Análisis Matemático para Inteligencia Artificial

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

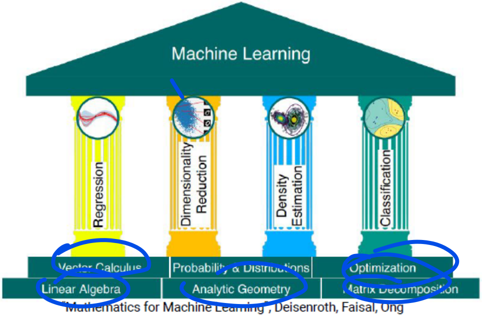
Especialización en Inteligencia Artificial

23/6/2023

¿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?

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

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

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

¿qué ocurre si tomamos $\tilde{v} = \underbrace{(-2, 1)}_{\mathbb{R}^2}$?

$\tilde{v} = (-1) \cdot (2, -1) = (-1)\, v$

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

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?

$(1, \frac{1}{2}, \frac{3}{4})$ $\in \mathcal{P}_2(x)$ $(2,-1,0)$

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

1. $p(x) + q(x) = (p + q)(x) = (3) + \left(-\frac{1}{2}\right)x + \left(\frac{3}{4}\right)x^2 = (1+2) + \left(\frac{1}{2}-1\right)x + \left(\frac{3}{4}+0\right)x^2 \in \mathcal{P}_2$

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

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

$\Updownarrow$ biyección con $\mathbb{R}^3$

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

$(1+\frac{1}{2}x+\frac{3}{4}x^2)(2-x) = 2 + x + \frac{3}{2}x^2 - x - \frac{1}{2}x^2 - \frac{3}{4}x^3$!! $\notin \mathcal{P}_2$

↳ NO es cerrado

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

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

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

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 → $\exists e \in \mathbb{V} \,/\, \forall a \in \mathbb{V}$ $a+e = e+a = a$
3. $+$ tiene elemento inverso → $\forall v \in \mathbb{V}$ $\exists \tilde{v} \in \mathbb{V} \,/\, v+\tilde{v} = \tilde{v}+v = e$
4. $+$ es conmutativa → $\forall a,b \in \mathbb{V}$ $a+b = b+a$
5. $\alpha \bullet (v + w) = \alpha \bullet v \overset{\mathbb{V}}{+} \alpha \bullet w, \ \forall \alpha \in \mathbb{K}, \forall v, w \in \mathbb{V}$ distrib. del · wrt. la + de $\mathbb{V}$
6. $(\alpha \underset{\mathbb{K}}{+} \beta) \bullet v = \alpha \bullet v + \beta \bullet v, \ \forall \alpha, \beta \in \mathbb{K}, \forall v \in \mathbb{V}$ distrib del · wrt la + del $\mathbb{K}$
7. $\bullet$ tiene elemento neutro: $1 \bullet v = v, \forall v \in \mathbb{V}$
8. $\bullet$ es asociativa:

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

⊛ grupo

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

Subespacios triviales

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

$0 \in S$? ✓

$0+0=0 \in S$ ✓

$k \cdot 0 = 0 \in S$ ✓

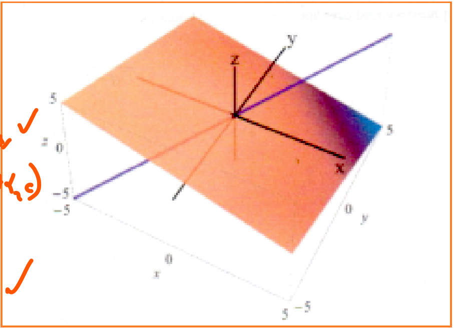
$v=(a,b,c)$ fijo

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 $\alpha=0$, $0\cdot(a,b,c)=(0,0,0)=0_V$ ✓

2) $\alpha_1\cdot v+\alpha_2 v=(\alpha_1 a+\alpha_2 a, \alpha_1 b+\alpha_2 b, \alpha_1 c+\alpha_2 c)$
$=\underbrace{(\alpha_1+\alpha_2)}_{\mathbb{R}}(a,b,c)$ ✓

