Beyond Rank-1: Discovering Rich Community Structure in Multi-Aspect Graphs

1 SUPPLEMENT INFORMATION

1.1 Block Term Decomposition

BTD provides a tensor decomposition in a sum of Tucker terms.

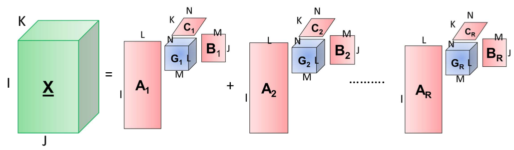

Figure 1: BTD -(L, M, N) for a third-order tensor $\underline{\mathbf{X}} \in \mathbb{R}^{I \times J \times K}$.

Tucker Decomposition[7] : A decomposition of a 3-mode tensor $\underline{\mathbf{X}} \in \mathbb{R}^{I \times J \times K}$ with Rank P, Q, and R is defined as the sum of outer product rank-1 components and one small core tensor $\mathcal{G} \in \mathbb{R}^{P \times Q \times R}$:

$$\underline{\mathbf{X}} \approx \mathcal{G} \bullet_1 \mathbf{A} \bullet_2 \mathbf{B} \bullet_3 \mathbf{C} = \sum_{p=1}^{P} \sum_{q=1}^{Q} \sum_{r=1}^{R} g_{pqr} \mathbf{a}_p \circ \mathbf{b}_q \circ \mathbf{c}_r \quad (1)$$

BTD - (L, M, N) : De Lathauwer at et. [2, 3] introduce a new type of tensor decomposition that unifies the Tucker and the CP decomposition and refereed as Block Term Decomposition (BTD). The BTD of a 3-mode tensor $\underline{\mathbf{X}} \in \mathbb{R}^{I \times J \times K}$, shown in figure 1, is a sum of rank-(L, M, N) terms is a represented as:

$$\underline{\mathbf{X}} \approx \sum_{r=1}^{R} \mathcal{G}_r \bullet_1 \mathbf{A}_r \bullet_2 \mathbf{B}_r \bullet_3 \mathbf{C}_r \quad (2)$$

The factor matrices $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ is defined as $\mathbf{A} = [\mathbf{A}_1 \ \mathbf{A}_2 \ldots \mathbf{A}_R] \in \mathbb{R}^{I \times LR}$, $\mathbf{B} = [\mathbf{B}_1 \ \mathbf{B}_2 \ldots \mathbf{B}_R] \in \mathbb{R}^{J \times MR}$ and $\mathbf{C} = [\mathbf{C}_1 \ \mathbf{C}_2 \ldots \mathbf{C}_R] \in \mathbb{R}^{K \times NR}$. The small core tensors $\mathcal{G}_r \in \mathbb{R}^{L \times M \times N}$ are full rank-(L, M, N). If R=1, then Block-term and Tucker decompositions are same.

1.2 Convergence of RichCom

Here we demonstrate the convergence of Algorithm 1 for cLL1 on three real datasets i.e **Football[5]**, **European ATN**[4] and **EU-Core**[8] network that we use for evaluation. Figure 2 summarizes the convergence of the algorithm, showing the approximated fitness as a function of the number of iterations. It is clear that the algorithm converges to a very good approximation within $40 - 50$ iterations.

1.3 Qualitative Analysis of RichCom

Football[5]: Figure 4 provide visualization of Top-10 community structures discovered by RichCom and we plot these nodes using original football graph and mapped them to ground truth communities provided in literature. Football dataset is characterized

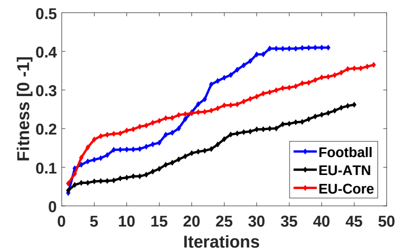

Figure 2: Fitness vs. number of iterations. For each dataset, computation cost was average 47 sec/iteration.

by multiple cliques and near (cliques) structures. Interestingly, we found 10 conferences forming near cliques (in literature, total 12 conferences are given as ground truth) and few of the conferences teams had games with other conferences groups that result in formation of near bipartite and star relation.

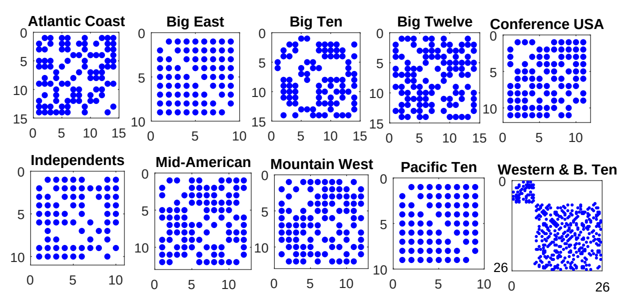

Figure 3: Top 10 structures of football teams found by RichCom.

