New smooth tensor definition and corresponding theorem

Chanwoo Lee, July 28, 2021

Definition 1 ((α,β)-smooth tensor)**.** The tensor Θ is called (α,β)-smooth if there exists α-Hölder function $f\colon \boldsymbol{S}^\beta\times\cdots\times\boldsymbol{S}^\beta\rightarrow\mathbb{R}$ and $\boldsymbol{x}_{\ell i_\ell}\in\boldsymbol{S}^\beta$ for $\ell\in[m], i_\ell\in[d]$ such that

$$\Theta(i_1,\ldots,i_m)=f(\boldsymbol{x}_{1i_1},\ldots,\boldsymbol{x}_{mi_m}),\text{ for all }(i_1,\ldots,i_m)\in[d]^m.$$

Remark 1. New definition of smooth tensor in Definition 1 incorporates most existing tensor models, including hypergraphon model, low-rank tensors, single index models,and GLM models.

Example 1 (Hypergraphon model)**.** Tensor Θ generated from α-smooth hypergraphon satisfies

$$\Theta(i_1,\ldots,i_m)=f\left(\frac{i_1}{d},\ldots,\frac{i_m}{d}\right),\text{ for all }(i_1,\ldots,i_m)\in[d]^m,$$

where $f\colon[0,1]\times\cdots\times[0,1]\rightarrow[0,1]$ is $\alpha-$smooth function. Then by definition, we check that Θ is $(\alpha,1)$-smooth tensor by setting $\boldsymbol{x}_{\ell i_\ell}=i_\ell/d$ for all $\ell\in[m], i_\ell\in[d]$.

Example 2 (Low rank model(CP))**.** Low rank tensor Θ with rank r is $(1,r)$-smooth tensor by the following three steps of construction.

S1: Fix an m-variate function

$$\begin{aligned}
&f\colon \boldsymbol{S}^r\times\cdots\times\boldsymbol{S}^r\rightarrow[0,1]\\
&\quad(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_m)\mapsto f(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_m)=\sum_{i=1}^r\lambda_i\boldsymbol{x}_1(i)\cdots\boldsymbol{x}_m(i).
\end{aligned}$$

By the multilinearity of f and boundness of $\boldsymbol{S}^r$, f is $\bar\lambda$-Lipschitz function, where $\bar\lambda=\max_{i\in[r]}\lambda_i$.

S2: Draw samples $\boldsymbol{x}_{11},\ldots,\boldsymbol{x}_{1d},\ldots,\boldsymbol{x}_{m1},\ldots,\boldsymbol{x}_{md}$ from $\boldsymbol{S}^r$.

S3: Define the signal tensor,

$$\Theta(i_1,\ldots,i_m)=f(\boldsymbol{x}_{1i_1},\ldots,\boldsymbol{x}_{mi_m}),\text{ for all }(i_1,\ldots,i_m)\in[d]^m.$$

Example 3 (Low rank model(Tucker))**.** Low rank tensor Θ with rank $\boldsymbol{r}=(r_1,\ldots,r_m)$ is $(1,\bar r)$-smooth tensor where $\bar r=\max_{k\in[m]}r_k$ from the following steps.

S1: For $\mathcal{S}\in\mathbb{R}^{r_1\times\cdots\times r_m}$, we define $\bar{\mathcal{S}}$ as

$$\bar{\mathcal{S}}(i_1,\ldots,i_m)=\begin{cases}0 & \text{if there exists } k\in[m] \text{ such that } i_k>r_k,\\ \mathcal{S}(i_1,\ldots,i_m) & \text{otherwise,}\end{cases}\quad\text{for all }(i_1,\ldots,i_m)\in[\bar r]^m.$$

S2: Fix an m-variate function

$$\begin{aligned}
&f\colon \boldsymbol{S}^{\bar r}\times\cdots\times\boldsymbol{S}^{\bar r}\rightarrow\mathbb{R}\\
&\quad(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_m)\mapsto f(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_m)=\bar{\mathcal{S}}\times_1\boldsymbol{x}_1\times\cdots\times_m\boldsymbol{x}.
\end{aligned}$$

By the multilinearity of f and boundness of $\boldsymbol{S}^{\bar{r}}$, f is $\max_{k\in[m]}\|\bar{S}_{(k)}\|_{\text{sp}}$-Lipschitz function.

S3: Draw samples $\boldsymbol{x}_{11},\ldots,\boldsymbol{x}_{1d}\in\boldsymbol{S}^{r_1},\ldots,\boldsymbol{x}_{m1},\ldots,\boldsymbol{x}_{md}\in\boldsymbol{S}^{r_m}$ and obtain $\tilde{\boldsymbol{x}}_{\ell k}=(\underbrace{\boldsymbol{x}_{\ell k}}_{r_k},\underbrace{0}_{\bar{r}-r_k})^T\in\boldsymbol{S}^{\bar{r}}$ for all $\ell\in[d],k\in[m]$.

