Artificial Intelligence Lab 7: Linear and Logistic Regression

(Q1) [25 Marks] Data in the file *linear* is given in the form $(x_i, y_i)_{i=1}^n$, where $x_i \in \mathbb{R}$, and $y_i \in \mathbb{R}$. Let $w = (w(1), w(2)) \in \mathbb{R}^2$. Learn the optimal w_* for loss function $L(w) = \sum_i L_i(w)$, where $L_i(w) = (w(1)x_i + w(2) - y_i)^2$.

(Q2) [25 Marks] Data in the file *logistic* is given in the form $(x_i, y_i)_{i=1}^n$, where $x_i \in \mathbb{R}^2$, and $y_i \in \{-1, +1\}$. Let $w = (w(1), w(2), w(3)) \in \mathbb{R}^3$. Let $\sigma(w^\top x) = \frac{1}{1+exp(-(w(1)x(1)+w(2)x(2)+w(3)))}$ be the likelihood that example x belongs to class $+1$. Learn the optimal w_* for loss function $L(w) = \sum_i L_i(w)$, where $L_i(w) = -\log \sigma(y_i w^\top x_i)$. [25 Marks]

(Q3) [50 Marks] `import` the MNIST dataset (link: http://yann.lecun.com/exdb/mnist/), and use logistic regression to classify digits 4 versus 7.