109021115 吳嘉濬

1. Write down your regression equation in basic part:

   My regression equation in **basic** part is

   $X\_ploy = w_0 + w_1 x + w_2 x^2$, where x means the input feature **weight**

2. Briefly describe the variables and the regression equation you used in the advanced part:

   My regression equation in **advanced** part is

   $X\_poly = w_0 + w_1 x_0 + w_2 x_1 + w_3 x_2 + \cdots + w_7 x_6 + w_8 x_0^2 + w_9 x_1^2 + w_{10} x_2^2 + \cdots + w_{14} x_6^2$, where

   x_0 means the input variable **age**, x_1 means the input variable **gender**,

   x_2 means the input variable **height**, x_3 means the input variable **weight**,

   x_4 means the input variable **bodyFat**, x_5 means the input variable **diastolic**,

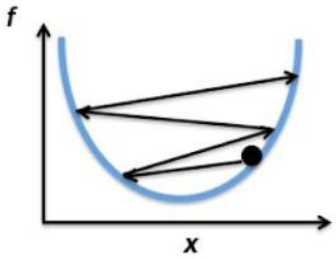
   x_6 means the input variable **systolic**

3. Briefly describe the difficulty you encountered:

   When I on the first try of gradient descent on the regression model, I found that the parameter is jumping between positive and negative with larger and larger magnitude, then in no more than 10000 rounds, the parameters are divergent and returned error in the middle of process.

4. Summarize how you solve the difficulty and your reflections:

   I found the phenomena about the parameters are positive in this round, then negative in next round, and back to positive and so forth. It occurred to me that this is very likely happen when learning rate is set to large(see the graph), so I immediately decrease the learning rate and find out that it works!!!



   After the difficulty is solved, I have learned that I need to be extremely careful about every hyperparameter, and it turns out that I spend a lot more time on tuning the hyperparameters and training model, rather than coding the training process and model structure.