Sequential Models in Data Science

Markov Chains and Endomorphism Reduction

Ariel Fuxman & Gilad Zusman

June 2023

1 Some Questions on Endomorphisms

1. Before moving on to the proof, we need to prove a Lemma.

Lemma 1. *The matrix $(I - \omega M)$ is invertible.*

Proof. Since $\frac{1}{\omega}$ is not an eigenvalue of M,

$$\det(I - \omega M) = \det\left(\omega\left(\frac{1}{\omega}I - M\right)\right) = \omega^n \det\left(\frac{1}{\omega}I - M\right) \neq 0$$

so $(I - \omega M)$ is invertible. □

Denote

$$S = I_n + \omega M + \ldots + \omega^{p-1}M^{p-1}$$

Remembering that ω is a scalar,

$$\omega MS = \omega M\left(I_n + \omega M + \ldots + \omega^{n-1}M^{n-1}\right) = \omega M + \ldots + \omega^{p-1}M^{p-1} + \underbrace{\omega^p M^p}_{I_n} = S$$

$$S - \omega MS = 0$$

$$(I - \omega M)\,S = 0$$

By Lemma 1, the matrix $(I - \omega M)$ is invertible. So we can multiply by its inverse:

$$(I - \omega M)^{-1}(I - \omega M)\,S = 0 \cdot (I - \omega M)^{-1}$$

$$S = 0$$

2. u is a linear transformation since for all $a, b \in \mathbb{R}$ and $P, Q \in E_n$,

$$\begin{aligned} u\left(aP + bQ\right) &= \left(x-a\right)\left(x-b\right)\left(aP+bQ\right)' - nx\left(aP+bQ\right) = \\ &= a\left(\left(x-a\right)\left(x-b\right)P' - nxP\right) + b\left(\left(x-a\right)\left(x-b\right)Q' - nxQ\right) = \\ &= au(P) + bu(Q) \end{aligned}$$

Since $\{1, x, \ldots, x^n\}$ is a basis for E_n, there exists a unique sequence of coefficients $c_0, \ldots, c_n$, such that:

$$\begin{aligned} P &= \sum_{k=0}^{n} c_k x^k \\ P' &= \sum_{k=0}^{n} k c_k x^{k-1} \\ \left(x-a\right)P' &= \sum_{k=0}^{n} k c_k x^k - a\sum_{k=0}^{n} k c_k x^{k-1} \\ \left(x-b\right)\left(x-a\right)P' &= \sum_{k=0}^{n} k c_k x^{k+1} - a\sum_{k=0}^{n} k c_k x^k - b\sum_{k=0}^{n} k c_k x^k + ab\sum_{k=0}^{n} k c_k x^{k-1} \end{aligned}$$

The only monomial with power greater than x^n is x^{n+1} with coefficient nc_n. In addition,

$$nxP = n\sum_{k=0}^{n} c_k x^{k+1}$$

has x^{n+1} with coefficient nc_n. Meaning that $u\left(P\right) \in E_n$, so $u \in \mathrm{End}\left(E_n\right)$.

Now that we showed that $u \in \mathrm{End}\left(E_n\right)$ we can move on to the eigenvalues and eigenspaces.

Proposition 1. *Assume $a \neq b$. Then, there are $n+1$ different eigenvalues. Moreover, the (eigenvalue, eigenvector) pairs are:*

$$\left(-\left(n-k\right)a - kb, \left(x-a\right)^k\left(x-b\right)^{n-k}\right), \quad k = 0, \ldots, n$$

Proof. We directly show the polynomial equality:

$$\left(x-a\right)\left(x-b\right)P' = \left(nx+\lambda\right)P$$

Denote

$$
\begin{aligned}
P &= (x-a)^k\,(x-b)^{n-k} \\
\implies P' &= k\,(x-a)^{k-1}\,(x-b)^{n-k} + (n-k)\,(x-a)^k\,(x-b)^{n-k-1} \\
\implies (x-a)\,(x-b)\,P' &= (x-a)^k\,(x-b)^{n-k}\,(k(x-b)+(n-k)(x-a)) = \\
&= \left(nx + \underbrace{-\,(n-k)\,a - kb}_{\lambda}\right)\underbrace{(x-a)^k\,(x-b)^{n-k}}_{P}
\end{aligned}
$$

We know that we didn't miss any other eigenvalues or eigenvectors. This is because of two Theorems for endomorphism on finite-dimensional spaces:

1. The geomertic multiplicity of an eigenvalue is smaller or equal to its algebraic multiplcity.
2. The sum of the algebraic multiplicities of the eigenvalues is equal to the dimension of the space.

