摘要

张量是矩阵和向量的高阶推广，现实世界中的很多数据都有着自然的张量形式。然而在实际应用中，受多种因素的影响，原始的张量数据很可能是不完备的，或者有某些元素是缺失的。所以我们需要根据张量已知的元素和一些对张量结构的先验假设来补全整个张量数据。这其中，一个常用的假设是原始的张量是低秩的，或者是近似低秩的。在这种情况下我们的目的就变成了在给定的线性约束下去优化张量的某种秩。然而，高阶张量的秩的定义并不是唯一的，并且每一种秩都对应着某种特定的张量分解方法。在这些张量分解方法中，CP (CANDECOMP/PARAFAC)分解的复杂度和张量的阶数呈理想的线性关系。但CP 分解有着两个理论上的主要缺点：一是 CP 分解的计算是 NP 困难的，目前依然没有有效的计算方法；二是 CP 秩不是下半连续的函数，在实际计算中是不稳定的。相比于 CP 分解，张量的 Tucker 分解是一种数值稳定的分解方式，但 Tucker 分解的复杂度会随着张量阶数的增长呈指数增长。为了克服两种分解各自的缺点，本文考虑利用最近提出的张量 Tensor Train (TT)分解。TT 分解的计算同样是数值稳定的，并且其复杂度和张量的阶数呈比较理想的线性关系，从而在一定程度上解决了前面两种分解存在的问题。

在假设原始张量是 TT 分解意义下的低秩张量后，我们提出了基于 TT 分解的低秩张量补全模型。我们借鉴了低秩矩阵补全理论中的方法，即用矩阵核范数来逼近矩阵的秩。矩阵核范数是矩阵的秩在单位球上最紧的凸包，所以用矩阵核范数来逼近矩阵的秩是矩阵补全问题中一个非常有效的方法。张量的 TT 核范数是矩阵核范数的一种推广，在本文中我们利用张量的 TT 核范数来逼近张量的 TT 秩，从而将原始的秩优化模型松弛为了一个凸优化模型。我们分别基于 Douglas-Rachford 分裂方法和 ADMM 设计了两种高效的一阶算法。数值算例也证明了我们所提出的算法在实际问题中的有效性。最后，在文章的结尾处我们讨论了一些可以用于加速我们算法的技巧和策略。

关键词：张量补全；低秩张量；Tensor Train (TT)分解；凸优化

Abstract

Tensors are higher order generalizations of matrices and vectors. Many real-world data have a natural tensor form. However, in many applications the tensor of interest may be incomplete or may have corrupted elements. And one often needs to estimate the missing elements based on the partially observed data and some presumed structures of the original tensor. One commonly used assumption is that the tensor is of low rank, or approximately so. In such cases, to complete the partially observed tensor we aim at optimizing the tensor rank under given linear constraints. However, the definitions of tensor rank are not unique for order higher than 3. And each definition corresponds to a certain tensor decomposition. The CP (CANDECOMP/PARAFAC) decomposition, although has the ideally linear complexity with the tensor order, suffers from the theoretical deficits that its computation is NP-hard and it is not a lower semi-continuous function. The Tucker decomposition provides an alternate numerically stable way to decompose a tensor, but its complexity grows exponentially with the tensor order. To tackle these issues, we resort to the recently proposed tensor train decomposition which shares the merits of the CP and the Tucker decompositions, as it is numerically stable and its complexity scales linearly with the tensor order.

By assuming a low rank prior of the original tensor under the Tensor Train decomposition, we propose a low rank tensor completion model and derive two algorithms to solve this model. We borrow the idea from the low rank matrix completion theory in which the matrix rank is approximated by matrix nuclear norm, which is the tightest convex envelope for the matrix rank on the unit ball. This relaxation is also proved to be a power approach for matrix completion. As an extension of matrix nuclear norm, the tensor train nuclear norm is used to approximate the tensor train rank and the original model in thus converted into a convex program. We then design two algorithms that are based on the Douglas-Rachford splitting technique and the classical ADMM to efficiently solve this convex program. Numerical experiments validated the effectiveness of the proposed algorithms. And in the final part we end our discussion with some possible speed-up strategies to further improve algorithm performance.

Key Words: tensor completion; low rank tensors; tensor train decomposition; convex optimization.

目录