

## Vector and Matrix Norms

♠ **Vector Norm.** Given an  $n$ -dimensional vector  $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ , a vector norm  $\|x\|$ , is a non-negative real number, defined such that

1.  $\|x\| > 0$  when  $x \neq 0$  and  $\|x\| = 0$  if and only if  $x = \theta$ , the zero vector,
2.  $\|kx\| = |k| \times \|x\|$  for any scalar  $k$ ,
3.  $\|x + y\| \leq \|x\| + \|y\|$ ,

For  $p = 1, 2, \dots, \infty$ , the  $p$ -vector norm  $\|x\|_p$  is defined as

$$\|x\|_p = \left[ \sum_i |x_i|^p \right]^{1/p}.$$

The special case  $\|x\|_\infty$  is defined as

$$\|x\|_\infty = \max_i |x_i|.$$

The most commonly encountered vector norm (often simply called "the norm" of a vector) is the L2-norm, given by

$$\|x\|_2 = \|x\| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}.$$

This and other types of vector norms are summarized in the following table, together with the value of the norm for the example vector  $v = (1, 2, 3)^t$ .

norm	symbol	value	approx.
$L^1$	$\ x\ _1$	6	6.000
$L^2$	$\ x\ _2$	$\sqrt{14}$	3.742
$L^3$	$\ x\ _3$	$6^{2/3}$	3.302
$L^4$	$\ x\ _4$	$2^{1/4}\sqrt{7}$	3.146
$L^\infty$	$\ x\ _\infty$	3	3.000

The concept of unit circle (the set of all vectors of norm 1) is different in different norms: for the 1-norm the unit circle in  $\mathbb{R}^2$  is a square with vertices at  $(1, 0)$ ,  $(0, 1)$ ,  $(-1, 0)$ , and  $(0, -1)$ ; for the 2-norm (Euclidean norm) it is the well-known unit circle, while for the infinity norm it is a square with vertices at  $(1, 1)$ ,  $(-1, 1)$ ,  $(-1, -1)$ , and  $(1, -1)$ . Note that due to the definition of the norm, the unit circle is always convex and centrally symmetric (therefore, for example, the unit ball may be a rectangle but cannot be a triangle).

Two norms  $\|\cdot\|_\alpha$  and  $\|\cdot\|_\beta$  on a vector space  $V$  are called equivalent if there exist positive real numbers  $\mathcal{R}$  and  $\mathcal{S}$  such that

$$\mathcal{R}\|x\|_\alpha \leq \|x\|_\beta \leq \mathcal{S}\|x\|_\alpha,$$

for all  $x \in V$ . On a finite-dimensional vector space all norms are equivalent. For instance, the  $L^1$ ,  $L^2$ , and  $L^\infty$  norms are all equivalent on  $\mathbb{R}^n$ :

$$\begin{aligned}\|x\|_2 &\leq \|x\|_1 \leq \sqrt{n} \|x\|_2 \\ \|x\|_\infty &\leq \|x\|_2 \leq \sqrt{n} \|x\|_\infty \\ \|x\|_\infty &\leq \|x\|_1 \leq n \|x\|_\infty.\end{aligned}$$

Equivalent norms define the same notions of continuity and convergence and for many purposes do not need to be distinguished. To be more precise the uniform structure defined by equivalent norms on the vector space is uniformly isomorphic.

Given any norm  $\|\cdot\|_\alpha$  and any non-singular matrix  $A$ , we can define a vector norm  $\|\cdot\|_\beta$  such that  $\|x\|_\beta = \|Ax\|_\alpha$ .

**♠ Matrix Norm.** Given a complex or real matrix  $A = (a_{ij})$ , a matrix norm  $\|A\|$  is a non-negative number associated with  $A$  having the properties:

1.  $\|A\| > 0$  when  $A \neq 0$  and  $\|A\| = 0$  if and only if  $A = Z$ , the zero matrix,
2.  $\|kA\| = |k| \times \|A\|$  for any scalar  $k$ ,
3.  $\|A + B\| \leq \|A\| + \|B\|$ ,

For square matrices, a sub-multiplicative matrix norm also satisfies:

4.  $\|AB\| \leq \|A\| \times \|B\|$ .

If a matrix norm is not sub-multiplicative, then it is called a generalized matrix norm.

