Algebra Lineal: Aplicaciones a la Física, Notas Curso 2014

5. Transformaciones lineales

Una transformación lineal (TL) (también denominada aplicación lineal o mapeo lineal) es una función $F : V \to W$ entre dos espacios vectoriales V, W *sobre el mismo cuerpo* K que satisface

$$i)\ \ F(v_1 + v_2) = F(v_1) + F(v_2)\ \ \forall\, v_1, v_2 \in V$$

$$ii)\ \ F(\alpha v) = \alpha F(v)\ \ \forall\, \alpha \in K,\ v \in V$$

es decir, $F(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 F(v_1) + \alpha_2 F(v_2)\ \forall\ \alpha_1, \alpha_2 \in K$ y $v_1, v_2 \in V$. F es entonces un homomorfismo entre espacios vectoriales. Si $W = V$, F es un endomorfismo, y se lo denomina también *operador lineal.*

Propiedades básicas:
I) $F(0) = 0$ (el primer 0 es el vector nulo de V y el segundo el vector nulo de W)
Dem.: $F(0) = F(0 + 0) = F(0) + F(0) \Rightarrow F(0) = 0$
II) $F(-v) = -F(v)$
Dem.: $0 = F(0) = F(v + (-v)) = F(v) + F(-v)$, de donde $F(-v) = F(v)$.
(También pueden demostrarse utilizando ii))

Ejemplos (probar como ej.):
1) $F : V \to V$ definida por $F(v) = \alpha v$, con $\alpha \in K$, es TL $\forall \alpha \in K$ (incluso $\alpha = 0$).
Si $\alpha = 1 \Rightarrow F$ es la *identidad* ($F(v) = v\ \forall v \in V$), denotada por I.
Si $\alpha = 0 \Rightarrow F$ es la TL *nula* ($F(v) = 0\ \forall\ v \in V$).
2) $F : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $F(x, y) = (x + y, 2y + x)$ es TL (verificar).
También lo es $F : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $F(x, y) = (ax + by, cx + dy)$, con $a, b, c, d \in \mathbb{R}$ (fijos). Toda TL de $\mathbb{R}^2$ en $\mathbb{R}^2$ puede escribirse de esta forma (probar!).
3) $F : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $F(x, y) = (x + y^2, 2y + x)$ no es TL. Tampoco lo es $F(x, y) = (1 + x + y, 2y)$.
5) $F : \mathbb{R}^n \to \mathbb{R}^m$ dada por $F(x_1, \ldots, x_n) = (a_{11}x_1 + \ldots + a_{1n}x_n, \ldots, a_{m1}x_1 + \ldots + a_{mn}x_n)$ es TL $\forall\ a_{ij} \in \mathbb{R}$.
Toda TL de $\mathbb{R}^n$ en $\mathbb{R}^m$ puede escribirse de esta forma (demostrarlo!).
6) $F : \mathbb{R}^{2\times 2} \to \mathbb{R}^{2\times 2}$ dada por $F(A) = A^t$ (traspuesta) es TL.
También lo es $F(A) = B \cdot A$, con B matriz real fija de $n \times n$, y $G(A) = B \cdot A + A \cdot C$.
7) $F : C(\mathbb{R} \to C\mathbb{R}$ ($C(\mathbb{R})$ es el subespacio de funciones reales continuas) dada por $[F(f)](x) = \int_0^x f(t)dt$, es TL.
8) Si $C_1(\mathbb{R}) \subset \mathbb{R}^{\mathbb{R}}$ es el subespacio de funciones reales derivables, $F : C(\mathbb{R}) \to \mathbb{R}^{\mathbb{R}}$ dada por $F(f) = f'$, es también TL.

Ppropiedades fundamentales:
III) Si S es un subespacio de V y $F : V \to W$ es TL, la imagen de S por f, $F(S) = \{F(v)|v \in S\}$, es un subespacio de W.
En particular, si $S = V$, el subespacio $F(V)$ se denomina *imagen* de F y se denota $I(F)$.
Dem.: Si $v_1, v_2 \in S \Rightarrow F(v_1) + F(v_2) = F(v_1 + v_2) \in F(S)$ pues $v_1 + v_2 \in S$.
Si $\alpha \in K$ y $v \in S \Rightarrow \alpha F(v) = F(\alpha v) \in F(S)$ pues $\alpha v \in S$.
En particular, $0 \in F(S)$ ($F(0) = 0$)

IV) La pre-imagen de un subespacio S' de W, $S = \{v \in V|F(v) \in S'\}$ es un subespacio de V.
Dem.: $0 \in S$ pues $F(0) = 0 \in S'$.
Si $v_1, v_2 \in S \Rightarrow v_1 + v_2 \in S$ pues $F(v_1 + v_2) = F(v_1) + F(v_2) \in S'$
Si $\alpha \in K$ y $v \in S \Rightarrow \alpha v \in S$ pues $F(\alpha v) = \alpha F(v) \in S'$.
En particular, la pre-imagen de $\{0\}$ (subespacio nulo de W) se denomina **núcleo** o **espacio nulo** de F y se denotará como $N(F)$: $N(F) = \{v \in V|F(v) = 0\}$.

V) Si $v = \sum_{i=1}^m \alpha_i v_i$, con $\alpha_i \in K$ y $v_i \in V \Rightarrow F(v) = F(\sum_{i=1}^m \alpha_i v_i) = \sum_{i=1}^m \alpha_i F(v_i)$
Esto puede demostrarse fácilmente por inducción (para los que no lo ven obvio).
Esta propiedad implica que la imagen del subespacio $\overline{C}$ generado por un subconjunto de vectores $C = \{v_1, \ldots, v_m\} \subset V$ es el subespacio $\overline{F(C)}$ generado por la imagen $F(C) = \{F(v_1), \ldots, F(v_m)\} \subset W$:

$$F(\overline{C}) = \overline{F(C)}$$

En particular, si V es finitamente generado y $B=\{b_1,\ldots,b_n\}$ es una base de V, $I(F)=F(\overline{B})=\overline{F(B)}$. En otras palabras, la imagen $I(F)$ es el subespacio generado por los n vectores $F(b_1),\ldots,F(b_n)$ de W. Esto implica que una transformación lineal **queda completamente determinada por los vectores que asigna a los elementos de una base**, es decir, por los n vectores $F(b_1),\ldots,F(b_n)$:
Si $v\in V \Rightarrow v=\sum_{i=1}^n \alpha_i b_i$ y $F(v)=\sum_{i=1}^n \alpha_i F(b_i)$.
Nótese también que si $v_1,\ldots,v_m$ son LD $\Rightarrow$ los vectores $F(v_1),\ldots,F(v_m)$ son también LD: Si $0=\sum_{i=1}^m \alpha_i v_i$, con algún $\alpha_i\neq 0 \Rightarrow 0=F(\sum_{i=1}^m \alpha_i v_i)=\sum_{i=1}^m \alpha_i F(v_i)$.

VI) Si V es finitamente generado entonces

$$\dim N(F)+\dim I(F)=\dim V$$

Dem.: Sea $\{b_1,\ldots,b_m,b_{m+1},\ldots,b_n\}$ base de V tal que $\{b_1,\ldots,b_m\}$ sea base de $N(F)$ ($F(b_i)=0$ si $i\leq m$). Si $v=\sum_{i=1}^n \alpha_i b_i\in V \Rightarrow$

$$F(v)=\sum_{i=1}^n \alpha_i F(b_i)=\sum_{i=m+1}^n \alpha_i F(b_i)$$

pertenece al espacio generado por $F(b_{m+1})\ldots,F(b_n)$. Además, $F(b_{m+1}),\ldots,F(b_n)$ son LI pues si

$$0=\sum_{i=m+1}^n \alpha_i F(b_i)=F(\sum_{i=m+1}^n \alpha_i b_i)$$

el vector $\sum_{i=m+1}^n \alpha_i b_i\in N(F)$ y por tanto, $\sum_{i=m+1}^n \alpha_i b_i=\sum_{i=1}^m \alpha_i b_i$. Pero por independencia lineal de los b_i, debe ser $\alpha_i=0$ para $i=1,\ldots,n$, por lo que $F(b_{m+1}),\ldots,F(b_n)$ son LI.
La dimensión de la imagen es por lo tanto $n-m$, y se cumple entonces $\dim N(F)+\dim I(F)=m+(n-m)=n=\dim V$. La dimensión de la imagen $I(F)$ se denomina **rango** de F y la dimensión del espacio nulo $N(F)$ **nulidad** de F.

Ejemplos simples:
1) $F:V\rightarrow V$ dada por $F(v)=\alpha v$
Si $\alpha=0$, $N(F)=V$, $I(F)=\{0\}$. Si dim $V=n$, dim $N(F)$+dim $I(F)=n+0=n$.
Si $\alpha\neq 0$, $N(F)=\{0\}$, $I(F)=V$. Si dim $V=n$, dim $N(F)$+dim $I(F)=0+n=n$.
2) $F:\mathbb{R}^2\rightarrow\mathbb{R}^2$ dada por $F(x,y)=(x,0)$. $N(F)=\{(0,y)|y\in\Re\}$, $I(F)=\{(x,0)|x\in\Re\}$.
dim $N(F)$+dim $I(F)$=1+1=2=dim V.
3) $F:\mathbb{R}^{2\times 2}\rightarrow\mathbb{R}^{2\times 2}$ dada por $F(A)=A^t$. $N(F)=\{0\equiv(\begin{smallmatrix}0&0\\0&0\end{smallmatrix})\}$, $I(F)=\mathbb{R}^2$; dim $N(F)$+dim$I(F)$=0+4=4.

Propiedades Geométricas:
1) Una transformación lineal $F:\mathbb{R}^n\rightarrow\mathbb{R}^m$ transforma segmentos paralelos ($s_i=\{\mathbf{r}_i+t\mathbf{n},\ t\in[0,1]\}$, $i=1,2$) en segmentos paralelos ($F(s_i)=\{F(\mathbf{r}_i)+tF(\mathbf{n}),\ t\in[0,1]\}$, $i=1,2$). (Probar!). Por lo tanto, F transforma, por ej., un paralelogramo en un paralelogramo (que puede degenerar en un segmento o punto).

2) Una transformación lineal $F:V\rightarrow W$ transforma conjuntos convexos en conjuntos convexos (probar!, recordar dem. dada en clase). Recordemos que un conjunto C en un espacio V sobre $K=\mathbb{R}$ o $\mathbb{C}$ es convexo sii contiene el segmento que une dos puntos del conjunto: $u,v\in C\Rightarrow u+(v-u)t\in C\ \forall\ t\in[0,1]$.

5.1 Representación matricial de funciones lineales:

Sea $F:V\rightarrow W$, con V, W finitamente generados de dimensión n y m respect. Sea $B=(b_1,\ldots,b_n)$ una base ordenada de V y $B'=(b'_1,\ldots,b'_m)$ una base ordenada de W. Podemos escribir

$$F(b_i)=\sum_{j=1}^m T_{ji}b'_j,\ \ i=1,\ldots,n$$

donde $T_{ji}\in K$ son las coordenadas de $F(b_i)$ en la base B' de W. Por lo tanto, si $v\in V$, $v=\sum_{i=1}^n \alpha_i b_i$ y

$$F(v)=\sum_{i=1}^n \alpha_i F(b_i)=\sum_{i=1}^n \alpha_i\sum_{j=1}^m T_{ji}b'_j=\sum_{j=1}^m(\sum_{i=1}^n T_{ji}\alpha_i)b'_j=\sum_{j=1}^m \alpha'_j b'_j$$

con $\alpha'_j = \sum_{i=1}^n T_{ji}\alpha_i$, $j = 1, \ldots, m$. Es decir, en forma matricial,

$$\begin{pmatrix} \alpha'_1 \\ \ldots \\ \alpha'_m \end{pmatrix} = \begin{pmatrix} T_{11} & \ldots & T_{1n} \\ \ldots & \ldots & \ldots \\ T_{m1} & \ldots & T_{mn} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \ldots \\ \alpha_n \end{pmatrix},$$

que se escribe en forma concisa como

$$[F(v)]_{B'} = [F]_{B'}^B [v]_B$$

donde

$$[F(v)]_{B'} = \begin{pmatrix} \alpha'_1 \\ \ldots \\ \alpha'_m \end{pmatrix}, \quad [v]_B = \begin{pmatrix} \alpha_1 \\ \ldots \\ \alpha_n \end{pmatrix},$$

son las coordenadas de $F(v)$ en la base B' y de v en la base B, y

$$[F]_{B'}^B = \begin{pmatrix} T_{11} & \ldots & T_{1n} \\ \ldots & \ldots & \ldots \\ T_{m1} & \ldots & T_{mn} \end{pmatrix} = \begin{pmatrix} . & . & . \\ [F(b_1)]_{B'} & \ldots & [F(b_n)]_{B'} \\ . & . & . \end{pmatrix}$$

es la matriz de $m \times n$ que representa la transformación lineal F respecto de las bases B de V y B' de W. Esta matriz depende de las bases elegidas, pero una vez elegida las bases es claramente *única*, ya que la función lineal queda completamente determinada por los vectores que asigna a los elementos de una base. En particular, si $V = W$, podemos elegir $B' = B$, y

$$[F(v)]_B = [F]_B^B [v]_B$$

Notemos que la identidad $I : V \to V$ definida por $I(v) = v \ \forall \ v$ queda así representada por la matriz identidad I_n: $[I]_B^B = I_n \ \forall \ B$ de V. Por simplicidad, denotaremos a $[F]_B^B$ también como $[F]_B$ cuando quede claro que estamos trabajando con operadores lineales representados en una misma base.

Ejemplo 1: Sea $F : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $F(x, y) = (2x + y, 4y + 3x)$. En la base canónica $B = (e_1, e_2)$, $e_1 = (1, 0)$, $e_2 = (0, 1)$, tenemos $F(e_1) = (2, 3) = 2e_1 + 3e_2$, $F(e_2) = (1, 4) = e_1 + 4e_2$, y la matriz que representa a F en esta base es

$$[F]_B^B = \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix} \tag{1}$$

Ejemplo 2: Reflexión respecto del eje x en $\mathbb{R}^2$. Si $F(v)$ es el vector obtenido al reflejar v respecto del eje x, tenemos (recordar dibujo) $F(e_1) = e_1$, $F(e_2) = -e_2$ y por lo tanto

$$[[F]_B^B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2}$$

Ejemplo 3: Reflexión respecto de la recta de ec. $y = x$ en $\mathbb{R}^2$: Si $F(v)$ es el vector obtenido al reflejar v respecto de la recta $y = x$, tenemos $F(e_1) = e_2$, $F(e_2) = e_1$ y por lo tanto

$$[F]_B^B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{3}$$

Ejemplo 4: Rotación de ángulo θ en $\mathbb{R}^2$. Si $F(v)$ es el vector obtenido al rotar v un ángulo θ antihorario, tenemos (recordar dibujo) $F(e_1) = \cos(\theta)e_1 + \sin(\theta)e_2$, $F(e_2) = -\sin(\theta)e_1 + \cos(\theta)e_2$ y por lo tanto

$$[F]_B^B = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \tag{4}$$

Ejemplo 5: Rotación de ángulo θ en $\mathbb{R}^3$ alrededor del eje z. Si $F(v)$ es el vector obtenido al rotar v un ángulo θ antihorario alrededor del eje z, tenemos, en la base canónica de $\mathbb{R}^3$, $B = ((1, 0, 0), (0, 1, 0), (0, 0, 1))$, $F(b_1) = \cos(\theta)b_1 + \sin(\theta)b_2$, $F(b_2) = -\sin(\theta)b_1 + \cos(\theta)b_2$ y $F(b_3) = b_3$. Por lo tanto

$$[F]_B^B = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{5}$$

Ejemplo 6: Sea P_n el espacio de polinomios de grado $\leq n$, y sea $D : P_2 \to P_1$ el operador derivación restringido a polinomios de grado ≤ 2 con codominio P_1. Sea $B(1, t, t^2)$ la base "canónica" de P_2 y $(1, t)$ la de P_1. Tenemos $D(1) = 0$, $D(t) = 1$, $D(t^2) = 2t$ y por lo tanto

$$[D]^B_{B'} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \tag{6}$$

Notemos que estas representaciones implican $F(x, y) = (x, -y)$ en (2), $F(x, y) = (y, x)$ en (3), $F(x, y) = (x\cos(\theta) + y\sin(\theta), -x\sin(\theta) + y\cos(\theta))$ en (4), $F(x, y, z) = (x\cos(\theta) + y\sin(\theta), -x\sin(\theta) + y\cos(\theta), z)$ en (5) y $D(x1 + yt + zt^2) = y1 + 2zt$ en (6).

5.2 Cambio de base

Consideremos primero el caso de endomorfismos $F : V \to V$, y sean $B = (b_1, \ldots, b_n)$, $\tilde{B} = (\tilde{b}_1, \ldots, \tilde{b}_n)$ dos bases ordenadas de V. Tenemos $[v]_B = S[v]_{\tilde{B}}$, $[F(v)]_B = S[F(v)]_{\tilde{B}}$, siendo S la matriz de cambio de base (su columna i es el vector de coordenadas $[\tilde{b}_i]_B$). Por lo tanto, $\forall\ v \in V$,

$$[F(v)]_{\tilde{B}} = S^{-1}[F(v)]_B = S^{-1}([F]^B_B[v]_B) = S^{-1}([F]^B_B S[v]_{\tilde{B}}) = (S^{-1}[F]^B_B S)[v]_{\tilde{B}}$$

es decir, comparando con $[F(v)]_{\tilde{B}} = [F]^{\tilde{B}}_{\tilde{B}}[v]_{\tilde{B}}$,

$$[F(v)]^{\tilde{B}}_{\tilde{B}} = S^{-1}[F]^B_B S$$

que implica también $[F]^B_B = S[F]^{\tilde{B}}_{\tilde{B}}S^{-1}$.
La matrices que representan a un endomorfismo F en diferentes bases son entonces *semejantes* (A de $n \times n$ es semejante a B de $n \times n$ si $A = S^{-1}BS$, con S de $n \times n$ no singular). Las matrices semejantes poseen **el mismo determinante** y **la misma traza**:

$$\mathrm{Det}[A] = \mathrm{Det}[S^{-1}BS] = \mathrm{Det}[S^{-1}]\mathrm{Det}[B]\mathrm{Det}[S] = \mathrm{Det}[B]$$

$$\mathrm{Tr}[A] = \mathrm{Tr}[S^{-1}BS] = \mathrm{Tr}[BSS^{-1}] = \mathrm{Tr}[B]$$

Recordemos que la traza se define como **la suma de todos los elementos diagonales de una matriz**:

$$\mathrm{Tr}[A] = \sum_{i=1}^{n} A_{ii}$$

y satisface $\mathrm{Tr}[AB] = \mathrm{Tr}[BA]\ \forall\ A, B$ de $n \times n$: $\sum_{i,j} A_{ij}B_{ji} = \sum_{i,j} B_{ij}A_{ji}$.

Las matrices que representan a un mismo operador lineal en distintas bases tienen pues *el mismo determinante y la misma traza.* Estas cantidades **permanecen invariantes** frente a cambios de fase y constituyen pues propiedades del operador, no dependientes de la representación.

La matriz S de cambio de base puede reescribirse en este contexto como la matriz que representa al operador identidad I ($I(v) = v\ \forall V \in V$) entre dos bases diferentes:

$$[v]_{\tilde{B}} = [I(v)]_{\tilde{B}} = [I]^B_{\tilde{B}}[v]_B$$

Comparando con $[v]_{\tilde{B}} = S^{-1}[v]_B$, tenemos pues

$$S^{-1} = [I]^B_{\tilde{B}}, \quad S = [I]^{\tilde{B}}_B$$

Por lo tanto, en el caso de endomorfismos podemos escribir

$$[F]^{\tilde{B}}_{\tilde{B}} = [I]^B_{\tilde{B}}[F]^B_B[I]^{\tilde{B}}_B$$

Nótese también que la transformación lineal $G : V \to V$ definida por $G[b_i] = \tilde{b}_i$, $i = 1, \ldots, n$, puede representarse en la base B por la matriz

$$[G]^B_B = [I]^{\tilde{B}}_B = S$$

mientras que $[G]^B_{\tilde{B}}$ es obviamente la matriz identidad I_n.

