Normas Matriciales

Alan Reyes-Figueroa

Métodos Numéricos II (Aula 01) 05.julio.2022

Introducción al Curso

Este curso es continuación de Métodos Numéricos I, esto quiere decir que estudiaremos métodos computacionales para hacer matemática numérica. El enfoque será mezcla entre curso teórico y práctico.

Estudiamos tres grandes temas:

1. Álgebra lineal numérica:
   - Sistemas lineales y factoración de matrices.
   - Cálculo de autovalores y autovectores.
2. Optimización numérica (continua):
   - Optimización no restricta:
     - métodos de gradiente, punto interior, CG, quase-Newton.
   - Optimización con restricciones:
     - programación lineal (LP).
3. Optimización discreta y combinatoria.

Introducción al Curso

Los métodos numéricos son un área que integra muchas ramas de la matemática. Por ejemplo, haremos uso extensivo de

- cálculo vectorial
- análisis real en $\mathbb{R}^n$
  - topología de espacios métricos,
  - propiedades de funciones continuas,
  - convergencia de secuencias y series, Taylor.
- álgebra lineal
- grafos y algoritmos
- programación

Normas de Vectores:

Definición

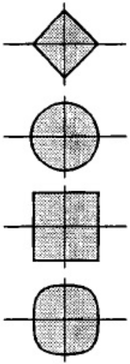
*Una **norma** (de vectores) en $\mathbb{R}^n$ es una función $||\cdot|| : \mathbb{R}^n \to \mathbb{R}$ que satisface las siguientes propiedades*

1. $||\mathbf{x}|| \geq 0$, $\forall \mathbf{x} \in \mathbb{R}^n$, *y* $||\mathbf{x}|| = 0$ *si y sólo si* $\mathbf{x} = \mathbf{0}$.
2. $||\mathbf{x} + \mathbf{y}|| \leq ||\mathbf{x}|| + ||\mathbf{y}||$, $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.
3. $||\alpha \mathbf{x}|| = |\alpha| \, ||\mathbf{x}||$, $\forall \alpha \in \mathbb{R},\ \mathbf{x} \in \mathbb{R}^n$.

Las normas más populares en $\mathbb{R}^n$ corresponden a la familia de normas-p ó p-**normas**. A continuación se muestran algunos ejemplos de p-normas, así como del disco unitario $\mathbb{D} = \{\mathbf{x} \in \mathbb{R}^2 : ||\mathbf{x}|| \leq 1\}$:

Conceptos de Álgebra Lineal

$$
\begin{aligned}
\|x\|_1 &= \sum_{i=1}^{m} |x_i|, \\
\|x\|_2 &= \left(\sum_{i=1}^{m} |x_i|^2 \right)^{1/2} = \sqrt{x^* x}, \\
\|x\|_\infty &= \max_{1 \le i \le m} |x_i|, \\
\|x\|_p &= \left(\sum_{i=1}^{m} |x_i|^p \right)^{1/p} \quad (1 \le p < \infty).
\end{aligned}
$$



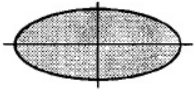
Aparte de las *p*-normas, otra de las más comunes es la familia de las *p*-**normas pesadas**, en donde en cada una de las coordenadas, las entradas se combinan mediante pesos.

Estos pesos vienen dados por una matriz diagonal $W \in \mathbb{R}^{n\times n}$:

$$||\mathbf{x}||_W = ||W\mathbf{x}||.$$

Por ejemplo, una 2-norma pesada en $\mathbb{R}^2$ se ve de la siguiente forma

$$\|x\|_W = \left(\sum_{i=1}^{m} |w_i x_i|^2\right)^{1/2}.$$



Normas de matrices inducidas por normas vectoriales:

Una matriz $A \in \mathbb{R}^{m\times n}$ puede verse como un vector en $\mathbb{R}^{mn}$ mediante el mapa

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \longrightarrow (a_{11}, \dots, a_{1n}, \dots, a_{m1}, \dots, a_{mn}) \in \mathbb{R}^{mn}.$$

Normas Matriciales

Tiene sentido entonces definir normas matriciales, para medir el "tamaño" de estas matrices.

Sin embargo, en los espacios de matrices, ciertas normas especiales son más útiles que las p-normas discutidas anteriormente. Estas son las normas inducidas, y se definen en términos del comportamiento de la matriz como operador $A : \mathbb{R}^n \to \mathbb{R}^m$.

