Math 308
5-4-1994

Solutions to Problem Set 8.

1. Let V be an inner product space with inner product $f : V \times V \to F$, where $F = \mathbb{R}$ or $F = \mathbb{C}$.

(a) Let $T : V \to V$ be a linear transformation such that

$$f(T(x), T(y)) = f(x, y)$$

for all $x, y \in V$. Prove that T is injective. Conclude that if V is finite-dimensional, then T is an isomorphism.

Suppose $T(v) = 0$. Then

$$0 = f(T(v), T(v)) = f(v, v).$$

By positive definiteness, $v = 0$. Hence, $\ker T = \{0\}$, so T is injective.

An injective map from a finite-dimensional vector space to itself is an isomorphism, by Problem 1 on Problem Set 7. $\square$

An isomorphism of vector spaces $T : V \to V$ is **unitary** if

$$f(T(x), T(y)) = f(x, y)$$

for all $x, y \in V$. ("**Unitary**" is usually used when $F = \mathbb{C}$; if $F = \mathbb{R}$, T is **orthogonal**.)

(b) Prove that the set of unitary (orthogonal) transformations $V \to V$ forms a group under function composition.

I need to assume V is finite-dimensional.

The identity transformation I is unitary, since

$$f(I(v), I(w)) = f(v, w).$$

Suppose S and T are unitary. Then

$$f(ST(v), ST(w)) = f(T(v), T(w)) = f(v, w),$$

the first equality because S is unitary and the second because T is unitary. Therefore, ST is unitary.

Finally, since V is finite-dimensional, unitary maps are isomorphisms, by part (a). Hence, if T is unitary, it has an inverse T^{-1}. I must show T^{-1} is unitary.

Since T is unitary,

$$f(v, w) f\left(TT^{-1}(v), TT^{-1}(w)\right) = f\left(T^{-1}(v), T^{-1}(w)\right).$$

Therefore, T^{-1} is unitary. Hence, the set of unitary maps forms a group. $\square$

The group of unitary transformations $\mathbb{C}^n \to \mathbb{C}^n$ relative to the standard inner product

$$\langle a_1, \ldots, a_n \rangle \cdot \langle b_1, \ldots, b_n \rangle = a_1 \overline{b_1} + \cdots + a_n \overline{b_n}$$

is the **unitary group** $U_n(\mathbb{C})$.

The group of orthogonal transformations $\mathbb{R}^n \to \mathbb{R}^n$ relative to the standard inner product

$$\langle a_1, \ldots, a_n\rangle \cdot \langle b_1, \ldots, b_n\rangle = a_1 b_1 + \cdots + a_n b_n$$

is the **orthogonal group** $O_n(\mathbb{R})$.

(c) Show that an orthogonal transformation T preserves lengths: $|Tx| = |x|$. Show that an orthogonal transformation preserves angles: The angle between x and y is the same (in magnitude) as the angle between Tx and Ty.

Suppose T is orthogonal. Then

$$|T(x)|^2 = T(x) \cdot T(x) = x \cdot x = |x|^2.$$

Hence, $|T(x)| = |x|$.
It follows that

$$\frac{T(x) \cdot T(y)}{|T(x)||T(y)|} = \frac{x \cdot y}{|x||y|}.$$

This says that the cosine of the angle between x and y is the same as the cosine of the angle between $T(x)$ and $T(y)$. If I only consider positive angles — hence, angles in the range $0 \le \theta < \pi$ — this means that the angle between x and y is the same (in magnitude) as the angle between $T(x)$ and $T(y)$. $\square$

A $n \times n$ real matrix is **orthogonal** if it is the matrix of an orthogonal transformation $\mathbb{R}^n \to \mathbb{R}^n$ relative to some basis on $\mathbb{R}^n$.

(d) Show that the rows of an orthogonal matrix form an orthonormal set. Show that the same is true of the columns.

That is, if A is orthogonal and A^T is the transpose of A, then $AA^T = I = A^T A$ — i.e. $A^T = A^{-1}$.

If e_i is the i-th standard basis vector, then Ae_i gives the i-th column of A. If A is orthogonal, then

$$(Ae_i) \cdot (Ae_j) = e_i \cdot e_j = \delta(i, j).$$

This proves that the columns of an orthogonal matrix form an orthonormal set. In matrix form, this means that $A^T A = I$. Multiplying on the right by A^{-1} — which I know exists because orthogonal transformations are isomorphisms — I find $A^T = A^{-1}$. Then $AA^T = AA^{-1} = I$. But this matrix equation says that the rows of A form an orthonormal set. $\square$

(e) Let A be an orthogonal matrix. Prove that $\det A = \pm 1$.

