ELIMINATION AND INSERTION OPERATIONS FOR FINITE
MARKOV CHAINS

Abstract. A Markov chain (MC) observed only when it is outside of a subset \(D\) is again a MC
with a well-known transition matrix \(P_D\). This matrix can be obtained also in a few iterations, each
requiring \(O(n^2)\) operations, when the states from \(D\) are "eliminated" one at a time. We modify these
iterations to allow for a state previously eliminated to be "reinserted" into the state space in one
iteration. This modification sheds a new light on the relationship between an initial and censored
MC, and introduces a new operation - "insertion" into the theory of MCs.

Key words. Markov chain, censored Markov chain, State Elimination Algorithm.

1. Introduction. Let \(X\) be a finite state space, \(|X| = n, P = \{p(x,y)\}\) a
stochastic matrix indexed by elements of \(X\), and \((Z_n)\) a Markov chain (MC) defined
by \(X, P\) with some initial distribution. Let \(D \subset X, S = X \setminus D\). It is well
known that a MC observed only at visits to the subset \(S\) is again a MC (sometimes
called a censored or embedded MC) with a new transition matrix \(P_D\), see formula
(2.4) below. This formula can be found e.g. in the classical text [3]. We say that set
\(D\) is "eliminated." Such a transition matrix has an especially simple form when
\(D\) consists of only one point, \(D = \{z\}\). In this case, if we denote
\(P_{\{z\}} = P'\), it is easy to
see that
\[
\tag{1.1}
p'(x, y) = p(x, y) + p(x, z)n(z)p(z, y), x \in X, y \in S,
\]
where \(n(z) = \sum_{n=0}^{\infty} p^n(z, z) = 1/(1 - p(z, z))\). According to formula (1.1), each row-
vector of the new stochastic matrix \(P'\) is a linear combination of two rows of \(P\) (with
the \(z\)-column deleted). This transformation corresponds formally to one step of the
Gaussian elimination and requires \(O(n^2)\) operations. We call such a transformation
of a matrix an iteration. If \(|D| = k\) then the transition matrix \(P_D\) can be obtained in
\(k\) iterations.

Censored MCs have had numerous applications in Probability Theory and Linear
Algebra, see e.g. [11] and [4]. A relatively recent method for recursively calculating
many important characteristics of MCs based on formula (1.1), is the so called state
reduction (SR) approach. This was initiated by the papers [1] and [6], where the so-
called GTH/S algorithm to calculate the invariant distribution for an ergodic Markov
chain was introduced. See also [5], where similar ideas were analyzed mainly from the
algebraic point of view.

Briefly this approach can be described as follows. If an initial Markov model
\(M = (X, P)\) is finite, \(|X| = n\), and only one point is eliminated each time, then a
sequence of stochastic matrices \((P_k), k = 1, \ldots, n - 1\) can be calculated recursively
on the basis of formula (1.1) or (2.4) below. Such a sequence of stochastic matrices
provides an opportunity to calculate many characteristics of the initial Markov model
\(M\) recursively starting from some reduced model \(M_s, 1 < s \leq n\).

Another application of formula (1.1) in the area of the Optimal Stopping (OS)
of MCs was started in [8], where the so called the State Elimination Algorithm was
introduced. According to this algorithm all points that do not belong to an optimal
stopping set are eliminated one by one, or more generally in the countable case at some
steps a subset may be eliminated. The order in which states are eliminated is defined
by some auxiliary procedure. Another algorithm based on the State Elimination
algorithm was used to calculate the Generalized Gittins index in [9].

