Math 525: Lecture 17

March 14, 2018

For the remainder, it should be understood that by "Markov chain", we mean a Markov chain with finite state space S and which is stationary (and hence admits a transition matrix $P = (P_{ij})$). We will, without loss of generality, take $S = \{1, \dots, n\}$ for the entirety of this lecture.

1 Irreducible matrices

It turns out that the reducibility of the transition matrix P has a very useful interpretation. Let's first recall the notion of reducibility. To do that, we'll need the notion of a permutation matrix:

Definition 1.1. A bijective function $\pi : S \to S$ is called a *permutation* of $S = \{1, \dots, n\}$. The matrix $A = (A_{ij})$ with entries

$$A_{ij} = \begin{cases} 1 & \text{if } \pi(j) = i \\ 0 & \text{otherwise} \end{cases}$$

is the *permutation matrix* associated with π.

Example 1.2. The matrix

$$K = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

is a permutation matrix corresponding to the permutation $\pi(2) = 1$, $\pi(3) = 2$, and $\pi(1) = 3$.

1. Left-multiplication:

$$K \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}.$$

   This corresponds to a "reordering" of the vector by the permutation π (i.e., the i-th entry of the old vector becomes the $\pi(i)$-th entry of new vector).

2. Similarity transform:

$$K \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} K^{\intercal} = \begin{pmatrix} 4 & 5 & 6 \\ 7 & 8 & 9 \\ 1 & 2 & 3 \end{pmatrix} K^{\intercal} = \begin{pmatrix} 5 & 6 & 4 \\ 8 & 9 & 7 \\ 2 & 3 & 1 \end{pmatrix}.$$

This corresponds to a "simultaneous reordering" of the rows and columns by the permutation π (i.e., the (i, j)-th entry of the old matrix becomes the $(\pi(i), \pi(j))$-th entry of the new matrix).

Definition 1.3. Two matrices A and B are said to be *permutation similar* if we can find a permutation matrix K such that

$$KAK^{\intercal} = B.$$

Definition 1.4. A square matrix $A = (A_{ij})$ is said to be *reducible* if it is permutation similar to a block upper triangular matrix:

$$KAK^{\intercal} = \begin{pmatrix} A^{(1)} & A^{(2)} \\ 0 & A^{(3)} \end{pmatrix}$$

where $A^{(1)}$ and $A^{(3)}$ are square matrices (of order at least 1). The matrix A is *irreducible* if it is not reducible. We say a Markov chain is reducible (resp. irreducible) if its transition matrix is reducible (resp. irreducible).

Definition 1.5. Let $A = (A_{ij})$ be a matrix. We say that there is a *walk* from $1 \le u \le n$ to $1 \le v \le n$ if we can find a (nonempty) finite sequence of nonzero entries of A

$$A_{i_1 i_2},\ A_{i_2 i_3},\ \ldots,\ A_{i_{k-1} i_k} \tag{1}$$

such that $i_1 = u$ and $i_k = v$. For brevity, we denote this walk by

$$u = i_1 \dashrightarrow i_2 \dashrightarrow \cdots \dashrightarrow i_k = v.$$

Since we may not want to write i_1, i_2, etc., we will sometimes simply write

$$u \to v$$

to mean that a walk from u to v exists. A walk of length one (e.g., $u \dashrightarrow v$) is called an *edge*.

Remark 1.6. The statements "there is a walk from u to v", "$u \to v$", "v is *reachable* from u", and "v is *accessible* from u" are synonyms.

