Chapter 3

NORMS

1. Vector norms

In this chapter the notions of vector norm and matrix norm will be introduced, with some of their properties as well. The notion of norm generalizes the notion of length of a vector $\mathbf{x} \in \mathbf{R}^n$, given by the value

$$\sqrt{x_1^2 + \cdots + x_n^2}.$$

3.1 Definition. A function from $\mathbf{C}^n$ to $\mathbf{R}$

$$\mathbf{x} \to \|\mathbf{x}\|$$

which enjoys the following properties:

a) $\|\mathbf{x}\| \geq 0$ and $\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = \mathbf{0}$,

b) $\|\alpha\mathbf{x}\| = |\alpha|\, \|\mathbf{x}\|$ for every $\alpha \in \mathbf{C}$,

c) $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ for every $\mathbf{y} \in \mathbf{C}^n$,

is called *vector norm*. ■

The property c), if the length of a vector is taken as norm, is the well known *triangular inequality*, which states that the sum of the lengths of two sides of a triangle is greater than or equal to the length of the third side.

Generally a norm is denoted by the symbol $\|\,.\,\|$. An index is added if a particular norm is referred. In the following some of the norms commonly used will be introduced.

3.2 Definition. Let $\mathbf{x} \in \mathbf{C}^n$; we define the following three norms:

$$\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i| \qquad \text{1-norm}$$

$$\|\mathbf{x}\|_2 = \sqrt{\mathbf{x}^H\mathbf{x}} = \sqrt{\sum_{i=1}^n |x_i|^2} \qquad \text{2-norm}$$

$$\|\mathbf{x}\|_\infty = \max_{i=1,\ldots,n} |x_i| \qquad \infty\text{-norm}$$

The 2-norm is just the euclidean length of the vector $\mathbf{x}$. ■

Here we show, for example, that the ∞-norm satisfies the properties a), b) and c) in definition 3.1:

a) since $|x_i| \geq 0$ for $i = 1, \ldots, n$, it follows that $\max\limits_{i=1,\ldots,n} |x_i| \geq 0$ so $\|\mathbf{x}\|_\infty \geq 0$; moreover if $\max\limits_{i=1,\ldots,n} |y_i| = 0$, then $|x_i| = 0$ for $i = 1, \ldots, n$; also the converse implication is true, therefore $\|\mathbf{x}\|_\infty = 0$ if and only if $\mathbf{x} = \mathbf{0}$;

b) $\|\alpha\mathbf{x}\|_\infty = \max\limits_{i=1,\ldots,n} |\alpha x_i| = \max\limits_{i=1,\ldots,n} |\alpha|\,|x_i| = |\alpha| \max\limits_{i=1,\ldots,n} |x_i| = |\alpha|\,\|\mathbf{x}\|_\infty$;

c) $\|\mathbf{x}+\mathbf{y}\|_\infty = \max\limits_{i=1,\ldots,n} |x_i + y_i| \leq \max\limits_{i=1,\ldots,n} (|x_i| + |y_i|)$
$\leq \max\limits_{i=1,\ldots,n} |x_i| + \max\limits_{i=1,\ldots,n} |y_i| = \|\mathbf{x}\|_\infty + \|\mathbf{y}\|_\infty$.

The proofs for the other two norms are analogous, in detail the property c) for $\| \,.\, \|_2$ can be proven by using the Cauchy-Schwarz inequality (see (1), chap.1).

The sets:

$$
\begin{aligned}
C_1 &= \{\mathbf{x} \in \mathbf{R}^2 : \|\mathbf{x}\|_1 \leq 1\},\\
C_2 &= \{\mathbf{x} \in \mathbf{R}^2 : \|\mathbf{x}\|_2 \leq 1\},\\
C_\infty &= \{\mathbf{x} \in \mathbf{R}^2 : \|\mathbf{x}\|_\infty \leq 1\},
\end{aligned}
$$

are the unitary balls in $\mathbf{R}^2$ with respect to 1-, 2- and ∞- norms (see fig.3.1).

