



Mateo Andrés Manosalva Amaris

Edgar Santiago Ochoa Quiroga

Sergio Alejandro Bello Torres

Ejercicio 1.

Sea $A \in \mathbb{R}^{m \times n}$. Entonces se satisface:

(a) $\|A\|_2 = \|A^T\|_2 \leq \|A\|_F = \|A^T\|_F$

Demostración. Para mostrar que $\|A\|_2 = \|A^T\|_2$ veremos que todo valor propio no nulo de $A^T A$ lo es también de $A A^T$: Sea v un vector propio de $A^T A$ y λ el valor propio de $A^T A$ asociado a v , entonces

$$\begin{aligned} A A^T (A v) &= A (A^T A v) \\ &= A (\lambda v) \\ &= \lambda (A v) \end{aligned}$$

De esto se sigue que λ es un valor propio de $A A^T$ asociado al vector propio $A v$. De manera análoga, se concluye que si w es un vector propio de $A A^T$, con γ el valor propio de $A A^T$ asociado a w , entonces γ es un valor propio de $A A^T$ asociado al vector propio $A^T w$.

Para ver que $\|A\|_F = \|A^T\|_F$ basta notar que

$$\|A\|_F^2 = \sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 = \sum_{j=1}^n \sum_{i=1}^m |a_{ij}|^2 = \sum_{j=1}^n \sum_{i=1}^m |a_{ji}|^2 = \|A^T\|_F^2$$

Por último, note que las componentes de la diagonal de $A^T A$ son de la forma $u_{jj} = \sum_{i=1}^m a_{ij}^2$, por lo tanto la traza de $A^T A$ es

$$\sum_{j=1}^n \sum_{i=1}^m |a_{ij}|^2 = \|A\|_F^2$$

y ya que la traza de una matriz es igual a la suma de sus valores propios, y como $A^T A$ es semidefinida positiva, obtenemos que:

$$\begin{aligned} \lambda_{\max}(A^T A) &\leq \text{tr}(A^T A) \\ \sqrt{\lambda_{\max}(A^T A)} &\leq \sqrt{\text{tr}(A^T A)} \\ \|A\|_2 &\leq \|A\|_F \end{aligned}$$

□

(b) $\|A\|_\infty \leq \sqrt{n} \|A\|_2$

Demostración. Primero note que $\|Ae_j\|_2^2 = \sum_{i=1}^m |a_{ij}|^2 \leq \|A\|_2^2$, por lo tanto, $\sum_{j=1}^n \|Ae_j\|_2^2 \leq n\|A\|_2^2$. Ahora, usando la desigualdad de Cauchy-Schwartz obtenemos el siguiente resultado:

$$\begin{aligned}\|A\|_\infty^2 &= \left(\max_{1 \leq i \leq m} \sum_{j=1}^n |a_{ij}| \right)^2 \\ &\leq \max_{1 \leq i \leq m} \sum_{j=1}^n |a_{ij}|^2 \\ &\leq \sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 \\ &= \sum_{j=1}^n \sum_{i=1}^m |a_{ij}|^2 \\ &= \sum_{j=1}^n \|Ae_j\|_2^2 \\ &\leq n\|A\|_2^2\end{aligned}$$

Tomando raíz cuadrada llegamos a que $\|A\|_\infty \leq \sqrt{n}\|A\|_2$. □

(c) $\|A\|_2 \leq \sqrt{m}\|A\|_\infty$

Demostración. Note que

$$\begin{aligned}\|Ax\|_2^2 &= \sum_{j=1}^m \left| \sum_{i=1}^n a_{ij}x_j \right|^2 \\ &\leq \sum_{j=1}^m \left(\max_{1 \leq i \leq m} \sum_{j=1}^n |a_{ij}|^2 \right) x_j^2 \\ &= m \max_{1 \leq i \leq m} \sum_{j=1}^n |a_{ij}|^2 \\ &= m\|A\|_\infty^2\end{aligned}$$

Tomando raíz cuadrada se obtiene que $\|Ax\|_2 \leq \sqrt{m}\|Ax\|_\infty$ y por lo tanto $\|A\|_2 \leq \sqrt{m}\|A\|_\infty$ □

(d) $\|A\|_2 \leq \sqrt{\|A\|_1 \|A\|_\infty}$.

