

```

In [ ]: import pandas as pd
import numpy as np
import networkx as nx
import matplotlib.pyplot as plt
from scipy.sparse.linalg import svds
from scipy import sparse
import seaborn as sns
from matplotlib.axis import Axis

def GetMatrixImage(fig,fignum,vt,vt_name):
    for i in range(len(vt)-1):
        ax = fig.add_subplot(9,2,i*2+fignum)
        # Set tick font size
        for label in (ax.get_xticklabels() + ax.get_yticklabels()):
            label.set_fontsize(16)
        # Show ticks in the left and lower axes only
        Axis.set_label_coords(ax.yaxis,0, 0.5)
        Axis.set_label_coords(ax.xaxis,0.5, 0)
        ax.spines['bottom'].set_position(('data', 0))
        ax.spines['left'].set_position(('data', 0))
        g = sns.scatterplot( x="v{}".format(i+1), y="v{}".format(i+2),data=pd
            "v{}".format(i+1):vt[i],
            "v{}".format(i+2):vt[i+1]
        )),palette="Set2")
        img_title = "Spectral Subspace Plots of {0}{1} and {0}{2} ".format(vt
        g.set_title(img_title,fontsize = 20)
        ax.set_xlabel(vt_name+str(i+1),fontsize = 20)
        ax.set_ylabel(vt_name+str(i+2),fontsize = 20)
        for p in ax.patches:
            height = p.get_height()
            ax.text(p.get_x()+p.get_width()/2.,
                height + 3,
                '{:1.2f}%'.format(100*height/len(vt)),
                ha="center")

```

## 1.第一个数据 (yelp.edgelist)

```

In [ ]: filepath = './yelp.edgelist'
data = pd.read_csv(filepath, sep=' ',header = None,names = ["用户id","饭店id",'
data.head()

```

Out[ ]:

	用户id	饭店id	权重
0	0	0	1
1	0	198	1
2	0	11036	1
3	0	40573	1
4	0	81	1

```

In [ ]: G = nx.DiGraph()
G.add_weighted_edges_from([tuple(x) for x in data.values])
# G2=nx.betweenness_centrality(G,k=2)
# nx.draw_networkx(G2)
# plt.show()
print (G.number_of_edges())

