Instituto Tecnológico de Costa Rica

Programa de Capacitación Profesional en Ciencias de los Datos

Curso: Matemática para Ciencias de los Datos

Informe de Trabajo Práctico 0

Realizado por:

Felipe Alberto Mejías Loría, 201231682

María Mora,

Profesor:

Saúl Calderón Ramírez

Fecha: San José, Mayo 15, 2019

1. Sistemas lineales (20 puntos)

1. Demuestre si los siguientes sistemas $L\left\{u(t)\right\}$ (con entrada $u(t)$ y salida $g(t) = L\left\{u(t)\right\}$, y $h(t)$ una función cualquiera) son lineales o no lineales:

   a) $g\left(t\right) = u(t) + 2$

   Se tiene que cumplir que:

   $$L\left\{\alpha_1 u_1(t) + \alpha_2 u_2(t)\right\} = \alpha_1 L\left\{u_1(t)\right\} + \alpha_2 L\left\{u_2(t)\right\}$$

   Desarrollando la propiedad anterior, se tiene que:

   $$\alpha_1 u_1(t) + \alpha_2 u_2(t) + 2 \neq \alpha_1(u_1(t) + 2) + \alpha_2(u_2(t) + 2)$$
   $$\alpha_1 u_1(t) + \alpha_2 u_2(t) + 2 \neq \alpha_1 u_1(t) + \alpha_1 2 + \alpha_2 u_2(t) + \alpha_2 2$$

   por lo tanto el sistema no es lineal.

   b) $g\left(t\right) = u(t)h(t)$

   Se tiene que cumplir que:

   $$L\left\{\alpha_1 u_1(t) + \alpha_2 u_2(t)\right\} = \alpha_1 L\left\{u_1(t)\right\} + \alpha_2 L\left\{u_2(t)\right\}$$

   Desarrollando la propiedad anterior, se tiene que:

   $$(\alpha_1 u_1(t) + \alpha_2 u_2(t))h(t) = \alpha_1(u_1(t)h(t)) + \alpha_2(u_2(t)h(t))$$
   $$\alpha_1 u_1(t)h(t) + \alpha_2 u_2(t)h(t) = \alpha_1 u_1(t)h(t) + \alpha_2 u_2(t)h(t)$$

   por lo tanto el sistema es lineal.

   c) $g\left(t\right) = \text{máx}(u(t))$ Se tiene que cumplir que:

   $$L\left\{\alpha_1 u_1(t) + \alpha_2 u_2(t)\right\} = \alpha_1 L\left\{u_1(t)\right\} + \alpha_2 L\left\{u_2(t)\right\}$$

   Desarrollando la propiedad anterior, se tiene que:

   $$máx(\alpha u_1(t) + \alpha_2 u_2(t)) = \alpha_1(máx(u_1(t)) + \alpha_2(máx(u_2(t))$$
   $$máx(\alpha u_1(t) + \alpha_2 u_2(t)) \neq \alpha_1 máx(u_1(t)) + \alpha_2 máx(u_2(t))$$

   por lo tanto el sistema no es lineal.

d) $g\left(t\right) = u'(t)$

Se tiene que cumplir que:

$$L\left\{\alpha_1 u_1(t) + \alpha_2 u_2(t)\right\} = \alpha_1 L\left\{u_1(t)\right\} + \alpha_2 L\left\{u_2(t)\right\}$$

Desarrollando la propiedad anterior, se tiene que:

$$u'(\alpha_1 u_1(t) + \alpha_2 u_2(t)) = \alpha_1 u'(u_1(t)) + \alpha_2 u'(u_2(t))$$

$$\alpha_1 u'(u_1(t)) + \alpha_2 u'(u_2(t)) = \alpha_1 u'(u_1(t)) + \alpha_2 u'(u_2(t))$$

por lo tanto el sistema es lineal.

e) $g\left(t\right) = |u(t)|$ Se tiene que cumplir que:

$$L\left\{\alpha_1 u_1(t) + \alpha_2 u_2(t)\right\} = \alpha_1 L\left\{u_1(t)\right\} + \alpha_2 L\left\{u_2(t)\right\}$$

Desarrollando la propiedad anterior, se tiene que:

$$|\alpha_1 u_1(t) + \alpha_2 u_2(t)| = \alpha_1 |u_1(t)| + \alpha_2 |u_2(t)|$$

$$|\alpha_1||u_1(t)| + |\alpha_2||u_2(t)| \neq \alpha_1 |u_1(t)| + \alpha_2 |u_2(t)|$$

por lo tanto el sistema no es lineal.

2. Demuestre si los siguientes sistemas con múltiples variables de entrada $d\left(\vec{u}, \vec{v}\right)$ (con entrada vectorial $\vec{u}, \vec{v} \in \mathbb{R}^n$ y salida escalar $s \in \mathbb{R}$), cumplen la condición de homogeneidad absoluta $d\left(\alpha_1 \vec{u}, \alpha_2 \vec{v}\right) = |\alpha_1|\,|\alpha_2|\, d\left(\vec{u}, \vec{v}\right)$ de superposición $d\left(\vec{u} + \vec{x}, \vec{v} + \vec{y}\right) = d\left(\vec{u}, \vec{v}\right) + d\left(\vec{x}, \vec{y}\right)$.

   a) La Norma de Manhattan l_1

$$\|\overrightarrow{x}\|_1 = \sum_i^n |x_i|$$

Se tiene que cumplir que:

$$L\left\{\alpha \overrightarrow{x}\right\} = |\alpha| L\left\{\overrightarrow{x}\right\}$$

