Obligatorio 1

Cesar Agustin Cortondo Landache - 4.966.059-0
Jorge Miguel Machado Ottonelli - 4.876.616-9
Eber Manuel Rodríguez Gonzalez - 5.097.757-4
Jorge Daniel Perez Kranilovich - 5.132.466-7
Grupo 63 - Tutor: Francisco Carballal
Obligatorio 1 - Métodos Numéricos 2021.

Setiembre 2021

Facultad de Ingeniería. Universidad de la República
Montevideo, Uruguay

Resumen

Resolución de problemas de búsqueda de matrices con ciertas condiciones mediante proyecciones en conjuntos, implementados mediante algoritmos iterativos.

Keywords: Matrices, Hilbert, Proyecciones, Convergencia, Schur, Frobenius, Simetría, Estocástico.

Índice

1. Introducción

Se consideran los siguientes dos problemas:

- **(P1)** Dada una lista de n de números complejos $\lambda = \{\lambda_1, \ldots, \lambda_n\}$, encontrar una matriz no-negativa $n \times n$ con auto-valores λ (en caso de que exista).
- **(P2)** Dada una lista de n números reales $\lambda = \{\lambda_1, \ldots, \lambda_n\}$, encontrar una matriz simétrica no-negativa $n \times n$ con auto-valores λ (en caso de que exista).

Siendo una matriz real $n \times n$ no-negativa si cada una de sus entradas son no-negativas.

1.1. Interpretación del problema

Una alternativa posible a (P2) es considerar el siguiente problema de optimización.

$$\min_{Q^TQ=I, R=R^T} \frac{1}{2} \left\| Q^T \Lambda Q - R \circ R \right\|^2 \tag{1}$$

Donde Λ es una matriz diagonal constante con el espectro deseado, y $\circ$ representa el producto Hadamard ($(A \circ B)_{ij} = (A)_{ij}\ (B)_{ij}$).

Se vera ahora como trabaja esta optimización para constatar si realmente es una alternativa a (P2).

1.1.1. Proposición. *Todas las matrices simétricas de dimensión $n \times n$, con n auto-valores reales $\lambda = \{\lambda_1, \ldots, \lambda_n\}$ son exactamente los elementos del conjunto:*
$\Omega = \{Q^T \Lambda Q : Q \in \mathbb{R}^{n \times n}; Q^T Q = I\}$, *donde Λ es una matriz diagonal constante con el espectro deseado.*

Demostración. Se define el conjunto de todas las matrices simétricas reales con valores propios reales $\lambda = \{\lambda_1, \ldots, \lambda_n\}$, $K = \{A \in M_{n \times n} : A = A^T, Av = \lambda_i v, V \in \mathbb{R}^n, \forall i = 1, .., n\}$ se quiere probar una doble inclusión, es decir si se toma un elemento de Ω ver que este pertenece a K y viceversa para probar que los conjuntos son los mismos.

1. $(K \subset \Omega)$. Sea S una matriz simétrica del espectro (tiene los λ autovalores). La condición necesaria y suficiente para que una matriz cuadrada sea diagonalizable es que exista una base del espacio vectorial formada por vectores propios del endomorfismo asociado a la matriz dada. En particular esa base esta asegurada para las matrices simétricas.

   Al ser S simétrica es diagonalizable, entonces por definición:

   $$S = PDP^{-1} \tag{2}$$

   Ahora se debe verificar que los sub-espacios propios correspondientes a autovalores distintos son ortogonales entre sí.

   Se consideran $\overrightarrow{x}$ y $\overrightarrow{y}$ autovectores de S correspondientes a los autovalores λ_1 y λ_2 respectivamente. Se calcula el producto:

   $$\overrightarrow{x}^T S \overrightarrow{y} = \overrightarrow{x}^T \lambda_2 \overrightarrow{y} = \lambda_2 \overrightarrow{x}^T \overrightarrow{y} \tag{3}$$

Como S simétrica se puede escribir lo anterior como:

$$\overrightarrow{x}^T S \overrightarrow{y} = \overrightarrow{x}^T S^T \overrightarrow{y} = (\overrightarrow{x} S)^T \overrightarrow{y} = (\lambda_1 \overrightarrow{x})^T \overrightarrow{y} = \lambda_1 \overrightarrow{x}^T \overrightarrow{y} \tag{4}$$

Igualando los últimos términos de (3) y (4) se obtiene:

$$\lambda_2 \overrightarrow{x}^T \overrightarrow{y} = \lambda_1 \overrightarrow{x}^T \overrightarrow{y} \Rightarrow (\lambda_2 - \lambda_1) \overrightarrow{x}^T \overrightarrow{y} = 0 \tag{5}$$

Como $\lambda_2 \neq \lambda_1$ entonces $\overrightarrow{x}^T \overrightarrow{y} = 0$ verificándose que los vectores propios de S son ortogonales entre si.

Dividiendo cada vector de P por su norma se obtiene una matriz Q ortogonal.

Debido a que S tiene los autovalores λ buscados, D es una matriz diagonal con los autovalores deseados por lo cual es igual que Λ en Ω. Sustituyendo ambos resultados en (2) se obtiene:

$$S = Q\Lambda Q^{-1} \text{ , donde } Q \text{ es ortogonal } (Q^{-1} = Q^T) \tag{6}$$

Operando en (6) se obtiene:

$$S = Q\Lambda Q^{-1} = Q\Lambda Q^T = Q^T \Lambda^T Q^{TT} = Q^T \Lambda Q \tag{7}$$

2. $(\Omega \subset K)$. Sea $A \in \Omega$, entonces $A = Q^T \Lambda Q$, $Q^T Q = I$.

   Entonces $Q^T = Q^{-1}$ y $A = Q^{-1} \Lambda Q$, por lo tanto A es equivalente a una matriz diagonal con valores propios λ, basta probar que A es simétrica para que pertenezca a K.

   Si transponemos A obtenemos:

$$A^T = (Q^T \Lambda Q)^T = Q^T \Lambda^T Q^{TT} = Q^T \Lambda Q^{TT} = Q^T \Lambda Q = A \tag{8}$$

□

1.1.2. Proposición. *Todas las matrices simétricas no-negativas de dimensión $n \times n$ son exactamente los elementos del conjunto* $\Gamma = \{R \circ R : R \in \mathbb{R}^{n \times n};\ R = R^T\}$.

Demostración. Se define el conjunto de todas las matrices simétricas no-negativas:
$Z = \{A \in \mathbb{R}^{n \times n} : A = A^T, (A)_{ij} \geq 0, \forall i, j : 1, .., n\}$. Se quiere probar una doble inclusión, es decir si se toma un elemento de Γ ver que este pertenece a Z y viceversa para probar que los conjuntos son los mismos.

1. $(\Gamma \subset Z)$. Sea $S \in \mathbb{R}^{n \times n}$ simétrica. Por construcción $(S \circ S) \in \Gamma$, basta probar que es simétrica y no-negativa.

   Haciendo el producto de Hadamard de S con sí misma:

$$(S \circ S)_{ij} = a_{ij}^2 \tag{9}$$

Como $a_{ij} \in \mathbb{R}$, $a_{ij}^2 \geq 0 \; \forall i, j = 1, \ldots, n$, se tiene que la matriz $(S \circ S)$ es no-negativa.

Se define $(S \circ S)_{ij} = \; b_{ij}$, y se tiene que:

$$b_{ij} = a_{ij} \times a_{ij} \; ; \text{ análogamente } b_{ji} = a_{ji} \times a_{ji} \tag{10}$$

Como S es simétrica $a_{ij} = a_{ji} \; \forall \; i, j = 1, \ldots, n$, entonces de (10) $b_{ij} = b_{ji}$, por lo tanto $(S \circ S)$ es simétrica. Entonces $(S \circ S) \in Z$, por lo tanto $\Gamma \subset Z$.

2. $(Z \subset \Gamma)$. Sea $R \in Z$, R es simétrica y no-negativa, se define $B, C \in \mathbb{R}^{n \times n}$ tal que $(B)_{ij} = \sqrt{(R)_{ij}}$ y $(C)_{ij} = -\sqrt{(R)_{ij}}$.

   Se puede ver que $(B)_{ij}$ y $(C)_{ij}$ están bien definidos en los reales debido a que $(R)_{ij} \geq 0$ $\forall i, j = 1, \ldots, n$, Por otro lado B es simétrica puesto que:

$$(B)_{ij} = \sqrt{(R)_{ij}} = \sqrt{(R)_{ji}} = (B)_{ji} \tag{11}$$

   Análogamente se tiene que C es simétrica. Y, haciendo $(B \circ B)$, se tiene que:

$$(B \circ B)_{ij} = \left(\sqrt{(R)_{ij}}\right)^2 = (R)_{ij} \; \forall ij = 1, \ldots, n \tag{12}$$

   De (12) se deduce que $(B \circ B) = R$ y se tiene $B = B^T$ (B es simétrica), entonces $R \in \Gamma$, entonces $Z \subset \Gamma$.

□

El problema (P2), se trata de encontrar una matriz que sea simétrica no-negativa y con los valores esperados, por lo tanto se busca una matriz que pertenezca a Ω y a Γ al mismo tiempo, en otras palabras $A \in \mathbb{R}^{n \times n} : A \in \Omega \cap \Gamma$.

Entonces, una matriz que pertenece a Ω es igual a una matriz que pertenece a Γ mas un error $\varepsilon_{QR} \in \mathbb{R}^{n \times n}$:

$$Q^T \Lambda Q = R \circ R + \varepsilon_{Q,R} \Rightarrow \varepsilon_{Q,R} = Q^T \Lambda Q - R \circ R \tag{13}$$

Tomando norma al cuadrado de ambos lados:

$$\|\varepsilon_{QR}\|^2 = \left\|Q^T \Lambda Q - R \circ R\right\|^2 \tag{14}$$

En general, el problema de encontrar una matriz que pertenezca a la intersección de los conjuntos es equivalente a que al minimizar la expresión (14) la norma del error sea cero.

