Math 2940 Worksheet

Matrix Operations and Inverses

Week 5

September 26th, 2019

This worksheet covers material from **Section 2.1-2.3**. Please work in collaboration with your classmates to complete the following exercises - this means sharing ideas and asking each other questions.

Question 1. Assuming that all the relevant multiplications exist, which of the following are always true? If not, can you find a counterexample?

(a) $A(B+C) = AB + AC$

(b) $AB = BA$

(c) $(AB)^T = B^T A^T$

(d) $A^n B = A^k (A^{n-k} B)$

(e) If $AB = AC$, then $B = C$

(f) If $AB = 0$, then $A = 0$ or $B = 0$.

Question 2. Suppose that A is a 4×4 matrix, and $\boldsymbol{x}$ is a vector in $\mathbb{R}^4$. There are two possible ways of computing $A^2\boldsymbol{x}$:

- $(AA)\boldsymbol{x}$
- $A(A\boldsymbol{x})$

Which method is faster? (You can just count the number of multiplications and ignore additions)

Question 3. Which of the following matrices are or are not invertible? Try to use as few calculations as possible, but still justify your answers. (You do not need to actually find the inverses.)

(a) $\begin{bmatrix} 5 & 7 \\ -3 & -6 \end{bmatrix}$

(b) $\begin{bmatrix} -7 & 0 & 4 \\ 3 & 0 & -1 \\ 2 & 0 & 9 \end{bmatrix}$

(c) $\begin{bmatrix} 5 & 0 & 0 \\ -3 & -7 & 0 \\ 8 & 5 & -1 \end{bmatrix}$

(d) $\begin{bmatrix} -4 & 6 \\ 6 & -9 \end{bmatrix}$

Question 4. Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 5 \\ 0 & 1 & 2 \end{bmatrix}$

(a) Compute A^{-1}

(b) Use your answer from part (a) to solve the system of equations $A\boldsymbol{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

(c) Which of the following do you think is the better way of solving a linear system $A\boldsymbol{x} = \boldsymbol{b}$?

- Computing the inverse like you just did in parts (a) and (b)
- Row reducing an augmented matrix $[A \,|\, \boldsymbol{b}]$

Question 5. Remember the question in last week's worksheet on perturbation. In the last comment, it said that the shear transformation matrix is ill-conditioned. But what do 'ill-conditioned' and 'well-conditioned' mean? We'll explore the idea of the *condition number* of a matrix in the following problems.

(a) To define conditioning, we first need to define matrix *norms*. For now, you can consider a norm as an operator $||\cdot|| : \mathbb{R}^{n\times n} \to \mathbb{R}$. That means, the operator takes in an $n \times n$ matrix, and returns a single real number. Intuitively, you can think of norms as describing the "length" of a matrix.

There are actually many possible norms we can use for matrices, some commonly used ones are the (1-norm, ∞-norm, Frobenius norm):

$$||A||_1 = \max_{1\leq j\leq n} \sum_{i=1}^{m} |A_{i,j}|, \quad ||A||_\infty = \max_{1\leq i\leq n} \sum_{j=1}^{m} |A_{i,j}|, \quad ||A||_F = \left(\sum_{i=1}^{n} \sum_{j=1}^{n} A_{i,j}^2 \right)^{1/2}.$$

Using the definition, determine the norms of the following matrices:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 3 \\ 2 & -1 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

(b) A condition number represents how susceptible a matrix is to roundoff errors. The condition number of a matrix A is defined by $\kappa(A) = ||A|| \cdot ||A^{-1}||$. Using a norm of your choice, find the condition number of the following matrix, where $0 < \varepsilon < 1$:

$$A = \begin{bmatrix} 1 & 0 \\ 1 & \varepsilon \end{bmatrix}.$$

Explain what happens to the condition number when ε varies.

(c) Suppose we want to reduce a matrix $A = \begin{bmatrix} 1 & 3 \\ 5 & 6 \end{bmatrix}$ into echelon form. By now, we know how to do this with a row replacement matrix R_1, or a Givens rotation matrix R_2. Find the condition number of these two matrices, using the Frobenius norm.

(*Hint*: for the Givens rotation matrix, you don't need to calculate the angle of rotation to compute its Frobenius norm.)

(d) If other 2×2 matrices are used in place of A, are your condition numbers from part (c) going to change?

Answer to Question 1.

(a) $\boxed{\text{True}}$, the distributive property holds for matrix multiplication.

