UNIVERSIDAD NACIONAL DE INGENIERÍA
FACULTAD DE CIENCIAS
ESCUELA PROFESIONAL DE MATEMÁTICA

Descomposición del valor singular y sus aplicaciones

por

Marlene Jenny Soldevilla Olivares

Tésis para Optar
el Título Profesional de
LICENCIADO en MATEMÁTICA

Prof. William Carlos Echegaray Castillo
Asesor

UNI, enero del 2004.

RESUMEN

El presente trabajo de **Descomposición de Valores Singulares** de matrices $A \in \mathbb{R}(m, n)$ es una de las herramientas más importantes y de una amplia variedad de aplicaciones del álgebra y del análisis numérico, que solamente requieren el conocimiento de rango de matrices; conjuntos ortonormales, valores y vectores propios de matrices de orden $n \times n$, entre otros.

En el capítulo 2 hacemos un breve estudio del cálculo de los valores y vectores propios asociados a matrices de onden $n \times n$.

En el capítulo 3 hacemos el estudio de la Descomposición del Valor Singular de matrices reales de orden $m \times n$, usando el capítulo 2, (esto se hace calculando los valores propios de la matrices $A^T A$).

En el capítulo 4 hacemos algunas aplicaciones de la Descomposición de los Valores Singulares, tales como Solución de Ecuaciones Lineales "grandes". El Problema de Almacenamiento de matrices de ordenes grandes (en la memoria del computador), también producto de matrices, etc.

A mis padres Santiago y Cristina
que siempre me alentaron para
culminar mi carrera y este trabajo

Agradezco al Profesor William Carlos Echegaray Castillo por la orientación y sus sabios consejos para la culminación del presente trabajo y también deseo agradecer a María Aliaga por su amistad y estimulo, y a todas aquellas personas que de una u otra forma me ayudaron a terminar este trabajo.

ÍNDICE

1 INTRODUCCIÓN

El principal objetivo del presente trabajo, es usar las propiedades de los valores propios de matrices cuadradas $A \in \mathbb{R}(n,n)$ (ó $A \in \mathbb{C}(n,n)$), donde $\mathbb{R}(n,n)$ (ó $\mathbb{C}(n,n)$) representa el conjunto de matrices reales (ó complejas). En general $\mathbb{R}(m,n)$ (ó $\mathbb{C}(m,n)$) representa el conjunto de matrices de orden $m \times n$ reales (ó complejas.)

En el capítulo 2 se ha desarrollado: localización de valores propios, también el cálculo de valores propios de matrices $A \in \mathbb{R}(n,n)$, así como sus respectivos vectores propios asociados.

En el capítulo 3 desarrollamos la descomposición de los valores singulares de matrices $A \in \mathbb{R}(m,n)$, también la sensibilidad de los valores singulares, es decir, comparar los valores singulares de A y $A+E$, donde $E \in \mathbb{R}(m,n)$ es la matriz de error, utilizando los conceptos del capítulo 2.

En el capítulo 4 desarrollamos algunas aplicaciones tales como: normas, solución de sistemas de ecuaciones lineales, el problema de almacenamiento (es decir, matrices de orden muy grande), producto de matrices, etc. usando la descomposición de los valores singulares de $A \in \mathbb{R}(m,n)$.

A continuación damos algunas definiciones que serán usadas en el presente trabajo.

1.1 Conceptos Preliminares

Definición 1.1. *Sea $A \in \mathbb{C}(n,n)$ una matriz, A es llamada* **UNITARIA** *si, y sólo si $A^*A = AA^* = I$, donde $A^* = (\overline{A})^T$*

Definición 1.2. *Sean $A \in \mathbb{R}(m,n)$ una matriz, $x \in \mathbb{R}^n$ un vector, definimos la siguientes normas de A:*

1. $\|A\|_2 = \sup\limits_{x\neq 0}\left(\left\{\frac{\|Ax\|_2}{\|x\|_2}\right\}\right)$, *llamada la norma 2 de A, donde*

   $\|x\|_2 = \sqrt{x_1^2 + \cdots + x_n^2}$.

2. $\|A\|_F = \sqrt{\left(\sum\limits_{j=1}^{n}\sum\limits_{i=1}^{m} a_{ij}^2\right)}$, *llamada la norma de Frobenius.*

3. $\|A\|_\infty = \sup\limits_{x\neq 0}\left(\left\{\frac{\|Ax\|_\infty}{\|x\|_\infty}\right\}\right)$, *llamada la norma infinita de A, donde*

   $\|x\|_\infty = \max\limits_{1\leq i\leq n}(\{|x_i|\})$.

Definición 1.3. *Sea $A \in \mathbb{R}(n,n)$ una matriz no-singular, definimos la Condición de A con respecto a la norma 2 al número:*

$$Cond_2(A) = \|A\|_2\|A^{-1}\|_2. \tag{1.1}$$

A continuación presentamos algunas propiedades

Proposición 1.1. *Sea A una matriz no singular de orden $n \times n$, Entonces*

1. $Cond(A) = Cond(A^T)$.
2. $Cond(AB) \leq Cond(A)Cond(B)$.
3. $Cond(\alpha A) = Cond(A)$, $\forall\ \alpha \in \mathbb{R}\backslash\{0\}$.

Prueba:

1. Sabemos que: $$\begin{aligned} Cond(A) &= \|A\|\|A^{-1}\| \\ &= \|A^T\|\|(A^{-1})^T\| \\ &= \|A^T\|\|(A^T)^{-1}\| \\ &= Cond(A^T). \end{aligned}$$

   Entonces $\qquad Cond(A) = Cond(A^T)$.

2. $$\begin{aligned} Cond(AB) &= \|AB\|\|(AB)^{-1}\| \\ &= \|AB\|\|B^{-1}A^{-1}\| \\ &\leq \|A\|\|B\|\|A^{-1}\|\|B^{-1}\| = \\ &= \|A\|\|A^{-1}\|\|B\|\|B^{-1}\| \\ &= Cond(A)Cond(B). \end{aligned}$$

   Entonces $$Cond(AB) \leq Cond(A)Cond(B).$$

3. $$\begin{aligned} Cond(\alpha A) &= \|\alpha A\|\|(\alpha A)^{-1}\| \\ &= |\alpha|\|A\|\|(\alpha)^{-1}A^{-1}\|,\ \forall\ \alpha \in \mathbb{R}\backslash\{0\} \\ &= |\alpha|\|A\||\alpha|^{-1}\|A^{-1}\|,\ \forall\ \alpha \in \mathbb{R}\backslash\{0\} \\ &= \|A\|\|A^{-1}\| \\ &= Cond(A). \end{aligned}$$

   Entonces $$Cond(\alpha A) = Cond(A),\ \forall\ \alpha \in \mathbb{R}\backslash\{0\}$$ ■

Definición 1.4. *Sea A una matriz no-singular, diremos que el sistema $Ax = b$ está* **mal condicionado** *si $Cond(A)$ es bastante "grande", en otro caso diremos que está* **bién condicionado**.

Ejemplo 1.1. *Consideremos el sistema $Ax = b$, donde:*

$$A = \begin{pmatrix} 1 & 0.9999 \\ 0.9999 & 1 \end{pmatrix},\ b = \begin{pmatrix} 1 \\ 2 \end{pmatrix},\ x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Notamos que $A^{-1} = 10^3 \begin{pmatrix} 5.00025001250243 & -4.999974998750118 \\ -4.99997499870118 & 5.00025001250243 \end{pmatrix}.$

Además $$\begin{aligned} \|A\| &= 1.9999, \\ \|A^{-1}\| &= 1.000000000000361 \times 10^4 \\ Cond(A) &= 1.999900000000723 \times 10^4. \end{aligned}$$

Observamos que es un número "grande" entonces cualquier sistema que formemos con la matriz A, diremos que dicho sistema está mal condicionado.

En particular con el sistema dado tenemos que dicha solución es:

$$x = \begin{pmatrix} 5000.75003750368 \\ -4999.24996249993 \end{pmatrix}. \qquad \square$$

Además de eso necesitamos algunos conceptos de orden, consideraremos también $\mathbb{R}^+ = \{x \in \mathbb{R}/x > 0\}$,

Definición 1.5. *Sea* $f : \mathbb{N} \longrightarrow \mathbb{R}^+$. *El orden de* f *se define como* $O(f) = \{g : \mathbb{N} \longrightarrow \mathbb{R}^+/\exists c \in \mathbb{R}^+ \text{ y } \exists n_0 \in \mathbb{N}, g(n) \leq cf(n), \forall n \geq n_0\}$.

Existe otra alternativa para verificar que $g \in O(f)$:

$$g \in O(f) \text{ sí } \lim_{n\to\infty} \frac{g(n)}{f(n)} = c, \text{ para algún } c \in \mathbb{R}^* = \mathbb{R}^+ \cup \{0\}, \text{ con } c \neq \infty$$

Ejemplo 1.2. *Sea* $f(n) = \dfrac{n^3}{2}$ *y* $g(n) = 37n^2 + 120n + 17$. *Probaremos que* $g \in O(f)$, *pero* $f \notin O(g)$.
Observe que $\forall n \geq 78$ *tenemos que* $g(n) \leq 1.f(n)$ *entonces* $g \in O(f)$.
Sí aplicamos la definición alternativa se obtiene:

$$\lim_{n\to\infty} \frac{g(n)}{f(n)} = \lim_{n\to\infty} \frac{37n^2 + 120n + 17}{n^3/2} = 0 \Longrightarrow c = 0$$

entonces $g \in O(f)$.
Ahora, veamos que $f \notin O(g)$:

$$\lim_{n\to\infty} \frac{f(n)}{g(n)} = \lim_{n\to\infty} \frac{n^3/2}{37n^2 + 120n + 17} = \infty$$

Esto nos lleva a la siguiente conclusión, entonces vamos a suponer que $f \in O(g)$, *entonces* $\exists c \in \mathbb{R}^+$ *y* $\exists n_0 \in \mathbb{N}$ *tal que* $\forall n \geq n_0$ *se tiene:*

$$\frac{n^3}{2} \leq (37n^2 + 120n + 17)c$$

Así,

$$\frac{n}{2} \leq 37c + \frac{120c}{n} + \frac{17c}{n^2} \leq 174c$$

entonces, $\dfrac{n}{2} \leq 174c$, *esto es una contradicción, debido a que* $\mathbb{N}$ *estaría acotado.* $\square$

Ejemplo 1.3. *Sean* $f(n) = n^2$ *y* $g(n) = nLn(n)$. *Mostraremos que* $g \in O(f)$, *pero* $f \notin O(g)$, *es decir,*

$$\lim_{n\to\infty} \frac{g(n)}{f(n)} = \lim_{n\to\infty} \frac{nLn(n)}{n^2} = \lim_{n\to\infty} \frac{Ln(n)}{n} \overset{L'Hospital}{=} \lim_{n\to\infty} \frac{1/n}{1} = 0,$$

entonces $g \in O(f)$,

además, como

$$\lim_{n\to\infty} \frac{f(n)}{g(n)} = \infty,$$

tenemos que $f \notin O(g)$. $\square$

El orden O posee las siguientes propiedades que son dadas en el siguiente:

Teorema 1.1. *Sean* $f, g, h : \mathbb{N} \longrightarrow \mathbb{R}^*$*:*

1. *Transitividad: Sí* $f \in O(g)$ *y* $g \in O(h)$, *entonces* $f \in O(h)$.
2. $O(f + g) = O(max(\{f, g\}))$.
3. $O(\alpha f) = O(f)$, $\forall\ \alpha > 0$

Prueba:

1. Sean $c_1 \in \mathbb{R}^+$ y $n_1 \in \mathbb{N}$ tales que $f(n) \leq c_1 g(n)$, $\forall\ n \geq n_1$,

   $c_2 \in \mathbb{R}^+$ y $n_2 \in \mathbb{N}$ tales que $g(n) \leq c_2 h(n)$, $\forall\ n \geq n_2$,

   definamos $n_3 = max(\{n_1, n_2\})$, entonces

   $f(n) \leq c_3 h(n)$, $\forall\ n \geq n_3$, donde $c_3 = c_1 c_2$;

   entonces $f \in O(h)$.

2. Sí $h \in O(f+g) \Longrightarrow \exists c \in \mathbb{R}^+, n_1 \in \mathbb{N}\ /\ h(n) \leq c(f(n)+g(n)), \forall\ n \geq n_1$

   pero $f(n) + g(n) \leq 2max(\{f, g\})(n)$

   $$\Longrightarrow h(n) \leq 2c\ max(\{f, g\})(n),\ \forall\ n \geq n_1$$
   $$\Longrightarrow h \in O(max(\{f, g\})).$$

Sí $h \in O(max(\{f,g\})) \Longrightarrow \exists c \in \mathbb{R}^+, n_1 \in \mathbb{N} \;/\; h(n) \leq c\; max(\{f,g\})(n)$

pero $max(\{f,g\})(n) \leq f(n) + g(n)$

$$\Longrightarrow h(n) \leq c\; (f(n) + g(n)),\; \forall\; n \geq n_1$$

$$\Longrightarrow h \in O(f+g).$$

Luego $O(f+g) = O(max(\{f,g\}))$.

3. Sea $g \in O(\alpha f) \Longrightarrow \exists c \in \mathbb{R}^+$, $\exists n_0 \in \mathbb{N}$ tal que

$$g(n) \leq c(\alpha f(n)) = (c\alpha) f(n),\; \forall\; n \geq n_0$$

hagamos $c' = c\alpha > 0$

$$\Longrightarrow g(n) \leq c' f(n),\; \forall\; n \geq n_0$$

$$\Longrightarrow g \in O(f)$$

$$\Longrightarrow O(\alpha f) \subset O(f) \qquad (i)$$

Sea $g \in O(f) \Longrightarrow \exists c \in \mathbb{R}^+$, $\exists n_0 \in \mathbb{N}$ tal que:

$$g(n) \leq c f(n) = \left(\frac{c}{\alpha}\right) \alpha f(n),\; \forall\; n \geq n_0,\; \alpha > 0$$

hagamos $c' = \dfrac{c}{\alpha} > 0$

$$\Longrightarrow g(n) \leq c'(\alpha f(n)),\; \forall\; n \geq n_0,\; \alpha > 0$$

$$\Longrightarrow g \in O(\alpha f)$$

$$\Longrightarrow O(f) \subset O(\alpha f) \qquad (ii)$$

Luego de (i) y (ii) tenemos:

$$O(\alpha f) = O(f), \qquad \forall\; \alpha > 0 \qquad \square$$

2 EL PROBLEMA DE VALORES PROPIOS DE UNA MATRIZ

En este capítulo desarrollaremos el estudio numérico del problema de valores propios de una matriz. Estos problemas son muy importantes y ofrecen una variedad de aplicaciones en matemática, ingeniería, estadística, economía, etc.

2.1 Conceptos Básicos

Definición 2.1. *Sea $A \in \mathbb{C}(m,n)$, denotamos a la conjugada de A, mediante* $\overline{A} \in \mathbb{C}(m,n)$.

La transpuesta de la conjugada de A es denotada por A^*, es decir, $A^* = (\overline{A})^T$.

Si $A \in \mathbb{C}(n,n)$ y $A^* = A$, entonces A es llamada matriz HERMITIANA,

A es llamada matriz UNITARIA si $A^*A = AA^* = I$.

Una matriz hermitiana A es DEFINIDA POSITIVA si $x^*Ax > 0, \forall x \in \mathbb{C}\backslash 0$.

A es semidefinida positiva si $x^*Ax \geq 0, \forall x \in \mathbb{C}$

2.2 Valores Propios y Vectores Propios

Definición 2.2. *Sea $A \in \mathbb{C}(n,n)$ una matriz. Entonces $\lambda \in \mathbb{C}$ es un VALOR PROPIO de A, si existe un vector $x \in \mathbb{C}^n\backslash\{0\}$ tal que:*

$$\begin{aligned} Ax &= \lambda x \qquad\qquad \text{ó} \\ (A - \lambda I)x &= 0 \end{aligned}$$

Al vector x también es llamado VECTOR PROPIO a la derecha de A asociado con el VALOR PROPIO λ.

(λ, x) es llamado el PAR PROPIO de A.

Un vector $y \in \mathbb{C}^n \backslash \{0\}$ *que satisface* $y^T A = \lambda y^T$ *es llamado VECTOR PROPIO a la izquierda de* A *asociado con el valor propio* λ.

Definición 2.3. *El polinomio* $P_A(\lambda) = det(A - \lambda I)$ *de grado* n *es llamado POLINOMIO CARACTERÍSTICO DE* A.
Asi, las raices de $P_A(\lambda)$ *son los valores propios de* A.