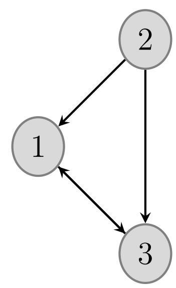
Proposition 1.7. *The square matrix $A = (A_{ij})$ is irreducible if and only if it is "strongly connected". That is, for every pair $1 \le u, v \le n$, we can find a walk from u to v.*

Before we give a proof of Proposition 1.7, let's try to understand the strongly connected property:

Example 1.8. Consider the matrix

$$P = \begin{pmatrix} & & 1 \\ 1/2 & & 1/2 \\ 1 & & \end{pmatrix}.$$

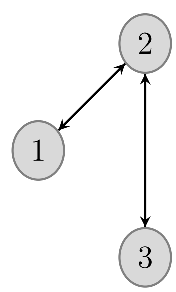
This matrix **does not** satisfy connectedness property since there is no “walk” from vertex $u = 1$ (or $u = 3$) to vertex $v = 2$:



What about the matrix

$$P = \begin{pmatrix} & 1 & \\ 1/2 & & 1/2 \\ & 1 & \end{pmatrix}?$$

Clearly, this matrix satisfies the connectedness property:



Note, in particular, that if there is a walk from u to v, then we can find some positive integer n such that

$$\mathbb{P}(X_n = v \mid X_0 = u) > 0$$

(indeed, n can be taken to be the number of edges in the walk).

Let's return to the proof of Proposition 1.7.

Proof. We prove the reverse direction by contrapositive (i.e., by establishing that reducible $\implies$ not strongly connected). Suppose the matrix is reducible. Then,

$$KAK^{\intercal} = \begin{pmatrix} A^{(1)} & A^{(2)} \\ 0 & A^{(3)} \end{pmatrix} \qquad \text{where } A^{(1)} \in \mathbb{R}^{m \times m} \text{ and } A^{(3)} \in \mathbb{R}^{(n-m)\times(n-m)}. \tag{2}$$

Let $B = KAK^{\intercal}$. It follows that for $u = n$ and $v = 1$, we are unable to find a walk

$$u = i_1 \dashrightarrow i_2 \dashrightarrow \cdots \dashrightarrow i_k = v$$

in B. Equivalently, we are unable to find a walk

$$\pi^{-1}(u) = \pi^{-1}(i_1) \dashrightarrow \pi^{-1}(i_2) \dashrightarrow \cdots \dashrightarrow \pi^{-1}(i_k) = \pi^{-1}(v).$$

in A. Equivalently, we are unable to find a walk

$$\pi^{-1}(u) = i_1' \dashrightarrow i_2' \dashrightarrow \cdots \dashrightarrow i_k' = \pi^{-1}(v).$$

in A. Therefore, A is not strongly connected.

As for the forward direction, we also proceed by contrapositive. The ideas are similar to the previous paragraph, so sketch them briefly. Suppose the matrix is not strongly connected. Then, we can find u and v such that there is no walk from u to v. Let M be the set of all rows **not** reachable by a walk from u. Then, by our assumptions, $v \in M$ so that $|M| = m > 0$. Let π be a permutation under which $\pi(M) = \{1, \ldots, m\}$. Let K be the permutation matrix associated with π. Then, (2) is satisfied, as desired. $\square$

Definition 1.9. If $i \to j$ and $j \to i$, we say that i and j *communicate*, written $i \leftrightarrow j$.

Proposition 1.10. *Communication is an equivalence relation.*

Due to the above, we can divide the state space into equivalence classes $C_1, \ldots, C_k$ such that if $i \in C_r$ and $j \in C_s$, then $i \leftrightarrow j$ if and only if $r = s$.

Exercise 1.11. Show that if the Markov chain is irreducible, the whole state space is a single equivalence class.

2 The period of a nonnegative matrix

Definition 2.1. Let $A = (A_{ij})$ be a nonnegative matrix (i.e., $A_{ij} \geq 0$ for all i, j). Denote by A^n the n-th power of A. The *period of i* is the greatest common divisor of the set $\{n \geq 1 \colon (A^n)_{ii} > 0\}$. We denote the period of i by $d(i)$. If $d(i) = 1$, we say that i is *aperiodic*.

Remark 2.2. The period is also called the *index of imprimitivity* or the *order of cyclicity.*

Before we prove some properties of periods, let's try to grasp the intuition:

Example 2.3. Consider the $n \times n$ transition matrix

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Note that this transition matrix defines a completely deterministic Markov chain. If $X_1 = 1$, then $X_n = n$. In particular, it returns to its original starting place in n steps, so that $d(1) = n$ (more generally, $d(i) = n$ for all i).