(b) $\boxed{\text{False}}$, for a counterexample, let $A = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$, and $B = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$. Then,

$$AB = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

but

$$BA = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$$

(c) $\boxed{\text{True}}$, the (i, j)-th entry of $(AB)^T$ would be given by multiplying the j-th row of A with the i-th column of B. The same thing would be true for finding the (i, j)-th entry of $B^T A^T$.

(d) $\boxed{\text{True}}$, matrix multiplication is associative, so

$$A^k(A^{n-k}B) = (A^k A^{n-k})B = A^n B$$

(e) $\boxed{\text{False}}$, for a counterexample, let $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$, and $C = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. Then,

$$AB = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

and

$$AC = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

but clearly $B \neq C$.

(f) $\boxed{\text{False}}$, for a counterexample see A and B from the part above.

Answer to Question 2. In order to compute $(AA)\boldsymbol{x}$, we first have to do the matrix multiplication AA:

$$AA = \begin{bmatrix} \square & \square & \square & \square \\ \square & \square & \square & \square \\ \square & \square & \square & \square \\ \square & \square & \square & \square \end{bmatrix} \begin{bmatrix} \square & \square & \square & \square \\ \square & \square & \square & \square \\ \square & \square & \square & \square \\ \square & \square & \square & \square \end{bmatrix} = \begin{bmatrix} \square & \square & \square & \square \\ \square & \square & \square & \square \\ \square & \square & \square & \square \\ \square & \square & \square & \square \end{bmatrix}$$

The resulting matrix has $4^2 = 16$ entries, and each entry requires 4 multiplications, requiring 64 multiplications in total.
Once we have AA, we just compute the matrix-vector product:

$$\begin{bmatrix} \square & \square & \square & \square \\ \square & \square & \square & \square \\ \square & \square & \square & \square \\ \square & \square & \square & \square \end{bmatrix} \begin{bmatrix} \square \\ \square \\ \square \\ \square \end{bmatrix} = \begin{bmatrix} \square \\ \square \\ \square \\ \square \end{bmatrix}$$

and with 4 entries that require 4 multiplications each, this takes 16 multiplications. Putting this together,

$$\boxed{\text{Computing } (AA)\boldsymbol{x} \text{ requires 80 multiplications.}}$$

However, to compute $A(A\boldsymbol{x})$, we first have to compute the matrix-vector product $A\boldsymbol{x}$:

$$\begin{bmatrix} \square & \square & \square & \square \\ \square & \square & \square & \square \\ \square & \square & \square & \square \\ \square & \square & \square & \square \end{bmatrix} \begin{bmatrix} \square \\ \square \\ \square \\ \square \end{bmatrix} = \begin{bmatrix} \square \\ \square \\ \square \\ \square \end{bmatrix}$$

which requires 16 multiplications. Once we have the vector $A\boldsymbol{x}$, we do another matrix-vector product $A(A\boldsymbol{x})$, which again takes 16 multiplications, so all together

$$\boxed{\text{Computing } A(A\boldsymbol{x}) \text{ requires 32 multiplications.}}$$

which is clearly faster than computing $(AA)\boldsymbol{x}$.

Answer to Question 3.

(a) The matrix $\begin{bmatrix} 5 & 7 \\ -3 & -6 \end{bmatrix}$ is $\boxed{\text{invertible}}$, since the columns are linearly independent, which we can tell because there are only two of them and they are not multiples of each other.

(b) The matrix $\begin{bmatrix} -7 & 0 & 4 \\ 3 & 0 & -1 \\ 2 & 0 & 9 \end{bmatrix}$ is $\boxed{\text{not invertible}}$, since there is no way of row reducing this matrix to the identity, because of the column of all zeros.

(c) The matrix $\begin{bmatrix} 5 & 0 & 0 \\ -3 & -7 & 0 \\ 8 & 5 & -1 \end{bmatrix}$ is $\boxed{\text{invertible}}$, since it can be row reduced to the identity matrix. Because the matrix is lower triangular, we see that we can use the entries on the main diagonal as pivots.

(d) The matrix $\begin{bmatrix} -4 & 6 \\ 6 & -9 \end{bmatrix}$ is $\boxed{\text{not invertible}}$, because the first column is a multiple of the second, so the columns are not linearly independent.

Answer to Question 4.