Desarrollando la propiedad anterior, se tiene que:

$$|\alpha| \sum_i^n |x_i| = |\alpha| \sum_i^n |x_i|$$

por lo tanto cumple con la homogeneidad absoluta. También se tiene que cumplir que:

$$L\{\overrightarrow{x}+\overrightarrow{y}\}=L\{\overrightarrow{x}\}+L\{\overrightarrow{y}\}$$

Desarrollando la propiedad anterior, se tiene que:

$$\sum_i^n |x_i+y_i| \neq \sum_i^n |x_i| + \sum_i^n |y|$$

por lo tanto no cumple la superposición y el sistema no es lineal.

b) Norma Euclidiana l_2

$$\|\overrightarrow{x}\|_2 = \sqrt{\sum_i^n |x_i|^2}$$

Se tiene que cumplir que:

$$L\{\alpha\overrightarrow{x}\}=|\alpha|L\{\overrightarrow{x}\}$$

Desarrollando la propiedad anterior, se tiene que:

$$|\alpha|\sqrt{\sum_i^n |x_i|^2} = |\alpha|\sqrt{\sum_i^n |x_i|^2}$$

por lo tanto cumple con la homogeneidad absoluta. También se tiene que cumplir que:

$$L\{\overrightarrow{x}+\overrightarrow{y}\}=L\{\overrightarrow{x}\}+L\{\overrightarrow{y}\}$$

Desarrollando la propiedad anterior, se tiene que:

$$\sqrt{\sum_i^n |x_i+y_i|^2} \neq \sqrt{\sum_i^n |x_i|^2} + \sqrt{\sum_i^n |y|^2}$$

por lo tanto no cumple la superposición y el sistema no es lineal.

c) Norma L∞ Se tiene que cumplir que:

$$L\{\alpha\overrightarrow{x}\}=|\alpha|L\{\overrightarrow{x}\}$$

Desarrollando la propiedad anterior, se tiene que:

$$\max\left(|\alpha x_i|\right) = |\alpha| \max\left(|x_i|\right)$$

por lo tanto cumple con la homogeneidad absoluta. También se tiene que cumplir que:

$$L\left\{\overrightarrow{u} + \overrightarrow{v}\right\} = L\left\{\overrightarrow{u}\right\} + L\left\{\overrightarrow{v}\right\}$$

Desarrollando la propiedad anterior, se tiene que:

$$\max\left(|u_i + v_i|\right) \neq \max\left(|v_i|\right) + \max\left(|u_i|\right)$$

por lo tanto no cumple la superposición y el sistema no es lineal.

d) Para una función multivariable $g(x_1, x_2, ..., x_n) = g(\overrightarrow{x})$ con dominio en $\mathbb{R}^n \to \mathbb{R}$:
$g(\overrightarrow{x}) = \overrightarrow{w} \cdot \overrightarrow{x} + b = \overrightarrow{w}^T \overrightarrow{x} + b = w_1 x_1 + w_2 x_2 + ... + w_n x_n + b$
con $\overrightarrow{w} \in \mathbb{R}^n$ y $b \in \mathbb{R}$ coeficientes conocidos, demuestre si es un sistema lineal respecto al arreglo de entradas $\overrightarrow{x}$. Se tiene que cumplir que:

$$L\left\{\alpha \overrightarrow{x}\right\} = |\alpha| L\left\{\overrightarrow{x}\right\}$$

Desarrollando la propiedad anterior, y con $\overrightarrow{u} = \alpha \overrightarrow{x}$ se tiene que:

$$\overrightarrow{w} \cdot \alpha \overrightarrow{x} + b = \alpha(\overrightarrow{w} \cdot \overrightarrow{x} + b)$$

$$\alpha \overrightarrow{w}^T \overrightarrow{x} + b = \alpha(\overrightarrow{w}^T \overrightarrow{x} + b)$$

$$\alpha(w_1 x_1 + w_2 x_2 + ... + w_n x_n) + b \neq \alpha(w_1 x_1 + w_2 x_2 + ... + w_n x_n + b)$$

por lo tanto no cumple con la homogeneidad absoluta. También se tiene que cumplir que:

$$L\left\{\overrightarrow{u} + \overrightarrow{v}\right\} = L\left\{\overrightarrow{u}\right\} + L\left\{\overrightarrow{v}\right\}$$

Desarrollando la propiedad anterior, se tiene que:

$$\overrightarrow{w} \cdot (\vec{u} + \vec{v}) + b \neq (\overrightarrow{w} \cdot (\vec{u}) + b) + (\overrightarrow{w} \cdot (\vec{v}) + b)$$

$$\overrightarrow{w}^T (\vec{u} + \vec{v}) + b \neq (\overrightarrow{w}^T \cdot (\vec{u}) + b) + (\overrightarrow{w}^T \cdot (\vec{v}) + b)$$

$$w_1(u_1 + v_1) + w_2(u_2 + v_2) + ... + w_n(u_n + v_n) + b \neq w_1(u_1 + v_1) + w_2(u_2 + v_2) + ... + w_n(u_n + v_n) + 2b$$

por lo tanto no cumple la superposición y el sistema no es lineal.

d.1.) Para que el sistema sea lineal se debe agregar un elemento más al vector w, el elemento debe tener un valor de uno. Adicionalmente, los arreglos de entrada deben contener un elemento más con valor b, con esto se elimina b de la ecuación original.

2. Vectores (20 puntos)

1. Graficación y propiedades de los vectores

   a) Usando Python grafique los siguientes vectores $\vec{v}_1 = \begin{bmatrix} -0{,}3 \\ 0{,}8 \\ 0{,}1 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} 0{,}5 \\ 0{,}2 \\ 0{,}4 \end{bmatrix}$ y $\vec{v}_3 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \\ 0 \end{bmatrix}$, usando la función quiver.

