What Is a Unitarily Invariant Norm?

Nicholas J. Higham*

February 2, 2021

A norm on $\mathbb{C}^{m\times n}$ is unitarily invariant if $\|UAV\| = \|A\|$ for all unitary $U \in \mathbb{C}^{m\times m}$ and $V \in \mathbb{C}^{n\times n}$ and for all $A \in \mathbb{C}^{m\times n}$. One can restrict the definition to real matrices, though the term unitarily invariant is still typically used.

Two widely used matrix norms are unitarily invariant: the 2-norm and the Frobenius norm. The unitary invariance follows from the definitions. For the 2-norm, for any unitary U and V, using the fact that $\|Uz\|_2 = \|z\|_2$, we obtain

$$\begin{aligned}\|UAV\|_2 &= \max_{x\neq 0} \frac{\|UAVx\|_2}{\|x\|_2} = \max_{x\neq 0} \frac{\|AVx\|_2}{\|x\|_2}\\ &= \max_{x\neq 0} \frac{\|Ay\|_2}{\|V^*y\|_2} \quad (y = Vx)\\ &= \max_{y\neq 0} \frac{\|Ay\|_2}{\|y\|_2} = \|A\|_2.\end{aligned}$$

For the Frobenius norm, using $\|A\|_F^2 = \mathrm{trace}(A^*A)$,

$$\begin{aligned}\|UAV\|_F^2 &= \mathrm{trace}(V^*A^*U^*\cdot UAV)\\ &= \mathrm{trace}(V^*A^*AV)\\ &= \mathrm{trace}(A^*A) = \|A\|_F^2,\end{aligned}$$

since the trace is invariant under similarity transformations.

More insight into unitarily invariant norms comes from recognizing a connection with the singular value decomposition

$$A = P\Sigma Q^*, \quad P^*P = I_m, \quad Q^*Q = I_n, \quad \Sigma = \mathrm{diag}(\sigma_i), \quad \sigma_1 \geq \cdots \geq \sigma_q \geq 0.$$

Clearly, $\|A\| = \|\Sigma\|$, so $\|A\|$ depends only on the singular values. Indeed, for the 2-norm and the Frobenius norm we have $\|A\|_2 = \sigma_1$ and $\|A\|_F = (\sum_{i=1}^q \sigma_i^2)^{1/2}$. Here, and throughout this article, $q = \min(m,n)$. Another implication of the singular value dependence is that $\|A\| = \|A^*\|$ for all A for any unitarily invariant norm.

There is a beautiful characterization of unitarily invariant norms in terms of symmetric gauge functions, which are functions $f : \mathbb{R}^q \to \mathbb{R}^q$ such that f is an absolute norm on $\mathbb{R}^q$ and $f(Px) = f(x)$ for any permutation matrix P and all x. An absolute norm is one with the property that $\|x\| = \||x|\|$ for all x, and this condition is equivalent to the monotonicity condition that $|x| \leq |y|$ implies $\|x\| \leq \|y\|$ for all x and y.

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

Theorem. *A norm on $\mathbb{C}^{m\times n}$ is unitarily invariant if and only if $\|A\| = f(\sigma_1,\ldots,\sigma_q)$ for all A for some symmetric gauge function f, where $\sigma_1,\ldots,\sigma_q$ are the singular values of A.*

The matrix 2-norm and the Frobenius norm correspond to f being the vector ∞-norm and the 2-norm, respectively. More generally, we can take for f any vector p-norm, obtaining the class of Schatten p-norms:

$$\|A\|_p^S = \left(\sum_{i=1}^{q} \sigma_i^p\right)^{1/p}, \quad 1 \le p \le \infty.$$

The Schatten 1-norm is the sum of the singular values, $\|A\|_1^S = \sum_{i=1}^{q} \sigma_i$, which is called the trace norm or nuclear norm. It can act as a proxy for the rank of a matrix. The trace norm can be expressed as $\|A\|_1^S = \mathrm{trace}(H) = \mathrm{trace}\big((A^*A)^{1/2}\big)$, where $A = UH$ is a polar decomposition.

Another class of unitarily invariant norms is the Ky Fan k-norms

$$\|A\|_{(k)} = \sum_{i=1}^{k} \sigma_i, \quad 1 \le k \le n.$$

We have $\|A\|_{(n)} = \|A\|_1^S$ and $\|A\|_{(1)} = \|A\|_2$.

Among unitarily invariant norms, the 2-norm and the Frobenius norm are widely usd in numerical analysis and matrix analysis. The nuclear norm is used in problems involving matrix rank minimization, such as matrix completion problems.

The benefit of the concept of unitarily invariant norm is that one can prove certain results for this whole class of norms, obtaining results for the particular norms of interest as special cases. Here are three important examples.

- For $A \in \mathbb{C}^{n\times n}$, the matrix $(A + A^*)/2$ is the nearest Hermitian matrix to A in any unitarily invariant norm.
- For $A \in \mathbb{C}^{m\times n}$ $(m \ge n)$, the unitary polar factor is the nearest matrix with orthonormal columns to A in any unitarily invariant norm.
- The best rank-k approximation to $A \in \mathbb{C}^{m\times n}$ in any unitarily invariant norm is obtained by setting all the singular values beyond the kth to zero in the SVD of A. This result, which generalizes the Eckart-Young theorem, which covers the 2- and Frobenius norm instances, is an easy consequence of the following result of Mirsky (1960).

  Theorem. *Let $A, B \in \mathbb{C}^{m\times n}$ have SVDs with diagonal matrices $\Sigma_A, \Sigma_B \in \mathbb{R}^{m\times n}$, where the diagonal elements are arranged in nonincreasing order. Then $\|A - B\| \ge \|\Sigma_A - \Sigma_B\|$ for every unitarily invariant norm.*

We also give a useful matrix norm inequality. For any matrices A, B, and C for which the product ABC is defined,

$$\|ABC\| \le \|A\|_2 \, \|B\| \, \|C\|_2$$

holds for any unitarily invariant norm, and in fact, any two of the norms on the right-hand side can be 2-norms.

References

This is a minimal set of references, which contain further useful references within.

- Roger Horn and Charles Johnson, Topics in Matrix Analysis, Cambridge University Press, 1991. Section 3.5.
- Roger A. Horn and Charles R. Johnson, Matrix Analysis, second edition, Cambridge University Press, 2013. Sections 5.6, 7.4. My review of the second edition.
- L. Mirsky, Symmetric Gauge Functions and Unitarily Invariant Norms, Quart. J. Math. 11, 50–59, 1960.
- Benjamin Recht, Maryam Fazel, and Pablo Parrilo, Guaranteed Minimum-Rank Solutions of Linear Matrix Equations via Nuclear Norm Minimization, SIAM Rev. 52 (3), 471–501, 2010.

Related Blog Posts

- What Is the Nearest Symmetric Matrix? (2020)
- What Is the Polar Decomposition? (2020)
- What Is the Singular Value Decomposition? (2020)

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/higham/what-is`.