Norma de Frobenius

Estos apuntes están escritos por Darío Coutiño Aquino y Egor Maximenko.

Objetivos. Dada una matriz $A \in \mathcal{M}_{m\times n}(\mathbb{C})$, su *norma de Frobenius* (llamada también la norma de Hilbert–Schmidt) se define como

$$\|A\|_F \coloneqq \big(\operatorname{tr}(A^*A)\big)^{1/2}$$

o, de manera equivalente, como

$$\|A\|_F = \left(\sum_{j=1}^{m}\sum_{k=1}^{n}|A_{j,k}|^2\right)^{1/2},$$

donde tr es la función traza. Estudiaremos varias representaciones de esta función y sus propiedades principales.

Requisitos. Normas en $\mathbb{C}^n$ y $\mathbb{R}^n$, normas matriciales inducidas por normas vectoriales, notación para renglones y columnas de una matriz, desigualdad de Cauchy, definición formal del producto de matrices, propiedades de la multiplicación de matrices, producto interno en un espacio vectorial complejo, propiedades de la traza de matrices.

El producto interno canónico en el espacio de matrices

1 Proposición (el producto interno canónico en el espacio de matrices)**.** *La función* $\mathcal{M}_{m\times n}(\mathbb{C}) \times \mathcal{M}_{m\times n}(\mathbb{C}) \to \mathbb{C}$ *definida como*

$$\langle A, B\rangle \coloneqq \operatorname{tr}(A^*B),$$

es un producto interno en $\mathcal{M}_{m\times n}(\mathbb{C})$*. En otras palabras, esta función es lineal con respecto al segundo argumento, hermítica y positiva definida.*

Demostración. En los siguientes cálculos suponemos que $A, B, C \in \mathcal{M}_{m\times n}(\mathbb{C})$ son matrices arbitrarias y $\lambda \in \mathbb{C}$ es un complejo arbitrario.

I. Verificamos que la función $\langle\cdot,\cdot\rangle$ es lineal con respecto al segundo argumento:

$$\begin{aligned}\langle A, B + C\rangle &= \operatorname{tr}(A^*(B + C)) = \operatorname{tr}(A^*B + A^*C)\\ &= \operatorname{tr}(A^*B) + \operatorname{tr}(A^*C) = \langle A, B\rangle + \langle A, C\rangle,\end{aligned}$$

$$\begin{aligned}\langle A, \lambda B\rangle &= \operatorname{tr}(A^*(\lambda B)) = \operatorname{tr}(\lambda A^*B)\\ &= \lambda \operatorname{tr}(A^*B) = \lambda\langle A, B\rangle.\end{aligned}$$

II. La función $\langle\cdot,\cdot\rangle$ es hermítica:

$$\overline{\langle B, A\rangle} = \overline{\operatorname{tr}(B^*A)} = \operatorname{tr}((B^*A)^*) = \operatorname{tr}(A^*B) = \langle A, B\rangle.$$

III. Mostremos que $\langle\cdot,\cdot\rangle$ es definida positiva. Sea $A \in \mathcal{M}_{m\times n}(\mathbb{C})$, $A \neq 0_{m\times n}$. Entonces

$$\langle A, A\rangle = \operatorname{tr}(A^*A) = \sum_{k=1}^{n}(A^*A)_{k,k} = \sum_{k=1}^{n}\left(\sum_{j=1}^{m}(A^*)_{k,j}(A)_{j,k}\right) = \sum_{k=1}^{n}\sum_{j=1}^{m}|A_{j,k}|^2. \quad (1)$$

Los sumandos de la suma doble en el lado derecho de (1) son no negativos. Además, existe un par de índices $p \in \{1,\ldots,m\}$, $q \in \{1,\ldots,n\}$ tal que $A_{p,q} \neq 0$ y por lo tanto $|A_{p,q}|^2 > 0$, lo cual demuestra que en el lado derecho de (1) existe al menos un sumando no nulo, y así $\langle A^*, A\rangle > 0$. □

2 Observación. En álgebra lineal numérica es cómodo el convenio que el producto interno es homogéneo con respecto al *segundo argumento*. Cambiando $\operatorname{tr}(A^*B)$ por $\operatorname{tr}(B^*A)$ obtenemos una función homogénea con respecto al primer argumento.

