

Texto da palestra

Teorema de Gerschgorin...
ou será Geršgorin?
ou ainda Geršhgorin?
Gersgorin? Geršgorin?
Gershgorin? Geršgorin?
...

Jaime Gaspar*

2004/09/19

1 Introdução¹

Vamos falar um pouco sobre o Teorema de Gerschgorin, generalizações deste resultado e teoremas de não-singularidade de matrizes, intimamente relacionados com os resultados anteriores.

2 Grafias do nome Gerschgorin

Nos livros e revistas, o nome Gerschgorin aparece escrito de muitas formas diferentes. Temos no acetato uma lista de algumas formas como o nome Gerschgorin surge. Indicamos a vermelho as variações em relação à primeira forma da lista. Apresentamos no acetato um “top +” com as formas mais vulgares, indicando-se a frequência relativa das formas. A forma “Gerschgorin”, que surge 35% das vezes, é a mais vulgar, pelo que é a forma que adoptamos.

3 Teorema de Gerschgorin

Vamos precisar da seguinte notação: dada uma matriz complexa tipo $n \times n$, definimos os R_i como sendo a soma dos módulos dos elementos não diagonais da i -ésima linha e, analogamente, definimos os S_i como sendo a soma dos módulos dos elementos não diagonais de i -ésima coluna.

*www.jaimegaspar.com.

¹Cada secção é referente a um acetato com o mesmo número da secção.

O Teorema de Gerschgorin diz-nos que os valores próprios de uma matriz complexa tipo $n \times n$ “caem” dentro dos n círculos de Gerschgorin de centro a_{ii} e raio R_i .

É natural perguntar como se distribuem os valores próprios pelos círculos. Por exemplo, terá cada círculo de Gerschgorin de ter um valor próprio? A resposta é não. No entanto, prova-se que em cada conjunto de k círculos, disjuntos dos restantes $n - k$, “caem” exactamente k valores próprios (contando multiplicidades). Resulta que um círculo de Gerschgorin disjunto dos restantes círculos tem um e um só valor próprio (contando multiplicidades).

4 Exemplo do Teorema de Gerschgorin

Vejamos um exemplo da aplicação do Teorema de Gerschgorin. Consideremos a matriz do acetato. O Teorema de Gerschgorin diz-nos que os valores próprios da matriz pertencem à união dos círculos na figura do acetato. Por exemplo, Z_1 é o círculo de centro $a_{ii} = -4 + 2i$, e raio $R_1 = |1| + |1|$, representado a vermelho no acetato. As cruces na figura representam os valores próprios.

Podemos ver que como diz o Teorema de Gerschgorin, os valores próprios “caem” dentro dos círculos. Na região $Z_1 \cup Z_2$ (disjunta do restante círculo), “caem” exactamente dois valores próprios, e no círculo Z_3 (disjunto dos restantes círculos) “cai” exactamente um valor próprio.

5 Mais exemplos do Teorema de Gerschgorin

Vejamos dois casos extremos da distribuição dos valores próprios pelos círculos de Gerschgorin. Na figura de cima do acetato, temos um exemplo da aplicação do Teorema de Gerschgorin a uma matriz, surgindo dois círculos de Gerschgorin sem qualquer valor próprio. Na figura de baixo, temos outro exemplo, onde cada círculo de Gerschgorin têm exactamente um valor próprio (porque os círculos são disjuntos).

6 Exemplo do Teorema de Gerschgorin aplicado à matriz transposta

Uma vez que a matriz A e a matriz transposta de A , A^T , têm os mesmos valores próprios, podemos aplicar o Teorema de Gerschgorin a A^T e obter uma versão do teorema onde em vez de intervirem as linhas da matriz A , intervêm as colunas, isto é, em vez de intervirem os R_i , intervêm os S_i .

Na figura do acetato, representamos a vermelho as regiões dadas para o Teorema de Gerschgorin para A e a amarelo as regiões dadas pelo Teorema de Gerschgorin para A^T . Como os valores próprios de A pertencem à reunião das regiões a vermelho e à união das regiões a amarelo, então pertencem à intersecção destas duas uniões, o que nos dá regiões menores.

7 Motivação para as generalizações

Gostaríamos de ter resultados análogos ao Teorema de Gerschgorin, mas que nos dessem regiões menores. Por exemplo, na figura do acetato, se o Teorema de Gerschgorin nos desse as regiões a vermelho, um resultado que o melhorasse podia dar as regiões a amarelo, e um outro resultado ainda melhor podia dar as regiões a verde. Vamos ver duas generalizações do Teorema de Gerschgorin que fazem exactamente isso.

