Lecture 10:

Time for Problem Session: sometime during week of Feb. 9.

Recall Defn: An MPT T is *mixing* if for all $A, B \in \mathcal{A}$,

$$\mu(T^{-n}(A) \cap B) \to \mu(A)\mu(B).$$

Recall:

$$Q := \lim_{n\to\infty} (1/n)(\sum_{k=0}^{n-1} P^k)$$

exists.

Note: even for irreducible P, $\lim_{n\to\infty} P^n$ need not exist:

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

Mixing for Markov chains:

Defn: P is *primitive* if $P^n > 0$ (entry-wise) for some n.

Graph interpretation: there is a uniform time n in which you can get from any state to any state.

Examples:

$$P = \begin{matrix} 0 & 1 & 0 \\ 1/3 & 0 & 2/3 \\ 0 & 1 & 0 \end{matrix}$$

P^n is never positive: look at graph.

$$P = \begin{matrix} 0 & 1/2 & 1/2 \\ 1/3 & 0 & 2/3 \\ 0 & 1 & 0 \end{matrix}$$

$P^4 > 0$.

Prop: P is primitive iff P is irreducible and aperiodic (i.e., gcd of cycle lengths of $G(P)$ is 1).

(check examples)

Proof: gcd $=$ gcd$\{n : \text{trace}(P^n) > 0\}$.

Only If: Clearly, primitivity implies irreducibility.

$P^n > 0$ implies $P^{n+1} > 0$ because no row is all zeros.

Aperiodic since $\gcd(n, n+1) = 1$.

If: Special case: there exists a self-loop, i.e., for some state i, $P_{ii} > 0$.

By irreducibility, connect i to j via the self-loop at k.

In general, use a combination of cycle lengths that are relatively prime (exercise).

TFAE

1. MC is mixing.
2. $(P^k)_{ij} \to \pi_j$.
3. P is primitive.

So, Example 1 is ergodic but not mixing, and Example 2 is mixing.

Proof:

$1 \Leftrightarrow 2$: just like the proof in analogous result for ergodicity:

Let A_i and A_j be initial "thin" cylinder sets: e.g., $A_i = \{x : x_0 = i\}$).

Recall: $\pi_i(P^n)_{ij} = \mu(T^{-n}(A_j) \cap A_i)$.

So, $(P^n)_{ij} \to \pi_j$ iff $\mu(T^{-n}(A_j) \cap A_i) \to \mu(A_i)\mu(A_j)$.

So, 2 holds iff the mixing defn. holds for cylinder sets A_i, A_j. Extend to general cylinder sets as in the exercise in the analogous result for ergodicity.

$2 \Rightarrow 3$: Follows since $\pi > 0$.

$3 \Rightarrow 2$: geometric proof (alternatively, can use renewal theory or Perron-Frobenius theory)

Let $W = \{(x_1 \dots x_d) \in \mathbb{R}_+^d : \ \sum x_i = 1\}$.

$$P : W \to W, \ x \mapsto xP$$

Well-defined since $(xP) \cdot \underline{1} = x(P \cdot \underline{1}) = x \cdot \underline{1} = 1$,.

Note: $P^k(W)$ is nested decreasing.

Claim: $\Delta \equiv \cap_{k=0}^{\infty} P^k(W)$ is a single point $\{z\}$.

—- If true, then $z = \pi$ because:

———- $\pi = \pi P^k \in W$.

—- Thus, for all i, $e_i P^k \to \pi$ and so $(P^k)_{ij} \to \pi_j$.

For general d:

Δ is a closed interval in W. If not a single point, then endpoints x, y of Δ form a linearly independent set, and x and y are both fixed by P^2. Thus, P^2 is the identity, contrary to primitivity.

Proof of claim for general d:

Exercise: can adapt proof above OR:

Theorem: A contraction mapping S on a compact space X has a unique fixed point z, and for all $x \in X$, $S^n x \to z$.

