

Algebra matricial 2

Curso de Física Computacional

M. en C. Gustavo Contreras Mayén

9 de mayo de 2017

1 Métodos iterativos

- El método de Jacobi
 - Transformación de Similitud y Diagonalización
 - Rotación de Jacobi
 - Diagonalización de Jacobi

1 Métodos iterativos

- El método de Jacobi
 - Transformación de Similitud y Diagonalización
 - Rotación de Jacobi
 - Diagonalización de Jacobi

El método de Jacobi

El método de Jacobi es un procedimiento iterativo relativamente simple que extrae todos los valores propios y vectores propios de una matriz simétrica.

Es útil cuando se tienen matrices pequeñas (digamos, menos de 50×50), porque el esfuerzo computacional se incrementa muy rápidamente con el tamaño de la matriz.

La fuerza principal del método es su robustez.

Transformación de Similitud y Diagonalización

Consideremos el problema estándar de valores propios de la matriz

$$\mathbf{A} \mathbf{x} = \lambda \mathbf{x} \quad (1)$$

donde la matriz \mathbf{A} es simétrica. Apliquemos la transformación

$$\mathbf{x} = \mathbf{P} \mathbf{x}^* \quad (2)$$

donde \mathbf{P} es una matriz no singular.

Sustituyendo la ec. (2) en la ec. (1) y multiplicando por la izquierda de cada lado de la igualdad por P^{-1} , se obtiene

$$P^{-1} A P x^* = \lambda P^{-1} P x^*$$

Sustituyendo la ec. (2) en la ec. (1) y multiplicando por la izquierda de cada lado de la igualdad por \mathbf{P}^{-1} , se obtiene

$$\mathbf{P}^{-1} \mathbf{A} \mathbf{P} \mathbf{x}^* = \lambda \mathbf{P}^{-1} \mathbf{P} \mathbf{x}^*$$

o equivalentemente

$$\mathbf{A}^* \mathbf{x}^* = \lambda \mathbf{x}^* \quad (3)$$

donde $\mathbf{A}^* = \mathbf{P}^{-1} \mathbf{A} \mathbf{P}$

Dado que λ no se modifica/altera durante la transformación, los valores propios de A son los mismos valores propios de A^* .

Las matrices que tienen los mismos valores propios se consideran similares, y la transformación entre ellas se llama transformación de similitud.

Las transformaciones de similitud se utilizan con frecuencia para cambiar un problema de valores propios a una forma que es más fácil de resolver.

Supongamos que nos arreglamos por algo para encontrar una matriz P que diagonaliza A^* .

Las ecuaciones (3) son entonces

$$\begin{bmatrix} A_{11}^* - \lambda & 0 & \dots & 0 \\ 0 & A_{22}^* - \lambda & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_{nn}^* - \lambda \end{bmatrix} = \begin{bmatrix} x_1^* \\ x_2^* \\ \vdots \\ x_n^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Que tiene por soluciones

$$\lambda_1 = A_{11}^* \quad \lambda_2 = A_{22}^* \quad \lambda_n = A_{nn}^* \quad (4)$$

$$\mathbf{x}_1^* = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \mathbf{x}_2^* = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} \quad \dots \quad \mathbf{x}_n^* = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

Que tiene por soluciones

$$\lambda_1 = A_{11}^* \quad \lambda_2 = A_{22}^* \quad \lambda_n = A_{nn}^* \quad (4)$$

$$\mathbf{x}_1^* = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \mathbf{x}_2^* = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} \quad \dots \quad \mathbf{x}_n^* = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

o

$$\mathbf{X}^* = [\mathbf{x}_1^* \quad \mathbf{x}_2^* \quad \dots \quad \mathbf{x}_n^*] = \mathbf{I}$$

De acuerdo a la ec. (2), los valores propios de A son

$$X = P X^* = P I = P \quad (5)$$

Por lo tanto la matriz de transformación P contiene los vectores propios de A , y los valores propios de A son los elementos de la diagonal de A^* .

Una transformación especial de similitud es la rotación plana

$$\mathbf{x} = \mathbf{R} \mathbf{x}^* \quad (6)$$

donde

$$\mathbf{R} = \begin{matrix} & & k & & \ell & & & \\ \left(\begin{array}{cccccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & c & 0 & 0 & s & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -s & 0 & 0 & c & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right) & \begin{matrix} k \\ \\ \\ \ell \end{matrix} \end{matrix} \quad (7)$$

La matriz \mathbf{R} se llama *matriz de rotación de Jacobi*.

Nótese que \mathbf{R} es una matriz de identidad modificada por los términos $c = \cos \theta$ y $s = \sin \theta$ que aparecen en las intersecciones de columnas/filas k y ℓ , donde θ es el ángulo de rotación.

