8D *Trace: A Connection Between Matrices and Operators*

We begin this section by defining the trace of a square matrix. After developing some properties of the trace of a square matrix, we will use this concept to define the trace of an operator.

8.47 definition: *trace of a matrix*

Suppose A is a square matrix with entries in $\mathbf{F}$. The *trace* of A, denoted $\operatorname{tr} A$, is defined to be the sum of the diagonal entries of A.

8.48 example: *trace of a 3-by-3 matrix*

Suppose

$$A = \begin{pmatrix} 3 & -1 & -2 \\ 3 & 2 & -3 \\ 1 & 2 & 0 \end{pmatrix}.$$

The diagonal entries of A, which are shown in red above, are 3, 2, and 0. Thus $\operatorname{tr} A = 3 + 2 + 0 = 5$.

Matrix multiplication is not commutative, but the next result shows that the order of matrix multiplication does not matter to the trace.

8.49 *trace of AB equals trace of BA*

Suppose A is an m-by-n matrix and B is an n-by-m matrix. Then

$$\operatorname{tr}(AB) = \operatorname{tr}(BA).$$

Proof Suppose

$$A = \begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{m,1} & \cdots & A_{m,n} \end{pmatrix}, \quad B = \begin{pmatrix} B_{1,1} & \cdots & B_{1,m} \\ \vdots & & \vdots \\ B_{n,1} & \cdots & B_{n,m} \end{pmatrix}.$$

The j^{th} term on the diagonal of the m-by-m matrix AB equals $\sum_{k=1}^{n} A_{j,k}B_{k,j}$. Thus

$$\begin{aligned} \operatorname{tr}(AB) &= \sum_{j=1}^{m} \sum_{k=1}^{n} A_{j,k}B_{k,j} \\ &= \sum_{k=1}^{n} \sum_{j=1}^{m} B_{k,j}A_{j,k} \\ &= \sum_{k=1}^{n} \bigl(k^{\text{th}} \text{ term on diagonal of the } n\text{-by-}n \text{ matrix } BA\bigr) \\ &= \operatorname{tr}(BA), \end{aligned}$$

as desired. ■

We want to define the trace of an operator $T \in \mathcal{L}(V)$ to be the trace of the matrix of T with respect to some basis of V. However, this definition should not depend on the choice of basis. The following result will make this possible.

8.50 *trace of matrix of operator does not depend on basis*

Suppose $T \in \mathcal{L}(V)$. Suppose $u_1, \dots, u_n$ and $v_1, \dots, v_n$ are bases of V. Then

$$\operatorname{tr} \mathcal{M}\big(T, (u_1, \dots, u_n)\big) = \operatorname{tr} \mathcal{M}\big(T, (v_1, \dots, v_n)\big).$$

Proof Let $A = \mathcal{M}\big(T, (u_1, \dots, u_n)\big)$ and $B = \mathcal{M}\big(T, (v_1, \dots, v_n)\big)$. The change-of-basis formula tells us that there exists an invertible n-by-n matrix C such that $A = C^{-1}BC$ (see 3.84). Thus

$$\begin{aligned} \operatorname{tr} A &= \operatorname{tr}\Big(\big(C^{-1}B\big)C\Big) \\ &= \operatorname{tr}\Big(C\big(C^{-1}B\big)\Big) \\ &= \operatorname{tr}\Big(\big(CC^{-1}\big)B\Big) \\ &= \operatorname{tr} B, \end{aligned}$$

where the second line comes from 8.49. ■

Because of 8.50, the following definition now makes sense.

8.51 definition: *trace of an operator*

Suppose $T \in \mathcal{L}(V)$. The *trace* of T, denoted $\operatorname{tr} T$, is defined by

$$\operatorname{tr} T = \operatorname{tr} \mathcal{M}\big(T, (v_1, \dots, v_n)\big),$$

where $v_1, \dots, v_n$ is any basis of V.

Suppose $T \in \mathcal{L}(V)$ and λ is an eigenvalue of T. Recall that we defined the multiplicity of λ to be the dimension of the generalized eigenspace $G(\lambda, T)$ (see 8.23); we proved that this multiplicity equals $\dim \operatorname{null}(T - \lambda I)^{\dim V}$ (see 8.20). Recall also that if V is a complex vector space, then the sum of the multiplicities of all eigenvalues of T equals $\dim V$ (see 8.25).

