AUTOVALORES Y AUTOVECTORES

Claudio Careglio

Definición del problema

- Problema de valores y vectores propios:
  - Sistema de ecuaciones puede escribirse de la forma:

$$(\mathbf{A}-\lambda\, \mathbf{I})\mathbf{x}=\mathbf{0} \qquad (1)$$

$$\mathbf{A}\mathbf{x}= \lambda\, \mathbf{x} \qquad (1')$$

  - Interesan soluciones distintas de la trivial (**x** distinto de 0).
  - Se denominan autovalores λ y autovectores **x**, a los números y vectores no nulos respectivamente que son solución del sistema de ecuaciones (1) (*Problema estándar de autovalores*).
  - Interpretación: λ es un escalar que cambia el módulo del vector **x** cuya dirección permanece invariante.
  - Alternativamente, se tiene el denominado problema *generalizado de autovalores y autovectores*:

$$(\mathbf{K}-\lambda\, \mathbf{M})\mathbf{x}=\mathbf{0} \qquad (2)$$

Definición del problema

- Problema de valores y vectores propios:
  - Considerando el sistema de ecuaciones:

$$(\mathbf{A}-\lambda\ \mathbf{I})\mathbf{x}=\mathbf{0}$$

    y considerando que **x** es distinto de 0, entonces.

$$\det(\mathbf{A}-\lambda\ \mathbf{I})=0$$

  - Del determinante se obtiene un polinomio en función de λ de grado n.
  - Las raíces de este polinomio ("polinomio característico") son los valores propios.
  - Por cada valor propio existe al menos una dirección invariante dada por el autovector **x**.

Métodos para determinar valores y vectores propios

- **Resolución del polinomio característico**
  - Encontrar las raíces del polinomio característico
  - Resolver el sistema homogéneo $(\mathbf{A}-\lambda\,\mathbf{I})\mathbf{x}=\mathbf{0}$ para cada valor propio obtenido como raíz del polinomio característico.
  - A medida que el orden n del sistema crece, crece el orden del polinomio y la dificultad para encontrar raíces.
  - Ejemplo:

    Calcular los autovalores y autovectores de la matriz:

$$\begin{bmatrix} 3 & -1 \\ -1 & 2 \end{bmatrix}$$

Métodos para determinar valores y vectores propios

- **Tranformación**
  - Transforman la matriz de coeficientes A en una matriz diagonal.
  - Encuentran la totalidad de los valores y vectores propios del sistema.

Métodos para determinar valores y vectores propios

- **Iterativos**
  - Aproximar sucesivamente los valores y vectores propios.
  - Convergen a un valor y vector propio.
  - Para encontrar más de un vector y valor propio: técnicas de deflación (eliminar de los procesos iterativos los valores y vectores propios que ya se conocen).
  - Método de la Potencia: es un algoritmo iterativo que encuentra la dirección invariante, asociada al autovalor dominante.
    - Sin escalamiento
    - Con escalamiento
  - Método de la Potencia Inversa:
    - Sin escalamiento
    - Con escalamiento

Método de la Potencia

Dada una matriz **A,** se buscan las direcciones invariantes, soluciones de

$(\mathbf{A} - \lambda \cdot \mathbf{I})\underline{x} = \underline{0}$

o bien:

$$\mathbf{A} \cdot \underline{x} = \lambda \cdot \underline{x}$$

Esto es equivalente a encontrar un vector $\underline{y}$ que se puede obtener como:

$$\underline{y} = \mathbf{A} \cdot \underline{x}$$

y resulta igual a:

$$\underline{y} = \lambda \cdot \underline{x}$$

lo que significa que los vectores $\underline{x}$ e $\underline{y}$ son PARALELOS

El **método de la Potencia** es un algoritmo iterativo que propone un $\underline{y_0}$ arbitrario, luego:

a) Obtiene un nuevo vector como:

$$\underline{y_{k+1}} = \mathbf{A} \cdot \underline{y_k}$$

b) Sólo se detiene si se cumple que:

$\underline{y_{k+1}}$ es paralelo a $\underline{y_k}$

Método de la Potencia

- **Puede considerarse como norma alguna de las siguientes:**
  - Norma infinito,
  - Norma cuadrática,
  - “Pseudonorma” (Nota: si bien la asignación a cada vector de su primera componente no es una norma, en la práctica se puede utilizar la primera componente de cada vector en lugar de una norma determinada, en el proceso iterativo y su implementación computacional).

