Adam CZORNIK

ON CONTROL PROBLEMS FOR JUMP LINEAR SYSTEMS

WYDAWNICTWO POLITECHNIKI ŚLĄSKIEJ

GLIWICE 2003

Adam CZORNIK

ON CONTROL PROBLEMS FOR JUMP LINEAR SYSTEMS

WYDAWNICTWO POLITECHNIKI ŚLĄSKIEJ
GLIWICE 2003

AUTOMATYKA z. 137
Nr kol. 1579
ISSN 0434-0760

CONTENTS

SPIS TREŚCI

Introduction

Modern control systems must meet performance requirements and meantime acceptable behavior even in the presence of abrupt changes in their dynamics due, for instance, to random component failures or repairs, environmental disturbances, dramatic changes in subsystem interconnections, sudden changes in the operating point of nonlinear plant etc. This can be found, for instance, in control of solar systems, robotic manipulator systems, aircraft control systems, large flexible structures for space stations (such as antennas, solar arrays), etc. If abrupt changes have only a small influence on the system behavior, classical sensitivity analysis may provide an adequate assessment of the effects. When the variations caused by the dynamic changes significantly alter the behavior of the system, a stochastic model that gives a quantitative indication of the relative likelihoods of various possible scenarios would be preferable. In some cases the relevant stochastic model may consist of a linear system with coefficients depending on certain stochastic process. Such models are called jump linear systems. They first appeared in the literature in papers [70] and [49] for continuous time systems. Since then much has been done and a good summary of the results obtained up to 1990 may be found in monograph [82]. Nevertheless still there exists a number of open questions in this area and many results need refinement and improvement. To justify the interest in analysis of this class of control systems let us focus attention on the exemplary situations in which the theory of jump linear systems seems to be the most applicable approach for solving the problem. In order to increase the reliability in the presence of emergency of failures, the control system must provide some kind of fault tolerance. Since von Neuman, we know that redundancy is the basic ingredient in building a reliable systems. The fault-prone control systems have attracted a significant research effort and the basic structure of a reliable system with redundant actuators and sensors, which is now agreed upon is presented in the Figure 0.1.

To analyze the behavior of such a system we clearly need a description of the occurrence of failures and their influence on the process. Consider, for example, a duplex system where two redundant controllers, C_1 and C_2 are used in parallel to control the plant P. Excluding partial failures of a component, four regimes of operation, can be associated to Figure 0.2 depending on which controller has failed:

$$\begin{array}{lll} \text{regime 1} & C_1 & C_2 \\ \text{regime 2} & \underline{C_1} & C_2 \\ \text{regime 3} & C_1 & \underline{C_2} \\ \text{regime 4} & \underline{C_1} & \underline{C_2} \end{array}$$

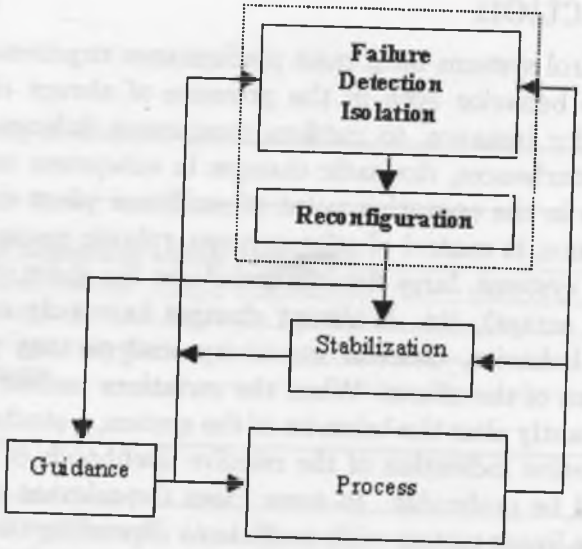

Fig. 0.1. A reliable control system

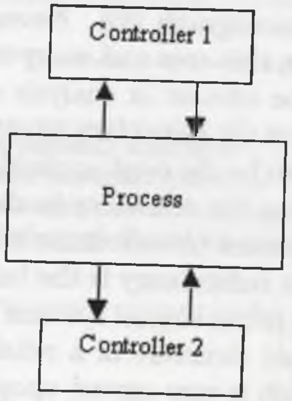

Fig. 0.2. A duplex control system

Failures thus appear as discrete events that cause a transition, or a jump, of the regime. These events being random, are characterized by transition probabilities $(p_1(n), p_2(n), p_3(n), p_4(n))$, where $p_i\ (n)$ is the probability of regime to be i at moment n. It is reasonable to assume that $p_i\ (n)$ depends on the state of regime at moment $n-1$ and does not depend on states of regimes at moment $k \leq n-2$. In this way we describe regimes as a Markov chain with transition probability

$$P = [p_{ij}]_{i,j=1,\ldots,4}\,,$$

where p_{ij} is the probability that the regime is i under condition that the previous regime was j. A classical model in reliability theory is given as

$$P = \begin{bmatrix} e^{-2\lambda\Delta} & e^{-\lambda\Delta} - e^{-2\lambda\Delta} & e^{-\lambda\Delta} - e^{-2\lambda\Delta} & 1 - 2e^{-\lambda\Delta} + e^{-2\lambda\Delta} \\ 0 & e^{-1} & 0 & 1 - e^{-\lambda\Delta} \\ 0 & 0 & e^{-\lambda\Delta} & 1 - e^{-\lambda\Delta} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

where $\lambda\Delta$ is the individual failure rate of C_1 and C_2 at time interval Δ. In writing P a simultaneous failure of C_1 and C_2 was excluded (the transition from 1 to 4 has zero probability) which corresponds to the realistic situation where failures are rare events and the occurrence of a simultaneous failure is highly unlikely. The failures manifest themselves through a modification of the actuator-process-sensors cascade in figure 0.2. Typically, the failure of a sensor introduces a bias or a drift in one of the measurement variables. Similarly, a failed actuator might produce a constant action on the process regardless of the command signal it receives. A physical fault in some parts of the plant also modifies the dynamics.

Yet another example of this kind has been discussed in [18]. This example deals with reliable control system design. The control system is described by

$$x\,(k+1) = Ax\,(k) + B\,(r(k))\,u\,(k)\,,$$

where

$$A = \begin{bmatrix} 2,71828 & 0 \\ 0 & 0.36788 \end{bmatrix}, \ B(1) = \begin{bmatrix} 1.71828 & 1.71828 \\ -0.63212 & 0.63212 \end{bmatrix}$$

$$B(2) = \begin{bmatrix} 0 & 1.71828 \\ 0 & 0.63212 \end{bmatrix}, \ B(3) = \begin{bmatrix} 1.71828 & 0 \\ -0.63212 & 0 \end{bmatrix}$$

$$B(4) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

This model captures the failure/repair events for a reliable system with two actuators, in which actuators may fail and need to be repaired. State 1 of $r(k)$ represents the case that both actuators work well, states 2 and 3 represent the case where one of actuators fails and has to be repaired, and state 4 represents the case where both actuators fail. Let p_f and p_r denotes the failure rate and repair rate, where the actuators repair and failure events are independent, then the probability transition matrix of Markov chain $r(k)$ is given by

$$P = \begin{bmatrix} (1-p_f)^2 & (1-p_f)\,p_r & (1-p_f)\,p_r & p_r^2 \\ (1-p_f) & (1-p_f)\,(1-p_r) & p_r & p_r\,(1-p_r) \\ (1-p_f)\,p_f & p_r p_f & (1-p_f)\,1-p_r & p_r 1-p_r \\ p_f^2 & (1-p_r)\,p_f & (1-p_r)\,p_f & (1-p_r)^2 \end{bmatrix}.$$

It is thus seen that fault-tolerant control systems are naturally described in terms of Cartesian product of a discrete random jump variable accounting for the occurrence of failures, and the more usual variables in continuous space representing the plant dynamics, i.e. in the terms of system with jumps in parameters.

Another example of this kind was discussed in [92]. Consider a linear system

$$\dot{x}(t) = Ax(t) + Bu(t)$$

with

$$A = \begin{bmatrix} \frac{1}{6} & \frac{1}{5} \\ -\frac{1}{5} & -\frac{1}{6} \end{bmatrix}, \; B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{0.1}$$

The control which minimizes the quadratic cost

$$\int_0^\infty (x'(t)Qx(t) + u'(t)Ru(t))\, dt$$

with

$$Q = \begin{bmatrix} \frac{2}{3} & 0 \\ 0 & \frac{4}{3} \end{bmatrix}, \; R = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

is given by

$$u(t) = -Kx(t),$$

where

$$K = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

The closed loop system is asymptotically stable. In real system the state vector x is not available. Instead of x we have y -the sensor observation of x which is related to x through the relation

$$y(t) = Lx(t),$$

where

$$L = \begin{bmatrix} l_1 & 0 \\ 0 & l_2 \end{bmatrix}$$

The system (0.1) remains stable for control

$$u(t) = -Ky(t)$$

if $l_1 > \frac{1}{2}$, $l_2 > \frac{1}{2}$. Assume that because of sensor failure there is no feedback from x_1 i.e $l_1 = 0$, and the system becomes unstable. Taking into account the possibility of failure of the sensors one can propose as a model the following system

$$\dot{x}(t) = Ax(t) + B(r(t))u(t),$$

where the random variable $r(t)$ takes one of four values 1, 2, 3, 4 and takes value 1 if both sensors work properly, 2 if the first sensor is broken, 3 if the second sensor is broken and 4 if both sensors are broken. The matrices $B(1)$, $B(2)$, $B(3)$, $B(4)$ are as follows

$$B(1) = B, \; B(2) = B \begin{bmatrix} 0 & 0 \\ 0 & l_2 \end{bmatrix}$$

$$B(3) = B \begin{bmatrix} l_1 & 0 \\ 0 & 0 \end{bmatrix}, \; B(4) = 0.$$

Again we are led to a hybrid model with continuous dynamics variables perturbed by random transitions of a regime variable which is discrete.

Further discussion of this kind of applications may be found in the following references [93], [94], [17], [95], [96].

In [98] an application of hybrid models of the similar structure has been proposed to control a solar thermal receiver. In this paper a solar 10 MWe electrical generating system has been described. This system has been build in California desert. A field of movable mirrors is used to focus the sun's energy on a central boiler. One of the most important control loops in the boiler is the steam temperature regulator which controls the feedwater flow rate to maintain the proper outlet steam temperature. The steam temperature regulator has been designed on the basis of a linear dynamic model which evolving from both analytical and empirical studies of the dynamic behavior of a solar-powered central receiver. A difficult problem is posed by the motion of clouds over the heliostats. On a partly cloudy day, the clouds tend to cover the helistats for a time that is quite short when compared to the dominant system time constants. These sudden changes in isolation may be frequent and are essentially unpredictable. The relevant system model, depends on the isolation level, thus changes in a discrete and apparently random fashion. Another anomaly is created by the cloud action. The perturbation variables in the system refer to a set of nominal operating points of the nonlinear equation of the system motion. The system variables are continuous across discontinouities in insolation, while the insolation level changes discretely.

In recent years concern with the safety of air traffic near crowded airports and tragic accidents have prompted research into more sophisticated radar tracking algorithms which could help air controllers monitor incoming and outgoing aircrafts. A specific difficulty appears where there exists a heavy traffic of small highly maneuverable private aircrafts interfering with commercial jet liners.

In a military context, a related problem is the so-called evasive target tracking problem where it is desired to keep track of an object which is maneuvering quickly in an attempt to evade its pursuer. The performance of the tracking system heavily depends on the accuracy and sophistication of the model used to describe the target dynamics.

The first element of the model is based on motion equations, relating variables like horizontal position, heading speed, bank angle or flight path angle. Physic provides a set of differential equations in R^n, where n is the number of variables we retain.

A well-known problem when multiple targets are to be tracked is associating of radar data to the various tracks, which might be difficult when a clutter of signals in a given region exists. However, another less usual phenomenon is of primary interest here, namely the effects of sudden changes in the acceleration of the target. The trajectory can be divided into several sequences when the aircraft flies with (almost) constant acceleration, bank angle and flight path angle; regimes of flight typically considered are ascending flight after take-off, turning, accelerated flight, uniform cruise motion or descending approach flight. The transitions between these regimes are discrete and depend primarily on the pilot decision which are in turn influenced by weather data, mission controller indications and on-board information such as fuel consumption, etc. or perception of the threat. Depending on the current regime of flight chosen by the pilot, the coefficients of the dynamic model have to be adjusted.

Of course evasive maneuvers are chosen to confuse the pursuer, and are therefore characterized by frequent, large, irregular and seemingly irrational acceleration changes. Figure 0.3 shows a typical realization where sudden acceleration is induced by the occurrence of a point process. From the tracker point of view, these transitions are perceived as random, and a model which makes the least use possible of other a priori information is a hybrid stochastic model with continuous dynamics perturbed by random transitions of a regime variable.

Further discussion of this problem is given in [8]-[10].

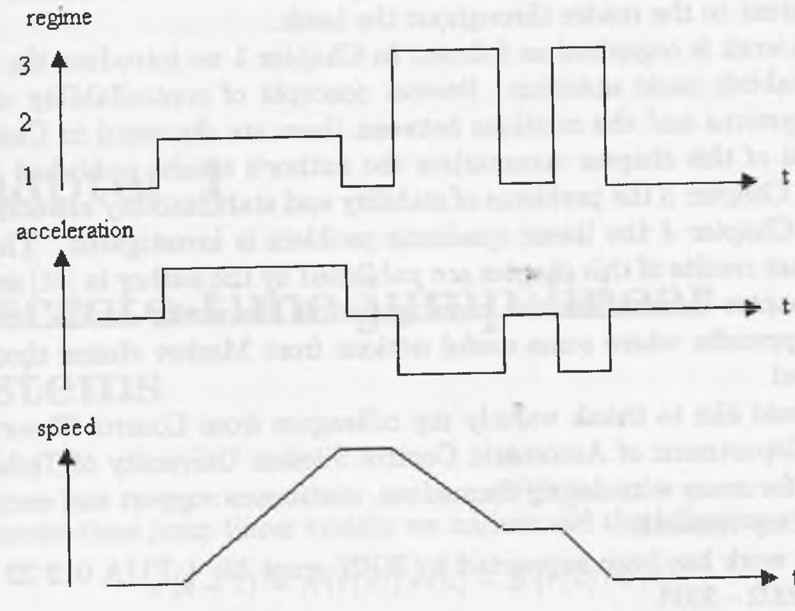

Fig. 0.3. Typical acceleration changes

Other important applications of jump systems are presented for example in the following papers: [85] (placement of failure-prone actuators on large space structures), [3], [16], [21] (control of manufacturing systems), [80], [100] (analysis of transient electrical power networks), [62], (economic policy planing). Athans in [7] suggested that this model setting also has the potential to become a basic framework in posing and solving control-related issues in Battle Management Command, Control, and Communications (BM/C^3) systems.

Most of the published results deal with continuous-time systems. This is natural because the process variables are continuous. However modern digital applications require discrete-time models. The purpose of this work is to present a comprehensive treatment of mathematical aspects of controllability, stability, and linear quadratic problem for discrete-time jump linear systems. In fact various topics for continuous and discrete-time systems are covered in parallel. It is often believed that results for continuous-time systems are also valid for discrete time systems, but it is not always true. Discrete-time systems have their own specification, and there exists a number of crucial and profound differences between the continuous and discrete-time systems.

Controllability belongs to this class of problems. The differences will be transparent to the reader throughout the book.

The work is organized as follows: In Chapter 1 we introduce the system and establish basic notation. Several concepts of controllability of jump linear systems and the relations between them are discussed in Chapter 2. Material of this chapter summarizes the author's results published in [27]-[29]. In Chapter 3 the problems of stability and stabilizability are considered and in Chapter 4 the linear quadratic problem is investigated. The most important results of this chapter are published by the author in [31] and [33]. Next chapter summarizes the investigation of the work. At the end there is an Appendix where some useful notions from Markov chains theory are presented.

I would like to thank warmly my colleagues from Control Theory Unit of the Department of Automatic Control Silesian University of Technology Gliwice for many stimulating discussions, continuous support and encouragement in my research.

This work has been supported by KBN grant No 4 T11A 012 22 in the period 2002 - 2003.

Chapter 1

Discrete-time jump linear systems

By discrete-time jump linear system we understand the following system:

$$x(k+1) = A(r(k))\,x(k) + B(r(k))\,u(k)\,, \tag{1.1}$$

where $x(k) \in R^n$ denotes the process state vector, $k = 0, 1, ...,$ $u(k) \in R^m$ is the control input, $r(k)$ is a Markov chain on a probability space $(\Omega, \mathcal{F}, P)$ which takes values in a finite set $S = \{1, 2, ..., s\}$ with transition probability matrix $P = [p(i,j)]_{i,j \in S}$ and initial distribution $\pi = [p(i)]_{i \in S}$. Furthermore, for $r(k) = i$, $A_i := A(i)$ and $B_i := B(i)$ are constant matrices of appropriate sizes. Denote by $x(k, x_0, \pi, u)$ the solution of (1.1) under the control u, with initial condition x_0 at time $k = 0$ and initial distribution of the Markov chain π. In the case when the initial distribution is Dirac it takes the form

$$p(i_0) = 1 \text{ and } p(i) = 0 \text{ for } i \neq i_0 \tag{1.2}$$

for some $i_0 \in S$, we will denote the solution of (1.1) by $x(k, x_0, i_0, u)$. In case of no control ($B(i) = 0$, $i \in S$) the appropriate solutions will be denoted by $x(k, x_0, \pi)$ and $x(k, x_0, i_0)$, respectively. We will consider only deterministic initial condition x_0. The control $u = (u(0), u(1), ...)$ is assumed to be such that $u(k)$ is measurable with respect to the $\sigma-$field generated by $r(0), r(1), ...,$ $r(k)$. That is, the control is causal. The assumption of causality of the control may be expressed alternatively by stating that the control $u(k)$ is of the form $f_k(r(0), r(1), ..., r(k))$. It is worth to notice that even if the control is of the form $u(k) = f_k(r(k))$ the solution $x(k, x_0, \pi, u)$ is not in general a Markov chain, however the joint process $(x(k, x_0, \pi, u), r(k))$, which takes values in $R^n \times S$, is a Markov chain for any control of the form $u(k) = f_k(r(k))$.

In this formulation the state of the system (1.1) is hybrid. It consists of two parts: x which is continuous and r which is discrete. For example, in target tracking problem x contains target location variables (position and velocity), and the discrete variable r represents presence of a manoeuvre, target classification (friend or foe), etc. In this example u will represent the tracker platform orientation command.

By I_n we denote the identity matrix of size n. For a square matrix A we denote by $\rho(A)$ the spectral radius. For a random variable X and a $\sigma-$field $\mathcal{F}_0$ we denote by EX the expectation and by $E\left(X|\,\mathcal{F}_0\right)$ the conditional expectation. When the $\sigma-$field $\mathcal{F}_0$ is generated by certain random variable Y we write $E\left(X|\,Y\right)$. Moreover for conditional expectation of the form $E\left(X|\,r(0)\right)$, where $r(0)$ has distribution π, we introduce symbol $E_\pi X$ and in special case when π is a Dirac distribution of the form (1.2) we write $E_{i_0}X$. We use similar convention for conditional probability denoted by $P\left(\cdot|\,A\right)$, when $A = \{r(0) = i_0\}$ we write $P_{i_0}(\cdot)$.

We introduce also the following notation which is used in formulation of controllability results in chapter 2.

$$F(k,k) = I_{n\times n}$$

$$F(k,l,i_{k-1},...,i_l) = A\left(i_{k-1}\right)A\left(i_{k-2}\right)...A\left(i_l\right),$$

for all $k > l \geq 0$, $i_{k-1},...,i_l \in S$

$$F(k,l) = A\left(r\left(k-1\right)\right)A\left(r\left(k-2\right)\right)...A\left(r\left(l\right)\right),\ k > l \geq 0,$$

$$\overline{F}_\pi(k,l) = E_\pi\left(F(k,l)\,|r\left(l-1\right)\right),k > l \geq 1$$

$$\overline{F}_\pi(k,0) = E_\pi F(k,0)$$

$$W_\pi\left(k\right) = E_\pi\sum_{t=0}^{k-1}\overline{F}_\pi(k,t+1)B\left(r(t)\right)B'\left(r(t)\right)\overline{F}'_\pi(k,t+1) \tag{1.3}$$

Using this notation we can write the solution of (1.1) in the following form

$$x\left(k,x_0,\pi,u\right) = F(k,0)x_0 + \sum_{t=0}^{k-1}F\left(k,t+1\right)B\left(r\left(t\right)\right)u\left(t\right), \tag{1.4}$$

or

$$x\left(k,x_0,\pi,u\right) = F(k,0,r(k-1),...,r(0))x_0 + \tag{1.5}$$

$$\sum_{t=0}^{k-1}F\left(k,t+1,r(k-1),...,r(t+1)\right)B\left(r\left(t\right)\right)u\left(t\right),k\geq 1.$$

We also use the following notation

$$\overline{S}_\pi^{(N)} = \{i_0,...,i_{N-1}\in S : p(i_0)p\left(i_0,i_1\right)...p\left(i_{N-2},i_{N-1}\right) > 0\} \tag{1.6}$$

$\overline{\overline{s}}_\pi^{(N)}$ is the number of elements of $\overline{S}_\pi^{(N)}$, and in the case when the initial distribution π is of the form $p(i_0) = 1$, and $p(i) = 0$ for $i \neq i_0$ we will write $\overline{S}_{i_0}^{(N)}$, and $\overline{\overline{s}}_{i_0}^{(N)}$ instead of $\overline{S}_\pi^{(N)}$, and $\overline{\overline{s}}_\pi^{(N)}$, respectively. It will be convenient to have the elements of $\overline{S}_{i_0}^{(N)}$ ordered in a sequence. In that purpose let order the elements of the set

$$S^N = \{(i_0,...,i_{N-1}) : i_0,...,i_{N-1}\in S\}$$

as follows

$$\begin{aligned}&\left(i_0,1,1,...,1,1\right),\left(i_0,1,1,...,1,2\right),...,\left(i_0,1,1,...,1,s\right),...\\&\left(i_0,1,1,...,2,1\right),\left(i_0,1,1,...,2,2\right),...,\left(i_0,1,1,...,2,s\right),...\\&\left(i_0,s,s,...,s,1\right),\left(i_0,s,s,...,s,2\right),...,\left(i_0,s,s,...,s,s\right).\end{aligned}$$

Withdraw all the elements $(i_0,...,i_{N-1})$ such that

$$p\left(i_0,i_1\right)...p\left(i_{N-2},i_{N-1}\right) = 0.$$

The sequence of all elements of $\overline{S}_{i_0}^{(N)}$ obtained in this way is called the natural order in the set $\overline{S}_{i_0}^{(N)}$ and it is denoted by $\widetilde{S}_{i_0}^{(N)}$. Fix a number $N > 0$ and a sequence $\left(i_0,i_1,...,i_{N-1}\right)$ of elements of S. Consider a matrix column blocks which are numbered successively by sequences: $i_0, \widetilde{S}_{i_0}^{(2)},...,\widetilde{S}_{i_0}^{(N)}$ and the block $\left(i_0,i_1,...,i_k\right)$, $k = 0,1,..,N-1$ is given by

$$F(N,k,i_{N-1},...,i_k)B_{i_k}$$

and the others are equal to 0. Denote the matrix obtained in this way by $C\left(i_0,i_1,...,i_{N-1}\right)$ and by $G(i_0)$ -the matrix consisting of all $C(i_0,i_1,...,i_{N-1})$ (as row blocks numbered by $\widetilde{S}_{i_0}^{(N)}$) for $\left(i_0,i_1,...,i_{N-1}\right)\in\overline{S}_{i_0}^{(N)}$. Moreover by $H\left(i_0\right)\in R^{n\overline{\overline{s}}_{i_0}^{(N)}\times m}$ let denote a matrix row blocks of which are numbered by the sequence $\widetilde{S}_\pi^{(N)}$, the block $\left(i_0,i_1,...,i_{N-1}\right)$, is given by $F(N,0,i_{N-1},...,i_0)$. For example in the case when $S = \{1,2\}$, $N = 3$ and $p\left(i,j\right) > 0$ for all $i,j\in S$ and $p(1) = 1$, we have

$$G(1) = \begin{bmatrix} C(1,1,1)\\ C(1,1,2)\\ C(1,2,1)\\ C(1,2,2)\end{bmatrix} =$$

$$\begin{array}{ccccccc} (1) & (1,1) & (1,2) & (1,1,1) & (1,1,2) & (1,2,1) & (1,2,2)\\ \lbrack A_1^2B_1 & A_1B_1 & 0 & B_1 & 0 & 0 & 0\rbrack\\ \lbrack A_2A_1B_1 & A_2B_1 & 0 & 0 & B_2 & 0 & 0\rbrack\\ \lbrack A_1A_2B_1 & 0 & A_1B_2 & 0 & 0 & B_1 & 0\rbrack\\ \lbrack A_2^2B_1 & 0 & A_2B_2 & 0 & 0 & 0 & B_2\rbrack \end{array}$$

and

$$H(1) = \begin{bmatrix} A_1^3 \\ A_2 A_1^2 \\ A_1 A_2 A_1 \\ A_2^2 A_1 \end{bmatrix}.$$

Moreover let $f_1^{(k)}, ..., f_n^{(k)} \in R^{nk}$ denote the vectors defined by

$$f_l = \left.\begin{bmatrix} e_l \\ e_l \\ \vdots \\ e_l \end{bmatrix}\right\} k\text{-times } e_l, \ l = 1, ..., n$$

where $e_1, ..., e_n$ is the standard base in R^n. Let us denote

$$\mathcal{A}_{i_0}^{(\delta)} = \left\{ X \subset \overline{S}_{i_0}^{(N)} : P_{i_0}\left((i_0, r(1), ..., r(N-1)) \in X\right) \geq \delta \right\}$$

and by $\overline{\overline{x}}$ denote the number of elements of $X \in \mathcal{A}_{i_0}^{(\delta)}$. For each $X \in \mathcal{A}_{i_0}^{(\delta)}$ let $G_X(i_0)$ be the submatrix of $G(i_0)$ that consists of the blocks $C(\alpha)$ for $\alpha \in X$ and for $\beta \in \overline{S}_{i_0}^{(N)}$

$$A_\beta = \{\omega \in \Omega : (r(0), r(1), ..., r(N-1)) = \beta\}. \tag{1.7}$$

and

$$\delta_0(i_0) = \min\left\{ P_{i_0}(A_\beta) : \beta \in \overline{S}_{i_0}^{(N)} \right\}. \tag{1.8}$$

With each $\alpha \in \overline{S}_\pi^{(N)}$, $\alpha = (i_0, ..., i_{N-1})$ we associate a deterministic time-varying system

$$x(k+1) = \overline{A}(k)x(k) + \overline{B}(k)u(k), \ N-1 \geq k \geq 0, \tag{1.9}$$

where

$$\left(\overline{A}(k), \overline{B}(k)\right) = (A(i_k), B(i_k)), \ N-1 \geq k \geq 0,$$

and we call it a deterministic system which corresponds to α.

Chapter 2

Controllability

Since the early work on state-space approaches to control systems analysis, it was recognized that certain nondegeneracy assumption were useful, in particular in the context of optimal control. However, it was until Kalman's work [63] that the property of controllability was isolated as of interest in and of itself, as it characterizes the degrees of freedom available when attempting to control a system.

The study of controllability for linear systems has spanned a great number of research directions, and topics such as testing degrees of controllability, and their numerical analysis aspects, are still a subject of intensive research.

The idea of controllability of jump linear systems has been already discussed in the following papers [27], [28], [29], [58], [66], in discrete time case and in [34], [82], [60], in continuous time case. The origin of idea of controllability discussed in the papers [66], [82], [60], namely the idea of $\varepsilon-$controllability with probability δ comes from papers [40], [65], [102] where general stochastic systems have been considered. In [58] the authors discussed the original idea of reaching given target in random time, and under different assumptions on the time they obtain different types of controllability. In fact they appear to be equivalent (see section 2.4). The problem of reaching target in the fixed time is disused in [27], [28], where in the first paper the target is a given value of expectation of the state and in the second one a vector. In this chapter we propose certain new ideas of controllability and we make a comparison with the existing ones.

This chapter is organized as follows. In Section 2.1 we study the weakest of the concepts of controllability for jump linear system namely the problem of controllability of the expectation of the final state (controllability with respect to the expectation). The main result of this section, Theorem 1 gives a necessary and sufficient condition for this type of controllability. Most of the results are from [27]. In the next section our attention is focussed on the possibility of reaching any deterministic target value from given deterministic initial condition in given time with prescribed probability (stochastic controllability with probability δ). We have also investigated several variations of this problem such as: the case when the target or initial condition are zero (so called controllability to zero or from zero), and the case when we want to achieve only certain neighborhood of the target. Section 2.3 is devoted to the special case of stochastic controllability with probability δ namely to the case $\delta = 1$. Most of the analysis tools presented here were taken from [28]. Whereas in Sections 2.1, 2.2 and 2.3 the control horizon is fixed, in Section 2.4 we consider the concepts of controllability when the time of achieving the target value can be a random variable. The results of this section are partially published in [29]. Finally in Section 2.5 the relationships between the introduced types of controllability are explained as well as the comparison with existing results is made.

2.1 Controllability with respect to expectation

In this chapter we propose a definition of controllability indicating the possibility of reachability of any given value of the expectation of the final state in given time.

We start from the following theorem:

Theorem 1 *For a fixed initial distribution π of the Markov chain the following conditions are equivalent*

1. *Matrix $W_\pi(N)$ is invertible (definition of $W_\pi(N)$ is given by (1.3)).*
2. *For all $x_1 \in R^n$ there exists a control u such that*
$$E_\pi x(N, 0, \pi, u) = x_1. \tag{2.1}$$
3. *For all $x_0,\ x_1 \in R^n$ there exists a control u such that*
$$E_\pi x(N, x_0, \pi, u) = x_1. \tag{2.2}$$

Proof. $(1 \Rightarrow 2)$ Assume that the matrix $W_\pi(N)$ is invertible. Fix the target vector $x_1 \in R^n$ and consider the control in the following form

$$u(k) = B'(r(k))\,\overline{F}'_\pi(N, k+1) W_\pi^{-1}(N)\, x_1. \tag{2.3}$$

This control is causal and using (1.4) the solution of (1.1) can be expressed as

$$\begin{aligned}
&E_\pi x(N, 0, u) = \\
&E_\pi \sum_{t=0}^{N-1} F(N, t+1)\, B(r(t))\, B'(r(t))\, \overline{F}'(N, t+1) W_\pi^{-1}(N)\, x_1 = \\
&E_\pi \sum_{t=0}^{N-1} E\left(F(N, t+1)\, B(r(t))\, B'(r(t))\, \overline{F}'(N, t+1) W_\pi^{-1}(N)\, x_1 \middle|\, r(t)\right) = \\
&E_\pi \sum_{t=0}^{N-1} E\left(F(N, t+1) \middle|\, r(t)\right) B(r(t))\, B'(r(t))\, \overline{F}'(N, t+1) W_\pi^{-1}(N)\, x_1 = \\
&\left(E_\pi \sum_{t=0}^{N-1} \overline{F}(N, t+1) B(r(t))\, B'(r(t))\, \overline{F}'(N, t+1)\right) W_\pi^{-1}(N)\, x_1 = x_1,
\end{aligned} \tag{2.4}$$

where the following properties of conditional expectation were used

$$E(E(\xi|\,\eta)) = E\xi,\quad E(\xi f(\eta)|\,\eta) = E(\xi|\,\eta)\, f(\eta). \tag{2.5}$$

The equality (2.4) shows that the control given by (2.3) is such that (2.1) is satisfied.

$(2 \Rightarrow 3)$ Suppose that (2.1) is satisfied. Fix $x_0,\ x_1 \in R^n$ and let u be such a control that

$$E_\pi x(N, 0, \pi, u) = x_1 - E_\pi F(N, 0) x_0$$

by (1.4) it is equivalent to

$$x_1 - E_\pi F(N, 0) x_0 = E_\pi \sum_{t=0}^{N-1} F(N, t+1)\, B(r(t))\, u(t)$$

and consequently

$$\begin{aligned}
x_1 &= E_\pi\left(F(N, 0) x_0 + \sum_{t=0}^{N-1} F(N, t+1)\, B(r(t))\, u(t)\right) \\
&= E_\pi x(N, x_0, \pi, u).
\end{aligned}$$

The last equality shows that (2.2) is satisfied.

$(3 \Rightarrow 1)$ For the purpose of getting contradiction suppose that 3 is satisfied but there exists nonzero vector $g \in R^n$ such that

$$g' W(N) g = 0. \tag{2.6}$$

Using the definition of $W(N)$ we have

$$g'W(N)g = g'\left(E\sum_{t=0}^{N-1}\overline{F}(N,t+1)B\left(r(t)\right)B'\left(r(t)\right)\overline{F}'(N,t+1)\right)g =$$

$$\sum_{t=0}^{N-1}E\left(g'\overline{F}(N,t+1)B\left(r(t)\right)B'\left(r(t)\right)\overline{F}'(N,t+1)g\right) =$$

$$\sum_{t=0}^{N-1}E\left\|B'\left(r(t)\right)\overline{F}'(N,t+1)g\right\|^2.$$

The last equality together with (2.6) implies that

$$E\left\|B'\left(r(t)\right)\overline{F}'(k,t+1)g\right\|^2 = 0,\ t = 0,...,N-1$$

and consequently

$$P\left(B'\left(r\left(t\right)\right)\overline{F}'(N,t+1)g = 0\right) = 1,\ t = 0,...,N-1. \tag{2.7}$$

By (2.2) there exists a causal control $u(k)$, $k = 0,...,N-1$ such that

$$g = E\sum_{t=0}^{N-1}F\left(N,t+1\right)B\left(r\left(t\right)\right)u\left(t\right) =$$

$$E\sum_{t=0}^{N-1}\overline{F}\left(N,t+1\right)B\left(r\left(t\right)\right)u\left(t\right).$$

To obtain the last equality property (2.5) and causality of control law has been used. This together with (2.7) implies

$$g'g = g'E\sum_{t=0}^{N-1}\overline{F}\left(N,t+1\right)B\left(r\left(t\right)\right)u\left(t\right) =$$

$$E\sum_{t=0}^{N-1}g'\overline{F}\left(N,t+1\right)B\left(r\left(t\right)\right)u\left(t\right) = 0$$

and consequently $g = 0$. This contradicts the assumption that $g \neq 0$. ■

Definition 1 *If one of the conditions 1-3 of Theorem 1 is satisfied then we call the system (1.1) $\pi-$controllable with respect to the expectation at time N ($\pi-CWRE$ at time N). If the system (1.1) is $\pi-CWRE$ at time N for each initial distribution then we say that it is controllable with respect to the expectation at time N ($CWRE$ at time N).*

Theorem 1 gives the necessary and sufficient conditions for $\pi-$CWRE at time N, moreover the proof is constructive in the sense that the control which governs the expectation to the desired value is explicitly given but the disadvantage is that the condition is difficult to check. A sufficient condition easier to check but non constructive in the sense that it does not give the control is presented now.

Corollary 1 *If there exists a sequence $(i_0,...,i_{N-1}) \in \overline{S}_\pi^{(N)}$ such that*

$$rank\begin{bmatrix}B_{i_{N-1}} & A_{i_{N-1}}B_{i_{N-2}} & ... & \prod_{j=1}^{N-1}A_{i_j}B_{i_0}\end{bmatrix} = n \tag{2.8}$$

then the system is $\pi-CWRE$ at time N. ($A_i = A(i)$, $B_i = B(i)$).

