Lecture 30: More Markov Chains (§5.3)

Markov Chain:

Finite set of states: $\{s_1, \dots, s_m\}$

At each time step, there is a fixed probability of transitioning from s_j to s_i. [No memory.]

State at time n: $p_n \in \mathbb{R}^m$ with $(p_n)_i = \begin{pmatrix}\text{prob in} \\ \text{state } s_i\end{pmatrix}$.

Define $A \in M_{m\times m}(\mathbb{R})$ by $A_{ij} = \begin{pmatrix}\text{prob of} \\ s_j \longrightarrow s_i\end{pmatrix}$.

Then $p_{n+1} = A p_n = A^n p_0$

[Seen three examples that stabilized; mention generalizations.]

Def: Probability vector: $p \in \mathbb{R}^n$ with all $p_i \geq 0$ and $\sum p_i = 1$.

Transition matrix: $A \in M_{n\times n}(\mathbb{R})$ with all $A_{ij} \geq 0$ where each column sums to 1.

To understand long-term behavior of a Markov chain need to understand $\lim_{n\to\infty} A^n$ where A is a transition matrix.

<u>Thm</u>: Suppose A is a transition matrix where there is a $d \geq 1$ with all entries of A^d positive. Then

a) 1 is an eigenvalue for A and $\dim E_1 = 1$. Moreover, E_1 can be spanned by a probability vector u.

b) Any other eigenvalue λ has $|\lambda| < 1$.

c) $\lim_{n \to \infty} A^n = \begin{pmatrix} | & & | \\ u & \cdots & u \\ | & & | \end{pmatrix}$

<u>Cor</u>: No matter what the initial state of the corresp. Markov chain, the p_n limit on u.

<u>Non-Ex</u>: $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

[Can't prove whole theorem as need to deal with non-diagonalizable matrices.]

<u>Thm</u>: If A is a transition matrix, then

a) 1 is an eigenvalue of A.

b) Any eigenvalue λ has $|\lambda| \leq 1$.

<u>Proof</u>: By HW, A and A^t have the same eigenvalues. <u>Key</u>: <u>Rows</u> of A^t sum to 1.

a) Setting $u = (1, \dots, 1) \in \mathbb{R}^m$ we have

$$A^t u = \begin{pmatrix} \text{sum of row } 1 \\ \vdots \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = u.$$

So 1 is an eigenvalue of A^t and hence of A.

b) Suppose $v = (v_1, v_2, \dots, v_m)$ is an eigenvector of A^t with eigenvalue λ. Suppose $|v_k| = \max_i |v_i|$. Then $(A^t v)_k = (\lambda v)_k = \lambda v_k$ and $(A^t v)_k = \sum_{j=1}^{m} A^t_{kj} v_j$. Hence

$$|\lambda v_k| = | \quad | \le \sum_{j=1}^{m} |A^t_{kj}||v_j|$$

$$\le \sum_{j=1}^{m} |A^t_{kj}||v_k| = \left(\sum_{j=1}^{m} A^t_{kj}\right)|v_k|$$

$$= |v_k|$$

As $v_k \neq 0$ (since v is an eigenvector!) this gives $|\lambda| \le 1$ as desired. $\blacksquare$

Thm: Suppose A is a transition matrix where every $A_{ij} > 0$. Then $\dim E_1 = 1$ and any eigenvalue $\lambda \neq 1$ has $|\lambda| < 1$.

Proof: Follow setup of (b) from last theorem, where now v is an eigenvector of A^t with eigenvalue λ with $|\lambda| = 1$. Will show $\lambda = 1$ and $v = c \cdot (1, 1, \ldots, 1)$. Now is $|v_k|$ is maximal:

$$\underbrace{|\lambda v_k|}_{\overset{\parallel}{|v_k|}} = \left| \sum_{j=1}^{m} A^t_{kj} v_j \right| \leq \sum_{j=1}^{m} A^t_{kj} |v_j| \leq \sum_{j=1}^{m} A^t_{kj} |v_k| = |v_k| .$$

Thus the two inequalitres must actually be equalities. The second one gives $|v_j| = |v_k|$ for all j. The first one forces all the v_j to have the same argument, and so $v_j = v_k$ for all j. Hence $v = c(1, \ldots, 1)$ for $c \in \mathbb{C}$, and $\lambda = 1$ as needed. $\blacksquare$

Thm: Suppose A is a transition matrix where all $A_{ij} > 0$ and which is diagonalizable. Then

$$\lim_{n\to\infty} A^n = \begin{pmatrix} | & & | \\ u & \cdots & u \\ | & & | \end{pmatrix}$$ where u is a prob. vector which is an eigenvector of A with eigenvalue 1.

Pf: Know $A = QDQ^{-1}$ where $D = \begin{pmatrix} 1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_m \end{pmatrix}$ and all $|\lambda_i| < 1$ for $i \geq 2$. Thus

$$\lim_{n\to\infty} A^n = \lim_{n\to\infty} QD^nQ^{-1} = Q\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & O & \\ 0 & & & \end{pmatrix}Q^{-1}.$$

and let's call the limit L. Now each A^n is a transition matrix (excercise) and hence L is as well. Notice that $L = AL$ as $\lim_{n\to\infty} A^n = A \cdot \lim_{n\to\infty} A^{n-1} = AL$. Consequently, the columns of L must be eigenvectors of A with eigenvalue 1. As $\dim(E_1) = 1$ and each column of L is a prob. vector, we conclude that all columns of L must be the same, as needed. ▨

Generalized diagonalization: Jordan Canonical Form.

Any $A \in M_{n\times n}(\mathbb{C})$ is similar to one of the form

$$\begin{pmatrix} \boxed{B_1} & & & & \\ & \boxed{B_2} & & 0 & \\ & & \boxed{\ddots} & & \\ & 0 & & \ddots & \\ & & & & \boxed{B_k} \end{pmatrix}$$

where each block (which may have diff sizes) is of the form $\begin{pmatrix} \lambda_i & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{pmatrix}$

Ex:
$$\begin{pmatrix} 2 & 1 & 0 & & & & & & \\ 0 & 2 & 1 & & & & 0 & & \\ 0 & 0 & 2 & & & & & & \\ & & & 2 & & & & & \\ & & & & 3 & 1 & & & \\ & 0 & & & & 3 & & & \\ & & & & & & 5 & & \\ & & & & & & & 0 & 1 \\ & & & & & & & & 0 \end{pmatrix}$$