基于张量分解的卫星遥测缺失数据预测算法

马 友 贾树泽 赵现纲 冯小虎 范存群* 朱爱军
(国家卫星气象中心 北京 100081)

摘 要：卫星健康状况监测是卫星安全保障的重要基础，而卫星遥测数据又是卫星健康状况分析的唯一数据来源。因此，卫星遥测缺失数据的准确预测是卫星健康分析的重要前瞻性手段。针对极轨卫星多组成系统、多仪器载荷以及多监测指标形成的高维数据特点，该文提出一种基于张量分解的卫星遥测缺失数据预测算法(TFP)，以解决当前数据预测方法大多面向低维数据或只能针对特定维度的不足。所提算法将遥测数据中的系统、载荷、指标以及时间等多维因素作为统一的整体进行张量建模，以完整、准确地表达数据的高维特征；其次，通过张量分解计算数据模型的成分特征，通过成分特征可对张量模型进行准确重构，并在重构过程中对缺失数据进行准确预测；最后，提出一种高效的优化算法实现相关的张量计算，并对算法中最优参数设置进行严格的理论推导。实验结果表明，所提算法的预测准确度优于当前大部分预测算法。

关键词：极轨卫星；遥测数据；缺失数据预测；张量分解
中图分类号：TN927; TP391 **文献标识码**：A **文章编号**：1009-5896(2020)02-0403-07
DOI: 10.11999/JEIT180728

Missing Telemetry Data Prediction Algorithm via Tensor Factorization

MA You JIA Shuze ZHAO Xiangang FENG Xiaohu
FAN Cunqun ZHU Aijun
(*National Satellite Meteorological Center*, *Beijing* 100081, *China*)

Abstract: Satellite health monitoring is an important concern for satellite security, for which satellite telemetry data is the only source of data. Therefore, accurate prediction of missing data of satellite telemetry is an important forward-looking approach for satellite health diagnosis. For the high-dimensional structure formed by the satellite multi-component system, multi-instrument and multi-monitoring index, the Tensor Factorization based Prediction (TFP) algorithm for missing telemetry data is proposed. The proposed algorithm surpasses most existing methods, which can only be applied to low-dimensional data or specific dimension. The proposed algorithm makes accurate predictions by modeling the telemetry data as a Tensor to integrally utilize its high-dimensional feature; Computing the component matrixes via Tensor Factorization to reconstruct the Tensor which gives the predictions of the missing data; An efficient optimization algorithm is proposed to implement the related tensor calculations, for which the optimal parameter settings are strictly theoretically deduced. Experiments show that the proposed algorithm has better prediction accuracy than the most existing algorithms.

Key words: Satellite; Telemetry data; Missing data prediction; Tensor factorization

1 引言

卫星健康监测是卫星业务中重要的基础性工作，在各种相关业务中处于最高等级。遥测数据是分析卫星健康状况的唯一数据来源，其实时记录了卫星平台及载荷的功能和性能参数，完整地反应了卫星的健康状况。遥测数据通过卫星的数据发射机传送给地面测控站，并通过相应的规则来分析卫星平台以及各载荷的健康状况。数据完整性是数据分析的前提，然而针对极轨卫星进行遥测数据分析时，数据的缺失是必须解决的问题，这是因为：(1)极轨卫星不能和地面进行全天候的实时通信，这类卫星相对于地面总是在改变位置，只有当这些卫星飞入地面测控站的接收范围时才具备星地通信的条件；(2)对于某一测控站来说，由于接收范围有限，每次通信只有约15 min的时间，通信结束后需要再等待约90 min才能进入下一次通信；(3)地

收稿日期：2018-07-19；改回日期：2019-04-20；网络出版：2019-09-27
*通信作者： 范存群 fancq@cma.gov.cn
基金项目：国家自然科学基金(61602126)，国家863计划项目(2011AA12A104)
Foundation Items: The National Natural Science Foundation of China (61602126), The National 863 Plan Project (2011AA12A104)

