A STOCHASTIC EXTENSION OF THE STRUCTURAL PROPERTIES OF COMPARTMENT MODELS VIA A KRONECKER ALGEBRA

By

RICHARD JAY WEIDNER
Bachelor of Science in Electrical Engineering
Oklahoma State University
Stillwater, Oklahoma
1977

Master of Science
Oklahoma State University
Stillwater, Oklahoma
1979

Submitted to the Faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
DOCTOR OF PHILOSOPHY
May, 1981

A STOCHASTIC EXTENSION OF THE STRUCTURAL

PROPERTIES OF COMPARTMENT MODELS

VIA A KRONECKER ALGEBRA

Thesis Approved:

Robert J. Mulholland
Thesis Adviser

Craig S. Sims

James R. Rowland

Marvin S. Keener

Norman N. Durham
Dean of Graduate College

ACKNOWLEDGMENTS

I would like to express my heart-felt thanks to Dr. Robert J. Mulholland, my thesis adviser. Without his continuing support and advice, my graduate education would not have been possible. I would also like to express my appreciation to the members of my doctoral committee, Dr. C. S. Sims, Dr. J. R. Rowland, Dr. R. Yarlagadda, and Dr. M. Keener, for their critical guidance of my research as it developed and for their immense contributions to my education.

Financial support was received under NSF Research Grant ENG 75-05341-A01, EPA Grant CR-80630-01-0, and an OSU Water Research Center Grant. They are gratefully acknowledged. Dr. Robert J. Mulholland was the principal investigator on these grants.

I would like to acknowledge the Oklahoma State Computing Center for their aid in both supporting the computer research and in developing the word processing capabilities upon which the preliminary drafts were written. Their patience and efforts are often overlooked.

Finally I would like to express my thanks to the people who lent me spiritual support during my doctoral education. First, to my parents for accepting my wishes even if they did not understand them, and second, to Dr. Yarlagadda and Dr. Mulholland for believing in me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The relationships of sections of energy or matter are the basic foundation of engineering science. They are clearly defined by the second and third laws of Thermodynamics. However, as a more practical form, the relationships may be expressed as state equations. As a result, a nonlinear matrix relation can be formed. The terms in the matrices have direct meaning in the physical world.

Chang and Fitzgerald (1977) have illustrated the relationship of compartmental matrix models with stirred tank chemical reaction networks. The matrix terms are composed of the fractional input flow coefficients. Barkstresser et al. (1979) has shown one application of compartmental methods in renal function identification. The compartments are the volume of the tracer distributions. Matis and Wehrly (1979) examined the application of stochastic compartment models in drug kinetics. Not least, is the ever expanding use of the compartment model in ecological systems as illustrated by Gowdy (1978).

The list of applications for compartment models has grown at a rapid rate, until the science of compartment model mathematics has become a study within itself, as illustrated by Sandberg (1978). Pre-dominant in compartmental methods is the linear donor-controlled compartment model and the resultant column diagonally dominant state coefficient matrix. The laws of mass or energy conservation find equivalent expression in the

relationships of the diagonal terms with the off-diagonal terms in the coefficient matrix.

Such matrices have established a body of structural properties which are unique to matrices of their class. Diagonal dominance is just one property of the group. Concepts of connectivity and reducibility grow from the direct relationship to physical systems. There are many other properties implied by the unique structure of such matrices. Chapter II recaps the work of Gowdy (1978) in examining the body of structural properties for deterministic compartment models.

Compartment models are based on empirical observation, as evidenced in their use in parametric identification (Campello and Cobelli [1978], Mulholland and Gowdy [1977], Mulholland and Weidner [1980]). This fact, plus their close relationship to lumped models of the real world, necessitates the use of stochastic methods in determining the models and their structural properties. Chapter III examines the stochastic moments of compartmental states considered as random variables. Note that this study does not examine the relation commonly known as a stochastic compartment model. Rather, it examines what has been considered as a deterministic relation. Except, now, this study adds the practical effects of uncertainty. In particular, Chapter III considers the effects of input uncertainty in compartment model identification. A Monte Carlo study is performed, and a clear suggestion of behavioral properties for the variance is observed. The sample rate for identification is shown to fall within strict bounds established by the variance of the states.

The dynamic variance obeys an equation of the form of the Lyapunov derivative matrix equation. While the equation is familiar, there is some difficulty in examining the properties of the variance, as opposed

to examining terms with the more conventional simple state matrix formulation. Chapter IV examines a Kronecker algebra as a matrix to vector transformation. The nice structure of the Kronecker product and sum yield many simplifying forms for the matrix systems. An appendix is included to list some of the more interesting algebraic identities. As a result of the Kronecker algebra, the Lyapunov equation is converted to the simple matrix form, allowing ready structural analysis. As an addition, Chapter IV examines the special case of compartment "stochastic" systems, otherwise known as Markov processes. The Kronecker algebra is shown to have beneficial structure for multiple-unit Markov system decomposition.

The reformulation of the Lyapunov equation via the Kronecker algebra establishes an analytic approach for examining the preservation of structural properties of the state mean in the state variance. Chapter V examines a number of proofs utilizing the Kronecker algebra and establishing the sought after structural properties. Matrix properties such as diagonal dominance, irreducibility, and boundedness are a few of the properties examined.

The result of these new properties is a revised identification theory. Now dynamic confidence intervals may be constructed, and a new sample theory proposed. Chapter VI examines the specifics of the dynamic bounds on the variance of compartment models. Gershgorin (1931) theory, as well as the techniques presented by Gowdy (1978), play an important role in identification of the variance bounds. Finally, guidelines are proposed for continuing this research in the construction of a new sample theory for identification of compartment models.

CHAPTER II

REVIEW OF BACKGROUND LITERATURE

Introduction

The analysis of the structural configuration of a specific system, and the properties implied by that configuration, can lend valuable information toward the application of scientific principles in the identification and subsequent manipulation of that system. Importantly, the general structural properties of a class of systems can similarly lend information for the analysis of any particular system from that class. The general field of compartment modeled systems is such a class which has had great success in the past in applying this principle.

This chapter will review pertinent literature in order to establish a sufficient background in the structural properties of compartment models. The first section will be a brief compilation of the historical development of compartment models. The next section will develop in more detail the properties of compartment models which are important to this study.

It should be noted that the structural properties of compartment models are drawn from arbitrary system boundaries. Thus, the casual observer of the physical system may not clearly visualize the structural identity.

History

The main concept of compartment models involves the dynamic relationship of distinct units of mass or energy. Though a large number of systems models satisfy this rule, researchers in the area of compartment models restrict the definition more severely. Commonly, a compartment model is a group of lumped interconnected entities governed by the laws of mass or energy conservation.

The structure of compartment models is inferred from samples of the compartments combined with "common sense" knowledge of the system. Therefore, a compartment is assumed to be composed of one homogeneous substance, and correspondingly, the flow into the compartment is instantaneously mixed to keep the compartment content homogeneous. Obviously, this assumption is not exactly valid, but the nature of the system observations renders the assumption acceptable, and often unavoidable.

Another assumption commonly applied is the property of "donor control." That is, the outflow from a compartment is dependent only on the content of that compartment. The state of the system is assumed to be below saturation. This type of compartment model represents the ecological condition of ultimate resource limitation (Patten, 1975).

The above are the general concepts applied to most compartment models in the literature. It is for this type of model that Sheppard (1948) first coined the term "compartment." However, the concepts were used much earlier in the twentieth century.

Different areas of science can date the use of compartment models in their areas to different sources in the early 1900's. Hevesy (1923) was the first to use compartment models with tracer studies in the area of Biology. He used the concept when he demonstrated the uptake and loss

of lead ions in plant roots. Kostitzin (1935) applied the compartmental analysis approach to ecological systems. Teorell (1937) founded the use of compartmental analysis in pharmacokinetics to study the kinetics of drug absorption, distribution, metabolism, and excretion.

Still the mathematical contributions could be traced back even farther to the later part of the nineteenth century. The matrix properties of diagonal dominance can be traced back to Levy (1881) and Desplanques (1886). The relationships of positive quantities exhibit the properties of positive matrices, in particular, the dominant eigenvalue theory of Perron (1907) and Frobenius (1912). The analysis of the relationships of those eigenvalues uses the circle theory of Gershgorin (1931) to great effect. Finally, the influence of Taussky (1949), in promoting these then underemphasized results to the scientific community, cannot be left unmentioned.

The analysis of the structural properties of compartment models relies heavily on the mathematical theory of these pioneers married with the connective structure of the compartments, as illustrated by Hearon (1963) and Thron (1972). In particular, the identification theory of ecosystems has gained from this "marriage" (Mulholland and Gowdy, 1977).

Matrix Compartment Model Structural Properties

Possibly, the foremost property that compartment models must exhibit is the conservation of mass. Sandberg (1978) expresses this property mathematically in the nonlinear equation

$$\dot{x}_i = f_{io} + \sum_{j=1}^{n}{}' (f_{ij} - f_{ji}) - f_{oi}, \quad i = 1, \ldots, n \tag{2.1}$$

where there are n compartments in the system, the prime notation indicates

summation for all $j \neq 1$, f_{rs} is the mass flow rate to compartment r from compartment s, subscript o indicates the environment, x_i represents the content of compartment i, and $\dot{x}_i$ represents the time rate of change of x. A hypothetical three compartment system governed by Equation (2.1) would be represented by Figure 1.

Equation (2.1) represents several types of systems (Brown, 1980). Many systems are accurately modeled by nonlinear, time varying relationships. In particular, ecosystems and chemical kinetics are highly nonlinear. Ecosystems are often state and time dependent as discussed in Smith (1970). Chemical kinetics tend to fail the homogeneous substance assumption, giving rise to multiple sub-compartments for each compartment and rendering nonlinear relationships (Brown, 1980).

However, an important method of observing the states of compartment models is through the injection of tracers. Ideally a tracer is a radioactive or stable substance which is kinetically indistinguishable from the compartmental contents. It also must be limited amounts in order to not perturb the observed system. Lastly, it must be uniformly distributed throughout the compartment. Various information about tracers may be obtained from Sheppard (1962), Jacquez (1972), and Shipley and Clark (1972).

The analysis of these tracer experiments is valid with the assumption of linearity (Gowdy, 1978). The tracer states are related similarly to the states in the linear compartment model, except the new states are inversely proportional to the original states of the system. If the nonlinear rate coefficients are expressed by $k_{ij}f_j(\bar{x}_j)$, the tracers are related by

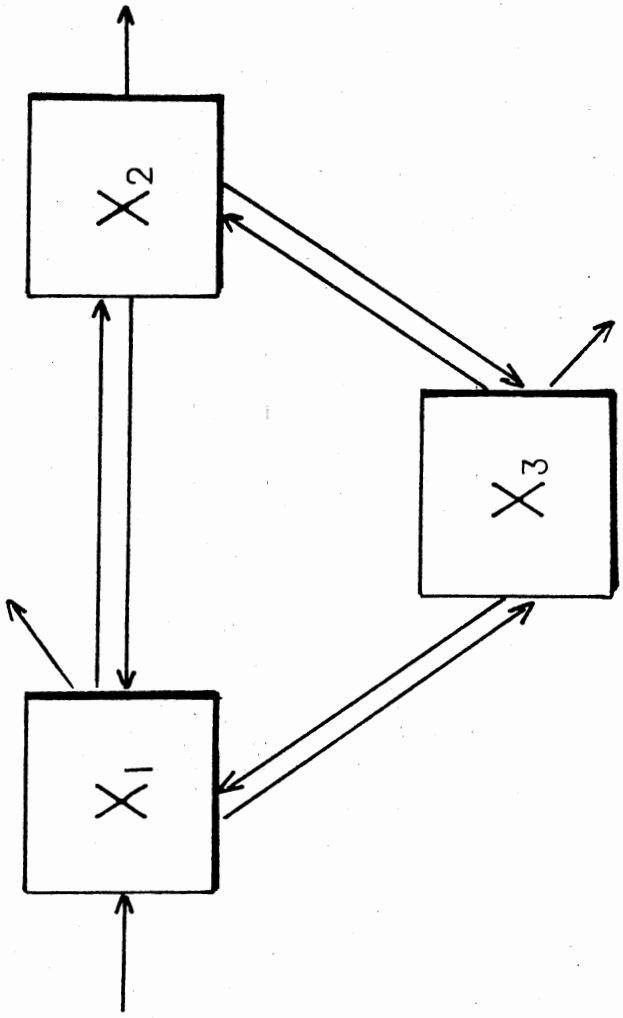

Figure 1. A Hypothetical Three-Compartment System

$$\dot{\delta}_i = \sum_{j=1}^{n}{}' [k_{ij} f_j(\bar{x}_j)\ \delta_j/\bar{x}_j - k_{ji} f_i(\bar{x}_i)\ \delta_i/\bar{x}_i]$$

$$- k_{oi} f_i(\bar{x}_i)\ \delta_i/\bar{x}_i \qquad (2.2)$$

where δ_i is the amount of tracer in compartment i, $\bar{x}_i$ is the steady state content of compartment i, $\dot{\delta}_i$ is the time rate of change of δ_i, and $\delta_i << \bar{x}_i$.

Another class of compartment models is the stochastic compartment model. Actually, this is two distinct types of models. First, the amount of material in the compartment can be considered as a finite number of particles, and a probability--such as the Markov process--can be derived to describe the model. In the second case the rate coefficients can be considered as random variables. When classical statistical techniques are applied to the identification of compartment models, the estimates of the rate coefficients may be considered as random variables with associated probability distributions. The first type of compartment model tends in the limit, as the number of particles approaches infinity, to approach the deterministic linear compartment model. The second type of compartment model is a facet of real world identification, which all compartment models exhibit. Unfortunately, the sample size is often insufficient to identify the probability distribution and the related probability moments.

Due to the wealth of information on linear systems, the most common compartment model is the linear constant coefficient model.

$$\dot{x}_i = f_{io} + \sum_{j=1}^{n}{}' (k_{ij} x_j - k_{ji} x_i) - k_{oi} x_i \qquad (2.3)$$

where the terms are defined as before, except the k_{ij} coefficients are

now time and state independent. Equation (2.3) can be restructured to form (Thron, 1972)

$$\dot{x} = f + A\,x \tag{2.4}$$

where x is a n vector of the compartment states, f is the n vector of the flows from the environment to the compartments, and the coefficient matrix A is the matrix of transfer rate coefficients ($a_{ij} = +k_{ij}$). Now

$$-a_{ii} = a_{oi} + \sum_{j=1}^{n}{}' \; a_{ji} \tag{2.5}$$

The main emphasis of this study will be on compartment models of the form of Equation (2.4).

It is important to note that when the flow rate from a compartment to the environment is zero, the sum of the terms in the column of A corresponding to that compartment is also zero. This is known as column diagonal dominance.

Many of the important mathematical properties of compartment models are derived from whether the compartments are open (material can pass from the system, $a_{oi} \neq 0$, or not), or whether it is strongly connected (material can pass from one compartment to any other compartment without leaving the system), or whether it contains any cycles (materials can loop around in three or more compartments). Both Hearon (1963) and Thron (1972) have written extensively on this subject. Since this study utilizes these results, their work will be largely recreated here, though in a much briefer form.

It is important to note that the matrix A is by definition composed of real elements, and that the diagonal elements are of opposite sign from the off-diagonal elements. Since the matrix is structured to

exhibit diagonal dominance, when none of the compartments is closed (i.e., $a_{oi} \neq 0$, $i = 1, \ldots, n$) the determinant of A is not zero (Taussky, 1949). This has important connotations. It implies that none of the eigenvalues of A is zero for a completely open system. It also implies that A is nonsingular and that its principal minors are nonsingular, under this assumption. The singularity of A has serious implications in controllability, observability, identifiability, etc.

When some of the compartments, considered as a set, are cut off from the rest of the system by null transfer rates, they are said to be not "strongly connected." The coefficient matrix A for such a system is said to be reducible to a diagonal form (Gantmacher, 1959)

$$\begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix} \tag{2.6}$$

where A_{11} and A_{22} are square matrices.

Extending the concept of reducibility with connectiveness for compartment models, Thron (1972) has shown that matrix A can be permuted to the form of Equation (2.6) where

$$A_{11} = \begin{bmatrix} A_{1,1} & \cdots & 0 \\ \vdots & & \vdots \\ A_{p,1} & \cdots & A_{p,p} \end{bmatrix}$$

$$A_{22} = \begin{bmatrix} A_{p+1,p+1} & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & A_{n,n} \end{bmatrix}$$

$$A_{21} = \begin{bmatrix} A_{p+1,1} & \cdots & A_{p+1,p} \\ \vdots & & \vdots \\ A_{n,1} & \cdots & A_{n,p} \end{bmatrix}$$

The partitions separate the singular and nonsingular sets of compartments along the diagonal. The diagonal partitions are all irreducible and the diagonal blocks in A_{11} are all nonsingular, while the diagonal blocks in A_{22} are singular. From the diagonal dominance property, the column sums related to A_{11} are nonzero. It is obvious from the form of Equation (2.6) that the singular subset of the compartment model can be disconnected from the nonsingular section, and the scientific manipulation of the nonsingular subsystem can be performed.

Before proceeding with the evolution of structural behavior for linear compartment models, it should be noted that many similar results exist for nonlinear compartment models, given certain key assumptions (Sandberg, 1978). The nonlinear forms are more powerful, though their applications are not useful to this study.

A practicing systems scientist would appreciate knowledge of many of the behavioral characteristics of a particular compartment system. Many of the characteristics are well founded in common sense knowledge of the system. Stability and monotonic responsiveness are just a couple of these behaviors. The evaluation of the properties of the matrix A has added mathematical reinforcement to many of these empirical characteristics, as well as suggesting many characteristics which are not clear.

The circle theory of Gershgorin (1931) adds a great amount of visual insight into the questioned properties. Let A be a nxn complex matrix. Then the spectrum of the eigenvalues of A is contained within the union

of the n circles in the complex plane

$$|z - a_{ii}| \leq \sum_{j=1}^{n}{}' \; |a_{ji}| \tag{2.7}$$

A consequence of this theory is the inequality

$$|\lambda| \leq \max_{i} \{ \sum_{j=1}^{n} |a_{ji}| \} \tag{2.8}$$

where λ is the characteristic value associated with A.

The concept of diagonal dominance can be pictorially represented in the complex plane using the Gershgorin theory. That is, for $a_{jj} < 0$ and

$$|a_{jj}| \geq \sum_{i=1}^{n}{}' \; |a_{ij}| = r_j \; ; \quad j = 1, \ldots, n \tag{2.9}$$

let

$$\delta_j = |a_{jj}| - r_j \; ; \qquad j = 1, \ldots, n \tag{2.10}$$

then $\delta_j \geq 0$ defines the distance between the jth Gershgorin circle and the imaginary axis in the complex λ-plane. If $\delta_j = 0$, then the eigenvalue map corresponds to a closed compartment. If all $\delta_j \neq 0$, $j = 1, \ldots, n$, then the system is open and the eigenvalues are nonzero. None of the Gershgorin circles intersect the imaginary axis.

Since the a_{jj} are all negative, $|a_{jj}| \geq r_j$, then all of the circles are in the left half plane, and correspondingly compartment models are stable.

