NASA CONTRACTOR REPORT 166300

Stochastic Control System Parameter Identifiability

Chih-hsiao Lee
Charles J. Herget

CONTRACT NCAR-~~504~~ 340-402
May 1975

NASA

NASA CONTRACTOR REPORT 166300

Stochastic Control System Parameter Identifiability

Chih-hsiao Lee
Charles J. Herget
Engineering Research Institute
Iowa State University
Ames, Iowa 50010

Prepared for
Ames Research Center
under Contract NCAR 340-402

National Aeronautics and
Space Administration

Ames Research Center
Moffett Field, California 94035

TABLE OF CONTENTS

1. INTRODUCTION

The problem of system identification has been studied intensively and has become one of the most active fields in engineering research. Some investigators have obtained theoretical results on parameter identifiability, convergence properties of parameter estimates [1,16,18,19], and regions of parameter identifiability [7]. There are also investigators working on developing computational algorithms for identifying certain control system parameters, for example, the stability and control derivatives of an aircraft from actual flight test data [17].

There are two main steps in solving the system identification problem: (1) Determine if the system parameters are identifiable, and (2) if the parameters are identifiable, work out an algorithm for estimating the parameters. To answer (1), we have to establish the definition of identifiability first. The most significant recent work on parameter identifiability when there are stochastic disturbances present in the system are Tse [18] and Tse and Anton [19]. There are a number of algorithms that have been proposed for solving (2).

Generally, if the system parameters are identifiable, they are only locally identifiable, i.e., we must have a sufficiently good initial estimate of the parameters such that the iterative estimation sequence will converge to the true parameters. Herget [7] provided a procedure for computing explicit regions in the parameter space in which the Gauss-Newton method will converge to a unique solution. The systems considered by him were deterministic.

The definition of parameter identifiability when there are stochastic disturbances present has been given by Tse and Anton [19]. They said the parameters are identifiable if there exists a sequence of estimates which is consistent in probability. They also established the necessary and sufficient conditions for the unknown parameters to be identifiable under some uniformity assumptions on the conditional density parametrized by the unknown parameters. Tse [18] also gave the definition of local parameter identifiability and proved that the positive definiteness of the average information matrix implies local parameter identifiability. The definition of local parameter identifiability is that there exists an open region containing the vector of unknown parameters as an interior point and there exists a local estimation sequence in the closure of the region which is consistent in probability. Staley and Yue [16] established a similar concept on stochastic parameter identifiability. They stated that the parameters are identifiable if every asymptotically efficient estimator converges to the true parameters in mean square (which is stronger than convergence in probability.)

Wald [20,21] considered the consistency and asymptotic properties of the maximum likelihood estimation sequence. He showed that under certain restrictions on the joint probability distribution of the observations, the maximum likelihood equation has at least one set of roots which is a consistent estimate of the unknown parameters. He also showed that any root of the maximum likelihood equation which is a consistent estimate of the parameters is asymptotically efficient.

Based mainly on Wald-Kendall-Åström theory, Aoki and Yue [1] examined the asymptotic properties of the maximum likelihood estimates of unknown parameters of a class of linear, stable, constant, discrete-time dynamic systems form where plant noise and observation noise are present. The systems considered by them were restricted to have certain cononical structure and were single-input and single-output.

For the identification of linear dynamical systems, Glover and Willems [6] established the concept of parametrization and developed sufficient conditions for local and global identifiability from the transfer function. Bellman and Åström [2] also provided an algorithm-oriented least-square identifiability. It can be shown that for single-input, zero state systems, local least-square identifiability is equivalent to local transfer-function identifiability under some assumptions on the least square criterion. The remaining second step of the system identification is to identify the system parameter exactly from input-output sequences for deterministic systems or to construct a consistent estimation sequence by using the constrained maximum likelihood method for stochastic systems. In both cases, it becomes an optimization problem, i.e., we first set a performance criterion which is a function of the unknown parameters and then find best estimates such that the cost function will reach its minimum.

In general, the cost function and its derivatives are nonlinear and an iterative procedure must be used to find the estimate. The best known method of solving a set of simultaneous nonlinear equations in which the increment in each iteration is computed as a linear combination

of the residuals is the Newton method. Kantorovich's Theorem [12] states certain sufficient conditions for the convergence of the Newton iteration sequence. However, these sufficient conditions are generally highly restrictive and are not easily examined. Moreover, in each iteration, we have to calculate the Hessian matrix of the cost function which includes the calculation of a bilinear form. These are the two main disadvantages of using the Newton method from the point of view of practical computation.

Later researchers have developed some modified versions of the Newton method. The most significant for the problem of interest here are by Ben-Israel [3] and Pereyra [13]. The main features of both of their algorithms are: first, we only need to compute the first order approximation of the Hessian matrix of the cost function; second, the sufficient conditions for the convergence of the iteration sequence are much easier to examine. Although Ben-Israel's algorithm and Pereyra's algorithm are the same if the first order approximation of the Hessian matrix of the cost function has full rank, Pereyra's sufficient conditions are preferable again from the point of view of practical calculation [13].

In the deterministic system identification problem, the solution for the cost function is the true parameter, i.e., the cost function will reach its global minimum, zero, if the output sequence generated by the estimate matches the measured output sequence. For a known system structure, i.e., a given parametrization, there may be more than one isolated point in the parameter space that will generate the same output

sequence for a given input sequence, i.e., the solution to the identification criterion is generally not unique globally. Herget [7] provided a modified version of Pereyra's theorem and a computation procedure employing interval arithmetic to find explicitly the regions centered at each local solution in which the solution is unique and hence is locally identifiable. The other feature of his work is the use of bilinear operators to represent the linear system model. In doing so, the identification problem of linear systems is equivalent to the initial-state observation problem of quadratic-in-the-state bilinear systems.

This dissertation considers the parameter identification problem of general discrete-time, nonlinear, multiple-input/multiple-output dynamic systems with Gaussian-white distributed measurement errors. The knowledge of the system parametrization is assumed to be known. Concepts of local parameter identifiability and local constrained maximum likelihood parameter identifiability are established. A set of sufficient conditions for the existence of a region of parameter identifiability is proposed. A computation procedure employing interval arithmetic is derived for finding the regions of parameter identifiability. It is shown that if the vector of the true parameters is locally constrained maximum likelihood identifiable, then with probability one, the vector of true parameters is a unique maximal point of the maximum likelihood function in the region of parameter identifiability and the constrained maximum likelihood estimation sequence will converge to the true parameters.

Chapter 2 is a review of Wald's theory [20,21] on the consistency of maximum likelihood estimates, the concepts of parameter identifiability

and local parameter identifiability established by Tse and Anton [19] and by Tse [18] and the application of these concepts to the system identification problems [16].

Chapter 3 is a survey of Bellman and Åström's [2] definition of least-square parameter identifiability, and Glover and Willem's [6] concepts of system parametrization and identifiability from the transfer-function. The author's contributions in this chapter are as follows. Sufficient condition for local least-square parameter identifiability is derived by employing the constant rank theorem [10]. The Theorem of Glover and Willems is modified to provide a sufficient condition for local parameter identifiability of minimal dimensional linear dynamic systems whose initial states are unknown, and a theorem is established to show that for single-input, zero-state linear systems, local least-square parameter identifiability is equivalent to local parameter identifiability from the transfer function if some constant rank assumptions on the impulse response matrix and its derivatives are satisfied.

Chapter 4 is a survey of the Newton-Kantorovich theory [12] on the convergence of the Newton iteration method, Pereyra's theory [13] on solving nonlinear least-square problems, ℓ_∞-norms, interval arithmetic [11], and Herget's results on regions of parameter identifiability [7] with application to the parameter identification problem of deterministic dynamic systems. A numerical example is provided by the author with the computer program listed in the Appendix.

Chapter 5 contains the principal new results of this dissertation. It considers the parameter identification problem of general discrete-time multiple-input/multiple-output dynamic systems with Gaussian-white distributed measurement errors. It is mainly a modification of Herget's results [7] on the parameter identification problem of deterministic systems and a generalization of Aoki and Yue's result [1] on the parameter identification problem of single-input/single-output canonical-form linear dynamic systems with measurement noise. A numerical example is included to illustrate the computation procedure for finding the regions of CML parameter identifiability.

Chapter 6 gives the conclusions of this dissertation and suggestions for further research.

2. MAXIMUM LIKELIHOOD ESTIMATION, PARAMETER IDENTIFIABILITY AND LOCAL PARAMETER IDENTIFIABILITY

Relevant past investigations of maximum likelihood estimation parameter identifiability, and local parameter identifiability are discussed here. The first result is A. Wald's theory [20,21] on the asymptotic properties and the consistency of the maximum likelihood estimate of an unknown parameter of a discrete process. Then we will discuss Tse and Anton's [19] definition on stochastic identifiability, the necessary and sufficient conditions for the unknown parameter to be identifiable under some uniformity assumptions on the conditional density parametrized by the unknown parameter, and Tse's [18] definition of local identifiability. An identification problem of a class of linear, stable, constant, discrete-time, single-input/single-output dynamical systems discussed by Aoki and Yue [1] will also be presented.

2.1 Preliminary Concepts

Definition 2.1

Let X denote an arbitrary nonempty set. A family of subsets R of X is called a sigma field if

(i) for every $A \in R$, then also $A^c \in R$ where A^c is the complement of A

(ii) if $A_1, A_2, \cdots, A_n, \cdots$ is a countable sequence of elements of R, then $\bigcup_{n=1}^{\infty} A_n \in R$, and

(iii) $\phi \in R$ where ϕ denotes the empty set.

(iv) Elements of R are called events.

Definition 2.2

A probability Pr. is a measure over a measurable space (X,R); that is, Pr. is a real-valued function which assigns to every $A \in R$ a number Pr.(A) such that

(i) $Pr.(A) \geq 0$ for every $A \in R$

(ii) $Pr.(X) = 1$, and

(iii) if $\{A_n\}_{n=1}^{\infty}$ is any countable union of disjoint events, then

$$Pr.\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} Pr.(A_n)$$

Definition 2.3

Let X be a set, R be a sigma field and Pr. a probability measure defined on R, then the triplet (X,R,Pr.) is called a probability space.

Definition 2.4

A random variable x is a real-valued function whose domain is X and which is R-measurable, that is, for every real number λ,

$$\{\omega \in X \mid x(\omega) \leq \lambda\} \in R.$$

Definition 2.5

Let $x_1, x_2, \cdots, x_n, \cdots$ be a sequence of random variables, if there exists a random variable $x(\omega)$ such that

$\lim_{n \to \infty} x_n(\omega) = x(\omega)$ for almost all ω, we say that

$\lim_{n \to \infty} x_n = x$ with probability one

Definition 2.6

Let $x_1, x_2, \cdots, x_n, \cdots$, be a sequence of random variables, we say the sequence $\{x_n\}_{n=1}^{\infty}$ converges to x in probability or converges stochastically to x if

$$\lim_{n \to \infty} \Pr.\{|x_n(\omega) - x(\omega)| \geq \epsilon\} = 0$$

is satisfied for every $\epsilon > 0$

Definition 2.7

A sequence of random variables $\{x_n\}_{n=1}^{\infty}$ is said to converge to x in the mean square sense if

(i) $E\{|x_n|^2\} < \infty$ for all n,

(ii) $E\{|x|^2\} < \infty$, and if

(iii) $\lim_{n \to \infty} E\{|x - x_n|^2\} = 0$.

This is written

$$\text{l.i.m.}_{n \to \infty} x_n = x .$$

Definition 2.8

If x is a random variable, its distribution function F_x is defined by

$$F_x(\xi) = \Pr.[x \leq \xi] \text{ for all } \xi \in (-\infty, \infty)$$

Definition 2.9

A distribution function F is said to be absolutely continuous if there exists a Borel measurable function p over $(-\infty, \infty)$ such that

$$F(\xi) = \int_{-\infty}^{\xi} p(t)dt$$

for all ξ. The function p is called a density of F.

Definition 2.10

Let $x_1,\cdots,x_n$ be random variables, $n \geq 1$. The joint distribution function of $x_1,\cdots,x_n$, or the distribution function of the random vector $x = (x_1,\cdots,x_n)$, is defined to be

$$F_x(\xi_1,\cdots,\xi_n) = \Pr.(\bigcap_{i=1}^{n}[x_i \leq \xi_i])$$

where $-\infty < \xi_i < \infty$, $1 \leq i \leq n$.

Definition 2.11

Let $\{x_\lambda, \lambda \in \Lambda\}$ be a family of random variables. They are said to be independent if for every positive integer n and every n distinct elements $\lambda_1,\cdots,\lambda_n$ in Λ, then

$$F_{x_{\lambda_1},\cdots,x_{\lambda_n}}(\xi_1,\cdots,\xi_n) = \prod_{j=1}^{n} F_{x_{\lambda_j}}(\xi_j)$$

for all $\underline{\xi} \in R^n$.

If $F_{x_{\lambda_i}}(\xi_i) = F_{x_{\lambda_j}}(\xi_j)$ for all $\lambda_i, \lambda_j \in \Lambda$, then $\{x_\lambda, \lambda \in \Lambda\}$ are said to be independently and identically distributed.

Let $x_1, x_2, \cdots, x_n, \cdots$ be a sequence of random variables with joint probability denstiy function $p(x_1, x_2, \cdots, x_n; \underline{\theta})$, $n = 1,2,\cdots$, which is of known functional form but $p(x_1,\cdots,x_n; \underline{\theta})$ depends upon an unknown vector of parameters $\underline{\theta}$ that may have any value in a set Ω. This set will be called the parameter space. Thus we are confronted with a family of joint probability density functions denoted by $\{p(x_1,\cdots,x_n; \underline{\theta}): \underline{\theta} \in \Omega, n = 1,2,\cdots\}$. To each value of $\underline{\theta}$, $\underline{\theta} \in \Omega$, there corresponds one member of the family denoted by $\{p(x_1,\cdots,x_n; \underline{\theta}): n = 1,2,\cdots\}$ which is a sequence of joint probability density functions parametrized by $\underline{\theta}$. Let $\{p(x_1,\cdots,x_n; \underline{\theta}_0)$:

$n = 1,2,---\}$ be a member of the family and let $\underline{\theta}_0$ be unknown, $\underline{\theta}_0 \in \Omega$. An estimate of $\underline{\theta}_0$ based on the observation sequence $x_1, x_2, ---x_n$, $n = 1,2,---$ is a measurable function of $\{x_1, x_2, ---, x_n\}$, $n = 1,2,---$, and is denoted by

$$\underline{\hat{\theta}}_n = \underline{\hat{\theta}}_n(x_1, ---, x_n), \quad n = 1,2,---.$$

$\{\underline{\hat{\theta}}_n\}_{n=1}^{\infty}$ is then called an estimation sequence.

Definition 2.12

Any estimation sequence $\{\underline{\hat{\theta}}_n\}_{n=1}^{\infty}$ of $\underline{\theta}_0 \in \Omega$ which converges stochastically to $\underline{\theta}_0$ is called a consistent estimate for $\underline{\theta}_0$.

Lemma 2.1.a [The Strong Law of Large Numbers] [4]

Let $x_1, x_2, ---, x_k, ---$, be independent random variables such that $E\, x_k = 0$, $E\, x_k^2 < \infty$. Let $b_n \geq 0$ converges up to $+\infty$.

If $\sum_{k=1}^{\infty} E\, x_k^2 / b_k^2 < \infty$, then

$$\Pr.\{\lim_{n \to \infty} \frac{x_1 + --- + x_n}{b_n} = 0\} = 1.$$

Lemma 2.1.b

Let $x_1, x_2, ---, x_k, ---$ be independent and identically distributed random variables.

If $E|x_1| < \infty$, then

$$\Pr.\{\lim_{n \to \infty} \frac{x_1 + --- + x_n}{n} = E\, x_1\} = 1$$

The concepts introduced above can be found in references [4], [5], and [9].

2.2 The Consistency of Maximum Likelihood Estimates

This section summarizes the theory given by Wald in [20,21]. Let $\{z_n\}_{n=1}^{\infty}$ be a sequence of independent, identically distributed random variables with joint probability density function $p(z_1,\cdots,z_n;\ \underline{\theta})$, $n = 1,2,\cdots$, parametrized by the unknown parameter $\underline{\theta}\in\Omega\subset R^p$, where Ω is the parameter space. Let $\|\cdot\|$ be a norm on R^p. Let $p(z;\ \underline{\theta})$ denote the probability density function and $F(z,\underline{\theta})$ denote the corresponding cumulative distribution function of z_i, i.e., $F(z;\theta) = Pr.\{z_i \le z\}$.

The following assumptions are made.

Assumption 1

$F(z;\ \underline{\theta})$ is either discrete or is absolutely continuous for all $\underline{\theta}\in\Omega$.

For the next assumption, we introduce the following notation: for $\underline{\theta}\in\Omega$ and $\rho > 0$ let $p(z;\ \underline{\theta},\rho)$ be

$$p(z;\ \underline{\theta},\rho) = \sup_{\|\underline{\theta} - \underline{\theta}'\| \le \rho} p(z;\ \underline{\theta}')$$

For any $r > 0$, let $\psi(x,r)$ be

$$\psi(z,r) = \sup_{\|\underline{\theta}\| > r} p(z;\underline{\theta})$$

Furthermore, let

$$p^*(z;\underline{\theta},\rho) = \begin{cases} p(z;\underline{\theta},\rho) & \text{if } p(z;\underline{\theta},\rho) > 1 \\ 1 & \text{otherwise} \end{cases}$$

Similarly, let

$$\psi^*(z,r) = \begin{cases} \psi(z,r) & \text{if } \psi(z,r) > 1 \\ 1 & \text{otherwise} \end{cases}$$

Assumption 2

For sufficiently small ρ and for sufficiently large r,

$$\int_{-\infty}^{\infty} \log p^*(z;\underline{\theta},\rho)\, dF(z;\underline{\theta}_0) < \infty$$

and

$$\int_{-\infty}^{\infty} \log \psi^*(z,r)\, dF(z;\underline{\theta}) < \infty \quad \text{for all } \underline{\theta}\in\Omega$$

where $\underline{\theta}_0$ is the true parameter point.

Assumption 3

If $\lim_{i\to\infty} \underline{\theta}_i = \underline{\theta}$, then $\lim_{i\to\infty} p(z;\underline{\theta}_i) = p(z;\underline{\theta})$ for all z except perhaps on a set whose probability measure is zero according to the probability distribution corresponding to $\underline{\theta}_0$.

Assumption 4

If $\underline{\theta}_1 \neq \underline{\theta}_0$, then $F(z;\underline{\theta}_1) \neq F(z;\underline{\theta}_0)$ for at least one value of z.

Assumption 5

If $\lim_{i\to\infty} \|\underline{\theta}_i\| = \infty$, then $\lim_{i\to\infty} p(z;\underline{\theta}_i) = 0$ for every z except perhaps on a fixed set whose probability measure is zero according to $\underline{\theta}_0$.

Assumption 6

$$\int_{-\infty}^{\infty} \left|\log p(z;\underline{\theta}_0)\right| dF(z;\underline{\theta}_0) < \infty$$

Assumption 7

Ω is a closed subset of R^p.

Assumption 8

$p(z;\underline{\theta},\rho)$ is a measurable function of z for $\underline{\theta}\in\Omega$ and $\rho > 0$.

Lemma 2.2

If $\underline{\theta} \neq \underline{\theta}_0$, then

$$E \log p(z; \underline{\theta}) < E \log p(z; \underline{\theta}_0) \tag{2.1}$$

Lemma 2.3

$$\lim_{\rho \to 0} E \log p(z; \underline{\theta}, \rho) = E \log p(z; \underline{\theta}_0) \tag{2.2}$$

Lemma 2.4

$$\lim_{r \to \infty} E \log \psi(z,r) = -\infty \tag{2.3}$$

By the above lemmas, we can prove the following theorems.

Theorem 2.5

Let $W \subset \Omega$ be a closed subset of Ω. If $\underline{\theta}_0$ does not belong to W, then

$$\Pr.\left\{\lim_{n \to \infty} \frac{\sup_{\underline{\theta} \in W} (z_1, z_2, \cdots, z_n; \underline{\theta})}{p(z_1, z_2, \cdots, z_n; \underline{\theta}_0)} = 0\right\} = 1. \tag{2.4}$$

Proof: By Lemma 2.4 we can choose $r_0 > 0$ such that

$$E \log \psi(z, r_0) < E \log p(z; \underline{\theta}) \tag{2.5}$$

Let W_1 be the subset of W such that

$$W_1 = \{\underline{\theta}: \|\underline{\theta}\| \leq r_0, \ \underline{\theta} \in W\}$$

For each $\underline{\theta} \in W_1$, we can choose a $\rho_{\underline{\theta}} > 0$ such that

$$E \log p(z; \underline{\theta}, \rho_{\underline{\theta}}) < E \log p(z; \underline{\theta}_0) \tag{2.6}$$

The existence of $\rho_{\underline{\theta}}$ is guaranteed by Lemmas 2.1 and 2.2. The set W_1 is closed and bounded; hence it is compact. Thus there exists a finite number of points $\underline{\theta}_1,\cdots,\underline{\theta}_k$ in W_1 such that the union of the spheres with center θ_i and radious $\rho_{\underline{\theta}_i}$, $i = 1,\cdots,k$, $\bigcup_{i=1}^{k} S(\underline{\theta}_i,\rho_{\underline{\theta}_i})$ covers W_1.

We see that

$$0 \le \sup_{\underline{\theta}\in} p(z_1,z_2,\cdots,z_n;\underline{\theta}) \le \sum_{i=1}^{k} p(z_1;\underline{\theta}_i,\rho_{\underline{\theta}_i}) \cdots p(z_n;\underline{\theta}_i,\rho_{\underline{\theta}_i})$$

$$+ \psi(z_1,r_0)\cdots\psi(z_n,r_0).$$

Thus we are going to show

$$\Pr.\{\lim_{n\to\infty} \frac{p(z_1;\underline{\theta}_i,\rho_{\underline{\theta}i})\cdots p(z_n;\underline{\theta}_i,\rho_{\underline{\theta}i})}{p(z_1;\underline{\theta}_0)\cdots (z_n;\underline{\theta}_0)} = 0\} = 1,\ i = 1\cdots,k \quad (2.7)$$

and

$$\Pr.\{\lim_{n\to\infty} \frac{\psi(z_1,r_0)\cdots\psi(z_n,r_0)}{p(z_1;\underline{\theta}_0)\cdots p(z_n;\underline{\theta}_0)} = 0\} = 1 \quad (2.8)$$

which is equivalent to showing that

$$\Pr.\{\lim_{n\to\infty} \sum_{j=1}^{n}[\log p(z_j;\underline{\theta}_i,\rho_{\underline{\theta}i}) - \log p(z_j;\underline{\theta}_0)] = -\infty\} = 1$$

$$i = 1,\cdots,k \quad (2.9)$$

and

$$\Pr.\{\lim_{n\to\infty} \sum_{j=1}^{n}[\log \psi(z_j,r_0) - \log p(z_j;\underline{\theta}_0)] = -\infty\} = 1. \quad (2.10)$$

But (2.9) and (2.10) follow immediately from (2.5), (2.6) and the strong law of large numbers.

Theorem 2.6

Let $\hat{\underline{\theta}}_n(z_1,\cdots,z_n)$ be a function of the observations such that

$$\frac{p(z_1,\cdots,z_n;\ \underline{\theta}_n)}{p(z_1,\cdots,z_n;\ \underline{\theta}_0)} \geq c>0 \quad \text{for all } n \text{ and for all } z_1,\cdots,z_n \tag{2.11}$$

Then

$$\Pr.\{\lim_{n\to\infty} \hat{\underline{\theta}}_n = \underline{\theta}_0\} = 1.$$

Proof: Let Ⓗ denote the set of limit points of $\{\hat{\underline{\theta}}_n\}_{n=1}^{\infty}$, then it suffices to show that for any $\epsilon>0$,

$$\sup\{\|\hat{\underline{\theta}} - \underline{\theta}_0\|:\ \hat{\underline{\theta}} \in Ⓗ\} \leq \epsilon \text{ with probability one} \tag{2.12}$$

Suppose that there exists a $\hat{\underline{\theta}} \in Ⓗ$ such that $\|\hat{\underline{\theta}} - \underline{\theta}_0\| > \epsilon$, then

$$\sup_{\|\theta - \theta_0\| \geq \epsilon} p(z_1,\cdots,z_n;\ \underline{\theta}) \geq p(z_1,\cdots,z_n;\ \hat{\underline{\theta}}_n)$$

for infinitely many n. But this implies

$$\frac{\sup_{\|\theta - \theta_0\| \geq \epsilon} p(z_1,\cdots,z_n;\ \underline{\theta})}{p(z_1,\cdots,z_n;\ \underline{\theta}_0)} \geq c>0 \tag{2.13}$$

for infinitely many n by (2.11).

By Theorem 2.5, (2.13) is an event with probability zero, thus (2.12) holds with probability one. We recall that the maximum likelihood estimate $\hat{\underline{\theta}}_n$ is obtained by

$$p(z_1,\cdots,z_n;\ \hat{\underline{\theta}}_n) = \operatorname*{Max}_{\underline{\theta}\in\Omega} p(z_1,\cdots,z_n;\ \underline{\theta}) \quad \text{for all } n \tag{2.14}$$

If $\hat{\underline{\theta}}_n$ exists, then

$$\frac{p(z_1,\cdots,z_n;\ \hat{\underline{\theta}}_n)}{p(z_1,\cdots,z_n;\ \underline{\theta}_0)} \geq 1 \quad \text{for all } n \text{ and for all } x_1,\cdots,x_n.$$

By Theorem 2.6, the maximum likelihood estimate is consistent.

2.3 Parameter Identifiability and Local Parameter Identifiability

This section summarizes the work by Tse and Anton in [19] and by Tse in [18].

Let $\{z_n\}_{n=1}^{\infty}$ be a sequence of observation statistics with joint probability density function $p(z_1,---,z_n;\underline{\theta})$, $n=1,2,---$, parametrized by the unknown parameter $\underline{\theta}\in\Omega\subset R^p$, where Ω is the parameter space which is a compact subset of R^p. Let $\|\cdot\|$ be a norm on R^p. The true parameter $\underline{\theta}_0$ is known to lie in the interior of Ω. The parameter identification problem is to estimate the true parameter $\underline{\theta}_0$ based on the observation sequence $\{z_n\}_{n=1}^{\infty}$.

Definition 2.13

The parameter $\underline{\theta}_0$ is said to be identifiable if there exists a sequence of estimates $\{\hat{\underline{\theta}}_n\}_{n=1}^{\infty}$ which is consistent in probability, i.e., for any δ, ϵ arbitrarily small, there exists an $N(\delta,\epsilon)$ such that for $n>N(\delta,\epsilon)$

$$\Pr.\{\|\hat{\underline{\theta}}_n - \underline{\theta}_0\| > \delta\} < \epsilon. \tag{2.15}$$

For brevity, we let

$$p(Z_n;\underline{\theta}) \triangleq p(z_1,---,z_n;\theta), \tag{2.16}$$

and

$$p(z_n|Z_{n-1};\theta) = p(Z_n;\underline{\theta})/p(Z_{n-1};\underline{\theta}) \quad \text{for } n=1,2,--- \tag{2.17}$$

For $\underline{\theta}\in\Omega$ and $\rho>0$, let

$$p(z_n,\rho|Z_{n-1};\underline{\theta}) = \sup_{\|\underline{\theta}-\underline{\theta}'\| \le \rho} p(z_n|Z_{n-1};\underline{\theta}') \tag{2.18}$$

The following assumptions are made.

Assumption 1

$p(Z_n,\underline{\theta})$ is measurable in Z_n with respect to $p(Z_n;\underline{\theta}_0)\,dZ_n$ and continuous in $\underline{\theta}\in\Omega$ for Z_n almost everywhere, i.e., for $\epsilon>0$ and $\underline{\theta}\in\Omega$, there exists a $\delta(\epsilon)>0$ such that for all $\underline{\theta}'\in\Omega$ with $\|\underline{\theta}-\underline{\theta}'\|<\delta$ we have $|p(Z_n;\underline{\theta}) - p(Z_n;\underline{\theta}')|<\epsilon$ for Z_n almost everywhere.

Assumption 2

$$\int_{R^n} \log p(z_n,\rho|Z_{n-1};\underline{\theta})\, p(Z_n;\underline{\theta}_0)\, dZ_n < \infty \quad \text{for some } \rho>0 \text{ and } \underline{\theta}\in\Omega \tag{2.19}$$

and

$$\int_{R^n} \log p(z_n|Z_{n-1};\underline{\theta}_0)\, p(Z_n;\underline{\theta}_0)\, dZ_n < \infty \quad \text{for all } n=1,2,\text{---} \tag{2.20}$$

Assumption 3

For all $\underline{\theta}\in\Omega$ and some $\rho_0>0$,

$$\mathrm{Var.}\{\sum_{k=1}^{n} \log p(z_k,\rho|Z_{k-1};\underline{\theta})\} = O(n^2) \tag{2.21}$$

for all $0\le\rho\le\rho_0$, where $O(n^2)$ is such that

$$\lim_{n\to\infty} \frac{O(n^2)}{n^2} = 0$$

Assumption 4

Let the set $B_n(\underline{\theta})$ be

$$B_n(\underline{\theta}) \equiv \{Z_n:\ p(Z_n;\theta)=0\}$$

then for all $\underline{\theta}_1, \underline{\theta}_2 \in \Omega$, we have

$$B_n(\underline{\theta}_1) = B_n(\underline{\theta}_2) \quad \text{for } n = 1,2,--- \tag{2.22}$$

Since the only information about $\underline{\theta}_0$ is the observation sequence statistics $\{z_n\}_{n=1}^{\infty}$ with their joint density function $p(Z_n; \underline{\theta})$, $n = 1,2,---$, if there are two points $\underline{\theta}_1, \underline{\theta}_2 \in \Omega$, $\underline{\theta}_1 \neq \underline{\theta}_2$, such that

$$p(Z_n; \underline{\theta}_1) = p(Z_n; \underline{\theta}_2) \tag{2.23}$$

or

$$p(z_n | Z_{n-1}; \underline{\theta}_1) = p(z_n | Z_{n-1}; \underline{\theta}_2) \quad \text{for all } n \tag{2.24}$$

we are not able to distinguish $\underline{\theta}_1$ and $\underline{\theta}_2$ in Ω.

Definition 2.14

Two parameters $\underline{\theta}_1, \underline{\theta}_2 \in \Omega$, $\underline{\theta}_1 \neq \underline{\theta}_2$ are said to be unresolvable if the equality

$$p(z_n | Z_{n-1}; \underline{\theta}_1) = p(z_n | Z_{n-1}; \underline{\theta}_2) \tag{2.25}$$

holds with probability one for all except a finite number of integers $n > 0$, i.e., if (2.25) holds with respect to the measure $p(Z_n; \underline{\theta}_1)\, dZ_n$ as well as $p(Z_n; \underline{\theta}_2)\, dZ_n$.

