WORKING PAPER

IDENTIFICATION - A THEORY OF GUARANTEED ESTIMATES

A.B. Kurzhanski

July 1988
WP-88-55

IIASA
International Institute
for Applied Systems Analysis

IDENTIFICATION – A THEORY OF GUARANTEED ESTIMATES

A.B. Kurzhanski

July 1988
WP-88-55

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

PREFACE

This paper gives an introduction to the theory of parameter identification and state estimation for systems subjected to uncertainties with set-membership bounds on the unknowns.

The situation under discussion may often turn to be more a propos since here the system and the environment are modelled as truly uncertain rather than noisy. The described approach is purely deterministic.

On the other hand the techniques involved here for the treatment of systems with nonquadratic constraints on the unknowns are proved to have some nontrivial interrelations with those developed in stochastic estimation theory. This may lead to some further estimation schemes that would combine the deterministic and the stochastic models of uncertainty.

The recurrence procedures of this paper are devised into relations that would allow numerical simulations.

Alexander B. Kurzhanski
System and Decision Sciences Program

CONTENTS

IDENTIFICATION – A THEORY OF GUARANTEED ESTIMATES

A.B. Kurzhanski

1. Introduction

A crucial issue in the process of *mathematical modelling* on the basis of *available observations* is the problem of system *parameter identification* under observation noise. The conventional area of applied mathematics within which the problem is usually discussed is *mathematical statistics* [1, 2]. The uncertainties in the system parameters and the observation noise are taken here to be described by stochastic mechanisms. The informational scheme for the identification process usually assumes that there exists an adequate statistical description for the unknowns. Within this framework a fairly complete theory has been developed for linear systems with disturbances modelled by gaussian noise and with quadratic criteria of optimality for the estimates [3, 4]. A large number of investigations is devoted to statistical identification under more general assumptions.

However, the statistical methods are not the only mathematical tools for the treatment of system modelling.

This paper gives an introduction to the theory of *guaranteed identification*. It demonstrates for example that the classical system parameter estimation problem under measurement noise may be posed in a deterministic setting rather than in a traditional probabilistic framework. The adopted model assumes that there is no statistical description for the measurement "noise" or for the disturbances in the system and that the only information on these is restricted to a *set-membership constraint* on their admissible values or realizations. A considerable number of applications in engineering and systems analysis are treated under informational assumptions that justify this approach (see e.g. [5–10]).

The basic techniques that are necessary for the treatment of the problems given here are based on set-valued calculus so that the solutions are formulated in the form of set-valued estimators. This approach also assures numerical robustness for the respective approximation schemes. Other results related to the topic of this paper may be found in [11–18].

Let us start with a trivial example. Suppose one is to identify a vector $c \in \mathbf{R}^2$ on the basis of observations $y(k) = c + \xi(k)$, $k = 1, \ldots, N, \ldots$ corrupted by "noise" $\xi(k)$.

Contrary to the conventional approach we will at first assume that there is no statistical data on $\xi(k)$ being available in advance. However we will suppose that a restriction

$$\xi(k) \in Q(k)$$

is given with set $Q(k)$ being known. We will assume that $Q(k)$, $k \geq k_0$ is a convex compact set.

Every single measurement $y(k)$ gives us some information on c, namely it indicates that the following inclusion is true

$$c \in y(k) - Q(k) \tag{1.2}$$

Having had m observations $y(1), \ldots, y(m)$, we observe that inclusion (1.2) should be true for every $k = 1, \ldots, m$. Hence, after m observations we will have

$$c \in \bigcap_{k=1}^{m} (y(k) - Q(k)) = C[1, m]$$

where the set $C[1, m]$ is the "guaranteed estimate" for c after m observations.

It is thus clear that every "new" measurement $y(m+1)$ introduces an innovation into the estimation process by means of an intersection of the previous estimate $C[1, m]$ with a "new" set $\{ y(m+1) - Q(m+1) \}$, so that

$$C[1\ ,\ m+1] = C[1\ ,\ m] \bigcap \{(y(m+1) - Q\ (m+1)\} \qquad (1.3)$$

Relation (1.3) is a ***recurrence equation*** which describes the evolution of the estimate $C[1\ ,\ m]$ in m. (Figures 1 and 2 demonstrate set $C[1\ ,\ m]$ for $m = 4$ with Q being (1) - a square, (2) - a circle; c^* stands for the unknown value to be estimated). The "accuracy" of the estimate will now depend on the behaviour of the "noise" $\xi(k)$. Let us trace this fact more precisely.

Assume c^* , $\xi^*(k)$ are the unknown actual values of c , $\xi(k)$, $k \in [1\ ,\ldots,\ m]$ so that the available measurement is

$$y(k) = c^* + \xi^*(k)$$

Then the estimate

$$C[1\ ,\ m] = \bigcap_{k=1}^{m} (c^* - (Q(k) - \xi^*(k))) =$$

$$= c^* + \bigcap_{k=1}^{m} (\xi^*(k) - Q(k))$$

where

$$\bigcap_{k=1}^{m} (\xi^*(k) - Q(k)) = R^*(m)$$

is ***the "error set" of the estimation process.*** It obviously depends on the behaviour of the "noise" $\xi^*(k)$, $k = 1\ ,\ldots,\ m$.

Let us examine the "worst case" solution (from the point of view of the observer). Suppose

$$\xi^*(k) \equiv 0 \qquad (1.4)$$

$$Q(k) \equiv Q\ ,\ Q = -Q \quad (k = 1\ ,\ldots,\ m)$$

("noise" constant, and "Q is stationary" and symmetric about the origin). Then, clearly

$$R^*(m) = \bigcap_{k=1}^{m} (-Q(k)) = -Q = Q$$

and the range of the error of estimation is precisely Q. The "guaranteed" error is

$$\max \{\ ||\ q\ ||\ |\ q \in Q\}$$

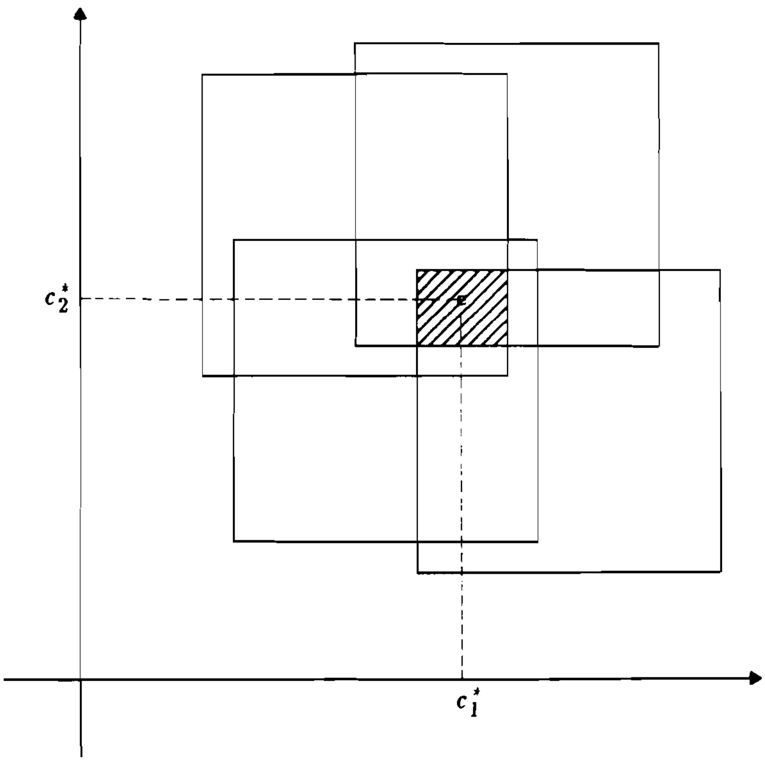

FIGURE 1

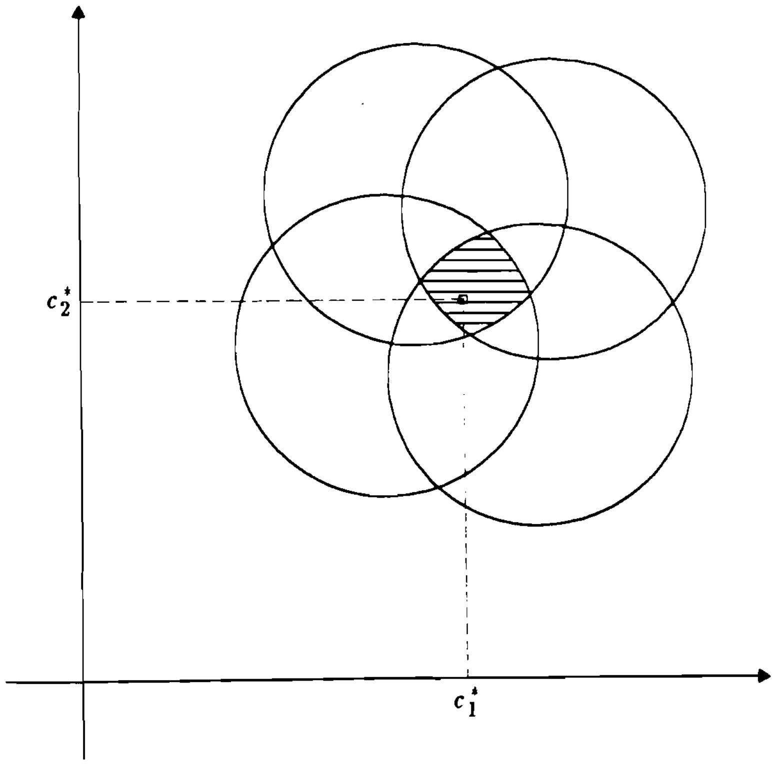

FIGURE 2

It is obvious here that none of the new measurements do bring any innovation into the estimation process.

In contrast an "adequate" behaviour of $\xi(k)$ may considerably improve the estimation. For example, assume that Q is a square: $Q = \mathbf{S}$,

$$\mathbf{S} = \{q : |q_i| \leq 1 , i = 1,2\}$$
$$c^* = \begin{bmatrix} c_1^* \\ c_2^* \end{bmatrix}$$

is the unknown vector to be identified.

If

$$\xi^*(1) = (1 , 1)$$
$$\xi^*(2) = (-1 , -1)$$

then the error set

$$R^*(2) = \{\xi^*(1) - Q\} \bigcap \{\xi^*(2) - Q\} = \{0\}$$

and the estimation is exact, (Figure 3).

For another example take $Q = S(0)$ to be a unit circle, $m = 3$, $\xi^*(1) = (1,0)$, $\xi^*(2) = (0,1)$, $\xi^*(3) = (0, -1)$, and $c_1^* = c_2^* = 2.5$ (Figure 4).

Let us now suppose that *the noise $\xi(k)$ is governed by a random mechanism.* Namely suppose that $\xi(k)$ is a random variable uniformly distributed in $Q = \mathbf{S}$ for any $k = 1 ,\ldots, \infty$ and that all the vectors $\xi(k)$ are jointly independent.

Taking two points $\xi^{(1)} = (1,1)$, $\xi^{(2)} = (-1,-1)$, consider two sets

$$Q^{(1)}(\epsilon) = Q \bigcap(\xi^{(1)} + \mathbf{S}_\epsilon(0))$$
$$Q^{(2)}(\epsilon) = Q \bigcap(\xi^{(2)} + \mathbf{S}_\epsilon(0))$$

where

$$\mathbf{S}_\epsilon(0) = \{q : | q_i | \leq \epsilon , i = 1 , 2 \}$$

For a random sequence

$$\xi[\cdot] \in \{\xi(k), k = 1,\ldots,\infty\}$$

consider the event $A_\epsilon(k)$ that

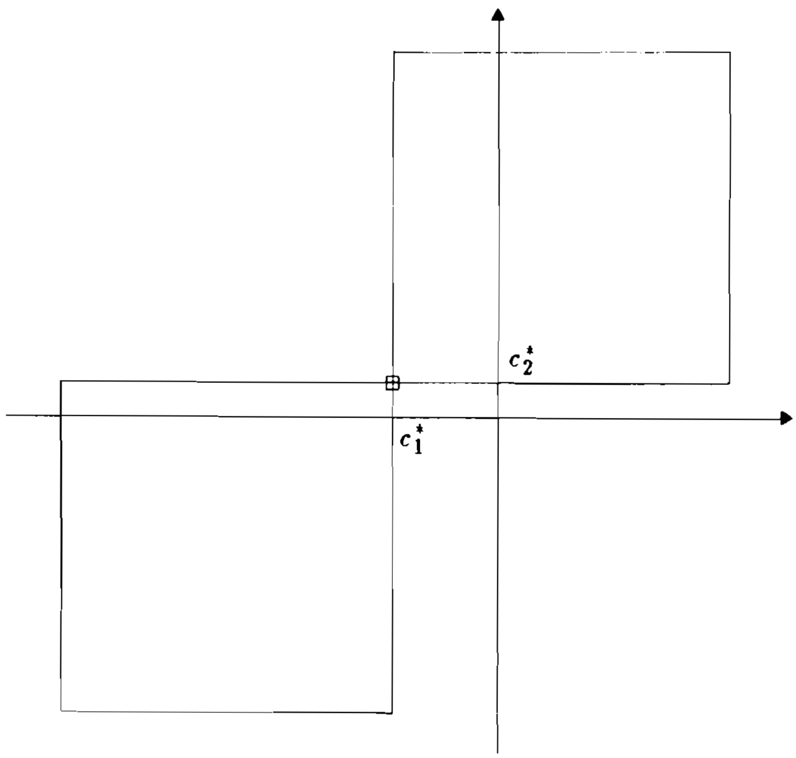

FIGURE 3

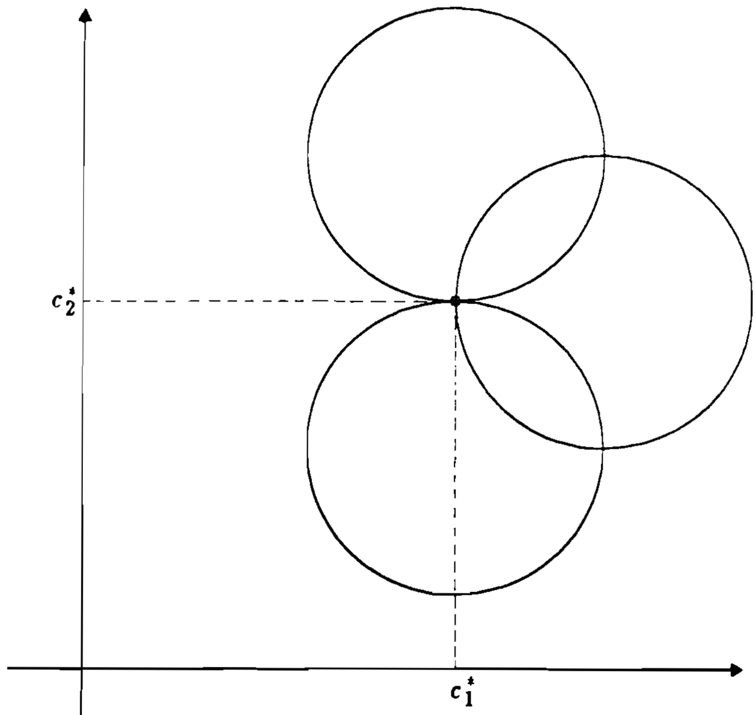

FIGURE 4

$$\xi(k) \notin Q^{(1)}(\epsilon) \bigcup Q^{(2)}(\epsilon)$$

for a given k. Denote A_ϵ to be the event that

$$\xi(k) \notin Q^{(1)}(\epsilon) \bigcup Q^{(2)}(\epsilon), \forall k$$

Then

$$A_\epsilon = \bigcap_k A_\epsilon(k)$$

and

$$P(A_\epsilon(k)) < p(\epsilon) < 1, \forall k \ .$$

Due to the joint independence of $\xi(k)$, we have

$$P(A_\epsilon) = \prod_{k=1}^{\infty} P\ (A_\epsilon(k)) = 0 \tag{1.5}$$

If we denote

$$A = \bigcup_i \{A_{\epsilon_i}\}, \ \epsilon_i > 0\ , \forall\ i = 1\ , \ldots, \infty,\ \epsilon_i \longrightarrow 0\ ,\ i \longrightarrow \infty$$

and A^c, A_ϵ^c to be the complements of A, A_ϵ, then, obviously, $A^c \subseteq A_{\epsilon_i}^c$ for any $\epsilon_i > 0$ and $A^c = \bigcap_i \{A_{\epsilon_i}^c\}$, so that

$$P(A_\epsilon^c) = 1,\ P(A^c) = \prod_i P\{A_{\epsilon_i}^c\} = 1 \tag{1.6}$$

Hence for any $\epsilon > 0$ the sequence $\xi[\cdot] = \{\xi(k),\ k = 1, \ldots, \infty\}$ will satisfy the inclusions

$$\xi(k') \in Q^{(1)}(\epsilon)$$
$$\xi(k'') \in Q^{(2)}(\epsilon)$$

with probability 1 for some $k = k'$, $k = k''$. (Otherwise, we would have $\xi[\cdot] \in A_{\epsilon_i}$).

Thus for any $\epsilon_i > 0$, for "almost all" sequences $\xi[\cdot]$ there exists an $M > 0$ (depending on the sequence) such that for $m > M$ the error set

$$R^*(m) \subseteq (Q^{(1)}(\epsilon) - Q) \bigcap (Q^{(2)}(\epsilon) - Q) = S_\epsilon(0)$$

or otherwise

$$\lim_{m \longrightarrow \infty} h\ (R^*(m)\ ,\ \{0\}) \leq \epsilon_i$$

where

$$h(R^*, \{0\}) = \max\{\|z\| \mid z \in R^*\}$$

and $\|z\|$ is the Euclidean norm of vector $z \in \mathbf{R}^2$.

It follows that with probability 1 we have

$$h(R^*(m)\ ,\ \{0\}) \underset{m \longrightarrow \infty}{\longrightarrow} 0$$

where $\{0\}$ is a singleton – the null element of $\mathbf{R}^2$.

Therefore, ***under the randomness assumptions*** of the above ***the estimation process is consistent with probability 1.*** Under the same assumptions it is clear that ***the "worst case"*** noise (1.4) ($\xi^*(k) \equiv 0\ ,\ k = 1\ ,\ldots,\ \infty$) ***may appear only with probability 0.***

The few elementary facts stated in this introduction develop into a theory of "guaranteed identification" which appears relevant to the treatment of parameter estimation, to dynamic state estimation problems, to the identification of systems with unmodelled dynamics and even to the solution of inverse problems for distributed systems [19]. It may also be propagated to the treatment of some problems for nonlinear systems [20].

The first part of the present paper deals with the simplest identification problem for a linear model describing the respective guaranteed estimates. Here the basic results are those that yield the recurrence relations for the estimates. They also lead to the discussion of the problem of consistency of the identification process.

The second part, written in a more compact form, deals with the "guaranteed" state estimation problem for discrete time linear systems with unknown but bounded inputs. This is followed by an introduction into the basic facts of "guaranteed nonlinear filtering".

The paper mainly deals with nonquadratic constraints on the unknowns. It also deals with nonlinearity and nonstationarity. This is partly done with the aim of reminding the reader that identification and state estimation problems are not merely linear-quadratic and stationary as it may seem from most of the available literature.

A special item discussed in the sequel is the relation between guaranteed and stochastic estimation procedures in the case of non-quadratic constraints on the unknowns.

2. Notations

Here we list some conventional notations adopted in this paper:

$\mathbf{R}^n$ will stand for the n-dimensional vector space, while $\mathbf{R}^{m \times n}$ - for the space of $m \times n$ - dimensional matrices, I_n will be the unit matrix of dimension n, $A \otimes B$ - the Kronecker product of matrices A , B, so that

$(A \otimes B)$ will be the matrix of the form

$$\begin{bmatrix} a_{11}B \ , & \cdots, & a_{1n}B \\ \cdots & \cdots & \cdots \\ a_{n1}B \ , & \cdots, & a_{nn}B \end{bmatrix}$$

The prime will stand for the transpose and $\bar{A}$ - for an mn - dimensional vector obtained by stacking the matrix $A = \{a^{(1)}, \ldots, a^{(n)}\}$, with columns $a^{(i)} \in \mathbf{R}^m$ $(a_j^{(i)} = a_{ij})$, so that $a_{(i-1)h+j} = a_j^{(i)}$, $(i = 1, \ldots, n)$, $(j = 1, \ldots, m)$, or in other terms

$$\bar{A} = \sum_{i=1}^{n} (e^{(i)} \otimes (A \ e^{(i)}))$$

where $e^{(i)}$ is a unit orth within $\mathbf{R}^n$ $(e_j^{(i)} = \delta_{ij}$, with δ_{ij} the Kronecker delta : $\delta_{ij} = 1$ for $i = j$, $\delta_{ij} = 0$ for $i \neq j)$.

If $\mathbf{C} = \{C\}$ is a set of $(m \times n)$-matrices C, then $\bar{\mathbf{C}}$ will stand for the respective set of mn-vectors $\bar{C} : \bar{\mathbf{C}} = \{\bar{C}\}$.

The few basic operations used in this paper are as follows:

If $<A \ , B> = tr \ AB'$ is the *inner product of matrices* A , $B \in \mathbf{R}^{m \times n}$ and $(p \ , q)$ - the *inner product* of vectors p , $q \in \mathbf{R}^n$, then for $x \in \mathbf{R}^n$, $y \in \mathbf{R}^m$ we have

$$y \otimes x' = yx' \in \mathbf{R}^{m \times n}$$
$$<A \ , y \otimes x' > = (A \ x \ , y) \qquad (2.1)$$

Other matrix equalities used here are

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

(A , B are $n \times n$ dimensional and their determinants $|A| \neq 0$, $|B| \neq 0$)

$$(A \otimes B)' = A' \otimes B'$$
$$(A \otimes B)\,\bar{K} = \overline{BKA'} \qquad (2.2)$$

A sequence of integers $i = k, \ldots, s$ will be $[k, s]$. A finite sequence of vectors $\{\xi(i) : i = k, \ldots, s\}$ will be denoted as $\xi\,[k, s]$, while an infinite one $\{\xi(i), i = s, \ldots, \infty\}$ as $\xi\,[s\,;\,\cdot]$ with $\xi\,[1, \cdot] = \xi[\cdot]$. Similar notations will be used for sequences of sets. For example $R[k, s]$ will stand for a sequence of sets $R(i)$ $k \leq i \leq s$.

Symbols *conv* $\mathbf{R}^n$ and *co* $\mathbf{R}^n$ will denote the varieties of all *convex compact* and *closed convex* subsets of $\mathbf{R}^n$ respectively,

$$\rho(\ell \mid Q) = \sup\,\{(\ell, q) \mid q \in Q\}$$

will be the *support function* of set $Q \subseteq \mathbf{R}^n$.

