STAT 206A: Gibbs Measures — Elchanan Mossel

Lecture 18

Lecture date: October 26 — *Scribe: Partha S. Dey*

In the previous lecture we defined the "loopy belief propagation" (LBP) algorithm for a factor graph G as the following iteration of messages. For every variable node v and every factor node f connected to v in G,

$$m_{v\to f}^{(t+1)}(x_v) = \frac{1}{Z}\prod_{\substack{f'\sim v\\ f'\neq f}} m_{f'\to v}^{(t)}(x_v), \tag{1}$$

$$m_{f\to v}^{(t+1)}(x_v) = \frac{1}{Z}\sum_{\substack{x_{w_1},x_{w_2},\ldots,x_{w_{k-1}}\\ f=(w_1,\ldots,w_{k-1},v)}} \varphi(x_{w_1},x_{w_2},\ldots,x_{w_{k-1}},x_v)\prod_{i=1}^{k-1} m_{w_i\to f}^{(t)}(x_{w_i}). \tag{2}$$

where the Z's are normalizing constants and the messages are initialized to $\{m_{v\to f}^{(0)}\}_{f\sim v}$. We also proved that if the factor graph is a tree, then for any set of initial messages the LBP algorithm converges to a unique set of messages $\{m_{v\to f}^{(\infty)}\}_{f\sim v}$ and the marginal at a variable node v is given by,

$$\mathbf{P}(x_v) = \frac{1}{Z}\prod_{f\sim v} m_{f\to v}^{(\infty)}(x_v),$$

In this lecture we are going to prove convergence of LBP algorithm under some conditions for graphs with cycles.

1 Self-repetition avoiding tree

Definition 1 *Given a factor graph G and a variable node $v\in V(G)$ the self-repetition avoiding tree (SRAT) (T,v) of G rooted at v, is defined as the infinite factor tree consisting of all paths in the factor graph G starting at $v_0=v$, $(v_0,f_1,v_2,f_3,v_4,\ldots)$ that never backtrack,* i.e. $(v_i,f_{i+1}),(f_i,v_{i+1})$ *are connected in G and $v_i\neq v_{i+2}$ and $f_i\neq f_{i+2}$ for all i.*

Figure one shows an example of a self-repetition avoiding tree (upto depth four) rooted at variable node v. In the figure we have labelled each node in the tree with the associated node in the original graph. The following exercises easily follow from the definition of SRAT.

Exercise 2 (1 pt) *Prove that (T,v) is finite iff G is a tree.*

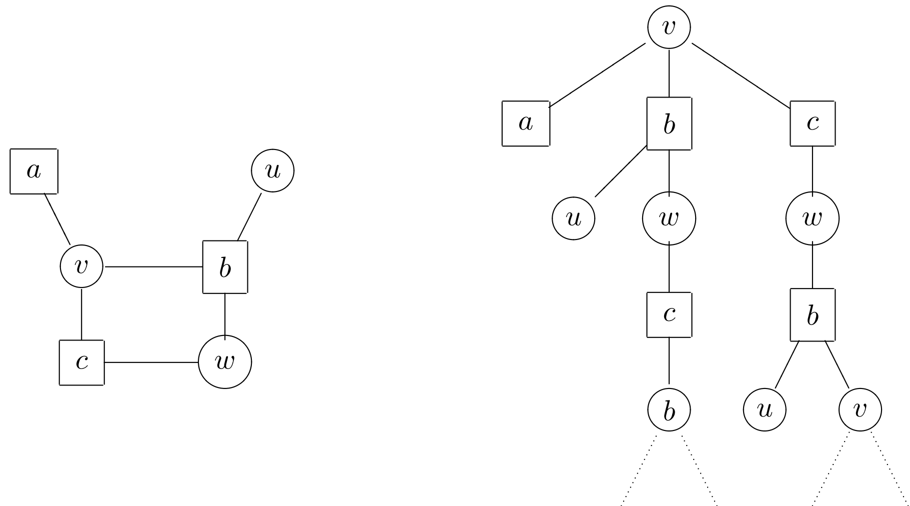

Figure 1: Example of a self-repetition avoiding tree.

