Lecture Notes for Math 623
Matrix Analysis

Michael E. O'Sullivan
mosullivan@mail.sdsu.edu
www-rohan.sdsu.edu/ ˜ mosulliv

April 18, 2013

1 Trace of an Operator

Let V be an n-dimensional vector space over $\mathbb{C}$.

Definition 1.1. Let $\varphi : V \longrightarrow V$ be a linear transformation with eigenvalues $\lambda_1, \ldots, \lambda_n$. The *trace* of φ is $\mathrm{Tr}(\varphi) = \sum_{i=1}^n \lambda_i$.

Let A be an $n \times n$ matrix the *trace* of A is $\mathrm{Tr}(A) = \sum_{i=1}^n a_{ii}$

We will show that for any basis $\mathcal{U}$ of V, the trace of ${}_{\mathcal{U}}[\varphi]_{\mathcal{U}}$ is equal to $\mathrm{Tr}\,\varphi$, but for now, the trace of a matrix is defined independently of the trace of an operator.

Proposition 1.2. *Let* $A, B \in \mathcal{M}_n$. $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.

Proof. $(AB)_{ik} = \sum_{j=1}^n a_{ij} b_{jk}$ so

$$\begin{aligned} \mathrm{Tr}(AB) = \sum_{i=1}^n (AB)_{ii} &= \sum_{i=1}^n \sum_{j=1}^n a_{ij} b_{ji} \\ &= \sum_{i=1}^n \sum_{j=1}^n b_{ji} a_{ij} = \mathrm{Tr}(BA) \end{aligned}$$

□

Corollary 1.3. *If A and B are similar,* $\mathrm{Tr}(A) = \mathrm{Tr}(B)$.

Proof. Let $A = S^{-1}BS$. Then

$$\mathrm{Tr}(A) = \mathrm{Tr}\left((S^{-1}B)S\right) = \mathrm{Tr}\left(S(S^{-1}B)\right) = \mathrm{Tr}\left(SS^{-1})B)\right) = \mathrm{Tr}(B)$$

□

Now we can link the two notions of trace.

Corollary 1.4. *The trace of a matrix is the sum of the eigenvalues of the matrix.*

Let φ be a linear transformation on V. For any basis $\mathcal{U}$ of the vector space V, ${}_{\mathcal{U}}[\varphi]_{\mathcal{U}} = \mathrm{Tr}(\varphi)$.

Proof. Let A be an $n \times n$ matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$. Let S be such that $S^{-1}AS$ is in Jordan form. Then $\mathrm{Tr}(A) = \mathrm{Tr}(S^{-1}AS)$ by the previous corollary. The Jordan matrix is upper triangular with the eigenvalues $\lambda_1, \ldots, \lambda_n$ on the diagonal. latter is the sum of the eigenvalues of A.

Let φ be a linear transformation on V. Different bases for V will yield different matrices for φ, but they will all be similar, with the same eigenvalues as φ. □

Proposition 1.5. *Trace is a linear operator:* $\mathrm{Tr}(A + B) = \mathrm{Tr}(A) + \mathrm{Tr}(B)$ *for* $A, B \in \mathcal{M}_n$.

Proof. Clear from the trace formula for matrices. □

Corollary 1.6. *For* $A, B \in \mathcal{M}_n$, $\mathrm{Tr}(AB - BA) = 0$. *Consequently, it is impossible for* $AB - BA = I$.

Proof. We have $\mathrm{Tr}(AB - BA) = \mathrm{Tr}(AB) - \mathrm{Tr}(BA) = 0$ by linearity and an earlier corollary. □

The expression $AB - BA$ is called the *commutator of A and B* (when it is nonzero A, B don't commute, so it is a measure of their noncommutability). This from

`{\tt http://en.wikipedia.org/wiki/Uncertainty_principle}`

"In matrix mechanics, observables such as position and momentum are represented by self-adjoint operators. When considering pairs of observables, one of the most important quantities is the commutator." For example, there is an operator for position and an operator for momentum and their commutator gives Heisenberg's uncertainty principle. Axler says that the previous corollary has important consequences in quantum theory.

2 Determinants

Definition 2.1. The determinant of a matrix is defined to be $\mathrm{Det}(A) = \sum_{\sigma\in S_n}(-1)^{\mathrm{sgn}(\sigma)}\prod_{i=1}^n a_{i,\sigma(i)}$.

The following propositions is in most textbooks.

Proposition 2.2. *There are three fundamental properties of determinants.*

(1) $\mathrm{Det}(I) = 1$ *for the identity matrix* I.

(2) Det *is alternating. If we transpose to columns (or rows) of* A *the determinant of the new matrix has determinant* $-\mathrm{Det}(A)$.

(3) Det *is multilinear on the columns (or rows). This means, relative to the first column (and similarly for the other columns),*

$$\begin{aligned}&\mathrm{Det}\left(\begin{bmatrix}u_1+av_1 & u_2 & u_3 & \dots & u_n\end{bmatrix}\right)\\ &\qquad = \mathrm{Det}\left(\begin{bmatrix}u_1 & u_2 & u_3 & \dots & u_n\end{bmatrix}\right) + a\,\mathrm{Det}\left(\begin{bmatrix}v_1 & u_2 & u_3 & \dots & u_n\end{bmatrix}\right)\end{aligned}$$

These properties uniquely define the determinant: the only function on matrices with these properties is the determinant.

An important consequence of the alternating property is that a matrix with two rows that are equal has determinant 0.

