ฟังก์ชันการแจกแจงภาวะการติดเชื้อเอดส์

นายรัชนิกร ชลไชยะ

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาคณิตศาสตร์ ภาควิชาคณิตศาสตร์

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2546

ISBN 974-17-4087-5

THE AIDS INCUBATION DISTRIBUTION FUNCTION

Mr. Ratchanikorn Chonchaiya

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of Master of Science in Mathematics

Department of Mathematics

Faculty of Science

Chulalongkorn University

Academic Year 2003

ISBN 974-17-4087-5

Thesis Title	The AIDS Incubation Distribution Function
By	Mr. Ratchanikorn Chonchaiya
Field of Study	Mathematics
Thesis Advisor	Associate Professor Kritsana Neammanee, Ph.D.
Thesis Co–advisor	Sajee Pianskool, Ph.D.

Accepted by the Faculty of Science, Chulalongkorn University in Partial Fulfillment of the Requirements for the Master's Degree.

.... Dean of the Faculty of Science

(Professor Piamsak Menasveta, Ph.D)

Thesis Committee

.... Chairman

(Assistant Professor Wicharn Lewkeeratiyutkul, Ph.D)

.... Thesis Advisor

(Associate Professor Kritsana Neammanee, Ph.D)

.... Thesis Co-advisor

(Sajee Pianskool, Ph.D)

.... Member

(Paisan Nakmahachalasint, Ph.D)

4572447323 : MAJOR MATHEMATICS
KEY WORDS : AIDS INCUBATION DISTRIBUTION FUNCTION/ MARKOV CHAIN
RATCHANIKORN CHONCHAIYA : AIDS INCUBATION DISTRIBUTION FUNCTION. THESIS ADVISOR : ASSOC. PROF. KRITSANA NEAMMANEE, Ph.D., THESIS CO–ADVISOR : SAJEE PIANSKOOL, Ph.D. 49 pp.
ISBN 974-17-4087-5

In this work, we derive the distribution functions of the AIDS incubation period, by considering two models which are more complicated than the ones usually studied. Model 1 is a Markovian model with probabilities of moving back and forth from one state to the next, with the possible onset of AIDS at any stages. Model 2, of a similar kind, has an added state before AIDS state. This follows a good deal of matrix algebra, ending with distribution function of the AIDS incubation period which are (different) weighted sums of Gamma distribution function for both models.

Department **Mathematics** Student's signature..............................

Field of study **Mathematics** Advisor's signature..............................

Academic year **2003** Co-advisor's signature.........................

รัชนิกร ชลไชยะ : ฟังก์ชันการแจกแจงภาวะการติดเชื้อเอดส์
(THE AIDS INCUBATION DISTRIBUTION FUNCTION)
อ.ที่ปรึกษา : รศ.ดร. กฤษณะ เนียมมณี, อ.ที่ปรึกษาร่วม : อ.ดร. ศจี เพียรสกุล
49 หน้า ISBN 974-17-4087-5

ในงานวิจัยนี้ เราได้หาฟังก์ชันการแจกแจงความน่าจะเป็นของช่วงเวลาของภาวะการติดเชื้อเอดส์ โดยพิจารณาจาก 2 ตัวแบบซึ่งเป็นตัวแบบซับซ้อนกว่าตัวแบบอื่นๆ ที่เคยศึกษาก่อนหน้านี้ ตัวแบบที่ 1 เป็นตัวแบบมาร์คอฟ ซึ่งมีโอกาสที่จะมีการเปลี่ยนแปลงแบบย้อนกลับสู่สภาวะการติดเชื้อก่อนหน้า และการเปลี่ยนแปลงแบบไปข้างหน้าสู่สภาวะเอดส์ ตัวแบบที่ 2 มีความคล้ายกับตัวแบบที่ 1 แต่เพิ่มสภาวะก่อนที่จะสู่ภาวะการติดเชื้อเอดส์ ในงานวิจัยนี้เราใช้วิธีการทางพีชคณิตเชิงเมทริกซ์ในการหาฟังก์ชันการแจกแจงภาวะการติดเชื้อเอดส์ซึ่งผลลัพธ์ที่ได้สำหรับทั้งสองตัวแบบ อยู่ในรูปของผลบวกของการแจกแจงแกมมา

ภาควิชา**คณิตศาสตร์**	ลายมือชื่อนิสิต..............................
สาขาวิชา**คณิตศาสตร์**	ลายมือชื่ออาจารย์ที่ปรึกษา.................
ปีการศึกษา **2546**	ลายมือชื่ออาจารย์ที่ปรึกษาร่วม.............

ACKNOWLEDGEMENTS

I would like to express my profound gratitude and deep appreciation to my thesis advisor Associate Professor Dr. Kritsana Neammanee and my thesis co–advisor Dr. Sajee Pianskool for their advice and encouragement. Sincere thanks and deep appreciation are also extended to Assistant Professor Dr. Wicharn Lewkeeratiyutkul, the chairman, and Dr. Paisan Nakmahachalasint, committee member, for their comments and suggestions. Also, I thanks all teachers who have taught me all along.

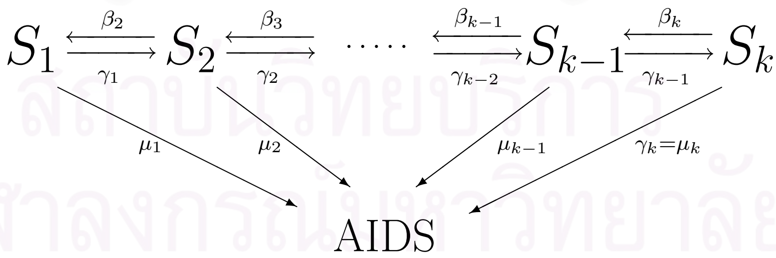
CONTENTS

CHAPTER I

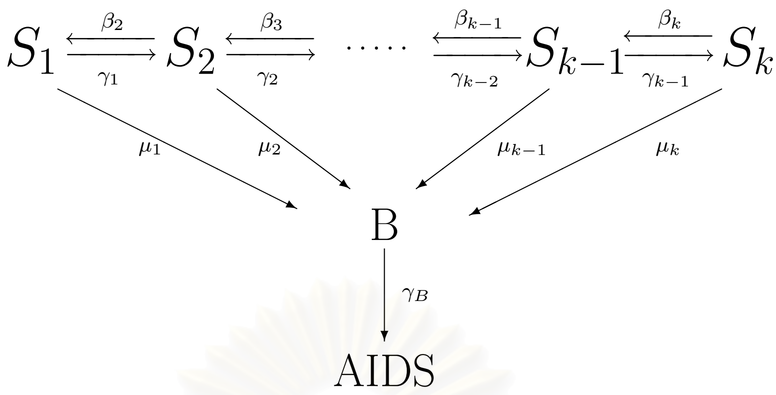
Introduction

In the studies of incubation of Acquired Immunodeficiency Syndrome (AIDS), an important problem is the characterization of the distribution function of the random time between infection with Human Immunodeficiency Virus (HIV) to AIDS onset. This distribution function has been referred to as the AIDS incubation distribution function. We also know that the mean value of this distribution function is usually very large taking a value of about 10 years for people between age 20-50 (Anderson et al., 1989).

In this study, we will derive the probability distributions in two models as follows.



Model 1



Model 2

The model 1 shows for the HIV epidemic in which β_i, $\gamma_i(> 0)$ and μ_i denote the transition rates. In this model, S_1 corresponds to the exposure stage of the Walter-Reed staging system (WRO stage, see Redfield and Burke, 1988) and is the HIV infected but antibody negative stage defined in Longini et al.(1989a, 1989b, 1991, 1992). In Tan and Hsu (1989), S_1 was referred to as the latent stage (L-stage) to account for the latency of HIV provirus. In model 1, AIDS denotes the AIDS stage and S_i denotes the $(i-1)$st substage of the infective stage $(i = 2, 3, \ldots, k)$. The Walter-Reed staging system assumed $k = 5$; Hethcote et al.(1991) assumed $k = 6$ whereas Longini et al.(1991) assumed $k = 7$ and $\beta_i = 0$ for $2 \leq i \leq k$ and if $\mu_j = 0$ for $1 \leq j \leq k-1$. Model 1 is more general than most of the models in the literature in the following two aspects:

(a) We assume that it is possible to have backward transition from S_i to S_{i-1}. The data reported by Nagelkerke et al.(1990) suggests that this is possible and hence should be taken into account.

(b) We assume also that it is possible to develop AIDS from any substages of infective stages, i.e., $S_i \rightarrow$ AIDS for $i \leq k$. The MACS data (Multicenter AIDS Cohort Studies) reported by Zhou et al.(1993) and the new revised

1993 AIDS defined by CDC (Center for Disease Control, Atlanta, GA) suggest that it is possible to have $S_i \rightarrow$ AIDS for $i < k$ and hence should be taken into account.

Notice that if $\beta_i = 0$ for $2 \leq i \leq k$ and if $\mu_j = 0$ for $1 \leq j \leq k-1$, then the model 1 reduces to the model considered by Longini et al.(1989a, 1989b, 1991, 1992).

Model 1 is assumed that AIDS can be developed directly from any substages of the infective stages. Since the average incubation period is usually very long, intuitively it is difficult to imagine that AIDS would be developed directly from early substages of the infective stage. Because of this consideration, we will postulate a state B for AIDS-related illness and consider a modified model 2. Notice that in the CDC staging system (see CDC Report, 1986), the stage B has been suggested as a stage between the infective stage and the AIDS stage. Hence the model 2 is in essence a reformulation of the CDC staging system of the HIV epidemic.

We will obtain the probability function in form $\displaystyle\sum_{j=0}^{k-1}\sum_{i=1}^{k} c_{ij} t^j e^{-b_i t}$ for both models.

CHAPTER II

Preliminaries

In our work, we need some basic knowledges from matrix algebra and probability theory as follows.

2.1 Basic knowledges in matrix algebra

A **diagonal matrix** is a square matrix in which the elements off the main diagonal are zero.

A square matrix A is **diagonalizable** if there is an invertible matrix P such that the matrix $P^{-1}AP$ is a diagonal matrix.

A square matrix A is said to be **nilpotent** if $A^r = 0$, where 0 is the zero matrix, for some natural number r.

