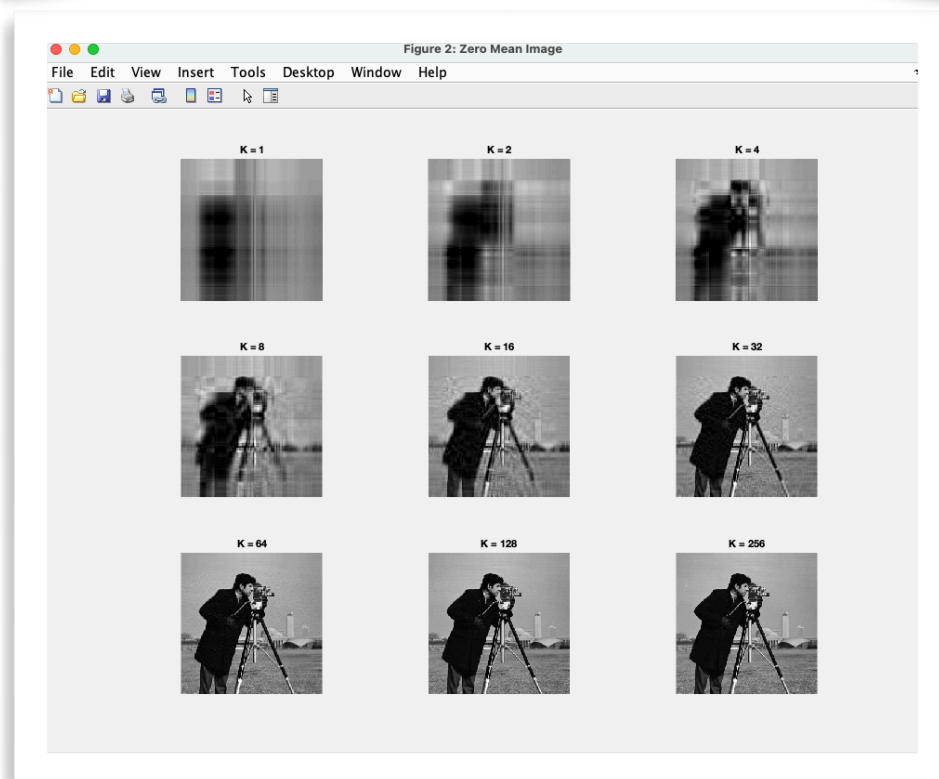
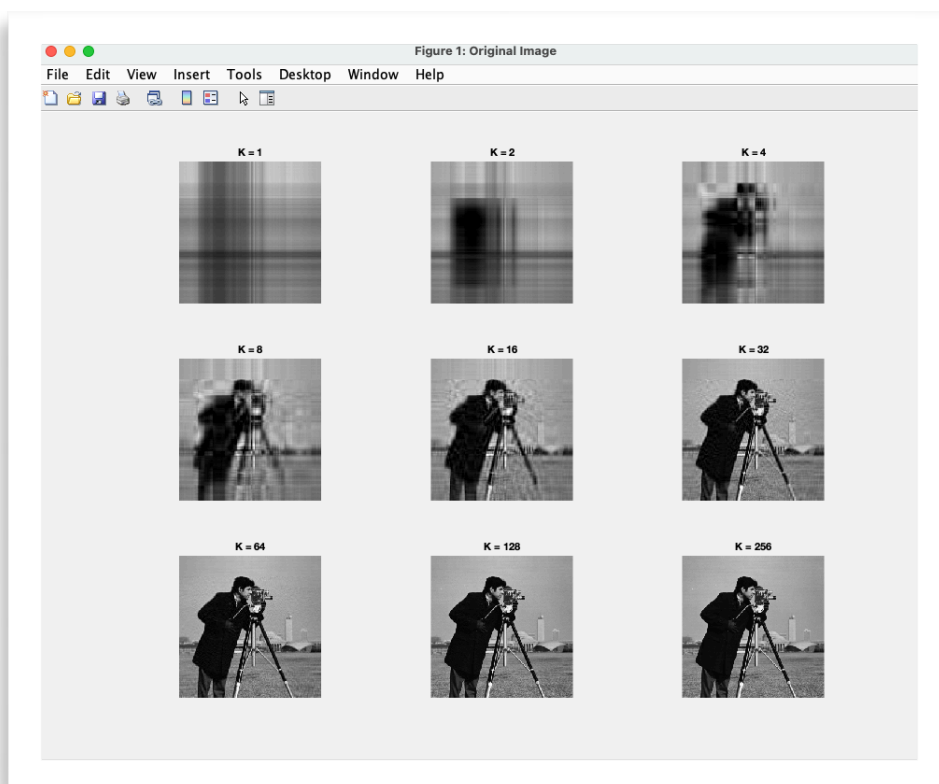
