Espacios invariantes. Núcleos iterados.

Rommel Andrés Bustinza Pariona

Facultad de Ciencias Físicas y Matemáticas
Universidad de Concepción

June 8, 2021

Espacios invariantes

Sea V un $\mathbb{K}$-espacio vectorial (no necesariamente de dimensión finita), $W \subseteq V$ un subespacio vectorial, y $T \in \mathcal{L}(V)$. Se dice que W es invariante con respecto a T si $T(W) \subseteq W$, i.e. $T(W)$ es subespacio vectorial de W.

OBSERVACIÓN: Para establecer que $T(W) \subseteq W$, puede pensarse en su caracterización usual:

$$\forall\, z \in V : z \in T(W) \Rightarrow z \in W\,,$$

la que en este contexto, también es equivalente con: $\forall\, x \in W : T(x) \in W$.

Ejemplos:

1. $W = \{\theta\}$ y $W = V$ son espacios invariantes (llamados triviales) de cualquier $T \in \mathcal{L}(V)$, pues $T(\{\theta\}) = \{\theta\}$ y $T(V) \subseteq V$.
2. Sea $T \in \mathcal{L}(V)$. $\mathrm{Ker}(T)$ e $\mathrm{Im}(T)$ son subespacios invariantes con respecto a T.
3. Sea $T \in \mathcal{L}(V)$ con espectro no vacío. Entonces $\forall\, \lambda \in \sigma(T) : T(S_\lambda) \subseteq S_\lambda$, ya que se demuestra: $\forall\, \lambda \in \sigma(T) : \forall\, v \in S_\lambda : T(v) = \lambda\, v \Rightarrow T(v) \in S_\lambda$.
4. Sea $T \in \mathcal{L}(\mathcal{P}(\mathbb{R}))$, definida para cada $p \in \mathcal{P}(\mathbb{R})$, por $T(p) := p'$. Entonces $\mathcal{P}_5(\mathbb{R})$ es un subespacio de $\mathcal{P}(\mathbb{R})$, es invariante con respecto a T.

CARACTERÍSTICA IMPORTANTE DE LOS SUBESPACIOS T-INVARIANTES: Dados V un $\mathbb{K}$-espacio vectorial, $T \in \mathcal{L}(V)$ y $W \subseteq V$ un subespacio T-invariante de V, entonces $T|_W : W \to W$ (la restricción de T a W) es un endomorfismo de W, i.e. $T|_W \in \mathcal{L}(W)$.

Proposición: Sea V un $\mathbb{K}$-espacio vectorial, W un subespacio vectorial de V, B_W una base de W, y $T \in \mathcal{L}(V)$. Luego W es invariante con respecto a T si y sólo si $T(B_W) \subseteq W$.

Demostración:

($\Rightarrow$) Hipótesis: $T(W) \subseteq W$. Como $B_W \subseteq W$, se tiene que $T(B_W) \subseteq T(W) \subseteq W$, y se concluye.

($\Leftarrow$) Hipótesis: $T(B_W) \subseteq W$. Sea $x \in T(W)$. Esto implica que $\exists\, z \in W \,:\, T(z) = x$. En vista que B_W es una base de W, $\exists\,\{w_j\}_{j=1}^{m} \subseteq B_W$, $\exists\,\{\alpha_j\}_{j=1}^{m} \subseteq \mathbb{K}$ tales que $z = \displaystyle\sum_{j=1}^{m} \alpha_j\, w_j$. Esto conduce a afirmar que $x = T(z) = \displaystyle\sum_{j=1}^{m} \alpha_j\, T(w_j) \in W$, pues W es subespacio vectorial de V. De esta forma, como $x \in T(W)$ es fijo pero arbitrario, se deduce que $T(W) \subseteq W$, i.e. W es invariante con respecto a T.

Definiciones equivalentes de Suma Directa Generalizada

TEOREMA: Sean V un $\mathbb{K}$-espacio vectorial, y $\{W_j\}_{j=1}^m$ una familia finita de subespacios de V. Las siguientes afirmaciones son EQUIVALENTES:

(a) $V = \displaystyle\bigoplus_{j=1}^{m} W_j$.

(b) $V = \displaystyle\sum_{j=1}^{m} W_j \quad \wedge \quad \forall\, z \in V : \exists\,!(u_j)_{j=1}^m \in \prod_{j=1}^{m} W_j : z = \sum_{j=1}^{m} u_j$.

(c) $$\begin{cases} V = \displaystyle\sum_{j=1}^{m} W_j \\ \forall\,(u_j)_{j=1}^m \in \displaystyle\prod_{j=1}^{m} W_j : \left(\sum_{j=1}^{m} u_j = \theta_V \Rightarrow \forall\, j \in \{1, \dots, m\} : u_j = \theta_V\right) \end{cases}.$$

PROOF: revisar documento disponible en INFODA.

Definición: Sea V un $\mathbb{K}$-espacio vectorial. Una familia $\{W_j\}_{j=1}^m$ de subespacios vectoriales de V, se llama una DESCOMPOSICIÓN INVARIANTE DE V con respecto a $T \in \mathcal{L}(V)$ si

$$\forall j \in \{1, ..., m\} \,:\, T(W_j) \subseteq W_j \quad \wedge \quad V = \bigoplus_{j=1}^{m} W_j \,.$$

Ejemplo: Sea V un $\mathbb{K}$-espacio vectorial finito dimensional, y $T \in \mathcal{L}(V)$ diagonalizable, con $\sigma(T) = \{\lambda_j\}_{j=1}^m$ (conjunto de valores propios distintos). Entonces $\{S_{\lambda_j}\}_{j=1}^m$ es una DESCOMPOSICIÓN INVARIANTE de V con respecto a T. En otras palabras:

$$\forall j \in \{1, ..., s\} \,:\, T(S_{\lambda_j}) \subseteq S_{\lambda_j} \quad \wedge \quad V = \bigoplus_{j=1}^{m} S_{\lambda_j} = \bigoplus_{j=1}^{m} \mathrm{Ker}(T - \lambda_j \tilde{I}) \,.$$

¿Qué pasa si T no es diagonalizable?

