

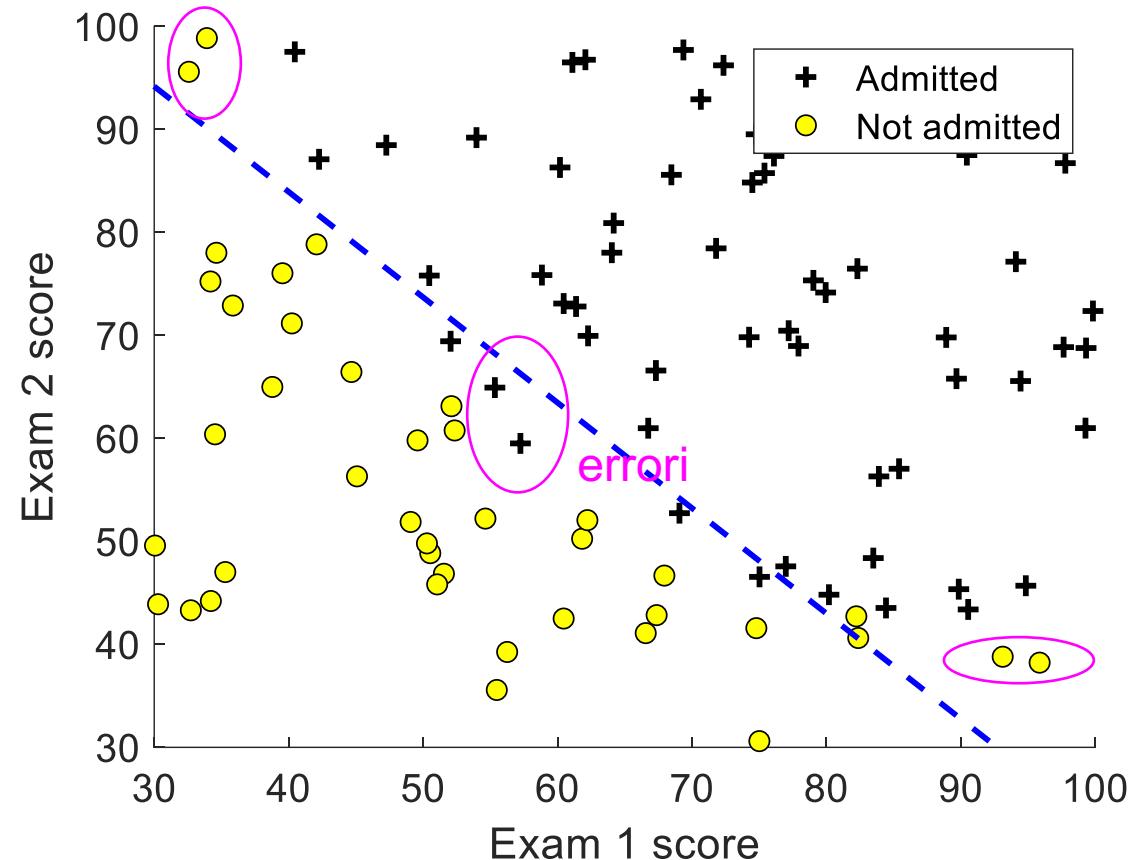
Esercizio: stima ammissione studenti

Vogliamo stimare la **probabilità di ammissione** $P(y = 1)$ di uno studente (o studentessa) all'università, visti i risultati di due esami (φ_1, φ_2), tramite una regressione logistica

Il dataset consiste di $N = 100$ studenti con
 $\varphi_1(i), \varphi_2(i)$ e $y(i) \in \{0,1\}$, per $i = 1, \dots, N$
e il vettore banale 1

$$P(y = 1|\boldsymbol{\varphi}) = s(a) = s(\boldsymbol{\varphi}^\top \boldsymbol{\theta}) = \frac{1}{1 + e^{-\boldsymbol{\varphi}^\top \boldsymbol{\theta}}}$$

