Análisis Matemático para Inteligencia Artificial

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

Especialización en Inteligencia Artificial

1/7/2022

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:

1. Para cada $\alpha \in \mathbb{R}$ (*o* $\mathbb{C}$), $u, v, w \in \mathbb{V}$.
   - $\Phi(u + v, w) = \Phi(u, w) + \Phi(v, w)$
   - $\Phi(\alpha \bullet u, v) = \alpha \bullet \Phi(u, v)$
2. $\Phi(u, v) = \overline{\Phi(v, u)}$ conjugado
3. $\Phi(v, v) \geq 0$, y $\Phi(v, v) = 0$ sii $v = 0$

análog $\Phi(u, v+w)$

$\langle p(x) + q(x), r(x) \rangle =$

$(p(x) + q(x)) . r(x) =$

$p(x) \cdot r(x) + q(x) \cdot r(x) =$

$\langle p(x), r(x) \rangle + \langle q(x), r(x) \rangle$

$\in \mathbb{R}$ $(x = x_0)$

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 $\mathbb{C}^n$).

$u \cdot w = (a, b) \cdot (c, d) = (ac + bd)$

$\hookrightarrow \in V : \mathbb{R}^2$

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.

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

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 $f(x) = 0, \forall x$

$\mathrm{Re}(z)$, $\mathrm{Im}(z)$

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

$\bar{z} = a - bi$

$\bar{\bar{z}} = a + bi \Rightarrow \bar{\bar{z}} = z$

$|z| = \sqrt{a^2+b^2}$

$z \cdot \bar{z} = a^2 + b^2 = |z|^2$

$\int_{-1}^{1} f\bar{f}\,dx = \int_{-1}^{1} f^2 dx \geq 0$

Dem: $\langle f+g, h\rangle \underset{\text{def } \langle,\rangle}{=} \int_{-1}^{1} (f(x)+g(x))\overline{h(x)}\,dx \overset{\text{distr}}{=} \int_{-1}^{1} \left[\left(f(x)\overline{h(x)}\right) + \left(g(x)\overline{h(x)}\right)\right] dx$

$\underset{\text{prop integr}}{=} \int_{-1}^{1} f(x)\overline{h(x)}\,dx + \int_{-1}^{1} g(x)\overline{h(x)}\,dx \underset{\text{def } \langle,\rangle}{=} \langle f, h\rangle + \langle g, h\rangle$

$C[-1,1]$ cont.

(2) $\langle f, g\rangle \overset{\text{def } \langle,\rangle}{=} \int_{-1}^{1} f(x)\overline{g(x)}\,dx \underset{\bar{\bar{z}} = z}{=} \int_{-1}^{1} \overline{\overline{f(x)} \cdot g(x)}\,dx = \int_{-1}^{1} \overline{\overline{f(x)}\, g(x)}\,dx =$

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

$\overline{z_1 \cdot z_2} = \bar{z}_1 \cdot \bar{z}_2$

Definición de Norma

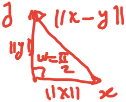
Sea $(\mathbb{V}, \langle ., . \rangle)$ un e.v. real (complejo) con p.i.. Sea $v \in \mathbb{V}$, se define la norma de v asociada a $\langle ., . \rangle$.

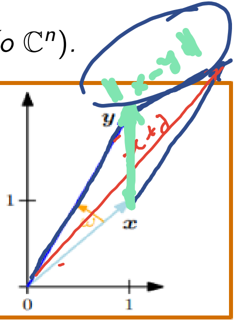
Notación: $||v|| = \langle v, v \rangle^{1/2}$

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

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





def norma

dem) $\|x-y\|^2 = \langle x-y, x-y \rangle = (\langle x, x-y \rangle) + (\langle y, x-y \rangle) =$

$f+g$ $\quad h$

$= (\langle x, x \rangle - \langle x, y \rangle) - (\langle y, x \rangle + \langle y, y \rangle) = \|x\|^2 - 2\langle x, y \rangle + \|y\|^2 =$