With $Q \in conv\ \mathbf{R}^n$ the operation of *sup* in the definition of $\rho(\ell \mid Q)$ may be substituted for *max*. Further on *int* Q will be the set of all *interior points* of Q.

$$S_r(x_0) = \{x : \|\, x - x_0 \,\| \leq r\ ;\ x, x_0 \in \mathbf{R}^n\}$$

will denote the *Euclidean ball* with center x_0 and radius r, ($\|\, x \,\| = (x, x)^{1/2}$), while $h(P, Q)$ will stand for the *Hausdorff distance* between sets $P, Q \in conv\ \mathbf{R}^n$. Namely

$$h(P, Q) = \min\{\ r : P \subseteq Q + r\,S(0)\ ,\ \ Q \subseteq P + r\,S(0)\ \}\ .$$

The symbol *epi f* stands for the *epigraph*

$$epi\ f = \{z = \{x, y\} : y \geq f(x)\ ,\ z \in \mathbf{R}^{n+1}\}$$

of function f - a subset of $\mathbf{R}^{n+1}$ and *co* Q stands for the *convex hull* of set Q with $\overline{co}\ Q$ being the *closure* of *co* Q.

For a given set $P \subseteq \Omega$ the symbol P^c will stand for the *complement* P^c of P.

The basic scheme will be first interpreted through the following "elementary" parameter estimation problem.

3. The Basic Problem

Consider a system

$$y(k) = C\,p(k) + \xi(k) \qquad (3.1)$$
$$k \in T_N = \{1 ,\ldots, N\}$$

where $y(k)$ is the *available measurement*, $p(k)$ is a *given input*, C is the *matrix parameter to be identified* and $\xi(k)$ is the *unknown disturbance*. We further assume $p \in \mathbf{R}^n$, $y \in \mathbf{R}^m$. Hence $\xi \in \mathbf{R}^m$, $C \in \mathbf{R}^{m \times n}$, (where $\mathbf{R}^{m \times n}$ stands for the space of real matrices of dimensions $m \times n$.)

The available additional information on C, $\xi\,[1\,,\,s]$ is given through restrictions on these values which are taken to be specified in advance.

The types of simple restrictions on C, $\xi\,[1\,,\,s]$ to be considered in the sequel are as follows:

$$(\bar{C}-\bar{C}^*)'\,\mathbb{L}(\bar{C}-\bar{C}^*) + \sum_{k=1}^{s} (\xi(k) - \xi^*(k))'\,N(k)(\xi(k) - \xi^*(k)) \leq 1 \qquad \text{(III.A)}$$

where $\mathbb{L} > 0$, $N(k) > 0$ ($\mathbb{L} \in \mathbf{R}^{mn \times mn}$, $N(k) \in \mathbf{R}^{m \times m}$) . (This is *the joint quadratic constraint*), or

$$(\bar{C} - \bar{C}^*)'\,\mathbb{L}(\bar{C} - \bar{C}^*) \leq 1 \qquad \text{(III.B)}$$
$$\sum_{1}^{s} (\xi\,[k] - \xi^*[k])'\,N(k)\,(\xi[k] - \xi^*[k]) \leq 1$$

which is *the separate quadratic constraint*, or

$$C \in \mathbf{C}_0\,,\ \xi(k) \in Q(k) \qquad \text{(III.C)}$$

which is *the geometrical or instantaneous constraint*. Here $\mathbf{C}_0$, $Q(k)$ are assumed to be convex and compact in $\mathbf{R}^{m \times n}$ and $\mathbf{R}^m$ respectively.

The restriction on the pair $\{C\,,\,\xi[1\,,\,s]\} = \varsigma[1\,,\,s]$ (whether given in the form (III.A), (III.B) or (III.C)) will be denoted by a unified relation as

$$\varsigma[1\,,\,N] \in \sum \qquad (3.2)$$

With measurement $y[1\ ,\ s]$ given, the aim of the solution will be to find the set of all pairs $\varsigma[1\ ,\ s]$ consistent with (3.1), (3.2) and with given $y\ [1\ ,\ s]$. More precisely the solution will be given through the notion of *the informational domain.*

Definition 3.1. The informational domain $\mathbf{C}[s] = \mathbf{C}[1\ ,\ s]$ *consistent with measurement* $y[1\ ,\ s]$ *and restriction (3.2) will be defined as the set of all matrices* C *for each of which there exists a corresponding sequence* $\xi[1\ ,\ s]$ *such that the pair* $\varsigma[1\ ,\ s] = \{C\ ,\ \xi[1\ ,\ s]\}$ *satisfies both restriction (3.2) and equation (3.1) (for the given* $y[1\ ,\ s]$ *).*

Hence the idea of the solution of the estimation problem is to find the set $\mathbf{C}[1\ ,\ s]$ of all the possible values of C each of which (together with an adequate $\xi[1\ ,\ s]$) could generate the given measurement sequence $y[1\ ,\ s]$.

It is obvious that set $\mathbf{C}[s] = \mathbf{C}[1\ ,\ s]$ now contains the unknown actual value $C = C^\circ$ which is to be estimated.

With set $\mathbf{C}[s]$ being known, one may also construct a *minmax estimate* $C_*[s]$ of C° - for example through the solution of the problem

$$\max\ \{d(C_0[s]\ ,\ Z)\ \mid\ Z \in \mathbf{C}[s]\} = \tag{3.3}$$

$$= \min_{C} \left\{ \max\{d(C\ ,\ Z)\ \mid\ Z \in \mathbf{C}[s]\},\ C \in \mathbf{C}[s] \right\} = \epsilon(s)\quad ,$$

where $d(\cdot\ ,\cdot)$ is some metric in the space $\mathbf{R}^{m\ \times\ n}$.

The element $C_0[s]$ is known as the *Chebyshev center* for set $\mathbf{C}[s]$.

Once $C_0[s]$ is specified, the estimation error $d(C_0[s]\ ,\ C^\circ) \leq \epsilon(s)$ ***is guaranteed*** by the procedure.

However, for many purposes, especially under a nonquadratic constraint (III.C), it may be convenient to describe *the whole set* $\mathbf{C}[s]$ rather than the minmax estimate $C_*[s]$.

If s varies and even $s \to \infty$ it makes sense to consider the *evolution* of $\mathbf{C}[s]$ and its *asymptotic behaviour* in which case the estimation process may turn to be *consistent,* i.e.

$$\lim_{s\ \to\ \infty} \mathbf{C}[s] = \{C^\circ\} \tag{3.4}$$

The convergence here is understood in the sense that

$$\lim_{s \to \infty} h \ (\mathbf{C}[s] \ , C^{\circ}) = 0$$

where $h(\mathbf{C}' \ , \mathbf{C}'')$ is the *Hausdorff metric* (see Introduction), and C° is a singleton in $\mathbf{R}^{m \times n}$.

In some particular cases the equality (3.4) may be achieved in a finite number s_0 of stages s when for example

$$\mathbf{C}[s] = C^{\circ}, \ s^{\circ} > 1,$$

The main discussion will be further concerned with the nonquadratic geometrical constraint (III.C). However it is more natural to start with the simplest "quadratic" restriction (III.A). In this case, as we shall see, the set $\mathbf{C}[s]$ turns to be an ellipsoid and the respective equations for $\mathbf{C}[s]$ arrive in explicit form.

4. The Joint Quadratic Constraint. Recurrence Equations

As equation (3.1) yields

$$\xi(k) = y(k) - Cp(k)$$

the set $\mathbf{C}[s]$ consists of all matrices C that satisfy (III.A), i.e.

$$(\bar{C} - \bar{C}^*)' \, \mathbb{L} \, (\bar{C} - \bar{C}^*) + \sum_{k=1}^{s} (y(k) - Cp(k) - \xi^*(k))' \, N(k)(y(k) - Cp(k) - \xi^*(k)) \le 1 \tag{4.1}$$

In view of the equality (2.2) which here turns into

$$I_m \ C \ p = (p' \otimes I_m) \ \bar{C}$$

we may rewrite (4.1) as

$$(\bar{C} - \bar{C}^*)' \, \mathbb{P}[s](\bar{C} - \bar{C}^*) - 2(\mathbb{D}[s] \ , \bar{C} - \bar{C}^*) + \gamma^2[s] \le 1$$

where

$$\mathbb{P}[s] = \mathbb{L} + \sum_{k=1}^{s} P(k)$$

$$P(k) = (p(k) \otimes I_m)\ N(k)(p'(k) \otimes I_m)$$

$$\mathbb{D}[s] = \sum_{k=1}^{s} D(k)$$

$$D'(k) = y^{*'}(k)\ N(k)\ (p'(k) \otimes I_m)$$

$$\gamma^2(s) = \sum_{k=1}^{s} y^{*'}(k)\ N(k)\ y^*(k) \tag{4.2}$$

$$y^*(k) = y(k) - C^* p(k) - \xi^*(k) \tag{4.3}$$

Hence the result is given by

Theorem 4.1. The set $\mathbf{C}[s]$ *is an ellipsoid defined by the inequality*

$$((\bar{C} - \bar{C}^* - \mathbb{P}^{-1}[s]\,\mathbb{D}[s])', \mathbb{P}[s]\,(\bar{C} - \bar{C}^* - \mathbb{P}^{-1}[s]\,\mathbb{D}[s])) \leq 1 - h^2[s] \tag{4.4}$$

with center

$$\bar{C}_o[s] = \mathbb{P}^{-1}[s]\ \mathbb{D}[s] + \bar{C}^*$$

Here

$$h^2[s] = \gamma^2(s) - (\mathbb{D}[s]\ ,\ \mathbb{P}^{-1}[s]\ \mathbb{D}[s]) \tag{4.5}$$

$$\mathbb{P}[s] = \mathbb{P}[s-1] + P(s)\ ,\ \mathbb{D}[s] = \mathbb{D}[s-1] + D(s) \tag{4.6}$$

$$\gamma^2(s) = \gamma^2(s-1) + y^{*'}(s)\ N(s)\ y^*(s)\ ,\ \gamma(0) = 0 \tag{4.7}$$

$$\mathbb{P}[0] = \mathbb{L}\ ,\ \mathbb{D}(0) = 0$$

$$\mathbb{P}^{-1}[s] = \mathbb{P}^{-1}[s-1] - \mathbb{P}^{-1}[s-1]\ G(s-1)\ K^{-1}(s-1)\ G'(s-1)\ \mathbb{P}[s-1] \tag{4.8}$$

$$G(s-1) = p(s-1) \otimes I_m$$

$$K(s-1) = N^{-1}(s-1) + G'(s-1)\ \mathbb{P}[s-1]\ G(s-1)$$

Relations (4.4) - (4.8) are *evolutionary equations* that describe the dynamics of the set $\mathbf{C}[s]$ (which is an *ellipsoid*) and its *center* $C_0[s]$ which *coincides* precisely *with the min-max estimate* $C_*[s]$ for $\mathbf{C}[s]$ (assuming $d(C\ ,\ Z)$ of (3.3) is taken to be the Euclidean metric).

Remark 6.1 A standard problem of *statistical estimation* is to find the *conditional distribution* of the values of a matrix C after s measurements due to equation (3.1) where $\xi(k)$, $k \in [1, \infty)$ are non correlated gaussian variables with given mean values $E\xi(k) = \xi^*(k)$ and covariance matrices

$$E\xi(k)\xi'(k) = N^{-1}(k)\ .$$

The initial gaussian distribution for the vector $\bar{C}$ is taken to be given with $E\bar{C} = \bar{C}^*$, $E\bar{C}\bar{C}' = \mathbb{L}^{-1}$.

A standard application of the least-square method or of some other conventional (e.g. bayesian or maximal likelihood) techniques yields an estimate

$$\bar{C}_*[s] = \mathbf{P}^{-1}[s]\mathbb{D}[s] + \bar{C}^*$$

with $\mathbf{P}[s]$, $\mathbf{D}[s]$ governed by equations (4.6), (4.8) [4]. The estimate is therefore similar to that of theorem 4.1: $\bar{C}_*[s]$ coincides with $\bar{C}_0[s]$. Here, however, the analogy ends – equations (4.5), (4.7) are specific only for the guaranteed estimates. The estimation errors for the stochastic and for the guaranteed deterministic solutions are defined through different notions and are therefore calculated through different procedures.

The next step is to specify the "worst case" and "best case" disturbances for the estimation process. From the definition (4.3) of $y^*(k)$ it is clear that if the actual values $\varsigma^\circ[1,s] = \{\xi^\circ[1,s], C^\circ\}$ for $\varsigma[1,s] = \{\xi[1,s], C\}$ are taken to be

$$\varsigma^\circ[1,s] = \varsigma^*[1,s]\ ,\ C^\circ = C^* \tag{4.9}$$

then

$$y^*[1,s] \equiv 0\ ,\ \mathbb{D}[s] \equiv 0$$

and therefore

$$h^2[s] = 0 \tag{4.10}$$

The ellipsoid $C[1,s]$ is then the "largest" possible in the sense that it includes all the ellipsoids derived through other measurements than the "worst" one

$$y_w(k) = C^*\, p(k) + \xi^*(k)\ ,\ k \in [1,s]$$

(Note that whatever are the admissible values of $y[1,s]$, all the respective ellipsoids $C[s]$ have ***one and the same center*** $C_o[s]$ and matrix $\mathbb{P}[s]$. They differ only through $h[s]$ in the right hand part of (4.4)).

The "smallest" possible ellipsoid is the one that turns to be a singleton. It is derived through the "best possible" measurement $y^{(b)}[1,s]$. The latter is defined by the pair

$$\{C^{(b)}, \xi^{(b)}[1, s]\}$$

where $C^{(b)} = C^*$ and $\xi^{(b)}[1, s]$ satisfies conditions

$$\sum_{k=1}^{s} (\xi^{(b)}(k) - \xi^*(k))' N(k)(p'(k) \otimes I_m) = 0 \tag{4.11}$$

$$\sum_{k=1}^{s} (\xi^{(b)}(k) - \xi^*(k))' N(k)(\xi^{(b)}(k) - \xi^*(k)) = 1 \tag{4.12}$$

With $C^{(b)} = C^*$ and with (4.11), (4.12) fulfillled we have

$$\begin{aligned} y(k) &= C^* p(k) + \xi^{(b)}(k) \\ y^*(k) &= \xi^{(b)}(k) - \xi^*(k) \end{aligned} \tag{4.13}$$

which yield $\mathbf{D}(k) \equiv 0$, $k \in [1, s]$ and further on, due to (4.5), (4.12), (4.11)

$$h^2[s] = \gamma^2[s] = 1$$

Hence from (4.4) it follows that $\mathbf{C}(s)$ is a singleton

$$\mathbf{C}(s) = C_o[s]$$

It is worth to observe that the set $\Xi_b(\cdot)$ of disturbances $\xi^{(b)}[1, s]$ which satisfy (4.11), (4.12) is nonvoid. Indeed, to fulfill (4.12) it suffices that $s > m$, det $N \neq 0$ and

$$(\eta_i[1, s], p_j[1, s]) = 0$$

for any $i,j \in [1, m]$. Here

$$\eta'(k) = (\xi^{(b)}(k) - \xi^*(k))' N(k)$$

Relation (4.11) defines a linear subspace $\mathbf{L}_\eta^{(k)}$ generated by vectors $\eta(k)$ and therefore also a linear subspace $\mathbf{L}_\xi$ generated by respective "vectors"

$$\bar{\xi}[1, s] = \xi^{(b)}[1, s] - \xi^*[1, s]$$

due (4.14). The required values

$$\bar{\xi}^{(b)}[1, s] = \xi^{(b)}[1, s] - \xi^*[1, s]$$

are then determined through the relation

$$\bar{\xi}^{(b)}[1, s] \in \mathbf{L}_\xi \cap \sigma_N(1)$$

where $\sigma_N(1)$ is the sphere

$$\sum_{k=1}^{s} \bar{\xi}'(k) N(k) \bar{\xi}(k) = 1$$

The last results may be given in the form of

Lemma 4.1. (a) The "worst case" estimate given by the "largest" ellipsoid $\mathbf{C}[s]$ *is generated by the measurement*

$$y_W[1\ ,\ s] = C^* p[1\ ,\ s] + \xi^*[1\ ,\ s]$$

(b) *The "best case" estimate given by a singleton* $\mathbf{C}[s] = C_0$ *is generated by the measurement*

$$y^{(b)}[1\ ,\ s] = C^*\ p[1\ ,\ s] + \xi^{(b)}[1\ ,s]$$

where $\xi^{(b)}[1\ ,\ s]$ *is any sequence* $\xi[1\ ,\ s]$ *that satisfies (4.11), (4.12).*

Case (b) indicates that *exact identifiability is possible even in the presence of disturbances.*

The terms used in the relations of the above are also relevant for exact identifiability in the absence of disturbances.

5. Exact Identifiability in the Absence of Disturbances

The equation

$$y(k) = Cp(k) \tag{5.1}$$

may be rewritten as

$$y(k) = (p'\ (k) \otimes I_m)\bar{C}$$

which yields

$$(p(k) \otimes I_m)\ N(k)\ y(k) = (p(k) \otimes I_m)\ N(k)\ (p'(k) \otimes I_m)\bar{C}$$

for $k \in [1\ ,\ s]$. This leads to equation

$$\mathbb{D}[s] = \mathbb{P}(s)\ \bar{C} \tag{5.2}$$

Hence for resolving (5.2) it suffices for the matrix $\mathbb{P}(s)$ to be invertible.

The matrix $\mathbb{P}[s]$ may be rewritten as

$$\mathbf{P}[s] = \sum_{k=1}^{s} N(k) \otimes p(k)\ p'(k) = \sum_{k=1}^{s} (p(k)\ p'(k) \otimes N(k))$$

The invertibility of $\mathbb{P}[s]$ with $N(k) = I_m$ is then ensured if $W[s] = \sum_{k=1}^{s} p(k)p'(k)$ is nonsingular.

Lemma 5.1 For the exact identifiability of matrix C in the absence of disturbances it is sufficient that

$$det\ \mathbb{P}[s] \neq 0$$

where $\mathbb{P}[s]$ is an $m^2 \times m^2$ matrix.

With $N(k) = I_m$ it is sufficient that

$$det\ W[s] \neq 0$$

where $W[s]$ is $m \times m$ dimensional.

In traditional statistics $W[s]$ is known as the *informational matrix*. We shall now proceed with the treatment of other types of constraints.

6. Separate Quadratic Constraints

Let us treat constraints (III.B) by substituting them with an equivalent system of joint constraints.

$$\begin{aligned} &\alpha\,(\bar{C} - \bar{C}^*)'\,\mathbb{L}(\bar{C} - \bar{C}^*) + \\ &+ (1-\alpha) \sum_{k=1}^{s} (\xi[k] - \xi^*[k])'\,N(k)(\xi[k] - \xi^*[k]) \leq 1 \end{aligned} \tag{6.1}$$

which should be true for any $\alpha \in (0\ ,\ 1]$.

For any given $\alpha \in (0\ ,\ 1]$, the respective domain $\mathbf{C}_\alpha[s]$ will be an ellipsoid of type (4.4) with $\mathbb{L}$ substituted for $\mathbb{L}_\alpha = \alpha\mathbb{L}$ and $N(k)$ for $N_\alpha = (1-\alpha)N(k)$. The actual domain $\mathbf{C}[s]$ for constraint (III.B) should therefore satisfy the equality

$$\mathbf{C}[s] = \{\bigcap \mathbf{C}_\alpha\,[s] \mid 0 < \alpha \leq 1\} \tag{6.2}$$

The latter formula shows that the calculations for $\mathbf{C}[s]$ may be *decoupled* into those for a series of ellipsoids governed by formulae of type (4.4)-(4.8) in which the matrices $\mathbb{L}$, $N(s)$ are substituted for $\mathbb{L}_\alpha$, $N_\alpha(s)$ respectively, each with a specific value of

$\alpha \in (0\ ,\ 1]$.

Thus each array of relations (4.4)-(4.8), $\mathbb{L} = \mathbb{L}_\alpha$, $N[1\ ,\ s] = N_\alpha[1\ ,\ s]$, produces an ellipsoid $\mathbf{C}_\alpha[s]$ that includes $\mathbf{C}[s]$. An approximation $\mathbf{C}^{(r)}\ [s]$ to $\mathbf{C}[s]$ from above may be reached through an intersection of any finite number of ellipsoids

$$\mathbf{C}^{(r)}\ [s] = \bigcap_{j=1}^{r} \mathbf{C}_{\alpha_j}[s] \tag{6.3}$$

where α_j runs through a fixed number of r preassigned values $\alpha_j \in (0\ ,\ 1]$; $j = 1\ ,\ldots,\ r$.

By intersecting over *all the values of* $\alpha \in (0\ ,\ 1]$ we will reach the *exact solution* (6.2). These facts may be summarized in

Lemma 6.1 The set $\mathbf{C}[s]$ *for constraint (6.1) may be presented as an intersection (6.2) of ellipsoids* $\mathbf{C}_\alpha[s]$ *each of which is given by relations (4.4)-(4.8) with* $\mathbb{L}$, $N\ [1\ ,\ s]$ *substituted for* $\mathbb{L}_\alpha$, $N_\alpha[1\ ,\ s]$.

Restricting the intersection to a finite number r *of ellipsoids* $\mathbf{C}_{\alpha_j}[s]$ *as in (6.3), one arrives at an approximation of* $\mathbf{C}[s]$ *from above:*

$$\mathbf{C}[s] \subseteq \mathbf{C}^{(r)}[s]\ .$$

It is not difficult to observe that for obtaining the exact solution $\mathbf{C}[s]$ it suffices to have only a denumberable sequence of values α_j , $j = 1\ ,\ldots,\ \infty$.

The relations given here are trivial. However they indicate that the calculation of $\mathbf{C}[s]$ may be done by ***independent parallel calculations*** for each of the ellipsoids $\mathbf{C}_\alpha\ [s]$.

This suggestion may be further useful for the more complicated and less obvious problems of the sequel.

Another option is to approximate $\mathbf{C}[s]$ by a *polyhedron*. This may require the knowledge of the projections of set $\mathbf{C}[s]$ on some preassigned directions $\ell^{(i)} \in \mathbf{R}^n$.

Since $\mathbf{C}[s]$ is obviously a convex compact set, it may also be described by its ***support function***, [21]

$$\rho(\ell \mid \mathbf{C}[s]) = \max\ \{(\ell\ ,\ \bar{C}) \mid \bar{C} \in \bar{\mathbf{C}}[s]\},\ \ell \in \mathbf{R}^{mn}\ ,$$

Denote

$$f(\ell) = inf\,\{\rho(\ell \mid \bar{\mathbf{C}}_\alpha[s]) \mid \alpha \in (0\,,\,1]\}$$

The function $f(\ell)$, being positively homogeneous, may turn to be ***nonconvex***.

We may convexify it by introducing $(co\ f)(\ell)$ - a closed convex function such that

$$\overline{co}\ (epi\ f) = epi\ (co\ f).$$

The support function may now be calculated as follows.

Theorem 6.1 Assume $f(0) = 0$. *Then* $\rho(\ell \mid \bar{\mathrm{C}}[s]) = (co\ f)\ (\ell)$.