(a) To compute A^{-1}, we want to row reduce the following augmented matrix:

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 3 & 5 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 & 0 & 1 \end{array}\right]$$

until the left hand side becomes the identity matrix. Using the top left corner as a pivot,

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & 1 & 2 & 0 & 0 & 1 \end{array}\right] \begin{array}{l} \\ \leftarrow \text{Row 2 } + (-2) \cdot (\text{ Row 1}) \\ \\ \end{array}$$

Swapping rows two and three,

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 & 1 \\ 0 & -1 & -1 & -2 & 1 & 0 \end{array}\right] \begin{array}{l} \\ \leftarrow \text{Row } 3 \\ \leftarrow \text{Row } 2 \end{array}$$

Using the middle entry as a pivot,

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 & 1 \\ 0 & 0 & 1 & -2 & 1 & 1 \end{array}\right] \leftarrow \text{Row } 3 \ + \ \text{Row } 2$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & 0 & -2 \\ 0 & 1 & 2 & 0 & 0 & 1 \\ 0 & 0 & 1 & -2 & 1 & 1 \end{array}\right] \begin{array}{l} \leftarrow \text{Row } 1 \ + (-2) \cdot (\text{ Row } 2\ ) \\ \\ \\ \end{array}$$

Then using the bottom right as a pivot,

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 1 & -1 \\ 0 & 1 & 0 & 4 & -2 & -1 \\ 0 & 0 & 1 & -2 & 1 & 1 \end{array}\right] \begin{array}{l} \leftarrow \text{Row } 1 \ + \cdot (\text{ Row } 3\ ) \\ \leftarrow \text{Row } 2 \ + (-2) \cdot (\text{ Row } 3\ ) \\ \\ \end{array}$$

Where we see that

$$\boxed{A^{-1} = \begin{bmatrix} -1 & 1 & -1 \\ 4 & -2 & -1 \\ -2 & 1 & 1 \end{bmatrix}}$$

(b) To solve $A\boldsymbol{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, we just multiply both sides on the left by A^{-1} to get:

$$\boldsymbol{x} = A^{-1} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 & 1 & -1 \\ 4 & -2 & -1 \\ -2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

computing this product, our solution is

$$\boxed{\boldsymbol{x} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}}$$

(c) While both methods will still work, $\boxed{\text{directly row reducing the augmented matrix is much faster.}}$

For this problem, we would have only needed to reduce an augmented 3×4 matrix. With the inverse method, we had to first reduce an augmented 3×6 matrix, and then compute a whole extra matrix-vector product on top.

Answer to Question 5.

(a) To compute the 1-norms, we need to sum up all the absolute values in each column, and then pick the maximum of those. For the two given matrices,

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$
$$4 \quad \textcircled{6}$$

and

$$\begin{bmatrix} 1 & 0 & 3 \\ 2 & -1 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$
$$\textcircled{4} \quad 1 \quad \textcircled{4}$$

so $\boxed{||A||_1 = 6}$ and $\boxed{||B||_1 = 4}$.

For the ∞-norm, we now sum up all the absolute values in each row, and then pick the maximum of those.

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{matrix} 3 \\ \textcircled{7} \end{matrix}, \qquad \begin{bmatrix} 1 & 0 & 3 \\ 2 & -1 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{matrix} \textcircled{4} \\ \textcircled{4} \\ 1 \end{matrix}$$

so $\boxed{||A||_\infty = 7}$ and $\boxed{||B||_\infty = 4}$.

For the Frobenius norm, we take each entry in the matrix, square it, add them all together, and then take the square root:

$$||A||_F = \sqrt{1^2 + 2^2 + 3^2 + 4^2}, \qquad ||B||_F = \sqrt{1^2 + 3^2 + 2^2 + 1^2 + 1^2 + 1^2}$$

which simplifies to $\boxed{||A||_F = \sqrt{30}}$ and $\boxed{||B||_F = \sqrt{17}}$.

(b) Since we have the matrix $A = \begin{bmatrix} 1 & 0 \\ 1 & \varepsilon \end{bmatrix}$, we can compute its inverse as follows:

$$\left[\begin{array}{cc|cc} 1 & 0 & 1 & 0 \\ 1 & \varepsilon & 0 & 1 \end{array}\right]$$
$$\left[\begin{array}{cc|cc} 1 & 0 & 1 & 0 \\ 0 & \varepsilon & -1 & 1 \end{array}\right]$$
$$\left[\begin{array}{cc|cc} 1 & 0 & 1 & 0 \\ 0 & 1 & -1/\varepsilon & 1/\varepsilon \end{array}\right]$$

so $A^{-1} = \begin{bmatrix} 1 & 0 \\ -1/\varepsilon & 1/\varepsilon \end{bmatrix}$.

