Embedded matrices for finite Markov chains

EINDHOVEN UNIVERSITY OF TECHNOLOGY

Department of Mathematics and Computing Science

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 $X_k \notin S$ for all $k \geq n$.

For an arbitrary subset $A \subset S$ 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).

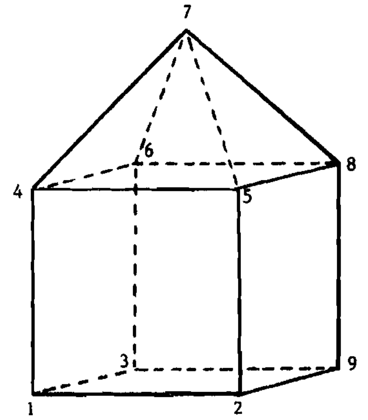

Figure 1

The vertices 8 and 9 are absorbing and in all other vertices edges are choosen 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

(0) $\quad I_j := e_j e_j' \qquad (e_j'$ is the transpose of $e_j)$.

For every subset $A \subset S$ we define

$$e_A := \sum_{j \in A} e_j ,$$

$$I_A := \sum_{j \in A} I_j ,$$

$$I := I_S \qquad (s \times s \text{ identity matrix}) ,$$

$$e := e_S \qquad (\text{column vector with all components one}) ,$$

$$o := e_\phi \qquad (\text{zero column vector}) .$$

The following relations can easily be verified.

(1) The matrix element $\quad p_{jk} = e_j' P e_k , \quad 1 \le j \le s$ and $1 \le k \le s$

(2) $\quad I_j e = e_j \quad , \quad 1 \le j \le s$

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

(4) $\quad I_A I_B = I_{A \cap B} \quad , \quad A \subset S$ 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.

$$e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad I_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad e_A = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \text{ and } I_A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

It is well-known that the probability distribution of a Markov chain is determined by its transition matrix P and an initial distribution π of X_0 on S. The corresponding probability distribution of the Markov chain is denoted by $\mathbb{P}_\pi$. An initial distribution on S is denoted by a column vector π with non-negative components summing to one, i.e.

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

$$\pi' e = 1 .$$

If $\pi = e_j$, we write $\mathbb{P}_j := \mathbb{P}_{e_j}$.

Expectations with respect to $\mathbb{P}_\pi$ or $\mathbb{P}_j$ are denoted by the symbol $\mathbb{E}_\pi$ or $\mathbb{E}_j$.

2. Embedded matrices

We start with a proposition that is the key to calculating probabilities with respect to $\mathbb{P}_\pi$.

<u>Proposition 1</u>. For every integer $n \geq 0$, every initial distribution π on S and every (n+1)-tuple subsets $A_0, A_1, \ldots, A_n \subset S$

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

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

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

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

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

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

$$= \sum_{j \in A_n} \sum_{k \in A_{n+1}} \pi' I_{A_0} P I_{A_1} \ldots P I_{A_{n-1}} P I_j\, e\, p_{jk} \qquad \text{(use (1))}$$

$$= \sum_{j \in A_n} \sum_{k \in A_{n+1}} \pi' I_{A_0} P I_{A_1} \ldots P I_j\, e\, e_j'\, P e_k \qquad \text{(use (2))}$$

$$= \sum_{j \in A_n} \sum_{k \in A_{n+1}} \pi' I_{A_0} P I_{A_1} \ldots P\, e_j\, e_j'\, P e_k \qquad \text{(use (0) and (2))}$$

$$= \sum_{j \in A_n} \sum_{k \in A_{n+1}} \pi' I_{A_0} P I_{A_1} \ldots P I_j\, P I_k\, e = \pi' I_{A_0} P I_{A_1} \ldots P I_{A_n} P I_{A_{n+1}} e \ . \qquad \square$$

<u>Remark</u>. If $\pi = e_j$ then

$$\mathbb{P}_j (X_1 \in A_1, \ldots, X_n \in A_n) = e_j' I_S P I_{A_1} \ldots P I_{A_n} e = e_j' P I_{A_1} \ldots P I_{A_n} e \ .$$

For the definition of embedded matrices we need the following lemma.