Furthermore, except for notable exceptions (Wang, 1976), the eigenvalues of compartment models are assumed to be real (Hearon, 1963). Thus, the circles are centered on the negative real axis. This concept of the jth Gershgorin circle of a compartment model is illustrated in Figure 2.

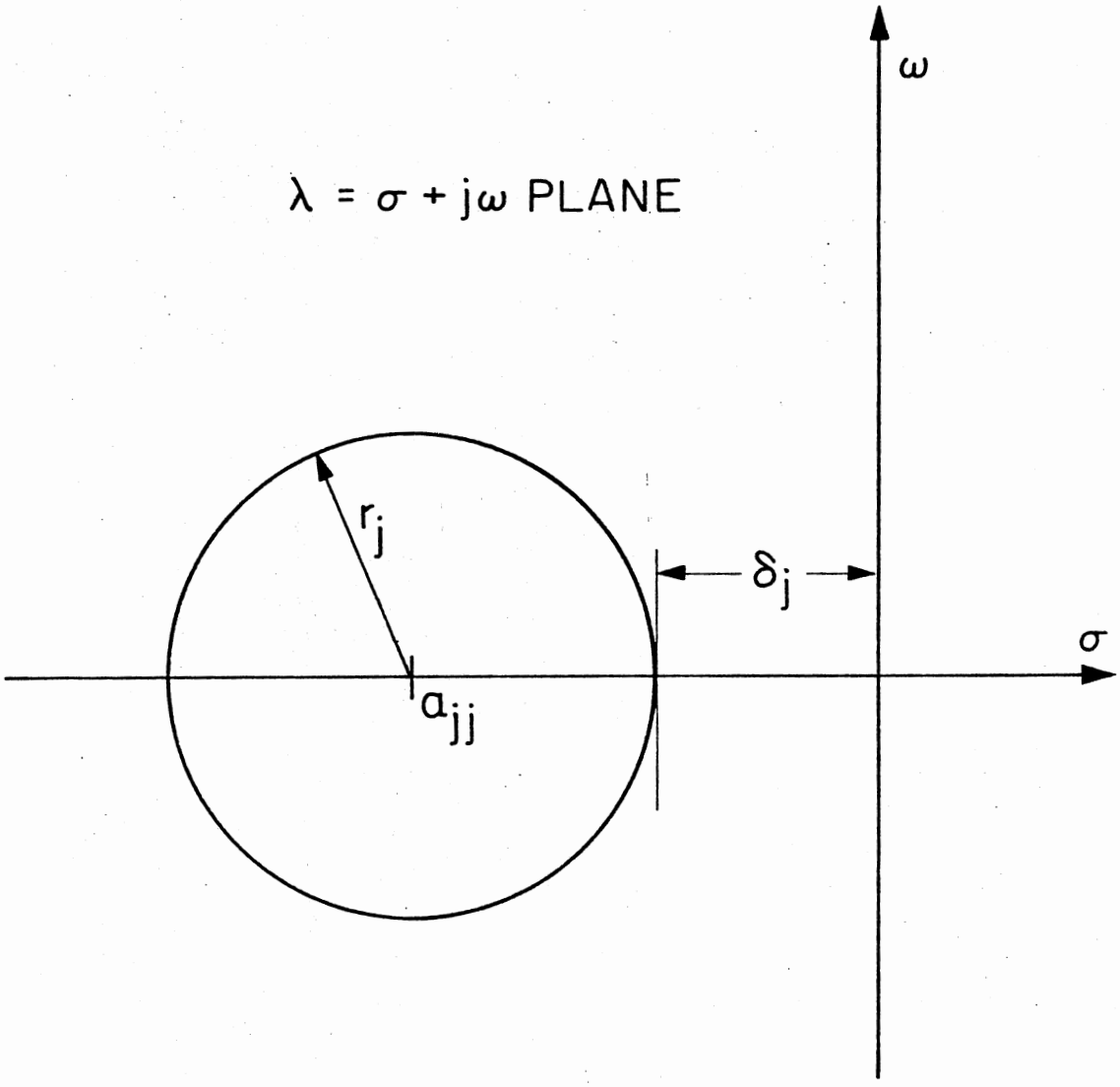

Figure 2. The jth Gershgorin Circle

One natural extension of the concept of the spectral graph of the circles is the idea of a bound on the responsive behavior of compartment models. Though necessary for stability, its identity is fundamentally important to many applications of this theory.

Sandberg (1978) has shown that compartment models have unique equilibrium points, if certain hypotheses are valid. More directly, Thron (1972) has shown that linear completely open compartmental systems have steady states, in response to certain driving input. Also under certain conditions, closed systems may have steady state responses.

With this concept of state the idea of bounds can be defined as bounds on the state trajectories between steady states. Completely open systems, in particular, obey the solution of Equation (2.4) (Bellman, 1960)

$$x(t) = \exp(At)\ x(0) \tag{2.11}$$

Therefore, it is obvious that the dynamic trajectories between non-negative (non-positive) steady states remain non-negative (non-positive). Coupled with the concept of superposition, it is plain that the trajectories do not overshoot the final steady state value. Otherwise a change of sign would occur. Carried one step further, it is clear that the transitions from one steady state to another are monotonic (Mulholland and Keener, 1974). A similar property has been proven by Sandberg (1978) for nonlinear compartment models.

Hence, the transitions between steady states are bounded by the final states. Still, the Gershgorin spectrum can be utilized to derive even better bounds on compartment models.

Frobenius (1912) has shown that any non-negative matrix A ($a_{ij} \geq 0$) has at least one dominant eigenvalue. This eigenvalue is referred to as

the Frobenius-Perron eigenvalue of A. Since the coefficient matrix of Equation (2.11) is non-negative (Bellman, 1960), there is a dominant eigenvalue for the state transition matrix associated with A. Once carried through the similarity transformation, it becomes evident that the matrix A has a dominant eigenvalue, and it is the least negative eigenvalue ($\delta = 0$). Gowdy (1978) has derived the same result as above through a linear shift of the spectrum associated with matrix A. Since the matrix A is finite, there also exists a minimum eigenvalue

$$\alpha = \max_j [r_j + |a_{jj}|] \tag{2.12}$$

together with the Frobenius-Perron eigenvalue

$$\delta = \min_j (\delta_j) \tag{2.13}$$

These bounds are expressed by

$$-\alpha < \lambda_1 < \lambda_2 < \ldots < \lambda_{n-1} < \lambda_n < -\delta \tag{2.14}$$

where λ is an eigenvalue of A. Combining these bounds with the concept of a taxicab norm

$$|x| = \sum_{j=1}^{n} |x_j| \tag{2.15}$$

it is easily shown that the solution of Equation (2.4) is bounded in norm by (Gowdy, 1978)

$$\exp(-\alpha t)\ |x(0)| \leq |x(t)| < \exp(-\delta t)\ |x(0)| \tag{2.16}$$

Thus the bounds suggested by the property of monotonicity can be conservatively identified as a function of time once the spectral bounds are known. This is possibly a better bound than the final states argued

earlier. These bounds are extremely useful in describing a range of behavior associated with only partially identified systems.

Another important property of compartment models is identifiability. Bellman and Astrom (1970) indicate that a matrix A is identifiable only if any change in the coefficients results in a change in the observations. Thus the properties of controllability and observability are also necessary.

Controllability and observability are widely accepted well known properties throughout systems science; therefore, they will not be discussed here. Suffice it to say, that the rank of the coefficient matrix is related to these quantities, and hence closed systems are uncontrollable (Johnson, 1976). They are also unidentifiable. Fortunately, the closed section of any compartment system can be separated from the open section and the identification and manipulation of the open section is feasible. Unfortunately, controllability and observability are necessary but not sufficient conditions for identifiability. The property of identifiability cannot be established (Gowdy, 1978).

Even without the sufficient conditions for a proof of identifiability, a great amount of research has still been conducted in this area. Many factors affect the observation of a system. For example, fixed digit observations add uncertainty in measurements. They also affect the sample rate.

Commonly, the sampling theory of Shannon (1948) establishes bounds on the sample period in the identification of systems. However, the unique properties of compartment models allow for a different sample theory to evolve. Specifically, sample bounds shown to have some meaning (Gowdy, 1978) are given by

$$-1/\delta \ln (1 - 10^{-d}) < \tau < (-1/\alpha n \ln (10^{-d}) \qquad (2.17)$$

where δ and α are eigenvalue bounds derived earlier, d is the number of significant digits in the observation, and τ is the sample period.

Summary

This chapter has discussed some of the rich historical background of compartment models. In particular, the structural properties of certain types of compartment models have been examined. Among these are diagonal dominance, monotonicity, finite bounds, and stability. Certainly, there are many other properties to discuss, but in the interest of moving ahead to consider compartment models with stochastic properties, the mentioned properties are sufficient.

The next chapter will examine compartment models with the added dimension of random variables. This type of model is unlike any mentioned in the literature, but it is felt that it more closely emulates the real world than the previous approximations. The structural properties of these models are then examined in Chapter V.

CHAPTER III

UNCERTAINTY AND IDENTIFICATION

Introduction

The initial hypothesis of compartment models is based upon identification considerations. The concept of the exchange of substance between entities is correct, but the consideration of models and representations of entities is limited by the accuracy of the identification of the system.

Chapter II has illustrated many theoretical rules for compartment models. For the most part, they consist of deterministic models for the systems. As such they are incomplete, if not incorrect.

Perhaps the most puzzling and frustrating experience of a novice experimentalist is the virtual impossibility of verifying deterministic physical laws via observations of the "real world." A journeyman learns the shady, not quite true, connection between the incomplete theory and the actual observations. An easy approach would be to blame the discrepancies on insignificant errors. Correction factors, experimental error are all widely recognized, but incorrectly their significance is overlooked. The practice of applying physical laws which observations verify, more or less, is a problem which must be resolved as thoroughly as possible. It is for this reason that real systems are analyzed by statistical means, if possible. Stochastic analysis is a must, not a frill.

A common practice of some scientists is to characterize observational discrepancies as the effect of higher order nonlinear terms not carried in the system model. There is no doubt that this analysis is partially correct in some systems. There is also a practice of accounting the discrepancies as faulty equipment errors, which is partially true in some systems as well. At any rate, whatever the reason, observations reflect an uncertain property of nature. This uncertainty is not random, but its effects are indistinguishable from the random properties of stochastic systems. Not just on sub-atomic levels, but on classical structure levels are the true descriptions of segments of matter fuzzy.

It is incorrect to call this uncertainty an error, if by error there is an implication of a true description. The term error implies the existence, with adequate machinery and capable experimentalists, of observations which could exactly verify the constructs, mathematical and mental, proposed by scientists. However, a different implication of error is as a measure of uncertainty. It is this second definition that is the norm for this work.

This work would argue that the deterministic rules of Chapter II are correct, but incomplete. Some statement of the uncertain qualities of nature must be reflected in the study of compartment models.

It is not suggested that all compartment models are stochastic. Indeed, the exception is more often true than the rule. However, the observations of compartment models are indistinguishable from the observations of the random experiment establishing stochastic systems. As a compromise, it is suggested that compartment models are the deterministic models suggested in the literature, but with properties found in stochastic systems. The compartmental states may be modeled successfully as

random variables, with distributions, moments, etc. Still, the relationships involving structure and conservation hold.

The effect of the stochastic properties is manifest most dramatically in the identification of the system, since it utilizes the observations directly. Since all connections with the system are after the identification, the stochastic properties must be reflected in controllers, filters, etc.

It is clear that the term "identification" must be defined. The meaning of a compartment model with stochastic properties must also be explored. This chapter will be one primarily of definition, and of exploration of the implication of uncertainty in identification. The vehicle for the exploration is an attempt at combining statistics with some deterministic theory to derive sample period bounds for compartment models.

The chapter will begin with an attempt at laying the groundwork for the definition of the compartment model. The concept of state and system will be re-examined with uncertainty in mind. From these the dynamical models for moments and terms necessary for identification will be surveyed. It is assumed that the concept of probability distributions is well known, and will be mentioned only in reference to the concept of the stochastic moments.

The chapter will end with an examination of a particular Monte Carlo identification experiment. The results give surprising statistical implications which are the motivation for the stochastic property analysis of this work.

Definitions

The mental identification of a system by an engineer is at first thought very clear. A finite set of machines, with interactions with the external environment, is not thought to be heuristically arbitrary.

However, mathematical models of that system are usually not clear. Simplicity begs the use of deterministic linear relationships with a sixth sense feeling for making a project useful. Sometimes the system is not easily constrained to constructs examined in theory, and the sixth sense is the only guide.

Possibly a helpful scientific approach would be in a three stage loop: observation, hypothetical identification, and manipulation. Engineers are often content if a hypothetical model, though possibly incorrect, renders the desired range of observation after manipulation.

Rarely are observations from inside a system available. Environmental compartment models may be an exception. Usually, the true knowledge of a system is a collection of connected data from observations of input influences and output effects.

This collection of input-output pairs is defined as a system by Zadeh and Polak (1969). Associated with this system is a bundle of relations known as an aggregate. Then a state is defined as a tag associated with the aggregate.

The definitions suggested by Zadeh and Polak (1969) are very mathematically pure. There is no arbitrary nature except for the misinterpretation of the scientist. This work would extend the definitions from a mathematician's idealness to possibly an engineer's compromise.

It could be suggested that observations are samples from an uncertain experiment. The observations are in turn uncertain. The definition

of the system is the connection of a set of finite samples from a group of uncertain events. (Notice the events are not random, merely unknown.) Associated with this set are statistical estimates from observations. The identity of the system validates the estimates. This is all that the observation establishes. From insight the scientist establishes a hypothetical model. Then the model is verified as fully as possible with mathematics. The terms associated with the estimates are known as states. The only connection is insight verified with mathematics.

This definition of a system is mathematically imprecise. However, it is a better representation of empirical systems. Observations must validate statistical significance arguments. Statements like "A state is the minimal information necessary to describe a system in all future time" are not necessary. Rather, a state can be considered as a random variable. The probability distribution of that state lends insight into the probability distribution of that state in future time.

It is clear that the connection between deterministic theory and stochastic theory for nonlinear systems is not known. For linear systems one could visualize the deterministic theory describing the deterministic moments from the stochastic nature of the system. The same is not true for nonlinear systems in general. Rather, it is believed that each particular nonlinear system must be examined for stochastic properties. Caution must be exhibited in applying stochastic theory for linear systems to linear approximations of nonlinear systems.

Once constrained to linear systems it becomes proper to discuss the moments of the state probability distributions, in particular the first few moments; mean, second moment, variance, and the variance of the variance are important, if the states are assumed to be Gaussian.

Defining the expectation operator by

$$E\{x^n\} = \int_{-\infty}^{\infty} \alpha^n \, p_x(\alpha) d\alpha \tag{3.1}$$

where $p_x(\alpha)$ is the probability density function of x, the first few moments are given by (Sage and Melsa, 1971)

$$\eta_x \triangleq E\{x\} = \int_{-\infty}^{\infty} \alpha \, p_x(\alpha) d\alpha \tag{3.2}$$

$$\psi_x \triangleq E\{xx'\} = \int_{-\infty}^{\infty} \alpha \, \alpha' \, p_x(\alpha) d\alpha \tag{3.3}$$

$$P_x \triangleq E\{(x - \eta_x)(x - \eta_x)'\} = \int_{-\infty}^{\infty} (\alpha - \eta_x)(\alpha - \eta_x)' \, p_x(\alpha) d\alpha \tag{3.4}$$

The next few moments are clearly tensors and difficult to express by matrix terms, as yet. In addition to the variance, the variance of the variance is useful, and is also a tensor.

Proceeding as in stochastic systems the continuous process is often described by the first order differential equation

$$\dot{x}(t) = A(t) \, x(t) + B(t) \, u(t) + G(t) \, w(t) \tag{3.5}$$

where x(t) is the state vector of the system, u(t) is the deterministic input vector to the system, w(t) is the random input vector to the system, and $\dot{x}(t)$ is the first derivative of the states x with respect to time. The coefficient matrices are deterministic and possibly time-varying. (Note that random and the notion uncertain are used interchangeably, since the effects are indistinguishable.)

The two input vectors u(t) and w(t) are a way of describing any type of stochastic input as a deterministic bias plus a zero-mean noise term. The term w(t) does not need to be white noise, but can be considered as white, if the dynamic model of any colored noise is appended to

the system model. Since the mathematics are simplified by the white noise assumption, it will be adopted with the previous comment in mind.

The dynamic transformations of the deterministic moments are described by

$$\dot{\eta}_x(t) = A(t)\,\eta_x(t) + B(t)\,u(t) \tag{3.6}$$

$$\dot{\Psi}_x(t) = A(t)\,\Psi_x(t) + \Psi_x(t)\,A'(t) + B(t)\,u(t)\,\eta_x'(t) + \eta_x(t)\,u'(t)\,B'(t) + G(t)\,V_w\,G'(t) \tag{3.7}$$

where $E'\{w(t)\,w(\tau)'\} = V_w\,\delta_d(t-\tau)$, and $E\{w(t)\} = 0$ due to the whiteness property of $w(t)$. The dynamic trajectory of the variance is described by

$$\dot{P}(t) = A(t)\,P(t) + P(t)\,A'(t) + G(t)\,V_w\,G'(t) \tag{3.8}$$

The terms ψ_x and P are matrices with the diagonal terms representing the notion of second moment and variance. The off-diagonal terms are expectation relations between different states. For this reason the matrix P is often known as the covariance of x. However, this is a source of confusion since

$$C_{xx}(t, \tau) = E\{(x(t) - \eta_x(t))(x(t+\tau) - \eta_x(t+\tau))'\} \tag{3.9}$$

is known as the covariance function of x, while

$$R_{xx}(t, \tau) = E\{x(t)\,x(t+\tau)\} \tag{3.10}$$

is known as the correlation function of x. The problem is one of convention, and the distinction between the covariance $P(t)$ and the covariance function $C_{xx}(t, \tau)$ is left to context.

It is not inconsistent to define

$$P(t) = c_{xx}(t, \tau) \qquad \text{for } \tau = 0 \tag{3.11}$$

Equations (3.9) through (3.11) are especially meaningful for stationary systems where the covariance and correlation functions are functions of the lag τ only.

The ideal moments are deterministic quantities. However, like the states, they have an uncertain aspect. They can only be known by estimates. In turn, the estimates can be considered as random events according to the rules of probability. The variance of the mean estimate is a function of the true variance. The variance of the variance estimate is a function of a term known as the variance of the variance. This term is a tensor if derived in a manner similar to the variance. However, it can be derived as a matrix, as will be suggested later.

The sources of the uncertainty in observations are of three types. They are state noise, input noise, and measurement noise. The clear separation of each source is often difficult, but heuristically they can be identified somewhat.

Measurement noise is a measure of the uncertainty associated with imperfect observation of a system, as opposed to the uncertainty which would be found using an ideal observor. That is, observations of a system could take the form

$$y = C\,x + v \tag{3.12}$$

where y is a linear combination of the states plus a random factor stemming from imperfect measurement, such as fixed digit samples, observer bias, etc.

It is assumed that v is white noise with any known biases or colored dynamics removed through calibration or compensatory techniques. Therefore,

$$E\{v\} = 0 \tag{3.13a}$$

$$E\{v(t)\ v'(\tau)\} = V_v \delta_d(t-\tau) \tag{3.13b}$$

The measurement noise is the unavoidable error of observations. Such errors do not affect the state of the system unless included in some type of feedback.