In our case, there are $n+1$ eigenvalues, each with algebraic and geometric multiplicity of 1. □

Now we need to consider the case $a = b$.

Proposition 2. *Assume $a = b$. Then, there is only one eigenvalue $\lambda = -na$ with eigenvector $(x-a)^n$.*

Proof. We use the method of induction. In the base case $n = 0$, u is the zero transformation, and hence $\lambda = 0$ is the only eigenvalue with eigenvector $P(x) = (x-a)^0 = 1$. Now, let $n > 0$ and assume the proposition is true for $n-1$. We need to find scalars $\lambda \in \mathbb{R}$ and non-zero polynomials P such that:

$$
(x-a)^2\,P' = (nx+\lambda)\,P \tag{1}
$$

This equality must be true for all $x \in \mathbb{R}$. In particular, for $x = a$, we must have:

$$
(na+\lambda)\,P(a) = 0
$$

Assume $P(a) \neq 0$. Then, we can divide both sides by $P(a)$ and conclude that $\lambda = -na$. Returning back to equation (1),

$$
(x-a)^2\,P' = n\,(x-a)\,P
$$

$$
(x-a)\,P' = nP
$$

Substituting $x = a$ again, we must have that:

$$nP(a) = 0$$

we assumed $n > 0$ so we can divide both sides by n and conclude that $P(a) = 0$, which a contradiction. Therefore, a must be a root of $P(x)$. So $P(x) = (x - a)\, Q(x)$, with $Q \in E_{n-1}$. Under this representation,

$$P' = Q(x) + Q'(x)\,(x - a)$$

returning to equation (1) again,

$$(x - a)^2 \big(Q(x) + Q'(x)\,(x - a)\big) = (nx + \lambda)\,(x - a)\, Q(x)$$

$$(x - a)\big(Q(x) + Q'(x)\,(x - a)\big) = (nx + \lambda)\, Q(x)$$

$$(x - a)^2\, Q'(x) = ((n - 1)x + \lambda + a)\, Q(x)$$

Denote $\mu = \lambda + a$. Then, we arrive at:

$$(x - a)^2\, Q'(x) = ((n - 1)x + \mu)\, Q(x)$$

by induction, this equation has the unique eigenvalue $\mu = -\,(n - 1)$ and eigenvector $Q(x) = (x - a)^{n-1}$. So the only solution to the initial problem is $\lambda = -na$ and $P(x) = (x - a)^n$. □

To find $\ker(u)$, we need to know if $\lambda = 0$ is an eigenvalue. If $\lambda = 0$ is not an eigenvalue, then the kernel is trivial. Otherwise, it is the eigenspace corresponding to $\lambda = 0$. This leads to the following Corollary:

Corollary. *If $a = 0$, then*

$$\ker(u) = span\{(x - b)^n\},$$

and if $b = 0$, then

$$\ker(u) = span\{(x - a)^n\}.$$

3. We begin by finding the eigenvalues of

$$A = \begin{pmatrix} -8 & 1 & 5 \\ 2 & -3 & 1 \\ -4 & 1 & 1 \end{pmatrix}$$

$$P_A(\lambda) = \begin{vmatrix} -8-\lambda & 1 & 5 \\ 2 & -3-\lambda & -1 \\ -4 & 1 & 1-\lambda \end{vmatrix} = \begin{vmatrix} -8-\lambda & 1 & 5 \\ 2 & -3-\lambda & -1 \\ 4+\lambda & 0 & -4-\lambda \end{vmatrix} =$$

$$= \begin{vmatrix} -8-\lambda & 1 & -3-\lambda \\ 2 & -3-\lambda & 1 \\ 4+\lambda & 0 & 0 \end{vmatrix} = (4+\lambda) \begin{vmatrix} 1 & -3-\lambda \\ -3-\lambda & 1 \end{vmatrix} = -(4+\lambda)^2(\lambda+2)$$