Lemma 1. For every subset $A \subset S$ the sequence of matrices

$$\left\{ \sum_{n=0}^{N} (PI_{A^*})^n \, PI_A \right\}_{N \geq 0}$$

is convergent.

Proof. For every integer $N \geq 0$ define the matrix $Q_N := \sum_{n=0}^{N} (PI_{A^*})^n PI_A$.

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

$$\begin{aligned} Q_N(j,k) &= e_j' \, Q_N \, e_k = \sum_{n=0}^{N} e_j' (PI_{A^*})^n \, PI_{A \cap \{k\}} \, e \\ &= \sum_{n=1}^{N+1} e_j' (PI_{A^*})^{n-1} \, PI_{A \cap \{k\}} \, e \\ &= \mathbb{P}_j (\exists_{1 \leq n \leq N+1} : X_1 \in A^* , \ldots , X_{n-1} \in A^* , X_n \in A \cap \{k\}) , \end{aligned}$$

which is nondecreasing and bounded by one.

Hence

$$\lim_{N \to \infty} Q_N(j,k) = \mathbb{P}_j (\exists_{n \geq 1} : X_1 \in A^* , \ldots , X_{n-1} \in A^* , X_n \in A \cap \{k\}) . \qquad \square$$

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

Definition 1. For every subset $A \subset S$ the embedded matrix P_A is defined by

$$P_A := \sum_{n=0}^{\infty} (PI_{A^*})^n \, PI_A \; .$$

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.

<u>Lemma 2.</u> For every subset $A \subset S$

$$(I - PI_{A^*})P_A = PI_A .$$

<u>Proof.</u>

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

$$PI_A + PI_{A^*} \sum_{n=0}^{\infty} (PI_{A^*})^n PI_A = PI_A + PI_{A^*}P_A . \qquad \square$$

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

$$(5) \qquad P_A = (I - PI_{A^*})^{-1} PI_A .$$

The following lemma supplies a probabilistic characterization of the presence of an eigenvalue one of the matrix PI_A.

<u>Lemma 3.</u> Let $A \subset S$. The matrix PI_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 PI_A has an eigenvalue one. Then there exists a columnvector $v \neq 0$ such that $PI_A v = v$.
Consider the vectors v^+ and v^- with components $v^+(i) := \max(v(i),0)$ and $v^-(i) = \max(-v(i),0)$. Then $v = v^+ - v^- = PI_A v^+ - PI_A v^-$. Since the vectors $PI_A v^+$ and $PI_A v^-$ are nonnegative $PI_A v^+ \geq v^+$ and $PI_A v^- \geq v^-$.
(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 $PI_A w \geq w$. Let j-th component of w be positive then from Proposition 1 with $\pi = e_j$ it follows that

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

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

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 π on S, all subsets $A, B \subset S$ and all integers $n \geq 1$

$$\mathbb{P}_\pi(\exists_{n\geq 1} : X_1 \in A^*, \ldots, X_{n-1} \in A^*, X_n \in A \cap B) = \pi' P_A e_B . \tag{6}$$

$$\mathbb{P}_\pi(X_k \in A \text{ for at least n different } k \geq 1) = \pi' P_A^n e . \tag{7}$$

(8) $\mathbb{P}_\pi(X_k \in A \text{ for exactly } n \text{ different } k \geq 1) = \pi'(P_A^n e - P_A^{n+1} e)$.

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

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

<u>Proof</u>. We only prove (6); the other assertions can be proved similarly.

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

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

$$= \sum_{k=0}^{\infty} \pi'(PI_{A^*})^k PI_{A\cap B}\, e \qquad \text{(use (4) and (3))}$$

$$= \pi' P_A e_B . \qquad \square$$

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

<u>Lemma 4</u>. For every subset $A \subset S$ the sequence of columnvectors $\{P_A^n e\}_{n\geq 0}$ is nonincreasing.