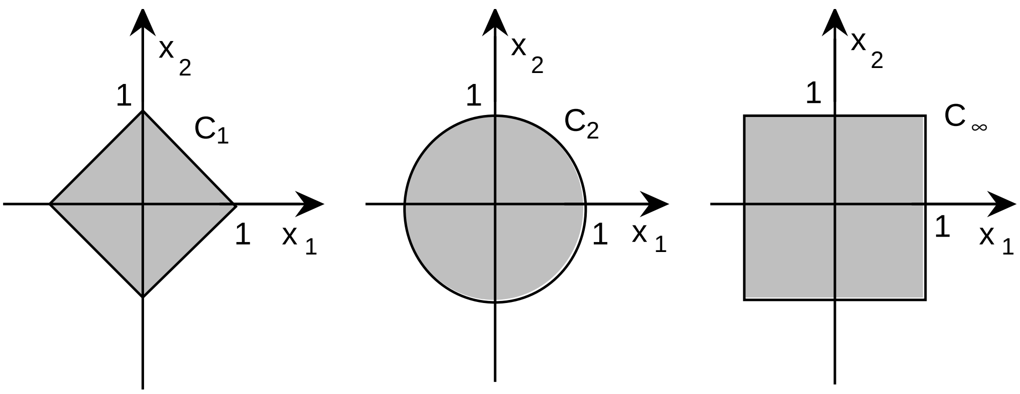

Fig. 3.1 - Unitary balls in $\mathbf{R}^2$ with respect to 1-, 2- and ∞- norms.

The following theorems give some important properties of vector norms.

3.3 Theorem. *The function* $\mathbf{x} \to \|\mathbf{x}\|$, $\mathbf{x} \in \mathbf{C}^n$, *is uniformly continuous.*

Proof. Let $\mathbf{x}, \mathbf{y} \in \mathbf{C}^n$. By the property c) of norms we have:

$$\|\mathbf{x}\| = \|\mathbf{x} - \mathbf{y} + \mathbf{y}\| \leq \|\mathbf{x} - \mathbf{y}\| + \|\mathbf{y}\|,$$

and

$$\|\mathbf{x}\| - \|\mathbf{y}\| \leq \|\mathbf{x} - \mathbf{y}\|. \tag{1}$$

In addition

$$\|\mathbf{y}\| = \|\mathbf{x}+\mathbf{y}-\mathbf{x}\| \le \|\mathbf{y}-\mathbf{x}\| + \|\mathbf{x}\| = \|\mathbf{x}-\mathbf{y}\| + \|\mathbf{x}\|,$$

which implies

$$-(\|\mathbf{x}\| - \|\mathbf{y}\|) \le \|\mathbf{x}-\mathbf{y}\|. \tag{2}$$

From (1) and (2) we have

$$|\,\|\mathbf{x}\| - \|\mathbf{y}\|\,| \le \|\mathbf{x}-\mathbf{y}\|. \tag{3}$$

By setting

$$\mathbf{x}-\mathbf{y} = \sum_{i=1}^{n} (x_i - y_i)\mathbf{e}_i,$$

where $\mathbf{e}_i$ is the i-th vector of the canonical basis of $\mathbf{C}^n$, from (3) and the properties b) and c) we have:

$$|\,\|\mathbf{x}\| - \|\mathbf{y}\|\,| \le \sum_{i=1}^{n} |x_i - y_i|\,\|\mathbf{e}_i\| \le \max_{i=1,\ldots,n} |x_i - y_i| \sum_{i=1}^{n} \|\mathbf{e}_i\|.$$

Since the number $\alpha = \sum\limits_{i=1}^{n} \|\mathbf{e}_i\|$ is positive and does not depend either on $\mathbf{x}$ or on $\mathbf{y}$, we find that if

$$\max_{i=1,\ldots,n} |x_i - y_i| \le \frac{\epsilon}{\alpha}, \qquad \text{we have} \quad |\,\|\mathbf{x}\| - \|\mathbf{y}\|\,| \le \epsilon. \qquad \blacksquare$$

Additional important properties are given by the following theorems.

3.4 Theorem (equivalence of norms). *Let* $\|\,.\,\|'$ *and* $\|\,.\,\|''$ *be two vector norms. Then they are topologically equivalent: this means that two constants* α *and* $\beta \in \mathbf{R}$ *exist,* $0 < \alpha \le \beta$*, such that for every* $\mathbf{x} \in \mathbf{C}^n$ *the following inequalities hold:*

$$\alpha\|\mathbf{x}\|'' \le \|\mathbf{x}\|' \le \beta\|\mathbf{x}\|''. \tag{4}$$

Proof. It is sufficient to prove (4) when the norm $\|\,.\,\|''$ is the ∞-norm. In the general case (4) follows by comparison. If $\mathbf{x} = \mathbf{0}$, the inequalities are trivially true. If $\mathbf{x} \ne \mathbf{0}$, let us consider the set

$$S = \{\, \mathbf{y} \in \mathbf{C}^n \;:\; \|\mathbf{y}\|_\infty = 1 \,\},$$

which is closed and bounded, because it contains all the vectors whose entries have moduli smaller than or equal to one, with at least one entry of modulus one. Since $\| \, . \, \|'$ is a continuous function, then it has a maximum and a minimum in S:

$$\alpha = \min_{\mathbf{y} \in S} \|\mathbf{y}\|' \qquad \text{and} \qquad \beta = \max_{\mathbf{y} \in S} \|\mathbf{y}\|', \qquad 0 < \alpha \le \beta.$$

