Embedded matrices for finite Markov chains

Citation for published version (APA):

Document status and date:
Published: 01/01/1983

Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.
Memorandum COSOR 83-19

Embedded matrices for finite Markov chains

by

D.A. Overdijk

Eindhoven, the Netherlands
December 1983
EMBEDDED MATRICES FOR FINITE MARKOV CHAINS

D.A. Overdijk

Abstract. For an arbitrary subset A of the finite state space S of a Markov chain the so-called embedded matrix $P_A$ is introduced. By use of these matrices formulas expressing recurrence probabilities can be written down almost automatically and derivations can be given very systematically.

Keywords: finite Markov chain, embedded matrix.

0. Introduction and summary

We consider a Markov chain $X_0, X_1, \ldots$ on the state space $S = \{1, 2, \ldots, s\}$. The corresponding matrix of transition probabilities is denoted by $P$. We do not exclude the case where $P$ is sub-Markov i.e. the case where the elements of $P$ are nonnegative with row sums less than one. For the Markov chain this means that for some time $n \geq 1$ we may have $X_k \not\in S$ for all $k \geq n$.

For an arbitrary subset $A \subseteq S$ we introduce the so-called embedded matrix $P_A$.

By use of these embedded matrices calculations can be performed very systematically, and the derivation and interpretation of results become quite transparent. Most of our results are not new and can be found in e.g. the standard reference for calculations in finite Markov chains Kemeny and Snell (1976).

The novelty lies in the ease with which the results are obtained. The idea of the embedded process can be found in e.g. Foguel (1969), Revuz (1975), Simons and Overdijk (1979).
In Section 1 some basic matrices are defined. In Section 2 we introduce the embedded matrix $P_A$, and explain how calculations can be performed using these matrices. In Section 3 the well-known partition of the state space in transient and nontransient states is derived by means of embedded matrices. As an example, in Section 4 we give detailed calculations for the random walk along the edges of the cube with roof (see Figure 1).

![Figure 1](image)

The vertices 8 and 9 are absorbing and in all other vertices edges are chosen with equal probability.

The following quantities are calculated.

i) The probability that absorption takes place in a given absorbing state (cf. problem 90 in Statistica Neerlandica).

ii) The mean and variance of the first entrance time in one of the absorbing states (cf. problem 59 in Statistica Neerlandica).

iii) The mean and variance of the number of different states of a given transient set visited by the random walk (cf. problem 54 in Statistica Neerlandica).
Except for the variance in iii) these quantities are also calculated in Kemeny and Snell (1976). In Section 4 we give general formulas for such quantities; the calculations for the random walk on the "cube" in Figure 1 are performed with the aid of a computer.

1. **Basic matrices**

For every \( j \in S \) the \( s \times 1 \)-matrix (column vector), whose \( j \)-th component is one and all other components zero, is denoted by \( e_j \) and we define the \( s \times s \)-matrix

\[
I_j := e_j e_j' \quad (e_j' \text{ is the transpose of } e_j).
\]

For every subset \( A \subset S \) we define

\[
e_A := \sum_{j \in A} e_j, \\
I_A := \sum_{j \in A} I_j, \\
I := I_S \quad (s \times s \text{ identity matrix}), \\
e := e_S \quad (\text{column vector with all components one}), \\
o := e_\emptyset \quad (\text{zero column vector}).
\]

The following relations can easily be verified.

1. The matrix element \( p_{jk} = e_j P e_k \), \( 1 \leq j \leq s \) and \( 1 \leq k \leq s \)

2. \( I_j e = e_j \), \( 1 \leq j \leq s \)

3. \( I_A e = e_A \), \( A \subset S \)

4. \( I_A I_B = I_{A \cap B} \), \( A \subset S \) and \( B \subset S \).
All equalities (inequalities) between matrices have to be interpreted componentwise and convergence of a sequence of matrices is convergence componentwise.

For every subset $A \subset S$ we write $A^* := S \setminus A$.

**Example 1.** Put $S = \{1,2,3\}$ and $A = \{2,3\}$ then we have e.g.

\[
\begin{align*}
    e_2 &= \begin{pmatrix}
        0 \\
        1 \\
        0
    \end{pmatrix},
    I_2 = \begin{pmatrix}
        0 & 0 & 0 \\
        0 & 1 & 0 \\
        0 & 0 & 0
    \end{pmatrix},
    e_A = \begin{pmatrix}
        0 \\
        1 \\
        1
    \end{pmatrix},
    I_A = \begin{pmatrix}
        0 & 0 & 0 \\
        0 & 1 & 0 \\
        0 & 0 & 1
    \end{pmatrix}.
\end{align*}
\]

It is well-known that the probability distribution of a Markov chain is determined by its transition matrix $P$ and an initial distribution $\pi$ of $X_0$ on $S$.

The corresponding probability distribution of the Markov chain is denoted by $P_{\pi}$. An initial distribution on $S$ is denoted by a column vector $\pi$ with non-negative components summing to one, i.e.

\[
\pi' e_j \geq 0 \quad \text{for all } 1 \leq j \leq s
\]

\[
\pi' e = 1.
\]

If $\pi = e_j$, we write $P_j := P_{e_j}$.