In the definition below, the sum of the eigenvalues "with each eigenvalue included as many times as its multiplicity" means that if $\lambda_1, \dots, \lambda_m$ are the distinct eigenvalues of T with multiplicities $d_1, \dots, d_m$, then the sum is

$$d_1\lambda_1 + \cdots + d_m\lambda_m.$$

Or if you prefer to work with a list of not-necessarily-distinct eigenvalues, with each eigenvalue included as many times as its multiplicity, then the eigenvalues could be denoted by $\lambda_1, \dots, \lambda_n$ (where n equals $\dim V$) and the sum is

$$\lambda_1 + \cdots + \lambda_n.$$

8.52 *on complex vector spaces, trace equals sum of eigenvalues*

Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Then $\operatorname{tr} T$ equals the sum of the eigenvalues of T, with each eigenvalue included as many times as its multiplicity.

Proof There is a basis of V with respect to which T has an upper-triangular matrix with the diagonal entries of the matrix consisting of the eigenvalues of T, with each eigenvalue included as many times as its multiplicity—see 8.37. Thus the definition of the trace of an operator along with 8.50, which allows us to use a basis of our choice, implies that $\operatorname{tr} T$ equals the sum of the eigenvalues of T, with each eigenvalue included as many times as its multiplicity. ■

8.53 example: *trace of an operator on* $\mathbf{C}^3$

Suppose $T \in \mathcal{L}(\mathbf{C}^3)$ is defined by

$$T(z_1, z_2, z_3) = (3z_1 - z_2 - 2z_3, 3z_1 + 2z_2 - 3z_3, z_1 + 2z_2).$$

Then the matrix of T with respect to the standard basis of $\mathbf{C}^3$ is

$$\begin{pmatrix} 3 & -1 & -2 \\ 3 & 2 & -3 \\ 1 & 2 & 0 \end{pmatrix}.$$

Adding up the diagonal entries of this matrix, we see that $\operatorname{tr} T = 5$.

The eigenvalues of T are 1, $2 + 3i$, and $2 - 3i$, each with multiplicity 1, as you can verify. The sum of these eigenvalues, each included as many times as its multiplicity, is $1 + (2 + 3i) + (2 - 3i)$, which equals 5, as expected by 8.52.

The trace has a close connection with the characteristic polynomial. Suppose $\mathbf{F} = \mathbf{C}$, $T \in \mathcal{L}(V)$, and $\lambda_1, \dots, \lambda_n$ are the eigenvalues of T, with each eigenvalue included as many times as its multiplicity. Then by definition (see 8.26), the characteristic polynomial of T equals

$$(z - \lambda_1) \cdots (z - \lambda_n).$$

Expanding the polynomial above, we can write the characteristic polynomial of T in the form

$$z^n - (\lambda_1 + \cdots + \lambda_n) z^{n-1} + \cdots + (-1)^n (\lambda_1 \cdots \lambda_n).$$

The expression above immediately leads to the next result. Also see 9.65, which does not require the hypothesis that $\mathbf{F} = \mathbf{C}$.

8.54 *trace and characteristic polynomial*

Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Let $n = \dim V$. Then $\operatorname{tr} T$ equals the negative of the coefficient of z^{n-1} in the characteristic polynomial of T.

The next result gives a nice formula for the trace of an operator on an inner product space.

8.55 *trace on an inner product space*

Suppose V is an inner product space, $T \in \mathcal{L}(V)$, and $e_1, \dots, e_n$ is an orthonormal basis of V. Then

$$\operatorname{tr} T = \langle Te_1, e_1 \rangle + \cdots + \langle Te_n, e_n \rangle.$$

Proof The desired formula follows from the observation that the entry in row k, column k of $\mathcal{M}\big(T, (e_1, \dots, e_n)\big)$ equals $\langle Te_k, e_k \rangle$ [use 6.30(a) with $v = Te_k$]. ■

The algebraic properties of the trace as defined on square matrices translate to algebraic properties of the trace as defined on operators, as shown in the next result.

