Numerical Methods in Soft Matter
Chapter 4: Markov chains

Isotton Gloria

June 17, 2024

Exercise 4.1

We have seen that the state probability vector of a Markov chain μ_n satisfies the recurrence relation $\mu_n = \mu_{n-1}\mathcal{P}$. Show that this equation is equivalent to write:

$\mu_n(i) = (1 - \Sigma_j p_{j,i})\mu_{n-1}(i) + \Sigma_j p_{i,j}\mu_{n-1}(j)$

Let us analyze each term in the equation above:

- $\mu_n(i)$, the system is in state i at time n;
- $(1 - \Sigma_j p_{j,i})$ is the probability of being in state i at previous time-step and not jumping in state j for each $j \neq i$;
- $\Sigma_j p_{j,i}$ represent the probability of being in state j at previous time-step and jumping in i.

This equation tells us that the distribution of the chain at time n is i and it is given by the probability of being in i at previous time and not jumping away plus the probability of being in j at previous time step and then jumping in i. The demonstration of that relation is the following:

$$\mu_n(i) = \sum_j \mu_{n-1}(j)p_{ij} = p_{ii}\mu_{n-1} + \sum_{j\neq i} p_{ij}\mu_{n-1}(j) \overset{*}{=} \left(1 - \sum_{j\neq i} p_{ij}\right)\mu_{n-1}(i) + \sum_j \mu_{n-1}(j)p_{ij}$$

where in $(*)$ we used the property of the column normalizaion of the Stochastic matrix:

$$\sum_j p_{ij} \overset{!}{=} 1 \Rightarrow \sum_j p_{ij} = \sum_{j\neq i} p_{ij} + p_{ii} \overset{!}{=} 1 \iff p_{ii} = 1 - \sum_{j\neq i} p_{ij}.$$

Exercise 4.2

Draw the digraphs and classify the states of the Markov chains defined by the following transition matrices.

$$\mathcal{P}_A = \begin{bmatrix} 0 & 0.5 & 0.5 \\ 0.5 & 0 & 0.5 \\ 0.5 & 0.5 & 0 \end{bmatrix} \quad \mathcal{P}_B = \begin{bmatrix} 0 & 0 & 0.5 & 0.5 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad \mathcal{P}_C = \begin{bmatrix} 0.3 & 0.4 & 0 & 0 & 0.3 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.6 & 0.4 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

Transition matrix A

Each state j is accessible from state i (meaning each state communicate with the other $i \leftrightarrow j$) so the Markov chain is irreducible. In particular because of the chain is defined on a finite space, it is called recurrent irreducible.

In this Markov chain, all states are aperiodic. For exemple let us consider state 1 and 2 path: $2 \to 1 \to 2$ of lenght 2, $2 \to 1 \to 3 \to 2$ of lenght 3. Since $gcd(2,3) = 1$ we conclude state 1 is aperiodic. For state 2 and state 3, similar analyses can be performed, and one can find that they are also aperiodic.

A finite irreducible and aperiodic Markov chain is called regular.

Transition matrix B

Each state j is accessible from state i, $i \leftrightarrow j$:

- State 1 has transitions to states 3 and 4 with probability 0.5;
- State 2 has a transition to state 1;
- State 3 and 4 have a transition to state 2 with probability 1;

Therefore, there is only one communicating class, so the Markov chain is irreducible. In particular because of the chain is defined on a finite space, it is called recurrent irreducible.
To study the periodicity of the chain let us consider state 2. We can find 2 different path to go from 2 to 2: $2 \to 1 \to 3 \to 2$, $2 \to 1 \to 4 \to 2$. The lenght of the two path is the same ($l = m = 3$) and the common greather divisor between 3 and 3 is not 1 so state 1 is periodic.
For state 1, 4 and state 3, similar analyses can be performed, and one can find that they are also periodic.

Transition matrix C
In this diagraph not all the states communicates with each other and we can identify 3 classes of equivalence:

- class 1 = state 5, state 3, state 4;
- class 2 = state 2;
- class 3 = state 1.

