THE STUDY OF "LOOP" MARKOV CHAINS

by

Perry L. Gillespie Jr.

A dissertation submitted to the faculty of
the University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Applied Mathematics

Charlotte

2011

Approved by:

Dr. Stanislav Molchanov

Dr. Isaac Sonin

Dr. Joe Quinn

Dr. Terry Xu

ABSTRACT

PERRY L. GILLESPIE JR.. The study of "loop" markov chains.
(Under the direction of DR. STANISLAV MOLCHANOV AND ISAAC SONIN)

The purpose of my research is the study of "Loop" Markov Chains. This model contains several loops, which could be connected at several different points. The focal point of this thesis will be when these loops are connected at one single point. Inside each loop are finitely many points. If the process is currently at the position where all the loops are connected, it will stay at its current position or move to the first position in any loop with a positive probability. Once within the loop, the movement can be deterministic or random. We'll consider the Optimal Stopping Problem for finding the optimal stopping set and the Spectral Analysis for our "Loop" system. As a basis, we'll start with the Elimination Algorithm and a modified version of the Elimination Algorithm. None of these algorithms are efficient enough for finding the optimal stopping set of our "Loop" Markov Chain; therefore, I propose a second modification of the Elimination Algorithm, the Lift Algorithm. The "Loop" Markov Chain will dictate which algorithm or combination of algorithms, would be appropriate to find the optimal set. The second part of this thesis will examine the Analysis of the Covariance Operator for our "Loop" Markov Chain system.

ACKNOWLEDGMENTS

I would like to show my gratitude to these two distinguished and talented mathematicians, my co-advisors, Dr. Isaac Sonin and Stanislav Molchanov. They are astute in their area and very proficient in others. I thank you for your patience and encouragement throughout my tenure here at University of North Carolina at Charlotte. Thank you Dr. Sonin for granting me permission to use your paper "Markov Chains, Optimal Stopping and Applied Probability Models" as a model for this dissertation. Without you, I would not been able to achieve this degree. Thank for Dr. Molchanov for your support, time and patience in helping me to achieve this goal. I would also like to think Dr. Joe Quinn for his assistance with creating these excel programs and his valuable comments on this presentation. Thank you Dr. Arvin for your words of wisdom and encouragement.

To the Fayetteville State University, the administration, College of Arts and Science, Math and Computer Science department, I thank you for all your support. I appreciate the opportunity that you allow me to work on and obtain this terminal degree. To my former and current colleagues, thank you for the confidences you had in me to pursue and complete this objective.

To Pastor and Lady Swinney and the New Bethel Church and Pastor James Chruchwell and the Emmanuel Church of Christ family, thank you for your prayers and support. The Bible tells us that "Life and Death lies in the power of the tongue". Thank you for speaking this day into existence, i.e. to be called Dr. Perry L. Gillespie, Ph D.

To my parents, brothers, sisters, and family members, thank you for being an example to me. Because of you blazing a path that I could follow, it gives me the confidence to create others paths for those coming after me (the Lord being me help).

To my lovely wife, Aarona of whom I love dearly, thank you for your inspiration, encouraging words, and support to complete this task that was set before me. Remember, this is our accomplishment.

Finally, this thesis would not have been possible without God's strength and wisdom. It was in His will that I complete this assignment; I'm glad I was obedience to Him. To God be the glory for the things He has done in MY LIFE.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Many decision makers use a certain type of random process to make decisions. The main feature of this type of process is that it's "memoryless" of the past. Such a process is called a Markov Process. First proposed by Andrey Markov in 1906 and modified by Andrey Kolmogorov in 1936, Markov Processes are Stochastic Processes (which are defined as a family of random variables) with the Markov Property. The Markov Property implies that the next transition depends only on the current position and is independent of the past. In other words, the present state contains all the information needed to determine any future movement of the process. Markov Processes are used in various disciplines, such as Biological, Physical, and Social Sciences along with Business and Engineering, just to name a few. Markov Chains are special cases of Markov Processes.

This research addresses the topic of "Loop" Markov Chains. Let's look at a brief description of the system. A loop is defined as a finite set of points contained in a circle. The state space of our Markov Chains consist of finitely many loops joined at a single point called the orign. Inside each loop, there are finitely many positions called states. If the process is currently at the origin, then it will stay or move to the first state on some loop with a positive probability. Transitions within the loops can be either deterministic or random.

The next chapter starts with the framework and an introduction of terms for Markov Chains. In particular, we'll address the following question: "After observing the movement of the chain, does a pattern, if one exists, converge?" Furthermore, the author will examine the different types of chains. The Fundamental Matrices for the Transient and Ergodic Chains will be explored, and we'll look at why these matrices are different. I'll state some basic properties that utilize these fundamental matrices.

There will be a review of some basic facts about the Central Limit Theorem, Hilbert Space, and its application to Markov Chains. We'll discuss the Homological (Poisson)

Equation and later find the invariant distribution for our "Loop" Markov Chain in L^2-space (deterministic only). The solution for our "Loop" Markov Chain system will be given in chapter 5.

Markov Decision Problems employ dynamical models based on well-understood stochastic processes and performance criteria based on established theory in operations research, economics, combinatorial optimization, and social science (Putterman 1994). A Markov Decision Problem is a Markov Decision Process together with a performance criterion. A Markov Decision Process is a controlled stochastic process satisfying the Markov Property with costs assigned to each state transition. The process can be described as follows: Let's assume the chain starts at an initial state. Each state has a set of actions assigned to them. The decision maker chooses any action from the action set for the current state. The process moves randomly to the next state and a reward is collected. The process continues until the decision maker chooses the option "quit". Chapter 3 will give a formal description of this process.

The Optimal Stopping Problem is a special case of the Markov Decision Problem. If the action space for each state contains only two actions (continue, quit), then the Decision Problem reduces to the Optimal Stopping Problem. The theory of optimal stopping is concerned with the problem of choosing a time to take a given action based on sequentially observed random variables in order to maximize an expected payoff or to minimize an expected cost. A famous example of an Optimal Stopping Problem is the Secretary Problem.

There are various methods used to solve optimal stopping problems for Markov Chains. Numerical Solution, Linear Programming and Dynamic Programming are just a few examples of these methods. The first algorithm, proposed by Isaac Sonin, that will be used to solve Optimal Stopping Problems is the "Elimination Algorithm". After giving a detailed description of this method, we will look at two examples implementing this algorithm.

Even though the "Elimination Algorithm" works well, it poses some limitations with our "Loop" model. Chapter 4 begins with a modification of this algorithm. The modified version of the Elimination Algorithm, called the Full Size Matrix Algorithm, has some advantages over the "Elimination Algorithm", but it is not efficient enough for our "Loop" Markov Chains model; therefore, I propose the "Lift" Algorithm. We'll see that this new

modification along with the first modified version of the Elimination Algorithm is a better fit for our problem. One goal of this thesis is to illustrate the combination of these two algorithms on the "Loop" Markov Chain Model.

The author was also responsible for creating Excel programs to aid in the computation of the 3 algorithms mentioned above. All examples in this thesis used one of these programs to generate the optimal stopping set.

The second part of my thesis is devoted to the Spectral Analysis of the Covariance Operator. I'll discuss the Doëblin Condition and its connection to the invariant distribution. Also, by applying the classical Central Limit Theorem for the Martingale-difference, we'll obtain the Covariance Operator. The final chapter describes this Covariance Operator and how it applies to our "Loop" Markov Chain Model.

CHAPTER 2: BASIC NOTATION AND MAIN PROPERTIES FOR MARKOV CHAINS

2.1 Basic Definitions and Classification of States

A Stochastic Process, in discrete time, is a sequence of random variables. The stochastic process $\{Z_n; n = 0, 1, \ldots\}$ with discrete state space S is a homogeneous Markov Chain if the following holds for each $j \in E$ and $n = 0, 1, \ldots$

$$P\{Z_{n+1} = j | Z_n = i, \ldots, Z_1 = i_1, Z_0 = i_0\} = P\{Z_{n+1} = j | Z_n = i\} = p_{ij} \qquad (2.1)$$

for any states $i_0, \cdots, i_n$ in the state space and $i_n = i$.

Matrix $P = \{p_{ij}\}$ is called the one-step Transition Probability Matrix. The elements in this square matrix P must satisfy the following conditions: The p_{ij}s are non-negative and $\sum_j p_{ij} = 1$.

Markov Chains are generally depicted using graphs. Circles and/or ovals represent states. Directed arrows represent the one-step transitions. Often these directed arrows are labeled by their transition probability values (see figure 2.1).

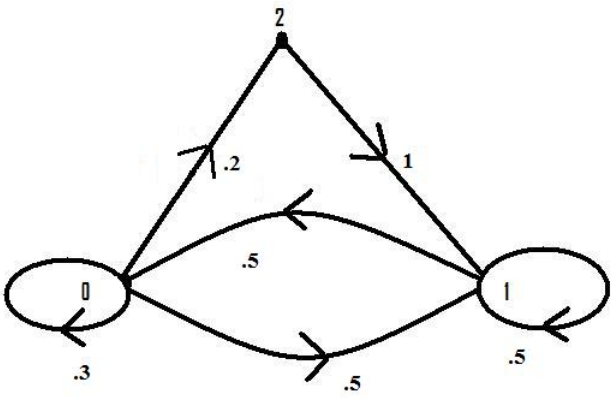


Figure 2.1: Transition Graph

The transition matrix of figure 2.1 is $P = \begin{bmatrix} .3 & .5 & .2 \\ .5 & .5 & 0 \\ 0 & 1 & 0 \end{bmatrix}$.

A n-step transition probability, denoted as $P^{(n)} = \left\{p_{ij}^{(n)}\right\}$, is the probability that a process in state i will be in state j after n additional transitions. It is well known that the discrete homogeneous Markov Chains n-step become the n-th power, i.e., $P^{(n)} = P^n$.

Notice we can compute $P^{(n)}$ from the following recursive formula, called the Chapman-Kolmogorov Equation:

$$p_{ij}^{(n+m)} = \sum_{allk} p_{ik}^{(n)} p_{kj}^{(m)} \tag{2.2}$$

.

The matrix notation of equation 2.2 is:

$$P^{n+m} = P^n \cdot P^m \tag{2.3}$$

Let's start by introducing some important definitions concerning individual states. A state j is accessible (reachable) from state i if there exists a path from i to j, which is denoted by $i \to j$ and $P\{i \to j\} > 0$. If i is accessible from j and j is accessible from i, then i and j communicate, denoted by $i \leftrightarrow j$. A communicating class is **closed** if the probability of leaving the class is zero. States are recurrent, if the process is guaranteed, with a probability of 1, to return to state j after it leaves. A recurrent state j is called positive recurrent, if starting at state j, the expected time until the process returns to state j is finite. If this expected time is infinite, then the recurrent state is called null recurrent. States are considered transient if there is a positive probability that the Markov Chain will never return to state j after it leaves j. Please note, a transient state can be visited only a finite number of times. State i is an absorbing state if the path from i to i has a probability of 1. When a process enters an absorbing state, it will never leave this state; it's trapped. If all states communicate with each other in a class, then it's called an irreducible class. A Markov Chain is said to be a regular Markov Chain if for some power, n, the transition matrix, P^n, has only positive elements. Finally, the period of state j is

defined as $d(j) = \gcd\left\{n \geq 1 : p_{jj}^{(n)} > 0\right\}$, the gcd (greatest common divisor) of lengths of cycles through which it is possible to return to j. In other words, if after leaving state j it's possible to return in a finite number of transitions that is a multiple of the integer $k > 1$. If $d(j) = 1$, then the state j is aperiodic: otherwise it's **periodic**. A Markov Chain is called an ergodic if it's positive recurrent and aperiodic. Some textbooks refers to Ergodic Markov Chains as Irreducible.

Let's consider figure 2.1. We have only one communicating classes, which is $\{0, 1, 2\}$. The class is closed, which means you cannot exit from this class. Because we can return to every state in $\{0, 1, 2\}$, it is also a recurrent class. Since gcd for the number of transitions from i back to i is 1, this example is aperiodic.

2.2 Transient and Embedded Markov Chains

Now, let's consider the general construction of the absorbing(transient) Markov Chains (see figure 2.2).

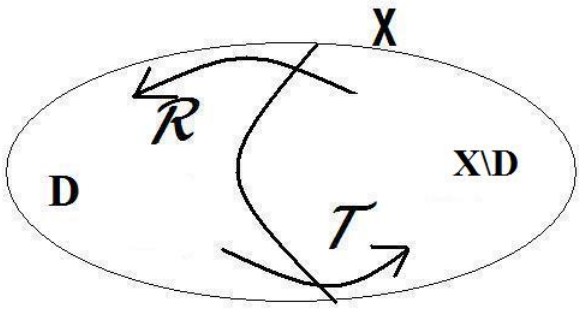


Figure 2.2: Fundamental Matrix Diagram

Formally speaking, let D be a subset of X and $M = (X, P)$ be a Markov model which will visit D with a probability of 1 from any point inside $X \backslash D$. The transition matrix P is

$$P = \begin{bmatrix} Q & T \\ R & P' \end{bmatrix} \tag{2.4}$$

where $Q = \{p(i, j), i, j \in D\}$ is a substochastic matrix which describes the transition

inside of D, P' is the transition inside of $X\backslash D$, R is the transition probability from $X\backslash D$ to D, and T is the transition probability from D to $X\backslash D$.

The classical way of calculating N, which was proposed by Kemeny [13], is as follows.

$$N = \sum_{n=0}^{\infty} Q^n = (I - Q)^{-1} \tag{2.5}$$

where I is a $k \times k$ identity matrix. Matrix N is known as the *Fundamental Matrix*. One special feature of this matrix is N only needs one part of P, that is Q, which is a sub-stochastic matrix. This takes transitions that leaves and returns to transient states. The definition of the fundamental matrix for a regular (ergodic) Markov chain is different please refer to 2.12.

From 2.5, we will get the following equality.

$$N = I + NQ = I + QN \tag{2.6}$$

When N goes through set D, then the first visit to $X\backslash D$ from this distribution of the Markov Chain is given by the matrix

$$U = NT \tag{2.7}$$

The equation 2.7 has the following interpretation. For any state $x \in X\backslash D$, a Markov Chain can either jump directly to state $y \in X\backslash D$ in one step or move to state $z \in D$ and make several cycles before going to $y \in X\backslash D$, see figure 2.3.

Let's look at some interesting quantities that can be expressed in terms of the fundamental matrix. If a process starts and ends in set D, then the following equations are true: $N = (I - Q)^{-1}$ is known as the Mean and $N_2 = N(2N_{dg} - I) - N_{sq}$ is known as the Variance. N_{dg} is the matrix that contains only the diagonal elements from the fundamental matrix and all other entries are zero. N_{sq} is the square of the fundamental matrix. If the process starts in set D and ends in set $X\backslash D$, then the following are true: $\tau = N\xi$ is called the Mean. ξ is the unit vector. $\tau_2 = (2N - I)\tau - \tau_{sq}$ is the Variance.

Let's consider when the Markov Chain can be reduced. We will denote this as Em-

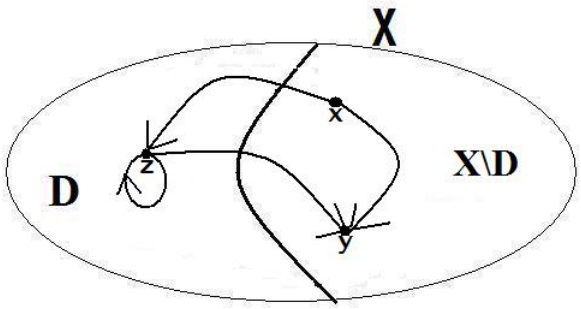


Figure 2.3: Fundamental Probability Diagram

bedded Markov Chain. There is a lot of information on computing many characteristics of Markov Chains. We will be referencing the text "Probability, Markov Chain, Queues, and Simulation" by William J. Stewart (see [32]) for a linear algebra approach.