Input noise is the uncertain aspect of the driving forces. Any control or input like the system states have random properties, thus the white noise driving term w(t).

Finally, state noise is a random fluctuation of the states due to internal system sources. Such noise is typically thermal noise, eddy currents, or such. State noise terms are often minor in effect, or at least indistinguishable from the uncertainty effects of the driving forces. Therefore, they are often ignored.

Unfortunately, the individual contributions of the types of noise are difficult to identify in observations of the states. Indeed, the whole area of identification suffers drastically from the uncertainties of the observations. Signal-to-noise ratios reflect the relative magnitude of measurement noise. Very low SNR's render identification impossible, and therefore noise separation is also impossible. On the other hand, very high SNR's coupled with low system noise also render the random effects inseparable. Still, the region where such problems occur has noise magnitudes which are too high to ignore completely.

The separation of noise factors is just one problem area in identification. Often other difficulties are so severe as to overshadow the separation problem. At any rate, the identification of the system is a very important area of compartmental analysis.

For these reasons, the next section will discuss identification and the implication of uncertainties.

Identification and Deterministic Sample Theory

The term identification has several meanings. Appreciation of the complexity and uncertainty of real systems has rendered many simplifying connotations. For example, a method of forecasting future events from time series requires a great deal less information about the system originating the series than a method of forecasting future behavior in response to different conditions. In turn the system identifications required by each method may vary vastly. Yet, given the same design conditions the two models should agree to an acceptable significance. The two need not exactly agree because the system is after all uncertain and cannot be exactly specified.

Identification for this work will be constrained to finding the best linear model of the form of Equation (3.5) for a compartment model. The term "best" indicates the minimal variance model where a minimal variance model suggests the minimal spread of the model parameters around the estimate means.

Since the observations of systems are "estimates" of the system moments, the resultant identified model is an estimate of the actual system. Accordingly, these model estimates can be considered as random events with accompanying distributions. The variance of the estimates is one indicator of the "goodness" of the model.

Many factors affect the estimates. Instead of considering every factor, this work will consider the effect of sample frequency on the variance of the model estimates. Many of the other factors are tied in

with the sample frequency, and thus some insight will be gained in those areas as well.

The major constraints of sample theory have been the aliasing effects noted in the Nyquist sample bound for continuous band limited signals. Once the bandwidth, B, is known, a sample frequency of 2B can be shown by Fourier analysis to minimize modulation interference known as aliasing.

However, the effect of finite digit quantization is to create new sample bounds. Consider a simple exponential such as in Figure 3. It is a continuous-time continuous-amplitude signal which is not bandlimited. The effect of roundoff quantization of 3 binary bits is shown in Figure 4. Clearly the effect of changing the sample period is to increase the error in representing the exponential in identification. Smaller sample periods increase the correlation of consecutive samples, as shown in Figure 5. Larger sample periods increase the error by the loss of meaningful points in the representation, as shown in Figure 6.

These quantization bounds have nothing in common with the Nyquist bounds. Indeed the Fourier anlaysis lends no proof of their existence. Though the effect of quantization is appreciable in the Fourier spectra, the connection of frequency to amplitude is nonexistent.

However, as an indirect result, the discrete Fourier transform resolution is tied to the quantization. The DFT is calculated by assuming the signal once sampled, is modulated, and thus periodic with period n points. The finite Fourier transform is represented by

$$X(f_k, T) = T A_k \qquad k = \pm 1, \pm 2, \ldots \tag{3.14}$$

where

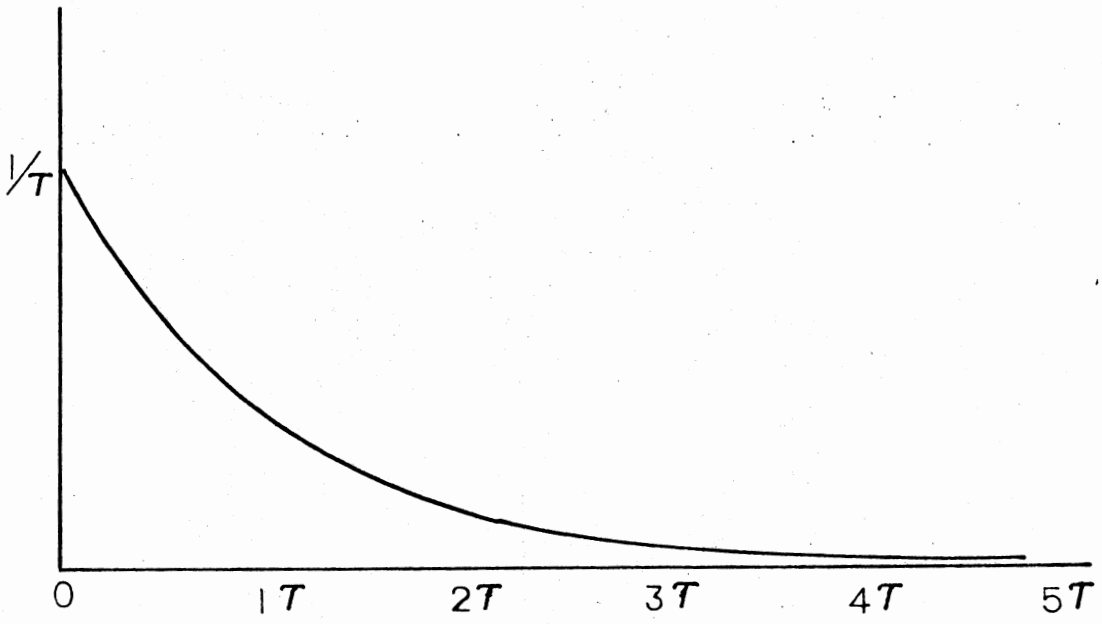

Figure 3. Continuous-Time Continuous-Amplitude Exponential

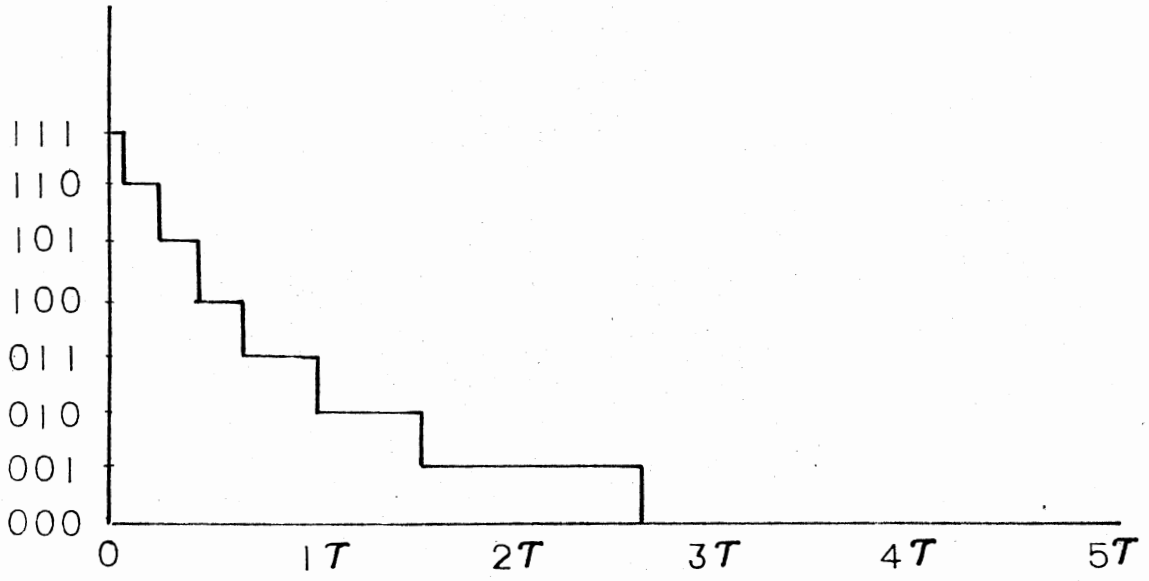

Figure 4. Continuous-Time Quantized Exponential

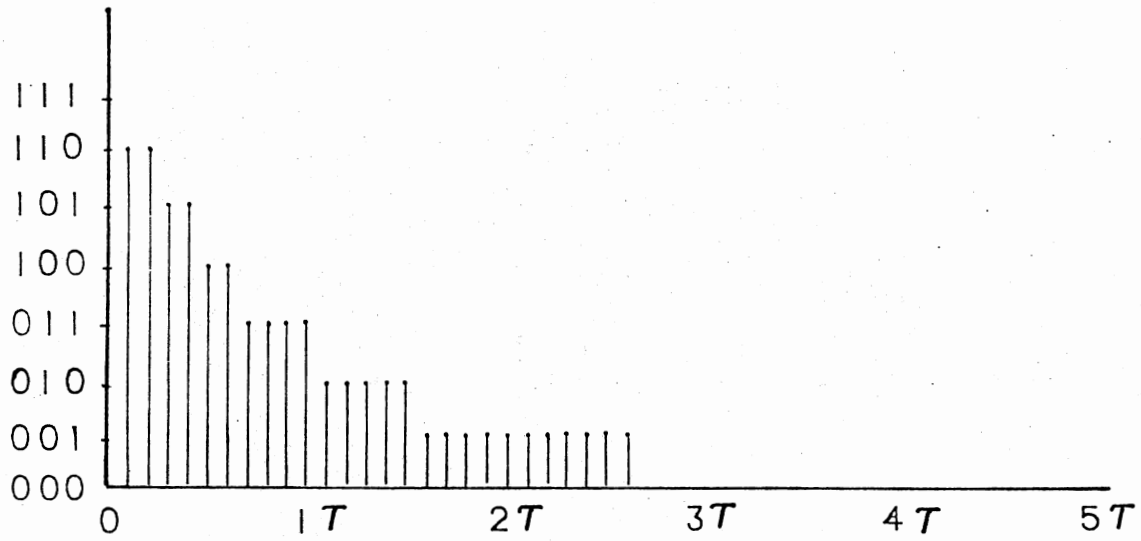

Figure 5. A Small Sample Period

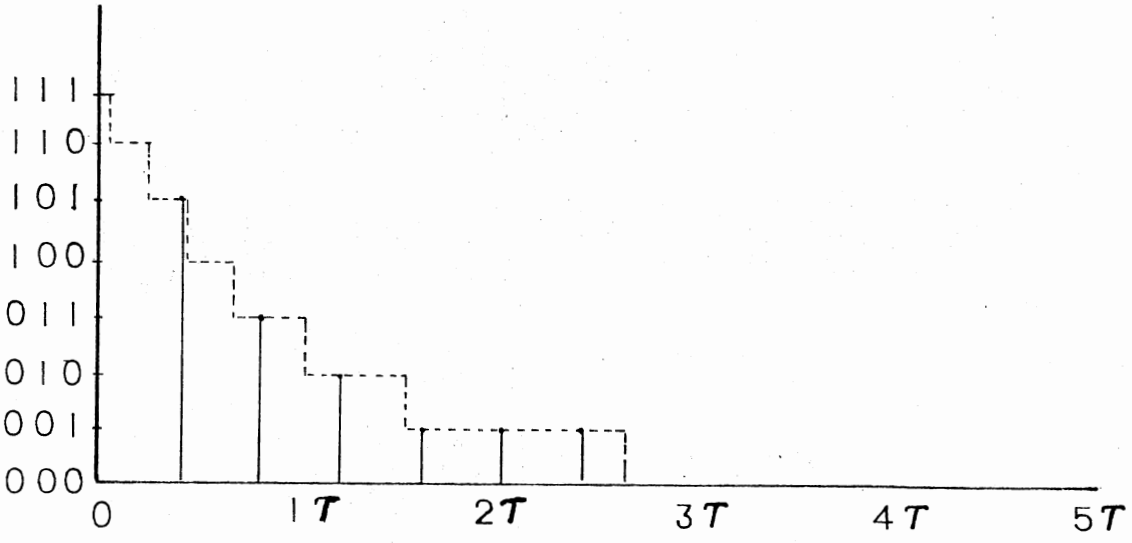

Figure 6. A Large Sample Period

$$X(f, T) = \int_o^T x(t) \exp(-j2\pi ft)\, dt$$

and

$$A_k = (1/T) \int_o^T x(t) \exp(-j2\pi f_k t)\, dt$$

Changing the sample period Δt where

$$T = n\, \Delta t \tag{3.15}$$

has the effect of changing the interval of the Fourier spectrum, and thus alters the finite Fourier transform.

Therefore, expanding the sample period with corresponding loss of significant points from the time series results in a corresponding loss of points in the Fourier spectrum. It has been argued that this bound is the Nyquist bound. Assuming continuous amplitude samples, it would appear to be so. However, discrete amplitude samples would render bounds other than the Nyquist frequency. In particular, an upper bound on the sample frequency would appear.

A method for testing the previous hypothesis was proposed by Gowdy (1978). Simple regression identification of a compartment model could be perturbed by noise, and bounds on the variance of the model estimates could be examined statistically.

Ecosystem tracer compartment models have many nice properties for identification. Typically, samples from each compartment of the system are available. The dynamics of the system are slow enough to make identification of the sample bounds feasible. Also, a priori information is usually available on bounds for the flow rates; that is, on the maximum and minimum eigenvalues.

Given bounds of α and δ on the eigenvalues of a particular compartmental ecosystem, Gowdy has proposed possible sample period bounds of

$$\tau^{+} = (-1/\alpha n) \ln 10^{-d} \tag{3.16}$$

and

$$\tau^{-} = (-1/\delta) \ln (1 - 10^{-d}) \tag{3.17}$$

where d is the precision of the samples, $\alpha/2 = t_o^{-1}$, δ is the minimum transfer out of the system, and t_o is the maximum turnover of the system.

These bounds are derived from a set of deterministic norm bounds on the states of the system given by

$$\exp(-\alpha t) \, |x(0)| < |x(t)| < \exp(-\delta t) \, |x(0)| \tag{3.18}$$

where $|x(t)|$ is the taxicab norm of $x(t)$. More will be said of the norm bounds and the derivation later. It is sufficient to note that the properties given in the previous chapter culminate in possible identification sample bounds, as proposed by Gowdy (1978).

Given these bounds an identification experiment could be proposed. An ecosystem could be brought to steady state with a constant infusion of tracer material in the input. Then the input with tracer could be reduced by a set amount giving a response of the type of Figure 3 for each compartment. Again the type of response is known from the body of structural properties established in Chapter II.

With superposition a deterministic model of the form of

$$\dot{x} = A \, x \tag{3.19}$$

could be proposed where n is the order of the system and the rank of A.

The solution of Equation (3.19) is well known by

$$x(t) = \Phi(t)\ x(0)$$

where $\Phi(t) = \exp(At)$.

In general, consecutive samples of the states would be related by

$$x(i) = \Phi(\tau)\ x((i-1)\tau)$$

where i is the index of the sample taken at uniform intervals.

Since τ is constant, n equations could be formed from the n+1 set of observations.

$$x(\tau) = \Phi(\tau)\ x(0) \tag{3.22a}$$

$$x(2\tau) = \Phi(\tau)\ x(\tau) \tag{3.22b}$$

$$\vdots$$

$$x(n\tau) = \Phi(\tau)\ x((n-1)\tau) \tag{3.22c}$$

The n^2 equations of Equation (3.22) could be expressed by the relation

$$X_2 = \Phi(\tau)\ X_1 \tag{3.23}$$

where

$$X_1 = \{x(0),\ x(\tau),\ \ldots,\ x((n-1)\tau)\}$$

$$X_2 = \{x(\tau),\ x(2\tau),\ \ldots,\ x(n\tau)\}$$

$$\Phi(\tau) = X_2 X_1^{-1} \tag{3.24}$$

Equation (3.24) is simple nth-order zero-mean regression.

The model parameter A is found from

$$A = (1/\tau)\ \ln\ (X_2 X_1^{-1}) \tag{3.25}$$

where $\ln\ (X_2 X_1^{-1})$ is found by similarity transformations similar to $\exp(At)$. Since the ecosystem is linear and probably contains real eigenvalues, the matrix A is unique.

The regression of Equation (3.25) was built into a Monte Carlo simulation package. The samples observed from a typical system could be perturbed by zero-mean uniformly-distributed white noise, and model estimates could then be examined.

The hypothetical ecosystem chosen for the simulation was first described by Smith (1970). Though Smith's model is not a real system model, it is built upon realistic numerical parameters. The system response mimics an aquatic ecosystem with phosphorous loading.

Smith's model consists of three compartments: water (x_1), aquatic plants (x_2), and a herbivore population (x_3). The ecosystem input (u_1) is the phosphorous in the water flowing through the system. The output is the phosphorous in the exiting water, and the herbivores leaving the system through migration. The system interaction is illustrated in Figure 7.

Smith presents typical data for such a system, and from these data the rate coefficients were directly computed. The result is the following model:

$$\dot{x}_1 = 5x_2 + 5x_3 - 16x_1 + u_1 \tag{3.26a}$$

$$\dot{x}_2 = 14x_1 - 95x_2 \tag{3.26b}$$

$$\dot{x}_3 = 90x_2 - 14x_3 \tag{3.26c}$$

The input rate u_1 is 100 mg P per day in steady state. The rate coefficients are in terms of reciprocal days.

The steady states were calculated for model Equation (3.26), and then Smith's model was simulated on the IBM-370 computer using the CSMP simulation language package. The initial conditions were the steady

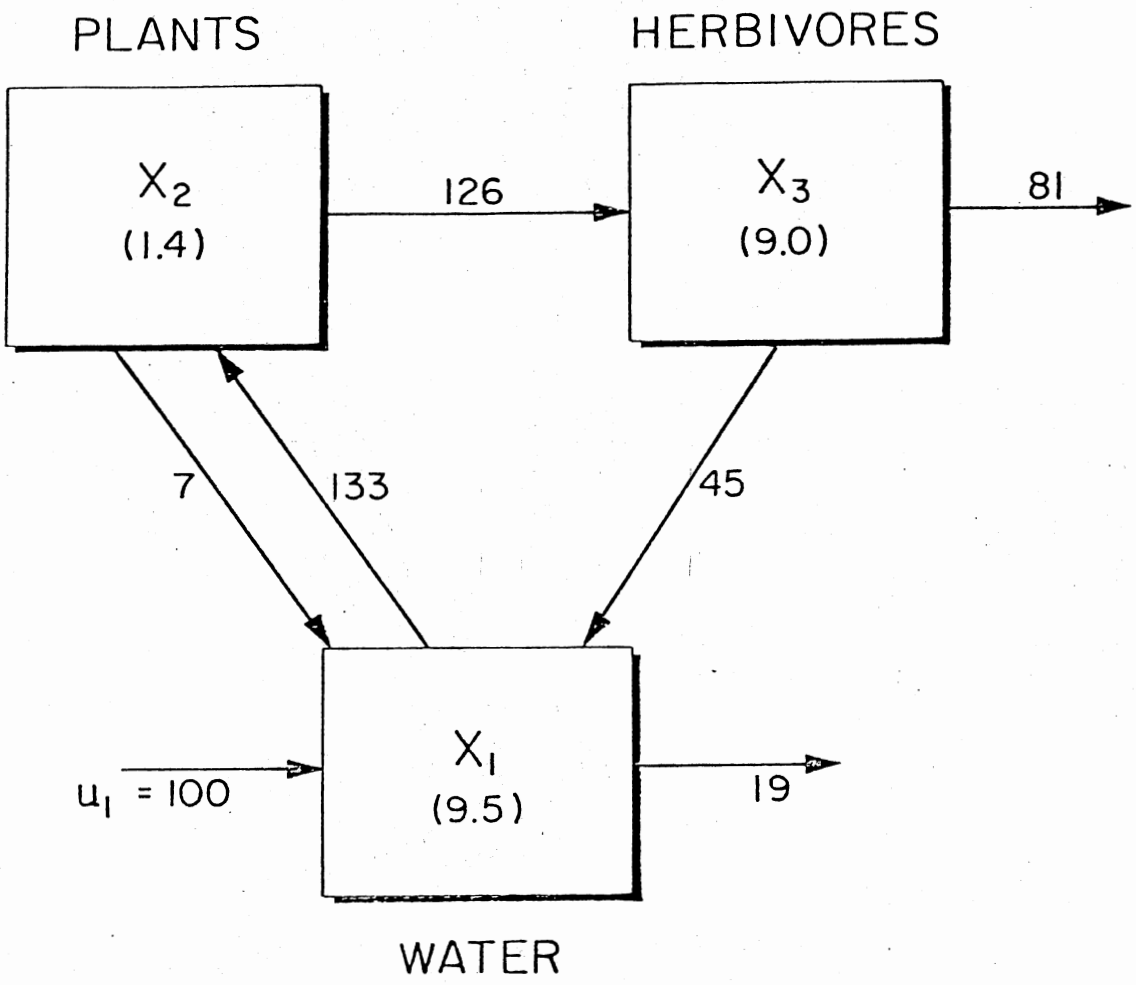

Figure 7. Smith's Model

states for a constant input (u_1) of 100. The input of the system ($t > 0$) was $u_1 = 0$. This emulates the basic experiment design.