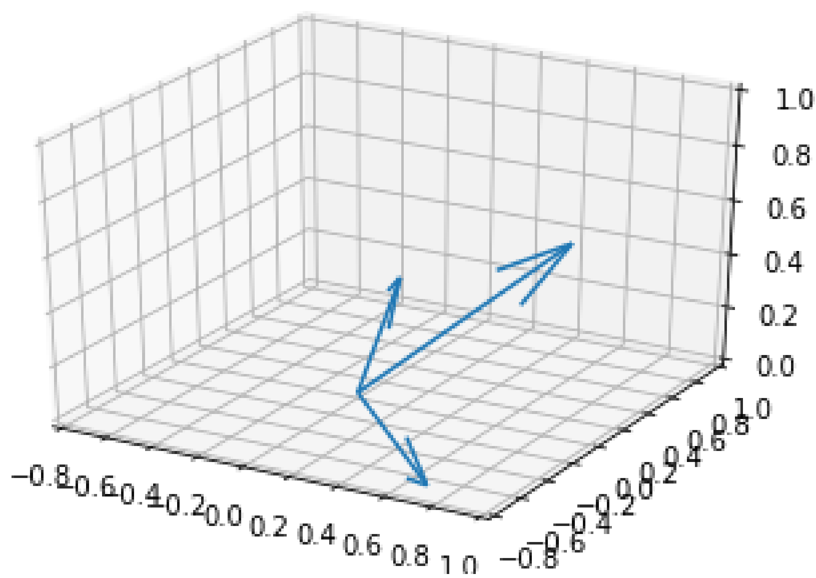


Figura 1: Gráfica de los vectores $\vec{v}_1, \vec{v}_2, \vec{v}_3$

   b) Escriba los vectores anteriores en términos de los vectores unitarios $\hat{i}, \hat{j}, \hat{k}$

   b.1)Hay que calcular el vector unitario $\hat{i}$=$\frac{\vec{v_1}}{\|\vec{v_1}\|}$

   Se calcula la $\|\vec{v_1}\|$

$$\sqrt{\sum_{i=1}^{2}|v_i|^2} = \sqrt{(-0.3)^2+(0.8)^2+(0.1)^2} = \sqrt{0.09+0.64+0.01} = \sqrt{0.74} \sim 0.86$$

Se calcula $\frac{\vec{v_1}}{\|\vec{v_1}\|}$

$$\hat{i} = \begin{bmatrix} \frac{-0{,}3}{,086} \\ \frac{0{,}8}{0{,}86} \\ \frac{0{,}1}{0{,}86} \end{bmatrix} = \begin{bmatrix} -0{,}35 \\ 0{,}93 \\ 0{,}12 \end{bmatrix}$$

Entonces

$$\vec{v}_1 = 0{,}86 \begin{bmatrix} -0{,}35 \\ 0{,}93 \\ 0{,}12 \end{bmatrix}$$

b.2) Hay que calcular el vector unitario $\hat{j}$=$\frac{\vec{v_2}}{\|\vec{v_2}\|}$

Se calcula la $\|\vec{v_2}\|$

$$\sqrt{\sum_{i=1}^{2}|v_i|^2} = \sqrt{(0.5)^2+(0.2)^2+(0.4)^2} = \sqrt{0.25+0.04+0.16} = \sqrt{0.45} \sim 0.67$$

Se calcula $\frac{\vec{v_2}}{\|\vec{v_2}\|}$

$$\hat{j} = \begin{bmatrix} \frac{0{,}5}{0{,}67} \\ \frac{0{,}2}{0{,}67} \\ \frac{0{,}4}{0{,}67} \end{bmatrix} = \begin{bmatrix} 0{,}75 \\ 0{,}30 \\ 0{,}60 \end{bmatrix}$$

Entonces

$$\vec{v}_2 = 0{,}67 \begin{bmatrix} 0{,}75 \\ 0{,}30 \\ 0{,}60 \end{bmatrix}$$

b.3) Hay que calcular el vector unitario $\hat{k}$=$\frac{\vec{v_3}}{\|\vec{v_3}\|}$

Se calcula la $\|\overrightarrow{v_3}\|$

$$\sqrt{\sum_{i=1}^{2}|v_i|^2} = \sqrt{\left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{-1}{\sqrt{2}}\right)^2 + (0)^2} = \sqrt{0.5+0.5+0} = \sqrt{1} = 1$$

$$\hat{k} = \vec{v}_3 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \\ 0 \end{bmatrix}$$

2. Demuestre cuales de los vectores anteriores son unitarios.

   a) Para el vector $\vec{v}_1 = \begin{bmatrix} -0{,}3 \\ 0{,}8 \\ 0{,}1 \end{bmatrix}$:

   Hay que demostrar que:

$$\|\vec{v}_1\| = 1$$

$$\sqrt{\sum_{i=1}^{2}|v_i|^2} = \sqrt{(-0.3)^2 + (0.8)^2 + (0.1)^2} = \sqrt{0.09+0.64+0.01} = \sqrt{0.74} \sim 0.86$$

   $\therefore$ No es unitario.

   1) Para el vector $\vec{v}_2 = \begin{bmatrix} 0{,}5 \\ 0{,}2 \\ 0{,}4 \end{bmatrix}$:

   Hay que demostrar que:

$$\|\vec{v}_2\| = 1$$

$$\sqrt{\sum_{i=1}^{2}|v_i|^2} = \sqrt{(0.5)^2 + (0.2)^2 + (0.4)^2} = \sqrt{0.25+0.04+0.16} = \sqrt{0.45} \sim 0.67$$

   $\therefore$ No es unitario.