Método de la Potencia

- **Convergencia:**

Dada una matriz **A** ε $\mathbf{R}^{\mathbf{NxN}}$**,** tal que sus autovectores $\underline{v}_1$, $\underline{v}_2$, …$\underline{v}_N$, y sus autovalores

$\lambda_1, \lambda_2, \lambda_3, \ldots.\ \lambda_N$ verifican que:

$\mathbf{A} \,.\, \underline{v}_k = \lambda_k \,.\, \underline{v}_k$ con k=1 a N

Es diagonalizable: es decir tiene N autovectores linealmente independientes que forman base de R^N.

Es posible ordenar los autovalores de mayor a menor considerando sus valores absolutos en la forma:

$$|\lambda_1| > |\lambda_2| \quad > |\lambda_3| > \ldots\ldots\ldots\ldots |\lambda_N|$$

Entonces el algoritmo del método de la potencia converge.

- Demostración.

Método de la Potencia

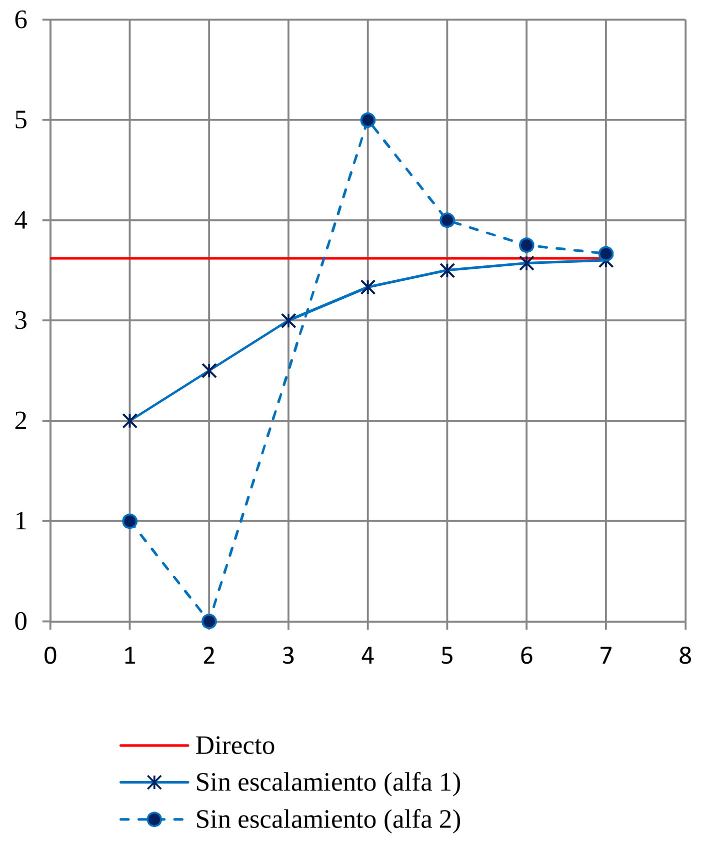
- **Ejemplo**

Datos:

