Chapter 10

Trace and Determinant

10.A Trace

10/27:

- To study the trace and determinant, we'll need to know how $\mathcal{M}(T,(v_1,\dots,v_n))$ (for $T\in\mathcal{L}(V)$) changes as $v_1,\dots,v_n$ changes.
- **n-by-n identity matrix**: The matrix of the identity operator $I\in\mathcal{L}(V)$. *Denoted by* $\boldsymbol{I}$. *Given by*

$$\begin{pmatrix}1 & & 0\\ & \ddots & \\ 0 & & 1\end{pmatrix}$$

  - $\mathcal{M}(I)$ is the same with respect to every basis of V.
- **Invertible** (matrix A): A square matrix A for which there exists a square matrix B of identical size such that $AB=BA=I$. *Also known as* **nonsingular**.
- **Inverse** (of an invertible matrix A): The unique matrix B in the above definition. *Denoted by* $\boldsymbol{A^{-1}}$.
  - The "unique" part of this definition follows from a proof symmetric to that of Theorem 3.12.
- **Singular** (matrix A): A matrix A that is not invertible. *Also known as* **noninvertible**.
- The following result is connected to Theorem 3.11.

  Theorem 10.1. *Suppose $u_1,\dots,u_n$, $v_1,\dots,v_n$, and $w_1,\dots,w_n$ are all bases of V. Suppose $S,T\in\mathcal{L}(V)$. Then*

$$\mathcal{M}(ST,(u_1,\dots,u_n),(w_1,\dots,w_n))=\mathcal{M}(S,(v_1,\dots,v_n),(w_1,\dots,w_n))\mathcal{M}(T,(u_1,\dots,u_n),(v_1,\dots,v_n))$$

- We now discuss the matrix of the identity operator with respect to two bases.

  Theorem 10.2. *Suppose $u_1,\dots,u_n$ and $v_1,\dots,v_n$ are bases of V. Then the matrices*

$$\mathcal{M}(I,(u_1,\dots,u_n),(v_1,\dots,v_n))\qquad\qquad\mathcal{M}(I,(v_1,\dots,v_n),(u_1,\dots,u_n))$$

  are invertible, and each is the inverse of the other.

  Proof. It follows from Theorem 10.1 that

$$I=\mathcal{M}(I,(u_1,\dots,u_n),(v_1,\dots,v_n))\mathcal{M}(I,(v_1,\dots,v_n),(u_1,\dots,u_n))$$

  and

$$I=\mathcal{M}(I,(v_1,\dots,v_n),(u_1,\dots,u_n))\mathcal{M}(I,(u_1,\dots,u_n),(v_1,\dots,v_n))$$

  as desired. ∎

- It follows that the above matrices change the coordinates of a vector in V from one basis to another.
- We now discuss change of basis for an operator.

Theorem 10.3. *Suppose $T \in \mathcal{L}(V)$. Let $u_1, \dots, u_n$ and $v_1, \dots, v_n$ be bases of V. Let*

$$A = \mathcal{M}(I, (u_1, \dots, u_n), (v_1, \dots, v_n))$$

Then

$$\mathcal{M}(T, (u_1, \dots, u_n)) = A^{-1}\mathcal{M}(T, (v_1, \dots, v_n))A$$

Proof. We have that

$$\begin{aligned}
\mathcal{M}(T, (u_1, \dots, u_n)) &= \mathcal{M}(IT, (u_1, \dots, u_n), (u_1, \dots, u_n)) \\
&= \mathcal{M}(I, (v_1, \dots, v_n), (u_1, \dots, u_n))\mathcal{M}(T, (u_1, \dots, u_n), (v_1, \dots, v_n)) && \text{Theorem 10.1} \\
&= A^{-1}\mathcal{M}(T, (u_1, \dots, u_n), (v_1, \dots, v_n)) && \text{Theorem 10.2}
\end{aligned}$$