8 Grafo $D(A)$

Para podermos ver os resultados que generalizam o Teorema de Gerschgorin, precisamos de três definições.

A cada matriz A complexa tipo $n \times n$ podemos associar um grafo orientado (essencialmente o mesmo que um grafo, mas com a diferença de darmos uma orientação às arestas) $D(A)$ com n vértices v_1, \dots, v_n , e que tem uma aresta orientada de v_i para v_j se e só se a entrada a_{ij} da matriz é não nula.

Por exemplo, para a matriz A tipo 3×3 do acetato, o grafo $D(A)$ tem 3 vértices e tem uma aresta a ir de v_1 para v_2 porque a entrada $a_{12} = -1$ é não nula.

9 Circuito de um grafo orientado

Um circuito de comprimento m de um grafo orientado é uma sequência de m aresta sucessivamente adjacentes, com $m \geq 1$, que não repete as arestas nem os vértices por onde passa o circuito, com a excepção do primeiro e último vértice, que exigimos que sejam o mesmo.

Uma vez que um circuito fica determinado pelos vértices por onde passa e pela ordem pela qual o faz, usamos a notação $v_{i_0} \rightarrow \dots \rightarrow v_{i_m}$ para denotar o circuito $(v_{i_0}, v_{i_1}), \dots, (v_{i_{m-1}}, v_{i_m})$.

Mais à frente vamos querer escrever produtórios onde o índice do produtório varia no conjunto dos índices dos vértices de um circuito. Assim, será útil escrever $k \in \gamma$ para indicar que v_k é um vértice por onde passa o circuito γ .

Vamos denotar os conjunto dos circuitos com comprimento pelo menos dois de um grafo orientado G por $\mathcal{C}_2(G)$.

10 Exemplo de circuitos

Consideremos o grafo do acetato. Apresentamos no acetato a lista dos circuitos do grafo. Justifiquemos que $v_1 \rightarrow v_2 \rightarrow v_3 \rightarrow v_1$ é um circuito de comprimento três. Tem início em v_1 , segue para v_2 respeitando a orientação da aresta, depois para v_3 também

seguindo a orientação de uma aresta e volta a v_1 , novamente respeitando a orientação de uma aresta. Não repete arestas e não repete os vértices, com excepção do primeiro e último vértice (que é v_1), que tem de ser igual. O circuito $v_1 \rightarrow v_2 \rightarrow v_3 \rightarrow v_1$ tem comprimento três porque passa por exactamente três arestas.

11 Matriz fracamente irredutível

Finalmente, a última das três definições. Dizemos que uma matriz A complexa tipo $n \times n$ é fracamente irredutível se e só se todo o vértice de $D(A)$ pertence a algum circuito de comprimento pelo menos dois.

Retomemos a matriz e o grafo orientado anteriores, apresentada no acetato. Como qualquer dos vértices v_1 , v_2 e v_3 pertence, por exemplo, ao circuito $v_1 \rightarrow v_2 \rightarrow v_3 \rightarrow v_1$ de comprimento 3, então a matriz é fracamente irredutível.

12 Primeira série de teoremas

Vamos ver duas generalizações do Teorema de Gerschgorin, que juntamente com este formam a nossa primeira série de teoremas.

O Teorema de Brauer é semelhante ao Teorema de Gerschgorin, com as diferenças de exigirmos $n \geq 2$ e em vez da condição “simples” $|z - a_{ii}| \leq R_i$, temos os produtos $|z - a_{ii}||z - a_{jj}| \leq R_i R_j$. As regiões do Teorema de Brauer são tipicamente ovais.

O terceiro teorema faz intervir vários factores em $\prod_{i \in \gamma} |z - a_{ii}| \leq \prod_{i \in \gamma} R_i$, tantos quanto o comprimento do circuito γ . Por esta razão as regiões dadas por este teorema tendem a ser complexas. Neste teorema fazemos a união com γ a percorrer $\mathcal{C}_2(D(A))$, o conjunto dos circuitos de comprimento pelo menos dois do grafo $D(A)$.

13 Exemplo do Teorema de Brauer

Consideremos a matriz do acetato. O Teorema de Brauer diz-nos que os valores próprios da matriz pertencem à união das regiões apresentadas no acetato. Por exemplo, Z_{12} , representada a vermelho na figura do acetato, surge nas “proximidades” de 4 e de $-3 + 3i$, pois na sua definição surge o produto $|z - 4||z - (-3 + 3i)|$.