The function $f(\ell)$ defines a convex compact set $\mathbf{C}[s]$ as one that consists of all those $\bar{C} \in \mathbf{R}^{mn}$ that satisfy

$$(\ell\,,\,\bar{C}) \leq f(\ell)\ ,\ \forall\ \ell \in \mathbf{R}^{mn} \tag{6.4}$$

or in other words

$$\mathbf{C}[s] = \{\,C : (\ell\,,\,\bar{C}) \leq \rho(\ell \mid \bar{\mathbf{C}}_\alpha\,[s])\ ,\ \forall\ \alpha \in (0\,,\,1],\, \ell \in \mathbf{R}^n\,\}$$

However (6.4) is equivalent to

$$(\ell\,,\,C) \leq (co\,f)\ (\ell)\ ,\ \forall\ \ell \in \mathbf{R}^{mn}$$

according to the definition of *co f*. Being closed, convex and positively homogeneous, *co f* turns to be the support function for $\mathbf{C}[s]$.

This result shows that provided $\mathbf{C}[s]$ is nonvoid, $(f(0) = 0)$, the function $\rho(\ell \mid \bar{\mathrm{C}}[s])$ may be estimated through a direct minimization of $\rho(\ell \mid \bar{\mathrm{C}}_\alpha[s])$ over α - rather than through the procedure of calculating the "infimal convolution" of the supports $\rho(\ell \mid \bar{\mathrm{C}}_\alpha[s])$ as required by conventional theorems of convex analysis.

The knowledge of $\rho(\ell \mid \bar{\mathrm{C}}[s])$ allows to construct some ***approximations*** from above for $\mathbf{C}[s]$. Taking, for example r directions $\ell^{(i)} \in \mathbf{R}^{mn}$, $(i = 1\ ,\ldots\ r)$ we may solve optimization problems in $\alpha \in (0\,,\,1]$: $\rho_i[s] = inf\,\{\,\rho(\ell^{(i)} \mid \bar{\mathrm{C}}_\alpha[s]) \mid \alpha \in (0\,,\,1]\,\}$

Denoting

$$L_i[s] = \{C : (\ell^{(i)}\,,\,\bar{C}) \leq \rho_i[s]\}$$

we may observe

$$\mathbf{C}[s] \subseteq \{ \bigcap L_i[s] \mid 1 \leq i \leq r \} = \mathbf{L}_r[s]$$

Where $\mathbf{L}_r[s]$ is *an mn-dimensional polyhedron with r faces.*

7. Geometrical Constraints

Returning to equation (3.1) assume that the restrictions on $\xi(k)$ and C that are given in advance are taken to be *geometrical* (i.e. of type III (C)). Namely

$$\xi(k) \in Q(k) \ , \ k \in [1 \ , s] \tag{7.1}$$

$$C \in \mathbf{C}_o \tag{7.2}$$

where $Q(k)$, $\mathbf{C}_o$ are convex compact sets in $\mathbf{R}^m$ and $\mathbf{R}^{m \times n}$ respectively. The *informational set* $\mathbf{C}[s]$ will now consist of all those matrices C that satisfy (7.2) and also generate the measured value $y[1 \ , s]$ together with some disturbance $\xi[1 \ , s]$ that satisfies (7.1).

Using standard techniques of convex analysis and matrix algebra we come to the following sequence of operations.

The system equations (3.1), (7.1) may be transformed into

$$y(k) \in (p'(k) \otimes I_m)\bar{C} + Q(k) \ ;$$

since $I_m \ C \ p = (p' \otimes I_m) \ \bar{C}$ according to (2.2).

The set $\mathbf{C}[s]$ will then consist of all matrices C such that for every $k \in [1 \ , s]$ we have

$$\psi'(k)(p'(k) \otimes I_m) \ \bar{C} \leq (\psi(k) \ , y(k)) + +\rho(\psi(k) \mid - \ Q(k)) \ , \tag{7.3}$$

together with

$$(\bar{\Lambda} \ , \bar{C}) \leq \rho(\bar{\Lambda} \mid \mathbf{C}_o) \tag{7.4}$$

for any $\psi(k) \in \mathbf{R}^m$, $\bar{\Lambda} \in \mathbf{R}^{mn}$.

(Recall that symbol $\rho(\psi \mid Q)$ stands for the value of the support function

$$\rho(\psi \mid Q) = sup \ \{(\psi \ , q) \mid q \in Q\}$$

of the set Q at point ψ.)

This leads to the inequality

$$\sum_{k=1}^{s} \psi'(k)(p'(k) \otimes I_m)\bar{C} + (\bar{\wedge}\ , \bar{C}) \le$$
$$\le \sum_{k=1}^{s} \{(\psi(k)\ , y(k)) + \rho(\psi(k) \mid -\ Q(k))\} + \rho(\bar{\wedge} \mid \bar{\mathbf{C}}_o)$$

for any $\psi(k) \in \mathbf{R}^m$, $\bar{\wedge} \in \mathbf{R}^{mn}$

Therefore, with $\bar{\wedge} \in \mathbf{R}^{mn}$ given we have*

$$(\bar{\wedge}\ , \bar{C}) \le \rho(\bar{\wedge}' - \sum_{k=1}^{s} \psi'(k)(p'(k) \otimes I_m) \mid \bar{\mathbf{C}}_o) + \tag{7.5}$$
$$+ \sum_{k=1}^{s} ((\psi(k)\ , y(k)) + \rho(\psi(k) \mid -\ Q(k)))$$

For an element $C \in \mathbf{C}[s]$ it is necessary and sufficient that relation (7.5) is true for any $\psi(k) \in \mathbf{R}^m$, $k \in [1\ , s]$.

Hence we come to

Lemma 7.1. The informational set $\mathbf{C}[s]$ *consistent with measurement* $y[1\ , s]$ *and with restrictions (7.1), (7.2) is defined by the following support function.*

$$\rho(\wedge \mid \mathbf{C}\ [s]) = f(\wedge) \tag{7.6}$$

where

$$f(\wedge) = inf\ \{\rho(\bar{\wedge}' - \sum_{k=1}^{s} \psi'(k)(p'(k) \otimes I_m) \mid \mathbf{C}_o) +$$
$$+ \sum_{k=1}^{s} \psi'(k)\ y(k) + \rho(\psi(k) \mid -\ Q(k)) \mid \psi(k) \in \mathbf{R}^m\ , k = [1\ , s]\ \}$$

The proof of Lemma 7.1 follows from (7.5) and from the fact that $f(\wedge)$ is a convex, positively homogeneous function, [21].

A special case arrives when there is no information on C at all and therefore $\mathbf{C}_o = \mathbf{R}^{m \times n}$. Following the previous schemes we come to

---

* When using the symbol $\rho(p \mid Q)$ for the support function of set Q at point p we will not distinguish a vector-column p from a vector-row p'.

Lemma 7.2. Under restrictions (7.1), $\mathbf{C}_0 = \mathbf{R}^{m \times n}$, *the set* $\mathbf{C}[s]$ *is given by the support function.*

$$\rho(\wedge \mid \mathbf{C}[s]) = \tag{7.7}$$
$$= inf\left\{\sum_{k=1}^{s} \{\rho(-\psi(k) \mid Q(k)) + \psi'(k)\, y(k)\}\right\}$$

over all vectors $\psi(k)$ that satisfy

$$\sum_{k=1}^{s} \psi'(k)\,(p'(k) \otimes I_m) = \bar{\wedge}' \tag{7.8}$$

A question may however arise which is whether in the last case the set $\mathbf{C}[s]$ will be bounded.

Lemma 7.3. Suppose $\mathbf{C}_0 = \mathbf{R}^{m \times n}$ *and the matrix* $\{p(1), \ldots, p(s)\} = P(s)$ *for* $s \geq n$ *is nonsingular. Then the set* $\mathbf{C}[s]$ *is bounded.*

Taking equation (7.8) it is possible to solve it in the form

$$\psi(k) = (p'(k) \otimes I_m)\,(I_m \otimes W(s))^{-1}\,\bar{\wedge} \tag{7.9}$$

where as before

$$W[s] = \sum_{k=1}^{s} (I_m \otimes p(k))(p'(k) \otimes I_m)$$

Indeed (7.8) may be transposed into

$$\sum_{k=1}^{s} (I_m \otimes p(k))\psi(k) = \bar{\wedge} \tag{7.10}$$

and the solution may be sought for in the form

$$\psi(k) = (p'(k) \otimes I_m)\ell \tag{7.11}$$

In view of (7.8) this yields equation

$$(I_m \otimes W[s])\ell = \bar{\wedge} \tag{7.12}$$

where the matrix $W[s]$ is invertible (the latter condition is ensured by the linear independence of vectors $p(k)$, $k = 1 \ldots s$, $s \geq n$). Equations (7.10)–(7.11) produce the solution (7.9).

Substituting $\psi(k)$ of (7.9) into (7.7) it is possible to observe that the support function $\rho(\wedge \mid \mathbf{C}[s])$ is equibounded in $\bar{\wedge}$ over all $\bar{\wedge} \in \mathbf{S}_1^{mn}(0)$ where $\mathbf{S}_1^{mn}(0)$ is a unit ball in $\mathbf{R}^{mn}$. This proves the boundedness of $\mathbf{C}[s]$.

Remark 7.1 Assuming that $\xi[s]$ is bounded by a quadratic constraint (III.B) with $\mathbb{L} = 0$ (so that there is no initial bound on C), and that $P(s)$ is nonsingular, the set $\mathbf{C}[s]$ again remains bounded.

The result of Lemma 7.3 therefore remains true when the geometrical constraint on $\xi[k]$ is substituted by a quadratic constraint on $\xi[\cdot]$. It is not difficult to observe that the result still remains true when $\xi[\cdot]$ is bounded in the metric of space ℓ_p:

$$\sum_{i=1}^{s} ((\xi[k] - \xi^*[k])' N(k)(\xi[k] - \xi^*[k]))^{p/2} \leq 1$$

with $1 \leq p \leq \infty$,

8. Recurrence Equations for Geometrical Constraints

One could already observe that equations (4.4)-(4.8) of theorem 4.1 are given in a recurrent form so that they would describe the evolution of the set $\mathbf{C}[s]$ that estimates the unknown matrix C. The next step will be to derive recurrence evolution equations for the case of geometrical constraints.

Starting with relation (7.5), substitute

$$\psi'(k) = \bar{\wedge}' M(k)$$

where $M(k) \in \mathbf{R}^{mn \times m}$, $1 \leq k \leq s$.

Then (7.5) will be transformed into the following inequality

$$(\bar{\wedge} , \bar{C}) \leq \rho(\bar{\wedge}' \mid (I_{mn} - \sum_{k=1}^{s} M(k)(p'(k) \otimes I_m))\bar{\mathbf{C}}_o) + \tag{8.1}$$
$$+ \sum_{k=1}^{s} (\bar{\wedge}' , M(k)\, y(k)) + \rho(\bar{\wedge} \mid M(k)(-Q(k)))$$

Denote the sequence of matrices $M(k) \in \mathbf{R}^{mn \times m}$, $k \in [1,\ldots,s]$ as $M[1\ ,\ s]$.

Lemma 8.1 In order that $C \in \mathbf{C}[s]$ *it is necessary and sufficient that (8.1) would hold for any* $\bar{\wedge} \in \mathbf{R}^{mn}$, *and any sequence* $M[1\ ,\ s] \in \mathbf{M}[1\ ,\ s]$.

The proof is obvious from (7.5), (8.1) and Lemma 7.1. Hence in view of the properties of support functions for convex sets we come to the following assertion.

Lemma 8.2 In order that the inclusion

$$C \in \mathbf{C}[s]$$

would be true it is necessary and sufficient that

$$\bar{C} \in C(s\ ,\ \bar{\mathbf{C}}_0\ ,\ M[1\ ,\ s])$$

for any sequence $M[1\ ,\ s] \in \mathbf{M}[1\ ,\ s]$ *where*

$$C(s\ ,\ \bar{\mathbf{C}}_0\ ,\ M[1\ ,\ s]) = (I_{mn} - \sum_{k=1}^{s} M(k)\ (p'(k) \otimes I_m))\ \bar{\mathbf{C}}_0 + \\ + \sum_{k=1}^{s} M(k)\ (y(k) - Q(k))$$

From Lemma 8.2 it now follows

Lemma 8.3. The set $\mathbf{C}[s]$ *may be defined through the equality*

$$\bar{\mathbf{C}}[s] = \bigcap \{\, C(s\ ,\ \bar{\mathbf{C}}_0\ ,\ M[1\ ,\ s])\ |\ M[1\ ,s] \in \mathbf{M}[1\ ,\ s]\ \}$$

In a similar way, assuming the process starts from set $C[s]$ at instant s, we have

$$\bar{\mathbf{C}}[s + 1] \subseteq (I_n - M(s + 1)\,(p'(s + 1) \otimes I_m))\ \bar{\mathbf{C}}[s] + \\ + M(s + 1)(y(s + 1) - Q(s + 1)) = C(s + 1\ ,\ \bar{\mathbf{C}}[s]\ ,\ M(s + 1)) \tag{8.2}$$

for any $M(s + 1) \in \mathbf{R}^{mn \times n}$ and further on

$$\bar{\mathbf{C}}[s + 1] = \bigcap \{\, C(s + 1\ ,\ \bar{\mathbf{C}}[s]\ ,\ M)\ |\ M \in \mathbf{R}^{mn \times n}\} \tag{8.3}$$

This allows us to formulate

Theorem 8.1 The set $\mathbf{C}[s]$ *satisfies the recurrence inclusion*

$$\bar{\mathbf{C}}\,[s + 1] \subseteq C(s + 1\ ,\ \bar{\mathbf{C}}[s]\ ,\ M),\ \mathbf{C}[0] = \mathbf{C}_0 \tag{8.4}$$

- whatever is the matrix $M \in \mathbf{R}^{mn \times n}$ *- and also the recurrence equation (8.3).*

The relations of the above allow to construct numerical schemes for approximating the solutions to the guaranteed identification problem.

Particularly, (8.4) may be decoupled into a variety of systems

$$\bar{\mathbf{C}}_M\,[s+1] \subseteq C(s+1\,,\bar{\mathbf{C}}_M[s]\,,M(s))\,,\mathbf{C}[0]=\mathbf{C}_0 \tag{8.5}$$

each of which depends upon a sequence $M[1\,,s]$ of "decoupling parameters". It therefore makes sense to consider

$$\mathbf{C}_U\,[s] = \{\bigcap \mathbf{C}_M[s] \mid M[1\,,s]\} \tag{8.6}$$

Obviously $\mathbf{C}\,[s] \subseteq \mathbf{C}_U\,[s]$

From the linearity of the right-hand side of (8.2) and the convexity of sets $\mathbf{C}_0$, $Q(s)$ it follows that actually $\mathbf{C}[s] = \mathbf{C}_U[s]$.

Lemma 8.4 The set $\mathbf{C}[s] = \mathbf{C}_U[s]$ *may be calculated through an intersection (8.6) of solutions* $\mathbf{C}_M[s]$ *to a variety of independent inclusions (8.5) parametrized by sequences* $M[1\,,s]$.

This fact indicates that $\mathbf{C}[s]$ may be reached by *parallel computations* due to equations (8.5). The solution to each of these equations may further be substituted by approximative set-valued solutions with ellipsoidal or polyhedral values. The precise techniques for these approximations however lie beyond the scope of this paper.

An important question to be studied is whether the estimation procedures given here may be consistent. It will be shown in the sequel that there exist certain classes of identification problems for which the answer to this question is affirmative.

9. Geometrical Constraints. Consistency Conditions

We will discuss this problem assuming $\mathbf{C}_0 = \mathbf{R}^{m \times n}$. Then the support function $\rho(\wedge) \mid \mathbf{C}[s])$ for set $\mathbf{C}[s]$ is given by (7.7), (7.8).

The measurement $y(k)$ may be presented as

$$y(k) = (p'(k) \otimes I_m)\, \bar{C}^* + \xi^*(k)\ , \ (k = 1, \ldots, s) \tag{9.1}$$

where $\bar{C}^*$ is the actual vector to be identified, $\xi^*(k)$ is the unknown actual value of the disturbance.

Substituting (9.1) into (7.7), (7.8) we come to

$$\rho(\wedge \mid \mathbf{C}[s]) =$$
$$= inf \left\{ \sum_{k=1}^{s} \rho(\psi(k) \mid \xi^*(k) - Q(k)) + \right.$$
$$\left. + \sum_{k=1}^{s} \psi'(k)(p'(k) \otimes I_m)\, \bar{C}^* \right\},$$

over all vectors $\psi(k)$ that satisfy

$$\psi[1\ , s] \in \Psi[s\ , \wedge] \tag{9.2}$$

where

$$\Psi[s\ , \wedge] = \{\psi[1\ , s] : \sum_{k=1}^{s} \psi'(k)(p'(k) \otimes I_m) = \bar{\wedge}'\}$$

This is equivalent to

$$\rho(\wedge \mid \mathbf{C}[s]) = (\bar{\wedge}, \bar{C}^*) + \rho(\wedge \mid R^*[s])\ ,$$

where

$$\rho(\wedge \mid R^*[s]) =$$
$$= inf \left\{ \sum_{k=1}^{s} \rho(\psi(k) \mid \xi^*(k) - Q(k)) \mid \psi[1\ , s] \in \Psi[s\ , \wedge] \right\} = \varphi(\wedge) \tag{9.3}$$

In other terms

$$\bar{\mathbf{C}}[s] \subseteq \bar{C}^* + R^*[s]$$

where $R^*[s]$ is the ***error set*** for the estimation process. The support function for $R^*[s]$ is given by (9.3).

Since $\xi^*(k) \in Q(k)$ we have

$$\rho(\bar{\wedge} \mid R^*[s]) \geq 0 \;, \forall \; \bar{\wedge} \in \mathbf{R}^{m \times n}$$

Hence every sequence $\psi^0 \; [1 \; , \; s] \in \Psi \; (s \; , \wedge)$ that yields

$$\sum_{k=1}^{s} \rho(\psi(k) \mid \xi^*(k) - Q(k)) = 0$$

will be a minimizing element for problem (9.3).

The estimation process will be consistent within the interval $[1 \; , \; s]$ if

$$R^*[s] = \{0\}$$

or, in other terms, if

$$\rho(\wedge \mid R^*[s]) = 0 \;, \forall \wedge \in \mathbf{R}^{m \times n} \tag{9.4}$$

Lemma 9.1 In order that $\rho(\bar{\wedge} \mid R^*[s]) = 0 \;, \forall \; \bar{\wedge} \in \mathbf{R}^{m \times n}$ *it is necessary and sufficient that there would exist* $mn + 1$ *vectors* $\bar{\wedge}^{(i)} \in \mathbf{R}^{mn} \;, \; i = 1 \;, \ldots, \; mn$*, such that*

$$\sum_{i=1}^{mn+1} \alpha_i \; \bar{\wedge}^{(i)} \neq 0 \;, \; \{\forall \; \alpha : (\alpha \;, \; \alpha) \neq 0 \;, \; \alpha_i \geq 0 \;, \; \forall \; i \in [1 \;, \ldots, \; mn + 1]\} \tag{9.5}$$

$$(\alpha = \alpha_1 \;, \ldots, \alpha_{mn+1})$$

and

$$\rho(\wedge^{(i)} \mid R^*[s]) = 0 \;, \; \forall \; i \in [1 \;, \ldots, \; mn + 1]$$

Vectors $\bar{\wedge}^{(i)}$ that satisfy (9.5) are said to form a *simplicial basis* in $\mathbf{R}^{mn}$.

Every vector $\bar{\wedge} \in \mathbf{R}^{mn}$ may then be presented as

$$\bar{\wedge} = \sum_{i=1}^{mn+1} \alpha_i \; \bar{\wedge}^{(i)} \;, \; \alpha_i \geq 0$$

Hence for any $\bar{\wedge} \in \mathbf{R}^{mn}$ we have

$$\rho(\wedge \mid R^*[s]) = \rho\left(\sum_{i=1}^{mn+1} \alpha_i \; \bar{\wedge}^{(i)} \mid R^*[s]\right) \leq$$

$$\leq \sum_{i=1}^{mn+1} \alpha_i \rho(\bar{\wedge}^{(i)} \mid R^*[s]) = 0$$

In view of (9.4) this yields $R^*[s] = \{0\}$.

We will now indicate some particular classes of problems when the inputs and the disturbances are such that they ensure the conditions of Lemma 9.1 to be fulfilled.

Condition 9.A

(i) *The disturbances $\xi^*(k)$ are such that they satisfy the equalities*

$$(\xi^*(k)\ ,\ \psi^*(k)) = \rho(\psi^*(k)\ |\ Q(k)) \tag{9.6}$$

for a certain r-periodic function $\psi^*(k)\ (r \geq m)$ that yields

Rank $\{\psi^*(1)\ ,\ldots,\ \psi^*(r)\} = m$.

(ii) *The input function $p(k)$ is q-periodic with $q \geq n + 1$*

Among the vectors $p(k)$, $(k = 1\ ,\ldots,\ q)$ one may select a simplicial basis in $\mathbf{R}^n$, i.e. for any $x \in \mathbf{R}^n$ there exists an array of numbers $\alpha_k \geq 0$ such that

$$x = \sum_{k=1}^{q} \alpha_k\ p(k)$$

(iii) *Numbers r and q are relative prime.*

Lemma 9.2 Under Condition 9.A the error set $R^*[s] = 0$ for $s \geq rq$.

We will prove that $R^*[s_0] = 0$ for $s_0 = rq$. The condition $R^*[s] = 0$ for $s \geq s_0$ will then be obvious.

Due to (9.3), the objective is to prove that under Condition 9.A there exists for every $\wedge \in \mathbf{R}^{m \times n}$ a set of vectors $\psi^0(k)$, $k = 1\ ,\ldots,\ s^0$, such that

$$\sum_{k=1}^{s_0} \rho(\psi^0(k)\ |\ \xi^*(k)\ -\ Q(k)) = 0\ , \tag{9.7}$$

$$\psi^0\ [1\ ,\ s_0] \in \Psi[s_0\ ,\ \wedge]\ .$$

Condition 9.A implies that there exists such a one-to-one correspondence $k = k(i\ ,\ j)$ between pairs of integers $\{i\ ,\ j\}$ $(i \in [1\ ,\ldots,\ r]\ ,\ j \in [1\ ,\ldots,\ q])$ and integers $k \in [1\ ,\ldots,\ s_0]$ that

$$p(k) = p(i)\ ,\ \psi(k) = \psi(j) \tag{9.8}$$

Indeed, if k^* is given, then it is possible to find a pair i^* , j^*, so that

$$k^* = i^* + \gamma r \ , \ k^* = j^* + \sigma q \ ,$$

where γ , σ are integers. Then we assume $p(k^*) = p(i^*)$, $\psi(k^*) = \psi(j^*)$.

The latter representation is unique in the sense that pair i^* , j^* may correspond to no other number k^{**} than k^*.