Definition 2.15

The set Ω is said to be identifiable if no two elements in Ω are unresolvable.

By using the constrained maximum likelihood method, the identification problem is: find $\hat{\underline{\theta}}_n \in \Omega$ as an estimate of $\underline{\theta}_0$ such that

$$p(Z_n; \hat{\underline{\theta}}_n) = \max_{\underline{\theta} \in \Omega} p(Z_n; \underline{\theta}), \quad n = 1,2,--- \tag{2.26}$$

Since Ω is compact and hence is closed and bounded, and $p(Z_n; \underline{\theta})$ is continuous in $\underline{\theta}$ by assumption 1, a solution to (2.26) exists and the estimate sequence $\{\underline{\hat{\theta}}_n\}_{n=1}^{\infty}$ is a consistent estimate for $\underline{\theta}_0$ if $\underline{\theta}_0$ is unique. This follows from Theorem 2.6. However, if there are two points $\underline{\theta}_1, \underline{\theta}_2 \in \Omega$, $\underline{\theta}_1 \neq \underline{\theta}_2$, such that

$$\lim_{n \to \infty} p(Z_n; \underline{\theta}_1) = \lim_{n \to \infty} p(Z_n; \underline{\theta}_2) \tag{2.27}$$

then it is obvious that $\{\underline{\hat{\theta}}_n\}_{n=1}^{\infty}$ will fail to converge.

Definition 2.16

Two parameters $\underline{\theta}_1, \underline{\theta}_2 \in \Omega$, $\underline{\theta}_1 \neq \underline{\theta}_2$, are said to be CML unresolvable if

$$\lim_{n \to \infty} p(z_n | Z_{n-1}; \underline{\theta}_1) = \lim_{n \to \infty} p(z_n | Z_{n-1}; \underline{\theta}_2) \tag{2.28}$$

with probability one.

Definition 2.17

The set Ω is said to be CML identifiable if no two elements in Ω are CML unresolvable.

The following theorem was given by Tse and Anton in [19].

Theorem 2.7

If for all $\theta_1, \theta_2 \in \Omega$, $\underline{\theta}_1 \neq \underline{\theta}_2$, there exists an countably infinite set $L \subset I^+$ (I^+ = the set of positive integers) such that

$$p(z_n | Z_{n-1}; \underline{\theta}_1) \neq p(z_n | Z_{n-1}; \underline{\theta}_2)$$

with nonzero probability with respect to $\underline{\theta}_1$ and $\underline{\theta}_2$ uniformly in $n \in L$, then Ω is CML identifiable.

The above concepts on parameter identifiability can be applied to the system identification problem.

Consider a linear discrete-time system described by:

$$\underline{x}(k+1) = F\underline{x}(k) + G\underline{u}(k)$$
$$\underline{z}(k) = Hx(k) + \underline{v}(k) \qquad (2.29)$$

where F is the $(n \times n)$ state transition matrix

G is the $(n \times q)$ input matrix

H is the $(r \times n)$ output matrix

$\underline{x}(k)$ is the n-state vector

$\underline{z}(k)$ is the r-output vector

$\underline{u}(k)$ is the q-input vector

$\underline{v}(k)$ is a Gaussian white noise with zero mean and convariance matrix Q.

Let the initial state $\underline{x}(0) \equiv \underline{x}_0$. The parameter $\underline{\theta}_0 = \{\underline{x}_0, F_0, G_0, H_0, Q_0\}$ is to be identified. We assume $\underline{\theta}_0 \in \Omega \subset R^p$ where Ω is a compact subset of R^p. Furthermore, we assume that

(1) the system is stable for all $\underline{\theta} \in \Omega$,

(2) the system is completely controllable, i.e.,
$\text{rank}[G, FG, \cdots, F^{n-1}G] = n$ for all $\underline{\theta} \in \Omega$, and

(3) the system is completely observable, i.e,
$\text{rank}[H^T, (HF)^T, \cdots, (HF^{n-1})^T] = r$ for all $\underline{\theta} \in \Omega$.

The assumption of controllability and observability implies that the system is of minimal dimension and equivalent systems for (2.29) exist. The joint probability density function of the observation sequence $\{\underline{z}(0), \cdots, \underline{z}(M)\}$, $M = 0, 1, 2, \cdots$, is given by

$$p(\underline{z}(0),\cdots,\underline{z}(M);\underline{\theta}) = \prod_{k=0}^{M} p_{v(k)}(\underline{z}(k) - H\underline{x}(k);\underline{\theta})$$

$$= \text{Const.}\ \exp\{-\tfrac{1}{2}\sum_{k=0}^{M}[\underline{z}(k) - H\underline{x}(k)]^T Q^{-1}[\underline{z}(k) - H\underline{x}(k)]\} \tag{2.30}$$

where $\underline{x}(k)$ is the solution to (2.29) for a given $\underline{\theta}\in\Omega$. The CML estimation method is then to find $\underline{\hat{\theta}}_n$ as an estimate of $\underline{\theta}_0$ such that

$$p(\underline{z}(0),\cdots,\underline{z}(M);\underline{\hat{\theta}}_M) = \max_{\underline{\theta}\in\Omega} p(\underline{z}(0),\cdots,\underline{z}(M);\underline{\theta})$$

$$M = 0,1,2,\cdots \tag{2.31}$$

If there are $\underline{\theta}_1,\underline{\theta}_2\in\Omega$, $\underline{\theta}_1 \neq \underline{\theta}_2$, that will both generate the same $H\underline{x}(k)$ sequence when applied with a given input sequence, then

$$\prod_{k=0}^{M} p_{v(k)}(\underline{z}(k) - H\underline{x}(k);\underline{\theta}_1) = \prod_{k=0}^{M} p_{v(k)}(\underline{z}(k) - H\underline{x}(k);\underline{\theta}_2)$$

$$M = 0,1,2,\cdots \tag{2.32}$$

for the same measurement noise distribution. Thus by definitions 2.14 and 2.16, $\underline{\theta}_1$ and $\underline{\theta}_2$ are both unresolvable and CML unresolvable. Since system (2.29) is minimal, the following theorem, which provides the sufficient condition for unresolvability, follows immediately.

Theorem 2.8

Let $\underline{\theta}_1 = \{\underline{x}_{01}, F_1, G_1, H_1, Q_1\}$ and

$\underline{\theta}_2 = \{\underline{x}_{02}, F_2, G_2, H_2, Q_2\}$, $\underline{\theta}_1,\underline{\theta}_2\in\Omega$, $\underline{\theta}_1 \neq \underline{\theta}_2$.

$\underline{\theta}_1$ and $\underline{\theta}_2$ are both unresolvable and CML unresolvable if there exists a nonsingular, nonidentity matrix P such that

$$H_1 = H_2 P^{-1}$$

$$F_1 = P F_2 P^{-1}$$

$$G_1 = P G_2$$

$$\underline{x}_{10} = P \underline{x}_{20}$$

$$Q_1 = Q_2$$

Proof: Let the state vectors generated by $\underline{\theta}_1$ and $\underline{\theta}_2$ be $\underline{x}_1(k)$ and $\underline{x}_2(k)$ respectively. Then $\underline{x}_1(k) = P\underline{x}_2(k)$ $H_1\underline{x}_1(k) = H_2P^{-1}P\underline{x}_2(k) = H_2\underline{x}_2(k)$ for any input, and the two systems parametrized by $\underline{\theta}_1$ and $\underline{\theta}_2$ respectively are equivalent.

Therefore

$$\prod_{k=0}^{M} p_{v(k)}(\underline{z}(k) - H_1\underline{x}_1(k); \underline{\theta}_1)$$

$$= \text{Const.} \cdot \exp\{-\tfrac{1}{2} \sum_{k=0}^{M} [z(k) - H_1\underline{x}_1(k)]^T Q_1^{-1} [z(k) - H_1\underline{x}_1(k)]\}$$

$$= \text{Const.} \cdot \exp\{-\tfrac{1}{2} \sum_{k=0}^{M} [\underline{z}(k) - H_2\underline{x}_2(k)]^T Q_2^{-1} [\underline{z}(k) - H_2\underline{x}_2(k)]\}$$

$$= \prod_{k=0}^{M} p_{v(k)}(\underline{z}(k) - H_2\underline{x}_2(k); \underline{\theta}_2), \quad M = 0,1,2,\text{---}$$

Hence $\underline{\theta}_1$ and $\underline{\theta}_2$ are both unresolvable and CML unresolvable. The above theorem is a modification of the one given by Tse and Anton in [19].

To illustrate Theorem 2.8, we have the following example. Consider the system

$$\underline{x}(k+1) = \begin{bmatrix} \theta_3 & 1.0 \\ 0 & \theta_4 \end{bmatrix} \underline{x}(k) + \begin{bmatrix} 0 \\ \theta_5 \end{bmatrix} u(k) \qquad (2.34)$$

$$\underline{z}(k) = [\,1 \quad 0\,]\, \underline{x}(k) + v(k)$$

$$x(0) = \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}$$

with $\underline{\theta} = [\theta_1 \ \theta_2 \ \theta_3 \ \theta_4 \ \theta_5]^T$ to be identified.

By Theorem 2.8, the following equations are obtained

$$\begin{bmatrix} \theta_3' & 1.0 \\ 0 & \theta'_4 \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} \theta_3 & 1.0 \\ 0 & \theta_4 \end{bmatrix} \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}^{-1}$$

$$\begin{bmatrix} 0 \\ \theta_5' \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} 0 \\ \theta_5 \end{bmatrix}$$

$$[1 \quad 0] = [1 \quad 0] \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}^{-1} \tag{2.35}$$

$$\begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}$$

and the solutions are

$$P_{11} = 1.0$$

$$P_{12} = 0$$

$$P_{21} = \theta_3 - \theta_4$$

$$P_{22} = 1.0$$

$$\theta_3' + \theta_4' = \theta_3 + \theta_4$$

$$\theta_5' = \theta_5$$

$$\theta_1' = \theta_1$$

$$\theta_2' = \theta_2 + P_{21} \ \theta_1$$

Specifically, if we let

$$\underline{\theta}_1 = [0.5,\ 0.1,\ 0.3,\ 0.7,\ 1.0], \text{ then}$$

$$\underline{\theta}_2 = [0.5,\ -0.1,\ 0.7,\ 0.3,\ 1.0]$$

and

$$P = \begin{bmatrix} 1.0 & 0 \\ -0.4 & 1.0 \end{bmatrix}.$$

I.e., the following two systems

$$\underline{x}(k+1) = \begin{bmatrix} 0.3 & 1.0 \\ 0 & 0.7 \end{bmatrix} \underline{x}(k) + \begin{bmatrix} 0 \\ 1.0 \end{bmatrix} u(k) \tag{2.36}$$

$$\underline{z}(k) = [\ 1 \quad 0\ ]\ \underline{x}(k) + v(k)$$

$$\underline{x}(0) = [0.5\ ,\ 0.1]^T$$

and

$$\underline{x}(k+1) = \begin{bmatrix} 0.7 & 1.0 \\ 0 & 0.3 \end{bmatrix} \underline{x}(k) + \begin{bmatrix} 0 \\ 1.0 \end{bmatrix} u(k) \tag{2.37}$$

$$z(k) = [\ 1 \quad 0\ ]\ x(k) + v(k)$$

$$\underline{x}(0) = [0.5\ ,\ -0.1]^T$$

are unresolved in any compact subset of R^5 containing $\underline{\theta}_1$ and $\underline{\theta}_2$.

If there is more than one vector of parameters in Ω that will generate the same observation sequence joint density function, the parameters are not globally identifiable. However, if there exist regions around each point and if there exists a local estimation sequence in each region, we

are still able to identify those parameters by some identification algorithms of local variation type. The following concept is established by Tse in [18].

Definition 2.18

The parameter $\underline{\theta}_0 \in \Omega$ is said to be locally identifiable if

(i) there exists an open set S_0 such that θ_0 is an interior point of S_0, and

(ii) there exists a consistent estimate

$$\{\underline{\hat{\theta}}_n\}_{n=1}^{\infty} \text{ in } \bar{S}_0 \text{ where } \bar{S}_0 \text{ is the closure of } S_0.$$

We will call S_0 the region of parameter identifiability. By imposing the same assumptions as the above on the joint density function of the observation sequence, Tse [18] has the following theorem.

Theorem 2.9

If for all $n = 1,2,---$, there exists a $\lambda^2 > 0$ such that

$$J_{n,n}(\underline{\theta}_0) \triangleq E_{\underline{\theta}_0}\{[\frac{\partial \log p(z_n|Z_{n-1};\underline{\theta}_0)}{\partial \underline{\theta}_0}][\frac{\partial \log p(z_n|Z_{n-1};\underline{\theta}_0)}{\partial \underline{\theta}}]^T\}$$

$$> \lambda^2 I_{p\times p} \tag{2.38}$$

where $E_{\underline{\theta}_0}$ represents the expectation with respect to the density function $p(Z_n;\underline{\theta}_0)$, then $\underline{\theta}_0$ is locally identifiable.

Another weaker sufficient condition for local identifiability was also established. Define

$$J_{k,j}(\underline{\theta}) \triangleq E_{\underline{\theta}}\{[\frac{\partial \log p_{k,j}(\underline{\theta})}{\partial \underline{\theta}}][\frac{\partial \log p_{k,j}(\underline{\theta})}{\partial \underline{\theta}}]^T\} \tag{2.39}$$

where

$$p_{k,j}(\underline{\theta}) \equiv p(z_k, z_{k+1}, \cdots, z_j | z_{k-1}; \underline{\theta}) \tag{2.40}$$

Noting that

$$-E_{\underline{\theta}}\{\frac{\partial^2 \log p_{k,j}(\underline{\theta})}{\partial \underline{\theta}^2}\} = -E_{\underline{\theta}}\{[\frac{\partial \log p_{k,j}(\underline{\theta})}{\partial \underline{\theta}}][\frac{\partial \log p_{k,j}(\underline{\theta})}{\partial \underline{\theta}}]^T\} \tag{2.41}$$

and

$$\begin{aligned} E_{\underline{\theta}}\{\frac{\partial^2 \log p_{k,j}(\underline{\theta})}{\partial \underline{\theta}^2}\} &= E_{\underline{\theta}}\{\sum_{j=k}^{j} \frac{\partial^2 \log p_{i,i}(\underline{\theta})}{\partial \underline{\theta}^2}\} \\ &= \sum_{i=k}^{j} E_{\underline{\theta}}\{\frac{\partial^2 \log p_{i,i}(\underline{\theta})}{\partial \underline{\theta}^2}\} \\ &= -\sum_{i=k}^{j} J_{i,i}(\underline{\theta}) \end{aligned} \tag{2.42}$$

We have

$$J_{k,j}(\underline{\theta}) = \sum_{i=k}^{j} J_{i,i}(\underline{\theta}) \tag{2.43}$$

and

$$J_{1,n}(\underline{\theta}_0) = \sum_{i=1}^{n} J_{i,i}(\underline{\theta}_0) \tag{2.44}$$

Theorem 2.10

If there exists a $\lambda^2 > 0$ such that

$$\lim_{n\to\infty} \sum_{i=1}^{n} J_{i,i}(\underline{\theta}_0) = \lim_{n\to\infty} J_{1,n}(\underline{\theta}_0) \geq \lambda^2 I \tag{2.45}$$

then $\underline{\theta}_0$ is locally identifiable.

Definition 2.19

A subset $S \subset R^p$ is said to be locally identifiable if all the elements in S are locally identifiable.

Theorem 2.11

A sufficient condition for a subset $S \subset R^p$ to be locally identifiable is that

$$\lim_{n \to \infty} J_{1,n}(\underline{\theta}) \geq \lambda^2(\underline{\theta}) I_{p \times p};\ \lambda^2(\underline{\theta}) > 0 \text{ for all } \underline{\theta} \in S \tag{2.46}$$

In the next section we will present a system identification example given by Aoki and Yue [1].

2.4 An Example of System Identification

Aoki and Yue [1] examined the asymptotic properties of the constrained maximum likelihood estimate of the unknown parameters of a class of linear, stable, constant, discrete-time systems with observation and plant noise. The system considered by them is in the completely observable companion form and is single-input and single-output, hence the system representation is unique. Therefore local identifiability will imply global identifiability for the class of systems considered by them. It is obvious that global identifiability implies local identifiability.

Consider the dynamic system represented by

$$\begin{aligned} \underline{x}(k+1) &= F\underline{x}(k) + Gu(k) \\ y(k) &= H\underline{x}(k) \\ z(k) &= H\underline{x}(k) + v(k) \end{aligned} \tag{2.47}$$

where F is an $(n \times n)$ matrix, G is an n-component vector, $H = [1\,0 \cdots 0]$ is a $(1 \times n)$ matrix, and v(k) is a sequence of independent and identically distributed random variables with zero mean and finite variance σ^2, i.e., $v(k) \sim N(0,\sigma^2)$, $k = 0,1,2,\cdots$. F has the following completely controllable companion form:

$$F = \begin{bmatrix} -a_1 & 1 & 0 & \cdots & 0 \\ -a_2 & 0 & 1 & 0 \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -a_n & 0 & \cdots & \cdots & 0 \end{bmatrix}$$

and $G = [b_1, b_2, \cdots, b_n]^T$

The initial state $\underline{x}(0)$ is an unknown n-vector. Our purpose is to identify

$$\underline{\theta} = [a_1, \cdots, a_n, b_1, \cdots, b_n]^T, \text{ a } 2n\text{-vector},$$

and the unknown initial state $\underline{x}(0) \equiv \underline{x}_0$.

The input sequence u(k), $k = 0,1,2,\cdots$ is known and is assumed to be uniformly bounded.

Suppose we take M observations. Define

$$\begin{aligned} \underline{U}_M &= [u(0), u(1), \cdots, u(M-1)]^T \\ \underline{V}_M &= [v(0), v(1), \cdots, v(M-1)]^T \\ \underline{Y}_M &= [y(0), y(1), \cdots, y(M-1)]^T \\ \underline{Z}_M &= [z(0), z(1), \cdots, z(M-1)]^T \end{aligned} \tag{2.48}$$

Then

$$A_M \underline{Y}_M = B_M \underline{U}_M + E_M \underline{x}(0) \tag{2.49}$$

where

$$A_M = I_{M\times M} + \sum_{i=1}^{n} a_i S^i$$

$$B_M = \sum_{i=1}^{n} b_i S^i$$

$$E_M = \begin{pmatrix} I_{n\times n} \\ - - - - \\ 0_{M-n,n} \end{pmatrix} \tag{2.50}$$

and S is the M×M shift matrix with element $S_{ij} = \delta_{i,j+1}$. Another way to express the input-output relation of (2.47) is as follows:

$$\underline{Y}_M = \underline{H}_M \underline{\theta} + E_M \underline{x}(0) \tag{2.51}$$

where

$$\underline{H}_M = (-S\underline{Y}_M, -S^2\underline{Y}_M, \cdots, -S^n\underline{Y}_M, S\underline{U}_M, S^2\underline{U}_M, \cdots, S^n\underline{U}_M) \tag{2.52}$$

which is an M x 2n matrix.

We assume that the true parameter $\underline{\theta}_0$ is an interior point of Ω where Ω is a compact subset of R^{2n}. Furthermore, we assume that the system (2.47) is stable for every $\theta \in \Omega$. By (2.47) and (2.49), the output sequence Z_M can be expressed as:

$$\underline{Z}_M = \underline{V}_M + \underline{Y}_M = \underline{V}_M + A_M^{-1}(B_M \underline{U}_M + E_M \underline{x}_0) \tag{2.53}$$

and the joint probability density function of $\underline{Z}_M$ is

$$p(Z_M; \underline{\theta}, \underline{x}_0) = \text{Const.} \cdot \exp\left(\frac{-1}{2\sigma^2} \|\underline{Z}_M - A_M^{-1}(B_M \underline{U}_M + E_M \underline{x}_0)\|^2\right) \tag{2.54}$$

The constrained likelihood estimate of $\underline{\theta}$ and $\underline{x}_0$, denoted by $\hat{\underline{\theta}}_M$ and $\hat{\underline{x}}_{0M}$ respectively, are obtained by

$$\log p(\underline{Z}_M; \hat{\underline{\theta}}_M, \hat{\underline{x}}_{0M}) = \max_{\underline{\theta} \in \Omega, \underline{x}_0 \in R^n} \log p(\underline{Z}_M; \underline{\theta}, \underline{x}_0) \tag{2.55}$$

For any $\underline{\theta} \in \Omega$, $\max_{\underline{x}_0 \in \Omega R^n} \log p(\underline{Z}_M; \underline{\theta}, \underline{x}_0)$ is obtained by

$$\tilde{\underline{x}}_{0M}(\underline{\theta}) = [E_M^T (A_M^T)^{-1} A_M^{-1} E_M]^{-1} E_M^T (A_M^T)^{-1} [\underline{Z}_M - A_M^{-1} B_M \underline{U}_M] \tag{2.56}$$

Then $\hat{\underline{\theta}}_M$ is obtained by

$$\begin{aligned} &\min_{\underline{\theta} \in \Omega} \|\underline{Z}_M - A_M^{-1}(B_M \underline{U}_M + E_M \tilde{\underline{x}}_{0M}(\underline{\theta}))\|^2 \\ &= \|\underline{Z}_M - A_M^{-1}(B_M \underline{U}_M + E_M \tilde{\underline{x}}_{0M}(\hat{\underline{\theta}}_M)\|^2 \end{aligned} \tag{2.57}$$

and

$$\hat{\underline{x}}_{0M} = \tilde{\underline{x}}_{0M}(\hat{\underline{\theta}}_M) \tag{2.58}$$

The following theorems were given by Aoki and Yue in [1].

Theorem 2.12

If the system (2.47) is completely controllable, and G is not a zero vector, then the constrained maximum likelihood estimate $\hat{\underline{\theta}}$ converges to the true parameter $\underline{\theta}_0$ with probability one if and only if

$$\lim_{M \to \infty} \frac{1}{M} \bar{\underline{U}}_{M,2n}^T \bar{\underline{U}}_{M,2n} > 0$$

where

$$\bar{\underline{U}}_{M,2n} = (S\underline{U}_M, S^2 \underline{U}_M, \cdots, S^{2n} \underline{U}_M) \tag{2.59}$$

Theorem 2.13

If the system (2.47) is completely controllable, and G is not a zero vector, then the constrained maximum likelihood estimate $\hat{\underline{\theta}}_M$ converges to the true parameter $\underline{\theta}_0$ if and only if

$$\lim_{M\to\infty} \frac{1}{M} \underline{H}_M \underline{H}_M^T > 0 \quad \text{for all } \underline{\theta}\in\Omega.$$

We note that the positive definiteness of the two matrices in Theorem 12 and 13 are sensitive to the input sequence $\underline{U}_M$; therefore, input synthesis is an important factor for the identifiability of the system parameters.

3. LOCAL LEAST-SQUARE PARAMETER IDENTIFIABILITY AND LOCAL PARAMETER IDENTIFIABILITY FROM THE TRANSFER FUNCTION OF LINEAR DYNAMIC SYSTEMS

In this chapter we will discuss specifically the parameter identifiability of linear dynamic systems. We will establish the concept of parameter identifiability without considering the identification algorithm, the algorithm-oriented least-square identifiability [2], and the transfer-function identifiability for linear dynamic systems [6]. Furthermore we will discuss the relation between the local least-square identifiability and the local transfer-function identifiability.

3.1 Notation

We will present specifically the manipulation of the bilinear operators in this section which is necessary for the approach later on.

For $\psi(\underline{\theta})$ a real-valued function of $\underline{\theta}\in\Omega\subset R^p$, i.e., $\psi:\Omega\to R$, if $\psi(\theta)$ is differentiable with respect to $\underline{\theta}\in\Omega$, we define

$$\psi'(\theta) \triangleq [\frac{\partial\psi(\underline{\theta})}{\partial\theta_1} ,---, \frac{\partial\psi(\underline{\theta})}{\partial\theta_p}]$$

where

$$\underline{\theta} = [\theta_1,---,\theta_p]^T.$$

If $\underline{\psi}(\underline{\theta})$ is an m-component vector-valued function of $\underline{\theta}\in\Omega$, we define

$$\underline{\psi}'(\underline{\theta}) = \begin{bmatrix} \frac{\partial \psi_1(\underline{\theta})}{\partial \theta_1} & \cdots & \frac{\partial \psi_1(\underline{\theta})}{\partial \theta_p} \\ \vdots & & \\ \frac{\partial \psi_m(\underline{\theta})}{\partial \theta_1} & \cdots & \frac{\partial \psi_m(\underline{\theta})}{\partial \theta_p} \end{bmatrix}$$

$$= \left[\frac{\partial \psi_i(\underline{\theta})}{\partial \theta_j} \right]$$

and

$$\underline{\psi}''(\underline{\theta}) = \left[\frac{\partial^2 \psi_i(\underline{\theta})}{\partial \theta_j \partial \theta_k} \right] \qquad \begin{matrix} i = 1,\cdots,m \\ j = 1,\cdots,p \\ k = 1,\cdots,p \end{matrix}$$

which is an $(m \times p \times p)$ bilinear operator [14]. For $A(\underline{\theta})$ an $(m \times p)$ matrix, $A(\underline{\theta}) = [A_{ij}(\underline{\theta})]$, we define

$$A'(\underline{\theta}) = \left[\frac{\partial A_{ij}(\underline{\theta})}{\partial \theta_k} \right] \qquad \begin{matrix} i = 1,\cdots,m \\ j = 1,\cdots,p \\ k = 1,\cdots,p \end{matrix}$$

which is an $(m \times p \times p)$ bilinear operator. If B is an $(m \times p \times p)$ bilinear operator, then the product of B with a p-vector $\underline{\theta}$ is an $(m \times p)$ matrix whose i,j element is

$$(B\underline{\theta})_{ij} = \sum_{k=1}^{p} B_{ijk}\theta_k$$

Moreover, $B\underline{\theta}\underline{\theta}$ is a m-vector defined by

$$B\underline{\theta}\underline{\theta} = (B\underline{\theta})\underline{\theta}$$

$$(B\underline{\theta}\underline{\theta})_i = \sum_{j=1}^{p} \left(\sum_{k=1}^{p} B_{ijk}\theta_k \right) \theta_j$$

We denote the permutation of B as B^*, where

$$(B^*)_{ijk} = (B)_{ikj}$$

B is said to be symmetric if

$$B = B^{*}$$

The transpose of B is denoted by B^{T}, where

$$(B^{T})_{ijk} = (B)_{kij}$$

The product of a bilinear operator B with a $(p \times q)$ matrix is an $(m \times p \times q)$ bilinear whose i,j,k element is

$$(BA)_{ijk} = \sum_{\ell=1}^{p} B_{ij\ell} A_{\ell k}$$

The product of a $(q \times m)$ matrix with an $(m \times p \times p)$ bilinear operator is as $(q \times p \times p)$ bilinear operator whose i,j,k element is

$$(AB)_{ijk} = \sum_{\ell=1}^{m} A_{i\ell} B_{\ell jk}$$

Given two matrices

$$A = [a_{ij}] \ , \quad i,j = 1,---,n \ , \text{ and}$$

$$B = [b_{ij}] \ , \quad i,j = 1,---,m \ ,$$

we form a new matrix C with elements $c_{ij;k\ell}$ obtained by multiplying each element of A by each element of B in the following way:

$$C = [c_{ij;k\ell}] = [a_{ik} b_{j\ell}]$$

Here, the pair of integers (k,j) act as the first index, and the pair of integers (k,ℓ) act as the second index, where

$$i,k = 1,2,---,n,$$

$$j,\ell = 1,2,---,m.$$

The matrix C is called the Kronecker product [15] of the matrices A and B and is denoted by

$$C = A \otimes B$$

For example, let A and B be second order matrices. Then their Kronecker product is a fourth order matrix, which can be written as

$$C = \begin{bmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22} \\ a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{12} \\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{21} \end{bmatrix}$$

$$= \begin{bmatrix} c_{11;11} & c_{11;12} & c_{11;21} & c_{11;22} \\ c_{12;11} & c_{12;12} & c_{12;21} & c_{12;22} \\ c_{21;11} & c_{21;12} & c_{21;21} & c_{21;22} \\ c_{22;11} & c_{22;12} & c_{22;21} & c_{22;22} \end{bmatrix}$$

For an $(n \times m)$ matrix $A = [a_{ij}]$, $\bar{A}$ is defined by

$$\bar{A} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \\ a_{12} \\ \vdots \\ a_{n2} \\ \vdots \\ a_{1n} \\ \vdots \\ a_{nm} \end{bmatrix}$$

3.2 Local Least-Square Parameter Identifiability

A general continuous-time deterministic dynamic system is described by

$$\dot{\underline{x}}(t) = \underline{\psi}(\underline{x}(t), \underline{u}(t), t; \underline{\theta})$$
$$\underline{y}(t) = \underline{\xi}(\underline{x}(t), \underline{u}(t), t; \underline{\theta}) \tag{3.1}$$

where $\underline{x}(t)$ is the n-state vector, $\underline{u}(t)$ is the q-input vector, $\underline{y}(t)$ is the r-output vector, $t\in[0,\infty) = R^+$. $\underline{\psi}:R^n \times R^q \times R^+ \times \Omega \to R^n$, and $\underline{\xi}:R^n \times R^q \times R^+ \times \Omega \to R^r$. $\underline{\theta}\in\Omega\subset R^p$ is the unknown parameter to be identified. If the system is discrete-time, then it is described by

$$\underline{x}(k+1) = \underline{\psi}(\underline{x}(k), \underline{u}(k), k; \underline{\theta})$$
$$\underline{y}(k) = \underline{\xi}(\underline{x}(k), \underline{u}(k), k; \underline{\theta}) \tag{3.2}$$
$$k = 0,1,2,\text{---}$$

We assume that Ω is a compact subset of R^p and the system is stable for all $\underline{\theta}\in\Omega$. We must note here that $\underline{\theta}$ may stand for the system model coefficients only, e.g., the F, G, H matrices of the linear systems, or it may include the unknown initial state. To distinguish the above two cases, we have the following definitions. Let Λ denote the set of all admissible inputs, let $\underline{h}(t;\underline{\theta})$ or $\underline{h}(k;\underline{\theta})$ denotes the output generated by $\underline{\theta}$ when applied with a $\underline{u}\in\Lambda$

Definition 3.1.a

Let $\underline{\theta}$ stand for the system model coefficients only, then for the continuous-time systems, $\underline{\theta}_1, \underline{\theta}_2 \in \Omega$, $\underline{\theta}_1 \neq \underline{\theta}_2$ are said to be unresolvable in Ω if

$$\underline{h}(t;\underline{\theta}_1) = \underline{h}(t;\underline{\theta}_2)$$

for all $\underline{x}_0 \in R^n$, $\underline{u} \in \Lambda$, and $t \in R^+$.

