DEPARTAMENTO DE COMPUTACION
Facultad de Ciencias Exactas y Naturales - UBA

Métodos Numéricos modo virtual
(pandemia COVID-19)
Material Complementario

Autovalores - versión 1.0

En este documento presentamos algunas propiedades de los autovalores de una matriz que necesitamos para fundamentar métodos numéricos relacionados con el cálculo de autovalores así como también en el contexto de SVD que veremos más adelante. Es material complementario de las diapos de la clase de autovalores usadas durante el dictado virtual (pandemia-COVID-19).
No pretende ser un documento que desarrolle toda la teoría de autovalores. Aquellos interesados en profundizar pueden consultar la extensa bibliografía que existe sobre el tema.

Comenzamos recordando la definición.

Definición: *Dada $A \in \mathbb{C}^{n\times n}$, $\lambda \in \mathbb{C}$ es autovalor si existe $v \in \mathbb{C}^n$, $v \neq 0$ tal que $Av = \lambda v$.*

El escalar λ podría ser un número real o complejo y está asociado a las raíces del polinomio característico $P(\lambda) = det(A - \lambda I)$. A continuación presentamos y demostramos algunas propiedades.

Proposición: *Sea λ autovalor de A y v autovector asociado. Entonces:*

1. *$\lambda - \alpha$ es autovalor de $A - \alpha I$ y v autovector asociado*
2. *λ^k es autovalor de A^k y v autovector asociado*
3. *Si A es matriz ortogonal $\Rightarrow |\lambda| = 1$*
4. *λ es autovalor de A^t*
5. *Si $\lambda_1, \dots, \lambda_k$ son autovalores distintos, entonces $v_1, \dots, v_k$ autovectores asociados son linealmente independientes.*

Demostración:

1. Basta hacer el producto: $(A - \alpha I)v = Av - \alpha v = \lambda v - \alpha v = (\lambda - \alpha)v$
   Resulta entonces que $(\lambda - \alpha)$ es autovalor y v autovector asociado.

2. Por inducción.

   Caso base k=1: es válido ya que λ es autovalor de A.
   Paso Inductivo: supongamos que λ^k es autovalor de A^k, queremos ver que λ^{k+1} es autovalor de A^{k+1}.

$$A^{k+1}v = A(A^k v) \underset{\substack{\downarrow \\ \text{por hip induc}}}{=} \lambda^k A v \overset{\substack{\text{por ser autovector} \\ \uparrow}}{=} \lambda^k \lambda v = \lambda^{k+1} v$$

3. Por ser ortogonal sabemos que $||Av||_2 = ||v||_2$, pero $||Av||_2 = ||\lambda v||_2 = |\lambda|||v||_2$. Entonces $|\lambda| = 1$.

4. Sabemos que los autovalores de A son las raíces de $P(\lambda) = det(A - \lambda I)$. Pero $det(A - \lambda I) = det((A - \lambda I)^t) = det(A^t - \lambda I)$. Entonces el polinomio característico de A y de A^t es el mismo y sus raíces también lo son.

5. Por inducción.

   Caso base k=1: es válido ya que v_1 no es el vector nulo.

   Paso inductivo: asumimos que $v_1, \ldots, v_k$ son linealmente independientes, queremos ver que $v_1, \ldots, v_{k+1}$ también lo son. Supongamos que no, entonces v_{k+1} se puede escribir como combinación lineal de $v_1, \ldots, v_k$

$$v_{k+1} = \sum_{j=1}^{k} c_j v_j$$

   donde los coeficientes c_j no son todos nulos ya que $v_{k+1} \neq 0$ por definición de autovector. Multiplicamos a ambos lados por A:

$$Av_{k+1} = \sum_{j=1}^{k} c_j A v_j$$

   Como $v_1, \ldots, v_{k+1}$ son autovectores, tenemos que

$$\lambda_{k+1} v_{k+1} = \sum_{j=1}^{k} c_j \lambda_j v_j$$

   Por otro lado, si multiplicamos por λ_{k+1} a la expresión de v_{k+1} en función de $v_1, \ldots, k$, tenemos