Oservación 2.1. *De las definiciones anteriores, obtenemos:*

1. *Si* (λ, x) *es un PAR PROPIO de una matriz no singular* A*, entonces* $\left(\frac{1}{\lambda}, x\right)$ *es el par propio de* A^{-1}.
   En efecto:
   como (λ, x) *es un PAR PROPIO de* $A \Rightarrow Ax = \lambda x$
   Siendo A *no singular se tiene:*
   $A^{-1}Ax = x = A^{-1}(\lambda x) = \lambda A^{-1}x$
   $\Longrightarrow \lambda A^{-1}x = x, \lambda \neq 0$
   $\Longrightarrow A^{-1}x = \frac{1}{\lambda}x$
   $\Longrightarrow \left(\frac{1}{\lambda}, x\right)$ *es un par propio de* A^{-1}

2. *Si* (λ, x) *es un par propio de* A*, entonces* $(\lambda - \sigma, x)$ *es un par propio de* $A - \sigma I$*, donde* $\sigma \in \mathbb{C}$
   En efecto:
   como (λ, x) *es un par propio de* $A \Longrightarrow Ax = \lambda x$*, además*

$$\begin{aligned} (A - \sigma I)x &= Ax - \sigma I x \\ &= \lambda x - \sigma x \\ &= (\lambda - \sigma)x \end{aligned}$$

$\Longrightarrow (\lambda - \sigma, x)$ *es un par propio de* $A - \sigma I$.

Definición 2.4. *El conjunto de valores propios de una matriz* A *es llamado el ESPECTRO de* A.

Ahora presentamos algunos resultados básicos de valores propios y vectores propios.

Teorema 2.1. *Los valores propios de una matriz triangular A son los elementos de su diagonal.*

PRUEBA: Sin pérdida de generalidad podemos suponer que A es una matriz triangular superior.

$\Longrightarrow A - \lambda I$ también es una matriz triangular superior,

$\Longrightarrow det(A - \lambda I) = \prod_{i=1}^{n}(a_{ii} - \lambda)$

$\Longrightarrow a_{11}, a_{22}, ..., a_{nn}$ son los valores propios de A.

En forma similar se obtiene el mismo resultado si A es una matriz triangular inferior. ■

Teorema 2.2 (GENERALIZACIÓN DE TEOREMA 2.1). *Sea*

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} & \cdots & A_{1k} \\ 0 & A_{22} & A_{23} & \cdots & A_{2k} \\ 0 & 0 & A_{33} & \cdots & A_{3k} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & A_{kk} \end{bmatrix}$$

donde cada $A_{ii} \in \mathbb{C}(l, l)$ es una matriz. Entonces el espectro de A es igual a la unión de los espectros de $A_{ii}, i = 1, 2, ..., k$.

PRUEBA: Se observa que A es una matriz del orden $kl \times kl$, entonces

$$det(A - \lambda I_{kl}) = det(A_{11} - \lambda I_l) \cdots det(A_{kk} - \lambda I_l)$$

Luego, el conjunto de ceros del polinomio característico de A es igual a la unión del conjunto de ceros de los polinomios característicos de $A_{11}, \cdots, A_{kk}$. ■

Teorema 2.3. *Dos matrices similares tienen los mismos valores propios.*

PRUEBA: Sean A, B dos matrices similares

$\Longrightarrow \exists T$ una matriz no-singular tal que, $B = TAT^{-1}$

Veamos que A y B tienen los mismos valores propios:

$$\begin{aligned}
det(B-\lambda I) &= det(TAT^{-1}-\lambda I)\\
&= det(TAT^{-1}-\lambda TT^{-1})\\
&= det(T(A-\lambda I)T^{-1})\\
&= det(T)det(A-\lambda I)det(T^{-1})\\
&= det(TT^{-1})det(A-\lambda I)\\
&= det(A-\lambda I)
\end{aligned}$$

$\Longrightarrow$ *A y B* tienen los mismos polinomios característicos y por tanto los mismos valores propios. ■

Nota 2.1. *El recíproco del teorema (2.3) es falso.*

Veamos el siguiente contraejemplo: sean las matrices

$$A=\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix} \quad y \quad B=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix},$$

entonces

$$A-\lambda I=\begin{bmatrix}1-\lambda & 1\\ 0 & 1-\lambda\end{bmatrix}, \quad B-\lambda I=\begin{bmatrix}1-\lambda & 0\\ 0 & 1-\lambda\end{bmatrix}$$

$det(A-\lambda I)=(1-\lambda)^2=det(B-\lambda I)$

se observa que tienen los mismos valores propios,

ahora supongamos que existe una matriz $T=\begin{bmatrix}t_{11} & t_{12}\\ t_{21} & t_{22}\end{bmatrix}$, *no singular tal que*

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

$$\Longrightarrow A=\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}=T\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}T^{-1}=TT^{-1}=I=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}=B$$

$\Longrightarrow A=B$ *contradicción.* ■

Teorema 2.4. *Sea A una matriz de orden $m \times m$. Entonces los vectores propios asociados con los valores propios distintos son linealmente independientes.*

PRUEBA: La demostración la haremos por inducción sobre m.

Sean $\lambda_1, \lambda_2, \cdots, \lambda_m$ los valores propios de A tales que $\lambda_i \neq \lambda_j$ $\forall$ $i \neq j$ y sus correspondientes vectores propios $x_1, x_2, \cdots, x_m$

Supongamos que para $(m-1)$ se satisface el teorema.

Ahora probaremos que también se cumple para m.

Consideremos $\{c_1, c_2, \cdots, c_m\} \subset \mathbb{R}$ constantes tales que:

$$\sum_{i=1}^{m} c_i x_i = 0 \tag{2.1}$$

Veamos que $c_1 = c_2 = \cdots = c_m = 0$:

En (2.1) multiplicamos por $(\lambda_1 I - A)(\lambda_2 I - A)\cdots(\lambda_{m-1}I - A)$ obteniéndose:

$$(\lambda_1 I - A)(\lambda_2 I - A)\cdots(\lambda_{m-1}I - A)\sum_{i=1}^{m} c_i x_i = 0 \tag{2.2}$$

observar que: $$(\lambda_i I - A)(\lambda_j I - A) = (\lambda_j I - A)(\lambda_i I - A)\ \forall i \neq j. \tag{2.3}$$

Luego para $i = 1, \cdots, m-1$ y sabiendo $Ax_i = \lambda_i x_i$ $(\equiv (\lambda_i I - A)x_i = 0)$ tenemos:

$$c_i(\lambda_1 I - A)\cdots(\lambda_{m-1}I - A)x_i = c_i \left[\prod_{\substack{j=1 \\ j\neq i}}^{m-1}(\lambda_j I - A)\right](\lambda_i I - A)x_i = 0.$$

Luego de (2.2) se tiene:

$$c_m(\lambda_1 I - A)\cdots(\lambda_{m-2}I - A)(\lambda_{m-1}I - A)x_m = 0$$

$$\xRightarrow{Ax_m = \lambda_m x_m} \quad c_m(\lambda_1 I - A)\cdots(\lambda_{m-2}I - A)(\lambda_{m-1} - \lambda_m)x_m = 0$$

y asi sucesivamente obtenemos:

$$c_m(\lambda_1 - \lambda_m)\cdots(\lambda_{m-2} - \lambda_m)(\lambda_{m-1} - \lambda_m)x_m = 0$$

Como $\lambda_i \neq \lambda_j \quad \forall\, i \neq j$ y $x_m \neq 0$ se tiene que:

$$c_m = 0.$$

Luego de (2.1) se tiene:

$$\sum_{i=1}^{m} c_i x_i = \sum_{i=1}^{m-1} c_i x_i = 0$$

Como $x_1, \cdots, x_{m-1}$ son linealmente independientes por hipótesis de inducción obtenemos que:

$$c_1 = c_2 = \cdots = c_{m-1} = c_m = 0$$

y de esta forma $x_1, \cdots, x_{m-1}, x_m$ son linealmente independientes. ■

Teorema 2.5. *Sea A una matriz de orden $n \times n$, entonces A es similar a una matriz diagonal D si, y sólo si, los vectores propios $x_1, \cdots .x_n$ de A son linealmente independientes.*

PRUEBA: Sean $\lambda_1, \cdots, \lambda_n$ los valores propios de A asociados a los vectores propios $x_1, \cdots .x_n$ respectivamente.

Definamos $X = (x_1, x_2, \cdots .x_n)$, entonces

$$\begin{aligned} AX = A(x_1, x_2, \cdots .x_n) &= (Ax_1, Ax_2, \cdots, Ax_n) \\ &= (\lambda_1 x_1, \lambda_2 x_2, \cdots, \lambda_n x_n) \\ &= (x_1, x_2, \cdots, x_n)\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \\ &= XD \end{aligned}$$

donde $D = diag(\lambda_1, \cdots, \lambda_n)$

Si $\{x_1, \cdots, x_n\}$ es linealmente independiente $\Longleftrightarrow$ X es no singular

$\Longleftrightarrow D = X^{-1}AX \Longleftrightarrow A$ es similar a D ∎

Corolario 2.1. *Si todos los valores propios de una matriz A de orden $n \times n$ son distintos, entonces A es similar a una matriz diagonal D.*

2.3 El Teorema de Triangularización de SCHUR y sus Aplicaciones

Teorema 2.6. *Para cualquier matriz compleja A de orden nxn existe una matriz unitaria U tal que*

$$U^*AU = T$$

es una matriz triangular superior. Los valores propios de A son los elementos de la diagonal de T.

PRUEBA: Probaremos usando inducción sobre n:

Para $n = 1$ el teorema es verdadero.

Supongamos que para $(n-1)$ el teorema es verdadero, $n \geq 2$.

Veamos que para n el teorema también es cierto:

Sea u un vector propio de A con $\|u\|_2 = 1$ asociado al valor propio λ_1. Entonces podemos elegir una matriz V de orden $n \times (n-1)$ tal que:

$$U_1 = (u, V)$$

sea unitaria.

Como $AU_1 = A(u, V) = (Au, AV) = (\lambda_1 u, AV)$ tenemos:

$$A_1 = U_1^* A U_1 = \begin{pmatrix} u^* \\ V^* \end{pmatrix} (\lambda_1 u, AV) = \left(\begin{array}{c|ccc} \lambda_1 & * & * & * \\ \hline 0 & & & \\ \vdots & & \hat{A} & \\ 0 & & & \end{array}\right) \tag{2.4}$$

donde $\hat{A} = V^*AV$ es una matriz de orden $(n-1) \times (n-1)$.

Se observa que los valores propios de A y $\hat{A}$ son los mismos, excepto λ_1. Por hipótesis existe una matriz unitaria V_1 de orden $(n-1) \times (n-1)$ tal que:

$$\hat{T} = V_1^* \hat{A} V_1$$

sea triangular superior.

Luego, definimos

$$U_2 = \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 0 & & & \\ \vdots & & V_1 & \\ 0 & & & \end{array}\right)$$

Como V_1 es unitaria, entonces U_2 es unitaria y

$$\begin{aligned} U_2^* A_1 U_2 &\overset{(2.4)}{=} U_2^* U_1^* A U_1 U_2 \\ &= U^* A U \end{aligned}$$

donde $U = U_1 U_2$.

Luego,

$$U^*AU = \left(\begin{array}{c|ccc} \lambda_1 & * & \cdots & * \\ \hline 0 & & & \\ \vdots & & V_1^* \hat{A} V_1 = \hat{T} & \\ 0 & & & \end{array}\right) = T$$

donde $\hat{T}$ es triangular superior, entonces U^*AU también es triangular superior.

Como los valores propios de una matriz triangular son los elementos de su diagonal que son justamente los valores propios de A. ∎

Teorema 2.7. *Sean* $\lambda_1, \cdots, \lambda_n$ *los valores propios de una matriz* A. *Entonces los valores propios de* A^m *son* $\lambda_1^m, \cdots, \lambda_n^m$. *En general, si*

$$p(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_k x^k$$

es un polinomio de grado k, *entonces los valores propios de* $p(A)$ *son* $p(\lambda_1). \cdots, p(\lambda_n)$.

PRUEBA: Del teorema 2.6 tenemos:

$$T = U^* A U$$

donde U es una matriz unitaria y T es una matriz triangular.

Luego

$$\begin{aligned} T^2 &= TT = (U^*AU)(U^*AU) \\ &= (U^*A)(UU^*)(AU) \qquad \text{pero} \quad UU^* = I \\ &= U^*AIAU \\ &= U^*A^2U \end{aligned}$$

$$\begin{aligned} \Longrightarrow \quad T^3 &= T^2T = (U^*A^2U)(U^*AU) \\ \Longrightarrow \quad &= U^*A^3U \end{aligned}$$

En forma sucesiva tenemos:

$$T^m = U^* A^m U \tag{2.5}$$

donde T^m es una matriz triangular cuyos elementos en la diagonal son $\lambda_1^m, \cdots, \lambda_n^m$ que también son sus valores propios. Como A^m y T^m son matrices similares, entonces ellos tienen los mismos valores propios.

Luego

$$\begin{aligned} p(A) &= c_0 I + c_1 A + c_2 A^2 + \cdots + c_k A^k &&= \sum_{i=0}^{k} c_i A^i \\ \Longrightarrow \quad U^* p(A) U &= \sum_{i=0}^{k} c_i U^* A^i U &&\overset{(2.5)}{=} \sum_{i=0}^{k} c_i T^i \end{aligned}$$

donde $c_i T^i$ es una matriz triangular (debido a que T^i lo es) cuyos elementos en la diagonal son

$$c_i\lambda_1^i, c_i\lambda_2^i, \cdots, c_i\lambda_n^i$$

$\Longrightarrow \sum_{i=0}^{k} c_i T^i$ también es una matriz triangular cuyos elementos en la diagonal son $p(\lambda_1). \cdots, p(\lambda_n)$, y como $p(A)$ y $\sum_{i=0}^{k} c_i T^i$ son matrices similares y por el teorema 2.6 tienen los mismos valores propios. ∎

Teorema 2.8. *Sea A una matriz hermitiana. Entonces*

1. *Existe una matriz unitaria U tal que $U^* AU = D$ es una matriz diagonal.*

2. *Los valores propios de A son reales.*

3. *Los vectores propios de A son ortonormales.*

PRUEBA:

1. Por el teorema 2.6 tenemos:

$$U^* AU = T \tag{2.6}$$

donde U es una matriz unitaria y T una matriz triangular superior.

Veamos que T es diagonal:

de (2.6) tenemos:

$$
\begin{aligned}
(U^*AU)^* &= T^* \\
\Longrightarrow \quad U^*A^*(U^*)^* &= T^* \\
\xrightarrow{A^*=A} \quad U^*AU &= T^* \\
T &= T^*
\end{aligned}
$$

Entonces T es una matriz Hermitiana y como

$$T^* = \left(\overline{T}\right)^T$$

entonces, se tiene que T^* es triangular inferior, por tanto T es una matriz diagonal.

2. Los valores propios de A son elementos de la diagonal de T y como T es diagonal y siendo además $T = \left(\overline{T}\right)^T = \overline{T}$, entonces los valores propios de A son reales.

3. Sea $U = (u_1, \cdots, u_n)$ donde los u_i son los vectores columnas y como

$$
\begin{aligned}
U^*AU &= T = D \qquad \text{por } (1) \\
\Longrightarrow \quad AU &= UD \\
\Longrightarrow \quad A(u_1, \cdots, u_n) &= (u_1, \cdots, u_n) diag(d_1, \cdots, d_n) \\
\Longrightarrow \quad Au_i &= d_i u_i
\end{aligned}
$$

entonces u_i son los vectores propios asociados a d_i $\forall\, i = 1, \cdots, n$.
Como U es unitaria,
entonces los vectores propios de A son ortonormales. ■

Corolario 2.2. *Si A es una matriz real simétrica, entonces existe una matriz ortogonal U tal que*

$$U^T AU = D,$$

donde D es una matriz diagonal.

PRUEBA: A es una matriz real $\Longrightarrow A = \overline{A}$,

además, A es simétrica $\Longrightarrow A = A^T$

$\Longrightarrow A = \left(\overline{A}\right)^T = A^*$

$\Longrightarrow A$ es una matriz Hermitiana real,

$\xRightarrow{teorema\ 2.8}$ existe una matriz unitaria U tal que

$$U^*AU = D$$

donde D es una matriz diagonal cuyos elementos son reales,

además, $U^* = \left(\overline{U}\right)^T = U^T$

$\Longrightarrow U^TAU = D$ ■

Teorema 2.9. *Los valores propios de una matriz A Hermitiana definida positiva son positivos. Recíprocamente, si una matriz A Hermitiana tiene todos sus valores propios positivos, entonces A debe ser definida positiva.*

PRUEBA: Sea (λ, x) un par propio de A.

$$\begin{aligned}
&\Longrightarrow && Ax = \lambda x \\
&\Longrightarrow && x^*Ax = \lambda x^*x \\
&\xRightarrow{A\ Hermitiana} && \lambda \in \mathbb{R} \\
&\xRightarrow{A\ Def.\ positiva} && x^*Ax > 0
\end{aligned}$$

Como $x \neq 0$, tenemos que:

$$x^*x = |x_1|^2 + \cdots + |x_n|^2 > 0 \quad \text{donde} \quad x = (x_1, \cdots, x_n)^T \in \mathbb{R}^n$$

Entonces,

$$\lambda = \frac{x^*Ax}{x^*x} > 0$$

Para probar el recíproco, como A es Hermitiana y usando el teorema 2.8 tenemos:

$$U^*AU = D = diag(\lambda_1, \cdots, \lambda_n) \tag{2.7}$$

Definamos para cualquier $x \in \mathbb{R}^n \backslash \{0\}$,

$$y = U^* x. \tag{2.8}$$

Entonces $y \neq 0$.