Veremos que **siempre** es posible determinar una descomposición invariante no trivial de V con respecto a T.

Ejemplo: Sea V un $\mathbb{K}$-espacio vectorial, y $T \in \mathcal{L}(V)$. Consideremos W un subespacio vectorial de V de dimensión uno, tal que W es invariante con respecto a T. Caracterizar W.

Sea $z \in V \backslash \{\theta\}$ tal que $W = \langle\{z\}\rangle$. Como $T(W) \subseteq W$, resulta que $T(z) \in W$. Esto implica que $\exists\, \alpha \in \mathbb{K}$ tal que $T(z) = \alpha\, z$, lo cual nos dice que (α, z) es un autopar de T. En consecuencia $W = S_\alpha$, i.e. W es un subespacio propio de T.

Ejemplo: Sea $T \in \mathcal{L}(\mathbb{R}^2)$ definida por $\forall\,(x,y) \in \mathbb{R}^2 \,:\, T(x,y) := (2x+y, 2y)$. Determine los subespacios vectoriales de $\mathbb{R}^2$ que son invariantes con respecto a T.

Desarrollo: Si W es subespacio vectorial de $\mathbb{R}^2$, de dimensión uno, invariante con respecto a T, entonces por lo anterior, W es un subespacio propio de T, de dimensión uno. Esto motiva calcular los valores y vectores propios de T. Consideremos la base canónica $B := \{e_1, e_2\}$ de $\mathbb{R}^2$, con $e_1 = (1,0)$ y $e_2 = (0,1)$. Luego, la matriz asociada de T con respecto a B es $[T]_B^B := \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$. De aquí se deduce que $\sigma(T) = \sigma([T]_B^B) = \{2\}$. El espacio propio asociado a $\lambda = 2$ es

$$S_2 := \mathrm{Ker}([T]_B^B - 2I) = \cdots = \left\langle \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\} \right\rangle ,$$

el cual es de dimensión uno. En consecuencia, $W := \langle \{(1,0)\} \rangle$ es el único subespacio invariante de $\mathbb{R}^2$ con respecto a T, de dimensión uno. Por lo tanto, los subespacios invariantes de $\mathbb{R}^2$ con respecto a T son $\{(0,0)\}$, $W = \langle \{(1,0)\} \rangle$ y $V := \mathbb{R}^2$.

Observación: Sea V un $\mathbb{K}$-espacio vectorial de dimensión finita n, y $T \in \mathcal{L}(V)$. Suponga que $\{W_j\}_{j=1}^m$ es una descomposición invariante de V con respecto a T. Entonces, puede probarse que existe una base B de V tal que $[T]_B^B$ es una matriz diagonal por bloques.

PROPOSICIÓN: Sean V un $\mathbb{K}$-espacio vectorial de dimensión finita, $T \in \mathcal{L}(V)$, y W_1, W_2 dos subespacios de V. Entonces son equivalentes las siguientes afirmaciones:

1. $V = W_1 \oplus W_2$, con W_1, W_2 dos subespacios T-invariantes de V.
2. Existe una base B de V, tal que $[T]_B^B$ es *matriz diagonal por bloques*, i.e. $[T]_B^B = \begin{pmatrix} A_1 & \Theta \\ \Theta & A_2 \end{pmatrix}$.

Este resultado puede generalizarse incluso.

PROPOSICIÓN: Sean V un $\mathbb{K}$-espacio vectorial de dimensión finita, $T \in \mathcal{L}(V)$, y $\{W_j\}_{j=1}^m$ una familia finita de subespacios de V. Entonces son equivalentes las siguientes afirmaciones:

1. $V = W_1 \oplus W_2 \cdots \oplus W_m$, con $\{W_j\}_{j=1}^m$ una familia finita de subespacios T-invariantes de V.
2. Existe una base B de V, tal que $[T]_B^B$ es *matriz diagonal por bloques*, i.e.

$$[T]_B^B = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_m \end{pmatrix}.$$

Núcleos iterados

Definición: Sea $A \in \mathcal{M}_n(\mathbb{K})$, ($\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$), y $\lambda \in \sigma(A)$. Se define el núcleo iterado de A con respecto a λ, al subespacio vectorial de $\mathbb{K}^{n\times 1}$

$$E_j(\lambda) := \mathrm{Ker}(A - \lambda I_n)^j, \quad \forall j \in \mathbb{N}.$$

Convención: $E_0(\lambda) := \mathrm{Ker}(I_n) = \{\theta\}$. Nótese que $E_1(\lambda) := \mathrm{Ker}(A - \lambda I_n) = S_\lambda$.

Propiedades (elementales) de los núcleos iterados: Sea $A \in \mathcal{M}_n(\mathbb{K})$, ($\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$), y $\lambda \in \sigma(A)$. Entonces

1. $\forall j \in \mathbb{N} : E_j(\lambda) \subseteq E_{j+1}(\lambda)$.
2. Si para algún $l \in \mathbb{N}$, $E_l(\lambda) = E_{l+1}(\lambda)$, entonces $\forall m \geq l : E_m(\lambda) = E_l(\lambda)$.
3. $\forall j \in \mathbb{N} : E_j(\lambda)$ es invariante con respecto a A (i.e. $\forall v \in E_j(\lambda) : A v \in E_j(\lambda)$).