$$\lambda_{k+1} v_{k+1} = \sum_{j=1}^{k} c_j \lambda_{k+1} v_j$$

   Restando ahora las dos últimas identidades, obtenemos:

$$0 = \sum_{j=1}^{k} c_j (\lambda_j - \lambda_{k+1}) v_j$$

   Por hipótesis inductiva $v_1, \ldots v_k$ son linealmente independientes, entonces la única manera de que una combinación lineal resulte en el vector nulo es que $c_j(\lambda_j - \lambda_{k+1}) = 0$ para todo $j = 1, \ldots, k$. Pero sabemos que existe al menos un coeficiente c_j no nulo lo que implica que $\lambda_j - \lambda_{k+1} = 0$ para algún j, que contradice que los autovalores son distintos. Por lo tanto $v_1, \ldots, v_{k+1}$ son linealmente independientes.

■

La siguiente propiedad nos ayuda para tener idea de cotas superiores e inferiores de los autovalores de una matriz. Puede ser útil para acotar el radio espectral de una matriz (máximo módulo de los autovalores de la matriz). Esto resultará de interés cuando abordemos método iterativos para resolver sistemas de ecuaciones lineales.

Proposición: *Sea* $A \in \mathbb{C}^{n\times n}$, $r_i = \sum_{k=1,k\neq i}^{n} |a_{ik}|$ *para* $i = 1,\dots,n$ *y los discos de Gershgorin*

$$D_i = \{x \in \mathbb{C} : |x - a_{ii}| \leq r_i\} \text{ para } i = 1,\dots,n$$

Si λ *es autovalor de* $A \Longrightarrow \lambda \in D_i$ *para algún* $i = 1,\dots,n$

Demostración:
Sea v el autovector asociado a λ. Como $v \neq 0$ entonces $||v||_\infty = \max\limits_{k=1,\dots,n} |v_k| \neq 0$. Sea k_0 el índice de una coordenada donde de v tal que $||v||_\infty = |v_{k_0}|$. Sabemos que $Av = \lambda v$. En particular, considerando la fila k_0, tenemos que:

$$\sum_{j=1}^{n} a_{k_0 j} v_j = \lambda v_{k_0}$$

Separando de la sumatoria el término k_0

$$\sum_{j=1,j\neq k_0}^{n} a_{k_0 j} v_j + a_{k_0 k_0} v_{k_0} = \lambda v_{k_0}$$

$$\sum_{j=1,j\neq k_0}^{n} a_{k_0 j} v_j = (\lambda - a_{k_0 k_0}) v_{k_0}$$

Tomamos módulo

$$\sum_{j=1,j\neq k_0}^{n} |a_{k_0 j}||v_j| \geq |\sum_{j=1,j\neq k_0}^{n} a_{k_0 j} v_j| = |(\lambda - a_{k_0 k_0})||v_{k_0}|$$

Como $|v_{k_0}| \neq 0$ podemos pasar dividiendo

$$\sum_{j=1,j\neq k_0}^{n} |a_{k_0 j}|\frac{|v_j|}{|v_{k_0}|} \geq |(\lambda - a_{k_0 k_0})|$$

Como $||v||_\infty = |v_{k_0}|$, entonces $\frac{|v_j|}{|v_{k_0}|} \leq 1$ y por lo tanto

$$\sum_{j=1,j\neq k_0}^{n} |a_{k_0 j}| \geq \sum_{j=1,j\neq k_0}^{n} |a_{k_0 j}|\frac{|v_j|}{|v_{k_0}|} \geq |(\lambda - a_{k_0 k_0})|$$

Concluimos entonces que $\lambda \in D_{k_0}$.

■

En las diapos hemos visto ejemplos de matrices para las cuales podíamos identificar una base de autovectores y otras matrices para las cuales no existe una base de autovectores. Dadas las propiedades que probamos más arriba, sabemos que en el caso que todos los autovalores sean distintos tendremos una base. Sin embargo, cuando existen autovalores de multiplicidad al menos 2, puede ocurrir que no exista una base de autovectores. A continuación presentaremos propiedades que nos permiten identificar si una matriz tiene base de autovectores.
Comenzamos analizando matrices semejantes y su inter-relación respecto a los autovalores.