Luego

$$x^* A x \overset{(2.7)}{=} x^* U D U^* x \overset{(2.8)}{=} y^* D y = \sum_{i=1}^{n} \lambda_i |y_i|^2 > 0, \qquad \forall\, x \in \mathbb{R}^n \backslash \{0\}$$

Entonces A es definida positiva. ■

Hemos visto que una matriz arbitraria A de orden $n \times n$ es similar a una matriz diagonal si, y sólo si sus n vectores propios son linealmente independientes. Luego en particular, una matriz Hermitiana, una matriz real simétrica y una matriz que tiene distintos valores propios son similares a una matriz diagonal. Esto es, dichas matrices son **NO-DEFECTIVAS.** El siguiente teorema muestra que una matriz arbitraria es también similar a una matriz diagonal por bloques.

Teorema 2.10 (La Forma Canónica de Jordan). *Sea A una matriz de orden $n \times n$, entonces existe una matriz X no-singular tal que:*

$$X^{-1} A X = \begin{bmatrix} J_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & J_m \end{bmatrix}$$

donde

$$J_i = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 & 0 \\ 0 & 0 & \lambda_i & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_i & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda_i \end{bmatrix}$$

donde λ_i son los valores propios de A.

PRUEBA: *ver [12], [13], [14]*

2.4 Localización de los Valores Propios

En varias aplicaciones prácticas es importante conocer la región (regiones) donde se encuentra (encuentran) un (algunos) valor (valores) propio (propios) o dada alguna región saber cuantos valores se encuentran en dicha región. Existen varias formas de localizar los valores propios de una determinada matriz A de orden $n \times n$. Comenzaremos con el método de Geršgorin.

Teorema 2.11 (Primer Teorema de Geršgorin). *Sea A una matriz de orden $n \times n$. Definamos*

$$r_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|, \qquad i = 1, \cdots, n$$

Entonces cada valor propio λ de A satisface al menos una de las siguientes desigualdades:

$$|\lambda - a_{ii}| \leq r_i, \qquad i = 1, \cdots, n$$

PRUEBA: Sean λ un valor propio de A y sea x un vector propio asociado a λ,

$$\begin{aligned}
\Longrightarrow \qquad Ax &= \lambda x \\
\Longrightarrow \qquad (\lambda I - A)x &= 0 \\
\Longrightarrow \qquad (\lambda - a_{ii})x_i + \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij}x_j &= 0, \qquad i = 1, \cdots, n \\
\Longrightarrow \qquad (\lambda - a_{ii})x_i &= -\sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij}x_j, \quad i = 1, \cdots, n \qquad (2.9)
\end{aligned}$$

donde x_i es la i−ésima componente del vector x. Sea $|x_k| = \max\limits_{1\leq j\leq n}(\{|x_j|\})$. Entonces, tenemos $\dfrac{|x_j|}{|x_k|} \leq 1$ para todo $j \neq k$.
Luego de (2.9) tenemos:

$$\begin{aligned} |\lambda - a_{kk}| &\leq \sum_{\substack{j=1\\ j\neq k}}^{n} |a_{kj}|\frac{|x_j|}{|x_k|} \\ &\leq \sum_{\substack{j=1\\ j\neq k}}^{n} |a_{kj}| = r_k \end{aligned}$$

Luego $$\lambda \in \{\lambda / |\lambda - a_{kk}| \leq r_k\}$$ ■

Definición 2.5. *Los discos* $R_i = \{z/|z - a_{ii}| \leq r_i\}$, $i = 1, \cdots, n$ *son llamados* **discos de Geršgorin** *en el plano complejo.*

Ejemplo 2.1.

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 4 & 9 \\ 1 & 1 & 1 \end{pmatrix}$$

Entonces tenemos: $r_1 = \sum\limits_{\substack{j=1\\ j\neq 1}}^{3} |a_{1j}| = 5,\ \ a_{11} = 1$

$$r_2 = \sum_{\substack{j=1\\ j\neq 2}}^{3} |a_{2j}| = 12,\ a_{22} = 4$$

$$r_3 = \sum_{\substack{j=1\\ j\neq 3}}^{3} |a_{3j}| = 2,\ \ a_{33} = 1$$

Los discos de Geršgorin son:

$$R_1 = \{z/|z-1| \leq 5\}$$
$$R_2 = \{z/|z-4| \leq 12\}$$
$$R_3 = \{z/|z-1| \leq 2\}$$ □

Teorema 2.12 (Segundo Teorema de Geršgorin). *Supongamos que se tiene* $R_i = \{z/|z - a_{ii}| \leq r_i\}$, $r_i = 1, 2, \cdots, r$ *discos de Geršgorin tales que* $R_i \cap R_j = \emptyset$, $\forall\ i \neq j$, *entonces el valor propio de* A, $\lambda_i \in R_i\ \forall\ i = 1, 2, \cdots, r$.

Ejemplo 2.2. *Sea* $A = \begin{pmatrix} 1 & 0.1 & 0.2 \\ 0.2 & 4 & 0.3 \\ 0.4 & 0.5 & 8 \end{pmatrix}$

$$\begin{aligned} \textit{Entonces}\quad R_1 &= \{z/|z-1| \leq 0.3\} \\ R_2 &= \{z/|z-4| \leq 0.5\} \\ R_3 &= \{z/|z-8| \leq 0.9\} \end{aligned}$$

notar que $R_i \cap R_j = \emptyset\ \forall\ i \neq j\ i, j = 1, 2, 3$, *además los valores propios de* A *son:*

$\lambda_1 = 0.98336253767999 \in R_1$,

$\lambda_2 = 3.96709236337626 \in R_2$,

$\lambda_3 = 8.04954509894374 \in R_3$. □

Además podemos localizar los valores propios de una matriz cuadrada A de orden $n \times n$ mediante los siguientes teoremas:

Teorema 2.13. *Sea* λ *un valor propio de una matriz* A. *Entonces se tiene*

$$|\lambda| \leq \|A\|.$$

En particular $\rho(A) \leq \|A\|$

PRUEBA: Sea $x \in \mathbb{R}^n \backslash \{0\}$ un vector propio asociado a λ,

$$\begin{array}{llcl} \text{entonces} & Ax & = & \lambda x \\ \Longrightarrow & \|Ax\| & = & \|\lambda x\| = |\lambda|\|x\| \qquad (2.10) \\ \text{Como} & \|Ax\| & \leq & \|A\|\|x\| \\ \xrightarrow{(2.10)} & |\lambda|\|x\| & \leq & \|A\|\|x\| \\ \xrightarrow{\|x\|\neq 0} & |\lambda| & \leq & \|A\| \\ \Longrightarrow & \max\limits_{\lambda \text{v.p.} A}(\{|\lambda|\}) & = & \rho(A) \leq \|A\| \end{array}$$

■

Corolario 2.3. $\rho(A) \leq \|A^T\|$

PRUEBA: Como los valores propios de A^T son los mismos que los de A

$$\begin{aligned} \text{entonces} \quad \rho(A) &= \rho(A^T) \\ \xrightarrow{teo\ 2.13} \quad \rho(A) &\leq \|A^T\| \end{aligned}$$ ■

Teorema 2.14.

$$\rho(A) \leq \min\left(\left\{\max_{1\leq i\leq n}\left(\left\{\sum_{j=1}^{n}|a_{ij}|\right\}\right), \max_{1\leq j\leq n}\left(\left\{\sum_{i=1}^{n}|a_{ij}|\right\}\right)\right\}\right)$$

PRUEBA: Sabemos que $\|A\|_\infty = \max\limits_{1\leq i\leq n}\left(\left\{\sum\limits_{j=1}^{n}|a_{ij}|\right\}\right)$

$$\|A^T\|_\infty = \max_{1\leq j\leq n}\left(\left\{\sum_{i=1}^{n}|a_{ij}|\right\}\right)$$

Como $\rho(A) \leq \|A\|_\infty$ por el teorema 2.13 y

$\rho(A^T) \leq \|A^T\|_\infty$ por el corolario 2.3

$$\Longrightarrow \rho(A) \leq \min(\{\|A\|_\infty, \|A^T\|_\infty\})$$ ■

Teorema 2.15. *Sean $\lambda_1, \cdots, \lambda_n$ los valores propios de A, entonces*

$$\sum_{i=1}^{n}|\lambda_i|^2 \leq \|A\|_F^2$$

PRUEBA: Por el teorema 2.6 de triangularización existe una matriz unitaria U tal que:

$$U^*AU = T$$

donde T es una matriz triangular superior

$$\begin{aligned} \Longrightarrow \quad TT^* &= (U^*AU)(U^*AU)^* \\ &= U^*AUU^*A^*U \\ &= U^*AA^*U \end{aligned}$$

entonces AA^* es una matriz semejante unitaria a TT^* .

Además, se sabe que las matrices semejantes tienen las mismas trazas, entonces

$$tr(TT^*) = tr(AA^*) = \|A\|_F^2$$

Como T es una matriz triangular superior, entonces se tiene:

$$\begin{aligned} tr(TT^*) &= \sum_{i=1}^{n}\sum_{j=1}^{n}|t_{ij}|^2 \\ \Longrightarrow \sum_{i=1}^{n}\sum_{j=1}^{n}|t_{ij}|^2 &= \|A\|_F^2 \end{aligned}$$

Además, los elementos de la diagonal de T son los valores propios de A

$$\begin{aligned} \text{entonces} \quad \sum_{i=1}^{n}|t_{ii}|^2 &= \sum_{i=1}^{n}|\lambda_i|^2 \\ \Longrightarrow \sum_{i=1}^{n}|\lambda_i|^2 &= \sum_{i=1}^{n}|t_{ii}|^2 \leq \sum_{i=1}^{n}\sum_{j=1}^{n}|t_{ij}|^2 = \|A\|_F^2 \\ \Longrightarrow \quad & \sum_{i=1}^{n}|\lambda_i|^2 \leq \|A\|_F^2 \end{aligned}$$

■

2.5 Cálculo de Valores Propios

Describiremos brevemente algunos métodos clásicos para calcular los valores propios dominantes y los correspondientes vectores propios.

2.5.1 Método de la Potencia

Este método es usado para calcular el mayor valor propio en módulo y el correspondiente vector propio de una matriz A de orden $n \times n$.

2.5.1.1 Búsqueda del primer valor propio más grande en módulo y su correspondiente vector propio

Sean los valores propios $\lambda_1, \lambda_2, \cdots, \lambda_n$ de la matriz A.

Supongamos que los correspondientes vectores propios $v^1, v^2, \cdots, v^n$ sean linealmentes independientes. Para resolver nuestro problema tenemos dos casos:

[**CASO** 1] Entre los valores propios de A existe uno cuyo valor en módulo es el más grande, es decir: $|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$. Consideremos $y \in \mathbb{R}^n$ un vector arbitrario, entonces este vector puede ser expresado como:

$$y = \sum_{j=1}^{n} c_j v^j$$

donde c_j $(j = 1, 2, \cdots, n)$ son constantes. entonces

$$\begin{aligned} & Ay = \sum_{j=1}^{n} c_j A v^j \\ \xrightarrow{Av^j = \lambda_j v^j} \quad & Ay = \sum_{j=1}^{n} c_j \lambda_j v^j \\ \xrightarrow{Av^j = \lambda_j v^j} \quad & A^m y = \sum_{j=1}^{n} c_j \lambda_j^m v^j \end{aligned} \tag{2.11}$$

donde $m \in \mathbb{Z}$.

Sea $e_1, e_2, \cdots, e_n$ la base canónica de $\mathbb{R}^n$, entonces tenemos:

$$v^j = \sum_{i=1}^{n} x_{ij} e_i \tag{2.12}$$

sustituyendo (2.12) en (2.11), obtendremos:

$$z^m = A^m y = \sum_{j=1}^{n} c_j \lambda_j^m \sum_{i=1}^{n} x_{ij} e_i$$

cambiando el orden de las sumas tenemos:

$$z^m = \sum_{i=1}^{n} e_i \sum_{j=1}^{n} c_j x_{ij} \lambda_j^m$$

esta expresión también puede ser expresado de la siguiente forma:

$$z_i^m = \sum_{j=1}^{n} c_j x_{ij} \lambda_j^m$$

de igual forma se tiene:

$$z_i^{m+1} = \sum_{j=1}^{n} c_j x_{ij} \lambda_j^{m+1}$$

de donde obtenemos:

$$\frac{z_i^{m+1}}{z_i^m} = \frac{c_1 x_{i1} \lambda_1^{m+1} + \cdots + c_n x_{in} \lambda_n^{m+1}}{c_1 x_{i1} \lambda_1^m + \cdots + c_n x_{in} \lambda_n^m} \tag{2.13}$$

Supongamos que $c_1 \neq 0$ y $x_{i1} \neq 0$. (Esto se logra escogiendo en forma apropiada el vector inicial y).

Transformando la última expresión se obtiene:

$$\frac{z_i^{m+1}}{z_i^m} = \lambda_1 \frac{1 + \dfrac{c_2 x_{i2}}{c_1 x_{i1}} \left(\dfrac{\lambda_2}{\lambda_1}\right)^{m+1} + \cdots + \dfrac{c_n x_{in}}{c_1 x_{i1}} \left(\dfrac{\lambda_n}{\lambda_1}\right)^{m+1}}{1 + \dfrac{c_2 x_{i2}}{c_1 x_{i1}} \left(\dfrac{\lambda_2}{\lambda_1}\right)^{m} + \cdots + \dfrac{c_n x_{in}}{c_1 x_{i1}} \left(\dfrac{\lambda_n}{\lambda_1}\right)^{m}}$$

De aquí tomando límite cuando $m \to \infty$ y teniendo en cuenta que los valores propios satisface $|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$ se tiene:

$$\lim_{m \to \infty} \frac{z_i^{m+1}}{z_i^m} = \lambda_1 \tag{2.14}$$

(esto es debido a que $\lim\limits_{m\to\infty} \left(\dfrac{\lambda_i}{\lambda_1}\right)^m = 0$ para $i > 1$), o aproximadamente,

$$\lambda_1 \approx \frac{z_i^{m+1}}{z_i^m} \qquad (i = 1, 2, \cdots, n)$$

en forma exacta tenemos:

$$\lambda_1 = \frac{z_i^{m+1}}{z_i^m} + O\left(\frac{\lambda_2}{\lambda_1}\right)^m$$

Nota para acelerar la convergencia del proceso de iteración de (2.14) es conveniente a veces formar la sucesión de matrices:

$$\begin{aligned}
A^2 &= AA \\
A^4 &= A^2A^2 \\
A^8 &= A^4A^4 \\
\cdots &\cdots \cdots \\
A^{2^k} &= A^{2^{k-1}}A^{2^{k-1}}
\end{aligned}$$

de la cual se tiene:

$$z^m = A^m y$$

donde $m = 2^k$. Supongamos por consiguiente, como es usual,

$$\lambda_1 \approx \frac{z_i^{m+1}}{z_i^m} \qquad (i = 1, 2, \cdots, n)$$

El vector $z^m = A^m y$ es aproximadamente el vector característico de la matriz A correspondiente al valor propio λ_1.

En efecto, de (2.11) tenemos:

$$A^m y = c_1 \lambda_1^m v^1 + \sum_{j=2}^{n} c_j \lambda_j^m v^j$$

donde v^j, $j = 1, 2, \cdots, n$ son los vectores característicos de la matriz A.

De esta forma se tiene:

$$A^m y = c_1 \lambda_1^m \left\{ v^1 + \sum_{j=1}^{n} \frac{c_j}{c_1} \left(\frac{\lambda_j}{\lambda_1}\right)^m v^j \right\}$$

como $\left(\frac{\lambda_j}{\lambda_1}\right)^m \to 0$ cuando $m \to \infty$ para todo $j > 1$, entonces para m suficientemente grande tendremos

$$A^m y \approx c_1 \lambda_1^m v^1.$$

Practicamente $A^m y$ es paralelo al vector v^1.

Ejemplo 2.3. *Hallar el valor propio más grande de la matriz:*

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

y su correspondiente vector propio.

Solucion: *Elijamos el vector inicial* $y = (1, 1, 1)^T$, *luego haciendo los cálculos hasta la iteración* 10 *tenemos:* $A^9 y = z^9 = \begin{bmatrix} 905238 \\ 417987 \\ 121248 \end{bmatrix}$ *y* $A^{10} y = z^{10} = \begin{bmatrix} 4038939 \\ 1862460 \\ 539235 \end{bmatrix}$ *y de esta forma se tendrá:*

$$\frac{z_1^{10}}{z_1^9} = \frac{4038939}{905238} = 4.4617426577$$
$$\frac{z_2^{10}}{z_2^9} = \frac{1862460}{417987} = 4.4557845100$$
$$\frac{z_3^{10}}{z_3^9} = \frac{539235}{121248} = 4.4473723278$$

En consecuencia, podemos tener aproximadamente,

$$\lambda_1 = \frac{1}{3}(4.4617426577 + 4.4557845100 + 4.4473723278) \approx 4.4549664985$$

Para el primer vector característico de la matriz A podemos elegir:

$$A^{10} y = \begin{bmatrix} 4038939 \\ 1862460 \\ 539235 \end{bmatrix}$$

Normalizándolo, obtenemos

$$v^1 = \begin{bmatrix} 0.9015003776 \\ 0.4157053110 \\ 0.1203584793 \end{bmatrix}$$

Así, hemos obtenido en forma aproximada el mayor valor propio, en módulo, de la matriz A y su correspondiente vector propio.

