

# Capítulo 2

## Normas de Vectores y Matrices.

### 2.1. Introducción

En este capítulo repasaremos brevemente el concepto de norma de un vector para centrarnos en el estudio de las normas de matrices o, si se quiere, en las normas de los operadores lineales. El estudio de las normas de matrices es importante por varias razones. Es necesario, por ejemplo, para definir de forma precisa conceptos tales como series de potencias de matrices; y desde luego es básico para precisar lo que se entiende por proximidad o lejanía entre matrices, aspectos fundamentales en el análisis de algoritmos en la computación numérica. Un par de ejemplos pueden servir para ilustrar estas ideas.

Es conocido que si  $x$  es un número complejo de módulo menor que 1 entonces

$$(1 - x)^{-1} = 1 + x + x^2 + x^3 + \dots$$

Esto sugiere la fórmula

$$(I - A)^{-1} = I + A + A^2 + A^3 + \dots$$

para calcular la inversa de la matriz  $I - A$ . Pero ¿cuándo es tal fórmula válida?. Resulta que es suficiente que una norma de la matriz  $A$  sea menor que 1, y además cualquier norma sirve. De forma parecida se puede ver que bajo ciertas condiciones relativas a la norma de  $A$  la serie

$$\sum_{k=0}^{\infty} \frac{1}{k} A^k$$

es convergente y sirve para definir la función matricial  $e^A$ .

Por otra parte, el cálculo numérico con matrices que proceden de datos experimentales, no es exacto; por lo general matrices están sometidas, bien sea por errores de redondeo o por imprecisión en las mediciones, a pequeñas perturbaciones. Cuán pequeñas son estas perturbaciones, o lo que es lo mismo, cuán lejos está la matriz verdadera de la calculada son conceptos que se pueden hacer precisos utilizando normas.

En todo este capítulo supondremos que  $\mathbb{F}$  es el cuerpo  $\mathbb{R}$  de los números reales o el cuerpo  $\mathbb{C}$  de los números complejos.

## 2.2. Normas de Vectores

### 2.2.1. Definición y Ejemplos

Como es bien sabido, el concepto de norma de un vector es una generalización del concepto de valor absoluto o módulo de un número complejo.

**Definición 2.1** .-Sea  $V$  un espacio vectorial sobre  $\mathbb{F}$  ( $\mathbb{R}$  o  $\mathbb{C}$ ). Una función  $\nu : V \longrightarrow \mathbb{R}$  es una norma en  $V$  si  $\nu$  satiface las siguientes propiedades:

$$(i) \quad x \neq 0 \Rightarrow \nu(x) > 0.$$

$$(ii) \quad \nu(\alpha x) = |\alpha| \nu(x), \quad \forall \alpha \in \mathbb{F}.$$

(iii)  $\nu(x + y) \leq \nu(x) + \nu(y)$ ,  $\forall x \in V$  (**desigualdad triangular**)

Tres importantes propiedades se siguen de forma inmediata de la Definición 2.2. Para cualquier norma  $\nu$

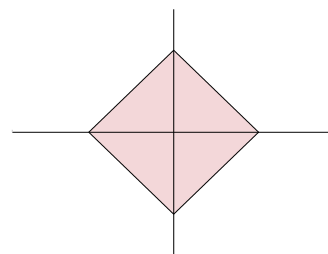
1.  $\nu(0) = 0$  porque  $\nu(0) = \nu(0x) = 0\nu(x) = 0$ .
2.  $\nu(-x) = \nu(x)$  porque  $\nu(-x) = |-1|\nu(x) = \nu(x)$
3.  $|\nu(x) - \nu(y)| \leq \nu(x - y)$ . En efecto, hay dos posibilidades:
  - a)  $\nu(x) \geq \nu(y)$ . En este caso  $|\nu(x) - \nu(y)| = \nu(x) - \nu(y)$  y por la desigualdad triangular  $\nu(x) = \nu((x - y) + y) \leq \nu(x - y) + \nu(y)$ .
  - b)  $\nu(x) < \nu(y)$ . Entonces  $|\nu(x) - \nu(y)| = \nu(y) - \nu(x)$  y de nuevo por la desigualdad triangular  $\nu(y) \leq \nu(y - x) + \nu(x) = \nu(x - y) + \nu(x)$ .

**Ejemplo 2.2** Se puede definir una infinidad de normas en  $\mathbb{F}^n$ , sin embargo las más utilizadas son las llamadas normas de Hölder o normas  $\ell_p$ . En lo que sigue supondremos que  $x = (x_1, x_2, \dots, x_n) \in \mathbb{F}^n$  es un vector de  $\mathbb{F}^n$ .

Para cada una de las normas dibujamos, a modo de ilustración la correspondiente bola unidad en  $\mathbb{R}^2$ ; i.e., el conjunto  $\bar{\mathcal{B}}(0, 1) = \{x \in \mathbb{R}^2 \mid \|x\| \leq 1\}$ .

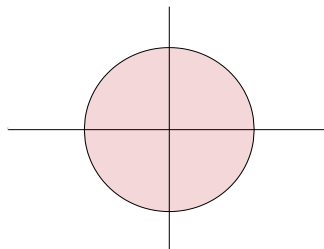
(a) La norma  $\ell_1$ :

$$\|x\|_1 = \sum_{i=1}^n |x_i|.$$



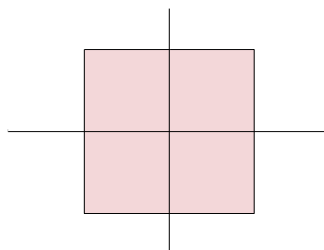
(b) La norma  $\ell_2$  o norma euclídea:

$$\|x\|_2 = \sqrt{\sum_{i=1}^n |x_i|^2}.$$



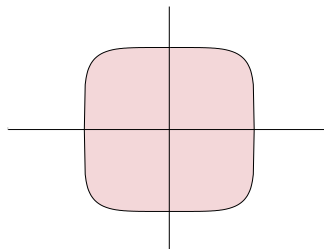
(c) La norma  $\ell_\infty$ :

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|.$$



(d) La norma  $\ell_p$  general ( $p \geq 2$ ):

$$\|x\|_p = \left( \sum_{i=1}^n |x_i|^p \right)^{1/p}.$$



Demostrar que las normas  $\ell_p$  son, en efecto, normas es fácil salvo la desigualdad triangular que, en el caso general, se conoce como desigualdad de Minkowsky. ésta a su vez, es consecuencia de otra desigualdad importante; la desigualdad de Hölder: Si  $p$  y  $q$  son números reales tales que  $\frac{1}{p} + \frac{1}{q} = 1$  entonces

$$|x^*y| \leq \|x\|_p \|y\|_q.$$

En el caso particular en que  $p = 2$  (y entonces también  $q = 2$ ) la desigualdad de Hölder se convierte en una desigualdad bien conocida:

$$|x^*y| \leq \|x\|_2 \|y\|_2,$$

la desigualdad de Cauchy-Schwartz.

