Análisis Matemático para Inteligencia Artificial

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

Especialización en Inteligencia Artificial

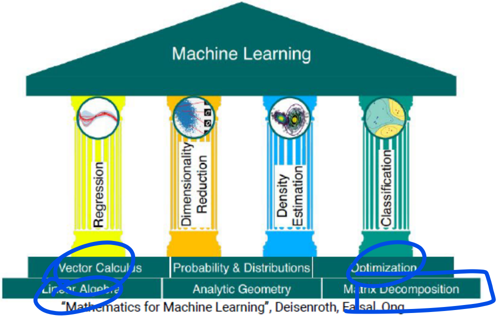
3/~~2~~3/2023

Presentación de la Materia

¿Por qué estudiar Análisis Matemático?

A medida que Machine Learning se vuelve más común, y los paquetes de software se vuelven más simples de usar, uno se abstrae cada vez más de los detalles técnicos que hay detrás. **Modelo de caja negra.**

Esto trae el **peligro** de desconocer las decisiones de diseño y las limitaciones de cada algoritmo.

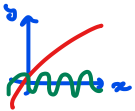


Bibliografía Recomendada: **Mathematics for Machine Learning**
Marc Peter Deisenroth, A. Aldo Faisal, and Cheng Soon Ong.
Published by Cambridge University Press (2020).
Está disponible gratis en `http://mml-book.github.io/`

Motivación (I)

Operaciones vectorizadas y GPUs: Early colab!

Motivación (II)



- Regresión logística asume superposición entre clases
- K-Means asume clusters esféricos
- Árboles de decisión tienen fronteras de decisión en forma de hiperplanos
- ¿Por qué las Redes Neuronales se entrenan más rápido usando GPUs?
- ¿Por qué en las redes neuronales importa la escala y en los árboles de decisión no?
- En kNN ¿Es lo mismo maximizar producto interno que minimizar distancia euclidea?

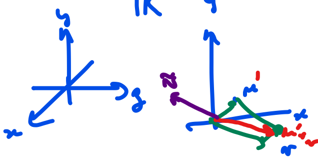
Clase 1: Espacios Vectoriales

Empecemos considerando los vectores $u = \underbrace{(1, \frac{1}{2}, \frac{1}{3})}_{\mathbb{R}^3}$ y $v = \underbrace{(2, -1)}_{\mathbb{R}^2}$, ¿podemos sumar los vectores?

MAL$_1$: $(1, \frac{1}{2}, \frac{1}{3}) + (2, -1, 0)$

MAL$_2$: $(1, \frac{1}{2}, \frac{1}{3}) + (0, 2, -1)$

$u' = (1, \frac{1}{2})$



¿qué ocurre si tomamos $\tilde{v} = (-2, 1)$?

$\tilde{v} = (-1) \cdot v$ se puede $u' + \tilde{v} \iff$ se puede $u' + v$

$\therefore$ Para definir correctamente $u + v$ deben estar en el mismo conjunto, y es posible cambiar el sentido y tamaño del vector. Es decir, que si el espacio vectorial es $\mathbb{R}^n$ podemos realizar estas operaciones:

1. $x \in \mathbb{R}^n, y \in \mathbb{R}^n \rightarrow x + y \in \mathbb{R}^n$
2. $x \in \mathbb{R}^n, k \in \mathbb{R} \rightarrow kx \in \mathbb{R}^n$

Algunas definiciones...

Sea $\mathbb{V} \neq \emptyset$, se define una operación (o suma) a una función $+ : \mathbb{V} \times \mathbb{V} \to \mathbb{V}$.

Esta operación se espera que cumpla con las siguientes propiedades:

1. Asociativa: $(x + y) + z = x + (y + z), \forall x, y, z \in \mathbb{V}$.
2. Elemento Neutro: $\forall x \in \mathbb{V}, \exists e \in \mathbb{V}$ tal que $x + e = e + x = x$.
3. Opuesto: $\forall x \in \mathbb{V}, \exists \tilde{x} \in \mathbb{V}$ tal que $x + \tilde{x} = \tilde{x} + x = e$.
4. Conmutativa: $\forall x, y \in \mathbb{V}, x + y = y + x$.

