



دانشکده علوم ریاضی

بهینه‌سازی محدب ۱

نیم‌سال دوم ۱۴۰۲-۱۴۰۱

مدرس: دکتر یاسایی

تمرین سری اول

شماره دانشجویی: ۹۸۱۰۰۴۱۸

نام و نام خانوادگی: آیلار خرسندنیا

پرسش ۱ مقادیر ویژه و بردارهای ویژه

(۱)

استفاده از تجزیه‌ی QR به صورت iterative روشی برای پیدا کردن مقادیر ویژه و بردارهای ویژه یک ماتریس می‌باشد. این روش از دو ویژگی ماتریس‌ها استفاده می‌کند:

۱. ماتریس‌های متشابه دارای مقادیر ویژه و بردارهای ویژه‌های متناظر یکسان هستند. دو ماتریس A و B متشابه هستند اگر:

$$A = C^{-1}BC$$

۲. در روش QR ماتریس A به دو ماتریس Q (که یک ماتریس با ستون‌های متعامد یکه) و R (یک ماتریس بالا مثلثی) می‌رسیم. به ازای هر ماتریس Q با ستون‌های متعامد یکه داریم:

$$Q^{-1} = Q^T$$

تجزیه‌ی ماتریس A برابر با $A = QR$ می‌شود. به کمک ویژگی دوم داریم:

$$Q^T A = R$$

حال ماتریس A_1 را به گونه‌ای می‌نویسیم که با ماتریس $(A = A_0)$ متشابه باشند تا مقادیر ویژه تغییری نکنند.

$$A_1 = Q^T A Q = R Q$$

به همین ترتیب با به دست آوردن تجزیه‌ی QR ماتریس A_{k+1} را به صورت

$$A_{k+1} = R_k Q_k = Q_k^{-1} Q_k R_k Q_k = Q_k^T A_{k+1} Q_k$$

قرار می‌دهیم. در نتیجه تمام A_K ها متشابه‌اند و مقادیر ویژه یکسان دارند. هر چقدر این روند را به صورت iterative انجام دهیم به یک ماتریس بالا مثلثی می‌رسیم که درآیه‌های قطر اصلی شامل مقادیر ویژه‌ی ماتریس A می‌شوند. (۲)

کد دقیقاً مطابق الگوریتم شرح داده شده در بخش قبل زده شده است. (۳)

با انجام الگوریتم شرح داده شده در بخش اول، ماتریسی می‌سازیم که درآیه‌های قطر اصلی (یا همان مقادیر ویژه‌ی ماتریس بالا مثلثی) حقیقی هستند. اما ماتریس‌هایی داریم که مقادیر ویژه‌ی آن‌ها مختلط غیر حقیقی می‌باشند و با این الگوریتم نمی‌توان مقادیر ویژه آن‌ها را یافت.

از طرفی این الگوریتم زمانی که قدر مطلق مقادیر ویژه یکسان نباشند به ماتریس بالا مثلثی تبدیل می‌شود. چرا که در این صورت درآیه‌های زیر قطر اصلی A_K به 0 میل می‌کند. اما اگر دو مقدار ویژه λ_i و λ_j با اندازه‌های برابر داشته باشیم، درآیه‌ی ماتریس a_{ij} ماتریس A_K برابر 1 می‌شود. زیرا:

$$|a_{ij}^{(K)}| = O\left(\left|\frac{\lambda_i}{\lambda_j}\right|^K\right) \quad |\lambda_i| < |\lambda_j|$$

روش دیگر برای پیدا کردن مقادیر ویژه power iteration می‌باشد. به این صورت که بردار دلخواه a_0 را در نظر می‌گیریم و:

$$a_{K+1} = \frac{Aa_K}{\|Aa_K\|}$$

در نتیجه a_K ها به بردار ویژه متناظر با بزرگ‌ترین مقدار ویژه میل می‌کند و از روی آن می‌توان λ_1 را پیدا کرد. محدودیت روشن power iteration:

اگر تکرار یزگ‌ترین مقدار ویژه بیشتر از یک باشد نمی‌توان از این روشن استفاده کرد.