Aparentemente la norma  $\ell_\infty$  no es una norma  $\ell_p$ . Sin embargo se la puede considerar como tal porque para cualquier  $x \in \mathbb{F}^n$  se tiene que

$$\|x\|_\infty = \lim_{p \rightarrow \infty} \|x\|_p.$$

Algo de esto ya se intuye en la forma que tienen las bolas unidad en  $\mathbb{R}^2$  para las normas  $\ell_p$ . También admite, claro está, una demostración rigurosa: Sea

$$\|x\|_p = (|x_1|^p + \cdots + |x_n|^p)^{1/p},$$

y supongamos que

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i| = |x_{i_1}| = \cdots = |x_{i_q}| = m.$$

Es decir, que las componentes  $i_1, \dots, i_q$  son, en módulo, mayores que todas las demás, y que en todas ellas el valor de dicho módulo es el mismo e igual a  $m$ . Así, si  $\{j_1, \dots, j_{n-q}\} = \{1, \dots, n\} \setminus \{i_1, \dots, i_q\}$  tenemos que

$$\begin{aligned} \|x\|_p &= (qm^p + |x_{j_1}|^p + \cdots + |x_{j_{n-q}}|^p)^{1/p} = \\ &= m \left( q + \left| \frac{x_{j_1}}{m} \right|^p + \cdots + \left| \frac{x_{j_{n-q}}}{m} \right|^p \right)^{1/p}. \end{aligned}$$

Como  $\left| \frac{x_{j_k}}{m} \right| < 1$  concluimos que  $\lim_{p \rightarrow \infty} \left| \frac{x_{j_k}}{m} \right|^p = 0$  y  $\lim_{p \rightarrow \infty} \|x\|_p = m = \|x\|_\infty$ , tal y como se deseaba demostrar.

### 2.2.2. Equivalencia de normas

Las normas son herramientas básicas para definir y analizar la convergencia de una sucesión de vectores en espacios vectoriales. Como es habitual, una sucesión  $\{x_k\} \subset V$  se dice que converge a  $x$  (y se escribe  $x_k \rightarrow x$ ) si la sucesión de números reales  $\{\nu(x_k - x)\}$  converge a cero,  $\nu(x_k - x) \rightarrow 0$ , siendo  $\nu$  una norma definida en  $V$ . Esta definición depende de la norma  $\nu$ . En principio podría suceder que una sucesión de vectores convergiera para una norma pero no para otra. De hecho, esto es perfectamente posible en espacios de dimensión infinita, pero no en espacios de dimensión finita. Ello es consecuencia de que todas las normas en un espacio vectorial de dimensión finita son equivalentes en el siguiente sentido:

**Definición 2.3** Sean  $\mu$  y  $\nu$  normas definidas en  $V$ , espacio vectorial sobre  $\mathbb{F}$ . Se dice que  $\mu$  y  $\nu$  son equivalentes si existen números reales positivos  $c_1$  y  $c_2$  tales que

$$c_1 \leq \frac{\nu(x)}{\mu(x)} \leq c_2, \quad \forall x \in V.$$

Debe notarse que así se define, en efecto, una relación de equivalencia en el conjunto de todas las normas definibles en un espacio vectorial. (Se verá en los ejercicios que hay una infinidad de normas definibles en cualquier espacio vectorial).

Nuestro objetivo es demostrar que todas las normas en un espacio de dimensión finita son equivalentes. Para ello debemos recordar algunos resultados básicos de topología. En todo espacio normado,  $(M, \|\cdot\|)$  podemos definir una distancia:

$$d(x, y) = \|x - y\|,$$

que hace de  $(M, d)$  un espacio métrico, y por lo tanto topológico con una base de abiertos dada por las bolas abiertas centradas en cada punto de  $M$ . Por consiguiente, si  $V$  es un espacio vectorial en el que tenemos definida una norma, automáticamente tenemos en  $V$  estructuras de espacio métrico y topológico. Si en  $M$  tenemos definidas dos normas, las bolas abiertas definidas por ambas normas pueden ser diferentes (véase el Ejemplo 2.2). A pesar de que las bases de abiertos determinadas por dichas normas sean diferentes, las topologías determinadas por éstos pueden ser iguales. La propiedad de que las normas son equivalentes determina que, en efecto, las topologías inducidas por ellas son la misma; es decir, todo abierto respecto de una de las normas lo es respecto de la otra.

En efecto, recordemos que si  $\nu$  es una norma en  $M$ , un conjunto  $A \subseteq M$  es abierto si para cada  $x \in A$  existe un número real  $\rho > 0$  tal que  $\mathcal{B}(x, \rho) = \{y \in V | \nu(x - y) < \rho\}$  está contenido en  $A$ . Ahora, si en  $M$  hay definidas dos normas equivalentes, digamos  $\nu$  y  $\mu$ , entonces existen constantes positivas  $c_1$  y  $c_2$  tales que  $c_1\mu(x) \leq \nu(x) \leq c_2\mu(x)$  para todo  $x \in M$ . Supongamos que  $A$  es abierto respecto de la norma  $\nu$ . Esto significa que para todo  $x \in A$  existe  $r > 0$  tal que  $\nu(x - y) < r$  implica que  $y \in A$ . Veamos que  $A$  es abierto respecto de  $\mu$ . Sea  $x \in A$  un elemento cualquiera y definamos  $s$  de modo que  $0 < s < \frac{r}{c_2}$ . Si  $\mu(x - y) < s$  tenemos que

$$\nu(x - y) \leq c_2\mu(x - y) < c_2 \frac{r}{c_2} = r.$$

Por lo tanto  $y \in A$  y  $A$  es abierto respecto a  $\mu$ . La demostración de que todo abierto respecto a  $\mu$  lo es respecto a  $\nu$  se hace igual.

Así pues, todas las normas equivalentes definidas en un conjunto  $M$  inducen la misma topología en dicho conjunto. En consecuencia, todos los invariantes topológicos (compacidad, conexión, ...) se mantienen. Es decir, si un conjunto es compacto o conexo respecto de una de las normas lo es respecto de la otra.

Estas propiedades tienen un interés sobre todo teórico, pero nos ahorra tener que estar prestando atención continuamente a las normas empleadas cuando se prueban resultados topológicos. En cualquier caso, una primera consecuencia es que la convergencia de sucesiones de vectores en espacios de dimensión finita es independiente de la norma elegida.

La demostración de que todas las normas definidas en un espacio vectorial de dimensión finita son equivalentes se puede hacer de diversas formas. La que adoptamos aquí no es la más directa pero tiene la virtud de poner de manifiesto una propiedad fundamental de los espacios vectoriales de dimensión finita: estos espacios vectoriales están caracterizados por el hecho de que las esferas unidades respecto de cualquier norma son conjuntos compactos. Demostraremos esta importante propiedad (que usaremos de manera significativa posteriormente) en el siguiente Lema. En la prueba usaremos que las normas  $\ell_1$  y  $\ell_2$  en  $\mathbb{F}^n$  son equivalentes. En efecto, tal y como se propone demostrar en los ejercicios, se tiene la siguiente desigualdad para todo  $x \in \mathbb{F}^n$ :

$$\|x\|_2 \leq \|x\|_1 \leq \sqrt{n}\|x\|_2. \quad (2.1)$$

Una última observación: nótese que la acotación no se conserva por homeomorfismo (el intervalo  $(0, 1)$  es homeomorfo a  $\mathbb{R}$ ) pero si dos normas son equivalentes y un conjunto es acotado respecto de una de ellas lo es, por la definición de equivalencia de normas, respecto de la otra.

**Lema 2.4** (a) *Todas las normas definidas en  $V$ , espacio vectorial sobre  $\mathbb{F}$ , son funciones continuas cuando en  $V$  se considera la topología inducida por  $\nu$  y en  $\mathbb{R}$  la topología habitual (la inducida por el valor absoluto).*

(b) *Sea  $V$  un espacio vectorial de dimensión finita y  $\nu$  una norma en  $V$ . La esfera unidad de  $V$  respecto de la norma  $\nu$ ,  $S_\nu = \{x \in V \mid \nu(x) = 1\}$ , es un conjunto compacto con la topología en  $V$  inducida por dicha norma.*

**Demostración.-**

- (a) Sea  $\nu$  una norma definida en  $V$  y consideremos en  $V$  la topología inducida por esta norma. Sea  $\varepsilon > 0$  un número real arbitrario. Escogemos  $0 < \delta < \varepsilon$ . Entonces, como para todo  $x, y \in V$

$$|\nu(x) - \nu(y)| \leq \nu(x - y),$$

resulta que si  $\nu(x - y) < \delta$  entonces  $|\nu(x) - \nu(y)| < \varepsilon$ . Esto demuestra que  $\nu$  es continua.