For the discrete-time systems, $\underline{\theta}_1$ and $\underline{\theta}_2$ are said to be unresolvable in Ω if

$$\underline{h}(k;\underline{\theta}_1) = h(k;\underline{\theta}_2)$$

for all $\underline{x}_0 \in R^n$, $\underline{u} \in \Lambda$, and $k \in I^+$.

Definition 3.1.b

If $\underline{\theta}$ includes the unknown initial state $\underline{x}_0$, then for the continuous-time systems, $\underline{\theta}_1, \underline{\theta}_2 \in \Omega$, $\underline{\theta}_1 \neq \underline{\theta}_2$ are said to be unresolvable in Ω if

$$\underline{h}(t;\underline{\theta}_1) = \underline{h}(t;\underline{\theta}_2)$$

for all $\underline{u} \in \Lambda$ and $t \in R^+$.

For the discrete-time systems, $\underline{\theta}_1$ and $\underline{\theta}_2$ are said to be unresolvable if

$$\underline{h}(k;\underline{\theta}_1) = \underline{h}(k;\underline{\theta}_2)$$

for all $\underline{u} \in \Lambda$ and $k \in I^+$.

Definition 3.2

A parameter $\underline{\theta}_0 \in \Omega$ is said to be locally identifiable if there exists an open sphere $S(\underline{\theta}_0, \rho)$ with radius $\rho > 0$ centered at $\underline{\theta}_0$ such that there is no other $\underline{\theta} \in S(\underline{\theta}_0, \rho) \cap \Omega$, $\underline{\theta} \neq \underline{\theta}_0$ which is unresolvable from $\underline{\theta}_0$.

The above definition is made independent of the method for recovering $\underline{\theta}_0$. However, Bellman and Åström [2] established an algorithm-oriented definition which is called the least-square identifiability. Specifically,

they first set a least-square type criterion parametrized by the unknown parameters and sought the unique local (or global) minimum of the criterion. Parameter identifiability was then implied by the uniqueness of the minimum of the criterion.

Consider the criterion given by

$$J_T(\underline{\theta}) = \int_0^T \|\underline{h}(t;\underline{\theta}) - \underline{y}(t)\|^2 dt, \quad T>0 \tag{3.3}$$

for the continuous-time system, or

$$J_M(\theta) = \sum_{k=0}^{M} \|\underline{h}(k;\underline{\theta}) - \underline{y}(k)\|^2, \quad M = 0,1,--- \tag{3.4}$$

for the discrete-time system. $\|\cdot\|$ denotes the norm, $\underline{y}(t)$ and $\underline{y}(k)$ are the measured outputs of the continuous-time systems and the discrete-time systems, respectively. The following definition was given by Bellman and Åström in [2].

Definition 3.3

Let $\underline{\theta}_0$ be the true parameter of a control system parametrized by the unknown parameter $\underline{\theta}$. Then $\underline{\theta}_0$ is said to be locally least-square identifiable if the criterion $J_T(\underline{\theta})$ or $J_M(\underline{\theta})$ has a local minimum at $\underline{\theta} = \underline{\theta}_0$. If the minimum is global, $\underline{\theta}_0$ is said to be globally identifiable.

To establish the sufficient conditions for $\underline{\theta}_0$ to be (locally) identifiable, we proceed as follows. We consider the discrete-time system. To find the local minimum of $J_M(\theta)$, we take the derivative of $J_M(\underline{\theta})$ with respect to $\underline{\theta}$ if $J_M(\underline{\theta})$ is a continuously differentiable function of θ. Since $J_M(\theta)$ is a scalar, $J_M'(\underline{\theta})$ is a p-component vector function and $J_M''(\underline{\theta})$ is a $(p \times p)$ matrix function. If $\underline{\theta}_0$ is the true parameter, then $\underline{h}(k;\underline{\theta}_0) = \underline{y}(k)$,

$k = 0,1,\cdots$, and $J_M(\underline{\theta}_0) = 0$, $M = 0,1,2,\cdots$. Since $J_M(\underline{\theta})$ is nonnegative, $\underline{\theta}_0$ is a minimal point of $J_M(\underline{\theta}_0)$ and hence $J_M'(\underline{\theta}_0) = 0$. However, if $\underline{\theta}_0$ is to be locally identifiable, $\underline{\theta}_0$ must be the unique minimal point for $J_M(\underline{\theta})$ in some neighborhood of $\underline{\theta}_0$, i.e., there must exist an open sphere $S(\underline{\theta}_0,\rho) \subset \Omega$ with radius $\rho > 0$ centered at θ_0 such that if $\theta \in S(\theta_0,\rho)$, $\theta \neq \theta_0$, then $J_M'(\underline{\theta}) \neq 0$ and hence $\underline{\theta}$ is not a minimal point of $J_M(\underline{\theta})$. If we can establish a sufficient condition such that $J_M'(\underline{\theta})$ is an injective function (a one-to-one mapping) in some neighborhood of $\underline{\theta}_0$, then this condition will imply that $\underline{\theta}_0$ is locally identifiable. We first state a result given by Narasimhan in [10].

Lemma 3.1

Let S be an open set in R^p and $\psi: S \to R^\ell$ be a C^k mapping (a k-times continuously differentiable function) with $k \geq 1$. Then if $\psi'(\theta)$ has constant rank j in a neighborhood of $\theta_0 \in S$, ψ is locally injective at θ_0 if and only if $j = p$.

Theorem 3.2

A sufficient condition for $\underline{\theta}_0$ to be locally identifiable is that there exists an open sphere $S(\underline{\theta}_0,\rho) \subset \Omega$ with radius $\rho > 0$ centered at $\underline{\theta}_0$ such that the $(p \times p)$ matrix $J_M''(\underline{\theta})$ is nonsingular for all $\theta \in S(\theta_0,\rho)$.

Proof: Since $J_M''(\underline{\theta})$ is nonsingular for all $\underline{\theta} \in S(\underline{\theta}_0,\rho)$, it has constant rank p for all $\underline{\theta} \in S(\underline{\theta}_0,\rho)$. By Lemma 3.1 $J_M'(\underline{\theta})$ is locally injective on $S(\underline{\theta}_0,\rho)$. Let $\underline{\theta} \in S(\underline{\theta}_0,\rho)$, $\underline{\theta} \neq \underline{\theta}_0$, then $J_M'(\underline{\theta}) \neq 0$ hence $J_M(\underline{\theta}) \neq 0$, $M = 0,1,2,.$ Thus $\underline{\theta}_0$ is the unique minimal point for $J_M(\underline{\theta})$, $M = 0,1,2,\cdots$, in $S(\underline{\theta}_0,\rho)$.

The above theorem is an immediate result of Lemma 3.1. The least-square identifiability and the identification algorithm for finding the

region of parameter identifiability will be studied extensively in Chapter 4.

3.3 Local Parameter Identifiability from the Transfer Function

In this section the systems considered are discrete-time. We will first briefly introduce the realization theory established by Ho and Kalman in [8] and explain the distinction between the realization and identification. Specifically, the parameter identification from the transfer function can be viewed as the realization from the transfer function restricted to the given parametrization.

The quadruplet {F, G, H, D} defines the internal description of a system, which we shall denote by Σ, via the equations:

$$\begin{aligned} \underline{x}(k+1) &= F\underline{x}(k) + G\underline{u}(k) \\ \underline{y}(k) &= H\underline{x}(k) + D\underline{u}(k) \end{aligned} \tag{3.6}$$

where F is the $n \times n$ state transition matrix

G is the $n \times q$ input matrix

H is the $r \times n$ output matrix

D is the $r \times q$ direct-coupling matrix

$\underline{x}(k)$ is the n-state vector

$\underline{y}(k)$ is the r-output vector

$\underline{u}(k)$ is the q-input vector

The external description of the system Σ is the zero-state impulsive response description, namely the description in terms of an impulse input and the corresponding output. There are two ways to represent the external

description of Σ. One of them is the time domain description. From equation (3.6), it can be easily seen that the impulse response of Σ is given by

$$\begin{aligned} W_0 &= D \\ W_1 &= HG + D \\ &\vdots \\ W_M &= \left\{\sum_{i=1}^{M-1} HF^iG\right\} + D \\ &\vdots \end{aligned} \quad , \tag{3.7}$$

and the impulse response matrix is given by

$$\bar{W} = \begin{bmatrix} W_0 \\ W_1 \\ \vdots \\ W_M \\ \vdots \end{bmatrix} = \begin{bmatrix} HG^D + D \\ HFG + HG + D \\ \vdots \\ \left\{\sum_{i=1}^{M-1} HF^iG\right\} + D \\ \vdots \end{bmatrix} . \tag{3.8}$$

By the frequency domain description, the input is related to the output by the transfer function T(z) such that

$$Y(z) = T(z)U(z) \tag{3.9}$$

where

$$T(z) = \left\{H(zI - F)^{-1}G + D\right\} \tag{3.10}$$

$z \in C$ (the field of complex numbers).

Through the concept of realization, the external description of a system can be related to the internal description of a system. The problem of realization can thus be stated as follows [8]: construct $\{F,G,H,D\}$ such that the identity (3.7) holds, i.e., given a sequence of constant $(r \times q)$ matrices $\{M_i\}_{i=0}^{\infty}$, find a quadruplet $\{F,G,H,D\}$ of constant matrices such that

$$M_0 = D$$

$$M_1 = HG$$

$$\vdots$$

$$M_M = HF^{M-1}G$$

$$\vdots$$

The sequence $\{M_i\}_{i=0}^{\infty}$ is called the Markov parameters of the system Σ. The dimension of Σ is defined by

$$\dim(\Sigma) \triangleq \dim(F)$$

We say that the realization $\{F,G,H,D\}$ is minimal if the dimension of F is less than or equal to the dimension of any other realization of Σ. From the linear system theory, we know that a realization is minimal if and only if Σ is both completely controllable and completely observable, i.e., if and only if

$$\text{rank}[G, FG, \cdots, F^{n-1}G] = n \quad \text{(completely controllable)}$$

$$\text{and} \quad \text{rank}[H^T, (HF)^T, \cdots, (HF^{n-1})^T] = n \quad \text{(completely observable)}$$

Furthermore, given an external description, two minimal realizations $\Sigma_1 = \{F_1, G_1, H_1, D_1\}$ and $\Sigma_2 = \{F_2, G_2, H_2, D_2\}$ are equivalent if and only if there exists a nonsingular $n \times n$ matrix P such that

$$\begin{aligned} F_2 &= PF_1P^{-1} \\ G_2 &= PG_1 \\ H_2 &= H_1P^{-1} \\ D_2 &= D_1 \end{aligned} \tag{3.9}$$

These two equivalent internal descriptions differ only in the coordination of their state spaces.

Even though we have the knowledge of the external description and the minimal dimension of Σ, generally we are not able to determine the quadruplet $\{F,G,H,D\}$ uniquely unless the structure of $\{F,G,H,D\}$ is constrained in some specified form. We can illustrate the above statement by the following example [2].

Consider

$$\begin{aligned} x(k+1) &= \begin{bmatrix} -(\theta_1+\theta_2) & \theta_3 \\ \theta_2 & -(\theta_3+\theta_4) \end{bmatrix} x(k) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u(k) \\ y(k) &= \begin{bmatrix} 1 & 0 \end{bmatrix} x(k) \end{aligned} \tag{3.10}$$

The transfer function of (3.10) is

$$T(z) = \frac{z + \theta_3 + \theta_4}{z^2 + z(\theta_1+\theta_2+\theta_3+\theta_4) + (\theta_1+\theta_2)(\theta_3+\theta_4) - \theta_2\theta_3} \tag{3.11}$$

Thus we have only three equations to solve the four unknowns $\{\theta_1,\theta_2,\theta_3,\theta_4\}$:

$$\theta_3+\theta_4 = a_1$$
$$\theta_1+\theta_2+\theta_3+\theta_4 = a_2$$
$$(\theta_1+\theta_2)(\theta_3+\theta_4) - \theta_2\theta_3 = a_3 \tag{3.12}$$

where a_1, a_2, and a_3 are known constants. For this under determined set of equations, one of the four unknowns has to be dependent on the other three, hence the representation of (3.10) is not unique. However, if the system representation is constrained to have the canonical form:

$$\underline{x}(k+1) = \begin{bmatrix} 0 & 1 \\ \theta_1 & \theta_2 \end{bmatrix} \underline{x}(k) + \begin{bmatrix} \theta_3 \\ \theta_4 \end{bmatrix} u(k)$$
$$y(k) = \begin{bmatrix} 1 & 0 \end{bmatrix} \underline{x}(k) \tag{3.13}$$

then

$$T(z) = \frac{\theta_3 z + \theta_4 - \theta_2\theta_3}{z^2 - \theta_2 z - \theta_1} = \frac{a_3 z + a_4}{z^2 + a_2 z + a_1}, \tag{3.14}$$

and

$$\theta_1 = a_1$$
$$\theta_2 = a_2$$
$$\theta_3 = a_3$$
$$\theta_4 = a_4 - a_2 a_3 \tag{3.15}$$

is the unique representation for system (3.13). The identification

(realization) of system (3.13) is an example of the canonical parameter identification problem which has been thoroughly studied. By canonical parameter identification, we mean finding a certain set of parameter in a given canonical parametrization which when applied with the input sequence from a given set of input-output sequence will generate a set of output data which will match the given output sequence within some well defined degree of accuracy. The primary reason for finding canonical parameters is to obtain a model which gives a good match to the measured input-output data. However, it may not be desirable to use a canonical form for a given physical system. That is, it may be desirable to identify specified parameters in a given parametrization. Hence the parameters identified in a specified canonical form may have little or no recognizable relationship to desired physical parameters. To expound the above statement, we give the following example.

Given the frequency domain external description of a second order zero-state system Σ:

$$T(z) = \frac{c}{z^2 + az + b} \tag{3.16}$$

The canonical parametrization:

$$\underline{x}(k+1) = \begin{bmatrix} 0 & 1 \\ \theta_1 & \theta_2 \end{bmatrix} \underline{x}(k) + \begin{bmatrix} 0 \\ \theta_3 \end{bmatrix} u(k) \tag{3.17}$$

$$y(k) = \begin{bmatrix} 1 & 0 \end{bmatrix} \underline{x}(k)$$

can be uniquely determined by

$$\theta_1 = -b$$
$$\theta_2 = -a$$
$$\theta_3 = c \tag{3.18}$$

hence the canonical parametrization is globally identifiable. However, if we are given the physical parametrization which is the same as the example in section 2.3:

$$\underline{x}(k+1) = \begin{bmatrix} \theta_1 & 1.0 \\ 0.0 & \theta_2 \end{bmatrix} \underline{x}(k) + \begin{bmatrix} 0 \\ \theta_3 \end{bmatrix} u(k) \tag{3.19}$$

$$y(k) = [\ 1 \quad 0]\ \underline{x}(k)$$

then we get the following set of equations:

$$\theta_1 + \theta_2 = -a$$
$$\theta_1 \theta_2 = b$$
$$\theta_3 = c \tag{3.20}$$

It is obvious that θ_1 and θ_2 can be interchanged without affecting the transfer function. Hence this physical parametrization is not globally identifiable but only locally identifiable since the two parameters:

$$\underline{\theta}_1 = (\theta_1, \theta_2, \theta_3) \text{ and } \underline{\theta}_2 = (\theta_2, \theta_1, \theta_3)$$

will both generate the same transfer function even though they are isolated in the parameter space if $\underline{\theta}_1 \neq \underline{\theta}_2$.

Definition 3.4

A parametrization of the system matrices $\{F,G,H,D\}$ is a continuously differentiable function which maps Ω, the parameter space, to the $R^{n(n+q+r)+rq}$ space; i.e., a parametrization is a $c^{\prime}$ function

$$(F,G,H,D)(\underline{\theta}):\Omega\subset R^{p}\rightarrow R^{n(n+q+r)+rq}\ .$$

For brevity, we let $n(n+q+r)+rq=\ell$. The above definition was given by Glover and Willems in [6].

Definition 3.5

Given a parametrization of a system Σ, two parameters $\underline{\theta}_1,\underline{\theta}_2\in\Omega$, $\underline{\theta}_1\neq\underline{\theta}_2$ are said to be unresolvable from the transfer function if

$$H(\underline{\theta}_1)(zI-F(\underline{\theta}_1))^{-1}G(\underline{\theta}_1)+D(\underline{\theta}_1)=H(\underline{\theta}_2)(zI-F(\underline{\theta}_2))^{-1}G(\underline{\theta}_2)+D(\underline{\theta}_2) \tag{3.21}$$

for all $z\in C$ and $z\neq(\lambda(F(\underline{\theta}_1)),\ \lambda(F(\underline{\theta}_2)))$ where $\lambda(\cdot)$ denotes the eigenvalues of the corresponding matrix, or equivalently,

$$\begin{aligned} &D(\underline{\theta}_1)=D(\underline{\theta}_2)\\ &H(\underline{\theta}_1)F^{i}(\underline{\theta}_1)G(\underline{\theta}_1)=H(\underline{\theta}_2)F^{i}(\underline{\theta}_2)G(\underline{\theta}_2)\quad i=0,1,2,\text{---} \end{aligned} \tag{3.22}$$

The following definition is similar to the one given by Glover and Willems in [6].

Definition 3.6

A parametrization is said to be locally identifiable from the transfer function at $\underline{\theta}_0\in\Omega$ if there is an open sphere $S(\underline{\theta}_0,\rho)\subset\Omega$ with radius $\rho>0$ such that there is no $\underline{\theta}\in S(\underline{\theta}_0,\rho)$, $\underline{\theta}\neq\underline{\theta}_0$, which is unresolvable from $\underline{\theta}_0$ i.e., if there is a $\underline{\theta}\in S(\underline{\theta}_0,\rho)$ such that

$$D(\underline{\theta}) = D(\underline{\theta}_0)$$
$$H(\underline{\theta})F^i(\underline{\theta})G(\underline{\theta}) = H(\underline{\theta}_0)F^i(\underline{\theta}_0)G(\underline{\theta}_0), \quad i = 1,2,\cdots \tag{3.23}$$

then $\underline{\theta} = \underline{\theta}_0$.

By Lemma 3.1, an immediate result follows.

Theorem 3.3

Let $(F,G,H,D)(\underline{\theta})$: $\Omega \to R^{\ell}$ be a parametrization of Σ, then the parametrization is locally identifiable from the transfer function at $\theta_0 \in \Omega$ if the gradient of the Markov parameter matrix $\bar{M}(\underline{\theta})$ with respect to $\underline{\theta}$ has constant rank p in an open sphere $S(\underline{\theta}_0,\rho)$ with radius $\rho > 0$ centered at $\underline{\theta}_0$ where

$$\bar{M}(\underline{\theta}) = \begin{bmatrix} D(\underline{\theta}) \\ H(\underline{\theta})G(\underline{\theta}) \\ H(\underline{\theta})F(\underline{\theta})G(\underline{\theta}) \\ \vdots \\ H(\underline{\theta})F^M(\underline{\theta})G(\underline{\theta}) \\ \vdots \end{bmatrix}$$

If a parametrized system is of minimal dimension, then it is related to its zero-state and zero-input equivalent systems by similarity transformations. For the physical parameter identification problem, we are interested in the equivalent systems which have the same parametrization, i.e., we wish to investigate if there is any transformation matrix which will transform a parametrized system to an equivalent system with different parameter values but with same parametrization. Specifically,

we shall investigate the solution $(P,\underline{\theta})$, $P\in GL(n)$, the space of nonsingular $(n \times n)$ matrices, $\underline{\theta}\in\Omega$, of the following set of equations:

$$\begin{aligned} PF(\underline{\theta})P^{-1} &= F(\underline{\theta}_0) \\ PG(\underline{\theta}) &= G(\underline{\theta}_0) \\ H(\underline{\theta})P^{-1} &= H(\underline{\theta}_0) \\ D(\underline{\theta}) &= D(\underline{\theta}_0) \end{aligned} \tag{3.24}$$

where $\underline{\theta}_0$ is the true parameter. It is obvious that if there is an open sphere $S(\underline{\theta}_0,\rho)\subset\Omega$ such that $(I_{n\times n},\underline{\theta}_0)$ is the unique solution of (3.24) in $GL(n)\times S(\underline{\theta}_0,\rho)$, then the parametrization is locally identifiable from the transfer function at $\underline{\theta}_0$. The following theorem provides the sufficient condition for (3.24) to have unique solution locally which was given by Glover and Willems in [6].

Theorem 3.4

Let $\{F,G,H,D\}: \Omega \to R^{\ell}$ be a given parametrization of the system matrices $\{F,G,H,D\}$ and suppose $\{F,G,H,D\}$ is minimal.

Let

$$\mathcal{C}(P,\theta) = \begin{bmatrix} PF(\underline{\theta})P^{-1} \\ PG(\underline{\theta}) \\ H(\underline{\theta})P^{-1} \\ D(\underline{\theta}) \end{bmatrix}$$

If there exists an open sphere $S(\underline{\theta}_0,\rho)$ with radius $\rho>0$ centered at θ_0 such that $\nabla_{(P,\underline{\theta})}\mathcal{C}(P,\underline{\theta})$ has constant rank n^2+P at $P=I$ and for all $\underline{\theta}\in S(\underline{\theta}_0,\rho)$, then the parametrization is locally identifiable from the transfer function at $\underline{\theta}_0$. The matrix $\nabla_{(P,\underline{\theta})}\mathcal{C}(P,\theta)$ evaluated at the point

$(I,\underline{\theta})$ is given by

$$\nabla_{(P,\underline{\theta})}\mathcal{E}(P,\underline{\theta})\Big|_{(I,\underline{\theta})} = \begin{bmatrix} (I_{n\times n} \otimes F^T(\underline{\theta}) - F(\underline{\theta}) \otimes I_{n\times n})_{n^2\times n^2} & \bar{F}'(\underline{\theta})_{n^2\times p} \\ (I_{n\times n} \otimes G^T(\underline{\theta}))_{nq\times n^2} & \bar{G}'(\underline{\theta})_{nq\times p} \\ (-H(\underline{\theta}) \otimes I_{n\times n})_{rn\times n^2} & \bar{H}'(\underline{\theta})_{rn\times p} \\ \bigcirc_{rq\times n^2} & \bar{D}'(\underline{\theta})_{rq\times p} \end{bmatrix} \tag{3.25}$$

which is an $(n^2+nq+rn+rq)$ by (n^2+p) matrix.

We give an example to illustrate the above theorem. Consider the system parametrized by

$$\begin{aligned} F &= \begin{bmatrix} 0 & 1 \\ \theta_1 & \theta_2 \end{bmatrix}, \\ G &= \begin{bmatrix} 0 \\ \theta_3 \end{bmatrix}, \\ H &= [\,1 \quad 0\,], \\ D &= [\,0\,] \\ \underline{\theta} &= [\theta_1 \quad \theta_2 \quad \theta_3]^T, \end{aligned} \tag{3.26}$$

then

$$\nabla_{(P,\underline{\theta})}\mathcal{E}(P,\theta)\Big|_{(I,\underline{\theta})} = \left[\begin{array}{c} \left[\begin{array}{cccc} 0 & \theta_1 & 0 & 0 \\ 1 & \theta_2 & 0 & 0 \\ 0 & 0 & 0 & \theta_1 \\ 0 & 0 & 1 & \theta_2 \end{array}\right] - \left[\begin{array}{cccc} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \theta_1 & 0 & \theta_2 & 0 \\ 0 & \theta_1 & 0 & \theta_2 \end{array}\right] \quad \begin{array}{ccc} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{array} \\ \begin{array}{cccc} 0 & \theta_3 & 0 & 0 \\ 0 & 0 & 0 & \theta_3 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \quad \begin{array}{ccc} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array} \end{array}\right]$$

$$= \begin{bmatrix} 0 & \theta_1 & -1 & 0 & 0 & 0 & 0 \\ 1 & \theta_2 & 0 & -1 & 1 & 0 & 0 \\ -\theta_1 & 0 & -\theta_2 & \theta_1 & 0 & 0 & 0 \\ 0 & -\theta_1 & 1 & 0 & 0 & 1 & 0 \\ 0 & \theta_3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \theta_3 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad (3.27$$

which is of rank $2^2+3=7$ for all $\underline{\theta}\in R^3$ hence the parametrization is globally identifiable from the transfer function.

We will extend the above theorem for the case that $\underline{\theta}$ includes the unknown initial state, i.e., $\underline{\theta} = \begin{bmatrix} \underline{x}_0 \\ \underline{\pi} \end{bmatrix}$, where $\underline{\pi}$ is the unknown system parameter vector contained in $\{F,G,H,D\}$ with a specified parametrization.

Then the augmented parametrization is a mapping from the parameter space Ω to the $R^{n(n+q+r)+rq+n}$ space. Let $\ell' = n(n+q+r)+rq+n$. We note that

$$\underline{x}_0 = [I_{nxn} \vdots 0_{nx(p-n)}]\underline{\theta}$$
$$\text{and } \underline{\eta} = [0_{(p-n)xn} \vdots I_{(p-n)x(p-n)}]\underline{\theta} \tag{3.28}$$

Let $\underline{\theta}_0 = \begin{bmatrix} \underline{x}_{00} \\ \underline{\eta}_0 \end{bmatrix}$ denote the true parameter. The solution $(\underline{P},\underline{\theta})$ of the following set of equations is sought

$$\begin{aligned} P\underline{x}_0 &= \underline{x}_{00} \\ PF(\underline{\eta})P^{-1} &= F(\underline{\eta}_0) \\ PG(\underline{\eta}) &= G(\underline{\eta}_0) \\ H(\underline{\eta})P^{-1} &= H(\underline{\eta}_0) \\ D(\underline{\eta}) &= D(\underline{\eta}_0) \end{aligned} \tag{3.29}$$

The following theorem is established immediately.

Theorem 3.5

Let $\{\underline{x}_0,F,G,H,D\}(\underline{\theta}):\Omega \to R^{\ell'}$ be a given parametrization of a system Σ and suppose $\{F,G,H,D\}$ is minimal. Let

$$\mathcal{E}^*(P,\underline{\theta}) = \begin{bmatrix} P\underline{x}_0 \\ PF(\underline{\eta})P^{-1} \\ PG(\underline{\eta}) \\ H(\underline{\eta})P^{-1} \\ D(\underline{\eta}) \end{bmatrix}$$

If there exists an open sphere $S(\underline{\theta}_0,\rho) \subset \Omega$ with radius $\rho > 0$ centered at $\underline{\theta}_0$ such that $\nabla_{(P,\underline{\theta})}\mathcal{E}^*(P,\underline{\theta})$ has constant rank n^2+p at $P = I_{n\times n}$ and for all $\underline{\theta} \in S(\underline{\theta}_0,\rho)$, then $\underline{\theta}_0$ is locally identifiable (by definition 3.2).

$\nabla_{(P,\underline{\theta})}\mathcal{E}^*(P,\underline{\theta})$ evaluated at $(I_{n\times n},\underline{\theta})$ is given by

$$\nabla_{(P,\underline{\theta})}\mathcal{E}^*(P,\underline{\theta})\Big|_{(I,\underline{\theta})} = \begin{bmatrix} (I_{n\times n} \otimes \underline{x}_0^T)_{n\times n^2} & (\nabla_{\underline{\theta}} \bar{\underline{x}}_0)_{n\times p} \\ (I_{n\times n} \otimes F^T(\eta) - F(\eta) \otimes I_{n\times n})_{n^2\times n^2} & (\nabla_{\underline{\theta}} \bar{F}(\eta))_{n^2\times p} \\ (I_{n\times n} \otimes G^T(\eta))_{nq\times n^2} & (\nabla_{\underline{\theta}} \bar{G}(\eta))_{nq\times p} \\ (-H(\eta) \otimes I_{n\times n})_{rn\times n^2} & (\nabla_{\underline{\theta}} \bar{H}(\eta))_{rn\times p} \\ \bigcirc_{rq\times n^2} & (\nabla_{\underline{\theta}} \bar{D}(\eta))_{rq\times p} \end{bmatrix} \tag{3.30}$$

which is an ℓ' by (n^2+p) matrix. We note here that $\underline{x}_0 = \bar{\underline{x}}_0$ since $\underline{x}_0$ is a column vector and hence $\nabla_{\underline{\theta}} \bar{\underline{x}}_0 = [I_{n\times n} \vdots\ 0_{n\times(p-n)}]$ by (3.28).

Comparing Definition 3.2 and 3.6, and Theorem 3.4 and 3.5, we see that identifiability from the transfer function is equivalent to the zero-state parameter identifiability according to Definition 3.2 which is a more general definition.

3.4 Region of Least-Square Identifiability and Transfer-Function Identifiability

In this section we will discuss the relation between the least-square identifiability and the transfer-function identifiability. We consider a parametrized single-input, zero-state linear system Σ with system matrices $\{F,G,H,D\}$. Let the true parameter be $\underline{\theta}_0$. Its measured impulse response matrix is given by

$$\underline{Y}_M = \begin{bmatrix} \underline{y}(0) \\ \underline{y}(1) \\ \underline{y}(2) \\ \vdots \\ \underline{y}(M) \end{bmatrix} = \begin{bmatrix} D(\underline{\theta}_0) \\ H(\underline{\theta}_0)G(\underline{\theta}_0) + D(\underline{\theta}_0) \\ H(\underline{\theta}_0)F(\underline{\theta}_0)G(\underline{\theta}_0) + H(\underline{\theta}_0)G(\underline{\theta}_0) + D(\underline{\theta}_0) \\ \vdots \\ \{\sum_{i=0}^{M} H(\underline{\theta}_0)F^i(\underline{\theta}_0)G(\underline{\theta}_0)\} + D(\underline{\theta}_0) \end{bmatrix} \tag{3.31}$$

which is an $r(M+1)$-component vector where $r(M+1) \geq p$.

The output sequence generated by the unknown parameter $\underline{\theta}$ when applied with the impulse input is given by

$$\underline{h}_M(\underline{\theta}) = \begin{bmatrix} \underline{h}(0;\underline{\theta}) \\ \underline{h}(1;\underline{\theta}) \\ \underline{h}(2;\underline{\theta}) \\ \vdots \\ \underline{h}(M;\theta) \end{bmatrix} = \begin{bmatrix} D(\underline{\theta}) \\ H(\underline{\theta})G(\underline{\theta}) + D(\underline{\theta}) \\ H(\underline{\theta})F(\underline{\theta})G(\underline{\theta}) + H(\underline{\theta})G(\underline{\theta}) + D(\underline{\theta}) \\ \vdots \\ \{\sum_{i=0}^{M} H(\underline{\theta})F^i(\underline{\theta})G(\underline{\theta})\} + D(\underline{\theta}) \end{bmatrix} \tag{3.31}$$

Let the identification criterion be

$$J_M(\underline{\theta}) \triangleq \sum_{k=0}^{M}[\underline{h}(k;\underline{\theta}) - y(k)]^T[\underline{h}(k;\underline{\theta}) - y(k)]$$
$$= [\underline{h}_M(\underline{\theta}) - \underline{Y}_M]^T[\underline{h}_M(\underline{\theta}) - \underline{Y}_M] \qquad (3.32)$$

Then

$$\underline{\phi}_M(\underline{\theta}) \triangleq 1/2\ J'_M(\underline{\theta}) = [h'_M(\underline{\theta})]^T[\underline{h}_M(\underline{\theta}) - \underline{Y}_M] \qquad (3.33)$$

which is a p-component vector function, and

$$\underline{\phi}'_M(\underline{\theta}) \triangleq [\underline{h}'_M(\underline{\theta})]^T[\underline{h}'_M(\underline{\theta})] + [\underline{h}''_M(\underline{\theta})]^T[\underline{h}_M(\underline{\theta}) - \underline{Y}_M] \qquad (3.34)$$

which is a (pxp) matrix function of $\underline{\theta}$.