1.2. Convexidad

Se define C convexa si $\theta x_1 + (1 - \theta) x_2 \in C$, $\forall x_1, x_2 \in C$, $\forall \theta \in [0, 1]$.

1.2.1. Convexidad de Ω

Sea $X^{(1)}, X^{(2)} \in \Omega$ para $\theta X^{(1)} + (1-\theta)X^{(2)} = R$, si $R \in \Omega \ \forall \theta \in [0,1]$, entonces Ω es convexa.

Para esto R debe cumplir simetría y tener como autovalores a $\lambda = \{\lambda_1, \ldots, \lambda_n\}$.

1. **Simetría de** R: Por la suma usual de matrices y sabiendo que $X^{(1)}$ y $X^{(2)}$ son simétricas tenemos:
$$R_{ij} = \theta X^{(1)}_{ij} + (1-\theta)\, X^{(2)}_{ij} = \theta X^{(1)}_{ji} + (1-\theta)\, X^{(2)}_{ji} = R_{ji} \tag{15}$$
Entonces R es simétrica.

   1.2.1.1. Observación. *La suma de matrices diagonales da como resultado una matriz diagonal. Es un resultado directo de aplicar la suma usual de matrices en dos matrices que solo tienen elementos distinto de nulo en su diagonal.*

2. **Espectro de** R: Por otro lado dependiendo del espectro buscado, Ω puede ser o no convexo. Cuando los auto valores cumplen que $\lambda_1 = \cdots = \lambda_n = k$, entonces, Ω es convexo.

   Demostración. Como $X^{(1)}$ y $X^{(2)}$ pertenecen a Ω entonces se pueden escribir como $X^{(1)} = Q^T \Lambda Q$, $X^{(2)} = P^T \Lambda P$. Y además $Q^T = Q^{-1}$ y $P^T = P^{-1}$.

   Entonces aplicando la ecuación de convexidad y operando se tiene:
$$\theta X^{(1)} + (1-\theta)X^{(2)} = \theta(Q^T \Lambda Q) + (1-\theta)(P^T \Lambda P) = \tag{16}$$
$$\theta(Q^{-1} \Lambda Q) + (1-\theta)(P^{-1} \Lambda P) = \theta(Q^{-1} k I Q) + (1-\theta)(P^{-1} k I P) = \tag{17}$$
$$\theta k(Q^{-1}Q) + (1-\theta)k(P^{-1}P) = \theta k I + (1-\theta)kI = kI \tag{18}$$
   Donde kI es una matriz diagonal, que cuenta con el espectro buscado, entonces es convexo. □

   Por otro lado se tiene que:

   Si se toma $X^{(1)}$ de manera que $\begin{cases} X^{(1)}_{i,j} = 0, i \neq j \\ X^{(1)}_{i,j} = \lambda_i, i = j \end{cases}$, de donde $X^{(1)}_{1,1} = \lambda_1$, $X^{(1)}_{2,2} = \lambda_2$, es decir es la matriz con los valores propios en la diagonal y el resto de entradas 0.

   Y se toma $X^{(2)}$ de manera que $\begin{cases} X^{(1)}_{i,j} = 0, i \neq j \\ X^{(1)}_{i,j} = \lambda_i, i = j \end{cases}$, pero invirtiendo el orden de los primeros dos auto valores, es decir $X^{(2)}_{1,1} = \lambda_2$, $X^{(2)}_{2,2} = \lambda_1$, es decir es la matriz igual que $X^{(1)}$ pero con las entradas (1,1) y (2,2) intercambiadas.

   Entonces Ω no es convexo.

Demostración. Basta con encontrar un par de $X^{(1)}, X^{(2)} \in \Omega : \theta X^{(1)} + (1-\theta) X^{(2)} \notin \Omega$ $\forall \theta \in [0,1]$. Sin perdida de generalidad se puede suponer que los dos auto valores distintos son $\lambda_1 \neq \lambda_2$.

Dadas las anteriores matrices en la condición de convexidad obtenemos:

$$\theta X^{(1)} + (1-\theta) X^{(2)} = R \tag{19}$$

Como $X^{(1)}, X^{(2)}$, son diagonales por la Observación (1.2.1.1), R es diagonal. Por lo anterior, para que R pertenezca a Ω, ésta tiene que tener los valores propios en la diagonal. Operando sobre la condición de convexidad se obtiene:

$$\theta \left(X^{(1)} - X^{(2)} \right) + X^{(2)} = R \tag{20}$$

Debido a que $X^{(1)}_{i,j} = X^{(2)}_{i,j}$ $\forall i, j = 3, \ldots, k$, entonces por la ecuación (20), $R_{i,j} = X^{(2)}_{i,j}$ $\forall i, j = 3, \ldots, k$; solo resta ver que sucede con $R_{1,1}$ y $R_{2,2}$:

$$R_{1,1} = \theta (\lambda_1 - \lambda_2) + \lambda_2 \tag{21}$$

Como R es diagonal entonces $R_{1,1}$ tiene que ser igual a λ_1 o λ_2, puesto que los demás autovalores ya se asignaron a las demás posiciones de la diagonal, por construcción. Y esto se tiene que cumplir $\forall \theta \in [0,1]$.

De la ecuación (21) y tomando $\theta = \frac{1}{2}$, se tiene que:

$$R_{1,1} = \frac{1}{2} (\lambda_1 - \lambda_2) + \lambda_2 = \frac{1}{2} (\lambda_1 + \lambda_2) \tag{22}$$

De (22) se tiene que $R_{1,1} = \lambda_1$ o $R_{1,1} = \lambda_2$ sii $\lambda_1 = \lambda_2$. Y, por hipótesis, se tiene que es absurdo puesto que $\lambda_1 \neq \lambda_2$.

Análogamente $R_{2,2} \neq \lambda_1$ o λ_2, por lo tanto R tiene dos autovalores distintos al conjunto de autovalores λ, entonces $R \notin \Omega$. □

Entonces no se puede asegurar que el conjunto Ω sea convexo.

1.2.2. Convexidad de Γ

Sea $X^{(1)}, X^{(2)} \in \Gamma$, tal que $\theta X^{(1)} + (1-\theta) X^{(2)} = R, \theta \in [0,1]$. Si $R \in \Gamma$ $\forall \theta \in [0,1]$, entonces Γ es un conjunto convexo.

Demostración. Basta probar que R es no-negativa y simétrica. Por la suma usual de matrices y sabiendo que $X^{(1)}$ y $X^{(2)}$ son simétricas tenemos:

$$R_{ij} = \theta X^{(1)}_{ij} + (1-\theta) X^{(2)}_{ij} = \theta X^{(1)}_{ji} + (1-\theta) X^{(2)}_{ji} = R_{ji} \tag{23}$$

Entonces R es simétrica.

Por otro lado dado $X^{(1)}$ y $X^{(2)}$ son no-negativas. Y $\theta \in [0,1]$ por lo tanto $\theta, (1-\theta) \geq 0$, entonces $R_{ij} = \theta X^{(1)}_{ij} + (1-\theta) X^{(2)}_{ij} \geq 0$ $\forall i, j$ (R es no-negativa). □

1.3. Matrices estocásticas

Una $A \in \mathbb{R}^{n \times n}$ se dice estocástica si, A cumple que:

1. $a_{ij} \geq 0$
2. $\sum_j a_{ij} = 1, \forall\ i = 1, \ldots, n.$

Se define (P1)* como:

- **(P1)***: Dada una lista de n números complejos $\lambda = \{\lambda_1, \ldots, \lambda_n\}$, encontrar una matriz estocástica $n \times n$ con auto-valores λ (en caso de que exista).

Se puede estudiar una equivalencia entre (P1) y (P1)*.

1.3.1. Equivalencia entre (P1) y (P1)*

Notemos S como una matriz estocástica.

1.3.1.1. Proposición. 1 *es un auto-valor de* S.

Demostración. Se define λ como autovalor si cumple que dado $S \in \mathbb{R}^{n \times n}$, entonces $\lambda \in \mathbb{R}$ es un autovalor de S sii $\exists v \in \mathbb{R}^{n \times 1}$ no nulo tal que $Sv = \lambda v$.

Un equivalente a esto es que λ es un auto valor de A, si cumple: $|A - \lambda I| = 0$.

Luego si A es una matriz cuadrada, $|A| = \left|A^T\right|$ (Capítulo 2 de *Álgebra y geometría* [1]).

Como $(S - I) = R$ es una matriz cuadrada. Si se transpone R se tiene que las antiguas filas de R ahora son las columnas de R^T (sin modificar el determinante). Entonces ahora la suma de las entradas de las columnas de R^T es igual a cero. Pues:

$$\sum_{i=1}^{n} R_{ij}^T = \sum_{j=1}^{n} s_{ij} - \lambda = 1 - 1 = 0 \tag{24}$$

Si a la ultima fila de R se le suman todas sus filas predecesoras, se obtiene una fila de ceros por (24). (Este proceso no modifica el determinante).

Además, si una matriz tiene una fila de ceros, su determinante es nulo (Proposición 1 - Pagina 74 de Álgebra y geometría [1]).

Entonces $\lambda = 1$ es valor propio de S. □

1.3.1.2. Proposición. $\rho(S) = 1$, *donde* $\rho(.)$ *es el radio espectral.*

Demostración. Se define radio espectral como: Dada A una matriz cuadrada $A \in \mathbb{R}^{n \times n}$. Su radio espectral ρ es definido como el máximo de los valores absolutos de sus valores propios. $\rho(A) = \text{máx}_i\, |\lambda_i|$ con $Av_i = \lambda_i v_i\ \forall i = 1, \ldots, n$

Además se tiene que el radio espectral esta acotado por la norma operador de la forma $\rho(A) \leq \|A\|$.

Sea λ valor propio de A tal que $|\lambda| = \rho(A)_1$ y v correspondiente vector propio de norma 1. Entonces $\|A\| = \text{máx}_{\|x\|=1} \|Ax\| \leq \|Av\| = \|\lambda v\| = |\lambda| = \rho(A)$

Sea la Norma 1 de A definida como $\|A\|_1 = \text{máx}_j \sum_{i=1}^{n} |a_{ij}|$. Como el $|A| = |A^T|$, entonces los valores propios de A coinciden con los de A^T, por conclusión $\rho(A) = \rho(A^T)$.