Ejemplo 1: La matriz que representa a una reflexión F respecto de la recta de ec. $y = x$ en $\mathbb{R}^2$, obtenida anteriormente, se relaciona con aquella que representa a la reflexión respecto del eje x mediante un cambio de base, y son por lo tanto semejantes. Si $B = (e_1, e_2)$ es la base canónica, respecto de la base $\tilde{b}_1 = (e_1+e_2)/\sqrt{2}$, $\tilde{b}_2 = (-e_1 + e_2)/\sqrt{2}$ (vectores unitarios paralelos a las rectas de ec. $y = x$ y $y = -x$) tenemos $F(\tilde{b}_1) = \tilde{b}_1$, $F(\tilde{b}_2) = -\tilde{b}_2$. Por lo tanto,

$$[F]_{\tilde{B}}^{\tilde{B}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

La base $\tilde{B}$ se relaciona con la base canónica mediante la matriz

$$S = [I]_B^{\tilde{B}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

con $S^{-1} = S^t = [I]_{\tilde{B}}^B$. Por lo tanto,

$$[F]_B^B = S[F]_{\tilde{B}}^{\tilde{B}} S^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

que es el resultado obtenido anteriormente. Nótese que la matriz no es diagonal en la base canónica, pero si lo es en la base $\tilde{B}$.

Ejemplo 2: Construir la matriz que representa a una reflexión F respecto de una recta inclinada un ángulo θ (antihorario) respecto del eje x, en R^2. Respecto de la base formada por $\tilde{b}_1 = \cos(\theta)e_1 + \sin(\theta)e_2$, $\tilde{b}_2 = -\sin(\theta)e_1 + \cos(\theta)e_2$, tenemos nuevamente, por definición de reflexión,

$$[F]_{\tilde{B}}^{\tilde{B}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

La base $\tilde{B}$ se relaciona con la base canónica B mediante la matriz de rotación

$$S = [I]_B^{\tilde{B}} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

con $S^{-1} = S^t = [I]_{\tilde{B}}^B$. Por lo tanto,

$$[F]_B^B = S[f]_{\tilde{B}}^{\tilde{B}} S^{-1} = \begin{pmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{pmatrix}$$

Nótese que existe una base ($\tilde{B}$) donde la transformación queda representada por una simple matriz diagonal.

Caso general: Llamemos $\tilde{B} = (\tilde{b}_1, \ldots, \tilde{b}_n)$ una nueva base de V y $\tilde{B}' = (\tilde{b}'_1, \ldots, \tilde{b}'_m)$ una nueva base de W, definidas por matrices de cambio de base S y S' respectivamente ($S = [I]_B^{\tilde{B}}$, $S' = [I]_{B'}^{\tilde{B}'}$). Dado que $[v]_B = S[v]_{\tilde{B}}$ y $[F(v)]_{B'} = S'[F(v)]_{\tilde{B}'}$, tenemos

$$[F(v)]_{\tilde{B}'} = S'^{-1}[F(v)]_{B'} = S'^{-1}[F]_{B'}^B(S[v]_{\tilde{B}}) = (S'^{-1}[F]_{B'}^B S)[v]_{\tilde{B}}$$

y por lo tanto, $[F(v)]_{\tilde{B}'} = [F]_{\tilde{B}'}^{\tilde{B}}[v]_{\tilde{B}}$, con

$$[F]_{\tilde{B}'}^{\tilde{B}} = S'^{-1}[F]_{B'}^B S = [I]_{\tilde{B}'}^{B'}[F]_{B'}^B[I]_B^{\tilde{B}}$$

Nótese que $[F]_{\tilde{B}'}^{\tilde{B}}$ y $[F]_{B'}^B$ son de $m \times n$, S' es de $m \times m$ y S de $n \times n$.

Ejemplo: Sea $D : P_2 \to P_1$ la función derivación restringida a polinomios de grado ≤ 2 con codominio P_1 y sea $B = (1, t, t^2)$ la base canónica de P_2, $B' = (1, t)$ la base canónica de P_1, $\tilde{B} = (1, 1 + t, 1 + t + t^2/2)$ y $\tilde{B}' = (1, 1 + t)$. Tenemos

$$[D]_{B'}^B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, S = [I]_B^{\tilde{B}} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1/2 \end{pmatrix}, S' = [I]_{B'}^{\tilde{B}'} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

con $S'^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$. Por lo tanto,

$$[D]_{\tilde{B}'}^{\tilde{B}} = S'^{-1}[D]_{B'}^{B} S = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

lo cual es también obvio a partir de la definición de D.

5.3 Composición (Producto) de operadores lineales

Sea $F : V \to W$ y $G : W \to U$ dos transformaciones lineales. La composición o producto $(GF) : V \to U$ se define por

$$(GF)(v) = (G \circ F)(v) = G(F(v))$$

El producto de transformaciones lineales es una transformacíion lineal:

$$(GF)(v_1 + v_2) = G(F(v_1 + v_2)) = G(F(v_1) + F(v_2)) = G(F(v_1)) + G(F(v_2)) = (GF)(v_1) + (GF)(v_2)$$

$$(GF)(\alpha v) = G(F(\alpha v)) = G(\alpha F(v)) = \alpha G(F(v)) = \alpha (GF)(v)$$

Para espacios finitamente generados, la matriz $[GF]_{B''}^{B}$ que representa a GF en las bases B, B'' de V y U, es el *producto* de las matrices $[G]_{B''}^{B'}$ y $[F]_{B'}^{B}$ que representan a F y G en bases B', B'' y B, B', siendo B' una base de W:

$$[GF]_{B''}^{B} = [G]_{B''}^{B'}[F]_{B'}^{B}$$

Notemos que si las dimensiones de V, W, U son n, m, p respect. $\Rightarrow$ $[GF]_{B''}^{B}$ es de $p \times n$, $[G]_{B''}^{B'}$ es de $p \times m$ y $[F]_{B'}^{B}$ es de $m \times n$.
Dem.:

$$[(GF)(v)]_{B''} = [G(F(v))]_{B''} = [G]_{B''}^{B'}[F(v)]_{B'} = [G]_{B''}^{B'}([F]_{B'}^{B}[v]_B) = ([G]_{B''}^{B'}[F]_{B'}^{B})[v]_B$$

En particular, si $V = V' = V''$, con $B = B' = B''$,

$$[(GF)]_{B}^{B} = [G]_{B}^{B}[F]_{B}^{B}$$

El producto de funciones así definido es obviamente asociativo ($H(GF) = (HG)F$ para $F : V \to W$, $G : W \to U$ y $H : U \to T$), pero en general, *no conmutativo*, aún cuando $V = W = U = T$.

Ejemplo: Consideremos la composición en $\mathbb{R}^2$ de una rotación F de $\pi/2$ antihoraria seguida de una reflexión G respecto del eje x: Tenemos, en la base canónica $B = ((1, 0), (0, 1))$:

$$[GF]_{B}^{B} = [G]_{B}^{B}[F]_{B}^{B} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} = -[H]_{B}^{B}$$

con H la reflexión respecto de la recta de ec. $y = x$ ($-H$ es la reflexión respecto de la recta de ec. $y = -x$)
Por otro lado, la composición en sentido inverso, es decir una reflexión respecto del eje x seguida de una rotación de $\pi/2$, da como resultado

$$[FG]_{B}^{B} = [F]_{B}^{B}[G]_{B}^{B} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = [H]_{B}^{B}$$

Este sencillo ejemplo muestra que el producto de operadores lineales no es en general conmutativo.

En general, se define el *conmutador* de dos operadores lineales $F : V \to V$, $G : V \to V$ como

$$[F, G] = FG - GF$$

La matriz que representa el conmutador es el conmutador de las matrices que representan F y G:

$$[[F, G]]_{B}^{B} = [F]_{B}^{B}[F]_{B}^{B} - [G]_{B}^{B}[F]_{B}^{B}$$

En el ejemplo anterior, $[F, G] = 2H$ ya que $[[F, G]]_{B}^{B} = 2[H]_{B}^{B}$.

5.4 El espacio vectorial de transformaciones lineales

Consideremos dos transformaciones lineales $F: V \to W$, $G: V \to W$. La suma $(F+G): V \to W$ se define como

$$(F+G)(v) = F(v) + G(v)$$

y es claramente una transformación lineal:

$$(F+G)(v_1+v_2) = F(v_1+v_2)+G(v_1+v_2) = F(v_1)+F(v_2)+G(v_1)+Gg(v_2) = (F+G)(v_1)+(F+G)(v_2)$$

$$(F+G)(\alpha v) = F(\alpha v) + G(\alpha v) = \alpha F(v) + \alpha G(v) = \alpha(F(v)+G(v)) = \alpha(F+G)(v)$$

Es fácil verificar que la suma es conmutativa ($F+G = G+F$) y asociativa ($(F+G)+H = F+(G+H)$). Existe además un elementro neutro 0, que es la función nula definida por $0(v) = 0 \; \forall \; v \in V$, con $F+0 = 0+F = F$. El elemento opuesto de F es entonces $-F$, definido por $-F(v) = -(F(v))$, que es también lineal. El conjunto de las funciones lineales $\{F: V \to V', \; F \text{ lineal}\}$ es pues un grupo abeliano con la operación de suma.

El producto por un escalar $\alpha \in K$, $(\alpha F): V \to V'$ se define obviamente como

$$(\alpha F)(v) = \alpha F(v)$$

y es también una función lineal:

$$(\alpha F)(v_1+v_2) = \alpha F(v_1+v_2) = \alpha(F(v_1)+F(v_2)) = \alpha F(v_1) + \alpha F(v_2) = (\alpha F)(v_1) + (\alpha F)(v_2)$$

$$(\alpha F)(\beta v) = \alpha F(\beta v) = \alpha(\beta F(v)) = (\alpha\beta)F(v) = (\beta\alpha)F(v) = \beta(\alpha F)(v)$$

Es fácil verificar además que $\alpha(\beta F) = (\alpha\beta)F$, $(\alpha+\beta)F = \alpha F + \beta F$, $\alpha(F+G) = \alpha F + \alpha G$, $1F = F$.

El conjunto de todas las transformaciones lineales de V a W es entonces *un espacio vectorial* sobre K, denominado $Hom(V,W)$ (Homomorfismos de V en W).

Notemos también que con respecto al producto (composición) de funciones, la suma verifica las propiedades distributivas $(G+H)F = GF + HF$ para $F: V \to W$, y $G, H: W \to U$ y $H(F+G) = HF + HG$ para $H: W \to U$ y $F, G: V \to W$. Además, por ser lineales, $\alpha(GF) = (\alpha G)F = G(\alpha F)$ para $\alpha \in K$.

La matriz que representa a $(F+G)$ en bases B, B' de V y W es claramente la suma de matrices,

$$[F+G]^B_{B'} = [F]^B_{B'} + [G]^B_{B'}$$

y la que representa a (αF) es obviamente

$$[\alpha F]^B_{B'} = \alpha[F]^B B'$$

En efecto, $[(F+G)(v)]_{B'} = [F(v)+g(v)]_{B'} = [F(v)]_{B'}+[G(v)]_{B'} = [F]^B_{B'}[v]_B+[G]^B_{B'}[v]_B = ([F]^B_{B'}+[G]^B_{B'})[v]_B$.
$[(\alpha F)(v)]_{B'} = [\alpha F(v)]_{B'} = \alpha[F(v)]_{B'} = \alpha([F]^B_{B'}[v]_B) = (\alpha[F]^B_{B'})[v]_B$.

6. Monomorfismos, Epimorfismos e Isomorfismos

I) Un monomorfismo es una TL $F : V \to W$ inyectiva (o sea, $F(v_1) \neq F(v_2)$ si $v_1 \neq v_2$).
F es un monomorfismo si y sólo si $N(F) = \{0\}$.
Dem.: Si F es un monomorfismo y $v \neq 0$, $F(v) \neq F(0) = 0 \Rightarrow N(F) = \{0\}$.
Si $N(F) = \{0\} \Rightarrow F(v_1) - F(v_2) = F(v_1 - v_2) \neq 0$ si $v_1 - v_2 \neq 0$, o sea, $F(v_1) \neq F(v_2)$ si $v_1 \neq v_2$.

Como consecuencia, dim $N(F) = 0$. Por lo tanto, si V es de dimensión finita, dim $I(F) =$ dim V.
Y como $I(F) \subset W$, F puede ser un monomorfismo sólo si dim $V \leq$ dim W.
Los monomorfismos **conservan la independencia lineal**: Si $\{v_1, \ldots, v_m\}$ son vectores LI de V y F es un monomorfismo $\Rightarrow \{F(v_1), \ldots, F(v_m)\}$ son vectores LI de W. Dem.: Si

$$0 = \sum_{i=1}^{m} \alpha_i F(v_i) = F(\sum_{i=1}^{m} \alpha_i v_i)$$

entonces $\sum_{i=1}^m \alpha_i v_i \in N(F)$. Como $N(F) = \{0\} \Rightarrow \sum_{i=1}^m \alpha_i v_i = 0$, lo que implica $\alpha_i = 0$ para $i = 1, \ldots, m$ por ser los v_i LI. Por lo tanto, $\{F(v_1), \ldots, F(v_m)\}$ son LI.
En particular, si $B = (b_1, \ldots, b_n)$ es una base de un espacio V finitamente generado y $F : V \to V'$ es un monomorfismo, $(F(b_1), \ldots, F(b_n))$ es una **base** de $I(F)$.

II) Un epimorfismo es una TL $F : V \Rightarrow W$ sobreyectiva ($I(F) = W$). Si W es de dimensión finita $\Rightarrow F$ es un epimorfismo si y sólo si dim $I(F) =$ dim W.
Como dim $I(F) \leq$ dim V, F puede ser un epimorfismo sólo si dim $W \leq$ dim V.

III) Un isomorfismo es una TL biyectiva, es decir, es a la vez un monomorfismo y un epimorfismo.
En espacios de dimensión finita, un isomorfismo puede entonces existir sólo si dim V=dim W (pues por ser monomorfismo, dim $V \leq$ dim W y por ser epimorfismo, dim $W \leq$ dim V.)
Si $B = (b_1, \ldots, b_n)$ es una base de V y $F : V \to W$ es un isomorfismo, $\{F(b_1), \ldots, F(b_n)\}$ es una base de W, pues son LI (por ser F monomorfismo) y generan W (por ser F epimorfismo). Los isomorfismos transforman entonces bases en bases.
Si $W = V$, un operador que es un isomorfismo se dice **no singular** o **automorfismo**. En caso contrario se dice **singular**.

Dados dos espacios V, W sobre el mismo cuerpo K, se dice que V es *isomorfo* a W si existe un isomorfismo $F : V \to W$. En tal caso existe una correspondencia uno a uno entre los elementos de V y W.
Dos espacios vectoriales V, W de dimensión finita sobre el mismo cuerpo K son isomorfos *si y sólo si* dim $V =$ dim W.
Dem.: Ya hemos demostrado que si $\exists$ un isomorfismo entre V y $W \Rightarrow$ dim V=dim W.
Por otro lado, si dim $V =$ dim $W = n$ y $B = \{b_1, \ldots, b_n\}$, $B' = \{b'_1, \ldots, b'_n\}$ son bases de V y W, la TL definida por

$$F(b_i) = b'_i, \quad i = 1, \ldots, n$$

(o sea $F(B) = B'$) es un isomorfismo. En efecto, es sobreyectiva, pues $I(F) = F(\overline{B}) = \overline{F(B)} = \overline{B'} = W$ (o sea, si $w \in W \Rightarrow w = \sum_{i=1}^n \alpha_i b'_i = \sum_{i=1}^n \alpha_i F(b_i) = F(\sum_{i=1}^n \alpha_i b_i) \in I(F)$), por lo que $I(F) = W$).
Esto implica que F es también un monomorfismo pues dim $N(F) =$ dim V-dim $I(F)$ =dim V-dim $W = 0$, lo que implica $N(F) = \{0\}$.

Notemos finalmente que si dim $V =$ dim $W = n$ y $F : V \to W$ es una TL, son equivalentes:
a) F es un isomorfismo; b) F es monomorfismo; c) F es epimorfismo
como consecuencia de la relación $\dim N(F) + \dim I(F) = \dim V$. En efecto, a) implica b)+c) (y viceversa) por definición, b) implica c) ya que en tal caso $N(F) = \{0\}$ y por lo tanto dim $I(F) = n$, y c) implica b) por la misma propiedad (pues si dim $I(F) = n \Rightarrow$ dim $N(F) = 0$ y por lo tanto $N(F) = \{0\}$).

No obstante, si V es de dimensión infinita, un operador lineal $F : V \to W$ puede ser monomorfismo sin ser epimorfismo y viceversa. Pro ejemplo, si P es el espacio de polinomios reales, $D : P \to P$ definido por $D(p(x)) = p'(x)$ (derivada) no es monomorfismo ($D(1) = 0$) pero es epimorfismo (probar como ejercicio). Y el operador $S : P \to P$ definido por $S(p(x)) = \int_0^x p(t)dt$ es monomorfismo (pues $N(S) = \{0\}$) pero no es epimorfismo (probar como ej.).

IV) Si $F : V \to W$ es un isomorfismo $\Rightarrow$ la transformación inversa $F^{-1} : W \to V$, definida por $F^{-1}(w) = v$, con v el único vector $\in V$ tal que $F(v) = w$, es lineal y es un isomorfismo.
Dem.: Si F es isomorfismo, la inversa F^{-1} es obviamente una función bien definida.
Si $F(v_1) = w_1$, $F(v_2) = w_2 \Rightarrow F(v_1 + v_2) = F(v_1) + F(v_2) = w_1 + w_2$, lo que implica $F^{-1}(w_1 + w_2) = v_1 + v_2 = F^{-1}(w_1) + F^{-1}(w_2)$.
Si $F(v) = w$ y $\alpha \in K \Rightarrow F(\alpha v) = \alpha F(v) = \alpha w$, lo que implica $F^{-1}(\alpha w) = \alpha v = \alpha F^{-1}(w)$.
F^{-1} es por lo tanto TL.
Además, F^{-1} es un monomorfismo, pues $N(F^{-1}) = \{0\}$ y es un epimorfismo pues si $v \in V$, $v = F^{-1}(w)$, con $w = F(v)$, por lo que $I(F^{-1}) = V$.
Como consecuencia de la definición, $F(F^{-1}(w)) = w \ \forall\ w \in W$ y $F^{-1}(F(v)) = v \ \forall\ v \in V$. Por lo tanto,

$$FF^{-1} = I_W, \quad F^{-1}F = I_V$$

donde $I_W : W \to W$ denota la identidad en W e $I_V : V \to V$ aquella en V. Para V y W de dimensión finita, las matrices correspondientes satisfacen

$$[FF^{-1}]_{B'} = [F]^B_{B'}[F^{-1}]^{B'}_B = I_n, \qquad [F^{-1}F]_B = [F^{-1}]^{B'}_B[F]^B_{B'} = I_n$$

donde $I_n = [I_W]_{B'} = [I_V]_B$ es la matriz identidad de $n \times n$. Por lo tanto, $[F^{-1}]^{B'}_B$ *es la matriz inversa* de $[F]^B_{B'}$:

$$[F^{-1}]^{B'}_B = ([F]^B_{B'})^{-1}$$

Una TL $F : V \to w$ entre espacios de dimensión finita es pues un isomorfismo *si y sólo si* está representada por matrices $[F]^B_{B'}$ cuadradas *no singulares* ($\text{Det}[F]^B_{B'} \neq 0$).
Dem.: Si es isomorfismo, por lo visto anteriormente $[F]^B_{B'}$ es cuadrada e invertible y por lo tanto no singular. Y si $[F]^B_{B'}$ es cuadrada no singular, la única solución de $[F]^B_{B'}[v]_B = 0$ es $[v]_B = 0$, es decir, $v = 0$. Esto implica que $N(F) = \{0\}$ y $\Rightarrow F$ es monomorfismo y por lo tanto isomorfismo.

Si $V = W$, F es un operador no singular sii $[F]_B$ es una matriz no singular. En tal caso $[F^{-1}]_B = [F]^{-1}_B$.

Recordemos que la inversa de una matriz no singular A ($\text{Det}[A] \neq 0$) puede obtenerse como

$$A^{-1} = C/\text{Det}[A], \ \text{ con } \ C_{ij} = (-1)^{i+j}\text{Det}[M_{ji}]$$

donde Det denota el determinante y C la matriz de cofactores traspuesta, siendo M_{ji} la matriz de $(n-1) \times (n-1)$ obtenida al suprimir la fila j y columna i de A.
Por ejemplo, si A es de 2×2,

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \Rightarrow \quad A^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Ejemplo 1): Si V es un espacio vectorial sobre el cuerpo K con dim $V = n$ y $B = (b_1, \ldots, b_n)$ es una base ordenada de V, la función $R : V \to K^n$ dada por

$$R(v) = [v]_B = \begin{pmatrix} \alpha_1 \\ \ldots \\ \alpha_n \end{pmatrix}$$

donde $v = \sum_{i=1}^n \alpha_i b_i$, es un *isomorfismo*.
En efecto, hemos demostrado anteriormente que $R(v)$ es lineal ($[\alpha v]_B = \alpha[v]_B$ y $[v_1 + v_2]_B = [v_1]_B + [v_2]_B$). Además $N(R) = \{0\}$ (pues $[v]_B = 0$ (vector columna nulo) si y sólo si $v = 0$). Por lo tanto es un isomorfismo. Esto implica en particular que un conjunto de vectores $v_1, \ldots, v_m \in V$ serán LI si y sólo si los correspondientes vectores columna $[v_1]_B, \ldots, [v_m]_B \in K^n$ son LI.

