Parte 3. Vectores y valores propios

Gustavo Montero

Escuela Universitaria Politécnica
Universidad de Las Palmas de Gran Canaria

Curso 2004-2005

Valores y vectores propios

Valor propio

Se dice que el número λ, real o complejo, es un valor propio A si existe un vector no nulo u, real o complejo tal que $Au = \lambda u$, es decir $(A - \lambda I)u = 0$

Valores y vectores propios

Valor propio

Se dice que el número λ, real o complejo, es un valor propio A si existe un vector no nulo u, real o complejo tal que $Au = \lambda u$, es decir $(A - \lambda I)u = 0$

Vector propio

El vector u se denomina vector propio de A asociado al valor propio λ.

Valores y vectores propios

Valor propio

Se dice que el número λ, real o complejo, es un valor propio A si existe un vector no nulo u, real o complejo tal que $Au = \lambda u$, es decir $(A - \lambda I)u = 0$

Vector propio

El vector u se denomina vector propio de A asociado al valor propio λ.

Polinomio característico

En general, el polinomio que resulta de desarrollar $|A - \lambda I|$, cuyos ceros son precisamente los valores propios de A, se denomina polinomio característico.

$$P(\lambda) = (-1)^n \lambda^n + a_1 \lambda^{n-1} + \cdots + a_n$$

- $\sum_{i=1}^{n} \lambda_i = -\dfrac{a_1}{(-1)^n} = \sum_{i=1}^{n} a_{ii} = trA$
- $\prod_{i=1}^{n} \lambda_i = a_n = |A|$

Valores y vectores propios

Valor propio

Se dice que el número λ, real o complejo, es un valor propio A si existe un vector no nulo u, real o complejo tal que $Au = \lambda u$, es decir $(A - \lambda I)u = 0$

Vector propio

El vector u se denomina vector propio de A asociado al valor propio λ.

Polinomio característico

En general, el polinomio que resulta de desarrollar $|A - \lambda I|$, cuyos ceros son precisamente los valores propios de A, se denomina polinomio característico.

$$P(\lambda) = (-1)^n \lambda^n + a_1 \lambda^{n-1} + \cdots + a_n$$

- $\sum_{i=1}^{n} \lambda_i = -\dfrac{a_1}{(-1)^n} = \sum_{i=1}^{n} a_{ii} = trA$
- $\prod_{i=1}^{n} \lambda_i = a_n = |A|$

Radio espectral

Se denomina radio espectral ρ de una matriz A al valor

$$\rho(A) = \max |\lambda_i|$$

Propiedades

Valores propios de una matriz cualquiera

- Si λ es complejo, entonces u es complejo.
- Los valores propios de $B = C^{-1}AC$ son los mismos de A. Si x es el vector propio asociado a λ, entonces Cx es un vector propio de B asociado a λ.

Propiedades

Valores propios de una matriz cualquiera

- Si λ es complejo, entonces u es complejo.
- Los valores propios de $B = C^{-1}AC$ son los mismos de A. Si x es el vector propio asociado a λ, entonces Cx es un vector propio de B asociado a λ.

Valores propios de matrices simétricas

- Si D es la matriz diagonal cuyos elementos diagonales son los valores propios de A, entonces existe una matriz ortogonal Q tal que $D = Q^{-1}AQ = Q^tAQ$.
- Asimismo, existen n vectores propios de A que forman un conjunto ortonormal, y coinciden con las columnas de la matriz ortogonal Q.
- Todos los valores propios de A son reales.
- A es definida positiva si y sólo si todos los valores propios de A son positivos.

Teorema de Cayley-Hamilton

Sea A una matriz cuyo polinomio característico es,

$$P(\lambda) = (-1)^n \lambda^n + a_1 \lambda^{n-1} + \cdot + a_n$$

entonces la matriz A verifica,

$$(-1)^n A^n + a_1 A^{n-1} + \cdot + a_n I = 0$$

Teorema de Schur

Sea A una matriz $n \times n$ cualquiera. Entonces existe una matriz U ortogonal tal que

$$T = U^{-1}AU$$

siendo T una matriz triangular superior cuyos elementos diagonales son los valores propios de A.

Teorema de Gerschgorin

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

$$R_i = \{z \in C \quad \text{tal que} |z - a_{ii}| \leq \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|\}$$

donde C denota el conjunto de los números complejos. Entonces los valores propios de A están contenidos en $R = \bigcup_{i=1}^{n} R_i$. Es más, si la unión de k de estos círculos no se corta con los restantes $n-k$, entonces dicha unión contiene k valores propios (incluyendo los valores propios múltiples)