[**CASO** 2] Multiplicidad de valor propio más grande, en módulo, de la matriz A. Sea s la multiplicidad del valor propio más grande en módulo, es decir,

$$\lambda_1 = \lambda_2 = \cdots = \lambda_s$$

para algún $s > 1, \quad s \in \mathbb{N}$ y

$$|\lambda_1| > |\lambda_k| \quad \text{para } k \in \{s+1, s+2, \cdots, n\}$$

de la relación (2.13) tenemos:

$$\begin{aligned} \frac{z_i^{m+1}}{z_i^m} &= \frac{c_1 x_{i1}\lambda_1^{m+1} + \cdots + c_s x_{is}\lambda_1^{m+1} + c_{s+1}x_{i,s+1}\lambda_{s+1}^{m+1} + \cdots + c_n x_{in}\lambda_n^{m+1}}{c_1 x_{i1}\lambda_1^m + \cdots + c_s x_{is}\lambda_1^m + \cdots + c_{s+1}x_{i,s+1}\lambda_{s+1}^m + \cdots + c_n x_{in}\lambda_n^m} \\ &= \lambda_1 \frac{c_1 x_{i1} + \cdots + c_s x_{is} + c_{s+1}x_{i,s+1}\left(\dfrac{\lambda_{s+1}}{\lambda_1}\right)^{m+1} + \cdots + c_n x_{in}\left(\dfrac{\lambda_n}{\lambda_1}\right)^{m+1}}{c_1 x_{i1} + \cdots + c_s x_{is} + c_{s+1}x_{i,s+1}\left(\dfrac{\lambda_{s+1}}{\lambda_1}\right)^{m} + \cdots + c_n x_{in}\left(\dfrac{\lambda_n}{\lambda_1}\right)^{m}} \end{aligned}$$

de donde, si $c_1 x_{i1} + \cdots + c_s x_{is} \neq 0$ y teniendo en cuenta que:

$$\left(\frac{\lambda_k}{\lambda_1}\right)^m \to 0 \text{ para } m \to \infty \text{ y } k > s$$

tenemos

$$\lim_{m\to\infty} \frac{z_i^{m+1}}{z_i^m} = \lambda_1, \qquad (i = 1, 2, \ldots, n)$$

o, en forma exacta se tiene:

$$\lambda_1 = \frac{z_i^{m+1}}{z_i^m} + O\left(\left(\frac{\lambda_{s+1}}{\lambda_1}\right)^m\right)$$

Así, el método anterior para el cálculo de λ_1 también es aplicable en este caso y $z^m = A^m y$ es uno de los vectores propios asociados de la matriz A.

Otra forma de determinar aproximadamente el valor propio dominante de una matriz A, está en el siguiente algoritmo.

Denotemos $\max(\{x\}) = \max\limits_{1\leq i\leq n}(\{|x_i|\})$, donde $x = (x_1, \cdots, x_n)^T \in \mathbb{R}^n$.

Asumiremos que A es diagonalizable.

ALGORITMO (MÉTODO DE LA POTENCIA):

PASO 1 Elijamos $x^0 \in \mathbb{R}^n \backslash \{0\}$

PASO 2 Para $k = 1, 2, \cdots$ hasta que converja

$$\hat{x}^k = Ax^{k-1}$$
$$x^k = \frac{\hat{x}^k}{\max(\{\hat{x}^k\})}$$

END

Teorema 2.16. $\max(\{\hat{x}^k\}) \longrightarrow \lambda_1$

$x^k \longrightarrow w^1$ *múltiplo de* v^1 *cuando* $k \to \infty$

PRUEBA: Aplicando sucesivamente el algoritmo tenemos:

$k = 1:$ $\quad \hat{x}^1 = Ax^0 \qquad x^1 = \dfrac{Ax^0}{max(\{Ax^0\})}$

$k = 2:$ $\quad \hat{x}^2 = Ax^1 = \dfrac{A^2x^0}{max(\{Ax^0\})} \qquad x^2 = \dfrac{A^2x^0}{max(\{Ax^0\})max(\{\frac{A^2x^0}{max(\{Ax^0\})}\})}$

$$= \frac{A^2x^0}{max(\{A^2x^0\})}$$

$k = 3:$ $\quad \hat{x}^3 = Ax^2 = \dfrac{A^3x^0}{max(\{A^2x^0\})} \qquad x^3 = \dfrac{A^3x^0}{max(\{Ax^0\})max(\{\frac{A^3x^0}{max(\{Ax^0\})}\})}$

$$= \frac{A^3x^0}{max(\{A^3x^0\})}$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$\hat{x}^k = \frac{A^kx_0}{\max(\{A^{k-1}x_0\})} \qquad x^k = \frac{A^kx_0}{\max(\{A^kx_0\})}$$

Como A es diagonalizable, entonces los vectores propios v^i, $i = 1, \cdots, n$ asociados con λ_i, $i = 1, \cdots, n$ pueden ser elegidos linealmente independientes. Luego x^0 puede ser escrito de la forma:

$$\begin{aligned}
x^0 &= \alpha_1 v^1 + \alpha_2 v^2 + \cdots + \alpha_n v^n, \quad \alpha_1 \neq 0 \\
\Longrightarrow \quad A^k x^0 &= A^k(\alpha_1 v^1 + \alpha_2 v^2 + \cdots + \alpha_n v^n) \\
&\overset{Av^i = \lambda_i v^i}{=} \alpha_1 \lambda_1^k v^1 + \alpha_2 \lambda_2^k v^2 + \cdots + \alpha_n \lambda_n^k v^n \\
\Longrightarrow \quad A^k x^0 &= \lambda_1^k \left(\alpha_1 v^1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k v^2 + \cdots + \alpha_n \left(\frac{\lambda_n}{\lambda_1} \right)^k v^n \right)
\end{aligned}$$

Como $|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$. Entonces $\lim_{k\to\infty} \left(\frac{\lambda_i}{\lambda_1} \right)^k = 0$, $i = 2, \cdots, n$.
Luego:

$$\begin{aligned}
x^k = \frac{A^k x^0}{max(\{A^k x^0\})} &= \frac{\lambda_1^k(\alpha_1 v^1 + \alpha_2 (\frac{\lambda_2}{\lambda_1})^k v^2 + \cdots + \alpha_n (\frac{\lambda_n}{\lambda_1})^k v^n)}{max(\{\lambda_1^k(\alpha_1 v^1 + \alpha_2 (\frac{\lambda_2}{\lambda_1})^k v^2 + \cdots + \alpha_n (\frac{\lambda_n}{\lambda_1})^k v^n\})} \\
&= \frac{\alpha_1 v^1 + \alpha_2 (\frac{\lambda_2}{\lambda_1})^k v^2 + \cdots + \alpha_n (\frac{\lambda_n}{\lambda_1})^k v^n}{max(\{\alpha_1 v^1 + \alpha_2 (\frac{\lambda_2}{\lambda_1})^k v^2 + \cdots + \alpha_n (\frac{\lambda_n}{\lambda_1})^k v^n\})} \to cv^1
\end{aligned}$$

donde c es una constante y

$$\begin{aligned}
max(\{\hat{x}^k\}) = max\left(\left\{ \frac{A^k x^0}{\max(\{A^{k-1}x^0\})} \right\} \right) &= \frac{max(\{A^k x^0\})}{max(\{A^{k-1}x^0\})} = \\
&= \frac{\lambda_1^k max(\{\alpha_1 v^1 + \alpha_2 (\frac{\lambda_2}{\lambda_1})^k v^2 + \cdots + \alpha_n (\frac{\lambda_n}{\lambda_1})^k v^n\})}{\lambda_1^{k-1} max(\{\alpha_1 v^1 + \alpha_2 (\frac{\lambda_2}{\lambda_1})^{k-1} v^2 + \cdots + \alpha_n (\frac{\lambda_n}{\lambda_1})^{k-1} v^n\})} \\
&= \frac{\lambda_1 max(\{\alpha_1 v^1 + \alpha_2 (\frac{\lambda_2}{\lambda_1})^k v^2 + \cdots + \alpha_n (\frac{\lambda_n}{\lambda_1})^k v^n\})}{max(\{\alpha_1 v^1 + \alpha_2 (\frac{\lambda_2}{\lambda_1})^{k-1} v^2 + \cdots + \alpha_n (\frac{\lambda_n}{\lambda_1})^{k-1} v^n\})} \to \lambda_1.
\end{aligned}$$

■

Ejemplo 2.4.

$$Sea\ A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix}$$

y consideremos $x^0 = (1,1,1)^T \in \mathbb{R}^3 \backslash \{0\}$

$$k{=}1\colon\ \hat{x}^1 = Ax^0 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 9 \\ 12 \end{pmatrix}$$

$$\max(\{\hat{x}^1\}) = 12$$

$$x^1 = \frac{\hat{x}^1}{\max(\{\hat{x}^1\})} = \begin{pmatrix} 1/2 \\ 3/4 \\ 1 \end{pmatrix} = \begin{pmatrix} 0.50 \\ 0.75 \\ 1.00 \end{pmatrix}$$

$$k{=}2\colon\ \hat{x}^2 = Ax^1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix}\begin{pmatrix} 1/2 \\ 3/4 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 29/4 \\ 19/2 \end{pmatrix} = \begin{pmatrix} 5.00 \\ 7.25 \\ 9.50 \end{pmatrix}$$

$$\max(\{\hat{x}^2\}) = 19/2 = 9.50$$

$$x^2 = \frac{\hat{x}^2}{\max(\{\hat{x}^2\})} = \begin{pmatrix} 10/19 \\ 29/38 \\ 1 \end{pmatrix} \approx \begin{pmatrix} 0.5263 \\ 0.7632 \\ 1.0000 \end{pmatrix}$$

$$k{=}3\colon\ \hat{x}^3 = Ax^2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix}\begin{pmatrix} 10/19 \\ 29/38 \\ 1 \end{pmatrix} = \begin{pmatrix} 96/19 \\ 279/38 \\ 183/19 \end{pmatrix} \approx \begin{pmatrix} 5.0526 \\ 7.4321 \\ 9.6316 \end{pmatrix}$$

$$\max(\{\hat{x}^3\}) = 183/19 \approx 9.6316$$

$$x^3 = \frac{\hat{x}^3}{\max(\{\hat{x}^3\})} = \begin{pmatrix} 32/61 \\ 93/122 \\ 1 \end{pmatrix} \approx \begin{pmatrix} 0.5246 \\ 0.7623 \\ 1.0000 \end{pmatrix}$$

$$k{=}4\text{:}\ \hat{x}^4 = Ax^3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix} \begin{pmatrix} 32/61 \\ 93/122 \\ 1 \end{pmatrix} = \begin{pmatrix} 308/61 \\ 895/122 \\ 587/61 \end{pmatrix} \approx$$

$$\approx \begin{pmatrix} 5.049180 \\ 7.336066 \\ 9.622951 \end{pmatrix}$$

$$\max(\{\hat{x}^4\}) = 587/61 \approx 9.622951$$

$$x^4 = \frac{\hat{x}^4}{\max(\{\hat{x}^4\})} = \begin{pmatrix} 308/587 \\ 895/1174 \\ 1 \end{pmatrix} \approx \begin{pmatrix} 0.52470187 \\ 0.76235094 \\ 1.00000000 \end{pmatrix}$$

$$k{=}5\text{:}\ \hat{x}^5 = Ax^4 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix} \begin{pmatrix} 32/61 \\ 93/122 \\ 1 \end{pmatrix} = \begin{pmatrix} 308/587 \\ 895/1174 \\ 1 \end{pmatrix} \approx$$

$$\approx \begin{pmatrix} 5.0494037479 \\ 7.3364565588 \\ 9.6235093697 \end{pmatrix}$$

$$\max(\{\hat{x}^5\}) = 5649/587 \approx 9.6235093697$$

$$x^5 = \frac{\hat{x}^5}{\max(\{\hat{x}^5\})} = \begin{pmatrix} 988/1883 \\ 2871/3766 \\ 1 \end{pmatrix} \approx \begin{pmatrix} 0.524694636219 \\ 0.762347318109 \\ 1.000000000000 \end{pmatrix}$$

Normalizandolo tenemos: $\dfrac{x^5}{\|x^5\|} = \begin{pmatrix} 0.3850895499 \\ 0.5595101709 \\ 0.7339307919 \end{pmatrix}.$ $\square$

2.5.1.2 Convergencia del método de la Potencia

La razón de convergencia del método de potencia está determinado por $\frac{\lambda_2}{\lambda_1}$ como veremos enseguida: de la prueba del teorema 2.16 (método de la potencia) tenemos:

$$\begin{aligned}
\|x^k - \alpha_1 v^1\| &= \left\| \alpha_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k v^2 + \cdots + \alpha_n \left(\frac{\lambda_n}{\lambda_1}\right)^k v^n \right\| \\
&\leq |\alpha_2| \left|\frac{\lambda_2}{\lambda_1}\right|^k \|v^2\| + \cdots + |\alpha_n| \left|\frac{\lambda_n}{\lambda_1}\right|^k \|v^n\| \\
&\leq \left|\frac{\lambda_2}{\lambda_1}\right|^k (|\alpha_2|\|v^2\| + \cdots + |\alpha_n|\|v^n\|)
\end{aligned}$$

(debido a que $\frac{\lambda_i}{\lambda_1} \leq \frac{\lambda_2}{\lambda_1}$, $i = 3, 4, \cdots, n$.)

Luego tenemos:

$$\|x^k - \alpha_1 v^1\| \leq \alpha \left|\frac{\lambda_2}{\lambda_1}\right|^k$$

donde $\alpha = |\alpha_2|\|v^2\| + \cdots + |\alpha_n|\|v^n\|$.

Notar que esta prueba es debido a que x^k se apróxima a $\alpha_1 v^1$ dependiendo de la rapidez como $\left|\frac{\lambda_2}{\lambda_1}\right|^k$ va hacia cero. Si λ_2 está cercano a λ_1, entonces la convergencia se hace lenta. ■

2.5.1.3 Método de los productos escalares para hallar el primer valor propio más grande en módulo

Este método es usado para calcular el valor propio dominante y el correspondiente vector propio de una matriz. Este nombre es debido a que está basado en la construcción implícita de la matriz A real.

Sean $\lambda_1, \lambda_2, \cdots, \lambda_n$ los valores propios de A tal que:

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

es decir, λ_1 es el valor propio dominante de A.

Ahora elijamos $x^0 \in \mathbb{R}^n \backslash \{0\}$ cualquiera,
calculemos las iteraciones:

$$x^k = Ax^{k-1}, \quad k = 1, 2, \cdots \tag{2.15}$$

Formemos también las iteraciones

$$y^k = A^T y^{k-1} \quad k = 1, 2, \cdots , \tag{2.16}$$

donde $y^0 = x^0$
Elijamos dos bases $\{\bar{x}^j\}$, $\{\bar{y}^j\}$ tales que

$$\langle \bar{x}^i, \bar{y}^j \rangle = \delta_{ij}$$

donde $A\bar{x}^i = \lambda_i \bar{x}^i$, $A^T \bar{y}^j = \lambda_j^* \bar{y}^j$ $(i, j = 1, 2, \cdots, n)$.
Luego tenemos que existen $a_1, a_2, \cdots, a_n \in \mathbb{R}$, $b_1, b_2, \cdots, b_n \in \mathbb{R}$ tales que:

$$x^0 = \sum_{i=1}^{n} a_i \bar{x}^i \quad \text{y} \quad y^0 = \sum_{j=1}^{n} b_j \bar{y}^j$$

Por (2.15) tenemos:

$$x^k = A^k x^0 = \sum_{i=1}^{n} a_i \lambda_i^k \bar{x}^i, \quad k = 1, 2, \cdots$$

y de (2.16) se tiene:

$$y^k = (A^T)^k y^0 = \sum_{j=1}^{n} b_j (\lambda_j^*)^k \bar{y}^j, \quad k = 1, 2, \cdots$$

Entonces de las dos últimas expresiones se obtiene:

$$\langle x^k, y^k\rangle = \langle A^k x^0, (A^T)^k y^0\rangle = \langle x^0, (A^T)^{2k} y^0\rangle = \left\langle \sum_{i=1}^{n} a_i \bar{x}^i, \sum_{j=1}^{n} b_j (\lambda_j^*)^{2k} \bar{y}^j \right\rangle$$

Debido a la ortonormalización se tiene:

$$\langle x^k, y^k\rangle = \sum_{i=1}^{n} a_i b_i (\lambda_i^*)^{2k}$$

Análogamente

$$\langle x^{k-1}, y^k\rangle = \sum_{i=1}^{n} a_i b_i (\lambda_i^*)^{2k-1}$$

Para $a_1 b_1 \neq 0$ tenemos:

$$\frac{\langle x^k, \bar{x}^k\rangle}{\langle x^{k-1}, \bar{x}^k\rangle} = \frac{a_1 b_1 (\lambda_1^*)^{2k} + a_2 b_2 (\lambda_2^*)^{2k} + \cdots + a_n b_n (\lambda_n^*)^{2k}}{a_1 b_1 (\lambda_1^*)^{2k-1} + a_2 b_2 (\lambda_2^*)^{2k-1} + \cdots + a_n b_n (\lambda_n^*)^{2k-1}}$$

$$\frac{\langle x^k, \bar{x}^k\rangle}{\langle x^{k-1}, \bar{x}^k\rangle} \to \lambda_1 + O\left(\left(\frac{\lambda_2}{\lambda_1}\right)^{2k}\right)$$

De este modo se tiene:

$$\lambda_1 \approx \frac{\langle x^k, \bar{x}^k\rangle}{\langle x^{k-1}, \bar{x}^k\rangle} = \frac{\langle A^k x^0, (A^T)^k x^0\rangle}{\langle A^{k-1} x^0, (A^T)^k x^0\rangle}$$

Notar que si A es simétrica entonces la relación anterior se convierte en:

$$\lambda_1 \approx \frac{\langle A^k x^0, A^k x^0\rangle}{\langle A^{k-1} x^0, A^k x^0\rangle}$$

Ejemplo 2.5. *Sea*

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

utilizar el método anterior para hallar el valor propio más grande en módulo de A.