روشن divide and conquer: در این روشن ماتریس سه قطری را برابر حاصل جمع ماتریس قطری بلوکی و یک ماتریس با رتبه‌ی یک قرار می‌دهیم. حال ماتریس قطری بلوکی را به دو ماتریس کوچک‌تر با اندازه‌های برابر تقسیم می‌کنیم و آن‌ها را نیز با همین روشن بازگشتی حل می‌کنیم تا جایی که پیدا کردن مقادیر ویژه ماتریس راحت‌تر باشد. مقادیر ویژه ماتریس اصلی از محاسبه‌ی این مسایل کوچک‌تر حل می‌شوند.

محدودیت روشن divide and conquer:

این روشن فقط برای ماتریس‌های سه قطری متقاضی حقیقی جواب می‌دهد.

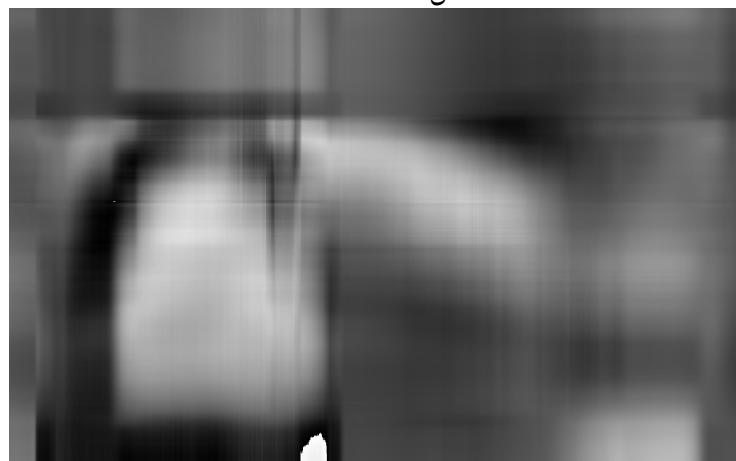
پرسش ۲ پردازش تصویر

(۱)

در بخش اول سوال ۲ ابتدا تجزیه‌ی SVD ماتریس را حساب می‌کنیم. حال هدف آن بود که با نگه داشتن پایه‌های مهم‌تر و حذف اطلاعات کم اهمیت‌تر عکس را فشرده کنیم. از آنجایی که مقادیر تکین بزرگ‌تر از اهمیت بیشتری برخوردارند پس باید در بازسازی دوباره ماتریس تصویر حضور دارند. با نگه داشتن K ستون ابتدایی از پایه‌های یافته شده و مقادیر تکین و بردارهای ترکیبی (به ازای های k مختلف) شکل را بازسازی می‌کنیم و با تصویر اولیه به کمک معیار PSNR مقایسه می‌کنیم. همانطور که از نمودار معلوم است، هر چه رنک بیشتر، تصویر بازسازی شده به عکس اصلی نزدیک‌تر می‌شود. اما روند صعودی در بازه‌ی $0 \dots 30$ خیلی بیشتر است و شبیه نمودار بعد از آن کمتر می‌شود. این به آن معناست که با افزایش رنک، تغییر نامحسوسی در شکل ایجاد می‌شود و در همان رنک‌های کمتر اطلاعات مهم نیز نشان داده می‌شود.

(۲)

