Capítulo 5 - Autovalores y Autovectores (3ra. parte) [1]

Facultad de Cs. Exactas, Ingeniería y Agrimensura
Universidad Nacional de Rosario

[1] Siguiendo *Linear Algebra and its applications*, G. Strang.

Outline

La clase anterior vimos elconcepto de *transformación de semejanza* y la relación de equivalencia de *semejanza entre matrices*.

Definición: Dada una matriz inversible M, la transformación que a toda matriz A la lleva a $M^{-1}AM$ es una *transformación de similaridad o semejanza*. Decimos que A es *semejante* a B si existe M inversible tal que $B = M^{-1}AM$.

La propiedad más importante que relaciona a matrices semejantes es que tienen el mismo polinomio caraterístico y por ende, los mismos autovalores. Los autovectores se relacionan a través de la matriz que define la transformación de semejanza entre ellas.

Si $B = M^{-1}AM$ entonces A y B tienen los mismos autovalores. Si x es un autovector de A asociado a λ, $M^{-1}x$ es un autovector de B asociado a λ.

A lo fines de muchas aplicaciones prácticas, dada una matriz A nos interesa saber cuál de todas sus matrices semejantes tiene la *estructura más sencilla*.

Dada una matriz A, la matriz de *estructura más sencilla* que comparte sus autovalores con A es la matriz diagonal que definimos como Λ. Pero no todas las matrices son semejantes a su matriz Λ. Ésas son las matrices *diagonalizables*.

Vimos también que el Lema de Schur nos asegura que toda matriz es semejante a una matriz triangular. Más aún, la matriz de la tranformación de semejanza U que lleva cualquier matriz A a una triangular T es una matriz unitaria. (matriz compleja con sus columnas ortonormales según el producto interno estándar en $\mathbb{C}^n$).

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

Observar que T tienelos autovalores de A en su diagonal. (Justificar).

Veamos ahora la prueba del Lema de Schur.

Prueba del Lema de Schur:
Debemos probar que, para toda matriz A de tamaño $n \times n$, existe una matriz unitaria U_n tal que $U_n^{-1}AU = T_n$, con T_n matriz triangular. La prueba es por inducción en n.

Si $n = 1$, la propiedad es trivial y la matriz diagonalizante U_1 es la identidad 1×1. Supongamos que para toda matriz A de tamaño $k \times k$ existe una matriz unitaria U_k tal que $U_k^{-1}AU_k = T_k$, con T_k triangular.

Sea A una matriz de tamaño $(k+1) \times (k+1)$. Tomamos λ_1, autovalor de A, y x^1 un autovector (normalizado) asociado a λ_1. Sea U^1 una matriz unitaria $(k+1) \times (k+1)$ cuya primer columna es x^1 (siempre existe por G-S) y $M = U_1^{-1}AU_1$. Si M^1 es la primer columna de M tenemos:

$$M^1 = Me^1 = (U_1^H A U_1)e^1 = U_1^H A(U_1 e^1) = U_1^H A x^1 = \lambda_1 (U_1^H x^1).$$

Como las filas de U_1^H son las columnas de U_1 conjugadas, y las columnas de U_1 son ortonormales, la primer componente de $U_1^H x^1$ es la norma de x^1 (que es 1) y las restantes componentes son cero.

Así, existen M_k matriz $k \times k$ y $b = (b_1, \ldots, b_k)$ tales que

$$U_1^{-1} A U_1 = \left[\begin{array}{cccc} \lambda_1 & b_1 & \ldots & b_k \\ 0 & & & \\ \vdots & & M_k & \\ 0 & & & \end{array} \right]$$

Por hipótesis de inducción, existe U_k unitaria tal que $U_k^{-1} M_k U_k = T_k$, con T_k triangular.

Sea U_2 la matriz $(k+1) \times (k+1)$ cuya primer columna y primera fila son el vector $e^1 \in \mathbb{R}^{k+1}$ y U_k es su submatriz $k \times k$ en su esquina inferior derecha. Esto es,

$$U_2 = \left[\begin{array}{cccc} 1 & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & U_k & \\ 0 & & & \end{array} \right]$$

Es fácil verificar que U_2 es unitaria. (Justificar). Además, U_2^{-1} tiene en su primer columna y en su primera fila al vector $e^1 \in \mathbb{R}^{k+1}$ y la submatriz $k \times k$ en su esquina inferior derecha es U_k^{-1}. Esto es,

$$U_2^{-1} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & U_k^{-1} & \\ 0 & & & \end{bmatrix}$$