2) Para el vector $\vec{v}_3 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \\ 0 \end{bmatrix}$:

Hay que demostrar que:

$$\|\vec{v}_3\| = 1$$

$$\sqrt{\sum_{i=1}^{2} |v_i|^2} = \sqrt{\left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{-1}{\sqrt{2}}\right)^2 + (0)^2} = \sqrt{0.5 + 0.5 + 0} = \sqrt{1} = 1$$

$\therefore$ Sí es unitario.

b) Calcule el ángulo en grados, entre los vectores $\vec{v}_1$y $\vec{v}_2$, $\vec{v}_2$ y $\vec{v}_3$ y $\vec{v}_1$ y $\vec{v}_3$, implementando la fórmula en **Pytorch**.

Salida de Jupyter:

a) Angulo en grados entre v1=tensor([-0.3000, 0.8000, 0.1000]) y v2=tensor([0.5000, 0.2000, 0.4000]): tensor(85.0293)

b) Angulo en grados entre v2=tensor([0.5000, 0.2000, 0.4000]) y v3=tensor([0.7071, -0.7071, 0.0000]): tensor(71.5650)

c) Angulo en grados entre v1=tensor([-0.3000, 0.8000, 0.1000]) y v3=tensor([0.7071, -0.7071, 0.0000]): tensor(154.7150)

c) Calcule la distancia en ℓ_1, ℓ_2, y ℓ_∞ entre los vectores $\vec{v}_1$y $\vec{v}_2$, $\vec{v}_2$ y $\vec{v}_3$ y $\vec{v}_1$ y $\vec{v}_3$, implementando la fórmula en **Pytorch**.

Salida de Jupyter:

1) Distancia entre los vectores $\vec{v}_1$y $\vec{v}_2$

Distancia Euclideana entre v1=tensor([-0.3000, 0.8000, 0.1000]) y v2=tensor([0.5000, 0.2000, 0.4000]): tensor(1.0440)
Distancia Euclideana usando torch.norm entre v1=tensor([-0.3000, 0.8000, 0.1000]) y v2=tensor([0.5000, 0.2000, 0.4000]): tensor(1.0440)

Distancia Manhattan entre v1=tensor([-0.3000, 0.8000, 0.1000]) y v2=tensor([0.5000, 0.2000, 0.4000]): tensor(1.7000)
Distancia Manhattan usando torch.norm entre v1=tensor([-0.3000, 0.8000, 0.1000]) y v2=tensor([0.5000, 0.2000, 0.4000]): tensor(1.7000)

Distancia Infinita entre v1=tensor([-0.3000, 0.8000, 0.1000]) y v2=tensor([0.5000, 0.2000, 0.4000]): tensor(0.8000)

2) Distancia entre los vectores $\vec{v}_2$y $\vec{v}_3$

Distancia Euclideana entre v2=tensor([0.5000, 0.2000, 0.4000]) y v3=tensor([0.7071, -0.7071, 0.0000]): tensor(1.0128)

Distancia Euclideana usando torch.norm entre v2=tensor([0.5000, 0.2000, 0.4000]) y v3=tensor([0.7071, -0.7071, 0.0000]): tensor(1.0128)

Distancia Manhattan entre v2=tensor([0.5000, 0.2000, 0.4000]) y v3=tensor([0.7071, -0.7071, 0.0000]): tensor(1.5142)
Distancia Manhattan usando torch.norm entre v2=tensor([0.5000, 0.2000, 0.4000]) y v3=tensor([0.7071, -0.7071, 0.0000]): tensor(1.5142)

Distancia Infinita entre v2=tensor([0.5000, 0.2000, 0.4000]) y v3=tensor([0.7071, -0.7071, 0.0000]): tensor(0.9071)

3) Distancia entre los vectores $\vec{v}_1$y $\vec{v}_3$
Distancia Euclideana entre v1=tensor([-0.3000, 0.8000, 0.1000]) y v3=tensor([0.7071, -0.7071, 0.0000]): tensor(1.8154)
Distancia Euclideana usando torch.norm entre v1=tensor([-0.3000, 0.8000, 0.1000]) y v3=tensor([0.7071, -0.7071, 0.0000]): tensor(1.8154)

Distancia Manhattan entre v1=tensor([-0.3000, 0.8000, 0.1000]) y v3=tensor([0.7071, -0.7071, 0.0000]): tensor(2.6142)
Distancia Manhattan usando torch.norm entre v1=tensor([-0.3000, 0.8000, 0.1000]) y v3=tensor([0.7071, -0.7071, 0.0000]): tensor(2.6142)

Distancia Infinita entre v1=tensor([-0.3000, 0.8000, 0.1000]) y v3=tensor([0.7071, -0.7071, 0.0000]): tensor(1.5071)

3. Propiedades del producto punto: demuestre lo siguiente. Además, muestrelo con una implementación en Pytorch, usando como entrada un arreglo de 50 arreglos generados al azar, adjunte un pantallazo con la salida de la comparación del resultado a ambos lados de la igualdad, o en su defecto, demuestre el no cumplimiento de la propiedad con un contraejemplo.

   a) Conmutatividad del producto punto $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$

