Universidade Estadual de Campinas

Instituto de Matemática, Estatística e Computação Científica.

EP1 - Domínio Numérico de Matrizes

MT404

Vinícius Oliveira Martins
RA: 206853

Setembro de 2023

1 Algumas Definições

Dado um subconjunto $\Omega \subset \mathbb{C}$, dizemos que esse subconjunto é **convexo** se para quaisquer dois pontos $x, y \in \Omega$ temos $tx + (1-t)y \subset \Omega$ para todo $t \in [0,1]$.

Para uma matriz $A \in \mathbb{C}^{n\times n}$ podemos definir alguns subconjuntos do plano complexo que nos dão informações sobre essa matriz. A multiplicação de A por um vetor $x \in \mathbb{C}^n$ pode ser pensada como um operador de $\mathbb{C}^n$ em $\mathbb{C}^n$, ou seja

$$\begin{aligned} T : \mathbb{C}^n &\to \mathbb{C}^n \\ x &\mapsto Ax \ , \ T(x) = Ax \ . \end{aligned}$$

Assim, os vetores de $\mathbb{C}^n$ que são levados a algum de seus múltiplos pelo operador T são ditos autovetores de A. Assim, temos a equação de autovetor

$$Av = \lambda v \ , \tag{1}$$

onde λ é um número complexo denominado autovalor. Dizemos ainda que v é o autovetor de A associado ao autovalor λ. Se A possuir m autovetores, $\{v_1, ..., v_m\}$, teremos $n \leq m$ autovalores, $\sigma(A) = \{\lambda_1, ..., \lambda_n\}$, com $\lambda_1 < ... < \lambda_n$, e $\sigma(A)$ é denominado **espectro** de A.

Além do espectro de A podemos também definir o **domínio numérico** de A como o subconjunto

$$W(A) = \{x^*Ax : x \in \mathbb{C}^n \ e \ x^*x = 1\} \subset \mathbb{C} \ . \tag{2}$$

. Neste texto iremos estudar a relação entre o domínio numérico e o espectro para uma matriz A, e como construir o domínio numérico usando um algoritmo simples.

2 Domínio Numérico de Matrizes Normais

Dada uma matriz $A \in \mathbb{C}^{n\times n}$ normal, temos sua **diagonalização**

$$A = UDU^* \ , \tag{3}$$

onde U é uma matriz unitária formada pelos autovetores de A e $D = diag(\lambda_1, ..., \lambda_n)$.

Dado, $x \in \mathbb{C}^n$ tal que $x^*x = 1$, temos

$$x^*Dx = \sum_{j=1}^{n} |x_j|^2 \lambda_j \ , \tag{4}$$

e

$$x^*x = \sum_{j=1}^{n} |x_j|^2 = 1 \Rightarrow |x_j|^2 \geq 0 \ , \quad \forall j = 1, ..., n \ . \tag{5}$$

Logo, x^*Dx é uma combinação convexa de $L = \{\lambda_1, ..., \lambda_n\}$. Assim, $W(D)$ é o invólucro convexo de L. Pela **P6**, temos que $W(A) = W(\Lambda)$. Portanto, $W(A)$ é o invólucro convexo dos autovalores de A.

Considere agora $A \in \mathbb{C}^{n\times n}$ hermitiana. Assim, se (λ, v) é um autopar de A, temos que (λ, v) também é autopar de A^*. Logo,

$$Av = \lambda v \Rightarrow v^*Av = \lambda v^*v = \lambda||v||^2 \ . \tag{6}$$

Além disso,

$$v^*Av = v^*A^*v = (Av)^*v = \bar{\lambda}v^*v = \bar{\lambda}||v||^2 \ . \tag{7}$$

Assim, de (6) e (7) temos que $\lambda = \bar{\lambda}$, ou seja, $\lambda \in \mathbb{R}$. Assim, $\sigma(A) \subset \mathbb{R}$. Como A hermitiana, temos que $AA^* = A^*A = \mathbb{I}$, logo é um caso particular de A normal. Assim, temos que $W(A)$ continua sendo o invólucro convexo de $\sigma(A)$, que é o intervalo fechado $[\lambda_1; \lambda_n] \subset \mathbb{R}$. Portanto, $W(A) = [\lambda_1; \lambda_n]$.