A square matrix A is called **tri-diagonal** if one in which nonzero entries appear only on the main diagonal and the two adjacent diagonals, i.e., $a_{ij} = 0$ if $|i-j| > 1$.

Let A be an $n \times n$ matrix with real entries. The number λ (real or complex) is called an **eigenvalue** of A if there exists a nonzero vector $v \in \mathbb{C}^n$ such that

$$Av = \lambda v.$$

The vector $v \neq 0$ is called an **eigenvector** of A corresponding to the eigenvalue λ.

A square matrix A is said to be **positive definite** if all eigenvalues of A are positive.

Let A and A' be any two square matrices. We say that A is **similar** to A' if there exists an invertible matrix P such that $A' = P^{-1}AP$.

Theorem 2.1.1 Let A be a $k \times k$ matrix with eigenvalues $b_1, b_2, \ldots, b_k$. Then A is similar to a $k \times k$ matrix J which is of the form

$$J = \begin{bmatrix} J_1 & 0 & 0 & \cdots & 0 \\ 0 & J_2 & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & J_{n-1} & 0 \\ 0 & \ldots\ldots & & 0 & J_n \end{bmatrix} \tag{2.1}$$

for some $n \in \mathbb{N}$ and for $l = 1, 2, \ldots, n$, J_l is of the form

$$J_l = \begin{bmatrix} b_i & 1 & 0 & \cdots & 0 \\ 0 & b_i & 1 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & b_i & 1 \\ 0 & \ldots\ldots & & 0 & b_i \end{bmatrix}.$$

The matrix J is called the **Jordan canonical form corresponding to** A.

Furthermore, we know that the Jordan canonical form J of A can be represented in the form

$$\begin{bmatrix} b_1 & d_1 & 0 & \cdots & 0 \\ 0 & b_2 & d_2 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & b_{k-1} & d_{k-1} \\ 0 & \ldots\ldots & & 0 & b_k \end{bmatrix}$$

where $d_i \in \{0, 1\}$.

Proof See, e.g., [6] Chapter 7.

Let $P(t) = [p_{ij}(t)]$ be a matrix with variable t and assume that $p_{ij}(t)$ is differentiable for all i, j. Define

$$P'(t) = \left[\frac{d}{dt} p_{ij}(t)\right].$$

For any square matrix A of real numbers,

$$e^A = \sum_{j=0}^{\infty} \frac{1}{j!} A^j.$$

It is well known that e^A is well-defined for any square matrix A.

Proposition 2.1.2 Let X and Y be arbitrary $n \times n$ matrices. Then

1. $e^0 = I_n$, where 0 is the zero $n \times n$-matrix and I_n is the identity matrix,
2. if $XY = YX$, then $e^{X+Y} = e^X \cdot e^Y = e^Y \cdot e^X$,
3. if Y is invertible, then $e^{YXY^{-1}} = Ye^X Y^{-1}$,
4. $\frac{d}{dt} e^{tX} = Xe^{tX} = e^{tX} X$, for all $t \in \mathbb{R}$.

Theorem 2.1.3 Let A be a tri-diagonal $k \times k$ matrix of the form

$$A = \begin{bmatrix} \lambda_1 & -\gamma_1 & 0 & \dots\dots\dots & & & 0 \\ -\beta_2 & \lambda_2 & -\gamma_2 & 0 & \dots\dots & & 0 \\ 0 & -\beta_3 & \lambda_3 & -\gamma_3 & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots\dots & & 0 & -\beta_{k-1} & \lambda_{k-1} & -\gamma_{k-1} \\ 0 & \dots\dots\dots & & & 0 & -\beta_k & \lambda_k \end{bmatrix},$$

where $\beta_i \geq 0$, $\gamma_i > 0$ and $\lambda_i \geq \gamma_i + \beta_i$ for $i = 1, 2, \dots, k$. Then A has the positive determinant.

Proof We prove this theorem by induction. Let $P(n)$ be the statement := An

$n \times n$ matrix of the form

$$A_n = \begin{bmatrix} \lambda_1 & -\gamma_1 & 0 & \dots\dots\dots & & 0 \\ -\beta_2 & \lambda_2 & -\gamma_2 & 0 & \dots\dots & 0 \\ 0 & -\beta_3 & \lambda_3 & -\gamma_3 & 0 \quad \cdots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots \quad \ddots & \vdots \\ 0 & \dots\dots & 0 & -\beta_{n-1} & \lambda_{n-1} & -\gamma_{n-1} \\ 0 & \dots\dots & & 0 & -\beta_n & \lambda_n \end{bmatrix},$$

where $\beta_i \geq 0$, $\gamma_i > 0$ and $\lambda_i = \gamma_i + \beta_i + \epsilon_i$ for some $\epsilon_i \geq 0$ for $i = 1, 2, \dots, n$ has a positive determinant.

Basis Step. Since $\det([\lambda_1]) = \lambda_1 > 0$, $P(1)$ is true.

Inductive Step. Let $q \in \mathbb{N}$ be such that $P(1), P(2), \dots, P(q)$ are true. To prove $P(q+1)$, we need the fact that

$$\det A_l \geq (\beta_1 + \epsilon_1) \begin{vmatrix} \lambda_2 & -\gamma_2 & 0 & \dots\dots & & 0 \\ -\beta_3 & \lambda_3 & -\gamma_3 & 0 & \dots\dots & 0 \\ \vdots & & & & & \vdots \\ 0 & \dots\dots & & -\beta_{l-1} & \lambda_{l-1} & -\gamma_{l-1} \\ 0 & \dots\dots & & 0 & -\beta_l & \lambda_l \end{vmatrix} + \gamma_1\gamma_2\dots\gamma_l$$

for $l = 2, 3, \dots, q+1$. (2.2)

We will prove (2.2) by finite induction on $\{2, 3, \dots, q+1\}$. Observe that

$$\det A_2 = \det \begin{bmatrix} \lambda_1 & -\gamma_1 \\ -\beta_2 & \lambda_2 \end{bmatrix} = (\gamma_1 + \beta_1 + \epsilon_1)\lambda_2 - \gamma_1\beta_2$$

$$\geq (\beta_1 + \epsilon_1)\lambda_2 + \gamma_1\gamma_2$$

which implies that (2.2) holds for $l = 2$. Next, suppose that (2.2) is true for $l = 2, 3, \dots, m$ where $m \in \{2, 3, \dots, q\}$. Then

$$\det\ A_{m+1} = (\gamma_1+\beta_1+\epsilon_1)\begin{vmatrix} \lambda_2 & -\gamma_2 & 0 & \dots\dots\dots\dots & & 0 \\ -\beta_3 & \lambda_3 & -\gamma_3 & 0 & \dots\dots\dots & 0 \\ \vdots & & & & & \vdots \\ 0 & \dots\dots\dots & & -\beta_m & \lambda_m & -\gamma_m \\ 0 & \dots\dots\dots & & 0 & -\beta_{m+1} & \lambda_{m+1} \end{vmatrix}$$

$$+ \quad \gamma_1\begin{vmatrix} -\beta_2 & -\gamma_2 & 0 & \dots\dots\dots\dots & & 0 \\ 0 & \lambda_3 & -\gamma_3 & 0 & \dots\dots\dots & 0 \\ \vdots & & & & & \vdots \\ 0 & \dots\dots\dots & & -\beta_m & \lambda_m & -\gamma_m \\ 0 & \dots\dots\dots & & 0 & -\beta_{m+1} & \lambda_{m+1} \end{vmatrix}$$

$$= (\beta_1+\epsilon_1)\begin{vmatrix} \lambda_2 & -\gamma_2 & 0 & \dots\dots\dots\dots & & 0 \\ -\beta_3 & \lambda_3 & -\gamma_3 & 0 & \dots\dots\dots & 0 \\ \vdots & & & & & \vdots \\ 0 & \dots\dots\dots & & -\beta_m & \lambda_m & -\gamma_m \\ 0 & \dots\dots\dots & & 0 & -\beta_{m+1} & \lambda_{m+1} \end{vmatrix}$$

$$+ \quad \gamma_1\begin{vmatrix} \lambda_2 & -\gamma_2 & 0 & \dots\dots\dots\dots & & 0 \\ -\beta_3 & \lambda_3 & -\gamma_3 & 0 & \dots\dots\dots & 0 \\ \vdots & & & & & \vdots \\ 0 & \dots\dots\dots & & -\beta_m & \lambda_m & -\gamma_m \\ 0 & \dots\dots\dots & & 0 & -\beta_{m+1} & \lambda_{m+1} \end{vmatrix}$$

$$
- \quad \gamma_1\beta_2 \begin{vmatrix} \lambda_3 & -\gamma_3 & 0 & \dots\dots\dots\dots & & 0 \\ -\beta_4 & \lambda_4 & -\gamma_4 & 0 & \dots\dots\dots & 0 \\ \vdots & & & & & \vdots \\ 0 & \dots\dots\dots & & -\beta_m & \lambda_m & -\gamma_m \\ 0 & \dots\dots\dots & & 0 & -\beta_{m+1} & \lambda_{m+1} \end{vmatrix}
$$

$$
\geq \quad (\beta_1+\epsilon_1) \begin{vmatrix} \lambda_2 & -\gamma_2 & 0 & \dots\dots\dots\dots & & 0 \\ -\beta_3 & \lambda_3 & -\gamma_3 & 0 & \dots\dots\dots & 0 \\ \vdots & & & & & \vdots \\ 0 & \dots\dots\dots & & -\beta_m & \lambda_m & -\gamma_m \\ 0 & \dots\dots\dots & & 0 & -\beta_{m+1} & \lambda_{m+1} \end{vmatrix}
$$

$$
+\gamma_1 \left((\beta_2+\epsilon_2) \begin{vmatrix} \lambda_3 & -\gamma_3 & 0 & \dots\dots\dots\dots & & 0 \\ -\beta_4 & \lambda_4 & -\gamma_4 & 0 & \dots\dots\dots & 0 \\ \vdots & & & & & \vdots \\ 0 & \dots\dots\dots & & -\beta_m & \lambda_m & -\gamma_m \\ 0 & \dots\dots\dots & & 0 & -\beta_{m+1} & \lambda_{m+1} \end{vmatrix} + \gamma_2\gamma_3\cdots\gamma_{m+1} \right)
$$

$$
- \quad \gamma_1\beta_2 \begin{vmatrix} \lambda_3 & -\gamma_3 & 0 & \dots\dots\dots\dots & & 0 \\ -\beta_4 & \lambda_4 & -\gamma_4 & 0 & \dots\dots\dots & 0 \\ \vdots & & & & & \vdots \\ 0 & \dots\dots\dots & & -\beta_m & \lambda_m & -\gamma_m \\ 0 & \dots\dots\dots & & 0 & -\beta_{m+1} & \lambda_{m+1} \end{vmatrix}
$$

$$\geq (\beta_1+\epsilon_1)\begin{vmatrix} \lambda_2 & -\gamma_2 & 0 & \dots\dots\dots\dots & & 0 \\ -\beta_3 & \lambda_3 & -\gamma_3 & 0 & \dots\dots\dots & 0 \\ \vdots & & & & & \vdots \\ 0 & \dots\dots\dots & & -\beta_m & \lambda_m & -\gamma_m \\ 0 & \dots\dots\dots & & 0 & -\beta_{m+1} & \lambda_{m+1} \end{vmatrix} + \gamma_1\gamma_2\gamma_3\cdots\gamma_{m+1}$$

where we have used the truth of $P(m-1)$ in the last inequality. Thus (2.2) is proved.