$= \|x\|^2 + \|y\|^2 - 2\|x\|\|y\|\cos w$

Ley de cos: $a^2 + b^2 - 2ab\cos\alpha = c^2$

Propiedades de la Norma

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

$$\underbrace{|\langle u, v\rangle| \leq ||u||\ ||v||}$$

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

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

dem ③: Definiendo $w = u - \underbrace{\frac{\langle u, v\rangle}{||v||^2}}_{k} \cdot v$. Luego

$$0 \leq ||w||^2 = \langle u - k\cdot v,\ u - k\cdot v\rangle \underset{props}{=} \langle u, u\rangle - \langle u, k\cdot v\rangle - \langle k\cdot v, u\rangle + \langle k\cdot v, k\cdot v\rangle$$

$$= ||u||^2 - 2k\langle u, v\rangle + k^2 ||v||^2 = ||u||^2 - 2\underbrace{\frac{\langle u, v\rangle}{||v||^2}\langle u, v\rangle}_{\langle u,v\rangle^2} + \frac{\langle u, v\rangle^2}{||v||^4}||v||^2$$

$$= ||u||^2 - \frac{\langle u, v\rangle^2}{\underbrace{||v||^2}_{k}} \geq 0 \iff \sqrt{||u||^2 ||v||^2} \geq \sqrt{\langle u, v\rangle^2}$$

$$||w||^2 \geq 0$$

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

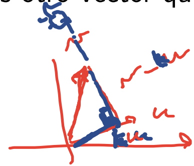
Def: $(\mathbb{V}, \langle ., . \rangle)$ un $\mathbb{K}$-EV (con $\mathbb{K} = \mathbb{R}$ *o* $\mathbb{C}$) con p.i. dos vectores $u, v \in \mathbb{V}$ se dicen ortogonales si $\langle u, v \rangle = 0$.

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

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:

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



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

Si $\langle v_i, v_i \rangle \neq 1$

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

ent $u_i = \dfrac{v_i}{\|v_i\|} \to \|u_i\| = \left\| \dfrac{v_i}{\|v_i\|} \right\| = 1 = \dfrac{\|v_i\|}{\|v_i\|}$

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:

- $k_1 = v_1$
- $k_2 = v_2 - Proy_{k_1}(v_2)$

$k_3 = v_3 - \left(Proy_{k_1}(v_3) + Proy_{k_2}(v_3) \right)$

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

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

Complemento Ortogonal

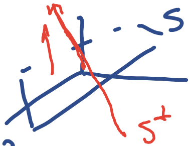
Sea $\mathbb{V}$ un EV de dimensión $n < \infty$ y $S \subset \mathbb{V}$ un SEV de dimensión $m \leq n$. El complemento ortogonal $(S^{\perp})$ es un SEV de de dimensión $n - m$ que satisface:

$$S \cup S^{\perp} = \mathbb{V} \quad \wedge \quad S \cap S^{\perp} = \emptyset \quad (\text{sólo } 0 \in S \cap S^{\perp})$$

Ejemplo: Sea $S = \text{gen}\left\{ \begin{pmatrix} 1\\1\\1 \end{pmatrix}, \begin{pmatrix} 1\\2\\0 \end{pmatrix} \right\}$ si: $\left\langle \begin{pmatrix} 1\\1\\1 \end{pmatrix}, \begin{pmatrix} 1\\2\\0 \end{pmatrix} \right\rangle = 1.1 + 1.2 + 1.0 \neq 0$ l.i.c

$\dim S = 2$

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



$\therefore S^{\perp} = \text{gen}\{(-2, -1, 1)^T\}$ $\quad \dim S^{\perp} = 1$

$\dim S + \dim S^{\perp} = \dim \mathbb{V} \quad (\mathbb{V} = \mathbb{R}^3)$

2) Por Gram-Schmidt $\quad k_1 = \begin{pmatrix} 1\\1\\1 \end{pmatrix}, \; k_2 = v - \dfrac{\overset{3}{\langle u, v \rangle}}{\underset{3}{\langle u, u \rangle}} \cdot u = \begin{pmatrix} 0\\1\\-1 \end{pmatrix}$