Solución: *Notar que A es una matriz simétrica, entonces basta formar solamente* *las iteraciones* $A^k x^0 (k = 1, 2 \cdots)$ *escojamos* $x^0 = (1,1,1)^T$ *como vector inicial, al final de las primeras seis iteraciones tenemos:*

$$A^5 x^0 = \begin{bmatrix} 2268 \\ 1089 \\ 333 \end{bmatrix} \quad y \quad A^6 x^0 = \begin{bmatrix} 10161 \\ 4779 \\ 1422 \end{bmatrix}$$

por consiguiente se obtiene:

$$\langle A^5 x^0, A^6 x^0 \rangle = 28723005 \quad y \quad \langle A^6 x^0, A^6 x^0 \rangle = 128106846$$

y por tanto

$$\lambda_1 \approx \frac{\langle A^6 x^0, A^6 x^0 \rangle}{\langle A^5 x^0, A^6 x^0 \rangle} = \frac{128106846}{28723005} \approx 4.460078115086. \qquad \square$$

2.5.1.4 Búsqueda del segundo valor propio más grande en módulo y su correspondiente vector propio

Supongamos que tenemos los valores propios $\lambda_i \; i = 1, 2, \cdots, n$ de una matriz A tales que:

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

En este caso podemos usar el método descrito en la sección (2.5.1) para aproximar el segundo valor propio (λ_2) de A y su correspondiente vector propio v^2.
De la relación (2.11) de la subsection (2.5.1.1) tenemos:

$$A^m y = c_1 \lambda_1^m v^1 + c_2 \lambda_2^m v^2 + \cdots + c_n \lambda_n^m v^n \tag{2.17}$$

y

$$A^{m+1} y = c_1 \lambda_1^{m+1} v^1 + c_2 \lambda_2^{m+1} v^2 + \cdots + c_n \lambda_n^{m+1} v^n \tag{2.18}$$

Luego de (2.18) $-\ \lambda_1$(2.17) tenemos:

$$A^{m+1}y - \lambda_1 A^m y = c_2\lambda_2^m(\lambda_2 - \lambda_1)v^2 + \cdots + c_n\lambda_n^m(\lambda_n - \lambda_1)v^n \tag{2.19}$$

Denotemos

$$\Delta_\lambda A^m y = A^{m+1}y - \lambda A^m y \tag{2.20}$$

Si $c_2 \neq 0$, entonces el primer término del segundo miembro de la relación (2.19) es su primer término principal cuando $m \to \infty$, obteniéndose

$$\Delta_{\lambda_1} A^m y \approx c_2\lambda_2^m(\lambda_2 - \lambda_1)v^2 \tag{2.21}$$

de donde

$$\Delta_{\lambda_1} A^{m-1} y \approx c_2\lambda_2^{m-1}(\lambda_2 - \lambda_1)v^2 \tag{2.22}$$

Como $A^m y = z^m = (z_1^m, z_2^m, \cdots, z_n^m)^T$ y de (2.20) y (2.21) tenemos:

$$\lambda_2 \approx \frac{\Delta_{\lambda_1} z_i^m}{\Delta_{\lambda_1} z_i^{m-1}} = \frac{z_i^{m+1} - \lambda_1 z_i^m}{z_i^m - \lambda_1 z_i^{m-1}} \qquad (i = 1, 2, \cdots, n) \tag{2.23}$$

Utilizando la relación (2.23), podemos aproximar el segundo valor propio. Notar que en la práctica resulta a veces mejor (debido a la pérdida de exactitud al restar números casi iguales) tomar un número de iteraciones k más pequeño para determinar λ_1; es decir, es aconsejable establecer

$$\lambda_2 \approx \frac{\Delta_{\lambda_1} z_i^m}{\Delta_{\lambda_1} z_i^{m-1}} = \frac{z_i^{k+1} - \lambda_1 z_i^k}{z_i^k - \lambda_1 z_i^{k-1}} \qquad (k < m) \tag{2.24}$$

donde k es el menor número para los cuales λ_2 comienza a dominar los valores propios subsiguientes.

Y para el vector propio asociado a λ_2 tenemos de (2.21)

$$v^2 \approx \Delta_{\lambda_1} z^k. \qquad (2.25)$$

Ejemplo 2.6. *Determinar el segundo valor propio más grande en módulo y su correspondiente vector propio de la matriz*

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

Solución: *Aplicando el método descrito en el ejercicio (2.5) con $k = 8$ y eligiendo $y = (1,1,1)^T$ tenemos:*

$$A^7y = \begin{bmatrix} 45433 \\ 21141 \\ 6201 \end{bmatrix} \quad A^8y = \begin{bmatrix} 202833 \\ 93906 \\ 27342 \end{bmatrix} \quad A^9y = \begin{bmatrix} 905238 \\ 417987 \\ 121248 \end{bmatrix}$$

formando las diferencias utilizando la relación

$$\Delta_{\lambda_1} z_i^j = z_i^{j+1} - \lambda_1 z_i^j \qquad (j = 1,2,3)$$

donde $z^j = A^j y$.

Los valores para el valor propio mayor en módulo λ_1 ($\lambda_1 = 4.462$, $\lambda_1 = 4.456$, $\lambda_1 = 4.447$) aproximadamente hallado en el ejercicio (2.3) y que son utilizados en la siguiente tabla:

CÁLCULO DEL SEGUNDO VALOR PROPIO

A^8y	$\lambda_1 A^7y$	$\Delta_{\lambda_1} A^7y$	A^9y	$\lambda_1 A^8y$	$\Delta_{\lambda_1} A^8y$
202833	**202722**	**111**	**905238**	**905041**	**197**
93906	**94204**	**-298**	**417987**	**418445**	**-458**
27342	**27576**	**-234**	**121248**	**121590**	**-342**

De aquí obtenemos:

$$\frac{\Delta_{\lambda_1} z_1^8}{\Delta_{\lambda_1} z_1^7} = \frac{197}{111} \approx 1.774774775,$$
$$\frac{\Delta_{\lambda_1} z_2^8}{\Delta_{\lambda_1} z_2^7} = \frac{-458}{-298} \approx 1.536912752,$$
$$\frac{\Delta_{\lambda_1} z_3^8}{\Delta_{\lambda_1} z_3^7} = \frac{-342}{-234} \approx 1.461538462$$

y de esta forma tenemos aproximadamente:

$$\lambda_2 \approx \frac{1}{3}(1.774774775 + 1.536912752 + 1.461538462) \approx 1.591075329$$

y para el segundo vector propio de A tenemos:

$$\Delta_{\lambda_1} A^8 y = \begin{bmatrix} 197 \\ -458 \\ -342 \end{bmatrix}.$$

Normalizando este vector, obtenemos:

$$v^2 \approx \begin{bmatrix} 0.3258371171 \\ -0.7575299473 \\ -0.5656664672 \end{bmatrix}.$$

El tercer valor propio de A la encontramos de la siguiente forma:

$$\lambda_1 + \lambda_2 + \lambda_3 = tr(A) = 4 + 2 + 1 = 7$$

de donde:

$$\lambda_3 \approx 0.953958173$$

y el tercer vector propio v^3 correspondiente lo encontramos usando las condiciones de ortogonalidad, es decir,

$$v^{1^T} x^3 = 0$$
$$v^{2^T} x^3 = 0$$

finalmente $v^3 = \dfrac{x^3}{\|x^3\|}$. $\square$

2.5.1.5 Método del Agotamiento

Supongamos que la matriz A es real y tiene valores propios $\lambda_1, \lambda_2, \cdots, \lambda_n$ tales que:

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

Además, consideremos la matriz

$$A_1 = A - \lambda_1 X_1 X_1^T \tag{2.26}$$

donde λ_1 es el primer valor propio de A y X_1 es el correspondiente vector propio normalizado, donde

$$X_1 = \begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{bmatrix}.$$

Proposición 2.1. *Los vectores propios* X_j $(j = 1, 2, \cdots, n)$ *ortonormalizados de la matriz* A *son también los vectores propios de* A_1*, conservando los valores propios correspondientes, a excepción* λ_1 *que toma el valor de cero.*

PRUEBA: De la relación (2.26) tenemos:

$$A_1 X_1 = AX_1 - \lambda_1(X_1 X_1^T)X_1 = \lambda_1 X_1 - \lambda_1 X_1(X_1^T X_1) = \lambda_1 X_1 - \lambda_1 X_1 = 0$$

debido a que $X_1^T X_1 = 1$, y de donde se tiene que:

$$A_1 X_1 = 0X_1$$

y, por tanto, cero es el valor propio de la matriz A_1 y X_1 es su correspondiente vector propio.

Como los vectores están normalizados, entonces tenemos:

$$X_j^T X_1 = 0 \qquad j = 2, 3, \cdots, n$$

usando la relación (2.26), obtenemos

$$A_1 X_j = AX_j - \lambda_1 (X_1 X_1^T) X_j = \lambda_j X_j - \lambda_1 X_1 (X_1^T X_j) = \lambda_j X_j \qquad j = 2, 3, \cdots, n$$

y de esta manera hacemos uso de los métodos anteriores para determinar el valor propio λ_2 de A que viene a ser el mayor valor propio en módulo de A_1 y después hallamos su correspondiente vector propio X_2 que justamente es v^2 el vector propio asociado a λ_2 de la matriz A.

Volvemos a usar este método para calcular el siguiente valor propio λ_3 y su correspondiente vector propio de A y así sucesivamente hasta obtener todos los valores propios y sus correspondiente vectores propios de A. A este método se le conoce como el **Método de Agotamiento**.

2.5.2 Valores y Vectores Propios de una Matriz Simétrica Definida Positiva

Consideremos A una matriz simétrica definida positiva, entonces sabemos que:

1. los valores propios $\lambda_1, \lambda_2, \cdots, \lambda_n$ son reales y positivos.

2. los vectores propios $x^j = (x_1^j, x_2^j, \cdots, x_n^j)^T$ (donde $j = 1, 2, \cdots, n$) correspondientes a $\lambda_1, \lambda_2, \cdots, \lambda_n$ satisfacen la condición de ortogonalidad.

De la ecuación característica $det(A - I\lambda) = 0$ sabemos que:

$$\begin{aligned}
(a_{11} - \lambda_1)x_1^1 + a_{12}x_2^1 + \cdots + a_{1n}x_n^1 &= 0 \\
a_{21}x_1^1 + (a_{22} - \lambda_1)x_2^1 + \cdots + a_{2n}x_n^1 &= 0 \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\
a_{n1}x_1^1 + a_{n2}x_2^1 + \cdots + (a_{nn} - \lambda_1)x_n^1 &= 0
\end{aligned}$$

es equivalentemente a:

$$\begin{aligned}
x_1^1 &= \frac{1}{\lambda_1}(a_{11}x_1^1 + a_{12}x_2^1 + \cdots + a_{1n}x_n^1) \\
x_2^1 &= \frac{1}{\lambda_1}(a_{21}x_1^1 + a_{22}x_2^1 + \cdots + a_{2n}x_n^1) \\
\cdots & \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
x_{n-1}^1 &= \frac{1}{\lambda_1}(a_{n-1,1}x_1^1 + a_{n-1,2}x_2^1 + \cdots + a_{n-1,n}x_n^1) \\
\lambda_1 &= \frac{1}{x_n^1}(a_{n1}x_1^1 + a_{n2}x_2^1 + \cdots + a_{nn}x_n^1)
\end{aligned} \tag{2.27}$$

esto nos permite hallar el vector propio x^1 asociado a λ_1 de la matriz A.
Como las componentes de los vectores propios están determinados por el factor de proporcionalidad, una de ellas es arbitrario; por ejemplo, excepto para un caso especial, podemos hacer $x_n^1 = 1$. Luego el sistema (2.27) puede resolverse por el método iterativo eligiendo valores iniciales adecuados $x_i^{1,0}$ y λ_1^0 y considerando

$$\begin{aligned}
x_i^{1,k+1} &= \frac{1}{\lambda_1^k}\left(\sum_{j=1}^{n-1} a_{ij}x_j^{1,k} + a_{in}\right) & i &= 1, 2, \cdots, n-1 \\
\lambda_1^{k+1} &= \sum_{j=1}^{n-1} a_{nj}x_j^{1,k+1} + a_{nn} & k &= 0, 1, 2, \cdots
\end{aligned}$$

De este modo, hallamos el primer valor propio de A en forma aproximada

$$\lambda_1 \approx \lambda_1^k$$

y el primer vector propio asociado

$$x^1 \approx \begin{bmatrix} x_1^{1,k} \\ x_2^{1,k} \\ \vdots \\ x_{n-1}^{1,k} \\ 1 \end{bmatrix}.$$

Para hallar el valor propio λ_2 de la ecuación caracterítica y el segundo vector propio x^2, escribamos el sistema apropiado de ecuaciones:

$$\lambda_2 x_i^2 = \sum_{j=1}^{n} a_{ij} x_j^2 \qquad i = 1, 2, \cdots, n \tag{2.28}$$

Eliminamos de la relación de ortogonalidad

$$\sum_{j=1}^{n} x_j^1 x_j^2 = 0 \tag{2.29}$$

unas de las incógnitas x_j^2, digamos, x_n^2, entonces el sistema (2.28) es equivalente a:

$$\begin{aligned} x_i^2 &= \frac{1}{\lambda_2} \sum_{j=1}^{n-1} a_{ij}^2 x_j^2 \qquad i = 1, 2, \cdots, n-2 \\ \lambda_2 &= \frac{1}{x_{n-1}^2} \sum_{j=1}^{n-1} a_{n-1,j}^2 x_j^2 \end{aligned} \tag{2.30}$$

Estableciendo $x_{n-1}^2 = 1$ resolvemos el sistema (2.30) y de esta forma hallamos λ_2, continuando este proceso podemos hallar los valores propios λ^j y sus correspondientes vectores propios x^j de la matriz A para $j = 3, \cdots, n$.

