Tensor Regression

Other titles in Foundations and Trends® in Machine Learning

Minimum-Distortion Embedding
Akshay Agrawal, Alnur Ali and Stephen Boyd
ISBN: 978-1-68083-888-6

Graph Kernels: State-of-the-Art and Future Challenges
Karsten Borgwardt, Elisabetta Ghisu, Felipe Llinares-López, Leslie O'Bray and Bastian Rieck
ISBN: 978-1-68083-770-4

Data Analytics on Graphs Part III: Machine Learning on Graphs, from Graph Topology to Applications
Ljubiša Stanković, Danilo Mandic, Miloš Daković, Miloš Brajović, Bruno Scalzo, Shengxi Li and Anthony G. Constantinides
ISBN: 978-1-68083-982-16

Data Analytics on Graphs Part II: Signals on Graphso
Ljubiša Stanković, Danilo Mandic, Miloš Daković, Miloš Brajović, Bruno Scalzo, Shengxi Li and Anthony G. Constantinides
ISBN: 978-1-68083-982-1

Data Analytics on Graphs Part I: Graphs and Spectra on Graphs
Ljubiša Stanković, Danilo Mandic, Miloš Daković, Miloš Brajović, Bruno Scalzo, Shengxi Li and Anthony G. Constantinides
ISBN: 978-1-68083-982-1

Tensor Regression

Jiani Liu
University of Electronic Science and Technology of China
jianiliu@std.uestc.edu.cn

Ce Zhu
University of Electronic Science and Technology of China
eczhu@uestc.edu.cn

Zhen Long
University of Electronic Science and Technology of China
zhenlong@std.uestc.edu.cn

Yipeng Liu
University of Electronic Science and Technology of China
yipengliu@uestc.edu.cn

the essence of knowledge
Boston — Delft

Foundations and Trends® in Machine Learning

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
United States
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is

J. Liu, C. Zhu, Z. Long and Y. Liu. *Tensor Regression*. Foundations and Trends® in Machine Learning, vol. 14, no. 4, pp. 379–565, 2021.

ISBN: 978-1-68083-887-9

Foundations and Trends® in Machine Learning
Volume 14, Issue 4, 2021

Editorial Scope

Topics

Foundations and Trends® in Machine Learning publishes survey and tutorial articles in the following topics:

- Adaptive control and signal processing
- Applications and case studies
- Behavioral, cognitive and neural learning
- Bayesian learning
- Classification and prediction
- Clustering
- Data mining
- Dimensionality reduction
- Evaluation
- Game theoretic learning
- Graphical models
- Independent component analysis
- Inductive logic programming
- Kernel methods
- Markov chain Monte Carlo
- Model choice
- Nonparametric methods
- Online learning
- Optimization
- Reinforcement learning
- Relational learning
- Robustness
- Spectral methods
- Statistical learning theory
- Variational inference
- Visualization

Information for Librarians

Foundations and Trends® in Machine Learning, 2021, Volume 14, 6 issues. ISSN paper version 1935-8237. ISSN online version 1935-8245. Also available as a combined paper and online subscription.

Contents

Tensor Regression

Jiani Liu[1], Ce Zhu[1], Zhen Long[1] and Yipeng Liu[1]

[1]*School of Information and Communication Engineering, University of Electronic Science and Technology of China (UESTC), China, jianiliu@std.uestc.edu.cn, eczhu@uestc.edu.cn, zhenlong@std.uestc.edu.cn, yipengliu@uestc.edu.cn*

ABSTRACT

The presence of multidirectional correlations in emerging multidimensional data poses a challenge to traditional regression modeling methods. Traditional modeling methods based on matrix or vector, for example, not only overlook the data's multidimensional information and lower model performance, but also add additional computations and storage requirements. Driven by the recent advances in applied mathematics, tensor regression has been widely used and proven effective in many fields, such as sociology, climatology, geography, economics, computer vision, chemometrics, and neuroscience. Tensor regression can explore multidirectional relatedness, reduce the number of model parameters and improve model robustness and efficiency. It is timely and valuable to summarize the developments of tensor regression in recent years and discuss promising future directions, which will help accelerate the research process of tensor regression, broaden the research direction, and provide tutorials for researchers interested in high dimensional regression tasks.

The fundamentals, motivations, popular algorithms, related applications, available datasets, and software resources for

Jiani Liu, Ce Zhu, Zhen Long and Yipeng Liu (2021), "Tensor Regression", Foundations and Trends® in Machine Learning: Vol. 14, No. 4, pp 379–565. DOI: 10.1561/2200000087.

tensor regression are all covered in this monograph. The first part focuses on the key concepts for tensor regression, mainly analyzing existing tensor regression algorithms from the perspective of regression families. Meanwhile, the adopted low rank tensor representations and optimization frameworks are also summarized. In addition, several extensions in online learning and sketching are described. The second part covers related applications, widely used public datasets and software resources, as well as some real-world examples, such as multitask learning, spatiotemporal learning, human motion analysis, facial image analysis, neuroimaging analysis (disease diagnosis, neuron decoding, brain activation, and connectivity analysis) and chemometrics. This survey can be used as a basic reference in tensor-regression-related fields and assist readers in efficiently dealing with high dimensional regression tasks.

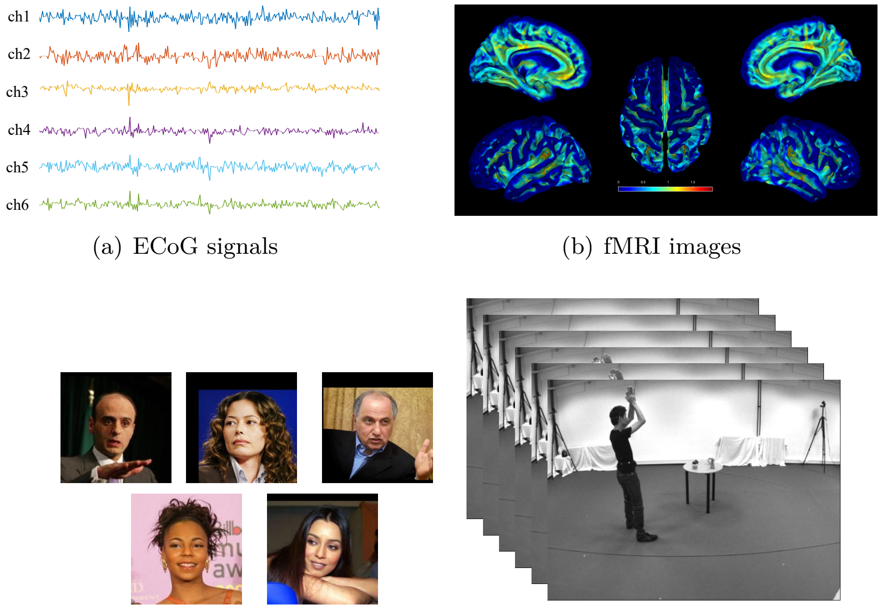
1

Introduction

Regression analysis is a key area of interest in the field of data analysis and machine learning. Regression analysis is devoted to exploring the dependencies between variables, thereby exploring the objective causal relationship of things and further predicting the possible state of things in the future.

Every day, a considerable amount of high dimensional data is collected and stored due to the rapid growth of digital imaging and sensing technologies. For example, numerous neuroimaging technologies, such as electroencephalography (EEG), electrocorticography (ECoG), and functional magnetic resonance imaging (fMRI), generate a lot of multidimensional data. Color images, hyperspectral images, or video sequences in computer vision can also be considered as three-dimensional or four-dimensional tensors. Data in climatology is usually organized in three modes: location, time, and variable type. Social network data also has many indexes, such as user, destination, and time. Fig. 1.1 provides some examples of high dimensional datasets, which are selected from some public datasets.

The emergence of high dimensional data has brought challenges to traditional data representation methods. Tensors, as high order ex-



(a) ECoG signals

(b) fMRI images

(c) Facial images

(d) Video sequences

Figure 1.1: Some examples of high dimensional datasets.

tensions of vectors, are considered as natural representations of high order data (Kolda and Bader, 2009; Cichocki *et al.*, 2015; Sidiropoulos *et al.*, 2017). Exploring the internal structure information of tensors can help people better represent high order data with fewer parameters, extract and interpret their attributes, and achieve data compression, storage, processing, and analysis. Tensor-based learning approaches have been proven to be more effective than classic vector-based methods, and have gotten a lot of attention in recent years, such as tensor completion (Long *et al.*, 2019), tensor principal component analysis (Lu *et al.*, 2016; Feng *et al.*, 2020), tensor sparse representation (Qi *et al.*, 2016; Xie *et al.*, 2017), tensor subspace learning (Lu *et al.*, 2011; Makantasis *et al.*, 2019), tensor clustering (Sun and Li, 2019; Poullis, 2019), tensor regression and classification (Tao *et al.*, 2005; Stoudenmire and Schwab, 2016; Wimalawarne *et al.*, 2016).

For regression tasks with high dimensional predictors or responses, how to effectively explore the multidirectional correlation of these high dimensional data has become an important research topic in the field of image processing and machine learning (Zhou *et al.*, 2013). Most classical regression statistical models are modeled based on vectors as regression coefficients and are not suitable for high dimensional regression problems. There are two main disadvantages for simply employing the traditional regression methods. On the one hand, using traditional regression methods requires first performing vectorization operations on the multiway data, which will ignore the inherent multiway structural information contained in the high dimensional data, resulting in a degradation in terms of model performance. On the other hand, a huge vector-based model will require a large number of parameters, suffering from storage and computational burden and undesired numerical instability.

The objective of this review is to provide a systematic study and analysis of tensor-based regression models and their applications in recent years. Here we give some examples with multidimensional inputs or outputs that will be discussed in this review.