شكل ۱: $rank = 4$



شكل ۲: $rank = 15$



شكل = 32 : ٣



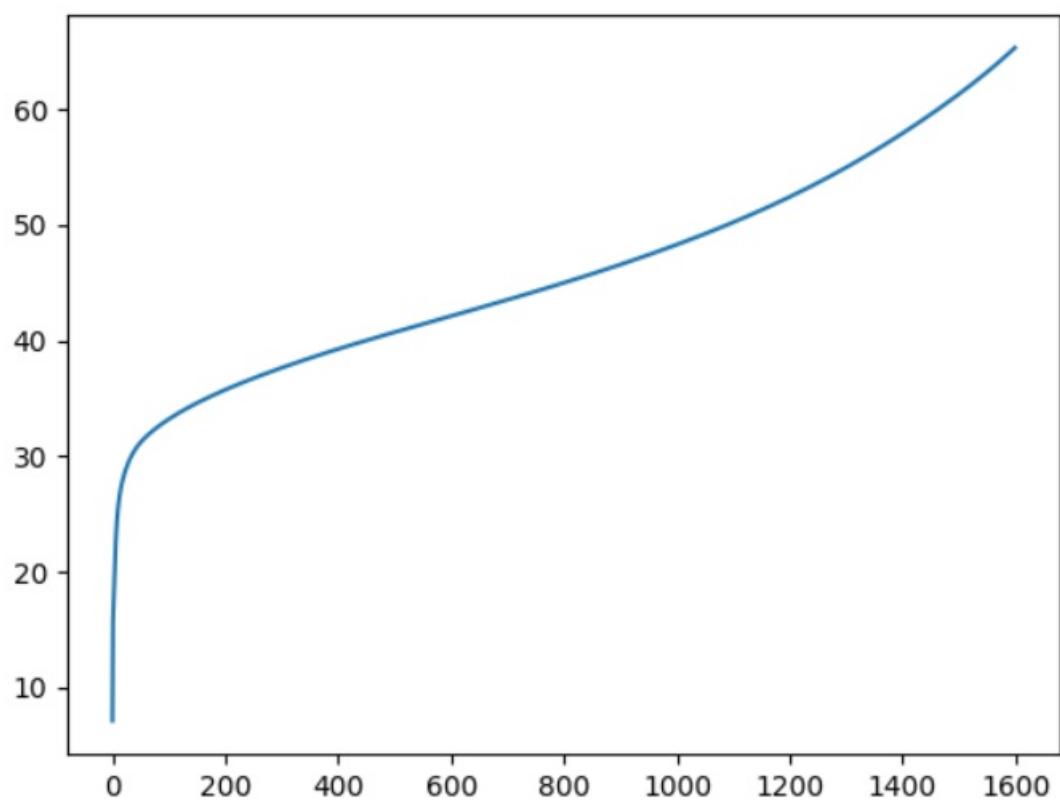
شكل = 60 : ٤



شكل = 175 : ٥



شكل ٦ PSNR :

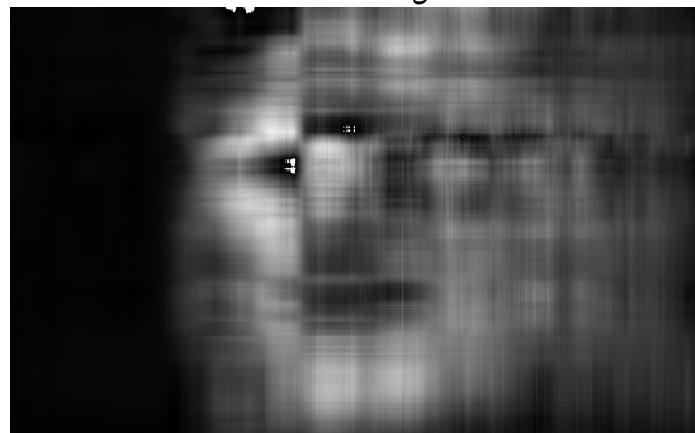


با عکس دیگری نیز این کار را تکرار می‌کنیم و با توجه به نمودار PSNR آن به سختی می‌شود بهترین رنک را انتخاب کرد ولی دیدن تصاویر می‌یابیم رنک ۸۰ اطلاعات خوبی راجع به عکس می‌یابیم و رنک‌های بالاتر جزیات خیلی کمی به تصویر اضافه می‌کند.

شکل ۷:



شکل ۸: $rank = 4$



شکل ۹: $rank = 35$



شكل ١٠ : $rank = 50$



شكل ١١ : $rank = 80$



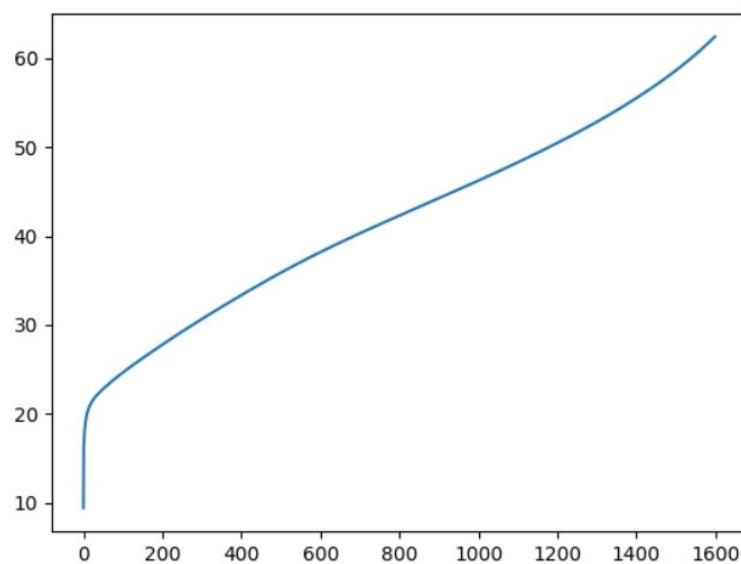
شكل ١٢ : $rank = 150$



شكل : ۱۳ rank = 200

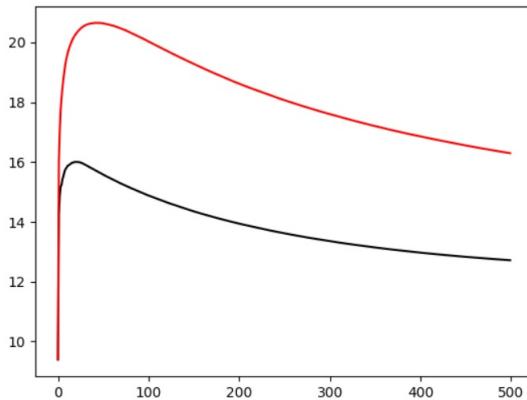


شكل : ۱۴ lionPSNR



در بخش دوم سوال ۲ دو نویز Gaussian و Salt and pepper را جداگانه به عکس اضافه می‌کنیم و تصاویر خروجی را بررسی می‌کنیم.
با رسم نمودار PSNR برای هر دو نویز می‌یابیم که دینویز کردن هر دو نویز با رنک نزدیک به ۵۰ بهترین حالت عکس بازسازی می‌شود.

