

Curso de Métodos Numéricos

DEMAT, Universidad de Guanajuato

Clase 13: Cálculo de eigenvectores y eigenvalores

- Método de la potencia inversa.
- Método iterativo de Jacobi.
- Iteración del cociente de Rayleigh.

MAT-251

Dr. Joaquín Peña Acevedo
CIMAT A.C.

e-mail: joaquin@cimat.mx

Método de la potencia

Dado un vector inicial \mathbf{v}^0 y fijando $k = 0$, repetimos los siguientes pasos hasta que $\|\mathbf{A}\mathbf{v}^k - \lambda_k \mathbf{v}^k\| < \tau$:

$$\mathbf{y}^{k+1} = \mathbf{A}\mathbf{v}^k$$

$$\mathbf{v}^{k+1} = \mathbf{y}^{k+1} / \|\mathbf{y}^{k+1}\|$$

$$\lambda_{k+1} = (\mathbf{v}^{k+1})^T \mathbf{A}\mathbf{v}^{k+1}$$

$$k = k + 1$$

- La normalización de los vectores que generan con este método es un paso importante en la implementación computacional.
- La convergencia del método depende de si hay un único eigenvalor que es el más grande en valor absoluto.
- Si los eigenvalores cumplen con $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$, la rapidez con la que converge el método tiene que ver con la razón $|\lambda_2/\lambda_1|$.

Observaciones (I)

El método de la potencia, cuando converge, permite obtener el eigenvector asociado al eigenvalor más grande en valor absoluto.

Se puede tratar de obtener otros eigenpares con esta metodología pero hay que revisar otros conceptos.

Definición

Sea $\mathbf{A} \in \mathbb{C}^{n \times n}$ y $\mathbf{u} \in \mathbb{R}^n$. Se dice que \mathbf{u} es un *eigenvector izquierdo* de \mathbf{A} asociado al eigenvalor μ si

$$\mathbf{u}^T \mathbf{A} = \mu \mathbf{u}^T. \quad (1)$$

Cuando se cumple que $\mathbf{A} \mathbf{v} = \lambda \mathbf{v}$, se dice que \mathbf{v} es un *eigenvector derecho* de \mathbf{A} asociado al eigenvalor λ .

Si calculamos la transpuesta de ambos miembros de la ecuación (1), $\mathbf{A}^T \mathbf{u} = \mu \mathbf{u}$, se ve que un eigenvector izquierdo es un eigenvector derecho de \mathbf{A}^T . Como \mathbf{A}^T y \mathbf{A} tienen los mismos eigenvalores, no se hace alguna diferencia entre eigenvalores izquierdos o derechos.

Observaciones (II)

Proposición 1

Sean μ y λ dos eigenvalores de \mathbf{A} , $\mu \neq \lambda$, y sean \mathbf{u} y \mathbf{v} los eigenvectores izquierdo y derecho asociados a μ y λ , respectivamente. Entonces \mathbf{u} y \mathbf{v} son ortogonales.

Dem. Puesto que $\mathbf{u}^T \mathbf{A} = \mu \mathbf{u}^T$ y $\mathbf{A} \mathbf{v} = \lambda \mathbf{v}$, entonces

$$\lambda \mathbf{u}^T \mathbf{v} = \mathbf{u}^T \mathbf{A} \mathbf{v} = \mu \mathbf{u}^T \mathbf{v} \implies (\lambda - \mu) \mathbf{u}^T \mathbf{v} = 0 \implies \mathbf{u}^T \mathbf{v} = 0.$$

Proposición 2

Sean $\mathbf{v}_1, \dots, \mathbf{v}_k$ los eigenvectores de $\mathbf{A} \in \mathbb{C}^{n \times n}$ asociados a los k eigenvalores distintos $\lambda_1, \dots, \lambda_k$. Entonces los eigenvectores $\mathbf{v}_1, \dots, \mathbf{v}_k$ son linealmente independientes. En particular, si todos los eigenvalores de la matriz \mathbf{A} son diferentes, entonces sus eigenvectores forman una base en \mathbb{C}^n y \mathbf{A} es diagonalizable.