With the Strong Markov Property and probabilistic reasoning, we can get the following lemma of the State Reduction Approach (SR).

Suppose $M_1 = (X_1, P_1)$ is a finite Markov model and let's assume that (Z_n) $n = 1, 2, \ldots$ be a Markov Chain specified by the model M_1. Let $X_2 \subset X_1$ and let $\tau_1, \tau_2, \ldots, \tau_n, \ldots$ be the sequence of Markov times of first, second, and etc on visits of (Z_n) to the set of X_2, so that $\tau_1 = min\{k > 0 : Z_k \in X_2\}$, $\tau_{n+1} = min\{k : \tau_n < k, Z_k \in X_2\}$, $0 < \tau_1 < \tau_2 < \ldots$. Let $u_1^{X_2}$ be the distribution of the Markov Chain (Z_n) for the initial model M_1 at the moment τ_1 of the first visit to set X_2. Let's consider the random sequence $Y_n = Z_{\tau_n}$, $n = 1, 2, \ldots$.

Lemma 2.1. *(Kolmogorov-Doëblin)*

1. *The random sequence (Y_n) is a Markov Chain in a model $M_2 = (X_2, P_2)$, where*

2. *The transition matrix $P_2 = \{p_2(x, y)\}$ is given by the equation*

$$p_2(x, y) = p_1(x, y) + \sum_{z \in D} p_1(x, z)\, u_1(z, y), \quad (x, y \in X_2) \tag{2.8}$$

where $u_1(z, y)$ is the probability from set D to $X_2 = X_1$ D. This formula can be found in Kemeny and Snell [13]

The strong Markov property for (Z_n) is implied immediately in part 1, meanwhile, the proof for 2 is obvious, see figure 2.3. We will present equation (2.8) in matrix form. If P_1 is decomposed as follows

$$P = \begin{bmatrix} Q_1 & T_1 \\ R_1 & P_1' \end{bmatrix}, \tag{2.9}$$

then

$$P_2 = P_1' + R_1 U_1 = P_1' + R_1 N_1 T_1. \tag{2.10}$$

One would notice that matrix P has the same form as matrix 2.4.

The matrix U_1 gives the distribution of Markov Chain at the moment of the first exit from D (the exit probability matrix), $N_1 = (I - Q)^{-1}$ is the fundamental matrix for the substochastic matrix Q_1, and I is the $|D| \times |D|$ identity matrix. If one is given set D, then matrices N_1 and U_1 are connected by $U_1 = N_1 T_1$.

M_1 and M_2 (which is the X_2-reduced model) will be known as *adjacent* models in this thesis. Let's take a look at the case when set D consists of one non-absorbing point z. We have the formula (2.8) which will be converted to the following form:

$$p_2(x, y) = p_1(x, \cdot) + p_1(x, z)\, n_1(z)\, p_1(z, \cdot), \quad (x \in X_2) \tag{2.11}$$

where $n_1(z) = \frac{1}{1 - p_1(z,z)}$.

By implementing the formula above, each row-vector of the new stochastic matrix P_2 is a linear combination of P_1, with the z-column deleted. This procedure is similar to the one step Gaussian elimination method for solving a linear system.

Suppose that the initial Markov model $M_1 = (X_1, P_1)$ is finite, that is $|X_1| = k$, and one state is eliminated at a time, then we have a sequence of stochastic matrices (P_n), $n = 2, ..., k$. This sequence can be calculated recursively using formula (2.8) as a basis. Equation (2.8) can be written as follows

$$p_{n+1}(x, y) = p_n(x, \cdot) + p_n(x, z)\, n_n(z)\, p_n(z, \cdot), \quad (x \in X_{n+1})$$

After calculating these probabilities, we will be able to calculate various characteristic

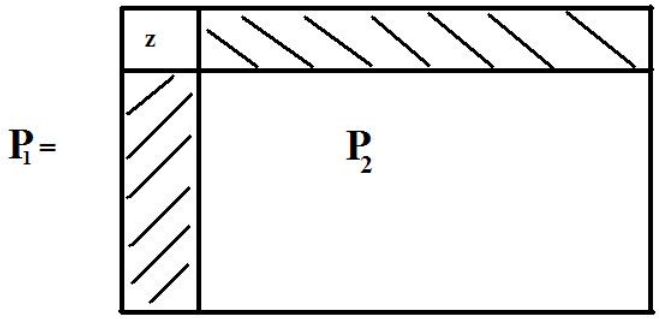


Figure 2.4: Gaussian Elimination Table

of the initial Markov model M_1 recursively starting from a point in the reduced model M_s, where $1 < s \leq k$. There are two conditions that must be satisfied: First, we need to identify a relationship between the characteristic in the reduced model. Secondly, find a model with one more point. This relationship could be obvious or very complicated.

2.3 Ergodic Markov Chain and Its Fundamental Matrix

Let's $\pi_i(n)$ denote the probability that a Markov Chain is at a given state i at step n. State probabilities at any time step n may be obtained from knowledge of the initial distribution and the transition matrix. Formally, we have $\pi_i(n) = \sum_k P\{Z_n = i | Z_0 = k\} \pi_k(0) = \sum \pi_k\{0\} p_{ki}$, that is, $\pi(n) = \pi(0) P^{(n)} = \pi(0) P^n$, where $\pi(n)$ is the *Distribution*, at time n and $\pi(0)$ is the initial state distribution. Since we're dealing with homogeneous discrete-time Markov Chains only, $P^{(n)} = P^n$.

Let $Z = \{Z_n; n = 0, 1, \ldots\}$ be a Markov Chain with a finite space state E and let P be the transition probability matrix. Then the vector π is the Stationary (Invariant) Distribution if this vector is the solution to the following system of equations: $\pi = \pi P$ and $\sum_{i \in E} \pi(i) = 1$.

Let's recall the definition from section 2.1 of Regular and Ergodic Markov Chain. After stating a theorem, we'll see how these two chains relate to one another.

Theorem 2.2. *If P is a regular transition matrix, then the following statements are valid:*

1. *$P^n \to A$, where A is known as the Limiting Matrix.*

2. *Each row of matrix A contains the same probability vector $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)$, i.e. $A = \xi\alpha$, where ξ is a column unit vector.*

3. *All components of α are positive.*

Figure 2.1 is a regular Markov chain and the limiting matrix is $A = \begin{bmatrix} \frac{5}{13} & \frac{7}{13} & \frac{1}{13} \\ \frac{5}{13} & \frac{7}{13} & \frac{1}{13} \\ \frac{5}{13} & \frac{7}{13} & \frac{1}{13} \end{bmatrix}$.

A Markov Chain is Ergodic, if there's a path leading from every state to every other state with a positive probability. Unlike regular Markov chains, if P is any Ergodic matrix, then not necessarily all $P^n \to A$. In this case, the sequence of powers P^n is Euler-Summable to the limiting matrix A. If it's possible to go from any state to any other state in n steps, then the chain is regular and also ergodic (see definition 2.1). The converse is not not true. Suppose transition matrix $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Then one can clearly see from the matrices below that matrix P is ergodic but not regular.

$$P = P^3 = P^5 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad P^2 = P^4 = P^6 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

We have seen in section (2.2) that the fundamental matrix for the Absorbing Markov Chain contained a substochastic matrix, Q. Since Q^n approaches 0 as n approaches infinity, $(I - Q)^{-1}$ is invertible. This is not the case for the Ergodic Markov Chain. So, one need to modify the fundamental matrix definition that would apply to Ergodic Markov Chains.

Definition 2.1. The Fundamental Matrix for the regular (ergodic) Markov Model is

$$Z = (I - (P - A))^{-1} = (I - P + A)^{-1} \tag{2.12}$$

$$= I + \sum_{n=1}^{\infty} (P - A)^n = I + \sum_{n=1}^{\infty} (P^n - A) \tag{2.13}$$

Some other properties which uses the fundamental matrix are the following:

1. *Mean matrix* is $M = (I - Z + EZ_{dg})\, D$, where I is the identity matrix, Z_{dg} is the fundamental diagonal matrix, D is diagonal matrix with elements $d_{ii} = \frac{1}{a_i}$, and E is

a unit matrix, that is, a matrix that has every position as 1.

2. *Variance matrix is* $M_2 = W - M_{sq}$, where $W = M\left(2Z_{dg}D - I\right) + 2\left(ZM - E\left(ZM\right)_{dg}\right)$ and M_{sq} represents squaring the mean.

3. *Covariance matrix*, C, formula is as follows: $\varphi = a_i\left(2z_{ii} - 1 - a_i\right)$ is the calculations for the diagonal and $c_{ij} = a_i z_{ij} + a_j z_{ji} - a_i d_{ij} - a_i a_j$ is the calculation for the off diagonal. It will look like the following: $Cov = \begin{bmatrix} \varphi_{11} & \cdots & c_{13} \\ \cdots & \cdots & \cdots \\ c_{n1} & \cdots & \varphi_{nn} \end{bmatrix}$.

2.4 Basic Facts about the Central Limit Theorem

Let's recall briefly basic facts about Central Limit Theorem (CLT) for Markov Chains. If x_t, $t = 1, 2, \ldots$ is an ergodic aperiodic Markov Chain or (in much more general situation) Markov Chains on the measurable space X satisfying the Doëblin condition, then there exist an unique invariant distribution $\pi(y)$ and

$$|p(t, x, y) - \pi(y)| \leq ce^{-\gamma t}$$

for appropriate $c, \gamma > 0$ (finite dimensional). One can introduce the Hilbert space

$$L^2(X, \pi) = \{f \in X \to R^1 \vdots \int f(y)\pi(y) = 0, \int f^2\pi(dy) \neq \|f\|_\pi^2 < \infty\} \tag{2.14}$$

Markovian process x_t $t \geq 0$ with initial distribution π is ergodic and the stationary one.

Remark. Among several forms (see [15]) of the Doëblin condition, we select the following one $[D_2]$. Let $(X, B(X))$ be a measurable space, $P(x, \Gamma)$ $x \in X$, $\Gamma \in B(X)$ be a transition function. $[D_2]$ holds if there are reference probabilistic measure $\mu(\Gamma)$ on $B(X)$, integer $k_0 > 0$ and constant $\delta > 0$ such that for any x, Γ

$$P(k_0, x, \Gamma) \geq \delta\mu(\Gamma)$$

One can prove now that under D_2 there exist unique invariant probabilistic measure

$$\pi(\cdot) : \pi(\Gamma) = \int_x \pi(dx) P(x, \Gamma)$$

and for $n \to \infty$

$$\sup_{x,\Gamma} |P(n, x, \Gamma) - \pi(\Gamma)| \leq (1 - \delta)^{\left[\frac{n}{k_0}\right]}$$

The following series (fundamental operator, fundamental matrix) converges on $L^2(X, \pi)$ in the strong operator sense (see [13])

$F = I + P + P^2 + \cdots$

Assume that P^* is the operator conjugated to P in $L^2(X, \pi)$. It means that for $\forall(\Psi_1, \Psi_2 \in L^2(\cdot))$

$(P\Psi_1 \cdot \Psi_2)_\pi = (\Psi_1 \cdot P^*\Psi_2)$

For finite or countable Markov Chains

$$P^*(x, y) = \frac{\pi(y)P(y, x)}{\pi(x)} \tag{2.15}$$

Note that P^* is a stochastic matrix (or a stochastic operator in the general case) and one can introduce (Markov Chains) $x^*(t)$, $t = 0, 1, ...$ with probabilities P^*. It is easy to understand that $x^*(t)$ is the initial chain $x(t)$ but in the inverted time. Invariant measure of $x^*(t)$ is again π.

The Fundamental Matrix is closely related to the martingale approach to the Central Limit Theorem. Let's recall the equation (2.14) and fix $f \in L^2(X, \pi)$. The following is known as the homological(Poisson) equation,

$$f = g - Pg, \quad g \in L^2(\cdot) \tag{2.16}$$

Formally,

$$g = (I - P)^{-1} f = \sum_{k=0}^{\infty} P^k f = Ff \tag{2.17}$$

but due to the estimation of (*), the last series converges in $L^2(\cdot)$ norm.

Let's consider the additive functional

$$S_n = f(x_0) + \cdots + f(x_{n-1}) = g(x_0) - (Pg)(x_0) + \cdots + g(x_{n-1}) - (Pg)(x_{n-1})$$

$$= \underbrace{[g(x_0) - (Pg)(x_0)]}_{m_0} + \underbrace{[g(x_1) - (Pg)(x_1)]}_{m_1} + \cdots + \underbrace{[g(x_{n-1}) - (Pg)(x_{n-1})]}_{m_{n-1}}$$

The first term is bounded and the sequence (m_1, F_1), (m_2, F_2), $\ldots$, (m_{n-1}, F_{n-1}) where $F_k = \sigma(x_0, \ldots, x_k)$, $k = 0, 1, 2, \ldots$ is a stationary ergodic martingale difference (as

$$E\left[g(x_k) - (Pg)(x_{k-1}) \mid F_{k-1}\right] = E[g(x_k) \mid F_{k-1}] - (Pg)(x_{k-1}) = (Pg)(x_{k-1}) - (Pg)(x_{k-1}) = 0)$$

We can use now the classical Central Limit Theorem for the martingale-difference (see [2]). It gives the following:

$$\frac{S_n}{\sqrt{n}} = \frac{m_0}{\sqrt{n}} + \frac{m_1 + \cdots + m_{n-1}}{\sqrt{n}} \stackrel{\text{law}}{\to} N(0, \sigma^2) \tag{2.18}$$

where (after simple calculations)

$$\sigma^2 = Em_1^2 = E\left[g(x_1) - (Pg)(x_0)\right]^2 = E(g^2(x_1)) - E((Pg)^2(x_0)) = (g \cdot g)_\pi - (Pg \cdot Pg)_\pi \tag{2.19}$$

$$= (g - Pg; g + Pg)_\pi = (f, f + 2Pf + 2P^2 f + \cdots)_\pi = (f, f + Pf + P^* f + \cdots + P^k f + (P^*)^k f + \cdots)_\pi$$

$$= (f, (F + F^* - I)f)_\pi \tag{2.20}$$

Operator $B = F + F^* - I = I + PF + P^*F^*$ will be called the covariance operator for the Markov Chain x_t and quadratic form $\sigma^2(f) = (f \cdot Bf)_\pi$ gives the limiting variance in the CLT:

$$\frac{S_n}{\sqrt{n}} \stackrel{\text{law}}{\to} N(0, \sigma^2(f))$$

2.5 "Loop" Markov Chain

This idea of "Loop" Markov Chains (LMC) was proposed first by Kai Lai Chung, see [5] in a sightly different setting then what I will be considering.

Let's consider a fixed number $k \geq 2$ and the integers $n_1, \ldots, n_k$ such that $\text{GCD}(n_1, \ldots, n_k) = 1$. The phase space X has the following structure: there is a common point O of all k loops, each loop l_j consist of successive points.

$$l_j = (0, 1_j, 2_j, \ldots, n_j - 1), j = 1, 2, \ldots, k \tag{2.21}$$

Transition probabilities of loop Markov Chains have a simple structure:

a. $p(0, 1_j) = p_j > 0, \;\; j = 1, 2, \ldots, k \qquad \sum_{j=1}^{k} p_j = 1$

b. Inside each loop, the motion is deterministic: $p(1_j, 2_j) = 1, \ldots, p(n_j - 1, 0) = 1$, $j = 1, 2, \ldots, k$ This chain is ergodic and aperiodic (due to arithmetic condition on $(n_1, \ldots, n_k)$)

It means that transition probability $P^n = [p(n, x, y]$, $x, y \in X$ are strictly positive for $n \geq n_0$ and for appropriate $\gamma > 0$, $c > 0$

$$\|p(n, x, y) - \pi(y)\| \leq ce^{-\gamma n}$$

.