Proof. We have the following property of the integral

$$E\xi \geq \int_A \xi\left(\omega\right)P\left(d\omega\right), \tag{2.9}$$

for any nonnegative random variable ξ and any measurable set A. Let consider (2.9) with $A = A_\alpha$ (see, (1.7) for definition of A_α). Then by the definition of $W_\pi(N)$ and (2.9) we have

$$W(N) = E\sum_{t=0}^{k-1}\overline{F}(N,t+1)B\left(r(t)\right)B'\left(r(t)\right)\overline{F}'(N,t+1) =$$

$$\sum_{t=0}^{k-1}\int_\Omega \overline{F}(N,t+1)B\left(r(t)\right)B'\left(r(t)\right)\overline{F}'(N,t+1)P\left(d\omega\right) \geq$$

$$\sum_{t=0}^{k-1}\int_{A_\alpha}\overline{F}(N,t+1)B\left(r(t)\right)B'\left(r(t)\right)\overline{F}'(N,t+1)P\left(d\omega\right) =$$

$$\begin{bmatrix}B_{i_{N-1}} & A_{i_{N-1}}B_{i_{N-2}} & ... & \prod_{j=1}^{N-1}A_{i_j}B_{i_0}\end{bmatrix}\times$$

$$\begin{bmatrix}B_{i_{N-1}} & A_{i_{N-1}}B_{i_{N-2}} & ... & \prod_{j=1}^{N-1}A_{i_j}B_{i_0}\end{bmatrix}'.$$

The last inequality together with the assumption (2.8) implies that $W_\pi(N)$ is invertible and the Corollary follows from point 1 of Theorem 1. ■

The next example shows that deterministic controllability of each pair (A_i, B_i) is not a necessary condition for system (1.1) to be CWRE at time N.

Example 1 *Consider system (1.1) with a two-state form structure. Let* $p(1) = p(2) = 0.5$, $p(1,1) = p(2,2) = p(2,1) = p(1,2) = 0.5$, *and*

$$A_1 = \begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix},\ A_2 = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix},\ B_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix},\ B_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

If we put $i_0 = i_1 = ... = i_{N-1} = 1$ *for* $N \geq 2$, *then we have*

$$p(i_0)\, p(i_0, i_1) p(i_1, i_2) ... p(i_{N-2}, i_{N-1}) = (0.5)^N$$

and

$$rank \left[B_{i_{N-1}} \middle| A_{i_{N-1}} B_{i_{N-2}} \ldots \prod_{j=1}^{N-1} A_{i_j} B_{i_0} \right] = 2.$$

By Corollary 1, we conclude that the system is CWRE at time N for each $N \geq 2$.

Theorem 2 *The following conditions are equivalent*

1. *Matrix* $W_\pi(N)$ *is invertible (definition of* $W_\pi(N)$ *is given by (1.3)) for each initial distribution* π.
2. *For all* $x_1 \in R^n$ *and all initial distribution* π *there exists a control sequence* $u(k)$, $k = 0, ..., N-1$ *such that*
$$E_\pi x(N, 0, \pi, u) = x_1. \tag{2.10}$$
3. *For all* $x_0,\ x_1 \in R^n$ *and all initial distributions* π *there exists a control sequence* $u(k)$, $k = 0, ..., N-1$ *such that*
$$E_\pi x(N, x_0, \pi, u) = x_1. \tag{2.11}$$
4. *For all* $x_0,\ x_1 \in R^n$ *and all* $i_0 \in S$ *there exists a control sequence* $u(k)$, $k = 0, ..., N-1$ *such that*
$$E_{r(0)=i_0} x(N, x_0, i_0, u) = x_1. \tag{2.12}$$
5. *For all* $x_1 \in R^n$ *and all* $i_0 \in S$ *there exists a control sequence* $u(k)$, $k = 0, ..., N-1$ *such that*
$$E_{r(0)=i_0} x(N, 0, i_0, u) = x_1. \tag{2.13}$$
6. *Matrix* $W_{r(0)=i_0}(N)$ *is invertible for all* $i_0 \in S$.

Proof. The equivalence of the conditions 1-3 follows from Theorem 1. The equivalence of conditions 4-6 may be established in the same way as the equivalence of conditions 1-3 in proof of Theorem 1. Moreover implication ($1 \Rightarrow 6$) is straightforward. Therefore to complete the proof it is enough to show that ($6 \Rightarrow 1$). For this purpose fix initial distribution π and choose the element i_0 of S such that $p(i_0) > 0$. By the theorem of total probability we have

$$W_\pi(N) = \sum_{i \in S} W_{r(0)=i}(N)\, p(i) \geq W_{r(0)=i_0}(N)\, p(i_0).$$

To obtain the last inequality the fact that matrices W_i are nonnegative definite for all $i \in S$ has been used. This inequality implies that $W_\pi(N)$ is positively definite because $W_{i_0}(N)$ has such a property and $p(i_0) > 0$. ■

The next example shows that deterministic controllability of each pair (A_i, B_i) is not a sufficient condition for $\pi-$CWRE at time N. We have already shown (Example 1) that deterministic controllability of each pair (A_i, B_i) is not a necessary condition for $\pi-$CWRE at time N. However, Corollary 1 guarantees that controllability of at least one deterministic system corresponding to $\alpha \in \overline{S}_\pi^{(N)}$ is a sufficient condition for $\pi-$CWRE at time N but it is not a necessary condition. It is possible (see, Example 3) that the deterministic system that corresponds to each $\alpha \in \overline{S}_\pi^{(N)}$ is not controllable in the deterministic sense and the system is $\pi-$CWRE at time N.

Example 2 *Consider system (1.1) with a two-state form structure. Let* $p(1) = 1$, $p(2) = 0$, $p(1,1) = p(2,2) = 0$, $p(2,1) = p(1,2) = 1$, *and*

$$A_1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},\ A_2 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix},\ B_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix},\ B_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

To test $\pi-$*CWRE at time N observe that* $\overline{F}(N,l)$ *is in this case a constant random variable and*

$$E\overline{F}(N,l)B(r(l))B'(r(l)) =$$

$$\overline{F}(N,l)E(B(r(l))B'(r(l))) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Consequently $W_\pi(N) = 0$ *and the system is not* $\pi-$*CWRE at time N for each N though each pair* (A_i, B_i), $i = 1, 2$ *is controllable, and consequently it is not CWRE at time N for each N.*

Example 3 *Consider system (1.1) with a two-state form structure. Let* $p(1)=1, p(2)=0, p(1,1)=p(2,2)=p(2,1)=p(1,2)=0.5$, *and*

$$A_1=\begin{bmatrix}1&0\\0&1\end{bmatrix},\ A_2=\begin{bmatrix}0&0\\0&0\end{bmatrix},\ B_1=\begin{bmatrix}0\\1\end{bmatrix},\ B_2=\begin{bmatrix}1\\1\end{bmatrix}.$$

Then $\overline{S}_\pi^{(2)}=\{(1,1),(1,2)\}$ *and neither the deterministic system that corresponds to the element* $(1,1)$ *nor the one which corresponds to the element* $(1,2)$ *is controllable in the deterministic sense. However,*

$$W_\pi(2)=\begin{bmatrix}1&1\\1&3\end{bmatrix}>0$$

and this system is $\pi-CWRE$ *at time 2.*

2.2 Stochastic Controllability

In this paragraph we examine a concept of controllability idea of which is to steer any deterministic initial condition to a given deterministic target value at given time with prescribed probability. We have the following definition:

Definition 2 *We say that system (1.1) is stochastically controllable with probability* δ *at time* N *(SCWP* δ *at time* N *) if, for all* $x_0,x_1\in R^n$ *there exists a control* u *such that*

$$P_\pi\left(x\left(N,x_0,\pi,u\right)=x_1\right)\geq\delta. \tag{2.14}$$

Analogically, we say that system (1.1) is SCWP δ *at time* N *to zero (from zero) if, for all* $x_0\in R^n$ $(x_1\in R^n)$ *there exists a control* u *such that*

$$P_\pi\left(x\left(N,x_0,\pi,u\right)=0\right)\geq\delta \qquad \left(P_\pi\left(x\left(N,0,\pi,u\right)=0\right)\geq\delta\right) \tag{2.15}$$

In the case when $\delta=1$ *we say that system (1.1) is directly controllable (DC) at time* N *(DC at time* N *to zero, DC at time* N *from zero).*

The next theorem reduces problems of SCWP δ at time N and DC at time N for system (1.1) with initial distribution of the Markov chain π, to problems of SCWP δ at time N and DC at time N for system (1.1) with Dirac initial distribution.

Theorem 3 *Suppose that for each* $i\in S$ *system (1.1) with* $P(r(0)=i)=1$ *is SCWP* $\delta(i)$ *at time* N *(SCWP* $\delta(i)$ *at time* N *to zero, SCWP* $\delta(i)$ *at time* N *from zero) then for initial distribution* π *of the form* $P(r(0)=i)=p(i)$, $i\in S$ *it is SCWP* δ *at time* N *(SCWP* δ *at time* N *to zero, SCWP* δ *at time* N *from zero), where*

$$\delta=\sum_{i\in S}p(i)\delta(i).$$

Moreover system (1.1) is DC at time N *(DC at time* N *to zero, DC time* N *from zero) for all initial distributions* π *if and only if it is DC at time* N *(DC at time* N *to zero, DC time* N *from zero) for all Dirac initial distributions.*

Proof. The proof is a straightforward consequence of the theorem of total probability and we will show only the part about SCWP δ at time N. Fix $x_0,x_1\in R^n$. For each $i\in\overline{S}_\pi^{(1)}$ let

$$u^{(i)}=\left(u^{(i)}(0),u^{(i)}(1),...u^{(i)}(N-1)\right)$$

be such a control that $u^{(i)}(k)$ is of the form $f_k\left(i,r\left(1\right),...,r\left(k\right)\right)$ and

$$P\left(x\left(N,x_0,i,u^{(i)}\right)=x_1\middle|r(0)=i\right)\geq\delta(i). \tag{2.16}$$

Let define a new control $u=(u(0),u(1),...u(N-1))$ by

$$u(k)=f_k\left(r(0),r\left(1\right),...,r\left(k\right)\right).$$

Then from the theorem of total probability and (2.16) we have

$$P\left(\left\|x\left(N,x_0,\pi,u\right)-x_1\right\|>0\right)=$$

$$\sum_{i\in\overline{S}_\pi^{(1)}}P\left(\left\|x\left(N,x_0,\pi,u^{(i)}\right)-x_1\right\|>0\middle|r(0)=i\right)p(i)\geq\sum_{i\in S}p(i)\delta\left(i\right).$$

The last inequality completes the proof. Following the same line of reasoning the remaining part of the Theorem may be proved. ■

Having in mind the previous result we may restrict our considerations to system (1.1) with initial distribution of the Markov chain being a Dirac one, without losing generality and in the remainder of the section it is assumed that the distribution is of the form: $P\left(r(0)=i_0\right)=1$.

The next theorem contains necessary and sufficient conditions for SCWP δ at time N as well as SCWP δ at time N from zero and to zero.

Theorem 4 *System (1.1) is SCWP* δ *at time* N *from zero if and only if there exists* $X\in\mathcal{A}_{i_0}^{(\delta)}$ *such that*

$$rankG_X(i_0)=rank\left[G_X(i_0)\quad f_l^{(\overline{\pi})}\right],\ \text{for all } l=1,...,n. \tag{2.17}$$

System (1.1) is SCWP δ at time N to zero if and only if there exists $X \in \mathcal{A}_{i_0}^{(\delta)}$ such that

$$\operatorname{Im} H_X(i_0) \subset \operatorname{Im} G_X(i_0), \tag{2.18}$$

and it is SCWP δ at time N if and only if there exists $X \in \mathcal{A}_{i_0}^{(\delta)}$ such that

$$rank G_X(i_0) = rank \begin{bmatrix} G_X(i_0) & f_l^{(\overline{\overline{x}})} \end{bmatrix}, \text{ for all } l = 1, ..., n, \tag{2.19}$$

and

$$\operatorname{Im} H_X(i_0) \subset \operatorname{Im} G_X(i_0). \tag{2.20}$$

Before entering the formal proof let us briefly discuss the main idea. Since the set $\overline{S}_{i_0}^{(N)}$ - of all possible sample paths of the Markov chain at time N is finite, therefore the family $\mathcal{A}_{i_0}^{(\delta)}$ consisting of all subsets of sample paths with probability of being taken greater or equal δ is finite too. Therefore the question about SCWP δ at time N can be reformulated, similarly as for deterministic system, as a question about existence of a solution of a finite set of linear equations. Nevertheless now we must take into account the constrain that control $u(k)$ at time k may depend only on the variables $r(0), ...r(k)$ and should be independent of $r(k+1),..., r(N)$. In the proof we obtain this by the proper definition of matrices $G_X(i_0)$ and $H_X(i_0)$.

Proof. Suppose that system (1.1) is SCWP δ at time N from zero. Then for each $y \in R^n$ there exists a control sequence $u(0), ..., u(N-1)$ such that

$$u(k) = g_k(i_0, r(1), ..., r(k)), k = 0, ..., N-1,$$

and

$$P_{i_0}(x(N, 0, i_0, u) = y) \geq \delta, \tag{2.21}$$

where g_k is a function from $\overline{S}_{i_0}^{(k)}$ to $R^m, k = 0, ..., N-1$. Let us define

$$\mathcal{B}(y) = \{\omega \in \Omega : x(N, 0, i_0, u) = y\}$$

and

$$X(y) = \left\{(i_0, r(1), ..., r(N-1)) \in \overline{S}_{i_0}^{(N)} : \omega \in \mathcal{B}(y)\right\}$$

then $X(y) \in \mathcal{A}_{i_0}^{(\delta)}$ and for $(i_0, ..., i_{N-1}) \in X(y)$ the following holds

$$\sum_{t=0}^{N-1} F(N, t+1, i_{N-1}, ..., i_{t+1}) B(i_t) g_t(i_0, ..., i_t) = y.$$

It means that the system of equations

$$G_{X(y)}(i_0)v = z, \tag{2.22}$$

where

$$z = \left.\begin{bmatrix} t \\ \vdots \\ t \end{bmatrix}\right\} \overline{\overline{x}}(y) \text{ times}$$

has a solution for $t = y$. Denote by $V(X(y))$ the subspace of all $t \in R^n$ such that the system of equations (2.22) has a solution. From SCWP δ at time N from zero we obtain that

$$\bigcup_{y \in R^n} V(X(y)) = R^n. \tag{2.23}$$

Since the set $\mathcal{A}_{i_0}^{(\delta)}$ is finite, then consequently

$$\{X(y) : y \in R^n\}$$

and

$$\{V(X(y)) : y \in R^n\}$$

are finite as well. Then (2.23) implies that R^n is a union of finite number of linear subspaces, but it is possible if and only if one of them is the whole R^n. Therefore there exists $y_0 \in R^n$ such that $V(X(y_0)) = R^n$. This clearly forces that the system of equations

$$G_{X(y_0)}(i_0)v = z,$$

where

$$z = \left.\begin{bmatrix} t \\ \vdots \\ t \end{bmatrix}\right\} \overline{\overline{x}} \text{ times}$$

has a solution for each $t \in R^n$. Since vectors $f_l^{(\overline{\overline{x}}(y_0))}$, $l = 1, ..., n$ form a base of the space

$$\left\{\begin{bmatrix} t \\ \vdots \\ t \end{bmatrix} \in R^{n\overline{\overline{x}}(y_0)} : t \in R^n\right\},$$

Kronecker Capeli Theorem (see e.g.[14]) implies that (2.17) holds with $X = X(y_0)$.

Suppose now that (2.17) holds for some $X \in \mathcal{A}_{i_0}^{(\delta)}$. Again by Kronecker Capeli Theorem the set of equations

$$G_X(i_0)v = z$$

where

$$z = \left[\begin{array}{c} y \\ \vdots \\ y \end{array}\right] \Bigg\} \, \overline{\overline{x}} \text{ times}$$

has a solution for each $y \in R^n$. This fact implies that for each $y \in R^n$ and each $(i_0, ..., i_{N-1}) \in X$ there exists a sequence $g_k(i_0, ..., i_k)$, $k = 0, ..., N-1$ such that

$$\sum_{t=0}^{N-1} F(N, t+1, i_{N-1}, ..., i_{t+1}) B(i_t) g_t(i_0, ..., i_t) = y.$$

If we define the control

$$u(k) = \begin{cases} g_k(i_0, r(1)..., r(k)) & \text{if } (i_0, r(1)..., r(k)) \in X_k \\ 0 & \text{if } (i_0, r(1)..., r(k)) \notin X_k \end{cases}$$

where

$$X_k = \{(i_0, ..., i_k) : (i_0, ..., i_k, i_{k+1}, ..., i_{N-1}) \in X \text{ for some } i_{k+1}, ..., i_{N-1} \in S\}$$

then

$$P_{i_0}(x(N, 0, i_0, u) = y) \geq \delta$$

since $X \in \mathcal{A}_{i_0}^{(\delta)}$. Consequently system (1.1) is SCWP δ at time N from zero.

Suppose that the system (1.1) is SCWP δ at time N to zero. Then for each $y \in R^n$ there exists a control sequence $u(0), ..., u(N-1)$ such that

$$u(k) = g_k(i_0, r(1), ..., r(k)), k = 0, ..., N-1,$$

and

$$P_{i_0}(x(N, y, i_0, u) = 0) \geq \delta, \tag{2.24}$$

where g_k is a function from $\overline{S}_{i_0}^{(k)}$ to $R^m, k = 0, ..., N-1$. Let us define

$$\mathcal{B}(y) := \{\omega \in \Omega : x(N, y, i_0, u) = 0\}$$

and

$$X(y) = \{(i_0, r(1), ..., r(N-1)) : \omega \in \mathcal{B}(y)\}$$

then $X(y) \in \mathcal{A}_{i_0}^{(\delta)}$ and according to (1.5) for $(i_0, ..., i_{N-1}) \in X$ the following holds

$$\sum_{t=0}^{N-1} F(N, t+1, i_{N-1}, ..., i_{t+1}) B(i_t) g_t(i_0, ..., i_t) = -F(N, 0, i_{N-1}, ..., i_0) y.$$

The last implies that

$$-H_{X(y)}(i_0) y \in \operatorname{Im} G_{X(y)}(i_0).$$

Let us denote by $V(X(y))$the linear subspace of all $t \in R^n$ such that

$$H_{X(y)}(i_0) t \in \operatorname{Im} G_{X(y)}(i_0).$$

From the SCWP δ at time N to zero we know that

$$\bigcup_{y \in R^n} V(X(y)) = R^n.$$

Following the line of reasoning after (2.23) we conclude that there exists $y_0 \in R^n$ such that $V(X(y_0)) = R^n$ and (2.18) holds with $X = X(y_0)$. Assume now that the condition (2.18) holds for some $X \in \mathcal{A}_{i_0}^{(\delta)}$. This means that for each $x_0 \in R^n$ there exists v such that

$$-H_X(i_0) x_0 = G_X(i_0) v$$

This in turn implies that for each $(i_0, ..., i_{N-1}) \in X$ there exists a sequence $g_k(i_0, ..., i_k)$, $k = 0, ..., N-1$ such that

$$\sum_{t=0}^{N-1} F(N, t+1, i_{N-1}, ..., i_{t+1}) B(i_t) g_t(i_0, ..., i_t) = -F(N, 0, i_{N-1}, ..., i_0) x_0.$$

If we define the control

$$u(k) = \begin{cases} g_k(i_0, r(1)..., r(k)) & \text{if } (i_0, r(1)..., r(k)) \in X_k \\ 0 & \text{if } (i_0, r(1)..., r(k)) \notin X_k \end{cases}$$

then

$$P_{i_0}(x(N, x_0, i_0, u) = 0) \geq \delta$$

since $X \in \mathcal{A}_{i_0}^{(\delta)}$. Consequently system (1.1) is SCWP δ at time N to zero. In the same way we can show the part about SCWP δ at time N. ■

Remark 1 *Suppose that the system is SCWP δ at time N and let $X \in \mathcal{A}_{i_0}^{(\delta)}$ be such that (2.19) and (2.20) hold. It easy to conclude that for each $\alpha \in X$ the deterministic system that corresponds to α is controllable. In fact conditions (2.19) and (2.20) are much stronger. They imply that if we fix $x_0, x_1 \in R^n$, and $\alpha, \beta \in X$ of the following form $\alpha = (i_0, i_1, ..., i_{N-1})$, $\beta = (j_0, j_1, ..., j_{N-1})$, $i_l = j_l$ for $l = 0, ..., k$ then it is possible to construct the controls $u_\alpha = (u_\alpha(0), ..., u_\alpha(N-1))$, $u_\beta = (u_\beta(0), ..., u_\beta(N-1))$ which steer x_0 and x_1 in deterministic systems corresponding to α and β, respectively and the controls are such that $u_\alpha(l) = u_\beta(l)$ for $l = 0, ..., k$.*

In the previous papers devoted to $\varepsilon-$controllability with probability δ the definition is formulated as follows (see [40], [65], [66], [81], [102]):

For all $x_0, x_1 \in R^n$ there exists a control u such that

$$P(\|x(N, x_0, i_0, u) - x_1\| < \varepsilon) \geq \delta.$$

The first impression is that this definition is less restrictive than SCWP δ at time N. The following theorem establishes the equivalence between this definition for $x_0 = 0$ and SCWP δ at time N from zero.

Theorem 5 *Fix $\varepsilon > 0$. If for all $x_1 \in R^n$ there exists a control u such that $P(\|x(N, 0, i_0, u) - x_1\| < \varepsilon) \geq \delta$ then system (1.1) is SCWP δ at time N from zero.*

Proof. We will prove that if for all $x_1 \in R^n$ there exists a control u such that

$$P(\|x(N, 0, i_0, u) - x_1\| < \varepsilon) \geq \delta, \tag{2.25}$$

then there exists a control u such that

$$P(x(N, 0, i_0, u) = x_1) \geq \delta, \tag{2.26}$$

The proof of the rest is similar. Suppose that (2.26) does not hold. By Theorem 4 it means that for each $X \in \mathcal{A}_{i_0}^{(\delta)}$ there exists $y_X \in R^n$ such that

$$t_X \notin \operatorname{Im} G_X(i_0),$$

where

$$t_X = \left.\begin{bmatrix} y_X \\ \vdots \\ y_X \end{bmatrix}\right\} \overline{\overline{x}} \text{ times}.$$

Since the family $\mathcal{A}_{i_0}^{(\delta)}$ is finite and $\operatorname{Im} G_X(i_0)$ is a closed linear subspace of $R^{\overline{\overline{x}}}$ (which cannot be dense because $R^{\overline{\overline{x}}}$ is finite dimensional) then there exists $y \in R^n$ such that

$$dist\left[z_X, \operatorname{Im} G_X(i_0)\right] > \varepsilon k, \tag{2.27}$$

where

$$z_X = \left.\begin{bmatrix} y \\ \vdots \\ y \end{bmatrix}\right\} \overline{\overline{x}} \text{ times}$$

and $k = \max\left\{\overline{\overline{x}} : X \in \mathcal{A}_{i_0}^{(\delta)}\right\}$. By the assumption (2.25) we know that there exists a control u such that

$$P_{i_0}(\mathcal{B}) \geq \delta, \tag{2.28}$$

where

$$\mathcal{B} = \{\omega \in \Omega : \|x(N, 0, i_0, u) - y\| < \varepsilon\}.$$

>From the definition of $\mathcal{A}_{i_0}^{(\delta)}$ and (2.28) we see that

$$X := \{(r(0), r(1), ..., r(N-1)) : \omega \in B\} \in \mathcal{A}_{i_0}^{(\delta)}.$$

Now let enumerate the elements of X according to the natural order in $\overline{S}_{i_0}^{(N)}$, so

$$X = \left\{\alpha_{k_1}, ..., \alpha_{k_{\overline{\overline{x}}}}\right\},$$

and let $x_{\alpha_{k_i}}$ be the value of $x(N, 0, i_0, u)$ on $A_{\alpha_{k_i}}$ ($x(N, 0, i_0, u)$ is constant on A_{α_i}). Then

$$v \in \operatorname{Im} G_X(i_0),$$

where

$$v = \begin{bmatrix} x_{\alpha_{k_1}} \\ \vdots \\ x_{\alpha_{k_{\overline{\overline{x}}}}} \end{bmatrix}$$

and

$$\|v - z_X\| = \sqrt{\sum_{i=1}^{\overline{\overline{x}}} \left\|x_{\alpha_{k_i}} - y\right\|^2} \leq \varepsilon \overline{\overline{x}}.$$

However, this contradicts (2.27). ■

Remark 2 *In a very similar way we can show that if for all $x_0, x_1 \in R^n$ ($x_0 \in R^n$) there exists a control u such that*

$$P(\|x(N, x_0, i_0, u) - x_1\| < \varepsilon) \geq \delta$$

($P(\|x(N, x_0, i_0, u)\| < \varepsilon) \geq \delta$,) then system (1.1) is SCWP δ at time N, (SCWP δ at time N to zero).

2.3 Direct controllability

The case $\delta = 1$ of SCWP δ at time N deserves special attention. Therefore we have introduced the notion DC for this type of controllability and now we will focus our attention on it.

Since $\overline{S}_{i_0}^{(N)}$ is a finite set, there exists a number $\delta_0 > 0$ such that

$$\overline{S}_{i_0}^{(N)} = \mathcal{A}_{i_0}^{(\delta)} \text{ for all } \delta \geq \delta_0. \tag{2.29}$$

Having that in mind we obtain necessary and sufficient conditions for DC at time N from Theorem 4. They are given in the following Corollary.

Corollary 2 *System (1.1) is DC to zero at time N if and only if*

$$\operatorname{Im} H_{\overline{S}_{i_0}^{(N)}}(i_0) \subset \operatorname{Im} G_{\overline{S}_{i_0}^{(N)}}(i_0) \tag{2.30}$$

System (1.1) is DC from zero at time N if and only if

$$rank G_{\overline{S}_{i_0}^{(N)}}(i_0) = rank \begin{bmatrix} G_{\overline{S}_{i_0}^{(N)}}(i_0) & f_l^{\left(\overline{s}_{i_0}^{(N)}\right)} \end{bmatrix}, \tag{2.31}$$

for all l=1,...,n. System (1.1) is DC at time N if and only if

$$rank G_{\overline{S}_{i_0}^{(N)}}(i_0) = rank \begin{bmatrix} G_{\overline{S}_{i_0}^{(N)}}(i_0) & f_l^{\left(\overline{s}_{i_0}^{(N)}\right)} \end{bmatrix}, \tag{2.32}$$

and

$$\operatorname{Im} H_{\overline{S}_{i_0}^{(N)}}(i_0) \subset \operatorname{Im} G_{\overline{S}_{i_0}^{(N)}}(i_0) \tag{2.33}$$

for all l=1,...,n.

From this theorem it is clear that (1.1) is DC at time N if and only if it is simultaneously DC to zero and from zero.

When we consider the system without jumps it is well known that controllability from zero implies the controllability to zero and that inverse implication is not true. The next example shows that for the system with jumps the $\pi-$direct controllability from zero does not imply the $\pi-$direct controllability to zero.

Example 4 *Consider the system (1.1) with $S = \{1,2\}$, $N = 2$,*

$$A_1 = \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix}, A_2 = \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}, B_1 = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, B_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$p(i) > 0,\ p(i,j) > 0,\ i,\ j \in S.$$

According to the notation we have (for simplicity we omit the index $\overline{S}_{i_0}^{(2)}$ in $G(i)$ and $H(i)$)

$$G(1) = \begin{bmatrix} C(1,1) \\ C(1,2) \end{bmatrix} = \begin{bmatrix} A_1B_1 & B_1 & 0 \\ A_2B_1 & 0 & B_2 \end{bmatrix} = \begin{bmatrix} 4 & 0 & 0 \\ 2 & 2 & 0 \\ 4 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$

and

$$G(2) = \begin{bmatrix} C(2,1) \\ C(2,2) \end{bmatrix} = \begin{bmatrix} A_1B_2 & B_1 & 0 \\ A_2B_2 & 0 & B_2 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 2 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$

and it is easy to check that condition (2.31) is satisfied. The control which steers zero initial condition to $\begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \end{bmatrix}$ at time $N = 2$ is given by

$$u(0) = \begin{cases} \frac{x_1^{(0)}}{4} & if \quad r(0) = 1 \\ \frac{x_1^{(0)}}{2} & if \quad r(0) = 2 \end{cases}$$

$$u(1) = \begin{cases} \frac{x_2^{(0)}}{2} - \frac{x_1^{(0)}}{4} & if \quad r(0) = 1, r(1) = 1 \\ x_2^{(0)} + \frac{x_1^{(0)}}{4} & if \quad r(0) = 1, r(1) = 2 \\ x_2^{(0)} + \frac{x_1^{(0)}}{2} & if \quad r(0) = 2, r(1) = 2 \\ \frac{x_2^{(0)}}{2} - \frac{x_1^{(0)}}{4} & if \quad r(0) = 2, r(1) = 1 \end{cases}.$$

From the other hand the system is not $\pi-DC$ to zero at time 2. In fact we have

$$H(1) = \begin{bmatrix} 7 & 4 \\ 6 & 7 \\ 5 & 0 \\ -2 & 1 \end{bmatrix}, H(2) = \begin{bmatrix} 3 & -4 \\ -2 & 3 \\ 1 & 0 \\ -2 & 5 \end{bmatrix}$$

and

$$\begin{bmatrix} 11 \\ 13 \\ 5 \\ 1 \end{bmatrix} \in H(1), \begin{bmatrix} -1 \\ 1 \\ 1 \\ 3 \end{bmatrix} \in H(2)$$

but

$$\begin{bmatrix} 11 \\ 13 \\ 5 \\ 1 \end{bmatrix} \notin G(1), \begin{bmatrix} -1 \\ 1 \\ 1 \\ 3 \end{bmatrix} \notin G(2).$$

Example 5 *Consider the system (1.1) with $S = \{1,2\}$, $N = 3$,*

$$A_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, A_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, B_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, B_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

$$p\ [p(i,j)]_{i,j=1,2} = \begin{bmatrix} 0.3 & 0.7 \\ 0 & 1 \end{bmatrix}.$$

If the initial distribution is of the form $\pi : p(2) = 1$, *then the problem is trivial because* $\overline{S}_2^{(3)} = \{(2,2,2)\}$. *Consider the case when* $p(1) = 1$. *We have*

$$\overline{S}_1^{(3)} = \{(1,1,1),(1,1,2),(1,2,2)\}$$

and

$$P_1\left(A_{(1,1,1)}\right) = 0.09,\ P_1\left(A_{(1,1,2)}\right) = 0.21,\ P_1\left(A_{(1,2,2)}\right) = 0.7.$$

Moreover for $0.09 \geq \delta > 0$

$$\mathcal{A}_1^{(\delta)} = \Big\{\overline{S}_1^{(3)}, \{(1,1,2),(1,2,2)\}, \{(1,1,1),(1,2,2)\}, \{(1,2,2)\},$$
$$\{(1,1,1),(1,1,2)\}, \{(1,1,2)\}, \{(1,1,1)\}\Big\},$$

for $0.21 \geq \delta > 0.09$

$$\mathcal{A}_1^{(\delta)} = \Big\{\overline{S}_1^{(3)}, \{(1,1,2),(1,2,2)\}, \{(1,1,1),(1,2,2)\}, \{(1,2,2)\},$$
$$\{(1,1,1),(1,1,2)\}, \{(1,1,2)\}\Big\},$$

for $0.3 \geq \delta > 0.21$

$$\mathcal{A}_1^{(\delta)} = \Big\{\overline{S}_1^{(3)}, \{(1,1,2),(1,2,2)\}, \{(1,1,1),(1,2,2)\}, \{(1,2,2)\},$$
$$\{(1,1,1),(1,1,2)\}\Big\},$$

for $0.7 \geq \delta > 0.3$

$$\mathcal{A}_1^{(\delta)} = \Big\{\overline{S}_1^{(3)}, \{(1,1,2),(1,2,2)\}, \{(1,1,1),(1,2,2)\}, \{(1,2,2)\}\Big\},$$

for $0.79 \geq \delta > 0.7$

$$\mathcal{A}_1^{(\delta)} = \Big\{\overline{S}_1^{(3)}, \{(1,1,2),(1,2,2)\}, \{(1,1,1),(1,2,2)\}\Big\}$$

for $0.91 \geq \delta > 0.79$

$$\mathcal{A}_1^{(\delta)} = \Big\{\overline{S}_1^{(3)}, \{(1,1,2),(1,2,2)\}\Big\}$$

for $1 \geq \delta > 0.91$

$$\mathcal{A}_1^{(\delta)} = \Big\{\overline{S}_1^{(3)}\Big\}.$$

$$C(1,1,1) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \end{bmatrix},$$

$$C(1,1,2) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$C(1,2,2) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$G(1) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$H(1) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix},$$

Let us first discuss the problem of DC at time 3. From the structure of the matrix $C(1,2,2)$ *we see that conditions (2.17) and (2.18) can not be satisfied with* X *such that* $(1,2,2) \in X$ *and therefore the system is not DC at time 3. Moreover it is not SCWP* δ *at time 3 for all* $\delta > 0.7$, *because for each such* δ *and* $X \in \mathcal{A}_1^{(\delta)}$ *we have* $(1,2,2) \in X$. *Now for* $0.7 \leq \delta < 0.79$ *we see that conditions (2.17) and (2.18) could be satisfied only with* $X = \{(1,1,1),(1,1,2)\}$. *However, we have*

$$G_X(1) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \textit{with } \operatorname{rank} G_X(1) = 3$$

and

$$\textit{rank} \begin{bmatrix} G_X(1) & f_1^{(\overline{\pi})} \end{bmatrix} = 4.$$

So the system is not SCWP δ *at time 3 from zero for* $\delta = 0.7$. *Moreover we have*

$$\operatorname{Im} H_X(1) = \operatorname{Im} \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \subset \operatorname{Im} G_X(1)$$

therefore the system is SCWP δ at time 3 to zero for $\delta = 0.7$. Finally notice that conditions (2.17) and (2.18) are satisfied with $X = \{(1,1,2)\}$ and therefore the system is SCWP δ at time 3 for all $\delta < 0.79$, because for such δ we have $\{(1,1,2)\} \in \mathcal{A}_1^{(\delta)}$.

2.4 Controllability at random time

In this paragraph we will discuss the problem of controllability at random time. The idea of the next definition is taken from [58].

Definition 3 *The system (1.1) is $\pi-$weakly controllable, if for all $x_0, x_1 \in R^n$ there exists a control u and a random time τ a.s. finite such that*

$$P_\pi\left(x\left(\tau, x_0, \pi, u\right) = x_1\right) > 0; \tag{2.34}$$

$\pi-$controllable, if this probability can be made equal to one; $\pi-$strongly controllable if it is weakly $\pi-$controllable and $ET_{x_0,x_1} < \infty$, for each $x_0, x_1 \in R^n$ where

$$T_{x_0,x_1} = \min\left\{k : x\left(k, x_0, \pi, u\right) = x_1\right\}. \tag{2.35}$$

Analogically, we introduce the concepts of π-weak controllability, π- controllability and π-strong controllability to zero and from zero. As usually in the case of Dirac initial distribution π of the form $p(i_0) = 1$ we will say about i_0-weak controllability, i_0-controllability and i_0-strong controllability. If system (1.1) is π-weakly controllable, π-controllable or π-strongly controllable (to zero, from zero) for all initial distribution π we will call it weakly controllable, controllable or strongly controllable (to zero, from zero), respectively.