14 Exemplo do teorema 3

Consideremos a matriz do acetato. O teorema 3 diz-nos que os valores próprios da matriz pertencem à união das regiões apresentadas no acetato. As regiões são as apresentadas porque os circuitos de comprimento pelo menos dois de $D(A)$ são, a menos do vértice onde começam, $v_1 \rightarrow v_2 \rightarrow v_3 \rightarrow v_1$ e $v_1 \rightarrow v_3 \rightarrow v_1$. (Circuitos que passam pelos mesmos vértices pela mesma ordem, mas que começam em vértices diferentes, dão origem às mesmas regiões, pelo que é suficiente considerar os dois circuitos que surgem

indicados como índice γ nos Z_γ). O circuito $v_1 \rightarrow v_2 \rightarrow v_3 \rightarrow v_1$ está representado a vermelho na figura do acetato, e surge nas “proximidades” das entradas diagonais $-3-2i$, $3+2i$ e $-2+2i$, porque estas surgem no produto da desigualdade que define esta região. O circuito $v_1 \rightarrow v_3 \rightarrow v_1$ está representado a verde e surge nas “proximidades” das entradas diagonais $-3-2i$ e $-2+2i$.

15 Implicações entre a primeira série de teoremas

Vamos comparar os teoremas da primeira série, no sentido de dizer “quem implica quem”. Diremos que um primeiro teorema implica um segundo teorema se e só se para toda a matriz satisfazendo as hipóteses de ambos os teoremas, temos que a união das regiões dadas pelo primeiro teorema está contida na união das regiões dadas pelo segundo teorema.

Neste sentido, os dois novos teoremas implicam o Teorema de Gerschgorin. O terceiro teorema também implica o segundo. As implicações recíprocas das referidas são falsas. Podemos dizer que o teorema 3 é o mais “forte”, seguido pelo Teorema de Brauer e depois pelo Teorema de Gerschgorin, sendo portanto este último o mais fraco.

16 Teorema 1 não implica o teorema 2

Vejamos exemplos das não implicações.

Para a matriz do acetato, que satisfaz as hipóteses dos teoremas de Gerschgorin e Brauer, temos que as regiões dadas pelo Teorema de Gerschgorin, a vermelho no acetato, não estão contidas nas regiões dadas pelo Teorema de Brauer, a amarelo, logo o Teorema de Gerschgorin não implica o Teorema de Brauer.

17 Teorema 1 não implica o teorema 3

Para a matriz do acetato, que satisfaz as hipóteses dos teoremas de Gerschgorin e do teorema 3, temos que as regiões dadas pelo Teorema de Gerschgorin, a vermelho no acetato, não estão contidas nas regiões dadas pelo teorema 3, a verde, logo o Teorema de Gerschgorin não implica o teorema 3.

18 Teorema 2 não implica o teorema 3

Finalmente, para a matriz do acetato, que satisfaz as hipóteses dos teoremas de Brauer e 3, temos que as regiões dadas pelo Teorema de Brauer, a amarelo no acetato, não estão contidas nas regiões dadas pelo teorema 3, a verde, logo o Teorema de Brauer não implica o teorema 3.

19 Segunda série de teoremas

Vejam agora três teoremas de não-singularidade de matrizes, intimamente relacionados com os três teoremas anteriores.

Dizemos que uma matriz complexa tipo $n \times n$ é de diagonal dominante por linhas se e só se cada entrada diagonal é, em módulo, maior que a soma dos módulos das restantes entradas da sua linha, isto é, maior que o correspondente R_i . Analogamente definimos matriz de diagonal dominante por colunas, em cuja definição em vez dos R_i surgem os S_i .

O Teorema de Lévy-Desplanques diz-nos que uma matriz de diagonal dominante por linhas é invertível. O mesmo vale para matrizes de diagonal dominante por colunas. (basta aplicar o Teorema de Lévy-Desplanques, na “versão para linhas”, à matriz A^T).

O teorema B, também chamado Teorema de Brauer, é semelhante ao teorema A, com a diferença que “acrescenta” um produto $|a_{ii}||a_{jj}| > R_i R_j$ às desigualdades $|a_{ii}| > R_i$ do teorema de Lévy-Desplanques, da mesma forma que o Teorema 2 (de Brauer) da primeira série acrescenta um produto às condições do Teorema de Gerschgorin. Analogamente ao que acontece no terceiro teorema da série anterior, este teorema C, apenas válido para matrizes fracamente irredutíveis, tem nas desigualdades tantos factores quanto o comprimento do circuito γ .