   Para el vector $\vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ . \\ . \\ u_n \end{bmatrix}$ y el vector $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ . \\ . \\ v_n \end{bmatrix}$:

$$\overrightarrow{u}^T \cdot \overrightarrow{v} = \overrightarrow{v}^T \cdot \overrightarrow{u}$$

$$\begin{bmatrix} u_1 & u_2 & . & . & u_n \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \\ . \\ . \\ v_n \end{bmatrix} = \begin{bmatrix} v_1 & v_2 & . & . & v_n \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \\ . \\ . \\ u_n \end{bmatrix}$$

$$\sum_{i=1}^{n} u_i \cdot v_i = \sum_{i=1}^{n} v_i \cdot u_i$$

Se logra observar como $u_i \cdot v_i = v_i \cdot u_i$

b) No asociatividad del producto punto $\vec{u} \cdot (\vec{v} \cdot \vec{w}) \neq (\vec{u} \cdot \vec{v}) \cdot \vec{w}$
Tomando $\vec{u} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \vec{v} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}, \vec{w} = \begin{bmatrix} 5 \\ 6 \\ 7 \end{bmatrix}$:
Obteniendo el producto punto de $\vec{v} \cdot \vec{w}$ y $\vec{u} \cdot \vec{v}$:

$$\vec{v} \cdot \vec{w} = \sum_{i=1}^{n} v_i \cdot w_i = (2 \cdot 5) + (3 \cdot 6) + (4 \cdot 7) = 56$$

$$\vec{u} \cdot \vec{v} = \sum_{i=1}^{n} u_i \cdot v_i = (1 \cdot 2) + (2 \cdot 3) + (3 \cdot 4) = 20$$

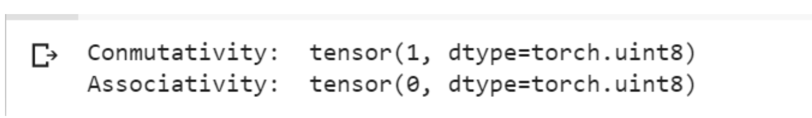
Calculando $\vec{u} \cdot (\vec{v} \cdot \vec{w})$ y $(\vec{u} \cdot \vec{v}) \cdot \vec{w}$:

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \cdot 56 = 20 \cdot \begin{bmatrix} 5 \\ 6 \\ 7 \end{bmatrix}$$

$$\begin{bmatrix} 56 \\ 112 \\ 168 \end{bmatrix} \neq \begin{bmatrix} 100 \\ 120 \\ 140 \end{bmatrix}$$

$\therefore$ Se observa como ambos vectores son diferentes y se demuestra la no asociatividad del producto punto.

La salida de la demostración en Python es:



```
Conmutativity:  tensor(1, dtype=torch.uint8)
Associativity:  tensor(0, dtype=torch.uint8)
```

Figura 2: Salida de la demostración de conmutatividad y asociatividad en Python

3. La distancia de Minkowski como índice de error

Explique la diferencia entre las métricas, analizándolo desde el punto de vista de la sensibilidad a los valores atípicos. La Tabla 1 muestra 3 particiones del set de datos, las cuales le ayudarán a responder entonces la pregunta ¿cuál métrica es más sensible a los valores atípicos? Use como apoyo el cálculo en Pytorch del MAE y el RMSE de los datos pendientes (?) en la tabla 1. Realice el cálculo de forma matricial, prescindiendo al máximo de estructuras de repetición.

i	t_i	$\widetilde{t_i}$	$\lvert\widetilde{t_i} - t_i\rvert$	$(\widetilde{t_i} - t_i)^2$
1	4	2	2	4
2	6	4	2	4
3	5	3	2	4
4	6	4	2	4
5	8	6	2	4
6	10	8	2	4
7	7	5	2	4
8	4	2	2	4
9	2	4	2	4
10	8	10	2	4

Cuadro 1: Primer conjunto de muestras

Resultados de Python para el Primer Set de Datos:

$$MAE = 2.0$$

$$RMSE = 2.0$$

$$\sigma_{MAE} = 0$$

$$\sigma_{RMSE} = 0$$

i	t_i	$\widetilde{t_i}$	$\lvert\widetilde{t_i} - t_i\rvert$	$(\widetilde{t_i} - t_i)^2$
11	5	4	1	1
12	3	2	1	1
13	2	3	1	1
14	4	5	1	1
15	20	21	1	1
16	32	29	3	9
17	5	2	3	9
18	4	7	3	9
19	7	4	3	9
20	41	38	3	9

Cuadro 2: Segundo conjunto de muestras

Resultados de Python para el Segundo Set de Datos:

$$MAE = 2.0$$

$$RMSE = 2.2361$$

$$\sigma_{MAE} = 1.0541$$

$$\sigma_{RMSE} = 4.2164$$

i	t_i	$\widetilde{t_i}$	$\lvert\widetilde{t_i} - t_i\rvert$	$(\widetilde{t_i} - t_i)^2$
21	6	6	0	0
22	20	20	0	0
23	31	31	0	0
24	41	41	0	0
25	50	50	0	0
26	62	62	0	0
27	73	73	0	0
28	4	4	0	0
29	7	7	0	0
30	40	20	20	400

Cuadro 3: Tercer conjunto de muestras

Resultados de Python para el Tercer Set de Datos:

$$MAE = 2.0$$

$$RMSE = 6.3246$$

$$\sigma_{MAE} = 6.3246$$

$$\sigma_{RMSE} = 126.4911$$

Al calcular tanto el MAE como el RMSE en los tres conjuntos de datos se observa como el MAE es constante con un valor de 2.0 y el RMSE aumenta conforme aumenta la variabilidad de las magnitudes de error. Esto último se observa analizando la magnitud de los errores en los tres conjuntos de datos; por ejemplo, en el primer set de datos el error es únicamente de 2, mientras que en el segundo set los errores tienen un valor de 1 y 3, sin embargo, el tercer set de datos presenta una magnitud de error de 20, y esto es lo que hace que se dispare el RMSE. Por lo tanto, se observa que el RMSE no necesariamente aumenta con la varianza de los errores, sino que aumenta con la variabilidad de las magnitudes de error. Por último, el MAE es adecuado para describir errores distribuidos uniformemente, mientras que el RMSE es una mejor medida para modelos que presenten una distribución normal.