Data were produced using a fourth-order Runge-Kutta integration routine with variable step size. Samples were produced at 0.001 time intervals for ten units of time. The integration routine was checked with a matrix exponential routine developed by Jessy Grizzle from a program written by W. Emmanuel at Oak Ridge National Laboratories.

The resultant data were in floating point with a mantissa of the form of x.xxxx. This gives approximately five normalized significant digits (d). Since the eigenvalue bounds for the model were $\alpha = 120$ and $\delta = 2$, the sample bounds were calculated to be $\tau^- = 5 \times 10^{-5}$ and $\tau^+ = 0.02$. Finally, 20 data sets were obtained at sample periods between 0.005 and 0.120.

The identification method requires quantized samples of the system states. The effect of this quantization is to subtrace an unknown quantity of value

$$0 \leq u \leq 10^{-4}$$

Therefore, a pseudo-random number generator was designed to simulate random numbers, uniformly distributed between zero and 10^{-4}. These uncertainties were added to the data to simulate the quantization effect. This provided the perturbations necessary for the statistical analysis. Ensemble averages of the perturbed parameters would yield the statistical sensitivity of the parameters. This was the basis of the Monte Carlo simulation.

The random quantities were denoted by w, where

$$0 \leq w \leq 10^{-4} \tag{3.27}$$

Using standard statistical tests with a 0.05 level of significance, it was determined that 400 samples were required to estimate the mean and variance of w. Four hundred runs of the identification algorithm were computed for each sample period. The moment estimates were computed from

$$\eta = (1/n) \sum_{i=1}^{n} a_i \tag{3.28}$$

and

$$\sigma^2 = (1/(n-1)) \sum_{i=1}^{n} (a_i - \eta)^2 \tag{3.29}$$

Two entries in the coefficient matrix were determined to exemplify the range of behavior of the parameter ensembles. The mean and variance were therefore calculated for entries a_{11} and a_{12} in the matrix A. The results are tabulated in Tables I and II. Plots of the means and variances of coefficients a and a versus sample period are graphed in Figures 8 through 11.

A Chi-squared goodness-of-fit test of normalcy was performed on the identified parameter ensembles. The region of acceptance for the sample statistic was

$$\chi^2 < \chi^2_{n,\alpha} \tag{3.30}$$

where the degrees of freedom (n) was 17, and the level of significance (α) was 0.05. The results are also presented in Tables I and II.

Examinations of the tables and graphs reveals that the mean does not differ from acceptable limits of over the sample period range. However, the parameter variance estimates do exhibit the sample period bounds from the theory. The bounds exhibited by the plots are

TABLE I

MONTE CARLO RESULTS FOR COEFFICIENT a_{11}

τ	$\hat{\mu}_a$	$\hat{\sigma}_a^2$	X^2
0·005	−16·024	$1·39 \times 10^{-2}$	32·4
0·006	−15·973	$7·90 \times 10^{-3}$	31·4
0·007	−15·964	$4·98 \times 10^{-3}$	36·2
0·008	−16·057	$3·43 \times 10^{-3}$	35·3
0·009	−16·039	$2·51 \times 10^{-3}$	41·7
0·010	−16·010	$1·95 \times 10^{-3}$	41·2
0·011	−16·010	$1·57 \times 10^{-3}$	31·1
0·012	−15·983	$1·31 \times 10^{-3}$	25·7
0·013	−16·022	$1·12 \times 10^{-3}$	23·6
0·020	−16·016	$6·38 \times 10^{-4}$	8·2
0·030	−15·992	$6·39 \times 10^{-4}$	7·0
0·040	−16·050	$9·89 \times 10^{-4}$	13·7
0·050	−15·997	$1·86 \times 10^{-3}$	25·2
0·060	−15·965	$3·95 \times 10^{-3}$	10·8
0·070	−16·007	$9·67 \times 10^{-3}$	15·1
0·080	−15·780	$2·02 \times 10^{-2}$	20·9
0·090	−16·249	$2·14 \times 10^{-2}$	6·8
0·100	−16·190	$3·40 \times 10^{-2}$	14·5
0·110	−16·178	$5·57 \times 10^{-2}$	19·7
0·120	−16·368	$9·33 \times 10^{-2}$	55·7

TABLE II

MONTE CARLO RESULTS FOR COEFFICIENT a_{12}

τ	$\hat{\mu}_a$	$\hat{\sigma}_a^2$	X^2
0·005	5·147	2·655	27·2
0·006	4·793	1·194	23·3
0·007	4·541	0·627	20·6
0·008	5·578	0·372	19·2
0·009	5·404	0·241	21·2
0·010	5·035	0·168	18·2
0·011	5·184	0·124	16·5
0·012	4·888	0·096	16·9
0·013	5·165	0·077	13·9
0·020	5·113	0·033	5·7
0·030	4·947	1·028	20·8
0·040	5·322	0·040	14·3
0·050	4·984	0·076	22·8
0·060	4·786	0·145	14·4
0·070	5·046	0·345	12·5
0·080	3·680	0·701	19·5
0·090	6·516	0·781	9·1
0·100	6·160	1·262	12·9
0·110	6·081	2·096	17·7
0·120	7·306	3·485	45·1

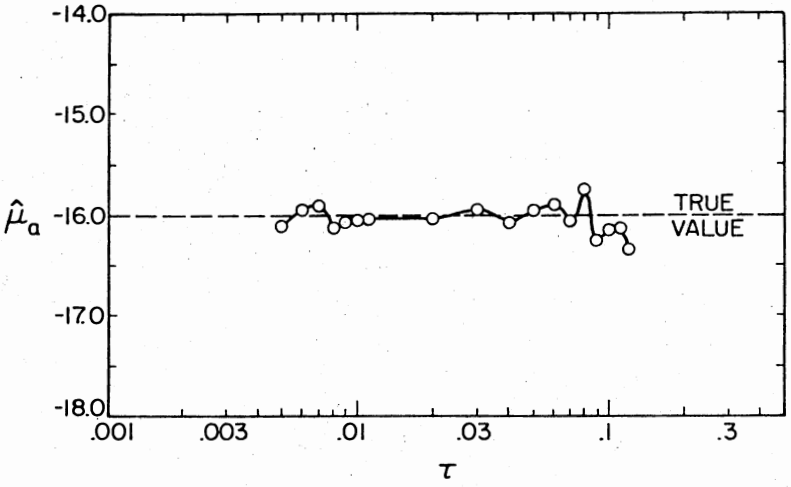

Figure 8. Estimated Mean of a_{11} Versus Sample Period

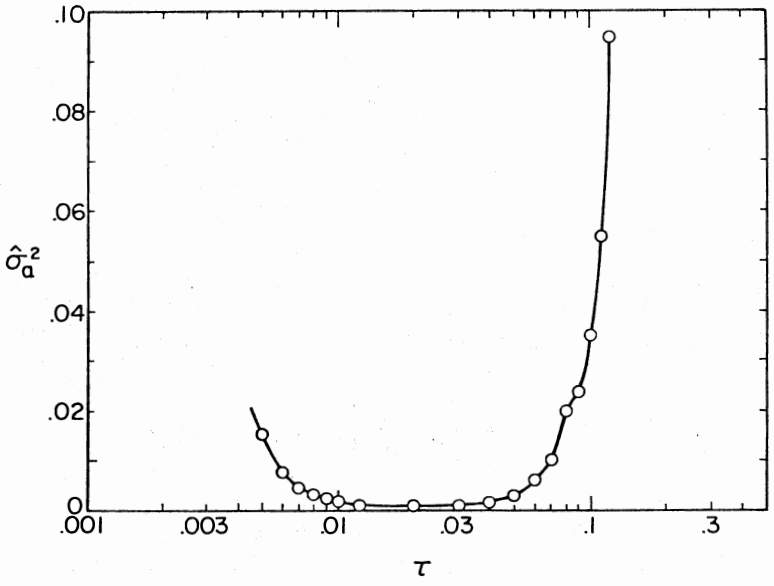

Figure 9. Estimated Variance of a_{11} Versus Sample Period

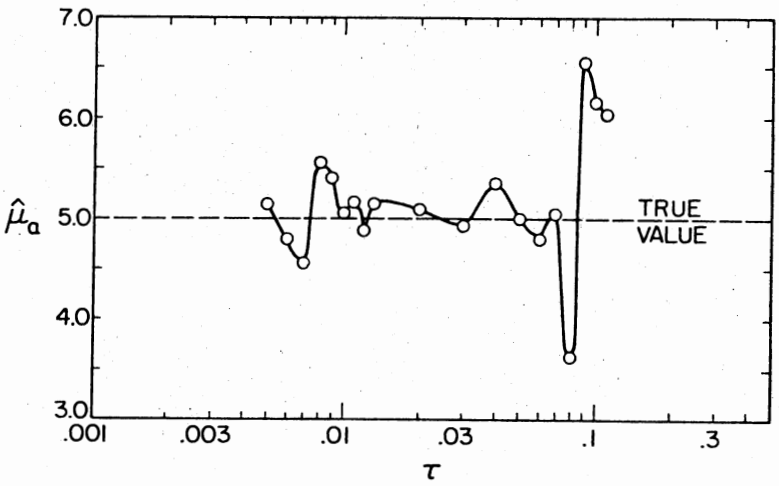

Figure 10. Estimated Mean of a_{12} Versus Sample Period

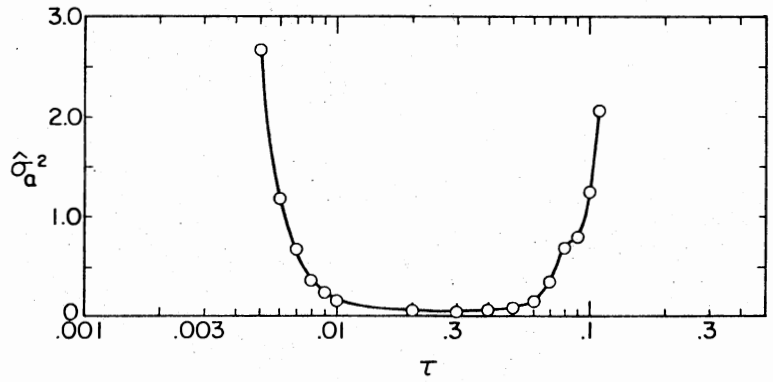

Figure 11. Estimated Variance of a_{12} Versus Sample Period

approximately 0.007 and 0.07. Since the random nature of the real world is not considered by the sampling theory, the shift of the sample period bounds is not inconsistent.

The spectral plots of Liff and Wolf (1966) illustrate a similar sample period range. They performed identification with z-transforms using a normalized integral squared error for a parameter fit criterion. Uncertainty was introduced in the response through Gaussian white noise superimposed on the output. The NIS was plotted against the sample period as in the graphs presented here.

Astrom examines a similar topic using an exponential fitting method. Since quantization was not in effect, the bounds were not noticeable in his graphs, other than the Nyquist related bound.

Summary

This chapter has been one of definition. The definition of a compartmental system has been examined from the viewpoint of an uncertain world.

The meaning of a system and the states associated with the system has been examined in hopes of relating observation with mathematics. The result is a compromise without the preciseness typical of modern deterministic system theory. But, the uncertainty could be approached as being a stochastic property with resultant moments and probability theory.

The motivation for such definitions has been the connection of compartment models with the identification of compartmental systems. Therefore, the identification properties have been stressed heavily. In particular, the existence of sample period bounds from the quantization

effects plus behavior bounds was examined. Spectacularly enough, it was found that upper and lower bounds on the sample period exist. The approach of identifying the bounds from deterministic theory was not satisfactory, though it did suggest a meaningful argument for their location.

The following chapters will examine the compartment models with stochastic properties defined in this chapter. In particular, stochastic bounds on the moments will be related to sample period bounds in an attempt at a more satisfactory examination of the sample bounds' identity.

CHAPTER IV

A KRONECKER ALGEBRA WITH APPLICATIONS

Introduction

The extension of the structural properties of compartment models relies chiefly on a matrix-to-vector transformation. This transformation is simply the result of rewriting the matrix equations into vector equations. Mathematically, this is represented by a Kronecker algebra.

The key structure of a Kronecker algebra is the Kronecker product. All of the identities involved in the algebra are derived from the product and its properties. Bear in mind that the Kronecker algebra is so termed to embrace a set of matrix equation identities. No connection with a mathematically defined "algebra" is intended.

Several interesting Kronecker identities already exist or are easy extensions of the known identities. Since they are crucial to this study, they are listed in the Appendix. This is not an exhaustive compilation, but rather a collection which has some remote connection to the general area of compartmental analysis.

There is some need for a set of definitions for the Kronecker identities. Particularly, the Kronecker sum is in need of a set form.

This chapter will consist of an examination of the basic Kronecker structures, and then apply these structures to certain applications which are more general than compartment models. The principle of invariant imbedding can then be applied. These applications include the Kronecker

form for the general linear matrix differential equation, and an extension of the stochastic matrix theory first proposed by Bellman (1960). In particular, the Kronecker series will be established to simplify the decomposition of multiple unit Markov matrices.

The Kronecker Structure

There are mainly two Kronecker structures which are important to this study. They are the Kronecker product and the Kronecker sum. A Kronecker product is defined to be

$$A \otimes B = [a_{ij} B] \tag{4.1}$$

where A is a nxn matrix and B is a mxm matrix, and the Kronecker product is a nmxnm matrix. For clarity, the product for two 2x2 matrices would be expressed by

$$A \otimes B = \begin{bmatrix} a_{11} b_{11} & a_{11} b_{12} & a_{12} b_{11} & a_{12} b_{12} \\ a_{11} b_{21} & a_{11} b_{22} & a_{12} b_{21} & a_{12} b_{22} \\ a_{21} b_{11} & a_{21} b_{12} & a_{22} b_{11} & a_{22} b_{12} \\ a_{21} b_{21} & a_{21} b_{22} & a_{22} b_{21} & a_{22} b_{22} \end{bmatrix}$$

A distinct pattern is visible in the resultant matrix. Any row or column builds in a set positional pattern. If the elements of A and B are assigned modulo values, a numeric code is visible from the first row or column.

The Kronecker sum is defined by

$$A \oplus B' = A \otimes I_m + I_n \otimes B \tag{4.2}$$

where A is again an nxn matrix, B is a mxm matrix, and I_j represents a

jth-order identity matrix. An example of a Kronecker sum for 2x2 matrices would be

$$A \oplus B' = \begin{bmatrix} a_{11} + B_{11} & b_{12} & a_{12} & 0 \\ b_{21} & a_{11} + b_{22} & 0 & a_{12} \\ a_{21} & 0 & a_{22} + b_{11} & b_{12} \\ 0 & a_{21} & b_{21} & a_{22} + b_{22} \end{bmatrix}$$

Again, if the elements of the Kronecker sum were assigned modulo values, a numeric code would be apparent. This code will be useful in the study of Markov equations.

Once the structures are defined, it is simple to derive many of the identities shown in the Appendix. From these identities, many important relations with familiar matrix equations can be derived.

The General Linear Matrix Equation

The vector matrix equation

$$\dot{x}(t) = A(t)\ x(t) \tag{4.3}$$

where $x(t)$ is a n vector, A is a possibly time varying nxn matrix, and $\dot{x}(t)$ is the time rate of change of $x(t)$; has found great use in modern engineering systems science. In general, Equation (4.3) may be considered to be a variation of the linear matrix equation illustrated in Brockett (1970) by

$$\dot{X}(t) = A(t)\ X(t) + X(t)\ B(t) \tag{4.4}$$

where X, A, and B are now time-varying nxn matrices, and $\dot{X}$ is the time rate of change of matrix X.

The Kronecker product and sum are useful in analyzing Equation (4.4).

