Solución del problema de valores propios

FCEN
FACULTAD DE CIENCIAS EXACTAS Y NATURALES

Temario

- Repaso problema de autovalores
- Métodos numéricos para el problema de autovalores
- Método de la potencia
- Escalamiento
- Algoritmo y programa de ejemplo
- Ejemplos
- Método de la potencia inversa
- Método de Hotelling

Problema de autovalores

Sea **A** una matríz de NxN y **X** un vector de Nx1.

El producto **Y=A X** se puede ver como una transformación lineal del espacio N dimensional en sí mismo y se buscan escalares λ para los que exista un vector no nulo **X** tales que **A X**=λ**X**

Cuando esto ocurre se dice que **X** es un autovector correspondiente al autovalor λ.

La ecuación anterior se puede expresar como un sistema de ecuaciones lineal $(A-\lambda I)X=0$

En este problema, este sistema tiene soluciones no triviales sí y solo si el determinante de $(A-\lambda I)$ es nulo.

La solución del determinante nulo da origen a un **polinomio característico de grado N,** cuyas raíces son los autovalores de la matriz **A**.

Soluciones numéricas para el problema de autovalores

Los métodos disponibles se pueden agrupar en tres:

- Métodos de solución del polinomio característico
- Métodos de transformación
- Métodos iterativos

Encontrar las **raíces de polinomios** de forma **analítica** es posible **únicamente hasta grado 4**. Se necesitan métodos numéricos para resolver matrices mayores a 4x4.

En los métodos de transformación se **aplican rotaciones y traslaciones** para **transformar la matriz** en otra diagonal **semejante**, de manera que la obtención de los autovalores sea más simple.

Un método iterativo es aquél que produce una **secuencia de aproximaciones** que convergen a la solución del problema.

Métodos iterativos

FCEN
FACULTAD DE CIENCIAS EXACTAS Y NATURALES

En principio la cantidad de aproximaciones necesarias es infinita. Además del método **se debe definir un criterio de convergencia** para aceptar una aproximación como la solución.

Para evaluar el **error** en cada iteración **se utiliza una norma**.

Además de asegurar la convergencia del método **es necesario determinar la tasa de la convergencia**. No es útil disponer de un método que alcance la convergencia en millones de iteraciones.

Existen **varios métodos** iterativos. En los más simples se producen secuencias de vectores que convergen a los autovectores de la matriz. Otros métodos más avanzados producen secuencias de matrices que poseen los mismos autovalores de la matriz original y convergen a una forma más simple de la matriz.

Método de la potencia fundamento

Sea **A** una matriz de NxN y que posee un conjunto de N autovectores **V** linealmente independientes que a su vez forman una base.

Sean $\lambda_1, \lambda_2, \ldots, \lambda_n$ los autovalores asociados a los autovectores de **A**.

Suponemos que los autovalores están ordenados. $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_n|$

Si $|\lambda_1| > |\lambda_2|$ entonces se denomina autovalor dominante.

Desarrollo del método

El método consiste en **generar una secuencia de vectores** que se aproximen al autovector asociado al autovalor dominante.

Como los autovectores **V** de **A** forman una base, se puede expresar el vector propuesto **X** como una combinación lineal de la base. $X = c_1 \mathbf{V_1} + c_2 \mathbf{V_2} + \ldots + c_n \mathbf{V_n}$

Premultiplicando por la matriz: $AX = c_1 A \mathbf{V_1} + c_2 A \mathbf{V_2} + \ldots + c_n A \mathbf{V_n}$

Reemplazando por el problema de autovalores: $AX = c_1 \lambda_1 \mathbf{V_1} + c_2 \lambda_2 \mathbf{V_2} + \ldots + c_n \lambda_n \mathbf{V_n}$

Método de la potencia
fundamento

Premultiplicando j veces: $(A)^j X = c_1 \lambda_1^j \mathbf{V_1} + c_2 \lambda_2^j \mathbf{V_2} + \ldots + c_n \lambda_n^j \mathbf{V_n}$

