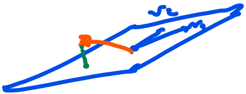
Análisis Matemático para Inteligencia Artificial

Martín Errázquin (merrazquin@fi.uba.ar)

Especialización en Inteligencia Artificial

Clase 2



Espacios con Producto Interno: Definición

$+ : \mathbb{V} \times \mathbb{V} \to \mathbb{V}$

$\cdot : \mathbb{K} \times \mathbb{V} \to \mathbb{V}$

$\Phi : \mathbb{V} \times \mathbb{V} \to \mathbb{K}$

$\Phi(\alpha v + \beta w, u) = \alpha \cdot \Phi(v, u) + \beta \cdot \Phi(w, u)$

Sea $\mathbb{V} - \mathbb{K}$ e.v., donde $\mathbb{K} = \mathbb{R}$ *o* $\mathbb{C}$, un producto interno sobre $\mathbb{V}$ es una función $\Phi : \mathbb{V} \times \mathbb{V} \to \mathbb{R}$ (*o* $\mathbb{C}$) que satisface:

$z = a + bi \in \mathbb{C}$

$\bar{z} = a + i(-b)$

1. Para cada $\alpha \in \mathbb{R}$ (*o* $\mathbb{C}$), $u, v, w \in \mathbb{V}$.
   - $\Phi(u + v, w) = \Phi(u, w) + \Phi(v, w)$
   - $\Phi(\alpha \bullet u, v) = \alpha \bullet \Phi(u, v)$

   linealidad
2. $\Phi(u, v) = \overline{\Phi(v, u)}$ $\to \in \mathbb{R}$ $\Phi(u, v) = \Phi(v, u)$ simetría
3. $\Phi(v, v) \geq 0$, y $\Phi(v, v) = 0$ sii $v = 0$

si $v = 0_V \Rightarrow \Phi(v, v) = 0_{\mathbb{K}}$

$v \neq 0_V \Rightarrow \Phi(v, v) > 0_{\mathbb{K}}$

Notación: $\Phi(u, v) = \langle u, v \rangle$

Definición: A un espacio vectorial real (complejo) provisto de un producto interno se lo llama espacio euclídeo (espacio unitario).

Obs: El p.i. es una generalización del producto escalar en $\mathbb{R}^n$ (o $\mathbb{C}^n$).

$v, w \in \mathbb{R}^n \Rightarrow v \cdot w = \sum_{i=1}^{n} v_i \cdot w_i$

Sea $\mathcal{V}$ el espacio de las funciones continuas de valor real (o complejo) en el intervalo $-1 \le x \le 1$ (se nota $\mathcal{C}([-1,1])$) con p.i.

$z \in \mathbb{C}$

$z = \overline{\overline{z}}$

$$\langle f, g\rangle = \int_{-1}^{1} f(x)\overline{g(x)}dx$$

1) a) $<f+g,h> = \int_{-1}^{1} [f(x)+g(x)]\cdot\overline{h(x)}\,dx = \int_{-1}^{1} f(x)\overline{h(x)} + g(x)\overline{h(x)}\,dx =$

$= \int_{-1}^{1} f(x)\overline{h(x)}dx + \int_{-1}^{1} g(x)\overline{h(x)}\,dx = <f,h> + <g,h>$ ✓

$z\cdot\overline{z} = \|z\|^2 = a^2+b^2 \in \mathbb{R}_0^+$

b) $<k\cdot f, g> = \int_{-1}^{1} (k\cdot f(x))\overline{g(x)}\,dx = \int_{-1}^{1} k\cdot f(x)\cdot\overline{g(x)}\,dx = k\cdot\int_{-1}^{1} f(x)\overline{g(x)}\,dx =$

$= k\cdot <f,g>$ ✓

$(a+bi)(a-bi) = a^2 \cancel{-abi+abi} - b^2i^2 = \underbrace{a^2}_{\ge 0} + \underbrace{b^2}_{\ge 0} \ge 0$

2) $<f,g> = \int_{-1}^{1} f(x)\overline{g(x)}\,dx = \overline{\overline{\int_{-1}^{1} f(x)\overline{g(x)}\,dx}} = \overline{\int_{-1}^{1} \overline{f(x)\cdot\overline{g(x)}}\,dx} =$