Exercise 3 (0.5 pt) *Prove that if all variable nodes and factor nodes have degree* ≥ 2*, then the self-repetition avoiding tree has no leaves.*

Exercise 4 (1 pt) *Show that the tree (T,v) is periodic in the following sense: For every edge* (u, f) *directed away from the root* v*, the tree below* f *is always the same.*

Claim 1 (Tatikonda-Jordon [1]) *The marginal at the factor node* v *calculated by LBP after* t *iterations equals the marginal of* v *at the tree* (T, v) *where all variable nodes* w *at distance exactly* t *from* v *in* (T, v) *are conditioned to have the value* $m^{(0)}_{w\to f}(x_w)$, i.e. $\{X_w\}$*'s are conditioned to have the product measure with marginals* $m^{(0)}_{w\to f}(x_w)$.

Proof: The proof follows easily from the iterative relation given in equation (1) & (2) . □

Corollary 5 *if (T,v) has uniqueness then the messages* $(m^{(t)}_{v\to f}, m^{(t)}_{f\to v})$ *converge to limit* $(m^{(\infty)}_{v\to f}, m^{(\infty)}_{f\to v})$ *in total variation distance and the limiting messages do not depend on the initial messages* $\{m^{(0)}_{v\to f}\}$.

Proof: **[Idea of the proof]** Uniqueness of the tree implies that marginal in any ball $B(v, l)$ converges to the same value independent of the initial values. Now the proof follows from the following exercise. □

Exercise 6 (2 pts) *Show that convergence of marginals implies the convergence of messages.*

Corollary 7 *If (T, v) has exp-uniqueness then after t iterations we have*

$$\left\| m_{v\to f}^{(t)} - m_{v\to f}^{(\infty)} \right\|_{TV} \le e^{-\Omega(t)}.$$

Exercise 8 (2 pts) *For any periodic tree (T, v) which is not coming from a cyclic factor graph, there exists (soft-)potentials such that uniqueness does not hold.*

Corollary 9 *Suppose $v \in V(G)$ is not contained in any cycle of length $\le l$ and that (T, v) has exp-uniqueness with parameter $(1-\epsilon)$, then LBP will calculate the marginals at v after l iterations with at most $(1-\epsilon)^l$ error in total variation distance.*

Proof: Expressing the marginal at v as a linear combination of the conditional distributions of X_v given $X_{B(v,l)}$, we have the true marginal at v

$$\mathbf{P}(X_v = x) = \sum_i \alpha_i \mathbf{P}(X_v = x | X_{B(v,l)} = \sigma_i) \tag{3}$$

and the marginal at v calculated by LBP after l iterations

$$\mathbf{P}^{(l)}(X_v = x) = \sum_j \beta_j \mathbf{P}(X_v = x | X_{B(v,l)} = \sigma_j). \tag{4}$$

where $\sum_i \alpha_i = \sum_j \beta_j = 1$ and β_i depends on the initial messages $\{m_{w\to f}^{(0)}\}$.

Now by exp-uniqueness all probabilities in (3) and (4) differ in total variation distance by at most $(1-\epsilon)^l$. *i.e.*

$$\left\| \mathbf{P}(X_v = x | X_{B(v,l)} = \sigma_i) - \mathbf{P}(X_v = x | X_{B(v,l)} = \sigma_j) \right\|_{TV} \le (1-\epsilon)^l \quad \forall\ i, j.$$

and

$$\mathbf{P}(X_v = x) - \mathbf{P}^{(l)}(X_v = x) = \sum_{i,j} \alpha_i \beta_j \left[\mathbf{P}(X_v = x | X_{B(v,l)} = \sigma_i) - \mathbf{P}(X_v = x | X_{B(v,l)} = \sigma_j) \right].$$

This completes the proof. □

2 Discussion

The results above still leave many open problem about the behavior of LBP for graphs with cycles. First of all, uniqueness in (T, v) is too strong for convergence. One subtle difference is that by construction, the initial messages are periodic. Moreover, in practice the initial messages are also random.

In fact for the well understood case of coding theory uniqueness arguments are almost never applicable as each codeword defines a different Gibbs measure on the infinite tree.

References

[1] S. Tatikonda and M. Jordan, "Loopy belief propagation and Gibbs measures", *Proc. Uncertainty in Artificial Intell.*, 18, pp. 493–500, 2002.