Simplicity and Complexity of Belief-Propagation

Elchanan Mossel[1]

[1]MIT

Dec 2020

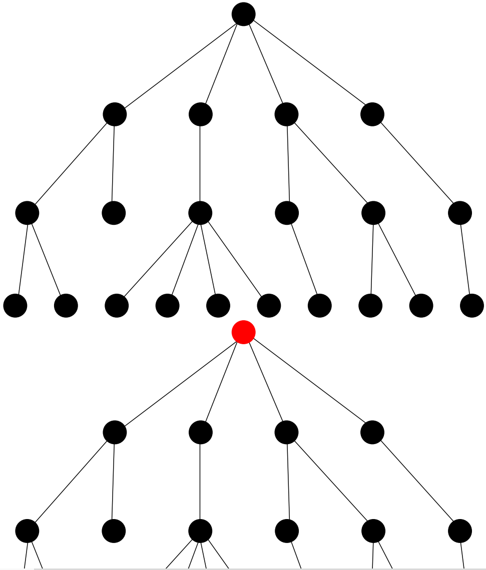
Markov Random Fields and Information Flow on Trees

Consider the following process on a tree.

Color the root randomly.

Repeat: Copy color of parent with probability θ. Otherwise, chose color $\sim U[q]$.

Will mostly consider full d-ary tree / Branching process trees.



Inference - Machine Learning Perspective

- Inference problem (V1): Infer root color from leaf colors?
- A: No! Even in one level, $\exists$ randomness in root given children.
- Inference Problem (V2): How much can infer on the root color from leaf colors?
- Machine Learning: We can compute the posterior exactly!
- Moreover: Belief-Propagation does it in linear time.
- Remark: Belief Propagation is often applied to non-tree graphs [Pearl 82]. Applications to trees where known in biology and statistical physics before [Hidden 1970, Preston 1974].

Part 1: LINEAR THEORY

$q = 2$

Inference - Asymptotic Perspective

- Let $q = 2$, fix the d-ary tree of h levels and call the two colors $+1, -1$. Let X_v be the color of node v. Let X_0 denote the root color and X_h are labels at level h of the tree.
- Inference Problem (V2): How much can infer on the root color from leaf colors?
- Q1: Can we analyze the optimal estimator (BP)?
- Q2: Is Majority $= sgn(\sum_i X_h(i))$ a good estimator as $h \to \infty$?
- Question asked in Statistical Physics in terms of the extremality of the free measure of the Ising/Potts model on the Bethe lattice.

The Majority Estimator

- Let $S_h = \sum_i X_h(i)$.
- Exercise 1. $\mathbb{E}[S_h|X_0] = (d\theta)^h X_0$ and

$$\frac{\mathbb{E}[S_h|X_0]^2}{Var[S_h|X_0]} = \frac{(d\theta)^{2h}}{Var[S_h|X_0]} \to \left\{ \begin{array}{ll} C(\theta), & d\theta^2 > 1, \\ 0, & d\theta^2 \le 1 \end{array} \right\}$$

- $\implies \lim_{h\to\infty} d_{\mathrm{TV}}(S_h|X_0 = +, S_h|X_0 = -) > 0$ if $d\theta^2 > 1$.
- $\implies \lim_{h\to\infty} \mathbb{E}[sgn(S_h)X_0] > 0$ if $d\theta^2 > 1$.
- Analyzing Fourier Transform of S_h Kesten-Stigum-66 proved:
- $d\theta^2 \le 1 \implies S_h \to_{h\to\infty}$ a normal law independent of X_0 $(*)$.
- $d\theta^2 > 1 \implies S_h \to_{h\to\infty}$ a non-normal law dependent on X_0.
- $d\theta^2 = 1$ is referred to as the Kesten-Stigum threshold.
- Exercise: Apply the martingale CLT, to prove the normal case.

Perspectives United:

- Thm for $q = 2$: If $d\theta^2 \leq 1$ then $\mathbb{P}[BP \to_{h\to\infty} (0.5, 0.5)] = 1$.
- $\implies$ BP infers non-trivially **iff** Majority infers non-trivially.
- Multiple proofs: Bleher, Ruiz, and Zagrebnov (95), Ioffe (96), Evans-Kenyon-Peres-Schulmann (00), Borg, Tour-Chayes, M, Roch (06), etc. (also: Chayes, Chayes, , Sethna, Thouless, (1986)).
- EKPS: Also for random trees, where d is the average degree.
- All use some *concavity* of the functionals of the distribution.
- Next: A proof sketch and applications areas of the linear theory.

Recursion of Random Variables for Binary Tree

- Let P_T^+ denote the measure of X_h when root is $+$, let $P^T = 0.5(P_T^- + P_T^+)$.
- $M := M_T := P_T[X_0 = +|X_h] - P_T[X_0 = -|X_h]$.
- <u>Ex:</u> (Bayes): $\frac{dP_T^\pm}{dP_T} = 1 \pm M, \quad E_T^+[M] = E_T[M^2] = E_T^+[M^2]$.
- <u>Claim 1:</u> If $T = 0 -> S, M = M_T, N = M_S$:
- $M = \theta N,\ E_T^+[N] = \theta E_S^+[N],\ E_T^+[N^2] = \theta E_S^+[N^2] + (1-\theta)E_S[N^2]$.
- <u>Claim 2:</u>: If T_1, T_2 are two trees joined at the root to form T and $N_i = X_{T_i}$ then:

$$M = \frac{N_1 + N_2}{1 + N_1 N_2}.$$

$$\implies \text{Belief Propagation Recursion} : M_{n+1} = \theta \frac{M_n + M_n'}{1 + \theta^2 M_n M_n'}.$$

$2\theta^2 < 1 \implies E[M_n^2] \to 0$, i.e., asymptotic indpendence

$$M = \theta \frac{N_1 + N_2}{1 + \theta^2 N_1 N_2}, \quad \frac{1}{1+r} = 1 - r + \frac{r^2}{1+r} \implies$$

$$M = \theta(N_1 + N_2) - \theta^3 N_1 N_2 (N_1 + N_2) + \theta^4 N_1^2 N_2^2 M$$

$$M \leq \theta(N_1 + N_2) - \theta^3 N_1 N_2 (N_1 + N_2) + \theta^4 N_1^2 N_2^2$$

$\implies$ (taking E_T^+ recalling $E_T^+[M] = E_T[M^2]$)

$$E_T[M^2] \leq 2\theta^2 E_S[N^2] - \theta^4 E_{S^+}[N^2]$$

$\implies$

$$E[M_n^2] \leq (2\theta^2)^n$$

Application 1: The Phylogenetic Inference Problem