(If, on the contrary, there would exist a $k^{**} \geq k^*$ such that

$$k^{**} = i^* + \gamma_0 r \ , \ k^{**} = j^* + \sigma_0 q \ ,$$

then we would have

$$k^{**} - k^* = (\gamma_0 - \gamma) r$$
$$k^{**} - k^* = (\sigma_0 - \sigma) q$$

and $k^{**} - k^*$ would be divided by $s_0 = rq$ without a remainder. Since $k^{**} - k^* < s_0$, it follows that $k^{**} = k^*$).

As the number of pairs $\{i \ , j\}$ is so and as each pair $\{i \ , j\}$ corresponds to a unique integer $k \in [1 \ , s_0]$, the function $k = k(i \ , j)$ is a one-to-one correspondence.

Thus if $\wedge \in \mathbf{R}^{m \times n}$ and sequence $\psi^* \ [1 \ , s]$ satisfies Condition 9.A (i), then there exists a sequence $x[1 \ , s_0]$, $(x(k) \in \mathbf{R}^n)$, such that

$$\sum_{i=1}^{r} \psi^*(i) \ x'(i) = \wedge$$

Due to Condition 9.A (ii)

$$x(i) = \sum_{j=1}^{q} \alpha_{ij} \ p(j)$$

for some values $\alpha_{ij} \geq 0$.

Therefore

$$\sum_{i=1}^{r} \ \sum_{j=1}^{q} \alpha_{ij} \ \psi^*(i) \ p'(j) = \wedge \tag{9.9}$$

Assigning to every pair $\{i \ , j\}$ the value $k = k(i \ , j)$ we may renumerate the values α_{ij} with one index, substituting ij for $k = k(i \ , j)$. Having in mind (9.8), we may rewrite (9.9) as

$$\sum_{k=1}^{s_0} \alpha_k \, \psi^*(k) \, p'(k) = \wedge \tag{9.10}$$

The transition from (9.9) to (9.10) is unique. Hence, for each $\wedge \in \mathbf{R}^{m \times n}$ there exists a sequence $\alpha[1 \, , \, s_0]$ of nonnegative elements $\alpha_k \geq 0$ such that

$$\sum_{k=1}^{s_0} \alpha_k \, \psi^{*\prime}(k) \, (p'(k) \otimes I_m) = \bar{\wedge}' \tag{9.11}$$

Substituting $\psi^0(k) = \alpha_k \psi^*(k)$ and taking into account equalities (9.6) we observe that (9.7) is fulfilled. Namely

$$\sum_{k=1}^{s_0} \rho(\alpha_k \, \psi^*(k) \mid \xi^*(k) - Q(k)) = 0$$

while (9.11) yields $\psi^0 \, [1 \, , s] \in \Psi[s_0 \, , \wedge]$. Lemma 9.2 is thus proved.

A second class of problems that yield consistency is described by

Condition 9.B.

(i) *function $p(k)$ is periodic with period $q \leq n$. The matrix $W[q] = \sum_{k=1}^{q} p(k) \, p'(k)$ is nonsingular,*

(ii) *the disturbances $\xi(k)$ are such that if $\{\bar{\wedge}^{(i)}\}$, $i = 1 , \ldots , mn + 1$ is a given simplicial basis in $\mathbf{R}^{mn}$ and vectors $\psi^{(i)}(k) \in \mathbf{R}^m$ are those that yield*

$$\sum_{k=1}^{q} \psi^{(i)\prime}(k)(p'(k) \otimes I_m) = \bar{\wedge}^{(i)} \tag{9.12}$$

then the sequence $\xi(j)$, $j = 1 , \ldots , q(mn + 1)$ does satisfy conditions

$$(\xi(k + i) \, , \, \psi^{(i)}(k)) = \rho(\psi^{(i)}(k) \mid Q(k)) \tag{9.13}$$
$$(k = 1 , \ldots , q \; ; \; i = 1 , \ldots , m \, (n + 1))$$

Lemma 9.3 Under Condition 9.B the set $R[s] = \{0\}$ for $s \geq q(mn + 1)$

The proof of this Lemma follows from Lemma 7.1 and from direct substitution of (9.12), (9.13) into (9.3) (since the required set of vectors $\psi^{(i)}(k)$ does always exist due to condition $\mid W(q) \mid \neq 0$)

A simple particular case when Lemma 9.3 works is when C is a vector $(C \in \mathbf{R}^n)$ and when the restriction on $\xi(k)$ is $|\ \xi(k)\ | \leq \mu$.

Then $\wedge^{(i)} \in \mathbf{R}^n$ and (9.12) turns into

$$\sum_{k=1}^{q} \psi^{(i)}(k)\ p'\ (k) = \wedge^{(i)}$$

where $\psi^{(i)}(k)$ are scalars.

Relations (9.13) now yield

$$\xi(k + i) = \mu\ sign\ \psi^{(i)}(k) \tag{9.14}$$

Therefore the "best" disturbance $\xi(j) = \pm\mu$ now depends only upon the signs of $\psi^{(i)}(k)$, $j = i + k$. Here *the order of pluses and minuses is predetermined* by relation (9.14). However a natural question does arise. This is whether the consistency condition would still hold (at least asymptotically, with $h\,(R\,[s]\ ,\ \{0\}) \longrightarrow 0\ ,\ s \longrightarrow \infty))$ if $\xi(j)$ would attain its values at random.

The answer to the last question is given below.

Condition 9.C

(i) *function* $p(k)$, $k = 1\ ,\ldots,\ \infty$, *is periodic with period* $q \leq n$; *the matrix* $W(q)$ *is non-singular.*

(ii) *the sequence* $\xi(i)$ *is formed of jointly independent random variables with identical nondegenerate probabilistic densities, concentrated on the set*

$$Q(k) \equiv Q\ ,\ Q \in comp\ \mathbf{R}^m\ ,\ int\ Q \neq \emptyset$$

Condition (ii) means in particular that for every convex compact subset $Q_\epsilon \subseteq Q$, $(Q_\epsilon \in comp\ \mathbf{R}^m)$ of measure $\epsilon > 0$ the probability

$$P\{\xi(k) \in Q_\epsilon\} = \delta > 0\ ,\ \forall\ k \in [1\ ,\ \infty]$$

At the same time it will not be necessary for values of the distribution densities of the variables $\xi(i)$ to be known.

Lemma 9.4 Under Condition 9.C the relation

$$h(R^*[s]\ ,\ \{0\}) \longrightarrow 0\ ,\ s \longrightarrow \infty$$

holds with probability 1.

We will prove that for every $\epsilon > 0$ with probability 1 for a sequence $\xi[\,\cdot\,]$ there exists a number $N > 0$ such that for $s \geq N$ one has

$$h(R^*[s]\ ,\ \{0\}) \leq \epsilon \tag{9.15}$$

Since $W(q)$ is nonsingular, there exists for a given $\wedge \in \mathbf{R}^{m \times n}$ a sequence $\psi^0\ [1\ ,\ q]$ such that

$$\sum_{k=1}^{q} \psi^0(k)\ p'(k) = \wedge$$

Let $\xi^0(k) \in Q$ denote a respective sequence of elements that satisfy the relations

$$(\xi^0(k)\ ,\ \psi^0(k)) = \rho(\psi^0(k)\ |\ Q) \tag{9.16}$$

It is clear that elements $\xi^0(k)$ belong to the *boundary* ∂Q of set Q. Without loss of generality we may assume that all the vectors $\xi^0(k)$ are chosen among the *extremal points* of Q.

(A point $\xi^0 \in Q$ is said to be *extremal* for Q if it cannot be presented in the form

$$\xi^0 = \alpha\ \xi^{(1)} + (1 - \alpha)\ \xi^{(2)}\ , 0 < \alpha < 1\ ,$$

for any pair of elements $\xi^{(1)}\ ,\ \xi^{(2)} \in Q$.)

Hence each $\xi^0(k)$ of (9.16) is either already extremal - if (9.16) gives a unique solution, - or could be chosen among the extremal points for set $\Xi_\wedge = \{\xi : (\xi\ ,\ \psi^0(k)) = \rho(\psi^0(k)\ |\ Q)$ which yields extremality of $\xi^0(k)$ relative to Q).

Consider a sequence of Euclidean balls $S_\delta\ (\xi^0(k))$ with centers at $\xi^0(k)$ and radii $\delta > 0$. Denote

$$Q_\delta(k) = Q \bigcap S_\delta(\xi^0(k))$$

Then with *int* $Q \neq \emptyset$ the measure $\mu(Q_\delta(k)) > 0$ for any $\delta > 0$.

Let us consider q infinite sequences

$$\xi(qj + k) \ , \qquad (9.17)$$
$$(j = 0 , \ldots , \infty ; \ k = 1 , \ldots , q)$$

generated by the "noise" variable $\xi(i)$.

Denote $A_\delta(k)$ to be the event that

$$\xi \ (qj + k) \notin Q_\delta(k) \ , \ (j = 1 , \ldots , \infty)$$

and

$$A(k) = \bigcup \{ A_{\delta_i}(k) \mid \delta_i > 0 \ , \ \delta_i \longrightarrow 0 \ , \ i \longrightarrow \infty \}$$

Then obviously $P(\xi \, [\cdot] \in A_{\delta_i}(k)) = 0$ for any $\delta_i > 0$ (due to the joint independence of the variables $\xi(i)$) and due to a Lemma by Borel and Cantelli [22] we have (for any $k = 1 , \ldots , q$)

$$P(\xi \, [\cdot] \in A^c(k)) = 1$$

Hence with probability 1 for a sequence $\xi[\cdot]$ there exists a number $j(k)$ such that

$$\xi(qj(k) + k) \in Q_\delta(k) \qquad (9.18)$$

Denoting $\bigcap_{k=1}^{q} \bar{A}(k) = B$, we observe

$$P(\xi \, [\cdot] \in B) = P(\xi[\cdot] \in \bigcap_{k=1}^{q} A^c(k)) = \qquad (9.20)$$
$$= \prod_{k=1}^{q} P(\xi \, [\cdot] \in A^c \, (k)) = 1$$

due to the joint independence of the random variables $\xi(i)$.

Hence each sequence $\xi^*[\cdot]$ may be decoupled into q nonintersecting subsequences (9.17) each of which, with probability 1, satisfies for any $\delta > 0$ the inclusion (9.18) for some $i = qj(k) + k$ (due to (9.20)).

Therefore, with $\delta > 0$ given, we may select

$$\psi^*(i) = \psi^0(k)$$
$$\text{for } i = qj(k) + k \ , \ k = 1 , \ldots , q \ ,$$

$$\psi^*(i) = 0 \ , \ i \neq qj(k) + k \ , \qquad (9.21)$$
$$N = qj(q) + q$$

Substituting $\psi^*(i)$, $\xi^*(i)$ into (9.3) and using the periodicity of $p(i)$ $(p(qj + k) = p(k) \ , \ j = 1 \ , \ldots, \infty \ ; \ k = 1 \ , \ldots, \ q)$

we have

$$\rho(\wedge \mid R[N]) = \sum_{i=1}^{N} \rho(-\psi^*(i) \mid \xi^*(i) - Q) =$$

$$= \sum_{k=1}^{q} \rho(\psi^*(qj(k) + k) \mid \xi^*(qj(k) + k) - Q) \qquad (9.22)$$

with

$$\sum_{i=1}^{N} \psi^*(i) \ p'(i) = \sum_{k=1}^{q} \psi^*(qj(k) + k) \ p'(qj(k) + k) = \wedge$$

$$\xi^*(qj(k) + k) \in Q_\delta(k)$$

In view of (9.16), (9.21), (9.22) and the definition of $Q_\delta(k)$ one may observe

$$\rho(\wedge \mid R[N]) =$$
$$= \sum_{k=1}^{q} (\rho(\psi^0(k) \mid \xi^0(k) - Q) +$$
$$+ \rho(\psi^0(k) \mid \xi^*(qj(k) + k) - \xi^0(k))) \leq$$
$$\leq \delta \sum_{k=1}^{q} || \ \psi^0(k) \ ||$$

Therefore, with $\wedge$, σ given, one may select $\psi^0 \ [1 \ , \ q]$, δ, so that

$$\delta \sum_{k=1}^{q} || \ \psi^0 \ (k) \ || \leq \sigma$$

Summarizing the discussion of the above we observe that for every $\wedge \in \mathbf{R}^{m \times n}$, $\sigma > 0$, there exists a number $N(\wedge \ , \sigma)$ that ensures $\rho(\wedge \mid R[s]) \leq \sigma \ , \ s \geq N, \ N = N(\wedge \ , \sigma)$.

If $\bar{\wedge}_0^{(i)} = e^{(i)}$ is an orthonomal basis in $\mathbf{R}^{mn}$ $(e_j^{(i)} = \delta_{ij} \ ; \ j = 1 \ , \ldots, mn)$ and

$$N_0(\sigma) = \max\{N(\bar{\wedge}_0^{(i)} \ , \sigma) \ , \ N(-\bar{\wedge}_0^{(i)} \ , \sigma) \ \} \ , \ (i = 1 \ , \ldots, \ mn) \ ,$$

then

$$\rho(\pm e^{(i)} \mid R[s]) \leq \sigma \ , \ (\forall \ i = 1 \ ,..., mn) \ , \ s \geq N_0 \ (\sigma)$$

and

$$h\{R[s] \ , \{0\}\} \leq \sqrt{mn} \ \sigma$$

Taking $\epsilon = \sqrt{mn} \ \sigma$, $N = N_0 \sigma$ we arrive at the relation (9.15). Lemma 9.4 is now proved.

The examples given in Cases A and C indicate two important classes of disturbances $\xi(k)$ of which one consists of *periodic functions* and the other of *a sequence of equidistributed independent random variables*. In both cases one may ensure consistency of the identification process. However this requires some additional assumptions on the inputs $p(k)$. Basically this means that function $p(k)$ should be periodic and its informational matrix should be nondegerate as indicated in the precise formulations, (see also [23, 24]).

10. Identification of the Coefficients of a Linear Autonomous Dynamic System

Consider a dynamic process governed by a linear system

$$x(k+1) = Ax(k) + B \ u(k) + \xi(k) \tag{10.1}$$
$$k \in [0 \ , s]$$

The *input* $u(k)$ and the *output* $y = x(k)$ are taken here to be *given*, the constant *coefficients A, B are to be identified* and the *input noise* $\xi(k)$ is taken to be *unknown but bounded* by a geometrical constraint

$$\xi(k) \in Q(k) \ , \ k \in [0 \ , s] \tag{10.2}$$

Here as usual $x \in \mathbf{R}^n$, $u \in \mathbf{R}^p$, $v \in \mathbf{R}^q$, $A \in \mathbf{R}^{n \times n}$, $B \in \mathbf{R}^{n \times p}$ and there is some additional information on A, B. Namely it is assumed that

$$A \in \mathbf{A} \ , B \in \mathbf{B} \ , \tag{10.3}$$

where A , B are convex and compact sets in the matrix space of respective dimensions.

We will derive a recurrence equation for the related informational domains. These are given by the following definition.

Definition 10.1 The informational domain $A[s] \times B[s] = H[s]$ consistent with system (10.1), restrictions (10.2), (10.3) and measurement $x(k)$, $k \in [0 , s]$ is the set of all matrix pairs $\{A , B\}$ for each of which there exists a sequence $\xi[0 , s] \in Q[0 , s]$ such that relations (10.1)-(10.3) would be fulfilled.

Since the input $u[0 , s]$ is taken to be given, the domain $H[s]$ will obviously depend upon $u[0 , s]$:

$$H[s] = H[s , u[0 , s]) = H(s , \cdot)$$

In order to solve the estimation problem we introduce a matrix C and a vector $p(k)$.

$$C = [A, B], p(k) = \begin{bmatrix} x(k) \\ u(k) \end{bmatrix}$$

Then taking

$$y(k) = x(k + 1) ,$$

we come to the standard measurement equation of § 3:

$$y(k) = Cp(k) + \xi(k)$$

Applying the recurrence equation of (8.2) we come to the relations that describe the dynamics of set $H(s , u[0 , s]) = H[s]$.

The consistency theorems of § 9 may be applied if there is some additional information on A , B and on the known inputs $u[0 , s]$ that would ensure that the conditions of these theorems would be fulfilled.

Another formal scheme for obtaining a recurrence equation for $H[s]$ may be presented as follows. Introducing a vector

$$z = \begin{bmatrix} \bar{A} \\ \bar{B} \end{bmatrix}$$

and an $n \times n(n + m)$- matrix

$$\mathbf{G}(k) = (x'(k) \otimes I_n , u'(k) \otimes I_n)$$

we arrive at the system

$$z(k+1) = z(k) \, , \tag{10.4}$$

$$y(k) = \mathbf{G}(k)\; z(k) + \xi(k) \; , \; 0 \le k \le s \; , \tag{10.5}$$

where the aim is to identify the *informational domain* $Z(s) = H[s]$ of the states of system (10.4) consistent with measurement $y[0\;,\;s]$ and constraints (10.2), (10.3).

Following formally the results of § 13 (formula (13.6) for the one-stage process) and rewriting them in terms of the notations of this paragraph we come to the recurrence relation

$$Z(k+1) \subseteq \bigcap_{M} \{(I - M'\,\mathbf{G}(k))\; Z(k) + \tag{10.6}$$

$$+ \; M(y(k) - Q(k))\} \; , \; Z(0) = \begin{bmatrix} \bar{A} \\ \bar{B} \end{bmatrix}$$

$$Z \in \mathbf{R}^{n(n+m)} \; , \; M \in \mathbf{R}^{n(n+m) \times n}$$

which at each stage is true for any matrix $M \in \mathbb{M}^{n(n+m) \times n}$. According to the conventional scheme we arrive at

Lemma 10.1 The set-valued estimate for the vector C of coefficients for system (10.1) *is given by the solution* $Z(s) = H(s)$ *for equation* (10.6).

It is now natural to consider in greater detail the issue of *state estimation* for linear systems with unknown but bounded measurement noise and input disturbances. We will start with the first case.

11. The Observation Problem

Consider a recurrence equation

$$x(k+1) = A(k)\; x(k) \; , \; x(k_0) = x^0 \; , \tag{11.1}$$

$$x \in \mathbf{R}^n \; , \; A(k) \in \mathbf{R}^{n \times n} \; , \; k \ge k_0 \; ,$$

together with a measurement equation

$$y(k) = g'(k)\; x(k) + \xi(k) \; , \; k \ge k_0 + 1$$

with vector $g(k) \in \mathbf{R}^n$ and "noise" $\xi(k)$ restricted by a geometrical constraint.

$$\xi(k) \in Q(k) \; , \; Q(k) \in comp\; \mathbf{R}^m$$

The objective is to estimate the initial vector x^0 by processing a given measurement $y[1\ ,\ s]$, taking $A(k)\ ,\ g(k)\ ,\ Q(k)$ to be given in advance. We will further call this ***the observation problem*** (in the presence of unknown but bounded "noise" with set-membership bounds on the unknowns).

Observing that $x(s) = S(s)\ x^0$, where $S(s)$ is the solution to the matrix equation

$$S(k+1) = A(k)\ S(k)\ ,\ S(k_o) = I_n$$

we may denote

$$p'(k) = g'(k)\ S(k) \tag{11.2}$$

transforming our problem to the conventional form of § 3 with

$$y(k) = p'(k)\ x^0 + \xi(k)$$

and with x^0 replacing the unknown C.

The condition for the identifiability of x^0 in the absence of "noise" now turns to be again $|\ W(s)\ | \neq 0$ with

$$W(s) = \sum_{k=k_o}^{s} S'(k)\ g(k)\ g'(k)\ S(k) \tag{11.3}$$

The latter relation is known as the *observability condition* [3, 4] for system (11.1) with measurement

$$y(k) = g'(k)\ x(k) \tag{11.4}$$

Condition $|W(s)| \neq 0$ is obviously ensured if vectors $p(k) = S'(k)\ g(k)$, $(k = 1\ ,\ldots,\ k)$ are linearly independent.

The general solution will now consist in constructing the informational domains $\mathbf{X}^o[s]$ for the vector x^o. They are the direct substitutes for $\mathbf{C}[s]$.

Following (8.2), (11.2) we will have a system of recurrence relations

$$\begin{gathered} X_M^o\,(k+1) \subseteq (I_n - M(k+1)\ g'(k+1)\ S(k+1))X_M^o(k) + \\ +M(k+1)(y(k+1) - Q(k+1))\ ,\ X(k_o) = X^o \\ S(k+1) = A(k)\ S(k)\ ,\ S(k_o) = I_n \end{gathered} \tag{11.5}$$

which are true for any sequence $M[k_o + 1\ ,\ s]$.

The results of the previous paragraph then leads us to

Lemma 11.1 The solution x^0 to the observation problem may be estimated from above by

$$\mathbf{X}^o[s] = \{ \bigcap X_M^o(s) \mid M[k_o + 1\ ,\ s]\} \tag{11.6}$$

Namely

$$x^o \in \mathbf{X}^o[s]\ ,\ \forall\ s > k_o + 1$$

The solution will be consistent with

$$h\ \{ X_M^o\ (k)\ ,\ x^o\} \longrightarrow 0\ ,\ k \longrightarrow \infty \tag{11.7}$$

if for example the problem falls under one of the conditions 9A - 9C of the previous paragraph.

Particularly, for an autonomous system (11.1), this will be ensured if

(a) the function $p(k) = g'S(k)$ is n-periodic,

(b) the vectors

$$g'\ ,\ g'A\ ,\ldots,\ g'A^{n\ -\ 1}$$

are linearly independent (the system (11.1), (11.4) is completely observable).

(c) the noise is uniformly distributed in the interval $Q(k) \equiv Q = -\ Q$.

Lemma 11.2 Under conditions (a) - (c) *the solution* $\mathbf{X}^o[s]$, (11.5), (11.6) *to the observation problem is consistent in the sense of* (11.7).

A simple example, when the conditions of Lemma 11.1 are satisfied, is given by a system (11.1) in $\mathbf{R}^3$

$$g' = (1\ ,0\ ,0)\ ,\ A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix},\ \mid \xi(k) \mid\ \leq 1$$

Here

(a) $p(k) = p(3i + j)$ is periodic with period 3, $j = 1\ ,\ 2\ ,\ 3$; $i = 0\ ,\ldots,\ \infty$; $3i + j = k$

(b) $p(j) = e^{(j)} = \delta_{kj}$, $k = 1$, 2 , 3 so that $p(1)$, $p(2)$, $p(3)$ are linearly independent,

(c) $\xi\,(k)$ is taken to be equidistributed in the interval [-1,1].

