Modern Discrete Probability

I - Introduction

Review and Some Fundamental Models

Sébastien Roch
UW–Madison
Mathematics

September 25, 2014

1. **Preliminaries**
    - Review of graph theory
    - Review of Markov chain theory

2. Some fundamental models
    - Random walks on graphs
    - Percolation
    - Some random graph models
    - Markov random fields
    - Interacting particles on finite graphs

Graphs

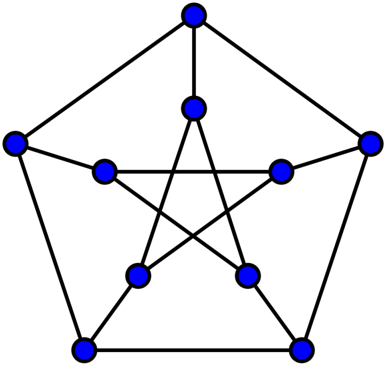
Definition (Undirected graph)

An *undirected graph* (or graph for short) is a pair $G = (V, E)$ where V is the set of *vertices* (or nodes, sites) and

$$E \subseteq \{\{u, v\} : u, v \in V\},$$

is the set of *edges* (or bonds). The V is either finite or countably infinite. Edges of the form $\{u\}$ are called *loops*. We do *not* allow E to be a multiset.

An example: the Petersen graph



Basic definitions

A vertex $v \in V$ is *incident* with an edge $e \in E$ if $v \in e$. The incident vertices of an edge are sometimes called *endvertices*. Two vertices $u, v \in V$ are *adjacent*, denoted by $u \sim v$, if $\{u, v\} \in E$. The set of adjacent vertices of v, denoted by $N(v)$, is called the *neighborhood* of v and its size, i.e. $\delta(v) := |N(v)|$, is the *degree* of v. A vertex v with $\delta(v) = 0$ is called *isolated*. A graph is called d*-regular* if all its degrees are d. A countable graph is *locally finite* if all its vertices have a finite degree.

Example

All vertices in the Petersen graph have degree 3, i.e., it is 3-regular. In particular there is no isolated vertex.

Adjacency matrix

Let $G = (V, E)$ be a graph with $n = |V|$. The *adjacency matrix* A of G is the $n \times n$ matrix defined as $A_{xy} = 1$ if $\{x, y\} \in E$ and 0 otherwise.

Example

The adjacency matrix of a *triangle* (i.e. 3 vertices with all non-loop edges) is

$$\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

Paths, cycles, and spanning trees I

Definition (Subgraphs)

A *subgraph* of $G = (V, E)$ is a graph $G' = (V', E')$ with $V' \subseteq V$ and $E' \subseteq E$. The subgraph G' is said to be *induced* if

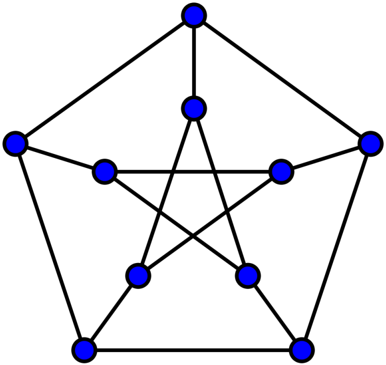
$$E' = \{\{x, y\} \, : \, x, y \in V', \; \{x, y\} \in E\},$$

i.e., it contains all edges of G between the vertices of V'. In that case the notation $G' := G[V']$ is used. A subgraph is said to be *spanning* if $V' = V$. A subgraph containing all non-loop edges between its vertices is called a *complete subgraph* or *clique*.

Example

The Petersen graph contains no triangle, induced or not.

An example: the Petersen graph



Paths, cycles, and spanning trees II

A *path* in G (usually called a “walk” but that term has a different meaning in probability) is a sequence of (not necessarily distinct) vertices $x_0 \sim x_1 \sim \cdots \sim x_k$. The number of edges, k, is called the *length* of the path. If the *endvertices* x_0, x_k coincide, i.e. $x_0 = x_k$, we call the path a *cycle*. If the vertices are all distinct (except possibly for the endvertices), we say that the path (or cycle) is *self-avoiding*. A self-avoiding path or cycle can be seen as a (not necessarily induced) subgraph of G. We write $u \leftrightarrow v$ if there is a path between u and v. Clearly $\leftrightarrow$ is an equivalence relation. The equivalence classes are called *connected components*. The length of the shortest self-avoiding path connecting two distinct vertices u, v is called the *graph distance* between u and v, denoted by $\rho(u, v)$.