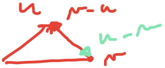
$k_3 = v_3 - P_{k_1}(v_3) - P_{k_2}(v_3)$, digo $v_3 = \begin{pmatrix} 0\\0\\1 \end{pmatrix}$

$= \begin{pmatrix} 0\\0\\1 \end{pmatrix} - \dfrac{\langle k_1, v_3 \rangle}{\langle k_1, k_1 \rangle} \cdot k_1 - \dfrac{\langle k_2, v_3 \rangle}{\langle k_2, k_2 \rangle} \cdot k_2 = \begin{pmatrix} 0\\0\\1 \end{pmatrix} - \dfrac{1}{3}\begin{pmatrix} 1\\1\\1 \end{pmatrix} - \dfrac{(-1)}{2}\begin{pmatrix} 0\\1\\-1 \end{pmatrix} =$

Distancia

$= \begin{pmatrix} -1/3 \\ 1/6 \\ 1/6 \end{pmatrix} = \frac{1}{6}\begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}$

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 $d(u, v) = ||u - v||$. $= ||v - u||$



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.

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

$$\underbrace{x^T}_{1\times n}\underbrace{A}_{n\times n}\underbrace{x}_{n\times 1} > 0, \ \forall x \in \mathbb{R}^n - \{0\} \quad \to 1\times 1$$

$A = A^T$

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

Ejemplo: $A = \begin{bmatrix} 2 & 3 \\ 3 & 9 \end{bmatrix} = A^T$, calculo $x^T A x$ siendo $x = \begin{pmatrix} x \\ y \end{pmatrix}$

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

$$= \underbrace{2x^2}_{\geq 0} + \underbrace{6xy}_{?} + \underbrace{(3y)^2}_{\geq 0} = x^2 + x^2 + 2.x(3y) + (3y)^2 =$$

$2(3y)x$

$(a+b)^2 = a^2 + 2ab + b^2$

$$= \underbrace{x^2}_{\geq 0} + \underbrace{(x+3y)^2}_{\geq 0} \geq 0$$

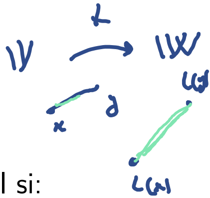
Teorema: Sea $\mathbb{V}$ un EV de dimensión finita, y B una base de $\mathbb{V}$, vale que es un p.i. sii existe una matriz definida positiva tal que:

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

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:

- 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:

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

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

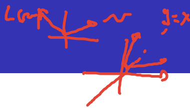
$= \begin{pmatrix} \alpha^2 x_1 y_1 + \alpha\beta x_2 y_1 + \alpha\beta x_1 y_2 + \beta^2 x_2 y_2 \\ \alpha y_1 + \beta y_2 \end{pmatrix}$

$\neq \alpha L\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \beta\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$

- 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.

Representaciones

1) Define $L(v) = (\quad)$
2) Gráficamente: rotación
3) Indicar $L(\text{Base})$



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:

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

$$p(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 \overset{\text{isom}}{\Longleftrightarrow} (a_0, a_1, a_2, a_3)$$

$$L(v) = L\left(\alpha_1 \begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix} + \alpha_2 \begin{pmatrix}0\\1\\0\\\vdots\end{pmatrix} + \dots + \alpha_n \begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}\right) \underset{L \text{ es T.L}}{=} \alpha_1 L(e_1) + \alpha_2 L(e_2) + \dots + \alpha_n L(e_n)$$

$$= \underbrace{(\alpha_1, \dots, \alpha_n)}_{\text{coord.}} \underbrace{\begin{bmatrix} \downarrow & & \downarrow \\ L(e_1) & \dots & L(e_n) \\ & & \end{bmatrix}}_{A = [L]_{B_c}^{B_c}}$$

$\tilde{x}$ ó $\tilde{y}$

↳ representación matricial

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.**

- 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...

Seguimos con el ejemplo...

Conclusiones ...

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:

$$\boxed{r(A) + n(A) = n}$$