Since $\mathbf{y} \neq \mathbf{0}$, it follows that $\alpha \neq 0$, therefore for every $\mathbf{y} \in S$ we have

$$0 < \alpha \le \|\mathbf{y}\|' \le \beta. \tag{5}$$

For every $\mathbf{x} \in \mathbf{C}^n, \mathbf{x} \neq \mathbf{0}$, let us consider the vector

$$\mathbf{y} = \frac{\mathbf{x}}{\|\mathbf{x}\|_\infty};$$

we have $\|\mathbf{y}\|_\infty = 1$, so $\mathbf{y} \in S$ and

$$\|\mathbf{y}\|' = \left\| \frac{\mathbf{x}}{\|\mathbf{x}\|_\infty} \right\| = \frac{\|\mathbf{x}\|'}{\|\mathbf{x}\|_\infty},$$

by the property b) of vector norms. By replacing in (5), we have

$$\alpha \le \frac{\|\mathbf{x}\|'}{\|\mathbf{x}\|_\infty} \le \beta,$$

and finally

$$\alpha\|\mathbf{x}\|_\infty \le \|\mathbf{x}\|' \le \beta\|\mathbf{x}\|_\infty \qquad \blacksquare$$

The constants α e β verifying (4), for the norms defined in 3.2, are exhibited in next theorem.

3.5 Theorem. *For every* $\mathbf{x} \in \mathbf{C}^n$ *we have*

1. $$\|\mathbf{x}\|_\infty \le \|\mathbf{x}\|_2 \le \sqrt{n}\, \|\mathbf{x}\|_\infty;$$
2. $$\|\mathbf{x}\|_2 \le \|\mathbf{x}\|_1 \le \sqrt{n}\, \|\mathbf{x}\|_2;$$
3. $$\|\mathbf{x}\|_\infty \le \|\mathbf{x}\|_1 \le n\, \|\mathbf{x}\|_\infty.$$

Proof. For the inequalities in line 1, the proof of theorem 3.4 can be straightly referred. Let

$$S = \{\, \mathbf{x} \in \mathbf{C}^n \; : \; \|\mathbf{x}\|_\infty = 1 \,\}.$$

In S $\| \, . \, \|_2$ takes its minimum for those vectors $\mathbf{x}$ with only one nonzero entry having modulus one, and takes its maximum for those vectors $\mathbf{x}$ with all entries having modulus one, i.e. $|x_i| = 1, \; i = 1, \ldots, n$. Then

$$\alpha = \min_{\mathbf{x} \in S} \|\mathbf{x}\|_2 = 1, \qquad \beta = \max_{\mathbf{x} \in S} \|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^{n} 1} = \sqrt{n}.$$

The first inequality in line 2 is obtained by noticing that

$$\|\mathbf{x}\|_2^2 = \sum_{i=1}^n |x_i|^2 \le \sum_{i=1}^n |x_i|^2 + \sum_{\substack{i,j=1\\ i\neq j}}^n |x_i|\,|x_j| = \left[\sum_{i=1}^n |x_i|\right]^2 = \|\mathbf{x}\|_1^2.$$

For the second inequality in line 2, let us consider the vector $\mathbf{y}$ defined as

$$y_j = \begin{cases} \dfrac{|x_j|}{\overline{x}_j} & \text{if } x_j \neq 0, \\ 0 & \text{if } x_j = 0. \end{cases}$$

Then by the Cauchy-Schwarz inequality (see (1), chap. 1), we have:

$$|\mathbf{x}^H\mathbf{y}| \le \sqrt{\mathbf{x}^H\mathbf{x}}\,\sqrt{\mathbf{y}^H\mathbf{y}},$$

and since

$$|\mathbf{x}^H\mathbf{y}| = \left|\sum_{j=1}^n \overline{x}_j\, y_j\right| = \sum_{j=1}^n |x_j| = \|\mathbf{x}\|_1 \quad \text{e} \quad \mathbf{y}^H\mathbf{y} \le n,$$

it turns out that $\|\mathbf{x}\|_1 \le \sqrt{n}\,\|\mathbf{x}\|_2$.

The inequalities in line 3 are consequences of the ones in lines 1 and 2.

