Clase 4 (Parte 1)

06/10/2023

Resumen clase anterior:

- En la clase anterior vimos métodos y clases en Python.
- Revisamos la intuición detrás del método de Gradiente Descendiente.

• Presentaciones (Rúbrica en el computador)

Feature Scaling / Escalando Características

• El método de Gradiente descendiente se puede ver afectado por las dimensiones de las características.

• El método se alimenta de los datos para mapear la función de pérdida. Por ejemplo, estudiemos con un ejemplo como el tamaño de los θ puede afectar el resultado.

Ejemplo de Motivación :

Modelo : $f_\theta(x) = \theta_2 x_2 + \theta_1 x_1 + \theta_0$

Caso 1 : X_1 con un rango amplio $\longrightarrow \theta_1$ mejor pequeño

X_2 con un rango pequeño $\longrightarrow \theta_2$ mejor grande

- Supongamos que el nivel medio a 1 metro de distancia de una batería en dB es dado por el siguiente modelo :

$$l_1 = \theta_0 + \theta_1 \underbrace{X_1}_{\text{Volumen de la fuente}} + \theta_2 \underbrace{X_2}_{\text{Grosor de la membrana}}$$

- Asumamos que físicamente una batería oscila entre los 90 dB y 130 dB.

Ejemplo:

$$\underbrace{89\,dB}_{\theta_0} + \underbrace{0{,}01}_{\theta_1}\frac{dB}{cm^3}\cdot 500\,cm^3 + \underbrace{40\frac{dB}{mm}}_{\theta_2}\cdot 0{,}1\,mm$$

$$\parallel$$

$$5\,dB + 4\,dB + 89\,dB$$

$$\parallel$$

$$98\,dB \text{ a } 1 \text{ metro}$$

Caso contrario: aumentemos θ_1 levemente

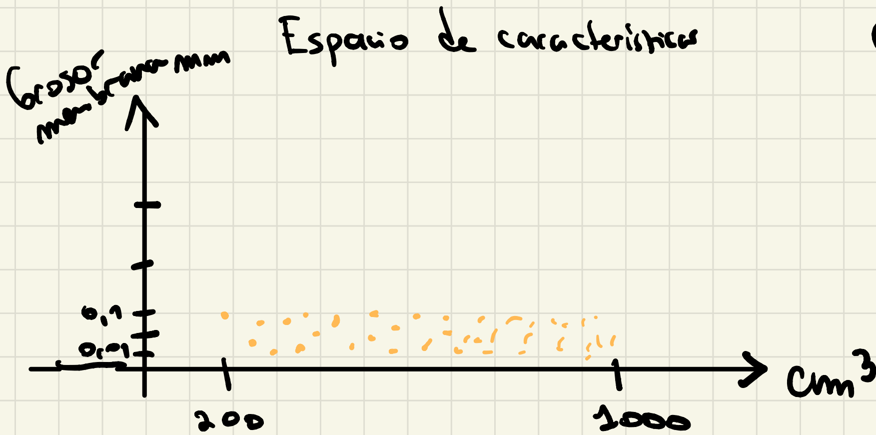
$$89\,dB + 0{,}1\frac{dB}{cm^3}\,500\,cm^3 + 40\frac{dB}{mm}\cdot 0{,}1\,mm$$

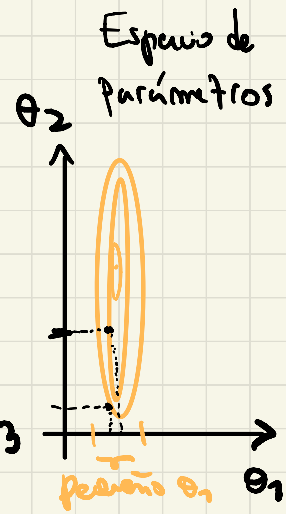
$$89\,dB + 50\,dB + 4\,dB = 143\,dB\ (> 130\,dB)$$

- Esta sería una mala aproximación: hemos superado los límites físicos.

- Un pequeño cambio en θ_1 alteró mucho el resultado!