```

2601677

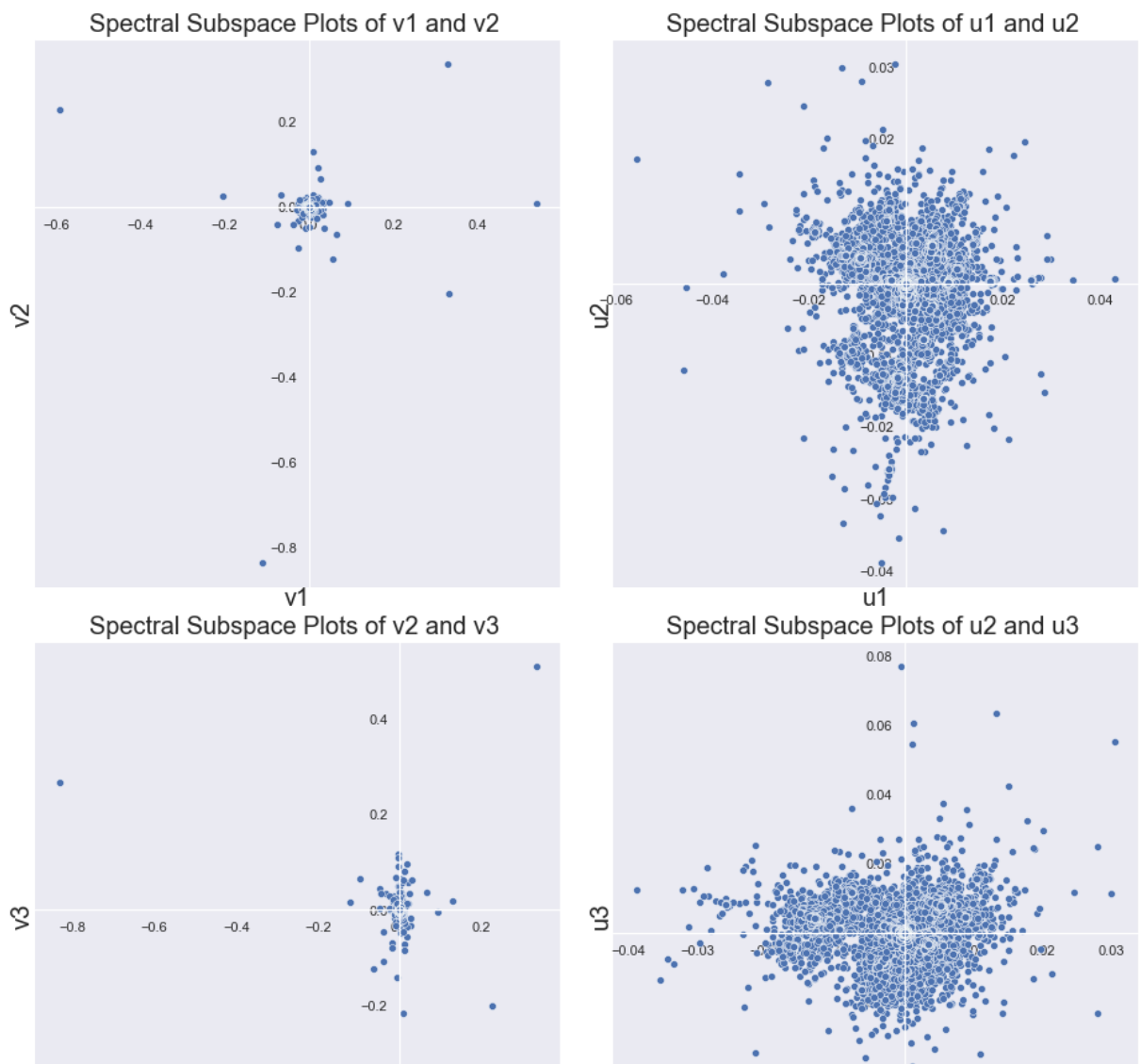
```
In [ ]: #
yelp = np.array(data)
yelp_row = yelp[:,0] # 行索引
yelp_col = yelp[:,1] # 列索引
yelp_data = yelp[:,2] # 索引对应的数值
yelp_matrix = sparse.coo_matrix((yelp_data, (yelp_row, yelp_col)), dtype=float)
print(yelp_matrix.shape)
print(type(yelp_matrix))
u, s, vt = svds(yelp_matrix, k=10, which = 'LM')

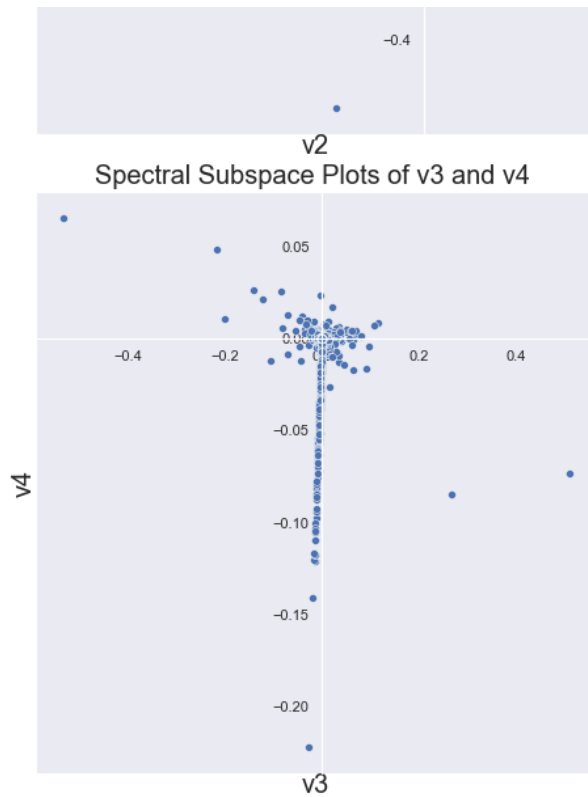
(686556, 85539)
<class 'scipy.sparse.coo.coo_matrix'>
```

```
In [ ]: # Spectral Subspace Plots
sns.set(style="dark")
fig = plt.figure(figsize=(16, 80))
fig_title = 'U and V Plot of yelp.edgelist'
# fig.suptitle(fig_title, fontsize=20)
fig.tight_layout() # 调整整体空白
plt.subplots_adjust(wspace=0.1, hspace=0.1) # 调整子图间距

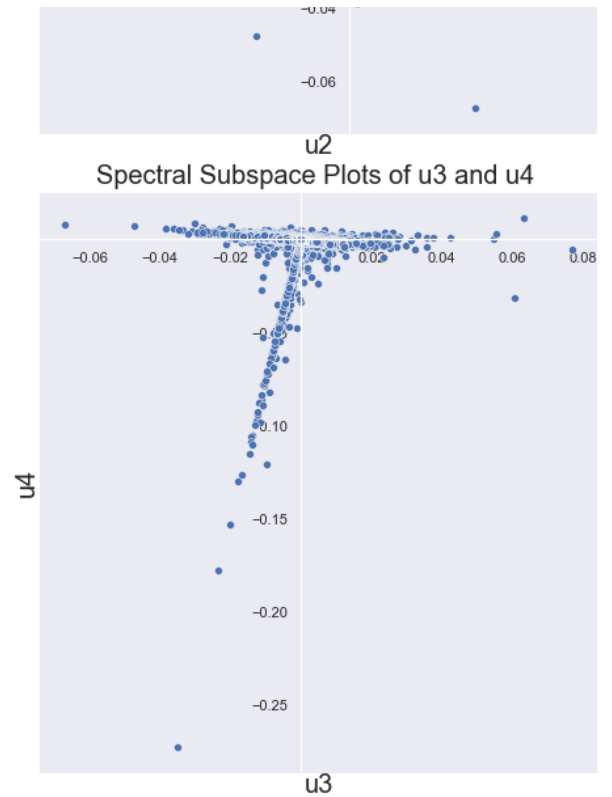
GetMatrixImage(fig, 1, vt, 'v')
GetMatrixImage(fig, 2, u.T, 'u')

plt.savefig('img/' + fig_title + '.jpg')
plt.show()
```