We note that

$$J_M(\underline{\theta}_0) = 0\ ,$$
$$\underline{\phi}_M(\underline{\theta}_0) = 0\ ,$$
$$\text{and}\quad \underline{\phi}'_M(\underline{\theta}_0) = [\underline{h}'_M(\underline{\theta}_0)]^T[\underline{h}'_M(\underline{\theta}_0)] \qquad (3.35)$$

Recalling that the Markov parameter matrix is defined by

$$\overline{M}(\underline{\theta}) \triangleq \begin{bmatrix} D(\underline{\theta}) \\ H(\underline{\theta})G(\underline{\theta}) \\ H(\underline{\theta})F(\underline{\theta})G(\underline{\theta}) \\ \vdots \\ H(\underline{\theta})F^{M-1}(\underline{\theta})G(\underline{\theta}) \end{bmatrix}$$

By performing row reduction on $\underline{h}_M(\underline{\theta})$, its easily seen that

$$\text{rank}[\overline{M}(\underline{\theta})] = \text{rank}[\underline{h}_M(\underline{\theta})] \quad \text{for all } \underline{\theta}\in\Omega$$
$$\text{and}\quad \text{rank}[\overline{M}'(\underline{\theta})] = \text{rank}[\underline{h}'_M(\underline{\theta})] \quad \text{for all } \underline{\theta}\in\Omega. \qquad (3.36)$$

Moreover, from the matrix theory we know that

$$\text{rank}[\underline{h}_M'(\underline{\theta})] = p \text{ if and only if } \text{rank}[(\underline{h}_M'(\underline{\theta}))^T(\underline{h}_M'(\underline{\theta}))] = p.$$

For brevity, let

$$\underline{f}_M(\underline{\theta}) \triangleq \underline{h}_M(\underline{\theta}) - \underline{Y}_M \tag{3.37}$$

and $$N_M(\underline{\theta}) \triangleq [\underline{h}_M'(\underline{\theta})]^T[\underline{h}_M'(\underline{\theta})]. \tag{3.38}$$

Our purpose is to find a region $S(\underline{\theta}_0,\rho)$ such that $\underline{\theta}_0$ is both least-square identifiable and identifiable from the transfer function in $S(\underline{\theta}_0,\rho)$. We first state the following Lemma.

Lemma 3.6

Let B and C be two $(n \times n)$ matrices. Let $\|\cdot\|$ be a norm on the space of $(n \times n)$ matrices.

If (i) B is nonsingular,

(ii) $\|B^{-1}\| \leq \alpha$,

(iii) $\|C - B\| \leq \delta$, and

(iv) $\alpha\delta < 1$,

then C is nonsingular and $\|C^{-1}\| \leq \dfrac{\alpha}{1 - \alpha\delta}$

Theorem 3.7

Let Σ be a zero-state, single-input linear system parametrized by the unknown parameter $\underline{\theta}$. Let the true parameter be $\underline{\theta}_0$. Let $\|\cdot\|$ be a norm on the respective spaces. If

(1) $N_M(\underline{\theta}_0)$ is nonsingular,

(2) there exists an open sphere $S(\underline{\theta}_0,\rho)$ with a radius $\rho > 0$ centered at $\underline{\theta}_0$ and a set of positive numbers (λ,γ,μ) such that

(i) $\|N_M^{-1}(\underline{\theta}_0)\| \le \lambda$,

(ii) $\|[\underline{f}_M''(\underline{\theta})]^T \underline{f}_M(\underline{\theta})\| \le \gamma$ for all $\underline{\theta} \in S(\underline{\theta}_0, \rho)$,

(iii) $\|N_M'(\underline{\theta})\| \le \mu$ for all $\underline{\theta} \in S(\underline{\theta}_0, \rho)$,

(iv) $\lambda(\mu\rho + \gamma) < 1$,

then $\underline{\theta}_0$ is both least-square identifiable and identifiable from the transfer function in $S(\underline{\theta}_0, \rho)$.

Proof: By 2(iii) we have

$$\|N_M(\underline{\theta}) - N_M(\underline{\theta}_0)\| \le \rho\mu \quad \text{for all } \underline{\theta} \in S(\underline{\theta}_0, \rho)$$

Moreover, $\|N_M^{-1}(\underline{\theta}_0)\| \le \lambda$ by 2(i). Applying Lemma 3.6 to this situation, we have $\alpha = \lambda$, $\delta = \rho\mu$, and $\lambda\rho\mu < 1$ by 2(iv). Thus $N_M(\underline{\theta})$ is nonsingular and

$$\|N_M^{-1}(\underline{\theta})\| \le \frac{\lambda}{1 - \lambda\mu\rho} \quad \text{for all } \underline{\theta} \in S(\underline{\theta}_0, \rho)$$

Since $N_M(\underline{\theta}) = [\underline{h}_M'(\underline{\theta})]^T[\underline{h}_M'(\underline{\theta})]$ and $N_M(\underline{\theta})$ has constant rank p in $S(\underline{\theta}_0, \rho)$, therefore $\underline{h}_M'(\underline{\theta})$ has constant rank p in $S(\underline{\theta}_0, \rho)$ and $\bar{M}(\underline{\theta})$ has constant rank p in $S(\underline{\theta}_0, \rho)$, hence $\underline{\theta}_0$ is locally identifiable from the transfer function by Theorem 3.3. We now show that $J_M''(\underline{\theta})$ is nonsingular for all $\underline{\theta} \in S(\underline{\theta}_0, \rho)$ and $J_M'(\underline{\theta})$ is locally injective on $S(\underline{\theta}_0, \rho)$.

By (3.34),

$$\underline{\phi}_M'(\underline{\theta}) = N_M(\underline{\theta}) + [\underline{f}_M''(\underline{\theta})]^T \underline{f}_M(\underline{\theta}),$$

therefore

$$\|N_M(\underline{\theta}) - \underline{\phi}_M'(\underline{\theta})\| = \|[\underline{f}_M'(\underline{\theta})^T \underline{f}_M(\underline{\theta})\| \le \gamma$$

for all $\underline{\theta} \in S(\underline{\theta}_0, \rho)$ by 2(ii). Applying Lemma 3.6 again, we have $\alpha = \frac{\lambda}{1 - \lambda\mu\rho}$,

$\delta = \gamma$, and $\alpha\delta = \frac{\lambda\gamma}{1 - \lambda\mu\rho} < 1$. Thus $\underline{\phi}_M'(\underline{\theta})$ is nonsingular, i.e., $\underline{\phi}_M'(\underline{\theta})$ has constant rank p for all $\underline{\theta} \in S(\underline{\theta}_0, \rho)$. This implies that $\underline{\theta}_0$ is locally least-square identifiable by Theorem 3.2.

A computation procedure for finding explicitly $S(\underline{\theta}_0, \rho)$ will be presented in the next chapter.

4. REGIONS OF PARAMETER IDENTIFIABILITY FOR DETERMINISTIC LINEAR DYNAMIC SYSTEMS

In this chapter we will study the identification algorithm extensively. We will first study Pereyra's [13] theory on the modified Newton method, which we shall call the Gauss-Newton method, for solving nonlinear least square problems. Then we will present the identification algorithm proposed by Herget [7] for finding the regions of parameter identifiability.

4.1 Nonlinear Least-Square Problems and the Gauss-Newton Method

In this section we will study the sufficient conditions for the convergence of the Newton iteration sequence and the Gauss-Newton iteration sequence.

Definition 4.1

A real-valued function $\|\cdot\|$ defined on the R^n space is called a norm if

(i) $\|\underline{x}\| \geq 0$ for all $\underline{x} \in R^n$,

(ii) $\|\underline{x}\| = 0$ if and only if $\underline{x} = 0$,

(iii) $\|\underline{x} + \underline{y}\| \leq \|\underline{x}\| + \|\underline{y}\|$ for all $\underline{x}, \underline{y} \in R^n$, and

(iv) $\|\alpha\underline{x}\| = |\alpha|\|\underline{x}\|$ for all $\alpha \in R$ and all $\underline{x} \in R^n$.

Definition 4.2

A mapping $\psi: \Omega \subset R^p \to R^m$ is said to be Frechet-differentiable at $\underline{\theta}$, where $\underline{\theta}$ is an interior point of Ω, if there is an $(m \times p)$ linear operator A such that

$$\lim_{\|\underline{k}\| \to 0} (1/\|\underline{k}\|)\|\psi(\underline{\theta} + \underline{k}) - \psi(\underline{\theta}) - A\underline{k}\| = 0$$

for all $\underline{k}$ such that $\underline{\theta}+\underline{k}\in\Omega$. A is denoted by $\psi'(\underline{\theta})$ and is called the F-derivative of ψ at $\underline{\theta}$.

By a least square problem, we mean given a nonlinear transformation $\underline{h}:\Omega\subset R^p \to R^m$ between the set $\Omega\subset R^p$ and the R^m space ($p \leq m$ in general), and the sequence of observations $\underline{Y}\in R^m$, find a parameter $\underline{\theta}_0\in\Omega$ such that the Euclidean norm of $\underline{h}(\underline{\theta}) - \underline{Y}$ is minimized at $\underline{\theta}_0$, i.e., we want to find the stationary points of the criterion $\|\underline{h}(\underline{\theta}) - \underline{Y}\|^2$.

For brevity, let $\underline{f}(\underline{\theta}) = \underline{h}(\underline{\theta}) - \underline{Y}$. If $\underline{f}(\underline{\theta})$ is at least twice Frechet-differentiable on Ω, then we can differentiate the criterion to find the minimal point of the criterion. Let

$$J(\underline{\theta}) = [\underline{h}(\underline{\theta}) - \underline{Y}]^T[\underline{h}(\underline{\theta}) - \underline{Y}] = \underline{f}^T(\underline{\theta})\underline{f}(\underline{\theta})$$

which is the square of the Euclidean norm of $\underline{h}(\underline{\theta}) - \underline{Y}$. Then

$$\underline{\phi}(\underline{\theta}) = 1/2\ J'(\underline{\theta}) = [\underline{f}'(\underline{\theta})]^T\underline{f}(\underline{\theta}) \tag{4.1}$$

where $\underline{f}'(\underline{\theta})$ is an $(m \times p)$ matrix function of $\underline{\theta}$ and $\underline{\phi}(\underline{\theta})$ is a p-component vector function of $\underline{\theta}$.

To find the stationary points of $J(\underline{\theta})$, we let

$$\underline{\phi}(\underline{\theta}) = 0 \tag{4.2}$$

Many questions arise. (1) Does there exist some $\underline{\theta}_0\in\Omega$ such that $\underline{\phi}(\underline{\theta}_0) = 0$? (2) If $\underline{\theta}_0$ exists, is $\underline{\theta}_0$ unique locally or globally? (3) If $\underline{\theta}_0$ is unique locally in some region, can we find explicitly the region? (4) If $\underline{\theta}_0$ exists, how do we construct an iteration sequence which will converge to $\underline{\theta}_0$?

To answer these questions, we proceed as follows. The standard Newton iteration sequence for solving the equation $\underline{\phi}(\underline{\theta}) = 0$ is defined by

$$\tilde{\underline{\theta}}_{n+1} = \tilde{\underline{\theta}}_n - [\underline{\phi}'(\tilde{\underline{\theta}}_n)]^{-1}\underline{\phi}(\tilde{\underline{\theta}}_n), \quad n = 0,1,\cdots, \tag{4.3}$$

where

$$\underline{\phi}'(\underline{\theta}) = [\underline{f}'(\underline{\theta})]^T\underline{f}'(\underline{\theta}) + [\underline{f}''(\underline{\theta})]^T\underline{f}(\underline{\theta}), \tag{4.4}$$

and $\underline{f}''(\underline{\theta})$ is an $m \times p \times p$ bilinear operator defined by

$$\underline{f}''(\underline{\theta}) = \left[\frac{\partial^2 f_i(\theta)}{\partial\theta_j\partial\theta_k}\right] \qquad \begin{matrix} i = 1,\cdots,m \\ j = 1,\cdots,p \\ k = 1,\cdots,p \end{matrix} \tag{4.5}$$

The manipulation of the bilinear operator follows that given in Sec. 3.1.

The Newton-Kantorovich Theorem provides sufficient conditions for the convergence of the iteration sequence (4.3) and the uniqueness of the solution for (4.2) in a region.

Theorem 4.1 [Newton-Kantorovich] [12]

Assume that $\underline{\phi}:\Omega \subset R^p \to R^p$ is F-differentiable on a convex set $\Omega_0 \subset \Omega$ and that

$$\|\underline{\phi}'(\underline{\theta}_1) - \underline{\phi}'(\underline{\theta}_2)\| \le \gamma\|\underline{\theta}_1 - \underline{\theta}_2\| \quad \text{for all } \underline{\theta}_1, \underline{\theta}_2 \in \Omega_0.$$

Suppose that there exists an $\underline{\theta}_0 \in \Omega_0$ such that $\|\underline{\phi}'(\underline{\theta}_0)\| \le \beta$ and $\alpha = \beta\gamma\eta \le 1/2$ where $\|[\underline{\phi}'(\tilde{\underline{\theta}}_0)]^{-1}\underline{\phi}(\tilde{\underline{\theta}}_0)\| \le \eta$.

Let

$$\rho_1 = (\beta\gamma)^{-1}[1 - (1 - 2\alpha)^{\frac{1}{2}}], \quad \rho_2 = (\beta\gamma)^{-1}[1 + (1 - 2\alpha)^{\frac{1}{2}}],$$

and assume that $\bar{S}(\tilde{\underline{\theta}}_0, \rho_1) \subset \Omega_0$ where $\bar{S}(\tilde{\underline{\theta}}_0, \rho_1)$ is the closure of the sphere

$S(\tilde{\underline{\theta}}_0,\rho_1)$ with radius ρ_1 centered at $\tilde{\underline{\theta}}_0$. Then the iteration sequance

$$\tilde{\underline{\theta}}_{n+1} = \tilde{\underline{\theta}}_n - [\underline{\phi}'(\tilde{\underline{\theta}}_n)]^{-1}\underline{\phi}(\tilde{\underline{\theta}}_n), \quad n = 0,1,---,$$

is well-defined, remains in $\bar{S}(\tilde{\underline{\theta}}_0,\rho_1)$ and converges to a solution $\underline{\theta}_0$ of $\underline{\phi}(\underline{\theta}) = 0$ which is unique in $S(\tilde{\underline{\theta}}_0,\rho_2) \cap \Omega_0$. Moreover, the error estimate

$$\|\underline{\theta}_0 - \tilde{\underline{\theta}}_n\| \leq (\beta\gamma 2^n)^{-1}(2\alpha)^{2n}, \quad n = 0,1,---,$$

holds.

By (4.4), we see that by using Newton's method, we must calculate $\underline{f}''(\underline{\theta})$ which is a bilinear operator in each iteration. This is a computational inconvenience. Pereyra [13] developed a modified Newton method which we shall call the Gauss-Newton method. The modification is that $\underline{\phi}'(\underline{\theta})$ is replaced by its first order approximation $[\underline{f}'(\underline{\theta})]^T\underline{f}'(\underline{\theta})$ and the modified iteration sequence is defined by

$$\tilde{\underline{\theta}}_{n+1} = \tilde{\underline{\theta}}_n - [(\underline{f}'(\tilde{\underline{\theta}}_n)]^{-1}\underline{\phi}(\tilde{\underline{\theta}}_n), \quad n = 0,1,--- \tag{4.6}$$

Due to the approximation on $\underline{\phi}'(\underline{\theta})$, another set of sufficient conditions for the convergence of (4.6) and the uniqueness of the solution of (4.2) was established by Pereyra [13].

Theorem 4.2

Assume $\phi:\Omega \to R^p$ is γ-differentiable. For brevity, let $N(\underline{\theta}) \equiv [\underline{f}'(\underline{\theta})]^T\underline{f}(\underline{\theta})$. Let $\tilde{\underline{\theta}}_0 \in \Omega$. If

(1) $N(\tilde{\underline{\theta}}_0)$ is nonsingular,

(2) there exists a sphere $S(\tilde{\underline{\theta}}_0,\rho) \subset \Omega$ of radius ρ centered at $\tilde{\underline{\theta}}_0$ such that

(i) $\|N^{-1}(\tilde{\underline{\theta}}_0)\| \leq 1/2\lambda$,

(ii) $\|[\underline{f}''(\underline{\theta})]^T \underline{f}(\underline{\theta}) - [\underline{f}''(\tilde{\underline{\theta}}_0)]^T \underline{f}(\tilde{\underline{\theta}}_0)\| \leq \gamma$ for all $\underline{\theta} \in S(\tilde{\underline{\theta}}_0, \rho)$,

(iii) $\|[\underline{f}''(\tilde{\underline{\theta}}_0)]^T \underline{f}(\tilde{\underline{\theta}}_0)\| \leq \gamma$

(iv) $\|N'(\underline{\theta})\| \leq \mu$ for all $\underline{\theta} \in S(\tilde{\underline{\theta}}_0, \rho)$

(v) $\lambda(\mu\rho + \gamma) < 1/2$,

(vi) $\eta \leq \rho$, where $\eta = \dfrac{\lambda\|\underline{g}(\tilde{\underline{\theta}}_0)\|}{1 - 2\lambda(\mu\rho + \gamma)}$.

Then the iteration sequence defined by (4.6) converges to the unique solution $\underline{\theta}_0$ of $\underline{g}(\underline{\theta}) = 0$ in the sphere $S(\tilde{\underline{\theta}}_0, \eta)$. Moreover, the error estimate is given by

$$\|\underline{\theta}_0 - \tilde{\underline{\theta}}_n\| \leq k^n \gamma,$$

where $k = 2\lambda(\mu\rho + \gamma)$.

The above theorem provides a convergence region centered at the initial iteration point such that the solution of (4.2) is unique in that region. If we have the knowledge of the solution of (4.2), then we are able to find a region centered at the solution $\underline{\theta}_0$ such that $\underline{\theta}_0$ is the unique solution of (4.2) in that region by modifying the above theorem. This will be studied in the next section.

4.2 Regions of Parameter Identifiability

After introducing the theory and algorithm for solving the nonlinear least-square problems, we are now going to apply it for the control system parameter identification problem. We shall first introduce the ℓ_∞-norm which will be employed by the identification algorithm developed in this section.

The well-known class of norms on R^n space is the ℓ_p-norm defined by

$$\|\underline{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} , \quad 1 < p < \infty$$

When $p = 2$, the ℓ_2-norm is usually called the Euclidean norm.

The limiting case of the ℓ_p-norm is the ℓ_∞-norm defined by

$$\|\underline{x}\|_\infty = \max_{1 \le i \le n} |x_i|$$

We can thus now induce a norm on the space of linear operators from R^n to R^m. We denote this space by $L(R^n, R^m)$. Given any two norms $\|\cdot\|$ and $\|\cdot\|'$ on R^n and R^m respectively, and any $A \in L(R^n, R^m)$, the norm of A with respect to $\|\cdot\|$ and $\|\cdot\|'$ is defined by

$$\|A\| = \sup_{\|\underline{x}\| = 1} \|A\underline{x}\|'$$

Such a matrix norm satisfies the properties:

(i) $\|A\| \ge 0$ for all $A \in L(R^n, R^m)$,

(ii) $\|A\| = 0$ only if $A = 0$,

(iii) $\|\alpha A\| = |\alpha| \, \|A\|$ for all $\alpha \in L(R^n, R^m)$,

(iv) $\|A + B\| \le \|A\| + \|B\|$ for all $A, B \in L(R^n, R^m)$.

The ℓ_1-, ℓ_2-, and ℓ_∞-norms are the most useful in numerical analysis work. For the system identification problem concerned in this section, we will employ the ℓ_∞-norm. We denote the ℓ_∞-norm of a matrix A by $\|A\|_\infty$. The following theorem provides the explicit expression of $\|A\|_\infty$ [12].

Theorem 4.3

Let $A \in L(R^n, R^m)$ where both R^n and R^m are normed by the ℓ_∞-norm. Denote $A = [a_{ij}]$.

Then

$$\|A\|_\infty = \max_{1 \le i \le m} \sum_{j=1}^{n} |a_{ij}| \tag{4.7}$$

Proof: For any $\underline{x} \in R^n$,

$$\|A\underline{x}\|_\infty = \max_{1 \le i \le m} |(A\underline{x})_i|$$

$$= \max_{1 \le i \le m} |\sum_{j=1}^{n} a_{ij} x_j|$$

$$\le \max_{1 \le i \le m} \sum_{j=1}^{n} |a_{ij}||x_j|$$

$$\le \max_{1 \le i \le m} \sum_{j=1}^{n} |a_{ij}| \cdot (\max_{1 \le j \le n} |x_j|)$$

$$= (\max_{1 \le i \le m} \sum_{j=1}^{n} |a_{ij}|) \|\underline{x}\|_\infty \tag{4.8}$$

It sufficies to show that there exists an $\underline{x}^* \in R^n$ such that the equality is attained in (4.8). Let k be the index such that

$$\max_{1 \le i \le m} \sum_{j=1}^{n} |a_{ij}| = \sum_{j=1}^{n} |a_{kj}|$$

Define $\underline{x}^*$ by

$$x_j^* = \begin{cases} a_{kj}/|a_{kj}| & , a_{kj} \neq 0 \\ 1 & , a_{kj} = 0 \end{cases} \qquad j = 1, \cdots, n$$

Then

$\|\underline{x}^*\|_\infty = 1$, and

$$\|A\underline{x}^*\|_\infty = \max_{1 \le i \le m} |(A\underline{x}^*)_i|$$

$$= \max_{1 \le i \le m} |\sum_{j=1}^{n} a_{ij} x_j^*|$$

If $i \neq k$, then

$$|\sum_{j=1}^{n} a_{ij} x_j^*| \le \sum_{j=1}^{n} |a_{ij}||x_j^*|$$

$$\le \sum_{j=1}^{n} |a_{ij}| |\max_{1 \le j \le n} x_j^*|$$

$$= \sum_{j=1}^{n} |a_{ij}|$$

$$\le \sum_{j=1}^{n} |a_{kj}|$$

If $i = k$, then

$$|\sum_{j=1}^{n} a_{kj} x_j^*| = \sum_{j=1}^{n} |a_{kj}|$$

Thus $\|A\underline{x}^*\|_\infty = \sum_{j=1}^{n} |a_{kj}| = \max_{1 \le i \le m} \sum_{j=1}^{n} |a_{ij}|$, and the equality in (4.8) is attained.

We can also induce an ℓ_∞-norm on an $(m \times p \times p)$ bilinear operator B defined by

$$\|B\|_\infty = \sup_{\|\underline{x}\| = 1} \|B\underline{x}\|_\infty$$

We now consider the problem of identifying the vector of unknown parameters $\underline{\eta}$, and the unknown initial state $\underline{w}(0)$ of a parametrized deterministic system whose state at time k, is the vector, $\underline{w}(k)$, where $k = 0,1,2---$. Let $\underline{x}(k) = \begin{bmatrix} \underline{w}(k) \\ \underline{\eta} \end{bmatrix}$, which we shall call the augmented state

vector, and let $\underline{x}(k)$ be a p-component vector. Then $\underline{\theta} \triangleq \underline{x}(0) = \begin{bmatrix} \underline{w}(0) \\ \underline{\eta} \end{bmatrix}$ is to be identified. We assume that the true parameter $\underline{\theta}_0$ is an interior point of a known compact subset $\Omega \subset R^p$, and the system is stable for all admissible inputs and $\underline{\theta} \in \Omega$. We note here that after the augmentation, the identification problem of the original system is equivalent to the initial-state observation problem of the augmented system. We assume that the function $\underline{g}(\underline{x},k)$ is known as a function of $\underline{x}$ and k and that

$$\underline{x}(k+1) = \underline{g}(\underline{x}(k),k) \tag{4.9}$$

We also assume that the inputs to the system are known implicitly in $\underline{g}(\underline{x},k)$.

Furthermore, we assume that observations of the state can be written in the form

$$\underline{y}(k) = C\underline{x}(k) \tag{4.10}$$

where C is a known $(r \times p)$ constant matrix.

Let

$$\underline{Y}_M \triangleq \begin{bmatrix} \underline{y}(0) \\ \underline{y}(1) \\ \cdot \\ \cdot \\ \cdot \\ \underline{y}(M) \end{bmatrix} \tag{4.11}$$

be the m-component vector of observations, where $m = r(M+1) \geq p$. Let

$$\underline{h}_M(\underline{\theta}) \triangleq \begin{bmatrix} C\underline{x}(0) \\ C\underline{x}(1) \\ \cdot \\ \cdot \\ \cdot \\ C\underline{x}(M) \end{bmatrix} \tag{4.12}$$

where $\underline{x}(k)$ is the solution to (4.9) when $\underline{x}(0) = \underline{\theta}$. Moreover, let

$$\underline{f}_M(\underline{\theta}) \triangleq \underline{h}_M(\underline{\theta}) - \underline{Y}_M \tag{4.13}$$

We assume that we have obtained a solution, say $\underline{\theta}_0$, such that

$$\underline{f}_M(\underline{\theta}_0) = 0 \tag{4.14}$$

i.e., $\underline{\theta}_0$ is the parameter value which when applied with the input sequence from a given set of input-output sequence will generate a set of output data which will match the given output sequence. In order that $\underline{\theta}_0$ is identifiable, $\underline{\theta}_0$ must be an interior point of an open sphere $S(\underline{\theta}_0,\rho) \subset \Omega$ with radius $\rho > 0$ centered at $\underline{\theta}_0$ such that $\underline{\theta}_0$ is the unique solution to the equation

$$\underline{f}_M(\underline{\theta}) = 0 \tag{4.15}$$

in $S(\underline{\theta}_0,\rho)$. We are now going to establish the sufficient conditions for the existence of $S(\underline{\theta}_0,\rho)$. Let the identification criterion be

$$J_M(\underline{\theta}) \triangleq [\underline{f}_M(\underline{\theta})]^T \underline{f}_M(\underline{\theta}) \tag{4.16}$$

We see immediately that $J_M(\underline{\theta}) = 0$ if and only if $\underline{f}_M(\underline{\theta}) = 0$. Since $J_M(\underline{\theta})$ is nonnegative for all $\underline{\theta}$, therefore its minimal value is zero. If we have

the knowledge of $\underline{\theta}_0$, then $J_M(\underline{\theta}_0) = \underline{f}_M(\underline{\theta}_0) = 0$ and $\underline{\theta}_0$ is the minimal point of $J_M(\underline{\theta})$. Let

$$\underline{\phi}_M(\underline{\theta}) = 1/2\ J_M'(\underline{\theta}) = [\underline{f}_M'(\underline{\theta})]^T \underline{f}_M(\underline{\theta}), \tag{4.17}$$

then

$$\underline{\phi}_M'(\underline{\theta}) = [\underline{f}_M'(\underline{\theta})]^T \underline{f}_M'(\underline{\theta}) + [\underline{f}_M''(\underline{\theta})]^T \underline{f}_M(\underline{\theta}) \tag{4.18}$$

Let

$$N_M(\underline{\theta}) = [\underline{f}_M'(\underline{\theta})]^T \underline{f}_M'(\underline{\theta}) \tag{4.19}$$

Note that

$$\underline{\phi}_M(\underline{\theta}_0) = 0, \tag{4.20}$$

and

$$\underline{\phi}_M'(\underline{\theta}_0) = [\underline{f}_M'(\underline{\theta}_0)]^T \underline{f}_M'(\underline{\theta}_0) = N_M(\underline{\theta}_0) \tag{4.21}$$

Thus if there exists an open sphere $S(\underline{\theta}_0,\rho)$ such that $\underline{\phi}_M'(\underline{\theta})$ is non-singular for all $\underline{\theta} \in S(\underline{\theta}_0,\rho)$, then $\underline{\phi}_M(\underline{\theta})$ is locally injective on $S(\underline{\theta}_0,\rho)$ and thus $\underline{\theta}_0$ is locally identifiable by Theorem 3.2. The following theorem, established by Herget [7], provides sufficient conditions for the existence of such $S(\underline{\theta}_0,\rho)$.

Theorem 4.4

Let R^p and R^m be the vector spaces of p and m-tuples respectively over R. Let $\|\cdot\|$ denotes any norm on the respective spaces. Let $\underline{f}_M(\underline{\theta})$ be an m-vector function of $\underline{\theta} \in R^p$ which is twice F-differentiable on $\Omega \subset R^p$.

If (1) there is $\underline{\theta}_0 \in \Omega$ such that $\underline{f}_M(\underline{\theta}_0) = 0$,

(2) $N_M(\underline{\theta}_0)$ is nonsingular,

(3) there is a sphere $S(\underline{\theta}_0, \rho)$ with radius $\rho > 0$ centered at $\underline{\theta}_0$ and a set of positive numbers $(\lambda, \gamma, \mu, \eta)$ such that

(i) $\|N_M^{-1}(\underline{\theta}_0)\| \le \lambda/2$,

(ii) $\|[\underline{f}_M''(\underline{\theta})]\underline{f}_M(\underline{\theta})\| \le \gamma$ for all $\underline{\theta} \in S(\underline{\theta}_0, \rho)$,

(iii) $\|N_M'(\underline{\theta})\| \le \mu$ for all $\underline{\theta} \in S(\underline{\theta}_0, \rho)$,

(iv) $\|\underline{\phi}_M(\underline{\theta})\| \le \eta$, for all $\underline{\theta} \in S(\underline{\theta}_0, \rho)$,

(v) $\lambda\eta \le \rho/4$,

(vi) $\lambda(2\mu\rho + \gamma) \le 1/2$,

then $\underline{\theta}_0$ is locally identifiable in $S(\underline{\theta}_0, \rho)$. Moreover, for any $\tilde{\underline{\theta}}_0 \in S(\underline{\theta}_0, \rho/2)$, the iteration sequence

$$\tilde{\underline{\theta}}_{n+1} = \tilde{\underline{\theta}}_n - N_M^{-1}(\tilde{\underline{\theta}}_n)\underline{\phi}_M(\tilde{\underline{\theta}}_n), \quad n = 0, 1, 2, \cdots$$

converges to the unique solution $\underline{\theta}_0$ of $\underline{f}_M(\underline{\theta}) = 0$ in $S(\underline{\theta}_0, \rho/2)$.

