Functional Bayesian Tucker Decomposition for Continuous-indexed Tensor Data

Shikai Fang, Xin Yu, Zheng Wang, Shibo Li, Robert M. Kirby, Shandian Zhe

KAHLERT SCHOOL OF COMPUTING
THE UNIVERSITY OF UTAH

www.sci.utah.edu

ICLR

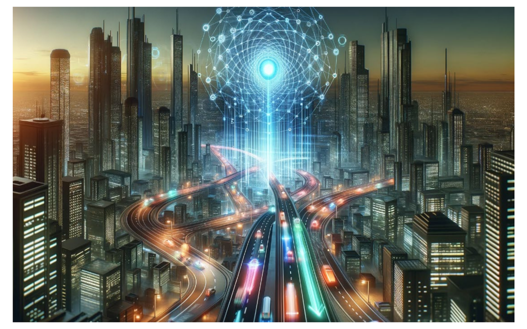
Github Repo

➤ Regular Tensor data and decomposition:

multi-dim array for high-order structural data

(user,user, message)

(item, group, site, device)



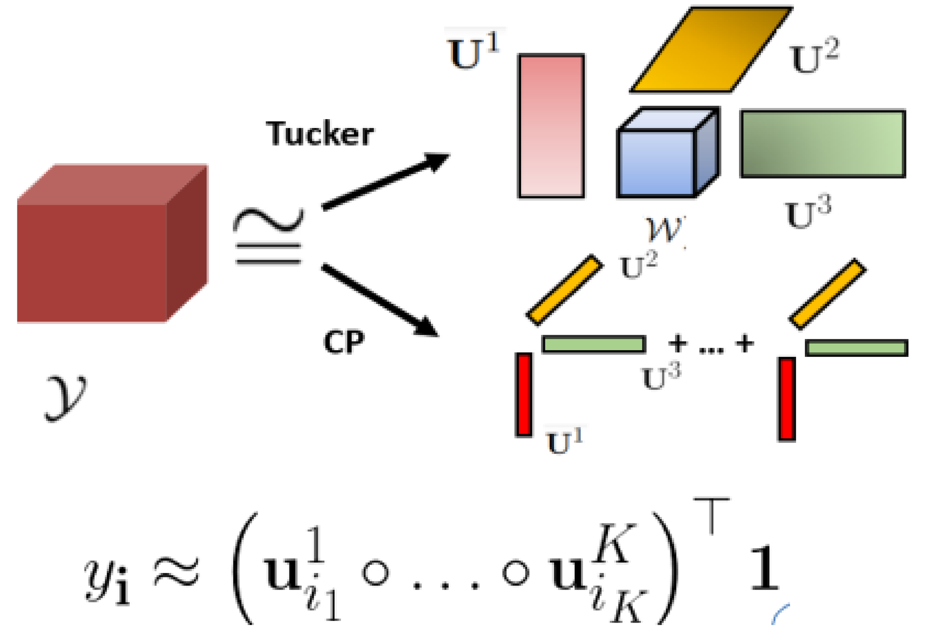
(city, road, section)

Each entry: (index1, index2, index3..)-> value

⇔ Interaction of multiple objects

Interger index: object #

$$y_{\mathbf{i}} \approx \text{vec}(\mathcal{W})^\top \left(\mathbf{u}^1_{i_1} \otimes \ldots \otimes \mathbf{u}^K_{i_K}\right)$$



