Math 456: Mathematical Modeling

Thursday, March 29th, 2018

Decomposition of Markov Chains

Thursday, March 29th, 2018

Last time

1. Counting visiting times.
2. The behavior of $\mathrm{p}^n(x, y)$ as $n \to \infty$ when y is transient.
3. Convergence theorem for irreducible, aperiodic chains (statement only).

Today

1. Examples of use of the convergence theorem.
2. Inequalities for hitting times
3. Decomposition of the state space into transient and closed+irreducible sets.

Last time

Convergence Theorem

Last time, we stated the following theorem

Theorem

Consider an irreducible, aperiodic chain, and let $\pi(y)$ *denote its stationary distribution.*

Then, for any $y \in S$, *we have*

$$\lim_{n\to\infty} \mathrm{p}^n(x,y) = \pi(y) \ \forall\, x \in S.$$

Also last time...

We also proved that if y is transient, then for every x,

$$\sum_{n=1}^{\infty} \mathrm{p}^n(x,y) < \infty$$

and therefore

$$\lim_{n\to\infty} \mathrm{p}^n(x,y) = 0$$

Thus, we have two ways of computing limits of $\mathrm{p}^n(x,y)$, depending on the nature of y.

Example

Problem: Consider the chain with transition matrix

$$\mathrm{p} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2/3 & 0 & 1/3 & 0 \\ 1/8 & 1/4 & 5/8 & 0 & 0 \\ 0 & 1/6 & 0 & 5/6 & 0 \\ 1/3 & 0 & 1/3 & 0 & 1/3 \end{pmatrix}$$

and compute $\lim_{n\to\infty} \mathrm{p}^n(x,y)$ for every x and y.

Example

Solution:

$$\mathrm{p} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2/3 & 0 & 1/3 & 0 \\ 1/8 & 1/4 & 5/8 & 0 & 0 \\ 0 & 1/6 & 0 & 5/6 & 0 \\ 1/3 & 0 & 1/3 & 0 & 1/3 \end{pmatrix}$$

We note that $\mathrm{p}(1, y) = 0$ for all states $y \neq 1$ and $\mathrm{p}(1, 1) = 1$. It follows 1 is a **recurrent state**.

Example

Solution:

$$p = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2/3 & 0 & 1/3 & 0 \\ 1/8 & 1/4 & 5/8 & 0 & 0 \\ 0 & 1/6 & 0 & 5/6 & 0 \\ 1/3 & 0 & 1/3 & 0 & 1/3 \end{pmatrix}$$

We note that $\mathrm{p}(1, y) = 0$ for all states $y \neq 1$ and $\mathrm{p}(1, 1) = 1$. It follows 1 is a **recurrent state**.
Plus, since all the other states communicate with 1, it follows that $\{2, 3, 4, 5\}$ are all **transient states**.

Example

Solution:

$$
\mathrm{p} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2/3 & 0 & 1/3 & 0 \\ 1/8 & 1/4 & 5/8 & 0 & 0 \\ 0 & 1/6 & 0 & 5/6 & 0 \\ 1/3 & 0 & 1/3 & 0 & 1/3 \end{pmatrix}
$$

Accordingly, if $y = 2, 3, 4, 5$ then

$$
\lim_{n\to\infty} \mathrm{p}^n(x, y) = 0 \text{ for every } x
$$

At the same time,

$$
\lim_{n\to\infty} \mathrm{p}^n(x, 1) = 0 \text{ for every } x
$$

Example

(Computing limits **using** the Convergence Theorem)

Problem: Consider the chain with transition matrix

$$p = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1/3 & 0 & 2/3 & 0 \end{pmatrix}$$

and compute $\lim_{n\to\infty} p^n(x,y)$ for every x and y.

Example

Solution:

$$p = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1/3 & 0 & 2/3 & 0 \end{pmatrix}$$

First, it is not difficult to see that every states communicates to each other, so the chain is **irreducible**.

Example

Solution:

$$p = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1/3 & 0 & 2/3 & 0 \end{pmatrix}$$

First, it is not difficult to see that every states communicates to each other, so the chain is **irreducible**.
Secondly, it is clear that $\mathrm{p}^3(1,1) > 0$ and $\mathrm{p}^4(1,1) > 0$, so $x = 1$ has period 1. Then, by irreducibility, all states have period 1.

