Linear Algebra Markov Chain Example

Introduction

Here is an example of a Markov chain suitable for presentation in elementary Linear Algebra. It uses basic results from Math 304 to understand the long-term behavior of a discrete model where a sequence of states, $X(k)$, $0 \leq k \in \mathbb{Z}$, is determined by a square matrix, $A \in \mathbb{R}^4_4$, by

$$X(k+1) = AX(k) = A^k X(0).$$

Each state vector $X(k) \in \mathbb{R}^4$ has coordinates which add up to 1, so the j^{th} coordinate means the probability that a "particle" is in the j^{th} position, $1 \leq j \leq 4$.

We want to understand the long-term behavior of this model as $k \to \infty$. The idea is to diagonalize A, that is, find an invertible matrix P such that $P^{-1}AP = D = diag(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ is diagonal. Then $D^k = diag(\lambda_1^k, \lambda_2^k, \lambda_3^k, \lambda_4^k)$ and $A = PDP^{-1}$ so $A^k = PD^kP^{-1}$.

Let $S = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$ be the standard basis of $\mathbb{R}^4$. These are possible state vectors, where $\mathbf{e}_j$ corresponds to the particle being located in position j with probability 1. In the following matrix, $A\mathbf{e}_j = Col_j(A)$ comes from assuming that a particle in position j at step k has probability of being in another position at step $k+1$ given by the coordinates of $Col_j(A)$. Let

$$A = \begin{bmatrix} 0 & \frac{1}{2} & 0 & 0 \\ 1 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & 1 \\ 0 & 0 & \frac{1}{2} & 0 \end{bmatrix}$$

so $\mathbf{e}_1 \to \mathbf{e}_2$ with probability 1, $\mathbf{e}_2 \to \mathbf{e}_1$ or $\mathbf{e}_2 \to \mathbf{e}_3$ each with probability $\frac{1}{2}$, $\mathbf{e}_3 \to \mathbf{e}_2$ or $\mathbf{e}_3 \to \mathbf{e}_4$ each with probability $\frac{1}{2}$, and $\mathbf{e}_4 \to \mathbf{e}_3$ with probability 1.

We want to calculate the characteristic polynomial of A, whose roots will be the eigenvalues. One of those eigenvalues must be 1 for the following reasons.

Note that the transpose of A has obvious eigenvector $\begin{bmatrix}1\\1\\1\\1\end{bmatrix}$ with eigenvalue 1 because the entries of each row of A^T add up to 1. Also, $char_A(t) = char_{A^T}(t)$ because

$$\det(A - tI_4) = \det((A - tI_4)^T) = \det(A^T - tI_4)$$

so A and A^T have the same eigenvalues. We find, after $tRow_2 + Row_1 \to Row_1$, and then doing cofactor expansion along column 1,

$$char_A(t) = \det\begin{bmatrix}-t & \frac{1}{2} & 0 & 0\\ 1 & -t & \frac{1}{2} & 0\\ 0 & \frac{1}{2} & -t & 1\\ 0 & 0 & \frac{1}{2} & -t\end{bmatrix} = \det\begin{bmatrix}0 & \frac{1}{2} - t^2 & \frac{t}{2} & 0\\ 1 & -t & \frac{1}{2} & 0\\ 0 & \frac{1}{2} & -t & 1\\ 0 & 0 & \frac{1}{2} & -t\end{bmatrix} =$$

$$-\det\begin{bmatrix}\frac{1}{2} - t^2 & \frac{t}{2} & 0\\ \frac{1}{2} & -t & 1\\ 0 & \frac{1}{2} & -t\end{bmatrix} = \frac{1}{2}\det\begin{bmatrix}\frac{1}{2} - t^2 & 0\\ \frac{1}{2} & 1\end{bmatrix} - (-t)\det\begin{bmatrix}\frac{1}{2} - t^2 & \frac{t}{2}\\ \frac{1}{2} & -t\end{bmatrix} =$$

$$\frac{1}{2}\left(\frac{1}{2}-t^2\right)+t\left(\left(\frac{1}{2}-t^2\right)(-t)-\frac{t}{4}\right)=\frac{1}{4}-\frac{t^2}{2}+t\left(t^3-\frac{t}{2}-\frac{t}{4}\right)=$$
$$t^4-\left(\frac{1}{2}+\frac{1}{2}+\frac{1}{4}\right)t^2+\frac{1}{4}=t^4-\frac{5}{4}t^2+\frac{1}{4}=\left(t^2-\frac{1}{4}\right)(t^2-1)=$$
$$\left(t-\frac{1}{2}\right)\left(t+\frac{1}{2}\right)(t-1)(t+1).$$

We used another cofactor expansion of the 3×3 matrix along row 3 to get two 2×2 matrices.