面测控站的数量有限，各站之间不能进行无时隙的通信接力，造成通信的空窗期，此期间地面人员无法获取卫星遥测数据。

因此如何尽量准确地预测通信空窗期的卫星遥测数据，以补全数据完整性，是后续遥测数据分析的重要基础。虽然针对卫星遥测数据预测的相关研究尚不多见，但由于该问题在数学本质上是缺失数据的预测问题，其他很多领域已经进行过类似的研究，比如推荐系统中的缺失评分预测[1]，电子商务中的销量预测[2]以及Web服务中的QoS预测等[3,4]。然而，已有的预测方法很难应用于极轨卫星遥测数据的预测，这是因为：(1)卫星是一个复杂的系统工程，包含多个子系统，每个子系统又包含多个监测对象，监测对象又有多个监测指标，并考虑到时间因素，这总体构成了一个多维度的数据模型；(2)多监测目标、不同指标之间又有一定的相关性，不能割裂开来进行分析，这就决定了针对卫星遥测数据的预测需要在多个维度上进行统一、完整以及整体的考虑，以充分利用多维数据的整体特征进行准确分析；(3)部分已有方法虽然也考虑到了数据不同维度的特征，但没有将各维度作为一个统一的整体进行考虑，而是针对某一特定的维度进行，从而忽视了各个维度之间的关系。这些方法包括诸如时间感知的、地点感知的以及环境感知的预测方法等[5–7]。这种针对某一特定维度的预测方法难以扩展到另一维度上，如果将针对所有维度的预测方法强行组合，将导致预测方法非常臃肿复杂从而难以使用。

针对以上问题，本文提出了一种基于张量分解的卫星遥测缺失数据预测算法(Tensor Factorization based Prediction algorithm for missing telemetry data, TFP)。TFP算法首先利用“张量”概念对高维遥测数据进行建模，对遥测数据的各个维度进行了统一、完整的分析；然后通过张量分解计算出遥测数据张量模型的成分矩阵，这些成分矩阵可以对遥测数据张量模型进行准确的重构；重构的过程即进行了缺失遥测数据的预测。

2 TFP算法

TFP算法考虑高维遥测数据的整体结构，以充分利用数据的维度特征实现各维度上缺失数据的准确预测。为达到此目标，两个核心问题需要解决：(1)选择何种模型对高维遥测数据进行建模，(2)如何用所建模型实现缺失数据的预测。

TFP算法通过张量的概念对高维遥测数据进行建模，张量在表现形式上是多维数组，使其非常适合于数据的高维建模；把传统预测方法中矩阵分解的概念扩展到张量后，发现张量的分解与重构操作同样具有数据预测的效果。

2.1 张量的表示形式

令$\mathcal{X}\in\mathbb{R}^{I_1\times I_2\times\cdots\times I_N}$表示一个$N$维张量，其第$n$维的长度为$I_n$ ($1\le n\le N$)。例如，本文中的卫星遥测数据需要4个维度来描述：卫星整体包含I_1个子系统，每个子系统有I_2个监测目标，每个监测目标包含I_3个监测指标，每个监测指标都在I_4个时刻进行数据采样。这样，卫星遥测数据通过4维张量$\mathcal{X}\in\mathbb{R}^{I_1\times I_2\times I_3\times I_4}$进行建模，$\mathcal{X}_{i_1i_2i_3i_4}$是该张量中的一个元素，表示第$i_1$个系统中第$i_2$个目标的第$i_3$个指标在第$i_4$个时刻的监测值。

2.2 TFP算法原理

TFP算法的基本原理为：(1)通过张量分解计算张量的成分矩阵，即使张量中包含缺失数据也同样可以分解；(2)通过成分矩阵对张量进行重构，缺失数据也同时获得重构，即完成了数据预测。

2.2.1 张量分解

张量分解首先需要确定张量的秩，在介绍张量秩的概念之前，首先给出秩1 (rank one)张量的定义如下：

定义 1 秩1张量：有N维张量$\mathcal{X}\in\mathbb{R}^{I_1\times I_2\times\cdots\times I_N}$，若$\mathcal{X}=\boldsymbol{a}^{(1)}\circ\boldsymbol{a}^{(2)}\circ\cdots\circ\boldsymbol{a}^{(N)}$，其中，$\boldsymbol{a}^{(i)}$表示长度为$I_i$的向量，符号“$\circ$”表示向量的外积运算，则$\mathcal{X}$是一个秩1张量，且$\mathcal{X}_{i_1i_2\cdots i_N}=\boldsymbol{a}^{(1)}_{i_1}\boldsymbol{a}^{(2)}_{i_2}\cdots\boldsymbol{a}^{(N)}_{i_N}$。

定义1说明：若$\mathcal{X}$是一个N维的秩1张量，则$\mathcal{X}$可表示为N个向量的外积，这是最简单的张量分解形式。

实际中大部分张量的秩都大于1，不能按照定义1分解，因此，给出秩R张量的概念如下：

定义 2 秩R张量：若N维张量$\mathcal{X}\in\mathbb{R}^{I_1\times I_2\times\cdots\times I_N}$是$R$个秩1张量的和，即

$$\mathcal{X}=\sum_{r=1}^{R}\mathcal{X}_r \tag{1}$$

其中，每个$\mathcal{X}_r\,(1\le r\le R)$是秩1张量，且$\mathcal{X}_r\in\mathbb{R}^{I_1\times I_2\times\cdots\times I_N}$，则$\mathcal{X}$为秩$R$张量，$R$称为张量的秩。