So the eigenvalues are -2 with algebraic multiplicity of 1, and -4 with algebraic multiplicity of 2. For $\lambda = -2$ we find $\operatorname{Ker}(A - \lambda I)$:

$$\begin{pmatrix} -6 & 1 & 5 \\ 2 & -1 & -1 \\ -4 & 1 & 3 \end{pmatrix} \Longrightarrow \begin{pmatrix} 2 & -1 & -1 \\ -6 & 1 & 5 \\ -4 & 1 & 3 \end{pmatrix} \Longrightarrow \begin{pmatrix} 2 & -1 & -1 \\ 0 & -2 & 2 \\ 0 & -1 & 1 \end{pmatrix} \Longrightarrow \begin{pmatrix} 2 & -1 & -1 \\ 0 & -2 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

So the eigenvector is:

$$v^{(1)} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Now we find $\operatorname{Ker}(A - \lambda I)$ for $\lambda = -4$:

$$\begin{pmatrix} -4 & 1 & 5 \\ 2 & 1 & -1 \\ -4 & 1 & 5 \end{pmatrix} \Longrightarrow \begin{pmatrix} -4 & 1 & 5 \\ 2 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} \Longrightarrow \begin{pmatrix} 2 & 1 & -1 \\ -4 & 1 & 5 \\ 0 & 0 & 0 \end{pmatrix} \Longrightarrow \begin{pmatrix} 2 & 1 & -1 \\ 0 & 3 & 3 \\ 0 & 0 & 0 \end{pmatrix}$$

So there is only one eigenvector:

$$v^{(2)} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

so the number of jordan blockes corresponding to $\lambda = -4$ is 1. We can already conclude that the Jordan normal form of A is:

$$J = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -4 & 1 \\ 0 & 0 & -4 \end{pmatrix}$$

Now we find a generalized vector $v^{(3)}$ such that:

$$(A + 4I)\, v^{(3)} = v^{(2)}$$

$$\left(\begin{array}{ccc|c} -4 & 1 & 5 & 1 \\ 2 & 1 & -1 & -1 \\ -4 & 1 & 5 & 1 \end{array}\right) \Longrightarrow \left(\begin{array}{ccc|c} 2 & 1 & -1 & -1 \\ -4 & 1 & 5 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right) \Longrightarrow \left(\begin{array}{ccc|c} 2 & 1 & -1 & -1 \\ 0 & 3 & 3 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

There are no restrictions on z so we can choose $z = 0$. Then, $x = y = -\frac{1}{3}$. So

$$v^{(3)} = \begin{pmatrix} -1/3 \\ -1/3 \\ 0 \end{pmatrix}$$

$$P = \begin{pmatrix} 1 & 1 & -1/3 \\ 1 & -1 & -1/3 \\ 1 & 1 & 0 \end{pmatrix}$$

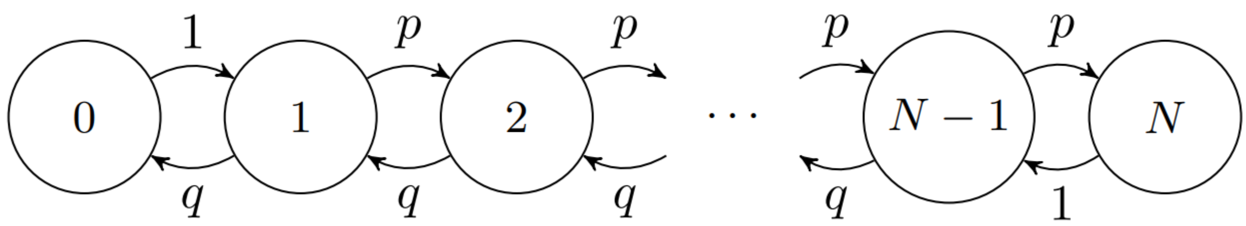
and

$$A = \begin{pmatrix} 1 & 1 & -1/3 \\ 1 & -1 & -1/3 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} -2 & 0 & 0 \\ 0 & -4 & 1 \\ 0 & 0 & -4 \end{pmatrix} \begin{pmatrix} 1 & 1 & -1/3 \\ 1 & -1 & -1/3 \\ 1 & 1 & 0 \end{pmatrix}^{-1}$$

2 A Standard Question on Markov Chains

1.

