Max-Product Belief Propagation for Linear Programming

Sejun Park and Jinwoo Shin (Department of Electrical Engineering, KAIST)

Introduction and Goal

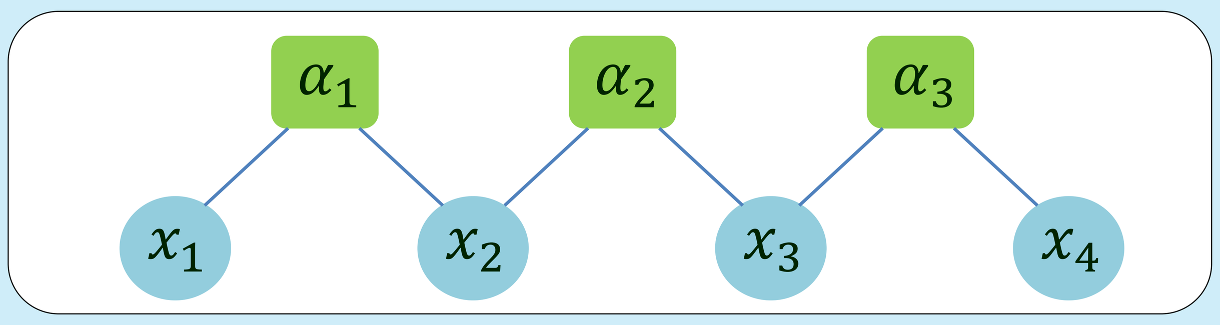
- **Max-product belief propagation (BP)** is a popular distributed heuristic for approximating a maximum a posteriori (MAP) assignment given a graphical model (GM).
- BP has shown remarkable performances in computer vision, speech recognition, error correcting code, natural language processing, etc. There have been made extensive research efforts to understand BP behind its empirical success.
- Several characterizations about BP fixed point has been proposed (Weiss et al. 2001, Vinyals et al. 2010).
- It has been studied about BP convergence to the correct answer under few classes loopy GM including shortest path (Ruozzi et al. 2008), matching (Sanghavi et al. 2008), perfect matching (Bayati et al. 2011) and network flow (Gamarnik et al. 2012). **However, prior proofs about BP convergence are highly rely on the problem setup and it cannot be extended to others.**
- We obtain a generic criteria that BP converges correctly under a relation with its associated Linear Programming (LP).

Belief Propagation and Graphical Model

A joint distribution of n-dimensional binary random vector $X = [X_i] \in \{0,1\}^n$ is called GM if it factorizes as follows:

$$\Pr(X = x) \propto \prod_{i \in [n]} \psi_i(x_i) \prod_{\alpha \in F} \psi_\alpha(x_\alpha)$$

where $\{\psi_i, \psi_\alpha\}$ are non-negative functions called factors and a collection $F = \{\alpha_1, \dots, \alpha_k\} \subset 2^{[n]}$ for $|\alpha_i| \geq 2$.



BP is an message passing algorithm for approximating the MAP assignment of GM. BP might oscillate or converge incorrectly.

Main Result

We consider the following LP:

$$\begin{aligned} &\textit{maximize} && w \cdot x \\ &\textit{subject to} && Ax \geq b \\ & && x = [x_i] \in [0,1]^n \end{aligned}$$

We can formulate the probability distribution using GM whose MAP assignment is a integral solution of LP

$$\Pr(X = x) = \prod_i e^{w_i x_i} \prod_\alpha \mathbf{1}_{A_\alpha x_\alpha \geq b_\alpha}$$

where x_α is a subvector of x with indices α and A_α / b_α is a corresponding submatrix/subvector of A/b.

Theorem [Park and Shin 2015]

Max-product belief propagation on GM converges to the solution of LP if the following conditions holds

C1: The optimal solution x^* of LP is unique and integral

C2: $|\{\alpha : i \in \alpha\}| \leq 2$ for all $i \in [n]$

C3: For every α, for every $x_\alpha \in \{0,1\}^{|\alpha|}$ with $A_\alpha x_\alpha \geq b_\alpha$ and every $i \in \alpha$ with $x_i \neq x_i^*$, there exists $\gamma \in \alpha$ such that

$$|\{j \in \{i\} \cup \gamma : |\{\alpha' \in F : j \in \alpha'\}| = 2\}| \leq 2$$

