

به نام خداوند بخشنده مهربان



دانشگاه صنعتی امیرکبیر

پروژه نهایی درس ریاضی مهندسی

استاد درس: دکتر امیرمزلقانی

موعد تحویل: یکشنبه یکم بهمن ماه ۱۳۹۶

سال تحصیلی ۹۷ - ۹۶

قسمت اول

یک مسأله‌ی دو کلاسه را در نظر گرفته و ۱۰۰۰ نمونه تصادفی با توزیع گوسی با بردارهای میانگین و ماتریس کوواریانس داده شده در زیر برای هر کدام تولید کنید.

$$\mu_1 = [10, 10]^T, \mu_2 = [22, 10]^T, \Sigma_1 = \Sigma_2 = \begin{bmatrix} 4 & 4 \\ 4 & 9 \end{bmatrix}$$

الف) خطی را که PCA داده‌ها را روی آن تصویر می‌کند محاسبه و رسم کنید.

ب) همه‌ی نقاط داده را روی خط PCA حاصله تصویر کرده و نتیجه را نمایش دهید.

ج) آیا چیزی که مشاهده می‌کنید مورد انتظارتان بود؟ توضیح دهید.

د) نقاط داده را در فضای دوبعدی بازسازی کرده و خطای بازسازی را محاسبه کنید.

قسمت دوم

در این قسمت شما روش تصویر ویژه (Eigenface) را برای شناسایی چهره انسان پیاده‌سازی خواهید کرد. تصاویر چهره از Yale Face Database B بوده که در آن ۶۴ تصویر تحت شرایط نور متفاوت از هر یک، جمعاً ۶۴۰ تصویر، وجود دارد. با پیاده‌سازی شما قدرت تجزیه مقدار منفرد (SVD) را در نمایش تصاویر درخواهید یافت.

الف) پس از این که faces.zip را unzip کردید، پوشه‌ای را می‌یابید که حاوی همه‌ی تصاویر آموزش و تست است؛ train.txt و test.txt به ترتیب، مجموعه‌ی آموزش و تست را مشخص می‌کند، هر خطی مسیر تصویر و برچسب متناظر آن را ارائه می‌دهد.

ب) تصاویر مجموعه آموزش را در یک ماتریس X بارگذاری کنید: در مجموع ۵۴۰ تصویر آموزش وجود دارد که هر کدام ۵۰×۵۰ پیکسل است و باید به یک بردار ۲۵۰۰ بعدی تبدیل شود. بنابراین سائز X باید ۵۴۰×۲۵۰۰ باشد که هر سطری یک تصویر چهره تخت شده است. یک تصویر چهره از ماتریس X را انتخاب کرده و آن را

نمایش دهید. مراحل گفته شده را برای مجموعه تست نیز انجام دهید. سائز ماتریس X_{test} برای مجموعهی تست باید 2500×100 باشد.

(ج) تصویر میانگین: تصویر میانگین μ از با استفاده از همهی مجموعهی آموزش محاسبه کنید. تصویر میانگین را نمایش دهید.

(د) تفریق میانگین: تصویر میانگین μ را از همهی سطریهای ماتریس X کم کنید. یک تصویر چهره را بعد تفریق میانگین انتخاب کرده و آن را نمایش دهید. همین کار را با استفاده از بردار میانگین محاسبه شده توسط تصاویر آموزش در قسمت (ج) برای تصاویر تست انجام دهید.

(ه) تصویر ویژه: تجزیه مقدار منفرد را روی مجموعه $X (X=U\Sigma V^T)$ اجرا کنید تا ماتریس V^T حاصل شود، که هر سطر ماتریس V^T همان بعد تصاویر چهره را دارد. ما v_i (امین سطر V^T) را به عنوان امین تصویر ویژه در نظر می‌گیریم. ۱۰ تصویر ویژه اول را نمایش دهید.

(و) تقریب مرتبه پایین: از آن جا که Σ یک ماتریس قطری با ترتیب غیرافزایشی است، می‌توانیم r عنصر اول Σ را با r ستون اول U و r سطر اول V^T برای تقریب \hat{X}_r استفاده کنیم. ماتریس \hat{X}_r تقریب مرتبه r ماتریس X نامیده می‌شود. نمودار خطای تقریب مرتبه r را به عنوان تابعی از r رسم کنید. ($r=1, 2, 3, \dots, 200$)

(ز) ویژگی‌های تصویر ویژه: r تصویر ویژه بالای تصاویر ویژه یک زیر فضای خطی r بعدی فضای تصاویر اصلی به نام فضای چهره را اسپن می‌کند، که مرکز آن تصویر میانگین و محورهای آن تصاویر ویژه $\{v_1, v_2, \dots, v_r\}$ هستند. بنابراین با استفاده از r تصویر ویژه بالا می‌توانیم یک تصویر چهره z ۲۵۰۰ بعدی را به عنوان یک بردار ویژگی r بعدی $F=V^T$ نمایش دهیم. تابعی بنویسید که ماتریس ویژگی r بعدی F و F_{test} را به ترتیب برای تصاویر آموزش X و تصاویر تست X_{test} تولید کند. (برای تولید F ماتریس X را در ترانهادهی r سطر اول V^T ضرب کنید، F باید همان تعداد سطر X و r ستون داشته باشد؛ به طور مشابه برای X_{test})

(ح) شناسایی تصویر: ویژگی‌های آموزش و تست را به ازای $r=10$ استخراج کنید. یک مدل رگرسیون خطی با استفاده از F آموزش دهید و روی F_{test} امتحان کنید. دقت طبقه‌بندی را روی مجموعه گزارش دهید. نمودار دقت طبقه‌بندی روی مجموعه تست را به عنوان تابعی از r رسم کنید. ($r=1, 2, \dots, 200$)

- گزارش پروژه را به صورت **report_stdNum.pdf** نامگذاری کنید.
- به ازای هر بخش یک پوشه شامل پیاده‌سازی و تحلیل و گزارش ایجاد کنید.
- پروژه خود را قبل از زمان مشخص شده در مدل آپلود کنید.