In that exemple not each state j is accessible from state i so the Markov chain is reducible. Let us study the periodicity of the classes:

- class 1: let us consider the state 5, one can build 2 path $5 \to 3 \to 5$ or $5 \to 3 \to 4 \to 5$. The lenghts of the path are 3 and 2, that are co prime, meaning that the state 5 and (similarly) its class is aperiodic;
- class 2 and 3 are aperiodic: they have a self-transition and the number 1 is co-prime to every integer.

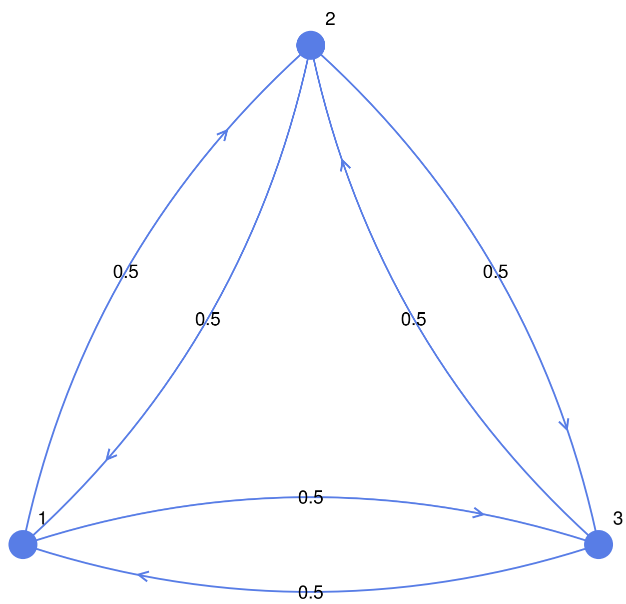


Figure 1: A

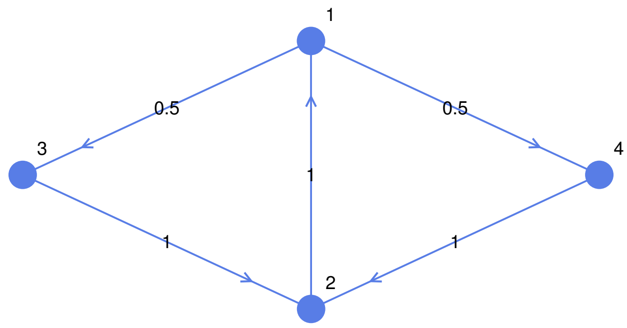


Figure 2: B

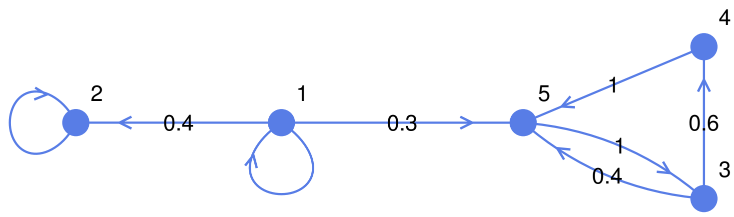


Figure 3: C

Exercise 4.3

Given the two Markov chains defined by the following transition matrices:

$$\mathcal{P}_1 = \begin{bmatrix} 0.5 & 0.5 \\ 1 & 0 \end{bmatrix} \qquad \mathcal{P}_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0.25 & 0 & 0.75 \end{bmatrix}$$

verify whether they are irreducible and in case compute their period. Moreover, for both of them compute $\mathcal{P}_n$ and the limit at infinite (or the limit of its average).

Transition matrix 1
The Markov chain described by $\mathcal{P}_1$ [4] is irreducible, because state 1 can communicates with state 2 and state 2 can communicates with state 1. In that case, there is a self loop that indicates that the chain is aperiodic. The Markov chain is aperiodic and irridicibile (i.e. regular) therefore from the theorem we know that for any initial probability distribution μ_0,

$$\mu_n = \mu_0 \mathcal{P}_n \text{ as } n \to \infty.$$

If we try to compute the limit at infinite of $\mathcal{P}_n$ we obtain: $\underset{n\to\infty}{lim} \mathcal{P}_1^n \to \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$