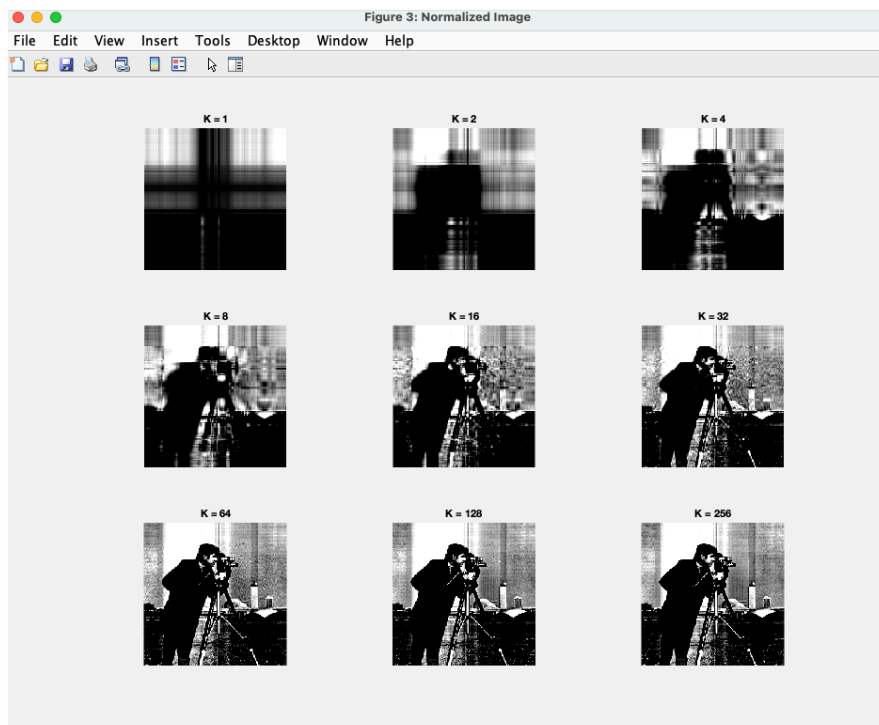
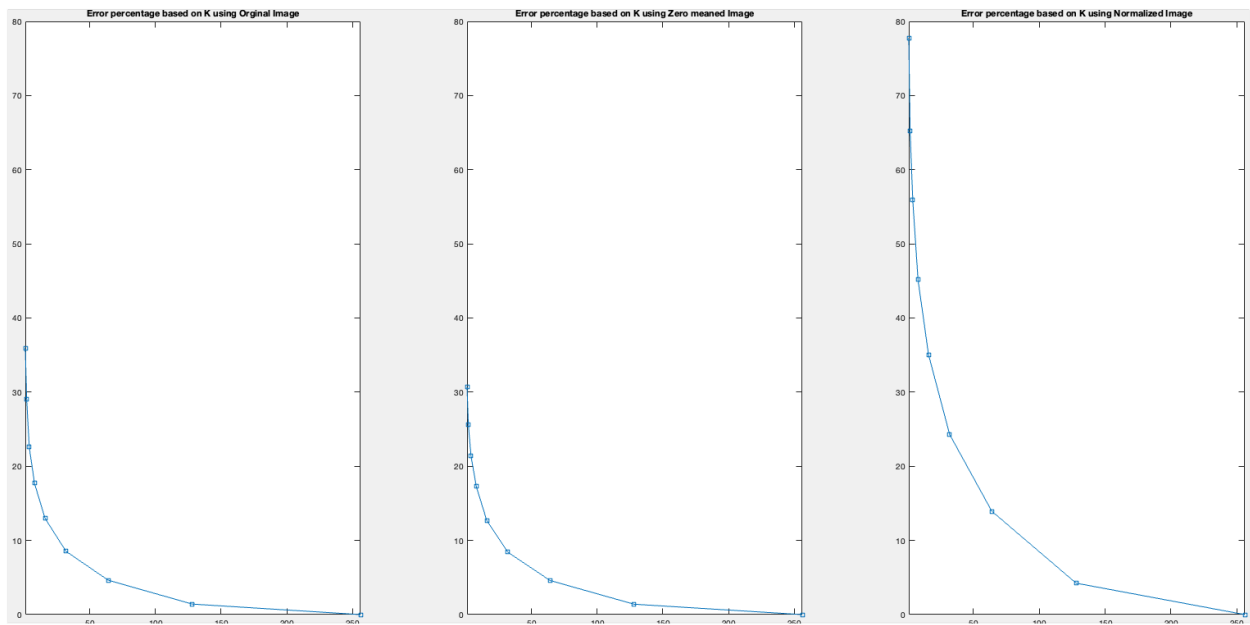