Remark 3 *From the definition it is clear that $\pi-$controllability (to zero, from zero) implies $\pi-$ weak controllability (to zero, from zero). It is also true that $\pi-$strong controllability implies $\pi-$controllability. To prove this statement suppose that (1.1) is $\pi-$strong controllable. Fix $x_0, x_1 \in R^n$ and let control u be such that $ET_{x_0,x_1} < \infty$ with T_{x_0,x_1} given by (2.35). Suppose that*

$$P_\pi\left(x\left(T_{x_0,x_1}, x_0, \pi, u\right) = x_1\right) < 1.$$

Then

$$P_\pi(A) > 0, \tag{2.36}$$

where

$$A = \left\{\omega \in \Omega : x\left(T_{x_0,x_1}, x_0, \pi, u\right) \neq x_1\right\}.$$

Moreover for $\omega \in \Omega$ we have

$$T_{x_0,x_1} = \infty. \tag{2.37}$$

(2.36) together with (2.37) implies that $ET_{x_0,x_1} = \infty$. This contradicts the assumption about $\pi-$strong controllability. Of course the same is true for $\pi-$strong controllability to zero (from zero) and $\pi-$controllability to zero (from zero).

The next theorems show that the problem of $\pi-$weak controllability (to zero, from zero), can be reduced to the problem of i_0-weak (to zero, from zero) for i_0 such that $p(i_0) > 0$.

Theorem 6 *System (1.1) is $\pi-$weakly controllable (to zero, from zero) if and only if there exists $i_0 \in S_\pi^{(1)}$, such that (1.1) is i_0-weakly controllable (to zero, from zero).*

Proof. We will prove only the part about $\pi-$weak controllability the proof of the rest is very similar. Suppose that there exists $i_0 \in S_\pi^{(1)}$, such that (1.1) is i_0-weakly controllable. Fix $x_0, x_1 \in R^n$, then there exists a control u and a random time τ_{i_0} a.s. finite such that

$$P_{i_0}\left(x\left(\tau_{i_0}, x_0, i_0, u\right) = x_1\right) > 0.$$

The control is of the form

$$u = (u(0), u(1), ...),$$

where $u(0)$ is a constant random variable, and $u(k)$ for $k \geq 1$ is of the form $f_k(i_0, r(1), ..., r(k))$. Let us define a new control sequence $\overline{u} = (\overline{u}(0), \overline{u}(1), ...)$ by

$$\overline{u}(k) = \begin{cases} f_k(i_0, r(1), ..., r(k)) & \text{if} \quad r(0) = i_0 \\ 0 & \text{if} \quad r(0) \neq i_0 \end{cases}.$$

Then by the theorem of total probability we have

$$P_\pi\left(x\left(\tau_{i_0}, x_0, \pi, \overline{u}\right) = x_1\right) \geq P_\pi\left(x\left(\tau_{i_0}, x_0, \pi, \overline{u}\right) = x_1 \middle| r(0) = i_0\right) p(i_0) =$$
$$P_{i_0}\left(x\left(\tau_{i_0}, x_0, i_0, u\right) = x_1\right) p(i_0) > 0.$$

It means that system (1.1) is $\pi-$weakly controllable. Suppose now that (1.1) is $\pi-$weakly controllable. Fix $x_0, x_1 \in R^n$ and let a control u and a random time τ are such that

$$P_\pi\left(x\left(\tau, x_0, \pi, u\right) = x_1\right) > 0.$$

Again by the theorem of total probability we have

$$0 < P_\pi\left(x\left(\tau, x_0, \pi, u\right) = x_1\right) =$$

$$\sum_{i \in S_\pi^{(1)}} P_\pi\left(x\left(\tau, x_0, \pi, u\right) = x_1 \middle| r(0) = i\right) p(i) =$$

$$\sum_{i \in S_\pi^{(1)}} P_i\left(x\left(\tau, x_0, i, u\right) = x_1\right) p(i).$$

From the last inequality we conclude that at least one of the number

$$P_i\left(x\left(\tau, x_0, i, u\right) = x_1\right), i \in S_\pi^{(1)}$$

is positive, say i_0, and therefore (1.1) is i_0-weakly controllable. ■

The next two theorems establish the relationships between $\pi-$ controllability (to zero, from zero), $\pi-$ strong controllability (to zero, from zero) and $i-$ controllability (to zero, from zero), $i-$ strong controllability (to zero, from zero).

Theorem 7 *System (1.1) is $\pi-$ controllable (to zero, from zero) if and only if for all $i \in S_\pi^{(1)}$ system (1.1) is $i-$controllable (to zero, from zero).*

Proof. As previously we will prove only the part about $\pi-$ controllability. Suppose that (1.1) is $\pi-$ controllable. Fix x_0, $x_1 \in R^n$ and let a control u and a random time τ are such that

$$P_\pi\left(x\left(\tau, x_0, \pi, u\right) = x_1\right) = 1. \tag{2.38}$$

By theorem of total probability we have

$$P_\pi\left(x\left(\tau, x_0, \pi, u\right) = x_1\right) = \sum_{i \in S_\pi^{(1)}} P_\pi\left(x\left(\tau, x_0, \pi, u\right) = x_1 \middle| r(0) = i\right) p(i) =$$

$$\sum_{i \in S_\pi^{(1)}} P_i\left(x\left(\tau, x_0, i, u\right) = x_1\right) p(i).$$

If one of the numbers $P_i\left(x\left(\tau, x_0, i, u\right) = x_1\right)$, $i \in S_\pi^{(1)}$ was strictly less than 1, then the right hand side of the last equality would be less than 1 too. However the left hand side is by (2.38) equal to 1, therefore

$$P_i\left(x\left(\tau, x_0, i, u\right) = x_1\right) = 1$$

for all $i \in S_\pi^{(1)}$. This means that system (1.1) is $i-$controllable for all $i \in S_\pi^{(1)}$. Suppose now that (1.1) is $i-$controllable for all $i \in S_\pi^{(1)}$. Fix x_0, $x_1 \in R^n$ and let a control $u_i, i \in S_\pi^{(1)}$ and a random time τ_i are such that

$$P_i\left(x\left(\tau_i, x_0, i, u\right) = x_1\right) = 1.$$

The control u_i is of the form

$$u_i = \left(u_i(0), u_i(1), \ldots\right),$$

where $u_i(0)$ is a constant random variable, and $u_i(k)$ for $k \geq 1$ is of the form $f_k(i, r(1), ..., r(k))$. Let us define a new control sequence $\overline{u} = (\overline{u}(0), \overline{u}(1), \ldots)$ by

$$\overline{u}(k) = f_k(i, r(1), ..., r(k)) \text{ if } r(0) = i \text{ for } i \in S_\pi^{(1)}.$$

Define a random variable τ by

$$\tau = \min\left\{k : x\left(k, x_0, \pi, \overline{u}\right) = x_1\right\}.$$

By the theorem of total probability we have

$$P\left(\tau < \infty\right) = \sum_{i \in S_\pi^{(1)}} P\left(\tau < \infty \middle| r(0) = i\right) p(i) =$$

$$\sum_{i \in S_\pi^{(1)}} P\left(\tau_i < \infty\right) p(i) = 1$$

and

$$P_\pi\left(x\left(\tau, x_0, \pi, \overline{u}\right) = x_1\right) =$$

$$\sum_{i \in S_\pi^{(1)}} P_\pi\left(x\left(\tau, x_0, \pi, \overline{u}\right) = x_1 \middle| r(0) = i\right) p(i) =$$

$$\sum_{i \in S_\pi^{(1)}} P_i\left(x\left(\tau, x_0, i, \overline{u}\right) = x_1\right) p(i) = 1.$$

It means that (1.1) is $\pi-$ controllable.

Theorem 8 *System (1.1) is $\pi-$strongly controllable (to zero, from zero) if and only if it is $i-$strongly controllable (to zero, from zero) for all $i \in S_\pi^{(1)}$.*

We omit the proof, because it can be done using a very similar technique as in the proofs of the previous two theorems.

Having in mind the previous three theorems and the Definition 3 we can formulate the following remark.

Remark 4 *System (1.1) is weakly controllable (controllable, strongly controllable) if and only if it is $i-$weakly controllable ($i-$controllable, $i-$strongly controllable) for all $i \in S$. The same is true for controllability of all these types to zero and from zero.*

The next theorem shows that to check weak controllability or weak controllability to zero it is enough to check $i-$weak controllability or $i-$weak controllability to zero for recurrent state $i \in S$ (see, Definition 12 in Appendix).

Theorem 9 *System (1.1) is weakly controllable (to zero) if only if for each recurrent state i (1.1) is $i-$weakly controllable (to zero)*

Proof. We prove only the part about weak controllability, the proof for weak controllability to zero is very similar. If the system (1.1) is weakly controllable then according to the definition it is $i-$weakly controllable for each $i \in S$ and in particular for all i recurrent.

Suppose now that (1.1) is $i-$weakly controllable for all recurrent i. Fix $\overline{x}_0$, $x_1 \in R^n$, an initial distribution π and $j \in S_\pi^{(1)}$. Then according to Theorem 55 in Appendix there exists recurrent state i_0 which is accessible from j in a.s. finite time. Due to our assumption for each x_0, $x_1 \in R^n$ there exists a control $u_{x_0,x_1} = (u_{x_0,x_1}(0), u_{x_0,x_1}(1), ...)$ and a random variable $\overline{\tau}$ a.s. finite such that

$$P_{i_0}\left(x\left(\overline{\tau}, x_0, i_0, u_{x_0,x_1}\right) = x_1\right) > 0.$$

Next let τ_{i_0} be the time of the first visit in state i_0 (see Remark 9 in the Appendix). We know that τ_{i_0} is a.s finite. Define new control sequence $u = (u(0), u(1), ...)$ by

$$u(k) = \begin{cases} 0 \text{ for } k < \tau_{i_0} \\ u_{\overline{x}_0,x_1}(k - \tau_{i_0}) \text{ for } k \geq \tau_{i_0} \end{cases},$$

where $\overline{x}_0 = x\left(\tau_{i_0}, x_0, i_0, 0\right)$, then it is clear that

$$P_\pi\left(x\left(\overline{\tau} + \tau_{i_0}, x_0, \pi, u\right) = x_1\right) > 0$$

and $\overline{\tau} + \tau_{i_0}$ is a.s. finite, so (1.1) is weakly controllable. ■

Notice that the proof does not work for controllability from zero. The reason is that even if we take $x_0 = 0$ we cannot guarantee that $x\left(\tau_{i_0}, x_0, i_0, 0\right) = 0$.

Our next goal is to show that weak controllability, controllability and strong controllability are equivalent. We will show the equivalence in the following way: first we will show a necessary condition for weak controllability (Theorem 10) and next we will prove that this condition is sufficient for strong controllability (Theorem 11). According to Remark 3 this means that the three concepts are equivalent.

Theorem 10 *The necessary condition for weak controllability and weak controllability from zero of the system (1.1) is that for all closed communicating classes C of S there exists a sequence $(i_0, ..., i_{T-1}) \in C^T$ such that*

$$p(i_0, i_1)...p(i_{T-2}, i_{T-1}) > 0$$

and

$$rank\ L = n,$$

where

$$L = \left[F(T,T)B_{i_{T-1}},\ F(T,T-1,i_{T-1})B_{i_{T-2}},\ F(T,1,i_{T-1},...,i_1)B_{i_0}\right].$$

Proof. We prove only the part about weak controllability the proof for weak controllability from zero is very similar. Suppose that the system is weakly controllable, fix a closed communicating class C of S and $\overline{i}_0 \in C$. For each $T = 1, 2, ...$ denote

$$C_T = \left\{(i_0, ...i_{T-1}) \in C^T : p(i_0, i_1)...p(i_{T-2}, i_{T-1}) > 0\right\}$$

so C_T is the set of all possible sample paths of the Markov chain which start from i_0. For $c_T = (i_0, ...i_{T-1}) \in C_T$ define a subspace

$$Q\left(c_T\right) =$$

$$\text{Im}\left[F(T,T)B_{i_{T-1}}, F(T,T-1,i_{T-1})B_{i_{T-2}},\ F(T,1,i_{T-1},...,i_1)B_{i_0}\right]$$

of R^n. Assume that for all $T = 1, 2, ...$ and all $c_T \in C_T$

$$\dim Q\left(c_T\right) < n. \tag{2.39}$$

By the Cayley-Hamilton Theorem the set

$$\left\{Q\left(c_T\right) : T = 1, 2, ..., c_T \in C_T\right\}$$

is finite. This together with (2.39) implies that there exists

$$\overline{x} \in R^n \setminus \bigcup_{T=1,2,...} \bigcup_{c_T \in C_T} Q\left(c_T\right). \tag{2.40}$$

The i_0-weak controllability means that for $\overline{x}$ there exist a control u and a random time τ a.s. finite such that

$$P\left(A\right) > 0,$$

where

$$A=\left\{\omega\in\Omega:\sum_{t=0}^{\tau-1}F\left(\tau,t+1,r(t+1),...,r(\tau-1)\right)B\left(r\left(t\right)\right)u\left(t\right)=\overline{x}\right\}$$

Denote

$$\Omega_T=\{\omega\in\Omega:\tau\left(\omega\right)=T\},T=1,2,...$$

Since τ is a.s. finite then there exists T_0 such that

$$0<P\left(\Omega_{T_0}\cap A\right)=$$

$$P\left(\sum_{t=0}^{T_0-1}F\left(T_0,t+1,r(T_0-1),...,r(t+1)\right)B\left(r\left(t\right)\right)u\left(t\right)=\overline{x}\right)=$$

$$\sum_{c_{T_0}=(i_0,...i_{T_0-1})\in C_{T_0}}P\left(\sum_{t=0}^{T_0-1}F\left(T_0,t+1,r(T_0-1),...,r(t+1)\right)\right|$$

$$B\left(r\left(t\right)\right)u\left(t\right)=\overline{x}|\,r(1)=i_1,...r(T_0-1)=i_{T_0-1})\times$$
$$P\left(r(1)=i_1,...r(T_0-1)=i_{T_0-1}\right)$$

This implies that there exists $\overline{c}_{T_0}=\left(\overline{i}_0,...\overline{i}_{T_0-1}\right)\in C_{T_0}$ such that

$$\sum_{t=0}^{T_0-1}F\left(T_0,t+1,\overline{i}_{T_0-1},...,\overline{i}_t\right)B\left(r\left(t\right)\right)u\left(t\right)=\overline{x}.$$

It contradicts (2.40). ■

Theorem 11 *If for all closed communicating classes C of S there exists a sequence $(i_0,...,i_{T-1})\in C$ such that $p(i_0,i_1)...p(i_{T-2},i_{T-1})>0$ and*

$$rank\ L=n \tag{2.41}$$

where

$$L=\left[F(T,T)B_{i_{T-1}},\ F(T,T-1,i_{T-1})B_{i_{T-2}},\ F(T,1,i_{T-1},...,i_1)B_{i_0}\right]$$

then the system (1.1) is strongly controllable.

Proof. Fix a closed communicating class C of S, $(x_0,i_0)\in R^n\times C$ and $x\in R^n$ and assume that the condition (2.41) holds. According to Theorem 9 it is enough to show that (1.1) is i_0-strongly controllable. By assumptions there exists T_0 and $\overline{c}_{T_0}=\left(\overline{i}_0,...\overline{i}_{T_0-1}\right)\in C_{T_0}$ such that

$$rankQ\left(\overline{c}_{T_0}\right)=n.$$

It implies that $Q\left(\overline{c}_{T_0}\right)Q'\left(\overline{c}_{T_0}\right)$ is nonsingular and for all $x_0\in R^n$ we can define the sequence of vectors $u_{i_0}(x_0),...,u_{i_{T-1}}(x_0)\in R^m$ by

$$\left[u'_{i_0}(x_0),...,u'_{i_{T-1}}(x_0)\right]'=$$
$$\left(Q\left(\overline{c}_{T_0}\right)Q'\left(\overline{c}_{T_0}\right)\right)^{-1}Q'\left(\overline{c}_{T_0}\right)\left(x-F(k,0,\overline{i}_0,...\overline{i}_{T_0-1})x_0\right).$$

Consider the control $\overline{u}$ given by the following algorithm

Step 1: If $r(n)=i_0$ put $l=n,\overline{u}(n)=u_{i_0}\left(x(l)\right)$ and go to step 2, otherwise put $\overline{u}(n)=0$ and go to step 1

Step $k:(k=2,...,T-1)$ If $r(n)=i_{k-1}$ put $\overline{u}(n)=u_{i_{k-1}}\left(x(l)\right)$ and go to step $k+1$, otherwise put $\overline{u}(n)=0$ and go to step 1

Step T : If $r(n)=i_{T-1}$ put $\overline{u}(n)=u_{i_{T-1}}(x(l))$ and stop, otherwise put $\overline{u}(n)=0$ and go to step 1

Consider the random time

$$\eta_{i_0,...i_{T-1}}=\min\{l\geq T:r(l-1)=i_{T-1},...,r(l-T)=i_0\}.$$

By Theorem 62 in Appendix we know that

$$E\left(\eta_{i_0,...i_{T-1}}\middle|\,r(0)=j_0\right)<\infty$$

and

$$P\left(x\left(\eta_{i_0,...i_{T-1}},x_0,j_0,\overline{u}\right)=x\right)\geq\sum_{l=T}^{\infty}P\left(\eta_{i_0,...i_{T-1}}=l\right)=1,$$

because on each set $\left\{\eta_{i_0,...i_{T-1}}=l\right\}$ we have

$$x\left(\eta_{i_0,...i_{T-1}},x_0,j_0,\overline{u}\right)=x$$

■

As we have mentioned before from Theorem 10 and Theorem 11 the following theorem follows:

Theorem 12 *For system (1.1) the weak controllability, controllability and strong controllability are equivalent. Moreover each of these conditions is equivalent to the following: for all closed communicating classes C of S there exists a sequence $(i_0,...,i_{T-1})\in C$ such that $p(i_0,i_1)...p(i_{T-2},i_{T-1})>0$ and*

$$rank\ L=n. \tag{2.42}$$

where

$$L=\left[F(T,T)B_{i_{T-1}},\ F(T,T-1,i_{T-1})B_{i_{T-2}},\ F(T,1,i_{T-1},...,i_1)B_{i_0}\right].$$

Remark 5 *From Theorem 10 we know that condition (2.42) is a necessary condition for weak controllability from zero and from Theorem 11 that it is a sufficient condition for strong controllability, and in particular for strong controllability from zero. Therefore weak controllability from zero, controllability from zero and strong controllability from zero are equivalent, and they are equivalent to (2.42) too.*

Remark 6 *Having in mind the interpretation of condition (2.42) for deterministic time-varying system that corresponds to the sequence*

$$(i_0, ..., i_{T-1})$$

(see, (1.1)) we can reformulate the condition for weak controllability, controllability and strong controllability by saying that for all closed communicating classes C of S there exists a sequence

$$(i_0, ..., i_{T-1}) \in C$$

such that

$$p(i_0, i_1)...p(i_{T-2}, i_{T-1}) > 0$$

and the deterministic time varying system that corresponds to

$$(i_0, ..., i_{T-1})$$

is controllable in the deterministic sense.

We can use the same arguments as in the proofs of Theorem 10 and Theorem 11 to show that for system (1.1) weak controllability to zero, controllability to zero and strong controllability to zero are equivalent and each of them is equivalent to the following: for all closed communicating classes C of S there exists a sequence $(i_0, ..., i_{T-1}) \in C$ such that $p(i_0, i_1)...p(i_{T-2}, i_{T-1}) > 0$ and the deterministic time varying system that corresponds to $(i_0, ..., i_{T-1})$ is controllable to zero in deterministic sense. Therefore we have the following result.

Theorem 13 *For system (1.1) the weak controllability to zero, controllability to zero and strong controllability to zero are equivalent. Moreover each of these conditions is equivalent to the following: for all closed communicating classes C of S there exists a sequence $(i_0, ..., i_{T-1}) \in C$ such that $p(i_0, i_1)...p(i_{T-2}, i_{T-1}) > 0$ and*

$$\operatorname{Im} F(T, 0, i_{T-1}, ..., i_0) \subset \operatorname{Im}\left[F(T, T) B_{i_{T-1}} ... F(T, 1, i_{T-1}, ..., i_1 B_{i_0}\right].$$

It appears that the properties of i-strong controllability and i-strong controllability to zero are properties of class of state. More precisely we have the following theorem.

Theorem 14 *If C is a closed commutating class of state and system (1.1) is i-strongly controllable (i-strongly controllable to zero) for certain $i \in C$ then it is j-strongly controllable (j-strongly controllable to zero) for all $j \in C$.*

Proof. Let τ_j denotes the time of the first visit to state j starting from i defined as follows $\tau_j = \min\{k \geq 1 : r(k) = j\}$ for $P(r(0) = i) = 1$. Since i and j are recurrent, then (see Remark 9)

$$E(\tau_j | r(0) = i) < \infty. \tag{2.43}$$

Fix $x_0, x_1 \in R^n$. From the assumption about i-strong controllability we known that for each $\overline{x}_0 \in R^n$ there exists a control $\overline{u}_{\overline{x}_0}$ and a random time $\tau_{\overline{x}_0}$ a.s. finite such that

$$P_i(x(\tau, \overline{x}_0, i, u) = x_1) = 1$$

and $ET_{\overline{x}_0, x_1} < \infty$, for each $x_0, x_1 \in R^n$ where

$$T_{\overline{x}_0, x_1} = \min\{k : x(k, \overline{x}_0, i, u) = x_1\}. \tag{2.44}$$

Define a new control u by

$$u(k) = \begin{cases} 0 \text{ for } k < \tau_j \\ \overline{u}_{x(\tau_j, \overline{x}_0, i, \overline{u})}(k - \tau_j) \text{ for } k \geq \tau_j \end{cases}.$$

Then it is clear that

$$P_j(x(\tau_j + \tau, x_0, i, u) = x_1) = 1$$

and

$$T_{x_0, x_1} = \min\{k : x(k, x_0, j, u) = x_1\} \leq \tau_j + T_{\overline{x}_0, x_1}.$$

Therefore by (2.43) and (2.44) we have $ET_{x_0, x_1} < \infty$. The proof for i-strong controllability to zero can be done in similar way. ■

The equivalence enables us to consider in the remaining part only the concept of strong controllability or strong controllability to zero.

To demonstrate our considerations we present an example.

Example 6 *Consider the system (1.1) with $S = \{1,2,3,4,5\}$,*

$$A_1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},\ A_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},\ A_3 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

$$A_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, A_5 = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$$

$$B_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix},\ B_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, B_3 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, B_4 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, B_5 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$[p(i,j)]_{i,j=1,2,3,4} = \begin{bmatrix} p(1,1) & p(1,2) & p(1,3) & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & p(3,3) & p(3,4) & p(3,5) \\ 0 & 0 & p(4,3) & p(4,4) & p(4,5) \\ 0 & 0 & 0 & p(5,4) & p(5,5) \end{bmatrix}.$$

All elements not denoted by 0 are nonzero entries. The state space of this chain consists of two closed communicating classes of recurrent states $\{2\}$ and $\{3,4,5\}$. The state 1 is a transient state.

The system is not strongly controllable. It follows from Theorem 12. If we take $C = \{2\}$ then there is a unique sequence $(i_0, ..., i_{T-1}) \in C^T$ such that $p(i_0, i_1)...p(i_{T-2}, i_{T-1}) > 0$, namely $(2, ..., 2)$ and by Caley-Hamilton Theorem

$$rank\,[F(T,T)B_2 \quad \cdots \quad F(T,1,2,...,2)B_2] = rank\,[B_2 \qquad A_2B_2],$$

but

$$rank\,[B_2 \qquad A_2B_2] = 1.$$

For the same reason the system is not 2-strongly controllable. According to Theorem 14 problems of i-strong controllability for $i = 3,4,5$ are equivalent. Using this theorem for sequence $(3,3) \in S_3^{(2)}$ we have

$$rank\,[B_3 \qquad A_3B_3] = 2$$

and the system is i-strongly controllable for $i = 3,4,$ and 5.

Now let us discuss the problem of strong controllability to zero. We use theorem Theorem 13 with $(2,2) \in S_2^{(2)}$ and obtain

$$\operatorname{Im} F(2,0,i_1 = 1,...,i_0 = 1) = \operatorname{Im} A_2^2 = \{0\}.$$

Moreover because $i-$strong controllability implies $i-$strong controllability to zero we know that the system is $i-$strongly controllable to zero for $i = 2,3,4,5$ and by Theorem 13 it is strongly controllable to zero.

2.5 Comparison and discussion

Assume that the initial distribution π is given as $p(i_0) = 1$ and try to discuss the relationships between $\pi-$CWRE at time N, SCWP δ at time N and $\pi-$strong controllability. First notice that if the system is SCWP δ at time N for each recurrent i_0 then according to Remark 1 there exists $\alpha \in \overline{S}_{i_0}^{(N)}$ such that the deterministic system which corresponds to α is controllable. Therefore by Corollary 1 the system is CWRE at time N and strongly controllable. Example 3 shows that CWRE at time N does not imply strong controllability. It also demonstrates that neither strong controllability nor CWRE at time N implies DC at time N. It is so because the necessary condition for DC at time N is that all deterministic systems corresponding to each element of $\overline{S}_{i_0}^{(N)}$ are controllable and we have systems without this property which are CWRE at time N (Example 3). Moreover, when we consider a Markov chain with one class of transient states then the existence of one element in $\overline{S}_{i_0}^{(N)}$,such that the corresponding deterministic system is controllable, is enough for the strong controllability. Finally, when we have a system which is strongly controllable then there exist natural N and $\delta > 0$ such that it is SCWP δ at time N and, according to the above considerations, CWRE at time N. In fact for fixed $x_1,\ x_2 \in R^n$ from the strong controllability we conclude that there are a control u and an almost sure finite random variable T such that

$$P\left(x\left(T, x_0, i_0, u\right) = x_1\right) = 1. \tag{2.45}$$

Since T is a.s. finite there exists at least one natural number N such that

$$P\left(T = N\right) := \delta > 0. \tag{2.46}$$

From (2.45) and (2.46) we conclude that the system is SCWP δ at time N.

Chapter 3

Stability

Stability is a qualitive property crucial for functioning of all systems both natural and man-made, and it is usually the first requirement to be considered in practical applications. Therefore much more results are available for stability of jump linear systems than for controllability.

A natural approach to study stability of stochastic systems is to adopt techniques used for checking stability of deterministic systems such as the second Lyapunov's method. A stochastic version of the Lyapunov's second method was developed almost simultaneously in [15] and [64]. In [72] and [55] the stability properties of general stochastic systems has been systematically investigated. Kozin in [68] clarified many confusing stability concepts and gave a nice explanation of the relationships among various stochastic stability concepts. Certain generalization of the Lyapunov's second method has been proposed in [11], [12]. Basing on this extension some sufficient conditions for almost sure sample stability of continuous time linear differential equations with random time-varying communication delays can be found. This approach has been used to study the stochastic stability of jump linear systems in [71], [88], [89], [61], [46], [82].

It is noted in [68] that almost sure stability does not imply moment stability. The opposite statement is true. However second moment stability criteria, which are easy to check, for almost sure stability are to conservative. This has been illustrated in [84] and [69], where it is shown that the regions of second moment stability are considerably smaller than the ones for almost sure stability. Practically we are interested in the stability of individual sample paths of the system rather than the stability of their moments. Therefore, almost sure stability is a more useful concept than moment stability. In [5], [6] the Lyapunov exponent method has been suggested as the best method to obtain tightest criteria for almost sure stability for general stochastic system. This approach has been adopted to jump linear systems in [47], [44], [48]. It seems from this approach that the almost sure stability problem for jump linear systems can be completely solved by determining the sign of the largest Lyapunov exponent. However the determination of the sign of the top Lyapunov exponent, is a very complicated and computationally difficult task as demonstrated in [45]. Thus, the estimation of the largest Lyapunov exponent becomes a very important research area.

In [4] relationships between sample path and moment stability for a special class of continuous-time jump linear systems was described. Similar results for discrete-time jump linear systems with jumps forming sequence of independent identically distributed random variables was announced in [39], however as shown in [41], the main result is incorrect. The possibility of using the approach from [4] to discrete-time jump linear systems was discussed in [78], where it was heuristically shown that the approach from [4] cannot be used for discrete-time jump linear systems. The main result of [4] states that the δ-moment largest Lyapunov exponent is differentiable and its derivative at zero is just the largest sample path Lyapunov exponent. For one dimensional jump linear systems it was in [38] and [46] proved that almost sure stability is equivalent to $\delta-$moment stability for certain δ. The generalizations of this result for multidimensional discrete-time jump linear systems was presented in [44]. These results are very important, however they do not help to check almost sure stability of a particular system. It is so because there is no explicit formula for the largest $\delta-$moment Lyapunov exponent and consequently the value of its derivative at zero cannot be evaluated. The lack of conditions to check the almost sure stability as well as $\delta-$moment stability is undoubtedly the most important challenge of further investigation.

From the point of view of LQ theory, which is discussed in the next chapter, $\delta-$moment stability for $\delta = 2$ is crucial. The $\delta-$moment stability for $\delta = 2$, called also mean square stability is the best developed concept of stability for jump linear systems. In the literature we may find many necessary and sufficient conditions for this type of stability (see [24]). This conditions are given in terms of spectral radius of certain matrix to be less than one or in the terms of existence of non-negative definite solution of a set of linear matrix equations. The set of equations is called a coupled Lyapunov equation. Numerical properties such as sensitivity of coupled Lyapunov equations are discussed in [35] and [37]. Another characterization of the mean square stability is in terms of spectral radius of a certain matrix to be less than one. However, the large size this matrix causes problems. In [36] an upper bound for the spectral radius is found. The computation of the bound are numerically simpler but gives only sufficient conditions for mean square stability.

This chapter is organized as follows. In the next section we introduce different types of stability and discuss their basic properties. In Section 3.2 the results for one dimensional systems are presented. The best developed concept of stability, namely mean square stability, is discussed in Section 3.3. Sections 3.4 and 3.5 are devoted to almost sure and $\delta-$moment stability, respectively. Finally in Section 3.6 the relationships between the introduced types of stability as well as general discussion are made.

3.1 Different concepts of stability

We define different concepts of stability for the uncontrolled discrete jump linear system:

$$x(k+1) = A(r(k))\,x(k)\,, k \geq 0 \tag{3.1}$$

We discuss the following concepts of stability.

Definition 4 *System (3.1) is said to be*

1. $\pi-$almost surely stable ($\pi-$ASS), if for all $x_0 \in R^n$,
$$P_\pi\left(\lim_{N\to\infty} \|x(N, x_0, \pi)\| = 0\right) = 1$$
If it is $\pi-$ASS for all initial distributions π then we say that it is almost surely stable (ASS).

2. $\pi, \delta-$moment stable ($\pi, \delta-$ MS), if for all $x_0 \in R^n$,
$$\lim_{N\to\infty} E_\pi \|x(N, x_0, \pi)\|^\delta = 0$$
If it is $\pi, \delta-$MS for all initial distributions π then we say that it is $\delta-$moment stable ($\delta-$MS).

3. $\pi-$mean stable, if for all $x_0 \in R^n$,
$$\lim_{N\to\infty} E_\pi x(N, x_0, \pi) = 0\ .$$
If it is $\pi-$mean stable for all initial distributions π then we say that it mean stable.

The next theorem reduces investigation of stability (where stability is defined in one of the ways 1-3) to $\pi-$stability for each Dirac initial distribution π.

Theorem 15 *System (3.1) is*

1. ASS if and only if it is $\pi-$ASS for all Dirac initial distributions π.
2. $\delta-$MS if and only if it is $\pi,\delta-$MS for all Dirac initial distributions π.
3. mean stable if and only if it is $\pi-$mean stable for all Dirac initial distributions π.

Proof. We prove only the first point. The proof of the rest is very similar. Implication $\Rightarrow$ follows from the definition of ASS. To prove the opposite implication let fix an initial distribution π and an initial condition $x_0 \in R^n$. From the theorem of total probability we have

$$P_\pi\left(\lim_{N\to\infty}\|x(N,x_0,\pi)\|=0\right)=$$

$$\sum_{i\in S}P\left(\lim_{N\to\infty}\|x(N,x_0,i)\|=0\Big|\,r(0)=i\right)p(i)=\sum_{i\in S}p(i)=1.$$

■

It is interesting that problems of ASS and $\delta-$MS may be reduced to problems $\pi-$ASS and $\pi,\delta-$MS, respectively, for one initial distribution π, as it is shown in the next theorem.

Theorem 16 *System (3.1) is*

1. ASS if and only if it is $\pi-$ASS for certain initial distribution π such that $P(r(0)=i)>0$ for all $i\in S$.
2. $\delta-$MS if and only if it is $\pi,\delta-$MS for certain initial distribution π such that $P(r(0)=i)>0$ for all $i\in S$.

Proof. We prove only the first point. The proof of the rest is very similar. Implication $\Rightarrow$ follows from the definition of ASS. To prove the opposite implication let assume that system (3.1) is $\pi-$ASS for certain initial distribution π such that $P(r(0)=i)>0$. Again by the theorem of total probability we have

$$1=P_\pi\left(\lim_{N\to\infty}\|x(N,x_0,\pi)\|=0\right)=$$

$$\sum_{i\in S}P\left(\lim_{N\to\infty}\|x(N,x_0,i)\|=0\Big|\,r(0)=i\right)p(i) \qquad (3.2)$$

By Theorem 15 it is enough to show that

$$P\left(\lim_{N\to\infty}\|x(N,x_0,i)\|=0\Big|\,r(0)=i\right)=1$$

for all $i\in S$. Suppose that there exists $i_0\in S$ such that

$$P\left(\lim_{N\to\infty}\|x(N,x_0,i_0)\|=0\Big|\,r(0)=i_0\right)<1.$$

Then by the assumption that $P(r(0)=i)>0$ for all $i\in S$ we have

$$\sum_{i\in S}P\left(\lim_{N\to\infty}\|x(N,x_0,i)\|=0\Big|\,r(0)=i\right)p(i)<\sum_{i\in S}p(i)=1.$$

This contradicts with (3.2). ■

In the next example we show that Theorem 16 is not true for mean stability.

Example 7 *Consider system (3.1) with $S=\{1,2\}$,*

$$P=\begin{bmatrix}1&0\\0&1\end{bmatrix},\ A_1=1,\ A_2=-1.$$

Then for initial distribution $\pi: P(r(0)=1)=P(r(0)=2)=\frac{1}{2}$, we have

$$E_\pi x(N,x_0,\pi)=0,\ N\geq 1$$

and this system is $\pi-$mean stable. Whereas for initial distribution

$$\tilde{\pi}: P(r(0)=1)=1, P(r(0)=2)=0,$$

we have

$$E_{\tilde{\pi}}x(N,x_0,\tilde{\pi})=x_0,\ N\geq 1$$

so the system is not mean stable even though it is $\pi-$mean stable for certain positive π.

The $\delta-$MS has a kind of monotony property given by the following Theorem.