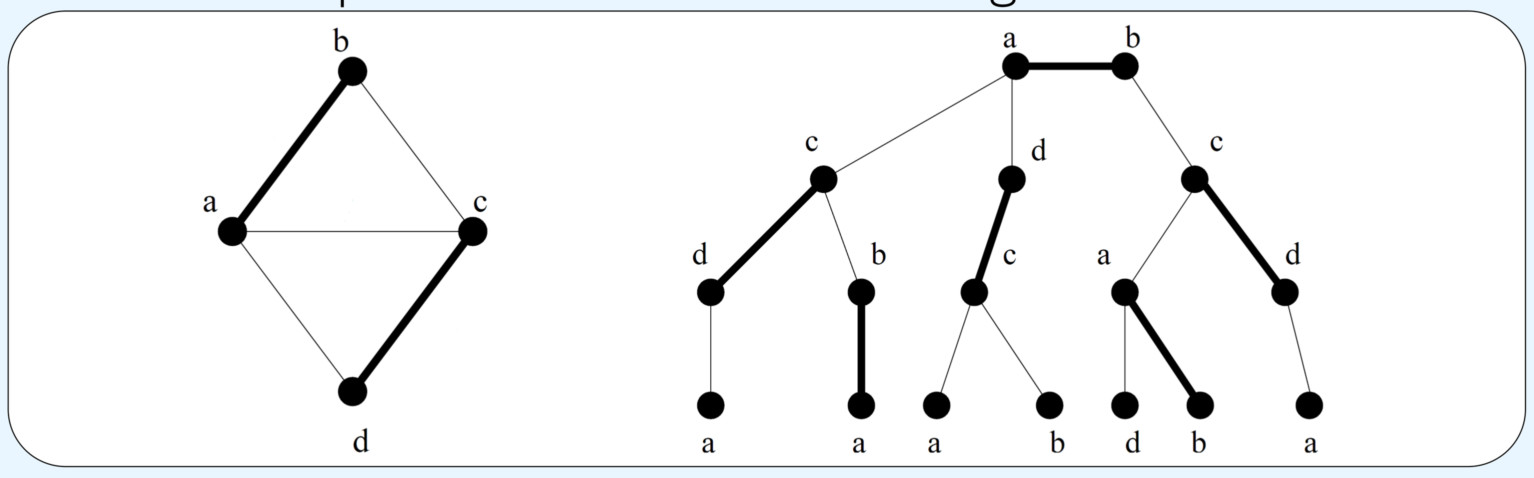
$$A_\alpha x'_\alpha \geq b_\alpha \text{ where } x'_k = \begin{cases} x_k & \textit{if } k \notin \{i\} \cup \gamma \\ x_k^* & \textit{otherwise} \end{cases}$$

$$A_\alpha x''_\alpha \geq b_\alpha \text{ where } x''_k = \begin{cases} x_k & \textit{if } k \in \{i\} \cup \gamma \\ x_k^* & \textit{otherwise} \end{cases}$$

C1, C2, C3 are easy to check given LP and GM !

Proof Sketch

- BP always find a MAP assignment if GM is tree structured.
- Instead of analyzing a loopy GM directly, we consider a tree structured GM called computational tree. The BP decision on the computational tree and the original GM are identical.



- We use the following procedure, introduced by (Sanghavi et al. 2008), for analyzing BP convergence.
- Suppose the BP decision for a variable i is different from the optimal solution of LP after large enough BP iterations.

  Step 1. Flip the root value to the optimum value of i.
  Step 2. Using C3, flip the values of children of the root so that the factors associate with the root are not violated.
  Step 3. Repeat Step 2 inductively by replacing the root to its child variable until it reaches to leaf or stop. Under C2, it forms an alternating path.
  Step 4. Show that BP decision on i is not identical to the MAP assignment in the computational tree. It requires to show that some assignment is a feasible solution of LP.

- In prior works, showing the feasibility in Step 4 highly relies on the problem setup and it cannot be generalized to others.
- **Our main contribution is to resolve the feasibility issue by handling a polytope of a feasible solution space of LP.**

Corollaries: LP Exmaples solvable by BP

Following LPs satisfy C1, C2 and C3, i.e., BP converges to their optimal solutions.

- Shortest path :
$$\begin{aligned} &\textit{minimize} && w \cdot x \\ &\textit{subject to} && \textstyle\sum_{e \in \delta^o(v)} x_e - \sum_{e \in \delta^i(v)} x_e = \begin{cases} 1 & \textit{if } v = s \\ -1 & \textit{if } v = t \\ 0 & \textit{otherwise} \end{cases} \\ & && x = [x_e] \in [0,1]^{|E|} \end{aligned}$$
- Perfect matching :
$$\begin{aligned} &\textit{maximize} && w \cdot x \\ &\textit{subject to} && \textstyle\sum_{e \in \delta(v)} x_e = 1 \quad \forall v \in V \\ & && x = [x_e] \in [0,1]^{|E|} \end{aligned}$$
- Network Flow :
$$\begin{aligned} &\textit{minimize} && w \cdot x \\ &\textit{subject to} && \textstyle\sum_{e \in \delta^o(v)} x_e - \sum_{e \in \delta^i(v)} x_e = d_v \quad \forall v \in V \\ & && x = [x_e] \in \mathbf{R}_+^{|E|} \end{aligned}$$
- Traveling salesman :
$$\begin{aligned} &\textit{minimize} && w \cdot x \\ &\textit{subject to} && \textstyle\sum_{e \in \delta(v)} x_e = 2 \quad \forall v \in V \\ & && x = [x_e] \in [0,1]^{|E|} \end{aligned}$$

Following LPs satisfy C2, C3, i.e., BP converges to their optimal solutions if C1 is satisfied

- Cycle packing :
$$\begin{aligned} &\textit{maximize} && w \cdot x \\ &\textit{subject to} && \textstyle\sum_{e \in \delta(v)} x_e = 2y_v \quad \forall v \in V \\ & && x = [x_e] \in [0,1]^{|E|}, y = [y_v] \in [0,1]^{|V|} \end{aligned}$$
- Vertex cover :
$$\begin{aligned} &\textit{maximize} && \textstyle\sum_e x_e \\ &\textit{subject to} && \textstyle\sum_{e \in \delta(v)} x_e \leq b_v \quad \forall v \in V \\ & && x = [x_e] \in [0,1]^{|E|} \end{aligned}$$

Conclusion

- In this work, we provide generic sufficient conditions that BP correctly converges to an optimal solution of LP. Our results not only reveal the BP success theoretically, but also provide new directions on efficient solvers for large-scale LPs.