Fact: If $P > 0$, then $S := P|_W$ (i.e. $w \mapsto wP$) is a contraction mapping w.r.t. Hilbert metric on $int(W)$:

$$\rho(v, w) = \max_{i,j} \log \left(\frac{w_i / w_j}{v_i / v_j} \right)$$

i.e., there is a constant $0 < c < 1$, s.t. for all $v, w \in int(W)$, $\rho(Sv, Sw) \leq c\rho(v, w)$.

If P is primitive, then for some n, $P^n(W)$ is a compact subset of $int(W)$.

Apply contraction mapping theorem:

——— $P|_W$ has a unique fixed point (which must be π) and for all $x \in P^n(W)$, $xP^k \to \pi$

apply to each $x = e_i P^n$. $\square$

Note: π is the unique eigenvector (and it exists). Using P-F theory, can extend uniqueness and existence to irreducible stochastic matrices and then existence to general stochastic matrices.

Note: Mixing is often called Strong Mixing.

Note: There are many variants on the notion of Strong Mixing:

Mixing of all orders $\Rightarrow$ Mixing $\Rightarrow$ Mild Mixing $\Rightarrow$ Weak Mixing $\Rightarrow$ Totally Ergodic $\Rightarrow$ Ergodic

$\mathbb{Z}_+^d$ and $\mathbb{Z}^d$-actions:

Let T be an IMPT on a probability space $(M, \mathcal{A}, \mu)$. Then T generates a collection $\mathcal{C} := \{T^i : i \in \mathbb{Z}\}$ of IMPT's acting on $(M, \mathcal{A}, \mu)$. In fact, this forms a group under composition:

$$T^i \circ T^j = T^{i+j}$$

– clearly associative
– Identity: $T^0 = I$
– Inverses: $(T^i)^{-1} = T^{-i}$

Then the map $\mathbb{Z} \to \mathcal{C}$ $i \mapsto T^i$ is a group homomorphism (whose image either $\mathbb{Z}$ or $\mathbb{Z}_n$)

Defn: IMPT$(M, \mathcal{A}, \mu)$ is the set of all IMPT's from $(M, \mathcal{A}, \mu)$ to itself (we identify two such IMPT's if they agree off a set of measure zero).

Fact: IMPT$(M, \mathcal{A}, \mu)$ is a group under composition.

Defn: A *measure-preserving group action* (MP) of a group G on $(M, \mathcal{A}, \mu)$ is a group homomorphism $G \to IMPT(X, \mathcal{A})$ (the latter is the group of all invertible MPT's)

– $g \mapsto T^g$ satisfying $T^{g_1 g_2} = T^{g_1} \circ T^{g_2}$

An MP $\mathbb{Z}$-action is essentially the same as an IMPT.

Lecture 11:

Select time for problem session.

Recall:

Defn: IMPT$(M, \mathcal{A}, \mu)$ is the set of all IMPT's from $(M, \mathcal{A}, \mu)$ to itself (we identify two such IMPT's if they agree off a set of measure zero).

Fact: IMPT$(M, \mathcal{A}, \mu)$ is a group under composition.

Defn: A *measure-preserving group action* (MP) of a group G on $(M, \mathcal{A}, \mu)$ is a group homomorphism $G \to IMPT(X, \mathcal{A})$; this means:

$g \mapsto T^g$ satisfies $T^{g_1 g_2} = T^{g_1} \circ T^{g_2}$

Sometimes, we identify the group action with its image $\mathcal{T} = \{T^g : g \in G\}$.

An MP $\mathbb{Z}$-action is essentially the same as an IMPT.

More generally, an MP $\mathbb{Z}^d$ action is essentially the same as d *commuting* IMPT's: $T_1, \ldots, T_d$:

Given an MP $\mathbb{Z}^d$ action, $T_i = T^{e_i}$, $i = 1, \ldots, d$, are invertible and commute.

Conversely, given a set of commuting IMPT's $T_1, \ldots, T_d$,

$$\sum_i a_i e_i \mapsto T_1^{a_1} \circ T_2^{a_2} \circ \cdots T_d^{a_d} \ (a_i \in \mathbb{Z})$$

is a $\mathbb{Z}^d$ action.

Defn: A group action is *faithful* if the map $g \mapsto T^g$ is one-to-one.

