Análisis Numérico

Primer cuatrimestre 2022

Clases de aproximación de autovalores

Outline

Outline

1. Localización de autovalores

2. Aproximación de autovalores

Localización de autovalores

Teorema de Gershgorin

El conjunto de autovalores de una matriz $A \in \mathbb{C}^{n \times n}$ está contenido en la unión de n discos $D_i, i = 1, \ldots, n$, en el plano complejo, definidos por

$$D_i = \left\{ z : |z - a_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| \right\} \qquad \text{(discos de Gershgorin)}$$

Ejemplo

$$A = \begin{pmatrix} -1+i & 0 & \frac{1}{4} \\ \frac{1}{4} & 1 & \frac{1}{4} \\ 1 & 1 & 3 \end{pmatrix}$$

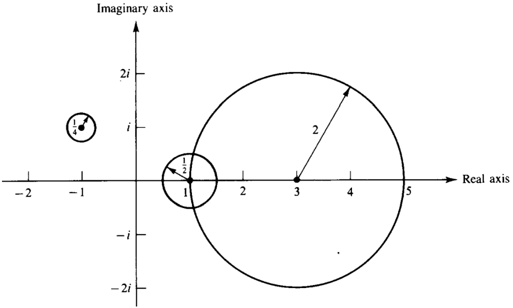

Figure: Discos de Gershgorim para A

Localización de autovalores

Demostración (Gershgorin):

- λ autovalor de A, x autovector con

$$Ax = \lambda x, \quad \|x\|_\infty = 1, \qquad |x_i| = \|x\|_\infty$$

Localización de autovalores

Demostración (Gershgorin):

- λ autovalor de A, x autovector con

$$Ax = \lambda x, \quad \|x\|_\infty = 1, \qquad |x_i| = \|x\|_\infty$$

- Como $(Ax)_i = \lambda x_i$ tenemos

$$\lambda x_i = \sum_{j=1}^{n} a_{ij} x_j$$

Localización de autovalores

Demostración (Gershgorin):

- λ autovalor de A, x autovector con

$$Ax = \lambda x, \quad \|x\|_\infty = 1, \qquad |x_i| = \|x\|_\infty$$

- Como $(Ax)_i = \lambda x_i$ tenemos

$$\lambda x_i = \sum_{j=1}^{n} a_{ij} x_j$$

- Entonces

$$(\lambda - a_{ii})\, x_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij} x_j$$

Localización de autovalores

Demostración (Gershgorin):

- λ autovalor de A, x autovector con

$$Ax = \lambda x, \quad \|x\|_\infty = 1, \qquad |x_i| = \|x\|_\infty$$

- Como $(Ax)_i = \lambda x_i$ tenemos

$$\lambda x_i = \sum_{j=1}^{n} a_{ij} x_j$$

- Entonces

$$(\lambda - a_{ii})\, x_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij} x_j \qquad \overset{|x_i|=1,\ |x_j| \leq 1}{\Longrightarrow} \qquad |\lambda - a_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|$$

Localización de autovalores

Demostración (Gershgorin):

- λ autovalor de A, x autovector con

$$Ax = \lambda x, \quad \|x\|_\infty = 1, \qquad |x_i| = \|x\|_\infty$$

- Como $(Ax)_i = \lambda x_i$ tenemos

$$\lambda x_i = \sum_{j=1}^{n} a_{ij} x_j$$

- Entonces

$$(\lambda - a_{ii})\, x_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij} x_j \qquad \overset{|x_i|=1,\ |x_j|\leq 1}{\Longrightarrow} \qquad |\lambda - a_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|$$

- Entonces $\lambda \in D_i$

Localización de autovalores

Observaciones:

Localización de autovalores

Observaciones:

- Como A y A^t tienen los mismos autovalores, también tenemos que los autovalores de A están en la unión de los discos E_j con

$$E_j = \left\{ z : |z - a_{jj}| \leq \sum_{\substack{i=1 \\ i \neq j}}^{n} |a_{ij}| \right\}, \quad j = 1, \dots, n$$

Localización de autovalores

Observaciones:

- Como A y A^t tienen los mismos autovalores, también tenemos que los autovalores de A están en la unión de los discos E_j con

$$E_j = \left\{ z : |z - a_{jj}| \leq \sum_{\substack{i=1 \\ i \neq j}}^{n} |a_{ij}| \right\}, \quad j = 1, \ldots, n$$

- Por lo tanto,

$$\{\lambda : \lambda \text{ autovalor de } A\} \subset (\cup_{i=i}^{n} D_i) \cap \left(\cup_{j=1}^{n} E_j\right)$$

Localización de autovalores

Observaciones:

- Como A y A^t tienen los mismos autovalores, también tenemos que los autovalores de A están en la unión de los discos E_j con

$$E_j = \left\{ z : |z - a_{jj}| \leq \sum_{\substack{i=1 \\ i \neq j}}^{n} |a_{ij}| \right\}, \quad j = 1, \ldots, n$$

- Por lo tanto,

$$\{\lambda : \lambda \text{ autovalor de } A\} \subset (\cup_{i=i}^{n} D_i) \cap \left(\cup_{j=1}^{n} E_j\right)$$

- Si uno de los discos D_i o E_i degenera a un punto $z = a_{ii}$, entonces a_{ii} es un autovalor de A. Es equivalente a decir que si la fila i o columna i de A contiene solo el elemento i-ésimo, a_{ii}, distinto de 0, entonces a_{ii} es autovalor de A

Outline

Método de potencias

Método de potencias

Objetivo: aproximar el autovalor de mayor valor absoluto de $A \in \mathbb{C}^{n \times n}$

Método de potencias

Objetivo: aproximar el autovalor de mayor valor absoluto de $A \in \mathbb{C}^{n\times n}$

Hipótesis

Método de potencias

Objetivo: aproximar el autovalor de mayor valor absoluto de $A \in \mathbb{C}^{n\times n}$

Hipótesis

- $\lambda_1, \ldots, \lambda_n$ autovalores de A con

$$|\lambda_1| > |\lambda_2| \geq \ldots \geq |\lambda_n|$$

notar: único autovalor dominante y simple

Método de potencias

Objetivo: aproximar el autovalor de mayor valor absoluto de $A \in \mathbb{C}^{n\times n}$

Hipótesis

- $\lambda_1, \ldots, \lambda_n$ autovalores de A con

$$|\lambda_1| > |\lambda_2| \geq \ldots \geq |\lambda_n|$$

notar: único autovalor dominante y simple

- Existe una base de autovectores:

$$Au^{(i)} = \lambda_i u^{(i)}, \qquad i = 1, \ldots, n$$

con $\left\{u^{(1)}, \ldots, u^{(n)}\right\}$ l.i.

Método de potencias

Objetivo: aproximar el autovalor de mayor valor absoluto de $A \in \mathbb{C}^{n \times n}$

Hipótesis

- $\lambda_1, \ldots, \lambda_n$ autovalores de A con

$$|\lambda_1| > |\lambda_2| \geq \ldots \geq |\lambda_n|$$

notar: único autovalor dominante y simple

- Existe una base de autovectores:

$$Au^{(i)} = \lambda_i u^{(i)}, \qquad i = 1, \ldots, n$$

con $\left\{u^{(1)}, \ldots, u^{(n)}\right\}$ l.i.

- tenemos una funcional lineal

$$\phi : \mathbb{C}^n \to \mathbb{C}$$

tal que $\phi(u^{(1)}) \neq 0$

Método de potencias

Objetivo: aproximar el autovalor de mayor valor absoluto de $A \in \mathbb{C}^{n\times n}$

Hipótesis

- $\lambda_1, \ldots, \lambda_n$ autovalores de A con

$$|\lambda_1| > |\lambda_2| \geq \ldots \geq |\lambda_n|$$

notar: único autovalor dominante y simple

- Existe una base de autovectores:

$$Au^{(i)} = \lambda_i u^{(i)}, \qquad i = 1, \ldots, n$$

con $\left\{u^{(1)}, \ldots, u^{(n)}\right\}$ l.i.