La matriz de rotación tiene la propiedad útil de ser ortogonal, lo que significa que

$$\mathbf{R}^{-1} = \mathbf{R}^T \quad (8)$$

Una consecuencia de la ortogonalidad es que la transformación en la ecuación (2) tiene la característica esencial de una rotación: *conserva la magnitud del vector*, es decir $|\mathbf{x}| = |\mathbf{x}^*|$

La transformación de similitud correspondiente al plano de rotación en la ec. (6) es

$$\mathbf{A}^* = \mathbf{R}^{-1} \mathbf{A} \mathbf{R} = \mathbf{R}^T \mathbf{A} \mathbf{R} \quad (9)$$

La transformación de similitud correspondiente al plano de rotación en la ec. (6) es

$$\mathbf{A}^* = \mathbf{R}^{-1} \mathbf{A} \mathbf{R} = \mathbf{R}^T \mathbf{A} \mathbf{R} \quad (9)$$

La matriz \mathbf{A}^* no solo tiene los mismos valores propios de la matriz original \mathbf{A} , sino que también por la ortogonalidad de \mathbf{R} , es simétrica.

La transformación en la ec. (9) cambia solo los renglones/columna k y ℓ de A .

Las fórmulas para esos cambios son

$$\begin{aligned}A_{kk}^* &= c^2 A_{kk} + s^2 A_{\ell\ell} - 2 c s A_{k\ell} \\A_{\ell\ell}^* &= c^2 A_{\ell\ell} + s^2 A_{kk} + 2 c s A_{k\ell} \\A_{k\ell}^* &= A_{\ell k}^* = (c^2 - s^2) A_{k\ell} + c s (A_{kk} - A_{\ell\ell}) \\A_{ki}^* &= A_{ik}^* = c A_{ki} - s A_{\ell i} \quad i \neq k, i \neq \ell \\A_{\ell i}^* &= A_{i\ell}^* = c A_{\ell i} + s A_{ki} \quad i \neq k, i \neq \ell\end{aligned} \tag{10}$$

Diagonalización de Jacobi

El ángulo θ en la matriz de rotación de Jacobi se puede elegir de modo que $A_{k\ell}^* = A_{\ell k}^* = 0$.

Esto sugiere la siguiente idea: ¿Por qué no diagonalizar A haciendo un bucle a través de todos los términos por fuera de la diagonal y cero uno por uno?

Esto es exactamente lo que hace el método de Jacobi.

Sin embargo, hay un inconveniente importante: la transformación que aniquila un término fuera de la diagonal deshace algunos de los ceros previamente creados.

Afortunadamente, resulta que los términos fuera de la diagonal que reaparecen serán menores que antes.

Así, el método de Jacobi es un procedimiento iterativo que aplica repetidamente las rotaciones de Jacobi hasta que los términos fuera de la diagonal se anulan.

La matriz final de transformación P es la acumulación de rotaciones individuales R_i

$$P = R_1 \cdot R_2 \cdot R_2 \dots \quad (11)$$

Las columnas de P terminan siendo los vectores propios de A , y los elementos diagonales de $A^* = P^T A P$ se convierten en los valores propios.

Veamos un poco más a detalle la rotación de Jacobi.

De la ec. (10) vemos que $A_{k\ell}^* = 0$ si

$$(c^2 - s^2) A_{k\ell} + c s (A_{kk} - A_{\ell\ell}) = 0 \quad (12)$$

Usando las identidades trigonométricas

$$c^2 - s^2 = \cos^2 \theta - \sin^2 \theta = \cos 2\theta$$

$$c s = \cos \theta \sin \theta = \frac{1}{2} \sin 2\theta$$

La ec. (12) queda como

$$\tan 2\theta = -\frac{2A_{k\ell}}{A_{kk} - A_{\ell\ell}} \quad (13)$$

la cual podría ser resuelta para θ , seguida por el cálculo de $c = \cos \theta$ y $s = \sin \theta$.

Sin embargo, el procedimiento descrito a continuación conduce a un mejor algoritmo.

Introducimos la notación

$$\phi = \cot 2\theta = -\frac{A_{kk} - A_{\ell\ell}}{2 A_{k\ell}} \quad (14)$$

y usando la identidad trigonométrica

$$\tan 2\theta = \frac{2}{(1 - t^2)}$$

donde $t = \tan \theta$

La ec. (13) puede escribirse como

$$t^2 + 2\phi t - 1 = 0$$

la cual tiene raíces

$$t = -\phi \pm \sqrt{\phi^2 + 1}$$

Se ha encontrado que la raíz $|t| \leq 1$, que corresponde a $|\theta| \leq 45^\circ$, conduce a la transformación más estable.

Por lo tanto, elegimos el signo más si $\phi > 0$ y el signo menos si $\phi \leq 0$.