Ejemplo 2.7. *Sea la matriz* $A = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 5 & 1 \\ 2 & 1 & 6 \end{bmatrix}$ *determinar sus valores propios y sus correspondientes vectores propios.*

solución: *La matriz A es simétrica y definida positiva debido a que:*

$$
\begin{aligned}
\Delta_1 &= 4 > 0 \\
\Delta_2 &= \begin{vmatrix} 4 & 2 \\ 2 & 5 \end{vmatrix} = 16 > 0 \\
\Delta_3 &= det(A) = 80 > 0
\end{aligned}
$$

El sistema asociado es de la forma:

$$
\left.\begin{array}{rcl}
\lambda_j x_1^j &=& 4x_1^j + 2x_2^j + 2x_3^j \\
\lambda_j x_2^j &=& 2x_1^j + 5x_2^j + x_3^j \\
\lambda_j x_3^j &=& 2x_1^j + x_2^j + 6x_3^j
\end{array}\right\} \quad j = 1, 2, 3 \tag{2.31}
$$

estableciendo $j = 1$ *y* $x_3^1 = 1$, *obtenemos:*

$$
\left.\begin{array}{rcl}
x_1^1 &=& \frac{1}{\lambda_1}(4x_1^1 + 2x_2^1 + 2) \\
x_2^1 &=& \frac{1}{\lambda_1}(2x_1^1 + 5x_2^1 + 1) \\
\lambda_1 &=& 2x_1^1 + x_2^1 + 6
\end{array}\right\} \tag{2.32}
$$

Resolviendo el sistema (2.32) y eligiendo los valores iniciales

$$x_1^{1,0} = 1 \quad y \quad x_2^{1,0} = 1$$

Se obtiene de la última ecuación del sistema (2.32) $\lambda_1^0 = 9$. *En la tabla siguiente se muestran los cálculos obtenidos por el método descrito, usando estos datos:*

UTILIZACIÓN DEL MÉTODO DE ITERACIÓN PARA EL CÁLCULO DE VALORES PROPIOS Y VECTORES PROPIOS DE UNA MATRIZ CORRESPONDIENTE A LA PRIMERA RAÍZ DE LA ECUACIÓN CARACTERÍSTICA

k	$x_1^{1,k}$	$x_2^{1,k}$	$x_3^{1,k}$	λ_1^k	k	$x_1^{1,k}$	$x_2^{1,k}$	$x_3^{1,k}$	λ_1^k
00	1	1	1	9	06	0.806	0.771	1	8.383
01	0.89	0.89	1	8.67	07	0.807	0.771	1	8.385
02	0.85	0.83	1	8.53	08	0.8074	0.7715	1	8.3863
03	0.83	0.80	1	8.46	09	0.8076	0.7717	1	8.3869
04	0.81	0.78	1	8.40	10	0.8076	0.7719	1	8.3871
05	0.805	0.770	1	8.38	11	0.8077	0.7720	1	8.3874

Podemos tomar

$$\lambda_1 = 8.3874 \quad y \quad x^1 = \begin{bmatrix} 0.8077 \\ 0.7720 \\ 1.0000 \end{bmatrix}$$

Pongamos ahora $j = 2$ en el sistema (2.31). De la condición de ortogonalidad para los vectores x^1 y x^2 tenemos

$$0.8077x_1^2 + 0.7720x_2^2 + x_3^2 = 0$$

de donde

$$x_3^2 = -0.8077x_1^2 - 0.7720x_2^2 \tag{2.33}$$

Sustituyendo esta expresión en el sistema (2.31) y haciendo $x_2^2 = 1$ tenemos:

$$\left.\begin{aligned} x_1^2 &= \frac{1}{\lambda_2}(2.3846x_1^2 + 0.4560) \\ \lambda_2 &= 1.1923x_1^2 + 4.2280 \end{aligned}\right\} \tag{2.34}$$

Resolviendo el sistema (2.34) se obtiene:

$$x_1^{(2,0)} = 1 \quad y \quad \lambda_2^0 = 5.42$$

En la tabla siguiente se hallan los cálculos correspondientes:

UTILIZACIÓN DEL MÉTODO DE ITERACIÓN PARA EL CÁLCULO DE VALORES PROPIOS Y VECTORES PROPIOS DE UNA MATRIZ CORRESPONDIENTE A LA SEGUNDA DE LA RAÍZ DE LA ECUACIÓN CARACTERÍSTICA

k	x_1^{2k}	x_2^{2k}	λ_2^k	k	x_1^{2k}	x_2^{2k}	λ_2^k
00	1	1	5.42	06	0.223	1	4.494
01	0.52	1	4.85	07	0.220	1	4.490
02	0.53	1	4.64	08	0.218	1	4.488
03	0.28	1	4.56	09	0.2174	1	4.487
04	0.25	1	4.53	10	0.2171	1	4.4868
05	0.23	1	4.500	11	0.2170	1	4.4867

Podemos tomar $\lambda_2 = 4.4867$ *y* $x_1^2 = 0.2170$, $x_2^2 = 1$.
La tercera coordenada se determina a partir de las relaciones ortogonales (2.33):

$$x_3^2 = -0.9473$$

y por tanto

$$x^2 = \begin{bmatrix} 0.2170 \\ 1 \\ -0.9473 \end{bmatrix}$$

El tercer valor propio se determina directamente de las dos relaciones ortogonales:

$$\left.\begin{aligned} 0.8077x_1^3 + 0.7720x_2^3 + x_3^3 &= 0 \\ 0.2170x_1^3 + x_2^3 - 0.9473x_3^3 &= 0 \end{aligned}\right\}$$

Haciendo $x_1^3 = 1$, *tendremos* $x_2^3 = -0.5673$, $x_3^3 = -0.3698$. *En consecuencia*

$$x^3 = \begin{bmatrix} 1 \\ -0.5673 \\ -0.3698 \end{bmatrix}$$

De la última ecuación del sistema (2.31) hallamos también, para $j = 3$,

$$\lambda_3 = 2.1260 \qquad \square$$

2.5.3 Método de la Potencia Inversa

El siguiente método iterativo es conocido como la iteración inversa, es un método efectivo para calcular un vector propio cuando se tiene en forma razonable una buena aproximación a un valor propio conocido.

ALGORITMO(MÉTODO DE LA POTENCIA INVERSA): Sea σ una aproximación al valor propio real λ_1 tal que: $|\lambda_1 - \sigma| \leq |\lambda_i - \sigma|$ $(i \neq 1)$

PASO 1 Elijamos x^0.

PASO 2 Para $k = 1, 2, \cdots$ hacemos

Resolvemos $(A - \sigma I)\hat{x}^k = x^{k-1}$, $\hat{x}^k$ es la incógnita a encontrar usando el método de la eliminación gaussiana con pivot parcial.

$$x^k = \frac{\hat{x}^k}{\max(\{\hat{x}^k\})}$$

Parar si $\|(A - \sigma I)x^k\|_\infty < c\mu\|A\|_\infty$ donde c es una constante.

Teorema 2.17. *La sucesión* $\{x^k\}$ *converge a la dirección del vector propio correspondiente al valor propio* λ_1

PRUEBA: Los valores propios de $(A - \sigma I)^{-1}$ son

$$(\lambda_1 - \sigma)^{-1}, (\lambda_2 - \sigma)^{-1}, \cdots, (\lambda_n - \sigma)^{-1}$$

y los vectores propios $v^1, v^2, \cdots, v^n$ son los mismos como los de A. Luego, como en el caso del método de la potencia, podemos escribir:

$$\begin{aligned}\hat{x}^k &= \frac{c_1}{(\lambda_1-\sigma)^k}v^1 + \frac{c_2}{(\lambda_2-\sigma)^k}v^2 + \cdots + \frac{c_n}{(\lambda_n-\sigma)^k}v^n \\ &= \frac{1}{(\lambda_1-\sigma)^k}\left[c_1v^1 + c_2\left(\frac{\lambda_1-\sigma}{\lambda_2-\sigma}\right)^k v^2 + \cdots + c_n\left(\frac{\lambda_1-\sigma}{\lambda_n-\sigma}\right)^k v^n\right]\end{aligned}$$

como λ_1 es bién cercano a σ que cualquier otro valor propio, entonces el primer término del lado derecho $\dfrac{1}{(\lambda_1-\sigma)}$ es el dominante.

Entonces $x^k \to \alpha v^1$ (múltiplo) cuando $k \to \infty$. ■

2.6 Perturbación de los valores propios

Los valores propios de matrices A reales simétricas estás bién condicionadas, es decir, pequeños cambios en los elementos de A causa pequeños cambios en los valores propios de A. El teorema de perturbación está basado en el siguiente resultado:

Teorema 2.18 (Teorema de Courant-Fischer Minmax). *Sean los valores propios de una matriz real simétrica A tales que: $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Entonces*

$$\lambda_i = \min_{S} \max_{0\neq x\in S}\left(\frac{x^TAx}{x^Tx}\right)$$

donde el mínimo es tomado sobre todos los subespacios S de dimensión $(n-i+1)$ y el máximo es tomado sobre todos los vectores x no nulos en el espacio S.

Prueba: *ver [15].*

Teorema 2.19 (Perturbación de los Valores Propios). *Sean $A, E \in \mathbb{R}(n,n)$ matrices simétricas, donde E es la matriz de perturbación de la matriz A, definamos $A' = A + E$ y consideremos $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ y $\lambda'_1 \geq \lambda'_2 \geq \cdots \geq \lambda'_n$ los valores* *propios de A y A' respectivamente. Entonces*

$$\lambda_i - \|E\|_2 \leq \lambda'_i \leq \lambda_i + \|E\|_2, \quad i = 1, 2, \cdots, n$$

Prueba:

Aplicando el teorema (2.18) para A', tenemos

$$\begin{aligned}
\lambda'_i &= min\left(\left\{ max\left(\left\{\frac{x^T A' x}{x^T x}\right\}\right)\right\}\right) \\
&= min\left(\left\{ max\left(\left\{\frac{x^T A x}{x^T x} + \frac{x^T E x}{x^T x}\right\}\right)\right\}\right) \\
&\leq min\left(\left\{ max\left(\left\{\frac{x^T A x}{x^T x}\right\}\right) + max\left(\left\{\frac{x^T E x}{x^T x}\right\}\right)\right\}\right) \\
&= min\left(\left\{ max\left(\left\{\frac{x^T A x}{x^T x}\right\}\right)\right\}\right) + min\left(\left\{ max\left(\left\{\frac{x^T E x}{x^T x}\right\}\right)\right\}\right) \\
&= min\left(\left\{ max\left(\left\{\frac{x^T A x}{x^T x}\right\}\right)\right\}\right) + \|E\|_2 \\
&= \lambda_i + \|E\|_2
\end{aligned}$$

en forma similar tenemos que $\lambda_i - \|E\|_2 \leq \lambda'_i$, de donde obtenemos:

$$\lambda_i - \|E\|_2 \leq \lambda'_i \leq \lambda_i + \|E\|_2, \quad i = 1, 2, \cdots, n$$

■

3 DESCOMPOSICIÓN DEL VALOR SINGULAR

3.1 Introducción

Sabemos que toda matriz $A \in \mathbb{C}(m,n)$ puede ser descompuesta en la forma $A = UDV^*$, donde U y V son matrices unitarias, D es una matriz diagonal. Si $A \in \mathbb{R}(m,n)$, entonces U y V son matrices ortogonales reales.
Vamos a suponer en adelante que $A \in \mathbb{R}(m,n)$. Esta descomposición es llamada *Descomposición del Valor Singular de A*

3.2 Existencia de la Descomposición del Valor Singular

Teorema 3.1. *Sea $A \in \mathbb{R}(m,n)$ una matriz. Entonces existen matrices ortogonales $U \in \mathbb{R}(m,m)$ y $V \in \mathbb{R}(n,n)$ tales que:*

$$U^T AV = \begin{pmatrix} (D_1)_{r\times r} & 0_{r\times(n-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{pmatrix} = D$$

donde D_1 es una matriz diagonal no singular, 0 son matrices nulas. Los elementos de la diagonal de D_1 son no-negativos y estos pueden, sin pérdida de generalidad, ser ordenados en forma no creciente. El número de elementos en la diagonal de D_1 es igual a $rang(A)$ (rango de A).

Prueba:

Consideremos la matriz $B = A^T A$

$\Longrightarrow$ B es una matriz simétrica semidefinida positiva.

$\Longrightarrow$ los valores propios de B son no-negativos.

Sean $\lambda_1 = \sigma_1^2, \lambda_2 = \sigma_2^2, \cdots, \lambda_n = \sigma_n^2$ los valores propios de B y supongamos que $\sigma_1 \geq \sigma_2 \geq \cdots \sigma_r > 0$ y $\sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_n = 0$

Sean $v_1, v_2, \cdots, v_n$ los vectores propios asociados a $\lambda_1, \lambda_2, \cdots, \lambda_n$.

Como B es simétrica, por el corolario (2.2) se tiene que $\{v_1, v_2, \cdots, v_n\}$ es un conjunto ortonormal y satisface:

$$
\begin{array}{rrcll}
 & Bv_i & = & \lambda_i v_i & i = 1, 2, \cdots, n \\
\Longrightarrow & A^T A v_i & = & \sigma_i^2 v_i & i = 1, 2, \cdots, n \\
\Longrightarrow & v_i^T (A^T A) v_i & = & \sigma_i^2 v_i^T v_i & i = 1, 2, \cdots, n \\
\Longrightarrow & v_i^T (A^T A) v_i & = & \sigma_i^2 & i = 1, 2, \cdots, n \\
\Longrightarrow & v_i^T (A^T A) v_i & = & \sigma_i^2 > 0 & i = 1, 2, \cdots, r \qquad (3.1) \\
\text{y} & v_i^T (A^T A) v_i & = & 0 & i = r+1, \cdots, n \qquad (3.2) \\
\text{además} & v_i^T (A^T A) v_j & = & 0 & i, j = 1, 2, \cdots, r; \quad i \neq j \qquad (3.3)
\end{array}
$$

Definamos:

$V_1 = (v_1, v_2, \cdots, v_r)$ donde v_i son vectores propios asociados a λ_i, $i = 1, 2, \cdots, r$.

$V_2 = (v_{r+1}, \cdots, v_n)$ donde v_i son vectores propios asociados a $\lambda_i = 0$, $i = r+1, \cdots, n$.

Entonces

$$
\begin{array}{rrcll}
 & V_2^T A^T A V_2 & = & V_2^T (A^T A)(v_{r+1}, \cdots, v_n) & \\
 & & = & V_2^T (A^T A v_{r+1}, \cdots, A^T A v_n) & \\
 & & = & V_2^T (0, 0, \cdots, 0) & \\
\Longrightarrow & V_2^T A^T A V_2 & = & 0 & \\
\Longrightarrow & \|A V_2\| & = & 0 & \\
\Longrightarrow & A V_2 & = & 0 & \\
\Longrightarrow & A v_k & = & 0 & k = r+1, \cdots, n \qquad (3.4)
\end{array}
$$

$$\text{Definamos } u_i = \frac{1}{\sigma_i} A v_i, \quad i = 1, \cdots, r \tag{3.5}$$

Veamos que $\{u_i\}$, $i = 1, \cdots, r$ es un conjunto ortonormal

$$\begin{aligned} u_i^T u_j &= \frac{1}{\sigma_i}(Av_i)^T \frac{1}{\sigma_j}(Av_j) \\ &= \frac{1}{\sigma_i \sigma_j} v_i^T (A^T A) v_j \\ &= \begin{cases} 0, & \text{si } i \neq j \text{ por } (3.3) \\ 1, & \text{si } i = j \text{ por } (3.1) \end{cases} \end{aligned} \tag{3.6}$$

$\Longrightarrow$ $\{u_i\}$, $i = 1, \cdots, r$ es un conjunto ortonormal.

Definamos: $U_1 = (u_1, u_2, \cdots, u_r)$ y elijamos $U_2 = (u_{r+1}, u_{r+2}, \cdots, u_m)$ tal que $U = (U_1, U_2)$ sea una matriz ortogonal.

Entonces, para cualquier $k > r$, tenemos:

$$\begin{aligned} u_k^T A v_i &= \sigma_i u_k^T u_i & i &= 1, \cdots, r & &\text{por } (3.5) \\ &= 0 & & & &\text{por } (3.6) \\ \Longrightarrow u_k^T A v_i &= 0 & i &= 1, \cdots, r & & \\ \text{y} \quad u_k^T A v_i &= 0 & i &= r+1, \cdots, r & &\text{por } (3.4) \end{aligned}$$

Sea $V = (V_1, V_2)$

$$U^T A V = \begin{pmatrix} u_1^T \\ u_2^T \\ \vdots \\ u_m^T \end{pmatrix} A(v_1, v_2, \cdots, v_n)$$

$$
= \begin{pmatrix} \frac{1}{\sigma_1} v_1^T A^T \\ \frac{1}{\sigma_2} v_2^T A^T \\ \vdots \\ \frac{1}{\sigma_r} v_r^T A^T \\ u_{r+1}^T \\ \vdots \\ u_m^T \end{pmatrix} A(v_1, v_2, \cdots, v_n)
$$

$$
= \begin{pmatrix} \sigma_1 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \sigma_3 & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \sigma_r & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}_{m\times n} = \begin{pmatrix} D_1 & 0 \\ 0 & 0 \end{pmatrix}_{m\times n} = D
$$

entonces

$$
U^T AV = \begin{pmatrix} (D_1)_{r\times r} & 0_{r\times(n-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{pmatrix} = D_{m\times n}
$$

donde $D_1 = diag(\sigma_1, \cdots, \sigma_r)$ es una matriz diagonal

Veamos que $rang(A) = rang(D) = r$:

$$
\begin{array}{lrcll}
\text{Como} & U^T AV & = & D & \\
\Longrightarrow & UU^T AV & = & UD & \\
& AV & = & UD & (UU^T = I_{m\times m} \text{ ortogonal}) \\
\Longrightarrow & AVV^T & = & UDV^T & \\
\Longrightarrow & A & = & UDV^T & (VV^T = I_{n\times n} \text{ ortogonal}) \\
\Longrightarrow & rang(A) & = & rang(UDV^T) = rang(D) = r & \blacksquare
\end{array}
$$

Definición 3.1. *Las componentes de la diagonal de la matriz* D_1 *del teorema precedente son llamadas Valores Singulares de* A *y los números* $\sigma_1, \cdots, \sigma_r$ *son llamados valores singulares positivos.*

Definición 3.2. *Las columnas de* U *son llamadas* **Vectores Singulares por la Izquierda de** A *y los de* V **Vectores Singulares por la Derecha**.

