به دلیل پیچیده تر بودن این روش ، ما از روش اول استفاده میکنیم .

حال با تکرار روش LMS می توانیم مؤلفه های ماتریس W را تا دقت خوبی پیدا کنیم . و با داشتن این ماتریس میتوانیم تقریب منابع را بدست آوریم .

مرور :

کار هایی که در این روش انجام دادیم را میتوانیم به صورت زیر خلاصه کنیم .



1− جمع آوری سیگنال های ضبط شده

2− انجام عمل سفید سازی

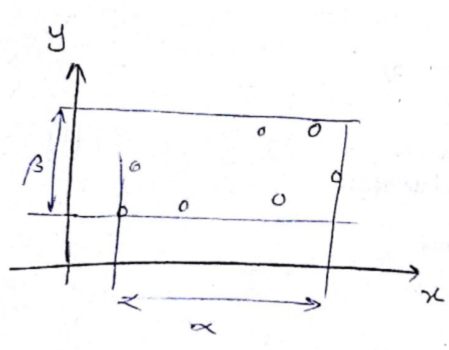
3− مقداردهی اولیه ضرایب فیلتر W

4− اجرای الگوریتم تکراری بهینه سازی

5− محاسبه سیگنال های منبع

۴ روش PCA

در ابتدا به توضیح کلی روش PCA که مخفف عبارت $((Principal\ Component\ Analysis))$ می باشد میپردازیم . معنای فارسی نام این روش " تحلیل مؤلفه اصلی " می باشد .
روش PCA : در واقع روش تحلیل مؤلفه اصلی یک تبدیل خطی است که داده های موجود را از فضای $n-$ دیمانسیون به فضای دیگری که با تعداد برابر دیمانسیون تبدیل میکند و این موضوع را نیز در نظر دارد که داده های اصلی (تبدیل نشده) دارای چند وابستگی در فضا چه از نظر متغیر و چه از نظر دیمانسیون و چه از نظر ویژگی می باشد و نتیجه ای که ما از این تبدیل میگیریم (داده های جدید) در فضای هدف دارای این ویژگی می باشد که بین داده ها وابستگی وجود ندارد و محور ها داده های جدید بر هم عمود خواهند بود
از جمله کاربرد های این روش میتوان به موضوع های زیر اشاره کرد :
$1-$ کاهش دیمانسیون
$2-$ استخراج ویژگی
توضیح روش اگر از ما بپرسند که کدام یک از متغیر های در شکل زیر را انتخاب می کنید روش هوشمندانه این است که متغیری که واریانس بیشتری دارد را انتخاب کنیم . در این مثال ما متغیر x را انتخاب میکنیم زیرا $(\alpha > \beta)$



اما اگر اجازه داشته باشیم که بیش از یک متغیر را انتخاب کنیم ، مثلا ترکیب خطی از x و y مثل z که z یک متغیر جدید است و $z = c_1x + c_2y$ ، آنگاه به ازای چه مقادیری برای c_1 و c_2 ، بیشترین واریانس ممکن را برای متغیر z خواهیم داشت ؟! مسئله را به حالت کلی تر تعمیم میدهیم . پس داریم :

$$z = u_1x_1 + u_2x_2 + \cdots + u_Dx_D$$

اگر بخواهیم به شکل ماتریسی بیان کنیم :

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_D \end{bmatrix}, u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_D \end{bmatrix}, Z = u.x = u^T x = x^T u$$

حال سوال پیدا کردن u به نحوی است که واریانس z ماکزیمم شود .

مسئله را به فضای R^D میبریم .

$$X = \{\overrightarrow{x_1}, \quad \overrightarrow{x_2}, \quad \ldots, \quad \overrightarrow{x_N}\}, \overrightarrow{x_i} = \begin{bmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{iD} \end{bmatrix} \in R^D$$

که $\overrightarrow{x_i}$ سمپل هایی هستند که به صورت بردار در فضایی با بعد D داده شده اند .

$$\overline{X} = \frac{1}{N}\Sigma_{i=1}^{N}\overrightarrow{x_i}$$

$$COV(X) = \frac{1}{N}\Sigma_{i=1}^{N}x_i.x_i^T = C$$

ماتریس C ماتریس کواریانس سمپل ها می باشد که یک ماتریس با سایز $D \times D$ است که همه مقادیر ویژه آن غیرمنفی هستند.