Theorem 17 *If (3.1) is $\delta-$MS then it is $\delta'-$MS for all $\delta'<\delta$.*

Proof. Fix $i_0\in S$, and consider initial distribution π as follows

$$P(r(0)=i_0)=1.$$

If $g: R\to R$ is convex and X is a random variable such that EX exists then by Jensen inequality we have

$$E_{i_0}g(X)\geq g(E_{i_0}X).$$

Using this inequality with $g(x) = x^{\frac{\delta}{\delta'}}$ and $X = \|x(N, x_0, i_0)\|^{\delta'}$ we obtain

$$E_{i_0} \|x(N, x_0, i_0)\|^{\delta} \geq \left(E_{i_0} \|x(N, x_0, i_0)\|^{\delta'}\right)^{\frac{\delta}{\delta'}}$$

and consequently if

$$E_{i_0} \|x(N, x_0, i_0)\|^{\delta}$$

tends to zero then, so does

$$E_{i_0} \|x(N, x_0, i_0)\|^{\delta'}$$

■

3.2 Stability of scalar systems

In this chapter we consider the scalar version of system (3.1). In this simplest case we are able to present necessary and sufficient conditions for all introduced types of stability.

Fix a positive number δ and introduce the following notations

$$\varphi_i^{(\delta)}(k) = E\left(|x(k)|^{\delta} 1_{\{r(k)=i\}}\right)$$

$$\varphi^{(\delta)}(k) = \begin{bmatrix} \varphi_1^{(\delta)}(k) \\ \vdots \\ \varphi_s^{(\delta)}(k) \end{bmatrix}$$

$$D_{\delta} = diag\left[|A(i)|^{\delta}\right]_{i=1,\ldots,s}.$$

The next Lemma presents a recurrent formula for $\varphi(k)$.

Lemma 1 *For every natural k we have*

$$\varphi^{(\delta)}(k+1) = P' D_{\delta} \varphi^{(\delta)}(k). \tag{3.3}$$

Proof. For fixed $i \in S$ and natural k we have

$$\varphi_i^{(\delta)}(k+1) = E\left(|x(k+1, x_0, i_0)|^{\delta} 1_{\{r(k+1)=i\}}\right) =$$

$$E\left(|A(r(k))|^{\delta} |x(k, x_0, i_0)|^{\delta} 1_{\{r(k+1)=i\}}\right) =$$

$$\sum_{j \in S} E\left(|A(r(k))|^{\delta} |x(k, x_0, i_0)|^{\delta} 1_{\{r(k+1)=i\}} 1_{\{r(k)=j\}}\right) =$$

$$\sum_{j \in S} |A(j)|^{\delta} E\left(|x(k, x_0, i_0)|^{\delta} 1_{\{r(k+1)=i\}} 1_{\{r(k)=j\}}\right) =$$

$$\sum_{j \in S} |A(j)|^{\delta} E\left(|x(k, x_0, i_0)|^{\delta} 1_{\{r(k)=j\}}\right) p_{ji} =$$

$$\sum_{j \in S} |A(j)|^{\delta} \varphi_j^{(\delta)}(k) p_{ji}.$$

This justifies (3.3). ■

This technical Lemma enables us to formulate necessary and sufficient conditions for δ–MS of scalar systems.

Theorem 18 *The scalar system (3.1) is δ–MS if and only if*

$$\rho(P' D_{\delta}) < 1. \tag{3.4}$$

Proof. Because of the equality

$$E|x(k, x_0, i_0)|^{\delta} = \sum_{j \in S} \varphi_j^{(\delta)}(k)$$

and positivity of $\varphi_j^{(\delta)}(k)$ the δ–MS is equivalent to the following condition

$$\lim_{k \to \infty} \varphi^{(\delta)}(k) = 0.$$

But by (3.3) it is possible if and only if $\rho(P' D_{\delta}) < 1$. ■

An immediate consequence of the proof is the following remark.

Remark 7 *If scalar system (3.1) is δ–MS then for each $x_0 \in R$ there exist constants $c > 0$ and $0 < a < 1$ such that*

$$E|x(k, x_0, i_0)|^{\delta} \leq c a^k.$$

Remark 8 *Since for any square matrices X, Y of the same size we have $\rho(XY) = \rho(YX)$ and $\rho(X) = \rho(X')$, therefore $\rho(P' D_{\delta}) = \rho(P D_{\delta})$.*

Another characterization of δ–MS is given by the next theorem.

Theorem 19 *The following conditions are equivalent to δ–MS of scalar system (3.1)*

1. *for all positive numbers $Q(i)$, $i \in S$ there exist positive numbers $R(i)$, $i \in S$ such that*

$$(I_s - P D_{\delta}) \begin{bmatrix} R(1) \\ \vdots \\ R(s) \end{bmatrix} = \begin{bmatrix} Q(1) \\ \vdots \\ Q(s) \end{bmatrix} \tag{3.5}$$

2. *there exist positive numbers $Q(i)$, $i \in S$ and $R(i)$, $i \in S$ such that (3.5) holds.*

The proof of this theorem is an immediate consequence of the following general result about matrices with nonnegative entries.

Theorem 20 *If $A = [a_{ij}]_{i,j=1,...,n}$ is such that $a_{ij} \geq 0$, $i, j = 1, ..., n$ then the following conditions are equivalent*

1. $\rho(A) < 1$
2. for all positive numbers $Q(i)$, $i \in S$ there exist positive numbers $R(i)$, $i \in S$ such that

$$(I_s - PD_\delta)\begin{bmatrix} R(1) \\ \vdots \\ R(s) \end{bmatrix} = \begin{bmatrix} Q(1) \\ \vdots \\ Q(s) \end{bmatrix} \tag{3.6}$$

3. there exist positive numbers $Q(i)$, $i \in S$ and $R(i)$, $i \in S$ such that (3.5) holds.

Proof. Suppose that $\rho(A) < 1$. Fix positive numbers $Q(i)$, $i \in S$ and define

$$Q = \begin{bmatrix} Q(1) \\ \vdots \\ Q(s) \end{bmatrix}. \tag{3.7}$$

According to the assumption that $(I - A)$ is invertible and it implies that

$$(I - A)^{-1} = \sum_{n=0}^{\infty} A^n.$$

If we define

$$R = \begin{bmatrix} R(1) \\ \vdots \\ R(s) \end{bmatrix} \in R^s, \tag{3.8}$$

by

$$R = \sum_{n=0}^{\infty} A^n Q,$$

then $R(i) > 0$, $i \in S$ and

$$(I - A)R = Q.$$

The implication $(2 \Rightarrow 3)$ is obvious. Suppose now that there exist positive numbers $Q(i)$, $i \in S$ and $R(i)$, $i \in S$ such that (3.5) holds. Then with the notation (3.8) and (3.7) we have

$$(I - A)R = Q.$$

and consequently

$$AR = R - Q < R, \tag{3.9}$$

where the last inequality is understood as inequality for all coordinates. Now we use the following fact about a matrix with nonnegative entries (see [56]). If A is a matrix with nonnegative entries and $x > 0$ is such that $Ax < \alpha x$, then $\rho(A) < \alpha$. This fact together with (3.9) implies that $\rho(A) < 1$. ■

In the next theorem we present necessary and sufficient conditions for ASS for ergodic Markov chain.

Theorem 21 *Suppose that $r(k)$ is ergodic with limit distribution $\pi = (\pi_i)_{i \in S}$. Then the scalar system (3.1) is ASS if and only if*

$$|A(1)|^{\pi_1} \dots |A(s)|^{\pi_s} < 1. \tag{3.10}$$

Moreover, if (3.10) holds then $x(k, x_0, i_0)$ tends to zero at an exponential rate.

Proof. Denote by $\tau_i(k)$ the time of the Markov chain $r(k)$ being in state $i \in S$ up to moment k, i.e.

$$\tau_i(k) = \sum_{l=0}^{k} 1_{\{r(l)=i\}}.$$

By Theorem 59 we know that

$$\lim_{k\to\infty} \frac{\tau_i(k)}{k} = \pi_i.$$

Therefore for each $i_0 \in S$

$$\lim_{k\to\infty} \frac{1}{k} \ln |x(k, x_0, i_0)| = \lim_{k\to\infty} \frac{1}{k} \ln \left(|A(1)|^{\tau_1(k)} \dots |A(s)|^{\tau_s(k)}\right) =$$

$$\lim_{k\to\infty} \sum_{j\in S} \frac{\tau_j(k)}{k} \ln(|A(j)|) = \sum_{j\in S} \pi_j \ln(|A(j)|)$$

and it is clear that if (3.10) holds then $\lim_{k\to\infty} |x(k, x_0, i_0)| = 0$ a.s. and the convergence is at an expotentional rate. Moreover, if $|A(1)|^{\pi_1} \dots |A(s)|^{\pi_s} > 1$, then $\lim_{k\to\infty} |x(k, x_0, i_0)| = \infty$. To complete the proof it is enough to show that if

$$|A(1)|^{\pi_1} \dots |A(s)|^{\pi_s} = 1 \tag{3.11}$$

then

$$\limsup_{k\to\infty} |x(k, x_0, i_0)| = \infty.$$

Suppose that (3.10) holds. Then, we have

$$\sum_{j\in S} \pi_j \ln(|A(j)|) = 0 \tag{3.12}$$

and also

$$x(k+1) = |A(1)|^{\tau_1(k)} \dots |A(s)|^{\tau_s(k)} |x(0)| =$$

$$|x(0)| \exp\left(\sum_{j\in S}\sum_{l=0}^{k} 1_{\{r(l)=j\}} \ln(|A(j)|)\right) =$$

$$|x(0)| \exp\left(\sum_{l=0}^{k}\sum_{j\in S} 1_{\{r(l)=j\}} \ln(|A(j)|)\right). \tag{3.13}$$

By Theorem 60 we conclude from (3.12) that

$$\limsup_{N\to\infty} \frac{\left(\sum_{l=0}^{k}\sum_{j\in S} 1_{\{r(l)=j\}} \ln(|A(j)|)\right)}{\sqrt{aN \lg\lg N}} = 1, \text{ a.s.},$$

for certain positive constant a. Thus

$$\limsup_{N\to\infty} \sum_{l=0}^{N}\sum_{j\in S} 1_{\{r(l)=j\}} \ln(|A(j)|) = \infty, \text{a.s.},$$

and consequently by (3.13) we conclude that

$$\limsup_{k\to\infty} |x(k, x_0, i_0)| = \infty.$$

■

In the rest of this section we assume that $r(k)$ is ergodic with limit distribution $\pi = (\pi_i)_{i\in S}$. Denote by $\overline{r}(k)$, $k = 0, 1\dots$ sequence of independent identically distributed random variables such that

$$P(\overline{r}(0) = i) = \pi_i, \ i \in S.$$

Of course $\overline{r}(k)$ is also an ergodic Markov chain with transition probability matrix

$$\overline{P} = \begin{bmatrix} \pi_1 & \dots & \pi_s \\ \vdots & \ddots & \vdots \\ \pi_1 & \dots & \pi_s \end{bmatrix}$$

and limit distribution $\pi = (\pi_i)_{i\in S}$. Together with (3.1) we will consider a scalar system governed by $\overline{r}(k)$ and given by

$$\overline{x}(k+1) = A(\overline{r}(k))\,\overline{x}(k), k \geq 0. \tag{3.14}$$

Theorem 21 says that (3.1) is ASS if and only if (3.14) is ASS. Such an equivalence is not true for δ–MS as the following example shows.

Example 8 *Consider system (3.1) with $S = \{1, 2\}$,*

$$P = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{3}{4} & \frac{1}{4} \end{bmatrix}, \ A_1 = a, \ A_2 = b.$$

The limit distribution is $\pi : \pi_1 = \frac{9}{17}$, $\pi_2 = \frac{8}{17}$. For (3.14) we have

$$E_\pi |\overline{x}(N, x_0, \pi)|^\delta = \left(E_\pi |A(\overline{r}(0))|^\delta\right)^{N-1} x_0 = \left(|a|^\delta \pi_1 + |b|^\delta \pi_2\right)^{N-1} x_0,$$

and therefore (3.14) is δ–MS if and only if

$$|a|^\delta \frac{9}{17} + |b|^\delta \frac{8}{17} < 1. \tag{3.15}$$

By Theorem 18 system (3.1) is δ–MS if and only if

$$1 > \rho(PD_\delta) = \rho\left(\begin{bmatrix} \frac{1}{3}a^\delta & \frac{2}{3}b^\delta \\ \frac{3}{4}a^\delta & \frac{1}{4}b^\delta \end{bmatrix}\right) =$$

$$\frac{1}{6}|a|^\delta + \frac{1}{8}|b|^\delta + \frac{1}{24}\sqrt{\left(16|a|^{2\delta} + 264|b|^\delta |a|^\delta + 9|b|^{2\delta}\right)}.$$

For $\delta = 1, a = 2, b = 0.1$ we have $|a|^\delta \frac{9}{17} + |b|^\delta \frac{8}{17} = 1.1 > 1$ and

$$\frac{1}{6}|a|^\delta + \frac{1}{8}|b|^\delta + \frac{1}{24}\sqrt{\left(16|a|^{2\delta} + 264|b|^\delta |a|^\delta + 9|b|^{2\delta}\right)} = 0.7963 < 1.$$

Therefore (3.1) is δ–MS whereas (3.14) is not δ–MS.

The next theorem describes the relationships between δ–MS and ASS.

Theorem 22 *Suppose that $r(k)$ is ergodic. The scalar system (3.1) is ASS if and only if it is $\delta-$MS for certain $\delta > 0$.*

In the proof of this Theorem we will need the following Lemma from [46].

Lemma 2 *For $n \geq 2$, let*

$$H = \begin{bmatrix} -h_1 & h_{12} & \cdots & h_{1n} \\ h_{21} & -h_2 & \cdots & h_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ h_{n1} & h_{n2} & \cdots & -h_n \end{bmatrix} = \begin{bmatrix} H_1 & H_2 & \cdots & H_n \end{bmatrix},$$

where H_j is the jth column of H. Let $b = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}' \in R^n$. Suppose that the entries of H satisfy $h_i > 0$, $h_{ij} \geq 0$, and $\sum_{m \neq i} h_{im} \leq h_i$ for all $i, j = 1, ..., n$ with $\sum_{m \neq i} h_{im} < h_i$ for some i. Then

1. $\det H = (-1)^n \eta$ *for some* $\eta > 0$

2. $\det \begin{bmatrix} H_1 & \ldots & H_{j-1} & b & H_{j+1} & \ldots & H_n \end{bmatrix} = (-1)^{n-1}\gamma_j$ *for some* $\gamma_j > 0$ *and all* $j = 1, ..., n$.

Proof. Suppose that the system (3.1) is $\delta-$MS then according to Remark 7

$$\sum_{k=0}^{\infty} E\left|x\left(k, x_0, i_0\right)\right|^{\delta} < \infty.$$

Let

$$\xi = \limsup_{k \to \infty} \left|x\left(k, x_0, i_0\right)\right|,$$

then from Markov's inequality, we have that for any $c > 0$, the following holds

$$P\left(\xi \geq c\right) = P\left(\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} \left\{\omega : \left|x\left(m, x_0, i_0\right)\right| \geq c\right\}\right) \leq$$

$$\sum_{m=n}^{\infty} P\left(\bigcup_{m=n}^{\infty} \left\{\omega : \left|x\left(m, x_0, i_0\right)\right| \geq c\right\}\right) \leq$$

$$\sum_{m=n}^{\infty} P\left(\left\{\omega : \left|x\left(m, x_0, i_0\right)\right| \geq c\right\}\right) \leq \frac{1}{c^{\delta}} \sum_{m=n}^{\infty} E\left|x\left(m, x_0, i_0\right)\right|^{\delta} \overset{n\to\infty}{\to} 0.$$

Thus with any $c > 0$, $P\left(\xi \geq c\right) = 0$ and therefore $P\left(\xi = 0\right) = 1$. This provides that system (3.1) is ASS. Now we show that ASS implies $\delta-$MS for certain $\delta > 0$. For that purpose we will point 3 of Theorem 20. Namely we

show that for certain $\delta > 0$ equation (3.6) with $Q(i) = |A(i)|^{\delta}$ has a positive solution $R(i)$, $i \in S$. Multiply from right equation (3.6) by D_{δ}^{-1} to obtain

$$\sum_{j \in S} \left(\delta_{ij} \left|A(i)\right|^{-\delta} - p_{ij}\right) R(j) = 1, i \in S. \tag{3.16}$$

If we denote

$$o_i(\delta) = 1 - \delta \ln|A(i)| - |A(i)|^{-\delta}$$

then we have

$$\lim_{\delta \to 0} \frac{o_i(\delta)}{\delta} = 0$$

and equation (3.16) may be rewritten as follows

$$\sum_{j \in S} \left(q_{ij} + \left(\delta \ln|A(i)| + o_i(\delta)\right)\delta_{ij}\right) = -1, \tag{3.17}$$

where $q_{ij} = p_{ij} - \delta_{ij}$. Since

$$P' \begin{bmatrix} \pi_1 \\ \vdots \\ \pi_s \end{bmatrix} = \begin{bmatrix} \pi_1 \\ \vdots \\ \pi_s \end{bmatrix},$$

therefore

$$Q' \begin{bmatrix} \pi_1 \\ \vdots \\ \pi_s \end{bmatrix} = 0, \tag{3.18}$$

where $Q = [q_{ij}]_{i,j=1,\ldots,s}$, and consequently $\det Q = 0$. Denote by Q_i the matrix obtained by striking out the i-th column and i-th row of the matrix Q. Equation (3.18) implies that $rank Q = s - 1$, and therefore there exists $i_0 \in S$ such that $\alpha_{i_0} \neq 0$, where $\alpha_i = \det Q_i$. Notice that equation (3.18) may be rewritten as follows

$$Q'_{i_0} \begin{bmatrix} \pi_1 \\ \vdots \\ \pi_{i_0-1} \\ \pi_{i_0+1} \\ \vdots \\ \pi_s \end{bmatrix} = \pi_{i_0} \begin{bmatrix} -q_{i_0 1} \\ \vdots \\ -q_{i_0\, i_0-1} \\ -q_{i_0\, i_0+1} \\ \vdots \\ q_{i_0\, s} \end{bmatrix},$$

and by Cramer formulas we have

$$\pi_j = \frac{\pi_{i_0}}{\alpha_{i_0}} \alpha_j,\ j \in S.$$

The ergodicity assumption implies that $\pi_j > 0$ and this in turn implies in light of the above formula that $\alpha_j \neq 0$ for all $j \in S$ and that the ratio

$$\frac{\pi_i}{\alpha_i} = \text{constant} \tag{3.19}$$

is constant. Observe that matrices Q_j, $j \in S$ satisfy the assumption of Lemma 2 and therefore

$$\alpha_i = (-1)^{s-1} c_i.$$

This together with (3.19) implies

$$\alpha_i = (-1)^{s-1} \pi_i c, \; i \in S$$

for certain positive c. Using this equality and the properties of determinant we obtain

$$\det\left[q_{ij} + (\delta \ln |A(i)| + o_i(\delta))\, \delta_{ij}\right]_{i,j \in S} = \sum_{i \in S} (\delta \ln |A(i)| + o_i(\delta))\, \alpha_i =$$

$$\delta c \sum_{i \in S} \pi_i \ln |A(i)| + o(\delta), \tag{3.20}$$

where

$$\lim_{\delta \to 0} \frac{o(\delta)}{\delta} = 0.$$

ASS implies that $\sum_{i \in S} \pi_i \ln |A(i)| < 0$ and by (3.20)

$$\det\left[q_{ij} + (\delta \ln |A(i)| + o_i(\delta))\, \delta_{ij}\right]_{i,j \in S} \neq 0$$

for small δ, therefore (3.17) has a unique solution given by

$$R(i) = \frac{-\det F_i(\delta)}{\det\left[q_{ij} + (\delta \ln |A(i)| + o_i(\delta))\, \delta_{ij}\right]_{i,j \in S}},$$

where $F_i(\delta)$ is the matrix obtained by replacing ith column of

$$\left[q_{ij} + (\delta \ln |A(i)| + o_i(\delta))\, \delta_{ij}\right]_{i,j \in S}$$

by $\begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}'$. Denote by G_i the matrix obtained by replacing ith column of Q by $\begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}'$. Again by the properties of determinant we have

$$\det F_i(\delta) = \det G_i + o(1).$$

Moreover by Lemma 2 we have

$$\det G_i = (-1)^{s-1} \gamma_i$$

for positive γ_i and therefore

$$R(i) = \frac{-\left((-1)^{s-1}\gamma_i + o(1)\right)}{\det\left[q_{ij} + (\delta \ln |A(i)| + o_i(\delta))\, \delta_{ij}\right]_{i,j \in S}} =$$

$$\frac{-\left((-1)^{s-1}\gamma_i + o(1)\right)}{\delta c \sum_{i \in S} \pi_i \ln |A(i)| + o(\delta)} > 0$$

for small δ. ■

3.3 Mean square stability

In this paragraph we study the special case of δ–MS, namely the case of $\delta = 2$ called in the literature mean square stability. This type of stability deserves special attention for the following two reasons. Firstly, mean square stability is the only case of δ–MS for which there exists testable necessary and suficcient conditions due to possibility of using the stochastic version of the Lyapunov's second method. Secondly, mean square stability play a crucial role in one of the most important optimal control problems namely in linear quadratic control problem. This is demonstrated in the next chapter.

The following necessary and sufficient conditions for second moment stability of jump linear system have been proved in [61].

Theorem 23 *For system (3.1) the following conditions are equivalent*

1. system (3.1) is mean square stable
2. for all $x_0 \in R^n$ and there exists a constant c such that
$$\sum_{N=1}^{\infty} E_{r(0)=i_0} \left\| x(N, x_0, i_0) \right\|^2 < c, \text{ for all } i_0 \in S$$
3. for all $x_0 \in R^n$ there exist constants $c > 0$ and $0 < a < 1$ such that
$$E_{r(0)=i_0} \left\| x(N, x_0, i_0) \right\|^2 < c a^N, \text{ for all } i_0 \in S \text{ and } N = 0, 1, \ldots$$
4. for each symmetric positive definite matrices Q_i, $i \in S$ there exists solution H_i, $i \in S$ of the following coupled Lyapunov equation
$$H_i - A_i' \left(\sum_{j=1}^{s} p_{ij} H_j \right) A_i = Q_i, \; i \in S. \tag{3.21}$$

The third point of the theorem brings testable conditions for the second moment stability. Next theorem, which is taken from [24] describes mean stability in terms of spectral radius of the following matrix

$$\mathcal{A}_d = (P' \otimes I_{n^2})\,(diag\,(A_i \otimes A_i))\,,$$

where $X \otimes Y$ is the Kronecker product of matrices X and Y.

Theorem 24 *System (3.1) is mean square stable if and only if*

$$\rho(\mathcal{A}_d) < 1. \tag{3.22}$$

Condition (3.22) requires calculating of eigenvalues of the matrix which is of very high dimension (sn^2). Therefore one may be interested in sufficient conditions which are given in terms of matrices of lower sizes. The next theorem supplies such a condition. Denote by $\lambda_1(X)$ the largest eigenvalue of symmetric matrix X. Denote the set of all eigenvalues of a square matrix X by $\sigma(X)$. The following properties of Kronecker product are used in our further consideration. The proof may be found in [75].

Theorem 25 *If A_1 and A_2 are $m \times m$ and B_1, B_2 are $n \times n$ then we have*

$$(A_1A_2) \otimes (B_1B_2) = (A_1 \otimes B_1)\,(A_2 \otimes B_2)\,. \tag{3.23}$$

Moreover, for a complex polynomial $p(x,y) = \sum_{k,l=0}^{p} c_{kl}x^ky^l$ of two variables and for square matrices A and B we have

$$\sigma\left(\sum_{k,l=0}^{p} c_{kl}A^k \otimes B^l\right) = \{p(x,y) : x \in \sigma(A)\,,\; y \in \sigma(B)\}\,. \tag{3.24}$$

Theorem 26 *The discrete time jump linear system is stable if*

$$\lambda_1(PP')\sum_{i=1}^{s}\lambda_1^2(A_iA_i') < 1.$$

In the proof we need the following lemma from [83].

Lemma 3 *Let $X, Y, Z \in R^{n\times n}$ with $X = X'$, $Y = Y'$. Then the following inequalities hold*

$$\lambda_1(X+Y) \le \lambda_1(X) + \lambda_1(Y)\,, \tag{3.25}$$

$$\max_{i=1,\dots,s} |\lambda_i(Z)| \le \sqrt{\lambda_1(ZZ')}. \tag{3.26}$$

Proof. From the properties (3.23) of Kronecker product we have

$$(A_iA_i') \otimes (A_iA_i') = (A_i \otimes A_i)\,(A_i \otimes A_i)' \tag{3.27}$$

This together with definition of $\mathcal{A}_d$ implies

$$\mathcal{A}_d\mathcal{A}_d' = (PP') \otimes \left(\sum_{i=1}^{s}((A_iA_i') \otimes (A_iA_i'))\right).$$

By (3.24) the eigenvalues of $X \otimes Y$ are equal to $\lambda_i(X)\lambda_i(Y)$, and therefore the spectrum $\sigma(\mathcal{A}_d\mathcal{A}_d')$ has the following form

$$\left\{\lambda_k(PP')\lambda_l\left(\sum_{i=1}^{s}((A_iA_i') \otimes (A_iA_i'))\right) : 1 \le k \le n, 1 \le l \le n^2\right\}. \tag{3.28}$$

(3.27) shows that $(A_iA_i') \otimes (A_iA_i')$ is symmetric and nonnegative definite so we can use (3.25) to obtain

$$\lambda_1\left(\sum_{i=1}^{s}((A_iA_i') \otimes (A_iA_i'))\right) \le \sum_{i=1}^{s}\lambda_1\left((A_iA_i') \otimes (A_iA_i')\right) \tag{3.29}$$

Applying (3.24) again we get

$$\lambda_1\left((A_iA_i') \otimes (A_iA_i')\right) = \lambda_1^2\,(A_iA_i') \tag{3.30}$$

Using (3.28), (3.29) and (3.30) $\lambda_1(\mathcal{A}_d\mathcal{A}_d')$can be bounded as follows

$$\lambda_1(\mathcal{A}_d\mathcal{A}_d') \le \lambda_1(PP')\sum_{i=1}^{s}\lambda_1^2\,(A_iA_i')\,. \tag{3.31}$$

Now from (3.26) and (3.31) we obtain

$$\max_{i=1,\dots,s}|\lambda_i(\mathcal{A}_d)| \le \lambda_1(\mathcal{A}_d\mathcal{A}_d') \le \sqrt{\lambda_1(PP')\sum_{i=1}^{s}\lambda_1^2\,(A_iA_i')}\,.$$

The proof is complete. ■

The coupled Lyapunov equation (3.21) plays a crucial role in mean square stability of jump linear systems. Therefore now we present some of its properties.

The first result shows that (3.21) can be derived in four equivalent forms. This theorem has been shown in [24]. To present this theorem it is convenient to introduce the following notation. For any matrices L_i, $i \in S$ of size $n \times n$ let

$$T_{1j}(L_1,\dots,L_s) = \sum_{i=1}^{s}p_{ij}A_iL_iA_i'$$

$$T_{2j}(L_1,...,L_s)=\sum_{j=1}^{s}p_{ij}A_i'L_jA_i$$

$$T_{3j}(L_1,...,L_s)=\sum_{i=1}^{s}p_{ij}A_jL_iA_j'$$

$$T_{4j}(L_1,...,L_s)=\sum_{j=1}^{s}p_{ij}A_j'L_jA_j$$

Theorem 27 *The following conditions are equivalent to mean square stability of system (3.1)*

1. *For some positive definite H_i, $i\in S$ and Q_i, $i\in S$ and some integer $1\leq\eta\leq 4$, we have*

$$H_i-T_{\eta i}(H_1,...,H_s)=Q_i.$$

2. *For any positive definite Q_i, $i\in S$ and any integer $1\leq\eta\leq 4$ there are unique positive definite H_i, $i\in S$ such that*

$$H_i-T_{\eta i}(H_1,...,H_s)=Q_i.$$

For $\eta=2$ we obtain equation (3.21).

In real world the model parameters of the system are not known precisely, for example they come from certain estimation procedures. In such situation we are interested in the difference between the solution with true parameters and the solution with approximate ones. In the next theorem we present upper bounds for the norm of this difference in terms of degree of accuracy of the coefficients. This theorem is taken from [35]. Similar results for continuous time case may be found in [37]. To present this result let introduce the following notations. In the space made up of all s-length sequences of matrices $H=(H_1,...,H_s)$ of size $n\times n$ we consider the following three norms $\|H\|_1=\sum_{i=1}^{s}\|H_i\|$, $\|H\|_2=\sqrt{\sum_{i=1}^{s}\|H_i\|^2}$, $\|H\|_\infty=\max\{\|H_i\|:i=1,...,s\}$, where $\|H_i\|$ is the operator norm of H_i.

Theorem 28 *Suppose that system (3.1) is mean square stable. Let $H=(H_1,...,H_s)$ be the solution of (3.21) for certain positive definite matrices $Q=(Q_1,...,Q_s)$. Suppose that $A=(A_1,...,A_s)$ and $Q=(Q_1,...,Q_s)$ are perturbed by matrices $\Delta A=(\Delta A_1,...,\Delta A_s)$ and $\Delta Q=(\Delta Q_1,...,\Delta Q_s)$, respectively, such that $Q_i+\Delta Q_i$, $i\in S$ are positive definite and system*

$$x(k+1)=(A(r(k))+\Delta A(r(k)))\,x(k),\ k\geq 0$$

is mean square stable. Then the disturbed Lyapunov equation

$$\overline{H}_i-(A_i+\Delta A_i)'\left(\sum_{j=1}^{s}p_{ij}\overline{H}_j\right)(A_i+\Delta A_i)=Q_i+\Delta Q_i,\ i\in S.$$

has a solution $\overline{H}_i=H_i+\Delta H_i$ and we have

$$\frac{\|\Delta H\|_1}{\|H+\Delta H\|_1}\leq\left\|\widetilde{H}\right\|_\infty s\left(K_1\frac{\|\Delta Q\|_1}{\|Q+\Delta Q\|_1}+A\right),$$

$$\frac{\|\Delta H\|_2}{\|H+\Delta H\|_2}\leq\left\|\widetilde{H}\right\|_\infty\sqrt{s}\left(K_2\frac{\|\Delta Q\|_2}{\|Q+\Delta Q\|_2}+B\right),$$

$$\frac{\|\Delta H\|_\infty}{\|H+\Delta H\|_\infty}\leq\left\|\widetilde{H}\right\|_\infty\left(K_\infty\frac{\|\Delta Q\|_\infty}{\|Q+\Delta Q\|_\infty}+C\right),$$

where

$$A=\sum_{i=1}^{s}\|\Delta A_i\|\left(2\|A_i\|+\|\Delta A_i\|\right)\max_{j=1,...,s}\{p_{ij}\}$$

$$B=\sqrt{\sum_{i=1}^{s}\left(\|\Delta A_i\|\left(2\|A_i\|+\|\Delta A_i\|\right)\sqrt{\sum_{j=1}^{s}p_{ij}^2}\right)^2}$$

$$C=\max_{i=1,...,s}\|\Delta A_i\|\left(2\|A_i\|+\|\Delta A_i\|\right)$$

and $\widetilde{H}_i$, $i\in S$ is the unique solution of (3.21) for $Q_i=I$, $i\in S$ and

$$K_1=\sum_{i=1}^{s}\left(1+\|A_i+\Delta A_i\|^2\right),$$

$$K_2=\sqrt{\sum_{i=1}^{s}\left(1+\|A_i+\Delta A_i\|^2\right)^2},\ K_\infty=\max_{i=1,...,s}\left(1+\|A_i+\Delta A_i\|^2\right).$$

For the standard Lyapunov equation many bounds for the solution are proposed in the literature. The surveys of such results can be found in [87], [67], [74], [30]. The reasons that the problem to estimate upper and lower bounds of these equations has become an attractive topic are that the bounds are also applied to solve many control problems such as stability analysis [77], [91], time-delay system controller design [86], estimation of the minimal cost and the suboptimal controller design [76], convergence of numerical algorithms [2], robust stabilization problem [20]. Eigenvalue bounds can be also used to determine whether or not the system under consideration possesses the singularly perturbed structure. An excellent motivation to study the bounds together with a survey of results for Lyapunov equation is given in [52] (Section 2.2). The authors advocated the results in this area by saying that sometimes we are just interested in the general behavior of the underlying system and then the behavior can be determined by examining certain bounds on the parameters of the solution instead of the full solution. The result of this kind for coupled Lyapunov equation is presented now. Here we assume that for a symmetric matrix X of the size $n \times n$ the eigenvalues $\lambda_k(X)$ are numerated in such a way that

$$\lambda_1(X) \geq \lambda_2(X) \geq \ldots \geq \lambda_n(X).$$

Theorem 29 *For the eigenvalues $\lambda_k(P_i)$, $k = 1, \ldots, n, i \in S$ of positive definite solution $P_i, i \in S$ of (3.21), the following inequalities hold*

$$\sum_{k=1}^{l} \lambda_k(P_i) \leq \sum_{k=1}^{l} \lambda_k(Q_i) +$$

$$\left(\max_{j\in S} p_{ij}\right) \lambda_1(A_i A_i') \frac{\sum_{i\in S}\sum_{k=1}^{l} \lambda_k(Q_i)}{1 - \max_{j\in S} \lambda_1\left(A_j A_j'\right) \max_{j\in S} \sum_{i\in S} p_{ij}} \tag{3.32}$$

for $l = 1, \ldots, n$, if $\max_{j\in S} \lambda_1\left(A_j A_j'\right) \max_{j\in S} \sum_{i\in S} p_{ij} < 1$, and

$$\sum_{k=1}^{l} \lambda_{n-k+1}(P_i) \geq \sum_{k=1}^{l} \lambda_{n-k+1}(Q_i) +$$

$$\left(\min_{j\in S} p_{ij}\right) \min_{j\in S} \lambda_n\left(A_j A_j'\right) \frac{\sum_{i\in S}\sum_{k=1}^{l} \lambda_{n-k+1}(Q_i)}{1 - \min_{j\in S} \lambda_n\left(A_j A_j'\right) \min_{j\in S} \sum_{i\in S} p_{ij}}, \tag{3.33}$$

for $l = 1, \ldots, n$, if $\min_{j\in S} \lambda_n\left(A_j A_j'\right) \min_{j\in S} \sum_{i\in S} p_{ij} < 1$.

In the proof we need the following result from [83].