4. Funciones multivariable (15 puntos)

1. Un hiperplano f definido en un espacio $\mathbb{R}^{n+1}$ se puede expresar como una función con dominio $\vec{x} \in \mathbb{R}^n$ y codominio en $\mathbb{R}$ como sigue: $z = f(\vec{x}) = \vec{x} \cdot \vec{w}$, con $\vec{w} \in \mathbb{R}^n$ el arreglo de coeficientes de tal funcional.

   a) Tomese $\vec{w}_1 = \begin{bmatrix} 0{,}5 \\ 0{,}2 \end{bmatrix}$ para la función f_1 y $\vec{w}_1 = \begin{bmatrix} -0{,}1 \\ 0{,}05 \end{bmatrix}$ para la función f_2, (funciones con dominio en $\mathbb{R}^2$ y codominio en). Grafique ambos planos en Pytorch.

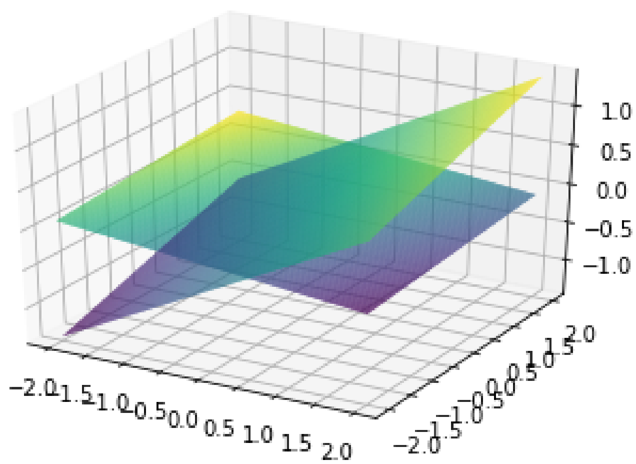


Figura 3: Gráfica de los planos f_1, f_2

   b) Exprese ambos planos en su forma vectorial $\vec{n} \cdot \overrightarrow{P_0 P_1} = 0$,con $\vec{n}$ el vector normal a tal hiperplano, P_0 y P puntos sobre el mismo. Un vector normal $\vec{n} \in \mathbb{R}^m$ a dos vectores $\vec{v} \in \mathbb{R}^m$ y $\vec{u} \in \mathbb{R}^m$ se obtiene en términos del producto cruz como $\vec{n} = \vec{v} \times \vec{u}$, el cual se define a continuación (con $A_{\setminus i}$ el resultado de eliminar la columna i de la matriz A):

$$\vec{n} = \det\left(A_{\setminus 1}\right)\hat{i} - \det\left(A_{\setminus 2}\right)\hat{j} + \det\left(A_{\setminus 3}\right)\hat{k}$$

   con $A = \begin{bmatrix} - & \vec{v} & - \\ - & \vec{u} & - \end{bmatrix}$

Para la función f_1:

$$\vec{n_1} \cdot \overrightarrow{P_0P_1} = 0,$$

Para obtener $\vec{n_1}$, evalúo la función $f_1$en (1,1), (2,1) y (2,2):

$$f_1(1,1) = 0{,}5(1) + 0{,}2(1) = 0{,}7$$

$$f_1(2,1) = 0{,}5(2) + 0{,}2(1) = 1{,}2$$

$$f_1(2,2) = 0{,}5(2) + 0{,}2(2) = 1{,}4$$

Obtengo los puntos:

$$P_0 = (1, 1, 0{,}7)$$

$$P_1 = (2, 1, 1{,}2)$$

$$P_2 = (2, 2, 1{,}4)$$

Calculo los vectores:

$$\overrightarrow{P_0P_1} = \begin{bmatrix} 2-1 \\ 1-1 \\ 1{,}2-0{,}7 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0{,}5 \end{bmatrix}$$

$$\overrightarrow{P_0P_2} = \begin{bmatrix} 2-1 \\ 2-1 \\ 1{,}4-0{,}7 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0{,}7 \end{bmatrix}$$

Sea la matriz $A = \begin{bmatrix} \overrightarrow{P_0P_1} \\ \overrightarrow{P_0P_2} \end{bmatrix}$:

$$\vec{n_1} = \det\left(A_{\backslash 1}\right)\hat{i} - \det\left(A_{\backslash 2}\right)\hat{j} + \det\left(A_{\backslash 3}\right)\hat{k}$$

$$A = \begin{bmatrix} 1 & 0 & 0{,}5 \\ 1 & 1 & 0{,}7 \end{bmatrix}$$

$$\vec{n_1} = \det\left(\begin{bmatrix} 0 & 0{,}5 \\ 1 & 0{,}7 \end{bmatrix}\right)\hat{i} - \det\left(\begin{bmatrix} 1 & 0{,}5 \\ 1 & 0{,}7 \end{bmatrix}\right)\hat{j} + \det\left(\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\right)\hat{k}$$