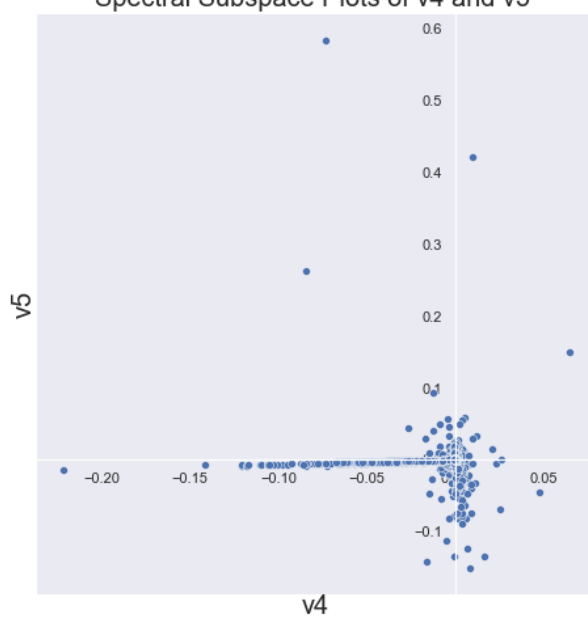




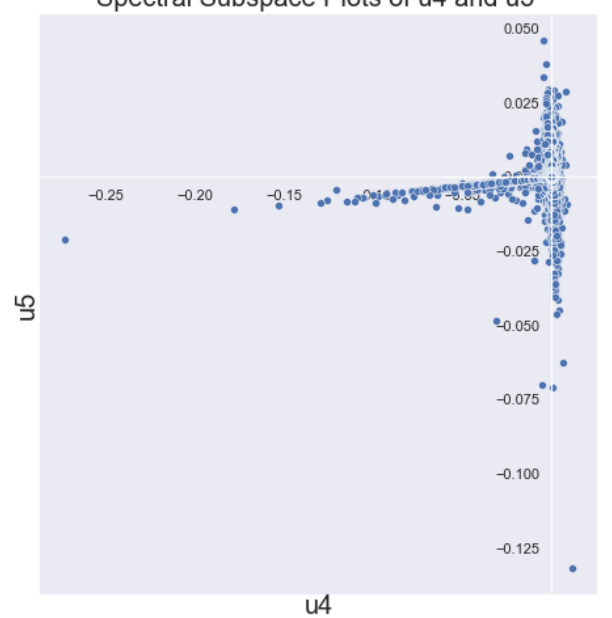
Spectral Subspace Plots of  $v_3$  and  $v_4$



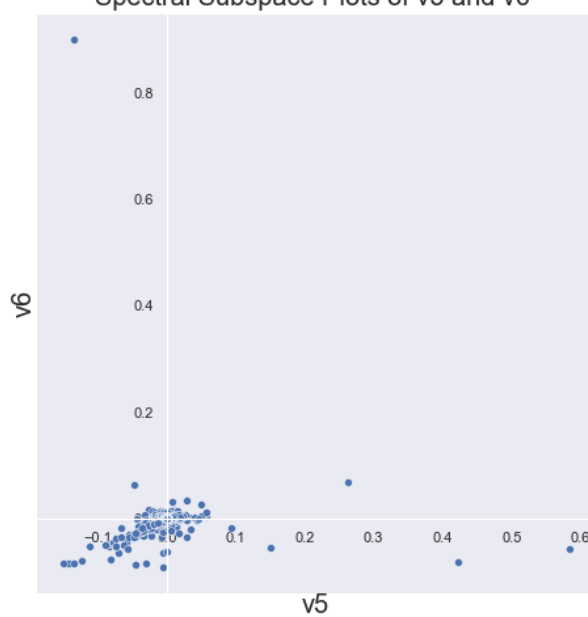
Spectral Subspace Plots of  $u_3$  and  $u_4$



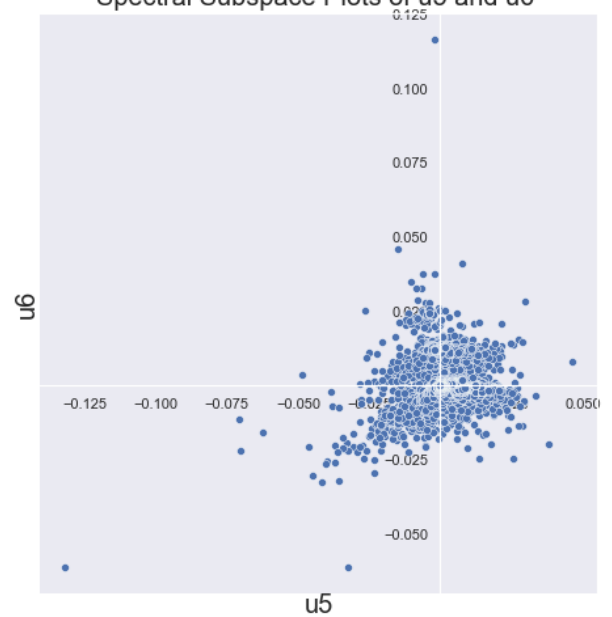
Spectral Subspace Plots of  $v_4$  and  $v_5$



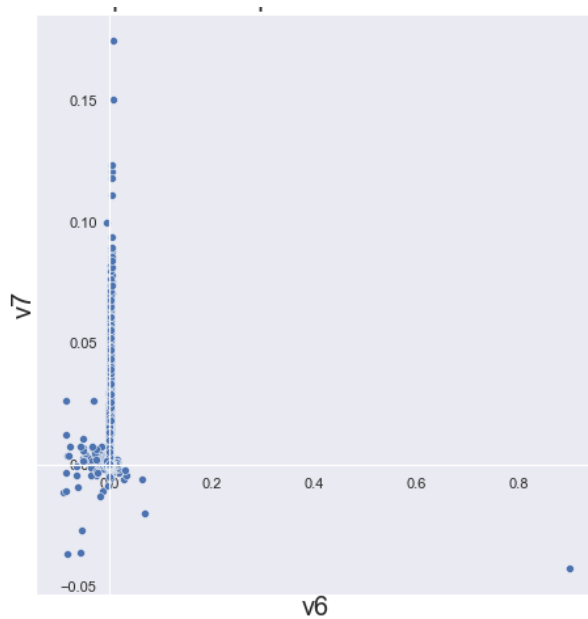
Spectral Subspace Plots of  $u_4$  and  $u_5$



