Rank-one Boolean tensor factorization and the multilinear polytope

Jan 2023

Aida Khajavirad
Lehigh University

Joint work with Alberto Del Pia (University of Wisconsin-Madison)

Boolean tensor factorization: introduction

- A tensor of order N is an N-dimensional array. Factorizations of high-order tensors, i.e., $N \geq 3$, as products of low-rank matrices, have applications in signal processing, numerical linear algebra, computer vision, data mining, neuroscience, and elsewhere.

- We consider the problem of factorizing a high-order tensor with binary entries, referred to as a binary tensor.

- In Boolean tensor factorization (BTF), the binary tensor is approximated by products of low rank binary matrices using Boolean algebra.

- Applications include neuro-imaging, recommendation systems, topic modeling, and sensor network localization

- BTF is NP-hard in general; all existing methods rely on heuristics without any guarantee on the quality of the solution.

Boolean tensor factorization: problem statement

- For simplicity, we focus on tensors of order three.

- The (Boolean) rank of a binary tensor $\mathcal{G}$ is the smallest integer r such that there exist $3r$ binary vectors x^p, y^p, z^p, for $p \in [r]$, with

$$\mathcal{G} = \bigvee_{p=1}^{r} (x^p \otimes y^p \otimes z^p),$$

  where $\vee$ denotes the component-wise "or" operation, and $\otimes$ denotes the vector outer product.

- The rank-r BTF is the problem of finding the closest rank-r binary tensor to a binary tensor. Given a $n \times m \times l$ binary tensor $\mathcal{G}$ and an integer r, find $3r$ binary vectors $x^p \in \{0,1\}^n$, $y^p \in \{0,1\}^m$, $z^p \in \{0,1\}^l$, for all $p \in [r]$, that minimize

$$\left\| \mathcal{G} - \bigvee_{p=1}^{r} x^p \otimes y^p \otimes z^p \right\|^2,$$

  where the Frobenius norm of $\mathcal{G}$ is defined as $\|\mathcal{G}\| := \sqrt{\sum_{i,j,k} w_{ijk}^2}$.

Rank-one Boolean tensor factorization

- Rank-one BTF: the simplest case of BTF with $r = 1$:

$$\begin{aligned} \min \quad & \left\| \mathcal{G} - x \otimes y \otimes z \right\|^2 \\ \text{s.t.} \quad & x \in \{0,1\}^n, \; y \in \{0,1\}^m, \; z \in \{0,1\}^l. \end{aligned}$$

- Rank-one BTF is NP-hard and no algorithm with theoretical guarantees is known for this problem.

- Define $S_0 := \{(i,j,k) \in [n] \times [m] \times [l] : g_{ijk} = 0\}$, $S_1 := \{(i,j,k) \in [n] \times [m] \times [l] : g_{ijk} = 1\}$, and $w_{ijk} := x_i y_j z_k$, for $i \in [n]$, $j \in [m]$, $k \in [l]$.

- Rank-one BTF can be written, in an extended space, as the problem of minimizing a linear function over a highly structured multilinear set:

$$\begin{aligned} \min \quad & \sum_{(i,j,k) \in S_0} w_{ijk} + \sum_{(i,j,k) \in S_1} (1 - w_{ijk}) \\ \text{s.t.} \quad & w_{ijk} = x_i y_j z_k, \qquad \forall i \in [n], j \in [m], k \in [l] \\ & x \in \{0,1\}^n, \; y \in \{0,1\}^m, \; z \in \{0,1\}^l. \end{aligned}$$

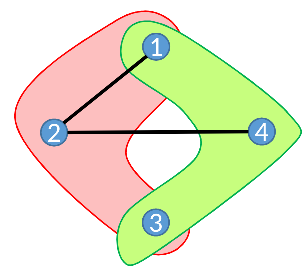
Multilinear sets and polytopes

- With any hypergraph $G = (V, E)$, we associate a Multilinear set $\mathcal{S}_G$ defined as:

$$\mathcal{S}_G = \{(u, w) \in \{0,1\}^{|V|} \times \{0,1\}^{|E|} : w_e = \prod_{v \in e} u_v, \ e \in E\},$$

