MATH 076 - Computational Inverse Problems

Carter Kruse, July 12, 2023

Homework 1

Instructions

This homework assignment is split into two sections: written questions and computational questions. For the written questions, you are **not** permitted to use a calculator or any other computational tools. For the computational questions, you are asked to use a coding language of your choice to perform the tasks requested. You may make use any code posted to our course's Canvas page. **Please show all of your work.** If you have any questions or uncertainties, please reach out to your instructor.

Written Questions

1) Find the SVD and compute the condition number of

$$A = \begin{bmatrix} -1.5 & 2.5 \\ 2.5 & -1.5 \end{bmatrix}$$

Let us define the singular value decomposition (SVD). Assuming $m \geq n$, for $A \in \mathbb{R}^{m \times n}$, the SVD of A is

$$A = U \Sigma V^T = \sum_{i=1}^{n} u_i \sigma_i v_i^T$$

Σ is an $n \times n$ diagonal matrix with the *singular values* $\sigma_1, \ldots, \sigma_n$:

$$\Sigma = \text{diag}\left(\sigma_1, \ldots, \sigma_n\right), \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$$

Matrices $U \in \mathbb{R}^{m \times n}$ and $V \in \mathbb{R}^{n \times n}$ consist of the left and right *singular vectors*:

$$U = [u_1, u_2, \ldots, u_n] \qquad V = [v_1, v_2, \ldots, v_n]$$

The matrices have orthonormal columns, i.e. $u_i^T u_j = v_i^T v_i = \delta_{i,j}$, or simply $U^T U = V^T V = \mathbb{I}_n$, where

$$\delta_{i,j} \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

To find the SVD, let us use linear algebra. The eigenvectors of $A^T A$ make up the columns of V. The eigenvectors of AA^T make up the columns of U. The square root of eigenvalues of either $A^T A$ or AA^T make up Σ.

In this case, we have the following:

$$A^T A = AA^T = \begin{bmatrix} -1.5 & 2.5 \\ 2.5 & -1.5 \end{bmatrix} \begin{bmatrix} -1.5 & 2.5 \\ 2.5 & -1.5 \end{bmatrix}$$

$$A^T A = AA^T = \begin{bmatrix} (-1.5)^2 + (2.5)^2 & (-1.5)(2.5) + (2.5)(-1.5) \\ (2.5)(-1.5) + (-1.5)(2.5) & (2.5)^2 + (-1.5)^2 \end{bmatrix}$$

$$A^TA = AA^T = \begin{bmatrix} 8.5 & -7.5 \\ -7.5 & 8.5 \end{bmatrix}$$

To compute the eigenvalues, we determine the values of λ for which $\det(A - \lambda I) = 0$, as follows:

$$(8.5 - \lambda)^2 - (-7.5)^2 = 0$$

$$(8.5)^2 - 17\lambda + \lambda^2 - (7.5)^2 = 0$$

$$\lambda^2 - 17\lambda + 16 = 0$$

When we factor to determine the appropriate value of λ, as $(\lambda - 16)(\lambda - 1) = 0$, we find that the eigenvalues of the matrix $A^TA = AA^T$ are $\lambda = 1, 16$.

According to the SVD, this means that Σ (the $n \times n$ diagonal matrix) with singular values $\sigma_1, \ldots, \sigma_n$ is given as follows:

$$\Sigma = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}$$

Now, to determine the eigenvectors, we aim to solve the expression $Av = \lambda v$, which is equivalent to solving $v(A - \lambda I) = 0$.

$$\begin{bmatrix} 8.5 - \lambda & -7.5 \\ -7.5 & 8.5 - \lambda \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Using the value $\lambda = 1$, this produces the following system of equations:

$$7.5x - 7.5y = 0$$

$$-7.5x + 7.5y = 0$$

So, we may consider a vector $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ which satisfies the above, which, normalized, is $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$.