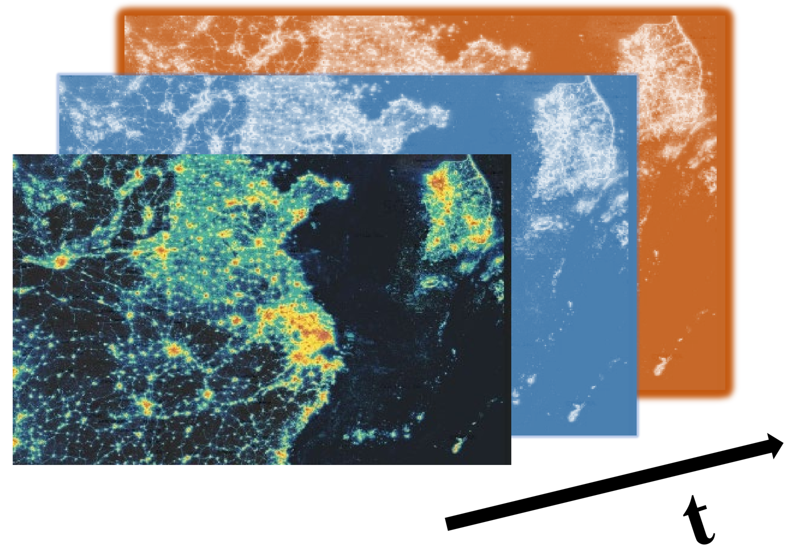
$$y_{\mathbf{i}} \approx \left(\mathbf{u}^1_{i_1} \circ \ldots \circ \mathbf{u}^K_{i_K}\right)^\top \mathbf{1}$$

Object 2 Vector!

Low-rank factors

$$\mathbf{U}^k = [\mathbf{u}^k_1 \ldots \mathbf{u}^k_{d_k}]$$

➤ General case: "Continuous-index" tensor data



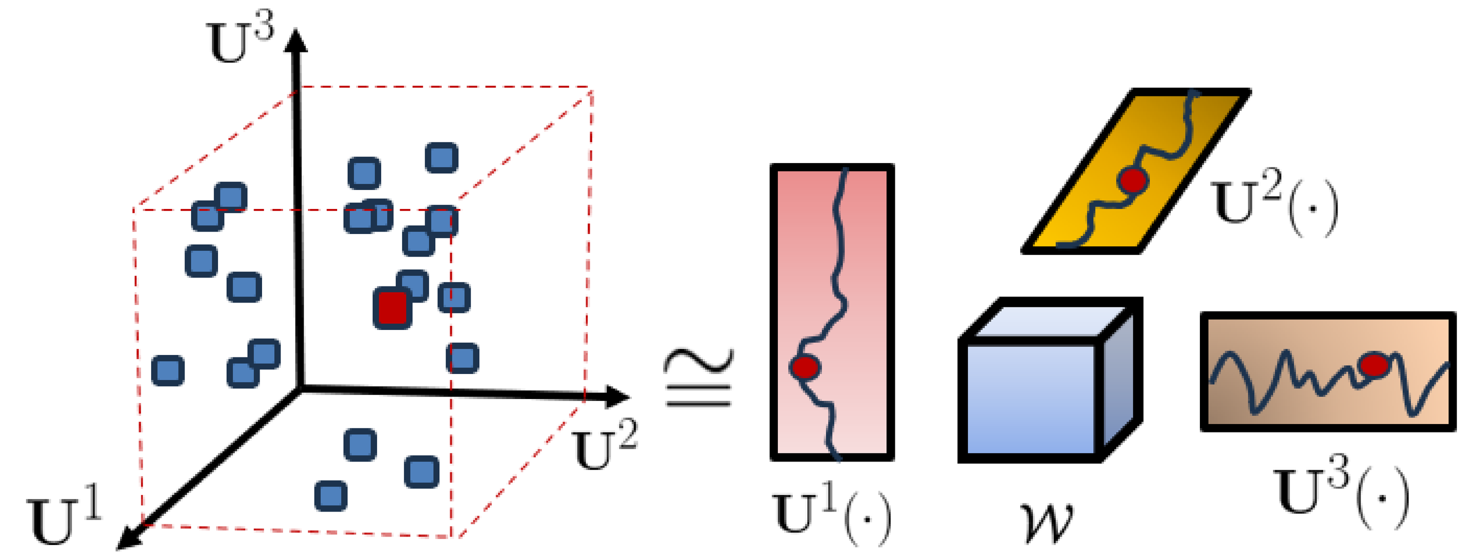
(latitude, longitude, height, time)

