Math 33A
Linear Algebra and Applications

Discussion for July 11-14, 2022

Problem 1.

Here is an infinite-dimensional version of Euclidean space: In the space of all infinite sequences, consider the subspace ℓ_2 of square-summable sequences (namely, those sequences $(x_1, x_2, \dots)$ for which the infinite series $x_1^2 + x_2^2 + \cdots$ converges). For x and y in ℓ_2, we define

$$||\vec{x}|| = \sqrt{x_1^2 + x_2^2 + \cdots} \quad \text{and} \quad \vec{x} \cdot \vec{y} = x_1 y_1 + x_2 y_2 + \cdots .$$

A preliminary question is, why do $||\vec{x}||$ and $\vec{x} \cdot \vec{y}$ make sense, that is, why are they finite real numbers?

(a) Check that $\vec{x} = (1, 1/2, 1/4, 1/8, 1/16, \dots)$ is in ℓ_2, and find $||\vec{x}||$. Recall the formula for the geometric series: $1 + a + a^2 + a^3 + \cdots = 1/(1-a)$ if $-1 < a < 1$.

(b) Find the angle between $(1, 0, 0, 0, \dots)$ and $(1, 1/2, 1/4, 1/8, \dots)$.

(c) Give an example of a sequence $(x_1, x_2, \dots)$ that converges to 0 ($\lim_{n\to\infty} x_n = 0$) but does not belong to ℓ_2.

(d) Let L be the subspace of ℓ_2 spanned by $(1, 1/2, 1/4, 1/8, \dots)$. Find the orthogonal projection of $(1, 0, 0, 0, \dots)$ onto L.

The Hilbert space ℓ_2 was initially used mostly in physics: Werner Heisenberg's formulation of quantum mechanics is in terms of ℓ_2. Today, this space is used in many other applications, including economics. See, for example, the work of the economist Andreu Mas-Colell of the University of Barcelona.

Solution:

(a) Using the formula for the geometric series $||\vec{x}||^2 = 4/3$ so $||\vec{x}|| = 2/\sqrt{3}$.

(b) Set $\vec{x} = (1, 0, 0, 0, \dots)$ and $\vec{y} = (1, 1/2, 1/4, 1/8, \dots)$, then

$$\theta = \arccos\left(\frac{\vec{x} \cdot \vec{y}}{||\vec{x}|| \cdot ||\vec{y}||}\right) = \arccos\left(\frac{1}{2/\sqrt{3}}\right) = \arccos\left(\frac{\sqrt{3}}{2}\right) = \frac{\pi}{6}.$$

(c) Consider $\vec{x} = (1, 1/\sqrt{2}, 1/\sqrt{3}, 1/\sqrt{4}, \dots)$, then

$$||\vec{x}||^2 = \sqrt{1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots} = \sqrt{\sum_{n=1}^{\infty} \frac{1}{n}}$$

which diverges since the harmonic series $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges.

(d) Let $\vec{x} = (1, 0, 0, 0, \dots)$ and $\vec{y} = (1, 1/2, 1/4, 1/8, \dots)$, we want the orthogonal projection of $\vec{x}$ onto $L = \operatorname{span}(\vec{y})$. For this, we first find a vector of length one in the direction of $\vec{y}$, namely

$$\vec{u} = \frac{\vec{y}}{||\vec{y}||} = \frac{\sqrt{3}}{2}\left(1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \dots\right)$$

and now we compute

$$\operatorname{proj}_L(\vec{x}) = (\vec{x}\cdot\vec{u})\vec{u} = \left(\frac{\sqrt{3}}{2}\right)\frac{\sqrt{3}}{2}\left(1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \dots\right) = \left(\frac{3}{4}, \frac{3}{8}, \frac{3}{16}, \frac{3}{32}, \dots\right).$$

Problem 2.

Give an algebraic proof for the triangle inequality

$$||\vec{v} + \vec{w}|| \leq ||\vec{v}|| + ||\vec{w}||.$$

Draw a sketch.

Solution: Note that

$$\begin{aligned} ||\vec{v} + \vec{w}||^2 &= (\vec{v} + \vec{w}) \cdot (\vec{v} + \vec{w}) = \vec{v}\cdot\vec{v} + \vec{v}\cdot\vec{w} + \vec{w}\cdot\vec{v} + \vec{w}\cdot\vec{w} = \\ &= ||\vec{v}||^2 + 2(\vec{v}\cdot\vec{w}) + ||\vec{w}||^2 \leq ||\vec{v}||^2 + 2(||\vec{v}||\cdot||\vec{w}||) + ||\vec{w}||^2 = (||\vec{v}|| + ||\vec{w}||)^2 \end{aligned}$$

where we have used the Cauchy-Schwarz inequality. Thus $||\vec{v} + \vec{w}|| \leq ||\vec{v}|| + ||\vec{w}||$.

Problem 3.

(a) Consider a vector $\vec{v}$ in $\mathbb{R}^n$, and a scalar k. Show that $||k\vec{v}|| = |k|||\vec{v}||$.

(b) Show that if $\vec{v}$ is a nonzero vector in $\mathbb{R}^n$, then $\vec{u} = \frac{\vec{v}}{||\vec{v}||}$ is a unit vector.

Solution:

(a) Note that

$$||k\vec{v}||^2 = (k\vec{v})\cdot(k\vec{v}) = k^2(\vec{v}\cdot\vec{v}) = k^2||\vec{v}||^2$$

and thus taking square roots $||k\vec{v}|| = |k|||\vec{v}||$ since $|k| = \sqrt{k^2}$.

(b) We compute

$$||\vec{u}|| = \left|\left|\frac{\vec{v}}{||\vec{v}||}\right|\right| = \left|\left|\frac{1}{||\vec{v}||}\vec{v}\right|\right| = \frac{1}{||\vec{v}||}||\vec{v}|| = 1$$

using what we just proved.

Problem 4($\star$).

Can you find a line L in $\mathbb{R}^n$ and a vector $\vec{x}$ in $\mathbb{R}^n$ such that $\vec{x} \cdot \text{proj}_L \vec{x}$ is negative? Explain, arguing algebraically.

Solution: No. Let $\vec{x} = \vec{x}^{||} + \vec{x}^{\perp}$ be the decomposition of $\vec{x}$ into the components parallel and perpendicular to L. In particular $\vec{x}^{||} = \text{proj}_L \vec{x}$ and $\vec{x}^{\perp} \cdot \vec{x}^{||} = 0$. Now

$$\vec{x} \cdot \text{proj}_L \vec{x} = (\vec{x}^{||} + \vec{x}^{\perp}) \cdot \vec{x}^{||} = \vec{x}^{||} \cdot \vec{x}^{||} + \vec{x}^{\perp} \cdot \vec{x}^{||} = ||\vec{x}^{||}||^2 \geq 0.$$

Problem 5($\star$).

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 6.

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 7.

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 $A^T A\vec{x} = A^T\vec{b}$. 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 $A^T A\vec{x} = A^T\vec{b}$.

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