Numerical analysis: Homework 1

Pratyush Sudhakar (ps2245)
Submission Date: February 7, 2024

Question 2:

For this problem, let $A \in \mathbb{R}^{n \times n}$ be a square matrix and $x \in \mathbb{R}^n$ be a vector of length n. Prove the following:

1. $\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\|x\|_\infty$

   Solution:

$$\|x\|_\infty = \max|x_i| = |x_a| \quad \text{for some } a \in 1, 2, 3 \ldots n,$$
$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

Since $\|x\|_\infty$ is the maximum absolute value among all elements of x, for each term in the vector x, we have:

$$|x_i| \leq \|x\|_\infty = |x_a| \quad \forall i \in \{1, \ldots, n\}.$$

Squaring both sides of the inequality, we get (we can do it because both sides are non-negative):

$$x_i^2 \leq x_a^2 \quad \forall i \in \{1, \ldots, n\}.$$

Summing this inequality over all i gives us:

$$\sum_{i=1}^{n} x_i^2 \leq \sum_{i=1}^{n} x_a^2 = n \cdot x_a^2.$$

Taking the square root on both sides (we can do this because both sides are non-negative), we have:

$$\sqrt{\sum_{i=1}^{n} x_i^2} \leq \sqrt{n \cdot x_a^2} = \sqrt{n} \cdot |x_a|.$$

Since $|x_a| = \|x\|_\infty$, this simplifies to:

$$\|x\|_2 \leq \sqrt{n} \cdot \|x\|_\infty.$$

Hence, proved.

On the other hand, for the first inequality, we focus on the fact that the square of the infinity norm is less than or equal to the sum of the squares of all components, which is due to the definition of the infinity norm being the largest component and it is present on the R.H.S with other positive quantities:

$$|x_a|^2 \leq \sum_{i=1}^{n} x_i^2.$$

Taking the square root on both sides (both sides are non-negative), we obtain:

$$|x_a| \leq \sqrt{\sum_{i=1}^{n} x_i^2} = \|x\|_2.$$

Since $|x_a|$ was chosen as the maximum absolute value of the components of x (i.e., $\|x\|_\infty$), this shows that:

$$\|x\|_\infty \leq \|x\|_2.$$

Combining the results, we conclude:

$$\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n} \cdot \|x\|_\infty.$$

Thus, both parts of the inequality are proven.

2. $\|A\|_2 \leq \sqrt{n}\|A\|_\infty$

   Solution: By the definition of norms

   $\|A\|_2 = \max\limits_{\|x\|_2=1} \|Ax\|_2$, we have:

$$\|A\|_2 = \max_{||x||_2 \neq 0} \frac{||Ax||_2}{||x||_2} \leq \max_{||x||_2 \neq 0} \frac{\sqrt{n}\|Ax\|_\infty}{||x||_2} \; (\because \textit{part a})$$

   Also, from part a we know that $||x||_\infty \leq ||x||_2$. Therefore, the denominator is more than $||x||_\infty$. Hence,

$$\max_{x \neq 0} \frac{||Ax||_2}{||x||_2} \leq \max_{||x||_2 \neq 0} \frac{\sqrt{n}\|Ax\|_\infty}{||x||_2} \leq \max_{|||x||2 \neq 0} \frac{\sqrt{n}\|Ax\|_\infty}{||x||_\infty} \leq \max_{|||x||\infty \neq 0} \frac{\sqrt{n}\|Ax\|_\infty}{||x||_\infty} (\because ||x_2|| \neq 0 \implies ||x||_\infty \neq 0)$$

   This is equal to $\sqrt{n} \cdot ||A||_\infty$

   Hence, proved.

$$\|A\|_2 \leq \sqrt{n}\|A\|_\infty$$

3. For any orthogonal matrix $Q \in \mathbb{R}^{n\times n}$: $\|Qx\|_2 = \|x\|_2$.

   Solution:

   To show that for any orthogonal matrix $Q \in \mathbb{R}^{n\times n}$, the 2-norm of Qx is equal to the 2-norm of x, we proceed as follows:

   Given that Q is an orthogonal matrix, one of its key properties is that the transpose of Q, denoted by Q^T, is also its inverse. That is, $Q^TQ = I$, where I is the identity matrix.

   The 2-norm of a vector x, denoted by $\|x\|_2$, is defined as the square root of the sum of the squares of its components. This can be represented in matrix form as $\|x\|_2 = \sqrt{x^Tx}$.