■

We remark that if A is unitary we have

$$\|A\mathbf{x}\|_2 = \|\mathbf{x}\|_2 \quad \text{for every } \mathbf{x} \in \mathbf{C}^n, \tag{6}$$

since

$$\|A\mathbf{x}\|_2 = \sqrt{(A\mathbf{x})^H(A\mathbf{x})} = \sqrt{\mathbf{x}^H A^H A\mathbf{x}} = \sqrt{\mathbf{x}^H\mathbf{x}} = \|\mathbf{x}\|_2.$$

2. Matrix norms

The notion of norm can be easily applied to square matrices. In addition to the three properties required by the definition of vector norm, a fourth property involving the product of matrices is included, as shown by the following definition.

3.6 Definition. A function from $\mathbf{C}^{n\times n}$ to $\mathbf{R}$

$$A \to \|A\|$$

which verifies the following properties

a) $\|A\| \geq 0$ e $\|A\| = 0$ if and only if $A = O$,

b) $\|\alpha A\| = |\alpha| \, \|A\|$ for every $\alpha \in \mathbf{C}$,

c) $\|A + B\| \leq \|A\| + \|B\|$ per ogni $B \in \mathbf{C}^{n\times n}$,

d) $\|AB\| \leq \|A\| \, \|B\|$ per ogni $B \in \mathbf{C}^{n\times n}$,

is called *matrix norm.* ∎

Also for matrix norms the same notation used for vector norms will be used. Since the properties a), b) and c) of matrix norms are the same as those of vector norms, it follows that also matrix norms are uniformly continuous functions, and that for them an equivalence theorem analogous to theorem 3.4 holds.

Now we will show how a particular matrix norm can be related to a given vector norm. First of all, we remark that, due to the continuity of vector norms, the set

$$\{\, \mathbf{x} \in \mathbf{C}^n \;:\; \|\mathbf{x}\| = 1 \,\}$$

is closed; moreover, by theorem 3.4 α exists such that $\|\mathbf{x}\|_\infty \leq \alpha\|\mathbf{x}\|$, i.e. $\max\limits_{i=1,\ldots,n} |x_i| \leq \alpha$, so the set S is also bounded. Since a continuous function admits maximum and minimum over a closed and bounded set of $\mathbf{C}^n$, the value

$$\max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\|$$

exists.

3.7 Theorem. *Let* $\| \,.\, \|$ *be a vector norm. The function*

$$A \to \max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\|, \quad A \in \mathbf{C}^{n\times n}, \quad \mathbf{x} \in \mathbf{C}^n$$

is a matrix norm.

Proof. We prove the properties a), b), c), d), of definition 3.6.

a) $\|A\mathbf{x}\| \geq 0$, therefore $\max\limits_{\|\mathbf{x}\|=1} \|A\mathbf{x}\| \geq 0$. Moreover if $A = O$, then $\|A\mathbf{x}\| = 0$ for every $\mathbf{x}$; conversely if $\max\limits_{\|\mathbf{x}\|=1} \|A\mathbf{x}\| = 0$, then $\|A\mathbf{x}\| = 0$ for every $\mathbf{x}$ such that $\|\mathbf{x}\| = 1$, therefore $A = O$.

b) $\max\limits_{\|\mathbf{x}\|=1} \|\alpha A\mathbf{x}\| = \max\limits_{\|\mathbf{x}\|=1} |\alpha| \, \|A\mathbf{x}\| = |\alpha| \max\limits_{\|\mathbf{x}\|=1} \|A\mathbf{x}\|$.

c) $\max\limits_{\|\mathbf{x}\|=1} \|(A + B)\mathbf{x}\| = \max\limits_{\|\mathbf{x}\|=1} \|A\mathbf{x} + B\mathbf{x}\| \leq \max\limits_{\|\mathbf{x}\|=1} (\|A\mathbf{x}\| + \|B\mathbf{x}\|)$

$$\le \max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\| + \max_{\|\mathbf{x}\|=1} \|B\mathbf{x}\|.$$

d) If $AB = O$, then $\|AB\| = 0$; thus the inequality is verified. If $AB \neq O$, then a vector $\mathbf{y}$ exists, with $\|\mathbf{y}\| = 1$, such that

$$\|AB\mathbf{y}\| = \max_{\|\mathbf{x}\|=1} \|AB\mathbf{x}\| \neq 0.$$

By setting $\mathbf{z} = B\mathbf{y}$, we have that $\mathbf{z} \neq \mathbf{0}$ (if $\mathbf{z} = \mathbf{0}$, we would have $(AB)\mathbf{y} = \mathbf{0}$ and then $AB = O$). Thus

$$\|(AB)\mathbf{y}\| = \|A(B\mathbf{y})\| = \|A\mathbf{z}\| = \|\mathbf{z}\|\frac{\|A\mathbf{z}\|}{\|\mathbf{z}\|} = \|B\mathbf{y}\|\left\|\frac{A\mathbf{z}}{\|\mathbf{z}\|}\right\|.$$

