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:

$z = a + ib$, $a, b \in \mathbb{R}$

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

1. Para cada $\alpha \in \mathbb{R}$ (*o* $\mathbb{C}$), $u, v, w \in \mathbb{V}$.
   $$\left.\begin{array}{l} \bullet\ \Phi(u + v, w) = \Phi(u, w) + \Phi(v, w) \\ \bullet\ \Phi(\alpha \bullet u, v) = \alpha \bullet \Phi(u, v) \end{array}\right\} \text{lineal}$$
2. $\Phi(u, v) = \overline{\Phi(v, u)}$
3. $\Phi(v, v) \geq 0$, y $\Phi(v, v) = 0$ sii $v = 0$

Si $\mathbb{K} = \mathbb{R} \Rightarrow \Phi(u, v) = \Phi(v, u)$

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 $\mathbb{C}^n$).

$$\left\langle \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} \right\rangle = 1 \cdot 4 + 2 \cdot 5 + 3 \cdot 6 = 4 + 10 + 18 = 32$$

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.

$! \; z \cdot \bar{z} = \|z\|^2 = a^2 + b^2 \geq 0$

$z = a + b\,i \quad \forall a, b \in \mathbb{R}$

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

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

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

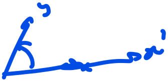
2) $\langle f,g\rangle = \int_{-1}^{1} f(x)\cdot\overline{g(x)}\,dx = \int_{-1}^{1} \overline{\overline{f(x)}\cdot g(x)} = \overline{\int_{-1}^{1} \overline{f(x)}\cdot g(x)\,dx} = \overline{\int_{-1}^{1} g(x)\cdot\overline{f(x)}\,dx} = \overline{\langle g,f\rangle}$ ✓

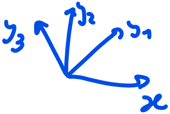
3) $\langle f,f\rangle = \int_{-1}^{1} f(x)\cdot\overline{f(x)}\,dx = \int_{-1}^{1} \underbrace{\|f(x)\|^2}_{\geq 0}dx \geq 0$

$\Rightarrow \langle f,f\rangle = 0 \iff \|f(x)\|^2 = 0 \;\forall x \Updownarrow f(x) = 0 \;\forall x$

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 $\langle ., . \rangle$.



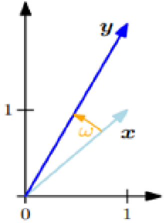


Notación: $||v|| = \langle v, v \rangle^{1/2}$

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

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

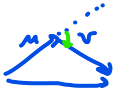


$$\langle x,y \rangle = ||x|| \, ||y|| \cdot \underbrace{\cos(w)}_{\in [-1;1]}$$

Propiedades de la Norma

③ de P.I.

1. $\forall v \in \mathbb{V},\ ||v|| \geq 0$, y $||v|| = 0$ sii $v = 0$.
2. Sean $\alpha \in \mathbb{R}(o\ \mathbb{C})$, $v \in \mathbb{V}$, $||\underbrace{\alpha \bullet v}_{\mathbb{K}\times\mathbb{V}}|| = |\underbrace{\alpha| \cdot ||v}_{\mathbb{R}/\mathbb{C}\times\mathbb{R}}||$. $\longleftarrow$ $\langle \alpha\cdot v, \alpha v\rangle = \alpha\bar{\alpha}\cdot\langle v,v\rangle = \|\alpha\|^2\cdot\|v\|^2$
3. Desigualdad de Cauchy Schwartz: si $u, v \in \mathbb{V}$ entonces
$$|\langle u, v\rangle| \leq ||u||\ ||v||$$
$|\langle u,v\rangle| = \big|\, \|u\|\cdot\|v\| \cos(\omega) \big| = \|u\|\cdot\|v\|\cdot\underbrace{|\cos(\omega)|}_{\in[0,1]} \leq \|u\|\cdot\|v\|$
4. Desigualdad Triangular: si $u, v \in \mathbb{V}$ entonces
$$||u + v|| \leq ||u|| + ||v||$$



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

$\Updownarrow$

u y v no colineales

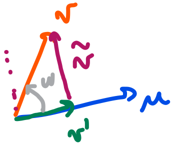
Ortogonalidad

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

$v = \tilde{v} + v' \quad \tilde{v} \perp v'$

$||v||^2 = ||\tilde{v}||^2 + ||v'||^2$



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.

$\tilde{v} = v - v' = v - P_u(v)$

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

$$\underbrace{||v||.\cos(w)}_{\frac{\langle u, v \rangle}{||u||}} \cdot \underbrace{\tilde{u}}_{\frac{u}{||u||}}$$

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

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:

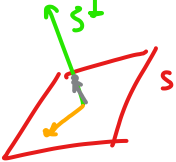
$$k_3 = v_3 - Proy_{k_1}(v_3) - Proy_{k_2}(v_3)$$

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

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: $S = \langle \underbrace{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}}_{u}, \underbrace{\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}}_{v} \rangle$ con prod. escalar.