3 Notación (“vectorización” de una matriz)**.** Dada una matriz $A \in \mathcal{M}_{m\times n}(\mathbb{C})$, denotemos por $\operatorname{vec}(A)$ al vector obtenido de la matriz A de la siguiente manera:

$$\operatorname{vec}(A) = \left[A_{(s-1)\div n+1,\, s \bmod n}\right]_{s=1}^{mn},$$

donde los signos $\div$ y mod denotan las operaciones de la división entera y del residuo, respectivamente. En otras palabras,

$$(\operatorname{vec}(A))_{(j-1)n+k} = A_{j,k} \qquad (j \in \{1,\ldots,m\},\ k \in \{1,\ldots,n\}).$$

Por ejemplo, si $A \in \mathcal{M}_{3\times 2}(\mathbb{C})$, entonces

$$A = \begin{bmatrix} A_{1,1} & A_{1,2} & A_{1,3} \\ A_{2,1} & A_{2,2} & A_{2,3} \end{bmatrix} \qquad \text{y} \qquad \operatorname{vec}(A) = \begin{bmatrix} A_{1,1} \\ A_{1,2} \\ A_{1,3} \\ A_{2,1} \\ A_{2,2} \\ A_{2,3} \end{bmatrix}.$$

4 Proposición. *La función* $\operatorname{vec}\colon \mathcal{M}_{m\times n}(\mathbb{C}) \to \mathbb{C}^{mn}$ *es un isomorfismo de espacios vectoriales complejos.*

Idea de demostración. Las propiedades aditiva y homogénea se verifican fácilmente, y la función inversa actúa por la regla

$$v \mapsto \left[v_{(j-1)n+k}\right]_{j,k=1}^{m,n}.$$

□

5 Proposición. *Sean* $A, B \in \mathcal{M}_{m\times n}(\mathbb{C})$*. Entonces*

$$\langle A, B\rangle = \operatorname{vec}(A)^* \operatorname{vec}(B). \quad (2)$$

Idea de demostración. Ambos lados de la fórmula (2) son iguales a

$$\sum_{j=1}^{m}\sum_{k=1}^{n}\overline{A_{j,k}}B_{j,k}.$$ □

6 Observación. Con ayuda de la fórmula (2) podríamos demostrar la Proposición 1 de otra manera, usando la Proposición 4 y el hecho que el producto punto en $\mathbb{C}^n$ es un producto interno.

7 Ejercicio. El producto interno canónico de dos matrices $A, B \in \mathcal{M}_{m\times n}(\mathbb{C})$ se preserva al multiplicar A y B por la misma matriz unitaria del lado izquierdo o del lado derecho. En otras palabras, si $U \in \mathcal{M}_m(\mathbb{C})$ tal que $U^*U = I_m$ y $V \in \mathcal{M}_n(\mathbb{C})$ tal que $V^*V = I_n$, entonces

$$\langle UA, UB\rangle = \langle A, B\rangle, \qquad \langle AV, BV\rangle = \langle A, B\rangle.$$

Varias definiciones equivalentes de la norma de Frobenius

8 Definición (definición de la norma de Frobenius a través del producto interno canónico de matrices)**.** Sea $A \in \mathcal{M}_{m\times n}(\mathbb{C})$. Entonces la *norma de Frobenius* de la matriz A se define como

$$\|A\|_F \coloneqq \sqrt{\langle A, A\rangle}, \tag{3}$$

esto es,

$$\|A\|_F \coloneqq \sqrt{\operatorname{tr}(A^*A)}. \tag{4}$$

9 Observación. Se sabe que la traza del producto de matrices no depende del orden de los factores, por eso (4) se pueden escribir también como

$$\|A\|_F = \sqrt{\operatorname{tr}(AA^*)}. \tag{5}$$

10 Proposición (expresión de la norma de Frobenius de una matriz en términos de sus entradas)**.** *Sea $A \in \mathcal{M}_{m\times n}(\mathbb{C})$. Entonces*

$$\|A\|_F = \left(\sum_{j=1}^{m}\sum_{k=1}^{n}|A_{j,k}|^2\right)^{1/2}. \tag{6}$$

