Open problem session

Berkeley students in probability

April 3, 2024

Contents

1 Reconstruction Problem (Vilas Winstein) **1**

2 Markov chains on triangle-rich graphs (Mriganka Basu Roy Chowdhury) **1**

3 Hegselmann-Krause model (Adam Quinn Jaffe) **2**

1 Reconstruction Problem (Vilas Winstein)

- A transition matrix M on a finite state-space, and sample path $X_1 \to X_2 \to X_3 \cdots \to X_n$. If you see X_n, what can you say about X_1 ?
- But what if you have a tree, where each X_i spawns d i.i.d. copies of X_{i+1} according to (one-step of) the Markov chain started at X_i.
- Given *all the leaves*, can you tell what the root is?
- If

$$M = \begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix},$$

where p encodes the temperature and the probabilities of transferring to the other state, what is the dependence on p of the root reconstructability?

- Let λ be the second eigenvalue of M. Kesten-Stigum shows that $\lambda^2 d > 1$, the root is reconstructable. Here d is the branching factor of the tree.
- On the other hand, if $\lambda d < 1$, reconstruction is impossible. The question is what happens when $\lambda^2 d < 1 < \lambda d$.
- The problem above is called the *symmetric binary channel*. For this model, Kesten-Stigum is known to be tight, i.e. reconstruction is impossible if $\lambda^2 d < 1$.
- Now, you may generalize this to something like the q-Potts model. Then Kesten-Stigum is *not sharp* for $q > 4$ (Mossel). That is, reconstruction is possible below the Kesten-Stigum threshold.
- **Question:** What happens if you take the XY-model, where the spins are on the unit circle, and the probability of a different spin is proportional to $f(\sigma') \propto \exp(\langle \sigma', \sigma \rangle)$.

2 Markov chains on triangle-rich graphs (Mriganka Basu Roy Chowdhury)

- Let $\mathcal{G}(n, p)$ be the usual random graph. The core theme is to consider mixing on $\mathcal{G}(n, p)$ conditioned to have many triangles (although this is only a prototype tilt). The chain here is given by Glauber dynamics.
- The simplest instance is given a uniformly chosen graph with $\geq n^3 r^3/6$ triangles where $r > 1/2$.
- To this end, it is perhaps helpful to remember the typical behavior (or stationary for our chain) behavior here, which is that of $\mathcal{G}(n, p)$ conditioned to have $\geq n^3 r^3/6$ triangles where $r > p$.

- The description here is the result of substantial progress in nonlinear large deviations, yielding that the typical conditional behavior is actually of $\mathcal{G}(n, r)$, i.e., each edge individually picks up more probability, for most values of r and p. This regime is called replica-symmetric.
- However there is a small region $p < p_c$ and $r_c^-(p) < r < r_c^+(p)$ for critical p_c and $r_c^-(p), r_c^+(p)$ (the latter interval decreasing with p), where the symmetry is broken, and instead the dominant mechanism is via *clustering*. We certainly expect these aspects to show up in the mixing behavior.
- A related question is that of the mixing of exponential random graph model (ERGM), where the measure is given by $\mu^\beta(G) \propto \exp(\beta\Delta(G))$ (say) where Δ is the number of triangles. It is known this is fast mixing at high temperature, and also that (as established recently), that at low temperature, fast mixing happens in *metastable wells*. This is in line with the heuristic that many slow mixing Glauber dynamics mix fast in locally optimal wells, and then take a lot of time to move between wells.
- Our program would look at more general Hamiltonians. For instance the above question is the extreme case of $H = 0$ when the graph has enough triangles, and $-\infty$ elsewhere.
- A related question is that of restricting the graph to have between $n^3r^3/6$ to $n^3r^3/6 + f(n)$ triangles, where $f(n) \ll n^3$, i.e., a narrow band. A basic starting point might be to answer for what $f(n)$ is the Glauber dynamics irreducible? The harder question is to answer mixing. Observe that the value of the Hamiltonians here is not that important since the number of triangles are essentially the same throughout.

3 Hegselmann-Krause model (Adam Quinn Jaffe)

- The general discipline is *opinion dynamics*.
- *Hegselmann-Krause model*: n individuals $1, \dots, n$, each of them has an opinion X_i. Let us assume that the opinions are real valued.
- Individuals opinions are placed i.i.d. on $\mathbb{R}$, suffices to take bounded.
- For every X_i, draw a ball of radius ε (which is another parameter), and X_i moves to the average of all the X_j that falls in the ball. These moves all happen simultaneously.
- This can be thought of applying a kernel average on the dataset.
- The individuals move closer together, until they all are in a ball of radius ε, and then they coalesce. This is called the *consensus time*. This need not happen for all starting configurations.
- Formally we start with $X_1(0), \dots, X_n(0)$ and $X_i(t) = \frac{1}{\#N_i(t)} \sum_{j \in N_i(t)} X_j(t-1)$ where $N_i(t) = \{j : X_j(t-1) \in B_\varepsilon(X_i(t-1))\}$.
- The main questions are, *when does consensus occur*, and *what is its probability*?
- You can also do this in other geometries where averaging is defined, and ask the same questions.
- You can also replace the kernel $\mathbf{1}_{[-\varepsilon,\varepsilon]}$ with a smoother kernel of bandwidth ε. Consensus might need to be redefined or considered at infinity.
- It is known that if the initial points comes from some density, then you can lower bound the probability of consensus by properties of the density. Most things known about it are quite simple (arxiv – physics and social dynamics).