   Now, consider the 2-norm of the vector Qx:

   $$\|Qx\|_2 = \sqrt{(Qx)^T(Qx)}.$$

   By the property of transposition of a product of matrices, we have:

   $$(Qx)^T = x^TQ^T.$$

   Substituting this into our expression for $\|Qx\|_2$, we get:

   $$\|Qx\|_2 = \sqrt{x^TQ^TQx}.$$

   Using the property of orthogonal matrices $Q^TQ = I$, this simplifies to:

   $$\|Qx\|_2 = \sqrt{x^TIx}.$$

   The property of the identity matrix is such that $Ix = x$. Therefore, we can further simplify the expression:

   $$\|Qx\|_2 = \sqrt{x^Tx}.$$

   Recognizing that the right side of the equation is the definition of the 2-norm of x, we finally have:

   $$\|Qx\|_2 = \|x\|_2.$$

   This proves that the 2-norm of a vector is invariant under orthogonal transformations. Q.E.D.

4. $\|A\|_2 = \sigma_{\max}$, where $\sigma_{\max}$ is the largest singular value of A.

   Solution: We start by defining the matrix 2-norm (also known as the spectral norm), which is given by

   $$\|A\|_2 = \sqrt{\max_{\|x\|_2=1} \|Ax\|_2^2}.$$

   This can be rewritten using the property that for any matrix A, $\|Ax\|_2^2 = x^TA^TAx$. Hence,

   $$\|A\|_2^2 = \max_{\|x\|_2=1} x^TA^TAx.$$

Since A^TA is symmetric, we can diagonalize it using its singular value decomposition as $A^TA = V\Lambda V^T$, where V is an orthogonal matrix ($V^TV = I$) and Λ is a diagonal matrix with non-negative real numbers on the diagonal (since A^TA is diagonal matrix with singular values). This means we can write

$$x^TA^TAx = x^TV\Lambda V^Tx.$$

Given that V is orthogonal, we can substitute $y = V^Tx$ (where y also has a 2-norm of 1), and thus we have

$$x^TV\Lambda V^Tx = y^T\Lambda y = \sum_{i=1}^{n} \lambda_i y_i^2,$$

where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$. To maximize this expression under the constraint that $\|y\|_2 = 1$, we choose y to be the first standard basis vector e_1, which corresponds to choosing x to be the first column of V, v_1. Then we get

$$\|A\|_2^2 = \lambda_{\max}(A^TA),$$

where $\lambda_{\max}(A^TA)$ is the largest eigenvalue of A^TA, and thus

$$\|A\|_2 = \sqrt{\lambda_{\max}(A^TA)}.$$

We can prove that the singular values of A are the square roots of the eigenvalues of A^TA,

The singular value decomposition (SVD) of a matrix A is given by $A = U\Sigma V^T$, where:

Considering the matrix A^TA, we have:

$$A^TA = (U\Sigma V^T)^T(U\Sigma V^T) = V\Sigma^TU^TU\Sigma V^T$$

Since U is orthogonal, we have $U^TU = I$, where I is the identity matrix. Thus:

$$A^TA = V\Sigma^T\Sigma V^T$$

The matrix $\Sigma^T\Sigma$ is a diagonal matrix with the squares of the singular values of A as its diagonal entries. Therefore, the diagonal entries of $\Sigma^T\Sigma$ are the eigenvalues of A^TA, which means the singular values of A are the square roots of these eigenvalues.

Hence, we can write:

$$\|A\|_2 = \sigma_{\max} = \sqrt{\lambda_{\max}(A^TA)}$$

where $\|A\|_2$ is the 2-norm of A, $\sigma_{\max}$ is the largest singular value of A, and $\lambda_{\max}(A^TA)$ is the largest eigenvalue of A^TA.

$$\|A\|_2 = \sigma_{\max},$$

which is the desired result.

5. For any orthogonal matrix $Q \in \mathbb{R}^{n\times n}$: $\|QA\|_2 = \|A\|_2$.

   Solution:

$$\begin{aligned}
\|QA\|_2 &= \max_{\|x\|_2=1} \|QAx\|_2, \\
&= \max_{\|x\|_2=1} \sqrt{(QAx)^T(QAx)}, \\
&= \max_{\|x\|_2=1} \sqrt{(x^TA^TQ^TQAx)}, \because Property\ of\ Tranposition \\
&= \max_{\|x\|_2=1} \sqrt{(x^TA^TAx)}, \because Property\ of\ OrthogonalMatrices \\
&= \max_{\|x\|_2=1} \|Ax\|_2, \\
&= \|A\|_2. \quad \text{(Q.E.D.)}
\end{aligned}$$

Question 3:

For this problem, let $V \in \mathbb{R}^{m\times n}$ with $m > n$ be a matrix with linearly independent columns, prove that

1. V^TV is positive definite
2. VV^T is positive semi-definite but not positive definite

Solution: First, I will state some properties of the given matrix V. If $m > n$ and all the columns are linearly independent that it means that the rank is n.