- (b) Supongamos que  $\dim V = n$  y sea  $\{v_1, \dots, v_n\}$  una base de  $V$ . Definimos la siguiente función:

$$\begin{aligned} g : \mathbb{F}^n &\rightarrow V \\ a = (a_1, \dots, a_n) &\rightarrow v = a_1 v_1 + \dots + a_n v_n \end{aligned}$$

donde la topología en  $V$  es la derivada de la norma  $\nu$ .

Esta aplicación es un isomorfismo de espacios vectoriales; es decir, es una aplicación lineal y biyectiva. Y desde el punto de vista topológico es también un homeomorfismo cuando en  $V$  y  $\mathbb{F}^n$  se consideran las topologías derivadas de las normas  $\nu$  y  $\ell_1$ . Es decir,  $g$  es una aplicación continua con inversa continua.

No obstante, para llegar al resultado deseado sólo necesitamos que  $g$  sea continua. Aunque, tal y como se acaba de decir, éste es un hecho conocido, lo demostramos porque la demostración es muy sencilla y esta propiedad se usará con frecuencia en este curso. Que la inversa de  $g$  también es continua requiere más trabajo. Aunque no se va a utilizar, demostraremos esta propiedad al final de la sección.

Sea  $\varepsilon > 0$  un número real dado. Tenemos que encontrar un número real  $\delta > 0$  tal que si  $\|a - b\|_1 < \delta$  entonces  $\nu(g(a) - g(b)) < \varepsilon$ . Ahora bien

$$\begin{aligned} \nu(g(a) - g(b)) &= \nu\left(\sum_{i=1}^n (a_i v_i - b_i v_i)\right) \leq \\ &\leq \sum_{i=1}^n |a_i - b_i| \nu(v_i) \\ &\leq M \|a - b\|_1, \end{aligned}$$

donde  $M = \max\{\nu(v_1), \dots, \nu(v_n)\}$ . Observamos que  $M > 0$  porque  $v_i \neq 0$  para todo  $i = 1, \dots, n$  por ser vectores de una base de  $V$ . Basta escoger  $0 < \delta < \frac{\varepsilon}{M}$  para obtener el resultado deseado.

Sea  $S_\nu \subset V$  la esfera unidad respecto de la norma  $\nu$ :  $S_\nu = \{x \in V | \nu(x) = 1\}$ . Este conjunto es acotado respecto de la métrica determinada por  $\nu$ . Y



también es cerrado porque es la anteimagen por la aplicación continua  $\nu$  del cerrado  $\{1\}$  de  $\mathbb{R}$ . Como  $g$  es continua,  $g^{-1}(S_\mu)$  es cerrado. El conjunto  $g^{-1}(S)$  también es acotado. Esto es, en realidad, una consecuencia inmediata de un resultado que veremos en la próxima lección. Sin dicho resultado, necesitamos una demostración. Ya hemos visto que  $g : \mathbb{F}^n \rightarrow V$  y  $\nu : V \rightarrow \mathbb{R}$  son funciones continuas. Entonces, la composición de ambas

$$\begin{aligned} f : \mathbb{F}^n &\longrightarrow \mathbb{R} \\ (a_1, \dots, a_n) &\rightsquigarrow \nu(a_1 v_1 + \dots + a_n v_n) \end{aligned}$$

también lo es. Sea  $S_1^{n-1} = \{x \in \mathbb{F}^n \mid \|x\|_1 = 1\}$  la esfera unidad de  $\mathbb{F}^n$  respecto de la norma  $\ell_1$ . Este conjunto es cerrado y acotado respecto de la norma  $\ell_1$ . Por (2.1) también es cerrado y acotado respecto de la norma euclídea. Ahora bien, sabemos (Teorema de Heine-Borel) que ser cerrado y acotado en  $\mathbb{F}^n$  ( $\mathbb{F} = \mathbb{R}$  o  $\mathbb{C}$ ) respecto de la norma euclídea equivale a ser compacto. Así pues  $S_1^{n-1}$  es compacto. Consecuentemente,  $f$  restringida a  $S_1^{n-1}$  alcanza su mínimo. Sea  $m = \min_{y \in S_1^{n-1}} f(y)$ . Notemos que  $m > 0$  porque si  $y \in S_1^{n-1}$  entonces  $y \neq 0$ , por lo que  $v_y = y_1 v_1 + \dots + y_n v_n \neq 0$  y  $f(y) = \nu(v_y) > 0$ .

Sea ahora  $a = (a_1, \dots, a_n) \in g^{-1}(S_\nu)$ . Esto significa que  $x = a_1 v_1 + \dots + a_n v_n \in S_\nu$  y así

$$\begin{aligned} 1 = \nu(x) &= \nu\left(\sum_{i=1}^n a_i v_i\right) = \\ &= \|a\|_1 \nu\left(\sum_{i=1}^n \frac{a_i}{\|a\|_1} v_i\right) \end{aligned}$$

Pero como  $\left(\frac{a_1}{\|a\|_1}, \dots, \frac{a_n}{\|a\|_1}\right) \in S_1^{n-1}$  resulta que

$$\nu\left(\sum_{i=1}^n \frac{a_i}{\|a\|_1} v_i\right) = f\left(\frac{a_1}{\|a\|_1}, \dots, \frac{a_n}{\|a\|_1}\right) \geq m.$$

Por lo tanto

$$1 = \nu(x) \geq \|a\|_1 m,$$

y como  $m > 0$  obtenemos  $\|a\|_1 \leq \frac{1}{m}$ . Así pues,  $g^{-1}(S_\nu)$  está acotado en la norma  $\ell_1$  (y por lo tanto en la norma  $\ell_2$ ). Y como también es cerrado en ambas normas por ser equivalentes, y  $\mathbb{F} = \mathbb{R}$  o  $\mathbb{C}$ ,  $g^{-1}(S_\nu)$  es compacto. Finalmente, como  $g$  biyectiva,  $g(g^{-1}(S_\nu)) = S_\nu$ . Ahora bien, la imagen de un compacto por una aplicación continua es compacto. Concluimos entonces que  $S_\nu$  es compacto, tal y como se deseaba demostrar. ■

El recíproco del apartado (b) del lema anterior también es verdadero. Aunque no lo vamos a utilizar lo exponemos por completitud. La demostración usa el siguiente resultado de Riesz, que no demostramos:

**Lema 2.5 (Lemma de Riesz)** *Sea  $E$  un espacio vectorial y  $\nu$  una norma definida en él. Sea  $S \subset E$  un subespacio propio y  $0 < \alpha < 1$ . Existe  $x_\alpha \in E \setminus S$  tal que  $\nu(x_\alpha) = 1$  y  $\nu(s - x_\alpha) > \alpha$  para cada  $s \in S$ .*

**Teorema 2.6** *Un espacio vectorial  $V$  es de dimensión finita si y sólo si la esfera unidad en  $V$ , respecto de cualquier norma, es un conjunto compacto.*

**Demostración.**— La necesidad de que la esfera unidad sea un conjunto compacto para que  $V$  sea de dimensión finita se ha probado en el Lema 2.4. Demostremos ahora la suficiencia.