The solution to this problem may be given by a polyhedral approximation so that, assuming $X^o[k]$ given, we will seek for an approximation of $X^o[k]$ by a polytope $X^o[k+1]$ through the formula

$$\rho(\ell \mid \mathbf{X}^o[k+1]) = inf\,\{H(\ell\,,\,m\,,\,\mathbf{X}^o[k]) \mid m\}$$
$$H(\ell\,,\,m\,,\,\mathbf{X}^o[k]) = \{\rho(\ell \mid (I_n - m'\,p(k+1))\,\mathbf{X}^o[k]) + (\ell\,,\,m)\;y(k+1) +$$
$$+\,\rho(-\ell \mid m'\;Q(k+1))\}\;,\,\ell \in \mathbf{R}^3,\,m \in \mathbf{R}^3,$$

taking for each step a set of orthonormal vectors $\{e^{(i)}\}$ with a set of vectors $\{-e^{(i)}\}$, and assuming $\ell = e^{(i)}$, $\ell = -\;e^{(i)}$, $(i = 1\;,\ldots,\;3)$

Therefore, in order to define $\mathbf{X}^o[k+1]$ with $\mathbf{X}^o[k]$ given, we will have to solve 6 independent unconstrained minimization problems, in 3 variables each, so that the vertices of $\mathbf{X}^o[k+1]$ would be given by 3 coordinates each, selected from the variety of numbers

$$\rho(+e^{(i)} \mid \mathbf{X}[k+1])\;,\,-\rho\,(-e^{(i)} \mid \mathbf{X}[k+1])\;,\,(i = 1\;,\,2\;,\,3).$$

A simpler algorithm involves only one optimization problem (in three variables, the coordinates of m) so that one should minimize in m the function

$$V_A\,(m\;,\,k+1) = \prod_{i=1}^{3}\left[H(e^{(i)}\;,\,m\;,\,\mathbf{X}^o[k]) + H(-e^{(i)}\;,\,m\;,\,\mathbf{X}^o\;[k])\right]$$

which for a given m, is equal to the volume of a polyhedron $\mathbf{X}(m\;,\,k+1) \supseteq \mathbf{X}[k+1]$ The last inclusion is true for any $m \in \mathbf{R}^3$ and one should therefore seek for the optimal m. The projections of $X[k]$ on the axes $\{x_1,\,x_2\}$, $\{x_1,\,x_3\}$ are shown in Figure 5.

A separate issue is the construction of an ellipsoidal approximation for $\mathbf{X}[k+1]$.

A more complicated problem is to estimate the state of a linear system with unknown input on the basis of measurement corrupted by noise. We will therefore deal with the problem of guaranteed state estimation for a linear system subjected to unknown but bounded disturbances with nonquadratic restrictions on the unknowns.*

* The treatment of quadratic constraints is known well enough and may be found in references [15, 16]

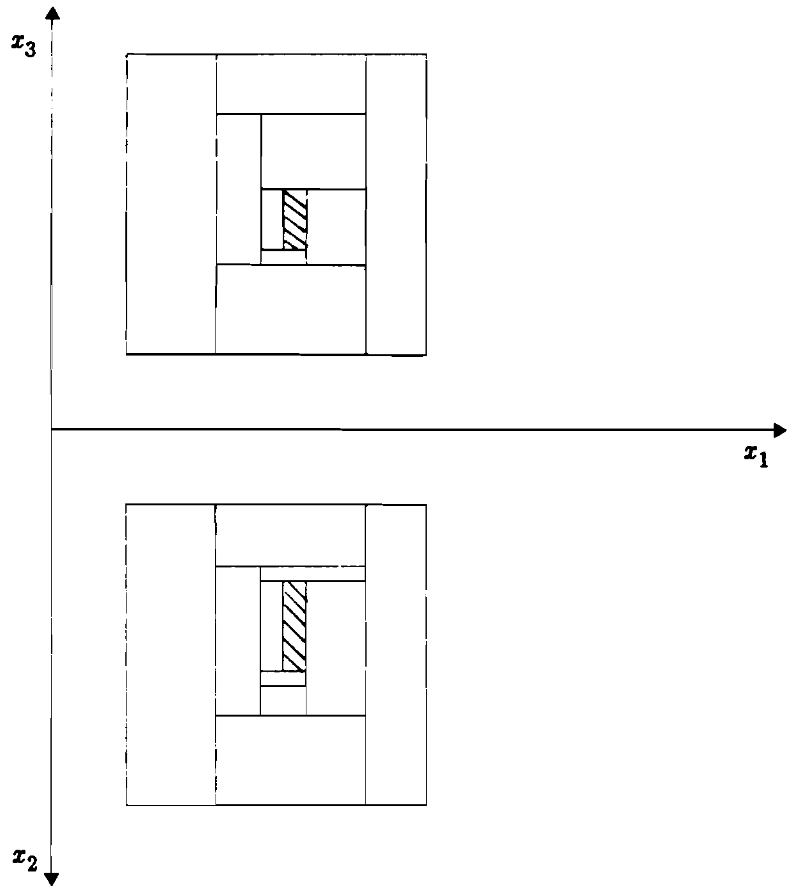

FIGURE 5

12. Uncertain Dynamic Systems

An uncertain dynamic system is understood here to be a discrete-time multistage process, described by an n-dimensional equation

$$x(k+1) = A(k)\, x\,(k) + B(k)\, v\,(k) \tag{12.1}$$

where $A(k)$, $B(k)$, $k = 0, \ldots, s$ are given matrices. The input $v(k)$, and the initial stage x^0 are vectors of finite-dimensional spaces $\mathbf{R}^p$ and $\mathbf{R}^n$ respectively. They are assumed to be unknown being restricted in advance by instantaneous "geometric" constraints

$$x(0) = x^0 \in X^0 \,,\ v(k) \in P(k) \,,\ k = 0, \ldots, s \ , \tag{12.2}$$

where X^0 , $P(k)$ are given convex and compact sets. It is further assumed that direct measurements of the state $x(k)$ are impossible, the available information on the process dynamics being generated by the equation

$$y(k) = G(k)\, x\,(k) + \xi(k)\ ;\ k = 1\ ,\ldots,s \tag{12.3}$$

with measurement vector $y(k) \in \mathbf{R}^m$ and matrix $G(k)$ given. The disturbances $\xi(k)$ are unknown and restricted as before by an inclusion

$$\xi(k) \in Q(k) \tag{12.4}$$

with convex compact set $Q(k) \in \mathbf{R}^m$ given in advance.

We will use the symbol $x(k\ ,\ v\,[0\ ,\ k\ -\ 1]\ ,\ x^0)$ to denote the end of the trajectory $x(j)$ for system (12.1) formed for $[0\ ,\ k]$ with $v[0\ ,\ k\ -\ 1]\ ,\ x^0$ given.*

Let us assume that after s stages of system operation there appeared a measurement sequence $y[1\ ,\ s]$, generated due to relations (12.1)-(12.4).

The knowledge of $y[1\ ,\ s]$ will allow us to consider the following construction.

Definition 12.1 An informational domain $X[s] = X(s\ ,\ 0\ ,\ X^0)$ *will be defined as the set that consists of the ends* $x(s\ ,\ v[0\ ,\ s\ -\ 1]\ ,\ x^0)$ *of all those trajectories* $x(j)$ *formed for the interval* $j \in [0\ ,\ s]$ *that could generate the measured sequence* $y[1\ ,\ s]$ *under constraints* (12.2)-(12.4).

More generally, with $y[k+1\ ,\ \ell]\ ,\ (k+1 \le \ell)$ *and* $F \in co\ \mathbf{R}^n$ *given,* $X(\ell\ ,\ k\ ,\ F)$ *will be the set of the ends* $x(\ell\ ,\ v[k\ ,\ \ell - 1]\ ,\ x^*)$ *of the trajectories* $x(j)$ *of system* (12.1) *that start at stage* k *from state* $x(k) = x^*$ *and are consistent with realization* $y[k+1\ ,\ \ell]$ *due to equation* (12.3) *with constraints*

$$x^* \in F\ ,\ v(i) \in P(i)\ ,\ k \le i \le \ell - 1\ ,$$
$$\xi(j) \in Q(j)\ ,\ k+1 \le j \le \ell\ ,$$

The dynamics of the total system (12.1)-(12.3) will now be determined by the evolution of sets $X[s]$. It is clear that set $X[s]$ includes the unknown actual state of system

* In order to simplify some further notations of this paragraph we will generally start the process at stage $k_0 = 0$ instead of arbitrary $k_0 = k^*$, although the basic system is *nonstationary*.

(12.1).

In particular $X[s] = X(s, 0, X^0)$.

From the definitions of the above it is possible to verify the following assertions.

Lemma 12.1 Assume F, $P(k)$, $Q(k)$ to be convex compact sets in spaces $\mathbf{R}^n$, $\mathbf{R}^p$, $\mathbf{R}^m$ respectively. Then each of the sets $X(s, \ell, F)$ will be convex and compact.

Lemma 12.2 Whatever is the set $F \subseteq \mathbf{R}^n$, the following equality is true $(s \geq \ell \geq k)$

$$X(s, k, F) = X(s, \ell, X(\ell, k, F)) \tag{12.5}$$

Condition (12.5) indicates that the transformation $X(s, k, F)$ possesses a *semi-group property generating a generalized dynamic system* in the space of convex compact subsets of $\mathbf{R}^n$. The generalized system will then absorb all the informational and dynamic features of the total process. Here each $X[s]$ contain all the prehistory of the process and the process evolution for $r > s$ depends only upon $X[s]$ but not upon the previous $X[i]$, $i < s$.

The general description of $X[s]$ requires a rather cumbersome procedure which does not follow directly from § § 7,8. Our objective is to obtain a description of sets $X[s]$ which are the *set-valued state estimators* for the system (12.1)-(12.4). The situation therefore justifies the consideration of approximation techniques based on solving some auxiliary deterministic or even stochastic estimation problems. In order to explain the procedures, we will start with an elementary one-stage solution.

13. Guaranteed State Estimation. The One-Stage Problem

Consider the system

$$z = Ax + Bv, \ y = Gz + \xi \tag{13.1}$$

where

$$x, z \in \mathbf{R}^n, v \in \mathbf{R}^p, \xi \in \mathbf{R}^m,$$

and the matrices A , B , G are given. Knowing the constraints

$$x \in X \ , v \in P \ , \xi \in Q \ , \tag{13.2}$$

where

$$X \in comp\ \mathbf{R}^n \ , P \in comp\mathbf{R}^p \ , Q \in comp\ \mathbf{R}^q$$

and knowing the value y, one has to determine the set Z of vectors z consistent with equations (13.1) and inclusions (13.2).

Denote

$$Z_s = AX + BP$$
$$Z_y = \{z : y - Gz \in Q\}$$

Then obviously

$$Z = Z_s \bigcap Z_y \tag{13.3}$$

Standard considerations yield a relation for the support function

$$\rho(\ell \mid Z) = \max \{(\ell , z) \mid z \in Z\}$$

Applying the convolution formula of convex analysis [21]

$$\rho(\ell \mid Z) = inf\ \{\rho(\ell^* \mid Z_s) + \rho\ (\ell^{**} \mid Z_y) \mid \ell^* + \ell^{**} = \ell\}$$

Lemma 13.1 The support function $\rho(\ell \mid Z) = \psi(\ell)$ *where*

$$\psi(\ell) = inf\ \{\Phi(\ell , p) \mid p \in \mathbf{R}^m\} \tag{13.4}$$
$$\Phi(\ell , p) = \rho(A'\ \ell - A'\ G'\ p \mid X) + \rho(B'\ \ell - B'\ G'\ p \mid P)$$
$$+ \rho(-p \mid Q) + (p , y) ,$$

The set Z may be given in another form. Indeed whatever the vectors $\ell , p , \ell \neq 0$ are, it is possible to represent (13.4) $p = M\ell = p[\ell , M]$ where matrix $M \in \mathbf{R}^{m \times n}$. Relation (13.4) will then turn into

$$\psi(\ell) = inf\ \{\Phi(\ell , p[\ell , M]) \mid M \in \mathbf{R}^{m \times n}\} \tag{13.5}$$

Problem (13.5) will be referred to as the *dual problem* for (13.3). The latter relation yields the inclusion

$$Z \subseteq (I_n - M'\ G)\ (AX + BP) + M'(y - Q) = R(M) \tag{13.6}$$

which will be true for any matrix M.

Equality (13.5) thus leads to *set-valued duality relations* in the form of (13.6) and further on in the form of

Lemma 13.2 The following equality is true

$$Z = \{\bigcap R(M) \mid M\} \tag{13.7}$$

over all matrices $M \in \mathbf{R}^{m \times n}$. *Here set* Z *is a "guaranteed" estimate for* z *which may be calculated due to* (13.5).

The necessity of solving (13.5) gives rise to the question of whether it is possible to calculate $\rho(\ell \mid Z)$ in some other way, for example, by the variation of the relations for some kind of *stochastic estimation problem.* A second question is whether there exist any general relations between the solutions to the guaranteed and to the stochastic filtering problems.

In fact it is possible to obtain an inclusion that would combine the properties of both (13.6) and of conventional relations for the linear-quadratic Gaussian estimation problem.

14. Relation Between Guaranteed and Stochastic Estimation. The One-Stage Problem

Having fixed a certain triplet $h = \{x \, , \, v \, , \, \xi\}$ that satisfies (13.2) (the set of all such triplets will be further denoted as H), consider the system

$$w = A(x + q) + Bv \; , \; y = Gw + \xi + \eta \; , \tag{14.1}$$

where q , η are independent Gaussian stochastic vectors with zero means

$$Eq = 0 \quad , \quad E\eta = 0 \, ,$$

and with covariance matrices

$$Eqq' = L \qquad E\eta\eta' = N$$

where L , N are positive definite. Assume that after one random event the vector y has appeared due to system (14.1). The conditional expectation $E(w \mid y)$ may then be determined for example by means of a Bayesian procedure or by a least-square method. We have

$$E(w \mid y) = Ax + A\mathbf{P}A'\, G'\, N^{-1}(y - GAx - GBv - \xi) + Bv\ , \qquad (14.2)$$
$$\mathbf{P}^{-1} = L^{-1} + A'\, G'\, N^{-1}\, GA$$

or in accordance with a conventional matrix transformation [25].

$$\mathbf{P} = L - LA'\, G'K^{-1}\, GAL\ , \qquad (14.3)$$
$$K = N + GALA'\, G'\ ,$$

an equivalent condition

$$\bar{w}_y = E(w \mid y) = Ax + ALA'\, G'\, K^{-1}\,(y - GAx - (GBv + \xi)) + Bv \qquad (14.4)$$

We observe that the conditional variance

$$E((w - \bar{w}_y)(w - \bar{w}_y)' \mid y) = A\mathbf{P}A' \qquad (14.5)$$

does not depend upon h and is determined only by pair

$$\wedge = \{L\ ,\ N\}$$

where $L > 0$, $N > 0$. (In the latter case further we will write $\wedge > 0$.)

Therefore we may consider the set of all conditional mean values

$$W(\wedge) = \{\bigcup \bar{w}_y \mid h \in H\}$$

that correspond to all possible $h \in H$. Here

$$W(\wedge) = (I_n - ALA'\, G'\, K^{-1}\, G)\,(AX + BP) + ALA'\, G'\, K^{-1}\,(y - Q) \qquad (14.6)$$

Having denoted

$$\Psi(\wedge) = K^{-1}\, GALA'$$

we come to

Lemma 14.1 The set $W(\wedge)$ is convex and compact: $W(\wedge) \in$ comp $\mathbf{R}^n$. The following equality is true

$$\rho(\ell \mid W(\wedge)) = \Phi(\ell\ ,\ p(\ell\ ,\ \wedge)) \qquad (14.7)$$

where

$$p(\ell\ ,\ \wedge) = \Psi(\wedge)\ \ell$$

We may now observe that function $\Phi(\ell , p(\ell , \wedge))$ differs from $\Phi(\ell , p[\ell , M])$ used in (13.5) by a mere substitution of $p(\ell , \wedge)$ by $p[\ell , M]$. Comparing (14.7) and (13.5), we conclude

Lemma 14.2 Whatever is the pair $\wedge > 0$, the inclusion

$$Z \subseteq W(\wedge) \tag{14.8}$$

is true.

We will see that by varying $\wedge$ in (14.8) it is possible to achieve an exact description of set Z.

In order to prove this conjecture some standard assumptions are required.

Assumption 14.1 The matrix GA is of rank m.

We shall also make use of the following relation:

Lemma 14.3 Under assumption 14.1 take $\wedge = \wedge(1 , \alpha) = \{I_n , \alpha I_m\}$. Then $\Psi(\wedge(1 , \alpha))\ G' \longrightarrow I_m$ with $\alpha \longrightarrow 0$.

The given relation follows from equality $\Psi(\wedge(1 , \alpha))\ G' = (\alpha I_m + D)^{-1} D$ where matrix $D = GALAG'$ is nonsingular, $L = I_n$.

Theorem 14.1 The inclusion $z \in Z$ is true if and only if for any $\ell \in \mathbf{R}^n$, $\wedge > 0$ we have

$$(\ell , z) \leq \rho(\ell \mid W(\wedge)) = f(\ell , \wedge) \tag{14.9}$$

Inequality (14.9) follows immediately from the inclusion $z \in Z$ due to Lemma 14.2. Therefore it suffices to show that (14.9) yields $z \in Z$. Suppose that for a certain z^* the relation (14.9) is fulfilled, however $z^* \bar{\in} Z = Z_s \bigcap Z_y$. First assume that $z^* \bar{\in} Z_y$. Then there exists an $\epsilon > 0$ and a vector p^* such that

$$(-p^* , y) + (G' p^* , z^*) > \rho(-p^* \mid Q) + \epsilon \tag{14.10}$$

Now we will show that it is possible to select a pair of values ℓ^* , $\wedge^*$ that depend upon p^* and are such that

$$(\ell^* , z^*) > \rho(\ell^* \mid W(\wedge^*)) = f(\ell^* , \wedge^*) \tag{14.11}$$

Indeed, taking $\ell^* = G' p^*$, $\wedge(1 , \alpha) = \{I_n , \alpha I_m\}$ we have

$$f(\ell^* , \wedge(1 , \alpha)) = \Phi(\ell , \wedge(1 , \alpha)) \pm ((p^* , y) + \rho(-p^* \mid Q)) \qquad (14.12)$$

From Lemma 14.3 and condition

$$p(\ell^* , \wedge(1 , \alpha)) = K^{-1}(\alpha) \; GAI_n \, A' \; G' \; p^* \; , \; K(\alpha) = \alpha I_m + GAA'G'$$

it follows that

$$p(\ell^* , \wedge(1 , \alpha)) \longrightarrow p^* \; , \; \alpha \longrightarrow 0 \qquad (14.13)$$

But then from condition (14.13), from Lemma 14.2 and from the properties of function $f(\ell , \wedge)$ it also follows that for any $\epsilon > 0$ there exists an $\alpha_0 \, (\epsilon)$ such that for $\alpha \leq \alpha_0(\epsilon)$ the inequality

$$\mid f(\ell^* , \wedge(1 , \alpha)) - ((p^* , y) + \rho(-p^* \mid Q)) \mid \leq \epsilon/2 \qquad (14.14)$$

is true.

Comparing (14.10), (14.12), (14.14) we observe that for $\alpha \leq \alpha_0(\epsilon)$.

$$(\ell^* , z^*) = (G' \; p^* , z^*) \geq f(\ell^* , \wedge(1 , \alpha)) + \epsilon/2 \; .$$

Therefore, with $\wedge^* = \wedge(1 , \alpha^*)$, $\alpha < \alpha_0(\epsilon)$ the pair $\{\ell^* , \wedge^*\}$ yields the inequality (14.11).

Now assume $z^* \bar{\in} Z_s$. Then there exists a vector ℓ^0 for which

$$(\ell^0 , z^*) \geq \varsigma(\ell^0) + \sigma \; , \; \sigma > 0 \; .$$

where

$$\varsigma(\ell) = \rho(A' \; \ell \mid X) + \rho(B' \; \ell \mid P)$$

Taking $\ell = \ell^0$, $\wedge = \wedge(1 , \alpha)$ we find:

$$\Psi(\wedge(1 , \alpha)) \longrightarrow 0 \; , \; \alpha \longrightarrow \infty \; .$$

But then for any $\sigma \longrightarrow 0$ there exists a number $\alpha^0(\sigma)$ such that

$$\mid f(\ell^0 , \wedge(1 , \alpha)) - \varsigma(\ell^0) \mid \leq \sigma/2$$

provided $\alpha > \alpha^0(\sigma)$. Hence, for $\alpha > \alpha^0(\sigma)$ we have

$$(\ell^0 , z^*) \geq f(\ell^0 , \alpha) + \sigma/2$$

contrary to (14.9). The theorem is thus proved.

From the given proof it follows that Theorem 14.1 remains true if we restrict ourselves to the one parametrical class

$$\wedge^{(1)} = \{\wedge(1\ , \alpha)\}\ , \wedge(1\ , \alpha) = \{I_n\ , \alpha I_m\}$$

Therefore, the theorem yields:

Corollary 14.1 Under the conditions of Theorem 14.1 the inclusion $z \in Z$ is true if and only if for any $\ell \in \mathbf{R}^n$ we have

$$(\ell\ , z) \leq f_1(\ell)\ , \tag{14.15}$$

where

$$f_1(\ell) = inf\ \{f(\ell\ , \wedge(1\ , \alpha))\ |\ \alpha > 0\}$$

Being positively homogeneous, the function $f_1(\ell)$ may, however, turn out to be non-convex, its lower convex bound being the second conjugate $f_1^{**}(\ell)$ where, [21],

$$g^*(q) = sup\ \{(\ell\ , q) - g(\ell)\}\ , g^{**}\ (\ell) = (g^*)^*(\ell)$$

The convexification of $f_1(\ell)$ in (14.15) will not violate this inequality. In other words, (14.15) will yield

Corollary 14.2 Under the conditions of Theorem 14.1, we have

$$\rho(\ell\ |\ Z) = f_1^{**}(\ell) \leq f_1(\ell) \tag{14.16}$$

However, if we move on to a broader class $\wedge^{(2)} = \{L\ , N\}$ where $L > 0$ and $N > 0$ depend together on at least m independent parameters it is possible to achieve a direct equality immediately, i.e.

$$\rho(\ell\ |\ Z) = f_2(\ell) \tag{14.17}$$

where

$$f_2(\ell) = inf\ \{f(\ell\ , \wedge)\ |\ \wedge \subseteq \wedge^{(2)}\} = f_1^{**}(\ell)\ , \tag{14.18}$$

Problem (14.18) will be called *the stochastically dual* for (13.5). The following assertion is true.

Theorem 14.2 Under assumption 14.1 relations (14.17), (14.18) *are true, where the infimum is taken over all* $L > 0$, $N > 0$.

The proof of Theorem 14.2 is rather long and will be omitted in this text. It may be found in paper [26].

The stochastic dual problem (14.18) may therefore replace (13.6).

On the other hand we may again turn to set-valued duality, now in terms of a stochastic problem. Due to Corollary 14.1 the set of inequalities (14.15) will lead us to

Lemma 14.3 The following equality is true

$$Z = \{\bigcap W(\wedge) \mid \wedge \in \wedge^{(1)}\} \tag{14.19}$$

The relations of this paragraph indicate that set Z may be described by deterministic relations (13.7) as well as by approximations (14.19) generated due to the stochastic estimation problems of the above.

The results of this paragraph allow to devise solutions to multistage problems.