Paths, cycles, and spanning trees III

Definition (Connectivity)

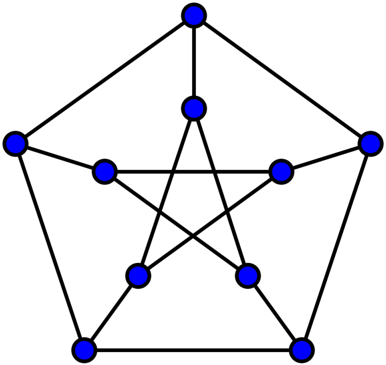
A graph is *connected* if any two vertices are linked by a path, i.e., if $u \leftrightarrow v$ for all $u, v \in V$. Or put differently, if there is only one connected component.

Example

The Petersen graph is connected.

A *forest* is a graph with no self-avoiding cycle. A *tree* is a connected forest. Vertices of degree 1 are called *leaves*. A *spanning tree* of G is a subgraph which is a tree and is also spanning.

An example: the Petersen graph



Examples of finite graphs

- Complete graph K_n
- Cycle C_n
- Rooted b-ary trees $\widehat{\mathbb{T}}_b^\ell$
- Hypercube $\{0,1\}^n$

Examples of infinite graphs

- d-ary tree $\mathbb{T}_d$
- Lattice $\mathbb{L}^d$

Transitive graphs

Definition (Graph automorphisms)

An *automorphism* of a graph $G = (V, E)$ is a bijection ϕ of V to itself that preserves the edges, i.e., such that $\{x, y\} \in E$ if and only if $\{\phi(x), \phi(y)\} \in E$. A graph $G = (V, E)$ is *vertex-transitive* if for any $u, v \in V$ there is an automorphism mapping u to v.

Example

Any "rotation" of the Petersen graph is an automorphism.

Example

$\mathbb{T}_d$ is vertex-transitive. $\widehat{\mathbb{T}}_b^\ell$ has many automorphisms but is not vertex-transitive.

Flows I

Definition (Flow)

Let $G = (V, E)$ be a connected graph with two distinguished, distinct vertex sets, a *source-set* $A \subseteq V$ and a *sink-set* Z. Let $c : E \to \mathbb{R}_+$ be a *capacity* function. A *flow* on the *network* (G, c) from source A to sink Z is a function $f : V \times V \to \mathbb{R}$ such that:

F1 *(Antisymmetry)* $f(x, y) = -f(y, x), \forall x, y \in V$.

F2 *(Capacity constraint)* $|f(x, y)| \leq c(e), \forall e = \{x, y\} \in E$, and $f(x, y) = 0$ otherwise.

F3 *(Flow-conservation constraint)*

$$\sum_{y : y \sim x} f(x, y) = 0, \qquad \forall x \in V \backslash (A \cup Z).$$

Flows II

For $U, W \subseteq V$ and $F \subseteq E$, let $f(U, W) := \sum_{u \in U, w \in W} f(u, w)$ and $c(F) := \sum_{e \in F} c(e)$. The *strength* of f is $|f| := f(A, A^c)$.

Definition (Cutset)

Let $F \subseteq E$. We call F a *cutset* separating A and Z if all paths connecting A and Z include an edge in F. Let A_F be the set of vertices not separated from A by F, and similarly for Z_F.

Lemma (Max flow $\leq$ min cut): For any cutset F, $|f| \leq c(F)$.