Finalmente, si $U_{k+1} = U_1 U_2$, resulta U_{k+1} unitaria (Justificar).
Sólo resta probar que $U_{k+1}^{-1} A U_{k+1} = T_{k+1}$, con T_{k+1} matriz triangular.
Tenemos

$$U_{k+1}^{-1} A U_{k+1} = (U_2^{-1} U_1^{-1}) A (U_1 U_2) = U_2^{-1} (U_1^{-1} A U_1) U_2 = U_2^{-1} M U_2.$$

Por lo tanto,
$U_{k+1}^{-1}AU_{k+1} =$

$$= \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & U_k^{-1} & \\ 0 & & & \end{bmatrix} \begin{bmatrix} \lambda_1 & b_1 & \dots & b_k \\ 0 & & & \\ \vdots & & M_k & \\ 0 & & & \end{bmatrix} \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & U_k & \\ 0 & & & \end{bmatrix} =$$

$$= \begin{bmatrix} \lambda_1 & b_1 & \dots & b_k \\ 0 & & & \\ \vdots & & U_k^{-1}M_k & \\ 0 & & & \end{bmatrix} \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & U_k & \\ 0 & & & \end{bmatrix} =$$

$$= \begin{bmatrix} \lambda_1 & bU_k \\ 0 & \\ \vdots & U_k^{-1}M_kU_k \\ 0 & \end{bmatrix} = \begin{bmatrix} \lambda_1 & bU_k \\ 0 & \\ \vdots & T_k \\ 0 & \end{bmatrix}$$

Por lo tanto $U_{k+1}^{-1}AU_{k+1} = T_{k+1}$ es una matriz triangular. $\square$

El Lema de Schur permitió probar que toda matriz hermitiana (resp. simétrica) es diagonalizable por una matriz unitaria (resp. ortogonal). O sea, las matrices hermitianas tienen n autovectores ortonormales.

Esta propiedad nos lleva a la *descomposición espectral* de estas matrices. Esto es, toda matriz hermitiana puede ser escrita como combinación lineal de las matrices de proyección sobre los autoespacios de sus autovectores diferentes.

A hermitiana (simétrica), $\lambda_i, i = 1, \ldots, k$ autovalores diferentes de A,
$A = \sum_{i=1}^{k} \lambda_i P_i$ (P_i matriz proyección sobre autoespacio asociado a λ_i)

Cerramos la clase anterior preguntándonos qué otras matrices son diagonalizables por una matriz unitaria (equivalentemente, admiten una descomposición espectral). Esas matrices son las *matrices normales*.

Definición: N es una *matriz normal* si conmuta con su hermitiana. Esto es, si $NN^H = N^H N$.

Observación: las matrices hermitianas y unitarias son normales. (Justificar).

Ejercicio(en la práctica): Sea T una matriz triangular. Entonces T es normal si y solo si T es diagonal.

Teorema: Sea A matriz $n \times n$. Entonces, existe U unitaria tal que $U^{-1}AU = \Lambda$ si y sólo si A es normal.

Prueba: Sea A una matriz normal. Por el Lema de Schur, existe U, matriz unitaria, tal que $U^{-1}AU = U^H AU = T$, con T triangular. Veremos que T es hermitiana y por ende diagonal. Tenemos,

$$TT^H = (U^H AU)(U^H A^H U) = U^H (AA^H)U = U^H (A^H A)U =$$

$$= U^H A^H (UU^H) AU = (U^H A^H U)(U^H AU) = T^H T.$$

Por lo tanto, T es normal. Como T es triangular, T es diagonal y $T = \Lambda$.

Recíprocamente, sea U matriz unitaria tal que $U^{-1}AU = \Lambda$. Debemos probar que A es normal. (Ejercicio). $\square$

Volvemos a preguntarnos, ¿cuál es la matriz más sencilla que es semejante a una matriz (no normal) dada?

El Lema de Schur, nos permite llegar hasta una matriz triangular a través de una transformación de semejanza definida por una matriz unitaria. Si nos permitimos transformaciones de semejanza con cualquier matriz no singular M, ¿cuál será la forma *más diagonal posible* a la que podemos llevar a A? Ésta es la denominada *Forma de Jordan*.

Si A es una matriz $n \times n$ con s autovectores l.i. entonces A es semejante a una matriz *diagonal por bloques* J, con s *bloques de Jordan*, $J_1, \ldots, J_s$.

Cada bloque de Jordan $J_i, i = 1, \ldots, s$, es una matriz triangular superior que tiene un sólo autovalor λ_i con autoespacio de dimensión 1. Las entradas en la diagonal son iguales a λ_i, las entradas en la segunda diagonal superior son iguales a 1 y el resto de las entradas son nulas.