- tenemos una funcional lineal

$$\phi : \mathbb{C}^n \to \mathbb{C}$$

tal que $\phi(u^{(1)}) \neq 0$

Algoritmo

Método de potencias

Objetivo: aproximar el autovalor de mayor valor absoluto de $A \in \mathbb{C}^{n \times n}$

Hipótesis

- $\lambda_1, \ldots, \lambda_n$ autovalores de A con

$$|\lambda_1| > |\lambda_2| \geq \ldots \geq |\lambda_n|$$

notar: único autovalor dominante y simple

- Existe una base de autovectores:

$$Au^{(i)} = \lambda_i u^{(i)}, \qquad i = 1, \ldots, n$$

con $\left\{ u^{(1)}, \ldots, u^{(n)} \right\}$ l.i.

- tenemos una funcional lineal

$$\phi : \mathbb{C}^n \to \mathbb{C}$$

tal que $\phi(u^{(1)}) \neq 0$

Algoritmo

Dados
A: matriz $n \times n$
M: # iteraciones
$x^{(0)}$: aproximación inicial

Aproximación de autovalores

Método de potencias

Objetivo: aproximar el autovalor de mayor valor absoluto de $A \in \mathbb{C}^{n\times n}$

Hipótesis

- $\lambda_1, \ldots, \lambda_n$ autovalores de A con

$$|\lambda_1| > |\lambda_2| \geq \ldots \geq |\lambda_n|$$

notar: único autovalor dominante y simple

- Existe una base de autovectores:

$$Au^{(i)} = \lambda_i u^{(i)}, \qquad i = 1, \ldots, n$$

con $\left\{u^{(1)}, \ldots, u^{(n)}\right\}$ l.i.

- tenemos una funcional lineal

$$\phi : \mathbb{C}^n \to \mathbb{C}$$

tal que $\phi(u^{(1)}) \neq 0$

Algoritmo

Dados
A: matriz $n \times n$
M: # iteraciones
$x^{(0)}$: aproximación inicial

```
for i = 1 : M
  y = A * x;
  r = Phi(y) / Phi(x);
  x = y; %o: x=y/norm(y);
endfor
```

Aproximación de autovalores

Método de potencias

Objetivo: aproximar el autovalor de mayor valor absoluto de $A \in \mathbb{C}^{n\times n}$

Hipótesis

- $\lambda_1, \ldots, \lambda_n$ autovalores de A con

$$|\lambda_1| > |\lambda_2| \geq \ldots \geq |\lambda_n|$$

notar: único autovalor dominante y simple

- Existe una base de autovectores:

$$Au^{(i)} = \lambda_i u^{(i)}, \qquad i = 1, \ldots, n$$

con $\left\{u^{(1)}, \ldots, u^{(n)}\right\}$ l.i.

- tenemos una funcional lineal

$$\phi : \mathbb{C}^n \to \mathbb{C}$$

tal que $\phi(u^{(1)}) \neq 0$

Algoritmo

Dados
A: matriz $n \times n$
M: # iteraciones
$x^{(0)}$: aproximación inicial

```
for i = 1 : M
  y = A * x;
  r = Phi(y) / Phi(x);
  x = y; %o: x=y/norm(y);
endfor
```

r: aprox autovalor λ_1
x: aprox autovector $u^{(1)}$

Convergencia del método de potencias

Suponiendo las hipótesis anteriores, generamos una sucesión

$$x^{(0)} \text{ dado}, \quad x^{(k+1)} = Ax^{(k)}, \quad k = 1, 2, \ldots$$

con $x^{(0)} = a_1 u^{(1)} + a_2 u^{(2)} + \ldots + a_n u^{(n)}$ con $a_1 \neq 0$.
Si definimos

$$r_k = \frac{\phi\left(x^{(k+1)}\right)}{\phi\left(x^{(k)}\right)}$$

entonces tenemos

$$\lim_{k\to\infty} r_k = \lambda_1$$

Además $x^{(k)}/\|x^{(k)}\|$ aproxima a la dirección de $u^{(1)}$

Aproximación de autovalores

Convergencia del método de potencias

Suponiendo las hipótesis anteriores, generamos una sucesión

$$x^{(0)} \text{ dado}, \quad x^{(k+1)} = Ax^{(k)}, \quad k = 1, 2, \ldots$$

con $x^{(0)} = a_1 u^{(1)} + a_2 u^{(2)} + \ldots + a_n u^{(n)}$ con $a_1 \neq 0$.
Si definimos