Proof: $f(A, A^c) \stackrel{(F3)}{=} f(A, A^c) + \sum_{u \in A_F \setminus A} f(u, V) \stackrel{(F1)}{=} f(A_F, A_F^c) \stackrel{(F2)}{\leq} c(F)$. ■

Flows III

Theorem (Max-Flow Min-Cut Theorem)

$$\max\{|f| \,:\, \textit{flow } f\} = \min\{c(F) \,:\, \textit{cutset } F\}.$$

Proof: Let f be an optimal flow. (The sup is achieved by compactness.) An *augmentable path* is a self-avoiding path $x_0 \sim \cdots \sim x_k$ with $x_0 \in A$, $x_i \notin A \cup Z$ for all $i \neq 0, k$, and $f(x_{i-1}, x_i) < c(\{x_{i-1}, x_i\})$ for all i. By optimality of f there cannot be such a path with $x_k \in Z$, otherwise we could push more flow through it. Let $B \subseteq V \backslash (A \cup Z)$ be the set of all possible final vertices in an augmentable path. Let F be the edge set between B and B^c. Note that $f(x, y) = c(e)$ for all $e = \{x, y\} \in F$ with $x \in B$ and $y \in B^c$, and that F is a cutset. So we have equality in the previous lemma with $B = A_F$. ∎

Directed graphs

Definition

A *directed graph* (or digraph for short) is a pair $G = (V, E)$ where V is a set of *vertices* (or nodes, sites) and $E \subseteq V^2$ is a set of *directed edges*.

A *directed path* is a sequence of vertices $x_0, \ldots, x_k$ with $(x_{i-1}, x_i) \in E$ for all $i = 1, \ldots, k$. We write $u \to v$ if there is such a path with $x_0 = u$ and $x_k = v$. We say that $u, v \in V$ *communicate*, denoted by $u \leftrightarrow v$, if $u \to v$ and $v \to u$. The $\leftrightarrow$ relation is clearly an equivalence relation. The equivalence classes of $\leftrightarrow$ are called the *(strongly) connected components* of G.

Markov chains I

Definition (Stochastic matrix)

Let V be a finite or countable space. A *stochastic matrix* on V is a nonnegative matrix $P = (P(i,j))_{i,j \in V}$ satisfying

$$\sum_{j \in V} P(i,j) = 1, \qquad \forall i \in V.$$

Let μ be a probability measure on V. One way to construct a *Markov chain* (X_t) on V with transition matrix P and initial distribution μ is the following. Let $X_0 \sim \mu$ and let $(Y(i,n))_{i \in V, n \geq 1}$ be a mutually independent array with $Y(i,n) \sim P(i,\cdot)$. Set inductively $X_n := Y(X_{n-1}, n)$, $n \geq 1$.

Markov chains II

So in particular:

$$\mathbb{P}[X_0 = x_0, \ldots, X_t = x_t] = \mu(x_0)P(x_0, x_1)\cdots P(x_{t-1}, x_t).$$

We use the notation $\mathbb{P}_x, \mathbb{E}_x$ for the probability distribution and expectation under the chain started at x. Similarly for $\mathbb{P}_\mu, \mathbb{E}_\mu$ where μ is a probability measure.

Example (Simple random walk)

Let $G = (V, E)$ be a finite or countable, locally finite graph. *Simple random walk* on G is the Markov chain on V, started at an arbitrary vertex, which at each time picks a uniformly chosen neighbor of the current state.

Markov chains III

The *transition graph* of a chain is the directed graph on V whose edges are the transitions with nonzero probabilities.

Definition (Irreducibility)

A chain is *irreducible* if V is the unique connected component of its transition graph, i.e., if all pairs of states communicate.

Example

Simple random walk on G is irreducible if and only if G is connected.

Markov property I

Let (X_t) be a Markov chain with transition matrix P and initial distribution μ. Let $\mathcal{F}_t = \sigma(X_0, \ldots, X_t)$. A fundamental property of Markov chains known as the *Markov property* is that, given the present, the future is independent of the past. In its simplest form: $\mathbb{P}[X_{t+1} = y \,|\, \mathcal{F}_t] = \mathbb{P}_{X_t}[X_{t+1} = y] = P(X_t, y)$. More generally, let $f : V^\infty \to \mathbb{R}$ be bounded, measurable and let $F(x) := \mathbb{E}_x[f((X_t)_{t\geq 0})]$, then (see [D, Thm 6.3.1]):

Theorem (Markov property)

$$\mathbb{E}[f((X_{s+t})_{t\geq 0}) \,|\, \mathcal{F}_s] = F(X_s) \qquad a.s.$$

We will come back to the “strong” Markov property later.