Example 1: Weather forecasting

Weather forecasting is a common task to predict the state of weather at a given space based on historical recordings. The data for this example comes from the U.S. Historical Climatology Network (USHCN). It records the monthly measurements of four climate variables including maximum, minimum, and average temperatures over a month, and total monthly precipitations of approximately 1200 USHCN stations from 1915 to 2014. Table 1.1 lists the recordings for the four variables of some selected stations at a specific time. It records the temperature in degrees Fahrenheit and precipitation in hundreds of inches.

The objective is to design a regressor that could predict the possible value of these four variables at timepoint t based on the previous recordings at $t-1, \cdots, t-K$. Let $\mathcal{X}(:, t, m) \in \mathbb{R}^P$ represent the recording of m-th variable at time t of all the stations, then we aim to construct

Table 1.1: The monthly average of four climate variables for some selected stations in U.S. We have listed the recording for January 2014.

Stations	Maximum temperature (F)	Minimum temperature (F)	Average temperature (F)	Total monthly precipitations
BREWTON 3 SSE	56.3	32.0	44.2	3.15
FAIRHOPE 2 NE	55.5	30.0	42.7	2.34
GAINESVILLE LOCK	50.3	24.3	37.3	1.90
GREENSBORO	50.0	25.2	37.6	1.47
HIGHLAND HOME	50.9	23.0	36.9	3.35
MUSCLE SHOALS AP	47.8	22.9	35.3	1.71
SAINT BERNARD	44.3	20.1	32.2	2.99
···	···	···	···	···
WORLAND	31.1	5.1	18.1	0.38
YELLOWSTONE PK MAMMOTH	34.3	15.8	25.0	0.62

a data generation process as follows

$$\mathcal{X}(:,t,m) = \sum_{k=1}^{K} \mathcal{B}_k(:,:,m)\mathcal{X}(:,t-k,m) + \mathcal{E}(:,t,m) \tag{1.1}$$

where $\mathcal{B}_k \in \mathbb{R}^{P\times P\times M}$, $m = 1,\cdots,M$, $M = 4$ denotes the number of climate variables, $\mathcal{E}$ is the bias. In order to characterize both the spatial and temporal correlations of the climate data, the low rank constraint is commonly used over the coefficient tensor $\mathcal{B} \in \mathbb{R}^{P\times KP\times M}$ which is obtained by putting all the $\{\mathcal{B}_k\}_{k=1}^{K}$ together along the second mode. We have discussed a number of different methods for tackling this learning problem in this monograph.

Example 2: Age estimation

Estimating age automatically from facial images or brain images has attracted a lot of attention in recent years. It is a challenging problem since the impact factors for the aging process are complex. Meanwhile, the facial images and brain images are indexed by multiple dimensions, commonly represented as a 3rd order tensor. Therefore, a lot of parameters are needed to predict the age as in the following model

$$y_n = \langle \mathcal{X}_n, \mathcal{B} \rangle + \epsilon_n, \tag{1.2}$$

where $\mathcal{X}$ is the employed brain images like MRI or fMRI images for n-th sample, y_n is the corresponding sample age, the coefficient tensor $\mathcal{B}$ is in the same size as the predictor, ϵ_n is the bias for n-th sample.

Here we consider the UK biobank dataset for the age estimation task. The T1-weighted 3D structured MRI images and corresponding age labels are exploited to train the regression model. The MRI images are all sized at $182 \times 218 \times 182$. Thus, the parameters needed for model (1.2) will reach $7,000,000$, which makes the learning problem intractable.

In order to control model complexity and explore multidirectional relatedness, the low rank constraint is naturally considered, namely low rank tensor regression, which we will discuss later in this review.

Example 3: A study for autism spectrum disorder

One major issue in neuroimaging analysis is to infer clinical assessments or human traits from neuroimages, such as the age estimation from the fMRI images described above. In addition to this, through modeling the data from a reverse perspective, we can obtain a tensor response regression task that aims to generate corresponding brain images in response to changes in disease state or age. In this way, it is easy to find out the activity pattern between different groups of subjects.

Specifically, here we consider the data from the Autism Brain Imaging Data Exchange (ABIDE) for autism spectrum disorder (ASD). ASD is a series of complex neurodevelopmental disorders that can affect children's social behavior and communication abilities. Studying the different patterns of information processing in the brains of people with ASD and normal children can better help the diagnosis and treatment of ASD. In other words, it is an important research topic to compare and analyze the neural images of normal children and children with ASD through existing data, and to find the brain regions or activity patterns that are most likely to distinguish the two groups. We could view this task as a tensor response regression problem, with the fMRI images being the response and three predictors including the ASD status, age, and sex.

Since there is a high dimensionality in the response, like the fractional amplitude of low-frequency fluctuations (fALFF) data sized at $91 \times 109 \times 91$ in ABIDE, the number of model parameters required will be very large, far exceeding the number of samples, which makes the problem difficult to solve. Therefore, how to explore the multidirectional correlation between data, and how to accurately reduce the dimensionality while ensuring that the model fitting error is small, become very important. This is also an issue that this review aims to deal with.

Contributions

The main contributions of this review are summarized as follows:

- This is the first thorough overview of the fundamentals, motivations, popular algorithms, strategies for efficient implementation, related applications, available datasets, and software resources for tensor-based regression analysis. There is presently no comprehensive overview for tensor-based regression methods. Only some reviews for tensor-based analysis (Sun *et al.*, n.d.; Cichocki *et al.*, 2017) include a subsection for regression analysis, covering only a part of popular algorithms. And there is no classification and discussion of existing methods from multiple perspectives, nor a systematic summary of algorithms, applications, datasets, etc.

- This review groups and illustrates the existing tensor-based regression methods mainly from the perspective of different regression families, such as simple linear tensor regression, generalized tensor regression, penalized tensor regression, Bayesian tensor regression, quantile tensor regression, projection based tensor regression, kernelized tensor regression, tensor Gaussian process regression, tensor additive models, random forest based tensor regression and deep tensor regression. Meanwhile, a discussion of these methods is also given in terms of modeling (tensor-on-vector regression, vector-on-tensor regression, tensor-on-tensor regression), tensor representations (CP, Tucker, t-SVD, TT, TR), and optimization frameworks (ADMM, AM, GD).

- Another highlight of this overview is the use of some real cases in both Chapter 1 and 7 to help readers better understand the application background of tensor regression analysis, such as the forecasting task of spatio-temporal data, human motion trajectories reconstruction from corresponding image sequences, age estimation and disease diagnosis from brain images, a study of the difference between the neural images of children with ASD and normal children, and decoding electroencephalogram signals.

Who should read this review

This review intends to give non-specialists and students interested in tensor regression analysis a decent starting point. The basics, core ideas, and theoretical characteristics of most tensor-based regression methods are covered, but it may be beneficial for readers to have some background in convex optimization, machine learning, and statistics.

We attempt to construct a more comprehensive overview covering most existing tensor regression methods. Moreover, we may pay more attention to the core ideas of the algorithms rather than the specific details and theoretical properties of the algorithms. This is our focus because we want to help readers understand the possible solutions to tensor regression tasks and the main ideas, advantages, and disadvantages of these solutions conceptually rather than mathematically.

In addition, those who are interested in starting related works can also read this review to learn how to use existing tensor-based regression methods to solve specific regression tasks with multiway data, what datasets can be selected, and what software packages are available to start related work as soon as possible.

Organization of this review

The remaining part of the review proceeds as follows: Chapter 2 provides the notations, basic tensor-related operations, and some popular decomposition methods used in this review. The traditional regression methods are introduced in Chapter 3 and extended into the tensor field in Chapter 4 and Chapter 5. Meanwhile, Chapter 4 gives a comprehensive illustration, discussion, comparison, and summary of existing popular linear tensor regression methods. Chapter 5 illustrates the key concepts and algorithms of nonlinear tensor regression models. Then some strategies for efficient implementation, including sketching and online learning, are discussed in Chapter 6. Chapter 7 summarizes the common application fields and gives some practical examples in order to guide the research on tensor regression and validate the effectiveness of some tensor regression approaches. Finally, Chapter 8 concludes the open-source software, and Chapter 9 draws some conclusions and makes recommendations for possible future research works.

References

Acar, E., C. A. Bingol, H. Bingol, R. Bro, and B. Yener. (2007). "Seizure recognition on epilepsy feature tensor". In: *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE. 4273–4276.

Acar, E., Y. Levin-Schwartz, V. D. Calhoun, and T. Adali. (2017). "Tensor-based fusion of EEG and FMRI to understand neurological changes in schizophrenia". In: *2017 IEEE International Symposium on Circuits and Systems (ISCAS)*. 1–4. DOI: 10.1109/ISCAS.2017.8050303.

Ahmed, T., H. Raja, and W. U. Bajwa. (2019). "Tensor Regression Using Low-rank and Sparse Tucker Decompositions". *arXiv preprint arXiv:1911.03725*.

Amigo, J. M., J. Coello, and S. Maspoch. (2005). "Three-way partial least-squares regression for the simultaneous kinetic-enzymatic determination of xanthine and hypoxanthine in human urine". *Analytical and bioanalytical chemistry*. 382(6): 1380–1388.

Arbaiy, N., H. A. Rahman, M. Z. M. Salikon, and P. C. Lin. (2018). "Workplace Safety Risk Assessment Model Based on Fuzzy Regression". *Advanced Science Letters*. 24(3): 1656–1659.

Argyriou, A., T. Evgeniou, and M. Pontil. (2008). "Convex multi-task feature learning". *Machine Learning*. 73(3): 243–272.