But, first a structural tool is necessary for reforming the matrices into vectors. A stacking operator is defined as a matrix-to-vector transformation which rewrites the elements of a matrix X into a vector x. A row stacking operator would form a column vector x from the n rows of X concatenated in descending order. The notation for the row stacking operator would be x_r, and the result for a 2x2 matrix would be

$$x_r = (x_{11} \quad x_{12} \quad x_{21} \quad x_{22}) \tag{4.5}$$

Similarly, a column stack would be columns concatenated into a column vector.

$$x_c = (x_{11} \quad x_{21} \quad x_{12} \quad x_{22}) \tag{4.6}$$

With these operators, Bellman (1960) has illustrated the Kronecker form of the matrix equation

$$A\,X + X\,B = C \tag{4.7}$$

is

$$(A \oplus B')\; x_r = c_r \tag{4.8}$$

Barnett (1973) has carried the work one step farther. He has shown that the time invariant form for Equation (4.4) can be represented by

$$\dot{x}_r(t) = (A \oplus B')\; x_r(t) \tag{4.9}$$

Equation (4.9) is another notation for Equation (4.3). The solution of the homogeneous Equation (4.9) is

$$x_r(t) = \exp((A \oplus B')t)\; x_r(0) \tag{4.10}$$

where $x_r(0)$ is the initial condition vector for $x_r(t)$, with

$$X(t) = \exp(At)\; X(0)\; \exp(Bt)' \tag{4.11}$$

Theorem 4.1

The work of Bellman and Barnett can be viewed as special cases of the Kronecker form of Equation (4.4) given by

$$\dot{x}_r(t) = (A(t) \oplus B'(t))\, x_r(t) \tag{4.12}$$

Theorem 4.2

The state transition matrix of Equation (4.12) is the Kronecker product of the state transition matrices of the subsystems corresponding to matrices A and B. That is,

$$\Psi(t) = \Phi_A(t) \otimes \Phi_B(t) \tag{4.13}$$

where

$$x_r(t) = \Psi(t)\, x_r(0) \tag{4.14}$$

and

$$\dot{\Phi}_A = A(t)\, \Phi_A \tag{4.15}$$

$$\dot{\Phi}_B = B(t)\, \Phi_B \tag{4.16}$$

Proof

Differentiating Equation (4.13) gives

$$\dot{\Psi} = \dot{\Phi}_A \otimes \Phi_B + \Phi_A \otimes \dot{\Phi}_B$$

From Equations (4.15) and (4.16)

$$\dot{\Psi} = (A\, \Phi_A) \otimes \Phi_B + \Phi_A \otimes (B\, \Phi_B)$$

Since

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$$

$$\dot{\Psi} = (A\ \Phi_A) \otimes (I\ \Phi_B) + (I\ \Phi_A) \otimes (B\ \Phi_B)$$

$$= (A \otimes I)(\Phi_A \otimes \Phi_B) + (I \otimes B)(\Phi_A \otimes \Phi_B)$$

$$= (A \otimes I + I \otimes B)(\Phi_A \otimes \Phi_B)$$

Finally,

$$\dot{\Psi} = (A \oplus B')\ \Psi$$

which is satisfied for the state transition matrix of Equation (4.12). Furthermore,

$$\Psi(t_o, t_o) = I \tag{4.17}$$

from observation of the structure of the Kronecker product and the definition of Φ_A and Φ_B. Therefore, the theorem is proven. It is again left to the Appendix to establish the equality between Equation (4.13) and

$$X(t) = \Phi_A(t)\ X(0)\ \Phi_B'(t) \tag{4.18}$$

Once the solution of the homogeneous form (Equation (4.4)) has been found, it is easy to establish the full linear matrix differential equation

$$\dot{X}(t) = A(t)\ X(t) + X(t)\ B(t) + D(t)\ C(t)\ E'(t) \tag{4.19}$$

with Kronecker form

$$\dot{x}_r(t) = (A(t) \oplus B'(t))\ x_r(t) + (D(t) \otimes E(t))\ c_r(t) \tag{4.20}$$

The solution of Equation (4.19)

$$X(t) = \Phi_A(t)\ X(0)\ \Phi_B'(t) + \int_o^t \Phi_A(t,\tau)\ D(\tau)\ C(\tau)\ E'(\tau)\ \Phi_B'(t,\tau)\ d\tau \tag{4.21}$$

has its alternate form

$$x_r(t) = (\Phi_A(t) \otimes \Phi_B(t))\ x_r(0)$$
$$+ (\Phi_A(t,\tau) \otimes \Phi_B(t,\tau))\ (D(\tau) \otimes E(\tau))\ c_r(\tau)\ d\tau \qquad (4.22)$$

where D, E, and C are compatible matrices.

The Lyapunov Derivative

An important variation of Equation (4.4) is the Lyapunov derivative given by

$$\dot{P}(t) = A\ P(t) + P(t)\ A' + G(t)\ Q(t)\ G'(t) \qquad (4.23)$$

where P is a nxn symmetric matrix. Clearly, the row or column stack operators can be used, and the solution of Equation (4.23) is of the form of Equation (4.21) or Equation (4.22). The Lyapunov derivative will be useful in propagating the second moment for compartment models.

Notice should be made of the eigenvalues of the Kronecker product and sum. The eigenvalues of the Kronecker product of two n-dimension matrices A and B are the n-squared values formed by the combinations of the products of the eigenvalues of A with the eigenvalues of B. The eigenvalues of the Kronecker sum are similarly the combinations of the sums of the eigenvalues of A with the eigenvalues of B. Clearly, the eigenvalues of the Lyapunov equation are repetitive.

The Markov Process

The matrix structure of a Kronecker sum can be viewed as being analogous to the transition maps of sequential theory. One application of this characteristic is in the decomposition of multiple unit Markov processes.

The states defined by Markov models are finite probabilities associated with specific process conditions. Examples of such states would be the probabilities of certain numbers of population members in a set area, or the probabilities of certain configurations of health states in a population.

The Markov equations describe the transitions from one state to another. Though the canonical notation differs, the equations are a form of the compartment model. The classical Markov process notation would be

$$\dot{p}(t) = A'(t)\ p(t) \tag{4.24}$$

where p is the n vector of probabilities associated with the finite n conditions of the process, $a_{ij}(t)$ is the time-varying transition rate from state i to state j. Since the states are finite probabilities, the strict diagonal negative equality holds:

$$a_{ii}(t) = -\sum_{j=1}^{n}{}' a_{ij}(t) \tag{4.25}$$

The solution of Equation (4.24) is well known by

$$p(t) = \Psi(t)\ p(0) \tag{4.26}$$

where $\Psi(t)$ is the state transition matrix associated with $A'(t)$, and $p(0)$ is the initial condition vector of $p(t)$.

Commonly, Markov processes are applied to describing the probability relations of a group of population members. Given that each member j may have n_j health conditions and that there are m members in the population, there are L states in the full Markov model, and

$$L = \prod_{j=1}^{m} n_j \tag{4.27}$$

Obviously

$$L = n_m \prod_{j=1}^{m-1} n_j \tag{4.28}$$

or

$$Lm = L/n_m = \prod_{j=1}^{m-1} n_j \tag{4.29}$$

Neglecting epidemics or catastrophes, the members of a population can be considered to be mutually independent. The probability of a simultaneous change in health among the population can be assumed to be zero. For this study, the states are assumed to be communicating. The effects of an irreversible change such as death are assumed to be an extension beyond this work.

The state assignments are arbitrary. For convenience, they can be set into an order of decreasing health for the entire population. Certain states may be considered as equivalent and merged, but for this study the full unmerged model is desired. The merged form could always be found from permutations of the results.

A specific state assignment is achieved if each condition of member j is assigned a modulo n_j value with the order of decreasing value with decreasing health. A numeric code is then formed if each unit is assigned a positional value building to the right in the state designation. Though the position of a particular member is arbitrary, the form of the Markov equation becomes very clear.

Theorem 4.3

Given a m unit Markov process with communicating states, and that

unit m of the process is independent of the rest of the process with a Markov process relation of its n_m states

$$\dot{p}(t) = F'_m(t)\ p(t)$$

the full process coefficient matrix can be represented by

$$A'_L(t) = F'_m(t)\ A_{Lm}(t) \tag{4.30}$$

where $A'_{Lm}(t)$ is the coefficient matrix of the Markov relation for the system without unit m.

Proof

Let the Markov relation be ordered as proposed above. Then matrix $A'_L(t)$ can be separated into two matrices

$$A'_L(t) = B(t) + C(t)$$

where B is the matrix of coefficients corresponding to unit n's changes in status. Matrix C is the matrix of coefficients corresponding to the rest of the process.

Matrix B is of the form of n-squared partitions of size (Lm)x(Lm). Each partition is diagonal due to the state order and the restriction on the possible status changes for the system. Each off-diagonal partition is of the form

$$B_{ij} = f_{ki}\ I_{Lm}$$

where f_{ki} is the transition rate corresponding to a particular change in the health of unit m from state i to state k.

Due to the diagonal equality, the diagonal partitions are

$$B_{ii} = f_{ii}\ I_{Lm}$$

Therefore, matrix B corresponding to the addition of unit m to the population is

$$B(t) = F'_m(t) \otimes I_{Lm}$$

by definition of the Kronecker product.

Matrix C is block diagonal. Each (Lm)x(Lm) partition of C corresponds to the system without unit m. Obviously the off-diagonal blocks are zero given the state ordering. The n_m diagonal blocks are equal and correspond to changes in any of the other members in the population while keeping unit m at a constant health. Therefore,

$$C(t) = I_{n_m} \otimes A'$$

by definition of the Kronecker product. Finally, the theorem is proven by definition of the Kronecker sum.

Corollary 4.3a

Given a m unit Markov process with communicating states, and that every unit j of the process is mutually independent, the process coefficient matrix can be represented by

$$A'_L(t) = \sum_{j=1}^{m} \oplus F'_j(t) \tag{4.31}$$

where

$$\sum_{j=1}^{m} \oplus F'_j = F'_m \oplus F'_{m-1} \oplus \ldots F'_2 \oplus F'_1$$

Proof

The corollary follows naturally from continuing the theorem through the separation of each member.

Theorem 4.4

Given the hypothesis of Theorem 4.3, the state transition matrix of the Markov chain corresponding to a n-unit Markov process can be represented by

$$\Psi_L(t) = \phi_{n_m}(t) \otimes \Psi_{Lm}(t) \tag{4.32}$$

Proof

Given the Kronecker sum form of Equation (4.29), Equation (4.32) follows directly from the so-called "trivial" result of Theorem 4.2. The identity matrices must be kept consistent with the operation rank, but otherwise the proof is identical to the proof of Theorem 4.2.

Corollary 4.4a

Given the hypothesis of Corollary 4.3a, the Markov chain state tranition matrix can be represented by

$$\Psi_L(t) = \prod_{i=1}^{m} \otimes \Phi_i(t) \tag{4.33}$$

where

$$\prod_{i=1}^{m} \otimes \Phi_i = \Phi_m \otimes \Phi_{m-1} \otimes \cdots \Phi_2 \otimes \Phi_1$$

Proof

Again the corollary is just an extension of the main theorem results.

An example has been included to clarify the arguments presented for the Markov decomposition. For simplicity, there are three binary state

units in the population. They are exponentially distributed with constant hazard models

$$\dot{p}_j(t) = \begin{bmatrix} -\lambda_j & \mu_j \\ \lambda_j & -\mu_j \end{bmatrix} p_j(t)$$

where a typical transition diagram would be represented in Figure 12.

The state assignments are given in Table III, and the subsequent coefficient matrix is represented by Figure 13. The decomposition structures of matrices B and C are clear from Figures 14 and 15.

The i,j entry of the state transition matrix for the entire process would be given by (Henry and Weidner, 1980)

$$\Psi_n(i,j) = \prod_{k=1}^{n} \Phi_k(g(i,k),\ g(j,k)) \qquad (4.34)$$

where n = 3 and

$$\Phi_k(t,t_o) = 1/(\lambda_k + \mu_k) \begin{bmatrix} \mu_k + \lambda_k \Delta_k & \mu_k - \mu_k \Delta_k \\ \lambda_k - \lambda_k \Delta_k & \lambda_k + \lambda_k \Delta_k \end{bmatrix}$$

where

$$\Delta_k = \exp[-(\lambda_k + \mu_k)(t - t_o)]$$

and the indices are determined from

$$g(i,k) = [\,i/2^{k-1}\,] \bmod 2 + 1$$

where $\lfloor a \rfloor$ is the greatest integer strictly less than a, and b mod 2 is the remainder of b, congruence modulo 2.

It should be noted that the full Markov matrices are never needed in full force. The storage savings on large problems can be immense. Table

TABLE III

STATE ASSIGNMENTS FOR A TRI-UNIT SYSTEM

State	Unit			Binary Weight Notation		
p_i		2		w_1	w_2	w_3
1	up	up	up	1	1	1
2	down	up	up	0	1	1
3	up	down	up	1	0	1
4	down	down	up	0	0	1
5	up	up	down	1	1	0
6	down	up	down	0	1	0
7	up	down	down	1	0	0
8	down	down	down	0	0	0

TABLE IV

STORAGE SAVINGS OF TRI-UNIT DECOMPOSITION

Number of Units	System Order	Memory Utilization	
		Kronecker	Eigenvalue
1	2	6	10
2	4	12	28
3	8	22	88
4	16	40	304
5	32	74	1,120
6	64	140	4,288
7	128	270	16,768
8	256	528	66,304
9	512	1,042	263,680
10	1,024	2,068	1,051,648

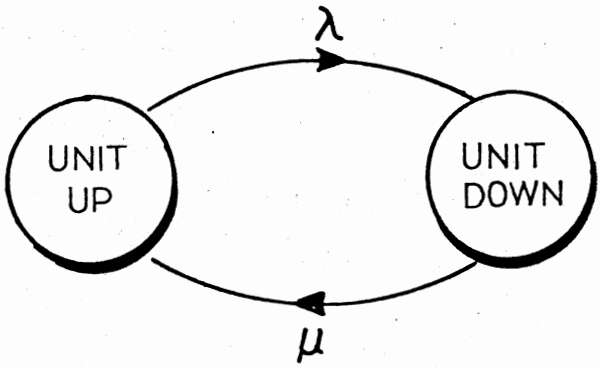

Figure 12. Transition Diagram for a Binary State Unit

$$
\begin{bmatrix}
-(\lambda_1+\lambda_2+\lambda_3) & \mu_1 & \mu_2 & 0 & \mu_3 & 0 & 0 & 0 \\
\lambda_1 & -(\mu_1+\lambda_2+\lambda_3) & 0 & \mu_2 & 0 & \mu_3 & 0 & 0 \\
\lambda_2 & 0 & -(\lambda_1+\mu_2+\lambda_3) & \mu_1 & 0 & 0 & \mu_3 & 0 \\
0 & \lambda_2 & \lambda_1 & -(\mu_1+\mu_2+\lambda_3) & 0 & 0 & 0 & \mu_3 \\
\lambda_3 & 0 & 0 & 0 & -(\lambda_1+\lambda_2+\mu_3) & \mu_1 & \mu_2 & 0 \\
0 & \lambda_3 & 0 & 0 & \lambda_1 & -(\mu_1+\lambda_2+\mu_3) & 0 & \mu_2 \\
0 & 0 & \lambda_3 & 0 & \lambda_2 & 0 & -(\lambda_1+\mu_2+\mu_3) & \mu_1 \\
0 & 0 & 0 & \lambda_3 & 0 & \lambda_2 & \lambda_1 & -(\mu_1+\mu_2+\mu_3)
\end{bmatrix}
$$

Figure 13. Coefficient Matrix for a Tri-Unit System

$-\lambda_3$	0	0	0	μ_3	0	0	0
0	$-\lambda_3$	0	0	0	μ_3	0	0
0	0	$-\lambda_3$	0	0	0	μ_3	0
0	0	0	$-\lambda_3$	0	0	0	μ_3
λ_3	0	0	0	$-\mu_3$	0	0	0
0	λ_3	0	0	0	$-\mu_3$	0	0
0	0	λ_3	0	0	0	$-\mu_3$	0
0	0	0	λ_3	0	0	0	$-\mu_3$

Figure 14. Decomposition Structure B for a Tri-Unit System

$-(\lambda_1+\lambda_2)$	μ_1	μ_2	0	0	0	0	0
λ_1	$-(\mu_1+\lambda_2)$	0	μ_2	0	0	0	0
λ_2	0	$-(\lambda_1+\mu_2)$	μ_1	0	0	0	0
0	λ_2	λ_1	$-(\mu_1+\mu_2)$	0	0	0	0
0	0	0	0	$-(\lambda_1+\lambda_2)$	μ_1	μ_2	0
0	0	0	0	λ_1	$-(\mu_1+\lambda_2)$	0	μ_2
0	0	0	0	λ_2	0	$-(\lambda_1+\mu_2)$	μ_1
0	0	0	0	0	λ_2	λ_1	$-(\mu_1+\mu_2)$

Figure 15. Decomposition Structure C for a Tri-Unit System

IV contains some of the storage savings of the proposed decomposition.

The entire matrix can be found from $n2^n$ multiplications. The full matrix may be formed by less sophisticated techniques, but it is noted that often the full matrix is often not necessary for many applications. This technique allows a method for escaping forming the full state transition matrix.

Summary

This chapter has dealt chiefly with the Kronecker forms and the applications of a Kronecker algebra. An efficient form for the solution of the general linear differential matrix equation has been established. As a result of this solution, the time-varying-coefficient multiple-unit Markov process can be decomposed, and an extremely efficient algorithm proposed for solving the Markov differential equation.

Clearly, the Kronecker algebra is of great value. In addition, the Kronecker algebra may be used in evaluating variance equations for compartment models. That will be covered in the next chapter.

CHAPTER V

STRUCTURAL PROPERTIES OF THE VARIANCE

Introduction

Previous chapters have been pieces from an overall puzzle concerning the stochastic properties of compartment models. Chapter II has established a wealth of matrix structural properties for compartment models, while Chapter III has examined the relationship between the deterministic theory and the stochastic moments. Finally, Chapter IV proposed a new tool for examining the moment equations, namely a Kronecker algebra. This chapter will combine the three chapters into a set of matrix structural properties for the variance of the states of compartment models.

It will be shown that the deterministic equation for the variance obeys the structural features of the compartment model. Hence, the features of the spectral range--irreducibility, monotonic behavior, etc.--can be transferred. The concepts of open and closed system structures do not preserve any meaning and are ignored.

First, the basic structure of the variance differential equation is examined. This is followed by a set of theorems for the structural properties of the variance coefficient matrix. Finally, partitions of the covariance equations are examined for an argument which establishes the behavioral properties of the variance for compartmental systems.

The Variance Differential Equation

Consider a system composed of n compartments where the flows out of the compartments are related by the stochastic relation.

$$\dot{x} = A\,x + B\,w \tag{5.1}$$

The terms are defined by

$$E\{w(t)\} = 0$$

$$E\{w(t)\ w'(\tau)\} = Q\ \delta_d(t-\tau)$$

$$E\{x(t)\} = \eta(t)$$

$$Cov\{x(t)\ x'(t)\} = P(t)$$

The mean η is related dynamically by

$$\dot{\eta} = A\,\eta \tag{5.2}$$

The variance P is related by

$$\dot{P} = A\,P + P\,A' + B\,Q\,B' \tag{5.3}$$

An example for a second order system is illustrated by Figure 16.

Matrix A is nxn with strictly real values. The diagonal terms are non-positive, and the off-diagonal terms are non-negative and related by

$$|a_{ii}| \geq \sum_{j}^{n\,\prime} a_{ji} \tag{5.4}$$

Therefore, the column sums are non-positive. This is the negative Minkowski matrix. The eigenvalues of A also have non-positive real parts. Though it should be noted that for the rest of this study, the eigenvalues are assumed to be real.

Assume that the system modeled by Equation (5.1) is strongly connected, hence irreducible (Thron, 1972). Since the strongly connected

$$\begin{bmatrix} \dot{p}_{11} & \dot{p}_{12} \\ \dot{p}_{21} & \dot{p}_{22} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} + \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}$$

$$+ \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{21} \\ b_{12} & b_{22} \end{bmatrix}$$

Figure 16. Covariance Equation for a Second-Order System

portions of a system can be separated from the rest of a not-strongly connected system, this assumption is acceptable. Second, assume the system is completely open. This implies that the inequalities of Equation (5.4) are all strictly greater than in sense. Therefore, the column sums are negative.

By convention a matrix A is defined as negative(positive), if every element a is negative(positive). Similarly, a matrix A is non-positive (non-negative), if every element is less than or equal to zero (greater than or equal to zero). Alternatively, a symmetric matrix is positive definite(negative definite), if the quadratic form x'Ax is positive(negative) for all values of x except x = 0. A symmetric matrix A is positive semi-definite(negative semi-definite), if the quadratic form is positive (negative) or zero for all $x \neq 0$. It is important to note that P is positive semi-definite, but not necessarily non-negative (Searle, 1971).