Demostración. Tenemos que $\|Ax\|_\infty = \max_{1 \leq i \leq n} \left| \sum_{j=1}^n a_{ij}x_j \right|$ y también que $\|Ax\|_1 =$

$\sum_{i=1}^m \left| \sum_{j=1}^n a_{ij} x_j \right|$. De esta manera

$$\begin{aligned} \|Ax\|_2^2 &= \sum_{i=1}^m \left| \sum_{j=1}^n a_{ij} x_j \right|^2 \\ &= \sum_{i=1}^m \left(\left| \sum_{j=1}^n a_{ij} x_j \right| \left| \sum_{j=1}^n a_{ij} x_j \right| \right) \\ &\leq \left(\sum_{i=1}^m \left| \sum_{j=1}^n a_{ij} x_j \right| \right) \max_{1 \leq i \leq m} \left| \sum_{j=1}^n a_{ij} x_j \right| \\ &= \|Ax\|_1 \|Ax\|_\infty \end{aligned}$$

Así, $\|Ax\|_2 \leq \sqrt{\|Ax\|_1 \|Ax\|_\infty}$, en consecuencia $\|A\|_2 \leq \sqrt{\|A\|_1 \|A\|_\infty}$ □

Ejercicio 2.

Sea $\|\cdot\|$ una norma en \mathbb{R}^n y A una matriz invertible de tamaño $n \times n$. Pruebe que:

Si $Ax = b$, $(A + \delta A)(x + \delta x) = b + \delta b$ y $\|A^{-1}\| \|\delta A\| < 1$, entonces $A + \delta A$ es invertible y se cumple que:

$$\frac{\|\delta x\|}{\|x\|} \leq \frac{\text{cond}(A)}{1 - \|A^{-1}\| \|\delta A\|} \left(\frac{\|\delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|b\|} \right).$$

Demostración. Primero mostremos que $A + \delta A$ es invertible. Notemos que como A es invertible tenemos

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

Luego por hipótesis y usando que la norma matricial es submultiplicativa tenemos que

$$\| -A^{-1}\delta A \| \leq \|A^{-1}\| \|\delta A\| < 1.$$

Por lo que $I - A^{-1}\delta A$ es invertible por el teorema de la serie de Neumann. Así $A + \delta A$ es invertible, pues es el producto de dos matrices invertibles.

Ahora probemos la desigualdad. Como por hipótesis $Ax = b$, notemos que

$$\begin{aligned} (A + \delta A)(x + \delta x) &= Ax + A\delta x + \delta Ax + \delta A\delta x \\ &= b + A\delta x + \delta Ax + \delta A\delta x. \end{aligned}$$

Reemplazando en la hipótesis tenemos que

$$b + A\delta x + \delta Ax + \delta A\delta x = b + \delta b,$$

luego

$$A\delta x = \delta b - \delta Ax - \delta A\delta x,$$

y como A es invertible

$$\delta x = A^{-1} \delta b - A^{-1} \delta A x - A^{-1} \delta A \delta x.$$

Si tomamos la norma en \mathbb{R}^n en ambos lados, por la desigualdad triangular y el hecho de que la norma matricial inducida es compatible y submultiplicativa tenemos que

$$\begin{aligned} \|\delta x\| &= \|A^{-1} \delta b - A^{-1} \delta A x - A^{-1} \delta A \delta x\| \\ &\leq \|A^{-1} \delta b\| + \|A^{-1} \delta A x\| + \|A^{-1} \delta A \delta x\| \\ &\leq \|A^{-1}\| \|\delta b\| + \|A^{-1}\| \|\delta A\| \|x\| + \|A^{-1}\| \|\delta A\| \|\delta x\|. \end{aligned}$$

Luego restando el ultimo sumando y factorizando $\|\delta x\|$ tenemos

$$\|\delta x\| (1 - \|A^{-1}\| \|\delta A\|) \leq \|A^{-1}\| \|\delta b\| + \|A^{-1}\| \|\delta A\| \|x\|.$$

Como A es invertible, es distinta de 0, luego $\|A\| > 0$, así tenemos que por la definición del número de condición $\|A^{-1}\| = \frac{\text{cond}(A)}{\|A\|}$. Además tenemos que $\|b\| = \|Ax\| \leq \|A\| \|x\|$, por tanto $\frac{1}{\|A\| \|x\|} \leq \frac{1}{\|b\|}$. Con todo esto notemos que para el lado derecho de la desigualdad tenemos que

$$\begin{aligned} \|A^{-1}\| \|\delta b\| + \|A^{-1}\| \|\delta A\| \|x\| &= \frac{\text{cond}(A)}{\|A\| \|x\|} \|\delta b\| \|x\| + \frac{\text{cond}(A)}{\|A\|} \|\delta A\| \|x\| \\ &\leq \frac{\text{cond}(A)}{\|b\|} \|\delta b\| \|x\| + \frac{\text{cond}(A)}{\|A\|} \|\delta A\| \|x\| \\ &= \text{cond}(A) \|x\| \left(\frac{\|\delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|b\|} \right). \end{aligned}$$