Using the value $\lambda = 16$, this produces the following system of equations:

$$-7.5x - 7.5y = 0$$

$$-7.5x - 7.5y = 0$$

So, we may consider a vector $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ which satisfies the above, which, normalized, is $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$

These eigenvectors correspond to the columns of U and V. The following is the solution, which provides the SVD of the matrix A:

$$A = U\Sigma V^T = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}^T$$

The condition number is defined as

$$\text{cond}(A) = \left|\left|A^{-1}\right|\right|_2 ||A||_2 = \frac{\sigma_1}{\sigma_n}$$

Thus, to compute the condition number, we simply take the ratio between the greatest singular value and the smallest, as follows

$$\text{cond}(A) = \frac{\sigma_1}{\sigma_n} = 4$$

2) The solution x_{LS} to the linear least squares problem $\min_x ||Ax - b||_2$ is formally given as

$$x_{LS} = \left(A^T A\right)^{-1} A^T b$$

under the assumption that $A \in \mathbb{R}^{m\times n}$ has more rows than columns and $A^T A$ is full rank. Use this expression together with the SVD to show that x_{LS} has the *same* SVD expansion as when A is square and non-singular, i.e. prove

$$x_{LS} = \sum_{i=1}^{n} \frac{u_i^T b}{\sigma_i} v_i$$

Let us define the singular value decomposition (SVD). Assuming $m \geq n$, for $A \in \mathbb{R}^{m\times n}$, the SVD of A is

$$A = U\Sigma V^T = \sum_{i=1}^{n} u_i \sigma_i v_i^T$$

Σ is an $n \times n$ diagonal matrix with the *singular values* $\sigma_1, \ldots, \sigma_n$:

$$\Sigma = \operatorname{diag}\left(\sigma_1, \ldots, \sigma_n\right), \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$$

Matrices $U \in \mathbb{R}^{m\times n}$ and $V \in \mathbb{R}^{n\times n}$ consist of the left and right *singular vectors*:

$$U = [u_1, u_2, \ldots, u_n] \qquad V = [v_1, v_2, \ldots, v_n]$$

The matrices have orthonormal columns, i.e. $u_i^T u_j = v_i^T v_i = \delta_{i,j}$, or simply $U^T U = V^T V = \mathbb{I}_n$, where

$$\delta_{i,j} \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Assumptions

$$\begin{aligned} U^T U = UU^T = I \qquad & \left(U^T\right)^{-1} = U \\ V^T V = VV^T = I \qquad & \left(V^T\right)^{-1} = V \\ \Sigma^T \Sigma = \Sigma\Sigma^T = \Sigma^2 \end{aligned}$$

Using the SVD with the solution x_{LS} to the linear least squares problem $\min_x ||Ax - b||_2$, under the assumption that $A \in \mathbb{R}^{m\times n}$ has more rows than columns and $A^T A$ is full rank produces the following:

$$\begin{aligned} x_{LS} &= \left(A^T A\right)^{-1} A^T b \\ &= \left(\left(U\Sigma V^T\right)^T \left(U\Sigma V^T\right)\right)^{-1} \left(U\Sigma V^T\right)^T b \\ &= \left(\left(V\Sigma^T U^T\right)\left(U\Sigma V^T\right)\right)^{-1} \left(U\Sigma V^T\right)^T b \\ &= \left(V\Sigma^T \Sigma V^T\right)^{-1} \left(U\Sigma V^T\right)^T b \\ &= \left(V\Sigma^2 V^T\right)^{-1} \left(U\Sigma V^T\right)^T b \\ &= \left(V^T\right)^{-1} \left(\Sigma^{-2}\right) \left(V\right)^{-1} \left(U\Sigma V^T\right)^T b \\ &= \left(V^T\right)^{-1} \left(\Sigma^{-2}\right) \left(V\right)^{-1} \left(V\Sigma^T U^T\right) b \\ &= \left(\left(V^T\right)^{-1} \left(\Sigma^{-2}\right) \Sigma^T U^T\right) b \\ &= \left(\left(V^T\right)^{-1} \left(\Sigma^{-1}\right) U^T\right) b \\ &= \left(V\Sigma^{-1} U^T\right) b \end{aligned}$$

