

# Problemas de valores propios

## Curso de Física Computacional

M. en C. Gustavo Contreras Mayén

Facultad de Ciencias - UNAM

24 de abril de 2018



1. Problemas de valores propios

2. Diagonalización de matrices

3. Método de Jacobi

4. El método de Jacobi

# 1. Problemas de valores propios

## 1.1 Introducción

## 1.2 Definición

# 2. Diagonalización de matrices

# 3. Método de Jacobi

# 4. El método de Jacobi

Muchos problemas fundamentales de la física e ingeniería se formulan naturalmente como problemas de valores propios.

En términos generales, una ecuación de autovalores es un operador homogéneo o ecuación matricial, que presenta un parámetro, cuyos valores admisibles, los autovalores, deben determinarse simultáneamente con las soluciones correspondientes, denominadas autovectores o funciones propias.

Por ejemplo, los vectores propios del tensor de inercia permiten colocar un cuerpo rígido en su sistema de ejes principales, lo que simplifica enormemente la descripción de su dinámica.

La caracterización de los modos propios de vibración de las estructuras complejas desempeña un papel clave en diversas áreas, que van desde la ingeniería estructural hasta la espectroscopia molecular.

Implica resolver el problema del valor propio de la llamada rigidez o matriz dinámica.

Incluso las relaciones de dispersión de fonones para las redes cristalinas son el resultado de problemas de autovalores similares.



La ecuación de Schrödinger, que es central para la mecánica cuántica, es en su tiempo independiente de una ecuación diferencial de autovalores.

Al expandir las funciones propias en términos de conjuntos completos de funciones de base, se transforma en un problema de valor propio para la matriz hamiltoniana simétrica real.

La estructura electrónica de los sistemas moleculares se calcula en química cuántica precisamente a partir de las soluciones de dichos problemas de autovalores de matriz.

Un problema matricial de valores propios es una representación de ecuaciones de operador.

Un vector no nulo  $\xi$  es llamado **valor propio (eigenvector)** de un operador lineal  $\mathcal{A}$  si, por la acción del operador, se transforma en un vector colineal

$$\mathcal{A} \xi = \lambda \xi \tag{1}$$

$$\mathcal{A} \xi = \lambda \xi$$

El escalar  $\lambda$  se le llama **valor propio** del operador  $\mathbf{A}$  correspondiente al **vector propio (eigenvector)**  $\xi$ .

Aparte de su magnitud, los valores propios son invariantes a la acción del operador.

Usando un conjunto completo de vectores base  $v_i, i = 1, \dots, b$ , se puede representar en  $\mathbb{R}^n$  el operador  $\mathcal{A}$  mediante la matriz asociada de  $n \times n$

$$\mathbf{A} = [a_{ij}]_{nn}$$

Teniendo como elementos el producto escalar  $a_{ij} = \langle v_i, \mathcal{A} v_j \rangle$ , y el vector propio  $\xi$ , por el vector columna  $\mathbf{x} = [x_i]_n$ , con  $x_i = \langle v_i, \xi \rangle$ .

# Problema de valores propios

La ecuación (1) se representa por el problema matricial de valores propios:

$$\mathbf{A} \cdot \mathbf{x} = \lambda \mathbf{x} \quad (2)$$

# Problema de valores propios

La ecuación (1) se representa por el problema matricial de valores propios:

$$\mathbf{A} \cdot \mathbf{x} = \lambda \mathbf{x} \quad (2)$$

o equivalentemente por el sistema lineal

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_n \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad (3)$$



# Re-escribiendo la ecuación

Re-escribiendo la ecuación (2) como

$$(\mathbf{A} - \lambda \mathbf{E}) \cdot \mathbf{x} = 0 \quad (4)$$

se enfatiza la *matrix característica*  $\mathbf{A} - \lambda \mathbf{E}$ , y el hecho de que el problema de valores propios, es un sistema lineal homogéneo.

# Sistema lineal homogéneo

El sistema tiene soluciones no triviales, si y sólo si, su determinante, conocido como *determinante característico*, es igual a cero:

$$\det(\mathbf{A} - \lambda \mathbf{E}) = 0 \quad (5)$$

# Ecuación característica

Entonces la *ecuación característica (secular)* tiene la forma:

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{21} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix} = 0 \quad (6)$$

# La ecuación secular

Siendo equivalente a una ecuación polinomial de orden  $n$  en  $\lambda$ , la ecuación secular tiene en principio  $n$  raíces reales o complejas  $\lambda_1, \lambda_2, \dots, \lambda_n$ , distintas o no, que son precisamente los valores propios de la matriz  $A$ .