Método de Krylov

Se basa en el teorema de Cayley-Hamilton ($P(A) = 0$):

$$(-1)^n A^n + a_1 A^{n-1} + \cdot + a_n I = 0$$

Multiplicando por x resulta,

$$(-1)^n A^n x + a_1 A^{n-1} x + \cdot + a_n x = 0$$

Si denotamos $A^{n-i} x = v_i$, para $i = 1, 2, \ldots n$, obtenemos el sistema lineal

$$\sum_{i=1}^{n} a_i v_i = (-1)^{n+1} A^n x$$

cuya resolución nos proporciona los coeficientes a_i de $P(\lambda)$

Método de Leverrier modificado

Fórmula de Newton

Se basa en la fórmula de Newton. Sea el polinomio de grado n

$$P_n(x) = \prod_{i=1}^{n}(x - x_i) = \sum_{k=0}^{n} a_k x^{n-k},\ a_0 = 1$$

y la suma

$$S_k = \sum_{i=1}^{n} x_i^k, \qquad k = 1, 2, \ldots, n$$

La fórmula de Newton nos da la siguiente relación entre S_k y a_k,

$$ka_k + S_k + \sum_{i=1}^{k-1} S_i a_{k-i} = 0$$

Método de Leverrier modificado

Fórmulas de Leverrier o de Faddeev

$$
\begin{aligned}
a_k &= \frac{1}{k} tr(AB_{k-1}), \qquad & k = 1, 2, \ldots, n \\
B_k &= -AB_{k-1} + a_k I, \qquad & k = 1, 2, \ldots, n
\end{aligned}
$$

siendo $B_0 = I$ y $B_n = 0$.

Método de Leverrier modificado

Fórmulas de Leverrier o de Faddeev

$$
\begin{aligned}
a_k &= \frac{1}{k} tr(AB_{k-1}), & k = 1, 2, \ldots, n \\
B_k &= -AB_{k-1} + a_k I, & k = 1, 2, \ldots, n
\end{aligned}
$$

siendo $B_0 = I$ y $B_n = 0$.

Algoritmo de Leverrier Modificado

$$
\begin{array}{lll}
A_1 = A & a_1 = tr(A_1) & B_1 = A_1 - a_1 I \\
A_2 = B_1 A & a_2 = \frac{1}{2} tr(A_2) & B_2 = A_2 - a_2 I \\
\vdots & \vdots & \vdots \\
A_n = B_{n-1} A & a_n = \frac{1}{n} tr(A_n) & B_n = A_n - a_n I
\end{array}
$$

donde los a_i son los coeficientes del polinomio característico.

Método de Leverrier modificado

Fórmulas de Leverrier o de Faddeev

$$\begin{aligned} a_k &= \frac{1}{k} tr(AB_{k-1}), && k = 1, 2, \ldots, n \\ B_k &= -AB_{k-1} + a_k I, && k = 1, 2, \ldots, n \end{aligned}$$

siendo $B_0 = I$ y $B_n = 0$.

Algoritmo de Leverrier Modificado

$$\begin{array}{lll} A_1 = A & a_1 = tr(A_1) & B_1 = A_1 - a_1 I \\ A_2 = B_1 A & a_2 = \frac{1}{2} tr(A_2) & B_2 = A_2 - a_2 I \\ \vdots & \vdots & \vdots \\ A_n = B_{n-1} A & a_n = \frac{1}{n} tr(A_n) & B_n = A_n - a_n I \end{array}$$

donde los a_i son los coeficientes del polinomio característico.

Aplicación al cálculo de la inversa

Como $B_n = 0$, se obtiene que $A^{-1} = \dfrac{1}{a_n} B_{n-1}$, obteniendo así un método para calcular la inversa de una matriz.

Métodos de las potencias

Sucesión de los vectores del subspacio de Krylov

Supongamos que los valores propios de una matriz A cumplen,

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$$

Consideremos cualquier vector v, que puede ser escrito

$$v = \sum_{i=1}^{n} \alpha_i u_i = v_0$$

Entonces multiplicando por la matriz A se tiene,

$$\begin{array}{lcl} v_1 & = & Av_0 = Av = \sum_{i=1}^{n} \lambda_i \alpha_i u_i \\ v_2 & = & Av_1 = A^2 v = \sum_{i=1}^{n} \lambda_i^2 \alpha_i u_i \\ \cdots & \cdots & \cdots\cdots\cdots \\ v_k & = & Av_{k-1} = A^k v = \sum_{i=1}^{n} \lambda_i^k \alpha_i u_i \end{array}$$

$$v_k = \lambda_1^k \left(\alpha_1 u_1 + \sum_{i=2}^{n} \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^k u_i \right)$$

Métodos de las potencias

Paso al límite

Luego el límite del cociente de las componentes r-ésimas de dos vectores consecutivos resulta,

$$\lim_{k\to\infty} \frac{\left(v_{k+1}\right)_r}{\left(v_k\right)_r} = \lambda_1 \frac{\alpha_1\left(u_1\right)_r + \sum_{i=2}^n \alpha_i \left(\frac{\lambda_i}{\lambda_1}\right)^{k+1}\left(u_i\right)_r}{\alpha_1\left(u_1\right)_r + \sum_{i=2}^n \alpha_i \left(\frac{\lambda_i}{\lambda_1}\right)^{k}\left(u_i\right)_r} = \lambda_1$$

si $\alpha_1 \neq 0$ y $\left(u_1\right)_r \neq 0$.

