2. (6 points) [THINKING LOGISTIC-ALLY...] Consider the scenario in which a user maintains a dataset consisting of songs that he has downloaded over a period of time. He also tracks the likes (-1)/dislikes(+1) for each song along with a set of features X_1, X_2. X_1 is a binary variable that takes value 1 if the song is sung by his favorite singer, and X_2 corresponds to song duration in minutes. This dataset with 10 data points is given below:

$$[X_1 \ X_2] = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 5 & 10 & 13 & 2 & 3 & 5 & 2 & 10 & 10 & 3 \end{bmatrix}^{\mathsf{T}}$$

$$y = \begin{bmatrix} -1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}^{\mathsf{T}}$$

(a) (3 points) Train a logistic regression model on this dataset by performing the gradient descent algorithm steps by setting the initial weights to 0. No bias (intercept term) is required and the step size $\eta = 1$. Report the updated weights at the end of two iterations.