Since the vector $\mathbf{u} = \dfrac{\mathbf{z}}{\|\mathbf{z}\|}$ is such that $\|\mathbf{u}\| = 1$, then we have

$$\max_{\|\mathbf{x}\|=1} \|AB\mathbf{x}\| = \|A\mathbf{u}\| \, \|B\mathbf{y}\| \le \max_{\|\mathbf{v}\|=1} \|A\mathbf{v}\| \max_{\|\mathbf{w}\|=1} \|B\mathbf{w}\|. \qquad \blacksquare$$

3.8 Definition. The matrix norm

$$\|A\| = \max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\|,$$

is called *matrix norm induced* by the vector norm $\| \, . \, \|$. ■

3.9 Theorem. *The following induced matrix norms are obtained from the corresponding vector norms defined in 3.2:*

$$\|A\|_1 = \max_{j=1,\dots,n} \sum_{i=1}^{n} |a_{ij}| \qquad \textit{1-norm}$$

$$\|A\|_2 = \sqrt{\rho(A^H A)} \qquad \textit{2-norm}$$

$$\|A\|_\infty = \max_{i=1,\dots,n} \sum_{j=1}^{n} |a_{ij}| \qquad \infty\textit{-norm}$$

Proof. 1-norm - Let $\mathbf{x} \in \mathbf{C}^n$, such that $\|\mathbf{x}\|_1 = 1$. Then

$$\|A\mathbf{x}\|_1 = \sum_{i=1}^{n} \left| \sum_{j=1}^{n} a_{ij} x_j \right| \le \sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}||x_j| = \sum_{j=1}^{n} |x_j| \sum_{i=1}^{n} |a_{ij}|$$

$$\le \left[\max_{j=1,\dots,n} \sum_{i=1}^{n} |a_{ij}| \right] \sum_{j=1}^{n} |x_j| = \max_{j=1,\dots,n} \sum_{i=1}^{n} |a_{ij}|,$$

and therefore

$$\max_{\|\mathbf{x}\|_1=1} \|A\mathbf{x}\|_1 \le \max_{j=1,\dots,n} \sum_{i=1}^{n} |a_{ij}|.$$

Now a vector $\mathbf{x}$, $\|\mathbf{x}\|_1 = 1$, must be found, such that

$$\|A\mathbf{x}\|_1 = \max_{j=1,\dots,n} \sum_{i=1}^{n} |a_{ij}|.$$

This vector exists, because if k is the index of a column of A where the sum over all the entries' moduli attains its maximum, i.e.

$$\sum_{i=1}^{n} |a_{ik}| = \max_{j=1,\dots,n} \sum_{i=1}^{n} |a_{ij}|,$$

the vector $\mathbf{x} = \mathbf{e}_k$ is such that $\|\mathbf{x}\|_1 = 1$ and

$$\|A\mathbf{x}\|_1 = \|A\mathbf{e}_k\|_1 = \|(a_{1k}, \dots, a_{nk})^T\|_1 = \sum_{i=1}^{n} |a_{ik}|.$$

2-norm - Since the matrix $A^H A$ is hermitian, by theorem 2.26 we have

$$A^H A = U D U^H,$$

where U is unitary and D diagonal with the eigenvalues of $A^H A$ as principal entries. If $A = O$, then $\rho(A^H A) = 0$, and conversely, if $\rho(A^H A) = 0$, we have $D = O$ and $A = O$. If $A \ne O$, then

$$\mathbf{x}^H A^H A \mathbf{x} \ge 0 \quad \text{per } \mathbf{x} \ne \mathbf{0}.$$

Reasoning in the same way as in the proof of theorem 2.31, it turns out that the eigenvalues of $A^H A$ are nonnegative and that at least one of them, corresponding to the spectral radius of $A^H A$, satisfies the equation:

$$\lambda_1 = \rho(A^H A) > 0.$$

Let $\mathbf{x}$ such that $\|\mathbf{x}\|_2 = 1$ and $\mathbf{y} = U^H \mathbf{x}$; since U is unitary, from (6) we obtain $\|\mathbf{y}\|_2 = 1$ and therefore

$$\begin{aligned}
\max_{\|\mathbf{x}\|_2=1} \|A\mathbf{x}\|_2 &= \max_{\|\mathbf{x}\|_2=1} \sqrt{\mathbf{x}^H A^H A \mathbf{x}} = \max_{\|\mathbf{y}\|_2=1} \sqrt{\mathbf{y}^H D \mathbf{y}} \\
&= \max_{\|\mathbf{y}\|_2=1} \sqrt{\sum_{i=1}^{n} \lambda_i \, |y_i|^2} \le \max_{\|\mathbf{y}\|_2=1} \sqrt{\lambda_1 \sum_{i=1}^{n} |y_i|^2} \\
&= \sqrt{\lambda_1} = \sqrt{\rho(A^H A)}.
\end{aligned}$$