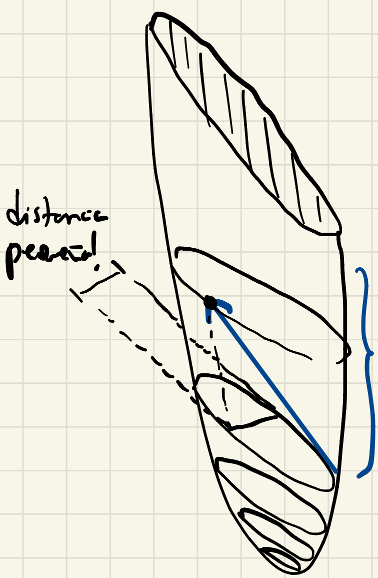
Visualmente :





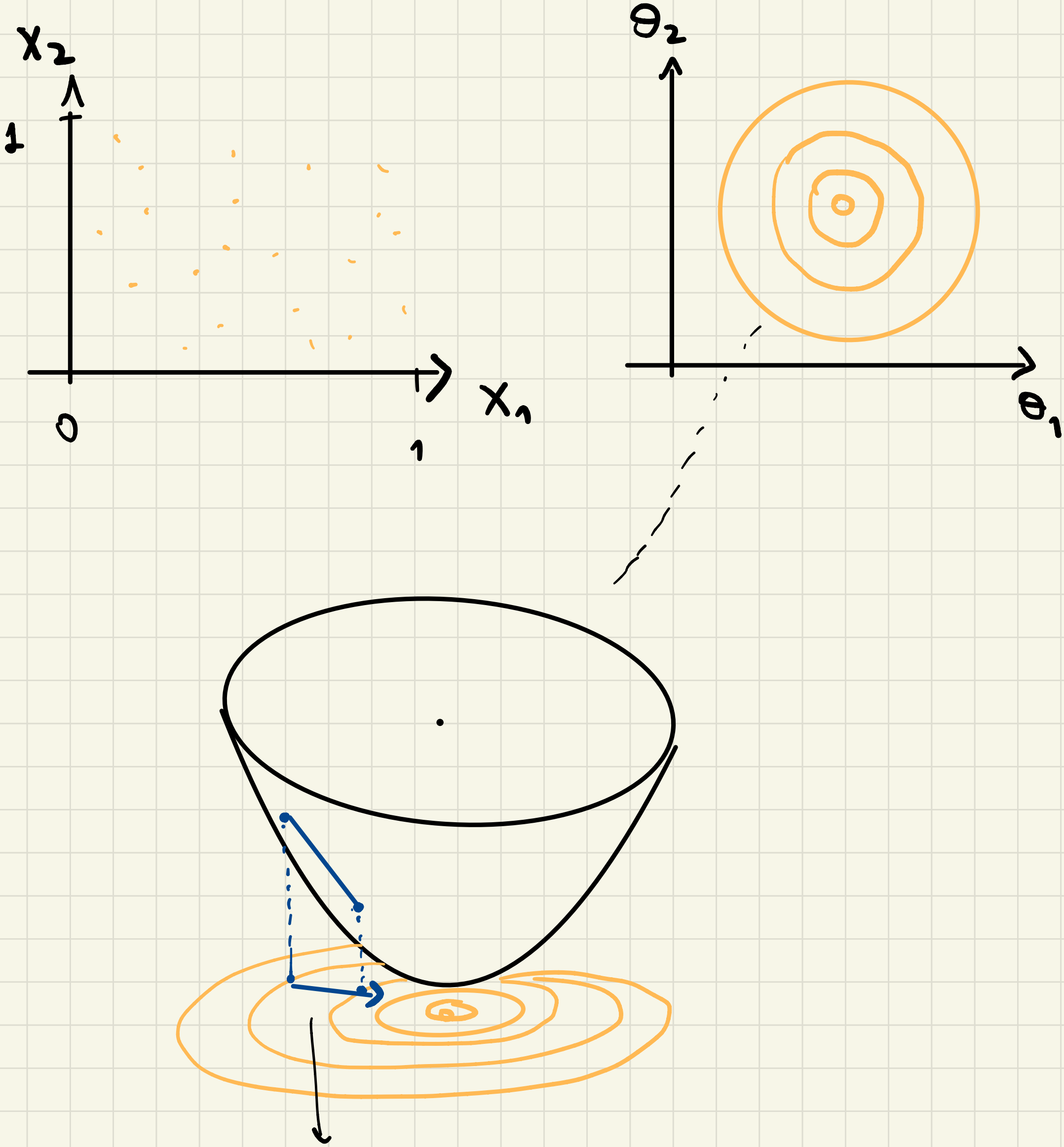
- Por ser $\mathcal{L}(\theta, (x,y))$ muy abrupta,

Como :



un pequeño paso se puede traducir en un gran cambio!