Ba, L. J. and R. Caruana. (2013). "Do deep nets really need to be deep?" *arXiv preprint arXiv:1312.6184*.

Ba, S. O. and J.-M. Odobez. (2005). "Evaluation of multiple cues head pose tracking algorithm in indoor environments". *Tech. rep.*

Bahadori, M. T., Q. R. Yu, and Y. Liu. (2014). "Fast multivariate spatio-temporal analysis via low rank tensor learning". In: *Advances in neural information processing systems*. 3491–3499.

Baltrušaitis, T., P. Robinson, and L.-P. Morency. (2012). "3D constrained local model for rigid and non-rigid facial tracking". In: *2012 IEEE conference on computer vision and pattern recognition*. IEEE. 2610–2617.

Bates, D. M. and D. G. Watts. (1988). *Nonlinear regression analysis and its applications*. Vol. 2. Wiley New York.

Beck, N., G. King, and L. Zeng. (2000). "Improving quantitative studies of international conflict: A conjecture". *American Political science review*: 21–35.

Bell, P., M. J. Gales, T. Hain, J. Kilgour, P. Lanchantin, X. Liu, A. McParland, S. Renals, O. Saz, M. Wester, *et al.* (2015). "The MGB challenge: Evaluating multi-genre broadcast media recognition". In: *2015 IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU)*. IEEE. 687–693.

Bergant, K. and L. Kajfež-Bogataj. (2005). "N–PLS regression as empirical downscaling tool in climate change studies". *Theoretical and Applied Climatology*. 81(1): 11–23.

Beutel, A., P. P. Talukdar, A. Kumar, C. Faloutsos, E. E. Papalexakis, and E. P. Xing. (2014). "Flexifact: Scalable flexible factorization of coupled tensors on hadoop". In: *Proceedings of the 2014 SIAM International Conference on Data Mining*. SIAM. 109–117.

Billio, M., R. Casarin, S. Kaufmann, and M. Iacopini. (2018). "Bayesian Dynamic Tensor Regression". *University Ca'Foscari of Venice, Dept. of Economics Research Paper Series No*. 13.

Blagoveschensky, P. and A. H. Phan. (2020). "Deep convolutional tensor network". *arXiv preprint arXiv:2005.14506*.

Boltia, S. A., A. S. Fayed, A. Musaed, and M. A. Hegazy. (2019). "Bilinear and trilinear algorithms utilizing full and selected variables for resolution and quantitation of four components with overlapped spectral signals in bulk and syrup dosage form". *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*. 222: 117219.

Boqué, R. and A. K. Smilde. (1999). "Monitoring and diagnosing batch processes with multiway covariates regression models". *AIChE Journal*. 45(7): 1504–1520.

Breiman, L. (2001). "Random forests". *Machine learning*. 45(1): 5–32.

Bro, R. (1996). "Multiway calibration. multilinear pls". *Journal of chemometrics*. 10(1): 47–61.

Bro, R. (1997). "PARAFAC. Tutorial and applications". *Chemometrics and intelligent laboratory systems*. 38(2): 149–171.

Bro, R. (1999). "Exploratory study of sugar production using fluorescence spectroscopy and multi-way analysis". *Chemometrics and Intelligent Laboratory Systems*. 46(2): 133–147.

Bro, R. (2006). "Review on multiway analysis in chemistry—2000–2005". *Critical reviews in analytical chemistry*. 36(3-4): 279–293.

Bro, R., Å. Rinnan, and N. K. M. Faber. (2005). "Standard error of prediction for multilinear PLS: 2. Practical implementation in fluorescence spectroscopy". *Chemometrics and Intelligent Laboratory Systems*. 75(1): 69–76.

Bro, R., A. K. Smilde, and S. de Jong. (2001). "On the difference between low-rank and subspace approximation: improved model for multilinear PLS regression". *Chemometrics and Intelligent Laboratory Systems*. 58(1): 3–13.

Cai, D., X. He, J.-R. Wen, J. Han, and W.-Y. Ma. (2006). "Support tensor machines for text categorization". *Tech. rep.*

Camarrone, F. and M. M. Van Hulle. (2016). "Fully Bayesian tensor-based regression". In: *2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP)*. IEEE. 1–6.

Camarrone, F. and M. M. Van Hulle. (2018). "Fast multiway partial least squares regression". *IEEE Transactions on Biomedical Engineering*. 66(2): 433–443.

Cao, X. and G. Rabusseau. (2017). "Tensor regression networks with various low-rank tensor approximations". *arXiv preprint arXiv:1712.09520*.

Carroll, J. D. and J.-J. Chang. (1970). "Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition". *Psychometrika*. 35(3): 283–319.

Chen, C., K. Batselier, W. Yu, and N. Wong. (2020). "Kernelized support tensor train machines". *arXiv preprint arXiv:2001.00360*.

Chen, H., G. Raskutti, and M. Yuan. (2019). "Non-convex projected gradient descent for generalized low-rank tensor regression". *The Journal of Machine Learning Research*. 20(1): 172–208.

Chen, J. and J.-H. Yen. (2003). "Three-way data analysis with time lagged window for on-line batch process monitoring". *Korean Journal of Chemical Engineering*. 20(6): 1000–1011.

Chen, Y., X. Jin, B. Kang, J. Feng, and S. Yan. (2018). "Sharing Residual Units Through Collective Tensor Factorization To Improve Deep Neural Networks." In: *IJCAI*. 635–641.

Cheng, J.-H. and D.-W. Sun. (2017). "Partial least squares regression (PLSR) applied to NIR and HSI spectral data modeling to predict chemical properties of fish muscle". *Food engineering reviews*. 9(1): 36–49.

Chien, J.-T. and Y.-T. Bao. (2017). "Tensor-factorized neural networks". *IEEE transactions on neural networks and learning systems*. 29(5): 1998–2011.

Chow, E., D. Ebrahimi, J. J. Gooding, and D. B. Hibbert. (2006). "Application of N-PLS calibration to the simultaneous determination of Cu 2+, Cd 2+ and Pb 2+ using peptide modified electrochemical sensors". *Analyst*. 131(9): 1051–1057.

Cichocki, A., D. Mandic, L. De Lathauwer, G. Zhou, Q. Zhao, C. Caiafa, and H. A. Phan. (2015). "Tensor decompositions for signal processing applications: From two-way to multiway component analysis". *IEEE Signal Processing Magazine*. 32(2): 145–163.

Cichocki, A., N. Lee, I. Oseledets, A.-H. Phan, Q. Zhao, D. P. Mandic, *et al.* (2016). "Tensor networks for dimensionality reduction and large-scale optimization: Part 1 low-rank tensor decompositions". *Foundations and Trends® in Machine Learning*. 9(4-5): 249–429.

Cichocki, A., A.-H. Phan, Q. Zhao, N. Lee, I. V. Oseledets, M. Sugiyama, and D. Mandic. (2017). "Tensor networks for dimensionality reduction and large-scale optimizations. part 2 applications and future perspectives". *Foundations and Trends® in Machine Learning*. 9(6): 431–673.

Clarkson, K. L. and D. P. Woodruff. (2017). "Low-rank approximation and regression in input sparsity time". *Journal of the ACM (JACM)*. 63(6): 1–45.

Cohen, M. B., S. Elder, C. Musco, C. Musco, and M. Persu. (2015). "Dimensionality reduction for k-means clustering and low rank approximation". In: *Proceedings of the forty-seventh annual ACM symposium on Theory of computing*. 163–172.

Cohen, N., O. Sharir, Y. Levine, R. Tamari, D. Yakira, and A. Shashua. (2017). "Analysis and design of convolutional networks via hierarchical tensor decompositions". *arXiv preprint arXiv:1705.02302*.

Das, S., S. Roy, and R. Sambasivan. (2018). "Fast Gaussian process regression for big data". *Big data research*. 14: 12–26.

Daubechies, I., R. DeVore, M. Fornasier, and C. S. Güntürk. (2010). "Iteratively reweighted least squares minimization for sparse recovery". *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*. 63(1): 1–38.

De Almeida, A. L. and A. Y. Kibangou. (2014). "Distributed large-scale tensor decomposition". In: *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE. 26–30.

De Jong, S. and C. J. Ter Braak. (1994). "Comments on the PLS kernel algorithm". *Journal of chemometrics*. 8(2): 169–174.

De Lathauwer, L. (2008). "Decompositions of a higher-order tensor in block terms—Part II: Definitions and uniqueness". *SIAM Journal on Matrix Analysis and Applications*. 30(3): 1033–1066.

De Lathauwer, L., B. De Moor, and J. Vandewalle. (2000). "A multilinear singular value decomposition". *SIAM journal on Matrix Analysis and Applications*. 21(4): 1253–1278.

Diakonikolas, I., W. Kong, and A. Stewart. (2019). "Efficient algorithms and lower bounds for robust linear regression". In: *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms*. SIAM. 2745–2754.

Efron, B., T. Hastie, I. M. Johnstone, R. Tibshirani, H. Ishwaran, K. Knight, J. Loubes, P. Massart, D. Madigan, G. Ridgeway, *et al.* (2004). "Least angle regression". *Annals of Statistics*. 32(2): 407–499.

Eliseyev, A. and T. Aksenova. (2013). "Recursive N-way partial least squares for brain-computer interface". *PloS one*. 8(7).

Eliseyev, A. and T. Aksenova. (2016). "Penalized multi-way partial least squares for smooth trajectory decoding from electrocorticographic (ECoG) recording". *PloS one*. 11(5).

Eliseyev, A., C. Moro, T. Costecalde, N. Torres, S. Gharbi, C. Mestais, A. L. Benabid, and T. Aksenova. (2011). "Iterative N-way partial least squares for a binary self-paced brain–computer interface in freely moving animals". *Journal of neural engineering*. 8(4): 046012.