CHAPTER 3: THE OPTIMAL STOPPING PROBLEM FOR MARKOV CHAINS AND KNOWN ALGORITHMS

3.1 General Markov Decision Problem and Basic Notation

Since the Optimal Stopping Problem is a special case of the Markov Decision Problem, we'll start with a review of the general structure of this problem. The general Markov Decision Problem is described by a tuple $M = (X, A, A(x), P^a, r(x,a), \beta, L)$, where X is the state space, A is the action space, $A(x)$ is the set of actions available at each state $x \in X$, $P^a = \{p(x, y|a)\}$ is the transition matrix, $r(x,a)$ is the (current) reward function and β is the discount factor, $0 < \beta \leq 1$. P^a describes the transitions of each state x and the actions a assigned to them. Finally, L is a functional which is defined on all trajectories in this system. We consider the important case of total discounted reward, i.e. $L(\cdot) = \sum r(x_i a_i .....)\, \beta^i$.

The time parameters t, s, and $n \in N$ and the sequence $x_0 a_0 x_1 a_1 .....$ represents the trajectory. $H_\infty = (X \times A)^\infty$ denotes the set of trajectories, while $h_n = x_0 a_0 x_1 a_1 .....x_n a_n$ describes the trajectory history of length n. Let $H_n = X \times (X \times A)^{n-1}$ be the space of histories up to $n \in N$.

A non-randomized policy is a sequence of measurable function ϕ_n, $n \in N$, from H_n to A such that $\phi_n(x_0 a_0 x_1 a_1 .....x_n a_n) \in A(x_n)$. If the ϕ_n depend only on X_n for each n, we will say the policy ϕ is *Markov*. ϕ_n is stationary if ϕ is a Markov Policy and does not depend on n. Formally speaking, a *stationary policy* is a single measurable mapping $\phi : X \longrightarrow A$ such that $\phi(x) \in A(x)$ for every $x \in X$.

It is possible to increase the set of policies by selecting actions at random. A *random policy*, π, is as sequence of transition probabilities $\pi_n(da_n|h_n)$ from H_n to A, where $n \in N$, such that $\pi_n(A(x_n)|x_0 a_0 ... x_{n-1} a_{n-1} x_n) = 1$. A policy π is known as a *randomized Markov* if $\pi_n(\cdot|x_0 a_0 ... x_{n-1} a_{n-1} x_n) = \pi_n(\cdot|x_n)$. If $\pi_m(\cdot|x_m) = \pi_n(\cdot|x_n)$ for every $m, n \in N$, then the randomized Markov policy π is a *randomized stationary policy*.

Let x be the initial state and π a policy. P_x^π is the induced probability and E_x^π is the expectation of this measure.

Now, we will introduce the reward operators:

$$P^a f(x) = E\left[f(x_1) | x_0 = x,\ a_0 = a\right], \qquad F^a f(x) = r(x,a) + \beta P^a f(x)$$

$\beta = 1$ for the averaging and the total cost operators.

The optimality operators are as follows:

$$Pf(x) = \sup_{a \in A(x)} P^a f(x), \qquad Ff(x) = \sup_{a \in A(x)} F^a f(x)$$

Let $\pi = (\pi(\cdot|h_n))$. $w^\pi(x)$ is the value of π at an initial state x, that is, $w^\pi(x) = E_x^\pi\left[\sum_{n=0}^{\infty} \beta^n r(Z_n, A_n)\right]$. E_x^π is the expectation with respect to π, while Z_n and A_n are the random state and action taken at moment n. The *value function* $v(x)$, defined by $v(x) = \sup_\pi w^\pi(x)$, and satisfies the Bellman (optimality) equation

$$v(x) = \sup_{a \in A(x)} \left(r(x,a) + \beta P^a v(x)\right)$$

where $P^a f(x) = \sum_y p(x,y|a) f(y)$ is the *averaging operator* on transition P^a.

3.2 Optimal Stopping Problem For Markov Chains

Let's review some basic notation for the Optimal Stopping Problem (OSP). The Optimal Stopping Problem is a special case of Markov Decision Problem which was considered in section 3.1, where the action space contains just two actions,for all x, $A(x) = \{stop, continue\}$. The Optimal Stopping Problem is defined as a tuple $M = (X, P, c, g, \beta)$ where X is the state space, P is the transition (probability) matrix, $c(x) = r(x, stop)$ is the one step cost function, $g(x) = r(x, continue)$ is the terminal reward, and β is the discount factor, $0 \leq \beta \leq 1$. A tuple that doesn't contain a terminal reward, $M = (X, P, c, \beta)$, is called a Reward Model. A tuple that doesn't have a terminal reward, cost function, and discount factor, $M = (X, P)$, is called a Markov Model. The value function, $v(x)$, for an

Optimal Stopping Problem is defined as

$$v(x) = \sup_{\tau \geq 0} E_x \left[\sum_{i=0}^{\tau-1} \beta^i c(Z_i) + \beta^\tau g(Z_\tau) \right]$$

where the *sup* is taken over all stopping times τ, $\tau \leq \infty$If $\tau = \infty$ with positive probability, we will assume that $g(Z_\infty) = 0$. $Pf(x)$ is the averaging operator, $Pf(x) = \sum p(x,y) f(y)$ and $Ff(x) = c(x) + \beta Pf(x)$ be the reward operator. Finally, $S = \{x : v(x) = g(x)\}$ is the optimal stopping set.

If $\beta < 1$, it will be better handled by converting to the model $\beta = 1$. Using this technique, one would need to introduce another state, the absorbing state, which will be denoted as e. We will multiply the existing transition (probability) matrix by the discount factor, which will be denoted by β, that is, $p^\beta(x,y) = \beta p(x,y)$ for $x, y \in X$. Next, we will add another row and column which would represents the absorbing state. The added row represents the transition from the absorbing state to any of the other states in the system, $p^\beta(e,x) = 0$. The added column represents the transition from any other state in the model to the absorbing state. The probability is one minus the discount factor, $p^\beta(x,e) = 1 - \beta$. Finally, the transition probability from the absorbing state back to the absorbing state is one, $p^\beta(e,e) = 1$. It's convenient to consider a more general setting when β is replaced by a function $\beta(x)$, $0 \leq \beta(x) \leq 1$, the probability of "survival", $\beta(x) = P_x(Z_1 \neq e)$.

Assume $G \subseteq X$. Let's denote τ_G be the moment of first visit to G, i.e., $\tau_G = \min(n \geq 0 : x_n \in G)$. The following is an important result of OSP with countable X.

Theorem 3.1. *(Shiryayev 1969, 2008)*

(a) *The value function $v(x)$ is minimal solution of the Bellman (optimality) equation $v = \max(g, c + Pv)$, i.e. the minimal function satisfying $v(x) \geq g(x)$, $v(x) \geq Fv(x)$ for all $x \in X$;*

(b) $v(x) = \lim_n v_n(x)$, *where $v_n(x)$ is the value function for the optimal stopping problem on a finite time interval of length n; satisfying the recurrent relations $v_{n+1} = \max(g, c + Pv_n)$, $v_0 = g$*

(c) *for any $\epsilon > 0$ the random time $\tau_\epsilon = \min\{n \geq 0 : g(Z_n) \geq v(Z_n) - \epsilon\}$, is an ϵ-optimal*

stopping time;

(d) *if* $P_x(\tau_0 < \infty) = 1$ *then the random time* $\tau_0 = \min\{n \geq 0 : g(Z_n) = v(Z_n)\}$ *is an optimal stopping time;*

(e) *if state space* X *is finite then set* $S = \{x : g(x) = v(x)\}$ *is not empty and* τ_0 *is an optimal stopping time.*

3.3 The Elimination Algorithm

The objective of the Optimal Stopping Problem is to find the set of optimality, if it exists. This task could be very difficult directly; however, a lot easier indirectly. Let identify the state(s) that don't belong to the stopping set. Next, we will "eliminate" them from the state space. This idea was proposed by Isaac Sonin, see [25].

Assume the state space is finite and the states that don't belong to the stopping set is removed one at a time. Also, let $M_1 = (X_1, P_1, c_1, g)$ be an Optimal Stopping Problem with finite $X_1 = \{x_1, ..., x_k\}$ and F_1 be corresponding averaging operator. The first step is to calculate the difference $g(x_i) - Fg(x_i)$, $i = 1, 2, ..., k$ until the first state is identified whose difference is negative. If all differences are nonnegative, then the following statement is true.

Proposition 3.2. *Let* $M = (X, P, g)$ *be an optimal stopping problem, and* $g(x) \geq Fg(x)$ *for all* $x \in X$*. Then* X *is the optimal stopping set in the problem* M*, and* $v(x) = g(x)$ *for all* $x \in X$*.*

If $g(x) \geq Fg(x)$ for all $x \in X$ is not true, then there exist a z where $g(z) < F(z)$. This implies that $g(z) < v(z)$ and z is not a part of the stopping set. The next step is to "reduced" model of OSP $M_2 = (X_2 P_2, c_2, g)$ with state set $X_2 = (X_1 \setminus \{z\})$ and transition probabilities $p_2(x, y)$, $x, y \in X_2$, recalculated by $P_2 = P_1' + R_1 U_1 = P_1' + R_1 N_1 T_1$. Notice this is the same formula used in lemma 2.1.

The following theorem guarantees the stopping set in the reduced model M_2 coincides with the optimal stopping set in the initial model M_1.

Theorem 3.3. *(Elimination Algorithm Theorem, [25]) Let* $M_1 = (X_1, P_1, c_1, g)$ *be an optimal stopping model* $D \subseteq C_1 = \{z \in X_1 : g(z) < Fg(z)\}$*. Consider an Optimal stopping*

model $M_2 = (X_2, P_2, c_2, g)$ *with* $X_2 = X_1 \setminus D$, $p_2(x,y)$ *defined by* $P_2 = P_1' + R_1U_1 = P_1' + R_1N_1T_1$, *and* c_2 *is defined by* $c_2 = c_{1,x_2} + R_1N_1c_{1,D}$. *Let* S *be the optimal stopping set in* M_2. *Then*

1. S *is the optimal stopping set in* M_1 *also,*

2. $v_1(x) = v_2(x) \equiv v(x)$ *for all* $x \in X_2$, *and for all* $z \in D$

$$v_1(z) = E_{1,z}\left[\sum_{n=0}^{\tau-1} c_1(Z_n) + v(Z_\tau)\right] = \sum_{w\in D} n_1(z,w)\, c_1(w) + \sum_{y\in X_2} u_1(z,y)\, v(y), \quad (3.1)$$

where $u(z,\cdot)$ *is the distribution of a MC at the moment* τ *of first visit to* X_2, *and* $N_1 = \{n_1(z,w), z, w \in D\}$ *is the fundamental matrix for the substochastic matrix* Q_1.

After some set is eliminated, we need to check the differences $g(x) - F_2g(x)$ for $x \in X_2$, where F_2 is an averaging operator for the stochastic matrix P_2 and so on. This process continues for no more than k steps, producing the model $M_k = (X_kP_k, c_k, g)$, where $g(x) - F_kg(x) \geq 0$ for all $x \in X_k$. This implies X_k is the stopping set in this and all previous models.

Finally, the recursively calculate of the values of $v(x)$ for all $x \in X_1$, using the formula $v_1(z) = n_1(z)\left[c_1(z) + \sum_{y\in X_2} p_1(z,y)\, v(y)\right]$, starting with the equalities $v(x) = g(x)$ for $x \in S = X_k$, where k is the iteration number in the reduction stage of the algorithm where it stopped.

The examples below show how the elimination algorithm is implemented.

Example 3.1. From figure 3.1, one sees that $X_1 = \{1,2,3\}$, the transition probabilities are given by a matrix P_0. Assume the cost function vector $c(x)$ is: $c_0(x) =$ | −.5 | .9 | 1 |. Also the terminal reward vector $g(x)$ is: $g =$ | 2 | −1.5 | 4 |. The discount factor is $\beta = 0.9$. Because we have a discount factor, an absorbing state, e, must be introduced. The cost function and terminal reward for the absorbing state is 0, that is, $c(e) = g(e) = 0$. Now, our transition matrix becomes matrix P_1. The cost and terminal reward are $c =$ | −.5 | .9 | 1 | 0 |

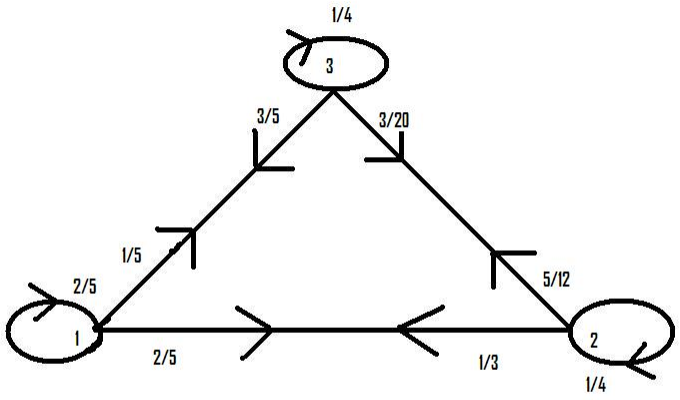


Figure 3.1: Triangle

and $g = \boxed{2}\boxed{-1.5}\boxed{4}\boxed{0}$ respectively. Step one, we need to calculate $g(x) - F_1 g(x) \equiv g(x) - (c_1(x) + P_1 g(x))$ and obtain that $g(2) - (c_1(2) + F_1 g(2)) = -4.1625 < 0$; therefore, state 2 must be eliminated. After some simple calculations, we have a new transition matrix (rounded), P_2 and the cost function values are: $c_2 = \boxed{-.08194}\boxed{1.156774}\boxed{0}$. After calculating $g(x) - F_1 g(x)$, we obtain that $g(1) - (c_2 + F_2 g(1)) = -.33355 < 0$; therefore, state 1 must be eliminated. Again, after some simple calculations, we have a new transition matrix, P_3 and the cost function value is: $c_3 = \boxed{1.059845}\boxed{0}$. If we execute this equation $g(x) - F_1 g(x)$ one more time, the result will be positive, i.e. $g(3) - (c_3(3) + F_3 g(3)) = .102835 > 0$. The optimal stopping set is $S = \{3, e\}$ and $v(3) = g(3) = 4$. By applying equation $f(z) = n_n(z)\left[\sum_{y \in X_2} p_n(z,y) f(y) + c_n(z)\right]$ for $n = 1$, we get $v(1) = f(1) = \left(\frac{1}{.500645}\right)[.354194 * 4 - .08194] = 2.666237$. By applying equation $f(z) = n_n(z)\left[\sum_{y \in X_2} p_n(z,y) f(y) + c_n(z)\right]$ again for $n = 2$, we obtain $v(2) = f(2) = \left(\frac{1}{.775}\right)[.3 * 2.666237 + .375 * 4 + .9] = 4.128866$.

