Generalized tensor response regression with multilinear covariates

Anonymous Author 1
Unknown Institution 1

Anonymous Author 2
Unknown Institution 2

Anonymous Author 3
Unknown Institution 3

Abstract

We consider the problem of learning higher-order tensor with side information on a set of modes. Such data problems arise frequently arise in applications such as neuroimaging, network analysis, and ... We propose a new family of tensor response regression models that incorporate covariate information.

1 Introduction

Many contemporary scientific and engineering studies collect multi-way array data, a.k.a. tensor, accompanied by additional covariates. For example, in neuro-imaging analysis, researchers measure brain connections from a sample of individuals with the goal to identifying the brain edges affected by individual covariates. In social network analysis, explain the connection (comminitity) by attributable of both nodes. ... (add two pictures; one for estimating network population; another for estimating link prediction) In this article, we provide a general treatment to these seemingly different problems.

2 General Model

Let $\mathcal{Y} = [\![y_{i_1,\ldots,i_K}]\!] \in \mathbb{R}^{d_1\times\cdots\times d_K}$ denote an order-K data tensor of interest. In addition, suppose we observe covariate $\boldsymbol{X}_k = [\![x_{pi}^{(k)}]\!] \in \mathbb{R}^{p_k\times d_k}$ on the mode-k, where $x_{pi}^{(k)}$ is the p-th covariate of entry i along the mode k. We propose the following multilinear structure in the mean of the tensor. Specifically,

$$\begin{aligned}&\mathbb{E}(\mathcal{Y}|\boldsymbol{X}_1,\ldots,\boldsymbol{X}_K) = f(\Theta), \text{ where}\\ &\Theta = \mathcal{B}\times\{\boldsymbol{X}_1,\ldots,\boldsymbol{X}_K\},\end{aligned}$$

where $f(\cdot)$ is a known link function, $\Theta \in \mathbb{R}^{d_1\times\cdots\times d_K}$ is the linear predictor, $\mathcal{B}\in\mathbb{R}^{p_1\times\cdots p_K}$ is the parameter tensor of interest, and $\boldsymbol{X}_k \in \mathbb{R}^{d_k\times p_k}$ are known covariate matrices, and $\times$ denotes the tensor Tucker product. We give three examples of multi-covariates tensor regression model arises in literature.

Example 1 (Spatio-temporal growth model)**.** Let $\mathcal{Y} = [\![y_{ijk}]\!] \in \mathbb{R}^{d\times m\times n}$ denote the pH measurements of d lakes at m levels of depth and for n time points. Suppose the sampled lakes belong to q types, with p lakes in each type. Let $\{\ell_j\}_{j\in[m]}$ denote the depth levels and $\{t_k\}_{k\in[n]}$ the time points. Assume the expected pH trend in depth is a polynomial of order $r-1$ and that the expected trend in time is a polynomial of order $s-1$. Then, a classical spatio-temporal growth model can be represented as

$$\mathbb{E}(\mathcal{Y}|\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3) = \mathcal{B}\times\{\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{X}_3\},$$

where $\mathcal{B}\in\mathbb{R}^{p\times r\times s}$ is the coefficient tensor of interest, $\boldsymbol{X}_1\in\{0,1\}^{d\times p}$ is the design matrix for lake types,

$$\boldsymbol{X}_2 = \begin{pmatrix} 1 & \ell_1 & \cdots & \ell_1^{r-1}\\ 1 & \ell_2 & \cdots & \ell_2^{r-1}\\ \vdots & \vdots & \ddots & \vdots\\ 1 & \ell_m & \cdots & \ell_m^{r-1}\end{pmatrix}, \quad \boldsymbol{X}_3 = \begin{pmatrix} 1 & t_1 & \cdots & t_1^{s-1}\\ 1 & t_2 & \cdots & t_2^{s-1}\\ \vdots & \vdots & \ddots & \vdots\\ 1 & t_n & \cdots & t_n^{s-1}\end{pmatrix}$$

are the design matrices for spatial and temporal effects, respectively.