Example

Solution:

$$\mathrm{p} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1/3 & 0 & 2/3 & 0 \end{pmatrix}$$

Since the chain is irreducible and aperiodic, the **convergence theorem** says that

$$\lim_{n\to\infty} \mathrm{p}^n(x,y) = \pi(y) \ \ \forall\, x, y.$$

Let us then find $\pi(y)$ directly.

Example

Solution:

$$p = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1/3 & 0 & 2/3 & 0 \end{pmatrix}$$

Solving the eigenfunction problem defining the vector π yields

$$\pi^t = (\tfrac{3}{11}, \tfrac{3}{11}, \tfrac{2}{11}, \tfrac{3}{11})$$

Decomposition of the State Space

A different perspective on hitting times

Often computing the exact probabilities

$$\mathbb{P}_x[T_x = n], \quad n = 1, 2, \ldots$$

is too difficult and not necessary.

What if for our purposes, it's enough to know $\mathbb{P}_x[T_x = n]$ is not very large for some n?

Inequalities are always easier to obtain than formulas (they give us less information), but is still useful information.

Decomposition of the State Space

Example

Consider a 3 state chain with transition amtrix

$$\mathrm{p} = \begin{pmatrix} 0.7 & 0.2 & 0.1 \\ 0.3 & 0.5 & 0.2 \\ 0.2 & 0.4 & 0.4 \end{pmatrix}$$

What can we say about, for instance

$$\mathbb{P}_3[T_3 > n]??$$

Decomposition of the State Space

Example

Consider a 3 state chain with transition amtrix

$$p = \begin{pmatrix} 0.7 & 0.2 & 0.1 \\ 0.3 & 0.5 & 0.2 \\ 0.2 & 0.4 & 0.4 \end{pmatrix}$$

Observe that, no matter what x is, we always have

$$\mathbb{P}[X_{n+1} = 3 \mid X_n = x] \geq 0.1$$

Equivalently,

$$\mathbb{P}[X_{n+1} \neq 3 \mid X_n = x] \leq 0.9.$$

Decomposition of the State Space

Example

Consider a 3 state chain with transition amtrix

$$\mathrm{p} = \begin{pmatrix} 0.7 & 0.2 & 0.1 \\ 0.3 & 0.5 & 0.2 \\ 0.2 & 0.4 & 0.4 \end{pmatrix}$$

In this case, we can think of estimating a Bernoulli variable: "have we reached 3 yet y/n?". That is, we have the inequality

$$\mathbb{P}_x[X_n \neq 3, X_{n-1} \neq 3, \ldots, X_1 \neq 3] \leq (0.9)^n.$$

Therefore,

$$\mathbb{P}_x[T_x > n] \leq (0.9)^n.$$

Decomposition of the State Space

This last trick can be generalized:

Lemma (See Lemma 1.3 in Durrett)

If there is a state $y \in S$*, and* $\alpha \in (0,1)$ *and* $k_0 \in \mathbb{N}$ *such that*

$$\mathrm{P}_x[T_y \leq k_0] \geq \alpha \ \ \forall\, x \in S,$$

then for every $n \in \mathbb{N}$*,*

$$\mathrm{P}_x[T_y > nk_0] \leq (1-\alpha)^n.$$

Decomposition of the State Space

The way to think about this lemma is:

if regardless of the initial point you have some at least some chance of reaching the state y *in at most* k_0 *steps, then the chance that you have not reached* y *in* n *steps decreases exponentially with* n*.*

Decomposition of the State Space

Theorem

If $C \subset \mathcal{S}$ is closed and irreducible, then all of the states in C are recurrent.

Decomposition of the State Space

Proof

Recall that if C is closed and irreducible and has N states, then

$$\text{given } x, y \in C \text{ there is } k \leq N \text{ such that } \mathrm{p}^k(x, y) > 0$$

Decomposition of the State Space

Proof

In terms of T_x this means there is a $\delta > 0$ such that

$$\mathrm{P}_x(T_x \leq N) \geq \delta \text{ for every } x \in C$$

Decomposition of the State Space

Proof

In terms of T_x this means there is a $\delta > 0$ such that

$$\mathrm{P}_x(T_x \le N) \ge \delta \text{ for every } x \in C$$

$$(\text{Say } C = \{x_1, x_2, \ldots, x_n\}, \text{ then take } \delta = \min_{1 \le i \le N} \mathrm{P}_{x_i}(T_{x_i} \le N)$$
$$\text{which must be} > 0)$$