$$\vec{n_1} = -0{,}5\hat{i} - 0{,}2\hat{j} + \hat{k}$$

$$\vec{n_1} = \begin{bmatrix} -0{,}5 \\ -0{,}2 \\ 1 \end{bmatrix}$$

La ecuación de un plano que pasa por $P_0 = (x_0, y_0, z_0)$, normal a $\vec{n} = A\hat{i} + B\hat{j} + C\hat{k}$es:

$$\vec{n_1} \cdot \overrightarrow{P_0P} = 0$$

$$\left[\begin{array}{ccc} A & B & C \end{array}\right] \cdot \left[\begin{array}{c} x-x_0 \\ y-y_0 \\ z-z_0 \end{array}\right] = 0$$

$$A(x-x_0)+B(y-y_0)+B(z-z_0)=0$$

$$-0{,}5(x-x_0)+-0{,}2(y-y_0)+(z-z_0)=0$$

2. Para cada una de las siguientes funciones multivariable: grafique su superficie, calcule el vector gradiente manualmente, evaluelo y grafique el vector unitario en la dirección del gradiente para los dos puntos especificados (en la misma figura de la superficie) y calcule la magnitud de tal vector gradiente en cada punto. Además calcule la matriz Hessiana. En general, investigue ¿qué indica la magnitud del vector gradiente y la matriz Hessiana?

   a) $z = f(x,y) = x^3y^2+1$, evaluación del gradiente en los puntos $P_0 = (0,0)$ y $P_1 = (7{,}4, -6{,}3)$.
   Calculando el vector gradiente manualmente:

   $$\nabla f = \frac{\delta f}{\delta x}\hat{\imath} + \frac{\delta f}{\delta y}\hat{\jmath} = (3x^2y^2)\hat{\imath} + (2x^3y)\hat{\jmath}$$

   Evaluando en P_0:
   $$\nabla f_{(0,0)} = (0)\hat{\imath} + (0)\hat{\jmath}$$

   Evaluando en P_1:

   $$\nabla f_{(7{,}4,-6{,}3)} = (6520{,}27)\hat{\imath} + (-5105{,}82)\hat{\jmath}$$

   Calculando la magnitud de los gradientes evaluados:

   $$\left\|\nabla f_{(0,0)}\right\| = 0$$

   $$\left\|\nabla f_{(7{,}4,-6{,}3)}\right\| = \sqrt{(6520{,}27)^2 + (-5105{,}82)^2} = 8281{,}50$$

   Normalizando el resultado de $\nabla f_{(7{,}4,-6{,}3)}$:

   $$\left[\begin{array}{c} \frac{6520{,}27}{8281{,}50} \\ \frac{-5105{,}82}{8281{,}50} \end{array}\right] = \left[\begin{array}{c} 0{,}787329 \\ -0{,}616533 \end{array}\right]$$

   La matriz Hessiana sería:

   $$H_{f(x,y)} = \left[\begin{array}{cc} \frac{\delta f}{\delta x^2} & \frac{\delta f}{\delta x \delta y} \\ \frac{\delta f}{\delta y \delta x} & \frac{\delta f}{\delta y^2} \end{array}\right] = \left[\begin{array}{cc} 6xy^2 & 6x^2y \\ 6x^2y & 2x^3 \end{array}\right]$$

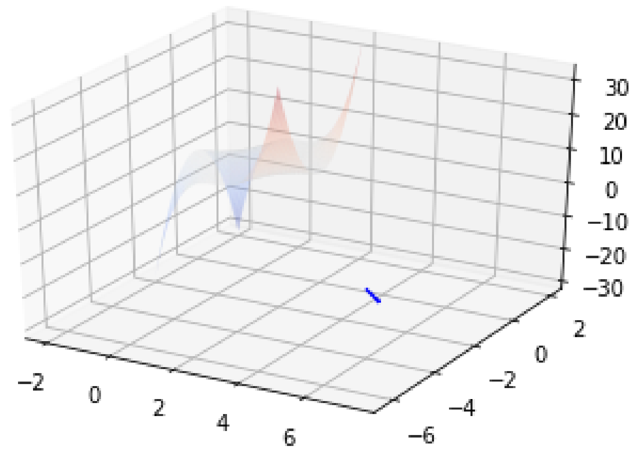


Figura 4: Gráfica de la primera función multivariable

b) $z = f(x, y) = e^x + e^y + 2x^2 + 4y + 3$, evaluación del gradiente en los puntos $P_0 = (3{,}8, 1{,}8)$ y $P_1 = (6{,}2, 7)$
Calculando el vector gradiente manualmente:

$$\nabla f = \frac{\delta f}{\delta x}\hat{\imath} + \frac{\delta f}{\delta y}\hat{\jmath} = (e^x + 4x)\hat{\imath} + (e^y + 4)\hat{\jmath}$$

Evaluando en P_0:

$$\nabla f_{(3{,}8,1{,}8)} = (59{,}90)\hat{\imath} + (10{,}05)\hat{\jmath}$$

Evaluando en P_1:

$$\nabla f_{(6{,}2,7)} = (517{,}55)\hat{\imath} + (1100{,}63)\hat{\jmath}$$

Calculando la magnitud de los gradientes evaluados:

$$\left\|\nabla f_{(3{,}8,1{,}8)}\right\| = \sqrt{(59{,}90)^2 + (10{,}05)^2} = 60{,}74$$

$$\left\|\nabla f_{(6{,}2,7)}\right\| = \sqrt{(517{,}55)^2 + (1100{,}63)^2} = 1216{,}24$$

Normalizando el resultado de $\nabla f_{(3{,}8,1{,}8)}$:

$$\begin{bmatrix} \frac{59{,}90}{60{,}74} \\ \frac{10{,}05}{60{,}74} \end{bmatrix} = \begin{bmatrix} 0{,}986170 \\ 0{,}165459 \end{bmatrix}$$

