

# Matrix norm

From Wikipedia, the free encyclopedia

In mathematics, a **matrix norm** is a natural extension of the notion of a vector norm to matrices.

## Contents

- 1 Definition
- 2 Induced norm
- 3 "Entrywise" norms
  - 3.1 L2,1 and Lp,q norms
  - 3.2 Frobenius norm
  - 3.3 Max norm
- 4 Schatten norms
- 5 Consistent norms
- 6 Compatible norms
- 7 Equivalence of norms
  - 7.1 Examples of norm equivalence
- 8 Notes
- 9 References

## Definition

In what follows,  $K$  will denote the field of real or complex numbers. Let  $K^{m \times n}$  denote the vector space containing all matrices with  $m$  rows and  $n$  columns with entries in  $K$ . Throughout,  $A^*$  denotes the conjugate transpose of matrix  $A$ .

A matrix norm is a vector norm on  $K^{m \times n}$ . That is, if  $\|A\|$  denotes the norm of the matrix  $A$ , then,

- $\|A\| \geq 0$
- $\|A\| = 0$  iff  $A = 0$
- $\|\alpha A\| = |\alpha| \|A\|$  for all  $\alpha$  in  $K$  and all matrices  $A$  in  $K^{m \times n}$

- $\|A + B\| \leq \|A\| + \|B\|$  for all matrices  $A$  and  $B$  in  $K^{m \times n}$ .

Additionally, in the case of square matrices (thus,  $m = n$ ), some (but not all) matrix norms satisfy the following condition, which is related to the fact that matrices are more than just vectors:

- $\|AB\| \leq \|A\|\|B\|$  for all matrices  $A$  and  $B$  in  $K^{n \times n}$ .

A matrix norm that satisfies this additional property is called a **submultiplicative norm** (in some books, the terminology *matrix norm* is used only for those norms which are submultiplicative). The set of all  $n \times n$  matrices, together with such a submultiplicative norm, is an example of a Banach algebra.

## Induced norm

If vector norms on  $K^m$  and  $K^n$  are given ( $K$  is the field of real or complex numbers), then one defines the corresponding *induced norm* or *operator norm* on the space of  $m$ -by- $n$  matrices as the following suprema:

$$\begin{aligned}\|A\| &= \sup\{\|Ax\| : x \in K^n \text{ with } \|x\| = 1\} \\ &= \sup\left\{\frac{\|Ax\|}{\|x\|} : x \in K^n \text{ with } x \neq 0\right\}.\end{aligned}$$

The operator norm corresponding to the  $p$ -norm for vectors is:

$$\|A\|_p = \sup_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_p}.$$

These are different from the entrywise  $p$ -norms and the Schatten  $p$ -norms for matrices treated below, which are also usually denoted by  $\|A\|_p$ .

In some cases, the norms can be computed or estimated by

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^m |a_{ij}|, \text{ which is simply the maximum absolute column sum of the matrix.}$$

$\|A\|_\infty = \max_{1 \leq i \leq m} \sum_{j=1}^n |a_{ij}|$ , which is simply the maximum absolute row sum of the matrix

$\|A\|_2 \leq \left( \sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 \right)^{1/2} = \|A\|_F$ , where the right hand side is the Frobenius norm or  $L_{2,2}$  norm. The equality holds if and only if the matrix  $A$  is a rank-one matrix or a zero matrix.

For example, if the matrix  $A$  is defined by

$$A = \begin{bmatrix} -3 & 5 & 7 \\ 2 & 6 & 4 \\ 0 & 2 & 8 \end{bmatrix},$$

then we have  $\|A\|_1 = \max(|-3|+2+0, 5+6+2, 7+4+8) = \max(5, 13, 19) = 19$ . and  $\|A\|_\infty = \max(|-3|+5+7, 2+6+4, 0+2+8) = \max(15, 12, 10) = 15$ .