Do the same for semigroups instead of groups to get MP semigroup actions; in particular, an $\mathbb{Z}_+^d$ action is essentially the same as d commuting MPT's.

Formally:

Defn: A *semigroup* is a set S with a binary operation $\star$, $S \times S \to S$, $s_1 \star s_2 \in S$ satisfying associativity.

We usually assume that S has an identity element (in which case it is usually called a *monoid*).

Defn: MPT$(M, \mathcal{A}, \mu)$ is the set of all MPT's from $(M, \mathcal{A}, \mu)$ to itself (we identify two MPT's if they agree off a set of measure zero).

MPT$(M, \mathcal{A}, \mu)$ is a semigroup under composition.

Defn: A *measure-preserving semigroup action* of a semigroup G is a semigroup homomorphism $G \to MPT(X, \mathcal{A})$ (preserves multiplication and the identity)

A $\mathbb{Z}_+^d$ semigroup action is generated by d commuting MPT's.

Examples:

1. The $\mathbb{Z}^2$ group action generated by two rotations of the circle:

$$\{T_\alpha^i \circ T_\beta^j : (i, j) \in \mathbb{Z}^2\}$$

(with $\mu =$ normalized Lebesgue measure)

If $\{\alpha, \beta, 1\}$ is linearly independent over the rationals then it is a faithful action; example $\alpha = \sqrt{2}, \beta = \sqrt{3}$.

2. The $\mathbb{Z}_+^2$ semigroup action generated by the doubling map T and the tripling map S on $[0, 1)$:

$$\{T^i \circ S^j\}$$

$T^i \circ S^j(x) = 2^i 3^j \bmod 1$.

(with $\mu =$ Lebesgue measure)

Conjecture: Lebesgue measure is the only nontrivial (to be defined later) Borel probability measure for which this action is MP.

3. The $\mathbb{Z}^d$ group action (or $\mathbb{Z}_+^d$ semi-group action) determined by a d-parameter stationary process $\{X_{\bar{i}} : \bar{i} \in \mathbb{Z}^d\}$ (or $\bar{i} \in \mathbb{Z}_+^d$).

Here $M = F^{\mathbb{Z}^d}$, $\mathcal{A}$ is the Borel σ-algebra, μ is the law of the process, which is determined by probability measures on cylinder sets: for a finite subset S of $\mathbb{Z}^d$ and a configuration $w \in F^S$

$$A_{S,w} = \{x \in M : x_s = w_s, s \in S\}$$

For $\bar{i} \in \mathbb{Z}^d$, $T^{\bar{i}}$ is the shift by the vector $\bar{i}$:

$$(T^{\bar{i}}(x))_{\bar{j}} = x_{\bar{j}+\bar{i}}$$

The $\mathbb{Z}^d$-action: $\bar{i} \mapsto T^{\bar{i}}$ is called the *shift action.*

For $d = 2$, the action is generated by T^{e_1}, the horizontal left-shift and T^{e_2}, the vertical down-shift. Sometimes we write $T = \sigma$.

Some specific shift actions:

a. *IID* X_{ij}

b. *Ledrappier 3-dot example*: let

$$M = \{x \in \{0,1\}^{\mathbb{Z}^2} : x_{ij} + x_{i+1,j} + x_{i,j+1} = 0 \bmod 2\}$$

We will define a shift-invariant measure μ whose support is M (the (topological) support of a probability measure is the smallest closed set with measure 1; equivalently the cylinder sets which have strictly positive measure are those which intersect M).

μ is the uniform measure on M "subject to the constraint."

Formally:

Let H_+ denote the (closed) upper half-plane in $\mathbb{Z}^2$. Let M_0 denote the set of configurations on H_+ given by the same constraint as M. Subject to this constraint, an arbitrary configuration on the x-axis uniquely determines the configuration in the entire half-plane: the mapping from M_0 to $\{0,1\}^{x-axis}$, given by $z \mapsto z|_{x-axis}$, is a bijection.

Define μ_0, a Borel probability measure on M_0 as the "pull back" of the measure iid$(1/2, 1/2)$ on $\{0,1\}^{x-axis}$.