Demostración 1): Sea $j \in \mathbb{N}$, fija pero arbitraria, y $v \in E_j(\lambda) := \mathrm{Ker}(A - \lambda I_n)^j$, también fijo pero arbitrario. Esto significa que

$$\begin{aligned}(A - \lambda I_n)^j v = \theta &\Rightarrow (A - \lambda I_n)(A - \lambda I_n)^j v = (A - \lambda I_n)\theta = \theta \\ &\Rightarrow (A - \lambda I_n)^{j+1} v = \theta \quad \Rightarrow v \in \mathrm{Ker}(A - \lambda I_n)^{j+1} =: E_{j+1}(\lambda).\end{aligned}$$

Como $v \in E_j(\lambda)$ es fijo pero arbitrario, se infiere que $E_j(\lambda) \subseteq E_{j+1}(\lambda)$. En vista que $j \in \mathbb{N}$ es fija pero arbitraria, se concluye que $\forall j \in \mathbb{N} : E_j(\lambda) \subseteq E_{j+1}(\lambda)$.

Demostración 2): Probemos que $E_{l+1}(\lambda) = E_{l+2}(\lambda)$, por doble inclusión. Por la propiedad 1), tenemos que $E_{l+1}(\lambda) \subseteq E_{l+2}(\lambda)$.
Resta probar entonces que $E_{l+2}(\lambda) \subseteq E_{l+1}(\lambda)$. Sea $v \in E_{l+2}(\lambda)$, fijo pero arbitrario. Esto significa que

$$(A - \lambda I_n)^{l+2} v = \theta \quad \Leftrightarrow (A - \lambda I_n)^{l+1} \underbrace{(A - \lambda I_n) v}_{=: w} = \theta .$$

Esto motiva introducir $w := (A - \lambda I_n) v \in E_{l+1}(\lambda) = E_l(\lambda)$, por hipótesis. Esto implica que $(A - \lambda I_n)^l w = \theta$, es decir $(A - \lambda I_n)^{l+1} v = \theta$, de donde se infiere que $v \in E_{l+1}(\lambda)$.
De esta manera, se prueba que $\forall v \in \mathbb{K}^{n\times 1} : v \in E_{l+2}(\lambda) \Rightarrow v \in E_{l+1}(\lambda)$, i.e.
$E_{l+2}(\lambda) \subseteq E_{l+1}(\lambda)$, y por la doble inclusión se concluye que $E_{l+2}(\lambda) = E_{l+1}(\lambda) = E_l(\lambda)$.
Finalmente, aplicando el Principio de Inducción Matemática (SE DEJA AL LECTOR) se demuestra que $\forall k \in \mathbb{N} : E_{l+k}(\lambda) = E_l(\lambda)$, y concluimos la demostración.

Demostración 3): Sea $j \in \mathbb{N}$, fijo pero arbitrario. Hay que probar que
$\forall v \in E_j(\lambda) : A v \in E_j(\lambda)$.
Sea $v \in E_j(\lambda)$, fijo pero arbitrario. Tenemos $A v = (A - \lambda I_n) v + \lambda v$, de donde para concluir el resultado sólo necesitamos probar que $(A - \lambda I_n) v \in E_j(\lambda)$. En efecto,

$$(A - \lambda I_n)^j (A - \lambda I_n) v = (A - \lambda I_n)^{j+1} v = (A - \lambda I_n) \underbrace{(A - \lambda I_n)^j v}_{= \theta} = \theta ,$$

de donde se infiere que $(A - \lambda I_n) v \in E_j(\lambda)$, y por tanto
$A v = (A - \lambda I_n) v + \lambda v \in E_j(\lambda)$.
De esta manera, se concluye que $\forall j \in \mathbb{N} : E_j(\lambda)$ es invariante.

Otras propiedades de los núcleos iterados

Sea $A \in \mathcal{M}_n(\mathbb{K})$, ($\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$), y $\lambda \in \sigma(A)$. Entonces

4) $\forall\, j \in \mathbb{N} \,:\, \dim(E_{j+1}(\lambda)) - \dim(E_j(\lambda)) \,\leq\, \dim(E_j(\lambda)) - \dim(E_{j-1}(\lambda))$.

5) Sea $\sigma(A) = \{\lambda_1, \ldots, \lambda_k\}$. Entonces, $\forall \{j_1, \ldots, j_k\} \subseteq \mathbb{N} \,:\, E_{j_1}(\lambda_1) \oplus \ldots \oplus E_{j_k}(\lambda_k)$, i.e.

$$\forall\, l, m \in \{1, \ldots, k\}, l \neq m \,:\, E_{j_l}(\lambda_l) \cap \sum_{m \in \{1,\ldots,k\} \setminus \{l\}} E_{j_m}(\lambda_m) = \{\theta\}.$$

6) La dimensión máxima de los núcleos iterados de A con respecto a un valor propio λ, es la multiplicidad algebraica m_λ de λ, i.e. $\forall\, \lambda \in \sigma(A) \,:\, \exists\, l \in \mathbb{N} \,:\, E_l(\lambda) = E_{l+1}(\lambda)$ y $\dim(E_l(\lambda)) = m_\lambda$.