Definición

Dadas normas $||\cdot||_{(n)}$ y $||\cdot||_{(m)}$ en el dominio y el rango de una matriz $A \in \mathbb{R}^{m\times n}$, decimos que $||\cdot||_{(m,n)}$ es la ***norma inducida*** *por $||\cdot||_{(n)}$ y $||\cdot||_{(m)}$ si $||\cdot||_{(m,n)}$ es el menor número C que satisface*

$$||A\mathbf{x}||_{(m)} \leq C||\mathbf{x}||_{(n)}, \;\; \forall \mathbf{x} \in \mathbb{R}^n. \tag{1}$$

UVG UNIVERSIDAD DEL VALLE DE GUATEMALA

Normas Matriciales

Si $\mathbf{x} \neq \mathbf{0}$, podemos pasar el término $||\mathbf{x}||_{(n)}$ dividiendo al lado izquierdo de (1), de modo que

$$\frac{||A\mathbf{x}||_{(m)}}{||\mathbf{x}||_{(n)}} \leq C, \ \forall \mathbf{x} \in \mathbb{R}^n, \ \mathbf{x} \neq \mathbf{0}.$$

Así,

$$||A||_{(m,n)} = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{||A\mathbf{x}||_{(m)}}{||\mathbf{x}||_{(n)}}, \tag{2}$$

de modo que $||A||_{(m,n)}$ es el mayor factor por el cual el operador A ampliado un vector. De la ecuación (3) en la definición de norma, la acción de A se determina por su acción sobre vectore sunitarios. Luego (2) puede reducirse a

$$||A||_{(m,n)} = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{||A\mathbf{x}||_{(m)}}{||\mathbf{x}||_{(n)}} = \sup_{||\mathbf{x}||_{(n)}=1} ||A\mathbf{x}||_{(m)}. \tag{3}$$

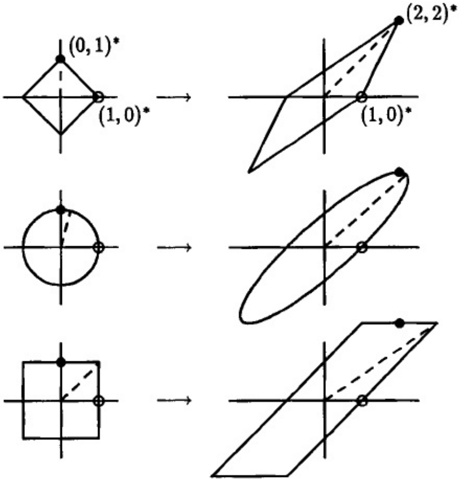
En el caso particular que $A \in \mathbb{R}^{m\times n}$, entonces la p**-norma matricial, inducida por la** p**-norma** en $\mathbb{R}^n$ es

$$||A||_p = \sup_{\mathbf{x}\neq\mathbf{0}} \frac{||A\mathbf{x}||_p}{||\mathbf{x}||_p} = \sup_{||\mathbf{x}||_p=1} ||A\mathbf{x}||_p. \tag{4}$$

Ejemplo: Considere la matriz $\begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix}$, $A : \mathbb{R}^2 \to \mathbb{R}^2$. (También la podemos ver como un mapa $A : \mathbb{C}^2 \to \mathbb{C}^2$, pero $\mathbb{R}$ es más conveniente para hacer ilustraciones).

La siguiente imagen muestra el resultado de algunas p-normas para A:

Normas Matriciales



$\|A\|_1 = 4$

$\|A\|_2 \approx 2.9208$

$\|A\|_\infty = 3$

1-norma de una matriz:
Sea $A \in \mathbb{R}^{m\times n}$. Escribimos A en sus columnas

$$A = \begin{pmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_n \end{pmatrix},$$

donde cada $\mathbf{a}_i \in \mathbb{R}^m$ es un vector. Considere la bola unitaria en la 1-norma, dada por $B_1(0) = \{\mathbf{x} \in \mathbb{R}^n : \sum_{j=1}^n |x_j| \leq 1\}$. Entonces, de la

ecuación (4), para todo $\mathbf{x} = (x_1, \dots, x_n)^T \in \mathbb{R}^n$, se tiene

$$||A\mathbf{x}||_1 = \left|\left|\sum_{i=1}^n x_i \mathbf{a}_i\right|\right|_1 \leq \sum_{i=1}^n |x_i|\ ||\mathbf{a}_i||_1 \leq \sum_{i=1}^n |x_i| \max_{1\leq i\leq n} ||\mathbf{a}_i||_1 \leq \max_{1\leq i\leq n} ||\mathbf{a}_i||_1,\ \forall ||\mathbf{x}||_1 \leq 1.$$