Since $\det A = \det A^T$, $AA^T = I$ implies

$$1 = \det I = \det(AA^T) = (\det A)(\det A^T) = (\det A)^2.$$

Therefore, $\det A = \pm 1$. $\square$

(f) Show that the matrix of an orthogonal transformation $T \in O_2(\mathbb{R})$ has the form

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}.$$

Conclude that an orthogonal transformation of $\mathbb{R}^2$ is either a rotation or a reflection.

The rows of an orthogonal matrix form an orthonormal set. So if

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

is orthogonal, I have

$$a^2 + b^2 = 1, \quad c^2 + d^2 = 1, \quad ac + bd = 0.$$

Since $a^2 + b^2 = 1$, there is an angle θ such that $a = \cos\theta$ and $b = \sin\theta$. Then $ac + bd = 0$ gives

$$c(\cos\theta) + d(\sin\theta) = 0.$$

I can satisfy this equation by setting $c = -\sin\theta$ and $d = \cos\theta$ or by setting $c = \sin\theta$ and $d = -\cos\theta$. Both of these satisfy $c^2 + d^2 = 1$; moreover, these two cases are the only solutions, since a line $c(\cos\theta) + d(\sin\theta) = 0$ can intersect a circle $c^2 + d^2 = 1$ in at most two points. These two solutions yields the matrices

$$\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}.$$

These are the same as the matrices in the problem statement: Simply replace θ with $-\theta$ in the first matrix. $\square$

---

2. An $n \times n$ matrix A is **upper triangular** if $A_{ij} = 0$ for $i > j$. (Thus, the elements below the main diagonal are 0.) Prove by induction that the determinant of an upper triangular matrix is the product of the entries on the main diagonal.

The result is trivial if $n = 1$. Take $n > 1$, and assume the result is true for $(n-1) \times (n-1)$ matrices. Let $A \in M_n(F)$ be upper triangular. Expand $|A|$ by cofactors of row n. The only nonzero term is $(-1)^{n+n} A_{nn} |A(n \mid n)| = A_{nn} |A(n \mid n)|$. Now $A(n \mid n)$ is upper triangular and $(n-1) \times (n-1)$; by induction,

$$|A(n \mid n)| = A_{11} \cdots A_{n-1,n-1}.$$

Therefore,

$$|A| = A_{nn} |A(n \mid n)| = A_{11} \cdots A_{n-1,n-1} A_{nn}.$$

Therefore, the result is true by induction. $\square$

---

3. (a) Let I be the $m \times m$ identity matrix, 0 the $m \times n$ zero matrix, C an $n \times m$ matrix, and D an $n \times n$ matrix. Use cofactor expansion to prove that

$$\begin{vmatrix} I & 0 \\ C & D \end{vmatrix} = |D|.$$

Induct on m, the size of the identity matrix in the upper block. If $m = 1$, the matrix is

$$\begin{bmatrix} 1 & 0 \\ C & D \end{bmatrix}.$$

Expanding by cofactors of the first row, I find that

$$\begin{vmatrix} 1 & 0 \\ C & D \end{vmatrix} = |D|.$$

Take $m > 1$, and assume the result is true for $m-1$. Let I_m be the $m \times m$ identity matrix, 0 the $m \times n$ zero matrix, C an $n \times m$ matrix, and D an $n \times n$ matrix. Expanding by cofactors of the first row,

$$\begin{vmatrix} I_m & 0 \\ C & D \end{vmatrix} = 1 \cdot \begin{vmatrix} I_{m-1} & 0 \\ C(1 \mid 1) & D \end{vmatrix} = 1 \cdot |D| = |D|.$$

(The second equality follows by the induction hypothesis.) This establishes the result for all m. □

(b) Let A be an $m \times m$ matrix, 0 the $m \times n$ zero matrix, C an $n \times m$ matrix, and D an $n \times n$ matrix. Fix C and D, and define a function $T : M_m(F) \to F$ by

$$T(A) = \begin{vmatrix} A & 0 \\ C & D \end{vmatrix}.$$

Prove that T is alternating and linear in each row.

