FTML Exercices 4

Pour le 7 avril 2023

TABLE DES MATIÈRES

1 SGD 1
2 Regression logistique 1
3 Convexité 2

1 SGD

Essayer de faire converger SGD sur l'exercice du TP4 sur la régression logistique en expérimentant avec le taux d'apprentissages γ et les autres hyperparamètres.

2 REGRESSION LOGISTIQUE

We consider the classical logistic regression problem, in the following setting :

— $\mathcal{X} = \mathbb{R}^d$

— $\mathcal{Y} = \{-1, 1\}$

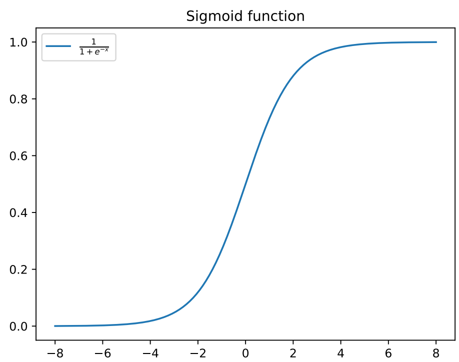
— Logistic loss :

$$l(\hat{y}, y) = \log(1 + e^{-\hat{y}y}) \tag{1}$$

We also define the following function :

Definition 1. Sigmoid function
$\sigma : \mathbb{R} \to \mathbb{R}$.

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{2}$$



1] Show that σ is differentiable and that

$$\forall z, \sigma'(z) = \sigma(z)\sigma(-z) \tag{3}$$

2] Show that $l(\hat{y}, y)$ is strictly convex in its first argument, which means for fixed y, $\hat{y} \mapsto l(\hat{y}, y)$ is strictly convex. Studying the derivatives will be helpful.

3] Without regularization, the empirical risk writes :

$$R_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} l(x_i^T\theta, y_i) \tag{4}$$

Compute the gradient $\nabla_\theta R_n(\theta)$ of the empirical risk $R_n(\theta)$ in this setting.

Note that during the practical session, we did introduce a regularization $\mu||\theta||^2$, but it is straightforward to deduce the gradient of the regularized risk from the gradient of the non-regularized one.

3 CONVEXITÉ

Connaître la définition de fonctions convexes, strictement convexes et fortement convexes et des exemples pour chaque cas. Vous pouvez trouver les définitions dans lectures notes.pdf (section 2.3).