Now a vector $\mathbf{x}$, $\|\mathbf{x}\|_2 = 1$, must be found, such that

$$\|A\mathbf{x}\|_2 = \sqrt{\rho(A^H A)}.$$

This vector is $\mathbf{x}_1$, eigenvector of $A^H A$ for the eigenvalue λ_1, normalized, i.e. $\|\mathbf{x}_1\|_2 = 1$. In fact we have:

$$\mathbf{x}_1^H A^H A \mathbf{x}_1 = \lambda_1 \mathbf{x}_1^H \mathbf{x}_1 = \lambda_1 = \rho(A^H A).$$

∞-norm - As we did in the proof for the 1-norm, we can write:

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

Now a vector $\mathbf{x}$, $\|\mathbf{x}\|_\infty = 1$, must be found, such that

$$\|A\mathbf{x}\|_\infty = \max_{i=1,\ldots,n} \sum_{j=1}^{n} |a_{ij}|.$$

If $A = O$ then $\mathbf{x} = \mathbf{e}_1$ can be chosen, if $A \neq O$ the vector $\mathbf{x}$ can be chosen in the following way:

$$x_j = \begin{cases} \dfrac{|a_{kj}|}{a_{kj}} & \text{if } a_{kj} \neq 0 \\ 0 & \text{otherwise,} \end{cases}$$

where k is the index of the row of A where the sum over all the entries' moduli attains its maximum. ∎

If A is hermitian, then we have

$$\begin{aligned} &\|A\|_1 = \|A\|_\infty \\ &\|A\|_2 = \sqrt{\rho(A^H A)} = \sqrt{\rho(A^2)} = \sqrt{\rho^2(A)} = \rho(A), \end{aligned}$$

and if A is also positive definite then we have

$$\|A\|_2 = \lambda_{\max},$$

where $\lambda_{\max}$ is the largest eigenvalue of A.

Another matrix norm, frequently used because it can be computed straightly, is the following:

$$\|A\|_F = \sqrt{\sum_{i,j=1}^{n} |a_{ij}|^2} = \sqrt{\operatorname{tr}(A^H A)}, \tag{7}$$

which is called *Frobenius (or Schur)* norm of A.

The function (7) verifies the properties a), b) and c) in definition 3.6, as the entries of A can be seen as the entries of a vector $\mathbf{a} \in \mathbf{C}^m$, with $m = n^2$, so that $\|A\|_F = \|\mathbf{a}\|_2$. Concerning the property d), let $C = AB$, i.e.

$$c_{ij} = \mathbf{a}_i^T \mathbf{b}_j = \overline{\mathbf{a}}_i^H \mathbf{b}_j,$$

where $\mathbf{a}_i^T \in \mathbf{C}^{1\times n}$ is the i-th row of A and $\mathbf{b}_j \in \mathbf{C}^n$ is the j-th column of B. By using the Cauchy-Schwarz inequality (see (1), chap. 1) we have

$$|c_{ij}|^2 \le (\overline{\mathbf{a}}_i^H \overline{\mathbf{a}}_i)\ (\mathbf{b}_j^H \mathbf{b}_j) = (\mathbf{a}_i^H \mathbf{a}_i)\ (\mathbf{b}_j^H \mathbf{b}_j),$$

and consequently

$$\|C\|_F^2 = \sum_{i,j=1}^{n} |c_{ij}|^2 \le \sum_{i=1}^{n} \mathbf{a}_i^H \mathbf{a}_i \ \sum_{j=1}^{n} \mathbf{b}_j^H \mathbf{b}_j = \|A\|_F^2 \|B\|_F^2.$$