From this, we deduce that x_{LS} has the *same* SVD expansion as when A is square and non-singular. This is done by simply using the expansion of the expression above. $\square$

$$x_{LS} = \sum_{i=1}^{n} \frac{u_i^T b}{\sigma_i} v_i$$

3) When solving inverse problems, it is common to want to enforce some kind of *sparsity* condition. This means that we want our solution vector $x \in \mathbb{R}^n$ to have relatively few non-zero values. We will define the 0-"norm" $||x||_0$ to be

$$||x||_0 = \# \text{ non-zero values of } x$$

For example, if $v = [0, 27, 0]^T$, then $||v||_0 = 1$. While this would seem to be the most natural way to discuss sparsity, the 0-"norm" as defined is not a real norm. Why is this the case?

The definition of a *vector* norm is given as $||x||$ with $\mathbb{R}^n \to \mathbb{R}$ with the following properties:

- $||x|| \geq 0$
- $||x|| = 0 \implies x = 0$
- $||\alpha x|| = |\alpha| \, ||x||$ with $\alpha \in \mathbb{R}$
- $||x + y|| \leq ||x|| + ||y||$ with $y \in \mathbb{R}^n$

Thus, to be defined as a real norm, any potential proposal must satisfy all of the properties. For a solution vector $x \in \mathbb{R}^n$, the 0-"norm" defined as $||x||_0 = \#$ non-zero values of x satisfies the following properties:

- $||x|| \geq 0$
- $||x|| = 0 \implies x = 0$
- $||x + y|| \leq ||x|| + ||y||$ with $y \in \mathbb{R}^n$

However, it does *not* satisfy the property $||\alpha x|| = |\alpha| \, ||x||$ with $\alpha \in \mathbb{R}$, as demonstrated by the following simple counterexample.

Suppose $v = [0, 1]^T \in \mathbb{R}^2$. Given the definition of the 0-"norm", we have $||v|| = 1$. Let $\alpha = 4$. Then, we have $\alpha v = [0, 4]^T \in \mathbb{R}^2$. The following equality does not hold:

$$||\alpha v||_0 \neq |\alpha| \, ||v||$$

$$\left(\left|\left| [0, 4]^T \right|\right|_0 = 1 \right) \neq \left(4 = |4| \left|\left| [0, 1]^T \right|\right|_0 \right)$$

Thus, the 0-"norm" does not satisfy the necessary properties to be considered a real norm, even though it may be the most natural way to discuss sparsity.

4) Suppose $x \in \mathbb{R}^n$. Prove the following.

2-Norm (Euclidean Norm)

$$||x||_2 = \left(\sum_{i=1}^{n} x_i^2 \right)^{\frac{1}{2}} = \sqrt{\overline{x}^T \overline{x}}$$

∞-Norm (Max Norm)

$$||x||_\infty = \max_{i=1,\ldots,n} |x_i|$$

1-Norm

$$||x||_1 = \sum_{i=1}^{n} |x_i|$$

(A) $||x||_\infty \leq ||x||_2$

In the case where $x \in \mathbb{R}$, $||x||_\infty = ||x||_2$. This is due to the fact that the vector x is a singular value.

$$||x||_\infty = \max_{i=1,\ldots,n} |x_i| = |x_1|$$

$$||x||_2 = \left(\sum_{i=1}^{n} x_i^2 \right)^{\frac{1}{2}} = \left((x_1)^2 \right)^{\frac{1}{2}} = |x_1|$$

Thus, $||x||_2 = ||x||_\infty$.