Observaciones (III)

Dem. Supongamos que solo los eigenvectores $\mathbf{v}_1, \dots, \mathbf{v}_p$ son linealmente independientes, para $p < k$. Entonces podemos encontrar los coeficientes c_1, \dots, c_k , no todos cero, tales que

$$\mathbf{v}_{p+1} = c_1 \mathbf{v}_1 + \dots + c_p \mathbf{v}_p \quad (2)$$

Multiplicando por \mathbf{A} ambos miembros de (2), se tiene

$$\lambda_{p+1} \mathbf{v}_{p+1} = \mathbf{A} \mathbf{v}_{p+1} = c_1 \lambda_1 \mathbf{v}_1 + \dots + c_p \lambda_p \mathbf{v}_p \quad (3)$$

Restando el resultado de multiplicar (2) por λ_{p+1} a (3) se obtiene

$$c_1(\lambda_1 - \lambda_{p+1})\mathbf{v}_1 + \dots + c_p(\lambda_p - \lambda_{p+1})\mathbf{v}_p = \mathbf{0}$$

Por ser linealmente independientes, todos los coeficientes deben ser 0, debe haber al menos un índice i para el cual $c_i \neq 0$, de modo que $\lambda_i - \lambda_{p+1} = 0$, pero esto contradice que los eigenvalores son distintos. Así, se debe cumplir que $p = k$.

Note que si \mathbf{A} tiene n eigenvalores diferentes, entonces todas las columnas de la matriz $\mathbf{V} = [\mathbf{v}_1 \ \cdots \ \mathbf{v}_n]$ son linealmente independientes y por lo tanto es invertible. Como $\mathbf{AV} = \mathbf{VD}$, con $\mathbf{D} = \text{diag}(\lambda_1, \dots, \lambda_n)$, tenemos

$$\mathbf{A} = \mathbf{VDV}^{-1},$$

es decir, \mathbf{A} es similar a una matriz diagonal.

Deflación en el cálculo de eigenvalores (I)

Supongamos que tenemos procedimiento que nos permite calcular un eigenpar $(\lambda_1, \mathbf{v}_1)$ de la matriz \mathbf{A} . Para calcular otros eigenpares podemos usar la técnica de *deflación*, es decir, construimos una matriz \mathbf{A}_1 en la que el eigenvalor λ_1 es reemplazado por 0 y el resto de los eigenvalores no cambian.

La siguiente proposición indica como se puede obtener otros eigenpares.

Proposición 3

Sea $\mathbf{A} \in \mathbb{C}^{n \times n}$ con eigenvalores $\lambda_1, \dots, \lambda_n$, con eigenvectores izquierdos $\mathbf{u}_1, \dots, \mathbf{u}_n$ y eigenvectores derechos $\mathbf{v}_1, \dots, \mathbf{v}_n$. Supongamos que λ_1 tiene multiplicidad 1 y definamos

$$\mathbf{A}_1 = \mathbf{A} - \mu \mathbf{v}_1 \mathbf{z}^T,$$

donde \mathbf{z} es un vector tal que $\mathbf{z}^T \mathbf{v}_1 = 1$. Entonces \mathbf{A}_1 tiene por eigenvalores $\lambda_1 - \mu, \lambda_2, \dots, \lambda_n$.

Deflación en el cálculo de eigenvalores (II)

Para obtener los eigenvalores mediante el método de la potencia, lo usual es normalizar \mathbf{v}_1 de modo que $\|\mathbf{v}_1\|_2 = 1$, tomar $\mathbf{z} = \mathbf{v}_1$ y $\mu = \lambda_1$. Esto es el método de deflación de Hotelling

Ejemplo 1 (I)

$$\mathbf{A} = \begin{bmatrix} 4.6023708 & -0.6484326 & 2.6800333 & 0.1378698 & 0.3655997 \\ -0.3480484 & -4.9229298 & 0.0876574 & -1.2066205 & -1.2046782 \\ 1.0992412 & 0.0206325 & -4.2138133 & -0.3166074 & -1.3973391 \\ 0.5966447 & 2.3193512 & -2.5578133 & 4.6455689 & -0.1206493 \\ -0.0702810 & -1.8568523 & 0.7597306 & -0.1774737 & 4.4888034 \end{bmatrix}$$