3) $k\cdot\alpha\cdot(a,b,c)= (k\cdot\alpha)(a,b,c)$ ✓



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 el caso 2

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

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

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

commut. + asoc.

$$v+w = s_v + t_v + s_w + t_w = \underbrace{(s_v + s_w)}_{\in S} + \underbrace{(t_v + t_w)}_{\in T} = s' + t' \in S+T$$

3) ¿$v \in S+T, k \in \mathbb{K} \Rightarrow k \cdot v \in S+T$? $k \cdot v = k \cdot (s + t) = \underbrace{k \cdot s}_{\in S} + \underbrace{k \cdot t}_{\in T} \in S+T$

Por 1), 2) y 3) $S+T$ es subesp. de $\mathbb{V}$

Sistemas generadores

Definición: Sea $\mathcal{V}$ un espacio vectorial, y $G = \{v_1, \dots, v_r\} \subseteq \mathbb{V}$. 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 $i = 1, \dots, r$.

$$3\begin{pmatrix}0\\0\\-1\end{pmatrix} + \pi\begin{pmatrix}e^2\\e^{-1}\\e^0\end{pmatrix} + (-5)\begin{pmatrix}0\\0\\0\end{pmatrix}$$

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: $\langle G \rangle = \mathbb{V}$.

$$\mathbb{R}^2 = \left\langle \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix} \right\rangle = \left\langle \begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\-1\end{pmatrix}, \begin{pmatrix}2\\0\end{pmatrix}, \begin{pmatrix}3\\4\end{pmatrix} \right\rangle$$

Ejemplo

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

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?

$\exists k_1, k_2 \,/\, k_1\begin{pmatrix}1\\1\\1\end{pmatrix} + k_2\begin{pmatrix}2\\1\\0\end{pmatrix} = \begin{pmatrix}4\\3\\2\end{pmatrix}$

$\Downarrow$

$0 = k_1\begin{pmatrix}1\\1\\1\end{pmatrix} + k_2\begin{pmatrix}2\\1\\0\end{pmatrix} + (-1)\begin{pmatrix}4\\3\\2\end{pmatrix}$

$\Downarrow$

$(k_1, k_2, -1)$

$\exists a, b, c \in \mathbb{R} \,/\, a\begin{pmatrix}1\\1\\1\end{pmatrix} + b\begin{pmatrix}2\\1\\0\end{pmatrix} + c\begin{pmatrix}4\\3\\2\end{pmatrix} = \begin{pmatrix}x\\y\\z\end{pmatrix} \quad \forall x, y, z \in \mathbb{R}^3$

$$\begin{array}{l} ① \\ ② \\ ③ \end{array}\begin{cases} a + 2b + 4c = x \\ a + b + 3c = y \\ a + 2c = z \end{cases}$$

de ③ $\rightarrow a = z - 2c$

en ② $z - 2c + b + 3c = y$

$b = y - z - c$

en ① $z - \cancel{2c} + 2y - 2z \cancel{-2c} + \cancel{4c} = x$

$2y - z = x$ $\rightarrow$ sólo puedo resolver/genera los $\begin{pmatrix}2y-z\\y\\z\end{pmatrix}$

$\hookrightarrow$ 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.

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.

$E = \left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\}$ "base canónica de $\mathbb{R}^2$"

$B = \left\{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\-1\end{pmatrix}\right\}$

$B'' = \left\{\begin{pmatrix}2\\1\end{pmatrix}, \begin{pmatrix}3\\1\end{pmatrix}\right\}$

$B' = \left\{\begin{pmatrix}1/\sqrt{2}\\1/\sqrt{2}\end{pmatrix}, \begin{pmatrix}1/\sqrt{2}\\-1/\sqrt{2}\end{pmatrix}\right\}$

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 $v \in \mathbb{V}$.