根据定义1和定义2，对秩R张量进行分解时，首先把该张量看作R个秩1张量的和，然后对每一个秩1张量$\mathcal{X}_r\,(1\le r\le R)$按式(2)进行分解

$$\mathcal{X}_r=\boldsymbol{a}_r^{(1)}\circ\boldsymbol{a}_r^{(2)}\circ\cdots\circ\boldsymbol{a}_r^{(N)} \tag{2}$$

因此，N维秩R张量$\mathcal{X}\in\mathbb{R}^{I_1\times I_2\times\cdots\times I_N}$可按式(3)分解为

$$\mathcal{X}=\sum_{r=1}^{R}\boldsymbol{a}_r^{(1)}\circ\boldsymbol{a}_r^{(2)}\circ\cdots\circ\boldsymbol{a}_r^{(N)} \tag{3}$$

若令$\boldsymbol{a}_r^{(j)}$是一个列向量，则对应的R个列向量$\boldsymbol{a}_1^{(j)},\boldsymbol{a}_2^{(j)},\cdots,\boldsymbol{a}_R^{(j)}$构成了一个$I_j\times R$矩阵，令$\boldsymbol{A}^{(j)}$表示该矩阵。将$\boldsymbol{a}_r^{(j)}$改写为$\boldsymbol{A}_{:r}^{(j)}$，则式(3)可改写为式(4)

$$\mathcal{X}=\sum_{r=1}^{R}\boldsymbol{A}_{:r}^{(1)}\circ\boldsymbol{A}_{:r}^{(2)}\circ\cdots\circ\boldsymbol{A}_{:r}^{(N)} \tag{4}$$

式(4)即为本文所使用的张量分解形式，并称$\boldsymbol{A}^{(j)}(1\le j\le N)$为张量$\mathcal{X}$的成分矩阵。

由于计算机的浮点精度有限，式(4)中的“=”一般不能严格成立，因此用“≈”代替。式(4)右端可看作近似的重构张量，并令$\hat{\mathcal{X}}$表示这个近似于$\mathcal{X}$的重构张量。

2.2.2 张量重构

按式(4)将张量分解后，包括缺失值在内的$\mathcal{X}$中的所有数据，可按式(5)近似计算为

$$\mathcal{X}_{i_1i_2\cdots i_N}\approx\hat{\mathcal{X}}_{i_1i_2\cdots i_N}=\sum_{r=1}^{R}\boldsymbol{A}_{i_1r}^{(1)}\cdot\boldsymbol{A}_{i_2r}^{(2)}\cdot\cdots\cdot\boldsymbol{A}_{i_Nr}^{(N)} \tag{5}$$

式(5)即为张量重构。

2.2.3 缺失遥测数据预测

基于张量的分解和重构操作，TFP算法的预测步骤分为两步：首先按式(4)分解遥测数据的张量模型以获得成分矩阵；其次按式(5)重构张量模型，缺失数据的预测值即包含在重构的张量中。

2.3 TFP算法实现

按2.2节所述，TFP算法的主要工作在于张量分解，张量重构的操作可由式(5)直接完成。和传统的矩阵分解一样，张量分解是一个典型的优化问题，为此需确定合适的优化目标和优化策略。

2.3.1 张量分解优化目标

准确的数据预测应满足：对于$\mathcal{X}$中的已知数据使得

$$||\mathcal{X}-\hat{\mathcal{X}}||^2=\sum_{\mathcal{X}_{i_1i_2\cdots i_N}\text{已知}}e_{i_1i_2\cdots i_N}^2\le\varepsilon \tag{6}$$

其中，$e_{i_1i_2\cdots i_N}=\mathcal{X}_{i_1i_2\cdots i_N}-\hat{\mathcal{X}}_{i_1i_2\cdots i_N}$，$\varepsilon$表示业务允许的预测误差。由于数据噪声的影响，简单地选择式(6)作为优化目标会造成过拟合。因此，在其基础上将优化目标调整为

$$L=\sum_{\mathcal{X}_{i_1i_2\cdots i_N}\text{已知}}\Big[e_{i_1i_2\cdots i_N}^2+\lambda\Big(\left\|\boldsymbol{A}_{i_1:}^{(1)}\right\|^2+\left\|\boldsymbol{A}_{i_2:}^{(2)}\right\|^2+\cdots+\left\|\boldsymbol{A}_{i_N:}^{(N)}\right\|^2\Big)\Big]\le\varepsilon \tag{7}$$