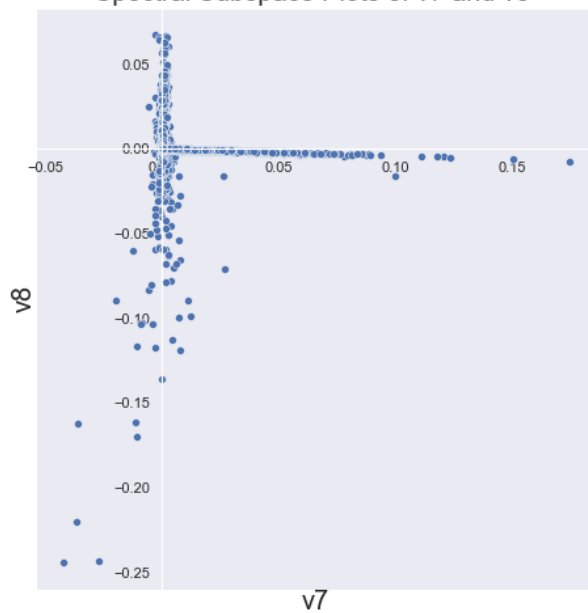
Spectral Subspace Plots of  $v_5$  and  $v_6$



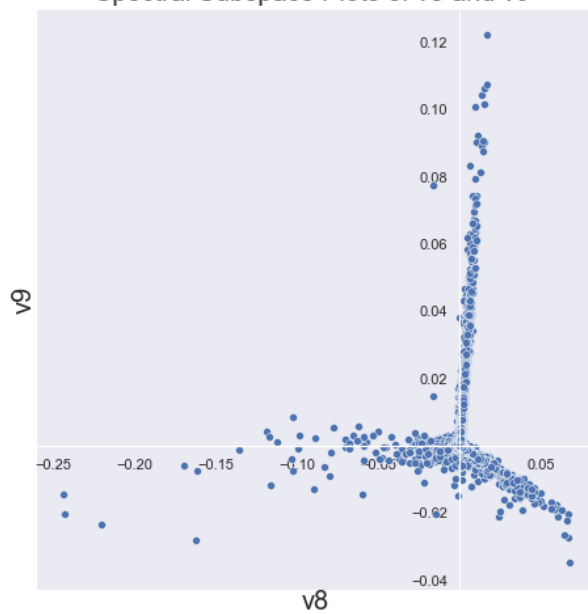
Spectral Subspace Plots of  $u_5$  and  $u_6$