Ejemplo 2): Si dim $V = n$, dim $W = m$ y B, B' son bases ordenadas de V y W, la función $H : Hom(V, W) \to K^{m\times n}$ dada por $H(F) = [F]^B_{B'}$ es un isomorfismo.
En efecto, H es lineal (ya que $[(F + G)]^B_{B'} = [F]^B_{B'} + [G]^B_{B'}$ y $[\alpha F]^B_{B'} = \alpha[F]^B_{B'}$). Además, H es inyectiva, pues $N(H) = \{0\}$ (0 denota la función nula, representada en cualquier base por la matriz nula de $m \times n$ (sus elementos son todos 0)) y es también suryectiva, pues una matriz arbitraria $T \in K^{m\times n}$ corresponde a la transformación lineal definida por $F(b_i) = \sum_{j=1}^m T_{ji}b'_j$, $i = 1, \ldots, n$. Es decir, $I(H) = K^{m\times n}$. Esto implica que dim $Hom(V, W) = $ dim $K^{m\times n} = m.n$.

La dimensión de la imagen de F puede calcularse evaluando el **rango** de la matriz $T = [F]_{B'}^{B}$ de $m \times n$ (es decir, el número de columnas (o filas) LI) en *cualquier* par de bases B, B', ya que esto será equivalente al número de vectores $F(b_i)$ LI. Del mismo modo, la imagen $I(F)$ puede obtenerse a partir del espacio columna de T (es decir, el espacio generado por las columnas de T) y el núcleo $N(F)$ a partir del espacio nulo de T (este último es el conjunto de vectores $[v]_B \in K^{n\times 1}$ que satisfacen $T[v]_B = 0$, y que son por tanto ortogonales a todas las filas de T).
Como base de $K^{m\times n}$ pueden elegirse las matrices E^{ij} cuyo único elemento no nulo es el ij, definidas por $E^{ij}_{kl} = 1$ si $k = i$ y $l = j$ y $E^{ij}_{kl} = 0$ en caso contrario. Como base de $Hom(V, V')$ pueden elegirse las correspondientes transformaciones lineales F^{ij} definidas por $F^{ij}(e_l) = 0$ si $l \neq j$ y $F^{ij}(e_l) = e'_i$ si $l = j$, tal que $[F^{ij}]_{e'}^{e} = E^{ij}$. Aquí e y e' denotan las bases canónicas de K^n y K^m respectivamente.

6.1 Rango, espacio columna y espacio fila de una matriz.

El **espacio columna** una matriz T ($EC(T)$) de $m \times n$ es el subespacio $S_C \subset K^{m\times 1}$ generado por las n columnas de T, y el **espacio fila** de T ($EF(T)$) el subespacio $S_F \subset K^{1\times n}$ generado por las m filas de T.

Estos subespacios son en general diferentes, pero **tienen siempre la misma dimensión** (véase demostración abajo). El **rango** de una matriz T es la dimensión del EC o EF de la misma.

El **espacio nulo** de T es el subespacio $S_N \subset K^{n\times 1}$ formado por las soluciones $X \in K^{n\times 1}$ de la ecuación homogénea $TX = 0$: $N(T) = \{X \in \mathbb{R}^{n\times 1}, TX = 0\}$. Su dimensión se denomina nulidad de T.

Consideremos ahora las relaciones entre F y la matriz $T = [F]_{B'}^{B}$ de $m \times n$ que representa a F en bases $B = (b_1, \ldots, b_n)$, $B' = (b'_1, \ldots, b'_m)$ de V y W. La columna i de T es la matriz columna de coordenadas $[F(b_i)]_{B'}$, con $[F(v)]_{B'} = T[v]_B$.

a) dim $I(F) =$ rango(T).
En efecto, $I(F) = \overline{\{F(b_1), \ldots, F(b_n)\}}$, por lo que su dimensión es el número de vectores $F(b_i)$ LI. Pero esto coincide con el número de columnas $[F(b_i)]_{B'}$ LI de T (véase Ej. 1), que es la dimensión del $EC((T)$, es decir su rango.

b) Como dim $I(F)$ es independiente de las bases B y B' $\Rightarrow$ rango(T)=rango(T') si $T' = S'^{-1}TS$, con S' de $m \times m$ y S de $n \times n$ no singulares, ya que T' corresponde a la representación de F en otro par de bases ($T' = [F]_{\tilde{B}'}^{\tilde{B}}$, con $S = [I]_{B}^{\tilde{B}}$, $S' = [I]_{B'}^{\tilde{B}'}$).

c) dim $N(F) =$ nulidad (T). En efecto, $N(F)$ es el subespacio formado por los vectores $v \in V$ que satisfacen $F(v) = 0$, y por lo tanto, $[F(v)]_{B'} = T[v]_B = 0$, de modo que $[v]_B$ pertence al espacio nulo de T. Y si x pertenece al espacio nulo de T ($Tx = 0$) entonces $F(v) = 0$, con $[v]_B = x$. Ambos subespacios son pues isomorfos y tienen la misma dimensión. Esto también implica que nulidad(T) =nulidad(T') si $T' = S'^{-1}TS$ con S' y S no singulares.
La relación dim $N(F)+$ dim $I(F) =$ dim $V = n$ implica entonces nulidad (T)+rango $(T) = n$.

d) Las dimensiones del EC. y del EF de una matriz arbitraria T de $m \times n$ son coincidentes.
Dem.: Podemos considerar a T como la representación de una TL $F : V \to W$ en bases B y B' de V y W, con dim $V = n$, dim $W = m$, tal que $T = [F]_{B'}^{B}$. En nuevas bases definidas por $\tilde{B} = (\tilde{b}_1, \ldots, \tilde{b}_k, \tilde{b}_{k+1}, \ldots, \tilde{b}_n)$ tal que $(\tilde{b}_{k+1}, \ldots, \tilde{b}_n)$ sea base de $N(F)$, y $\tilde{B}' = (\tilde{b}'_1, \ldots, \tilde{b}'_k, \tilde{b}'_{k+1}, \ldots, \tilde{b}'_m)$ tal que $\tilde{b}'_i = F(b_i)$ para $i \leq k$ (recordar que estos son LI !) se tendrá $[F(\tilde{b}_i)]_{\tilde{B}'} = 0$ si $i > k$ y $([F(\tilde{b}_i)]_{\tilde{B}'})_j = \delta_{ij}$ si $i \leq k$, por lo que la matriz $T' = [F]_{\tilde{B}'}^{\tilde{B}} = S'^{-1}TS$, con $S = [I]_{B}^{\tilde{B}}$ y $S' = [I]_{B'}^{\tilde{B}'}$ no singulares, contendrá una submatriz identidad de $k \times k$ en las primeras k filas, siendo las restantes nulas. La dim. del EC. y del EF de esta matriz es por lo tanto k.
Entonces k es la dimensión de la imagen de F, por lo que la dim. del EC de $T = S'T'S^{-1}$ será k.
A su vez, el EF de T es el EC de T^t. Como la dim. del EC de T'^t es también k, la dim. del EC de $T^t = S^{-1\,t}T'^tS'^t$ es entonces k, por ser S'^t, $S^{-1\,t}$ no singulares. Pero esta es la dim. del EF de T.

Una demostración alternativa que permite hallar una base de $EF(T)$ y $EC(T)$ es la siguiente: Aplicando un número finito de operaciones elementales por fila, se puede llevar T a la forma de escalonada de Gauss-

Jordan,

$$T' = S'^{-1}T = \begin{pmatrix} 1 & x & \dots & 0 & x & \dots & 0 & x & \dots \\ 0 & 0 & \dots & 1 & x & \dots & 0 & x & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots & 1 & x & \dots \\ & & \dots & & & \dots & & & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots \\ & & \dots & & & \dots & & & \dots \end{pmatrix}$$

donde x representa elementos no necesariamente nulos, y S'^{-1} una matriz no singular que es el producto de las operaciones elementales. T' posee k filas no nulas que son LI y por lo tanto la dimensión de $EF(T')$ (idéntico $EF(T)$ por ser las filas de T' combinaciones lineales de las de T) es k. Una base de $EF(T)$ son pues las k filas no nulas de T'.

k es también el número de columnas LI de T', ya que las columnas con pivotes (primer elemento no nulo de c/fila no nula) son LI y generan $EC(T')$. Por lo tanto, la dimensión de $EC(T')$ es también k.
Pero esta es entonces la dimensión de $EC(T)$ por ser S'^{-1} no singular (véase b)).

Considerando a T como la representación de una TL $F: V \to W$ en bases B, B' de V y W tal que $T = [F]_{B'}^{B}$, la matriz $T' = S'^{-1}T = [F]_{\tilde{B}'}^{B}$ corresponde a un cambio de base en W, con $S' = [I]_{B'}^{\tilde{B}'}$. Las columnas con pivotes de T', $[F(b_{i_p})]_{\tilde{B}'}$, forman una base de $EC(T')$, por lo que los correspondientes vectores $F(b_{i_p})$ forman una base de $I(F)$. Las correspondientes columnas de T, $[F(b_{i_p})]_{B'}$, forman entonces una base de $EC(T)$. de T. Nótese que en general, $EC(T) \neq EC(T')$, aunque las dimensiones sean iguales, pero $EF(T') = EF(T)$.

Ejemplo 3) Sea $D: P_2 \to P_2$ el operador derivada restringido al subespacio de polinomios de grado ≤ 2. Es obvio que $N(D) = P_0$ (el subespacio de polinomios de grado 0, es decir, constantes) y que $I(D) = P_1$ (el subespacio de polinomios de grado ≤ 1), con dim $N(D)+$ dim $I(D) = 1 + 2 = 3 =$dim P_2.
En forma matricial, en la base "canónica" $B = (1, t, t^2)$, tenemos

$$[D]_B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

El rango de esta matriz es 2, ya que posee dos filas (o columnas) LI, que coincide con dim $I(D)$.

Además, el espacio columna de $[D]_B$ es $\overline{\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}\}} = \overline{\{[e_1]_e, 2[e_2]_e\}}$, de modo que $I(D)$ será el subespacio generado por $e_1 = 1$ y $e_2 = t$, es decir, P_1.

El espacio nulo de $[D]_B$ es $\overline{\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\}} = \overline{\{[e_1]_e\}}$, y $N(D)$ es por lo tanto el subespacio generado por e_1, es decir, P_0.

Ejemplo 4) Sea $F: \mathbb{R}^2 \to \mathbb{R}^2$ dada por $F(x, y) = (2x + y, 3x - y)$. Mostrar que F es no singular y hallar su inversa.

F es un isomorfismo pues $I(F) = \{x(2, 3) + y(1, -1),\ x, y \in \mathbb{R}\} = \mathbb{R}^2$, ya que $(2, 3)$ y $(1, -1)$ son LI y por lo tanto base de $\mathbb{R}^2$. Puede llegarse al mismo resultando notando que $N(F) = \{(0, 0)\}$. Y también, notando que la matriz que representa a F en la base canónica $B = ((1, 0), (0, 1))$ es

$$[F]_B = \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix}$$

Como $[F]_B$ es no singular (Det$[[F_B] = -5 \neq 0$), $[F]_B$ es invertible y por lo tanto F es un isomorfismo. La matriz que representa a su inversa es

$$[F^{-1}]_B = ([F]_B)^{-1} = \frac{1}{5}\begin{pmatrix} 1 & 1 \\ 3 & -2 \end{pmatrix}$$

por lo que la función inversa está dada por $F^{-1}(x, y) = (x + y, 3x - 2y)/5$.

Ejemplo 5) Sea $F : \mathbb{R}^3 \to \mathbb{R}^3$ dada por $F(x, y, z) = (x + y + z, 2x + z, 3x - y + z)$. Hallar $N(F)$ y $I(F)$.

En este caso conviene pasar directamente a la representación matricial. La matriz que representa a F en la base canónica $B = ((1, 0, 0), (0, 1, 0), (0, 0, 1))$ es

$$[F]_B = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 0 & 1 \\ 3 & -1 & 1 \end{pmatrix}$$

F no es un isomorfismo pues $[F]_B$ es una matriz singular ($\text{Det}[[F]_B] = 0$) y por lo tanto no posee inversa. Las columnas C_i de $[F]_B$ están vinculadas por $C_3 = (C_2+C_1)/2$ (y las filas F_i de $[F]_B$ por $F_3 = 2F_2-F_1$), siendo C_2 y C_1 LI. El rango de $[F]_B$ es por lo tanto 2. Esto implica que dim $I(F) = 2$ y que dim $N(F) = 3-2 = 1$.

Para hallar $N(F)$ se resuelve el sistema de ecuaciones que resulta de $F(x, y, z) = (0, 0, 0)$, o sea, $x+y+z = 0$, $2x + z = 0$, $3x - y + z = 0$, que puede reescribirse en forma matricial como

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & 0 & 1 \\ 3 & -1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

es decir, $[F]_B[v]_B = 0$ (vector columna nulo). Puede verse fácilmente que $[F]_B$ es equivalente por filas a la matriz $\begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 0 \end{pmatrix}$. La solución al sistema homogéneo está entonces dada por $x = -z/2$, $y = -z/2$, con z arbitrario, por lo que el espacio nulo de la matriz es el conjunto $\overline{\{(-1/2, -1/2, 1)^t\}}$.
El núcleo de F es pues el subespacio generado por el vector $v_0 = (-1/2, -1/2, 1)$
Una base del espacio columna de $[F]_B$ es, por ej., el conj. formado por las dos primeras columnas.
Por lo tanto, $I(F)$ es el espacio generado por $v_1 = (1, 2, 3) = f(e_1)$, $v_2 = (1, 0, -1) = f(e_2)$.
Podemos escribir entonces $V = \overline{\{e_1, e_2\}} \oplus \overline{v_0}$ (la barra sobre vectores indica el espacio generado por dichos vectores), con v_0 base del núcleo y $(F(e_1), F(e_2))$ base de $I(F)$.

Ejemplo 6) Sea $f : \mathbb{R}^2 \to \mathbb{R}^2$, definida por $F(1, 1) = (2, 1)$, $F(1, -1) = (-1, 0)$. Hallar $F(x, y)$.

Los datos alcanzan para definir F pues $b'_1 = (1, 1)$, $b'_2 = (1, -1)$ son LI y por lo tanto base de $\mathbb{R}^2$ (y toda transformación lineal queda completamente determinada por los vectores que asigna a los elementos de una base). Tenemos, a partir de los datos, si B es la base canónica y $B' = (b'_1, b'_2)$,

$$[F]_B^{B'} = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}$$

Además,

$$[I]_B^{B'} = S = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad [I]_{B'}^{B} = S^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

Por lo tanto,

$$[F]_B^B = [F]_B^{B'}[I]_{B'}^B = [F]_B^{B'}S^{-1} = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}/2 = \frac{1}{2}\begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix}$$

que implica

$$F(x, y) = (x + 3y, x + y)/2$$

Puede llegarse al mismo resultado a partir de la relación $e_1 = (b'_1 + b'_2)/2$, $e_2 = (b'_1 - b'_2)/2$, con $Ff[e_1] = [F(b'_1)+F(b'_2)]/2 = (1, 1)/2$, $F[e_2] = [F(b'_1)-F(b'_2)]/2 = (3, 1)/2$ y por lo tanto $F(x, y) = xF(e_1)+yF(e_2) = (x+3y, x+y)/2$. El método matricial es, no obstante, más directo, y adecuado para ser aplicado a sistemas de grandes dimensiones.

7. Inversas a Izquierda y Derecha

Sea $F : V \to W$ una transformación lineal. $G : W \to V$ lineal se denomina inversa a izquierda de F si

$$GF = I_V$$

donde $I_V : V \to V$ es la identidad en V. En tal caso F es la inversa a derecha de G.

Teorema: Una transformación lineal posee inversa a izquierda si y sólo si es un monomorfismo, e inversa a derecha si y sólo si es un epimorfismo.
(Recordar esquema gráfico hecho en clase).
Dem.: a) Si F es un monomorfismo $\Rightarrow \forall\, v' \in I(F)\ \exists$ un y sólo un $v \in V$ tal que $F(v) = v'$. Podemos escribir en general $W = I(F) \oplus Q$, donde Q es un suplemento de $I(F)$. Todo vector $v' \in W$ puede pues escribirse en forma *única* como $v' = v'_1 + v'_2$, donde $v'_1 = F(v_1) \in I(F)$ y $v'_2 \in Q$. Definimos entonces $G : V' \to V$ como

$$G(v') = v_1$$

De esta forma, $G(v'_1) = v_1$ y $G(v'_2) = 0$ si $v'_1 \in I(F)$ y $v'_2 \in Q$. Es fácil comprobar que G es lineal (pues $G(v' + u') = v_1 + u_1 = G(v') + G(u')$ si $v' = v_1 + v_2$ y $u' = u_1 + u_2$, y $G(\alpha v') = \alpha v_1 = \alpha G(v')$). Es además un epimorfismo y satisface $GF = I_V$.
Notemos que si $I(F) \neq W$, la inversa a izquierda no es única, pues podemos sumar a G cualquier función lineal $H : W \to V$ no nula que satisfaga $H(v') = 0$ si $v' \in I(F)$ (o sea, $I(F) \subset N(H)$) tal que $HF = 0$ y por lo tanto $(G + H)F = GF$.
Además, si F no es monomorfismo, $\exists\ v \in V$, con $v \neq 0$, tal que $F(v) = 0$ y por lo tanto, $(GF)(v) = G(F(v)) = G(0) = 0 \neq v$, por lo que F no puede tener inversa a izquierda en tal caso.

b) Si $G : W \to V$ es un epimorfismo, sea $N(G)$ su espacio nulo y sea Q un suplemento tal que $W = N(G) \oplus Q$. G restringido a Q ($G : Q \to V$) es un isomorfismo, ya que sigue siendo epimorfismo y además, si $v' \in Q$ y $v' \neq 0$, $G(v') \neq 0$. Definamos $F : V \to V'$ tal que $F(v)$ es el único vector v' de Q que satisface $G(v') = v$ ($I(F) = Q$). Es fácil ver que F es lineal, es monomorfismo y satisface $GF = I_V$.
No obstante, la inversa a derecha no es única si $N(G) \neq \{0\}$, pues podemos sumar a F cualquier función no nula $H : V \to W$ con $I(H) \subset N(G)$, tal que $GH = 0$.
Además, si G no es epimorfismo, $\exists\, v \in V$ tal que v no pertence a $I(G)$ y por lo tanto $(GF)(v) = G(F(v)) \neq v$, pues $G(F(v)) \in I(G)$. No puede pues existir inversa a derecha en este caso.

Si V es de dimensión n y W de dimensión m, las matrices que representan a F y G en bases ordenadas B, B' de V y W satisfacen

$$[G]_B^{B'}[F]_{B'}^B = I_n$$

con I_n la identidad de $n \times n$, $[F]_{B'}^B$ de $m \times n$ y $[G]_B^{B'}$ de $n \times m$. La matriz $[G]_B^{B'}$ se dice que es inversa a izquierda de la matriz $[F]_{B'}^B$, y $[F]_{B'}^B$ la matriz inversa a derecha de $[G]_B^{B'}$. Si F posee una inversa a izquierda G y a derecha H debe ser entonces un isomorfismo, con $m = n$ en el caso de dimensión finita. En tal caso $G = H$, ya que $G = GI_W = G(FH) = (GF)H = I_V H = H$.

Este teorema implica que una matriz A de $m \times n$ (m filas, n columnas) **tiene inversa a izquierda** B ($BA = I_n$, con B de $n \times m$) **si y sólo si Rango** $(A) = n$ (y por lo tanto $m \geq n$), en cuyo caso A representa a un **monomorfismo**. Y una matrix B de $m \times n$ tiene **inversa a derecha** A ($BA = I_m$, con A de $n \times m$) **si y sólo si Rango** $(B) = m$ (y por lo tanto $m \leq n$), en cuyo caso representa un epimorfismo.
Si una matriz posee inversa izquierda y a derecha entonces debe ser necesariamente cuadrada y representar un isomorfismo, siendo pues no singular. En tal caso la inversa a izquierda y a derecha coinciden.