Eliseyev, A., C. Moro, J. Faber, A. Wyss, N. Torres, C. Mestais, A. L. Benabid, and T. Aksenova. (2012). "L1-penalized N-way PLS for subset of electrodes selection in BCI experiments". *Journal of neural engineering*. 9(4): 045010.

Emparán, M., R. Simpson, S. Almonacid, A. Teixeira, and A. Urtubia. (2012). "Early recognition of problematic wine fermentations through multivariate data analyses". *Food control*. 27(1): 248–253.

Escandar, G. M., A. C. Olivieri, N. K. M. Faber, H. C. Goicoechea, A. M. de la Peña, and R. J. Poppi. (2007). "Second-and third-order multivariate calibration: data, algorithms and applications". *TrAC Trends in Analytical Chemistry*. 26(7): 752–765.

Faber, N. K. M. and R. Bro. (2002). "Standard error of prediction for multiway PLS: 1. Background and a simulation study". *Chemometrics and Intelligent Laboratory Systems*. 61(1-2): 133–149.

Fan, J., L. Ding, Y. Chen, and M. Udell. (2019). "Factor Group-Sparse Regularization for Efficient Low-Rank Matrix Recovery". In: *Advances in Neural Information Processing Systems*. 5105–5115.

Fanelli, G., T. Weise, J. Gall, and L. Van Gool. (2011). "Real time head pose estimation from consumer depth cameras". In: *Joint pattern recognition symposium*. Springer. 101–110.

Feng, L., Y. Liu, L. Chen, X. Zhang, and C. Zhu. (2020). "Robust block tensor principal component analysis". *Signal Processing*. 166: 107271.

Feng, X., T. Li, X. Song, and H. Zhu. (2019). "Bayesian Scalar on Image Regression With Nonignorable Nonresponse". *Journal of the American Statistical Association*: 1–24.

Fernández-Delgado, M., M. Sirsat, E. Cernadas, S. Alawadi, S. Barro, and M. Febrero-Bande. (2018). "An extensive experimental survey of regression methods". *Neural Networks*.

Fishman, M., S. R. White, and E. M. Stoudenmire. (2020). "The ITensor Software Library for Tensor Network Calculations". arXiv: 2007.14822.

Gao, C. and X.-j. Wu. (2012). "Kernel Support Tensor Regression". *Procedia Engineering*. 29: 3986–3990. DOI: https://doi.org/10.1016/j.proeng.2012.01.606.

Gao, Q., L. O. Dragsted, and T. Ebbels. (2019). "Comparison of Bi- and Tri-Linear PLS Models for Variable Selection in Metabolomic Time-Series Experiments". *Metabolites*. 9(5): 92.

Garipov, T., D. Podoprikhin, A. Novikov, and D. Vetrov. (2016). "Ultimate tensorization: compressing convolutional and fc layers alike". *arXiv preprint arXiv:1611.03214*.

Ge, H., K. Zhang, M. Alfifi, X. Hu, and J. Caverlee. (2018). "DisTenC: A distributed algorithm for scalable tensor completion on spark". In: *2018 IEEE 34th International Conference on Data Engineering (ICDE)*. IEEE. 137–148.

Geladi, P., Y.-L. Xie, A. Polissar, and P. Hopke. (1998). "Regression on parameters from three-way decomposition". *Journal of chemometrics*. 12(5): 337–354.

Gómez, V. and M. P. Callao. (2008). "Analytical applications of second-order calibration methods". *Analytica chimica acta*. 627(2): 169–183.

Goodarzi, M. and M. P. De Freitas. (2010). "MIA-QSAR modelling of activities of a series of AZT analogues: bi-and multilinear PLS regression". *Molecular Simulation*. 36(4): 267–272.

Gourier, N., D. Hall, and J. L. Crowley. (2004). "Estimating face orientation from robust detection of salient facial structures". In: *FG Net workshop on visual observation of deictic gestures*. Vol. 6. FGnet (IST–2000–26434) Cambridge, UK. 7.

Goyal, R., S. Ebrahimi Kahou, V. Michalski, J. Materzynska, S. Westphal, H. Kim, V. Haenel, I. Fruend, P. Yianilos, M. Mueller-Freitag, *et al.* (2017). "The" something something" video database for learning and evaluating visual common sense". In: *Proceedings of the IEEE International Conference on Computer Vision*. 5842–5850.

Granger, C. W. (1988). "Some recent development in a concept of causality". *Journal of econometrics*. 39(1-2): 199–211.

Grasedyck, L. (2010). "Hierarchical singular value decomposition of tensors". *SIAM Journal on Matrix Analysis and Applications*. 31(4): 2029–2054.

Guhaniyogi, R., S. Qamar, and D. B. Dunson. (2017). "Bayesian tensor regression". *The Journal of Machine Learning Research*. 18(1): 2733–2763.

Guo, W., I. Kotsia, and I. Patras. (2012). "Tensor learning for regression". *IEEE Transactions on Image Processing*. 21(2): 816–827.

Gurden, S. P., J. A. Westerhuis, R. Bro, and A. K. Smilde. (2001). "A comparison of multiway regression and scaling methods". *Chemometrics and Intelligent Laboratory Systems*. 59(1-2): 121–136.

Hackbusch, W. and S. Kühn. (2009). "A new scheme for the tensor representation". *Journal of Fourier analysis and applications*. 15(5): 706–722.

Hainmueller, J. and C. Hazlett. (2014). "Kernel regularized least squares: Reducing misspecification bias with a flexible and interpretable machine learning approach". *Political Analysis*: 143–168.

Hansen, P., J. Jørgensen, and A. Thomsen. (2002). "Predicting grain yield and protein content in winter wheat and spring barley using repeated canopy reflectance measurements and partial least squares regression". *The Journal of Agricultural Science*. 139(3): 307–318.

Hao, B., B. Wang, P. Wang, J. Zhang, J. Yang, and W. W. Sun. (2021). "Sparse Tensor Additive Regression". *Journal of Machine Learning Research*. 22(64): 1–43. URL: http://jmlr.org/papers/v22/19-769.html.

Harshman, R. A. *et al.* (1970). "Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multimodal factor analysis".

Hasegawa, K., M. Arakawa, and K. Funatsu. (1999). "3D-QSAR study of insecticidal neonicotinoid compounds based on 3-way partial least squares model". *Chemometrics and Intelligent Laboratory Systems*. 47(1): 33–40.

Hasegawa, K., M. Arakawa, and K. Funatsu. (2000). "Rational choice of bioactive conformations through use of conformation analysis and 3-way partial least squares modeling". *Chemometrics and Intelligent Laboratory Systems*. 50(2): 253–261.

Hastie, T., R. Tibshirani, and J. Friedman. (2009). *The elements of statistical learning: data mining, inference, and prediction*. Springer Science and Business Media.

Hawkins, C. and Z. Zhang. (2021). "Bayesian tensorized neural networks with automatic rank selection". *Neurocomputing*. 453: 172–180.

He, L., K. Chen, W. Xu, J. Zhou, and F. Wang. (2018). "Boosted sparse and low-rank tensor regression". In: *Advances in Neural Information Processing Systems*. 1009–1018.

Hewing, L., J. Kabzan, and M. N. Zeilinger. (2019). "Cautious model predictive control using Gaussian process regression". *IEEE Transactions on Control Systems Technology*.

Hoerl, A. E. and R. W. Kennard. (2000). "Ridge regression: biased estimation for nonorthogonal problems". *Technometrics*. 42(1): 80–86.

Hoff, P. D. (2015). "Multilinear tensor regression for longitudinal relational data". *The Annals of Applied Statistics*. 9(3): 1169.

Hou, M. and B. Chaib-Draa. (2015). "Hierarchical Tucker tensor regression: Application to brain imaging data analysis". In: *2015 IEEE International Conference on Image Processing (ICIP)*. IEEE. 1344–1348.

Hou, M. and B. Chaib-draa. (2016). "Online incremental higher-order partial least squares regression for fast reconstruction of motion trajectories from tensor streams". In: *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on*. IEEE. 6205–6209.

Hou, M. and B. Chaib-draa. (2017). “Fast Recursive Low-rank Tensor Learning for Regression.” In: *IJCAI*. 1851–1857.

Hou, M., Y. Wang, and B. Chaib-draa. (2015). “Online local gaussian process for tensor-variate regression: Application to fast reconstruction of limb movements from brain signal”. In: *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE. 5490–5494.

Houthuys, L. and J. A. Suykens. (2021). “Tensor-based restricted kernel machines for multi-view classification”. *Information Fusion*. 68: 54–66.

Hsu, C.-Y., P. Indyk, D. Katabi, and A. Vakilian. (2019). “Learning-Based Frequency Estimation Algorithms.” In: *International Conference on Learning Representations*.

Hu, W., D. Tao, W. Zhang, Y. Xie, and Y. Yang. (2016). “The twist tensor nuclear norm for video completion”. *IEEE transactions on neural networks and learning systems*. 28(12): 2961–2973.

Idé, T. (2019). “Tensorial Change Analysis Using Probabilistic Tensor Regression”. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 33. No. 01. 3902–3909.

Idelbayev, Y. and M. A. Carreira-Perpinán. (2020). “Low-rank compression of neural nets: Learning the rank of each layer”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 8049–8059.

Imaizumi, M., T. Maehara, and K. Hayashi. (2017). “On tensor train rank minimization: Statistical efficiency and scalable algorithm”. In: *Advances in Neural Information Processing Systems*. 3930–3939.

Imaizumi, M. and K. Hayashi. (2016). “Doubly decomposing nonparametric tensor regression”. In: *International Conference on Machine Learning*. PMLR. 727–736.