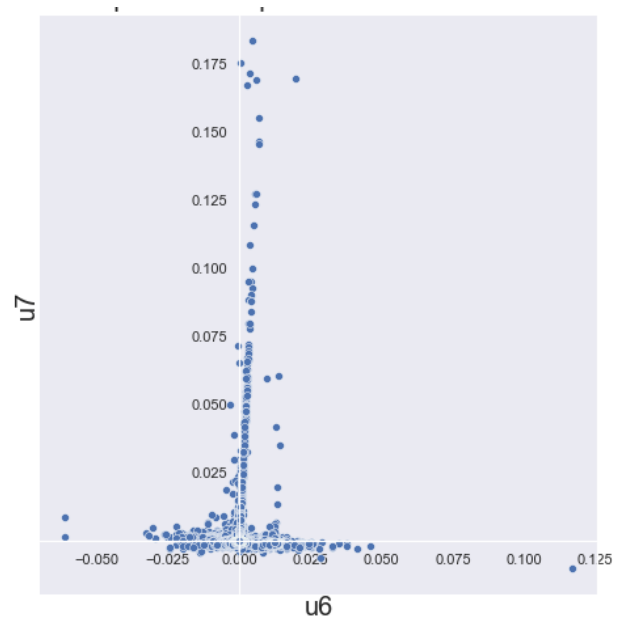
Spectral Subspace Plots of v7 and v8



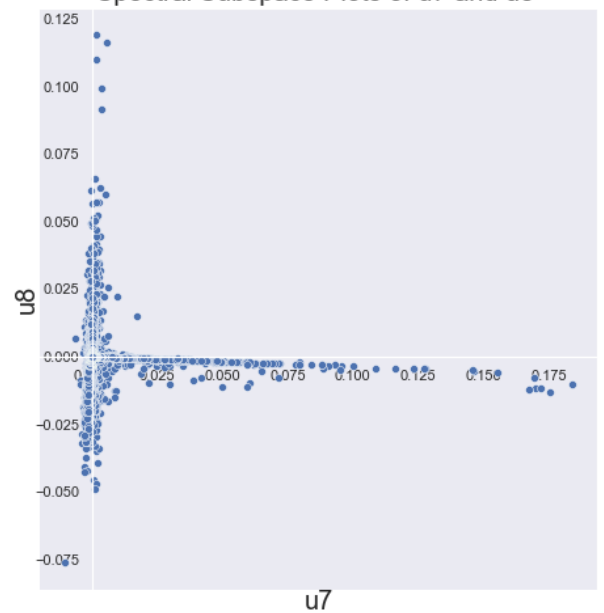
Spectral Subspace Plots of v8 and v9



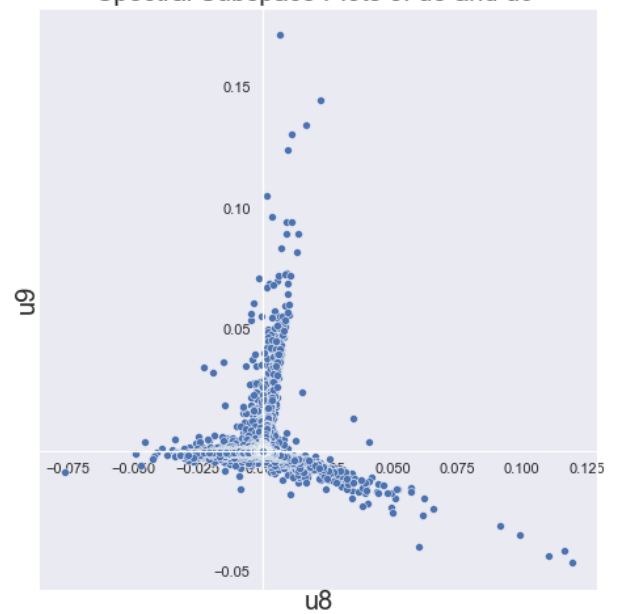
Spectral Subspace Plots of v9 and v10



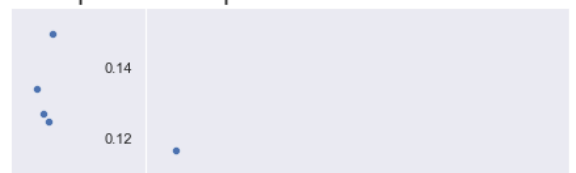
Spectral Subspace Plots of u7 and u8

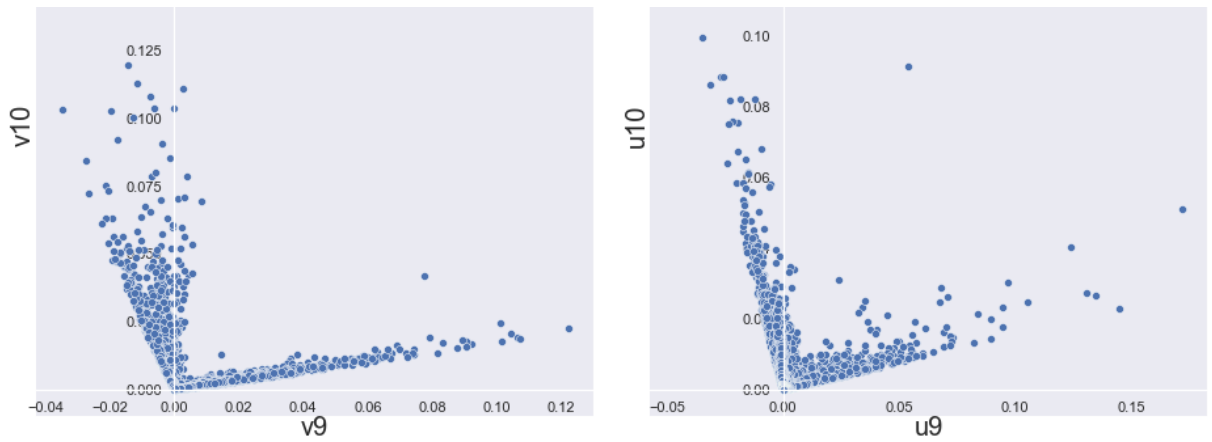


Spectral Subspace Plots of u8 and u9



Spectral Subspace Plots of u9 and u10





## 2.第二幅数据(bookcrossing)

```
In [ ]: filepath = './book.txt'
book_data = pd.read_csv(filepath, sep=';') # 制表符分隔tab
book_data.head()
```

```
Out[ ]:   User-ID  ISBN  Book-Rating
0  276725  034545104X         0.0
1  276726  0155061224         5.0
2  276727  0446520802         0.0
3  276729  052165615X         3.0
4  276729  0521795028         6.0
```

```
In [ ]: # 转换为分类数据
df = book_data.astype('category') # 可以指定特定的列转为分类数据 df['col1'] = df['
# 将标签数据转换为编码
df_code = pd.DataFrame({col: df[col].cat.codes for col in df}, index=df.index)
df_code.head()
```

```
Out[ ]:   User-ID  ISBN  Book-Rating
0  104433  57188         0
1  104434  29750         5
2  104435  107392         0
3  104436  127253         3
4  104436  127287         6
```

```
In [ ]: yelp2 = np.array(df_code)
yelp_row2 = yelp2[:,0] # 行索引
yelp_col2 = yelp2[:,1] # 列索引
yelp_matrix2 = sparse.coo_matrix((np.ones(len(yelp2)), (yelp_row2, yelp_col2)))
print(yelp_matrix2.shape)
print(type(yelp_matrix2))
u, s, vt = svds(yelp_matrix2, k=10, which = 'LM')
## Plots
# Spectral Subspace Plots
sns.set(style="dark")
fig = plt.figure(figsize=(16, 80))
```

```

fig_title = 'U and V Plot of bookcrossing'
# fig.suptitle(fig_title, fontsize=20)
fig.tight_layout()#调整整体空白
plt.subplots_adjust(wspace =0.1, hspace =0.1)#调整子图间距

GetMatrixImage(fig,1,vt,'v')
GetMatrixImage(fig,2,u.T,'u')

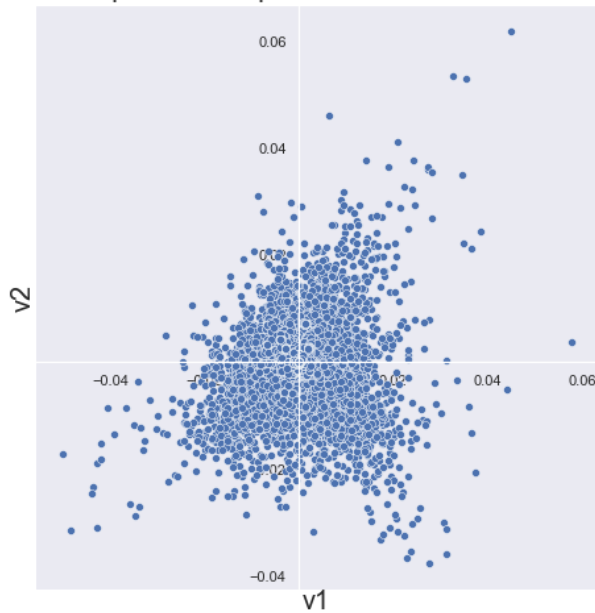
plt.savefig('img/'+ fig_title + '.jpg')
plt.show()