Equation (5.3) can be rewritten as a set of scalar equations

$$\dot{p}_{kj} = \sum_{i=1}^{n} (a_{ji}\, p_{ki} + a_{ki}\, p_{ij}) + \sum_{\ell}^{n} \sum_{i}^{n} b_{k\ell}\, b_{ji}\, q_{\ell i} \qquad (5.5)$$

where the diagonal terms are

$$\dot{p}_{kk} = \sum_{i=1}^{n} a_{ki}\, (p_{ki} + p_{ik}) + \sum_{\ell}^{n} \sum_{i}^{n} b_{k\ell}\, b_{ki}\, q_{\ell i} \qquad (5.6)$$

or as a vector matrix equation

$$\dot{p}_r = F\, p_r + G\, q_r \qquad (5.7)$$

where $F = (A \oplus A)$, $G = (B \otimes B)$, p_r is the column vector composed of the rows of P, q_r is the similar row stack of Q, $\oplus$ is the Kronecker sum operator, and $\otimes$ is the Kronecker product operator. Matrices F and G are of dimension $n^2 x n^2$, and p_r and q_r are n^2 vectors. Second order examples of Equations (5.5) and (5.7) are shown in Figures 17 and 18.

$$\dot{p}_{11} = a_{11}p_{11} + a_{11}p_{11} + a_{12}p_{12} + a_{12}p_{21} + (BQB')_{11}$$

$$\dot{p}_{12} = a_{21}p_{11} + a_{11}p_{12} + a_{22}p_{12} + a_{12}p_{22} + (BQB')_{12}$$

$$\dot{p}_{21} = a_{11}p_{21} + a_{21}p_{11} + a_{12}p_{22} + a_{22}p_{21} + (BQB')_{21}$$

$$\dot{p}_{22} = a_{21}p_{21} + a_{21}p_{12} + a_{22}p_{22} + a_{22}p_{22} + (BQB')_{22}$$

Figure 17. Second-Order Example of Equation (5.5)

$$\begin{bmatrix} \dot{p}_{11} \\ \dot{p}_{12} \\ \dot{p}_{21} \\ \dot{p}_{22} \end{bmatrix} = \begin{bmatrix} a_{11}+a_{11} & a_{12} & a_{12} & 0 \\ a_{21} & a_{22}+a_{11} & 0 & a_{12} \\ a_{21} & 0 & a_{11}+a_{22} & a_{12} \\ 0 & a_{21} & a_{21} & a_{22}+a_{22} \end{bmatrix} \begin{bmatrix} p_{11} \\ p_{12} \\ p_{21} \\ p_{22} \end{bmatrix}$$

$$+ \begin{bmatrix} b_{11}b_{11} & b_{11}b_{12} & b_{12}b_{11} & b_{12}b_{12} \\ b_{11}b_{21} & b_{11}b_{12} & b_{12}b_{21} & b_{12}b_{22} \\ b_{21}b_{11} & b_{21}b_{12} & b_{22}b_{11} & b_{22}b_{12} \\ b_{21}b_{21} & b_{21}b_{22} & b_{22}b_{21} & b_{22}b_{22} \end{bmatrix} \begin{bmatrix} q_{11} \\ 0 \\ 0 \\ q_{22} \end{bmatrix}$$

Figure 18. Second-Order Example of Equation (5.7)

The first-order vector Equation (5.6) is a straightforward equation which could be though to model a hypothetical system. The "pseudo-system" has structural properties implied by F and G in much the same manner as A implies the properties of the physical compartmental system.

Matrix F Properties

Theorem 5.1

Given vector Equation (5.6), where matrix A is diagonally dominant as in Equation (5.4) and is a negative Minkowski matrix, Matrix F is a diagonally dominant negative Minkowski matrix.

Proof

The terms of F can be expressed by

$$f_{rc} = a_{jk}\,\delta_{hi} + a_{hi}\,\delta_{jk} \tag{5.8}$$

where $j = (r - 1) \bmod n + 1$, $k = (c - 1) \bmod n + 1$, $h = \lfloor r/n \rfloor + 1$, $i = \lfloor c/n \rfloor + 1$, and $\lfloor a \rfloor$ is the greatest integer strictly less than a, n is the dimension of A, δ_{jk} is the Kronecker delta defined by

$$\delta_{jk} = \begin{cases} 1, & j = k \\ 0, & j \neq k \end{cases}$$

and a mod b is the remainder, g, when $a/b = f + g/b$ (all terms a, b, f, g are integers).

Therefore, the diagonal terms of F are

$$f_{rr} = a_{jj} + a_{hh} \tag{5.9}$$

since $j = k$, $h = i$.

The sum of any column l from matrix F is

$$f_{\ell\ell} + \sum_{r=1}^{n}{}' f_{r\ell} = a_{jj} + \sum_{i=1}^{n}{}' a_{ij} + a_{hh} + \sum_{i=1}^{n}{}' a_{ih} \tag{5.10}$$

Therefore, matrix F is diagonally dominant with its column sums composed of the n^2 sums of the columns of matrix A. Hence F is a negative Minkowski matrix and the proof is complete.

It is well known (Bellman, 1960) that the eigenvalues of $F(\lambda_i)$ are the sums of the eigenvalues of $A(\mu_j)$. That is

$$\lambda_i = \mu_j + \mu_k \tag{5.11}$$

for all j,k in A.

Lemma 5.2

The spectral range is now defined by

$$\lambda_{min} = 2\mu_{min} \leq \lambda_i \leq \lambda_{max} = 2\mu_{max} \tag{5.12}$$

where λ_i is a proper value of F.

Corollary 5.2a

The spectral radius $\rho(A)$ is defined by $\rho(A) = \max|\lambda_i|$; therefore,

$$\rho(F) = 2\rho(A) \tag{5.13}$$

Corollary 5.2b

Since $\rho(A) \leq \|A\|$ for any matrix norm corresponding to a vector norm, then

$$2\rho(A) \leq \|F\| \tag{5.14}$$

Proof (Householder, 1964)

Let $A x = \lambda x$, $x \neq 0$, λ a proper value of A, then

$$\|\lambda x\| = |\lambda| \, \|x\| \leq \|x\| \, \text{lub}(A) \leq \|x\| \, \|A\| \tag{5.15}$$

for any consistent pairs of vector and matrix norms.

Corollary 5.2c

Since $F p = \lambda p$, $p \neq 0$, then

$$|\lambda| = \|F p\| / \|p\| \tag{5.16}$$

for any λ a proper value of F. More directly,

$$|\mu_i| + |\mu_j| = \|F p\| / \|p\|$$

for any vector norm.

From Barnett (1973), for a square matrix A, I the identity matrix

$$A \otimes I = P(I \otimes A)P$$

where P is defined by

$$V_c = P V_r \quad \text{or} \quad V_r = P V_c$$

where V_c is the column stack of matrix V and V_r is the row stack of matrix V. Also

$$P = P^{-1} = P'$$

Definition: A square matrix is reducible if a permutation of the rows and columns of A permutes A into the form

$$G = \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}$$

where A_{11} and A_{22} are square matrices, and 0 is the null matrix.

Theorem 5.2

If A is an irreducible square matrix, then F is an irreducible square matrix.

Proof

The proof is one of contradiction. Assume F is reducible, then

$$MFM' = H$$

where H is a lower triangular matrix, M is an arbitrary permutation matrix.

$$\begin{aligned} MFM' &= M(A \otimes I + I \otimes A)M' \\ &= M(A \otimes I)M' + M(I \otimes A)M' \\ &= M(A \otimes I)M' + MP(A \otimes I)PM' \end{aligned}$$

unless: I. $MP = M$, or II. $MP = N$, where $N(A \otimes I)N'$ is lower triangular, then F is not reducible.

I. Since $P \neq I$, unless A is reducible, then condition I. fails.

II. Similarly, $M \neq I$, $N \neq I$, and $M \neq P$. Substituting $M = NP$

$$\begin{aligned} MFM' &= NP(A \otimes I)PN' + N(A \otimes I)N' \\ &= N(I \otimes A)N' + N(A \otimes I)N' \end{aligned}$$

This requires

$$MFM' = NFN'$$

or $M = N$. But $P \neq I$. So $M \neq N$, and condition II. fails. Thus F is not reducible, unless A is reducible.

Corollary 5.2.1

If A is a reducible matrix, F is a reducible matrix.

Proof

Since A is reducible, there exists a permutation matrix P such that

$$PAP' = G$$

The Kronecker sum of G is

$$H = G \oplus G'$$

By inspection, H is of the form

$$H = \begin{bmatrix} H_{11} & 0 \\ H_{21} & H_{22} \end{bmatrix}$$

where H_{11} and H_{22} are square matrices, since G is of reduced form. Then,

$$\begin{aligned} H &= G \otimes I + I \otimes G \\ &= PAP' \otimes I + I \otimes PAP' \\ &= PAP' \otimes PP' + PP' \otimes PAP' \\ &= (P \otimes P)(A \otimes I) + (I \otimes A))(P \otimes P)' \\ &= MFM' \end{aligned}$$

Since $(P \otimes P) = M$ is a permutation matrix, F is reducible.

Lemma 5.3

If A is irreducible and singular, then F is irreducible and singular.

Proof

Since A is irreducible and singular, it has all zero column sums

(Taussky, 1949). Therefore, F is irreducible and singular from Equation (5.10) and Theorem 5.2.

Corollary 5.3a

If A is irreducible and nonsingular, F is irreducible and nonsingular.

Proof

The proof follows directly from Equation (5.10) and Taussky's (1949) Theorem II.

Lemma 5.4

Matrix F is nonsingular, if and only if the compartment system is completely open, hence A is nonsingular.

Proof

The proof follows directly from Equation (5.11) and Thron's (1972) Theorem 2.

Definition: A matrix F composed of a function of one or more other matrices A, B, . . . is symmetric if

$$f_{k\ell} = f(a_{ij}, b_{nm}, \ldots)$$

has a corresponding cross diagonal element

$$f_{\ell k} = f(a_{ji}, b_{mn}, \ldots)$$

Lemma 5.5

The terms of the Kronecker Sum ($A \otimes I$ and $I \otimes A$) are index symmetric and therefore F is index symmetric.

As a direct result of Lemma 5.5, the sum of column ℓ in F corresponds to the sum of row ℓ in F via the indices of the elements a_{ji} in the column sum are reversed to a_{ij} in the row sum.

Covariance Partition Equations and Behavioral Properties

The variance terms along the diagonal are usually the main interest of a study of the second moment equation. With that purpose in mind, it is possible to separate the matrix variance equation into a set of scalar equations for the off-diagonal covariance terms and a set of scalar equations for the diagonal variance terms.

The off-diagonal terms of Equation (5.3) may be written as

$$\dot{p}_{\ell m} = (a_{\ell\ell} + a_{mm})\, p_{\ell m} + \sum_{i \neq \ell, m}^{n} (a_{mi}\, p_{\ell i} + a_{\ell i}\, p_{im}) + a_{m\ell}\, p_{\ell\ell} + a_{\ell m}\, p_{mm} \tag{5.18}$$

with no forcing term, assuming the driving term BQB' is diagonal. Consider a vector p_3 composed of the off-diagonal terms of matrix P in row-stacked order, and a vector p_1 composed of the diagonal terms of matrix P in row stack order. Equation (5.18) could be written as

$$\dot{p}_3 = F^{(3)} p_3 + F^{(4)} p_1 \tag{5.19}$$

A third-order example of Equation (5.19) is shown in Figure 19. The diagonal terms of $F^{(3)}$ are by inspection

$$\begin{bmatrix} \dot{p}_{12} \\ \dot{p}_{13} \\ \dot{p}_{21} \\ \dot{p}_{23} \\ \dot{p}_{31} \\ \dot{p}_{32} \end{bmatrix} = \begin{bmatrix} a_{11}+a_{22} & a_{23} & 0 & 0 & 0 & a_{13} \\ a_{32} & a_{11}+a_{33} & 0 & a_{12} & 0 & 0 \\ 0 & 0 & a_{11}+a_{22} & a_{13} & a_{23} & 0 \\ 0 & a_{21} & a_{31} & a_{22}+a_{33} & 0 & 0 \\ 0 & 0 & a_{32} & 0 & a_{33}+a_{11} & a_{12} \\ a_{31} & 0 & 0 & 0 & a_{21} & a_{33}+a_{22} \end{bmatrix} \begin{bmatrix} p_{12} \\ p_{13} \\ p_{21} \\ p_{23} \\ p_{31} \\ p_{32} \end{bmatrix}$$

$$+ \begin{bmatrix} a_{21} & a_{12} & 0 \\ a_{31} & 0 & a_{13} \\ a_{21} & a_{12} & 0 \\ 0 & a_{32} & a_{23} \\ a_{31} & 0 & a_{13} \\ 0 & a_{32} & a_{23} \end{bmatrix} \begin{bmatrix} p_{11} \\ p_{22} \\ p_{33} \end{bmatrix}$$

Figure 19. Third-Order Example of Equation (5.19)

$$f_{ii}^{(3)} = a_{\ell\ell} + a_{mm} \tag{5.20}$$

Matrix $F^{(3)}$ like F is index symmetric. Thus the sum of column k is

$$f_{kk}^{(3)} + \sum_{j=1}^{n}{}' f_{jk}^{(3)} = a_{\ell\ell} + \sum_{i\neq\ell,m}^{n} a_{im} + a_{mm} + \sum_{i\neq\ell,m}^{n} a_{i} \tag{5.21}$$

Hence, $F^{(3)}$ is diagonally dominant. Furthermore, if $a_{\ell m} \neq 0$ or $a_{m\ell} \neq 0$,

$$|f_{kk}^{(3)}| > \sum_{j=1}^{n^2-n}{}' f_{jk}^{(3)} \tag{5.22}$$

regardless of whether A has zero column sums or not.

Since $p_{ki} = p_{ik}$, the set of equations in Equation (5.19) is composed of two equivalent sets of equations. Let p_2 be the $(n^2 - n)/2$ vector composed of the terms of p_3 with the equivalent pairs replaced as single terms. Now, Equation (5.19) can be replaced by

$$\dot{p}_2 = F^{(2)} p_2 + F^{(1)} p_1 \tag{5.23}$$

where $F^{(2)}$ and $F^{(1)}$ are $(n^2 - n)/2 \times (n^2 - n)/2$. Matrix $F^{(2)}$ is again index symmetric. The column sums of $F^{(2)}$ corresponding to element p_{ik} are exactly the column sums of $F^{(3)}$ corresponding to term p_{ik} or p_{ki}. Hence $F^{(2)}$ is also diagonally dominant, and has the structural form of A or F. A third-order example of Equation (5.23) is shown in Figure 20.

The diagonal terms of P, the variances, can be expressed by

$$\dot{p}_{kk} = 2a_{kk} p_{kk} + \sum_{i=1}^{n}{}' a_{ki} (p_{ki} + p_{ik}) + \sum_{i=1}^{n} b_{ki} b_{ki} q_{ii} \tag{5.24}$$

or

$$\dot{p}_1 = F^{(5)} p_1 + F^{(6)} p_2 + Q_1 \tag{5.25}$$

$$\begin{bmatrix} p_{12} \\ p_{13} \\ p_{23} \end{bmatrix} = \begin{bmatrix} a_{11}+a_{22} & a_{23} & a_{13} \\ a_{32} & a_{11}+a_{33} & a_{12} \\ a_{31} & a_{21} & a_{22}+a_{33} \end{bmatrix} \begin{bmatrix} p_{12} \\ p_{13} \\ p_{23} \end{bmatrix}$$

$$+ \begin{bmatrix} a_{21} & a_{12} & 0 \\ a_{31} & 0 & a_{13} \\ 0 & a_{32} & a_{23} \end{bmatrix} \begin{bmatrix} p_{11} \\ p_{22} \\ p_{33} \end{bmatrix}$$

Figure 20. Third-Order Example of Equation (5.23)

Matrix $F^{(5)}$ is diagonal with

$$f^{(5)}_{kk} = 2a_{kk} \tag{5.26}$$

A third-order example of Equation (5.24) is illustrated in Figure 21.

By definition, the terms of p_1 are non-negative. By inspection, matrices $F^{(1)}$ and $F^{(6)}$ are non-negative. Unlike the general compartmental state, the covariance of the states is not defined as being non-negative. However, if considering the donor-controlled (washout) compartment model, all non-negative covariance terms may be practical.

Mathematically, the covariance (p_1 and p_2) can be considered as non-negative if the initial conditions on P are non-negative and the driving terms are all non-negative (Thron, 1972). Specifically, this could be expressed as an extension of Thron's Theorem 7.

Lemma 5.6

For any matrix F of the form of Equation (5.7), the solution of $dp/dt = F\ p$ is non-negative if the initial values are non-negative.

Proof

Since the driving terms in Equations (5.23) and (5.25) are non-negative, then p_1 and p_2 are non-negative if $p_2(0)$ is non-negative.

The driving term in Equation (5.3) may be considered to be diagonal due to the practical limitations of the experiments. The off-diagonal covariance terms may be considered as non-negative if the experiment is begun from the null steady state where the states are practically independent. Lemma 5.6 has interesting applications on the behavior of the variance terms.

$$\begin{bmatrix} p_{11} \\ p_{22} \\ p_{33} \end{bmatrix} = \begin{bmatrix} 2a_{11} & 0 & 0 \\ 0 & 2a_{22} & 0 \\ 0 & 0 & 2a_{33} \end{bmatrix} \begin{bmatrix} p_{11} \\ p_{22} \\ p_{33} \end{bmatrix} + \begin{bmatrix} 2a_{12} & 2a_{13} & 0 \\ 2a_{21} & 0 & 2a_{23} \\ 0 & 2a_{31} & 2a_{32} \end{bmatrix} \begin{bmatrix} p_{12} \\ p_{13} \\ p_{23} \end{bmatrix}$$

$$+ \begin{bmatrix} b_{11}^2 & 0 & 0 \\ 0 & b_{22}^2 & 0 \\ 0 & 0 & b_{33}^2 \end{bmatrix} \begin{bmatrix} q_{11} \\ q_{22} \\ q_{33} \end{bmatrix}$$

Figure 21. Third-Order Example of Equation (5.24)

Corollary 5.6a

In any multicompartment system with a steady state solution, if the initial value is a steady state, the covariance of each compartment state changes from its initial to its final state without overshoot, if in every compartment the initial covariance is not less than the final steady state quantity or if in every compartment it is not greater.

Corollary 5.6b

When a system initially in steady state with input covariance Q responds to a non-negative step function covariance ΔQ, which leads to a new steady state, the transitions of the covariance to the final steady state are monotonic in all compartments.

Summary

This chapter has brought together the concepts of the deterministic moments with the structural properties of the compartment model. An extension of the structural properties has resulted in properties for the variance of the compartmental states.

The general equations for the variance have been examined at great detail. Resulting from a Kronecker algebra, the properties of the matrix A could be transferred to matrix F. Hence, the properties of the mean have influence on the properties of the variance. Finally, a few of the behavioral properties were examined for specifically the variance and the off-diagonal covariance terms.

The next chapter will utilize these properties in proposing a new identification experiment, and possibly a meaningful extension of the sampling theory.

CHAPTER VI

VARIANCE BOUNDS AND PROPOSED FUTURE RESEARCH

Introduction

The previous chapters have proposed various structural properties for the statistical variance of the compartment model states. The variance has been shown to behave similarly to the mean in monotonic responsiveness, spectral range, etc. This chapter will be devoted to the application of the structural properties in a particular area, namely variance bounds for identification.