Lemma 4 *Let $X, Y \in R^{n\times n}$ with $X = X'$, $Y = Y'$, $X, Y \geq 0$. Then the following inequalities hold*

$$\lambda_{i+j-1}(XY) \leq \lambda_i(X)\lambda_j(Y), \text{ if } i + j \leq n + 1 \tag{3.34}$$

$$\lambda_{i+j-n}(XY) \geq \lambda_i(X)\lambda_j(Y), \text{ if } i + j \geq n + 1 \tag{3.35}$$

$$\sum_{k=1}^{l} \lambda_k(X + Y) \leq \sum_{k=1}^{l} \lambda_k(X) + \sum_{k=1}^{l} \lambda_k(Y) \tag{3.36}$$

$$\sum_{k=1}^{l} \lambda_{n-k+1}(X + Y) \geq \sum_{k=1}^{l} \lambda_{n-k+1}(X) + \sum_{k=1}^{l} \lambda_{n-k+1}(Y). \tag{3.37}$$

Proof. Denote

$$F_i = \sum_{j\in S} p_{ij} P_j \tag{3.38}$$

From (3.21) it follows, by using (3.36) and (3.34), that

$$\sum_{k=1}^{l} \lambda_k(P_i) \leq$$

$$\sum_{k=1}^{l} \lambda_k(Q_i) + \sum_{k=1}^{l} \lambda_k(A_i' F_i A_i) = \sum_{k=1}^{l} \lambda_k(Q_i) + \sum_{k=1}^{l} \lambda_k(F_i A_i A_i')$$

$$\leq \sum_{k=1}^{l} \lambda_k(Q_i) + \lambda_1(A_i A_i') \sum_{k=1}^{l} \lambda_k(F). \tag{3.39}$$

Applying (3.36) to (3.38) implies

$$\sum_{k=1}^{l} \lambda_k(F_i) \leq \sum_{j\in S} \left(p_{ij} \sum_{k=1}^{l} \lambda_k(P_j)\right). \tag{3.40}$$

Combining (3.39) with (3.40) yields to

$$\sum_{k=1}^{l} \lambda_k(P_i) \leq$$

$$\sum_{k=1}^{l} \lambda_k(Q_i) + \max_{j\in S} \lambda_1\left(A_j A_j'\right) \sum_{j\in S} \left(p_{ij} \sum_{k=1}^{l} \lambda_k(P_j)\right). \tag{3.41}$$

Summing up the above inequality over $i \in S$ we have

$$\sum_{i\in S}\sum_{k=1}^{l} \lambda_k(P_i) \leq \sum_{i\in S}\sum_{k=1}^{l} \lambda_k(Q_i) +$$

$$\max_{j\in S} \lambda_1\left(A_j A_j'\right) \sum_{i\in S}\sum_{j\in S} \left(p_{ij} \sum_{k=1}^{l} \lambda_k(P_j)\right) =$$

$$\sum_{i\in S}\sum_{k=1}^{l} \lambda_k(Q_i) + \max_{j\in S} \lambda_1\left(A_j A_j'\right) \sum_{j\in S} \left(\left(\sum_{i\in S} p_{ij}\right) \sum_{k=1}^{l} \lambda_k(P_j)\right) \leq$$

$$\sum_{i\in S}\sum_{k=1}^{l} \lambda_k(Q_i) + \max_{j\in S} \lambda_1\left(A_j A_j'\right) \left(\max_{j\in S} \sum_{i\in S} p_{ij}\right) \sum_{i\in S}\sum_{k=1}^{l} \lambda_k(P_i).$$

Solving this inequality with respect to $\sum_{i\in S}\sum_{k=1}^{l}\lambda_k(P_i)$ and taking into account that

$$\max_{j\in S}\lambda_1\left(A_jA_j'\right)\max_{j\in S}\sum_{i\in S}p_{ij}<1$$

we obtain

$$\sum_{i\in S}\sum_{k=1}^{l}\lambda_k(P_i)\le\frac{\sum_{i\in S}\sum_{k=1}^{l}\lambda_k(Q_i)}{1-\max_{j\in S}\lambda_1\left(A_jA_j'\right)\max_{j\in S}\sum_{i\in S}p_{ij}}. \tag{3.42}$$

(3.39) implies also that

$$\begin{aligned}\sum_{k=1}^{l}\lambda_k(P_i)&\le\\ \sum_{k=1}^{l}\lambda_k(Q_i)&+\lambda_1(A_iA_i')\left(\sum_{j\in S}\left(p_{ij}\sum_{k=1}^{l}\lambda_k(P_j)\right)\right)\le\\ \sum_{k=1}^{l}\lambda_k(Q_i)&+\left(\max_{j\in S}p_{ij}\right)\lambda_1(A_iA_i')\sum_{j\in S}\sum_{k=1}^{l}\lambda_k(P_j).\end{aligned} \tag{3.43}$$

Applying (3.42) to the right hand side of (3.43) we have (3.32).

To proove (3.33) observe that the use of (3.27) and (3.35) to (3.21) gives

$$\begin{aligned}\sum_{k=1}^{l}\lambda_{n-k+1}(P_i)&\ge\sum_{k=1}^{l}\lambda_{n-k+1}(Q_i)+\sum_{k=1}^{l}\lambda_{n-k+1}(A_i'F_iA_i)=\\ &\sum_{k=1}^{l}\lambda_{n-k+1}(Q_i)+\sum_{k=1}^{l}\lambda_{n-k+1}(F_iA_iA_i')\ge\\ &\sum_{k=1}^{l}\lambda_{n-k+1}(Q_i)+\lambda_n(A_iA_i')\sum_{k=1}^{l}\lambda_{n-k+1}(F_i)\ge\\ &\sum_{k=1}^{l}\lambda_{n-k+1}(Q_i)+\min_{j\in S}\lambda_n\left(A_jA_j'\right)\sum_{k=1}^{l}\lambda_{n-k+1}(F_i).\end{aligned} \tag{3.44}$$

Summing (3.44) over $i\in S$ we have

$$\begin{aligned}&\sum_{i\in S}\sum_{k=1}^{l}\lambda_{n-k+1}(P_i)\ge\\ &\sum_{i\in S}\sum_{k=1}^{l}\lambda_{n-k+1}(Q_i)+\min_{j\in S}\lambda_n\left(A_jA_j'\right)\sum_{i\in S}\sum_{k=1}^{l}\lambda_{n-k+1}(F_i).\end{aligned} \tag{3.45}$$

Applying (3.27) to (3.38) leads to

$$\sum_{k=1}^{l}\lambda_{n-k+1}(F_i)\ge\sum_{j\in S}p_{ij}\sum_{k=1}^{l}\lambda_{n-k+1}(P_j). \tag{3.46}$$

Combining (3.45) with (3.46) yields to

$$\begin{aligned}\sum_{i\in S}\sum_{k=1}^{l}\lambda_{n-k+1}(P_i)&\ge\sum_{i\in S}\sum_{k=1}^{l}\lambda_{n-k+1}(Q_i)+\\ &\min_{j\in S}\lambda_n\left(A_jA_j'\right)\sum_{i\in S}\left(\sum_{j\in S}p_{ij}\sum_{k=1}^{l}\lambda_{n-k+1}(P_j)\right)=\\ &\sum_{i\in S}\sum_{k=1}^{l}\lambda_{n-k+1}(Q_i)+\\ &\min_{j\in S}\lambda_n\left(A_jA_j'\right)\sum_{j\in S}\left(\left(\sum_{i\in S}p_{ij}\right)\sum_{k=1}^{l}\lambda_{n-k+1}(P_j)\right)\ge\\ &\sum_{i\in S}\sum_{k=1}^{l}\lambda_{n-k+1}(Q_i)+\\ &\min_{j\in S}\lambda_n\left(A_jA_j'\right)\min_{j\in S}\sum_{i\in S}p_{ij}\sum_{i\in S}\sum_{k=1}^{l}\lambda_{n-k+1}(P_i).\end{aligned}$$

Solving this inequality with respect to $\sum_{i\in S}\sum_{k=1}^{l}\lambda_{n-k+1}(P_i)$ and taking into account that

$$\min_{j\in S}\lambda_n\left(A_jA_j'\right)\min_{j\in S}\sum_{i\in S}p_{ij}<1$$

we obtain

$$\sum_{i\in S}\sum_{k=1}^{l}\lambda_{n-k+1}(P_i)\ge\frac{\sum_{i\in S}\sum_{k=1}^{l}\lambda_{n-k+1}(Q_i)}{1-\min_{j\in S}\lambda_n\left(A_jA_j'\right)\min_{j\in S}\sum_{i\in S}p_{ij}}. \tag{3.47}$$

Combining (3.44) and (3.46) we conclude that

$$\begin{aligned}&\sum_{k=1}^{l}\lambda_{n-k+1}(P_i)\ge\sum_{k=1}^{l}\lambda_{n-k+1}(Q_i)+\\ &\left(\min_{j\in S}p_{ij}\right)\min_{j\in S}\lambda_n\left(A_jA_j'\right)\sum_{i\in S}\sum_{k=1}^{l}\lambda_{n-k+1}(P_i).\end{aligned}$$

Applying (3.47) to the right hand side of the above inequality we obtain (3.33). ■

3.4 Almost sure stability

It is well known that in practical applications, what is observed is sample path behavior rather than moment behavior, therefore almost sure stability is much more desirable than any moment stability. However, the analysis of almost sure stability is much more difficult than moment stability, this is why in the literature there are much more results for moment stability (especially second moment stability). Although moment stability, as we will see, implies almost sure stability, the stability criteria for high moments (say, second moment) as almost sure stability criteria are too conservative to be useful in practical applications. From our discussions in Section 2.2 about one dimensional systems, we know that almost sure stability is equivalent to stability of certain small moments. Therefore it is reasonable to use lower moment stability to study almost sure stability. In this section, we will devote our effort to the study of almost sure stability.

When relations between ASS and stability of matrices A_i are considered the first impression could be that if each matrix A_i is stable then system (3.1) is ASS. The next example shatters this expectation.

Example 9 *Consider system (3.1) with $n = 2$, $s = 2$,*

$$A_1 = \begin{bmatrix} \frac{1}{2} & 1 \\ 0 & \frac{1}{3} \end{bmatrix}, A_2 = \begin{bmatrix} \frac{4}{5} & 0 \\ 1 & \frac{1}{5} \end{bmatrix}.$$

Suppose that the Markov chain has the transition matrix of the form

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Then for initial distribution $\pi : P(r(0) = 1) = 1$ we have

$$x(2N, x_0, \pi) = \left(\begin{bmatrix} \frac{1}{2} & 1 \\ 0 & \frac{1}{3} \end{bmatrix} \begin{bmatrix} \frac{4}{5} & 0 \\ 1 & \frac{1}{5} \end{bmatrix} \right)^N x_0 =$$

$$\begin{bmatrix} \frac{7}{5} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{15} \end{bmatrix}^N x_0.$$

Because matrix

$$\begin{bmatrix} \frac{7}{5} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{15} \end{bmatrix}$$

has eigenvalue $\frac{11}{15} + \frac{1}{15}\sqrt{115} > 1$, therefore

$$\lim_{N \to \infty} \|x(2N, x_0, \pi)\| = \infty,$$

for $x_0 \neq 0$ and by Theorem 15 this system is not ASS, although both A_1 and A_2 are stable.

Quite opposite situation is also possible e.i. each matrices A_i are unstable but the system (3.1) is stable. This is demonstrated in the next example.

Example 10 *Consider system (3.1) with $n = 2$, $s = 2$,*

$$A_1 = \begin{bmatrix} 2 & 1 \\ 6 & 3 \end{bmatrix}, A_2 = \begin{bmatrix} 2 & 0 \\ -4 & 0 \end{bmatrix}.$$

Suppose that the Markov chain has the transition matrix of the form

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix},$$

where $p_{ij} > 0$ for all $i, j \in \{1, 2\}$. Observe that

$$A_1 A_2 = 0$$

Then for initial distribution $\pi : P(r(0) = 1) = \pi_1 > 0$, $P(r(0) = 2) = \pi_2 > 0$ we have

$$P_\pi (x(N + 1, x_0, \pi) = 0) = 1 \text{ for } N \geq \tau,$$

where

$$\tau = \min\{k : r(k + 1) = 2, r(k) = 1\}.$$

Since

$$P_\pi (\tau < \infty) = 1,$$

therefore

$$P_\pi \left(\lim_{N \to \infty} x(N, x_0, \pi) = 0 \right) = 1$$

that means that the system is π-ASS. By Theorem 16 it is ASS, even though both A_1 and A_2 are unstable.

To understand the possible difficulties arising in the analysis of ASS better consider the following example.

Example 11 *Consider system (3.1) with $n = 2$, $s = 2$,*

$$A(1) = \begin{bmatrix} 0 & a \\ b & 0 \end{bmatrix}, A(2) = \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix}.$$

Suppose that the Markov chain has the transition matrix of the form

$$P = \begin{bmatrix} 0 & 1 \\ p & 1 - p \end{bmatrix}.$$

Then

$$P^n = \begin{bmatrix} \frac{p+(-p)^n}{1+p} & -\frac{-1+(-p)^n}{1+p} \\ -p\frac{-1+(-p)^n}{1+p} & \frac{1+p(-p)^n}{1+p} \end{bmatrix}$$

and

$$P^n \underset{n\to\infty}{\longrightarrow} \begin{bmatrix} \frac{p}{1+p} & \frac{1}{1+p} \\ \frac{p}{1+p} & \frac{1}{1+p} \end{bmatrix}.$$

Consider the initial distribution $\pi_1 : P(r(0)=1)=1$. *Let* $0=\tau_1<\tau_2<\ldots$ *be the times* n *with* $r(n)=1$ *and*

$$\eta_k = \tau_{k+1} - \tau_k - 1, k = 1, 2 \ldots .$$

In this notation τ_k *is the moment when* $r(n)$ *visited state 1 for the k-th time and* η_k *is the time of being in state 2 after k-th visit in state 1. Then for each* $l \in N$ *there exists exactly one* $k(l)$ *such that* $\tau_{k(l)} \le l < \tau_{k(l)+1}$. *With this hint we have*

$$X(l) := \prod_{i=0}^{l} A(r(i)) = A(1)A^{\eta_1}(2)A(1)\ldots A^{\eta_{k-1}}(2)A(1)A^{l-\tau_{k(l)}}(2),$$

because, according to the structure of P *we know that after each visit in state 1 the chain goes to state 2. Observe that*

$$A(1)A^m(2) = \begin{bmatrix} 0 & a\beta^m \\ b\alpha^m & 0 \end{bmatrix}$$

and hence

$$X(l) = \begin{cases} \begin{bmatrix} 0 & d_1(l) \\ d_2(l) & 0 \end{bmatrix} & \text{for } k(l) = 2s(l)+1 \text{ for some } s \in N \\ \begin{bmatrix} d_3(l) & 0 \\ 0 & d_4(l) \end{bmatrix} & \text{for } k(l) = 2s(l) \quad \text{for some } s \in N \end{cases}$$

where

$$d_1(l) = a^{s(l)+1} b^{s(l)} \alpha^{\sum_{i=1}^{s(l)} \eta_{2i}} \beta^{\sum_{i=1}^{s(l)} \eta_{2i-1}} \beta^{l-\tau_{k(l)}}$$

$$d_2(l) = a^{s(l)} b^{s(l)} \alpha^{\sum_{i=1}^{s(l)-1} \eta_{2i}} \beta^{\sum_{i=1}^{s(l)} \eta_{2i-1}} \alpha^{l-\tau_{k(l)}}$$

$$d_3(l) = a^{s(l)} b^{s(l)} \alpha^{\sum_{i=1}^{s(l)-1} \eta_{2i}} \beta^{\sum_{i=1}^{s(l)} \eta_{2i-1}} \alpha^{l-\tau_{k(l)}}$$

$$d_4(l) = a^{s(l)} b^{s(l)} \alpha^{\sum_{i=1}^{s(l)} \eta_{2i-1}} \beta^{\sum_{i=1}^{s(l)-1} \eta_{2i}} \beta^{l-\tau_{k(l)}}$$

From Theorem 59 in the Apenndix with function f *given by* $f(1)=0$ *and* $f(2)=1$ *we obtain*

$$\frac{\sum_{i=1}^{k} \eta_i}{k} \underset{k\to\infty}{\overset{a.c}{\longrightarrow}} \frac{1}{1+p} \tag{3.48}$$

and therefore

$$\frac{\sum_{i=1}^{s(l)} \eta_{2i-1}}{l} \underset{l\to\infty}{\overset{a.c}{\longrightarrow}} \frac{1}{2(1+p)} \tag{3.49}$$

$$\frac{\sum_{i=1}^{s(l)-1} \eta_{2i}}{l} \underset{l\to\infty}{\overset{a.c}{\longrightarrow}} \frac{1}{2(1+p)}. \tag{3.50}$$

We also know ([50], Lemma 12, p.85)

$$\frac{k(l)}{l} \underset{l\to\infty}{\overset{a.c}{\longrightarrow}} \frac{p}{1+p}$$

and

$$\frac{\tau_{k(l)}}{l} \underset{l\to\infty}{\overset{a.c}{\longrightarrow}} 1,$$

therefore

$$\frac{s(l)}{l} \underset{l\to\infty}{\overset{a.c}{\longrightarrow}} \frac{p}{2(1+p)} \tag{3.51}$$

$$\frac{l-\tau_{k(l)}}{l} \underset{l\to\infty}{\overset{a.c}{\longrightarrow}} 0 \tag{3.52}$$

Using these we get

$$\frac{1}{l}\ln\left| a^{s(l)+1} b^{s(l)} \alpha^{\sum_{i=1}^{s(l)} \eta_{2i}} \beta^{\sum_{i=1}^{s(l)} \eta_{2i-1}} \beta^{l-\tau_{k(l)}} \right| \underset{l\to\infty}{\overset{a.c}{\longrightarrow}}$$

$$\frac{p}{2(1+p)} \ln|ab| + \frac{1}{2(1+p)} \ln|\alpha\beta|$$

and the same limit have all the other logarithms of the nonzero entries of $X(l)$. *So the sufficient condition for* π_1-ASS *is*

$$|ab|^{\frac{p}{2(1+p)}} |\alpha\beta|^{\frac{1}{2(1+p)}} < 1.$$

It is also clear that if

$$|ab|^{\frac{p}{2(1+p)}} |\alpha\beta|^{\frac{1}{2(1+p)}} > 1$$

then $\|X(l)\| \underset{l\to\infty}{\overset{a.c}{\longrightarrow}} \infty$ *and the system is not* π_1-ASS. *The case*

$$|ab|^{\frac{p}{2(1+p)}} |\alpha\beta|^{\frac{1}{2(1+p)}} = 1$$

requires further investigations.

Consider now the initial distribution π_2 *is* $P(r(0)=2)=1$.

$$X(l) := \prod_{i=0}^{l} A(r(i)) = A^{\eta_1}(2)A(1)A^{\eta_2}(2)A(1)\ldots,$$

because of the structure of P. Now we have

$$A^m(2)A(1) = \begin{bmatrix} 0 & \alpha^m a \\ \beta^m b & 0 \end{bmatrix}$$

and hence

$$X(l) = \begin{cases} \begin{bmatrix} 0 & \widetilde{d}_1(l) \\ \widetilde{d}_2(l) & 0 \end{bmatrix} & \textit{for } k(l) = 2s(l) \quad \textit{for some } s \in N \\ \begin{bmatrix} \widetilde{d}_3(l) & 0 \\ 0 & \widetilde{d}_4(l) \end{bmatrix} & \textit{for } k(l) = 2s(l)+1 \textit{ for some } s \in N \end{cases}$$

where

$$\widetilde{d}_1(l) = a^{s(l)+1}b^{s(l)}\alpha^{\sum_{i=1}^{s(l)}\eta_{2i-1}}\beta^{\sum_{i=1}^{s(l)}\eta_{2i}}\beta^{l-\tau_{k(l)}}$$

$$\widetilde{d}_2(l) = a^{s(l)}b^{s(l)+1}\alpha^{\sum_{i=1}^{s(l)-1}\eta_{2i}}\beta^{\sum_{i=1}^{s(l)}\eta_{2i-1}}\alpha^{l-\tau_{k(l)}}$$

$$\widetilde{d}_3(l) = a^{s(l)}b^{s(l)}\alpha^{\sum_{i=1}^{s(l)-1}\eta_{2i-1}}\beta^{\sum_{i=1}^{s(l)}\eta_{2i}}\alpha^{l-\tau_{k(l)}}$$

$$\widetilde{d}_4(l) = a^{s(l)}b^{s(l)}\alpha^{\sum_{i=1}^{s(l)}\eta_{2i}}\beta^{\sum_{i=1}^{s(l)-1}\eta_{2i-1}}\beta^{l-\tau_{k(l)}}$$

Using (3.48)-(3.52) we get

$$\frac{1}{l}\ln\left|a^{s(l)}b^{s(l)+1}\alpha^{\sum_{i=1}^{s(l)-1}\eta_{2i}}\beta^{\sum_{i=1}^{s(l)}\eta_{2i-1}}\alpha^{l-\tau_{k(l)}}\right| \underset{l\to\infty}{\overset{a.c}{\to}}$$

$$\frac{p}{2(1+p)}\ln|ab| + \frac{1}{2(1+p)}\ln|\alpha\beta|$$

and the same limit have all the other logarithms of the nonzero entries of $\widetilde{X}(l)$. So the sufficient conditions for π_2-ASS is the some as for π_1-ASS. We also know that if

$$|ab|^{\frac{p}{2(1+p)}}|\alpha\beta|^{\frac{1}{2(1+p)}} > 1$$

then the system is neither π_1-ASS nor π_2-ASS. The case

$$|ab|^{\frac{p}{2(1+p)}}|\alpha\beta|^{\frac{1}{2(1+p)}} = 1$$

requires further investigations. By Theorem 15 we conclude that the system is ASS if

$$|ab|^{\frac{p}{2(1+p)}}|\alpha\beta|^{\frac{1}{2(1+p)}} < 1.$$

Now we present several sufficient conditions for ASS. To present the first result let introduce the following notation. For a positive integer v we define the stochastic process $\widetilde{r}_v(k)$, taking values in S^v, as

$$\widetilde{r}_v(k) = (r(kv+v-1), ..., r(kv)),\ k = 0, 1, ...$$

Clearly $\widetilde{r}_v(k)$ is a Markov chain and for $\widetilde{i}_v, \widetilde{j}_v \in S^v$,

$$\widetilde{i}_v = (i_{v-1}, ..., i_0), \widetilde{j}_v = (j_{v-1}, ..., j_0),$$

the transition probability of $\widetilde{r}_v(k)$ are

$$\widetilde{p}(i_v, j_v) = P\left(\widetilde{r}_v(k+1) = \widetilde{j}_v \,\middle|\, \widetilde{r}_v(k) = \widetilde{i}_v\right) =$$

$$p(i_{v-1}, j_0)\, p(j_0, j_1) ... p(j_{v-2}, j_{v-1}).$$

Suppose that the Markov chain is ergodic then there exists a limit distribution:

$$\lim_{n\to\infty} p(n, i, j) = \pi_j$$

and therefore

$$\lim_{n\to\infty} \widetilde{p}(n, \widetilde{i}_v, \widetilde{j}_v) = \pi_{\widetilde{j}_v},$$

where

$$\pi_{\widetilde{j}_v} = \pi_{j_0} p(j_0, j_1) ... p(j_{v-2}, j_{v-1}).$$

Now for all $\widetilde{i}_v \in S^v$, $\widetilde{i}_v = (i_{v-1}, ..., i_0)$ we define

$$\widetilde{A}_{i_v} = A_{i_{v-1}} ... A_{i_0}.$$

Theorem 30 *[24]Suppose that the Markov chain is ergodic. If for some positive integer v and some matrix norm $\|*\|$ we have*

$$\prod_{\widetilde{i}_v \in S^v} \left\|\widetilde{A}_{\widetilde{i}_v}\right\|^{\pi_{\widetilde{i}_v}} < 1$$

then system (3.1) is ASS.

The next result is taken from [42].

Theorem 31 *If there exist positive definite matrices Q_i, $i \in S$ such that*

$$\sup_{\|x\|=1} \prod_{j=1}^{s} \left(\frac{x'A_i'Q_jA_ix}{x'Q_ix}\right)^{p_{ij}} < 1,$$

for all $i \in S$, then system (3.1) is ASS.

The last theorem has been generalized.. To present this generalization we introduce the following notations. For positive integer m, sequence $Q = (Q_1, ..., Q_s)$ of positive definite matrices of size $n \times n$, vector $x \in R^n$ and states $i, j_1, ..., j_m \in S$ denote

$$\mu(i, m, , j_1, ..., j_m, Q, x) =$$

$$\left(\frac{x' A_i' A_{j_1}' ... A_{j_{m-1}}' Q_{j_m} A_{j_{m-1}} ... A_{j_1} A_i x}{x' Q_i x} \right)^{p_{ij_1} ... p_{j_{m-1} j_m}}$$

and

$$\Gamma(i, m, Q, x) = \prod_{j_1=1}^{s} \prod_{j_2=1}^{s} ... \prod_{j_m=1}^{s} \mu(i, m, , j_1, ..., j_m, Q, x).$$

The following generalization of Theorem 31 has been proved in [79].

Theorem 32 *If there exist positive definite matrices Q_i, $i \in S$ and positive integer m such that*

$$\sup_{\|x\|=1} \Gamma(i, m, Q, x) < 1,$$

for all $i \in S$, then system (3.1) is ASS.

The fact that Theorem 32 is a generalization of Theorem 31 is demonstrated in the next example.

Example 12 *Consider system (3.1) with $n = 2$, $s = 2$,*

$$A(1) = \begin{bmatrix} \sqrt{2} & 0 \\ 0 & a \end{bmatrix}, A(2) = \begin{bmatrix} a & 0 \\ 0 & \sqrt{2} \end{bmatrix},$$

with $a > 0$. Suppose that the Markov chain has the transition matrix of the form

$$P = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}.$$

Let try to use Theorem 31 with $Q_1 = Q_2 = I$. Then

$$\sup_{\|x\|=1} \prod_{j=1}^{s} \left(\frac{x' A_1' A_1 x}{x' x} \right)^{p_{1j}} = \sup_{\|x\|=1} \frac{x' \begin{bmatrix} 2 & 0 \\ 0 & a^2 \end{bmatrix} x}{x' x} \geq$$

$$\frac{\begin{bmatrix} 1 \\ 0 \end{bmatrix}' \begin{bmatrix} 2 & 0 \\ 0 & a^2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}}{\begin{bmatrix} 1 \\ 0 \end{bmatrix}' \begin{bmatrix} 1 \\ 0 \end{bmatrix}} = 2$$

and we can not say that the system is stable. When we use Theorem 32 for $m = 2$ with the same Q_1 and Q_2 we have

$$\sup_{\|x\|=1} \prod_{j=1}^{2} \prod_{l=1}^{2} \left(\frac{x' A_1' A_l' A_l A_1 x}{x' x} \right)^{p_{1l} p_{lj}} = 2\sqrt{2} a$$

and

$$\sup_{\|x\|=1} \prod_{j=1}^{2} \prod_{l=1}^{2} \left(\frac{x' A_2' A_l' A_l A_2 x}{x' x} \right)^{p_{2l} p_{lj}} = 2\sqrt{2} a.$$

Therefore this system is stable for

$$a < \frac{\sqrt{2}}{4}. \tag{3.53}$$

Notice also that in this case Theorem 32 gives also necessary conditions for almost sure stability. It is not difficult to observe that stability of the system is equivalent to stability of one dimensional system

$$x(k+1) = a(r(k)) x(k)$$

with $a(1) = a$ and $a(2) = \sqrt{2}$. Because in our case the Markov chain is ergodic with limit distribution

$$\pi : \pi_1 = \pi_2 = 0.5,$$

therefore by Theorem 21 it is stable if and only if

$$1 > |a(1)|^{\pi_1} |a(2)|^{\pi_2} = (a)^{0.5} \left(\sqrt{2}\right)^{0.5}$$

and this is possible only if (3.53) holds.

Another sufficient conditions for ASS may be obtained from our consideration about one dimensional systems. It is given in the next theorem.

Theorem 33 *Suppose that the Markov chain is ergodic with limit distribution*

$$(\pi_1, ..., \pi_s)$$

and suppose that for certain matrix norm $\|*\|$ with the property $\|XY\| \leq \|X\| \|Y\|$ we have

$$\|A_1\|^{\pi_1} \|A_2\|^{\pi_2} ... \|A_s\|^{\pi_s} < 1.$$

Then system (3.1) is ASS.

Proof. By the property of the norm we have

$$\|A(r(N))A(r(N-1))...A(r(0))\| \le \|A_1\|^{\tau_1(N)} \|A_2\|^{\tau_2(N)} ... \|A_s\|^{\tau_s(N)},$$

where

$$\tau_i(k) = \sum_{l=0}^{k} \mathbf{1}_{\{r(l)=i\}}.$$

Therefore

$$\lim_{N\to\infty} \frac{1}{N} \ln \|A(r(N))A(r(N-1))...A(r(0))\| \le$$

$$\lim_{N\to\infty} \frac{1}{N} \ln \|A_1\|^{\tau_1(N)} \|A_2\|^{\tau_2(N)} ... \|A_s\|^{\tau_s(N)} =$$

$$\sum_{i=1}^{s} \pi_i \ln \|A_i\|.$$

In the last step we use Theorem 59. By the assumption the right hand side is negative and therefore

$$\|A(r(N))A(r(N-1))...A(r(0))\| \underset{N\to\infty}{\rightarrow} 0, \text{a.s.}$$

this implies ASS of (3.1). ■

We can also use the consideration about one dimensional systems to obtain necessary and sufficient conditions for ASS of (3.1) in case when matrices $A(i)$, $i \in S$ can be simultaneously diagonalized. This is done in the next theorem.

Theorem 34 *If matrices $A(i)$, $i \in S$ can be simultaneously diagonalized to the following form*

$$T^{-1}A(i)T = \begin{bmatrix} \lambda_1(i) & & 0 \\ & \ddots & \\ 0 & & \lambda_n(i) \end{bmatrix}$$

then

1. *system (3.1) is δ−MS if and only if*

$$\rho\left(P'D_\delta^{(v)}\right) < 1, \; v = 1, ..., n$$

where

$$D_\delta^{(v)} = diag\left[|\lambda_v(i)|^\delta\right]_{i=1,...,s}.$$

2. *if in addition $r(k)$ is ergodic with limit distribution $\pi = (\pi_i)_{i\in S}$ then system (3.1) is ASS if and only if it is δ−MS for all $\delta < \delta_0$ for certain $\delta_0 > 0$ and this is equivalent to the following*

$$|\lambda_v(i)|^{\pi_1} ... |\lambda_v(i)|^{\pi_s} < 1, \; v = 1, ..., n.$$

Proof. Under the assumption of the Theorem δ−MS (ASS) of (3.1) is equivalent to δ−MS (ASS) of each scalar systems

$$x(k+1) = \lambda_v(r(k))x(k), \; v = 1, ..., n.$$

Therefore the conclusion follows from Theorem 18, Theorem 21, and Theorem 22. ■

We have obtained this result as a simple consequence of our one dimensional consideration. It can be generalized, however it requires deeper analysis. It is natural to expect that if the matrices pairwise commute then the stability problem should be easier. If the matrices $A(i)$, $i \in S$ can be simultaneously diagonalized then they pairwise commute, therefore the last theorem includes the case of commuting matrices. However if $A(i)$, $i \in S$ pairwise commute then it does not guarantee simultaneous transformation to a diagonal form, but then according to [56], there exists a unitary matrix T such that they can be transformed by this similarity transformation to the upper triangular forms. The next theorem deals with such a situation.

Theorem 35 *Suppose that the Markov chain $r(k)$ is ergodic with limit distribution $\pi = (\pi_i)_{i\in S}$ and that the matrices $A(i)$, $i \in S$ can be simultaneously transformed by a similarity transformation T to upper triangular form*

$$T^{-1}A(i)T = \begin{bmatrix} \lambda_i(1) & * & \cdots & * \\ 0 & \lambda_i(2) & \cdots & * \\ \vdots & & \ddots & \\ 0 & \cdots & 0 & \lambda_i(n) \end{bmatrix}, \; i \in S. \tag{3.54}$$

Then, a necessary and sufficient condition for almost sure stability of system (3.1) is

$$|\lambda_1(l)|^{\pi_1} |\lambda_2(l)|^{\pi_2} ... |\lambda_s(l)|^{\pi_s} < 1, \tag{3.55}$$

for all $l = 1, ..., n$.

For the first time this theorem appears in [43] as a necessary sufficient of ASS in case of $r(k)$ being a sequence of independent and identically distributed discrete random variables. In this paper the necessity of this conditions for $n = 2$ has also been shown. Next in [41] the necessity of this conditions for $n > 2$ has also been shown in the case of a sequence of independent and identically distributed discrete random variables. It appears that this proof works also with slight modifications for ergodic Markov chain. This is done below.

Proof. We start with the proof of necessity of condition (3.55). If the system (3.1) is ASS, then all elements of matrix

$$A\left((r(k)\right) A\left((r(k-1)\right) \ldots A\left((r(0)\right)$$

tend to zero almost surely when k tends to infinity. In particular all diagonal elements do so. Therefore ASS of (3.1) implies ASS of each of scalar systems

$$y_i(k+1) = \lambda_i(r(k)) y_i(k), \ i \in S.$$

This in light of Theorem 21 implies (3.55).

Now we show the sufficiency of (3.55). Without loss of generality, we can assume that all $A(i)$, $i \in S$ are in upper diagonal form (3.54). Let b be the upper bound of the absolute values of the off-diagonal elements of $A(i)$, $i \in S$. Consider the following jump linear system

$$\overline{x}(k+1) = \overline{A}(r(k))\overline{x}(k), \tag{3.56}$$

where

$$\overline{A}(i) = \begin{bmatrix} |\lambda_i(1)| & b & \cdots & b \\ 0 & |\lambda_i(2)| & \cdots & b \\ \vdots & & \ddots & \\ 0 & \cdots & 0 & |\lambda_i(n)| \end{bmatrix}.$$

It is straightforward to verify that ASS of (3.56) implies ASS of (3.1). We show that (3.55) implies ASS of (3.56). Let

$$G(k) = \overline{A}\left((r(k)\right) \overline{A}\left((r(k-1)\right) \ldots \overline{A}\left((r(0)\right) =$$

$$\begin{bmatrix} g_{11}(k) & g_{12}(k) & \cdots & g_{1n}(k) \\ 0 & g_{22}(k) & & g_{2n}(k) \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & g_{nn}(k) \end{bmatrix}.$$

The ASS of (3.56) is equivalent to the fact that all elements of $G(k)$ tends almost surely to zero when k tends to infinity. To proceed, we use induction on the n. If $n = 1$, $G(k)$ tends almost sure to zero by Theorem 21. Suppose that it is true for $n-1$. Because of the triangular structure of $G(k)$ and

the induction hypotesis, it is sufficient to show that for n the elements of the last column of $G(k)$ are converging to zero almost surely. From $G(k) = \overline{A}\left((r(k)\right) G(k-1)$ we obtain recursive formulas for the elements of the last column

$$g_{mn}(k) = \left|\lambda_{r(k)}(m)\right| g_{mn}(k-1) + b \sum_{l=m}^{n} g_{l+1,n}(k-1), \ l = 1, 2, , n-1 \tag{3.57}$$

$$g_{nn}(k) = \left|\lambda_{r(k)}(n)\right| g_{nn}(k-1)$$

Again by Theorem 21 assumption (3.55) implies that $g_{nn}(k)$ tends almost surely to zero at an exponential rate. Therefore, there exists $0 < \rho_n < 1$ and random variable M_n such that for all k

$$g_{nn}(k) \leq M_n \rho_n^k.$$

Now, we use an induction argument on the index m of $g_{mn}(k)$ to show that for each $1 \leq m \leq n$, there exists a $0 < \rho_m < 1$ and random variable M_m which is a polynomial in the variable k such that for all k

$$g_{mn}(k) \leq M_m \rho_m^k. \tag{3.58}$$

From this, we can conclude that $g_{mn}(k)$ tends to zero almost surely for all $1 \leq m \leq n$. We proceed as follows: Suppose that for some $1 < m < n$, there exists $M_m, M_{m+1}, \ldots, M_n$ and $\rho_m, \rho_{m+1}, \ldots, \rho_n < 1$, such that (3.58) holds. We show that (3.58) holds for $m-1$. From (3.57), we have

$$\begin{aligned} g_{m-1n}(k) = \left|\lambda_{r(k)}(m-1)\right| g_{m-1n}(k-1) + b \sum_{l=m-1}^{n} g_{l+1,n}(k-1) = \\ \left|\lambda_{r(k)}(m-1)\right| \left|\lambda_{r(k-1)}(m-1)\right| \ldots \left|\lambda_{r(1)}(m-1)\right| g_{m-1n}(0) + \\ b \sum_{l=1}^{k-1} \left|\lambda_{r(k)}(m-1)\right| \left|\lambda_{r(k-1)}(m-1)\right| \ldots \\ \left|\lambda_{r(l+1)}(m-1)\right| \left(g_{mn}(l-1) + \ldots g_{nn}(l-1)\right) + \\ b\left(g_{mn}(k-1) + \ldots g_{nn}(k-1)\right). \end{aligned} \tag{3.59}$$

We see that the last term almost surely converges to zero at an exponential rate because of the induction hypotesis. Actually, let

$$\overline{M}_{m-1} = \max_{m \leq l \leq n} M_l \text{ and } \overline{\rho}_{m-1} = \max_{m \leq l \leq n} \rho_l.$$

We have

$$b\left(g_{mn}(k-1)+...g_{nn}(k-1)\right) \leq nb\overline{M}_{m-1}\left(\overline{\rho}_{m-1}\right)^{k}. \tag{3.60}$$

By Theorem 21 we also know that the first term in (3.59) tends to zero almost surely at an exponential rate so there exists $\widetilde{M}_{m-1}>0$ and $0<\widetilde{\rho}_{m-1}<1$ such that

$$\left|\lambda_{r(k)}(m-1)\right|\left|\lambda_{r(k-1)}(m-1)\right|...\left|\lambda_{r(1)}(m-1)\right| g_{m-1n}(0) \leq$$

$$\widetilde{M}_{m-1}\left(\widetilde{\rho}_{m-1}\right)^{k}. \tag{3.61}$$

Now we consider the second term in (3.59). Again by Theorem 21 there exists $\widehat{M}_{m-1}>0$ and $0<\widehat{\rho}_{m-1}<1$ such that

$$\left|\lambda_{r(k)}(m-1)\right|\left|\lambda_{r(k-1)}(m-1)\right|...\left|\lambda_{r(l+1)}(m-1)\right| \leq \widehat{M}_{m-1}\left(\widehat{\rho}_{m-1}\right)^{k-l}.$$

Let

$$\rho_{m-1}=\max\left\{\overline{\rho}_{m-1},\widetilde{\rho}_{m-1},\widehat{\rho}_{m-1}\right\}$$

and

$$M'_{m-1}=\max\left\{\overline{M}_{m-1},\widetilde{M}_{m-1},\widehat{M}_{m-1}\right\}.$$

Then, it follows that the second term in (3.59) satisfies the inequality

$$b\sum_{l=1}^{k-1}\left|\lambda_{r(k)}(m-1)\right|\left|\lambda_{r(k-1)}(m-1)\right|...$$

$$\left|\lambda_{r(l+1)}(m-1)\right|\left(g_{mn}(l-1)+...g_{nn}(l-1)\right) \leq$$

$$b\sum_{l=1}^{k-1}\widehat{M}_{m-1}\left(\widehat{\rho}_{m-1}\right)^{k-l}\left(M_m\rho_m^l+...M_n\rho_n^l\right) \leq$$

$$nb\left(M'_m\right)^2 k\left(\rho_{m-1}\right)^{k}. \tag{3.62}$$

Combining (3.60), (3.61) and (3.62), we conclude that there exists $0<\rho_m<1$ and random variable M_{m-1} which is a polynomial in the variable k such that for all k

$$g_{m-1n}(k) \leq M_{m-1}\rho_{m-1}^{k}.$$

This completes the proof. ■

From this theorem we see that under the assumption on the simultaneously upper triangularisation stability of each matrices $A(i)$, $i\in S$ implies

ASS of the system. We have already noticed (see Example 9) that in general it is not true.