# Cálculo de las raíces

El cálculo de las raíces del determinante característico expandido se puede realizar mediante cualquier método general para ecuaciones polinomiales.

El conjunto de todos los valores propios forma el *espectro* de la matriz  $A$ .

# Cálculo de los vectores propios

Una vez que se conocen los valores propios, para obtener los vectores propios se logra de la manera más elemental mediante la sustitución en secuencia cada valor propio  $\lambda_j$  en el sistema homogéneo lineal (3)

# Cálculo de los vectores propios

La solución obtenida (expresada en términos de un componente arbitrario) representa el vector propio correspondiente  $\mathbf{x}^{(j)}$  de la matriz  $\mathbf{A}$ .

Se puede proceder de manera análoga para todos los autovalores, calculando los vectores propios respectivos.

Pero en la práctica, sin embargo, hay enfoques numéricos mucho más eficientes.

Denotando por  $\mathbf{X}$  la matriz que tiene como columnas los vectores propios  $\mathbf{x}^{(j)}$ , las  $n$  ecuaciones matriciales (3) resultantes de la sustitución de los valores propios  $\lambda_j$  y de los vectores propios  $\mathbf{x}^{(j)}$  ( $j = 1, 2, \dots, n$ ) se puede escribir sintéticamente como:



$$\mathbf{A} \cdot \mathbf{X} = \mathbf{X} \cdot \mathbf{\Lambda} \quad (7)$$

# Matriz modal

$$\mathbf{A} \cdot \mathbf{X} = \mathbf{X} \cdot \Lambda \quad (7)$$

donde

$$\mathbf{X} = \begin{vmatrix} \mathbf{x}_1^{(1)} & \mathbf{x}_1^{(2)} & \mathbf{x}_1^{(n)} \\ \mathbf{x}_2^{(1)} & \mathbf{x}_2^{(2)} & \mathbf{x}_2^{(n)} \\ \mathbf{x}_n^{(1)} & \mathbf{x}_n^{(2)} & \mathbf{x}_n^{(n)} \end{vmatrix}, \Lambda = \begin{vmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_n \end{vmatrix} \quad (8)$$

La matriz  $\mathbf{X}$  se le llama **matriz modal** y la ecuación (7) es llamada **ecuación modal**.

# Propiedad de la matriz modal

Para una matriz  $A$  de “bien comportada”, es sencillo demostrar que la matriz modal  $X$  está compuesta de columnas linealmente independientes (como vectores propios de  $A$ ), por lo tanto es no singular ( $\det X \neq 0$ ) y admite una matriz inversa:  $X^{-1}$ .

# Propiedad de la matriz modal

En consecuencia, la ecuación modal puede ser escribir como

$$\mathbf{X}^{-1} \cdot \mathbf{A} \cdot \mathbf{X} = \mathbf{\Lambda} \quad (9)$$

# Propiedad de la matriz modal

En consecuencia, la ecuación modal puede ser escribir como

$$\mathbf{X}^{-1} \cdot \mathbf{A} \cdot \mathbf{X} = \Lambda \quad (9)$$

Siempre que se conozcan efectivamente los vectores propios de  $\mathbf{A}$ , los valores propios correspondientes son los elementos diagonales de la matriz

$$\mathbf{X}^{-1} \cdot \mathbf{A} \cdot \mathbf{X}$$

# 1. Problemas de valores propios

## 2. Diagonalización de matrices

### 2.1 Diagonalización por transformaciones

## 3. Método de Jacobi

## 4. El método de Jacobi

Las matrices asociadas al mismo operador lineal  $\mathcal{A}$  relativas a bases diferentes, se les llama *similares*.



