DEVELOPMENT OF A MODEL TO STUDY THE CHARACTERISTICS OF AN AIR TRAFFIC CONTROL RADAR

BY

MD. ABDUL MANNAN

A THESIS SUBMITTED TO THE DEPARTMENT
OF ELECTRICAL AND ELECTRONIC ENGINEERING, BUET, DHAKA,
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE
OF MASTER OF SCIENCE IN ELECTRICAL AND ELECTRONIC ENGINEERING

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
BANGLADESH UNIVERSITY OF ENGINEERING AND TECHNOLOGY, DHAKA

JANUARY, 1995

DEVELOPMENT OF A MODEL TO STUDY THE CHARACTERISTICS OF AN AIR TRAFFIC CONTROL RADAR

A thesis submitted by

MD. ABDUL MANNAN
Roll no 901319F, Registration no 84200,
in partial fulfillment for the degree of
Master of Science in Electrical and Electronic Engineering.
Examination held on January 1, 1995 and approved as to style and contents.

BOARD OF EXAMINERS

Dr. A. B. M. Siddique Hossain
Professor and Head
Department of Electrical and Electronic Engineering
BUET, Dhaka-1000, Bangladesh

Chairman
(Supervisor)

Dr. Syed Fazl-i Rahman
Professor
Department of Electrical and Electronic Engineering
BUET, Dhaka-1000, Bangladesh

Member

Dr. Md. Abdul Matin
Professor and Dean
Faculty of Electrical and Electronic Engineering
BUET, Dhaka-1000, Bangladesh

Member

Dr. Gias Uddin Ahmad
Professor
Department of Physics
BUET, Dhaka-1000, Bangladesh

Member
(External)

DECLARATION

I hereby declare that the work presented in this thesis is done by me under the supervision of Dr. A. B. M. Siddique Hossain, Professor and Head, Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology, Dhaka. It is also declared that neither this thesis nor any part thereof has been submitted elsewhere for the award of any degree or diploma.

Countersigned by the supervisor

Dr. A. B. M. Siddique Hossain

Signature of the candidate

Md. Abdul Mannan

ACKNOWLEDGEMENT

All praises are for almighty Allah.

The author would like to express his sincere gratitude and profound indebtedness to his supervisor Dr. A. B. M. Siddique hossain, Professor and Head, Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology (BUET), Dhaka, for his constant guidance, helpful advice, valuable suggestions, encouragement and endless patience throughout the progress of this work, without which the work could not have been completed. The author would also like to thank him for directing me toward this interesting subject in the first place.

The author gratefully recalls the spontaneous co-operation and encouragement obtained from Md. Sanaul Hoque, Lecturer, Department of Computer Science and Engineering, BUET, Dhaka, during the progress of this work. The author also expresses his gratitude to Dr. Gias Uddin Ahmad, Professor, Department of Physics, BUET, Dhaka, for kindly acting as external examiner.

The author conveys his sincere thanks to Mr. Mohshin Uddin Anwar along with other friends for their support. I am also grateful for the help received from Mrs. Nurun Nahar and express my sincerest and special thanks to her for the all out cooperation.

Finally, I would like to dedicate this work to my most beloved parents, Md. Habez Uddin and Mrs. Halima Begum.

ABSTRACT

The availability of digital computers has stipulated the use of digital signal processing in many diverse fields covering engineering, medicine and economics. The aim of this thesis is to develop a model for detection and tracking purposes of **Air Traffic Control Radar**. Neyman-Pearson criterion has been used for detection problem and **Kalman filtering** has been used for the estimation of random signals to extract the pertinent information.

Kalman and Bucy proposed an extremely powerful recursive state estimation technique, commonly described today as **Kalman filtering**. This thesis is concerned with the application of Kalman filtering technique to radar signal processing. The selection of appropriate states to configure the algorithm for use in radar signal processing is also considered. A new practical approach is presented to aid the evaluation of radar systems.

Various techniques are available for the estimation of random signals in the presence of noise and in doing so the need for solving sets of algebraic equations simultaneously arises. This corresponds to inverting a matrix whose order is that of the number of simultaneous equations involved. For the problem of this category, a convenient technique is one in which previously determined estimates are simply updated as new data come in, rather than solving the problem all over again. The recursive estimation technique (**Kalman filtering**) is exactly such a scheme where simultaneous estimates (filtered or predicted) of a number of signal components by minimizing the mean-square error of each signal component simultaneously are looked for. In radar tracking problem one wants to estimate the range, range rate, bearing angle and bearing rate at each time the radar measurement is available. These signal variables will be arranged in a column to be defined as the signal vector. Actually all practical signal processing problems are multidimensional and involve the collection of several signals together.

The other basic materials that covers the nature of radar, the simple radar equation and the propagation of radar waves and how it is contaminated by the atmosphere and other deleterious effects. The thesis deals with the filtering of noisy data in order to extract the signal from noise in an optimum (minimum mean-square error) sense. Initially the tracking problem is highlighted and existing structures are discussed. The random signal and purely additive noise components are assumed to be statistically independent. To show the performance of the proposed algorithm real data has been used that obtained from the Radars at **Zia International Airport, Dhaka**. Before using the real data, the model was tested by known available data and found to be working well. Graphical representation of complete results are also included.

CONTENTS

LIST OF FIGURES

NOTATIONS

a	Time constant
a(k)	Time-varying coefficient that acts as a weight of the previous estimate x(k)
A	System matrix
A_t	Cross section of reflecting object
A_e	Effective area of the receiving antenna
b(k)	Variable gain factor
B	Bandwidth
B_n	Noise bandwidth
c	Observation or measurement parameter
C	Observation matrix
d	Threshold value of the signal
e	Error between the estimate and actual value
erfx	Error function of x
E	Received energy
E/n_o	Signal-to-noise ratio
$E_i(n)$	Integration efficiency
f	Radar signal frequency
f(k)	Input sampled-data signal
f_p	Pulse repetition frequency
$f(x_1 \| 0)$	Probability density function with signal absent
$f(x_1 \| s)$	Probability density function with signal present
F_n	Noise figure of the receiver
F_s	System noise figure
g(k)	Output sampled-data signal
g_i	Cross correlation between the random variables x and y(j)

G	Antenna gain
G(k)	Kalman Predictor gain
h(i)	Coefficients for each measurement sample y(i) to minimize the mean-square error
k	Number of samples used (observation time instant)
K(k)	Kalman Filter gain
$\mathbf{K}_{11}(k)$	First diagonal element K(k)
ℓ	Spacing between samples
$\ell(x_1)$	Likelihood ratio
m	Number of data samples
M	Maximum acceleration of the aircraft
n	Refractive index of atmosphere
n_B	Number of pulses return from a target at interval T
n_o	White noise spectral density
N	Refractivity
ΔN	Additional noise introduced by the network itself
N_i	Input noise power
N_o	Output noise power
p	Barometric pressure
p_e	Mean-square error
p(k)	Mean-square error for k samples
p(k+1|k)	Mean-square prediction error
P_1	Probability of occurrence of signals
P_e	Overall probability of error
P_{e0}	Probability of mistaken noise for signal
P_{e1}	Probability of missing the signal when it is present because of the noise
P_t	Radar transmitted power
P_r	Radar received power
P_n	False-alarm probability
$\mathbf{P}_1$	Probability that aircraft will proceed at constant radial and angular velocities

$\mathbf{P}_2$	Probability that aircraft acceleration (deceleration) in either of the two orthogonal directions (ρ and θ) is at maximum value M
P_d	Probability of detection
$\mathbf{P}(k)$	Error covariance matrix
$\mathbf{P}(k\|k-1)$	Predicted covariance matrix
$\mathbf{P}(k+1\|k)$	Prediction mean-square error
$\mathbf{P}_{11}(k+1\|k)$	First diagonal element of $\mathbf{P}(k+1\|k)$ (mean-square range prediction error)
$\mathbf{P}_{33}(k+1\|k)$	Third diagonal element of $\mathbf{P}(k+1\|k)$ (mean-square bearing prediction error)
$\mathbf{Q}(k)$	System noise covariance matrix
$R+\rho(k)$	Radar Range at time k
R	Average range
$R_y(k)$	Autocorrelation function
R_{ij}	Data autocorrelation between y(i) and y(j)
R_{max}	Maximum radar range
$\mathbf{R}(k)$	Observation noise covariance matrix
S_i	Input signal power
S_o	Output signal power
S_{min}	Minimum detectable signal
$(S/N)_1$	Signal-to-noise ratio of a single pulse
$(S/N)_n$	Signal-to-noise ratio per pulse when n pulses are integrated
t	Time
T	Scan time
T_0	Temperature in °k
T_e	Effective noise temperature
T_s	System noise temperature
Δt	Time interval for the radio wave to travel to the object and back
u(k)	Acceleration of the aircraft
$u_1(k)$	Change in radial velocity
$u_2(k)$	Change in bearing rate

$v(k)$	Additive white noise
$w(k)$	White noise
x	Constant signal
$x(k)$	Time-varying signal
$\hat{x}(k)$	Estimate of the signal process $x(k)$ at time k
$\mathbf{x}(k)$	Signal vector
x_0	Mean value of random signal x
$y(k)$	Data signals
$\mathbf{y}(k)$	Measurement (data) vector
$y(i)$	Measurement signal samples
λ	Wavelength of radar pulse
σ_x^2	Variance of random signal x
σ_v^2	Variance of additive noise $v(k)$
σ_w^2	Variance of white noise process $w(k)$
$\sigma_\rho^2(k)$	Variance of additive noise for the range measurement
$\sigma_\theta^2(k)$	Variance of additive noise for the angle measurement
$\sigma_\rho(k)$	r.m.s. noise in the range sensor
$\sigma_\theta(k)$	r.m.s. noise in the bearing sensor
$\rho(k)$	Deviation from the average range
$\dot{\rho}(k)$	Radial velocity
$\theta(k)$	Bearing (azimuth) at time k
θ_B	Antenna beamwidth
$\dot{\theta}_s$	Antenna scanning rate (deg/s)
ω	Antenna scan rate (rpm)
$\dot{\theta}(k)$	Bearing rate (angular velocity) at time k
δ_{ij}	Kronecker delta
$\beta(k)$	Predictor gain
γ	Noise-to-signal ratio

INTRODUCTION

General:

Considerable research has been undertaken in the field of optimal estimation theory in relation to target tracking using noisy radar data. Target tracking is the determination of the present (and often future) position and velocity (states) of a moving object from noisy measurements of its present states. This is of interest in both military and civilian applications. Various algorithms, using the **Kalman filter** have been derived and by solving the Kalman filter equations, a general solution can be obtained which minimizes the mean-square error; this method is recursive in nature [1-3,20]. In radar systems, detection procedures involve the comparison of the received signal with a certain threshold. But when the background noise power fluctuates, it is difficult to maintain a constant false-alarm rate with a fixed threshold detection scheme [1,30].

The dynamics of a moving target is usually represented in terms of a state space formulation. The target states are its position, velocity and acceleration. They can be estimated by using state estimation procedures based on measurements provided by the radar. When the exact a priori knowledge of the system model, and the statistics of system noise and measurement error are available, the **Kalman filter** provides the minimum mean-square error estimate [1,2]. The Kalman filter has been applied successfully in a variety of target tracking problems.

The Nature of Radar:

Radar is a contraction of the words **Radio Detection and Ranging**. It was first developed as a detection device to warn of the approach of hostile aircraft and of directing antiaircraft weapons. Presently radar is widely used in military, navigation, satellite service, air traffic control etc. Radar is an electromagnetic system for the detection and location of objects. It operates by transmitting a particular type of wave form and detects the nature of the echo signal. Purpose of the air traffic control system is the safe and efficient operation of aircraft flying in the vicinity of airports, aircraft flying enroute from one airport to another, and the aircraft and vehicles on the ground at airports. In the air traffic control radar systems the basic problems are those of first detecting an aircraft when it appears and the estimating its range, velocities and azimuth, among other parameters [1,7].

Several different radars are employed for the control of air traffic. The Airport Surveillance Radar (ASR) provides the information on the location and movement of all aircraft flying within the vicinity of airports. The Air Route Surveillance Radar (ARSR) detect and monitor aircraft enroute. The ASR and ARSR also provide the air traffic controller with the location of hazardous or uncomfortable weather that aircraft should avoid. The ARSR radars for enroute air traffic control have a nominal range of 350 km and an altitude coverage of about 18 km. ASR radar monitors aircraft within 100 km of the airport at which it is located. Both carry out detection and tracking functions. Once the transmitted pulse is emitted by the radar a sufficient length of time must elapse to allow any echo signals to return and be detected before the next pulse may be transmitted. Therefore, the rate at which the pulses may be transmitted is determined by the longest range at which targets are expected [1,7].

Presently radars are classified from another point of view as primary radar and secondary radar. The original primary radar only revealed its total potential when associated with secondary radar so that now air traffic control systems can rely on the integrity of the information provided by the primary and secondary radars [8,31].

Secondary Radar:

Secondary radar is radio location system which measures time, but which, in contrast to normal radar techniques, instead of using the passive echo reflected from a target, uses an active answering device (called transponder) which is located in the target aircraft. Although a secondary radar system will obviously give a position in terms of range and bearing, it is usually used in conjunction with primary radar. The reason for this is obvious that a secondary radar system requires cooperation, and assumes that a transponder is available.

In comparison with the primary radar system, the introduction of a transponder provides substantial advantages such that

(1) In contrast with the primary radar system where, as the range R, increases the power of the echo signal decreases by a factor of $1/(4\pi)^2R^4$, the power of the transponder reply only decreases by a factor of $1/4\pi R^2$. Consequently, it is possible to work with a much lower transmitter power.

(2) The interrogation and the response can be transmitted at two different frequencies, thus avoiding any undesirable echoes such as ground clutter, permanent echoes (caused by rain clouds and other meteorological phenomena).

(3) The transponder equipment is a receiver and transmitter of coded messages. Thus an exchange of information can be obtained as well as information about location.

The secondary surveillance radar system is so designed that a ground station can monitor a space having a maximum radius of 370 km and a height of some 15 km above the radar horizon. In this space only a few aircraft, relative to the size of the area, and separated at a great distances from one another, will be moving at high speeds. In the radial direction the location of an aircraft must be accurate to within some 10 m and, in azimuth, must be accurate to within a few degrees so that measurements can be correlated with the findings of the primary radar equipment. The aim of a secondary radar is to receive information from aircraft that allows the ground station to locate and recognize them. Contrary to primary radar, secondary radar needs aircraft to participate in the recognition. For this purpose, aircraft is equipped with a device called a **transponder** [8,31].

Simple Radar Tracking System:

The radar continuously emits high frequency bursts (pulses) of electromagnetic energy into space. A portion of the transmitted signal is intercepted by a reflecting object (target) and is reradiated in all directions. The receiving antenna collects the returned energy and delivers it to a receiver, where it is processed to detect the presence of the target and to extract its location and relative velocity.

The fig. I shows the set of ideal transmitted and received pulses together with a typical received pulse. A radar beam is used to determine the range and velocity of an object at a distance x from the transmitter. The information required is the time interval Δt representing the time passed for the radio wave to travel to the object and back. The typical received signal is not of ideal shape due to various disturbances and noise adding in with the return pulse confuses the decision as to exactly when the pulse begins, i.e. arrives at the antenna. So, we measure $\Delta t_1 \neq \Delta t$. The range estimate $x=c\Delta t_1/2$ (c is the velocity of pulse propagation in space) from one measurement can therefore cause large errors. To reduce the error, a periodic sequence of pulses is transmitted every T seconds, as indicated in the fig. I, which produces a sequence of measured values of range x(0), x(1),..., x(k). The radar energy is directed into space by the radar antenna. The antenna beam width in these air traffic control radars is of the order of 1 to 1.5° in azimuth, i.e., an angle measured in a plane tangent to the earth's surface, so that aircraft location in azimuth is provided by the direction of the antenna beam when the aircraft intercepts the radar pulse. The radar antenna normally turns, or scans, at a fixed rate, so that aircraft can be detected anywhere in azimuth [2].

Radar Receiver:

Radar receiver accepts the echo signal reflected from the target via the antenna, the rotating joint and the transmitter-receiver switch. It is designed as a heterodyne receiver, in which the oscillator is automatically controlled by the transmitter frequency. The function of any radar system is to give information as to the behavior of a target environment. The ability of a radar receiver to detect a weak echo signal is limited by the noise energy that occupies the same portion of the frequency spectrum as does the signal energy. The weakest signal the receiver can detect is called the **minimum detectable signal**. The spurious undesired signals always present in signalling systems and their components are usually called noise. Since noise is the chief factor that limits receiver sensitivity, it is necessary to obtain some means of describing it quantitatively. This noise reduces the amount of information that can be transmitted with a given signal power. The noise present in a radar receiving system arises both from internally originating sources and from externally originating sources. This externally originating noise enters the receiving antenna along with the desired signal. The function of radar receiver is to detect that desired echo signals in the presence of these noise, clutter, or interference. To do this function it must separate the wanted

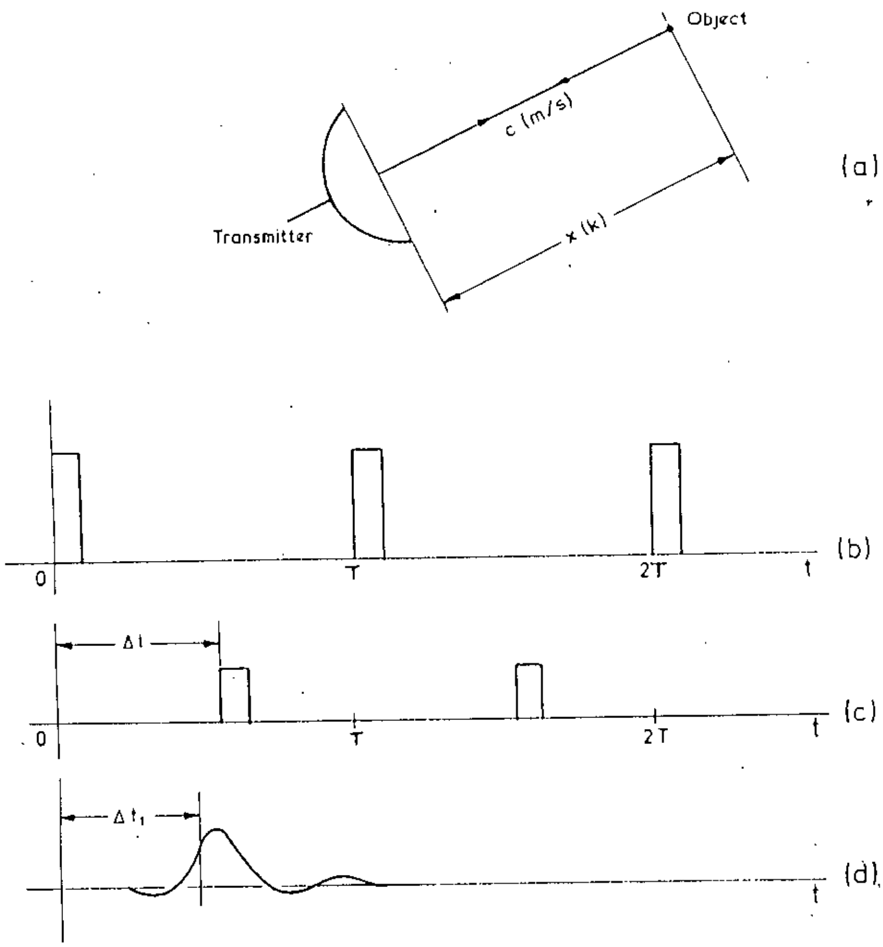

Fig. I (a) Simplified radar tracking system, (b) ideal transmitted pulses, (c) ideal received pulses, (d) typical received pulses from unwanted signals, and amplify the wanted signals to a level where target information can be displayed to an automatic data processor. The block diagram of a typical radar receiver is shown in the Fig. II.