Los eigenvalores de esta matriz son:

$$-4.90000, -4.50000, 4.4000000, 4.60000, 5.00000$$

Aplicando el método de la potencia con una tolerancia $\tau = 10^{-8}$, en 1025 iteraciones y $\|\mathbf{A}\mathbf{v}_1 - \lambda_1\mathbf{v}_1\| = 9.96 \times 10^{-9}$ se obtiene

$$\lambda_1 = 5, \quad \mathbf{v}_1 = \begin{pmatrix} 0.907136 \\ -0.098688 \\ 0.064364 \\ 0.345457 \\ 0.209478 \end{pmatrix}.$$

Ejemplo 1 (II)

Aplicando el método de la potencia a la matriz

$$\mathbf{A}_1 = \mathbf{A} - \lambda_1 \mathbf{v}_1 \mathbf{v}_1^T,$$

en 312 iteraciones se obtiene

$$\lambda_2 = -4.9, \quad \mathbf{v}_2 = \begin{pmatrix} -0.10503 \\ 0.93676 \\ 0.23286 \\ -0.19511 \\ 0.13839 \end{pmatrix}, \quad \|\mathbf{A}_1 \mathbf{v}_2 - \lambda_2 \mathbf{v}_2\| = 9.90 \times 10^{-9}.$$

Aplicando el método de la potencia a la matriz

$$\mathbf{A}_2 = \mathbf{A} - \lambda_1 \mathbf{v}_1 \mathbf{v}_1^T - \lambda_2 \mathbf{v}_2 \mathbf{v}_2^T,$$

se obtiene

$$\lambda_3 = 4.6, \quad \mathbf{v}_3 = \begin{pmatrix} -0.177104 \\ -0.044042 \\ -0.152952 \\ -0.169554 \\ 0.956322 \end{pmatrix}, \quad \|\mathbf{A}_2 \mathbf{v}_3 - \lambda_3 \mathbf{v}_3\| = 9.92 \times 10^{-9}.$$

Ejemplo 1 (III)

Aplicando el método de la potencia a la matriz

$$\mathbf{A}_3 = \mathbf{A} - \lambda_1 \mathbf{v}_1 \mathbf{v}_1^T - \lambda_2 \mathbf{v}_2 \mathbf{v}_2^T - \lambda_3 \mathbf{v}_3 \mathbf{v}_3^T,$$

se obtiene

$$\lambda_4 = -4.5, \quad \mathbf{v}_4 = \begin{pmatrix} -0.29289 \\ -0.15833 \\ 0.83438 \\ 0.29142 \\ -0.32867 \end{pmatrix}, \quad \|\mathbf{A}_3 \mathbf{v}_4 - \lambda_4 \mathbf{v}_4\| = 9.85 \times 10^{-9}.$$

Aplicando el método de la potencia a la matriz

$$\mathbf{B}_4 = \mathbf{A} - \lambda_1 \mathbf{v}_1 \mathbf{v}_1^T - \lambda_2 \mathbf{v}_2 \mathbf{v}_2^T - \lambda_3 \mathbf{v}_3 \mathbf{v}_3^T - \lambda_4 \mathbf{v}_4 \mathbf{v}_4^T,$$

se obtiene

Ejemplo 1 (IV)

$$\lambda_5 = 4.4, \quad \mathbf{v}_4 = \begin{pmatrix} -0.373325 \\ -0.351160 \\ 0.065192 \\ 0.855864 \\ -0.023715 \end{pmatrix} \quad \|\mathbf{B}_4 \mathbf{v}_5 - \lambda_4 \mathbf{v}_5\| = 2.59 \times 10^{-10}.$$

NOTA: Los vectores obtenidos $\mathbf{v}_2, \mathbf{v}_3$, etc. no son eigenvectores de \mathbf{A} . Sólo obtenemos los eigenvalores de la matriz.

Descomposición espectral (I)

Si $\mathbf{V} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \cdots \ \mathbf{v}_n]$ es la matriz que tiene por columnas los eigenvectores de \mathbf{A} , y $\mathbf{D} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ la matriz diagonal con los eigenvalores de \mathbf{A} , entonces

$$\mathbf{AV} = \mathbf{VD}.$$

A esto se llama la *descomposición espectral* de \mathbf{A} .