Demostración 4): Sea $j \in \mathbb{N}$, fijo pero arbitrario. Por la propiedad 1), se sabe que $E_{j-1}(\lambda) \subseteq E_j(\lambda) \subseteq E_{j+1}(\lambda)$

$$\Rightarrow \dim(E_{j-1}(\lambda)) \leq \dim(E_j(\lambda)) \leq \dim(E_{j+1}(\lambda))$$

$$\Rightarrow \dim(E_j(\lambda)) - \dim(E_{j-1}(\lambda)) \geq 0 \wedge \dim(E_{j+1}(\lambda)) - \dim(E_j(\lambda)) \geq 0.$$

Probaremos que si $E_{j-1}(\lambda) \subset E_j(\lambda) \subset E_{j+1}(\lambda)$ (inclusión estricta), entonces

$$0 < \dim(E_{j+1}(\lambda)) - \dim(E_j(\lambda)) \leq \dim(E_j(\lambda)) - \dim(E_{j-1}(\lambda)).$$

Consideremos bases de los espacios involucrados (estamos en dimensión finita):

$$B_{E_{j-1}(\lambda)} := \{v_1, \ldots, v_k\} \text{ base de } E_{j-1}(\lambda),$$
$$B_{E_j(\lambda)} := \{v_1, \ldots, v_k, v_{k+1}, \ldots, v_{k+l}\} \text{ base de } E_j(\lambda),$$
$$B_{E_{j+1}(\lambda)} := \{v_1, \ldots, v_k, v_{k+1}, \ldots, v_{k+l}, v_{k+l+1}, \ldots, v_{k+l+m}\} \text{ base de } E_{j+1}(\lambda),$$

Definimos ahora $\forall\, i \in \{1, ..., m\} \ : \ w_i \ := \ (A - \lambda\, I_n)\, v_{k+l+i}$.

AFIRMACIÓN 1: $\forall\, i \in \{1, ..., m\} \ : \ w_i \in E_j(\lambda) \backslash E_{j-1}(\lambda)$.

Demostración: Sea $i \in \{1, ..., m\}$ cualquiera. Entonces,

$$(A - \lambda\, I_n)^j\, w_i \ = \ (A - \lambda\, I_n)^j\, (A - \lambda\, I_n)\, v_{k+l+i} \ = \ (A - \lambda\, I_n)^{j+1} \underbrace{v_{k+l+i}}_{\in E_{j+1}(\lambda)} \ = \theta \ \Rightarrow \ w_i \in E_j(\lambda)\,.$$

Probemos ahora que $w_i \notin E_{j-1}(\lambda)$. Por REDUCCIÓN AL ABSURDO, supongamos que $w_i \in E_{j-1}(\lambda)$, i.e.

$$\begin{aligned} (A - \lambda\, I_n)^{j-1}\, w_i \ = \ \theta \ &\Rightarrow (A - \lambda\, I_n)^{j-1}\, (A - \lambda\, I_n)\, v_{k+l+i} \ = \ \theta \\ &\Rightarrow (A - \lambda\, I_n)^j\, v_{k+l+i} \ = \ \theta \ \Rightarrow \ v_{k+l+i} \in E_j(\lambda)\,, \end{aligned}$$

lo cual implica que v_{k+l+i} es c.l. de $B_{E_j(\lambda)}$. Esto conduce a afirmar entonces que $B_{E_{j+1}(\lambda)}$ es l.d. $(\rightarrow\leftarrow)$. De esta manera, sucede que $w_i \notin E_{j-1}(\lambda)$, y se concluye la afirmación.

AFIRMACIÓN 2: $\{w_i\}_{i=1}^m$ ES L.I.

Demostración: Consideremos la combinación lineal para el vector nulo:

$$\begin{aligned} \sum_{i=1}^m \alpha_i\, w_i \ = \ \theta \ &\Rightarrow (A - \lambda\, I_n) \sum_{i=1}^m \alpha_i\, v_{k+l+i} \ = \ \theta \\ &\Rightarrow \sum_{i=1}^m \alpha_i\, v_{k+l+i} \in E_1(\lambda) \ \cap \ \langle \{v_{k+l+i}\}_{i=1}^m \rangle \ \subseteq \ E_j(\lambda) \ \cap \ \langle \{v_{k+l+i}\}_{i=1}^m \rangle \ = \ \{\theta\} \\ &\Rightarrow \sum_{i=1}^m \alpha_i\, v_{k+l+i} \ = \ \theta \ \Rightarrow \ \forall\, i \in \{1, ..., m\} \ : \ \alpha_i \ = \ 0 \ \Rightarrow \ \{w_i\}_{i=1}^m \ \text{es l.i.} \end{aligned}$$

CONSECUENCIA: $\forall\, i \in \{1, \dots, m\} \,:\, w_i \in \langle\{v_{k+1}, \dots, v_{k+l}\}\rangle$. Luego,

$$\langle\{w_1, \dots, w_m\}\rangle \subseteq \langle\{v_{k+1}, \dots, v_{k+l}\}\rangle \Rightarrow \underbrace{\dim(\langle\{w_1, \dots, w_m\}\rangle)}_{=m} \leq \underbrace{\dim(\langle\{v_{k+1}, \dots, v_{k+l}\}\rangle)}_{=l}$$

$$\Rightarrow 0 < \dim(E_{j+1}(\lambda)) - \dim(E_j(\lambda)) = m \leq l = \dim(E_j(\lambda)) - \dim(E_{j-1}(\lambda)) .$$

...los otros casos posibles:

a) Si $E_{j-1}(\lambda) = E_j(\lambda)$, entonces, por propiedad vista antes, $E_j(\lambda) = E_{j+1}(\lambda)$, y la propiedad se valida.

b) Si $E_{j-1}(\lambda) \subset E_j(\lambda) = E_{j+1}(\lambda)$, la propiedad también es cierta.

Observación: Si $E_{j-1}(\lambda) \subset E_j(\lambda)$ (inclusión estricta), entonces $\exists\, \{v_{k+p}\}_{p=1}^{l} \subseteq E_j(\lambda) \backslash E_{j-1}(\lambda)$ tal que $E_j(\lambda) = E_{j-1}(\lambda) \bigoplus \langle\{v_{k+p}\}_{p=1}^{l}\rangle$.