8.56 *trace is linear*

The function $\operatorname{tr}\colon \mathcal{L}(V) \to \mathbf{F}$ is a linear functional on $\mathcal{L}(V)$ such that

$$\operatorname{tr}(ST) = \operatorname{tr}(TS)$$

for all $S, T \in \mathcal{L}(V)$.

Proof Choose a basis of V. All matrices of operators in this proof will be with respect to that basis. Suppose $S, T \in \mathcal{L}(V)$.

If $\lambda \in \mathbf{F}$, then

$$\operatorname{tr}(\lambda T) = \operatorname{tr} \mathcal{M}(\lambda T) = \operatorname{tr}\big(\lambda \mathcal{M}(T)\big) = \lambda \operatorname{tr} \mathcal{M}(T) = \lambda \operatorname{tr} T,$$

where the first and last equalities come from the definition of the trace of an operator, the second equality comes from 3.38, and the third equality follows from the definition of the trace of a square matrix.

Also,

$$\operatorname{tr}(S+T) = \operatorname{tr} \mathcal{M}(S+T) = \operatorname{tr}\big(\mathcal{M}(S)+\mathcal{M}(T)\big) = \operatorname{tr} \mathcal{M}(S) + \operatorname{tr} \mathcal{M}(T) = \operatorname{tr} S + \operatorname{tr} T,$$

where the first and last equalities come from the definition of the trace of an operator, the second equality comes from 3.35, and the third equality follows from the definition of the trace of a square matrix. The two paragraphs above show that $\operatorname{tr}\colon \mathcal{L}(V) \to \mathbf{F}$ is a linear functional on $\mathcal{L}(V)$.

Furthermore,

$$\operatorname{tr}(ST) = \operatorname{tr} \mathcal{M}(ST) = \operatorname{tr}\big(\mathcal{M}(S)\mathcal{M}(T)\big) = \operatorname{tr}\big(\mathcal{M}(T)\mathcal{M}(S)\big) = \operatorname{tr} \mathcal{M}(TS) = \operatorname{tr}(TS),$$

where the second and fourth equalities come from 3.43 and the crucial third equality comes from 8.49. ■

The equations $\operatorname{tr}(ST) = \operatorname{tr}(TS)$ and $\operatorname{tr} I = \dim V$ uniquely characterize the trace among the linear functionals on $\mathcal{L}(V)$—see Exercise 10.

The equation $\operatorname{tr}(ST) = \operatorname{tr}(TS)$ leads to our next result, which does not hold on infinite-dimensional vector spaces (see Exercise 13). However, additional hypotheses on S, T, and V lead to an infinite-dimensional generalization of the result below, with important applications to quantum theory.

The statement of the next result does not involve traces, but the short proof uses traces. When something like this happens in mathematics, then usually a good definition lurks in the background.

8.57 *identity operator is not the difference of ST and TS*

There do not exist operators $S, T \in \mathcal{L}(V)$ such that $ST - TS = I$.

Proof Suppose $S, T \in \mathcal{L}(V)$. Then

$$\operatorname{tr}(ST - TS) = \operatorname{tr}(ST) - \operatorname{tr}(TS) = 0,$$

where both equalities come from 8.56. The trace of I equals $\dim V$, which is not 0. Because $ST - TS$ and I have different traces, they cannot be equal. ■

Exercises 8D

1 Suppose V is an inner product space and $v, w \in V$. Define an operator $T \in \mathcal{L}(V)$ by $Tu = \langle u, v \rangle w$. Find a formula for $\operatorname{tr} T$.

2 Suppose $P \in \mathcal{L}(V)$ satisfies $P^2 = P$. Prove that

$$\operatorname{tr} P = \dim \operatorname{range} P.$$

3 Suppose $T \in \mathcal{L}(V)$ and $T^5 = T$. Prove that the real and imaginary parts of $\operatorname{tr} T$ are both integers.

4 Suppose V is an inner product space and $T \in \mathcal{L}(V)$. Prove that

$$\operatorname{tr} T^* = \overline{\operatorname{tr} T}.$$

5 Suppose V is an inner product space. Suppose $T \in \mathcal{L}(V)$ is a positive operator and $\operatorname{tr} T = 0$. Prove that $T = 0$.