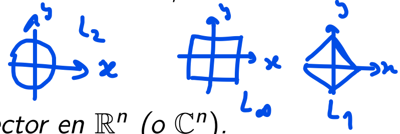
$= \overline{\int_{-1}^{1} \overline{f(x)}\cdot g(x)\,dx} = \overline{\int_{-1}^{1} g(x)\overline{f(x)}\,dx} = \overline{<g,f>}$ ✓

3) $<f,f> = \int_{-1}^{1} f(x)\cdot\overline{f(x)}\,dx = \int_{-1}^{1} \underbrace{\|f(x)\|^2}_{\ge 0\ \forall x}\,dx \ge 0$ y además $= 0$ si y sólo si $f(x) = 0\ \forall x \iff f(x) = 0$

$\Phi \Rightarrow \Phi(v,v) \geq 0 \Rightarrow \exists \sqrt{\Phi(v,v)} = \|v\|$

Sea $(\mathbb{V}, \langle ., . \rangle)$ un e.v. real (complejo) con p.i.. Sea $v \in \mathbb{V}$, se define la norma de v asociada a $\langle ., . \rangle$.

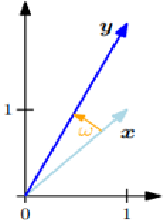
Notación: $||v|| = \langle v, v \rangle^{1/2}$



Es la generalización de la longitud de un vector en $\mathbb{R}^n$ (o $\mathbb{C}^n$).

Def: A partir de un p.i. se puede se puede definir el ángulo w entre dos vectores x, y

$$\cos(\omega) = \frac{\langle x, y \rangle}{\underbrace{\|x\|}_{\sqrt{\langle x,x \rangle}} \; \underbrace{\|y\|}_{\sqrt{\langle y,y \rangle}}} \in [-1; 1]$$



$\langle x, y \rangle = \|x\| \, \|y\| \cdot \cos(\omega)$

Propiedades de la Norma

1. $\forall v \in \mathbb{V},\ ||v|| \geq 0$, y $||v|| = 0$ sii $v = 0$. $\leftarrow \|v\|^2 = \langle v, v \rangle = 0 \iff v = 0_V$
2. Sean $\alpha \in \mathbb{R}(o\ \mathbb{C})$, $v \in \mathbb{V}$, $||\alpha \bullet v|| = |\alpha|\ ||v||$.
3. Desigualdad de Cauchy Schwartz: si $u, v \in \mathbb{V}$ entonces

$$|\langle u, v \rangle| \leq ||u||\ ||v||$$

4. Desigualdad Triangular: si $u, v \in \mathbb{V}$ entonces

$$||u + v|| \leq ||u|| + ||v||$$

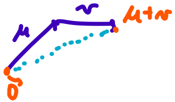
$$\|\alpha \cdot v\|^2 = \langle \alpha \cdot v, \alpha \cdot v \rangle = \alpha \langle v, \alpha \cdot v \rangle = \alpha \cdot \overline{\langle \alpha v, v \rangle} = \alpha \cdot \bar{\alpha}\, \overline{\langle v, v \rangle} = |\alpha|^2 \langle v, v \rangle$$

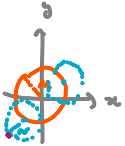
$$|\alpha|^2 \cdot \|v\|^2$$

3 $$|\langle u, v \rangle| = |\, \|u\| \cdot \|v\| \cdot \cos(\omega) | = \|u\| \cdot \|v\| \cdot \underbrace{|\cos(\omega)|}_{\in [0;1]}$$

$$\leq \|u\| \cdot \|v\| \cdot 1 = \|u\| \cdot \|v\| \ \checkmark$$

4





Def: $(\mathbb{V}, \langle.,.\rangle)$ un $\mathbb{K}$-EV (con $\mathbb{K} = \mathbb{R}$ *o* $\mathbb{C}$) con p.i. dos vectores $u, v \in \mathbb{V}$ se dicen ortogonales si $\langle u, v\rangle = 0$.