$$Var(Z) = \frac{1}{N}\Sigma_{i=1}^{N}(Z_i - \overline{Z})$$

که Z_i یک اسکالر است و مقدار آن برابر است با :

$$Z_i = u^T X_i$$

$$\overline{Z} = \frac{1}{N}\Sigma_{i=1}^{N}Z_i = \frac{1}{N}\Sigma_{i=1}^{N}u^T X_i = u^T(\frac{1}{N}\Sigma_{i=1}^{N}X_i) = u^T\overline{X}$$

$$Var(Z) = \frac{1}{N}\Sigma_{i=1}^{N}(u^T X_i - u^T\overline{X})(X_i^T u - \overline{X}^T u) = \frac{1}{N}\Sigma_{i=1}^{N}u^T(X_i - \overline{X})(X_i - \overline{X})^T u =$$

$$u^T[\frac{1}{N}\Sigma_{i=1}^{N}(X_i - \overline{X})(X_i - \overline{X})^T]u = u^T COV(X)u = u^T C u \implies Var(Z) = u^T C u$$

حالا میخواهیم این عبارت را ماکزیمم کنیم . میتوان u را بی نهایت در نظر گرفت پس باید محدودیتی روی u در نظر بگیریم پس در نظر میگیریم.

$$uu^T = 1 \;\; or \;\; ||u|| = 1$$

$$\begin{cases} max \;\; u^T C u \\ subject \; to \;\; : u^T u = 1 \end{cases} \implies J = u^T C u - \lambda(u^T u - 1) \implies \frac{\mathrm{d}J}{\mathrm{d}u} = 0$$

$$\implies Cu - \lambda u = 0 \implies Cu = \lambda u$$

در عبارت آخر به تعریف بردار ویژه و مقدار ویژه رسیدیم. بنابراین پاسخ سؤال ما برای u که واریانس Z را ماکزیمم کند ، یکی از بردار های ویژه ماتریس کواریانس نمونه هاست ولی اکنون این سوال پیش می آید : کدام یک از بردار های ویژه ؟!

$$Var(Z) = u^T C u = u^T \lambda u = \lambda u^T u = \lambda$$

بنابراین مقدار واریانس Z با مقدار ویژه ماتریس کواریانس برابر است و بنابراین برای رسیدن به مقدار ماکزیمم باید بردار ویژه ی متناظر با بزرگ ترین مقدار ویژه را به عنوان u در نظر بگیریم .

$$\lambda = \lambda_{max} \implies \boxed{u} \checkmark$$

رابطه $Z = u^T X$ یک تابع تبدیل خطی است که فضای D بعدی نمونه ها را به فضای یک بعدی متناظر میکند .

$$X \in R^D \implies Z \in R^1$$

در حالت کلی تر ، $Z = U^T X$ که فضای D بعدی نمونه ها را به فضای m بعدی متناظر میکند و U یک ماتریس است .

$$Z_{m\times1}, U_{D\times m} \implies U^T_{m\times D}, X_{D\times1}$$

$$U = \{\overrightarrow{u_1},\ \ \overrightarrow{u_2},\ \ \ldots,\ \ \overrightarrow{u_m}\} \in R^{D\times m}, \overrightarrow{u_j} \in R^D$$

پاسخ مسئله این است که $\overrightarrow{u_1}$ تا $\overrightarrow{u_m}$ همگی بردار ویژه ی ماتریس کواریانس (C) هستند که به ترتیب نزولی از بردار ویژه های متناظر با مقدار ویژه های بیشتر تا بردار ویژه های متناظر با مقدار ویژه های کمتر ، مرتب شده اند این U میتواند واریانس Z را ماکزیمم کند.
به $\overrightarrow{u_1}$ تا $\overrightarrow{u_m}$ ، $Principal\ \ Component\ \ of\ \ Z$ می گویند .

$$C\begin{cases}\lambda_1 > \lambda_2 > \lambda_3 > \cdots > \lambda_D\\ u_1 \quad u_2 \quad u_3 \quad \ldots \quad u_D\end{cases}$$

برای مثال در حالت مقصد دو بعدی داریم :

$$X \longrightarrow R^2, \overrightarrow{Z} = U^T\overrightarrow{X}, \overrightarrow{Z} \in R^2, \overrightarrow{X} \in R^5$$

$$X \implies COV(X) = C \implies \begin{cases}\lambda_1, u_1\\ \lambda_2, u_2\end{cases} \implies U = [u_1, u_2]$$

با انتخاب U به روشی که قبلا توضیح داده شده ، ما مطمئن می شویم که دو مؤلفه از Z بیشترین واریانس ممکن را از بین دیگر ترکیب های ممکن برای ساخت Z دارند و این ایده اصلی PCA است .