$$P_0 = \begin{array}{|c|c|c|} \hline \frac{2}{5} & \frac{2}{5} & \frac{1}{5} \\ \hline \frac{1}{3} & \frac{1}{4} & \frac{5}{12} \\ \hline \frac{3}{5} & \frac{3}{20} & \frac{1}{4} \\ \hline \end{array} \qquad P_1 = \begin{array}{|c|c|c|c|} \hline .36 & .36 & .18 & .1 \\ \hline .3 & .225 & .375 & .1 \\ \hline .54 & .135 & .225 & .1 \\ \hline 0 & 0 & 0 & 1 \\ \hline \end{array} \qquad P_2 = \begin{array}{|c|c|c|} \hline .499355 & .354194 & .146452 \\ \hline .592258 & .290323 & .117419 \\ \hline 0 & 0 & 1 \\ \hline \end{array}$$

$$P_3 = \begin{array}{|c|c|} \hline .70933 & .29067 \\ \hline 0 & 1 \\ \hline \end{array}$$

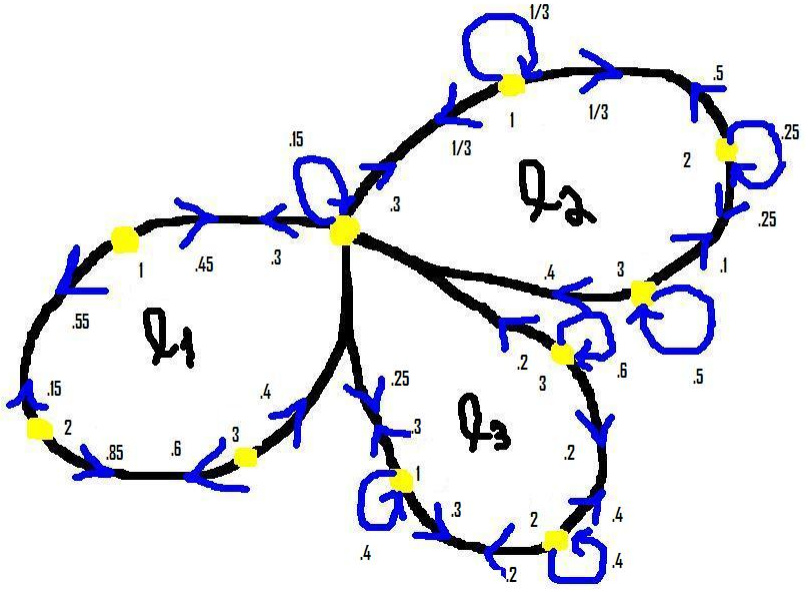


Figure 3.2: Loop Example

Example 3.2. Elimination Algorithm From figure 3.2, the state space is

$$X_1 = \{1_1, 1_2, 1_3, 2_1, 2_2, 2_3, 3_1, 3_2, 3_3\},$$

where the first (larger) number represents the loop number and the second (smaller) number represents the state. The transition probability matrix is

$$P_0 = \begin{array}{|c|c|c|c|c|c|c|c|c|c|}\hline .15 & .3 & 0 & 0 & .3 & 0 & 0 & .25 & 0 & 0 \\ \hline .45 & 0 & .55 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & .15 & 0 & .85 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline .4 & 0 & .6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & .5 & .25 & .25 & 0 & 0 & 0 \\ \hline .4 & 0 & 0 & 0 & 0 & .1 & .5 & 0 & 0 & 0 \\ \hline .3 & 0 & 0 & 0 & 0 & 0 & 0 & .4 & .3 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & .4 & .4 & .2 \\ \hline .2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .2 & .6 \\ \hline \end{array}$$

Assume that the cost vector $c(x)$ is zero, i.e. $c \equiv 0$ and the terminal reward vector $g(x)$ is: $g = \begin{array}{|c|c|c|c|c|c|c|c|c|c|}\hline 1 & 2 & 3.1 & -.5 & 3 & 1 & 2 & -1 & 2 & 3.5 \\ \hline \end{array}$.

Just as in the previous example, when considering a discount factor an absorbing state, e, must be introduced into the model. With the terminal reward value for the absorbing state being zero, that is, $g(e) = 0$ and the new transition matrix becomes

$$P_0 = \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|}\hline .13 & .27 & 0 & 0 & .27 & 0 & 0 & .23 & 0 & 0 & .1 \\ \hline .405 & 0 & .495 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .1 \\ \hline 0 & .135 & 0 & .765 & 0 & 0 & 0 & 0 & 0 & 0 & .1 \\ \hline .36 & 0 & .54 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .1 \\ \hline .3 & 0 & 0 & 0 & .3 & .3 & 0 & 0 & 0 & 0 & .1 \\ \hline 0 & 0 & 0 & 0 & .45 & .22 & .23 & 0 & 0 & 0 & .1 \\ \hline .36 & 0 & 0 & 0 & 0 & .09 & .45 & 0 & 0 & 0 & .1 \\ \hline .27 & 0 & 0 & 0 & 0 & 0 & 0 & .36 & .27 & 0 & .1 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & .36 & .36 & .18 & .1 \\ \hline .18 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .18 & .54 & .1 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \hline \end{array}.$$

After calculating $g(x) - P_0 g(x)$ and realizes that $g(0) - P_0 g(0) \approx -.26 < 0$, state 0 must be eliminated (see figure 3.3).

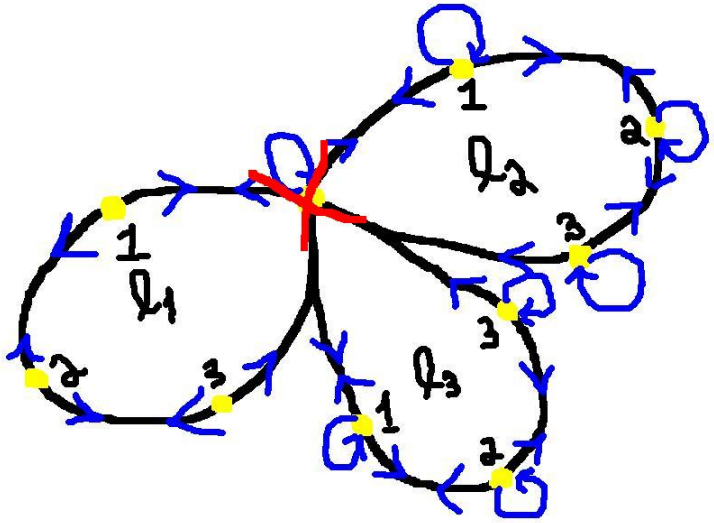


Figure 3.3: First Iteration

Again, utilizing the excel program to calculate $g(x) - P_1 g(x)$, we obtain that $g(1_1) - P_1 g(1_1) \approx -.06 < 0$; therefore, state 1_1 is the next candidate for elimination (see figure 3.4).

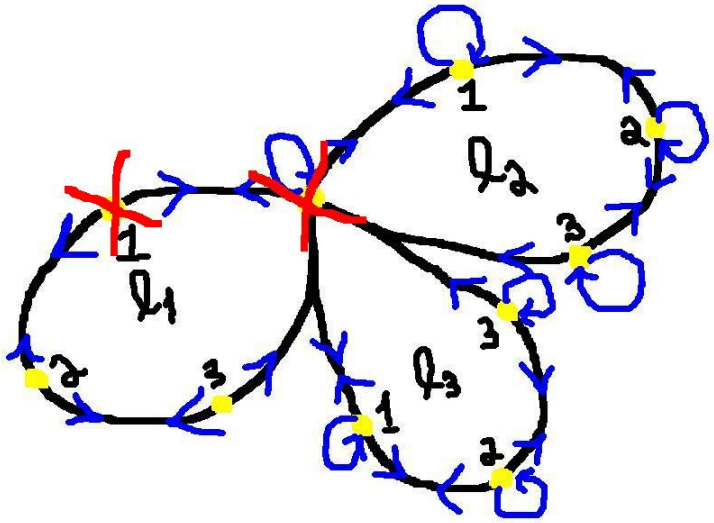


Figure 3.4: Second Iteration

Executing the excel program again, $g(1_3) - P_2 g(1_3) \approx -2.15 < 0$ which means state 1_3 is eliminated (see figure 3.5).

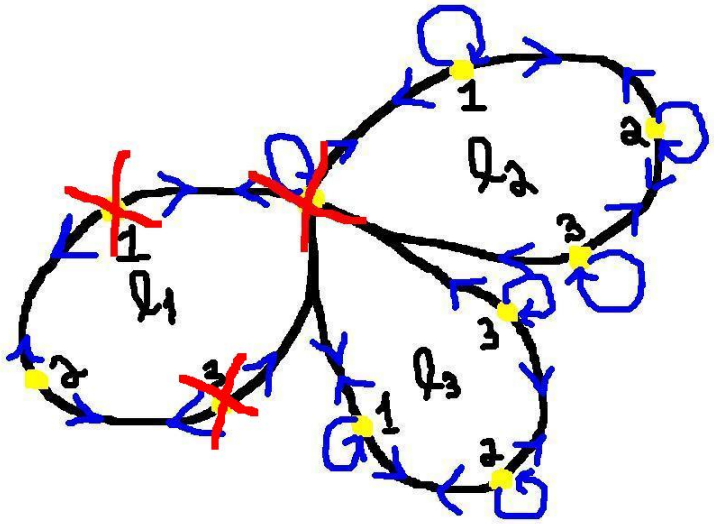


Figure 3.5: Third Iteration

Recalculating the states left in the space state, $g(2_1) - P_3 g(2_1) \approx -1.03 < 0$ state 2_1 needs to be eliminated (see figure 3.6).

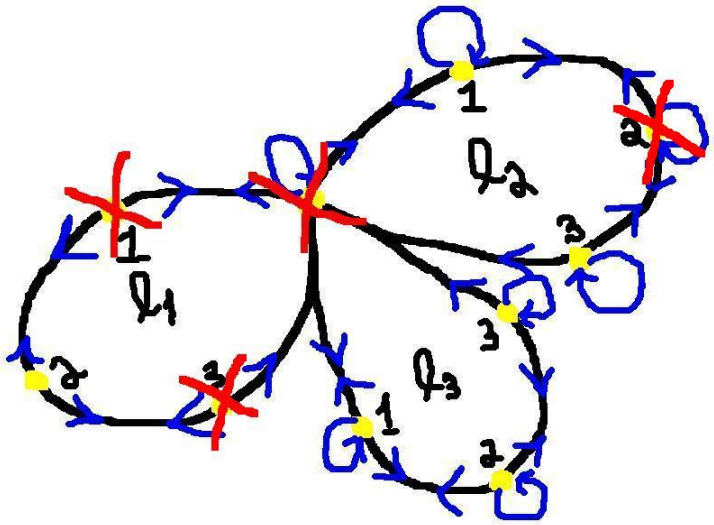


Figure 3.6: Fourth Iteration

Finally, state 3_1 is eliminated because $g(3_1) - P_4 g(3_1) \approx -1.54 < 0$ (see figure 3.7).

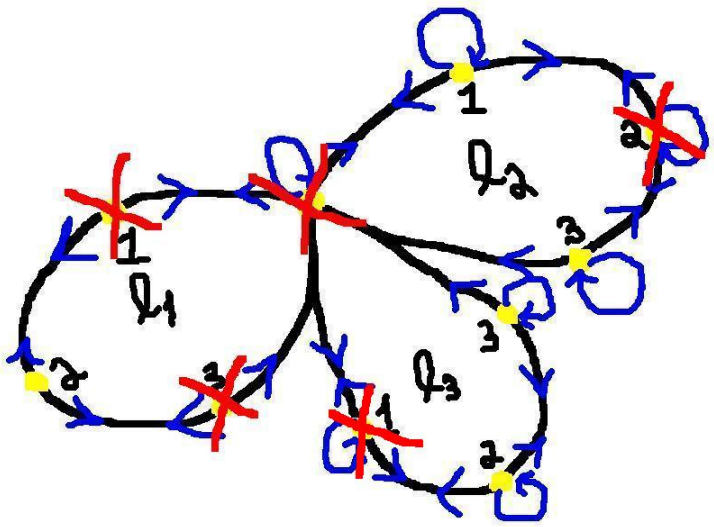


Figure 3.7: Final Iteration

Because $g(x) - P_5 g(x) < 0$ for all $x \in X_5$, we are optimal. The optimal stopping set is $X_5 = \{1_2, 2_1, 2_3, 3_2, 3_3\}$.

Using a excel program (a backward recursive procedure), the results for the Bellman's equation, $v(x)$ is

0	$1,1$	$1,2$	$1,3$	$2,1$	$2,2$	$2,3$	$3,1$	$3,2$	$3,3$	e
2.14	2.40	3.1	2.44	3	2.32	2	1.74	2	3.5	0

CHAPTER 4: A MODIFICATION OF THE ELIMINATION ALGORITHM AND LIFT ALGORITHM

4.1 A Modified Version of Elimination Algorithm with Using Full Size Matrices Technique

Because of tedious calculations for these stochastic matrices of equal size, it was suggested to develop an excel program. Three different programs were created for each algorithm; Elimination Algorithm (which was introduced in section 3.3), a modification to the Elimination Algorithm, which I will call the Full Size Matrix Algorithm and the Lift Algorithm (which will be introduce in this chapter). After the development of these excel programs, E. Presman suggested the following modification to the Elimination Algorithm. Consider the Markov Chain (Y_n) a MC with the same state space X, but with a new transition matrix. For clarity, we will use the following notation: D is the deleted set, $X \setminus D = S$, two diagonal characteristic matrices I_D and I_S. I_D is a diagonal matrix with $d_i = 1$ if $i \in D$ and 0 otherwise. I_S is a diagonal matrix with $S_i = 1$ if $i \in S$ and 0 otherwise. Remember, multiplication of diagonal matrix on the right is equivalent to multiplication of column, and on the left – rows. In the initial model, we removed index 1, that is, $P_1 = P$. In the new model M_2 we're replacing index 2 with D, i.e. $P_2 = P_D$. Now, P, P_D, N_D, I_D, I_S are full size $|X| \times |X|$ square matrices.

The (Kolmogorov, Doëblin) lemma 2.1 remains true, but now we assume the (Y_n) is a Markov Chain with the same *state space* X, specifically, we allow the initial point x be in D as well as in $S = X \setminus D$, even though after the first step MC is always in S. In addition to $P_2 = P_1' + R_1U_1 = P_1' + R_1N_1T_1$ for $x, y \in S$, we have the equation $T + QNT = (I + QN)\,T = NT$ for $x \in D$, $y \in S$. The last equation is true by $N = I + QN = I + NQ$. Instead of $P_2 = P_1' + R_1U_1 = P_1' + R_1N_1T_1$, we obtain the following stochastic matrix for

a MC (Y_n)

$$P_D = PI_S + PI_D N_D PI_S = (I + PI_D N_D)\, PI_S = N_D PI_S = \begin{bmatrix} 0 & N_1 T_1 \\ 0 & P_1^0 + R_1 N_1 T_1 \end{bmatrix} \quad (4.1)$$

where in formula $P_2 = P_1' + R_1 U_1 = P_1' + R_1 N_1 T_1$, P_1^0 is replaced by PI_S, R_1 is replaced by PI_D, T_1 is replaced by PI_S, and $N_1 = (I - Q_1)^{-1}$ is replaced with N_D, where $N_D = (I - PI_D)^{-1}$,$Q_1 N_1 = N_1 - I$

$$PI_D N_D = \begin{bmatrix} Q_1 N_1 & 0 \\ R_1 N_1 & 0 \end{bmatrix},\ I + PI_D N_D = N_D = \begin{bmatrix} N_1 & 0 \\ R_1 N_1 & I \end{bmatrix} \quad (4.2)$$

The Full Size Matrix analog of $c_2 = c_{1,x_2} + R_1 N_1 c_{1,D}$ will be

$$c_D = c + PI_D N_D c = (I + PI_D N_D)\, c = N_D c = \begin{bmatrix} N_1 c_{1,D} \\ R_1 N_1 c_{1,D} + c_{1,S} \end{bmatrix} \quad (4.3)$$

where $c_{1,D}$ and $c_{1,S}$ are apart of vector $c = c_1$ with coordinates in D and S respectively. Now c and c_D are both full vectors defined on the complete state space $X = X_1$. As in $c_2(x) = E_x \sum_{n=0}^{\tau-1} c_1(Z_n) = c_1(x) + \sum_{z \in D} p_1(x,z) \sum_{w \in D} n_1(z,w)\, c_1(w)$ function c_D can be described as

$$c_D(x) = E_{1x} \sum_{n=0}^{\tau-1} c_1(Z_n),\ x \in X \quad (4.4)$$

where $E_{1x} = E_x$ is an initial expectation, $\tau = \tau_S^+$ is the moment of first *return* to $S = X \setminus D$ if $x \in S$.