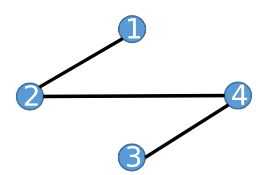
  and the Multilinear polytope MP_G is the convex hull of $\mathcal{S}_G$.

$$w_{12} = u_1 u_2$$
$$w_{24} = u_2 u_4$$
$$w_{123} = u_1 u_2 u_3$$
$$w_{134} = u_1 u_3 u_4$$



- For quadratic sets, we obtain the graph representation of the Boolean quadric polytope (Padberg, 89)

$$w_{12} = u_1 u_2$$
$$w_{24} = u_2 u_4$$
$$w_{34} = u_3 u_4$$



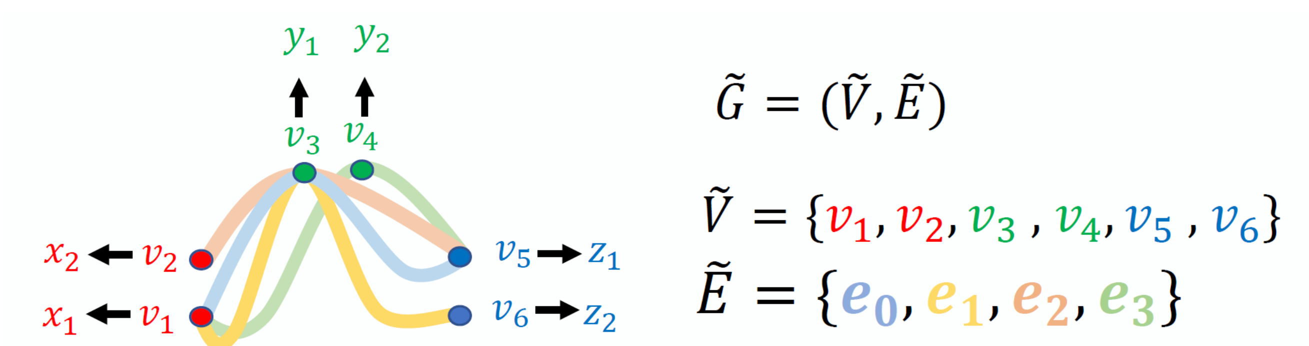
- There are interesting connections between the complexity of MP_G and the acyclicity degree of its hypergraph.

Standard linearization of Rank-one BTF

- A simple LP relaxation of Rank-one BTF can be obtained by replacing each multilinear term $w_{ijk} = x_i y_j z_k$, $x_i, y_j, z_k \in \{0,1\}$, by its convex hull:

$$
\begin{aligned}
\min \quad & \sum_{(i,j,k)\in S_0} w_{ijk} + \sum_{(i,j,k)\in S_1} (1 - w_{ijk}) && \text{(sLP)} \\
\text{s.t.} \quad & w_{ijk} \le x_i,\ w_{ijk} \le y_j,\ w_{ijk} \le z_k, && \forall (i,j,k) \in S_1 \\
& w_{ijk} \ge 0,\ w_{ijk} \ge x_i + y_j + z_k - 2, && \forall (i,j,k) \in S_0 \\
& (x,y,z) \in [0,1]^{n+m+l}.
\end{aligned}
$$

- Stronger LP relaxations can be obtained by convexifying multiple multilinear terms at a time.



Rank-one BTF and the multilinear polytope

- Proposition: The facet description of $\text{MP}_{\tilde{G}}$ is given by:

$$
\begin{aligned}
& w_e \leq u_v, \quad \forall e \in \tilde{E},\ \forall v \in e \\
& w_e \geq 0, \quad \sum_{v \in e} u_v - w_e \leq 2, \quad \forall e \in \tilde{E} \\
& u_v \leq 1, \quad \forall v \in \tilde{V} \\
& u_{e_0 \setminus e} + w_e - w_{e_0} \leq 1, \quad \forall e \in \tilde{E} \setminus \{e_0\} \\
& u_{e \setminus e_0} - w_e + w_{e_0} \leq 1, \quad \forall e \in \tilde{E} \setminus \{e_0\} \\
& -u_{e \cap e'} + w_e + w_{e'} - w_{e_0} \leq 0, \quad \forall e \neq e' \in \tilde{E} \setminus \{e_0\} \\
& \sum_{v \in V} u_v - \sum_{e \in \tilde{E} \setminus \{e_0\}} w_e + w_{e_0} \leq 4.
\end{aligned}
$$