Ahora se busca la Norma 1 para S^T (como S es estocástica la suma de las columnas de S^T es 1), entonces $\left\|S^T\right\| = 1$. Por lo cual $\rho(S^T) = \rho(S) \leq 1$, pero además como 1 es valor propio de S. Entonces $\rho(S) = 1$. □

1.3.1.3. Proposición. *Sea λ que satisface las condiciones anteriores para ser el espectro de S, y sea A una matriz no-negativa con espectro λ. Si x es un autovector de A correspondiente al autovalor 1, con entradas positivas. Entonces $D^{-1}AD$ es una matriz estocástica con espectro λ, donde $D = diag(x)$.*

Demostración. Se debe de probar la no negatividad de $D^{-1}AD$, que la suma de las entradas de sus filas es 1 y que contiene el espectro λ.

Sea $D^{-1}AD = R$, como $D = diag(x)$:

$$D = \begin{cases} d_{ij} = x_i & \text{si i=j} \\ 0 & \text{si i} \neq \text{j} \end{cases} \Rightarrow D^{-1} = C = \begin{cases} c_{ij} = 1/x_i & \text{si i=j} \\ 0 & \text{si i} \neq \text{j} \end{cases} \tag{25}$$

Como x tiene todas sus entradas positivas, entonces D y C son no-negativas. Dado que A también es no-negativa, $D^{-1}AD = CAD = R$ es no-negativa.

Por la asociativa de matrices $D^{-1}AD = D^{-1}(AD)$, y sea $(AD) = Z$ entonces:

$$z_{ij} = a_{ij}x_j \text{ , con x : } Ax = 1x = x \Rightarrow \sum_{j=1}^{n} a_{ij}x_j = x_i \ \forall i = 1, \ldots, n \tag{26}$$

Luego, como $D^{-1}AD = R$ y por (26):

$$r_{ij} = a_{ij}x_j/x_i \Rightarrow \sum_{j=1}^{n} a_{ij}x_j/x_i = x_i/x_i = 1 \tag{27}$$

Dado que R es una matriz no-negativa y por (27), se cumple que R es una matriz estocástica. Ahora resta probar que R tiene como valores propios a los elementos de λ.

Como A tiene los autovalores de λ, entonces A es diagonalizable y por lo tanto $\exists, P \in \mathbb{R}^{n\times n} : A = P^{-1}\Lambda P$, donde $\Lambda = diag(\lambda)$. Entonces $R = D^{-1}AD = D^{-1}(P^{-1}\Lambda P)D$.

Falta probar que $(D^{-1}P^{-1}) = (PD)^{-1}$

$$PP^{-1} = D^{-1}D = I \Rightarrow DPP^{-1} = D \Rightarrow DPP^{-1} = ID \Rightarrow DPP^{-1}D^{-1} = I \tag{28}$$

Asociando queda:

$$(DP)(P^{-1}D^{-1}) = I \Rightarrow (DP)^{-1} = (P^{-1}D^{-1}) \tag{29}$$

Entonces R tiene a los elementos de λ como valores propios. □

1.3.1.4. Proposición. *Sea S una matriz estocástica con valores propios λ, entonces se puede encontrar una matriz no-negativa con valores propios λ.*

Demostración. Sea $x \in \mathbb{R}^n$ no-negativo, se define $D \in \mathbb{R}^{n\times n} : D = diag(x)$, la idea de la demostración se basa en construir una matriz usando el reciproco de la Proposición 1.3.1.3.

Como S tiene λ valores propios entonces es diagonalizable, y por lo tanto se la puede reescribir como $S = P^{-1}\Lambda P$, con $\Lambda = diag(\lambda)$.

Sea la matriz $R = D^{-1}SD = D^{-1}P^{-1}\Lambda PD$, basta ver que R es no-negativa y tiene valores propios λ.

R es el producto de matrices no-negativas, por lo tanto es no-negativa. Por otro lado, anteriormente se demostró que $(PD)^{-1} = (D^{-1}P^{-1})$. Sea $M = PD$, si se reescribe $R = M\Lambda M^{-1}$, se tiene que R es diagonalizable y sus valores propios son los de λ.

En general, a partir de una matriz estocástica se obtiene una matriz no-negativa con el mismo espectro. □

Con lo visto en las proposiciones anteriores, dada una matriz no-negativa con el espectro λ, donde $1 \in \lambda$ y el vector propio asociado a 1 es no-negativo, se puede encontrar una matriz estocástica con el mismo espectro. Por otro lado, dada una matriz estocástica con dicho espectro, se puede hallar una matriz no-negativa con las mismas características.

Por lo tanto, se puede concluir que bajo esas condiciones, los problemas (P1) y (P1)* son equivalentes.

1.4. Proyecciones

H es un espacio de Hilbert si es completo con respecto a la norma $\|x\| = \sqrt{< x, x >}$. Completo en este contexto significa que cualquier sucesión de Cauchy de elementos del espacio converge a un elemento en el espacio, en el sentido que la norma de las diferencias tiende a 0.

Sea x un elemento de un espacio de Hilbert H, y sea C un sub-conjunto cerrado de H. Cualquier $c_0 \in C$ tal que $\|x - c_0\| \leq \|x - c\|$, $\forall\ c \in C$, será denominado como una proyección de x en C; $c_0 = P_C(x)$.

1.4.1. Teorema. *Sean $C_1, \ldots, C_N$ conjuntos convexos cerrados en un espacio de Hilbert H de dimensión finita. Se supone que $\bigcap_{C_i}^{N} \neq \phi$. Entonces, la secuencia:*

$$x_{i+1} = P_{C_{\phi(i)}}(x_i) \ \text{,donde } \phi(i) = (i mod N) + 1 \tag{30}$$

converge a un elemento de $\bigcap_{i=1}^{N} C_i$, comenzando por un x_0 arbitrario.

1.4.2. Teorema. *Sean C_1 y C_2 conjuntos cerrados no-vacíos en un espacio de Hilbert H de dimensión finita. Para cualquier valor inicial $y_0 \in C_2$, con $x_1 = P_{C1}(y_0)$, $y_1 = P_{C2}(x_1)$ y $x_2 = P_{C1}(y_1)$, entonces:*

$$\|x_2 - y_1\| \leq \|x_1 - y_1\| \leq \|x_1 - y_0\| \tag{31}$$

Demostración. Por definición de proyección se tiene que $x_1 = P_{C1}(y_0)$, se cumple que:

$$\|y_0 - x_1\| \leq \|y_0 - x\| \ \ \forall x \in C_1 \tag{32}$$

Sí $y_1 = P_{C2}(x_1)$, se tiene que:

$$\|x_1 - y_1\| \leq \|x_1 - y\| \ \ \forall y \in C_2 \tag{33}$$

Como $y_0 \in C_2$ se cumple:

$$\|x_1 - y_1\| \leq \|x_1 - y_0\| \tag{34}$$

Además, sí $x_2 = P_{C1}(y_1)$, entonces:

$$\|y_1 - x_2\| \leq \|y_1 - x\prime\| \ \ \forall x\prime \in C_1 \tag{35}$$

En especifico, sí $x_1 \in C_1$ se cumple:

$$\|y_1 - x_2\| \leq \|y_1 - x_1\| \tag{36}$$

Luego por definición de $\|.\|$ se sabe que:

$$\|a - b\| = \|b - a\| \tag{37}$$

Pues, $\|a-b\| = \sqrt{< a-b, a-b >} = \sqrt{a^2 - 2ab + b^2} = \sqrt{< a-b, a-b >} = \|b-a\|$.

Por (34), (36), (37) y transitividad se obtiene:

$$\|x_2 - y_1\| = \|y_1 - x_2\| \leq \|y_1 - x_1\| = \|x_1 - y_1\| \leq \|x_1 - y_0\| \tag{38}$$

En consecuencia, $\|x_2 - y_1\| \leq \|x_1 - y_1\| \leq \|x_1 - y_0\|$ □

2. Resolución de (P2)

2.1. Preliminares

Se considerará en esta sección que:

- $< A, B >= tr(AB)$
- $\mathcal{S}^n$ como el conjunto de todas las matrices simétricas reales, es decir: $\mathcal{S}^n = \{A \in \mathbb{R}^{n \times n} : A = A^t\}$

2.1.1. Observación. *Si $A \in \mathcal{S}^n$ entonces la norma:* $\|A\| = \sqrt{\langle A, A \rangle} = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} (A_{ij})^2 \right)$

2.1.2. Proposición. *El conjunto $\mathcal{S}^n$ es un espacio de Hilbert.*

Demostración. Basta probar que toda sucesión de Cauchy del conjunto $\mathcal{S}^n$ converge a un elemento de $\mathcal{S}^n$.

Sea $\{A_k\}_{k \in \mathbb{N}}$ una sucesión de Cauchy en $\mathcal{S}^n$, cada entrada ${a^{(k)}}_{i,j}$ es una sucesión en los reales. Entonces $\forall \varepsilon > 0 \ \exists k_0 : \ \forall k > k_0$ que cumple:

$$\left\| A^{(k+1)} - A^{(k)} \right\| < \varepsilon \iff \sqrt{\langle A^{(k+1)} - A^{(k)}, A^{(k+1)} - A^{(k)} \rangle} < \varepsilon \tag{39}$$

Como $A^{(k+1)} - A^{(k)} \in \mathcal{S}^n$, por Observación (2.1.1), se tiene:

$$\sum_{j=1}^{n} \left(\sum_{i=1}^{n} \left(a_{ij}^{(k+1)} - a_{ij}^{(k)} \right)^2 \right) < \varepsilon \tag{40}$$

En particular, es una sumatoria de reales positivos. Por lo tanto, cada término tiene que ser menor que ε:

$$\left(a_{ij}^{(k+1)} - a_{ij}^{(k)}\right)^2 < \varepsilon \tag{41}$$

Elevando al cuadrado cada término y llamando $\varepsilon' = \sqrt{\varepsilon}$, se puede reescribir (41) como:

$$\left|a_{ij}^{(k+1)} - a_{ij}^{(k)}\right| < \varepsilon', \forall k > k_0 \tag{42}$$

Se deduce de la expresión (42), que la sucesión real $\{a_{ij}^{(k)}\}_{k\in\mathbb{N}}$ es de Cauchy, y como el conjunto $\mathbb{R}$ es completo, entonces converge a un real l_{ij}.