3 Teorema de Hausdorff-Toeplitz

Podemos enunciar o **Teorema de Hausdorff-Toeplitz** como:

Teorema 3.1. *Dada uma matriz $A \in \mathbb{C}^{n\times n}$, temos que seu domínio numérico é um subconjunto convexo de $\mathbb{C}$.*

O primeiro fato que podemos observar sobre $W(A)$ é que, sendo um subconjunto do plano complexo, se $W(A)$ for convexo, teremos

$$tv + (1-t)w = t(x^*Ax) + (1-t)(y^*Ay) \in W(A) \subset \mathbb{C} \ \forall t \in (0, 1) \ , \tag{8}$$

onde

$$\begin{aligned} &x, y \in \mathbb{C}^n \\ &v = x^*Ax \ , \quad w = y^*Ay \\ &x^*x = y^*y = 1 \ . \end{aligned}$$

Dados dois vetores ortonormais $x, y \in \mathbb{C}^n$. O conjunto $\{x, y, e_1, ..., e_n\}$, onde e_j são os vetores da base canônica de $\mathbb{C}^n$, é tal que $span(\{x, y, e_1, ..., e_n\}) = \mathbb{C}^n$. Reduzindo a matriz $M = [x|y|e_1|...|e_n]$ a sua forma canônica por linhas temos a matriz $\bar{M} = [x|y|\xi_1|...|\xi_{n-2}]$, onde $\xi_j = e_k$ para algum k, e $\{x, y, \xi_1, ..., \xi_{n-1}\}$ é uma base de $\mathbb{C}^n$. Aplicando o processo de Gram-Schmidt nos vetores dessa base temos a matriz unitária U, tal que $x = Uv$ e $y = Uw$, onde $v = (v_1, v_2, 0, ..., 0)^t$ e $w = (w_1, w_2, 0, ..., 0)^t$.

Assim, temos

$$\begin{aligned} tx^*Ax + (1-t)y^*Ay &= tv^*U^*AUv + (1-t)w^*U^*AUw \\ &= tv^*Bv + (1-t)w^*Bw \\ &= t\xi^*B_2\xi + (1-t)\eta^*B_2\eta \ , \end{aligned} \tag{9}$$

onde B_2 é a submatriz principal 2×2 de B e $\xi, \eta \in \mathbb{C}^2$ e $\xi, \eta \in \mathbb{C}^2$ são os dois primeiros elementos de v e w, respectivamente.

Além disso, pela propriedade de translação podemos considerar apenas matrizes com traço nulo já que uma matriz A de traço não nulo terá um domínio numérico com a mesma fronteira que o domínio numérico de $A - \frac{1}{2}Tr(A)\mathbb{I}$, ou seja, $W(A)$ é apenas $W(A - \frac{1}{2}Tr(A)\mathbb{I})$ transladado.

Sabemos que o traço de uma matriz é igual a soma de seus autovalores. Assim, dada uma matriz $A \in \mathbb{C}^{2\times 2}$, tal que $Tr(A) = 0$, seus autovalores serão $\pm z$, $z \in \mathbb{C}$. Sendo a_1 o autovetor unitário associado a $+z$ e a_2 o autovetor unitário associado a $-z$, o vetor $b = e^{i\phi}a_1 + a_2$ é não nulo para qualquer $\phi \in \mathbb{R}$. Logo

$$\begin{aligned} b^*Ab &= z(e^{i\phi}a_2^*a_1 - e^{-i\phi}a_1^*a_2) \\ &= 2zIm(e^{i\phi}a_1^*a_2)i \ . \end{aligned}$$

