FunBaT: Functional Bayesian Tucker Decomposition for Continuous-indexed Tensor Data

ICLR 2024

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

Github: github.com/xuangu-fang/Functional-Bayesian-Tucker-Decomposition

Regular Tensor Data

- Multi-dim array for high-order structural data

Entry: (index1, index2..)-> value ⇔ Interaction of multiple objects

integer coordinates!

Social Network

(user,user, message)

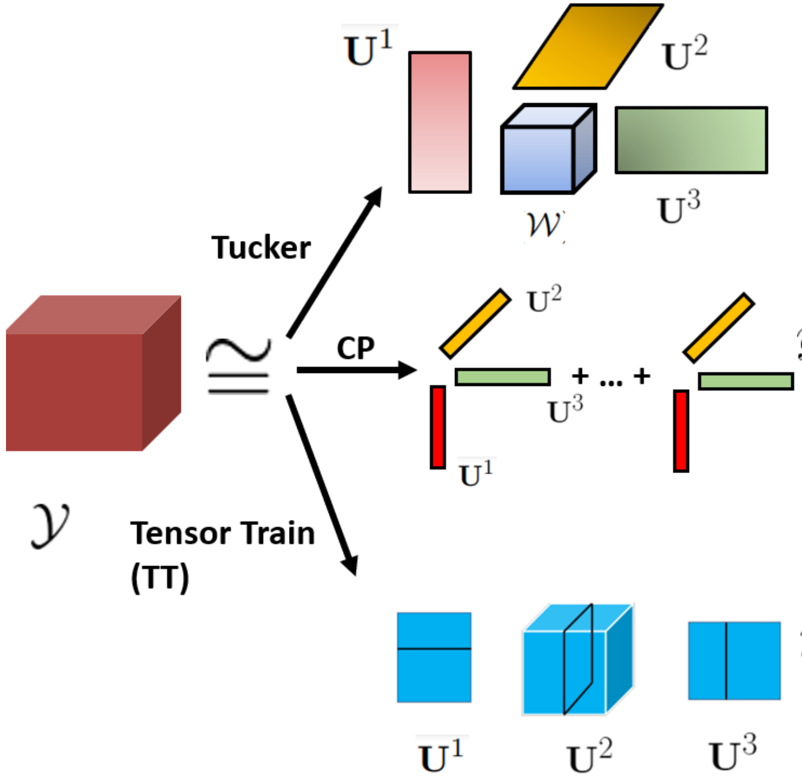
Online Ads

(item, group, site, device)

Traffic Flow

(city, road, population, period)

Regular Tensor Decomposition



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

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

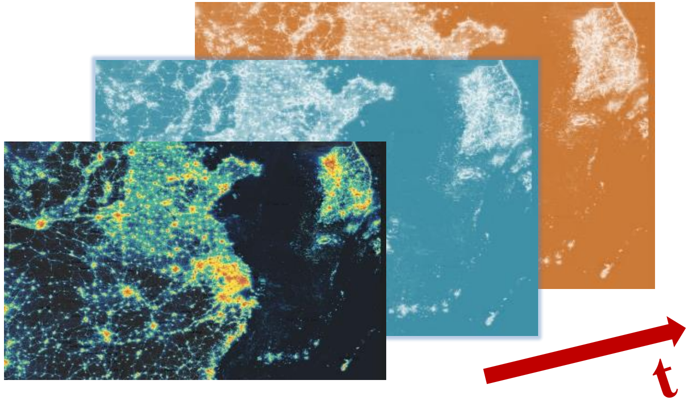
$$y_{\mathbf{i}} \approx \underbrace{\mathbf{u}^1_{i_1}}_{r_0 \times r_1} \underbrace{\mathbf{u}^2_{i_2}}_{r_1 \times r_2} \cdots \underbrace{\mathbf{u}^K_{i_K}}_{r_{K-1} \times r_K}$$