<u>Proof</u>. For the j-th component $P_A^n e(j)$ of the vector $P_A^n e$ we have (use (7))

$$P_A^n e(j) = e_j' P_A^n e = \mathbb{P}_j(X_k \in A \text{ for at least } n \text{ different } k \geq 1) . \qquad \square$$

3. <u>Transient sets</u>

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 $A = \{j\}$)

$$e_j' P_j e < 1 .$$

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 $0 \le q < 1$ such that for all $n \ge 1$

$$P_j^n e \le q^{n-1} e .$$

Proof. Suppose $j \in S$ is transient then by definition $q := e_j' P_j e < 1$.
Hence $I_j P_j e \le q e$.
We proceed by induction. For $n = 1$ the proposition is trivial. Suppose the proposition has been proved for $0 \le k \le n$.
We have

$$\begin{aligned} P_j^{n+1} e &= P_j^n P_j e && (\text{use } P_j^n = P_j^n I_j \text{ for } n \ge 1) \\ &= P_j^n I_j P_j e && (\text{use } I_j P_j e \le q e) \\ &\le P_j^n q e = q P_j^n e \le q^n e . \end{aligned}$$

□

Proposition 3. For every nontransient state $j \in S$

$$e_j' P_j^n e = 1 \text{ for all } n \ge 1 .$$

Hence $\mathbb{P}_j(X_n = j \text{ infinitely often}) = 1$ (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

$$e_j' P_j^{n+1} e = e_j' P_j^n P_j e \qquad \text{(use } P_j^n = P_j^n I_j \text{ for } n \geq 1\text{)}$$

$$= e_j' P_j^n I_j P_j e$$

$$= e_j' P_j^n e_j = e_j' P_j^n e = 1 . \qquad \square$$

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))

$$e_j' \lim_{n\to\infty} P_A^n e = 0 \text{ for all } j \in S .$$

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

$$e_j' \lim_{n\to\infty} P_A^n e \leq \sum_{k\in A} e_j' \lim_{n\to\infty} P_k^n e .$$

Since every $k \in A$ is transient we conclude from Lemma 5 that $\lim_{n\to\infty} e_j' P_A^n e = 0$.
Conversely suppose $A \subset S$ is not transient then there exists $j \in A$ such that $e_j' P_j^n e = 1$ for all $n \geq 1$ (Proposition 3). Since $\mathbb{P}_j(X_n = j$ infinitely often) $\leq \mathbb{P}_j(X_n \in A$ infinitely often) we have $\lim_{n\to\infty} e_j' P_A^n e \geq \lim_{n\to\infty} e_j' P_j^n e = 1$. Hence $\lim_{n\to\infty} e_j' P_A^n e \neq 0$. $\square$

If the subset $A \subset S$ is transient Proposition 4 states that $\lim_{n\to\infty} P_A^n e = 0$. 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 $0 \le r < 1$ such that

$$P_A^n e \le r^n e \text{ for all } n \ge n_0 .$$

Proof. From Proposition 4 and Lemma 4 we know: $P_A^n e \downarrow 0$. Hence there exist an integer n_0 and a number $0 \le w < 1$ such that

$$P_A^n e \le w e \text{ for all } n \ge n_0 .$$

Suppose $n \ge 2n_0$ then $n \ge n_0\left[\frac{n}{n_0}\right]$ ([a]:= largest integer not exceeding a). Hence

$$P_A^n e \le P_A^{n_0\left[\frac{n}{n_0}\right]} e \le w^{\left[\frac{n}{n_0}\right]} e \le w^{\frac{n}{2n_0}} e .$$

Put $r := w^{\frac{1}{2n_0}}$ then $0 \le r < 1$ and $P_A^n e \le r^n e$ for all $n \ge 2n_0$. □

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

$$e_j' P_A^n e_k \le e_j' P_A^n e \le r^n \text{ for n sufficiently large.}$$

(11) Hence the sequence of matrices $\{P_A^n\}_{n\ge 0}$ 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 $P_A^n v \to o (n\to\infty)$ for every vector v (use (11)) and therefore the matrix P_A does not have an eigenvalue one.
Conversely, suppose $A \subset S$ is not transient, then it follows from Proposition 4 that $v = \lim_{n\to\infty} P_A^n e \neq o$. Since $P_A v = v$ we conclude that P_A has an eigenvalue one. □

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

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

Proof. The following identities are easily verified by induction

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

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

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

The last proposition of this section states that the chain can not 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.