Además

$$\|A\|_F = \|\operatorname{vec}(A)\|_2. \tag{7}$$

Demostración. Se obtiene de la Proposición 5 poniendo $B = A$. □

11 Proposición (expresión de la norma de Frobenius en términos de las normas euclidianas de los renglones o columnas)**.** *Sea* $A \in \mathcal{M}_{m\times n}(\mathbb{C})$*. Entonces*

$$\|A\|_F = \left(\sum_{j=1}^{m} \|A_{j,*}\|_2^2 \right)^{1/2}, \tag{8}$$

$$\|A\|_F = \left(\sum_{k=1}^{n} \|A_{*,k}\|_2^2 \right)^{1/2}. \tag{9}$$

Demostración. La fórmula (8) se obtiene de (6) al notar que

$$\sum_{k=1}^{n} |A_{j,k}|^2 = \|A_{j,*}\|_2^2.$$

Para deducir (9) de (6), hay que intercambiar el orden de las sumas y notar que

$$\sum_{j=1}^{m} |A_{j,k}|^2 = \|A_{*,k}\|_2^2. \qquad \square$$

12 Ejemplo. Consideremos una matriz real 2×3:

$$A = \begin{bmatrix} 2 & 1 & -4 \\ 5 & 0 & -1 \end{bmatrix}.$$

Apliquemos la fórmula (4):

$$A^*A = \begin{bmatrix} 29 & 2 & -13 \\ 2 & 1 & -4 \\ -13 & -4 & 17 \end{bmatrix}, \qquad \operatorname{tr}(A^*A) = 29 + 1 + 17 = 47, \qquad \|A\|_F = \sqrt{47} \approx 6.86.$$

Por la fórmula (6),

$$\|A\|_F = (4 + 1 + 16 + 25 + 0 + 1)^{1/2} = \sqrt{47}.$$

Por la fórmula (8),

$$\|A\|_F = \left(\|A_{1,*}\|_2^2 + \|A_{2,*}\|_2^2\right)^{1/2} = (21 + 26)^{1/2} = \sqrt{47}.$$

Por la fórmula (9),

$$\|A\|_F = \left(\|A_{*,1}\|_2^2 + \|A_{*,2}\|_2^2 + \|A_{*,3}\|_2^2\right)^{1/2} = (29 + 1 + 17)^{1/2} = \sqrt{47}.$$

13 Ejercicio. Encontrar comandos que calculan la norma de Frobenius de la matriz dada en varios sistemas de álgebra computacional, por ejemplo, en Sagemath, GNU Octave, R, Wolfram Mathematica, MATLAB.

14 Ejercicio. En algún lenguaje de programación escribir una función que tenga un argumento matricial A y que calcule y devuelva la norma de Frobenius de la matriz dada A. Se puede usar el siguiente esquema con dos ciclos encajados:

```
function FrobeniusNorm(A):
   s <- 0;
   m <- number of rows of A;
   n <- number of columns of A;
   for j <- 1, ..., m:
      for k <- 1, ..., n:
         s <- s + abs(A[j, k]) ^ 2;
   return sqrt(s);
```

15 Ejemplo (la norma de Frobenius de un vector)**.** Sea $v = \left[v_j\right]_{j=1}^n \in \mathbb{C}^n$. Identificamos el vector v con la matriz-columna $W \in \mathcal{M}_{n\times 1}(\mathbb{C})$, poniendo

$$W_{j,1} = v_j \qquad (j \in \{1, \ldots, n\}).$$

Entonces la norma de Frobenius de la matriz W coincide con la norma euclidiana del vector v:

$$\|W\|_F = \left(\sum_{j=1}^n |v_j|^2\right)^{1/2} = \|v\|_2.$$

16 Proposición. *La norma de Frobenius en efecto es una norma en el espacio vectorial complejo $\mathcal{M}_{m\times n}(\mathbb{C})$.*

Demostración. Es un hecho general que si $\langle\cdot,\cdot\rangle$ es un producto interno, entonces la función definida mediante la fórmula (3) es una norma. Se dice que esta norma es *inducida* por el producto interno $\langle\cdot,\cdot\rangle$. Para los estudiantes que olvidaron el concepto de *norma inducida por un producto interno* recordemos el esquema de demostración.

La propiedad subaditiva se obtiene de la desigualdad de Schwarz para este producto interno:

$$\begin{aligned}\|A+B\|_F^2 &= \langle A+B, A+B\rangle = \langle A, A\rangle + \langle A, B\rangle + \langle B, A\rangle + \langle B, B\rangle \\ &= \|A\|_F^2 + 2\operatorname{Re}(\langle A, B\rangle) + \|B\|_F^2 \le \|A\|_F^2 + 2|\langle A, B\rangle| + \|B\|_F^2 \\ &\le \|A\|_F^2 + 2\|A\|_F\,\|B\|_F + \|B\|_F^2 = (\|A\|_F + \|B\|_F)^2.\end{aligned}$$

La propiedad homogénea absoluta se verifica directamente:

$$\|\lambda A\|_F = \sqrt{\langle\lambda A, \lambda A\rangle} = \sqrt{|\lambda|^2\,\langle A, A\rangle} = |\lambda|\,\|A\|_F.$$