Theorem 4.1. *(Modified Elimination Theorem) Let $M_1 = (X_1, P_1, c_1, g)$ be an Optimal Stopping (OS) model, $D \subseteq C_1 = \{z \in X_1 : g(z) < F_1 g(z)\}$. Consider an OS model $M_D = (X_D, P_D, c_D, g)$ with $p_D(x,y)$ defined by 4.1 and c_D is defined by 4.3. Let S be the optimal stopping set in M_2. Then*

1. *S is the optimal stopping set in M_1.*

2. $v_1(x) = v_2(x) \equiv v(x)$ *for all* $x \in X_1$*, and for all* $z \in D$

$$v_1(z) = E_{1,z}\left[\sum_{n=0}^{\tau-1} c_1(Z_n) + v(Z_\tau)\right] = \sum_{w\in D} n_1(z,w)\, c_1(w) + \sum_{y\in X_2} u_{1(z,y)v(y)} \tag{4.5}$$

where $u_1(z,\cdot)$ *is the distribution of a MC at the moment* τ *of first visit to* X_2*, and* $N_1 = \{n_1(z,w), z, w \in D\}$ *is the fundamental matrix for the substochastic matrix* Q_1*.*

Remark. The significant feature of State Elimination Algorithm (SEA) is that the elimination of set D_1 and D_2 in two steps is equivalent of a set $D_1 \cup D_2$ in one step. When using State Elimination Algorithm in Excel, we only need the formulas for one step initial elimination: set D will serve as input, c_D, P_D and the output, new set $D' = \{x : (g(x) - (c'(x) + P'g(x))) < 0\}$ serves as input for D. We can initiate an iterative process, where $D_0 = \emptyset$ and $D_n \subset D_{n+1}$, $n = 0, 1, ...$ any time after. The stopping condition is when $D_n = D_{n+1}$. We will implement a recursive procedure to calculate the value function. More specifically, the value function will start with $v_{(n)} = g$ on $S_n = X \setminus D_n$ and implement $v_{(n)} = c_n + P_n g$ on D_n recursively. We also will have $v_{(n)} \nearrow v$ and $v_{(n)} = v$ when the iterations stop.

4.2 The Lift Algorithm

The previous algorithms, State Reduction Elimination and the modification of the Elimination Algorithm "Full Size Matrix Algorithm", work when finite Markov Chain is small. However, when the length and the number of loops are substantial, it could pose a serious drawback to our loop system. Inside each loop, no matter the amount of states, both algorithms work efficiently. A potential problem arises when the origin becomes a candidate for elimination. Let's refer to example 3.2. Notice in both examples, the origin is eliminated. There is really no sufficient difference when the State Elimination Algorithm is used verses the Full Size Matrix Algorithm because the system is small. Suppose the system has about a thousand loops connected at the origin. If the origin become a candidate for elimination, then this could become a serious problem. When the origin is eliminated, it creates a new (complicated) connections, see figures 4.3, 4.4, and 4.5. The first and last states of each loop will be connected to one another. The amount of new connections made are $\binom{2k}{2}$, where k denotes the number of loops. After the introduction of the next algorithm, I'll use example

4.2 to further explain this problem. In order to avoid this situation, I propose the following algorithm.

Suppose we have the same optimal stopping model that was described earlier in this chapter, that is, $M_1 = (X_1, P_1, c_1, g)$ and set $D \subseteq C_1 = \{z \in X_1 : g(z) < F_1 g(z)\}$ which satisfied the condition for elimination. Let's introduce new notation to denote the two previous models and the third one that will be in presented in this section as $M_{D,i}$ where $i = 1, 2, 3$. Let's review

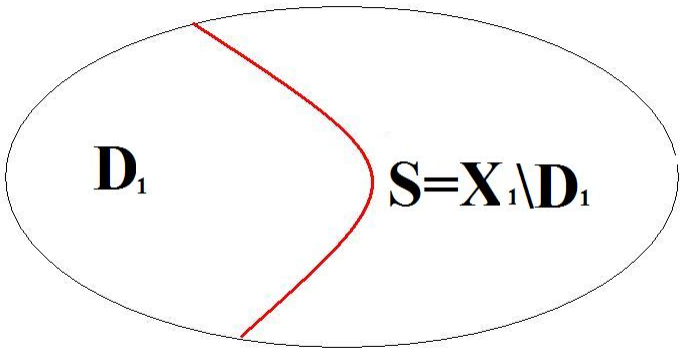


Figure 4.1: $M_{D,1} = (X_2, P_2, c_2, g)$

1. The optimal stopping model $M_{D,1} = (X_2, P_2, c_2, g)$ (see figure 4.1) and new reduced state space is $X_2 = X_1 \setminus D$. This model also has a new *transition matrix*, $p_2(x, y)$ defined by 2.10 and a new *cost matrix*,

$$c_2 = c_{1,x_2} + R_1 N_1 c_{1,D}, \tag{4.6}$$

   but the *terminal reward function*, g stays the same. These two models M_1 and $M_{D,1}$ are equivalent to Theorem 1.

2. Again, assume we have a optimal stopping model $M_{D,2} = (X_1, P_D, c_D, g)$ but this time, the state space stays the same. The *transition matrix*, P_D and *cost function*,

c_D changes. The *terminal reward* function remains that same. These two models M_1 and $M_{D,2}$ are equivalent to Theorem 2.

Now using this transition matrix P_D we consider one more interpretation which also has some practical applications.

Let us introduce also function $g_D(x) = E_x[\sum_{n=0}^{\tau-1} c_1(Z_n) + g(Z_\tau)]$, where $\tau = \tau_S$ is the moment of first *return* to $S = X \setminus D$. In other words, g_D is the expected reward if we stop at S. (This is the same function mentioned in Remark 3M). Then

$$g_D(x) = g(x) \text{ if } \ x \in S, \ \ g_D(x) = c_D(x) + P_D g(x), \text{if } x \in D. \tag{4.7}$$

In matrix form we have

$$\mathbf{g}_D = I_D N_D \mathbf{c} + (I_D P_D + I_S) g. \tag{4.8}$$

In other words the computation of function $g_D(x)$ can be obtained instead of P_D using also matrix $I_D P_D + I_S = P_D$,

$$P_D = \begin{bmatrix} 0 & N_1 T_1 \\ 0 & I \end{bmatrix}. \tag{4.9}$$

The new matrix in a reduced form is $P_{0D} = P_{10} + R_1 N_1 T_1$

Now we can introduce one more model

3. Consider an OS model $M_{D,3} = (X_1, P_1, c_2, g_D)$ with the *same state space* X_1 and the *same initial transition matrix* P_1 , with new function $\mathbf{c}_D = \mathbf{c} I_S$ and with *new elevated terminal reward function* $g_D = v_D$.

Now we can present

Theorem 4.2. *(The Triple equality Theorem) Let* $M_1 = (X_1, P_1, c_1, g)$ *be an initial OS model,* $D \subseteq C_1 = \{z \in X_1 : g(z) < F_1 g(z)\}$, *and* $M_{D,i}, i = 1, 2, 3$ *be three models described above. Let* S *be the optimal stopping set in* M_1. *Then*

(a) *S is the optimal stopping set in all four models,* $M_1, M_{D,i}, i = 1, 2, 3$

(b) *value function is the same in all four models.*

The equivalence of the first three models, i.e M_1 and $M_{D,i}, i = 1, 2$ was proved by Sonin in cite [25] and [26]

So we need only to prove that model $M_{D,4}$, i.e. "lift" model has the same solution.

First we prove two Lemmas. To simplify our presentation we assume that $c(x) = 0$ for all x.

Lemma 4.3. *Let $M = (X, P, g)$ be a finite OS model, $D \subset X$, $S \subset X$ be an absorbing set and $g(x) < Pg(x)$ for all $x \notin S$. Then the value function $v_D(x)$satisfies*

$$v_D(x) = E_x g(Z_\tau) > g(x) \text{ for all } x \notin S, \tag{4.10}$$

where $\tau = \tau_S$ is the moment of the first visit to set S.

Proof. For all $x \in S$, we obviously have $v_D(x) = g(x)$. According to Theorem 3.1, the value function $v_D(x) = \lim v_{D,n}(x)$, where the sequence of nondecreasing functions $v_{D,n}(x)$ defined as: $v_{D,0}(x) = g(x)$, $v_{D,n}(x) = \max(g(x), Pv_{D,n-1}(x)), n = 1, 2, ...$The assumption of the lemma implies that $v_{D,1}(x) = \max(g(x), Pg(x)) = Pg(x) > g(x)$ for all $x \notin S$. Therefore, for all such x, we have $v_D(x) \geq Pg(x) > g(x)$, where set S is an optimal stopping set and formula (4.10) holds. *Q.E.D.*

Note: The function v_D coincides with function $g_D(x)$ which was defined by equation 4.7.

Lemma 4.4. *Let $M = (X, P, c, g)$ be a finite OS model, S be an optimal stopping set and $v(x)$ be the value function in this model. Let function $g'(x)$ satisfies inequality*

$$g(x) \leq g'(x) \leq v(x) \tag{4.11}$$

for all $x \in X$. Then set S is an optimal stopping set in OS model $M' = (X, P, c, g')$ and both models have the same value function.

Proof. The value function $v(x)$ satisfies Bellman equation $v(x) = \max(g(x), c + Pv(x))$. Let us show that $v(x)$ satisfies also Bellman equation for model M'.

$$v(x) = \max(g'(x), c + Pv(x)). \tag{4.12}$$

Since obviously $v(x) \leq v'(x)$, where $v'(x)$ is the value function in model M', this will show that $v(x)$ is the minimal solution for Bellman equation (4.12) and therefore $v(x) = v'(x)$. If $x \in S = \{x : g(x) = v(x)\}$ then (4.11) implies that $g(x) = g'(x) = v(x)$ and $v(x) = g(x) \geq c(x) + Pv(x)$ and hence $g'(x) = v(x) \geq c(x) + Pv(x)$, i.e. for such x function $v(x)$ satisfies (4.12). If $x \notin S$ then $v(x) = c(x) + Pv(x) \geq g(x), v(x) \geq g'(x)$ and again function $v(x)$ satisfies (4.12). Lemma 4.4 is proved. *Q.E.D.*

It is possible that in the model M', there are other new optimal stopping sets S' larger than S. The reason for this is, in general, $S = \{x : g(x) = v(x)\} \subset S' = \{x : g'(x) = v'(x)\}$. Simply put, $v = v'$.

Now using lemmas 4.3 and 4.4, we can prove Theorem 4.2.

Proof. It sufficient to prove that function $g_D(x)$ satisfies the inequalities

$$g(x) \leq g_D(x) \leq v(x)). \tag{4.13}$$

Then by Lemma 4.4 we will obtain that $v_D(x) = v(x)$ and therefore an optimal stopping set S in the initial model is also an optimal stopping set in model $M_{3,D}$. By definition of set D, we have $g(x) < Pg(x)$ and therefore by Lemma 4.3 applied to a situation when set S we may consider as an absorbing set, we have $g(x) < g_D(x)$ for all $x \in D$. On the other hand, $g_D(x) = E_x g(Z_\tau)$, where $\tau = \tau_S$ is the moment of the first visit to set $S = X \setminus D$ and therefore $g_D(x) \leq v(x) = \sup_\tau E_x g(Z_\tau)$. Thus $g_D(x)$ satisfies (4.13) and Theorem is proved. *Q.E.D.*

Theorem 4.2 provides a new computational algorithm to solve Optimal Stopping Problem for finite Markov Chains. As in the State Elimination Algorithm at each state, when a new terminal reward function g' is calculated, there are two possibilities. Ei-

ther $g'(x) - Pg'(x) \geq 0$ for all $x \in X$, and then $g'(x) = v(x)$ for all $x \in X$. The optimal stopping set in this model is the whole space X, and the optimal stopping set S_* in the initial model is the set of all points where the terminal reward function, $g(x)$ was not change at all, i.e. $S_* = \{x : g(x) = v(x)\}$. The other possibility is when set $D' = \{x : g'(x) - Pg'(x) < 0\} \neq \emptyset$. For this reason, function g' is lifted to g'_D. According to equation 4.7, set D' must have as least one point not in set D. Having this requirement satisfied will ensure each time the terminal function changes, it changes at least one point. In the case of the State Elimination Algorithm, the lifts will stop in no more than m steps, where $m = |X|$.

4.3 Examples

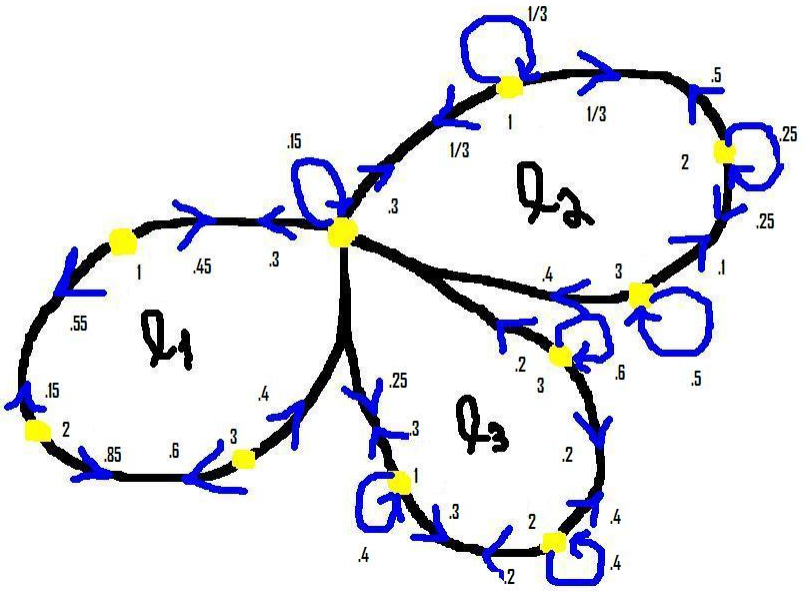


Figure 4.2: Loop Example

Example 4.1. Full Size Matrix The begin of this example will mimic example 3.2. From figure 1, the state space is $X_1 = \{1_1, 1_2, 1_3, 2_1, 2_2, 2_3, 3_1, 3_2, 3_3\}$, where the first (larger) number represents the loop number and the second (smaller) number represents the state. The transition probability matrix is

$$P_0 = \begin{array}{|c|c|c|c|c|c|c|c|c|c|}\hline .15 & .3 & 0 & 0 & .3 & 0 & 0 & .25 & 0 & 0 \\ \hline .45 & 0 & .55 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & .15 & 0 & .85 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline .4 & 0 & .6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & .5 & .25 & .25 & 0 & 0 & 0 \\ \hline .4 & 0 & 0 & 0 & 0 & .1 & .5 & 0 & 0 & 0 \\ \hline .3 & 0 & 0 & 0 & 0 & 0 & 0 & .4 & .3 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & .4 & .4 & .2 \\ \hline .2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .2 & .6 \\ \hline \end{array}$$

Assume that the cost vector $c(x)$ is zero, i.e. $c \equiv 0$ and the terminate reward vector $g(x)$ is: $g = \begin{array}{|c|c|c|c|c|c|c|c|c|c|}\hline 1 & 2 & 3.1 & -.5 & 3 & 1 & 2 & -1 & 2 & 3.5 \\ \hline \end{array}$.