Hence, by definition of Singular Value Decomposition, there must be n non-zero singular values. Let's denote the singular values as

$$\sigma_1, \sigma_2, \dots, \sigma_n.$$

The matrix V can be decomposed as $V = U\Sigma W^T$, where U and W are orthogonal matrices and Σ is a diagonal matrix with the singular values on the diagonal.

Here, U and W are square matrices of size m and n respectively. The matrix U is of size $m \times m$ and W is of size $n \times n$.

1. **Proof for part 1:**

   We need to show that V^TV is positive definite. The matrix V^TV is of size $n\times n$. Let's denote it as $A = V^TV$.

   We can express A as:

$$A = V^TV = (U\Sigma W^T)^T(U\Sigma W^T) = W\Sigma^TU^TU\Sigma W^T = W\Sigma^T\Sigma W^T.$$

   Since W is an orthogonal matrix, $W^T = W^{-1}$. Therefore, we can simplify the expression for A as:

$$A = W\Sigma^T\Sigma W^T = W\Sigma^T\Sigma W^{-1}.$$

Let's denote $\Sigma^T\Sigma$ as Λ. Then, we can express A as:

$$A = W\Lambda W^{-1}.$$

We know that Σ has n non-zero value on the diagonal. Hence, $\Sigma^T\Sigma$ is an $n \times n$ diagonal matrix with n non-zero values on the diagonal (the squares of the singular values). This means that Λ is diagonal matrix with n postive values on the diagonal.

We also know that W is an orthogonal matrix. Hence, $W^{-1} = W^T$.

Therefore, $\forall x \in \mathbb{R}^n$, we have:

$$\begin{aligned} x^T Ax &= x^T W\Lambda W^T x \\ &= (W^T x)^T \Lambda (W^T x) \end{aligned}$$

Let's denote $y = W^T x$. Then, we have:

$$x^T Ax = y^T \Lambda y.$$

We can expand the expression for $y^T\Lambda y$ as:

$$y^T \Lambda y = \sum_{i=1}^{n} \lambda_i y_i^2.$$

Since Λ is a diagonal matrix with n positive values on the diagonal, we have:

$$\lambda_i y_i^2 > 0 \quad \forall i \in \{1, \ldots, n\}.$$

This means that the sum of the terms in the expansion of $y^T\Lambda y$ is positive. Hence, we have:

$$y^T \Lambda y > 0.$$

Since $y^T\Lambda y = x^T Ax$, we have:

$$x^T Ax > 0 \quad \forall x \in \mathbb{R}^n.$$

This means that A is positive definite. Q.E.D.

2. **Proof for part 2:**

   We need to show that VV^T is positive semi-definite but not positive definite. The matrix VV^T is of size $m \times m$. Let's denote it as $B = VV^T$.

   We can express B as:

   $$B = VV^T = (U\Sigma W^T)(U\Sigma W^T)^T = U\Sigma W^T W\Sigma^T U^T = U\Sigma\Sigma^T U^T.$$

   Let's denote $\Sigma\Sigma^T$ as Λ. Then, we can express B as:

   $$B = U\Lambda U^T.$$

   We know that Σ has n non-zero value on the diagonal. Hence, $\Sigma\Sigma^T$ is an $m \times m$ diagonal matrix with the first n non-zero values on the diagonal (the squares of the singular values) with the rest of the diagonal entires being zero. This means that Λ is diagonal matrix with n postive values on the diagonal and the rest of the diagonal entires being zero.

   Therefore, $\forall x \in \mathbb{R}^m$, we have:

   $$\begin{aligned} x^T Bx &= x^T U\Lambda U^T x \\ &= (U^T x)^T \Lambda (U^T x) \end{aligned}$$

   Let's denote $y = U^T x$. Then, we have:

   $$x^T Bx = y^T \Lambda y.$$

   We can expand the expression for $y^T \Lambda y$ as:

   $$y^T \Lambda y = \sum_{i=1}^{n} \lambda_i y_i^2.$$

   Since Λ is a diagonal matrix with n positive values on the diagonal and the rest of the diagonal entires being zero, we have:

   $$\lambda_i y_i^2 \geq 0 \quad \forall i \in \{1, \ldots, m\}.$$

   It implies that the sum of all the terms in the expansion of $y^T \Lambda y$ is non-negative. Hence, we have:

   $$y^T \Lambda y \geq 0.$$

   Equality at $y^T \Lambda y = 0$ is achieved when the first n entries of y are zero.
   Since $y^T \Lambda y = x^T Bx$, we have:

   $$x^T Bx \geq 0 \quad \forall x \in \mathbb{R}^m.$$

   Equality when the first n entries of $U^T x$ are zero. This means that B is positive semi-definite.