Por otro lado, como las matrices de la sucesión $\{A_k\}_{k\in\mathbb{N}}$ son simétricas se cumple que $a_{ij}^{(k)} = a_{ji}^{(k)}$ $\forall i, j = 1, \ldots, n$, entonces convergen al mismo real. Como $l_{ij} = l_{ji}$, se tiene que la sucesión $\{A_k\}_{k\in\mathbb{N}}$ converge a una matriz $L \in \mathcal{S}^n$, pues tiene como entradas los l_{ij}. □

A continuación, se probará que los conjuntos Γ y Ω son cerrados. Se observa que para demostrar que un conjunto es cerrado, basta con mostrar que toda sucesión convergente en el conjunto converge a un elemento del mismo.

2.1.3. Proposición. *El conjunto* Ω *es cerrado.*

Demostración. Dada $\{A_k\}_{k\in\mathbb{N}}$ sucesión convergente en Ω, y sea $\Xi = \{W \in \mathbb{R}^{n\times n} : W^T W = I\}$. Se nota que, por definición de Ω, $A_k = Q_k^T \Lambda Q_k$, donde Q_k es una sucesión convergente del conjunto Ξ.

Se observa que como Λ es fijo, basta con probar que el conjunto Ξ es cerrado, es decir, $\lim_{k\to\infty} Q_k = R \in \Xi$.

Por sucesión de matrices reales, se tiene:

$$\lim_{k\to\infty} q_{ij}^{(k)} \to r_{ij} \tag{43}$$

Por otro lado, $(Q_k^T Q_k) = I$ $\forall k \in \mathbb{N}$, es decir:

$$(Q_k^T Q_k)_{ij} = \sum_{t=1}^{n} (q^T)_{it}^{(k)} q_{tj}^{(k)} = \sum_{t=1}^{n} (q_{tj}^{(k)})^2 = \begin{cases} 1 & i = j \\ 0 & si\ no. \end{cases} \tag{44}$$

De (43) y al aplicar el límite $k \to \infty$ en (44), se llega:

$$\lim_{k\to\infty} (Q_k^T Q_k)_{ij} = \lim_{k\to\infty} \sum_{t=1}^{n} (q_{tj}^{(k)})^2 = \sum_{t=1}^{n} (r_{tj})^2 = (R^t R)_{ij} = \begin{cases} 1 & i = j \\ 0 & si\ no. \end{cases} \tag{45}$$

Por lo tanto, se ha llegado a que $Q_k \to R \in \Xi$, concluyendo que Ξ es un conjunto cerrado. Y, en consecuencia:

$$\lim_{k\to\infty} \{A_k\}_{k\in\mathbb{N}} = \lim_{k\to\infty} Q_k^T \Lambda Q_k = R^T \Lambda R \in \Omega \tag{46}$$

Entonces, Ω es un conjunto cerrado. □

2.1.4. Proposición. *El conjunto* Γ *es cerrado.*

Demostración. Dado el conjunto $\Gamma = \{R \circ R : R \in \mathbb{R}^{n\times n}; R = R^T\}$ es el conjunto de todas las matrices simétricas no-negativas.

Sea $\{A_k\}_{k\in\mathbb{N}}$ una sucesión convergente en el conjunto Γ, y el límite de la sucesión es la matriz L. Basta probar que L es simétrica y no-negativa.

Sea $\{B_k\}_{k\in\mathbb{N}}$ una sucesión convergente de matrices simétricas, entonces cada entrada $b_{i,j}^{(k)}$ de la matriz B es una sucesión convergente de números reales; por convergencia de sucesiones reales $b_{i,j}^{(k)}$ converge a un $r_{ij} \in \mathbb{R}$ $\forall i, j = 1, \ldots, n$.

Por otro lado, como las matrices de la sucesión son simétricas cumplen que $a_{i,j}^{(k)} = a_{j,i}^{(k)}$ $\forall i, j = 1, \ldots, n$, y por lo tanto convergen al mismo valor real $r_{ij} = r_{ji}$.

Entonces, se deduce que la sucesión $\{B_k\}_{k\in\mathbb{N}}$ converge a una matriz R con entradas $r_{ij} \in \mathbb{R}$ tal que $R = R^T$.

Por definición del conjunto Γ, A_k cumple que $A_k = B_k \circ B_k$, donde B_k es una sucesión convergente de matrices simétricas que converge a una matriz R simétrica. Entonces:

$$\lim_{k\to\infty} A_k = \lim_{k\to\infty} B_k \circ B_k = R \circ R \tag{47}$$

Como R es simétrica ($R = R^T$), y $\lim_{k\to\infty} A_k = R \circ R \in \Gamma$, entonces Γ es cerrado. □

2.1.5. Teorema. *Dada $A \in \mathcal{S}^n$, sea $A = V diag(\mu_1, \ldots, \mu_n) V^T$ donde V es una matriz real ortogonal, y $\mu_1 \geq \cdots \geq \mu_n$. Si $\Lambda = diag(\lambda_1, \ldots, \lambda_n)$, con $\lambda_1 \geq \cdots \geq \lambda_n$, entonces $V\Lambda V^T$ es la mejor aproximación de Ω a A (bajo la norma Frobenius).*

2.1.6. Teorema. *Dada $A \in \mathcal{S}^n$, $A_+ \in \mathcal{S}^n$ es la mejor aproximación a A en Γ (bajo la norma Frobenius), donde $(A_+)_{ij} = max\{A_{ij}, 0\}, \forall\ 1 \leq i, j \leq n$.*

Ahora se verá que el problema (P2) (definido en la sección 1) es un caso particular de (P):

- (P) : Encontrar $X \in \Omega \cap \Gamma$

Se observa que, en la sección 1.1 se presentó una alternativa posible para la resolución del problema (P2), en la cual dado un $\Lambda \in \mathbb{R}^{n\times n}$ matriz diagonal constante con el espectro deseado, se debe encontrar $Q \in \mathbb{R}^{n\times n}$ y $R \in \mathbb{R}^{n\times n}$ soluciones factibles del siguiente problema de optimización:

$$\min_{Q^TQ=I, R=R^T} \frac{1}{2} \left\| Q^T\Lambda Q - R \circ R \right\|^2 \tag{48}$$

donde $\circ$ representa el producto Hadamard.

Dado $A \in \Gamma$ y $B \in \Omega$, donde Γ y Ω son subconjuntos cerrados y no vacíos de $\mathcal{S}^n$ (y, por lo tanto, de un espacio de Hilbert H). Y sea $A = V diag(\mu_1, \ldots \mu_n) V^T$ la descomposición de auto-vectores (V) y auto-valores ($diag(\mu_1, \ldots \mu_n)$) de A. Se tiene que, por el Teorema 2.1.5, $P_\Omega(A) = V\Lambda V^T$, dado que $V\Lambda V^T$ es la mejor aproximación de A en Ω. Por lo tanto se debe cumplir que:

$$\left\| V\Lambda V^T - A \right\| \leq \|\omega - A\| \quad \forall \omega \in \Omega \tag{49}$$

Sea $B = V\Lambda V^T$, por construcción se tiene que $B \in \mathcal{S}^n$ debido a que $B \in \Omega$. Se tiene de aplicar el Teorema 2.1.6 que $P_\Gamma(B) = B_+ = (V\Lambda V^T)_+$, esto a causa de que B_+ es la mejor aproximación de B en Γ. Entonces, por la definición de proyección:

$$\|B - B_+\| \leq \|B - \gamma\| \quad \forall \gamma \in \Gamma \tag{50}$$

Como (50) se cumple para cualquier $\gamma \in \Gamma$ se cumple en especial para $A \in \Gamma$, por lo tanto con lo anterior, (49) y (50) llegamos a que:

$$\|B - B_+\| \leq \|B - A\| \leq \|\omega - A\| \quad \forall \omega \in \Omega \tag{51}$$

Utilizando esta estrategia de aplicar proyecciones alternadas entre Γ y Ω, siempre se puede encontrar un $A' = P_\Gamma(B)$ o $B' = P_\Omega(A)$ que cumpla con (51). Con esta estrategia se esta buscando al $A \in \Gamma$ y $B \in \Omega$ que acoten lo mejor posible la diferencia entre ellos, es decir que $||B - A||$ sea el valor mínimo posible.

La mejor solución posible para lo planteado anteriormente es encontrar $A' \in \Gamma$ y $B' \in \Omega$, tal que $A' = B'$, pues $||B' - A'|| = 0$. Por lo tanto, es equivalente a encontrar una matriz simétrica X tal que $X \in \Gamma \cap \Omega$ ($X = A' = B'$). Cabe notar que para el caso en el que Ω no es convexo, se tiene que esta solución no converge. En está situación el problema de optimización no tendría una tendencia a 0 (caso en el que (P2) no tiene solución).