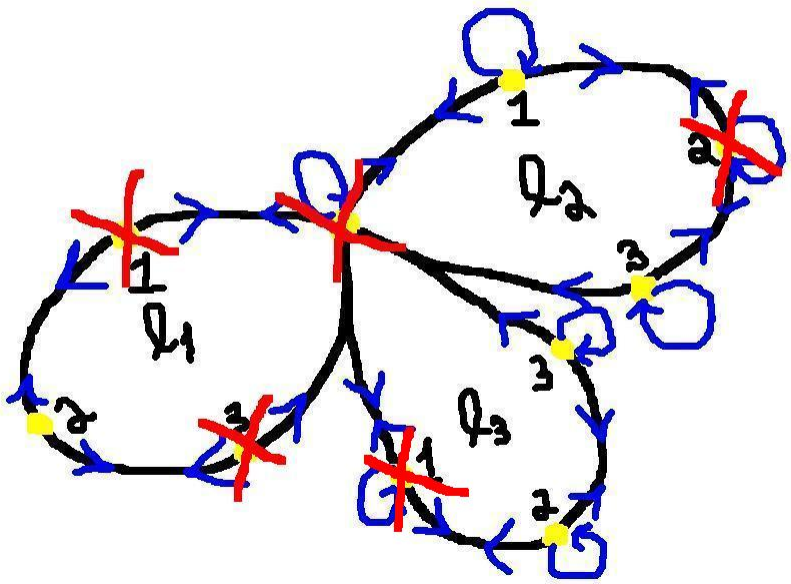
Just as in the previous example, when considering a discount factor an absorbing state, e, must be introduce into the model. With the terminal reward value for the absorbing state being zero, that is, $g(e) = 0$ and the new transition matrix becomes

$$P_0 = \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|}\hline .13 & .27 & 0 & 0 & .27 & 0 & 0 & .23 & 0 & 0 & .1 \\ \hline .405 & 0 & .495 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .1 \\ \hline 0 & .135 & 0 & .765 & 0 & 0 & 0 & 0 & 0 & 0 & .1 \\ \hline .36 & 0 & .54 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .1 \\ \hline .3 & 0 & 0 & 0 & .3 & .3 & 0 & 0 & 0 & 0 & .1 \\ \hline 0 & 0 & 0 & 0 & .45 & .22 & .23 & 0 & 0 & 0 & .1 \\ \hline .36 & 0 & 0 & 0 & 0 & .09 & .45 & 0 & 0 & 0 & .1 \\ \hline .27 & 0 & 0 & 0 & 0 & 0 & 0 & .36 & .27 & 0 & .1 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & .36 & .36 & .18 & .1 \\ \hline .18 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .18 & .54 & .1 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \hline \end{array}.$$

Again, our goal is to find the state or set of states that satisfy the following condition: $g(x) - Pg(x) < 0$. After implementing the excel program to calculate $g(x) - Pg(x)$, we

notice that several states must be eliminated. These states are $\{0, 1_1, 1_3, 2_2, 3_1\}$. This will leave us with the following result (the columns that are highlighted represent the states there were eliminated):

	0	**1_1**	**1_2**	**1_3**	**2_1**	**2_2**	**2_3**	**3_1**	**3_2**	**3_3**	**e**	**g**	**g – Pg**
0	0	0	0	0	0	0	0	0	0	0	0	0	0
1, 1	0	0	0	0	0	0	0	0	0	0	0	0	0
1, 2	0	0	.54	0	.13	0	0	0	.05	0	.27	3.1	.91
1, 3	0	0	0	0	0	0	0	0	0	0	0	0	0
2, 1	0	0	.06	0	.60	0	.09	0	.04	0	.21	3	.76
2, 2	0	0	0	0	0	0	0	0	0	0	0	0	0
2, 3	0	0	.07	0	.20	0	.48	0	.05	0	.20	2	.12
3, 1	0	0	0	0	0	0	0	0	0	0	0	0	0
3, 2	0	0	.03	0	.06	0	0	0	.53	.18	.19	2	.02
3, 3	0	0	.03	0	.07	0	0	0	.21	.54	.14	3.5	.86
e	0	0	0	0	0	0	0	0	0	0	1	0	0



As you can see, the optimal stopping set is $\{1_2, 2_1, 2_3, 3_2, 3_3\}$.

The result for the Bellman's Equation, $v(x)$ is

0	1_1	1_2	1_3	2_1	2_2	2_3	3_1	3_2	3_3	e
2.14	2.40	3.1	2.44	3	2.32	2	1.74	2	3.5	0

Example 4.2. Lift Algorithm Again, let's re-equate ourselves with figure 3.2 to illustrate this algorithm. We will follow the same procedure outlined in either of the previous algorithms for eliminating states until the origin (center point) becomes a candidate for elimination. Instead of eliminating the origin, I propose to "Lift" the terminal reward value $g(x)$. This will ensure that the origin is kept and more importantly, we can avoid our system generating many different connections. Please refer to figures 4.3, 4.4, 4.5 for pictorial explanation.

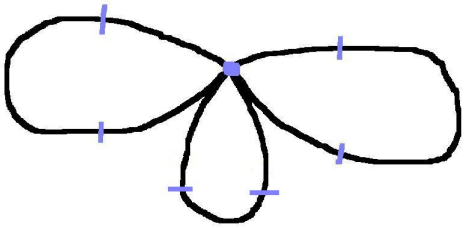


Figure 4.3: Initial Model

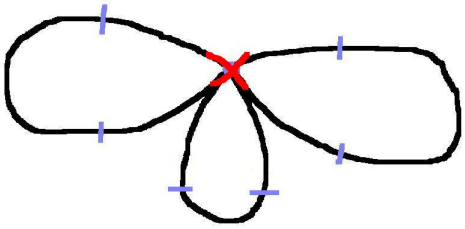


Figure 4.4: Origin Eliminated

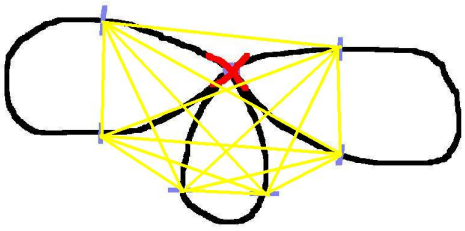


Figure 4.5: New Connections

The next question should be "How much can $g(x)$ be raised?". Answer: The value of $g(x)$ can't exceed of $Pg + c$. By allowing $g = Pg + c$, we will preserve the origin and maintain a "manageable" system. Please notice when $g(0) - (Pg(0) + c(0))$, the result will

be 0.

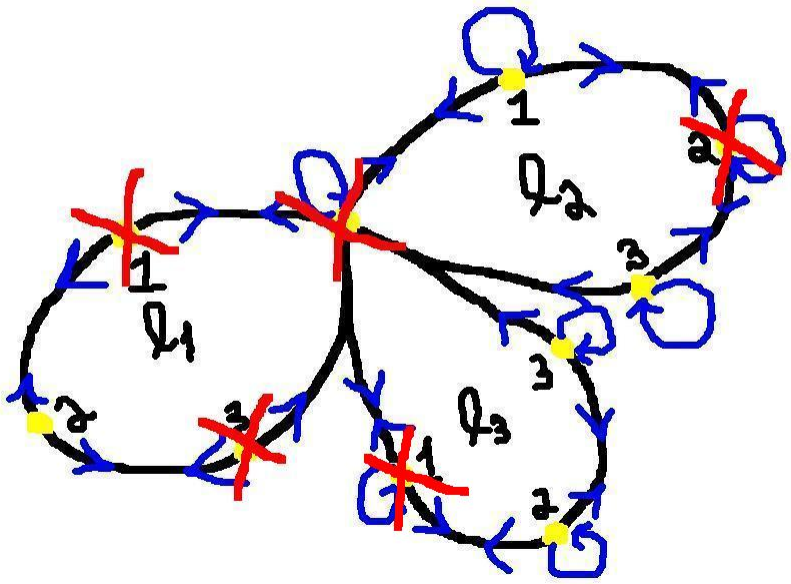
If we look back at example 4.1, one would notice that the origin is the first state eliminated. Instead, let's lift the value, $g(0)$, and avoid the decomposition of our model. $g(0)$ will be "Lifted" from -.26 to .73 and $g - Pg$ must be recalculated.

	0	$\mathbf{1_1}$	$\mathbf{1_2}$	$\mathbf{1_3}$	$\mathbf{2_1}$	$\mathbf{2_2}$	$\mathbf{2_3}$	$\mathbf{3_1}$	$\mathbf{3_2}$	$\mathbf{3_3}$	**e**	**g**	$\mathbf{g - Pg}$
0	.13	.27	0	0	.27	0	0	.23	0	0	.1	1	.73
1, 1	.405	0	.495	0	0	0	0	0	0	0	.1	2	−2.40
1, 2	0	.135	0	.765	0	0	0	0	0	0	.1	3.1	.113
1, 3	.36	0	.54	0	0	0	0	0	0	0	.1	−.5	−2.44
2, 1	.3	0	0	0	.3	.3	0	0	0	0	.1	3	1.16
2, 2	0	0	0	0	.45	.22	.23	0	0	0	.1	1	−1.03
2, 3	.36	0	0	0	0	.09	.45	0	0	0	.1	2	.24
3, 1	.27	0	0	0	0	0	0	.36	.27	0	.1	−1	−1.76
3, 2	0	0	0	0	0	0	0	.36	.36	.18	.1	2	1.01
3, 3	.18	0	0	0	0	0	0	0	.18	.54	.1	3.5	.86
e	0	0	0	0	0	0	0	0	0	0	1	0	0

After the "Lift" algorithm is implemented, the origin is no longer a candidate for elimination. We will execute the "Full Size Matrix" method to get the final result.

Comment: The Lift method can be used more than once, but no more than the number of states minus one.

	0	$\mathbf{1_1}$	$\mathbf{1_2}$	$\mathbf{1_3}$	$\mathbf{2_1}$	$\mathbf{2_2}$	$\mathbf{2_3}$	$\mathbf{3_1}$	$\mathbf{3_2}$	$\mathbf{3_3}$	**e**
0	.34	0	.13	0	.27	0	0	0	.10	0	.16
1, 1	.405	0	.495	0	0	0	0	0	0	0	.1
1, 2	.33	0	.48	0	0	0	0	0	0	0	.19
1, 3	.36	0	.54	0	0	0	0	0	0	0	.1
2, 1	.3	0	0	0	.47	0	.09	0	0	0	.14
2, 2	0	0	0	0	.58	0	.29	0	0	0	.13
2, 3	.36	0	0	0	.05	0	.48	0	0	0	.11
3, 1	.42	0	0	0	0	0	0	0	.42	0	.16
3, 2	.15	0	0	0	0	0	0	0	.51	.18	.16
3, 3	.18	0	0	0	0	0	0	0	.18	.54	.1
e	0	0	0	0	0	0	0	0	0	0	1



The results for the Bellman's Equation, $v(x)$ is

0	1_1	1_2	1_3	2_1	2_2	2_3	3_1	3_2	3_3	e
2.14	2.40	3.1	2.44	3	2.32	2	1.74	2	3.5	0

Remark: I would like to emphases the major differences between the Elimination and Full Size Matrix Algorithm. The first algorithm eliminates states one at a time. Each time a state is removed, it also reduces that state space. However, the latter algorithm can eliminate a state or set of states. In this case, the state space is preserved.

Comment: Keep in mind, if the parameters values change, we may need to wait until Markov Chain enters the absorbing state. This implies it is not optimal to stop.

Beside the Bellman and the Poisson Equations, the Elimination Algorithm become the basis for the recursive procedure. Along with a probabilistic interpretation, this is the driving force for the calculation the *Gittins index* $\gamma(x)$.

$$\gamma(x) = \sup_{\tau>0} \frac{E_x \sum_{n=0}^{\tau-1} \beta^n c(Z_n)}{E_x \sum_{n=0}^{\tau-1} \beta^n}$$

where β is a discount factor, $0 < \beta < 1$, and τ is the stopping time.

CHAPTER 5: SPECTRAL ANALYSIS OF THE COVARIANCE OPERATOR OF "LOOP" MARKOV CHAINS

I will start this chapter with calculating the invariant distribution for the "Loop" Markov Chain Model. Let's assume that $\pi = \pi(y)$ is a stationary (or invariant) distribution, i.e. an unique probabilistic solution of the equation $\pi P = \pi$.

Lemma 5.1. *: Invariant distribution for the LMC is given by*

$$\pi(0) = \frac{1}{\sum\limits_{i=1}^{k} p_i n_i}; \pi(i,j) = \frac{p_i}{\sum\limits_{i=1}^{k} p_i n_i} = \frac{p_i}{m_0} \tag{5.1}$$

where $i = 1, 2, \ldots, k$ and $j = 1, 2, \ldots, n_i - 1$. Here $m_0 = \sum p_i n_j = E_0 \tau_0$ and τ_0 is the time of the first return from $0 \to 0$.

Proof. Let $p_1, p_2, \ldots, p_k$ be the probability of entering a certain loop. The GCF $(n_1, \ldots, n_k) = 1$ therefore chain is aperiodic.

$$\sum_{i,x} \pi_i(x) = 1 \quad \Rightarrow \quad \pi(0) + \sum_{i=1}^{k} \pi(0) p_i (n_i - 1) = 1$$

Finally

$$\pi(0) = \frac{1}{1 + \sum\limits_{i=1}^{k} p_i (n_i - 1)} = \frac{1}{1 + \sum\limits_{i=1}^{k} (p_i n_i - p_i)} = \frac{1}{1 + \sum\limits_{i=1}^{k} p_i n_i - 1} = \frac{1}{\sum\limits_{i=1}^{k} p_i n_i}$$

$$\Rightarrow \pi(0) = \frac{1}{\sum\limits_{i=1}^{k} p_i n_i} \text{ and } \pi_i(x) = \frac{p_i}{\sum\limits_{i=1}^{k} p_i n_i}$$

Q.E.D.

The spectral analysis of the covariance operator B is a natural and interesting problem with the potential for statistical applications. If instead of observation of the chain $x(t)$ we can observe only the additive functional $S_n = \sum_{s=0}^{n-1} f(x_s)$ then (for fixed L^2- norm $\|f\|_\pi = 1$) the biggest amount of the information about chain will provide the top eigenfunction Ψ_0 of $B : \lambda_0\Psi_0 = B\Psi_0$, $\lambda_0 = \max \lambda_i(B)$ (it is the leading factor in the statistical analysis). From another side, function $\Psi \in \ker B = \{\Psi : B\Psi = 0\}$ are also interesting for such functions as S_n which are bounded in probability if $n \to \infty$. The following theorem gives the complete characterization of $\ker B$.

Theorem 5.2. *In the situation in section 2, ker* $B = Span(f : f = g - Pg)$ *and* $g = P^*Pg$, *i.e.* $g \in \ker(I - P^*P)$.

Let's stress that $I - P^*P$ is symmetric, bounded, and nonnegative on $L^2(x, \pi)$ due to the fact that P^*P is symmetric and stochastic (with top eigenvalue of 1).