15. A Multi-Stage System

Returning to system (12.1)-(12.4) let us seek for $X[s] = X(s , k_0 , X^0)$. We further introduce notations

$$Y(k) = \{x : y(k) - G(k)\, x \in Q(k)\}$$

and $X^*(s , j , F)$ is the solution $X(s)$ to the equation

$$X(k + 1) = A(k)\, X(k) + B(k)\, P(k) \,,\; j \leq k < s - 1 \tag{15.1}$$

with $X(j) = F$. Then it is possible to verify the following recurrent equation similar to (13.3).

Lemma 15.1 Assume $y[k_0 + 1 , k]$ *to be the realization for the measurement vector* y *of system (12.3), (12.1). Then the following condition is true.*

$$X[k] = X(k , k_0 , X^0) = X^* (k , k - 1 \mid X[k - 1]) \bigcap Y(k) \tag{15.2}$$

Formula (15.2) indicates that the innovation introduced by the k-th measurement $Y(k)$ appears in the form of an intersection. Therefore $X^* (k , k - 1 | X[k - 1])$ is the estimate for the state of the system on stage k *before* the arrival of the k-th measurement while $X[k]$ is the estimate obtained *after* its arrival.

Relations (15.2) may be interpreted as a *recurrence equation*. One may rewrite them in a somewhat different way, namely through (13.6) and (13.7). Applying (13.7) for each stage k we come to

Lemma 15.2 The set $X[k]$ satisfies the following recurrence equation

$$X[k + 1] = \{\bigcap (I_n - M'G(k))(A(k)X[k] + B(k)\mathbf{P}(k)) + \\ + M'(y(k) - Q(k))|M\}$$
$$X[k_0] = X^0$$

A nonlinear version of this scheme is given further in § § 18–20. However, the topic of this paragraph is another procedure. It is the scheme of *stochastic filtering approximation* which follows from the results of § 14, (Theorem 14.1). Together with (12.1, (12.3) consider the system (involving almost sure equalities)

$$w(k + 1) = A(k)\, w(k) + B(k)\, v(k) + C(k)\, u(k) \tag{15.3}$$
$$k = k_0 , 1 ,\dots, s - 1 ;\; w(k_0) = x^0 + w^0 ,$$
$$z(k) = G(k)\, w(k) + \xi(k) + \eta(k) ,\; u(k) \in \mathbf{R}^q , \tag{15.4}$$

where the inputs x^0 , $v(k)$, $\xi(k)$ are deterministic, subjected to "instantaneous" constraints

$$x^0 \in X^0 ,\; v(k) \in \mathbf{P}(k) ,\; \xi(k) \in Q(k) ,$$

while w^0 , $u(k)$, $\eta(k)$ are independent stochastic Gaussian vectors with

$$\begin{gathered} \bar{w}^0 = Ew^0 = 0 ,\; \bar{u}(k) = Eu(k) = 0 , \\ \bar{\eta}(k) = E\eta(k) = 0 ,\; Ew^0\, w^{0\prime} = P^0 , \\ Eu(k)\, u'(k) = L(k) ,\; E\eta(k)\, E\eta(k) = N(k) , \end{gathered} \tag{15.5}$$

where L , N are positive definite.

Suppose that after $k - k_0$ stages for system (15.3), (15.4) measurement $z[k_0 , k] \in \mathbf{R}^{m(k - k_0)}$ has been realized. Having fixed the triplet

$$\xi[0\ ,\ k] = \{x^0\ ,\ v[k_0\ ,\ k-1]\ ,\ \xi[k_0\ ,\ k]\}$$

and having denoted $\omega(k) = \{\, v(k-1)\ ,\ \xi(k)\,\}$, $D(k) = \{\, P(k-1)\ ,\ Q(k)\,\}$ we may find a recursion for the conditional mean value

$$\bar{w}(k+1) = E\{w(k+1) \mid \omega(k)\ ,\ \bar{w}(k)\ ,\ z(k+1)\}$$

Define

$$\begin{aligned} W[k+1\ ,\ F] &= W(k+1\ ,\ L(k)\ ,\ N(k+1)\ ,\ F) \\ &= \bigcup \{\, \bar{w}[k+1] \mid \omega(k) \in D(k)\ ,\ \bar{w}(k) \in F\,\} \end{aligned}$$

From Theorems 14.1, 14.2 and Lemma 14.3 we come to the following propositions

Theorem 15.1 Suppose Assumption 14.1 holds for $A = A(k)$, $G = G(k+1)$, $k \in [k_0\ ,\ s]$ *and the sequence of observations* $y[k_o\ ,\ s]$, $z[k_o\ ,\ s]$ *for system* (12.1), (12.3) *and* (15.3), (15.4) *coincide:* $y[k_o\ ,\ s] = z[k_o\ ,\ s]$. *Then the following relation is true*

$$\begin{aligned} X[s] = \{ \bigcap W(s\ ,\ L\ ,\ N\ ,\ X[s-1]) \mid \wedge \in \wedge^{(1)} \}\ ,\ s > k_0\ , \\ X[k_0] = X^0\ ,\ \wedge = \{L\ ,\ N\}\ ,\ P^0 = 0\ , \end{aligned} \tag{15.6}$$

moreover, with $P^0 = 0$ *and*

$$f_i(\ell\ ,\ s) = inf\ \{\rho\ (\ell \mid W(s\ ,\ L\ ,\ N\ ,\ X[s-1])\}$$

over all $(L\ ,\ N) = \wedge \subset \wedge^{(i)}$, $i = 1\ ,\ 2$, *we have*

$$\rho(\ell \mid X[s]) = f_1^{**}\ (\ell\ ,\ s),\ \rho(\ell|X[s]) = f_2(\ell\ ,\ s)\ ,$$

where the second conjugate is taken in the variable ℓ.

Theorem 15.2 Under the condition of Theorem 15.1 for each positive definite matrix pair $\{\, L(k-1)\ ,\ N(k)\,\} = \wedge(k)$, *the following inclusions are valid*

$$\begin{aligned} X[k+1] &\subseteq W(k+1\ ,\ L(k)\ ,\ N(k+1)\ ,\ X[k]) \\ &= R(k+1\ ,\ \wedge(k+1)\ ,\ X[k])\ ,\ k \geq 0\ , \end{aligned} \tag{15.7}$$

where

$$\begin{gathered} R(k+1\ ,\ \wedge(k+1)\ ,\ X[k]) = (I_n - H(k+1)\ G(k+1))\ (A(k)\ X[k] + B(k)P(k) + \\ + H(k+1)\ (y(k+1) - Q(k+1))\ , \\ X[0] = X^0\ , \\ H(k+1) = C(k)\ L(k)\ C'(k)\ G'(k+1)\ K^{-1}\ (k+1)\ , \\ K(k+1) = N(k+1) + G(k+1)\ C(k)\ L(k)\ C'(k)\ G'(k+1)\ , \end{gathered}$$

The recurrence relations (15.7) thus allow a complete description of $X[s]$ through equation (15.6). Solving the system

$$W(k+1) = R(k+1 \, , \wedge(k+1) \, , \, W(k)) \, , \\ W(0) = X^0$$

we find

$$X[k+1] \subseteq W(k+1)$$

where

$$\rho(\ell \mid X[k+1]) = inf \, \{\rho(\ell \mid W(k+1)) \mid \wedge(j+1) \, ; \, j = k_0 \, , \ldots , \, k \, ; \, P^0 = 0\}$$

with each pair $\wedge(j+1) = \{ L(j) \, , \, N(j+1) \}$ belonging to the class $\wedge^{(2)}$. The total number of parameters over which the minimum is sought for does not exceed km.

The procedure given above is similar to the one given in (14.2). It is justified if the sets $X[k]$ are to be known for each $k > 0$. Note that in any way with *arbitrary* $L(j) \, , \, N(j+1) \, , \, j = 0 \, , \ldots , \, k-1$, the set $W(k)$ always *includes* $X[k]$.

Let us now assume that the desired estimate is to be found for only a fixed stage $s > k_0$. Taking $z[k_0 \, , s]$ to be known and triplet $\xi[k_0 \, , s]$ for system (15.3), (15.4) to be fixed, we may find the conditional mean values

$$\bar{w}(k) = E\{ w(k) \mid z[k_0 + 1 \, , k] \, , \xi[k_0 \, , k]\}$$

and the conditional covariance

$$\mathbf{P}(k) = E\{ w(k) - \bar{w}(k)) \, (w(k) - \bar{w}(k))' \mid z[k_o + 1 \, , k] \, , \xi[k_0 \, , k]\}$$

where

$$Ew(k_0) = x^0 \, , \, \mathbf{P}(k_0) = P^0$$

Denoting

$$\bar{w}[k \, , j \, , F] = E\{ w(k) \mid z[j+1 \, , k] \, , v[j \, , k-1] \, , \xi[j+1 \, , k] \, , \bar{w}(j)\} \\ \bar{W}[k \, , j \, , F] = \bigcup E\{ w(k) \mid z[j+1 \, , k] \, , v[j \, , k-1] \in P[j \, , k-1] \, , \\ \xi[j+1 \, , k] \in Q[j+1 \, , k] \, , \bar{w}(j) \in F\} \\ \bar{W}[k \, , k_0 \, , X^0] = \bar{W}(k) \, ,$$

and having in view the Markovian property for the process (15.3), (15.4) it is possible to conclude the following:

Lemma 15.3 The equality

$$\bar{W}(k) = \bar{W}[k\ ,\ j\ ,\ \bar{W}(j)] \tag{15.8}$$

holds for any $j\ ,\ k\ ,\ j \leq k$.

The corresponding formulae that generalize (14.2), (14.3) have the form

$$\begin{aligned} \bar{W}(k+1) &= (E - S(k+1)\ G(k+1))\ (A(k)\ \bar{W}(k) + B(k)P) \\ &\quad + S(k+1)\ (z(k+1) - Q)\ , \\ S(k+1) &= D(k+1)\ G'(k+1)\ K^{-1}\ (k+1)\ , \\ P(k+1) &= D(k) - D(k)\ G'(k+1)\ K^{-1}(k+1)\ G(k+1)D(k)\ , \\ D(k) &= A(k)\ P(k)\ A'(k) + C(k)\ L(k)\ C'(k) \\ K(k+1) &= N(k+1) + G(k+1)\ D(k)\ G'(k+1) \\ P(k_{0)} &= L\ , \end{aligned} \tag{15.9}$$

If we again suppose $z[k_0\ ,\ s] = y[k_0\ ,\ s]$, then due to the inclusions

$$\bar{W}(k+1) \supseteq \bar{W}[k+1\ ,\ k\ ,\ X[k]]\ ,\ k > k_0$$

that follow from Lemma 14.2 and to the monotonicity property

$$\bar{W}[k+1\ ,\ k\ ,\ F_1] \subseteq \bar{W}[k+1\ ,\ k\ ,\ F_2]\ ,\ F_1 \subseteq F_2\ ,$$

that follows from (15.9) we obtain in view of (15.8)

$$X[k] \subseteq \bar{W}(k)\ ,\ \text{for } k > 1 \tag{15.10}$$

Consider the following condition:

Assumption 15.1 The system (12.1), (12.3), $v\,[0\ ,\ s-1] = 0\ ,\ \xi[1\ ,\ s] = 0$ *is completely observable on* $[k_o\ ,\ s]$.

The given property is defined for example in [4].

In the latter case the following proposition is true:

Theorem 15.3 Under the conditions of Theorem 15.1 and assumption 15.1 assume $y[k_0\ ,\ s] = z[k_0\ ,\ s]$. *Then the equality*

$$X[s] = \{\bigcap \bar{W}(s)\ |\ P^0\ ,\ N(k{+}1)\ ,\ L(k)\ ,\ k = k_0\ ,\ldots,\ s-1\} \tag{15.11}$$

is true for any $P^0 > 0$ *and any diagonal* $N(k) > 0$, $L(k) > 0$. *Moreover for the given class of matrices we have*

$$\rho(\ell \mid X[s]) = f^{**}(\ell, s), f^{**}(\ell, s) = \inf \{ \rho(\ell \mid \bar{W}(s)) \mid P^0, L > 0, N > 0, k \in [k_0, s] \} \tag{15.12}$$

Therefore, the precise estimate is again attained here through a minimization procedure.

Remark 15.1 The relations (15.9), (15.10) may therefore be treated as follows

(a) In the case of a *set-membership* description of uncertainty as in (12.2), (12.4) with $u(k) \equiv 0$, $\eta(k) \equiv 0$, equations (15.9), (15.10) contain *complete information* on $X[k+1]$ as stated in Theorem 15.3.

(b) In the case of *both* set-membership and stochastic uncertainty, as in (15.3)-(15.5), equation (15.9) describes *the evolution of the set of the mean values of the estimates.*

(c) In the case of pure *stochastic* uncertainty with sets X^0, $P(k)$, $Q(k)$ consisting of one element $(x^0$, $p(k)$, $q(k))$ each, the relation (15.9) turns out to be an *equality* which coincides with the conventional equations of *Kalman's filtering theory.*

Remark 15.2 Following the scheme of Theorem 14.1 it is possible to demonstrate that relation (15.11) holds for P^0 , $N(k)$, $L(k)$ selected as follows:

$$P^0 = \beta \, I_n \, , \, N(k) = \alpha(k) \, I_m \, , \, L(k) = \beta(k) \, I_n$$

where

$$\beta > 0 \quad , \quad \alpha(k) > 0 \quad , \quad \beta(k) \geq 0 \quad , \quad k \in [k_0 \, , \, s]$$

Example

Consider a two-dimensional system

$$x(k+1) = \begin{bmatrix} 1, & \epsilon \\ -\epsilon\omega^2, & 1 \end{bmatrix} x(k) \tag{15.13}$$

with a scalar observation

$$y = x_1 + \xi, \quad \xi \in Q = \{\xi : |\xi| \leq \nu\} \; . \tag{15.14}$$

The initial state $x^0 \in X^0$ where $X^0 = x^*(0) + \mathrm{S}$, x^* is given and $\mathrm{S} = \{x : |x_i| \le 1;\ i = 1, 2\}$ is a square.

The aim is to estimate the state $x(k)$ at each stage k. Making use of formula (13.6) at each stage k, we will estimate $X[k+1] = X(k+1, k, X[k])$ by a rectangle $\boldsymbol{X}[k]$ oriented along the axes $\{x_1, x_2\}$. Here the calculations are as follows.

If X is a rectangle such that $X = x^* + \boldsymbol{X}$ where

$$\boldsymbol{X} = \{x : |x_1| \le \mu_1, |x_2| \le \mu_2\} \ ,$$

then

$$\rho(l|X) = (l, x^*) + \mu_1|l_1| + \mu_2|l_2| \tag{15.15}$$

Thus we may calculate some values of the function $\rho(l|X(k+1, k, X(k)))$ with $X(k)$ given. Using formula (13.6) for our example we have

$$F(M) = (I_n - M'G)A = \begin{pmatrix} 1 - m_1, & 0 \\ -\ m_2, & 1 \end{pmatrix} A = \begin{bmatrix} 1 - m_1, & \epsilon(1 - m_1) \\ -m_2 - \epsilon\omega^2, & -\epsilon m_2 + 1 \end{bmatrix}$$

$$M = (m_1, m_2)$$

Therefore

$$\begin{aligned} &\rho(l|X(k+1, k, X[k])) = \\ &= \inf\{\rho(l'F(M)|X[k]) + \rho(l'M'|y(k) - Q)\} \ , \end{aligned} \tag{15.16}$$

Starting with rectangle X^0 and calculating $\rho(l|X[1])$ for

$$l = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

due to formulae (15.15), (15.16), we define a rectangle $X[1] \supseteq X[1]$ - the "smallest" rectangle that includes $X[1]$ and is oriented along the axes $\{x_1, x_2\}$. Further on, taking $\boldsymbol{X}[1]$ instead of $X[1]$, and repeating the procedure, we come to a rectangle $\boldsymbol{X}[2]$ etc. Thus, after k stages, we will find a rectangle

$$\boldsymbol{X}[k] \supseteq X(k, 0, X^0) = X[k]$$

which is an upper estimate for $X[k]$.

The respective calculations were done for a system described by relations (15.13), (15.14) with $y(k)$ being an actual realization of the system generated by an initial vector $x^* \in X^0$ unknown to the observer and by an unknown "noise" $\xi(k)$ that attains either of the values $+\mu$ or $-\mu$ due to a random mechanism.

The results of the simulations for several starting sets X^0 are given in Figures 6–8 with $\epsilon = 0.2$, $\omega^2 = 1.2$, $\nu = 0.5$. In Figure 9 we have the same problem with an additional "horizontal" input disturbance

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} v(k)$$

added to the right hand part of (15.13), assuming $v(k)$ being unknown, random and uniformly distributed in the interval $-0.25 \leq v(k) \leq 0.25$. The calculations are the same as before except that due to (13.6) we have to substitute $\rho(l'F(M)|X(k))$ by

$$\rho(l'F(M)|X(k)) + \rho(l'(I_2 - M'G)|BP)$$

where

$$BP = \{p : p_1 = 0, |p_2| \leq 0.25\} \ .$$

The ideas of the above allow to approach *nonlinear systems*. Some of the basic facts related to guaranteed nonlinear filtering are given in the sequel.

16. Nonlinear Uncertain Systems

Consider a multistage process described by an n-dimensional *recurrence inclusion*

$$x(k+1) \in F(k, x(k)), \qquad k \geq k_0 \geq 0 \tag{16.1}$$

where $k \in [k_0, \infty)$, $x(k) \in \mathbb{R}^n$, $F(k, x(k))$ is a given multivalued map from $[k_0, \infty) \times \mathbf{R}^n$ into comp $\mathbf{R}^n$.

As before suppose the initial state $x(k_0) = x^0$ of the system is confined to a preassigned set:

$$x^0 \in X^0 \ , \tag{16.2}$$

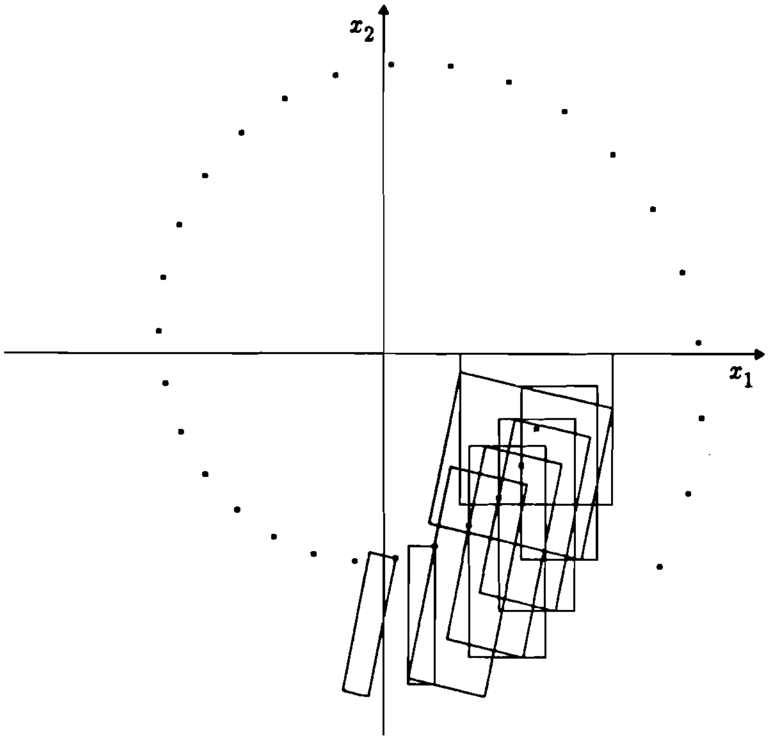

FIGURE 6

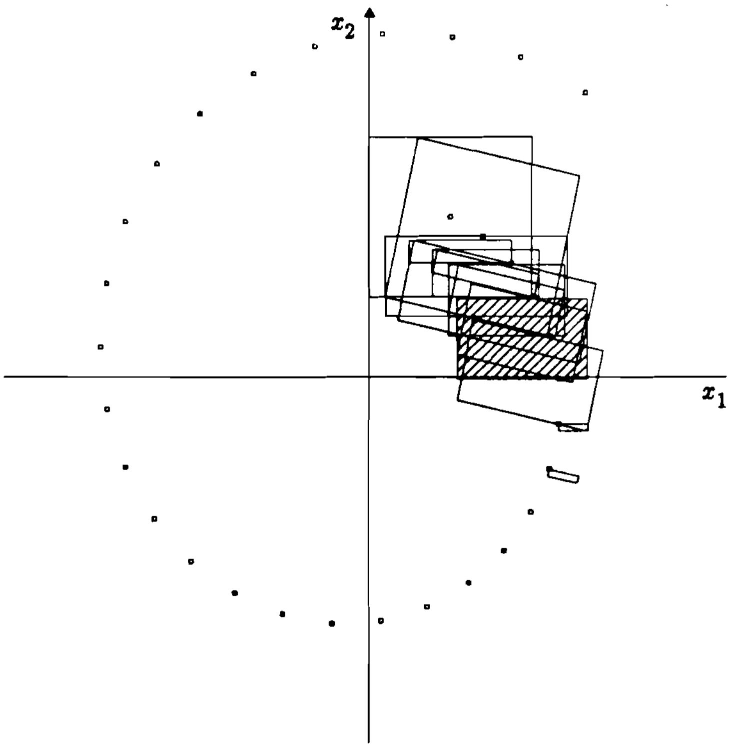

FIGURE 7

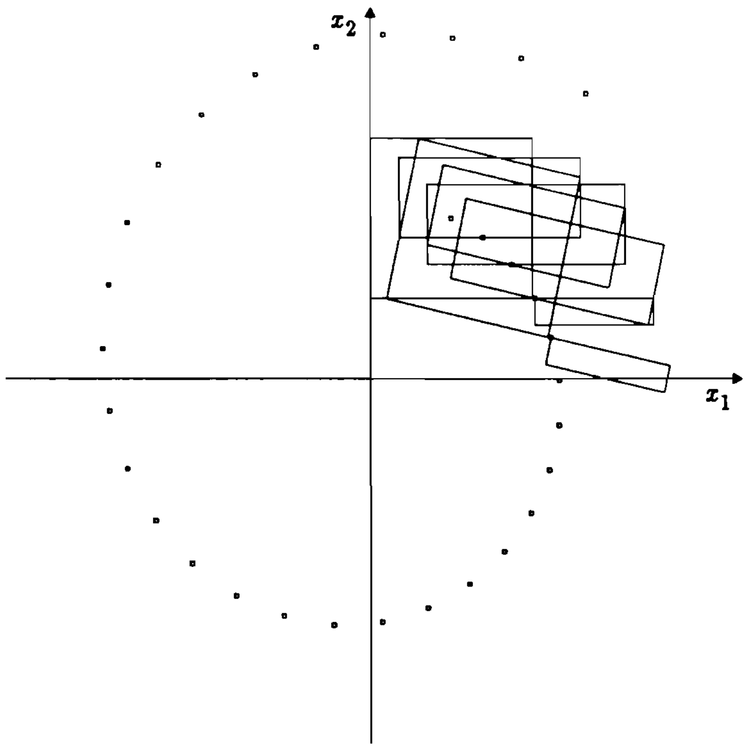

FIGURE 8

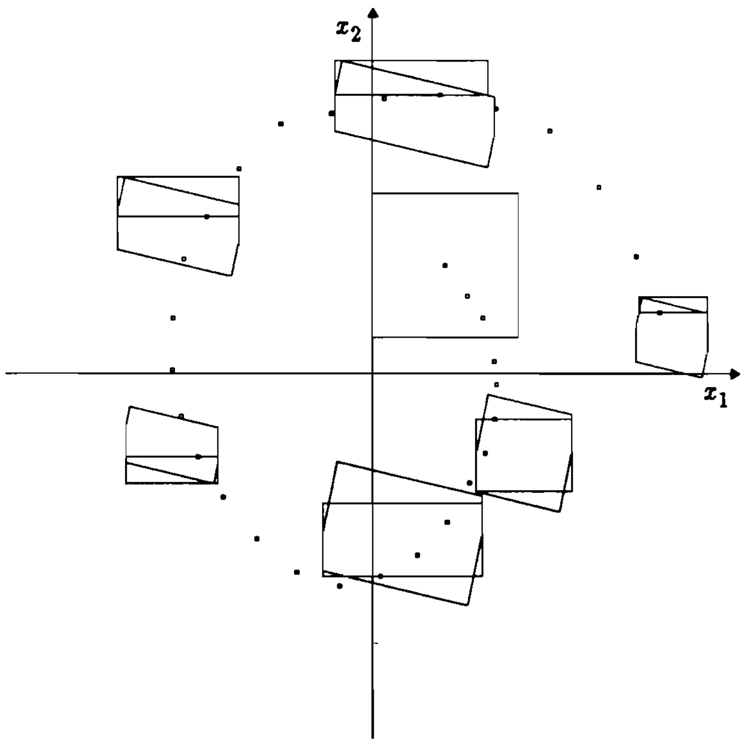

FIGURE 9

Let $Q(k)$ be a multivalued map from $[k_0\ ,\infty)$ into comp $\mathbb{R}^m$ and $G(k)$ - a single-valued map from $[k_0\ ,\infty]$ into the set of $m\times n$-matrices. The pair $G(k)$, $Q(k)$, ***introduces a state constraint***

$$G(k)x(k)\in Q(k)\,,\quad k\geq k_0{+}1\ , \tag{16.3}$$

on the solutions of system (16.1).