Teorema de Pitágoras: Si $u, v \in \mathbb{V}$ son ortogonales entonces $||u+v||^2 = ||u||^2 + ||v||^2$.

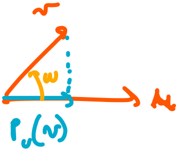
$M_{i,j} = \langle v_i, v_j \rangle \in \mathbb{K}^{r\times r}$

si $\perp$ ortogon.
$M = \begin{pmatrix} ||v_1||^2 & & 0 \\ & ||v_2||^2 & \\ 0 & & ||v_r||^2 \end{pmatrix}$

Def: $(\mathbb{V}, \langle.,.\rangle)$ un $\mathbb{K}$-EV (con $\mathbb{K} = \mathbb{R}$ *o* $\mathbb{C}$) con p.i.. Se dice que $\{v_1, ..., v_r\} \subset \mathbb{V}$ es un conjunto ortogonal si $\langle v_i, v_j\rangle = 0, \ \forall i \neq j$. Si $||v_i|| = 1, \ \forall i$ se dice que es un conjunto ortonormal.

si $\perp$ ortonorm.
$M = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} = I_{r\times r}$

La proyección ortogonal del vector v sobre el vector u es otro vector que notamos como $P_u(v)$, y se define:



$$P_u(v) = \frac{\langle u, v\rangle}{||u||}\underbrace{\frac{u}{||u||}}_{\check{u}} = \frac{\langle u, v\rangle}{\langle u, u\rangle}u$$

$\dfrac{||u|| \cdot ||v|| \cdot \cos(w)}{||u||}$

Def: Una base ortonormal (BON) de un E.V. es una base $B = \{v_1, ..., v_n\}$ que satisface:

$$\langle v_i, v_j \rangle = 0, \ \forall i \neq j$$
$$\langle v_i, v_i \rangle = 1, \ \forall i$$

Obs: Si sólo se cumple que $\langle v_i, v_j \rangle = 0, \ \forall i \neq j$ se dice que es una base ortogonal.

Para transformar una base en una base ortonormal usamos el proceso de Gram-Schmidt:

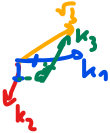
$B = \{v_1, v_2, v_3, \dots, v_N\}$

$B' = \{k_1, k_2, k_3, \dots, k_n\}$ BOG

$\tilde{B} = \left\{\frac{k_1}{\|k_1\|}, \frac{k_2}{\|k_2\|}, \dots, \frac{k_n}{\|k_n\|}\right\}$ BON

$$k_1 = v_1$$
$$k_2 = v_2 - Proy_{k_1}(v_2)$$
$$\vdots$$
$$k_n = v_n - \sum_{i=1}^{n-1} Proy_{k_i}(v_n)$$

$k_3 = v_3 - Proy_{k_1}(v_3) - Proy_{k_2}(v_3)$

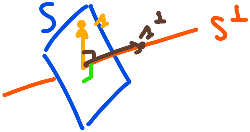


Y así, $\tilde{B} = \{k_1, ..., k_n\}$ pidiendo que $||k_i|| = 1$ resulta una BON.

Complemento Ortogonal

Sea $\mathbb{V}$ un EV de dimensión $n < \infty$ y $S \subset \mathbb{V}$ un SEV de dimensión $m \leq n$. El complemento ortogonal ($S^{\perp}$) es un SEV de de dimensión $n - m$ que satisface:

$$S \cup S^{\perp} = \mathbb{V} \quad \wedge \quad S \cap S^{\perp} = \emptyset$$



Ejemplo:

$$S = \left\langle \underbrace{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}}_{u}, \underbrace{\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}}_{v} \right\rangle$$

$\dim(S) = 2$

$\dim(\mathbb{R}^3) = 3 \Rightarrow \dim(S^{\perp}) = 1$

$$u \times v = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ 1 & 1 & 1 \\ 1 & 2 & 0 \end{vmatrix} = (-2, 1, 1) = w$$

$$\Rightarrow S^{\perp} = \left\langle \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \right\rangle$$

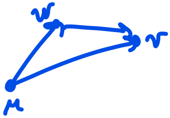
EV + PI → norma → distancia $\qquad \mathbb{R}_0^+$