其中，$\lambda\left(\left\|\boldsymbol{A}_{i_1:}^{(1)}\right\|^2+\left\|\boldsymbol{A}_{i_2:}^{(2)}\right\|^2+\cdots+\left\|\boldsymbol{A}_{i_N:}^{(N)}\right\|^2\right)$是防止过拟合的正则化项，$\lambda$为正则化参数，称$L$为该优化问题的损失函数。

2.3.2 张量分解优化算法

针对式(7)所代表的优化问题，梯度下降是最常见的优化算法，但该类算法具有越靠近极值点收敛越慢的缺点，为避免该缺点，TFP算法使用iRPROP$^+$[8]作为张量分解的优化算法。iRPROP$^+$具有以下优点：迭代步长可根据迭代状态自适应调整以不断加快收敛速度。成分矩阵中的元素$\boldsymbol{A}_{ijr}^{(j)}$的迭代过程为

步骤 1　计算L关于$\boldsymbol{A}_{ijr}^{(j)}$的1阶偏导数

$$g_{ijr}^{(j)}=\frac{\partial L}{\partial\boldsymbol{A}_{ijr}^{(j)}}=\sum_{\mathcal{X}_{i_1i_2\cdots i_N}\text{已知}}\Big(-2e_{i_1i_2\cdots i_N}\boldsymbol{A}_{i_1r}^{(1)}\boldsymbol{A}_{i_2r}^{(2)}\cdots\boldsymbol{A}_{i_{j-1}r}^{(j-1)}\boldsymbol{A}_{i_{j+1}r}^{(j+1)}\cdots\boldsymbol{A}_{i_Nr}^{(N)}+2\lambda\boldsymbol{A}_{ijr}^{(j)}\Big) \tag{8}$$

令$g_{ijr}^{(j)}|_t$表示第t–1次迭代刚结束时$g_{ijr}^{(j)}$的值。

步骤 2　当第t–1次迭代计算完成后，第t次迭代过程如下

(1) 若$g_{ijr}^{(j)}|_t\cdot g_{ijr}^{(j)}|_{t-1}>0$，则第$t$次迭代和第$t$–1次迭代的收敛方向一致，为加快收敛速度，加大迭代步长如式(9)

$$\delta_{ijr}^{(j)}|_t=\delta_{ijr}^{(j)}|_{t-1}\cdot\eta^+ \tag{9}$$

其中，$\delta_{ijr}^{(j)}|_t$表示第t次迭代的步长，$\delta_{ijr}^{(j)}$通过增长因子$\eta^+(>1)$变大。令$\boldsymbol{A}_{ijr}^{(j)}|_t$表示第$t$–1次迭代刚结束时$\boldsymbol{A}_{ijr}^{(j)}$的值，则第$t$次迭代按式(10)计算

$$\boldsymbol{A}_{ijr}^{(j)}|_{t+1}=\boldsymbol{A}_{ijr}^{(j)}|_t-\text{sign}\left(g_{ijr}^{(j)}|_t\right)\cdot\delta_{ijr}^{(j)}|_t \tag{10}$$

其中，$\text{sign}(x)=\begin{cases}1, & x>0\\ 0, & x=0\\ -1, & x<0\end{cases}$。

(2) 若$g_{ijr}^{(j)}|_t\cdot g_{ijr}^{(j)}|_{t-1}<0$，则说明第$t$–1次迭代由于步长过大从而错过了$L$的极值点，因此，第$t$次迭代应当按相反的方向收敛。令$L|_t$表示第$t$–1轮迭代结束后$L$的值，如果

$$L|_t>L|_{t-1} \tag{11}$$

则说明第t–1轮迭代使损失函数L变大，不符合优化目标。此时，忽略第t–1次迭代得到的$\boldsymbol{A}_{ijr}^{(j)}|_t$，回溯至第$t$–2次迭代重新计算，即

$$\boldsymbol{A}_{ijr}^{(j)}|_{t+1}=\boldsymbol{A}_{ijr}^{(j)}|_{t-1} \tag{12}$$

若式(11)不成立，说明第t–1次迭代虽然越过了极值点，但相比上次迭代更靠近极值点，此时不需回溯，而是将迭代步长缩小后按式(10)计算第t次迭代。按式(13)缩小迭代步长

$$\delta_{ijr}^{(j)}|_t=\delta_{ijr}^{(j)}|_{t-1}\times\eta^- \tag{13}$$

其中，$\eta^-(0<\eta^-<1)$称为减小因子。

(3) 若$g_{i_jr}^{(j)}|_t \cdot g_{i_jr}^{(j)}|_{t-1}=0$，则不需要调整迭代步长，即$\delta_{i_jr}^{(j)}|_t=\delta_{i_jr}^{(j)}|_{t-1}$，并按式(10)计算第$t$次迭代。