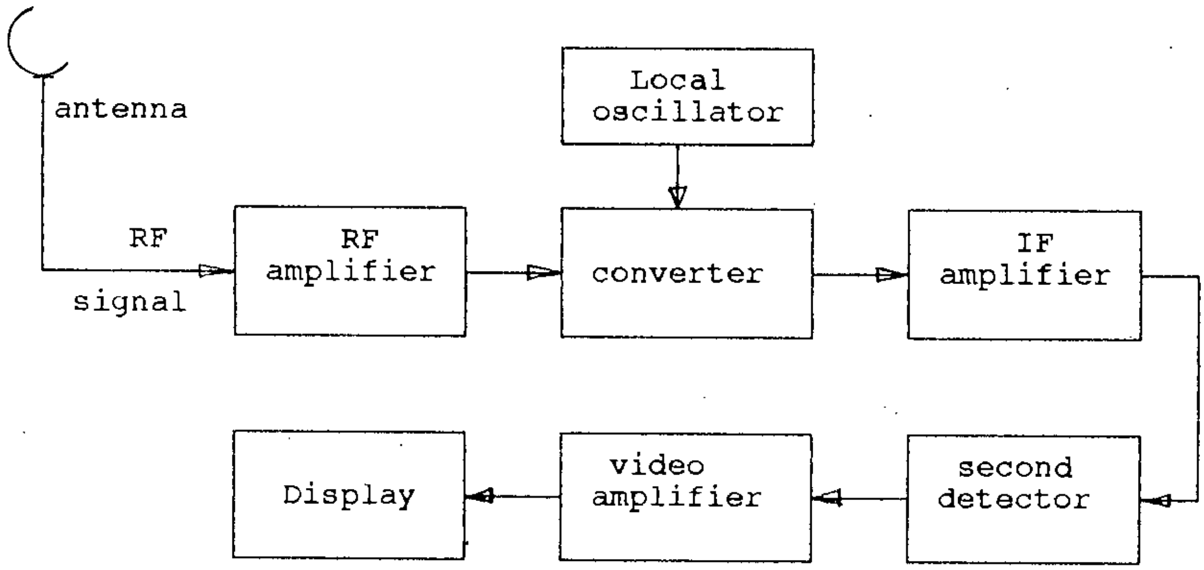

Fig. II Typical radar receiver

At the microwave frequencies usually used for radar, the external noise which enters via the antenna is generally quite low so that the receiver sensitivity is usually set by the internal noise generated within the receiver. The quantitative measure of receiver noise temperature and the quantitative measure of externally originating noise is available in Berkowitz [16] and Skolnik [7].

An unavoidable component of noise generated by the thermal motion of the conduction electrons in the ohmic portions of the receiver input stages. Since it increases with the temperature of the source generating it, it is called thermal noise and is directly proportional to the temperature of the ohmic portions of the receiver circuit and the receiver bandwidth. The available thermal noise power is generated by a receiver of bandwidth B_n (in hz) at a temperature T_o (in °k) is equal to KT_oB_n.

The noise power in practical receivers is often greater than can be accounted for by thermal noise alone. The additional noise components are due to mechanisms other than the thermal agitation of the conduction electrons. No matter whether the noise is generated by a thermal mechanism or by some other mechanism, the total noise at the output of the receiver may be considered to be equal to the thermal noise power obtained from an ideal receiver multiplied by a factor called the **noise figure** which is formulated in the first chapter [7].

There are many other forms of radar signal contamination and distortion that can be identified with the known physical systems [17]. These include:

(1) slow phase path variations caused by travelling ionospheric disturbances;
(2) discrete multimode propagation;
(3) high angle rays, with strongly range-dependent Doppler shift;
(4) impulsive noise from distant thunderstorms;
(5) manmade radio frequency interference;
(6) echoes from meteors, auroras;
(7) ground clutter received through side lobes, back lobes.

Detector:

Detector characteristics also play an important role on radar performance. An elaborate description of detector characteristics are available in Skolnik [7]. The portion of the radar receiver which extracts the modulation from the carrier is called the detector. It includes that portion of the radar receiver from the output of the IF amplifier to the input of the indicator or data processor. One form of detector is the envelope detector, which recognizes the presence of the signal on the basis of the amplitude of the carrier envelope where all phase information is destroyed. Another detector called zero crossing detector which counts the zero crossings of the received wave form. The coherent detector is an example of one which uses both phase and amplitude. The coherent detector does not destroy phase information as does the envelope detector, nor does it destroy amplitude as does the zero crossings detector. Since it utilizes more information than either the envelope detector or the zero crossings detector, it is not surprising that the signal-to-noise ratio from the coherent detector is better than from the other two.

In practice, target fluctuations introduce random phase variations in the high frequency signal received back from the target. The high frequency pulses cannot be simply heterodyned down to the baseband signals and then added. This is the problem always encountered in adding pulses of the same frequency but different phase. Sometimes the resultant sum adds up, and sometimes it decreases, because of phase cancellations. The lack of phase coherence pulse to pulse means that the individual pulses must be envelope detected before processing. Therefore, an integration loss is a common phenomenon in the detection processes [30].

Importance of the Study:

Target tracking is an important problem with wide applications in both military and nonmilitary areas such as fire control systems, satellite orbit determination, maritime surveillance and air traffic control [29]. The need for radar systems evaluation under real conditions has grown because of the great increase of radar systems used for air traffic control over the last few years [14]. One of the most interesting features is the radar tracking by prediction-correction fashion and this thesis is devoted to the random signal processing of radar tracking problem using Kalman filter which is a technique like this. The main advantage of Kalman filter is that it offers the best estimate of random signals for simultaneous estimation of vector (multidimensional) signals. Random signal correction is a signal processing technique to improve signal quality by removing contamination and distortion. In the processing of signals by using Kalman filter no matrix inversion is required and many calculations can be done offline [3-6,11,20].

Air traffic control consists mainly of detecting aircraft, identifying them, and estimating their location and speed. This task is made more and more difficult because of greater aerial traffic. The range of signal processing algorithms for radar remote sensing is constantly being expanded as new techniques are developed and refined [8,10].

Kalman filter or the extended Kalman filter can be applied in a wide varieties of fields such as

(1) For the estimation of attitude angles of an orbiting satellite such as weather satellites, environmental sensing satellites, communication satellites, orbiting astronomical observatories and many other satellites.

(2) For the estimation of the position and velocity of a space vehicle on circumlunar mission.

(3) For the estimation of the number of cars traversing a given section of a highway and their velocities. This information is needed in developing any automatic traffic control system.

(4) In meteorological department for weather forecasting.

We have also presented the feature of secondary radar since the probability that two aircraft fall in the same radar beam of a civil airport is no longer negligible. Secondary radar is very well matched to this task because aircraft play an active role in the sense that they themselves emit the requested information [3,31].

Thesis Organization and Objective:

This thesis comprises eight chapters which is structured as follows:

Chapter one presents the radar equation and the brief description of the sources of noise and interfering signals.

Chapter two focus on the simple signal processing problem, i.e. detection of the presence of a signal in noise. The chapter starts with the discussion on detection criteria and finally present the curves relating the probability of detection and signal-to-noise ratio.

Chapter three begins with a brief discussion of digital filtering theory relevant to estimation theory and presents the concept of an estimator. In this chapter filters are classified as nonrecursive and first-order recursive types for the estimation of signals in noisy data.

The theoretical basis of estimation method is presented in chapter four based on the minimization of mean-square error. Here we motivated a methodology that describes the estimation procedure to find the best possible linear filter.

In chapter five we have investigated the optimum recursive estimation and we get a set of equations referred to as the scalar Kalman filter.

Chapter six actually has sought to provide an organized mathematical technique for radar data processing algorithms. The first-order Kalman filter equations specified in chapter five are modified in this chapter to take higher-order filters into account. Both filtering and prediction problems have been introduced in this chapter.

In chapter seven we apply the method of vector signal processing specifically for radar tracking problem. A computer program has been developed based on the Kalman filtering and prediction equations for radar tracking. Results of computer calculations have been introduced in this chapter.

Finally, in chapter eight we make conclusion of our findings and recommend some issues for further research in this direction.

This thesis is an attempt to process the random signals in an optimum sense, i.e. from a given set of data samples how to determine whether a signal of known characteristics is present or not. The estimates that minimize the mean-square error are taken as the 'best' or optimum estimates.

The problem of detecting the presence or absence of signal in the continuous presence of noise is often encountered in the broad area of signal processing. In detection problems occurring in practice several types of signals may be present simultaneously. The problem is then is to see whether the signal of interest is present in this group. The other signals are then labeled interfering signals and may be lumped in with the ever present noise.

There are two specific functions of radar, aircraft detection and tracking. Once the presence of an aircraft is detected, tracking begins. These two functions of aircraft detection and estimation of its characteristics, for tracking purposes, involved signal processing problems because of the unavoidable presence of noise, as well as other possible deleterious effects. This may be useful in tracking the target to ensure that it is following an appropriate sky path, to prevent air collisions, or to help in the approach to an airport. Kalman filter has applied for this purpose and a computer model has been developed based on the equations that constitute the vector Kalman filter and vector Kalman predictor for radar signal processing. A set of curves have been produced from computer calculations to show the validity the model.

CHAPTER 1

THE SIMPLE FORM OF RADAR EQUATION

1.0 Introduction

A simplified form of the radar equation has been derived in this chapter. In section 1.2, we discussed generally the atmospheric effect on the propagation of radar waves. The noise introduced by the receiver can be found by calculating the **noise figure** of it which is also discussed. Integration of radar pulses which is a technique to improve the performance of radar system is presented here. The other feature presented in this chapter is the brief description of the sources of noise, interfering signals and limitations in detection problem.

1.1 The Simple Form of Radar Equation

The radar equation relates the range of a radar to the characteristics of the transmitter, receiver, antenna, target, and environment. If the power of the radar transmitter is denoted by P_t, and an isotropic antenna is used (one which radiates uniformly in all directions), the power density (w/m^2) at a distance R m from the radar would be $P_t/4\pi R^2$. The radar antenna serves to focus the beam in a desired direction, however, this focussing effect being given by its gain G over the uniform (or isotropic) power density distribution in space. The power density at the target from an antenna with a transmitting gain G is thus $P_tG/4\pi R^2$ [1,7].

The target intercepts a portion of the incident power and reradiates it in all directions. An ideal reflecting object of cross section A_t, located R m from the transmitter will reflect back to the transmitter $P_tGA_t/4\pi R^2$ w. This power in turn spreads out in space as it moves back to the radar antenna. The reflected power density in w/m^2 back at the antenna is thus the power reflected from the target divided by the surface area $4\pi R^2$ encompassed, or just $(P_tGA_t/4\pi R^2)/4\pi R^2$. If the effective area of the receiving antenna is denoted by A_e, the power received by the radar is

$$P_r = \frac{P_t G A_t}{4\pi R^2} \frac{A_c}{4\pi R^2} = \frac{P_t G A_t A_c}{(4\pi)^2 R^4} \quad \text{..................(1.1)}$$

The maximum radar range R_{max} is the distance beyond which the target cannot be detected. It occurs when the received echo signal power P_r just equals the minimum detectable signal S_{min}. Therefore

$$R_{max} = \left[\frac{P_t G A_t A_c}{(4\pi)^2 S_{min}}\right]^{\frac{1}{4}} \quad \text{......................(1.2)}$$

This is the fundamental form of the **radar equation**. Ideally, for an antenna of aperture area A_c m^2, G is given by

$$G = \frac{4\pi A_c}{\lambda^2} \quad \text{.................(1.3)}$$

with λ the wavelength of the high frequency energy.
Since radars generally use the same antenna for both transmission and reception, we have

$$R_{max} = \left[\frac{P_t G^2 \lambda^2 A_c}{(4\pi)^3 S_{min}}\right]^{\frac{1}{4}} \quad \text{.......................(1.4)}$$

The radar equation states that if long ranges are desired, the transmitted power must be large, the radiated energy must be concentrated into a narrow beam (high transmitting antenna gain), the received echo energy must be collected with large antenna aperture (also synonymous with high gain), and the receiver must be sensitive to weak signals [7].

This simplified form of the radar equation does not adequately describe the performance of a practical radar. Many important factors that affect range are not explicitly included. In practice, the observed maximum radar ranges are usually much smaller than what would be predicted by the above equation. There are many reasons for the failure of the simple radar equation to correlate with actual performance. A complete and detailed discussion of all the factors that influence the performance of a radar to predict range and bearing is not the subject matter of this thesis. However, the sources of errors and limitations in detection and tracking of a practical radar system to be treated only lightly. More detailed information will be found in the references listed at the end of this thesis [7,16].

1.2 Propagation of Radar Waves and Atmospheric Effects on it

The radar is an electromagnetic system that operates by transmitting a particular type of electromagnetic waves. The propagation of radar waves is affected by the earth's surface and its atmosphere. The earth is almost a perfect absorber at microwave frequencies and can be treated as a black body at a physical temperature and noise from the earth will usually enter via the antenna. Moreover, scattering of electromagnetic energy from the surface of the earth, refraction caused by an inhomogeneous atmosphere, and attenuation by the gases constituting the atmosphere affect on the propagation of radar waves. The regions of the atmosphere which affect the propagation of electromagnetic waves are the troposphere and the ionosphere. The absorption of radar waves in the lower atmosphere is the result of the presence of both free molecules and suspended particles such as dust grains and water drops condensed in fog and rain. In a noncondensed atmosphere, oxygen and water vapor are the substances which cause absorption. This kind of absorption is called tropospheric attenuation. In the upper atmosphere, electron collisions occur due to the presence of neutral particles and heavy ions. The ultimate effect of collisions is to cause absorption of energy from the electromagnetic waves traversing the ionized medium. This type of absorption of signal energy is called ionospheric attenuation. Moreover, because of the nonisotropic characteristics of the troposphere and ionosphere, radar waves on their passage through the atmosphere experience an angular deviation. The radar is also affected by the reflection, or back scatter, of energy from the earth's surface and from rain, snow, birds and other clutter objects. The radar waves propagating within the atmosphere do not travel in straight lines but are generally bent or refracted. Bending, or refraction, of radar waves in the atmosphere is caused by the variation of the index of refraction. One effect of refraction is to extend the distance to the horizon, thus increasing the radar coverage, fig.(1.1a). Another effect is the introduction of errors in the management of elevation angle, fig.(1.1 b). In the troposphere, the index of refraction, which is a function of meteorological variables, such as water vapor, air temperature, and air pressure, can be represented by

$$(n-1)x10^6 = N = \frac{77.6p}{T} + \frac{3.73x10^5 e}{T^2} \qquad \text{.............................(1.5)}$$

where n = refractive index
T = air temperature (°k)
p = barometric pressure (mbar)
e = Partial pressure of water vapor (mbar)

The parameter $N=(n-1)x10^6$ is the "Scaled up" index of refraction and is called refractivity.

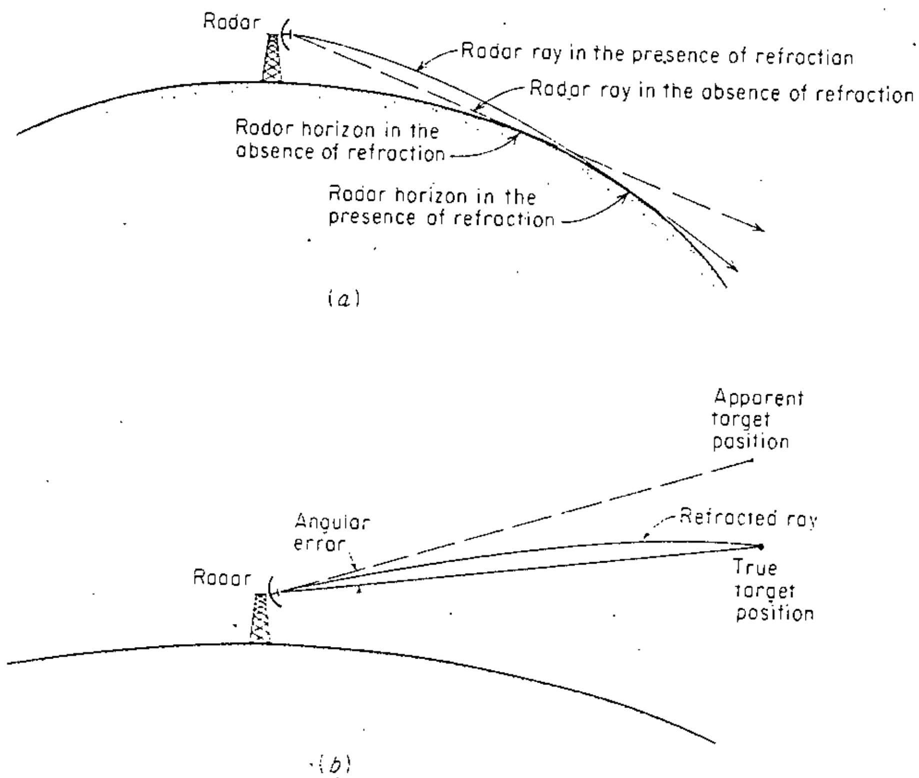

Fig. 1.1 (a) Extension of the radar horizon due to refraction of radar waves by the atmosphere, (b) angular error caused by refraction

The prime difference between optical and microwave refraction is that water vapor has negligible effect on the former; consequently the second term of eqn.(1.5) may be neglected at optical frequencies. Since the barometric pressure and the water vapor content decrease rapidly with height, while the temperature T decreases slowly with height, the index of refraction normally decreases with increasing altitude. A typical value of the index of refraction near the earth surface is 1.0003. In a standard atmosphere the index decreases at the rate of about $4x10^{-8}$ m of altitude. The decrease in refractive index with altitude means that the velocity of propagation increases with altitude, causing radar waves to bend downward. The result is on increase in the effective radar range as was illustrated in fig.(1.1a) [7,16].

Another mechanism that permits radar coverage to be extended beyond the geometrical horizon is diffraction. Radar waves are diffracted around the curved earth in same manner that light is diffracted by a straight edge.

These are the causes for which time delays or range errors are always inherent in radar target measurement. Again, when the antenna looks into space and searching for aircraft, it continuously picks up electromagnetic energy from the earth's atmosphere, radiation from the earth itself and from the sun if the antenna beam intercepts the sun's range, and the noise like electromagnetic radiation which arrives from the extraterrestrial sources as our own galaxy, extragalactic sources, and radio stars which is known as cosmic noise. In general, this cosmic noise level will vary directly as the square of wavelength [1,7,16].

The chief effect of weather on radar performance is the back scatter or clutter, from precipitation within the radar resolution cell. In general radars at the lower frequencies are not bothered by meteorological or weather effects, but at the higher frequencies, weather echoes may be quite strong and mask the desired target signals just as any other unwanted clutter signal. The other important factors that affect the radar performance and those are mentioned in this section can be found with mathematical formulation and graphical representation in the references.

1.3 Noise Figure

The noise figure of a receiver is a measure of the noise produced by a practical receiver as compared with the noise of an ideal receiver. The noise figure F_n of a linear network may be defined as

$$F_n = \frac{S_i/N_i}{S_o/N_o} = \frac{N_o}{KT_oB_nG} \qquad \text{.......................(1.6)}$$

where S_i = input signal power
N_i = input noise power (equal to KT_oB_n)
S_o = output signal power
N_o = output noise power
K = Boltzman's constant = 1.38×10^{-23} j/°K
G = gain = S_o/S_i
T_o = standard temperature
B_n = noise bandwidth

The noise figure may also be written as

$$F_n = \frac{KT_oB_nG + \Delta N}{KT_oB_nG} = 1 + \frac{\Delta N}{KT_oB_nG} \qquad \text{.......................(1.7)}$$

where ΔN is the additional noise introduced by the network itself. Therefore, the noise figure F_n is essentially a measure of how much additional noise the receiving system introduces over and above the thermal noise picked up by the antenna, i.e. the noise figure may be interpreted as a measure of the degradation of signal-to-noise ratio as the signal passes through the receiver.

The effective noise temperature, T_e is defined as

$$T_e = (F_n - 1)T_o \qquad \text{.......................(1.8)}$$

and the system noise temperature, T_s is defined as the effective noise temperature of the receiver system including the effects of antenna temperature, T_a such that

$$T_s = T_e + T_a \qquad \text{................(1.9)}$$

or, $$T_s = T_oF_s \qquad \text{................(1.10)}$$

where F_s is the system noise figure including the effects of antenna temperature. The effective noise temperature of a receiver consisting of a number of networks in cascade is

$$T_c = T_1 + \frac{T_2}{G_1} + \frac{T_3}{G_1 G_2} + \ldots\ldots \qquad \ldots\ldots\ldots\ldots(1.11)$$

where T_i and G_i are the effective noise temperature and gain of the ith network.

The effective noise temperature and the noise figure both describe the same characteristics of a network. In general, the effective noise temperature has been preferred for describing low noise devices, and the noise figure is preferred for the conventional receivers. For radar receivers the noise figure is the more widely used term. We may now introduce this into the simple radar equation that derived in section 1.1. Eqn.(1.6) can be written as

$$S_i = \frac{KT_0 B_n F_n S_0}{N_0} \qquad \ldots\ldots\ldots\ldots(1.12)$$

If the minimum detectable signal s_{min} is that value of S_i corresponding to the minimum ratio of output signal-to-noise ratio $(S_o/N_o)_{min}$ necessary for detection, then

$$S_{min} = kT_0 B_n F_n \left(\frac{S_0}{N_0}\right)_{min} \qquad \ldots\ldots\ldots\ldots(1.13)$$

Substituting eqn.(1.13) into the simple form of the radar eqn.(1.2), derived in section 1.1, we have

$$R_{max}^4 = \frac{P_t G A_e A_t}{(4\pi)^2 KT_0 B_n F_n (S_0/N_0)_{min}} \qquad \ldots\ldots\ldots\ldots(1.14)$$

Therefore, the receiver should be designed to generate as little internal noise as possible, especially, in the input stages where the desired signals are the weakest. There are some other losses that occur throughout the radar system and reduce the signal-to-noise ratio at the receiver output. The antenna beam-shape loss, collapsing loss, and losses in the microwave plumbing are example of losses which cannot be ignored in any serious prediction of radar performance [16,7].

