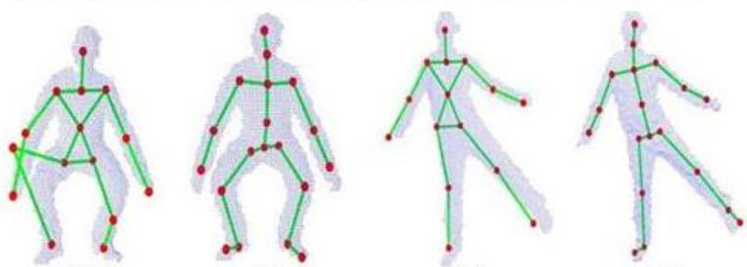
بسمه تعالی

دانشگاه صنعتی شریف

دانشکده مهندسی برق

گروه سیستم‌های دیجیتال



آزمایشگاه یادگیری و بینایی ماشین

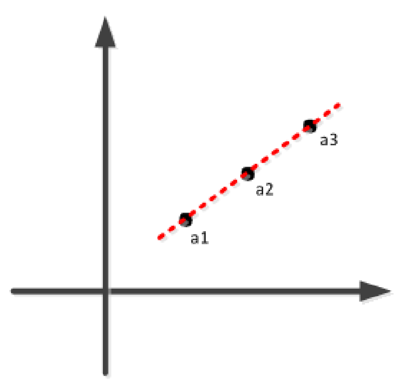
دستور کار آزمایش اول: تحلیل مؤلفه های اساسی

زمان لازم برای انجام آزمایش: حداکثر دو جلسه

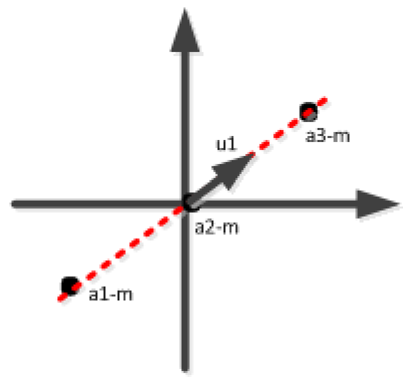
قسمت اول: کاهش بُعد

وقتی که داده دارای ابعاد بزرگ است، بهتر است ابعاد آن را برای انتقال یا مسائل دیگر، کاهش دهیم. با این کار، قطعا بخشی از اطلاعات حذف می‌شود. ولی ما قصد داریم که با حذف بخش بزرگی از ابعاد داده، فقط حدود ۱٪ از اطلاعات آن دور ریخته شود. شاید فکر کنید این انتظار خیلی بزرگی است؛ ولی بدانید که شدنی است! برای این که تصوری از یک داده با ابعاد بزرگ داشته باشید، یک تصویر RGB با ابعاد 800*400 در نظر بگیرید. کل ابعاد این تصویر بخاطر وجود ۳ کانال، برابر با 800*400*3 می شود که برابر با 960000 است. همان‌طور که می‌بینید، دارای ابعاد بسیار بزرگی است!

برای توضیح روش کاهش بعد، ابتدا برای سادگی فرض کنید که داده فقط دو بعد دارد. نقاطی نمونه از این داده را در نمودار روبه‌رو مشاهده می‌فرمایید. فرض کنید که داده‌ها روی یک زیرفضای یک بعدی (یک خط) قرار دارند.



در ابتدا، ما نقاط را از میانگین آنها کم می‌کنیم تا میانگین‌شان صفر شود.



اکنون هر نقطه را به جای این که با دو عدد (دو بعد) نمایش دهیم، می خواهیم با یک عدد (یک بعد) نمایش دهیم:

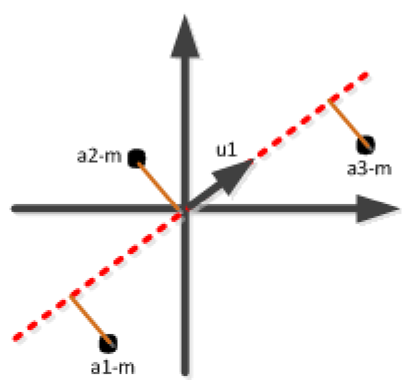
$$\alpha_1 = (a_1 - m)^T u_1$$

$$\alpha_2 = (a_2 - m)^T u_1$$

$$\alpha_3 = (a_3 - m)^T u_1$$

همان‌طور که می‌بینید، داده‌های بالا را از دو بعد به یک بعد کاهش دادیم. به عبارتی دیگر، آنها را روی بهترین زیرفضای ممکن برای نمایش آنها تصویر کردیم. به همین خاطر بیشتر اطلاعات آنها در تصویر شده آنها باقی ماند.