Así, cada bloque de Jordan es de la forma

$$J_i = \begin{bmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ & & & & \lambda_i \end{bmatrix}$$

Decimos que J es *diagonal por bloques* porque sus bloques *se ubican en la diagonal* y, por fuera de ellos, las entradas son todos nulas. Esto es, toda matriz con s autovectores l.i. es semejante a una matriz J de la forma:

$$J = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s \end{bmatrix}$$

La Forma de Jordan es un importante resultado desde el punto de vista teórico ya que nos proporciona la forma *más diagonal* (matriz más *rala*) a la que podemos aspirar entre las matrices semejantes a una dada.
Además, la Forma de Jordan *caracteriza* a las matrices semejantes. Esto es, dos matrices son semejantes si y solo si tienen la misma forma de Jordan.

Sin embargo, carece de valor práctico ya que su construcción es extremadamente técnica e inestable (pequeños cambios en una entrada de la matriz original pueden llevar a Formas de Jordan totalmente diferentes).

No todas las matrices diagonalizables son normales.

- A es diagonalizable si y solo si A tiene n autovectores l.i..
- A es normal si y solo si A tiene n autovectores ortogonales.

Ejercicio: Probar que $A = \begin{bmatrix} 1 & 1 \\ 3 & 3 \end{bmatrix}$ es diagonalizable y no es normal.

¿Cuándo A posee n autovectores l.i.?
Sabemos que si todos sus autovalores tienen multiplicidad igual a 1 (n autovalores diferentes), los autovectores son l.i..

¿Qué sucede cuándo A tiene autovectores con multiplicidad mayor que 1?.

Si A es digonalizable y λ es un autovalor de A de multiplicidad p, A tiene (al menos) p autovectores l.i. asociados a λ. (Justificar). O sea, el autoespacio asociado a λ debe tener dimensión al menos p.

Entonces, Si A es diagonalizable, la dimensión del autoespacio asociado a cada uno de sus valores diferentes es al menos la multiplicidad del autovalor $\longrightarrow$ *condición necesaria*.

¿Es condición suficiente?

El siguiente teorema, válido para autovectores y autovalores de cualquier transformación lineal de un espacio vectorial en sí mismo, nos asegura que si.

Teorema: Sea V un espacio vectorial sobre $\mathbb{K}$ y T una transformación lineal de V en sí mismo. Sean $\lambda_i \in \mathbb{K}, i = 1, \ldots, p$, autovalores diferentes T. Para cada $i = 1, \ldots, p$, sea $\mathscr{V}_i$ un conjunto de vectores l.i. en el autoespacio correspondiente a λ_i. Entonces, $\mathscr{V} = \cup_{i=1}^{p} \mathscr{V}_i$ es un conjunto de vectores l.i. de V.

Prueba:
Consideremos una combinación lineal nula de vectores de $\mathscr{V}$

$$\sum_{i=1}^{p}\sum_{j=1}^{t(i)} \alpha_{ij} v_j^i = 0 \tag{1}$$

donde para cada $i = 1, \ldots, p$, $v_j^i \in \mathscr{V}_i$ para todo $j = 1, \ldots t(i)$.

Para cada $i = 1, \ldots, p$, sea $u^i = \sum_{j=1}^{t(i)} \alpha_{ij} v_j^i$. Claramente, u^i es un elemento del autoespacio asociado a λ_i.
Así, (1) puede expresarse como $\sum_{i=1}^{p} u^i = 0$.
Como los autovectores correspondientes a autovalores diferentes son l.i. no es posible que su suma sea el vector nulo. Por lo tanto, $u^i = 0$ para todo $i = 1, \ldots, p$.

Prueba: (continuación)
Así, para todo $i = 1, \ldots, p$ tenemos

$$u^i = \sum_{j=1}^{t(i)} \alpha_{ij} v_j^i = 0.$$

Como los vectores $v_j^i, j = 1, \ldots, t(i)$ son l.i., $\alpha_{ij} = 0$ para todo $j = 1, \ldots, t(i)$.
Por lo tanto, los vectores de $\mathscr{V}$ son l.i.. $\square$

Tenemos entonces como corolario:
Corolario: Una matriz A es diagonalizable si y solo si la multiplicidad de todo autovalor es menor o igual a la dimensión de su autoespacio asociado.

Prueba: Ejercicio.

La multiplicidad de un autovalor (multiplicidad como raíz del polinomio característico) se suele llamar *multiplicidad algebraica* para contraponerla con su *multiplicidad geométrica*.