Sea  $\nu$  una norma en  $V$  y supongamos que  $V$  es de dimensión infinita. Sea  $x_1 \in V$  cualquier vector tal que  $\nu(x_1) = 1$ . Sea  $V_1 = \langle x_1 \rangle$  el subespacio generado por  $x_1$ . Como  $V$  es de dimensión infinita,  $V_1$  es un subespacio propio de  $V$ , por el Lema de Riesz, existe  $x_2 \in V$  tal que  $\nu(x_2) = 1$  y  $\nu(x_2 - y) \geq \frac{1}{2}$  para todo  $y \in V_1$ .

Sea  $V_2 = \langle x_1, x_2 \rangle$  el subespacio generado por  $x_1$  y  $x_2$ . De nuevo,  $V_2 \subset V$  es un subespacio propio de  $V$ . Aplicando otra vez el Lema de Riesz, existe  $x_3 \in V$  tal que  $\nu(x_3) = 1$  y  $\nu(x_3 - y) \geq \frac{1}{2}$  para todo  $y \in V_2$ .

Siguiendo así sucesivamente, construimos una sucesión de vectores  $\{x_n\}$  con las siguientes propiedades:

- (a)  $\nu(x_n) = 1$  para  $n = 1, 2, \dots$
- (b)  $\nu(x_n - x_m) \geq \frac{1}{2}$  para  $n \neq m$ ,  $n, m \in \mathbb{N}$ .

Por lo tanto hemos construido una sucesión de vectores en la esfera unidad de  $V$ , respecto de la norma  $\nu$ , que no admite una subsucesión convergente. Esto significa que la esfera unidad no es un conjunto compacto en  $V$ , contradiciendo la hipótesis. ■

Demostremos finalmente el resultado sobre la equivalencia de normas que buscábamos.

**Teorema 2.7** *Todas las normas definidas en un espacio vectorial de dimensión finita son equivalentes.*

**Demostración.** Sea  $V$  un espacio vectorial de dimensión finita sobre  $\mathbb{F}$ , y sean  $\nu$  y  $\mu$  dos normas definidas en  $V$ . Sea  $S_\mu = \{x \in V \mid \mu(x) = 1\}$ .

Por una parte, este conjunto es compacto con la topología en  $V$  inducida por la norma  $\mu$ . Hemos visto en la demostración del Lema 2.4 que si

$$\begin{aligned} g : \mathbb{F}^n &\rightarrow V \\ a = (a_1, \dots, a_n) &\rightarrow v = a_1 v_1 + \dots + a_n v_n \end{aligned}$$

entonces  $g^{-1}(S_\mu)$  es compacto en  $\mathbb{F}^n$  con la topología habitual (la derivada de la norma euclídea o de la  $\ell_1$ ). También se ha demostrado  $g$  es continua cuando en  $V$  se considera la topología derivada de cualquier norma, en particular, de la norma  $\nu$ . Como la imagen de un compacto por una aplicación continua es un compacto y  $g$  es biyectiva, concluimos que  $S_\mu$  es compacto en  $V$  con la topología inducida por la norma  $\nu$ .

Así pues, cuando en  $V$  consideramos la topología inducida por  $\nu$  tenemos que  $\nu$  es una aplicación continua y  $S_\mu$  un compacto. Por lo tanto,  $\nu$  alcanza en  $S_\mu$  su máximo y su mínimo. Pongamos  $c_2 = \min_{y \in S_\mu} \nu(y)$  y  $c_1 = \max_{y \in S_\mu} \nu(y)$ . Entonces, para

$x \in V$ , tenemos que  $\frac{x}{\mu(x)} \in S_\mu$  y

$$\nu(x) = \mu(x) \nu\left(\frac{x}{\mu(x)}\right) \geq \mu(x) \min_{y \in S_\mu} \nu(y) = c_2 \mu(x).$$

Y de la misma forma

$$\nu(x) = \mu(x) \nu\left(\frac{x}{\mu(x)}\right) \leq \mu(x) \max_{y \in S_\mu} \nu(y) = c_1 \mu(x). \quad \blacksquare$$

Una consecuencia inmediata de la equivalencia de normas en espacios de dimensión finita es que los isomorfismos algebraicos de espacios vectoriales normados son isomorfismos topológicos u homeomorfismos:

**Corolario 2.8** *Sean  $V_1$  y  $V_2$  dos espacios vectoriales de dimensión finita  $n$  sobre  $\mathbb{F}$  y sea  $f : V_1 \rightarrow V_2$  un isomorfismo. Sean  $\nu$  y  $\mu$  normas definidas en  $V_1$  y  $V_2$ , respectivamente. Entonces  $f$  y  $f^{-1}$  son aplicaciones continuas cuando en  $V_1$  y  $V_2$  se consideran las topologías inducidas por estas normas.*

**Demostración** Probamos la continuidad de  $f$ , la de  $f^{-1}$  se demuestra igual.

En primer lugar, es fácil demostrar que la aplicación  $\eta : V_1 \rightarrow \mathbb{R}$  definida por  $\eta(x) = \mu(f(x))$  es una norma en  $V_1$ . Como  $V_1$  es de dimensión finita,  $\eta$  y  $\nu$  son equivalentes. Existe una constante positiva  $c_2$  tal que  $\eta(x) \leq c_2\nu(x)$  para todo  $x \in V_1$ .

Sea ahora  $\varepsilon > 0$  un número real dado. Escojamos  $\delta$  de modo que  $0 < \delta < \frac{\varepsilon}{c_2}$ . Entonces, si  $x, y \in V_1$  cumplen que  $\nu(x - y) < \delta$ , se sigue

$$\mu(f(x) - f(y)) = \mu(f(x - y)) = \eta(x - y) \leq c_2\nu(x - y) < c_2\delta < \varepsilon.$$

Esto demuestra que  $f$  es uniformemente continua en  $V_1$ . ■

Una consecuencia inmediata es el siguiente resultado anunciado en la demostración del Lema 2.4: la inversa de

$$\begin{array}{ccc} g : & \mathbb{F}^n & \rightarrow V \\ & a = (a_1, \dots, a_n) & \rightarrow v = a_1v_1 + \dots + a_nv_n \end{array}$$

es una aplicación continua.

## 2.3. Normas de Matrices

### 2.3.1. Definiciones, Ejemplos y Primeras Propiedades

Dado que  $\mathbb{F}^{m \times n}$  es un espacio vectorial todo lo dicho en la sección anterior es de aplicación a las matrices de tamaño  $m \times n$  con elementos en el cuerpo de los números reales o complejos. En particular, todas las normas definidas en  $\mathbb{F}^{m \times n}$  son equivalentes; i.e. generan la misma topología, y son funciones continuas.

Debemos observar que las propiedades que definen una norma tienen en cuenta las operaciones propias de la estructura de espacio vectorial sobre el que se definen: el producto por escalares y la suma de vectores (desigualdad triangular). Sin embargo las matrices, en algunos casos, se pueden multiplicar. Una norma sobre  $\mathbb{F}^{n \times n}$  se dirá que es una norma de matriz cuando respecto al producto se verifica una propiedad similar a la desigualdad triangular para la suma. Tomar en consideración esta propiedad de las matrices nos conduce a la siguiente definición:

**Definición 2.9** .- Sean  $\mu$ ,  $\nu$  y  $\rho$  normas definidas en  $\mathbb{F}^{m \times n}$ ,  $\mathbb{F}^{n \times p}$  y  $\mathbb{F}^{m \times p}$ , respectivamente. Diremos que  $\mu$ ,  $\nu$  y  $\rho$  son **consistentes** si para todas matrices  $A \in \mathbb{F}^{m \times n}$  y  $B \in \mathbb{F}^{n \times p}$  se verifica

$$\rho(AB) \leq \mu(A)\nu(B).$$

En particular una norma  $\nu$  definida en  $\mathbb{F}^{n \times n}$  se dice que es consistente si  $\nu(AB) \leq \nu(A)\nu(B)$  para todas  $A, B \in \mathbb{F}^{n \times n}$ .