Indyk, P., A. Vakilian, and Y. Yuan. (2019). “Learning-Based Low-Rank Approximations”. *NeurIPS*.

Izenman, A. J. (1975). “Reduced-rank regression for the multivariate linear model”. *Journal of multivariate analysis*. 5(2): 248–264.

Jong, S. de. (1998). “Regression coefficients in multilinear PLS”. *Journal of Chemometrics: A Journal of the Chemometrics Society*. 12(1): 77–81.

Kaandorp, M. L. and R. P. Dwight. (2020). "Data-driven modelling of the Reynolds stress tensor using random forests with invariance". *Computers & Fluids*. 202: 104497.

Kai, S., G. Yan-chen, W. Hai-qing, and L. Ping. (2004). "GMPLS based intelligent quality control for internal rubber mixing process". In: *Fifth World Congress on Intelligent Control and Automation (IEEE Cat. No. 04EX788)*. Vol. 4. IEEE. 3295–3298.

Kanagawa, H., T. Suzuki, H. Kobayashi, N. Shimizu, and Y. Tagami. (2016). "Gaussian process nonparametric tensor estimator and its minimax optimality". In: *International Conference on Machine Learning*. PMLR. 1632–1641.

Kang, F., X. Liu, and J. Li. (2019). "Concrete Dam Behavior Prediction Using Multivariate Adaptive Regression Splines with Measured Air Temperature". *Arabian Journal for Science and Engineering*. 44(10): 8661–8673.

Karahan, E., P. A. Rojas-Lopez, M. L. Bringas-Vega, P. A. Valdes-Hernandez, and P. A. Valdes-Sosa. (2015). "Tensor analysis and fusion of multimodal brain images". *Proceedings of the IEEE*. 103(9): 1531–1559.

Kasiviswanathan, S. P., N. Narodytska, and H. Jin. (2018). "Network Approximation using Tensor Sketching." In: *IJCAI*. 2319–2325.

Kaymak, S. and I. Patras. (2014). "Multimodal random forest based tensor regression". *IET Computer Vision*. 8(6): 650–657.

Kia, S. M., C. F. Beckmann, and A. F. Marquand. (2018). "Scalable multi-task Gaussian process tensor regression for normative modeling of structured variation in neuroimaging data". *arXiv preprint arXiv:1808.00036*.

Kiers, H. A. and A. K. Smilde. (1998). "Constrained three-mode factor analysis as a tool for parameter estimation with second-order instrumental data". *Journal of Chemometrics*. 12(2): 125–147.

Kilmer, M. E., K. Braman, N. Hao, and R. C. Hoover. (2013). "Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging". *SIAM Journal on Matrix Analysis and Applications*. 34(1): 148–172.

Kilmer, M. E. and C. D. Martin. (2011). "Factorization strategies for third-order tensors". *Linear Algebra and its Applications*. 435(3): 641–658.

Kilmer, M. E., C. D. Martin, and L. Perrone. (2008). "A third-order generalization of the matrix svd as a product of third-order tensors". *Tufts University, Department of Computer Science, Tech. Rep. TR-2008-4*.

Kim, Y.-D., E. Park, S. Yoo, T. Choi, L. Yang, and D. Shin. (2015). "Compression of deep convolutional neural networks for fast and low power mobile applications". *arXiv preprint arXiv:1511.06530*.

Kistler, R., E. Kalnay, W. Collins, S. Saha, G. White, J. Woollen, M. Chelliah, W. Ebisuzaki, M. Kanamitsu, V. Kousky, *et al.* (2001). "The NCEP–NCAR 50-year reanalysis: monthly means CD-ROM and documentation". *Bulletin of the American Meteorological society*. 82(2): 247–268.

Kolbeinsson, A., J. Kossaifi, Y. Panagakis, A. Bulat, A. Anandkumar, I. Tzoulaki, and P. M. Matthews. (2021). "Tensor dropout for robust learning". *IEEE Journal of Selected Topics in Signal Processing*. 15(3): 630–640.

Kolda, T. G. and B. W. Bader. (2006). "MATLAB Tensor Toolbox". *Tech. rep.* Sandia National Laboratories.

Kolda, T. G. and B. W. Bader. (2009). "Tensor decompositions and applications". *SIAM review*. 51(3): 455–500.

Kolda, T. G. (2006). "Multilinear operators for higher-order decompositions." *Tech. rep.* Citeseer.

Kossaifi, J., A. Bulat, G. Tzimiropoulos, and M. Pantic. (2019a). "T-net: Parametrizing fully convolutional nets with a single high-order tensor". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 7822–7831.

Kossaifi, J., Z. C. Lipton, A. Kolbeinsson, A. Khanna, T. Furlanello, and A. Anandkumar. (2020). "Tensor regression networks". *Journal of Machine Learning Research*. 21: 1–21.

Kossaifi, J., Y. Panagakis, A. Anandkumar, and M. Pantic. (2019b). "TensorLy: Tensor Learning in Python". *Journal of Machine Learning Research*. 20(26): 1–6. URL: http://jmlr.org/papers/v20/18-277.html.

Kressner, D. and C. Tobler. (2012). “htucker—A MATLAB toolbox for tensors in hierarchical Tucker format”. *Mathicse, EPF Lausanne.*

Kumar, N., A. C. Berg, P. N. Belhumeur, and S. K. Nayar. (2009). “Attribute and simile classifiers for face verification”. In: *2009 IEEE 12th international conference on computer vision*. IEEE. 365–372.

La Cascia, M., S. Sclaroff, and V. Athitsos. (2000). “Fast, reliable head tracking under varying illumination: An approach based on registration of texture-mapped 3D models”. *IEEE Transactions on pattern analysis and machine intelligence*. 22(4): 322–336.

Lawrence, K. D. (2019). *Robust regression: analysis and applications*. Routledge.

Learned-Miller, E., G. B. Huang, A. RoyChowdhury, H. Li, and G. Hua. (2016). “Labeled faces in the wild: A survey”. In: *Advances in face detection and facial image analysis*. Springer. 189–248.

Lebedev, V., Y. Ganin, M. Rakhuba, I. V. Oseledets, and V. S. Lempitsky. (2015). “Speeding-up Convolutional Neural Networks Using Fine-tuned CP-Decomposition”. In: *ICLR (Poster)*.

Li, C., Z. Sun, J. Yu, M. Hou, and Q. Zhao. (2019a). “Low-rank Embedding of Kernels in Convolutional Neural Networks under Random Shuffling”. In: *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE. 3022–3026.

Li, L. and X. Zhang. (2017). “Parsimonious tensor response regression”. *Journal of the American Statistical Association*: 1–16.

Li, W., J. Lou, S. Zhou, and H. Lu. (2019b). “Sturm: Sparse tubal-regularized multilinear regression for fmri”. In: *International Workshop on Machine Learning in Medical Imaging*. Springer. 256–264.

Li, X., D. Xu, H. Zhou, and L. Li. (2013a). “Tucker tensor regression and neuroimaging analysis”. *Statistics in Biosciences*: 1–26.

Li, X., J. Haupt, and D. Woodruff. (2017). “Near optimal sketching of low-rank tensor regression”. In: *Advances in Neural Information Processing Systems*. 3466–3476.

Li, Y., J. H. Gilmore, D. Shen, M. Styner, W. Lin, and H. Zhu. (2013b). “Multiscale adaptive generalized estimating equations for longitudinal neuroimaging data”. *NeuroImage*. 72: 91–105.

Li, Y., H. Zhu, D. Shen, W. Lin, J. H. Gilmore, and J. G. Ibrahim. (2011). "Multiscale adaptive regression models for neuroimaging data". *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 73(4): 559–578.

Li, Z., H.-I. Suk, D. Shen, and L. Li. (2016). "Sparse Multi-Response Tensor Regression for Alzheimer's Disease Study With Multivariate Clinical Assessments". *IEEE Transactions on Medical Imaging*. 35(8): 1927–1936.

Liaw, A., M. Wiener, *et al.* (2002). "Classification and regression by randomForest". *R news*. 2(3): 18–22.

Lim, J., G. Kim, C. Mo, M. S. Kim, K. Chao, J. Qin, X. Fu, I. Baek, and B.-K. Cho. (2016). "Detection of melamine in milk powders using near-infrared hyperspectral imaging combined with regression coefficient of partial least square regression model". *Talanta*. 151: 183–191.

Liu, J., P. Musialski, P. Wonka, and J. Ye. (2012). "Tensor completion for estimating missing values in visual data". *IEEE transactions on pattern analysis and machine intelligence*. 35(1): 208–220.

Liu, J., C. Zhu, and Y. Liu. (2020a). "Smooth compact tensor ring regression". *IEEE Transactions on Knowledge and Data Engineering*.

Liu, J., C. Zhu, Z. Long, H. Huang, and Y. Liu. (2021). "Low-rank tensor ring learning for multi-linear regression". *Pattern Recognition*. 113: 107753.

Liu, J., P. C. Cosman, and B. D. Rao. (2017a). "Robust Linear Regression via ℓ_0 Regularization". *IEEE Transactions on Signal Processing*. 66(3): 698–713.

Liu, S., T. Liu, A. Vakilian, Y. Wan, and D. P. Woodruff. (2020b). "On Learned Sketches for Randomized Numerical Linear Algebra". *arXiv preprint arXiv:2007.09890*.

Liu, Y., J. Liu, and C. Zhu. (2020c). "Low-Rank Tensor Train Coefficient Array Estimation for Tensor-on-Tensor Regression". *IEEE Transactions on Neural Networks and Learning Systems*.