$$\mathbb{P}_j (X_1 \in T) = e_j' P e_T = 0 \text{ for all } j \in T^* .$$

Proof. Suppose $j \in S$ is nontransient and $e_j' P e_T = r > 0$. 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 $\sum_{n=1}^{\infty} \mathbb{P}_j(A_n) = \infty$ that $\mathbb{P}_j(A_n \text{ infinitely often}) = 1$.
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) $$N := (I - Q)^{-1} .$$

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

$$\mathbb{P}_j\,(\text{first entrance in } T^* \text{ takes place in } k) =$$
$$e_j'\,P_{T^*}\,e_k = (I - PI_T)^{-1}\,PI_{T^*}\,e_k\ .$$

Proof. Let j be a transient and k a nontransient state.
From (6) we obtain $\mathbb{P}_j$ (first entrance in T^* takes place in k) $= e_j'\,P_{T^*}\,e_k$.
Using Lemma 3 and (5) we get $P_{T^*} = (I - PI_T)^{-1}\,PI_{T^*}$. □

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

For every $j \in S$ we have $\mathbb{E}_j U_A = e_j'(I - P_A)^{-1}\,e_A$.
For every $j \in A$ we have $\mathbb{E}_j U_A^2 = e_j'(2(I - P_A)^{-1} - I)(I - P_A)^{-1}\,e$.
For every $j \in A^*$ we have $\mathbb{E}_j U_A^2 = e_j'(2(I - P_A)^{-1} - 2I - P_A)(I - P_A)^{-1}\,e$.

Proof.

$$\mathbb{E}_j U_A = \mathbb{P}_j\,(X_0 \in A) + \sum_{n=1}^{\infty} n\,\mathbb{P}_j(X_k \in A \text{ for exactly } n \text{ different } k \geq 1) \qquad \text{(use (8))}$$

$$= e_j'(e_A + \sum_{n=1}^{\infty} n(P_A^n\,e - P_A^{n+1}\,e)) \qquad \text{(use (11))}$$

$$= e_j'(e_A + \sum_{n=1}^{\infty} n\,P_A^n\,e - \sum_{n=1}^{\infty} (n-1)P_A^n\,e) \qquad (\text{use } P_A^n\,e = P_A^n\,e_A \text{ for } n \geq 1)$$

$$= e_j'(e_A + \sum_{n=1}^{\infty} P_A^n\,e_A) = \quad \text{(use Proposition 6)} \quad e_j'(I - P_A)^{-1}\,e_A .$$

For every $j \in A$

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

$$= e_j' \sum_{n=1}^{\infty} (n+1)^2 (P_A^n e - P_A^{n+1} e) = e_j' \left(\sum_{n=1}^{\infty} (n+1)^2 P_A^n e - \sum_{n=2}^{\infty} n^2 P_A^n e \right)$$

$$= e_j' \left(2 \sum_{n=1}^{\infty} n P_A^n e + \sum_{n=0}^{\infty} P_A^n e \right) \qquad \text{(use Proposition 6)}$$

$$= e_j'(2(I - P_A)^{-2} P_A e + (I - P_A)^{-1} e) = e_j'(2(I - P_A)^{-1} - I)(I - P_A)^{-1} e .$$

For every $j \in A^*$ we have

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

and the proof is similar. □

<u>Corollary 1</u>. For U_T, the number of visits to the transient part T at times $n \geq 0$, we have for every $j \in T$

$$\mathbb{E}_j U_T = e_j' N e \text{ and } \mathbb{E}_j U_T^2 = e_j'(2N - I) N e$$

(cf. *Kemeny* and *Snell* (1976) p. 51) .

<u>Proof</u>. 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 $e_A = e_T = e$. □

In the last proposition we consider transient chains i.e. chains where $T = S$.