We will write $m \mid n$ to mean that the integer m divides the integer n.

Lemma 2.4. *If $a \mid b$ and $a \mid b + c$, then $a \mid c$.*

Proof. If $a \mid b$, we can find an integer k such that $ak = b$. If $a \mid b + c$, we can find an integer k' such that $ak' = b + c$. Then,

$$a\left(k' - k\right) = ak' - ak = b + c - b = c$$

and hence $a \mid c$. □

Lemma 2.5. *If $a \mid b$ and $b \mid a$, then $a = \pm b$.*

Proof. If $a \mid b$, we can find an integer k such that $ak = b$. If $b \mid a$, we can find an integer k' such that $bk' = a$. Then,

$$bk'k = b$$

and hence it must be the case that either $k' = k = 1$ or $k' = k = -1$. □

Proposition 2.6. *If $i \leftrightarrow j$, then $d(i) = d(j)$.*

Proof. Since $i \to j$, we can find a positive integer n such that

$$(P^n)_{ij} = \mathbb{P}(X_n = j \mid X_0 = i) > 0.$$

Similarly, since $j \to i$, we can find a positive integer m such that

$$(P^m)_{ji} > 0.$$

Therefore,

$$(P^{n+m})_{ii} = (P^n P^m)_{ii} = \sum_k (P^n)_{ik}(P^m)_{ki} \geq (P^n)_{ij}(P^m)_{ji} > 0.$$

This implies $d(i) \mid n + m$.

Now, if $(P^r)_{jj} > 0$ for some r, the same reasoning yields

$$(P^{n+r+m})_{ii} = (P^n P^r P^m)_{ii} = \sum_k \sum_{k'} (P^n)_{ik}(P^r)_{kk'}(P^m)_{k'i} \geq (P^n)_{ij}(P^r)_{jj}(P^m)_{ji} > 0.$$

Therefore, $d(i) \mid n + r + m$, and hence $d(i) \mid r$. But note that the above is trivially satisfied with $r = d(j)$, and hence $d(i) \mid d(j)$.

Symmetrically, we can establish $d(j) \mid d(i)$ to conclude that $d(i) = d(j)$. □

Corollary 2.7. *For an irreducible Markov chain, $d(i) = d(j)$ for all i and j.*

Remark 2.8. Due to the above, we refer to the period of any state in an irreducible Markov chain as **the** period of the Markov chain.

3 Primitive matrices

Definition 3.1. A nonnegative matrix A is *primitive* if there exists a positive integer m such that A^m is positive (i.e., $(A^m)_{ij} > 0$ for all i, j).

We will use the following number-theoretic fact without proof:

Lemma 3.2 (Semigroup lemma)**.** *Any set of non-negative integers which is closed under addition and which has greatest common divisor 1 must contain all but finitely many of the non-negative integers.*

Proposition 3.3. *Let A be an irreducible and aperiodic matrix. Then, A is primitive.*

Proof. Let A be an irreducible and aperiodic matrix. Note that $J_i = \{n \colon (A^n)_{ii} > 0\}$ satisfies $\gcd(J_i) = 1$ by the presumed aperiodicity. Therefore, by lemma 3.2, we can find $M(i)$ such that for all $n \geq M(i)$, we have $(A^n)_{ii} > 0$. Since A is irreducible, we can find $m(i,j)$ such that $(A^{m(i,j)})_{ij} > 0$. Therefore, for $n \geq M(i)$,

$$(A^{n+m(i,j)})_{ij} = \sum_k (A^n)_{ik}(A^{m(i,j)})_{kj} \geq (A^n)_{ii}(A^{m(i,j)})_{ij} > 0.$$

Let $M = \max_{i,j} M(i) + m(i,j)$. Then, $(A^M)_{ij} > 0$ for all i, j as desired. $\square$

Definition 3.4. A Markov chain whose transition matrix is primitive is called *regular.*