We also have that

$$\begin{aligned}
\mathcal{M}(T, (u_1, \dots, u_n), (v_1, \dots, v_n)) &= \mathcal{M}(TI, (u_1, \dots, u_n), (v_1, \dots, u_n)) \\
&= \mathcal{M}(T, (v_1, \dots, v_n), (v_1, \dots, v_n))\mathcal{M}(I, (u_1, \dots, u_n), (v_1, \dots, v_n)) \\
& && \text{Theorem 10.1} \\
&= \mathcal{M}(T, (v_1, \dots, v_n))A
\end{aligned}$$

Substituting the second equation into the first gives the desired results. ■

- **Trace** (of $T \in \mathcal{L}(V)$, V complex): The sum of the eigenvalues of T with each eigenvalue repeated according to its multiplicity. *Denoted by* $\textbf{trace}\,\boldsymbol{T}$.
- **Trace** (of $T \in \mathcal{L}(V)$, V real): The sum of the eigenvalues of $T_{\mathbb{C}}$ with each eigenvalue repeated according to its multiplicity. *Denoted by* $\textbf{trace}\,\boldsymbol{T}$.
- Suppose $T \in \mathcal{L}(V)$. Let $n = \dim V$. Then $\operatorname{trace} T$ equals the negative of the coefficient of z^{n-1} in the characteristic polynomial of T.
- **Trace** (of A): The sum of the diagonal entries of a square matrix A. *Denoted by* $\textbf{trace}\,\boldsymbol{A}$.
- We now build up to proving that $\operatorname{trace} T = \operatorname{trace} \mathcal{M}(T, (v_1, \dots, v_n))$ where $v_1, \dots, v_n$ is an arbitrary basis of V.

Theorem 10.4. *If A and B are square matrices of the same size, then*

$$\operatorname{trace}(AB) = \operatorname{trace}(BA)$$

Proof. Let

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

The j^{th} diagonal entry of AB is by the definition of matrix multiplication $\sum_{k=1}^{n} A_{j,k}B_{k,j}$. Thus,

$$\begin{aligned}
\operatorname{trace}(AB) &= \sum_{j=1}^{n}\sum_{k=1}^{n} A_{j,k}B_{k,j} \\
&= \sum_{k=1}^{n}\sum_{j=1}^{n} B_{k,j}A_{j,k} \\
&= \operatorname{trace}(BA)
\end{aligned}$$

as desired. ■

- We now prove that the trace of a matrix is unique up to change of basis.

Theorem 10.5. *Let $T \in \mathcal{L}(V)$. Suppose $u_1, \dots, u_n$ and $v_1, \dots, v_n$ are bases of V. Then*

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

Proof. Let $A = \mathcal{M}(I, (u_1, \dots, u_n), (v_1, \dots, v_n))$. Then

$$\begin{aligned}
\mathcal{M}(T, (u_1, \dots, u_n)) &= \operatorname{trace}(A^{-1}(\mathcal{M}(T, (v_1, \dots, v_n)))A) && \text{Theorem 10.3} \\
&= \operatorname{trace}((\mathcal{M}(T, (v_1, \dots, v_n)))A^{-1}A) && \text{Theorem 10.4} \\
&= \operatorname{trace} \mathcal{M}(T, (v_1, \dots, v_n))
\end{aligned}$$

as desired. ■

- We can now prove the main result.

Theorem 10.6. *Suppose $T \in \mathcal{L}(V)$. Then $\operatorname{trace} T = \operatorname{trace} \mathcal{M}(T)$.*

Proof. By Theorem 10.5, $\operatorname{trace} \mathcal{M}(T)$ is independent of which basis of V we choose. Thus, to prove that $\operatorname{trace} T = \operatorname{trace} \mathcal{M}(T)$, it will suffice to prove the equality for any basis of V.

Let $v_1, \dots, v_n$ be the basis of V specified by Theorem 8.13. It follows that $\operatorname{trace} \mathcal{M}(T) = d_1\lambda_1 + \cdots + d_m\lambda_m$ where $\lambda_1, \dots, \lambda_m$ are the eigenvalues of T and $d_1, \dots, d_m$ are there respective multiplicities. But this is just $\operatorname{trace} T$ if V is complex and $\operatorname{trace} T_{\mathbb{C}}$ if V is real, as desired. ■

- Note that the statement of Theorem 10.6 does not specify a basis because trace is invariant under change of basis, as proven in Theorem 10.5.