A		y_0	Tolerancia
3	-1	1	0,01
-1	2	1	

Método de la Potencia (sin escalamiento): $y_{k+1} = \mathbf{A} \cdot y_k$

k	y_{k+1}		α	Error	¿Iterar?
0	y_1=	2	2	-	-
		1	1		
1	y_2=	5	2,5	0,5	Continuar
		0	0	-1	Continuar
2	y_3=	15	3	0,5	Continuar
		-5	-	-	-
3	y_4=	50	3,3333333	0,3333333	Continuar
		-25	5	-	-
4	y_5=	175	3,5	0,1666667	Continuar
		-100	4	-1	Continuar
5	y6=	625	3,5714286	0,0714286	Continuar
		-375	3,75	-0,25	Continuar
6	y7=	2250	3,6	0,0285714	Continuar
		-1375	3,6666667	-0,0833333	Continuar



Método de la Potencia con escalamiento

Dada una matriz **A,** se buscan las *Direcciones Invariantes*, soluciones de

$$(\mathbf{A}-\lambda\cdot\mathbf{I})\underline{x}=\underline{0}$$

o bien:

$$\mathbf{A}\cdot\underline{x}=\lambda\cdot\underline{x}$$

El **método de la Potencia con escalamiento** es un algoritmo iterativo que propone un $\underline{y_0}$ arbitrario y se *escala* para obtener el versor $\underline{x_0}$, con el que se inicia el proceso iterativo, luego:

a) Obtener un nuevo vector como $\underline{y_{k+1}}=\mathbf{A}\cdot\underline{x_k}$

b) Sólo se detiene si se cumple que $\underline{y_{k+1}}$ es paralelo a $\underline{x_k}$

Método de la Potencia con escalamiento

- **Ejemplo:**

Datos:

A		y_0	x_0	Tolerancia
3	-1	1	1	0,01
-1	2	1	1	

Método de la Potencia (con escalamiento): $\underline{y_{k+1}} = \mathbf{A} \cdot \underline{x_k}$

k		y_{k+1}		x_{k+1}	α	Error	¿Iterar?
0	y_1=	2	x_1=	1	2	-	-
		1		0,5	1		
1	y_2=	2,5	x_2=	1	2,5	0,5	Continuar
		0		0	0	-1	Continuar
2	y_3=	3	x_3=	1	3	0,5	Continuar
		-1		-0,3333	-	-	-
3	y_4=	3,3333	x_4=	1	3,3333333	0,3333333	Continuar
		-1,6667		-0,5	5	-	-
4	y_5=	3,5	x_5=	1	3,5	0,1666667	Continuar
		-2		-0,5714	4	-1	Continuar
5	y6=	3,5714	x6=	1	3,5714286	0,0714286	Continuar
		-2,1429		-0,6	3,75	-0,25	Continuar
6	y7=	3,6	x7=	1	3,6	0,0285714	Continuar
		-2,2		-0,6111	3,6666667	-0,0833333	Continuar

Ejemplo: Método de la Potencia sin escalamiento vs. escalamiento

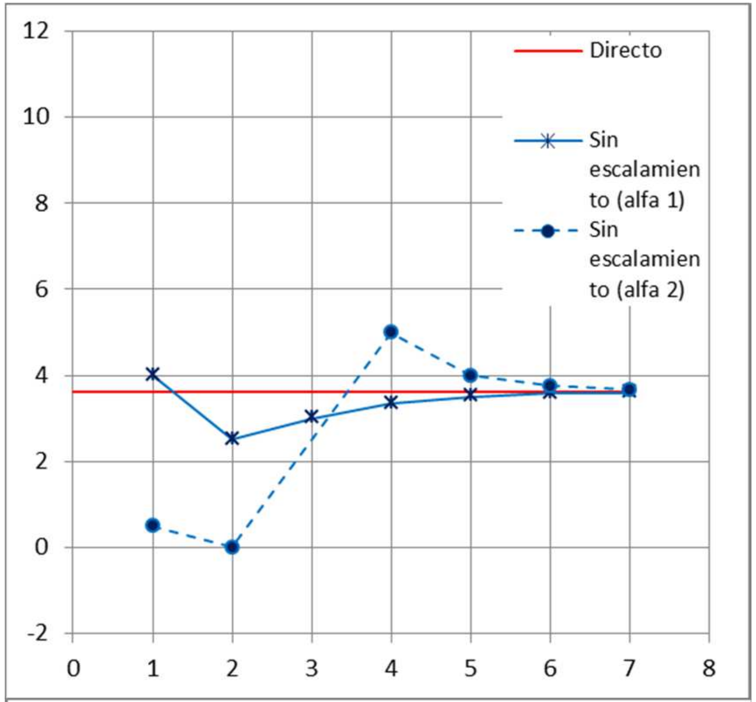
A		y_0	x_0	Tolerancia
3	-1	1	0,5	0,01
-1	2	1	2	