Para demostrar la Propiedad 5), necesitamos introducir la noción de funciones polinomiales matriciales. Dada la función polinomial (escalar) $p \in \mathcal{P}_k(\mathbb{K})$, caracterizada por $\mathbb{K} \ni x \mapsto p(x) := \sum_{j=0}^{k} \alpha_j x^j$, con $\{\alpha_j\}_{j=0}^{k} \subseteq \mathbb{K}$, éste induce la función (polinomial) matricial $p : \mathcal{M}_n(\mathbb{K}) \to \mathcal{M}_n(\mathbb{K})$ dado por $p(A) := \sum_{j=0}^{k} \alpha_j A^j$, para cualquier $A \in \mathcal{M}_n(\mathbb{K})$. Tener presente que $A^0 := I_n$.

Con esta noción, invocamos ahora el siguiente resultado técnico, cuya demostración sigue el mismo espíritu del resultado conocido en $\mathbb{Z}$.

Lema 1. Dados $p, q \in \mathcal{P}(\mathbb{K}) \backslash \{\theta\}$, existen $r, s \in \mathcal{P}(\mathbb{K})$ tal que $\forall\, x \in \mathbb{K} : r(x)\, p(x) + s(x)\, q(x) = \mathrm{m.c.d.}\{p, q\}(x)$.

Aquí, $\mathrm{m.c.d.}\{p, q\}$ denota el polinomio máximo común divisor de p y q.

Como consecuencia, se puede establecer que

Lema 2. Sean $A \in \mathcal{M}_n(\mathbb{K})$, y $p, q \in \mathcal{P}(\mathbb{K})$ tales que para algún vector $v \in \mathbb{K}^{n\times 1}$ se cumple: $p(A)\, v = \theta$, y $q(A)\, v = \theta$. Entonces, $d(A)\, v = \theta$, donde $d := \mathrm{m.c.d.}\{p, q\}$.

Demostración: Aplicando el Lema 1, tenemos que existen $r, s \in \mathcal{P}(\mathbb{K})$ tal que $\forall\, x \in \mathbb{K} : r(x)\, p(x) + s(x)\, q(x) = d(x)$. Esto induce la igualdad matricial:

$$r(A)\, p(A) + s(A)\, q(A) = d(A) \Rightarrow d(A)\, v = r(A)\, p(A)\, v + s(A)\, q(A)\, v = \theta\,.$$

Corolario: si en el Lema 2, además se tiene que $p, q \in \mathcal{P}(\mathbb{K})$ son primos relativos entre sí (i.e. $\mathrm{m.c.d.}\{p, q\} = 1$), entonces se concluye que $v = \theta$.

Demostración de Propiedad 5): Consideraremos el caso de dos Núcleos Iterados asociados a valores propios distintos de A.

Sean $j_1\,, j_2 \in \mathbb{N}$, $\lambda_1\,, \lambda_2 \in \sigma(A)$. Por probar que $\lambda_1 \neq \lambda_2 \Rightarrow E_{j_1}(\lambda_1) \cap E_{j_2}(\lambda_2) = \{\theta\}$. Sea $v \in E_{j_1}(\lambda_1) \cap E_{j_2}(\lambda_2)$, es decir $(A - \lambda_1\, I_n)^{j_1}\, v = \theta$ y $(A - \lambda_2\, I_n)^{j_2}\, v = \theta$. Esto equivale a expresarlo como $p(A)\, v = \theta$ y $q(A)\, v = \theta$, donde $\forall\, x \in \mathbb{K} : \; p(x) := (x - \lambda_1)^{j_1}$, $\forall\, x \in \mathbb{K} : \; q(x) := (x - \lambda_2)^{j_2}$. Así, p y q son polinomios primos relativos entre sí. Luego, aplicando el Lema 2, se infiere que $v = \theta$. Finalmente, se concluye que $E_{j_1}(\lambda_1) \cap E_{j_2}(\lambda_2) = \{\theta\}$.

Demostración de Propiedad 6): Primero, recordemos que dado $\lambda \in \sigma(A)$, se cumple que $\forall\, j \in \mathbb{N} \,:\, E_j(\lambda) \subseteq \mathbb{K}^{n\times 1} \wedge E_j(\lambda) \subseteq E_{j+1}(\lambda)$. Esto garantiza la existencia de $l \in \mathbb{N}$ tal que $E_l(\lambda) = E_{l+1}(\lambda)$.

Probemos ahora que $\dim(E_l(\lambda)) = m_\lambda$. Se demostrará por doble desigualdad.

Afirmación 1: $\forall\, j \in \mathbb{N} \,:\, \dim(E_j(\lambda)) \leq m_\lambda$. Sea $j \in \mathbb{N}$. Definimos $k := \dim(E_j(\lambda))$. Esto implica que existe una base $B_k := \{v_1, ..., v_k\}$ de $E_j(\lambda) \subseteq \mathbb{K}^{n\times 1}$. Se puede completar B_k para inducir una base $\tilde{B} := \{v_1, ..., v_k, v_{k+1}, ..., v_n\}$ de $\mathbb{K}^{n\times 1}$. Esto permite definir la aplicación lineal $T_A : \mathbb{K}^{n\times 1} \to \mathbb{K}^{n\times 1}$ tal que $\forall\, x \in \mathbb{K}^{n\times 1} \,:\, T_A(x) := A\,x$. Considerando $B := \{e_j\}_{j=1}^n$, base canónica de $\mathbb{K}^{n\times 1}$, se tiene que $[T_A]_B^B = A$. A su vez, se sabe que $[T_A]_B^B \sim [T_A]_{\tilde{B}}^{\tilde{B}}$. Esto significa que $\exists\, P := \left[\tilde{I}\right]_{\tilde{B}}^{B} \in \mathcal{M}_n(\mathbb{K})$ (matriz de paso o de cambio de base), tal que $[T_A]_{\tilde{B}}^{\tilde{B}} = P^{-1}\,A\,P$.

