Smooth tensor estimation and completion

Miaoyan Wang, April 26, 2020

1 Model

Let $\mathcal{Y} = [\![y_{i_1,\ldots,i_K}]\!] \in \{0,1\}^{d\times\cdots\times d}$ be an order-K, $(d,\ldots,d)$-dimensional binary tensor. Let $\boldsymbol{\xi}^{(k)} = (\xi_1^{(k)},\ldots,\xi_d^{(k)}) \in [0,1]^d$ be random vectors following (unknown) distributions $\mathbb{P}^{(k)}$ for all $k\in[K]$, and $\boldsymbol{\xi}^{(k)}$ and $\boldsymbol{\xi}^{(k')}$ are mutually independent for $k\neq k'\in[K]$. Assume that the entries of $\mathcal{Y}$ are independent sub-Gaussian random variables conditional on $\{\boldsymbol{\xi}^{(k)}\}$.

$$\mathbb{E}\left(y_{i_1,\ldots,i_K}|\{\boldsymbol{\xi}^{(k)}\}\right) = f\left(\xi_{i_1}^{(1)},\ldots,\xi_{i_K}^{(K)}\right),\quad \text{for all } (i_1,\ldots,i_K)\in[d]\times\cdots\times[d],$$

where $f\colon [0,1]^K \mapsto [0,1]$ is an unknown multivariate function belonging to a function class $f\in\mathcal{F}_\alpha(M)$. Specifically, the function class is defined as

$$\mathcal{F}_\alpha(R) = \{f\colon \mathrm{Im}(f)\in[0,1] \text{ and } \|f\|_{\mathcal{H}_\alpha}\leq R\},$$

where $\alpha>0$ is the smoothness parameter and $R>0$ is the Hölder norm bound for the functions in the class.

Recall that the function Hölder norm $\|f\|_{\mathcal{H}_\alpha}$ is defined as

$$\|f\|_{\mathcal{H}_\alpha} \stackrel{\text{def}}{=} \max_{|\boldsymbol{i}|\leq\lfloor\alpha\rfloor}\sup_{\boldsymbol{x}\in\mathcal{D}}|\nabla_{\boldsymbol{i}}f(\boldsymbol{x})| + \max_{|\boldsymbol{i}|=\lfloor\alpha\rfloor}\sup_{\boldsymbol{x}\neq\boldsymbol{x}'\in\mathcal{D}}\frac{\nabla_{\boldsymbol{i}}|f(\boldsymbol{x})-\nabla_{\boldsymbol{i}}f(\boldsymbol{x}')|}{\|\boldsymbol{x}-\boldsymbol{x}'\|_1^{\alpha-\lfloor\alpha\rfloor}},$$

where we have used the short-hand notion

$$\nabla_{\boldsymbol{i}}f(\boldsymbol{x}) = \frac{\partial^{i_1+\cdots+i_K}}{\partial x_1^{i_1}\cdots\partial x_K^{i_K}}f(x_1,\ldots,x_k),$$

for multi-indices $\boldsymbol{i}=(i_1,\ldots,i_K)$ with $|\boldsymbol{i}| = i_1+\cdots+i_K$, and $\boldsymbol{x}=(x_1,\ldots,x_K)$ in the function domain.

2 Estimation

Define the objective function

$$L(\mathcal{C},\{\boldsymbol{M}_k\}) = \|\mathcal{Y}-\mathcal{C}\times_1\boldsymbol{M}_1\times\cdots\times_K\boldsymbol{M}_K\|_F^2.$$

Denote $\Theta = \mathcal{C} \times_1 \boldsymbol{M}_1 \times \cdots \times_K \boldsymbol{M}_K$ and $\boldsymbol{r} = (r_1, \ldots, r_K)$. Then the feasible domain is

$$\mathcal{P}(\boldsymbol{r}) = \Big\{ \Theta \in \mathbb{R}^{d_1 \times \cdots \times d_K} : \Theta = \mathcal{C} \times_1 \boldsymbol{M}_1 \times \cdots \times_K \boldsymbol{M}_K, \text{ where } \mathcal{C} \in \mathbb{R}^{r_1 \times \cdots \times r_K} \text{ and } \boldsymbol{M}_k \in \{0,1\}^{d_k \times r_k} \text{ are membership matrices for all } k \in [K] \Big\}.$$

We propose an adaptive smooth (?) estimation,

$$\hat{\Theta} = \arg\min_{\Theta \in \mathcal{P}(\boldsymbol{r}*)} L(\Theta), \quad \text{with} \quad \boldsymbol{r}^* = (r_1^*, \ldots, r_K^*),$$
$$\text{and} \quad r_k^* = \lceil d_k^{1/(\alpha \wedge 1 + 1)} \rceil \text{ for all } k \in [K].$$

Theorem 2.1. *Consider a function class $\mathcal{F}_\alpha(R)$ with $\alpha > 0$ and $M > 0$. We have*

$$\sup_{f \in \mathcal{F}_\alpha(R)} \sup_{\boldsymbol{\xi}^{(k)} \sim \mathbb{P}^{(k)}, k \in [K]} \frac{1}{d^K} \mathbb{E}\left(\|\hat{\Theta} - f(\xi_{i_1}^{(1)}, \ldots, \xi_{i_K}^{(K)})\|_F^2 \right) \leq C \left(d^{-K\alpha/(\alpha+1)} + \frac{\log d}{d^{K-1}} \right),$$

where the constant $C > 0$ epends only on R, and the expectation is taken jointly over $\mathcal{Y}$, $\{\boldsymbol{\xi}^{(k)}\}$ for all $k \in [K]$.

Phase transition at $\alpha = 1$ only for $K \geq 3$??