By (2.2),

$$\det A_{q+1} \geq (\beta_1+\epsilon_1)\begin{vmatrix} \lambda_2 & -\gamma_2 & 0 & \dots\dots\dots\dots & & 0 \\ -\beta_3 & \lambda_3 & -\gamma_3 & 0 & \dots\dots\dots & 0 \\ \vdots & & & & & \vdots \\ 0 & \dots\dots\dots & & -\beta_q & \lambda_q & -\gamma_q \\ 0 & \dots\dots\dots & & 0 & -\beta_{q+1} & \lambda_{q+1} \end{vmatrix}$$

$$+\gamma_1\gamma_2\dots\gamma_{q+1}$$

$$\geq \gamma_1\gamma_2\dots\gamma_{q+1}$$

$$> 0$$

where we have used the fact that $P(q)$ holds in the second inequality. Hence, by the induction hypothesis, A_{q+1} has a positive determinant. Then the theorem is proved. #

Theorem 2.1.4 If $P(t)$ and $Q(t)$ are matrices having differentiable entries, then

$$\frac{d}{dt}(P(t)Q(t)) = \frac{d}{dt}(P(t))\,Q(t) + P(t)\frac{d}{dt}(Q(t)).$$

2.2 Basic knowledges in Probability Theory

A **probability space** is a measure space $(\Omega, \Im, P)$ in which P is a positive measure such that $P(\Omega) = 1$. The set Ω will be referred to as a **sample space**. The element of $\Im$ are called **events.** For any event A, the value $P(A)$ is called the **probability of A**.

A real-valued function X from a probability space $(\Omega, \Im, P)$ is said to be a **random variable** if for each $x \in \mathbb{R}$, $\{\omega \in \Omega \mid X(\omega) \leq x\} \in \Im$.

Let X be a random variable on a probability space $(\Omega, \Im, P)$. The **expected value** $E(X)$ is defined by

$$E(X) = \int_\Omega X dP.$$

We often refer to the expected value as the **mean** or **average** or **first moment of** X.

For a random variable X, we define $F : \mathbb{R} \to [0, 1]$ as the **distribution function of** X by

$$F(x) = P(\{w |\ X(w) \leq x\}).$$

It is convenient to abbreviate this expression by

$$F(x) = P(X \leq x).$$

If there exists $f : \mathbb{R} \to [0, \infty)$ such that the distribution function F of X can be represented in the form of

$$F(x) = \int_{-\infty}^{x} f(t)dt,$$

we will say that X is a **continuous random variable** and f is the **probability function** of X.

A continuous random variable X will be said to have a **gamma distribution**

function with parameters (α, β) if its probability function of X is given by

$$f(x) = \begin{cases} 0, & \text{if } x \leq 0, \\ \dfrac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}}, & \text{if } x > 0, \end{cases}$$

where α and β are positive constants, and the Gamma function $\Gamma : \mathbb{R}^+ \to \mathbb{R}$ is defined by

$$\Gamma(r) = \int_0^{\infty} x^{r-1} e^{-x} dx \qquad \text{for all } r \in \mathbb{R}^+.$$

A **stochastic process** is a family of random variables $\{X_t |\ t \in I\}$, where t is a **parameter** running over an index set I and a probability space $(\Omega, \mathcal{F}, P)$, which take values in a measurable space $(S, \mathcal{S})$, called the state space. I is called a **parameter space**.

If I is a countable set, then a stochastic process $\{X_t |\ t \in I\}$ is said to be a **discrete stochastic process** and in the case that I is an interval, $\{X_t |\ t \in I\}$ is said to be a **continuous stochastic process**.

In this research, we assume that every Markov process has a finite state space.

A continuous stochastic process $\{X_t \mid t \in [0, \infty)\}$ is said to be a **Markov process** if for all $s, t, u \geq 0$ such that $u \leq s$ and $i, j, k \in S$

$$P\Big(\{X_{s+t} = j \mid X_s = i, X_u = k\}\Big) = P\Big(\{X_{s+t} = j \mid X_s = i\}\Big).$$

Let $\{X_t \mid t \in [0, \infty)\}$ be a Markov process with a state space S. For any $s, t \geq 0$ and $i, j \in S$, we call a real number $p_{ij}^{(s)}(t)$, defined by

$$p_{ij}^{(s)}(t) = P(X_{s+t} = j \mid X_s = i),$$

the **t-step transition probability at time *s*** and call the matrix

$$P^{(s)}(t) = [p_{ij}^{(s)}(t)]_{i,j \in S}$$

the ***t*-step transition matrix at time *s*** .

A state i of a Markov process is called **absorbing** if it is imposible to leave it (i.e., $p_{ii}^{(s)}(t) = 1$ for every $s, t \geq 0$). A Markov process is **absorbing** if it has at least one absorbing state, and if from every state it is possible to go to an absorbing state (not necessarily in one step).

In an absorbing Markov process, a state which is not absorbing is called **transient**.

Let $t > 0$. If all $p_{ij}^{(s)}(t)$ are equal for every $s > 0$, then $\{X_t \mid t \in [0, \infty)\}$ is said to be a **homogeneous Markov process** or **Markov process with stationary transition probability**. We also denote $p_{ij}^{(s)}(t)$ and $P^{(s)}(t)$ by $p_{ij}(t)$ and $P(t)$, respectively.

Let $\{X_t \mid t \in [0, \infty)\}$ be a homogeneous continuous–time Markov process with state space S and transition matrix $[p_{ij}(t)]$.

The matrix $Q = [v_{ij}]_{i,j \in S}$ which satisfies the following conditions:

1. for $i \neq j$, the entry v_{ij} = rate at which X_t jumps from i to j,
   i.e., $p_{ij}(t) = v_{ij}t + o(t) \quad$ for $t \to 0$ where $\lim\limits_{t \to 0} \dfrac{o(t)}{t} = 0$,
2. $v_{ii} = -\sum\limits_{j \neq i} v_{ij}$ for each i,

is called a **generator matrix** or a **rate matrix** of $\{X_t \mid t \in [0, \infty)\}$.

Note that a generator matrix Q satisfies the following properties:

1. $v_{ij} \geq 0$ for $i \neq j$,
2. $\sum\limits_{j \in S} v_{ij} = 0$ for $i \in S$.

For a homogeneous Markov process $\{X_t \mid t \in [0, \infty)\}$ and $i, k \in S$, if $i \neq k$, the **first passage time** of i is defined to be

$$T_{ik} = \inf\{t \geq 0 : X_t = k \mid X_0 = i\};$$

the **mean first passage time** is

$$\mu_{ik} = E(T_{ik}).$$

Proposition 2.2.1 Let $\{X_t \mid t \in [0,\infty)\}$ be a homogeneous Markov process with finite state space. Then

$$P(t) = e^{Qt},$$

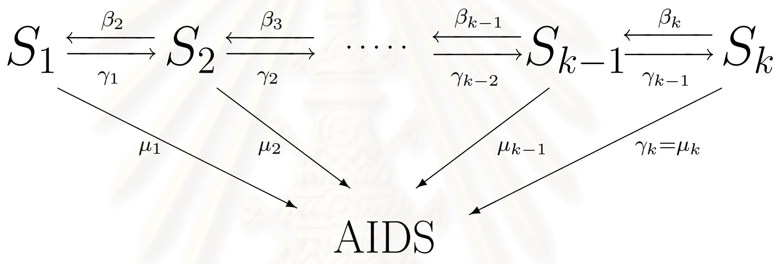
where P and Q are the transition and generator matrices of $\{X_t \mid t \in [0,\infty)\}$, respectively.

Proof See, e.g., [17] page 388.

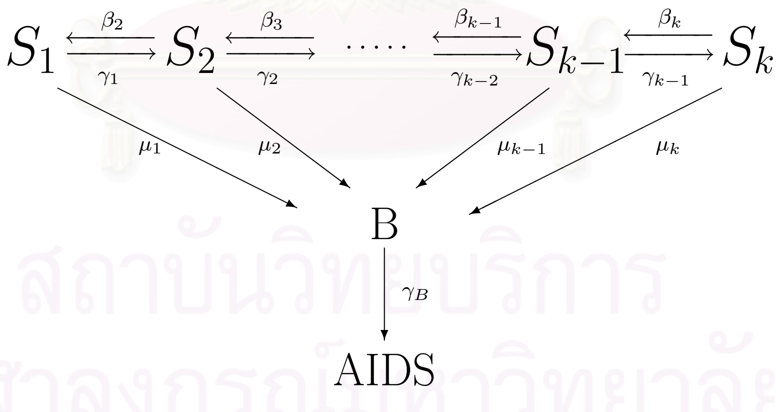
CHAPTER III

Main Results

The purpose of this chapter is to find the probability function of the first passage time of each state of AIDS incubation, i.e., the HIV incubation function. We will consider the HIV epidemics in two models as follows.



