# Week4-Thinking

### Thinking1 ALS都有哪些应用场景

1.ALS是一种重要的数据拟合技术，可应用于线形回归、非线形回归

2.ALS是一种数学的优化方法，可以解决最优化问题。

在推荐的场景中使用较多。

ALS分为显示反馈和隐式反馈两种：显示反馈ALS、隐式反馈ALS-WR。

ALS适用于解决由明确评分矩阵的应用场景；

ALS-WR通过置信度权重c来解决用户没有明确反馈对物品的偏好，即对于更确信用户偏好的项赋予较大的权重，对于没有反馈的项，赋予较小的权重。

### Thinking2 ALS进行矩阵分解的时候，为什么可以并行化处理

ALS是先固定X求解Y，再固定Y求解X，在矩阵求解过程中，比如固定Y，求解X的时候，目标评分矩阵A，X的每一行可以独立求解，X的第I行和Y计算得到A的第i行，这样对每一步的X或者Y的行或者列都可以独立并行求解，故ALS可以进行并行化计算。

### Thinking3 梯度下降法中的批量梯度下降（BGD），随机梯度下降（SGD），和小批量梯度下降有什么区别（MBGD）

区别在于每次迭代使用的样本量不同

1.BGD是最原始的形式，每次迭代时使用所有样本来进行梯度的更新。

优点：

1）一次迭代是对所有样本进行计算，比较稳定，利用矩阵进行操作，实现了并行；

2）由全数据集确定的方向能够更好地代表样本总体，能更准确地朝极值所在的方向。当目标函数是凸函数时，BGD能够得到全局最优。

缺点：

当样本数目m很大时，每迭代一步都需要对所有样本计算，训练过程很慢，收敛慢。

2.SGD是每次迭代从m个样本随机抽取一个样本来对参数进行更新，使得训练速度加快。优点：

1）由于不是在全部训练数据上的损失函数，而是在每轮迭代中，随机优化某一条训练数据上的目标函数，这样每一轮参数的更新速度大大加快，收敛快，最终解在全局最优解附近。

缺点：

1）准确度下降，即使在目标函数为强凸函数的情况下，SGD仍旧无法做到线性收敛；2）可能会收敛到局部最优，由于单个样本并不能代表全体样本的趋势；

3）不易于并行实现。

3.MBGD是综合了BGD和SGD的优缺点，每次迭代只使用一部分样本batch\_size来进行更新。

优点：

1）通过矩阵运算，每次在一个batch上优化参数并不会比单个数据慢太多；

2）每次使用一个batch可以大大减小收敛所需要的迭代次数；

3）可实现并行化。

缺点：batch\_size的不当选择可能会带来的一些问题

### Thinking4 你阅读过和推荐系统/计算广告/预测相关的论文么？有哪些论文是你比较推荐的，可以分享到微信群中

之前对于这方面的论文阅读确实很少，推荐大家一个我之前只是粗略阅读过的文章，题目是《Variational Autoencoders for Collaborative Filtering 》

A close up of a newspaper

Description automatically generated