Sean $\mathbb{V} \neq \emptyset, \mathbb{K} \neq \emptyset$, se define una operación (o producto escalar) a una función $\bullet : \mathbb{K} \times \mathbb{V} \to \mathbb{V}$. Este conjunto $\mathbb{K}$ es generalmente $\mathbb{R}$ *o* $\mathbb{C}$ es un cuerpo de escalares.

Comentario*: Un cuerpo es un conjunto con algunas operaciones sobre los elementos de éste, que se comportan como la adición, sustracción, multiplicación y división que cumplen con las propiedades que conocemos. Para no especificar el cuerpo se usa la palabra **escalar**.*

¿Estudiar espacios vectoriales sólo sirve para vectores?

$\nearrow (1, \frac{1}{2}, \frac{3}{4})$ $\quad \nearrow (2,-1,0)$

Sean $p(x) : 1 + \frac{1}{2}x + \frac{3}{4}x^2$ y $q(x) : 2 - x$, ¿valen 1 y 2?

1. $p(x) + q(x) = (p + q)(x) = (1+2)x^0 + (\frac{1}{2} - 1)x^1 + \left(\frac{3}{4} + 0\right)x^2 \in \mathcal{P}_2[x]$ $= 3 + (-\frac{1}{2})x + \frac{3}{4}x^2$
2. $\overset{\in \mathbb{R}}{k} \cdot p(x) = (kp)(x) = k + \frac{k}{2}x + \frac{3}{4}k\, x^2 \quad \in \mathcal{P}_2[x]$

hay una biyección e/ $\mathbb{R}^3$ y $\mathcal{P}_2[x]$

Repasemos el producto de polinomios: $p(x) \cdot q(x) =$

$(1 + \frac{1}{2}x + \frac{3}{4}x^2)(2 - x) = 2 - x + x - \frac{1}{2}x^2 + \frac{3}{2}x^2 - \underbrace{\frac{3}{4}x^3}_{!!}$

↳ prod NO es cerrado

Sean $A, B \in \mathbb{R}^{n\times n}$, ¿valen 1 y 2?

- $A + B = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ \vdots & & \ddots & \vdots \\ a_{n1}+b_{n1} & \cdots & & a_{nn}+b_{nn} \end{pmatrix} \in \mathbb{R}^{n\times n}$

- $A \cdot B \in \mathbb{R}^{n\times n}$

↳ prod SÍ es cerrado

- $k \cdot A = \begin{pmatrix} k \cdot a_{11} & \cdots & k \cdot a_{1n} \\ \vdots & \ddots & \\ k \cdot a_{n1} & \cdots & k \cdot a_{nn} \end{pmatrix} \in \mathbb{R}^{n\times n}$

wrt = with respect to

Diremos que $\mathcal{V} = (\mathbb{V}, +, \mathbb{K}, \bullet)$ es un espacio vectorial si $\mathbb{K}$ y $\mathbb{V}$ son conjuntos no vacíos y la operación $+$ en $\mathbb{V}$, y la acción $\bullet$ de $\mathbb{K}$ en $\mathbb{V}$ cumplen:

1. $+$ es asociativa — $\forall a,b,c \in \mathbb{V}$ $(a+b)+c = a+(b+c)$
2. $+$ tiene elemento neutro — $\exists e \in \mathbb{V} \,/\, \forall v \in \mathbb{V}$ $v+e = e+v = v$
3. $+$ tiene elemento inverso — $\forall v \in \mathbb{V}$ $\exists \tilde{v} \in \mathbb{V}$ $\tilde{v}+v = v+\tilde{v} = e$
4. $+$ es conmutativa — $\forall v,w \in \mathbb{V}$ $v+w = w+v$
5. $\alpha \bullet (v+w) = \alpha \bullet v + \alpha \bullet w, \ \forall \alpha \in \mathbb{K}, \forall v, w \in \mathbb{V}$ — distr. del escalar wrt la suma
6. $(\alpha + \beta) \bullet v = \alpha \bullet v + \beta \bullet v, \ \forall \alpha, \beta \in \mathbb{K}, \forall v \in \mathbb{V}$ — ojo — distr. de la suma de escalares wrt al prod. x escalar
7. $\bullet$ tiene elemento neutro: $1 \bullet v = v, \forall v \in \mathbb{V}$
8. $\bullet$ es asociativa:

$$\alpha \bullet (\beta \bullet v) = (\underbrace{\alpha\beta}_{\text{ojo}}) \bullet v, \ \forall \alpha, \beta \in \mathbb{K}, \forall v \in \mathbb{V}$$

(*) grupo conmutativo

Sea $\mathcal{V} = (\mathbb{V}, +, \mathbb{K}, \bullet)$ un espacio vectorial, un subconjunto $S \subseteq \mathbb{V}$, $S \neq \emptyset$ se dice que es un subespacio de $\mathbb{V}$ si la suma y el producto por escalares de $\mathbb{V}$ son una operación y una acción en S que lo convierten en un $\mathbb{K}$-espacio vectorial.

Condiciones necesarias y suficientes para caracterizar subespacios

S es un subespacio en un $\mathbb{K}$-espacio vectorial sii:

1. $S \neq \emptyset \quad (0 \in S)$
2. $v, w \in S \rightarrow v + w \in S$ → suma es cerrada en S
3. $\alpha \in \mathbb{K}, v \in S \rightarrow \alpha \bullet v \in S$ → prod. x escalar es cerrado en S

Subespacios triviales

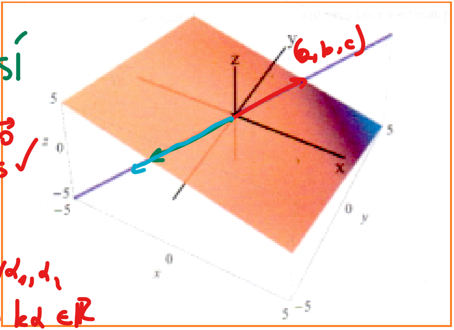
- $\{0\} \subseteq \mathbb{V}$
- $\mathbb{V} \subseteq \mathbb{V}$

1) $0 \in \mathbb{V}$ ✓

2) $0+0=0 \in \mathbb{V}$ ✓

3) $k \cdot 0 = 0 \in \mathbb{V}$ $\forall k \in \mathbb{K}$ ✓

Sea $\mathcal{V} = (\mathbb{R}^3, +, \mathbb{R}, \bullet)$, sea
$S = \{\alpha \bullet (a, b, c), \alpha \in \mathbb{R}\}$,
¿es un subespacio vectorial? SÍ



1) si $\alpha = 0 \Rightarrow 0(a,b,c) = (0,0,0) = \vec{0} \in S$ ✓

2) $(\alpha_1 a, \alpha_1 b, \alpha_1 \cdot c) + (\alpha_2 a, \alpha_2 b, \alpha_2 c)$
$= ((\alpha_1+\alpha_2)a, (\alpha_1+\alpha_2)b, (\alpha_1+\alpha_2)c)$
$= (\alpha_1+\alpha_2)(a,b,c) \in S \quad \forall \alpha_1, \alpha_2$

3) $k \cdot (\alpha \cdot (a,b,c)) = (k\alpha)(a,b,c) \in S$ con $k\alpha \in \mathbb{R}$

Algunos ejemplos más ...