2.2. Algoritmo

Como se mostró anteriormente, una posible forma de encontrar la mejor solución a los problemas de tipo (P2) es encontrar una matriz $X \in \Gamma \cap \Omega$ producto de proyecciones alternadas entre Γ y Ω, si existe. Para esto, se muestra un posible pseudocódigo capaz de computar una solución:

```
    #Algoritmo 1 - Algoritmo para (P2).
    algoritmo1() {
1)      Ingresar 'maxIter' el limite máximo de iteraciones del algoritmo.
2)      Ingresar 'lambda' valor propios deseados.
3)      Generar 'Y' matriz simétrica no-negativa cualquiera con dimensiones
          iguales a la cantidad de valores propios en 'lambda'.
4)      Inicializar 'iter' = 0 //'iter' es el contador de iteraciones.
5)      do {
6)          Obtener 'V' y 'mu' auto-vectores y auto-valores (respectivamente)
              de la matriz 'Y'.
7)          Hacer X = V * diag(lambda) * traspuesta(V).
8)          Hacer X = (1/2) * (X + traspuesta(X)).
9)          Hacer Y = maxCero(X).
10)         Incrementar 'iter' en un valor.
11)     } until ( (||X-Y|| < eps) || (iter > maxIter) )
        // eps representa al valor épsilon de maquina.
12)     return 'Y' posible solución del problema.
    } //end.
```

```
    #Función auxiliar - maxCero.
    maxCero(A matriz cuadrada) {
1)      Inicializar 'n' igual a la dimensión de 'A'.
2)      for (i = 1 : n) {
3)          for (j = 1 : n) {
4)              Hacer A(i,j) =  máximo(0, real(A(i,j)));
            } //endfor
        } //endfor
5)      return 'A'.
    } //end.
```

Antes que nada, se destaca que en archivo (*./Algoritmos/algoritmo1.m*) del Anexo se incluye una implementación de este algoritmo para la plataforma de Octave.

Se aclararan secciones del pseudocódigo para evitar posibles dudas con la interpretación de éste.

Se observa que desde la línea 1 hasta la 4 se realiza el ingreso de los datos ha evaluar por el algoritmo. Se nota que se hace el ingreso de una variable *maxIter*, la cual cumplirá el rol de romper el bucle en el caso de que no se cumpla la primera condición de parada y el algoritmo ya haya iterado *maxIter* veces.

A continuación, se ingresa al ciclo del *do*. El ciclo comienza descomponiendo la matriz Y (línea 6) de la siguiente forma: $Y = V^T \mu V$, donde $V \in \mathbb{R}^{n \times n}$ sus auto-vectores y $\mu = diag(\mu_1 \dots \mu_n)$ sus auto-valores.

En la línea 7 se declara la matriz auxiliar X, la cual es el resultado de proyectar la matriz Y en el conjunto Ω, bajo las hipótesis del Teorema 2.1.1.

Para asegurar la simetría de la matriz auxiliar X, en la línea 8, se le suma su traspuesta, y para mantener sus valores propios se la multiplica por $\frac{1}{2}$.

El último paso del ciclo del *do* (línea 9) se reescribe a la matriz Y como una proyección de X en el conjunto Γ, bajo las hipótesis del Teorema 2.1.2. Se denota, que se implemento una función auxiliar para esta operación (*maxCero*), en la cual dada una matriz A cuadrada, se modifica las entradas $A_{i,j} < 0$ por 0.

Se observa que hasta aquí que, se tiene a X como la mejor aproximación en el conjunto Ω, y a Y como la mejor del conjunto Γ. Como el algoritmo trata de encontrar la matriz que cumpla pertenecer a ambos conjuntos, se busca que X e Y sean lo mas cercano posible, es decir minimizar $||X - Y||$. Lo cual será una condición de parada del ciclo (línea 11).

Cuando el algoritmo termina, retorna la matriz Y (línea 12), que es una posible solución del problema (P2), en caso de que sea posible.

Con respecto al algoritmo, se cuenta con dos limitantes, que pueden llegar a generar un problema. La primera limitante es que para que el algoritmo retorne la solución buscada, se deberán ingresar los valores propios *lambda* deseados de forma ordenada.

En caso contrario, la norma Frobenius no funciona de manera adecuada, pues, el algoritmo computa los auto-valores de forma ordenada, y si los ingresados no poseen orden esta norma puede terminar siendo muy grande.

La segunda limitante, es el hecho de que alguno de los valores propios de *lambda* sea negativo, pues es posible que este valor sea reescrito como 0 por causa de la invocación de la función auxiliar *maxCero*. Esto genera que ese valor propio no pertenezca a la matriz Y, por lo cual $Y \notin \Omega$ y la norma Frobenius no convergería a 0.

Ha causa de estos problemas es que se define una segunda condición de parada, que es la cota de iteraciones del ciclo.

3. Resolución del problema general (P1)

3.1. Preliminares

Se considerará en esta sección que: $< A, B >= tr(AB^*) = \sum_{i,j} A_{ij}\bar{B}_{ij}$, donde la norma asociada es la norma de Frobenius.

3.1.0.1. Teorema. *Dada $A \in \mathbb{C}^{n\times n}$ con auto-valores $\mu_1, \ldots, \mu_2$. Entonces existe una matriz unitaria $U \in \mathbb{C}^{n\times n}$ y una matriz triangular superior $T \in \mathbb{C}^{n\times n}$ tal que $A = UTU^*$, y $T_{ii} = \mu_i$, $i = 1, \ldots, n$.*

En esta sección se trabajara con conjuntos Ψ y Υ, donde:

- Ψ el conjunto de todas las matrices complejas con espectro λ, es decir, $\Psi = \{UTU^* : UU^* = I, T \in \Theta\}$, donde $\Theta = \{T \in \mathbb{C}^{n\times n} : T$ es triangular superior con espectro $\lambda\}$, $\lambda = \{\lambda_1, \ldots, \lambda_n\}$ lista de auto-valores complejos.
- Υ el conjunto de todas las matrices reales positivas, $\Upsilon = \{R \circ R : R \in \mathbb{R}^{n\times n}\}$

3.1.0.2. Proposición. *El conjunto de todas las matrices reales $\mathbb{R}^{n\times n}$ son un espacio de Hilbert.*

Demostración. Basta con probar que toda sucesión de Cauchy del conjunto $\mathbb{R}^{n\times n}$ converge a un elemento de $\mathbb{R}^{n\times n}$.

Sea $\{A_k\}_{k\in\mathbb{N}}$ una sucesión de Cauchy en $\mathbb{R}^{n\times n}$, cada entrada $a_{i,j}^{(k)}$ es una sucesión en los reales. Entonces $\forall \varepsilon > 0\ \exists k_0 :\ \forall k > k_0$ que cumple:

$$\left\| A^{(k+1)} - A^{(k)} \right\| < \varepsilon \iff \sqrt{\langle A^{(k+1)} - A^{(k)}, A^{(k+1)} - A^{(k)} \rangle} < \varepsilon \tag{52}$$

Se puede observar que el producto interno utilizado en en esta sección sobre las matrices reales queda una expresión igual que en la Proposición (2.1.2), por lo tanto cada sucesión de reales $a_{ij}^{(k)}$ converge a un real. A diferencia de la proposición anterior $a_{ij}^{(k)}$ no tiene porque ser igual $a_{ji}^{(k)}$, por lo tanto convergen pero no necesariamente al mismo real.

En conclusión, $\mathbb{R}^{n\times n}$ es un espacio de Hilbert. □

3.1.0.3. Observación. *Una sucesión c_k en los complejos $\mathbb{C}$, se puede reescribir como: $c_k = a_k + ib_k$, donde a_k, b_k son sucesiones en los reales, en particular si a_k y b_k convergen en los reales, entonces c_k converge en los complejos.*

3.1.0.4. Proposición. *El conjunto de todas las matrices complejas* $\mathbb{C}^{n\times n}$ *son un espacio de Hilbert.*

Demostración. Basta con probar que toda sucesión de Cauchy del conjunto $\mathbb{C}^{n\times n}$ converge a un elemento de $\mathbb{C}^{n\times n}$.

Sea $\{C_k\}_{k\in\mathbb{N}}$ una sucesión de Cauchy en $\mathbb{C}^{n\times n}$, cada entrada $c_{ji}^{(k)}$ es una sucesión en los complejos.

Por la Observación 3.1.0.3 $c_k = a_k + ib_k$, donde a_k, b_k son sucesiones en los reales, por lo tanto se pueden reescribir la sucesión de matrices complejas como:

$$C_k = A_k + iB_k \tag{53}$$

Con A_k, B_k sucesiones de Cauchy en $\mathbb{R}^{n\times n}$, por la Proposición 3.1.0.2 estas convergen a matrices $R_A, R_B \in \mathbb{R}^{n\times n}$, entonces tomando el limite cuando $k \to \infty$ en (53):

$$\lim_{k\to\infty} C_k = \lim_{k\to\infty} A_k + i \lim_{k\to\infty} B_k = R_A + iR_B \tag{54}$$

Como $R_A + iR_B \in \mathbb{C}^{n\times n}$, entonces $\mathbb{C}^{n\times n}$ es un espacio de Hilbert. □

3.1.0.5. Observación. *Sea* $C \in \mathbb{C}$ *tal que* $C = a + bi$*, donde* $a, b \in \mathbb{R}$*.*
El producto $C\bar{C} = (a + bi)(a - bi) = a^2 + b^2$*.*

3.1.0.6. Proposición. *El conjunto* Ψ *es cerrado.*

Demostración. Dada $\{A_k\}_{k\in\mathbb{N}}$ sucesión convergente en Ψ, y sea $\mathcal{U} = \{U \in \mathbb{C}^{n\times n} : UU^* = I\}$ el conjunto de matrices unitarias complejas. Basta probar que los conjuntos $\mathcal{U}$ y Θ son cerrados para que Ψ también lo sea.

Por definición, $A_k = U_k T_k U_k^*$, donde U_k y T_k son sucesiones convergentes de $\mathcal{U}$ y Θ respectivamente.

1. ($\mathcal{U}$ es cerrado). Por sucesión de matrices complejas, se tiene que:

$$\lim_{k\to\infty} u_{ij}^{(k)} = r_{ij} \in \mathbb{C} \tag{55}$$

Como $(U_k U_k^* = I)\ \forall k \in \mathbb{N}$, y al aplicar la Observación 3.1.0.5, se tiene:

$$(U_k U_k^*)_{ij} = \sum_{t=1}^{n} (U_k)_{it}(U_k^*)_{tj} = \sum_{t=1}^{n} (U_k)_{it}(\bar{U}_k^T)_{tj} = \sum_{t=1}^{n} (U_k)_{it}(\bar{U}_k)_{it} = \tag{56}$$

$$= \sum_{t=1}^{n} re((U_k)_{it})^2 + im((U_k)_{it})^2 = \begin{cases} 1 & i = j \\ 0 & si\ no. \end{cases} \tag{57}$$

De (55) al aplicar el límite $k \to \infty$ en (56) y (57), se obtiene:

$$\lim_{k\to\infty} (U_k U_k^*)_{ij} = \lim_{k\to\infty} \sum_{t=1}^{n} re((U_k)_{it})^2 + im((U_k)_{it})^2 = \sum_{t=1}^{n} re(r_{it})^2 + im(r_{it})^2 = (RR^*)_{ij} \tag{58}$$

Entonces, de (57) y (58) se tiene que $\mathcal{U}$ es cerrado.

