Markov Chains and Dynamical Systems, Spring 2024

Log: a brief summary of the classes

February 6

Introduction:

Dynamical systems: phase space M , map $T : M \to M$. Orbit, periodicity, asymptotic behavior. Simple examples on $M = \mathbb{R}$ or $M = [0, 1)$ with regular vs. chaotic patterns. Billiards.

Markov chains: state space, transition probabilities. A simple example: the weather chain. Stationary distribution and its relation to the following questions: on the average, what is the long term proportion of sunny days? If today it is sunny, what is the expected number of days that it is sunny again? Formula for the stationary distribution for chains with two states. *Random walks* on $\mathbb{Z}$ and $\mathbb{Z}^2$

Relation of the two topics: evolution at different time scales.

Gambler's ruin problem.

February 8

Probability recap: $(\Omega, \mathcal{F}, \mathbb{P})$ - sample spaces, σ-algebra of events, axioms of probability. Examples: roll two dice, infinite sequences of coin tosses. Inclusion-exclusion formula.

Conditional probability. Multiplication rule. Independence. Pairwise and complete independence of events $A_1, A_2, ...$

Discrete random variables: mass function, expected value $\mathbb{E}X$. The binomial and the geometric distributions. Two methods for computing $\mathbb{E}X$ – sums of indicators and $\sum\limits_{k\geq 1} \mathbb{P}(X \geq k)$.

Independent random variables, independent sequences.

Definition of a (homogeneous)*Markov chain.* Transition matrix and its properties.

February 13

Recap of setting: state space S, Markov chain, transition probabilities. Stochastic matrices ($\sum_j p(i, j) = 1\ \forall i$).

Examples: Gambler's ruin, Ehrenfest chain, Weather chain, Social mobility chain. Description of the transition graphs and the matrices $p(i, j)$ for these examples.

Absorbing sates. First return time T_i for $i \in S$, the return probabilities $\rho_{ii} = \mathbb{P}(T_i < \infty \,|\, X_0 = i)$. Transient ($\rho_{ii} < 1$) and recurrent ($\rho_{ii} = 1$) states. Discussion of the above examples from this perspective.

Multistep transition probabilities. Chapman-Kolmogorov equations. $p_n(i, j) = (p^n)_{ij}$. Inspection of p^n for the above examples.

February 15

Recap of Markov chains: transition matrix $p(i, j)$, multistep transition probabilities. Evolution of distributions.

Stationary distribution $\pi(i)$ as a (row) eigenvector corresponding to the eigenvalue 1 for multiplication by $p(i, j)$ from the right. Computing π for chains with two states.

Computing π for chains with more than two states. Question of uniqueness.

Recap of notations: $P_x(\cdot) = \mathbb{P}(\cdot|X_0 = x)$, $T_x = \min\{n \geq 1 | X_n = x\}$ (first hitting/return time), $\rho_{xy} = P_x(T_y < \infty)$.
Transience ($\rho_{xx} < 1$) – recurrence ($\rho_{xx} = 1$) dichotomy.

Relation $x \to y$ and equivalence relation $x \leftrightarrow y$. Closed and irreducible classes. Proposition: if C is a finite, closed and irreducible class, then all states $x \in C$ are recurrent.
Corollary – classification of states: $S = T \cup C_1 \cup \ldots C_K$, where T is the collection of transient states, while C_i are closed irreducible classes.

February 20

Application of classification of states: how to determine the classes for a given transition matrix,

Recap of notations: $P_x(\cdot) = \mathbb{P}(\cdot|X_0 = x)$, $T_x^1 = T_x = \min\{n \geq 1 | X_n = x\}$ (first hitting/return time). Notion of a stopping time. $T_x^k = T_x = \min\{n \geq T_x^{k-1} | X_n = x\}$ (kth hitting/return time), $\rho_{xy} = P_x(T_y < \infty)$, $N_x = \#\{n \geq 1 | X_n = x\}$. Transience ($\rho_{xx} < 1$) – recurrence ($\rho_{xx} = 1$) dichotomy.