شکل ۱۵ : Gaussian and sp PSNR



شکل ۱۶ : Gaussian noise added



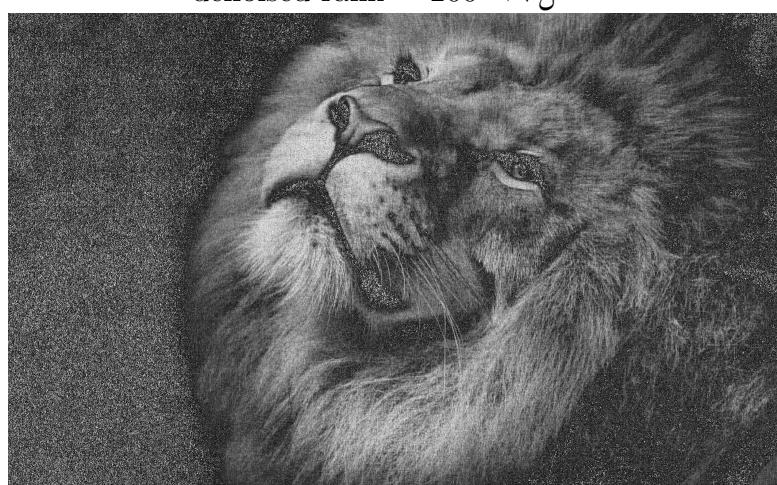
denoised rank = 20 : شکل ۱۷



denoised rank = 50 : شکل ۱۸



denoised rank = 200 : شکل ۱۹



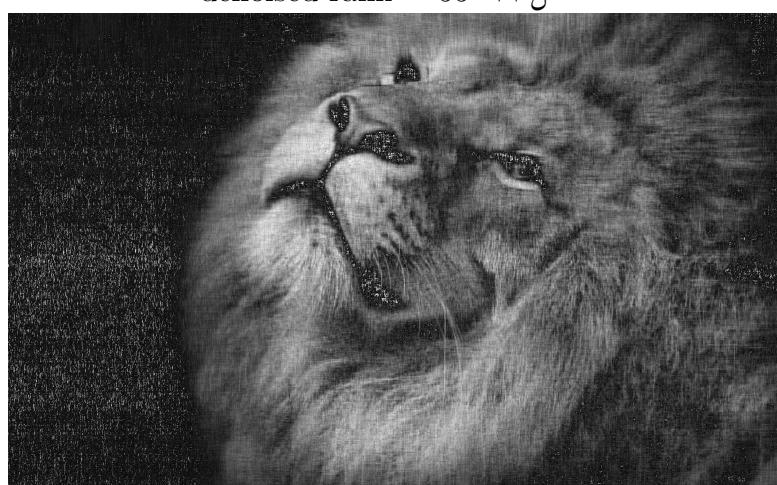
شکل ۲۰ : salt and pepper added



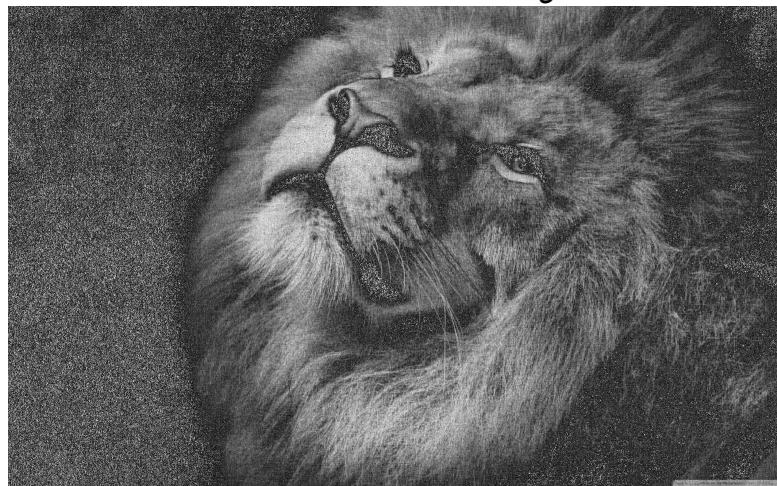
شکل ۲۱ : denoised rank = 30



شکل ۲۲ : denoised rank = 55

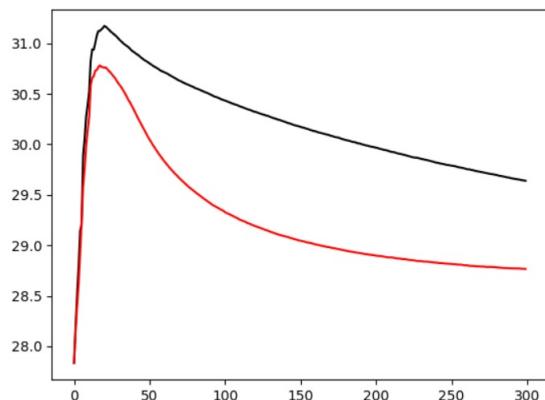


denoised rank = 200 : ٢٣ شكل



در شکل پرنده با توجه به نمودار، PSNR می‌یابیم که بهترین حالت دینویز کردن با رنک ۲۵ انجام می‌شود.

شکل Gaussian and sp PSNR : ۲۴



شکل Gaussian added : ۲۵



شکل denoised rank = 13 : ۲۶



denoised rank = 25 : ٢٧



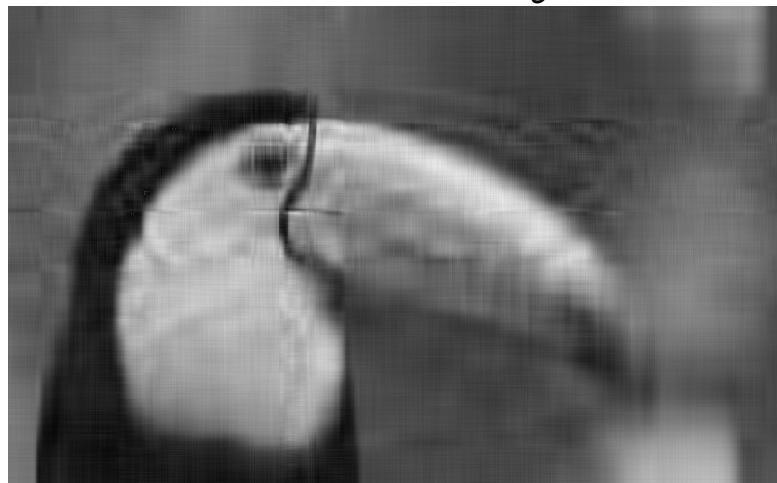
denoised rank = 150 : ٢٨



salt and pepper added : ٢٩



denoised rank = 13 : شکل ۳۰



denoised rank = 25 : شکل ۳۱



denoised rank = 150 : شکل ۳۲



پرسش ۳ کاهش بعد داده با روش PCA

(۱)

از تعریف ماتریس کوواریانس به راحتی می‌توان دید که تغییر میانگین داده‌ها تاثیری روی ماتریس کوواریانس ندارد زیرا در فرمول آن باید ماتریس میانگین ماتریس را از ماتریس اصلی کم کنیم.