- The trace is additive.

Theorem 10.7. *Suppose $S, T \in \mathcal{L}(V)$. Then $\operatorname{trace}(S + T) = \operatorname{trace} S + \operatorname{trace} T$.*

Proof. We have that

$$\begin{aligned}
\operatorname{trace}(S + T) &= \operatorname{trace} \mathcal{M}(S + T) && \text{Theorem 10.6} \\
&= \operatorname{trace}(\mathcal{M}(S) + \mathcal{M}(T)) \\
&= \operatorname{trace} \mathcal{M}(S) + \operatorname{trace} \mathcal{M}(T) \\
&= \operatorname{trace} S + \operatorname{trace} T && \text{Theorem 10.6}
\end{aligned}$$

as desired. ■

- We now state a curious consequence of the previous theorems that has important applications to quantum theory.

Theorem 10.8. *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

$$\begin{aligned}
\operatorname{trace}(ST - TS) &= \operatorname{trace}(ST) - \operatorname{trace}(TS) && \text{Theorem 10.7} \\
&= \operatorname{trace} \mathcal{M}(ST) - \operatorname{trace} \mathcal{M}(TS) && \text{Theorem 10.6} \\
&= \operatorname{trace} \mathcal{M}(S)\mathcal{M}(T) - \operatorname{trace} \mathcal{M}(T)\mathcal{M}(S) && \text{Theorem 3.11} \\
&= \operatorname{trace} \mathcal{M}(S)\mathcal{M}(T) - \operatorname{trace} \mathcal{M}(S)\mathcal{M}(T) && \text{Theorem 10.4} \\
&= 0
\end{aligned}$$

Since $\operatorname{trace} I > 0$ necessarily, $\operatorname{trace}(ST - TS) \neq \operatorname{trace} I$. It follows that $ST - TS \neq I$, as desired. ■

10.B Determinant

- **Determinant** (of $T \in \mathcal{L}(V)$, V complex): The product of the eigenvalues of T with each eigenvalue repeated according to its multiplicity. *Denoted by* $\mathbf{det}\,\boldsymbol{T}$.
- **Determinant** (of $T \in \mathcal{L}(V)$, V real): The product of the eigenvalues of $T_{\mathbb{C}}$ with each eigenvalue repeated according to its multiplicity. *Denoted by* $\mathbf{det}\,\boldsymbol{T}$.
- If $\lambda_1, \dots, \lambda_m$ are the distinct eigenvalues of T (or $T_{\mathbb{C}}$ if V is real) with corresponding multiplicities $d_1, \dots, d_m$, then
$$\det T = \lambda_1^{d_1} \cdots \lambda_m^{d_m}$$
- Suppose $T \in \mathcal{L}(V)$. Let $n = \dim V$. Then $\det T$ is $(-1)^n$ times the constant term of the characteristic polynomial of T.
- Invertibility and determinant.

  Theorem 10.9. *An operator on V is invertible if and only if its determinant is nonzero.*

  Proof. Let $T \in \mathcal{L}(V)$. We divide into two cases (V is complex and V is real). Let's begin.

  Suppose first that V is complex. By Theorem 5.7, there is a basis of V with respect to which $\mathcal{M}(T)$ is upper triangular. By Theorem 5.8, T is invertible iff all diagonal entries of $\mathcal{M}(T)$ are nonzero. By Theorem 5.9, all diagonal entries of $\mathcal{M}(T)$ are nonzero iff all eigenvalues of T are nonzero. But this is true iff the product of the eigenvalues of T, i.e., $\det T$ is nonzero, as desired.

  Now suppose that V is real. As before, T is invertible iff 0 is not an eigenvalue of $T_{\mathbb{C}}$. But by Theorem 9.4, it follows in both directions that 0 is not an eigenvalue of T, so $\det T \neq 0$ in this case too, as desired. ∎

- Characteristic polynomial and determinant.

  Theorem 10.10. *Suppose $T \in \mathcal{L}(V)$. Then the characteristic polynomial of T equals* $\det(zI - T)$.