Una norma  $\nu$  definida en  $\mathbb{F}^{n \times n}$  consistente también se dice que es **multiplicativa** o **submultiplicativa**.

Una norma definida en  $\mathbb{F}^{n \times n}$  se dice que es una **norma de matriz** si es consistente.

Un caso particular importante es el de las normas consistentes (también llamadas en este caso compatibles) con las normas de vector: Si  $\mu$  es una norma definida en  $\mathbb{F}^{n \times n}$  y  $\nu$  es una norma definida en  $\mathbb{F}^n$  entonces se dice que  $\mu$  es consistente o compatible con la norma de vector  $\nu$  si para toda matriz  $A \in \mathbb{F}^{n \times n}$  y todo vector  $x \in \mathbb{F}^n$  se tiene que

$$\nu(Ax) \leq \mu(A)\nu(x)$$

**Ejemplo 2.10** 1. - La norma euclídea en  $\mathbb{F}^{m \times n}$  será

$$\|A\| = \left( \sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 \right)^{\frac{1}{2}}$$

A esta norma se le llama también norma de Frobenius y se suele representar por  $\|A\|_F$ . En algunos sitios se le llama también norma de Schur o de Hilbert-Schmidt. Nosotros utilizaremos el nombre de norma euclídea o de Frobenius.

La norma de Frobenius en  $\mathbb{F}^{n \times n}$  es una **norma de matriz**. Más general, las normas de Frobenius en  $\mathbb{F}^{m \times n}$ ,  $\mathbb{F}^{n \times p}$  y  $\mathbb{F}^{m \times p}$  son consistentes. En efecto, supongamos que  $A \in \mathbb{F}^{m \times n}$  y  $B \in \mathbb{F}^{n \times p}$ . Entonces  $AB \in \mathbb{F}^{m \times p}$  y, llamando  $a'_j$  a la  $j$ -ésima columna de  $A^*$ , tenemos:

$$\|AB\|_F^2 = \sum_{i=1}^m \sum_{j=1}^p |a^i b_j|^2 = \sum_{i=1}^m \sum_{j=1}^p |a'_i{}^* b_j|^2$$

Usamos ahora la desigualdad de Cauchy-Schwartz:  $|x^*y| \leq \|x\|_2 \|y\|_2$ . Así

$$\begin{aligned} \|AB\|_F^2 &\leq \sum_{i=1}^m \sum_{j=1}^p \|a'_i\|_2^2 \|b_j\|_2^2 \\ &= \sum_{k=1}^n \sum_{i=1}^m |a'_{ki}|^2 \sum_{k=1}^n \sum_{j=1}^p |b_{kj}|^2 \\ &= \sum_{i=1}^m \sum_{k=1}^n |a_{ik}|^2 \sum_{k=1}^n \sum_{j=1}^p |b_{kj}|^2 = \|A\|_F^2 \|B\|_F^2 \end{aligned}$$

donde hemos usado que para un número complejo  $|z| = |\bar{z}|$ .

2. - La norma  $\ell_\infty$  en  $\mathbb{F}^{m \times n}$  será:

$$\|A\| = \max_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} |a_{i,j}|$$

Esta norma en  $\mathbb{F}^{n \times n}$  no es una norma de matriz como lo demuestra el siguiente ejemplo:

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad A^2 = AA = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$$

y

$$\|AA\| = 2 > 1 = \|A\| \|A\|$$

3. - En general la norma  $\ell_p$  en  $\mathbb{F}^{m \times n}$  se definirá como:

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

Utilizando la desigualdad de Hölder se puede demostrar que en  $\mathbb{F}^{n \times n}$  esta norma es una norma de matriz si y sólo si  $1 \leq p \leq 2$ .

4. - Una pequeña modificación en la definición de la norma  $\ell_\infty$  nos permite obtener una norma de matriz en  $\mathbb{F}^{n \times n}$ :

$$\|A\| = n \max_{1 \leq i, j \leq n} |a_{i,j}|$$

En efecto, si  $a = \max_{1 \leq i, j \leq n} |a_{ij}|$  y  $b = \max_{1 \leq i, j \leq n} |b_{ij}|$  entonces

$$\begin{aligned} \|AB\| &= n \max_{1 \leq i, j \leq n} \left| \sum_{k=1}^n a_{ik} b_{kj} \right| \leq n \max_{1 \leq i, j \leq n} \sum_{k=1}^n |a_{ik} b_{kj}| \leq \\ &\leq n \max_{1 \leq i, j \leq n} \sum_{k=1}^n ab = na nb = \|A\| \|B\| \end{aligned}$$

La norma de Frobenius se puede escribir alternativamente de cualquiera de las dos formas siguientes:

$$\text{a) } \|A\|_F = \sqrt{\text{tr}(A^*A)}$$

$$\text{b) } \|A\|_F^2 = \|a_1\|_2^2 + \|a_2\|_2^2 + \dots + \|a_n\|_2^2, \text{ donde } A = [a_1 \ a_2 \ \dots \ a_n] \in \mathbb{F}^{n \times n} \text{ está escrita en términos de sus vectores columna.}$$

### 2.3.2. Normas de Matriz Inducidas

Una técnica general para obtener normas de matriz a partir de normas vectoriales es la siguiente: Sea  $A \in \mathbb{F}^{m \times n}$  y pensemos en  $A$  como una transformación u operador lineal:

$$\begin{aligned} A &: \mathbb{F}^n \rightarrow \mathbb{F}^m \\ x &\rightarrow Ax \end{aligned}$$

Consideremos en  $\mathbb{F}^n$  y  $\mathbb{F}^m$  normas  $\|\cdot\|_n$  y  $\|\cdot\|_m$ , respectivamente. Es un hecho conocido que las aplicaciones lineales entre espacios vectoriales normados de dimensión finita son operadores acotados. Que un operador  $h: E_1 \rightarrow E_2$  entre dos espacios normados es acotado significa que  $\forall x \in E_1$  existe una constante  $M > 0$  tal que  $\|h(x)\| \leq M\|x\|$  donde hemos representado con el mismo signo,  $\|\cdot\|$ , y sin posibilidad de confusión, las dos normas definidas en  $E_1$  y  $E_2$ . Lo que sigue es una demostración de esta propiedad. Si  $E_1$  y  $E_2$  son espacios vectoriales de dimensión finita, hemos visto en el Corolario 2.8 que las aplicaciones lineales son continuas. Si  $x \in E_1$  entonces  $\frac{x}{\|x\|}$  está en la esfera unidad  $S_1$  de  $E_1$  que, por el Lema 2.4, es un conjunto compacto en  $E_1$ . Por consiguiente, si  $h$  es lineal entonces, como las normas son siempre funciones continuas,  $\|\cdot\| \circ h$  restringido a  $S_1$  alcanza su máximo y mínimo. Sea  $M = \max_{y \in S_1} \|h(y)\|$ . Entonces  $\|h(\frac{x}{\|x\|})\| \leq M$ . Pero  $\|h(\frac{x}{\|x\|})\| = \frac{1}{\|x\|} \|h(x)\|$ , de donde se sigue que  $\|h(x)\| \leq M\|x\|$ .