$$r_k = \frac{\phi\left(x^{(k+1)}\right)}{\phi\left(x^{(k)}\right)}$$

entonces tenemos

$$\lim_{k \to \infty} r_k = \lambda_1$$

Además $x^{(k)}/\|x^{(k)}\|$ aproxima a la dirección de $u^{(1)}$

Nota: Como $\phi : \mathbb{C}^n \to \mathbb{C}$ podemos tomar por ejemplo la evaluación de la componente i-ésima: $\phi(x) = x_i$, suponiendo que $u_1^{(i)} \neq 0$.

Demostración (método de potencias):

Aproximación de autovalores

Demostración (método de potencias):

- Supongamos $x^{(0)} = a_1 u^{(1)} + a_2 u^{(2)} + \ldots + a_n u^{(n)}$ con $a_1 \neq 0$

Aproximación de autovalores

Demostración (método de potencias):

- Supongamos $x^{(0)} = a_1 u^{(1)} + a_2 u^{(2)} + \ldots + a_n u^{(n)}$ con $a_1 \neq 0$
- Entonces $x^{(k)} = A^k x^{(0)} = a_1 \lambda_1^k u^{(1)} + a_2 \lambda_2^k u^{(2)} + \ldots + a_n \lambda_n^k u^{(n)}$

Demostración (método de potencias):

- Supongamos $x^{(0)} = a_1 u^{(1)} + a_2 u^{(2)} + \ldots + a_n u^{(n)}$ con $a_1 \neq 0$
- Entonces $x^{(k)} = A^k x^{(0)} = a_1 \lambda_1^k u^{(1)} + a_2 \lambda_2^k u^{(2)} + \ldots + a_n \lambda_n^k u^{(n)}$
- Podemos reescribirlo como

$$x^{(k)} = \lambda_1^k \left[a_1 u^{(1)} + a_2 \frac{\lambda_2^k}{\lambda_1^k} u^{(2)} + \ldots + a_n \frac{\lambda_n^k}{\lambda_1^k} u^{(n)} \right] = \lambda_1^k \left[a_1 u^{(1)} + \varepsilon_k \right]$$

con $\varepsilon_k \to 0$ si $k \to \infty$ siendo que $|\lambda_i / \lambda_k| < 1$ para $i = 2, \ldots, n$

Demostración (método de potencias):

- Supongamos $x^{(0)} = a_1 u^{(1)} + a_2 u^{(2)} + \ldots + a_n u^{(n)}$ con $a_1 \neq 0$
- Entonces $x^{(k)} = A^k x^{(0)} = a_1 \lambda_1^k u^{(1)} + a_2 \lambda_2^k u^{(2)} + \ldots + a_n \lambda_n^k u^{(n)}$
- Podemos reescribirlo como

$$x^{(k)} = \lambda_1^k \left[a_1 u^{(1)} + a_2 \frac{\lambda_2^k}{\lambda_1^k} u^{(2)} + \ldots + a_n \frac{\lambda_n^k}{\lambda_1^k} u^{(n)} \right] = \lambda_1^k \left[a_1 u^{(1)} + \varepsilon_k \right]$$

con $\varepsilon_k \to 0$ si $k \to \infty$ siendo que $|\lambda_i / \lambda_k| < 1$ para $i = 2, \ldots, n$

- Como ϕ es lineal y continuo, tenemos

$$\frac{\phi(x^{k+1})}{\phi(x^k)} = \frac{\phi \left\{ \lambda_1^{k+1} \left[a_1 u^{(1)} + \varepsilon_{k+1} \right] \right\}}{\phi \left\{ \lambda_1^k \left[a_1 u^{(1)} + \varepsilon_k \right] \right\}} = \lambda_1 \frac{\phi \left[a_1 u^{(1)} + \varepsilon_{k+1} \right]}{\phi \left[a_1 u^{(1)} + \varepsilon_k \right]} \longrightarrow \lambda_1$$

donde usamos que $a_1 \neq 0$ y $\phi(u^{(1)}) \neq 0$

Aproximación de autovalores

Demostración (método de potencias):