Markov property II

Let (X_t) be a Markov chain with transition matrix P. We define $P^t(x,y) := \mathbb{P}_x[X_t = y]$.

Theorem (Chapman-Kolmogorov)

$P^t(x,z) = \sum_{y \in V} P^s(x,y) P^{t-s}(y,z), \qquad s \in \{0,1,\ldots,t\}.$

Proof: $\mathbb{P}_x[X_t = z \,|\, \mathcal{F}_s] = F(X_s)$ with $F(y) := \mathbb{P}_y[X_{t-s} = z]$ and take $\mathbb{E}_x$ on each side. ■

If we write μ_s for the law of X_s as a row vector, then

$$\mu_s = \mu_0 P^s$$

where here P^s is the matrix product of P by itself s times.

Aperiodicity

Definition (Aperiodicity)

A chain is said to be *aperiodic* if for all $x \in V$

$$\gcd\{t \,:\, P^t(x,x) > 0\} = 1.$$

Example (Lazy walk)

A *lazy, simple random walk* on G is a Markov chain such that, at each time, it stays put with probability $1/2$ or chooses a uniformly random neighbor of the current state otherwise. Such a walk is aperiodic.

Stationary distribution I

Definition (Stationary distribution)

Let (X_t) be a Markov chain with transition matrix P. A *stationary measure* π is a measure such that

$$\sum_{x\in V} \pi(x) P(x,y) = \pi(y), \qquad \forall y \in V,$$

or in matrix form $\pi = \pi P$. We say that π is a *stationary distribution* if in addition π is a probability measure.

Example

The measure $\pi \equiv 1$ is stationary for simple random walk on $\mathbb{L}^d$.

Stationary distribution II

Theorem (Existence and uniqueness: finite case)

If P is irreducible and has a finite state space, then it has a unique stationary distribution.

Definition (Reversible chain)

A transition matrix P is *reversible* w.r.t. a measure η if $\eta(x)P(x,y) = \eta(y)P(y,x)$ for all $x, y \in V$. By summing over y, such a measure is necessarily stationary.

By induction, if (X_t) is reversible w.r.t. a stationary distribution π

$$\mathbb{P}_\pi[X_0 = x_0, \ldots, X_t = x_t] = \mathbb{P}_\pi[X_0 = x_t, \ldots, X_t = x_0].$$

Stationary distribution III

Example

Let (X_t) be simple random walk on a connected graph G. Then (X_t) is reversible w.r.t. $\eta(v) := \delta(v)$.

Example

The Metropolis algorithm modifies a given irreducible symmetric chain Q to produce a new chain P with the same transition graph and a prescribed positive stationary distribution π. The definition is of the new chain is:

$$P(x,y) := \begin{cases} Q(x,y)\left[\frac{\pi(y)}{\pi(x)} \wedge 1\right], & \text{if } x \neq y, \\ 1 - \sum_{z \neq x} P(x,z), & \text{otherwise.} \end{cases}$$

Proof of Metropolis chain reversibility

Proof: Suppose $x \neq y$ and $\pi(x) \geq \pi(y)$. Then, by the definition of P, we have

$$\begin{aligned} \pi(x)P(x,y) &= \pi(x)Q(x,y)\frac{\pi(y)}{\pi(x)} = Q(x,y)\pi(y) \\ &= Q(y,x)\pi(y) = P(y,x)\pi(y), \end{aligned}$$

where we used the symmetry of Q. Moreover $P(x,z) \leq Q(x,z)$ so $\sum_{z \neq x} P(x,z) \leq 1$. ∎

Convergence

Theorem (Convergence to stationarity)

Suppose P is irreducible, aperiodic and has stationary distribution π. Then, for all x, y, $P^t(x, y) \to \pi(y)$ as $t \to +\infty$.