We claim that μ_0 defines a stationary process on H_+, i.e, μ_0 is preserved under horizontal (left or right) shifts and vertical (up) shifts.

For instance, any "valid" configuration w on a right isosceles triangle with vertices $(0,0), (0, m-1), (m-1, 0)$ is uniquely determined by its values on its bottom edge.

So, $\mu(w) = 2^{-m}$.

Now, let Δ be a translate of this triangle, with vertices $(i, j), (i, j+m-1), (i+m-1, j)$, with $j \geq 0$.

Extend the vertical edge and hypotenuse of Δ to form a larger triangle Δ', whose horizontal edge is on the x-axis. For each valid configuration w on Δ and each arbitrary configuration u on the part of the hypotenuse that is in $\Delta' \setminus \Delta$, w extends to a unique valid configuration on Δ'. The size of the bottom edge of Δ' is $m+j$, and there are 2^j configurations on T. Thus, $\mu(w) = 2^j 2^{-m-j} = 2^{-m}$.

Thus, for such triangle configurations, μ_0 is invariant under vertical (up) shifts. It is clearly invariant horizontal shifts (left or right).

Configurations on such triangles generate the σ-algebra on M_0. In fact, in M_0, cylinder sets on rectangles are the same as cylinder sets on such triangles.

Now, define the analogous measure μ_n on configurations on the plane $j \geq n$.

Observe that μ_n is consistent with μ_{n-1} for the same reasons as above: if w is a right isosceles triangle sitting on $y = n$, with base length m, then w gives rise to exactly two configurations, w', w'', on the triangle sitting on $y = n-1$, with base length $m+1$, and so

$$\mu_{n-1}(w) = \mu_{n-1}(w') + \mu_{n-1}(w'') = 2^{-(m+1)} + 2^{-(m+1)} = 2^{-m} = \mu_n(w)$$

Thus, the limit of μ_n as $n \to -\infty$ defines a shift-invariant measure μ.

Lecture 12

Were considering MP $\mathbb{Z}^2$ shift actions, equivalent to $\mathbb{Z}^2$ stationary processes.

Example c. *Ising Model*

($\mathbb{Z}^2$, Ferromagnetic, no external field, coupling strength = 1.)

Alphabet: $F = \{\pm 1\}$

For $\bar{i}, \bar{j} \in \mathbb{Z}^2$, $\bar{i} \sim \bar{j}$ means that $\bar{i}$ and $\bar{j}$ are adjacent (horizontally or vertically), and we write $[\bar{i}, \bar{j}]$ for the edge that they form.

A configuration w on set S of lattice sites in $\mathbb{Z}^2$ is an element of F^S. It models the atoms of a magnetizable material, each magnetized in one of two possible directions.

For a finite set $S \subset \mathbb{Z}^2$ and a configuration $x \in F^S$, define the *energy* of x:

$$U(x) = \sum_{[\bar{i},\bar{j}] \subset S} -x_{\bar{i}} x_{\bar{j}}$$

"Like bonds" contribute lower energy. Magnetization is preferred.

For two configurations x, y on disjoint sets S and T, $x \wedge y$ denotes the "concatenated" configuration on $S \cup T$, i.e., $x \wedge y \in F^{S \cup T}$.

Define the (external) boundary of S:

$$\partial S := \{\bar{k} \in S^c : \bar{k} \sim \bar{i} \text{ for some } \bar{i} \in S\}.$$

Given a finite set $S \subset \mathbb{Z}^2$ and a boundary configuration $\delta \in F^{\partial S}$, define the measure on F^S: for $x \in F^S$,

$$\mu_{S,\delta}(x) := \frac{e^{-U(x \wedge \delta)}}{Z} = \frac{e^{\sum_{[i,j] \subset S \cup \partial S} \quad (x \wedge \delta)_{\bar{i}} (x \wedge \delta)_{\bar{j}}}}{Z}$$

where $Z := Z_{S,\delta} = \sum_{w \in F^S} e^{-U(w \wedge \delta)}$ is the normalizing factor.

Ways to think about this formula:

– For a given configuration x on F^S, Numerator is e^{L-U} where L is the number of "like bonds" and U is the number of "unlike bonds" in $S \cup \partial S$.