Proof: We shall first prove that $N_M(\underline{\theta})$ is nonsingular and $\|N_M^{-1}(\underline{\theta})\| \le \lambda$ for all $\underline{\theta} \in S(\underline{\theta}_0, \rho)$. By 3(iii), we have

$$\|N_M(\underline{\theta}) - N_M(\underline{\theta}_0)\| \le \mu\rho \quad \text{for all } \underline{\theta} \in S(\underline{\theta}_0, \rho)$$

Applying Lemma 3.6 to this situation, we have $\alpha = \lambda/2$, $\delta = \mu\rho$ and $\alpha\delta = \lambda\mu\rho/2 \le 1/8$. Therefore $N_M(\underline{\theta})$ is nonsingular and $\|N_M^{-1}(\underline{\theta})\| \le \dfrac{\lambda/2}{1 - \frac{\lambda\mu\rho}{2}} = \dfrac{\lambda}{1 + (1 - \lambda\mu\rho)} < \lambda$ for all $\underline{\theta} \in S(\underline{\theta}_0, \rho)$.

A result due to Bartle in Herget [7] is: If $\|\underline{\phi}_M'(\underline{\theta}) - \underline{\phi}_M'(\underline{\theta}_0)\| \le \beta$ for all $\underline{\theta} \in S(\underline{\theta}_0, \rho)$, then $\|\underline{\phi}_M(\underline{\theta}_1) - \underline{\phi}_M(\underline{\theta}_2) - \underline{\phi}_M'(\underline{\theta}_0)(\underline{\theta}_1 - \underline{\theta}_2)\| \le \beta\|\underline{\theta}_1 - \underline{\theta}_2\|$ for all $\underline{\theta}_1, \underline{\theta}_2 \in S(\underline{\theta}_0, \rho)$. In our case, $\underline{\phi}_M'(\underline{\theta}_0) = N_M(\underline{\theta}_0)$ by (4.21). Thus by 3(ii) and

(iii)

$$\|\underline{\phi}_M'(\underline{\theta}) - \underline{\phi}_M'(\underline{\theta}_0)\| = \|N_M(\underline{\theta}) - N_M(\underline{\theta}_0) + [\underline{f}_M''(\underline{\theta})]^T \underline{f}_M(\underline{\theta})\|$$

$$\leq \mu\rho \quad \text{for all } \underline{\theta} \in S(\underline{\theta}_0, \rho)$$

Therefore we have

$$\|\underline{\phi}_M(\underline{\theta}_1 - \underline{\phi}_M(\underline{\theta}_2) - \underline{\phi}_M'(\underline{\theta}_0)(\underline{\theta}_1 - \underline{\theta}_2)\| \leq (\mu\rho + \gamma)\|\underline{\theta}_1 - \underline{\theta}_2\|$$

for all $\underline{\theta}_1, \underline{\theta}_2 \in S(\underline{\theta}_0, \rho)$.

Also, look at

$$N_M(\underline{\theta}) - \underline{\phi}_M'(\underline{\theta}_0) = N_M(\underline{\theta}) - N_M(\underline{\theta}_0)$$

Thus

$$\|N_M(\underline{\theta}) - \underline{\phi}_M'(\underline{\theta}_0)\| \leq \mu\rho \quad \text{for all } \underline{\theta} \in S(\underline{\theta}_0, \rho)$$

Let $\tilde{\underline{\theta}}_0 \in S(\underline{\theta}_0, \rho/2)$, and

$$\tilde{\underline{\theta}}_{n+1} = \tilde{\underline{\theta}}_n - N_M^{-1}(\tilde{\underline{\theta}}_n)\phi_M(\tilde{\underline{\theta}}_n), \quad n = 0,1,2,---$$

We want to show that $\tilde{\underline{\theta}}_n \in S(\underline{\theta}_0, \rho/2)$ for all n. We prove it by induction.

$$\tilde{\underline{\theta}}_1 = \tilde{\underline{\theta}}_0 - N_M^{-1}(\tilde{\underline{\theta}}_0)\phi_M(\tilde{\underline{\theta}}_0)$$

Thus $\|\tilde{\underline{\theta}}_1 - \tilde{\underline{\theta}}_0\| \leq \lambda\eta \leq \rho/4$. Hence $\tilde{\underline{\theta}}_1 \in S(\underline{\theta}_0, \rho/2)$ and $\tilde{\underline{\theta}}_1 \in S(\underline{\theta}_0, \rho)$. Now assume $\tilde{\underline{\theta}}_\nu \in S(\underline{\theta}_0, \rho/2)$ for $\nu = 1,---,n$. Then we have

$$-N_M(\tilde{\underline{\theta}}_{\nu-1})(\tilde{\underline{\theta}}_\nu - \tilde{\underline{\theta}}_{\nu-1}) = \underline{\phi}_M(\tilde{\underline{\theta}}_{\nu-1})$$

Thus $\underset{\sim}{\phi}_M(\tilde{\underset{\sim}{\theta}}_\nu) = \underset{\sim}{\phi}_M(\tilde{\underset{\sim}{\theta}}_\nu) - \underset{\sim}{\phi}_M(\tilde{\underset{\sim}{\theta}}_{\nu-1}) + \underset{\sim}{\phi}_M(\tilde{\underset{\sim}{\theta}}_{\nu-1})$

$$= \underset{\sim}{\phi}_M(\tilde{\underset{\sim}{\theta}}_\nu) - \underset{\sim}{\phi}_M(\tilde{\underset{\sim}{\theta}}_{\nu-1}) - N_M(\tilde{\underset{\sim}{\theta}}_{\nu-1})(\tilde{\underset{\sim}{\theta}}_\nu - \tilde{\underset{\sim}{\theta}}_{\nu-1})$$

$$= \underset{\sim}{\phi}_M(\tilde{\underset{\sim}{\theta}}_\nu) - \underset{\sim}{\phi}_M(\tilde{\underset{\sim}{\theta}}_{\nu-1}) - \underset{\sim}{\phi}_M'(\underset{\sim}{\theta}_0)(\tilde{\underset{\sim}{\theta}}_\nu - \tilde{\underset{\sim}{\theta}}_{\nu-1})$$

$$+ [\underset{\sim}{\phi}_M'(\underset{\sim}{\theta}_0) - N_M(\tilde{\underset{\sim}{\theta}}_{\nu-1})](\tilde{\underset{\sim}{\theta}}_\nu - \tilde{\underset{\sim}{\theta}}_{\nu-1})$$

Therefore $\|\underset{\sim}{\phi}_M(\tilde{\underset{\sim}{\theta}}_\nu)\| \le (\mu\rho + \gamma)\|\tilde{\underset{\sim}{\theta}}_\nu - \tilde{\underset{\sim}{\theta}}_{\nu-1}\| + \mu\rho\|\tilde{\underset{\sim}{\theta}}_\nu - \tilde{\underset{\sim}{\theta}}_{\nu-1}\| = k\|\tilde{\underset{\sim}{\theta}}_\nu - \tilde{\underset{\sim}{\theta}}_{\nu-1}\|$

where $k = 2\mu\rho + \gamma$.

Now $\|\tilde{\underset{\sim}{\theta}}_\nu - \tilde{\underset{\sim}{\theta}}_{\nu-1}\| \le \lambda\|\underset{\sim}{\phi}_M(\tilde{\underset{\sim}{\theta}}_{\nu-1})\|$

which implies

$$\|\tilde{\underset{\sim}{\theta}}_\nu - \tilde{\underset{\sim}{\theta}}_{\nu-1}\| \le \lambda k\|\tilde{\underset{\sim}{\theta}}_{\nu-1} - \tilde{\underset{\sim}{\theta}}_{\nu-2}\|$$

$$\le (\lambda k)^{\nu-1}\|\tilde{\underset{\sim}{\theta}}_1 - \tilde{\underset{\sim}{\theta}}_0\|$$

Look at

$$\tilde{\underset{\sim}{\theta}}_{n+1} = \tilde{\underset{\sim}{\theta}}_n - N_M^{-1}(\tilde{\underset{\sim}{\theta}}_n)\phi_M(\tilde{\underset{\sim}{\theta}}_n)$$

$$\|\tilde{\underset{\sim}{\theta}}_{n+1} - \tilde{\underset{\sim}{\theta}}\| \le \lambda\|\phi_M(\tilde{\underset{\sim}{\theta}}_n)\|$$

$$\le (\lambda k)\|\tilde{\underset{\sim}{\theta}}_n - \tilde{\underset{\sim}{\theta}}_{n-1}\|$$

Therefore

$$\|\tilde{\underset{\sim}{\theta}}_{n+1} - \tilde{\underset{\sim}{\theta}}_n\| \le (\lambda k)^n\|\tilde{\underset{\sim}{\theta}}_1 - \tilde{\underset{\sim}{\theta}}_0\|.$$

Note that

$$\tilde{\underset{\sim}{\theta}}_{n+1} - \tilde{\underset{\sim}{\theta}}_0 = \sum_{\nu=0}^{n}[\tilde{\underset{\sim}{\theta}}_{\nu+1} - \tilde{\underset{\sim}{\theta}}_\nu]$$

Thus

$$\begin{aligned}\|\tilde{\underline{\theta}}_{n+1} - \tilde{\underline{\theta}}_0\| &\le \sum_{\nu=0}^{n} \|\tilde{\underline{\theta}}_{\nu+1} - \tilde{\underline{\theta}}_\nu\| \\ &\le \sum_{\nu=0}^{n} (\lambda k)^\nu \|\tilde{\underline{\theta}}_1 - \tilde{\underline{\theta}}_0\| \\ &\le \sum_{\nu=0}^{\infty} (\lambda k)^\nu \|\tilde{\underline{\theta}}_1 - \tilde{\underline{\theta}}_0\| \\ &= \frac{1}{1-\lambda k} \|\tilde{\underline{\theta}}_1 - \tilde{\underline{\theta}}_0\| \\ &\le \left(\frac{1}{1-\lambda k}\right) \frac{\rho}{4} \\ &\le \frac{1}{(1/2)} \cdot \frac{\rho}{4} = \rho/2\end{aligned}$$

which implies $\tilde{\underline{\theta}}_{n+1} \in S(\underline{\theta}_0, \rho/2)$, thus the induction proof is completed. Next, we shall show that $\{\tilde{\underline{\theta}}_n\}_{n=0}^{\infty}$ is a Cauchy sequence. Let $m > n$ and $m = n + j$, then

$$\tilde{\underline{\theta}}_m - \tilde{\underline{\theta}}_n = \tilde{\underline{\theta}}_{n+j} - \tilde{\underline{\theta}}_n = \sum_{\nu=n}^{n+j-1} [\tilde{\underline{\theta}}_{\nu+1} - \tilde{\underline{\theta}}_\nu].$$

Thus

$$\begin{aligned}\|\tilde{\underline{\theta}}_{n+j} - \tilde{\underline{\theta}}_n\| &\le \sum_{\nu=n}^{n+j-1} \|\tilde{\underline{\theta}}_{\nu+1} - \tilde{\underline{\theta}}_\nu\| \\ &\le \sum_{\nu=n}^{n+j-1} (\lambda k)^\nu \|\tilde{\underline{\theta}}_1 - \tilde{\underline{\theta}}_0\| \\ &= (\lambda k)^n \left[\sum_{\nu=0}^{j-1} (\lambda k)^\nu\right] \|\tilde{\underline{\theta}}_1 - \tilde{\underline{\theta}}_0\| \\ &\le (\lambda k)^n \left[\sum_{\nu=0}^{\infty} (\lambda k)^\nu\right] \|\tilde{\underline{\theta}}_1 - \tilde{\underline{\theta}}_0\| \\ &= \frac{(\lambda k)^n}{1-\lambda k} \|\tilde{\underline{\theta}}_1 - \tilde{\underline{\theta}}_0\|\end{aligned}$$

$$\leq \frac{(\lambda k)^n}{1-\lambda k} \left(\frac{\rho}{4}\right)$$

$$\leq \frac{(\lambda k)^n}{1/2} \left(\frac{\rho}{4}\right)$$

$$= (\lambda k)^n \left(\frac{\rho}{4}\right)$$

Since $\lambda k \leq 1/2$, therefore $\lim_{n \to \infty} (\lambda k)^n \left(\frac{\rho}{2}\right) = 0$ and hence $\{\tilde{\underline{\theta}}_n\}_{n=0}^{\infty}$ is a Cauchy sequence.

Since R^p is a complete normed space, $\{\tilde{\underline{\theta}}_n\}_{n=0}^{\infty}$ converges to some point $\underline{\theta}_1 \epsilon S(\underline{\theta}_0, \rho)$. We must show that $\underline{\theta}_0 = \underline{\theta}_1$. On the contrary suppose $\underline{\theta}_1 \neq \underline{\theta}_0$, then we can write

$$\underline{\theta}_0 - \underline{\theta}_1 = N_M^{-1}(\underline{\theta}_0) N_M(\underline{\theta}_0)(\underline{\theta}_0 - \underline{\theta}_1)$$

$$= N_M^{-1}(\underline{\theta}_0)[\underline{\phi}_M(\underline{\theta}_1) - \underline{\phi}_M(\underline{\theta}_0) - \underline{\phi}_M'(\underline{\theta}_0)(\underline{\theta}_1 - \underline{\theta}_0)$$

$$+ (N_M(\underline{\theta}_0) - \underline{\phi}_M'(\underline{\theta}_0))(\underline{\theta}_0 - \underline{\theta}_1)]$$

since $\underline{\phi}_M(\underline{\theta}_0) = \phi_M(\underline{\theta}_1) = 0$ and $N_M(\underline{\theta}_0) = \underline{\phi}_M(\underline{\theta}_0)$.

Thus

$$\|\underline{\theta}_0 - \underline{\theta}_1\| \leq \lambda(\mu\rho + \gamma + \mu\rho)\|\underline{\theta}_0 - \underline{\theta}_1\|$$

$$\leq 1/2 \|\underline{\theta}_0 - \underline{\theta}_1\|$$

which is a contradiction. Thus $\underline{\theta}_0 = \underline{\theta}_1$ and the proof is completed.

Remarks:

(1) $S(\underline{\theta}_0, \rho)$ is the region of parameter identifiability, i.e., $\underline{\theta}_0$ is the unique solution for $\underline{f}_M(\underline{\theta}) = 0$ in $S(\underline{\theta}_0, \rho)$, but the convergence

of the Gauss-Newton iteration sequence is assured in $S(\underline{\theta}_0,\rho/2)$. This is an added feature. However, our primary interest is the uniqueness of the solution of $\underline{f}_M(\underline{\theta}) = 0$ in some region in Ω. We have assumed we already know $\underline{\theta}_0$.

(2) If we are concerned about the uniqueness of the solution only, then conditions 3(i), (ii), (iii) of Theorem 4.4 and $\lambda(\mu\rho+\gamma) < 2$ suffice for the existence of $S(\underline{\theta}_0,\rho)$. To prove this, we suppose that there is a $\underline{\theta}_1 \in S(\underline{\theta}_0,\rho)$ such that $\underline{f}_M(\underline{\theta}_1) = 0$, then $\underline{\phi}_M(\underline{\theta}_1) = 0$. Thus we can write

$$\underline{\theta}_0 - \underline{\theta}_1 = N_M^{-1}(\underline{\theta}_0)N_M(\underline{\theta}_0)(\underline{\theta}_0 - \underline{\theta}_1)$$

$$= N_M^{-1}(\underline{\theta}_0)(\underline{\phi}_M(\underline{\theta}_1) - \underline{\phi}_M(\underline{\theta}_0) - \underline{\phi}_M'(\underline{\theta}_0)(\underline{\theta}_1 - \underline{\theta}_0)$$

$$+ (N_M(\underline{\theta}_0) - \underline{\phi}_M'(\underline{\theta}_0))(\underline{\theta}_0 - \underline{\theta}_1))$$

Thus

$$\|\underline{\theta}_0 - \underline{\theta}_1\| \leq \lambda/2(\mu\rho+\gamma)\|\underline{\theta}_0 - \underline{\theta}_1\|$$
$$< \|\underline{\theta}_0 - \underline{\theta}_1\|$$

which is a contradiction. The region implied by these set of sufficient conditions will generally be of larger radius since $\lambda(\mu\rho+\gamma) \leq 1/2$ is more conservative than $\lambda(\mu\rho+\gamma) < 2$.

(3) Theorem 3.2 is implied by Theorem 4.4. To prove this, we first recall that $N_M(\underline{\theta})$ is nonsingular and $\|N_M^{-1}(\underline{\theta})\| < \lambda$ for all $\underline{\theta} \in S(\underline{\theta}_0,\rho)$, and $\underline{\phi}_M'(\underline{\theta}) = N_M(\underline{\theta}) + [\underline{f}_M''(\underline{\theta})]^T\underline{f}_M(\underline{\theta})$. Therefore

$$\|N_M(\underline{\theta}) - \underline{\phi}_M'(\underline{\theta})\| = \|[\underline{f}_M''(\underline{\theta})]^T\underline{f}_M(\underline{\theta})\| \leq \gamma$$

for all $\underline{\theta} \in S(\underline{\theta}_0, \rho)$. Apply Lemma 3.6 again, we have $\alpha = \lambda$, $\delta = \gamma$ and $\lambda\gamma \leq 1/2 < 1$. Thus $\underline{\phi}_M'(\underline{\theta})$ is nonsingular for all $\underline{\theta} \in S(\underline{\theta}_0, \rho)$ and $\underline{\phi}_M(\underline{\theta})$ is locally injective on $S(\underline{\theta}_0, \rho)$.

We now derive a set of recursive formulas for computationally verifying condition 3(i) to 3(vi) of Theorem 4.4 for the case of dynamic systems.

Recalling that

$$\underline{f}_M(\underline{\theta}) = \underline{h}_M(\underline{\theta}) - Y_M$$

$$= \begin{bmatrix} C\underline{x}(0) \\ C\underline{x}(1) \\ \vdots \\ C\underline{x}(M) \end{bmatrix} - \begin{bmatrix} \underline{y}(0) \\ \underline{y}(1) \\ \vdots \\ \underline{y}(M) \end{bmatrix}$$

$$= \begin{bmatrix} C\underline{x}(0) \\ C\underline{g}(\underline{x}(0), 0) \\ \vdots \\ C\underline{g}(\underline{x}(M-1), M-1) \end{bmatrix} - \begin{bmatrix} \underline{y}(0) \\ \underline{y}(1) \\ \vdots \\ \underline{y}(M) \end{bmatrix} \qquad (4.22)$$

Thus

$$\underline{f}_M'(\underline{\theta}) = \begin{bmatrix} C & & \\ C\underline{g}'(\underline{x}(0), 0) & & \\ \vdots & & \\ C\underline{g}'(\underline{x}(M-1), M-1) & \cdots & \underline{g}'(\underline{x}(0), 0) \end{bmatrix}$$

$$= \begin{bmatrix} C\Gamma_0(\underline{\theta}) \\ C\Gamma_1(\underline{\theta}) \\ \vdots \\ C\Gamma_M(\underline{\theta}) \end{bmatrix} \tag{4.23}$$

where $\Gamma_k(\underline{\theta}) = g'(\underline{x}(M-1), M-1)\ \Gamma_{k-1}(\underline{\theta})$, $k = 1, \cdots, M$ (4.24)

and $\Gamma_0(\underline{\theta}) = I_{p \times p}$ = $(p \times p)$ identity matrix, (4.25)

$$\underline{f}''_M(\underline{\theta}) = \begin{bmatrix} C\Gamma'_0(\underline{\theta}) \\ C\Gamma'_1(\underline{\theta}) \\ \vdots \\ C\Gamma'_M(\underline{\theta}) \end{bmatrix} \tag{4.26}$$

where $\Gamma'_k(\underline{\theta}) = [g''(\underline{x}(k-1), k-1)\ \Gamma_{k-1}(\underline{\theta})]^* \Gamma_{k-1}(\underline{\theta})$

$$+ g'(\underline{x}(k-1), k-1)\ \Gamma'_{k-1}(\underline{\theta}),\ k = 1, \cdots, M \tag{4.27}$$

and

$$\Gamma'_0(\underline{\theta}) = 0 \tag{4.28}$$

To compute the bounds on the norms given in the hypothesis of Theorem 4.4, we choose the ℓ_∞-norm since the procedure is relatively straightforward if we employ the interval arithmetic [11].

Let

$\mathcal{I}$ = {set of all finite closed intervals $[a,b]: a,b \in R, a \leq b$}

The interval $I = [a,a]$ is called a degenerate interval. The interval

arithmetic operations are defined by

$$I * J \triangleq \{x * y : x \in I, y \in J\} \quad \text{for all } I, J \in \mathcal{I} \tag{4.29}$$

where the symbol "*" indicates one of the arithmetic operations +,-,·, and /, except that I/J is not defined if $0 \in J$.

For example,

$$[a,b] + [c,d] = [a+c, b+d]$$

$$[a,b] \cdot [c,d] = [\min(ac,ad,bc,bd),\ \max(ac,ad,bc,bd)]$$

$$[a,b]/[c,d] = [a,b] \cdot [1/d, 1/c], \text{ provided } 0 \notin [c,d].$$

If the real number ξ is to belong to the closed interval $[\xi_L, \xi_R]$ on the real numbers, $\xi_L \leq \xi_R$, we denote this interval by

$$[\xi] \triangleq [\xi_L, \xi_R] \quad \text{for brevity.}$$

If $\psi(x)$ is a continuous, real-valued function of $x \in R$, then the interval function $[\psi([x])]$ is defined by

$$[\psi([x])] \triangleq \{y : y = \psi(x), x \in [x]\}$$

An interval function will be called a rational function if it is defined and can be expressed as a rational interval arithmetic expression in the interval variable and a finite set of constant coefficient intervals. For simplicity, we shall assume that all of our functions of $\underline{\theta} \in R^p$ are rational functions so that interval arithmetic suffices to evaluate their norms. It is always true that the true interval function is a proper subset of the computed interval, i.e.,

$$\bigcup_{x\in[x]} \psi([x,x]) = [\psi([x])] \subsetneq I\psi([x])$$

where $I\psi([x])$ is the computed interval obtained by replacing x by [x] and evaluating ψ by interval arithmetic instead of ordinary arithmetic. The following Theorem in [11] proves the above statement.

Theorem 4.5

Let $I = [a,b] \in \mathcal{I}$ and

$$\mathcal{I}_I = \{J \in \mathcal{I} : J \subset I\}$$

Let $\underline{\psi} = \{\psi : \psi = \mathcal{I}_I \to \mathcal{I}, \ \psi \text{ is continuous on } \mathcal{I}_I\}$, and d be a metric on $\underline{\psi}$ such that $(\underline{\psi}, d)$ is a complete metric space. For any rational function $\psi \in (\underline{\psi}, d)$ and arbitrary $J \in \mathcal{I}_I$,

$$\bigcup_{x\in[x]} \psi([x,x]) \subsetneq I\psi(J)$$

Proof: From (4.29), it follows that if $I,J,k,L \in \mathcal{I}$, $I \subset k$ and $J \subset L$, then

$$I * J \subset k * L$$

provided in the case of division that $0 \notin L$. This property of interval arithmetic is called "monotonic inclusion". Hence the result is obvious from the monotonic inclusion property and the definition of the rational interval function. Since a finite number of these operations is involved and since for every $x \in J$ $[x,x] \subset J$, $[x,x] \in \mathcal{I}_I$, then $\psi([x,x]) \subset I\psi(J)$. To prove that equality need not be attained, it suffices to give an example.

Let $\psi(x) = x^2$, $J = [-1/2, 1]$

Then $I\psi(J) = J^2 = [-1/2, 1]$,

but $[\psi([x])] = [0,1]$,

thus $[-1/2,1] \not\subset [0,1]$.

We can now apply the interval arithmetic to the evaluations of norms required in Theorem 4.4.

The closed sphere $\bar{S}(\underline{\theta}_0,\rho)$ is the vector interval given by

$$[\underline{\theta}_0] = \bar{S}(\underline{\theta}_0,\rho) = \prod_{i=1}^{p} [\theta_{0i}] = \prod_{i=1}^{p} [\theta_{0i} - \rho, \theta_{0i} + \rho] \tag{4.30}$$

which is the Cartesian product of closed intervals. We consider condition 3(iv) of Theorem 4.4 first. We wish to have a bound on

$$\sup_{\underline{\theta} \in S(\underline{\theta}_0,\rho)} \|\underline{\phi}_M(\underline{\theta})\| = \sup_{\underline{\theta} \in S(\underline{\theta}_0,\rho)} \operatorname{Max}_{1 \le i \le p} |\phi_{M_i}(\underline{\theta})|.$$

Let

$$[\phi_{M_i}([\underline{\theta}_0])] = [\phi_{M_{L_i}}([\underline{\theta}_0]), \phi_{M_{R_i}}([\underline{\theta}_0])],$$

then

$$\sup_{\underline{\theta} \in S(\underline{\theta}_0,\rho)} \|\underline{\phi}_M(\underline{\theta})\| = \operatorname{Max}_{1 \le i \le p} \{\operatorname{Max}[|\phi_{M_{L_i}}([\underline{\theta}_0])|, |\phi_{M_{R_i}}([\underline{\theta}_0])|]\}.$$

The computation of $I\underline{\phi}_M([\underline{\theta}_0])$ is generally much easier than that of the $[\underline{\phi}_M([\underline{\theta}_0])]$, moreover, by Theorem 4.5,

$$\sup_{\underline{\theta} \in S(\underline{\theta}_0,\rho)} \|\underline{\phi}_M(\underline{\theta})\| \le \operatorname{Max}_{1 \le i \le p} \{\operatorname{Max}[|I_L\phi_{M_i}([\underline{\theta}_0])|, |I_R\phi_{M_i}([\underline{\theta}_0])|]\} \tag{4.31}$$

Hence we will compute the right hand side of (4.31) as the bound of the norm of $\underline{\phi}_M(\underline{\theta})$.

For condition 3(ii) of Theorem 4.4, we let

$$A(\underline{\theta}) = [a_{ij}(\underline{\theta})] \equiv [\underline{f}''_M(\underline{\theta})]^T \underline{f}_M(\underline{\theta}) \tag{4.32}$$

for brevity.

Then

$$\sup_{\underline{\theta} \in S(\underline{\theta}_0, \rho)} \|A(\underline{\theta})\| \leq \max_{1 \leq i \leq p} \left[\sum_{j=1}^{p} \max\{|I_L a_{ij}([\underline{\theta}_0])|, |I_R a_{ij}([\underline{\theta}_0])|\}\right] \tag{4.33}$$

For condition 3(iii) of Theorem 4.4, we let

$$B(\underline{\theta}) = [b_{ijk}(\underline{\theta})] \equiv N'_M(\underline{\theta}) \tag{4.34}$$

where

$$N'_M(\underline{\theta}) = [\underline{f}'_M(\underline{\theta})]^T \underline{f}''_M(\underline{\theta}) + ([\underline{f}''_M(\underline{\theta})]^T \underline{f}'_M(\underline{\theta}))^T \tag{4.35}$$

Then

$$\sup_{\underline{\theta} \in S(\underline{\theta}_0, \rho)} \|B(\underline{\theta})\| = \sup_{\underline{\theta} \in S(\underline{\theta}_0, \rho)} \left[\sup_{\|\underline{\xi}\| = 1} \|B(\underline{\theta})\underline{\xi}\| \right] \tag{4.36}$$

We note that $\|\underline{\xi}\| = 1$ if and only if $\xi_i = \pm 1$ for some $i = 1, \cdots, p$. Hence we introduce the following notation. For $i, n = 1, \cdots, p$, let

$$[\xi_i]_n = [-1, +1] \quad \text{if } i \neq n$$

and $$[\xi_i]_i = [+1, +1]$$

For $i = 1, \cdots, p$ and $n = p+1, \cdots, 2p$, let

$$[\xi_i]_n = [-1, +1] \quad \text{if } i \neq n - p$$

$$[\xi_i]_n = [-1, -1] \quad \text{if } i = n - p$$

Then we have

$$\sup_{\underline{\theta} \in S(\underline{\theta}_0, \rho)} \|B(\underline{\theta})\| \leq \max_{1 \leq i \leq p} \left[\max_{1 \leq n \leq 2p} \sum_{j=1}^{p} \sum_{k=1}^{p} \max\{|I_L b_{ijk}([\underline{\theta}_0]) \cdot [\xi_k]_n|, |I_R b_{ijk}([\underline{\theta}_0]) \cdot [\xi_k]_n|\}\right] \tag{4.37}$$

We summarize this algorithm by the flow graph in Fig. 1.

4.3 An Example of Computing the Region of Parameter Identifiability

If a system is linear, time-invariant, and its parametrization is known, then its augmented system has the quadratic-in-the-state bilinear representation.

$$\begin{aligned} \underline{x}(k+1) &= [F + D\underline{x}(k)]\underline{x}(k) + [G + E\underline{x}(k)]\underline{u}(k) \\ \underline{y}(k) &= C\underline{x}(k) \quad , \end{aligned} \tag{4.38}$$

i.e.,

$$\underline{g}(\underline{x},k) = [F + D\underline{x}]\underline{x} + [G + E\underline{x}]\underline{u}(k), \tag{4.39}$$

where $\underline{u}(k)$ is a q-vector, F is a $(p \times p)$ matrix, G is a $(p \times q)$ matrix, D is a $(p \times p \times p)$ bilinear operator, and E is a $(p \times q \times p)$ bilinear operator. Then

$$\underline{g}'(\underline{x},k) = F + (D + D^{*})\underline{x} + E^{*}\underline{u}(k), \tag{4.40}$$

and

$$\underline{g}''(\underline{x},k) = D + D^{*} \tag{4.41}$$

We now give an example to illustrate the algorithm of Theorem 4.4. Consider the parametrized system which has already been given in Sec. 2.3.

$$\begin{bmatrix} w_1(k+1) \\ w_2(k+1) \end{bmatrix} = \begin{bmatrix} \eta_1 & 1 \\ 0 & \eta_2 \end{bmatrix} \begin{bmatrix} w_1(k) \\ w_2(k) \end{bmatrix} + \begin{bmatrix} 0 \\ \eta_3 \end{bmatrix} u(k)$$

$$\underline{y}(k) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} w_1(k) \\ w_2(k) \end{bmatrix}, \qquad k = 0,1,2,\text{---} \tag{4.42}$$

The block diagram of this system is depicted in Fig. 2. We note that by reversing the positions of $\frac{1}{z-\eta_2}$ and $\frac{1}{z-\eta_1}$, and by transforming the initial conditions by a similarity transformation, the resulting equivalent system will generate the same output sequence. The equivalent transformation matrix is given by

$$P = \begin{bmatrix} 1 & 0 \\ \eta_1 - \eta_2 & 1 \end{bmatrix}$$

which has already been obtained in Sec. 2.3. The parameter $\underline{\theta} = [w_1(0), w_2(0), \eta_1, \eta_2, \eta_3]^T$ is to be identified. Letting $\underline{x}(k) = [w_1(k), w_2(k), \eta_1, \eta_2, \eta_3]^T$, this system can be written in the quadratic-in-the-state bilinear system form as given in (4.38) if we let

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

$$G = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

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

$$E = \begin{bmatrix} 0\,0\,0\,0\,0 \\ 0\,0\,0\,0\,1 \\ 0\,0\,0\,0\,0 \\ 0\,0\,0\,0\,0 \\ 0\,0\,0\,0\,0 \end{bmatrix},$$

and

$$C = [1\ 0\ 0\ 0\ 0]$$

We have used the notation in [14] for representing the bilinear form D, and E degenerated to an ordinary matrix since u(k) is a scalar.