S4: Define the signal tensor,

$$\Theta(i_1,\ldots,i_m)=f(\tilde{\boldsymbol{x}}_{1i_1},\ldots,\tilde{\boldsymbol{x}}_{mi_m}),\text{ for all }(i_1,\ldots,i_m)\in[d]^m.$$

Example 4 (Tensor block model)**.** Tensor block model assumes a checkerboard structure among tensor entries under marginal index permutation. The signal tensor Θ takes at most r distinct values, where r is the total number of multiway blocks. Our model incorporates TBM with $(1,r)$-smoothness.

Example 5 (Single index model)**.** Single index model is a flexible semi-parametric model proposed in economics and high-dimensional statistics. The SIM assumes the existence of a (unknown) monotonic function $g\colon\mathbb{R}\to\mathbb{R}$ such that $g(\Theta)$ has rank r. As long as g is Lipchitz function, we can easily see that SIM is included in $(1,r)$-smooth tensor.

Example 6 (Generalized linear model)**.** Let $\mathcal{Y}$ be a binary tensor from a logistic model with mean $\Theta=\text{logit}(\mathcal{Z})$, where $\mathcal{Z}$ is a latent low-rank tensor with rank r. Since logistic link is 1-Lipchitz function and Example 2, This model generates $(1,r)$-smooth tensor. Same conclusion holds for general exponential-family models with a (known) link function.

Theorem 0.1 (Block approximation to (α,β)-smooth tensors)**.** Let $\Theta\in[0,1]^m$ be an order-m (α,β)-smooth tensor defined in (1). Then, for every $K\in\mathbb{N}_+$, there exists a block-K^m tensor $\bar{\Theta}$ such that

$$\|\Theta-\bar{\Theta}\|_F\lesssim\frac{d^m}{K^{2\alpha/\beta}}.$$

Proof. Notice that the covering number $\frac{1}{\epsilon^\beta}\leq N(\epsilon,\boldsymbol{S}^\beta,\|\cdot\|_F)\leq\left(\frac{2}{\epsilon}+1\right)^\beta$, which implies $\boldsymbol{S}^r$ can be covered by at most K number of $2K^{-1/\beta}$-balls. Therefore, there exists $\boldsymbol{y}_1,\ldots,\boldsymbol{y}_K\in\boldsymbol{S}^\beta$ satisfying

$$\boldsymbol{S}^\beta\subset\cup_{k\in[K]}B(\boldsymbol{y}_k,2K^{-1/\beta}).$$

Then, for every $x_{\ell i_\ell}$, we can always find some $\boldsymbol{y}\in\{\boldsymbol{y}_1,\ldots,\boldsymbol{y}_K\}$ such that

$$\|\boldsymbol{x}_{\ell i_\ell}-\boldsymbol{y}\|_2\leq\frac{2}{K^{1/\beta}}.$$

Let $z_\ell\colon[d]\to[K]$ be ℓ-th membership function that satisfies

$$\|\boldsymbol{x}_{\ell i_\ell}-\boldsymbol{y}_{z_\ell(i_\ell)}\|_2\leq\frac{2}{K^{1/\beta}},\quad\text{for all }\ell\in[m],i_\ell\in[d].$$

Define a block-K^m tensor $\bar{\Theta}$ by

$$\bar{\Theta}(i_1,\ldots,i_m):=f(\boldsymbol{y}_{z_1(i_1)},\ldots,\boldsymbol{y}_{z_m(i_m)}).$$

By definition of $\bar{\Theta}$, we have

$$
\begin{aligned}
\|\Theta - \bar{\Theta}\|_F^2 &= \sum_{(i_1,\ldots,i_m)\in[d]^m} \left| f(\boldsymbol{x}_{1i_1},\ldots,\boldsymbol{x}_{mi_m}) - f(\boldsymbol{y}_{z_1(i_1)},\ldots,\boldsymbol{y}_{z_m(i_m)}) \right|^2 \\
&\le \sum_{(i_1,\ldots,i_m)\in[d]^m} \sum_{\ell\in[m]} \|\boldsymbol{x}_{\ell i_\ell} - \boldsymbol{y}_{z_\ell(i_\ell)}\|_2^{2\alpha} \\
&\le \frac{4^\alpha m d^m}{K^{2\alpha/\beta}},
\end{aligned}
$$

where the second inequality uses the smoothness of f.

□