1.4 Integration of Radar Pulses

For the signal pulses radiated by the radar are of finite energy and must themselves compete with the noise after reflection by the aircraft. The radar beam spreads steadily as it propagates into space, providing an unavoidable $1/(\text{range})^2$ decrease in power in each direction, or $1/(\text{range})^4$ decrease total. The signal power returned depends on the reflecting properties of the aircraft as well as its distance, and sometime can be of the order of the noise or even less. In practice, to reduce the error, the probability of detecting a signal at one particular range and azimuth setting is enhanced by combining several successive signal-pulse returns from the same target. This is possible in a single radar scan where each location in space is illuminated by several successive pulses. Actually the radar beam in rotating through space covers a typical target long enough to have several pulses returned from it [7,16].

If the n number of pulses returned from a point target as the radar antenna scans through its beamwidth, then

$$n_B = \frac{\theta_B f_P}{\dot{\theta}_s} = \frac{\theta_B f_P}{6\omega_m} \qquad \text{.........................(1.15)}$$

where θ_B = antenna beamwidth, deg.

f_p = pulse repetition frequency, hz.

$\dot{\theta}_s$ = antenna scanning rate, deg/s

ω_m = antenna scan rate, rpm

The pulse repetition frequency for primary radar 633 hz, beamwidth 1.5° and antenna scan rate 15 rpm. These parameters result in 10.55 hits from a point target on each scan. The process of summing all the radar echo pulses for the purpose of improvement in detection is called integration. Many techniques might be employed for accomplishing integration. All practical integration techniques employ some sort of storage device.

Integration may be accomplished in the radar receiver either before the second detector (in the IF) or after the second detector (in the video). Integration before the detector is called predetection or coherent integration, while integration after the detector is called postdetection or incoherent integration.

The pulse repetition frequency (prf) is determined primarily by the maximum range at which targets are expected. If the prf is made too high, the likelihood of obtaining target echoes from the wrong pulse transmission is increased. Echo signals received after an interval exceeding the pulse repetition period can result in confusing range measurements. A high prf is desired to provide maximum average power on the target.

Predetection integration requires that the phase of the echo signal should be preserved if full benefit is to be obtained from the summing process. On the other hand phase information is destroyed by the second detector in postdetection. For this reason, postdetection integration is not as efficient as predetection integration.

If n pulses, all of the same signal-to-noise ratio, were integrated by an ideal predetection integrator, the resultant or integrated signal-to-noise (power) ratio would be exactly n times that of a signal pulse. If the same n pulses were integrated by an ideal postdetection device, the resultant signal-to-noise ratio would be less than n times that of a single pulse. This loss in integration efficiency is caused by the nonlinear action of the second detector, which converts some of the signal energy to noise energy in the rectification process. Although postdetection integration is not as efficient as predetection integration, it is easier to implement in most applications. Postdetection integration is therefore preferred, even though the integrated signal to noise ratio may not be as great. The integration efficiency may be defined as follows:

$$E_i(n) = \frac{(S/N)_1}{n(S/N)_n} \qquad \text{.....................(1.16)}$$

where n = number of pulses integrated of equal amplitude

$(S/N)_1$ = value of signal-to-noise ratio of a single pulse required to produce given probability of detection (for n =1)

$(S/N)_n$ = value of signal-to-noise ratio per pulse required to produce same probability of detection when n pulses are integrated.

The improvement in the signal-to-noise ratio when n pulses are integrated in postdetection is $nE_i(n)$ and is the integration improvement factor. The improvement with ideal predetection would be equal to n.

In practice, an integration loss will be encountered that can be estimated as a function of n and the type of integrator used. Integration loss in decibels is defined as $L_i(n)=10\log[1/E_i(n)]$. The integration improvement factor (or the integration loss) is not a sensitive function of either the probability of detection or the probability of false-alarm. Details about integration type and loss due to this are available in Schwartz [1] and Skolnik [7].

CHAPTER 2

DETECTION OF SIGNALS IN NOISE

2.0 Introduction

In first chapter we discussed briefly the nature of radar associated with the air traffic control system. In this chapter we focus on the detection of the presence of desired signal in the continuous presence of noise, one of the two signal processing tasks. In section 2.1, detection theory is presented where the signal is considered to be constant amplitude and the noise is assumed to be gaussian and independent of the signal. In section 2.2, the Neyman-Pearson criterion is presented and the results obtained were used to assess the performance of the air traffic control radars at **Zia International Airport**.

2.1 Detection Theory

The detection of weak signals in the presence of noise is equivalent to deciding whether the receiver output is due to noise alone or to signal-plus-noise. Assuming the signal of constant amplitude A whose presence we are interested in detecting in the continuous presence of noise v(k), independent of the signal. The composite received waveshape that we must process to determine the presence of signal is then

$$x(k)=A+v(k) \qquad \text{..................(2.1)}$$

$$\text{or,} \quad x(k)=v(k)$$

The signal when it does appear results in a change in the dc level or bias of x(k) as shown in fig.(2.1). The signal is shown raising the level of x(k) by A units during the T second it is present. Assume that we continually sample x(k) m times in T second interval as shown in fig.(2.2a).

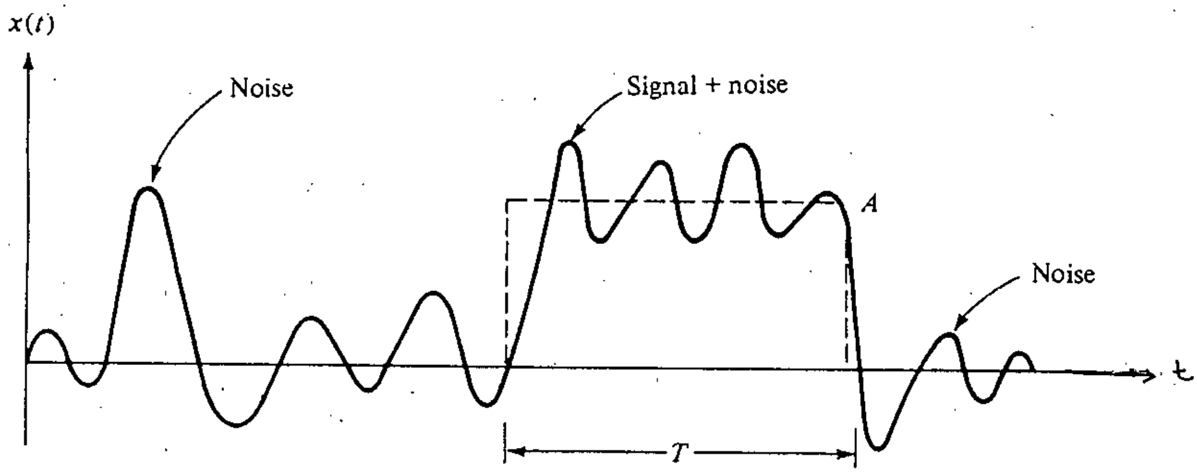

Fig. 2.1 Received analog waveshape

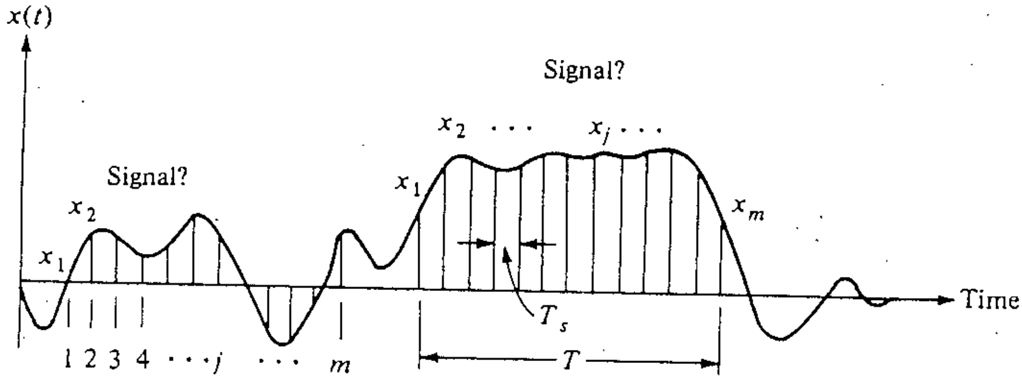

Fig. 2.2a Sampled form of received signal

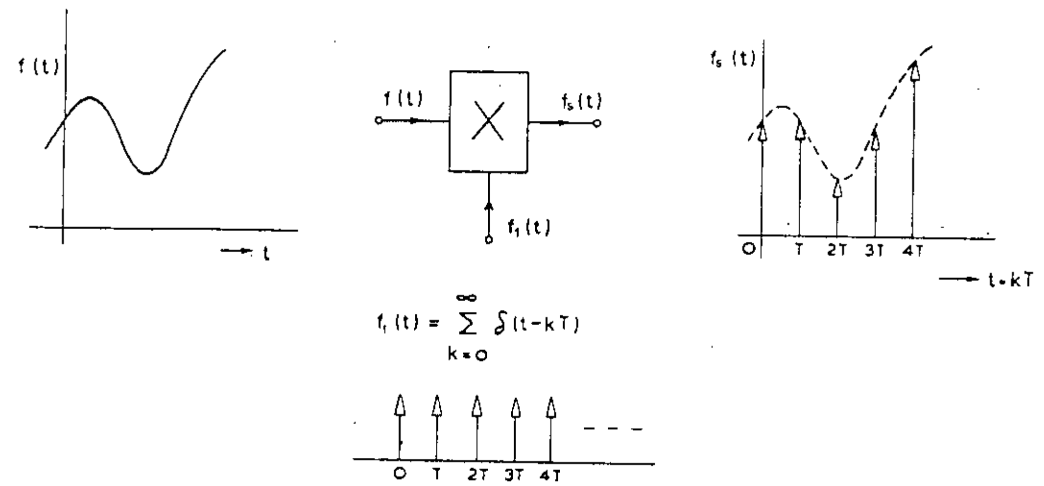

Fig. 2.2b Modulation of a signal by a sequence of impulses

In communications and control systems, the discrete-time sequences x(k) and y(k), or in the general notation f(k), are samples of a continuous-time waveform. These can be interpreted as a form of modulation of the signal f(t) by a sequence of impulses as in fig.(2.2b). A signal f(t) is sampled every T seconds output wave will be

$$f_s(t) = f(t)f_1(t)$$

and can be represented as

$$f_s(t) = f(t)\sum_{k=0}^{\infty} \delta(t-kT)$$

or,

$$f_s(t) = \sum_{k=0}^{\infty} f(kT)\ \delta(t-kT)$$

where the subscript s denotes the sampled signal. We are now interested in processing the m samples x_j, j=1,2,3,...,m, in any T second interval to determine the presence or absence of signal. We assume that the noise before entering the system is white or at least that its bandwidth is much higher than that of the signal with which we have to deal. We also assume that there is some bandwidth B sufficient to pass the signal essentially undistorted [13]. In most detection systems x(k) is filtered to bandwidth B prior to the sampling process to reduce the effect of the noise. For white noise, however, the mean noise power is proportional to the bandwidth, and so restricting the noise bandwidth as much as possible can only help in the detection process. We shall assume that v(k) is bandlimited white noise of bandwidth B.

We know from Nyquist sampling theorem that with the signal f(t) band-limited B hertz, it is then readily shown that sampling the signal does not destroy any information content, provided that the sampling rate $f_c \geq 2B$. The minimum sampling rate of 2B times per second is called the Nyquist sampling rate and 1/2B the Nyquist sampling interval. So the larger the number of uncorrelated samples of noise or signal-plus-noise we can collect, the better our chance of detecting a signal when it appears. Therefore we may select the sampling internal T exactly 1/2B, with a corresponding maximum of m=2BT samples in any T second interval.

In detection problem usually two types of error may occur:

(1) the signal may be missed when it is present because of the noise;

or (2) the noise may be detected mistakenly for signal when the signal is absent (often called a false-alarm).

The overall probability of error is due to both. We denote the probability of first type of error by P_{e1} and the second by P_{e0}. An ideal radar signal detector would be one which maximizes the probability of detecting the signal when it appears and minimizes the probability of mistaking noise for signal when it is absent. Unfortunately, both probabilities cannot be optimized simultaneously. The best we can do to keep the noise probability (or false-alarm probability) at some tolerable level and maximize the corresponding signal probability. This is the performance criterion usually adopted for radar systems. The radar case is an example of a detection problem in which the signal appears relatively infrequently, with no predetermined (a priori) statistics, where as the noise is always present [1,30].

Now both P_{e1} and P_{e0} are conditional probabilities since they are conditioned on the signal's being present and absent, respectively. Since these errors are due to mutually exclusive events, we can find the overall probability of error P_e by adding the two error probabilities after unconditioning them by multiplying by the appropriate probabilities of occurrence of the two events. So,

$$P_e = P_1 P_{e1} + (1-P_1)P_{e0} \qquad \text{.............(2.2)}$$

where P_1 is the (assumed known) probability of occurrence of signals.

It is obvious that both P_{e1} and P_{e0} depend on the respective distributions of x_1 in the two cases of signal present or signal absent, so that the minimization of P_e depends on these statistics as well. Considering a typical case as shown in fig.(2.3), where we denote the probability density function of x_1 when the signal is absent, $f(x_1|0)$ and that when the signal is present, $f(x_1 | s)$. We have

$$f(x_1 | 0) = f_v(x_1) \qquad \text{...............(2.3)}$$

if the noise is gaussian then we can write

$$f_v(v) = \frac{e^{-v^2/2\sigma_v^2}}{\sqrt{2\pi\sigma_v^2}} \qquad \text{......................(2.4)}$$

So,

$$f(x_1|0) = \frac{e^{-x_1^2/2\sigma_v^2}}{\sqrt{2\pi\sigma_v^2}} \qquad \text{......................(2.5)}$$

Similarly, with signal present x_1=A+v, we obtain

$$f(x_1|s) = \frac{e^{-(x_1-A)^2/2\sigma_v^2}}{\sqrt{2\pi\sigma_v^2}} \qquad \text{.......................(2.6)}$$

The processor for one sample x_1, must have a decision rule which specifies two separate ranges of the possible values of x_1, the range of values X_1 corresponding to signal present hypothesis and the range of values X_0 corresponding to the signal absent hypothesis. The probability of error P_e depends on the particular choice of these two joint regions, and it is the object of our analysis to find those ranges which minimize P_e.

If we assume that the two regions are simply chosen as shown in fig.(2.3), a decision level d is picked; all values of $x_1>d$ correspond to signal present (X_1), and all values of $x_1<d$ correspond to signal absent (X_0). In general, the two regions X_1 and X_0 incorporate all points on the line corresponding to possible values of x_1 and may cover several decision regions, as shown in fig.(2.4).

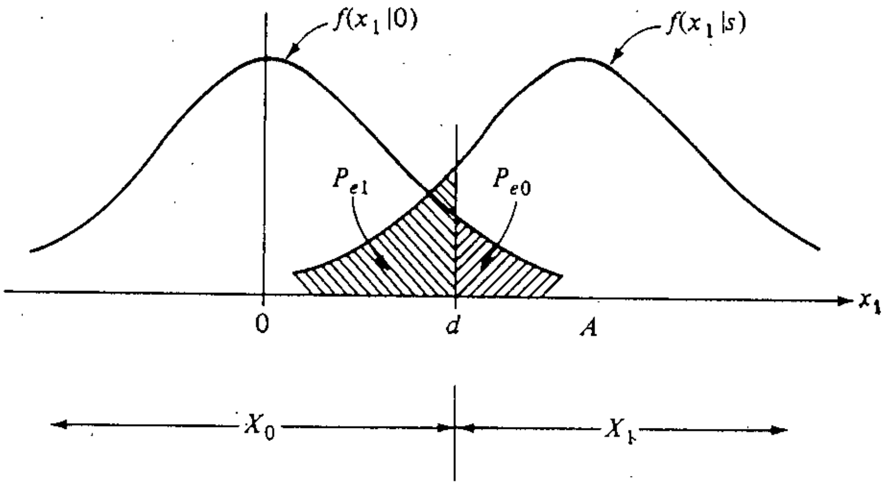

Fig. 2.3 Decisions region

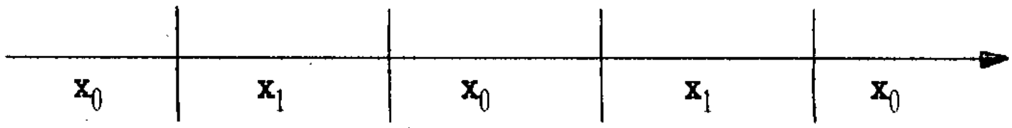

Fig. 2.4 More general decision regions

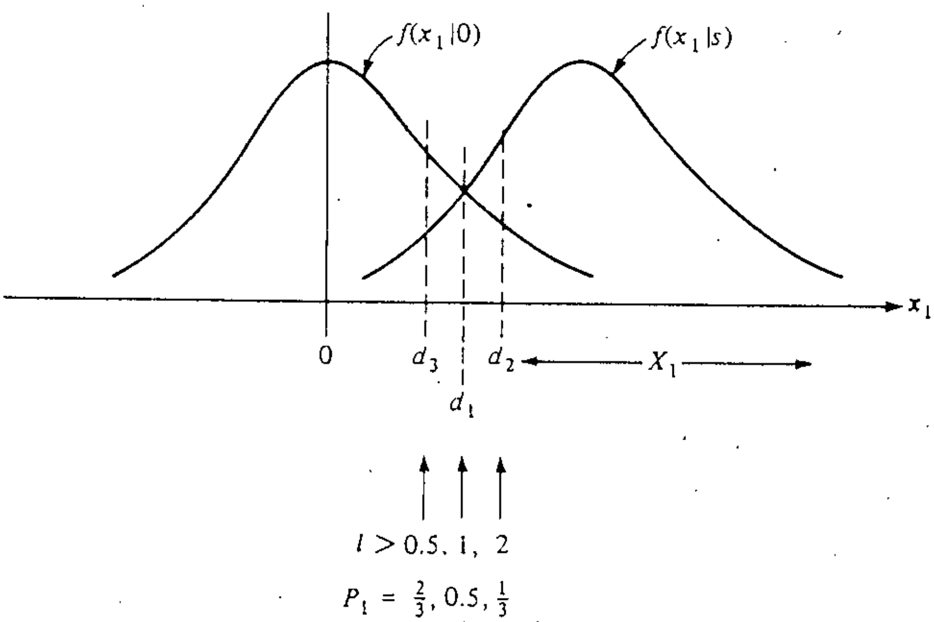

Fig. 2.5 Decision regions for different values of P_1

It is apparent from the definitions that P_{e0} corresponds to the probability that x_1 will fall into region X_1 given that only noise is present while P_{e1} corresponds to the probability that x_1 will fall into region X_0 given that a signal is present. These probabilities in turn are found by integrating the appropriate density functions over the respective regions. Specifically, for the example shown in fig.(2.3), we have

$$P_{co} = \int_{X_1} (f x_1 | o) dx_1 \qquad \text{..........................(2.7)}$$

and $$P_{e1} = \int_{X_0} f(x_1|s)\,dx_1 \qquad \text{.........................(2.8)}$$

From eqn.(2.2), the overall probability of error is given by

$$P_e = P_1\int_{X_0} f(x_1|s)\,dx_1 + (1-P_1)\int_{X_1} f(x_1|0)\,dx_1 \qquad \text{.........................(2.9)}$$

Since X_0 and X_1 cover the entire space of x_1, we can eliminate X_0 using

$$\int_{X_0+X_1} f(x_1|s)\,dx_1 = 1$$

We have

$$P_e = P_1 + \int_{X_1}\left[(1-P_1)f(x_1|0) - P_1 f(x_1|s)\right]dx_1 \qquad \text{.........................(2.10)}$$

Since P_1, $1-P_1$, and both density functions are all of necessity positive, it is obvious that for minimum P_e, we can choose the inequality

$$P_1 f(x_1|s) > (1-P_1)\,f(x_1|0) \qquad \text{.........................(2.11)}$$

Therefore, in this case the integral over X_1 is as negative as one can make it, and hence P_e is as small as possible. The values of x_1 satisfying inequality (2.11) thus correspond to region X_1, while those which reverse the inequality correspond to X_0.

Again, from eqn.(2.11), we can formulate the so-called likelihood ratio as

$$l(x_1) \equiv \frac{f(x_1|s)}{f(x_1|o)} > \frac{1-P_1}{P_1} \qquad \text{.........................(2.12)}$$

If P_1 increases, the X_1 region increases and the X_0 decreases corresponding to an increased likelihood that signal will be detected. Similarly, as P_1 decreases, the X_1 region is reduced in size as shown in fig.(2.5).