Ejemplo: Considere la matriz

$$\mathbf{A} = \begin{bmatrix} 0.995813 & 0.438443 & 0.340460 & 0.375729 & 0.471884 \\ 0.399457 & 0.085561 & 0.530025 & 0.351771 & 0.242472 \\ 0.648080 & 0.886231 & 0.950510 & 0.947743 & 0.484992 \\ 0.061177 & 0.572064 & 0.811057 & 0.136243 & 0.321605 \\ 0.773554 & 0.283946 & 0.340117 & 0.799517 & 0.119569 \end{bmatrix}$$

Entonces la matriz \mathbf{V} formada con los eigenvectores de \mathbf{A} y la matriz \mathbf{D} son

Descomposición espectral (II)

$$\mathbf{V} = \begin{bmatrix} 0.432871 & 0.777414 & 0.126983 & -0.068752 & -0.293537 \\ 0.301977 & -0.039717 & 0.094261 & 0.777720 & -0.193608 \\ 0.668416 & -0.397584 & -0.214973 & -0.526157 & -0.252219 \\ 0.357399 & -0.459252 & 0.559358 & 0.173162 & 0.214484 \\ 0.383301 & 0.158321 & -0.784790 & -0.289133 & 0.875634 \end{bmatrix}$$
$$\mathbf{D} = \begin{bmatrix} 2.55546 & 0 & 0 & 0 & 0 \\ 0 & 0.67344 & 0 & 0 & 0 \\ 0 & 0 & -0.51639 & 0 & 0 \\ 0 & 0 & 0 & -0.32015 & 0 \\ 0 & 0 & 0 & 0 & -0.10466 \end{bmatrix}$$

Se puede ver que

$$\|\mathbf{AV} - \mathbf{VD}\|_2 \approx 1.8 \times 10^{-15},$$

Además $\mathbf{A} = \mathbf{VDV}^{-1}$, donde

$$\mathbf{V}^{-1} = \begin{bmatrix} 0.509985 & 0.429247 & 0.547913 & 0.459126 & 0.311231 \\ 0.895673 & -0.172922 & -0.507643 & -0.153925 & 0.153501 \\ 0.484998 & -0.621600 & -0.416033 & 1.108661 & -0.366251 \\ -0.216539 & 1.098243 & -0.346404 & -0.126657 & 0.101484 \\ -0.022007 & -0.351102 & -0.635311 & 0.778671 & 0.683293 \end{bmatrix}.$$

Descomposición espectral (III)

Esto último muestra que los eigenvectores no son ortonormales.

La ortogonalidad de los eigenvectores se puede garantizar cuando \mathbf{A} es simétrica:

Supongamos que $\mathbf{A} = \mathbf{A}^T$ y sean (λ, \mathbf{v}) y (μ, \mathbf{u}) son eigenpares de \mathbf{A} , con $\lambda \neq \mu$. Entonces

$$\mathbf{u}^T \mathbf{A} \mathbf{v} = (\mathbf{u}^T \mathbf{A} \mathbf{v})^T = \mathbf{v}^T \mathbf{A}^T \mathbf{u} = \mathbf{v}^T \mathbf{A} \mathbf{u}.$$

Además,

$$\begin{aligned}\mathbf{u}^T \mathbf{A} \mathbf{v} &= \lambda \mathbf{u}^T \mathbf{v} \\ \mathbf{v}^T \mathbf{A} \mathbf{u} &= \mu \mathbf{v}^T \mathbf{u} = \mu \mathbf{u}^T \mathbf{v}\end{aligned}$$

Así,

$$\lambda \mathbf{u}^T \mathbf{v} = \mu \mathbf{u}^T \mathbf{v} \implies (\lambda - \mu) \mathbf{u}^T \mathbf{v} = 0 \implies \mathbf{u}^T \mathbf{v} = 0.$$

Haciendo que los eigenvectores tengan norma Euclidiana unitaria, se ve $\mathbf{V}^T \mathbf{V} = \mathbf{I}$, entonces $\mathbf{A} = \mathbf{V} \mathbf{D} \mathbf{V}^T$. Esto es, la matriz \mathbf{A} es *diagonalizable*.