Expectations with respect to $P_{\pi}$ or $P_j$ are denoted by the symbol $E_{\pi}$ or $E_j$.

**2. Embedded matrices**

We start with a proposition that is the key to calculating probabilities with respect to $P_{\pi}$. 
Proposition 1. For every integer $n \geq 0$, every initial distribution $\pi$ on $S$ and every $(n+1)$-tuple subsets $A_0, A_1, \ldots, A_n \subseteq S$

$$\mathbb{P}_\pi (X_0 \in A_0, X_1 \in A_1, \ldots, X_n \in A_n) = \pi A_0 \pi A_1 \ldots \pi A_n \epsilon .$$

Proof. We proceed by induction. Let $\pi$ be an initial distribution on $S$ and let $A_0 \subseteq S$. Since $\mathbb{P}_\pi (X_0 \in A_0) = \pi e A_0 \pi A_0$, (use (3)), the proposition is true for $n = 0$.

Suppose the proposition has been proved for $0 \leq k \leq n$. Then we have

$$\mathbb{P}_\pi (X_0 \in A_0, \ldots, X_{n+1} \in A_{n+1})$$

$$= \sum_{j \in A_n} \mathbb{P}_\pi (X_0 \in A_0, \ldots, X_{n-1} \in A_{n-1}, X_n = j) \mathbb{P}_j (X_1 \in A_{n+1})$$

$$= \sum_{j \in A_n} \sum_{k \in A_{n+1}} \mathbb{P}_\pi (X_0 \in A_0, \ldots, X_n = j) \mathbb{P}_j (X_1 = k)$$

$$= \sum_{j \in A_n} \sum_{k \in A_{n+1}} \pi' A_0 \pi A_1 \ldots \pi A_{n-1} \pi_j e P_{jk}$$

(use 1))

$$= \sum_{j \in A_n} \sum_{k \in A_{n+1}} \pi' A_0 \pi A_1 \ldots \pi_j e e_j P_{jk}$$

(use 2))

$$= \sum_{j \in A_n} \sum_{k \in A_{n+1}} \pi' A_0 \pi A_1 \ldots \pi_j e e_j P_{jk}$$

(use 0) and (2))

$$= \sum_{j \in A_n} \sum_{k \in A_{n+1}} \pi' A_0 \pi A_1 \ldots \pi_j \pi_k e = \pi' A_0 \pi A_1 \ldots \pi A_n \pi A_{n+1} \epsilon . \quad \square$$

Remark. If $\pi = e_j$ then

$$\mathbb{P}_j (X_1 \in A_1, \ldots, X_n \in A_n) = e_j A_1 \pi A_2 \ldots \pi A_n = e_j A_1 \ldots \pi A_n \epsilon .$$

For the definition of embedded matrices we need the following lemma.
Lemma 1. For every subset $A \subseteq S$ the sequence of matrices

$$\left\{ \sum_{n=0}^{N} (P_{I} A)_{n}^{n} P_{I} A \right\}_{N \geq 0}$$

is convergent.

Proof. For every integer $N \geq 0$ define the matrix $Q_{N} := \sum_{n=0}^{N} (P_{I} A)_{n}^{n} P_{I} A$.

Using Proposition 1 with $\pi = e_{j}$ we get for the entries $Q_{N}(j,k)$ ($1 \leq j \leq s$ and $1 \leq k \leq s$) of the matrix $Q_{N}$

$$Q_{N}(j,k) = e_{j}^{T} Q_{N} e_{k} = \sum_{n=0}^{N} e_{j}^{T} (P_{I} A)_{n}^{n} P_{I} A \cap \{ k \} e_{k}$$

$$= \sum_{n=1}^{N+1} e_{j}^{T} (P_{I} A)_{n-1}^{n-1} P_{I} A \cap \{ k \} e_{k}$$

$$= \mathbb{P}_{j} (\exists_{1 \leq n \leq N+1} : X_{1} \in A^{*} , \ldots , X_{n-1} \in A^{*} , X_{n} \in A \cap \{ k \}) ,$$

which is nondecreasing and bounded by one.

Hence

$$\lim_{N \to \infty} Q_{N}(j,k) = \mathbb{P}_{j} (\exists_{n \geq 1} : X_{1} \in A^{*} , \ldots , X_{n-1} \in A^{*} , X_{n} \in A \cap \{ k \}) . \quad \Box$$

We are now ready to introduce the class of embedded matrices.

Definition 1. For every subset $A \subseteq S$ the embedded matrix $P_{A}$ is defined by

$$P_{A} := \sum_{n=0}^{\infty} (P_{I} A)_{n}^{n} P_{I} A .$$

If $A = \{ j \}$ with $j \in S$ we write $P_{j} := P_{\{ j \}}$. 
The probabilistic interpretation of the entries of the matrix $P_A$ follows from the proof of Lemma 1:

$$P_A(j,k) = \mathbb{P}_j(\exists_{n \geq 1} : X_1 \in A^*, \ldots, X_{n-1} \in A^*, X_n \in A \cap \{k\}) .$$

For the actual calculation of embedded matrices the following two lemmas are useful.