每一轮迭代结束后，一旦$L\le\varepsilon$则说明达到优化目标，迭代结束。这样就得到了遥测数据张量模型的成分矩阵$\boldsymbol{A}^{(j)}(j=1, 2, \cdots, N)$，然后按式(5)即可直接算出相关的缺失数据。算法1详细总结了TFP算法的工作过程。

算法1：TFP算法

输入：数据集$\mathcal{X}\in\mathbb{R}^{I_1\times I_2\times\cdots\times I_N}$；

输出：训练后的成分矩阵$\boldsymbol{A}^{(j)}$ (j=1 to N)

随机初始化成分矩阵$\boldsymbol{A}^{(j)}(j=1 \text{ to } N)$

Repeat

For each $\boldsymbol{A}_{i_jr}^{(j)}$ $(1\le j\le N, 1\le i_j\le I_j, 1\le r\le R)$

If $g_{i_jr}^{(j)}|_t \cdot g_{i_jr}^{(j)}|_{t-1}>0$

$\delta_{i_jr}^{(j)}|_t=\min\left(\delta_{i_jr}^{(j)}|_{t-1}\cdot\eta^+, \text{MaxSize}\right)$

$\boldsymbol{A}_{i_jr}^{(j)}|_{t+1}=\boldsymbol{A}_{i_jr}^{(j)}|_t-\text{sign}\left(g_{i_jr}^{(j)}|_t\right)\cdot\delta_{i_jr}^{(j)}|_t$

Else If $g_{i_jr}^{(j)}\cdot g_{i_jr}^{(j)\prime}<0$

$\delta_{i_jr}^{(j)}|_t=\max\left(\delta_{i_jr}^{(j)}|_{t-1}\cdot\eta^-, \text{MinSize}\right)$

If $L|_t>L|_{t-1}$

$\boldsymbol{A}_{i_jr}^{(j)}|_{t+1}=\boldsymbol{A}_{i_jr}^{(j)}|_t+\text{sign}\left(g_{i_jr}^{(j)}|_{t-1}\right)\cdot\delta_{i_jr}^{(j)}|_{t-1}$

$L|_t=0$

End If

Else

$\delta_{i_jr}^{(j)}|_t=\delta_{i_jr}^{(j)}|_{t-1}$

$\boldsymbol{A}_{i_jr}^{(j)}|_{t+1}=\boldsymbol{A}_{i_jr}^{(j)}|_t-\text{sign}\left(g_{i_jr}^{(j)}|_t\right)\cdot\delta_{i_jr}^{(j)}|_t$

End If

End For

Until $L\le\varepsilon$ or maximum iterations exhausted

2.4 TFP算法时间复杂度分析

如算法1所述，张量分解算法的循环边界为：$1\le j\le N\&\&1\le i_j\le I_j\&\&1\le r\le R$，其中$N$指训练集$\mathcal{X}$的维度，$I_j$指第$j$维的长度，$R$表示$\mathcal{X}$的秩。因此，表面上每次循环的计算量为$R\cdot\prod_{j=1}^{N}I_j$，其中$\prod_{j=1}^{N}I_j$等于$\mathcal{X}$的容量(即已知数据和缺失数据的总个数)。实际上，由式(7)可知，算法1只根据$\mathcal{X}$中的已知数据进行计算，所以算法1每次循环的时间复杂度为$O(C\cdot R)$，其中C指训练集$\mathcal{X}$中已知数据的个数。因此TFP算法的时间复杂度为$O(S\cdot C\cdot R)$，其中S为达到收敛到目标所需的迭代次数。TFP算法使用iRPROP$^+$作为优化算法，虽然所有优化算法的迭代次数都无法确定，但可以确定的是，相比于梯度下降、随机梯度下降等方法，iRPROP$^+$是迭代速度最快的优化算法之一[8]。

3 实验

通过风云三号B气象卫星遥测数据，对TFP算法进行了验证，并和其它有代表性的方法进行了对比。

3.1 数据集介绍

数据采用风云三号B气象卫星2018年5月1日的遥测数据。数据包含卫星16个子系统，每个子系统包含3～16个电子仪器，每个电子仪器包含21～129个监测指标。因此，数据整体上是包含系统、仪器、指标以及时间在内的4维结构。

3.2 实验方法

以完整的、不含缺失值的数据为基础，随机挖掉其中的某些数据，对挖掉的数据进行预测，然后通过预测值与原实际值的对比来评价预测的准确性。

3.3 TFP算法预测结果以及与其它算法的对比

将TFP算法进行最优参数设置后(设置方法将在3.4节进行阐述)，与其它5个重要算法进行了对比，这5个算法是：(1)时间感知的预测方法(TA)[5]；(2)非负矩阵分解(NMF)[9]；(3)概率矩阵分解(PMF)[10]；(4)基于用户的协同过滤(UPCC)[11]；(5)基于物品的协同过滤(IPCC)[12]。