Let $\underline{\theta} = [w_1(0),\ w_2(0),\ \eta_1,\ \eta_2,\ \eta_3]^T$. It can be easily seen that $\underline{\theta}^* = [w_1(0),\ w_2(0) + (\eta_1 + \eta_2)w_1(0),\ \eta_2,\ \eta_1,\ \eta_3]^T$ is a point in Ω which is unresolvable from $\underline{\theta}$ by the equivalent transformation. Hence the solution to the equation $\underline{f}_M(\underline{\theta}) = 0$ is not unique in R^5 and only local identifiability can be imposed on the parameters of systems (4.42). We applied the algorithm of Theorem 4.4 to this example with $\underline{\theta}_0 = [0.5,\ 0.1,\ 0.3,\ 0.7,\ 1.0]^T$. We note that $\underline{\theta}_1 = [0.5,\ -0.1,\ 0.7,\ 0.3,\ 1.0]^T$ will give exactly the same set of $\underline{y}(k)$ sequence for any input sequence u(k) and hence is unresolvable from $\underline{\theta}_0$. However, the distance from $\underline{\theta}_0$ to $\underline{\theta}_1$ is 0.4 by using the ℓ_∞-norm. Therefore $\underline{\theta}_0$ and $\underline{\theta}_1$ are each locally identifiable in the

spheres $S(\underline{\theta}_0,\rho)$ and $S(\underline{\theta}_1,\rho)$ respectively where $\rho < 0.4$. By applying Theorem 4.4 and the algorithm developed in the above section, we are able to find two spheres centered at $\underline{\theta}_0$ and $\underline{\theta}_1$ respectively such that $\underline{\theta}_0$ and $\underline{\theta}_1$ are locally identifiable in those two spheres respectively. To demonstrate this situation, we use an input sequence $u(k) = 100 \sin(k\pi/4)$, for $k = 0,1,---, 19$, we found that with twenty observation, the sphere centered at $\underline{\theta}_0$ is of radius $\rho = 0.21$ and

$$\lambda = 0.1314 \times 10^{-7}$$

$$\eta = 0.3646 \times 10^{7}$$

$$\gamma = 0.2284 \times 10^{8}$$

$$\mu = 0.2124 \times 10^{7}$$

$$\lambda\eta = 0.0479 < 0.0525 = \rho/4$$

$$\lambda(2\mu\rho + \gamma) = 0.312 < 0.5 \ .$$

Thus the hypotheses of Theorem 4.4 are satisfied, and we conclude that $\underline{\theta}_0$ is locally identifiable in $S(\underline{\theta}_0, 0.21)$ and the Gauss-Newton sequence will converge to $\underline{\theta}_0$ starting from any point in the sphere $S(\underline{\theta}_0, 0.105)$. We also tested the conditions of Theorem 4.4 about the point $\underline{\theta}_1$ and found that $\rho = 0.31$. Hence $\underline{\theta}_1$ is locally identifiable in $S(\underline{\theta}_1, 0.31)$ and the Gauss-Newton sequence will converge to $\underline{\theta}_1$ starting from any point in the sphere $S(\underline{\theta}_1, 0.155)$.

Since Theorem 4.4 gives sufficient conditions for convergence, and because of the upper bounding implied by the use of interval analysis, the question of whether these results are overly conservative naturally arises.

However, this example illustrates that the computed sphere of convergence is of reasonable size in view of the distance from $\underline{\theta}_0$ to $\underline{\theta}_1$, i.e., the theoretical radius of the region of identifiability.

The computer program for testing the conditions in Theorem 4.4 for the above specific example is listed in the Appendix.

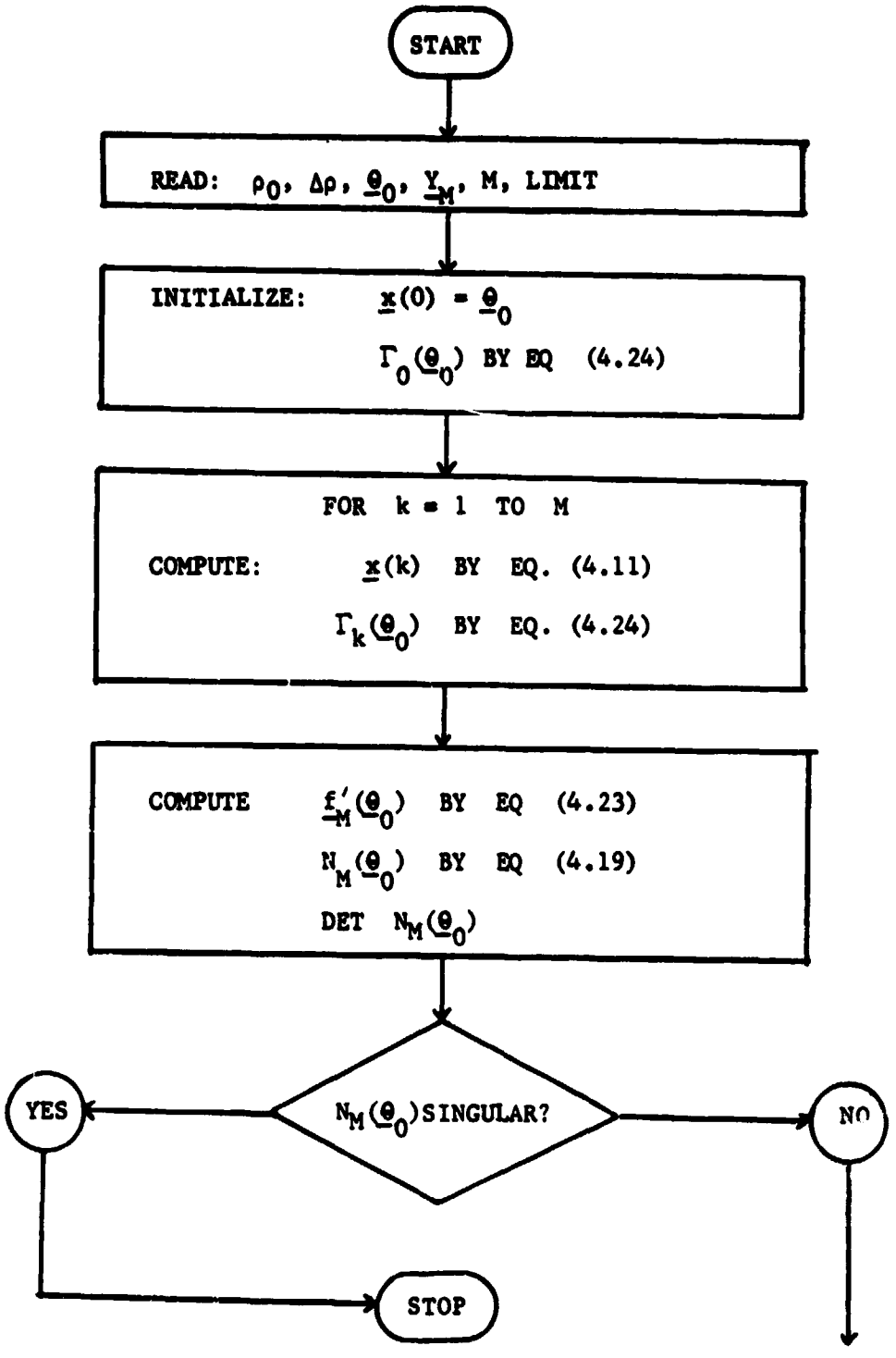

Figure 1. Flow Graph for Computing Regions of Parameter Identifiability

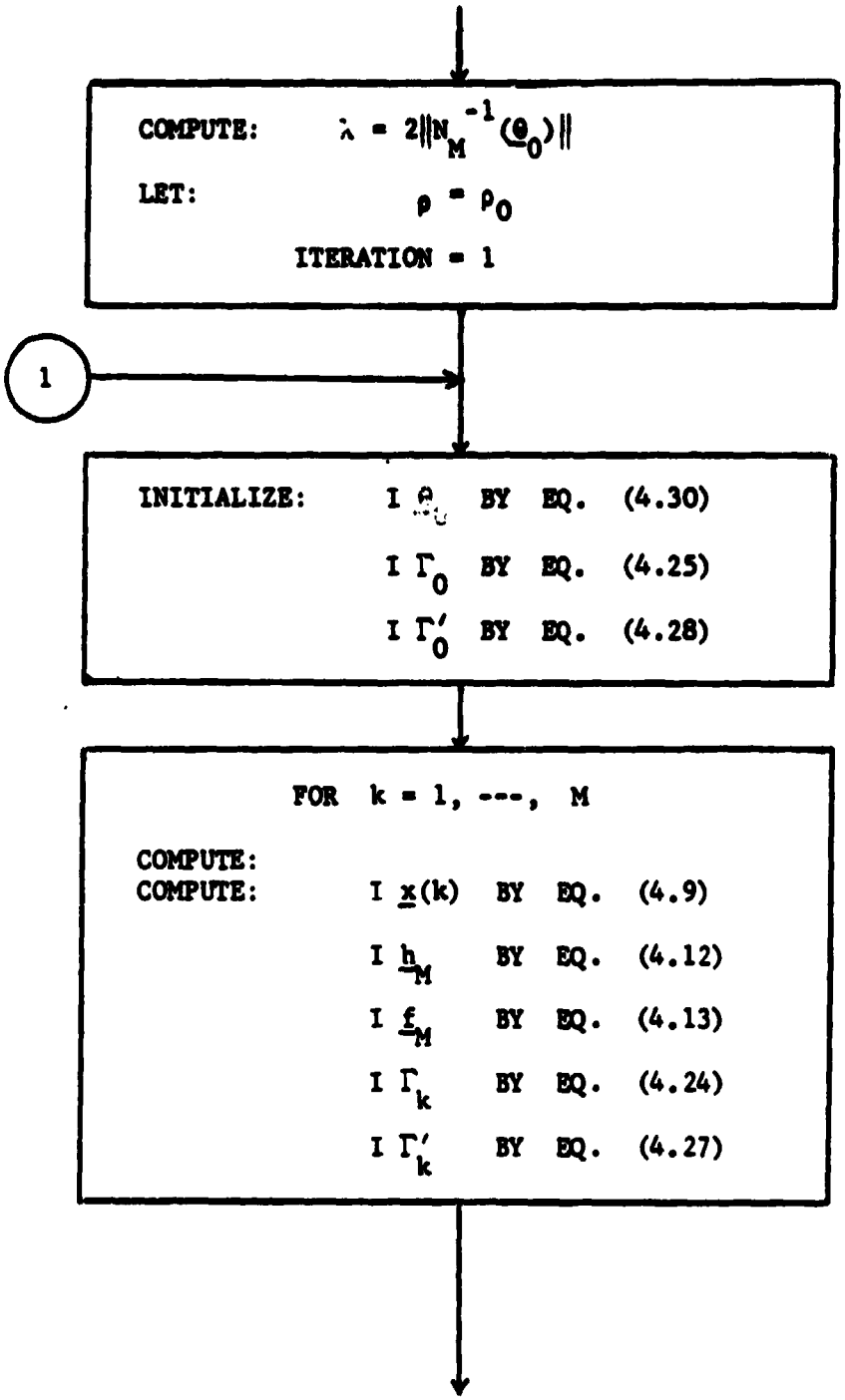

Figure 1. Continued

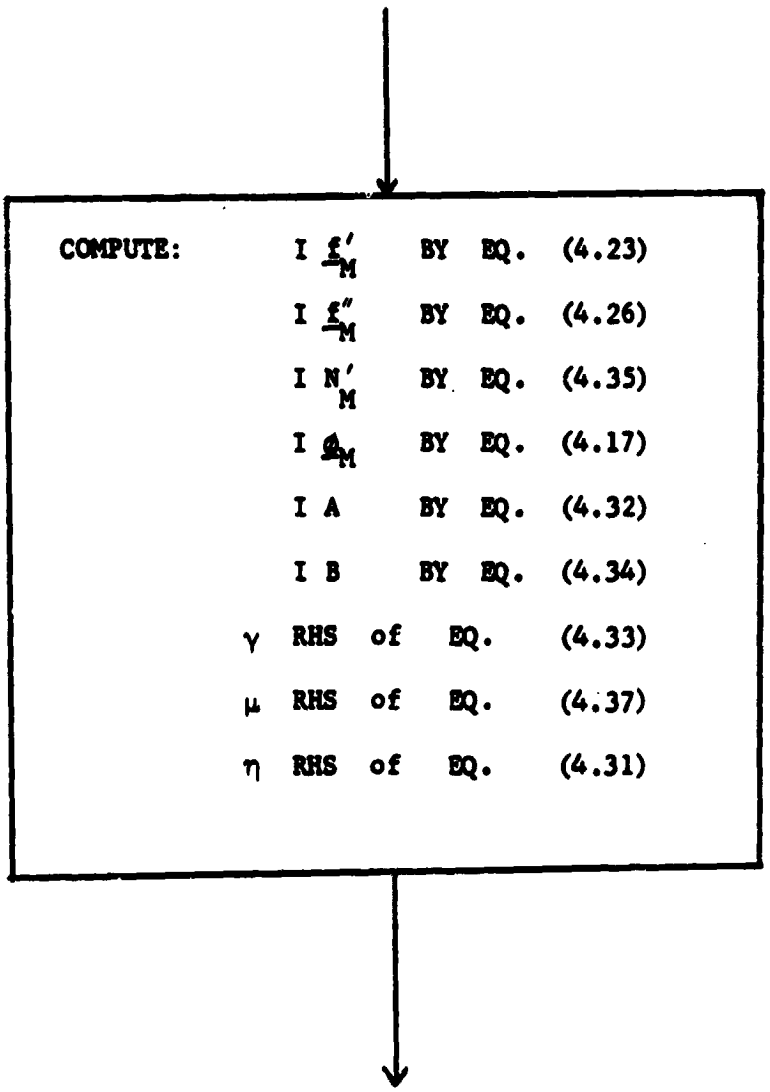

Figure 1. Continued

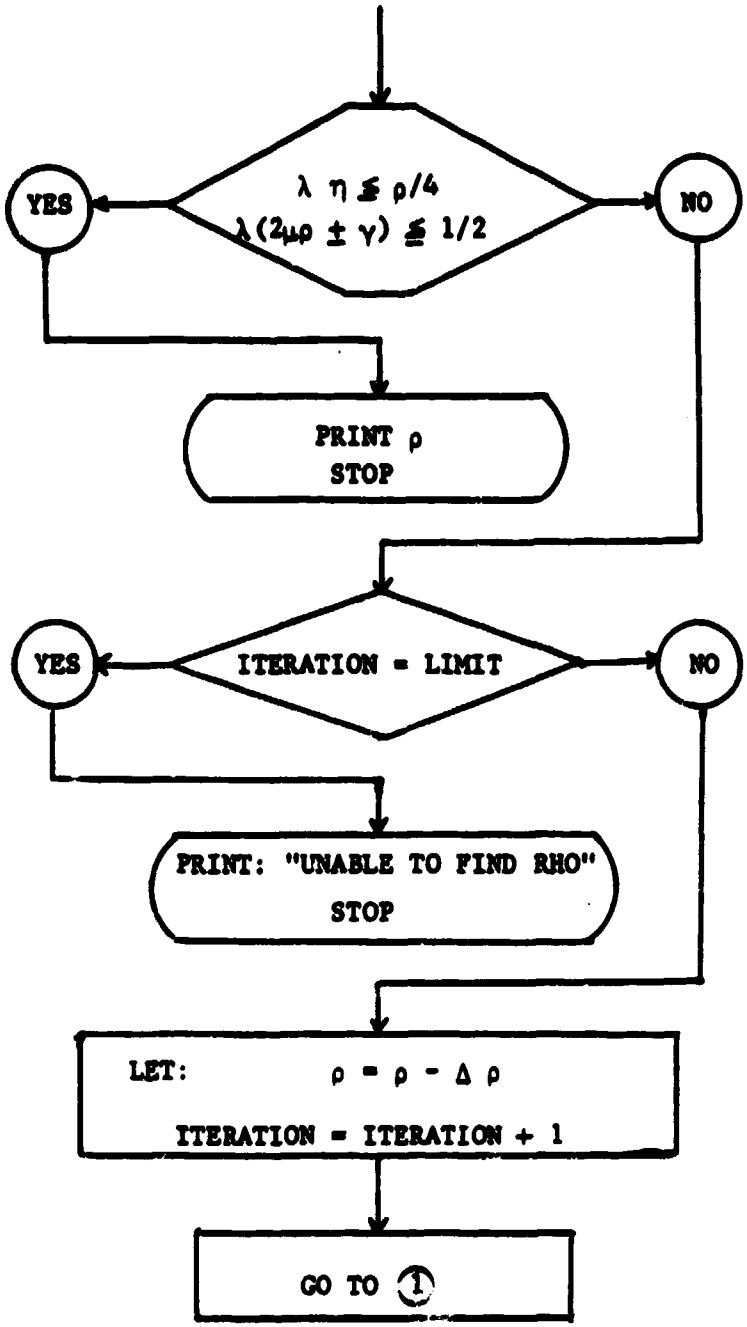

Figure 1. Continued

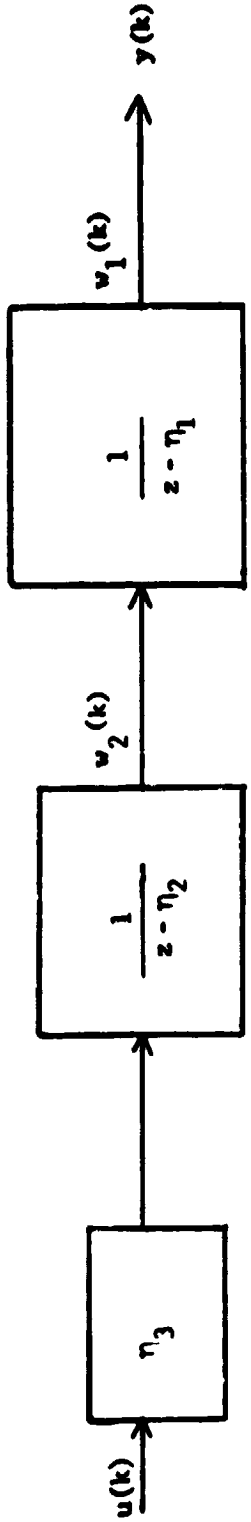

Figure 2. Block Diagram of System (4.42)

5. LOCAL PARAMETER IDENTIFIABILITY AND LOCAL CML PARAMETER IDENTIFIABILITY OF DYNAMIC SYSTEMS WITH MEASUREMENT NOISE

In this chapter we will study the parameter identification problem of general discrete-time, multiple-input/multiple-output dynamic systems with measurement noise. Specifically, we will establish the concept of local parameter identifiability and provide a computation procedure for finding the explicit regions of parameter identifiability. Moreover, we will show that the constrained maximum likelihood estimation sequence converges to the locally identifiabile parameters with probability one.

5.1 Local Parameter Identifiability and Local CML Parameter Identifiability

Let $\{\underline{z}_i\}_{i=0}^{\infty}$ be a sequence of random vectors, which we shall call the observations, with joint probability density function $p(\underline{z}_0,---,\underline{z}_M;\underline{\theta})$, $M=0,1,2,---$, parameterized by the unknown parameter $\underline{\theta}\in\Omega\subset R^p$. R^p is the space of real p-tuples with an arbitrary norm denoted by $\|\cdot\|$, and Ω is a compact subset of R^p. We let $\underline{Z}_M=(\underline{z}_0,\underline{z}_1,---,\underline{z}_M)$, and we assume that the true parameter $\underline{\theta}_0$ lies in the interior of Ω. Furthermore, we assume that $p(\underline{Z}_M;\underline{\theta})$ is continuous with respect to $\underline{\theta}\in\Omega$ for $\underline{Z}_M$ almost everywhere, i.e., for $\epsilon>0$ and $\underline{\theta}\in\Omega$, there exists a $\delta(\epsilon)>0$ such that for all $\underline{\theta}'\in\Omega$ with $\|\underline{\theta}-\underline{\theta}'\|<\delta$ we have $|p(Z_M;\underline{\theta})-p(Z_M;\underline{\theta}')|<\epsilon$ for Z_M almost everywhere. Following Tse and Anton [19], we make the following definitions.

Definition 5.1

Two parameters $\underline{\theta}_1,\underline{\theta}_2\in\Omega$, $\underline{\theta}_1\neq\underline{\theta}_2$, are said to be unresolvable if the equality

$$p(Z \quad) = p(\underline{Z};\underline{\theta}_2) \tag{5.1}$$

holds with probability one with respect to $\underline{\theta}_1$ and $\underline{\theta}_2$ for all except a finite number of integers $n>0$, i.e., for all except a finite number of integers $n>0$, (5.1) holds with respect to the measure $p(\underline{Z}_M;\underline{\theta}_1)d\underline{Z}_M$ as well as $p(\underline{Z}_M;\underline{\theta}_2)d\underline{Z}_M$.

Definition 5.2

A parameter $\underline{\theta}_0 \in \Omega$ is said to be locally identifiable if there exists an open sphere $S(\underline{\theta}_0,\rho) \subset \Omega$ with radius $\rho>0$ centered at $\underline{\theta}_0$ such that there is no other $\underline{\theta} \in S(\underline{\theta}_0,\rho)$, $\underline{\theta} \neq \underline{\theta}_0$, which is unresolvable from $\underline{\theta}_0$.

The problem of constrained maximum likelihood (CML) estimation is as follows: find $\hat{\underline{\theta}}_M \in \Omega$, an estimate of the true parameter $\underline{\theta}_0$, where $\hat{\underline{\theta}}_M$ is such that

$$p(Z_M;\hat{\underline{\theta}}_M) = \underset{\underline{\theta}\in\Omega}{\text{Max}}\ p(Z_M;\underline{\theta}), \quad M=0,1,2,\text{---} \tag{5.2}$$

Since Ω is closed and bounded and $p(\underline{Z}_M;\underline{\theta})$ is continuous on Ω for $\underline{Z}_M$ almost everywhere, a solution to (5.2) exists. However, if $\underline{\theta}_1$ and $\underline{\theta}_2$ are unresolvable in Ω, then they cannot be identified by the CML estimation method constrained to Ω. Therefore the following definition is established.

Definition 5.3

A parameter $\underline{\theta}_0 \in \Omega$ is said to be locally CML identifiable if there exists an open sphere $S(\underline{\theta}_0,\rho)$ with radius $\rho>0$ centered at $\underline{\theta}_0$ such that the sequence $\{\hat{\underline{\theta}}_M\}_{M=0}^{\infty}$ converges to $\underline{\theta}_0$ with probability one, where $\{\hat{\underline{\theta}}_M\}_{M=0}^{\infty}$ is constructed by

$$p(\underline{Z}_M;\hat{\underline{\theta}}_M) = \max_{\underline{\theta}\in\bar{S}(\underline{\theta}_0,\rho)} p(\underline{Z}_M;\underline{\theta}), \quad M = 0,1,2,\cdots \tag{5.3}$$

$\bar{S}(\underline{\theta}_0,\rho)$ denotes the closure of $S(\underline{\theta}_0,\rho)$. We will call $S(\underline{\theta}_0,\rho)$ given above the region of parameter identifiability.

We now consider the problem of identifying the vector of unknown parameters, $\underline{\eta}$, of a system whose state at time k is the vector $\underline{w}(k)$ where $k = 0,1,2,\cdots$. Let $\underline{x}(k) = \begin{bmatrix} \underline{w}(k) \\ \underline{\eta} \end{bmatrix}$ and assume $\underline{x}(k)$ is a p-component vector which we shall call the augmented state vector. Let the initial state $\underline{w}(0)$ also be unknown, then $\underline{\theta} = \underline{x}(0) = \begin{bmatrix} \underline{w}(0) \\ \underline{\eta} \end{bmatrix}$ is the parameter vector to be identified. The identification problem is then equivalent to the initial-state observation problem of the augmented system. We assume that the function $\underline{g}(\underline{x},k)$ is known as a function of $\underline{x}$ and k and that

$$\underline{x}(k+1) = \underline{g}(\underline{x}(k),k) \qquad k = 0,1,2,\cdots. \tag{5.4}$$

Observations of the state are taken which we will assume can be written in the form

$$\begin{aligned} \underline{y}(k) &= C\underline{x}(k) \\ \underline{z}(k) &= \underline{y}(k) + \underline{v}(k) \qquad k = 0,1,2,\cdots. \end{aligned} \tag{5.5}$$

where C is a known $(r \times p)$ matrix and $\underline{v}(k)$ is a Gaussian-white noise vector with r components which has zero mean and covariane matrix $Q = \sigma^2 I_{r\times r}$, $\sigma^2 < \infty$. The observation sequence of system (5.4) and (5.5) is

$$\underline{Z}_M = [\underline{z}^T(0),\underline{z}^T(1),\cdots,\underline{z}^T(M)]^T$$

and the joint probability density function is given by

$$p(\underline{Z}_M;\underline{\theta}) = \text{Const.}\cdot\exp\left[-\frac{1}{2}\sum_{k=0}^{M}(\underline{z}(k) - C\underline{x}(k))^T Q^{-1}(\underline{z}(k) - C\underline{x}(k))\right] \quad (5.6)$$

where $\underline{x}(k)$ is the solution to (5.4) when $\underline{x}(0) = \underline{\theta}$, and $M = 0,1,2,---$.

We assume that the structure of $\underline{g}(\underline{x},k)$ is such that the following assumptions hold.

(1) The inputs to the system are known implicitly in $\underline{g}(\underline{x},k)$.

(2) For every admissible input sequence and all $\underline{\theta}\in\Omega$, the states $\underline{x}(k)$ and the deterministic part of the observations, $\underline{w}(k)$, generated by $\underline{\theta}$ when applied with the input sequence are bounded.

(3) $\underline{g}(\underline{x},k)$ is at least twice continuously differentiable with respect to $\underline{\theta}$, and hence so is $p(\underline{Z}_M;\underline{\theta})$.

5.2 Regions of Parameter Identifiability

Let the assumptions given in Section 5.1 hold, and let us define

$$L_M(\underline{\theta}) = \frac{1}{M+1}\sum_{k=0}^{M}[\underline{z}(k) - C\underline{x}(k)]^T[\underline{z}(k) - C\underline{x}(k)]$$
$$\text{for } M = 0,1,2,---. \quad (5.7)$$

Then the CML estimation method in Eq. (2) is equivalent to finding

$\underset{\underline{\theta}\in\Omega}{\text{Min}}\ L_M(\underline{\theta}) = L_M(\hat{\underline{\theta}}_M)$. $M = 0,1,2,---$.

Following the notation in Chapter 4, we let

$$\underline{h}_M(\underline{\theta}) = \begin{bmatrix} C\underline{x}(0) \\ \cdot \\ \cdot \\ \cdot \\ C\underline{x}(M) \end{bmatrix} \quad (5.8)$$

be the m-component vector where $m = r(M+1) \geq p$ and $\underline{x}(k)$ is the solution to equation (4) when $\underline{x}(0) = \underline{\theta}$.

Let

$$\underline{Y}_M = \begin{bmatrix} \underline{y}(0) \\ \vdots \\ \underline{y}(M) \end{bmatrix}, \tag{5.9}$$

$$\underline{V}_M = \begin{bmatrix} \underline{v}(0) \\ \vdots \\ \underline{v}(M) \end{bmatrix}, \tag{5.10}$$

$$\underline{f}_M(\underline{\theta}) = \underline{h}_M(\underline{\theta}) - \underline{Y}_M, \tag{5.11}$$

and

$$\underline{s}_M(\underline{\theta}) = \underline{h}_M(\underline{\theta}) - \underline{Z}_M. \tag{5.12}$$

Then

$$L_M(\underline{\theta}) = \frac{1}{M+1}\, \underline{s}_M^T(\underline{\theta})\underline{s}_M(\underline{\theta}). \tag{5.13}$$

Let

$$\underline{\phi}_M^*(\underline{\theta}) = \frac{1}{2}\, L_M'(\underline{\theta}).$$

Then

$$\underline{\phi}_M^*(\underline{\theta}) = \frac{1}{M+1}[\underline{s}_M'(\underline{\theta})]^T \underline{s}_M(\underline{\theta}), \tag{5.14}$$

and

$$\underline{\phi}_M^{*\prime}(\underline{\theta}) = \frac{1}{M+1}\{[\underline{s}_M'(\underline{\theta})]^T[\underline{s}_M'(\underline{\theta})] + [\underline{s}_M''(\underline{\theta})]^T \underline{s}_M(\underline{\theta})\}. \tag{5.15}$$

We note that

$$\underline{s}_M'(\underline{\theta}) = \underline{f}_M'(\underline{\theta}) \tag{5.16}$$

and

$$\underline{s}_M''(\underline{\theta}) = \underline{f}_M''(\underline{\theta}). \tag{5.17}$$

Let

$$N_M^*(\underline{\theta}) = \frac{1}{M+1}[\underline{f}_M'(\underline{\theta})]^T[\underline{f}_M'(\underline{\theta})] \tag{5.18}$$

By eqs. (5.5), (5.11), (5.12), and (5.13),

$$L_M(\underline{\theta}) = \frac{1}{M+1}[\underline{f}_M(\underline{\theta}) - \underline{v}_M]^T[\underline{f}_M(\underline{\theta}) - \underline{v}_M]$$

or

$$J_M(\underline{\theta}) = \frac{1}{M+1}\underline{f}_M^T(\underline{\theta})\underline{f}_M(\underline{\theta}) - \frac{2}{(M+1)}[\underline{f}_M(\underline{\theta})]^T\underline{v}_M + \frac{1}{M+1}\underline{v}_M^T\underline{v}_M \tag{5.19}$$

Let us define

$$L^*(\underline{\theta}) = \lim_{M\to\infty} E[L_M(\underline{\theta})].$$

The following Theorem is a generalization of Aoki and Yue's Theorem given in [1].

Theorem 5.1

Let the assumptions given in Section 5.1 hold.