۵ استفاده از روش $SVD(PCA)$ در حل مسئله

در حل مسئله $cocktail\ \ party$ فرض ما بر این بود که صوت های بدست آمده در واقع یک ترکیب خطی از صوت های منبع می باشد یعنی به صورت :

$$X = AS$$

می باشد که در آن ماتریس A ماتریس ترکیب کننده و S نشان گر منابع است . در این مسئله هر دو A و S مجهول هستند اگر ماتریس A را طبق SVD به صورت

$$A = USV^T$$

تعریف کنیم . طبق SVD ماتریس های U, V ماتریس های چرخشی ومتعامد خواهند بود . در واقع SVD که مخفف عبارت $Singular\ \ Value\ \ Decomposition$ است . یک راه برای تجزیه ماتریس A به شکل $U\Sigma V^T$ می باشد که در آن ماتریس های U, V ماتریس های چرخشی و متعامد هستند و ماتریس Σ یک ماتریس قطری . حال اگر ما بتوانیم از طریقی ماتریس وارون را حساب کنیم کار تمام شده است زیرا میتوان ماتریس منابع را به صورت زیر بدست آورد :

$$Y = A^{-1} \implies S = YX$$

حال با فرض $A = U\Sigma V^T$ اگر ماتریس کواریانس X را تشکیل دهیم داریم :

$$E[XX^T] = E[(AS)(AS)^T] = E[(U\Sigma V^T S)(U\Sigma V^T S)^T] =$$

$$U\Sigma V^T E[SS^T] V\Sigma U^T$$

جون منابع مستقل هستند(البته فرض هم کردیم که روی منابع عملیات سفید سازی انجام شده است) پس داریم :

$$E[XX^T] = U\Sigma V^T V\Sigma U^T$$

و چون V یک ماتریس متعامد($VV^T = I$) است پس :

$$E[XX^T] = U\Sigma^2 U^T$$

طبق جبر خطی تنها صورتی که یک ماتریس را می توان به صورت $K = UGU^T$ نوشت ، این است که ماتریس U ماتریس بردار ویژه های ماتریس K باشد و ماتریس G یک ماتریس قطری باشد که المان های روی قطر آن برابر با مقدار ویژه های ماتریس K باشد .(المان های روی قطر ماتریس G باید متناظر با بردارد ویژه های ماتریس U باشد .)
تاکنون ماتریس Σ را که برابر با $G^{\frac{1}{2}}$ است یافتیم و ماتریس U را نیز که برابر با بردار ویژه های ماتریس کواریانس X است پیدا کردیم حال تنها ماتریس V مانده است . یافتن ماتریس V اندکی از دیگر ماتریس ها سخت تر می باشد . فرض کنید روی ماتریس X عملیات سفید سازی انجام شده است برای یافتن ماتریس V باید مراحل زیر بر روی ماتریس X پیاده شود .

$1-$ ابتدا عناصر ستونی ماتریس X (ماتریس X باید سفید باشد .)را به توان ۲ رسانده و آن ها را باهم جمع کرده و در یک سطر می گذاریم ، مثلا فرض کنید ماتریس X به صورت

$$X = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$$

باشد در پایان این مرحله به صورت

$$X_1 = \begin{bmatrix} 35 & 56 \end{bmatrix}$$

در می آید .

$2-$ ماتریس X_1 بدست آمده را به صورت سطری به تعداد سطر های ماتریس X تکرار کرده و ماتریس جدید را تشکیل میدهیم . در این مرحله ماتریس X به صورت

$$X_2 = \begin{bmatrix} 35 & 56 \\ 35 & 56 \\ 35 & 56 \end{bmatrix}$$

در می آید .