Figure 4 provide visualization for the structures found by RichCom and we plot these nodes using original football graph and mapped them to ground truth communities provided in literature.

1.4 RichCom and supported ADMM Solver

Given $\underline{\mathbf{X}}$, this section provides the pseudo code of constrained LL1 decomposition in order to factorize the multi-aspect graph or tensor into its constituent community-revealing factors and provide community structure's encoding formulation.

The alternating direction method of multipliers (ADMM)[1] is an algorithm that solves optimization problems given in Equ. (3) by breaking them into smaller pieces.The pseudo code of RichCom and ADMM solver for solving Equ. (3) is given by:

$$\{\mathbf{A}, \mathbf{B}, \mathbf{C}\} \leftarrow \underset{\mathbf{A}, \mathbf{B}, \mathbf{C}}{\operatorname{argmin}} \mathcal{LS}(\underline{\mathbf{X}}, \mathbf{A}, \mathbf{B}, \mathbf{C}) + r(\mathbf{A}) + r(\mathbf{B}) + r(\mathbf{C}) \quad (3)$$

WSDM '20, February 3-7, 2020, Houston, Texas
2020. ACM ISBN xxx-x-xxxx-xxxx-x/xx/xx…$15.00
https://doi.org/10.1145/1122445.1122456

REFERENCES

[1] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, Jonathan Eckstein, et al. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine learning*, pages 1–122, 2011.

[2] Lieven De Lathauwer. Decompositions of a higher-order tensor in block terms–part i: Lemmas for... In *SIAM J. Matrix Anal. Appl.* Citeseer, 2008.

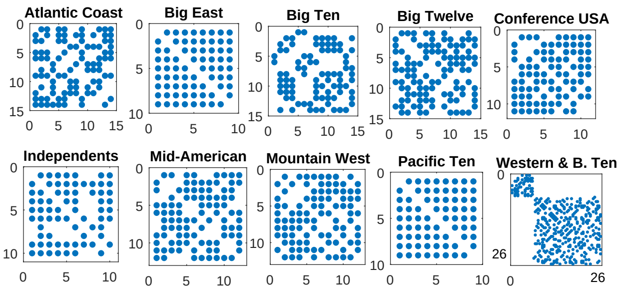

Figure 4: Top 10 structures of football teams found by RichCom.

Algorithm 1: RichCom: Discovering Rich Community Structure

Input: $\underline{\mathbf{X}} \in \mathbb{R}^{I \times J \times K}$, $L \in \mathbb{R}^R$, Max iterations I_{max}.
Output: Factor matrices **A**, **B**, **C**, Structures $\mathcal{S}$.

$(\mathbf{A}, \mathbf{B}, \mathbf{C}) \leftarrow \text{cLL1}(\underline{\mathbf{X}}, L, I_{max})$
$\mathbf{D}_r \leftarrow (\mathbf{A}_r \cdot \mathbf{B}_r^T) \quad \forall r \in R$
$\{Y_{nodes}, Y_{comm}\} \leftarrow communityDetection(\mathbf{D})$
for i = 1 : total communities **do**
  $m \leftarrow Y_{nodes}(find(Y_{comm} == i))$
  $\underline{\mathbf{T}}_i \leftarrow \underline{\mathbf{X}}(m, m, m)$
  $\mathcal{S}_i \leftarrow encode(\underline{\mathbf{T}}_i)$ ▷ **using section ??**
end for
Visualize $\mathcal{S}$ ▷ **using section ??**
Return $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathcal{S})$

