Análisis Matemático para Inteligencia Artificial

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

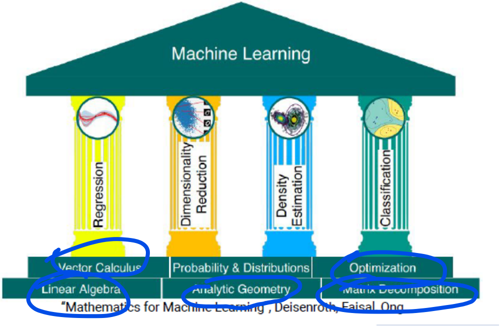
Especialización en Inteligencia Artificial

28/4/2023

¿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

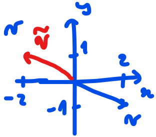
Empecemos considerando los vectores $u = \underbrace{(1, \frac{1}{2}, \frac{1}{3})}_{\mathbb{R}^3}$ y $v = \underbrace{(2, -1)}_{\mathbb{R}^2}$, ¿podemos sumar los vectores?

$u + v' = (1, \frac{1}{2}, \frac{1}{3}) + (2, -1, 0) =$

$u + v'' = (1, \frac{1}{2}, \frac{1}{3}) + (0, 2, -1) =$

? ✗

$u' = \underbrace{(1, \frac{1}{2})}_{\mathbb{R}^2}$ $\exists\ u' + v$



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

$\tilde{v} = \underbrace{(-1)}_{\mathbb{R}} \cdot \underbrace{(2, -1)}_{\mathbb{R}^2} = (-1) \cdot v$

escalar

puedo sumar $\tilde{v}$ $\iff$ puedo sumar v

prod. x escalar en ambos, en $\mathbb{R}^2$ $\iff$

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

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

$\mathbb{V} \quad \mathbb{V} \quad \mathbb{V}$

$+(x,y) = z$

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

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****.*

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

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

1. $p(x) + q(x) = (p + q)(x) = (1+2) + (\frac{1}{2}-1)x + (\frac{3}{4}+0)x^2 = 3 + (-\frac{1}{2})x + \frac{3}{4}x^2$

   $(3, -\frac{1}{2}, \frac{3}{4}) \in \mathcal{P}_2[x]$

2. $k \cdot p(x) = (kp)(x) = k + \frac{k}{2}x + \frac{3}{4}k\, x^2$ $\quad k(1, \frac{1}{2}, \frac{3}{4})$ $\in \mathcal{P}_2[x]$

Hay una biyección e/ $\mathbb{R}^3$ y $\mathcal{P}_2$

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

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

Prod. e/elementos NO es cerrado

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

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

$A \cdot B = (\longrightarrow)(\downarrow) \cdots \in \mathbb{R}^{n\times n}$

es cerrado ! ✓

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

Definición de **Espacio Vectorial**

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:

G_1

1. $+$ es asociativa $\forall a,b,c \in \mathbb{V} \quad a+(b+c) = (a+b)+c$
2. $+$ tiene elemento neutro $\exists e \in \mathbb{V} \,/\, \forall v \in \mathbb{V} \quad v+e = e+v = v$
3. $+$ tiene elemento inverso $\forall v \in \mathbb{V} \quad \exists \tilde{v} \in \mathbb{V} \,/\, \tilde{v}+v = v+\tilde{v} = e$
4. G_2 $+$ es conmutativa $\forall a,b \in \mathbb{V} \quad a+b = b+a$
5. $\alpha \bullet (v + w) = \alpha \bullet v + \alpha \bullet w, \;\; \forall \alpha \in \mathbb{K}, \forall v, w \in \mathbb{V}$ dist. del prod. x esc. wrt a la suma
6. $(\alpha \underbrace{+}_{\mathbb{K}} \beta) \bullet v = \alpha \bullet v \underbrace{+}_{\mathbb{V}} \beta \bullet v, \;\; \forall \alpha, \beta \in \mathbb{K}, \forall v \in \mathbb{V}$ dist. de suma de escalar wrt a el prod. x esc.
7. $\bullet$ tiene elemento neutro: $1 \bullet v = v, \forall v \in \mathbb{V}$
8. $\bullet$ es asociativa:

$$\alpha \bullet \overbrace{(\beta \bullet v)}^{\mathbb{V}} = (\alpha\beta) \bullet v, \;\; \forall \alpha, \beta \in \mathbb{K}, \forall v \in \mathbb{V}$$

G_1 grupo

G_2 grupo conmutativo

Subespacios Vectoriales: definición

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. $v, w \in S \to v + w \in S$
3. $\alpha \in \mathbb{K}, v \in S \to \alpha \bullet v \in S$