Sea $\mathbb{V}$-$\mathbb{K}$, ($\mathbb{R}$ *o* $\mathbb{C}$) EV con p.i. $\langle .,. \rangle$ se define la distancia $d : \mathbb{V} \times \mathbb{V} \to \mathbb{K}$ como $d(u,v) = ||u - v||$. $= \sqrt{\langle u-v, u-v \rangle}$

0 si $u=v$

Propiedades:

1. $d(u,v) \geq 0, \ \forall u,v \in \mathbb{V}$
2. $d(u,v) = 0 \Leftrightarrow u = v \Leftrightarrow u - v = 0_V$
3. $d(u,v) = d(v,u), \ \forall u,v \in \mathbb{V}$
4. $d(u,v) \leq d(u,w) + d(w,v), \ \forall u,v,w \in \mathbb{V}$



Observación: Existen distancias que no están asociadas a ninguna norma.

Matrices definidas positivas

$A^T = A$

Una matriz $A \in \mathbb{R}^{n\times n}$ se dice definida positiva si es simétrica y vale que:

forma cuadrática $\longrightarrow$ $$\underbrace{x^T A x}_{\mathbb{R}} > 0, \ \forall x \in \mathbb{R}^n - \{0\}$$

Si vale que $x^T A x \geq 0$ se la llama semi definida positiva.

Ejemplo: $A = \begin{pmatrix} 2 & 3 \\ 3 & 9 \end{pmatrix} \in \mathbb{R}^2$

$x = \begin{pmatrix} a \\ b \end{pmatrix}$ $\quad x^T A x = (a \ \ b)\begin{pmatrix} 2 & 3 \\ 3 & 9 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = (a \ \ b)\begin{pmatrix} 2a+3b \\ 3a+9b \end{pmatrix} = 2a^2+3ab+3ab+9b^2 =$

$= 2a^2+6ab+9b^2 = \underbrace{a^2}_{\geq 0} + \underbrace{(a+3b)^2}_{\geq 0} \geq 0$ y > 0 $\forall x \neq (0,0)$

!

si $a \neq 0$ $\Rightarrow > 0$

si $a = 0, b \neq 0$ $\Rightarrow > 0$

Teorema: Sea $\mathbb{V}$ un EV de dimensión finita, y B una base de $\mathbb{V}$, vale que es un p.i. sii existe una matriz definida positiva tal que:

$$\langle x, y\rangle = \tilde{x}^T A \tilde{y}$$

$\tilde{x} = \begin{bmatrix} a \\ b \\ \vdots \\ n \end{bmatrix} \Leftrightarrow x = a\cdot v_1 + \dots + n\cdot v_n$

donde $\tilde{x}, \ \tilde{y}$, son las representaciones de $x, \ y$ en la base B. $B = \{v_1, \dots, v_n\}$

Sea $T : \mathbb{V} \to \mathbb{W}$ una transformación, donde $\mathbb{V}, \ \mathbb{W}$ son dos conjuntos arbitrarios. Se dice que T es:

- Inyectiva: si $\forall x, \ y \in \mathbb{V} : T(x) = T(y) \to x = y$ $\quad \forall x \neq y \Rightarrow T(x) \neq T(y)$
- Suryectiva: si $T(\mathbb{V}) = \mathbb{W}$ $\quad T(\mathbb{V}) = \mathrm{Im}(T) = \{T(v) \ \forall v \in \mathbb{V}\} \subseteq \mathbb{W}$
- Biyectiva: si es inyectiva y suryectiva.

Transformaciones Lineales

Sean $\mathbb{V}, \ \mathbb{W}$ dos EV, $L : \mathbb{V} \to \mathbb{W}$ es una transformación lineal si:

$$L(\alpha x + \beta y) = \alpha L(x) + \beta L(y) \ \ \forall \alpha, \beta \in \mathbb{K}, \ \forall x, y \in \mathbb{V}$$

- Isomorfismo: $L : \mathbb{V} \to \mathbb{W}$ es lineal y biyectiva.
- Endomorfismo: si $L : \mathbb{V} \to \mathbb{V}$ es lineal.
- Automorfismo: $L : \mathbb{V} \to \mathbb{V}$ es lineal y biyectiva.