  Proof. Suppose first that V is complex. We know that λ is an eigenvalue of T iff $z - \lambda$ is an eigenvalue of $zI - T$:
$$-(T - \lambda I) = 0 = (zI - T) - (z - \lambda)I$$
  Raising both sides to the $\dim V$ power and taking null spaces proves that the multiplicity of λ wrt. T equals the multiplicity of $z - \lambda$ wrt. $zI - T$. It follows that
$$\det(zI - T) = (z - \lambda_1)^{d_1} \cdots (z - \lambda_n)^{d_n}$$
  which is the characteristic polynomial, as desired.

  The real case follows from applying the complex case to $T_{\mathbb{C}}$. ∎

- **Permutation** (of $(1, \dots, n)$): A list $(m_1, \dots, m_n)$ that contains each of the numbers $1, \dots, n$ exactly once.
- $\mathbf{perm}\,\boldsymbol{n}$: The set of all permutations of $(1, \dots, n)$.
- **Sign** (of a permutation $(m_1, \dots, m_n)$): The number 1 if the number of pairs of integers (j, k) with $1 \leq j < k \leq n$ such that j appears after k in the permutation is even, and the number -1 otherwise (e.g., if the number of such pairs is odd). *Denoted by* $\mathbf{sign}\,\boldsymbol{n}$. *Also known as* **signum**.
  - "In other words, the sign of a permutaton equals 1 if the natural order has been changed an even number of times and equals -1 if the natural order has been changed an odd number of times" (Axler, 2015, p. 313).

- We now prove a connection between the sign and transpositions.

  Theorem 10.11. *Interchanging two entries in a permutation multiplies the sign of the permutation by* -1.

- **Determinant** (of A): The following quantity. *Denoted by* $\det \boldsymbol{A}$. *Given by*

$$\det A = \sum_{(m_1,\dots,m_n)\in \operatorname{perm} n} (\operatorname{sign}(m_1,\dots,m_n))A_{m_1,1}\cdots A_{m_n,n}$$

- We now build up to proving that the determinant of A is invariant with respect to basis.
- Interchanging two columns.

  Theorem 10.12. *Suppose A is a square matrix and B is the matrix obtained from A by interchanging two columns. Then*

$$\det A = -\det B$$

  Proof. Notice that the same products appear in the sum defining the determinants of both matrices. However, the terms appear in different orders; in fact, each term has a unique transposition. Thus, every term of $\det B$ is -1 times the corresponding term in $\det A$ by Theorem 10.11. It follows by factoring out the -1's that $\det A = -\det B$. ■

- If $\mathcal{M}(T)$ has two equal columns, then T is not injective hence not invertible, so $\det = 0$. Similarly...

  Theorem 10.13. *If A is a square matrix that has two equal columns, then* $\det A = 0$.

  Proof. By the definition of A, interchanging the two equal columns of A gives A. But by Theorem 10.12, this implies that

$$\begin{aligned}\det A &= -\det A\\ 2\det A &= 0\\ \det A &= 0\end{aligned}$$

  as desired. ■

- We now generalize Theorem 10.12.

  Theorem 10.14. *Suppose $A = \begin{pmatrix} A_{\cdot,1} & \cdots & A_{\cdot,n}\end{pmatrix}$ is an $n\times n$ matrix and $(m_1,\dots,m_n)$ is a permutation. Then*

$$\det\begin{pmatrix} A_{\cdot,m_1} & \cdots & A_{\cdot,m_n}\end{pmatrix} = (\operatorname{sign}(m_1,\dots,m_n))\det A$$

  Proof. Change A into $\begin{pmatrix} A_{\cdot,m_1} & \cdots & A_{\cdot,m_n}\end{pmatrix}$ iteratively, one column switch at a time, and apply Theorems 10.12 and 10.11. ■

- The determinant is linear.

