5th International Workshop on Climate Informatics
September 24-25, 2015
CI 2015
Hosted by the National Center for Atmospheric Research in Boulder, CO

ACCELERATED LOW-RANK TENSOR ONLINE LEARNING FOR MULTI-MODEL ENSEMBLE

Rose Yu , Dehua Cheng, Yan Liu

***Abstract*—Low-rank tensor learning has many applications in machine learning. A series of batch learning algorithms have achieved great successes. However, in climate data analysis, we are confronted with large-scale tensor streams, which pose significant challenges to existing solutions. In this paper, we propose an accelerated low-rank tensor online learning algorithm (ALTO) and apply our method to climate multi-model ensemble task. Experiment results show that our method achieves comparable predictive accuracy with significant speed-up.**

I. MOTIVATION

Large-scale climate data come in streams. Batch learning algorithms would suffer from computational bottleneck, especially facing the challenge of short response time. Therefore, effective and fast online learning algorithms are a must for enabling real-time large-scale climate analysis.

The multi-model ensemble problem arises in climate modeling. In the past decades, numerous climate models have been developed to generate large simulation data sets of future climate projections [1]. Physical models share similar representations of the ocean-atmosphere and land-ice processes but have different parameter uncertainty levels. Learning the correlation between model simulations and the observations can help quantify uncertainty in climate models and prompt the design of more accurate models. The multi-model ensemble task aims to learn such correlation by combining climate model outputs into a more accurate description of the observations.

Existing methods such as model coupling [2] and super model [3] have been studied. Unfortunately, learning the parameters of those statistical models is computationally expensive. In this paper, we represent the climate data as tensors, which are multilinear generalization of the matrices. We formulate the problem as an online low-rank tensor learning task and develop a novel framework, namely the **A**ccelerated **L**ow-rank **T**ensor **O**nline Learning (ALTO) algorithm to address the problem. Experimental results show that our algorithm achieve competitive accuracy with significant speed-up.

Corresponding author: Rose Yu, University of Southern California, CA US qiyu@usc.edu

II. METHOD

Suppose we have gathered the model simulation outputs from S models of M climate variables in P locations over time period T. At the same time, we are given access to the actual observations of the same variables, locations and time. As in the forecasting problem setting, we can represent the observation measurements using a three-mode tensor $\mathcal{X} \in \mathbb{R}^{P\times T\times M}$. Similarly, we encode the model outputs with a four-mode tensor $\mathcal{Y} \in \mathbb{R}^{P\times T\times M\times S}$. Those model outputs serve as "experts" for the climate prediction. Incorporating those experts' advice can reduce the uncertainty of the forecasts.

We start with a simple linear model as $\mathcal{X} = \mathcal{W}_{:,:,m}\mathbf{Y}_{t,m}$, where $\mathbf{Y}_{t,m} = [\mathcal{Y}_{:,t,m,1}^\top, \ldots, \mathcal{Y}_{:,t,m,S}^\top]^\top$ denotes the concatenation of S model outputs at time t for variable m, and $\mathcal{W} \in \mathbb{R}^{P\times PS\times M}$ characterizes the "importance" of various models in climate predictions. We formulate the multi-model ensemble task as follows:

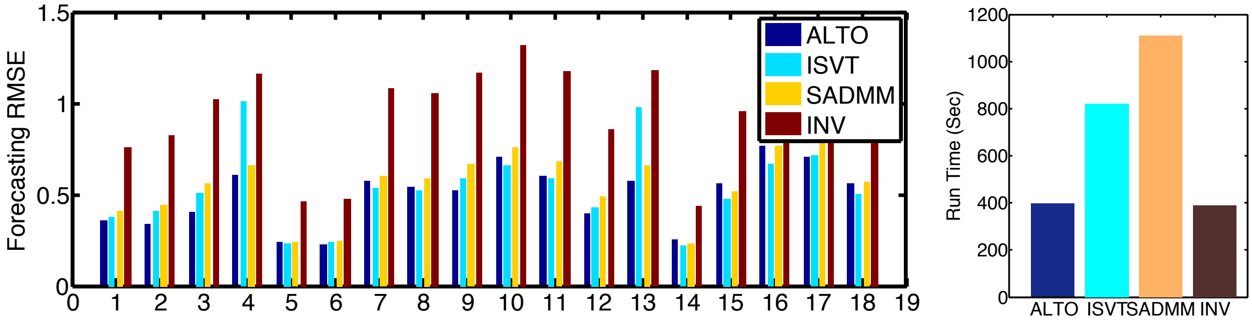
$$\widehat{\mathcal{W}} = \underset{\mathcal{W}}{\text{argmin}} \left\{ \|\widehat{\mathcal{X}} - \mathcal{X}\|_F^2 + \mu \sum_{m=1}^{M} \text{tr}(\widehat{\mathcal{X}}_{:,:,m}^\top \mathbf{L} \widehat{\mathcal{X}}_{:,:,m}) \right\}$$
$$\text{s.t. } \widehat{\mathcal{X}}_{:,t,m} = \mathcal{W}_{:,:,m}\mathbf{Y}_{t,m}, \ \sum_{n=1}^{N} \text{rank}(\mathcal{W}_{(n)}) \le R$$

Where the Laplacian matrix $\mathbf{L}$ accounts for the spatial proximity of observations. With change of variables, both the online forecasting and the multi-model ensemble problem can be reformulated into the low-rank tensor learning framework in Equation 1.

$$\widehat{\mathcal{W}} = \underset{\mathcal{W}}{\text{argmin}} \left\{ \textstyle\sum_{t=1}^{T} \sum_{m=1}^{M} \|\mathcal{W}_{:,:,m}\mathcal{Z}_{:,t,m} - \mathcal{X}_{:,t,m}\|_F^2 \right\}$$
$$\text{s.t.} \quad \text{rank}(\mathcal{W}) \le R. \tag{1}$$

We propose to solves the above problem via a simple two-step approach: (1) solving the unconstrained optimization problem given the new data, (2) updating the solution with the low-rank constraint, i.e. projecting the solution to the space of low-rank tensors.

