linear regression:

h（theta）： 预测函数

cost function： J(theta): 预测值与实际结果的差距



grade： 对cost function在当前点对theta求导。

min J(theta): 梯度下降

gradient descend work well even n (number of features) is big.

gradient descent(kn2) k = number of iterations

normal equation(n3)

注意：

1.X是原数据和在第一列加一列1。对应theta1，x0。

2.第一次写的时候，犯了个经典的错误，用theta直接接受theta下降的值，导致没有做到同时theta 下降。theta1立刻改变后会影响theta2。

A screenshot of a social media post

Description automatically generated

A close up of text on a whiteboard

Description automatically generated