Recently in [10] we presented an algorithm for finding an optimal strategy and the
value function for a Markov Decision Process (MDP) model where at each moment of
discrete time a decision maker (DM) can apply one of three possible actions - continue
when MC evolves according to the transition matrix \( P \), quit when the evolution of
a MC is stopped, and restart when the MC is moved to one of a finite number \( m \)
of fixed "restart" points. A decision at state \( x \) brings a corresponding reward,
positive or negative, \( c(x) \), \( q(x) \) or \( r_i(x), i = 1, \ldots, m \). The goal of a DM is to maximize
the total expected discounted reward. Such a model is a generalization of a model of
Katehakis and Veinott in [2], where a restart to a unique point was allowed without
any fee and the quit action was absent. Both models are related to the well-known
Gittins index. For the case \( m = 1 \), a recursive algorithm to solve this model by
performing \( O(n^3) \) operations was proposed. An important part of this algorithm
is a sequence of recursive steps when a transition matrix is transformed. It turns
out that at some steps the points are eliminated but on other steps they need to
be included (inserted) back. Assume, for example, that three points, \( z_1, z_2 \) and \( z_3 \),
are subsequently eliminated. Reversing formula (1.1), it is easy to restore point \( z_3 \).
This requires one iteration step. What if we wish to restore point \( z_1 \), i.e. to have
only points \( z_2 \) and \( z_3 \) be eliminated? At first glance it seem that we either have to
keep the matrix \( P_1 \) in memory and eliminate points \( z_2 \) and \( z_3 \) or restore (in three
iteration steps) points \( z_3, z_2 \) and \( z_1 \) and then eliminate points \( z_2 \) and \( z_3 \). It turns out
that we can insert point \( z_1 \) in one iteration if we will keep in memory a
nonstochastic matrix \( W_D \) similar to matrix \( P_D \). This matrix is also obtained by iterations. This
result was given by Theorem 3 in [10]. The proof was somewhat tedious, part of it
was given in Appendix and we wrote that “we fail to find a simpler proof though one
likely exists”. The main goals of this note are: first, to describe this new operation of
insertion and to present corresponding formulas, and second, to give simpler, shorter
and more transparent proof of this theorem, based on two new lemmas, keeping the
formulation of the theorem the same. We also note that these “new elimination” steps
allow us as a byproduct to obtain a recursive algorithm to calculate a fundamental
matrix \( N = N_D = (I - Q)^{-1} \) corresponding to any transient MC with substochastic
matrix \( Q \). In Section 3 we prove this theorem and we discuss the transformation of
transition matrices of MCs under elimination and insertion. At the end we give two
small numerical examples.

2. Censored MC and Elimination. An important and traditional tool for
the study of Markov chains (MCs) is the notion of a Censored (Embedded) MC. Two
operations on stochastic and related matrices are introduced in this section. They
serve as building blocks for the algorithm in [10].

A pair \( M = (X, P) \), where \( X \) is a state space and \( P \) is a stochastic matrix is
called Markov model. Let us assume that a Markov model \( M = (X, P) \) is given and
\( D \subset X, S = X \setminus D \). Then the matrix \( P = \{p(x, y)\} \) can be decomposed as follows

\[
P = \begin{bmatrix}
Q & T \\
R & P_0
\end{bmatrix},
\]

where the substochastic matrix \( Q \) describes the transitions inside of \( D \), \( P_0 \) describes
the transitions inside of \( S \) and so on. Let us introduce the sequence of Markov
times \( \tau_0, \tau_1, \ldots, \tau_n, \ldots \) where \( \tau_0 = 0 \), and \( \tau_n, n \geq 1 \) are the times of first, and so on, return of the MC \((Z_n)\) to the set \( S \), i.e., \( \tau_{n+1} = \min\{k > \tau_n, Z_k \in S\}\). Let us consider the random sequence \( Y_n = Z_{\tau_n}, n = 0, 1, 2, \ldots \) and assume that \( Z_0 = x \in S \). The strong Markov property and standard probabilistic reasoning imply the following basic lemma of the SR approach which should probably be credited to Kolmogorov and Doeblin.

**Lemma 2.1 (Elimination Lemma).** (a) The random sequence \((Y_n)\) is a Markov chain in a model \( M'_D = (S, P'_D) \), where

(b) the transition matrix \( P'_D = \{p'(x,y), x,y \in S\} \) is given by the formula

\[
P'_D = P_0 + RU = P_0 + RNT.
\]

Here \( U \) is the matrix of the distribution of the MC at the time of the first visit to \( S \) starting from \( x \in D \) and \( N = N(D) \) is the fundamental matrix for the substochastic matrix \( Q \), i.e. \( N = \sum_{n=0}^\infty Q^n = (I - Q)^{-1} \), where \( I \) is the \(|D| \times |D|\) identity matrix. This representation is given, for example, in the classical text [3].