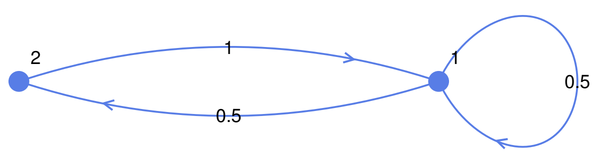


Figure 4: Transition matrix 1

Transition matrix 2
The Markov chain described by $\mathcal{P}_2$ [5] is not irreducible, because state 2 does not communicate with state 1 or 3. In this chain we can identify 2 communicating classes (meaning a set of nodes such that only members of the same class communicate with each other):

- class 1: state 1 and 3;
- class 2: state 2.

In the chain there are also self loops that indicate that the chain is aperiodic. Therefore this Markov chain is aperiodic and not irriducibile.

If we try to compute the limit at infinite of $\mathcal{P}_n$ we obtain: $\underset{n\to\infty}{lim}\mathcal{P}_2^n \to \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

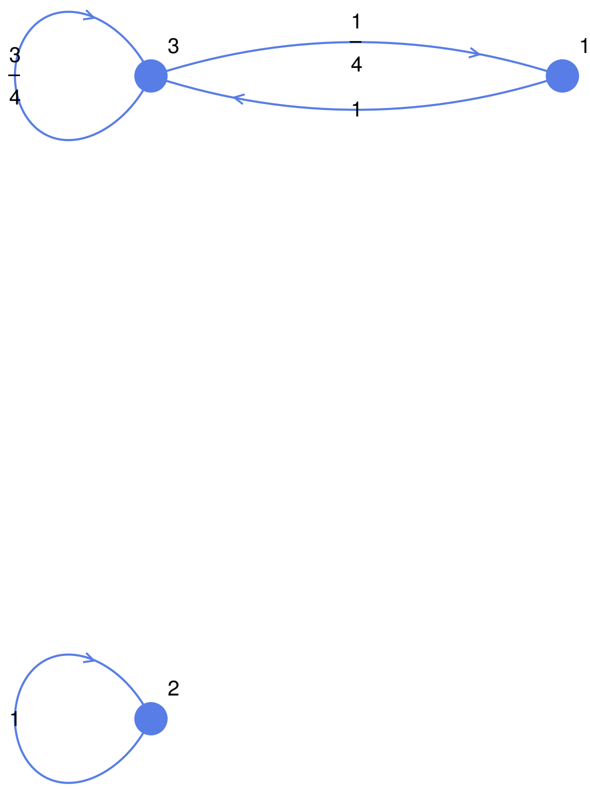
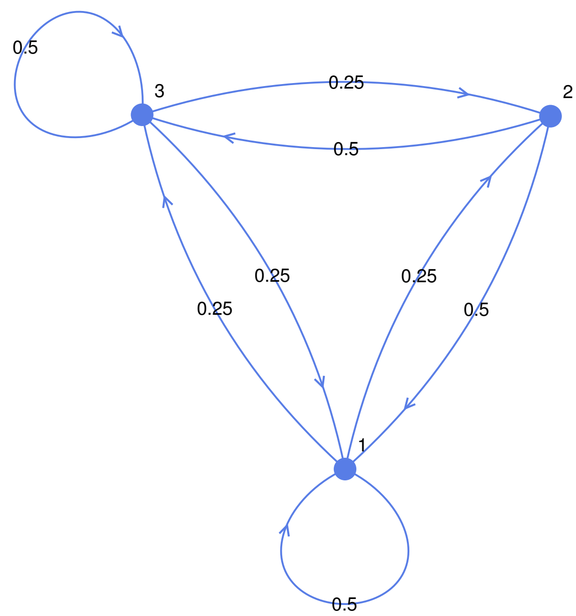


Figure 5: Transition matrix 2

Exercise 4.4

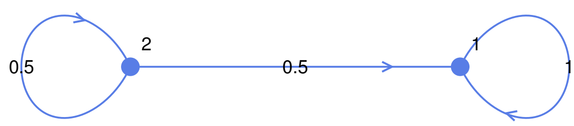
Given the Markov chain defined by $\begin{bmatrix} 0.5 & 0.25 & 0.25 \\ 0.5 & 0 & 0.5 \\ 0.25 & 0.25 & 0.5 \end{bmatrix}$ show that is regular.

