

# 用于极大规模低秩矩阵的快速奇异值分解处理的系统和方法



信息和通信

计算机/人工智能/数据处理和信息技术

## 机会

奇异值分解是降维、数据压缩和特征提取的基础工具。然而，为大规模矩阵计算奇异值分解成本极高。一个100万×100万的稠密矩阵需要7.27 TB内存——远超典型桌面计算机的能力。传统奇异值分解算法仍需存储整个矩阵、执行大型矩阵乘法，且时间复杂度非常高。即使矩阵具有低秩结构（在图像集或时间序列数据等实际应用中很常见），现有方法在典型内存约束下也无法扩展到百万级矩阵。因此，需要一种仅使用部分数据、无需将完整矩阵加载到内存中即可处理极大规模矩阵的奇异值分解算法。

## 技术

本专利提出CUR-SVD方法，通过仅操作一小部分行和列来计算超大矩阵的奇异值分解，其核心创新在于让内存无法装下的大矩阵也能在普通电脑上完成SVD分解，解决了传统方法需要将整个矩阵加载到内存中的根本限制。该方法首先从完整矩阵中随机采样少量列形成矩阵C，同时随机采样少量行形成矩阵R。不同于单纯依赖随机选取，它进一步应用离散经验插值方法从这些采样中智能筛选出信息量最大的行和列，从而保证极高的近似精度。获得紧凑的C和R矩阵后，CUR-SVD对它们进行奇异值分解，并通过一个轻量级中间矩阵恢复出完整的奇异值和奇异向量，全程无需直接处理原始大矩阵。该方法的一个关键突破在于能够自动处理未知秩的情况：它从一小块行和列开始计算，检查奇异值是否已趋于稳定，若未稳定则使用增量奇异值分解逐块增加行列直至收敛，整个过程无需从头重新计算。在整个计算过程中，完整矩阵永远不会被加载到内存中，仅从磁盘读取被选中的行和列。这种内存效率具有变革性意义：专利中展示了在一个百万乘百万的矩阵上计算奇异值分解仅需约3GB内存，而传统方法需要超过7TB，这使得大型矩阵分解在普通台式计算机上即可实现，广泛应用于图像处理、主成分分析和机器学习等领域。

## 优势

- 极端可扩展性：计算100万×100万矩阵的奇异值分解仅需3 GB内存——传统方法无法实现。
- 快速运行时间：百万矩阵的前200个奇异值/向量在约82秒内计算完成。
- 低内存占用：内存需求仅正比于选定的行/列数，而非完整矩阵维度。
- 可比的准确性：对低秩矩阵实现约万亿分之一量级的相对误差，与标准奇异值分解精度匹配。
- 支持未知秩：固定精度版本使用收敛停止准则自动估计秩。
- 在实际数据上验证：在大规模人脸图像数据集上，主成分分析速度比标准奇异值分解快12倍；重建的人脸在视觉上与原图无法区分。

### 备注

CIMDA: P00128

### IP状态

已申请专利



技术成熟度等级 (TRL) ?

4

### 发明人

严洪教授

张泽松教授

郑琦宣博士

Mr. Muhammad Ashraf Abdelhamid  
ABDELGAWAD

查询: kto@cityu.edu.hk



## 应用

- 主成分分析：对无法放入内存的大规模图像数据集进行降维。
- 推荐系统：对百万级用户-物品矩阵进行协同过滤的奇异值分解。
- 时间序列分析：从流体动力学、气候数据或大型传感器网络中提取主导模式。
- 压缩感知与数据压缩：科学计算中大规模矩阵的低秩近似。
- 机器学习预处理：在模型训练前减少大规模文本或基因组数据的特征空间。