Así, en el caso de matrices simétricas que además son no singulares, se puede usar su descomposición espectral para resolver sistemas de ecuaciones lineales:

Método de la potencia inversa (I)

- En lugar de aplicar las iteraciones a la matriz \mathbf{A} , el método de la potencia inversa opera con la matriz \mathbf{A}^{-1} .
- De este modo, debe converger al eigenvalor más pequeño en valor absoluto si

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_{n-1}| > |\lambda_n|$$

- Para no tener que calcular la inversa, lo que se prefiere es resolver un sistema de ecuaciones lineales mediante algún método numérico:

Método de la potencia inversa

Dar un vector inicial $\mathbf{x}^{(0)}$, con $\|\mathbf{x}^{(0)}\| = 1$, $k = 0$ y una tolerancia $\tau > 0$. Ejecutar los siguientes pasos:

- 1 Revolver $\mathbf{A}\mathbf{y}^{(k)} = \mathbf{x}^{(k)}$
- 2 Calcular $\mathbf{x}^{(k+1)} = \mathbf{y}^{(k)} / \|\mathbf{y}^{(k)}\|$.
- 3 Calcular $\lambda_{k+1} = (\mathbf{x}^{(k+1)})^T \mathbf{A}\mathbf{x}^{(k+1)}$.
- 4 Calcular $e = \|\mathbf{A}\mathbf{x}^{(k+1)} - \lambda_{k+1}\mathbf{x}^{(k+1)}\|$
- 5 Si $e > \tau$, hacer $k = k + 1$ y regresar al paso 1.

Una ventaja del método de la potencia inversa es que se puede adaptar para obtener otro eigenvalor:

- Se aplica a la matriz $(\mathbf{A} - \delta\mathbf{I})^{-1}$.
- El método converge al eigenvalor más cercano a δ .

Método de la potencia inversa (III)

Método de la potencia inversa

Dado un vector inicial \mathbf{x}^0 , el escalar δ que define la traslación, y fijando $k = 0$, repetimos hasta que $\|\mathbf{r}\| < \tau$ para obtener el eigenpar (μ, \mathbf{x}^k) :

$$\text{Resolver } (\mathbf{A} - \delta \mathbf{I})\mathbf{y} = \mathbf{x}^k$$

$$\hat{\mathbf{x}} = \mathbf{y}/\|\mathbf{y}\|$$

$$\mathbf{w} = \mathbf{x}^k/\|\mathbf{y}\|$$

$$\rho = \hat{\mathbf{x}}^\top \mathbf{w}$$

$$\mu = \delta + \rho$$

$$\mathbf{r} = \mathbf{w} - \rho \hat{\mathbf{x}}$$

$$\mathbf{x}^{k+1} = \hat{\mathbf{x}}$$

$$k = k + 1$$

Convergencia del método de la potencia inversa

El argumento es el mismo que el usado para el método de la potencia. Note que

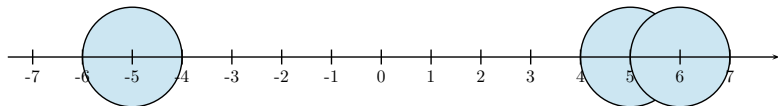
- $\mathbf{A} - \delta \mathbf{I}$ tiene los mismos eigenvectores que \mathbf{A} .
- Si λ_i es un eigenvalor de \mathbf{A} , entonces $\lambda_i - \delta$ es un eigenvalor de $\mathbf{A} - \delta \mathbf{I}$.
- Los eigenvalores de $(\mathbf{A} - \delta \mathbf{I})^{-1}$ son de la forma

$$\frac{1}{\lambda_i - \delta}$$

Eligiendo de manera apropiada δ podemos hacer que $\frac{1}{\lambda_i - \delta}$ sea el eigenvalor dominante.

Ejemplo (I)

$$\mathbf{A} = \begin{bmatrix} 5 & 1 & -1 \\ 0 & 6 & 1 \\ 1 & 0 & -5 \end{bmatrix}$$



Eigenvalores: -4.88959807 , 4.81228115 , 6.07731691 .