Now we will use the Lyapunov exponent method to study ASS of (3.1). The main idea of this method is to examine formulae

$$\lim_{k\to\infty}\frac{1}{k}\ln\left\|A\left((r(k)\right)A\left((r(k-1)\right)...A\left((r(0)\right)\right\|$$

or

$$\lim_{k\to\infty}\frac{1}{k}\ln\left\|x\left(k,x_0,\pi\right)\right\|.$$

The first question is if the limits exist, next how they depend on the initial distribution of the Markov chain, the choice of matrix norm, and (in the second case) on the initial condition x_0. Another problem is that even though all matrices $A(i)$, $i\in S$ are nonzero the product $A\left((r(k)\right)A\left((r(k-1)\right)...A\left((r(0)\right)$ may be a zero matrix (see Example 10). To deal with the last problem we extend $\ln(*)$ as follows

$$\ln(x)=\begin{cases}\ln(x) \text{ for } x>0\\ -\infty \text{ for } x=0\\ \infty \text{ for } x=\infty.\end{cases}$$

The next theorem collects main results about Lyapunov exponents presented in [48], [90], and [51].

Theorem 36 *Suppose that all matrices $A(i)$, $i\in S$ are non singular, and the Markov chain $r(k)$ is ergodic. Then the limits*

$$\lim_{k\to\infty}\frac{1}{k}\ln\left\|A\left((r(k)\right)A\left((r(k-1)\right)...A\left((r(0)\right)\right\| \tag{3.63}$$

and

$$\lim_{k\to\infty}\frac{1}{k}E_\pi\ln\left\|A\left((r(k)\right)A\left((r(k-1)\right)...A\left((r(0)\right)\right\| \tag{3.64}$$

exist and are finite constants, they do not depend on the matrix norm and initial distribution π and they are equal. Moreover if the limit is negative then (3.1) is ASS and if it is greater then 1 then (3.1) is not ASS. In addition there exists a proper subspace L of R^n such that for all $x_0\in R^n\backslash L$ the limit

$$\lim_{k\to\infty}\frac{1}{k}\ln\left\|x\left(k,x_0,\pi\right)\right\| \tag{3.65}$$

exists and does not depend on the initial distribution π as well as on the vector norm and is equal to (3.63) and

$$\lim_{k\to\infty}\frac{1}{k}\ln\left\|A\left((r(k)\right)A\left((r(k-1)\right)...A\left((r(0)\right)\right\|=$$

$$\max_{x_0 \neq 0} \lim_{k \to \infty} \frac{1}{k} \ln \|x(k, x_0, \pi)\| .$$

Definition 5 *The common value of (3.63), (3.64), and (3.65) is called the largest sample path Lyapunov exponent of system (3.1).*

Theorem 36, even though very interesting from theoretical point of view, is useless in practical development of stability unless we can calculate the value of the largest Lyapunov exponent or even its sign. Some results in this direction are presented in the next section in the context of relations between the Lyapunov exponent just introduced and $\delta-$moment Lyapunov exponent. There is one more problem with this approach to almost sure stability. The above theorem leaves as unsolved the case of zero Lyapunov exponent. It is worth to notice that our consideration about one dimensional systems is based on the idea of Lyapunov exponent. However in this case we were able to present explicit formulas for the largest Lyapunov exponent as well as we were able to deal with the case of zero Lyapunov exponent.

3.5 $\delta-$moment stability

This section is devoted to study of the concept of $\delta-$MS of system (3.1). We start with the following property of $\delta-$MS.

Theorem 37 *The following conditions are equivalent to $\delta-$MS of system (3.1).*

1. *there exist constant $M > 0$ and $0 < \rho < 1$ such that*

$$E_\pi \|x(k, x_0, \pi)\|^\delta \leq M\rho^k \tag{3.66}$$

2. *the series $\sum_{k=0}^{\infty} E_\pi \|x(k, x_0, \pi)\|^\delta$ converges for all $x_0 \in R^n$ and all initial distributions π.*

Proof. It is clear that (3.66) implies that $\sum_{k=0}^{\infty} E_\pi \|x(k, x_0, \pi)\|^\delta$ converges and the convergence in turn implies $\delta-$MS of system (3.1). Suppose now that system (3.1) is $\delta-$MS. Let π_i, $i \in S$ be the initial distribution of the Markov chain which is defined by the i-th row of the matrix P and let π be any initial distribution. From the $\delta-$MS, we have

$$\lim_{k \to \infty} \max_{1 \leq i \leq s} E_{\pi_i} \|A((r(k)) A((r(k-1)) \ldots A((r(0))\|^\delta = 0.$$

Then for any $0 < \eta < 1$ given, there exists an integer m such that

$$\max_{1 \leq i \leq s} E_{\pi_i} \|A((r(m-1)) A((r(m-2)) \ldots A((r(0))\|^\delta \leq \eta.$$

By the time homogeneous property, there exists an $M > 0$ such that for any $0 \leq q < m$ and k, we have

$$\max_{1 \leq i \leq s} E_{\pi_i} \|A((r(k+q)) A((r(k-1)) \ldots A((r(0))\|^\delta \leq M.$$

Let $k = pm + q$, where $0 \leq q < m$, then we obtain, using the time homogeneous property again,

$$\begin{aligned}
& E_\pi \|A((r(k)) A((r(k-1)) \ldots A((r(0))\|^\delta \leq \\
& E_\pi^\delta \|A((r(pm+q)) \ldots A((r(pm))\|^\delta \ldots \\
& \|A((r(pm-1)) \ldots A((r((p-1)m))\|^\delta \times \\
& \|A((r(m-1)) A((r(k-1)) \ldots A((r(0))\|^\delta = \\
& \sum_{i_0, \ldots, i_{pm+q}} p_{i_0} p_{i_0 i_1} \cdots p_{i_{pm+q-1} i_{pm+q}} \|A(i_{pm+q}) \ldots A(i_{pm})\|^\delta \times \\
& \left\|A(i_{pm-1}) \ldots A\left(i_{(p-1)m}\right)\right\|^\delta \times \ldots \|A(i_{m-1}) \ldots A(i_0)\|^\delta = \\
& \sum_{i_0, \ldots, i_{pm-1}} p_{i_0} p_{i_0 i_1} \cdots p_{i_{pm-2} i_{pm-1}} \left\|A(i_{pm-1}) \ldots A\left(i_{(p-1)m}\right)\right\|^\delta \times \ldots \\
& \|A(i_{m-1}) \ldots A(i_0)\|^\delta \times \\
& \left(\sum_{i_{pm}, \ldots, i_{pm+q}} p_{i_{pm-1}, i_{pm}} \cdots p_{i_{pm+q-1} i_{pm+q1}} \|A(i_{pm+q}) \ldots A(i_{pm})\|^\delta \right) = \\
& \sum_{i_0, \ldots, i_{pm-1}} p_{i_0} p_{i_0 i_1} \cdots p_{i_{pm-2} i_{pm-1}} \left\|A(i_{pm-1}) \ldots A\left(i_{(p-1)m}\right)\right\|^\delta \times \ldots \\
& \|A(i_{m-1}) \ldots A(i_0)\|^\delta \times \\
& E_{\pi_{i_{pm-1}}} \|A(r(q)) \ldots A(r(0))\|^\delta \leq \\
& \max_{1 \leq i \leq s} E_{\pi_i} \|A((r(q)) \ldots A((r(0))\|^\delta \times \\
& \sum_{i_0, \ldots, i_{pm-1}} p_{i_0} p_{i_0 i_1} \cdots p_{i_{pm-2} i_{pm-1}} \left\|A(i_{pm-1}) \ldots A\left(i_{(p-1)m}\right)\right\|^\delta \times \ldots \\
& \|A(i_{m-1}) \ldots A(i_0)\|^\delta .
\end{aligned}$$

We can repeat this procedure p times to obtain

$$E_\pi \|A((r(k))A((r(k-1))...A((r(0))\|^\delta \leq M\eta^p \leq M_1 r_1^k,$$

where $M_1 = Mr^{-\frac{q}{m}}$ and $r_1 = r^{\frac{1}{m}}$. This shows that δ–MS of system (3.1) implies (3.66). ■

The theorem enables us to show the following relationship between δ–MS and ASS of system (3.1).

Theorem 38 *If system (3.1) is δ–MS then it is ASS.*

Proof. Let $\xi = \limsup_{k\to\infty} \|x(k,x_0,\pi)\|$, then from Markov inequality, we have that for any $c > 0$, the following holds

$$P(\xi > c) \leq P\left(\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty}\{\|x(k,x_0,\pi)\| > c\}\right) \leq$$

$$P\left(\bigcup_{m=n}^{\infty}\{\|x(k,x_0,\pi)\| > c\}\right) \leq \sum_{m=n}^{\infty} P(\|x(k,x_0,\pi)\| > c) \leq$$

$$\frac{1}{c^\delta}\sum_{m=n}^{\infty} E_\pi \|x(m,x_0,\pi)\|^\delta \overset{n\to\infty}{\Longrightarrow} 0$$

The last step is true because of point 2 of Theorem 37. Thus $P(\xi \geq c) = 0$ and consequently $P(\xi = 0) = 1$. ■

Further relations between δ–MS and ASS are given in terms of δ- moment Lyapunov exponents. Before we present the formal definitions of the δ- moment Lyapunov exponents we need the following analogue of Theorem 36.

Theorem 39 *Suppose that all matrices $A(i)$, $i \in S$ are non singular, and the Markov chain $r(k)$ is ergodic. Then the limit*

$$g(\delta) = \lim_{k\to\infty} \frac{1}{k} \ln E_\pi \|A((r(k))A((r(k-1))...A((r(0))\|^\delta \tag{3.67}$$

exists and is a finite constant, it does not depend on the matrix norm and initial distribution π. Moreover if the limit is negative then (3.1) is δ–MS and if it is positive then (3.1) is not δ–MS. In addition there exists a proper subspace L of R^n such that for all $x_0 \in R^n \backslash L$ the limit

$$\lim_{k\to\infty} \frac{1}{k} \ln E\|x(k,x_0,\pi)\|^\delta \tag{3.68}$$

exists and does not depend on the initial distribution π as well as on the vector norm and is equal to (3.67) and

$$\lim_{k\to\infty} \frac{1}{k} \ln E_\pi \|A((r(k))A((r(k-1))...A((r(0))\|^\delta =$$

$$\max_{x_0\neq 0} \lim_{k\to\infty} \frac{1}{k} \ln E_\pi \|x(k,x_0,\pi)\|^\delta.$$

The proof can be found in [39]. This theorem, similarly as Theorem 36, is very interesting from theoretical point of view, however is useless in practical development of stability unless we can calculate the value of the largest Lyapunov exponent or even its sign. The important difference between this theorem and Theorem 36 is that now the case of zero Lyapunov exponent is not left as unsolved. This is because of Theorem 37.

Definition 6 *The common value $g(\delta)$ of (3.67) and (3.68) is called the largest δ–moment Lyapunov exponent of system (3.1).*

Now we present a connection between δ–moment Lyapunov exponent and sample path Lyapunov exponent.

Theorem 40 *Suppose that all matrices $A(i)$, $i \in S$ are non singular, and the Markov chain $r(k)$ is ergodic. Then the function $g(\delta)$ is differentiable from the right at $\delta = 0$ and $g'(0+)$ is largest sample path Lyapunov exponent of system (3.1).*

The proof of this theorem can be found in [44].

We end this section with certain sufficient condition of δ–MS which is a straightforward consequence of Theorem 18.

Theorem 41 *System (3.1) is δ–MS if there exists a matrix norm $\|*\|$ such that $\|XY\| \leq \|X\|\|Y\|$ for all matrices X and Y and*

$$\rho(P'D_\delta) < 1,$$

where

$$D_\delta = diag\left[\|A(i)\|^\delta\right]_{i=1,...,s}.$$

3.6 Comparison and discussion

In this paragraph we give a summary of our results on stability of jump linear systems. We have discussed three types of stability: ASS, $\delta-$MS and mean stability. On the base of Example 7 we can say that idea of mean stability is rather useless from practical point of view. The relations between ASS and $\delta-$MS are not explained completely. We know from Theorem 38 that $\delta-$MS implies ASS. Moreover by Theorem 40 if all matrices $A(i)$, $i \in S$ are non singular, and the Markov chain $r(k)$ is ergodic and the largest sample path Lyapunov exponent is negative then ASS implies $\delta-$MS for small δ. However we do not know, except of the one dimensional case, if ASS implies that the largest sample path Lyapunov exponent is negative. We belive that it is the case. Unfortunately we do not have testable necessary and sufficient conditions neither for $\delta-$MS nor for ASS. There are two exceptions: one dimensional system and mean square stability ($\delta-$MS for $\delta = 2$). For one dimensional systems the easy to check conditions for ASS are given by Theorem 21 and for $\delta-$MS by Theorem 18 and Theorem 19. The necessary and sufficient conditions for mean square stability are presented in Theorem 23, Theorem 24. In the special case when all matrices can be simultaneously transformed to upper triangular form we can also present necessary and sufficient conditions for $\delta-$MS and for ASS, this is done in Theorem 35. Under more restrictive assumption that all matrices can be simultaneously transformed to diagonal form we are able to explain completely the relations between $\delta-$MS and ASS as in Theorem 34. Interesting approach of Lyapunov exponent to develop ASS and $\delta-$MS is discussed in Theorems 36 and 39. However these results are useless in practical development of stability unless we can calculate the value of the largest Lyapunov exponent or at least its sign.

Chapter 4

The jump linear quadratic regulator

The objective of this chapter is to study one of the optimization problems for jump linear systems, namely optimization of quadratic cost functional. Such problem is called JLQ problem (J for jump, L for linear dynamics, Q for quadratic cost). It is well know that the solution of standard LQ problem has a model solution. The optimal control has a feedback form and the feedback matrix can be easily found by solving Riccati difference or Riccati algebraic equations. The conditions for existence of a solution of these problems are also well known and they are given in the form of easily checkable conditions of controllability, stabilizability, observability or detectability of suitable subsystems. One can expect similar results for JLQ.

Historically, the first papers about JLQ problems are [70], [97] and [101]. These papers dealt with continuous time systems under diferent assumptions. One of the first solutions to the discrete time JLQ was presented in [13] for finite time interval and in [23] for infinite time interval.. Further discussion of this problem may be found in [1], [25], [58], [53]. In all these papers time invariant version of JLQ problem has been considered. The first papers studing time-varying JLQ problem are probably [31] and [33].

When we compare the results on the standard LQ with the ones on JLQ we can find many similarities. The solutions has the feedback form, the feedback matrices are defined in terms of a certain type of quadratic matrix equation which is called coupled Riccati equation. The sufficient conditions for existence of the solution are given in terms of properly defined stabilizability and detectability. The closed loop system is stable under certain conditions though stability is understood in the mean square sense. The problem is that from the practical point of view we are interested in almost sure stability and as we know from our previous considerations mean square stability implies almost sure stability however opposite statement is not true. It means that the JLQ technique can be too conservative to find a regulator which ensures almost sure stability. This is also a reason why we should look not only for sufficient conditions for existence of JLQ problem but also necessary. Such a new condition is proposed in Definition 7 and Corollary 3. In the next section we present different formulations of JLQ problems. In Section 2 the simplest JLQ problem, namely JLQ problem on finite time interval is considered. Next, in Section 3, we consider noise free JLQ problem on infinite time interval. It is shown that a solution of this problem exists if and only if certain coupled difference Riccati equation has a global and bounded nonnegative definite solution and this in turns is equivalent to optimalizablity (see Definition 7). This result is a significant extension of previous ones where only sufficient conditions for existence of the solution are presented and only in time invariant case. Because of the significancy of the coupled Riccati equation in JLQ theory we study its properties in Section 4. In this section a very important question about stability of the optimal closed loop systems is addressed also (see Theorem 45). In the case of time invariant systems the role of the coupled difference Riccati equation is played by a coupled algebraic Riccati equation. Properties of this equation are studied in Section 5. Section 6 is devoted to JLQ problem for systems with additive disturbances. Basing on the nonuniquness of the optimal control in this case it is shown that for time varying systems with coefficients tending to certain limits the optimal control may be realized in the form of time invariant feedback. For the standard systems results of this kind has been discovered in [26]. A discussion of presented results is given in the last section.

4.1 Problem formulation

The system under study in this chapter is described by the following state equation:

$$x(k+1) = A_k(r(k))x(k) + B_k(r(k))u(k) + C_k(r(k))w(k), \qquad (4.1)$$

where, as previously, the state $x(k) \in R^n$, the control $u(k) \in R^m$. $w(k)$, $k = 0, 1...$, is a second order independent identically distributed sequence of $n-$dimensional random variables with $Ew(k) = 0$ and $Ew(k)w'(k) = I$. They represent disturbances. We also assume that processes $r(k)$, $k = 0, 1, ...$ and $w(k)$, $k = 0, 1, ...$ are independent. The sequences of matrices of appropriate sizes $A_k(i)$, $B_k(i)$ and $C_k(i)$ are bounded for each $i \in S$. In the case of $C_k(i) = 0$ for all $i \in S$ and $k = 0, 1...$ the system (4.1) is referred as a noise free system

and in the opposite case as a noise system. To introduce the cost criteria consider sequences $Q_k(i) \in R^{n\times n}$ and $R_k(i) \in R^{m\times m}$, $i \in S$, $k = 0, 1, ...$ of symmetric nonnegative and symmetric positive definite matrices, respectively and symmetric nonnegative definite matrices $K(i) \in R^{n\times n}, i \in S$. We assume that sequences $Q_k(i)$, $R_k(i)$ and $R_k^{-1}(i)$ are bounded for each $i \in S$. In this chapter we consider only the initial distribution π of the following form $P(r(0) = i_0) = 1$ for certain $i_0 \in S$. The optimal control problems we consider in this chapter are defined as follows.

Problem 1: JLQ problem on finite time interval. For given N find a control sequence $u = (u(0), ..., u(N-1))$ such that the cost functional

$$J(x_0, \pi, u, N) = \langle K(r(N))x(N), x(N)\rangle +$$

$$E\left[\sum_{k=0}^{N-1} \langle Q_k(r(k))\, x(k), x(k)\rangle + \langle R_k(r(k))\, u(k), u(k)\rangle\right], \qquad (4.2)$$

takes the minimum value.

Problem 2: Noise free JLQ problem on infinite time interval. For the noise free system (4.1) find a control sequence $u = (u(0), u(1), ...)$ such that cost functional

$$J_{nf}(x_0, \pi, u) =$$

$$\lim_{N\to\infty} E\left[\sum_{k=0}^{N} \langle Q_k(r(k))\, x(k), x(k)\rangle + \langle R_k(r(k))\, u(k), u(k)\rangle\right], \qquad (4.3)$$

takes the minimal value.

Problem 3: Noise JLQ problem on infinite time interval. For the noise system (4.1) find a control sequence $u = (u(0), u(1), ...)$ such that cost functional

$$J(x_0, \pi, u) =$$

$$\lim_{N\to\infty} \frac{1}{N} E\left[\sum_{k=0}^{N} \langle Q_k(r(k))\, x(k), x(k)\rangle + \langle R_k(r(k))\, u(k), u(k)\rangle\right], \qquad (4.4)$$

takes the minimal value.

We also consider the time invariant cases when the coefficient of (4.1) and (4.2)-(4.4) do not depend on k.

It is clear that to solve any of the Problems 1-3 it is enough to solve the appropriate problem with π being Dirac distribution, therefore in our further considerations only such problems are discussed.

4.2 JLQ problem on finite time interval

The solution of the JLQ problem on finite time interval is given by the following theorem. The proof for the noise free system can be found in [23] and for the noise system in [59].

Theorem 42 *The optimal control law for the JLQ problem on finite time interval is given by*

$$\widetilde{u}(k) = -L_k(r(k))x(k), i \in S, k = 0, ..., N, \tag{4.5}$$

where for each $i \in S, k = 0, ..., N$ the optimal gain is defined as

$$L_k(i) = \left(R_k(i) + B_k'(i)F_{k+1}(i)B_k(i)\right)^{-1} B_k'(i)F_{k+1}(i)A_k(i), \tag{4.6}$$

where

$$F_{k+1}(i) = \sum_{j\in S} p_{ij}P_{k+1}(j) \tag{4.7}$$

and the sequence of sets of symmetric nonnegative definite matrices

$$\{P_k(i) : i \in S\}, k = 0, ..., N-1$$

satisfies the equation

$$P_k(i) = A_k'(i)F_{k+1}(i)\left(A_k(i) - B_k(i)L_k(i)\right) + Q_k(i), k = 0, ..., N-1, i \in S \tag{4.8}$$

with terminal conditions $P_N(i) = K(i), i \in S$.

The value of the optimal cost is given by

$$\langle P_0(i_0)x_0, x_0\rangle + q_0(i_0) \tag{4.9}$$

where $q_k(i), k = 0, ..., N$ satisfies the recurrent formula:

$$q_k(i) = \sum_{j\in S} q_{k+1}(j)p_{ij} + \sum_{j\in S} tr\left(C_k'(i)P_{k+1}(j)C_k(i)\right) \tag{4.10}$$

with terminal conditions $q_N(i) = 0, i \in S$.

For convenience we introduce the following notation: The right hand side of (4.8) is denoted by $f(k, P_*, i)$ i.e.

$$f(k, P_*, i) = A_k'(i)F_{k+1}(i)\left(A_k(i) - B_k(i)L_k(i)\right) + Q_k(i).$$

The notation P_* in $f(k, P_*, i)$ means that f depends on all matrices $P_j, j \in S$. Moreover, let

$$P_k^{(N)}(i, K(*)) = P_{N-k}(i), k = 1, ..., N, i \in S \tag{4.11}$$

where $P_k(i)$ is given by (4.8). Again the notation $K(*)$ means that $P_k^{(N)}(i, K(*))$ depends on all matrices $K(j), j \in S$. Using this notation we can rewrite equation (4.8) as

$$P_k^{(N)}(i, K(*)) = f(N-k, P_{k-1}^{(N)}(*, K(*), i), k = 0, ..., N \tag{4.12}$$

or equivalently

$$\overline{P}_l^{(N)}(i, K(*)) = f(l, \overline{P}_{l+1}^{(N)}(*, K(*), i), l = N, ..., 0 \tag{4.13}$$

with

$$P_k^{(N)}(i, K(*)) = \overline{P}_{N-k}^{(N)}(i, K(*))$$

and initial (terminal) conditions

$$P_0^{(N)}(i, K(*)) = K(i), \ (\overline{P}_N^{(N)}(i, K(*)) = K(i)), i \in S$$

Moreover, the minimal value of the cost functional (4.2) is equal to

$$\left\langle P_N^{(N)}(r_0, K(*))\, x_0, x_0\right\rangle + \widetilde{q}_N(r_0), \tag{4.14}$$

where $\widetilde{q}_k(i), k = 0, ..., N$ is given by

$$\widetilde{q}_k(i) = \sum_{j\in S} \widetilde{q}_{k-1}(j)p_{ij} + \sum_{j\in S} tr\left(C_{N-k}'(i)P_{k-1}^{(N)}(j, K(*))C_{N-k}(i)\right) \tag{4.15}$$

with terminal conditions $\widetilde{q}_0(i) = 0$. From the proof of Theorem 42 in [23] the following formula in the noise free case can be deduced:

$$\begin{aligned} J(x_0, i_0, u, N) = {} & \left\langle P_N^{(N)}(i_0, K(*))\, x_0, x_0\right\rangle - \\ & E\left(\langle K(r(N))x(N), x(N)\rangle\right) + \\ E\left[\sum_{k=0}^{N-1} \langle R_k(r(k))(u(k) - L_k(r(k))x(k)), (u(k) - L_k(r(k))x(k))\rangle\right] & \end{aligned} \tag{4.16}$$

and

$$E\left[\sum_{k=l}^{N-1} \langle Q_k(r(k))\, x(k), x(k)\rangle + \langle R_k(r(k))\, \widetilde{u}(k), \widetilde{u}(k)\rangle\right]$$

$$r(l) = i, x(l) = x_l] =$$

$$\left\langle P_{N-l}^{(N)}(i, K(*))x_l, x_l \right\rangle - E \left\langle K(r(N))x(N), x(N) \right\rangle, \tag{4.17}$$

where $\widetilde{u}$ is given by (4.5), $L_k(i)$ by (4.6) and u is any admissible control. Introduce also the following terminology. If there exists a sequence of matrices $P(k,i)$, $k = 0, 1, ..., i \in S$ such that

$$P(k,i) = A_k'(i)F_{k+1}(i)\left(A_k(i) - B_k(i)L_k(i)\right) + Q_k(i) \tag{4.18}$$

where

$$L_k(i) = \left(R_k(i) + B_k'(i)F_{k+1}(i)B_k(i)\right)^{-1} B_k'(i)F_{k+1}(i)A_k(i),$$

and

$$F_k(i) = \sum_{j \in S} p_{ij}P(k,j)$$

then equation (4.18) is called a coupled difference Riccati equation and the sequence $P(k,i)$, $k = 0, 1, ..., i \in S$ will be called its global solution. Moreover, if this sequence is bounded then we call it a global and bounded solution. If a solution has the property $P(k,i) \leq \widetilde{P}(k,i)$, $(P(k,i) \geq \widetilde{P}(k,i))$ $k = 0, 1, ...,$ $i \in S$ for any other solution $\widetilde{P}(k,i)$ then it is called a minimal (maximal) solution.

4.3 Noise free JLQ problem on infinite time interval

The primary concern of this section is to establish sufficient and necessary conditions for the existence of optimal control for the time-varying noise free JLQ problem on infinite time interval. For this purpose we introduce the following definition.

Definition 7 *The noise free system (4.1) with cost functional (4.3) is called optimizable if, for all* $(i_0, x_0) \in S \times R^n$ *there exists control* u *such that* $J_{nf}(x_0, i_0, u) < \infty$. *In such situation we say that*

$$(A_k(r(k)), B_k(r(k)), Q_k(r(k)), R_k(r(k)), r(k))$$

is optimizable.

The concept of optimizability is an extenstion of concept of stabilizability in the sense that stabilizability implies optimizability and oposite implication is not true. It will also appear that optimizability is a necessery and sufficient condtion for existence of the solution of time-varying noise free JLQ problem on infinite time interval

Theorem 43 *If noise free system (4.1) with cost functional (4.3) is optimizable then the limits*

$$\lim_{N \to \infty} \overline{P}_k^{(N)}(i, 0) = P(k,i) \tag{4.19}$$

exist for all $k = 0, 1, ...$ *and* $i \in S$. $P(k,i)$ *is a global and bounded solution of (4.18), and* $P(k,i)$ *is symmetric and nonnegative, and* $P(k,i)$ *is the minimal nonnegative definite global and bounded solution of (4.18). The optimal control is given by*

$$\widetilde{u}(k) = -L_k(r(k))x(k), i \in S, k = 0, ..., N, \tag{4.20}$$

where

$$F_k(i) = \sum_{j \in S} p_{ij}P(k,j)$$

$$L_k(i) = \left(R_k(i) + B_k'(i)F_{k+1}(i)B_k(i)\right)^{-1} B_k'(i)F_{k+1}(i)A_k(i) \tag{4.21}$$

and $J_{nf}(x_0, i_0, \widetilde{u}) = \langle P(0, i_0)x_0, x_0 \rangle$. *On the other hand, if there exists nonnegative definite global and bounded solution of (4.18) then the noise free system (4.1) with cost functional (4.3) is optimizable.*

Proof. For $0 \leq N_1 \leq N_2$ fix $(i_0, x_0) \in \times S \times R^n$ and consider the cost functionals $J(x_0, i_0, u, N_1)$ and $J(x_0, i_0, u, N_2)$ both with $K(i) = 0$, $i \in S$. Then it follows easily from the form of the cost functional that $J(x_0, i_0, u, N_1) \leq J(x_0, i_0, u, N_2)$ and then from (4.17) we conclude that

$$\left\langle \overline{P}_k^{(N_1)}(i, 0)x_0, x_0 \right\rangle \leq \left\langle \overline{P}_k^{(N_2)}(i, 0)x_0, x_0 \right\rangle, \tag{4.22}$$

for $k = 0, ..., N_1$. By (4.14) and the optimalizability conditions we conclude that there is a constant $c > 0$ such that

$$\left\| \overline{P}_k^{(N)}(i, 0) \right\| < c \tag{4.23}$$

for all N, $k = 0, ...N$, and $i \in S$, N. From (4.22) and (4.23) we conclude that the limit $P(k,i)$ in (4.19) indeed exists and $P(k,i) = P'(k,i)$, $P(k,i) \geq 0$. Moreover, since the constant in (4.23) does not depend on k and N, $P(\cdot, i)$ is bounded. Taking the limit in both sides of (4.13) we see that $P(\cdot, i)$ satisfies this equation so $P(k,i)$ is a bounded and global solution of coupled difference Riccati equation (4.18).

To show that $P(k,i)$ is the minimal nonnegative global and bounded solution of (4.18) we introduce another nonnegative global and bounded solution of $L(k,i)$, and consider $\widetilde{P}_k^{(N)}(i, L(N,*))$ the solution of (4.12) with initial conditions

$$P_0^{(N)}(i, L(N,*)) = L(N,i),\ i \in S.$$

Since the solution is unique and, since $L(\cdot, i)$ satisfies (4.12), we have

$$\widetilde{P}_k^{(N)}(i, L(N,*)) = L(k,i),\ i \in S$$

and $k = 0,...N$. Furthermore, it follows that if we have two solutions $P_k^{(N)}(i, K_1(*))$ and $P_k^{(N)}(i, K_2(*))$ of (4.12) and if $K_1(i) \le K_2(i)$ then

$$P_k^{(N)}(i, K_1(*)) \le P_k^{(N)}(i, K_2(*)),\ i \in S.$$

This property implies that $P(k,i) \le L(k,i)$. So $P(k,i)$ is the minimal nonnegative global and bounded solution of (4.18).

To solve the optimal control problem fix $(i_0, x_0) \in S \times R^n$ and consider the cost functional $J^{(1)}(x_0, i_0, u, N)$ with $K(i) = 0,\ i \in S$ and $J^{(2)}(x_0, i_0, u, N)$ with $K(i) = P(N,i), i \in S$. Then apply control (4.20) and use the fact that

$$J^{(1)}(x_0, i_0, u, N) \le J^{(2)}(x_0, i_0, u, N)$$

and that $\widetilde{u}$ is optimal for $J^{(2)}(x_0, i_0, u, N)$. We see that

$$J^{(1)}(x_0, i_0, \widetilde{u}, N) \le J^{(2)}(x_0, i_0, \widetilde{u}, N) = \langle P(N, i_0)\, x_0, x_0 \rangle,$$

The right hand side is independent of N, so

$$J(x_0, i_0, \widetilde{u}) = \lim_{N\to\infty} J^{(1)}(x_0, i_0, \widetilde{u}, N) \le \langle P(N, i_0)\, x_0, x_0 \rangle. \tag{4.24}$$

On the other hand we have

$$J(x_0, i_0, u) = \lim_{N\to\infty} J^{(1)}(x_0, i_0, u, N) \ge$$

$$\lim_{N\to\infty} \left\langle P_0^{(N)}(i_0, 0)\, x_0, x_0 \right\rangle = \langle P(0, i_0)\, x_0, x_0 \rangle. \tag{4.25}$$

(4.24) and (4.25) show the optimality of $\widetilde{u}$.

Now suppose that there exists nonnegative definite global and bounded solution $P(k,i)$ of (4.18). Fix $(i_0, x_0) \in S \times R^n$ and consider again the cost functional $J^{(1)}(x_0, i_0, u, N)$ and $J^{(2)}(x_0, i_0, u, N)$. Then applying the control (4.20) we conclude from (4.24) that $J(x_0, i_0, \widetilde{u}) < \infty$. It means that the system is optimizable. ■

The following corollary is a straightforward consequence of this theorem.

Corollary 3 *The solution of Problem 2 exists if and only if the system is optimizable.*

In the time invariant case Theorem 43 reduces to the following result.

