Metric Factorization: Recommendation beyond Matrix Factorization度量分解

Shuai Zhang，UNSW and Data61, CSIRO，Sydney, NSW, Australia

Lina Yao，University of New South Wales，Sydney, NSW, Australia

Yi Tay，Nanyang Technological University，Singapore

Xiwei Xu，Data61, CSRIO，Sydney, NSW, Australia

Xiang Zhang，University of New South Wales，Sydney, NSW, Australia

Liming Zhu，Data61, CSRIO，Sydney, NSW, Australia

CSRIO：联邦科学与工业研究组织

1. 主要内容

本文假设用户与项目（物品）可放置在低维空间中，可测量明确接近度。基于矩阵分解的推荐模型中采用的点积不满足不等性，可能会限制表达性，导致次优解。

基于欧几里得距离作为相似性度量的度量学习方法。本文通过分解来学习度量空间中的低秩结构。因子化度量学习（FML）：包含距离而不是评级得分的因子化偏好矩阵，将偏好转换为距离，欧几里得距离替换点积，将距离矩阵分解为用户和项目密集嵌入。矩阵因子分解常过度拟合大尺寸的潜在因素，限制灵活性和能力。

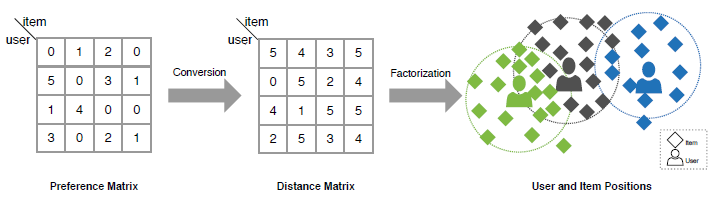
1. 创新点

提出了一种新颖的技术，即分解度量学习（FML），其中文中探索了在度量向量空间中寻找低秩结构的概念。在本文的方法中，用户和项目被表示为多维坐标系统中的点（即，度量向量空间）。建议基于用户和项目的接近程度，在度量标准空间中定义。然而，与其他度量学习方法不同，本文模型背后的关键新颖性是它对度量空间进行了分解。

本文指定两种FML变体来解决两个经典且成熟的推荐任务：评级预测和项目排名。本文的模型可以有效地了解两种设置中用户和项目的位置。

1. 模型架构/思路

FML的高级概述。：



M个用户，N个项目，偏好（评级或隐式交互）矩阵由R∈RM×N表示，Rui表示用户u对项目i的偏好，值可以是明确的评级或隐式反馈。

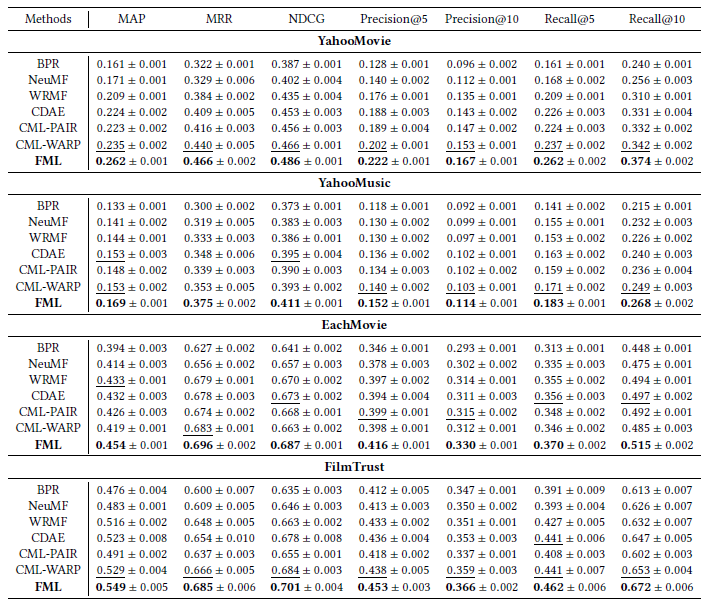
将用户偏好转换为距离：，距离函数满足1.非负性；2.不可分辨的；3.同一性；4.对称性；5.三角不等式。用户与项目之间的距离：。之后将距离矩阵分解并学习用户和项目的位置，将隐式反馈转换为距离：，若Rui=0,则Yui=a，若Rui=1，则Yui=0。排名任务的损失函数定义为逐点丢失：。

评级预测中，将评级矩阵转换为距离：，Rmax为最大评分，Rui为用户评级，得分函数为，D(u,i)为用户与项目之间的距离，损失函数为，同时为每个评分添加置信度，置信机制为，更高的置信给予更可靠的评级，预测评分：。使用Adagrad [4]优化评级估计和排名学习模型，降低了学习率的需求。

1. 实验

隐式反馈项目排名表现评估：

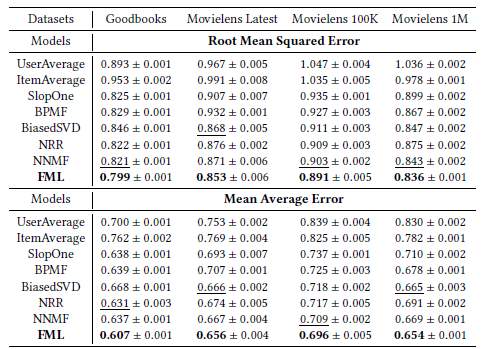
数据集：Yahoo Research Dataset，FilmTrust，EachMovie，评价指标：Recall，Precision，MAP，MRR，NDCG，后三项为质量评估。



评级估计评估：

数据集：Goodbooks，Movielens Datasets

评价指标：RMSE，MAE，评估预测误差



1. 总结
2. FML可以更准确地反映用户和项目之间的距离关系
3. 基于欧几里得距离的模型比基于点积的推荐器和神经网络模型表现更好
4. 增加非线性是不够的
5. FML将用户和项目视为度量向量空间中的点，通过分解了解点的位置