**Lemma 2.** For every subset $A \subseteq S$

$$(I - P_I \bullet_{A^*})P_A = P_I A .$$

**Proof.**

$$P_A = \sum_{n=0}^{\infty} (P_I \bullet_{A^*})^n P_I A = P_I A + \sum_{n=1}^{\infty} (P_I \bullet_{A^*})^n P_I A =$$

$$P_I A + P_I \bullet_{A^*} \sum_{n=0}^{\infty} (P_I \bullet_{A^*})^n P_I A = P_I A + P_I \bullet_{A^*} P_A .$$

If for some subset $A \subseteq S$ the matrix $P_I \bullet_{A^*}$ has no eigenvalue one, i.e. the matrix $I - P_I \bullet_{A^*}$ is regular, then the embedded matrix $P_A$ can easily be calculated. We then have

$$P_A = (I - P_I \bullet_{A^*})^{-1} P_I A .$$

The following lemma supplies a probabilistic characterization of the presence of an eigenvalue one of the matrix $P_I A$.

**Lemma 3.** Let $A \subseteq S$. The matrix $P_I A$ has an eigenvalue one iff there exists a state $j \in S$ such that

$$\mathbb{P}_j(X_n \in A \text{ for all } n \geq 1) > 0 .$$
Proof. First suppose the matrix $P_A$ has an eigenvalue one. Then there exists a column vector $v \neq 0$ such that $P_A v = v$.

Consider the vectors $v^+$ and $v^-$ with components $v^+(i) = \max(v(i), 0)$ and $v^-(i) = \max(-v(i), 0)$. Then $v = v^+ - v^- = P_A v^+ - P_A v^-$. Since the vectors $P_A v^+$ and $P_A v^-$ are nonnegative $P_A v^+ \geq v^+$ and $P_A v^- \geq v^-$. (Use the fact that $v = v^+ - v^- = b - c$, where $b$ and $c$ are nonnegative vectors, implies $b \geq v^+$ and $c \geq v^-).$ At least one of the vectors $v^+$ and $v^-$ is nonzero and therefore a vector $w$ exists such that $w \neq 0$, $0 \leq w \leq e$ and $P_A w \geq w$. Let j-th component of $w$ be positive then from Proposition 1 with $\pi = e_j$ it follows that

$$\mathbb{P}_j (X_n \in A \text{ for all } n \geq 1) = \lim_{n \to \infty} e_j (P_A^n e) w > 0.$$ 

Conversely, suppose a state $j \in S$ exists such that

$$\mathbb{P}_j (X_n \in A \text{ for all } n \geq 1) > 0.$$ 

Since $\mathbb{P}_j (X_n \in A \text{ for all } n \geq 1) = e_j \lim_{n \to \infty} (P_A^n e) > 0$ the vector $v := \lim_{n \to \infty} (P_A^n e) \neq 0$ and $P_A v = v$. \(\square\)

Recurrence probabilities with respect to a subset $A \subset S$ can be calculated using embedded matrices. We give some useful examples.

**Proposition 2.** For every initial distribution $\pi$ on $S$, all subsets $A, B \subset S$ and all integers $n \geq 1$

\begin{align*}
(6) & \quad \mathbb{P}_\pi (\exists n \geq 1 : X_1 \in A^*, \ldots, X_{n-1} \in A^*, X_n \in A \cap B) = \pi^t P_A e_B.
(7) & \quad \mathbb{P}_\pi (X_k \in A \text{ for at least } n \text{ different } k \geq 1) = \pi^t P_A^n e.
\end{align*}
\[ \text{(8)} \quad \mathbb{P}_{\pi}(X_k \in A \text{ for exactly } n \text{ different } k \geq 1) = \pi'(p^ne - p^{n+1}e). \]

\[ \text{(9)} \quad \mathbb{P}_{\pi}(X_k \in A \text{ for some } k > n) = \pi' P^n e. \]

\[ \text{(10)} \quad \mathbb{P}_{\pi}(X_k \in A \text{ infinitely often}) = \pi' \lim_{n \to \infty} P^n e. \]

**Proof.** We only prove (6); the other assertions can be proved similarly.

\[ \mathbb{P}_{\pi}(\exists n \geq 1 : X_1 \in A^*, \ldots, X_{n-1} \in A^*, X_n \in A \cap B) \]

\[ = \sum_{k=0}^{\infty} \mathbb{P}_{\pi}(X_1 \in A^*, \ldots, X_k \in A^*, X_{k+1} \in A \cap B) \quad \text{(Proposition 1)} \]

\[ = \sum_{k=0}^{\infty} \pi'(p_I)_A^k p_I A \cap B e \]

\[ = \pi' P_A e_B. \]

We conclude this section with a lemma to be used later.

**Lemma 4.** For every subset \( A \subset S \) the sequence of columnvectors \( \{p^n_A e\}_{n \geq 0} \) is nonincreasing.

**Proof.** For the \( j \)-th component \( p^n_A e(j) \) of the vector \( p^n_A e \) we have (use (7))

\[ p^n_A e(j) = e_j' p^n_A e = \mathbb{P}_j (X_k \in A \text{ for at least } n \text{ different } k \geq 1). \]

3. **Transient sets**

We start with the definition of a transient state. Our definition yields the usual partition of the state space \( S \) in transient and nontransient states (see e.g. Kemeny and Snell (1976)).
Definition 2. A state \( j \in S \) is called transient if the probability that the chain, when started in \( j \), will ever return to \( j \) is less than one i.e. if (see (6) with \( B = S \) and \( A = \{j\} \))

\[
e_j^\prime p_j e < 1.
\]

Furthermore a subset \( A \subset S \) consisting of transient states only is called transient and the set of all transient states of \( S \) is called its transient part and denoted by \( T \).