In the case where $x \in \mathbb{R}^n$, where $n > 1$, let us consider the value of i for which $|x_i|$ is the maximum of all possible values, say j.

$$||x||_\infty = \max_{i=1,\ldots,n} |x_i| = |x_j|$$

To determine the value of $||x||_2$, select x_j to remove from the summation, as follows:

$$||x||_2 = \left(\sum_{i=1}^{n} x_i^2 \right)^{\frac{1}{2}} = \left(x_j^2 + \sum_{i=1, i \neq j}^{n} x_i^2 \right)^{\frac{1}{2}}$$

We may consider the sum as a residual term $\varepsilon > 0$, as $x_i^2 > 0$ for all i.

$$||x||_2 = \left(\sum_{i=1}^{n} x_i^2 \right)^{\frac{1}{2}} = \left(x_j^2 + \varepsilon \right)^{\frac{1}{2}}$$

Thus, we know the following:

$$||x||_2 = \left(\sum_{i=1}^{n} x_i^2 \right)^{\frac{1}{2}} \geq |x_j^2|$$

Thus, $||x||_\infty \leq ||x||_2$. $\square$

(B) $||x||_2 \leq ||x||_1$

Given that $||x||_2 \geq 0$ and $||x||_1 \geq 0$, we may prove the statement via the following:

$$||x||_2^2 \leq ||x||_1^2$$

The definitions for $||x||_2$ and $||x||_1$ may be used as follows:

$$||x||_2^2 = \sum_{i=1}^{n} x_i^2 \qquad ||x||_1^2 = \left(\sum_{i=1}^{n} |x_i|\right)^2$$

To demonstrate the inequality $||x||_2^2 \leq ||x||_1^2$, we may show the following:

$$\sum_{i=1}^{n} x_i^2 \leq \left(\sum_{i=1}^{n} x_i^2 + 2\sum_{i,j,i<j} |x_i|\,|x_j|\right) = \left(\sum_{i=1}^{n} |x_i|\right)^2$$

Thus, $||x||_2 \leq ||x||_1$. $\square$

(C) $||x||_1 \leq n\,||x||_\infty$

Let us consider the value of i for which $|x_i|$ is the maximum of all possible values, say j.

$$||x||_\infty = \max_{i=1,\ldots,n} |x_i| = |x_j|$$

Thus, we know $n\,||x||_\infty = n\,|x_j|$.

Further, we have the following:

$$||x||_1 = \sum_{i=1}^{n} |x_i|$$

As $|x_j|$ was selected to be the maximum element (in absolute terms) in the vector x, there are no values $|x_k|$ for which $|x_j| < |x_k|$. While there may be values for which $|x_k| = |x_j|$, there is no exceeding this limit. Thus, we have the following bound:

$$\sum_{i=1}^{n} |x_i| \leq n\,|x_j|$$

(The upper bound assumes that the values $|x_i|$ are equivalent for all $i \in 0, \ldots n$.)

Thus, $||x||_1 \leq n\,||x||_\infty$. $\square$

(D) $||x||_2 \leq \sqrt{n}\,||x||_\infty$

Given that $||x||_2 \geq 0$ and $||x||_\infty \geq 0$, we may prove the statement via the following:

$$||x||_2^2 \leq n\,||x||_\infty^2$$

The definitions for $||x||_2$ and $||x||_1$ may be used as follows:

$$||x||_2^2 = \sum_{i=1}^{n} x_i^2 \qquad ||x||_\infty^2 = \left(\max_{i=1,\ldots,n} |x_i|\right)^2$$

To demonstrate the inequality $||x||_2^2 \leq n\,||x||_\infty^2$, we may show the following:

$$\sum_{i=1}^{n} x_i^2 \leq n\left(\max_{i=1,\ldots,n} x_i^2\right) = n\left(\max_{i=1,\ldots,n} |x_i|\right)^2$$

(This follows similarly to the previous argument, with x_i^2 substituted for $|x_i|$.)

Thus, $||x||_2 \leq \sqrt{n}\,||x||_\infty$. $\square$

Computational Questions

View *Matlab* Code