由于以上5个方法并不能直接用于多维数据结构，必须先将遥测数据张量拆分成2维矩阵才能执行这些算法，而TFP算法直接运行在多维张量上。根据3.4节介绍的实验方法，把需要预测的数据挖掉后，剩下的数据称为训练集(training data)，将训练集数据量占原数据集数据量的比例称为数据密度。实验结果如表1所示，其中，MAE为平均绝对误差，RMSE为均方根误差。实验结果表明：数据密度越大预测误差越小，和常见的预测算法相比，所提TFP算法具有更高的准确性。

卫星的全部检测指标有近10^4个，以其中的一个重要指标(太阳阵电流)为例，将预测值误差的概率分布通过图1进行了统计。在50%数据密度下有1440个预测值，最小误差0.035，最大误差0.235，将误差范围分割为50个等长区间，即每个区间的长度为0.004。图1展示了误差在不同区间的分布频度。通过将该分布频度进行归一化拟合后，近似趋近于$N\left(0.12, 0.31^2\right)$的正态分布，和表1中训练集为50%数据密度时的预测结果吻合。

3.4 TFP算法的最优参数设置

3.4.1 ε设置

根据式(7)，ε控制张量分解的精度，即张量分解的结束条件。业务期望的预测精度决定了ε设

表 1　TFP算法与其它5个方法的对比

方法	数据密度5%		数据密度10%		数据密度20%		数据密度50%	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
NMF	0.6175	1.5789	0.6007	1.5485	0.5986	1.5233	0.4870	1.4847
PMF	0.5687	1.4792	0.4984	1.2842	0.4492	1.1855	0.4006	1.0820
UPCC	0.6204	1.4010	0.5513	1.3139	0.4875	1.2343	0.3114	1.0749
IPCC	0.6886	1.4278	0.5908	1.3245	0.4454	1.2094	0.2895	1.1724
TA	0.6239	1.4058	0.5360	1.3045	0.4496	1.2030	0.2106	1.0988
TFP	**0.3815**	**0.9469**	**0.3073**	**0.7597**	**0.2270**	**0.5619**	**0.1235**	**0.3150**

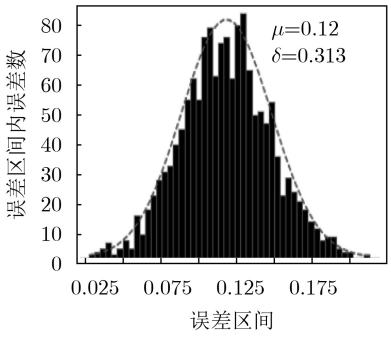

图 1 预测误差在不同区间的分布

置。令P表示业务期望的预测精度，M表示训练集中数据的个数，即$P=\frac{1}{M}\sum_{\mathcal{X}_{i_1i_2\cdots i_N}\text{已知}}e^2_{i_1i_2\cdots i_N}$，并让式(7)中损失函数的两组成部分相等(将在3.4.2节阐述这样做的原因)，即

$$\sum e^2_{i_1i_2\cdots i_N}=\sum_{\mathcal{X}_{i_1i_2\cdots i_N}\text{已知}}\lambda\left(\left\|\boldsymbol{A}^{(1)}_{i_1:}\right\|^2+\left\|\boldsymbol{A}^{(2)}_{i_2:}\right\|^2+\cdots+\left\|\boldsymbol{A}^{(N)}_{i_N:}\right\|^2\right)=\frac{\varepsilon}{2}\tag{14}$$

则由式(14)可得

$$\varepsilon=2\cdot P\cdot M\tag{15}$$

实验中设置P=0.01，即期望预测误差在0.01之内，则ε可根据相应训练集中已知元素的个数得出。

3.4.2　λ设置

根据式(7)，λ控制损失函数L两组成部分的比例，文献[13,14]证明了这两部分应成正比例关系

$$k=\frac{\sum_{\mathcal{X}_{i_1i_2\cdots i_N}\text{已知}}e^2_{i_1i_2\cdots i_N}}{\sum_{\mathcal{X}_{i_1i_2\cdots i_N}\text{已知}}\lambda\left(||\boldsymbol{A}^{(1)}_{i_1:}||^2+||\boldsymbol{A}^{(2)}_{i_2:}||^2+\cdots+||\boldsymbol{A}^{(N)}_{i_N:}||^2\right)}\tag{16}$$