For probability measures μ, ν on V, let their *total variation distance* be $\|\mu - \nu\|_{\mathrm{TV}} := \sup_{A \subseteq V} |\mu(A) - \nu(A)|$.

Definition (Mixing time)

The *mixing time* is

$$t_{\mathrm{mix}}(\varepsilon) := \min\{t \geq 0 \,:\, d(t) \leq \varepsilon\},$$

where $d(t) := \max_{x \in V} \|P^t(x, \cdot) - \pi(\cdot)\|_{\mathrm{TV}}$.

Proofs of total variation distance properties I

Lemma: $\|\mu - \nu\|_{\mathrm{TV}} = \frac{1}{2}\sum_{x\in V} |\mu(x) - \nu(x)|$.
Proof: Let $B := \{x\ :\ \mu(x) \geq \nu(x)\}$. Then, for any $A \subseteq V$,

$$\mu(A) - \nu(A) \leq \mu(A \cap B) - \nu(A \cap B) \leq \mu(B) - \nu(B),$$

and similarly $\nu(A) - \mu(A) \leq \nu(B^c) - \mu(B^c)$. The two bounds are equal so $|\mu(A) - \nu(A)| \leq \mu(B) - \nu(B)$, which is achieved at $A = B$. Also

$$\mu(B) - \nu(B) = \frac{1}{2}\left[\mu(B) - \nu(B) + \nu(B^c) - \mu(B^c)\right] = \frac{1}{2}\sum_{x\in V} |\mu(x) - \nu(x)|.$$

■

Proofs of total variation distance properties II

Lemma: $d(t)$ is non-increasing in t.
Proof:

$$\begin{aligned}
d(t+1) &= \max_{x\in V}\sup_{A\subseteq V}|P^{t+1}(x,A)-\pi(A)| \\
&= \max_{x\in V}\sup_{A\subseteq V}\left|\sum_z P(x,z)(P^t(z,A)-\pi(A))\right| \\
&\leq \max_{x\in V}\sum_z P(x,z)\sup_{A\subseteq V}|P^t(z,A)-\pi(A)| \\
&\leq \max_{z\in V}\sup_{A\subseteq V}|P^t(z,A)-\pi(A)|
\end{aligned}$$

■

A little linear algebra I

Assume V is finite and $n := |V|$.

Theorem

Any real eigenvalue λ of P satisfies $|\lambda| \leq 1$.

Proof: $Pf = \lambda f \implies |\lambda| \|f\|_\infty = \|Pf\|_\infty = \max_x |\sum_y P(x,y) f(y)| \leq \|f\|_\infty$ ∎

Assume further that P is reversible w.r.t. π. Define

$$\langle f, g \rangle_\pi = \sum_{x \in V} \pi(x) f(x) g(x), \qquad \|f\|_\pi^2 = \langle f, f \rangle_\pi.$$

Theorem

There is an orthonormal basis of $(\mathbb{R}^n, \langle \cdot, \cdot \rangle_\pi)$ of real right eigenvectors $\{f_j\}_{j=1}^n$ of P with real eigenvalues $\{\lambda_j\}_{j=1}^n$.

A little linear algebra II

Proof: Let D_π be the diagonal matrix with π on the diagonal. By reversibility,

$$M(x,y) := \sqrt{\frac{\pi(x)}{\pi(y)}}P(x,y) = \sqrt{\frac{\pi(y)}{\pi(x)}}P(y,x) =: M(y,x).$$

So $M = (M(x,y))_{x,y} = D_\pi^{1/2} P D_\pi^{-1/2}$, as a symmetric matrix, has real eigenvectors $\{\phi_j\}_{j=1}^n$ forming an orthonormal basis of $\mathbb{R}^n$ with corresponding eigenvalues $\{\lambda_j\}_{j=1}^n$. Define $f_j := D_\pi^{-1/2}\phi_j$. Then

$$Pf_j = PD_\pi^{-1/2}\phi_j = D_\pi^{-1/2}D_\pi^{1/2}PD_\pi^{-1/2}\phi_j = D_\pi^{-1/2}M\phi_j = \lambda_j D_\pi^{-1/2}\phi_j = \lambda_j f_j,$$

and

$$\langle f_i, f_j\rangle_\pi = \langle D_\pi^{-1/2}\phi_i, D_\pi^{-1/2}\phi_j\rangle_\pi = \sum_x \pi(x)[\pi(x)^{-1/2}\phi_i(x)][\pi(x)^{-1/2}\phi_j(x)] = \langle \phi_i, \phi_j\rangle.$$