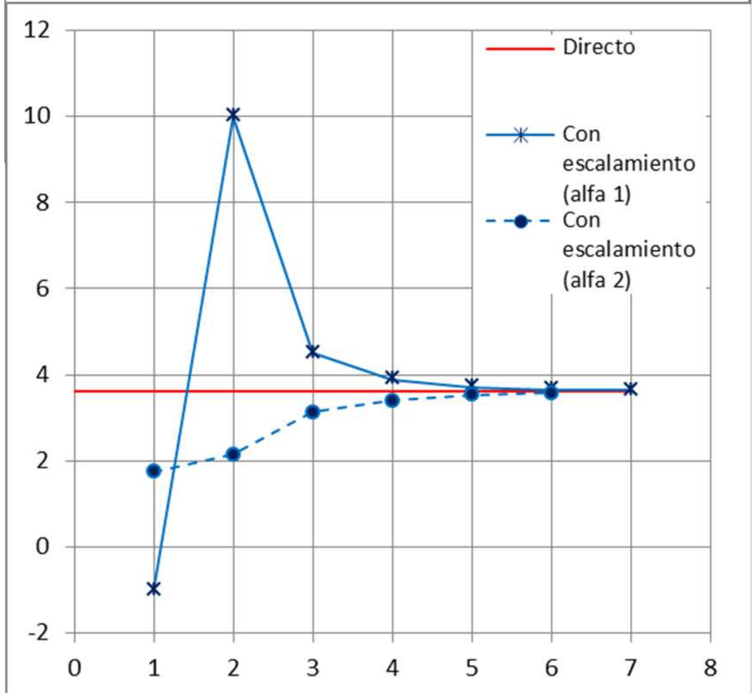
Método de la Potencia (sin escalamiento):

k		y_{k+1}	α	Error	¿Iterar?
0	y_1=	2	4	-	-
		1	0,5		
1	y_2=	5	2,5	-1,5	Continuar
		0	0	-0,5	Continuar
2	y_3=	15	3	0,5	Continuar
		-5	-	-	-
3	y_4=	50	3,3333333	0,3333333	Continuar
		-25	5	-	-
4	y_5=	175	3,5	0,1666667	Continuar
		-100	4	-1	Continuar
5	y6=	625	3,5714286	0,0714286	Continuar
		-375	3,75	-0,25	Continuar
6	y7=	2250	3,6	0,0285714	Continuar
		-1375	3,6666667	-0,0833333	Continuar

Método de la Potencia (con escalamiento):

k		y_{k+1}		x_{k+1}	α	Error	¿Iterar?
0	y_1=	-0,5	x_1=	1	-1	-	-
		3,5		-7	1,75		
1	y_2=	10	x_2=	1	10	11	Continuar
		-15		-1,5	2,1428571	0,3928571	Continuar
2	y_3=	4,5	x_3=	1	4,5	-5,5	Continuar
		-4		-0,8889	-	-	-
3	y_4=	3,8889	x_4=	1	3,8888889	-0,6111111	Continuar
		-2,7778		-0,7143	3,125	-	-
4	y_5=	3,7143	x_5=	1	3,7142857	-0,1746032	Continuar
		-2,4286		-0,6538	3,4	0,275	Continuar
5	y6=	3,6538	x6=	1	3,6538462	-0,0604396	Continuar
		-2,3077		-0,6316	3,5294118	0,1294118	Continuar
6	y7=	3,6316	x7=	1	3,6315789	-0,0222672	Continuar
		-2,2632		-0,6232	3,5833333	0,0539216	Continuar