文献[15]又在此基础上指出k的恰当取值为1，即令损失函数两端相等。假设张量$\mathcal{X}$已经被准确地分解，则由式(5)可得：对于$\mathcal{X}$中的任意元素，有$\mathcal{X}_{i_1i_2\cdots i_N}\approx\sum_{r=1}^{R}\boldsymbol{A}^{(1)}_{i_1r}\cdot\boldsymbol{A}^{(2)}_{i_2r}\cdot\cdots\cdot\boldsymbol{A}^{(N)}_{i_Nr}$。因此，$\boldsymbol{A}^{(j)}_{i_jr}$可被近似估计为$\boldsymbol{A}^{(j)}_{i_jr}\approx\sqrt[N]{\frac{\mathcal{X}_{i_1i_2\cdots i_N}}{R}}$，这样，对于每一$\mathcal{X}_{i_1i_2\cdots i_N}\in\mathcal{X}$，有

$$||\boldsymbol{A}^{(1)}_{i_1:}||^2+||\boldsymbol{A}^{(2)}_{i_2:}||^2+\cdots+||\boldsymbol{A}^{(N)}_{i_N:}||^2\approx NR\left(\frac{\mathcal{X}_{i_1i_2\cdots i_N}}{R}\right)^{\frac{2}{N}}\tag{17}$$

在分解精度很高的要求下，意味着张量$\mathcal{X}$可以被准确地分解为成分矩阵，即式(17)满足，又因为式(16)中已经设置$k=1$，则根据式(14)，式(16)以及式(17)可共同推导出

$$\lambda=\frac{\varepsilon}{2NR\sum\limits_{\mathcal{X}_{i_1i_2\cdots i_N}\text{已知}}\left(\frac{\mathcal{X}_{i_1i_2\cdots i_N}}{R}\right)^{\frac{2}{N}}}\tag{18}$$

3.4.3　R设置

根据式(1)，R表示张量的秩，一般张量的秩很难精确判定[16]，对于一个张量$\mathcal{X}\in\mathbb{R}^{I_1\times I_2\times\cdots\times I_N}$，其秩由其形状即$I_1,I_2,\cdots,I_N$确定[17]，一般通过实验的方法寻找最优的$R$。在期望预测精度$P$=0.01的基础上，测试了不同$R$取值对实际预测精度的影响，并通过最优实际预测精度确定相应的R值为最优。随机取原数据集20%的数据为训练集，不同R取值下，实际预测精度如图2所示。在相似的预测精度下TFP算法选择最小的R取值，以降低预测算法的时间复杂度。

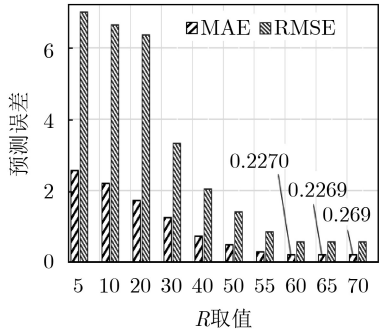

图 2 R取值对预测精度的影响

图2中，在不同的R值下分解张量，随着R的值增加，MAE和RMSE逐渐降低，说明随着R值的增加预测精度逐渐提高。预测精度在R=60, 65以及70时非常接近，并且随着R的增加预测准确度提升有限，因此选择60作为张量的秩，因为这样可以在保证预测准确度的基础上减小算法的复杂度。

3.4.4 其它设置

实验表明：设置$\eta^+ = 1.0001$, $\eta^- = 0.5$，张量分解可达到较快的收敛速度。初始迭代步长δ可随意取一个很小的值，实验中设置其为1E-6。初始迭代步长对收敛速度没有影响，因为增长因子η^+可对迭代步长动态调整，以加快收敛速度。

4 结束语

针对卫星遥测数据的高维结构特点，本文提出了基于张量分解的卫星遥测缺失数据预测算法，以充分考虑数据的整体特征进行准确地缺失数据预测。所提算法通过张量概念对高维遥测数据进行建模；通过张量分解计算其成分矩阵；然后通过成分矩阵对张量进行重构以准确地预测缺失数据；设计了高效的优化算法实现了相关的张量计算；并对算法中的重要参数设置进行严格的理论推导。实验结果表明，所提算法比已有的预测算法表现出更高的预测准确性。

参 考 文 献

[1] 李平, 张路遥, 曹霞, 等. 基于潜在主题的混合上下文推荐算法[J]. 电子与信息学报, 2018, 40(4): 957–963. doi: 10.11999/JEIT170623.

LI Ping, ZHANG Luyao, CAO Xia, *et al*. Hybrid context recommendation algorithm based on latent topic[J]. *Journal of Electronics & Information Technology*, 2018, 40(4): 957–963. doi: 10.11999/JEIT170623.