In the special case of  $p = 2$  (the Euclidean norm), the induced matrix norm is the *spectral norm*. The spectral norm of a matrix  $A$  is the largest singular value of  $A$  i.e. the square root of the largest eigenvalue of the positive-semidefinite matrix  $A^*A$ :

$$\|A\|_2 = \sqrt{\lambda_{\max}(A^*A)} = \sigma_{\max}(A)^{[1]}$$

where  $A^*$  denotes the conjugate transpose of  $A$ .

More generally, one can define the subordinate matrix norm on  $K^{m \times n}$  induced by  $\|\cdot\|_\alpha$  on  $K^n$ , and  $\|\cdot\|_\beta$  on  $K^m$  as:

$$\|A\|_{\alpha, \beta} = \max_{x \neq 0} \frac{\|Ax\|_\beta}{\|x\|_\alpha}.$$

Subordinate norms are consistent with the norms that induce them, giving

$$\|Ax\|_\beta \leq \|A\|_{\alpha, \beta} \|x\|_\alpha.$$

For  $\alpha = \beta$ , any induced operator norm is a sub-multiplicative matrix norm since  $\|ABx\| \leq \|A\|\|Bx\| \leq \|A\|\|B\|\|x\|$  and  $\max_{\|x\|=1} \|ABx\| = \|AB\|$ .

Any induced norm satisfies the inequality

$$\|A^r\|^{1/r} \geq \rho(A),$$

where  $\rho(A)$  is the spectral radius of  $A$ . For a symmetric or hermitian matrix  $A$ , we have equality for the 2-norm, since in this case the 2-norm is the spectral radius of  $A$ . For an arbitrary matrix, we may not have equality for any norm. Take

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$

the spectral radius of  $A$  is 0, but  $A$  is not the zero matrix, and so none of the induced norms are equal to the spectral radius of  $A$ .

Furthermore, for square matrices we have the spectral radius formula:

$$\lim_{r \rightarrow \infty} \|A^r\|^{1/r} = \rho(A).$$

## "Entrywise" norms

These vector norms treat an  $m \times n$  matrix as a vector of size  $mn$ , and use one of the familiar vector norms.

For example, using the  $p$ -norm for vectors, we get:

$$\|A\|_p = \|\text{vec}(A)\|_p = \left( \sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^p \right)^{1/p}$$

This is a different norm from the induced  $p$ -norm (see above) and the Schatten  $p$ -norm (see below), but the notation is the same.

The special case  $p = 2$  is the Frobenius norm, and  $p = \infty$  yields the maximum norm.

## L2,1 and Lp,q norms

Let  $(a_1, \dots, a_n)$  be the columns of matrix  $A$ . The  $L_{2,1}$  norm<sup>[2]</sup> is the sum of the Euclidean norms of the columns of the matrix:

$$\|A\|_{2,1} = \sum_{j=1}^n \|a_j\|_2 = \sum_{j=1}^n \left( \sum_{i=1}^m |a_{ij}|^2 \right)^{1/2}$$

In this norm, the two indices  $i$  and  $j$  of  $A_{i,j}$  are treated differently; all matrix norms introduced prior to the  $L_{2,1}$  norm treat the two indices symmetrically. The  $L_{2,1}$  norm is used in robust data analysis and sparse coding for feature selection.

The  $L_{2,1}$  norm can be generalized to the  $L_{p,q}$  norm, defined by

$$\|A\|_{p,q} = \left[ \sum_{j=1}^n \left( \sum_{i=1}^m |a_{ij}|^p \right)^{q/p} \right]^{1/q}$$

## Frobenius norm

When  $p = q = 2$  for the  $L_{p,q}$  norm, it is called the **Frobenius norm** or the **Hilbert–Schmidt norm**, though the latter term is used more frequently in the context of operators on (possibly infinite dimensional) Hilbert space. This norm can be defined in various ways:

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

where  $A^*$  denotes the conjugate transpose of  $A$ ,  $\sigma_i$  are the singular values of  $A$ , and the trace function is used. The Frobenius norm is similar to the Euclidean norm on  $K^n$  and comes from the Frobenius inner product on the space of all matrices.