In Step 1, the unconstrained optimization problem is equivalent to an ordinary linear regression, which has closed-form solution. In Step 2, we project the solution from Step 1 to the low-rank tensor space using ALTO in



(a) RMSE

(b) Runtime(Sec)

Fig. 1. Per variable forecasting RMSE for 18 variables (a) and overall run time (b) comparison of multi-model ensemble for ALTO and baselines using 90 % training data, with 7 different models over 20 years.

Algorithm 1 **A**ccelerated **L**ow rank **T**ensor **On**line Learning (ALTO)

$[\mathcal{W}^{\text{new}}, \mathbf{U}^{\text{new}}] = \text{ALTO}(\mathcal{W}, \mathbf{U}, R, K)$:

Input: original tensor $\mathcal{W}$ and projection matrices $\mathbf{U}_i, i = 1, 2, 3$. rank R, augmentation factor K

Output: updated tensor $\mathcal{W}^{\text{new}}$ and projection matrices $\mathbf{U}_i^{\text{new}}, i = 1, 2, 3$.

1 Augment, orthogonalize and normalize $\mathbf{U}_i, i = 1, 2, 3$ to $\mathbf{V}_i, i = 1, 2, 3$ with $R + K$ columns.

2 Project $\mathcal{W} \to \mathcal{S}' = \mathcal{W} \times_1 \mathbf{V}_1^\top \times_2 \mathbf{V}_2^\top \times_3 \mathbf{V}_3^\top$.

3 Find the rank-R approximation to $\mathcal{S}'$ with TSM: $\text{TSM}(\mathcal{S}', R) = \mathcal{S} \times_1 \mathbf{V}_1' \times_2 \mathbf{V}_2' \times_3 \mathbf{V}_3'$.

4 Return $\mathbf{U}_i^{\text{new}} = \mathbf{V}_i \mathbf{V}_i', i = 1, 2, 3$ and $\mathcal{W}^{\text{new}} = \mathcal{S} \times_1 \mathbf{U}_1^{\text{new}} \times_2 \mathbf{U}_2^{\text{new}} \times_3 \mathbf{U}_3^{\text{new}}$.

Algorithm 1, where *low-rank Tensor Sequential Mapping* (TSM) is a procedure that sequentially maps the unfolded tensor into the rank-R subspace.

III. EVALUATION

We compare our algorithm with the following baselines on the multi-model ensemble task:

- INV: closed form solution of *Exact Update* for VAR model without low-rank constraint.
- SADMM: stochastic alternating direction method of multipliers [5] adapted for tensor nuclear norm regularizer.
- ISVT: iterative singular value thresholding [6] generalized to tensor mode-n rank constraint.
- GREEDY: greedy sequential rank-1 approximation [4] for low-rank tensor learning in batch setting.

For observation series, we collect the monthly measurements from NCEP-DOE Reanalysis 2 [7]. For model outputs, 7 different model simulation data are taken from the World Climate Research Programme's (WCRP's) CMIP3 multi-model dataset and processed with CDO software. [1] We align the variables of observation series with the model output series. 19 variables are selected with 252 time points from 1979 to 1999.

We use model outputs to predict the observation measurements. 90% of the time series are used for online training. We examine the forecasting error for each variable separately using the learned model. Figure 1 shows the forecasting RMSE for each variables and overall run time in second. ALTO is not only more accurate but also much faster than baselines.

IV. ACKNOWLEDGMENT

The research was sponsored by NSF research grants IIS-1134990, IIS-1254206 and U.S. Defense Advanced Research Projects Agency (DARPA) under Social Media in Strategic Communication (SMISC) program, Agreement Number W911NF-12-1-0034.

REFERENCES

[1] C. Tebaldi and R. Knutti, "The use of the multi-model ensemble in probabilistic climate projections," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 365, no. 1857, pp. 2053–2075, 2007.

[2] L. Van den Berge, F. Selten, W. Wiegerinck, and G. Duane, "A multi-model ensemble method that combines imperfect models through learning," *Earth System Dynamics*, vol. 2, pp. 161–177, 2011.

[3] W. Wiegerinck and F. Selten, "Supermodeling: Combining imperfect models through learning," 2011.

[4] M. T. Bahadori, Q. R. Yu, and Y. Liu, "Fast multivariate spatio-temporal analysis via low rank tensor learning," in *Advances in Neural Information Processing Systems*, pp. 3491–3499, 2014.

[5] H. Ouyang, N. He, L. Tran, and A. Gray, "Stochastic alternating direction method of multipliers," in *Proceedings of the 30th International Conference on Machine Learning*, pp. 80–88, 2013.

[6] P. Jain, R. Meka, and I. S. Dhillon, "Guaranteed rank minimization via singular value projection," in *Advances in Neural Information Processing Systems*, pp. 937–945, 2010.

[7] P. W. Jones, "First-and second-order conservative remapping schemes for grids in spherical coordinates," *Monthly Weather Review*, vol. 127, no. 9, pp. 2204–2210, 1999.

[1] CDO 2015: Climate Data Operators.
Available at: http://www.mpimet.mpg.de/cdo