Example 2 (Network population model)**.** Network response model is a very recent model in neuroimanig. The model studies the relationship between the network-valued response with the individual covariates. Suppose we observe n i.i.d. observation $\{(\boldsymbol{Y}_i,\boldsymbol{x}_i):\ i=1,\ldots,n\}$, where $\boldsymbol{Y}_i\in\{0,1\}^{d\times n}$ is the brain connectivity network on the i-th individual and $\boldsymbol{x}_i\in\mathbb{R}^p$ is the subject covariate such as age, gender. The network-response model is of the form

$$\operatorname{logit}(\mathbb{E}(\boldsymbol{Y}_i|\boldsymbol{x}_i)) = \mathcal{B}\times_3\boldsymbol{x}_i, \quad \text{for } i=1,\ldots,n \tag{1}$$

where $\mathcal{B}\in\mathbb{R}^{d\times d\times p}$ is the coefficient tensor of interest. In fact, the model (1) is a special case of our multilinear tensor-response model. To see this, let $\mathcal{Y}\in\{0,1\}^{d\times d\times n}$ denote the response tensor by stacking $\{\boldsymbol{Y}_i\}$ together along the 3$^{\text{rd}}$ mode and $\boldsymbol{X}=[\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n]\in\mathbb{R}^{p\times n}$, then model (1) can be expressed as

$$\operatorname{logit}(\mathbb{E}(\mathcal{Y}|\boldsymbol{X})) = \mathcal{B}\times_3\boldsymbol{X} = \mathcal{B}\times\{\boldsymbol{I}_d,\boldsymbol{I}_d,\boldsymbol{X}\},$$

Preliminary work. Under review by AISTATS 2020. Do not distribute.

where $\boldsymbol{I}_d$ denotes the identity matrix of dimension d.

Example 3 (Link model with node attributes)**.** Let $V = [n]$ be a set of vertices and explanatory variable $x_i \in \mathbb{R}^p$ associated to each $i \in V$. The network $G = (V, E)$ is described by the following matrix model. The edge connects the two vertices i and j independently of the others is modeled as

$$\operatorname{logit}(\mathbb{P}((i,j) \in E) = \boldsymbol{x}_i^T \boldsymbol{B} \boldsymbol{x}_j = \langle \boldsymbol{B}, \boldsymbol{x}_i^T \boldsymbol{x}_j \rangle.$$

Let $\mathcal{Y} = [\![y_{ij}]\!], wehrey_{ij} = \mathbb{1}_{(i,j)\in E}$. Define $\boldsymbol{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n] \in \mathbb{R}^{p \times n}$. Then the above model can be expressed as

$$\operatorname{logit}(\mathbb{E}(\boldsymbol{Y})) = \boldsymbol{B} \times_1 \boldsymbol{X} \times_2 \boldsymbol{X}$$

In the above three example and many other studies, researchers are interested in identifying the regions of tensor that are associated to the covariates.

2.1 Second Level Heading

Second level headings are initial caps, flush left, bold, and in point size 10. Use one line space before the second level heading and one-half line space after the second level heading.

2.1.1 Third Level Heading

Third level headings are flush left, initial caps, bold, and in point size 10. Use one line space before the third level heading and one-half line space after the third level heading.

Fourth Level Heading Fourth level headings must be flush left, initial caps, bold, and Roman type. Use one line space before the fourth level heading, and place the section text immediately after the heading with no line break, but an 11 point horizontal space.

2.2 CITATIONS, FIGURES, REFERENCES

2.2.1 Citations in Text

Citations within the text should include the author's last name and year, e.g., (Cheesman, 1985). References should follow any style that you are used to using, as long as their style is consistent throughout the paper. Be sure that the sentence reads correctly if the citation is deleted: e.g., instead of "As described by (Cheesman, 1985), we first frobulate the widgets," write "As described by Cheesman (1985), we first frobulate the widgets."