Esto muestra que $\sup_{||\mathbf{x}||_1=1} ||A\mathbf{x}||_1 \leq \max_{1\leq i\leq n} ||\mathbf{a}_i||_1$

Sea $j = \arg\max_{1\leq i\leq n} ||\mathbf{a}_i||_1$, esto es $||\mathbf{a}_j||_1 = \max_{1\leq i\leq n} ||\mathbf{a}_i||_1$. Tomando el vector $\mathbf{x} = \mathbf{e}_j$, obtenemos que

$$||A\mathbf{x}||_1 = ||A\mathbf{e}_j||_1 = ||\mathbf{a}_j||_1 = \max_{1\leq i\leq n} ||\mathbf{a}_i||_1,$$

y el supremo se alcanza en un vector unitario. Luego, $||A||_1 = \max\limits_{1\leq i\leq n} ||\mathbf{a}_i||_1$.

<u>∞-norma de una matriz</u>:

Si ahora denotamos A por sus filas, $A = \begin{pmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_m^T \end{pmatrix}$, $\mathbf{a}_i \in \mathbb{R}^n$. Un argumento muy similar al anterior se utiliza para mostrar que

$$||A||_\infty = \sup_{||\mathbf{x}||_\infty = 1} ||A\mathbf{x}||_\infty = \max_{1\leq i\leq m} ||\mathbf{a}_i^T||_1.$$

UNIVERSIDAD DEL VALLE DE GUATEMALA

Normas Matriciales

Calcular p-normas matriciales, con $p \neq 1, \infty$ es más difícil. Para ello, observemos que los productos internos acotan el producto de normas vectoriales.
Sean $1 \leq p, q \leq \infty$, tales que $\frac{1}{p} + \frac{1}{q} = 1$. Tenemos las siguientes desigualdades:

$$
\begin{aligned}
|\mathbf{x}^T\mathbf{y}| &\leq ||\mathbf{x}||_p||\mathbf{y}||_q \quad \text{(desigualdad de Hölder)} \\
|\mathbf{x}^T\mathbf{y}| &\leq ||\mathbf{x}||_2||\mathbf{y}||_2 \quad \text{(desigualdad de Cauchy-Schwarz)}
\end{aligned}
$$

Con esta información, vamos a calcular la

<u>2-norma de una matriz fila</u>:
Sea $A \in \mathbb{R}^{1\times n}$ una matriz de una sola fila. Podemos escribir $A = \mathbf{a}^T$, con $\mathbf{a}^T \in \mathbb{R}^n$. Entonces, de la desigualdad de Cauchy-Schwarz

Normas Matriciales

$$||A\mathbf{x}||_2 = ||\mathbf{a}^T\mathbf{x}||_2 \leq ||\mathbf{a}||_2||\mathbf{x}||_2,$$

luego $\displaystyle\sup_{||\mathbf{x}||_2=1} \frac{||A\mathbf{x}||_2}{||\mathbf{x}||_2} \leq ||\mathbf{a}||_2$.

Esta cota se alcanza. Haciendo $\mathbf{x} = \mathbf{a}$, se tiene que $||A\mathbf{a}||_2 = ||\mathbf{a}||_2^2$, y se tiene que $||A||_2 = ||\mathbf{a}||_2$.

<u>2-norma de un producto exterior:</u>

Sea $A \in \mathbb{R}^{m\times n}$ una matriz de rango 1, digamos $A = \mathbf{u}\mathbf{v}^T$, con $\mathbf{u} \in \mathbb{R}^m, \mathbf{v} \in \mathbb{R}^n$. Para todo $\mathbf{x} \in \mathbb{R}^n$,

$$||A\mathbf{x}||_2 = ||\mathbf{u}\mathbf{v}^T\mathbf{x}||_2 = ||\mathbf{u}||_2|\mathbf{v}^T\mathbf{x}| \leq ||\mathbf{u}||_2||\mathbf{v}||_2||\mathbf{x}||_2.$$

De ahí que $||A||_2 \leq ||\mathbf{u}||_2||\mathbf{v}||_2$. De nuevo, esta desigualdad se alcanza haciendo $\mathbf{x} = \mathbf{v}$, de modo que $||A||_2 = ||\mathbf{u}||_2||\mathbf{v}||_2$.