Now for m random variables $x_1,x_2,\ldots,x_j,\ldots,x_m$ and with geometric properties of them, we may define an m-dimensional probability density function $f(x_1,x_2,\ldots,x_m)$ integrable over the entire m-dimensional space. Once an m-dimensional region X_1 is chosen, an error will occur with the signal absent if the composite sample group falls into region X_1. The probability of this happening is given by the appropriate integration over X_1 such that

$$P_{e0} = \int_{X_1}\cdots\int\prod_{j=1}^{m}\left[f(x_j|0)\right]dx_1\ldots dx_m \qquad \text{.....................(2.13)}$$

which is just the m-dimensional extension of eqn.(2.7). Proceeding exactly as in the one-dimensional case, it can be shown that the detection procedure appropriate to minimizing the overall probability of error consists of deciding the signal is present if

$$l(x_1,\ldots,x_m) \equiv \frac{\prod_{j=1}^{m} f(x_j|s)}{\prod_{j=1}^{m} f(x_j|0)} > \frac{1-P_1}{P_1} \qquad \text{.........................(2.14)}$$

Taking the natural logarithm of both sides of eqn.(2.14), we have

$$\ln l \equiv \sum_{j=1}^{m}\ln\frac{f(x_j|s)}{f(x_j|0)} > \ln\frac{1-P_1}{P_1} \qquad \text{.......................(2.15)}$$

The detection procedure thus consists of operating on each sample, summing the resultant numbers, and checking to see whether this sum is greater than a specified threshold. Now if the noise is gaussian we can use the eqns.(2.5) and (2.6) to find the ratio of density functions required by eqn.(2.15) and taking the natural logarithm, we get a decision rule which decides that the signal is present when the values of x_j, j=1,2,...,m, satisfy the inequality

$$\sum_{j=1}^{m}\left[x_j^2-(x_j-A)^2\right] > 2\sigma_v^2\ln\frac{1-P_1}{P_1} \qquad \text{....................(2.16)}$$

From which we obtain

$$\sum_{j=1}^{m} x_j > \frac{mA}{2} + \frac{\sigma_v^2}{A}\ln\frac{1-P_1}{P_1} \qquad \text{.......................(2.17)}$$

The optimum processor that we obtain here consists of a summer which simply adds the samples and checks to see whether the sum exceeds a specified level d.

If we divide through by the constant m, we have an alternate optimum processor where we decide a signal is present if

$$y = \frac{1}{m}\sum_{j=1}^{m} x_j > \frac{A}{2} + \left(\frac{\sigma_v^2}{mA}\ln\frac{1-P_1}{P_1}\right) \qquad \text{.......................(2.18)}$$

Hence we can say that for gaussian noise with fixed signal amplitude A the sample mean is the appropriate number to be calculated in detecting the presence of the signal. An example of the overall probability of error P_e, from eqn.(2.2) is shown in Schwartz [1] for the special case $P_{e0}=P_{e1}$ such that

$$P_e = \frac{1}{2}\left(1 - erf\frac{\sqrt{m}A}{2\sqrt{2}\sigma_v}\right) \qquad \text{...................(2.19)}$$

with the error function of x defined as

$$erf\,x \equiv \frac{2}{\sqrt{\pi}}\int_o^x e^{-y^2}\,dy \qquad \text{.......................(2.20)}$$

From which we get the performance curve as shown in fig.(2.6). It is apparent that if we desire a probability of error P_e less than 10^{-4} we require $\sqrt{m}A/\sigma_v > 7.3$. For $\sqrt{m}A/\sigma_v = 10$, $P_e = 2.87\times10^{-6}$, for $\sqrt{m}A/\sigma_v = 11$, $P_e = 1.9\times10^{-7}$ and for $\sqrt{m}A/\sigma_v = 12$, $P_e = 1.0\times10^{-9}$. So small changes in A/σ_v, the so-called signal to noise ratio result in significant changes in the probability of error. This is due to the exponentially decreasing tails in the gaussian distribution. For this reason, increasing the signal-to-noise ratio are considered so significant in signal processing systems and one tries to reduce the noise and increase the signal power received as much as possible. To improve the equivalent signal-to-noise ratio the number of samples m may also be increased. Since the samples are spaced a fixed time interval apart to ensure statistical independence, the time for detection increases with m [1].

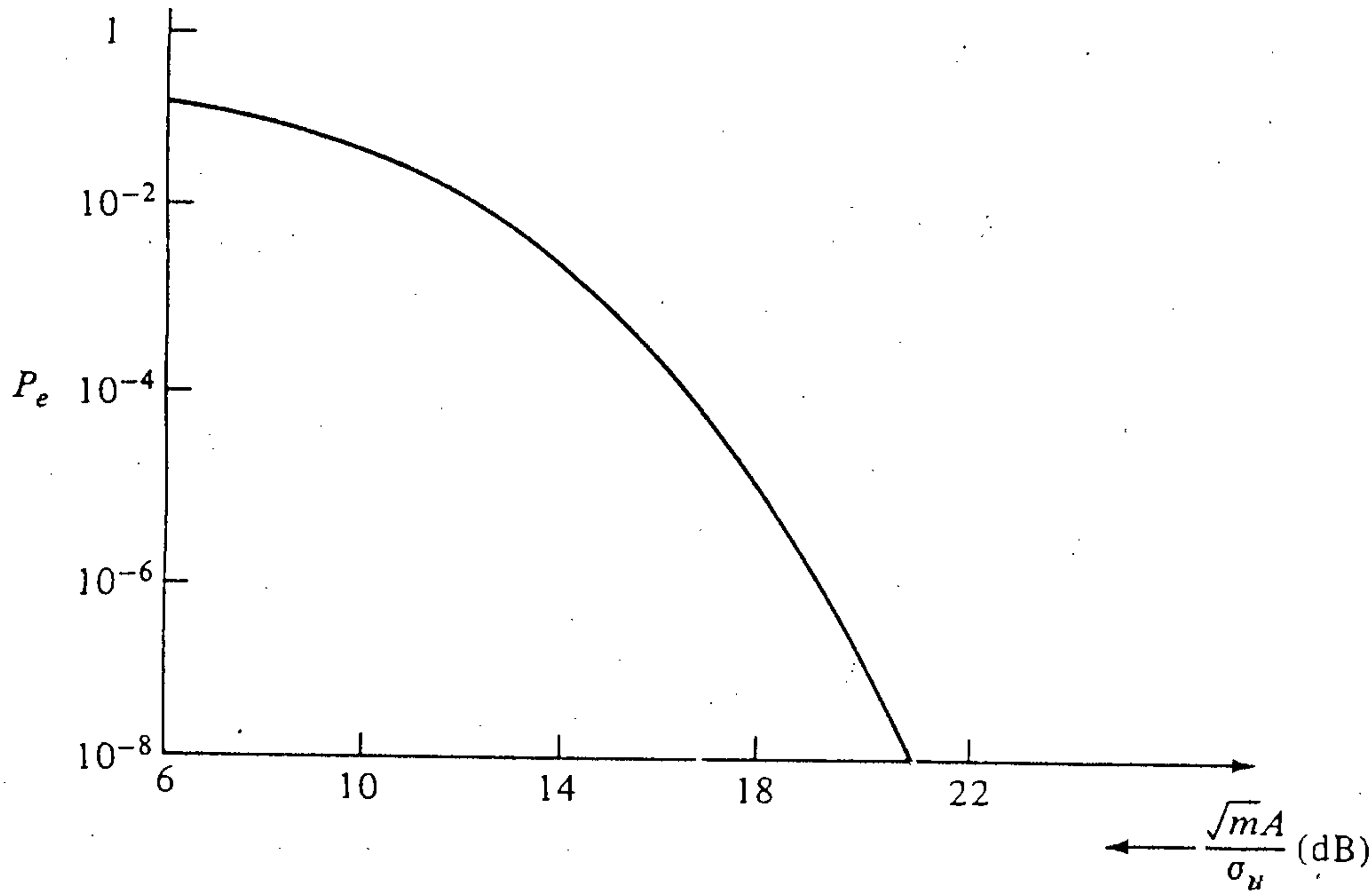

Fig. 2.6 Performance curve of the optimum detector (for $P_1=0.5$)

2.2 Neyman-Pearson Theory for Radar

In the detection problem an appropriate performance criterion is one that maximizes detection probability while keeping the false-alarm probability fixed at some tolerable value. This criterion will be discussed in detail and we shall use the results obtained here to assess the performance of the air traffic control radars (primary and secondary radar) at **Zia International Airport**.

The Neyman-Pearson criterion, after the two statisticians who first explored its properties, minimizes the probability of signal loss, P_{e1} with the probability of mistaking noise for signal, P_{e0}, held fixed at some tolerable level. In radar case, we would like to maximize the probability of detection P_d= 1-P_{e1} with a false-alarm probability P_n= P_{e0} specified. For an m-dimensional space with the m independent samples x_1, x_2,..., x_m, we have, as in eqn.(2.13)

$$1-P_{e1} = P_d = \int_{X_1}\cdots\int \prod_{j=1}^{m} [f(x_j|s)]\, dx_1 \ldots dx_m \qquad \text{.................... (2.21)}$$

and

$$P_{e0} = P_n = \int_{X_1}\cdots\int \prod_{j=1}^{m} [f(x_j|0)]\, dx_1 \ldots dx_m \qquad \text{........................ (2.22)}$$

It is apparent that we cannot simultaneously adjust the signal region X_1 to maximize the detection probability P_d and minimize the false-alarm probability P_n. Therefore, in Neyman-Pearson test P_n is kept fixed at some tolerable level and then search for the region X_1 that maximizes P_d. We consider a signal of amplitude A appearing in a gaussian noise background. Then for one sample x_1 we can show that the signal is declared present if

$$x_1 > d = \frac{A}{2} + \frac{\sigma_v^2}{A}\ln k \qquad \text{..................... (2.23)}$$

where k is a constant. This is shown in fig.(2.7). Upon normalizing, and using the error function definition, for gaussian noise we have

$$P_n = \int_d^{\infty} \frac{e^{-x_1^2/2\sigma_v^2}}{\sqrt{2\pi\sigma_v^2}} dx_1 = \int_{d\sqrt{2}\sigma_v}^{\infty} \frac{e^{-x^2}}{\sqrt{\pi}} dx = \frac{1}{2}(1 - erf\frac{d}{\sqrt{2}\sigma_v}) \qquad ....(2.24)$$

Eqn.(2.24) shows the relation between the threshold d (or, equivalently, the constant k) to the false-alarm probability P_n. It is apparent from eqn.(2.24) that reducing P_n increases d, and vice versa. Moreover, the threshold level d is independent of the signal amplitude A and it depends solely on the noise standard deviation. But eqn.(2.23) indicates the dependency of d on A. The constant k obviously must adjust itself to make d independent of A. In the case of m independent samples of a signal of amplitude A appearing in a gaussian noise background, the signal is declared present if

$$y = \sum_{j=1}^{m} x_j > d \qquad ....(2.25)$$

Then with signal absent, E(y)=0 and $\sigma_y^2 = m\sigma_v^2$ the variance of each of the m noise samples. The density function of y with signal absent is then given by

$$f(y|o) = \frac{e^{-y^2/2\sigma_y^2}}{\sqrt{2\pi\sigma_y^2}} \qquad ....(2.26)$$

using eqn.(2.26), we can express the false-alarm probability by

$$P_n = \int_d^{\infty} f(y|0)\, dy = \int_c^{\infty} e^{-\frac{x^2}{\sqrt{\pi}}} dx = \frac{1}{2}(1 - erfc) \qquad ....(2.27)$$

where c is a parameter which is defined as

$$c \equiv \frac{d}{\sqrt{2}\sigma_y} = \frac{d}{\sqrt{2m}\sigma_v}$$

The false-alarm probability and its relation to the decision level d is shown in fig.(2.8) which is identical to the single-sample case with σ_y by replacing σ_v. Although the parameter d is proportional to $\sqrt{m}$, the peak of the f(y|s) curve is proportional to m, as shown. Thus as m increases, there is an improvement in detection.

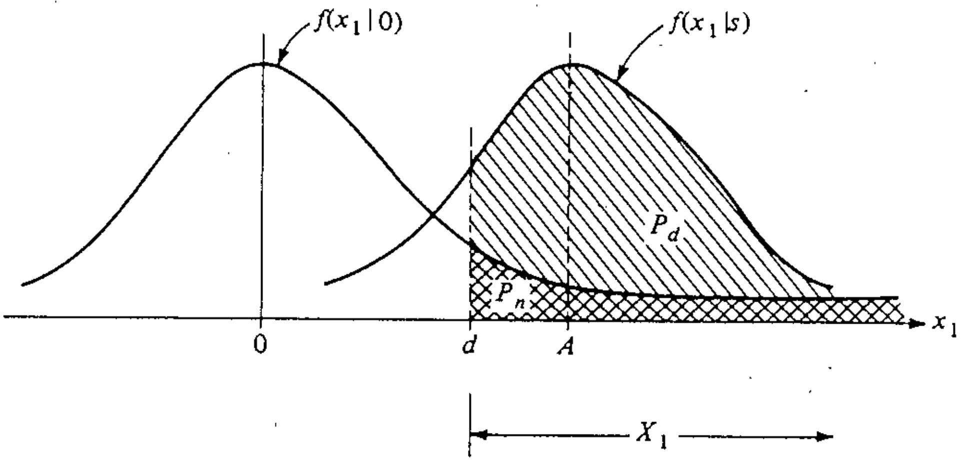

Fig. 2.7 Neyman-Pearson test of signal for one sample

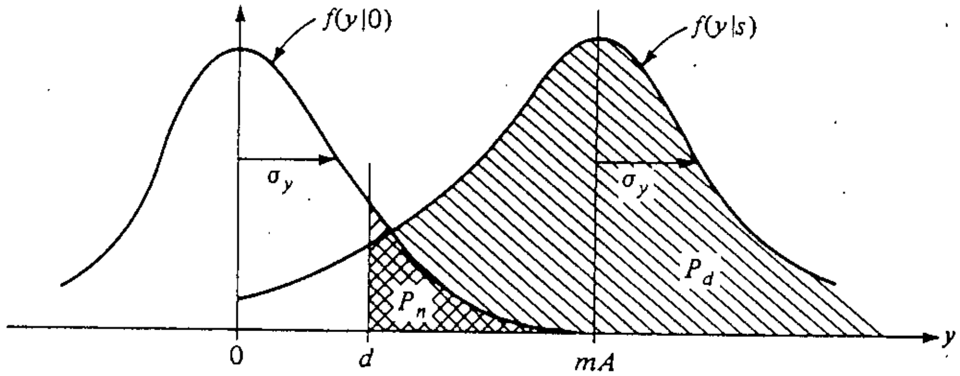

Fig. 2.8 Neyman-Pearson test of signal for m independent samples

The false-alarm probabilities of practical radars are quite small. The reason for this is that the false-alarm probability is the probability that noise pulse will cross the threshold during an interval of time approximately equal to the reciprocal of the bandwidth. Values of $P_n=10^{-8}$ to 10^{-10} and even smaller are much more realistic choices because radar signal pulses are commonly of the order of 2/μs in width or less. For $P_n=10^{-8}$ a false-alarm would occur on the average of once every 100 sec. For $P_n=10^{-10}$ the average false-alarm rate is once every 10^4 sec.

To solve the eqn.(2.27) for c when P_n is very small we have to proceed integration by parts and for small P_n the integral of eqn.(2.27) can be represented by a so-called asymptotic series (for $c \geq 3$ and $P_n \ll 1$)

$$P_n \approx \frac{e^{-c^2}}{2\sqrt{\pi}c}, \qquad for\ c^2 > 3 \qquad \text{.......................(2.28)}$$

If we take natural logarithm in the above equation and apply trial and error calculations we can obtain the values of c for the different values of P_n. But here eqn.(2.27) is selected to find the values of c (using known values of P_n) which is used in eqn.(2.32) to plot the curves showing the variation of the probability of detection with signal-to-noise ratio, fig.(2.10).

The probability of detection of signal P_d which corresponds to the probability that $y = \sum_{j=1}^{m} x_j > d$ with the signal present. Hence with the signal present E(y)=mA and $\sigma_y^2 = m\sigma_v^2$, as in the noise only case. This is indicated in fig.(2.8). Therefore, P_d is given by

$$P_d = \int_{c-\sqrt{m}A/\sqrt{2}\sigma_v}^{\infty} \frac{e^{-x^2}}{\sqrt{\pi}} dx = \frac{1}{2}\left[1 - erf\left(c - \frac{\sqrt{m}A}{\sqrt{2}\sigma_v}\right)\right] \qquad \text{..................(2.29)}$$

For $\sqrt{m}A/\sigma_v = \sqrt{2}c$ the probability of detection is 0.5. As $\sqrt{m}A/\sigma_v$ increases beyond this value, P_d increases, approaching 1. For a given probability of detection and probability of false-alarm, one may find the required signal-to-noise ratio A/σ_v or its equivalent power ratio A^2/σ_v^2. Increasing the number of samples m used is equivalent to increasing the effective A^2/σ_v^2 by m as well. The reason for this is apparent from fig.(2.8). As m increases, the width of the curves shown, or the standard deviation σ_y, increases as $\sqrt{m}$. The decision level d increases as $\sqrt{m}$ as well. But the expected value of the variable y with signal increases directly with m. The overlap of the two density functions decreases, and the probability of detection increases with m. The physical reason for this is as more signal samples are taken, their amplitudes add directly, providing the mA expected value of y, whereas the noise samples being added are random. This is the usual phenomenon in signal-detection problems; adding more and more independent samples makes the signal rise up out of the noise.

It was mentioned that the maximum number of uncorrelated noise samples available in a time interval T second is m=2BT, where bandwidth B prior to sampling and samples taken at a minimum spacing of 1/2B second intervals. For gaussian noise these samples are independent. Since the mean noise power for bandlimited white noise is proportional to the bandwidth B, so we can write

$$\sigma_v^2 = n_0 B \qquad \text{.....................(2.30)}$$

where n_0=white noise spectral density = FKT_0 (watt/hz). Here F is the noise figure, $K=1.38\times10^{-23}$ J/ °K and T_0 is the room temperature.

So, the effective signal-to-noise ratio,

$$\frac{mA^2}{\sigma_v^2} = \frac{2BTA^2}{n_o B} = \frac{2E}{n_o} \qquad \text{..........................(2.31)}$$

where E=the energy in a single pulse=A^2T.

Therefore, the detectability of the pulse is a function only of the effective signal-to-noise ratio, given by the ratio of its energy E to the noise spectral density n_0. This is an important result in detection theory. The ratio E/n_0 is used in determining the detectability of a signal in gaussian noise. It is obvious that the probability of detection can be improved by increasing the received signal energy E or by reducing the noise spectral density n_0. The energy is in turn given by the product of the signal power and the pulse duration T. Hence the detectability is increased either by increasing signal power or by lengthening the duration of the signal pulse. From this discussion probability of detection eqn.(2.29) can be reduced to the form

$$P_d = \int_{c-\sqrt{E/n_0}}^{\infty} \frac{e^{-x^2}}{\sqrt{\pi}} dx = \frac{1}{2}\left[1 - erf\left(c - \sqrt{\frac{E}{n_0}}\right)\right] \qquad \text{.................(2.32)}$$

This is a normalized form of the probability of detection which is applicable to the case of varying-amplitude signal pulses as well as the special case of rectangular pulses.

2.3 Application to Air Traffic Control Radar

In order to discuss the signal detection characteristics of the radars using the technique developed in this chapter, we need the radar equation that was developed in chapter 1, eqn.(1.4). We can now proceed to discuss both the primary and secondary radar. The pertinent specifications for the two radars (primary and secondary radar) at **Zia International Airport** are given in Table 2.1 and typical radar specifications for ASR and ARSR are given in Table 2.2. Relevant calculations and elaborate description of radar performance curves can be found in Schwartz & Shaw [1]. But the radar parameters for ASR and ARSR are given here in tabular form so that we can make a comparative study with primary and secondary radar specifications at a glance.