Por lo tanto tenemos que

$$\|\delta x\| (1 - \|A^{-1}\| \|\delta A\|) \leq \text{cond}(A) \|x\| \left(\frac{\|\delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|b\|} \right).$$

Como por hipótesis $0 < 1 - \|A^{-1}\| \|\delta A\|$, podemos dividir y mantener la desigualdad, así dividiendo también por $\|x\|$ concluimos que

$$\frac{\|\delta x\|}{\|x\|} \leq \frac{\text{cond}(A)}{1 - \|A^{-1}\| \|\delta A\|} \left(\frac{\|\delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|b\|} \right).$$

□□

Ejercicio 3.

Sea

$$A = \begin{pmatrix} \alpha & \alpha \\ \alpha & \alpha + \delta \end{pmatrix},$$

$\alpha > 0$ fijo, $\delta > 0$ variable.

- (a) Obtenga el número de condición de A . Para valores de δ muy pequeños o muy grandes, ¿podemos afirmar que el sistema $Ax = b$ está mal condicionado? Justifique su respuesta.

- (b) ¿Existe algún valor de δ que haga óptimo el número de condición de A ? ¿Cuál es este número de condición?

Ejercicio 4.

Al aproximar una función continua $f : [0, 1] \rightarrow \mathbb{R}$ mediante un polinomio $p(t) = a_n t^n + \dots + a_1 t + a_0$, el error de aproximación E se mide en la norma L^2 , es decir:

$$E^2 := \|p - f\|_{L^2}^2 = \int_0^1 [p(t) - f(t)]^2 dt.$$

- (a) Muestre que la minimización del error $E = E(a_0, \dots, a_n)$ conduce a un sistema de ecuaciones lineales $H_n a = b$, donde:

$$b = [b_0, \dots, b_n]^T \in \mathbb{R}^{n+1}, \quad b_i = \int_0^1 f(t) t^i dt, \quad i = 0, \dots, n,$$

y H_n es la matriz de Hilbert de orden n , definida como:

$$(H_n)_{i,j} = \frac{1}{i+j+1}, \quad i, j = 0, \dots, n.$$

El vector a representa los coeficientes del polinomio p .

Demostración. Sea $p(t) = \sum_{j=0}^n a_j t^j$, note que

$$\begin{aligned} E^2(a_0, \dots, a_n) &= \int_0^1 (p(t) - f(t))^2 dt \\ &= \int_0^1 \left(\sum_{j=0}^n a_j t^j - f(t) \right)^2 dt \\ &= \int_0^1 \left(\sum_{j=0}^n a_j t^j \right)^2 dt - 2 \int_0^1 f(t) \sum_{j=0}^n a_j t^j dt + \|f\|_{L^2}^2. \end{aligned}$$

Note que esta función depende de los coeficientes, por lo tanto para minimizar el error podemos derivar parcialmente con respecto a a_k , $0 \leq k \leq n$ e igualar a 0. Derivando obtenemos que

$$\frac{\partial E^2}{\partial a_k} = 2 \int_0^1 \left(\sum_{j=0}^n a_j t^{k+j} \right) dt - 2 \int_0^1 f(t) t^k dt = 0,$$

despejando de esta ecuación obtenemos que

$$\int_0^1 \sum_{j=0}^n a_j t^{j+k} dt = \sum_{j=0}^n a_j \int_0^1 t^{j+k} dt = \int_0^1 f(t) t^k dt,$$

más aún, como $\int_0^1 t^{k+j} dt = \frac{1}{k+j+1} = (H_n)_{k,j}$, obtenemos

$$\sum_{j=0}^n a_j (H_n)_{k,j} = (H_n)_k \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \int_0^1 f(t) t^k dt = b_k,$$

donde $(H_n)_k$ denota la k -ésima fila de la matriz (H_n) , que es lo mismo que $H_n a = b$.

□

(b) Muestre que H_n es simétrica y definida positiva.

Demostración. Note que la simetría es inmediata de que

$$(H_n)_{i,j} = \frac{1}{i+j+1} = \frac{1}{j+i+1} = (H_n)_{j,i}.$$

Para ver que la matriz es definida positiva note que

$$\begin{aligned} x^T H_n x &= (x_1, \dots, x_n) \begin{pmatrix} 1 & \cdots & \frac{1}{n+2} \\ \vdots & \ddots & \vdots \\ \frac{1}{n+2} & \cdots & \frac{1}{2n+1} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \\ &= \sum_{i=0}^n \sum_{j=0}^n \frac{x_j x_i}{i+j+1} \\ &= \sum_{i=0}^n \sum_{j=0}^n x_j x_i \int_0^1 t^i t^j dt \\ &= \int_0^1 \sum_{i=0}^n x_i t^i \sum_{j=0}^n x_j t^j dt \\ &= \left\| \sum_{j=0}^n x_j t^j \right\|_{L^2}^2 > 0 \end{aligned}$$

□

(c) Solucione el sistema $H_n x = b$, donde b tiene componentes $b_i = 1/(n+i+1)$, para $i = 0, \dots, n$. Para esto, use las factorizaciones LU ($[L, U] = \text{lu}(H)$) y Cholesky ($L = \text{chol}(H)$). Luego resuelva los dos sistemas triangulares.

