Markov Chain:

Finite set of states: \( \{s_1, \ldots, s_n\} \)

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

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

Define \( A \in M_{n \times n}(\mathbb{R}) \) by \( A_{ij} = (\text{prob of } s_j \to s_i) \).

Then \( p_{n+1} = Ap_n = A^np_0 \)

[See 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.
**Thm:** 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 $\lambda$ has $|\lambda| < 1$.

c) $\lim_{n \to \infty} A^n = \left( \frac{1}{d}, \cdots, \frac{1}{d} \right)$

**Cev:** No matter what the initial state of the corres. Markov chain, the $\lim_{n \to \infty}$ limit on $u$.

**Non-Ex:** $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

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

**Thm:** If $A$ is a transition matrix, then

a) 1 is an eigenvalue of $A$.

b) Any eigenvalue $\lambda$ has $|\lambda| \leq 1$. 
Proof: By HW, $A$ and $A^t$ have the same eigenvalues. Key: Rows of $A^t$ sum to 1.

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

$$A^t u = \left( \sum_{i=1}^{n} u_i \right) \hat{u} = \hat{u} = u.$$  

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

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

$$|\lambda v_k| = |v_k| = \sum_{j=1}^{n} |A^t_{kj}| |v_j| \leq \sum_{j=1}^{n} |A^t_{kj}| |v_k| = \left( \sum_{j=1}^{n} |A^t_{kj}| \right) |v_k| = |v_k|$$

As $v_k \neq 0$ (since $v$ is an eigenvector!) this gives $|\lambda| \leq 1$ as desired. \qed
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 \( \nu \) is an eigenvector of \( A^t \) with eigenvalue \( \lambda \) with \( |\lambda| = 1 \). Will show \( \lambda = 1 \) and \( \nu = c \cdot (1, 1, \ldots, 1) \). Now is \( |\nu_k| \) is maximal:

\[
|\lambda \nu_k| = |\sum_{j=1}^{n} A_{kj}^t \nu_j| \leq \sum_{j=1}^{n} A_{kj}^t |\nu_j| \leq \sum_{j=1}^{n} A_{kj}^t |\nu_k| = |\nu_k|
\]

Thus the two inequalities must actually be equalities. The second one gives \( |\nu_j| = |\nu_k| \) for all \( j \). The first one forces all the \( \nu_j \) to have the same argument, and so \( \nu_j = \nu_k \) for all \( j \). Hence \( \nu = c(1, \ldots, 1) \) for \( c \in \mathbb{C} \), and \( \lambda = 1 \) as needed. \( \square \)
Thm: Suppose $A$ is a transition matrix where all $A_{ij} > 0$ and which is diagonalizable. Then

$$\lim_{n \to \infty} A^n = (u_1 \ldots u_1)$$

where $u$ is a prob. vector which is an eigenvector of $A$ with eigenvalue $1$.

Pf: Know $A = QDQ^{-1}$ where $D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$

and all $|\lambda_i| < 1$ for $i \geq 2$. Thus

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

and let's call the limit $L$. Now each $A^n$ is a transition matrix (exercise) 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. \(\Box\)
Generalized diagonalization: Jordan Canonical Form.

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

\[
\begin{pmatrix}
B_1 & & & \\
& B_2 & & \\
& & \ddots & \\
& & & B_k
\end{pmatrix}
\]

where each block (which may have different sizes) is of the form

\[
\begin{pmatrix}
\lambda & 1 & 0 \\
0 & \ddots & 0 \\
0 & 0 & \lambda_i
\end{pmatrix}
\]

Ex:

\[
\begin{pmatrix}
2 & 1 & 0 \\
0 & 2 & 1 \\
0 & 0 & 2
\end{pmatrix}
\begin{pmatrix}
0 & & \\
& 2 & 3 \\
& & 3
\end{pmatrix}
\begin{pmatrix}
1 & 0 \\
0 & 1 \\
0 & 0
\end{pmatrix}
\]

\[
\begin{pmatrix}
0 & & \\
& 2 & 3 \\
& & 3
\end{pmatrix}
\begin{pmatrix}
0 & & \\
& 2 & 3 \\
& & 3
\end{pmatrix}
\begin{pmatrix}
1 & 0 \\
0 & 1 \\
0 & 0
\end{pmatrix}
\]