$3-$ ماتریس X_2 بدست آمده را با ماتریس X ضرب نقطه ای میکنیم. در این مرحله نتیجه به صورت زیر بدست می آید :

$$X_3 = X_2.X = \begin{bmatrix} 35 & 56 \\ 35 & 56 \\ 35 & 56 \end{bmatrix} . \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} = \begin{bmatrix} 35 & 112 \\ 105 & 224 \\ 175 & 336 \end{bmatrix}$$

$4-$ حال ماتریس بدست آمده را در ماتریس X^T ضرب می کنیم.

$$X_4 = X_3 \times X^T = \begin{bmatrix} 35 & 112 \\ 105 & 224 \\ 175 & 336 \end{bmatrix} \times \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix} = \begin{bmatrix} 259 & 553 & 847 \\ 553 & 1211 & 1869 \\ 847 & 1869 & 2891 \end{bmatrix}$$

$5-$ حال اگر ماتریس X_4 را با استفاده از SVD تجزیه کنیم و به صورت $X_4 = PQR^T$ بنویسیم ، ماتریس P همان ماتریس V خواهد بود . در واقع می توان ثابت کرد که ماتریس X_4 یک تخمین از طیف انرژی X است و با تجزیه آن می توان ماتریس V را یافت . حال که ماتریس V را نیز یافتیم داریم :

$$X = AS \implies X = U\Sigma V^T S \implies S = V\Sigma^{-1}U^{-1}X = V\Sigma^{-1}U^T X$$

پس ماتریس منابع را توانستیم بدست بیاریم و توانستیم مسئله را حل کنیم .

۶ مقایسه دو روش PCA و ICA

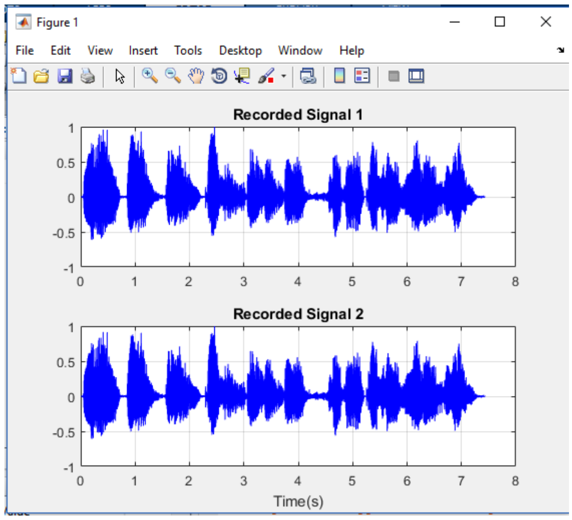
$\Leftarrow$ در هر دو روش داده ها را به دستگاه معادلی میبریم که بتوان منابع را بهتر از هم تفکیک کرد .
$\leftarrow$ در روش PCA : فرض میکنیم که اطلاعات حول بردار های ویژه ماتریس کواریانس متمرکز است و با این فرض از بین بردار های ویژه ، آن هایی را که بیشترین انرژی رو به خودشون اختصاص دادند را انتخاب میکنیم و کل داده ها را به زیر فضای تشکیل شده توسط این بردار ها منتقل میکنیم بدین ترتیب بُعد مسئله را نیز به این تعداد بعد کاهش دادیم .
$\leftarrow$ در حالی که در روش ICA ما به دنبال پیدا کردن یک نگاشت به فضای جدیدی هستیم که در آن فضا ، تا جای ممکن منابع از هم مستقل خطی باشند .
$\Leftarrow$ روش PCA برای جداسازی نویز با توزیع گوسی مناسب است .
$\Leftarrow$ روش ICA برای جداسازی نویز با توزیع غیر گوسی مناسب است .
$\Leftarrow$ در روش PCA : به طور صریح مولفه های انتخابی ما مولفه هایی با بیشترین انرژی هستند .
$\Leftarrow$ در روش ICA : منابع را با استفاده از صوت هایی که با ترکیب و $scale$ دلخواه از منابع تولید شده اند ، می توان استخراج کرد . در این فرآیند ما به مؤلفه های " استقلال - انرژی " برای پیدا کردن منابع نیاز داریم .
$\Leftarrow$ در روش ICA فرض ما مبنی بر خطی بودن ماتریس ترکیب است .
$\Leftarrow$ در روش ICA فرض می شود ماتریس ترکیب تغییر ناپذیر با زمان است .
$\Leftarrow$ در روش ICA فرآیند پیدا کردن صوت ها به صورت یک تابع از μ است .
$\Leftarrow$ در روش ICA ما باید به تعداد منابع ، صوت در اختیار داشته باشیم .
$\Leftarrow$ هر دو روش ، دو روش آماری هستند .
$\Leftarrow$ در روش PCA ما از آماره های مرتبه دو استفاده میکنیم .
$\Leftarrow$ در روش ICA ما از آماره های مرتبه بالاتر استفاده میکنیم .
$\Leftarrow$ هر دو روش در ضمینه های زیر به کار می روند :
$\leftarrow$ جدا سازی کور منبع
$\leftarrow$ استخراج ویژگی
$\leftarrow$ علم عصب شناسی
$\Leftarrow$ در روش PCA برخی مؤلفه ها مهم تر هستند (مؤلفه هایِ با انرژی بیشتر انتخاب می شوند)
$\Leftarrow$ در روش ICA تمامی مؤلفه ها مهم هستند .

