Análisis Matemático para Inteligencia Artificial

Martín Errázquin (merrazquin@fi.uba.ar)

Especialización en Inteligencia Artificial

Clase 1

Administrativo

TODO el material excepto las grabaciones de las clases está disponible en el aula virtual del campus!

$V_i \to$ long.
$V_i \to$ último día 100%.
V_i $I \cdot 0.9$
V_i $I \cdot 0.9^2 = 0.81$

Forma de evaluación:

1. Entrega 1 (individual): ejercicio de la guía 1
2. Entrega 2 (individual): ejercicio de la guía 2
3. Trabajo Final (grupos de 3 integrantes)
    - Componente de programación
    - Componente de interpretación de código
    - Componente de matemática

La nota final es un promedio ponderado de las notas individuales de cada entrega. Las dos entregas individuales tienen un límite de entrega *soft*.

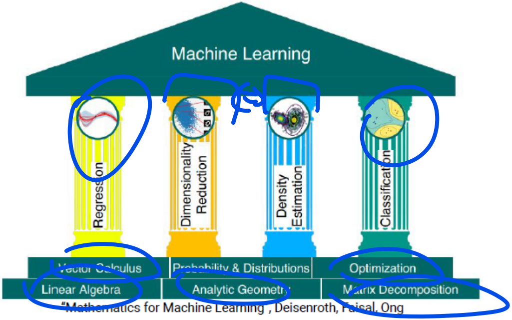
Los trabajos ~~se entregan en la carpeta~~ de drive individual compartida a cada alumno ~~posterior a la primera~~ clase.

↳ campus

¿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): performance

Operaciones vectorizadas y GPUs: Early colab!

Motivación (II): Más tipos de error

- Error de programación: error de sintaxis → compiler
- Error de programación: el código está mal pensado

Aparecen nuevos!

- Error de matemática: el modelo está mal planteado → debug lápiz y papel
- Error de planteo: no están dadas las condiciones para la aplicación del modelo seleccionado
- Error de traducción: el código no refleja el modelo planteado
  ↳ debug comparando

Motivación (III): entender los modelos

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

$$\left.\begin{array}{l} u + v' = (1+2, \frac{1}{2} - 1, \frac{1}{3} + 0) \\ u + v'' = (1+0, \frac{1}{2} + 2, \frac{1}{3} - 1) \end{array}\right\} \times$$

$\nexists\, u + v$

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

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

$\therefore$ 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 $+ : \mathbb{V} \times \mathbb{V} \to \mathbb{V}$.

$(x, y) \to x^y$

$2^1 \neq 1^2$

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?

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

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

1. $p(x) + q(x) = (p+q)(x) = (1+2) + (\frac{1}{2} - 1)x + (\frac{3}{4} + 0)x^2 \in \mathcal{P}_2$
2. $k \cdot p(x) = (kp)(x) = (k \cdot 1) + (k \cdot \frac{1}{2})x + (k \cdot \frac{3}{4})x^2 \in \mathcal{P}_2 \rightarrow \checkmark$

$\mathcal{P}_2$ Hay una biyección con $\mathbb{R}^3$

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

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

!!

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

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

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

$k \in \mathbb{R}$

$\underset{n \times n \quad n \times n}{A \cdot B} \in \mathbb{R}^{n \times n} \rightarrow$ en las mat. cuadradas el prod. también es cerrado!

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:

1. $+$ es asociativa $\forall v, w, u \in \mathbb{V} \quad (v+w)+u = v+(w+u)$
2. $+$ tiene elemento neutro $\exists\, e \in \mathbb{V} \,/\, \forall x \in \mathbb{V} \quad x+e = e+x = x$
3. $+$ tiene elemento inverso $\forall x \in \mathbb{V} \,/\, \exists\, \tilde{x} \in \mathbb{V} \,/\, x+\tilde{x} = \tilde{x}+x = e$
4. $+$ es conmutativa $\forall v, w \in \mathbb{V} \quad v+w = w+v$
5. $\alpha \bullet (v + w) = \alpha \bullet v + \alpha \bullet w, \;\; \forall \alpha \in \mathbb{K}, \forall v, w \in \mathbb{V}$ distrib. de • sobre + de $\mathbb{V}$
6. $(\alpha \underbrace{+}_{\mathbb{K}} \beta) \bullet v = \alpha \bullet v \underbrace{+}_{\mathbb{V}} \beta \bullet v, \;\; \forall \alpha, \beta \in \mathbb{K}, \forall v \in \mathbb{V}$ distrib. de • sobre + de $\mathbb{K}$
7. $\bullet$ tiene elemento neutro: $1 \bullet v = v, \forall v \in \mathbb{V}$
8. $\bullet$ es asociativa:

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

$\mathbb{K} \times \mathbb{V}$ $\qquad$ $\mathbb{K} \times \mathbb{K}$

⊛ grupo

⊛ 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.