```

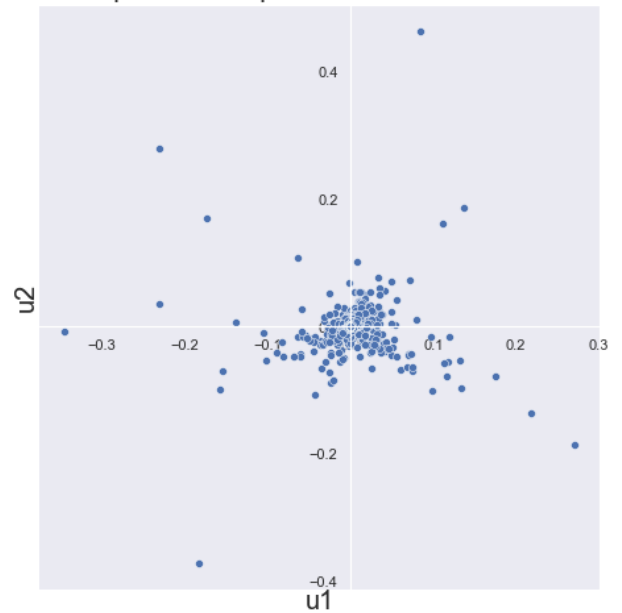
(105283, 340557)