Si $A \in \mathbb{C}^{m\times n}$ tiene rango n (lo que implica $m \geq n$) sus columnas son LI. Es fácil mostrar que $A \in \mathbb{C}^{m\times n}$ **tiene rango n si y sólo si la matriz $A^\dagger A \in \mathbb{C}^{n\times n}$ es no singular (**$\mathrm{Det}(A^\dagger A) \neq 0$**)**. Recordemos que $A^\dagger = (A^t)^*$, es decir, $A^\dagger_{ij} = A^*_{ji}\ \forall\ i, j$.
Dem.: Si las columnas son independientes, la única solución de $AX = 0$ (con $X \in \mathbb{C}^{n\times 1}$) es $X = 0$ (en otras palabras, A representa a un monomorfismo y por lo tanto su espacio nulo es $\{0\}$). Si existe X tal que $A^\dagger AX = 0$ entonces $X^\dagger A^\dagger AX = (AX)^\dagger(AX) = |AX|^2 = 0$ y por lo tanto $AX = 0$. Esto implica entonces $X = 0$, por lo que $A^\dagger A$ es no singular: $\mathrm{Det}(A^\dagger A) \neq 0$.

Análogamente, Si $A^\dagger A$, es no singular, la única solución de $A^\dagger AX = 0$ es $X = 0$, por lo que la única solución de $AX = 0$ es $X = 0$, indicando que las columnas de A son linealmente independientes, es decir, que A tiene rango n.
Esto permite pues construir en forma inmediata una inversa a izquierda B $\forall$ $A \in \mathbb{C}^{m\times n}$ con rango n:

$$B = (A^\dagger A)^{-1} A^\dagger \qquad (1)$$

ya que $BA = I_n$ (notar que $B \in \mathbb{C}^{n\times m}$).

Análogamente, si $B \in \mathbb{C}^{m\times n}$ tiene rango m, B representa a un epimorfismo. En tal caso $B^\dagger \in \mathbb{C}^{n\times m}$ tiene rango m y por lo tanto $BB^\dagger$ es no singular. Una inversa a derecha de B es pues

$$A = B^\dagger (BB^\dagger)^{-1} \qquad (2)$$

ya que $BA = I_m$ (notar que $A \in \mathbb{C}^{n\times m}$). Recordemos, no obstante, que ai $m \neq n$, existen otras inversas a izquierda y a derecha respectivamente. Por otro lado, si $m = n \Rightarrow B = A^{-1}$ en (1) y $A = B^{-1}$ en (2), como el lector puede fácilmente comprobar.

Ejemplo 1) Sea $D : P \to P$ la derivación considerada en el espacio vectorial P de todos los polinomios, de dimensión infinita. D no es un monomorfismo, pues $D(1) = 0$ (la derivada de cualquier polinomio de grado 0 es el polinomio nulo), por lo que $N(D) = \overline{\{1\}}$, pero sí es un epimorfismo, pues $\forall q(t) \in P\ \exists\ p(t) \in P$ tal que $D(p(t)) = q(t)$ (Por ejemplo, $p(t) = \int_0^t q(t')dt'$, que es un polinomio).
Por lo tanto, D tendrá inversa a derecha. Una inversa es precisamente la integración $S : P \to P$ definida por $S(q(t)) = \int_0^t q(t')dt'$, que es lineal y que satisface

$$DS = I_P$$

(I_P es la identidad en P). En efecto, $(DS)(t^n) = D(S(t^n)) = D(t^{n+1}/(n+1)) = t^n\ \forall\ n \geq 0$. No obstante, $SD \neq I_P$ pues $SD(1) = S(0) = 0 \neq 1$, o sea que S es inversa sólo a derecha de D.
S es un monomorfismo ($N(S) = \{0\}$), pero no un epimorfismo, ya que la imagen de S no contiene a los polinomios de grado 0.
Notemos también que S' definido por $S'(q(t)) = \int_a^t q(t')dt'$ es también una inversa a derecha de D para cualquier a real, y también lo es T dada por $T(q(t)) = \int_0^t q(t')dt' + cq(0)$, siendo c cualquier constante $\in K$.

Ejemplo 2) Sea $D : P_2 \to P_1$ la derivación restringida a P_2 (Polinomios de grado ≤ 2) con codominio P_1. Tenemos, en las bases $e = (1, t, t^2)$, $e' = (1, t)$ de P_2 y P_1 respectivamente,

$$[D]^e_{e'} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

ya que $D(e_1) = 0$, $D(e_2) = 1 = e'_1$, $D(e_3) = 2t = 2e'_2$. D así definido es claramente un epimorfismo, pues $I(D) = P_1$. Una inversa a derecha de D es la transformación $S : P_1 \to P_2$ definida como la integral $S(p(t)) = \int_0^t p(t')dt'$, con $S(e'_1) = t = e_2$, $S(e'_2) = t^2/2 = e_3/2$, representada por la matriz

$$[S]^{e'}_e = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1/2 \end{pmatrix}$$

Es fácil ver que S es un monomorfismo y que es una inversa a derecha de D, pues

$$[D]^e_{e'}[S]^{e'}_e = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1/2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

es decir,

$$DS = I_{P_1}$$

Precisamente, si $A = [S]^{e'}_e \Rightarrow$ la ec. (1) nos da $B = [D]^e_{e'}$. Y si $B = [D]^e_{e'}$, la ec. (2) nos da $A = [S]^{e'}_e$, como el lector puede fácilmente verificar.

No obstante, notemos que $[S]_e^{e'}[D]_{e'}^e = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \neq I_3$, por lo que $SD \neq I_{P_2}$.

Notemos también que $S' : P_1 \to P_2$ definida por $S'(e_1) = t + a$, $S'(e_2) = t^2/2 + b$, y representada por

$$[S']_e^{e'} = \begin{pmatrix} a & b \\ 1 & 0 \\ 0 & 1/2 \end{pmatrix}$$

es *también* inversa a derecha de $[D]_{e'}^e$ $\forall\, a, b$, como puede verificarse fácilmente, y que

$$[D']_{e'}^e = \begin{pmatrix} c & 1 & 0 \\ d & 0 & 2 \end{pmatrix}$$

es también una inversa a izquierda de $[S]_e^{e'}$ $\forall\, c, d$, de modo que éstas no son únicas.

7.1 Solución general de una ecuación lineal

Como aplicación fundamental, consideremos, para $F : V \to V'$ lineal, la ecuación general

$$F(v) = v'$$

donde se trata de encontrar el conjunto de vectores $v \in V$ que satisfacen $F(v) = v'$.
Si $v' = 0$, la ecuación se denomina homogénea y el conjunto de soluciones es el espacio nulo $N(F)$.
Si $v' \neq 0$, la ecuación se denomina no homogénea. En tal caso el conjunto de soluciones de la ecuación no homogénea *no es un espacio vectorial* (por ej. 0 no pertenece al conjunto pues $F(0) = 0 \neq v'$). Se cumple en cambio que si $F(v_1) = v_1'$ y $F(v_2) = v_2' \Rightarrow$

$$F(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 v_1' + \alpha_2 v_2'$$

o sea, la combinación lineal $\alpha_1 v_1 + \alpha_2 v_2$ es solución de la ecuación $F(v) = \alpha_1 v_1' + \alpha_2 v_2'$.

Es obvio además que existirá solución si y sólo si $v' \in I(F)$.
Supongamos ahora que existan dos soluciones v_1, v_2, tal que $F(v_1) = F(v_2) = v'$. Entonces

$$0 = F(v_1) - F(v_2) = F(v_1 - v_2)$$

por lo que $v_2 - v_1 \in N(F)$. Por lo tanto, $v_2 = v_1 + v_n$, donde $v_n \in N(F)$. Es decir, si v_p es una solución particular que satisface $F(v_p) = v' \Rightarrow$ cualquier otra solución v es de la forma

$$v = v_p + v_n$$

donde $v_n \in N(F)$ es un vector del espacio nulo de F, es decir, una solución de la ec. *homogénea* ($F(v_n) = 0$). La solución *general* estará dada entonces por la suma de una solución particular v_p de la ecuación no homogénea y de una solución v_n de la ecuación homogénea.

Resulta claro entonces que si $v' \in I(F)$, la solución será única si y sólo si $N(F) = 0$, o sea, si y sólo si F es un monomorfismo. En tal caso, la única solución de $F(v) = v'$ puede encontrarse como

$$v = G(v')$$

donde G *es una inversa a izquierda de* F. En efecto, si $F(v) = v' \Rightarrow G(F(v)) = (GF)(v) = I_V(v) = v = G(v')$. Puede utilizarse cualquier inversa a izquierda ya que difieren entre sí sólo para vectores que no pertenecen a $I(F)$ ($G_2(v') = G_1(v')$ si $v' \in I(F)$).

Por otro lado, si $F : V \to V'$ es un epimorfismo $\Rightarrow$ la ecuación $F(v) = v'$ tendrá siempre solución, pero no será única a no ser que $N(F) = \{0\}$ (en cuyo caso F es isomorfismo). La solución general será

$$v = G(v') + v_n$$

donde G es una inversa a derecha de F y v_n un vector arbitrario de $N(F)$. En efecto, $F(G(v')+v_n) = F(G(v'))+F(v_n) = (FG)(v')+0 = I_{V'}(v') = v'$. Aquí $G(v')$ representa la solución particular v_p, la cual, remarquemos, *es una función lineal* de v'. Puede utilizarse cualquier inversa a derecha pues estas difieren sólo en un vector de v_n ($G_2(v') = G_1(v') + v_n$ con $v_n \in N(F)$).

Finalmente, si F es un isomorfismo, existe una única solución $\forall\ v' \in V'$ dada por

$$v = F^{-1}(v')$$

con F^{-1} la (única) inversa de F, que es a la vez inversa a izquierda y derecha.

Ejemplo 1): Para el caso de sistemas de m ecuaciones lineales con n incógnitas, dados por

$$AX = Y$$

con A de $m \times n$, X de $n \times 1$, Y de $m \times 1$ y A y Y de elementos reales (que corresponde a la función $F : \mathbb{R}^{n\times 1} \to \mathbb{R}^{m\times 1}$ dada por $F(X) = AX$) los resultados anteriores implican que:
1) El sistema posee solución si y sólo si Y pertenece al espacio columna de A.
En tal caso, la solución general será de la forma

$$X = X_p + X_n$$

donde X_p es una solución particular ($AX_p = Y$) y X_n una solución general del sistema homogéneo ($AX_n = 0$, con 0 el vector columna nulo).

2) La solución será única si y sólo si el espacio nulo de A es el vector columna nulo ($X_n = 0$), es decir, si y sólo si Rango $(A) = n$ (y por lo tanto, $m \geq n$). En este caso F es un monomorfismo y la única solución (en el caso que Y pertenezca al espacio columna de A) puede encontrarse como $X = BY$, con B de $n \times m$ una inversa a izquierda de A ($BA = I_n$).

3) Si la dimensión del espacio columna es m (en cuyo caso Rango $(A) = m$ y por lo tanto, $m \leq n$) existirá solución para cualquier Y de $m \times 1$. En este caso F es un epimorfismo y la solución general puede escribirse como $X = CY + X_n$, con C de $n \times m$ una inversa a derecha de A ($AC = I_m$) y X_n solución del sistema homogéneo $AX_n = 0$.

4) Si $m = n$ y Rango $(A) = n \Rightarrow$ existe siempre una única solución dada por $X = A^{-1}Y$, con A^{-1} la inversa de A. En este caso F representa un isomorfismo.

Ejemplo 2): Resolver el sistema $x + y = a$, $4x + 2z = b$, en forma matricial. Corresponde a

$$\begin{pmatrix} 1 & 1 & 0 \\ 4 & 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}$$

Como el rango de la matriz es 2, existirá solución $\forall\, a, b$. Reduciendo por filas el sistema ampliado y llevándolo a la forma de Gauss-Jordan, se obtiene

$$\left(\begin{array}{ccc|c} 1 & 1 & 0 & a \\ 4 & 0 & 2 & b \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 0 & \frac{1}{2} & \frac{1}{4}b \\ 0 & 1 & -\frac{1}{2} & a - \frac{1}{4}b \end{array}\right)$$

de donde se lee la solución general $x = b/4 - z/2$, $y = a - b/4 + z/2$, con z libre. Podemos escrbir esta solución como

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{4} \\ 1 & -\frac{1}{4} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} + t \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 1 \end{pmatrix}$$

con $t \in \mathbb{R}$ arbitrario (parámetro libre), donde $B = \begin{pmatrix} 0 & \frac{1}{4} \\ 1 & -\frac{1}{4} \\ 0 & 0 \end{pmatrix}$ es precisamente *una inversa a derecha de la matriz de coeficientes* $A = \begin{pmatrix} 1 & 1 & 0 \\ 4 & 0 & 2 \end{pmatrix}$ y $X = t \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 1 \end{pmatrix}$ un elemento arbitrario del espacio nulo de A

$(AX = 0)$. Nótese que

$$\begin{pmatrix} 1 & 1 & 0 \\ 4 & 0 & 2 \end{pmatrix} \begin{pmatrix} 0 & \frac{1}{4} \\ 1 & -\frac{1}{4} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

pero

$$\begin{pmatrix} 0 & \frac{1}{4} \\ 1 & -\frac{1}{4} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 4 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 \end{pmatrix}$$

8. Operadores de Proyección

Recordemos que si S_1 es un subespacio de V y S_2 un suplemento tal que $V = S_1 \oplus S_2$, todo vector $v \in V$ puede escribirse en forma única como $v = v_1 + v_2$, con $v_1 \in S_1$, $v_2 \in S_2$.
El proyector P_{S_1/S_2} sobre S_1 en la dirección de S_2 (recordar la interpretación geométrica dada en clase) queda entonces definido por

$$P_{S_1/S_2}(v) = v_1$$

y satisface por lo tanto $P^2_{S_1/S_2} = P_{S_1/S_2}$, pues $P^2_{S_1/S_2}(v) = P_{S_1/S_2}(P_{S_1/S_2}(v)) = P_{S_1/S_2}(v_1) = v_1$. Obviamente, su núcleo e imagen son $N(P_{S_1/S_2}) = S_2$, $I(P_{S_1/S_2}) = S_1$. Depende de S_1 y S_2.

Análogamente, si $P^2 = P$, P es un proyector sobre $S_1 = I(P)$ en la dirección de $S_2 = N(P)$, con $S_1 \oplus S_2 = V$, es decir $P = P_{I(P)/N(P)}$.
En efecto, para cualquier $v \in V$, $v = P(v) + v - P(v) = v_1 + v_2$, con $v_1 = P(v) \in I(P)$ y $v_2 = v - P(v) \in N(P)$ pues $P(v - P(v)) = P(v) - P^2(v) = P(v) - P(v) = 0$. Además $I(P) \cap N(P) = \{0\}$ pues si $v = P(v')$ y $P(v) = 0 \Rightarrow 0 = P^2(v') = P(v') = v$. Por lo tanto $V = I(P) \oplus N(P)$.
Finalmente $P(v) = P(v_1 + v_2) = P(v_1) + P(v_2) = P^2(v) + 0 = P(v) = v_1$, por lo que P es proyector sobre $S_1 = I(P)$ en la dirección de $S_2 = N(P)$. Esto incluye los casos triviales $I(P) = V$, en cuyo caso $N(P) = \{0\}$ y por lo tanto $P_{S_1/S_2} = I_V$ (operador identidad), y $I(P) = \{0\}$, en cuyo caso $N(P) = V$ y $P = 0$ (operador nulo). Dado que $v = v_1 + v_2 = P_{S_1/S_2}(v) + P_{S_2/S_1}(v)$, se cumple siempre

$$P_{S_1/S_2} + P_{S_2/S_1} = I_V$$

Si $\dim V = n$, la representación matricial de P_{S_1/S_2} en una base ordenada $B = (b_1, \ldots, b_k, b_{k+1}, \ldots, b_n)$, en la que $(b_1, \ldots, b_k)$ es base de S_1 y $(b_{k+1}, \ldots, b_n)$ es base de S_2, es de la forma

$$[P_{S_1/S_2}]^B_B = \begin{pmatrix} I_{k\times k} & 0_{k\times m} \\ 0_{m\times k} & 0_{m\times m} \end{pmatrix}$$

donde $I_{k\times k}$ denota la matriz identidad de $k \times k$, $m = n - k$ y 0_{pq} la matriz nula de $p \times q$. Es claro que P_{S_1/S_2} es un operador singular ($\mathrm{Det}[P_{S_1/S_2}] = 0$), salvo en el caso $S_1 = V$ ($S_2 = \{0\}$, $P = I_V$).

Los proyectores ortogonales son aquellos en los que S_2 es el subespacio ortogonal a S_1, tema que veremos en detalle más adelante. En general, S_2 puede ser cualquier suplemento de S_1, no necesariamente el ortogonal, por lo que la matriz que representa a P_{S_1/S_2} puede no ser simétrica o hermítica en una base ortonormal.

Ej.: Consideremos, para $V = \mathbb{R}^2$, el proyector sobre el subespacio generado por $b_1 = (1, 0)$ en la dirección del subespacio generado por $b_2 = (1, 1)$. Como $P(b_1) = b_1$, $P(b_2) = 0$, en la base $B = (b_1, b_2)$ se obtiene

$$[P_{S_1/S_2}]^B_B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

En la base canónica $B_c = (e_1, e_2)$, con $e_1 = (1, 0)$, $e_2 = (0, 1)$, se obtiene entonces

$$[P_{S_1/S_2}]^{B_c}_{B_c} = S[P_{S_1/S_2}]^B_B S^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$$

donde $S = [I]^B_{B_c} = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ y $S^{-1} = \left(\begin{smallmatrix} 1 & -1 \\ 0 & 1 \end{smallmatrix}\right) = [I]^{B_c}_B$. Las columnas de $[P_{S_1/S_2}]^{B_c}_{B_c}$ son proporcionales a $[b_1]_{B_c} = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$.
Para $v = (x, y)$ obtenemos entonces $[P_{S_1/S_2}(v)]_{B_c} = [P_{S_1/S_2}]^{B_c}_{B_c}\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = \left(\begin{smallmatrix} x-y \\ 0 \end{smallmatrix}\right)$, es decir, $P_{S_1/S_2}(x, y) = (x - y, 0)$ en acuerdo con $(x, y) = (x - y)(1, 0) + y(1, 1)$ (Recordar dibujo).
El proyector ortogonal usual sobre el subespacio S_1 generado por $e_1 = b_1$ es $P_{S_1} \equiv P_{S_1/S_1^\perp}$, donde $S_1^\perp$ es el subespacio ortogonal, generado por ej. por e_2. Obtenemos $[P_{S_1}]^{B_c}_{B_c} = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ y por lo tanto $P(x, y) = (x, 0)$, de acuerdo a $(x, y) = x(1, 0) + y(0, 1)$.

9. Autovalores y Autovectores

Sea V un espacio vectorial sobre el cuerpo K y sea $F : V \to V$ un operador lineal. Un escalar $\lambda \in K$ es un *autovalor* de F si existe $v \in V$, con $v \neq 0$, tal que

$$F(v) = \lambda v \qquad (v \neq 0)$$

En tal caso v es un *autovector* de F correspondiente al autovalor λ. Sinónimo de autovalor es *valor propio* (en inglés "eigenvalue", donde "eigen" proviene del alemán y significa propio) y sinónimo de autovector es *vector propio* ("eigenvector" en inglés).
La acción de F sobre un autovector es pues la multiplicación por un escalar. Por ejemplo, si $V = \mathbb{R}^n$, esto implica que $v \neq 0$ es autovector de F si y sólo si $w = F(v)$ tiene la misma dirección que v (aunque no necesariamente el mismo sentido) o es nulo. Como veremos, el conocimiento del conjunto de los autovalores y autovectores de un operador permite conocer su estructura en detalle.

El concepto de autovalor y autovector de un operador lineal tiene una importancia fundamental en Física, especialmente en Mecánica Cuántica. Mencionemos que en la misma los observables físicos corresponden a operadores lineales (de características determinadas) en un cierto espacio (el de los vectores que representan los estados del sistema) y las cantidades medibles son precisamente los autovalores de dichos operadores. E inmediatamente después de una medición, el sistema cuántico queda en un estado que es autovector del operador correspondiente al autovalor medido.
El concepto de autovalor y autovector es también fundamental para el tratamiento de sistemas descriptos por ecuaciones diferenciales lineales acopladas, tales como osciladores armónicos acoplados, ya sean clásicos o cuánticos, donde permite entender el concepto de desacoplamiento y modo normal. Otras aplicaciones incluyen, como veremos, la determinación de los ejes principales de inercia en un cuerpo rígido y la resolución de sucesiones definidas mediante relaciones recursivas lineales, tales como la famosa sucesión de Fibonacci.