با توضیحات بالا مشخص است که ماتریس کوواریانس مجموعه داده‌ها برابر ماتریس کوواریانس ماتریس \tilde{X} است. زیرا این ماتریس همان ماتریس مجموعه داده‌هاست که میانگینش را صفر کرده‌ایم.

حال از جبرخطی می‌دانیم که برای هر ماتریس X دلخواه، ماتریس XX^T یک ماتریس متقارن و مثبت نیمه معین می‌باشد و برای چنین ماتریس‌هایی می‌دانیم که تجزیه SVD همان نتیجه‌ی تجزیه مقدار ویژه را به ما می‌دهد. (با صرف نظر از جابجایی سطر و ستون‌ها)

بنابراین ماتریس U ماتریسی است که ستون‌های آن، بردار ویژه‌ی متناظر با مقدار ویژه‌ی λ است. و ماتریس Σ نیز همان ماتریس Λ است. (البته در ماتریس Σ مقادیر ویژه به ترتیب آمده‌اند). و همچنین ماتریس V نیز برابر همان U می‌باشد. پس در نهایت داریم:

$$\text{ماتریس کوواریانس مجموعه داده‌ها} = \tilde{X}\tilde{X}^T = U\Sigma V^T = U\Sigma U^T$$

(۲)

باید ماتریس W به گونه‌ای باشد که با برگرداندن داده‌ی λ – بعدی به فضای n – بعدهی کمترین میزان از دست رفتن اطلاعات را داشته باشیم.

وقتی داده‌ی x_i را با ماتریس W کاهش بعد می‌دهیم، می‌توانیم با ماتریس W^T ، آن را به فضای \mathbb{R}^n برگردانیم و در این صورت میزان از دست رفتن اطلاعات این داده برابر $\|x_i - W^T(Wx_i)\|^2$ می‌شود.

حال به جای ماتریس W ابتدا فرض کنید می‌خواهیم بهترین بردار یکه‌ای را پیدا کنیم که بهترین نتیجه را به ما بدهد، یعنی میزان عبارت زیر را کمینه کند.

$$\begin{aligned} \sum_{i=1}^m \|x_i - u(u^T x_i)\|^2 &= \sum_{i=1}^m (\|x_i\|^2 + x_i^T uu^T uu^T - 2x_i^T uu^T x_i) \\ &= \sum_{i=1}^m \|x_i\|^2 - \sum_{i=1}^m (x_i^T uu^T x_i) = \sum_{i=1}^m \|x_i\|^2 - \sum_{i=1}^m (u^T x_i x_i^T u) = \sum_{i=1}^m \|x_i\|^2 - u^T \left(\sum_{i=1}^m x_i x_i^T \right) u \\ &= \sum_{i=1}^m \|x_i\|^2 - u^T X X^T u = \sum_{i=1}^m \|x_i\|^2 - \frac{u^T X X^T u}{u^T u} \end{aligned}$$

از آنجایی که عبارت اول ثابت است و ربطی به انتخاب u ندارد پس برای کمینه کردن کل عبارت باید عبارت دوم را بیشینه کنیم. طبق قضایایی که در درس خواندیم می‌دانیم این عبارت به ازای بردار ویژه متناظر با بزرگ‌ترین مقدار ویژه بیشینه می‌شود. پس بردار u باید همان ستون اول ماتریس U باشد.

با ادامه دادن این فرآیند در مرحله‌ی λ ام به بردار ویژه متناظر با λ امین مقدار ویژه می‌رسیم و در انتهای λ ستون اول ماتریس U می‌رسیم پس ادعای مطرح شده را ثابت کردیم.

(۳)

به همین گونه که در بخش قبل مطرح کردیم اگر هر ستون ماتریس U که هر کدامشان یک بردار ویژه برای ماتریس XX^T می‌باشد را در یک داده ضرب داخلی کنیم به یک عدد می‌رسیم که مربوط به آن داده می‌باشد. چون می‌خواستیم داده‌ها را به فضای \mathbb{R}^l ببریم پس هر داده را باید در λ بردار ویژه ضرب داخلی کنیم و بردار λ تایی حاصل را به عنوان داده‌ی جدید با بعد کمتر در نظر بگیریم.