Finalmente, la norma de Frobenius de cualquier matrix no nula es estrictamente positiva. Es una de las propiedades del producto interno que fue verificada en la demostración de la Proposición 1. $\square$

La norma de Frobenius y la multiplicación de matrices

17 Lema (desigualdad de Cauchy para las entradas del producto de matrices)**.** *Sean* $A \in \mathcal{M}_{m\times n}(\mathbb{C})$, $B \in \mathcal{M}_{n\times p}(\mathbb{C})$, $j \in \{1,\ldots,m\}$, $k \in \{1,\ldots,p\}$. *Entonces*

$$|(AB)_{j,k}|^2 \le \|A_{j,*}\|_2^2 \, \|B_{*,k}\|_2^2. \tag{10}$$

Demostración. Escribimos $(AB)_{j,k}$ por la definición del producto de matrices:

$$(AB)_{j,k} = \sum_{s=1}^{n} A_{j,s} B_{s,k},$$

aplicamos las propiedades subaditiva y multiplicativa del valor absoluto y la desigualdad de Cauchy:

$$|(AB)_{j,k}| \le \sum_{s=1}^{n} |A_{j,s}| \, |B_{s,k}| \le \left(\sum_{s=1}^{n} |A_{j,s}|^2\right)^{1/2} \left(\sum_{s=1}^{n} |B_{s,k}|^2\right)^{1/2} = \|A_{j,*}\|_2 \|B_{*,k}\|_2. \qquad \square$$

18 Teorema (propiedad submultiplicativa de la norma de Frobenius)**.** *Sea* $A \in \mathcal{M}_{m\times n}(\mathbb{C})$ *y sea* $B \in \mathcal{M}_{n\times p}(\mathbb{C})$. *Entonces*

$$\|AB\|_F \le \|A\|_F \|B\|_F. \tag{11}$$

Demostración. Aplicamos la fórmula (6) y el Lema 17:

$$\|AB\|_F^2 \le \sum_{j=1}^{m}\sum_{k=1}^{n} |(AB)_{j,k}|^2 \le \sum_{j=1}^{m}\sum_{k=1}^{n} \|A_{j,*}\|_2^2 \, \|B_{*,k}\|_2^2 = \left(\sum_{j=1}^{m} \|A_{j,*}\|_2^2\right)\left(\sum_{k=1}^{n} \|B_{*,k}\|_2^2\right).$$

Recordando las fórmulas (8) y (9) llegamos al producto $\|A\|_F^2 \|B\|_F^2$. $\square$

19 Corolario. *Sean* $A \in \mathcal{M}_{m\times n}(\mathbb{C})$, $v \in \mathbb{R}^n$. *Entonces*

$$\|Av\|_2 \le \|A\|_F \|v\|_2.$$

Demostración. Este resultado se obtiene del Teorema 18 y Ejemplo 15. $\square$

Recordemos que la norma matricial en $\mathcal{M}_{m\times n}(\mathbb{C})$ asociada a la norma euclidiana $\|\cdot\|_2$ en $\mathbb{C}^n$ se define como

$$\|A\|_{\mathrm{matr},2} = \sup_{v \in \mathbb{C}^n \setminus \{\mathbf{0}_n\}} \frac{\|Av\|_2}{\|v\|_2}. \tag{12}$$

20 Proposición (la norma de Frobenius es una cota superior para la norma matricial asociada a la norma euclidiana de vectores)**.** *Sea* $A \in \mathcal{M}_{m\times n}(\mathbb{C})$. *Entonces*

$$\|A\|_{\mathrm{matr},2} \le \|A\|_F.$$

Demostración. Se sigue del Corolario 19 y de la fórmula (12) que define $\|A\|_{\mathrm{matr},2}$. □