Definición: *Sean* $A, B \in \mathbb{C}^{n\times n}$. *Diremos que* A *y* B *son semejantes sii existe una matriz* P *inversible tal que*

$$A = PBP^{-1}$$

Proposición: *Si* $A, B \in \mathbb{C}^{n\times n}$ *son semejantes entonces tienen los mismos autovalores.*

Demostración:
Sea λ autovalor de A y v un autovector asociado. Queremos ver que λ es autovalor B. Sabemos que

$$Av = \lambda v$$

Multiplicando a ambos lados por P^{-1} obtenemos

$$P^{-1}Av = \lambda P^{-1}v$$

Por la relación de semejanza entre A y B

$$BP^{-1}v = \lambda P^{-1}v$$

Como P es inversible y v no es nulo, entonces $P^{-1}v \neq 0$ y podemos concluir que λ es autovalor de B con autovector asociado $P^{-1}v$.
Sea λ autovalor de B y v un autovector asociado. Queremos ver que λ es autovalor A. Sabemos que

$$Bv = \lambda v$$

Multiplicando a ambos lados por P obtenemos

$$PBv = \lambda Pv$$

Por la relación de semejanza entre A y B

$$APv = \lambda Pv$$

Como P es inversible y v no es nulo, entonces $Pv \neq 0$ y podemos concluir que λ es autovalor de A con autovector asociado Pv. ■
En el caso que una matriz A sea semejante a una matriz diagonal, podremos derivar propiedades muy importantes acerca de la existencia de base de autovectores. Analicemos esto a continuación.

Definición: *Sean $A \in \mathbb{C}^{n\times n}$. Diremos que A es diagonalizable por semejanza sii es semejante a una matriz D diagonal.*

Proposición: *Sea $A \in \mathbb{C}^{n\times n}$. A es diagonalizable por semejanza sii tiene base de autovectores.*

Demostración:
$\Rightarrow$)
Si A es diagonalizable por semejanza, entonces existe P matriz inversible tal que $A = PDP^{-1}$. Por la propiedad demostrada anteriormente sabemos que si v es autovector de D entonces Pv es autovector de A. Los autovectores de una matriz diagonal son los vectores canónicos $e_1, \ldots, e_n$. Entonces Pe_i es autovector de A. Pero Pe_i es la columna i-ésima de P y por ser inversible tiene sus columnas linealmente independientes. Entonces $Pe_1, \ldots, Pe_n$ son los autovectores de A y forman una base.
$\Leftarrow$)
Si A tiene base de autovectores, entonces existen $v_1, \ldots, v_n$ vectores linealmente independientes tal que $Av_i = \lambda_i v_i$ siendo λ_i el autovalor al cual están asociados. Entonces

$$A[v_1, \ldots, v_n] = [\lambda_1 v_1, \ldots, \lambda_n v_n]$$

$$A[v_1, \ldots, v_n] = [v_1, \ldots, v_n] \begin{bmatrix} \lambda_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ \ldots & \ldots & \lambda_n \end{bmatrix}$$

Si definimos $P = [v_1, \ldots, v_n]$ resulta ser una matriz inversible y

$$AP = PD$$
$$A = PDP^{-1}$$

Concluimos que A es diagonalizable por semejanza. ■

Con las propiedades enunciadas hasta el momento estamos en condiciones *teóricas* de establecer si una matriz tiene base de autovectores: necesitamos saber si es diagonalizable por semejanza.
Lo que quisiéramos es poder afirmar esto a partir de alguna otra propiedad *más sencilla* de verificar, al menos para alguna familia de matrices. Este es el propósito de lo que sigue a continuación.
A partir de ahora consideraremos matrices con coeficientes reales. Algunos resultados también son válidos para matrices con coeficientes complejos, pero simplicaremos el desarrollo al considerar valores reales.