Then $\lim_{M\to\infty} L_M(\underline{\theta}) = L^*(\underline{\theta})$ for all $\underline{\theta}\in\Omega$ with probability one. Furthermore,

$$L^*(\underline{\theta}) = \lim_{M\to\infty} \frac{1}{M+1}[\underline{f}_M(\underline{\theta})]^T\underline{f}_M(\underline{\theta}) + r\sigma^2. \tag{5.20}$$

Proof. By assumption (2), $\underline{h}_M(\underline{\theta})$ has bounded elements for all $\underline{\theta}\in\Omega$ and $\underline{Y}_M$ has bounded elements. Hence $\underline{f}_M(\underline{\theta}) = \underline{h}_M(\underline{\theta}) - \underline{Y}_M$ has bounded elements for all

$\underline{\theta} \in \Omega$. Let

$$\underline{f}_M(\underline{\theta}) = \begin{bmatrix} f_1(\underline{\theta}) \\ f_2(\underline{\theta}) \\ \vdots \\ f_m(\underline{\theta}) \end{bmatrix} .$$

Then there exists a $K < \infty$ such that $|f_k(\underline{\theta})| \leq K$ for all $\underline{\theta} \in \Omega$ and for all k. Thus

$$\lim_{M \to \infty} \frac{1}{M+1} [\underline{f}_M(\underline{\theta})]^T \underline{f}_M(\underline{\theta}) = \lim_{M \to \infty} \frac{1}{M+1} \sum_{k=1}^{m} f_k^2(\underline{\theta})$$

$$\leq \lim_{M \to \infty} \frac{r(M+1)}{(M+1)} K^2 = rK^2. \qquad (5.21)$$

Therefore the limit in (5.20) exists.

Consider the second term on the right hand side of (5.19). Let

$$\underline{V}_M = \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_m \end{bmatrix}$$

Then $\frac{1}{M+1} [\underline{f}_M(\underline{\theta})]^T \underline{V}_M = \frac{1}{M+1} \sum_{k=1}^{m} f_k(\underline{\theta}) V_k$.

Since $EV_k = 0$, $EV_k^2 = \sigma^2$ for all k, applying Lemma 2.1.a to our situation, we have $x_k = f_k(\underline{\theta}) V_k$, $b_M = M$, and

$$\sum_{k=1}^{\infty} \frac{Ex_k^2}{k^2} = \sum_{k=1}^{\infty} \frac{f_k^2(\underline{\theta}) EV_k^2}{k^2} \leq \sigma^2 K^2 \sum_{k=1}^{\infty} \frac{1}{k^2} < \infty$$

Thus

$$\frac{1}{M+1}\sum_{k=1}^{m} x_k = \frac{1}{M+1}\sum_{k=1}^{m} f_k(\underline{\theta})V_k \to 0 \quad \text{with probability one.} \qquad (5.22)$$

Consider the third term on the right hand side of (5.19).

$$\frac{1}{M+1}\underline{V}_M^T\underline{V}_M = \frac{1}{M+1}\sum_{k=1}^{m} V_k^2$$

Applying Lemma 2.1.b to this situation, we have $x_k = V_k^2$, $E|x_1| = EV_1^2 = \sigma^2 < \infty$. Thus

$$\frac{1}{M+1}\sum_{k=1}^{m} V_k^2 \to r\sigma^2 \quad \text{with probability one.} \qquad (5.23)$$

By (5.21), (5.22), and (5.23), the theorem is proved.

If $\underline{\theta}$ is the true parameter, then $L^*(\underline{\theta}_0) = r\sigma^2 = \min L^*(\underline{\theta})$. The following theorem provides sufficient conditions for the existence of a sphere $S(\underline{\theta}_0,\rho)$ such that $\underline{\theta}_0$ is the unique minimal point of $J^*(\underline{\theta})$ for all $\underline{\theta} \in S(\underline{\theta}_0,\rho)$. The form of the theorem was motivated by the work of Pereyra [13].

We first recall that the ℓ_2-norm, or the Euclidean norm on the R^n space is defined by

$$\|\underline{x}\|_2 = \left(\sum_{i=1}^{n} |x_i|^2\right)^{\frac{1}{2}}, \quad \text{for all } \underline{x} \in R^n.$$

Theorem 5.2

Let $\|\cdot\|_2$ denotes the Euclidean norm. For brevity, let

$$A = [a_{ij}] = \frac{1}{\sqrt{M+1}}\underline{f}'_M(\underline{\theta}).$$

Note that $A^TA = N_M(\underline{\theta})$ which is an $m \times p$ matrix.

If (1) there is a $\underline{\theta}_0 \in \Omega$ such that $\underline{f}_M(\underline{\theta}_0) = 0$ and hence

$$\lim_{M\to\infty} \frac{1}{M+1} \underline{f}_M^T(\underline{\theta}_0)\underline{f}_M(\underline{\theta}_0) = 0,$$

(2) $N_M^*(\underline{\theta})$ is nonsingular for all $M \in I^+$, and

(3) there is a sphere $S(\underline{\theta}_0,\rho) \subset \Omega$ with radius $\rho > 0$ centered at $\underline{\theta}_0$ and a set of positive numbers (λ,γ,μ,K) such that for all $M \in I^+$

(i) $\|N_M^{-1}(\underline{\theta}_0)\|_2 \le \lambda$,

(ii) $\{\sum_{i=1}^{p} \sum_{j=1}^{p} ([A^TA]_{ij})^2\}^{\frac{1}{2}} \le K$ for all $\underline{\theta} \in S(\underline{\theta}_0,\rho)$,

(iii) $\|\frac{1}{M+1}[\underline{f}''_M(\underline{\theta})]\|_2 \le \gamma$ for all $\underline{\theta} \in S(\underline{\theta}_0,\rho)$,

(iv) $\|N_M^*(\underline{\theta})\|_2 \le \mu$ for all $\underline{\theta} \in S(\underline{\theta},\rho)$, and

(v) $\lambda(\mu\rho+\gamma) < 1$.

Then (a) $\underline{\theta}_0$ is the unique point in $S(\underline{\theta}_0,\rho)$ such that $\underline{f}_M(\underline{\theta}_0) = 0$ for all $M \in I^+$ and is the unique minimal point for $\lim_{M\to\infty} \frac{1}{M+1} \underline{f}_M^T(\underline{\theta})\underline{f}_M(\underline{\theta})$ in $S(\underline{\theta}_0,\rho)$. I.e., $\underline{\theta}_0$ is the unique point in $S(\underline{\theta}_0,\rho)$ such that $L^*(\underline{\theta}) = r\sigma^2 = L^*(\underline{\theta}_0)$.

(b) $\underline{\theta}_0$ is the unique minimal point for $\lim_{M\to\infty} L_M(\underline{\theta})$ in $S(\underline{\theta}_0,\rho)$ with probability one.

Proof. A result due to Bartle in Herget [7] is: If $\|\underline{\phi}_M^{*\prime}(\underline{\theta}) - \underline{\phi}_M^{*\prime}(\underline{\theta}_0)\|_2 \le \beta$ for all $\underline{\theta} \in S(\underline{\theta}_0,\rho)$, then $\|\underline{\phi}_M^*(\underline{\theta}_1) - \underline{\phi}_M^*(\underline{\theta}_2) - \underline{\phi}_M^{*\prime}(\underline{\theta}_0)(\underline{\theta}_1 - \underline{\theta}_2)\|_2 \le \beta\|\underline{\theta}_1 - \underline{\theta}_2\|_2$ for all $\underline{\theta}_1,\underline{\theta}_2 \in S(\underline{\theta}_0,\rho)$. In our particular case, $\underline{\phi}_M^{*\prime}(\underline{\theta}) = N_M^*(\underline{\theta}) + \frac{1}{M+1}[\underline{f}''_M(\underline{\theta})]^T\underline{f}_M(\underline{\theta})$ and $\underline{\phi}_M^{*\prime}(\underline{\theta}_0) = N_M^*(\underline{\theta}_0)$.

Thus $\|\underline{\phi}_M^{*\prime}(\underline{\theta}) - \underline{\phi}_M^{*\prime}(\underline{\theta}_0)\|_2 \le \mu\rho+\gamma$ for all $\underline{\theta} \in S(\underline{\theta}_0,\rho)$ by 3(iii) and (iv). Therefore we have

$$\|\underline{\phi}_M^*(\underline{\theta}_1) - \underline{\phi}_M^*(\underline{\theta}_2) - \underline{\phi}_M^{*\prime}(\underline{\theta}_0)(\underline{\theta}_1 - \underline{\theta}_2)\|_2 \leq (\mu\rho + \gamma)\|\underline{\theta}_1 - \underline{\theta}_2\|_2$$

for all $\underline{\theta}_1, \underline{\theta}_2 \in S(\underline{\theta}_0, \rho)$.

Now suppose there is another point $\underline{\theta}_1 \in S(\underline{\theta}_0, \rho)$, $\underline{\theta}_1 \neq \underline{\theta}_2$, such that

$$\lim_{M\to\infty} \frac{1}{M+1} \underline{f}_M^T(\underline{\theta}_1)\underline{f}_M(\underline{\theta}_1) = 0,$$

i.e.

$$\lim_{M\to\infty} \left\|\frac{1}{\sqrt{M+1}} \underline{f}_M(\underline{\theta}_1)\right\|_2^2 = 0.$$

We can write

$$\begin{aligned}
\underline{\theta}_0 - \underline{\theta}_1 &= N_M^{*-1}(\underline{\theta}_0) N_M^*(\underline{\theta}_0)(\underline{\theta}_0 - \underline{\theta}_1) \\
&= N_M^{*-1}(\underline{\theta}_0)\,[\underline{\phi}_M^*(\underline{\theta}_1) - \underline{\phi}_M^*(\underline{\theta}_0)(\underline{\theta}_1 - \underline{\theta}_0) \\
&\qquad + [N_M^*(\underline{\theta}_0) - \underline{\phi}_M^{*\prime}(\underline{\theta}_0)](\underline{\theta}_0 - \underline{\theta}_1) \\
&\qquad - \underline{\phi}_M^*(\underline{\theta}_1) + \underline{\phi}_M^*(\underline{\theta}_0)] \quad \text{for all } M \in I^+.
\end{aligned}$$

We note that

$\underline{\phi}_M^{*\prime}(\underline{\theta}_0) = N_M^*(\underline{\theta}_0)$ and $\underline{\phi}_M^*(\underline{\theta}_0) = 0$. Therefore

$$\begin{aligned}
\|\underline{\theta}_0 - \underline{\theta}_1\|_2 &= \|N_M^{*-1}(\underline{\theta}_0)[\underline{\phi}_M^*(\underline{\theta}_1) - \underline{\phi}_M^*(\underline{\theta}_0)(\underline{\theta}_1 - \underline{\theta}_0) - \underline{\phi}_M^*(\underline{\theta}_1)]\|_2 \\
&\leq \|N_M^{*-1}(\underline{\theta}_0)\|_2 \|\underline{\phi}_M^*(\underline{\theta}_1) - \underline{\phi}_M^*(\underline{\theta}_0) - \underline{\phi}_M^{*\prime}(\underline{\theta}_0)(\underline{\theta}_1 - \underline{\theta}_0)\|_2 \\
&\quad + \|N_M^{*-1}(\underline{\theta}_0)\|_2 \|\underline{\phi}_M^*(\underline{\theta}_1)\|_2 \\
&\leq \lambda(\mu\rho + \gamma)\|\underline{\theta}_1 - \underline{\theta}_0\|_2 + \lambda\|\underline{\phi}_M^*(\underline{\theta}_1)\|_2 \\
&< \|\underline{\theta}_0 - \underline{\theta}_1\|_2 + \lambda\|\underline{\phi}_M^*(\underline{\theta}_1)\|_2 \quad \text{for all } M \in I^+.
\end{aligned}$$

We now want to show that

$$\lim_{M\to\infty} \|\underline{\phi}_M^*(\underline{\theta}_1)\|_2 = 0,$$

But $\|\underline{\phi}_M^*(\underline{\theta}_1)\|_2^2 = \underline{\phi}_M^{*T}(\underline{\theta}_1)\underline{\phi}_M^*(\underline{\theta}_1)$

$$= \frac{1}{(M+1)^2}\{[\underline{f}_M'(\underline{\theta}_1)]^T\underline{f}_M(\underline{\theta}_1)\}^T\{[\underline{f}_M'(\underline{\theta}_1)\}$$

$$= \frac{1}{(M+1)^2}\underline{f}_M^T(\underline{\theta}_1)[\underline{f}_M'(\underline{\theta}_1)][\underline{f}_M'(\underline{\theta}_1)]^T\underline{f}_M(\underline{\theta}_1)$$

$$\le \left[\frac{1}{(M+1)}\|[\underline{f}_M'(\underline{\theta}_1)][\underline{f}_M'(\underline{\theta}_1)]^T\|_2\right]\left[\frac{1}{M+1}\underline{f}_M^T(\underline{\theta}_1)\underline{f}_M(\underline{\theta}_1)\right].$$

The final inequality follows easily from Schwarz inequality and properties of a norm, i.e. $|\underline{x}^TA\underline{x}| \le \|\underline{x}\|_2\|A\underline{x}\|_2 \le \|A\|_2\cdot\|\underline{x}\|_2^2$.

We now state a result given in [14]. If B is a p x p matrix, $B = [b_{ij}]$, then $\|B\|_2 \le \left(\sum_{i=1}^{p}\sum_{j=1}^{p}(b_{ij})^2\right)^{\frac{1}{2}}$.

Now look at

$$\frac{1}{M+1}[\underline{f}_M'(\underline{\theta}_1)][\underline{f}_M'(\underline{\theta}_1)]^T$$

$$= \left[\frac{1}{\sqrt{M+1}}\underline{f}_M'(\underline{\theta}_1)\right]\left[\frac{1}{\sqrt{M+1}}\underline{f}_M'(\underline{\theta}_1)\right]^T$$

$$= AA^T$$

where $[AA^T]_{ij} = \sum_{k=1}^{p} a_{ik}a_{jk}$.

Thus $\|AA^T\|_2 \leq \{\sum_{i=1}^{m}\sum_{j=1}^{m}(\sum_{k=1}^{p}a_{ik}a_{jk})^2\}^{\frac{1}{2}}$

$$= \{\sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{k=1}^{p}\sum_{\ell=1}^{p}a_{ik}a_{jk}a_{i\ell}a_{j\ell}\}^{\frac{1}{2}}$$

$$= \{\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{k=1}^{m}\sum_{\ell=1}^{m}a_{ki}a_{\ell j}a_{\ell i}a_{kj}\}^{\frac{1}{2}}$$

$$= \{\sum_{i=1}^{p}\sum_{j=1}^{p}(\sum_{k=1}^{m}a_{ki}a_{kj})^2\}^{\frac{1}{2}}$$

$$= \{\sum_{i=1}^{p}\sum_{j=1}^{p}([A^TA]_{ij})^2\}^{\frac{1}{2}}$$

$$\leq K \quad \text{for all } \underline{\theta}\in S(\underline{\theta}_0,\rho) \text{ and } M\in I^+ \text{ by 3(ii).}$$

Therefore

$$\|\underline{\phi}_M^*(\underline{\theta}_1)\|_2^2 \leq K[\frac{1}{M+1}\underline{f}_M^T(\underline{\theta}_1)\underline{f}_M(\theta_1)] = K\|\frac{1}{\sqrt{M+1}}\underline{f}_M(\underline{\theta}_1)\|_2^2.$$

Thus $\lim_{M\to\infty}\|\underline{\phi}_M^*(\underline{\theta}_1)\|_2^2 = 0$, and we conclude that

$$\|\underline{\theta}_0 - \underline{\theta}_1\| < \|\underline{\theta}_0 - \underline{\theta}_1\|$$

which is a contradiction, and so we conclude $\underline{\theta}_1 = \underline{\theta}_0$. This completes the proof of part a.

To prove part (b), we see that

$$L_M(\underline{\theta}) \to L^*(\underline{\theta}) \quad \text{with probability one for all } \underline{\theta}\in\Omega$$

and $L^*(\underline{\theta}) < \infty$ for all $\underline{\theta}\in\Omega$ by Theorem 5.1.

Thus

$$L_M(\underline{\theta}_0) \to L^*(\underline{\theta}_0) \quad \text{with probability one.}$$

If $\underline{\theta} \neq \underline{\theta}_0$, then $L^*(\underline{\theta}) > L^*(\underline{\theta}_0)$, therefore there exists a $\delta(\underline{\theta}, \underline{\theta}_0) > 0$ such that

$$L^*(\underline{\theta}) - L^*(\underline{\theta}_0) > \delta > 0$$

Since $L_M(\underline{\theta}) \to L^*(\underline{\theta})$ with probability one, there exists a $M_0(\delta)$ such that

$$|L_M(\underline{\theta}) - L^*(\underline{\theta})| < \delta/2 \quad \text{with probability one}$$

and

$$|L_M(\underline{\theta}_0) - L^*(\underline{\theta}_0)| < \delta/2 \quad \text{with probability one for all } M > M_0(\delta)$$

Thus

$$|(L_M(\underline{\theta}) - L^*(\underline{\theta})) + (L_M(\underline{\theta}_0) - L^*(\underline{\theta}_0))| < \delta \quad \text{with probability one.}$$

Now look at

$$\begin{aligned} &|L_M(\underline{\theta}) - L_M(\underline{\theta}_0)| \\ &= |(L_M(\underline{\theta}) - L^*(\underline{\theta})) + L^*(\underline{\theta}) - L^*(\underline{\theta}_0) - (L_M(\underline{\theta}_0) - L^*(\underline{\theta}_0))| \\ &\geq |L^*(\underline{\theta}) - L^*(\underline{\theta}_0)| + |(L_M(\underline{\theta}) - L^*(\underline{\theta})) - (L_M(\underline{\theta}_0) - L^*(\underline{\theta}_0))| \\ &> 0 \quad \text{with probability one for all } M > M_0(\delta). \end{aligned}$$

This completes the proof of part (b).

The above theorem gives explicit forms for the desired sphere $S(\underline{\theta}_0, \rho)$; however, use of the Euclidean norm was needed in the proof of the theorem rather than an arbitrary norm. It is usually difficult to compute the indicated bounds using the Euclidean norm, and so we present in the

following theorem a set of conditions in terms of a more general norm. Again, let $\|\cdot\|_2$ denote the Euclidean norm on the particular real vector space under consideration. We will say that any other norm, $\|\cdot\|$, is subordinate to $\|\cdot\|_2$ if $\|\underline{x}\| \leq \|\underline{x}\|_2$ for all $\underline{x}$. We see that if we let $\underline{x} \in R^p$, and define

$$\|\underline{x}\|_\infty = \max_{1 \leq i \leq p} |x_i|,$$

then $\|\cdot\|_\infty$ is subordinate to $\|\cdot\|_2$. This fact makes the computation of all the required bounds particularly simple if we use interval arithmetic [11].

Theorem 5.3

Let $\|\cdot\|_2$ denotes the Euclidean norm, and let $\|\cdot\|$ be any norm which is subordinate to $\|\cdot\|_2$.

If (1) there is a $\underline{\theta}_0 \in \Omega$ such that $\underline{f}_M(\underline{\theta}_0) = 0$ for all $M \in I^+$ and hence

$$\lim_{M \to \infty} \frac{1}{M+1} \underline{f}_M^T(\underline{\theta}_0) \underline{f}_M(\underline{\theta}_0) = 0,$$

(2) $N_M^*(\underline{\theta}_0)$ is nonsingular for all $M \in I^+$,

(3) there is a sphere $S(\underline{\theta}_0, \rho)$ with radius $\rho > 0$ centered at $\underline{\theta}_0$ and a set of positive numbers $(\lambda, \gamma, \mu, \beta)$ such that for all $M \in I^+$

(i) $\|N_M^{-1}(\underline{\theta}_0)\| \leq \lambda$

(ii) $\|\frac{1}{\sqrt{M+1}} [\underline{f}_M'(\underline{\theta})]^T\| \leq \beta$ for all $\underline{\theta} \in S(\underline{\theta}_0, \rho)$

(iii) $\frac{1}{M+1} \|[\underline{f}_M''(\underline{\theta})]^T \underline{f}_M(\underline{\theta})\| \leq \gamma$ for all $\underline{\theta} \in S(\underline{\theta}_0, \rho)$

(iv) $\|N_M^*(\underline{\theta})\| \leq \mu$ for all $\underline{\theta} \in S(\underline{\theta}_0, \rho)$

(v) $\lambda(\mu\rho+\gamma)<1$

Then conclusions (a) and (b) of Theorem 5.2 are true.

Proof: As in the proof of Theorem 5.2, we suppose there is a $\underline{\theta}_1 \in S(\underline{\theta}_0,\rho)$, $\underline{\theta}_1 \neq \underline{\theta}_0$, such that

$$\lim_{M\to\infty} \frac{1}{M+1} \underline{f}_M^T(\underline{\theta}_1)\underline{f}_M(\underline{\theta}_1) = 0$$

Hence,

$$\lim_{M\to\infty} \left\| \frac{1}{\sqrt{M+1}} \underline{f}_M(\underline{\theta}_1) \right\|_2 = 0$$

Since $\|\cdot\|$ is a subordinate to $\|\cdot\|_2$,

$$\lim_{M\to\infty} \left\| \frac{1}{\sqrt{M+1}} \underline{f}_M(\underline{\theta}_1) \right\| = 0.$$

Again we have

$$\|\underline{\theta}_0 - \underline{\theta}_1\| < \|\underline{\theta}_0 - \underline{\theta}_1\| + \lambda\|\underline{\phi}_M^*(\underline{\theta}_1)\| \quad \text{for all } M \in I^+$$

Now look at

$$\|\underline{\phi}_M^*(\underline{\theta}_1)\| = \left\| \frac{1}{M+1} [\underline{f}_M'(\underline{\theta}_1)]^T \underline{f}_M(\underline{\theta}_1) \right\|$$

$$\leq \left\| \frac{1}{\sqrt{M+1}} [\underline{f}_M'(\underline{\theta}_1)]^T \right\| \left\| \frac{1}{\sqrt{M+1}} \underline{f}_M(\underline{\theta}_1) \right\|$$

$$\leq \beta \left\| \frac{1}{\sqrt{M+1}} \underline{f}_M(\underline{\theta}_1) \right\|$$

Hence $\lim_{M\to\infty} \|\underline{\phi}_M^*(\underline{\theta}_1)\| = 0$

Therefore we conclude that

$$\|\underline{\theta}_0 - \underline{\theta}_1\| < \|\underline{\theta}_0 - \underline{\theta}_1\|$$

which is a contradiction, and so $\underline{\theta}_1 = \underline{\theta}_0$. Part b is the same as in Theorem 5.2. This concludes the proof of the theorem.

Now let $S(\underline{\theta}_0,\rho)$ be a sphere such that $\underline{\theta} \in S(\underline{\theta}_0,\rho)$ and $L^*(\underline{\theta}) = L^*(\underline{\theta}_0)$ implies $\underline{\theta} = \underline{\theta}_0$, e.g. as provided in Theorems 5.2 and 5.3. Now consider the CML estimation problem: find $\hat{\underline{\theta}}_M \in S(\underline{\theta}_0,\rho)$ as an estimate of $\underline{\theta}_0$ where $\hat{\underline{\theta}}_M$ is constructed by

$$L_M(\hat{\underline{\theta}}_M) = \min_{\underline{\theta} \in \bar{S}(\underline{\theta}_0,\rho)} L_M(\underline{\theta}), \quad M = 0,1,2,\cdots \tag{5.27}$$

$\bar{S}(\underline{\theta}_0,\rho)$ denotes the closure of $S(\underline{\theta}_0,\rho)$.

To show that $\{\hat{\underline{\theta}}_M\}_{M=0}^{\infty}$ converges to $\underline{\theta}_0$ with probability one, we need the following Lemma.

Lemma 5.4 [Wald-Kendall-Åström] [Aoki and Yue, 1]

Let $\{\hat{\underline{\theta}}_M\}_{M=0}^{\infty}$ be constructed by (5.2), then $\{\hat{\underline{\theta}}_M\}_{M=0}^{\infty}$ converges to $\underline{\theta}^* \in \Omega^* \cap \Omega$ with probability one, where Ω^* is defined by

$$\Omega^* = \{\underline{\theta} : L^*(\underline{\theta}) = L^*(\underline{\theta}_0)\}.$$

Theorem 5.5

Let $S(\underline{\theta}_0,\rho)$ be given by Theorems 5.2 or 5.3 and let $\{\hat{\underline{\theta}}_M\}_{M=0}^{\infty}$ be constructed by (5.27). Then the CML estimation sequence $\{\hat{\underline{\theta}}_M\}_{M=0}^{\infty}$ converges to $\underline{\theta}_0$ with probability one.

Proof: Since $\underline{\theta}_0$ is the unique $\underline{\theta}$ in $S(\underline{\theta}_0,\rho)$ such that $L^*(\underline{\theta}) = L^*(\underline{\theta}_0)$, $\Omega^* \cap S(\underline{\theta}_0,\rho) = \{\underline{\theta}_0\}$ is a singleton.

Hence the result follows immediately from Lemma 5.4.

Corollary 5.6

Let $S(\underline{\theta}_0, \rho)$ be the given in Theorem 5.2 or 5.3, then $\underline{\theta}_0$ is locally CML identifiable.

5.3 An Example of Computing the Region of Parameter Identifiability

We now give an example to illustrate the algorithm of Theorem 5.3. Consider the system

$$\begin{bmatrix} w_1(k+1) \\ w_2(k+1) \end{bmatrix} = \begin{bmatrix} \eta_1 & 1 \\ 0 & \eta_2 \end{bmatrix} \begin{bmatrix} w_1(k) \\ w_2(k) \end{bmatrix} + \begin{bmatrix} 0 \\ \eta_3 \end{bmatrix} u(k)$$

$$y(k) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} w_1(k) \\ w_2(k) \end{bmatrix} \tag{5.28}$$

$$z(k) = y(k) + v(k) \quad , \quad k = 0,1,2,\text{---}$$

$v(k)$ is $N(0,\sigma^2)$, $\sigma^2 < \infty$.

The block diagram of this system is depicted in Fig. 3. The deterministic part of this system is the same as system (4.42) hence by reversing the positions of $\frac{1}{z-\eta_1}$ and $\frac{1}{z-\eta_2}$, and by transforming the initial conditions by a similarity transformation, the resulting equivalent system will generate the same $y(k)$ sequence.

Let $\underline{\theta} = [w_1(0), w_2(0), \eta_1, \eta_2, \eta_3]^T$. Recalling from Sec. 4.3, $\underline{\theta}^* = [y_1(0), w_2(0) + (\eta_1 - \eta_2)w_1(0), \eta_2, \eta_1, \eta_3]$ is a point in Ω which is unresolvable from $\underline{\theta}$ by the equivalent transformation. Hence the solution to the equation $\underline{f}_M(\underline{\theta}) = 0$ is not unique in R^5 and only local identifiability can be imposed on the parameters of system (5.28). We applied the

algorithm of Theorem 5.3 to this example with $\underline{\theta}_0 = (0.5, 0.1, 0.3, 0.7, 1.0)^T$. We note that $\underline{\theta}_1 = (0.5, -0.1, 0.7, 0.3, 1.0)^T$ will give exactly the same set of y(k) sequence for any input sequence u(k) and hence is unresolvable from $\underline{\theta}_0$. However, the distance from $\underline{\theta}_0$ to $\underline{\theta}_1$ is 0.4 by using the norm $\|\cdot\|_\infty$. Therefore $\underline{\theta}_0$ and $\underline{\theta}_1$ are each locally identifiable in the spheres $S(\underline{\theta}_0,\rho)$ and $S(\underline{\theta}_1,\rho)$ respectively where $\rho < 0.4$. Using an input sequence $u(k) = 100 \sin(k\pi/4)$ for $k = 0, 1, ---$, the following data listed in Table 1 are obtained.

We see that uniqueness is guaranteed in a sphere of radius $\rho = 0.33$. We know the true radius to be $\rho = 0.4$ in this example. Again, this example illustrates that the computed sphere of parameter identifiability is of reasonable size in view of the distance from $\underline{\theta}_1$ to $\underline{\theta}_0$. Moreover, the region size obtained by applying Theorem 5.3 is larger than that obtained by applying Theorem 4.4 since condition 3(vi) of Theorem 4.4 is more conservative than condition 3(v) of Theorem 5.3.

Table 1. Region of parameter identifiability of system (5.28)

No. of Observations	λ	β	γ	μ	ρ	$\lambda(\mu\rho+\gamma)$
20	0.1314×10^{-6}	0.9682×10^{4}	0.6828×10^{7}	0.2126×10^{6}	0.285	0.905
40	0.1554×10^{-8}	0.8936×10^{5}	0.5226×10^{9}	0.8136×10^{6}	0.330	0.813
60	0.2139×10^{-9}	0.2737×10^{6}	0.4047×10^{10}	0.5418×10^{6}	0.330	0.866
80	0.2107×10^{-9}	0.2882×10^{6}	0.2560×10^{10}	0.1854×10^{6}	0.315	0.539
100	0.2263×10^{-8}	0.1493×10^{6}	0.4018×10^{9}	0.1046×10^{6}	0.295	0.909

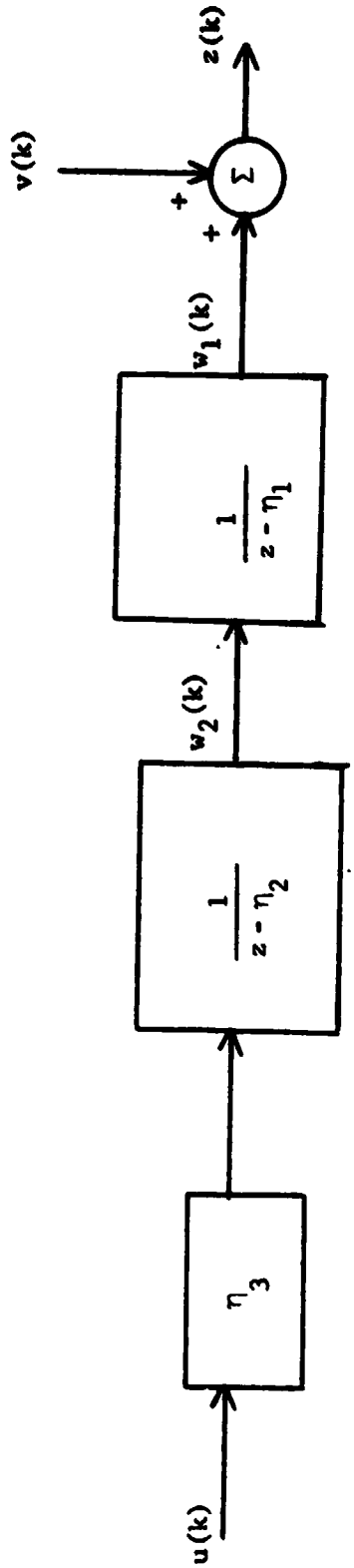

Figure 3. Block Diagram of System (5.28)

6. CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

Concepts of local identifiability and local CML identifiability of parameters which parametrizes the joint probability density function of the observation sequence are established. They are applicable to the identification problem of control systems where there are stochastic disturbances present. The local least-square identifiability and the transfer-function identifiability of parameters of deterministic linear dynamic systems are also introduced. Sufficient conditions for their identifiability are provided. It has been shown that for single-input, zero-state linear systems, we are able to find a region containing the true parameter as an interior point such that the true parameters are both locally least-square identifiable and identifiable from the transfer function in the same region under certain constant rank assumptions on the impulse response matrix and the derivatives of the identification criterion.