Sacando factor común: $(A)^j X = c_1 \lambda_1^j \left(\mathbf{V_1} + c_2 \left(\frac{\lambda_2}{c_1 \lambda_1} \right)^j \mathbf{V_2} + \ldots + c_n \left(\frac{\lambda_n}{c_1 \lambda_1} \right)^j \mathbf{V_n} \right)$

Como λ_1 es el autovalor dominante los cocientes dentro del paréntesis son menores a 1, por lo tanto, al premultiplicar repetidas veces todos los términos salvo el primero tienden a cero.

$$\lim_{j \to \infty} (A)^j X = c_1 \lambda_1^j \mathbf{V_1}$$

Geométricamente se observa que el vector solución es paralelo al autovector asociado al autovalor dominante.

Método de la potencia fundamento

¿Cómo se obtiene el autovalor asociado?

Dividiendo dos iteraciones sucesivas se obtiene

$$\frac{(A)^{(j+1)}X}{(A)^{(j)}X}=\frac{\lambda_1^{(j+1)}\left(c_1\mathbf{V_1}+c_2\left(\frac{\lambda_2}{\lambda_1}\right)^{(j+1)}\mathbf{V_2}+\ldots+c_n\left(\frac{\lambda_n}{\lambda_1}\right)^{(j+1)}\mathbf{V}_n\right)}{\lambda_1^{j}\left(c_1\mathbf{V_1}+c_2\left(\frac{\lambda_2}{\lambda_1}\right)^{j}\mathbf{V_2}+\ldots+c_n\left(\frac{\lambda_n}{\lambda_1}\right)^{j}\mathbf{V}_n\right)}$$

A medida que se produce la convergencia, los cocientes de autovalores tienden a cero y se obtiene:

$$\alpha^{(j+1)}=\frac{\lambda_1^{(j+1)}\left(c_1\mathbf{V_1}\right)}{\lambda_1^{j}\left(c_1\mathbf{V_1}\right)}=\lambda_1$$