Model 1



Model 2

The given models are for the HIV epidemic in which β_i is the backward transition rate from S_i to S_{i-1}, $\gamma_i > 0$ is the forward transition rate from S_i to S_{i+1} and μ_i is the transition rate from S_i to AIDS. In model 1, AIDS denotes the AIDS stage, S_1 is the HIV infected but antibody negative stage and S_i($i = 2, 3, \ldots, k$) denotes the $(i-1)$st substage of the infective stage.

Model 1 is more general than most of the models in the literature in the following two aspects :

(a) We assume that it is possible to have backward transition from S_i to S_{i-1} $(i = 2, 3, \ldots, k)$.

(b) We assume also that it is possible to develop AIDS from any substage of infective stages, i.e., $S_i \to$ AIDS for $i \leq k$.

Model 1 is assumed that AIDS can be developed directly from any substage of the infective stages. Since the average incubation period is usually very long, intuitively it is difficult to imagine that AIDS would be developed directly from early substages of the infective stages. Because of this consideration, we will postulate a state B for AIDS-related illness and consider a modified model 2. Notice that in the CDC staging system (see CDC Report, 1986), the stage B has been suggested as a stage between the infective stages and the AIDS stage. Hence model 2 is in essence a reformulation of the CDC staging system of the HIV epidemic.

3.1 Proof of Model 1

From model 1, for every $t > 0$, we let X_t be a random variable of which the value is the state of HIV epidemic at the time t. So the state space of $\{X_t \mid t \geq 0\}$ is $\{S_1, S_2, \ldots, S_{k+1}\}$ where S_{k+1} is the AIDS state. Hence the $(k+1)$st stage is an absorbing state and by the fact that $\gamma_i > 0$ for $i = 1, 2, \ldots, k$, we see that the states $S_1, S_2, \ldots, S_k$ are transition states. Starting at time $s = 0$, let T_i be the random time that S_i is absorbed into S_{k+1}, $i = 1, 2, \ldots, k$. Then T_i is referred to as the first passage time of S_i and $f_i(t)$, the probability function, the first passage probability of S_i. In what follows, we put

$$f(t) = [f_1(t), \ldots, f_k(t)]^T,$$

where T denotes the transpose operator.

In this work, we assume that $\{X_t | t \geq 0\}$ is a homogeneous continuous Markov process and let the transition matrix $P(t) = [p_{ij}(t)]$. By Chapter 5 of [17] we have that:

$$\left.\begin{aligned} p_{ij}(t) &= v_{ij}t + o(t) \quad \text{where } t \to 0 \text{ and } i \neq j, \\ p_{ii}(t) &= 1 - v_{ii}t + o(t) \quad \text{where } t \to 0, \\ v_{ij} &\geq 0, \quad v_{ii} = \sum_{j \neq i} v_{ij}, \end{aligned}\right\} \tag{3.1}$$

where v_{ij} is the transition rate at which X_t jumps from i to j.

From (3.1), we know that the generator matrix of (X_t) is of the form

$$\begin{bmatrix} -v_{11} & \ldots & v_{1k} & v_{1(k+1)} \\ \vdots & \ddots & \vdots & \vdots \\ v_{k1} & \ldots & -v_{kk} & v_{k(k+1)} \\ v_{(k+1)1} & \ldots & v_{(k+1)k} & -v_{(k+1)(k+1)} \end{bmatrix}.$$

Remark From model 1, for every $i, j = 1, 2, \ldots, k$, we see that $v_{ij} = 0$ for $|i - j| > 1$, $v_{i(i+1)} = \gamma_i$, $v_{i(k+1)} = \mu_i$ and $v_{i(i-1)} = \beta_i$ $(i \neq 1)$. Since the $(k+1)$st state is an absorbing state, $v_{(k+1)j} = 0$ for $j = 1, 2, \ldots, k$. So the generator matrix of $\{X_t \mid t \geq 0\}$ is $-Q$ where

$$Q = \begin{bmatrix} \lambda_1 & -\gamma_1 & \dots & 0 & 0 & -\mu_1 \\ -\beta_2 & \lambda_2 & -\gamma_2 & \dots & 0 & -\mu_2 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda_{k-1} & -\gamma_{k-1} & -\mu_{k-1} \\ 0 & 0 & \dots & -\beta_k & \lambda_k & -\gamma_k \\ 0 & 0 & \dots & 0 & 0 & 0 \end{bmatrix} \tag{3.2}$$

Proposition 3.1 Let $\lambda_i = \beta_i + \gamma_i + \mu_i$ for $i = 1, 2, \dots, k-1$, $\lambda_k = \beta_k + \gamma_k$ and $\beta_1 = 0$. Let

$$A = \begin{bmatrix} \lambda_1 & -\gamma_1 & \dots & 0 & 0 \\ -\beta_2 & \lambda_2 & -\gamma_2 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_{k-1} & -\gamma_{k-1} \\ 0 & 0 & \dots & -\beta_k & \lambda_k \end{bmatrix}.$$

If A is invertible, then

$$f(t) = e^{-At} A 1_k.$$

Proof

We observe that

$$Q = \begin{bmatrix} A & -\mu \\ 0_k & 0 \end{bmatrix}_{(k+1)\times(k+1)}$$

where $\mu = [\mu_1, \mu_2, \dots, \mu_k]^T$ and $0_k = \underbrace{(0, 0, \cdots, 0)}_{k}$,

$$Q^2 = Q \times Q = \begin{bmatrix} A^2 & -A\mu \\ 0_k & 0 \end{bmatrix}$$

and

$$Q^3 = Q^2 \times Q = \begin{bmatrix} A^3 & -A^2\mu \\ 0_k & 0 \end{bmatrix}.$$

Hence, by inductive reasoning, we have that

$$Q^m = \begin{bmatrix} A^m & -A^{m-1}\mu \\ 0_k & 0 \end{bmatrix}$$

where $m = 1, 2, \ldots$.

By the fact that

$$\begin{aligned} A^{-1}\{I_k - e^{-At}\}\mu &= -A^{-1}\{e^{-At} - I_k\}\mu \\ &= -A^{-1}\{\sum_{m=0}^{\infty}(-1)^m \frac{t^m}{m!}A^m - I_k\}\mu \\ &= -\sum_{m=1}^{\infty}(-1)^m \frac{t^m}{m!}A^{m-1}\mu, \end{aligned}$$

where $Q^0 = I_{k+1}$, and by Proposition 2.2.1, we have

$$\begin{aligned} P(t) &= e^{-Qt} \\ &= \sum_{j=0}^{\infty}(-1)^j \frac{t^j}{j!}Q^j \\ &= I_{k+1} + \sum_{j=1}^{\infty}(-1)^j \frac{t^j}{j!}\begin{bmatrix} A^j & -A^{j-1}\mu \\ 0_k & 0 \end{bmatrix} \\ &= I_{k+1} + \begin{bmatrix} \sum_{j=1}^{\infty}(-1)^j \frac{t^j}{j!}A^j & -\sum_{j=1}^{\infty}(-1)^j \frac{t^j}{j!}A^{j-1}\mu \\ 0_k & 0 \end{bmatrix} \\ &= I_{k+1} + \begin{bmatrix} e^{-At} - I_k & A^{-1}[I_k - e^{-At}]\mu \\ 0_k & 0 \end{bmatrix}, \end{aligned}$$

i.e.,

$$P(t) = \begin{bmatrix} e^{-At} & A^{-1}[I_k - e^{-At}]\mu \\ 0_k & 1 \end{bmatrix}. \tag{3.3}$$

Let

$$h^{(t)} = \begin{bmatrix} p_{1(k+1)}(t) & p_{2(k+1)}(t) & \cdots & \cdots & p_{k(k+1)}(t) \end{bmatrix}^T.$$

Hence, by (3.3) we have

$$h^{(t)} = A^{-1}[I_k - e^{-At}]\mu. \tag{3.4}$$

For $t > 0$ and $i = 1, 2, \ldots, k$ we see that

$[0, ..., 0, \underbrace{1}_{i^{th}}, 0, ..., 0](h^{(t)})$

$= p_{i(k+1)}(t)$

$= P(\{X_t = k+1 \mid X_0 = i\})$

$= P(\{X_t = k+1, T_i > t \mid X_0 = i\}) + P(\{X_t = k+1, T_i \leq t \mid X_0 = i\})$

$= P(\{X_t = k+1, T_i \leq t \mid X_0 = i\})$

$= P($getting into the $(k+1)$st stage when we start at the state $i,$ we spend time less than or equal to $t)$

$= P(T_i \leq t)$

$= F_i(t)$

where $F_i(t)$ is the distribution function of T_i. If $\frac{d}{dt}[0, ..., 0, \underbrace{1}_{i^{th}}, 0, ..., 0](h^{(t)})$ exists for every i, then

$$f(t) = \frac{d}{dt}(h^{(t)}). \tag{3.5}$$

By (3.4),(3.5) and the fact that

$$\begin{aligned} A1_k &= \begin{bmatrix} \lambda_1 & -\gamma_1 & 0 & \ldots\ldots\ldots & 0 \\ -\beta_2 & \lambda_2 & -\gamma_2 & \ldots\ldots\ldots & 0 \\ 0 & -\beta_3 & \lambda_3 & \ldots\ldots\ldots & 0 \\ \vdots & & & & \vdots \\ 0 & \ldots\ldots\ldots\ldots & 0 & -\beta_k & \lambda_k \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ \vdots \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} \mu_1 & \mu_2 & \cdots & \mu_k \end{bmatrix}^T \\ &= \mu, \end{aligned} \tag{3.6}$$

we have,

$$\begin{aligned} f(t) &= \frac{d}{dt}[A^{-1}[I_k - e^{-At}]\mu] \\ &= \frac{d}{dt}[A^{-1}[I_k - e^{-At}]A1_k] \\ &= \frac{d}{dt}[A^{-1} \cdot A \cdot 1_k - A^{-1}e^{-At} \cdot A \cdot 1_k] \\ &= -\frac{d}{dt}[e^{-At} \cdot 1_k] \\ &= e^{-At} \cdot A \cdot 1_k \ \text{ (by Theorem 2.1.2(4)).} \end{aligned}$$