الف) برای نرمالیزاسیون تصویر یکبار با صفر کردن میانگین عکس و جمع کردن مجدد میانگین بعد از اعمال SVD در بازگشت به فضای تصویر و بار دیگر با استفاده از دستور `normalize` متلب و محاسبه ی خطا نسبت به خود عکس نرمال شده نتایج به شرح زیر است.

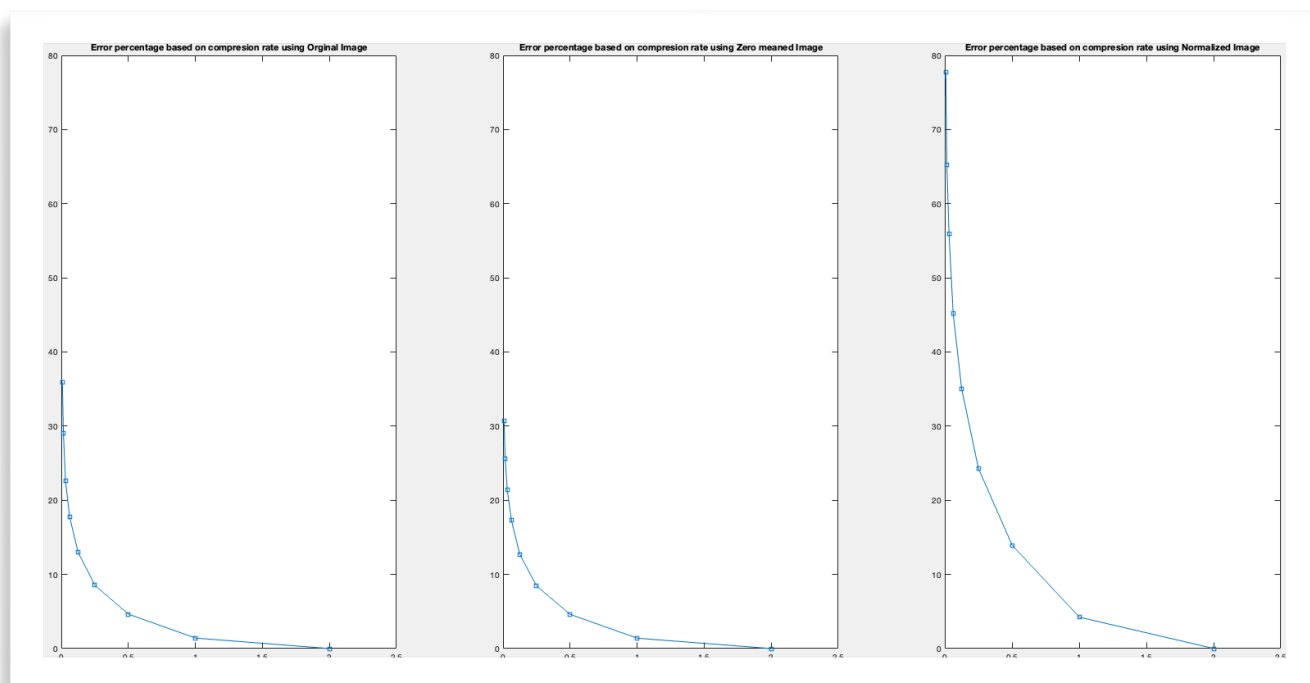




ب) نمودار درصد خطا برحسب رتبه به شرح زیر



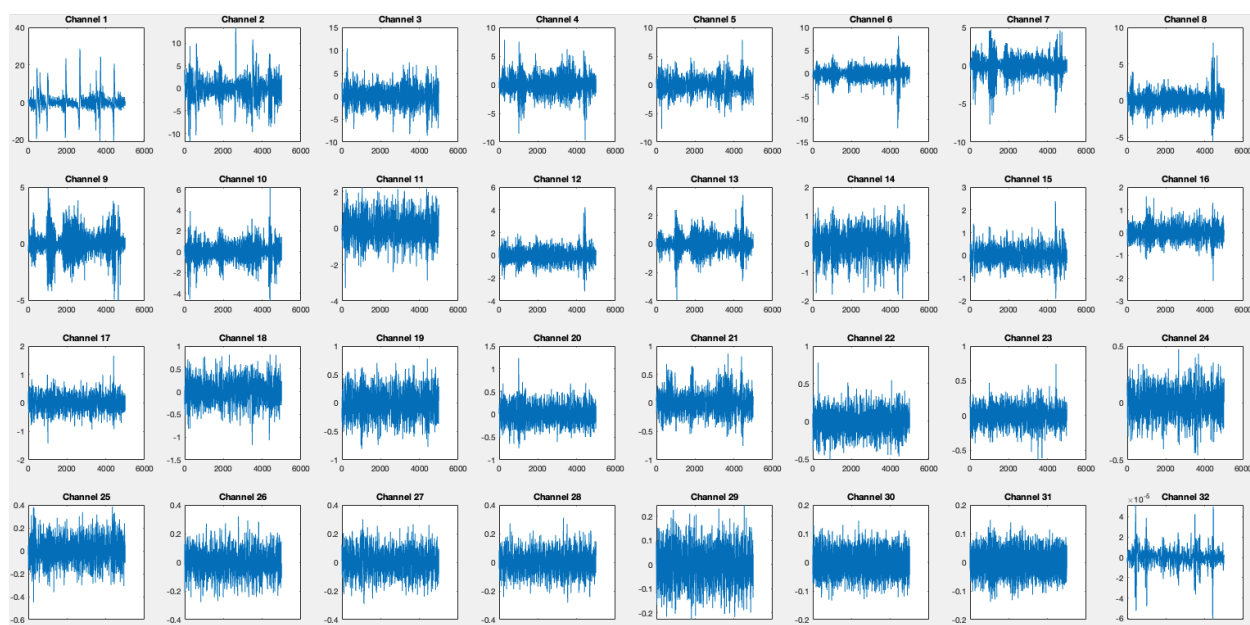
پ) و بر حسب فشردن شدن تصویر به شرح زیر می باشد.



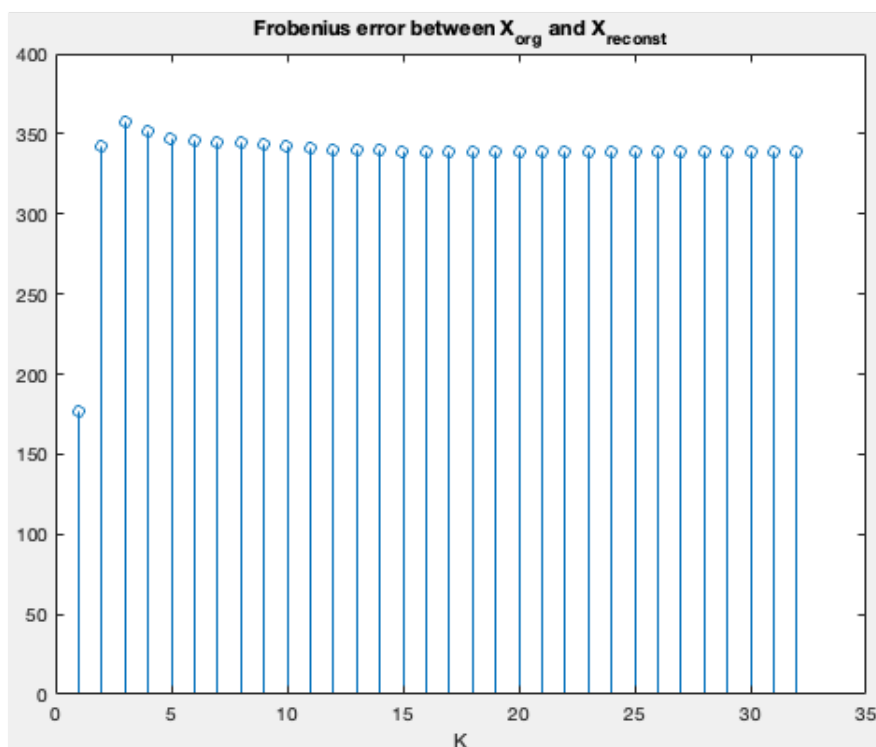
نتیجه ای که موارد بالا به ما نشان می دهد این است که زمانی که میانگین را کم می کنیم و سپس SVD کوتاه شده را اعمال می کنیم و مجدد میانگین را جمع می کنیم خطای بازسازی با تعداد عناصر یکسان کمتر می باشد. همچنین روند خطای تصویر بازسازی شده به شکل شبه لگاریتمی بر حسب تعداد عناصر انتخاب شده یا میزان فشردگی تصویر نزولی می شود که طبیعی هم هست چرا که عناصر اولیه ای که برمیداریم و لحاظ می کنیم مقادیر تکین بزرگتری دارد.

۲.

