Math 33A
Linear Algebra and Applications

Discussion 6

Problem 1($\star$).

The following is one way to define the quaternions, discovered in 1843 by the Irish mathematician Sir W. R. Hamilton. Consider the set H of all 4×4 matrices M of the form

$$M = \begin{bmatrix} p & -q & -r & -s \\ q & p & s & -r \\ r & -s & p & q \\ s & r & -q & p \end{bmatrix}$$

where p, q, r, s are arbitrary real numbers. We can write M more succinctly in partitioned form as

$$M = \begin{bmatrix} A & -B^T \\ B & A^T \end{bmatrix}$$

where A and B are rotation–scaling matrices.

(a) Show that H is closed under addition: If M and N are in H, then so is $M + N$.

(b) Show that H is closed under scalar multiplication: If M is in H and k is an arbitrary scalar, then kM is in H.

(c) The above show that H is a subspace of the linear space $\mathbb{R}^{4\times 4}$. Find a basis of H, and thus determine the dimension of H.

(d) Show that H is closed under multiplication: If M and N are in H, then so is MN.

(e) Show that if M is in H, then so is M^T.

(f) For a matrix M in H, compute $M^T M$.

(g) Which matrices M in H are invertible? If a matrix M in H is invertible, is M^{-1} necessarily in H as well?

(h) If M and N are in H, does the equation $MN = NM$ always hold?

Solution:

(a) When we add two matrices in H we obtain another matrix in H

$$\begin{bmatrix} A & -B^T \\ B & A^T \end{bmatrix} + \begin{bmatrix} C & -D^T \\ D & C^T \end{bmatrix} = \begin{bmatrix} (A+C) & -(B+D)^T \\ (B+D) & (A+C)^T \end{bmatrix}.$$

(b) When we multiply a matrix in H by a real scalar we obtain a matrix in H

$$k \begin{bmatrix} A & -B^T \\ B & A^T \end{bmatrix} = \begin{bmatrix} (kA) & -(kB)^T \\ (kB) & (kA)^T \end{bmatrix}.$$

(c) The general element of H has four arbitrary constants, so H has dimension 4. A basis is

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

(d) When we multiply two matrices in H we obtain another matrix in H

$$\begin{bmatrix} A & -B^T \\ B & A^T \end{bmatrix} \begin{bmatrix} C & -D^T \\ D & C^T \end{bmatrix} = \begin{bmatrix} (AC - B^T D) & -(BC + A^T D)^T \\ (BC + A^T D) & (AC - B^T D)^T \end{bmatrix}$$

where it is useful to notice that since all A, B, C, D are rotation–scaling matrices, they commute with each other.

(e) When we transpose a matrix in H we obtain another matrix in H

$$\begin{bmatrix} A & -B^T \\ B & A^T \end{bmatrix}^T = \begin{bmatrix} (A^T) & -(-B)^T \\ (-B) & (A^T)^T \end{bmatrix}.$$

(f) We expand $M^T M$ as

$$\begin{bmatrix} p & q & r & s \\ -q & p & -s & r \\ -r & s & p & -q \\ -s & -r & q & p \end{bmatrix} \begin{bmatrix} p & -q & -r & -s \\ q & p & s & -r \\ r & -s & p & q \\ s & r & -q & p \end{bmatrix} = (p^2 + q^2 + r^2 + s^2) \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

(g) If $M \neq 0$ then $p^2 + q^2 + r^2 + s^2 \neq 0$ so by the above

$$M^T M = (p^2 + q^2 + r^2 + s^2) \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

and thus

$$\left(\frac{1}{(p^2 + q^2 + r^2 + s^2)} M^T \right) M = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

so

$$M^{-1} = \frac{1}{(p^2 + q^2 + r^2 + s^2)} M^T = \frac{1}{(p^2 + q^2 + r^2 + s^2)} \begin{bmatrix} p & q & r & s \\ -q & p & -s & r \\ -r & s & p & -q \\ -s & -r & q & p \end{bmatrix}.$$

Problem 2.

Consider a consistent system $A\vec{x} = \vec{b}$.

(a) Show that this system has a solution $\vec{x_0}$ in $(\ker A)^\perp$. Justify why an arbitrary solution $\vec{x}$ of the system can be written as $\vec{x} = \vec{x_h} + \vec{x_0}$, where $\vec{x_h}$ is in $\ker(A)$ and

$\vec{x_0}$ is in $(\ker A)^{\perp}$.

(b) Show that the system $A\vec{x} = \vec{b}$ has only one solution in $(\ker A)^{\perp}$.

(c) If $\vec{x_0}$ is the solution in $(\ker A)^{\perp}$ and $\vec{x_1}$ is another solution of the system $A\vec{x} = \vec{b}$, show that $||\vec{x_0}|| < ||\vec{x_1}||$. The vector $\vec{x_0}$ is called the minimal solution of the linear system $A\vec{x} = \vec{b}$.