By modifying Glover and Willems' theorem in [6], a theorem is given to show that if a system is of minimal dimension, the system parameters (including the unknown initial state) are locally identifiable if it has unique equivalent system locally when the system is restricted to a given parametrization.

A brief survey on the theory of solving nonlinear least-square problems, ℓ_∞-norm, and interval arithmetic is given. Employing these techniques, a least-square type identification algorithm for finding explicitly the regions of parameter identifiability of general linear deterministic

dynamic systems is developed. A numerical example is included to illustrate this algorithm.

By modifying Herget's result [7], a theorem providing sufficient conditions for local CML identifiability of parameters of general dynamic systems with Gaussian-white measurement errors is established. A computation procedure is provided by the theorem for finding the regions of parameter identifiability. It has been shown that with probability one, the true parameter vector is the unique extremal point of the maximum likelihood function parametrized by the unknown parameter vector and the constrained maximum likelihood estimation sequence is consistent in the region of parameter identifiability. A numerical example is included to illustrate this computation procedure.

The system parameter identification problem of Gauss-Markov stochastic control systems driven by plant Gaussian-white noise and observed with Gaussian-white noise is an an area of further endeavor.

It has been shown the parameter identification problem of linear dynamic systems is equivalent to the initial-state observation problem of the quadratic-in-the-state bilinear systems. Hence the observability theory of quadratic-in-the-state bilinear systems needs to be studied more extensively.

Since the sufficient conditions for parameter identifiability are sensitive to the input sequence, further work in the area of optimal input synthesis for system identification may prove fruitful.

7. LITERATURE CITED

1. M. Aoki and P. C. Yue. "On certain convergence questions in system identification." SIAM J. Control, 8, (1970), 239-256.

2. R. Bellman and K. J. Åström. "On structural identifiability." Mathematical Biosciences, 7, (1970), 329-339.

3. Adi Ben-Israel. "A Newton-Raphson method for the solution of system of equations." J. Math. Analysis and Applications, 15, (1966), 243-252.

4. L. Breiman. Probability. Reading, Mass.: Addison Wesley, 1968.

5. J. L. Doob. Stochastic Processes. New York: Wiley, 1963.

6. K. Glover and J. C. Willems. "Parameterization of linear dynamical systems: canonical forms and identifiability." IEEE Trans. Automatic Control, AC-19 (Dec. 1974), 640-645.

7. C. J. Herget. "Regions of parameter identifiability." Engineering Research Institute Report ERI-74169, Iowa State University, Ames, Iowa, August 1974.

8. B. L. Ho and R. E. Kalman. "Effective construction of linear state-variable models from input/output data." Proc. 3rd. Allerton Conference, Urbana-Champaign, 1965.

9. R. V. Hogg and A. T. Craig. Introduction to Mathematical Statistics. London: Macmillan, 1970.

10. R. Narasimhan. Analysis on Real and Complex Manifolds. Amsterdam, The Netherlands: North-Holland, 1968.

11. E. P. Oppenheimer. "Application of interval analysis to problems of linear control systems." Ph.D. dissertation, Iowa State University, Ames, Iowa, 1974.

12. J. M. Ortega and W. C. Rheinholdt. Iterative Solution of Nonlinear Equations in Several Variables. New York: Academic Press, 1970.

13. V. Pereyra. "Iterative methods for solving nonlinear least square problems." SIAM J. Numerical Analysis, 20, (1967), 27-36.

14. L. B. Rall. Computational Solution of Nonlinear Operator Equations. New York: Wiley, 1969.

15. V. I. Smirnoff. Linear Algebra and Group Theory. New York: McGraw-Hill, 1961.

16. R. M. Staley and P. C. Yue. "On system parameter identifiability." Information Sciences, 2, (1970), 127-138.

17. L. W. Taylor and K. Iliff. "A modified Newton-Raphson method for determining stability derivatives from flight data." In Computing Methods in Optimization Problems, Vol. 2. L. Zadeh, L. W. Neustadt and A. V. Balakrishman, Eds. New York: Academic Press, 1969.

18. E. Tse. "Information matrix and local identifiability of parameters." Proc. of the 1973 Joint Automatic Control Conference, Columbus, Ohio, 1973.

19. E. Tse and J. Anton. "On the identifiability of parameters." IEEE Trans. Automatic Control, AC-17 (Oct. 1972), 637-646.

20. A. Wald. "Note on the consistency of the maximum likelihood estimate." Annals of Math. Statistics, 20, (1949), 595-601.

21. A. Wald. "Asymptotic properties of the maximum likelihood estimate of an unknown parameter of a discrete stochastic process." Annals of Math. Statistics, 19, (1948), 40-46.

8. ACKNOWLEDGMENTS

The author wishes to acknowledge the encouragement and guidance given to him by his major professor, Dr. C. J. Herget, during the preparation of this dissertation and throughout his graduate career at Iowa State.

Special thanks are offered to Dr. R. G. Brown, Dr. J. A. Dyer, and Dr. G. T. Chapman who as teachers made the author's graduate studies at Iowa State intellectually stimulating.

Additionally, the author expresses his gratitude to the NASA-Ames Research Center for providing financial support through the NASA-Ames University Consortium Joint Research Interchange Agreement with Iowa State University while preparing this document.

To his parents, General and Mrs. S. M. Lee, no acknowledgments can possibly convey the appreciation for their expressions of encouragement and confidence.

9. APPENDIX

This Appendix contains a listing of the computer program used to conduct the parameter identification example of the parameterized system given in Sec. 4.3. It has been written in Fortran language using double precision.

If actual input/output measurement data are available, "GENERATE INPUT AND OUTPUT SEQUENCES" in this computer program should be removed. Proper dimensioning of the matrix arrays should be noted.

```
      IMPLICIT COMP_EX(A-G,P-Q) , REAL*8(X-Y)
      DIMENSION XA(5,5),XB(5,5,5),XBSTA(5,5,5),XE(5,5),XC(1,5),XX(5),
     CXF(2,2),XG(2),XGAMM(5,5,5),XGP(5,5,5),X(2),XINFO(5,5),Y(50),
     CXFP(20,5),XFPT(5,20),HLA(5),A(5,5),B(5,5,5),BSTAR(5,5,5),C(1,5),
     CGDP(5,5,5),G(5,5),GAMMP(5,5,5,21),W(100,1),FE(2,2),GE(2),FT(2,2),
     CGT(2),ETAE(2),ETAT(2),AX(5),AYE(30),AYT(30),GAMMA(5,5,21),DE(2)
      DIMENSION DT(2),D2E(2),D2T(2),GP(5,5,21),GAMM(5,5,5,21),GGG(5,5,5)
     C,GG(5,5),FPT(20,5),FPP(20,5,5),FPPT(5,5,20),PHI(5),H1(5),PK(5),
     CFPPTF(5,5),AD(5),HDD(5),ANP(5,5,5),AA(5,5,5),AAT(5,5,5),HPK(5),
     CALEN(5,5),HLEN(5,5),HK(2),F(30),XU(30),GAMMM(5,5,5,21),U(30),
     CXD1(2),XD2(2),LL(5),MM(5),AAA(5,5),FP(20,5),ANORM(5,5,5)
      HPI04=7.85398E-01
      XPI04=7.853980-01
      NP=5
      NS=2
      NC=20
      XEP=1.0D-30
      XT=(10.0D 00)**2
      T=(10.0E 00)**2
      XYZ=DFLOAT(NO)
      XW=1.0D 00
C
C     COMPUTE LAMBDA
C
C
C     SPECIFY A MATRIX
C
      DO 140 I=1,NP
      DO 140 J=1,NP
  140 XA(I,J)=0.0D 00
      XA(1,2)=1.0D 00
      XA(3,3)=1.0D 00
      XA(4,4)=1.0D 00
      XA(5,5)=1.0D 00
C
C     SPECIFY B AND BSTAR MATRICES
C
```

```
      DO 121 I=1,NP
      DO 121 J=1,NP
      DO 121 K=1,NP
  121 XB(I,J,K)=0.0D 00
      XE(1,1,3)=1.0D 00
      XB(2,2,4)=1.0D 00
      DO 122 I=1,NP
      DO 122 J=1,NP
      DO 122 K=1,NP
  122 XBSTA(I,J,K)=XB(I,K,J)
C
C     SPECIFY E MATRIX
C
      DO 141 I=1,NP
      DO 141 J=1,NP
  141 XE(I,J)=0.0D 00
      XE(2,5)=1.0D 00
C
C     SPECIFY C MATRIX
C
      DO 142 I=1,NP



  142 XC(1,I)=0.0D 00
      XC(1,1)=1.0D 00
C
C     SPECIFY TRUE DATA
C
```

```
      X(1)=0.5D 00
      X(2)=0.1D 00
      XF(1,1)=0.3D 00
      XF(1,2)=1.0D 00
      XF(2,1)=0.0D 00
      XF(2,2)=0.7D 00
      XG(1)=0.0D 00
      XG(2)=1.0D 00
      XX(1)=X(1)
      XX(2)=X(2)
      XX(3)=XF(1,1)
      XX(4)=XF(2,2)
      XX(5)=XG(2)
      Y(1)=X(1)
C
C     SPECIFY GAMMA ZERO MATRIX
C
      DO 5 I=1,NP
      DO 5 J=1,NP
      IF (I.EQ.J) GO TO 10
      XGAMM(I,J,1)=0.0D 00
      GO TO 5
   10 XGAMM(I,J,1)=1.0D 00
    5 CONTINUE
C
C     SPECIFY G PRIME ONE MATRIX
C
      DO 128 I=1,NP
      DO 128 J=1,NP
      XGP(I,J,1)=0.0D 00
      DO 123 K=1,NP
  123 XGP(I,J,1)=XGP(I,J,1)+(XB(I,J,K)+XBSTA(I,J,K))*XX(K)
      XGP(I,J,1)=XGP(I,J,1)+XA(I,J)
  128 CONTINUE
C
C     GENERATE INPUT AND OUTPUT SEQUENCES
C
```

```
      DO 120 K=2,NO
      L1=K-1
      XU(L1)=DSIN(XPIO4*DFLOAT(L1-1))*XT
      CALL MATVEC(XF,X,XD1,NS,NS)
      DO 201 I=1,NS
  201 XD2(I)=XU(L1)*XG(I)
      CALL MATADD(XD1,XD2,X,NS,1)
      Y(K)=X(1)
      XX(1)=X(1)
      XX(2)=X(2)
C
C     GENERATE G PRIME MATRICES
C
      DO 124 I=1,NP
      DO 124 J=1,NP
      XGP(I,J,K)=0.0D 00

      DO 125 N=1,NP
  125 XGP(I,J,K)=XGP(I,J,K)+(XB(I,J,N)+XBSTA(I,J,N))*XX(N)
      XU(L1)=DSIN(XPIO4*DFLOAT(L1))
      XGP(I,J,K)=XGP(I,J,K)+XU(L1)*XE(I,J)*XT+XA(I,J)
  124 CONTINUE
C
C     GENERATE GAMMA MATRICES
C
      DO 130 I=1,NP
      DO 130 J=1,NP
      XGAMM(I,J,K)=0.0D 00
      DO 130 N=1,NP
      XGAMM(I,J,K)=XGAMM(I,J,K)+XGP(I,N,K-1)*XGAMM(N,J,K-1)
  130 CONTINUE
  120 CONTINUE
      DO 170 I=1,NO
      DO 170 J=1,NP
      XFP(I,J)=0.0D 00
      DO 170 N=1,NP
      XFP(I,J)=XFP(I,J)+XC(1,N)*XGAMM(N,J,I)
```

```
  170 CONTINUE
      PRINT 555
  555 FORMAT(1H1,'THE F PRIME MATRIX'//)
      DO 666 I=1,NO
  666 PRINT 270, (XFP(I,J),J=1,NP)
C
C     GENERATE INFORMATION MATRIX
C
      CALL TRNSPZ(XFP,XFPT,NO,NP)
      CALL MATMLT(XFPT,XFP,XINFO,NP,NO,NP)
      PRINT 480
  480 FORMAT(1H1,'THE N MATRIX'//)
      DO 481 I=1,NP
  481 PRINT 270, (XINFO(I,J),J=1,NP)
      DO 781 I=1,NP
      DO 781 J=1,NP
      XINFO(I,J)=XINFO(I,J)/XW
  781 CONTINUE
      CALL MATINV(XINFO,NP,NP,XEP,DETER,6999)
      PRINT 250
  250 FORMAT('0','THE INFORMATION MATRIX'//)
      DO 782 I=1,NP
      DO 782 J=1,NP
      XINFO(I,J)=XINFO(I,J)/XW
  782 CONTINUE
      DO 260 I=1,NP
  260 PRINT 270, (XINFO(I,J),J=1,NP)
  270 FORMAT(' ',5X,5D15.7)
      DO 300 I=1,NP
      HLA(I)=0.0
      DO 300 J=1,NP
      Z=XINFO(I,J)
      H=ABS(Z)
      HLA(I)=HLA(I)+H
  300 CONTINUE
      HLAM=2.0*AMAX1(HLA(1),HLA(2),HLA(3),HLA(4),HLA(5))
      HLAM=XYZ*HLAM
      PRINT 737, HLAM
```

```
  737 FORMAT('0',' LAMBDA=',E15.7)
C
C       SPECIFY A MATRIX IN INTERVAL FORM
C
        DO 20 I=1,NP
        DO 20 J=1,NP
   20   A(I,J)=CMPLX(0.0,0.0)
        A(1,2)=CMPLX(1.0,1.0)
        A(3,3)=CMPLX(1.0,1.0)
        A(4,4)=CMPLX(1.0,1.0)
        A(5,5)=CMPLX(1.0,1.0)
C
C       SPECIFY B AND BSTAR MATRICES IN INTERVAL FORM
C
        DO 16 I=1,NP
        DO 16 J=1,NP
        DO 16 K=1,NP
   16   B(I,J,K)=CMPLX(0.0,0.0)
        B(1,1,3)=CMPLX(1.0,1.0)
        B(2,2,4)=CMPLX(1.0,1.0)
        DO 15 I=1,NP
        DO 15 J=1,NP
        DO 15 K=1,NP
        BSTAR(I,J,K)=B(I,K,J)
   15   CONTINUE
C
C       SPECIFY G DOUBLE PRIME MATRIX
C
        DO 17 I=1,NP
        DO 17 J=1,NP
        DO 17 K=1,NP
        CALL INTADD(B(I,J,K),BSTAR(I,J,K),C1)
   17   GDP(I,J,K)=C1
C
C       SPECIFY E MATRIX IN INTERVAL FORM
C
```

```
      DO 19 I=1,NP
      DO 19 J=1,NP
   19 G(I,J)=CMPLX(0.0,0.0)
      G(2,5)=CMPLX(1.0,1.0)
C
C     SPECIFY C MATRIX IN INTERVAL FORM
C
      DO 99 I=1,NP
   99 C(1,I)=CMPLX(0.0,0.0)
      C(1,1)=CMPLX(1.0,1.0)
C
C     SPECIFY GAMMA 0, GAMMA 1, GAMMA' 0, AND GAMMA' 1 MATRICES
C
      READ (5,747) RHO,HDETA
  747 FORMAT(2F10.0)
      W(1,1)=0.0
      W(2,1)=0.0
      DO 600 NN=3,3
      RHO=RHO-HDETA
      W(NN,1)=RHO
      R1=0.5
      R2=0.1

      FE(1,1)=CMPLX(0.3-RHO,0.3+RHO)
      FE(1,2)=CMPLX(1.0,1.0)
      FE(2,1)=CMPLX(0.0,0.0)
      FE(2,2)=CMPLX(0.7-RHO,0.7+RHO)
      GE(1)=CMPLX(0.0,0.0)
      GE(2)=CMPLX(1.0-RHO,1.0+RHO)
      ETAE(1)=CMPLX(R1-RHO,R1+RHO)
      ETAE(2)=CMPLX(R2-RHO,R2+RHO)
      FT(1,1)=CMPLX(0.3,0.3)
      FT(1,2)=CMPLX(1.0,1.0)
      FT(2,1)=CMPLX(0.0,0.0)
      FT(2,2)=CMPLX(0.7,0.7)
      GT(1)=CMPLX(0.0,0.0)
      GT(2)=CMPLX(1.0,1.0)
```

```
      ETAT(1)=CMPLX(R1,R1)
      ETAT(2)=CMPLX(R2,R2)
      AX(1)=ETAE(1)
      AX(2)=ETAE(2)
      AX(3)=FE(1,1)
      AX(4)=FE(2,2)
      AX(5)=GE(2)
      AYE(1)=ETAE(1)
      AYT(1)=ETAT(1)
      CALL INTSUB(AYE(1),AYT(1),AC)
      F(1)=AC
      DO 210 I=1,NP
      DO 210 J=1,NP
      IF (I.EQ.J) GO TO 211
      GAMMA(I,J,1)=CMPLX(0.0,0.0)
      GO TO 210
  211 GAMMA(I,J,1)=CMPLX(1.0,1.0)
  210 CONTINUE
      DO 6 I=1,NP
      DO 6 J=1,NP
      GAMMA(I,J,2)=CMPLX(0.0,0.0)
      DO 8 L=1,NP
      CALL INTMUL(B(I,J,L),AX(L),C1)
      CALL INTADD(GAMMA(I,J,2),C1,C2)
    8 GAMMA(I,J,2)=C2
      DO 9 L=1,NP
      CALL INTMUL(BSTAR(I,J,L),AX(L),C1)
      CALL INTADD(GAMMA(I,J,2),C1,C2)
    9 GAMMA(I,J,2)=C2
      CALL INTADD(A(I,J),GAMMA(I,J,2),C1)
      GAMMA(I,J,2)=C1
      C3=G(I,J)*0.0
      CALL INTADD(C3,GAMMA(I,J,2),C2)
      GAMMA(I,J,2)=C2
    6 CONTINUE
      DO 21 I=1,NP
      DO 21 J=1,NP
      DO 21 K=1,NP
```

```
      GAMMP(I,J,K,1)=CMPLX(0.0,0.0)
   21 GAMMP(I,J,K,2)=GDP(I,J,K)
C
C     GENERATE H AND F MATRICES
C
      DO 100 K=2,NO

      L1=K-1
      U(L1)=SIN(HPIO4*FLOAT(L1-1))*T
      DO 1 I=1,NS
      DE(I)=CMPLX(0.0,0.0)
      DT(I)=CMPLX(0.0,0.0)
      DO 1 KK=1,NS
      CALL INTMUL(FE(I,KK),ETAE(KK),C1)
      CALL INTMUL(FT(I,KK),ETAT(KK),C2)
      CALL INTADD(DE(I),C1,C3)
      CALL INTADD(DT(I),C2,C4)
      DE(I)=C3
      DT(I)=C4
    1 CONTINUE
      DO 2 I=1,NS
      D2E(I)=U(L1)*GE(I)
      D2T(I)=U(L1)*GT(I)
    2 CONTINUE
      DO 3 I=1,NS
      CALL INTADD(DE(I),D2E(I),C1)
      CALL INTADD(DT(I),D2T(I),C2)
      ETAE(I)=C1
      ETAT(I)=C2
    3 CONTINUE
      AX(1)=ETAE(1)
      AX(2)=ETAF(2)
      AYE(K)=CMPLX(0.0,0.0)
      DO 301 I=1,NP
      CALL INTMUL(C(1,I),AX(I),C1)
      CALL INTADD(AYE(K),C1,C2)
      AYE(K)=C2
```

```
  301 CONTINUE
      AX(1)=ETAT(1)
      AX(2)=ETAT(2)
      AYT(K)=CMPLX(0.0,0.0)
      DO 302 I=1,NP
      CALL INTMUL(C(1,I),AX(I),C1)
      CALL INTADD(AYT(K),C1,C2)
      AYT(K)=C2
  302 CONTINUE
      CALL INTSUB(AYE(K),AYT(K),C1)
      F(K)=C1
C
C     GENERATE G' AND GAMMA MATRICES
C
      DO 50 I=1,NP
      DO 50 J=1,NP
      AX(1)=ETAE(1)
      AX(2)=ETAE(2)
      GP(I,J,K+1)=CMPLX(0.0,0.0)
      DO 51 L=1,NP
      CALL INTMUL(B(I,J,L),AX(L),C1)
      CALL INTADD(GP(I,J,K+1),C1,C2)
   51 GP(I,J,K+1)=C2
      DO 52 L=1,NP
      CALL INTMUL(BSTAR(I,J,L),AX(L),C1)
      CALL INTADD(GP(I,J,K+1),C1,C2)
   52 GP(I,J,K+1)=C2
      GG(I,J)=SIN(HPIO4*FLOAT(L1))*G(I,J)*T

      CALL INTADD(GG(I,J),GP(I,J,K+1),C2)
      GP(I,J,K+1)=C2
      CALL INTADD(A(I,J),GP(I,J,K+1),C2)
      GP(I,J,K+1)=C2
   50 CONTINUE
      DO 11 I=1,NP
      DO 11 J=1,NP
```

```
      GAMMA(I,J,K+1)=CMPLX(0.0,0.0)
      DO 11 L=1,NP
      CALL INTMUL(GP(I,L,K+1),GAMMA(L,J,K),C1)
      CALL INTADD(GAMMA(I,J,K+1),C1,C2)
      GAMMA(I,J,K+1)=C2
   11 CONTINUE
C
C     GENERATE GAMMA PRIME MATRICES
C
      DO 29 I=1,NP
      DO 29 J=1,NP
      DO 29 L=1,NP
      GAMM(I,J,L,K+1)=CMPLX(0.0,0.0)
      DO 29 M=1,NP
      CALL INTMUL(GDP(I,J,M),GAMMA(M,L,K),C1)
      CALL INTADD(GAMM(I,J,L,K+1),C1,C2)
      GAMM(I,J,L,K+1)=C2
   29 CONTINUE
      DO 22 I=1,NP
      DO 22 J=1,NP
      DO 22 L=1,NP
   22 GAMM(I,J,L,K+1)=GAMM(I,L,J,K+1)
      DO 23 I=1,NP
      DO 23 J=1,NP
      DO 23 L=1,NP
      GAMMM(I,J,L,K+1)=CMPLX(0.0,0.0)
      DO 23 M=1,NP
      CALL INTMUL(GAMM(I,J,M,K+1),GAMMA(M,L,K),C1)
      CALL INTADD(GAMMM(I,J,L,K+1),C1,C2)
      GAMMM(I,J,L,K+1)=C2
   23 CONTINUE
      DO 24 I=1,NP
      DO 24 J=1,NP
      DO 24 L=1,NP
      GGG(I,J,L)=CMPLX(0.0,0.0)
      DO 25 M=1,NP
```

```
      CALL INTMUL(GP(I,M,K+1),GAMMP(M,J,L,K),C1)
      CALL INTADD(GGG(I,J,L),C1,C2)
   25 GGG(I,J,L)=C2
      CALL INTADD(GAMMM(I,J,L,K+1),GGG(I,J,L),C1)
      GAMMP(I,J,L,K+1)=C1
   24 CONTINUE
  100 CONTINUE
C
C     GENERATE F PRIME AND F DOUBLE PRIME MATRICES
C
      DO 30 I=1,NO
      DO 30 J=1,NP
      FP(I,J)=CMPLX(0.0,0.0)
      DO 30 L=1,NP
      CALL INTMUL(C(I,L),GAMMA(L,J,I),C1)

      CALL INTADD(FP(I,J),C1,C2)
   30 FP(I,J)=C2
      DO 31 I=1,NO
      DO 31 J=1,NP
      DO 31 L=1,NP
      FPP(I,J,L)=CMPLX(0.0,0.0)
      DO 31 M=1,NP
      CALL INTMUL(C(I,M),GAMMP(M,J,L,I),C1)
      CALL INTADD(FPP(I,J,L),C1,C2)
      FPP(I,J,L)=C2
   31 CONTINUE
      DO 32 I=1,NP
      DO 32 J=1,NP
      DO 32 L=1,NO
   32 FPPT(I,J,L)=FPP(L,I,J)
      DO 33 I=1,NP
      DO 33 J=1,NO
   33 FPT(I,J)=FP(J,I)
C
C     COMPUTE K
C
```

```
      DO 498 I=1,NP
      PK(I)=CMPLX(0.0,0.0)
      DO 499 J=1,NO
      CALL INTABS(FPT(I,J),C1)
      CALL INTADD(C1,PK(I),C2)
  499 PK(I)=C2
      PZ=PK(I)
  498 HPK(I)=AIMAG(PZ)
      HQ=AMAX1(HPK(1),HPK(2),HPK(3),HPK(4),HPK(5))
      HQ=HQ/DSQRT(XYZ)
      PRINT 501,HQ
  501 FORMAT('0','K=',E15.7)
C
C     COMPUTE ETA
C
      DO 35 I=1,NP
      PHI(I)=CMPLX(0.0,0.0)
      DO 36 J=1,NO
      CALL INTMUL(FPT(I,J),F(J),C1)
      CALL INTADD(PHI(I),C1,C2)
   36 PHI(I)=C2
      CALL INTABS(PHI(I),C3)
      PHI(I)=C3
      AZ=PHI(I)
   35 H1(I)=AIMAG(AZ)
      HETA=AMAX1(H1(1),H1(2),H1(3),H1(4),H1(5))
      HETA=HETA/XYZ
      PRINT 738,HETA
  738 FORMAT('0','ETA=',E15.7)
C
C     COMPUTE GAMMA
C
      DO 37 I=1,NP
      DO 37 J=1,NP
      FPPT(I,J)=CMPLX(0.0,0.0)
      DO 37 L=1,NO
      CALL INTABS(FPPT(I,J),F(L),C1)
```

```
      CALL INTADD(FPPTF(I,J),C1,C2)
      FPPTF(I,J)=C2
   37 CONTINUE
      DO 38 I=1,NP
      AD(I)=CMPLX(0.0,0.0)
      DO 39 J=1,NP
      CALL INTABS(FPPTF(I,J),C1)
      CALL INTADD(C1,AD(I),C2)
   39 AD(I)=C2
      AZ=AD(I)
   38 HDD(I)=AIMAG(AZ)
      HGA=AMAX1(HDD(1),HDD(2),HDD(3),HDD(4),HDD(5))
      HGA=HGA/XYZ
      PRINT 739,HGA
  739 FORMAT('0','GAMMA=',E15.7)
C
C     COMPUTE K2
C
      DO 60 I=1,NP
      DO 60 J=1,NP
      DO 60 N=1,NP
      ANP(I,J,N)=CMPLX(0.0,0.0)
      DO 60 L=1,NP
      CALL INTMUL(FPT(I,L),FPP(L,J,N),C1)
      CALL INTADD(ANP(I,J,N),C1,C2)
      ANP(I,J,N)=C2
   60 CONTINUE
      DO 61 I=1,NP
      DO 61 J=1,NP
      DO 61 N=1,NP
      AA(I,J,N)=CMPLX(0.0,0.0)
      DO 61 L=1,NP
      CALL INTMUL(FPPT(I,J,L),FP(L,N),C1)
      CALL INTADD(AA(I,J,N),C1,C2)
      AA(I,J,N)=C2
```

```
61 CONTINUE
   DO 62 I=1,NP
   DO 62 J=1,NP
   DO 62 N=1,NP
62 AAT(I,J,N)=AA(N,I,J)
   DO 63 I=1,NP
   DO 63 J=1,NP
   DO 63 N=1,NP
   CALL INTADD(ANP(I,J,N),AAT(I,J,N),C1)
   ANP(I,J,N)=C1
63 CONTINUE
   DO 70 N=1,2
   IF (N.EQ.2) GO TO 71
   DO 72 I=1,NP
   DO 72 J=1,NP
   IF (I.EQ.J) GO TO 73
   AAA(I,J)=CMPLX(-1.0,1.0)
   GO TO 72
73 AAA(I,J)=CMPLX(1.0,1.0)
72 CONTINUE
   GO TO 77
71 DO 74 I=1,NP
   DO 74 J=1,NP

   IF (I.EQ.J) GO TO 75
   AAA(I,J)=CMPLX(-1.0,1.0)
   GO TO 74
75 AAA(I,J)=CMPLX(-1.0,-1.0)
74 CONTINUE
77 DO 78 L=1,NP
   DO 78 I=1,NP
   DO 78 J=1,NP
   ANORM(I,J,L)=CMPLX(0.0,0.0)
   DO 78 K=1,NP
   CALL INTMUL(ANP(I,J,K),AAA(L,K),C1)
   CALL INTADD(ANORM(I,J,L),C1,C2)
   ANORM(I,J,L)=C2
```

```
   78 CONTINUE
      DO 79 L=1,NP
      DO 79 I=1,NP
      ALEN(I,L)=CMPLX(0.0,0.0)
      DO 80 J=1,NP
      CALL INTABS(ANORM(I,J,L),C1)
      CALL INTADD(ALEN(I,L),C1,C2)
      ALEN(I,L)=C2
   80 CONTINUE
   79 HLEN(I,L)=AIMAG(ALEN(I,L))
      HK(N)=AMAX1(HLEN(1,1),HLEN(1,2),HLEN(1,3),HLEN(1,4),HLEN(2,1),HLEN
     C(2,2),HLEN(2,3),HLEN(2,4),HLEN(3,1),HLEN(3,2),HLEN(3,3),HLEN(3,4),
     CHLEN(4,1),HLEN(4,2),HLEN(4,3),HLEN(4,4))
      HK(N)=AMAX1(HK(N),HLEN(1,5),HLEN(2,5),HLEN(3,5),HLEN(4,5),HLEN(5,1
     C),HLEN(5,2),HLEN(5,3),HLEN(5,4),HLEN(5,5))
   70 CONTINUE
      HK2=AMAX1(HK(1),HK(2))
      HK2=HK2/XYZ
      PRINT 740,HK2
  740 FORMAT('0','K2=',E15.7)
C
C     SPECIFY THE STOPPING CRITERION
C
      HL=HLAM*HETA
      HBO=HLAM*(2.0*HK2*RHO+HGA)
      RHO4=RHO/4.0
      IF (HL.GT.RHO4) GO TO 608
      IF (HBO.GT.0.5) GO TO 608
      IF (W(NN,1).EQ.W(NN-2,1)) GO TO 400
      RHO=RHO+2.0*HDETA
```

```
608 MK=NN-2
    PRINT 602, MK
602 FORMAT('0',' I=',I3)
    PRINT 605,HL
605 FORMAT('0',' LAMBDA*ETA=',E15.7)
    PRINT 606, HBO
606 FORMAT('0',' LAMBDA(2.0*K2*RHO+GAMMA)=',E15.7)
    PRINT 601,RHO
601 FORMAT('0',' RHO=',E15.7)
600 CONTINUE
400 PRINT 500, RHO
500 FORMAT ('0',' THE CONVERGENCE SPHERE DIAMETER, RHO=',E15.7)
999 STOP
    END
```