dCMF: Learning interpretable evolving patterns from temporal multiway data

Christos Chatzis, Carla Schenker, Jeremy Cohen, Evrim Acar

Abstract

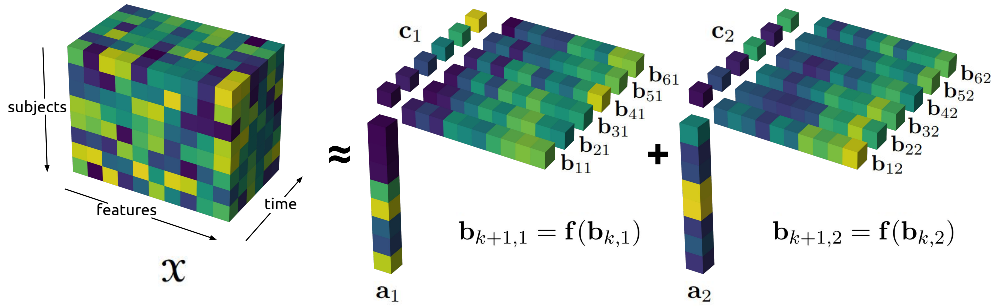

Figure 1: The dCMF model. The tensor $\mathcal{X}$ is factorized using factors $\mathbf{A} = \begin{bmatrix}\mathbf{a}_1 & \mathbf{a}_2\end{bmatrix}$, $\mathbf{C} = \begin{bmatrix}\mathbf{c}_1 & \mathbf{c}_2\end{bmatrix}$ and $\mathbf{B}_k = \begin{bmatrix}\mathbf{b}_{k1} & \mathbf{b}_{k2}\end{bmatrix}$ $\forall k \leq K$, where K is the number of time-steps. dCMF captures evolving patterns through the $\mathbf{B}_k$ factors, which are regularized to adhere to a linear dynamical system.

Multi-way datasets (also referred to as higher-order tensors) are commonly analyzed using unsupervised matrix and tensor factorizations to reveal the underlying patterns. Frequently, such datasets include timestamps and could correspond to, for example, health-related measurements of subjects collected over time. In their essence, these methods decompose the data into interpretable factors that reveal the underlying patterns. When one of the "ways" that the data evolves across is time, implying that patterns can change over time, the objective of such analyses often becomes the identification and tracking of the underlying evolving patterns.

Standard methods can be used off-the-shelf to analyze such temporal data. For example, the CANDECOMP/PARAFAC (CP) tensor model [1] can be used to extract a temporal factor that captures the evolution of pattern strength over time. However, the underlying structure remains fixed. PARAFAC2 [2] addresses this limitation by allowing factors that capture structural changes and has been previously used to capture evolving patterns [3], though specific structural constraints must hold. In contrast, coupled matrix factorizations (CMF) [4] impose no such constraints, offering greater flexibility but lacking uniqueness guarantees. These approaches, however, do not account for the inherent sequential nature of the time dimension. Taking the temporal aspect into account offers several advantages, including more accurate future predictions [5] and robustness to noise and missing data [6, 7]. There are numerous approaches incorporating time in the literature, with temporal regularization being a prevalent strategy [7–10]. Another approach to studying evolving patterns is through linear dynamical systems (LDS). For instance, many studies in hyperspectral imaging utilize LDS, e.g., [11], and the g(eneralized)LDS framework has recently been proposed to analyze multiple multivariate time series [5].

Analyzing temporal multiway data with the goal of capturing interpretable evolving patterns requires methods with specific properties: **(a) time-awareness** to ensure the sequential nature of time is captured, **(b) structural flexibility** to allow for changes in underlying patterns, and **(c) uniqueness** for interpretability. To the best of our knowledge, there is a lack of methods fulfilling these properties.

In this work [12], we bridge the gap between tensor factorizations and dynamical modeling and propose d(ynamical)CMF (see Figure 1). dCMF constrains the temporal evolution of the latent factors to adhere to a specific LDS structure, thus taking into account the order of the observations. We explore how the proposed method is related to CP, PARAFAC2 and t(emporal)PARAFAC2 [6,10]. We highlight also the fact that if an estimate of the transition matrix is known a-priori, the framework allows for promoting this specific structure on the evolving factors.

Using extensive experiments on synthetic data, we demonstrate the effectiveness of the proposed method in terms of recovering the underlying patterns accurately. We generate the ground truth factors, form the datasets, add noise (multiple levels are considered) and attempt to find the underlying patterns. Three different cases are designed. In the first, where the ground truth is smoothly changing and adheres to the PARAFAC2 structural constraint, we observe similar performance between dCMF, PARAFAC2 and tPARAFAC2. However, in the second case where the structural constraint is violated, dCMF outperforms alternative methods when the data is smoothly changing. Lastly, we create data where the evolving factors are evolving according to different transition matrices and we observe that utilizing this information helps dCMF achieve higher accuracy. We empirically discuss the uniqueness of the computed factors.

References

[1] R. A. Harshman, "Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multi-modal factor analysis," *UCLA Working Papers in Phonetics*, vol. 16, pp. 1–84, 1970.

[2] R. A. Harshman, "PARAFAC2: Mathematical and technical notes," *UCLA Working Papers in Phonetics*, no. 22, pp. 30–47, 1972.

[3] E. Acar, M. Roald, K. M. Hossain, V. D. Calhoun, and T. Adali, "Tracing evolving networks using tensor factorizations vs. ICA-based approaches," *Frontiers in Neuroscience*, vol. 16, p. 861402, 2022.

[4] A. P. Singh and G. J. Gordon, "Relational learning via collective matrix factorization," in *KDD'08*, pp. 650—658, 2008.

[5] Z. Liu and M. Hauskrecht, "Learning linear dynamical systems from multivariate time series: A matrix factorization based framework," in *SDM'16: Proc. SIAM Int. Conf. Data Mining*, pp. 810–818, 2016.

[6] C. Chatzis, C. Schenker, M. Pfeffer, and E. Acar, "tPARAFAC2: Tracking evolving patterns in (incomplete) temporal data," *arXiv*, 2024.

[7] H.-F. Yu, N. Rao, and I. S. Dhillon, "Temporal regularized matrix factorization for high-dimensional time series prediction," in *Advances in Neural Information Processing Systems*, vol. 29, pp. 847–855, 2016.

[8] J.-G. Jang, J. Lee, J. Park, and U. Kang, "Accurate PARAFAC2 decomposition for temporal irregular tensors with missing values," in *IEEE Int. Conf. Big Data*, pp. 982–991, 2022.

[9] A. P. Appel, R. L. Cunha, C. C. Aggarwal, and M. M. Terakado, "Temporally evolving community detection and prediction in content-centric networks," in *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 3–18, 2018.

[10] C. Chatzis, M. Pfeffer, P. Lind, and E. Acar, "A time-aware tensor decomposition for tracking evolving patterns," in *MLSP'23: Proc. IEEE Int. Workshop Machine Learning for Signal Processing*, pp. 1–6, 2023.

[11] S. Henrot, J. Chanussot, and C. Jutten, "Dynamical spectral unmixing of multitemporal hyperspectral images," *IEEE Transactions on Image Processing*, vol. 25, no. 7, pp. 3219–3232, 2016.

[12] C. Chatzis, C. Schenker, J. E. Cohen, and E. Acar, "dcmf: Learning interpretable evolving patterns from temporal multiway data," *arXiv preprint*, 2025. `https://arxiv.org/abs/2502.19367`.