Método de la Potencia Inversa

La matriz $A \in \Re^{NxN}$ es diagonizable y sus valores propios $\lambda_i \in \Re$ y $\underset{\sim}{v_i} \in \Re^N$ verifican

$$A\underset{\sim}{v_i} = \lambda_i \underset{\sim}{v_i}$$

Si la ecuación anterior se multiplica por la matriz inversa de A:

$$I\underset{\sim}{v_i} = \lambda_i \quad A^{-1} \quad \underset{\sim}{v_i}$$

o bien

$$A^{-1} \underset{\sim}{v_i} = \left(\frac{1}{\lambda_i}\right)\underset{\sim}{v_i}$$

es decir que $\eta_i = \frac{1}{\lambda_i}$ es autovalor dominante de A^{-1} asociado al mismo $\underset{\sim}{v_i}$.

El Método de la Potencia aplicado sobre la matriz inversa A^{-1} converge al autovalor dominante de A^{-1} ; esto es el mayor η_i tomado en valor absoluto. Pero según la relación entre los η_i y los λ_i , el mayor $|\eta_i|$ está asociado con el menor $|\lambda_i|$ de A.

El autovalor dominante de A^{-1}, es tal que:

$$|\eta_1| > |\eta_2| \geq |\eta_3| \geq \ldots \geq |\eta_N|$$

y se cumple la siguiente relación:

$$\eta_1 = \frac{1}{\lambda_N}$$

Método de la Potencia Inversa

- **Ejemplo: Método de la Potencia Inversa con escalamiento**

A^{-1}

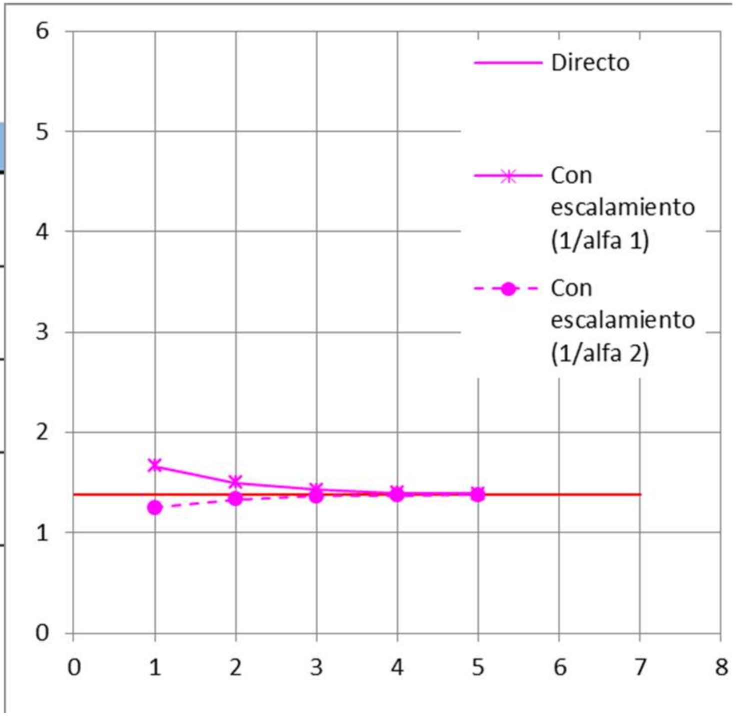
0,4	0,2
0,2	0,6

Tolerancia
0,01

Método de la Potencia Inversa (con escalamiento):