So we got four distinct eigenvalues, $\lambda_1=\frac{1}{2}$, $\lambda_2=-\frac{1}{2}$, $\lambda_3=1$ and $\lambda_4=-1$, each with algebraic multiplicity 1. For each of these we must find a basis vector for the 1-dimensional eigenspace, and put them together to get the columns of transition matrix, P.

Plug $\lambda_1 = \frac{1}{2}$ into $A - tI_4$. Row reduce

$$\left[\begin{array}{cccc|c} -\frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 1 & -\frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} & 1 & 0 \\ 0 & 0 & \frac{1}{2} & -\frac{1}{2} & 0 \end{array}\right] \text{ to } \left[\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \text{ so } \begin{array}{c} x_1 = -r \\ x_2 = -r \\ x_3 = r \\ x_4 = r \in \mathbb{R} \end{array}, \text{ then}$$

$$A_{\lambda_1} = \left\{ \begin{bmatrix} -r \\ -r \\ r \\ r \end{bmatrix} \in \mathbb{R}^4 \;\middle|\; r \in \mathbb{R} \right\} \text{ has basis } \left\{ w_1 = \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

Plug $\lambda_2 = -\frac{1}{2}$ into $A - tI_4$. Row reduce

$$\left[\begin{array}{cccc|c} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 1 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 1 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{array}\right] \text{ to } \left[\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \text{ so } \begin{array}{c} x_1 = r \\ x_2 = -r \\ x_3 = -r \\ x_4 = r \in \mathbb{R} \end{array}, \text{ then}$$

$$A_{\lambda_2} = \left\{ \begin{bmatrix} r \\ -r \\ -r \\ r \end{bmatrix} \in \mathbb{R}^4 \;\middle|\; r \in \mathbb{R} \right\} \text{ has basis } \left\{ w_2 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} \right\}.$$

Plug $\lambda_3 = 1$ into $A - tI_4$. Row reduce

$$\left[\begin{array}{cccc|c} -1 & \frac{1}{2} & 0 & 0 & 0 \\ 1 & -1 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & -1 & 1 & 0 \\ 0 & 0 & \frac{1}{2} & -1 & 0 \end{array}\right] \text{ to } \left[\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & -2 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \text{ so } \begin{array}{c} x_1 = r \\ x_2 = 2r \\ x_3 = 2r \\ x_4 = r \in \mathbb{R} \end{array}, \text{ then}$$

$$A_{\lambda_3} = \left\{ \begin{bmatrix} r \\ 2r \\ 2r \\ r \end{bmatrix} \in \mathbb{R}^4 \;\middle|\; r \in \mathbb{R} \right\} \text{ has basis } \left\{ w_3 = \begin{bmatrix} 1 \\ 2 \\ 2 \\ 1 \end{bmatrix} \right\}.$$

Plug $\lambda_3 = -1$ into $A - tI_4$. Row reduce

$$\left[\begin{array}{cccc|c} 1 & \frac{1}{2} & 0 & 0 & 0 \\ 1 & 1 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 1 & 1 & 0 \\ 0 & 0 & \frac{1}{2} & 1 & 0 \end{array}\right] \text{ to } \left[\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -2 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \text{ so } \begin{array}{c} x_1 = -r \\ x_2 = 2r \\ x_3 = -2r \\ x_4 = r \in \mathbb{R} \end{array}, \text{ then}$$

$$A_{\lambda_4} = \left\{ \begin{bmatrix} -r \\ 2r \\ -2r \\ r \end{bmatrix} \in \mathbb{R}^4 \;\middle|\; r \in \mathbb{R} \right\} \text{ has basis } \left\{ w_4 = \begin{bmatrix} -1 \\ 2 \\ -2 \\ 1 \end{bmatrix} \right\}.$$

Then we found an eigen-basis $T = \{w_1, w_2, w_3, w_4\}$ such that the linear map $L_A : \mathbb{R}^4 \to \mathbb{R}^4$ defined as usual by $L_A(X) = AX$, is represented with respect to the standard basis, S, by A, but is represented with respect to T by the diagonal matrix $D = diag(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$.