Método de la potencia escalamiento

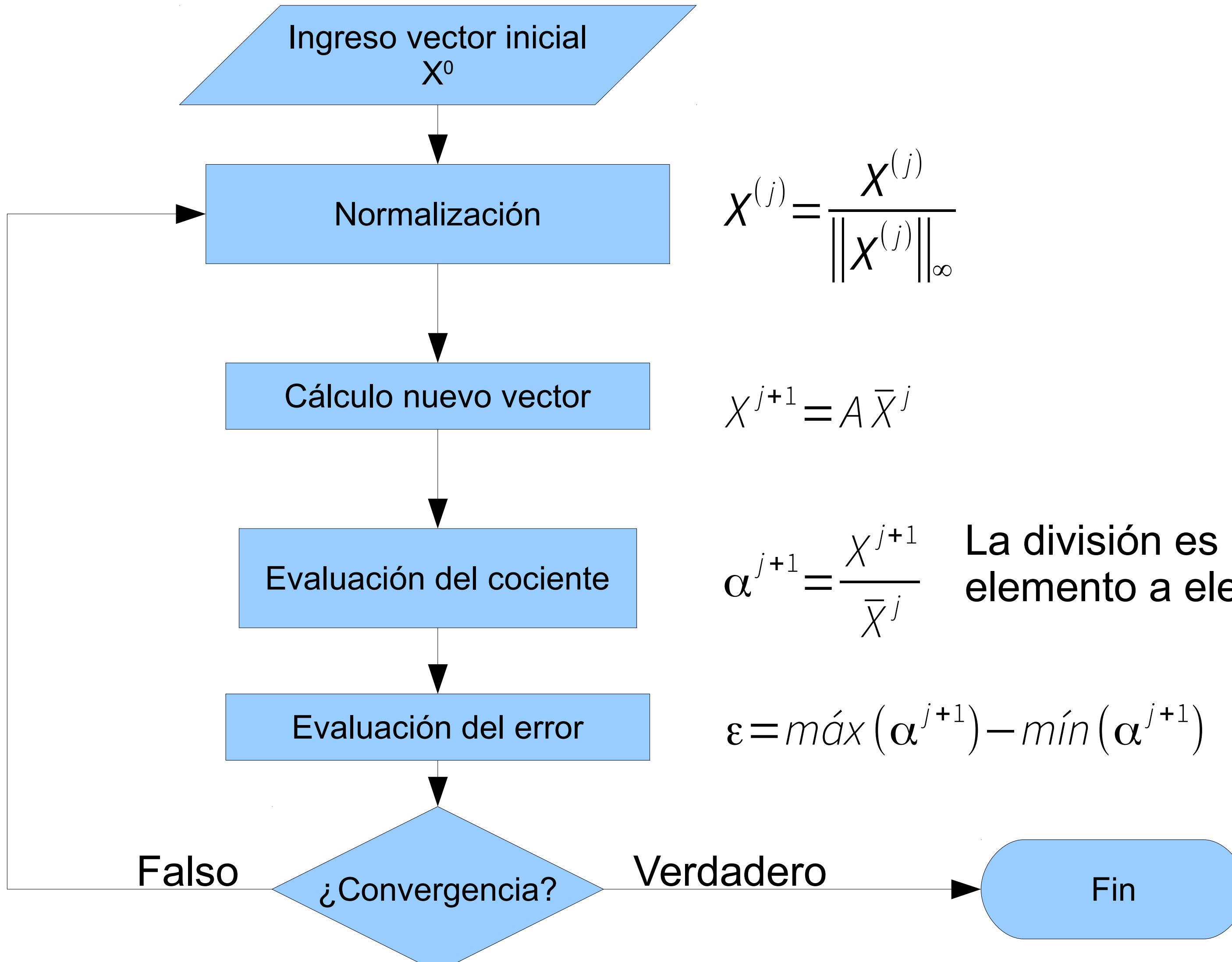
A medida que el método evoluciona, el vector calculado va aumentando o reduciendo su magnitud, según si el autovalor dominante es mayor o menor a 1.

Si se multiplica demasiadas veces, el valor del vector puede generar desbordamiento numérico. Para evitarlo se utiliza una "normalización" o "escalamiento", es decir

$$X^{(j)} = \frac{X^{(j)}}{\left\| X^{(j)} \right\|_{\infty}}$$

Se suele elegir la norma infinito ya que asegura que la secuencia converge a un vector cuya máxima componente valdrá 1.

Resumen método de la potencia



$$X^{(j)} = \frac{X^{(j)}}{\left\| X^{(j)} \right\|_\infty}$$

$$X^{j+1} = A\bar{X}^{j}$$

$$\alpha^{j+1} = \frac{X^{j+1}}{\bar{X}^{j}}$$

La división es elemento a elemento

$$\varepsilon = \text{máx}\left(\alpha^{j+1}\right) - \text{mín}\left(\alpha^{j+1}\right)$$

Programa

```
function [lambda,V]=power1(A,X0,tol,maxiter)
k=0;
err=1;
while (k<=maxiter)&&(err>tol)
   Xnorm=X0/max(abs(X0)); %escalamiento
   Y=A*Xnorm; %nuevo valor
   alfa=Y./Xnorm; %calculo de autovalores
   err=abs(max(alfa)-min(alfa)); %calculo del error
   X0=Y; %actualizacion
endwhile
lambda=mean(alfa);
V=Xnorm;
endfunction
```

Ejemplo

FCEN FACULTAD DE CIENCIAS EXACTAS Y NATURALES

Dada la matriz $\mathbf{A} = \begin{bmatrix} 9 & 1 \\ 1 & 2 \end{bmatrix}$ se desean obtener los autovectores de la matriz.