Usually the en route radars (ARSR) are designed to cover almost 3.4 times the range of the airport radars (ASR) and for this reason, they are provided with higher transmitter power. But since the secondary radar is equipped with the transponder, radar range extensively increase with lower transmitter power as the transponder acts as a midway receiving transmitting station. Considering a target having $2m^2$ reflecting area at the maximum range, the received powers are

$$P_r = \begin{cases} 7.15\times10^{-13} \text{ watt} & \text{for primary radar} \\ 2.91\times10^{-7} \text{ watt} & \text{for secondary radar (at the transponder)} \end{cases}$$

TABLE 2.1 **RADAR (Primary and Secondary RADAR) Parameters**

Radar Specifications	Primary Radar	Secondary Radar
Maximum Range R	160 km	370 km
Output Power P_t	3.5 MW	2.5 KW
Frequency Range	(1250-1350) MHz	(1030-1090) MHz
Pulse repetition freq.	633 Hz	316 Hz
Pulse duration	1.6 μsec	3 μsec
Pulse period	1.579 msec	3.16 msec
Wavelength	0.23 m	0.28 m
Antenna Gain	32 dB	20 dB
Antenna rotation	15 r.p.m.	15 r.p.m.
Antenna type	AOH-THD 286	AS 809
Intermediate frequency	30 MHz	60 MHz
Range resolution	120 meter	
Beamwidth	1.5°	1.35°

TABLE 2.2 **RADAR (ASR and ARSR) Parameters**

Radar Specification	**ASR (airport)**	**ARSR (en route)**
Maximum Range R	110 km	370 km
Output Power P_t	400 KW	4 MW
Frequency Range	2.7 GHz	1.3 GHz
Pulse repetition freq.	1200 Hz	360 Hz
Pulse duration	0.83 μs	2 μs
Pulse period	0.83 ms	2.8 ms
Wavelength	0.11 m	0.23 m
Antenna Gain	34 dB	34 dB
Antenna rotation	15 rpm	6 rpm
Beamwidth	1.5°	1.35°

From Table 2.1, it is obvious that the primary radar has 1400 times the transmitted power of the secondary radar. But the secondary radar can cover almost 2.3 times longer distance that of the primary radar since the secondary radar is equipped with the transponder. Since E is the energy in a rectangular pulse of amplitude A and width T, we therefore have $E=2P_tT$. The pulse widths for the two radars are $1.6\mu s$ for the primary radar and $3\mu s$ for the secondary radar respectively. Therefore we have

$$E = \begin{cases} 2.289\times10^{-18} \ \text{J} & \text{for primary radar} \\ 1.744\times10^{-12} \ \text{J} & \text{for secondary radar (at the transponder)} \end{cases}$$

As the wider pulse used for the secondary radar, target detection capability of it increase correspondingly since for the wider pulse more energy is received, but the minimum radial spacing of two targets required to distinguish them also increase. Fig.(2.9) shows how two airplanes can appear to be a single plane if the difference in propagation time to and from these targets is less than the pulse duration.

It can be mentioned here that for any signal, no matter how small its power, could be detected since it counts not the signal power or energy received but the ratio of signal-to-noise. The noise figure for both radar system is specified to be F=4dB or 2.512 means that 151 percent additional noise is introduced by the radar receiving system.

We may assume the usual room temperature on the ground is 300°K with F=4dB or 2.512 numerically. Then we have $n_0=1.04\times10^{-20}$ watt/hz and the E/n_0 ratios are

$$E/n_0 = \begin{cases} 23.43 \ \text{dB} & \text{for primary radar} \\ 82.24 \ \text{dB} & \text{for secondary radar (at the transponder)} \end{cases}$$

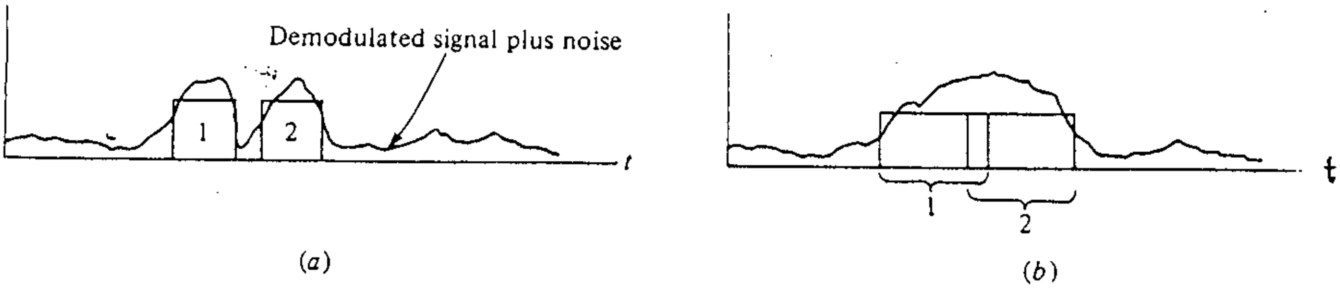

Fig.2.9 Radar resolution vs. pulse duration, (a) Narrow separated pulses, (b) Wide overlapping pulses

Upto this we have neglected several important effects and losses in this calculation and hence this result does not represent the whole picture. The power loss that was ignored in this system may be about 9dB and reducing the detection probability. But integration of radar pulses may consider as a compensating factor. Both the primary and the secondary radar rotate at a rate of 15 rpm, completing a 360° scan in 4s. The two radar beams have azimuth bandwidths of 1.5° and 1.35°, respectively. The number of pulses emitted per second are 633 and 316 pulses per second, respectively. This corresponds to 10.55 pulses reflected per target in the primary radar and 4.7 pulses per target in the secondary radar, providing that more samples in the signal processing or, equivalently, augmenting the received energy by the same amount. Thus the primary radar energy received should be 10.55E and the secondary radar energy 4.7E. Alternately, this results in E/n_0 improvement, of 10.23dB and 6.76dB, respectively.

Eqn.(2.32) is used to plot a family of curves relating probability of detection to signal-to-noise ratio. The signal-to-noise ratio needed to achieve a specified probability of detection without exceeding a specified false-alarm probability can be calculated from these graphs. A resultant set of curves describing the performance of the envelope-detector are labeled as noncoherent detection and the coherent-detection with c=4.5 to maintain $P_n=10^{-10}$ are given in Schwartz & Shaw [1].

Probability of detection vs. SNR

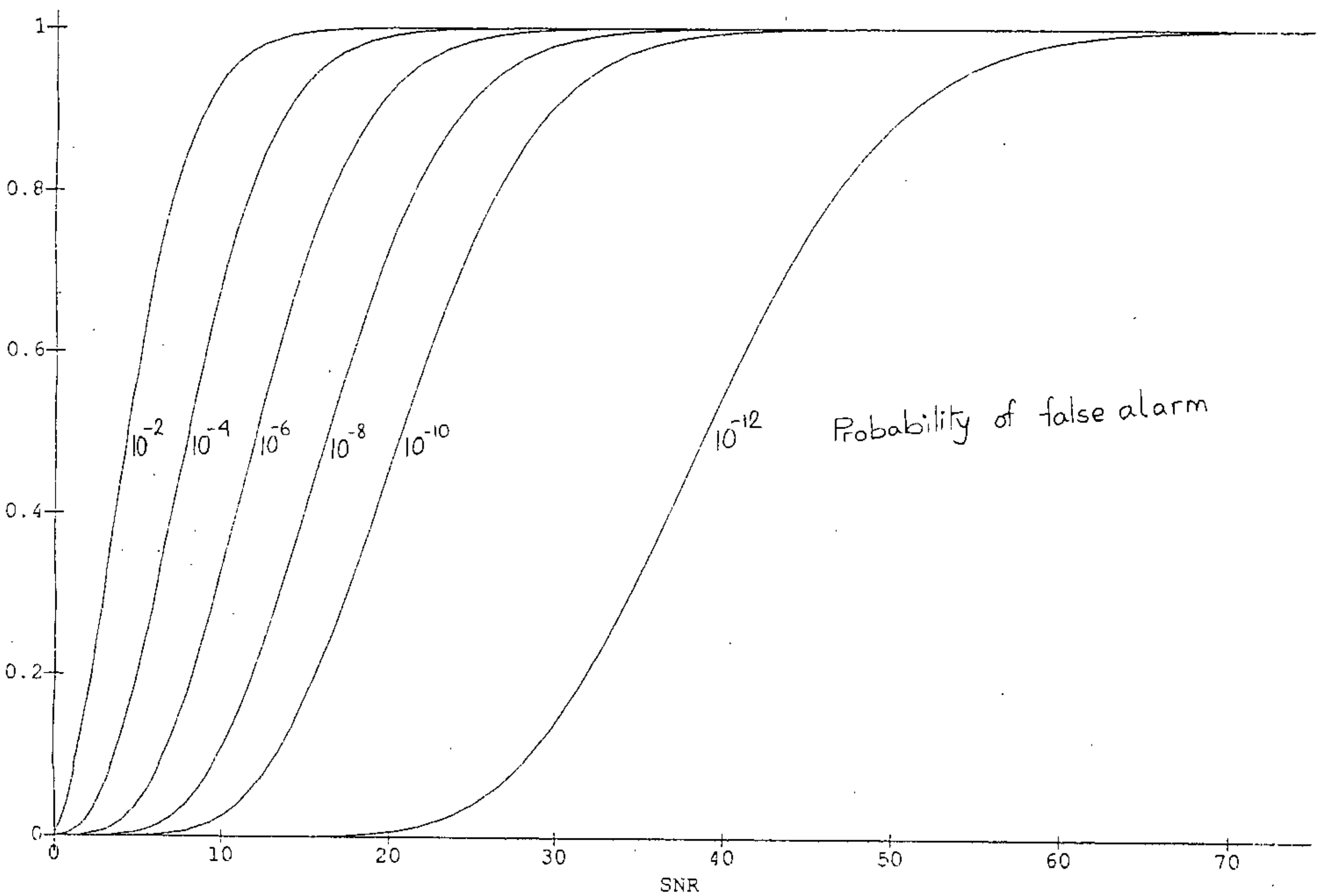

Fig. 2.10 The variation of probability of detection with signal-to-noise ratio

2.4 Summary

So far in this chapter we dealt with the detection of the presence of a signal in noise and in doing so we focused on two criteria primarily; (1) the minimization of the probability of error appropriate in digital communications, and (2) the Neyman-Pearson criterion appropriate to radar, sonar, and other systems in which the a priori probability of receiving a signal is not known. Detection capabilities of both primary and secondary radar are considered here. In the next chapter we focus on the estimation of random signals to extract the desired signal parameters.

CHAPTER 3

DIGITAL FILTERING OF NOISY DATA

3.0 Introduction

In this chapter we have discussed digital (i.e. discrete-time) filtering theory relevant to estimation theory as developed in the following chapters. Further on in chapter 3, we consider the nonrecursive and first-order recursive filter structures as estimators of signals in noisy data since they lead into discrete-time Kalman filter form to be developed in chapter 5. In this chapter and next, we approach the estimation problem in a step-by-step fashion. The mean-square error is used as a criterion to assess the degree of noise suppression by the filter estimators.

3.1 Estimation of Random Signals

In all practical problems the signal is present but because of noise, inaccuracies in the data samples, limited precision of instruments (these are often all modeled as additive noise), the data samples scatter about the actual signal values. The signal usually occupies a limited frequency range, while the noise is spread over a wide band of frequencies. So the signal is random since its precise value cannot be predicted in advance. Random signals are often called stochastic signals or stochastic processes. The purely random signal has a constant spectral density over all frequencies. This kind of signal is called white noise, by analogy to the property of white light containing all frequencies of visible spectrum. In order to remove at least partly the noise from the signal we would use some kind of filtering. The process of extracting the useful information from a signal and discarding the extraneous is called signal processing. There are a vast number of ways of processing signals. In most of the modern signal processing techniques one works with so-called discrete-time signal. Discrete-time implies that signals are defined only for discrete values of time, i.e. time is quantized. The process of picking values of an analog signal at a set of discrete times is called sampling. Such discrete-time signals often referred to as sampled-data signals. The widely-used term digital implies that both time and amplitude are quantized. A digital system is therefore one in which a signal is represented as a sequence of numbers which take on only a finite set of values. Sampling and quantization occur in all signal processing work involving digital computer. The combined operations are called analog-to-digital (A/D) conversion [1,2,9].

The filtering of random signals is referred to as estimation, because most estimation filters are statistical and estimation is a well-defined statistical technique. This text is concerned with the development of signal processing techniques to extract pertinent signal information from random signals utilizing any a priori information available. We call these techniques signal estimation, and the filters that use discrete-time algorithms are called signal estimator or just estimator. Sometimes estimators are called filters (e.g. Kalman filter) because they perform the same function as a deterministic filter except for random signals,i.e. they remove unwanted disturbances. Noisy measurements are processed by the estimator to produce filtered data [9].

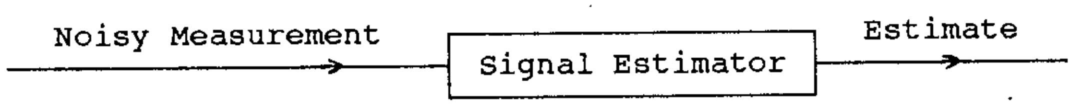

Fig. 3.1 Typical signal estimator structure

The amount of information available in the data is related to the precision (variance) of the particular measurement instrumentation used as well as to any signal processing devices or algorithms employed to improve the estimates. There are many different estimators and algorithms are available. We must have a reasonable measure to evaluate their performance, and decide which one is superior. The two primary statistical measures employed are the mean (accuracy) and the variance (precision). These measures lead to desirable estimator properties; i.e. we want estimators that are accurate or unbiased and precise. More formally, an unbiased estimator is one whose expected value is identical with the parameter being estimated.

Estimation can be thought of as a procedure made up of three primary parts:

(1) the specification of a criterion;
(2) the selection of models from a priori knowledge;
(3) the development and implementation of an algorithm.

Criterion functions are usually selected on the basis of information for which an estimator will be developed and can be classified as deterministic and probabilistic. For example, mean-square error criterion, maximum likelihood, minimum (error) variance etc. are some typical criterion functions.

Models represent a broad class of information formalizing the a priori knowledge about the process generating the signal, measurement instrumentation, noise characterization, probabilistic structure, etc. For example, a standard signal processing model is that of a signal in additive noise

Measurement = signal + noise

where the noise statistics are specified as well as the signal structure.

Finally, the algorithm or technique chosen to minimize (or maximize) the criterion can take many different forms depending on (1) the models, (2) the criterion, and (3) the choice of solution. The development of a particular algorithm is an interaction of selecting the appropriate criterion and models as shown in fig.(3.2). Thus the estimation procedure is a combination of these three major ingredients: criterion, models, and algorithm. Conceptually, this completes the discussion of the general estimation procedure [1,12,25].

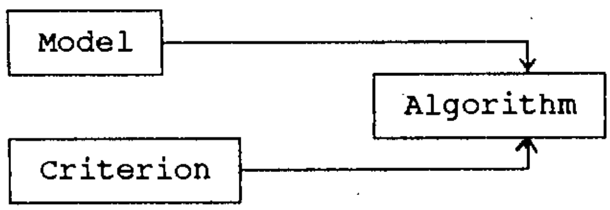

Fig. 3.2 Interaction of model and criterion in an estimation algorithm

3.2 Digital Transfer Function

In linear discrete-time system, the input f(k) and output g(k) sequences are related by linear difference equations with constant coefficients. In the theory of digital or discrete-time filters, the general equation is usually written in the following way:

$$g(k)+b_1g(k-1)+....+b_mg(k-m)= a_0f(k)+a_1f(k-1)+....+a_nf(k-n) \qquad ....(3.1)$$

where b_0 is taken, by convention, as unity. The interpretation of eqn.(3.1) is that at time k(t=kT), the output value can be computed from the current input and a linear combination of previous inputs and outputs. Taking the z-transform of eqn.(3.1) term-by-term, we obtain

$$G(z)\left(1+\sum_{j=1}^{m} b_j z^{-j}\right) = F(z)\sum_{i=0}^{n} a_i z^{-i} \qquad ....(3.2)$$

where $F(z) = \sum_{k=0}^{\infty} f(k)z^{-k}$

and $G(z)=\sum_{k=0}^{\infty} g(k)z^{-k}$

From eqn.(3.2) we can now define the discrete-time (or digital) transfer function as

$$H(z) = \frac{G(z)}{F(z)} = \frac{\sum_{i=0}^{n} a_i z^{-i}}{1+\sum_{j=1}^{m} b_j z^{-j}} \qquad ....(3.3)$$

which is a rational polynomial in z^{-1}. This transfer function is valid for zero initial conditions. From eqn.(3.3) we can write for the output

$$G(z)=H(z)X(z) \qquad ....(3.4)$$

and the output sequence y(k) is then obtained using the inverse z-transform.

From this discussion we have the output y(k) for the input x(k) as

$$y(k) = \sum_{i=0}^{n} a_i x(k-i) - \sum_{j=1}^{m} b_j y(k-j) \qquad \text{.........................(3.5)}$$

A hardware implementation of eqn.(3.5) is shown in fig.(3.3)

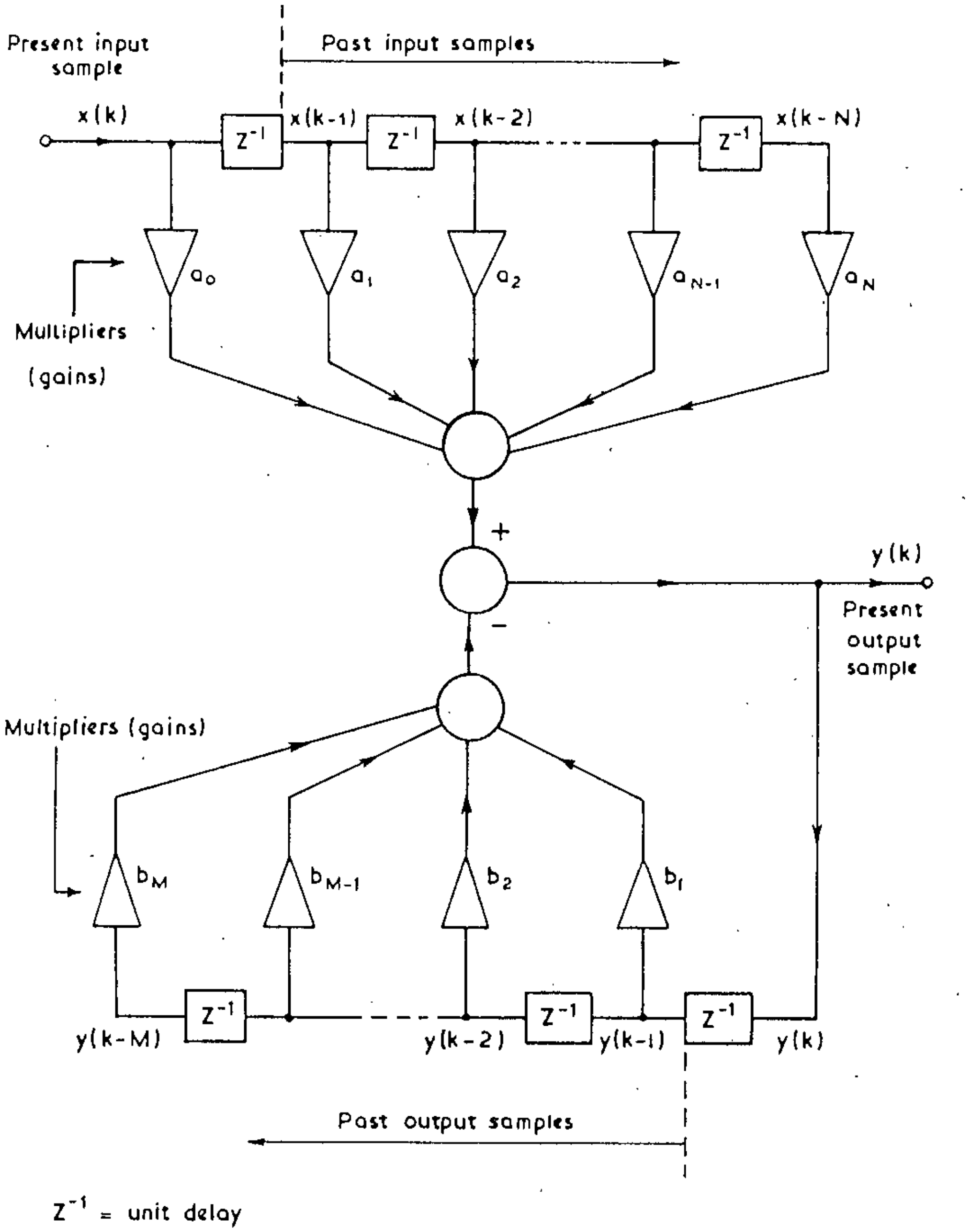

Fig. 3.3 Hardware implementation of a discrete-time filter

3.3 The Classification of Digital Filters

For the purpose of realization, digital filters are classified into nonrecursive and recursive types. The nonrecursive structure contains only the feed forward paths as shown in fig.(3.4). This is a special case of eqn.(3.3) in which all b_m coefficients are zero, i.e. the output is a sum of linearly weighted present and a number of previous samples of the input signal. The input and output sampled-data signals, denoted by f(k) and g(k) respectively, will be changed later to y and x where filters are considered as estimators. The output g(k) can be written as

$$g(k)=h(0)f(k)+h(1)f(k-1)+...+h(m-1)f(k-m+1)$$

or

$$g(k) = \sum_{i=0}^{m-1} h(i)f(k-i) \qquad \text{......................(3.6)}$$

which is the convolution summation in discrete-time. This represents a finite memory structure of finite number m of input samples.

In recursive filter structures the output depends both on the input and on the previous outputs, as shown in the general hardware realization of fig.(3.3), where we have both feed forward and feedback paths.