Solución. El teorema de Stone-Weierstrass nos dice que los polinomios son densos en las funciones continuas, esto nos da que la aproximación que nos piden es posible, al truncar el número de coeficientes del polinomio vimos que el problema de optimizarlos se reduce a resolver el sistema $H_n x = b$, esto podemos programarlo en Matlab, Python, etc. Para este

caso nosotros hemos hecho el trabajo en ambos lenguajes con el propósito de comparar resultados, antes de presentar los códigos en Matlab observemos que

$$b_i = \frac{1}{n+i+1} = \int_0^1 f(t)t^i dt,$$

y por lo tanto $f(t) = t^n$, en efecto estamos aproximando el polinomio t^n por polinomios, así pues, esperaríamos una solución numérica del estilo $\alpha = (0, \dots, 1)^T$. Para el caso de la factorización LU implementamos

```
1      n = 10;
2      H = hilb(n); %Genera la matriz de Hilbert de orden n
3      b = zeros(n, 1); %Crea un vector con n ceros
4      for i = 1:n
5          b(i) = 1 / (i + n - 1); %Cambia las entradas por las del
                                   ejercicio
6      end
7      [L, U] = lu(H);
8
9      %Solucionamos los sistemas
10
11     y = L \ b;
12     a_LU = U \ y;
13
14     disp(a_LU')
```

En el caso de la factorización de Cholesky

```
1      n = 10;
2      H = hilb(n);
3      b = zeros(n, 1);
4      for i = 1:n
5          b(i) = 1 / (i + n - 1);
6      end
7      L = chol(H);
8      y = L' \ b;
9      a_LU = L \ y;
10     disp(a_LU')
```

En ambos casos tomamos un tamaño $n = 10$ para probar el algoritmo, en donde la factorización LU nos arrojó el resultado $x = (0, 0, 0, 0, 0, 0, 0, 0, 0, 1)^T$ y la factorización de Cholesky nos dió el resultado

$$x = \begin{bmatrix} 4,4438 \times e^{-11} \\ -3,8598 \times e^{-9} \\ 8,2505 \times e^{-8} \\ -7,5194 \times e^{-7} \\ 3,5931 \times e^{-6} \\ -9,8904 \times e^{-6} \\ 1,6243 \times e^{-5} \\ -1,5708 \times e^{-5} \\ 8,2514 \times e^{-6} \\ 1 \end{bmatrix}$$

- (d) Para ambos métodos, ¿qué tan precisas son las soluciones numéricas \hat{x}_{approx} ? Tabule los errores de la solución:

$$e(n) = \|\hat{x}_{\text{approx}} - x_{\text{exact}}\|$$

como una función de $n = 2, \dots, 15$. Note que $x_{\text{exact}} = (0, \dots, 1)^T$. Puede graficar los errores en función de n utilizando la función `semilogy` de Matlab. Explique en detalle los resultados.

Solución. Toca preguntarle al coyoyo

Ejercicio 5.

- (a) Simplifique el algoritmo de eliminación de Gauss para resolver un sistema lineal $Ax = b$ donde A es una matriz tridiagonal.
- (b) Considere la ecuación de Poisson con término fuente f en el intervalo $(0, 1)$:

$$-T''(x) = f(x), \quad x \in (0, 1),$$

con condiciones de frontera $T(0) = T(1) = 0$. Aproximando la segunda derivada por diferencias finitas:

$$T''(x) \approx \frac{T(x-h) - 2T(x) + T(x+h)}{h^2},$$

y discretizando en $x_i = ih$, $i = 0, 1, \dots, n$, con $n = 1/h$, obtenemos el sistema:

$$-T_{i-1} + 2T_i - T_{i+1} = h^2 f(x_i), \quad i = 1, \dots, n-1.$$

Escriba este sistema en la forma $AT = f$, donde A es una matriz tridiagonal. Resuelva el sistema para $n = 1000$ y $f(x) = \sin(2\pi x)$. Compare su solución con $T(x) = \sin(2\pi x)/(4\pi^2)$.