CS/ECE/ME532 Assignment 5

1. Here we continue the problem studied in Activity 11. Let a 4-by-2 matrix $\boldsymbol{X}$ have SVD $\boldsymbol{X} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^T$ where $\boldsymbol{U} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix}$, $\boldsymbol{S} = \begin{bmatrix} 1 & 0 \\ 0 & \gamma \end{bmatrix}$, and $\boldsymbol{V} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$

Let $\boldsymbol{y} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$.

a) The ratio of the largest to the smallest singular values is termed the condition number of $\boldsymbol{X}$. Find the condition number if $\gamma = 0.1$, and $\gamma = 10^{-8}$. Solve $\boldsymbol{X}\boldsymbol{w} = \boldsymbol{y}$ for $\boldsymbol{w}$ and find $||\boldsymbol{w}||_2^2$ for these two values of γ.

b) A system of linear equations with a large condition number is said to be "ill-conditioned". One consequence of an ill-conditioned system of equations is solutions with large norms as you found in the previous part of this problem. A second consequence is that the solution is very sensitive to small errors in $\boldsymbol{y}$ such as may result from measurement error or numerical error. Suppose $\boldsymbol{y} = \begin{bmatrix} 1+\epsilon \\ 0 \\ 0 \\ 1 \end{bmatrix}$. Write $\boldsymbol{w} = \boldsymbol{w}_o + \boldsymbol{w}_\epsilon$ where $\boldsymbol{w}_o$ is the solution for arbitrary γ when $\epsilon = 0$ and $\boldsymbol{w}_\epsilon$ is the perturbation in that solution due to some error $\epsilon \neq 0$. How does the norm of the perturbation due to $\epsilon \neq 0$, $||\boldsymbol{w}_\epsilon||_2^2$, depend on the condition number? Find $||\boldsymbol{w}_\epsilon||_2^2$ for $\epsilon = 0.01$ and $\gamma = 0.1$ and $\gamma = 10^{-8}$.

c) Now consider a "low-rank" inverse. Instead of writing

$$(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T = \sum_{i=1}^{p} \frac{1}{\sigma_i}\boldsymbol{v}_i\boldsymbol{u}_i^T$$

where p is the number of columns of $\boldsymbol{X}$ (assumed less than the number of rows), we approximate

$$(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T \approx \sum_{i=1}^{r} \frac{1}{\sigma_i}\boldsymbol{v}_i\boldsymbol{u}_i^T$$

In this approximation we only invert the largest r singular values, and ignore all of them smaller than σ_r. Use $r = 1$ in the low-rank inverse to find $\boldsymbol{w} = \boldsymbol{w}_o + \boldsymbol{w}_\epsilon$

where $\boldsymbol{y} = \begin{bmatrix} 1+\epsilon \\ 0 \\ 0 \\ 1 \end{bmatrix}$ as in part b). Compare the results to part b).

①

a) $w = \frac{1}{\sqrt{2}} \begin{bmatrix} 1+\gamma^{-1} \\ 1-\gamma^{-1} \end{bmatrix}$ from Activity 11

$\|w\|_2^2 = \frac{1}{2}\left[(1+\gamma^{-1})^2 + (1-\gamma^{-1})^2\right] = 1+\gamma^{-2}$

$\underline{\gamma = 0.1}$

$$w = \frac{1}{\sqrt{2}} \begin{bmatrix} 1+10 \\ 1-10 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 11 \\ -9 \end{bmatrix}$$

* Condition number = 10

$$\|w\|_2^2 = \left(\frac{11}{\sqrt{2}}\right)^2 + \left(\frac{-9}{\sqrt{2}}\right)^2 = \frac{121}{2} + \frac{81}{2} = \frac{202}{2} = 101$$