Definición: Dada una matriz A y λ un autovalor de A, llamamos *multiplicidad geométrica de* λ a la dimensión del autoespacio asociado a λ.

Hemos visto que si la multiplidad geométrica de todo autovalor es al menos su multiplicidad algebraica, la matriz es diagonalizable.

La multiplicidad geométrica, ¿podría ser mayor a la algebraica? NO

Teorema La multiplicidad geométrica de un autovalor es siempre a lo sumo su multiplicidad algebraica.

Prueba: no la vemos.

Teorema:
Sea A una matriz $n \times n$ y $\lambda_i, i = 1, \ldots, p$, los autovalores distintos de A, con multiplicidad algebraica $ma(i)$, multiplicidad geométrica $mg(i)$ y $\mathscr{B}_i$ una base de su autoespacio. Entonces, las siguientes condiciones son equivalentes:

1. A es diagonalizable.
2. Para todo $i = 1, \ldots, p$, $ma(i) = mg(i)$.
3. $\sum_{i=1}^{p} mg(i) = n$.
4. $\cup_{i=1}^{p} \mathscr{B}_i$ es una base de $\mathbb{K}^n$.

Prueba:
$(1) \Rightarrow (2)$: Sea S diagonalizante de A. Para cada $i = 1, \ldots, p$ existen $ma(i)$ columnas de S que corresponden a autovectores asociados a λ_i. Como S es inversible, esos autovectores son l.i..Por lo tanto, existen $ma(i)$ vectores l.i. en el autoespacio asociado a λ_i lo que implica $ma(i) \leq mg(i)$. Como sabemos que $ma(i) \geq mg(i)$, resulta $ma(i) = mg(i)$.

$(2) \Rightarrow (3)$: Es trivial.

Prueba: (continuación)

$(3) \Rightarrow (4)$:

Ya hemos probado que $\cup_{i=1}^{p} \mathscr{B}_i$ es un conjunto de vectores l.i. de $\mathbb{K}^n$. Además, como $\mathscr{B}_i \cap \mathscr{B}_j = \emptyset$ si $i \neq j$ (justificar), resulta:

$$\left|\cup_{i=1}^{p} \mathscr{B}_i\right| = \sum_{i=1}^{p} |\mathscr{B}_i| = \sum_{i=1}^{p} mg(i) = n.$$

Por lo tanto, $\cup_{i=1}^{p} \mathscr{B}_i$ es una base de $\mathbb{K}^n$.

$(4) \Rightarrow (1)$: Si $\cup_{i=1}^{p} \mathscr{B}_i$ es una base de $\mathbb{K}^n$ entonces $\sum_{i=1}^{p} |\mathscr{B}_i| = n$ (justificar). Por lo tanto, tenemos:

$$n = \sum_{i=1}^{p} |\mathscr{B}_i| = \sum_{i=1}^{p} mg(i) \leq \sum_{i=1}^{p} ma(i) = n$$

y, para todo $i = 1, \ldots, p$, $|\mathscr{B}_i| = ma(i)$. (Justificar)

Sea S la matriz $n \times n$ que en sus columnas tiene los n vectores l.i. de $\cup_{i=1}^{p} \mathscr{B}_i$. Es fácil ver que $AS = S\Lambda$. $\square$

Dada una matriz A de tamaño $n \times n$, ¿qué herramientas tenemos hoy para saber si es diagonalizable?

Podemos seguir el siguiente método:

1. Calcular los autovalores diferentes de A: $\lambda_i, i = 1, \ldots, p$.
2. Para todo $i = 1, \ldots, p$, calcular una base $\mathscr{B}_i$ de $N(A - \lambda_i I)$.
3. Si $\sum_{i=1}^{p} |\mathscr{B}_i| < n$, A no es diagnalizable. PARAR.
4. A es diagonalizable y la matriz S que tiene por columnas los vectores en $\cup_{i=1}^{p} \mathscr{B}_i$ es matriz diagonalizante de A.

El paso complicado en este método es el paso 1, el cálculo de los autovalores de A.

Hasta ahora, sólo sabemos hacerlo calculando su polinomio característico y luego buscando sus raíces. Ambas cosas son muy pesadas computacionalmente.

El polinomio característico requiere del desarrollo de un determinante (que en matrices grandes es complejo).

Para la búsqueda de las raíces del polinomio no existen fórmulas explícitas a partir de polinomios de grado 5 y debemos hacer uso de métodos numéricos.

En realidad, los autovalores se calculan por una secuencia de transformaciones de similitud simples, llevándola a su forma triangular (Lema de Schur) donde los autovalores aparecen en la diagonal.