Each entry: (index1, index2, index3..)-> value

⇔ a multivariate functions

Real-valued index : input of function

➤ FunBaT: Tucker-form functional decomposition



$$f(\mathbf{i}) = f(i_1, \ldots i_K) \approx \text{vec}(\mathcal{W})^\top \left(\mathbf{U}^1(i_1) \otimes \ldots \otimes \mathbf{U}^K(i_K)\right)$$

continuous-index entry ⇔ interaction of **mode-wise latent functions**

➤ Model of latent function: State-Space Gaussian Process (SSGP)

$$\mathbf{U}^k(i_k) = [u^k_1(i_k), \ldots, u^k_{r_k}(i_k)]^T; \; u^k_j(i_k) \sim \mathcal{GP}\left(0, \kappa(i_k, i'_k)\right), j = 1 \ldots r_k$$

$$p(\mathbf{U}^k) = p(\mathbf{Z}^k) = p(\mathbf{Z}^k(i^1_k), \ldots, \mathbf{Z}^k(i^{N_k}_k)) = p(\mathbf{Z}^k_1) \prod_{s=1}^{N_k - 1} p(\mathbf{Z}^k_{s+1} | \mathbf{Z}^k_s),$$

where $p(\mathbf{Z}^k_1) = \mathcal{N}(\mathbf{Z}^k(i^1_k)|\mathbf{0}, \tilde{\mathbf{P}}^k_\infty)$; $p(\mathbf{Z}^k_{s+1}|\mathbf{Z}^k_s) = \mathcal{N}(\mathbf{Z}^k(i^{s+1}_k)|\tilde{\mathbf{A}}^k_s \mathbf{Z}^k(i^s_k), \tilde{\mathbf{Q}}^k_s)$.

➤ Efficient and scalable Inference by:

moment-matching + message merging + Bayesian Filter/Smoother

$$\mathcal{N}(y_n \mid \text{vec}(\mathcal{W})^\top \left(\mathbf{U}^1(i^n_1) \otimes \ldots \otimes \mathbf{U}^K(i^n_K)\right), \tau^{-1}) \approx Z_n f_n(\tau) f_n(\mathcal{W}) \prod_{k=1}^K f_n(\mathbf{Z}^k(i^n_k)),$$

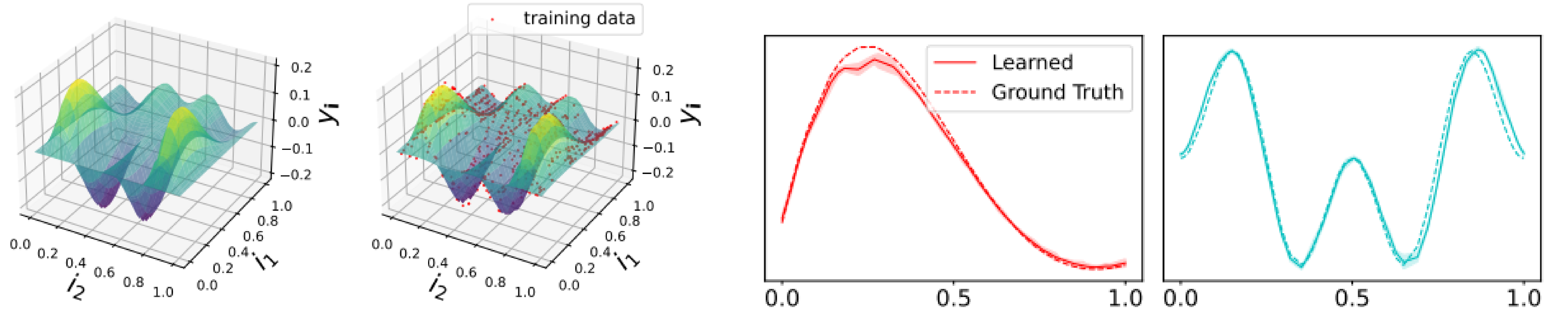
$$q(\mathcal{W}) = p(\mathcal{W}) \prod_{n=1}^N f_n(\mathcal{W}) = \mathcal{N}(\text{vec}(\mathcal{W}) \mid \mathbf{0}, \mathbf{I}) \prod_{n=1}^N \mathcal{N}(\text{vec}(\mathcal{W}) \mid \boldsymbol{\mu}_n, \mathbf{S}_n)$$