Si escalamos las características, podemos evitar este problema:



ahora el camino será más directo hacia el mínimo!

Escalando Datos

1) Absolute Maximum Scaling :

- no centra la data
- no destruye la esparsidad.

$$X_{scaled} = \frac{X_i}{\max\left(\left|\{X_i\}_{i=1}^{n}\right|\right)}$$

Ejemplo : $[-5, -2, 1, 0, 1, 5]$

$$\downarrow$$

$$-1 \quad -2/5 \quad 1/5 \quad 0 \quad 1/5 \quad 1$$

Rango : $[-1, 1]$

2) Min-Max Scaling

$$X_{scaled} = \frac{X_i - \min\left(\{X_i\}_{i=1}^{n}\right)}{\max\left(\{X_i\}_{i=1}^{n}\right) - \min\left(\{X_i\}_{i=1}^{n}\right)}$$

Ejemplo :

$[-5, -2, 1, 0, 1, 5]$

$$\downarrow$$

$$0 \quad 3/10 \quad 6/10 \quad 5/10 \quad 6/10 \quad 1$$

Rango : $[0, 1]$

3) Normalización:

$$X_{scaled} = \frac{X_i - \overline{X}}{\max(\{X_i\}) - \min(\{X_i\})}$$

4) Estandarización:

$$X_{scaled} = \frac{X_i - \overline{X}}{\sigma}$$

σ: desviación estandar

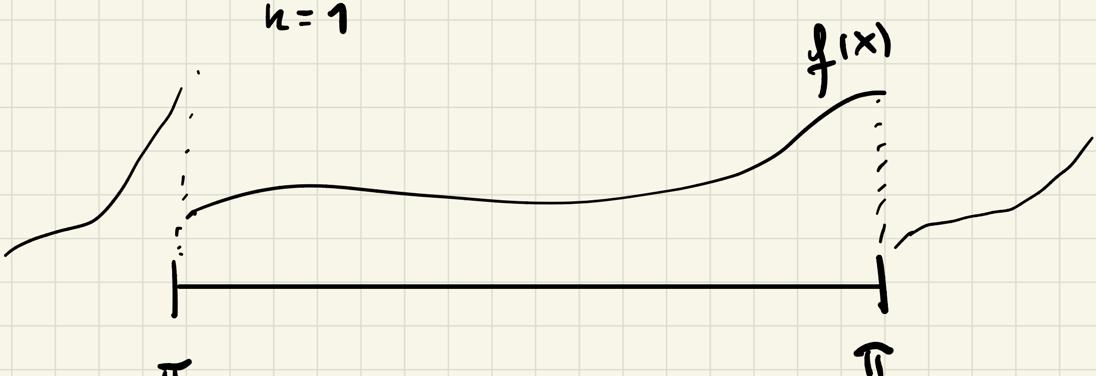
Features o caracteristicas en audio

1) Transformada de Fourier (Discreta) ¿Es una regresión?

- Queremos representar una señal $f(X)$ como una combinación lineal de coeficientes.
- Estamos sumamente interesados en los coeficientes!
- Es decir queremos realizar un "cambio de bases".

Si tenemos una función $f(x)$ estamos representado en el tiempo y deseamos entender su representación como suma de senos y cosenos! debemos saber algunas cosas. Asumamos que podemos representar $f(x)$ así:

$$f(x) = a_0 + \sum_{k=1}^{\infty} a_k \cos(kx) + b_k \operatorname{sen}(kx) \quad (1)$$



$$f(x) : \mathbb{R} \longrightarrow [-\pi, \pi]$$

* Necesitamos ∞ términos para representar f (Es un compromiso que asumimos)

* Cuando vemos $\underset{d/dx}{\Sigma}$ o $\int_{\partial/\partial x}$: Pensamos en matrices!

* Observemos el problema :
Es un problema de descomposición!

* $f(x)$ es una combinación lineal de
$a_0 \cdot 1$, senos y cosenos ! ☺

* De alguna manera tendremos un montón de $f(t_k)$ y queremos encontrar
a_0 , a_n y b_n para cada n .
Tal y como nuestro problema de regresión donde θ era nuestra incógnita.