A new basic experiment will be proposed with the concept of new dynamic bounds on the mean estimate resulting from the bounds on the variance. As is typical, the assumption of Gaussianly-distributed random variables will be carried throughout the discussion. The result of the examination will be guidelines for future endeavors in compartment model identifications and sampling.

The Revised Basic Experiment

Basically, the identification experiment is the same as in Gowdy (1978). However, modification must be performed to hopefully gain sufficient information to estimate the moment parameters. It is suggested that the experiment be of the form of Figure 22. Initially, the system states should be brought to a steady state region. This region should be of sufficient length to establish reasonable estimates on the mean and

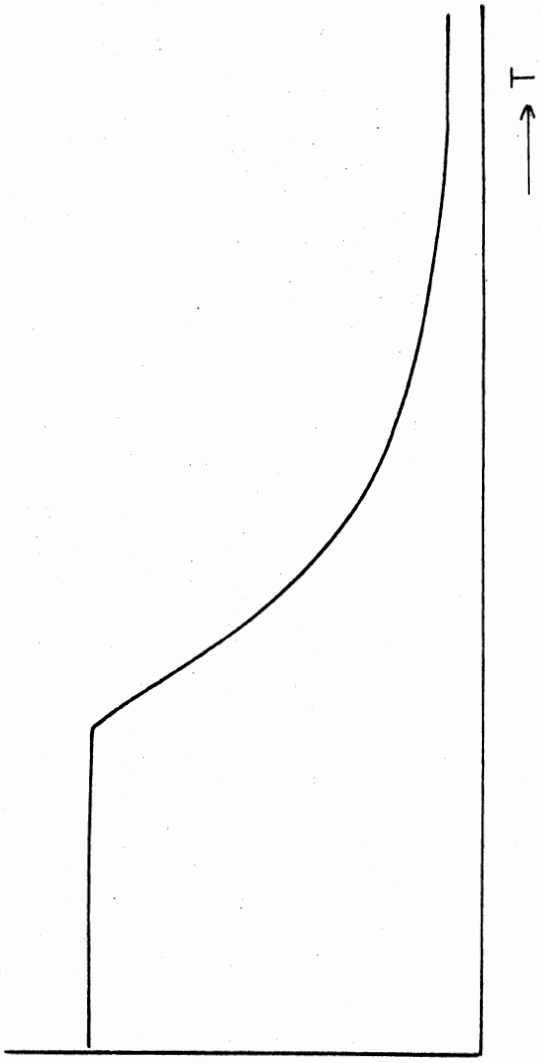

Figure 22. Revised Basic Experiment

variance of the states. Standard time series techniques can be utilized for this purpose. These estimates are the initial condition estimates for the transition region.

Similarly, a final steady state region is sampled to establish estimates of the final moments for the transition region. Superposition can then be applied to obtain the relationships of Chapter V.

Note the experiment is still a negative step response to avoid the precursor problem. Also, Lemma 5.6 and subsequent corollaries will apply for the transition region with a step function input.

From pilot experiments or similar a priori information, a sampling theory can be developed utilizing bounds on the sample mean. An enlarged view of the transition region with sampling could be represented by Figure 23.

Traditional techniques would attempt to remove the trend and analyze the data with assumptions of ergodicity and hence equal variance samples. However, the trend in the data in the desired information and cannot be easily estimated or neglected.

Referring to Figure 23, if the confidence intervals on the samples overlap, the samples are not statistically independent. Therefore, the sample period would be too small. On the other hand, if using a fixed interval regression, the samples could be taken too far apart, and again statistical significance will be lost.

For these reasons, the variances will be approximated by bounds at least in trace from the information in Chapter V and the eigenvalue spectrum suggested by Gershgorin (1931). Confidence intervals may result from such an analysis.

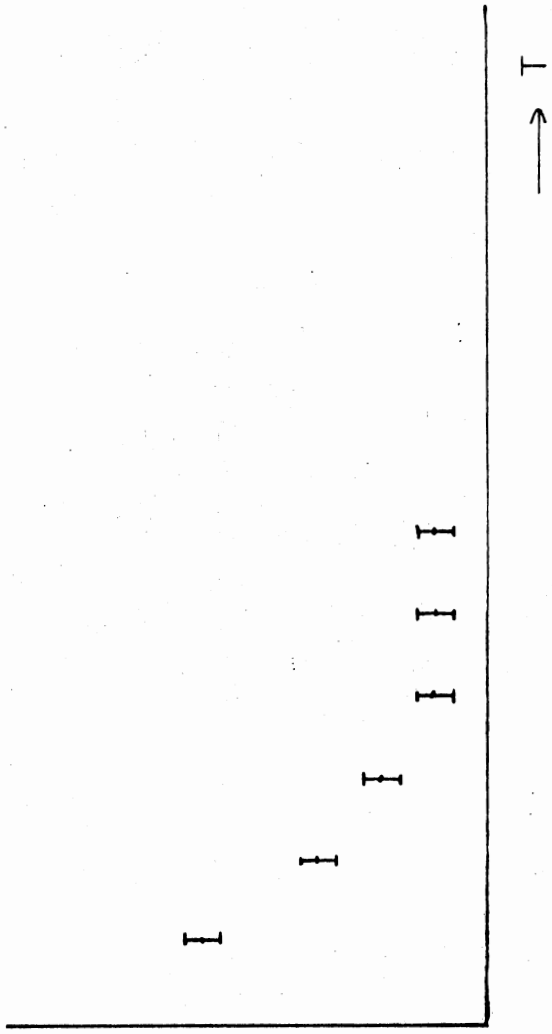

Figure 23. Transition Region of the Basic Experiment

Variance Bounds

It is evident that the variance is known only through estimates. Therefore, the variance is open to the same confidence interval argument as the mean. However, some use is found in utilizing the practical aspects of the experiment, and therefore analyzing the confidence intervals using the variance estimate as the actual variance. Future research may prove useful in determining the variance of the variance estimates.

The state transition matrix of Equation (5.7) is given by

$$\psi(t) = L\, \Xi(t)\, L^{-1} \tag{6.1}$$

where

$$\Xi(t) = \exp(\Lambda t) \tag{6.2}$$

L is an invertible $n^2 x n^2$ matrix of the eigenvectors of F, Λ is a diagonal $n^2 x n^2$ matrix of the eigenvalues such that

$$L^{-1}\, F\, L = \Lambda \tag{6.3}$$

The eigenvalues of F are defined to be the roots of

$$|I\, \lambda - F| = 0 \tag{6.4}$$

As was noted in Chapter V, the eigenvalues are related to the eigenvalues of A and hence to the compartment turnovers. This results from the Gershgorin spectrum.

It is known that for a nxn complex matrix A, the spectrum is contained within the union of the n circles in the complex plane described by

$$|\lambda - a_{jj}| < \sum_{i=1}^{n}{}' |a_{ij}| \qquad j = 1, \ldots, n \tag{6.5}$$

This is the circles theorem accredited to Gershgorin (1931). The

circles theorem is interpreted to mean the eigenvalues of a matrix A lie within the circles of the spectrum.

As a consequence of the Gershgorin theorem and the eigenvalue relation, Theorem 6.1 can be stated.

Theorem 6.1

For a complex matrix F, formed by $F = A \bigoplus A'$, the spectrum is contained in the union of the n^2 circles in the complex plane

$$|\lambda - a_{ii} - a_{hh}| < \sum_{j=1}^{n}{}' |a_{ji}| + \sum_{k=1}^{n}{}' |a_{kh}| \tag{6.6}$$

Proof

Theorem 6.1 is actually a restatement of Gershgorin's theorem with substitution for the terms in F formed by the Kronecker sum.

The circles aspect grows from a substitution of

$$r_{ih} = \sum_{i=1}^{n}{}' |a_{ji}| + \sum_{k=1}^{n}{}' |a_{kh}| \tag{6.7}$$

Then Equation (6.6) is a closed circular region in the λ-plane of radius r_{ih} and center at $a_{ii} + a_{hh}$.

The utility of the Gershgorin spectrum is realized by a theorem by Frobenius (1912).

Theorem 6.2

An irreducible non-negative matrix A always has a positive characteristic number r, which is a simple root of the characteristic equation. The moduli of all the other characteristic numbers are at least r. A

characteristic vector z with positive coordinates corresponds to the "dominant" characteristic number r.

Gowdy (1978) has illustrated how to apply this theorem to matrices of compartmental systems. In short, if a diagonal matrix zI, where $z = \max_i (a_{ii})$, is added to A, then Frobenius' theorem can be applied. Therefore, there is at least one real, simple eigenvalue which is dominant. Correspondingly, due to Relation (5.11), there is at least one real, simple least-negative eigenvalue of F which is dominant. It is

$$\lambda_{max} = 2\mu_{max} \tag{6.8}$$

If the system is completely open, none of the Gershgorin circles for matrix A touch the imaginary λ-axis. That is,

$$|a_{jj}| > \sum_{i=1}^{n}{}' |a_{ij}| = r_j \qquad j = 1, \ldots, n \tag{6.9}$$

Similarly, the Gershgorin circles for matrix F do not touch the imaginary axis if the system is completely open.

$$|a_{jj} + a_{hh}| > \sum_{i=1}^{n}{}' |a_{ij}| + \sum_{k=1}^{n}{}' |a_{kh}| = r_{jh} \tag{6.10}$$

The circle spectrum is represented in much the same manner as Figure 3.

Since all of the eigenvalues of F are negative real numbers, due to the eigenvalues of A being negative real numbers, they can be ordered and bounded as

$$-2\alpha < \lambda_1 < \lambda_2 < \ldots < \lambda_{n^2} < -2\delta \tag{6.11}$$

where the eigenvalue bounds are defined in Gowdy (1978) as

$$\delta = \max_j (r_j + |a_{jj}|) \tag{6.12}$$

and

$$\delta = \min_{j} \left(|a_{jj}| - r_j \neq 0 \right) \tag{6.13}$$

These bounds can be used to establish the range of behavior of the covariance trajectories.

Theorem 6.3

The covariance can be bounded by

$$\exp(-2\alpha(t))\ \mathrm{tr}(P(0)) \leq \mathrm{tr}(P(t)) \leq \exp(-2\ (t))\ (\mathrm{tr}(P(0)) + |p_2(0)|)$$

if $p_2(0) > 0$.

Proof

Consider Equation (5.24) with the effect of the input removed.

$$\dot{p}_{kk} = 2\ a_{kk}\ p_{kk} + \sum_{i=1}^{n}{}' a_{ki}\ (p_{ki} + p_{ik}) \tag{6.14}$$

where $a_{kk} < 0$; $a_{ki} > 0$, all $k \neq i$; $p_{ki} > 0$. For each k and t in [0,t),

$$\dot{p}_{kk}(t) - 2\ a_{kk}\ p_{kk}(t) \geq 0 \tag{6.15}$$

On the interval [0,t) for each k,

$$d/dt\ (\exp(-2\ a_{kk}(t))\ p_{kk}(t)) \geq 0 \tag{6.16}$$

and

$$p_{kk}(t) \geq p_{kk}(0)\ \exp(+2a_{kk}(t)) \tag{6.17}$$

The trace of P (sum of the diagonal variance terms)

$$\mathrm{tr}(P(t)) = \sum_{k=1}^{n} p_{kk}(t) \geq \exp(+2a_{11}(t))\ p_{11}(0)$$

$$+ \ldots + \exp(+2a_{nn}(t))\ p_{nn}(0) \tag{6.18}$$

By definition, $p_{kk}(t) > 0$; therefore,

$$\mathrm{tr}(p(t)) \geq \exp(-2\alpha(t))\ \mathrm{tr}(P(0)) \tag{6.19}$$

This gives a lower bound on the terms in the variance of the compartmental states.

An upper bound on the variance can be estimated, but is not as close as the lower bound. Consider the sum of Equations (5.18) for $m = 1, \ldots, n$; $l = m, \ldots, n$. This is $|p_1| + |p_2|$, where

$$|y| = \sum_{i=1}^{n} |y_i| \tag{6.20}$$

$$|\dot{p}_1| + |p_2| = \sum_{m=1}^{n} \sum_{\ell=1}^{n} ((a_{\ell\ell} + a_{mm})\ p_{\ell m}$$

$$+ \sum_{i \neq \ell, m}^{n} (a_{mi}\ p_{\ell i} + a_{\ell i}\ p_{im})$$

$$+ a_{m\ell}\ p_{\ell\ell} + a_{\ell m}\ p_{mm} \tag{6.21}$$

Equation (6.21) is equivalent to

$$|\dot{p}_1| + |p_2| = \sum_{m=1}^{n} \sum_{=m}^{n} ((a_{o\ell} + a_{om})p_{\ell m})$$

$$+ \sum_{i=1}^{n} (a_{ii} + a_{oi})\ p_{ii} \tag{6.22}$$

where

$$-a_{oi} = a_{ii} + \sum_{j=1}^{n} a_{ji}$$

Then

$$|\dot{p}_1| + |\dot{p}_2| < 2\, a_{oi}\, \max(|p_1| + |p_2|) \tag{6.23}$$

The maximum observed exogenous turnover can be estimated a priori and gives the bound δ. Therefore, the upper bound is

$$|\dot{p}_1| + |p_2| \leq 2\delta\,(|p_1| + |p_2|) \tag{6.24}$$

Proceeding as in the derivation of Equation (6.19), the upper bound on the covariance is

$$|p_1(t)| + |p_2(t)| \leq \exp(-2\delta\,(t))\ (|p_1(0)| + |p_2(0)|) \tag{6.25}$$

Note that $|p_1| = tr(P)$. Therefore,

$$tr(p(t)) + |p_2(t)| \leq \exp(-2\delta\,(t))\ (tr(P(0) + |p_2(0)|) \tag{6.26}$$

Since $|p_2| \geq 0$,

$$tr(P(t)) \leq \exp(-2\delta\,(t))\ (tr(P(0)) + |p_2(0)|) \tag{6.27}$$

The variance of the compartmental states are then bounded by

$$\exp(-2\alpha\,(t)\ tr(P(0)) \leq tr(P(t)) \leq \exp(-2\delta\,(t))\ (tr(P(0)) + |p_2(0)|) \tag{6.28}$$

This completes the proof.

If the eigenvalue bounds (α and δ) are known a priori, then these variance bounds are meaningful due to the presence of the estimates of P(0) at the beginning of the transient region in the basic experiment. Fortunately, many compartmental analysts have rough estimates for α and δ due to the relationship with the exogenous turnovers of the system. The minimal eigenvalue bound (α) can be estimated by twice the inverse of the maximum turnover rate. The maximum eigenvalue bound (δ) can simply be

the minimal Gershgorin range of the inverse of the minimal turnover rate.

Future Research Efforts

Given the bounds on the variance from the previous section, it is possible to construct confidence intervals on the transient region mean bounds proposed by Gowdy (1978). The combination of the two would yield a new basis for an identification sampling theory.

Gowdy's bounds would be useful due to the existence of the initial value mean estimates from the revised basic experiment. A confidence interval on the mean could be formed by approximating two standard deviations around the mean estimate. The two-standard deviation range would be found from the initial value estimates of the variance.

All statistical measures are inexact by definition. Therefore, great laxity is incorporated in the experimental procedure. Only experimental knowledge can solve the difficult problem of meaningfully utilizing the statistical information. This point, though felt to be unsatisfactory, is an inescapable part of systems identification. For this reason any theory, such as this work, can only yield guidelines for identification and not firm threshold decisions.

The bounds on the variance would yield dynamic confidence limits on the mean. Hopefully better confidence limits could be found than through the false assumption of ergodicity and its accompanying statistical procedures.

At this point there must be a discussion of a misconception in sampling theory, namely nonuniform sampling. Typically, statisticians convert raw experimental data into ergodic data through trend-removing transformations. If the data are simple functions, such as a single exponential

curve or a quadratic, this amounts to taking a log or a power of the data. It has been noted that the sampling theory changes with the transformation. Thus, nonuniform sample periods as a function of a log or a power are found meaningful. Sampling at log rates is known as μ-law sampling.

Data which are composed of complex combinations of exponentials cannot be organized into simple μ-law schemes. This is the common misconception. However, if there is a clearly dominant exponential factor, then it can be analyzed by μ-law methods. Typically, systems would require an eigenvalue which is dramatically separated from the rest of the eigenvalues. Some compartmental systems have this characteristic, but it is not known whether the trait is widespread enough to warrant such a general procedure.

The problem with applying μ-law to a set of exponentials is that much of the information identifying the nondominant eigenvalues is lost. Often, the other eigenvalues can be ignored as illustrated by the popularity of Box-Jenkins techniques, Karhunen-Loeve transformations, and principal value analysis. If the compartmental system has a separated dominant eigenvalue, then μ-law combined with Prony's method or some similar curve peeling method would find great use. The structural properties of this thesis would still be of great use in the alternate identification method.

Evidently, μ-law sampling would have adverse effects on the regression technique presented in this study. Therefore, this scheme has not been considered.

If uniform sampling is maintained, then minimal bounds on the sample period would be derived from the minimal period over which the samples significantly decrease. This would be the period over which the quantified data exhibit a change greater than the sum of two initial value

standard deviations and two standard deviations after one time period. The second standard deviation could be approximated by a meaningful adaptation of the variance bounds (Equation (6.28)).

The upper bound in Equation (6.28) is not satisfactorily tight. Further examination of the off-diagonal covariance terms could possibly improve the upper bound. Since the terms are statistical estimates, the bounds may be sufficient in their present state.

The maximum sample period will again be determined by a significance argument. The maximum sample period would be determined by the maximum time period over which at least n+1 significant samples of each compartment could be obtained and still be contained in the transient region. Clearly, it would be a function of the minimal eigenvalue. Fortunately, good approximations exist for the minimal eigenvalue.

These comprise the more meaningful suggestions for future research endeavors which utilize the structural properties established in this study. In addition, several interesting prospects have been encountered which could bear further investigation.

Of most immediacy would be the extension of the work to negative off-diagonal covariance initial values. Though heuristically it is meaningful to argue for positivity, a mathematical argument is felt to be necessary, if it exists.

Second, the extension of structural properties to partially closed systems is necessary. It should be noted that Smith's model in Chapter III is partially closed. Yet the structural properties of sampling bounds are observed. In turn the argument for variance structural properties for some partially closed systems is established.

Frequency spectrum techniques are of little use on this topic. The

effects of quantization and sampling are difficult to approach through the continuous Fourier transform. The entire basis of the frequency spectrum for aperiodic signals would have to be re-established.

Of more promise is the possible analysis through Walsh basis functions. Still the amplitudes are are continuous functions, not discrete. It is felt that discrete amplitude basis functions are one of the most important topics remaining for analysis in identification.

The difficulty with the previous efforts in discrete communication theory (signal processing) is the lack of causality. It is sufficient for communications identification to propose a model which will reconstruct the signal. Such models need not have any connection with the physical world. For this reason, the wealth of digital signal processing work must be examined for techniques which do have direct connection with the real-world system. It is those techniques which are of use in general identification.

Summary

This chapter has been one of application. The structural properties of compartment models have been utilized in examining the Gershgorin spectrum for the variance. Resulting from the eigenvalue spectrum has been an approach for dynamically bounding the trace of the variance of the compartmental states. A revised basic identification experiment has been proposed which would furnish statistical estimates for the bounds derived in Gowdy (1978) and for the new variance bounds.