$(S, +, \mathbb{K}, \cdot)$ es E.V.

Condiciones necesarias y suficientes para caracterizar subespacios

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

1. $S \neq \emptyset$ $(0^{e} \in S)$ → si $e \in S$ ✓
2. $v, w \in S \rightarrow v + w \in S$ → + cerrada en S
3. $\alpha \in \mathbb{K}, v \in S \rightarrow \alpha \bullet v \in S$ → • cerrada en S

Subespacios triviales

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

$0 \in S$ ✓

$0 + 0 = 0 \in S$ ✓

$\forall k \in \mathbb{K},\ k \cdot 0 = 0 \in S$ ✓

Ejemplos de subespacios propios

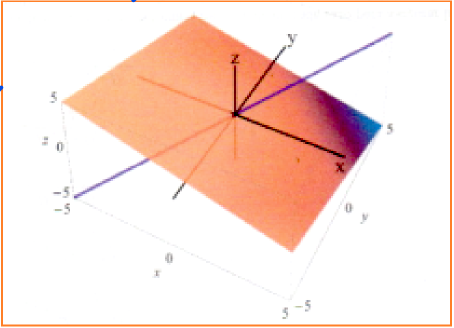
$v = (a, b, c) \in \mathbb{R}^3$ fijo

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

1) si $\alpha = 0$, $\alpha \cdot (a,b,c) = (0,0,0) \in S$ ✓

2) sean $\alpha_1, \alpha_2 \in \mathbb{R}$ $\alpha_1(a,b,c) + \alpha_2(a,b,c) =$
$(\alpha_1 a + \alpha_2 a, \alpha_1 b + \alpha_2 b, \alpha_1 c + \alpha_2 c) =$
$\underbrace{(\alpha_1 + \alpha_2)}_{\alpha_3}(a,b,c) \in S$ ✓

3) $k \cdot \alpha \cdot (a,b,c) = \underbrace{(k \cdot \alpha)}_{\alpha_2}(a,b,c) \in S$ ✓



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

$S = \{0, 1, 2\}$
$T = \{0, 3\}$ en $V = \mathbb{R}$

$S \cap T = \{0\}$
$S \cup T = \{0, 1, 2, 3\}$
$S + T = \{0, 3, 1, 2, 4, 5\}$

Demostremos el caso 2

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

1) ¿ $e\in S+T$? como S y T son subesp., por cond ④ $e\in S$, $e\in T$

$$e = \underset{s}{e} + \underset{t}{e} \in S+T \checkmark$$

$\boxed{S+T \text{ es subesp. de } \mathbb{V}}$

2) ¿ $\forall v_1,v_2\in S+T$ $v_1+v_2\in S+T$?

$v_1 = s_1 + t_1$ con $s_1, s_2\in S$

$v_2 = s_2 + t_2$ $\quad t_1, t_2\in T$

$v_1+v_2 = s_1+s_2+t_1+t_2$ donde $s_1+s_2\in S$ x cond ②, $t_1+t_2\in T$

$= s_3 + t_3 \in S+T \checkmark$

3) ¿ $\forall v\in S+T$ $k\cdot v\in S+T$?

$v = s+t$, x cond ⑤ de E.V.

$k\cdot v = k\cdot(s+t) = k\cdot s + k\cdot t$

$= s_1 + t_1 \in S+T \checkmark$

x cond ③

Sistemas generadores

$\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + \dots + \alpha_r v_r = v$

$G = \{ v_1, v_2, v_3, \dots, v_r \}$

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 $i = 1, \dots, r$.

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

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

$\langle G \rangle = \{$ todas las posibles CL de $G \}$

"G genera $\mathbb{V}$" $\iff \langle G \rangle = \mathbb{V}$

Ejemplo

$\cdot 2 + \cdot 1 =$

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?

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

$$\begin{cases} a+2b+4c=x & (1) \\ a+b+3c=y & (2) \\ a+0b+2c=z & (3) \end{cases}$$

de (3) $a = z-2c$ en (2) $z-2c+b+3c=y$

$b = y-c-z$

en (1)

$z-2c+2y-2c-2z+4c=x$

$2y-z=x$

Tiene sol. $\iff$ $x = 2y-z$

$\Downarrow$

NO Tiene sol $\forall x,y,z$

$\Rightarrow \langle G \rangle \neq \mathbb{R}^3$

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

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

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

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.

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

$\forall k \in \mathbb{R}$

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

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 sobra nada

no falta 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

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

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

$B'' = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ base

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 $v \in \mathbb{V}$.