Recall proposition: if C is a finite, closed and irreducible class, then all states $x \in C$ are recurrent. Proof: computing $E_x(N_y)$ in two different ways. Recurrence equivalent to $E_x(N_y) = \infty$. Two Lemmas: (i) if $x \in C$ recurrent, all $y \in C$ recurrent. (ii) *if C is finite*, there is at least one recurrent state in C.

Standing assumption (unless otherwise stated): irreducible chain on finite S. Proposition: there exist unique stationary distribution with $\pi_i > 0$ for all i. Reduce to the case of positive transition matrix.

February 22

Nonnegative and positive matrices. Irreducibility: for any i, j there exists n_0 such that $(p^{n_0})_{ij} > 0$. Question: when is it possible to choose the same n_0 for all i, j (that is, a power of p is positive)?

Recall proposition: there exists unique $\pi_j > 0$ for finite and irreducible S. Proof for *positive* p: consider the eigenspace $L = \{v_i | \sum_i v_i p_{ij} = v_j\}$: (i) $v_i \geq 0\ \forall i$; (ii) $v_i > 0\ \forall i$; (iii) L is simple. Extend to the *irreducible* case: (i) $r_{ij} = \frac{1}{2}(p_{ij} + \delta_{ij})$ (lazy chain) (ii) $\exists n_0$ such that r^{n_0} is positive (waste your time).

Some cases when the stationary distribution is easy to determine: *doubly stochastic chains* (uniform π_i), example: symmetric random walk with periodic boundary conditions. A glimpse at infinite S: symmetric random walk on $\mathbb{Z}$ has (uniform) stationary measure, but this cannot be scaled to a stationary probability distribution.

Detailed balance condition: $\pi_i p_{ij} = \pi_j p_{ji}$ for every $i \neq j$.

February 27

Examples of chains with *detailed balance*, that is, there exists π_i such that $\pi_i p_{ij} = \pi_j p_{ji}$ for every $i \neq j$. (Biased) random walk on $\{1, \ldots, K\}$ with reflecting boundary conditions; determination of π_i. Blowup of π_i when drift to the right and $K \to \infty$.

Example of an irreducible chain with three states that does not have detailed balance. Further examples for detailed balance: random walks on (undirected) graphs. Randomly hoping figures on the chessboard.

Birth and death chains. Example: stationary distribution for the Ehrenfest chain.

Period of state $i \in S$ defined as $d_i = gcd(I_i)$, with $I_i = \{n \geq 1 | p_{ii}^n > 0\}$. Examples.

February 29

Recall the notion of the period of state $i \in S$ defined as $d_i = gcd(I_i)$, with $I_i = \{n \geq 1 | p_{ii}^n > 0\}$. Main properties:

period is constant on irreducible classes,
if $d_i = 1$, then $\exists n_0(i)$ such that $p^n_{ii} > 0$ whenever $n \geq n_0$. Proof by number theoretic considerations.
Irreducible aperiodic chains: $\exists M \geq 1$ such that p^M is positive.

Two proofs of the convergence theorem: if irreducible, aperiodic, finite S, then $p^n_{ij} \to \pi_j$ for every $i, j \in S \times S$.

Linear algebraic – dynamical proof: letting $L = \{\underline{v} \in \mathbb{R}^{|S|} | \sum_{i \in S} v_i = 0\}$, and $F : L \to L$ multiplication by p_{ij} from right, then all eigenvalues of $F|_L$ are included in the interior of the unit disc. Key idea: letting $\Delta_0 = \Delta - \underline{\pi}$, we have that $\Delta_0 \subset L$ is compact with the origin included in $int\, \Delta_0$ such that $F(\Delta_0) \subset int\, \Delta_0$.

March 5

Infinite S: if the chain is irreducible and has stationary distribution with $\pi_i > 0$, then i is recurrent.

Recall T_i is the time of the first return/hit. Letting $N_n(i) = \#\{k = 1, \ldots, n | X_k = i\}$, we have: irreducibility and recurrence implies $\frac{N_n(i)}{n} \to (\mathbb{E}_i T_i)^{-1}$ almost surely. Proof based on (i) $\frac{R(k)}{k} \to \mathbb{E}_i T_i$ by law of large numbers where $R(k)$ is the time of the kth return/hit; (ii) $R(N_n(y)) \leq n < R(N_n(y) + 1)$.