k	y_{k+1}		x_{k+1}		α	Error	¿Iterar?
0	y_1=	0,6	x_1=	1	0,6	-	-
		0,8		1,33333333	0,8		
1	y_2=	0,66666667	x_2=	1	0,66666667	0,06666667	Continuar
		1		1,5	0,75	-0,05	Continuar
2	y_3=	0,7	x_3=	1	0,7	0,03333333	Continuar
		1,1		1,57142857	0,73333333	-0,01666667	Continuar
3	y_4=	0,71428571	x_4=	1	0,71428571	0,01428571	Continuar
		1,14285714		1,6	0,72727273	-0,00606061	Detener
4	y_5=	0,72	x_5=	1	0,72	0,00571429	Detener
		1,16		1,61111111	0,725	-0,00227273	Detener

λ_{min}= 1,38888889
1,37931034



Método de la Potencia Inversa

Puede ser preferible en algunos casos realizar una factorización de la matriz A =LU, al principio y por única vez. Luego, en **cada paso** del proceso iterativo del método de Potencias inversas se resuelve el sistema de ecuaciones $A\underset{\sim}{y}_{k+1} = \underset{\sim}{x}_k$:

A =LU

$\underset{\sim}{x}_k$ son conocidos

$\underset{\sim}{y}_{k+1}$ es el vector de incógnitas

(Lz = $\underset{\sim}{x}_k$

U $\underset{\sim}{y}_{k+1}$ =z

Ejemplos de aplicación

- **Ejemplo:**

  E.D.O. con condiciones de borde

$$10 \frac{d^2u}{dx^2} + \lambda\ u = 0 \qquad u(x=0) = 0 \qquad u(x=1) = 0$$

x	u(x)
0	$u_0=0$
0,25	u_1
0,5	u_2
0,75	u_3
1	$u_4=0$

$$\left(\begin{pmatrix} 320 & -160 & 0 \\ -160 & 320 & -160 \\ 0 & -160 & 320 \end{pmatrix} - \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} \right) \begin{Bmatrix} u_1 \\ u_2 \\ u_3 \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \\ 0 \end{Bmatrix}$$

- **Ejemplo:**

  Problema de autovalores generalizado

$$(K - \lambda\ M)\ \mathbf{x} = \mathbf{0}$$

$$\left(\begin{pmatrix} 2 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 2 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix} \right) \begin{Bmatrix} x_1 \\ x_2 \\ x_3 \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \\ 0 \end{Bmatrix}$$

Iteración mediante cociente de Rayleigh

- El escalar "Cociente de Rayleigh" de un vector $\mathbf{x}$ ε $\mathbf{R}$ se define como:

  $\rho = (\mathbf{x}^T A\, \mathbf{x}) / (\mathbf{x}^T\, \mathbf{x})$

- Considerando un $\mathbf{x}$ de A ε $\mathbf{R}^{nxn}$ la mejor estimación de λ se puede considerar un problema de mínimos cuadrados n x 1 dado por:

  $A\, \mathbf{x} = \lambda\, \mathbf{x}$

  y de sus ecuaciones normales:

  $\mathbf{x}^T \mathbf{x}\, \lambda = \mathbf{x}^T A\, \mathbf{x}$

  se obtiene:

  $\lambda = (\mathbf{x}^T A\, \mathbf{x}) / (\mathbf{x}^T\, \mathbf{x})$

  El cociente de Rayleigh se puede emplear para acelerar la convergencia de los métodos iterativos de autovalores.

Deflación: Convergencia a modos intermedios

Al analizar el proceso de convergencia del Método de las Potencias resulta claro observar que si los vectores $\underset{\sim}{x_k}$ no tienen componente en el autovector $\underset{\sim}{v_1}$ (es decir $a_1 = 0$), la convergencia será al segundo autovector $\underset{\sim}{v_2}$.