Lo que equivale a usar

$$t = \operatorname{sgn}(\phi) \left(-|\phi| + \sqrt{\phi^2 + 1} \right)$$

Lo que equivale a usar

$$t = \operatorname{sgn}(\phi) \left(-|\phi| + \sqrt{\phi^2 + 1} \right)$$

Para evitar un error excesivo por el redondeo si ϕ es grande, multiplicamos ambos lados de la ecuación por $|\phi| + \sqrt{\phi^2 + 1}$, lo que nos lleva a

$$t = \frac{\operatorname{sgn}(\phi)}{|\phi| + \sqrt{\phi^2 + 1}} \quad (15)$$

En el caso de que ϕ sea muy grande, debemos de sustituir la ec. (15) por la aproximación

$$t = \frac{1}{2\phi} \quad (16)$$

En el caso de que ϕ sea muy grande, debemos de sustituir la ec. (15) por la aproximación

$$t = \frac{1}{2\phi} \quad (16)$$

para prevenir un desborde en el cálculo de ϕ^2 .

Una vez calculado t , podemos usar la relación trigonométrica

$$\tan \theta = \sin \theta / \cos \theta = \sqrt{1 - \cos^2} / \cos \theta$$

Para obtener

$$c = \frac{1}{\sqrt{1 + t^2}}, \quad s = t c \quad (17)$$

Ahora podemos mejorar las fórmulas de transformación en las ecs. (10).

Resolviendo de la ec. (12) para $A_{\ell\ell}$, se obtiene

$$A_{\ell\ell} = A_{kk} + A_{k\ell} \frac{c^2 - s^2}{c s} \quad (18)$$

Re-emplazando todas las veces que aparece $A_{\ell\ell}$ de la ec. (18) y simplificando, las fórmulas de transformación de la ec. (10) se pueden escribir como

$$\begin{aligned}A_{kk}^* &= A_{kk} - t A_{k\ell} \\A_{\ell\ell}^* &= A_{\ell\ell} - t A_{k\ell} \\A_{k\ell}^* &= A_{\ell k}^* = 0 \\A_{ki}^* &= A_{ik}^* = A_{ki} - s A_{\ell i} + \tau A_{ki} \quad i \neq k, i \neq \ell \\A_{\ell i}^* &= A_{i\ell}^* = A_{\ell i} - s A_{ki} + \tau A_{\ell i} \quad i \neq k, i \neq \ell\end{aligned}\tag{19}$$

donde

$$\tau = \frac{s}{1 + c}\tag{20}$$

La introducción de τ nos permitió expresar cada fórmula en la forma, (valor original) + (cambio), que es útil para reducir el error por redondeo.

Al inicio del proceso de diagonalización de Jacobi, la matriz de transformación P se inicializa a la matriz de identidad.

Cada rotación de Jacobi cambia esta matriz de P a $P^* = P R$.

Los cambios correspondientes en los elementos de P puede demostrarse que son (sólo las columnas k y ℓ son afectadas)

$$\begin{aligned} P_{ik}^* &= P_{ik} - s(P_{i\ell} + \tau P_{ik}) \\ P_{i\ell}^* &= P_{i\ell} - s(P_{ik} + \tau P_{i\ell}) \end{aligned} \tag{21}$$

Todavía tenemos que decidir el orden en que los elementos fuera de la diagonal de A deben ser eliminados.

La idea original de Jacobi era atacar el elemento más grande porque hacerlo resulta en el menor número de rotaciones.

El problema aquí es que A se tiene que buscar para el elemento más grande antes de cada rotación, que es un proceso que consume mucho tiempo.

Si la matriz es grande, es más rápido recorrerla por filas o columnas y anular todos los elementos por encima de un valor **umbral**. En el siguiente barrido, se disminuye el umbral y el proceso se repite.

Hay varias maneras de elegir el **umbral**.

Calculamos la suma S de los elementos por encima de la diagonal principal de A :

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^n |A_{ij}| \quad (22)$$

Hay varias maneras de elegir el **umbral**.

Calculamos la suma S de los elementos por encima de la diagonal principal de A :

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^n |A_{ij}| \quad (22)$$

Dado que hay $n(n-1)/2$ de tales elementos, la magnitud promedio de los elementos fuera de la diagonal es

$$\frac{2S}{n(n-1)}$$

El **umbral** que usaremos es

$$\mu = \frac{0.5 S}{n(n-1)} \quad (23)$$

que representa la cuarta parte de la magnitud promedio de los elementos que están por fuera de la diagonal.

En resumen, el procedimiento de barrido de Jacobi (usa solamente la parte superior de la matriz), es el siguiente:

Calcular el **umbral** μ usando las ecs. (22) y (23)

Hacer un barrido de los elementos por fuera de la diagonal de **A**:

Si $|A_{ij}| \geq \mu$:

Calcular ϕ , t , c y s de las ecs. (14)-(17)

Calcular τ de la ec. (20)

Modificar los elementos de **A** de acuerdo a las ecs. (19)

Actualizar la matriz de transformación **P** usando las ecs. (21)

El barrido se realiza hasta que $\mu \leq \varepsilon$, donde ε es la tolerancia. Normalmente se requieren de 6-10 barridos para alcanzar la convergencia.