Lecture 23: Trace and determinants! (1) (Final lecture)

Travis Schedler

Thurs, Dec 9, 2010 (version: Monday, Dec 13, 3:52 PM)

Goals (2)

• Recall \( \chi_T(x) = (x-\lambda_1) \cdots (x-\lambda_n) = x^n - \text{tr}(T)x^{n-1} + \cdots + (-1)^n \det(T). \)

• \( \text{tr}(T) = \) sum of eigenvalues. Theorem: equals sum of diagonal entries of \( M(T), \) independent of basis.

• Deduce: \( \text{tr}(S+T) = \text{tr}(S) + \text{tr}(T). \)

• \( \det(T) = \) product of eigenvalues. Theorem: equals a unique sum formula on \( M(T) \) using sign of permutations.

• Real case: Volume of \( T(R) = |\det(T)| \) times volume of \( R. \)

• Theorem: \( \det(ST) = \det(S) \det(T). \)

• Define \( \chi_T(x) := \det(xI - M(T)) \) for general \( F, \) independent of basis.

• Formula for \( A^{-1} \) using \( \det(A); \) Cayley-Hamilton theorem.

Trace (3)

**Definition 1.** For a matrix \( A = (a_{ij}) \) over an arbitrary \( F, \) define \( \text{tr}(A) := a_{11} + \cdots + a_{nn}. \)

It is immediate that \( \text{tr}(A+B) = \text{tr}(A) + \text{tr}(B). \)

**Theorem 2** (Corollary 10.10). *For all \( T \in \mathcal{L}(V), \) \( \text{tr}(M(T)) \) does not depend on the basis.*

• For \( F = \mathbb{C}, \) deduce: \( \text{tr}(M(T)) = \lambda_1 + \cdots + \lambda_n = \text{tr}(T). \)

• For general \( F, \) define \( \text{tr}(T) := \text{tr}(M(T)). \)

• Immediate: \( \text{tr}(S+T) = \text{tr}(S) + \text{tr}(T). \)

**Lemma 3** (Proposition 10.9). \( \text{tr}(AB) = \text{tr}(BA). \)
The theorem follows immediately from the lemma and the change-of-basis formula: \( \text{tr}(SAS^{-1}) = \text{tr}(S^{-1}SA) = \text{tr}(A) \).

**Proof of Lemma.** \( \text{tr}(AB) = \sum_{i,j} a_{ij}b_{ji} = \sum_{i,j} b_{ji}a_{ij} = \text{tr}(BA). \) \( \square \)

### Determinant and volume (4)

Properties of \( \det(T) = \lambda_1 \cdots \lambda_n \):

- \( \det(T) \neq 0 \) iff \( T \) is invertible.
- Let \( F = \mathbb{R} \). Theorem 10.38: \( \text{vol}(T(R)) = |\det(T)| \text{vol}(R). \)
- Suppose \( T \) is self-adjoint. Then, \( T \) dilates in orthogonal directions by \( \lambda_1, \ldots, \lambda_n. \) Taking these as axes, we deduce: for any region \( R, \text{vol}(T(R)) = |\det(T)| \text{vol}(R). \)
- For general \( T \), write \( T = S\sqrt{T^*T} \) where \( S \) is an isometry.
- By det theorem, \( \det(T) = \det(S) \det(\sqrt{T^*T}). \)
- Since \( S \) preserves volume, \( \text{vol}(T(R)) = \text{vol}(\sqrt{T^*T}(R)). \) But \( \sqrt{T^*T} \) is positive (\( \Rightarrow \) self-adjoint)!
- \( \text{vol}(T(R)) = \det(\sqrt{T^*T}) \text{vol}(R) = \det(S^{-1}) \det(T) \text{vol}(R). \)
- Finally, \( \det(S^{-1}) = \pm 1 \) since \( S^{-1} \) is an isometry (its block diagonal matrix has \( \pm 1 \) and rotation matrices as blocks).
- So \( \text{vol}(T(R)) = |\det(T)| \text{vol}(R). \)