Lemma 5. For every transient state \( j \in S \) there exists a number \( 0 \leq q < 1 \) such that for all \( n \geq 1 \)

\[
p_j^n e \leq q^{n-1} e.
\]

Proof. Suppose \( j \in S \) is transient then by definition \( q := e_j^\prime p_j e < 1 \).

Hence \( I_j p_j e \leq q e \).

We proceed by induction. For \( n = 1 \) the proposition is trivial. Suppose the proposition has been proved for \( 0 \leq k \leq n \).

We have

\[
p_j^{n+1} e = p_j^n p_j e \quad \text{(use } p_j^n = p_j^n I_j \text{ for } n \geq 1)\]

\[
= p_j^n I_j p_j e \quad \text{(use } I_j p_j e \leq q e)\]

\[
\leq p_j^n q e = q p_j^n e \leq q^n e.
\]

Proposition 3. For every nontransient state \( j \in S \)

\[
e_j^\prime p_j^n e = 1 \text{ for all } n \geq 1.
\]

Hence \( p_j (X_n = j \text{ infinitely often}) = 1 \) (see (10)).
Proof. We proceed by induction. Let \( j \in S \) be nontransient then the proposition is true for \( n = 1 \) (use Definition 2).

Suppose the proposition has been proved for \( 0 \leq k \leq n \). We have

\[
e_j^1 P^{n+1} e = e_j^1 P^n P_j e = e_j^1 P^n I_j \quad (\text{use } P^n_j = P^n_j I_j \text{ for } n \geq 1)
\]

\[
e_j^1 P^n I_j P_j e = e_j^1 P^n e_j = e_j^1 P^n e = 1.
\]

Proposition 4. A subset \( A \subset S \) is transient iff the probability that the chain is in the set \( A \) infinitely often is zero for every starting position \( j \in S \) i.e. iff (cf. (10))

\[
e_j^1 \lim_{n \to \infty} P^n_A e = 0 \text{ for all } j \in S.
\]

Proof. First suppose \( A \subset S \) is transient. If the chain is infinitely often in the set \( A \) then, since the set \( A \) is finite, there exists \( k \in A \) such that \( X_n = k \) infinitely often.

Hence

\[
\mathbb{P}_j (X_n \in A \text{ infinitely often}) \leq \sum_{k \in A} \mathbb{P}_j (X_n = k \text{ infinitely often}).
\]

From (10) we obtain

\[
e_j^1 \lim_{n \to \infty} P^n_A e = \sum_{k \in A} e_j^1 \lim_{n \to \infty} P^n_A e.
\]

Since every \( k \in A \) is transient we conclude from Lemma 5 that \( \lim_{n \to \infty} e_j^1 P^n_A e = 0 \).

Conversely suppose \( A \subset S \) is not transient then there exists \( j \in A \) such that \( e_j^1 P^n_j e = 1 \) for all \( n \geq 1 \) (Proposition 3). Since \( \mathbb{P}_j (X_n = j \text{ infinitely often}) \leq \mathbb{P}_j (X_n \in A \text{ infinitely often}) \) we have \( \lim_{n \to \infty} e_j^1 P^n_A e \geq \lim_{n \to \infty} e_j^1 P^n_j e = 1 \). Hence \( \lim_{n \to \infty} e_j^1 P^n_A e \neq 0 \).
If the subset \( A \subset S \) is transient Proposition 4 states that \( \lim_{n \to \infty} P^n_A e = 0 \).

The following lemma expresses the fact that this convergence is exponential.

**Lemma 6.** If the subset \( A \subset S \) is transient then there exist an integer \( n_0 \) and a number \( 0 \leq r < 1 \) such that

\[
P^n_A e \leq r^n e \text{ for all } n \geq n_0.
\]

**Proof.** From Proposition 4 and Lemma 4 we know: \( P^n_A e \to 0 \). Hence there exist an integer \( n_0 \) and a number \( 0 \leq w < 1 \) such that

\[
P^n_A e \leq w^n e \text{ for all } n \geq n_0.
\]

Suppose \( n \geq 2n_0 \) then \( n \geq n_0 \left\lfloor \frac{n}{n_0} \right\rfloor \) (\( \lfloor a \rfloor \): largest integer not exceeding \( a \)). Hence

\[
P^n_A e \leq \frac{n}{n_0} P^n_A e \leq w^n e \leq w^{2n_0} e.
\]

Put \( r := \frac{w}{2n_0} \) then \( 0 \leq r < 1 \) and \( P^n_A e \leq r^n e \) for all \( n \geq 2n_0 \).

If \( A \subset S \) is transient then we conclude from Lemma 6 that

\[
e_j^T P^n_A e_k \leq e_j^T P^n_A e \leq r^n \text{ for } n \text{ sufficiently large.}
\]

(11) Hence the sequence of matrices \( \{P^n_A\}_{n \geq 0} \) converges exponentially to the zero matrix if the set \( A \) is transient.