Liu, Y., S. Wu, X. Huang, B. Chen, and C. Zhu. (2017b). "Hybrid CS-DMRI: Periodic Time-Variant Subsampling and Omnidirectional Total Variation Based Reconstruction". *IEEE Transactions on Medical Imaging*. 36(10): 2148–2159.

Lock, E. F. (2018). "Tensor-on-tensor regression". *Journal of Computational and Graphical Statistics*. 27(3): 638–647.

Loh, W.-Y. (2011). "Classification and regression trees". *Wiley interdisciplinary reviews: data mining and knowledge discovery*. 1(1): 14–23.

Long, X., L. Jin, and J. Joshi. (2012). "Exploring trajectory-driven local geographic topics in foursquare". In: *Proceedings of the 2012 ACM conference on ubiquitous computing*. 927–934.

Long, Z., Y. Liu, L. Chen, and C. Zhu. (2019). "Low rank tensor completion for multiway visual data". *Signal Processing*. 155(3): 301–316.

Lopes, J. and J. Menezes. (2003). "Industrial fermentation end-product modelling with multilinear PLS". *Chemometrics and Intelligent Laboratory Systems*. 68(1-2): 75–81.

Lu, C., J. Feng, Y. Chen, W. Liu, Z. Lin, and S. Yan. (2016). "Tensor robust principal component analysis: Exact recovery of corrupted low-rank tensors via convex optimization". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 5249–5257.

Lu, C., J. Feng, W. Liu, Z. Lin, S. Yan, *et al.* (2019). "Tensor robust principal component analysis with a new tensor nuclear norm". *IEEE transactions on pattern analysis and machine intelligence*.

Lu, H., K. N. Plataniotis, and A. N. Venetsanopoulos. (2011). "A survey of multilinear subspace learning for tensor data". *Pattern Recognition*. 44(7): 1540–1551.

Lu, W., Z. Zhu, and H. Lian. (2020). "High-dimensional quantile tensor regression". *Journal of Machine Learning Research*. 21(250): 1–31.

Lucey, P., J. F. Cohn, K. M. Prkachin, P. E. Solomon, and I. Matthews. (2011). "Painful data: The UNBC-McMaster shoulder pain expression archive database". In: *Face and Gesture 2011*. IEEE. 57–64.

Luo, L., J. Yang, J. Qian, and Y. Tai. (2015). "Nuclear-L1 norm joint regression for face reconstruction and recognition with mixed noise". *Pattern Recognition*. 48(12): 3811–3824.

Makantasis, K., A. Doulamis, N. Doulamis, and A. Voulodimos. (2019). "Common Mode Patterns for Supervised Tensor Subspace Learning". In: *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE. 2927–2931.

Malik, O. A. and S. Becker. (2018). "Low-rank tucker decomposition of large tensors using tensorsketch". In: *Advances in Neural Information Processing Systems*. 10096–10106.

Mei, Z., W. Zhang, L. Zhang, and D. Wang. (2019). "Real-time multistep prediction of public parking spaces based on Fourier transform–least squares support vector regression". *Journal of Intelligent Transportation Systems*: 1–13.

Mirzaei, M., A. Asif, and H. Rivaz. (2019). "Combining total variation regularization with window-based time delay estimation in ultrasound elastography". *IEEE Transactions on Medical Imaging*. 38(12): 2744–2754.

Mishra, A., D. K. Dey, and K. Chen. (2017). "Sequential co-sparse factor regression". *Journal of Computational and Graphical Statistics*. 26(4): 814–825.

Mitchell, T. M., S. V. Shinkareva, A. Carlson, K.-M. Chang, V. L. Malave, R. A. Mason, and M. A. Just. "Predicting Human Brain Activity Associated with Noun Meanings". *Citeseer. REFERENCES REFERENCES*.

Moly, A., A. Aksenov, A. L. Benabid, and T. Aksenova. (2020). "Online adaptive group-wise sparse NPLS for ECoG neural signal decoding". *arXiv preprint arXiv:2010.10353*.

Mostafapour, S. and H. Parastar. (2015). "N-way partial least squares with variable importance in projection combined to GC× GC-TOFMS as a reliable tool for toxicity identification of fresh and weathered crude oils". *Analytical and bioanalytical chemistry*. 407(1): 285–295.

Nelder, J. A. and R. W. Wedderburn. (1972). "Generalized linear models". *Journal of the Royal Statistical Society: Series A (General)*. 135(3): 370–384.

Nilsson, J., S. de Jong, and A. K. Smilde. (1997). "Multiway calibration in 3D QSAR". *Journal of Chemometrics: A Journal of the Chemometrics Society*. 11(6): 511–524.

Nomikos, P. and J. F. MacGregor. (1994). “Monitoring batch processes using multiway principal component analysis”. *AIChE Journal*. 40(8): 1361–1375.

Nomikos, P. and J. F. MacGregor. (1995). “Multi-way partial least squares in monitoring batch processes”. *Chemometrics and intelligent laboratory systems*. 30(1): 97–108.

Novikov, A., D. Podoprikhin, A. Osokin, and D. Vetrov. (2015). “Tensorizing neural networks”. *arXiv preprint arXiv:1509.06569*.

Nuttall, D. L., H. Goldstein, R. Prosser, and J. Rasbash. (1989). “Differential school effectiveness”. *International Journal of Educational Research*. 13(7): 769–776.

Oseledets, I. V. (2011). “Tensor-train decomposition”. *SIAM Journal on Scientific Computing*. 33(5): 2295–2317.

Oseledets, I. V. and S. V. Dolgov. (2012). “Solution of linear systems and matrix inversion in the TT-format”. *SIAM Journal on Scientific Computing*. 34(5): A2718–A2739.

Pagh, R. (2013). “Compressed matrix multiplication”. *ACM Transactions on Computation Theory (TOCT)*. 5(3): 1–17.

Papadogeorgou, G., Z. Zhang, and D. B. Dunson. (2019). “Soft Tensor Regression”. *arXiv preprint arXiv:1910.09699*.

Papalexakis, E. E., C. Faloutsos, and N. D. Sidiropoulos. (2012). “Parcube: Sparse parallelizable tensor decompositions”. In: *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer. 521–536.

Papastergiou, T. and V. Megalooikonomou. (2017). “A distributed proximal gradient descent method for tensor completion”. In: *2017 IEEE International Conference on Big Data (Big Data)*. IEEE. 2056–2065.

Pati, Y. C., R. Rezaiifar, and P. S. Krishnaprasad. (1993). “Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition”. In: *Proceedings of 27th Asilomar conference on signals, systems and computers*. IEEE. 40–44.

Pedersen, D. K., L. Munck, and S. B. Engelsen. (2002). “Screening for dioxin contamination in fish oil by PARAFAC and N-PLSR analysis of fluorescence landscapes”. *Journal of Chemometrics: A Journal of the Chemometrics Society*. 16(8-10): 451–460.

Penny, W. D., K. J. Friston, J. T. Ashburner, S. J. Kiebel, and T. E. Nichols. (2011). *Statistical parametric mapping: the analysis of functional brain images*. Elsevier.

Perez-Garcia, D., F. Verstraete, M. M. Wolf, and J. I. Cirac. (2007). "Matrix product state representations". *Quantum Information and Computation*. 7(5): 401–430.

Pettersson, Å. K. and B. Karlberg. (1997). "Simultaneous determination of orthophosphate and arsenate based on multi-way spectroscopic-kinetic data evaluation". *Analytica chimica acta*. 354(1-3): 241–248.

Phan, A. H., P. Tichavsky, and A. Cichocki. (2015). "Tensor Deflation for CANDECOMP/PARAFAC— Part I: Alternating Subspace Update Algorithm". *IEEE Transactions on Signal Processing*. 63(22): 5924–5938.

Poullis, C. (2019). "Large-scale urban reconstruction with tensor clustering and global boundary refinement". *IEEE transactions on pattern analysis and machine intelligence*.

Qi, N., Y. Shi, X. Sun, and B. Yin. (2016). "TenSR: Multi-dimensional tensor sparse representation". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 5916–5925.

Rabusseau, G. and H. Kadri. (2016). "Low-Rank Regression with Tensor Responses". In: *Advances in Neural Information Processing Systems*. 1867–1875.

Ragusa, E., P. Gastaldo, R. Zunino, and E. Cambria. (2019). "Learning with similarity functions: a tensor-based framework". *Cognitive Computation*. 11(1): 31–49.

Rai, P., Y. Wang, S. Guo, G. Chen, D. Dunson, and L. Carin. (2014). "Scalable Bayesian low-rank decomposition of incomplete multiway tensors". In: *International Conference on Machine Learning*. 1800–1808.

Raskutti, G. and M. Yuan. (2015). "Convex regularization for high-dimensional tensor regression". *arXiv preprint arXiv:1512.01215*. 639.

Rasmussen, C. E. (2003). "Gaussian processes in machine learning". *Lecture Notes in Computer Science*: 63–71.

Recht, B., M. Fazel, and P. A. Parrilo. (2010). “Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization”. *SIAM review*. 52(3): 471–501.

Reddy, C. K., Y. Li, and C. Aggarwal. (2015). “A Review of Clinical Prediction Models.” *Healthcare data analytics*. 36: 343–378.

Rendall, R., L. H. Chiang, and M. S. Reis. (2019). “Data-driven methods for batch data analysis–A critical overview and mapping on the complexity scale”. *Computers & Chemical Engineering*. 124: 1–13.

Romera-Paredes, B., H. Aung, N. Bianchi-Berthouze, and M. Pontil. (2013). “Multilinear multitask learning”. In: *International Conference on Machine Learning*. 1444–1452.

Rousseeuw, P. J. and M. Hubert. (2018). “Anomaly detection by robust statistics”. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*. 8(2): e1236.

