RESUMEN PARCIAL 3

COSTES

- ENTRENAMIENTO
  - KNN : $O(D.N)$ si normalizamos (si no, nada)
  - NB : $O(DN)$
  - RegLog : $O(\text{nepocas}.N.D)$
  - NN : $O(\text{nepocas}.N.D.J)$

- CLASIFICACIÓN / PREDICCIÓN
  - KNN : $O(DN)$
  - NB : $O(DK)$
  - RegLog : $O(D)$
  - NN : $O(DJ)$

MODELOS LINEALES

- Si $\mathbb{P}(C_1) = \mathbb{P}(C_2)$ y $\Sigma_1 = \Sigma_2 = \mathbb{I}$ entonces:
  - Si $(\bar{x} - \mu_2)^2 > (\bar{x} - \mu_1)^2 \Rightarrow C_1$
  - Si $(\bar{x} - \mu_2)^2 \leq (\bar{x} - \mu_1)^2 \Rightarrow C_2$

  $\bar{x}$ se clasifica C_1 si está más cerca de la media μ_1.

- Ahora $\mathbb{P}(C_1)$ puede ser diferente de $\mathbb{P}(C_2)$
  - Frontera de decisión: $\mathbb{P}(C_1|\bar{x}) = 0'5 \Rightarrow \sigma(a) = 0'5 \Rightarrow a = 0$

$$a = \frac{(\bar{x}-\mu_2)^2}{2} - \frac{(\bar{x}-\mu_1)^2}{2} + \ln\left(\frac{\mathbb{P}(C_1)}{\mathbb{P}(C_2)}\right) = 0$$

$$\Rightarrow \boxed{\underbrace{(\mu_1 - \mu_2)}_{\omega}\bar{x} + \underbrace{\frac{\mu_2^2 - \mu_1^2}{2} + \ln\left(\frac{\mathbb{P}(C_1)}{\mathbb{P}(C_2)}\right)}_{\omega_0 = b} = 0}$$ ECUACIÓN HIPERPLANO

  - Si $\mathbb{P}(C_1) = \mathbb{P}(C_2) \Rightarrow$ la frontera es la mediatriz de μ_1 y μ_2.
  - Si $\mathbb{P}(C_1) > \mathbb{P}(C_2) \Rightarrow$ la frontera se aleja de μ_1 hacia μ_2.

▶ $\mathbb{P}(\bar{x}|C_1) = \mathcal{N}(\mu_1, \Sigma)$ y $\mathbb{P}(\bar{x}|C_2) = \mathcal{N}(\mu_2, \Sigma)$

llegaríamos a
$$\begin{cases} w = \Sigma^{-1}(\mu_1 - \mu_2) \\ b = \frac{1}{2}\mu_1 \Sigma^{-1} \mu_1 + \frac{1}{2}\mu_2 \Sigma^{-1} \mu_2 + \ln\left(\frac{\mathbb{P}(C_1)}{\mathbb{P}(C_2)}\right) \end{cases}$$

→ muchos parámetros a estimar: $\mu_1, \mu_2, \Sigma, \mathbb{P}(C_1), \mathbb{P}$

$$\mu_1 \leftrightarrow D, \quad \mu_2 \leftrightarrow D, \quad \Sigma \leftrightarrow \frac{D(D+1)}{2}, \quad \mathbb{P}(C_1) \leftrightarrow 1$$

Sólo necesitamos w y $\underset{w_0}{b}$: $\boxed{D+1}$ $\quad \bar{w} = [w_0 \; w]$

$$\boxed{\mathbb{P}(C_1|\bar{x}) = \sigma(\bar{x}\bar{w})}$$

Obs: Si $\mathbb{P}(C_1|\bar{x}) = \sigma(\bar{x}\bar{w}) \Rightarrow$
$\Rightarrow \boxed{\mathbb{P}(C_2|\bar{x}) = 1 - \sigma(\bar{x}}$

• Frontera: $\sigma(\bar{x}\bar{w}) = 0'5 \Rightarrow \boxed{\bar{x}\bar{w} = \vec{0}}$

REGRESIÓN LOGÍSTICA

$t_j = 1$ si $\bar{x}_j$ es C_1, $t_j = 0$ si $\bar{x}_j$ es C_2

[...]

Para mejorar la verosimilitud del ejemplo $\bar{x}_j$ hay que mover la recta en sentido opuesto al gradiente $(\sigma_j - t_j)\bar{x}_j$ y proporcional a la constante de aprendizaje: $\bar{w} = \bar{w} - \eta(\sigma_j - t_j)\bar{x}_j$

▶ Algoritmo RL MV:

```
reg_log_mv(η, nepocas):
    generar w̄ ∈ [-0'5, 0'5]^(D+1):
    for ep in range(nepocas):
        for j in range(1, N):
            σ = σ(w̄x̄_j)
            w̄ = w̄ - η(σ - t_j)x̄_j
    return w̄
```

▶ Algoritmo RL MAP:

```
reg_log_map(η, nepocas, σ²):
    generar w̄ ∈ [-0'5, 0'5]^(D+1)
    for ep in range(nepocas):
        for j in range(1, N):
            σ = σ(w̄x̄_j)
            w̄ = w̄ - η(σ - t_j)x̄_j -
```

(*) En MAP suponemos que el módulo $|\bar{w}|$ viene dado por una gaussiana de media 0.

▶ Obs
→ RL solo trabaja con atributos numéricos
→ Se recomienda normalizar