Function cLL1 $(\underline{\mathbf{X}}, L, I_{max})$
  Initialize **A**, **B**, **C** randomly
  $s \leftarrow sum(L)$; $R \leftarrow length(L)$
  $\underline{\mathbf{X}}_{(1)} = tenmat(\underline{\mathbf{X}}, 1)$; $\underline{\mathbf{X}}_{(2)} = tenmat(\underline{\mathbf{X}}, 2)$; $\underline{\mathbf{X}}_{(3)} = tenmat(\underline{\mathbf{X}}, 3)$
  while $k < I_{max}$ or not-convergence **do**
    $\mathbf{G} \leftarrow AA^T$; $\mathbf{Y}_A^{(k)} \leftarrow (\mathbf{B}^{(k-1)} \odot \mathbf{C}^{(k-1)})^\dagger$
    $\mathbf{F} \leftarrow \left(\mathbf{Y}_A^{(k)} \cdot \underline{\mathbf{X}}_{(1)}\right)^T$; $\rho = min(10^{-3}, (||\mathbf{Y}_A^{(k)}||_F^2/s)$
    $\mathbf{A}^{(k)}, \hat{\mathbf{A}}^{(k)} \leftarrow \mathbf{ADMM}(\mathbf{A}^{(k-1)}, \hat{\mathbf{A}}^{(k-1)}, \mathbf{F}, \mathbf{G}, \rho)$
    ▷ **Algorithm ADMM step [6]**
    $\mathbf{G} \leftarrow BB^T$; $\mathbf{Y}_B^{(k)} \leftarrow (\mathbf{c}^{(k-1)} \odot \mathbf{A}^{(k)})^\dagger$
    $\mathbf{F} \leftarrow \left(\mathbf{Y}_B^{(k)} \cdot \underline{\mathbf{X}}_{(2)}\right)^T$; $\rho = min(10^{-3}, (||\mathbf{Y}_B^{(k)}||_F^2/s)$
    $\mathbf{B}^{(k)}, \hat{\mathbf{B}}^{(k)} \leftarrow \mathbf{ADMM}(\mathbf{B}^{(k-1)}, \hat{\mathbf{B}}^{(k-1)}, \mathbf{F}, \mathbf{G}, \rho)$
    $\mathbf{G} \leftarrow CC^T$; $\mathbf{Y}_C^{(k)} \leftarrow (\mathbf{A}^{(k)} \odot \mathbf{B}^{(k)})^\dagger =$
    $[(\mathbf{A}_1^{(k)} \otimes \mathbf{B}_1^{(k)})1_{L_1} \ (\mathbf{A}_2^{(k)} \otimes \mathbf{B}_2^{(k)})1_{L_2} \ \ldots \ (\mathbf{A}_R^{(k)} \otimes \mathbf{B}_R^{(k)})1_{L_R}]^\dagger$
    $\mathbf{F} \leftarrow \mathbf{Y}_C^{(k)} \cdot \underline{\mathbf{X}}_{(3)})^T$; $\rho = min(10^{-3}, (||\mathbf{Y}_C^{(k)}||_F^2/s)$
    $\mathbf{C}^{(k)}, \hat{\mathbf{C}}^{(k)} \leftarrow \mathbf{ADMM}(\mathbf{C}^{(k-1)}, \hat{\mathbf{C}}^{(k-1)}, \mathbf{F}, \mathbf{G}, \rho)$
  end while
  Return A, B, C
end Function

Algorithm 2: ADMM solver for Equ. (??).

Input: Residual matrices $\mathbf{R}_H$, $\mathbf{R}_U$, $\mathbf{R}_F$, $\mathbf{R}_G$, and ρ
Output: $\mathbf{R}_H$, $\mathbf{R}_U$

$\mathbf{L} \leftarrow$ Lower Cholesky decomposition$(\mathbf{R}_G + \rho\mathbf{I})$
while $iter < I_{ADMM}$ or $(r < \epsilon$ and $s < \epsilon)$ **do**
  $\widetilde{\mathbf{R}}_H \leftarrow (\mathbf{L}^T)^{-1}\mathbf{L}^{-1}(\mathbf{R}_F + \rho(\mathbf{R}_H + \mathbf{R}_U)$
  $\mathbf{R}_H^0 \leftarrow \mathbf{R}_H$
  $\mathbf{R}_H \leftarrow \text{argmin}_{\mathbf{R}_H} r(\mathbf{R}_H) + Tr(\mathbf{R}_G) + \frac{\rho}{2}||\mathbf{R}_H - \widetilde{\mathbf{R}}_H^T + \mathbf{R}_U||$
  $\mathbf{R}_U \leftarrow \mathbf{R}_U + \mathbf{R}_H - \widetilde{\mathbf{R}}_H$
  $r \leftarrow ||\mathbf{R}_H - \widetilde{\mathbf{R}}_H^T||_F^2/||\mathbf{R}_H||_F$
  $s \leftarrow ||\mathbf{R}_H - \widetilde{\mathbf{R}}_H^0||_F^2/||\mathbf{R}_U||_F$
end while
Return $(\mathbf{R}_H, \mathbf{R}_U)$

[3] Lieven De Lathauwer. Decompositions of a higher-order tensor in block termsâĂŤpart ii: Definitions and uniqueness. *SIAM Journal on Matrix Analysis and Applications*, pages 1033–1066, 2008.

[4] Jungeun Kim and Jae-Gil Lee. Community detection in multi-layer graphs: A survey. *ACM SIGMOD Record*, pages 37–48, 2015.

[5] Polina Rozenshtein, Nikolaj Tatti, and Aristides Gionis. Discovering dynamic communities in interaction networks. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 678–693. Springer, 2014.

[6] Shaden Smith, Alec Beri, and George Karypis. Constrained tensor factorization with accelerated ao-admm. In *2017 46th International Conference on Parallel Processing (ICPP)*, pages 111–120. IEEE, 2017.

[7] L.R. Tucker. Some mathematical notes on three-mode factor analysis. *Psychometrika*, pages 279–311, 1966.

[8] Hao Yin, Austin R Benson, Jure Leskovec, and David F Gleich. Local higher-order graph clustering. In *Proceedings of the 23rd ACM SIGKDD Int. Conf. on KDD*, pages 555–564. ACM, 2017.