Salzo, S., L. Rosasco, and J. Suykens. (2018). “Solving lp-norm regularization with tensor kernels”. In: *International Conference on Artificial Intelligence and Statistics*. PMLR. 1655–1663.

Sarlos, T. (2006). “Improved approximation algorithms for large matrices via random projections”. In: *2006 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS’06)*. IEEE. 143–152.

Schuldt, C., I. Laptev, and B. Caputo. (2004). “Recognizing human actions: a local SVM approach”. In: *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004*. Vol. 3. IEEE. 32–36.

Sena, M. M. and R. J. Poppi. (2004). “N-way PLS applied to simultaneous spectrophotometric determination of acetylsalicylic acid, paracetamol and caffeine”. *Journal of pharmaceutical and biomedical analysis*. 34(1): 27–34.

Serneels, S., P. Geladi, M. Moens, F. Blockhuys, and P. J. Van Espen. (2005). “Influence properties of trilinear partial least squares regression”. *Journal of Chemometrics: A Journal of the Chemometrics Society*. 19(8): 405–411.

Serneels, S., M. Moens, P. J. Van Espen, and F. Blockhuys. (2004). “Identification of micro-organisms by dint of the electronic nose and trilinear partial least squares regression”. *Analytica chimica acta*. 516(1-2): 1–5.

Serneels, S. and P. J. Van Espen. (2005). “Bootstrap confidence intervals for trilinear partial least squares regression”. *Analytica chimica acta*. 544(1-2): 153–158.

Shi, Y. and A. Anandkumar. (2020). “Higher-Order Count Sketch: Dimensionality Reduction that Retains Efficient Tensor Operations”. In: *2020 Data Compression Conference (DCC)*. IEEE. 394–394.

Sidiropoulos, N. D., L. De Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos. (2017). “Tensor decomposition for signal processing and machine learning”. *IEEE Transactions on Signal Processing*. 65(13): 3551–3582.

Sigal, L., A. O. Balan, and M. J. Black. (2010). “Humaneva: Synchronized video and motion capture dataset and baseline algorithm for evaluation of articulated human motion”. *International journal of computer vision*. 87(1-2): 4.

Signoretto, M., L. De Lathauwer, and J. A. Suykens. (2011). “A kernel-based framework to tensorial data analysis”. *Neural Networks*. 24(8): 861–874.

Signoretto, M., Q. T. Dinh, L. De Lathauwer, and J. A. Suykens. (2014). “Learning with tensors: a framework based on convex optimization and spectral regularization”. *Machine Learning*. 94(3): 303–351.

Sikorska, E., A. Gliszczyńska-Świgło, M. Insińska-Rak, I. Khmelinskii, D. De Keukeleire, and M. Sikorski. (2008). “Simultaneous analysis of riboflavin and aromatic amino acids in beer using fluorescence and multivariate calibration methods”. *Analytica chimica acta*. 613(2): 207–217.

Skup, M., H. Zhu, and H. Zhang. (2012). “Multiscale Adaptive Marginal Analysis of Longitudinal Neuroimaging Data with Time-Varying Covariates”. *Biometrics*. 68(4): 1083–1092.

Smilde, A., R. Bro, and P. Geladi. (2005). *Multi-way analysis: applications in the chemical sciences*. John Wiley & Sons.

Smilde, A. K. (1997). “Comments on multilinear PLS”. *Journal of Chemometrics*. 11(5): 367–377.

Smilde, A. K. and H. A. Kiers. (1999). “Multiway covariates regression models”. *Journal of Chemometrics: A Journal of the Chemometrics Society*. 13(1): 31–48.

Smilde, A. K., J. A. Westerhuis, and R. Boque. (2000). “Multiway multiblock component and covariates regression models”. *Journal of Chemometrics: A Journal of the Chemometrics Society*. 14(3): 301–331.

Smilde, A., R. Tauler, J. Henshaw, L. Burgess, and B. Kowalski. (1994). “Multicomponent determination of chlorinated hydrocarbons using a reaction-based chemical sensor. 3. Medium-rank second-order calibration with restricted Tucker models”. *Analytical Chemistry*. 66(20): 3345–3351.

Smola, A. J. and B. Schölkopf. (2004). “A tutorial on support vector regression”. *Statistics and computing*. 14(3): 199–222.

Soltani, R. and H. Jiang. (2016). “Higher order recurrent neural networks”. *arXiv preprint arXiv:1605.00064*.

Song, Q., H. Ge, J. Caverlee, and X. Hu. (2019). “Tensor completion algorithms in big data analytics”. *ACM Transactions on Knowledge Discovery from Data (TKDD)*. 13(1): 1–48.

Song, X. and H. Lu. (2017). “Multilinear regression for embedded feature selection with application to fmri analysis”. In: *Thirty-First AAAI Conference on Artificial Intelligence*.

Specht, D. F. *et al.* (1991). “A general regression neural network”. *IEEE transactions on neural networks*. 2(6): 568–576.

Spencer, D., R. Guhaniyogi, and R. Prado. (2019). “Bayesian Mixed Effect Sparse Tensor Response Regression Model with Joint Estimation of Activation and Connectivity”. *arXiv preprint arXiv:1904.00148*.

Srivastava, N., E. Mansimov, and R. Salakhudinov. (2015). “Unsupervised learning of video representations using lstms”. In: *International conference on machine learning*. PMLR. 843–852.

Stone, M. and R. J. Brooks. (1990). “Continuum regression: cross-validated sequentially constructed prediction embracing ordinary least squares, partial least squares and principal components regression”. *Journal of the Royal Statistical Society: Series B (Methodological)*. 52(2): 237–258.

Stoudenmire, E. and D. J. Schwab. (2016). “Supervised learning with tensor networks”. In: *Advances in Neural Information Processing Systems*. 4799–4807.

Su, J., W. Byeon, J. Kossaifi, F. Huang, J. Kautz, and A. Anandkumar. (2020). "Convolutional Tensor-Train LSTM for Spatio-Temporal Learning". *Advances in Neural Information Processing Systems*. 33.

Sudlow, C., J. Gallacher, N. Allen, V. Beral, P. Burton, J. Danesh, P. Downey, P. Elliott, J. Green, M. Landray, *et al.* (2015). "UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age". *Plos med*. 12(3): e1001779.

Sun, Q., W.-X. Zhou, and J. Fan. (2019). "Adaptive huber regression". *Journal of the American Statistical Association*: 1–24.

Sun, W. W., B. Hao, and L. Li. "Tensors in Modern Statistical Learning".

Sun, W. W. and L. Li. (2017). "STORE: sparse tensor response regression and neuroimaging analysis". *The Journal of Machine Learning Research*. 18(1): 4908–4944.

Sun, W. W. and L. Li. (2019). "Dynamic tensor clustering". *Journal of the American Statistical Association*: 1–28.

Suzuki, T. (2015). "Convergence rate of Bayesian tensor estimator and its minimax optimality". In: *International Conference on Machine Learning*. 1273–1282.

Suzuki, T., H. Kanagawa, H. Kobayashi, N. Shimizu, and Y. Tagami. (2016). "Minimax optimal alternating minimization for kernel nonparametric tensor learning". *Advances in neural information processing systems*. 29: 3783–3791.

Takada, M., H. Fujisawa, and T. Nishikawa. (2019). "HMLasso: lasso with high missing rate". *stat*. 1050: 19.

Tang, X., X. Bi, and A. Qu. (2020). "Individualized multilayer tensor learning with an application in imaging analysis". *Journal of the American Statistical Association*. 115(530): 836–851.

Tao, D., X. Li, W. Hu, S. Maybank, and X. Wu. (2005). "Supervised tensor learning". In: *Data Mining, Fifth IEEE International Conference on*. IEEE. 8–pp.

Tibshirani, R. (1996). "Regression shrinkage and selection via the lasso". *Journal of the Royal Statistical Society: Series B (Methodological)*. 58(1): 267–288.

Tomioka, R. and T. Suzuki. (2013). “Convex tensor decomposition via structured Schatten norm regularization”. In: *Proceedings of the 26th International Conference on Neural Information Processing Systems-Volume 1*. 1331–1339.

Trevisan, M. G. and R. J. Poppi. (2003). “Determination of doxorubicin in human plasma by excitation–emission matrix fluorescence and multi-way analysis”. *Analytica chimica acta*. 493(1): 69–81.

Tucker, L. R. (1963). “Implications of factor analysis of three-way matrices for measurement of change”. *Problems in measuring change*. 15: 122–137.

Valdes-Sosa, P. A., A. Roebroeck, J. Daunizeau, and K. Friston. (2011). “Effective connectivity: influence, causality and biophysical modeling”. *Neuroimage*. 58(2): 339–361.

Van der Aa, N., X. Luo, G.-J. Giezeman, R. T. Tan, and R. C. Veltkamp. (2011). “Umpm benchmark: A multi-person dataset with synchronized video and motion capture data for evaluation of articulated human motion and interaction”. In: *2011 IEEE international conference on computer vision workshops (ICCV Workshops)*. IEEE. 1264–1269.

Vargas-Govea, B., G. González-Serna, and R. Ponce-Medellın. (2011). “Effects of relevant contextual features in the performance of a restaurant recommender system”. *ACM RecSys*. 11(592): 56.

Vervliet, N., O. Debals, L. Sorber, M. Van Barel, and L. De Lathauwer. (2016). “Tensorlab user guide”. *Available on: http://www. tensorlab. net*.

Vinck, M., L. Huurdeman, C. A. Bosman, P. Fries, F. P. Battaglia, C. M. Pennartz, and P. H. Tiesinga. (2015). “How to detect the Granger-causal flow direction in the presence of additive noise?” *Neuroimage*. 108: 301–318.