در بخش قبل ثابت کردیم آن λ تا باید متناظر با λ تا بزرگ‌ترین مقدار ویژه‌ها باشند که از لحاظ شهودی نیز واضح است. زیرا برای تشکیل ماتریس اصلی این λ بردار در اعداد بزرگ‌تری ضرب می‌شوند و تاثیر بسیار بیشتری دارند پس درواقع λ مولفه‌ی اصلی هستند.

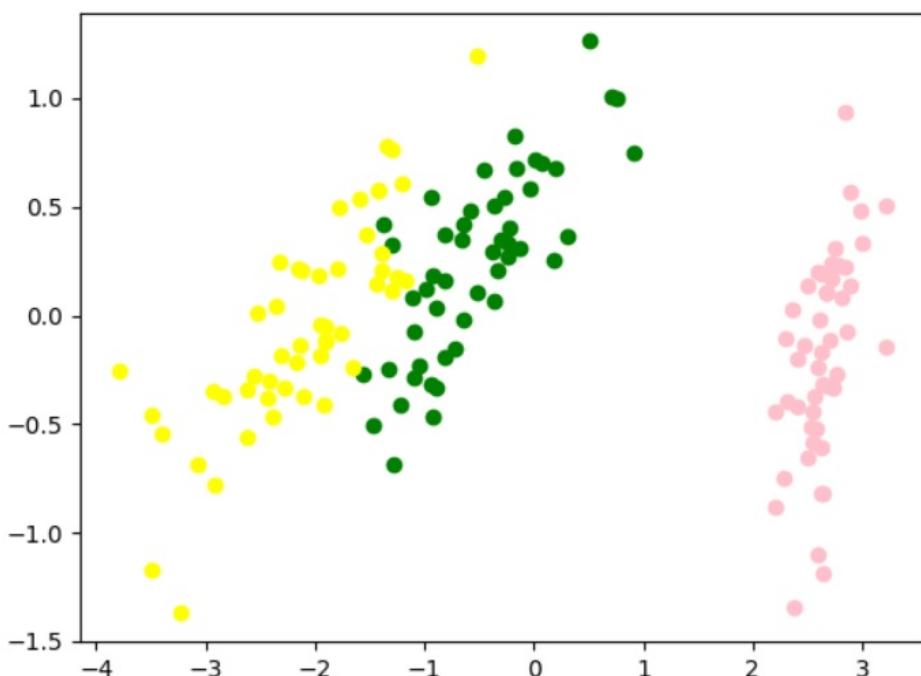
(۴)

یک الگوریتمی که می‌توانیم استفاده کنیم این است که به جای استفاده از تجزیه SVD با روش‌های تقریبی، بردارهای ویژه را محاسبه کنیم. مثلاً روش Power Iteration که در بخش سوم سوال ۱ نیز به آن کمی پرداختیم. با این کار می‌توانیم ۷ بردار ویژه مورد نظرمان را محاسبه کنیم و از اردر کمتری نسبت به SVD این کار را انجام دهیم.

(۵)

مانند چیزی که در بخش ۲ و ۳ به آن پرداختیم ابتدا میانگین داده‌ها را با ساختن \tilde{X} ، صفر می‌کنیم و سپس با استفاده از تجزیه SVD ماتریس U را به دست می‌آوریم و ۷ ستون اول آن را جدا می‌کنیم و با استفاده از این ماتریس، ماتریس داده‌های جدید با بعد کمتر را تشکیل می‌دهیم. داده‌های جدید که ۲ – بعدی هستند را روی نمودار رسم می‌کنیم و متناظر با هر نوع گل رنگ متفاوتی به آن نقاط می‌دهیم تا نمودار زیر حاصل شود.

شکل ۳۳: نمودار ویژگی گل‌ها با بعد کمتر



بله، این نمودار با نتیجه‌های که انتظار داشتیم مطابقت دارد. زیرا با کاهش بعد همچنان انواع گل‌های مختلف از هم جدا هستند و فقط در رنگ سبز و زرد میزان خیلی کمی تلاقی وجود دارد. اما در مجموع انواع مختلف گل‌ها با ۲ ویژگی اصلی به خوبی از هم جدا شده‌اند.