#

Next, we will consider the matrix A defined in Proposition 3.1. By Theorem 2.1.3, we see that A has a positive determinant and then A^{-1} exists. By Proposition 3.1 we have

$$f(t) = e^{-At}A1_k. \tag{3.7}$$

From Theorem 2.1.1, there exist a $k \times k$ Jordan canonical form J of A such that

$$J = \begin{bmatrix} b_1 & d_1 & 0 & \dots\dots\dots & & 0 \\ 0 & b_2 & d_2 & 0 & \cdots & 0 \\ 0 & 0 & b_3 & d_3 & \cdots & 0 \\ \vdots & & & \ddots & \ddots & \vdots \\ 0 & \dots\dots & & 0 & b_{k-1} & d_{k-1} \\ 0 & \dots\dots\dots\dots & & & 0 & b_k \end{bmatrix} \tag{3.8}$$

where $b_1, \dots, b_k$ are eigenvalues of A and $d_i \in \{0, 1\}$ for $i = 1, \dots, k$ and an invertible matrix P such that

$$A = PJP^{-1}.$$

Moreover, we know that there exists $n \in \mathbb{N}$ such that J can be written in the form,

$$\begin{bmatrix} J_1 & 0 & 0 & \dots & 0 \\ 0 & J_2 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & J_{n-1} & 0 \\ 0 & 0 & \dots & 0 & J_n \end{bmatrix} \tag{3.9}$$

where J_i's are submatrices of J defined in (2.1). For each $l = 1, 2, \dots, n$, if

$$J_l = \begin{bmatrix} b_i & 1 & 0 & \cdots & 0 \\ 0 & b_i & 1 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & b_i & 1 \\ 0 & \dots\dots\dots & & 0 & b_i \end{bmatrix},$$

let

$$S_l = \begin{bmatrix} b_i & 0 & 0 & \cdots & 0 \\ 0 & b_i & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & b_i & 0 \\ 0 & \ldots\ldots\ldots & & 0 & b_i \end{bmatrix} \quad \text{and} \quad N_l = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 1 \\ 0 & \ldots\ldots\ldots & & 0 & 0 \end{bmatrix},$$

and if $J_l = [b_i]$, let

$$S_l = [b_i] \qquad \text{and} \qquad N_l = [0].$$

Then

$$J_l = S_l + N_l \quad \text{and} \quad S_l N_l = N_l S_l \quad \text{for all } l = 1, 2, \ldots, n. \tag{3.10}$$

Let

$$N = \begin{bmatrix} N_1 & 0 & 0 & \cdots & 0 \\ 0 & N_2 & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & N_{n-1} & 0 \\ 0 & \ldots\ldots\ldots & & 0 & N_n \end{bmatrix} = \begin{bmatrix} 0 & d_1 & 0 & \cdots & 0 \\ 0 & 0 & d_2 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & d_{k-1} \\ 0 & \ldots\ldots\ldots & & 0 & 0 \end{bmatrix}$$

and

$$S = \begin{bmatrix} S_1 & 0 & 0 & \cdots & 0 \\ 0 & S_2 & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & S_{n-1} & 0 \\ 0 & \ldots\ldots\ldots & & 0 & S_n \end{bmatrix} = \begin{bmatrix} b_1 & 0 & 0 & \cdots & 0 \\ 0 & b_2 & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & b_{k-1} & 0 \\ 0 & \ldots\ldots\ldots & & 0 & b_k \end{bmatrix}.$$

From (3.8), (3.9) and (3.10), we see that

$$J = N + S \qquad \text{and} \qquad NS = SN.$$

Let

$$\overline{N} = PNP^{-1} \qquad \text{and} \qquad \overline{S} = PSP^{-1}.$$

Hence

$$\overline{N} + \overline{S} = P(N+S)P^{-1} = PJP^{-1} = A \tag{3.11}$$

and

$$\overline{N}\,\overline{S} = P(NS)P^{-1} = P(SN)P^{-1} = \overline{S}\,\overline{N} \tag{3.12}$$

Since

$$S^n = \begin{bmatrix} b_1^n & 0 & \dots & 0 & 0 \\ 0 & b_2^n & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & b_{k-1}^n & 0 \\ 0 & 0 & \dots & 0 & b_k^n \end{bmatrix},$$

we have

$$\begin{aligned} e^{-St} &= \sum_{n=0}^{\infty} \frac{(-St)^n}{n!} \\ &= \begin{bmatrix} e^{-b_1 t} & 0 & \cdots & 0 \\ 0 & e^{-b_2 t} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots\dots\dots & & e^{-b_k t} \end{bmatrix}. \end{aligned}$$

Observe that

$$N^2 = \begin{bmatrix} 0 & 0 & d_1 d_2 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 0 & d_2 d_3 & \cdots & \cdots & 0 \\ \vdots & & & & & \ddots & \vdots \\ 0 & 0 & \dots & \dots & \dots & \dots & d_{k-1} d_k \\ 0 & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & 0 & \dots & \dots & \dots & \dots & 0 \end{bmatrix}$$

and by using mathematical induction, for every m we can see that $N^m = \left[n_{ij}^{(m)}\right]$ where

$$n_{ij}^{(m)} = \begin{cases} \prod_{l=i}^{i+m-1} d_l & \text{if } j = i+m \text{ and } j < k, \\ 0 & \text{otherwise.} \end{cases}$$

Hence $N^m = 0$ for $m \geq k$ and

$$e^{-Nt} = \sum_{m=0}^{k-1} \frac{(-tN)^m}{m!} = \sum_{m=0}^{k-1} \frac{(-t)^m (n_{ij}^{(m)})}{m!}$$

$$= I_k - t \begin{bmatrix} 0 & d_1 & 0 & \dots & 0 \\ 0 & 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & d_{k-1} \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix} + \frac{t^2}{2!} \begin{bmatrix} 0 & 0 & d_1 d_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & d_{k-2} d_{k-1} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$+ \dots + \frac{(-t)^{k-1}}{(k-1)!} \begin{bmatrix} 0 & 0 & 0 & \dots & \prod_{i=1}^{k-1} d_i \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}$$

$:= [n_{ij}]$ where

$$n_{ij} = \begin{cases} n_{ij}^{(p)} \dfrac{(-t)^p}{p!} & \text{if } j > i \text{ and } j = i+p, \\ 1 & \text{if } j = i, \\ 0 & \text{if } j < i, \end{cases}$$

i.e.,

$$e^{-Nt} = \begin{bmatrix} 1 & -n_{12}^{(1)}t & n_{13}^{(2)}\frac{t^2}{2} & \cdots\cdots & n_{1k}^{(k-1)}\frac{(-t)^{k-1}}{(k-1)!} \\ 0 & 1 & -n_{23}^{(1)}t & \cdots\cdots & n_{2k}^{(k-2)}\frac{(-t)^{k-2}}{(k-2)!} \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots \ 1 & -n_{(k-1)k}^{(1)}t \\ 0 & 0 & 0 & \cdots \ 0 & 1 \end{bmatrix}.$$

By (3.6) and (3.7), for $i_0 \in \{1, 2, \ldots, k\}$ we have

$$\begin{aligned}
f_{i_0}(t) &= e_{i_0}^T f(t) \quad \text{where } e_{i_0}^T = [0, \ldots, 0, \underbrace{1}_{i_0^{\text{th}}}, 0, \ldots, 0] \\
&= e_{i_0}^T e^{-At} A 1_k \\
&= e_{i_0}^T e^{-At} \begin{bmatrix} \mu_1 & \mu_2 & \cdots & \mu_{k-1} & \mu_k \end{bmatrix}^T ; \quad \text{which by (3.11) and (3.12)} \\
&= e_{i_0}^T e^{-\overline{S}t} e^{-\overline{N}t} \begin{bmatrix} \mu_1 & \mu_2 & \cdots & \mu_k \end{bmatrix}^T ; \quad \text{which by Proposition 2.1.2(3)} \\
&= e_{i_0}^T P e^{-St} e^{-Nt} P^{-1} \begin{bmatrix} \mu_1 & \mu_2 & \cdots & \mu_k \end{bmatrix}^T \\
&= \begin{bmatrix} e^{-b_1 t} v_{i_0 1} & e^{-b_2 t} v_{i_0 2} & \cdots & e^{-b_k t} v_{i_0 k} \end{bmatrix} [n_{ij}] \begin{bmatrix} \sum_{i=1}^{k} v'_{1i}\mu_i & \sum_{i=1}^{k} v'_{2i}\mu_i & \cdots & \sum_{i=1}^{k} v'_{ki}\mu_i \end{bmatrix}^T \\
&\qquad \text{where } P = [v_{ij}] \text{ and } P^{-1} = [v'_{ij}] \\
&= \begin{bmatrix} e^{-b_1 t} v_{i_0 1} & \sum_{l=1}^{2} (-1)^{2-l} e^{-b_l t} v_{i_0 l} n_{l2}^{(2-l)} \frac{t^{2-l}}{(2-l)!} & \cdots & \sum_{l=1}^{k} (-1)^{k-l} e^{-b_l t} v_{i_0 l} n_{lk}^{(k-l)} \frac{t^{k-l}}{(k-l)!} \end{bmatrix} \\
&\qquad \begin{bmatrix} \sum_{i=1}^{k} v'_{1i}\mu_i & \sum_{i=1}^{k} v'_{2i}\mu_i & \cdots & \sum_{i=1}^{k} v'_{ki}\mu_i \end{bmatrix}^T \\
&= \sum_{j=0}^{k-1} \sum_{i=1}^{k} c_{ij}^{i_0} t^j e^{-b_i t} \qquad (3.13)
\end{aligned}$$

where

$$c_{ij}^{i_0} = \begin{cases} \frac{(-1)^j}{j!} v_{i_0 i} \prod_{l=i}^{i+j-1} d_l \sum_{r=1}^{k} \mu_r v'_{(i+j)r}, & \text{if } i+j \le k, \\ 0, & \text{otherwise,} \end{cases}$$

and $\prod_{l=i}^{j} d_l = 1$ for $j < i$.

By the fact that

$$\int_0^\infty t^n e^{bt}\,dt = \frac{n!}{b^{n+1}},$$

we see that the average time of AIDS incubation is

$$E(T_{i_0}) = \int_0^\infty t f_{i_0}(t)\,dt = \sum_{j=0}^{k-1}\sum_{i=1}^{k} c_{ij}\frac{(j+1)!}{b_i^{j+2}} \tag{3.14}$$

as the desired result.

