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).