$$q(\mathbf{Z}^k_s) = q(\mathbf{Z}^k_{s-1}) p(\mathbf{Z}^k_s | \mathbf{Z}^k_{s-1}) \prod_{n \in \mathcal{D}^k_s} f_n(\mathbf{Z}^k_s)$$

Time cost: $\mathcal{O}(NKR)$

Linear to mode, # entry, rank

➤ Synthetic Data: reconstruction of tensor surface



(a) Real tensor surface (b) Estimated tensor surface (c) $\mathbf{U}^1(i_1)$ (d) $\mathbf{U}^2(i_2)$

➤ Real-world task results & learned latent functions

BeijingAir:
(atmospheric-pressure, temperature, time)

Datasets	RMSE *PM2.5*	RMSE *PM10*	RMSE *SO2*
Resolution: 428 × 501 × 1461 (original)			
P-Tucker	1.256 ± 0.084	1.397 ± 0.001	0.963 ± 0.169
Tucker-ALS	1.018 ± 0.034	1.012 ± 0.021	0.997 ± 0.024
Tucker-SVI	1.891 ± 0.231	1.527 ± 0.107	1.613 ± 0.091
Methods using continuous indexes			
FTT-ALS	1.020 ± 0.013	1.001 ± 0.013	1.001 ± 0.026
FTT-ANOVA	2.150 ± 0.033	2.007 ± 0.015	1.987 ± 0.036
FTT-cross	0.942 ± 0.025	0.933 ± 0.012	0.844 ± 0.026
RBF-SVM	0.995 ± 0.015	0.955 ± 0.02	0.794 ± 0.026
BLR	0.998 ± 0.013	0.977 ± 0.014	0.837 ± 0.021
FunBaT-CP	0.296 ± 0.018	0.343 ± 0.028	$\mathbf{0.386 \pm 0.009}$
FunBaT	$\mathbf{0.288 \pm 0.008}$	$\mathbf{0.328 \pm 0.004}$	$\mathbf{0.386 \pm 0.01}$

US-Temperature: (latitude, longitude, time)

Mode-Rank	RMSE R=3	RMSE R=5	RMSE R=7
P-Tucker	1.306 ± 0.02	1.223 ± 0.022	1.172 ± 0.042
Tucker-ALS	> 10	> 10	> 10
Tucker-SVI	1.438 ± 0.025	1.442 ± 0.021	1.39 ± 0.09
FTT-ALS	1.613 ± 0.0478	1.610 ± 0.052	1.609 ± 0.055
FTT-ANOVA	5.486 ± 0.031	4.619 ± 0.054	3.856 ± 0.059
FTT-cross	1.415 ± 0.0287	1.312 ± 0.023	1.285 ± 0.052
RBF-SVM	2.374 ± 0.047	2.374 ± 0.047	2.374 ± 0.047
BLR	2.959 ± 0.041	2.959 ± 0.041	2.959 ± 0.041
FunBaT-CP	$\mathbf{0.805 \pm 0.06}$	$\mathbf{0.548 \pm 0.03}$	$\mathbf{0.551 \pm 0.048}$
FunBaT	1.255 ± 0.108	1.182 ± 0.117	1.116 ± 0.142

$\mathbf{U}^1(i_1)$ Laitude-mode

$\mathbf{U}^2(i_2)$ Longitude-mode

$\mathbf{U}^3(i_3)$ Time-mode