A simple first-order recursive filter structure is shown in fig.(3.5). Here the output consists of the present input and a weighted previous output, where the weight is denoted by a. The output-input relationship can be written directly, by inspection of fig.(3.5) as

$$g(k)=f(k)+ag(k-1) \qquad \text{.............(3.7)}$$

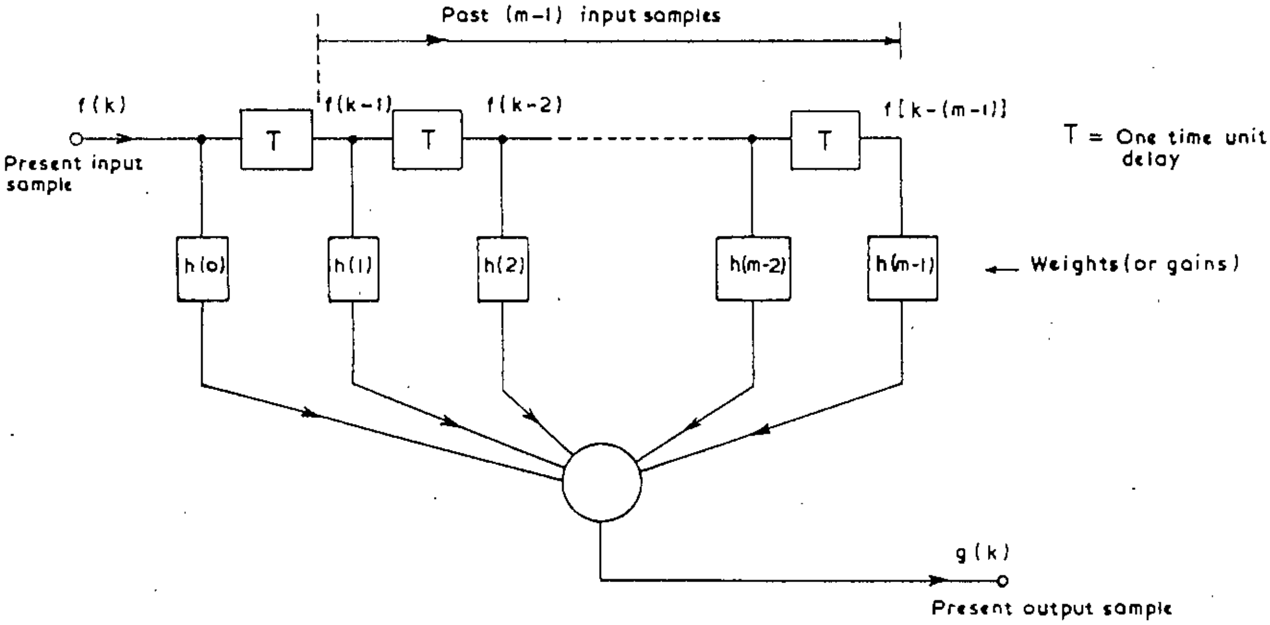

Fig. 3.4 Nonrecursive filter

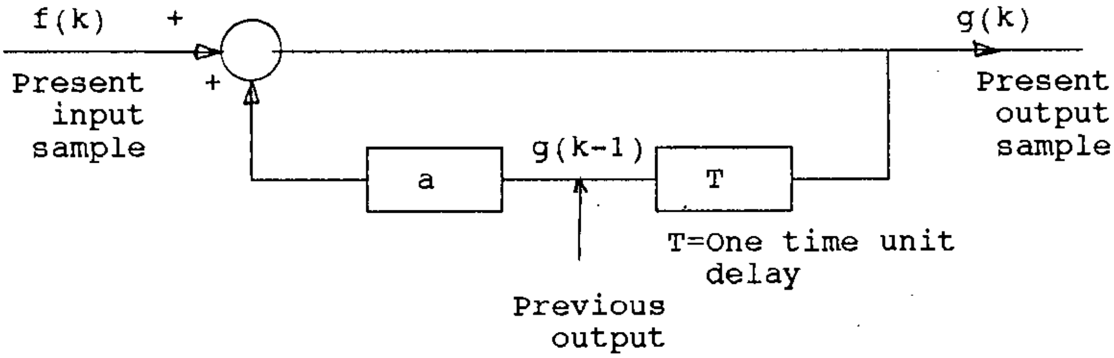

Fig. 3.5 Recursive (first-order) filter

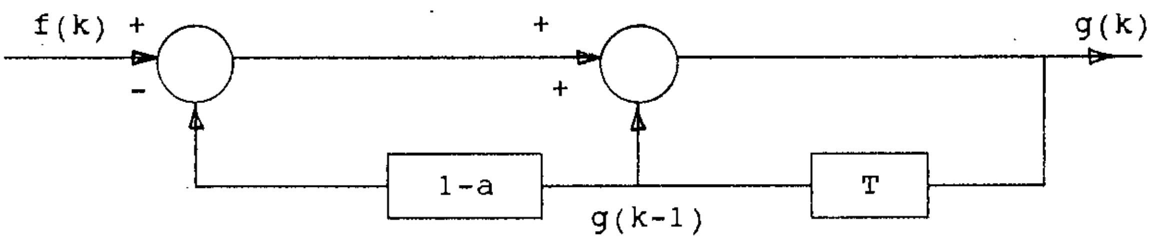

Fig. 3.6 An alternative form for fig. 3.5

Since the recursive filter is a feedback connection, so for stability we need $a<1$. An alternative form of the recursive filter is shown in fig.(3.6), for which we have

$$g(k)=g(k-1)+[f(k)-(1-a)g(k-1)] \qquad \text{...........(3.8)}$$

We shall come across this type of recursive structure in chapters 4,5 and 6.

These filters, nonrecursive and first-order recursive, are the structures on which estimation theory is based. They are often referred to both as filters and as batch and sequential processors. The first area, nonrecursive or batch processing, is also known as classical estimation theory, while the second area, recursive or sequential processing, can be called modern estimation theory.

The recursive structure has theoretically an infinite memory and hence it is referred to as an infinite impulse response (IIR) filter design. It is not unconditionally stable unless restrictions are placed on the values of the b coefficients.

Nonrecursive filter has only a limited memory, which is controlled by the number of delay stages and it results in the finite impulse response (FIR) filter design. It has only feedforward paths and is unconditionally stable. The input signal is delayed by a number of delay elements. The outputs of these time-delay elements are subsequently multiplied by a set of stored weights and the products summed to form the output signal. This implies that the output is given by the convolution of the input signal with the stored weights [2].

3.4 Nonrecursive Estimator

We use the notation x for a constant signal, and x(k) for the time-varying signal. Measurement of this signal is denoted by y(k), is linearly related to the signal x and an additive noise component v(k), introduced by random errors in measurements or any other causes. Therefore, we have

$$y(k)=x+v(k) \qquad \text{.............(3.9)}$$

The signal x is a random variable with some expected value $E(x)=x_0$ and variance σ_x^2. The noise samples are assumed to be of zero-mean with identical variances σ_v^2, and uncorrelated.

It is assumed that m data samples, as specified by eqn.(3.9), are to be processed using the nonrecursive filter structure of fig.(3.4), with all m weights equal to 1/m. The input f(k) is then y(k), and the output g(k) is taken as an estimate of parameter x, denoted by $\hat{x}$, as shown in fig.(3.7). Here data y(i), i=1,2,...,m, are available as a batch. They are stored, multiplied by equal weights, and the result is summed to produce the output

$$\hat{x} = \frac{1}{m} \sum_{i=1}^{m} y(i) \qquad \text{.........................(3.10)}$$

In general, the nonrecursive filter processor with different weights can be written as

$$\hat{x} = \sum_{i=1}^{m} h(i) y(i) \qquad \text{.........................(3.11)}$$

It is more common to use the square of the error $(\hat{x}-x)^2$ as a measure of the deviation between the signal and its estimate and the mean-square error $E(e^2)=E(\hat{x}-x)^2$ as the measure of goodness of the estimate. If we define the error between this estimate $\hat{x}$ and actual value x as $e=\hat{x}-x$, then the mean-square error is given by

$$p_e = E[e^2] = E(\hat{x}-x)^2 = E[x+v(k)-x]^2$$

or $$p_e = E[v^2(k)] = \sigma_v^2 \qquad \text{.............(3.12)}$$

Now for the estimate over m data samples, as expressed by eqn.(3.10), we have the mean-square error as given by

$$p_e = E(e^2) = E(\hat{x}-x)^2 = E\left\{\frac{1}{m}\sum_{k=1}^{m}[x+v(k)]-x\right\}^2$$

or

$$p_e = E\left\{\frac{1}{m}\left[\sum_{k=1}^{m} x + \sum_{k=1}^{m} v(k)\right] - x\right\}^2$$

or,

$$p_c = E\left[\frac{1}{m}\sum_{k=1}^{m} v(k)\right]^2$$

$$= \frac{1}{m^2}\sum_{i=1}^{m}\sum_{j=1}^{m} E[v(i)v(j)]$$

$$= \frac{1}{m^2}\sum_{i=1}^{m}\sum_{j=1}^{m} \sigma_v^2 \delta_{ij}$$

where δ_{ij} represents the Kronecker delta, i.e. $\delta_{ij}=1$ for i=j, and $\delta_{ij}=0$ for $i \neq j$. Therefore, we have

$$\sum_{i=1}^{m}\sum_{j=1}^{m} \delta_{ij} = \sum_{i=1}^{m}\left[\delta_{i1}+\delta_{i2}+\ldots+\delta_{im}\right]$$

$$\begin{aligned} = \quad & \delta_{11} + \delta_{12} + \ldots + \delta_{1m} \\ & + \delta_{21} + \delta_{22} + \ldots + \delta_{2m} \\ & + \delta_{m1} + \delta_{m2} + \ldots + \delta_{mm} \end{aligned}$$

In the above only δ_{ij} for i=j are equal to 1, all the others are zero. There are m such terms, therefore the result is then

$$p_e = \sigma_v^2/m \qquad \text{.................(3.13)}$$

which is an important relationship showing that as the number of samples m increases, the mean-square error p_e decreases. The sample mean is thus a good estimate of $\dot{x}$ in this sense.

The sample mean has another interesting property in estimating x. If we take the expectation of x in eqn.(3.10), which gives

$$E(\hat{x}) = \left\{\frac{1}{m}\sum_{i=1}^{m}[x+v(i)]\right\} = E(x) = x_0 \qquad \text{....................(3.14)}$$

Since, as stated earlier in this section, $E(x)=x_0$, and $E[v(i)]=0$.

Therefore, we have found that the estimate of x, on average, is the same as the average of the estimate. An estimator with this property is called an unbiased estimator, which, on average, produces the desired result.

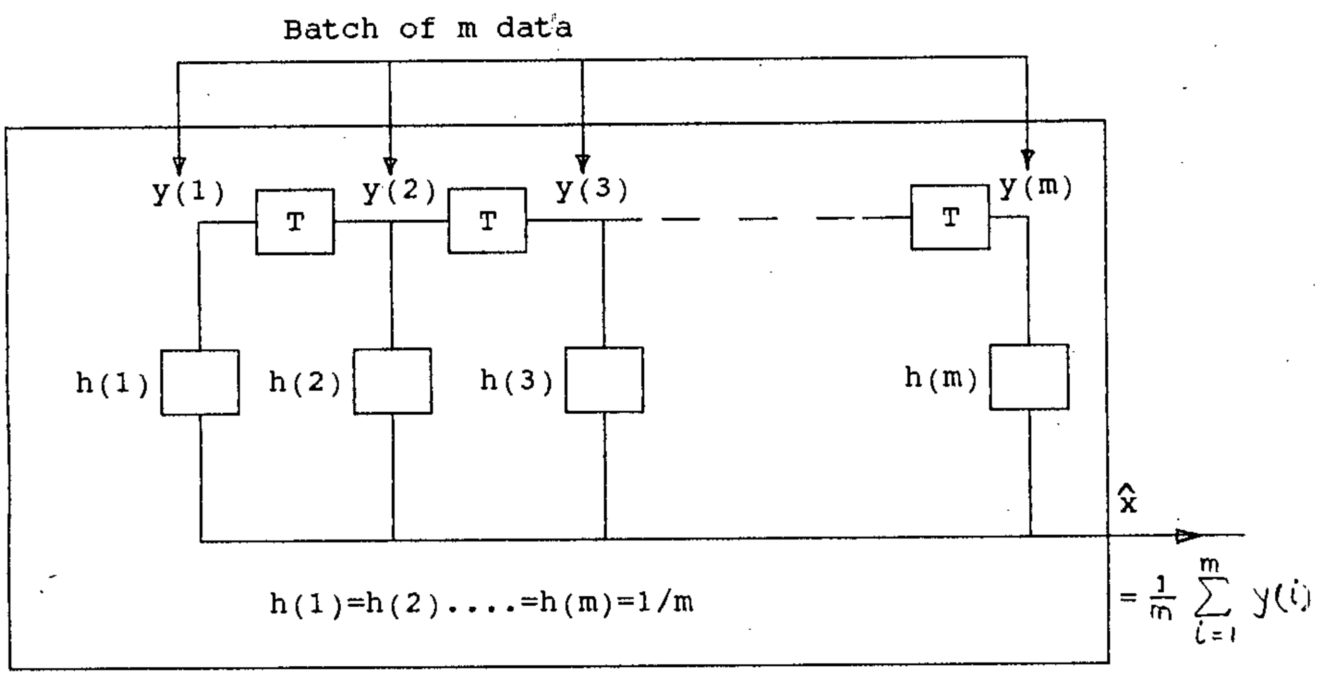

Fig. 3.7 Sample mean estimator

3.5 Recursive Estimator

We consider the simple first-order recursive filter shown in fig.(3.8), where y(k) and g(k) are the input and output sequences respectively. The input signal y(k) represents the measurements as expressed by eqn.(3.9), and the filter output is given by

$$g(k)=y(k)+ag(k-1) \qquad |a|<1 \qquad \text{............(3.15)}$$

This filter continually updates the output, adding a new data sample, y(k) to a fraction of the previous output ag(k-1). To find the result of such a process we consider sequentially input samples y(1), y(2),..., y(m), assuming g(k)=0 for k<1. From eqn.(3.15), we have

$$g(o) = y(o)+ag(-1) = 0$$

$$g(1) = y(1)+ag(o) = y(1)$$

$$g(2) = y(2)+ag(1) = y(2)+ay(1)$$

$$g(3) = y(3)+ag(2) = y(3)+ay(2)+a^2y(1)$$

.

.

$$g(m) = y(m)+ag(m-1) = y(m)+ay(m-1)+a^2y(m-2)+...+a^{m-2}y(2)+a^{m-1}y(1) \quad \text{...............(3.16)}$$

Substituting for y(k)=x+v(k), and separating the signal and noise terms, we have

$$g(m) = (1+a+a^2+...+a^{m-1})x+[v(m)+av(m-1)+...+a^{m-1}v(1)]$$

or, $$g(m) = \frac{1-a^m}{1-a}x+\sum_{i=1}^{m} a^{m-i}v(i) \quad \text{................(3.17)}$$

where the first term is the sum of the geometric series associated with x. For large m, $|a|^m<<1$, the signal part of g(m) approaches x/(1-a), while the variance due to the noise approaches $\sigma_v^2/(1-a^2)$. This indicate that a good estimate of x is given by $\hat{x}=(1-a)g(m)$ which leads to the important result

$$\hat{x} = (1-a^m)x+(1-a)\sum_{i=1}^{m} a^{m-i}v(i) \quad \text{....................(3.18)}$$

This means that the output (1-a)g(m) is taken as an estimate of signal x, after the mth input sample has been processed as shown in fig.(3.9).

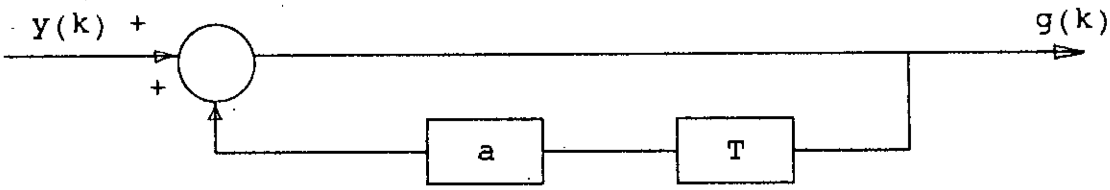

Fig. 3.8 Recursive filter as noisy data processor

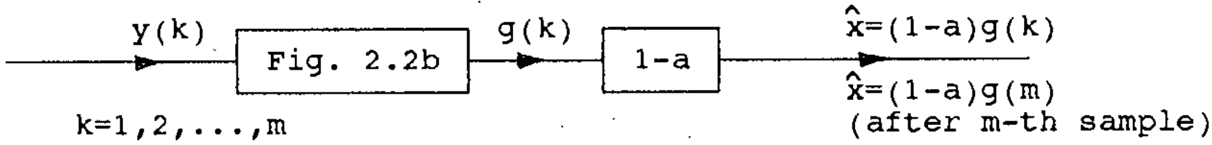

Fig. 3.9 Recursive filter as estimator

So by making the effective time constant a of the filter longer, we can reduce the mean-square error in the estimation of x. But this recursive device, although simple, is not the best estimate of x.

3.6 Summary

Two types of filters are discussed in this chapter. The nonrecursive filter has a finite memory and can have excellent linear phase characteristics, but it requires a large number of terms to obtain a relatively sharp cutoff frequency response. The recursive filter has an infinite memory and tends to have fewer terms, but its phase characteristics are not as linear as the nonrecursive ones. Nonrecursive filters have no feedback paths and hence no stability problem. Since the recursive filter is a feedback structure, the problem of stability must be considered. Moreover, both kind of filters associated with the error results from the quantization of the input data to a specified number of bits, and also due to the round-off in the multiplication and addition operations of the digital filters. To determine their effects on a digital filter, these sources of error are treated as random noise sources.

CHAPTER 4

OPTIMUM ESTIMATION OF SCALAR SIGNALS

4.0 Introduction

In the previous chapter we talked about the estimation of random signals. We have seen in sections 3.4 and 3.5 that the mean-square error is a useful criterion showing how good an estimation process is. In this chapter, the mean-square error is taken as the fundamental criterion.

The optimum nonrecursive estimator derived in section 4.1 is the scalar Wiener filter. Disadvantages of this batch type processor are summarized in section 4.2, where a recursive type (or sequential) processor is developed from the optimum nonrecursive type. The signal we consider in this chapter is a constant signal parameter with a random distribution of its values. In the next chapter we extend this technique to random time-varying single signal which is named as scalar signal [1-3].

4.1 Optimum Nonrecursive Estimator (Scalar Wiener Filter)

In this section we deal with the nonrecursive filter whose output is to be the signal estimate, i.e.

$$\hat{x} = \sum_{i=1}^{m} h(i)y(i) \qquad \text{...............(4.1)}$$

where y(1), y(2),...,y(m) are m data signals which is a linear batch processor. Now we choose the m coefficients, h(i), i=1,2,...,m, that minimizes the mean-square error, i.e. $p_e=E(e^2)= E(x-\hat{x})^2$ is minimized. Here x is desired signal and $\hat{x}$ is its estimate. To set up the problem we write

$$p_e = E\left[x - \sum_{i=1}^{m} h(i)y(i)\right]^2 \qquad \text{......................(4.2)}$$

We differentiate p_e with respect to each of the m parameters, setting each partial derivative equal to zero to obtain the required m equations from which to find the values of h(i) that minimizes the mean-square error. Thus we have

$$\frac{\partial p_e}{\partial h(j)} = -2E\left[x - \sum_{i=1}^{m} h(i)y(i)\right] y(j) = o \qquad \text{......................(4.3)}$$

or,

$$\sum_{i=1}^{m} h(i)\ E\left[y(i)y(j)\right] = E\left[xy(j)\right] \qquad \text{..................(4.4)}$$

where j=1,2,...,m. From eqn.(4.3) we can also write

$$E[ey(i)]=0 \quad \text{for } i=1,2,...,m \qquad \text{..............(4.5)}$$

where $e=x-\hat{x}$ is the error. This is called the orthogonality principle. It means that the product of the error $e=x-\hat{x}$ with each of the measured samples y(i) is equal to zero in an expected-value sense.

Probabilistic relationship which might exist between two random variables like y_n and y_{n+1} is called **correlation**. The correlation between pairs of signal samples may describe by the autocorrelation function defined as the expectation of the product of any two signal samples separated in time by k samples. For example, the estimator output y(i) has the autocorrelation function $R_y(k)$ such that

$$R_y(k)=E[y(i)y(i+k)]$$

if $R_y(k)=0$, i.e. zero autocorrelation function value, the variables are uncorrelated; while the larger the autocorrelation function values, the more correlated the variables are.