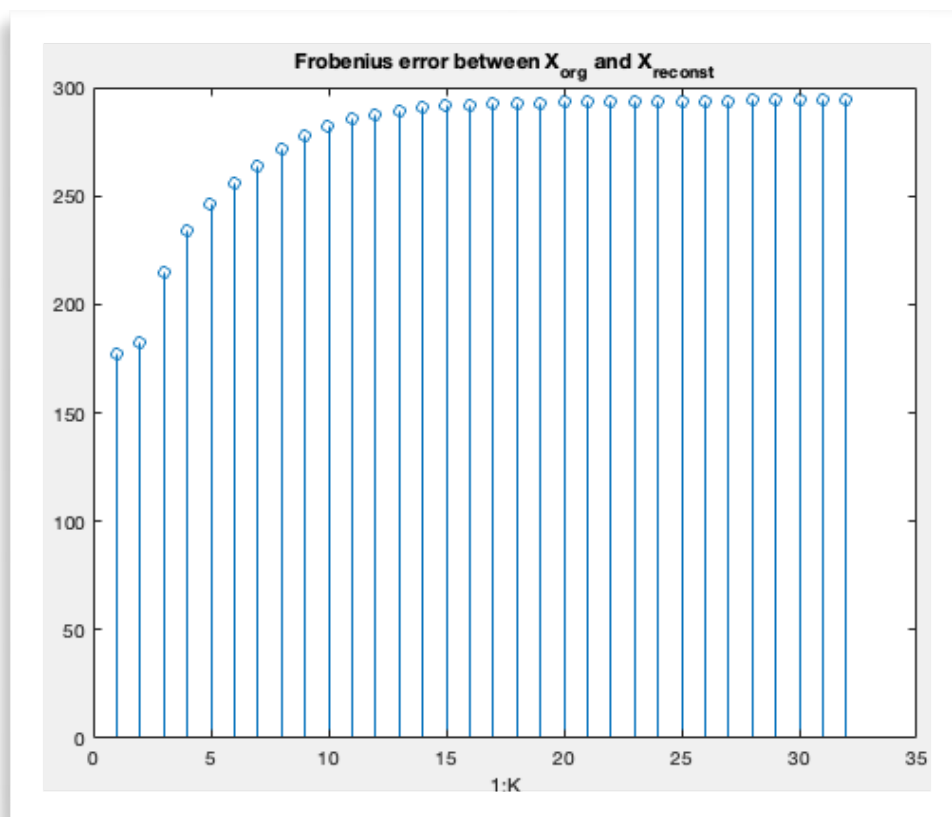
در ابتدا اگر بخواهیم منابع را رسم کنیم، ۳۲ منبع به ترتیب به شکل زیر در می آیند.



حال اگر نمودار خطا برحسب بازسازی سیگنال با استفاده از یک منبع را رسم کنیم به شرح زیر می شود:

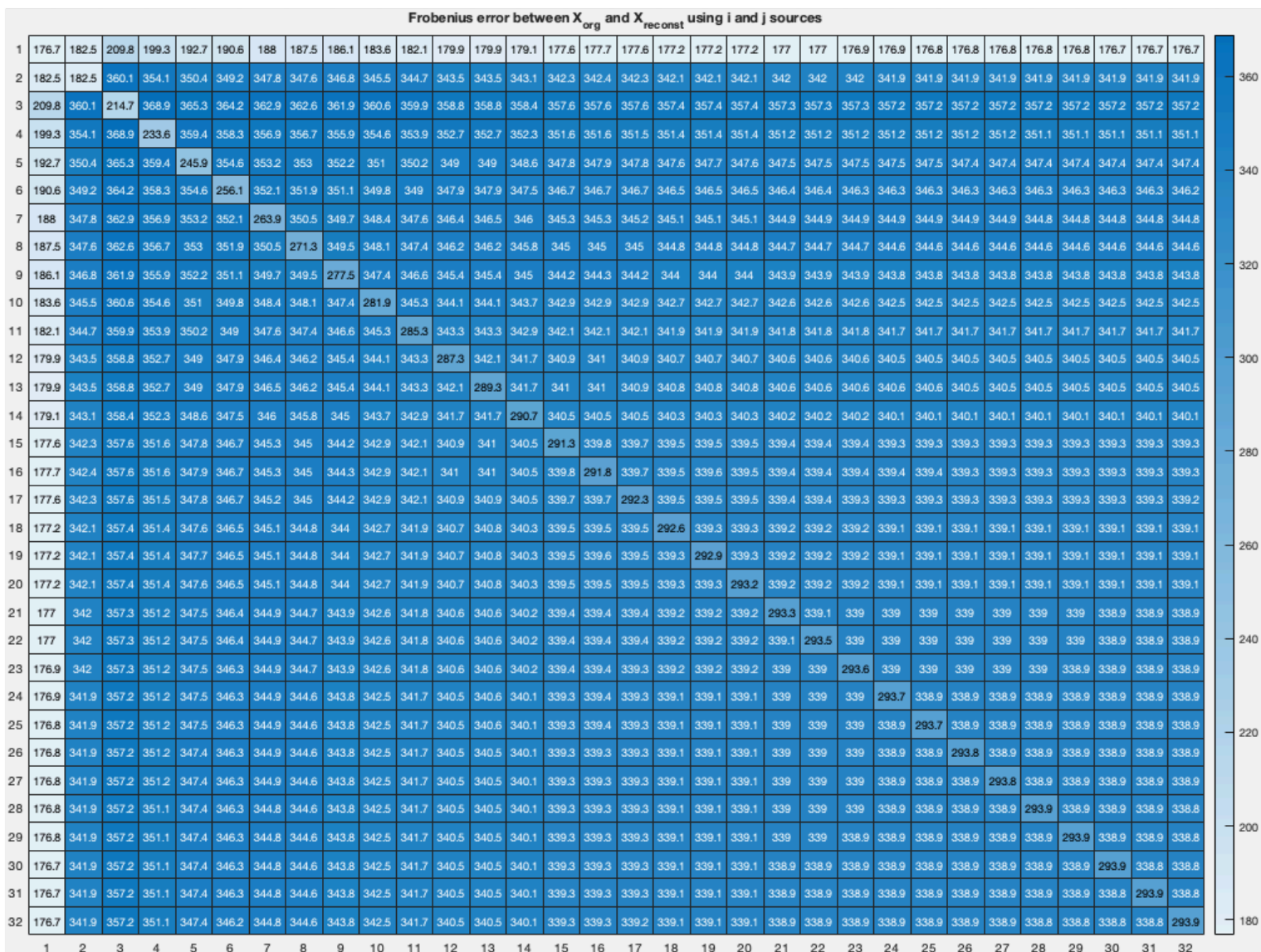


همانطور که مشاهده می شود، به صورت تک منبع با اختلاف خطای سیگنال بازسازی شده با منبع اول نویز کمتری نسبت به بقیه منابع دارد. از طرفی نمودار خطای سیگنال بازسازی شده نسبت به سیگنال اصلی بدون نویزی وقتی که از k منبع اول تجزیه استفاده می کنیم به صورت زیر می شود.