Esto se logra mediante una "deflación".
Dado un vector cualquiera $\underset{\sim}{x_k}$, se lo puede escribir como:

$$\underset{\sim}{x_k} = a_1 \underset{\sim}{v_1} + a_2 \underset{\sim}{v_2} + ... + a_N \underset{\sim}{v_N}$$

ya que los $\underset{\sim}{v_i}$ son una base de $\Re^N$ por ser A $\in \Re^{NxN}$ diagonizable.

Para obtener a_1 se puede proyectar $\underset{\sim}{x_k}$ en la dirección de $\underset{\sim}{v_1}$:

$$\left\langle \underset{\sim}{v_1^T} \bullet \underset{\sim}{x_k} \right\rangle = a_1 \left\langle \underset{\sim}{v_1^T} \bullet \underset{\sim}{v_1} \right\rangle + a_2 \left\langle \underset{\sim}{v_1^T} \bullet \underset{\sim}{v_2} \right\rangle + ... + a_N \left\langle \underset{\sim}{v_1^T} \bullet \underset{\sim}{v_N} \right\rangle$$

Deflación: Convergencia a modos intermedios

Si se acepta que A es una matriz simétrica, se sabe que sus autovectores constituyen una base de $\Re^N$ **ortogonal**:

$$\left\langle \underset{\sim}{v}_1^T \bullet \underset{\sim}{v}_1 \right\rangle = \left\| \underset{\sim}{v}_1 \right\|_2^2$$

$$\left\langle \underset{\sim}{v}_1^T \bullet \underset{\sim}{v}_j \right\rangle = 0 \qquad \forall \quad j \in [2, N]$$

así para A simétrica

$$a_1 = \left\langle \underset{\sim}{v}_1^T \bullet \underset{\sim}{x}_k \right\rangle \frac{1}{\left\| \underset{\sim}{v}_1 \right\|_2^2}$$

Al reemplazar en la combinación lineal resulta:

$$\underset{\sim}{x}_k = \underset{\sim}{v}_1 \, \underset{\sim}{v}_1^T \, \underset{\sim}{x}_k \frac{1}{\left\| \underset{\sim}{v}_1 \right\|_2^2} + a_2 \, \underset{\sim}{v}_2 + \ldots + a_N \, \underset{\sim}{v}_N$$

Deflación: Convergencia a modos intermedios

o bien:

$$\underset{\sim}{x_k} - \underset{\sim}{v_1}\,\underset{\sim}{v_1^T}\,\underset{\sim}{x_k}\,\frac{1}{\left\|\underset{\sim}{v_1}\right\|_2^2} = a_2\,\underset{\sim}{v_2} + \ldots + a_N\,\underset{\sim}{v_N}$$

$$\left[I - \underset{\sim}{v_1} \bullet \underset{\sim}{v_1^T}\,\frac{1}{\left\|\underset{\sim}{v_1}\right\|_2^2} \right] \underset{\sim}{x_k} = a_2\,\underset{\sim}{v_2} + \ldots + a_N\,\underset{\sim}{v_N}$$

Al mutiplicar la igualdad anterior por A se tiene:

$$A\left[I - \underset{\sim}{v_1} \bullet \underset{\sim}{v_1^T}\,\frac{1}{\left\|\underset{\sim}{v_1}\right\|_2^2} \right] \underset{\sim}{x_k} = a_2\;A\;\underset{\sim}{v_2} + \ldots + a_N\;A\;\underset{\sim}{v_N} == \lambda_2 a_2\,\underset{\sim}{v_2} + \ldots + \lambda_N a_N\,\underset{\sim}{v_N} = \lambda_2\left(a_2\,\underset{\sim}{v_2} + \ldots + \left(\frac{\lambda_N}{\lambda_2}\right) a_N\,\underset{\sim}{v_N} \right)$$