INY SURY

BIYEC. LINEAL V→V

ISOM. ENDOM.

AUTOM.

Teorema: Sea $\mathbb{V}$ y $\mathbb{W}$, dos espacios vectoriales de dimensión finita son un isomorfismo sii $dim(\mathbb{V}) = dim(\mathbb{W})$.

$n < \infty$

Teorema: Sea $\mathbb{V}$ un EV, $dim(\mathbb{V}) = n < \infty$ tiene un isomorfismo con $\mathbb{R}^n$. Si consideramos la base $B = \{v_1, ..., v_n\}$, todo $v \in \mathbb{V}$ puede escribirse como $v = \alpha_1 v_1 + ... + \alpha_n v_n$. Luego las coordenadas de v en la base B resulta:

$\mathcal{P}_2[x]$ es isom. de $\mathbb{R}^3$

$$\alpha = (\alpha_1, ..., \alpha_n)^T \in \mathbb{R}^n$$

$$v = 3 + 4x + 5x^2 \in \mathcal{P}_2[x]$$

$$v = 3 \cdot 1 + 4 \cdot x + 5 \cdot x^2 = \begin{pmatrix} | & | & | \\ 1 & x & x^2 \\ | & | & | \end{pmatrix} \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix}$$

$B = \{1, x, x^2\}$

$$[v]_B = (3, 4, 5) \in \mathbb{R}^3$$

Núcleo e Imagen de una transformación

Sea $L : \mathbb{V} \to \mathbb{W}$, se define:

- Núcleo (o Kernel) $Nu(L) = \{v \in \mathbb{V} : L(v) = 0_W\}$,
- Imagen $Im(L) = L(\mathbb{V}) = \{w \in \mathbb{W} : \exists v \in \mathbb{V},\ w = L(v)\}$

Teorema: **Toda transformación lineal se puede representar de forma matricial.** $L : \mathbb{R}^n \to \mathbb{R}^m \,/\, L(v) = A \cdot v$ con $A \in \mathbb{R}^{m \times n}$

- Espacio Nulo de A: es un subespacio de $\mathbb{R}^n$ formado por todas las soluciones del sistema lineal homogéneo $Av = \underbrace{0}_{\in \mathbb{R}^m}$, $A \in \mathbb{R}^{m \times n}$.

$$N(A) = \{v \in \mathbb{R}^{n \times 1} : Av = 0\}$$

$A = \begin{pmatrix} | & & | \\ A_1 & \cdots & A_n \\ | & & | \end{pmatrix}$

- Espacio columna de A: es el subespacio de $\mathbb{R}^m$ generado por los n vectores columna de A: $= \langle \{A_1, \ldots, A_n\} \rangle$

$$EC(A) = \{\alpha_1 (a_{11}, ..., a_{m1})^T + ... + \alpha_m (a_{1n}, ..., a_{mn})^T, \alpha_i \in \mathbb{R}\}$$

- Espacio fila de A: e el subespacio de R^n generado por los m vectores fila de A: $= \langle \{A^{(1)}, \ldots, A^{(m)}\} \rangle$

$A = \begin{pmatrix} - A^{(1)} - \\ \vdots \\ - A^{(m)} - \end{pmatrix}$

$$EF(A) = \{\alpha_1 (a_{11}, ..., a_{1n}) + ... + \alpha_m (a_{m1}, ..., a_{mn}), \alpha_i \in \mathbb{R}\}$$

Veamos un ejemplo...