در نتیجه به نظر می رسد که صرفاً منبع اول حاوی اطلاعات مفید در راستای بازسازی سیگنال بوده و با اضافه شدن منابع دیگر صرفاً نویز سیگنال بالاتر می رود.

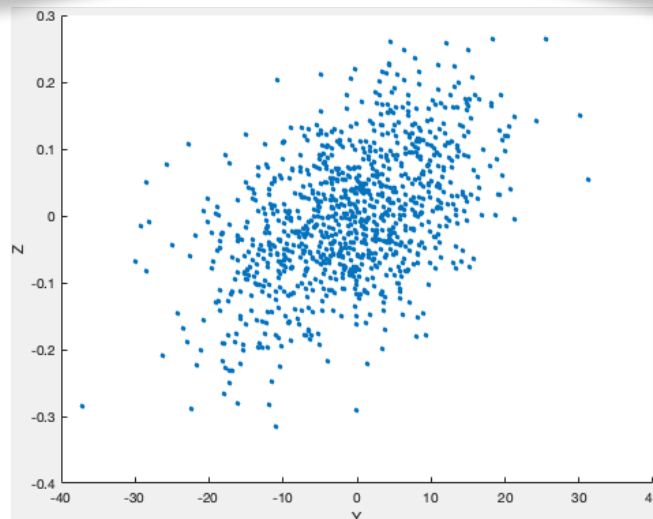
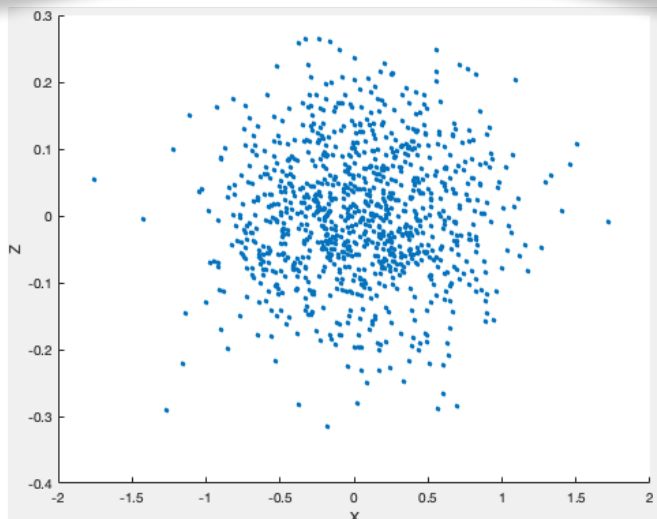
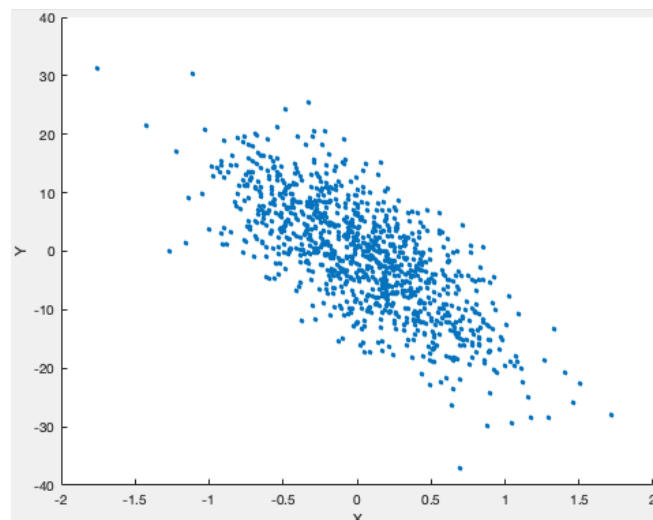
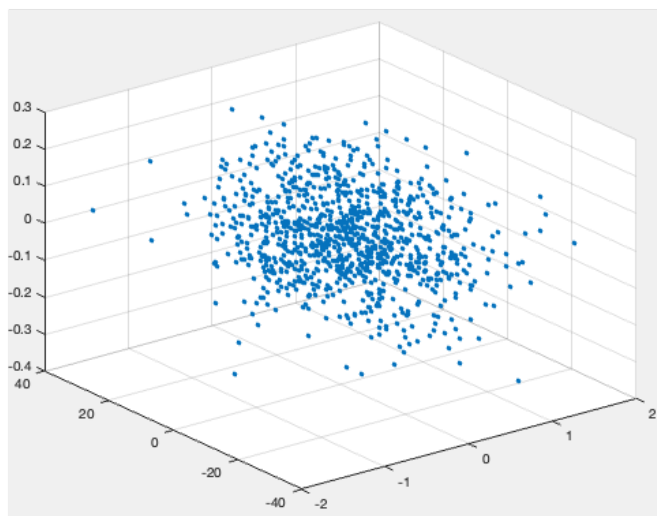
به عنوان آخرین تلاش، مقدار خطا برای حالات انتخاب دو منبع در تمامی جفت منابع را رسم کرده و همانطور که مشاهده میشود عنصر یک یک این ماتریس کمترین خطای ممکن را دارد که گویای بازسازی صرف تک منبع اول انتخاب شده SVD را نشان می دهد.



۳.