Therefore, returning to eqn.(4.4), we may introduce

$$E[y(i)y(j)] = R_{ij} \qquad \text{......... (4.6)}$$

which is the data autocorrelation between y(i) and y(j). Similarly, we introduce

$$E[xy(i)] = g_i \qquad \text{..........(4.7)}$$

which is the cross correlation between the random variables x and y(i). Using eqns.(4.6) and (4.7), we write eqn.(4.4) as

$$\sum_{j=1}^{m} h(j)R_{ij} = g_i \qquad \text{................(4.8)}$$

which is an important result. In expanded form this set of equations corresponds to

$$\begin{aligned} R_{11}h(1)+R_{12}h(2)+...+R_{1m}h(m) &= g_1 \\ R_{21}h(1)+R_{22}h(2)+...+R_{2m}h(m) &= g_2 \\ \cdots\cdots\cdots\cdots\cdots \\ R_{m1}h(1)+R_{m2}h(2)+...+R_{mm}h(m) &= g_m \end{aligned} \qquad \text{.......(4.9)}$$

More compactly, if we define the mxm correlation matrix R with elements R_{ij},i,j=1,2,...,m, a column vector **h** with elements h(1),h(2),...,h(m), and a column vector g with elements $g_1, g_2, ..., g_m$, we have as the set of equations in matrix-vector from,

$$\mathbf{Rh} = \mathbf{g} \qquad \text{.........(4.10)}$$

The formal solution of eqn.(4.10) is

$$\mathbf{h} = \mathbf{R}^{-1}\mathbf{g} \qquad \text{.........(4.11)}$$

with R^{-1} the inverse matrix of R. The estimate eqn.(4.1) can be written as

$$\hat{x} = \mathbf{h}^T\mathbf{y} \qquad \text{.........(4.12)}$$

where **h** and **y** are (mx1) column vectors, and $\mathbf{h}^T$ is a row vector. Substituting eqn.(4.11) into (4.12) we obtain for the estimate

$$\hat{x} = \mathbf{g}^T\mathbf{R}^{-1}\mathbf{y} \qquad \text{.........(4.13)}$$

and similarly for the least mean-square error

$$p_e = E(x^2)-\mathbf{g}^T\mathbf{R}^{-1}\mathbf{g} \qquad \text{.........(4.14)}$$

A filter of this type is often called a **scalar Wiener filter**, and eqn.(4.8) is known as the **scalar Wiener-Hopf equation.**

The solution vector **h** represents the optimum linear filter through which the data samples are to be passed. It should be noted that the relationship y(k)=x+v(k), eqn.(3.8), has not been used in the above derivations. Therefore, the result is more general than it appears. It states that if the data samples y(i), i=1,2,...,m, somehow contain the unknown random variable x, the signal, the best linear-filter operation carried out on the samples in order to estimate x is given by the Wiener filter [1].

If we consider the measured data and signal be related linearly by $y(k)=x+v(k)$, where $v(k)$ is additive noise, and the noise samples are zero-mean, with variance σ_v^2, uncorrelated with each other and with the signal x, then we have

$$E[v(j)v(k)] = \begin{cases} 0 & j \neq k \\ \sigma_v^2 & j=k \end{cases}$$

and $E[xv(j)]=0.$

Again, for the purpose of simplification, we assume $E(x)=0$, and hence $E(x^2)=\sigma_x^2$.
To solve this problem, we calculate first

$$\begin{aligned} R_{ij} &= E[y(i)y(j)] \\ &= E\{[x+v(i)][x+v(j)]\} \\ &= \sigma_x^2 + \sigma_v^2\delta_{ij} \end{aligned} \qquad \text{.........(4.15)}$$

where δ_{ij} is the Kronecker delta, i.e.

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

and $$g_i = E[xy(i)] = E[x^2] = \sigma_x^2 \qquad \text{.........(4.16)}$$

Substituting eqns.(4.15) and (4.16) into the set of eqns.(4.9) and summing both sides we have finally

$$\left(\sigma_v^2 + m\,\sigma_x^2\right)\sum_{i=1}^{m} h(i) = m\,\sigma_x^2$$

from which we have

$$\sum_{i=1}^{m} h(i) = \frac{m\,\sigma_x^2}{m\,\sigma_x^2 + \sigma_v^2}$$

and hence we can write

$$h(1) = h(2) = \ldots = h(m) = \frac{\sigma_x^2}{m\,\sigma_x^2 + \sigma_v^2} \qquad \ldots\ldots\ldots\ldots\ldots(4.17)$$

The least mean-squared estimate is thus given by

$$\hat{x} = \frac{1}{m + \gamma} \sum_{i=1}^{m} y(i) \qquad \ldots\ldots\ldots\ldots\ldots(4.18)$$

here $\gamma = \sigma_v^2/\sigma_x^2$. The corresponding least mean-square error value from eqns.(4.14) and (4.17) is given by

$$p_c = \frac{\sigma_v^2}{m + \gamma} \qquad \ldots\ldots\ldots\ldots\ldots(4.19)$$

The conclusion can be drawn is that for large signal-to-noise ratio ($\gamma \ll m$) the two errors are about the same and decreases as 1/m.

4.2 Recursive Estimator from the Optimum Nonrecursive Estimator

In the previous section we have derived the scalar Wiener filter equation. The difficulties with the Wiener filter are as follows:

(1) it requires previous knowledge (or stored estimates) of the autocorrelation matrix R ;

(2) the number of data samples m to be in the processing must specified beforehand;

(3) if m is changed for any reason (for example, more data may become available), the calculations must be done all over again;

(4) it requires the inversion of the (mxm) matrix R. If m is large, this can take substantial computer time.

To allow updating of the estimate as more information becomes available, and to save on digital processing cost, another processing scheme has been developed known as the recursive (or sequential) processor that continually generates a new estimate from the previous stored one plus the next data sample as it comes in. This will be the main theme of the following chapters, but first we derive a recursive algorithm from the nonrecursive solution obtained in the previous section.

For the recursive estimation of a signal parameter we can use the same problem specified as before defining the successive sampled as y(k)=x+v(k) which provide a linear estimator

$$\hat{x} = \sum_{i=1}^{k} h(i)y(i)$$

such that the mean-square error $p_e=E(x-\hat{x})^2$, is as small as possible. In section 4.1, we have the nonrecursive solution to this problem, and the results for k samples may be given as

$$\hat{x} = \hat{x}(k) = \sum_{i=1}^{k} h(i)y(i), \qquad where \quad h(i) = \frac{1}{k+\gamma} \qquad \text{.....................(4.20)}$$

with the corresponding mean-square error

$$p_e = p(k) = E\left[x-\hat{x}(k)\right]^2 = \frac{\sigma_v^2}{k+\gamma} \qquad \text{.....................(4.21)}$$

where $\gamma=\sigma_v^2/\sigma_x^2$ and $\hat{x}(k)$ represents to the kth estimate of the parameter x, i.e. the estimate after a batch of k samples have been processed.

For (k+1) samples the estimate and the corresponding mean-square error would be

$$\hat{x}(k+1) = \sum_{i=1}^{k+1} h(i)y(i), \qquad where \quad h(i) = \frac{1}{(k+1)+\gamma} \qquad \text{.....................(4.22)}$$

and,

$$p(k+1) = \frac{\sigma_v^2}{(k+1)+\gamma} \qquad \text{.....................(4.23)}$$

From eqns.(4.20) and (4.21), we have

$h(i)=p(k)/\sigma_v^2$ for k samples

and similarly from eqns.(4.22) and (4.23)

$h(i)=p(k+1)/\sigma_v^2$ for k+1 samples

From the above two equations we can form the ratio

$$\frac{p(k+1)}{p(k)} = \frac{k+\gamma}{k+1+\gamma} = \frac{1}{1+1/(k+\gamma)}$$

from which we can write

$$\frac{p(k+1)}{p(k)} = \frac{1}{1+p(k)/\sigma_v^2} \qquad \text{....................(4.24)}$$

This is a difference equation. Using this equation from the known value of p(k), we can find p(k+1), then, p(k+2), etc. Therefore this is a simple algorithm for finding the variation of mean-square error with sample size.

For the signal estimate $\hat{x}(k+1)$, after processing (k+1) samples, we may write from eqn.(4.22).

$$\hat{x}(k+1) = \frac{1}{k+1+\gamma}\sum_{i=1}^{k} y(i) + \frac{1}{k+1+\gamma} y(k+1)$$

from which using eqn.(4.20), we obtain

$$\hat{x}(k+1) = \frac{k+\gamma}{k+1+\gamma}\hat{x}(k) + \frac{1}{k+1+\gamma} y(k+1)$$

Using eqn.(4.24), we have

$$\hat{x}(k+1) = \frac{p(k+1)}{p(k)}\hat{x}(k) + \frac{p(k+1)}{\sigma_v^2} y(k+1) \qquad \text{....................(4.25)}$$

This is a recursive estimating equation which together with eqn.(4.24) forms the required recursive algorithm. We use eqn.(4.24) to find p(k+1) in terms of p(k). Then, from the stored previous value $\hat{x}(k)$ and the new data sample y(k+1), we can calculate $\hat{x}(k+1)$. This procedure has the property that it continually generates the best linear mean-square estimator of x, and at the same time it provides the corresponding mean-square error, p(k+1). From the error relation of eqn.(4.24), we have, $p(k) \rightarrow 0$ for k very large.

To start this recursive process we must calculate the first estimate $\hat{x}(1)$, based on a single observation, by nonrecursive methods.

If we now compare the recursive relationship of eqn.(4.25) with the recursive filter of eqn.(3.15) in section 3.5, we find that it is of the same form, but with time-varying coefficients. Denoting these coefficients by

$$a(k+1) = \frac{p(k+1)}{p(k)} \quad , \quad b(k+1) = \frac{p(k+1)}{\sigma_v^2} \qquad \text{.....................(4.26)}$$

eqn.(4.25) becomes

$$\hat{x}(k+1) = a(k+1)\hat{x}(k)+b(k+1)y(k+1) \qquad \text{.....................(4.27)}$$

Using eqns.(4.26) and (4.24), we can show that the parameters a(k+1) and b(k+1) are related in the following way:

$$a(k+1) = 1-b(k+1) \qquad \text{.....................(4.28)}$$

Therefore,eqn.(4.27) can be written as

$$\hat{x}(k+1) = \hat{x}(k) + b(k+1)[y(k+1)-\hat{x}(k)] \qquad \text{.....................(4.29)}$$

The interpretation of this result is quite interesting. It shows that the (k+1)th estimate is the same as the previous kth estimate plus a correction term involving the difference between the new sample value y(k+1) and the previous estimate. This correction term is multiplied by time-varying gain factor b(k+1) which continually decreases with k. So ultimately, the estimate stabilizes at some value depending on the data sample, and will be modified only if a new sample y(k+1) differs considerably from the previous estimate. The two forms of recursive estimator for eqns.(4.27) and (4.29) are shown in figs.(4.1) and (4.2) respectively.

This recursive filter has been derived from the nonrecursive filter solution for a particular case. It can be shown that the same result is obtained by starting with a recursive filter structure having two parameters a(k+1) and b(k+1) and using them to minimize the mean-square error of the estimate. This is the approach used in chapter 5 [2,3,25].

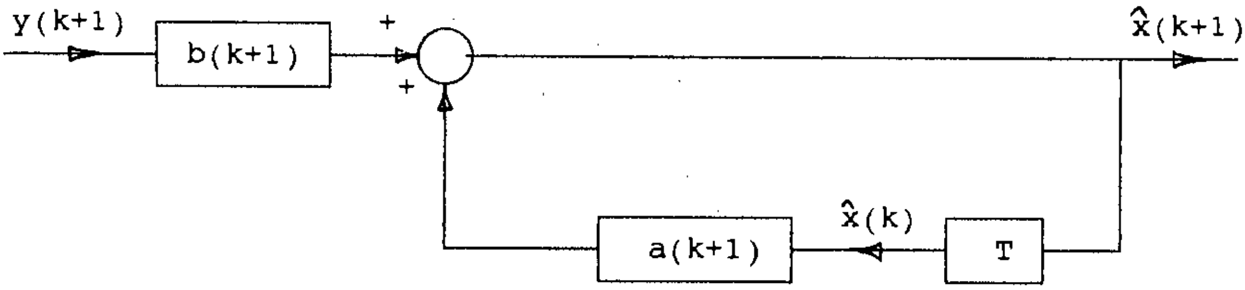

Fig. 4.1 Recursive filter

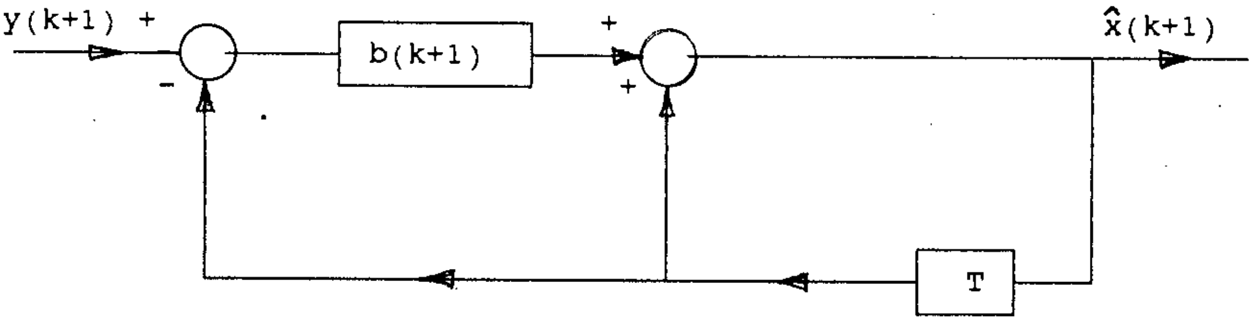

Fig. 4.2 An equivalent form of fig. 4.1

4.3 Summary

In this chapter estimation theory is presented in terms of both optimum nonrecursive and first-order recursive estimators based on minimization of the mean-square error. The scalar Wiener filter considered in section 4.1 is essentially a block process estimate which is best suited to the situation where only a finite block of data is available and they are typically applied in areas such as seismic surveying and image processing. In dealing with an infinite time series the Weiner filter would require a complete recalculation of all auto- and cross-correlation terms for each new input sample. The optimum recursive (or Kalman) estimator uses this new knowledge to update a recursive estimate [Kalman, Bozic, Kailath 1981]. Recursive means that we do not have to store the entire data vector which continues to grow as it advances, and entirely recompute the estimate at each stage. In a recursive filter, the estimate itself serves as a summary of all the past data. It is only necessary to modify the estimate if the new data is arrives. The recursive estimation technique that introduced in section 4.2 will be discussed in detail in the next chapter where the optimum scalar kalman filter will be derived and this will be extended in chapter 6 to cover the vector kalman formulation.

CHAPTER 5

KALMAN FILTERING

5.0 Introduction

The recursive estimation techniques were developed around 1960, most notably by Rudolph E. Kalman. For this reason, the processors devised at that time (as well as wide variety of generalizations and extensions to time-varying statistics, continuous-time signals, nonlinear dynamics, etc.) are referred to as Kalman filters. This chapter describes the Kalman filter in detail and we discuss recursive approach to the filtering and prediction of random processes. For simplicity we assume the processes to be stationary with time. The random signal and purely additive noise components are also assumed to be statistically independent [1-4,12,22].

5.1 Scalar Kalman Filter

In this section we generalize the analysis in the following ways:

(1) We deal with randomly time-varying signals or random processes;

(2) The observation (data) equation is changed by a factor c multiplying the signal. We need it to enable the generalization of results to vector signals;

(3) We derive the optimum estimate for a generalized first-order recursive filter. Results are arranged in so-called scalar Kalman filter form, suitable for a direct transformation into vector Kalman filter.

We assume that the random signal to be estimated can be modelled as a first-order recursive process driven by zero-mean white noise and can be expressed by the dynamical equation

$$x(k) = ax(k-1)+w(k-1) \qquad \text{..........(5.1)}$$

Fig.(5.1) is a block-diagram representation of eqn.(5.1). If we assume the initial sample values to be zero, i.e. $x(k)=0$ and $w(k)=0$ for $k<0$. The random drive is specified by

$$E[w(k)]=0$$

$$E[w(k)w(j)] = \begin{cases} 0 & k \neq j \\ \sigma_w^2 & k=j \end{cases} \qquad \text{..........(5.2)}$$

If $\sigma_w^2=0$, the white noise process will disappear.

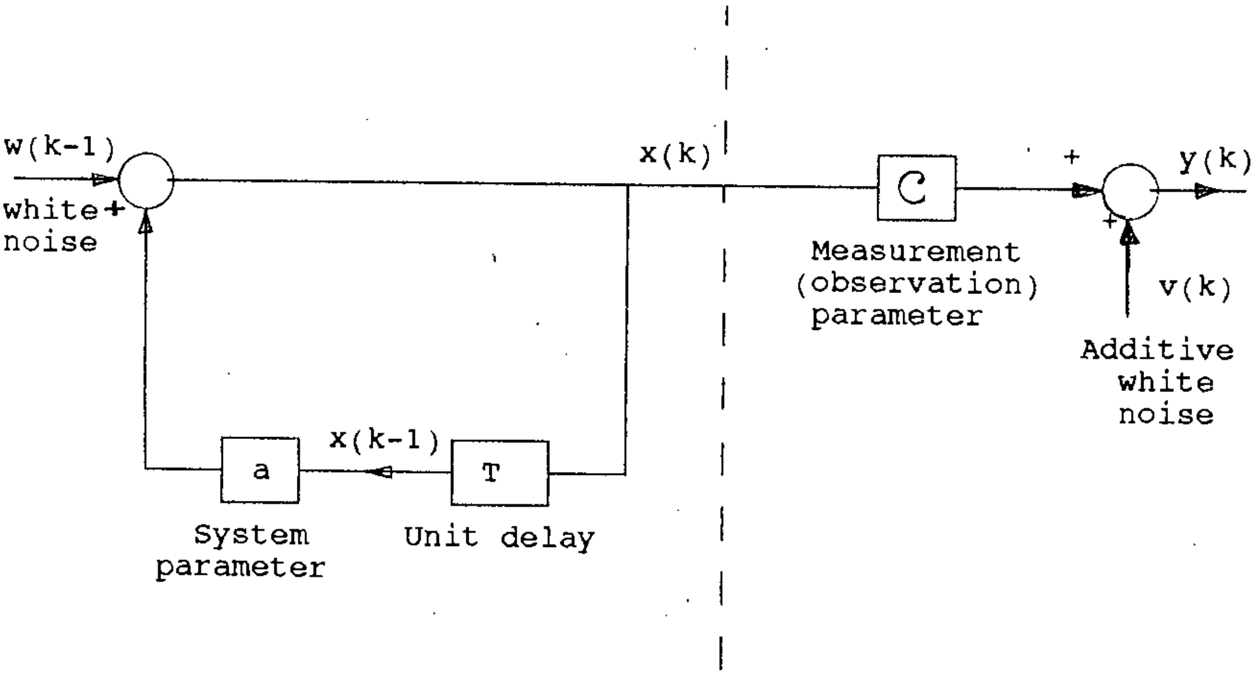

Fig. 5.1 Signal model

Fig.5.2 Measurement model

A random process defined by eqn.(5.1) is said to be an autoregressive process of the first order and the statistical parameter of x(k) are

$$E[x(k)] = 0$$
$$E[x^2(k)] = \sigma_x^2$$
$$E[x(k)x(k+\ell)] = R_x(\ell) = a^{|\ell|}\sigma_x^2 \qquad \text{..........(5.3)}$$

Here $R_x(\ell)$ is the autocorrelation of two samples of x(k) and ℓ represents the spacing between samples, σ_x^2 is the signal variance. The parameter 'a' plays the role of a time constant of the process. The larger a is (approaching 1), the more sluggish the process is, requiring a longer time interval (in terms of units of sample spacing T) to change significantly from its current value [1].

The widely used observation model for additive noise as shown in fig.(5.2), described by the equation

$$y(k) = cx(k)+v(k) \qquad \text{..........(5.4)}$$

where x(k) is the time-varying signal and the factor c represents an observation or measurement parameter. It will be seen later that this factor is useful for the transformation of results to vector signals. Here v(k) represents an independent additive white noise with zero-mean and variance σ_v^2.

5.2 Optimum Filter Derivation

We assume the recursive estimator is to be of the form

$$\hat{x}(k) = a(k)\hat{x}(k-1)+b(k)y(k) \qquad \text{..........(5.5)}$$

where the first term represents the weighted previous estimate and the second term is weighted current data sample. We want to determined the 'best' estimate using eqn.(5.5). By 'best' we mean the estimate which minimizes the mean-square error. In this case we have two parameters, a(k) and b(k), which are to be determined from minimization of the mean-square error

$$p(k) = E[e^2(k)] \qquad \text{..........(5.6)}$$

where $e(k)=\hat{x}(k)-x(k)$ is the error.