# Un par de teoremas

Se puede demostrar que:

## (Primer teorema)

*Dos matrices reales  $A$  de  $n \times n$  y  $A' \in \mathbf{M}_{\mathbb{R}}^{n \times n}$ , son similares si y sólo si, están relacionadas por una transformación de similaridad:*

$$A' = S^{-1} \cdot A \cdot S \quad (10)$$

*donde  $S \in \mathbf{M}_{\mathbb{R}}^{n \times n}$  es una matriz invertible.*

# Del primer teorema

De hecho, suponiendo que la acción del operador se describe con relación a una primera base por la matriz  $\mathbf{A}$  y con respecto a una segunda base por la otra matriz  $\mathbf{A}'$ , tenemos

$$\mathbf{y} = \mathbf{A} \cdot \mathbf{x}, \quad \mathbf{y}' = \mathbf{A}' \cdot \mathbf{x}' \quad (11)$$

# Del primer teorema

Donde  $x$  y  $x'$  de un lado, y por el otro  $y$ ,  $y'$ , son las representaciones de las columnas de la matriz con los mismos vectores relativos a la primera y segunda base, respectivamente.

# Del primer teorema

Si  $S$  es una matriz no singular que realiza el cambio de base, entonces:

$$\mathbf{x} = \mathbf{S} \cdot \mathbf{x}', \quad \mathbf{y} = \mathbf{S} \cdot \mathbf{y}' \quad (12)$$

# Del primer teorema

Sustituyendo las relaciones (12) en la primera de las ecuaciones (11), se sigue que

$$\mathbf{y}' = (\mathbf{S}^{-1} \cdot \mathbf{A} \cdot \mathbf{S}) \cdot \mathbf{x}'$$

al comparar éste resultado con la segunda de las ecuaciones (11), se obtiene la ec. (10), con lo que se cumple la condición necesaria y suficiente para que las dos matrices sean similares.

# Segundo teorema

Consideremos el siguiente teorema

(Segundo teorema)

*Dos matrices similares tienen valores propios idénticos.*

La demostración es sencilla, por lo que se omitirá

# Consecuencia de los teoremas

Como consecuencia de los teoremas tenemos que: dos matrices similares tienen valores propios idénticos y los vectores propios correspondientes se pueden deducir uno de otro por medio de una transformación de similaridad.

Específicamente, a partir de la reformulación (9) de la ecuación modal se deduce que las matrices  $\mathbf{A}$  y  $\mathbf{\Lambda}$  son similares.



# Consecuencia de los teoremas

Específicamente, a partir de la reformulación (9) de la ecuación modal se deduce que las matrices  $A$  y  $\Lambda$  son similares.

Las matrices similares con matrices diagonales se denominan **diagonalizables**.

# Consecuencia de los teoremas

La esencia de la diagonalización de matrices consiste en el hecho de que, siempre que se encuentre una matriz de transformación que diagonalice la matriz considerada, los valores propios buscados estarán en la diagonal principal de la matriz transformada.

# Consecuencia de los teoremas

Mientras que las columnas de la matriz de transformación (que pueden identificarse con la matriz modal) representarán con precisión los vectores propios correspondientes de la matriz original.

1. Problemas de valores propios

2. Diagonalización de matrices

3. Método de Jacobi

3.1 Aprovechando la simetría de las matrices

4. El método de Jacobi

# Aprovechando la simetría de las matrices

Como se mencionó anteriormente, la solución de problemas de valores propios para matrices simétricas reales es de gran interés en muchas áreas de la física y la ingeniería.

# Aprovechando la simetría de las matrices

La utilización efectiva de la simetría de las matrices involucradas conduce a una considerable simplificación de los métodos numéricos desarrollados, que de esta manera ganan estabilidad, precisión y eficiencia (tanto en términos de costo de operación como de uso de memoria).

# Aprovechando la simetría de las matrices

Una propiedad esencial de los vectores propios de las matrices simétricas reales es que son reales y ortogonales.

Como consecuencia, la matriz modal  $\mathbf{X}$  tiene columnas ortogonales.

