Math 456: Mathematical Modeling

Tuesday, February 27th 2018

Markov Chains:
More on Stationarity distributions

Tuesday, February 27th 2018

Last two lectures

1. Ehrenfest chain, Wright-Fisher model, etc
2. Weighted graphs and random walks.
3. Markov Chains with random initial state.
4. Stationary distributions.

Last time

5. Path from x to y: a sequence of states x_k such that

$$x = x_0,\ x_1, \dots, x_{m-1},\ x_m = y$$

with $\mathrm{p}(x_{k-1}, x_k) > 0$ for $k = 1, \dots, m$.

The number m is called the lenght of the path.

If there is a path from x to y, we write $x \mapsto y$
which is read as "x communicates with y".

Last time

6. Closedness and irreducibility.

$A \subset S$ is closed means that:
$\mathrm{p}(x, y) = 0$ whenever $x \in A,\ y \notin A$.

$A \subset S$ is irreducible means that:
$x \mapsto y$ whenever both $x, y \in A$

Today

1. Theorem on stationary distributions.
2. Classification of states: transient and recurrent.
3. The Strong Markov Property.

Irreducible chains and stationary distributions

Recall that if q is a N vector representation a distribution over the state space of a Markov chain, then

$$q\mathrm{p} \quad (\text{equivalently, } \mathrm{p}^t q)$$

represents the new distribution after one step of the chain.

As such, distributions which are unchanged by the chain correspond to the (statistical) equilibrium of the system. Thus, a **stationary distribution** is a distribution q such that

$$\mathrm{p}^t q = q$$

As it turns out, if the Markov chain is irreducible, there is always one, and only one, stationary distribution.

Theorem (Theorem 1.14 in Durrett)

For an irreducible Markov chain, there is one, and only stationary distribution π*. Moreover* $\pi(x) > 0$ *for each state* x*.*

Irreducible chains and stationary distributions

A Theorem on Stationary Distributions

Irreducibility: what did it mean again?

Irreducible chains and stationary distributions

A Theorem on Stationary Distributions

Irreducibility: what did it mean again?

Let us keep our convention of using S to denote the state space. A chain is irreducible if $x \mapsto y$ for any pair of states $x, y \in S$.

Irreducible chains and stationary distributions

A Theorem on Stationary Distributions

Irreducibility: what did it mean again?

Let us keep our convention of using S to denote the state space. A chain is irreducible if $x \mapsto y$ for any pair of states $x, y \in S$.

In terms of the transition matrix, irreducibility is equivalent to

$$\forall\, x, y \in S \text{ there is some } k \in \mathbb{N} \text{ such that } \mathrm{p}^k(x, y) > 0$$

Note: it could be that different pairs of x, y require different k (remember the Ehrenfest chain!)

Irreducible chains and stationary distributions

A Theorem on Stationary Distributions

An important fact –which you should try to prove on your own, before we prove it in the next class– is the following.

Lemma

If S is irreducible and $\mathrm{p}(x,x) > 0$ for every state x then

$$\mathrm{p}^{N-1}(x,y) > 0 \text{ for all states } x,y \in S.$$

Here, N is the number of states in the chain.

Note: The Ehrenfest chain **does not** have this property!.

Before the proof, a concrete example

$$\begin{pmatrix} 0.5 & 0.0 & 0.2 & 0.3 \\ 0.0 & 0.3 & 0.6 & 0.1 \\ 0.2 & 0.6 & 0.2 & 0.0 \\ 0.3 & 0.1 & 0.0 & 0.6 \end{pmatrix}$$

Check: $\mathrm{p}^2(i,j) > 0 \;\; \forall\, i, j$. This means the chain is irreducible.

Before the proof, a concrete example

$$\begin{pmatrix} 0.5 & 0.0 & 0.2 & 0.3 \\ 0.0 & 0.3 & 0.6 & 0.1 \\ 0.2 & 0.6 & 0.2 & 0.0 \\ 0.3 & 0.1 & 0.0 & 0.6 \end{pmatrix}$$

Note that the matrix is symmetric.
From here, it is easy to check that $(1, 1, 1, 1)$ is an eigenvector.

Irreducible chains and stationary distributions

Proof of the theorem

The proof has three steps

1. Show that transition matrix p^t must have at least one eigenvector q with eigenvalue equal to 1
2. Show that any eigenvector q of p^t must have coordinates which are either all positive, or all negative.
3. Show that the space of eigenvectors for 1 is a set of one dimension, and conclude that the unique π we want lies along this one dimensional space.

Proof of Step 1.

The entries in each column of the matrix p^t add up to 1.

This means the entries in each column of $\mathrm{p}^t - I$ add up to 0.

Accordingly, the image of $\mathrm{p}^t - \mathrm{I}$ is contained in the ortoghonal complement to $q = (1, 1, \ldots, 1)$

As such, $\mathrm{p}^t - \mathrm{I}$ is a $N \times N$ matrix with rank $\leq N - 1$. $\square$

Proof of Step 2

Let q be an eigenvector of p^t with eigenvalue 1

Then q is also an eigenvector with eigenvalue 1 for

$$A = \left(\frac{1}{2}\mathrm{I} + \frac{1}{2}(\mathrm{p}^t)\right)^{N-1}$$

Check: (*this is where irreducibility is used!)

A is a stochastic matrix with **strictly positive entries***:

$$\sum_{j=1}^{N} A_{ij} = 1 \text{ for } i = 1, \ldots, N$$

$$A_{ij} > 0 \text{ for } i, j = 1, \ldots, N.$$

Proof of Step 2 (continued)

If the entries of q are not all nonnegative nor all nonpositive, then, on account that $A_{ij} > 0$, we have

$$|q_j| = \left| \sum_{i=1}^{N} A_{ij} q_i \right| < \sum_{i=1}^{N} |A_{ij} q_i| = \sum_{i=1}^{N} A_{ij} |q_i|$$

for $j = 1, \ldots, N$.