<u>Proposition 10</u>. 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

$$\mathbb{E}_j V_A = e_j' \sum_{i\in A} P_i e$$

$$\mathbb{E}_j V_A^2 = e_j' \left\{ \sum_{i\in A} P_i e + 2 \sum_{\substack{i,k\in A \\ i<k}} (I - PI_{\{i,k\}^*})^{-1} (PI_i P_k e + PI_k P_i e) \right\} .$$

If $j \in A^*$ then $\mathbb{E}_j W_A = \mathbb{E}_j V_A$ and $\mathbb{E}_j W_A^2 = \mathbb{E}_j V_A^2$. For $j \in A$ we have

$$\mathbb{E}_j W_A = 1 + e_j' \sum_{i\in A\setminus\{j\}} P_i e$$

$$\mathbb{E}_j W_A^2 = 1 + e_j' \left\{ 3 \sum_{i\in A\setminus\{j\}} P_i e + 2 \sum_{\substack{i,k\in A\setminus\{j\} \\ i<k}} (I - PI_{\{i,k\}^*})^{-1} (PI_i P_k e + PI_k P_i e) \right\}.$$

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))

$$\mathbb{P}_j (M_i = 1) = \mathbb{P}_j(\exists_{n\geq 1} : X_n = i) = e_j' P_i e .$$

Hence $\mathbb{E}_j V_A = e_j' \sum_{i\in A} P_i e$.

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

The set $T = \{1,2,\dots,7\}$ is the transient part of the chain. For the embedded matrix P_{T^*} we obtain

$$P_{T^*} = \text{(use (5) and Lemma 3)} \quad (I - PI_T)^{-1}\, PI_{T^*} =$$

$$\frac{1}{1947}\begin{pmatrix} 3465 & 1485 & 1485 & 2112 & 1320 & 1320 & 1188 & 0 & 0 \\ 1485 & 2811 & 687 & 1276 & 1476 & 768 & 880 & 0 & 0 \\ 1485 & 687 & 2811 & 1276 & 768 & 1476 & 880 & 0 & 0 \\ 1584 & 957 & 957 & 3784 & 1716 & 1716 & 1804 & 0 & 0 \\ 990 & 1107 & 576 & 1716 & 3108 & 984 & 1452 & 0 & 0 \\ 990 & 576 & 1107 & 1716 & 984 & 3108 & 1452 & 0 & 0 \\ 891 & 660 & 660 & 1804 & 1452 & 1452 & 3124 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1947 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1947 \end{pmatrix} \frac{1}{12}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 12 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 12 \end{pmatrix} =$$

$$\frac{1}{1947}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 957 & 990 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 781 & 1166 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 781 & 1166 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1309 & 638 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1386 & 561 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1386 & 561 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1507 & 440 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1947 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1947 \end{pmatrix}.$$

Using Proposition 8 we can calculate all absorption probabilities e.g.

$\mathbb{P}_3$ (absorption takes place in vertex 8) =

$e_3' \, P_{T^*} \, e_8 = \frac{781}{1947} = 0.401$.

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

$$N = \frac{1}{1947} \begin{pmatrix} 3465 & 1485 & 1485 & 2112 & 1320 & 1320 & 1188 \\ 1485 & 2811 & 687 & 1276 & 1476 & 768 & 880 \\ 1485 & 687 & 2811 & 1276 & 768 & 1476 & 880 \\ 1584 & 957 & 957 & 3784 & 1716 & 1716 & 1804 \\ 990 & 1107 & 576 & 1716 & 3108 & 984 & 1452 \\ 990 & 576 & 1107 & 1716 & 984 & 3108 & 1452 \\ 891 & 660 & 660 & 1804 & 1452 & 1452 & 3124 \end{pmatrix} .$$

Hence $(N\,e)' = \frac{1}{1947}(12375, 9383, 9383, 12518, 9933, 9933, 10043)$ and $((2N - I)N\,e)' = \frac{1}{1947^2}(246584085, 178329437, 178329437, 249901058, 190164447, 19016447, 192874121)$.

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.

$\mathbb{E}_1$ (absorption time) $= e_1' \, N\,e = \frac{12375}{1947} = 6.356$,

second moment absorption time starting in 1 =

$e_1'(2N - I)N\,e = \frac{246584085}{1947^2} = 65.048$

var_1 (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) $(I - PI_{A^*})^{-1} PI_A$ =