Wahls, S., V. Koivunen, H. V. Poor, and M. Verhaegen. (2014). “Learning multidimensional Fourier series with tensor trains”. In: *2014 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*. IEEE. 394–398.

Wang, H. *et al.* (2012). “Bayesian graphical lasso models and efficient posterior computation”. *Bayesian Analysis*. 7(4): 867–886.

Wang, P., T. Shi, and C. K. Reddy. (2020). "Tensor-based Temporal Multi-Task Survival Analysis". *IEEE Transactions on Knowledge and Data Engineering*.

Wang, W., Y. Sun, B. Eriksson, W. Wang, and V. Aggarwal. (2018). "Wide compression: Tensor ring nets". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 9329–9338.

Wang, X., L. T. Yang, Y. Wang, X. Liu, Q. Zhang, and M. J. Deen. (2019). "A distributed tensor-train decomposition method for cyber-physical-social services". *ACM Transactions on Cyber-Physical Systems*. 3(4): 1–15.

Wang, Y., H. Tung, A. J. Smola, and A. Anandkumar. (2015). "Fast and guaranteed tensor decomposition via sketching". In: *Advances in Neural Information Processing Systems*. 991–999.

Wimalawarne, K., M. Sugiyama, and R. Tomioka. (2014). "Multitask learning meets tensor factorization: task imputation via convex optimization". In: *Advances in neural information processing systems*. 2825–2833.

Wimalawarne, K., R. Tomioka, and M. Sugiyama. (2016). "Theoretical and experimental analyses of tensor-based regression and classification". *Neural computation*. 28(4): 686–715.

Wood, S. N. (2011). "Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models". *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 73(1): 3–36.

Woodruff, D. P. *et al.* (2014). "Sketching as a tool for numerical linear algebra". *Foundations and Trends® in Theoretical Computer Science*. 10(1–2): 1–157.

Xie, Q., Q. Zhao, D. Meng, and Z. Xu. (2017). "Kronecker-basis-representation based tensor sparsity and its applications to tensor recovery". *IEEE transactions on pattern analysis and machine intelligence*. 40(8): 1888–1902.

Xiong, L., X. Chen, T.-K. Huang, J. Schneider, and J. G. Carbonell. (2010). "Temporal collaborative filtering with bayesian probabilistic tensor factorization". In: *Proceedings of the 2010 SIAM international conference on data mining*. SIAM. 211–222.

Xu, J., J. Zhou, P.-N. Tan, X. Liu, and L. Luo. (2019). "Spatio-temporal Multi-task Learning via Tensor Decomposition". *IEEE Transactions on Knowledge and Data Engineering*.

Xu, P. and H. Huang. (2015). "Modeling crash spatial heterogeneity: Random parameter versus geographically weighting". *Accident Analysis and Prevention*. 75: 16–25.

Xu, Y., Y. Li, S. Zhang, W. Wen, B. Wang, Y. Qi, Y. Chen, W. Lin, and H. Xiong. (2020). "TRP: Trained Rank Pruning for Efficient Deep Neural Networks". In: *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20*. Ed. by C. Bessiere. International Joint Conferences on Artificial Intelligence Organization. 977–983. DOI: 10.24963/ijcai.2020/136.

Yan, H., K. Paynabar, and M. Pacella. (2019). "Structured Point Cloud Data Analysis Via Regularized Tensor Regression for Process Modeling and Optimization". *Technometrics*. 61(3): 385–395.

Yan, X., B. Xie, L. Song, B. Boots, and G. EDU. (2015). "Large-scale Gaussian process regression via doubly stochastic gradient descent". In: *The ICML Workshop on Large-Scale Kernel Learning*.

Ou-Yang, L., X.-F. Zhang, and H. Yan. (2020). "Sparse regularized low-rank tensor regression with applications in genomic data analysis". *Pattern Recognition*. 107: 107516.

Yang, Y., D. Krompass, and V. Tresp. (2017). "Tensor-train recurrent neural networks for video classification". In: *International Conference on Machine Learning*. PMLR. 3891–3900.

Ye, J., G. Li, D. Chen, H. Yang, S. Zhe, and Z. Xu. (2020). "Block-term tensor neural networks". *Neural Networks*. 130: 11–21.

Yi, B.-K. and C. Faloutsos. (2000). "Fast time sequence indexing for arbitrary Lp norms". In: *VLDB*. Vol. 385. No. 394. 99.

Yin, M., S. Liao, X.-Y. Liu, X. Wang, and B. Yuan. (2020). "Compressing recurrent neural networks using hierarchical tucker tensor decomposition". *arXiv preprint arXiv:2005.04366*.

Yokota, T. and A. Cichocki. (2014). "Multilinear tensor rank estimation via sparse Tucker decomposition". In: *2014 Joint 7th International Conference on Soft Computing and Intelligent Systems (SCIS) and 15th International Symposium on Advanced Intelligent Systems (ISIS)*. IEEE. 478–483.

Yu, R. and Y. Liu. (2016). "Learning from multiway data: Simple and efficient tensor regression". In: *International Conference on Machine Learning*. 373–381.

Yu, R., D. Cheng, and Y. Liu. (2015). "Accelerated online low rank tensor learning for multivariate spatiotemporal streams". In: *International Conference on Machine Learning*. 238–247.

Yu, R., G. Li, and Y. Liu. (2018). "Tensor regression meets gaussian processes". In: *International Conference on Artificial Intelligence and Statistics*. PMLR. 482–490.

Yu, R., S. Zheng, A. Anandkumar, and Y. Yue. (2017). "Long-term forecasting using tensor-train rnns". *Arxiv*.

Yuan, L., Q. Zhao, L. Gui, and J. Cao. (2019). "High-order tensor completion via gradient-based optimization under tensor train format". *Signal Processing: Image Communication*. 73: 53–61.

Zha, Z., X. Yuan, B. Wen, J. Zhou, J. Zhang, and C. Zhu. (2020). "A Benchmark for Sparse Coding: When Group Sparsity Meets Rank Minimization". *IEEE Transactions on Image Processing*: 1–1. DOI: 10.1109/TIP.2020.2972109.

Zhang, A. R., Y. Luo, G. Raskutti, and M. Yuan. (2020). "ISLET: Fast and optimal low-rank tensor regression via importance sketching". *SIAM journal on mathematics of data science*. 2(2): 444–479.

Zhang, C. and P. C. Woodland. (2018). "High order recurrent neural networks for acoustic modelling". In: *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE. 5849–5853.

Zhang, L. and Z.-H. Zhou. (2018). "ℓ_1-regression with Heavy-tailed Distributions". In: *Advances in Neural Information Processing Systems*. 1076–1086.

Zhang, X. and L. Li. (2017). "Tensor Envelope Partial Least-Squares Regression". *Technometrics*: 1–11.

Zhao, Q., C. F. Caiafa, D. P. Mandic, Z. C. Chao, Y. Nagasaka, N. Fujii, L. Zhang, and A. Cichocki. (2013a). "Higher order partial least squares (HOPLS): a generalized multilinear regression method". *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 35(7): 1660–1673.

Zhao, Q., M. Sugiyama, and A. Cichocki. (2018). "Learning efficient tensor representations with ring structure networks". In: *Sixth International Conference on Learning Representations(ICLR Workshop)*.

Zhao, Q., G. Zhou, S. Xie, L. Zhang, and A. Cichocki. (2016). "Tensor ring decomposition". *arXiv preprint arXiv:1606.05535*.

Zhao, Q., L. Zhang, and A. Cichocki. (2015). "Bayesian CP factorization of incomplete tensors with automatic rank determination". *IEEE transactions on pattern analysis and machine intelligence*. 37(9): 1751–1763.

Zhao, Q., G. Zhou, T. Adali, L. Zhang, and A. Cichocki. (2013b). "Kernelization of tensor-based models for multiway data analysis: Processing of multidimensional structured data". *IEEE Signal Processing Magazine*. 30(4): 137–148.

Zhao, Q., G. Zhou, L. Zhang, and A. Cichocki. (2014). "Tensor-variate gaussian processes regression and its application to video surveillance". In: *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE. 1265–1269.

Zhe, S., W. Xing, and R. M. Kirby. (2019). "Scalable high-order gaussian process regression". In: *The 22nd International Conference on Artificial Intelligence and Statistics*. PMLR. 2611–2620.

Zhen, X., A. Islam, M. Bhaduri, I. Chan, and S. Li. (2015). "Direct and simultaneous four-chamber volume estimation by multi-output regression". In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 669–676.

Zhong, B., X. Yang, Y. Shen, C. Wang, T. Wang, Z. Cui, H. Zhang, X. Hong, and D. Chen. (2016). "Higher order partial least squares for object tracking: A 4D-tracking method". *Neurocomputing*. 215: 118–127.

Zhou, H., L. Li, and H. Zhu. (2013). "Tensor regression with applications in neuroimaging data analysis". *Journal of the American Statistical Association*. 108(502): 540–552.

Zhou, J., W. W. Sun, J. Zhang, and L. Li. (2020a). "Partially Observed Dynamic Tensor Response Regression". *arXiv preprint arXiv:2002.09735*.

Zhou, M., Y. Liu, Z. Long, L. Chen, and C. Zhu. (2019). "Tensor rank learning in CP decomposition via convolutional neural network". *Signal Processing: Image Communication*. 73: 12–21.

Zhou, Y., R. K. Wong, and K. He. (2020b). "Broadcasted nonparametric tensor regression". *arXiv preprint arXiv:2008.12927*.

Zhu, X., H.-I. Suk, S.-W. Lee, and D. Shen. (2016). "Canonical feature selection for joint regression and multi-class identification in Alzheimer's disease diagnosis". *Brain imaging and behavior*. 10(3): 818–828.