# Aprovechando la simetría de las matrices

La diagonalización de una matriz simétrica real se puede realizar por medio de una *matriz de transformación ortogonal*  $\mathbf{R}$ , definida por

$$\mathbf{R} \cdot \mathbf{R}^T = \mathbf{R}^T \cdot \mathbf{R} = \mathbf{E}$$



# Aprovechando la simetría de las matrices

El mayor beneficio de esta propiedad de ortogonalidad es que permite que la costosa operación para calcular la inversa de la matriz  $\mathbf{R}$ , sea reemplazada en la ecuación modal (9) por la operación de transposición significativamente más simple

$$\mathbf{R}^{-1} = \mathbf{R}^T$$

# Diagonalización de la matriz

Así pues, la diagonalización de la matriz  $\mathbf{A}$  se logra mediante la *transformación de similitud ortogonal*

$$\mathbf{R}^T \cdot \mathbf{A} \cdot \mathbf{R} = \Lambda \quad (13)$$

1. Problemas de valores propios

2. Diagonalización de matrices

3. Método de Jacobi

4. El método de Jacobi

4.1 Definición del método

# El método de Jacobi

Podemos considerar que el método de Jacobi es un algoritmo iterativo infinito.

Consiste esencialmente en aplicar una secuencia de transformaciones ortogonales a la matriz  $A$ , con la intención de reducir gradualmente sus elementos fuera de la diagonal hasta que desaparezcan numéricamente.

# El método de Jacobi

Se considera utilizar rotaciones bidimensionales, que son las transformaciones ortogonales más simples.

Las transformaciones elementales tienen la forma de rotaciones planas que implican pares de coordenadas y apuntan a la puesta a cero los elementos simétricos fuera de la diagonal.

# El método de Jacobi

Como un efecto secundario perjudicial, la acción de las rotaciones elementales no puede restringirse a los dos elementos dirigidos y sustituir los ceros fuera de la diagonal creados previamente.

Sin embargo, se produce una reducción general gradual de los elementos fuera de la diagonal, hasta que la matriz transformada se vuelve diagonal en el límite de los errores de redondeo y se puede identificar con la matriz  $\Lambda$ .

Al final, el producto de las sucesivas matrices de rotación proporciona la matriz modal, que tiene en las columnas los vectores propios de la matriz  $A$ , que a su vez corresponden a los valores propios ubicados en la diagonal de la matriz  $\Lambda$ .



# Desarrollo del método de Jacobi

Consideremos por simplicidad el caso de una matriz  $A$  de  $2 \times 2$ .

# Desarrollo del método de Jacobi

Consideremos por simplicidad el caso de una matriz  $\mathbf{A}$  de  $2 \times 2$ .

La transformación ortogonal se puede lograr mediante la matriz de rotación

$$\mathbf{R} = \begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix} \quad (14)$$

La matriz  $\mathbf{R}$  rota el sistema de los vectores base en un ángulo  $\varphi$  (vectores unitarios en el sistema cartesiano)

# Desarrollo del método de Jacobi

La matriz de rotación así definida  $\mathbf{R}$  es ortogonal independiente del ángulo  $\varphi$  y, por lo tanto, se puede imponer la condición adicional en  $\varphi$  que la transformación  $\mathbf{R}$  diagonaliza la matriz  $\mathbf{A}$ .

# Desarrollo del método de Jacobi

La matriz resultante puede escribirse formalmente

$$\mathbf{A}' = \mathbf{R}^T \cdot \mathbf{A} \cdot \mathbf{R}$$

y sus elementos equivalen a

$$a'_{11} = a_{11} \cos^2 \varphi + 2 a_{21} \sin \varphi \cos \varphi + a_{22} \sin^2 \varphi,$$

$$a'_{22} = a_{11} \sin^2 \varphi - 2 a_{21} \sin \varphi \cos \varphi + a_{22} \cos^2 \varphi,$$

$$a'_{21} = a_{21} (\cos^2 \varphi - \sin^2 \varphi) + (a_{22} - a_{11}) \sin \varphi \cos \varphi = a'_{12}$$

# Uso de una propiedad

Aquí hemos utilizado el hecho de que la transformación ortogonal conserva el carácter simétrico de la matriz.

# Forzando una condición

Hacinedo que los elementos simétricos fuera de la diagonal de la matriz  $A'$  se anulen ( $a'_{21} = a'_{12} = 0$ ), obtenemos la siguiente ecuación algrabica para  $\cot \varphi$ :

$$\cot^2 \varphi + \frac{a_{22} - a_{11}}{a_{21}} \cot \varphi - 1 = 0$$

Una de las formas matemáticamente equivalentes de expresar la solución, que revelará sus virtudes más adelante, es

$$\tan \varphi = \left[ \frac{a_{11} - a_{22}}{2 a_{21}} \pm \sqrt{\left( \frac{a_{11} - a_{22}}{2 a_{21}} \right)^2 + 1} \right]^{-1}$$



# Las matrices $A'$ y $R$

En lugar de expresar desde aquí el ángulo  $\varphi$  en sí mismo, uno puede determinar numéricamente (de manera más eficiente) los valores de la función  $\cos \varphi = (1 + \tan \varphi)^{-1/2}$  y  $\sin \varphi = \tan \varphi \cos \varphi$ , que son necesarios para especificar tanto las matrices  $R$  y  $A'$ .