۷ مثال ۱

به صوت های زیر که ترکیبی از دو منبع صوت هستند گوش کنید !

$mix1$ ، $mix2$

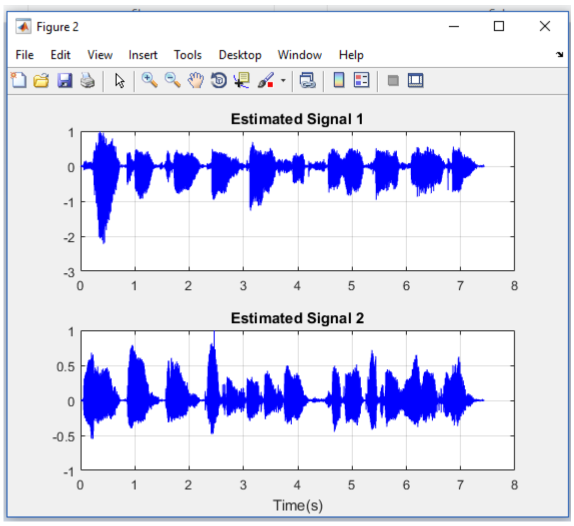
روش ICA در تصویر زیر ، تصویر این دو صوت را مشاهده میکنید



حال اگر با استفاده از این روش این صوت ها را از هم جدا کنیم ، صوت های زیر بدست می آیند :

$source1$ ، $source2$

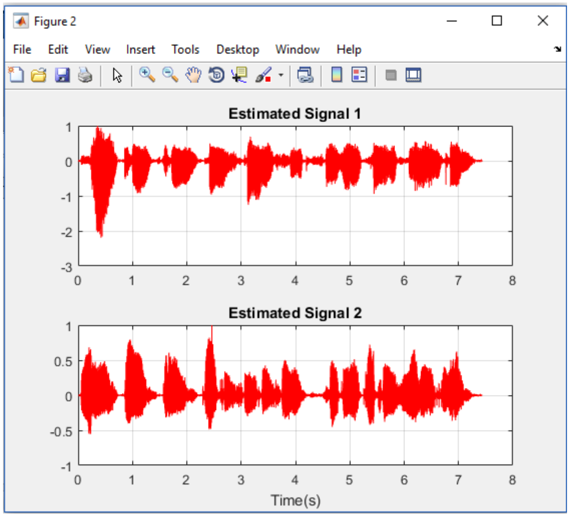
و تصاویر صوت های جدا شده به صورت زیر است :



حال اگر با استفاده از روش SVD این صوت ها را جدا کنیم صوت های زیر بدست می آیند :

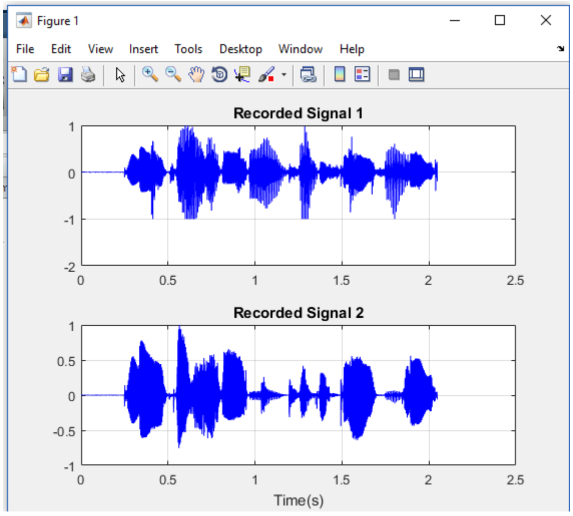
$source1$ ، $source2$

و تصاویر صوت های جدا شده به صورت زیر است :