- Proposition: All inequalities defining facets of $\text{MP}_{\tilde{G}}$ are facet-defining for the multilinear polytope of rank-one BTF, once we associate u_{v_1}, u_{v_2} to any two x variables, u_{v_3}, u_{v_4} to any two y variables, and u_{v_5}, u_{v_6} to any two z variables

LP relaxations for Rank-one BTF

- A stronger LP relaxation of Rank-one BTF is given by

$$
\begin{aligned}
\min \quad & \sum_{(i,j,k)\in S_0} w_{ijk} + \sum_{(i,j,k)\in S_1} (1 - w_{ijk}) && \text{(cLP)}\\
\text{s.t.} \quad & w_{ijk} \le x_i,\ w_{ijk} \le y_j,\ w_{ijk} \le z_k, \quad \forall (i,j,k) \in S_1\\
& w_{ijk} \ge 0,\ w_{ijk} \ge x_i + y_j + z_k - 2, \quad \forall (i,j,k) \in S_0\\
& w_{i'jk} - w_{ijk} \le 1 - x_i, \quad \forall (i,j,k) \in S_0,\ (i',j,k) \in S_1\\
& w_{ij'k} - w_{ijk} \le 1 - y_j, \quad \forall (i,j,k) \in S_0,\ (i,j',k) \in S_1\\
& w_{ijk'} - w_{ijk} \le 1 - z_k, \quad \forall (i,j,k) \in S_0,\ (i,j,k') \in S_1\\
& w_{ij'k} + w_{ijk'} - w_{ijk} \le x_i, \quad \forall (i,j,k) \in S_0,\ (i,j',k) \in S_1,\ (i,j,k') \in S_1\\
& w_{i'jk} + w_{ijk'} - w_{ijk} \le y_j, \quad (i,j,k) \in S_0,\ \forall (i',j,k) \in S_1,\ (i,j,k') \in S_1\\
& w_{i'jk} + w_{ij'k} - w_{ijk} \le z_k, \quad \forall (i,j,k) \in S_0,\ (i',j,k) \in S_1,\ (i,j',k) \in S_1\\
& x_i + x_{i'} + y_j + y_{j'} + z_k + z_{k'} + w_{ijk} - w_{i'jk} - w_{ij'k} - w_{ijk'} \le 4,\\
& \qquad \forall (i',j,k) \in S_0,\ (i,j',k) \in S_0,\ (i,j,k') \in S_0,\ (i,j,k) \in S_1\\
& (x,y,z) \in [0,1]^{n+m+l}.
\end{aligned}
$$

LP relaxations for Rank-one BTF

- We analyze the theoretical performance of the following relaxation of Problem (cLP):

$$
\begin{aligned}
\min \quad & \sum_{(i,j,k)\in S_0} w_{ijk} + \sum_{(i,j,k)\in S_1} (1 - w_{ijk}) && \text{(fLP)} \\
\text{s.t.} \quad & w_{ijk} \leq x_i,\; w_{ijk} \leq y_j,\; w_{ijk} \leq z_k, && \forall (i,j,k) \in S_1 \\
& w_{ijk} \geq 0, \quad w_{ijk} \geq x_i + y_j + z_k - 2, && \forall (i,j,k) \in S_0 \\
& w_{i'jk} - w_{ijk} \leq 1 - x_i, && \forall (i,j,k) \in S_0,\; (i',j,k) \in S_1 \\
& w_{ij'k} - w_{ijk} \leq 1 - y_j, && \forall (i,j,k) \in S_0,\; (i,j',k) \in S_1 \\
& w_{ijk'} - w_{ijk} \leq 1 - z_k, && \forall (i,j,k) \in S_0,\; (i,j,k') \in S_1 \\
& (x,y,z) \in [0,1]^{n+m+l}.
\end{aligned}
$$

Recovery under random models

- Given an LP relaxation of Rank-one BTF and a random model for noise in the input tensor, what is the maximum level of noise under which the relaxation recovers the ground truth with high probability? (probability tending to 1 as $n, m, l \to \infty$.)