Let $U \in \mathbf{C}^{n\times n}$ be a unitary matrix. Since $(UA)^H UA = A^H A$, it turns out that

$$\|A\|_2 = \|UA\|_2 \quad \text{and} \quad \|A\|_F = \|UA\|_F,$$

and, as $A^H A$ and $(AU)^H AU$ are similar, we have

$$\|A\|_2 = \|AU\|_2 \quad \text{e} \quad \|A\|_F = \|AU\|_F.$$

Since $A^H A$ and $(UAU^H)^H UAU^H = UA^H AU^H$ are similar as well, we have also

$$\|A\|_2 = \|UAU^H\|_2 \quad \text{and} \quad \|A\|_F = \|UAU^H\|_F.$$

3. Properties of matrix norms

– Let $A \in \mathbf{C}^{n\times n}$ and $\mathbf{x} \in \mathbf{C}^n$. If $\| \,.\, \|$ is the matrix norm induced by the vector norm $\| \,.\, \|$, then

$$\|A\mathbf{x}\| \leq \|A\| \, \|\mathbf{x}\|.$$

In fact, when $\mathbf{x} = \mathbf{0}$, the inequality is trivially satisfied; when $\mathbf{x} \neq \mathbf{0}$, we have

$$\|A\mathbf{x}\| = \|\mathbf{x}\| \, \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} = \|\mathbf{x}\| \, \|A\mathbf{y}\|,$$

where the vector $\mathbf{y} = \dfrac{\mathbf{x}}{\|\mathbf{x}\|}$ is such that $\|\mathbf{y}\| = 1$ and therefore

$$\|A\mathbf{y}\| \leq \max_{\|\mathbf{z}\|=1} \|A\mathbf{z}\| = \|A\|.$$

– Since $\|AB\| \leq \|A\|\|B\|$, for every matrix norm, we have also

$$\|A^m\| \leq \|A\|^m \qquad \text{for every positive integer } m.$$

– Since $\|I\| = \|I\,I\| \leq \|I\| \, \|I\|$, we have $\|I\| \geq 1$ for every matrix norm.

– If $\| \,.\, \|$ is an induced matrix norm, then $\|I\| = 1$. In fact, by definition of matrix norm:

$$\|I\| = \max_{\|\mathbf{x}\|=1} \|I\mathbf{x}\| = \max_{\|\mathbf{x}\|=1} \|\mathbf{x}\| = 1.$$

For this reason the Frobenius norm cannot be an induced norm as

$$\|I\|_F = \sqrt{n} \neq 1 \quad \text{for } n > 1.$$

– If $A \in \mathbf{C}^{n\times n}$ is nonsingular , then, as

$$\|I\| = \|A^{-1}A\| \leq \|A^{-1}\| \, \|A\|,$$

for every matrix norm we have:

$$\|A^{-1}\| \geq \frac{1}{\|A\|}.$$

3.10 Theorem. *For every induced matrix norm* $\| \,.\, \|$ *the following inequality holds:*

$$\rho(A) \leq \|A\|.$$

Proof. Let λ be an eigenvalue and $\mathbf{x}$ a corresponding eigenvector, normalized with respect to the norm $\| \,.\, \|$:

$$A\mathbf{x} = \lambda\mathbf{x}, \qquad \|\mathbf{x}\| = 1.$$

Then

$$|\lambda| = \|A\mathbf{x}\|,$$

and consequently

$$|\lambda| \leq \max_{\|\mathbf{v}\|=1} \|A\mathbf{v}\| = \|A\|.$$

This relation holds for every eigenvalue λ of A and therefore also for the eigenvalues with maximum modulus. ∎

3.11 Theorem. *The function*

$$A \to \|S^{-1}AS\|_\infty,$$

where S *is a nonsingular matrix, is an induced matrix norm.*

Proof. The function

$$\mathbf{x} \to \|S^{-1}\mathbf{x}\|_\infty, \tag{8}$$

since S^{-1} is nonsingular, verifies the properties a), b) e c) of definition 3.1, thus it is a vector norm. The matrix norm induced by (8) is given by

$$\|A\| = \max_{\|S^{-1}\mathbf{x}\|_\infty=1} \|S^{-1}A\mathbf{x}\|_\infty = \max_{\|\mathbf{y}\|_\infty=1} \|S^{-1}AS\mathbf{y}\|_\infty,$$

where $\mathbf{y} = S^{-1}\mathbf{x}$. ∎

3.12 Theorem. *Let* $A \in \mathbf{C}^{n\times n}$*; then for any* $\epsilon > 0$ *there an induced norm* $\| \,.\, \|$ *exists such that*

$$\|A\| \leq \rho(A) + \epsilon.$$

Proof. Let J the Jordan canonical form of A (see theorem 2.18):

$$A = TJT^{-1},$$

where J is a block diagonal matrix, with blocks shaped in this way:

$$C_i^{(j)} = \begin{bmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \lambda_i & 1 \\ & & & \lambda_i \end{bmatrix}.$$

where λ_i is an eigenvalue A. Given the matrix

$$E = \begin{bmatrix} 1 & & & & \\ & \epsilon & & & \\ & & \epsilon^2 & & \\ & & & \ddots & \\ & & & & \epsilon^{n-1} \end{bmatrix},$$

the matrix

$$E^{-1}JE = E^{-1}T^{-1}ATE$$

is block diagonal too and each block has the following form:

$$D_i^{(j)} = \begin{bmatrix} \lambda_i & \epsilon & & \\ & \ddots & \ddots & \\ & & \lambda_i & \epsilon \\ & & & \lambda_i \end{bmatrix}.$$