■

Random walk on a graph

Definition

Let $G = (V, E)$ be a finite or countable, locally finite graph. *Simple random walk* on G is the Markov chain on V, started at an arbitrary vertex, which at each time picks a uniformly chosen neighbor of the current state.

Questions:

- How often does the walk return to its starting point?
- How long does it take to visit all vertices once or a particular subset of vertices for the first time?
- How fast does it approach stationarity?

Random walk on a network

Definition

Let $G = (V, E)$ be a finite or countable, locally finite graph. Let $c : E \to \mathbb{R}_+$ be a positive edge weight function on G. We call $\mathcal{N} = (G, c)$ a *network*. Random walk on $\mathcal{N}$ is the Markov chain on V, started at an arbitrary vertex, which at each time picks a neighbor of the current state proportionally to the weight of the corresponding edge.

Any countable, reversible Markov chain can be seen as a random walk on a network (not necessarily locally finite) by setting $c(e) := \pi(x)P(x, y) = \pi(y)P(y, x)$ for all $e = \{x, y\} \in E$.

Bond percolation I

Definition

Let $G = (V, E)$ be a finite or countable, locally finite graph. The *bond percolation* process on G with density $p \in [0, 1]$, whose measure is denoted by $\mathbb{P}_p$, is defined as follows: each edge of G is independently set to *open* with probability p, otherwise it is set to *closed*. Write $x \Leftrightarrow y$ if $x, y \in V$ are connected by a path all of whose edges are open. The *open cluster* of x is

$$\mathcal{C}_x := \{y \in V \,:\, x \Leftrightarrow y\}.$$

Bond percolation II

We will mostly consider bond percolation on $\mathbb{L}^d$ or $\mathbb{T}_d$.

Questions:

- For which values of p is there an infinite open cluster?
- How many infinite clusters are there?
- What is the probability that y is in the open cluster of x?

Random graphs: Erdös-Rényi

Definition

Let $V = [n]$ and $p \in [0, 1]$. The *Erdös-Rényi graph* $G = (V, E)$ on n vertices with density p is defined as follows: for each pair $x \neq y$ in V, the edge $\{x, y\}$ is in E with probability p independently of all other edges. We write $G \sim \mathbb{G}_{n,p}$ and we denote the corresponding measure by $\mathbb{P}_{n,p}$.

Questions:

- What is the probability of observing a triangle?
- Is G connected? If not, how large are the components?
- What is the typical chromatic number (i.e., the smallest number of colors needed to color the vertices so that no two adjacent vertices share the same color)?

Random graphs: preferential attachment

Definition

The *preferential attachment process* produces a sequence of graphs $(G_t)_{t\geq 1}$ as follows. We start at time 1 with two vertices, denoted v_0 and v_1, connected by an edge. At time t, we add vertex v_t with a single edge connecting it to an old vertex, which is picked proportionally to its degree. We write $(G_t)_{t\geq 1} \sim \mathrm{PA}_1$.

Questions:

- How are the degrees distributed?
- What is the typical distance between two vertices?

Gibbs random fields I

Definition

Let S be a finite set and let $G = (V, E)$ be a finite graph. Denote by $\mathcal{K}$ the set of all cliques of G. A positive probability measure μ on $\mathcal{X} := S^V$ is called a *Gibbs random field* if there exist *clique potentials* $\phi_K : S^K \to \mathbb{R}$, $K \in \mathcal{K}$, such that

$$\mu(x) = \frac{1}{\mathcal{Z}} \exp\left(\sum_{K \in \mathcal{K}} \phi_K(x_K)\right),$$

where x_K is x restricted to the vertices of K and $\mathcal{Z}$ is a normalizing constant.