We now have the transition matrix, $P = {}_SP_T$ where $Col_j(P) = w_j$, and we get its inverse by row reducing $[P|I_4] \to [I_4|P^{-1}]$:

$$P = \begin{bmatrix} -1 & 1 & 1 & -1 \\ -1 & -1 & 2 & 2 \\ 1 & -1 & 2 & -2 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad \text{and} \quad P^{-1} = \frac{1}{6}\begin{bmatrix} -2 & -1 & 1 & 2 \\ 2 & -1 & -1 & 2 \\ 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & 1 \end{bmatrix}.$$

As explained before, we could now compute $A^k = PD^kP^{-1} =$

$$\begin{bmatrix} -1 & 1 & 1 & -1 \\ -1 & -1 & 2 & 2 \\ 1 & -1 & 2 & -2 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} (1/2)^k & 0 & 0 & 0 \\ 0 & (-1/2)^k & 0 & 0 \\ 0 & 0 & 1^k & 0 \\ 0 & 0 & 0 & (-1)^k \end{bmatrix} \frac{1}{6} \begin{bmatrix} -2 & -1 & 1 & 2 \\ 2 & -1 & -1 & 2 \\ 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & 1 \end{bmatrix}$$

but this expression is quite complicated. To simplify our calculations, we first compute

$$PD^k = [(1/2)^k Col_1(P)|(-1/2)^k Col_2(P)|Col_3(P)|(-1)^k Col_4(P)]$$

and see that as k gets very large, the first two columns go to the zero column 0_1^4. The third and fourth columns only depend on whether k is even or odd.

So we assume now that k is so large that we can take those first two columns to be zero columns and get a simplified computation of

$$PD^kP^{-1} = \begin{bmatrix} 0 & 0 & 1 & -(-1)^k \\ 0 & 0 & 2 & 2(-1)^k \\ 0 & 0 & 2 & -2(-1)^k \\ 0 & 0 & 1 & (-1)^k \end{bmatrix} \frac{1}{6} \begin{bmatrix} -2 & -1 & 1 & 2 \\ 2 & -1 & -1 & 2 \\ 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & 1 \end{bmatrix} =$$

$$\frac{1}{6} \begin{bmatrix} 1+(-1)^k & 1-(-1)^k & 1+(-1)^k & 1-(-1)^k \\ 2(1-(-1)^k) & 2(1+(-1)^k) & 2(1-(-1)^k) & 2(1+(-1)^k) \\ 2(1+(-1)^k) & 2(1-(-1)^k) & 2(1+(-1)^k) & 2(1-(-1)^k) \\ 1-(-1)^k & 1+(-1)^k & 1-(-1)^k & 1+(-1)^k \end{bmatrix}.$$

Finally, we get one answer when k is even:

$$\frac{1}{6} \begin{bmatrix} 2 & 0 & 2 & 0 \\ 0 & 4 & 0 & 4 \\ 4 & 0 & 4 & 0 \\ 0 & 2 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 1/3 & 0 & 1/3 & 0 \\ 0 & 2/3 & 0 & 2/3 \\ 2/3 & 0 & 2/3 & 0 \\ 0 & 1/3 & 0 & 1/3 \end{bmatrix}$$

and another answer when k is odd:

$$\frac{1}{6}\begin{bmatrix} 0 & 2 & 0 & 2 \\ 4 & 0 & 4 & 0 \\ 0 & 4 & 0 & 4 \\ 2 & 0 & 2 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1/3 & 0 & 1/3 \\ 2/3 & 0 & 2/3 & 0 \\ 0 & 2/3 & 0 & 2/3 \\ 1/3 & 0 & 1/3 & 0 \end{bmatrix}.$$

This means that after a large **even** number of steps, if the initial state vector $X(0)$ is either $\mathbf{e}_1$ or $\mathbf{e}_3$ then there is a probability of $1/3$ that the particle is in position 1 and probability of $2/3$ that it is in position 3, but if $X(0)$ is either $\mathbf{e}_2$ or $\mathbf{e}_4$ then it has a probability of $2/3$ of being in position 2 and probability $1/3$ of being in position 4.

After a large **odd** number of steps, if the initial state vector $X(0)$ is either $\mathbf{e}_1$ or $\mathbf{e}_3$ then there is a probability of $2/3$ that the particle is in position 2 and probability of $1/3$ that it is in position 4, but if $X(0)$ is either $\mathbf{e}_2$ or $\mathbf{e}_4$ then it has a probability of $1/3$ of being in position 1 and probability $2/3$ of being in position 3.