The matrix \( N = N(D) \) has the following well-known probabilistic interpretation, \( N = \{n(x,y), x,y \in D\} \), where \( n(x,y) = E_x \sum_{n=0}^\infty I_y (Z_n) \), and \( \tau \) is the time of the first visit to \( S \), i.e. \( \tau = \min(n \geq 0 : Z_n \in S) \). Thus \( n(x,y) \) is the expected number of visits to \( y \) starting from \( x \) until \( \tau \). The matrix \( N \) also satisfies the equalities

\[
N = I + QN = I + NQ.
\]

\( MC (Y_n) \) is called an embedded (censored) MC.

An important case is when the set \( D \) consists of one nonabsorbing point \( z \). In this case formula (2.2) takes the form (1.1), where \( n(z) = 1/(1-p(z,z)) \), is a "fundamental matrix".

In Lemma 2.1, the Markov model \( M'_D \) has a reduced state space \( S \). Sometimes, it is more convenient to have all stochastic matrices of equal full size. Then we consider a MC \((Y_n)\) with initial points \( x \) not only in \( S \) but also in \( D \), though after the first step the MC \((Y_n)\) is always in \( S \). Lemma 2.1 remains true but now we obtain a Markov model \( M_D = (X, P_D) \). In addition to (2.2) for \( x,y \in S \), we have the equality \( T + QNT = (I + QN)T = NT \) for \( x \in D, y \in S \). The last equality is true by (2.3). Thus instead of (2.2) we have the following full size transition matrix

\[
P_D = \begin{bmatrix}
0 & NT \\
0 & P_0 + RNT
\end{bmatrix}.
\]

Note that the rows of matrix \( P_D \) give the distribution of MC \((Z_n)\) at the time \( \tau_1 \) of the first return to set \( S \), i.e. \( \tau_1 = \min\{n > 0 : Z_n \in S\} \) and \( P(Z_{\tau_1} = y) = P_D(Y_1 = y) \) when \( x \in X, y \in S \). For \( x \in D \), the moment of the first return coincides with the moment of the first visit and this distribution is given by submatrix \( NT \). For the points from \( x \in S \) the corresponding distribution is given by submatrix \( P_0 + RNT \). If \( D = \{z\} \), i. e. when state \( z \) is eliminated, formula (2.4) is replaced by the one-state elimination formula, written here for columns, \((P_{|z|} = P')\),

\[
p'(\cdot, z) = 0, \quad p'(\cdot, y) = p(\cdot, y) + p(\cdot, z) \frac{p(z,y)}{1-p(z,z)}, y \neq z.
\]

We say that matrix \( P' \) is obtained from \( P \) in one iteration. Thus matrix \( P_D \) can be calculated directly by (2.4) or recursively using formula (2.5) in \(|D|\) iterations.
3. Elimination vs Insertion. Suppose that the set $D = \{ z_1, z_2, \ldots, z_k \}$ is eliminated in an initial model $M$, and $P_D$ is the corresponding matrix obtained recursively by formula (2.5). Let $z \in D$, say $z = z_1$. How can one obtain the transition matrix $P_{D \setminus z}$, i.e. when only the points $z_2, \ldots, z_k$ are eliminated? Of course we can obtain this matrix starting from an initial matrix $P$ and eliminating these points, performing $k - 1$ iterations. Is there a way to obtain this matrix in just one iteration?

The answer for this rhetorical question is Yes. In this case we say that point $z$ is inserted (restored). To do this, initially, instead of the stochastic matrices $P_1, P_2, \ldots, P_k = P_D$, we must calculate recursively similar but different nonstochastic (!) matrices $W_1, W_2, \ldots, W_k = W_D$. We do this by applying the second part of formula (2.5) to all columns, including previously eliminated states, i.e. when state $z$ is eliminated using the formula

\begin{equation}
(3.1) \quad w'(\cdot, y) = w(\cdot, y) + w(\cdot, z) \frac{w(z, y)}{1 - w(z, z)}, \ y \in X.
\end{equation}

Let us show immediately that this transformation of matrix $W$ into $W'$ can be reversed, i.e. the following statement holds