Tomamos $\mathbf{x}^0 = (1, 1, 1)^T$. Para diferentes valores de σ se obtiene lo siguiente para una tolerancia 10^{-8} y un máximo de 200 iteraciones:

δ	Iteraciones	Valor final	$\ r\ $
-6	9	-4.88959807	7.6868×10^{-9}
4	24	4.81228115	8.4095×10^{-9}
7	200	7.92268309	1.8454
5.5	84	6.07731692	8.4447×10^{-9}

Rotaciones de Givens

Una rotación de Givens, $G(i, j, \theta) = [g_{ij}] \in \mathbb{R}^{n \times n}$, es una matriz que coincide con la matriz identidad, excepto en cuatro entradas:

$$g_{ii} = g_{jj} = \cos \theta = c, \quad g_{ij} = \sin \theta = s, \quad g_{ji} = -\sin \theta = -s.$$

$$G(i, j, \theta) = \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & c & \cdots & s & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \vdots & -s & \cdots & c & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \vdots & 0 & \vdots & 0 & \vdots & 1 \end{bmatrix}$$

Método iterativo de Jacobi (I)

Sea $\mathbf{A}^{(0)} \in \mathbb{R}^{n \times n}$ una matriz simétrica y $\mathbf{G}(i, j, \theta)$ una rotación de Givens. Entonces $\mathbf{A}^{(1)} = \mathbf{G}^\top \mathbf{A}^{(0)} \mathbf{G}$ es una matriz simétrica similar a $\mathbf{A}^{(0)}$. Además, tenemos la siguiente relación entre las entradas de las matrices:

$$\begin{aligned}A_{ii}^{(1)} &= c^2 A_{ii}^{(0)} - 2sc A_{ij}^{(0)} + s^2 A_{jj}^{(0)} \\A_{jj}^{(1)} &= s^2 A_{ii}^{(0)} + 2sc A_{ij}^{(0)} + c^2 A_{jj}^{(0)} \\A_{ij}^{(1)} &= (c^2 - s^2) A_{ij}^{(0)} + sc(A_{ii}^{(0)} - A_{jj}^{(0)}) \\A_{ik}^{(1)} &= c A_{ik}^{(0)} - s A_{jk}^{(0)} \quad k \neq i, j \\A_{jk}^{(1)} &= s A_{ik}^{(0)} + c A_{jk}^{(0)} \quad k \neq i, j \\A_{kl}^{(1)} &= A_{kl}^{(0)} \quad k, l \neq i, j\end{aligned}$$

Escogemos θ de modo que $A_{ij}^{(1)} = A_{ji}^{(1)} = 0$. Entonces

$$\tan 2\theta = \frac{2A_{ij}^{(0)}}{A_{jj}^{(0)} - A_{ii}^{(0)}}.$$

Método iterativo de Jacobi (II)

- El proceso se repite de forma iterativa definiendo $\mathbf{A}^{(k+1)} = \mathbf{G}^T \mathbf{A}^{(k)} \mathbf{G}$.
- En cada iteración, para definir los índices i y j , lo ideal es escoger el elemento $A_{ij}^{(k)}$ de mayor magnitud para acelerar el proceso.
- En la práctica se van revisando cada elemento fuera de la diagonal, y si es menor que cierta tolerancia, se aplica el proceso anterior para hacerlo cero.
- Se tienen que dar varios 'barridos' a la matriz porque el proceso iterativo no preserva los ceros que previamente se han creado.

Al final se obtiene una matriz $\mathbf{A}^{(k)}$ que es diagonal y los elementos en ella corresponden a los eigenvalores de la matriz original.

Método iterativo de Jacobi (III)

Método de Jacobi

Dar una matriz simétrica $\mathbf{A}^{(0)} = [a_{ij}^{(0)}]$, una tolerancia $\tau > 0$ y un número M de iteraciones máximas. Inicializar $\mathbf{V} = \mathbf{I}$.

