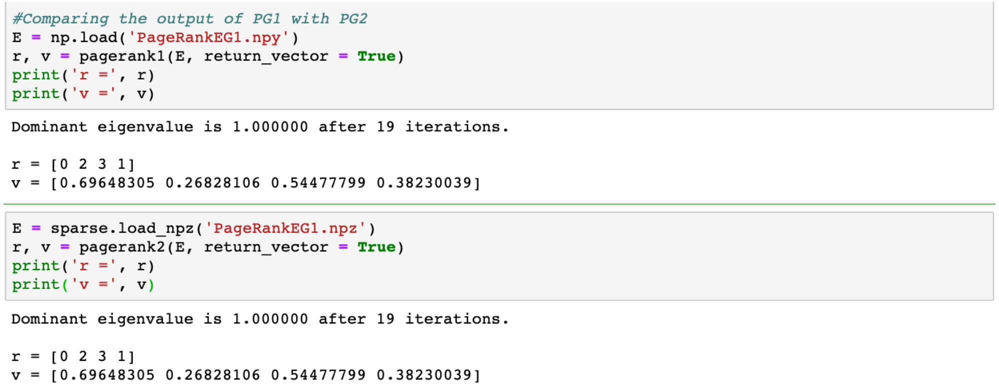
CS 111 (S19): Homework 6

Due by 6:00 PM, Wednesday, May 29

NAME and PERM ID No.: Chen Li, 5468137 (replace with yours)

UCSB EMAIL: chenli@ucsb.edu (replace with yours)

This is the comparison between PG1 and PG2 on EG1



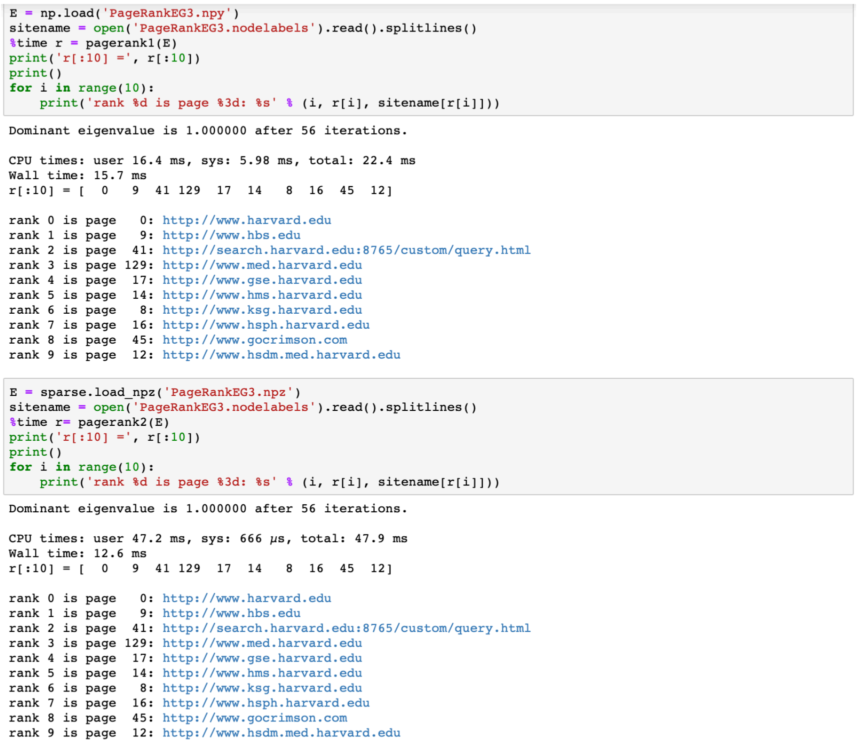
```
#Comparing the output of PG1 with PG2
E = np.load('PageRankEG1.npy')
r, v = pagerank1(E, return_vector = True)
print('r =', r)
print('v =', v)
```

```
Dominant eigenvalue is 1.000000 after 19 iterations.

r = [0 2 3 1]
v = [0.69648305 0.26828106 0.54477799 0.38230039]
```

```
E = sparse.load_npz('PageRankEG1.npz')
r, v = pagerank2(E, return_vector = True)
print('r =', r)
print('v =', v)
```

```
Dominant eigenvalue is 1.000000 after 19 iterations.

r = [0 2 3 1]
v = [0.69648305 0.26828106 0.54477799 0.38230039]
```

