Problem Set 2 {Supervised Learning II}

1. Logistic Regression: Training stability

Goal: Develop your skills debugging machine Learning Algorithms.

→ Implementation of Logistic regression is Provided in src/pol_lr.py.

→ Two labeled dataset:

① data/ds1_a.txt

② data/ds1_b.txt

(a) * On dataset A, the training Convergo in 2,78,103 iteration.

- * On dataset B, the training didnot Convergo until 50,00p,00 iteration

↳ Maybe it never Convergo

(b) x, red → y == 0

o, blue → y == 1

learning_rate = 1

$\|\Delta\theta\| < 10^{-15}$ for Convergence

Observation

1) for dataset A ~~algorithm~~ in the learning phase, θ tends to Converge to a fixed point

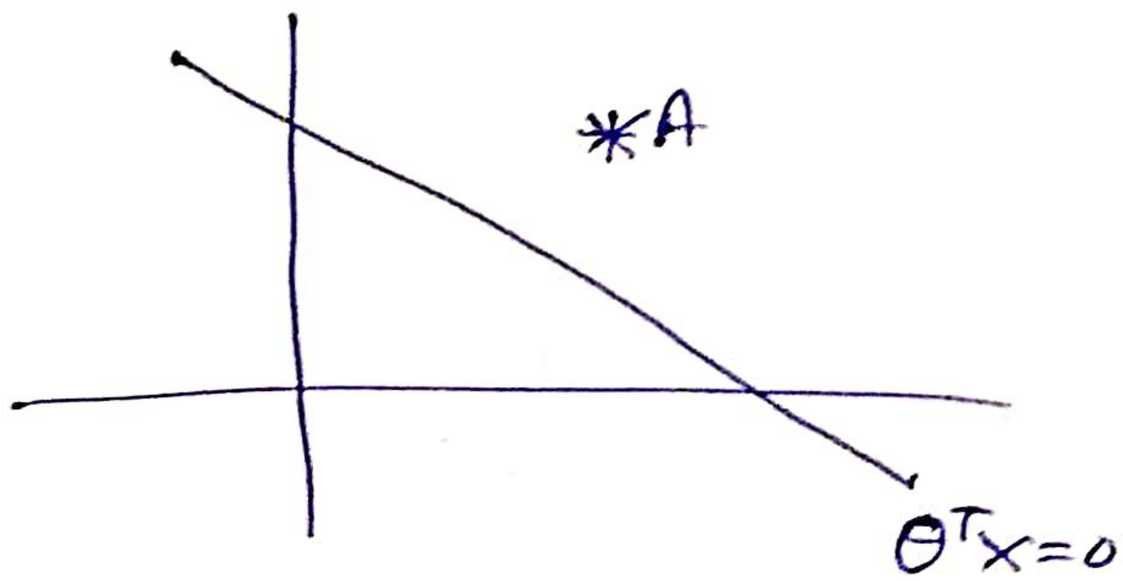
2) for dataset B, in the learning phase, θ tends to continuously increase.

⇒ If the dataset can be perfectly separated by a linear decision boundary, then the objective $l(\theta)$ can be arbitrarily maximized by just scaling the θ.

→ Multiplying θ by a factor does not change the decision boundary.

→ But probability being assigned to each data point can be arbitrarily changed

$$h_\theta(A) = \frac{1}{1+e^{-\theta^T A}}$$



If $\theta \to k\theta$

$$h_\theta^*(A) = \frac{1}{1+e^{-k\theta^T A}}$$

$$h_\theta(A) < h_\theta^*(A)$$

⇒ as $k \to \infty$ $h_\theta^*(A) \to 1$

- ds1_a.txt
  → Cannot be separated perfectly, so we are able to converge at some θ.
- ds2_b.txt
  → Can be seperated pefectly, so we are not able to converge at any θ.