### Determinant and alternating forms (5)

The determinant has the following essential properties:

- **Multiadditivity:** Let \( A \) and \( B \) be matrices that differ only in a fixed column. Let \( C \) be the matrix which is also the same as \( A \) and \( B \) except in this column, which is the sum of the columns appearing in \( A \) and \( B \). Then, \( \det(C) = \det(A) + \det(B). \)
  - Interpretation: For \( R = \) standard cube (coords 0 and 1), then \( C(R), A(R), \) and \( B(R) \) are parallelepipeds sharing a common base; the height of \( C(R) \) is the sum of the heights of \( A(R) \) and \( B(R) \).
- **Multihomogeneity:** Let \( B \) be obtained from \( A \) by multiplying a column by \( \lambda. \) Then \( \det(B) = \lambda \det(A). \)
  - Interpretation: As before, \( B(R) \) has the same base as \( A(R) \) and \( \lambda \) times the height.
- **Alternation:** Let \( A \) be a matrix with two identical columns. Then \( \det(A) = 0. \)
\(- A(R) \) is degenerate, so \( \text{vol}(A(R)) = 0 \). Alternatively, \( A \) is noninvertible \( \Rightarrow \det(A) = 0 \).

- **Theorem:** Our sum formula for determinant is the unique function with these properties, up to scaling.

**Theorem:** \( \det(AB) = \det(A) \det(B) \) \((6)\)

- Assume \( \det \) is the unique function with the above properties up to scaling; scale it so that \( \det(I) = 1 \).
- Let \( A \) be a fixed matrix. Define \( F_A(B) := \det(AB) \).
- Claim: \( F_A \) also has the above properties!
- Consequence: There exists \( \alpha \) such that \( \det(AB) = \alpha \det(B) \) for all \( B \).
- Plug in \( B = I \). Then \( \det(A) = \alpha \det(I) = \alpha \). So, \( \det(AB) = \det(A) \det(B) \)!
- Proof of claim: \( A \) is linear, so \( \det(A(v_1 \cdots v_n)) = \det(Av_1 \cdots Av_n) \) and column operations on \( v_1, \ldots, v_n \) become the same operations on \( (Av_1, \ldots, Av_n) \).
- Corollary: \( \det(AB) = \det(BA) \). Also, \( \det(SAS^{-1}) = \det(S^{-1}SA) = \det(A) \).
- So \( \det(T) := \det(\mathcal{M}(T)) \) makes sense (independent of basis)!
- So does \( \chi_T := \det(xI - \mathcal{M}(T)) \)!

**Proof of uniqueness theorem (7)**

- Let \( \dim V = n \) and suppose that \( F : V^n \to \mathbb{F} \) satisfies multilinearity (=multiadditivity+multihomogeneity) and alternation.
- Remark: Alternation implies skew-symmetry: \( 0 = F(u + v, u + v) - F(u, u) - F(v, v) = F(u, v) + F(v, u) \).
- Let \( v_1, \ldots, v_n \) be a basis of \( V \). Claim: \( F \) is uniquely determined by \( c := F(v_1, \ldots, v_n) \).
- Suppose \( w_j = \sum_i a_{i,j} v_i \). Then \( F(w_1, \ldots, w_n) = \sum_{\sigma : \{1, \ldots, n\} \to\{1, \ldots, n\}} F(a_{\sigma(1),1} v_{\sigma(1)}, \ldots, a_{\sigma(n),n} v_{\sigma(n)}) = \sum_{\sigma} a_{\sigma(1),1} \cdots a_{\sigma(n),n} F(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) \).
- Unless \( \sigma \) is a bijection, i.e., \( \sigma \in S_n := \text{permutations of } \{1, \ldots, n\} \), then two of these entries are equal so we get zero.
- By skew-symmetry, \( F(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) = (-1)^k c \) where \( k \) is the number of swaps of entries needed to reorder \( (v_{\sigma(1)}, \ldots, v_{\sigma(n)}) \) to \( (v_1, \ldots, v_n) \).
- We get uniqueness!
Existence and formula for determinant (8)

- Set \( F(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) = \text{sign}(\sigma)c \). We need sign(\( \sigma \)) = \( \pm 1 \) to be 1 if \( \sigma \) requires an even number of swaps and \(-1\) if it requires an odd number of swaps.