Corollary: if irreducible and there exists stationary distribution π_i, we have $\pi_i = (\mathbb{E}_i T_i)^{-1}$.

Extension (without proof): if irreducible, there exists stationary distribution π_i and
$f \in L^1(S)$ (ie. $\sum_{i \in S} |f(i)|\pi_i < \infty$) then $\frac{1}{n} \sum_{k=1}^n f(X_k) \to \mathbb{E}_\pi f = \sum_{i \in S} f(i)\pi_i$ almost surely.

Recall the convergence theorem: if irreducible, aperiodic, finite S, then $p^n_{ij} \to \pi_j$ for every $i, j \in S \times S$. Probabilistic proof: extends to (countable) infinite S supposed there exists stationary distribution. Core idea: coupling. Suffices to show that if $\rho^{(1)}_i$ and $\rho^{(2)}_i$ are arbitrary probability distributions on S, with X_n and Y_n the corresponding realizations of the Markov chain, then $|\mathbb{P}(X_n = j) - \mathbb{P}(Y_n = j)| \to 0$ for every j as $n \to \infty$. Letting $T = \min\{n \geq 1 | X_n = Y_n\}$, it is enough to show $\mathbb{P}(T > n) \to 0$ as $n \to \infty$. If the chain on S is irreducible and aperiodic, then the chain on $S \times S$ is recurrent. Hence $\mathbb{P}(T = \infty) = 0$.

March 7 (long class)

Chains with absorbing states. Examples: two years college, tennis at deuce. Computing the exit probabilities.

Gambler's ruin problem revisited on $\{0, 1, \ldots, N-1, N\}$. Obtaining recursion for the exit probabilities $h(i) = \mathbb{P}(T_N < T_0 | X_0 = i)$. Fair case: linear profile for $h(i)$, connection to boundary value problems (differential equations in the continuum limit).

Gambler's ruin, biased case: letting $\Theta = q/p$, obtain $h(i) = \dfrac{\Theta^i - 1}{\Theta^N - 1}$. For N even, $h(N/2) = (\Theta^{N/2} + 1)^{-1}$. Refer back to tennis at deuce.

Linear algebraic formulation for exit probabilities: letting $S = C \cup \{a, b\}$ where a and b are absorbing states, introduce $T_a = \min\{n \geq 1 | X_n = a\}$ and T_b analogously, and finally $h(i) = \mathbb{P}(T_b < T_a | X_0 = i)$ for $i \in C$, we have $h = (I - r)^{-1} w$. Here h and w are vectors with components labelled by $i \in C$, in particular $h(i)$ is as above, $w(i) = p(i, b)$ and r is the matrix obtained as the restriction of p to $C \times C$, while I is the unit matrix.

Exit times. Consider $S = C \cup A$ where the states in A are absorbing, let T_A denote the first hit of A and $g(i) = \mathbb{E}(T_A | X_0 = i)$. Obtaining $g(i)$ by recursion in the two year college and the tennis cases.

Linear algebraic formulation for exit times: with notations as above, $g = (I - r)^{-1}\mathbf{1}$, where $\mathbf{1}$ is the vector with all entries equal 1. Example: tennis (along with a simpler derivation).

Further examples for exit times: waiting time for TT and HT in sequences of fair coin tosses. Discussion of why $\mathbb{E}(T_{TT}) = 6$ while $\mathbb{E}(T_{HT}) = 4$.

Exit time for gambler's ruin, fair case, on $\{0, 1, \dots, N\}$. Derivation of $g_N(i) = i(N-i)$. Connection to the boundary value problem $f : [0,1] \to \mathbb{R}$, $f''(x) = -2$, $f(0) = f(1) = 0$. Discussion of the asymptotic of gambler's ruin, fair case: $h_N(1) = 1/N \to 0$ as $N \to \infty$, yet $g_N(1) = N - 1 \to \infty$ as $N \to \infty$.