Sean S, T subespacios de $\mathcal{V} = (\mathbb{V}, +, \mathbb{K}, \bullet)$. Probar si también los son:

1. $S \cap T = \{v \in \mathbb{V} : v \in S \wedge v \in T\} \subseteq \mathbb{V}$.
2. $S + T = \{v \in \mathbb{V} : v = s + t, s \in S, t \in T\} \subseteq \mathbb{V}$.
3. $S \cup T = \{v \in \mathbb{V} : v \in S \vee v \in T\} \subseteq \mathbb{V}$.

$S + T = \{v \in \mathbb{V} : v = s + t, s \in S, t \in T\} \subseteq \mathbb{V}$

1). ¿$\vec{0} \in S+T$? $\quad \vec{0} \in S, T$ por ser S.E. $\Rightarrow \underbrace{\vec{0}}_{S} + \underbrace{\vec{0}}_{T} = \vec{0} \Rightarrow \vec{0} \in S+T$

2) ¿$v+w \in S+T$ $\forall v, w \in S+T$?

$v \in S+T \Rightarrow \exists s_v \in S, t_v \in T \,/\, s_v + t_v = v$

$w \in S+T \Rightarrow \exists s_w \in S, t_w \in T \,/\, s_w + t_w = w$

por S.E. $s_v + s_w = s_{v+w} \in S$

$t_v + t_w = t_{v+w} \in T$

$v+w = s_v + t_v + s_w + t_w$

$= \underbrace{s_{v+w}}_{\in S} + \underbrace{t_{v+w}}_{\in T} \in S+T$

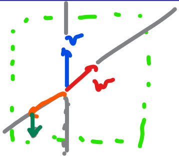
3) ¿$k.v \in S+T$ $\forall k \in \mathbb{K}, v \in S+T$?

$v = s_v + t_v$

$k.v = k.(s_v + t_v) = \overbrace{k.s_v}^{\in S} + \overbrace{k.t_v}^{\in T}$

$= s_{v'} + t_{v'} \in S+T$

Representación de subespacios



Sistemas generadores

Definición: Sea $\mathcal{V}$ un espacio vectorial, y $G = \{v_1, \ldots, v_r\} \subseteq \mathbb{V}$. Una combinación lineal de G es un elemento $v \in \mathbb{V}$ tal que $v = \sum_{i=1}^{r} \alpha_i \bullet v_i$, donde $\alpha_i \in \mathbb{K}$, para $i = 1, \ldots, r$. ↳ NOTA: v_i es CL de G $\forall v_i \in G$

Definición: Sea $\mathcal{V}$ un espacio vectorial, y $G \subseteq \mathbb{V}$. Se dice que G es un sistema de generadores de $\mathcal{V}$ si todo elemento de $\mathbb{V}$ es una combinación lineal de G.
Notación: $\langle G \rangle = \mathbb{V}$.

Ejemplo

$2\begin{pmatrix}1\\1\\1\end{pmatrix} + 1\begin{pmatrix}2\\1\\0\end{pmatrix} = \begin{pmatrix}4\\3\\2\end{pmatrix}$

Sea $G = \left\{ \begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}2\\1\\0\end{pmatrix}, \begin{pmatrix}4\\3\\2\end{pmatrix} \right\}$ dado un vector cualquiera $v \in \mathbb{R}^3$,

¿podemos escribirlo como combinación lineal de los vectores de G? $\Leftrightarrow \langle G \rangle = \mathbb{R}^3$?