– For a given configuration x on F^S, Numerator is a product of factors: e for each like bond and e^{-1} for each unlike bond in $S \cup \partial S$.

– Denominator is the sum of numerators for *all* configurations w on S.

By finite additivity, this measure defines probabilities on all configurations on all subsets of S.

Note: normally one does not include interactions on edges entirely contained in ∂S (i.e., within δ); we include such interactions, but this does not change the measure.

Notation: For a configuration x on a set V and a subset $W \subset V$, $x(W)$ denotes the restriction of x to W.

More generally, given finite subsets $S, T \subset \mathbb{Z}^2$ such that

$$\partial S \subseteq T \subset S^c$$

and a configuration $y \in F^T$, define the measure on F^S: for $x \in F^S$,

$$\mu_{S,y}(x) := \frac{e^{-U(x \wedge y)}}{Z}$$

where $x \in F^S$ and $Z := Z_{S,y} = \sum_{w \in F^S} e^{-U(w \wedge y)}$ is the normalizing factor.

Lemma:

$$\mu_{S,y} = \mu_{S,\delta}$$

where $\delta = y(\partial S)$

Proof:

$$\mu_{S,y} = \frac{e^{-U(x \wedge y)}}{\sum_{w \in F^S} e^{-U(w \wedge y)}}$$

$$= \frac{(e^{-U(x\wedge\delta)})(*)}{\sum_{w\in F^S}(e^{-U(w\wedge\delta)})(*)}$$

where

$$(*) = e^{\sum_{[\bar{i},\bar{j}]\subset S\cup T,\ [\bar{i},\bar{j}]\not\subset S\cup\partial S}\quad y_{\bar{i}}y_{\bar{j}}}$$

(because ∂S insulates S from $T \setminus (S \cup \partial S)$). $\square$

∂S insulates S from $T \setminus \partial S$.

Defn: A *Gibbs measure* μ for the Ising model on a subset K (finite or infinite) of $\mathbb{Z}^2$ is a probability measure on F^K s.t. for all finite subsets S, T with $\partial S \subseteq T \subset K \setminus S$, $z \in F^{S\cup T}$,

$$\mu(z(S) \mid z(T)) = \mu_{S,z(\partial S)}(z(S))$$

Note that this says two things:

$$1) \quad \mu(z(S) \mid z(T)) = \mu(z(S) \mid z(\partial S))$$

i.e., conditioned on the boundary, the inside and outside are independent (this is often called the *Markov Random Field (MRF) property*) -and-

2) the conditional probability has the (exponential) Gibbs form (this is often called the *Gibbs speciication*)

Proposition: for all finite K and $\eta \in F^{\partial K}$, $\mu_{K,\eta}$ is a Gibbs measure for the Ising model.

Proof: Cancellation!

Let S, T, z be as in the defn of Gibbs measure.

Let $x = z(S)$ and $y = z(T)$.

We claim that

$$\mu_{K,\eta}(x|y) = \mu_{S,y}(x) \tag{1}$$

Then by the Lemma, we get

$$\mu(z(S) \mid z(T)) = \mu_{K,\eta}(x|y) = \mu_{S,y}(x)$$

$$= \mu_{S,\delta}(x) = \mu_{S,z(\partial S)}(z(S))$$

where $\delta = y(\partial S) = z(\partial S)$.

Proof of (1):

$$\mu_{K,\eta}(x|y) = \frac{\mu_{K,\eta}(x \wedge y)}{\mu_{K,\eta}(y)}$$

$$= \frac{\sum_{v \in F^{K \setminus (S \cup T)}} e^{-U(x \wedge y \wedge v \wedge \eta)}}{\sum_{w \in F^S} \sum_{v \in F^{K \setminus (S \cup T)}} e^{-U(w \wedge y \wedge v \wedge \eta)}}$$

$$= \frac{e^{-U(x \wedge y)}}{\sum_{w \in F^S} e^{-U(w \wedge y)}} = \mu_{S,y}(x)$$

(because the effect of v and η on the energy is the same for x and all w, since T insulates S from $F \setminus (S \cup T)$).

$\square$