Markov chains with (countable) infinite state spaces – instructive example: reflecting random walk on $S = \{0, 1, 2, \dots\}$. The trichotomy of (i) transience – drift to the right; (ii) nullrecurrence (infinite expected return time) – symmetric case; (iii) positive recurrence (stationary distribution exists) – drift to the left.

March 12

Branching processes as Markov chains on $\{0, 1, \dots\}$, offspring distribution, 0 as absorbing state. Relevant questions: μ_n expected size of the nth generation, probability of extinction $r_n \nearrow r_\infty$.

Tools: Markov inequality, for ξ nonnegative integer valued: $\mathbb{P}(\xi \neq 0) \le \mathbb{E}\xi$. Conditional expectation and tower rule: $1 - r_n \le \mu_n = \mu^n$ where $\mu = \mathbb{E}Y$ is the expected value of the offspring distribution. Subcritical ($\mu < 1$), critical ($\mu = 1$) and supercritical ($\mu > 1$) branching.

Further tool: generating function for a nonnegative integer valued random variable. Examples, basic properties. Generating functions for sums with a random number of terms.

Branching processes: letting $g(z) = g_Y(z)$, we have $g_{X_n}(z) = g_n(z) = (g \circ \cdots \circ g)(z)$ for the generating function of the nth generation. $r_n = g_n(0)$; connection to dynamical systems: orbit of 0. Conclusion: r_∞ is the smallest nonnegative root of $z = g(z)$.

April 2

Recap of branching processes. Trichotomy and analogy: subcritical – positive recurrence; critical – nullrecurrence; supercritical – transience.

Examples: binary branching; branching with geometric offspring distribution

April 4

Simple Symmetric Random Walk (SSRW) on $\mathbb{Z}$; Formula for $\mathbb{P}(X_{2n} = 0)$. Stirling formula. Recurrence.

Outlook: SSRW on $\mathbb{Z}^d$: recurrence when $d = 2$, transience when $d \ge 3$.

Rotations of the circle. $\mathbb{S}^1$ as phase space. Invertibility, isometry, rigidity. Rational α: every point is periodic with the same period. Irrational α: every point has a dense orbit.

April 9

Midterm Exam

April 11

Invariant measures. Example: invariant probability measures associated to periodic orbits.

Lebesgue measure is the only invariant (Borel probability) measure (unique ergodicity) for irrational α (without proof).

Doubling map or $2x \pmod 1$. Non-invertibility, expansion, dyadic rationals are eventually fixed. Examples of periodic points, every rational $x \in \mathbb{S}^1$ is either periodic or eventually periodic.

One sided full shift with two symbols. $\Sigma^+ = \{0, 1\}^{\mathbb{N}}$ as a metric space – the symbolic metric, $\sigma : \Sigma^+ \to \Sigma^+$, the left shift.

Equivalence of dynamical systems, conjugacy.

The doubling map and the one sided full shift are (almost) conjugate, discussion of the conjugacy. Applications: characterization of periodic points and points with dense orbits.

Application: how to obtain the binary expansion for a rational number $x \in (0,1)$.

April 16

Lebesgue measure is invariant for the doubling map.

Invariant measures for the doubling map and the shift. Cylinder sets, description of the $1/2-1/2$ Bernoulli measure, when pushed forward, gives Lebesgue. Outlook: many further invariant measures for the shift.

Invariant sets (with respect to an invariant measure). Ergodicity as a relation of the dynamics and the measure.

Linear self-maps of the real line. Continuously differentiable maps of the real line: graphical analysis (cobweb plot), attracting and repelling fixed points.

Recursively defined sequences.

April 18

Time and space averages, ergodic measures. The ergodic theorem. Case of the doubling map: behavior of time averages for various points.

The case of the irrational rotation: as Lebesgue is the only invariant (Borel probability) measure, orbits equidistribute for every starting point.

Arnold's problem on the frequency of the first decimal digits in the sequence 2^n. Benford's law.

April 19

Calculus reminder: intermediate value theorem, mean value theorem.