**Lemma 3.1.** If matrix $W'$ is obtained from matrix $W$ by elimination formula (3.1) then matrix $W$ can be obtained from matrix $W'$ by insertion formula

\begin{equation}
(3.2) \quad w(\cdot, y) = w'(\cdot, y) - w'(\cdot, z) \frac{w'(z, y)}{1 + w'(z, z)}, \ y \in X.
\end{equation}

**Proof.** Using formula (3.1) for $x = z$ we obtain the equality $w'(z, y) = w(z, y)/(1 - w(z, z))$ and hence formula (3.1) can be written as

\begin{equation}
(3.3) \quad w(\cdot, y) = w'(\cdot, y) - w'(\cdot, z) w'(z, y), \ y \in X.
\end{equation}

Applying this formula for $y = z$ we obtain $w(\cdot, z) = w'(\cdot, z)/(1 + w'(z, z))$ and hence formula (3.3) can be written as (3.2). $\blacksquare$

Note that by the definition of matrices $W_D$ and $P_D$ their columns for $y \notin D$ coincide but the matrix $P_D$ has zero columns for $y \in D$. The interpretation of the nonzero columns in $W_D$ is given in Theorem 3.4.

Let us introduce the matrix $N^+ = N^+(D) = \{ n^+(x, y) | D \}$, $x \in X, y \in D$, where $n^+(x, y)\in D$ is the expected number of visits of a MC $(Z_n)$ to state $y$ after the initial moment until the moment of first return (visit) to set $S$. As with matrix $N = N(D)$ we usually will skip $D$. Note the differences between matrices $N^+$ and $N = \{ n(x, y) \}$, $x, y \in D$: first, $N^+$ is the $|X| \times |D|$ matrix, $N$ is the $|D| \times |D|$ matrix; second, $n^+(x, y)$ counts the number of visits to $y$ after the initial moment, while as $n(x, y)$ including the initial moment. We also have obvious equalities: if $x, y \in D$ and $x \neq y$ then $n^+(x, y) = n(x, y)$, and $n^+(x, x) = n(x, x) - 1$. If $y \in D, x \notin D$ then $n^+(x, y) = n(x, y) = \sum_{z \in D} p(x, z) n(z, y)$.

To describe the structure of matrix $N^+$ it is convenient to introduce also an auxiliary matrix, $P(D)$. It consists of the first $D$ columns of matrix $P$, see (2.1), i.e. has dimension $|X| \times |D|$ and contains blocks $Q$ and $R$. 

4
**Lemma 3.2.** (One point Lemma). Let $D \subseteq X$, and $N,N^+,P(D)$ are matrices defined above. Then

a) matrix $N^+ = P(D)N$, i.e. 

\[(3.4)\]

\[
N^+ = \begin{bmatrix} QN \\ RN \end{bmatrix} = \begin{bmatrix} N - I \\ RN \end{bmatrix}.
\]

b) the columns of $N^+$ can also be obtained by formula

\[(3.5)\]

\[
n^+(\cdot,y|D) = \frac{p_{D\setminus y}(\cdot,y)}{1 - p_{D\setminus y}(y,y)} = p_{D\setminus y}(\cdot,y)n(y,y|D), y \in D.
\]

First, we show that it is sufficient to consider only the case when $D$ contains only one point, $D = \{y\}$. The explains the name for this lemma. The reason is that all points in $D$ except $y$ one can be eliminated without changing $n(y,y|D)$ or $n^+(\cdot,y|D)$.

More precisely

**Proposition 3.3.** Let $D \subseteq X, y \in D$. Then

\[
n(\cdot,y|D) = n_{D\setminus y}(\cdot,y), n^+(\cdot,y|D) = n^+_{D\setminus y}(\cdot,y).
\]

**Proof.** Intuitively this statement is almost obvious: the expected number of visits to state $y$ starting in $y$ before exit to $S$ remains the same if model $M$ is transformed to model $M_{D\setminus y}$. The strict proof of this and more general statement about trajectories in an initial and reduced model was given in [8].

Now we can prove lemma 3.2. Let $D = \{y\}$. Then $p_{D\setminus y}(\cdot,y) = p(\cdot,y)$ and $n(y,y)$ is the expectation of a geometric random variable with the $P(\text{success}) = P(\text{exit from } D) = 1 - p(y,y)$ and hence $n(y,y) = 1/(1 - p(y,y))$. Correspondingly $n^+(y,y) = n(y,y) - 1 = p(y,y)/(1 - p(y,y)$ and $n^+(x,y) = p(x,y)n(y,y)$.