1) Polinomio característico: $\det|\mathbf{A} - \lambda \mathbf{I}| = \mathbf{0}$

$$\det\left|\begin{bmatrix} 9 & 1 \\ 1 & 2 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right| = \det\left|\begin{bmatrix} 9-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix}\right| = (9-\lambda)(2-\lambda) - 1 = 17 - 11\lambda + \lambda^2 = 0$$

Las raíces del polinomio son: 9,140055 y 1,859945. La tasa de convergencia es $\left|\frac{\lambda_2}{\lambda_1}\right|$=**0,203**

El autovector asociado al autovalor dominante es: $(\mathbf{A} - \lambda_1 \mathbf{I})\mathbf{v}_1 = 0$

$$\begin{bmatrix} -0{,}14006 & 1 \\ 1 & -7{,}14006 \end{bmatrix} \begin{pmatrix} v_{11} \\ v_{12} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \rightarrow \mathrm{v}_1 = \begin{pmatrix} 1 \\ 0{,}14006 \end{pmatrix}$$

El autovector asociado al segundo autovalor es: $(\mathbf{A} - \lambda_2 \mathbf{I})\mathbf{v}_2 = 0$

$$\begin{bmatrix} 7{,}14006 & 1 \\ 1 & 0{,}14006 \end{bmatrix} \begin{pmatrix} v_{11} \\ v_{12} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \rightarrow \mathrm{v}_1 = \begin{pmatrix} 1 \\ -7{,}14006 \end{pmatrix}$$

Ejemplo

Dada la matriz $\mathbf{A} = \begin{bmatrix} 9 & 1 \\ 1 & 2 \end{bmatrix}$ se desea obtener el autovector dominante de la matriz.

2) Método de la potencia **sin escalamiento**.

Iteración	q^j	$\lVert q^j \rVert$	$\overline{q^j}$	$\overline{q^{j+1}}$	α^{j+1}	error
0	[1 1]	1	[1 1]	[10 3]	[10 3]	7
1	[10 3]	1	[10 3]	[93 16]	[9,33 5,33]	3,967
2	[93 16]	1	[93 16]	[853 125]	[9,172 7,8125]	1,3595
3	[853 125]	1	[853 125]	[7802 1103]	[9,1465 8,824]	0,323
4	[7802 1103]	1	[7802 1103]	[71321 10008]	[9,1414 9,073]	0,068
5	[71321 10008]	1	[71321 10008]	[651897 91337]	[9,1403 9,1264]	0,014

Analítico: [1 0,14006]

Analítico: 9,140055

Ejemplo

FCEN
FACULTAD DE CIENCIAS EXACTAS Y NATURALES

Dada la matriz $\mathbf{A} = \begin{bmatrix} 9 & 1 \\ 1 & 2 \end{bmatrix}$ se desea obtener el autovector dominante de la matriz.

2) Método de la potencia **con escalamiento**.

Iteración	q^j	$\lVert q^j \rVert$	$\overline{q^j}$	$\overline{q^{j+1}}$	α^{j+1}	error
0	[1 1]	1	[1 1]	[10 3]	[10 3]	7
1	[10 3]	10	[1 0,3]	[9,3 1,6]	[9,3 5,33]	3,967
2	[9,3 1,6]	9,3	[1 0,172]	[9,172 1,344]	[9,172 7,813]	1,359
3	[9,172 1,344]	9,172	[1 0,1465]	[9,1465 1,293]	[9,1465 8,824]	0,3225
4	[9,1465 1,293]	9,1465	[1 0,14137]	[9,1414 1,2828]	[9,1414 9,0734]	0,0679
5	[9,1414 1,2828]	9,1414	[1 0,14032]	[9,1403 1,2806]	[9,1403 9,1264]	0,0139

Analítico: [1 0,14006]

Analítico: 9,140055

Ejemplo

FCEN
FACULTAD DE CIENCIAS EXACTAS Y NATURALES

Dada la matriz $\mathbf{A} = \begin{bmatrix} 9 & 1 \\ 1 & 2 \end{bmatrix}$ se desea obtener el autovector dominante de la matriz.