Recordamos que

$$[T_A]_{\tilde{B}}^{\tilde{B}} := \left([T_A(v_1)]_{\tilde{B}} \,|\, \cdots \,|\, [T_A(v_n)]_{\tilde{B}}\right) = \left([A\,v_1]_{\tilde{B}} \,|\, \cdots \,|\, [A\,v_n]_{\tilde{B}}\right) .$$

Como $E_j(\lambda)$ es invariante con respecto a A, resulta $\forall\, i \in \{1, ..., k\} \,:\, A\,v_i \in E_j(\lambda)$, es decir $\forall\, i \in \{1, ..., k\} \,:\, A\,v_i \in \langle\{v_1, ..., v_k\}\rangle$. Esto permite afirmar que

$$\tilde{A} := [T_A]_{\tilde{B}}^{\tilde{B}} = \begin{pmatrix} A_k & C \\ \Theta & B \end{pmatrix} , \quad \text{donde } A_k \in \mathcal{M}_k(\mathbb{K}) .$$

Veamos ahora que $\sigma(A_k) = \{\lambda\}$. Sea $(\mu, z) \in \mathbb{K} \times \mathbb{K}^{k\times 1} \backslash \{\theta\}$ un autopar de A_k, es decir $A_k\, z = \mu\, z$. Esto permite deducir que

$$P^{-1}\, A\, P \begin{pmatrix} z \\ \theta \end{pmatrix} = \tilde{A} \begin{pmatrix} z \\ \theta \end{pmatrix} = \begin{pmatrix} A_k & C \\ \Theta & B \end{pmatrix} \begin{pmatrix} z \\ \theta \end{pmatrix} = \begin{pmatrix} A_k\, z \\ \theta \end{pmatrix} = \mu \begin{pmatrix} z \\ \theta \end{pmatrix}.$$

Luego, el vector $w := P \begin{pmatrix} z \\ \theta \end{pmatrix} \in \mathbb{K}^{n\times 1}$ es tal que $A\, w = \mu\, w$. Esto nos dice que $\mu \in \sigma(A)$ y $w \in S_\mu$. Por otro lado, $w = \sum_{i=1}^{k} z_i\, v_i \in \langle \{v_i\}_{i=1}^{k} \rangle := E_j(\lambda)$. En consecuencia, $w \in S_\mu \cap E_j(\lambda)$.

- Si $\mu \neq \lambda$, entonces $w \in S_\mu \cap E_j(\lambda) = E_1(\mu) \cap E_j(\lambda) = \{\theta\}$, lo cual implica $w = \theta$, por ende $z = \theta$ $(\rightarrow\leftarrow)$.
- Por lo tanto, se deduce así que $\mu = \lambda$, es decir $\sigma(A_k) = \{\lambda\}$, de multiplicidad k. Esto nos dice que $\forall\, x \in \mathbb{K} : p_{A_k}(x) = (x - \lambda)^k$.
- Como $\tilde{A}$ y A son semejantes, poseen el mismo polinomio característico. Esto significa que $\forall\, x \in \mathbb{K} : p_A(x) = p_{A_k}(x)\, p_B(x)$, lo cual implica que $\dim(E_j(\lambda)) = k \leq m_\lambda$.
- Siendo $j \in \mathbb{N}$ fijo, pero arbitrario, se concluye la AFIRMACIÓN 1.

Afirmación 2: Si $E_l(\lambda) = E_{l+1}(\lambda)$, entonces $\dim(E_l(\lambda)) \geq m_\lambda$. Sea $\tilde{k} := \dim(E_l(\lambda))$. Esto garantiza la existencia de una base $B_{\tilde{k}} := \{u_1, ..., u_{\tilde{k}}\}$ de $E_l(\lambda)$, la cual puede extenderse a una base $\widetilde{B} := \{u_1, ..., u_{\tilde{k}}, u_{\tilde{k}+1}, ...u_n\}$ de $\mathbb{K}^{n\times 1}$. Así, procediendo como en la demostración de la Afirmación 1, se tiene la existencia de una matriz $Q \in \mathcal{M}_n(K)$ no singular, tal que

$$Q^{-1}\, A\, Q = \begin{pmatrix} A_{\tilde{k}} & \widetilde{C} \\ \Theta & \widetilde{B} \end{pmatrix} =: \widetilde{A}, \quad \text{donde } A_{\tilde{k}} \in \mathcal{M}_{\tilde{k}}(\mathbb{K}).$$

En lo que sigue, probaremos que $\widetilde{B}$ no tiene el valor propio λ. Para esto

a) probar (por Inducción Matemática) que
$\forall\, j \in \mathbb{N} : \forall\, u \in \mathbb{K}^{n\times 1} : u \in \mathrm{Ker}(\widetilde{A} - \lambda\, I_n)^j \Leftrightarrow Q\, u \in \mathrm{Ker}(A - \lambda\, I_n)^j$.

b) Así, $Q\, u \in \mathrm{Ker}(A - \lambda\, I_n)^l = E_l(\lambda) \Leftrightarrow Q\, u \in \langle\{u_i\}_{i=1}^{\tilde{k}}\rangle \Leftrightarrow u \in \langle\{e_i\}_{i=1}^{\tilde{k}}\rangle$, donde $\{e_i\}_{i=1}^{\tilde{k}}$ son los primeros $\tilde{k}$ vectores canónicos de $\mathbb{K}^{n\times 1}$.