En particular para la aplicación lineal  $A$  se tiene que  $\forall x \in \mathbb{F}^n$  existe una constante  $M > 0$  tal que  $\|Ax\|_m \leq M\|x\|_n$ . Y esto significa que el conjunto

$$\left\{ \frac{\|Ax\|_m}{\|x\|_n} : x \in \mathbb{F}^n, x \neq 0 \right\}$$

está acotado superiormente. Consecuentemente, tiene un supremo. Podemos así definir la siguiente función:

$$\begin{aligned} f &: \mathbb{F}^{m \times n} \rightarrow \mathbb{R} \\ A &\rightarrow f(A) \end{aligned}$$

donde

$$f(A) = \sup_{0 \neq x \in \mathbb{F}^n} \frac{\|Ax\|_m}{\|x\|_n}$$

Es claro que el cociente

$$\frac{\|Ax\|_m}{\|x\|_n}$$

es no negativo. Probaremos que  $f$  es una norma consistente con las normas de vector que sirven para definirla.

Comenzaremos observando que el *sup* se alcanza y que, por lo tanto, podemos sustituir *sup* por *max*. Para ello probamos, en primer lugar, que (supondremos siempre  $x \in \mathbb{F}^n$ ):

$$f(A) = \sup_{\|x\|_n=1} \|Ax\|_m.$$

En efecto si  $M = \sup_{0 \neq x} \frac{\|Ax\|_m}{\|x\|_n}$  entonces  $M \geq \frac{\|Ax\|_m}{\|x\|_n}, \forall x \neq 0$ . En particular  $M \geq \|Ax\|_m, \forall x \in \mathbb{F}^n$  tal que  $\|x\|_n = 1$ , con lo que  $M \geq \sup_{\|x\|_n=1} \|Ax\|_m$ . Recíprocamente, si  $M = \sup_{\|x\|_n=1} \|Ax\|_m$ , entonces  $M \geq \|Ax\|_m, \forall x \in \mathbb{F}^n$  tal que  $\|x\|_n = 1$ . Sea  $x \in \mathbb{F}^n$  un vector no nulo arbitrario; entonces  $\frac{x}{\|x\|_n}$  es un vector unitario, de modo que  $M \geq \|A \frac{x}{\|x\|_n}\|_m = \frac{1}{\|x\|_n} \|Ax\|_m$ . Por lo tanto  $M \geq \sup_{0 \neq x \in \mathbb{F}^n} \frac{\|Ax\|_m}{\|x\|_n}$ . En conclusión  $f(A) = \sup_{\|x\|_n=1} \|Ax\|_m$ .

Ahora bien,  $\|\cdot\|_m$  es una función continua respecto de cualquier norma de  $\mathbb{F}^n$  y también  $A$  es una función continua (por ser una función lineal). Por otra parte, la esfera unidad es un conjunto compacto (respecto de cualquier norma) de  $\mathbb{F}^n$ , y cualquier función continua alcanza sus valores máximo y su mínimo en ella. Así pues

$$f(A) = \max_{\|x\|_n=1} \|Ax\|_m = \max_{0 \neq x \in \mathbb{F}^n} \frac{\|Ax\|_m}{\|x\|_n}.$$

Demostraremos ahora que  $f(A)$  es una norma vectorial consistente con las normas  $\|\cdot\|_m$  y  $\|\cdot\|_n$ . A estas normas se les llama también **normas de operador**.



**Teorema 2.11** .- Sea  $\|\cdot\|_n$  y  $\|\cdot\|_m$  normas definidas en  $\mathbb{F}^n$  y  $\mathbb{F}^m$ , respectivamente.

a) La función real  $f$  definida en  $\mathbb{F}^{m \times n}$  por

$$f(A) = \max_{\|x\|_n=1} \|Ax\|_m = \max_{0 \neq x \in \mathbb{F}^n} \frac{\|Ax\|_m}{\|x\|_n}$$

es una norma que denotaremos con  $\|\cdot\|_{m,n}$ .

b) Si  $A \in \mathbb{F}^{m \times n}$  y  $B \in \mathbb{F}^{n \times p}$  entonces

$$\|AB\|_{m,p} \leq \|A\|_{m,n} \|B\|_{n,p}.$$

En particular, si  $m = n$  entonces la norma  $\|\cdot\|_{n,n}$  es una norma de matriz.

c) Las normas  $\|\cdot\|_{m,n}$ ,  $\|\cdot\|_m$  y  $\|\cdot\|_n$  son consistentes.

d) Si  $\mu$  es cualquier otra norma de matriz consistente con  $\|\cdot\|_m$  y  $\|\cdot\|_n$ , entonces  $\|A\|_{m,n} \leq \mu(A)$ ,  $\forall A \in \mathbb{F}^{m \times n}$ .

En lo sucesivo, y para evitar una notación excesivamente pesada, no escribiremos los subíndices que asocian cada norma al espacio en el que está definida y representaremos todas las normas con el símbolo  $\|\cdot\|$ . La razón es que en todo momento es claro en qué espacio está el vector o la matriz sobre el que actúa la norma. Así, escribiremos  $\|x\|$  y  $\|A\|$  en vez de  $\|x\|_n$  o  $\|A\|_{m,n}$  porque si  $A \in \mathbb{F}^{m \times n}$  entonces la norma en  $\mathbb{F}^{m \times n}$  es  $\|\cdot\|_{m,n}$  y no se necesita precisar los subíndices; y, además,  $Ax$  sólo tiene sentido si  $x \in \mathbb{F}^n$  y no se precisa especificar el subíndice en el correspondiente símbolo de la norma.

**Demostración.**- a) Debemos verificar todos los axiomas que definen las normas de matriz

(i) Si  $A \neq 0$  existe  $x \in \mathbb{F}^n$  tal que  $Ax \neq 0$ . Por lo tanto  $\|Ax\| > 0$  y  $\|A\| = \max_{\|x\|=1} \|Ax\| > 0$ .

(ii) Para todo  $\lambda \in \mathbb{F}$

$$\begin{aligned} \|\lambda A\| &= \max_{\|x\|=1} \|\lambda Ax\| = \max_{\|x\|=1} |\lambda| \|Ax\| = \\ &= |\lambda| \max_{\|x\|=1} \|Ax\| = |\lambda| \|A\|. \end{aligned}$$

(iii) Probamos ahora que si  $A, B \in \mathbb{F}^{m \times n}$  entonces  $\|A + B\| \leq \|A\| + \|B\|$ . En efecto

$$\begin{aligned} \|A + B\| &= \max_{\|x\|=1} \|(A + B)x\| = \max_{\|x\|=1} \|Ax + Bx\| \leq \\ &\leq \max_{\|x\|=1} (\|Ax\| + \|Bx\|) \leq \max_{\|x\|=1} \|Ax\| + \max_{\|x\|=1} \|Bx\| = \\ &= \|A\| + \|B\|. \end{aligned}$$

b) Sean  $A \in \mathbb{F}^{m \times n}$  y  $B \in \mathbb{F}^{n \times p}$ . Por definición

$$\|AB\| = \max_{\|x\|=1} \|ABx\|.$$

Sea  $x_0 \in \mathbb{F}^p$  un vector con  $\|x_0\| = 1$  donde se alcanza el máximo:  $\|AB\| = \|ABx_0\|$ . Ahora bien, para cada matriz  $X$

$$\|X\| = \max_{y \neq 0} \frac{\|Xy\|}{\|y\|}$$

de donde resulta que  $\|Xy\| \leq \|X\| \|y\|$ ,  $\forall y \neq 0$ . Así

$$\begin{aligned} \|AB\| &= \|A(Bx_0)\| \leq \|A\| \|Bx_0\| \leq \\ &\leq \|A\| \|B\| \|x_0\| = \|A\| \|B\|, \end{aligned}$$

porque  $\|x_0\| = 1$ ; que es lo que se quería demostrar.