Let  $\lambda_1, \lambda_2, \dots, \lambda_n$  be the eigenvalues of  $A$ , then

$$\frac{1}{\|A^{-1}\|} \leq |\lambda| \leq \|A\|. \quad (1)$$

The spectrum of a square matrix  $A$ , denoted by  $\sigma(A)$  is the set of all eigenvalues of  $A$ . The spectral radius of  $A$ , denoted by  $\rho(A)$  is defined as:

$$\rho(A) = \max\{|\lambda| : \lambda \in \sigma(A)\}$$

The spectral norm of  $m \times n$  matrix  $A$ , denoted by  $\|A\|_2$ , which is the square root of the maximum eigenvalue of the positive semi-definite matrix  $A^*A$  or  $AA^*$  (we choose the one in smaller size),

$$\|A\|_2 = \sqrt{\rho(A^*A)} = \sqrt{\rho(AA^*)} \quad (2)$$

is often referred to as "the" matrix norm.

The row norm is defined by

$$\|A\|_\infty = \max_i \sum_{j=1}^n |a_{ij}|. \quad (3)$$

$\|A\|_1$ ,  $\|A\|_2$ , and  $\|A\|_\infty$  satisfy the inequality

$$\|A\|_2^2 \leq \|A\|_1 \times \|A\|_\infty.$$

The Frobenius norm of an  $m \times n$  matrix  $A$  defined as the square root of the sum of the absolute squares of its elements; it is also equal to the square root of the trace of the positive semi-definite matrix  $AA^*$

$$\|A\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2} = \sqrt{\text{trace}(AA^*)}. \quad (4)$$

The Frobenius norm can also be considered as a vector norm.

♣ **Induced Norm.** If vector norms on  $\mathbb{K}^n$  are given ( $\mathbb{K}^n$  is field of real or complex numbers), then one defines the corresponding induced norm, natural norm, or operator norm on the space of matrices as the follows:

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

For example, the operator norm corresponding to the p-norm for vectors is:

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

In the case of  $p = 1$  and  $p = \infty$ , the norms can be computed as:

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^m |a_{ij}|,$$

as the column norm of the matrix  $A$  and

$$\|A\|_\infty = \max_{1 \leq i \leq m} \sum_{j=1}^n |a_{ij}|,$$

as the row norm of the matrix  $A$ . In the case of  $p = 2$  (the Euclidean norm), the induced matrix norm is the spectral norm.

**Theorem 1.** *All induced norms are sub-multiplicative.*

**Proof.** Let  $A$  and  $B$  be two matrices where  $AB$  is well defined, then

$$\|AB\| = \max_{x \neq \theta} \frac{\|ABx\|}{\|x\|} = \max_{x \neq \theta} \frac{\|ABx\|}{\|Bx\|} \frac{\|Bx\|}{\|x\|} \leq \left[ \max_{x \neq \theta} \frac{\|Ax\|}{\|x\|} \right] \left[ \max_{x \neq \theta} \frac{\|Bx\|}{\|x\|} \right] = \|A\| \times \|B\|. \quad \blacksquare$$

**Theorem 2.** If  $A$  is an  $n \times n$ , then  $\rho(A) \leq \|A\|$  for any sub-multiplicative matrix norm  $\|\cdot\|$ .

**Proof.** Let  $v$  be a eigenvector of  $A$  corresponding to the dominant eigenvalue  $\lambda$  (i.e.,  $Av = \lambda v$  with  $|\lambda| = \rho(A)$ ). Define a matrix  $V$ , where the first column of  $V$  is  $v$  and the next  $n - 1$  columns are zero vectors. Define the matrix

$$U = \frac{1}{\|V\|} V \quad \text{with the property} \quad AU = \lambda U;$$

then we have

$$\rho(A) = |\lambda| \cdot 1 = |\lambda| \|U\| = \|\lambda U\| = \|AU\| \leq \|A\| \|U\| = \|A\|. \quad \blacksquare$$

**Note.** The max norm  $\|A\|_{\max} = \max\{|a_{ij}|\}$  is not a sub-multiplicative matrix norm. Also the sub-multiplicative Frobenius norm is not an induced norm since the induced norm of any  $n \times n$  identity matrix  $I_n$  is one, but  $\|I_n\|_F = \sqrt{n}$ .

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

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

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

**♣ Equivalence of Norms.** For any two vector norms  $\|\cdot\|_\alpha$  and  $\|\cdot\|_\beta$ , we have

$$\mathcal{R} \|A\|_\alpha \leq \|A\|_\beta \leq \mathcal{S} \|A\|_\alpha$$

for some positive numbers  $\mathcal{R}$  and  $\mathcal{S}$ , for all matrices  $A \in \mathbb{K}^{m \times n}$ . Moreover, when  $A \in \mathbb{R}^{n \times n}$ , then for any vector norm  $\|\cdot\|$ , there exists a unique positive number  $\mathcal{K}$  such that  $\mathcal{L}\|A\|$  is a sub-multiplicative matrix norm for every  $\mathcal{L} \geq \mathcal{K}$ . A matrix norm  $\|\cdot\|_a$  is said to be minimal if there exists no other matrix norm  $\|\cdot\|_b$  satisfying  $\|\cdot\|_a \leq \|\cdot\|_b$ .