- Supongamos $x^{(0)} = a_1 u^{(1)} + a_2 u^{(2)} + \ldots + a_n u^{(n)}$ con $a_1 \neq 0$
- Entonces $x^{(k)} = A^k x^{(0)} = a_1 \lambda_1^k u^{(1)} + a_2 \lambda_2^k u^{(2)} + \ldots + a_n \lambda_n^k u^{(n)}$
- Podemos reescribirlo como

$$x^{(k)} = \lambda_1^k \left[a_1 u^{(1)} + a_2 \frac{\lambda_2^k}{\lambda_1^k} u^{(2)} + \ldots + a_n \frac{\lambda_n^k}{\lambda_1^k} u^{(n)} \right] = \lambda_1^k \left[a_1 u^{(1)} + \varepsilon_k \right]$$

  con $\varepsilon_k \to 0$ si $k \to \infty$ siendo que $|\lambda_i / \lambda_k| < 1$ para $i = 2, \ldots, n$

- Como ϕ es lineal y continuo, tenemos

$$\frac{\phi(x^{k+1})}{\phi(x^k)} = \frac{\phi \left\{ \lambda_1^{k+1} \left[a_1 u^{(1)} + \varepsilon_{k+1} \right] \right\}}{\phi \left\{ \lambda_1^k \left[a_1 u^{(1)} + \varepsilon_k \right] \right\}} = \lambda_1 \frac{\phi \left[a_1 u^{(1)} + \varepsilon_{k+1} \right]}{\phi \left[a_1 u^{(1)} + \varepsilon_k \right]} \longrightarrow \lambda_1$$

  donde usamos que $a_1 \neq 0$ y $\phi(u^{(1)}) \neq 0$

- Notemos que $\frac{x^{(k)}}{\lambda_1^k} \to a_1 u^{(1)}$

Método de potencias inverso

Método de potencias inverso

- La idea es aplicar el método de potencias a A^{-1}, así aproximamos el autovalor de mayor valor absoluto de A^{-1}, que es (usando la notación anterior) λ_n^{-1}, el inverso del autovalor de menor valor absoluto de A

Método de potencias inverso

- La idea es aplicar el método de potencias a A^{-1}, así aproximamos el autovalor de mayor valor absoluto de A^{-1}, que es (usando la notación anterior) λ_n^{-1}, el inverso del autovalor de menor valor absoluto de A
- Se asume que existe un único autovalor simple de menor valor absoluto, que no es nulo

Método de potencias inverso

- La idea es aplicar el método de potencias a A^{-1}, así aproximamos el autovalor de mayor valor absoluto de A^{-1}, que es (usando la notación anterior) λ_n^{-1}, el inverso del autovalor de menor valor absoluto de A
- Se asume que existe un único autovalor simple de menor valor absoluto, que no es nulo
- Es posible evitar invertir A

Aproximación de autovalores

Método de potencias inverso

- La idea es aplicar el método de potencias a A^{-1}, así aproximamos el autovalor de mayor valor absoluto de A^{-1}, que es (usando la notación anterior) λ_n^{-1}, el inverso del autovalor de menor valor absoluto de A
- Se asume que existe un único autovalor simple de menor valor absoluto, que no es nulo
- Es posible evitar invertir A

```
for i = 1 : M
  y = A^(-1) * x;
  r = Phi(y) / Phi(x);
  x = y; %o:x=y/norm(y);
endfor
```

No invertir A $\longrightarrow$

```
for i = 1 : M
  y = A \ x;
  r = Phi(y) / Phi(x);
  x = y; %o:x=y/norm(y);
endfor
```

Método de potencias inverso

- La idea es aplicar el método de potencias a A^{-1}, así aproximamos el autovalor de mayor valor absoluto de A^{-1}, que es (usando la notación anterior) λ_n^{-1}, el inverso del autovalor de menor valor absoluto de A
- Se asume que existe un único autovalor simple de menor valor absoluto, que no es nulo
- Es posible evitar invertir A

```
for i = 1 : M
  y = A^(-1) * x;
  r = Phi(y) / Phi(x);
  x = y; %o:x=y/norm(y);
endfor
```