(الف)

کشیدگی و پخش سه بعدی دیتا در حالت اصلی به شرح زیر:



x_tautness =

15.1745

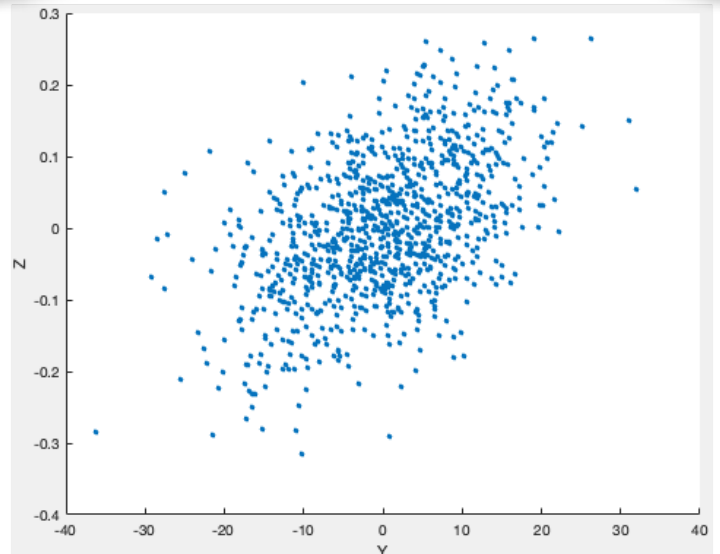
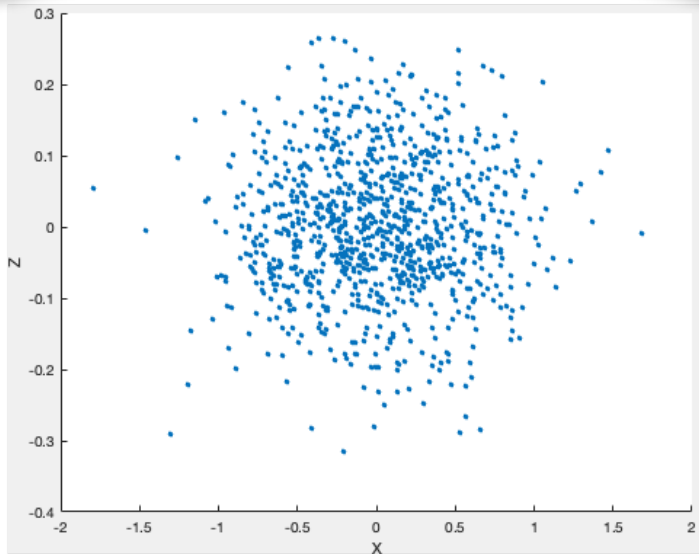
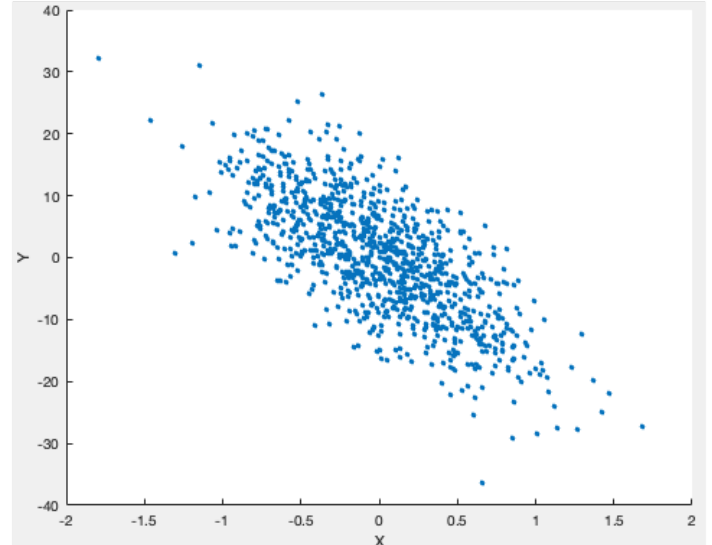
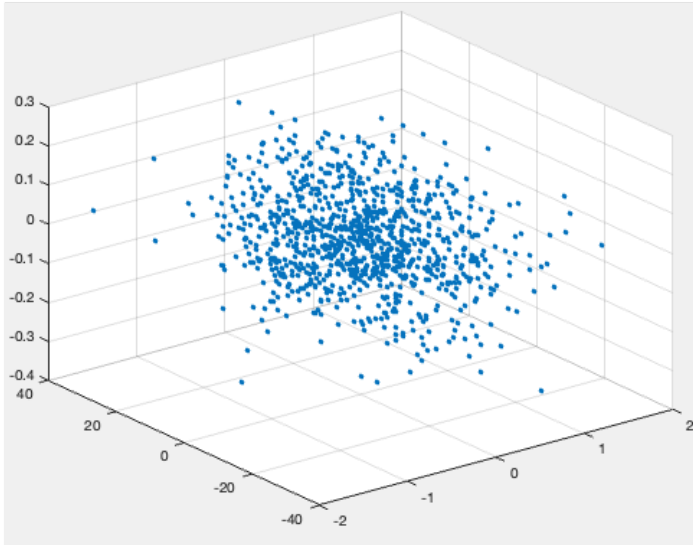
y_tautness =

304.8869

z_tautness =

3.0939

و با حذف میانگین به شرح زیر است:



x_tautness_zm =

15.1309

y_tautness_zm =

303.7244

z_tautness_zm =

3.0939

ب) با توجه به مفاهیم مطرح شده در درس تجزیه SVD را بر روی داده های میانگین صفر شده اعمال می کنیم. در نتیجه محورهای جدید ما ستون های ماتریس V هستند.

	1	2	3
1	-0.0366	-0.9842	0.1732
2	0.9993	-0.0351	0.0112
3	0.0049	-0.1734	-0.9848

مقادیر تکین همان کشیدگی در راستای محورهای جدید می باشند.

	1	2	3
1	303.9310	0	0
2	0	10.4305	0
3	0	0	2.0500

و با توجه به صفر بودن میانگین دیتا، کشیدگی به توان دو و قسم بر تعداد مشاهدات همان واریانس را به ما خواهد داد.

	1	2	3
1	92.3740	0	0
2	0	0.1088	0
3	0	0	0.0042

و در نهایت داده سفیدشده هم که همان $U \cdot S$ می شود که یک ماتریس 1000×3 می باشد.

پ) برای تبدیل PCA داریم :

جهت های جدید همان ستون های ماتریس coeff می باشد.

	1	2	3
1	-0.0366	0.9842	-0.1732
2	0.9993	0.0351	-0.0112
3	0.0049	0.1734	0.9848

که دقیقا با همان جهت های حاصله از svd همخوانی دارد.

واریانس در هر جهت به دست آمده همان مقادیر موجود در بردار latent است.

	1
1	92.4665
2	0.1089
3	0.0042

که دقیقا با مقادیر واریانس محاسبه شده در بخش قبل همخوانی دارد.

در نهایت به سادگی کشیدگی از واریانسی که اینجا محاسبه کردیم با ضرب در تعداد مشاهدات و جذر گرفتن به دست می آید.

