Análisis Matemático para Inteligencia Artificial

Verónica Pastor (vpastor@fi.uba.ar),
Martín Errázquin (merrazquin@fi.uba.ar)

Especialización en Inteligencia Artificial

Clase 2

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:

$z = a + ib$

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

1. Para cada $\alpha \in \mathbb{R}$ (*o* $\mathbb{C}$), $u, v, w \in \mathbb{V}$.
   - a) $\Phi(u + v, w) = \Phi(u, w) + \Phi(v, w)$ — Φ lineal
   - b) $\Phi(\alpha \bullet u, v) = \alpha \bullet \Phi(u, v)$ — $K \times K$
2. $\Phi(u, v) = \overline{\Phi(v, u)}$ — @$\mathbb{R}$ $\Phi(u, v) = \Phi(v, u)$
3. $\Phi(v, v) \geq 0$, y $\Phi(v, v) = 0$ sii $v = 0$

   si $v \neq 0_v \Rightarrow \Phi(v, v) > 0$

   si $v = 0_v \Rightarrow \Phi(v, v) = 0$

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 $\mathbb{C}^n$).

$v, w \in \mathbb{R}^n$ $\quad v \cdot w = \sum_{i=1}^{n} v_i \cdot w_i$ $\quad \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = 1 \cdot 4 + 2 \cdot 5 + 3 \cdot 6 = 32$

También hay otros espacios con productos internos ...

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

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

Verificar que cumple:

1. Para cada $\alpha \in \mathbb{R}$ (o $\mathbb{C}$), $f, g, h \in \mathcal{C}([-1,1])$.
   - $\Phi(f+g,h) = \Phi(f,h) + \Phi(g,h)$
   - $\Phi(\alpha \bullet f, g) = \alpha \bullet \Phi(f,g)$
2. $\Phi(f,g) = \overline{\Phi(g,f)}$
3. $\Phi(f,f) \geq 0$, y $\Phi(f,f) = 0$ sii $f(x) = 0, \forall x$

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

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) $<f,f> = \int_{-1}^{1} \underbrace{f(x)\overline{f(x)}}_{\|f(x)\|^2 \geq 0 \ \forall x} dx = \int_{-1}^{1} \|f(x)\|^2 dx \geq 0 \quad \forall f \in \mathcal{C}([-1,1])$

$= 0 \Leftrightarrow \|f(x)\|^2 = 0 \ \forall x \Leftrightarrow f(x) = 0$ ✓

Definición de Norma

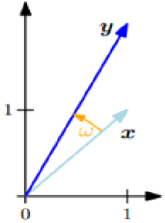
Sea $(\mathbb{V}, \langle .,. \rangle)$ un e.v. real (complejo) con p.i.. Sea $v \in \mathbb{V}$, se define la norma de v asociada a $\langle .,. \rangle$.

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

$$\cos(\omega) = \frac{\langle x,y \rangle}{||x|| \; ||y||} = \frac{\langle x,y \rangle}{\sqrt{\langle x,x \rangle \cdot \langle y,y \rangle}}$$



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

$\in [1; 1]$

Propiedades de la Norma

1. $\forall v \in \mathbb{V},\ ||v|| \geq 0$, y $||v|| = 0$ sii $v = 0$. ← heredado del p.i.
2. Sean $\alpha \in \mathbb{R}(o\ \mathbb{C})$, $v \in \mathbb{V}$, $||\alpha \bullet v|| = |\alpha|\ ||v||$.

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

   podría ser $\mathbb{C}$
3. Desigualdad de Cauchy Schwartz: si $u, v \in \mathbb{V}$ entonces

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

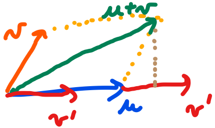
$\langle u, v\rangle = ||u|| \cdot ||v|| \cdot \cos(w)$, $\cos(w) \in [-1;1]$

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

$\leq ||u|| \cdot ||v||$

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

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



$||u+v|| = ||u|| + ||v||$

$\Updownarrow$

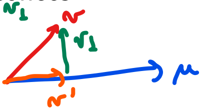
u, v son colineales

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 $||u + v||^2 = ||u||^2 + ||v||^2$.