Low-rank factors

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

Continuous-indexed Tensor

- **More general data form for real-world case**



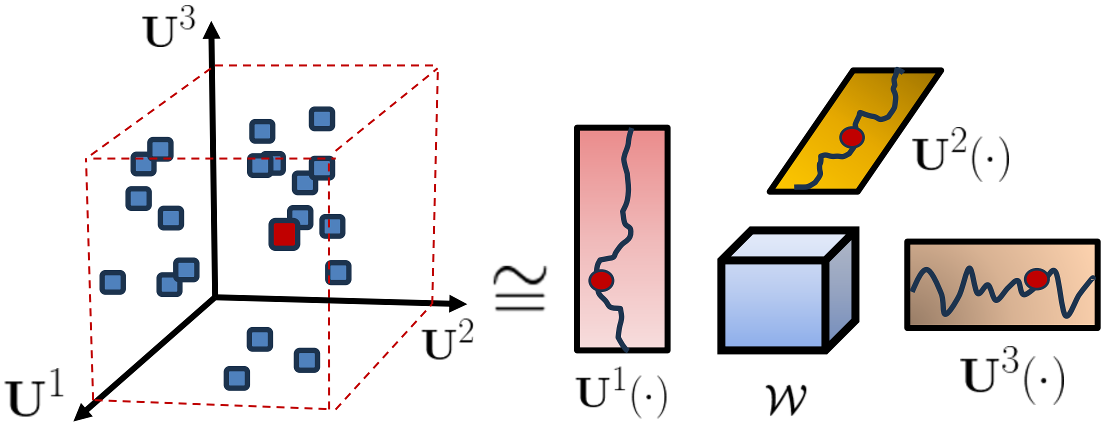
(latitude, longitude, height, time)

Tensor indexes : continuous & real-valued number!

Formulation of FunBaT

- Tensor entry ⇔ Tucker-form interaction of latent functions

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



Formulation of FunBaT

- Assign Gaussian Processes(GPs) prior to latent function

$$\mathbf{U}^k(i_k) = [u_1^k(i_k), \ldots, u_{r_k}^k(i_k)]^T;\ u_j^k(i_k) \sim \mathcal{GP}\left(0, \kappa(i_k, i_k')\right), j = 1 \ldots r_k$$

- Apply State-Space Gaussian Processes(SSGPs) to efficient representation

$$p(\mathbf{U}^k) = p(\mathbf{Z}^k) = p(\mathbf{Z}^k(i_k^1), \ldots, \mathbf{Z}^k(i_k^{N_k})) = p(\mathbf{Z}_1^k) \prod_{s=1}^{N_k-1} p(\mathbf{Z}_{s+1}^k | \mathbf{Z}_s^k),$$

$$\text{where } p(\mathbf{Z}_1^k) = \mathcal{N}(\mathbf{Z}^k(i_k^1) | \mathbf{0}, \tilde{\mathbf{P}}_\infty^k);\ p(\mathbf{Z}_{s+1}^k | \mathbf{Z}_s^k) = \mathcal{N}(\mathbf{Z}^k(i_k^{s+1}) | \tilde{\mathbf{A}}_s^k \mathbf{Z}^k(i_k^s), \tilde{\mathbf{Q}}_s^k).$$

Joint Prob. and Inference

- Joint prob., and approx. posterior:

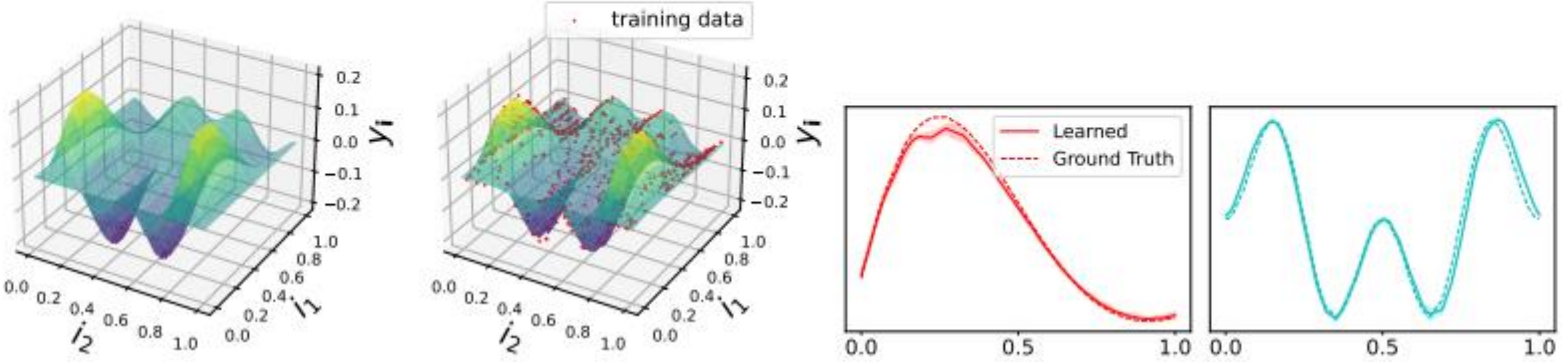
$$p(\mathcal{D}, \mathbf{\Theta}) = p(\mathcal{D}, \{\mathbf{Z}^k\}_{k=1}^K, \mathcal{W}, \tau) = p(\tau)p(\mathcal{W}) \prod_{k=1}^{K} [p(\mathbf{Z}_1^k) \prod_{s=1}^{N_k-1} p(\mathbf{Z}_{s+1}^k | \mathbf{Z}_s^k)] \prod_{n=1}^{N} l_n,$$

$$p(\mathbf{\Theta}|\mathcal{D}) \approx q(\mathbf{\Theta}) = q(\tau)q(\mathcal{W}) \prod_{k=1}^{K} q(\mathbf{Z}^k)$$

$$q(\mathbf{Z}_s^k) = q(\mathbf{Z}_{s-1}^k)p(\mathbf{Z}_s^k|\mathbf{Z}_{s-1}^k) \prod_{n \in \mathcal{D}_s^k} f_n(\mathbf{Z}_s^k),$$