<class 'scipy.sparse.coo.coo\_matrix'>

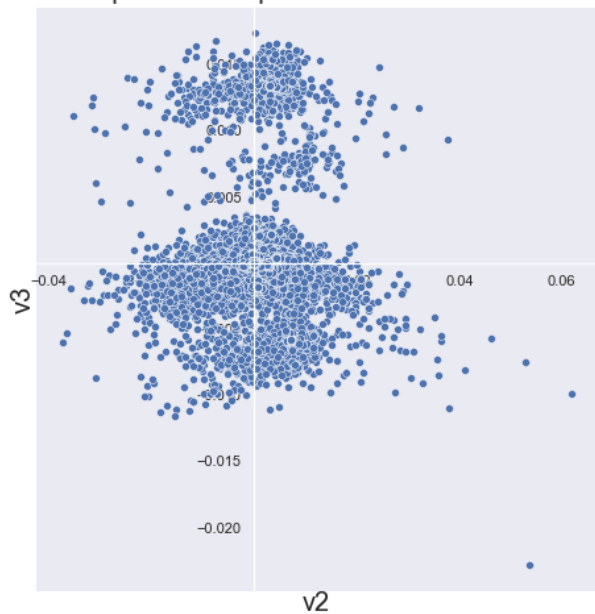
Spectral Subspace Plots of v1 and v2



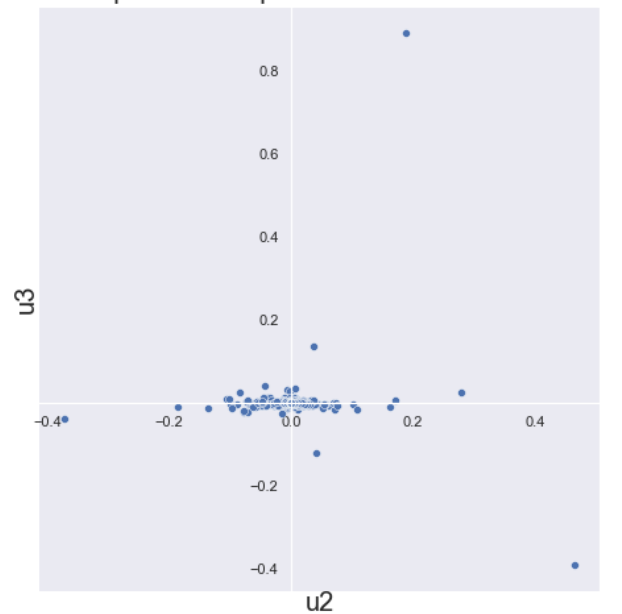
Spectral Subspace Plots of u1 and u2



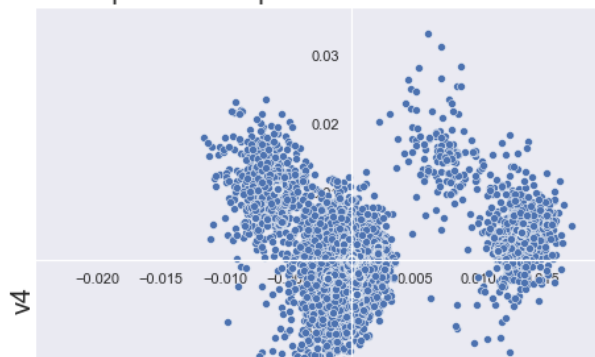
Spectral Subspace Plots of v2 and v3



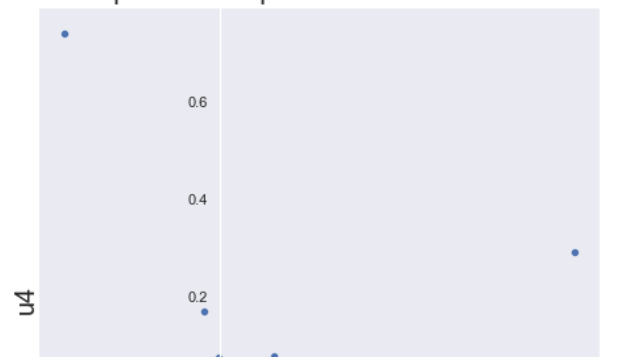
Spectral Subspace Plots of u2 and u3

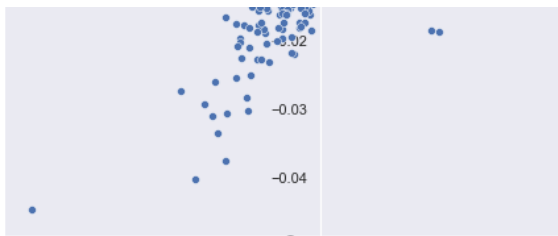


Spectral Subspace Plots of v3 and v4

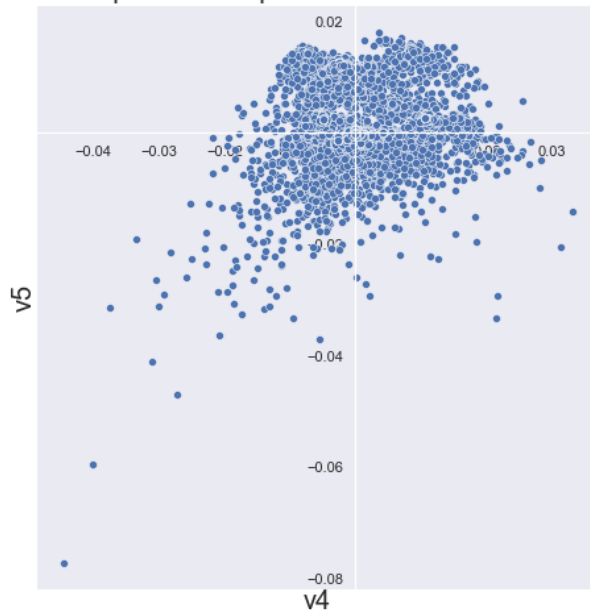


Spectral Subspace Plots of u3 and u4

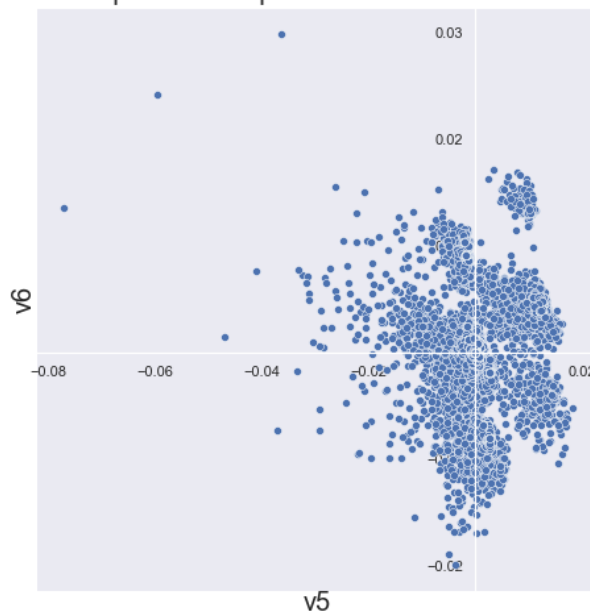




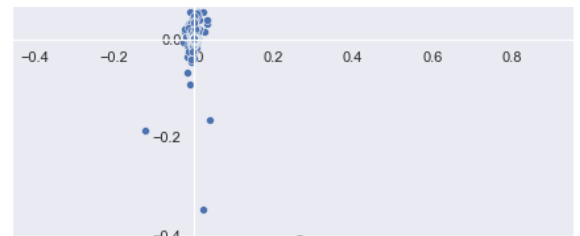
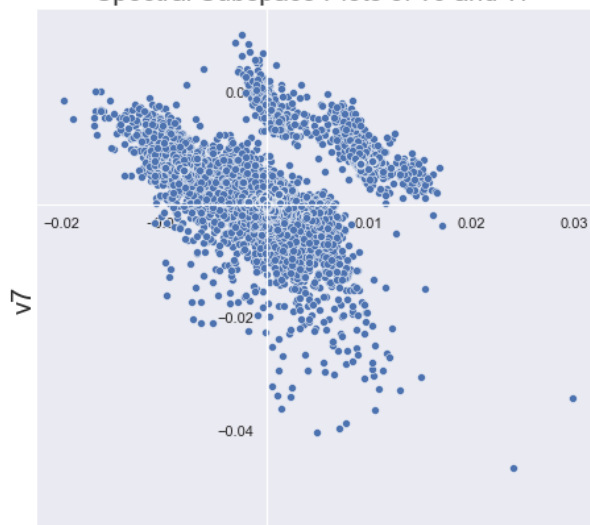
Spectral Subspace Plots of v4 and v5