**Proposition 5.** A subset \( A \subset S \) is transient iff the embedded matrix \( P_A \) does not have an eigenvalue one.
Proof. If $A \subset S$ is transient then $P^n_A v \to o(n+\infty)$ for every vector $v$ (use (11)) and therefore the matrix $P^n_A$ does not have an eigenvalue one.

Conversely, suppose $A \subset S$ is not transient, then it follows from Proposition 4 that $v = \lim_{n \to \infty} P^n_A e \neq o$. Since $P^n_A v = v$ we conclude that $P^n_A$ has an eigenvalue one.

**Proposition 6.** If $A \subset S$ is transient then

$$
\sum_{n=0}^{\infty} P^n_A = (I - P_A)^{-1} P_A \quad \text{and} \quad \sum_{n=1}^{\infty} n P^n_A = (I - P_A)^{-2} P_A .
$$

Proof. The following identities are easily verified by induction

$$(I - P_A) \sum_{n=0}^{k} P^n_A = I - P_A^{k+1} , \quad k \geq 0$$

$$(I - P_A)^2 \sum_{n=0}^{k} n P^n_A = k P_A^{k+2} - (k+1) P_A^{k+1} + P_A , \quad k \geq 0 .$$

Using Proposition 5 and (11) and taking $k \to \infty$ completes the proof.

The last proposition of this section states that the chain cannot enter the transient part $T$ from a nontransient state.

**Proposition 7.** For every nontransient state $j \in S$ the probability that the chain moves in one transition from $j$ into the transient part $T$ is zero i.e.

$$
P_j (X_1 \in T) = e^j P e_T = 0 \quad \text{for all } j \in T^* .$$
Proof. Suppose \( j \in S \) is nontransient and \( e_j^T \mathbb{P}_T = r > 0 \). It follows from Proposition 3 that the chain, when started in \( j \), will return to \( j \) infinitely many times. For every integer \( n \geq 1 \) consider the event \( A_n \) that after the \( n \)-th visit to \( j \) the chain moves immediately into \( T \). We have \( \mathbb{P}_j(A_n) = r \) for all \( n \geq 1 \). It follows from the Markov property that the events \( A_n \) are independent, and using the Borel-Cantelli lemma we conclude from
\[
\sum_{n \geq 1} \mathbb{P}_j(A_n) = \infty \quad \text{that} \quad \mathbb{P}_j(A \text{ infinitely often}) = 1.
\]
This contradiction with Proposition 4 completes the proof.

Suppose we have a Markov chain with a nonempty transient part \( T \). Since it is impossible for the chain to enter the transient part \( T \) from outside (Proposition 7), it is sometimes convenient to consider the Markov chain restricted to the transient part \( T \). The corresponding transition matrix is sub-Markov and its entries are the transition probabilities between the transient states in \( T \).

Following the terminology of Kemeny and Snell we denote this matrix by \( Q \) (see Kemeny and Snell (1976) p. 44). This chain with state space \( T \) and transition matrix \( Q \) has only transient states and therefore \( I - Q \) is regular (Proposition 5). As in Kemeny and Snell (1976) p. 46 we define the fundamental matrix \( N \) of the original chain
\[
(12) \quad N := (I - Q)^{-1}.
\]

4. A random walk

We start this section with three propositions that answer the questions mentionned in the introduction.
Proposition 8. For every transient starting position \( j \in T \) and every nontransient state \( k \in T^* \) we have

\[
P_j \text{ (first entrance in } T^* \text{ takes place in } k) = \sum_{T} \mathbf{e}_j^T \mathbf{P}^T \mathbf{e}_k = (I - \mathbf{P}_{T^*})^{-1} \mathbf{P}_{T^*} \mathbf{e}_k.
\]

Proof. Let \( j \) be a transient and \( k \) a nontransient state.

From (6) we obtain \( \mathbf{P}_j \text{ (first entrance in } T^* \text{ takes place in } k) = \sum_{T} \mathbf{e}_j^T \mathbf{P}^T \mathbf{e}_k \).

Using Lemma 3 and (5) we get \( \sum_{T} \mathbf{P}^T = (I - \mathbf{P}_{T^*})^{-1} \mathbf{P}_{T^*} \).

Proposition 9. For every transient subset \( A \subseteq S \) let \( U_A \) be the number of visits to the set \( A \) at times \( n \geq 0 \).

For every \( j \in S \) we have

\[
\mathbb{E}_j U_A = e_j^T (I - \mathbf{P}_A)^{-1} \mathbf{e}_A.
\]

For every \( j \in A \) we have

\[
\mathbb{E}_j U_A^2 = e_j^T (2(I - \mathbf{P}_A)^{-1} - I)(I - \mathbf{P}_A)^{-1} \mathbf{e}_A.
\]

For every \( j \in A^* \) we have

\[
\mathbb{E}_j U_A^2 = e_j^T (2(I - \mathbf{P}_A)^{-1} - 2I - \mathbf{P}_A)(I - \mathbf{P}_A)^{-1} \mathbf{e}_A.
\]

Proof.