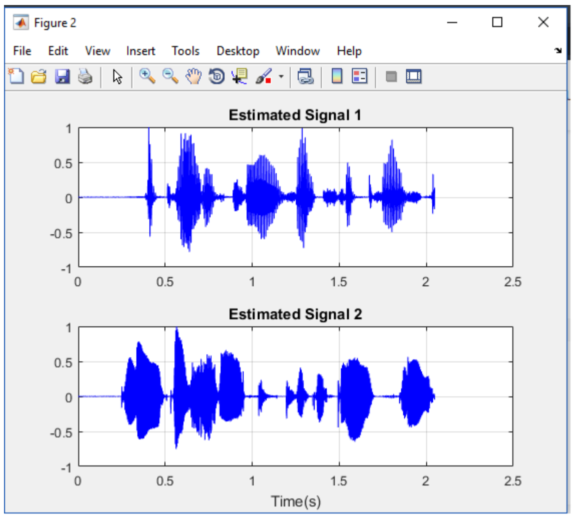
۸ مثال ۲

به صوت های زیر که ترکیبی از دو منبع صوت هستند گوش کنید !
$mix1$ ، $mix2$
روش ICA در تصویر زیر ، تصویر این دو صوت را مشاهده میکنید



حال اگر با استفاده از این روش این صوت ها را از هم جدا کنیم ، صوت های زیر بدست می آیند :

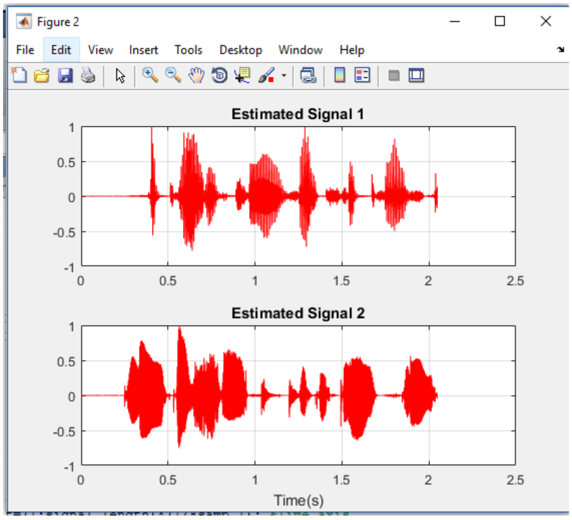
$source1$ ، $source2$
و تصاویر صوت های جدا شده به صورت زیر است :



حال اگر با استفاده از روش SVD این صوت ها را جدا کنیم صوت های زیر بدست می آیند :

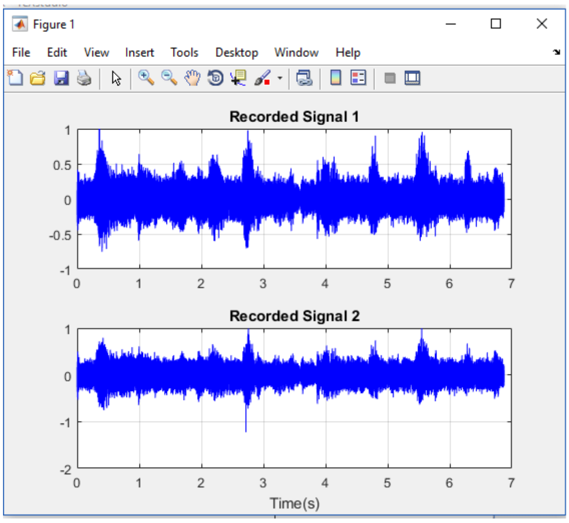
$source1$ ، $source2$

و تصاویر صوت های جدا شده به صورت زیر است :