Proposition 2.3. $\mathrm{Det}(AB) = \mathrm{Det}(A)\,\mathrm{Det}(B)$

Proof. Let $b_1, \dots, b_n$ be the columns of B and let $e_1, \dots, e_n$ be the standard basis vectors.

$$\mathrm{Det}(AB) = \mathrm{Det}\left(\begin{bmatrix}Ab_1 & Ab_2 & Ab_3 & \dots & Ab_n\end{bmatrix}\right)$$

Using multilinearity and $Ab_i = \sum_{i=1}^n b_{m_i,i}Ae_{m_i}$

$$\mathrm{Det}(AB) = \sum_{m_1=1}^n\sum_{m_2=1}^n\cdots\sum_{m_n=1}^n\prod_{i=1}^n b_{m_i,i}\,\mathrm{Det}\left(\begin{bmatrix}Ae_{m_1} & Ae_{m_2} & Ae_{m_3} & \dots & Ae_{m_n}\end{bmatrix}\right)$$

If any $m_i = m_j$ the determinant will be 0, so we may restrict to $m_1, \dots, m_n$ being a permutations of $1, \dots, n$

$$\begin{aligned}&= \sum_{\sigma\in S_n}\prod_{i=1}^n b_{\sigma(i),i}\,\mathrm{Det}\left(\begin{bmatrix}Ae_{\sigma(1)} & Ae_{\sigma(2)} & Ae_{\sigma(3)} & \dots & Ae_{\sigma(n)}\end{bmatrix}\right)\\ &= \sum_{\sigma\in S_n}\prod_{i=1}^n b_{\sigma(i),i}\,\mathrm{sgn}(\sigma)\,\mathrm{Det}(A)\\ &= \mathrm{Det}(A)\,\mathrm{Det}(B)\end{aligned}$$

□

Corollary 2.4. *If A and B are similar then* $\mathrm{Det}(A) = \mathrm{Det}(B)$*.*

Proof. Let $A = S^{-1}BS$. Then

$$\mathrm{Det}(A) = \mathrm{Det}\left((S^{-1}B)S\right) = \mathrm{Det}\left(S(S^{-1}B)\right) = \mathrm{Det}\left(SS^{-1})B)\right) = \mathrm{Det}(B)$$

□

Corollary 2.5. *Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of A, and let $p(x) = x^n + p_{n-1}x^{n-1} + \cdots + p_1 x + p_0 = \prod_{i=1}^n (x - \lambda_i)$ be the characteristic polynomial of A.*

$$\mathrm{Det}(A) = \prod_{i=1}^n \lambda_i = (-1)^n p_0$$

Proof. The characteristic polynomial has roots λ_i so factors as shown. The constant term is $\prod_{i=1}^n (-\lambda_i)$. Since A is similar to its Jordan form its determinant is the same as its Jordan from. The determinant of the Jordan form is the product of the eigenvalues, because the determinant of any upper triangular matrix is the product of the diagonal entries. □

Corollary 2.6. *The characteristic polynomial of A is $p_A(t) = \det(xI - A)$.*

Proof. For any complex number z, the eigenvalues of $zI - A$ are $z - \lambda_i$, so the constant term of the characteristic polynomial of $zI - A$ is $\prod_{i=1}^n (-1)^n (z - \lambda_i) = \mathrm{Det}(zI - A)$. Now consider the indeterminate x, it is clear that $\prod_{i=1}^n (-1)^n (x - \lambda_i) = \mathrm{Det}(xI - A)$, since this formula holds when x is substituted with any complex number. □

3 Generalizations of Trace and Determinant

Definition 3.1. Let $I, J \subseteq \{1, \ldots, n\}$ and let $k = |I|$ and $m = |J|$. By $A(I, J)$ we indicate the $k \times m$ submatrix of A formed from the rows in I and the columns in J. We are particularly interested in $A(I, I)$, which is called a *principal submatrix* of A. For fixed k there are $\binom{n}{k}$ principal $k \times k$ submatrices of A. The *principal minor* associated to I is $\mathrm{Det}(A(I, I)$.

Definition 3.2. Let

$$E_k(A) = \sum_{\substack{I \subseteq \{1,\ldots,n\} \\ |I|=k}} \mathrm{Det}(A(I, I))$$

Observe that $E_n(A) = \text{Det}(A)$, and $E_1(A) = \sum_{i=1}^n a_{ii} = \text{Tr}(A)$. Similarly E_{n-1} is the sum of the n principal submatrices obtained by eliminating the ith row and ith column of A. By convention, we set $E_0(A) = 1$.

Proposition 3.3. $\text{Det}(xI - A) = \sum_{k=0}^n x^{n-k}(-1)^k E_k(A)$

Proof. In class □

Definition 3.4. The kth elementary symmetric function of $\lambda_1, \ldots, \lambda_n$ is

$$S_k(\lambda_1, \ldots \lambda_n) = \sum_{\substack{I \subseteq \{1,\ldots,n\} \\ |I|=k}} \prod_{i \in I} \lambda_i \qquad = \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} \lambda_{i_1} \lambda_{i_2} \ldots \lambda_{i_k}$$

Proposition 3.5.

$$p_A(x) = \prod (x - \lambda_i) = \sum_{k=0}^n \lambda^{n-k}(-1)^k S_k(\lambda_1, \ldots, \lambda_n)$$

Corollary 3.6. *Let $A \in \mathcal{M}_n$ have eigenvalues $\lambda_1, \ldots, \lambda_n$. Then $E_k(A) = S_k(\lambda_1, \ldots, \lambda_n)$.*

Proof.

$$p_A(x) = \prod_{i=1}^n (x - \lambda_i) = \sum x^{n-k}(-1)^k S_k(\lambda_1, \ldots, \lambda_n)$$

But also

$$p_A(x) = \text{Det}(xI - A) = \sum x^{n-k}(-1)^k E_k(A)$$

Consequently $E_k(A) = S_k(\lambda_1, \ldots, \lambda_n)$. □