En cada iteración se elimina la componente en $\underset{\sim}{v_1}$ si se trabaja con la matriz B_1 :

$$B_1 = A\left[I - \underset{\sim}{v_1}\,\underset{\sim}{v_1^T}\,\frac{1}{\left\|\underset{\sim}{v_1}\right\|_2^2} \right] = \left[A - A\,\underset{\sim}{v_1}\,\underset{\sim}{v_1^T}\,\frac{1}{\left\|\underset{\sim}{v_1}\right\|_2^2} \right] = \left[A - \lambda\,\underset{\sim}{v_1}\,\underset{\sim}{v_1^T}\,\frac{1}{\left\|\underset{\sim}{v_1}\right\|_2^2} \right]$$

Deflación: Convergencia a modos intermedios

Así el proceso iterativo para modos intermedios es:

$$\underset{\sim}{y_k} \ \varepsilon\ \Re^N$$

$$\underset{\sim}{x_k} = \underset{\sim}{y_k} \frac{1}{\left\| \underset{\sim}{y_k} \right\|}$$

$$\underset{\sim}{y_{k+1}} = B_1 \underset{\sim}{x_k}$$

$$B_1 = A - \lambda_1 \underset{\sim}{v_1} \underset{\sim}{v_1^T} \frac{1}{\left\| \underset{\sim}{v_1} \right\|_2^2} = A \left[I - \underset{\sim}{v_1} \underset{\sim}{v_1^T} \frac{1}{\left\| \underset{\sim}{v_1} \right\|_2^2} \right]$$

$$(\alpha_{k+1})_i = \left(\underset{\sim}{y_{k+1}} \right)_i \frac{1}{\left(\underset{\sim}{x_k} \right)_i}$$

y cuando "k" sea suficientemente grande convergerá $(\alpha_{k+1})_i \rightarrow \lambda_2$ para cualquier componente i; $\underset{\sim}{y_k}$ se alineará con v_2 .

Métodos iterativos: Autovalores vs. S.E.L.

Se busca $\underline{x} \;\varepsilon\; R^N$ tal que:

	Autovalores	Sistema de Ecuaciones Lineales
	$\underline{f} = \mathbf{A} \cdot \underline{x} - \lambda \cdot \underline{x} = \underline{0}$	$\underline{f} = \mathbf{A} \cdot \underline{x} - \underline{b} = \underline{0}$
Inicialización	Se propone una solución inicial x_0	
Recurrencia	$\underline{y}_{k+1} = \mathbf{A} \cdot \underline{x_k}$ $\underline{x}_{k+1} = \underline{y}_{k+1} \cdot \left(1/\left\lVert \underline{y}_{k+1} \right\rVert\right)$	$\underline{x}^{(k+1)} = \mathbf{T} \cdot \underline{x}^{(k)} + \underline{c}$
Control de Detención	$\left\lvert f(x_{k+1}) \right\rvert \leq \varepsilon_f$ o $\left\lvert x_{k+1} - x_k \right\rvert \leq \varepsilon_a$ o $k \leq k\max$	
Actualización	Retiene la última solución aproximada	
CRITERIOS DE CONVERGENCIA	La matriz **A** debe ser Diagonalizable. Sus autovectores deben ser linealmente independientes	Mayor de los Valores Absolutos de **T** debe ser menor a 1 $\rho(\mathbf{T}) < 1$ Matriz **A** estrictamente diagonal dominante

Resumen

- Se ha discutido la importancia del problema de autovalores y autovectores en aplicaciones prácticas de ingeniería.
- Se ha destacado la importancia del uso de los métodos iterativos
- Se han realizada ejercicios en donde se ha implementado y aplicado tanto al problema de autovalores y autovectores estándar como generalizado:
  - El método de la Potencia (sin escalamiento)
  - El método de la Potencia (con escalamiento)
  - El método de la Potencia Inversa (con escalamiento)

Resumen

- El método de la Potencia converge al autovalor mayor.
- El método de la Potencia Inversa converge al autovalor menor.
- Es conveniente el empleo del escalamiento para evitar valores de y_k muy grandes (para para $|\lambda| > 1$) o muy chicos (para $|\lambda| < 1$) ya que son proporcionales a λ^k