

Cálculo de valores y vectores propios de una matriz

Pablo Andrés Úsuga Gómez, Juan Pablo Robledo Meza, Camilo Montoya Arango
Mateo Sebastián Mora Montero, Daniel Andrés Hernandez Pedraza

Diciembre 6 de 2024

Contexto del problema: Dada una matriz $A \in \mathbb{R}^{n \times n}$, queremos hallar vectores $x \neq 0$ y escalares $\lambda \in \mathbb{R}$ tales que

$$Ax = \lambda x$$

1. Método del polinomio característico

Geométricamente estos vectores se mapean de manera paralela a ellos. Para una solución de este tipo, decimos que λ es un valor propio de A , y que x es un vector propio de A asociado a λ . Supongamos que tenemos un vector $x \neq 0$ y un escalar λ que satisfacen dicha igualdad, entonces $(A - \lambda I_n)x = 0$, donde I_n es la matriz identidad de $n \times n$. De este modo, x está en el espacio nulo de la matriz $(A - \lambda I_n)$, este lo denotaremos por $E_\lambda := \text{Nul}(A - \lambda I_n)$. También nos referiremos a él como espacio del valor propio λ . Como $x \neq 0$ y $x \in E_\lambda$, entonces dicha matriz no es invertible, pues su espacio nulo no es trivial, esto equivale a decir $\det(A - \lambda I_n) = 0$. Por tanto, dado un escalar $\lambda \in \mathbb{R}$, este es valor propio de A si y solo si λ es solución a la ecuación $\det(A - \lambda I_n) = 0$, obteniendo así una caracterización algebraica de los valores propios de la matriz A , de hecho, sabemos que $\det(A - \lambda I_n)$ es un polinomio de grado n , y este recibe un nombre especial.

Polinomio característico de una matriz: Dada una matriz $A \in \mathbb{R}^{n \times n}$, definimos el polinomio característico de A como

$$p_A(\lambda) := \det(A - \lambda I_n)$$

Para hallar los valores propios de A resolvemos la ecuación característica $p_A(\lambda) = 0$, y las raíces reales de esta serán los valores propios de A . Un resultado inmediato, pero muy importante, es que una matriz $A \in \mathbb{R}^{n \times n}$ tiene a lo sumo n valores propios distintos, que corresponde a la cantidad de raíces que puede tener un polinomio de grado n con coeficientes en un cuerpo, en este caso el cuerpo es \mathbb{R} .

Suponga que ya encontró las raíces reales del polinomio característico $p_A(\lambda)$, el proceso para hallar los vectores propios de A , asociados a cada valor propio λ , consiste en hallar el conjunto E_λ de vectores que solucionan la ecuación $(A - \lambda I_n)x = 0$, y esto debe hacerse para cada uno de los valores propios para encontrar sus respectivos vectores propios asociados.

El proceso completo del método del polinomio característico para encontrar los valores y vectores propios de una matriz $A \in \mathbb{R}^{n \times n}$ se puede resumir como:

1. Encuentre las raíces reales del polinomio característico $p_A(\lambda)$, estas serán los valores propios de la matriz A .
2. Para cada valor propio λ encontrado en el paso anterior, resuelva el sistema $(A - \lambda I_n)x = 0$, los vectores x solución del sistema serán los vectores propios asociados al valor propio λ .

Análisis computacional del método del polinomio característico

Vimos que teóricamente el método es completamente correcto, pues el método surge a partir de las relaciones algebraicas que deben satisfacerse entre valores propios, vectores propios y la matriz A , sin embargo, en la práctica nos vamos a encontrar con muchos inconvenientes para llevarlo a cabo de una manera eficiente, computacionalmente hablando, más los posibles errores numéricos en las aproximaciones de los resultados.

De base, tenemos que hallar el polinomio característico de la matriz A , calculando el determinante de $(A - \lambda I_n)$, esto en sí mismo es una tarea costosa cuando tratamos de matrices de un gran tamaño, de hecho, más adelante estudiaremos el método de descomposición QR, que esencialmente intenta resolver el problema de hallar determinantes de un modo rápido factorizando la matriz de una manera conveniente.

Luego de hallar el polinomio característico, debemos emplear algún método numérico para hallar las raíces reales del mismo. Si el polinomio es de grado muy bajo, 2, 3, 4, existen fórmulas explícitas de las raíces, pero incluso con estas fórmulas habría que analizar con detalle la propagación de errores en las operaciones del cálculo. En un caso más general donde estas fórmulas no existen, uno puede intentar combinaciones de métodos como buscar cambios de signo del polinomio en intervalos pequeños y luego usar bisección o método de Newton con posibilidades de divergencia si la longitud del intervalo escogido es muy grande y por tanto la aproximación inicial pueda estar muy lejos de la raíz.

Otro problema ligado a encontrar las raíces del polinomio característico, es que, a priori, no sabemos cuántas raíces hay, solo tenemos una cota superior que es el grado del polinomio, pero pueden darse casos extremos donde tengamos un polinomio de grado muy alto con muy pocas raíces reales, y por tanto no tenemos una manera concreta de discriminar en un algoritmo de búsqueda cuándo dejar de buscar raíces.

Por último, las raíces que encontramos nos dan valores propios de la matriz, pero ahora nos falta hallar los vectores propios asociados a estos, y para esta tarea hay que resolver el sistema $(A - \lambda I_n)x = 0$, donde una vez más necesitamos de otro método numérico para resolver esta tarea específica, que en general también es muy costosa para matrices de gran tamaño.

2. Método de descomposición QR

La descomposición QR es un método numérico para encontrar los valores propios de una matriz cuadrada A . Consiste en descomponer A como:

$$A = QR,$$

donde Q es una matriz ortogonal ($Q^T Q = I$) y R es una matriz triangular superior.
Para calcular los valores propios:

1. Comienza con una matriz A .

2. Realiza la descomposición QR para encontrar Q y R .
3. Actualiza A usando la relación $A \leftarrow RQ$.
4. Repite este proceso iterativamente hasta que A converja a una matriz triangular superior. Los valores propios de A aparecerán entonces como las entradas diagonales de la matriz resultante.

2.1. Desplazamiento Espectral

El desplazamiento espectral es una técnica utilizada en el método QR para mejorar su convergencia hacia los valores propios. En lugar de aplicar directamente el proceso QR a la matriz A , se modifica iterativamente la matriz restando un múltiplo de la identidad:

$$A_k - \mu_k I,$$

donde μ_k es conocido como el desplazamiento espectral y I es la matriz identidad. Después de realizar la descomposición QR, se restaura el desplazamiento agregando $\mu_k I$ al resultado:

$$A_{k+1} = R_k Q_k + \mu_k I.$$

¿Por qué se utiliza el desplazamiento espectral? La idea detrás del desplazamiento espectral es acelerar la convergencia del método QR, concentrando la convergencia hacia valores propios dominantes o separando valores propios cercanos. Al elegir μ_k de manera adecuada, se minimizan las operaciones necesarias para triangularizar A .

2.1.1. Elección de μ_k :

Existen varias estrategias para seleccionar μ_k :

- Diagonal inferior derecha: Una opción común es elegir $\mu_k = A[-1, -1]$, el elemento diagonal inferior derecho de A . Esto ayuda a acercar la matriz hacia una forma triangular.
- Máximo de la diagonal: $\mu_k = \max(\text{diag}(A))$, lo cual puede ser útil cuando el valor propio dominante es conocido.
- Submatriz 2×2 : Para matrices reales y simétricas, a menudo se selecciona μ_k como uno de los valores propios de una submatriz 2×2 en la esquina inferior derecha de A . Esto se denomina desplazamiento de Wilkinson y mejora la precisión.

2.1.2. Beneficios del desplazamiento espectral:

- Reduce el número de iteraciones necesarias para la convergencia.
- Facilita la separación de valores propios cercanos.
- Mejora la estabilidad numérica del algoritmo QR.

2.1.3. Ejemplo del uso de desplazamiento espectral:

Supongamos que queremos calcular los valores propios de la matriz:

$$A = \begin{bmatrix} 6 & 2 \\ 2 & 3 \end{bmatrix}.$$

1. Seleccionamos un desplazamiento inicial $\mu_0 = A[-1, -1] = 3$. 2. Modificamos la matriz como:

$$A_0 = A - \mu_0 I = \begin{bmatrix} 6 & 2 \\ 2 & 3 \end{bmatrix} - 3 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 2 & 0 \end{bmatrix}.$$

3. Realizamos la descomposición QR de A_0 para obtener Q y R , luego actualizamos A_1 :

$$A_1 = RQ + \mu_0 I.$$

4. Repetimos este proceso hasta que A_k converja a una matriz triangular superior.

2.1.4. Limitaciones:

Aunque el desplazamiento espectral es muy útil, debe elegirse cuidadosamente para evitar problemas numéricos:

- Si μ_k se selecciona de manera inadecuada, la matriz puede divergir en lugar de converger.
- En algunos casos, un desplazamiento adaptativo puede ser necesario, evaluando μ_k en función del progreso de las iteraciones.

2.2. Métodos de Descomposición QR

Gram-Schmidt: El método Gram-Schmidt construye una base ortonormal a partir de las columnas de A . A partir de esta base se genera la matriz Q , mientras que R se obtiene como:

$$R = Q^T A.$$

Este método es sencillo de implementar, pero es susceptible a inestabilidades numéricas debido a la acumulación de errores de redondeo.

Householder: Este método utiliza reflexiones de Householder para triangular la matriz A , eliminando las entradas debajo de la diagonal de manera iterativa. Es más estable numéricamente que Gram-Schmidt, especialmente para matrices mal condicionadas.

Reflexión de Householder: La reflexión de Householder transforma un vector x en otro vector alineado con un eje específico (por ejemplo, $e_1 = [1, 0, 0, \dots]^T$), anulando las demás componentes. La matriz de reflexión se define como:

$$H = I - 2 \frac{vv^T}{v^T v},$$

donde:

- $v = x - \|x\|e_1$ es el vector de Householder.

- $\|x\|$ es la norma euclidiana del vector x .
- H es ortogonal ($H^T H = I$) y simétrica ($H^T = H$).

Algoritmo de Householder: El método de Householder se basa en las siguientes etapas:

1. Para cada columna k de la matriz A , selecciona el vector x , que corresponde a las entradas de la columna k por debajo (e incluyendo) de la diagonal.
2. Calcula el vector de Householder $v = x - \|x\|e_1$, y normalízalo:

$$v = \frac{v}{\|v\|}.$$

3. Construye la matriz de reflexión:

$$H_k = I - 2 \frac{vv^T}{v^T v}.$$

4. Aplica H_k a la matriz A :

$$A' = H_k A.$$

Esto elimina las entradas debajo de la diagonal en la columna k .

5. Repite este proceso para cada columna k , obteniendo una matriz triangular superior R y una acumulación de reflexiones que forman $Q = H_1 H_2 \cdots H_{n-1}$.

2.3. Transformación a Forma de Hessenberg

Antes de aplicar el método QR, muchas implementaciones convierten la matriz A a su forma de Hessenberg para reducir el costo computacional y acelerar la convergencia.

Definición: Una matriz H está en forma de Hessenberg si tiene ceros en todas las posiciones por debajo de la subdiagonal. Esto significa que H tiene la siguiente estructura:

$$H = \begin{bmatrix} h_{1,1} & h_{1,2} & h_{1,3} & \cdots & h_{1,n} \\ h_{2,1} & h_{2,2} & h_{2,3} & \cdots & h_{2,n} \\ 0 & h_{3,2} & h_{3,3} & \cdots & h_{3,n} \\ 0 & 0 & h_{4,3} & \cdots & h_{4,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & h_{n,n} \end{bmatrix}.$$

Proceso de transformación: La forma de Hessenberg se obtiene mediante transformaciones ortogonales, como las reflexiones de Householder. Estas transformaciones eliminan las entradas debajo de la subdiagonal de A , preservando sus valores propios.

El proceso es iterativo:

1. Para cada columna k de la matriz A , construimos un vector de reflexión v que anula todas las entradas debajo de la subdiagonal.
2. Construimos la matriz de reflexión $H_k = I - 2 \frac{vv^T}{v^T v}$.

3. Aplicamos H_k a A para obtener una nueva matriz con ceros debajo de la subdiagonal en la columna k .

Ventajas:

- La transformación a Hessenberg reduce significativamente el costo computacional del método QR.
- Facilita la convergencia, especialmente para matrices grandes.

2.4. Justificación Matemática

El proceso iterativo $A \leftarrow RQ$ transforma la matriz acercándola a una forma triangular o diagonal en cada iteración mientras preserve sus valores propios. Esto se debe a que A , Q , y R son matrices semejantes, es decir:

$$A_{k+1} = Q^T A_k Q.$$

Por lo tanto, los valores propios de A se conservan en A_{k+1} .

Cuando se utiliza la forma de Hessenberg, el número de operaciones necesarias para calcular QR y RQ se reduce, ya que muchas entradas son cero.

2.5. Suposiciones para la Convergencia

- La matriz A debe ser cuadrada.
- La convergencia está garantizada si A es simétrica o normal (i.e., $A^T A = A A^T$).
- Para matrices generales, reducir A a forma de Hessenberg ayuda a garantizar la estabilidad y eficiencia.
- Si se usa desplazamiento espectral, debe elegirse cuidadosamente para evitar problemas numéricos.

3. Método de las potencias

3.1. Teoría

La teoría espectral de matrices comenzó a tomar forma en el siglo XIX gracias a matemáticos como Augustin-Louis Cauchy, Arthur Cayley, James Joseph Silvestre, estos personajes sentaron las bases teóricas para que más adelante se trabajará en el desarrollo de este algoritmo. El método de las potencias se consolidó a principios del siglo XX, impulsado por la necesidad de resolver problemas en mecánica, física matemática y análisis estructural, donde los valores propios de matrices representan frecuencias naturales, tensiones, y otras propiedades. En particular, Google lo emplea para calcular el PageRank de los documentos en su motor de búsqueda.

Este método es útil para encontrar los valores propios extremos de una matriz, es decir, los valores propios que tienen el módulo más grande y más pequeño y de ahí llegar a sus vectores propios asociados.

3.2. Explicación detallada

3.2.1. Requerimientos

Para aplicar el método de las potencias, se supone que la matriz A de $n \times n$ tiene n valores característicos $\lambda_1, \lambda_2, \dots, \lambda_n$, y además, cada uno tiene sus vectores propios correspondientes linealmente independientes. También suponemos que alguno de esos λ es mayor que los demás.

3.2.2. Descripción del método

Es un método iterativo de aproximaciones sucesivas.

1. Tomamos una matriz $n \times n$ (cuadrada y, preferiblemente, diagonalizable). Además, debemos definir una tolerancia preestablecida.
2. El método nos propone empezar con un vector no nulo y multiplicarlo por la matriz original.
3. Multiplicamos la matriz $n \times n$ por el vector inicial, llamado \mathbf{x}_0 .
4. El resultado de esta multiplicación será un vector llamado \mathbf{x}_1 .
5. Normalizamos el vector resultante \mathbf{x}_1 .
6. Multiplicamos el vector \mathbf{x}_1 con la matriz original.
7. Repetimos el proceso iterativamente hasta que la diferencia entre dos aproximaciones cumpla con la tolerancia preestablecida.

De esta manera, llegaremos al valor propio más grande asociado a la matriz original $n \times n$.

¿Y para el menor valor propio?

Para encontrar el menor valor propio, utilizamos la matriz inversa $n \times n$ de la matriz original y seguimos los pasos descritos anteriormente

3.3. Formulación Matemática del Método de la Potencia

Consideraciones Iniciales Consideremos una matriz A de dimensión $n \times n$ con un conjunto de valores propios $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ y sus correspondientes vectores propios $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$, donde los valores propios están ordenados por magnitud:

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|$$

3.3.1. Fundamentos Teóricos

Descomposición Espectral: Todo vector inicial \mathbf{x}_0 puede representarse como una combinación lineal de los vectores propios:

$$\mathbf{x}_0 = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_n \mathbf{v}_n$$

Proceso Iterativo: Iteraciones sucesivas están dadas por:

$$\mathbf{x}_{k+1} = A \cdot \mathbf{x}_k$$

Después de múltiples iteraciones, los términos correspondientes a valores propios menores se anulan. Supongamos que:

$$\mathbf{x}_0 = [c_1 \mathbf{v}_1, c_2 \mathbf{v}_2, \dots, c_n \mathbf{v}_n]^T$$

Tras k iteraciones:

$$\mathbf{x}_k = c_1 \lambda_1^k \mathbf{v}_1 + c_2 \lambda_2^k \mathbf{v}_2 + \dots + c_n \lambda_n^k \mathbf{v}_n$$

Si $|\lambda_1| > |\lambda_2|$, entonces λ_1 dominará la iteración.

3.4. Observación

Cualquier vector inicial \mathbf{x}_0 puede expresarse como una combinación lineal de los vectores propios de la matriz A :

$$\mathbf{x}_0 = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \cdots + c_n \mathbf{v}_n$$

Al multiplicar \mathbf{x}_0 por la matriz A , utilizamos la propiedad de que A multiplicado por un vector propio \mathbf{v}_i da el valor propio λ_i asociado al vector propio \mathbf{v}_i :

$$A \cdot \mathbf{v}_i = \lambda_i \mathbf{v}_i$$

Esto nos lleva a:

$$\mathbf{x}_k = \sum_{i=1}^n c_i \lambda_i^k \mathbf{v}_i$$

Tras p iteraciones, obtenemos:

$$\mathbf{x}_p = A^p \cdot \mathbf{x}_0$$

Donde λ_1 (el valor propio de mayor magnitud) dominará la expresión conforme $p \rightarrow \infty$.

3.4.1. Ventajas del método de la potencia

- Es eficiente para grandes grafos dispersos, como el de la web, ya que solo requiere operaciones de multiplicación matriz-vector.
- La convergencia es rápida si el grafo está bien estructurado y cumple ciertas propiedades (conectividad y un nodo de salida global).

4. Referencias

1. Datta B. N. (2010). *Numerical Linear Algebra and Applications* (2nd ed., Chapter 7). Society for Industrial and Applied Mathematics.
2. Burden, R.L. and Faires, J.D. (2011) *Numerical Analysis*. 9th Edition, Brookscole, Boston.
3. Hoffman, J. D., Frankel, S. (2018). *Numerical methods for engineers and scientists*. CRC press.

"El secreto mejor guardado del universo es la convergencia del presente"