$$\frac{1}{100}\begin{pmatrix} 144 & 48 & 48 & 48 & 0 & 0 & 0 \\ 48 & 116 & 16 & 16 & 0 & 0 & 0 \\ 48 & 16 & 116 & 16 & 0 & 0 & 0 \\ 36 & 12 & 12 & 112 & 0 & 0 & 0 \\ 21 & 32 & 7 & 32 & 100 & 0 & 0 \\ 21 & 7 & 32 & 32 & 0 & 100 & 0 \\ 9 & 3 & 3 & 28 & 0 & 0 & 100 \end{pmatrix} \frac{1}{12}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 3 & 3 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 3 & 3 & 0 \end{pmatrix} =$$

$$\frac{1}{300}\begin{pmatrix} 0 & 0 & 0 & 0 & 84 & 84 & 36 \\ 0 & 0 & 0 & 0 & 128 & 28 & 12 \\ 0 & 0 & 0 & 0 & 28 & 128 & 12 \\ 0 & 0 & 0 & 0 & 96 & 96 & 84 \\ 0 & 0 & 0 & 0 & 56 & 31 & 99 \\ 0 & 0 & 0 & 0 & 31 & 56 & 99 \\ 0 & 0 & 0 & 0 & 99 & 99 & 21 \end{pmatrix} .$$

Using the embedded matrix P_A we can calculate all recurrence probabilities for the set A e.g.

$\mathbb{P}_1$ (first entrance in the set A takes place at vertex 7) = (use (6))

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

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

$$e_1' P_A e = \frac{204}{300} = 0.68 \ .$$

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

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

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

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

$$(5.750, 7.286, 7.286, 11.577, 9.992, 9{,}992, 11.087) \ .$$

Using Proposition 9 we can calculate the mean and variance of the number of visits to the set A. We obtain e.g.

$\mathbb{E}_7$ (number of visits to the set A starting position included) =

$$e_7'(I - P_A)^{-1} e_A = 3.0961$$

var_7 (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.

$$P_5 = \text{(use (5) and Lemma 3)} = (I - PI_{\{5\}^*})^{-1} PI_5 =$$

$$\frac{1}{1036}\begin{pmatrix} 0 & 0 & 0 & 0 & 440 & 0 & 0 \\ 0 & 0 & 0 & 0 & 492 & 0 & 0 \\ 0 & 0 & 0 & 0 & 256 & 0 & 0 \\ 0 & 0 & 0 & 0 & 572 & 0 & 0 \\ 0 & 0 & 0 & 0 & 387 & 0 & 0 \\ 0 & 0 & 0 & 0 & 328 & 0 & 0 \\ 0 & 0 & 0 & 0 & 484 & 0 & 0 \end{pmatrix}, \quad P_6 = \frac{1}{1036}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 440 & 0 \\ 0 & 0 & 0 & 0 & 0 & 256 & 0 \\ 0 & 0 & 0 & 0 & 0 & 492 & 0 \\ 0 & 0 & 0 & 0 & 0 & 572 & 0 \\ 0 & 0 & 0 & 0 & 0 & 328 & 0 \\ 0 & 0 & 0 & 0 & 0 & 387 & 0 \\ 0 & 0 & 0 & 0 & 0 & 484 & 0 \end{pmatrix}$$

$$\text{and } P_7 = \frac{1}{3124}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1188 \\ 0 & 0 & 0 & 0 & 0 & 0 & 880 \\ 0 & 0 & 0 & 0 & 0 & 0 & 880 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1804 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1452 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1452 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1177 \end{pmatrix} .$$

Hence e.g.

$$\mathbb{P}_5(\exists_{n\geq 1} : X_n = 5) = (\text{use } (7)) = e_5' P_5 e = \frac{387}{1036} = 0.374.$$

$$\mathbb{P}_1(\exists_{n\geq 1} : X_n = 7) = \frac{1188}{3124} = 0.380\ .$$

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

(use Proposition 10) $e_1'(P_5\, e + P_6\, e + P_7\, e) = \frac{880}{1036} + \frac{1188}{3124} = 1.230$.

References

Foguel, S.R., The ergodic theory of Markov processes.
Van Nostrand Mathematical Studies # 21, New York,
Van Nostrand Reinhold, 1969.

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

Revuz, D., Markov Chains, North-Holland Publishing Company, 1975.

Simons, F.H. and D.A. Overdijk, Recurrent and sweep-out sets for Markov
processes. Mh. Math. 86, 305-326 (1979).