* Asumamos que podemos discretizar el problema eliminando los ∞ (Asumimos un error de aproximación). [Convergencia fuera de la discusión].

$$f(x) \cong a_0 + \sum_{k=1}^{N} a_k \cos(kx) + \underbrace{\sum_{k=1}^{N} b_k \operatorname{sen}(kx)}_{\text{by Moivre could be complex!}}$$

- Nosotros necesitamos un producto interno para pensar como "vectores" nuestras funciones.

$$\langle f, g\rangle = \int_{-\pi}^{\pi} f(x)\, g(x)\, dx$$

- f está siendo representada por $2N+1$ bases.

Recuerdo lejano álgebra lineal:

- Observe que: $V = \alpha_1 V_1 + \alpha_2 V_2 + \alpha_3 V_3$

y los e_i, e_j tal que $\langle e_i, e_j\rangle = \begin{cases} 1 & i=j \\ 0 & i\neq j \end{cases}$ ortogonales

$$\Rightarrow \underbrace{\begin{bmatrix} \langle e_1, e_1\rangle & \langle e_1, e_2\rangle & \langle e_1, e_3\rangle \\ \langle e_2, e_1\rangle & \langle e_2, e_2\rangle & \langle e_2, e_3\rangle \\ \langle e_3, e_1\rangle & \langle e_3, e_2\rangle & \langle e_3, e_3\rangle \end{bmatrix}}_{\text{Matriz de Gram}} \underbrace{\begin{bmatrix} a_0 \\ a_k \\ b_k \end{bmatrix}}_{\text{Coef}} = \begin{bmatrix} \langle f, e_1\rangle \\ \langle f, e_2\rangle \\ \langle f, e_3\rangle \end{bmatrix}$$

Ej $\langle V, e_1\rangle = \alpha_1 \langle v_1, v_1\rangle + \alpha_2 \overset{0}{\cancel{\langle v_1, v_2\rangle}} + \alpha_3 \overset{0}{\cancel{\langle v_1, v_3\rangle}} = \alpha_1$

Ahora, serán $\cos(kx)$, $\operatorname{sen}(kx)$, 1 ortogonales:

$$\langle 1, \cos(kx)\rangle = \int_{-\pi}^{\pi} 1 \cdot \cos kx \, dx = 0 \ (\text{si}) \ \perp$$

$$\langle \cos kx, \operatorname{sen} kx\rangle = \int_{-\pi}^{\pi} \cos kx \operatorname{sen} kx \, dx = 0 \ \perp \ (\text{si})$$

$$\langle \cos kx, \cos lx\rangle = 0 \quad k \neq l$$

$$\langle \operatorname{sen} kx, \operatorname{sen} lx\rangle = 0 \quad k \neq l$$

Entonces:

1) $$\langle 1, f(x)\rangle = \langle 1, a_0\rangle + \sum a_k \langle 1, \cos(kx)\rangle + \sum b_k \langle 1, \operatorname{sen}(kx)\rangle$$

$$\langle 1, f(x)\rangle = a_0 \int_{-\pi}^{\pi} 1 \, dx$$

$$\boxed{\frac{\langle 1, f(x)\rangle}{\langle 1, 1\rangle} = a_0} = \frac{1}{2\pi} \int_{\pi}^{\pi} f(x)\, dx$$

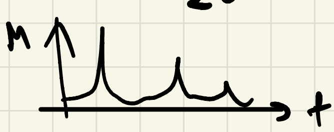
2) $a_k = \dfrac{\langle f(x), \cos(kx) \rangle}{\langle \cos kx, \cos kx \rangle} = \dfrac{1}{\pi}\displaystyle\int_{-\pi}^{\pi} f(x)\cos kx \, dx$

3) $b_k = \dfrac{\langle f(x), \sin(kx) \rangle}{\langle \sin(kx), \sin(kx) \rangle} = \dfrac{1}{\pi}\displaystyle\int_{-\pi}^{\pi} f(x)\sin(kx) \, dx$

a_k, b_k, a_0 : Amplitud de las frecuencias!

- El algoritmo va por otro camino pero la idea es calcular los coeficientes en ventanas temporales y calcular el valor absoluto de la representación de dichos coeficientes (Complejos).
- Vía moivre la representación de los coeff se combina y obtiene un coeff complejo!