21 Ejemplo (las normas de una matriz diagonal)**.** Sea $d \in \mathbb{C}^n$. Consideremos la matriz diagonal con entradas diagonales $d_1, \ldots, d_n$:

$$\operatorname{diag}(d) = \left[d_j \delta_{j,k}\right]_{j,k=1}^{n}.$$

Aplicamos las propiedades de la delta de Kronecker:

$$\|\operatorname{diag}(d)\|_F^2 = \sum_{j=1}^{n} \sum_{k=1}^{n} |d_j|^2 \delta_{j,k}^2 = \sum_{j=1}^{n} \sum_{k=1}^{n} |d_j|^2 \delta_{j,k} = \sum_{j=1}^{n} |d_j|^2 = \|d\|_2,$$

para concluir que

$$\|\operatorname{diag}(d)\|_F = \|d\|_2.$$

Por otro lado, es posible demostrar que

$$\|\operatorname{diag}(d)\|_{\mathrm{matr},2} = \|d\|_\infty.$$

22 Ejercicio (las normas de la matriz identidad)**.** Calcular $\|I_n\|_{\mathrm{matr},2}$ y $\|I_n\|_F$.

23 Ejercicio (las normas de una matriz básica)**.** Sean $p \in \{1, \ldots, m\}$, $q \in \{1, \ldots, n\}$. Denotemos por $E_{p,q}$ a la matriz de tamaño $m \times n$ cuya entrada (p, q) es igual a 1, y las demás entradas son cero:

$$E_{p,q} = \left[\delta_{p,j} \delta_{q,k}\right]_{j,k=1}^{m,n}.$$

Calcular $\|E_{p,q}\|_{\mathrm{matr},2}$ y $\|E_{p,q}\|_F$.

24 Proposición (la norma de Frobenius de la matriz transpuesta y de la matriz transpuesta conjugada)**.** *Sea* $A \in \mathcal{M}_{m\times n}(\mathbb{C})$*. Entonces*

$$\|A^\top\|_F = \|A^*\|_F = \|A\|_F.$$

25 Proposición (propiedad submultiplicativa con la norma de Frobenius y la norma inducida por la norma euclidiana)**.** *Sean* $A \in \mathcal{M}_{m\times n}(\mathbb{C})$*,* $B \in \mathcal{M}_{n\times p}(\mathbb{C})$*. Entonces*

$$\|AB\|_F \le \|A\|_{\mathrm{matr},2} \|B\|_F, \tag{13}$$

$$\|AB\|_F \le \|A\|_F \|B\|_{\mathrm{matr},2}. \tag{14}$$

Demostración. Para demostrar (13), recordemos que cada columna del producto AB se puede escribir como el producto de la matriz A por la columna correspondiente de la matriz B:

$$(AB)_{*,k} = AB_{*,k},$$

luego aplicamos la fórmula (9) y la propiedad principal de $\|A\|_{\mathrm{matr},2}$:

$$\|AB\|_F = \sum_{k=1}^{n} \|AB_{*,k}\|_2^2 \le \sum_{k=1}^{n} \|A\|_{\mathrm{matr},2}^2 \|B_{*,k}\|_2^2 = \|A\|_{\mathrm{matr},2}^2 \sum_{k=1}^{n} \|B_{*,k}\|_2^2 = \|A\|_{\mathrm{matr},2}^2 \|B\|_F.$$

La fórmula (14) se puede deducir de (13) usando la Proposición 24:

$$\|AB\|_F = \|(AB)^*\|_F = \|B^* A^*\|_F \le \|B^*\|_{\mathrm{matr},2} \|A^*\|_F = \|B\|_{\mathrm{matr},2} \|A\|_F.$$

Hemos utilizado el hecho que $\|B^*\|_{\mathrm{matr},2} = \|B\|_{\mathrm{matr},2}$. □

26 Proposición (la norma de Frobenius es invariante bajo la multiplicación por matrices unitarias)**.** *Sea* $A \in \mathcal{M}_{m\times n}(\mathbb{C})$*.*

1. *Si* $U \in \mathcal{M}_m(\mathbb{C})$ *y* $U^*U = I_m$*, entonces*

$$\|UA\|_F = \|A\|_F. \tag{15}$$

2. *Si* $V \in \mathcal{M}_n(\mathbb{C})$ *y* $V^*V = I_n$*, entonces*

$$\|AV\|_F = \|A\|_F. \tag{16}$$

Demostración. Esta proposición se puede ver como un corolario del Ejercicio 7. Escribamos la demostración directa, utilizando (4), propiedades de la operación $A \mapsto A^*$ y propiedades de la función traza:

$$\begin{aligned} \|UA\|_F^2 &= \mathrm{tr}((UA)^*(UA)) = \mathrm{tr}(A^*U^*UA) = \mathrm{tr}(A^*A) = \|A\|_F^2, \\ \|AV\|_F^2 &= \mathrm{tr}((AV)^*(AV)) = \mathrm{tr}(V^*A^*AV) = \mathrm{tr}(V^*VA^*A) = \mathrm{tr}(A^*A) = \|A\|_F^2. \end{aligned}$$ □