Proof of Step 2 (continued)

Adding these inequalities for $j = 1, \ldots, N$ we obtain

$$\sum_{j=1}^{N} |q_j| < \sum_{j=1}^{N} \sum_{i=1}^{N} A_{ij} |q_i|$$
$$= \sum_{i=1}^{N} \left(\sum_{j=1}^{N} A_{ij} \right) |q_i| = \sum_{i=1}^{N} |q_i|.$$

NOTE: We used that the entries of any row of A_{ij} add up to 1.

Proof of Step 2 (continued)

Then, we obtain that

$$\sum_{j=1}^{N} |q_j| < \sum_{i=1}^{N} |q_i|$$

clearly a contradiction.

We conclude all the entries of q are nonnegative or nonpositive.

Irreducible chains and stationary distributions

Proof of the theorem

Proof of Step 2 (finished).

In light of this, using that $A_{ij} > 0$ and the identity

$$q_j = \sum_{i=1}^{N} A_{ij} q_i$$

we conclude that each q_j is non-zero, so the entries of q are either all strictly positive, or strictly negative.

This finishes step 2.

$\square$

Proof of Step 3.

What if the eigenspace of p^t for 1 is of dimension > 1?

There are q and q' in this space which are **orthogonal**.

By Step 2, the entries of q and q' are all strictly positive or strictly negative, in particular

$$q \cdot q' = \sum_{i=1}^{N} q_i q'_i \neq 0$$

A contradiction!, therefore the eigenspace is 1-dimensional. $\square$

Example

1.Ehrenfest chain

In this case, the equation $\mathrm{p}^t\pi = \pi$, becomes

$$\pi(i-1)\mathrm{p}(i-1,i) + \pi(i+1)\mathrm{p}(i+1,i) = \pi(i)$$

For $i \neq 0, N$

$$\pi(i-1)\frac{N-i+1}{N} + \pi(i+1)\frac{i+1}{N} = \pi(i)$$

For $i = 0, N$

$$\pi(1)\frac{1}{N} = \pi(0), \;\; \pi(N-1)\frac{1}{N} = \pi(N)$$

Example

1.Ehrenfest chain

As it turns out… there is a unique stationary distribution,

$$\pi(i) = \binom{N}{i} 2^{-N} = \frac{N!}{(N-i)!i!} 2^{-N}$$

Thus, this is the **binomial distribution** $B(N, p)$ with

$$p = 0.5$$

Example

1.Ehrenfest chain

It is not difficult to check that this is a stationary distribution.

The Ehrenfest chain is **irreducible**, so by the previous theorem we know that **there are no other stationary distributions**.

Example

1.Ehrenfest chain

Now, how could someone have found this general formula for the Ehrenfest chain?

It is because in this case $\pi(i)$ happens to satisfy an even stronger condition than

$$\mathrm{p}^t\pi = \pi$$

it has what is known as a **detailed balance condition**.

Example

2.Gambler's ruin

The Gambler's ruin is **not** an irreducible chain.
Accordingly, there is more than one stationary distribution!

Any distribution of the form

$$(\lambda, \ldots, 1 - \lambda)$$

with $\lambda \in [0, 1]$, is a stationary distribution!

Example

3.Doubly stochastic chains

A matrix is said to be **doubly stochastic**, if

$$0 \leq a_{ij} \leq 1,$$
$$\sum_i a_{ij} = 1, \ \forall\, j,$$
$$\sum_j a_{ij} = 1, \ \forall\, i.$$

A chain is then, called doubly stochastic if its transition matrix is itself doubly stochastic.

Example

3.Doubly stochastic chains

A moment of reflection will make one thing clear:

a doubly stochastic chain admits the uniform distribution as a stationary distribution

Classification of states

States in a Markov chain fall in two categories, informally described as follows

1. Those states that are guaranteed to be visited infinitely many times as one lets time go by. These states are called **recurrent states**.
2. Those states for which there is some non zero probability that after some visit, they are never visited again. These states are called **transient states**.

Problem: Given a state x, what is the probability that after visiting state x, the system will eventually return to x at some later time?

Classification of states

Stopping Times

Answering this problem will take us to both an important concept and useful computational tool: **stopping times**.

Given a Markov Chain $X_0, X_1, X_2, \ldots$ a **stopping time** is a random variable T taking the values $0, 1, 2, \ldots$ and having the property that for every n,

$$\{T = n\} = \begin{array}{l}\text{is a event entirely determined by} \\ \text{the values of } X_0, X_1, \ldots, X_n\end{array}$$

The most basic stopping time is the **first time of return to** x or **first hitting time for** x, denoted T_x and defined by

$$T_x := \min\{n \geq 1 \mid X_n = x\}$$

Why is T_x a stopping time? Simply because for every n

$$\{T_x = n\} = \{X_1 \neq x, X_2 \neq x, \ldots, X_{n-1} \neq x, X_n = x\}$$

More generally:
Given a subset $A \subset S$, one has the **first hitting time for** A

$$T_A := \min\{n \geq 1 \mid X_n \in A\}$$

it should be clear that T_A is also a stopping time.

Essentially, the Strong Markov property says the following:

Theorem (Strong Markov Property)

Let X_n denote a Markov Chain. If T is a stopping time and $Y_n := X_{T+n}$, then Y_n is also a Markov chain. Moreover, the transition probability for Y_n is the same as the one for X_n.