6 Suppose V is an inner product space and $P, Q \in \mathcal{L}(V)$ are orthogonal projections. Prove that $\operatorname{tr}(PQ) \geq 0$.

7 Suppose $T \in \mathcal{L}(\mathbf{C}^3)$ is the operator whose matrix is

$$\begin{pmatrix} 51 & -12 & -21 \\ 60 & -40 & -28 \\ 57 & -68 & 1 \end{pmatrix}.$$

Someone tells you (accurately) that -48 and 24 are eigenvalues of T. Without using a computer or writing anything down, find the third eigenvalue of T.

8 Prove or give a counterexample: If $S, T \in \mathcal{L}(V)$, then $\operatorname{tr}(ST) = (\operatorname{tr} S)(\operatorname{tr} T)$.

9 Suppose $T \in \mathcal{L}(V)$ is such that $\operatorname{tr}(ST) = 0$ for all $S \in \mathcal{L}(V)$. Prove that $T = 0$.

10 Prove that the trace is the only linear functional $\tau \colon \mathcal{L}(V) \to \mathbf{F}$ such that

$$\tau(ST) = \tau(TS)$$

for all $S, T \in \mathcal{L}(V)$ and $\tau(I) = \dim V$.

Hint: Suppose that $v_1, \dots, v_n$ is a basis of V. For $j, k \in \{1, \dots, n\}$, define $P_{j,k} \in \mathcal{L}(V)$ by $P_{j,k}(a_1 v_1 + \dots + a_n v_n) = a_k v_j$. Prove that

$$\tau(P_{j,k}) = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases}$$

Then for $T \in \mathcal{L}(V)$, use the equation $T = \sum_{k=1}^{n} \sum_{j=1}^{n} \mathcal{M}(T)_{j,k} P_{j,k}$ to show that $\tau(T) = \operatorname{tr} T$.

11 Suppose V and W are inner product spaces and $T \in \mathcal{L}(V, W)$. Prove that if $e_1, \dots, e_n$ is an orthonormal basis of V and $f_1, \dots, f_m$ is an orthonormal basis of W, then

$$\operatorname{tr}\bigl(T^*T\bigr) = \sum_{k=1}^{n} \sum_{j=1}^{m} |\langle Te_k, f_j \rangle|^2.$$

The numbers $\langle Te_k, f_j \rangle$ are the entries of the matrix of T with respect to the orthonormal bases $e_1, \dots, e_n$ and $f_1, \dots, f_m$. These numbers depend on the bases, but $\operatorname{tr}(T^*T)$ does not depend on a choice of bases. Thus this exercise shows that the sum of the squares of the absolute values of the matrix entries does not depend on which orthonormal bases are used.

12 Suppose V and W are finite-dimensional inner product spaces.

(a) Prove that $\langle S, T \rangle = \operatorname{tr}\bigl(T^*S\bigr)$ defines an inner product on $\mathcal{L}(V, W)$.

(b) Suppose $e_1, \dots, e_n$ is an orthonormal basis of V and $f_1, \dots, f_m$ is an orthonormal basis of W. Show that the inner product on $\mathcal{L}(V, W)$ from (a) is the same as the standard inner product on $\mathbf{F}^{mn}$, where we identify each element of $\mathcal{L}(V, W)$ with its matrix (with respect to the bases just mentioned) and then with an element of $\mathbf{F}^{mn}$.

Caution: The norm of a linear map $T \in \mathcal{L}(V, W)$ as defined by 7.86 is not the same as the norm that comes from the inner product in (a) above. Unless explicitly stated otherwise, always assume that $\|T\|$ refers to the norm as defined by 7.86. The norm that comes from the inner product in (a) is called the ***Frobenius norm*** *or the* ***Hilbert–Schmidt norm****.*

13 Find $S, T \in \mathcal{L}\bigl(\mathcal{P}(\mathbf{F})\bigr)$ such that $ST - TS = I$.

Hint: Make an appropriate modification of the operators in Example 3.9.

This exercise shows that additional hypotheses are needed on S and T to extend 8.57 to the setting of infinite-dimensional vector spaces.