Métodos de las potencias

Normalización

Para evitar que los v_i sean muy grandes, se suele construir la sucesión normalizada,

$$
\begin{array}{ccc}
w_1 & = & \dfrac{Av}{||Av||_\infty} \\
w_2 & = & \dfrac{Aw_1}{||Aw_1||_\infty} \\
\cdots & \cdots & \cdots\cdots\cdots \\
w_k & = & \dfrac{Aw_{k-1}}{||Aw_{k-1}||_\infty}
\end{array}
$$

siendo, en este caso

$$\lim_{k\to\infty} \frac{(Aw_{k+1})_r}{(w_k)_r} = \lambda_1$$

Resumen

- El conocimiento del espectro de una matriz, esto es, el conjunto de sus valores propios, tiene mucho interés a la hora de estudiar su condicionamiento para aplicar métodos iterativos de resolución de sistemas de ecuaciones lineales. El teorema de Gerschgorin define la zona donde se encuentra dicho espectro.
- El método de Krylov, basado en el teorema de Cayley-Hamilton, conduce a la obtención de los coeficientes del polinomio característico. Sin embargo, para el caso de matrices de orden elevado suele resultar excesivamente costoso.
- El método de Leverrier modificado construye de forma recursiva los coeficientes del polinomio característico. Además, como resultado se puede calcular la inversa. Su coste para matrices de orden elevado puede ser prohibitivo.
- El método de las potencias permite aproximar el valor absoluto del mayor valor propio de una matriz, utilizando simplemente productos matriz-vector.

Resumen

- El conocimiento del espectro de una matriz, esto es, el conjunto de sus valores propios, tiene mucho interés a la hora de estudiar su condicionamiento para aplicar métodos iterativos de resolución de sistemas de ecuaciones lineales. El teorema de Gerschgorin define la zona donde se encuentra dicho espectro.
- El método de Krylov, basado en el teorema de Cayley-Hamilton, conduce a la obtención de los coeficientes del polinomio característico. Sin embargo, para el caso de matrices de orden elevado suele resultar excesivamente costoso.
- El método de Leverrier modificado construye de forma recursiva los coeficientes del polinomio característico. Además, como resultado se puede calcular la inversa. Su coste para matrices de orden elevado puede ser prohibitivo.
- El método de las potencias permite aproximar el valor absoluto del mayor valor propio de una matriz, utilizando simplemente productos matriz-vector.

Resumen

- El conocimiento del espectro de una matriz, esto es, el conjunto de sus valores propios, tiene mucho interés a la hora de estudiar su condicionamiento para aplicar métodos iterativos de resolución de sistemas de ecuaciones lineales. El teorema de Gerschgorin define la zona donde se encuentra dicho espectro.
- El método de Krylov, basado en el teorema de Cayley-Hamilton, conduce a la obtención de los coeficientes del polinomio característico. Sin embargo, para el caso de matrices de orden elevado suele resultar excesivamente costoso.
- El método de Leverrier modificado construye de forma recursiva los coeficientes del polinomio característico. Además, como resultado se puede calcular la inversa. Su coste para matrices de orden elevado puede ser prohibitivo.
- El método de las potencias permite aproximar el valor absoluto del mayor valor propio de una matriz, utilizando simplemente productos matriz-vector.

Resumen

- El conocimiento del espectro de una matriz, esto es, el conjunto de sus valores propios, tiene mucho interés a la hora de estudiar su condicionamiento para aplicar métodos iterativos de resolución de sistemas de ecuaciones lineales. El teorema de Gerschgorin define la zona donde se encuentra dicho espectro.
- El método de Krylov, basado en el teorema de Cayley-Hamilton, conduce a la obtención de los coeficientes del polinomio característico. Sin embargo, para el caso de matrices de orden elevado suele resultar excesivamente costoso.
- El método de Leverrier modificado construye de forma recursiva los coeficientes del polinomio característico. Además, como resultado se puede calcular la inversa. Su coste para matrices de orden elevado puede ser prohibitivo.
- El método de las potencias permite aproximar el valor absoluto del mayor valor propio de una matriz, utilizando simplemente productos matriz-vector.

Resumen

- El conocimiento del espectro de una matriz, esto es, el conjunto de sus valores propios, tiene mucho interés a la hora de estudiar su condicionamiento para aplicar métodos iterativos de resolución de sistemas de ecuaciones lineales. El teorema de Gerschgorin define la zona donde se encuentra dicho espectro.
- El método de Krylov, basado en el teorema de Cayley-Hamilton, conduce a la obtención de los coeficientes del polinomio característico. Sin embargo, para el caso de matrices de orden elevado suele resultar excesivamente costoso.
- El método de Leverrier modificado construye de forma recursiva los coeficientes del polinomio característico. Además, como resultado se puede calcular la inversa. Su coste para matrices de orden elevado puede ser prohibitivo.
- El método de las potencias permite aproximar el valor absoluto del mayor valor propio de una matriz, utilizando simplemente productos matriz-vector.