The subset of $\mathbb{R}^n$ that consists of all the points of $\mathbb{R}^n$ through which at stage $s\in[k_0,\infty)$ there passes at least one of the trajectories $x(k\ ,\ k_0\ ,\ x^0)$, that satisfy constraint (16.3) for $k\in[k_0,\tau]$, will be denoted as $X(s\mid\tau,k_0,x^0)$.

If set $Q(k)$ of (16.3) is of a specific type

$$Q(k)=y(k)-\tilde{Q}(k)$$

where $y(k)$ and $\tilde{Q}(k)$ are given, then (16.3) transforms into

$$y(k)\in G(k)x(k)+\tilde{Q}(k) \tag{16.4}$$

which could be interpreted as an *equation of observations* for the uncertain system (16.1) given above. Sets $X(s\mid\tau,k_0,X^0)$ therefore give us *guaranteed estimates* of the unknown states of system (16.1) on the basis of an observation of vector $y(k)$, $k\in[k_0,\tau]$ due to equation (16.4).

For various relations between s and τ this reflects the following situations

(a) for $s=\tau$ - the problem of *"guaranteed filtering"*

(b) for $s>\tau$ - the problem of *"guaranteed prediction"*

(c) for $s<\tau$ - the problem of *"guaranteed refinement"*

The aim of this paper will first be to study the informational sets $X(\tau|\tau\ ,\ k_0\ ,\ X^0)=X(\tau\ ,\ k_0\ ,\ X^0)$ similar to those of the above and their evolution in "time" τ.

The sets $X(k,k^0,x^0)$ may also be interpreted as *attainability domains* for system (16.1) under the state space constraint (16.3). The objective is therefore to describe the evolution of these domains. A further objective will be to describe the more complicated sets $X(s \mid \tau, k_0, x^0)$ and their evolution

17. A Generalized Nonlinear Dynamic System

From the definition of sets $X(s \mid \tau, k^0, x^0)$ it follows that the following properties are true.

Lemma 17.1. *Whatever are the instants* $t,s,k,$ $(t \geq s \geq k \geq 0)$ *and the set* $\mathbf{F} \in \mathrm{comp}\,\mathbb{R}^n$, *the following relation is true*

$$\boldsymbol{X}(t,k,\mathbf{F}) = \boldsymbol{X}(t,s,\boldsymbol{X}(s,k,\mathbf{F}))\,. \tag{17.1}$$

Lemma 17.2. *Whatever are the instants* $s,t,\tau,k,l(t \geq s \geq l;\ \tau \geq l \geq k;\ t \geq \tau)$ *and the set* $\mathbf{F} \in \mathrm{comp}\,\mathbb{R}^n$ *the following relation is true*

$$X(s \mid t,k,\mathbf{F}) = X(s \mid t,l,X(l \mid \tau,k,\mathbf{F}))\,. \tag{17.2}$$

Relation (17.1) shows that sets $\boldsymbol{X}(k,\tau,X)$ again satisfy a *semigroup property* which allows to define a *generalized dynamic system* in the space $2^{\mathbb{R}^n}$ of all subsets of $\mathbb{R}^n$. On the other hand, (17.2) is a more general relation which is true when the respective intervals of observation may overlap.

In general the sets $X(s \mid t,k,\mathbf{F})$ need not be either convex or connected. However, it is obvious that the following is true

Lemma 17.3. *Assume that the map F is linear in x:*

$$F(k\ ,\ x) = A(k)x + P$$

where $P \in \mathrm{conv}\,\mathbb{R}^n$. *Then for any set* $\mathbf{F} \in \mathrm{conv}\,\mathbb{R}^n$ *each of the sets* $X(s \mid t,k,\mathbf{F}) \in \mathrm{conv}\,\mathbb{R}^n (t \geq s \geq k \geq 0)$.

Therefore the next step will be to describe the evolution of the set $X[k] = X(k\ ,\ k_o\ ,\ X^o)$. This will be later given in the form of a decoupling procedure. However it is convenient to commence with a description of the one-stage problem.

18. The One-Stage Problem

Consider the system

$$z \in F(x), \quad Gz \in Q, \quad x \in X,$$

where $z \in \mathbf{R}^n$, $X \in \mathrm{comp}\,\mathbf{R}^n$, $Q \in \mathrm{conv}\,\mathbb{R}^m$, $F(\kappa)$ is a multivalued map from $\mathbf{R}^n$ into $\mathrm{conv}\,\mathbf{R}^n$, G is a linear (single-valued) map from $\mathbb{R}^n$ into $\mathbb{R}^m$.

It is clear that the sets $F(X) = \{ \bigcup F(x) \mid x \in X \}$ need not be convex.

Let Z, Z^* respectively denote the sets of all solutions to the following systems:

(a) $z \in F(X), \quad Gz \in Q,$

(b) $z^* \in \mathrm{co}\,F(X), \quad Gz^* \in Q,$

It is obvious that the following statement is true

Lemma 18.1. The sets Z, $\mathrm{co}\,Z$, Z^* satisfy the following inclusions

$$Z \subseteq \mathrm{co}\,Z \subseteq Z^* \tag{18.1}$$

Denote

$$\Phi(l,p,q) = (l - G'p, q) + \rho(-p \mid Q)$$

Then the function $\Phi(l,p,q)$ may be used to describe the sets $\mathrm{co}\,Z, Z^*$. The techniques of nonlinear analysis yield

Lemma 18.2. The following equalities are true

$$\rho(l \mid Z) = \rho(l \mid \mathrm{co}\,Z) = \sup_q \inf_p \Phi(l,p,q) \;, \quad q \in F(X),\, p \in \mathbb{R}^m \tag{18.2}$$

$$\rho(l \mid Z^*) = \inf_p \sup_q \Phi(l,p,q) \;, \quad q \in F(X),\, p \in \mathbb{R}^m \tag{18.3}$$

The sets $\mathrm{co}\ Z$, Z^* are convex due to their definition. However it is not difficult to give an example of a nonlinear map $F(x)$ for which Z is nonconvex and the functions $\rho(l \mid \mathrm{co}\,Z)$, $\rho(l \mid Z^*)$ do not coincide, so that the inclusions $Z \subset \mathrm{co}\,Z$, $\mathrm{co}\,Z \subset Z^*$ are strict.

Indeed, assume $X = \{0\}$, $x \in \mathbf{R}^2$

$$F(0) = \{x : 6x_1 + x_2 \leq 3 \;,\; x_1 + 6x_2 \leq 3 \;,\; x_1 \geq 0 \;,\; x_2 \geq 0\}$$
$$G = (0 \;,\; 1) \;,\; Q = (0 \;,\; 2) \;.$$

Then

$$Y = \{x : Gx \in Q\} = \{x : 0 \le x_2 \le 2\}$$

The set $F(0)$ is a nonconvex polyhedron O K D L in Figure 10a while set Y is a stripe. Here, obviously, set Z which is the intersection of $F(0)$ and Y, turns to be a nonconvex polyhedron O A B D L, while sets co Z , Z^* are convex polyhedrons O A B L and O A C L respectively (see Figures 10b and 10c). The corresponding points have the coordinates

A = (0 , 2), B = (1/2 , 2), C = (1 , 2), D = (3/7 , 3/7), K = (0 , 3), L = (3 , 0),
O = (0 , 0).

Clearly $Z \subset \text{co } Z \subset Z^*$.

This example may also serve to illustrate the existence of a "duality gap", [21] between (18.2) and (18.3).

For a linear-convex map $F(x) = Ax + P$ $(P \in \text{conv}\, \mathbb{R}^n)$ there is no distinction between Z, co Z, and Z^*:

Lemma 18.3 *Assume $F(x) = Ax + P$ where $P \in \text{conv}\, \mathbb{R}^n$, A is a linear map from $\mathbb{R}^n$ into $\mathbb{R}^n$. Then $Z = \text{co}\, Z = Z^*$.*

The description of Z, co Z, Z^* may however be given in a "decoupled" form which, allows to present all of these sets as the intersections of some parametrized varieties of convex multivalued maps of relatively simple structure.

19. The One Stage Problem - A Decoupling Procedure.

Whatever are the vectors $l, p (l \neq 0)$ it is possible to present $p = M'l$ where M belongs to the space $\mathbb{M}^{m \times n}$ of real matrices of dimension $m \times n$. Then, obviously,

$$\rho(l \mid Z) = \sup_q \inf_M \Phi(l\, , M'l\, , q) = \rho(l \mid \text{co}\, Z)\ ,\ q \in F(X),\ M \in \mathbb{M}^{n \times m}, \tag{19.1}$$

$$\rho(l \mid Z^*) = \inf_M \sup_q \Phi(l\, , M'l\, , q) \qquad q \in F(X),\ M \in \mathbb{M}^{n \times m}$$

or

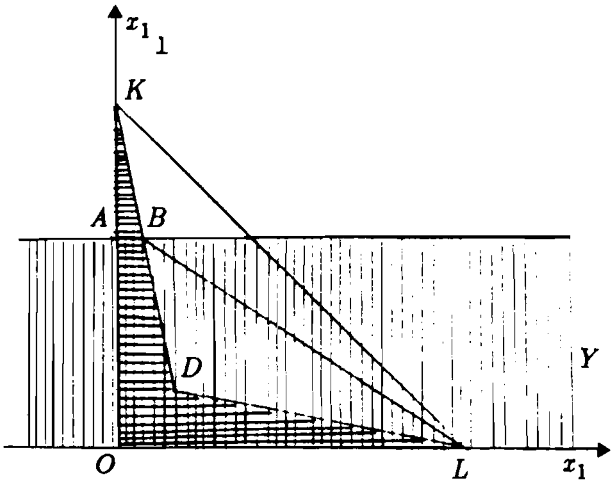

FIGURE 10a

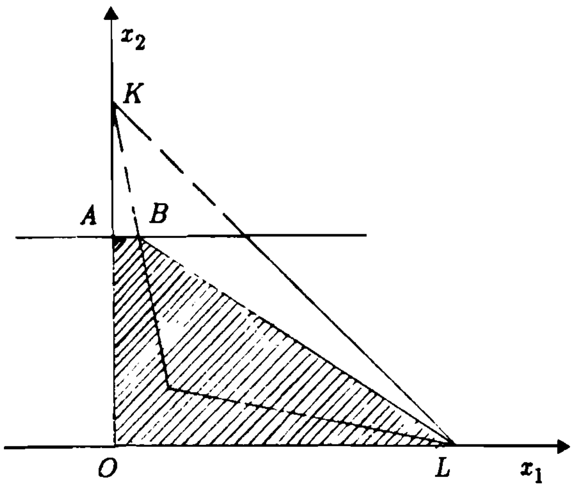

FIGURE 10b

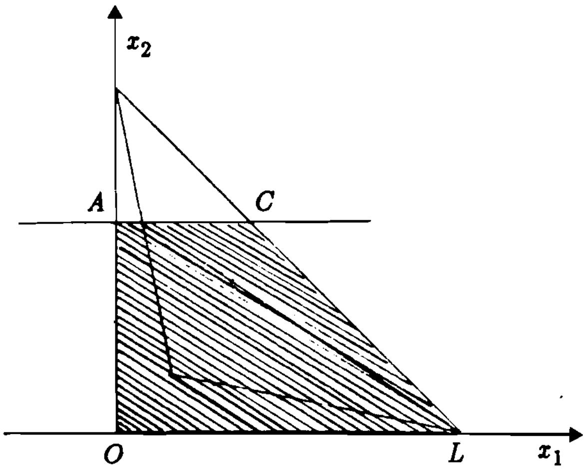

FIGURE 10c

$$\rho(l \mid Z^*) = \inf\{\Phi(l\,,\,M'l) \mid M \in \mathbf{M}^{n\times m}\}\,, \tag{19.2}$$

where

$$\Phi(l\,,\,M'l) = \bigcup\{\Phi(l\,,\,Ml\,,\,q) \mid q \in \mathrm{co}\,\boldsymbol{F}(x)\} =$$
$$= \rho((E - G'\;M')l \mid \mathrm{co}\,\boldsymbol{F}(X)) + \rho(-M'l \mid Q)\;\;.$$

From (19.1) it follows

$$Z \subseteq \bigcup_{q \in \boldsymbol{F}(X)} \bigcap_{M} R(M,q) \subseteq \bigcap_{M} \bigcup_{q \in \boldsymbol{F}(X)} R(M,q),\; M \in \mathbf{M}^{n\times m} \tag{19.3}$$

where

$$R(M,q) = (E_n - MG)q - MQ\,.$$

Similarly (19.2) yields

$$Z^* \subseteq \bigcap_{M} \bigcup_{q \in \mathrm{co}\,\boldsymbol{F}(X)} \{(E_n - MG)q - MQ\}\,. \tag{19.4}$$

Moreover a stronger assertion holds.

Theorem 19.1. *The following relations are true*

$$Z = Z(X) = \bigcup_{q \in \boldsymbol{F}(X)} \bigcap_{M} R(M,q) \tag{19.5}$$

$$Z* = Z*(X) = \bigcap_{M} R(M,\mathrm{co}\,\boldsymbol{F}(X)) \tag{19.6}$$

where $M \in \mathbf{M}^{m\times n}$.

Obviously for $\boldsymbol{F}(x) = AX + P,(X,P \in \mathrm{co}\,\mathbb{R}^n)$ we have $\boldsymbol{F}(X) = \mathrm{co}\,\boldsymbol{F}(X)$ and $Z = Z^* = \mathrm{co}\,Z$.

This *first* scheme of relations may serve to be a basis for constructing multistage procedures. Another procedure could be derived from the following *second* scheme. Consider the system

$$z \in \boldsymbol{F}(x) \tag{19.7}$$

$$Gx \in Q, \tag{19.8}$$

for which we are to determine the set of all vectors z consistent with inclusions (19.7), (19.8). Namely, we are to determine the restriction $\boldsymbol{F}_Y(x)$ of $\boldsymbol{F}(x)$ to set Y. Here we have

$$\boldsymbol{F}_Y(x) = \begin{cases} \boldsymbol{F}(x) & \text{if } x \in Y \\ \phi & \text{if } x \in Y \end{cases}$$

where as before $Y = \{ x : Gx \in Q \}$.

Lemma 19.1 *Assume $F(x) \in \mathrm{comp}\, \mathbb{R}^n$ for any x and $Q \in \mathrm{conv}\, \mathbb{R}^m$. Then*

$$F_Y(x) = \bigcap_L (F(x) - LGx + LQ)$$

over all $n \times m$ matrices L, $(L \in \mathbf{M}^{n \times m})$.

Denote the null vectors and matrices as $\{0\}_m \in \mathbf{R}^m$, $\{0\}_{m,n} \in \mathbf{R}^{m \times n}$, and the $(n \times m)$ matrix L_{mn} as

$$L_{mn} = \begin{bmatrix} I_m \\ \{0\}_{n-m,\,m} \end{bmatrix}$$

Suppose $x \in Y$. Then $\{0\}_m \in Q - Gx$ and for any $(n \times m)$ -matrix L we have $\{0\}_n \in L\,(Q - Gx)$. Then it follows that for $x \in Y$.

$$F(x) \subseteq \bigcap_L (F(x) + L(Q - Gx)) \subseteq F(x)$$

On the other hand, suppose $x \bar{\in} Y$.

Let us demonstrate that in this case

$$\bigcap_L \{ F(x) + L(Q - Gx) \} = \phi .$$

Denote $A = F(x)$, $B = Q - Gx$. For any $\lambda > 0$ we then have

$$\bigcap_L (A + LB) \subseteq (A + \lambda L_{mn} B) \cap (A - \lambda L_{mn} B)$$

Since $\{ 0 \}_m \notin B$ we have $\{ 0 \}_n \notin L_{mn} B$. Therefore there exists a vector $l \in \mathbb{R}^n$, $l \neq 0$ and a number $\gamma > 0$ such that

$$(l,x) \geq \gamma > 0 \quad \text{for any} \quad x \in L_m B ,$$

Denote

$$\mathbb{L} = \{ x : (l,x) \geq \gamma \} .$$

Then $\mathbb{L} \supseteq L_m B$ and

$$(A + \lambda L_{mn} B) \cap (A - \lambda L_{mn} B) \subseteq (A + \lambda \mathbb{L}) \cap (A - \lambda \mathbb{L})$$

Set A being bounded there exists a $\lambda > 0$ such that

$$(A + \lambda \mathbb{L}) \cap (A - \lambda \mathbb{L}) = \phi .$$

Hence

$$\bigcap_L (A + LB) = \phi$$

and the Lemma is proved.

If in addition to (19.7), (19.8) we have

$$x \in X \tag{19.9}$$

then the set Z_o consistent with (19.7)-(19.9) may be presented as

$$Z_0(X) = \bigcup_{x \in X} \bigcap_L (F(x) - LGx + LQ) \tag{19.10}$$

Therefore each of the sets $Z(x)$, $Z^o(x)$ $(x \in X)$ may be respectively decoupled into the calculation of either set-valued functions $R(M \, , \, q)$ or

$$R_o \, (L \, , \, x) = F(x) - LGx + LQ$$

according to (19.5), (19.10). It may be observed that each of these are also applicable when $Z(X)$, $Z_0(X)$ are *disconnected*.

In the linear-convex case

$$F(x) = Ax + P \, , \, P \in \text{conv } \mathbf{R}^n \, ,$$

we have

$$Z(X) = \bigcap_M \{(E - MG)(AX + P) + MQ\}$$

$$Z_0(x) = \bigcap_L \{(A - LG)X + P + LQ\}$$

20. Solution to the Problem of Nonlinear "Guaranteed" Filtering

Returning to system (16.1)–(16.3) we will look for the sequence of sets $\boldsymbol{X}[s] = \boldsymbol{X}(s, k_0, X^0)$ together with two other sequences of sets. These are

$$\boldsymbol{X}^*[s] = \boldsymbol{X}^*(s, k_0, X^0)$$

- the solution set for system

$$x(k+1) \in \operatorname{co} F(k, X^*[k]), \quad X^*[k_0] = X^0 \tag{20.1}$$

$$G(k+1)\; x(k+1) \in Q(k+1), \; k \geq k_0 \tag{20.2}$$

and $X_*[s] = X_*(s, k_0, X^0)$ which is obtained due to the following relations:

$$X_*[s] = \operatorname{co} Z[s] \tag{20.3}$$

where $Z[k+1]$ is the solution set for the system

$$z(k+1) \in F(k, X_*[k]), \quad Z[k^0] = X^0, \tag{20.4}$$

$$G(k+1) z(k+1) \in Q(k+1), \quad k \geq k_0 . \tag{20.5}$$

The sets $X_*[\tau]$, $X^*[\tau]$ are obviously convex. They satisfy the inclusions

$$X[\tau] \subseteq X_*[\tau] \subseteq X^*[\tau]$$

while each of the sets $X[\tau]$, $X_*[\tau]$, $X^*[\tau]$ lies within

$$Y(\tau) = \{ x : G(\tau) x \in Q(\tau) \}, \quad \tau \geq k_0 + 1 ,$$

The sets $X[\tau]$, $X_*[\tau]$, $X^*[\tau]$ may therefore be obtained by solving sequences of problems

$$x(k+1) \in F(k \; , x(k)) \tag{20.6}$$

$$G(k+1)\; x(k+1) \in Q(k) \; , \qquad\qquad k \geq k_o \tag{20.7}$$

for $X[s]$, (20.1), (20.2) for $X^*[s]$ and (20.3) - (20.5) for $X_*[s]$

In order to solve the "guaranteed" filtering problem with $Q(k) = y(k) - \tilde{Q}(k)$ one may follow *the first scheme* of § 19, considering the multistage system

$$Z(k+1) = (I_n - M(k+1) G(k+1)) F^0(k, S(k)) + M(k+1)(y(k+1) \; - \; \tilde{Q}(k+1)) \tag{20.8}$$

$$S(k) = \{ \cap Z(k) \mid M(k) \}, \;\; k > k_0 , \; S(k_0) = X^0, \tag{20.9}$$

where $M(k+1) \in \mathbb{R}^{n \times m}$.

From Theorem 19.1 one may now deduce the following result

Theorem 20.1 *The solving relations for $X[s]$, $X_*[s]$, $X^*[s]$ are as follows*

$$X[s] = S(s) \quad \text{for} \quad F^0(k, S(k)) = F(k, S(k)) \tag{20.10}$$

$$X^*[s] = S(s) \quad \text{for} \quad F^0(k, S(k)) = \operatorname{co} F(k, S(k)) \tag{20.11}$$

$$X_*[s] = \operatorname{co} S(s) \quad \text{for} \quad F^0(k, S(k)) = F(k, \operatorname{co} S(k)) \; . \tag{20.12}$$

It is obvious that $X[\tau]$ is the *exact solution* for the *guaranteed filtering* problem while $X_*[\tau]$, $X^*[\tau]$ are *upper convex majorants* for $X[\tau]$. It is clear that by interchanging and combining relations (20.11), (20.12) from stage to stage it is possible to construct a broad variety of other convex majorants for $\boldsymbol{X}[\tau]$. However for the linear case they will all coincide with $\boldsymbol{X}[\tau]$.