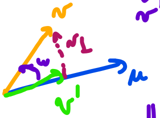
$v = v' + v_\perp$ con $v' \parallel u$

$\|v\| = \|v'\|^2 + \|v_\perp\|^2$



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.

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



$\check{v}' = \check{u}$

$\|v'\| = \|v\| \cdot \cos(\omega)$

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

Proceso de Ortogonalización de Gram Schmidt

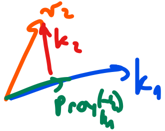
Def: Una base ortonormal (BON) de un E.V. es una base $B = \{v_1, ..., v_n\}$ que satisface:

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

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:

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



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

$\check{k}_i = \dfrac{k_i}{||k_i||}$

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

BOG

$\hookrightarrow \tilde{B} = \{\check{k}_1, ..., \check{k}_n\}$

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: $S = \left\langle \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \right\rangle$ con p.i.

Quiero $S^{\perp}$

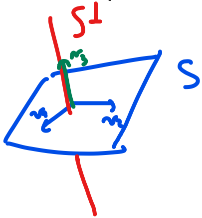
$\dim(S) = 2 \quad \dim(\mathbb{R}^3) = 3 \Rightarrow \dim(S^{\perp}) = 1 = 3 - 2$

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

$w \perp u$

$w \perp v$

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



$v_1 \perp v_3$

$v_2 \perp v_3$

$v_i \perp v_j$

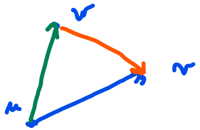
$\forall v_i \in S,$

$v_j \in S^{\perp}$

Distancia

$\Phi(\ ) \longrightarrow \|v\| = \sqrt{\Phi(v,v)}$

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



Propiedades:

1. $d(u, v) \geq 0, \ \forall u, v \in \mathbb{V}$
2. $d(u, v) = 0 \Leftrightarrow u = v$
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

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

$$x^T A x > 0, \ \forall x \in \mathbb{R}^n - \{0\}$$

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

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

$\vec{x} = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$

$$\vec{x}^T A \vec{x} = (x \ y)\begin{pmatrix} 2 & 3 \\ 3 & 9 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = (2x+3y, \ 3x+9y)\begin{pmatrix} x \\ y \end{pmatrix} = 2x^2+3xy+3xy+9y^2$$

$$= \underbrace{2x^2}_{} + \underbrace{6xy}_{} + \underbrace{9y^2}_{} = \underbrace{x^2}_{\geq 0} + \underbrace{(x+3y)^2}_{\geq 0} > 0 \quad \forall \vec{x} = \begin{pmatrix} x \\ y \end{pmatrix} \neq \vec{0}$$

$(x+3y)^2 = x^2 + 6xy + 9y^2$

si $x \neq 0 \Rightarrow x^2 > 0$

$x = 0, \ y \neq 0 \Rightarrow (x+3y)^2 > 0$

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:

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

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

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

$\forall x, y \in \mathbb{V},\ x \neq y \Rightarrow T(x) \neq T(y)$

- Inyectiva: si $\forall x,\ y \in \mathbb{V} : T(x) = T(y) \to x = y$
- Suryectiva: si $T(\mathbb{V}) = \mathbb{W}$ $T(\mathbb{V}) = \{T(v)\ \forall v \in \mathbb{V}\}$
- 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. iso + endo

$B_V = \{v_1, v_2, \dots, v_n\}$

$B_W = \{w_1, w_2, \dots, w_n\}$

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

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:

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

$\mathcal{P}_2[x]$

$v = 3x^2 - 5x + 10 = 10\begin{pmatrix}0\\0\\1\end{pmatrix} + (-5)\begin{pmatrix}0\\x\\0\end{pmatrix} + 3\begin{pmatrix}x^2\\0\\0\end{pmatrix}$

$B_{\mathcal{P}_2} = \left\{\begin{pmatrix}0\\0\\1\end{pmatrix}, \begin{pmatrix}0\\x\\0\end{pmatrix}, \begin{pmatrix}x^2\\0\\0\end{pmatrix}\right\}$

$1 \quad x \quad x^2$

$[v]_B = (10, -5, 3) \in \mathbb{R}^3$

es la img. del isomorfismo

$dim(\mathcal{P}_2) = 3 = \#B \longrightarrow \mathbb{R}^3$

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.** $\exists A \,/\, L(v) = A \cdot v$

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

$L : \mathbb{R}^n \to \mathbb{R}^m \,/\, L(v) = A \cdot v$

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

$A = \begin{pmatrix} | & & | \\ A_1 & \cdots & A_n \\ | & & | \end{pmatrix}$ $A_i \in \mathbb{R}^m$

- Espacio columna de A: es el subespacio de $\mathbb{R}^m$ generado por los n vectores columna de A:

con $A \in \mathbb{R}^{m\times n}$

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

- Espacio fila de A: e el subespacio de R^n generado por los m vectores fila de A:

$A = \begin{pmatrix} - A_1 - \\ \vdots \\ - A_m - \end{pmatrix}$ $A_i \in \mathbb{R}^n$

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

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? $L\left(\alpha\begin{pmatrix}x_1\\y_1\end{pmatrix}+\beta\begin{pmatrix}x_2\\y_2\end{pmatrix}\right)=L\begin{pmatrix}\alpha x_1+\beta x_2\\ \alpha y_1+\beta y_2\end{pmatrix}=$

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

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

$\begin{cases}x+y=0\\x-y=0\end{cases}\to x=y$ $\quad x+y=2x=0 \quad x=y=0$

$\Rightarrow Nu(L)=\left\{\begin{pmatrix}0\\0\end{pmatrix}\right\}$ → inyectiva

- $Im(L)=\{L(v)\ \forall v\in\mathbb{R}^2\}$

$L\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}1\\1\end{pmatrix}$

$L\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}1\\-1\end{pmatrix}$

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

L es biyectiva

L es isomorfismo

L es endomorfismo → L es automorfismo

Seguimos con el ejemplo...

si L es TL $\quad \exists A \in \mathbb{R}^{2\times 2} / L(v) = A \cdot v = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$

$\mathbb{R}^2 \to \mathbb{R}^2$

$\downarrow \quad \downarrow$

$n \quad m$

$$\begin{pmatrix} x+y \\ x-y \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix} \Rightarrow \begin{cases} x+y = ax+by \\ x-y = cx+dy \end{cases}$$

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

- $N(A) = \{ v \in \mathbb{R}^2 / Av = 0 \}$

$$\begin{cases} x+y=0 \\ x-y=0 \end{cases} \Longrightarrow N(A) = \left\langle \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\rangle = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} \to \dim(N(A)) = 0$$

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

$$\searrow \dim(EC(A)) = \dim(EF(A)) = 2$$

∇

- $EF(A) = \langle (1,1)^T, (1,-1)^T \rangle \nearrow$

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

Teorema: Sea $A \in \mathbb{R}^{m\times n}$, $dim(\overbrace{EC(A)}^{\in \mathbb{R}^m}) = dim(\overbrace{EF(A)}^{\in \mathbb{R}^n}) = r(A)$. 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:

$$\underbrace{r(A)}_{\substack{\#\,dim \\ \text{"útiles"}}} + \underbrace{n(A)}_{\substack{\#\,dim \\ \text{"perdidas"}}} = n \quad \#\,dim\ \text{entrada}$$

si $L: \mathbb{V} \to \mathbb{W}$ a TL

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