Buscamos $S^{\perp}$

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

$$w = u \times v = \begin{vmatrix} i & j & k \\ 1 & 1 & 1 \\ 1 & 2 & 0 \end{vmatrix} = (-2, 1, 1) \Rightarrow S^{\perp} = \left\langle \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \right\rangle$$

$w \perp u$

$w \perp v$

Distancia

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 $d(u,v) = ||u - v||$. $= \langle u-v, u-v \rangle^{1/2}$

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: $\underbrace{A = \begin{pmatrix} 2 & 3 \\ 3 & 9 \end{pmatrix}}_{\in \mathbb{R}^{2\times 2}} \rightarrow A^t = \begin{pmatrix} 2 & 3 \\ 3 & 9 \end{pmatrix} = A \rightarrow$ simétrico ✓

$\bar{x} = \begin{pmatrix} x \\ y \end{pmatrix}$

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

$$= 2x^2 + 6xy + 9y^2 =$$

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

$$= x^2 + (x^2 + 6xy + 9y^2) = \underbrace{x^2}_{\geq 0} + \underbrace{(x + 3y)^2}_{\geq 0} > 0 \quad \forall (x, y) \in \mathbb{R}^2 / \{0\}$$

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: o $x \neq y \Rightarrow T(x) \neq T(y)$

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

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

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:

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

$\mathcal{P}_2[x]$:

$v = 3x^2 - 5x + 10$

$[v]_B = \underbrace{(10, -5, 3)}_{\text{es la imagen}} \in \mathbb{R}^3$

de $L: \mathcal{P}_2[x] \to \mathbb{R}^3$

$B = \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ x \\ 0 \end{pmatrix}, \begin{pmatrix} x^2 \\ 0 \\ 0 \end{pmatrix} \right\}$

$1 \quad x \quad x^2$

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

Núcleo e Imagen de una transformación

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

$A \cdot v \in \mathbb{R}^n$

$v \in \mathbb{R}^n$

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.** $L(v) = A \cdot v$

$v_1, v_2 \in N(A) \Rightarrow A(\alpha v_1 + \beta v_2) = \alpha \cdot A \cdot v_1 + \beta \cdot A \cdot v_2 = \alpha \cdot 0 + \beta \cdot 0 = 0$

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

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

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

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

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

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

Veamos un ejemplo...

$$L:\mathbb{R}^2\to\mathbb{R}^2 \qquad L\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}x+y\\x-y\end{pmatrix}$$

- ¿es TL?

$$L(\alpha v_1+\beta v_2)=L\left(\alpha\begin{pmatrix}x_1\\y_1\end{pmatrix}+\beta\begin{pmatrix}x_2\\y_2\end{pmatrix}\right)=L\begin{pmatrix}\alpha x_1+\beta x_2\\ \alpha y_1+\beta y_2\end{pmatrix}=\begin{pmatrix}\alpha x_1+\beta x_2+\alpha y_1+\beta y_2\\ \alpha x_1+\beta x_2-\alpha y_1-\beta y_2\end{pmatrix}=$$

$$=\begin{pmatrix}\alpha(x_1+y_1)+\beta(x_2+y_2)\\ \alpha(x_1-y_1)+\beta(x_2-y_2)\end{pmatrix}=\alpha\begin{pmatrix}x_1+y_1\\x_1-y_1\end{pmatrix}+\beta\begin{pmatrix}x_2+y_2\\x_2-y_2\end{pmatrix}=\alpha L\begin{pmatrix}x_1\\y_1\end{pmatrix}+\beta L\begin{pmatrix}x_2\\y_2\end{pmatrix}=\alpha L(v_1)+\beta L(v_2)$$

✓

- $Nu(L)=\left\{v\in\mathbb{R}^2 / L(v)=\begin{pmatrix}0\\0\end{pmatrix}\right\}$

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

$\to$ L es inyectiva

- $Im(L)=$ $\quad L\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}1\\1\end{pmatrix},\quad L\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}1\\-1\end{pmatrix}$

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

LI

L es biyectiva

isomorfismo

automorfismo

Seguimos con el ejemplo...

A: L es TL, $\exists A \in \mathbb{R}^{2\times 2} / L(v) = A \cdot v$

$$A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \Rightarrow A \cdot v = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ x-y \end{pmatrix} = L(v)$$

$$N(A) = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} \longrightarrow \dim(N(A)) = 0$$

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

$$\|$$

$$\cdot\ EF(A) = \left\langle (1,1)^T, (1,-1)^T \right\rangle = \mathbb{R}^2 \rightarrow \dim(EF(A)) = 2$$

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

$$\underbrace{r(A)}_{\#\dim\ \text{"mapeables"}} + \underbrace{n(A)}_{\#\dim\ \text{"perdidos"}} = \underbrace{n}_{\#\ \dim\ \text{entrada}}$$

$$A' = \begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix}$$

(si L es TL

$$\dim(V) = \dim(Nu(L)) + \dim(Im(L))$$