The Frobenius norm is sub-multiplicative and is very useful for numerical linear algebra. This norm is often easier to compute than induced norms and has the useful property of being invariant under rotations, that is,  $\|A\|_F^2 = \|AR\|_F^2 = \|RA\|_F^2$  for any rotation matrix  $R$ . This property follows from the trace definition restricted to real matrices,

$$\|AR\|_F^2 = \text{trace}(R^T A^T AR) = \text{trace}(RR^T A^T A) = \text{trace}(A^T A) = \|A\|_F^2$$

and

$$\|RA\|_F^2 = \text{trace}(A^T R^T RA) = \text{trace}(A^T A) = \|A\|_F^2$$

where we have used the orthogonal nature of  $R$ , that is,  $R^T R = RR^T = \mathbf{I}$ , and the cyclic nature of the trace,  $\text{trace}(XYZ) = \text{trace}(ZXY)$ . More generally the norm is invariant under a unitary transformation for complex matrices.

## Max norm

The **max norm** is the elementwise norm with  $p = \infty$ :

$$\|A\|_{\max} = \max_{ij} |a_{ij}|. [3]$$

This norm is not sub-multiplicative.

## Schatten norms

The Schatten  $p$ -norms arise when applying the  $p$ -norm to the vector of singular values of a matrix. If the singular values are denoted by  $\sigma_i$ , then the Schatten  $p$ -norm is defined by

$$\|A\|_p = \left( \sum_{i=1}^{\min\{m, n\}} \sigma_i^p \right)^{1/p}.$$

These norms again share the notation with the induced and entrywise  $p$ -norms, but they are different.

All Schatten norms are sub-multiplicative. They are also unitarily invariant, which means that  $\|A\| = \|UAV\|$  for all matrices  $A$  and all unitary matrices  $U$  and  $V$ .

The most familiar cases are  $p = 1, 2, \infty$ . The case  $p = 2$  yields the Frobenius norm, introduced before. The case  $p = \infty$  yields the spectral norm, which is the matrix norm induced by the vector 2-norm (see above). Finally,  $p = 1$  yields the **nuclear norm** (also known as the *trace norm*, or the Ky Fan 'n'-norm), defined as

$$\|A\|_* = \text{trace} \left( \sqrt{A^*A} \right) = \sum_{i=1}^{\min\{m, n\}} \sigma_i.$$

(Here  $\sqrt{A^*A}$  denotes a positive semidefinite matrix  $B$  such that  $BB = A^*A$ . More precisely, since  $A^*A$  is a positive semidefinite matrix, its square root is well-defined.)

## Consistent norms

A matrix norm  $\|\cdot\|_{ab}$  on  $K^{m \times n}$  is called *consistent* with a vector norm  $\|\cdot\|_a$  on  $K^n$  and a vector norm  $\|\cdot\|_b$  on  $K^m$  if:

$$\|Ax\|_b \leq \|A\|_{ab} \|x\|_a$$

for all  $A \in K^{m \times n}, x \in K^n$ . All induced norms are consistent by definition.

## Compatible norms

A matrix norm  $\|\cdot\|_b$  on  $K^{n \times n}$  is called *compatible* with a vector norm  $\|\cdot\|_a$  on  $K^n$  if:

$$\|Ax\|_a \leq \|A\|_b \|x\|_a$$

for all  $A \in K^{n \times n}, x \in K^n$ . Induced norms are compatible by definition.

## Equivalence of norms

For any two vector norms  $\|\cdot\|_\alpha$  and  $\|\cdot\|_\beta$ , we have

$$r \|A\|_\alpha \leq \|A\|_\beta \leq s \|A\|_\alpha$$

for some positive numbers  $r$  and  $s$ , for all matrices  $A$  in  $K^{m \times n}$ . In other words, all norms on  $K^{m \times n}$  are *equivalent*; they induce the same topology on  $K^{m \times n}$ . This is true because the vector space  $K^{m \times n}$  has the finite dimension  $m \times n$ .