2) Método de la potencia. **Inicio con el segundo autovector**

Iteración	q^j	$\lVert q^j \rVert$	$\overline{q^j}$	$\overline{q^{j+1}}$	α^{j+1}	error
0	[1 -7,14006]	7,14006	[0,14005 -1]	[0,26049 -1,8599]	[1,85994 1,859945]	-5E-6

Analítico: [1 0,14006]

Analítico: 9,140055

¿Por qué no converge al dominante?

Al ser el vector inicial paralelo al segundo autovector, y los autovectores forman una base, el vector es ortogonal al primer autovector y siempre da 0 la componente c_1

Método iterativo de la potencia inversa

Mediante el método de la potencia se obtiene una aproximación al autovector y autovalor dominantes asociados a la matriz A.

Si la matriz no es singular, se puede utilizar la inversa para generar la secuencia de vectores, lo cual arrojará como resultado la inversa del mínimo autovalor de la matriz A.

Observación: el proceso es idéntico al del método de la potencia, sin embargo, la inversión de la matriz A no es eficiente del punto de vista numérico. Por lo tanto conviene resolver un SEL mediante la factorización LU.

$X^{j+1}=A^{-1}X^{j}\rightarrow$ Premultiplicando por A: $AX^{j+1}=AA^{-1}X^{j}\rightarrow AX^{j+1}=X^{j}$

$(LU)X^{j+1}=L(UX^{j+1})=X^{j}\rightarrow Ly=X^{j}$ siendo $UX^{j+1}=y$

De esta forma, se determinan las matrices L y U una única vez al principio y luego se resuelven sistemas triangulares, actualizando el valor del término independiente.

Ejemplo 4

Dada la matriz $\mathbf{A} = \begin{bmatrix} 9 & 1 \\ 1 & 2 \end{bmatrix}$ se desea obtener el mínimo autovector de la matriz.

Método de la **potencia inversa**.

Iteració n	q^j	$\|q^j\|$	$\overline{q^j}$	$\overline{q^{j+1}}$	α^{j+1}	error
0	[1 1]	1	[1 1]	[0,0588 0,4706]	[0,0588 0,4706]	-0,412
1	[0,0588 0,4706]	0,4706	[0,125 1]	[-0,0441 0,522]	[0,3529 0,522]	-0,875
2	[-0,0441 0,522]	0,522	[-0,0845 1]	[-0,06876 0,5344]	[0,8137 0,5344]	0,279
3	[-0,06876 0,5344]	0,5344	[-0,1287 1]	[-0,07396 0,53698]	[0,5748 0,53698]	0,0378
4	[-0,07396 0,53698]	0,53698	[-0,1377 1]	[-0,07502 0,53751]	[0,5447 0,5375]	0,0072
			[1 -7,262]		1,848	

Analítico: [1 -7,14006]

Analítico: 1,859945

Método para encontrar otros autovalores

Un método para encontrar los demás autovalores consiste en, primero eliminar del sistema original el autovalor dominante que ya se conoce, y luego repetir el método de la potencia para encontrar el siguiente autovalor dominante. Es una estrategia similar a la deflación de polinomios y se denomina **Método de Hotelling.**

Se remueve el autovalor calculado de A y se obtiene $B = A - \lambda_1 V_1 V_1^T$

Se resuelve el nuevo problema de autovalores $B V = \lambda V$

donde el autovalor dominante de A ahora vale cero

Idealmente, la deflación permite obtener todos los autovalores de la matriz A, sin embargo, la aritmética de punto flotante y el error de convergencia del proceso iterativo, van degradando la solución ya que el autovalor "conocido" en realidad no es exacto. Al eliminar el autovalor aproximado de la matriz, no se elimina de forma exacta y quedan residuos.

Por lo tanto, se pueden determinar solamente algunos de los autovalores restantes, hasta que la precisión numérica se degrada por completo.