Figure 1: The Transition Diagram of the Markov Chain



2. The chain is irreducible since there is a path from any state i to any state j (if we let the system run enough steps, there is a possibility to get to state j). This is realized as the graph is strongly connected and its edges are strictly positive.

3.

Definition 1. Define the finite sequence $(a_k)_{k=0}^{N}$ by:

$$a_k = \frac{p^{\max(0,k-1)}}{q^{\min(N-1,k)}}$$

or, equivalently, $a = \left(1, \frac{1}{q}, \frac{p}{q^2}, \frac{p^2}{q^3}, \dots, \frac{p^{N-2}}{q^{N-1}}, \frac{p^{N-1}}{q^{N-1}}\right)$.

Proposition 3. *Let $N \geq 1$. The unique stationary distribution $\pi = (\pi_0, \dots, \pi_N)$ of the Markov chain is given by the formula:*

$$\pi_k = \frac{a_k}{\sum_{k=0}^{N} a_k},$$

Proof. We begin by writing the $(N+1) \times (N+1)$ transition matrix:

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ q & 0 & p & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & q & 0 & p & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & q & 0 & p & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & q & 0 & p & 0 \\ 0 & 0 & 0 & 0 & 0 & \dots & q & 0 & p \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & 1 & 0 \end{pmatrix}$$

The stationary distribution satisfies $\pi = \pi P$. Matrix multiplication leads us to the following set of equations numbered from 0 to N:

$$\pi_0 = q\pi_1$$
$$\pi_1 = \pi_0 + q\pi_2$$
$$\pi_2 = p\pi_1 + q\pi_3$$
$$\pi_3 = p\pi_2 + q\pi_4$$
$$\vdots = \vdots$$
$$\pi_{N-2} = p\pi_{N-3} + q\pi_{n-1}$$
$$\pi_{N-1} = p\pi_{N-2} + \pi_N$$
$$\pi_N = p\pi_{N-1}$$

Looking at the last two equations, we replace π_N by $p\pi_{N-1}$:

$$\pi_{N-1} = p\pi_{N-2} + p\pi_{N-1}$$
$$\pi_N = p\pi_{N-1}$$

subtracting $p\pi_{N-1}$ from both sides of equation number $N-1$, we get:

$$\pi_{N-1} - p\pi_{N-1} = p\pi_{N-2}$$
$$\pi_N = p\pi_{N-1}$$

remembering that $p+q=1$,we get:

$$q\pi_{N-1} = p\pi_{N-2}$$
$$\pi_N = p\pi_{N-1}$$

if we keep performing this procedure, we get:

$$
\begin{aligned}
\pi_0 &= q\pi_1 \\
q\pi_1 &= \pi_0 \\
q\pi_2 &= p\pi_1 \\
q\pi_3 &= p\pi_2 \\
\vdots &= \vdots \\
q\pi_{N-2} &= p\pi_{N-3} \\
q\pi_{N-1} &= p\pi_{N-2} \\
\pi_N &= p\pi_{N-1}
\end{aligned}
$$

Now we can express $\pi_1, \ldots, \pi_N$ with π_0:

$$
\begin{aligned}
\pi_1 &= \frac{1}{q}\pi_0 \\
\pi_2 &= \frac{p}{q^2}\pi_0 \\
\vdots &= \vdots \\
\pi_{N-1} &= \frac{p^{N-2}}{q^{N-1}}\pi_0 \\
\pi_N &= \frac{p^{N-1}}{q^{N-1}}\pi_0
\end{aligned}
$$