20 Teorema de Gerschgorin implica o Teorema de Lévy-Desplanques

Vejam como partindo do Teorema de Gerschgorin, provamos facilmente o Teorema de Lévy-Desplanques. Tomemos uma qualquer matriz satisfazendo as hipóteses do Teorema de Lévy-Desplanques, isto é, uma matriz $A = [a_{ij}]$ complexa tipo $n \times n$ de diagonal dominante por linhas. Por definição de diagonal dominante por linhas, temos $\forall i \in \{1, \dots, n\}$, $|a_{ii}| > R_i$. Podemos rescrever estas desigualdades na forma $\forall i \in \{1, \dots, n\}$, $|0 - a_{ii}| > R_i$. Se, para todo o $i \in \{1, \dots, n\}$, temos $|0 - a_{ii}| > R_i$, então para nenhum $i \in \{1, \dots, n\}$, temos $|0 - a_{ii}| \leq R_i$. Ora, $|0 - a_{ii}| \leq R_i$ é a condição que define 0 pertencer ao i -ésimo círculo de Gerschgorin, Z_i . Então $\nexists i \in \{1, \dots, n\} : 0 \in Z_i$. Portanto, $0 \notin \bigcup_{i=1}^n Z_i$. Mas, pelo Teorema de Gerschgorin, temos que os valores próprios de A pertencem a $\bigcup_{i=1}^n Z_i$, e vimos que 0 não pertence a esta união, logo 0 não é valor próprio de A . Mas de 0 não ser valor próprio de A vem que $\det A \neq 0$, que é o que afirma a tese do Teorema de Lévy-Desplanques.

A implicação recíproca demonstra-se com igual facilidade e por meio de argumentos do mesmo género.

21 Equivalências entre a primeira e segunda séries de teoremas

Os teoremas da primeira e segunda séries formam pares em que os elementos de um mesmo par são equivalentes. Um dos elementos do par dá-nos regiões onde “caem” os valores próprios, e o outro dos elementos dá-nos uma condição suficiente para uma matriz ser invertível.

As equivalências mostram-se com argumentos semelhantes aos que usámos para provar que o Teorema de Gerschgorin implica o Teorema de Lévy-Desplanques, tornando-se apenas ligeiramente mais complicadas à medida que “avancamos” nas séries.

Notemos que para cada par, num dos elementos do par temos as desigualdades “módulos das entradas diagonais e com $>$ ” (por exemplo, $|a_{ii}||a_{jj}| > R_i R_j$), e no outro temos as desigualdades “módulos de z menos as entradas diagonais com \leq ” (por exemplo, $|z - a_{ii}||z - a_{jj}| \leq R_i R_j$).

22 Terceira série de teoremas

Apenas como curiosidade, vejamos rapidamente mais duas séries de teoremas.

Os teoremas anteriores tanto se podem aplicar à matriz A como à matriz A^T (que tem os mesmos valores próprios que A). Aplicando os teoremas à matriz A^T , obtemos novas versões dos teoremas onde em vez dos R_i surgem os S_i , porque em vez de intervirem as linhas da matriz, intervêm as colunas. Estas novas duas séries fazem intervir simultaneamente os R_i e os S_i , controlando-se o “peso” que os R_i e os S_i têm por meio de um valor p que está entre 0 e 1.

A terceira série apresentada neste acetato é difere da primeira série apenas por onde na primeira série aparece só R_i , nesta terceira série aparece $R_i^p S_i^{1-p}$.

23 Exemplo da terceira série

Consideremos a matriz do acetato. As regiões dadas pelo primeiro teorema da terceira série estão representadas, na figura do acetato, desde cinzento claro para $p = 0$ até preto para $p = 1$. Por exemplo, a região mais à esquerda aumenta com p , uma vez que para p próximo de 0 temos as regiões mais claras menores e para p próximo de 1 temos as regiões escuras maiores. Já com a região ao centro passa-se o contrário: diminui com o aumento de p .

24 Quarta série de teoremas

Esta quarta série está para a segunda série como a anterior está para a primeira: onde na segunda série aparece R_i , aparece agora $R_i^p S_i^{1-p}$.

25 Exemplo da quarta série

Para exemplificar a utilidade de podermos fazer intervir simultaneamente as linhas e colunas da matriz, vejamos um exemplo no qual o Teorema de Brauer não se aplica “por linhas” (isto é, à matriz A) nem “por colunas” (isto é, à matriz A^T), mas aplica-se se usarmos um teorema onde intervenham simultaneamente as linhas e as colunas. Para a matriz A apresentada no acetato, o Teorema de Brauer não pode ser aplicado nem a A nem a A^T , porque em qualquer dos casos uma desigualdade não é satisfeita. No entanto, o primeiro teorema da quarta série pode ser aplicado, usando $p = 1/2$ (isto é, dando igual “peso” às linhas e às colunas), concluindo-se que A é invertível.