Nonparametric Tensor Estimation and Completion via Hypergraphon Learning

Miaoyan Wang   MIAOYAN.WANG@WISC.EDU
University of Wisconsin – Madison

Abstract

We consider the problem of tensor estimation from noisy observations with possibly missing entries. A nonparametric approach to tensor estimation is developed based on K-uniform hypergraphons. The hypergraphon model captures the key features of conditional independence arising in the generative process of multiway tensor data. The nonparametric hypergraphon representation of tensors encompasses many existing tensor models—such as CP models, Tucker models, shape constrained tensor models—as special examples. We develop a rate-optimal nonparametric tensor estimator using the stochastic blockmodel approximation to the underlying hypergraphon. A surprising distinction between the hypergraphon-based tensor estimation and high-dimensional nonparametric regression is revealed. Furthermore, we establish the integrated risk bound for the hypergraphon estimation in specially structured functional classes. The result uncovers the joint contribution of the statistical bias-variance error and the agnostic discretization error. Numerical results demonstrate the robustness of our proposal over previous tensor methods and the attractive performance as the tensor order increases.

Keywords: List of keywords

1. Introduction

2. Nonparametric Tensor Model via Hypergraphons

Let $\mathcal{Y} = [\![y_{i_1,\ldots,i_K}]\!] \in \mathbb{R}^{d_1 \times \cdots \times d_K}$ be an order-K $(d_1,\ldots,d_K)$-dimensional data tensor. We propose the following generative process for the tensor entries $y_{i_1,\ldots,i_K}$. First
conditionally independent tensor model:

$$Y_{i_1,\ldots,i_K} | \boldsymbol{\xi}_{i_1,\ldots,i_K} \sim \text{Bernoulli}(\theta_\omega),$$
$$\theta_{i_1,\ldots,i_K} = f(\xi_{i_1}^{(1)},\ldots,\xi_{i_K}^{(K)}),$$

where $f \colon [0,1]^K \mapsto [0,1]$ is an unknown multivariate function of interest, and $\boldsymbol{\xi} \colon [d_1] \times \cdots \times [d_K] \mapsto [0,1]^K$ is the latent design variables at position $(i_1,\ldots,i_K)$.

Furthermore, we assume that the collection of latent variables at the k-th coordinate, $(\xi_1^{(k)},\ldots,\xi_{d_K}^{(k)})$, follow a d_k-dimensional distribution $\mathbb{P}^{(k)}$, and the distributions $\mathbb{P}^{(k)}$ and $\mathbb{P}^{(k')}$ are mutually independent for $k \neq k' \in [K]$.

Definition 1 (Uniform random design) *A uniform random design over the index set $[d]$ is a stochastic process $\{\xi_i \colon i \in [d]\}$, where ξ_i are i.i.d. sampled from $\text{Uniform}[0,1]$ for all $i \in [d]$.*

Definition 2 (Cartesian product of random designs) *An order-K uniform design over the index set $[d_1] \times \cdots \times [d_K]$ is defined as a cartesian product of uniform random designs over each marginal*

index set $[d_k]$ for $k \in [K]$. Specifically, the order-K uniform design is a collection of i.i.d random vectors $\{(\xi_{i_1}^{(1)}, \ldots, \xi_{i_K}^{(K)})\}$, where $\xi_{i_k}^{(k)}$ are i.i.d. sampled from Uniform$[0, 1]$, *for all $i_k \in [d_k]$ and $k \in [K]$.*

Definition 3 (Hypergraphon) *A K-uniform hypergraphon is a measureable function $f \colon [0, 1]^K \mapsto [0, 1]$.*

The function can be seen as a kernel functions for random tensor models.

Acknowledgments

We thank a bunch of people.

Appendix A. My Proof of Theorem 1

This is a boring technical proof.

Appendix B. My Proof of Theorem 2

This is a complete version of a proof sketched in the main text.