Proof. Assume that $\sigma^2(f) = 0 \Longrightarrow f = g - Pg$ and $(g \cdot g)_\pi - (Pg, Pg)_\pi = 0$, i.e. $(g, g)_\pi = (g,$ $P^*Pg) \Longrightarrow (g,\ (I - P^*P)g) = 0$. But $(I - P^*P)$ is symmetric and nonnegative and the spectral theorem for $(I - P^*P)$ gives immediately that $g \in \ker\ (I - P^*P)$. *Q.E.D.*

Let's apply this general result to the LMC.

In the future $f(x)$, $g(x)$ will be the generic function from $L^2\,(X, \pi)$, f_0, f_{ij}, g_0, g_{ij} will be the particular functions from $L^2\,(X, \pi)$. F will be the functions constant on the loops and G are functions linear on the loops.

Theorem 5.3. $\ker B = span(f_0; f_{ij}, i = 1, 2, .., k;\ j = 1, 2, ..., n_i - 2)$, $\dim \ker\,(B) = \sum_{i=1}^{k} n_i -$ $2k + 1$

Here

$$f_0 = g_0 - Pg_0,$$

$$g_0(y) = \delta_0(y) = \begin{cases} 1 & \text{if} \quad y = 0 \\ 0 & \text{if} \quad y \neq 0 \end{cases} \qquad f_0(y) = \begin{cases} 1 & \text{if} \quad x = 0 \\ -p_i & \text{if} \quad x = (i, n_i - 1) \\ 0 & \text{otherwise} \end{cases}$$

For $1 \leq j \leq n_i - 2, \;\; i = 1, 2, ..., k$

$$g_{ij}(y) = \delta_{(i,j)}(y), \qquad f_{ij} = \begin{cases} -1 & \text{if} \quad x = 0 \\ 1 & \text{if} \quad x = (i, j) \\ 0 & \text{otherwise} \end{cases}$$

Proof. Let's calculate the stochastic matrix $Q = P^*P$ which corresponds to the following Markov Chain. Assume that the initial chain $x(t)$ starts from $x \in X$. Let's make one step with transition probabilities $p^*(y, z)$ (negative direction at that time) and after one step with probabilities $p(x, y)$ (positive direction at that time). These two steps give the single transition of $\bar{x^+}$from x to z (with probability $q(x, y) = \sum\limits_{z \in X} p^*(x, z)p(z, y)$. Direct calculations using formula (4) gives

$$\begin{aligned} p^*((i, l), (i, l - 1)) &= 1 \qquad l = 2, ..., n_i - 1, \;\; i = 1, 2, ..., k; \\ p^*((i, 1), 0) &= 1 \qquad i = 1, 2, ..., k; \\ p^*(0, (i, n_i - 1)) &= p_i \qquad i = 1, 2, ..., k. \end{aligned}$$

Now from the formula $q(x, y) = \sum\limits_{z \in X} p^*(x, z)p(z, y)$ we get

$$\begin{aligned} q((i, l), (i, l)) &= 1 \qquad l = 2, ..., n_i - 1, \;\; i = 1, 2, ..., k; \\ q((i, 1), (j, 1)) &= p_j \qquad i, j = 1, , ..., k; \\ q(0, 0) &= 1 \end{aligned}$$

It means that for the chain $\bar{x^+}(t)$ with transition matrix Q the state 0; (i, l), $i = 1, 2, ..., k$;

$l = 1, 2, ...n_i - 2$ are absorbing ones and the remaining states set $\Gamma = \{(i, n_i - 1),\ i = 1, 2, ..., k\}$ are the transient ones. *Q.E.D.*

It is well-known (but in our particular can obvious) that all Q-harmonic function, i.e. the solutions of the equation $g = Qg$, or $(I - P^*P)g = 0$ are constant on the ergodic classes of the chain and $\dim\{g : g = Qg\}$ is equal to the number of such classes; i.e.

$$\ker(I - P^*P) = span(g = \delta_0(y), \quad g(y) = \delta_{(i,j)}(i, y), \quad i = 1, 2, ..., k; \quad j = 2, 3, ...n_i - 1).$$

As a result

$$\dim\ \ker(I - P^*P) = 1 + \sum_{i=1}^{k}(n_i - 2) = \sum_{i=1}^{k} n_i - 2k + 1$$

For each function $g_0, g_{(i,j)}$ one can calculate the corresponding f, such that $\sigma^2(f) = 0$. Namely

$$f_0(y) = (g_0 - Pg_0(y)) = \begin{cases} 1 & \text{if} \quad y = 0 \\ -1 & \text{if} \quad y = (i, n_i - 1), \quad i = 1, ..., k \end{cases}$$

$$f_{(i,j)}(y) = \begin{cases} 1 & \text{if} \quad y = (i, j) \\ -1 & \text{if} \quad y = (i, j - 1), \quad j = 2, ..., n_i - 1, \quad i = i, ..., k \end{cases}$$

Our goal now is the description of function $ImB = (\ker B)^{\perp}$ in $L^2(X, \pi)$, containing the functions $f : \sigma(f) = (Bf, f)_{\pi} > 0$.

Theorem 5.4. $\dim(\ker)^{\perp} = k - 1$. *It consists of the functions* Γ*, const on the loops*

$$F(x) = \begin{cases} \varphi_i & \textit{if} \quad x = (i, l), \quad l = 1, 2, ..., n_i - 1 \\ F(0) & \textit{if} \quad x = 0 \end{cases}$$

and restricted by the linear conditions $F(0) = \sum\limits_{i=1}^{k} p_i\varphi_i,\ \sum\limits_{i=1}^{k} p_i n_i \varphi_i = 0.$

Proof. Condition of orthogonality $(F \cdot f_{(i,l))_{\pi}} = 0$ gives $F(i, l-1) = F(i, l); l = 2, ...n_i - 1$. The additional condition of orthogonality $(F \cdot f_0)_{\pi} = 0$ provides the first relation $F(0) = \sum\limits_{i=1}^{k} p_i\varphi_i$.

Finally due to definition of $L^2(X,\pi)$, $\sum F(x)\pi(x) = 0$, i.e. $F(0) + \sum\limits_{i=1}^{n} (n_i - 1)\, p_i\varphi_i = 0$. Taking into account the previous relation we'll get $\sum\limits_{i=1}^{k} n_i p_i \varphi_i = 0$. *Q.E.D.*

For each function $F \in (\ker B)^{\perp}$ we can solve in $L^2(X,\pi)$ the homological equation $F = G - PG$ and calculate $\sigma^2(F) = (BF,F)_\pi = (G,G)_\pi - (PG,PG)_\pi$. Let

$$G_i = G(i,1)$$

then

$$G - PG = F$$

gives

$$G(i,1) - G(i,2) = \varphi_i, G(i,2) - G(i,3) = \varphi_i, ... \Rightarrow G(i,l) = G_i - (l-1)\varphi_i, i = 1,2,...,n_i - 1.$$

Since

$$G(i, n_i - 1) - G(0) = \varphi_i$$

we have

$$G(0) = G_i - (n_i - 1)\varphi_i$$

i.e.

$$G_i = G(0) + (n_i - 1)\varphi_i.$$

It is easy to see

$$G(0) - (PG)(0) = G(0) - \sum_{i=1}^{k} p_i \left(G_0 + (n_i - 1)\,\varphi_i\right) = G(0) - G(0) - \sum_{i=1}^{k} p_i n_i \varphi + \sum_{i=1}^{k} p_i \varphi = F(0).$$

The following diagram presents all three functions F, G, PG.

Diagram 1:$F(x) \quad \sum\limits_{i=1}^{k} \varphi_i p_i = F(0), \quad \sum\limits_{i=1}^{k} \varphi_i p_i n_i = 0$

F is constant on each horizontal edge.

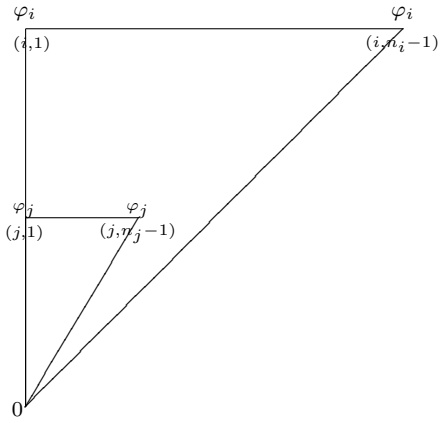
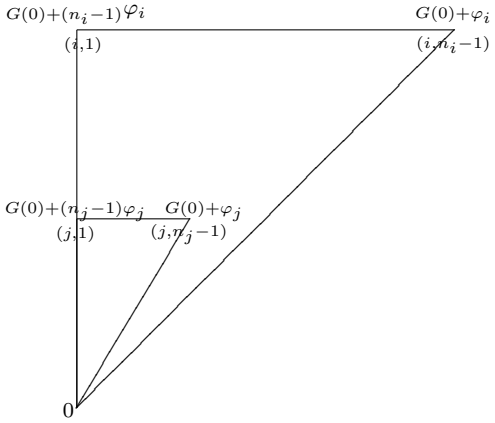


Diagram 2:$G(x)$

$G(.)$ is linear on each horizontal edge.



Since $G(x) \in L^2(x,\pi)$ for $G(0)$ we have the equation

$$G(0) + \sum_{i=1}^{k} p_i \sum_{j=1}^{n_i-1} (G(0) + j\varphi_i) = 0, \text{ i.e}$$

$$G(0) + \sum_{i=1}^{k} p_i G(0)(n_i - 1) + \sum_{i=1}^{k} \varphi_i p_i \frac{n_i(n_i-1)}{2} = 0,$$

$$G(0)m(0) + \sum_{i=1}^{k} \varphi_i p_i \frac{n_i(n_i-1)}{2},$$

$$G(0) = -\frac{1}{m(0)} \sum_{i=1}^{k} \frac{\varphi_i p_i n_i^2}{2}$$

(in the last row, we used the relation $\sum_{i=1}^{k} \varphi_i p_i n_i = 0$). Also

$$PG(0) = \sum_{i=1}^{k} p_i(G(0) + (n_i - 1)\varphi_i) = G(0) - \sum_{i=1}^{k} \varphi_i p_i = G(0) - F(0)$$

For $\sigma^2(F)$ we have

$$\sigma^2(F) = (G, G)_\pi - (PG, PG)_\pi$$

$$= \frac{G^2(0) - (G(0) - F(0))^2}{m(0)} + \frac{1}{m(0)} \sum_{i=1}^{k} p_i \sum_{l=1}^{n_i - 1} (G(0) + l\varphi_i)^2 - (G(0) + (l-1)\varphi_i)^2$$

$$= \frac{2G(0)F(0) - F^2(0)}{m(0)} + \frac{1}{m(0)} \sum_{i=1}^{k} p_i \sum_{i=1}^{k} \varphi_i \left(2G(0) + (2l-1)\,\varphi_i\right)$$

$$= \frac{2G(0)F(0) - F^2(0)}{m(0)} + \frac{1}{m(0)} \sum_{i=1}^{k} \varphi_i p_i \left(2G(0)(n_i - 1) + \varphi_i (n_i - 1)^2\right)$$

$$= \frac{2G(0)F(0) - F^2(0)}{m(0)} - \frac{2G(0)F(0)}{m(0)} + \frac{1}{m(0)} \sum_{i=1}^{k} \varphi_i^2 p_i \left(n_i - 1\right)^2$$

$$= \frac{1}{m(0)} \left[- \left(\sum_{i=1}^{k} \varphi_i p_i \right)^2 + \sum_{i=1}^{k} \varphi_i^2 p_i \left(n_i - 1\right)^2 \right]$$

It is better to present the final answer in the different form (using the relation $\sum\limits_{i=1}^{k} \varphi_i p_i n_i = 0$):

$$\sigma^2(F) = \frac{1}{m(0)} \left[\sum_{i=1}^{k} \varphi_i^2 p_i \left(n_i - 1\right)^2 - \left(\sum_{i=1}^{k} \varphi_i p_i \left(n_i - 1\right) \right)^2 \right]$$

$$= \frac{1}{m(0)} \sum_{i=1}^{k} p_i \left[l_i \left(n_i - 1\right) - l \right]^2 \quad \text{where } l = \sum_{i=1}^{k} \varphi_i p_i \left(n_i - 1\right) = -F(0)$$

One can see now that $\sigma^2(F) = 0$ if and only if $\varphi_i = \frac{F(0)}{n_i - 1}$, $i = 1, 2, ..., k$. Vector $\overrightarrow{\varphi}_0 = \left[-1, \frac{1}{n_i - 1}, ..., \frac{1}{n_k - 1} \right]$ is an eigenvector for the covariance matrix B with eigenvalue $\lambda_0 = 0$. The orthogonal complement to this vector if given by

$$(F, \varphi_0)_\pi = 0 \Rightarrow -F(0) + \sum_{i=1}^{k} p_i \frac{\varphi_i (n_i - 1)}{n_i - 1} \Rightarrow F(0) = \sum_{i=1}^{k} p_i \varphi_i$$

In other terms under conditions $F(0) = \sum\limits_{i=1}^{k} p_i \varphi_i$ the $\sigma^2(F) > 0$ and $rank\ B = (k-1)$.

Appendix A

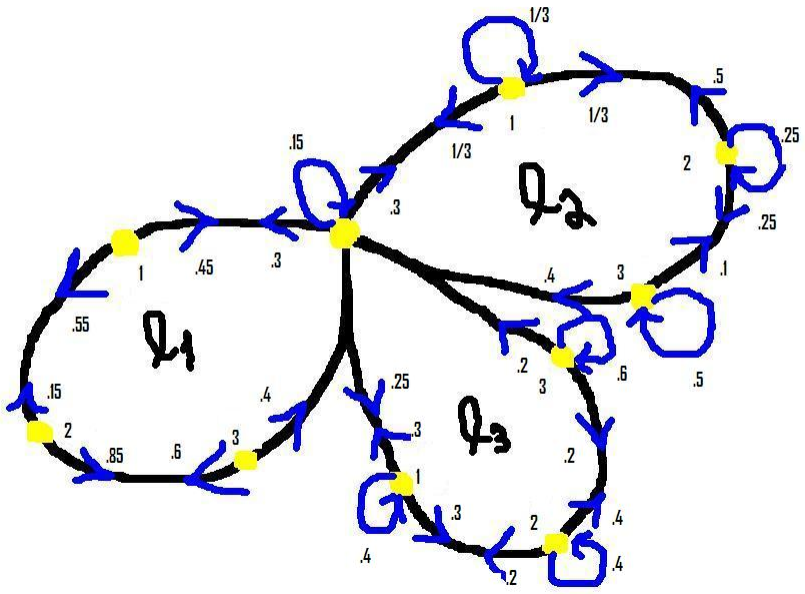


Figure 1: Loop Example

Example .1. Lift Algorithm

This is the original information in tableau form.

	0	**1, 1**	**1, 2**	**1, 3**	**2, 1**	**2, 2**	**2, 3**	**3, 1**	**3, 2**	**3, 3**	**g**
0	.15	.3	0	0	.3	0	0	.25	0	0	1
1, 1	.45	0	.55	0	0	0	0	0	0	0	2
1, 2	0	.15	0	.85	0	0	0	0	0	0	3.1
1, 3	.4	0	.6	0	0	0	0	0	0	0	$-.5$
2, 1	1/3	0	0	0	1/3	1/3	0	0	0	0	3
2, 2	0	0	0	0	.5	.25	.25	0	0	0	1
2, 3	.4	0	0	0	0	.1	.5	0	0	0	2
3, 1	.3	0	0	0	0	0	0	.4	.3	0	-1
3, 2	0	0	0	0	0	0	0	.4	.4	.2	2
3, 3	.2	0	0	0	0	0	0	0	.2	.6	3.5

We will consider $c = 0$ and $\beta = .9$. Remember an extra state is added to handle the

discount factor. The states' row which are shaded in the tableau or blacked out in figure 2 are the ones which should be eliminated.