Veamos algunas consecuencias inmediatas de tal definición:
10.1) El conjunto de autovectores de F correspondientes a un determinado autovalor λ, junto con el vector nulo 0, forma un subespacio de V denominado *espacio propio* o autoespacio ("eigenspace" en inglés) de F correspondiente al autovalor λ, que denotaremos como $V_F(\lambda)$ o simplemente, $V(\lambda)$.
En efecto, si $v \neq 0$ es autovector de F corresp. al autovalor λ, y $\alpha \in K$, $F(\alpha v) = \alpha F(v) = \alpha\lambda v = \lambda(\alpha v)$ $\forall$ α, por lo que αv es también autovector de F corresp. a λ si $\alpha \neq 0$
Y si v_1 y v_2 son autovectores corresp. al mismo autovalor λ, $F(v_1 + v_2) = F(v_1) + F(v_2) = \lambda v_1 + \lambda v_2 = \lambda(v_1 + v_2)$, por lo que $v_1 + v_2$ es también autovector corresp. al autovalor λ (si $v_1 + v_2 \neq 0$).
La dimensión de $V_F(\lambda)$ es como mínimo uno (ya que al menos existe un autovector no nulo).

9.2) El espacio propio $V_F(\lambda)$, con λ autovalor de F, es el núcleo del operador $F - \lambda I$:

$$V_F(\lambda) = N[F - \lambda I]$$

donde I denota el operador identidad en V ($I(v) = v \ \forall \ v \in V$).
En efecto, si $F(v) = \lambda v \Rightarrow 0 = F(v) - \lambda v = (F - \lambda I)(v)$, y si $(F - \lambda I)(v) = 0 \Rightarrow F(v) = \lambda v$.
Por lo tanto $\lambda \in K$ es autovalor de F si y sólo si

$$N[F - \lambda I] \neq \{0\}$$

En particular, $\lambda = 0$ es autovalor de F si y sólo si $N(F) \neq \{0\}$, es decir, si y sólo si F no es monomorfisomo. En tal caso $V_F(0) = N(F)$.

9.3) Si V es de dimensión finita n, $\lambda \in K$ es autovalor si y sólo si

$$\mathrm{Det}[F - \lambda I] = 0$$

En efecto, si λ es autovalor, $N[F - \lambda I] \neq \{0\}$, por lo que $F - \lambda I$ no es monomorfismo. La matriz que representa a $F - \lambda I$ debe ser entonces singular en cualquier base, y por lo tanto, su determinante nulo. Por otro lado, si $\mathrm{Det}[F - \lambda I] = 0$, $F - \lambda I$ no es monomorfismo, y por lo tanto existe $v \neq 0$ tal que $(F - \lambda I)(v) = 0$.
Recordemos que si B es una base de V,

$$\mathrm{Det}[F - \lambda I] = |[F]_B - \lambda I_n|$$

donde $[F]_B \equiv [F]_B^B$ es la matriz que representa a F en dicha base, $I_n = [I]_B$ es la matriz identidad de $n \times n$ y $|A| = \mathrm{Det}A$ denota el determinante de la matriz A. $\mathrm{Det}[F - \lambda I]$ es *independiente* de la base elegida: Si B' es otra base de V, y $S = [I]_B^{B'}$,

$$|[F]_{B'} - \lambda I_n| = |S^{-1}[F]_B S - \lambda I_n| = |S^{-1}([F]_B - \lambda I_n)S| = |S^{-1}||[F]_B - \lambda I_n||S| = |[F]_B - \lambda I_n|$$

ya que $[I]_{B'} = [I]_B = I_n$ y $|S^{-1}| = 1/|S|$. Los autovalores de F en un espacio de dimensión finita n se obtienen entonces como las raíces pertenecientes al cuerpo K del polinomio

$$P(\lambda) = \mathrm{Det}[F - \lambda I]$$

denominado *polinomio característico*, que es de grado n (pues $[F - \lambda I]_e$ es de $n \times n$). La ecuación $P(\lambda) = 0$ se denomina *ecuación caracerística* y posee, por lo tanto, a lo sumo n *raíces distintas*, que en general pueden ser complejas. Serán autovalores si pertenecen al cuerpo K. Si $K = \mathbb{C} \Rightarrow$ *toda raíz* de $P(\lambda)$ es autovalor.

9.4) *Teorema: Los autovectores de F correspondientes a autovalores distintos son LI*
Demostraremos el teorema por inducción. Para $n = 1$, v_1 autovector es LI pues es no nulo (mejor comprensión se logra comenzando con $n = 2$: Si $v_1 \neq 0$ y $v_2 \neq 0$ son autovectores correspondientes a autovalores *distintos* $\Rightarrow$ son LI pues de lo contrario $v_2 = \alpha v_1$, y correspondería por 9.1 al mismo autovalor que v_1). Supongamos ahora que $v_1, \ldots, v_{k-1}$ son autovectores LI, con autovalores $\lambda_1, \ldots, \lambda_{k-1}$, y que $v_k \neq 0$ es autovector con autovalor λ_k, siendo λ_k distinto a todos los anteriores. Si

$$0 = \alpha_1 v_1 + \ldots + \alpha_{k-1} v_{k-1} + \alpha_k v_k$$

entonces

$$0 = F(0) = \alpha_1 F(v_1) + \ldots + \alpha_{k-1} F(v_{k-1}) + \alpha_k F(v_k) = \alpha_1 \lambda_1 v_1 + \ldots \alpha_{k-1} \lambda_{k-1} v_{k-1} + \alpha_k \lambda_k v_k$$

Restando ahora la primera ecuación mult. por λ_k,

$$0 = (\lambda_1 - \lambda_k)\alpha_1 v_1 + \ldots + (\lambda_{k-1} - \lambda_k)\alpha_{k-1} v_{k-1}$$

que implica $\alpha_1 = \ldots = \alpha_{k-1} = 0$ por ser $v_1, \ldots v_{k-1}$ LI y $\lambda_k \neq \lambda_i$ para $i = 1, \ldots, k-1$. Por lo tanto, también $\alpha_k = 0$ y entonces $v_1, \ldots, v_k$ son L.I.

Nótese que la demostración es igualmente válida si los $\lambda_1, \ldots, \lambda_{k-1}$ no son todos distintos (pero sí distintos a λ_k) siempre que los $v_1, \ldots, v_{k-1}$ sean LI
Por lo tanto, *ningún elemento $v_k \neq 0$ de $V(\lambda_k)$ puede ser generado por autovectores correspondientes a autovalores distintos de λ_k.*

9.5) Si $\lambda_1 \neq \lambda_2 \Rightarrow V_F(\lambda_1) \cap V_F(\lambda_2) = \{0\}$
Es consecuencia inmediata del último párrafo. Si $v \in V_F(\lambda_2)$ y $v \in V_F(\lambda_1)$, entonces $F(v) = \lambda_1 v = \lambda_2 v$, por lo que $(\lambda_1 - \lambda_2)v = 0$ y por lo tanto $v = 0$. Esto implica que $V_F(\lambda_1) + V_F(\lambda_2) = V_F(\lambda_1) \oplus V_F(\lambda_2)$.
Del mismo modo, la intersección de $V_F(\lambda_k)$ con la suma $V_F(\lambda_1) + \ldots + V_F(\lambda_{k-1})$ es $\{0\}$ si λ_k es distinto a todos los autovalores anteriores. Si así no fuese existiría un vector $v \in V_F(\lambda_k)$ con $v \neq 0$ que puede ser escrito como combinación lineal de autovectores correspondientes a autovalores distintos, pero esto es imposible por el teorema 9.4 anterior.
Podemos entonces escribir la suma de espacios propios como suma directa $V_F(\lambda_1) \oplus \ldots \oplus V_F(\lambda_k)$.

9.6) Un operador F en un espacio V de dimensión finita se dice *diagonalizable* si existe una base formada por autovectores de F.
En tal caso, denotando esta base como $B' = (b'_1, \ldots, b'_n)$, con $F(b'_i) = \lambda_i b'_i$, $i = 1, \ldots, n$, la matriz que representa a F en dicha base es *diagonal*:

$$[F]_{B'} = \begin{pmatrix} & \ldots & \\ [F(b'_1)]_{B'} & \ldots & [F(b'_n)]_{B'} \\ & \ldots & \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ & & \ldots & \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}$$

Recíprocamente, si $[F]_{B'}$ es diagonal $\Rightarrow F(b'_i) = \lambda_i b'_i$ y B' es necesariamente una base de autovectores. Si F es diagonalizable y B es una base arbitraria de V, tenemos

$$[F]_{B'} = S^{-1}[F]_B S$$

con $[F]_{B'}$ *diagonal* y $S = [I]_B^{B'}$ la matriz de cambio de base, *por lo que existe una matriz no singular S tal que $S^{-1}[F]_B S$ es diagonal*. La columna i de S es el vector de coordenadas $[b'_i]_B$ del autovector b'_i corresp. al autovalor λ_i en la base original B. Esta es la forma de construir la matriz diagonalizante S.

Consecuencia Importante de 9.4: Si $P(\lambda)$ posee n raíces distintas $\lambda_i \in K$, $\Rightarrow F$ es diagonalizable.
En efecto, en tal caso existirán n autovectores LI (uno por cada autovalor) que serán por tanto base de V. La dimensión de cada espacio propio $V_F(\lambda_i)$ es en este caso 1, que es igual a la multiplicidad de cada raíz.

9.7) Teorema: *F es diagonalizable si y sólo si i) todos las raíces de $P(\lambda)$ pertenecen al cuerpo y ii) la dimensión del espacio propio $V_F(\lambda_i)$ correspondiente a la raíz λ_i es igual a la multiplicidad m_i de dicha raíz.*
La dimensión del espacio propio $d_i = \dim V_F(\lambda_i)$ es el máximo número de autovectores LI que pueden obtenerse para un mismo autovalor λ_i, y se denomina también *multiplicidad geométrica* de λ_i.
Demostración: Supongamos F diagonalizable. En una base B' en la que $[F]_{B'}$ es diagonal, tenemos

$$P(\lambda) = (\lambda_1 - \lambda) \dots (\lambda_n - \lambda)$$

La multiplicidad m_i de una raíz λ_i será pues igual al número de veces que λ_i se repite en la diagonal. Pero este número es igual al número de vectores de la base B' que tienen a λ_i como autovalor, que es precisamente la dimensión del espacio propio. Nótese también que d_i es el número de filas nulas de $[F]_{B'} - \lambda_i I_n$.
Por otro lado, si la dimensión de cada espacio propio es igual a la multiplicidad m_i de la raíz λ_i, la suma directa de todos los espacios propios correspondientes a autovalores *distintos*, $V_F(\lambda_1) \oplus \ldots \oplus V_F(\lambda_k)$ tendrá dimensión $d_1 + \ldots + d_k = m_1 + \ldots + m_k = n$ (ya que la suma de todas las multiplicidades es igual al grado del polinomio), por lo que será igual al espacio V. Existe entonces una base formada por autovectores de F.

9.8) En general, la dimensión del espacio propio $V_F(\lambda_i)$ puede ser igual o menor que la multiplicidad de la raíz λ_i: $\dim\, V_F(\lambda_i) \leq m_i$.
En efecto, eligiendo una base e donde los primeros $d_i = \dim V_F(\lambda_i)$ elementos formen una base de $V_F(\lambda_i)$, las primeras d_i columnas de $[F]_e$ tendrán elementos no nulos sólo en la diagonal y por lo tanto $P(\lambda) = \mathrm{Det}[[F]_e - \lambda I_n] = (\lambda_i - \lambda)^{d_i} Q(\lambda)$, con $Q(\lambda)$ un polinomio de grado $n - d_i$, por lo que m_i será como mínimo d_i ($m_i = d_i$ si $Q(\lambda_i) \neq 0$ y $m_i > d_i$ si $Q(\lambda_i) = 0$).
Si $d_i < m_i$, *F no es diagonalizable* (aún tomando como cuerpo $\mathbb{C}$).

Ejemplo 1 (Casos Triviales): Si $I : V \to V$ es el operador identidad $\Rightarrow I(v) = v \;\forall\; v \in V$ por lo que su único autovalor es 1. El espacio propio correspondiente es V ($V_F(1) = V$).
Si V es de dimensión finita n, puede llegarse a la misma conclusión notando que

$$P(\lambda) = \mathrm{Det}[I - \lambda I] = |I_n - \lambda I_n| = |(1 - \lambda)I_n| = (1 - \lambda)^n$$

$\lambda = 1$ es pués la única raíz de $P(\lambda)$, y posee multiplicidad n, que es igual a la dimensión del espacio propio (el mismo V). I es por lo tanto diagonalizable (caso trivial).

Si $0 : V \to V$ es el operador nulo $\Rightarrow 0(v) = 0 = 0v \;\forall\; v \in V$ por lo que su único autovalor es 0, y el espacio propio correspondiente es V ($V_F(0) = V$).
Si V es de dimensión $n \Rightarrow [0]_B = 0$ (la matriz nula) en cualquier base y por lo tanto $P(\lambda) = \mathrm{Det}[0 - \lambda I_n] = |-\lambda I_n| = (-\lambda)^n$, por lo que 0 es la única raíz, con multiplicidad n. 0 es también trivialm. diagonalizable.

Ejemplo 2: Sea $F : \mathbb{R}^2 \to \mathbb{R}^2$ la reflexión respecto del eje x. Si $B = (e_1, e_2)$ es la base canónica, con $e_1 = (1, 0)$, $e_2 = (0, 1)$, sabemos que $F(e_1) = e_1$, $F(e_2) = -e_2$. Por lo tanto e_1 es autovector de F con autovalor 1 y e_2 autovector de F con autovalor -1. No pueden existir otros autovalores pues la dimensión de V es 2. $V_F(1)$ es entonces el espacio generado por e_1, es decir, el conjunto de vectores $(x, 0)$, con $x \in \mathbb{R}$, sobre los que F actúa como identidad, y $V_F(-1)$ el generado por e_2, es decir, el conjunto de vectores $(0, y)$, con $y \in \mathbb{R}$, para los que la acción de F es la inversión de sentido.
Podemos obtener el mismo resultado a partir de la representación matricial

$$[F]_B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

que ya es diagonal, por lo que los autovalores son 1 y -1: Tenemos $P(\lambda) = |[F]_B - \lambda I_2| = (1 - \lambda)(-1 - \lambda)$, siendo entonces las raíces ± 1.

Ejemplo 3: Sea $F : \mathbb{R}^2 \to \mathbb{R}^2$ la reflexión respecto de la recta de ecuación $y = x$, dada por (recordar ejemplo dado) $F(x, y) = (y, x)$. Si $B' = (b'_1, b'_2)$, con $b'_1 = (1, 1)$, $b'_2 = (-1, 1)$, tenemos $F(b'_1) = b'_1$, $F(b'_2) = -b'_2$, por lo que los autovalores son nuevamente 1 y -1, con $V_F(1)$ el espacio generado por b'_1 y $V_F(-1)$ el espacio generado por b'_2. Se obtiene entonces

$$[F]_B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad [F]_{B'} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = S^{-1}[F]_B S\,, \quad S = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

El pol. característico es $P(\lambda) = |[F]_B - \lambda I_2| = \lambda^2 - 1 = (1 - \lambda)(-1 - \lambda) = |[F]_{B'} - \lambda I_2|$ y sus raíces ± 1.

Ejemplo 4: **Operador de Proyección:** Si $P^2 = P \Rightarrow$ los únicos autovalores posibles de P son 0 y 1, con $V_P(0) = N(P)$ si $N(P) \neq \{0\}$ y $V_P(1) = I(P)$ si $I(P) \neq \{0\}$.
Dem.: Hemos visto que si $P^2 = P \Rightarrow P$ es un proyector sobre $I(P)$ en la dirección de $N(P)$, con $I(P) \oplus N(P) = V$ y $P(v) = v$ si $v \in I(P)$ y $P(v) = 0 = 0v$ si $v \in N(P)$. Por lo tanto, los autovalores son: 1 (si $I(P) \neq \{0\}$, es decir, si $P \neq 0$) y 0 (si $N(P) \neq \{0\}$, es decir, si $P \neq I_V$). P es pues dagonalizable, siendo una base de autovectores la formada por la unión de una base de la imagen $I(P)$ y una del núcleo $N(P)$. P es pues siempre diagonalizable.
Estas propiedades pueden también demostrarse directamente: Si $P(v) = \lambda v$, con $v \neq 0 \Rightarrow P^2(v) = P(P(v)) = P(\lambda v) = \lambda P(v) = \lambda^2 v$, pero también $P^2(v) = P(v) = \lambda v$, por lo que $\lambda^2 = \lambda$, es decir $\lambda(\lambda - 1) = 0$. Esto implica $\lambda = 0$ o $\lambda = 1$. Si $\exists\, v \neq 0$ tal que $P(v) = 0 \Rightarrow 0$ es autovalor y $V_P(0) = N(P)$. Si $\exists\, v \neq 0$ tal que $P(v) = v \Rightarrow 1$ es autovalor y $V_P(1) = I(P)$. En efecto, si $v \neq 0$ y $P(v) = v \Rightarrow v \in I(P)$, y si $v \in I(P)$ $\Rightarrow v = P(v')$ y $P(v) = P(P(v')) = P^2(v') = P(v') = v$, por lo que $v \in V_P(1)$.

Ejemplo 6: **Potencias de operadores lineales.** Si v es autovector de F con autovalor $\lambda \Rightarrow v$ es también autovector de F^k con autovalor λ^k para cualquier $k > 0$ natural, y también para cualquier $k < 0$ entero si F es invertible (o sea, automorfismo), en cuyo caso $\lambda \neq 0$ (pues necesariamente $N(F) = \{0\}$; ver 9.4).
Dem.: Si $k = 2$, $F^2(v) = F(F(v)) = F(\lambda v) = \lambda F(v) = \lambda^2 v$. La demostración para $k > 2$ es análoga y puede hacerse fácilmente por inducción: $F^k(v) = F(F^{k-1}(v)) = F(\lambda^{k-1} v) = \lambda^{k-1} F(v) = \lambda^k v$.
Si $k = -1$, $F^k = F^{-1}$ es la inversa de F y $v = F^{-1}F(v) = F^{-1}(\lambda v) = \lambda F^{-1}(v)$, por lo que $F^{-1}(v) = \lambda^{-1} v$.
El resultado para $F^{-k} \equiv (F^{-1})^k$ y $k > 1$ es entonces obvio: $F^k(v) = \lambda^k v\ \forall\, k \in \mathbb{Z}$ (véase también 9.4)

Ejemplo 7: Si λ es un autovalor de $F \Rightarrow \alpha\lambda + c$ es autovalor del operador $\alpha F + cI$, con $\alpha, I \in K$ e I el op. identidad: En efecto, si $F(v) = \lambda v$, $v \neq 0 \Rightarrow (\alpha F + cI)(v) = \alpha F(v) + cv = (\alpha\lambda + c)v$.

Ejemplo 8: Sea $D^2 : V \to V$ el operador derivada segunda en el espacio V de funciones $f : [0, a] \to \mathbb{R}$, $a > 0$, que son derivables a cualquier orden y satisfacen $f(0) = f(a) = 0$ (V es de dimensión infinita).
La ecuación $D^2(f) = \lambda f$ conduce a la ecuación diferencial $f'' = \lambda f$, cuya solución es $f(x) = c_1 \cos(sx) + c_2 \sin(sx)$ con $\lambda = -s^2$. La condición de contorno $f(0) = f(1) = 0$ implica $c_1 = 0$ y $\sin(sa) = 0$, o sea, $s = n\pi/a$, con $n > 0$ natural. Por lo tanto, los autovalores son $\lambda_n = -n^2\pi^2/a^2$ y los autovectores (llamados en este caso autofunciones) $f_n(x) = c_n \sin(n\pi x/a)$, con $n = 1, 2, \ldots$ y $c_n \neq 0$ arbitrario.

Ejemplo 9: La **ecuación de Schrödinger** estacionaria de la mecánica cuántica es

$$H|\Psi\rangle = E|\Psi\rangle$$

donde H es un operador lineal, $|\Psi\rangle$ un vector no nulo que representa el estado del sistema y E la energía del estado. Es pues una ecuación de autovalores: La energía E representa un autovalor de H y el vector $|\Psi\rangle$ el autovector correspondiente de H. Para una partícula en una dimensión, H toma la forma

$$H = -\frac{\hbar^2}{2m} D^2 + V(x)$$

donde D es el operador derivada anterior, $\hbar = h/(2\pi)$, con h la constante de Planck, m la masa de la partícula y $V(x)$ el potencial, con $|\Psi\rangle \to \psi(x)$, siendo $\psi(x)$ la "función de onda" (en realidad, $\psi(x) = \langle x|\Psi\rangle$, y la forma anterior de H es su representación efectiva en la base de autoestados $|x\rangle$ del operador posición: $\langle x|H|\Psi\rangle = (-\frac{\hbar^2}{2m} D^2 + V(x))\psi(x))$. H es el operador energía pues el impulso está representado por el operador $p = -i\hbar D$, por lo que el primer término de H es el operador energía cinética $P^2/(2m)$.
Si $V(x) = 0$ para $x \in [0, a]$ y $V(x) = \infty$ para $x > a$ o $x < 0$, la Ec. de Schrödinger se reduce al problema del ej. 8: Tenemos $H = -\frac{\hbar^2}{2m} D^2$ para $x \in [0, a]$, con $\psi(x) = 0$ para $x \geq a$ o $x \leq 0$, debiendo ser continua. Los autovalores son por lo tanto $E_n = \frac{\hbar^2 n^2 \pi^2}{2ma^2}$ y los autovectores $\psi_n(x) = c_n \sin(n\pi x/a)$.