Proposición: *Sea $A \in \mathbb{R}^{n\times n}$. Las siguientes propiedades son válidas:*

1. *Si λ es un autovalor de A tal que $\lambda \in \mathbb{R}$ entonces tiene un autovector asociado con coeficientes reales.*
2. *Si A es simétrica sus autovalores son números reales.*
3. *Si A es simétrica y λ_1, λ_2 son autovalores distintos entonces los autovectores asociados son ortogonales.*

Demostración:

1. Un autovector v asociado a λ es solución del sistema homogéneo $(A - \lambda I)v = 0$. Este sistema puede resolverse usando, por ejemplo, eliminación gaussiana. Este procedimiento realiza operaciones elementales entre los coeficientes de A. Como $A \in \mathbb{R}^{n\times n}$, entonces el algoritmo opera con aritmética en los reales por lo tanto la solución que obtiene tendrá coeficientes reales.

2. Sea λ autovalor de A con $v \neq 0$ autovector asociado.

$$Av = \lambda v$$

Tomamos valor conjugado en ambos lados:

$$\bar{A}\bar{v} = \bar{\lambda}\bar{v}$$

Como $A \in \mathbb{R}^{n\times n}$, entonces $\bar{A} = A$

$$A\bar{v} = \bar{\lambda}\bar{v}$$

Multiplicamos a ambos lados por v^t

$$v^t A\bar{v} = \bar{\lambda}v^t\bar{v}$$

$$(A^t v)^t\bar{v} = \bar{\lambda}v^t\bar{v}$$

Como A es simétrica,

$$(Av)^t\bar{v} = \bar{\lambda}v^t\bar{v}$$

Como v es autovector asociado a λ,

$$\lambda v^t\bar{v} = \bar{\lambda}v^t\bar{v}$$

Como $v^t\bar{v} \neq 0$, entonces $\lambda = \bar{\lambda}$. Concluimos que λ es real.

3. Sean λ_1, λ_2 autovalores distintos de A y v_1, v_2 los autovectores asociados. Entonces:

$$Av_1 = \lambda_1 v_1$$
$$Av_2 = \lambda_2 v_2$$

A la primera identidad la multiplicamos por v_2^t y a la segunda por v_1^t:

$$v_2^t Av_1 = \lambda_1 v_2^t v_1$$

$$v_1^t A v_2 = \lambda_2 v_1^t v_2$$

Como A es simétrica,

$$v_2^t A v_1 = (A^t v_2)^t v_1 \underset{\substack{\downarrow \\ \text{A simétrica}}}{=} (Av_2)^t v_1 = (v_1^t A v_2)^t \underset{\substack{\downarrow \\ \text{es un número}}}{=} v_1^t A v_2$$

Entonces

$$\lambda_1 v_2^t v_1 = \lambda_2 v_1^t v_2$$

Como por hipótesis $\lambda_1 \neq \lambda_2$ y además $v_2^t v_1 = v_1^t v_2$, se deriva que para que valga la identidad deberá cumplirse que $v_2^t v_1 = v_1^t v_2 = 0$. Concluímos que los autovectores son ortogonales.

■

Finalmente daremos condiciones para que una matriz A sea diagonalizable por semejanza. Comenzaremos por una versión simplificada del teorema de Schur que establece la semejanza de una matriz a una matriz triangular superior. Este será un punto inicial para llegar, para ciertas matrices, a determinar la semejanza a matrices diagonales.