- Matrice dei dati $X \in \mathbb{R}^{100 \times 3}$
- Vettore delle label $Y \in \mathbb{R}^{100 \times 1}$
- Vettore dei parametri $\boldsymbol{\theta} \in \mathbb{R}^{3 \times 1}$



Attività 2 - Regressione Logistica

es: stima ammissione studenti

Probabilità di ammissione $P(y=1)$

due esami (φ_1, φ_2)

$N = 100$ studenti (dataset)

$\varphi_1(i), \varphi_2(i) \in \{0, 1\}$ e $y(i) \in \{0, 1\}$ per $i = 1, \dots, N$

$$P(y=1 | \bar{\varphi}) = s(\alpha) = s(\bar{\varphi}^T \bar{\theta}) = \frac{1}{1 + e^{-\bar{\varphi}^T \bar{\theta}}}$$

Matrice dei dati $X \in \mathbb{R}^{100 \times 3}$

Vettore delle label $\bar{Y} \in \mathbb{R}^{100 \times 1}$

Vettore dei parametri $\bar{\theta} \in \mathbb{R}^{3 \times 1}$

? quanto vale la stima del vettore $\bar{\theta}$

? quanto vale l'accuratezza del modello sui dati di training

Regressione Logistica: Funzione di costo

$$J(\theta) = - \sum_{i=1}^N (y(i) \cdot \ln \pi(i) + (1-y(i)) \cdot \ln [1-\pi(i)])$$

$$\text{con } \pi(i) \equiv P(y(i)=1 | \bar{\varphi}(i)) = \frac{1}{1 + e^{-\bar{\varphi}^T(i) \bar{\theta}}}$$

$$\text{gradiente } \nabla_{\theta} J(\theta) = \sum_{i=1}^N \varphi(i) \cdot (\pi(i) - y(i))$$

In MATLAB:

$$J = -y^T \cdot \log(\text{sigmoid}(X \cdot \theta)) - (1-y)^T \cdot \log(1 - \text{sigmoid}(X \cdot \theta));$$

$$\text{grad} = X^T \cdot [\text{sigmoid}(X \cdot \theta) - y];$$

$$\text{con sigmoid}(z) = (1 + \exp(-z))^{-1}$$

Ottimizzazione utilizzando fminunc $\rightarrow \bar{\theta}$

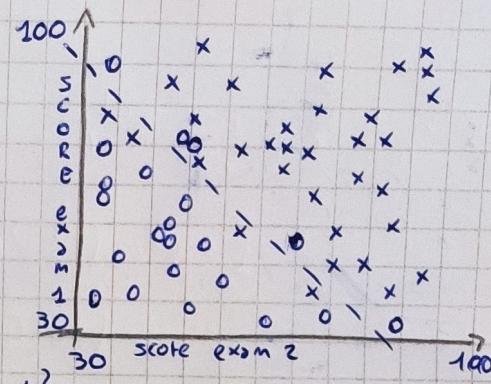
Preditzione: utilizzo sigmoid(z)

Preditzione di score1: 45 e score2: 85

$$\text{prob} = \text{sigmoid}([45 \ 85] \cdot \theta) \approx 0.776$$

Accuratezza: con la soglia dello 0,5 utilizzo la funzione round()

$$\text{acc} = \text{mean}(p == 1) \cdot 100 \quad \text{con } p = \text{round}(\text{sigmoid}(X \cdot \theta))$$



```
Plotting data with + indicating (y = 1) examples and o indicating (y = 0) examples.  
Cost at initial theta (zeros): 69.314718  
Gradient at initial theta (zeros):  
-10.000000  
-1200.921659  
-1126.284221  
Local minimum found.  
Optimization completed because the size of the gradient is less than  
the value of the optimality tolerance.  
<stopping criteria details>  
Cost at theta found by fminunc: 20.349770  
theta:  
-25.161343  
0.206232  
0.201472  
For a student with scores 45 and 85, we predict an admission probability of 0.776291  
Train Accuracy: 89.000000  
>>
```