3.2 Proof of Model 2.

In this section, we will derive the AIDS incubation function of model 2. The main idea is the same as that of model 1. From model 2, for every $t > 0$, we let X_t be a random variable whose value is the state of HIV epidemic at time t. So the state space of $\{X_t \mid t \geq 0\}$ is $\{S_1, S_2, \ldots, S_{k+1}, S_{k+2}\}$ where S_{k+1} is the B-state and S_{k+2} is the AIDS stage. Hence the state S_{k+2} is an absorbing state and by the fact that $\gamma_i > 0$ for $i = 1, 2, \ldots, k+1$, we see that the states $S_1, S_2, \ldots, S_k, S_{k+1}$ are transition states. Starting at time $s = 0$, let T_i be the random time that S_i is absorbed into S_{k+2}, $i = 1, 2, \ldots, k+1$. Then T_i is referred to as the first passage time of S_i and $g_i(t)$, the probability function, the first passage probability of S_i. In what follows, we put

$$g(t) = [g_1(t), \ldots, g_{k+1}(t)]^T.$$

In this work, we assume that $\{X_t \mid t \geq 0\}$ is a homogeneous continuous Markov process and let the transition matrix $P(t) = [p_{ij}(t)]$ satisfy (3.15) as follows :

$$\left.\begin{aligned} &p_{ij}(t) = v_{ij}t + o(t), \quad \text{where } t \to 0 \text{ and } i \neq j \\ &p_{ii}(t) = 1 - v_{ii}t + o(t), \quad \text{where } t \to 0 \\ &v_{ij} \quad \geq 0, \quad v_{ii} = \sum_{j\neq i} v_{ij}, \end{aligned}\right\} \tag{3.15}$$

when v_{ij} is the transition rate at which X_t jumps from i to j.

By using the same argument of Proposition 3.1, if we can show that

$$A = \begin{bmatrix} \lambda_1 & -\gamma_1 & 0 & \dots & -\mu_1 \\ -\beta_2 & \lambda_2 & \gamma & \dots & -\mu_2 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_k & -\mu_k \\ 0 & 0 & \dots & 0 & \lambda_{k+1} \end{bmatrix}$$

is invertible then we have

$$g(t) = e^{-At} A 1_{k+1}. \tag{3.16}$$

To show that A is invertible, we can observe that

$$A = \begin{bmatrix} R & -\mu \\ 0_k & \gamma_B \end{bmatrix}$$

where

$$R = \begin{bmatrix} \lambda_1 & -\gamma_1 & 0 & \dots\dots\dots\dots & & 0 \\ \beta_2 & \lambda_2 & -\gamma_2 & 0 & \cdots & 0 \\ \vdots & & & & & \vdots \\ 0 & \cdots & 0 & \beta_{k-1} & \lambda_{k-1} & -\gamma_{k-1} \\ 0 & \dots\dots\dots & & 0 & \beta_k & \lambda_k \end{bmatrix},$$

$$\mu = [\mu_1, \mu_2, \dots, \mu_k]^T \quad \text{and}$$

$$0_k = [\underbrace{0, 0, \dots, 0}_{k}].$$

By Theorem 2.1.3, R has a positive determinant. Since $\gamma_B > 0$, we have $\det A = \gamma_B \det R > 0$, i.e., A is invertible. So (3.16) holds.

Let

$$J = \begin{bmatrix} b_1 & d_1 & 0 & \cdots & 0 \\ 0 & b_2 & d_2 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & b_k & d_k \\ 0 & \ldots\ldots\ldots & & 0 & b_{k+1} \end{bmatrix}$$

where $b_i, i = 1, 2, \ldots, k+1$, are all eigenvalues of A and $d_i \in \{0, 1\}$ be the Jordan canonical form of A and $P = [v_{ij}], P^{-1} = [v'_{ij}]$ be $(k+1) \times (k+1)$ matrices such that

$$A = PJP^{-1}.$$

By (3.16) we also use the same argument in proving model 1 to find the AIDS incubation of model 2. The result is as follows for $i_0 \in \{1, 2, \ldots, k+1\}$,

$$g_{i_0}(t) = e_{i_0}^T e^{-At} A 1_{k+1} = \sum_{j=0}^{k} \sum_{i=1}^{k+1} d_{ij}^{i_0} t^j e^{-b_i t} \tag{3.17}$$

where

$$d_{ij}^{i_0} = \begin{cases} \dfrac{(-1)^j}{j!} \left(\prod_{l=i}^{i+j-1} d_l \right) v_{i_0 i} v'_{(i+j)(k+1)} \gamma_B, & \text{if } i + j \leq k+1, \\ 0, & \text{otherwise.} \end{cases}$$

We also know that the mean-time of model 2 is

$$E(T_{i_0}) = \sum_{j=0}^{k} \sum_{i=1}^{k+1} d_{ij}^{i_0} \frac{(j+1)!}{b_i^{j+1}}. \tag{3.18}$$

CHAPTER IV

Some Models on AIDS Incubation

In this chapter, we give a procedure to find AIDS incubation functions and the AIDS Incubation of some special cases. The results follows from the procedure given in 4.1.

4.1 Procedure to find an AIDS incubation distribution function.

From the proof of model 1, we can conclude the procedure to find AIDS incubation function as follows :

1.) Find all eigenvalues $b_1, b_2, \ldots, b_k$ of the matrix

$$A = \begin{bmatrix} \lambda_1 & -\gamma_1 & \ldots & 0 & 0 \\ -\beta_2 & \lambda_2 & -\gamma_2 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_{k-1} & -\gamma_{k-1} \\ 0 & 0 & \ldots & -\beta_k & \lambda_k \end{bmatrix}.$$

2.) Find an invertible matrix P and the Jordan canonical form

$$J = \begin{bmatrix} b_1 & d_1 & 0 & \ldots\ldots & & 0 \\ 0 & b_2 & d_2 & 0 & \cdots & 0 \\ 0 & 0 & b_3 & d_3 & \cdots & 0 \\ \vdots & & & \ddots & \ddots & \vdots \\ 0 & \ldots\ldots & & 0 & b_{k-1} & d_{k-1} \\ 0 & \ldots\ldots\ldots & & & 0 & b_k \end{bmatrix},$$

where $d_i \in \{0, 1\}$ (see Appendix), which satisfy the condition $A = PJP^{-1}$.

3.) By substitution all of the constants from step 1 and step 2 in the formula (3.13), we can solve the AIDS incubation distribution function.

In model 2, we also have the same procedure of model 1.

4.2 Examples for some models

Example 4.1. (Model 1 of Tan and Byers 1993) Assume that $\mu_i = 0$ for $i = 1, 2, \ldots, k-1$ and $\beta_i = 0$, $\gamma_i = \gamma$ for $i = 1, 2, \ldots, k$, then $\lambda_i = \gamma$.

$$S_1 \xrightarrow{\gamma} S_2 \xrightarrow{\gamma} S_3 \longrightarrow \cdots \longrightarrow S_k \xrightarrow{\gamma} \text{AIDS}$$

Model 1 of Tan and Byers 1993

Step 1 We have

$$A = \begin{bmatrix} \gamma & -\gamma & 0 & \ldots\ldots & & 0 \\ 0 & \gamma & -\gamma & 0 & \cdots & 0 \\ \vdots & & & & & \vdots \\ 0 & \ldots\ldots & & 0 & \gamma & -\gamma \\ 0 & \ldots\ldots\ldots & & & 0 & \gamma \end{bmatrix}$$

and γ is the only eigenvalue of matrix A.

Step 2

From Appendix, we can find an invertible matrix P,

$$P = \begin{bmatrix} (-\gamma)^{k-1} & 0 & \ldots\ldots & 0 \\ 0 & (-\gamma)^{k-2} & 0 \;\cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & 0 \;\; -\gamma & 0 \\ 0 & \ldots\ldots\ldots\ldots & 0 & 1 \end{bmatrix},$$

$$P^{-1} = \begin{bmatrix} \frac{1}{(-\gamma)^{k-1}} & 0 & \ldots\ldots & 0 \\ 0 & \frac{1}{(-\gamma)^{k-2}} & 0 \;\cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & 0 \;\; \frac{1}{-\gamma} & 0 \\ 0 & \ldots\ldots\ldots\ldots & 0 & 1 \end{bmatrix},$$

and

$$J = \begin{bmatrix} \gamma & 1 & 0 & \ldots\ldots & 0 \\ 0 & \gamma & 1 & 0 \;\cdots & 0 \\ \vdots & & & & \vdots \\ 0 & \ldots\ldots & 0 & \gamma & 1 \\ 0 & \ldots\ldots\ldots & & 0 & \gamma \end{bmatrix}$$

such that $A = PJP^{-1}$.

We observe that $d_i = 1$ for $i = 1, 2, \ldots, k-1$.

Step 3

$$c_{ij} = \begin{cases} \dfrac{\gamma^{k-i_0+1}}{(k-i_0)!}, & \text{if } i = i_0 \text{ and } j = k - i_0, \\ 0, & \text{otherwise.} \end{cases}$$

So $f_{i_0}(t) = \dfrac{\gamma^{k-i_0+1}}{(k-i_0)!}\, t^{k-i_0} e^{-\gamma t}$ which is Gamma density. #

Example 4.2. (Model 1 of Longini et al. (1989a, 1989b, 1991) and Anderson et al. (1989)) Assume that $\mu_i = 0$ for $i = 1, 2, \ldots, k-1$, $\beta_i = 0$ for $i = 1, 2, \ldots, k$ and γ_i's are all distinct.

$$S_1 \xrightarrow{\gamma_1} S_2 \xrightarrow{\gamma_2} S_3 \longrightarrow \cdots \longrightarrow S_k \xrightarrow{\gamma_k} \text{AIDS}$$

Model 1 of Longini et al. (1989a, 1989b, 1991) and Anderson et al. (1989)

Step 1

$$A = \begin{bmatrix} \gamma_1 & -\gamma_1 & 0 & \ldots\ldots\ldots & & 0 \\ 0 & \gamma_2 & -\gamma_2 & 0 & \cdots & 0 \\ \vdots & & & & & \vdots \\ 0 & \ldots\ldots\ldots & & 0 & \gamma_{k-1} & -\gamma_{k-1} \\ 0 & \ldots\ldots\ldots\ldots & & & 0 & \gamma_k \end{bmatrix},$$

which all eigenvalues are all entries in the main diagonal.