Theorem 44 *Consider the noise free system (4.1) with cost functional (4.3) under assumption that sequences $A_k(i)$, $B_k(i)$, $Q_k(i)$, and $R_k(i)$ are constant for each $i \in S$ and equal to $A(i)$, $B(i)$, $Q(i)$, and $R(i)$, respectively. This system is optimizable if and only if there exists a nonnegative definite solution of the coupled algebraic Riccati equation*

$$P(i) = A'(i)F(i)\left(A(i) - B(i)L(i)\right) + Q(i), \tag{4.26}$$

where

$$L(i) = \left(R(i) + B'(i)F(i)B(i)\right)^{-1} B'(i)F(i)A(i). \tag{4.27}$$

$$F(i) = \sum_{j\in S} p_{ij}P(j). \tag{4.28}$$

If such a solution exists then there exists a minimal solution $P_{\min}(i)$ and the optimal control is given by

$$\widetilde{u}(k) = -L(r(k))x(k), i \in S, k = 0, ..., N, \tag{4.29}$$

where

$$L(i) = \left(R(i) + B'(i)F(i)B(i)\right)^{-1} B'(i)F(i)A(i). \tag{4.30}$$

$$F(i) = \sum_{j\in S} p_{ij}P_{\min}(j). \tag{4.31}$$

and

$$J_{nf}(x_0, i_0, \widetilde{u}) = \langle P(i_0)\, x_0, x_0 \rangle. \tag{4.32}$$

4.4 Coupled difference Riccati equation

In this chapter we discuss several properties of the coupled difference Riccati equation. We start with the following extension of definition of stochastic stability.

Definition 8 *The system*

$$x(k+1) = A_k(r(k))x(k), \tag{4.33}$$

is called exponentially stochastically stable if $E\|x(k)\|^2 \leq ab^k$ *for certain constants* $a > 0$ *and* $0 < b < 1$ *and all initial conditions* $x_0 \in R^n$ *and all initial distributions* π. *It is called stochastically stable if*

$$\sum_{k=0}^{\infty} E\|x(k)\|^2 < \infty$$

In such situations $(A_k(r(k)), r(k))$ *is called exponentially stochastically stable and stochastically stable respectively.*

Definition 9 *The noise free system (4.1) is called exponentially stochastically stabilizable if there exists a feedback control*

$$u(k) = -L_k(r(k))x(k)$$

such that $L_k(i)$ *is bounded for each* $i \in S$ *and for the resulting closed loop system*

$$x(k+1) = (A_k(r(k)) - B_k(r(k))L_k(r(k)))\, x(k)$$

is exponentially stochastically stable. In such situation we will call $-L_k(r_k)$ *the exponentially stochastically stabilizable feedback and the triplet*

$$(A_k(r(k)), B_k(r(k)), r(k))$$

exponentially stochastically stabilizable.

It is clear that exponential stochastic stabilizability implies optimalizability, so from Theorem 43 it is a sufficient condition for the existence of a nonnegative definite global and bounded solution of coupled difference Riccati equation (4.18). However this assumption alone does not guarantee that the solution is unique. Moreover the optimal control given by (4.20) may lead to stochastically unstable system. To present a condition for uniqueness of the solution and for exponential stochastic stabilizablity of the optimal closed loop system we introduce the following definition.

Definition 10 *Consider the following system*

$$\begin{aligned} x(k+1) &= A_k(r(k))x(k), \\ y(k) &= C_k(r(k))x(k). \end{aligned}$$

It is called exponentially stochastically detectable if

$$(A_k'(r(k)), C_k'(r(k)), r(k))$$

is exponentially stochastically stabilizable, and the triplet

$$(A_k(r(k)), C_k(r(k)), r(k))$$

is called exponentially stochastically detectable in this case.

Now we are ready to formulate the main result about existence and uniqueness of the global and bounded nonnegative definite solution of coupled difference Riccati equation (4.18).

Theorem 45 *Suppose that*

$$(A_k(r(k)), B_k(r(k)), r(k))$$

is exponentially stochastically stabilizable and that

$$\left(A_k(r(k)), \sqrt{Q_k(r(k))}, r(k)\right) \tag{4.34}$$

is exponentially stochastically detectable. Then the coupled difference Riccati equation (4.18) has a unique global and bounded nonnegative definite solution. Moreover the optimal feedback gain for the noise free system (4.1) with cost functional (4.3) given by (4.20) is stochastically stabilizable feedback.

Before we present the proof notice that if $Q_k(i) = C_k'(i)C_k(i)$ for certain matrices $C_k(i)$, then the concept of detectability may be used to justify asumption (4.34).

In the proof we need the following technical Lemma. It may be justified in the same way as Proposition 3 in [54].

Lemma 5 *Suppose that* $(A_k(r(k)), r(k))$ *is exponentially stochastically stable and that the n-dimensional random variables* $f(k)$, $k = 0, 1, \ldots$ *are such that* $E\sum_{k=0}^{\infty}\|f(k)\|^2 = c < \infty$, *then for*

$$z(k+1) = A_k(r(k))z(k) + f(k), \;\; z(0) = z_0$$

we have

$$E\sum_{k=0}^{\infty}\|z(k)\|^2 \leq \alpha E\|z_0\|^2 + \beta,$$

for certain constants α *and* β.

We are now ready to prove Theorem 45

Proof. As we have already noticed stabilizability implies optimizablity, and by Theorem 43 we know that there exists nonnegative definite global and bounded solution $P(k,i)$ of (4.18) which is the minimal solution. We first show that under the control given by (4.20) the closed-loop system is exponentially stochastically stable. From detectability assumption there exists a bounded matrix sequence $M_k(i)$ such that

$$\left(A_k(r(k)) + M_k(r(k))\sqrt{Q_k(r(k))}, r(k)\right)$$

is exponentially stochastically stable. The solution $\widetilde{x}$ of (4.1) which corresponds to the control given by (4.20) satisfies

$$\widetilde{x}(k+1)=\left(A_k\left(r\left(k\right)\right)+M_k\left(r(k)\right)\sqrt{Q_k(r(k))}\right)\widetilde{x}\left(k\right)+f(k) \qquad (4.35)$$

where

$$f(k)=B_k\left(r(k)\right)\widetilde{u}(k)-M_k\left(r(k)\right)\sqrt{Q_k(r(k))}\widetilde{x}\left(t\right)$$

It follows that

$$\|f(k)\|^2\leq\|B_k\left(r(k)\right)\|^2\|\widetilde{u}(k)\|^2+\|M_k\left(r(k)\right)\|^2\left\|\sqrt{Q_k(r(k))}\widetilde{x}\left(k\right)\right\|^2\leq$$

$$\frac{\mu}{\sigma}\left\langle R_k\left(r(k)\right)\widetilde{u}(k),\widetilde{u}(k)\right\rangle+\mu\left\langle\sqrt{Q_k(r(k))}\widetilde{x}(t),\sqrt{Q_k(r(k))}\widetilde{x}(t)\right\rangle\leq$$

$$\delta\left(\left\langle R_k\left(r(k)\right)\widetilde{u}(k),\widetilde{u}(k)\right\rangle+\left\langle\sqrt{Q_k(r(k))}\widetilde{x}(t),\sqrt{Q_k(r(k))}\widetilde{x}(t)\right\rangle\right) \qquad (4.36)$$

where $\delta=\max\left(\frac{\mu}{\sigma},\mu\right)$, $\|B_k\left(i\right)\|<\mu$, $\|M_k\left(i\right)\|<\mu$, and $\sigma I<R_k\left(i\right)$. From (4.36) and by Theorem 43 we have

$$E\sum_{k=0}^{\infty}\|f(k)\|^2\leq\delta J_{nf}\left(x_0,i_0,\widetilde{u}\right)=\delta\left\langle P\left(0,i_0\right)x_0,x_0\right\rangle\leq\delta\nu\left\|x_0\right\|^2,$$

where $0\leq P\left(k,i\right)\leq\nu I$. Applying Lemma 5 to (4.35) leads to

$$E\sum_{k=0}^{\infty}\|\widetilde{x}(k)\|^2<\infty.$$

It means that the closed loop system is stochastically stable.

Next we show that under the detectability condition the stabilizing solution $P\left(k,i\right)$ is the maximal solution of (4.18). Fix $(x_0,i_0)\in R^n\times S$ and denote by U_{stab} the subset of all admissible control sequences consisting of such control u, that the corresponding solution x of (4.1) satisfies

$$E\sum_{k=0}^{\infty}\|\widetilde{x}(k)\|^2<\infty.$$

Let $\widetilde{P}\left(k,i\right)$ be any nonnegative definite global and bounded solution of (4.18). From (4.16), we have

$$J\left(x_0,i_0,u,N\right)=\left\langle\widetilde{P}\left(0,i_0\right)x_0,x_0\right\rangle-E\left(\left\langle\widetilde{P}\left(N,i\right)x(N),x(N)\right\rangle\right)+$$

$$E\left[\sum_{k=0}^{N-1}\left\langle R_k\left(r(k)\right)\left(u(k)-L_k(r(k))x(k)\right),\left(u(k)-L_k(r(k))x(k)\right)\right\rangle\right] \qquad (4.37)$$

Hence for $u\in U_{stab}$ it follows that

$$J(x_0,i_0,u)=\lim_{N\to\infty}J\left(x_0,i_0,u,N\right)=\left\langle\widetilde{P}\left(0,i_0\right)x_0,x_0\right\rangle+$$

$$E\left[\sum_{k=0}^{\infty}\left\langle R_k\left(r(k)\right)\left(u(k)-L_k(r(k))x(k)\right),\left(u(k)-L_k(r(k))x(k)\right)\right\rangle\right]$$

Thus

$$J(x_0,i_0,u)\geq\left\langle\widetilde{P}\left(0,i_0\right)x_0,x_0\right\rangle. \qquad (4.38)$$

Now consider $\widetilde{u}(t)$ given by (4.20). Substituting it into (4.37) gives

$$J(x_0,i_0,\widetilde{u})=\left\langle P\left(0,i_0\right)x_0,x_0\right\rangle. \qquad (4.39)$$

Combining (4.37) with (4.38) implies $\widetilde{P}\left(0,i_0\right)\leq P\left(0,i_0\right)$. As we have noticed in the proof of Theorem 43 it implies $\widetilde{P}\left(k,i_0\right)\leq P\left(k,i_0\right)$ for all k. Since the solution $P\left(\cdot,i\right)$ is simultaneously maximal and minimal, it is unique. ■

We demonstrated the use of this theorem on an example.

Example 13 *Consider the noise free system (4.1) $n=1$, $s=2$,*

$$A_k(1)=a_k,A_k(2)=0.$$

$$a_k=\sqrt{\frac{(6k+13)\,(2k+3)}{(k+1)\,(4k+9)}}$$

$$B_k(1)=B_k(2)=1$$

$$Q_k(1)=Q_k(2)=R_k(1)=R_k(2)=1$$

Suppose that the Markov chain has the transition matrix of the form

$$P=\begin{bmatrix}\frac{1}{2}&\frac{1}{2}\\1&0\end{bmatrix}.$$

Taking

$$L_k(1)=a_k$$

and

$$L_k(2)=0$$

in Definition 9 we see that the system is exponentially stochastically stabilizable and detectable. Therefore by Theorem 45, coupled Riccati equation (4.18) which takes form

$$P(k,1)=\frac{1}{2}a_k^2\left(P(k+1,1)+1\right)-\frac{1}{4}a_k^2\frac{\left(P(k+1,1)+1\right)^2}{\frac{3}{2}+\frac{1}{2}P(k+1,1)}+1$$

$$P(k,2) = 1,$$

has a unique global and bounded positive solution. It is matter of straightforward calculation to check that

$$P(k,1) = 3 + \frac{1}{k+1},$$

$$P(k,2) = 1,$$

satisfies this equation. Further the optimal feedback gain takes the following form

$$L_k(1) = \frac{a_k(4k+9)}{6k+13}$$

$$L_k(2) = 1.$$

In the proof of Theorem 43 we have shown (see (4.23)) that under optimalizability condition the solution $P_k^{(N)}(i,0)$ of coupled difference Riccati equation with zero initial conditions satisfies the following inequality

$$\left\| P_k^{(N)}(i,0) \right\| < c$$

for all N, $k = 0, ...N$, and $i \in S$, N and certain constant c. It appears that the result may be extended on the case of arbitrary initial conditions and this is done in the next Lemma.

Lemma 6 *If the system $(A_k(r(k)), B_k(r(k)), r(k))$ is stochastically stabilizable, then there exists a constant $c(\{K(i) : i \in S\})$ such that*

$$\left\| P_k^{(N)}(i, K(*)) \right\| < c(\{K(i) : i \in S\}) \tag{4.40}$$

for any N and $k = 1, ..., N$, where $P_k^{(N)}(i, K(*))$ is given by (4.12).

Proof. Consider the noise free system (4.1) with cost functional (4.3). Let $\tilde{u}$ be such that the closed loop system is stable and put

$$V_k^{(N)}(x,i) =$$

$$E\left[\sum_{\nu=k}^{N-1} \langle Q_\nu(r(\nu))x(\nu), x(\nu)\rangle + \langle R_\nu(r(\nu))u(\nu), u(\nu)\rangle \middle| x(\nu) = x, r(\nu) = i \right]$$

Then

$$\left\langle P_k^{(N)}(i_0, K(*))x_0, x_0 \right\rangle$$

$$= V_k^{(N)}(x,i) \le J_{nf}(x_0, r_0, \tilde{u}) + E\left(\langle K(r_0)x_N, x_N\rangle \mid r(0) = i_0, x(0) = x_0\right)$$

and consequently $\left\| P_k^{(N)}(i, K(*)) \right\| < c(\{K(i) : i \in S\})$. ■

Using this result for the time invariant system we can generaliz property (4.19) as follows (see [31] for details).

Theorem 46 *If the system $\{A(i), B(i), i \in S\}$ is stochastically stabilizable and the system $\left\{A(i), \sqrt{Q(i)}, i \in S\right\}$ is stochastically detectable, then for any initial values $\{P_0(i) : i \in S\}$ we have*

$$\lim_{k\to\infty} P_k(i, P_0(*)) = P(i),$$

where $P_k(i, P_0(i)), i \in S$ are given by time invariant coupled difference Riccati equation

$$P_k(i, P_0(*)) = A'(i)F_{k-1}(i)\left[A(i) - B(i)L_k(i)\right] + Q(i), \tag{4.41}$$

$$P_0(i, P_0(*)) = P_0(i)$$

where

$$F_{k-1}(i) = \sum_{j\in S} p_{ij}P_{k-1}(j, P_0(*))$$

$$L_k(i) = \left(R(i) + B'(i)F_{k-1}(i)B(i)\right)^{-1} B'(i)F_{k-1}(i)A(i) \tag{4.42}$$

for $k = 1, ..., N$. and $\{P(i) : i \in S\}$ is unique nonnegative solution of the coupled algebraic Riccati equation (4.26) and the convergence is uniform on the set

$$\{P_0(i) : \|P_0(i)\| < c, i \in S\}$$

for any positive constant $c > 0$.

The next theorem, which describes the asymptotic behavior of the coupled difference Riccati equation in the case when coefficients have limits, is the key result in the proof of the fact that the optimal control for the time varying noise system on infinite time interval can be realized, in the time invariant feedback form in some cases.

Theorem 47 *Assume that the sequence $(A_N(j), B_N(j), Q_N(j), R_N(j))$,*

$$A_N(j) \in R^{n\times n}, B_N(j) \in R^{n\times m}, C_N(j) \in R^{n\times n},$$

$$Q_N(j) \in R^{n\times n},\ R_N(j) \in R^{m\times m},\ Q_N(j) \geq 0,\ R_N(j) > 0,\ j \in S,$$

is such that the limits of $A_N(j)$, $B_N(j)$, $Q_N(j)$, $R_N(j)$ when N tends to infinity, exist for each $j \in S$ and $R(j) > 0$, $(A(i), B(i), i \in S)$ is stochastically stabilizable, $\left(A(i), \sqrt{Q(i)}, i \in S\right)$ is stochastically detectable, where

$$\begin{aligned} A(j) &= \lim_{N\to\infty} A_N(j),\ B(j) = \lim_{N\to\infty} B_N(j), &(4.43)\\ Q(j) &= \lim_{N\to\infty} Q_N(j), R(j) = \lim_{N\to\infty} R_N(j),\ j \in S, &(4.44)\end{aligned}$$

then

$$\lim_{N\to\infty} \frac{1}{N}\sum_{k=0}^{N} P_k^{(N)}(i, K(*)) = P(i), \tag{4.45}$$

for any initial condition $\{K(i) : K(i) \geq 0, i \in S\}$, where $P(i)$ is the unique solution of coupled algebraic Riccati equation (4.26).

Proof. For simplicity we denote

$$L(A,B,F,R,Q) = A'FA - A'FB\left(R + B'FB\right)^{-1} B'FA + Q.$$

Using this notation we can write

$$P_k^{(N)}(i, K(*)) = L(A_{N-k}(i), B_{N-k}(i), F_{k-1}(i), R_{N-k}(i), Q_{N-k}(i)), \tag{4.46}$$

where

$$F_{k-1}^{(N)}(i) = \sum_{j\in S} p_{ij} P_{k-1}^{(N)}(j, K(*)). \tag{4.47}$$

Together with (4.46) we consider the following time invariant difference Riccati equation

$$P_k(i, K(*)) = L(A(i), B(i), F_{k-1}(i), R(i), Q(i)), \tag{4.48}$$

where

$$F_{k-1}(i) = \sum_{j\in S} p_{ij} P_{k-1}(j, K(*)) \tag{4.49}$$

and

$$P_0(j, K(*)) = K(j).$$

Stochastic stabilizability of $\{A(i), B(i), i \in S, k \in N\}$ implies exponential stochastic stabilizability of $\{A_k(i), B_k(i), i \in S, k \in N\}$ (see Lemma 8 in [31]) so we conclude that there exists a constant

$$c(\{K(i) : i \in S\}) > 0,$$

such that

$$\left\|P_k^{(N)}(i, K(*))\right\| < c(\{K(i) : i \in S\}) \tag{4.50}$$

and

$$\left\|P_l(i, P_k^{(N)}(*, K(*)))\right\| < c(\{K(i) : i \in S\}), \tag{4.51}$$

for all k, l, N, $N \geq k \geq 1$. By (4.43), (4.50) and (4.51) for any $\varepsilon > 0$ we can find $N_0(\varepsilon)$ such that

$$\|\Delta(N,k)\| < \varepsilon, \tag{4.52}$$

for all $N \geq N_0(\varepsilon)$ and $k = 1, ..., N$, where

$$\begin{aligned}\Delta(N,k) = {} & L(A(i), B(i), F_{k-1}^{(N)}(i), R(i), Q(i)) - \\ & L(A_{N-k}(i), B_{N-k}(i), F_{k-1}^{(N)}(i), R_{N-k}(i), Q_{N-k}(i)). \qquad (4.53)\end{aligned}$$

It is easy to check that for any $\varepsilon > 0$ there exists $\zeta > 0$ such that for any symmetric nonnegative definite matrices $U, V \in R^{n\times n}$ if

$$\|U\| < s \times c(\{K(i) : i \in S\}) \max_{i,j\in S} p_{ij},$$

and $\|V\| < \varepsilon$, where s is the number of elements of the set S, then

$$\max_{i\in S} \|W(U,V,i)\| < \zeta\varepsilon, \tag{4.54}$$

where

$$W(U,V,i) = L(A(i), B(i), U+V, R(i), Q(i)) - L(A(i), B(i), U, R(i), Q(i)).$$

Now we prove, by induction with respect to k, that for any $\varepsilon > 0$ and k there exists a constant $c(k)$ such that for any $k, l, N, N - k - l \geq N_0(\varepsilon)$

$$\|Z(N,k,l,i)\| \leq c(k)\varepsilon, \tag{4.55}$$

where

$$Z(N,k,l,i) = P_{k+l}^{(N)}(i, K(*)) - P_k\left(i, P_l^{(N)}(*, K(*))\right).$$

From (4.52) it follows that (4.55) is true for $k = 1$. We will prove it for $k+1$ assuming that(4.55) holds for k. By (4.52):

$$\begin{aligned} & P_{k+1+l}^{(N)}(i, K(*)) = \\ & L(A_{N-k-l-1}(i), B_{N-k-l-1}(i), F_{k+l+1}(i), R_{N-k-l-1}(i), Q_{N-k-l-1}(i)) = \\ & L(A(i), B(i), F_{k+l+1}(i), R(i), Q(i)) - \Delta(N, k+l) \qquad (4.56)\end{aligned}$$

and
$$\|\Delta(N,k+l)\| < \varepsilon,$$
for any $k,l,N,N-k-l-1 \geq N_0(\varepsilon)$. The induction assumptions (4.55), (4.54) with ε replaced by $c(k)\varepsilon$ and (4.56) imply

$$P_{k+1+l}^{(N)}(i,K(*)) =$$
$$L\left(A(i),B(i),\sum_{j\in S} p_{ij}\left(P_k\left(i,P_l^{(N)}(i,K(*))\right)+Z(N,k,l,i)\right),R(i),Q(i)\right)$$
$$-\Delta(N,k+l) =$$
$$L\left(A(i),B(i),\sum_{j\in S} p_{ij}P_k\left(i,P_l^{(N)}(i,K(*))\right)+\sum_{j\in S} p_{ij}Z(N,k,l,i),R(i),Q(i)\right)$$
$$-\Delta(N,k+l) =$$
$$L\left(A(i),B(i),\sum_{j\in S} p_{ij}P_k\left(i,P_l^{(N)}(i,K(*))\right),R(i),Q(i)\right)+$$
$$W\left(\sum_{j\in S} p_{ij}P_k\left(i,P_l^{(N)}(i,K(*))\right),\sum_{j\in S} p_{ij}Z(N,k,l,i),i\right)-\Delta(N,k+l) =$$
$$P_{k+1}\left(i,P_l^{(N)}(*,K(*))\right)+Z(N,k+1,l,i),$$
where
$$Z(N,k+1,l,i) =$$
$$W\left(\sum_{j\in S} p_{ij}P_k\left(i,P_l^{(N)}(*,K(*))\right),\sum_{j\in S} p_{ij}Z(N,k,l,i),i\right)-\Delta(N,k+l)$$
and
$$\left\|W\left(\sum_{j\in S} p_{ij}P_k\left(i,P_l^{(N)}(*,K(*))\right),\sum_{j\in S} p_{ij}Z(N,k,l,i),i\right)-\Delta(N,k+l)\right\| \leq$$
$$c(k+1)\varepsilon,$$
for $c(k+1)=\zeta c(k)+1$. Hence (4.55) holds for $k+1$.

Now we are ready to prove (4.45). Fix a set of initial conditions
$$\{K(i) : K(i) \geq 0, i \in S\}$$

of (4.41) and a real constant $\delta > 0$. Then from Theorem 46 there is k_0, such that
$$\|P_k(i,P_0(*)) - P(i)\| < \frac{\delta}{2}, \tag{4.57}$$
for all $k > k_0, i \in S$, and each initial value $\{P_0(i) : \|P_0(i)\| < c(K(i))\}$. For $\varepsilon = \frac{\delta}{2c(k_0)}$ take N_0 according to (4.52). Then for each k,N such that $N-k > N_0+k_0$ from (4.55) and (4.57) we have
$$\left\|P_{k+k_0}^{(N)}(i,K(*)) - P(i)\right\| \leq$$
$$\left\|P_{k+k_0}^{(N)}(i,K(*)) - P_{k_0}\left(i,P_k^{(N)}(*,K(*))\right)\right\| +$$
$$\left\|P_{k_0}\left(i,P_k^{(N)}(*,K(*))\right) - P\right\| \leq \delta. \tag{4.58}$$

Moreover, by (4.50) we conclude
$$\limsup_{N\to\infty} \frac{1}{N}\left\|\sum_{k=0}^{N} P_k^{(N)}(i,K(*)) - P(i)\right\| \leq$$
$$\limsup_{N\to\infty} \frac{1}{N}\sum_{k=0}^{k_0}\left\|P_k^{(N)}(i,K(*)) - P(i)\right\|$$
$$+\limsup_{N\to\infty} \frac{1}{N}\sum_{k=k_0+1}^{N-N_0}\left\|P_k^{(N)}(i,K(*)) - P(i)\right\|$$
$$+\limsup_{N\to\infty} \frac{1}{N}\sum_{k=N-N_0+1}^{N}\left\|P_k^{(N)}(i,K(*)) - P(i)\right\| =$$
$$\limsup_{N\to\infty} \frac{1}{N}\sum_{k=k_0+1}^{N-N_0}\left\|P_k^{(N)}(i,K(*)) - P(i)\right\|,$$
but from (4.58) we see that
$$\limsup_{N\to\infty} \frac{1}{N}\sum_{k=k_0+1}^{N-N_0}\left\|P_k^{(N)}(i,K(*)) - P(i)\right\| \leq$$
$$\limsup_{N\to\infty} \frac{N-N_0-k_0}{N}\delta = \delta,$$
which yields the conclusion of the theorem. ■

4.5 Coupled algebraic Riccati equation

In this paragraph we study methods of solving coupled Riccati equation related to the time invariant systems. Having in mind the equivalence of exponential stochastic stability and stochastic stability for time invariant systems we have the following result from Theorem 45.

Theorem 48 *Consider the noise free system (4.1) with cost functional (4.3) under assumption that sequences $A_k(i)$, $B_k(i)$, $Q_k(i)$, and $R_k(i)$ are constant for each $i \in S$ and equal to $A(i)$, $B(i)$, $Q(i)$, and $R(i)$, respectively. Assume that*

$$(A(r(k)), B(r(k)), r(k))$$

is stochastically stabilizable and that

$$\left(A(r(k)), \sqrt{Q(r(k))}, r(k)\right)$$

is stochastically detectable. Then there exists exactly one nonnegative solution of the coupled algebraic Riccati equation (4.26), control given by (4.29) is optimal and the minimal value of the cost functional is given by (4.32).

Observe that coupled algebraic Riccati equation (4.26) can be rewritten as follows

$$Q_i + A_i'F_iA_i - A_i'F_iB_i\left(R_i + B_i'F_iB_i\right)^{-1}B_i'F_iA_i - P_i = 0. \tag{4.59}$$

where $F_i = \sum_{j\in S} p_{ij}P_j$. Assume that $p_{ii} > 0$. Under this assumption we can divide this equation by p_{ii} and consider a new coupled Riccati equation

$$Q_i + A_i'F_iA_i - A_i'\widehat{F}_iB_i\left(I + B_i'F_iB_i\right)^{-1}B_i'F_iA_i - P_i = 0 \tag{4.60}$$

where $F_i = P_i + \sum_{j\neq i} p_{ij}P_j$. Simply the following substittuion has been made

$$A_i \to \sqrt{p_{ii}}A_i, B_i \to \sqrt{p_{ii}}B_iR_i^{-1/2}, p_{ij} \to \frac{p_{ij}}{p_{ii}}, \tag{4.61}$$

The next Theorem taken from [32] presents lower bounds for the solution of (4.59). In order to present this theorem we introduce the following notation and constants: The eigenvalues $\lambda_i(X)$, $i = 1, ..., n$, of a symmetric matrix $X \in R^{n\times n}$ are assumed to be arranged such that $\lambda_1(X) \geq \lambda_2(X) \geq ... \geq \lambda_n(X)$. Consider the following scalar functions $f(a,b,c) = \left(-a + \sqrt{a^2 + bc}\right)/b$ for $b \neq 0$ and denote

$$p_d = \min_{j\neq i}\{p_{ij}\}$$

$$\overline{a}_1 = \min_{i\in S}\lambda_n\left(A_i'A_i\right) - 1,\ \overline{\pi}_1 = \min_{i\in S}\left\{\sum_{j\neq i} p_{ji}\right\},$$

$$\overline{r} = \max_{i\in S}\lambda_1\left(B_iB_i'\right),\ \overline{q}_1 = \sum_{i\in S}\sum_{k=1}^{l}\lambda_{n-k+1}\left(Q_i\right).$$

$$\overline{M}_i = \lambda_1\left(B_iB_i'\right)p_d\alpha_d + (p_d - 1)\left(\lambda_n\left(A_i'A_i\right) + \lambda_1\left(B_iB_i'\right)\sum_{k=1}^{l}\lambda_{n-k+1}\left(Q_i\right)\right)$$

$$\alpha_d = f\left(-\overline{a}_1 + 1 - \overline{a}_1\overline{\pi}_1 - \overline{r}\overline{q}_1\left(\overline{\pi}_1 + 1\right), 2\overline{r}\left(\overline{\pi}_1 + 1\right), 2\overline{q}_1\right)$$

$$\overline{N}_i = \left(\sum_{k=1}^{l}\lambda_{n-k+1}\left(Q_i\right)\right)\left(1 + \lambda_1\left(B_iB_i'\right)p_d\alpha_d\right) + \lambda_n\left(A_i'A_i\right)p_d\alpha_d$$

$$\theta(l,i) = f\left(\overline{M}_i, 2\lambda_1\left(B_iB_i'\right)(1 - p_d), 2\overline{N}_i\right)$$

Theorem 49 *Let the positive definite matrices P_i, $i \in S$ satisfy (4.60). Then*

$$\sum_{k=1}^{l}\lambda_{n-k+1}\left(P_i\right) \geq \theta(l,i) \tag{4.62}$$

for $l = 1, ..., n$ and $i \in S$.

When we consider $l = 1$ in (4.62) and take into account that for nonnegative definite matrix X we have $X \geq \lambda_n(X)I$, then we get the following matrix lower bound for solution of (4.62)

$$P_i \geq A_i'\left(\frac{I}{\sum_{j\neq i} p_{ij}\theta(1,j) + \theta(1,i)} + B_iB_i'\right)^{-1}A_i + Q_i, \tag{4.63}$$

$i \in S$. As an illustration of this result consider the following example.

Example 14 *Consider (4.59) with the following parameters $S = \{1, 2\}$,*

$$A_1 = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}, A_2 = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, B_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, B_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$(p_{ij})_{i,j\in S} = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}, Q_1 = Q_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Then the solution of (4.59) is

$$P_1 = \begin{bmatrix} 5.1425265 & 5.8765969 \\ 5.8765969 & 11.409165 \end{bmatrix} \text{ and } P_2 = \begin{bmatrix} 1.7757412 & -0.77574 \\ -0.77574 & 1.7757412 \end{bmatrix}$$

with

$$\lambda_1(P_1) = 14.935582, \lambda_2(P_1) = 1.616099, \lambda_1(P_2) = 2.5514824, \lambda_2(P_2) = 1$$

and (4.63) gives (taking into account substtiution (4.61))

$$\lambda_1(P_1) \geq 5.9028979, \lambda_2(P_1) \geq 1.3831878,$$
$$\lambda_1(P_2) \geq 1.8175436, \lambda_2(P_2) \geq 1.$$

Now we present numerical algorithm to solve (4.59). This algorithm has been proposed in [99] and we will present it in a form of theorem. For a sequence of matrices $X = (X_1, ..., X_s)$ denote

$$\Phi_i(X) = \sum_{j \in S} p_{ij} X_j$$

for all $i \in S$.

Theorem 50 *Consider coupled algebraic Riccati equation (4.59) under assumption that*

$$(A(r(k)), B(r(k)), r(k))$$

is stochastically stabilizable and that

$$\left(\sqrt{p_{r(k),r(k)}}A(r(k)), \sqrt{Q(r(k))}, r(k)\right)$$

is stochastically detectable. Then the following decoupled Riccati equation

$$P_i^{(\eta)} = Q_i + A_i'\Phi_i\left(\widetilde{P}_i^{(\eta)}\right)A_i + p_{ii}A_i'P_i^{(\eta)}A_i$$
$$-\left(p_{ii}A_i'P_i^{(\eta)}B_i + A_i'\Phi_i\left(\widetilde{P}_i^{(\eta)}\right)B_i\right)\times$$
$$\left(R_i + B_i'\Phi_i\left(\widetilde{P}_i^{(\eta)}\right)B_i + p_{ii}B_i'P_i^{(\eta)}B_i\right)^{-1}\times$$
$$\left(p_{ii}A_i'P_i^{(\eta)}B_i + A_i'\Phi_i\left(\widetilde{P}_i^{(\eta)}\right)B_i\right)',$$

where $P_i^{(0)} = 0$, and $\widetilde{P}_i^{(\eta)} = \left(\widetilde{P}_{i1}^{(\eta)}, ..., \widetilde{P}_{is}^{(\eta)}\right)$ is defined as follows

$$\widetilde{P}_{ii}^{(\eta)} = 0, \ \widetilde{P}_{ij}^{(\eta)} = P_j^{(\eta)} \text{ for } j = 1, ..., i-1$$

and

$$\widetilde{P}_j^{(\eta)} = P_j^{(\eta-1)} \text{ for } j = i+1, ..., s$$

has solution $P_i^{(\eta)}$ for all $i \in S$. Moreover

$$\lim_{\eta \to \infty} P_j^{(\eta)} = P_j,$$

where $(P_i, i \in S)$ is the unique nonnegative definite solution of coupled Riccati equation (4.59).

The recurrent algorithm desribed in the above theorem converges also when the initial condition $P_i^{(0)}$ is any nonegative definite matrix such that $P_i^{(0)} \leq P_i, i \in S$, where $(P_i, i \in S)$ is the unique nonnegative definite solution of coupled Riccati equation 4.59. To design such initial conditions one may use inequality (4.63).

4.6 Noise JLQ problem on infinite time interval

In this section we assume that the Markov chain $r(k)$ is ergodic with limit distribution $\pi_l(i)$. We start with the following counterparts of Theorem 44 and Theorem 45 for the noise system.

Theorem 51 *Consider noise system (4.1) with cost functional (4.4). If the noise free system*

$$(A_k(r(k)), B_k(r(k)), Q_k(r(k)), R_k(r(k)), r(k))$$

is optimizable then control given by (4.20) is optimal for (4.1) with cost functional (4.4) and the minimal value of the cost functional is

$$\sum_{i \in S}\sum_{j \in S} \pi_l(i) p_{ij} tr\left(C'(i)P(j)C(i)\right).$$

The proof of this theorem, which is very similar to the proof of Theorem 44 and Theorem 45, may be found in [31].

Now we consider time varying noise system (4.1) with cost functional (4.4) under the following assumptions

$$(A_N(j), B_N(j), Q_N(j), R_N(j)),$$
$$A_N(j) \in R^{n \times n}, B_N(j) \in R^{n \times m}, C_N(j) \in R^{n \times n},$$

$$Q_N(j) \in R^{n\times n}, R_N(j) \in R^{m\times m}, Q_N(j) \geq 0, R_N(j) > 0,\ j \in S,$$

the limits of $A_N(j)$, $B_N(j)$, $Q_N(j)$, $R_N(j)$ when N tends to infinity, exist for each $j \in S$ and $R(j) > 0$, $(A(i), B(i), i \in S)$ is stochastically stabilizable, $\left(A(i), \sqrt{Q(i)}, i \in S\right)$ is stochastically detectable, where

$$A(j) = \lim_{N\to\infty} A_N(j),\ B(j) = \lim_{N\to\infty} B_N(j),$$

$$Q(j) = \lim_{N\to\infty} Q_N(j),\ R(j) = \lim_{N\to\infty} R_N(j),\ j \in S,$$

The next Theorem shows that for the system with this property the optimal control for the time varying noise system can be realized in the form of a time invariant feedback gain.

Theorem 52 *Under the above formulated assumptions the optimal control law for the time-varying noise system (4.1) with cost functional (4.4) is given by*

$$\tilde{u}_k = -L(i)x_k, i \in S, \tag{4.64}$$

where

$$L(i) = \left(R(i) + B'(i)F(i)B(i)\right)^{-1} B'(i)F(i)A(i). \tag{4.65}$$

$$F(i) = \sum_{j\in S} p_{ij}P(j) \tag{4.66}$$

and the set of positive-semidefinite symmetric matrices $\{P(i) : i \in S\}$ is the unique solution of the coupled algebraic Riccati equation (4.26).

