**Exercise 1 in machine learning**

**崔颖琦 kyrie 1730004002**

1. Given the gradient for each of the following functions:

Given the following data points:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| x | -1.0 | -0.5 | 0.0 | 0.5 | 1.0 |
| y | 1.0 | 0.5 | 0.0 | 0.5 | 2.0 |

Questions:

1. What is the cost function for linear regression?

Answer: (m=5)

1. If we use the gradient descent algorithm to minimize the cost function for linear regression, what are the θ values and cost values in the first three iterations? Suppose the initial θ values are [1, 0.5]Tand the learning rate α is 0.1.

Answer:

Before the gradient descend:

The first iteration:

=1-0.1\*0.2\*1

=0.98

=0.5-0.1\*0.2\*(0.25)

=0.495

=0.1\*2.0846

=0.20846

The seconde iteration:

Theta=[0.98;0.495]

=0.98-0.1\*0.2\*0.9

=0.962

=0.495-0.1\*0.2\*(0.2375)

=0.49025

=0.1\*2.0516

=0.20516

The third iteration:

Theta=[0.962;0.49025];

=0.962-0.1\*0.2\*0.81

=0.9458

=0.49025-0.1\*0.2\*(0.22563)

=0.48574

=0.1\*2.0685

=0.20247

1. Will the gradient descent algorithm converge if the learning rateα is 1? Suppose the initial θ values are [1, 0.5]T.

Answer:

Before the gradient descent, the cost is 0.21250. After once gradient descent, the cost is 0.19063. And after 4 step of gradient descent, the cost of the hypothesis will stay at 0.190 Because the cost of the hypothesis is decreasing with the process of gradient descent, and the final cost is equal the cost when the learning rate is 0.1.Although there are only 4 steps to get to the minimum, the gradient descent algorithm will converge if the learning rate is 1.

1. What are the θ values if we use normal equation method to minimize the cost function for linear regression?

Answer:

X=[1 ,-1;1,-0.5;1, 0;1,0.5;1,1]

Y=[1;0.5;0;0.5;2]

Theta==

=[0.2,0;0,0.4][1,1,1,1,1; -1 , -0.5, 0 , 0.5 , 1][1; 0.5 ; 0 ; 0.5; 2]

=[0.2 , 0.2 , 0.2 , 0.2 , 0.2 ; -0.4 , -0.2 , 0 , 0.2 , 0.4][1 ; 0.5 ; 0 ; 0.5 ; 2]

=

So theta=