2. (Θ es cerrada). Sea $T = \Lambda + E \in \Theta$, donde $\Lambda \in \mathbb{C}^{n\times n}$ tal que $\Lambda = diag(\lambda)$; y $E \in \Theta'$ con $\Theta' = \{T' \in \mathbb{C}^{n\times n} : T' \text{ es triangular superior con su diagonal nula}\}$.

   Dado $\{T_k\}_{k\in\mathbb{N}} = \Lambda + E_k$ sucesión convergente en Θ, donde $\{E_k\}_{k\in\mathbb{N}}$ es sucesión convergente de Θ'. Como Λ es fijo, basta probar que Θ' es cerrado.

   Por sucesión de matrices complejas, entonces:

$$\forall i < j \ : \ \lim_{k\to\infty} e_{ij}^{(k)} = c_{ij} \in \mathbb{C} \Rightarrow (C_{ij}) \begin{cases} c_{ij} & i < j \\ 0 & si\ no. \end{cases} \tag{59}$$

   Por lo tanto, $C \in \Theta'$, se llega a que Θ' es cerrado. Y, en conclusión:

$$\lim_{k\to\infty} \{T_k\}_{k\in\mathbb{N}} = \lim_{k\to\infty} \Lambda + E_k = \Lambda + C \in \Theta \tag{60}$$

   Entonces, Θ es un conjunto cerrado.

En consecuencia:

$$\lim_{k\to\infty} \{A_k\}_{k\in\mathbb{N}} = \lim_{k\to\infty} U_k T_k U_k^* = R(\Lambda + C)R^* \in \Psi \tag{61}$$

Entonces, Ψ es cerrado. □

3.1.0.7. Proposición. *El conjunto* Υ *es cerrado.*

Demostración. Dada $\{A_k\}_{k\in\mathbb{N}}$ sucesión convergente en Υ. Se prueba primero para $\Upsilon \subset \mathbb{R}^{n\times n}$. Por sucesión de matrices reales, se llega a:

$$\lim_{k\to\infty} a_{ij}^{(k)} = p_{ij} \in \mathbb{R} \tag{62}$$

Por definición del conjunto Υ, $A_k = B_k \circ B_k$, donde $\{B_k\}_{k\in\mathbb{N}}$ es una sucesión convergente en $\mathbb{R}^{n\times n}$ que cumple:

$$\lim_{k\to\infty} b_{ij}^{(k)} = l_{ij} \in \mathbb{R} \Rightarrow \lim_{k\to\infty} (A_k)_{ij} = \lim_{k\to\infty} (B_k \circ B_k)_{ij} = l_{ij}^2 \geq 0 \tag{63}$$

Al aplicar (62) y (63), se tiene:

$$\lim_{k\to\infty} (A_k)_{ij} = p_{ij} = l_{ij}^2 \geq 0 \Rightarrow \lim_{k\to\infty} A_k = P \in \Upsilon \tag{64}$$

Entonces, se obtiene que Υ es cerrado en $\mathbb{R}^{n\times n}$.

Ahora, se necesita que el conjunto Υ sea cerrado en $\mathbb{C}^{n\times n}$. Tomando Υ' como la copia de Υ en $\mathbb{C}^{n\times n}$, es decir, $\Upsilon' = \{A + iO \in \mathbb{C}^{n\times n} : A \in \Upsilon \text{ y } O \in \mathbb{R}^{n\times n} \text{ la matriz nula}\}$. Como se llego a que Υ es cerrado en $\mathbb{R}^{n\times n}$, se deduce que Υ' es cerrado en $\mathbb{C}^{n\times n}$. □

De forma análoga al problema (P2), se ve que se puede reformular el problema (P1) como un caso particular de (P). Es decir, se puede realizar la estrategia de proyecciones alternadas entre los conjuntos Ψ y Υ.

Como se vio en las Proposición 3.1.0.4, $\mathbb{C}^{n\times n}$ es de Hilbert y son no vacíos. A su vez, los conjuntos Ψ , $\Upsilon \subset \mathbb{C}^{n\times n}$ cumplen que son cerrados en $\mathbb{C}^{n\times n}$, por las Proposiciones 3.1.0.6 y 3.1.0.7.

Con lo anterior, se cumplen las hipótesis del Teorema 1.4.2., se sabe que existen proyecciones que optimizan la solución. Entonces, el aplicar una proyección del uno de estos conjuntos en el otro, de manera reiterada alternando de conjunto en conjunto, es equivalente a encontrar un $X = \Psi \cap \Upsilon$, siempre y cuando $\Psi \cap \Upsilon \neq 0$. Al no asegurar la convergencia de estos dos conjuntos, no es posible asegurar la convergencia con esta estrategia (y encontrar X).

3.1.0.8. Definición. *Sea U unitaria y $\{\widehat{\lambda}_1, ..., \widehat{\lambda}_n\}$ una permutación de λ tal que minimiza $\sum_{i=1}^{n} |\widehat{\lambda}_i - T_{ii}|^2$ sobre todas las permutaciones posibles. Se define:*

$$P_\Psi(U,T) = U\widehat{T}U^* \text{ , donde } \widehat{T} \in \Theta \text{ y } \widehat{T}_{ij} = \begin{cases} \widehat{\lambda}_i & si\ i = j \\ T_{ij} & en\ otro\ caso \end{cases} \tag{65}$$

$P_\Psi(.,.)$ proyecta en el conjunto Ψ.

3.1.1. Unicidad

Se nota que no siempre se cumple la unicidad para el operador $P_\Psi(U,T)$. Se verá en este caso un ejemplo de esto. Se considera:

$$T_1 = \begin{pmatrix} 1 & 1 & 4 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix}, T_2 = \begin{pmatrix} 2 & -1 & 3\sqrt{2} \\ 0 & 1 & \sqrt{2} \\ 0 & 0 & 3 \end{pmatrix}, U = \tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix} \textit{ unitaria} \tag{66}$$

Sea $T_2 = UT_1U^*$, se busca probar que $P_\Psi(U,T_1) \neq P_\Psi(I,T_2)$, cuando $\lambda = \{0,0,0\}$.
Se observa en este caso, que:

$$UU^* = I \iff U^* = U^{-1} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{67}$$

Entonces verifica $UT_1U^* = T_2$.
Al calcular las proyecciones del operador para T_1 y T_2, es necesario calcular $\widehat{T}_1$ y $\widehat{T}_2$:

$$\widehat{T}_1 = \begin{pmatrix} 0 & 1 & 4 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}, \widehat{T}_2 = \begin{pmatrix} 0 & -1 & 3\sqrt{2} \\ 0 & 0 & \sqrt{2} \\ 0 & 0 & 0 \end{pmatrix} \tag{68}$$

Luego, por la definición del operador se tiene que:

$$P_\Psi(U,T_1) = U\widehat{T}_1U^* = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & 3\sqrt{2} \\ \frac{1}{2} & -\frac{1}{2} & \sqrt{2} \\ 0 & 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & -1 & 3\sqrt{2} \\ 0 & 0 & \sqrt{2} \\ 0 & 0 & 0 \end{pmatrix} = I\widehat{T}_2I^* = P_\Psi(I,T_2) \tag{69}$$

Que como se puede observar son distintas. Se verá que se puede concluir de manera más general.

3.1.1.1. Observación. *Dados* $U_1, U_2 \in \mathbb{C}^{n\times n}$ *y* $T_1, T_2 \in \Theta$ *tal que* $U_1T_1U_1^*, U_2T_2U_2^* \in \Psi$, *entonces sí* $U_1T_1U_1^* = U_2T_2U_2^*$ *no implica* $P_\Psi(U_1, T_1) = P_\Psi(U_2, T_2)$.

Como se acaba de ver en el ejemplo anterior si se toma $U_2 = I$ se verifica que $U_1T_1U_1^* = UT_1U^* = T_2 = IT_2I^* = U_2T_2U_2^*$, sin embargo las proyecciones de T_1 y T_2 para el sub-conjunto que contiene a $\lambda = \{0, 0, 0\}$ como valores propios, son distintas.

El resultado anterior muestra que no se verifica la unicidad. En este contexto, se tienen los siguientes teoremas.

3.1.2. Teorema. *Sea* $A = U_1T_1U_1^* = U_2T_2U_2^*$, *donde* $U_1, U_2 \in \mathbb{C}^{n\times n}$ *son unitarias, y* $T_1, T_2 \in \mathbb{C}^{n\times n}$ *son triangulares superiores. Entonces* $\|P_\Psi(U_1, T_1) - A\| = \|P_\Psi(U_2, T_2) - A\|$.

3.1.3. Teorema. *Sea* $A = UTU^* \in \mathbb{C}^{n\times n}$ *con* U *unitaria y* T *triangular. Entonces,* $\|P_\Psi(U, T) - A\| \leq \left\|U\tilde{T}U^* - A\right\|$, $\forall \tilde{T} \in \Theta$.

