

What Is the Trace of a Matrix?

Nicholas J. Higham*

January 24, 2023

The trace of an $n \times n$ matrix is the sum of its diagonal elements: $\text{trace}(A) = \sum_{i=1}^n a_{ii}$. The trace is linear, that is, $\text{trace}(A+B) = \text{trace}(A) + \text{trace}(B)$, and $\text{trace}(A) = \text{trace}(A^T)$.

A key fact is that the trace is also the sum of the eigenvalues. The proof is by considering the characteristic polynomial $p(t) = \det(tI - A) = t^n + a_{n-1}t^{n-1} + \dots + a_1t + a_0$. The roots of p are the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ of A , so p can be factorized

$$p(t) = (t - \lambda_1)(t - \lambda_2) \dots (t - \lambda_n),$$

and so $a_{n-1} = -(\lambda_1 + \lambda_2 + \dots + \lambda_n)$. The Laplace expansion of $\det(tI - A)$ shows that the coefficient of t^{n-1} is $-(a_{11} + a_{22} + \dots + a_{nn})$. Equating these two expressions for a_{n-1} gives

$$\text{trace}(A) = \sum_{i=1}^n \lambda_i. \quad (1)$$

A consequence of (1) is that any transformation that preserves the eigenvalues preserves the trace. Therefore the trace is unchanged under similarity transformations: $\text{trace}(X^{-1}AX) = \text{trace}(A)$ for any nonsingular X .

An example of how the trace can be useful, suppose A is a symmetric and orthogonal $n \times n$ matrix, so that its eigenvalues are ± 1 . If there are p eigenvalues 1 and q eigenvalues -1 then $\text{trace}(A) = p - q$ and $n = p + q$. Therefore $p = (n + \text{trace}(A))/2$ and $q = (n - \text{trace}(A))/2$.

Another important property is that for an $m \times n$ matrix A and an $n \times m$ matrix B ,

$$\text{trace}(AB) = \text{trace}(BA) \quad (2)$$

(despite the fact that $AB \neq BA$ in general). The proof is simple:

$$\begin{aligned} \text{trace}(AB) &= \sum_{i=1}^m (AB)_{ii} = \sum_{i=1}^m \sum_{k=1}^n a_{ik} b_{ki} = \sum_{k=1}^n \sum_{i=1}^m b_{ki} a_{ik} \\ &= \sum_{k=1}^n (BA)_{kk} = \text{trace}(BA). \end{aligned}$$

This simple fact can have non-obvious consequences. For example, consider the equation $AX - XA = I$ in $n \times n$ matrices. Taking the trace gives $0 = \text{trace}(AX) - \text{trace}(XA) =$

*Department of Mathematics, University of Manchester, Manchester, M13 9PL, UK (nick.higham@manchester.ac.uk).

$\text{trace}(AX - XA) = \text{trace}(I) = n$, which is a contradiction. Therefore the equation has no solution.

The relation (2) gives $\text{trace}(ABC) = \text{trace}((AB)C) = \text{trace}(C(AB)) = \text{trace}(CAB)$ for $n \times n$ matrices A , B , and C , that is,

$$\text{trace}(ABC) = \text{trace}(CAB). \quad (3)$$

So we can cyclically permute terms in a matrix product without changing the trace.

As an example of the use of (2) and (3), if x and y are n -vectors then $\text{trace}(xy^T) = \text{trace}(y^T x) = y^T x$. If A is an $n \times n$ matrix then $\text{trace}(xy^T A)$ can be evaluated without forming the matrix $xy^T A$ since, by (3), $\text{trace}(xy^T A) = \text{trace}(y^T Ax) = y^T Ax$.

The trace is useful in calculations with the Frobenius norm of an $m \times n$ matrix:

$$\|A\|_F = \left(\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 \right)^{1/2} = (\text{trace}(A^* A))^{1/2},$$

where $*$ denotes the conjugate transpose. For example, we can generalize the formula $|x + iy|^2 = x^2 + y^2$ for a complex number to an $m \times n$ matrix A by splitting A into its Hermitian and skew-Hermitian parts:

$$A = \frac{1}{2}(A + A^*) + \frac{1}{2}(A - A^*) \equiv B + C,$$

where $B = B^*$ and $C = -C^*$. Then

$$\begin{aligned} \|A\|_F^2 &= \|B + C\|_F^2 = \text{trace}((B + C)^*(B + C)) \\ &= \text{trace}(B^* B + C^* C) + \text{trace}(B^* C + C^* B) \\ &= \text{trace}(B^* B + C^* C) + \text{trace}(BC - CB) \\ &= \text{trace}(B^* B + C^* C) \\ &= \|B\|_F^2 + \|C\|_F^2. \end{aligned}$$

If a matrix is not explicitly known but we can compute matrix-vector products with it then the trace can be estimated by

$$\text{trace}(A) \approx x^T Ax,$$

where the vector x has elements independently drawn from the standard normal distribution with mean 0 and variance 1. The expectation of this estimate is

$$\begin{aligned} E(x^T Ax) &= E(\text{trace}(x^T Ax)) = E(\text{trace}(Axx^T)) = \text{trace}(E(Axx^T)) \\ &= \text{trace}(AE(xx^T)) = \text{trace}(A), \end{aligned}$$

since $E(x_i x_j) = 0$ for $i \neq j$ and $E(x_i^2) = 1$ for all i , so $E(xx^T) = I$. This stochastic estimate, which is due to Hutchinson, is therefore unbiased.

References

- Haim Avron and Sivan Toledo, Randomized Algorithms for Estimating the Trace of an Implicit Symmetric Positive Semi-definite Matrix, J. ACM 58, 8:1-8:34, 2011.

Related Blog Posts

- What Is a Matrix Norm? (2021)
- What Is an Eigenvalue? (2022)

This article is part of the “What Is” series, available from <https://nhigham.com/category/what-is> and in PDF form from the GitHub repository <https://github.com/nhigham/what-is>.