$\underline{\gamma = 10^{-8}}$

$$w = \frac{1}{\sqrt{2}} \begin{bmatrix} 1+10^{8} \\ 1-10^{8} \end{bmatrix}$$

* Condition number = 10^8

$$\|w\|_2^2 = \frac{(1+10^8)^2 + (1-10^8)^2}{2} = 1 + (10^{-8})^{-2} \approx 10^{16}$$

b) $y = \begin{bmatrix} 1+\varepsilon \\ 0 \\ 0 \\ 1 \end{bmatrix} \Rightarrow w_0 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1+\gamma^{-1} \\ 1-\gamma^{-1} \end{bmatrix}$

$$w_\varepsilon = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1/\gamma \end{bmatrix} \begin{bmatrix} \varepsilon/2 \\ \varepsilon/2 \end{bmatrix} = \frac{\varepsilon}{2 w_0}$$

Then, $\|w_\varepsilon\|_2^2 = \frac{\varepsilon^2}{4\|w_0\|_2^2}$ increases as the condition number increases.

$$w = w_0 + w_\varepsilon \Rightarrow w = V S^{-1} U^T (y_0 + y_\varepsilon)$$

$$\Rightarrow w_\varepsilon = V S^{-1} U^T y_\varepsilon = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \frac{1}{1-\gamma} \begin{bmatrix} \gamma & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1+\varepsilon \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\Rightarrow w_\varepsilon = \sum_{i=1}^{2} \frac{1}{\sigma_i} v_i (u_i^T y_\varepsilon) = \frac{1}{2\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1+\varepsilon \\ 0 \\ 0 \\ 1 \end{bmatrix} + \frac{1}{2\gamma\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \begin{bmatrix} 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1+\varepsilon \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\Rightarrow w_\varepsilon = \frac{1}{2\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} (2+\varepsilon) + \frac{1}{2\gamma\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} (2+\varepsilon)$$

$$\Rightarrow w_\varepsilon = \frac{2+\varepsilon}{2\sqrt{2}} \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} \gamma^{-1} \\ -\gamma^{-1} \end{bmatrix} \right) = \frac{2+\varepsilon}{2\sqrt{2}} \begin{bmatrix} 1+\gamma^{-1} \\ 1-\gamma^{-1} \end{bmatrix}$$

$\underline{\varepsilon = 0.01}$

$\underline{\gamma = 0.1}$

$$\|w_\varepsilon\|_2^2 = \frac{2+0.01}{8}(121) + \frac{2.001}{8}(81) = \frac{101(2.01)}{4} \approx 50.53$$

$$\|W_\varepsilon\|_2^2 = \frac{0.0101}{4}$$

$\underline{\gamma = 10^{-8}}$

$$\|w_\varepsilon\|_2^2 = \frac{2.001}{2\sqrt{2}}(1+10^8)^2 + \frac{2.001}{2\sqrt{2}}(1-10^8)$$

$$\|W_\varepsilon\|_2^2 \approx \frac{10^{12}}{4}$$

c)

in b) $\quad w = \frac{1}{\sqrt{2}}\begin{bmatrix} 1+\gamma^{-1} \\ 1-\gamma^{-1} \end{bmatrix} + \frac{2+\varepsilon}{2\sqrt{2}}\begin{bmatrix} 1+\gamma^{-1} \\ 1-\gamma^{-1} \end{bmatrix}$

with rank-1 approximation: (next page)

$$X\_1 = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \\ 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix} \Rightarrow w^T = x y^T = [1+\varepsilon \quad 1+\varepsilon]$$

Then

$$w = w_0 + w_\varepsilon = [1 \quad 1] + [\varepsilon \quad \varepsilon]$$

$$w_0 = \sum_{i=1}^{r} \frac{1}{\sigma_i} v_i u_i^T y = \frac{1}{\sigma_1} v_1 u_1^T y = \frac{1}{2\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

and

$$w_\varepsilon = \sum_{i=1}^{r} \frac{1}{\sigma_1} v_i u_i^T y_\varepsilon = \frac{1}{\sigma_1} v_1 u_1^T y_\varepsilon = \frac{1}{2\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} \varepsilon \\ 0 \\ 0 \\ 0 \end{bmatrix} = \frac{\varepsilon}{2\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Hence, in the low-rank case, $\|W_\varepsilon\|_2$ is proportional to ε and does not blow up as it does with the full-rank inverse. This case is much less sensitive to the error ε.