Para $k = 1, 2, \dots, M$, realizar los siguientes pasos:

- ❶ Encontrar los índices i, j del elemento de la matriz $\mathbf{A}^{(k-1)} = [a_{ij}^{(k-1)}]$ más grande, en valor absoluto, fuera de su diagonal.
- ❷ Terminar de iterar si $|a_{ij}^{(k-1)}| < \tau$.
- ❸ Calcular $\delta = (a_{jj}^{(k-1)} - a_{ii}^{(k-1)}) / (2a_{ij}^{(k-1)})$.
- ❹ $t = \text{sign}(\delta) / (|\delta| + \sqrt{1 + \delta^2})$.
- ❺ $c = \cos \theta = 1 / \sqrt{1 + t^2}$.
- ❻ $s = \sin \theta = ct$.
- ❼ Calcular la rotación de Givens $\mathbf{G} = \mathbf{G}(i, j, \theta)$ (usando c y s).
- ❽ Calcular $\mathbf{A}^{(k)} = \mathbf{G}^T \mathbf{A}^{(k-1)} \mathbf{G}$.
- ❾ Calcular $\mathbf{V} = \mathbf{V} \mathbf{G}$.

Ejemplo del uso del método de Jacobi (I)

$$\mathbf{A} = \begin{bmatrix} 2.0 & 1.0 & -1.0 \\ 1.0 & 5.0 & -2.0 \\ -1.0 & -2.0 & 4.0 \end{bmatrix} \xrightarrow{\mathbf{G}(3,2)} \begin{bmatrix} 5.3028 & -2.1764 & -0.3477 \\ -2.1764 & 4.0603 & 1 \times 10^{-16} \\ -0.3477 & 0.0000 & 1.6369 \end{bmatrix}$$

$$\xrightarrow{\mathbf{G}(2,1)} \begin{bmatrix} 6.9448 & 4 \times 10^{-16} & -0.278 \\ 2 \times 10^{-16} & 2.4182 & -0.209 \\ -0.278 & -0.209 & 1.637 \end{bmatrix} \xrightarrow{\mathbf{G}(3,1)} \begin{bmatrix} 6.9593 & 0.011 & 3 \times 10^{-17} \\ 0.011 & 2.418 & -0.209 \\ -1 \times 10^{-17} & -0.209 & 1.6224 \end{bmatrix}$$

$$\xrightarrow{\mathbf{G}(3,2)} \begin{bmatrix} 6.9593 & 0.0106 & 0.0026 \\ 0.0105 & 2.4699 & 3 \times 10^{-16} \\ 0.0026 & 0.0000 & 1.5708 \end{bmatrix} \xrightarrow{\mathbf{G}(2,1)} \begin{bmatrix} 6.9593 & 1 \times 10^{-16} & 0.0026 \\ 1 \times 10^{-18} & 2.4698323 & -6 \times 10^{-6} \\ 0.0026 & -6 \times 10^{-6} & 1.5708 \end{bmatrix}$$

Por lo que los eigenvalores estimados son:

$$\{1.5708263, 2.4698323, 6.9593414\}$$

Iteración del cociente de Rayleigh (I)

Dada \mathbf{A} una matriz simétrica y un vector \mathbf{u} diferente de cero, el *cociente de Rayleigh* se define como

$$\rho(\mathbf{u}, \mathbf{A}) = \frac{\mathbf{u}^\top \mathbf{A} \mathbf{u}}{\mathbf{u}^\top \mathbf{u}}.$$

Algunas propiedades que se tienen son:

- ❶ Para cualquier escalar α , se tiene que $\rho(\alpha \mathbf{u}, \mathbf{A}) = \rho(\mathbf{u}, \mathbf{A})$.
- ❷ Si \mathbf{u} es un eigenvector de \mathbf{A} , $\mathbf{A} \mathbf{u} = \lambda \mathbf{u}$,

$$\rho(\mathbf{u}, \mathbf{A}) = \frac{\mathbf{u}^\top \mathbf{A} \mathbf{u}}{\mathbf{u}^\top \mathbf{u}} = \frac{\lambda \mathbf{u}^\top \mathbf{u}}{\mathbf{u}^\top \mathbf{u}} = \lambda.$$

Proposición

Para cualquier vector \mathbf{x} tal que $\|\mathbf{x}\|_2 = 1$ se cumple que

$$\lambda_{\min}(\mathbf{A}) \leq \mathbf{x}^T \mathbf{A} \mathbf{x} \leq \lambda_{\max}(\mathbf{A}).$$

La igualdad se cumple si y sólo si \mathbf{x} es un eigenvector de \mathbf{A} .