Before formulating our main theorem, recall that both matrices $W_D$ and $P_D$ are defined recursively, starting from the same stochastic matrix $P$ and their columns for $y \notin D$ are the same, i.e. $P_D$ is a part of $W_D$. The rows of $P_D$ have a simple probabilistic interpretation as the distribution of MC (Z_n) at the moment of the first return to set $S = X \setminus D$. Thus there is no question whether the matrix $P_D$ for $D = \{z_1,...,z_k\}$ depends on the order of states in $D$, it does not. For the $y$-columns of matrix $W_D$ for $y \in D$ this is initially an open question, but formula (3.6) (or (3.8)) shows that as with $P_D$ the order is irrelevant. Let us denote by $P^0_D$ a matrix which consists only of non zero columns of matrix $P_D$, see (2.4), and thus has dimension $|X| \times |S|$ and contains blocks $NT$ and $P_0 + RNT$, and denote by $p_{D\setminus y}(\cdot,y)$ the columns of matrix $P_{D\setminus y}$.

**Theorem 3.4** (Insertion Theorem). Let $D \subset X, S = X \setminus D$, the transition matrix $P$ is decomposed as in (2.1) and $N,N^+,P_D^0$ are matrices defined above. Then

a) the matrix $W_D = [N^+|P_D^0]$, i.e. the first $|D|$ columns of $W_D$ for $y \in D$ coincide with columns of matrix $N^+ = N^+(D)$ and the remaining columns, for $y \in S$ coincide with columns of matrix $P_D^0$, i.e.

\[(3.6)\]

\[
W_D = \begin{bmatrix} QN & NT \\ RN & P_0 + RNT \end{bmatrix}.
\]
Let the right side of formula (3.9) with
and correspondingly for point
G, can be also obtained in one iteration.

Corollary 3.5. The equality (3.7) (the y-th column of matrix \( W_D \) for \( y \in D \)), can be described also by the formula

\[
(3.8) \quad w_D(y, \cdot) = n^+(\cdot, y|D) = \frac{p_{D\setminus y}(\cdot, y)}{1 - p_{D\setminus y}(y, y)}, \quad y \in D,
\]

this corollary follows immediately from point a) of Theorem 3.4 and formula (3.5).

Before to prove the theorem note that by formula (2.3) the submatrix
this corollary follows immediately from point a) of Theorem 3.4 and formula (3.5).

We can prove theorem 3.4. To prove point a) is equivalent to prove (3.8). Similarly to the proof of (3.2), using Proposition 3.3, it is sufficient to

\[
(3.9) \quad n^+(\cdot, y|G) = n^+(\cdot, y|D) + n^+(\cdot, z|G)n^+(z, y|D).
\]

Proof. Similarly to the proof of (3.2), using Proposition 3.3, it is sufficient to consider only the case when \( D = \{y\}, G = \{y, z\} \), thus the name of this lemma. In
this case \( p_{D\setminus y}(\cdot, \cdot) = p(\cdot, \cdot) \). The proofs of (3.9) for all three possible cases \( x = z, \quad x = y, \quad x \notin \{y, z\} \) are very similar so we consider only the more difficult case \( x = z \). In
this case \( n^+(z, y) = n(z, y) \) and formula (3.5) applied separately for the sets \( D \) and
\( G \), and correspondingly for point \( x = z \) and columns \( y \) and \( z \), implies the equalities

\[
n^+(z, y|D) = p(z, y)n(y, y)|D) = \frac{p(z, y)}{1 - p(y, y)}, \quad n^+(z, z|G) = \frac{p_D(z, z)}{1 - p_D(z, z)}.
\]

Then the right side of formula (3.9) with \( x = z \) can be written as equality

\[
(3.10) \quad p(z, y)n(y, y)|D)/(1 + n^+(z, z|G)) = \frac{p(z, y)}{(1 - p(y, y))(1 - p_D(z, z))}.
\]

Thus to prove lemma we need to prove that the left side of (3.9), i.e. \( n^+(z, y|G) \equiv n(z, y) \) coincides with (3.10).