- ML / 03 - Análisis y Visualización ...
- Correo rodlabs ...
- Fourier-transform.ipynb

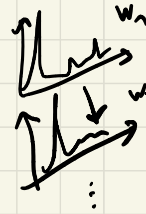


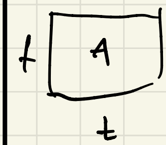
2) Espectrograma :

- Añadimos una dimensión más: el tiempo.
- Espectrograma
- short-time Fourier transform

3) Energía RMS :

Intuitivamente e informalmente "que tan fuerte suena un audio".





$|STFT(t,k)|^2$

$$E = \sum_{k=0}^{N} |X[k]|^2$$

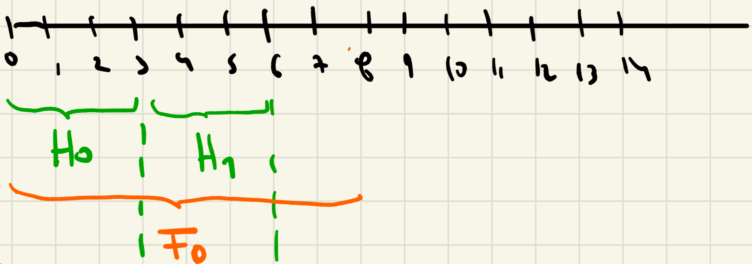
$$E_{RMS} = \sqrt{\sum_{n=0}^{N} |X[k]|^2}$$

Ejemplo :

hop-size = 3

frame-size = 8

len(s) = 14



$$E_0 = \sum_{k=0}^{7} |X_k|^2$$

$\vdots$

$$E_4 = \sum_{k=?}^{13} |X_k|^2$$

El vector será

$$\vec{E} = [E_0, E_1, E_2, E_3, E_4]$$

4) Zero Crossing Rate

$$ZCR = \frac{1}{N-1} \sum_{k=0}^{N-1} 1_{\mathbb{R}<0}(s_k \cdot s_{k-1})$$

$$1_{\mathbb{R}<0} = \begin{cases} 1 & \text{si } s_k \cdot s_{k-1} < 0 \\ 0 & \text{en otro caso} \end{cases}$$

$$[1,\ 0.5,\ -1,\ -2,\ 0.5,\ 1,\ 2]$$

$$\underbrace{0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 0}$$

$$\frac{1}{7-1} \sum \cdot 2 = 2/6 = 1/3$$

- Podemos aplicar por ventanas al igual que en el caso de la energía!
- Notebook: Zero Crossing Rate.

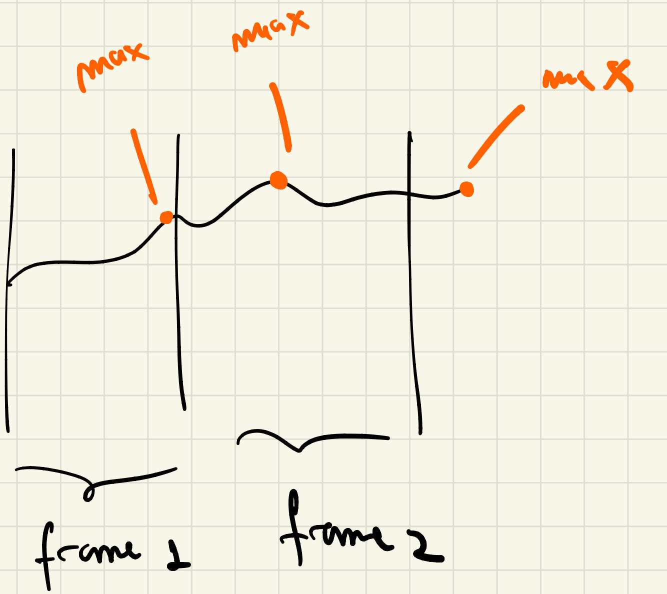
5) Amplitud o Envolvente

$$A(F_n) = \max_{\tilde{n} = t \cdot K}^{(t+1)(K-1)} s(k)$$

último sample del frame t

amplitud del k-esimo sample

Primer sample del frame t



$$A(F_0) = \max_{\tilde{n}=0}^{K-1} s(n)$$

- Entregado el código de Transformada de Fourier
- Energía, Zero Crossing Rate, Espectrograma, Librosa en general + audios snare/kicks.

(clase 5) : Problemas de Clasificación

Regresión Logística

Motivación: Supongamos queremos clasificar objetos como siendo A o B (Análogamente: $\{0, 1\}$; $\{-1, 1\}$ etc)

Esto se denomina clasificación binaria.