- If such a sign exists, then sign(\( \sigma \circ \tau \)) = sign(\( \sigma \)) sign(\( \tau \)) automatically. This is the essence of determinant: a multiplicative function on permutation matrices!

- Existence: define sign(\( \sigma \)) = \((-1)^{o(\sigma)}\), where \( o(\sigma) = \left| \left\{ (i, j) : i < j \text{ and } \sigma(i) > \sigma(j) \right\} \right| \).

- Lemma: If we swap adjacent entries of \((\sigma(1), \ldots, \sigma(n))\), then \( o(\sigma) \) goes up or down by one. So it changes parity (even to odd or vice-versa).

- Now if we swap entries \( i \) and \( j \), this takes \( 2|j - i| - 1 \) adjacent swaps, so \( o(\sigma) \) still changes parity!

- We deduce that \((-1)^{o(\sigma)} = (-1)^k\) for any sequence of \( k \) swaps which equals \( \sigma \). This proves existence!

Matrix formula for determinant (9)

- Fix a basis \( v_1, \ldots, v_n \) and let \( \text{det} \) be the unique multilinear alternating function so that \( \text{det}(v_1, \ldots, v_n) = 1 \).

- Given any \( T \in \mathcal{L}(V) \), define \( \text{det}(T) = \text{det}(w_1, \ldots, w_n) \) where \( T(v_i) = w_i \).

- Write \( A = M(v_j)(T) = (a_{i,j}) \). So \( w_j = \sum_i a_{i,j}v_i \).

- By our computation, \( \text{det}(T) = \sum_{\sigma \in S_n} \text{sign}(\sigma)a_{\sigma(1),1} \cdots a_{\sigma(n),n} \).

- Finally, for an upper triangular matrix, \( \text{det}(A) = \text{the product of the diagonal entries} \). So for \( F = C \), \( \text{det}(T) = \text{the product of the eigenvalues} \)!

Computing the determinant (and \( \chi_A(x) \)) (10)

- Gaussian elimination: to compute \( \text{det}(A) \) (or \( \text{det}(xI - A) \)), perform row operations.

- Multiplying a row by \( \lambda \) multiplies the determinant by \( \lambda \).

- If we swap two rows, that multiplies the determinant by \(-1\).

- If we add a multiple of a row to another row, that does not change the determinant.

- The determinant of an upper-triangular matrix is the product of the diagonal entries!
Inverses of matrices using \( \det \) (11)

- Note that \( \det(A) \neq 0 \) iff \( A \) is invertible: if it is invertible, \( \det(A) \det(A^{-1}) = \det(AA^{-1}) = 1 \) so \( \det(A) \neq 0 \); if it is not invertible, then its columns are linearly dependent so \( \det(A) = 0 \) as we explained.

- So can we get a formula for \( A^{-1} \) using \( \det \)? More precisely, we would like \( A^{-1} = \det(A) \cdot A' \) for some matrix \( A' \) whose entries are polynomial functions of \( A \).

- Indeed, let \( A' = (a'_{ij}) \) where \( a'_{ij} = (-1)^{i-j} \times \) the determinant of the \((n-1) \times (n-1)\) matrix obtained by striking out the \(i\)-th column and the \(j\)-th row.

- Claim: \( A'A = \det(A)I = AA' \). This proves the formula!

- Proof: \( (A')_i = \sum a'_{ik}a_{kj} = \det \) of the matrix obtained from \( A \) by replacing the \(i\)-th column with the \(j\)-th column.