Proof. From Theorem 51 and Theorem 52 we have for any control $u = (u_0, \ldots)$

$$J(x_0, u) \geq \lim_{N\to\infty} \frac{1}{N}\left(\left\langle P_N^{(N)}(r_0, K(*))\, x_0, x_0\right\rangle + \right.$$

$$\left.\sum_{j\in S} tr\left(C'_{N-k}(j)P_{k-1}^{(N)}(j, K(j))C_{N-k}(j)\right)\right) = \tag{4.67}$$

$$\sum_{i\in S}\sum_{j\in S} \pi(i)p_{ij}tr\left(C'(i)P(j)C(i)\right).$$

Since, as we know from Lemma 6, the family of matrices $P_{k-1}^{(N)}(j, K(*))$ is uniformly bounded with respect to N and $k = 1, \ldots, N$. Inequality (4.67) shows that the cost functional (3.27) does not take a value less than

$$\sum_{i\in S}\sum_{j\in S} \pi(i)p_{ij}tr\left(C'(i)P(j)C(i)\right).$$

The calculation similar to that in the proof of Theorem 2 in [25] shows that for the control given by (4.64)-(4.66) the cost functional takes the value

$$\sum_{i\in S}\sum_{j\in S} \pi(i)p_{ij}tr\left(C'(i)P(j)C(i)\right). \blacksquare$$

As an illustration of the above theorem consider again Example 13 but now for the noise case.

Example 15 *Consider the noise system (4.1) $n = 1$, $s = 2$,*

$$A_k(1) = a_k, A_k(2) = 0.$$

$$a_k = \sqrt{\frac{(6k+13)(2k+3)}{(k+1)(4k+9)}}$$

$$B_k(1) = B_k(2) = 1, C_k(1) = C_k(2) = 1 + \frac{1}{k+1}$$

$$Q_k(1) = Q_k(2) = R_k(1) = R_k(2) = 1$$

Suppose that the Markov chain has the transition matrix of the form

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{bmatrix}.$$

We see that the assumption of Theorem 52 are satisfied with

$$A(1) = \sqrt{3}, A(2) = 0.$$

$$B(1) = B(2) = 1, C(1) = C(2) = 1$$

$$Q(1) = Q(2) = R(1) = R(2) = 1.$$

According to Theorem 52 the optimal time invariant feedback takes the following form

$$L(1) = \frac{2\sqrt{3}}{3}$$

$$L(2) = 1.$$

4.7 Comparison and discussion

The main result of Section 3, Theorem 43, is the first result in the literature about time-varying JLQ problem on infinite time interval. Also Theorem 45, which gives conditions for stability of the optimal closed loop system and for uniquness of global and bounded solution of coupled difference Riccati equation, has never been published. Both these theorems are extensions of results in [54], where only parameters A_k in (4.1) are random and the random changes are independent and identically distributed. In this case the Riccati equation related to this problem is simpler for analysis because it consists of one equation only, whereas in our case it is a set of equations.

The time invariant JLQ on infinite time interval has been solved in many papers [58], [59], [24] however under much stronger assumptions than those from Theorem 4.4. Moreover assumptions of this Theorem are also necessary for the existence of optimal control and therefore cannot be weakened

In Section 4 we present several properties of coupled Riccati equation. The conditions for existence and uniqueness of nonnegative definite global and bounded solution given by Theorem 4.3 have not been published, whereas its time invariant counterpart, Theorem 48, has been already proved in [24]. Theorem 47 deserves special attention. It enables to show that for time-varying noise JLQ problem on infinite time interval the optimal control can be realized in the time invariant feedback form under assumptions that the coefficients of the system and the cost functional converge (see Theorem 4.8). Such a result for standard LQG problem has been shown in [26]. Results of this kind may by used in solving of adaptive control problems.

Appendix

We use the following definition of stationary Markov chain.

Definition 11 *Let $(\Omega, \mathcal{F}, P)$ be a probability space and let S be a finite set. A sequence $(X_n)_{n\in \mathcal{N}_0}$, where $X_n : \Omega \to S$ is called a Markov chain with initial distribution $\pi = (p(i))_{i\in S}$ if and only if π is the distribution of X_0 and for each natural n and any $i_0, i_1, ..., i_n \in S$*

$$P\left(X_n = i_n \middle| X_0 = i_0, ..., X_{n-1} = i_{n-1}\right) = P\left(X_n = i_n \middle| X_{n-1} = i_{n-1}\right),$$

whenever the left member is definite. If for fixed $i, j \in S$ the probability $P\left(X_n = i \middle| X_{n-1} = j\right)$ is the same for all n then the Markov chain is called stationary.

Since we consider only stationary Markov chains we omit the term stationary and we use the term Markov chain instead of stationary Markov chain.

The matrix $P = [p(i,j)]_{i,j\in S}$, where

$$p(i,j) = P\left(X_n = j \middle| X_{n-1} = i\right) = P\left(X_1 = j \middle| X_0 = i\right)$$

is called transition matrix of the Markov chain and the distribution π of X_0 is called initial distribution of the Markov chain. Let denote

$$p(n,i,j) = P\left(X_n = j \middle| X_0 = i\right) = P\left(X_{n+k} = j \middle| X_k = i\right).$$

It can be shown that

$$[p(n,i,j)]_{i,j\in S} = P^n.$$

Denote by $f(n,i,j)$ the probability that the Markov chain will be in j for the first time at n-th step, given that it starts from i it is

$$f(n,i,j) = P\left(X_k \neq j, 0 < k < n, X_n = j \middle| X_0 = i\right).$$

Now we present a list of definitions and theorems about stationary Markov chains. All of them are standard results and can be found in any textbook devoted to stationary Markov chains, for example [57], [50].

Definition 12 *(Classification of states) A state $j \in S$ is said to be accessible from state $i \in S$ if there exists n such that $p(n,i,j) > 0$. States $i, j \in S$ are said to communicate if each is accessible from the other. A subset of the set of all states S is called a communicating class of states if all the states are accessible from the other. A communicating class is called closed if no states outside the class is accessible from a state inside the class. A state i is called recurrent or persistent if $\sum_{n=1}^{\infty} f(n,i,i) = 1$, otherwise it is called transient.*

Definition 13 *A state j is called to have period d if d is the greatest common divisor of all those n's for which $p(n,j,j) > 0$. A Markov chain is called aperiodic if each state has the period equal to 1.*

Theorem 53 *If C is a closed class of states and $i \in C$ is a persistent state then each state $j \in C$ is persistent.*

Theorem 54 *Set S of all states of the Markov chain can be divided in an unique way into disjoint sets $T, C_1, ..., C_r$, where all the states in T are transient and each C_i is a communicating class of states. Moreover, all the states from C_i are recurrent and C_i are closed.*

Theorem 55 *For each state $i \in S$ there exists a recurrent state $j \in S$ and a number n such that $p(n,i,j) > 0$. Moreover, the expectation of the time of reaching j is finite.*

Definition 14 *A Markov chain is called irreducible if there does not exist a nonempty closed set other than S itself.*

Theorem 56 *A Markov chain is irreducible if and only if for all pairs i, j of the state i is accessible from j.*

Theorem 57 *Let η denote the time of the leaving of the set T of transient states, defined as follows to state*

$$\eta = \min\{k \geq 1 : r(k) \notin T\}.$$

Then

$$E(\eta|\, r(0) = i) < \infty$$

for all $i \in C$ in particular

$$P(\eta < \infty|\, r(0) = i) = 1$$

for all $i \in C$.

Remark 9 *Let τ_j denote the time of the first visit to state $j \in C_p$, defined as follows $\tau_j = \min\{k \geq 1 : r(k) = j\}$. If the state j is persistent then of course*

$$P(\tau_j < \infty|\, r(0) = j) = 1.$$

It can be shown that

$$P(\tau_j < \infty|\, r(0) = i) = 1$$

for all $i \in C_p$ and $E(\tau_j|\, r(0) = i) < \infty$.

Definition 15 *A Markov chain is called ergodic if there exists only one close class of communicating states and it is aperiodic.*

Theorem 58 *If the Markov chain is ergodic than $\lim_{n\to\infty} p(n,i,j) = \pi_j$ exists for all j independly of i and then the distribution $\pi = (\pi_i)_{i\in S}$ is called limit distribution. The matrix*

$$\overline{P} = \begin{bmatrix} \pi_1 & \dots & \pi_s \\ \vdots & \ddots & \vdots \\ \pi_1 & \dots & \pi_s \end{bmatrix}$$

satisfies equations

$$\overline{P}P = P\overline{P} = \overline{P} = \overline{P}^2.$$

Theorem 59 *Suppose that the Markov chain is ergodic with limit distribution $(\pi_i)_{i\in S}$ and let*

$$f : S \to R^n$$

be any function, then

$$\lim_{N\to\infty} \frac{1}{N} \sum_{s=0}^{N} f(r(s)) = \sum_{i\in S} \pi_i f(i), \text{ a.s..}$$

Theorem 60 *Suppose that the Markov chain is ergodic with limit distribution $(\pi_i)_{i\in S}$ and let $f : S \to R^n$ be any function, then there exists positive constant a such that*

$$\limsup_{N\to\infty} \frac{\sum_{s=0}^{N} f(r(s)) - N\sum_{i\in S} \pi_i f(i)}{\sqrt{aN \lg\lg N}} = 1, \text{ a.s.}$$

Theorem 61 *Let $r(n)$, $n = 0, 1, ...$ be a Markov chain. Fix a number k and consider a sequence of random variables $\overline{r}(n)$, $n = 0, 1, ..., \overline{r}(n) = (r(n), ..., r(k+n))$. Then $\overline{r}(n)$, $n = 0, 1, ...$ is a Markov chain with state space*

$$S_{\pi}^{(k)} = \left\{(i_0, ... i_{k-1}) \in S^k : p(i_0, i_1) ... p(i_{k-2}, i_{k-1}) > 0\right\},$$

and if the state space S of $r(n)$ consists of one close communicating class then so does the space of $\overline{r}(n)$.

In fact Theorem 61 in [57] is proved for $k = 2$ but the extension for $k \geq 2$ is straightforward.

From Theorem 61 and Remark 9 one can easily get the following theorem.

Theorem 62 *Suppose that the space of the Markov chain $r(n)$ consists of one close communicating class then for each $i_0, \ldots i_{k-1} \in S$ such that*

$$p(i_0, i_1) \ldots p(i_{k-2}, i_{k-1}) > 0$$

we have

$$E\left(\eta_{i_0, \ldots i_{k-1}} \middle| r(0) = i\right) < \infty$$

where

$$\eta_{i_0, \ldots i_{k-1}} = \min\{l \geq k : r(l-1) = i_{k-1}, \ldots, r(l-k) = i_0\}.$$

Bibliography

[1] H. Abou-Kandil, G. Freiling, G. Jank, On the solution of discrete-time Markovian jump linear quadratic control problems, Automatica, vol. 31, pp. 765-768, 1995.

[2] J. C. Allwright, A lower bound for the solution of the algebraic Riccati equation of optimal control and a geometric convergence for the Kleinman algorithm, IEEE Transactions on Automatic Control, vol. 25, 826-829, 1980.

[3] R. Akella, P. R. Kumar, Optimal control of production rate in failure prone manufacturing systems, IEEE Transactions on Automatic Control, vol. 31, pp. 116-126, 1986.

[4] L. Arnold, A formula connecting sample and moment stability of linear stochastic systems, SIAM Journal on Applied Mathematic, vol. 44, no. 3, pp. 793-802, 1984.

[5] L. Arnold, V. Wihstutz (eds.), Lyapunov Exponents, Lecture Notes on Mathematics, no. 1186, Springer-Verlag, New York, 1986.

[6] L. Arnold, H. Crauel, J. P. Eckmann (eds.), Lyapunov Exponents, Lecture Notes in Mathematics, no. 1486, Springer-Verlag, New York, 1991.

[7] M. Athans, Command and control (C2) theory: a challange to control science, IEEE Transactions on Automatic Control, vol. 32, no. 4, pp. 286-293,1987.

[8] Y. Bar Shalom, T. E. Fortman, Tracking and Data Association, Academic Press, New York, 1988

[9] Y. Bar Shalom, Tracking methods in a multitarget environment, IEEE Transactions on Automatic Control, vol. 23, pp. 618-628, 1978.

[10] Y. Bar Shalom, K. Birmiwal, Variable dimension for maneuvering target tracking, IEEE Transactions on Aero. Electr. Systems, AES-18,pp. 621-628, 1982.

[11] A. P. Belle Isle, F. Kozin, On the almost sure sample stability of systems with randomly time-varying delays, Automatica, vol. 8, pp. 755-763, 1972.

[12] A. P. Belle Isle, Stability of systems with nonlinear feedback through randomly time-varying delays, IEEE Transactions on Automatic Control, vol. 20, no. 1, pp. 67-75,1975.

[13] W. P. Jr. Blair, D. D. Sworder, Feedback control of a class of linear discrete systems with jump parameters and quadratic cost criteria, International Journal on Control, vol. 21, no 2, pp. 833-844, 1975.

[14] M. Bochner, Introduction to Higher Algebra, Mac Millan, New York, 1931.

[15] J. E. Bertran, P. E. Sarchik, Stability of circuits with randomly time-varying parameters, Transactions of IRE, PGIT-5, Special Supplement, pp. 260-270, 1959.

[16] T. Bielecki, P. R. Kumar, Necessary and sufficient conditions for a zero inventory policy to be optimal in an unreliable manufacturing systems, Proccidings of 25 th IEEE Conference on Decision And Control, Athens, 1986, pp. 248-250.

[17] J. G. Birdwel, On reliable control systems design, Ph. D. Disertation, Electrical Systems Lab., Mass. Inst. Technology, report no. ESL-TH-821, 1978.

[18] J. G. Birdwel, D. A. Castanon, M. Athans, On reliable control system designs, IEEE Transaction on Systems, Man and Cybernetics, vol. 16, no. 5, pp. 703-711, 1986.

[19] E. K. Boukas, A. Swierniak, H. Yang, On the robustness of jump linear quadratic control, International Journal of Robust and Nonlinear Control, vol. 10, no. 7, pp. 899-910, 1997.

[20] E. K. Boukas, A. Swierniak, K. Simek, H. Yang, Robust stabilization and guaranteed cost control of large scale linear systems with jumps, Kybernetika, vol. 33, no. 1, pp. 121-131, 1997

[21] E. K. Boukas,Z. K. Liu, Production and maintenance control for manufacturing systems,IEEE Transactions on Automatic Control, vol. 46, no. 9, pp. 1455-1460, 2001.

[22] H. F. Chen, Estimation and Control for Stochastic Systems, Wiley, New York, 1985.

[23] H. J. Chizeck, A.S. Willsky, D. Castanon, Discrete-time markovian-jump linear quadratic optimal control, International Journal of Control, vol 43, no. 1, pp. 213-231, 1986.

[24] M. D. Costa, M. D. Fragoso, Stability results for discrete-time linear systems with Markovian jumping parameters, Journal of Mathematical Analysis and Applications, vol. 179, no. 1, pp. 154-178, 1993.

[25] O. L. V. Costa, M. D. Fragoso, Discrete-time LQ-optimal control problems for infinite Markov jump parameter systems, IEEE Transactions on Automatic Control, vol. 40, no. 12, pp. 2076-2088,1995.

[26] A. Czornik, On discrete-time linear quadratic-control, Systems and Control Letters, vol. 36, pp. 101-107, 1999.

[27] A. Czornik, A. Swierniak, On controllability with respect to the expectation of discrete time jump linear systems, Journal of the Franklin Institute, vol. 338, pp. 443-453, 2001.

[28] A. Czornik, A. Swierniak, Controllability of discrete time jump linear systems, submitted to Dynamics of Continous, Discrete and Impulsive Systems.

[29] A. Czornik, A. Swierniak, On controllability of discrete time jump linear systems, Procedings of 4th World CSCC, Athens, Greece, pp. 3041-3044, 2000.

[30] A. Czornik, A. Nawrat, On the bounds on the solutions of the algebraic Lyapunov and Riccati equations, Archives of Control Sciences, vol. 10, no 3-4, pp. 197-244, 2000.

[31] A. Czornik, A. Swierniak, On the discrete JLQ and JLQG problems, Nonlinear Analysis, vol.47, pp. 423-434, 2001.

[32] A. Czornik, A Swierniak, Lower bounds on the solution of coupled algebraic Riccati equation, Automatica, vol. 37, no 4, pp. 619-624, 2001.

[33] A. Czornik, A. Swierniak, On the discrete time-varying JLQG problem, International Journal of Applied Mathematics and Computer Science, vol. 12, No. 2, pp 101-105, 2002.

[34] A. Czornik, A. Swierniak, Controllability of continuous time jump linear systems, Proc. of 10th Mediterranean Conference on Control and Automation, CD-ROM, Portugal, Lisbon, 2002.

[35] A. Czornik, A. Swierniak, On the sensitivity of the coupled discrete-time Lyapunov equation, IEEE Transactions on Automatic Control, vol. 46, no. 5, pp. 659–664, 2001.

[36] A .Czornik, A. Swierniak, On stability of jump linear systems, in"Systems and Control: Theory and Applications" , Editor: Nikos E. Mastorakis, World Scientific and Engineering Society Press, pp.156-168, 2001.

[37] A. Czornik, A. Swierniak, On the sensitivity of the coupled continous-time Lyapunov equation, IEEE Transactions on Automatic Control, vol. 47, no.7, pp. 1138–1142, 2002.

[38] A. Czornik, A. Nawrat, S. Rabsztyn, On the stability of one dimensional discrete-time jump linear systems, aceptted for publication in Archives of Control Sciences 2003.

[39] N. H. Du, T. V. Nhung, Relations between the sample and moment Lyapunov exponents, Stochastics and Stochastic Reports, vol. 37. pp. 201-211, 1991.

[40] M. Ehrhardt and W. Kliemann, Controllability of linear stochastic system, Systems and Control Letters, vol. 2, pp. 145-153, 1982.

[41] Y. Fang, Stability analysis of linear control systems with uncertain parameters, Ph. D. dissertation, Department of Systems, Control and Industrial Engineering, Case Western Reserve University, Cleveland, Ohio, 1994.

[42] Y. Fang, A new general condition for almost sure stability of jump linear systems, IEEE Transactions on Automatic Control, vol. 42, no. 3, pp. 378-382, 1997.

[43] Y. Fang, K. A. Loparo, X. Feng, Almost sure and δ-moment stability of jump linear systems, International Journal of Control, vol 59, no. 5, pp. 1281-1307, 1994.

[44] Y. Fang, K. A. Loparo, On the relathionship between the sample path and moment Lyapunov exponents for jump linear systems, IEEE Transactions on Automatic Control, vol. 47, no. 9, pp. 1556-1560, 2002.

[45] X. Feng, K. A. Loparo, Almost sure instability of the random harmonic oscillator, SIAM Journal on Applied Mathematic, vol. 50, no. 3, pp. 744-759, 1990.

[46] X. Feng, K. A. Loparo, Y. Ji, H. J. Chizeck, Stochastic stability properties of jump linear systems, IEEE Transactions on Automatic Control, vol. 37, no. 1, pp. 38-53, 1992.

[47] X. Feng, K. A. Loparo, A nonrandom spectrum for Lyapunov exponents of linear stochastic systems, Stochastic Analysis and Applications, vol. 9, no. 1, pp. 25-40, 1991.

[48] X. Feng, K. A. Loparo, Stability of discrete-time jump linear systems, Journal on Mathematical Systems, Estimation and Control, vol. 5, no. 3, pp. 275-321, 1995.

[49] J. J. Florentin, Optimal control of continuous-time Markov stochastic systems, Journal of Electronics Control, vol. 10, pp. 473-481, 1961.

[50] D. Freedman, Markov Chains, Springer-Verlag, New York, 1983.

[51] H. Furstenberg, H. Kesten, Products of random matrices, Ann. Math. Staitist, vol. 31, pp. 457-489, 1960.

[52] A. Gajic, M. Qureshi, Lyapunov Matrix Equation in System Stability and Control, Mathematics in Science and Engineering, vol. 195, Academic Press, London, 1995.

[53] B. E. Griffith, K. Loparo, Optimal control of jump- linear gaussian systems, International Journal of Control, vol 42, no. 4, pp. 791-819, 1985.

[54] A. Halanay, V. Ionescu, Time-Varying Discrete Linerar Systems, Birkhauser, Basel, 1994.

[55] R. Z. Hasminskii, Stochastic Stability of Differential Equations, Sijthoff&Noordhoff, Maryland, 1980.

[56] R. A. Horn, C. R. Johnson, Matrix Analysis, Cambridge University Press, 1985.

[57] M. Iosifescu, Skończone procesy Markowa i ich zastosowania, PWN Warszawa, 1998.

[58] Y. Ji, H. Chizeck, Controllability, observability and discrete-time markovian jump linear quadratic control, International Journal of Control, vol 48, no. 2, pp. 481-498, 1988.

[59] Y. Ji, H.J. Chizeck, Jump linear quadratic gausian control: steady state solution and testable conditions, Control-Theory and Advanced Technology, vol. 6, pp. 289-319, 1990.

[60] Y. Ji, H.J. Chizeck, Controllability, stability, and continuous-time Markovian jump linear quadratic control, IEEE Transactions on Automatic Control, vol. 35, pp. 777-788, 1990.

[61] Y. Ji, H. J. Chizeck, X. Feng, K. A. Loparo, Stability and control of discrete-time jump linear systems, Control Theory and Advanced Technology, vol. 7, pp.247-270, 1991.

[62] T. Kazangey, D. D. Sworder, Effective federal policies for regulating residential housing, Proc. Summer Comp. Simulation Conf., Los Angeles, pp.1120-1128, 1971.

[63] R. E. Kalman, On the general theory of control systems, Proc. First IFAC Congress Automatic Control, Moscow, Butterworths, London, vol.1, pp. 481-492, 1960.

[64] I. Kats, N. N. Krasovskii, Ob usticzewosti sistiem sa sluczajnymi paramietrami, Prikl. Matiem. i Miech., vol. 24, no.5, pp. 809-823, 1960

[65] J. Klamka and L. Socha, Some remarks about stochastic controllability, IEEE Transactions on Automatic Control, vol. 22, pp. 880-881, 1977.

[66] J. Klamka, L. S. Dong, Stochastic controllability of discrete time systems with jump Markov disturbances, Archiwum Automatyki i Telemechaniki, vol. 35, pp. 67-74, 1990.

[67] N. Komaroff, On Bounds for the solution of the Riccati equation for discrete-time control system, Control and Dynamics Systems, vol. 76, pp. 275-311, 1996.

[68] F. Kozin, A survey of stability of stochastic systems, Automatica, vol. 5, pp. 95-112, 1969.

[69] F. Kozin, S. Sugimoto, Relations between sample and moment stability for linear stochastic differential equations, Proceedings of the Conference on Stochastic Differential Equations, D. Mason, ed., Academic Press, New York, 1977.

[70] N.N. Krasovski, E. A. Lidski, Analytical design of controllers in systems with random attribute, Parts I-III, Automation and Remote Control, vol. 22, pp. 1021-1027 (Part I), pp. 1141-1146 (Part II), pp. 1289-1297 (Part III), 1961.

[71] R. Krtolica, U. Ozguner, H. Chan, H. Gotkas, J. Winkelman, M. Liubakka, Stability of linear feedback systems with random communication delays, Procidings of ACC 1991, Boston MA, June 26-28, 1991.

[72] H. J. Kushner, Stochastic Stability and Control, Academic Press, New York, 1967.

[73] H. Kwakernaak, R. Sivan, Linear optimal control systems, John Wiley & Sons, New York, 1972.

[74] W. H. Kwon, Y. S. Moon and S. C. Ahn, Bounds in Algebraic Riccati and Lyapunov Equations: Survey and Some New Results, International Journal of Control, vol 64, no. 3, 377-389, 1996.

[75] P. Lancaster and L. Rodman, Algebraic Riccati equation, Oxford Univ. Press, 1995.

[76] G. Langholz, A New Lower Bound on the Cost of Optimal Regulator, IEEE Transactions on Automatic Control, vol. 24, 353-354, 1979.

[77] C. H. Lee, T.-H.S. Li, and F. C. Kung, A New Approach for the Robust Stability of Perturbed Systems with Class of Non-Commensurate Time Delays, IEEE Transactions on Circuits and Systems -I, vol. 40, 605-608, 1995.

[78] A. Leizarowitz, Estimates and exect expressions for Lyapunov exponents of stochastic linear differential equations, Stochastics, vol. 24, pp. 335-356, 1988.

[79] Z. G. Li, Y. C. Soh, C. Y. Wen, Sufficient conditions for almost sure stability of jump linear systems, IEEE Transactions on Automatic Control, vol. 45, no. 7, pp. 1325-1329, 2000.

[80] R. Malhame, C. Y. Chong, Electric load model synthesis by diffusion aproximation in a high order hybrid state stochastic systems, IEEE Transactions on Automatic Control, vol. 30, pp. 854-861, 1985.

[81] M. Mariton, Almost sure and moment stability of jump linear systems, Systems and Control Letters, vol. 11, pp. 393-397, 1988.

[82] M. Mariton, Jump Linear Systems in Automatic Control, Marcel Dekker, New York and Besel, 1990.

[83] A. W. Marshall and I. Olkin, Inequalities: Theory of Majorization and its Applications, Academic Press, New York, 1979.

[84] R. R. Mitchell, F. Kozin, Sample stability of second order linear differential equation with wide band noise coefficients, Siam Journal of Applied Mathematics, vol. 27, pp. 5701-605, 1974.

[85] R. C. Montgomery, Reliability considerations in placement of control systems components, Proc. AIAA Guidance and Control Conference, Gatlinburg, 1983.

[86] T. Mori, E. Noldus, M. Kuwahara, A way to stabilizable linear system with delayed state, Automatica, vol. 19, pp. 571-574, 1983.

[87] T. Mori and I. A. Derese, A brief summary of the bounds on the solution of the algebraic matrix equations in control theory, International Journal of Control, vol 39, no. 2, 247-256, 1984.

[88] T. Morozan, Stabilization of some stochastic discrete-time control systems, Stochastic Analysis and Applications, vol. 1, no. 1, pp. 89-116, 1983.

[89] T. Morozan, Optimal stationary control for dynamic systems with Markov perturbations, Stochastic Analysis and Applications, vol. 3, no. 1, pp. 299-325, 1983.

[90] V. Z. Oseledec, A multiplicative ergodic theorem: Lyapunov characteristic numbers for dynamic systems, Trans. Moscow Math. Soc., vol. 19, pp. 197-231, 1968.

[91] R. V. Patel and M. Toda, Quantitative Measures of Robustness for Multivariable Systems, Proceedings of the Joint Automatic Control Conference, San Francisco, TP8-A, pp. 45-49, 1980.

[92] H .H. Rosenbrock, P. D. McMorran, Good, bad, or optimal?, IEEE Transactions on Automatic Control, vol.16, no.6, pp.552-554, 1971.

[93] D. D. Siljak, Reliable control using multiple control systems, International Journal on Control, vol. 31, no 2, pp. 303-329, 1980.

[94] D. D. Siljak, Dynamic reliability using multiple control systems, Proc. 2nd Lawrece Syphosium on Systems Decision Sciences, Berkley, 1978.

[95] A. Swierniak, K. Simek, A. Czornik, Fault tolerant control design for linear systems, 5th National Conference on Diagnostics of Industrial Processes. Tech. Univ. Press of Zielona Gora. 2001, pp.45-50. Zielona Gora, Poland.

[96] A. Swierniak, K. Simek, E. K. Boukas, Intelligent robust control of fault tolerant linear systems, Proceedings IFAC Symposium on Artificial Intelligence in Real-Time Control, Malaysia, Kuala Lumpur, Pergamon Press, pp. 245-248,1997.

[97] D. D. Sworder, Feedback control of a class of linear systems with jump parameters, IEEE Transactions on Automatic Control, vol. 14, no. 1, pp. 9-14, 1969.

[98] D. D. Sworder and R. O. Rogers, An LQ- solution to a control problem associated with solar thermal central receiver, IEEE Transactions on Automatic Control, vol. 28, pp. 971-978, 1983.

[99] J. B. R. do Val, J. C. Geromel and O. L. V. Costa, Uncoupled Riccati iterations for the linear quadratic control problem of discrete-time Markov jump linear systems, IEEE Transactions on Automatic Control, vol. 43, no 12, pp. 1727-1773, 1998.

[100] A. S. Willsky, B. C. Levy, Stochastic stability research for complex power systems, Lab. Inf. Decision Systems, Mass. Inst. Technology, report no. ET-76-C-01-2295, 1979.

[101] W. M. Wonham, Random Differential Equations in Control Theory, in: *"Probabilistic Methods in Applied Mathematics"*, Vol. 2, A. T. Bharucha-Reid, Ed. New York: Academic Press, 1971.

[102] J. Zabczyk, Controllability of stochastic linear systems, Systems and Control Letters, vol. 1, pp. 25-31, 1981.

On control problems for jump linear systems

Summary

In this book we consider the problems of controllability, stability and optimal control with quadratic index for discrete-time linear systems with randomly jumping parameters. In the analyzed model the parameters are functions of a Markov chain with finite state space.

First we study various concepts of controllability and deliberately illustrate the relationships between them. For all proposed types of controllability we present necessary and sufficient conditions as well as several methods of synthesis of control law that ensures reaching of required goal. A first impression, when we consider the problem of controllability for jump linear systems, may be to reduce it to a problem of controllability of linear systems with time-varying parameters. However, one important problem arise in this approach. When we consider deterministic time-varying systems and we want to find a control that drives certain initial conditions to a final state in given time then starting from the first moment we know values of all the parameters up to the final moment. Whereas for jump linear systems in each moment we know only the past values of coefficients and the future values could be predicated with given probability. This causes that for jump linear systems quite different approach must be used. The presented results significantly extend and complete the existing knowledge in the fild of controllability of jump linear systems.

Stability of jump linear systems is the next subject discussed in this book is. We focus on two types of stability: moment stability and almost sure stability. For one dimensional systems we present full description of both types of stability together with relationships between them. Such complete solution is nevertheless avaliable only for this class of systems. Next we present results on mean square stability. This special case of moment stability deserves special attention from the following two reasons. First, it is the only case of moment stability for which the necessary and sufficient conditions are known. Secondly, mean square stability is closely related to linear quadratic problem which is one of the most important optimization problems. It is also interesting that conditions for mean square stability can be expressed in terms of solutions of properly definite set of matrix linear equations. This set of equation called coupled Lyapunov equation is also investigated. Regarding almost sure stability, which is the most desirable from practical point of view, only partial results are available. We present several sufficient conditions, however only for special commuting structure of the matrix coefficients we can present necessary and sufficient conditions. Similar situation occurs for moment stability, i.e. in general, only sufficient conditions are known and some more specific results can be formulated under additional assumptions about commuting structure. We also discuss the Lyapunov exponent approach to stability problem. However, these results are purely theoretical unless methods for determining the sign of the Lyapunov exponent are developed.

The last problem discussed in this book is the problem of minimizing quadratic cost functional. It is called JLQ problem. The important difference between the results from the literature and those presented here is that we consider the situation when the coefficient of the systems depend also on time. We start with the JLQ problem on finite time interval. In this case the optimal control is given in the form of linear feedback with the feedback matrices depending on time and the state of Markow chain (the mode). The optimal feedback is given by a solution of a set of quadratic recurrent matrix equations. This set of equations is called recurrent coupled Riccati equation. Next we consider the situation of an infinite time interval. In the case the solution does not always exists. The existence of solution depends on the existence of a global and bounded solution of recurrent coupled Riccati equation. Therefore, next we investigate properties of this equation. If we consider the case when the coefficients of the system and cost functional does not explicitly depend on time the recurrent coupled Riccati equation changes into a set of algebraic quadratic matrix equations called coupled algebraic Riccati equation. Properties of this equation together with numerical algorithm of solving are also presented. We end our considerations with JLQ problem for jump linear system with additive disturbance. This problem is called noise JLQ problem. It is interesting that noise JLQ problem may have more than one solution. Basing on this property we show that for certain class of time varying systems the optimal control can be realized in the time invariant feedback form.

O problemach sterowania układami liniowymi ze skokowo zmieniającymi się parametrami

Streszczenie

W pracy omawia się zagadnienia sterowalności, stabilności i sterowania optymalnego z kwadratowym funkcjonałem kosztów dla dyskretnych układów liniowych ze skokowo zmieniającymi się parametrami.

W rozdziale 1 zebrano istniejące koncepcje sterowalności takich układów i zaproponowano pewne nowe definicje sterowalności. Rozważa się zarówno sterowalność w ustalonym czasie, jak i sterowalność w czasie losowym. Następnie przedyskutowano zależności między różnymi typami sterowalności i dla każdego z nich podano metody syntezy prawa sterowania zapewniającego osiągnięcie wymaganego celu. Wyniki tego rozdziału w pełni rozwiązują problem sterowalności dyskretnych układów liniowych ze skokowo zmieniającymi się parametrami.

Rozdział 2 poświęcony jest stabilności. Rozdział ten rozpoczyna się od wprowadzenia różnych typów sterowalności i dyskusji prostszych relacji między nimi. Następnie dla układów jednowymiarowych podane są warunki konieczne i wystarczające dla każdego typu stabilności i dokłady opis relacji między nimi. Jest to jedyna klasa układów, dla której taki kompletny opis udało się uzyskać. Stabilność średniokwadratowa została szczególnie wnikliwie opisana z dwóch powodów. Po pierwsze jest ona ściśle związana z jednym z najważniejszych zagadnień sterowania optymalnego, a mianowicie z problemem liniowo kwadratowym. Po drugie jest to jedyny typ stabilności, dla którego znane są efektywne warunki konieczne i wystarczające. Z punktu widzenia praktyki najbardziej pożądana jest stabilność z prawdopodobieństwem jeden. Niestety otrzymane wyniki nie rozwiązują w pełni tego problemu.

Rozdział 3 poświęcony jest problemowi sterowania optymalnego z kwadratowym wskaźnikiem jakości. W pierwszej części tego rozdziału przedstawiono znane w literaturze wyniki dotyczące przypadku sterowania na skończonym przedziale czasowym. Następnie przedstawiono nowe wyniki dotyczące nieskończonego horyzontu czasowego. Istotną nowością w porównaniu ze znanymi pracami jest rozpatrywanie sytuacji, w której zarówno współczynniki modelu, jak i wskaźnika jakości zależą od czasu. Rezulatаty te zostały osiągnięte poprzez analizę układu stowarzyszonych równań różnicowych Riccatiego.

Wydano za zgodą
Rektora Politechniki Śląskiej

Wydawnictwo Politechniki Śląskiej
ul. Akademicka 5, 44-100 Gliwice
tel./fax (032) 237-13-81
Dział Sprzedaży i Reklamy (032) 237-18-48

www.polsl.gliwice.pl/alma.mater/wps.html
wydawnictwo@polsl.gliwice.pl

Nakł. 100+50 Ark. wyd. 9 Ark. druk. 8,375 Papier offset. 70x100, 80 g
Oddano do druku 09.05.03 r. Podpis. do druku 09.05.03 r. Druk ukończ. w czerwcu 2003 r.

Druk wykonano w Zakładzie Graficznym Politechniki Śląskiej w Gliwicach, ul. Kujawska 1
zam. 179/03