Subespacios triviales

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

$$\left.\begin{array}{l} 0 \in S \ \checkmark \\ 0 + 0 = 0 \in S \ \checkmark \\ \alpha \cdot 0 = 0 \in S \ \checkmark \end{array}\right\} \{0\} \text{ es subesp. de } \mathbb{V}$$

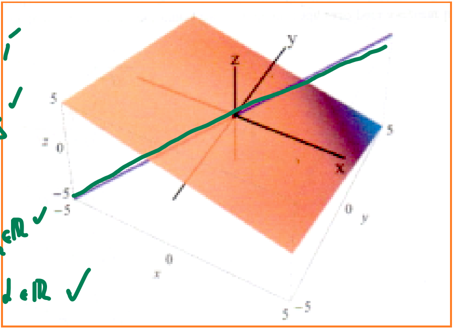
(a,b,c) fijos

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 $\alpha = 0$, $v = 0\cdot(a,b,c) = (0,0,0) \in S$

2) $(\alpha_1 a, \alpha_1 b, \alpha_1 c) + (\alpha_2 a, \alpha_2 b, \alpha_2 c) = ((\alpha_1+\alpha_2)a, (\alpha_1+\alpha_2)b, (\alpha_1+\alpha_2)c) = (\alpha_1+\alpha_2)(a,b,c)$ $\forall \alpha_1, \alpha_2 \in \mathbb{R}$ ✓

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



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. $S \cup T = \{v \in \mathbb{V} : v \in S \vee v \in T\} \subseteq \mathbb{V}$.

Demostremos 2 y 3

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

1) ¿$\vec{0} \in S+T$? $\vec{0} \in S, T$ por ser subesp. $\Rightarrow \underbrace{\vec{0}}_{s} + \underbrace{\vec{0}}_{t} = \underbrace{\vec{0}}_{v} \in S+T$

2) ¿$v, w \in S+T \Rightarrow v+w \in S+T$? si $v, w \in S+T \Rightarrow \exists\, s_v, s_w \in S$, $t_v, t_w \in T$ /

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

$\Rightarrow \underbrace{v+w}_{v} = s_v + t_v + s_w + t_w = \underbrace{(s_v + s_w)}_{s} + \underbrace{(t_v + t_w)}_{t} \Rightarrow$ ✓

si S y T son subesp.
$s_v + s_w \in S$
$t_v + t_w \in T$

3) ¿$k \cdot v \in S+T \ \forall v$? $v = s + t \Rightarrow \overbrace{k \cdot v}^{v} = \underbrace{k \cdot s}_{s} + \underbrace{k \cdot t}_{t}$ ✓

pues si S, T son subesp.
$\Rightarrow$ $k \cdot s \in S$
$k \cdot t \in T$

por 1, 2 y 3 $S+T$ también es subesp.

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 $i = 1, \ldots, r$.

$3 \cdot \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + \pi \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + (-4) \cdot \begin{pmatrix} 3 \\ 2 \\ 7 \end{pmatrix} + 0 \cdot \begin{pmatrix} 10 \\ 10 \\ 10 \end{pmatrix}$

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

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

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?

$\exists\, a, b, c \;/\; \forall v = (v_1, v_2, v_3) \in \mathbb{R}^3$

$$a\begin{pmatrix}1\\1\\1\end{pmatrix} + b\begin{pmatrix}2\\1\\0\end{pmatrix} + c\begin{pmatrix}4\\3\\2\end{pmatrix} = \begin{pmatrix}v_1\\v_2\\v_3\end{pmatrix}$$

Tiene sol.

$$\begin{cases} 1\cdot a + 2\cdot b + 4\cdot c = v_1 & (1)\\ 1\cdot a + 1\cdot b + 3\cdot c = v_2 & (2)\\ 1\cdot a + 0\cdot b + 2\cdot c = v_3 & (3) \end{cases}$$

$\rightarrow a = v_3 - 2c$

en (2) $\rightsquigarrow v_3 - 2c + b + 3c = v_2$

$b = v_2 - v_3 - c$

en (1) $\rightsquigarrow v_3 - \cancel{2c} + 2v_2 - 2v_3 - \cancel{2c} + \cancel{4c} = v_1$

$2v_2 - v_3 = v_1$!! $\rightarrow$ no tiene solución $\forall v_1, v_2, v_3$

$\Downarrow$

NO genera $\mathbb{R}^3$

Independencia Lineal

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

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

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.

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 $\langle v_\alpha\rangle_{\alpha\in I} = \mathbb{V}$.

no falta nada   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.

$\hookrightarrow \mathcal{P}$

$E = \{(1,0),(0,1)\}$ es base canónica de $\mathbb{R}^2$

$B_3 = \{(1,1),(0,1)\}$

$B_1 = \{(2,0),(0,1)\}$

$B_2 = \{(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}),(\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}})\}$

Variedad lineal

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