This is the comparison between PG1 and PG2 on EG3, on my machine, it took 15.7ms to run PG1 and took 12.6ms to run PG2. PG2 is slightly faster



```
E = np.load('PageRankEG3.npy')
sitename = open('PageRankEG3.nodelabels').read().splitlines()
%time r = pagerank1(E)
print('r[:10] =', r[:10])
print()
for i in range(10):
    print('rank %d is page %3d: %s' % (i, r[i], sitename[r[i]]))
```

```
Dominant eigenvalue is 1.000000 after 56 iterations.

CPU times: user 16.4 ms, sys: 5.98 ms, total: 22.4 ms
Wall time: 15.7 ms
r[:10] = [  0   9  41 129  17  14   8  16  45  12]

rank 0 is page   0: http://www.harvard.edu
rank 1 is page   9: http://www.hbs.edu
rank 2 is page  41: http://search.harvard.edu:8765/custom/query.html
rank 3 is page 129: http://www.med.harvard.edu
rank 4 is page  17: http://www.gse.harvard.edu
rank 5 is page  14: http://www.hms.harvard.edu
rank 6 is page   8: http://www.ksg.harvard.edu
rank 7 is page  16: http://www.hsph.harvard.edu
rank 8 is page  45: http://www.gocrimson.com
rank 9 is page  12: http://www.hsdm.med.harvard.edu
```

```
E = sparse.load_npz('PageRankEG3.npz')
sitename = open('PageRankEG3.nodelabels').read().splitlines()
%time r= pagerank2(E)
print('r[:10] =', r[:10])
print()
for i in range(10):
    print('rank %d is page %3d: %s' % (i, r[i], sitename[r[i]]))
```

```
Dominant eigenvalue is 1.000000 after 56 iterations.

CPU times: user 47.2 ms, sys: 666 µs, total: 47.9 ms
Wall time: 12.6 ms
r[:10] = [  0   9  41 129  17  14   8  16  45  12]

rank 0 is page   0: http://www.harvard.edu
rank 1 is page   9: http://www.hbs.edu
rank 2 is page  41: http://search.harvard.edu:8765/custom/query.html
rank 3 is page 129: http://www.med.harvard.edu
rank 4 is page  17: http://www.gse.harvard.edu
rank 5 is page  14: http://www.hms.harvard.edu
rank 6 is page   8: http://www.ksg.harvard.edu
rank 7 is page  16: http://www.hsph.harvard.edu
rank 8 is page  45: http://www.gocrimson.com
rank 9 is page  12: http://www.hsdm.med.harvard.edu
```

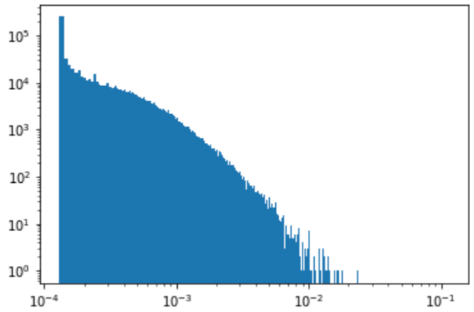
It took 10.4 second to run PG2 and only 5.42 second to run spla.eigs. max is 0.11427415903139658 and min is 0.00013008094286810808

```
#webGoogle
E = sparse.load_npz('webGoogle.npz')
```

```
%time r, v = pagerank2(E, return_vector = True)
```

```
Dominant eigenvalue is 1.000000 after 71 iterations.

CPU times: user 41.4 s, sys: 92.4 ms, total: 41.5 s
Wall time: 10.4 s
```

```
%matplotlib inline
plt.hist(v,bins="auto")
plt.gca().set_yscale("log")
plt.gca().set_xscale("log")
```



```
print("maxvalue:",v[r[0]])
print("minvalue:",v[r[-1]])
```

```
maxvalue: 0.11427415903139658
minvalue: 0.00013008094286810808
```

```
%time eigVal, eigVec = spla.eigs(E)
```

```
CPU times: user 20.6 s, sys: 203 ms, total: 20.8 s
Wall time: 5.42 s
```