En particular si  $m = n = p$  y  $A, B \in \mathbb{F}^{n \times n}$  entonces

$$\|AB\|_{n,n} \leq \|A\|_{n,n} \|B\|_{n,n}.$$

Es decir,  $\|\cdot\|_{n,n}$  es una norma de matriz.

c) Ya hemos visto que  $\forall A \in \mathbb{F}^{n \times n}$  y  $\forall x \in \mathbb{F}^n$ ,  $x \neq 0$ , se tiene que

$$\|Ax\|_m \leq \|A\|_{m,n} \|x\|_n.$$

Y para  $x = 0$  se obtiene la igualdad de forma evidente. Por consiguiente, las normas  $\|\cdot\|_{m,n}$ ,  $\|\cdot\|_m$  y  $\|\cdot\|_n$  son consistentes.

d) También hemos visto más arriba que existe  $x_0 \in \mathbb{F}^n$  con  $\|x_0\| = 1$  tal que

$$\|A\| = \|Ax_0\|.$$

Como  $\mu$  es consistente con  $\|\cdot\|_m$  y  $\|\cdot\|_n$  resulta que

$$\|A\| = \|Ax_0\|_m \leq \mu(A) \|x_0\|_n = \mu(A),$$

tal y como se deseaba demostrar. ■

**Ejemplo 2.12** .- Vamos a ver cómo se definen explícitamente las normas de matriz  $\ell_p$  para  $p = 1, 2, \infty$ .

1. - La norma de matriz  $\ell_1$ : Sea  $A \in \mathbb{F}^{m \times n}$  y escribamos esta matriz en función de sus columnas  $A = [a_1 \ a_2 \ \cdots \ a_n]$ . Vamos a probar que la norma de matriz inducida por la norma  $\ell_1$  es :

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

o, equivalentemente,

$$\|A\|_1 = \max_{1 \leq j \leq n} \|a_j\|_1$$

En efecto, pongamos  $M = \max_{1 \leq j \leq n} \|a_j\|_1$ . Vamos a demostrar que  $M = \max_{\|x\|_1=1} \|Ax\|_1$ .

Para ello probamos primero que  $M \geq \|Ax\|_1$  para todo  $x \in \mathbb{F}^n$  con  $\|x\|_1 = 1$ ; i.e. que  $M$  es una cota superior del conjunto

$$\{\|Ax\|_1 : \|x\|_1 = 1, x \in \mathbb{F}^n\},$$

Y a continuación que esta cota se alcanza. Es decir, que existe  $x_0 \in \mathbb{F}^n$  tal que  $\|x_0\|_1 = 1$  y  $M = \|Ax_0\|_1$ .

Sea  $x \in \mathbb{F}^n$  con  $\|x\|_1 = 1$ , y sean  $(x_1, \dots, x_n)$  sus componentes. Entonces

$$\begin{aligned} \|Ax\|_1 &= \|x_1 a_1 + \cdots + x_n a_n\|_1 \leq \\ &\leq (|x_1| \|a_1\|_1 + \cdots + |x_n| \|a_n\|_1) \leq \\ &\leq (|x_1| + \cdots + |x_n|) \max_{1 \leq j \leq n} \|a_j\|_1 = M \|x\|_1 = M \end{aligned}$$

Por lo tanto,  $M$  es una cota superior de  $\{\|Ax\|_1 : \|x\|_1 = 1, x \in \mathbb{F}^n\}$ . Ahora, si el  $\max_{1 \leq j \leq n} \|a_j\|_1$  se alcanza, digamos, para la columna  $k$  entonces  $M = \|a_k\|_1$ .

$$\begin{array}{c} k \\ \downarrow \end{array}$$

Tomando  $x_0 = e_k = (0, \dots, 0, 1, 0, \dots, 0) \in \mathbb{F}^n$ , tenemos que  $\|x_0\|_1 = 1$  y

$$M = \|a_k\|_1 = \|Ae_k\|_1 = \|Ax_0\|_1,$$

de modo que  $M$  es una cota superior de  $\{\|Ax\|_1 : \|x\|_1 = 1, x \in \mathbb{F}^n\}$  que se alcanza. Esto es

$$\max_{1 \leq j \leq n} \|a_j\|_1 = \max_{\|x\|_1=1} \|Ax\|_1 = \|A\|_1$$

2. - La norma de matriz  $\ell_\infty$ : De forma similar a lo que hemos visto más arriba, vamos a demostrar que la norma de matriz inducida por la norma  $\ell_\infty$  es

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

Pongamos  $M = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$ . Por una parte

$$\begin{aligned} \|Ax\|_\infty &= \max_{1 \leq i \leq n} \left| \sum_{j=1}^n a_{ij} x_j \right| \leq \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij} x_j| \leq \\ &\leq \max_{1 \leq i \leq n} \sum_{j=1}^n \left[ |a_{ij}| \max_{1 \leq k \leq n} |x_k| \right] = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| \|x\|_\infty = \\ &= M \|x\|_\infty. \end{aligned}$$

Así pues

$$\|A\|_\infty = \max_{\|x\|_\infty=1} \|Ax\|_\infty \leq M. \quad (2.2)$$

Para demostrar la desigualdad en sentido contrario vamos a probar que para cada  $i = 1, \dots, n$  existe un vector  $x_i$  con  $\|x_i\|_\infty = 1$  tal que

$$\sum_{j=1}^n |a_{ij}| = \left| \sum_{j=1}^n a_{ij} x_{ij} \right|, \quad (2.3)$$

donde  $x_{ij}$  es la  $j$ -ésima componente de  $x_i$ .

Supuesto esto demostrado es fácil ver que  $\|A\|_\infty \geq M$ . En efecto, para cada  $i = 1, \dots, n$

$$\begin{aligned}\|A\|_\infty &= \max_{\|x\|_\infty=1} \|Ax\|_\infty \geq \|Ax_i\| = \max_{1 \leq j \leq n} \left| \sum_{j=1}^n a_{ij} x_{ij} \right| = \\ &= \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| = M\end{aligned}$$

Así pues, debemos demostrar (2.3). Basta encontrar  $x_i$  tal que  $a_{ij} x_{ij} = |a_{ij}|$ . Pero esta misma identidad nos sirve de definición:

$$x_{ij} = \frac{\bar{a}_{ij}}{|a_{ij}|}$$

siempre que  $a_{ij} \neq 0$  y  $x_{ij} = 1$  si  $a_{ij} = 0$ . Con esta definición de  $x_i$  tenemos que  $\|x_i\|_\infty = 1$  y se verifica (2.3) porque cada sumando  $a_{ij} x_{ij} = |a_{ij}|$  es positivo.

En conclusión

$$\|A\|_\infty = \max_{\|x\|_\infty=1} \|Ax\|_\infty = M = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$$

tal y como se deseaba demostrar.

3. - La norma de matriz  $\ell_2$ : A la norma de matriz inducida por la norma  $\ell_2$  se le llama también **Norma Espectral**. Veremos en un tema posterior que tiene un significado muy importante que todavía no estamos en condiciones de comprender bien.

## 2.4. Sucesiones y Series de Matrices

Tal y como hemos dicho en la sección anterior, el uso de normas nos permite hablar de convergencia de sucesiones de vectores y, por lo tanto, de matrices. El objetivo de esta sección es introducir las serie de matrices y demostrar un resultado que necesitaremos en un tema posterior.