3.2.1 UNICIDAD DE LA DESCOMPOSICIÓN DEL VALOR SINGULAR

Existen $k = \min(\{m, n\})$ valores singulares de A.

Si $r = rang(A)$, entonces existen r valores singulares positivos.

Sean $\sigma_1 = \sqrt{\lambda_1}, \sigma_2 = \sqrt{\lambda_2}, \cdots, \sigma_r = \sqrt{\lambda_r} > 0$, donde $\lambda_1, \lambda_2, \cdots, \lambda_r$ son los valores propios de $A^T A$.

Si $r < k \Longrightarrow$ los $(k - r)$ valores singulares son ceros (teorema 3.1). Luego los valores singulares son únicos. Sin embargo los vectores singulares no son únicos, por ejemplo, si A tiene un valor singular $\sigma > 0$ de multiplicidad $m > 1$, entonces los correspondientes vectores columnas de V pueden ser elegidos del espacio generado por los vectores propios asociados a $\lambda = \sigma^2$ de $A^T A$.

Ejemplo 3.1. *Calcular los valores y vectores singulares de:*

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{pmatrix}$$

1. *Cálculo de Valores Singulares:*

   Sea

$$B = A^T A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 14 & 20 \\ 20 & 29 \end{pmatrix}$$

Entonces el polinomio característico de B es:

$$\begin{vmatrix} 14-\lambda & 20 \\ 20 & 29-\lambda \end{vmatrix} = (14-\lambda)(29-\lambda) - 400.$$

Luego los ceros del polinomio anterior son:

$$\lambda_1 = \frac{1}{2}(43 + 5\sqrt{73}) \quad y \quad \lambda_2 = \frac{1}{2}(43 - 5\sqrt{73}).$$

Entonces los valores singulares de A son:

$$\sigma_1 = \sqrt{\lambda_1} = \sqrt{\frac{1}{2}(43 + 5\sqrt{73})} \quad y \quad \sigma_2 = \sqrt{\lambda_2} = \sqrt{\frac{1}{2}(43 - 5\sqrt{73})}.$$

2. *Cálculo de V:*

La matriz V_1 compuesta por los vectores propios asociados a los valores propios de $B = A^T A$, en este caso $V = V_1$, debido que $r = rang(A) = 2$, donde $V_1 = (v_1, v_2) \in \mathbb{R}(2,2)$ y $v_1, v_2 \in \mathbb{R}^2$ son los vectores propios asociados a λ_1, λ_2 respectivamente.
Entonces

$$Bv_i = \lambda_i v_i, \qquad i = 1, 2$$

entonces

$$v_1 = \sqrt{\frac{73 - 3\sqrt{73}}{146}} \begin{pmatrix} 1 \\ \frac{\sqrt{73}+3}{8} \end{pmatrix}, \qquad v_2 = \sqrt{\frac{73 + 3\sqrt{73}}{146}} \begin{pmatrix} -1 \\ \frac{\sqrt{73}-3}{8} \end{pmatrix}$$

3. Cálculo de $U = (u_1, u_2, u_3)$

$$u_1 = \frac{1}{\sigma_1} A v_1 = \frac{1}{\sqrt{\frac{1}{2}(43 + 5\sqrt{73})}} \begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{pmatrix} \sqrt{\frac{73 + 3\sqrt{73}}{146}} \begin{pmatrix} 1 \\ \frac{\sqrt{73} + 3}{8} \end{pmatrix}$$

$$u_2 = \frac{1}{\sigma_2} A v_2 = \frac{1}{\sqrt{\frac{1}{2}(43 - 5\sqrt{73})}} \begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{pmatrix} \sqrt{\frac{73 - 3\sqrt{73}}{146}} \begin{pmatrix} -1 \\ \frac{\sqrt{73} + 3}{8} \end{pmatrix}$$

Luego elegimos u_3 tal que $U = (u_1, u_2, u_3)$ sea una matriz ortogonal, es decir, $u_1^T u_3 = 0$, $u_2^T u_3 = 0$ y $u_3^T u_3 = 1$, obteniéndose $\begin{pmatrix} \frac{\sqrt{6}}{6} \\ -\frac{2\sqrt{6}}{6} \\ \frac{\sqrt{6}}{6} \end{pmatrix}$.

Así tenemos

$$D = U^T A V = \begin{pmatrix} \sqrt{\frac{1}{2}(43 + 5\sqrt{73})} & 0 \\ 0 & \sqrt{\frac{1}{2}(43 - 5\sqrt{73})} \\ 0 & 0 \end{pmatrix}.$$

$\square$

Nota:

1. Asumiremos en adelante, sin pérdida de generalidad, que $m \geq n$; por que si $m < n$, consideraremos que la **Descomposición del Valor Singular** es para A^T, y si la descomposición del valor singular de A^T es UDV^T, entonces la Descomposición del Valor Singular de A es VD^TU^T.

2. Asumiremos que los valores singulares son no-crecientes.

   Luego $\sigma_{máx} = \sigma_1$ es el mayor valor singular y $\sigma_{min} = \sigma_r$ es el menor valor singular diferente de cero y denotemos

$vs(A) = \{\sigma \in \mathbb{R}/\sigma$ es un valor singular de $A\}$.

3.3 Relación entre la Descomposición del Valor Singular y la Descomposición Propia

El siguiente teorema prueba como la **Descomposición del Valor Singular** de A está relacionada con los valores propios de AA^T y A^TA.

Teorema 3.2. *Sea $A \in \mathbb{R}(m,n)$ una matriz, $A = UDV^T$ la descomposición del valor singular y sea $r = r(A)$. Entonces*

1. $V^T(A^TA)V = diag(\sigma_1^2, \sigma_2^2, \cdots, \sigma_r^2, 0, 0, \cdots, 0)_{n\times n}$

2. $U^T(AA^T)U = diag(\sigma_1^2, \sigma_2^2, \cdots, \sigma_r^2, 0, 0, \cdots, 0)_{m\times m}$

Prueba:

$$
\begin{aligned}
\text{Como } A &= UDV^T \\
\Longrightarrow A^TA &= (VD^TU^T)(UDV^T) \\
&= VD^TDV^T \\
\Longrightarrow A^TA &= V\tilde{D}V^T \qquad \tilde{D} = D^TD
\end{aligned}
$$

donde $\tilde{D} = diag(\sigma_1^2, \sigma_2^2, \cdots, \sigma_r^2, 0, 0, \cdots, 0)_{n\times n}$,
luego $\tilde{D} = V^TA^TAV$.
Para

$$
\begin{aligned}
AA^T &= (UDV^T)(VD^TU^T) \\
&= UDD^TU^T \\
\Longrightarrow AA^T &= U\hat{D}U^T \qquad \hat{D} = DD^T
\end{aligned}
$$

donde $\hat{D} = diag(\sigma_1^2, \sigma_2^2, \cdots, \sigma_r^2, 0, 0, \cdots, 0)_{m\times m}$,

luego $\hat{D} = U^T A A^T U$. ■

Del último Teorema se obtiene:

1. Los vectores singulares derechos $v_1, v_2, \cdots, v_n$ son los vectores propios de $A^T A$.

2. Los vectores singulares izquierdos $u_1, u_2, \cdots, u_m$ son los vectores propios de AA^T.

3. $\sigma_1^2, \sigma_2^2, \cdots, \sigma_r^2$ son los valores propios de AA^T y $A^T A$.

Ejemplo 3.2. *Sea*

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}$$

definamos $B_1 = A^T A = \begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix}$.

$\Longrightarrow |B_1 - \lambda I| = \lambda^2 - 6\lambda = 0 \Longrightarrow \lambda = 0,\ \lambda = 6.$

Definamos $B_2 = AA^T = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix}$.

$\Longrightarrow |B_2 - \lambda I| = 6\lambda^2 - 6\lambda^3 = 0 \Longrightarrow \lambda = 0,\ \lambda = 0,\ \lambda = 6.$

Los valores propios y vectores propios asociados de B_1 *son respectivamente:*

$$\lambda = 6, v_1 = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix}, \qquad \lambda = 0, v_2 = \begin{pmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix}$$

Los valores propios y vectores propios asociados de B_2 son respectivamente:

$$\lambda = 6, u_1 = \begin{pmatrix} \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \end{pmatrix}, \qquad \lambda = 0, u_2 = \begin{pmatrix} -\frac{2\sqrt{6}}{6} \\ \frac{\sqrt{6}}{6} \\ \frac{\sqrt{6}}{6} \end{pmatrix}, \qquad \lambda = 0, u_3 = \begin{pmatrix} 0 \\ -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix}$$

$$\Longrightarrow A = UDV^T = \begin{pmatrix} \frac{\sqrt{3}}{3} & -\frac{2\sqrt{6}}{6} & 0 \\ \frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{6} & -\frac{2\sqrt{2}}{2} \\ \frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{6} & \frac{\sqrt{2}}{2} \end{pmatrix} D \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix}$$

donde $D = \begin{pmatrix} \sqrt{\lambda} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \sqrt{6} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}.$

Se verifica que: $VA^TAV = diag(6,0) \in \mathbb{R}(2,2)$

$UAA^TU = diag(6,0,0) \in \mathbb{R}(3,3).$ □

Corolario 3.1. *Sea $A \in \mathbb{R}(n,n)$ una matriz simétrica con valores propios $\lambda_1, \lambda_2, \cdots, \lambda_n$. Entonces los valores singulares de A son $\sigma_i = |\lambda_i|$, $i = 1, 2, \cdots, n$.*

Prueba:

Como $A = A^T$

$\Longrightarrow A^TA = A^2$

Del teorema (3.2) tenemos que:

A tiene n valores singulares no-negativos que son las raíces de los n valores propios de A^2.

Debido a que los valores propios de A^2 son $\lambda_1^2, \lambda_2^2, \cdots, \lambda_n^2$

$\Longrightarrow \sigma_1 = |\lambda_1|, \sigma_2 = |\lambda_2|, \cdots, \sigma_n = |\lambda_n|$ ■

Corolario 3.2. *Una matriz $A \in \mathbb{R}(n,n)$ es no-singular si, y sólo si todos los valores singulares son diferentes de cero.*

Prueba:

Sabemos que $det(AA^T) = det(A)det(A^T) = det(A)^2$.

Luego A es una matriz no-singular $\Longleftrightarrow det(A) \neq 0 \Longleftrightarrow \lambda_i \neq 0$ (valores propios de A), $\forall\, i = 1,2,\cdots,n \overset{det(AA^T)\neq 0}{\Longleftrightarrow} \sigma_i \neq 0$ (valores singulares de A), $\forall\, i = 1,2,\cdots,n$ ■

Teorema 3.3. *Sea $A \in \mathbb{R}(m,n)$ $(m \geq n)$ y sea $C \in \mathbb{R}(m+n,m+n)$ una matriz definida por:*

$$C = \begin{pmatrix} 0_{m\times m} & A \\ A^T & 0_{n\times n} \end{pmatrix}$$

Sean $\sigma_1, \sigma_2, \cdots, \sigma_n$ los valores singulares de A. Entonces los valores propios de C son:

$$\begin{array}{llll} \lambda_1 = \sigma_1, & \lambda_2 = \sigma_2, & \cdots, & \lambda_n = \sigma_n, \\ \lambda_{n+1} = -\sigma_1, & \lambda_{n+2} = -\sigma_2, & \cdots, & \lambda_{2n} = -\sigma_n, \\ \lambda_{2n+1} = 0, & \lambda_{2n+2} = 0, & \cdots, & \lambda_{m+n} = 0. \end{array}$$

Prueba:

Sea $A = UDV^T$,

donde $U = (U_1, U_2)$, $U_1 \in \mathbb{R}(m,n)$, $U_2 \in \mathbb{R}(m,m-n)$, $D = \begin{pmatrix} D_1 \\ \tilde{0} \end{pmatrix}$, $D_1 \in \mathbb{R}(n,n)$, $\tilde{0} \in \mathbb{R}(m-n,n)$, $V \in \mathbb{R}(n,n)$.

Definamos:

$$P = \begin{pmatrix} \tilde{U}_1 & -\tilde{U}_1 & U_2 \\ \tilde{V} & \tilde{V} & \tilde{0} \end{pmatrix} \in \mathbb{R}(m+n,m+n)$$

donde $\tilde{U}_1 = \dfrac{1}{\sqrt{2}}U_1$, $\tilde{V} = \dfrac{1}{\sqrt{2}}V$

$$\Longrightarrow P^TCP = \begin{pmatrix} \tilde{U}_1^T & \tilde{V}^T \\ -\tilde{U}_1^T & \tilde{V}^T \\ \tilde{U}_2^T & \tilde{0}^T \end{pmatrix} \begin{pmatrix} 0_{m\times m} & A \\ A^T & 0_{n\times n} \end{pmatrix} \begin{pmatrix} \tilde{U}_1 & -\tilde{U}_1 & \tilde{U}_2 \\ \tilde{V} & \tilde{V} & \tilde{0} \end{pmatrix}$$

$$\Longrightarrow P^TCP = \begin{pmatrix} D_1 & 0 & 0 \\ 0 & -D_1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

donde $D_1 = diag(\sigma_1, \sigma_2, \cdots, \sigma_n)$.

Además $\sigma_1, \sigma_2, \cdots, \sigma_n, -\sigma_1, -\sigma_2, \cdots, -\sigma_n, \underbrace{0, 0, \cdots, 0}_{m-n}$ son los valores propios de C.

■

Ejemplo 3.3. *Sea*

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{pmatrix}$$

$$\Longrightarrow C = \begin{pmatrix} 0_{3\times 3} & A \\ A^T & 0_{2\times 2} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 3 & 4 \\ 1 & 2 & 3 & 0 & 0 \\ 2 & 3 & 4 & 0 & 0 \end{pmatrix} \in \mathbb{R}(3+2, 3+2).$$

Usando los resultados del ejemplo (3.1) se tiene:

$U = (U_1, U_2)$, $U_1 \in \mathbb{R}(3,2)$, $U_2 \in \mathbb{R}(3, 3-2)$, *luego*

$$\tilde{U}_1 = \frac{1}{\sqrt{2}} U_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} \left(\frac{\sqrt{73}+3}{4}\right) & \beta\left(\frac{\sqrt{73}+3}{4}\right) \\ \alpha\left(\frac{3\sqrt{73}+25}{8}\right) & \beta\left(\frac{3\sqrt{73}-25}{8}\right) \\ \alpha\left(\frac{\sqrt{73}+9}{2}\right) & \beta\left(\frac{\sqrt{73}-9}{2}\right) \end{pmatrix}$$

donde $\alpha = \sqrt{\dfrac{73-3\sqrt{73}}{73(43+5\sqrt{73})}}$ $\beta = \sqrt{\dfrac{73+3\sqrt{73}}{73(43-5\sqrt{73})}}$

$$U_2 = \begin{pmatrix} \dfrac{\sqrt{6}}{6} \\ -\dfrac{2\sqrt{6}}{6} \\ \dfrac{\sqrt{6}}{6} \end{pmatrix} = \frac{\sqrt{6}}{6}\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

$$\tilde{V} = \frac{1}{\sqrt{2}}V = \frac{1}{\sqrt{2}}\begin{pmatrix} \gamma & -\delta \\ \gamma\left(\dfrac{\sqrt{73}+3}{8}\right) & \delta\left(\dfrac{\sqrt{73}-3}{8}\right) \end{pmatrix}$$

donde $\gamma = \sqrt{\dfrac{73-3\sqrt{73}}{146}}$, $\delta = \sqrt{\dfrac{73+3\sqrt{73}}{146}}$

$\Longrightarrow P = \begin{pmatrix} \tilde{U}_1 & -\tilde{U}_1 & U_2 \\ \tilde{V} & \tilde{V} & \tilde{0} \end{pmatrix} \in \mathbb{R}(5,5)$ *donde* $\tilde{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

Entonces tenemos:

$$P^TCP = diag(t1, t2, -t1, -t2, 0)$$

donde $t1 = \sqrt{\frac{1}{2}(43+5\sqrt{73})}$ *y* $t2 = \sqrt{\frac{1}{2}(43-5\sqrt{73})}$ □

3.4 Sensibilidad del Valor Singular

Debido que el cuadrado de los valores singulares de A son justamente los valores propios de la matriz real simétrica A^TA y estudiar la sensibilidad de los valores singulares de A es estudiar la sensibilidad de los valores propios de A^TA y esto lo veremos a través del siguiente resultado:

Teorema 3.4 (Perturbación del Valor singular). *Consideremos las siguientes matrices* A, $B = A+E \in \mathbb{R}(m,n)$ $(m \geq n)$. *Sean* σ_i *y* $\tilde{\sigma}_i$, $i = 1, 2, \cdots, n$, *respectivamente los valores singulares de A y B, en orden no-decreciente. Entonces* $|\sigma_i - \tilde{\sigma}_i| \leq \|E\|_2$ *para todo* i.

