Análisis Matemático para Inteligencia Artificial

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

Especialización en Inteligencia Artificial

Clase 1

TODO el material excepto las grabaciones de las clases están disponibles en nuestro repositorio de github.

entrega 1:
27/10 → lanz.
3/11 → vencim
4/11-10/11 → 90%
11/10 - 17/11 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

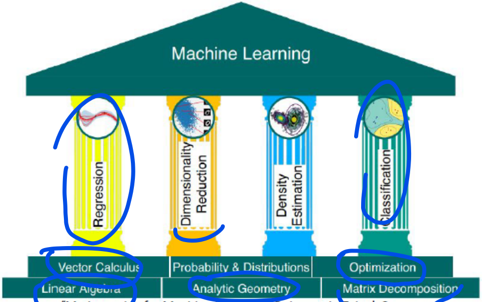
La nota final es un promedio ponderado de las notas individuales de cada entrega. Las dos entregas individuales tienen un límite de entrega *soft*.

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

¿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.



"Mathematics for Machine Learning", Deisenroth, Faisal, Ong

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 $u = \overbrace{(1, \frac{1}{2}, \frac{1}{3})}^{\mathbb{R}^3}$ y $v = \overbrace{(2, -1)}^{\mathbb{R}^2}$, ¿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 $\tilde{v} = (-2, 1)$? $\longrightarrow$ $\tilde{v} = (-1)\cdot v \in \mathbb{R}^2$ $\quad \nexists\, u+\tilde{v}$

$\therefore$ 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**.*

$(1, \frac{1}{2}, \frac{3}{4})$ $\in \mathcal{P}_2$ $(2, -1, 0)$ $\in \mathcal{P}_1$

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

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

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

$(1 + \frac{1}{2}x + \frac{3}{4}x^2)(2 - x) = 2 - x + x - \frac{1}{2}x^2 + \frac{3}{2}x^2 - \frac{3}{4}x^3 \to$ NO es cerrado en $\mathcal{P}_2$

$\notin \mathcal{P}_2$

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

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

Ⓖ 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.

$(S, +, \mathbb{K}, \cdot)$

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

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

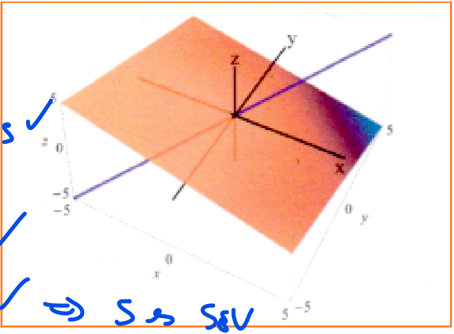
$0_v \in S$ ✓

$0 + 0 = 0 \in S$ ✓

$\alpha \cdot 0 = 0 \quad \forall \alpha \in \mathbb{K} \quad \in S$ ✓

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



1) $\alpha = 0 \Rightarrow \alpha.(a,b,c) = (0,0,0) = 0_V \in S$ ✓

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

3) $k \cdot \alpha.(a,b,c) = (k\alpha)(a,b,c) \in S$ ✓ $\Rightarrow$ S es SEV

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) $0_{\mathbb{V}} \in S+T$? como S, T son SEV ✓ $\Rightarrow 0_{\mathbb{V}} \in S, T$

$\Rightarrow 0_{\mathbb{V}} = \underbrace{0_{\mathbb{V}}}_{S} + \underbrace{0_{\mathbb{V}}}_{T} \Rightarrow 0_{\mathbb{V}} \in S+T$ ✓

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

$v_1 + v_2 \in S+T$? si $v_1, v_2 \in S+T \Rightarrow \exists s_1, s_2 \in S, t_1, t_2 \in T \,/\, \begin{array}{l} v_1 = s_1 + t_1 \\ v_2 = s_2 + t_2 \end{array}$

$$\begin{array}{rl} v_1 &= s_1 + t_1 \\ + \; v_2 &= s_2 + t_2 \\ \hline v_1 + v_2 &= \underbrace{(s_1 + s_2)}_{S} + \underbrace{(t_1 + t_2)}_{T} \in S+T \checkmark \end{array}$$

son SEV

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

si $v \in S+T \Rightarrow \exists s \in S, t \in T \,/\, v = s+t$

$k \cdot v = k \cdot (s+t) = k \cdot s + k \cdot t = \underbrace{k \cdot s}_{S} + \underbrace{k \cdot t}_{T} \in S+T$ ✓

como S y T son SEV

$k \cdot s \in S$

$k \cdot t \in T$

$\langle G \rangle = \{ v \in V / \exists \alpha_1, \dots, \alpha_r \in \mathbb{K}$
$v = \alpha_1 \cdot v_1 + \dots + \alpha_r v_r \}$

Sistemas generadores

$G = \{v_1, .., v_r\}$

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 $i = 1, \dots, r$.

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

$x = \begin{pmatrix} 1 \\ 7 \\ 3 \end{pmatrix}$

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

$\exists a, b, c \in \mathbb{R}$

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

Ejemplo

$(a+c)\begin{pmatrix}1\\1\\1\end{pmatrix}+(b+c)\begin{pmatrix}2\\1\\0\end{pmatrix}=a\cdot\begin{pmatrix}1\\1\\1\end{pmatrix}+b\cdot\begin{pmatrix}2\\1\\0\end{pmatrix}$

$2\cdot \quad + 1\cdot \quad =$

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?

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

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

de (3) $a = z-2c$

en (2)

$y = z-2c+b+3c$

$= z+b+c$

↳ $b = y-z-c$

en (1)

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

$\boxed{= 2y-z}$

no existe sol. válida $\forall x,y,z$

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

G 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 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.

Pero $\mathbb{R}^2$:

$E = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ base canónica

$B_2 = \left\{ \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}, \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix} \right\}$ BON

$B_1 = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$ BOG

$B_3 = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$ base

E, B_1, B_2, B_3 son LI + $\langle E\rangle = \langle B_1\rangle = \langle B_2\rangle = \langle B_3\rangle = \mathbb{R}^2$

Variedad lineal

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