\[
\mathbb{E}_j U_A = \mathbb{P}_j (X_0 \in A) + \sum_{n=1}^{\infty} n \mathbb{P}_j (X_n \in A \text{ for exactly } n \text{ different } k \geq 1) \quad (\text{use } (8))
\]

\[
= e_j^T (\mathbf{e}_A + \sum_{n=1}^{\infty} n (\mathbf{P}_A^n \mathbf{e} - \mathbf{P}_A^{n+1} \mathbf{e})) \quad (\text{use } (11))
\]

\[
= e_j^T (\mathbf{e}_A + \sum_{n=1}^{\infty} n \mathbf{P}_A^n \mathbf{e} - \sum_{n=1}^{\infty} (n-1)P_A^n \mathbf{e}) \quad (\text{use } \mathbf{P}_A^n \mathbf{e} = \mathbf{P}_A^n \mathbf{e} \text{ for } n \geq 1)
\]

\[
= e_j^T (\mathbf{e}_A + \sum_{n=1}^{\infty} P_A^n \mathbf{e}_A) \quad (\text{use Proposition 6})
\]

\[
= e_j^T (I - \mathbf{P}_A)^{-1} \mathbf{e}_A.
\]
For every \( j \in A \)

\[
\mathbb{E}_j U_A^2 = \sum_{n=1}^{\infty} (n+1)^2 P_j(X_k \in A \text{ for exactly } n \text{ different } k \geq 1) \quad \text{(use (8))}
\]

\[
= e_j \sum_{n=1}^{\infty} (n+1)^2 (P^n_A e - P^{n+1}_A e) = e_j \left( \sum_{n=1}^{\infty} (n+1)^2 P^n_A e - \sum_{n=2}^{\infty} n^2 P^n_A e \right)
\]

\[
= e_j \left( 2 \sum_{n=1}^{\infty} n P^n_A e + \sum_{n=0}^{\infty} P^n_A e \right) \quad \text{(use Proposition 6)}
\]

\[
= e_j (2(I - P_A)^{-2} P_A e + (I - P_A)^{-1} e) = e_j (2(I - P_A)^{-1} - I)(I - P_A)^{-1} e.
\]

For every \( j \in A^* \) we have

\[
\mathbb{E}_j U_A^2 = \sum_{n=1}^{\infty} n^2 P_j(X_k \in A \text{ for exactly } n \text{ different } k \geq 1)
\]

and the proof is similar. \( \square \)

**Corollary 1.** For \( U_T \), the number of visits to the transient part \( T \) at times \( n \geq 0 \), we have for every \( j \in T \)

\[
\mathbb{E}_j U_T = e_j Ne \quad \text{and} \quad \mathbb{E}_j U_T^2 = e_j (2N - I)Ne
\]


**Proof.** Consider the chain restricted to the transient part \( T \) with transition matrix \( Q \) (see (12)). In this case we have in Proposition 9 \( A = T, P_A = Q \) and \( e_A = e_T = e \). \( \square \)

In the last proposition we consider transient chains i.e. chains where \( T = S \).

**Proposition 10.** For a transient chain and an arbitrary subset \( A \subset S \) let \( V_A \)

be the number of different states in \( A \) visited by chain at times \( n \geq 1 \) and \( W_A \)

the number of different states in \( A \) visited by the chain at times \( n \geq 0 \).
For every $j \in S$ we have

$$E_j V_A = e_j \sum_{i \in A} P_i e$$

$$E_j V_A^2 = e_j \left\{ \sum_{i \in A} P_i e + 2 \sum_{i, k \in A, \{i, k\} \neq \emptyset} (I - PI_{i,k})^{-1}(PI_i P_k e + PI_k P_i e) \right\}.$$

If $j \in A^*$ then $E_j W_A = E_j V_A$ and $E_j W_A^2 = E_j V_A^2$. For $j \in A$ we have

$$E_j W_A = 1 + e_j \sum_{i \in A \setminus \{j\}} P_i e$$

$$E_j W_A^2 = 1 + e_j \left\{ 3 \sum_{i \in A \setminus \{j\}} P_i e + 2 \sum_{i, k \in A \setminus \{j\}, \{i, k\} \neq \emptyset} (I - PI_{i,k})^{-1}(PI_i P_k e + PI_k P_i e) \right\}.$$

**Proof.** We only prove the formulas for $V_A$. The formulas for $W_A$ can be proved similarly.

For $i \in A$ consider the random variable

$$M_i := \begin{cases} 1 & \text{if } X_n = i \text{ for some } n \geq 1 \\ 0 & \text{if } X_n \neq i \text{ for all } n \geq 1. \end{cases}$$

Evidently $V_A = \sum_{i \in A} M_i$.