Moreover, for every vector norm  $\|\cdot\|$  on  $\mathbb{R}^{n \times n}$ , there exists a unique positive real number  $k$  such that  $l\|\cdot\|$  is a sub-multiplicative matrix norm for every  $l \geq k$ .

A sub-multiplicative matrix norm  $\|\cdot\|_\alpha$  is said to be *minimal* if there exists no other sub-multiplicative matrix norm  $\|\cdot\|_\beta$  satisfying  $\|\cdot\|_\beta < \|\cdot\|_\alpha$ .

## Examples of norm equivalence

For matrix  $A \in \mathbb{R}^{m \times n}$  of rank  $r$ , the following inequalities hold:<sup>[4][5]</sup>

- $\|A\|_2 \leq \|A\|_F \leq \sqrt{r}\|A\|_2$
- $\|A\|_F \leq \|A\|_* \leq \sqrt{r}\|A\|_F$
- $\|A\|_{\max} \leq \|A\|_2 \leq \sqrt{mn}\|A\|_{\max}$
- $\frac{1}{\sqrt{n}}\|A\|_\infty \leq \|A\|_2 \leq \sqrt{m}\|A\|_\infty$
- $\frac{1}{\sqrt{m}}\|A\|_1 \leq \|A\|_2 \leq \sqrt{n}\|A\|_1$ .

Here,  $\|A\|_p$  refers to the matrix norm induced by the vector  $p$ -norm.

Another useful inequality between matrix norms is

$$\|A\|_2 \leq \sqrt{\|A\|_1 \|A\|_\infty},$$

which is a special case of Hölder's inequality.

## Notes

1. Carl D. Meyer, Matrix Analysis and Applied Linear Algebra, section 5.2, p281, Society for Industrial & Applied Mathematics, June 2000.
2. Ding, Chris; Zhou, Ding; He, Xiaofeng; Zha, Hongyuan (June 2006). "R1-PCA: Rotational Invariant L1-norm Principal Component Analysis for Robust Subspace Factorization". *Proceedings of the 23rd International Conference on Machine Learning*. ICML '06. Pittsburgh, Pennsylvania, USA: ACM. pp. 281–288.



doi:10.1145/1143844.1143880. ISBN 1-59593-383-2.

3. <http://mathworld.wolfram.com/MaximumAbsoluteRowSumNorm.html>

4. Golub, Gene; Charles F. Van Loan (1996). *Matrix Computations* – Third Edition. Baltimore: The Johns Hopkins University Press, 56–57. ISBN 0-8018-5413-X.

5. Roger Horn and Charles Johnson. *Matrix Analysis*, Chapter 5, Cambridge University Press, 1985. ISBN 0-521-38632-2.

## References

- James W. Demmel, *Applied Numerical Linear Algebra*, section 1.7, published by SIAM, 1997.
- Carl D. Meyer, *Matrix Analysis and Applied Linear Algebra*, published by SIAM, 2000. [1] (<http://www.matrixanalysis.com>)
- John Watrous, *Theory of Quantum Information*, 2.3 Norms of operators (<https://cs.uwaterloo.ca/~watrous/CS766/LectureNotes/02.pdf>), lecture notes, University of Waterloo, 2011.
- Kendall Atkinson, *An Introduction to Numerical Analysis*, published by John Wiley & Sons, Inc 1989

Retrieved from "[https://en.wikipedia.org/w/index.php?title=Matrix\\_norm&oldid=713509213](https://en.wikipedia.org/w/index.php?title=Matrix_norm&oldid=713509213)"

Categories: Norms (mathematics) | Linear algebra

- 
- This page was last modified on 4 April 2016, at 13:08.
  - Text is available under the Creative Commons Attribution-ShareAlike License; additional terms may apply. By using this site, you agree to the Terms of Use and Privacy Policy. Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc., a non-profit organization.