气象目标对雷达波后向散射能力的强弱通常称为气象目标的强度，常用的表示气象目标强度的参量有**反射率和反射率因子**。单位体积重云雨粒子后向散射截面的总和，称为气象目标的反射率。降水目标物单位体积中降水粒子直径6次方的总和称为雷达反射率因子，用Z表示，其常用单位为mm6 / m3。反射率因子Z值的大小，反映了气象目标内部降水粒子的尺度和数密度，常用来表示气象目标的强度。

i）当前降水量与雷达折射率之间的关系；

ii）雷达图包含当前目标站点及其周边地区的雷达反射率。需要考虑目标地点与周边地区之间的降水关系;

iii）最后，我们有不同时间跨度的雷达地图。根据历史数据，可能会发现降水量演变的一些模式。

大数据集下跑回归：

1. 分批训练

So what we can see is that running 20 regressions can take (a little) more time (from what we’ve seen earlier) than running only one on the whole dataset…. but it provides better estimates. So the tradeoff is not that simple, and maybe running several regressions on huge datasets can be a proper alternative.

1. Create a a distributed computing system

Mini\_batch 有助于防止陷入局部最小，如果是大的数据集训练，可能有多个局部低谷，会陷入局部最小，可以选择小一点

选择大一点，可以充分利用GPU,step\_size 尽量大，但不要超过 by the 1/L (smoothness) constraint

调节batch大小，使达到收敛

1. 全部数据作为一个batch和 B.每个采样作为一个batch:

B方法基于每个采样去修正，修正幅度大了以后，就容易跳出这些局部极值，避免过拟合发生。B方法一般不依赖于初始模型，所以可以用来训练初始的神经网络。之后再用A方法或者下面介绍的C方法优化

Batch = 1，基于SGD的batch不能取太大

batch数太小，而类别又比较多的时候，真的可能会导致loss函数震荡而不收敛，

Full Batch Learning 可以使用Rprop 只基于梯度符号并且针对性单独更新各权值。

以 Rprop 的方式迭代，会由于各个 Batch 之间的采样差异性，各次梯度修正值相互抵消，无法修正，才有了后来 **RMSProp** 的妥协方案

SGD regression

我们可以首先去除特征中不频繁的值，这样特征对应的取值减少，维数会降低。但是这种方法需要对数据进行预处理。至于PCA等常见的降维方法，由于数据量实在太大而不太适合使用。而hash trick是一种越来越受欢迎的降维方法。它不需要进行数据预处理，实现简单直接。

特征哈希法的目标是把原始的高维特征向量压缩成较低维特征向量，且尽量不损失原始特征的表达能力。

特征是水平归一化？？