First, I have to show that if I swap two rows of A, $T(A)$ is multiplied by -1. I will let A_k denote the k-th row of A. Swap rows i and j:

$$T\left(\begin{bmatrix} & \vdots & \\ \leftarrow & A_i & \rightarrow \\ & \vdots & \\ \leftarrow & A_j & \rightarrow \\ & \vdots & \end{bmatrix}\right) = \begin{vmatrix} & \vdots & & & \vdots & \\ \leftarrow & A_i & \rightarrow & \leftarrow & 0 & \rightarrow \\ & \vdots & & & \vdots & \\ \leftarrow & A_j & \rightarrow & \leftarrow & 0 & \rightarrow \\ & \vdots & & & \vdots & \\ & & & & & \\ & C & & & D & \\ & & & & & \end{vmatrix} =$$

$$- \begin{vmatrix} & \vdots & & & \vdots & \\ \leftarrow & A_j & \rightarrow & \leftarrow & 0 & \rightarrow \\ & \vdots & & & \vdots & \\ \leftarrow & A_i & \rightarrow & \leftarrow & 0 & \rightarrow \\ & \vdots & & & \vdots & \\ & & & & & \\ & C & & & D & \\ & & & & & \end{vmatrix} = -T\left(\begin{bmatrix} & \vdots & \\ \leftarrow & A_i & \rightarrow \\ & \vdots & \\ \leftarrow & A_j & \rightarrow \\ & \vdots & \end{bmatrix}\right).$$

This proves that T is alternating.

To show that T is linear, take $x, y \in F^m$ and $a \in F$. I'll put a linear combination $ax + y$ into row i of A and expand.

$$T\left(\begin{bmatrix} & \vdots & \\ \leftarrow & ax + y & \rightarrow \\ & \vdots & \end{bmatrix}\right) = \begin{vmatrix} & \vdots & & & \vdots & \\ \leftarrow & ax + y & \rightarrow & \leftarrow & 0 & \rightarrow \\ & \vdots & & & \vdots & \\ & & & & & \\ & C & & & D & \\ & & & & & \end{vmatrix} =$$

$$a \cdot \begin{vmatrix} & \vdots & & & \vdots & \\ \leftarrow & x & \rightarrow & \leftarrow & 0 & \rightarrow \\ & \vdots & & & \vdots & \\ & & & & & \\ & C & & & D & \\ & & & & & \end{vmatrix} + \begin{vmatrix} & \vdots & & & \vdots & \\ \leftarrow & y & \rightarrow & \leftarrow & 0 & \rightarrow \\ & \vdots & & & \vdots & \\ & & & & & \\ & C & & & D & \\ & & & & & \end{vmatrix} =$$

$$a \cdot T\left(\begin{bmatrix} & \vdots & \\ \leftarrow & x & \rightarrow \\ & \vdots & \end{bmatrix}\right) + T\left(\begin{bmatrix} & \vdots & \\ \leftarrow & y & \rightarrow \\ & \vdots & \end{bmatrix}\right).$$

This proves that T is linear in each row. $\square$

(c) Let A be an $m \times m$ matrix, 0 the $m \times n$ zero matrix, C an $n \times m$ matrix, and D an $n \times n$ matrix. Use (a) and (b) to prove that

$$\begin{vmatrix} A & 0 \\ C & D \end{vmatrix} = |A||D|.$$

By (b), T is alternating and linear in each row, so

$$T(A) = |A|T(I) = |A| \begin{vmatrix} I & 0 \\ C & D \end{vmatrix} = |A||D|. \square$$

(d) Let A be an $m \times m$ matrix, B an $m \times n$ matrix, 0 the $n \times m$ zero matrix, and D an $n \times n$ matrix. Use (c) to prove that

$$\begin{vmatrix} A & B \\ 0 & D \end{vmatrix} = |A||D|.$$

By (c),

$$\begin{vmatrix} A & B \\ 0 & D \end{vmatrix} = \left| \begin{bmatrix} A & 0 \\ C & D \end{bmatrix}^T \right| = \begin{vmatrix} A & 0 \\ C^T & D \end{vmatrix} = |A||D|. \square$$

"How can you say my life is not a success? Have I not for more than sixty years got enough to eat and escaped being eaten?" - Logan Pearsall Smith