Name: Tianxiu Ma

Student number: 230277302

Assignment number: Assignment 1

module code: ECS708U/ECS708P

Q5. What conclusion if any can be drawn from the weight values? How does gender and BMI affect blood sugar levels? What are the estimated blood sugar levels for the below examples? [2 marks]

1. Weights represent the effect of each feature on the predicted target (blood sugar levels in this case). A correlation is said to be positive or negative depending on the weights assigned to it.
2. The higher the weight, the greater the impact on the result.
3. The weight value of 'Sex' is `-11.4488` and the weight value of 'BMI' is `26.3047`. BMI has a greater impact on blood sugar levels.
4. The estimated blood sugar levels are `66.4351` and `209.3253`.

Now estimate the error on the test set. Is the error on the test set comparable to that of the train set? What can be said about the fit of the model? When does a model over/under fits?

```
pred=model(x_test)
cost = mean_squared_error(y_test, pred)
print(cost)
print(min(cost_lst))
```

```
tensor(2885.6191)
tensor(2890.4065)
```

1. The test and training set errors have very similar magnitudes, with the test error being somewhat lower than the training error (2890.4065 vs. 2885.6191). This shows that the error difference is pretty minimal and the model generalizes effectively to the test set.
2. Overfitting is typically observed when the test error is significantly higher than the training error. Underfitting occurs when both the training and test errors are high.

Q6. Try the code with a number of learning rates that differ by orders of magnitude and record the error of the training and test sets. What do you observe on the training error? What about the error on the test set? [3 marks]

The error result of the training and test sets with a number of learning rates that differ by orders of magnitude and record is shown below:

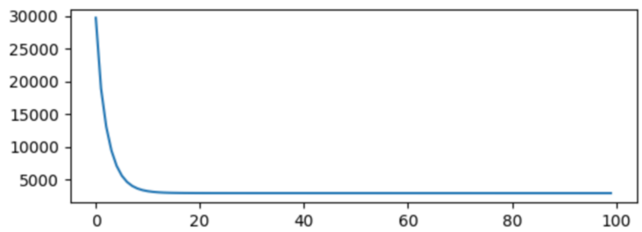
```
learning rate: 0.1
train error: 2890.406494140625
test error: 2885.619140625
learning rate: 0.01
train error: 3356.77734375
test error: 3431.068359375
learning rate: 0.001
train error: 20040.583984375
test error: 18534.30859375
```

```
learning rate: 0.0001
train error: 28468.08203125
test error: 25530.2265625
learning rate: 1e-05
train error: 29583.326171875
test error: 26443.8125
learning rate: 1
train error: 29711.322265625
test error: nan
learning rate: 10
train error: 29711.322265625
test error: nan
```

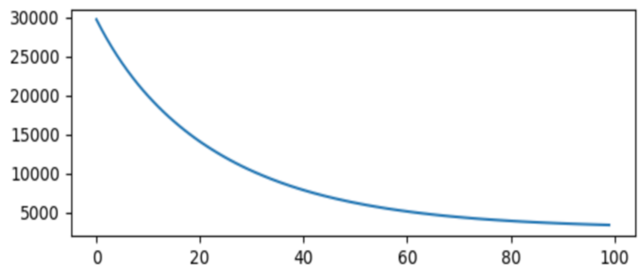
Observations:

- Both the training and test errors are poor when the learning rate is very high (10 and 1), and the model overfits.
- Training and test mistakes are closely spaced in a well-fitted model produced by a moderate learning rate (0.1).
- Overfitting results from lower learning rates (0.01, 0.001, 0.0001, 1e-05), where the training error stays low.

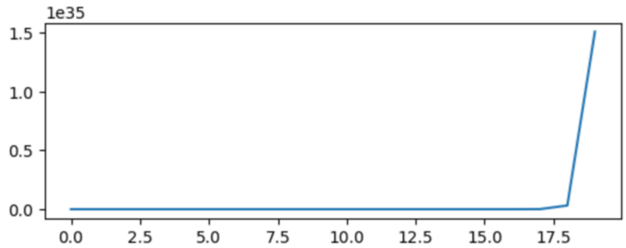
Learning rate:0.1



Learning rate:0.01



Learning rate: 1



Q8. First of all, find the best value of alpha to use in order to optimize best. Next, experiment with different values of λ and see how this affects the shape of the hypothesis. [3 marks]

This is the data for learning rate and Minimum cost:

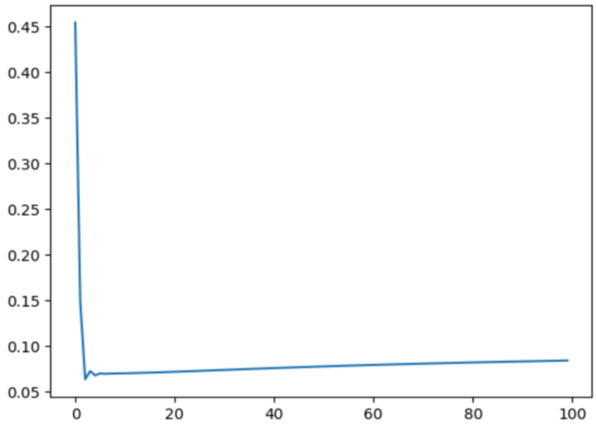
```
learning rate: 0.1
Minimum cost: 0.010472000576555729
learning rate: 0.01
Minimum cost: 0.06233194097876549
learning rate: 0.001
Minimum cost: 0.36066627502441406
learning rate: 1
Minimum cost: 0.008261877112090588
learning rate: 10
Minimum cost: 0.45409098267555237
```

Summary:

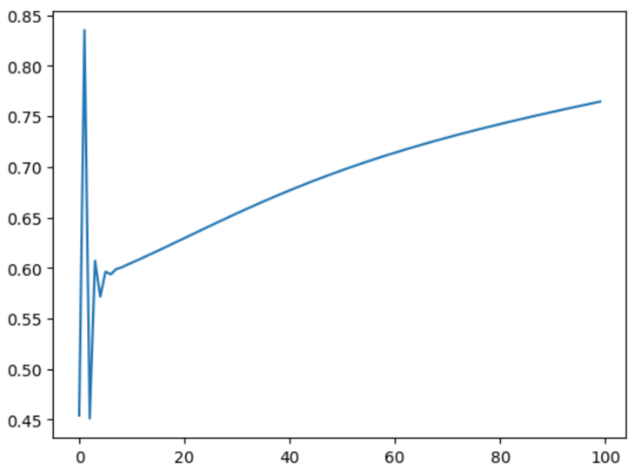
Judging from the influence of the above learning rate and loss function, when the learning rate(alpha) is 1, the corresponding loss is the smallest. So the learning rate of 1 can be used to optimize best.

the shape of the hypothesis with different values of λ :

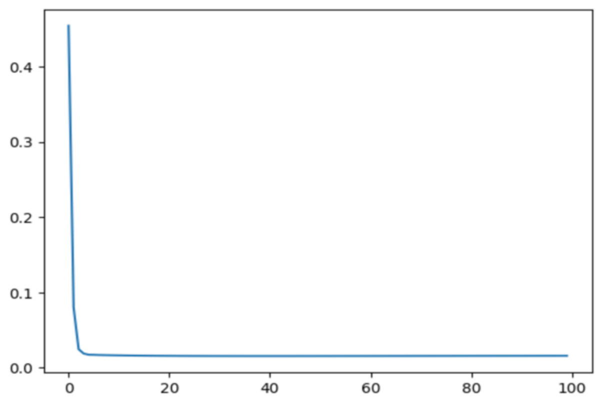
```
lam = 1
```



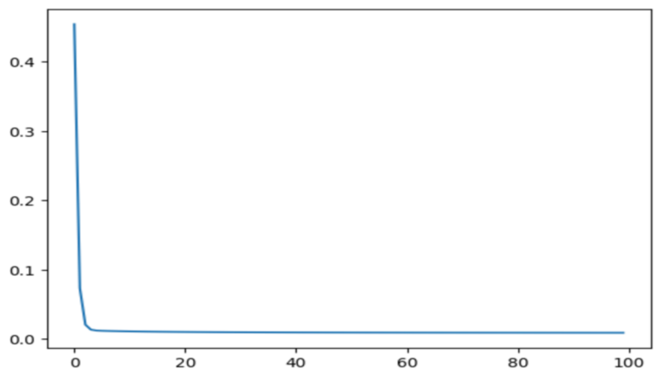

```
lam = 10
```



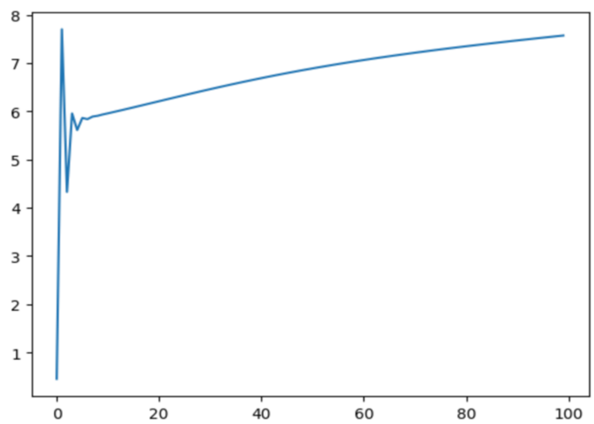

```
lam = 0.1
```




```
lam = 0.01
```




```
lam = 100
```



Lamda(λ) penalize the weight parameters in order to make the weight parameters as smooth as possible.