THE GEOMETRY OF THE HIGHER TRACES

KENT E. MORRISON

ABSTRACT. The higher traces of a matrix are the coefficients of its characteristic polynomial. We show that for real matrices these coefficients have a geometric interpretation in terms of expected values of oriented volumes, generalizing the volume interpretation of the determinant.

Let A be an $n \times n$ matrix over $\mathbf{R}$. The higher k-trace of A is the coefficient τ_k in the characteristic polynomial of A

$$\det(\lambda I - A) = \sum_{k=0}^{n} (-1)^k \tau_k \lambda^{n-k}.$$

Note that $\tau_1 = \operatorname{tr} A$ and $\tau_n = \det A$. (We write $\tau_k(A)$ when the dependence on A needs to be explicit.)

It is a familiar fact that τ_n is the oriented volume of the image under multiplication by A of a unit cube in $\mathbf{R}^n$, for example, the cube spanned by any orthonormal basis $u_1, \ldots, u_n$. The other coefficients τ_k likewise have an interpretation in terms of the change in k-dimensional volumes under multiplication by A.

Let $V(n,k)$ be the Stiefel manifold of orthonormal k-frames in $\mathbf{R}^n$ consisting of points $u = (u_1, \ldots, u_k)$ where the u_i are mutually orthogonal unit vectors in $\mathbf{R}^n$. Let $\operatorname{span} u$ be the k-dimensional subspace spanned by the u_i. The vectors $Au_1, \ldots, Au_k$ span a parallelepiped. Project that parallelepiped orthogonally onto $\operatorname{span} u$ and define $T_k(u)$ to be the oriented k-volume of the result. The aim of this article is to prove that as u varies over $V(n,k)$, the average value of $T_k(u)$ is $\binom{n}{k}\tau_k$.

Factor the characteristic polynomial of A over the complex numbers

$$\det(\lambda I - A) = (\lambda - \lambda_1) \cdots (\lambda - \lambda_n).$$

Then the coefficient of λ^k is

$$(-1)^k \sum_{i_1 < i_2 < \cdots < i_k} \prod_{j=1}^{k} \lambda_j$$

so τ_k is the kth elementary symmetric function of the eigenvalues $\lambda_1, \ldots, \lambda_n$.

Date: September 15, 1998. Revised April 12, 2010. Address updated 2015.

Lemma 1. *Let $\Lambda^k A$ be the kth exterior power of A. Then $\tau_k = \mathrm{tr}\,(\Lambda^k A)$.*

Proof. Over $\mathbf{C}$ the matrix A is similar to a matrix in Jordan form. Since $\tau_k(A)$ and $\mathrm{tr}\,(\Lambda^k A)$ are invariant under similarity, we can assume that A is in Jordan form. Then either $Ae_i = \lambda_i e_i$ or $Ae_i = \lambda_i e_i + e_{i-1}$. Now

$$\mathrm{tr}\,(\Lambda^k A) = \sum_{(i)} a_{(i)(i)}$$

where (i) is the sequence $i_1 < i_2 < \cdots < i_k$, $e_{(i)} = e_{i_1} \wedge \cdots \wedge e_{i_k}$, and $(\Lambda^k A)e_{(j)} = \sum_{(i)} a_{(i)(j)} e_{(i)}$. In the expansion of $Ae_{i_1} \wedge \ldots Ae_{i_k}$, the basis element $e_{(i)} = e_{i_1} \wedge \cdots \wedge e_{i_k}$ occurs with coefficient $\lambda_{i_1} \ldots \lambda_{i_k}$. Summing over all (i) we get $\tau_k(A)$. □

Lemma 2. *If A is an $n\times n$ matrix, then τ_k is the sum of all k by k sub-determinants gotten by choosing k rows and the corresponding columns.*

Proof. Let $(i) = (i_1, \ldots, i_k)$ where $1 \leq i_1 < i_2 < \cdots < i_k \leq n$, and let $e_{(i)} = e_{i_1} \wedge \cdots \wedge e_{i_k}$ be the basis vectors of $\Lambda^k \mathbf{R}^n$. In the expansion of $(\Lambda^k A)e_{(i)}$ the coefficient of $e_{(i)}$ is the determinant of the $k \times k$ submatrix consisting of the rows and columns indexed by (i). The trace of $\Lambda^k A$, which is τ_k, is the sum of these determinants over all (i). □

Define $\Delta_k(A)$ to be the determinant of the upper left $k \times k$ submatrix of A. For σ in the permutation group S_n let P_σ be the permutation matrix that has a 1 in the ij position when $\sigma(j) = i$ and 0 otherwise. Thus, P_σ is the matrix of the linear transformation that permutes the standard basis vectors $e_1, \ldots, e_n$ of $V = K^n$ sending e_i to $e_{\sigma(i)}$. We note that $P_{\sigma^{-1}} = P_\sigma^{-1} = {P_\sigma}^t$.