Tomando $\phi = k\pi$, $k \in \mathbb{Z}$, temos que $b^*Ab = 0$. Usando $\frac{b}{||b||}$ como a primeira coluna de uma matriz unitária W (obtida pelo processo de Gram-Schmidt do mesmo modo descrito anteriormente), temos $(W^*AW)_{11} = (W^*AW)_{22} = 0$. Assim, é suficiente considerarmos uma matriz

$$A = \begin{bmatrix} 0 & z_1 \\ z_2 & 0 \ . \end{bmatrix} \tag{10}$$

Ainda podemos simplificar ainda mais a matriz A. Sendo $z_1 = |z_1|e^{i\theta_1}$ e $z_2 = |z_2|e^{i\theta_2}$, podemos construir a mais uma matriz unitária

$$Y = \begin{bmatrix} 1 & 0 \\ 0 & e^{-i\theta} \end{bmatrix} \ , \tag{11}$$

tal que

$$\begin{bmatrix} 1 & 0 \\ 0 & e^{-i\theta} \end{bmatrix} \begin{bmatrix} 0 & z_1 \\ z_2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{bmatrix} = e^{i\psi} \begin{bmatrix} 0 & |z_1| \\ |z_2| & 0 \end{bmatrix} \ , \tag{12}$$

onde $\theta = \frac{1}{2}(\theta_2 - \theta_1)$ e $\psi = \frac{1}{2}(\theta_1 - \theta_2)$. Portanto, basta considerarmos uma matriz

$$A = \begin{bmatrix} 0 & m \\ n & 0 \end{bmatrix} \ , \tag{13}$$

onde m, n são reais não negativos, para provarmos o teorema de de Hausdorff-Toeplitz.

Assim, consideremos uma matriz A da forma (13). Pela propriedade do conjugado complexo podemos supor ainda que $m \geq n$. Agora, dado um vetor unitário $v \in \mathbb{C}^2$, podemos parametrizar este vetor como $v = e^{i\psi}(sin(\phi), e^{i\theta}cos(\phi))^t$, onde $\psi, \theta \in [0, 2\pi]$ e $\phi \in [0, \frac{\pi}{2}]$. Logo,

$$v^*Av = me^{i\theta}\sin(\phi)\cos(\phi) + ne^{-i\theta}\sin(\phi)\cos(\phi) \tag{14}$$

$$= \frac{1}{2}\sin(2\phi)(me^{i\theta} + ne^{-i\theta}) \ , \tag{15}$$

ou seja

$$W(a) = \left\{ z = x + iy : x = \frac{m+n}{2}\sin(2\phi)\cos(\theta),\ y = \frac{m-n}{2}\sin(2\phi)\sin(\theta) \right\} . \tag{16}$$

Da norma de z, temos que

$$\frac{2x^2}{(m+n)^2} + \frac{2y^2}{(m-n)^2} = \sin^2(2\phi) . \tag{17}$$

Logo, a região delimitada por $W(A)$ é, para $\phi = \frac{\pi}{4}$, uma elipse. Podemos também observar dois casos especiais. O primeiro e mais direto é o caso trivial em que $m = n = 0$, que implica em $W(A) = \{0\}$. Para o caso em que $m = n$ a matriz A se torna hermitiana e voltamos ao que foi discutido na seção anterior, nesse caso $W(A) = [-m, m]$.