Using the second equality in (2.3) the left side of (3.9) can be represented (skipping
\( G \)) as \( n(z, y) = n(z, y)p(y, y) + n(z, z)p(z, y) \) and hence \( n(z, y) = n(z, z)p(z, y)/(1 - p(y, y)) \). By point b) of (3.2) we have \( n(z, z) \equiv n(z, z|G) = 1/(1 - p_D(z, z)) \). Thus
\( n(z, y) \) coincides with the right side of (3.10). Lemma is proved.

Now we can prove theorem 3.4. To prove point a) is equivalent to prove (3.8). For the case when \( D = \{y\} \) by formula (3.1) with \( W' = P_D, W = P \) and \( z = y \) we have

\[
w_D(\cdot, y) = p(\cdot, y) + p(\cdot, y)\frac{p(y, y)}{1 - p(y, y)}, \quad y \in D,
\]
Thus our goal is to prove the first equality in the following formula

\[ W_D = W_1, W_G = W_2, p_{D\setminus y}(\cdot, y) = s_1(\cdot, y), p_{G\setminus y}(\cdot, y) = s_2(\cdot, y). \]

Thus our goal is to prove the first equality in the following formula

\[ w_2(\cdot, y) = n^+(\cdot, y|G) = \frac{s_2(\cdot, y)}{1 - s_2(y, y)}, y \in G, \]

where the second equality holds by (3.5).

If formula (3.8) holds for a set \( D \subseteq X \), then two equivalent statements are true

\[ w_1(\cdot, y) = n^+(\cdot, y|D) = \frac{s_1(\cdot, y)}{1 - s_1(y, y)}, s_1(\cdot, y) = \frac{w_1(\cdot, y)}{1 + w_1(y, y)}, y \in D. \]

The second formula follows from the first one if we apply the first formula for \( x = y \) to obtain the equality \( 1 - s_1(y, y) = 1/(1 + w_1(y, y)) \).

For \( y = z \) formula (3.11) can be checked directly: formula (3.1) with \( W' = W_2, W = W_1 \), state \( z \) eliminated, and with \( y = z \) implies \( w_2(\cdot, z) = w_1(\cdot, z)/(1 - w_1(z, z)) \); since \( z \notin D \) by definition of \( s_2 \) we have \( s_2(\cdot, z) = p_{D\setminus z}(\cdot, z) = p_{D}(\cdot, z) = s_1(\cdot, z) \), and by definition of \( s_1 \) we have \( s_1(\cdot, z) = w_D(\cdot, z) \).

For \( y \in D \) by definition of \( W_2 \), i.e. by formula (3.1) with \( W' = W_2, W = W_1 \), state \( z \) eliminated, using the first formula in (3.12), the equalities \( w_D(\cdot, z) = p_D(\cdot, z), w_D(z, z) = p_D(z, z) \) (since \( z \notin D \)), and the first formula in (3.12) for \( x = z \), we have

\[ w_2(\cdot, y) = n^+(\cdot, y|D) + \frac{p_D(\cdot, z)}{1 - p_D(z, z)} n^+(z, y|D), y \in D, z \notin D. \]

By formula (3.5) in Lemma 3.2 applied to the case when set \( D \) is replaced set \( G \) and \( y \) is replaced by \( z \), we have

\[ \frac{p_D(\cdot, z)}{1 - p_D(z, z)} = n^+(\cdot, z|G). \]

Therefore applying lemma 3.7, i.e. formula (3.9) the right side of (3.13) coincides with \( n^+(\cdot, z|G) \), i.e. formula (3.11) is proved.

**Remark 3.8.** In the MDP problems with a current reward function \( c(x) \) it is possible to introduce a transformation of the cost function \( c(x) \) (or any function \( f(x) \)) defined on \( X \) into the cost function \( c'_D(x) \) under transition from model \( M \) to model \( M'_D \) or correspondingly into function \( c_D(x) \) under transition to model \( M_D \). This formula was used first in [7] in the context of MDP. A corresponding transformation can be obtained for the insertion situation as well.

**4. Example.** Example shows elimination and insertion for the Markov chain with 5 states. Initially, the first 3 states are eliminated one by one, then states #2 and #3 are inserted.