**Examples of norm equivalence.** For a matrix  $A \in \mathbb{R}^{m \times n}$ , the following inequalities hold:

$$\begin{aligned} \|A\|_2 &\leq \|A\|_F \leq \sqrt{n} \|A\|_2 \\ \|A\|_{\max} &\leq \|A\|_2 \leq \sqrt{mn} \|A\|_{\max} \\ \frac{1}{\sqrt{n}} \|A\|_\infty &\leq \|A\|_2 \leq \sqrt{m} \|A\|_\infty \\ \frac{1}{\sqrt{m}} \|A\|_1 &\leq \|A\|_2 \leq \sqrt{n} \|A\|_1. \end{aligned}$$

**Example** Let

$$A = \begin{pmatrix} 1 & -2 & 3 \\ -4 & 5 & -6 \\ 7 & -8 & 9 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -1 & 0 \\ 2 & 3 & -1 \\ 5 & 0 & -2 \end{pmatrix}, \quad \text{and} \quad C = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix}.$$

Then for  $A$ , we obtain:

$$\rho(A) \approx |16.1168| = 16.1168,$$

$$\|A\|_{\infty} = \max\{|1| + |-2| + |3|, |-4| + |5| + |-6|, |7| + |-8| + |9|\} = 24,$$

$$\|A\|_1 = \max\{|1| + |-4| + |7|, |-2| + |5| + |-8|, |3| + |-6| + |9|\} = 18,$$

$$\|A\|_2 = \sqrt{\rho(AA^*)} \approx \sqrt{283.8585} \approx 16.8481,$$

$$\|A\|_F = \sqrt{|1|^2 + |-2|^2 + |3|^2 + |-4|^2 + |5|^2 + |-6|^2 + |7|^2 + |-8|^2 + |9|^2} = \sqrt{285} \approx 16.8819,$$

$$\|A\|_{max} = \max\{|1|, |-2|, |3|, |-4|, |5|, |-6|, |7|, |-8|, |9|\} = |9|,$$

with

$$\|A\|_{max} < \rho(A) < \|A\|_2 < \|A\|_F < \|A\|_1 < \|A\|_{\infty}.$$

Now for  $B$ , we obtain:

$$\rho(B) \approx |2.3772| = 2.3772,$$

$$\|B\|_{\infty} = \max\{|1| + |-1| + |0|, |2| + |3| + |-1|, |5| + |0| + |-2|\} = 7,$$

$$\|B\|_1 = \max\{|1| + |2| + |5|, |-1| + |3| + |0|, |0| + |-1| + |-2|\} = 8,$$

$$\|B\|_2 = \sqrt{\rho(BB^*)} \approx \sqrt{36.0994} \approx 6.0083,$$

$$\|B\|_F = \sqrt{|1|^2 + |-1|^2 + |0|^2 + |2|^2 + |3|^2 + |-1|^2 + |5|^2 + |0|^2 + |-2|^2} = \sqrt{45} \approx 6.7082,$$

$$\|B\|_{max} = \max\{|1|, |-1|, |0|, |2|, |3|, |-1|, |5|, |0|, |-2|\} = |5|,$$

$$\rho(B) < \|B\|_{max} < \|B\|_2 < \|B\|_F < \|B\|_{\infty} < \|B\|_1.$$

Finally, for  $C$ , we obtain:

$$\rho(C) = |4| = 4,$$

$$\|C\|_{\infty} = \max\{|1| + |1| + |2|, |2| + |1| + |1|, |1| + |2| + |1|\} = 4,$$

$$\|C\|_1 = \max\{|1| + |2| + |1|, |1| + |1| + |2|, |2| + |1| + |1|\} = 4,$$

$$\|C\|_2 = \sqrt{\rho(CC^*)} = \sqrt{16} = 4,$$

$$\|C\|_F = \sqrt{|1|^2 + |1|^2 + |2|^2 + |1|^2 + |1|^2 + |2|^2 + |1|^2 + |2|^2 + |1|^2} = \sqrt{18} \approx 4.2426,$$

$$\|C\|_{max} = \max\{|1|, |1|, |2|, |2|, |1|, |1|, |1|, |2|, |1|\} = |2|,$$

$$\|C\|_{max} < \rho(C) = \|C\|_1 = \|C\|_2 = \|C\|_{\infty} < \|C\|_F.$$