[2] CHEN I F and LU Chijie. Sales forecasting by combining clustering and machine-learning techniques for computer retailing[J]. *Neural Computing and Applications*, 2017, 28(9): 2633–2647. doi: 10.1007/s00521-016-2215-x.

[3] MA You, WANG Shangguang, HUNG P C K, *et al*. A highly accurate prediction algorithm for unknown Web service QoS values[J]. *IEEE Transactions on Services Computing*, 2016, 9(4): 511–523. doi: 10.1109/TSC.2015.2407877.

[4] 马友, 王尚广, 孙其博, 等. 一种综合考虑主客观权重的Web服务QoS度量算法[J]. 软件学报, 2014, 25(11): 2473–2485. doi: 10.13328/j.cnki.jos.004508.

MA You, WANG Shangguang, SUN Qibo, *et al*. Web service quality metric algorithm employing objective and subjective weight[J]. *Journal of Software*, 2014, 25(11): 2473–2485. doi: 10.13328/j.cnki.jos.004508.

[5] DING Shuai, LI Yeqing, WU Desheng, *et al*. Time-aware cloud service recommendation using similarity-enhanced collaborative filtering and ARIMA model[J]. *Decision Support Systems*, 2018, 107: 103–115. doi: 10.1016/j.dss.2017.12.012.

[6] KUANG Li, YU Long, HUANG Lan, *et al*. A personalized QoS prediction approach for CPS service recommendation based on reputation and location-aware collaborative filtering[J]. *Sensors*, 2018, 18(5): 1556. doi: 10.3390/s18051556.

[7] COLOMO-PALACIOS R, GARCÍA-PEÑALVO F J, STANTCHEV V, *et al*. Towards a social and context-aware mobile recommendation system for tourism[J]. *Pervasive and Mobile Computing*, 2017, 38: 505–515. doi: 10.1016/j.pmcj.2016.03.001.

[8] IGEL C and HÜSKEN M. Improving the Rprop learning algorithm[C]. The 2nd International Symposium on Neural Computation, Berlin, Germany, 2000: 115–121.

[9] GLIGORIJEVIĆ V, PANAGAKIS Y, and ZAFEIRIOU S. Non-negative matrix factorizations for multiplex network analysis[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019, 41(4): 928–940. doi: 10.1109/TPAMI.2018.2821146.

[10] MA Wenping, WU Yue, and GONG Maoguo. Local probabilistic matrix factorization for personal recommendation[C]. The 13th International Conference on Computational Intelligence and Security, Hong Kong, China, 2017: 97–101. doi: 10.1109/CIS.2017.00029

[11] SHAO Lingshuang, ZHANG Jing, WEI Yong, *et al*. Personalized QoS prediction for web services via collaborative filtering[C]. The IEEE International Conference on Web Services, Salt Lake City, USA, 2007: 439–446. doi: 10.1109/ICWS.2007.140.

[12] SARWAR B, KARYPIS G, KONSTAN J, *et al*. Item-based collaborative filtering recommendation algorithms[C]. The 10th International Conference on World Wide Web, Hong Kong, China, 2001: 285–295. doi: 10.1145/371920.372071.

[13] KANG M G and KATSAGGELOS A K. General choice of the regularization functional in regularized image restoration[J]. *IEEE Transactions on Image Processing*, 1995, 4(5): 594–602. doi: 10.1109/83.382494.

[14] KATSAGGELOS A K, BIEMOND J, SCHAFER R W, *et al*. A regularized iterative image restoration algorithm[J]. *IEEE Transactions on Signal Processing*, 1991, 39(4): 914–929. doi: 10.1109/78.80914.

[15] MILLER K. Least squares methods for ill-posed problems with a prescribed bound[J]. *SIAM Journal on Mathematical Analysis*, 1970, 1(1): 52–74. doi: 10.1137/0501006.

[16] KOLDA T G and BADER B W. Tensor decompositions and applications[J]. *SIAM Review*, 2009, 51(3): 455–500. doi: 10.1137/07070111X.

[17] COMON P, TEN BERGE J M, DE LATHAUWER L, *et al*. Generic and typical ranks of multi-way arrays[J]. *Linear Algebra and Its Applications*, 2009, 430(11/12): 2997–3007. doi: 10.1016/j.laa.2009.01.014.

马　友：男，1982年生，副研究员，主要研究方向为服务推荐与机器学习.

贾树泽：男，1982年生，高级工程师，主要研究方向为卫星故障诊断.

赵现纲：男，1979年生，研究员，主要研究方向为卫星通讯技术.

冯小虎：男，1973年生，研究员，主要研究方向为航天器精细化管理.

范存群：男，1986年生，高级工程师，主要研究方向为卫星资料同化.

朱爱军：男，1970年生，研究员，主要研究方向为卫星系统工程.