Homework 5

Li Donghao 20653877

Problem 1

1.a

The code are follows:

```
function [success] = problem1(p,r)
    R = randn(20,20);
    [U,S,V] = svds(R,10);

    A = U(:,1:r)*V(:,1:r)';

    E0 = rand(20);
    E = 1*abs(E0>(1-p));
    X = A + E;

    lambda = 0.25;
    cvx_begin
    variable L(20,20);
    variable S(20,20);
    variable W1(20,20);
    variable W2(20,20);
    variable Y(40,40) symmetric;
    Y == semidefinite(40);
    minimize(.5*trace(W1)+0.5*trace(W2)+lambda*sum(sum(abs(S))));
    subject to
    L + S >= X-1e-5;
    L + S <= X + 1e-5; Y == [W1, L';L W2];
    cvx_end
    % The difference between sparse solution S and E
    disp('$\-S-E\- \infty$:') ;
    norm(S-E,'inf');
    % The difference between the low rank solution L and A
    disp('\-A-L\-') ;
    norm(A-L);
    success=norm(S-E,'inf')<1e-4;
end
```

1.b

Changing p from 0.1 to 0.5, I find the success rete will decrease. When p=0.1 succes rate=0.7, When p=0.15 succes rate=0.3, When p>0.15 succes rate=0.

1.c

Changing r from 1 to 5, I find the success rete will decrease. When r=1 succes rate=0.7, When r=2 succes rate=0.2, When r>2 succes rate=0.

Problem 2

2.a

The covariance matrix is shown in the code:

```
V9=290*0.09+300*0.925*0.925+2;
V9V5=0.925*300;
V9V1=-0.3*290;
True_sigma=[
    291 , 290 , 290 , 290 , 0   , 0  , 0  , 0   , V9V1 , V9V1 ,
    290 , 291 , 290 , 290 , 0   , 0  , 0  , 0   , V9V1 , V9V1 ,
    290 , 290 , 291 , 290 , 0   , 0  , 0  , 0   , V9V1 , V9V1 ,
    290 , 290 , 290 , 291 , 0   , 0  , 0  , 0   , V9V1 , V9V1 ,
    0   , 0   , 0   , 0   , 301 , 300, 300, 300 , V9V5 , V9V5 ,
    0   , 0   , 0   , 0   , 300 , 301, 300, 300 , V9V5 , V9V5 ,
    0   , 0   , 0   , 0   , 300 , 300, 301, 300 , V9V5 , V9V5 ,
    0   , 0   , 0   , 0   , 300 , 300, 300, 301 , V9V5 , V9V5 ,
    V9V1, V9V1, V9V1, V9V1, V9V5,V9V5,V9V5, V9V5,V9    , V9-1 ,
    V9V1, V9V1, V9V1, V9V1, V9V5,V9V5,V9V5, V9V5,V9-1  , V9   ,
];
```

I have compared it with sample covariance matrix with sample size 1000 and found the difference is quite large. $||\Sigma - \hat{\Sigma}||_F = 51.4$. This large gap may due to number of sample is small. If we increase the number of sample to 100000, $||\Sigma - \hat{\Sigma}||_F = 3.27$.

2.b

I computed largetest four eigenvectors using matlab.

2.c

Code for sovling the SDP problem is:

```
function X=sPCA_(sigma,lambda)
R = sigma;
d = 10;
e = ones(d,1);
cvx_begin
variable X(d,d) symmetric;
X == semidefinite(d);
minimize(-trace(R*X)+lambda*(e'*abs(X)*e));
subject to
    trace(X)==1;
cvx_end
end
```

Comparing with PCA, when $\lambda = 0$, $||X_{PCA} - X_{SPCA}||_{inf} = 8.8620e-12$ which means the results are similiar. When $\lambda > 0$, $||X_{PCA} - X_{SPCA}||_{inf} = 7.4320e-4$ the results are slighly different. So the comclusion is that when $\lambda = 0$, Sparse PCA is the same with PCA. Which means my implimentation is correct.

2.d,e

Comparing the 1st 2nd 3rd and 4th principal components, I find the first two are similiar between $\lambda = 0$ and $\lambda > 0$. For the first pricipal component, $||X_{PCA} - X_{SPCA}||_{inf} = 7.4320e-4$. And $||X_{PCA} - X_{SPCA}||_{inf} = 0.0015$ for the second component. but the 3rd and 4th principal components are quite different. ($||X_{PCA} - X_{SPCA}||_{inf} = 0.2826$ and $||X_{PCA} - X_{SPCA}||_{inf} = 1.7568$ respectively) That may due to the Sparse PCA could handle sparse problems better and give accurate esimation.

In [ ]: