

Universidad Central de Venezuela
Facultad de Ingeniería
Escuela de Ingeniería Eléctrica
Departamento de Electrónica, Computación y Control

Autovalores y Autovectores

Cálculo Numérico (2514)

Br. Mendoza Romina

Marzo 2024

Resumen

En el siguiente documento contiene lo mínimo necesario para discutir y profundizar, en clase, de los temas: Autovalores y autovectores, y el método de potencia.

1 Autovalores y Autovectores

Para determinar los autovalores de una matriz $A_{n \times n}$, construimos el polinomio característico

$$p(\lambda) = \det(A - \lambda I) \quad (1)$$

y, entonces, se determinan sus ceros. Encontrar el determinante de una matriz $n \times n$ es caro desde el punto de vista computacional y hallar buenas aproximaciones para las raíces de $p(\lambda)$ también es difícil.

En el siguiente documento se hablará para resolver un sistema lineal hay una técnica iterativa que convergerá si todos los autovalores asociados con el problema tienen magnitud menor que 1. Los valores exactos de los autovalores en este caso no son muy importantes, sólo la región de un plano complejo en el que se encuentran.

1.1 Teorema 1: Círculo de Geršgorin

Sea A una matriz $n \times n$ y R_i denota el círculo en el plano complejo con centro a_{ii} y radio $\sum_{j=1, j \neq i}^n |a_{ij}|$; es decir,

$$R_i = \left\{ z \in \mathbb{C} : |z - a_{ii}| \leq \sum_{j=1, j \neq i}^n |a_{ij}| \right\}, \quad (2)$$

donde \mathbb{C} denota el plano complejo. Los autovalores de A están contenidos en la unión de estos círculos, $R = \bigcup_{i=1}^n R_i$. Además, la unión de cualquier k de los círculos que no cruzan el resto de $(n - k)$, contiene precisamente k (multiplicidades contadas) de los autovalores.

Ejemplo 1

Determine los círculos Geršgorin para la matriz

$$A = \begin{bmatrix} 4 & 1 & 1 \\ 0 & 2 & 1 \\ -2 & 0 & 9 \end{bmatrix}$$

y úselos para encontrar los límites del radio espectral de A .

Solución: Los círculos en el teorema de Geršgorin son:

$$R_1 = \{z \in \mathbb{C} : |z - 4| \leq 2\}, R_2 = \{z \in \mathbb{C} : |z - 2| \leq 1\}, R_3 = \{z \in \mathbb{C} : |z - 9| \leq 2\}.$$

Puesto que R_1 y R_2 están separados de R_3 , existen precisamente dos autovalores dentro de $R_1 \cup R_2$ y uno dentro de R_3 . Además, $p(A) = \max_{1 \leq i \leq 3} |\lambda_i|$, por lo que $7 \leq p(A) \leq 11$.

La determinación de las regiones en las que se encuentran es el primer paso para hallar las aproximaciones porque nos da aproximaciones iniciales.

1.2 Definición 1

Sea $\{\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \mathbf{v}^{(3)}, \dots, \mathbf{v}^{(k)}\}$ un conjunto de vectores. El conjunto es linealmente independiente si, siempre que

$$\mathbf{0} = \alpha_1 \mathbf{v}^{(1)} + \alpha_2 \mathbf{v}^{(2)} + \alpha_3 \mathbf{v}^{(3)} + \dots + \alpha_k \mathbf{v}^{(k)} \quad (3)$$

entonces, $\alpha_i = 0$, para cada $i = 1, 2, \dots, k$. De lo contrario, el conjunto de vectores es linealmente dependiente.

Observe que cualquier conjunto de vectores que contiene a el vector nulo es linealmente dependiente.

1.3 Teorema 2

Suponga que $\{\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \mathbf{v}^{(3)}, \dots, \mathbf{v}^{(k)}\}$ es un conjunto de n vectores linealmente independiente en \mathbb{R}^n . Entonces, para cualquier vector $\mathbf{x} \in \mathbb{R}^n$, existe un único conjunto de constantes $\beta_1, \beta_2, \dots, \beta_n$ con

$$\mathbf{x} = \beta_1 \mathbf{v}^{(1)} + \beta_2 \mathbf{v}^{(2)} + \beta_3 \mathbf{v}^{(3)} + \dots + \beta_k \mathbf{v}^{(k)} \quad (4)$$

1.4 Definición 2

Cualquier conjunto de n vectores linealmente independiente en \mathbb{R}^n recibe el nombre de base para \mathbb{R}^n .

Ejemplo 2

ESPACIO

1.5 Teorema 3

Si A es una matriz y $\lambda_1, \dots, \lambda_k$ son autovalores distintos de A con autovectores asociados $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}, \dots, \mathbf{x}^{(k)}$, entonces $\{\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}, \dots, \mathbf{x}^{(k)}\}$ es un conjunto linealmente independiente.

Ejemplo 3

ESPACIO

Ejemplo 4

ESPACIO

1.6 Definición 3

Un conjunto de vectores $\{\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \mathbf{v}^{(3)}, \dots, \mathbf{v}^{(k)}\}$, recibe el nombre de **ortogonal** si $(\mathbf{v}^{(i)})^t \mathbf{v}^{(j)} = 0$, para toda $i \neq j$. Si, además $(\mathbf{v}^{(i)})^t \mathbf{v}^{(i)} = 1$, para toda $i = 1, 2, \dots, n$. Entonces el conjunto recibe el nombre de **ortonormal**.

Puesto que $\mathbf{x}^t \mathbf{x} = \|\mathbf{x}\|_2^2$ para cualquier \mathbf{x} en \mathbb{R}^n , un conjunto de vectores ortogonales $\{\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \mathbf{v}^{(3)}, \dots, \mathbf{v}^{(k)}\}$ es ortonormal si y solo si

$$\|\mathbf{v}^{(i)}\|_2 = 1, \quad \text{para cada } i = 1, 2, \dots, n$$

Ejemplo 5

ESPACIO

1.7 Teorema 4

Un conjunto ortogonal de vectores diferentes a cero es linealmente independiente.

Existe un proceso, también conocido como Gram-Schmidt, que nos permite construir una base ortogonal para \mathbb{R}^n dado un conjunto de n vectores linealmente independiente en \mathbb{R}^n .

1.8 Teorema 5

Sea $\{\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}, \dots, \mathbf{x}^{(k)}\}$ un conjunto de k vectores linealmente independientes en \mathbb{R}^n . Entonces $\{\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \mathbf{v}^{(3)}, \dots, \mathbf{v}^{(k)}\}$ definido mediante

$$\begin{aligned}\mathbf{v}_1 &= \mathbf{x}_1, \\ \mathbf{v}_2 &= \mathbf{x}_2 - \left(\frac{\mathbf{v}_1^t \cdot \mathbf{x}_2}{\mathbf{v}_1^t \cdot \mathbf{v}_1} \right) \mathbf{v}_1, \\ \mathbf{v}_3 &= \mathbf{x}_3 - \left(\frac{\mathbf{v}_1^t \cdot \mathbf{x}_3}{\mathbf{v}_1^t \cdot \mathbf{v}_1} \right) \mathbf{v}_1 - \left(\frac{\mathbf{v}_2^t \cdot \mathbf{x}_3}{\mathbf{v}_2^t \cdot \mathbf{v}_2} \right) \mathbf{v}_2, \\ \mathbf{v}_k &= \mathbf{x}_k - \sum_{i=1}^{k-1} \left(\frac{\mathbf{v}_i^t \cdot \mathbf{x}_k}{\mathbf{v}_i^t \cdot \mathbf{v}_i} \right) \mathbf{v}_i\end{aligned}$$

es un conjunto de k vectores ortogonales en \mathbb{R}^n .

Observe que cuando el conjunto original de vectores construidos forman una base ortogonal para \mathbb{R}^n . A partir de esto podemos formar una base ortonormal $[\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n]$ simplemente al definir para cada $i = 1, 2, \dots, n$.

$$\mathbf{u}_i = \frac{\mathbf{v}_i}{\|\mathbf{v}_i\|_2} \quad (5)$$

Ejemplo 6

ESPACIO

2 El Método de Potencia

El método de potencia es una técnica iterativa que se usa para determinar el autovalor dominante de una matriz (es decir, el autovalor con la mayor magnitud). Al modificar ligeramente el método, también se puede usar para determinar otros autovalores. Una característica útil del método de potencia es que no sólo produce un autovalor, sino también un autovector asociado. De hecho, a menudo, el método de potencia se aplica para encontrar un autovector para un autovalor que es determinado por algunos otros medios.

Para aplicar el método de potencia, suponemos que la matriz $A_{n \times n}$ tiene n autovalores $\lambda_1, \lambda_2, \dots, \lambda_n$, con un conjunto asociado de autovectores linealmente independientes $\{\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \mathbf{v}^{(3)}, \dots, \mathbf{v}^{(n)}\}$. Además, suponemos que A tiene exactamente un autovalor, que es más grande en magnitud, por lo que

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n| \geq 0$$

Una matriz $n \times n$ no necesita tener n autovectores linealmente independientes. Cuando esto no es así, el método de potencia puede seguir siendo exitoso, pero no se garantiza que lo sea.

Si \mathbf{x} es cualquier vector en \mathbb{R}^n , el hecho de que $\{\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \mathbf{v}^{(3)}, \dots, \mathbf{v}^{(n)}\}$ es linealmente independiente implica que existen constantes $\beta_1, \beta_2, \dots, \beta_n$ con

$$\mathbf{x} = \sum_{j=1}^n \beta_j \mathbf{v}^{(j)}.$$

Al multiplicar ambos lados de esta ecuación por $A, A^2, \dots, A^k, \dots$ obtenemos

$$A\mathbf{x} = \sum_{j=1}^n \beta_j A\mathbf{v}^{(j)} = \sum_{j=1}^n \beta_j \lambda_j \mathbf{v}^{(j)}, A^2\mathbf{x} = \sum_{j=1}^n \beta_j \lambda_j A\mathbf{v}^{(j)} = \sum_{j=1}^n \beta_j \lambda_j^2 \mathbf{v}^{(j)}$$

y, en general, $A^k\mathbf{x} = \sum_{j=1}^n \beta_j \lambda_j^k \mathbf{v}^{(j)}.$

Si λ_1^k se factoriza a partir de cada término en el lado derecho de la última ecuación, entonces

$$A^k\mathbf{x} = \lambda_1^k \sum_{j=1}^n \beta_j \left(\frac{\lambda_j}{\lambda_1}\right)^k \mathbf{v}^{(j)}$$

Puesto que $|\lambda_1| > |\lambda_j|$, para todas $j = 2, 3, \dots, n$, tenemos $\lim_{k \rightarrow \infty} \left(\frac{\lambda_j}{\lambda_1}\right)^k = 0$, y

$$\lim_{k \rightarrow \infty} A^k\mathbf{x} = \lim_{k \rightarrow \infty} \lambda_1^k \beta_1 \mathbf{v}^{(1)} \quad (6)$$

La sucesión en la ecuación 6 converge a 0 si $|\lambda_1| < 1$ y diverge si $|\lambda_1| > 1$, siempre y cuando, $\beta_1 \neq 0$. Por consiguiente, las entradas en $A^k\mathbf{x}$ aumentarán con k si $|\lambda_1| > 1$, y tienden a 0 si $|\lambda_1| < 1$, tal vez, al resultar en desborde y subdesborde. Para cuidar esta posibilidad, escalamos las potencias de $A^k\mathbf{x}$ en una forma apropiada para garantizar que la cota en la ecuación 6 es finita y diferente de cero. El escalamiento comienza al seleccionar \mathbf{x} como vector unitario $\mathbf{x}^{(0)}$ relativo a $\|\cdot\|_\infty$ y seleccionar un componente $x_{p_0}^{(0)}$ de $\mathbf{x}^{(0)}$ con

$$x_{p_0}^{(0)} = 1 = \|\mathbf{x}^{(0)}\|_\infty$$

Sea $\mathbf{y}^{(1)} = A\mathbf{x}^{(0)}$ y defina $\mu_1 = y_{p_0}^{(1)}$. Entonces

$$\mu_1 = y_{p_0}^{(1)} = \frac{y_{p_0}^{(1)}}{x_{p_0}^{(0)}} = \frac{\beta_1 \lambda_1 v_{p_0}^{(1)} + \sum_{j=2}^n \beta_j \lambda_j v_{p_0}^{(j)}}{\beta_1 v_{p_0}^{(1)} + \sum_{j=2}^n \beta_j v_{p_0}^{(j)}} = \lambda_1 \left[\frac{\beta_1 v_{p_0}^{(1)} + \sum_{j=2}^n \beta_j \left(\frac{\lambda_j}{\lambda_1}\right) v_{p_0}^{(j)}}{\beta_1 v_{p_0}^{(1)} + \sum_{j=2}^n \beta_j v_{p_0}^{(j)}} \right]$$

Sea p_1 el entero mínimo tal que

$$|y_{p_1}^{(1)}| = \|\mathbf{y}^{(1)}\|_\infty$$

y defina $\mathbf{x}^{(1)}$ mediante

$$\mathbf{x}^{(1)} = \frac{1}{y_{p_1}^{(1)}} \mathbf{y}^{(1)} = \frac{1}{y_{p_1}^{(1)}} A\mathbf{x}^{(0)}$$

Entonces

$$\mathbf{x}_{p_1}^{(1)} = 1 = \left\| \mathbf{x}^{(1)} \right\|_{\infty}$$

Ahora defina

$$\mathbf{y}^{(2)} = A\mathbf{x}^{(1)} = \frac{1}{y_{p_1}^{(1)}} A^2 \mathbf{x}^{(0)}$$

y

$$\mu_2 = y_{p_1}^{(2)} = \frac{y_{p_1}^{(2)}}{x_{p_1}^{(1)}} = \frac{\left[\beta_1 \lambda_1^2 v_{p_1}^{(1)} + \sum_{j=2}^n \beta_j \lambda_j^2 v_{p_1}^{(j)} \right] / y_{p_1}^{(1)}}{\left[\beta_1 \lambda_1 v_{p_1}^{(1)} + \sum_{j=2}^n \beta_j \lambda_j v_{p_1}^{(j)} \right] / y_{p_1}^{(1)}} = \lambda_1 \left[\frac{\beta_1 v_{p_1}^{(1)} + \sum_{j=2}^n \beta_j \left(\frac{\lambda_j}{\lambda_1} \right)^2 v_{p_1}^{(j)}}{\beta_1 v_{p_1}^{(1)} + \sum_{j=2}^n \beta_j \left(\frac{\lambda_j}{\lambda_1} \right) v_{p_1}^{(j)}} \right]$$

Sea p_2 el entero más pequeño con

$$\left| y_{p_2}^{(2)} \right| = \left\| \mathbf{y}^{(2)} \right\|_{\infty}$$

y defina

$$\mathbf{x}^{(2)} = \frac{1}{y_{p_2}^{(2)}} \mathbf{y}^{(2)} = \frac{1}{y_{p_2}^{(2)}} A\mathbf{x}^{(1)} = \frac{1}{y_{p_2}^{(2)} \cdot y_{p_1}^{(1)}} A^2 \mathbf{x}^{(0)}$$

De manera similar, defina las sucesiones de vectores $\{\mathbf{x}^{(m)}\}_{m=0}^{\infty}$ y $\{\mathbf{y}^{(m)}\}_{m=1}^{\infty}$ y una sucesión de escalares $\{\mu^{(m)}\}_{m=1}^{\infty}$ de manera inductiva mediante

$$\mathbf{y}^{(m)} = A\mathbf{x}^{m-1}$$

$$\mu_m = y_{p(m-1)}^{(m)} = \lambda_1 \left[\frac{\beta_1 v_{p(m-1)}^{(1)} + \sum_{j=2}^n \beta_j \left(\frac{\lambda_j}{\lambda_1} \right)^m v_{p(m-1)}^{(j)}}{\beta_1 v_{p(m-1)}^{(1)} + \sum_{j=2}^n \beta_j \left(\frac{\lambda_j}{\lambda_1} \right)^{(m-1)} v_{p(m-1)}^{(j)}} \right] \quad (7)$$

y

$$\mathbf{x}^m = \frac{\mathbf{y}^m}{y_{p_m}^m} = \frac{A^m \mathbf{x}^{(0)}}{\prod_{k=1}^m y_{p_k}^{(k)}}$$

donde en cada paso, p_m se usa para representar el entero más pequeño para el que

$$\left| y_{p_m}^{(m)} \right| = \left\| \mathbf{y}^{(m)} \right\|_{\infty}$$

Al examinar la ecuación 7, observamos que dado $\left| \frac{\lambda_j}{\lambda_1} \right| < 1$, para cada $j = 2, 3, \dots, n$, $\lim_{m \rightarrow \infty} \mu^{(m)} = \lambda_1$, siempre y cuando $\mathbf{x}^{(0)}$ se seleccione de tal forma que $\beta_1 \neq 0$. Además, la sucesión de vectores $\{\mathbf{x}^{(m)}\}_{m=0}^{\infty}$ converge para un autovector asociado con λ_1 que tiene norma l_{∞} igual a uno.

El método de potencia tiene la desventaja de que es desconocido al principio si la matriz tiene un solo autovalor dominante. Tampoco se conoce cómo $\mathbf{x}^{(0)}$ debería seleccionarse para garantizar que su representación en términos de autovectores de la matriz contendrá una contribución diferente de cero de autovectores asociados con el autovalor dominante, si existiera.

Ejemplo 1

ESPACIO

2.1 Convergencia Acelerada

Al seleccionar el entero más pequeño p_m para el que $\|\mathbf{y}_{p_m}^{(m)}\| = \|\mathbf{y}^{(m)}\|$ garantizará, en general, que al final este índice se vuelve invariante. La velocidad a la que $\{\mu^{(m)}\}_{m=1}^{\infty}$ converge en λ_1 se determina mediante los radios $|\lambda_j/\lambda_1|^m$, para $j = 2, 3, \dots, n$, y en particular por medio $|\lambda_2/\lambda_1|^m$. La velocidad de convergencia es $O(|\lambda_2/\lambda_1|^m)$, por lo que existe una constante k , de tal forma que para m grande,

$$|\mu^{(m)} - \lambda_1| \approx k \left| \frac{\lambda_2}{\lambda_1} \right|^m,$$

lo que implica que

$$\lim_{m \rightarrow \infty} \frac{|\mu^{(m+1)} - \lambda_1|}{|\mu^{(m)} - \lambda_1|} \approx k \left| \frac{\lambda_2}{\lambda_1} \right| < 1.$$

La sucesión $\{\mu^{(m)}\}$ converge linealmente a λ_1 , por lo que el procedimiento Δ^2 de Aitkens se puede utilizar para acelerar la convergencia.

No es necesario que la matriz tenga diferentes autovalores para que el método de potencia converja. Si la matriz tiene un autovalor dominante único, λ_1 , con multiplicidad r superior a 1 y $\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \dots, \mathbf{v}^{(r)}$ son autovectores linealmente independientes asociados con λ_1 , el procedimiento seguirá convergiendo en λ_1 . En este caso, la sucesión de vectores $\{\mathbf{x}^{(m)}\}_{m=0}^{\infty}$, convergerá en un autovector λ_1 en la norma l_{∞} igual a uno que depende de la selección del vector inicial $\mathbf{x}^{(0)}$ y es una combinación lineal de $\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \dots, \mathbf{v}^{(r)}$.

Ejemplo 2

ESPACIO

2.2 Matrices Simétricas

Cuando A es simétrica, es posible hacer una variación en la selección de los vectores $\mathbf{x}^{(m)}$ y $\mathbf{y}^{(m)}$ y los escalares $\mu^{(m)}$ para mejorar significativamente el índice de convergencia de la sucesión $\{\mu^{(m)}\}_{m=1}^{\infty}$ para el autovalor dominante λ_1 . De hecho, a pesar de que el índice de convergencia del método de potencia general es $O(|\lambda_2/\lambda_1|^m)$, para las matrices simétricas es $O(|\lambda_2/\lambda_1|^{2m})$. Puesto que la sucesión $\{\mu^{(m)}\}$ sigue siendo convergente, también puede aplicarse el procedimiento Δ^2 de Aitkens.

Ejemplo 3

ESPACIO

Si λ es un número real que aproxima un autovalor de una matriz simétrica A y \mathbf{x} es un autovector asociado aproximado, entonces $A\mathbf{x} - \lambda\mathbf{x}$ es aproximadamente el vector cero. El siguiente teorema relaciona la norma de este vector para la precisión del autovalor λ

2.3 Teorema

Suponga que A es una matriz simétrica $n \times n$ con autovalores $\lambda_1, \lambda_2, \dots, \lambda_n$. Si $\|A\mathbf{x} - \lambda\mathbf{x}\|_2 < \epsilon$ para algunos números reales λ y vector \mathbf{x} con $\|\mathbf{x}\|_2 = 1$. Entonces

$$\min_{1 \leq j \leq n} |\lambda_j - \lambda| < \epsilon.$$

2.4 Método de Potencia Inversa

El **método de potencia inversa** es una modificación del método de potencia que da una convergencia más rápida. Se usa para determinar el autovalor de A que está más cerca de un número específico q .

Supongamos que la matriz A tiene autovalores $\lambda_1, \lambda_2, \dots, \lambda_n$ con autovectores linealmente independientes $\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \dots, \mathbf{v}^{(n)}$. Los autovalores de $(A - qI)^{-1}$, donde $q \neq \lambda_i$, para $i = 1, 2, \dots, n$, son

$$\frac{1}{\lambda_1 - q}, \frac{1}{\lambda_2 - q}, \dots, \frac{1}{\lambda_n - q},$$

con estos mismos autovectores $\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \dots, \mathbf{v}^{(n)}$.

Al aplicar el método de potencia a $(A - qI)^{-1}$ da

$$\mathbf{y}^{(m)} = (A - qI)^{-1} \mathbf{x}^{(m-1)},$$

$$\mu^{(m)} = \frac{\mathbf{y}_{p_{(m-1)}}^{(m)}}{\mathbf{x}_{p_{(m-1)}}^{(m-1)}} = \frac{\sum_{j=1}^n \beta_j \frac{1}{(\lambda_j - q)^m} v_{p_m}^{(j)}}{\sum_{j=1}^n \beta_j \frac{1}{(\lambda_j - q)^{m-1}} v_{p_m}^{(j)}}, \quad (8)$$

y

$$\mathbf{x}^{(m)} = \frac{\mathbf{y}^{(m)}}{y_{p_m}^{(m)}},$$

donde, en cada paso, p_m representa el entero más pequeño para el que $|\mathbf{y}_{p_m}^{(m)}| = \|\mathbf{y}^{(m)}\|_\infty$. La sucesión $\{\mu^{(m)}\}$ en la ecuación (8) converge en $1/(\lambda_k - q)$, donde

$$\frac{1}{|\lambda_k - q|} = \max_{1 \leq i \leq n} \frac{1}{|\lambda_i - q|}$$

y $\lambda_k \approx q + 1/\mu^{(m)}$ es el autovalor de A más cercano a q .

Conociendo k , la ecuación (8) se puede escribir como

$$\mu^{(m)} = \frac{1}{\lambda_k - q} \left[\frac{\beta_k v_{p(m-1)}^{(k)} + \sum_{j=1, j \neq k}^n \beta_j \left[\frac{\lambda_k - q}{\lambda_j - q} \right]^m v_{p(m-1)}^{(j)}}{\beta_k v_{p(m-1)}^{(k)} + \sum_{j=1, j \neq k}^n \beta_j \left[\frac{\lambda_k - q}{\lambda_j - q} \right]^{m-1} v_{p(m-1)}^{(j)}} \right]. \quad (9)$$

Por lo tanto, la selección de q determina la convergencia, siempre y cuando $1/(\lambda_k - q)$ sea un único autovalor dominante de $(A - qI)^{-1}$ (a pesar de que puede ser un autovalor múltiple). Mientras q está más cerca de un autovalor λ_k , más rápida será la convergencia ya que la convergencia es de orden

$$O \left(\left| \frac{(\lambda - q)^{-1}}{(\lambda_k - q)^{-1}} \right|^m \right) = O \left(\left| \frac{(\lambda_k - q)}{(\lambda - q)} \right|^m \right),$$

donde λ representa el autovalor de A que es el segundo más cercano a q .

El vector $y^{(m)}$ se obtiene al resolver el sistema lineal

$$(A - qI) \mathbf{y}^{(m)} = \mathbf{x}^{(m-1)}.$$

En general, se usa la eliminación gaussiana con pivoteo pero, como en el caso de la factorización LU, los multiplicadores se pueden guardar para reducir el cálculo. La selección de q puede basarse en el teorema del círculo de Geršgorin o en otros medios de localización de un autovalor.

Ejemplo 4

ESPACIO