Lemma 3. *If $A = (a_{ij})$, then the ij entry of $P_\sigma^{-1} A P_\sigma$ is $a_{\sigma(i),\sigma(j)}$.*

Proof. The ij entry of $P_\sigma A$ is $a_{i,\sigma(j)}$ and the ij entry of $P_\sigma^{-1} B$ is $b_{\sigma(i),j}$. Combining these we see that the ij entry of $P_\sigma^{-1} A P_\sigma$ is $a_{\sigma(i),\sigma(j)}$. □

Lemma 4.

$$\tau_k(A) = \frac{1}{k!(n-k)!} \sum_{\sigma \in S_n} \Delta_k(P_\sigma^{-1} A P_\sigma)$$

Proof. As σ ranges over all permutations each sub-determinant in Lemma 2 is counted $k!(n-k)!$ times in the sum $\sum_{\sigma \in S_n} \Delta_k(P_\sigma^{-1} A P_\sigma)$, because that is the number of σ that permute the first k indices among themselves and the remaining $n-k$ indices among themselves. □

Now to compute the average value of T_k on $V(n,k)$ we need to define a probability measure on $V(n,k)$. The orthogonal group $\mathbf{O}(n)$ acts transitively on $V(n,k)$ by

$$B \cdot (u_1, \ldots, u_k) = (Bu_1, \ldots, Bu_k),$$

with the stabilizer of a k-frame being a subgroup isomorphic to $\mathbf{O}(n-k)$. The projection π from $\mathbf{O}(n)$ onto $V(n,k)$ maps an orthogonal matrix to the frame consisting of the first k columns of the matrix. Let λ be normalized Haar measure on $\mathbf{O}(n)$. Then the pushforward measure $\pi_*\lambda$ is the probability measure on $V(n,k)$ that we use to define a random k-frame. Thus,

$$\int_{u \in V(n,k)} T_k(u)\, d(\pi_*\lambda) = \int_{B \in \mathbf{O}(n)} T_k(\pi(B))\, d\lambda.$$

Lemma 5. *Let B be an orthogonal matrix whose first k columns are $u_1, \ldots, u_k$. Then $T_k(u_1, \ldots, u_k) = \Delta_k(B^{-1}AB)$.*

Proof. Since $B^{-1} = B^t$, the ij entry of $B^{-1}AB$ is $u_i^t A u_j = \langle u_i, Au_j \rangle$ for $1 \leq i, j \leq k$. Thus,

$$\Delta_k(B^{-1}AB) = \det(\langle u_i, Au_j \rangle)_{1 \leq i,j \leq k},$$

which is exactly the determinant giving the value of the oriented k-volume in the definition of $T_k(u)$. □

Now we are ready to finish the proof of the main result.

Theorem 6.

$$\tau_k(A) = \binom{n}{k} \int_{u \in V(n,k)} T_k(u)\, d(\pi_*\lambda).$$

Proof. The characteristic polynomial is invariant under similarity transformations, and so $\tau_k(A) = \tau_k(B^{-1}AB)$. Thus,

$$\tau_k(A) = \int_{B \in \mathbf{O}(n)} \tau_k(B^{-1}AB)\, d\lambda.$$

From Lemma 4 applied to $B^{-1}AB$ we have

$$\tau_k(B^{-1}AB) = \frac{1}{k!(n-k)!} \sum_{\sigma \in S_n} \Delta_k(P_\sigma^{-1} B^{-1} A B P_\sigma),$$

and so

$$\tau_k(A) = \frac{1}{k!(n-k)!} \sum_{\sigma \in S_n} \int_{B \in \mathbf{O}(n)} \Delta_k(P_\sigma^{-1} B^{-1} A B P_\sigma)\, d\lambda.$$

Since Haar measure on $\mathbf{O}(n)$ is both left and right invariant and since P_σ and P_σ^{-1} are in $\mathbf{O}(n)$, it follows that

$$\int_{B\in\mathbf{O}(n)} \Delta_k(P_\sigma^{-1}B^{-1}ABP_\sigma)\,d\lambda = \int_{B\in\mathbf{O}(n)} \Delta_k(B^{-1}AB)\,d\lambda.$$

Therefore

$$\begin{aligned}
\tau_k(A) &= \frac{1}{k!(n-k)!}\sum_{\sigma\in S_n}\int_{B\in\mathbf{O}(n)} \Delta_k(B^{-1}AB)\,d\lambda \\
&= \frac{n!}{k!(n-k)!}\int_{B\in\mathbf{O}(n)} \Delta_k(B^{-1}AB)\,d\lambda \\
&= \binom{n}{k}\int_{B\in\mathbf{O}(n)} T_k(\pi(B))\,d\lambda \qquad \text{(Lemma 5)} \\
&= \binom{n}{k}\int_{u\in V(n,k)} T_k(u)\,d(\pi_*\lambda).
\end{aligned}$$

□

Note 7 (April 2010)**.** It has come to my attention that Eberlein [1] proved an almost identical result 30 years ago with the Grassmann manifold $G(n,k)$ of k-dimensional subspaces of $\mathbf{R}^n$ in place of the manifold of k-frames:

$$\tau_k(A) = \binom{n}{k}\int_{E\in G(n,k)} T_k(E)\,d(\pi_*\lambda).$$

In this formula, $T_k(E) = \det(\mathrm{pr}_E \circ A|E)$, and $\pi_*\lambda$ is the $\mathbf{O}(n)$-invariant measure on $G(n,k)$ obtained by pushing down Haar measure λ from $\mathbf{O}(n)$ using the natural projection $\pi : \mathbf{O}(n) \to G(n,k)$. The proof can proceed in exactly the same way as the proof of Theorem 6, although Eberlein's proof is somewhat different.

References

[1] Patrick Eberlein, A trace formula, *Linear and Multilinear Algebra* **9** (1980) 231–236.

California Polytechnic State University, San Luis Obispo, CA 93407

Current address: American Institute of Mathematics, 600 E. Brokaw Road, San Jose, CA 95112