Survey of some new convergence analysis techniques for sum and max belief propagation on loopy graphs

Alekh Agarwal

December 13, 2007

Basic setup

- Pairwise MRFs on binary variables
- Sum-product and max-product inference algorithms
- Associated computation tree – possibly infinite
- Can we truncate this tree?

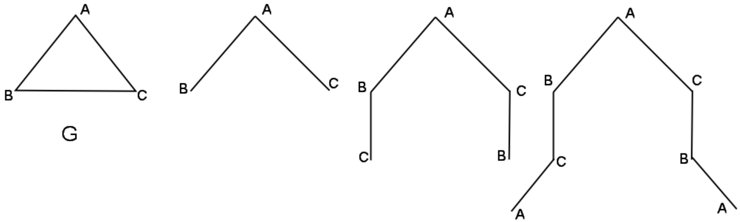

Figure: Sample graph and computation tree at first four time steps

Exact inference and Self-Avoiding Walks ([Weitz, 2006])

- Self-avoiding walk tree - a truncated computation tree.

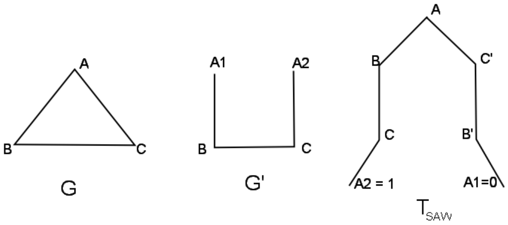

Figure: Self-Avoiding Walk tree for a toy graph

- A node occuring for a second time on a path behaves as a 0 or 1.
- Values fixed consistently with some scheme.
- Potentially exponential in size.
- Gives exact marginal and max-marginal at root.

Truncating the SAW tree

- Can approximate sum marginal arbitrarily well under certain conditions.
- Key idea - correlation decay or spatial mixing.
- Influence of a node on others should decay fast with distances in the graph.
- Can truncate the tree at a polynomial number of nodes (*i.e.* O(log n depth)) when decay is exponential.
- Post-truncation, just doing a computation tree. SAW tree an analytical tool.
- Also extended to multi-spin and non-pairwise systems.

Message passing and contractions ([Roosta, 2007], [Mooij and Kappen, 2007])

- Message passing iteratively solving fixed point system.
- Define $\nu = \log \frac{\mu(1)}{\mu(0)} \to$ log likelihood ratio.
- Message passing is a mapping $\tilde{\nu} = F(\nu)$.
- Can use Banach's fixed point theorem when F is a contraction.

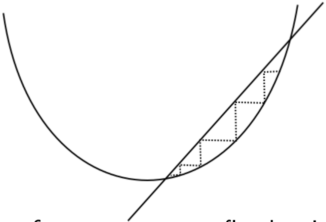

Figure: Illustration of convergence to fixed point for a contraction

- First analysis: guarantee $|F'(z)| < 1 \ \forall z$.
- Tighter analysis: guarantee $|F'(z)| < 1$ in range of F.

Analysis of max-product

- Above analyses don't carry over to max-product!
- Max-product inherently more unstable.

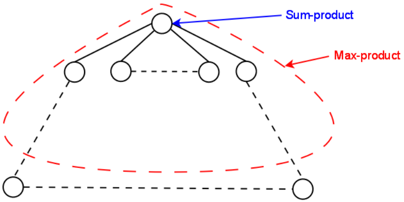

Figure: Illustration of what makes Max-product unstable

- Looking at the probability of a joint configuration, not that of a single node.
- No phenomenon of mixing happening on computation tree.
- Can try extending sum-product analysis via simulated annealing based approaches.
- Hard to prove convergence here.

Analysis of max-product

- Can write max-product as an integer LP.

$$\max_{\mathbf{x} \in \{0,1\}^n} \langle \theta, \phi(x) \rangle \equiv \max_{\mu \in \textsc{Marg}(G)} \langle \theta, \mu \rangle$$

- TRW is an LP relaxation [Kolmogorov and Wainwright, 2005].

$$\max_{\mu \in \textsc{Local}(G)} \langle \theta, \mu \rangle$$

- Better than a lot of SOCP relaxations.
- Clearly tight for tree structured constraints.
- Tight for submodular potentials via reduction to network flows [Boros and Hammer, 2002].
- Can we say something on average for near submodular potentials?

Near-submodularity and network flows

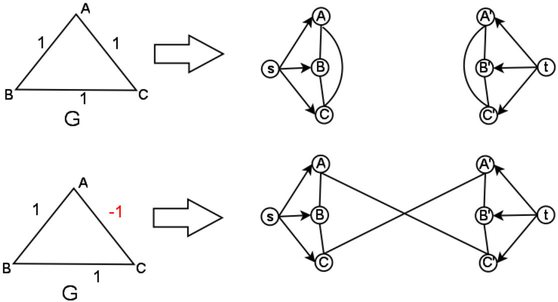

Figure: Flow networks for a submodular and a non-submodular problem

- Suppose all 1's is the MAP assignment.
- TRW finds max flow in networks shown above.
- The flow network has no cross-edges in the submodular problem, hence all 1's is trivially optimal min-cut too.
- Cross-edges in a non-submodular problem make analysis harder.

Some sufficient conditions and future work

- Can obtain simple sufficient conditions like $\theta_s \geq \sum_{x,y'} \theta_{xy'} \quad \rightarrow$ sum over cross-edges.
- Captures the effect of local fields.
- Better analysis gives $\theta_s + \sum_{u \in N^+(s)} \min \left\{ \theta_{su}, \theta_u \frac{\theta_{su}}{\sum_t \theta_{ut}} \right\} \geq \sum_{x,y'} \theta_{xy'} \quad \rightarrow$ sum over cross-edges.
- Allows you to borrow from rich neighbors!
- Future work to obtain sharper sufficient conditions.
- Do an average-case analysis for these conditions on random graphs.
- **Acknowledgement:** Would like to thank Prof. Martin Wainwright for his constant advice during this work.

References

Boros, E. and Hammer, P. L. (2002).
Pseudo-boolean optimization.
Discrete Applied Mathematics, 123(1-3):155–225.

Kolmogorov, V. and Wainwright, M. (2005).
On the optimality of tree-reweighted max-product message-passing.
In *Proceedings of the 21th Annual Conference on UAI 2005*.

Mooij, J. M. and Kappen, H. J. (2007).
Sufficient conditions for convergence of the sum-product algorithm.
arXiv preprint.

Roosta, T.; Wainwright, M. S. S. (15-20 April 2007).
Convergence analysis of reweighted sum-product algorithms.
ICASSP, 2:II–541–II–544.

Weitz, D. (2006).
Counting independent sets up to the tree threshold.
In *STOC '06: Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 140–149, New York, NY,