Se verá que el operador P_Ψ no siempre cumple con la definición de proyección, ejemplo: Se considera $\lambda = \{0, 0\}$ y las matrices:

$$U = \tfrac{1}{5}\begin{pmatrix} -3 & 4 \\ 4 & 3 \end{pmatrix},\ T = \begin{pmatrix} 1 & -3 \\ 0 & 2 \end{pmatrix},\ \tilde{U} = \tfrac{1}{5}\begin{pmatrix} -4 & 3 \\ 3 & 4 \end{pmatrix},\ \tilde{T} = \begin{pmatrix} 0 & -3 \\ 0 & 0 \end{pmatrix} \tag{70}$$

Al operar se obtiene:

$$U^* = U = \tfrac{1}{5}\begin{pmatrix} -3 & 4 \\ 4 & 3 \end{pmatrix},\ \widehat{T} = \begin{pmatrix} 0 & -3 \\ 0 & 0 \end{pmatrix},\ \tilde{U}^* = \tilde{U} = \tfrac{1}{5}\begin{pmatrix} -4 & 3 \\ 3 & 4 \end{pmatrix} \tag{71}$$

Entonces:

$$P_\Psi(U, T) = \tfrac{1}{25}\begin{pmatrix} 36 & 27 \\ -48 & -36 \end{pmatrix};\ UTU^* = \tfrac{1}{25}\begin{pmatrix} 77 & 39 \\ -36 & -2 \end{pmatrix};\ \tilde{U}\tilde{T}\tilde{U}^* = \tfrac{1}{25}\begin{pmatrix} 36 & 48 \\ -27 & -36 \end{pmatrix} \tag{72}$$

Luego:

$$P_\Psi(U, T) - UTU^* = \tfrac{1}{25}\begin{pmatrix} -41 & -12 \\ -12 & -34 \end{pmatrix};\ \tilde{U}\tilde{T}\tilde{U}^* - UTU^* = \tfrac{1}{25}\begin{pmatrix} -41 & 9 \\ 9 & -34 \end{pmatrix} \tag{73}$$

Para poder calcular la norma, primero se calcula:

$$(\tilde{U}\tilde{T}\tilde{U}^* - UTU^*)^2 = \tfrac{1}{625}\begin{pmatrix} 1762 & -675 \\ -675 & 1237 \end{pmatrix} \text{ y } (P_\Psi(U, T) - UTU^*)^2 = \tfrac{1}{625}\begin{pmatrix} 1825 & 900 \\ 900 & 1300 \end{pmatrix} \tag{74}$$

Entonces, se obtienen las normas:

$$\|P_\Psi(U, T) - UTU^*\| = \frac{4925}{625} = 7,88 \text{ y } \left\|\tilde{U}\tilde{T}\tilde{U}^* - UTU^*\right\| = \frac{1649}{625} = 2,64 \tag{75}$$

En particular, para este ejemplo $\|P_\Psi(U, T) - UTU^*\| \not\leq \left\|\tilde{U}\tilde{T}\tilde{U}^* - UTU^*\right\|$.

3.1.4. Observación. *Dada* $A = UTU^* \in \Psi'$, $P_\Psi(U, T)$ *no satisface que* $||P_\Psi(U, T) - A|| \leq ||M - A||$ $\forall$ $M \in \Psi$.

Para esto basta que $\exists M \in \Psi$ tal que $||M-A|| < ||P_\Psi(U,T)-A||$. Se observa que $P_\Psi(U,T)$ proyecta sobre el conjunto Ψ, pero no se asegura que la misma sea una proyección.

Por otro lado, como Ψ y Ψ' son cerrados en un espacio de Hilbert, existe una proyección de $A \in \Psi'$ en Ψ (se la llamara $P(A_\Psi)$), y esta cumple $\|A - P(A_\Psi)\| \leq \|A - c\|$, $\forall\ c \in \Psi$.

Por el punto anterior se tiene un caso en que se cumple

$$\|A - P(A_\Psi)\| \leq \left\|A - \tilde{U}\tilde{T}\tilde{U}^*\right\| \leq ||A - P_\Psi(U,T)|| \tag{76}$$

Por lo cual existe al menos un caso en el que $P(A_\Psi) \neq P_\Psi(U,T)$. Entonces no se puede asegurar que suceda $||P_\Psi(U,T) - A|| \leq ||M - A||\ \forall\ M \in \Psi$.

Se observa que el uso de $P_\Psi(U,T)$ es debido que a priori es difícil de encontrar $P(A_\Psi)$.

3.1.5. Observación. *La Observación 3.1.4 no contradice el Teorema 3.1.3.*

Esto sucede debido a que M puede no ser una $P_\Psi(U,T)$, como en el ejemplo que es de la forma $P_\Psi(\tilde{U},T)$, con $U \neq \tilde{U}$. El Teorema 3.1.3 asegura que la mejor proyectada conseguida al realizar una alteración en la descomposición $A = UTU^*$, modificando T, es cuando $P_\Psi(U,T) = U\widehat{T}U$, manteniendo intacto U y U^*.

3.2. Algoritmo

```
    #Algoritmo 2 - Algoritmo para (P1).
    algoritmo2() {
1)      Ingresar 'maxIter' el limite máximo de iteraciones del algoritmo.
2)      Ingresar 'lambda' valor propios deseados.
3)      Generar 'Y' matriz no-negativa cualquiera con dimensiones
          iguales a la cantidad de valores propios en 'lambda'.
4)      Inicializar 'iter' = 0 //'iter' es el contador de iteraciones.
5)      do {
6)          Obtener 'U' y 'T' matriz unitaria y triangular superior con
                autovalores iguales a 'Y' (respectivamente) de la matriz 'Y'.
7)          Hacer X = proyectPsi(U, T, lambda).
8)          Hacer Y = maxCero(X).
9)          Incrementar 'iter' en un valor.
10)     } until ( (||X-Y|| < eps) || (iter > maxIter) )
11)     return 'Y' posible solución del problema.
    } //end.
    # La función maxCero y la variable eps son los definidos en el Algoritmo1

    #Función auxiliar - proyectPsi.
    proyectPsi(U matriz unitaria, T matriz triangular superior, lambda) {
1)      Inicializar 'n' igual a la dimensión de 'T'.
2)      Hacer lambda = permutacion_minima(T, lambda);
3)      Sea T_aux = T.
4)      for (i = 1 : n) {
```

```
5)          Hacer T_aux(i,i) = lambda(i).
        } //endfor
6)      return U*(T_aux)*(U^(-1)).
    } //end.

    #Función auxiliar - permutacion_minima
    permutacion_minima(T matriz triangular superior, lambda) {
1)      Sea 'perLambda' la matriz que contiene todas las permutaciones
          posibles de 'lambda'. Y sea 'n' la dimensión del vector 'lambda'.
2)      Hacer tamanio_minimo = -1 y lambda_resultado = lambda.
3)      for (i = 1 : factorial(n)) {
4)          Hacer tamanio = 0.
5)          for (j = 1 : n) {
6)              Hacer tamanio += absoluto(perLambda(i,j) - T(j,j))^2.
            } //endfor
7)          if ((tamanio < tamanio_minimo) || (tamanio_minimo == -1)) {
8)              Hacer tamanio_minimo = tamanio.
9)              Hacer 'lambda_resultado' igual a la fila 'i' de 'perLambda'.
            } //endif
        } //endfor
10)     return 'lambda_resultado' mínima permutación con respecto a 'T'.
    } //end.
```

Al igual que el con el algoritmo 1, en el Anexo (*./Algoritmos/algoritmo2.m*) se incluye una implementación de este algoritmo para la plataforma de Octave.

Se observa que desde la línea 1 hasta la 4 se realiza el ingreso de los datos ha evaluar, al igual que con el algoritmo anterior.

A continuación, se ingresa al ciclo del *do*. El ciclo comienza descomponiendo la matriz Y mediante la descomposición de Schur (Sección 2.3 de *Matrix Analysis* [4]). (línea 6) de la siguiente forma: $Y = UTU^*$, donde $U \in \mathbb{C}^{n\times n}$ unitaria y $T \in \Theta$ triangular superior con auto-valores iguales a los de Y.

En la línea 7 se declara la matriz auxiliar X, la cual es el resultado de proyectar la matriz Y en el conjunto Ψ, mediante el operador $P_\Psi(U,T)$; mediante el uso de la función auxiliar (*proyectPsi*).

El último paso del ciclo del *do* (línea 8) se reescribe a la matriz Y como una proyección de X en el conjunto Υ, bajo las hipótesis del Teorema 3.1.6. Se denota, el uso de la función auxiliar (*maxCero*) (misma del algoritmo 1) para esta operación.

Hasta aquí se tiene a X como la mejor aproximación en el conjunto Ψ, y a Y como la mejor del conjunto Υ. Como el algoritmo 2 trata de encontrar la matriz que cumpla pertenecer a ambos conjuntos, al igual que el algoritmo 1, se busca minimizar $||X - Y||$. Lo cual será una condición de parada del ciclo (línea 10).

Cuando el algoritmo termina, retorna la matriz Y (línea 11), que es una posible solución del problema (P1), en caso de que sea posible.

Respecto a las limitantes de este algoritmo, se mantiene un problema recurrente del algoritmo 1, este es el uso de la función auxiliar *maxCero*. Esto es a causa de que se encarga de la proyección en Υ de la matriz X, por lo tanto se reescriben sus entradas (complejas) de manera tal de que quede solo con la parte real positiva de ellas, y en aquellas entradas con parte real negativa se reescriben como 0. Esto conlleva a que la matriz resultado Y no posea los valores propios no negativos y con parte imaginaria distinta de 0, resultando en la no convergencia de la condición del ciclo, y por lo tanto del resultado del algoritmo.

Al igual que con el algoritmo 1, ha causa de estos problemas es que se define la cota de iteraciones del ciclo.

4. Experimentación numérica

4.1. Experimentación en Problema (P2)

Se incorporaran ejecuciones experimentales del Algoritmo 1; con precisión de $eps(10^{-10})$ y un máximo de 1000 iteraciones.

1. Variando λ en cada simulación. Realizando 100 simulaciones para matrices cuadradas de dimensiones n, con $n \leq 100$; para 10 valores de n distintos.