The original question asked for the condition number in only one specific norm, but here I'll just compute everything in all three norms. For the 1-norm,

$$\begin{bmatrix} 1 & 0 \\ 1 & \varepsilon \end{bmatrix} \qquad\qquad \begin{bmatrix} 1 & 0 \\ -1/\varepsilon & 1/\varepsilon \end{bmatrix}$$

$$2 \quad \varepsilon \qquad\qquad 1+\frac{1}{\varepsilon} \quad \frac{1}{\varepsilon}$$

$$||A||_1 = 2 \qquad\qquad ||A^{-1}||_1 = 1+\frac{1}{\varepsilon}$$

$$\boxed{\kappa_1(A) = ||A||_1 \cdot ||A^{-1}||_1 = 2+\frac{2}{\varepsilon}}$$

For the ∞-norm,

$$\begin{bmatrix} 1 & 0 \\ 1 & \varepsilon \end{bmatrix} \begin{matrix} 1 \\ 1+\varepsilon \end{matrix} \qquad\qquad \begin{bmatrix} 1 & 0 \\ -1/\varepsilon & 1/\varepsilon \end{bmatrix} \begin{matrix} 1 \\ \frac{2}{\varepsilon} \end{matrix}$$

$$||A||_\infty = 1+\varepsilon \qquad\qquad ||A^{-1}||_\infty = \frac{2}{\varepsilon}$$

$$\boxed{\kappa_\infty(A) = ||A||_\infty \cdot ||A^{-1}||_\infty = 2+\frac{2}{\varepsilon}}$$

For the Frobenius norm,

$$||A||_F = \sqrt{1^2+1^2+\varepsilon^2} = \sqrt{2+\varepsilon^2} \qquad\qquad ||A^{-1}||_F = \sqrt{1^2+\frac{1}{\varepsilon^2}+\frac{1}{\varepsilon^2}} = \sqrt{1+\frac{2}{\varepsilon^2}}$$

$$\kappa_F(A) = ||A||_F \cdot ||A^{-1}||_F = \sqrt{2+\varepsilon^2}\sqrt{1+\frac{2}{\varepsilon^2}} = \varepsilon\left(1+\frac{2}{\varepsilon^2}\right)$$

$$\boxed{\kappa_F(A) = \varepsilon+\frac{2}{\varepsilon}}$$

In all three cases, we see that $\boxed{\kappa(A) \to +\infty \text{ as } \varepsilon \to 0}$.

(c) To row reduce this matrix to echelon form, we would replace the second row with (-5) times the first row. The corresponding row replacement matrix R_1 is:

$$R_1 = \begin{bmatrix} 1 & 0 \\ -5 & 1 \end{bmatrix}$$

we can calculate that its inverse is:

$$R_1^{-1} = \begin{bmatrix} 1 & 0 \\ 5 & 1 \end{bmatrix}$$

Their Frobenius norms are:

$$||R_1||_F = \sqrt{1^2+1^2+5^2} = \sqrt{27}, \qquad\qquad ||R_1^{-1}||_F = \sqrt{1^2+1^2+(-5)^2} = \sqrt{27}$$

so the condition number is $\boxed{\kappa(R_1) = 27}$, but this number may change depending upon the original matrix A.

For the Givens rotation, we are going to use a rotation matrix of the form:

$$R_2 = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

which corresponds to a counterclockwise rotation of an angle θ around the origin. To "undo" this linear transformation, we would just rotate the same angle in the opposite matrix, so its inverse matrix will be:

$$R_2^{-1} = \begin{bmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{bmatrix} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}$$

It turns out that we can compute these Frobenius norms without even knowing the angle θ, since:

$$||R_2||_F = \sqrt{\cos^2(\theta) + \sin^2(\theta) + \cos^2(\theta) + \sin^2(\theta)} = \sqrt{2}$$
$$||R_2^{-1}||_F = \sqrt{\cos^2(\theta) + \sin^2(\theta) + \cos^2(\theta) + \sin^2(\theta)} = \sqrt{2}$$

So the condition number is $\boxed{\kappa(R_2) = 2}$, which as we can see does not depend upon the choice of angle θ, so it will not depend on the original matrix A.