Looking at the chart, it's easy to understand that every state communicates with every other state. Since we can find a path of non-zero probability between any pair of states, the Markov chain is considered irreducible. At this point, according to the theorem, it is sufficient to find one aperiodic state to be able to say that the Markov chain is aperiodic.
To find one aperiodic state we can exploit the self-transitions. We know that a state is aperiodic if the greatest common divisor (gcd) of all the return times to the state is 1. Let us consider state 1. In this case, since the chain can return to state 1 in just one step, the gdc of the return times is always 1, making the state and the chain aperiodic.
Since we have shown that the Markov chain is aperiodic and irriducibile we can conclude that it is regular.



Exercise 4.4.B

Is the Markov chain defined by $\begin{bmatrix} 1 & 0 \\ 0.5 & 0.5 \end{bmatrix}$ regular?

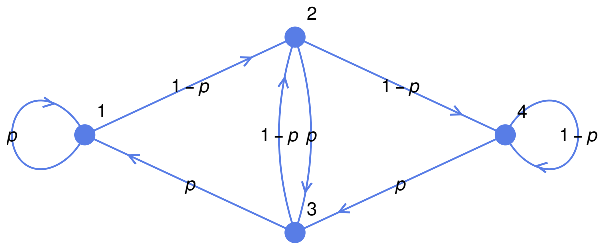


Since there is no direct transition from state 1 to state 2, state 1 and state 2 are not communicating directly. This means that the chain is reducible. It is not possible to reach state 2 from state 1 with a positive probability.
In addition, as we know a Markov chain can be considered aperiodic if all states have a period of 1. In this case, both states have a self-transition, which means that the period of both states is 1. However, because the chain is not irreducible (as stated above), it is not appropriate to discuss the aperiodicity of the entire chain.

Exercise 4.5.A

Let $0 < p < 1$ and let us consider a Markov chain defined on the finite space $S = 1, 2, 3, 4$ by the transition matrix: $\begin{bmatrix} p & 1-p & 0 & 0 \\ 0 & 0 & p & 1-p \\ p & 1-p & 0 & 0 \\ 0 & 0 & p & 1-p \end{bmatrix}$.

- Show that the Markov chain is recurrent irreducible;
- Show that the Markov chain is aperiodic (consider for instance the term p_{11});
- Compute the fixed point π that, given the first two points, is the invariant unique distribution of the Markov chain.



Looking at the diagraph one can see that it is possible to reach any state from any other state with positive probability. For instance, starting from state 1, one can reach state 2 in a single step and states 3 or 4 in two steps. Symmetrically, from states 2, 3, and 4, it is possible to reach state 1. Similarly it's trivial to notice that states 2 and 3 communicates with each other and so on for all the 4 states in the chain. Therefore, the Markov chain is **irreducible**.

To demonstrate recurrence, we need to show that every state in the Markov chain is recurrent.

Recurrence of a state implies that, given the starting point at that state, the process will return to that state with a probability of 1. It's important to observe that when the state space S where the Markov chain is defined is finite, the following result holds: all irreducible Markov chains defined on a finite space S are recurrent irreducible.
In our case $|S| = 4$ and so the Markov chain is **recurrent irreducible**.

To show that the Markov chain is aperiodic we need to demonstrate that at least one the states is aperiodic. We know that a state in a Markov chain is said to be aperiodic if the greatest common divisor (gcd) of the lengths of its cycles is 1. In simpler terms, a chain is aperiodic if there is no regular pattern or fixed interval in the occurrence of transitions between states.
Let us consider state 1. In this case, the chain can return to state 1 in just one step (cycle 1: $1 \to 1$) or in 3 steps (cycle 2: $1 \to 2 \to 3 \to 1$). Tthe gdc of 1 and 3 is 1, making the state and the **chain aperiodic**. This is a general rule: since the number 1 is co-prime to every integer, any state with a self-transition is aperiodic.