Step 2

We have

$$J = \begin{bmatrix} \gamma_1 & 0 & \ldots\ldots & & 0 \\ 0 & \gamma_2 & 0 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & \cdots & 0 & \gamma_{k-1} & 0 \\ 0 & \ldots\ldots & & 0 & \gamma_k \end{bmatrix},$$

$P = [v_{ij}]$ and $P^{-1} = [v'_{ij}]$ where v_{ij} and v'_{ij} are defined by :

$$v_{ij} = \begin{cases} \displaystyle\prod_{t=i}^{j-1} \frac{\gamma_t}{\gamma_t - \gamma_j}, & \text{if } i \le j, \\ 0, & \text{if } i > j, \end{cases}$$

and

$$v'_{ij} = \begin{cases} \dfrac{\prod_{t=i}^{j-1} \gamma_t}{\prod_{t=i+1}^{j} \left(\gamma_t - \gamma_i\right)}, & \text{if } i \leq j, \\ 0, & \text{if } i > j. \end{cases}$$

Hence $d_i = 0$ for $i = 1, 2, \ldots, k-1$.

Step 3

$$\begin{aligned} f_{i_0}(t) &= \sum_{i=1}^{k} c_{i0}^{i_0} e^{-\gamma_i t} \\ &= \sum_{i=1}^{k} \left(\prod_{j=1}^{i-1} \frac{\gamma_j}{\gamma_j - \gamma_i} \right) \left(\frac{\gamma_i \gamma_{i+1} \cdots \gamma_k}{\prod_{j=i+1}^{k} (\gamma_j - \gamma_i)} \right) e^{-\gamma_i t} \\ &= \sum_{i=1}^{k} \frac{\prod_{j=1}^{k} \gamma_j}{\prod_{\substack{j=1 \\ j \neq i}}^{k} (\gamma_j - \gamma_i)} e^{-\gamma_i t}. \end{aligned}$$

#

Appendix

Procedure to find an invertible matrix P which can transform a square matrix A into its Jordan canonical Form J

In what follows, the argument is considered in the field of complex numbers. Our details are referred from "Matrix Methods, an Introduction" by Richard Bronson in Chapter 8 *Jordan Canonical Forms*, page 176–207.

Let I be the identity matrix. To obtain an invertible matrix which can transform a square matrix into its Jordan canonical form, we need the following definitions and theorems.

Definition 1 Given a positive integer m and a square matrix A, a vector x_m is a **generalized eigenvector of rank** m corresponding to the matrix A and the eigenvalue λ if

$$(A-\lambda I)^m x_m = 0 \quad \text{but} \quad (A-\lambda I)^{m-1} x_m \neq 0.$$

We note that a generalized eigenvector of rank 1 is, in fact, an eigenvector corresponding to λ but a generalized eigenvector of rank $k \neq 1$ is not necessarily an eigenvector corresponding to λ.

Definition 2 Let x_m be a generalized eigenvector of rank m corresponding to a matrix A and an eigenvalue λ. The **chain generated by** x_m is a set of vectors $\{x_m, x_{m-1}, \ldots, x_1\}$ given by

$$\begin{aligned}
x_{m-1} &= (A-\lambda I)x_m \\
x_{m-2} &= (A-\lambda I)^2 x_m = (A-\lambda I)x_{m-1} \\
x_{m-3} &= (A-\lambda I)^3 x_m = (A-\lambda I)x_{m-2} \\
&\vdots \\
x_1 &= (A-\lambda I)^{m-1} x_m = (A-\lambda I)x_2.
\end{aligned}$$

Thus, in general

$$x_j = (A - \lambda I)^{m-j} x_m = (A - \lambda I) x_{j+1} \;\; (j = 1, 2, \ldots, m-1). \tag{1}$$

Theorem 3 The vector x_j(given in (1)) is a generalized eigenvector of rank j corresponding to the eigenvalue λ.

Theorem 4 The chain generated by x_m in Definition 2 is a linearly independent set of vectors.

Theorem 5 Every $n \times n$ matrix A possesses n linearly independent generalized eigenvectors. Generalized eigenvectors corresponding to distinct eigenvalues are linearly independent. If λ is an eigenvalue of A of multiplicity (the number of the repeated value of λ) ν, then A will have ν linearly independent generalized eigenvectors corresponding to λ.

Definition 6 A set of n linearly independent generalized eigenvectors is a **canonical basis** if it is composed entirely of chains.

Definition 7 Let A be an $n \times n$ matrix. A **generalized modal matrix** M for A is an $n \times n$ matrix whose columns, considered as vectors, form a canonical basis for A and appear in M according to the following rules:

M1 All chains consisting of one vector (that is, one vector in length) appear in the first columns of M.

M2 All vectors of one chain appear together in adjacent columns of M.

M3 Each chain appears in M in order of increasing rank (that is, the generalized eigenvectors of rank 1 appear before the generalized eigenvectors of rank 2 of the same chain, which appear before the generalized eigenvector of rank 3 of the same chain, etc.).

The important fact, however, is that for any arbitrary $n \times n$ matrix A, there exists at least one generalized modal matrix M corresponding to it. Furthermore, since the columns of M considered as vectors form a linearly independent set, it follows that the column rank of M is n, i.e., the rank of M is n. As a result the determinant of M is nonzero, i.e., M^{-1} exists.

Construction of an invertible matrix P

Let A be any square $n \times n$ matrix.

Step 1 From the characteristic matrix

$$C(x) = A - xI,$$

obtain the characteristic polynomial

$$f(x) = (x - \lambda_1)^{\nu_1}(x - \lambda_2)^{\nu_2} \dots (x - \lambda_p)^{\nu_p},$$

where $\lambda_1, \lambda_2, \dots, \lambda_p$ are all distinct eigenvalues of A.

Fix $i \in \{1, 2, \dots, p\}$.

Step 2 Find the ranks of the matrices $(A - \lambda_i I), (A - \lambda_i I)^2, (A - \lambda_i I)^3, \dots,$ $(A - \lambda_i I)^m$ where the integer m is determined to be the first integer for which $(A - \lambda_i I)^m$ has rank $n - \nu_i$.

Step 3 Now define

$$\rho_k = \operatorname{rank}(A - \lambda_i I)^{k-1} - \operatorname{rank}(A - \lambda_i I)^k \quad (k = 1, 2, \dots, m).$$

Then, ρ_k designates the number of linearly independent generalized eigenvectors of rank k corresponding to the eigenvalue λ_i that will appear in a canonical basis for A. Note that $\operatorname{rank}(A - \lambda_i I)^0 = \operatorname{rank}(I) = n$.

Step 4 Find the vectors defined in Definition 1 and Definition 2 explicitly to obtain a canonical basis for A.

Step 5 Find a generalized modal matrix M by applying Definition 7 and let $P = M^{-1}$.

Step 6 We can transform A to its Jordan canonical form J by setting $J = PAP^{-1}$.

Example 1. Let

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

We will find a generalized modal matrix M by the above procedure.

Step 1 Obtain the characteristic matrix

$$C(x) = \begin{bmatrix} 1-x & 1 & 0 & -1 \\ 0 & 1-x & 0 & 0 \\ 0 & 0 & 1-x & 1 \\ 0 & 0 & 0 & 1-x \end{bmatrix}.$$

Then $f(x) = (x-1)^4$ is its characteristic function. Thus $\lambda_1 = 1$ is the unique eigenvalue of A of multiplicity 4; hence, $n = 4, \nu = 4$ and $n - \nu = 0$.

Step 2 We can see that

$$(A - 1I) = \begin{bmatrix} 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

has rank 2 and

$$(A - 1I)^2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

has rank $0 = n - \nu$. Thus, $m = 2$.

Step 3 Then we know that $\rho_2 = \text{rank}(A - 1I) - \text{rank}(A - 1I)^2 = 2 - 0 = 2$ and $\rho_1 = \text{rank}(A - 1I)^0 - \text{rank}(A - 1I) = 4 - 2 = 2$; hence, a canonical basis for A will have two linearly independent generalized eigenvectors of rank 2 and two linearly independent generalized eigenvectors of rank 1.

Step 4 In order for a vector

$$\begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix}$$

to be a generalized eigenvector of rank 2, we will solve the equation

$$(A - 1I)[w\ x\ y\ z]^T \neq 0 \quad \text{and} \quad (A - 1I)^2[w\ x\ y\ z]^T = 0.$$

We can see that either x or z must be non–zero and w and y are arbitrary. If we first choose $x = 1, w = y = z = 0$ and then choose $z = 1, w = x = y = 0$, we obtain that two linearly independent generalized eigenvectors of rank 2 are

$$x_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad y_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Note that we could have chosen w, x, y, z in such a manner as to generate 4 linearly independent generalized eigenvectors of rank 2. The vectors

$$\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}$$

together with x_2 and y_2 form such a set. Thus we immediately have found a set of 4 linearly independent generalized eigenvectors corresponding to λ_1. This set, however, is **not** a canonical basis for A since it is not composed entirely of chains. In order to obtain a canonical basis for A, we use only two of these vectors (we will, in particular, use x_2 and y_2) and form chains from them.

By Definition 2, we obtain the two linearly independent generalized eigenvectors of rank 1 to be

$$x_1 = (A - I)x_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad y_1 = (A - I)y_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}.$$

Thus, a canonical basis for A is $\{x_2, x_1, y_2, y_1\}$, which consists of the two chains each containing two vectors $\{x_2, x_1\}$ and $\{y_2, y_1\}$.