Revisiting attracting and repelling fixed points: detailed arguments based on the mean value theorem, convergence of monotone and bounded sequences etc.

The logistic family $T_a x = ax(1-x)$, motivation from population dynamics. Fixed points: the case $a < 1$.

The implicit function theorem. A non-bifurcation result: a fixed point x_0 persists when perturbing the parameter a_0 unless $f'_{a_0}(x_0) = 1$.

The logistic family $T_a x = ax(1-x)$ for $a \geq 1$: orbits of $x \notin [0,1]$. Collision of fixed points for $a = 1$. Description of the attracting fixed point for $1 < a \leq 2$. Attracting fixed point for $2 < a < 3$: convergence with oscillation, the second iterate.

Recap of the logistic family $T_a x = ax(1-x)$ for $a < 3$. What happens at $a = 3$: discussion of the second iterate, inflection point, period-doubling bifurcation.

Outlook – complexity of the dynamics for $3 < a < 4$: period doubling cascades, parameters with attracting periodic orbit vs. parameters with an absolutely continuous invariant measure.

Discussion of logistic maps with $a > 4$ (for simplicity restrict to $a > 2+\sqrt{5}$). Intervals I_0 and I_1, inverse branches, construction of the invariant Cantor set. Topological conjugacy with the one sided full shift. Repeller.

April 23

Saddle-node bifurcation in one dimension. Examples.

Absolutely continuous measures, densities. Push-forward of a measure by a map Φ . Formula for the push-forward density for $\Phi : I \to J$ where I and J are intervals, and Φ is continuously differentiable and piecewise monotone.

Absolutely continuous in invariant measures (ie. invariant densities).

The transformation $\beta x \pmod 1$ with $\beta = \frac{2}{\sqrt{5}-1}$; piecewise constant invariant density.

April 25

The Gauss map and its invariant density.

Periodic points for continuous maps $T : \mathbb{R} \to \mathbb{R}$. Existence of a period 3 orbit implies existence of periodic orbits with (least) period n for every $n \in \mathbb{N}$. Statement of Sharkovsky's theorem (without proof).

April 30

Linear maps of the plane. Reminder: Jordan canonical form for real matrices. How the phase portrait is determined by the spectrum: source, sink, saddle, focus... Verifying stability by Lyapunov functions. Stable and unstable subspaces.

Two dimensional nonlinear maps: behavior in the vicinity of a fixed point. Hyperbolic fixed points and structural stability of their phase portraits.

Examples of non-hyperbolic fixed points with attracting/repelling behavior.

Hopf bifurcation.

May 2

The torus $\mathbb{T}^2$. $SL(2, \mathbb{Z})$ matrices. Toral automorphisms, definitions, invertibility, preservation of area (Lebesgue measure).

Discussion of an elliptic $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and a parabolic $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ example.

Hyperbolic toral automorphisms, discussed via the particular example of the CAT map: $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$. Sketch of domains. Stable and unstable leafs and foliations. Backward and forward asymptotic points, consequences of the irrational slope. Existence of a dense set of points homoclinic to the origin. Analogous phenomena for other periodic points, eg. $(1/2, 1/2)$.

Two equivalent characterizations of topological transitivity. Hyperbolic toral automorphisms are topologically transitive; proof based on the density of homoclinic points.

May 7

Baire category theorem. Proof of equivalence of the two characterizations of topological transitivity.

Smale's horseshoe map. The sets $H_0, H_1, V_0, V_1, H_{00}, H_{10}, H_{11}, H_{01}$ etc. $\Lambda = \Lambda^+ \cap \Lambda^- = \Lambda_1 \times \Lambda_2$, as the product of two Cantor sets.

May 9

Double sided full shift: product topology, separation metric, invertibility. Topological conjugacy with Smale's horseshoe. Periodic points, stable and unstable sets of a point, homoclinic points – geometric and symbolic characterization.

δ-chains and ε-shadowing. The shadowing property and its significance. Rotations do not have the shadowing property.

The doubling map has the shadowing property. The CAT map has the shadowing property.