Proposición: *Si $A \in \mathbb{R}^{n\times n}$ tiene todos sus autovalores reales, entonces existe $Q \in \mathbb{R}^{n\times n}$ ortogonal y $T \in \mathbb{R}^{n\times n}$ triangular superior tal que $A = QTQ^t$.*

Demostración:
Sea λ_1 autovalor de A y v_1 autovector asociado. Sin pérdida de generalidad podemos asumir $||v_1||_2 = 1$. Sabemos que existe H_1 una tranformación de Householder tal que $H_1 v_1 = e_1$ con e_1 el primer vector canónico. $H_1 A H_1^t$ es una matriz tal que su primera columna es $\lambda_1 e_1$. ¿Cómo vemos esto? La primera columna de una matriz es el resultado de multiplicarla por e_1. Entonces,

$$H_1 A H_1^t e_1 \underset{\substack{\downarrow \\ \text{por Householder}}}{=} H_1 A v \overset{\substack{\text{por ser autovector} \\ \uparrow}}{=} \lambda_1 H_1 v_1 \underset{\substack{\downarrow \\ \text{por Householder}}}{=} \lambda_1 e_1$$

Por definición, A es semejante a $H_1 A H_1^t$. Por lo tanto tienen los mismos autovalores. Por las características de la primera columna de $H_1 A H_1^t$, sabemos que uno de los autovalores es el coeficiente que se encuentra en primera fila y primera columna (notemos λ_1) y el resto de los autovalores corresponden a aquellos que surjan de la matriz $\tilde{A}$ de $(n-1)\times(n-1)$ conformada por las filas 2 a n y columnas 2 a n.
Si ahora aplicamos el mismo razonamiento a $\tilde{A}$, podemos afirmar que existe $\tilde{H}_2 \in \mathbb{R}^{(n-1)\times(n-1)}$ matriz ortogonal tal que $\tilde{H}_2 \tilde{A} \tilde{H}_2{}^t$ tiene la propiedad de que su primera columna tiene un autovalor en la primera posición (notemos λ_2) y ceros desde la segunda posición en adelante. La submatriz de $(n-2)\times(n-2)$ conformada por la filas y columnas de 2 a $n-1$ tiene el resto de los autovalores de A. Si expandimos a la matriz $\tilde{H}_2$ agregándole una primera fila y columna igual al primer vector canónico, obtenemos $H_2 \in \mathbb{R}^{n\times n}$ matriz ortogonal y se verifica que:

$$H_2 H_1 A H_1^t H_2^t = \begin{bmatrix} \lambda_1 & * & \dots & * \\ 0 & \lambda_2 & \dots & * \\ 0 & 0 & \dots & * \\ \vdots & \vdots & \dots & * \\ 0 & 0 & \dots & * \end{bmatrix}$$

Aplicando este proceso $n-1$ veces llegaremos a una matriz triangular superior. Como el producto de matrices ortogonales es ortogonal, entonces podemos afirmar que existe H tal que

$$HAH^t = T$$

$$A = H^t T H = Q T Q^t$$

■

Ya vimos que las matrices simétricas con coeficientes reales tienen autovalores reales. Por lo tanto, estas matrices cumplen las hipótesis de la versión simplificada del teorema de Schur y podemos afirmar que son semejantes a una matriz triangular. Aún más, por ser simétricas podemos afirmar que la matriz T es diagonal. Es el resultado que presentamos a continuación

Proposición: *Sea $A \in \mathbb{R}^{n\times n}$ matriz simétrica. Entonces A es diagonalizable por semejanza.*

Demostración:
Aplicando la versión simplificada el teorema de Schur a la matriz A, sabemos que existe Q matriz ortogonal y T matriz tringular superior tal que

$$A = Q T Q^t$$

Tomando la traspuesta a ambos lados

$$A^t = Q T^t Q^t$$

Como A es simétrica, concluimos que $QT^tQ^t = QTQ^t$ y por lo tanto $T^t = T$. Siendo T una matriz triangular superior, podemos afirmar que T es una matriz diagonal. Es decir, A es diagonalizable por semejanza. ■

Corolario: *Sea $A \in \mathbb{R}^{n\times n}$ matriz simétrica. Entonces A tiene base (ortonormal) de autovectores.*

Demostración:
Al ser simétrica, por la propiedad anterior sabemos que es diagonalizable por semejanza. Por la propiedad vista más arriba, esto implica que tiene base de autovectores. Sabemos que los autovectores corresponden a las columnas de la matriz Q que establece la semejanza, y como Q es ortogonal, concluimos que la base de autovectores es ortonormal.

■