Prueba

Haremos la prueba usando la relación entre los valores singulares de una matriz A y los valores propios de la matriz simétrica

$$\tilde{A} = \begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix} \in \mathbb{R}(m, n),$$

sean $\sigma_1, \sigma_2, \cdots, \sigma_k$ los valores singulares de A diferentes de cero, entonces el teorema (3.3) tenemos que:

$\sigma_1, \sigma_2, \cdots, \sigma_k, -\sigma_1, -\sigma_2, \cdots, -\sigma_k$ son los valores propios de $\tilde{A}$ distintos de cero y los otros $m - 2k$ valores propios son ceros. Definamos

$$\tilde{B} = \begin{pmatrix} 0 & B \\ B^T & 0 \end{pmatrix} \qquad \tilde{E} = \begin{pmatrix} 0 & E \\ E^T & 0 \end{pmatrix}$$

se observa que $\tilde{B} - \tilde{A} = \tilde{E}$.

Entonces volviendo a usar el teorema (3.3) tenemos que los valores propios de $\tilde{B}$ y $\tilde{E}$ están relacionadas con los valores singulares de B y E respectivamente, entonces de acuerdo al teorema (2.19) de perturbación de valores propios de matrices simétricas tenemos:

$$|\sigma_i - \tilde{\sigma}_i| \leq \|\tilde{E}\|_2 = \|E\|_2$$

■

Ejemplo 3.4. *Sean*

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \end{pmatrix} \qquad E = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0.0002 \end{pmatrix}$$

entonces los valores propios de:

$$A^TA = \begin{pmatrix} 1 & 3 & 6 \\ 2 & 4 & 7 \\ 3 & 5 & 8 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \end{pmatrix} = \begin{pmatrix} 46 & 56 & 66 \\ 56 & 69 & 82 \\ 66 & 82 & 98 \end{pmatrix}$$

son:

$$\begin{aligned} \lambda_1 &= \frac{213}{2} + \frac{1}{2}\sqrt{44457} \approx 211.92414334487142 \\ \lambda_2 &= \frac{213}{2} - \frac{1}{2}\sqrt{44457} \approx 1.075856655128756 \\ \lambda_3 &= 0 \qquad = 0.000000000000000 \end{aligned}$$

entonces los valores singulares de A son:

$$\begin{aligned} \sigma_1 &= \sqrt{\frac{213}{2} + \frac{1}{2}\sqrt{44457}} \approx 14.55761461726719 \\ \sigma_2 &= \sqrt{\frac{213}{2} - \frac{1}{2}\sqrt{44457}} \approx 1.03723510118418 \\ \sigma_3 &= \sqrt{0} \qquad = 0.000000000000000 \end{aligned}$$

Sea

$$B = A + E = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0.0002 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \\ 6 & 7 & 8.0002 \end{pmatrix}$$

entonces los valores propios de:

$$B^TB = \begin{pmatrix} 1 & 3 & 6 \\ 2 & 4 & 7 \\ 3 & 5 & 8.0002 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \\ 6 & 7 & 8.0002 \end{pmatrix} = \begin{pmatrix} 46 & 56 & 66.0012 \\ 56 & 69 & 82.0014 \\ 66.0012 & 82.0014 & 98.00320004 \end{pmatrix}$$

son:

$$\begin{aligned}\tilde{\lambda}_1 &\approx 211.927454560624\\ \tilde{\lambda}_2 &\approx 1.075745478674\\ \tilde{\lambda}_3 &\approx 0.000000000702\end{aligned}$$

entonces los valores singulares de B son:

$$\begin{aligned}\tilde{\sigma}_1 &\approx 14.55772834478732\\ \tilde{\sigma}_2 &\approx 1.03718150710189\\ \tilde{\sigma}_3 &\approx 0.00002649528260,\end{aligned}$$

además, tenemos que $\|E\|_2 = 0.0002$
luego tenemos:

$$\begin{aligned}|\tilde{\sigma}_1 - \sigma_1| &= 0.1137275201e - 03 < \|E\|_2 = 0.0002\\ |\tilde{\sigma}_2 - \sigma_2| &= 0.0535940823e - 03 < \|E\|_2 = 0.0002\\ |\tilde{\sigma}_3 - \sigma_3| &= 0.0264069522e - 03 < \|E\|_2 = 0.0002.\end{aligned}$$

□

Teorema 3.5. *Sea $A, E \in \mathbb{R}(m,n)$, σ_i, $\tilde{\sigma}_i$, $i = 1, \cdots, n$ los mismos como en el teorema (3.4). Entonces*

$$\sqrt{\sum_{i=1}^{n}(\sigma_i - \tilde{\sigma}_i)^2} \leq \|E\|_F.$$

4 APLICACIONES

La descomposición del valor singular puede ser usado para obtener propiedades muy importantes relacionadas a la estructura de las matrices, tales como la norma Euclideana, producto de matrices, almacenamiento de matrices de orden grande, entre otras propiedades que veremos en esta sección.

4.1 Normas

Teorema 4.1. *Sean* $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$ *los valores singulares de una matriz* $A \in \mathbb{R}(m,n)$. *Entonces*

1. $\|A\|_2 = \sigma_1 = \sigma_{max}$.

2. $\|A\|_F = (\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2)^{1/2}$.

3. $\|A^{-1}\|_2 = \dfrac{1}{\sigma_n} = \dfrac{1}{\sigma_{min}}$, *cuando* $A \in \mathbb{R}(n,n)$ *y no-singular.*

4. *Si* $A \in \mathbb{R}(n,n)$ *es una matriz no-singular, entonces* $Cond_2(A) = \dfrac{\sigma_1}{\sigma_n} = \dfrac{\sigma_{max}}{\sigma_{min}}$.

Prueba:

1. Sabemos que por el teorema (3.1) existen matrices ortogonales $U \in \mathbb{R}(m,m)$, $V \in \mathbb{R}(n,n)$ tales que: $U^TAV = \begin{pmatrix} D_1 & 0 \\ 0 & 0 \end{pmatrix} = D$, donde D_1 es una matriz no-singular, entonces:

$$\|A\|_2 = \|UDV^T\|_2 = \|D\|_2 = \max_i(\{\sigma_i\}) = \sigma_1 = \sigma_{max}.$$

2. De acuerdo al teorema (3.1) y la definición de $\|.\|_F$ tenemos:

$$\|A\|_F = \|UDV^T\|_F = \|D\|_F = (\sigma_1^2 + \cdots + \sigma_n^2)^{1/2}$$

3. En este caso $A \in \mathbb{R}(n,n)$ es una matriz no-singular, entonces existe $A^{-1} \in \mathbb{R}(n,n)$ y $\sigma_n = \sigma_{min} > 0$, luego $\dfrac{1}{\sigma_n}$ es el mayor valor singular de A^{-1} y aplicando el teorema (4.1)(1) tenemos:

$$\|A^{-1}\|_2 = \frac{1}{\sigma_n} = \frac{1}{\sigma_{min}}.$$

4. Por definición $Cond(A)_2 = \|A\|_2\|A^{-1}\|_2$, entonces por la parte (1) y (3) del presente teorema tenemos:

$$Cond(A)_2 = \|A\|_2\|A^{-1}\|_2 = \frac{\sigma_1}{\sigma_n} = \frac{\sigma_{max}}{\sigma_{min}}.$$

■

Ejemplo 4.1. *Sea* $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \\ 6 & 7 & 7.999 \end{pmatrix}$ *cuyo valores singulares son:*

$\sigma_1 = 14.557046000106184607$

$\sigma_2 = \ 1.037503124456149584$

$\sigma_3 = \ 0.000132424274849910$

y de acuerdo al teorema (4.1) tenemos:

1. $\|A\|_2 = \sigma_1 = 14.557046000106184607.$

2. $\|A\|_F = \sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_3^2} = 14.5939713921865.$

3. $\|A^{-1}\|_2 = \dfrac{1}{\sigma_3} = \dfrac{1}{0.000132424274849910} = 7551.48556796174.$

4. $Cond(A)_2 = \dfrac{\sigma_1}{\sigma_3} = \dfrac{14.557046000106184607}{0.000132424274849910} = 109927.322781155.$
Notamos con la última igualdad que la $Cond(A)$ *es un número "grande",*

por lo tanto cualquier sistema que formemos con la matriz A, diremos que dicho sistema está mal condicionado de acuerdo con la definición (1.4). □

4.2 Problema de Almacenamiento

Cuando se tiene una matriz $A \in \mathbb{R}(m,n)$ de gran tamaño, tenemos el problema de almacenamiento, entonces podemos hacer uso de los vectores singulares por la izquierda y por la derecha respectivamente de A, sabiendo que se tiene r ($r = rang(A)$) valores singulares positivos y $\sigma_{r+1} = \cdots = \sigma_n = 0$, y esto nos permite tener $(m+n)$ lugares en la memoria en vez de mn; lo cual queda justificado por el siguiente teorema.

Teorema 4.2. *Sean $A \in \mathbb{R}(m,n)$ una matriz y $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ los r valores singulares de A. Entonces:*

$$A = \sum_{j=1}^{r} \sigma_j u_j v_j^T$$

Prueba:

Debido al teorema (3.1) existen matrices ortogonales $U \in \mathbb{R}(m,m)$ y $V \in \mathbb{R}(n,n)$ tales que: $A = UDV^T$, entonces

$$A = (u_1, u_2, \cdots, u_m) \begin{pmatrix} \sigma_1 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \sigma_3 & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \sigma_r & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_n^T \end{pmatrix}$$

de donde obtenemos:

$$A = \sum_{j=1}^{r} \sigma_j u_j v_j^T$$

teniendo en cuenta que $\sigma_{r+1} = \cdots = \sigma_n = 0$. ■

Ejemplo 4.2. *Sea la matriz* $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \end{pmatrix}$, *cuyos valores singulares son:*

$\sigma_1 = 14.55761461726719$

$\sigma_2 = \ \ 1.03723510118419$

$\sigma_3 = \ \ 0.00000000000000$

de donde $r = 2$ *y de acuerdo al teorema (3.1) obtenemos:*

$$v_1 = \begin{pmatrix} 0.46250864690848 \\ 0.57056873257364 \\ 0.67862881823680 \end{pmatrix}, v_2 = \begin{pmatrix} -0.78703181947633 \\ -0.08823069048037 \\ 0.61057043851561 \end{pmatrix}$$

$$u_1 = \frac{1}{\sigma_1} A v_1 = \begin{pmatrix} 0.25000885532780 \\ 0.48517186008115 \\ 0.83791636721117 \end{pmatrix}, u_2 = \frac{1}{\sigma_2} A v_2 = \begin{pmatrix} 0.83705045665976 \\ 0.32667036802046 \\ -0.43889976493849 \end{pmatrix}$$

y de esta manera se tiene:

$$\sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T = A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \end{pmatrix}.$$

Como observamos, en la representación de A no hacemos uso de los vectores u_3 y v_3. □

4.3 Solución de Sistemas de Ecuaciones Lineales

La descomposición del valor singular de una matriz también puede ser utilizada para resolver sistemas de ecuaciones lineales de la forma:

$$Ax = b$$

donde $A \in \mathbb{R}(m, n)$, $b \in \mathbb{R}^m$ son dados y $x \in \mathbb{R}^n$ es la incóginita a encontrar; entonces hacemos lo siguiente: descomponemos $A = UDV^T$ y lo reemplazamos en el sistema original obteniendo

$$\begin{aligned} UDV^T x &= b \\ DV^T x &= U^T b \\ Dy &= b' = U^T b \end{aligned}$$

donde $y = V^T x$, ahora simplemente resolvemos la última ecuación que es mucho fácil, donde $D \in \mathbb{R}(m, n)$ es la matriz del teorema (3.1), pero no es recomendable para matrices de orden pequeño.

5 CONCLUSIONES

Observamos que la descomposición del valor singular de una matriz $A \in \mathbb{R}(m,n)$ abrevia los cálculos, esto es debido a la propiedad que tiene la matriz $A^T A$ (simétrica) con sólo calcular sus valores propios que son no-negativos. Geométricamente, la descomposición del valor singular de la matriz $A \in \mathbb{R}(m,n)$ puede interpretarse como una tranformación de la siguiente manera:

$$\begin{aligned} A \; : \; \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ x &\longmapsto y = A(x) \end{aligned}$$

donde $\|x\|_2 = 1$.

En el siguiente ejemplo muestra geométricamente la descomposición del valor singular de una matriz.

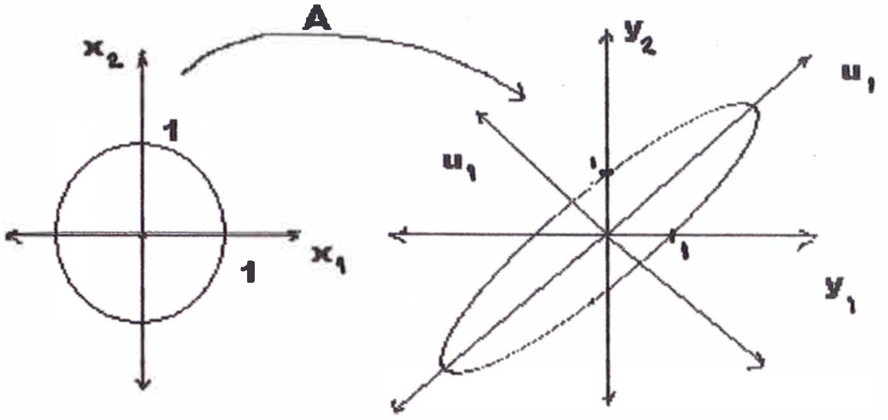
Ejemplo 5.1. *Sea la matriz* $A = \begin{pmatrix} 0.72 & 1.04 \\ 1.46 & 0.72 \end{pmatrix} \in \mathbb{R}(2,2)$, *en este caso* $m = n = 2$, *entonces los valores singulares de* A *son:* $\sigma_1 = 2.00$, $\sigma_2 = 0.50$; $U = \begin{pmatrix} 0.6 & -0.8 \\ 0.8 & 0.6 \end{pmatrix}$ *y* $V = \begin{pmatrix} 0.8 & 0.6 \\ 0.6 & -0.8 \end{pmatrix}$ *entonces tenemos:*

$$Ax = UDV^T x = \begin{pmatrix} 0.6 & -0.8 \\ 0.8 & 0.6 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 0.5 \end{pmatrix} \begin{pmatrix} 0.8 & 0.6 \\ 0.6 & -0.8 \end{pmatrix} x$$

consideremos $x = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$, *entonces*

$$y = Ax = UDV^T x = \begin{pmatrix} 0.052\sqrt{2} \\ -0.018\sqrt{2} \end{pmatrix}$$

graficamente se tiene: □



Finalmente podemos indicar que la *Descomposición del Valor Singular* de una matriz $A \in \mathbb{R}(m, n)$ es una de las herramientas más importantes que solamente está basado en el cálculo de los valores propios de $A^T A$ y que tiene muchas aplicaciones incluso en la biología y medicina la cual no tratamos en el presente trabajo.

BIBLIOGRAFIA

[1] R. L. Burden, Análisis Numérico, Grupo Editorial Iberoaméricana , 1985

[2] B. N. Datta, Numerical Linear Algebra and Applications, Books Publishing Company. 1995.

[3] R. A. DeCarlo, Linear Systems: A State Variable Approach with Numerical Implementation, Printice Hall, 1989.

[4] P. Deift, "The Bidiagonal Singular Value Decomposition and Hamiltonian Mechanics," SIAM J. Numer. Anal. v 1463-1516,1991.

[5] B. P. Deminovich y I. A. Maron, Cálculo Numérico Fundamental, Paraninfo- Madrid, 1985.

[6] W. Echegaray, "Enumeração, Isolamento e Cálculo de Zeros de Polinômios Complexos, Tésis de Maestría, POA-Brasil, 1996.

[7] G. E. Forsythe, M. A. Malcolm and C. B. Moler, Computer Methods for Mathematical Computations, Englewood Cliffs. N. J. Printice Hall.

[8] N. Gastinel, Análisis Numérico Lineal, Editorial Reverté. S.A.

[9] G. H. Golub and Van Loan, Matrix Computations, Johns Hopkins Univ. Press, Baltimore Maryland, New York, 1989

[10] D. Kincaid y W. Cheney, Análisis Numérico, Addison-Wesley Iberoamericana

[11] M. J. Maron, Análisis Numérico, CECSA México, 1995.

[12] E. D. Nering, Álgebra Lineal y Teoría de Matrices, Editorial Limusa, México.

[13] B. Noble y J. W. Daniel, Álgebra Lineal Aplicada, Prentice-Hall Hispanoamericana. S.A.

[14] A. Ralston, Análisis Numérico, un primer curso, McGraw-Hill Company.

[15] Y. Saad, Numerical Methods for Large Eigenvalue Problems: Theory and Algorithms. New York: John Wiley, 1992.

[16] G. W. Stewart, Matrix Algorithms. Volume I: Basic Decompositions, Siam 1998.

[17] J. Stoer and R. Bulirsch, Introduction to Numerical Analysis, Springer Verlag. 1992