Exercices 4 solutions

TABLE DES MATIÈRES

1 LOGISTIC REGRESSION GRADIENT

We use the following conventions :

— $\mathcal{X} = \mathbb{R}^2$

— $\mathcal{Y} = \{-1, 1\}$

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

Hence

$$\begin{aligned}
\frac{\partial l}{\partial \hat{y}}(\hat{y}, y) &= \frac{-ye^{-\hat{y}y}}{1 + e^{-\hat{y}y}} \\
&= \frac{-ye^{-\hat{y}y}e^{\hat{y}y}}{(1 + e^{-\hat{y}y})e^{\hat{y}y}} \\
&= \frac{-y}{1 + e^{\hat{y}y}} \\
&= -y\sigma(-\hat{y}y)
\end{aligned} \tag{2}$$

The empirical risk writes :

$$R_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} l(x_i^T\theta, y_i) \tag{3}$$

To compute the gradient, we can proceed as in Exercices 3 with only a difference in $\frac{\partial l}{\partial \hat{y}}(\hat{y}, y)$.

$$\nabla_\theta g_i = x_i(-y_i\sigma(-x_i^T\theta y_i)) \tag{4}$$

and

$$\begin{aligned}
\nabla_\theta R_n &= \frac{1}{n}\sum_{i=1}^{n} \nabla_\theta g_i \\
&= \frac{1}{n}\sum_{i=1}^{n} (-y_i\sigma(-x_i^T\theta y_i))x_i
\end{aligned} \tag{5}$$