To calculate π, we compute the left eigenvector associated with the eigenvalue $\lambda = 1$, because $\pi A = 1 \cdot \pi$. Let us verify that the matrix A has one eigenvalue equal to one. We begin by subtracting λ times the identity matrix from the stochastic matrix:

$$A - \lambda I = \begin{bmatrix} p-\lambda & 1-p & 0 & 0 \\ 0 & -\lambda & p & 1-p \\ p & 1-p & -\lambda & 0 \\ 0 & 0 & p & 1-p-\lambda \end{bmatrix}$$

Now we can calculate the determinant of this matrix by expanding with respect to one of the rows or columns. In this case, we expand with respect to the first column:

$$\det(A-\lambda I) = (p-\lambda)\cdot \det \begin{bmatrix} -\lambda & p & 1-p \\ 1-p & -\lambda & 0 \\ 0 & p & 1-p-\lambda \end{bmatrix} + p \cdot \det \begin{bmatrix} 1-p & 0 & 0 \\ -\lambda & p & 1-p \\ 0 & p & 1-p-\lambda \end{bmatrix} =$$

$$(p-\lambda)\cdot[-\lambda \cdot \det \begin{bmatrix} 1-p & 0 \\ p & 1-p-\lambda \end{bmatrix} - (1-p)\cdot \det \begin{bmatrix} p & 1-p \\ p & 1-p-\lambda \end{bmatrix}] + p(1-p)\cdot \det \begin{bmatrix} p & 1-p \\ p & 1-p-\lambda \end{bmatrix} =$$

$$(p-\lambda)\cdot(p\lambda - p^2\lambda + \lambda^2 - p\lambda^2 - \lambda^3) + p\cdot(-p\lambda + p^2\lambda) =$$

$$(p-\lambda)\left(-\lambda^3 + \lambda^2 - \lambda p^2 - \lambda^2 p + \lambda p\right) + p\left(\lambda p^2 - \lambda p\right) =$$

$$\lambda^4 - \lambda^3.$$

The eigenvalues are the values of λ for which holds that $\det(A - \lambda I) = 0$:

$$\lambda \to 0,$$
$$\lambda \to 0,$$
$$\lambda \to 0,$$
$$\lambda \to 1$$

We have verified that the matrix A has one eigenvalue equal to 1. The left eigenvector corresponding to the eigenvalue 1 of the matrix A is found by solving the system $v^T A = v^T$, where v^T is the transpose of the left eigenvector. One can find the left eigenvector by solving the system:

$$v^T \begin{bmatrix} p & 1-p & 0 & 0 \\ 0 & 0 & p & 1-p \\ p & 1-p & 0 & 0 \\ 0 & 0 & p & 1-p \end{bmatrix} = 1 \cdot v^T$$

.. or equivalently:

$$v^T \begin{bmatrix} p-1 & 1-p & 0 & 0 \\ 0 & -1 & p & 1-p \\ p & 1-p & -1 & 0 \\ 0 & 0 & p & -p \end{bmatrix} = \mathcal{O}^T$$

$$\begin{aligned} (p-1)v_1 + pv_3 &= 0 \\ (1-p)v_1 - v_2 + (1-p)v_3 &= 0 \\ pv_2 - v_3 + pv_4 &= 0 \\ (1-p)v_2 - pv_4 &= 0 \end{aligned}$$

Let us simplify the first and the last equations, expressing v_1, v_2 in function of v_3:

$$\begin{aligned} v_1 &= -v_3 p/(p-1) \\ (1-p)v_1 - v_4 + (1-p)v_3 + pv_4 &= v_4 \\ pv_2 - v_3 + pv_4 &= 0 \\ v_4 &= (1-p)/pv_3 \end{aligned}$$