Normalizando el resultado de $\nabla f_{(6{,}2,7)}$:

$$\begin{bmatrix} \frac{517{,}55}{1216{,}24} \\ \frac{1100{,}63}{1216{,}24} \end{bmatrix} = \begin{bmatrix} 0{,}425532 \\ 0{,}904944 \end{bmatrix}$$

La matriz Hessiana sería:

$$H_{f(x,y)} = \begin{bmatrix} \frac{\delta f}{\delta x^2} & \frac{\delta f}{\delta x \delta y} \\ \frac{\delta f}{\delta y \delta x} & \frac{\delta f}{\delta y^2} \end{bmatrix} = \begin{bmatrix} e^x + 4 & 0 \\ 0 & e^y \end{bmatrix}$$

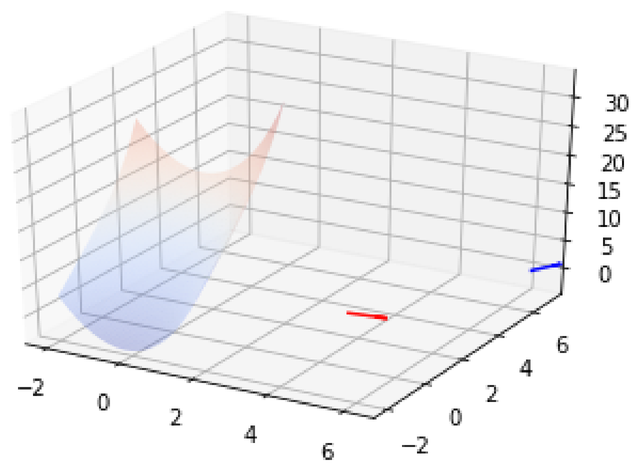


Figura 5: Gráfica de la segunda función multivariable

c) $z = f(x, y) = \ln\left(x^2 + y^2\right)$, evaluación del gradiente en los puntos $P_0 = (-2{,}4, 6{,}7)$ y $P_1 = (0, 4{,}3)$

Calculando el vector gradiente manualmente:

$$\nabla f = \frac{\delta f}{\delta x}\hat{i} + \frac{\delta f}{\delta y}\hat{j} = (\frac{2x}{x^2 + y^2})\hat{i} + (\frac{2y}{x^2 + y^2})\hat{j}$$

Evaluando en P_0:

$$\nabla f_{(-2{,}4,6{,}7)} = (-0{,}09)\hat{i} + (0{,}26)\hat{j}$$

Evaluando en P_1:

$$\nabla f_{(0,4{,}3)} = (0)\hat{i} + (0{,}46)\hat{j}$$

Calculando la magnitud de los gradientes evaluados:

$$\left\|\nabla f_{(-2{,}4,6{,}7)}\right\| = \sqrt{(-0{,}09)^2 + (0{,}26)^2} = 0{,}27$$

$$\left\|\nabla f_{(0,4{,}3)}\right\| = \sqrt{(0)^2 + (0{,}46)^2} = 0{,}46$$

Normalizando el resultado de $\nabla f_{(-2,4,6,7)}$:

$$\begin{bmatrix} \frac{-0{,}09}{0{,}27} \\ \frac{0{,}26}{0{,}27} \end{bmatrix} = \begin{bmatrix} -0{,}333333 \\ 0{,}962962 \end{bmatrix}$$

Normalizando el resultado de $\nabla f_{(0,4,3)}$:

$$\begin{bmatrix} 0 \\ \frac{0{,}46}{0{,}46} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

La matriz Hessiana sería:

$$H_{f(x,y)} = \begin{bmatrix} \frac{\delta f}{\delta x^2} & \frac{\delta f}{\delta x \delta y} \\ \frac{\delta f}{\delta y \delta x} & \frac{\delta f}{\delta y^2} \end{bmatrix} = \begin{bmatrix} \frac{2(x^2-y^2)}{(x^2+y^2)^2} & \frac{-4xy}{(x^2+y^2)^2} \\ \frac{-4xy}{(x^2+y^2)^2} & \frac{2(x^2-y^2)}{(x^2+y^2)^2} \end{bmatrix}$$

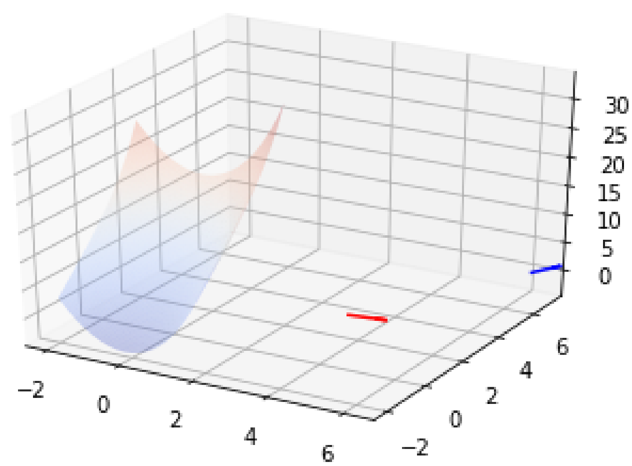


Figura 6: Gráfica de la tercera función multivariable

Es importante indicar que el vector gradiente denota la dirección para la cual una superficie definida por la función f cambia, y la magnitud del vector gradiente es la máxima razón de cambio de la función en el punto del gradiente. Además, la matriz Hessiana encuentra su uso principal en algunos algoritmos de aprendizaje automático, y se utiliza para encontrar el máximo y el mínimo de una curva.