c) Esto permite inferir que
$\forall\, i \in \{1, ..., \tilde{k}\} : Q\, e_i \in \langle\{u_i\}_{i=1}^{\tilde{k}}\rangle = E_l(\lambda) = \mathrm{Ker}(A - \lambda\, I_n)^l$. Por la parte a), se tiene que $Q\, e_i \in \mathrm{Ker}(A - \lambda\, I_n)^l \Leftrightarrow e_i \in \mathrm{Ker}(\widetilde{A} - \lambda\, I_n)^l$. Como resultado, se desprende que $\forall\, i \in \{1, ..., \tilde{k}\} : (\widetilde{A} - \lambda\, I_n)^l\, e_i = \theta$, lo cual indica que las primeras $\tilde{k}$ columnas de la matriz $(\widetilde{A} - \lambda\, I_n)^l$ son nulas. Como consecuencia, se tiene

$$(\widetilde{A} - \lambda\, I_n)^l = \begin{pmatrix} (A_{\tilde{k}} - \lambda\, I_{\tilde{k}})^l & \widehat{C} \\ \Theta & (\widetilde{B} - \lambda\, I_{n-\tilde{k}})^l \end{pmatrix} = \begin{pmatrix} \Theta & \widehat{C} \\ \Theta & (\widetilde{B} - \lambda\, I_{n-\tilde{k}})^l \end{pmatrix}.$$

d) Ahora, suponemos que $(\lambda, y) \in \mathbb{K} \times \mathbb{K}^{(n-\tilde{k})\times 1} \backslash \{\theta\}$ es un autopar de $\widetilde{B}$, es decir $(\widetilde{B} - \lambda\, I_{n-\tilde{k}})\, y \,=\, \theta$. Entonces

$$\begin{aligned}(\widetilde{A} - \lambda\, I_n)^{2l} \begin{pmatrix}\theta \\ y\end{pmatrix} &= \begin{pmatrix}\Theta & \widehat{C} \\ \Theta & (\widetilde{B} - \lambda\, I_{n-\tilde{k}})^l\end{pmatrix} \begin{pmatrix}\Theta & \widehat{C} \\ \Theta & (\widetilde{B} - \lambda\, I_{n-\tilde{k}})^l\end{pmatrix} \begin{pmatrix}\theta \\ y\end{pmatrix} \\ &= \begin{pmatrix}\Theta & \widehat{C} \\ \Theta & (\widetilde{B} - \lambda\, I_{n-\tilde{k}})^l\end{pmatrix} \begin{pmatrix}\widehat{C}\, y \\ (\widetilde{B} - \lambda\, I_{n-\tilde{k}})^l\, y\end{pmatrix} \\ &= \begin{pmatrix}\Theta & \widehat{C} \\ \Theta & (\widetilde{B} - \lambda\, I_{n-\tilde{k}})^l\end{pmatrix} \begin{pmatrix}\widehat{C}\, y \\ \theta\end{pmatrix} = \begin{pmatrix}\theta \\ \theta\end{pmatrix},\end{aligned}$$

lo cual implica que $w \,:=\, \begin{pmatrix}\theta \\ y\end{pmatrix} \in \mathrm{Ker}(\widetilde{A} - \lambda\, I_n)^{2l} \Leftrightarrow\, Q\, w \in \mathrm{Ker}(A - \lambda\, I_n)^{2l} \,=\, \mathrm{Ker}(A - \lambda\, I_n)^{l} \,=\, \langle \{u_i\}_{i=1}^{\tilde{k}} \rangle$, de donde se deduce que $w \in \langle \{Q^{-1}\, u_i\}_{i=1}^{\tilde{k}} \rangle \,=\, \langle \{e_i\}_{i=1}^{\tilde{k}} \rangle$ (Q es matriz de cambio de base). Esto implica que $y \,=\, \theta$ $(\rightarrow\leftarrow)$.

e) Lo anterior permite afirmar que $\lambda \notin \sigma(\widetilde{B})$. Así, $\forall\, x \in \mathbb{K} \,:\, p_A(x) \,=\, p_{A_{\tilde{k}}}(x)\, p_{\widetilde{B}}(x)$, con $p_{\widetilde{B}}(\lambda) \neq 0$, de donde se concluye que $\dim(E_l(\lambda)) \,=\, \tilde{k} \,=\, m_\lambda$.

Ejemplos: Aplique los núcleos iterados para mostrar que las siguientes matrices son semejantes a otra matriz casi diagonal. Indique la matriz de semejanza.

$$a)\ A = \begin{pmatrix} 0 & 2 \\ -2 & 4 \end{pmatrix}, \quad b) \qquad A = \begin{pmatrix} 0 & 3 & 1 \\ 2 & -1 & -1 \\ -2 & -1 & -1 \end{pmatrix}$$

Teorema. Sea $A \in \mathcal{M}_n(\mathbb{C})$ y $\sigma(A) = \{\lambda_1, ..., \lambda_k\} \subseteq \mathbb{C}$ (valores propios distintos). Entonces, $\exists \{j_1, ..., j_k\} \subseteq \mathbb{N}$ tal que

$$\mathbb{C}^{n\times 1} = \bigoplus_{s=1}^{k} E_{j_s}(\lambda_s) := E_{j_1}(\lambda_1) \oplus ... \oplus E_{j_k}(\lambda_k),$$

donde $\forall\, s \in \{1, ..., k\} : E_{j_s}(\lambda_s) = E_{j_s+1}(\lambda_s)$ y $\dim(E_{j_s}(\lambda_s)) = m_{\lambda_s}$.
Demostración: es consecuencia inmediata de las propiedades anteriores, considerando además que $\sum_{s=1}^{k} m_{\lambda_s} = n$. Se deja al lector hacer los detalles.