For every starting position $j \in S$ and every $i \in A$ we have (use (6))

$$P_j(M_i = 1) = P_j(\exists n \geq 1 : X_n = i) = e_j P_i e.$$

Hence $E_j V_A = e_j \sum_{i \in A} P_i e$. 
The set $T = \{1, 2, \ldots, 7\}$ is the transient part of the chain. For the embedded matrix $P^*_T$ we obtain

$$P^*_T = (\text{use (5) and Lemma 3}) (I - PL_T)^{-1} PI_T^* =$$

$$\begin{bmatrix}
3465 & 1485 & 1485 & 2112 & 1320 & 1320 & 1188 & 0 & 0 \\
1485 & 2811 & 687 & 1276 & 1476 & 768 & 880 & 0 & 0 \\
1485 & 687 & 2811 & 1276 & 1476 & 880 & 0 & 0 & 0 \\
1584 & 957 & 957 & 3784 & 1716 & 1716 & 1804 & 0 & 0 \\
990 & 1107 & 576 & 1716 & 3108 & 1452 & 0 & 0 & 0 \\
990 & 576 & 1107 & 1716 & 984 & 1452 & 0 & 0 & 0 \\
891 & 660 & 660 & 1804 & 1452 & 3124 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1947 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1947 \\
\end{bmatrix}$$

$$\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 12 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 12 \\
\end{bmatrix}$$

$$\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 12 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 12 \\
\end{bmatrix}$$

$$\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 957 & 990 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 781 & 1166 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 781 & 1166 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1309 & 638 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1386 & 561 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1386 & 561 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1507 & 440 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1947 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1947 \\
\end{bmatrix}$$
Using Proposition 8 we can calculate all absorption probabilities e.g.

\[ P_3 \text{ (absorption takes place in vertex 8)} = \]

\[ e_3^T P_T^* e_8 = \frac{781}{1947} = 0.401. \]

For the fundamental matrix \( N = (I - Q)^{-1} \) (see (12)) we obtain

\[
\begin{bmatrix}
3465 & 1485 & 1485 & .2112 & 1320 & 1320 & 1188 \\
1485 & 2811 & 687 & 1276 & 1476 & 768 & 880 \\
1485 & 687 & 2811 & 1276 & 768 & 1476 & 880 \\
1584 & 957 & 957 & 3784 & 1716 & 1716 & 1804 \\
990 & 1107 & 576 & 1716 & 3108 & 984 & 1452 \\
990 & 576 & 1107 & 1716 & 984 & 3108 & 1452 \\
891 & 660 & 660 & 1804 & 1452 & 1452 & 3124
\end{bmatrix}
\]

Hence \( (N e)' = \frac{1}{1947}(12375, 9383, 9383, 12518, 9933, 9933, 10043) \) and \( ((2N - I)N e)' = \frac{1}{1947^2}(246584085, 178329437, 178329437, 249901058, 190164447, 190164447, 192874121) \).

Since the random variable \( U_T \) in Corollary 1 equals the absorption time we can calculate, using Corollary 1, the mean and variance of the absorption time starting in an arbitrary transient state e.g.

\[ \mathbb{E}_1 \text{ (absorption time)} = e_1^T N e = \frac{12375}{1947} = 6.356, \]

second moment absorption time starting in 1 =

\[ e_1^T (2N - I)N e = \frac{246584085}{1947^2} = 65.048 \]

\[ \text{var}_1 \text{ (absorption time)} = 24.650. \]
From now we shall restrict the random walk to its transient part i.e.
we consider the transient chain on the state space \{1,2,\ldots,7\} with transition matrix

\[
P = \frac{1}{12} \begin{pmatrix}
    0 & 4 & 4 & 4 & 0 & 0 & 0 \\
    4 & 0 & 0 & 0 & 4 & 0 & 0 \\
    4 & 0 & 0 & 0 & 0 & 4 & 0 \\
    3 & 0 & 0 & 0 & 3 & 3 & 3 \\
    0 & 3 & 0 & 3 & 0 & 0 & 3 \\
    0 & 0 & 3 & 3 & 0 & 0 & 3 \\
    0 & 0 & 0 & 3 & 3 & 3 & 0
\end{pmatrix}.
\]

Let \( A = \{5,6,7\} \). We calculate the embedded matrix \( P_A = \) (use (5) and Lemma 3)

\[
(I - P_{A*})^{-1} P_A =
\]

\[
\begin{pmatrix}
    144 & 48 & 48 & 48 & 0 & 0 & 0 \\
    48 & 116 & 16 & 16 & 0 & 0 & 0 \\
    48 & 16 & 116 & 16 & 0 & 0 & 0 \\
    36 & 12 & 12 & 112 & 0 & 0 & 0 \\
    21 & 32 & 7 & 32 & 100 & 0 & 0 \\
    21 & 7 & 32 & 32 & 0 & 100 & 0 \\
    9 & 3 & 3 & 28 & 0 & 0 & 100
\end{pmatrix}
\]

\[
\begin{pmatrix}
    0 & 0 & 0 & 0 & 0 & 0 & 0 \\
    0 & 0 & 0 & 0 & 4 & 0 & 0 \\
    0 & 0 & 0 & 0 & 0 & 4 & 0 \\
    0 & 0 & 0 & 0 & 3 & 3 & 3 \\
    0 & 0 & 0 & 0 & 0 & 3 & 3 \\
    0 & 0 & 0 & 0 & 0 & 0 & 3 \\
    0 & 0 & 0 & 0 & 3 & 3 & 0
\end{pmatrix} =
\]

\[
\begin{pmatrix}
    0 & 0 & 0 & 0 & 84 & 84 & 36 \\
    0 & 0 & 0 & 0 & 128 & 28 & 12 \\
    0 & 0 & 0 & 0 & 28 & 128 & 12 \\
    0 & 0 & 0 & 0 & 96 & 96 & 84 \\
    0 & 0 & 0 & 0 & 56 & 31 & 99 \\
    0 & 0 & 0 & 0 & 31 & 56 & 99 \\
    0 & 0 & 0 & 0 & 99 & 99 & 21
\end{pmatrix}.
\]
Using the embedded matrix $P_A$ we can calculate all recurrence probabilities for the set $A$ e.g.