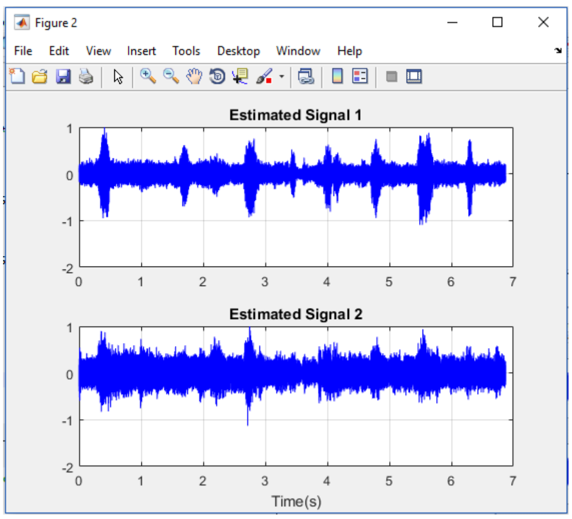
۹ مثال ۳

به صوت های زیر که ترکیبی از دو منبع صوت هستند گوش کنید !
$mix1$ ، $mix2$
روش ICA در تصویر زیر ، تصویر این دو صوت را مشاهده میکنید



حال اگر با استفاده از این روش این صوت ها را از هم جدا کنیم ، صوت های زیر بدست می آیند :

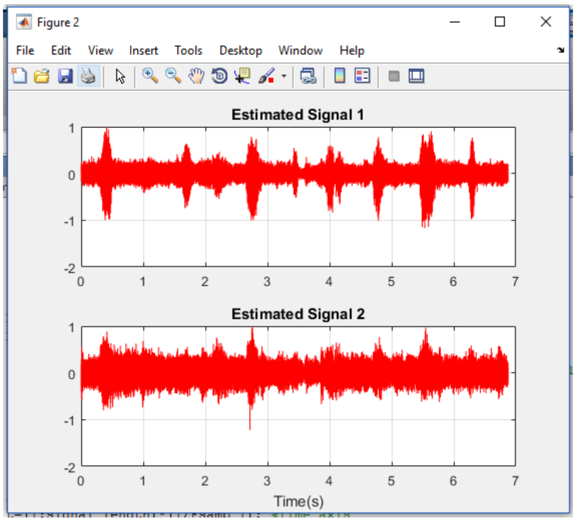
$source1$ ، $source2$
و تصاویر صوت های جدا شده به صورت زیر است :



حال اگر با استفاده از روش SVD این صوت ها را جدا کنیم صوت های زیر بدست می آیند :

$source1$ ، $source2$

و تصاویر صوت های جدا شده به صورت زیر است :



همانطور که مشاهده کردید متاسفانه هر دو برنامه نتوانستند صوت های اصلی را به شکل مطلوب بازسازی کنند ،این اتفاق می تواند چند دلیل داشته باشد :

$1-$ صوت ها دارای انعکاس و بازتاب (در هنگام ضبط شدن) هستند .

$2-$ ممکن است انعکاس نداشته باشیم اما ماتریس A ترکیب کننده دارای دترمینان صفر باشد و وارون پذیر نباشد .

$3-$ ممکن است درایه های ماتریس A ثابت نباشندو با زمان تغییر کنند.

در هر کدام از حالات بالا نمیتوانیم صوت های اصلی را با این برنامه ها بدست بیاوریم.

۱۰ ضمیمه (روش تجزیه SVD)

در این ضمیمه به توضیح تجزیه SVD می پردازیم . در واقع SVD یک ماتریس X را به شکل زیر تجزیه میکند .

$$X = USV^T, X_{M\times N}, U_{M\times M}, S_{M\times N}, V_{N\times N}$$

که در آن ماتریس U و V متعامد هستند و ماتریس S یک ماتریس قطری است . برای بدست آوردن مقادیر روی قطر ماتریس S کافی است ماتریس X^TX را تشکیل داده و مقادیر ویژه این ماتریس برابر خواهد بود با مربع مؤلفه های روی قطر S و عناصر ستونی ماتریس V برابر با بردار ویژه های ماتریس X^TX خواهد بود . البته مؤلفه های روی قطر ماتریس S با عناصر ستونی ماتریس V متناظر است. حال که ماتریس های و S و V را یافتیم . ماتریس U به صورت زیر میابیم :
ماتریس U یک ماتریس با عناصر ستونی متعامد است و اگر فرض کنیم ماتریس X یک ماتریس $M \times N$ باشد ، N مؤلفه اول ماتریس U به صورت زیر محاسبه می شود :

$$U : u_i = (S^{-1})_i X v_i, i = 1 : N$$

و بقیه ستون های آن طبق روش یافتن محور عمود $Gram - Schmidt$ پیدا می شوند .