10. Autovalores y Autovectores de Matrices

Todas las definiciones y propiedades anteriores se aplican igualmente al cálculo de autovalores y autovectores de matrices cuadradas, que pueden ser siempre consideradas como la representación de un cierto operador lineal (en un espacio vectorial de dimensión n) en una cierta base. Consideraremos en lo sucesivo $K = \mathbb{C}$.
Dada una matriz A de $n \times n$, λ es autovalor de A sii

$$|A - \lambda I_n| = 0$$

donde $|\dots|$ denota el determinante. El polinomio

$$P(\lambda) = |A - \lambda I_n| = \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ & & \dots & \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix} \qquad (10.1)$$

se denomina *polinomio característico* de la matriz A y es de grado n en λ, por lo que posee a lo sumo n raíces distintas.
Un vector columna $X = \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix}$ de $n \times 1$ es *autovector* de A correspondiente al autovalor λ si $X \neq 0$ y

$$AX = \lambda X$$

El conjunto de autovectores X corresp. a λ puede obtenerse resolviendo el sistema homogéneo

$$(A - \lambda I_n)X = 0$$

10.1) Las matrices semejantes poseen el mismo polinomio característico y por lo tanto los mismos autovalores. Recordemos que A es semejante o similar a B si $A = S^{-1}BS$, con S de $n \times n$ no singular.
En efecto $P_A(\lambda) = |A - \lambda I_n| = |S^{-1}BS - \lambda I_n| = |S^{-1}(B - \lambda I_n)S| = |B - \lambda I_n| = P_B(\lambda)$

Los autovectores no son en general los mismos: Si $AX = \lambda X$ y $A = S^{-1}BS$, el correspondiente autovector de B es SX, como el lector podrá fácilmente demostrar. Esto puede verse también directamente a partir del cambio de base asociado.

10.2) A y A^t (t denota traspuesta) poseen el mismo polinomio característico y por lo tanto los mismos autovalores (pero no necesariamente los mismos autovectores). Como $|B| = |B^t|$ $\forall$ B de $n \times n$,

$$P_{A^t}(\lambda) = |A^t - \lambda I_n| = |(A - \lambda I_n)^t| = |A - \lambda I_n| = P_A(\lambda)$$

10.3) Si A es real $\Rightarrow$ $P(\lambda)$ es real y por lo tanto sus raíces complejas aparecerán en pares conjugados:
Si λ es una raíz compleja, $0 = [P(\lambda)]^* = P(\lambda^*)$.
Además, si X es autovector de A con autovalor λ y A es real, entonces X^* es autovector de A con autovalor λ^*: Como $AX = \lambda X \Rightarrow (AX)^* = AX^* = \lambda^* X^*$.

10.4) A es una matriz singular si y sólo si A posee al menos un autovalor nulo.
Si A es singular $\Rightarrow$ $|A| = 0$ y por lo tanto, $|A - 0I_n| = 0$, por lo que 0 es autovalor. Análogamente, si $|A - 0I_n| = 0 \Rightarrow |A| = 0$ y por lo tanto, A es singular. Esto puede también deducirse directamente de 10.5

Un operador $F : V \to V$ en un espacio de dimensión finita tendrá pues un autovalor nulo si y sólo si no es un automorfismo. Notemos también que si $\mathrm{Det}[F] = 0$, el núcleo $N(F)$ no es otra cosa que el espacio propio correspondiente el autovalor 0: $N(F) = \{v | F(v) = 0\} = \{v | F(v) = 0.v\} = V_F(0)$

10.5) El determinante de una matriz es igual al producto de *todos* sus autovalores (reales y complejos, y elevados a sus respectivas multiplicidades):

$$|A| = \lambda_1 \lambda_2 \dots \lambda_n$$

En efecto, si $\lambda_1,\ldots,\lambda_n$ son las raíces de $P(\lambda)$, podemos escribir (utilizando la factorización en términos de raíces y notando que el término de grado n es $(-1)^n\lambda^n$)

$$P(\lambda)=|A-\lambda I_n|=(\lambda_1-\lambda)(\lambda_2-\lambda)\ldots(\lambda_n-\lambda) \qquad (10.2)$$

Por lo tanto, $|A|=P(0)=\lambda_1.\lambda_2\ldots\lambda_n$.
Esto implica que en un espacio de dimensión finita *el determinante de un operador lineal F es el producto de todos sus autovalores.*

10.6) La traza de una matriz es igual a la suma de todos sus autovalores (reales y complejos, y repetidos tantas veces como indica su multiplicidad):

$$\mathrm{Tr}A=\sum_{i=1}^{n}a_{ii}=\sum_{i=1}^{n}\lambda_i$$

A partir de la expresión (9.2) para $P(\lambda)$, vemos que el término de grado $n-1$ en λ es $(-\lambda)^{n-1}(\lambda_1+\ldots+\lambda_n)$, mientras que a partir de (9.1) vemos que el mismo es $(-\lambda)^{n-1}(a_{11}+\ldots+a_{nn})$. Como ambos son idénticos, se obtiene el resultado deseado.
Esto implica que la traza de un operador F es la suma de todos sus autovalores.

10.7 Diagonalización de Matrices

Una matriz A es diagonalizable si existe S no singular tal que

$$S^{-1}AS=A',\ \text{ con }\ A'\text{ diagonal}:\ A'=\begin{pmatrix}\lambda_1 & 0 & \ldots & 0\\ 0 & \lambda_2 & \ldots & 0\\ & & \ldots & \\ 0 & 0 & \ldots & \lambda_n\end{pmatrix}$$

es decir, si es semejante a una matriz diagonal.
En tal caso los elementos diagonales son necesariamente los autovalores de A, ya que

$$|A-\lambda I_n|=|A'-\lambda I_n|=(\lambda_1-\lambda)(\lambda_2-\lambda)\ldots(\lambda_n-\lambda)$$

y la columna i de S es autovector correspondiente al autovalor λ_i, pues $AS=SA'$:

$$S=\begin{pmatrix} & \ldots & \\ X_1 & \ldots & X_n\\ & \ldots & \end{pmatrix},\ \text{ con }\ AX_i=\lambda_i X_i,\ i=1,\ldots,n,\ \ X_i=\begin{pmatrix}x_{1i}\\ \ldots\\ x_{ni}\end{pmatrix}$$

Análogamente, si existen n vectores columna X_i LI tales que $AX_i=\lambda_i X_i,\quad i=1,\ldots,n$ entonces A es diagonalizable, con S la matriz de columnas X_i (que será invertible pues los X_i son LI y por lo tanto $|S|\neq 0$)

La dimensión del espacio propio correspondiente al autovalor λ_i es la dimensión del espacio nulo de $|A-\lambda_i I_n|$:

$$d_i=\dim V(\lambda_i)=\dim N[A-\lambda_i I_n]=n-R(A-\lambda_i I_n)$$

donde R denota el rango.
Notemos que si $P(\lambda)$ posee n raíces distintas $\Rightarrow$ A es diagonalizable, pues en tal caso existirán n vectores columna X_i LI tales que $AX_i=\lambda_i X_i$.
Si A es diagonalizable, resulta evidente que $|A|=|A'|=\lambda_1\ldots\lambda_n$ y que $\mathrm{Tr}A=\mathrm{Tr}A'=\lambda_1+\ldots+\lambda_n$, ya que el determinante y la traza de matrices semejantes son idénticas ($|S^{-1}AS|=|A|$, $\mathrm{Tr}S^{-1}AS=\mathrm{Tr}ASS^{-1}=\mathrm{Tr}A$).

Ejemplo 1: Consideremos la matriz

$$A=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}$$

que corresponde a la representación en la base canónica de la reflexión respecto de la recta $y=x$. Los autovalores se obtienen de la ecuación

$$|A-\lambda I_2|=\begin{vmatrix}-\lambda & 1\\ 1 & -\lambda\end{vmatrix}=1-\lambda^2=0$$

Por lo tanto, $\lambda = \pm 1$. El autovector X_1 correspondiente a $\lambda_1 = 1$ se obtiene resolviendo el sistema homogéneo $(A - \lambda_1 I_2)X_1 = 0$, es decir,

$$\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

que conduce a $x = y$. Los autovectores son entonces de la forma $x\binom{1}{1}$, con $x \neq 0$ y $V(1)$ es el espacio generado por $\binom{1}{1}$. Este corresponde al espacio generado por e'_1 en el ej. 3 anterior ($[e'_1]_e = \binom{1}{1}$).
Notemos que la matriz anterior tiene rango 1, por lo que $\dim V(1) = 2 - 1 = 1$.
El autovector correspondiente a $\lambda_2 = -1$ se obtiene resolviendo el sistema $(A - \lambda_2 I_2)X_2 = 0$, es decir,

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

que conduce a $x = -y$. Los autovectores son entonces de la forma $x\binom{1}{-1}$, con $x \neq 0$, y $V(-1)$ es el espacio generado por $\binom{1}{-1}$. Este corresponde al espacio generado por e'_2 en el ej. 3 anterior ($[e'_2]_e = \binom{1}{-1}$).
Una matriz de autovectores es entonces

$$S = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad \text{con} \quad S^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

y se verifica

$$S^{-1}AS = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

Notemos que se cumple $|A| = -1 = \lambda_1\lambda_2$ y $\mathrm{Tr}A = 0 = \lambda_1 + \lambda_2$.

Ejemplo 2: Consideremos

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

La ec. característica es

$$|A - \lambda I_2| = \begin{vmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{vmatrix} = (1-\lambda)^2 = 0$$

por lo que el único autovalor es $\lambda = 1$ con multiplicidad $m = 2$. No obstante, la matriz

$$A - 1I_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

posee rango 1, por lo que $\dim V(1) = 2 - 1 = 1 < 2$. Por lo tanto, esta matriz *no es diagonalizable*, ya que no existe una base de autovectores de la misma. La ecuación

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

conduce a $y = 0$, por lo que los autovectores son de la forma $x\binom{1}{0}$ y $V(1)$ es el espacio generado por $\binom{1}{0}$. No existe otro autovector LI de $\binom{1}{0}$. De todos modos, se cumple $|A| = 1 = 1.1$ y $\mathrm{Tr}A = 2 = 1 + 1$. Nótese que A es no singular ($|A| = 1 \neq 0$). La condición de no diagonalizable nada tiene que ver con la singularidad.

Cabe destacar, no obstante, que la matriz

$$B = \begin{pmatrix} 1 & 1 \\ \varepsilon & 1 \end{pmatrix}$$

es diagonalizable $\forall\ \varepsilon \neq 0$, ya que en tal caso la ecuación

$$|B - \lambda I_2| = (1-\lambda)^2 - \varepsilon = 0$$

posee siempre 2 raíces distintas: $\lambda = 1 \pm \sqrt{\varepsilon}$.
Esta conclusión es general: Si A no es diagonalizable podemos siempre encontrar una matriz B arbitrariamente próxima a A (es decir, cuyos elementos difieran de los de A en menos de ε, con $\varepsilon > 0$ arbitrario) tal que B es diagonalizable.

Ejemplo 3: Sea $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}$. Tenemos

$$|A - \lambda I_3| = \begin{pmatrix} 1-\lambda & 1 & 1 \\ 0 & 2-\lambda & 1 \\ 0 & 0 & 1-\lambda \end{pmatrix} = (1-\lambda)(2-\lambda)(1-\lambda)$$

por lo que las raíces de $P(\lambda)$ son $\lambda_1 = 1$ y $\lambda_2 = 2$, con multiplicidades 2 y 1 respectivamente. Si $\lambda = \lambda_1$,

$$|A - 1I_3| = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

posee rango 1, por lo que $\dim V(1) = \dim N(A - 1I_3) = 3 - 1 = 2$, igual a la multiplicidad de λ_1. La matriz es por lo tanto *diagonalizable* ya que necesariamente $\dim V(2) = 1$. El sistema $(A - 1I_3)X = 0$ conduce a $y + z = 0$, es decir, $y = -z$, con z y x arbitrarios, por lo que los autovectores para $\lambda_1 = 1$ son de la forma $\begin{pmatrix} x \\ -z \\ z \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$. Para $\lambda_2 = 2$,

$$|A - 2I_3| = \begin{pmatrix} -1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & -1 \end{pmatrix}$$

posee rango 2. El sistema $(A - 2I_3)X = 0$ conduce a $x = y$, con $z = 0$, por lo que los autovectores son de la forma $x \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$. Una matriz de autovectores es por lo tanto

$$S = \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \text{con} \quad S^{-1} = \begin{pmatrix} 1 & -1 & -1 \\ 0 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}$$

Se verifica entonces

$$A' = S^{-1}AS = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Nótese que el orden de los autovalores en A' corresponde al orden de los autovectores (columnas) en S.

Ejemplo 4: (Para hacer en la práctica) Sea $F : \mathbb{R}^2 \to \mathbb{R}^2$ el operador de rotación de ángulo θ (antihorario), con $[F]_{B_c} = \left(\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right)$, siendo B_c la base canónica. Es obvio que F no puede tener autovalores reales, ya que no existe $v \neq 0$ tal que $F(v) = \lambda v$. No es por lo tanto diagonalizable para $K = \mathbb{R}$.
Sin embargo, $[F]_{B_c}$ tiene autovalores complejos $\lambda = e^{\pm i\theta} = \cos\theta \pm i \sin\theta$ y es por lo tanto *diagonalizable* si se lo considera como $F : \mathbb{C}^2 \to \mathbb{C}^2$, con $K = \mathbb{C}$. Determine el lector los autovectores y compruebe que existe S tal que $S^{-1}[F]_{B_c}S$ es diagonal!

Ejemplo 5: Consideremos el operador de proyección ortogonal P sobre el plano $x + y = 0$ en $\mathbb{R}^3$ (o sea, proyección sobre este plano en la dirección del eje z). Si $B' = (b'_1, b'_2, b'_3)$, con $b'_1 = (e_1 + e_2)/\sqrt{2}$, $b'_2 = (-e_1 + e_2)/\sqrt{2}$, $b'_3 = b_3$, con $B_c = (e_1, e_2, e_3)$ la base canónica, entonces $P(b'_1) = 0$, $P(b'_2) = b'_2$, $P(b'_3) = b'_3$ y por lo tanto, $[P]_{B'} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Esta es pues una base de autovectores de P.
En la base canónica, obtenemos en cambio

$$[P]_{B_c} = S[P]_{B'}S^{-1} = \frac{1}{2}\begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad \text{con } S = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix}/\sqrt{2}$$

y $S^{-1}=S^t$. $[P]_{B_c}$ no es diagonal, aunque sigue cumpliendo que $[P]^2_{B_c}=[P]_{B_c}$. Se deja como ejercicio verificar explícitamente que los autovalores de la matriz $[P]_{B_c}$ son 0 y 1, y que una base de autovectores es precisamente B' (aunque por su puesto no es la única), de modo que S es una matriz diagonalizante de $[P]_{B_c}$, que verifica $S^{-1}[P]_{B_c}S=[P]_{B'}$, con $[P]_{B'}$ diagonal.

Ejemplo 6: **Autovalores de una matrix general de** 2×2. Si $A=\begin{pmatrix} a & b \\ c & d\end{pmatrix}=\frac{a+d}{2}I_2+\begin{pmatrix} \frac{a-d}{2} & b \\ c & -\frac{a-d}{2}\end{pmatrix}$, obtenemos fácilmente, a partir de $|A-\lambda I_2|=0$, que los autovalores son

$$\lambda_\pm=\frac{a+d}{2}\pm\sqrt{(\frac{a-d}{2})^2+bc}=\frac{1}{2}\mathrm{Tr}[A]\pm\sqrt{(\frac{\mathrm{Tr}[A]}{2})^2-\mathrm{Det}[A]}$$

donde $\mathrm{Tr}[A]=a+d$ es la traza de A y $\mathrm{Det}[A]=ad-bc$ su determinante. La última expresión puede obtenerse directamente de resolver el sistema $\left\{\begin{array}{rcl}\lambda_++\lambda_- &=& \mathrm{Tr}[A]\\ \lambda_+\lambda_- &=& \mathrm{Det}[A]\end{array}\right.$
Los dos autovalores quedan pues completamente determinados por la traza y el determinante.

Ejemplo 7: **Corrección de primer orden en los autovalores**.
Consideremos una matriz $B=A+\delta A$ de $n\times n$, siendo A diagonalizable y $\delta A=\varepsilon M$ una perturbación (ε es un parámetro suficientemente pequeño y M una matriz de $n\times n$ arbitraria).
Sea $S=(X_1,\ldots,X_n)$ una matriz de autovectores de A, tal que $S^{-1}AS=A'$ con A' diagonal ($A'_{ij}=\lambda_i\delta_{ij}$, con $AX_i=\lambda_i X_i$). Definiendo $\delta A'=S^{-1}(\delta A)S=\varepsilon S^{-1}MS$ (perturbación en la base en que A es diagonal) y notando que $|B-\lambda I|=|S^{-1}BS-\lambda I|$, obtenemos

$$|B-\lambda I|=|D+\delta A'-\lambda I|=\begin{vmatrix}\lambda_1+\delta A'_{11}-\lambda & \delta A'_{12} & \ldots & \delta A'_{1n}\\ \delta A'_{21} & \lambda_2+\delta A'_{22}-\lambda & \ldots & \delta A'_{2n}\\ & & \ldots & \\ \delta A'_{n1} & \delta A'_{n2} & \ldots & \lambda_n+\delta A'_{nn}-\lambda\end{vmatrix}$$

Consideremos primero el caso en que los autovalores λ_i de A son todos distintos. Para $\lambda=\lambda_i+\delta\lambda_i$ con $\delta\lambda_i$ una corrección de orden ε al autovalor λ_i, obtenemos entonces

$$|B-(\lambda_i+\delta\lambda_i)I|=(\delta A'_{ii}-\delta\lambda_i)\prod_{j\neq i}(\lambda_j-\lambda_i)+O(\varepsilon^2)$$

donde el primer término es el de mayor orden ($O(\varepsilon)$), y los restantes de orden $O(\varepsilon^2)$ o mayor. Por lo tanto, la ec. $|B-\lambda I|=0$ conduce a

$$\delta\lambda_i=\delta A'_{ii}+O(\varepsilon^2),\quad \text{con } \delta A'_{ii}=(S^{-1}\delta AS)_{ii}=\varepsilon\sum_{j,k}S^{-1}_{ij}M_{jk}S_{ki}$$

es decir, los $\delta\lambda_i$ son los términos diagonales de δA en la base en que A es diagonal. Como $\sum_j S^{-1}_{ij}S_{ji}=1$, si la columna i de S (el autovector X_i) se multiplica por α, la fila de i de S^{-1} se multiplica por $1/\alpha$, para mantener la igualdad anterior. Por lo tanto, la corrección $\delta A'_{ii}$ es, como debe ser, independiente de la base elegida del espacio propio, es decir de la elección del autovector $X_i\neq 0$ en el espacio propio.

En el caso general, si el espacio propio asociado a un autovalor λ_i tiene dimensión d_i (se dice entonces que tiene degeneración d_i), la corrección a λ_i son los *autovalores* de δA en el espacio propio asociado a λ_i, pues $|A-\lambda I|=|(\delta A')_i-\delta\lambda_i I_{d_i}|\prod_{\lambda_j\neq\lambda_i}(\lambda_j-\lambda_i)^{m_j}+O(\varepsilon^{d_i+1})$, con $\delta A'_i$ la matriz $\delta A'$ restringida al espacio propio asociado a λ_i y m_j la multiplicidad (algebraica) del autovalor λ_j. Se deben pues obtener los autovalores de $\delta A'_i$ (matriz de $d_i\times d_i$). El nivel degenerado λ_i se desdobla normalmente en varios niveles. Se dice entonces que se rompe la degeneración.
Es importante que A sea diagonalizable. De lo contrario, el ej. 2 anterior muestra que en el caso no-diagonalizable, la corrección puede ser por ej. de orden $\sqrt{\varepsilon}$.