Sea  $\{A_k\}_{k=0}^\infty$  una sucesión infinita de matrices con elementos en  $\mathbb{F}^{n \times m}$ , recordemos que  $\mathbb{F} = \mathbb{R}$  o  $\mathbb{C}$ . Con esta sucesión formamos otra, la de las sumas parciales

$$S_k = \sum_{j=0}^k A_j, \quad k \geq 0.$$

**Definición 2.13** La serie  $\sum_{j=0}^{\infty} A_j$  converge si la sucesión  $\{S_k\}$  es convergente; i.e. si existe el  $\lim_{k \rightarrow \infty} S_k$ . Además, si  $S = \lim_{k \rightarrow \infty} S_k$  entonces escribiremos

$$\sum_{j=0}^{\infty} A_j = S.$$

La convergencia de la serie  $\sum_{j=0}^{\infty} A_j$  se puede reducir a la de la serie numérica  $\sum_{j=0}^{\infty} \|A_j\|$  de la siguiente forma:

**Proposición 2.14** Si la serie numérica  $\sum_{j=0}^{\infty} \|A_j\|$  converge para alguna norma de matriz, también converge la serie matricial  $\sum_{j=0}^{\infty} A_j$ .

**Demostración.-** Sea  $\varepsilon > 0$  un número real dado. Veremos que existe un entero  $N > 0$  tal que si  $p, q \geq N$  entonces  $\|S_p - S_q\| < \varepsilon$ . Esto demuestra que la sucesión  $\{S_n\}$  de sumas parciales es una sucesión de Cauchy. Como  $\mathbb{F}^{m \times n}$  es completo,  $\{S_n\}$  converge.

En efecto

$$\|S_p - S_q\| = \left\| \sum_{j=p+1}^q A_j \right\| \leq \sum_{j=p+1}^q \|A_j\| = \left| \sum_{j=0}^p \|A_j\| - \sum_{j=0}^q \|A_j\| \right|.$$

Ahora bien, si  $\sum_{j=0}^{\infty} \|A_j\|$  converge, existe  $N > 0$  tal que si  $p, q \geq N$  entonces

$$\left| \sum_{j=0}^p \|A_j\| - \sum_{j=0}^q \|A_j\| \right| < \varepsilon,$$

que es lo que se quería demostrar. ■

En particular, para matrices cuadradas la serie de potencias  $\sum_{j=0}^{\infty} a_j A^j$  es convergente si lo es la serie numérica  $\sum_{j=0}^{\infty} |a_j| \|A\|^j$ . En realidad, este resultado no es un caso particular de la Proposición 2.14 pero se demuestra igual teniendo en cuenta que la norma elegida es una norma de matriz. En efecto, basta observar que si  $B_j = a_j A^j$  entonces

$$\|B_j\| \leq |a_j| \|A^j\| \leq |a_j| \|A\|^j,$$

donde la última desigualdad se debe a que las normas de matriz tienen la propiedad submultiplicativa.

Como muchas funciones escalares se pueden definir como series de potencias convergentes en ciertas regiones del plano complejo, la posibilidad de reducir la convergencia de una serie de potencias matriciales a la de sus correspondientes normas permite definir de forma sencilla algunas funciones matriciales.

Concretamente, sea  $f(z) = \sum_{j=0}^{\infty} a_j z^j$  en un entorno de  $z = 0$  con un radio de convergencia  $R$  y consideremos la serie matricial  $\sum_{j=0}^{\infty} a_j A^j$  con  $A \in \mathbb{F}^{n \times n}$ . Si hay alguna norma de matriz para la que  $\|A\| < R$ , la serie  $\sum_{j=0}^{\infty} |a_j| \|A\|^j$  converge y podemos definir

$$f(A) = \sum_{j=0}^{\infty} a_j A^j.$$

En particular, cada una de las siguientes funciones está perfectamente definidas:

$$\begin{aligned} e^A &= \sum_{j=0}^{\infty} \frac{1}{j!} A^j, \quad \forall A \in \mathbb{F}^{n \times n}. \\ \text{sen}(A) &= \sum_{j=0}^{\infty} \frac{(-1)^j}{(2j+1)!} A^{2j+1}, \quad \forall A \in \mathbb{F}^{n \times n}. \\ \text{cos}(A) &= \sum_{j=0}^{\infty} \frac{(-1)^j}{(2j)!} A^{2j}, \quad \forall A \in \mathbb{F}^{n \times n}. \end{aligned}$$

De la misma forma se podrían definir otras funciones matriciales como  $\ln A$ ,  $\tan(A)$ , etc. Nos interesa especialmente la serie geométrica  $\sum_{j=0}^{\infty} z^j$  que sabemos que converge

a  $\frac{1}{1-z} = (1-z)^{-1}$  si  $|z| < 1$ :

**Proposición 2.15** Si  $A \in \mathbb{F}^{n \times n}$  y  $\|A\| < 1$  para alguna norma de matriz, entonces  $I_n - A$  es invertible. Además, en tal caso  $(I_n - A)^{-1} = \sum_{j=0}^{\infty} A^j$  y  $\|(I_n - A)^{-1}\| \leq \frac{1}{1 - \|A\|}$ .

**Demostración.-** Si  $\|A\| < 1$  entonces la serie geométrica  $\sum_{j=0}^{\infty} \|A\|^j$  es convergente, y por la Proposición 2.14 la serie matricial  $\sum_{j=0}^{\infty} A^j$  converge. Sea  $B = \sum_{j=0}^{\infty} A^j$  y pongamos

$$S_k = \sum_{j=0}^k A^j.$$

Entonces

$$(I_n - A)S_k = (I_n - A)(I_n + A + \cdots + A^k) = I_n - A^{k+1}.$$

Como  $\lim_{k \rightarrow \infty} S_k = B$  resulta que

$$\lim_{k \rightarrow \infty} (I_n - A)S_k = (I_n - A)B,$$

y también

$$\lim_{k \rightarrow \infty} (I_n - A)S_k = \lim_{k \rightarrow \infty} (I_n - A^{k+1}).$$

Ahora bien,  $\lim_{k \rightarrow \infty} (I_n - A^{k+1}) = I_n$  porque  $\|I_n - A^{k+1} - I_n\| \leq \|A\|^{k+1}$  y la sucesión numérica  $\{\|A\|^k\}$  converge a cero por ser  $\|A\| < 1$ .

En consecuencia  $(I_n - A)B = I_n$  y

$$(I_n - A)^{-1} = B = \sum_{j=0}^{\infty} A^j.$$

Finalmente, como  $B = \sum_{j=0}^{\infty} A^j = \lim_{k \rightarrow \infty} \sum_{j=0}^k A^j$  y  $\|\cdot\|$  es continua,

$$\left\| \sum_{j=0}^{\infty} A^j \right\| = \|B\| = \lim_{k \rightarrow \infty} \left\| \sum_{j=0}^k A^j \right\| \leq \lim_{k \rightarrow \infty} \sum_{j=0}^k \|A^j\| \leq \lim_{k \rightarrow \infty} \sum_{j=0}^k \|A\|^j = \sum_{j=0}^{\infty} \|A\|^j,$$



donde hemos usado la desigualdad triangular de las normas y que  $\|\cdot\|$  una norma de matriz. Así pues

$$\|(I_n - A)^{-1}\| = \left\| \sum_{j=0}^{\infty} A^j \right\| \leq \sum_{j=0}^{\infty} \|A\|^j = \frac{1}{1 - \|A\|}. \blacksquare$$