Proposición

Para cualquier vector $\mathbf{x} \neq \mathbf{0}$ se cumple que

$$\lambda_{\min}(\mathbf{A}) \leq \rho(\mathbf{x}, \mathbf{A}) \leq \lambda_{\max}(\mathbf{A}),$$

y la igualdad se cumple si y sólo si \mathbf{x} es un eigenvector de \mathbf{A} .

Iteración del cociente de Rayleigh (IV)

Algoritmo de iteración del cociente de Rayleigh

Dar un vector inicial $\mathbf{x}^{(0)}$, con $\|\mathbf{x}^{(0)}\|_2 = 1$, una tolerancia $\tau > 0$ y un número máximo de iteraciones M .

- ❶ Calcular $\rho_0 = \rho(\mathbf{x}^{(0)}, \mathbf{A})$.
- ❷ Para $k = 1, \dots, M$ realizar los siguientes pasos:
 - (2.1) Resolver $(\mathbf{A} - \rho_{k-1}\mathbf{I})\mathbf{y}^{(k)} = \mathbf{x}^{(k-1)}$.
 - (2.2) $\mathbf{x}^{(k)} = \frac{\mathbf{y}^{(k)}}{\|\mathbf{y}^{(k)}\|_2}$.
 - (2.3) $\rho_k = \rho(\mathbf{x}^{(k)}, \mathbf{A})$.
 - (2.4) Terminar de iterar si $\|\mathbf{A}\mathbf{x}^{(k)} - \rho_k\mathbf{x}^{(k)}\| < \tau$.
- ❸ Devolver $\mathbf{x}^{(k)}$ y ρ_k como una estimación de un eigenpar de \mathbf{A} .

Observe que $\mathbf{A} - \rho_{k-1}\mathbf{I}$ es una matriz simétrica que no necesariamente es definida positiva, por lo que no se puede usar la factorización de Cholesky,

Iteración del cociente de Rayleigh (V)

pero sí se puede aplicar la factorización LDL^T en cada paso de la iteración del algoritmo.

Ejemplo. Considere la matriz

$$\mathbf{A} = \begin{bmatrix} 51.00000 & -1.00000 & 32.00000 & -17.00000 & 3.00000 \\ -1.00000 & 2.00000 & -9.00000 & -22.00000 & 9.00000 \\ 32.00000 & -9.00000 & -38.00000 & -23.00000 & -4.00000 \\ -17.00000 & -22.00000 & -23.00000 & 29.00000 & 29.00000 \\ 3.00000 & 9.00000 & -4.00000 & 29.00000 & -5.00000 \end{bmatrix},$$

la cual tiene como eigenvalores:

$$-55.9969, -30.7179, 9.6210, 39.1484, 76.9454.$$

Aplicando el algoritmo partiendo de $\mathbf{x}^{(0)} = (1/\sqrt{5}, 1/\sqrt{5}, 1/\sqrt{5}, 1/\sqrt{5}, 1/\sqrt{5})^T$ se obtiene:

Iteración del cociente de Rayleigh (VI)

Iteración	ρ_k	Error e
1	48.0351	1.718300
2	40.2691	6.464000
3	39.1494	0.203040
4	39.1484	0.000005
5	39.1484	6.97×10^{-15}

Tomando $\mathbf{x}^{(0)} = (0, 0, 1, 0, 0)^T$ se obtiene:

Iteración	ρ_k	Error e
1	-43.5572	13.787
2	-44.2690	12.630
3	-46.0626	12.347
4	-50.6017	10.357
5	-55.5018	3.5029
6	-55.9967	0.0713
7	-55.9969	5.69×10^{-7}
8	Matriz singular	

Tomando $\mathbf{x}^{(0)} = (0, 1, 0, 0, 0)^T$ se obtiene:

Iteración del cociente de Rayleigh (VII)

Iteración	ρ_k	Error e
1	9.1875	5.8551e+00
2	9.6210	6.3067e-02
3	9.6210	4.5007e-08

Tomando $\mathbf{x}^{(0)} = (1, 0, 0, 0, 0)^T$ se obtiene:

Iteración	ρ_k	Error e
1	-2.5526	2.0805e+01
2	6.2475	1.1577e+01
3	9.5905	1.1529e+00
4	9.6210	8.8385e-04
5	9.6210	3.7819e-13