Corolario: $A \in \mathcal{M}_n(\mathbb{C})$ es diagonalizable si y sólo si $\forall\, \lambda \in \sigma(A) : \dim(S_\lambda) = m_\lambda$.
Demostración: También se deja al lector hacer los detalles de la prueba.

Observación: El teorema previo sigue siendo válido en el cuerpo real, i.e. $A \in \mathcal{M}_n(\mathbb{R})$ con $\sigma(A) = \{\lambda_1, ..., \lambda_k\} \subseteq \mathbb{R}$, si además se verifica que $\sum_{s=1}^{k} m_{\lambda_s} = n$. En este caso se cumple que

$$\mathbb{R}^{n\times 1} = \bigoplus_{s=1}^{k} E_{j_s}(\lambda_s).$$

Observaciones sobre las dimensiones de los núcleos iterados

Denotando, para $j \geq 0$, $n_j := \dim(E_j(\lambda)) = \dim(\mathrm{Ker}(A - \lambda\, I_n)^j)$, las tres propiedades anteriores implican que esta sucesión de enteros $\{n_j\}_{j\geq 0}$ cumple las tres siguientes propiedades:

$$
\begin{aligned}
&(1)\ 0 = n_0 < n_1 < \ldots . < n_l = n_{l+1} = \ldots ,\\
&(2)\ n_l - n_{l-1} \leq n_{l-1} - n_{l-2} \leq \ldots \leq n_2 - n_1 \leq n_1 - n_0 = n_1 ,\\
&(3)\ n_l = m_\lambda .
\end{aligned}
$$

La sucesión de núcleos iterados queda entonces:

$$
\begin{array}{ccccccccccccc}
 & E_0(\lambda) & \subset & E_1(\lambda) & \subset & E_2(\lambda) & \subset & \ldots . & \subset & E_l(\lambda) & = & E_{l+1}(\lambda) & = & \ldots \\
\dim : & 0 = n_0 & < & n_1 & < & n_2 & < & \ldots & < & n_l & = & n_{l+1} & = & \ldots
\end{array}
$$

Recapitulando

1. Sea $A \in \mathcal{M}_n(\mathbb{K})$ (con $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$), y S un subespacio de $\mathbb{K}^{n\times 1}$. Se dice que S es invariante por A si $\forall\, u \in S \,:\, A\,u \in S$.
2. Sea $A \in \mathcal{M}_n(\mathbb{K})$ (con $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$), semejante a $B \in \mathcal{M}_n(\mathbb{K})$, con matriz de semejanza P, tal que $B = P^{-1}\,A\,P$. Entonces
$$\forall\, j \in \mathbb{N} \,:\, \forall\, \lambda \in \sigma(A) \,:\, \forall\, u \in \mathbb{K}^{n\times 1} \,:\, \Big(u \in \mathrm{Ker}((B-\lambda\, I_n)^j) \Leftrightarrow P\,u \in \mathrm{Ker}((A-\lambda\, I_n)^j)\Big)\,.$$
Por tanto, las dimensiones de los núcleos iterados son invariantes por semejanza.
3. Sea $A \in \mathcal{M}_n(\mathbb{K})$ (con $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$). Si $\mathbb{K}^{n\times 1} = S \oplus T$, con S y T invariantes por A, decimos que $\mathbb{K}^{n\times 1} = S \oplus T$ es una descomposición invariante por A. Esta definición se extiende a $\mathbb{K}^{n\times 1} = \bigoplus_{j=1}^{k} S_j$, donde para cada $j \in \{1, \ldots, k\}$, S_j es invariante por A.
4. Si $\mathbb{K}^{n\times 1} = S \oplus T$ es una descomposición invariante por $A \in \mathcal{M}_n(\mathbb{K})$, consideramos una base de $\mathbb{K}^{n\times 1}$ construida de la forma
$$\underbrace{s_1\, \ldots, s_l}_{\text{base de } S}\,,\ \underbrace{t_{l+1}, \ldots, t_n}_{\text{base de } T}\,,$$
con la cual se construye la matriz $P := \left(s_1 \,|\, \cdots \,|\, s_l \,|\, t_{l+1} \,|\, \cdots \,|\, t_n\right) \in \mathcal{M}_n(\mathbb{K})$ no singular. Entonces A es semejante a una matriz diagonal por bloques, i.e.
$$P^{-1}\,A\,P = \begin{pmatrix} A_1 & \Theta \\ \Theta & A_2 \end{pmatrix}, \quad \text{donde } A_1 \in \mathcal{M}_l(\mathbb{K})\,,\ A_2 \in \mathcal{M}_{n-l}(\mathbb{K})\,.$$

Recapitulando...

Como resultado, las descomposiciones invariantes producen matrices diagonales por bloques. Nótese que si una matriz $A \in \mathcal{M}_n(\mathbb{K})$ es tal que existe una matriz $P \in \mathcal{M}_n(\mathbb{K})$ no singular, con

$$P^{-1}\,A\,P \,=\, \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_s \end{pmatrix}, \quad \text{donde } \forall\, j \in \{1, .., s\} \,:\, A_j \text{ es cuadrada, entonces}$$

1. la partición correspondiente de columnas de P induce una descomposición invariante de $\mathbb{K}^{n\times 1}$ en s subespacios,
2. el polinomio característico de A es el producto de los polinomios característicos de las matrices A_j, $j \in \{1, ..., s\}$.