Assignment_5

March 30, 2022

[ ]:
```
# c)
import numpy as np
import matplotlib.pyplot as plt
```

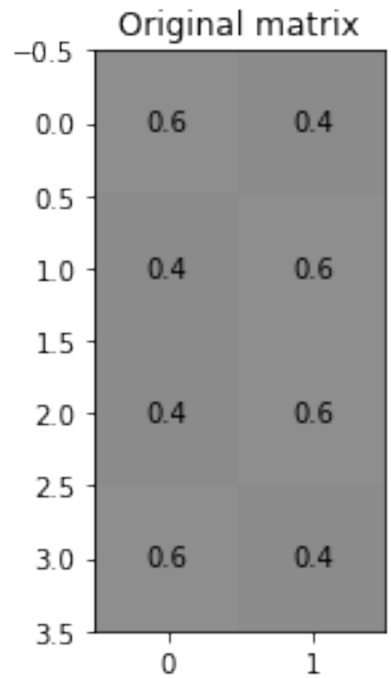
[ ]:
```
gamma = 0.1
U = (1/2)*np.array([[1,1],[1,-1],[1,-1],[1,1]])
S_matrix = np.array([[1,0],[0,gamma]])
V = (1/((2)*1/2))*np.array([[1,1],[1,-1]])
VT = V.transpose()
A = U@S_matrix@VT
```

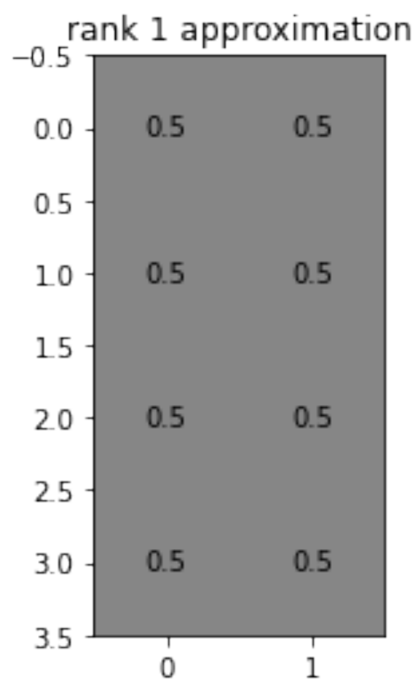
[ ]:
```
# display images of various rank approximations

# original matrix
plt.figure(num=None)
for (j,i),label in np.ndenumerate(A):
    plt.text(i,j,np.round(label,1),ha='center',va='center')
plt.imshow(A, vmin=-5, vmax=5, interpolation='none', cmap='gray')
plt.title('Original matrix' )

# rank-1 aprox
A_rank_r_approx = S_matrix[0,0]*U[:,0:1]@VT[0:1,:]
plt.figure(num=None)
for (j,i),label in np.ndenumerate(A_rank_r_approx):
    plt.text(i,j,np.round(label,1),ha='center',va='center')
plt.imshow(A_rank_r_approx, vmin=-10, vmax=10, interpolation='none',␣
 ↪cmap='gray')
plt.title('rank ' + str(1) + ' approximation'  )
```

[ ]:
```
array([[0.5, 0.5],
       [0.5, 0.5],
       [0.5, 0.5],
       [0.5, 0.5]])
```


Original matrix
−0.5
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
0.6
0.4
0.4
0.6
0.4
0.6
0.6
0.4
0
1


rank 1 approximation
−0.5
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
0.5
0.5
0.5
0.5
0.5
0.5
0.5
0.5
0
1

[ ]:
```python
print(U[:,0:1]@VT[0:1,:])
print(U[:,0:1])
print(VT[0:1,:])
```

```
[[0.5 0.5]
 [0.5 0.5]
 [0.5 0.5]
 [0.5 0.5]]
[[0.5]
 [0.5]
 [0.5]
 [0.5]]
[[1. 1.]]
```