	1
1	304.0830
2	10.4357
3	2.0510

داده سفید PCA همان ماتریس score 1000×3 می باشد که با $U \times S$ حاصل از SVD برابر است.

۴.

در ابتدا بدون در نظر گرفتن قیمت، بعد از صفر کردن میانگین داده صرف پخش شدگی و ویژگی های حاصل از توزیع داده توسط SVD برای ماتریس مقادیر تکین داریم :

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	3.9496e+03	0	0	0	0	0	0	0	0	0	0	0	0
2	0	1.7766e+03	0	0	0	0	0	0	0	0	0	0	0
3	0	0	642.8637	0	0	0	0	0	0	0	0	0	0
4	0	0	0	366.9821	0	0	0	0	0	0	0	0	0
5	0	0	0	0	158.6346	0	0	0	0	0	0	0	0
6	0	0	0	0	0	118.6498	0	0	0	0	0	0	0
7	0	0	0	0	0	0	90.0667	0	0	0	0	0	0
8	0	0	0	0	0	0	0	69.3635	0	0	0	0	0
9	0	0	0	0	0	0	0	0	40.6575	0	0	0	0
10	0	0	0	0	0	0	0	0	0	24.4208	0	0	0
11	0	0	0	0	0	0	0	0	0	0	11.3514	0	0
12	0	0	0	0	0	0	0	0	0	0	0	5.5091	0
13	0	0	0	0	0	0	0	0	0	0	0	0	1.2420

که مقادیر تکین اول به نسبت مقادیر بزرگتری داشته اما در مجموع یک روند ضریبی یکنواختی در کاهش مقادیر تکین مشاهده می شود.

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0.0293	-0.0067	-0.0115	0.0271	0.9523	-0.1691	0.2162	0.1244	0.0138	-0.0166	-0.0020	-0.0020	5.1297e-...
2	-0.0436	-0.0012	0.6329	0.7681	-0.0057	0.0468	0.0352	-0.0514	0.0235	0.0307	0.0055	1.9374e-...	1.2725e-...
3	0.0283	0.0049	-0.0883	-0.0128	-0.0571	0.3437	0.7859	-0.4908	0.0353	-0.0981	-0.0179	0.0081	-0.0042
4	-5.5585e-05	1.0268e-04	-9.7532e-04	8.5439e-04	-0.0022	-0.0035	0.0051	-0.0114	-0.0178	0.0016	-0.0159	-0.9996	-0.0080
5	4.4972e-04	-1.8207e-06	-0.0018	6.7966e-04	-1.6586e-04	0.0019	0.0039	-0.0044	-0.0139	0.0106	0.0115	-0.0078	0.9997
6	-0.0012	-3.6575e-04	0.0047	0.0057	-0.0133	-0.0727	0.0082	-0.0149	-0.0529	0.0510	-0.9942	0.0173	0.0104
7	0.0836	0.0057	-0.7559	0.6366	-0.0463	-0.1094	-0.0263	0.0176	0.0232	-0.0261	0.0055	8.9168e-...	-9.1602e-...
8	-0.0066	-3.5369e-04	0.0451	0.0026	-0.0065	0.0073	-0.0893	0.0654	0.2245	-0.9648	-0.0633	-0.0059	0.0148
9	0.0450	0.0086	0.0028	-0.0196	0.1723	-0.2094	-0.4324	-0.8544	-0.0704	-0.0389	0.0240	0.0087	-0.0025
10	0.9497	0.2924	0.0939	-0.0200	-0.0327	-0.0033	0.0025	0.0468	-0.0020	4.7869e-...	-7.7441e-...	-5.1754e-...	-7.3493e-...
11	0.0056	0.0025	-0.0112	-0.0320	0.0218	0.0410	-0.0713	-0.0600	0.9659	0.2299	-0.0430	-0.0161	0.0113
12	-0.2912	0.9562	-0.0236	0.0040	0.0170	0.0059	-0.0013	-0.0031	-0.0019	-2.1087e-...	-6.1575e-...	1.5488e-...	5.2749e-...
13	0.0229	-0.0058	-0.0928	0.0453	0.2364	0.8878	-0.3654	0.0182	-0.0769	0.0223	-0.0664	-0.0031	-9.0636e-...

مقادیر ضرایب هر کدام از ویژگی در بردارهای جهت SVD به صورت بالا است اما همانطور که مشاهده می شود ضرایب خیلی بر ربط و عملاً غیرقابل تفسیر هستند چرا که کشیدگی ها و اسکیل اعداد در جهت های مختلف باهم برابر نیست. در نتیجه داده را نرمالایز می کنیم :

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0.2510	-0.3153	0.2466	-0.0618	0.0822	-0.2197	0.7776	-0.1534	0.2604	-0.0194	-0.1096	-0.0868	0.0460
2	-0.2563	-0.3233	0.2959	-0.1287	0.3206	-0.3234	-0.2750	0.4027	0.3581	-0.2675	0.2628	0.0714	-0.0809
3	0.3467	0.1125	-0.0159	-0.0171	-0.0078	-0.0761	-0.3396	-0.1739	0.6444	0.3635	-0.3032	0.1132	-0.2511
4	0.0050	0.4548	0.2898	-0.8159	0.0865	0.1675	0.0741	0.0247	-0.0137	0.0062	0.0139	0.0040	0.0359
5	0.3429	0.2191	0.1210	0.1282	0.1369	-0.1530	-0.1996	-0.0801	-0.0185	-0.2311	0.1113	-0.8043	0.0436
6	-0.1892	0.1493	0.5940	0.2806	-0.4234	0.0593	0.0639	0.3268	0.0479	0.4314	0.0532	-0.1529	0.0456
7	0.3137	0.3120	-0.0177	0.1752	0.0167	-0.0717	0.1160	0.6008	-0.0676	-0.3628	-0.4592	0.2119	-0.0386
8	-0.3215	-0.3491	-0.0497	-0.2154	0.0986	0.0234	-0.1039	0.1218	-0.1533	0.1712	-0.6957	-0.3909	-0.0183
9	0.3198	-0.2715	0.2873	-0.1323	-0.2041	-0.1432	-0.1379	-0.0804	-0.4709	-0.0219	0.0365	0.1070	-0.6335
10	0.3385	-0.2395	0.2207	-0.1033	-0.1305	-0.1929	-0.3149	-0.0828	-0.1766	0.0352	-0.1048	0.2152	0.7202
11	0.2049	-0.3059	-0.3234	-0.2826	-0.5840	0.2732	0.0023	0.3179	0.2544	-0.1534	0.1745	-0.2096	0.0234
12	-0.2030	0.2386	-0.3001	-0.1685	-0.3456	-0.8035	0.0703	0.0049	-0.0449	0.0965	0.0193	-0.0417	-0.0045
13	0.3098	-0.0743	-0.2670	-0.0694	0.3946	-0.0532	0.0870	0.4244	-0.1952	0.6007	0.2714	-0.0552	0.0244

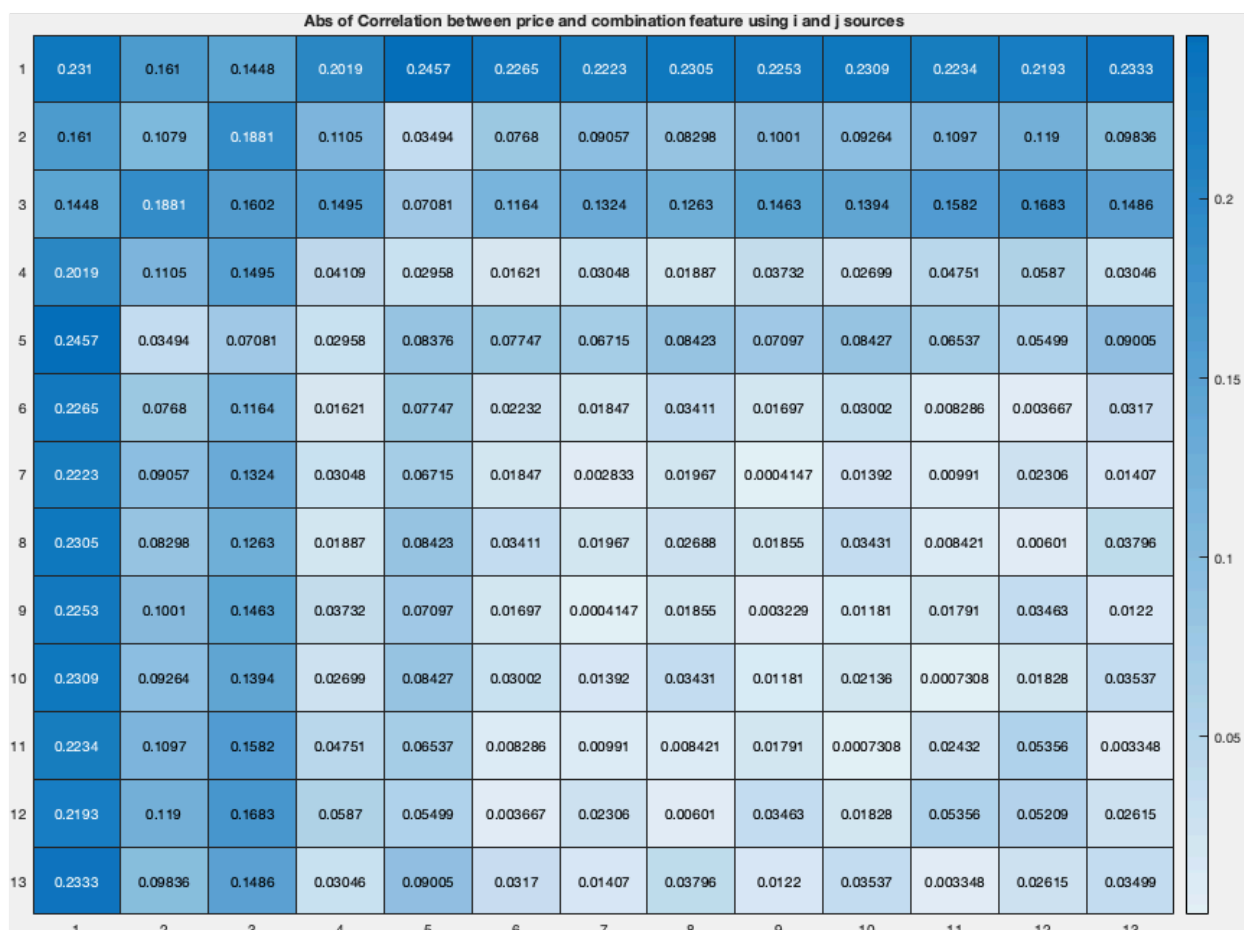
در نتیجه در بردار جهت اول که بیشترین مقدار تکین را دارد ترکیبی تقریباً با بزرگی یکسان از ویژگی های مختلف را داریم. (به جز ویژگی چهارم که احتمالاً چون ساختار ۰ و ۱ دارد به این ضریب رسیده است)

حال در ادامه می‌خواهیم ببینیم قیمت ها با بردارهای به دست آمده از SVD چه وضعیتی دارند. انتظار ما این است که با ترکیب ویژگی هایی که مقادیر تکین بیشتری داشتند همبستگی بیشتری داشته باشد. در نتیجه همبستگی داده سفید شده در فضای جدید را با مقادیر قیمت ها محاسبه میکنیم.

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	-0.2310	0.1079	0.1602	0.0411	-0.0838	-0.0223	-0.0028	-0.0269	0.0032	-0.0214	0.0243	0.0521	-0.0350

تقریباً همان روندی که انتظار داشتیم در همبستگی قیمت دیده می شود. این بدان معنا است که ویژگی هایی که ما انتخاب کردیم فارغ از بحث این که برای چه مساله ای می‌خواهیم استفاده کنیم، عملاً برای دسته بندی مساله ترکیب ویژگی مناسب می باشد که تقریباً با ضریب یکسان از تمام ویژگی های فضای اولیه ساخته می شود.

در نتیجه اگر بر فرض مثال بخواهیم دو ویژگی بهتر از فضای جدید که بیشترین همبستگی با قیمت را دارد انتخاب کنیم ماتریس بزرگی همبستگی به شکل زیر در می آید که همچنان بزرگ بودن اعداد سمت چپ تر و بالاتر این ماتریس گویای توانایی SVD در انتخاب ویژگی می باشد.



و در فضای ترکیب ویژگی ۱ و ۵ قیمت خانه به شرح زیر خواهد بود.