- Random corruption model for rank-one BTF: consider binary vectors $\bar{x} \in \{0,1\}^n$, $\bar{y} \in \{0,1\}^m$, $\bar{z} \in \{0,1\}^l$ and define the ground truth rank-one tensor $\bar{W} = (w_{ijk}) := \bar{x} \otimes \bar{y} \otimes \bar{z}$. Given $p \in [0,1]$, the noisy tensor $\mathcal{G}$ is as follows: for each $(i,j,k) \in [n] \times [m] \times [l]$, g_{ijk} is corrupted with probability p, i.e., $g_{ijk} := 1 - \bar{x}_i \bar{y}_j \bar{z}_k$, and g_{ijk} is not corrupted with probability $1-p$, i.e., $g_{ijk} := \bar{x}_i \bar{y}_j \bar{z}_k$. we focus on the case where p is a constant.

- Denote by $r_{\bar{x}}$ (resp. $r_{\bar{y}}$, $r_{\bar{z}}$) the ratio of ones in $\bar{x}$ (resp. $\bar{y}$, $\bar{z}$) to the number of elements in $\bar{x}$ (resp. $\bar{y}$, $\bar{z}$). Assume that $r_{\bar{x}}, r_{\bar{y}}, r_{\bar{z}}$ are positive constants and let $r_{\bar{w}} := r_{\bar{x}} r_{\bar{y}} r_{\bar{z}}$.

Information theoretic limits

- What is the corruption range, in terms of p, for which any algorithm, regardless of its computational complexity, can recover the ground truth with high probability.

- Theorem (Information theoretic lower bound): If $p \geq 1/2$, then the probability that rank-one BTF recovers the ground truth is at most $1/2$. Furthermore, if $r_{\bar{x}}, r_{\bar{y}}, r_{\bar{z}}, p$ are positive constants and $p > 1/2$, then with high probability rank-one BTF does not recover the ground truth.

- Theorem (Information theoretic upper bound): Assume that $r_{\bar{x}}, r_{\bar{y}}, r_{\bar{z}}$ are positive constants and

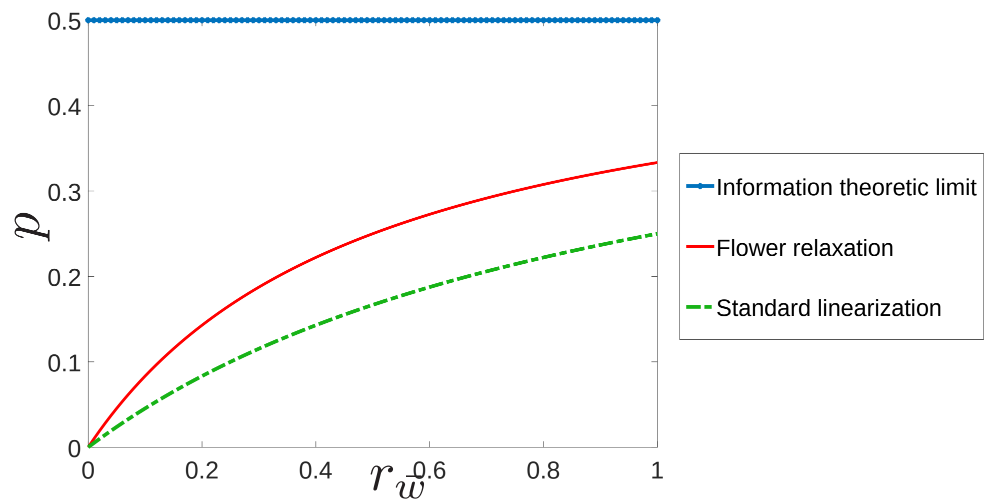
$$\lim_{n,m,l \to \infty} \frac{n+m+l}{\min\{nm, nl, ml\}} = 0.$$

  If p is a constant satisfying $p < 1/2$, then rank-one BTF recovers the ground truth with high probability.