Gibbs random fields II

Example

For $\beta > 0$, the *ferromagnetic Ising model* with inverse temperature β is the Gibbs random field with $S := \{-1, +1\}$, $\phi_{\{i,j\}}(\sigma_{\{i,j\}}) = \beta\sigma_i\sigma_j$ and $\phi_K \equiv 0$ if $|K| \neq 2$. The function $\mathcal{H}(\sigma) := -\sum_{\{i,j\}\in E} \sigma_i\sigma_j$ is known as the *Hamiltonian*. The normalizing constant $\mathcal{Z} := \mathcal{Z}(\beta)$ is called the *partition function*. The states $(\sigma_i)_{i\in V}$ are referred to as *spins*.

Questions:

- How fast is correlation decaying?

Interacting particles: Glauber dynamics I

Definition

Let μ_β be the Ising model with inverse temperature $\beta > 0$ on a graph $G = (V, E)$. The *(single-site) Glauber dynamics* is the Markov chain on $\mathcal{X} := \{-1, +1\}^V$ which at each time:

- selects a site $i \in V$ uniformly at random, and
- updates the spin at i according to μ_β conditioned on agreeing with the current state at all sites in $V \backslash \{i\}$.

Interacting particles: Glauber dynamics II

Specifically, for $\gamma \in \{-1,+1\}$, $i \in \Lambda$, and $\sigma \in \mathcal{X}$, let $\sigma^{i,\gamma}$ be the configuration σ with the spin at i being set to γ. Let $n = |V|$ and $S_i(\sigma) := \sum_{j \sim i} \sigma_j$. Because the Ising measure factorizes, the nonzero entries of the transition matrix are

$$Q_\beta(\sigma, \sigma^{i,\gamma}) := \frac{1}{n} \cdot \frac{e^{\gamma \beta S_i(\sigma)}}{e^{-\beta S_i(\sigma)} + e^{\beta S_i(\sigma)}}.$$

Theorem

The Glauber dynamics is reversible w.r.t. μ_β.

Question: How quickly does the chain approach μ_β?

Interacting particles: Glauber dynamics III

Proof of the theorem: This chain is clearly irreducible. For all $\sigma \in \mathcal{X}$ and $i \in V$, let $S_{\neq i}(\sigma) := \mathcal{H}(\sigma^{i,+}) + S_i(\sigma) = \mathcal{H}(\sigma^{i,-}) - S_i(\sigma)$. We have

$$\begin{aligned}
\mu_\beta(\sigma^{i,-})\, Q_\beta(\sigma^{i,-}, \sigma^{i,+}) &= \frac{e^{-\beta S_{\neq i}(\sigma)} e^{-\beta S_i(\sigma)}}{\mathcal{Z}(\beta)} \cdot \frac{e^{\beta S_i(\sigma)}}{n[e^{-\beta S_i(\sigma)} + e^{\beta S_i(\sigma)}]} \\
&= \frac{e^{-\beta S_{\neq i}(\sigma)}}{n\mathcal{Z}(\beta)[e^{-\beta S_i(\sigma)} + e^{\beta S_i(\sigma)}]} \\
&= \frac{e^{-\beta S_{\neq i}(\sigma)} e^{\beta S_i(\sigma)}}{\mathcal{Z}(\beta)} \cdot \frac{e^{-\beta S_i(\sigma)}}{n[e^{-\beta S_i(\sigma)} + e^{\beta S_i(\sigma)}]} \\
&= \mu_\beta(\sigma^{i,+})\, Q_\beta(\sigma^{i,+}, \sigma^{i,-}).
\end{aligned}$$

■

Interacting particles: voter model

Definition

Let $S := \{0, 1\}$ and let $G = (V, E)$ be a finite, connected graph. The *voter model* on G is the Markov chain (η_t) on S^V which, at each time t, picks a uniform site $i \in V$ as well as a uniform neighbor $j \sim i$ and sets $\eta_t(i) := \eta_{t-1}(j)$.

Questions:

- How long does it take to reach one of the two *absorbing states*, i.e., the all-0 and all-1 configurations?