Proposición

Sean $A \in \mathbb{R}^{\ell \times m}$, $B \in \mathbb{R}^{m \times n}$ matrices, y sean $||\cdot||_{(\ell)}$, $||\cdot||_{(m)}$, $||\cdot||_{(n)}$, normas matriciales en $\mathbb{R}^{\ell}, \mathbb{R}^{m}, \mathbb{R}^{n}$, respectivamente.
Entonces

$$||AB||_{(\ell,n)} \leq ||A||_{(\ell,m)}||B||_{(m,n)}.$$

Prueba:
Por definición, para cualquier $\mathbf{x} \in \mathbb{R}^n$ vale

$$||AB\mathbf{x}||_{(\ell)} \leq ||A||_{(\ell,m)}||B\mathbf{x}||_{(m)} \leq ||A||_{(\ell,m)}||B||_{(m,n)}||\mathbf{x}||_{(n)}.$$

Luego,

$$||AB||_{(\ell,n)} = \sup_{||\mathbf{x}||_{(n)} \neq 0} \frac{||AB\mathbf{x}||_{(\ell)}}{||\mathbf{x}||_{(n)}} \leq ||A||_{(\ell,m)}||B||_{(m,n)}. \quad \square$$

Normas Matriciales

Esta desigualdad no es una igualdad. Por ejemplo, para matrices cuadradas vale en general $||A^n||_p \leq ||A||_p^n$, para todo $n \geq 1$.
Sin embargo, no se cumple en general que $||A^n||_p = ||A||_p^n$, para $n \geq 2$.

Ejemplo:

Para $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $p = 1$, tenemos que $A^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$, $\forall n \in \mathbb{N}$. Luego,

$$||A^n||_1 = n + 1, \quad ||A||_1^n = 2^n.$$

Normas Matriciales

En general, no todas las normas matriciales son inducidas por una norma vectorial. Damos la siguiente definición general.

Definición

Una ***norma matricial*** *es una función* $||\cdot|| : \mathbb{R}^{m\times n} \to \mathbb{R}$ *que satisface:*

1. $||A|| \geq 0, \ \forall A \in \mathbb{R}^{m\times n}$.
2. $||A+B|| \leq ||A|| + ||B||$, $\forall A, B \in \mathbb{R}^{m\times n}$.
3. $||\alpha A|| = |\alpha| \, ||A||$, $\forall \alpha \in \mathbb{R}, \forall A \in \mathbb{R}^{m\times n}$.

La norma matricial más importante que no proviene de una norma vectorial es la **norma de Frobenius** o **norma de Hilbert-Schmidt**

$$||A||_F = \Big(\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 \Big)^{1/2}. \tag{5}$$

Si $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{R}^m$ denotan las columnas de A, entonces

$$||A||_F = \Big(\sum_{j=1}^{n} ||\mathbf{a}_j||_2^2 \Big)^{1/2}. \tag{6}$$

Proposición

Para $A \in \mathbb{R}^{m\times n}$, vale $||A||_F = \sqrt{\mathrm{tr}(A^TA)} = \sqrt{\mathrm{tr}(AA^T)}$.

<u>Prueba</u>: Si $A = \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \ldots & \mathbf{a}_n \end{pmatrix}$ son las columnas de A, $\mathbf{a}_j \in \mathbb{R}^m$, entonces

$$A^TA = \begin{pmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_n \end{pmatrix} \begin{pmatrix} \mathbf{a}_1 \ldots \mathbf{a}_n \end{pmatrix} = \big(\mathbf{a}_i^T\mathbf{a}_j\big) .$$

Luego, $\mathrm{tr}(A^TA) = \sum_{j=1}^{n} \mathbf{a}_j^T\mathbf{a}_j = \sum_{j=1}^{n} ||\mathbf{a}_i||_2^2 = \sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}^2 = \sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2 = ||A||_F^2.$ □