- Linear-cost Inference with messege passing and Bayesian Filter + Smoother

Numerical Results: Synthetic Data



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

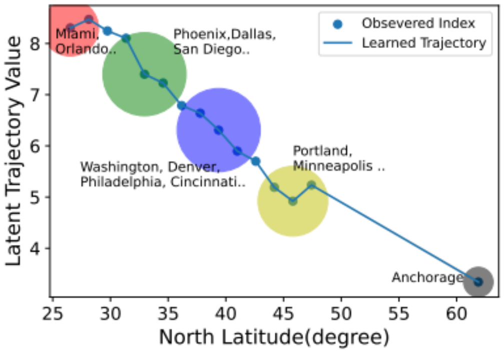
Figure 1: Results of Synthetic Data

Numerical Results: Real-world Data

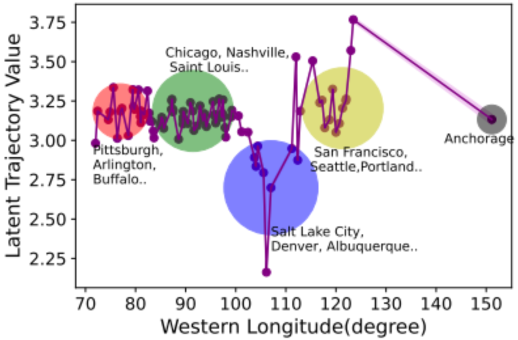
Datasets	RMSE *PM2.5*	RMSE *PM10*	RMSE *SO2*	MAE *PM2.5*	MAE *PM10*	MAE *SO2*
Resolution: $50 \times 50 \times 150$						
P-Tucker	0.805 ± 0.017	0.787 ± 0.006	0.686 ± 0.02	0.586 ± 0.003	0.595 ± 0.005	0.436 ± 0.011
Tucker-ALS	1.032 ± 0.049	1.005 ± 0.029	0.969 ± 0.027	0.729 ± 0.016	0.741 ± 0.007	0.654 ± 0.034
Tucker-SVI	0.792 ± 0.01	0.8 ± 0.026	0.701 ± 0.08	0.593 ± 0.01	0.605 ± 0.019	0.423 ± 0.031
Resolution: $100 \times 100 \times 300$						
P-Tucker	0.8 ± 0.101	0.73 ± 0.021	0.644 ± 0.023	0.522 ± 0.011	0.529 ± 0.013	0.402 ± 0.008
Tucker-ALS	1.009 ± 0.027	1.009 ± 0.026	0.965 ± 0.023	0.738 ± 0.01	0.754 ± 0.007	0.68 ± 0.011
Tucker-SVI	0.706 ± 0.011	0.783 ± 0.067	0.69 ± 0.086	0.509 ± 0.008	0.556 ± 0.031	0.423 ± 0.031
Resolution: $300 \times 300 \times 1000$						
P-Tucker	0.914 ± 0.126	1.155 ± 0.001	0.859 ± 0.096	0.401 ± 0.023	0.453 ± 0.002	0.366 ± 0.015
Tucker-ALS	1.025 ± 0.044	1.023 ± 0.038	1.003 ± 0.019	0.742 ± 0.011	0.757 ± 0.011	0.698 ± 0.007
Tucker-SVI	1.735 ± 0.25	1.448 ± 0.176	1.376 ± 0.107	0.76 ± 0.033	0.747 ± 0.028	0.718 ± 0.023
Resolution: $428 \times 501 \times 1461$ (original)						
P-Tucker	1.256 ± 0.084	1.397 ± 0.001	0.963 ± 0.169	0.451 ± 0.017	0.493 ± 0.001	0.377 ± 0.019
Tucker-ALS	1.018 ± 0.034	1.012 ± 0.021	0.997 ± 0.024	0.738 ± 0.005	0.756 ± 0.007	0.698 ± 0.011
Tucker-SVI	1.891 ± 0.231	1.527 ± 0.107	1.613 ± 0.091	0.834 ± 0.032	0.787 ± 0.018	0.756 ± 0.014
Methods using continuous indexes						
FTT-ALS	1.020 ± 0.013	1.001 ± 0.013	1.001 ± 0.026	0.744 ± 0.007	0.755 ± 0.007	0.696 ± 0.011
FTT-ANOVA	2.150 ± 0.033	2.007 ± 0.015	1.987 ± 0.036	1.788 ± 0.031	1.623 ± 0.014	1.499 ± 0.018
FTT-cross	0.942 ± 0.025	0.933 ± 0.012	0.844 ± 0.026	0.566 ± 0.018	0.561 ± 0.011	0.467 ± 0.033
RBF-SVM	0.995 ± 0.015	0.955 ± 0.02	0.794 ± 0.026	0.668 ± 0.008	0.674 ± 0.014	0.486 ± 0.026
BLR	0.998 ± 0.013	0.977 ± 0.014	0.837 ± 0.021	0.736 ± 0.007	0.739 ± 0.008	0.573 ± 0.009
FunBaT-CP	0.296 ± 0.018	0.343 ± 0.028	$\mathbf{0.386 \pm 0.009}$	$\mathbf{0.18 \pm 0.002}$	0.233 ± 0.013	0.242 ± 0.003
FunBaT	$\mathbf{0.288 \pm 0.008}$	$\mathbf{0.328 \pm 0.004}$	$\mathbf{0.386 \pm 0.01}$	0.183 ± 0.006	$\mathbf{0.226 \pm 0.002}$	$\mathbf{0.241 \pm 0.004}$

Table 1: Prediction error over *BeijingAir-PM2.5*, *BeijingAir-PM10*, and *BeijingAir-SO2* with $R = 2$, which were averaged over five runs. The results for $R = 3, 5, 7$ are in the supplementary.

Learned Latent Functions



(a) $\mathbf{U}^{1}(i_1)$: Mode of latitude



(b) $\mathbf{U}^{2}(i_2)$: Mode of longitude

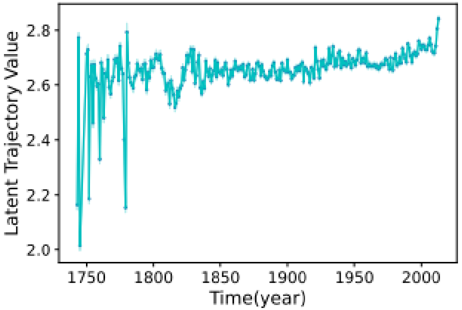


Figure 3: $\mathbf{U}^{3}(i_3)$: Mode of time

Thank you!

Q&A