  - This is zero unless \( i = j \), in which case we get \( \det \).

Properties of characteristic polynomial (12)

- Recall \( \chi_T := \chi_M(T) = \det(xI - M(T)) \).

- So, \( \chi_T(\lambda) = 0 \) iff \( \lambda \) is an eigenvalue (i.e., \( T - \lambda I \) is noninvertible). More generally:

  - Theorem: if \( f(x) \) is an irreducible polynomial, then \( f \) is a factor of \( \chi_T \) iff \( f(T) \) is noninvertible. See Optional Exercises #2 to get started.

Cayley-Hamilton theorem (13)

Theorem 4. \( \chi_T(T) = 0 \).

Proof (cf. Wikipedia)  

- \( \chi_A(x)I = \det(xI - A)I = (xI - A)'(xI - A) \).

- Tricky part: Now write \( (xI - A)' = B_0 + xB_1 + \ldots + x^{n-1}B_{n-1} \) where \( B_0, \ldots, B_{n-1} \) are all matrices with entries in \( F \), not polynomials.

- We find that \( \det(xI - A)I = B_0A + \sum_{i=1}^{n-1} (B_iA - B_{i-1})x^i + B_{n-1}x^n \).

- Write \( \chi_A(x) = \det(xI - A) = x^n + a_{n-1}x^{n-1} + \cdots + a_0 \).

- All the coefficients of both sides are equal: \( a_0I = B_0A, a_1I = B_1A - B_0, \) etc.

- Now we can plug in: \( \chi_A(A) = B_0A + \sum_{i=1}^{n-1} (B_iA - B_{i-1})A^i + B_{n-1}A^n \).

- All the terms on the RHS cancel, so \( \chi_A(A) = 0 \). \( \square \)
Further properties of characteristic polynomial (14)

- **Theorem:** The determinant of a block upper-triangular matrix is the product of the determinants of the diagonal blocks.

- **Proof:** Any permutation $\sigma$ with an entry $a_{\sigma(i),i}$ above the block diagonal has also an entry $a_{\sigma(j),j}$ below the block diagonal, which is zero.

- **Consequence:** if $U \subseteq V$ is $T$-invariant and $V = U \oplus U'$, then $\det(T) = \det(T|_U) \cdot \det(P_{U,U'}T|_{U'})$. Same for $\chi_T$.

- **Proof:** In basis for $V = U \oplus U'$, $\mathcal{M}(T) = \begin{pmatrix} \mathcal{M}(T|_U) & * \\ 0 & \mathcal{M}(P_{U',U}T|_{U'}) \end{pmatrix}$.

- In particular, if $U, U'$ are both $T$-invariant, $\det(T) = \det(T|_U) \det(T|_{U'})$ and $\text{tr}(T) = \text{tr}(T|_U) + \text{tr}(T|_{U'})$.

The general decomposition theorem (15)

- **Theorem:** $V = \bigoplus_f V(f)$, where $f$ ranges over irreducible polynomials such that $f(T)$ is noninvertible. Each $V(f)$ is a “generalized eigentuple,” and is $T$-invariant.

- **Proof:** Again, if $f(T)$ is noninvertible, then $V = \text{null} f^{\dim V}(T) \oplus \text{range} f^{\dim V}(T)$. Apply induction.

- **Restricted to** $V(f)$, $T$ has a block (lower)-triangular matrix with diagonal blocks all equal to

$$
\begin{pmatrix}
0 & 0 & \cdots & 0 & -a_0 \\
1 & 0 & \cdots & 0 & -a_1 \\
0 & 1 & \cdots & 0 & -a_2 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & 1 & -a_{k-1}
\end{pmatrix},
$$

where $f(x) = x^k + a_{k-1}x^{k-1} + \cdots + a_0$.

- **Jordan form:** We can make it so that all other entries are zero except for a single one in some corners between blocks.