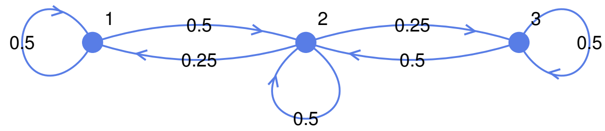
Combining all together the second equation becomes:

$$\begin{aligned} v_1 &= v_3 p/(1-p) \\ p(p-1)(1-p)v_3 - p/(1-p)v_4 + (1-p)v_3 &= 0 \to v_3 = p/(1-p)v_4 \\ pv_2 - v_3 + pv_4 &= 0 \\ v_4 &= (1-p)/pv_3 \end{aligned}$$

while the third equation is just $0 \cdot v_3 = 0$, which is true for every value of v_3. In other words, v_3 represent the free parameter that define the eigenvector in the associated eigenvector space. The left eigenvectors of eigenvalues $\lambda = 1$ are all the vectors belonging to the subspace generated by: $\left(\frac{p}{1-p}a, a, a, \frac{a(1-p)}{p}\right)$. Now, remembering that since we can write $\pi A = 1 \cdot \pi$, the pi is simply the left eigenvector of eigenvalue 1, that we computed before:

$$\pi = \mathcal{N} a \cdot \begin{bmatrix} \frac{p}{1-p} \\ 1 \\ 1 \\ \frac{1-p}{p} \end{bmatrix},$$

where the normalization constant is $\mathcal{N} = \sqrt{\left|a\left(\frac{1}{p}-1\right)\right|^2 + \left|\frac{ap}{1-p}\right|^2 + 2|a|^2}$.



Exercise 4.5.B

Find the stationary distribution π of the Markov chain with transition matrix $\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$.

Let us compute the eigenvalues. The characteristic equation of the matrix is given by

$$\det(A-\lambda I) = \det\left(\begin{bmatrix} \frac{1}{2}-\lambda & \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{2}-\lambda & \frac{1}{4} \\ 0 & \frac{1}{2} & \frac{1}{2}-\lambda \end{bmatrix}\right)$$

Simplifying,

$$= (\frac{1}{2}-\lambda)\left(\det\begin{bmatrix} \frac{1}{2}-\lambda & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2}-\lambda \end{bmatrix}\right) - \frac{1}{2}\left(\det\begin{bmatrix} \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{2}-\lambda \end{bmatrix}\right) = (\frac{1}{2}-\lambda)[(\frac{1}{2}-\lambda)(\frac{1}{2}-\lambda)-(\frac{1}{2})(\frac{1}{4})] - \frac{1}{2}[\frac{1}{4}(\frac{1}{2}-\lambda)]$$

Further simplification leads to

$$= (\frac{1}{2}-\lambda)\left[\frac{1}{4}-\lambda+\lambda^2-\frac{1}{8}\right] - \frac{1}{16}+\frac{1}{8}\lambda = \frac{1}{2}\left(-2\lambda^3+3\lambda^2-\lambda\right)$$

The eigenvalues are:

$$\frac{1}{2}(-\lambda+3\lambda^2-2\lambda^3) = 0 \longrightarrow \left\{\{\lambda \to 0\}, \{\lambda \to \frac{1}{2}\}, \{\lambda \to 1\}\right\}$$

We have verified that the matrix A has one eigenvalue equal to 1. The stationary distribution π can be found by computing the left eigenvector corresponding to the eigenvalue 1. The left eigenvector corresponding to the eigenvalue 1 of the matrix A is found by solving the system $v^T A = v^T$, where v^T is the transpose of the left eigenvector. One can find the left eigenvector by solving the system:

$$v^T \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} = 1 \cdot v^T$$

$$\begin{aligned} \frac{1}{2}v_1 + \frac{1}{4}v_2 &= v_1 \\ \frac{1}{2}v_1 + \frac{1}{2}v_2 + \frac{1}{2}v_3 &= v_2 \\ \frac{1}{4}v_2 + \frac{1}{2}v_3 &= v_3 \end{aligned}$$

$$\begin{aligned} v_2 &= 2v_1 \\ \frac{1}{2}v_1 + v_1 + \frac{1}{2}v_3 &= 2v_1 \to v_1 = v_3 \\ \frac{1}{2}v_1 + \frac{1}{2}v_3 &= v_3 \to v_1 = v_3 \end{aligned}$$

$$\pi = \begin{bmatrix} \frac{1}{4} \\ \frac{1}{2} \\ \frac{1}{4} \end{bmatrix}.$$