Step 5 Since this canonical basis has no chain consisting of one vector, **(M1)** is not applied. From **(M2)**, we assign either x_2 and x_1 to the first two columns of M and y_2 and y_1 to the last two columns of M or, alternatively, y_2 and y_1 to the first two columns of M and x_2 and x_1 to the last two columns of M. We can not, however, define $M = [x_1 \ y_1 \ y_2 \ x_2]$ since this alignment would split the $\{x_2, x_1\}$ chain and violate **(M2)**. Due to **(M3)**, x_1 must precede x_2 and y_1 must precede y_2; hence

$$M = [x_1\ x_2\ y_1\ y_2] = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

or

$$M = [y_1\ y_2\ x_1\ x_2] = \begin{bmatrix} -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

Example 2. Find a Jordan canonical form of

$$A = \begin{bmatrix} 4 & 0 & 1 & 0 \\ 2 & 2 & 3 & 0 \\ -1 & 0 & 2 & 0 \\ 4 & 0 & 1 & 2 \end{bmatrix}.$$

Step 1 The characteristic equation for A is $(\lambda - 3)^2(\lambda - 2)^2 = 0$; hence, $\lambda_1 = 3$ and $\lambda_2 = 2$ are both eigenvalues of multiplicity 2.

Step 2 For $\lambda_1 = 3$, we find that $n - \nu_1 = 2, m_1 = 2$. For $\lambda_2 = 2$, we find that $n - \nu_2 = 2, m_2 = 1$.

Step 3 For $\lambda_1 = 3$, we find that $\rho_2 = 1$, and $\rho_1 = 1$, so that a canonical basis for A has one generalized eigenvector of rank 2 and one generalized eigenvector of rank 1.

For $\lambda_2 = 2$, we find that $\rho_1 = 2$; hence, there are two generalized eigenvectors of rank 1.

Step 4 For $\lambda_1 = 3$, by Definition 1, we find that a generalized eigenvector of rank 2 is

$$x_2 = \begin{bmatrix} 1 \\ 3 \\ 0 \\ 1 \end{bmatrix}.$$

By Definition 2, we have that

$$x_1 = (A - 3I)x_2 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 3 \end{bmatrix}$$

is a generalized eigenvector of rank 1.

For $\lambda_2 = 2$, by Definition 1 and Definition 2, we obtain

$$y_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad z_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

as the required vectors. Thus a canonical basis for A is $\{x_2, x_1, y_1, z_1\}$ which consists of one chain containing two vectors $\{x_2, x_1\}$ and two chains containing one vector apiece $\{y_1\}$ and $\{z_1\}$. Note that once again, due to Theorem 5, we are guaranteed that $\{x_2, x_1\}$ are linearly independent of $\{y_1, z_1\}$ since they correspond to different eigenvalues.

Step 5 By Definition 7, the first two columns of M must be y_1 and z_1(however, any order) due to **(M1)** while the third and fourth columns must be x_1 and x_2, respectively, due to **M(3)**. Hence,

$$M = [y_1\ z_1\ x_1\ x_2] = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & -1 & 3 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 3 & 1 \end{bmatrix}$$

or

$$M = [z_1\ y_1\ x_1\ x_2] = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & -1 & 0 \\ 1 & 0 & 3 & 1 \end{bmatrix}.$$

For $M = [y_1\ z_1\ x_1\ x_2]$, we compute

$$P := M^{-1} = \begin{bmatrix} -3 & 1 & -4 & 0 \\ -1 & 0 & 2 & 1 \\ 0 & 0 & -1 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix}.$$

Step 6 Thus

$$\begin{aligned} PAP^{-1} &= \begin{bmatrix} -3 & 1 & -4 & 0 \\ -1 & 0 & 2 & 1 \\ 0 & 0 & -1 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 4 & 0 & 1 & 0 \\ 2 & 2 & 3 & 0 \\ -1 & 0 & 2 & 0 \\ 4 & 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} -3 & 1 & -4 & 0 \\ -1 & 0 & 2 & 1 \\ 0 & 0 & -1 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{bmatrix} := J \end{aligned}$$

is the Jordan canonical form of A.

Example 3. Find a matrix in Jordan canonical form that is similar to

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

Step 1 The characteristic equation of A is $(\lambda - 1)^7 = 0$, hence, $\lambda = 1$ is an eigenvalue of multiplicity 7.

Step 2 We find that $\text{rank}(A-1I) = 3, \text{rank}(A-1I)^2 = 1$ and $\text{rank}(A-1I)^3 = 0$ $= n - \nu$. Then $m = 3$.

Step 3 Thus we can find $\rho_3 = 1, \rho_2 = 2$ and $\rho_1 = 4$ which implies that a canonical basis for A will consist of one linearly independent generalized eigenvector of rank 3, two linearly independent generalized eigenvectors of rank 2 and four linearly independent generalized eigenvectors of rank 1.

Step 4 and Step 5 From Step 3 we know that a canonical basis for A consists of one chain of three vectors $\{x_3, x_2, x_1\}$, one chain of two vectors $\{y_2, y_1\}$, and two chains of one vector $\{z_1\}, \{w_1\}$ (we can find a canonical basis for A from the Definition 1 and Definition 2). Designating

$$M = [z_1 \ w_1 \ x_1 \ x_2 \ x_3 \ y_1 \ y_2],$$

we find that

$$M = \begin{bmatrix} 0 & 1 & -1 & 0 & 0 & -2 & 1 \\ 0 & 3 & 0 & 0 & 1 & 0 & 0 \\ -1 & 1 & 1 & 1 & 0 & 2 & 0 \\ -2 & 0 & -1 & 0 & 0 & -2 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & -1 & 0 \end{bmatrix}.$$

Thus,

$$M^{-1} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 1 & 4 & -2 & 2 \\ 0 & 0 & 1 & 1 & 3 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & -3 & 0 \\ 0 & 0 & -1 & -1 & -3 & 1 & -1 \\ 1 & 0 & 0 & -1 & -2 & -1 & 0 \end{bmatrix}.$$

Let $P := M^{-1}$.

Step 6 Thus,

$$J = PAP^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

References

1. Anderson, R.M., Blythe, S.P., Gupta, S. and Konings, E., The transmission dynamics of the human immunodeficiency virus type I in the male homosexual community in the United Kingdom: the influence of changes in sexual behavior, Philos. Trans. Roy. Soc. London Ser. B. Vol.325 (1989) : 7–60.

2. Bacchetti, P., Estimating the incubation period of AIDS by comparing population infection and diagnosis patterns, J. Amer. Statist. Assoc. Vol.85 (1990), 1002–1008.

3. CDC, Classification systems for human T-lymphotropic virus Type-III/lymphadenopathy-associated virus infections, MMWR. Vol.35 (1986), 334–339.

4. CDC, AIDS Public Information Data Set, Centers for Disease Control, Center for Infectious Disease, Division of HIV/AIDS (Atlanta, GA) (1991).

5. Chiang, C.L., Introduction to Stochastic Processes and their Applications. New York, Krieger, 2nd ed.,(1980).

6. Curtis, C.W., Linear Algebra : An Introductory Approach. New York, Springer-Verlag, (1984).

7. Glaz, J., Probabilities and moments for absorption infinite homogeneous birth-death processes, Biometrics. Vol.35(1979) : 813–816.

8. Hethcote, H.W., Van Ark, J.W. and Longini Jr., M., A simulation model of AIDS in San Francisco: I. Model formulation and parameter estimation, Math. Biosci. Vol.106(1991) : 203–222.

9. Longini Jr., I.M., Clark,W.S., Byers,R.H., et al., Statistical analysis of the stages of HIV infection using a Markov model, Statist.Medicine. Vol.8(1989a) : 831–843.

10. Longini Jr., I.M., Clark, W.S., Haber, M. and Horsburgh, R., The Stages of HIV infection: Waiting times and infection transmission probabilities, in: C. Castillo-Chavez, ed., Lecture Notes in Biomath. No. 83, Mathematical and Statistical Approaches to AIDS Epidemiology. New York, Springer,(1989b).

11. Longini Jr., I.M., Clark, W.S., Gardner, L.I. and Brundage, J.F., The dynamics of CD4+ T-lymphocyte decline in HIV-infected individuals: A Markov modeling approach, J. AIDS. Vol.4(1991) : 1141–1147.

12. Longini Jr., I.M., Byers, R.H., Hessol, N.A., and Tan, W.Y., Estimation of the stage-specific numbers of HIV infections via a Markov model and backcalculation, Statist. Medicine. Vol.11(1992) : 831–843.

13. Nagelkerke, N.J.D., et al., Transition dynamics of HIV disease in a cohort of African prostitutes: A Markov model approach, AIDS. Vol.4(1990) : 743–747.

14. Rao (1973), Linear Statistical Inference and its Applications. John Wiley & Sons, New York, 2nd ed.

15. Redfield, R.R. and D.S. Burke , HIV infection: The clinical picture, Sci. Amer. Vol.259(1988) : 90–99.

16. Richard, B. (1969), Matrix Methods. Academic Press, New York and London, 1st ed..

17. Sidney, R. (2002), Adventures in Stochastic Processes. Hamilton Printing Co., Rennselaer, New York, 3rd ed..

18. Tan, W.Y., On the absorption probability and absorption time of finite homogeneous birth-death processes, Biometrics. Vol.32(1976) : 745–752.

19. Tan, W.Y., On the distribution of second degree polynomial in normal random variables, Canad. J. Statist. Vol.5(1977) : 241–250.

20. Tan, W.Y. and H. Hsu, Some stochastic models of AIDS spread, Statist. Medicine. Vol.8(1989) : 121–136.

21. Tan, W.Y. and Byers, R.H., A stochastic model of the HIV epidemic and the HIV infection distribution in a homosexual population, Math. Biosci. Vol.113(1993) : 115–143.

22. Tan, W.Y., On the incubation distributions of the HIV epidemic, Statistics and Probability Letter. Vol.18(1993) : 279–287.

23. Zhou, S.Y., Kingsley, L.A. and Leung, S., Time from HIV seroconversion to CD4+ $\leq$ 200 using a Markov model, paper presented at the ASA-Biometric Society Winter Conference. , Ft. Lauderdale, FL, Jan (1993) : 3–5.

VITA

Mr. Ratchanikorn Chonchaiya was born on May 10, 1980 in Buriram, Thailand. He graduated with a Bachelor Degree of Science in Mathematics with second class honor from Chulalongkorn University in 2002. For his Master's degree, he has studied Mathematics at the Department of Mathematics, Faculty of Science, Chulalongkorn University. Upon his graduation , he will serve as a lecturer at the Faculty of Science, Burapa University.