10.8 Evaluación de Potencias y Series de Matrices

La diagonalización es muy conveniente para evaluar potencias y series de matrices (de $n\times n$) u operadores. En primer lugar, si

$$A = SA'S^{-1} \qquad (10.3)$$

(A semejante a A') se cumple, para k natural,

$$A^k = SA'^k S^{-1}$$

ya que $A^2 = (SA'S^{-1})(SA'S^{-1}) = SA'^2S^{-1}$ y en general (por inducción) $A^k = AA^{k-1} = SA'S^{-1}SA'^{k-1}S^{-1} = SA'^kS^{-1}$. Análogamente, para funciones definidas por series de potencias $f(u)=\sum_{k=0}^{\infty} a_k u^k$ convergentes $\forall\ u\in\mathbb{C}$,

$$f(A)=\sum_{k=0}^{\infty}a_kA^k=\sum_{k=0}^{\infty}a_kSA'^kS^{-1}=S[\sum_{k=0}^{\infty}a_kA'^k]S^{-1}=Sf(A')S^{-1}$$

Notemos que $f(A)$ está bien definido pues $|(A^k)_{ij}|\leq (mn)^k/n$, donde m el mayor elemento de la matriz ($|A_{ij}|\leq m\ \forall\ i,j$) y n la dimensión. Esto implica que la serie matricial converge absolutamente si la serie converge absolutamente $\forall\ u$ ($|(f(A))_{ij}|\leq f(mn)/n$). En particular,

$$\exp[At]=\sum_{k=0}^{\infty}\frac{(At)^k}{k!}=S\exp[A't]S^{-1}$$

Finalmente, si A es invertible (en cuyo caso A' es también invertible, como el lector debe reconocer inmediatamente) se cumple $A^{-1}=SA'^{-1}S^{-1}$ y en general

$$A^{-k}=SA'^{-k}S^{-1}$$

Si A es *diagonalizable*, $S^{-1}AS=A'$ con A' diagonal. Por lo tanto $A=SA'S^{-1}$. Podemos entonces utilizar las expresiones anteriores con A' diagonal y S una matriz de autovectores, en cuyo caso la evaluación resulta inmediata pues

$$A'=\begin{pmatrix}\lambda_1 & 0 & \ldots & 0\\ 0 & \lambda_2 & \ldots & 0\\ & & \ldots & \\ 0 & 0 & \ldots & \lambda_n\end{pmatrix}\quad\Rightarrow\quad (A')^k=\begin{pmatrix}\lambda_1^k & 0 & \ldots & 0\\ 0 & \lambda_2^k & \ldots & 0\\ & & \ldots & \\ 0 & 0 & \ldots & \lambda_n^k\end{pmatrix}$$

para k natural. Esto implica

$$f(A')=\sum_{k=0}^{\infty}a_k(A')^k=\begin{pmatrix}f(\lambda_1) & 0 & \ldots & 0\\ 0 & f(\lambda_2) & \ldots & 0\\ & & \ldots & \\ 0 & 0 & \ldots & f(\lambda_n)\end{pmatrix}$$

En particular,

$$\exp[A't]=\begin{pmatrix}e^{\lambda_1 t} & 0 & \ldots & 0\\ 0 & e^{\lambda_2 t} & \ldots & 0\\ & & \ldots & \\ 0 & 0 & \ldots & e^{\lambda_n t}\end{pmatrix}$$

Además, si A es invertible, sus autovalores son todos no nulos y es fácil ver que

$$(A')^{-1}=\begin{pmatrix}\lambda_1^{-1} & 0 & \ldots & 0\\ 0 & \lambda_2^{-1} & \ldots & 0\\ & & \ldots & \\ 0 & 0 & \ldots & \lambda_n^{-1}\end{pmatrix}$$

y por lo tanto

$$(A')^{-k}=\begin{pmatrix}\lambda_1^{-k} & 0 & \ldots & 0\\ 0 & \lambda_2^{-k} & \ldots & 0\\ & & \ldots & \\ 0 & 0 & \ldots & \lambda_n^{-k}\end{pmatrix}$$

Por ejemplo, en el caso del ej. 1 anterior se obtiene

$$\exp[At]=\exp[\begin{pmatrix}0&1\\1&0\end{pmatrix}t]=\frac{1}{2}\begin{pmatrix}1&1\\1&-1\end{pmatrix}\begin{pmatrix}e^t&0\\0&e^{-t}\end{pmatrix}\begin{pmatrix}1&1\\1&-1\end{pmatrix}=\begin{pmatrix}\cosh(t)&\sinh(t)\\\sinh(t)&\cosh(t)\end{pmatrix}$$

y en el ej. 3 anterior,

$$A^n=S\begin{pmatrix}1^n&0&0\\0&1^n&0\\0&0&2^n\end{pmatrix}S^{-1}=\begin{pmatrix}1&2^n-1&2^n-1\\0&2^n&2^n-1\\0&0&1\end{pmatrix}$$

$$\exp[At]=S\begin{pmatrix}e^t&0&0\\0&e^t&0\\0&0&e^{2t}\end{pmatrix}S^{-1}=\begin{pmatrix}e^t&e^{2t}-e^t&e^{2t}-e^t\\0&e^{2t}&e^{2t}-e^t\\0&0&e^t\end{pmatrix}$$

Ejemplo: **Sucesión de Fibonacci**. Está definida por la relación recursiva *lineal*

$$a_{n+1}=a_n+a_{n-1},\;\; n\geq 1$$

con $a_0=0$, $a_1=1$.
La expresión explícita de a_n puede obtenerse fácilmente planteando el problema en forma matricial. Resolveremos en realidad el problema para valores iniciales generales a_0, a_1. Tenemos, para $n\geq 1$,

$$\begin{pmatrix}a_{n+1}\\a_n\end{pmatrix}=\begin{pmatrix}1&1\\1&0\end{pmatrix}\begin{pmatrix}a_n\\a_{n-1}\end{pmatrix}$$

Por lo tanto, para $n\geq 1$ y definiendo $A=\begin{pmatrix}1&1\\1&0\end{pmatrix}$,

$$\begin{pmatrix}a_{n+1}\\a_n\end{pmatrix}=A^n\begin{pmatrix}a_1\\a_0\end{pmatrix}$$

La evaluación de A^n puede efectuarse mediante su diagonalización. Los autovalores de A son los números aureos

$$\lambda_\pm=\frac{1\pm\sqrt{5}}{2}$$

que satisfacen $\lambda^2=\lambda+1$, con autovectores $v_\pm\propto\binom{\lambda_\pm}{1}$. Podemos entonces escribir $A=SA'S^{-1}$ con $S=\binom{\lambda_+\,\lambda_-}{1\;\;\,1}$ y $A'=\binom{\lambda_+\;0}{0\;\lambda_-}$. Por lo tanto, $A^n=S(A')^nS^{-1}$ y se obtiene finalmente (se dejan las cuentas para el lector)

$$a_n=[(\lambda_+^n-\lambda_-^n)a_1-(\lambda_+^n\lambda_--\lambda_-^n\lambda_+)a_0]/\sqrt{5}$$

En el caso usual de Fibonacci, $a_0=0$, $a_1=1$ y $a_n=(\lambda_+^n-\lambda_-^n)/\sqrt{5}$. Como $\lambda_+=1.618$, $\lambda_-=-0.618$, el término dominante para n grande es el proporcional a λ_+^n.

Un tratamiento equivalente consiste en expresar el vector inicial $\binom{a_1}{a_0}$ como combinación lineal de los autovectores de A: $A^n\binom{a_1}{a_0}=A^n[c_+\binom{\lambda_+}{1}+c_-\binom{\lambda_-}{1}]=c_+\lambda_+^n\binom{\lambda_+}{1}+c_-\lambda_-^n\binom{\lambda_-}{1}$, de donde $a_n=\lambda_+^nc_++\lambda_-^nc_-$. Como $\binom{c_+}{c_-}=S^{-1}\binom{a_1}{a_0}$, se obtiene $c_+=a_1-\lambda_-a_0$, $c_-=-a_1+\lambda_+a_0$, obteniéndose el resultado anterior.

El mismo método se puede aplicar para toda sucesión definida por una relación recursiva fija *lineal*:

$$a_{n+1}=\alpha_0a_n+\alpha_1a_{n-1}+\ldots+\alpha_ka_{n-k}$$

para $n\geq k$, con $a_0,\ldots,a_k$ dados, que conduce a

$$\begin{pmatrix}a_{n+1}\\a_n\\\ldots\\a_{n-k+1}\end{pmatrix}=\begin{pmatrix}\alpha_0&\alpha_1&\ldots&\alpha_k\\1&0&\ldots&0\\0&1&\ldots&0\\&&\ldots&\\0&\ldots&1&0\end{pmatrix}\begin{pmatrix}a_n\\a_{n-1}\\\ldots\\a_{n-k}\end{pmatrix}=\begin{pmatrix}\alpha_0&\alpha_1&\ldots&\alpha_k\\1&0&\ldots&0\\0&1&\ldots&0\\&&\ldots&\\0&\ldots&1&0\end{pmatrix}^{n-k+1}\begin{pmatrix}a_k\\a_{k-1}\\\ldots\\a_0\end{pmatrix}$$

La sucesión geométrica elemental $a_n=\alpha_0^na_0$ corresponde al caso $k=0$ ($a_{n+1}=\alpha_0a_n$ si $n\geq 0$).

10.9 Desacoplamiento de Sistemas de Ecuaciones Diferenciales Lineales

Como otra aplicación, consideremos por ejemplo el sistema de ecuaciones diferenciales ordinarias lineales de primer orden

$$\frac{dX}{dt} = AX$$

con X de $n \times 1$ y A de $n \times n$, con elementos *constantes* (o sea, independientes del tiempo). Suponiendo A diagonalizable, tenemos $A = SA'S^{-1}$, con A' diagonal y S la matriz de autovectores. Por lo tanto $dX/dt = SA'S^{-1}X$, lo que implica

$$dX'/dt = A'X', \quad X' = S^{-1}X,$$

Como A' es diagonal, el sistema en las variables X' está *desacoplado*, y es de fácil resolución. Tenemos, para las componentes x'_i de X', las ecuaciones desacopladas

$$dx'_i/dt = \lambda_i x'_i, \quad i = 1, \ldots, n$$

donde λ_i son los autovalores de A, cuya solución es $x'_i = c_i e^{\lambda_i t}$. Finalmente, se obtiene

$$X(t) = SX'(t) = \sum_{i=1}^{n} c_i V_i e^{\lambda_i t},$$

donde V_i denota los autovectores de A (las columnas de S). Esto constituye la solución *general* del sistema de primer orden, conteniendo n constantes arbitrarias c_i que pueden determinarse a partir de las condiciones iniciales $x_i(0)$.

El procedimiento usualmente utilizado en Física e Ingeniería para llegar a esta solución es plantear una solución del tipo $X(t) = Ve^{\lambda t}$ con V constante. La ec. $dX/dt = AX$ implica entonces $\lambda V = AV$, por lo que V debe ser autovector de A con autovalor λ. La solución general se obtiene luego como combinación lineal arbitraria de estas soluciones particulares. Este procedimiento es en realidad correcto para encontrar la solución general sólo en el caso de matrices A diagonalizables.

Nótese también que el mismo método puede utilizarse para resolver sistemas análogos de segundo orden

$$\frac{d^2X}{dt^2} = AX$$

Sólo es necesario reemplazar $c_i e^{\lambda_i t}$ por $c_i^+ e^{\sqrt{\lambda_i}t} + c_i^- e^{-\sqrt{\lambda_i}t}$ en la solución general anterior.

Ejemplo 1: Consideremos el sistema de tres ecuaciones diferenciales acopladas de primer orden,

$$dx/dt = x + y + z, \quad dy/dt = 2y + z, \quad dz/dt = z$$

donde x, y, z son funciones de t, el cual puede escribirse en forma matricial como

$$\frac{d}{dt}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

o sea, $dv/dt = Av$, siendo A la matriz del ej. 3 anterior y $v = (x, y, z)^t$. Por lo tanto, utilizando las matrices S y S^{-1} de dicho ejemplo,

$$\frac{d}{dt}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = S\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}S^{-1}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

y mult. a izq. por S^{-1}, se llega a

$$\frac{d}{dt}\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}, \quad \text{con} \quad \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = S^{-1}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x - y - z \\ -z \\ y + z \end{pmatrix}$$

el cual es un sistema de tres ecuaciones dif. lineales *desacopladas*:

$$dx'/dt = x', \quad dy'/dt = y', \quad dz'/dt = 2z'$$

que es equivalente al original. La solución del sistema desacoplado es muy fácil de obtener:

$$x' = c_1 e^t,\ y' = c_2 e^t,\ z' = c_3 e^{2t}$$

Finalmente

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = S \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} x'+z' \\ -y'+z' \\ y' \end{pmatrix} = c_1 e^t \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + c_2 e^t \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} + c_3 e^{2t} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} c_1 e^t + c_3 e^{2t} \\ -c_2 e^t + c_3 e^{2t} \\ c_2 e^t \end{pmatrix}$$

Ejemplo 2: Sistema de dos resortes acoplados (recordar dibujo). Ecuación de movimiento:

$$\frac{d^2}{dt^2} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = -\frac{1}{m} \begin{pmatrix} k_1+k_2 & -k_2 \\ -k_2 & k_1+k_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Resuelto en clase. Detalles a cargo del lector. Sólo recordamos que las frecuencias propias $\omega_i = \sqrt{\lambda_i}$ (con λ_i los autovalores de la matriz $\left(\begin{smallmatrix} k_1+k_2 & -k2 \\ -k2 & k_1+k_2 \end{smallmatrix}\right)/m$ son $\omega_1 = \sqrt{(k_1+2k_2)/m}$, $\omega_2 = \sqrt{k_1/m}$, con $V_1 \propto \left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$, $V_2 \propto \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$.

11. Matrices hermíticas y reales simétricas

Un caso especial muy importante para la física es aquel de matrices de $n \times n$ *hermíticas* (o hermitianas o autoadjuntas), que son las que satisfacen $A^\dagger = A$, donde $A^\dagger \equiv A^{t*}$ denota la matriz traspuesta y conjugada (matriz adjunta):

$$A = A^\dagger \ \Leftrightarrow\ A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{12}^* & a_{22} & \dots & a_{2n} \\ & & \dots & \\ a_{1n}^* & a_{2n}^* & \dots & a_{nn} \end{pmatrix}$$

con a_{ii} real para $i = 1, \dots, n$. Si todos los elementos de A son *reales* $\Rightarrow A^\dagger = A^t$ y la condición de A hermítica equivale a A *simétrica* ($A^t = A$).

11.1) *Teorema:* Si A de $n \times n$ es una matriz hermítica sus autovalores λ_i son todos reales y los autovectores X_i correspondientes a autovalores distintos son ortogonales respecto del producto escalar usual para vectores complejos: Si $AX_i = \lambda_i X_i$, $AX_j = \lambda_j X_j \Rightarrow X_i^\dagger X_j = 0$ si $\lambda_i \neq \lambda_j$, donde

$$X_i^\dagger X_j = (x_{1i}^* \dots x_{ni}^*) \begin{pmatrix} x_{1j} \\ \dots \\ x_{nj} \end{pmatrix} = x_{1i}^* x_{1j} + \dots x_{ni}^* x_{nj}$$

Demostración: Sea X_i de $n \times 1$ autovector de A con autovalor λ_i (por lo tanto $X_i \neq 0$). Multiplicando la igualdad $AX_i = \lambda_i X_i$ a izquierda por $X_i^\dagger = X_i^{t*}$ se obtiene

$$X_i^\dagger A X_i = \lambda_i X_i^\dagger X_i \qquad (11.1)$$

con

$$X_i^\dagger X_i = (x_{1i}^* \dots x_{ni}^*) \begin{pmatrix} x_{1i} \\ \dots \\ x_{ni} \end{pmatrix} = x_{1i}^* x_{1i} + \dots x_{ni}^* x_{ni} = |x_{1i}|^2 + \dots + |x_{ni}|^2 > 0$$

Trasponiendo y conjugando la igualdad (11.1) se obtiene, notando que $(AB)^\dagger = B^\dagger A^\dagger$,

$$X_i^\dagger A^\dagger X_i = \lambda_i^* X_i^\dagger X_i \qquad (11.2)$$

Pero como $A^\dagger = A$, esto implica, comparando con (11.1), que $\lambda_i X_i^\dagger X_i = \lambda_i^* X_i^\dagger X_i$, o sea,

$$0 = (\lambda_i - \lambda_i^*) X_i^\dagger X_i$$

Como $X_i^\dagger X_i > 0 \Rightarrow \lambda_i = \lambda_i^*$, es decir, λ_i real.
Del mismo modo, si $AX_i = \lambda_i X_i$, $AX_j = \lambda_j X_j$, multiplicando a izquierda la primer ecuación por $X_j^\dagger$ y la segunda por $X_i^\dagger$ se obtiene

$$X_j^\dagger A X_i = \lambda_i X_j^\dagger X_i, \quad X_i^\dagger A X_j = \lambda_j X_i^\dagger X_j$$

Trasponiendo y conjugando la pimera de estas ecuaciones se obtiene $X_i^\dagger A^\dagger X_j = \lambda_i^* X_i^\dagger X_j$, es decir, $X_i^\dagger A X_j = \lambda_i X_i^\dagger X_j$ pues $A^\dagger = A$ y $\lambda_i = \lambda_i^*$ (ya demostrado). Por lo tanto, $\lambda_i X_i^\dagger X_j = \lambda_j X_i^\dagger X_j$, o sea,

$$0 = (\lambda_i - \lambda_j) X_i^\dagger X_j$$

por lo que $X_i^\dagger X_j = 0$ si $\lambda_i \neq \lambda_j$.

Puede demostrarse (lo haremos más adelante) que *toda* matriz hermítica A es *diagonalizable*, y que siempre existen n autovectores X_i LI y además ortogonales que satisfacen $AX_i = \lambda_i X_i$, con $X_i^\dagger X_j = 0$ si $i \neq j$ (aún cuando $\lambda_i = \lambda_j$).

En tal caso, eligiendo autovectores normalizados tales que $X_i^\dagger X_i = 1$ para $i = 1, \ldots, n$, se obtiene

$$X_i^\dagger X_j = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Por lo tanto la matriz de autovectores $S = (X_1 \ldots X_n)$ satisface

$$S^\dagger S = I_n$$

pues $(S^\dagger S)_{ij} = X_i^\dagger X_j = \delta_{ij}$. La inversa de S es pues directamente la matriz adjunta: $S^{-1} = S^\dagger$.

En resumen, si $A = A^\dagger$, $\exists$ S tal que $S^{-1} = S^\dagger$ y $S^\dagger A S = A'$, con A' diagonal y real: $A_{ij} = \lambda_i \delta_{ij}$

Matrices antihermíticas: Si $A^\dagger = -A$, se dice que A es antihermítica. En tal caso, $B = -iA$ resulta hermítica (pues $B^\dagger = iA^\dagger = -iA = B$), lo que implica, como $A = iB$, que A será también diagonalizable, con autovectores ortogonales si corresponden a autovalores distintos, pero con autovalores *imaginarios* en lugar de reales: Si $BX_i = \lambda_i X_i \Rightarrow AX_i = (i\lambda_i) X_i$.

Matrices reales simétricas: Para el caso particular de *matrices reales*, los resultados anteriores implican que *los autovalores de matrices reales simétricas ($A^\dagger = A^t = A$) son todos reales.* Los autovectores pueden pués elegirse reales, y por lo tanto, serán ortogonales respecto del producto escalar usual: Si $AX_i = \lambda_i X_i$ y $AX_j = \lambda_j X_j \Rightarrow$

$$X_i^t X_j = x_{1i} x_{1j} + \ldots + x_{ni} x_{nj} = 0 \text{ si } \lambda_i \neq \lambda_j$$

En tal caso, eligiendo autovectores normalizados (tales que $X_i^t X_i = 1$) la inversa de la matriz $S = (X_1, \ldots, X_n)$ será directamente la traspuesta:

$$S^t S = I_n \qquad (A = A^t \text{ real, } \ X_i \text{ real para } i = 1, \ldots, n)$$

En resumen, si $A = A^t$, con A real, $\exists$ S real tal que $S^{-1} = S^t$ y $S^t A S = A'$, con A' diagonal y real: $A'_{ij} = \lambda_i \delta_{ij}$. Más aun, puede elegirse siempre S tal que $\det S = +1$ (Si $S^{-1} = S^t \Rightarrow \det S = \pm 1$) en cuyo caso S corresponde a una rotación, como veremos más adelante

Matrices reales antisimétricas: Si A es real y $A^t = -A$, nuevamente $B = -iA$ resulta hermítica ($B^\dagger = iA^t = -iA = B$) y por lo tanto, $A = iB$ será diagonalizable en el cuerpo de los complejos, con autovalores *imaginarios* y autovectores ortogonales si corresponden a autovalores distintos.