حال، فرض کنید که این ۳ نقطه روی یک خط نباشند. پس از کم شدن از میانگین، داریم:



خط زیرفضای یافت شده همان خطی است که کمترین فاصله را از داده‌ها دارد و در جهت پراکندگی داده‌ها است.

در این صورت، واریانس داده‌ها به صورت زیر می‌شود:

$$var = \sum_{i=1}^{3}[(a_i - m)^T u_1]^2 =$$

$$= u_1^T[(a_1 - m)(a_1 - m)^T + (a_2 - m)(a_2 - m)^T + (a_3 - m)(a_3 - m)^T]u_1 =$$

$$= u_1^T C u_1$$

در رابطه بالا، ماتریس C همان ماتریس کواریانس است. اگر مساله بالا را با شرط این که اندازه بردار یکه u1 برابر با یک است، به صورت مساله بهینه‌سازی حل کنیم، خواهیم داشت:

$$Cu = \lambda u$$

(امتیازی): رابطه بالا را اثبات نمایید. (راهنمایی: مساله بهینه‌سازی بالا یک مساله ضرب‌کننده لاگرانژ است؛ به این صورت که شما می‌خواهید واریانس را که در بالا بدست آمد، حداکثر کنید تا زیرفضا در جهت حداکثر پراکندگی قرار گیرد و در عین حال، اندازه $(u_1^T u_1 - 1)$ حداقل شود تا اندازه u1 که برابر با $u_1^T u_1$ است، برابر با یک شود.)

در رابطه بالا، λ ها همان مقادیر ویژه ماتریس کواریانس، و u هم بردارهای ویژه ماتریس کواریانس هستند. در مثال ما که دو بعد داریم، دو مقدار و بردار ویژه موجود است که مقدار ویژه‌ی متناظر با $u1$ از دیگری، به مراتب بزرگتر است؛ به طوریکه اگر از بین K (در این مثال، ۲) تا بعد، K' (در این مثال، ۱) تا مقدار ویژه را فقط داشته باشیم، داریم:

$$\frac{\sum_{j=1}^{K'} \lambda_j}{\sum_{j=1}^{K} \lambda_j} = \frac{\lambda_1}{\lambda_1 + \lambda_2} = 0.99$$

به مقادیر ویژه استخراج شده، اصطلاحا Principle Components گفته می‌شود. و به برداشتن مقادیر ویژه برتر و دور ریختن بقیه مقادیر ویژه Principle Component Analysis (PCA) گفته می‌شود.

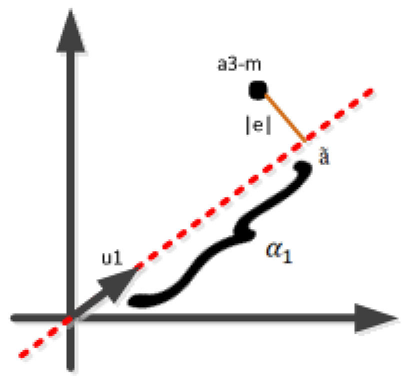
فرض کنید دوباره بخواهیم داده‌های دوبعدی را بازسازی (reconstruct) کنیم. مثلا برای بازسازی a1 داریم:

$$\alpha_1 u_1 + m = \tilde{a}_1$$

دقت نمایید که اگر K' بعد را نگاه می‌داشتیم، بازسازی به صورت زیر می‌شد:

$$\tilde{a}_1 = (\sum_{i=1}^{K'} \alpha_i u_i) + m$$

اکنون، به خطای بازسازی می‌پردازیم. فرض کنید که نمونه داده a را کاهش بعد داده و سپس بازسازی کرده‌ایم که ã است.



خطای بازسازی همان مقدار |e| است که در شکل بالا هم دیده می‌شود و به صورت زیر بیان می‌شود.

$$|e| = |a_1 - \tilde{a}|$$

انتظار می‌رود که این خطا خیلی کم (مثلا در حد ۱٪) باشد.