$E_1$ (first entrance in the set $A$ takes place at vertex 7) = (use (6))

$$e_1^T P_A e_7 = \frac{36}{300} = 0.12$$

$\mathbb{P}_1$ (the chain will ever visit the set $A$) = (use (7))

$$e_1^T P_A e = \frac{204}{300} = 0.68$$

From Proposition (5) we know that $I - P_A$ is regular. We find

$$\begin{pmatrix}
1 & 0 & 0 & 0 & 0.6780 & 0.6780 & 0.6102 \\
0 & 1 & 0 & 0 & 0.7581 & 0.3945 & 0.4520 \\
0 & 0 & 1 & 0 & 0.3945 & 0.7581 & 0.4520 \\
\end{pmatrix}
\begin{pmatrix}
1 - P_A \end{pmatrix}^{-1} =
\begin{pmatrix}
0 & 0 & 0 & 1 & 0.8814 & 0.8814 & 0.9266 \\
0 & 0 & 0 & 0 & 1.5963 & 0.5054 & 0.7458 \\
0 & 0 & 0 & 0 & 0.5054 & 1.5963 & 0.7458 \\
0 & 0 & 0 & 0 & 0.7458 & 0.7458 & 1.6045 \\
\end{pmatrix}$$

Hence (see Proposition 9)

$$\{ (I - P_A)^{-1} e_A \} ' = (1.966, 1.605, 1.605, 2.689, 2.8475, 2.8475, 3.0961)$$

and

$$\{ 2(I - P_A)^{-1} - I)(I - P_A)^{-1} e \} ' =

Using Proposition 9 we can calculate the mean and variance of the number of visits to the set $A$. We obtain e.g.
\[ E_7 (\text{number of visits to the set A starting position included}) = \]
\[ e_7' (I - P^*_A)^{-1} e_A = 3.0961 \]

\[ \text{var}_7 (\text{number of visits to the set A starting position included}) = \]
\[ 11.087 - (3.0961)^2 = 0.336. \]

We now calculate the embedded matrices \( P_5, P_6 \) and \( P_7 \).

\[ P_5 = (\text{use (5) and Lemma 3}) = (I - PI_{(5)^*})^{-1} PI_5 = \]
\[
\begin{pmatrix}
0 & 0 & 0 & 0 & 440 & 0 & 0 \\
0 & 0 & 0 & 0 & 492 & 0 & 0 \\
0 & 0 & 0 & 0 & 256 & 0 & 0 \\
0 & 0 & 0 & 0 & 572 & 0 & 0 \\
0 & 0 & 0 & 0 & 387 & 0 & 0 \\
0 & 0 & 0 & 0 & 328 & 0 & 0 \\
0 & 0 & 0 & 0 & 484 & 0 & 0 
\end{pmatrix}
\]
\[
\begin{pmatrix}
0 & 0 & 0 & 0 & 440 & 0 \\
0 & 0 & 0 & 0 & 256 & 0 \\
0 & 0 & 0 & 0 & 492 & 0 \\
0 & 0 & 0 & 0 & 572 & 0 \\
0 & 0 & 0 & 0 & 387 & 0 \\
0 & 0 & 0 & 0 & 328 & 0 \\
0 & 0 & 0 & 0 & 484 & 0 
\end{pmatrix}
\]

\[ P_6 = \frac{1}{1036} \]

\[ P_7 = \frac{1}{3124} \]

\[
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 1188 \\
0 & 0 & 0 & 0 & 0 & 880 \\
0 & 0 & 0 & 0 & 0 & 880 \\
0 & 0 & 0 & 0 & 0 & 1804 \\
0 & 0 & 0 & 0 & 0 & 1452 \\
0 & 0 & 0 & 0 & 0 & 1452 \\
0 & 0 & 0 & 0 & 0 & 1177 
\end{pmatrix}
\]
Hence e.g.

\[ 
\mathbb{P}_5 \left( \exists \, n \geq 1 \mid X_n = 5 \right) = \text{(use (7))} = e^5 \cdot \mathbb{P}_5 \cdot e = \frac{387}{1036} = 0.374. 
\]

\[ 
\mathbb{P}_1 \left( \exists \, n \geq 1 \mid X_n = 7 \right) = \frac{1188}{3124} = 0.380. 
\]

\( \mathbb{E}_1 \) (number of different states in A visited by the chain) =

\[ 
(\text{use Proposition 10}) \, e^1 \left( \mathbb{P}_5 \cdot e + \mathbb{P}_6 \cdot e + \mathbb{P}_7 \cdot e \right) = \frac{880}{1036} + \frac{1188}{3124} = 1.230. 
\]

References

Van Nostrand Mathematical Studies #21, New York,

Kemeny, J.G. and J. Snell, Finite Markov Chains,
Berlin etc., Springer Verlag, 1976.