  Theorem 10.15. *Suppose k,n are positive integers with $1\leq k\leq n$. Fix $n\times 1$ matrices $A_{\cdot,1},\dots,A_{\cdot,n}$ except $A_{\cdot,k}$. Then the function that takes an $n\times 1$ column vector $A_{\cdot,k}$ to*

$$\det\begin{pmatrix} A_{\cdot,1} & \cdots & A_{\cdot,k} & \cdots & A_{\cdot,n}\end{pmatrix}$$

  is a linear map from the vector space of $n\times 1$ matrices with entries in $\mathbb{F}$ to $\mathbb{F}$.

  Proof. The linearity follows from the definition, where each term in the sum contains precisely one entry from the k^{th} column of A. ■

- The determinant of the product of two matrices is equal to the product of the determinants[1].

  Theorem 10.16. *Suppose A, B are square matrices of the same size. Then*

  $$\det(AB) = \det(BA) = (\det A)(\det B)$$

  Proof. Given, but complicated. See LinAlgGIEPNotes on Browne. ■

- We can now prove that the determinant is independent of basis.

  Theorem 10.17. *Suppose $T \in \mathcal{L}(V)$. Suppose $u_1, \dots, u_n$ and $v_1, \dots, v_n$ are bases of V. Then*

  $$\det \mathcal{M}(T, (u_1, \dots, u_n)) = \det \mathcal{M}(T, (v_1, \dots, v_n))$$

  Proof. Invoke Theorem 10.16 in a paradigm symmetric to that used in the proof of Theorem 10.5. ■

- We can now prove that the determinant of an operator is equal to the determinant of any of its matrices.

  Theorem 10.18. *Suppose $T \in \mathcal{L}(V)$. Then $\det T = \det \mathcal{M}(T)$.*

  Proof. Invoke Theorem 10.16 in a paradigm symmetric to that used in the proof of Theorem 10.6. ■

- Like the trace is additive, the determinant is multiplicative.

  Theorem 10.19. *Suppose $S, T \in \mathcal{L}(V)$. Then*

  $$\det(ST) = \det(TS) = (\det S)(\det T)$$

  Proof. Invoke Theorems 10.18 and 10.16 in a paradigm symmetric to that used in the proof of Theorem 10.7. ■

- We now transition from discussing properties of the determinant to applications.
- Determinant of an isometry.

  Theorem 10.20. *Suppose V is an inner product space and $S \in \mathcal{L}(V)$ is an isometry. Then*

  $$|\det S| = 1$$

  Proof. Suppose first that V is complex. Then by Theorem 7.20, every eigenvalue of S has absolute value 1. Therefore, by the definition of the determinant as the product of the eigenvalues, we have that

  $$|\det S| = |\lambda_1| \cdots |\lambda_m| = 1$$

  as desired.

  Now suppose that V is real. Applying the complexification, we have that $|\det S_{\mathbb{C}}| = 1$ and $\det S = \det S_{\mathbb{C}}$, as desired. ■

- We have that $\det \sqrt{T^*T} \geq 0$ as a positive operator with all positive eigenvalues.
- We now further investigate the relation between T and $\sqrt{T^*T}$ with respect to the determinant.

  Theorem 10.21. *Suppose V is an inner product space and $T \in \mathcal{L}(V)$. Then*

  $$|\det T| = \det \sqrt{T^*T}$$

[1] The first proof of this theorem was given in 1812 by French mathematicians Jacques Binet and Augustin-Louis Cauchy.

Proof. We have by the Polar Decomposition that there exists an isometry $S \in \mathcal{L}(V)$ such that $T = S\sqrt{T^*T}$. Thus

$$\begin{aligned} |\det T| &= |\det S| \cdot \det \sqrt{T^*T} && \text{Theorem 10.16} \\ &= \det \sqrt{T^*T} && \text{Theorem 10.20} \end{aligned}$$

as desired. ■

- Axler (2015) now discusses applications of the determinant to volume in $\mathbb{R}^n$.
- If $\Omega \subset \mathbb{R}^n$, then the volume of $T(\Omega)$ (where T is a positive operator) equals $\det T$ times the volume of Ω.
- Isometries don't change volume.
- If T is an *arbitrary* operator, then the volume of $T(\Omega)$ equals $|\det T|$ times the volume of Ω.
- Integrals and derivatives are discussed.
- Talks about the Jacobian and change of coordinates.