	0	**1, 1**	**1, 2**	**1, 3**	**2, 1**	**2, 2**	**2, 3**	**3, 1**	**3, 2**	**3, 3**	**e**	**g**	**g − Pg**
0	.13	.27	0	0	.27	0	0	.23	0	0	.1	1	−.26
1, 1	.405	0	.495	0	0	0	0	0	0	0	.1	2	−1.9
1, 2	0	.135	0	.765	0	0	0	0	0	0	.1	3.1	.113
1, 3	.36	0	.54	0	0	0	0	0	0	0	.1	−.5	−2.03
2, 1	.3	0	0	0	.3	.3	0	0	0	0	.1	3	1.5
2, 2	0	0	0	0	.45	.22	.23	0	0	0	.1	1	−1.03
2, 3	.36	0	0	0	0	.09	.45	0	0	0	.1	2	.65
3, 1	.27	0	0	0	0	0	0	.36	.27	0	.1	−1	−1.45
3, 2	0	0	0	0	0	0	0	.36	.36	.18	.1	2	1.01
3, 3	.18	0	0	0	0	0	0	0	.18	.54	.1	3.5	1.07
e	0	0	0	0	0	0	0	0	0	0	1	0	0

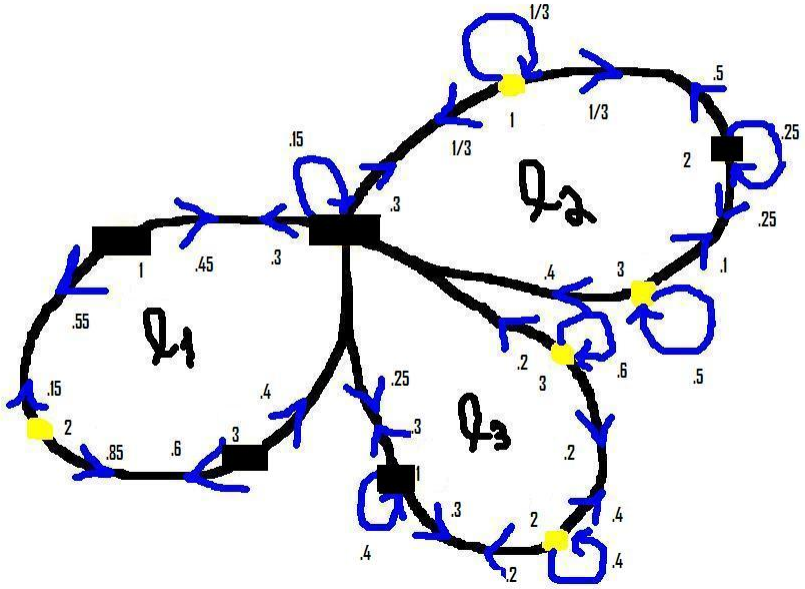


Figure 2: States that Need to be Eliminated

From the above tableau, the origin needs to be eliminated. In order to keep this state, we will "Lift" $g(0)$ from -.26 to .73 and recalculate $g - Pg$

	0	**1, 1**	**1, 2**	**1, 3**	**2, 1**	**2, 2**	**2, 3**	**3, 1**	**3, 2**	**3, 3**	**e**	**g**	**g − Pg**
0	.13	.27	0	0	.27	0	0	.23	0	0	.1	1	.73
1, 1	.405	0	.495	0	0	0	0	0	0	0	.1	2	−2.40
1, 2	0	.135	0	.765	0	0	0	0	0	0	.1	3.1	.113
1, 3	.36	0	.54	0	0	0	0	0	0	0	.1	−.5	−2.44
2, 1	.3	0	0	0	.3	.3	0	0	0	0	.1	3	1.16
2, 2	0	0	0	0	.45	.22	.23	0	0	0	.1	1	−1.03
2, 3	.36	0	0	0	0	.09	.45	0	0	0	.1	2	.24
3, 1	.27	0	0	0	0	0	0	.36	.27	0	.1	−1	−1.76
3, 2	0	0	0	0	0	0	0	.36	.36	.18	.1	2	1.01
3, 3	.18	0	0	0	0	0	0	0	.18	.54	.1	3.5	.86
e	0	0	0	0	0	0	0	0	0	0	1	0	0

After the "lift" algorithm is implemented, the origin is no longer a candidate for elimination. We will execute the "full size matrix" algorithm to get the final result.

	0	**1, 1**	**1, 2**	**1, 3**	**2, 1**	**2, 2**	**2, 3**	**3, 1**	**3, 2**	**3, 3**	**e**
0	.34	0	.13	0	.27	0	0	0	.10	0	.16
1, 1	.405	0	.495	0	0	0	0	0	0	0	.1
1, 2	.33	0	.48	0	0	0	0	0	0	0	.19
1, 3	.36	0	.54	0	0	0	0	0	0	0	.1
2, 1	.3	0	0	0	.47	0	.09	0	0	0	.14
2, 2	0	0	0	0	.58	0	.29	0	0	0	.13
2, 3	.36	0	0	0	.05	0	.48	0	0	0	.11
3, 1	.42	0	0	0	0	0	0	0	.42	0	.16
3, 2	.15	0	0	0	0	0	0	0	.51	.18	.16
3, 3	.18	0	0	0	0	0	0	0	.18	.54	.1
e	0	0	0	0	0	0	0	0	0	0	1

The results for the Bellman's equation, $v(x)$ are

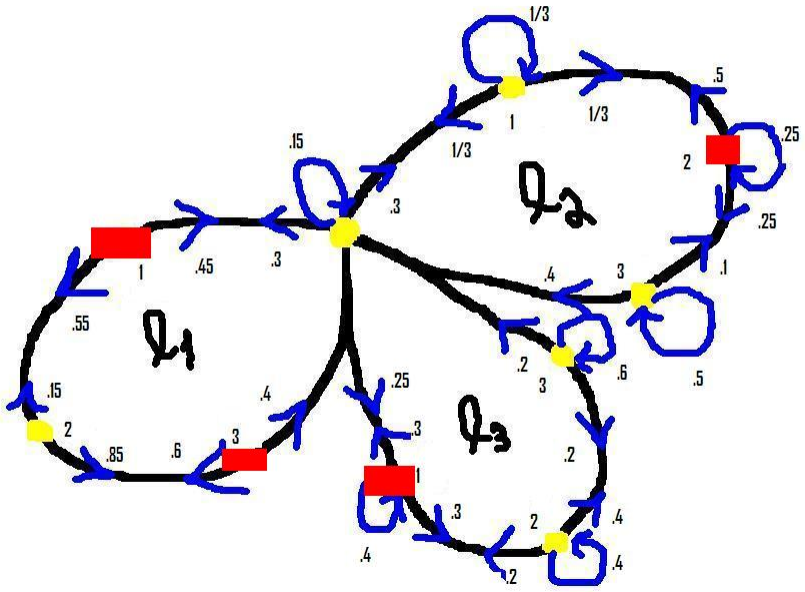


Figure 3: Final Solution

0	1, 1	1, 2	1, 3	2, 1	2, 2	2, 3	3, 1	3, 2	3, 3	e
2.14	2.40	3.1	2.44	3	2.32	2	1.74	2	3.5	0

Example .2. Full Size Matrix Algorithm

This is the original information in tableau form.

	0	**1, 1**	**1, 2**	**1, 3**	**2, 1**	**2, 2**	**2, 3**	**3, 1**	**3, 2**	**3, 3**	**g**
0	.15	.3	0	0	.3	0	0	.25	0	0	1
1, 1	.45	0	.55	0	0	0	0	0	0	0	2
1, 2	0	.15	0	.85	0	0	0	0	0	0	3.1
1, 3	.4	0	.6	0	0	0	0	0	0	0	−.5
2, 1	1/3	0	0	0	1/3	1/3	0	0	0	0	3
2, 2	0	0	0	0	.5	.25	.25	0	0	0	1
2, 3	.4	0	0	0	0	.1	.5	0	0	0	2
3, 1	.3	0	0	0	0	0	0	.4	.3	0	−1
3, 2	0	0	0	0	0	0	0	.4	.4	.2	2
3, 3	.2	0	0	0	0	0	0	0	.2	.6	3.5

We will consider $c = 0$ and $\beta = .9$. Remember an extra state is added to handle the

discount factor. The states in **red** are the ones which need to be eliminated.

	0	**1, 1**	**1, 2**	**1, 3**	**2, 1**	**2, 2**	**2, 3**	**3, 1**	**3, 2**	**3, 3**	**e**	**g**	**g − Pg**
0	.13	.27	0	0	.27	0	0	.23	0	0	.1	1	−.26
1, 1	.405	0	.495	0	0	0	0	0	0	0	.1	2	−1.9
1, 2	0	.135	0	.765	0	0	0	0	0	0	.1	3.1	.113
1, 3	.36	0	.54	0	0	0	0	0	0	0	.1	−.5	−2.03
2, 1	.3	0	0	0	.3	.3	0	0	0	0	.1	3	1.5
2, 2	0	0	0	0	.45	.22	.23	0	0	0	.1	1	−1.03
2, 3	.36	0	0	0	0	.09	.45	0	0	0	.1	2	.65
3, 1	.27	0	0	0	0	0	0	.36	.27	0	.1	−1	−1.45
3, 2	0	0	0	0	0	0	0	.36	.36	.18	.1	2	1.01
3, 3	.18	0	0	0	0	0	0	0	.18	.54	.1	3.5	1.07
e	0	0	0	0	0	0	0	0	0	0	1	0	0

We will execute the "full size matrix" algorithm and get the final result.

	0	**1, 1**	**1, 2**	**1, 3**	**2, 1**	**2, 2**	**2, 3**	**3, 1**	**3, 2**	**3, 3**	**e**
0	0	0	.20	0	.41	0	0	0	.14	0	.25
1, 1	0	0	.58	0	.17	0	0	0	.05	0	.2
1, 2	0	0	.55	0	.13	0	0	0	.05	0	.27
1, 3	0	0	.61	0	.15	0	0	0	.05	0	.19
2, 1	0	0	.06	0	.60	0	.09	0	.04	0	.21
2, 2	0	0	0	0	.58	0	.29	0	0	0	.13
2, 3	0	0	.07	0	.20	0	.48	0	.05	0	.20
3, 1	0	0	.09	0	.17	0	0	0	.48	0	.26
3, 2	0	0	.03	0	.07	0	0	0	.53	.18	.19
3, 3	0	0	.04	0	.07	0	0	0	.21	.54	.14
e	0	0	0	0	0	0	0	0	0	0	1

The results for the Bellman's equation, $v(x)$ are

0	$1,1$	$1,2$	$1,3$	$2,1$	$2,2$	$2,3$	$3,1$	$3,2$	$3,3$	e
2.14	2.40	3.1	2.44	3	2.32	2	1.74	2	3.5	0

REFERENCES

[1] Bellman, Richard E. (1957), *Dynamic Programming*, Princeton: Princeton University Press

[2] Billingsley, Patrick (1968), *Convergence of the Probability Measure*, J. Wiley, New York

[3] Bonnans, J. Frdric; Shapiro, Alexander (2000), *Perturbation analysis of optimization problems*, Springer Series in Operations Research. New York: Springer-Verlag.

[4] Boyd, Stephen and Vandenberghe, Lieven (2004). *Convex Optimization*, Cambridge University Press.

[5] Chung, Kai Lai (1960), *Markov Chains with Stationary Transition Probabilities*, section 1.15, Springer-Verlag

[6] Denardo, Eric V. (2002), *The Science of Decision Making: A Problem-Based Approach Using Excel*, John Wiley and Sons Inc., New York

[7] Feldman, Richard M. and Valdez-Flores, Ciriaco (1996), *Applied Probability and Stochastic Processes*, PWS Pubishing Company, Boston, MA

[8] Feller, William (1957), *An Introduction to Probability Theory and Its Applications Volume I*, John Wiley and Sons Inc., New York

[9] Ferguson, Thomas S. (1967), *Probability and Mathematical Statistics, A Series of Monographs and Textbooks, Mathematical Statistics, A decision Theoretic Approach*, Academic Press, New York and London

[10] Ibe, Oliver C. (2009), *Markov Processes for Stochastic Modeling*, Elsevier Academic Press

[11] Isaacson, Dean L. and Madsen, Richard W. (1979), *Markov Chains Theory and Applications*, John Wiley and Sons, New York, London and Toronto

[12] Karlin, Samuel and Taylor, Howard M. (1975), *A First Course In Stochastic Processes, 2nd Edition*, Academic Press, London

[13] J. G. Kemeny and J. L. Snell(1960), *Finite Markov Chains*, Van Nostrand Reinhold, Princeton, NJ.

[14] Mordecai Avriel (2003), *Nonlinear Programming: Analysis and Methods*, Dover Publishing.

[15] J. Neveu, *Foundation of Probability Theory*

[16] Nocedal, Jorge and Wright, Stephen J. (2006), *Numerical Optimization*, Springer.

[17] Norris, J. R.(1997), *Cambridge Series in Statistical and Probabilistic Mathematics*, Cambridge University Press, New York, New York

[18] Olofsson, Peter (2005), *Probability, Statistics, and Stochastic Processes*, J. Wiley, New York

[19] Puterman, Martin L., *Markov Decision Processes, Discrete Stochastic Dynamic Programming*, Wiley, New York

[20] Ross, Shelton M. (2003), *Probability Models, 8th Edition*, Academic Press, London

[21] T. J. Sheskin(1999), *State Reduction in a Markov Decision Process.*, Internet. J. Mathematics Eduation Science Technology, 30, 2, 167 - 185

[22] A.N. Shiryayev, *Statistical Seqential Analysis: Optimal Stopping Rules, (Russian) Izdat.* "Nauka", Moscow, pp. 231

[23] Sonin, Isaac (1995), *Two Simple Theorems in the Problem of Optimal Stopping*, INFORMS Applied Probability Coference, Atlanta, Georgia, pp. 27 -28

[24] Sonin, Isaac(1996), *Increasing the Reliablity of a Machine Reduces the Period of its Work*, J. Applied Probabilty 33, 1, 217 -223

[25] Sonin, Isaac(1999), *The Elimination Algorithm for the Problem of Optimal Stopping*, Mathematics Methods of Operation Research, 4, 1, 111 - 123

[26] Sonin, Isaac(1999), *The State Reduction and Related Algorithms and their applications to the study of Markov Chains, graph theory and the Optimal Stopping Problem*, Advances in Mathematics, 145, 159 - 188

[27] Sonin, Isaac(2006), *The Optimal stopping of Markov Chains and Recursive Solution of Poisson and Bellman Equations, In:, From Stochastic Calculus to Mathematical Finance*, The Shiryeav Festchrift, Kabanov Y., Lipster, R.; Stoyanov, J. (Eds), pp.609 - 621

[28] Sonin, Isaac and Thornton, John, *Recursive Algorithm for the Fundamental/Group Inverse Matrix of a Markov Chain from an Explicit Formula*, SIAM J. Matrix Anal. Appl., Vol. 23, No. 1, pp. 209 - 224

[29] Howard M. Taylor and Samuel Karlin(1998), *Introduction to Stochastic Modeling, 3rd Edition*, Academic Press, San Diego, California

[30] D. J. White(1993), *Markov Decision Processes*, J. Wiley, New York

[31] William J. Stewart(2009), *Probability, Markov Chains, Queues, and Simulation*, Princetion University, Princeton, New Jersey

[32] William J. Stewart, *Introduction to the Numerical Solution of Markov Chains*, Princetion University, Princeton, New Jersey

[33] Yang X.-S. (2008), *Introduction to Mathematical Optimization: From Linear Programming to Metaheuristics*, Cambridge Int. Science Publishing.