قسمت دوم: Eigenface

در قسمت قبل از این آزمایش، با مفهوم PCA و مقادیر ویژه‌ی برتر داده آماری آشنا شدید. در این آزمایش با استفاده از این مفهوم در کار با چهره روبه‌رو می‌شوید.

در هر روش تشخیص و تحلیل چهره‌ای تعدادی تصویر برای آموزش سیستم و تعدادی تصویر برای تست داریم. ممکن است که تصاویر تست از همان افراد تصاویر آموزش باشد و ممکن است هم نباشد که این مساله را سخت‌تر می‌کند. فرض کنید که تعدادی تصویر به صورت زیر برای آموزش (تشکیل فضای آماری) در اختیار داریم:

هر یک از این تصاویر دارای ابعاد مثلا N1*N2 است که اگر هر یک از آنها را به صورت یک بردار ستونی reshape کنیم، هر تصویر به صورت یک بردار 1*(N1*N2) در می‌آید.

حال اگر بین تصاویر آموزش میانگین بگیریم، تصویر شبح‌مانند m بدست می‌آید که در زیر مشاهده می‌شود:

حال، اگر هر تصویر آموزش Γ_i را از این تصویر میانگین کم کنیم، بردار $\varphi_i = \Gamma_i - m$ برای هر تصویر آموزش بدست می‌آید. فرض کنید M تصویر آموزش داریم. در این صورت، ماتریس کوواریانس به صورت زیر می‌شود:

$$C = \frac{1}{M}\sum_{i=1}^{M} \varphi_i \varphi_i^T = AA^T$$

در رابطه بالا، داریم: $A = [\varphi_1, \varphi_2, \dots, \varphi_M]$ که هر یک از φ_i ها یک بردار ستونی است.

اکنون، اگر مقادیر ویژه (λ) و بردارهای ویژه‌ی (u) این ماتریس کوواریانس را بدست آوریم:

$$AA^T u = \lambda u$$

هر تصویر صورتی را می‌توان به صورت یک ترکیب خطی (جمع وزن‌دار) از این بردارهای ویژه بیان کرد. چون این بردارهای ویژه یک فضای آماری برای بیان صورت را تشکیل می‌دهند، به آن‌ها Eigenface گفته می‌شود. به فضای ایجاد شده از این Eigenfaceها، face space گفته می‌شود.

دقت کنید که بعد اول A برابر با (N1*N2) است که خیلی بزرگ است. پس ابعاد AA^T هم خیلی بزرگ می‌شود. یافتن مقادیر و بردارهای ویژه آن بالطبع خیلی دشوار و زمان‌بر است. بهتر است برای کاهش محاسبات چاره‌ای اندیشید. اگر در رابطه زیر، بردارهای ویژه‌ی A^TA را v بنامیم، داریم:

$$A^T A v = \lambda v$$

در رابطه بالا، اگر طرفین رابطه را از چپ در A ضرب کنیم، خواهیم داشت:

$$(AA^T)(Av) = \lambda(Av)$$

با مقایسه سه رابطه قبل با این رابطه داریم:

$u = Av$

پس، می‌توان به جای محاسبه مقادیر و بردارهای ویژه AA^T که دارای ابعاد بزرگ (N1*N2)*(N1*N2) است، مقادیر و بردارهای ویژه A^TA را بدست آورد که دارای ابعاد M*M است؛ که این حجم محاسبات را بسیار پایین می‌آورد. پس از محاسبه بردارهای ویژه A^TA که v است، با استفاده از رابطه $u = Av$ بردارهای ویژه AA^T نیز بدست می‌آید.

حال دقت کنید که فضای آماری face space یک فضای (N1*N2) بعدی است که خیلی بزرگ است. صورتها ولیکن دارای مشخصات مشابه زیادی هستند. لذا اگر هر صورت را به منزله‌ی یک نقطه در این فضای (N1*N2) بعدی در نظر بگیریم، نقاط صورتها معمولا در یک محدوده کوچک از این فضا قرار می‌گیرند و به عبارتی، در یک زیرفضا (subspace) با ابعاد کمتر خلاصه می‌شوند. اگر M صورت آموزش داشته باشیم که M از بعد فضا (یعنی N1*N2) کمتر باشد، آنگاه فقط M-1 تا مقدار ویژه اولیه دارای مقدار قابل توجهی هستند و بقیه مقادیر ویژه کوچکتر دارای مقدار صفر یا نزدیک صفر هستند. لذا می‌توان با استفاده از PCA که در قسمت اول این آزمایش با آن آشنا شدید، فقط مقادیر ویژه برتر را نگاه داشت و با بردارهای ویژه متناظر آن مقادیر ویژه، فضای آماری با بعد کمتری را تشکیل داد. برای مثال، در مثال صورتهای بالا، بردارهای ویژه متناظر با ۷ مقدار ویژه برتر به صورت زیر هستند:

دقت کنید که هر چه مقدار ویژه یک بردار ویژه بزرگتر باشد یعنی در واریانس بین صورت‌ها آن بردار ویژه مهمتر است و نقش پررنگ‌تری دارد.

اکنون، به آزمایش می‌پردازیم. دیتاستی شامل چند زیر مجموعه از بانک داده Cohn-Kanade در اختیار شما قرار گرفته است. داده‌های مجموعه neutral را در یک حلقه با توابع کتابخانه opencv بخوانید. دقت نمایید که تصاویر به فرمت BGR هستند ولی لازم است شما آنها را به صورت gray ذخیره کنید. در صورت نیاز، برای نمایش تصویر می‌توانید از دستور imshow یا کتابخانه matplotlib استفاده نمایید. هر یک از تصاویر، به صورت یک ماتریس دوبعدی است. هر یک را به کمک دستور reshape به صورت یک بردار یک بعدی درآورید و به عنوان یک نمونه تصویر صورت در نظر بگیرید. در این مجموعه، تصاویر از حدود ۹۶ نفر موجود است که از هر فردی حداکثر دو تصویر قرار دارد. تصاویر ۷۰ فرد را برای آموزش و تشکیل فضای face space و تصاویر مابقی افراد را برای تست استفاده نمایید.

۱۰ مقدار ویژه برتر را در نظر بگیرید. تصاویر eigenface مربوط به این ۱۰ مقدار ویژه برتر را نمایش دهید. (راهنمایی: برای نمایش هر یک از eigenfaceها باید بردار ویژه مربوطه را به صورت یک ماتریس reshape N1*N2 کنید و سپس رسم کنید).

پنج عدد از تصاویری که در آموزش PCA استفاده کردید را به فضای Eigenface ساخته شده با ۱۰ بردار ویژه برتر تصویر نمایید. سپس، آن را از فضای Eigenface، بازسازی و رسم نمایید. دقت نمایید که حتما mean face را با مجموع وزندار بردار ویژه ها جمع کنید بازسازی را با تعداد ۲۰ بردار ویژه برتر تکرار کنید. و مشاهدات خود را بنویسید.

در این قسمت بازسازی را برای پنج نمونه از تصاویری که در آموزش PCA استفاده نکردید انجام دهید. چه مشاهده می‌کنید. آن را تحلیل نمایید و علت آن را بیان نمایید.

قسمت سوم: زیر فضاهای احساسات

زیر فضای احساسات خنثی، خوشحال، تعجب و انزجار را با استفاده از تصاویر مجموعه هر احساس بسازید. برای اینکه این زیرفضاها بیان کننده تغییرات اجزای صورت در احساسات مختلف باشند، در تمام تصاویر، قبل از شروع آموزش PCA، قسمتهای خارج از صورت مانند موها و پس زمینه را با استفاده از یک ماسک ثابت بپوشانید طوری که حتی الامکان فقط اجزای صورت مشخص باشند.

حال تعداد ۱۰ تصویر تصادفی از مجموعه خنثی انتخاب کنید و در سه زیرفضای دیگر آن ها را بازسازی کنید. مشاهدات خود را بنویسید و تحلیل کنید. آیا هویت افراد در بازسازی حفظ شده است؟ آیا احساس آن ها تغییر کرده است؟ کدام یک از اجزای صورت بهتر بازسازی شده اند؟

پیش گزارش

سه نقطهٔ (۱، ۱) و (۳، ۲) و(۲، ۳) را در یک فضای دوبعدی داریم.

- اگر بخواهیم این داده‌ها را به فضـای یک بعدی ببریم، مقدار ویژهٔ برتر و بردار ویژهٔ[1] نظیر آن را به دسـت آورید.
- داده‌ها را به فضای یک بعدی ببرید و خطای بازسازی را حساب کنید.

[1] First Principle Component