We have also

$$\|E^{-1}JE\|_\infty = \|E^{-1}T^{-1}ATE\|_\infty = \max_{i,j} \|D_i^{(j)}\|_\infty \leq \rho(A) + \epsilon,$$

where the strict inequality holds when the blocks $D_i^{(j)}$ related to the eigenvalues of maximum modulus have size 1, and ϵ is sufficiently small. By theorem 3.11, $\|E^{-1}T^{-1}ATE\|_\infty$ is an induced matrix norm, applied to A. ■

Notice that, if the eigenvalues of modulus $\rho(A)$ have the same algebraic and geometric multiplicities, then an induced matrix norm $\| \,.\, \|$ exists such that

$$\|A\| = \rho(A).$$

In particular this happens when A is diagonalizable.

3.13 Theorem. *Let* $\| \,.\, \|$ *be an induced matrix norm and* $A \in \mathbf{C}^{n\times n}$*, with* $\|A\| < 1$*. Then the matrix* $I + A$ *is nonsingular and the following inequality holds*

$$\|(I + A)^{-1}\| \leq \frac{1}{1 - \|A\|}.$$

Proof. Since $\|A\| < 1$, By theorem 3.10 we have $\rho(A) < 1$. Therefore the matrix $I + A$ cannot have eigenvalue 0, so it is nonsingular. From the equation

$$(I + A)\,(I + A)^{-1} = I$$

it follows that

$$(I + A)^{-1} = I - A(I + A)^{-1},$$

and since $\|I\| = 1$, from the properties c) and d) of matrix norms we obtain:

$$\|(I + A)^{-1}\| \le 1 + \|A\|\,\|(I + A)^{-1}\|,$$

and consequently

$$(1 - \|A\|)\,\|(I + A)^{-1}\| \le 1.$$

The thesis follows, taking into account that $\|A\| < 1$. ∎

4. Norm inequalities

The norms introduced in previous sections satisfy the following inequalities, which can be proven by using theorem 3.5 and the definition of induced matrix norm:

$$\begin{aligned}
&\frac{1}{\sqrt{n}}\,\|A\|_\infty \le \|A\|_2 \le \sqrt{n}\,\|A\|_\infty,\\
&\frac{1}{\sqrt{n}}\,\|A\|_1 \le \|A\|_2 \le \sqrt{n}\,\|A\|_1,\\
&\max_{i,j} |a_{ij}| \le \|A\|_2 \le n \max_{i,j} |a_{ij}|,\\
&\|A\|_2 \le \sqrt{\|A\|_1\,\|A\|_\infty}
\end{aligned}$$

For instance, in detail, the last inequality can be proven as follows: since the eigenvalues of the positive semidefinite matrix $A^H A$ are nonnegative, the maximum eigenvalue $\lambda_{\max}$ of $A^H A$ is $\rho(A^H A)$, and from the equation

$$A^H A\mathbf{x} = \lambda_{\max}\,\mathbf{x},$$

by applying norms, it follows that

$$\begin{aligned}
\rho(A^H A)\,\|\mathbf{x}\|_\infty &= \lambda_{\max}\,\|\mathbf{x}\|_\infty = \|A^H A\mathbf{x}\|_\infty\\
&\le \|A^H\|_\infty\,\|A\|_\infty\,\|\mathbf{x}\|_\infty = \|A\|_1\,\|A\|_\infty\,\|\mathbf{x}\|_\infty,
\end{aligned}$$

and finally

$$\|A\|_2^2 \le \|A\|_1\,\|A\|_\infty.$$

Moreover

$$\rho(A^H A) = \lambda_{\max} \leq \sum_{i=1}^{n} \lambda_i = \operatorname{tr}(A^H A) \leq \sum_{i=1}^{n} \rho(A^H A) = n\, \rho(A^H A),$$

so the following equivalence relation between the 2-norm and the Frobenius norm is obtained:

$$\|A\|_2 \leq \|A\|_F \leq \sqrt{n}\, \|A\|_2.$$