$L:\mathbb{R}^2 \to \mathbb{R}^2 / L\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}x+y\\x-y\end{pmatrix}$ ¿es TL? $L\left(\alpha\begin{pmatrix}x_1\\y_1\end{pmatrix}+\beta\begin{pmatrix}x_2\\y_2\end{pmatrix}\right) = L\begin{pmatrix}\alpha_1 x_1+\beta x_2\\ \alpha_1 y_1+\beta y_2\end{pmatrix} =$ lineal ✓

$= \begin{pmatrix}\alpha_1 x_1+\alpha_1 y_1+\beta x_2+\beta y_2\\ \alpha_1 x_1-\alpha_1 y_1+\beta x_2-\beta y_2\end{pmatrix} = \alpha\begin{pmatrix}x_1+y_1\\x_1-y_1\end{pmatrix}+\beta\begin{pmatrix}x_2+y_2\\x_2-y_2\end{pmatrix} = \alpha L\begin{pmatrix}x_1\\y_1\end{pmatrix}+\beta L\begin{pmatrix}x_2\\y_2\end{pmatrix}$ ✓

$Nu(L) = \{v \in \mathbb{R}^2 / L(v) = 0\}$

$\begin{cases}x+y=0\\x-y=0 \to x=y\end{cases} \quad 2x=0 \to x=0, \; y=0 \quad \Rightarrow Nu(L) = \left\{\begin{pmatrix}0\\0\end{pmatrix}\right\} \Rightarrow \dim(Nu(L)) = 0$

↓ inyectiva ✓

$Im(L) = \{L(v) \; \forall v \in \mathbb{R}^2\} \qquad B = \left\{\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}0\\1\end{pmatrix}\right\}$

isomorfismo

endomorfismo

automorfismo

↓ biyectiva

L lineal

$L\left(\alpha\begin{pmatrix}1\\0\end{pmatrix}+\beta\begin{pmatrix}0\\1\end{pmatrix}\right) = \alpha L\begin{pmatrix}1\\0\end{pmatrix}+\beta L\begin{pmatrix}0\\1\end{pmatrix} = \alpha\begin{pmatrix}1\\1\end{pmatrix}+\beta\begin{pmatrix}1\\-1\end{pmatrix}$

↑ suryectiva

Todo $v \in \mathbb{R}^2$ se puede esc. como $\alpha\begin{pmatrix}1\\0\end{pmatrix}+\beta\begin{pmatrix}0\\1\end{pmatrix}$

$\Rightarrow Im(L) = \left\langle \begin{pmatrix}1\\1\end{pmatrix},\begin{pmatrix}1\\-1\end{pmatrix}\right\rangle = \mathbb{R}^2 \Rightarrow \dim(Im(L)) = 2$

Seguimos con el ejemplo...

- L es TL $\Rightarrow \exists A \in \mathbb{R}^{2\times 2} / L(v) = A\cdot v$ $\qquad A\cdot\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ x-y \end{pmatrix}$ ✓

$$L\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ x-y \end{pmatrix} \Rightarrow \dots \Rightarrow A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

- $N(A) = \{v / Av = 0\} = \left\{\begin{pmatrix} 0 \\ 0 \end{pmatrix}\right\} \Rightarrow \dim(N(A)) = 0$

- $EC(A) = \left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\rangle = \mathbb{R}^2 \Rightarrow \dim(EC(A)) = \dim(EF(A)) = 2$

- $EF(A) = \left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\rangle = \mathbb{R}^2$

Conclusiones ...

$\subset \mathbb{R}^m$ $\qquad$ $\subset \mathbb{R}^n$ $\qquad$ $\leq \min(n,m)$

Teorema: Sea $A \in \mathbb{R}^{m\times n}$, $dim(EC(A)) = dim(EF(A)) = r(A)$. Donde $r(A)$ se denomina rango de la matriz.

Definición: Se denomina nulidad de una matriz A a la dimensión de su espacio nulo $N(A)$, $n(A) = dim(N(A))$, siendo $N(A) = \{v \in \mathbb{R}^n, Av = 0\}$

Teorema de Rango-Nulidad: Para toda matriz $A \in \mathbb{R}^{m\times n}$ se verifica:

$n, r(A), n(A) \geq 0$

$$\underbrace{r(A)}_{\#\dim\ \text{“aprovechadas”}} + \underbrace{n(A)}_{\#\dim\ \text{“perdidas”}} = \underbrace{n}_{\#\dim\ \text{entrada}\ \mathbb{V}}$$

sea $L: \mathbb{V} \to \mathbb{W}$ una T.L.

$$\dim(\mathbb{V}) = \dim(\mathrm{Im}(L)) + \dim(\mathrm{Nu}(L))$$