Solution:

(a) Since the system $A\vec{x} = \vec{b}$ is consistent, it has at least one solution $\vec{x}$. Let $\vec{x} = \vec{x}^{||} + \vec{x}^{\perp}$ be the decomposition of $\vec{x}$ into the components parallel and perpendicular to $V = \ker(A)$. In particular $\vec{x}^{\perp}$ is in $(\ker(A))^{\perp}$ and $\vec{x}^{||} = \text{proj}_V\vec{x}$ is in $\ker(A)$ so $A\vec{x}^{||} = \vec{0}$. Now

$$\vec{b} = A\vec{x} = A(\vec{x}^{||} + \vec{x}^{\perp}) = A\vec{x}^{||} + A\vec{x}^{\perp} = A\vec{x}^{\perp}$$

so $\vec{x_0} = \vec{x}^{\perp}$ is a solution of the system in $(\ker(A))^{\perp}$ and $\vec{x_h} = \vec{x}^{||}$ is in $\ker(A)$.

(b) Suppose that $A\vec{x} = \vec{b}$ has two solutions $\vec{x_1}$ and $\vec{x_2}$ in $(\ker(A))^{\perp}$. Since $(\ker(A))^{\perp}$ is a linear subspace, then $\vec{x_1}-\vec{x_2}$ is in $(\ker(A))^{\perp}$. Thus $A(\vec{x_1}-\vec{x_2}) = A\vec{x_1}-A\vec{x_2} = \vec{b}-\vec{b} = \vec{0}$ so $\vec{x_1}-\vec{x_2}$ is in $\ker(A)$. Now $\vec{x_1}-\vec{x_2}$ is both in $\ker(A)$ and $(\ker(A))^{\perp}$, but $\vec{0}$ is the only element in both subspaces, so $\vec{x_1}-\vec{x_2} = \vec{0}$. Thus $\vec{x_1} = \vec{x_2}$.

(c) Let $\vec{x_1} = \vec{x_1}^{||} + \vec{x_1}^{\perp}$ be the decomposition of $\vec{x_1}$ into the components parallel and perpendicular to $V = \ker(A)$. Now by the first part above we have that $\vec{x_1}^{\perp}$ is a solution of the system in $(\ker(A))^{\perp}$. Since $\vec{x_0}$ is also a solution of the system in $(\ker(A))^{\perp}$, by the second part above we have $\vec{x_1}^{\perp} = \vec{x_0}$. Since $\vec{x_1} \neq \vec{x_0}$ we have $\vec{x_1}^{||} \neq \vec{0}$, so $||\vec{x_1}^{||}|| > 0$ and by the Pythagoras theorem

$$||\vec{x_1}|| = ||\vec{x_1}^{||} + \vec{x_0}|| \geq ||\vec{x_1}^{||}|| + ||\vec{x_0}|| > ||\vec{x_0}||.$$

Problem 3.

Define the term minimal least-squares solution of a linear system. Explain why the minimal least-squares solution $\vec{x}^*$ of a linear system $A\vec{x} = \vec{b}$ is in $(\ker A)^{\perp}$.

Solution: We know that the least-squares solution of a linear system $A\vec{x} = \vec{b}$ are the exact solutions of the consistent linear system $A^TA\vec{x} = A^T\vec{b}$. In the previous problem we defined the term minimal solution of a consistent linear system. We then define the minimal least-squares solution of the linear system $A\vec{x} = \vec{b}$ to be the minimal solutions of the consistent linear system $A^TA\vec{x} = A^T\vec{b}$.

We first prove that $\ker(A) = \ker(A^TA)$, this will be useful. Let $\vec{v}$ be in $\ker(A)$, then $A^TA\vec{v} = A^T\vec{0} = \vec{0}$ so $\vec{v}$ is in $\ker(A^TA)$. Let $\vec{v}$ be in $\ker(A^TA)$, then $\vec{0} = A^TA\vec{v} =$

$A^T(A\vec{v})$ so $A\vec{v}$ is in $\ker(A^T)$. Now $A\vec{v}$ is in $\text{im}(A)$, and also in $\ker(A^T) = (\text{im}(A))^\perp$, but $\vec{0}$ is the only element in both subspaces, so $A\vec{v} = \vec{0}$, so $\vec{v}$ is in $\ker(A)$.

Now, let $\vec{x}^*$ be the minimal least-squares solution of the linear system $A\vec{x} = \vec{b}$. Then $\vec{x}^*$ is the minimal solutions of the consistent linear system $A^T A\vec{x} = A^T\vec{b}$, so by the previous exercise $\vec{x}^*$ is in $(\ker(A^T A))^\perp = (\ker(A))^\perp$.