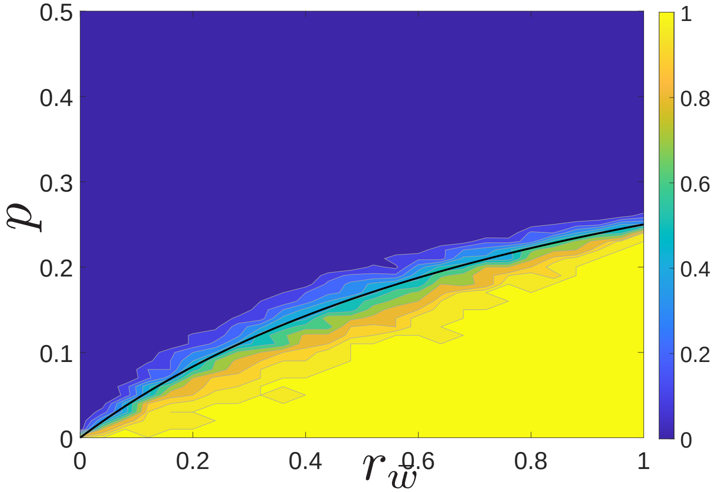
Recovery guarantees for LP relaxations

- Theorem: Assume that, as $n, m, l \to \infty$, we have $n \exp(-m)$, $n \exp(-l)$, $m \exp(-n)$, $m \exp(-l)$, $l \exp(-n)$, $l \exp(-m) \to 0$. If p is a constant satisfying $p < \frac{r_{\bar{w}}}{2(1+r_{\bar{w}})}$, then Problem (sLP) recovers the ground truth with high probability.

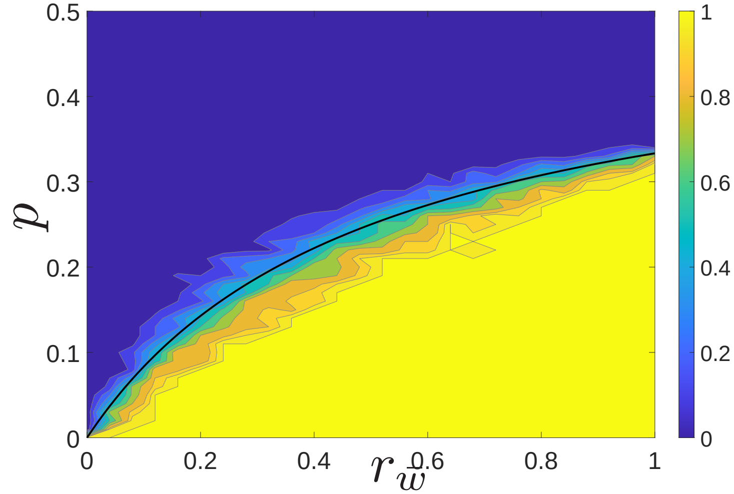
- Theorem: Assume that, as $n, m, l \to \infty$, we have $nml \exp(-n)$, $nml \exp(-m)$, $nml \exp(-l) \to 0$. If p is a constant satisfying $p < \frac{r_{\bar{w}}}{1+2r_{\bar{w}}}$, then Problem (fLP) recovers the ground truth with high probability.



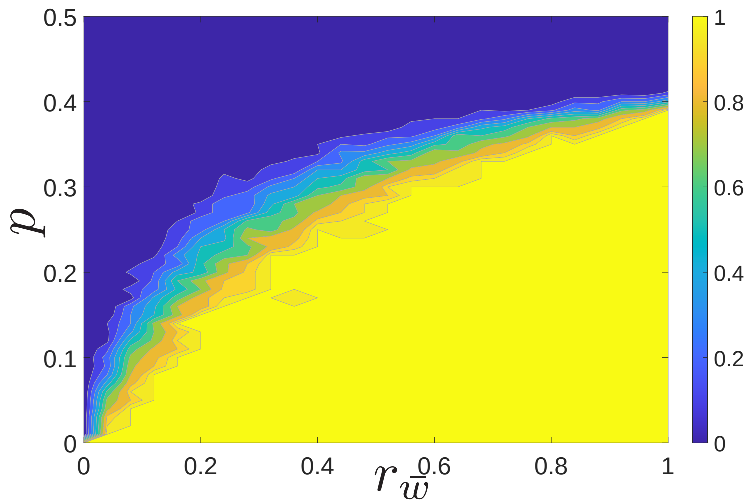
Numerical experiments



(a) Standard LP



(b) Flower LP



(c) Complete LP