2.2.2 Footnotes

Indicate footnotes with a number[1] in the text. Use 8 point type for footnotes. Place the footnotes at the bottom of the column in which their markers appear, continuing to the next column if required. Precede the footnote section of a column with a 0.5 point horizontal rule 1 inch (6 picas) long.[2]

2.2.3 Figures

All artwork must be centered, neat, clean, and legible. All lines should be very dark for purposes of reproduction, and art work should not be hand-drawn. Figures may appear at the top of a column, at the top of a page spanning multiple columns, inline within a column, or with text wrapped around them, but the figure number and caption always appear immediately below the figure. Leave 2 line spaces between the figure and the caption. The figure caption is initial caps and each figure should be numbered consecutively.

Make sure that the figure caption does not get separated from the figure. Leave extra white space at the bottom of the page rather than splitting the figure and figure caption.

This figure intentionally left non-blank

Figure 1: Sample Figure Caption

2.2.4 Tables

All tables must be centered, neat, clean, and legible. Do not use hand-drawn tables. Table number and title always appear above the table. See Table 1.

Use one line space before the table title, one line space after the table title, and one line space after the table. The table title must be initial caps and each table numbered consecutively.

Table 1: Sample Table Title

PART	DESCRIPTION
Dendrite	Input terminal
Axon	Output terminal
Soma	Cell body (contains cell nucleus)

[1] Sample of the first footnote.
[2] Sample of the second footnote.

3 SUPPLEMENTARY MATERIAL

If you need to include additional appendices during submission, you can include them in the supplementary material file.

4 INSTRUCTIONS FOR CAMERA-READY PAPERS

For the camera-ready paper, if you are using LaTeX, please make sure that you follow these instructions. (If you are not using LaTeX, please make sure to achieve the same effect using your chosen typesetting package.)

1. Download `fancyhdr.sty` – the `aistats2020.sty` file will make use of it.

2. Begin your document with

   ```
   \documentclass[twoside]{article}
   \usepackage[accepted]{aistats2020}
   ```

   The `twoside` option for the class article allows the package `fancyhdr.sty` to include headings for even and odd numbered pages. The option `accepted` for the package `aistats2020.sty` will write a copyright notice at the end of the first column of the first page. This option will also print headings for the paper. For the *even* pages, the title of the paper will be used as heading and for *odd* pages the author names will be used as heading. If the title of the paper is too long or the number of authors is too large, the style will print a warning message as heading. If this happens additional commands can be used to place as headings shorter versions of the title and the author names. This is explained in the next point.

3. If you get warning messages as described above, then immediately after `\begin{document}`, write

   ```
   \runningtitle{Provide here an
   alternative shorter version of the title
   of your paper}
   \runningauthor{Provide here the surnames
   of the authors of your paper, all
   separated by commas}
   ```

   Note that the text that appears as argument in `\runningtitle` will be printed as a heading in the *even* pages. The text that appears as argument in `\runningauthor` will be printed as a heading in the *odd* pages. If even the author surnames do not fit, it is acceptable to give a subset of author names followed by "et al."

4. Use the file sample_paper.tex as an example.

5. The camera-ready versions of the accepted papers are 8 pages, plus any additional pages needed for references.

6. If you need to include additional appendices, you can include them in the supplementary material file.

7. Please, don't change the layout given by the above instructions and by the style file.

Acknowledgements

Use the unnumbered third level heading for the acknowledgements. All acknowledgements go at the end of the paper.

References

References follow the acknowledgements. Use an unnumbered third level heading for the references section. Any choice of citation style is acceptable as long as you are consistent. Please use the same font size for references as for the body of the paper—remember that references do not count against your page length total.

References

J. Alspector, B. Gupta, and R. B. Allen (1989). Performance of a stochastic learning microchip. In D. S. Touretzky (ed.), *Advances in Neural Information Processing Systems 1*, 748–760. San Mateo, Calif.: Morgan Kaufmann.

F. Rosenblatt (1962). *Principles of Neurodynamics.* Washington, D.C.: Spartan Books.

G. Tesauro (1989). Neurogammon wins computer Olympiad. *Neural Computation* **1**(3):321–323.