3.1 Outra Atividade do Projeto

Lema 3.1. *Uma matriz $A \in \mathbb{C}^{2\times 2}$ de traço nulo é unitariamente equivalente a uma matriz de diagonal nula.*

Primeiro, consideramos a matriz unitária

$$U = \begin{bmatrix} \sin(\theta) & e^{i\phi}\cos(\theta) \\ -e^{-i\phi}\cos(\theta) & \sin(\theta) \end{bmatrix} , \tag{18}$$

onde $\phi \in [0, 2\pi]$ e $\theta \in [0, \frac{\pi}{2}]$. Logo, temos

$$UAU^* = U \begin{bmatrix} c & a \\ b & -c \end{bmatrix} U^* . \tag{19}$$

Assim, a diagonal da matriz UAU^* será

$$(UAU^*)_{11} = -(UAU^*)_{22} = \frac{1}{2}\sin(2\theta)(ae^{i\phi} + be^{-i\phi}) - c\cos(2\theta) . \tag{20}$$

Assim, $(UAU^*)_{11} = (UAU^*)_{22} = 0$ se

$$w = \frac{c}{ae^{i\phi} + be^{-i\phi}} , \tag{21}$$

for um número real e $\tan(2\theta) = 2w$. Podemos reescrever w como

$$w = \frac{c[(a+b)\sin(\phi) - (a-b)\cos(\phi)i]}{(a+b)^2\sin^2(\phi) + (a-b)^2\cos^2(\phi)} , \tag{22}$$

para que w seja real precisamos que $Im(w) = 0$, ou seja

$$Im(w) = \frac{c(a-b)\cos(\phi)}{(a+b)^2\sin^2(\phi) + (a-b)^2\cos^2(\phi)} = 0 \tag{23}$$

$$\Rightarrow c(a-b)\cos(\phi) = 0 \tag{24}$$

$$\phi = \frac{\pi}{2} + k\pi \ , \quad k \in \mathbb{Z} \ . \tag{25}$$

4 Aproximação Numérica de $W(A)$

A estratégia que usamos para determinar uma aproximação numérica para o domínio numérico de uma matriz $A \in \mathbb{C}^{n\times n}$ será calcular diversos pontos da fronteira de $W(A)$ e uni-los com uma poligonal. Essa estratégia é valida pois sabendo que $W(A)$ é convexo uma poligonal que une uma certa quantidade de pontos de $\partial W(A)$ é uma boa aproximação de $W(A)$. Para calcularmos esses pontos de $\partial W(A)$ primeiro devemos notar que sempre podemos separar A como

$$A = \frac{1}{2}(A + A^*) + \frac{1}{2}(A - A^*) = A_H + A_I, \tag{26}$$

onde A_H é hermitiana e A_I é anti-hermitiana. Como iA_I é uma matriz hermitiana, temos que o espectro de A_I é composto apenas por números imaginários puros. Como isso, temos

$$x^*Ax = x^*A_Hx + x^*A_Ix \ . \tag{27}$$

Portanto, o maior autovalor de A_H, λ_{max} é a coordenada do ponto mais a esquerda de $W(A)$. Se v_{max} é o autovetor associado a λ_{max}, então

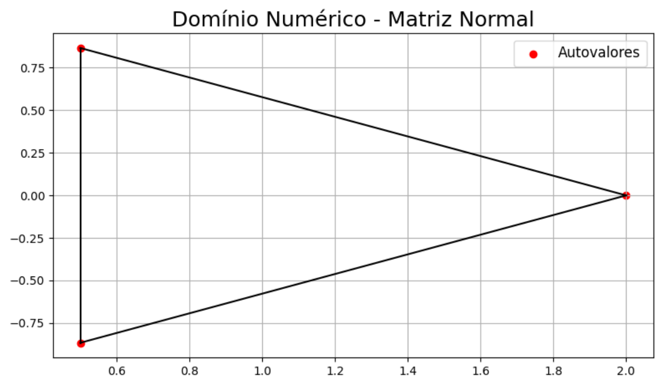
$$z_{max} = v_{max}^* A v_{max} \ , \tag{28}$$

é o ponto mais a esquerda de $W(A)$.