No invertir A $\longrightarrow$

```
for i = 1 : M
  y = A \ x;
  r = Phi(y) / Phi(x);
  x = y; %o:x=y/norm(y);
endfor
```

- Si A tiene una descomposición LU, solo se necesita calcularla una vez, y usar solo *backward substitution* en cada paso del algoritmo

Aproximación de autovalores

Método de potencias inverso

- La idea es aplicar el método de potencias a A^{-1}, así aproximamos el autovalor de mayor valor absoluto de A^{-1}, que es (usando la notación anterior) λ_n^{-1}, el inverso del autovalor de menor valor absoluto de A
- Se asume que existe un único autovalor simple de menor valor absoluto, que no es nulo
- Es posible evitar invertir A

```
for i = 1 : M
  y = A^(-1) * x;
  r = Phi(y) / Phi(x);
  x = y; %o:x=y/norm(y);
endfor
```

No invertir A $\longrightarrow$

```
for i = 1 : M
  y = A \ x;
  r = Phi(y) / Phi(x);
  x = y; %o:x=y/norm(y);
endfor
```

- Si A tiene una descomposición LU, solo se necesita calcularla una vez, y usar solo *backward substitution* en cada paso del algoritmo
- La aproximación de λ_n es $\frac{1}{r}$

Shifted inverse power method

Shifted inverse power method

- La idea es hallar el autovalor más cercano a un complejo dado μ (se supone que este es único y simple)

Aproximación de autovalores

Shifted inverse power method

- La idea es hallar el autovalor más cercano a un complejo dado μ (se supone que este es único y simple)
- Esto equivale a hallar el autovalor de menor valor absoluto de $A - \mu I$, por lo que podemos aplicar el método de potencias inversa a esta matriz (shifted matrix)

Shifted inverse power method

- La idea es hallar el autovalor más cercano a un complejo dado μ (se supone que este es único y simple)
- Esto equivale a hallar el autovalor de menor valor absoluto de $A - \mu I$, por lo que podemos aplicar el método de potencias inversa a esta matriz (shifted matrix)
- Algoritmo:

```
for i = 1 : M
  y = (A-mu) \ x;
  r = Phi(y) / Phi(x);
  x = y; %o: x=y/norm(y);
endfor
```

Aproximación de autovalores

Shifted inverse power method

- La idea es hallar el autovalor más cercano a un complejo dado μ (se supone que este es único y simple)
- Esto equivale a hallar el autovalor de menor valor absoluto de $A - \mu I$, por lo que podemos aplicar el método de potencias inversa a esta matriz (shifted matrix)
- Algoritmo:

```
for i = 1 : M
  y = (A-mu) \ x;
  r = Phi(y) / Phi(x);
  x = y; %o: x=y/norm(y);
endfor
```

- si λ es el autovalor buscado, ahora r da una aproximación de $\frac{1}{\lambda-\mu}$, por lo tanto una aproximación de λ es

$$\lambda \sim \frac{1}{r} + \mu$$

Proceso de deflación

Proceso de deflación

Objetivo: obtener aproximaciones de todos los autovalores de una matriz

Proceso de deflación

Objetivo: obtener aproximaciones de todos los autovalores de una matriz

- Se supone que los autovalores son simples y de módulos distintos

Proceso de deflación

Objetivo: obtener aproximaciones de todos los autovalores de una matriz

- Se supone que los autovalores son simples y de módulos distintos
- Idea: obtener aproximaciones de un autovalor de A y su correspondiente autovector, y con ellos construir una matriz de una dimensión menor, que tenga como autovalores los restantes autovalores de A

Proceso de deflación

Objetivo: obtener aproximaciones de todos los autovalores de una matriz

- Se supone que los autovalores son simples y de módulos distintos
- Idea: obtener aproximaciones de un autovalor de A y su correspondiente autovector, y con ellos construir una matriz de una dimensión menor, que tenga como autovalores los restantes autovalores de A
- Detalles: ver ejercicio 6 de la práctica 4