   Para esta parte se obtuvieron los datos de la siguiente tabla:

	Simulaciones en (P2) - Algoritmo 1		
Dimensión	Prom. de iteraciones (i)	Tiempo promedio (seg) (t)	Tasa de éxito (α)
3×3	29	0,00512	44 %
5×5	147	0,05423	32 %
6×6	192	0,11155	36 %
9×9	298	0,37937	23 %
10×10	440	0,69622	26 %
15×15	526	4,14000	7 %
20×20	752	6,1437	6 %
25×25	695	13,9528	4 %
45×45	NaN	NaN	0 %
50×50	NaN	NaN	0 %

Cuadro 1: Simulaciones en (P2).

   De estos datos, se puede concluir que al aumentar n la cantidad de iteraciones i tiene un comportamiento logarítmico creciente, el tiempo t tiene un comportamiento polinómico y la tasa de convergencia α lleva un comportamiento decreciente exponencial.

2. Se considerara el siguiente $\lambda = \{3 - K, 1 + K, -1, -1, -1, -1\}$ fijo, con $0 < K < 1$. Realizando 100 simulaciones para 10 valores de K elegidos de manera aleatoria y $n = 5$.

   Para esta parte se obtuvieron los datos de la siguiente tabla:

	Simulaciones en (P2) variando K - Algoritmo 1		
Valor de K	Prom. de iteraciones (i)	Tiempo promedio (seg) (t)	Tasa de éxito (α)
0,14087	298	0,17104	95 %
0,25515	333	0,19299	98 %
0,52537	294	0,17346	89 %
0,57979	280	0,16231	87 %
0,58039	283	0,16283	97 %
0,66064	294	0,17214	81 %
0,69976	251	0,14900	60 %
0,71775	244	0,14179	73 %
0,88822	221	0,12733	57 %
0,92171	220	0,12947	51 %

Cuadro 2: Simulaciones en (P2) variando K.

De estos datos se puede concluir que al aumentar K, t e i disminuyen con tendencia lineal y α disminuye con predisposición polinómica, con tendencias a el 50 % (tiende a 100 % cuando K tiende a 0).

Se brinda en el Anexo gráficas asociadas a las tablas anteriores (*./Datos de Simulaciones/Datos de Simulaciones.xlsx*), con el fin de proporcionar una visión mas clara de lo comentado.

4.2. Experimentación en Problema (P1)

Se incorporaran ejecuciones experimentales del Algoritmo 2 con precisión de $eps(10^{-10})$ y un máximo de 5000 iteraciones.

1. Variando λ en cada simulación. Realizando 100 simulaciones para cada una de las siguientes dimensiones de matrices cuadradas: 3×3, 4×4, 5×5, 6×6 y 7×7.

   Para esta parte se obtuvieron los datos de la siguiente tabla:

	Simulaciones en (P1) - Algoritmo 2		
Dimensión	Prom. de iteraciones (i)	Tiempo promedio (seg) (t)	Tasa de éxito (α)
3x3	37	0,01765	6 %
4x4	NaN	NaN	0 %
5x5	173	1,08472	2 %
6x6	NaN	NaN	0 %
7x7	NaN	NaN	0 %

Cuadro 3: Simulaciones en (P1).

Con los datos proporcionados por la tabla se puede observar que la convergencia del algoritmo 2 no es muy frecuente. La tasa de convergencia α presenta un comportamiento exponencial decreciente cuando se aumenta la dimensión estudiada.

Por otra parte para las dimensiones 3×3 y 5×5 en las que se obtuvieron convergencia en alguna de las simulaciones se aprecia que el tiempo y la cantidad de iteraciones tienen ambas un comportamiento exponencial creciente con respecto a la dimensión; por lo tanto, se puede concluir que a medida que incrementa la dimensión la convergencia es menos frecuente.

Se puede ver además que para las dimensiones 4×4, 6×6 y 7×7 no se obtuvieron convergencias con la cantidad de simulaciones realizadas. Entonces, se observa que las dimensiones para las cuales se obtuvo convergencia son impares, esto puede ser a causa de que tienen por lo menos un valor propio real. Por lo tanto, es posible intuir que para dimensiones pares al existir la posibilidad de que todos los valores propios sean complejos afecte la convergencia, disminuyendo la misma.

2. Se considera el siguiente problema: Encontrar una matriz estocástica X con espectro

$$\lambda = \{\lambda_1, \ldots, \lambda_5\} : \lambda_1 = 1,\ \lambda_i = \begin{cases} \text{mín}\{1, l_i\},\ si\ l_i \geq 0 \\ \text{máx}\{-1, l_i\},\ si\ l_i < 0 \end{cases},\ Sea:\ Z = \begin{pmatrix} . & . & 0 & 0 & . \\ . & . & . & 0 & 0 \\ 0 & . & . & . & 0 \\ 0 & 0 & . & . & . \\ . & 0 & 0 & . & . \end{pmatrix}$$

Donde $l_i \sim \mathcal{N}(\ 0,\ \ 1)$, para $i = 2, \ldots, 5$. Además X debe de cumplir, tener 0 en las mismas posiciones que Z.

La modificación realizada al Algoritmo 2 para que este resuelva el problema que se acaba de plantear, radica en el intercambio de la función auxiliar *maxCero*, por otra función auxiliar llamada *maxZeta* que se muestra a continuación:

```
    #Función auxiliar - maxZeta.
    maxZeta(A matriz cuadrada de dimensión 5) {
1)      Hacer A = maxCero(A).
2)      Hacer '0' a las entradas de 'A' de acuerdo al patrón Z (tridiagonal
             con excepción de las esquinas).
3)      for (i = 1 : 5) {
4)          Hacer 'sumaFila' igual a la sumatoria de la fila 'i' de 'A'.
5)          for (j = 1 : 5) {
6)              Hacer A(i,j) = A(i,j)/sumaFila.
            } //endfor
        } //endfor
7)      return 'A'.
    } //end.
    # La función maxCero es la definida en el Algoritmo1.
```

Para verificar que este algoritmo resuelve el problema se puede probar con los siguientes datos (valores aproximados, para los valores exactos dirigirse al anexo *./Datos de Simulaciones/Simulaciones (P1)aster estocastico/*), donde se presentan a la matriz de entrada Y y los valores propios λ, así como la matriz solución Y_{res}:

$$\lambda = \begin{pmatrix} 1; & -0{,}308; & 0{,}698; & -0{,}623; & 0{,}148 \end{pmatrix}$$

$$Y = \begin{pmatrix} 7{,}118 & 4{,}870 & 1{,}008 & 4{,}420 & 6{,}992 \\ 1{,}105 & 2{,}267 & 3{,}119 & 4{,}747 & 9{,}213 \\ 3{,}354 & 3{,}212 & 9{,}076 & 2{,}220 & 0{,}491 \\ 7{,}481 & 7{,}806 & 9{,}754 & 0{,}034 & 7{,}603 \\ 3{,}777 & 5{,}889 & 3{,}954 & 3{,}887 & 8{,}665 \end{pmatrix}; Y_{res} = \begin{pmatrix} 0{,}006 & 0{,}248 & 0 & 0 & 0{,}746 \\ 0{,}153 & 0{,}106 & 0{,}740 & 0 & 0 \\ 0 & 0{,}240 & 0{,}580 & 0{,}179 & 0 \\ 0 & 0 & 0{,}480 & 0 & 0{,}519 \\ 0{,}401 & 0 & 0 & 0{,}377 & 0{,}222 \end{pmatrix}$$

Se brinda en el anexo (*./Datos de Simulaciones/Simulaciones (P1)aster estocastico/*) aún mas casos de convergencia del algoritmo modificado.

5. Conclusiones

- A partir de los resultados obtenidos en las secciones 4.1 y 4.2, se puede concluir que en términos de tasa de éxito y tiempo promedio de cada simulación el algoritmo 1 es mas eficiente que el 2, en especial para valores de n mayores.

  En cuestión de la cantidad de iteraciones para aproximar la norma no se encontraron resultados concluyentes a pesar de la diferencia en la cantidad máxima de iteraciones sugerida. Cabe destacar que para las dimensiones pares el algoritmo 2 no se encontraron casos de convergencia por lo cual no se deduce un patrón. Esto no sucede en el caso del algoritmo 1, para el cual es indistinto a la paridad de las dimensiones.

- En cuestión de t la diferencia esta dada principalmente por el orden de la implementación del algoritmo 2, en particular la función *proyectPsi* es $O(n!)$.

  Por otra parte, la diferencia de α es consecuencia de dos factores. En primer lugar por la implementación mostrada es mas probable que se imponga un 0 en lugar del valor propio en la diagonal como resultado de la función *maxCero* para entradas complejas. En segundo lugar se aprecia empíricamente que las posibles soluciones del problema 1 es menos denso que el de las posibles soluciones del problema 2.

- En general, el algoritmo 2 no resulta ser un algoritmo efectivo para la resolución del problema 1. Esto se da especialmente porque se eligieron los λ de la matriz aleatoria a partir de una distribución uniforme dentro del rango de valores. De esta manera la probabilidad de tomar un λ real (parte imaginaria igual a 0) es 0 para cualquier matriz de dimensión par. Al tener mayor cantidad de complejos con parte imaginaria no nula es mas probable que *maxCero* imponga 0 en lugar del valor propio, reduciendo la tasa de éxito.

  En general el algoritmo 1 resulta ser un algoritmo bastante efectivo para la resolución del problema 2. Tiene una convergencia asintótica a 1000 con respecto a la cantidad de iteraciones necesarias para alcanzar un éxito, (que tiene sentido con el limite de iteraciones impuesto).

Referencias

[1] Eugenio Hernández, *Álgebra y geometría*. Addison Wesley, Universidad Autónoma de Madrid, 1era Edición, 1994. ISBN 0-201-62586-5.

[2] Ramón Bruzual, & Marisela Domínguez, *Espacios de Hilbert*. Facultad de Ciencias, Universidad Central de Venezuela, 1era Edición, 2005.

[3] Eleonora Catsigeras, *Elementos de topología usados en Cálculo*. Instituto de Matemática y Estadística Rafael Laguardia (IMERL), Fac. Ingeniería, Universidad de la República, Versión preliminar, 2004.

[4] Roger A. Horn, & Charles R. Johnson, *Matrix Analysis*. Cambridge University Press, 2da Edición, 1985. ISBN 0-521-38632-2.