Lemma 20.1 *Assume $F^0(k,S) = A(k)S + P(k)$ with $P(k)$, X^0 being convex and compact. Then $X[k] = X^*[k] = X_*[k]$ for any $k \geq k_0$.*

Consider the nonlinear system

$$\begin{aligned} Z(k+1) &= (I_n - M(k+1)G(k+1))F^0(k,Z(k)) \\ &\quad + M(k+1)(y(k+1) - \tilde{Q}(k+1)) \,, \qquad\qquad Z(k_0) = X^0 \,, \end{aligned}$$

having denoted its solution as

$$\begin{aligned} &Z(k;M_k(\cdot)) \quad \text{for} \quad F^0(k,Z) = F(k,Z) \\ &Z_*(k,M_k(\cdot)) \quad \text{for} \quad F^0(k,Z) = F(k,\text{co}\, Z) \\ &Z^*(k,M_k(\cdot)) \quad \text{for} \quad F^0(k,Z) = \text{co}\, F(k,Z) \end{aligned}$$

Then theorem 20.1 yields the following conclusion

Theorem 20.2 *Whatever is the sequence $M_s(\cdot)$, the following solving inclusions are true*

$$\begin{aligned} &\boldsymbol{X}[s] \subseteq Z(s,M_s(\cdot)) \\ &\boldsymbol{X}_*[s] \subseteq Z_*(s,M_s(\cdot)) \\ &\boldsymbol{X}^*[s] \subseteq Z^*(s,M_s(\cdot)) \,, \quad s > k_0 \,, \end{aligned} \tag{20.13}$$

with $Z(s,M_s(\cdot)) \subseteq Z_*(s,M_s(\cdot)) \subseteq Z^*(s,M_s(\cdot))$.

Hence we also have

$$\boldsymbol{X}[s] \subseteq \bigcap \{ Z(s,M_s(\cdot)) \mid M_s(\cdot)) \} \tag{20.14}$$

$$\boldsymbol{X}_*[s] \subseteq \bigcap \{ Z_*(s,M_s(\cdot)) \mid M_s(\cdot)) \} \tag{20.15}$$

$$\boldsymbol{X}^*[s] \subseteq \bigcap \{ Z^*(s,M_s(\cdot)) \mid M_s(\cdot)) \} \tag{20.16}$$

over all $M_s(s)$.

However a question arises which is whether (20.14)–(20.16) could turn into exact equalities.

Lemma 20.2 *Assume the system (16.1), to be linear: $F(k,x) = A(k)x + P(k)$ with sets $P(k)$, $Q(k)$ convex and compact. Then*

$$X[s] = X^*[s] = \bigcap \{ Z_s(\cdot, M_s(\cdot)) \} \tag{20.17}$$

where $Z_s(\cdot\ M_s(\cdot))$ is the solution tube for the equation

$$Z(k+1) = (I_n - M(k+1)G(k+1))(A(k)Z(k) + P(k)) + \\ + M(k+1)(y(k+1) - \tilde{Q}(k+1)),\ Z(k_o) = X^o \tag{20.18}$$

Hence in this case the intersections over $M(k)$ could be taken *either at each stage* as in (20.10), (20.11) *or at the final stage* as in (20.17).

Let us now follow *the second scheme* of § 19, considering the equation

$$x(k+1) \in \tilde{F}_{Y(k)}(k, x(k)),\quad x^0 = x(k_0),\quad x^0 \in X^0\ , \tag{20.19}$$

and denoting the set of its solutions that start at $x^0 \in X^0$ as $X^0(k, k_0, x^0)$ as

$$\bigcup \{ x^0(k, k_0, x^0) \mid x^0 \in X^0 \} = X^0(k, k_0, X^0) = X^0[k]\ .$$

According to Lemma 19.1 we may substitute (20.19) by equation

$$x(k+1) \in \bigcap_L (\tilde{F}\ (k, x(k)) - LG(k)x(k) + LQ(k))\ ,\quad x^0 \in X^0,$$

The calculation of $X^0[k]$ should hence follow the procedure of (19.10)

$$\tilde{X}[k+1] = \bigcup_{x \in \tilde{X}(k)} \bigcap_L (\tilde{F}(k,x) - LG(k)x + LQ(k)),\qquad X(k_0) = X^0\ . \tag{20.20}$$

Denote the "whole" solution tube for this solution $(k_0 \le k \le s)$ as $\tilde{X}^s_{k_0}[\cdot]$. Then the following assertion will be true.

Theorem 20.3 Assume $\tilde{X}^s_{k_0}[k]$ to be the cross-section of the tube $\tilde{X}^s_{k_0}[\cdot]$ at instant k and $X^o = X^o \bigcap Y(k_o)$. Then

$$X[s] = \tilde{X}^{s+1}_{k_0}[s] \text{ if } \tilde{F}(k,x) = F(k,x)\ ,$$
$$X^* = \tilde{X}^{s+1}_{k_0}[s] \text{ if } \tilde{F}(k,x) = \operatorname{co} F(k,x)$$

Here $\tilde{X}^s_{k_0}[s] \supseteq \tilde{X}^{s+1}_{k_0}[s]$ and the set $\tilde{X}^s_{k_0}[s]$ may not lie totally within $Y(s)$, while always $\tilde{X}^{s+1}_{k_o}\ [s] \subseteq Y(s)$.

Solving equation (20.19) is equivalent to finding all the solutions for the inclusion

$$x(k+1) \in \bigcap_L (\tilde{F}(k, x(k)) + L(y(k) - G(k)\ x(k) - \tilde{Q}(k))\ ,\ x(k_0) \in X^0 \tag{20.21}$$

Equation (20.21) may now be "decoupled"into a system of "simpler" inclusions

$$x(k+1) \in \tilde{F}(k,x(k)) + L(k)\ (y(k) - G(k)x(k)) - L(k)\ \tilde{Q}(k)\ , \quad x(k_0) \in X^0 \quad (20.22)$$

for each of which the solution set for $k_0 \leq k \leq s$ will be denoted as

$$\tilde{X}^s_{k_0}(\cdot, k_0, X^0, L(\cdot)) = \tilde{X}^s_{k_0}[\cdot, L(\cdot)]$$

Theorem 20.4 *The set $X^s_{k_0}[\cdot]$ of solutions to the inclusion*

$$x_{k+1} \in \tilde{F}(k,x(k))\ , \quad x(k_0) \in X^0$$
$$y(k) \in G(k)x(k) + \tilde{Q}(k), \quad k_0 \leq k \leq s$$

is the part of the solution tube

$$X^{s+1}_{k_0}[\cdot] = \bigcap_L \tilde{X}^{s+1}_{k_0}[\cdot, L]\ , [k_0, \ldots, s+1]$$

which is restricted to stages $[k_0\ , s]$. Here the intersection may be taken only over all constant matrices $L(k) \equiv L$.

This scheme also allows to calculate the cross sections $X^s_{k_0}[s]$. Obviously

$$X^s_{k_0} \subseteq \bigcap_{L[\cdot]} \tilde{X}^{s+1}_{k_0}[s, L[\cdot]] \quad (20.23)$$

over all sequences $L[\cdot] = \{ L(k_0), L(k_0+1), \ldots, L(s+1) \}$. Moreover the following proposition is true, and may be compared with [5, 9-11].

Theorem 20.5 *Assume $\tilde{F}(k,x)$ to be linear-convex: $\tilde{F}(k,x) = A(k)x + P(k)$, with $P(k)$, $Q(k)$ convex and compact. Then (20.23) turns to be an equality.*

The next estimation problems are those of "prediction" and "refinement".

21. The "Guaranteed Prediction" Problem

The solution to the guaranteed prediction problem is to specify set $X(s \mid t\ , k_o\ , X^o)$ for $s \geq t$ It may be deduced from the previous relations due to (17.2) since

$$X(s \mid t\ , k_o\ , X^o) = X(s \mid t\ , X(t\ , k_o\ , X^o))$$

Similarly we may introduce set

$$X^*(s \mid t\ , k_o\ , X^o) = X^*(s \mid t\ , X^*(s \mid t\ , X^*(t\ , k_o\ , X^o))$$

where $X^*(s \mid t, x)$ is the *attainability domain* for the inclusion

$$x(k+1) \in \text{co } F(k, x(k))$$

with $t \leq k \leq s$, $\quad x(t) = x$

The description of $X(s \mid t, k_o X^o)$, $X^*(s \mid t, k_o X^o)$ may be given through a modification of theorems 20.1 - 20.5, by the following assertion

Theorem 21.1 The solving relations for the prediction problem are

$$X(s \mid t, k_o, X^o) = X[s]$$

$$X^*(s \mid t, k_o, X^o) = X^*[s]$$

where $X[s]$, $X^*[s]$ are determined through (20.10), (20.12), (20.8), under the condition

$$S(k) = \{\bigcap Z(k) \mid M(k) \in \mathbf{R}^{n \times m}\}$$
$$S(k) = Z(k) \text{ for } k > t$$

For the linear convex case an alternative presentation is true. Denote $L_t^s(\cdot) = \{L(k_o), \ldots, L(s)\}$ to be a sequence of $(n \times m)$ - matrices $L(i)$, $k_o \leq i \leq s$, such that $L(i) \equiv 0$ for $t < i \leq s$.

Theorem 21.2 Assume $F(k, x) = A(k)x + P$ with P, X^o convex and compact. Then

$$X(s \mid t, k_o, X^o) = \{\bigcap \tilde{X}_{k_o}^s [s, L_s^t(\cdot)] \mid L_s^t(\cdot)\} \tag{21.1}$$

The solution to the prediction problem may therefore be decoupled into the calculation of the attainability domains $\tilde{X}_{k_o}^s [s, L_t^s(\cdot)]$ for the variety of systems

$$x(k+1) \in (A(k) - L(k)\, G(k))\, x(k) + L(k)\, y(k) + L(k)\, \tilde{Q}(k) + P(k) \tag{21.2}$$
$$L(k) \equiv 0 \text{ for } k > t$$

each of which starts its evolution from X^o.

The forthcoming "refinement" problem is a deterministic version of the interpolation problem of stochastic filtering theory.

22. The "Guaranteed" Refinement Problem

Assume the sequence $y[k\ ,\ t]$ to be fixed. Let us discuss the means of constructing sets $X(s \mid t\ ,\ k\ ,\ \mathbf{F})$, with $s \in [k\ ,\ t]$. From relation (17.2) one may deduce the assertion

Lemma 22.1 The following equality is true

$$X(s \mid t\ ,\ k\ ,\ \mathbf{F}) = X(s \mid s\ ,\ t\ ,\ X(t\ ,\ k\ ,\ \mathbf{F})) \tag{22.1}$$

Here the symbol $X(s \mid s\ ,\ t\ ,\ \mathbf{F})$, taken for $s \le t$, stands for the set of states $x(s)$ that serve as starting points for all the solutions $x(k\ ,\ s\ ,\ x(s))$ that satisfy the relations

$$x(k+1) \in F(k\ ,\ x(k))\ ,\ x(t) \in \mathbf{F}$$
$$x(k) \in Y(k)\ , \qquad s \le k \le t$$

Corollary 22.1 Formula (22.1) may be substituted for

$$X(s \mid t\ ,\ k\ ,\ \mathbf{F}) = X(s\ ,\ k\ ,\ \mathbf{F}) \bigcap X(s \mid s\ ,\ t\ ,\ \mathbf{K}) \tag{22.2}$$

where $\mathbf{K}$ *is any subset of* $\mathbf{R}^n$ *that includes* $X(t\ ,\ k\ ,\ \mathrm{F})$.

Thus the set $X(s \mid t\ ,\ k\ ,\ \mathbf{F})$ is described through the solutions of two problems the first of which is to define $X(s\ ,\ k\ ,\ \mathbf{F})$ (along the techniques of the above) and the second is to define $X(s \mid s\ ,\ t\ ,\ \mathbf{K})$. The solution of the second problem will be further specified for $\mathbf{F} \in \mathrm{comp}\, \mathbf{R}^n$ and for a closed convex Y.

The underlying elementary operation is to describe $\boldsymbol{X}$ - the set of all the vectors $x \in \mathbf{R}^n$ that satisfy the system

$$z \in F(x)\ , \quad z \in Y$$
$$(\boldsymbol{X} = \{x : F(x) \bigcap Y \ne \emptyset\})$$

Using suggestions similar to those applied in Lemma 19.1 we come to

Lemma 22.2 The set $\boldsymbol{X}$ *may be described as*

$$\boldsymbol{X} = \bigcup \{ \bigcap \{Ex - MF(x) + MY \mid M \in \mathbb{M}^{n \times n}\} \mid x \in \mathbf{R}^n\}$$

From here it follows:

Theorem 22.1 The set $X(s \mid s\ ,\ t\ ,\ \mathbf{R})$ *may be described as the solution of the multistage system (in backward "time")*

$$X[k] = Y(k) \bigcap X[k] \tag{22.3}$$

where

$$\boldsymbol{X}[k] = \bigcup \{ \bigcap \{Ex - MF(x) + MX[k+1] \mid M \in \mathbf{M}^{n \times n} \} \mid x \in \mathbf{R}^n \} ,$$
$$s \leq k \leq t , X[t] = Y[t] .$$

Finally we will specify the solution for the linear case

$$x(k+1) \in A(k)\, x(k) + P(k) , \; Y(k) = \{x : y(k) \in G(k)x + Q(k)\}$$

Assume

$$X = \{x : z \in Ax - P , x \in Y , z \in Z\} \; , \; Y = \{x : Gx \in Q - y\} \tag{22.4}$$

where $A \in \mathbf{M}^{n \times n}$, $G \in \mathbf{M}^{m \times n}$, P , Q , Z are convex and compact.

Lemma 22.3 The set X may be defined as

$$\rho(l \mid X) = \inf\{\rho(\lambda \mid P) + \rho(\lambda \mid Z) + \rho(p \mid Q - y)\}$$

over all the vectors $\lambda \in \mathbf{R}^n$, $p \in \mathbf{R}^m$ that satisfy the equality $l = A' \lambda + G'p$.

The latter relation yields:

Lemma 22.4 The set X may be defined as

$$X \subseteq L'(Z + P) + M'(Q - y) = H(L , M) \tag{22.5}$$

whatever are the matrices $L \in \mathbf{M}^{n \times n}$ and $M \in \mathbf{M}^{m \times n}$ that satisfy the equality $L' A + M'G = E_n$. Moreover the following equalities are true

$$X = \{\bigcap H(L , M) \mid L , M\} \tag{22.6}$$
$$\rho(l \mid X) = \inf \{\rho(l \mid H(L , M)) \mid L,M \}$$

over all $L \in \mathbf{M}^{n \times n}$, $M \in \mathbf{M}^{m \times n}$.

Corollary 22.2 Suppose $\mid A \mid \neq 0$. Then conditions (22.5), (22.6) may be substituted for

$$X \subseteq (E_n - M'G)\, A^{-1} (Z + P) + M'(Q - y) = H(M) ,$$
$$X = \bigcap\{H(M) \mid M\} , \rho(l \mid X) = \inf \{\rho(l \mid H(M)) \mid M\}$$

where

$$M \in \mathbf{M}^{m \times n} .$$

The latter relations may be used for recurrent procedures. These are either

$$X[k] = \bigcap\{H_k(L, M) \mid L A(k) + MG(k) = E_n\}, \tag{22.7}$$
$$H_k(L, M) = L'(X[k+1] + P(k)) + M'(y(k) - Q(k), X[t] = Y[t],$$
$$s \le k \le t$$

with

$$X(s \mid s, t, Y[t]) = X[s] \tag{22.8}$$

or

$$\boldsymbol{X}[k] \subseteq H_k(L(k), M(k)), \boldsymbol{X}[t] = Y[t] \tag{22.9}$$
$$s \le k \le t$$

with

$$X(s \mid s, t, Y(t)) = \bigcap \{\boldsymbol{X}[s] \mid L_s(\cdot), M_s(\cdot)\} \tag{22.10}$$

where

$$L_s(\cdot) = (L(s), \ldots, L(t)) ; M_s(\cdot) = (M(s), \ldots, M(t))$$

Theorem 22.2 The set $X(s \mid s, t, Y)$ may be derived due to either equations (22.7) or (22.9), (22.10).

Conclusion

This paper gives an introduction to the theory of *guaranteed identification and state estimation* under uncertainty with unknown but bounded observation "noise" and input disturbances. The whole problem is considered within a deterministic setting so that the results are given in the form of set-valued estimates the description of whose evolution is the objective of the solution schemes.

The mathematical techniques applied here are mainly those of convex analysis, set theory and related topics [21, 27]. A respective continuous version of the given problems would thus further lead us to the techniques of differential inclusions and viability theory of nonlinear analysis, [28, 29].

An important issue is that the purely deterministic solutions to the guaranteed filtering problems may be well approximated by solutions to related problems of stochastic filtering as shown in §§ 14, 15. (This idea also applies to the identification problems of §§ 7, 8.) Thus the well-developed computational techniques of stochastic estimation theory may be modified through some procedures of parallel computations to solve the problems of the above. Basically this gives a robust procedure for solving the specific class of ***inverse problems*** discussed in this paper (see also [30, 31]).

One may raise the question of what is more adequate in the analysis of systems – a stochastic or a "set-membership", deterministic description of uncertainty? The author's opinion is that this question is not correct – the specific informational assumptions for a given problem may require either of these approaches and techniques or perhaps a combination and interaction of both, [32–34]. It is the specific modelling problem that should dominate the tools.

The author wishes to thank F. Stettinger for the computer simulations.

References

[1] Cramer, H.: *Mathematical Methods of Statistics.* Princeton, 1946.

[2] Cox, D.R. and D.V. Hinkley: *Theoretical Statistics.* Chapman and Hall, London, 1974.

[3] Kalman, R.E.: A new approach to linear filtering and prediction problems, *Trans. ASME. Ser D* **82**(1960).

[4] Eykhoff, P.: *System Identification: Parameter and State Estimation*, J. Wiley & Sons, New York, 1974.

[5] Krasovski, N.N.: *The Control of a Dynamic System.* Nauka, Moscow, 1986.

[6] Leitmann, G., C.S. Lee and Y.H. Chen: Decentralized Control for an Uncertain Multi-Reach River System. Springer Lecture Notes on Control and Information Sciences. Optimal Control, **95**(1987).

[7] Chernousko, F.L.: Ellipsoidal Bounds for Sets of Attainability and Uncertainty in Control. *Optimal Control, Appl. & Methods* **3**(1982).

[8] Ackerman, J. and S. Turk: A common controller for a family of plant models. 21st IEEE Conf. on Decision & Control, 1982.

[9] Byrnes, C. and A. Isidori: Heuristics in Nonlinear Control and Modelling. Adaptive Control. Eds. C.I. Byrnes, A.B. Kurzhanski. Springer Lecture Notes in Control and Information Sciences **105**(1988).

[10] Usoro, P.B., F.C. Schweppe, D.N. Wormley and L.A. Gould: Ellipsoidal Set-Theoretic Control Synthesis. *Journal of Dynamic Systems, Measurement & Control*

104(1982).

[11] Krasovskii, N.N.: On the theory of controllability and observability of linear dynamic systems. *Prikl. Mat. Mech.* **28**, 1 (1964). (In Russian.)

[12] Witsenhausen, H.S.: Sets of possible states of linear systems given perturbed observations. *IEEE Trans. Automat. Control* **AC-3**(1968).

[13] Kurzanskii, A.B.: On the duality of the problems of control and observation, *Prikl. Mat. Mech.* **34**, 3 (1970). (In Russian.)

[14] Bertsekas, D. and J. Rhodes: Recursive state estimation for a setmembership description of uncertainty. *IEEE Trans. Automat. Control* **AC-16**, 2 (1971).

[15] Schweppe, F.C.: *Uncertain Dynamic Systems.* Prentice Hall, 1973.

[16] Kurzanskii, A.B.: *Control and Observation Under Conditions of Uncertainty.* Nauka, Moscow, 1977. (In Russian.)

[17] Pschenichnyi, B.N. and V.G. Pokotilo: The Minmax Approach to the Estimation of the Parameters of Linear Regression. *Izv. Akad. Nauk SSR, Tech. Cybernetics,* **2**(1983) (translated as "Engineering Cybernetics").

[18] Fogel, E.: System Identification via membership set constraints with energy constrained noise. *IEEE Trans. Automat. Control* **AC-24**, 5 (1979).

[19] Kurzhanski. A.B. and A. Yu. Hapalov: On the state estimation problem for distributed systems. Analysis & Optimization of Systems Proc. of the VII^{th} Int. INRIA Conf. Springer Lecture Notes in Control & Information Sciences **83**(1986).

[20] Kurzhanski, A.B. and Filippova: On the description of the set of viable trajectories of a differential inclusion. *Soviet Math. Doklady* (Engl. Transl.) **34**, 1, 1987 (**283**, 1, 1986).

[21] Rockafellar, R.T.: *Convex Analysis*, Princeton University Press, Princeton. 1970.

[22] Shiriayev, A.N.: *Probability.* Nauka, Moscow, 1980 (Engl. Transl.).

[23] Kurzhanski, A.B.: On evolution equations in estimation problems for systems with uncertainty. WP-82-49, 1982 IIASA, Laxenburg, Austria.

[24] Ustyuzhanin, A.M.: On the problem of matrix parameter identification. *Problems of Control & Information Theory* **15**(4), 1986.

[25] Albert, A.: *Regression and the Moor-Penrose Pseudo-inverse*, Academic Press, New York, 1972.

[26] Koscheev, A.S. and A.B. Kurzanskii: Adaptive estimation of multistage systems under uncertainty, *Izv. Akad. Nauk SSR, Tech. Cybernetics*, N2 1983 (translated as "Engineering Cybernetics").

[27] Kuratowski, K.: *Topology*, Vol. 1,2, Academic Press, New York, 1968.

[28] Aubin, J.-P. and A. Cellina: *Differential Inclusions.* Springer-Verlag, 1984.

[29] Aubin, J.-P. and J. Ekeland: *Applied Nonlinear Analysis.* Wiley-Interscience, 1984.

[30] Tikhonov, A.N. and V.A. Arsenin: *Methods of Solving Ill-Posed Problems.* Nauka, Moscow, 1979.

[31] Huber, P.J.: *Robust Statistics.* Wiley, 1981.

[32] Katz, I.J. and A.B. Kurzanskii: Minimax multistage filtering in statistically uncertain situations, *Automation and Remote Control* **11** (1978). (In Russian.)

[33] Polyak, B.T. and Ya. Z. Tsypkin: Robust Identificaion. *Automatica,* **16**, 1, 1980.

[34] Anan'ev, B.I. and A.B. Kurzhanski: The nonlinear filtering problem for a multistage system with statistical uncertainty. Second IFAC Symposium on Stochastic Control. Vilnius, USSR, 1986.