Decomposition of the State Space

Proof

Since $\mathrm{P}_x(T_x \leq N) \geq \delta$ for all x, the previous Lemma says that

$$\mathrm{P}_x(T_x \geq kN) \leq (1-\delta)^k$$

Decomposition of the State Space

Proof

Since $\mathrm{P}_x(T_x \leq N) \geq \delta$ for all x, the previous Lemma says that

$$\mathrm{P}_x(T_x \geq kN) \leq (1-\delta)^k$$

Since $\mathrm{P}_x(T_x = \infty) \leq P_x(T_x \geq kN)$ for every k, we conclude that

$$\mathrm{P}_x(T_x = \infty) \leq (1-\delta)^k$$

(for every $k \geq 1$)

Decomposition of the State Space

Proof

Since $\mathrm{P}_x(T_x \leq N) \geq \delta$ for all x, the previous Lemma says that

$$\mathrm{P}_x(T_x \geq kN) \leq (1-\delta)^k$$

Since $\mathrm{P}_x(T_x = \infty) \leq P_x(T_x \geq kN)$ for every k, we conclude that

$$\mathrm{P}_x(T_x = \infty) \leq (1-\delta)^k$$
$$(\text{for every } k \geq 1)$$

This means that $\mathrm{P}_x(T_x = \infty) = 0$ and thus x is recurrent.

Decomposition of the State Space

Now we know that closed, irreducible sets are an easy way to find recurrent states.
The following observation is good for detecting transient states.

Proposition

If x, y are such that $\rho_{xy} > 0$ and $\rho_{yx} < 1$, then x is transient.

Decomposition of the State Space

Now we know that closed, irreducible sets are an easy way to find recurrent states.
The following observation is good for detecting transient states.

Proposition

If x, y are such that $\rho_{xy} > 0$ and $\rho_{yx} < 1$, then x is transient.

Proof (sketch).

Heuristically, once y is reached, there is a non-zero probability of not reaching x ever again. Since it is possible to get from x to y, the probability of never returning to x is positive. □

Decomposition of the State Space

An immediate consequence of the previous theorem is the following

Corollary

If $x \mapsto y$ and x is recurrent, then $y \mapsto x$

Decomposition of the State Space

With these observations in hand, we can now classify all states in a Markov chain, decomposing the underlying state space $\mathcal{S}$

Theorem

For a Markov chain with finite state space S, we have the partition

$$\mathcal{S} = T \cup C_1 \cup C_2 \cup \ldots \cup C_m$$

Where T is the set of transient states and each C_k $(1 \leq k \leq m)$ are closed and irreducible sets comprised of recurrent states.

Decomposition of the State Space

Proof

Let C be the set of recurrent states, and $T = \mathcal{S} \setminus C$ the transient states.

Observe that on C, $x \mapsto y$ becomes a recurrence relation!

Why?

Well, we need to show the following three things

1. If $x \in C$ then $x \mapsto x$.
2. If $x, y \in C$ then $x \mapsto y$ if and only if $y \mapsto y$.
3. If $x, y, z \in C$ and $x \mapsto y, y \mapsto z$ then $x \mapsto z$.

Decomposition of the State Space

Proof (continued)

Since every $x \in C$ is recurrent, it is immediate that $x \mapsto x$.

By the previous corollary, we have that if x is recurrent and $x \mapsto y$, then $y \mapsto x$. So if $x, y \in C$ then $x \mapsto y$ implies that $y \mapsto x$ and vice versa.

Lastly, the third property is immediate by concatenating two appropriate paths of positive probability.

Decomposition of the State Space

Proof (continued).

Since $x \mapsto y$ is an equivalence relation, it defines equivalence classes

$$C_1, \dots, C_m$$

These subsets of C are disjoint from one another, and

$$C = C_1 \cup \dots \cup C_m$$

Then, it is clear that if $x, y \in C_k$ then $x \mapsto y$ and vice versa, and if x and y are in different C_k, then x does not communicate with y. This means each C_k is closed and irreducible, and

□

Decomposition of the State Space

This theorem says that a Markov chain is made out of subsystems which are irreducible Markov Chains that do not communicate between one another (they are close), plus some spurious transient states.

Therefore, a great deal of questions about Markov chains only need to be treated for irreducible chains.