# Los valores propios de $A$

Los valores propios de  $A$  se obtienen evaluando los elementos de la diagonal de la matriz  $A'$  transformada, usando los  $\cos \varphi$  y  $\sin \varphi$  determinados.

# Los vectores propios de $A$

Mientras que los vectores propios correspondientes son las columnas de la matriz de rotación  $R$ :

$$\lambda_1 = a'_{11}, \quad \mathbf{x}^{(1)} = \begin{bmatrix} \cos \varphi \\ \sin \varphi \end{bmatrix}$$
$$\lambda_2 = a'_{22}, \quad \mathbf{x}^{(2)} = \begin{bmatrix} -\sin \varphi \\ \cos \varphi \end{bmatrix}$$

# Los vectores propios de $A$

Mientras que los vectores propios correspondientes son las columnas de la matriz de rotación  $R$ :

$$\lambda_1 = a'_{11}, \quad \mathbf{x}^{(1)} = \begin{bmatrix} \cos \varphi \\ \sin \varphi \end{bmatrix}$$
$$\lambda_2 = a'_{22}, \quad \mathbf{x}^{(2)} = \begin{bmatrix} -\sin \varphi \\ \cos \varphi \end{bmatrix}$$

En resumen: la diagonalización de matrices  $(2 \times 2)$  se puede realizar exactamente, **mediante una única transformación ortogonal.**

# Generalización de la matriz de rotación

En el espacio  $n$ -dimensional, la matriz de rotación puede generalizarse para operar con respecto al eje definido por el par de coordenadas  $i$  y  $j$ :

# Generalización de la matriz de rotación

$$\mathbf{R}(i, j) = \begin{bmatrix} 1 & \vdots & & \vdots & 0 \\ \dots & \cos \varphi & \dots & -\sin \varphi & \dots \\ & \vdots & \ddots & \vdots & \\ \dots & \sin \varphi & \dots & \cos \varphi & \dots \\ 0 & \vdots & & \vdots & 1 \end{bmatrix} \quad (15)$$

columna  $i$     columna  $j$

# Elementos no nulos

Difiriendo de la matriz unitaria sólo por los cuatro elementos distintos de cero en  $a_{ii} = a_{jj} = \cos \varphi$  y en  $a_{ij} = -a_{ji} = -\sin \varphi$ , ubicado en la intersección de las filas  $i$  y las columnas  $j$ .

# La matriz de rotación

La matriz de rotación  $\mathbf{R}(i, j)$  está destinada para llevar a cabo la transformación ortogonal

$$\mathbf{A}' = \mathbf{R}^T(i, j) \cdot \mathbf{R}(i, j) \quad (16)$$

que hace que los elementos fuera de la diagonal  $a'_{ij}$  y  $a'_{ji}$  se anulen.



# Afectación de renglones

Los únicos elementos de  $A'$  que difieren de los de  $A$  se encuentran en las filas  $i$  y columnas  $j$ .

Ya que al multiplicar  $A$  por la izquierda con  $R^T(i, j)$ , sólo se modifican los renglones  $i$  y  $j$ .

# Afectación de columnas

Mientras que por la posterior multiplicación por la derecha con  $\mathbf{R}(i, j)$ , sólo las columnas  $i$  y  $j$  se ven afectadas.

# Afectación de columnas

Mientras que por la posterior multiplicación por la derecha con  $\mathbf{R}(i, j)$ , sólo las columnas  $i$  y  $j$  se ven afectadas.

Definiendo la matriz intermedia  $\tilde{\mathbf{A}}$ , podemos reescribir las transformaciones (16) como

$$\tilde{\mathbf{A}} = \mathbf{A} \cdot \mathbf{R}(i, j), \quad \mathbf{A}' = \mathbf{R}^T(i, j) \cdot \tilde{\mathbf{A}} \quad (17)$$