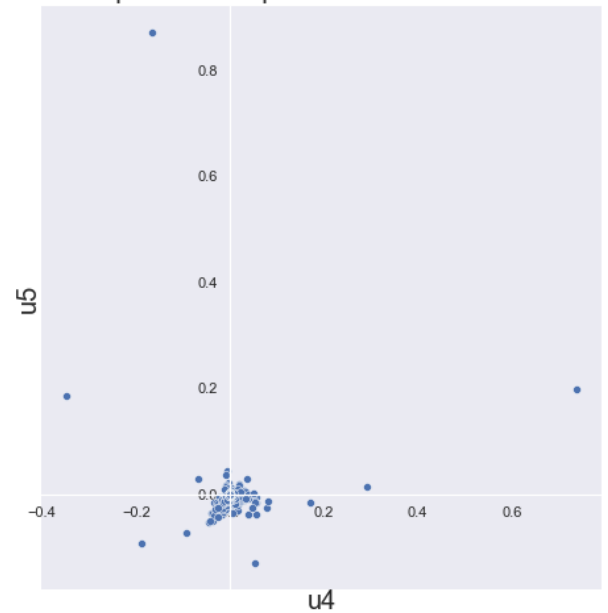
Spectral Subspace Plots of v5 and v6



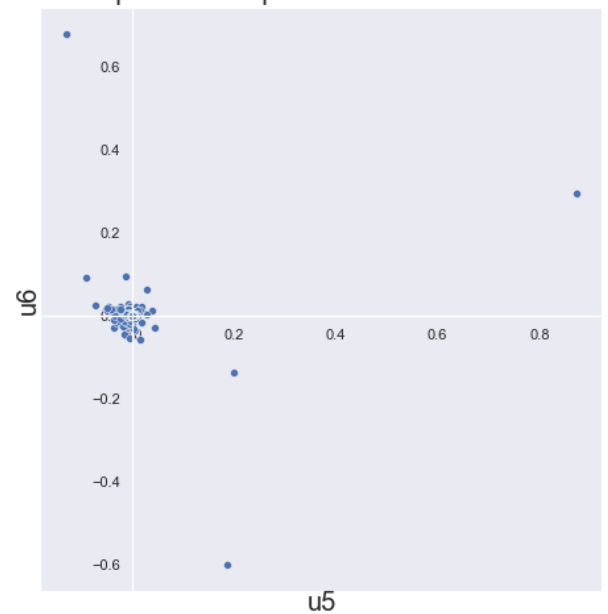
Spectral Subspace Plots of v6 and v7



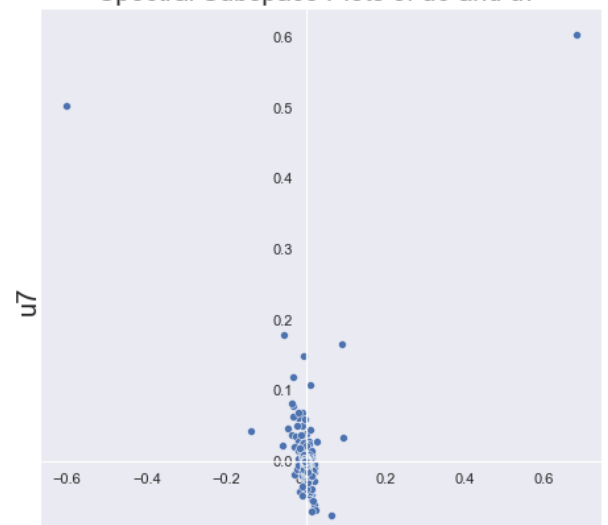
Spectral Subspace Plots of u4 and u5

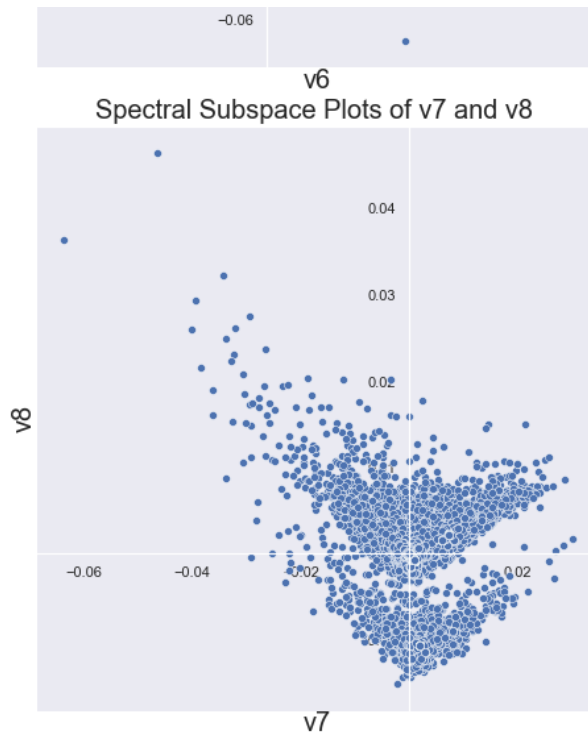


Spectral Subspace Plots of u5 and u6



Spectral Subspace Plots of u6 and u7





Spectral Subspace Plots of  $v_8$  and  $v_9$