Future research endeavors have been proposed for utilizing the variance bounds, hence the structural properties in a sampling theory for

compartment model identification. Additional research topics have been proposed which result from the examination presented here.

BIBLIOGRAPHY

Barnett, S. 1973. Matrix Differential Equations and Kronecker Products. SIAM J. Appl. Math. 24, 1.

Bellman, R. and Astrom, K. J. 1970. On Structural Identifiability. Math. Biosci 7, 329.

Bellman, R. 1959. Kronecker Products and the Second Method of Lyapunov. Math. Nachr. 20, 17.

Bellman, R. 1960. Introduction to Matrix Analysis. New York: McGraw-Hill.

Berkstresser, B. K., El-Asfouri, S., McConell, J. R., and McInnis, B. C. 1979. Identification Techniques for the Renal Function. Math. Biosci. 44, 157.

Brewer, J. W. 1978. Kronecker Products and Matrix Calculus in System Theory. IEEE Trans. on Circuits and Systems. CAS-25, 1.

Brockett, R. W. 1970. Finite Dimensional Linear Systems. New York: Wiley.

Brown, R. F. 1980. Compartmental System Analysis: State of the Art. I.E.E.E. Trans. Biomed. Engr. BME-27, 1.

Campello, L. and Cobelli, C. 1978. Parameter Estimation of Biological Stochastic Compartment Models-An Application. IEEE Trans. on Biomed. Engr. BME-25, 139.

Chang, F. and Fitzgerald, T. J. 1977. Discrete Flow Modeling: A General Discrete Time Compartmental Model. AICHE Journal. 23, 558.

Desplanques, J. 1886. Theoreme d'Algebra. J. Math. Spec. 1, 12.

Frobenius, G. 1912. Uber Matrizen aus Positiven Elementen. Sitzungsberichte. 8, 471.

Gantmacher, F. R. 1959. Applications of the Theory of Matrices. New York: Interscience Publishers.

Gershgorin, S. 1931. Uber die Abgrenzung der Eigenwerte einer Matrix. Izv. Akad. Nauk. SSR. 7, 794.

Gowdy, C. M. 1978. Ph.D. Dissertation. Oklahoma State University, Stillwater, Oklahoma.

Hearon, J. Z. 1963. Theorems of Linear Systems. An. N.Y. Acad. Sci. 108, 36.

Henry, W. C. and Weidner, R. J. 1980. Decomposition and Evaluation of Multiple Binary Unit Markov Processes. Proc. of 11th SW. Grad. Res. Conf. in Appl. Mech. 4, 23.

Hevesy, G. 1923. The Absorption and Translocation of Lead by Plants. Biochem. J. 17, 437.

Householder, A. S. 1964. The Theory of Matrices in Numerical Analysis, New York: Dover Publications.

Jacquez, J. A. 1972. Compartmental Analysis in Biology and Medicine. Amsterdam: Elsevier.

Johnson, L. E. 1976. Control and Controllability in Compartmental Systems. Math. Biosci. 30, 181.

Kostitzin, V. A. 1935. Evolution de l'Atmosphere. Paris: Herman.

Levy, L. 1881. Sur La Possibilite' de Pequilibre Electrique. Comptes Rendus Acad. Sci. Paris. 93, 706.

Liff, A. I. and Wolf, J. K. 1966. On the Optimum Sampling Rate for Discrete Time Modeling of Continuous Time Systems. IEEE Trans on Auto. Contr. 288.

Lynch, R. E., Rice, J. R., and Thomas, D. H. 1964. Tensor Product Analysis of Partial Differential Equations. Bull. Amer. Math. Soc. 70, 378.

Marcus, M. 1960. Basic Theorems in Matrix Theory. NBS AMS 57, 4. Washington, D.C., U.S. Government Printing Office.

Matis, J. H. and Wehrly, T. E. 1979. Stochastic Models of Gompartment Systems. Biometrics, 35, 199.

Mulholland, R. J. and Keener, M. S. 1974. Analysis of Linear Compartment Models for Ecosystems. J. Theor. Biol. 43, 105.

Mulholland, R. J. and Gowdy, C. M. 1977. Theory and Application of the Measurement of Structure and Determination of Function for Laboratory Ecosystems. Jour. Theor. Biol. 69, 321.

Mulholland, R. J. and Weidner, R. J. 1980. Sampling Strategies for the Identification of Compartment Models. Int. J. Sys. Sci. 11, 377.

Neudecker, H. 1969. A Note on Kronecker Matrix Products and Matrix Equation Systems. SIAM J. Appl. Math. 17, 603.

Patel, R. V. and Toda. 1978. On Norm Bounds for Algebraic Riccatti and Lyapuov Equations. IEEE Trans. on Auto. Contr. AC-23, 87.

Patten, B. C. 1975. Ecosystem Linearization: An Evolutionary Design Problem. Amer. Natur. 109, 529.

Perron, O. 1907. Zur Theorie der Matrics. Math Ann., 685, 248.

Sage, A. P. and Melsa, J. L. 1971. Estimation Theory with Application to Communications and Control. New York: McGraw-Hill.

Sandberg, I. W. 1978. On the Mathematical Foundations of Compartmental Analysis in Biology, Medicine, and Ecology. IEEE Trans. Circuits Syst. CAS-25, 273.

Searle, S. R. 1971. Linear Models. New York: John Wiley and Sons.

Shannon, C. E. 1948. A Mathematical Theory of Communication. Bell Sys. Tech Jour. 27, 379 and 623.

Sheppard, C. W. 1948. The Theory of the Study of Transfers within a Multi-compartment System using Isotropic Tracers. J. Appl. Phys. 19, 70.

Sheppard, C. W. 1962. Basic Principles of the Tracer Method: Introduction of Mathematical Tracer Kinetics. New York: John Wiley and Sons, Inc.

Shipley, R. A. and Clark, R. E. 1972. Tracer Methods for IN VIVO Kinetics. New York: Academic Press.

Smith, F. E. 1970. Analysis of Ecosystems: In Analysis of Temperate Forest Ecosystems. D. E. Reichle ed. New York: Springer-Verlag.

Taussky, O. 1949. A Recurring Theorem on Determinants, Amer. Math. Month. 56, 672.

Teorell, T. 1937. Kinetics of Distribution of Substance Administered to the Body. I. The Extravascular Modes of Administration, Arch. Int. Pharmacodynamic Therapie. 57, 205. II. The Intravascular Modes of Administration, Arch. Int. Pharmacodynamic Therapie. 57, 226.

Thron, C. D. 1972. Structure and Kinetic Behavior of Linear Multi-Compartment Systems. Bull. Math. Biophy. 34, 277.

Wang, C. 1977. Resonance Effect in Compartmental Analysis and Its Detection. Math. Biosci. 36, 109.

Zadeh, L. A., and Plolak, E. 1969. System Theory. New York: McGraw-Hill.

APPENDIX

A KRONECKER ALGEBRA CATALOG

Notation

$\otimes$	Kronecker Product, Direct Product, Tensor Product
$\oplus$	Kronecker Sum
$D(A)$	Determinant of A
$R(A)$	Rank of A
$\rho(A)$	Spectral range of A
$\lambda_i(A)$	Eigenvalue i of A
$x_i(A)$	Eigenvector i of A
$\Lambda(A)$	Diagonal matrix formed with the eigenvalues of A
T	Matrix formed of the eigenvectors of A
S	Matrix formed of the eigenvectors of B
$tr(A)$	Trace or spur of A
e_i	Unit vector (1 in element i, 0 elsewhere)
E_{ik}	Elementary matrix $(e_i\ e_k')$
U	Permutation matrix (exactly one 1 in each row and column, 0 elsewhere)
A'	Transpose of matrix A
A^{-1}	Inverse of matrix A
$A^{[k]}$	A Kronecker power of A
$f(A)$	An analytic function of A
$\pi\otimes$	A Kronecker product series building to the left

$\Sigma\oplus$ A Kronecker sum series building to the left

A_r A row stack of matrix A (a column vector formed by the row vectors of A concatenated in ascending numerical order), i.e., for A - 2x2

$$A_r = (a_{11}\ a_{12}\ a_{21}\ a_{22})$$

A_c A column stack of A (a column vector formed by the column vectors of A concatenated in ascending numerical order), i.e., for A - 2x2

$$A_c = (a_{11}\ a_{21}\ a_{12}\ a_{22})$$

M A particular permutation matrix defined by $A_r = M\ A_c$

n The rank of matrix A (A is assumed square; see Brewer [1978] for nonsquare matrices)

m The rank of matrix B

The Kronecker Product

A. $A \otimes B = a_{ij}\ B$

I.e., for A - 2x2 and B - 2x2

$$A \otimes B = \begin{bmatrix} a_{11}\ b_{11} & a_{11}\ b_{12} & a_{12}\ b_{11} & a_{12}\ b_{12} \\ a_{11}\ b_{21} & a_{11}\ b_{22} & a_{12}\ b_{21} & a_{12}\ b_{22} \\ a_{21}\ b_{11} & a_{21}\ b_{12} & a_{22}\ b_{11} & a_{22}\ b_{12} \\ a_{21}\ b_{21} & a_{21}\ b_{22} & a_{22}\ b_{21} & a_{22}\ b_{22} \end{bmatrix}$$

B. 1. $\lambda(A \otimes B)$ is the set of mn numbers (Bellman, 1960)

$$\lambda_i(a)\ \lambda_j(B) \qquad i = 1, \ldots, n;\ j = 1, \ldots, m$$

2. $x(A \otimes B) = x_i(A) \otimes x_j(B)$ (Bellman, 1960)

C. $R(A \otimes B) = R(A)\ R(B)$ (Marcus, 1960)

D. $D(A \otimes B) = D^{m}(A)\ D^{n}(B)$ (Marcus, 1960)

E. $tr(A \otimes B) = tr(A)\ tr(B)$ (Marcus, 1960)

F. $(A \otimes B)' = A' \otimes B'$ (Marcus, 1960)

G. If A and B are Hermitian, $A \otimes B$ is Hermitian (Marcus, 1960)

H. If A^{-1} and B^{-1} exist, $(A \otimes B)^{-1}$ exists and

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$$ (Marcus, 1960)

I. $B \otimes A = M\ (A \otimes B)M$, where $m = n$ (Barnett, 1973)

1. $M = \sum_{i} \sum_{k} E_{ik} \otimes E_{ki}$ (Brewer, 1978)

2. $M = M^{-1} = M'$ (Brewer, 1978)

J. $(T^{-1} \otimes S^{-1})(A \otimes B)(T \otimes S) = \Lambda(A) \otimes \Lambda(B)$

where $T^{-1}AT = \Lambda(A)$, $S^{-1}BS = \Lambda(B)$

K. $\Lambda(A) \otimes \Lambda(B) = \Lambda(A \otimes B)$

L. $(A \otimes B) \otimes C = A \otimes (B \otimes C)$ (Bellman, 1960)

M. $(A + C) \otimes B = A \otimes B + C \otimes B$ (Marcus, 1960)

N. $(A + B) \otimes (C + D) = A \otimes C + A \otimes D + B \otimes C + B \otimes D$ (Bellman, 1960)

O. $(A \otimes B)(C \otimes D) = AC \otimes BD$ (Bellman, 1960)

P. $(I \otimes A)(B \otimes I) = (B \otimes I)(I \otimes A)$ (Neudecker, 1969)

Q. $f(I \otimes A) = I \otimes f(A)$ (Barnett, 1973)

R. $f(A \otimes I) = f(A) \otimes I$ (Barnett, 1973)

S. $\exp(I \otimes A) = I \otimes \exp(A)$ (Barnett, 1973)

T. $\exp(A \otimes I) = \exp(A) \otimes I$ (Barnett, 1973)

U. $A^{[k+1]} = A \otimes A^{[k]}$ (Bellman, 1960)

V. $(AD)^{[k]} = A^{[k]}\ D^{[k]}$ (Bellman, 1960)

W. $\prod_{i}^{n} \otimes A_i = A_n \otimes A_{n-1} \otimes \ldots A_2 \otimes A_1$

X. $(AB)_c = (I \otimes A)B_c = (B' \otimes I)A_c = (B' \otimes A)I_c$ (Neudecker, 1969)

Y. $(AB)_r = (A \otimes I)B_r = (I \otimes B')A_r = (A \otimes B')I_r$

Z. $(ABC)_r = (A \otimes C')B_r$

AA. $(ABC)c = (C' \otimes A)B_c$

(Note: Lynch et al. [1964] and Brewer [1978] have good extensions of the product application into the solution of partial differential equations and a matrix calculus.)

BB. The derivative of the Kronecker Product is defined by

$$d/dt(A(t) \otimes B(t)) = \dot{A}(t) \otimes B(t) + A(t) \otimes \dot{B}(t)$$

Proof

$$d/dt(A(t) \otimes B(t)) = \lim_{h \to o}(A(t+h) \otimes B(t+h) - A(t) \otimes B(t))/h$$

$$= \lim_{h \to o}(A(t+h) \otimes B(t+h) - A(t) \otimes B(t) + A(t+h) \otimes B(t)$$

$$- A(t+h) \otimes B(t))/h$$

$$= \lim_{h \to o}(A(t+h) \otimes (B(t+h) - B(t)) + (A(t+h)$$

$$-A(t)) \otimes B(t))/h$$

if A(t) and B(t) are differentiable at t

$$= \dot{A}(t) \otimes B(t) + A(t) \otimes \dot{B}(t)$$

This ends the proof.

CC. The exponential of the Kronecker product is

1. $\exp(A \otimes B) = I_{nm} + (A \otimes B) + 1/2\ (A \otimes B)(A \otimes B)' + \ldots$

2. $\exp(A \otimes B) = (T \otimes S)(\exp(\Lambda(A) \otimes \Lambda(B))(T \otimes S)^{-1}$

The Kronecker Sum

A. $A \oplus B' = A \otimes I_m + I_n \otimes B$

i.e., for A - 2x2 and B - 2x2

$$A \oplus B' = \begin{bmatrix} a_{11}+b_{11} & b_{12} & a_{12} & 0 \\ b_{21} & a_{11}+b_{22} & 0 & a_{12} \\ a_{21} & 0 & a_{22}+b_{11} & b_{12} \\ 0 & a_{21} & b_{21} & a_{22}+b_{22} \end{bmatrix}$$

B. 1. $\lambda(A \oplus B')$ is the set of nm numbers (Bellman, 1960)

$\lambda_i(A) + \lambda_j(B)$ $\quad i = 1, \ldots, n;\ j = 1, \ldots, m$

2. $x(A \oplus B')$ is the set of nm vectors

$x_i(A) \otimes x_j(B)$

C. $Tr(A \oplus B') = m\ tr(A) + n\ tr(B)$

D. If A and B are hermitian, $A \oplus B'$ is hermitian.

E. $(A \oplus B')' = A' \oplus B$

F. $(A \oplus B) \oplus C = A \oplus (B \oplus C)$ $\quad$ iff $C = C'$

G. $\sum_{i}^{n} \oplus A_i = A_n \oplus (A_{n-1} \oplus \ldots \oplus (A_2 \oplus A_1)))$

where the identity matrices satisfy the structure.

H. $\sin(A \oplus B') = \sin(A) \otimes \cos(B) + \cos(A) \otimes \sin(B)$ (Barnett, 1973)

I. $\Lambda(A \oplus B') = \Lambda(A) \oplus \Lambda(B)$

J. $(T^{-1} \otimes S^{-1})(A \oplus B')(T \otimes S) = \Lambda(A) \oplus \Lambda(B)$ (Lynch, 1964)

where $T^{-1} AT = \Lambda(A)$ and $S^{-1} BS = \Lambda(B)$

Proof

$$A \oplus B' = A \otimes I + I \otimes B$$

$$= (T\Lambda(A)T^{-1} \otimes SS^{-1}) + (TT^{-1} \otimes S\Lambda(B)S^{-1})$$

$$= (T \otimes S)(\Lambda(A) \otimes I)(T \otimes S)^{-1}$$

$$+ (T \otimes S)(I \otimes \Lambda(B))(T \otimes S)^{-1}$$

$$= (T \otimes S)(\Lambda(A) \otimes I + I \otimes \Lambda(B))(T \otimes S)^{-1}$$

$$= (T \otimes S)(\Lambda(A) \oplus \Lambda(B))(T \otimes S)^{-1}$$

End of proof.

K. The exponential of the Kronecker sum is defined by

1. $\exp(A \oplus B') = I_{nm} + (A \oplus B') + 1/2\ (A \oplus B')(A \oplus B')' + \ldots$

2. $= (T \otimes S)\ \exp(\Lambda(A) \oplus \Lambda(B))(T \otimes S)^{-1}$

L. $\exp(A \oplus B') = \exp(A) \otimes \exp(B)$

1. $\exp(A \oplus B') = \exp(A \otimes I + I \otimes B)$

$$= (\exp(A \otimes I))(\exp(I \otimes B))$$

$$= (\exp(A) \otimes I)(I \otimes \exp(B))$$

$$= \exp(A) \otimes \exp(B)$$

2. $\exp(A \oplus B') = ((T \otimes S)(\exp(\Lambda(A) \oplus \Lambda(B))(T \otimes S)^{-1}$

since

$$\exp(\lambda_i(A) + \lambda_j(B)) = \exp(\lambda_i(A))\ \exp(\lambda_j(B))$$

$$\exp(\Lambda(A) \oplus \Lambda(B)) = \exp(\Lambda(A)) \otimes \exp(\Lambda(B))$$

$$\exp(A \oplus B') = (T \otimes S)(\exp(\Lambda(A)) \otimes \exp(\Lambda(B)))(T \otimes S)^{-1}$$

$$\exp(A \oplus B') = \exp(A) \otimes \exp(B)$$

3. Refer to the proof in Chapter IV.

M. The derivative of the Kronecker sum is defined by

$$d/dt(A(t) \oplus B'(t)) = (\dot{A}(t) \oplus \dot{B}'(t))$$

Proof

$$d/dt\ (A(t) \oplus B'(t)) = d/dt\ (A(t) \otimes I + I \otimes B(t))$$

$$= d/dt\ (A(t) \otimes I) + d/dt\ (I \otimes B(t))$$

$$= \dot{A}(t) \oplus \dot{B}'(t)$$

where A and B are differentiable at t.

2

VITA

Richard Jay Weidner

Candidate for the Degree of

Doctor of Philosophy

Thesis: A STOCHASTIC EXTENSION OF THE STRUCTURAL PROPERTIES OF COMPARTMENT MODELS VIA A KRONECKER ALGEBRA

Major Field: Electrical Engineering

Biographical:

Personal Data: Born in Woodward, Oklahoma, August 18, 1952, the son of Mr. and Mrs. Richard A. Weidner.

Education: Graduated from Woodward High School, Woodward, Oklahoma, in May, 1970; received the Bachelor of Science in Electrical Engineering degree from Oklahoma State University, Stillwater, Oklahoma, in December, 1977; received the Master of Science degree in Electrical Engineering from Oklahoma State University, Stillwater, Oklahoma, in May, 1979; completed requirements for the Doctor of Philosophy degree at Oklahoma State University, Stillwater, Oklahoma, in May, 1981.

Professional Experience: Research Assistant, Electrical Engineering, Oklahoma State University, Stillwater, Oklahoma, 1977 to present; Research Consultant, Fife Controls Corp., Oklahoma City, Oklahoma, summer of 1980.