QUESTION 4:

For differentiable functions $f(x)$, where the input x may be a vector $x = (x_1, x_2, \ldots, x_n)$, we define the relative condition number $\kappa_2(x)$ of computing $f(x)$ at x as

$$\kappa_2(x) \equiv \frac{\|J(x)\|_2}{\|f(x)\|_2/\|x\|_2},$$

where J is the Jacobian of f.

1. Compute $\kappa_2(x)$ for subtraction, *i.e.* $f(x) = x_1 - x_2$. When, if ever, is this an ill-conditioned problem?

2. Compute $\kappa_2(x)$ for multiplication, *i.e.* $f(x) = x_1 x_2$. When, if ever, is this an ill-conditioned problem?

1. **Solution for part 1:** To compute the relative condition number $\kappa_2(x)$ for the subtraction function $f(x) = x_1 - x_2$, we need to compute the Jacobian of f at x and then use it to compute the condition number.

   The Jacobian of f at x is a 1×2 matrix, given by:

   $$J(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} \end{bmatrix}. \tag{1}$$

   $$J(x) = \begin{bmatrix} 1 & -1 \end{bmatrix}. \tag{2}$$

   Next, the 2-norm of the Jacobian matrix is given by:

   $$\|J(x)\|_2 = \sqrt{\sum_{i=1}^{2} (J(x)_{1i})^2}. \tag{3}$$

   Next, the 2-norm of the vector x is gives as:

   $$\|x\|_2 = \sqrt{\sum_{i=1}^{2} x_i^2}. \tag{4}$$

   Finally,

   $$\|f(x)\| = |x_1 - x_2|. \tag{5}$$

   Therefore, the relative condition number $\kappa_2(x)$ is given by:

   $$\kappa_2(x) = \frac{\|J(x)\|_2}{\|f(x)\|_2/\|x\|_2} = \frac{\sqrt{1^2 + (-1)^2}}{|x_1 - x_2|/\sqrt{x_1^2 + x_2^2}} = \frac{\sqrt{2}}{|x_1 - x_2|/\sqrt{x_1^2 + x_2^2}}. \tag{6}$$

   We know that function is ill-conditioned when $\kappa_2(x)$ is large. This happens when the denominator is small. Therefore, the subtraction function is ill-conditioned when x_1 and x_2 are close to each other. In this case, the denominator is small and the relative condition number $\kappa_2(x)$ is large.

2. **Solution for part 2:** To compute the relative condition number $\kappa_2(x)$ for the multiplication function $f(x) = x_1x_2$, we need to compute the Jacobian of f at x and then use it to compute the condition number.

   The Jacobian of f at x is a 1×2 matrix, given by:

$$J(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} \end{bmatrix}. \tag{7}$$

$$J(x) = \begin{bmatrix} x_2 & x_1 \end{bmatrix}. \tag{8}$$

   Next, the 2-norm of the Jacobian matrix is given by:

$$\|J(x)\|_2 = \sqrt{\sum_{i=1}^{2}(J(x)_{1i})^2}. \tag{9}$$

   Next, the 2-norm of the vector x is gives as:

$$\|x\|_2 = \sqrt{\sum_{i=1}^{2} x_i^2}. \tag{10}$$

   Finally,

$$\|f(x)\| = |x_1x_2|. \tag{11}$$

   Therefore, the relative condition number $\kappa_2(x)$ is given by:

$$\kappa_2(x) = \frac{\|J(x)\|_2}{\|f(x)\|_2/\|x\|_2} = \frac{\sqrt{x_2^2 + x_1^2}}{|x_1x_2|/\sqrt{x_1^2 + x_2^2}} = \frac{\sqrt{x_2^2 + x_1^2}}{|x_1x_2|/\sqrt{x_1^2 + x_2^2}} = \frac{x_1^2 + x_2^2}{|x_1x_2|}. \tag{12}$$

   We know that function is ill-conditioned when $\kappa_2(x)$ is large. This happens when the denominator is small and the numerator large. It could happen when x_1 and x_2 are far away. In that case,

$$|x_2| >> |x_1| \tag{13}$$

$$\implies x_2^2 >> x_1^2 \tag{14}$$

$$\implies x_2^2 + x_1^2 \approx x_2^2 \tag{15}$$

$$\implies \frac{(x_2^2 + x_1^2)}{|x_1x_2|} \approx \frac{x_2^2}{|x_1x_2|} = \frac{|x_2|}{|x_1|} \tag{16}$$

   from equation (13), $\frac{|x_2|}{|x_1|} \to \infty$. Hence, the condition number is very large and the problem is ill-conditioned.