Remember that π is a distribution, so it must satisfy the normalizing condition:

$$\pi_0 + \ldots + \pi_N = 1$$

$$\pi_0 \left(1 + \frac{1}{q} + \ldots \frac{p^{N-2}}{q^{N-1}} + \frac{p^{N-1}}{q^{N-1}}\right) = 1$$

so it is evident that for all $k = 0, \ldots, N$,

$$\pi_k = \frac{a_k}{\sum_{k=0}^{N} a_k}$$

$\square$

3 The Gambler's Ruin

We first must obtain the probabilities for a finite version of this Markov chain before considering the limiting case. Let $N \geq 2$. Consider the chain defined by:

$$\begin{aligned} p_{00} &= p_{NN} = 1 \\ p_{i,i-1} &= q, \quad p_{i,i+1} = p, \text{ for } i = 1, \ldots, N-1 \end{aligned}$$

Here states 0 and N are absorbing. 0 resembles ruin and N resembles fortune. Let h_i be the probability that starting at state i, we achieve fortune. If we start at state $i = N$ then we already have achieved fortune and $h_N = 1$. If $i = 0$ then we are stuck in ruin and $h_N = 0$. Using exercise 6,

$$\begin{aligned} h_0 &= 0 \\ h_1 &= qh_0 + ph_2 \\ \vdots &= \vdots \\ h_{N-1} &= qh_{N-2} + ph_N \\ h_N &= 1 \end{aligned}$$

Remembering that $p + q = 1$, for all $i = 0, \ldots, N-1$,

$$qh_i + ph_i = qh_{i-1} + ph_{i+1}$$

$$\begin{aligned} \frac{q}{p}h_i + h_i &= \frac{q}{p}h_{i-1} + h_{i+1} \\ h_{i+1} - h_i &= \frac{q}{p}\left(h_i - h_{i-1}\right) \end{aligned}$$

In particular, $h_0 = 1$ implies that for $i = 1$:

$$h_2 - h_1 = \frac{q}{p}\left(h_1 - h_0\right) = \frac{q}{p}h_1$$

and generally, for all $i = 0, \ldots, N-1$,

$$h_{i+1} - h_i = h_1 \left(\frac{q}{p}\right)^i$$

Note that the sum $\sum_{k=1}^{i}(h_{k+1}-h_k)$ telescopes to $h_{i+1}-h_1$, so

$$h_{i+1}=h_{i+1}-h_1+h_1=h_1+\sum_{k=1}^{i}(h_{k+1}-h_k)=h_1\sum_{k=0}^{i}\left(\frac{q}{p}\right)^k$$

The sequence inside the sum is geometric only if $p\neq q$, so for all i (Since $i=0$ also fits the formula),

$$h_i=\begin{cases} h_1\frac{\left(\frac{q}{p}\right)^i-1}{\left(\frac{q}{p}\right)-1}, & p\neq q\\ h_1 i, & p=q\end{cases}$$

Since $h_N=1$, We must have that

$$h_1=\begin{cases} \frac{\left(\frac{q}{p}\right)-1}{\left(\frac{q}{p}\right)^N-1}, & p\neq q\\ \frac{1}{N}, & p=q\end{cases}$$

so finally

$$h_i=\begin{cases} \frac{\left(\frac{q}{p}\right)^i-1}{\left(\frac{q}{p}\right)^N-1}, & p\neq q\\ \frac{i}{N}, & p=q\end{cases}$$

Claim. For any starting position i, the probability of the game never ending (i.e never reaching state 0 or N) is 0.

Remark. Even though it is with probability 0, it is possible for the game to never end, for example if we alternate between two adjacent states forever ($N\geq 3$), $1,2,1,2,1,2,\ldots$

Proof. One way to show it is by using exercise 6 with the set $A=\{0,N\}$, this time getting the equations:

$$\begin{aligned} h_0&=1\\ h_1&=qh_0+ph_2\\ \vdots&=\vdots\\ h_{N-1}&=qh_{N-2}+ph_N\\ h_N&=1\end{aligned}$$

a similar derivation as before shows that $h_0=\ldots=h_N=1$, meaning that no matter where we start, we will always reach one of the barriers with a probability of 1. □

The claim implies that the probability of ruin is $1 - h_i$. To get the probability asked in the question, we consider the limiting case as $N \to \infty$. Calculating the limit, we get:

$$h_i = \begin{cases} \left(\frac{p}{q}\right)^i, & p > q \\ 1, & p \leq q \end{cases}$$

such that even if the game is fair, the probability of ruin is 1.