All calculations are performed with double floating-point precision (16 decimal significant digits), the values in tables are rounded to 4 digits after decimal point.
Table 4.1
The elimination of states #1, #2, #3.

<table>
<thead>
<tr>
<th>0.4000</th>
<th>0.3000</th>
<th>0.2000</th>
<th>0.1000</th>
<th>0.0000</th>
<th>0.6667</th>
<th>0.5000</th>
<th>0.3333</th>
<th>0.1667</th>
<th>0.0000</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>0.3000</td>
<td>0.2000</td>
<td>0.1000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.5000</td>
<td>0.6500</td>
<td>0.2000</td>
<td>0.1500</td>
<td>0.0000</td>
</tr>
<tr>
<td>0.2000</td>
<td>0.4000</td>
<td>0.1000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.3333</td>
<td>0.4000</td>
<td>0.4667</td>
<td>0.1333</td>
<td>0.0000</td>
</tr>
<tr>
<td>0.1000</td>
<td>0.1000</td>
<td>0.5000</td>
<td>0.2000</td>
<td>0.0000</td>
<td>0.1667</td>
<td>0.1500</td>
<td>0.1333</td>
<td>0.5167</td>
<td>0.2000</td>
</tr>
<tr>
<td>0.1000</td>
<td>0.2000</td>
<td>0.2000</td>
<td>0.2000</td>
<td>0.4000</td>
<td>0.1667</td>
<td>0.1500</td>
<td>0.2333</td>
<td>0.2167</td>
<td>0.4000</td>
</tr>
</tbody>
</table>

1.3810 1.4286 0.6190 0.3810 0.6667 0.5000 0.3333 0.1667 0.0000 1.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000
1.4286 1.8571 0.5714 0.4286 0.5000 0.6500 0.2000 0.1500 0.0000 1.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000
0.9048 1.1429 0.6952 0.3048 0.3333 0.4000 0.4667 0.1333 0.0000 1.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000
0.3810 0.4286 0.2190 0.2810 0.1667 0.1500 0.1333 0.5167 0.2000 1.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000
0.3810 0.4286 0.3190 0.2810 0.1667 0.1500 0.2333 0.2167 0.4000 1.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000

Table 4.2
The insertion of states #2 and #3.

<table>
<thead>
<tr>
<th>0.8750</th>
<th>0.7500</th>
<th>0.6250</th>
<th>0.2500</th>
<th>0.0000</th>
<th>0.6667</th>
<th>0.5000</th>
<th>0.3333</th>
<th>0.1667</th>
<th>0.0000</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>0.6250</td>
<td>0.8000</td>
<td>0.3750</td>
<td>0.2000</td>
<td>0.0000</td>
<td>0.5000</td>
<td>0.6500</td>
<td>0.2000</td>
<td>0.1500</td>
<td>0.0000</td>
</tr>
<tr>
<td>0.6250</td>
<td>0.7500</td>
<td>0.8750</td>
<td>0.2500</td>
<td>0.0000</td>
<td>0.3333</td>
<td>0.4000</td>
<td>0.4667</td>
<td>0.1333</td>
<td>0.0000</td>
</tr>
<tr>
<td>0.2500</td>
<td>0.2500</td>
<td>0.5500</td>
<td>0.2000</td>
<td>0.0000</td>
<td>0.1667</td>
<td>0.1500</td>
<td>0.1333</td>
<td>0.5167</td>
<td>0.2000</td>
</tr>
<tr>
<td>0.3125</td>
<td>0.3250</td>
<td>0.4375</td>
<td>0.2750</td>
<td>0.0000</td>
<td>0.1667</td>
<td>0.1500</td>
<td>0.2333</td>
<td>0.2167</td>
<td>0.4000</td>
</tr>
</tbody>
</table>

The original transition matrix and the result of consecutive elimination of states #1, #2, #3 is shown in Table 4.1. The Table 4.2 shows result of consecutive insertion of states #2 and #3. Bold columns correspond to non-zero columns in $P_D$, these columns form a stochastic matrix. Note, that the last matrix in Table 4.2 coincides with second matrix in Table 4.1 (with eliminated state #1), which is expected.

REFERENCES