Sabemos que $W(z_0A) = z_0W(A)$ para qualquer $z_0 \in \mathbb{C}^n$. Assim, a matriz $e^{-i\phi}A$ terá $W(e^{-i\phi}A) = e^{-i\phi}W(A)$, ou seja, seu domínio numérico será igual ao domínio numérico de A rotacionado pelo ângulo ϕ. Fazendo a separação de $e^{-i\phi}A$ em parte hermitiana e anti-hermitiana, encontramos o ponto mais a esquerda de $W(e^{-i\phi}A)$, $z_{max}(\phi)$, como descrito anteriormente. Assim, $z_{max}(\phi)$ é um outro ponto de $\partial W(A)$ com uma distância angular ϕ de z_{max}.

Fazendo $\phi \in [0, 2\pi)$, teremos que $z_{max}(0) = z_{max}$ e a sequência obtida $(z_{max})_\phi$ pertence a $\partial W(A)$. A qualidade dessa aproximação depende da quantidade de pontos que escolhemos para discretizar $[0, 2\pi)$. Essa quantidade de pontos pode mudar de matriz para matriz mas é possível determinar esse valor empiricamente.

Uma implementação deste algoritmo esta disponível neste repositório. No caso em que A é uma matriz normal não é necessário usar o algoritmo descrito. Como foi mostrado na seção (2) o domínio numérico de uma matriz normal é o envolucro convexo dos autovalores de A. Assim, basta calcularmos os autovalores de A para determinar $W(A)$. A implementação disponibilizada, lida com esse caso.



$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}.$$

Figura 1: Domínio numérico de uma matriz normal.

A seguir temos mais alguns domínios numéricos.

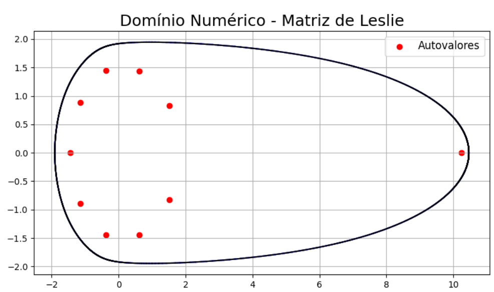


Figura 2: Domínio numérico de uma matriz de Leslie 10×10.

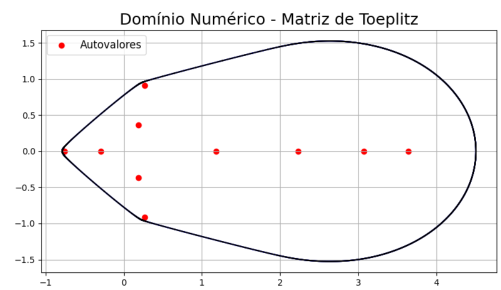


Figura 3: Domínio numérico de uma matriz de Toeplitz 10×10.

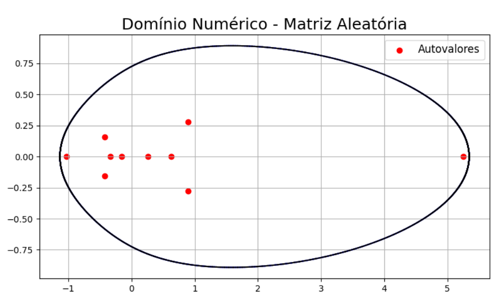


Figura 4: Domínio numérico de uma matriz aleatória 10×10 com elementos entre 0 e 1.

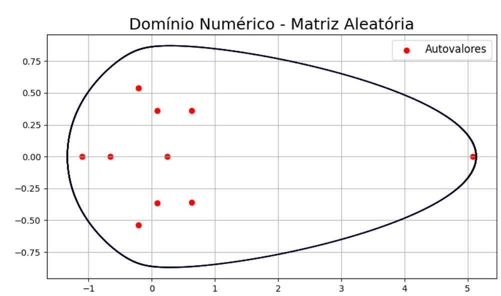


Figura 5: Domínio numérico de uma matriz aleatória 10×10 com elementos entre 0 e 1.

Referências

[1] Charles R. Johnson Roger A. Horn. *Topics in matrix analysis.* Cambridge University Press, cup edition, 1994.