$v = (x, y, z) \quad \exists \alpha_1, \alpha_2, \alpha_3 \,/\, \alpha_1\begin{pmatrix}1\\1\\1\end{pmatrix} + \alpha_2\begin{pmatrix}2\\1\\0\end{pmatrix} + \alpha_3\begin{pmatrix}4\\3\\2\end{pmatrix} = \begin{pmatrix}x\\y\\z\end{pmatrix}$

$\Leftrightarrow \begin{cases} \alpha_1 + 2\alpha_2 + 4\alpha_3 = x \\ \alpha_1 + \alpha_2 + 3\alpha_3 = y \\ \alpha_1 + 2\alpha_3 = z \to \alpha_1 = z - 2\alpha_3 \end{cases} \Leftrightarrow \begin{cases} z - 2\alpha_3 + 2\alpha_2 + 4\alpha_3 = x \\ z + \alpha_2 + \alpha_3 = y \quad \to \alpha_2 = y - z - \alpha_3 \\ \dots \end{cases}$

$\Leftrightarrow z - 2\alpha_3 + 2y - 2z - 2\alpha_3 + 4\alpha_3 = x \Leftrightarrow -z + 2y = x \to$ no funciona $\forall (x, y, z) \in \mathbb{R}^3$

$\Downarrow$

NO se puede escribir todo $v \in \mathbb{R}^3$ como CL de G

! $G' = \left\{ \begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}2\\1\\0\end{pmatrix} \right\}$

$\Rightarrow \langle G \rangle = \langle G' \rangle$

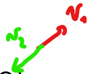
Dentro de los conjuntos generadores, nos interesan aquellos que son mínimos (menor cantidad de elementos).

Sea $S \subseteq \mathcal{V}$ un subespacio vectorial, y sea:

- $\{v_1, ..., v_n\} \subseteq \mathbb{V}$. Entonces $\langle v_1, ..., v_n \rangle \subseteq S$ sii $v_i \in S, \ \forall 1 \leq i \leq n$.
- $\{v_1, ..., v_n, v_{n+1}\} \subseteq \mathbb{V}$. Entonces $\langle v_1, ..., v_n, v_{n+1} \rangle = \langle v_1, ..., v_n \rangle$ sii $v_{n+1} \in \langle v_1, ..., v_n, \rangle$

Definición: Sea $\mathcal{V}$ un espacio vectorial, y sea $\{v_\alpha\}_{\alpha \in I}$ una familia de vectores en $\mathbb{V}$; se dice que $\{v_\alpha\}_{\alpha \in I}$ es linealmente independiente (l.i.) sii

$$\sum_{\alpha \in I} k_\alpha \bullet v_\alpha = \vec{0} \rightarrow k_\alpha = 0, \ \forall \alpha \in I$$



$0 \cdot v_1 + 0 \cdot v_2 + 0 \cdot v_3 = \vec{0}$

Observar:

- $\{0\}$ es linealmente dependiente (l.d.)
- si $v \neq 0$, $\{v\}$ es l.i.
- si $v_1 \propto v_2$ (colineales), $\{v_1, v_2\}$ es l.d.
- si v_1, v_2 no nulos, ni proporcionales, $\{v_1, v_2\}$ es l.i.

$\left\{ \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} \right\}$ L.I.

Bases y dimensión

Definición: Sea $\mathcal{V}$ un espacio vectorial, un conjunto $\{v_\alpha\}_{\alpha\in I}$ se llama base de $\mathcal{V}$ si $\{v_\alpha\}_{\alpha\in I}$ es un conjunto linealmente independiente de $\mathbb{V}$ que satisface $\langle v_\alpha \rangle_{\alpha\in I} = \mathbb{V}$.

tengo lo necesario + no sobra nada

Definición: Sean $\mathcal{V}$ un espacio vectorial, $B = \{v_1, ..., v_n\}$ una base de $\mathcal{V}$. Diremos que n es la dimensión de $\mathcal{V}$, donde $n < \infty$.

Comentario: Tener en cuenta que existen espacios vectoriales con dimensión infinita. $\rightarrow \mathcal{P}$

Variedad lineal

Sea $\mathcal{V}$ un espacio vectorial, M es una variedad lineal $M \subseteq \mathbb{V}$ es un conjunto de la forma $M = \{s + v,\ \textit{donde}\ s \in S\}$, siendo S subespacio de $\mathcal{V}$, y $v \in \mathbb{V}$.