Ejemplo 1: Sea

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

A es una matriz real simétrica si v es real. Tenemos

$$|A - \lambda I_2| = (1 - \lambda)^2 - v^2$$

por lo que los autovalores son $\lambda = 1 \pm v$, reales. Para $\lambda_1 = 1 + v$, puede verse fácilmente que el autovector es de la forma $X_1 = x_1 \binom{1}{1}$, mientras que para $\lambda_2 = 1 - v$, es de la forma $X_2 = x_2 \binom{-1}{1}$. Por lo tanto, podemos elegir $x_1 = x_2 = 1/\sqrt{2}$, para que $X_1^t X_1 = X_2^t X_2 = 1$. Se verifica además $X_1^t X_2 = \frac{1}{2}(1, 1)\binom{-1}{1} = 0$. Por lo tanto

$$S = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad S^{-1} = S^t = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

con

$$S^t AS = \begin{pmatrix} 1+v & 0 \\ 0 & 1-v \end{pmatrix}$$

Ejemplo 2: Sea

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

con v real. A es una matriz hermítica ($A^\dagger = A$). Tenemos

$$|A - \lambda I_2| = (1-\lambda)^2 - (iv)(-iv) = (1-\lambda^2) - v^2$$

por lo que los autovalores son nuevamente $\lambda = 1 \pm v$, reales. Para $\lambda_1 = 1+v$, puede verse fácilmente que el autovector es de la forma $X_1 = x_1\binom{i}{1}$, mientras que para $\lambda_2 = 1-v$, es de la forma $X_2 = x_2\binom{-i}{1}$. Por lo tanto, podemos elegir $x_1 = x_2 = 1/\sqrt{2}$, para que $X_1^\dagger X_1 = X_2^\dagger X_2 = 1$. Se verifica además $X_1^\dagger X_2 = \frac{1}{2}(-i,1)\binom{-i}{1} = (-1+1)/2 = 0$. Por lo tanto

$$S = \frac{1}{\sqrt{2}}\begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix}, \quad S^{-1} = S^\dagger = \frac{1}{\sqrt{2}}\begin{pmatrix} -i & 1 \\ i & 1 \end{pmatrix}$$

con

$$S^\dagger AS = \begin{pmatrix} 1+v & 0 \\ 0 & 1-v \end{pmatrix}$$

Ejemplo 3: Consideremos la ecuación

$$ax^2 + 2bxy + cy^2 = d$$

con coeficientes y variables reales. Podemos escribirla en forma matricial como

$$X^t AX = d, \quad X = \binom{x}{y}, \quad X^t = (x,y), \quad A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

La matriz A es real y simétrica, siendo por lo tanto siempre diagonalizable. Existe entonces una matriz ortogonal de autovectores S ($S^{-1} = S^t$), con $\mathrm{Det}\, S = 1$, tal que $S^{-1}AS = S^t AS = A'$, con A' diagonal: $A' = \left(\begin{smallmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{smallmatrix}\right)$, siendo $\lambda_\pm$ los autovalores de A (las raíces de $(a-\lambda)(b-\lambda) - b^2 = 0$). En tal caso, si $X = SX'$, tenemos

$$X^t AX = X'^t S^t ASX' = X'^t A' X' = \lambda_+ x'^2 + \lambda_- y'^2 = d$$

Si $d > 0$, vemos entonces que la gráfica de la ecuación en las variables x', y' será una elipse si $\lambda_\pm$ son ambos mayores que 0, y una hipérbola si $\lambda_+\lambda_- < 0$, con ejes principales x', y' en ambos casos. Como la transformación corresponde a una rotación (eligiendo el orden de autovectores tal que $\mathrm{Det} S = +1$), la ecuación original corresponderá si $|A| \neq 0$ a una elipse o hipérbola con ejes principales rotados (como consecuencia del término cruzado $2bxy$). El ángulo de inclinación θ entre los ejes x' y x puede obtenerse a partir de la matriz S escribiéndola en la forma $S = [I]_e^{e'} = \left(\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right)$. Para más detalles ver ejemplo resuelto en clase o en práctica.

Ejemplo 4: **Tensor de Inercia**. El tensor de inercia de un cuerpo rígido respecto de un origen O es

$$I_O = \sum_\nu m_\nu \begin{pmatrix} r_\nu^2 - x_\nu^2 & -x_\nu y_\nu & -x_\nu z_\nu \\ -y_\nu x_\nu & r_\nu^2 - y_\nu^2 & -y_\nu z_\nu \\ -z_\nu x_\nu & -z_\nu y_\nu & r_\nu^2 - z_\nu^2 \end{pmatrix} = \sum_\nu m_\nu (X_\nu^t X_\nu I_3 - X_\nu X_\nu^t)$$

donde $X_\nu^t = (x_\nu, y_\nu, z_\nu)$ y $r_\nu^2 = X_\nu^t X_\nu = x_\nu^2 + y_\nu^2 + z_\nu^2$. I_O queda pues representado por una matriz real *simétrica*. Frente a una rotación del sistema de coordenadas, $X_\nu = SX'_\nu$, con $\mathrm{Det} S = 1$, $S^t S = I_3$, se obtiene

$$I_O = \sum_\nu m_\nu (X'^t_\nu S^t S X'_\nu I_3 - SX'_\nu X'^t_\nu S^t) = S[\sum_\nu m_\nu (X'^t_\nu X'_\nu I_3 - X'_\nu X'^t_\nu)]S^t = SI'_O S^t$$

o sea, $I'_O = S^t I_O S$ con I'_O el tensor de inercia en el sistema rotado. Como I_O es real simétrica, existirá una matriz ortogonal de rotación S (matriz de autovectores normalizados y ordenados tal que $S^t S = I_3$ y $|S| = 1$) tal que I'_O sea *diagonal*. Esta matriz determinará los 3 *ejes principales de inercia*, y los autovalores de I_O serán los *momentos principales de inercia*. Si el vector velocidad angular Ω coincide con alguna de estas direcciones, el vector momento angular (dado en general por $L_O = I_O \Omega$) será proporcional a Ω.

Ejemplo 5: **Sistema general de n resortes acoplados**. El movimiento de tal conjunto esta descripto por un sistema de ecuaciones de segundo orden del tipo (recordar discusión de clase)

$$\frac{m_i d^2 x_i}{dt^2} = -\sum_{j=1}^{n} k_{ij} x_j, \quad i = 1, \ldots, n$$

donde x_i es la posición de la partícula i (medida desde la posición de equilibrio), $m_i > 0$ su masa y $k_{ij} = k_{ji}$. Podemos reescribir tal sistema en forma matricial como

$$M\frac{d^2 X}{dt^2} = -KX$$

donde M es una matriz diagonal de elementos m_i ($M_{ij} = m_i \delta_{ij}$), $X = (x_1, \ldots, x_n)^t$ y K la matriz de elementos k_{ij}. Definiendo la matriz diagonal $M^{1/2}$ de elementos $\sqrt{m_i}$ ($(M^{1/2})_{ij} = \sqrt{m_i}\delta_{ij}$) podemos reescribir tal sistema como $M^{1/2}M^{1/2}\frac{d^2X}{dt^2} = -KX$, y por lo tanto, multiplicando a izquierda por $M^{-1/2} = (M^{1/2})^{-1}$ (matriz diagonal de elementos $(M^{-1/2})_{ij} = \frac{1}{\sqrt{m_i}}\delta_{ij}$), como

$$\frac{d^2 Y}{dt^2} = -\tilde{K}Y, \text{ donde } \tilde{K} = M^{-1/2}KM^{-1/2}, \;\; Y = M^{1/2}X$$

(de forma que $y_i = \sqrt{m_i}x_i$). La ventaja esta forma matricial es que la matriz $\tilde{K}$ es real **simétrica** ($\tilde{K}^t = \tilde{K}$) y por lo tanto siempre **diagonalizable**. Existe entonces una matriz ortogonal S tal que $S^t\tilde{K}S = \tilde{K}'$, con K' **diagonal**, de elementos $\tilde{K}'_{ij} = \lambda_i \delta_{ij}$, siendo λ_i los autovalores de $\tilde{K}$. Por lo tanto, escribiendo $\tilde{K} = S\tilde{K}'S^t$, el sistema original resulta equivalente a

$$\frac{d^2 Y'}{dt^2} = -\tilde{K}'Y', \quad \text{donde } Y' = S^tY$$

Esto representa, dado que $\tilde{K}'$ es diagonal, un sistema de n resortes **desacoplados**:

$$\frac{d^2 y'_i}{dt^2} = -\lambda_i y'_i, \;\; i = 1, \ldots, n$$

La solución general de c/u de estas ecuaciones es, para $\lambda_i \neq 0$, $y'_i(t) = Ae^{i\omega_i t} + Be^{i\omega_i t} = C\cos(\omega_i t + \phi)$, donde $\omega_i = \sqrt{\lambda_i}$ **son las frecuencias propias de vibración del sistema**. Las variables $y'_i = \sum_{j=1}^{n} S_{ji}\sqrt{m_j}x_j$ (o sea, $y'_i = (S^tY)_i$) se denominan **modos normales de vibración**. Notemos que las frecuencias propias son las raíces de los autovalores de la matriz $M^{-1/2}KM^{-1/2}$, los cuales, en virtud de la propiedad 12.1 siguiente, coinciden con los de la matriz $M^{-1}K$. Por lo tanto, la conocida fórmula $\omega = \sqrt{k/m}$ para la frecuencia angular de un oscilador armónico se generaliza a $\omega_i = \sqrt{(M^{-1}K)_i}$, donde $(M^{-1}K)_i$ denota aquí el iésimo autovalor de la matriz $M^{-1}K$. Puede demostrarse que si la matriz K es definida positiva (definición que veremos luego y que corresponde a un sistema estable) entonces todos los autovalores de $\tilde{K}$ son positivos.

Ejemplo 6: **Problema generalizado de autovalores:** La ecuación

$$AX = \lambda BX$$

donde A y B son matrices de $n \times n$, B es no singular y $X \neq 0$ es un vector columna, define un problema generalizado de autovalores. Es obviamente equivalente al problema estándar $B^{-1}AX = \lambda X$, es decir, a la obtención de los autovalores y autovectores de la matriz $B^{-1}A$. Los autovalores pueden pues obtenerse como $|B^{-1}A - \lambda I| = 0$, equivalente a

$$|A - \lambda B| = 0$$

y los espacios propios pueden obtenerse como $\text{Nu}(B^{-1}A - \lambda I)$, equivalente a $\text{Nu}(A - \lambda B)$.

Si $A^\dagger = A$ y $B^\dagger = B$, $B^{-1}A$ no es general hermítica ($(B^{-1}A)^\dagger = AB^{-1}$). Si B es definida positiva (es decir, todos sus autovalores λ_i^B son positivos), es posible reescribir el problema generalizado como un problema de autovalores *hermítico y estándar*, mediante el mismo método utilizado en el ej. 5:
Si $AX = \lambda BX \Rightarrow B^{-1/2}AB^{-1/2}B^{1/2}X = \lambda B^{1/2}X$, por lo que el problema se reduce al problema estándar

$$\tilde{A}\tilde{X} = \lambda\tilde{X}$$

con $\tilde{A} = B^{-1/2}AB^{-1/2}$ una matriz *hermítica* ($\tilde{A}^\dagger = \tilde{A}$) y $\tilde{X} = B^{1/2}\tilde{X}$. Aqui, si $B = S_B B' S_B^\dagger$, con B' diagonal ($B'_{ij} = \lambda_i^B \delta_{ij}$, $\lambda_i^B > 0$), $B^{1/2} = S_B B'^{1/2} S_B^\dagger$, con $(B'^{1/2})_{ij} = \sqrt{\lambda_i^B}\delta_{ij}$, siendo $B^{-1/2}$ su inversa.

Por lo tanto, los autovalores λ serán *reales*, y pueden obtenerse de $|\tilde{A}-\lambda I|=0$, equivalente a $|A-\lambda B|=0$, mientras que los autovectores $\tilde{X}$ correspondientes a autovalores distintos serán ortogonales: $\tilde{X}_i^\dagger \tilde{X}_j=0$ si $\lambda_i\neq j\lambda_j$. Esto implica que los autovectores X del problema original serán ortogonales para *un producto escalar modificado dependiente de B:* $\tilde{X}_i^\dagger\tilde{X}_j=X_i^\dagger BX_j=0$ si $\lambda_i\neq\lambda_j$ y $\tilde{X}_i=B^{1/2}X_i$. Veremos en las próximas secciones la definición de producto escalar con más detalle. Al ser $\tilde{A}$ diagonalizable y $B^{1/2}$ no singular, tanto el conjunto $\{\tilde{X}_1,\ldots,\tilde{X}_n\}$ como $\{X_1,\ldots,X_n\}$ formarán una base de $\mathbb{C}^n$.

Más aun, dado que $\tilde{A}$ es hermítica, existirá una matriz no singular de autovectores normalizados y ortogonales $\tilde{S}=(\tilde{X}_1,\ldots,\tilde{X}_n)$ tal que $\tilde{S}^\dagger\tilde{S}=I$ y $\tilde{S}^\dagger\tilde{A}S=\tilde{A}'$ con $\tilde{A}'$ diagonal: $\tilde{A}'_{ij}=\lambda_i\delta_{ij}$. Esto implica que $S=B^{-1/2}\tilde{S}=(X_1,\ldots,X_n)$ satisface simultáneamente

$$S^\dagger AS=A',\quad S^\dagger BS=I$$

con $A'=\tilde{A}'$ diagonal: $A'_{ij}=\lambda_i\delta_{ij}$. Los autovalores generalizados λ_i pueden obtenerse directamente como las raíces de $|A-\lambda B|=0$, mientras que los correspondientes autovectores X_i (columnas de S) de la ecuación $(A-\lambda_i B)X_i=0$. La existencia de tal S queda pues garantizada en el caso $A^\dagger=A$ y $B^\dagger=B$, con B definida positiva ($\lambda_i^B>0\ \forall\ i$).

12. Otras propiedades importantes

12.1) Sean $F:V\to V$, $G:V\to V$ dos operadores sobre el mismo espacio V de dimensión finita. Entonces los autovalores de FG *son los mismos que los de* GF, aún cuando $FG\neq GF$
Demostración: En efecto, si $FG(v)=\lambda v$ con $v\neq 0\Rightarrow GF(G(v))=G(FG(v))=G(\lambda v)=\lambda G(v)$. Si $G(v)\neq 0\Rightarrow\lambda$ es también autovalor de GF con autovector $G(v)$. Si $G(v)=0\Rightarrow$ necesariamente $\lambda=0$ (pues $FG(v)=F(G(v))=F(0)=0$) y por lo tanto $0=\mathrm{Det}[FG]=\mathrm{Det}[GF]$. Esto implica que 0 es también autovalor de GF. Análogamente se muestra que todo autovalor de GF es autovalor de FG.
Esta propiedad es entonces también válida para matrices. Si A,B son dos matrices de $n\times n\Rightarrow$ los autovalores de AB son los mismos que los de BA, aún si $AB\neq BA$.

12.2) Si $[F,G]=FG-GF=0$ y $v\neq 0$ es autovector de F con autovalor $\lambda\Rightarrow G(v)\in V_F(\lambda)$, es decir, $G(v)$ es también autovector de F con el mismo autovalor si $G(v)\neq 0$.
Demostración: $F(G(v))=FG(v)=GF(v)=G(\lambda v)=\lambda G(v)$, por lo que $G(v)\in V_F(\lambda)$.
Además, si V es de dimensión finita n y los autovalores de F *son todos distintos* $\Rightarrow$ todo autovector de F es también autovector de G, pues en tal caso $V_F(\lambda)$ es de dimensión 1 para todo autovalor λ y por lo tanto, necesariamente $G(v)=\alpha v$.
Todos los operadores que conmutan con F quedan en este caso directamente representados por matrices *diagonales* en la base B' en que $[F]_{B'}$ es diagonal, y son por lo tanto diagonalizables.

Si F y G son ambos diagonalizables y $[F,G]=0\Rightarrow$ **existe una base común donde F y G son simultáneamente diagonales:** Por la propiedad anterior, sólo es necesario diagonalizar F, y luego diagonalizar G dentro de cada espacio propio de F.

13. Subespacios Invariantes

Sea $F : V \to V$ un operador lineal y sea $W \subset V$ un subespacio de V. W es un subespacio invariante bajo la acción de F (se dice también invariante bajo F o por F) si $F(W) \subset W$, es decir, si $\forall\, v \in W$, $F(v) \in W$.

Ejemplos triviales son $W = V$ y $W = \{0\}$, que son subespacios invariantes para todo operador lineal $F : V \to V$ (pues $F(V) \subset V$ y $F(0) = 0$).
También el núcleo $N(F)$ y la imagen $I(F)$ son siempre invariantes por F ($F(N(F)) = \{0\} \subset N(F)$, y si $v \in I(F)$, $F(v) \in I(F)$).
Resulta asimismo trivial reconocer que si $F = \alpha I$, con I el operador identidad $\Rightarrow$ cualquier subespacio $W \subset V$ es invariante por F, ya que si $v \in W$, $F(v) = \alpha v \in W$.
Como otro ejemplo común consideremos el proyector P sobre S_1 en la dirección de S_2, con $S_1 \oplus S_2 = V$.
Es obvio que S_1 es invariante por P pues si $v_1 \in S_1 \Rightarrow P(v_1) = v_1 \in S_1$ ($P(S_1) = S_1 = I(P)$). Más aún, cualquier subespacio de S_1 es también invariante por P.

13.1) Si λ es autovalor de $F \Rightarrow$ el espacio propio $V(\lambda)$ es un subespacio invariante por F: Si $v \in V(\lambda)$ $\Rightarrow F(v) = \lambda v \in V(\lambda)$.
13.2) La suma de espacios propios $V(\lambda_1) \oplus V(\lambda_2)$ de un mismo operador F es también invariante por F.
Si $v \in V(\lambda_1) \oplus V(\lambda_2)$, $v = v_1 + v_2$, con $F(v_i) = \lambda_i v_i$, $i = 1, 2$ y por lo tanto $F(v) = \lambda_1 v_1 + \lambda_2 v_2 \in V(\lambda_1) \oplus V(\lambda_2)$.
Análogamente, la suma de espacios propios $V(\lambda_1) \oplus \ldots \oplus V(\lambda_k)$ es invariante por F.
13.3) Si $[F, G] = 0$ y W es un subespacio invariante por F, $G(W)$ es también invariante por F.
En efecto, si $v \in W$, $w = F(v) \in W$ y por lo tanto, $F(G(v)) = FG(v) = GF(v) = G(F(v)) = G(w) \in G(W)$.
13.4) Si se escribe a V como suma directa de subespacios invariantes,
$V = S_1 \oplus S_2 \oplus \ldots \oplus S_k$, existe una base (aquella formada por la unión de las bases de cada subespacio invariante) en la que la matriz $[F]_B$ que representa a F en esta base estará definida por bloques de la forma

$$[F]_B = \begin{pmatrix} A_1 & 0 & \ldots & 0 \\ 0 & A_2 & \ldots & 0 \\ & & \ldots & \\ 0 & 0 & \ldots & A_k \end{pmatrix}$$

donde A_i es una matriz de dimensión $d_i \times d_i$, con $d_i = \text{dim} S_i$, $i = 1, \ldots, k$ y $\sum_{i=1}^{k} d_i = n$.
En efecto, si $(b_{i1}, \ldots, b_{id_i})$ es una base de S_i, $F(b_{ij}) \in S_i$, por lo que $F(b_{ij}) = \sum_{l=1}^{d_i} (A_i)_{lj} b_{il}$ para $j = 1, \ldots, d_i$.
Análogamente, si existe una base en la que $[F]_B$ tiene la forma de bloques anterior $\Rightarrow V = S_1 \oplus S_2 \oplus \ldots \oplus S_k$, con S_i invariante por F para $i = 1, \ldots, k$. Basta con considerar S_i como el subespacio generado por los vectores de la base correspondientes a cada bloque.

Los autovalores de F pueden pues obtenerse directamente diagonalizando cada bloque A_i:
Como $\text{Det}\, F = \prod_i \text{Det} A_i$ (demostración dada en clase), el polinomio característico resulta

$$\text{Det}(F - \lambda I) = \prod_{i=1}^{k} \text{Det}(A_i - \lambda I_{d_i})$$

por lo que sus racies serán las raíces de cada término, es decir, los autovalores de cada bloque A_i. Y los autovectores correspondientes pertenecerán al subespacio invariante asociado a A_i (detalles dados en clase). El conocimiento de subespacios invariantes posibilita pues grandes simplificaciones cuando se tiene que diagonalizar matrices de grandes dimensiones.