Substituting eqn.(5.5) for x(k), we have

$$p(k) = E[a(k)\hat{x}(k-1)+b(k)y(k)-x(k)]^2 \qquad \text{..........(5.7)}$$

Differentiating with respect to a(k) and b(k) and the results are equated to zero, we have

$$\frac{\partial p(k)}{\partial a(k)} = 2E\,[a(k)\,\hat{x}(k-1)+\ b(k)\,y(k)-x(k)]\ \hat{x}(k-1) = 0 \qquad \text{..............(5.8)}$$

and

$$\frac{\partial p(k)}{\partial b(k)} = 2E\,[a(k)\,\hat{x}(k-1)+\ b(k)\,y(k)-\hat{x}(k)]\,y(k) = 0 \qquad \text{..............(5.9)}$$

or alternately

$$E[e(k)\hat{x}(k-1)] = 0 \qquad \text{..........(5.10)}$$

and $$E[e(k)y(k)] = 0 \qquad \text{..........(5.11)}$$

which are orthogonality equations. A relationship between a(k) and b(k) may be derived using eqns.(5.10) and (5.5) such that

$$E\{[a(k)\hat{x}(k-1)+b(k)y(k)-x(k)]\hat{x}(k-1)\} = 0$$

Adding and subtracting a(k)x(k-1), the above becomes

$$E\{[a(k)[\hat{x}(k-1)-x(k-1)] + a(k)x(k-1)]\hat{x}(k-1)\} = E\{[x(k)-b(k)y(k)]\hat{x}(k-1)\}$$

Substituting the value for y(k) from eqn.(5.4) gives us

$$a(k)E[e(k-1)\hat{x}(k-1) + x(k-1)\hat{x}(k-1)] = E\{[x(k)[1-cb(k)]-b(k)v(k)]\hat{x}(k-1)\}$$

For the optimum estimator the orthogonality principle must hold and if we apply this we get the following relationships:

$$E[e(k-1)\hat{x}(k-1)] = 0$$

and, $E[v(k)\hat{x}(k-1)] = 0$

Therefore, the above equation reduces to

$$a(k)E[x(k-1)\hat{x}(k-1)] = [1-cb(k)]E[x(k)\hat{x}(k-1)]$$

From our signal generation model, $x(k) = ax(k-1)+w(k-1)$, and substituting this in above equation we obtain

$$a(k)E[x(k-1)\hat{x}(k-1)] = [1-cb(k)]E[ax(k-1)\hat{x}(k-1)+w(k-1)\hat{x}(k-1)]$$

Using the eqns.(5.5), (5.4) and (5.1) we can write

$$\hat{x}(k-1) = a(k-1)\hat{x}(k-2)+acb(k-1)x(k-2)+cb(k-1)w(k-2)+b(k-1)v(k-1)$$

Since all terms are uncorrelated with w(k-1), averages of all the products of the above with w(k-1) are zero, so we are now left with

$$a(k)E[x(k-1)\hat{x}(k-1)] = a[1-cb(k)]E[x(k-1)\hat{x}(k-1)]$$

This leads to the final relationship between a(k) and b(k) giving

$$a(k) = a[1-cb(k)] \qquad \text{.........(5.12)}$$

Substituting this into the eqn.(5.5), we have

$$\hat{x}(k) = a\hat{x}(k-1)+b(k)[y(k)-ac\hat{x}(k-1)] \qquad ....(5.13)$$

Eqn.(5.13) is the definition of the optimum first-order recursive estimator or **scalar Kalman filter**. The first term, $a\hat{x}(k-1)$, is a prediction of the current sample based on past observations. The second term is a correction term depending on the difference between the new data sample and the observation estimate, $\hat{y}(k)=ac\hat{x}(k-1)$, modified by the variable gain factor b(k), called the **Kalman gain**. The form of this filter is clearly illustrated in fig.(5.3). The aim of the optimum estimator in fig.(5.3) is to determine an optimum estimate $\hat{x}$ from the received signal x by using a priori knowledge.

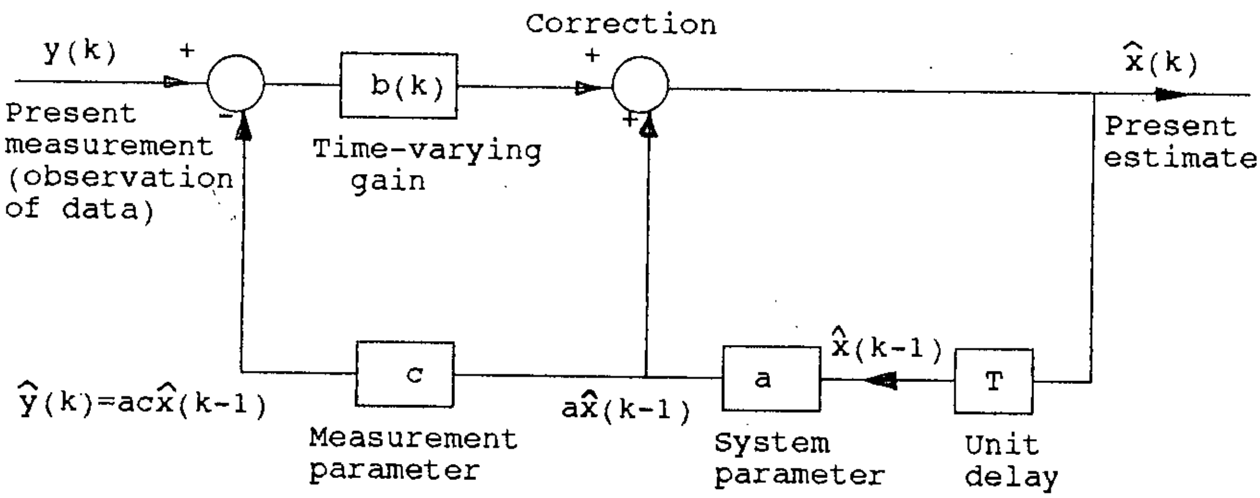

Fig. 5.3 Optimum recursive estimator (filter)

5.3 Derivation of the Kalman Gain

Having define the form of the Kalman filter the factor that remains to be derived is the time-varying Kalman gain term b(k).
From eqn.(5.6) we have

$$p(k) = E[e^2(k)] = E\{e(k)[\hat{x}(k)-x(k)]\}$$

Substituting for x(k) from eqn.(5.5), and using the orthogonality eqns.(5.10) and (5.11) we obtain

$$p(k) = -E[e(k)x(k)]$$

From eqns.(5.11) and (5.4) we have

$$cE[e(k)x(k)] = -E[e(k)v(k)$$

Hence the mean-square error is given by

$$p(k) = \frac{1}{c}E\,[e(k)\,v(k)]$$

Substituting for $e(k)=\hat{x}(k)-x(k)$, and using eqn.(5.5), we have

$$p(k) = \frac{1}{c}\;E\;[a(k)\,\hat{x}(k-1) + b(k)\,y(k) - x(k)]\,v(k)$$

But $E[\hat{x}(k-1)v(k)]$ and $Ex(k)v(k)]$, average to zero. So we are left with

$$p(k) = \frac{1}{c}\;b(k)\,E\;[y(k)\,v(k)] = \frac{1}{c}\,b(k)\sigma_v^2 \qquad \text{..............(5.14)}$$

so $$b(k) = cp(k)/\sigma_v^2 \qquad \text{.........(5.15)}$$

Now substituting eqn.(5.13) into the mean-square error equation, we have

$$p(k) = E\{a\hat{x}(k-1)+b(k)[y(k)-ac\hat{x}(k-1)]-x(k)\}^2$$

and using eqn.(5.4) and (5.13), we have

$$p(k) = E\{a[1-cb(k)]e(k-1)-[1-cb(k)]w(k-1)+b(k)v(k)\}^2$$

Since e(k-1), w(k-1) and v(k) are independent of each other, the cross products in the above expression average to zero. So, we have

$$p(k) = a^2[1-cb(k)]^2p(k-1)+[1-cb(k)]^2\sigma_w^2+b^2(k)\sigma_v^2$$

where $p(k-1) = E[e^2(k-1)]$

$$\sigma_w^2 = E[w^2(k-1)]$$

and $\sigma_v^2 = E[v^2(k)]$

From the above quadratic equation of b(k) we have the solution for b(k) such that

$$b(k)\left\{ \sigma_v^2 + c^2 \left[a^2 p(k-1) + \sigma_w^2\right] \right\} = c\left[a^2 p(k-1) + \sigma_w^2\right],$$ where we use $p(k)=b(k)\sigma_v^2/c$.

From which we have

$$b(k) = \frac{c\left[a^2 p(k-1) + \sigma_w^2\right]}{\sigma_v^2 + c^2 \sigma_w^2 + c^2 a^2 p(k-1)} \qquad \text{.............(5.16)}$$

The other solution of the quadratic equation for b(k) is b(k)=1/c. This solution is neglected because it is time-invariant as c is constant, while the first solution is time-varying through p(k-1). It is apparent that b(k) must be calculated first from a knowledge of p(k-1) and then p(k) is calculated from

$$p(k) = \frac{1}{c}\sigma_v^2\, b(k) \qquad \text{.............(5.17)}$$

i.e., $p(k)$ and $b(k)$ are directly related. Three equations given as 5.13, 5.16 and 5.17 constitute a complete computational algorithm. For the purpose of extending these results to vector signals in the next chapter, we arrange these equations and write them below as equations 5.18 to 5.21.

In this arrangement eqns.(5.16) and (5.17) are now written as three equations, 5.19 to 5.21, because we have introduced a new quantity, $p_1(k)$ as it has an important role and will be discussed later.

Recursive filter estimator:

$$\hat{x}(k)=a\hat{x}(k-1)+b(k)[y(k)-ac\hat{x}(k-1)] \tag{5.18}$$

Filter gain:

$$b(k)=cp_1(k)[c^2p_1(k)+\sigma_v^2]^{-1} \tag{5.19}$$

$$\text{where } p_1(k)=a^2p(k-1)+\sigma_w^2 \tag{5.20}$$

Mean-square error:

$$p(k)=p_1(k)-cb(k)p_1(k) \tag{5.21}$$

The above set of equations constitute the **scalar Kalman filter** for the signal model given by eqn.(5.1) and the measurement model given by eqn.(5.4).

5.4 Scalar Kalman Predictor

In the previous section we were concerned with the estimation of the current value of a random signal in additive white noise. But in many real life situations, it is often required, particularly in control systems, to predict ahead, if possible. Depending on how many steps of unit time ahead we want to predict, we distinguish one-step, two-step, or m-step prediction. Obviously the further in the future we want to predict, the larger the prediction error will be. We deal here only with one-step prediction, with the signal model as described by the eqn. (5.1) and the observation (or measurement) model as described by the eqn.(5.4). We would like to know the 'best' linear estimate of x(k+1), i.e. the signal at time k+1, given the data and previous estimate at time k. We denote this one-step prediction estimate as $\hat{x}(k+1|k)$. By 'best' we mean the predictor that minimizes the mean-square prediction error

$$p(k+1|k) = E[e^2(k+1|k)]$$
$$= E[x(k+1-\hat{x}(k+1|k)]^2 \qquad \text{.........(5.22)}$$

This is comparable to the mean-square error $p(k)=E[x(k)-\hat{x}(k)]^2$ in the filtering problem. By extension of the previous discussion on filtering it is apparent that the one-step linear predictor will be of the form

$$\hat{x}(k+1|k) = \alpha(k)\ \hat{x}(k|k-1) + \beta(k)y(k) \qquad \text{.............(5.23)}$$

The parameter $\alpha(k)$ and $\beta(k)$ are determined from the minimization of the mean-square prediction error given by eqn.(5.22) by using the appropriate orthogonality relations similar to those derived in the previous section

$$E[e(k+1|k)\hat{x}(k|k-1)] = 0 \qquad \text{.........(5.24)}$$

$$E[e(k+1|k)y(k)] = 0 \qquad \text{.........(5.25)}$$

The relationship between $\alpha(k)$ and $\beta(k)$ is determined using a similar way to the relationship between a(k) and b(k) derived for the filtering case such that

$$\alpha(k) = a - c\,\beta(k) \qquad \text{.............(5.26)}$$

Substituting this result into the prediction equation, we have

$$\hat{x}(k+1|k) = a\ \hat{x}(k|k-1) + \beta(k)\ [y(k) - c\ \hat{x}(k|k-1)] \qquad \text{.............(5.27)}$$

The variable gain term $\beta(k)$ is determined, together with $p(k+1|k)$ from eqns.(5.25) and (5.22), using a similar method as in the filtering case and we obtain

$$p(k+1|k) = \frac{a}{c}\sigma_v^2 \beta(k) + \sigma_w^2 \qquad \text{.............(5.28)}$$

and, $$\beta(k) = \frac{a\,c\,p(k|k-1)}{c^2 p(k|k-1) + \sigma_v^2} \qquad \text{.............(5.29)}$$

From eqn.(5.29) it is obvious that we can calculate $\beta(k)$ from the previous mean-square prediction error and then eqn.((5.28) gives us the mean-square prediction error for $p(k+1|k)$.

Again, the optimum processor consists of simply multiplying the previous estimate by a, and then adding a weighted correction term. This correction term consists of the difference between the new data sample y(k) and the previous prediction estimate $c\hat{x}(k|k-1)$ directly. In the filtering problem considered in section 5.3, the correction term involved y(k) minus a times the previous estimate.

If we assume that the random driving force in eqn.(5.1) is zero, the signal evolves according to the equation x(k)=ax(k-1). Therefore, given an estimate $\hat{x}(k)$ at time k, it seems reasonable to predict the estimate at time k+1 as

$$\hat{x}(k+1|k) = a\hat{x}(k) \qquad \text{.........(5.30)}$$

Using the equations 5.13 and 5.30 we can established the following relationship

$$\beta(k) = a\,b(k) \qquad \text{.............(5.31)}$$

i.e. the two gain factors b(k) and $\beta(k)$ are actually related by the parameter a [2].

The optimum one-step predictor is shown in fig.(5.4), and optimum filtering and prediction simultaneously are shown in fig.(5.5).

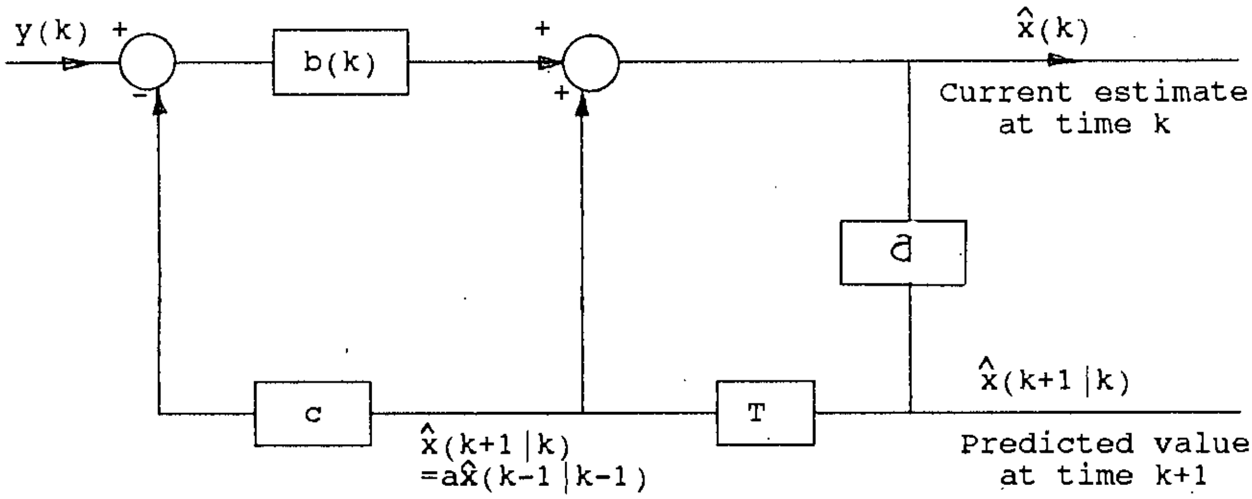

Fig. 5.5 Filtering and prediction simultaneously

5.5 Summary

In this chapter we have developed the scalar Kalman filter. The Kalman filter can be thought of as an estimator that produces three types of outputs given a noisy measurement sequence and the associated models:

(1) as a state estimator or reconstructor; i.e. it reconstructs estimates of the state x(k) from noisy measurements y(k);

(2) as a measurement filter that on input accepts the noisy sequence y(k), and on the output produces a filtered measurement sequence $\hat{x}(k)$;

(3) as a whitening filter that accepts noisy correlated measurements y(k) and produces uncorrelated or white equivalent measurements.

This chapter completes the various derivations of the Kalman filter. We summarized the filtering and prediction equations for the purpose of direct transformation into vector equations. In the next chapter we shall readily extend this results to include the estimation of time-varying random signals by using vector notation, that will enable us to consider simultaneous recursive estimation of several signals.

CHAPTER 6

OPTIMUM ESTIMATION OF VECTOR SIGNALS

6.0 Introduction

We have dealt so far with scalar random signals generated by a first-order autoregressive process. In chapter 5 we have optimized the first-order recursive filter and have got a set of equations referred to as the scalar Kalman filter. But the signal processing problems require simultaneous processing of several different signals and this chapter concern with such vector or multidimensional signals. It is shown in section 6.1 how vector equations are formulated in the case of simultaneous estimation of a number of signals. In section 6.2, the estimation problem for multidimensional systems is formulated in terms of vectors and matrices. Since there is an equivalence between scalar and matrix operations, all results in chapter 5 for scalar signals are transformed into vector and matrix equations in sections 6.3 and 6.4 [1-4,10,14,20,22].

6.1 Signal Vector

With the extension of the signal model to vector signals to include simultaneous estimation of several signals, we are now in a position to extend the scalar Kalman filter for the estimation of multidimensional signals. It is shown below that these multidimensional signals are conveniently represented by vector notation. In place of simple gain parameters we then have matrix operations on vector [1-3].

To demonstrate the formation of vector equations we consider that we have q independent signals to be estimated or predicted simultaneously. We denote samples of these signals at time k, as $x_1(k), x_2(k), \ldots, x_q(k)$. Again we assume that each one is generated by its own first-order autoregressive process. The jth signal is then formed according to the equation

$$x_j(k) = a_j x_j(k-1) + w_j(k-1) \qquad \text{..........(6.1)}$$

where j=1,2,...,q

Here each of the w_j processes is assumed to be white, zero-mean and independent of all others. We can define q-dimensional vectors made up of the q signals and q white noise driving processes as

$$\mathbf{x}(k) = \begin{bmatrix} x_1(k) \\ x_2(k) \\ \cdot \\ \cdot \\ \cdot \\ x_q(k) \end{bmatrix} \quad \text{and} \quad \mathbf{w}(k) = \begin{bmatrix} w_1(k) \\ w_2(k) \\ \cdot \\ \cdot \\ \cdot \\ w_q(k) \end{bmatrix} \qquad \text{..........(6.2)}$$

The q eqns.(6.1) can be written as the first-order vector equation (target state equation) in terms of these defined vectors such that

$$\mathbf{x}(k) = \mathbf{A}\mathbf{x}(k-1) + \mathbf{w}(k-1) \qquad \text{..........(6.3)}$$

where **x**(k), **x**(k-1) and **w**(k-1) are (qx1) column vectors and A is the state transition (qxq) matrix. In this case A matrix is diagonal and given by

$$\mathbf{A} = \begin{bmatrix} a_1 & 0 & \cdot & \cdot & 0 \\ 0 & a_2 & \cdot & \cdot & 0 \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ 0 & \cdot & \cdot & \cdot & a_q \end{bmatrix} \qquad \text{..........(6.4)}$$

6.2 Data Vector

The Kalman filter, we consider to discuss require a model for the measurement process of the vector signal in addition to the model for the generation of the signal which is mentioned in the previous section. Assuming that in estimating the signal vector $\mathbf{x}(k)$ we made r simultaneous noisy measurements at time k. These measurement samples are labelled $y_1(k), y_2(k), \ldots, y_r(k)$. So we have the following set of data.

$$
\begin{aligned}
y_1(k) &= c_1 x_1(k) + v_1(k) \\
y_2(k) &= c_2 x_2(k) + v_2(k) \\
&\;\;\vdots \\
y_r(k) &= c_r x_r(k) + v_r(k)
\end{aligned}
\qquad \text{.........(6.5)}
$$

where $v_j(k)$ terms represent additive noise and $c_1, \ldots, c_r$ are some measurement parameters which are similar to c introduced in eqn.(5.4). This set of equations can be put into vector form by defining r-component vectors $\mathbf{y}(k)$ and $\mathbf{v}(k)$. In terms of the previously defined q-component signal vector $\mathbf{x}(k)$, we then have the data vector (measurement equation):

$$\mathbf{y}(k) = \mathbf{C}\mathbf{x}(k) + \mathbf{v}(k) \qquad \text{..........(6.6)}$$

where y(k), x(k) and v(k) are (rx1) column vectors, and **C** is an (rxq) observation matrix. In this case **C** is (for r<q) given by

$$
\mathbf{C} = \overset{q \;\longrightarrow}{\underset{r \downarrow}{\begin{bmatrix} c_1 & 0 & \cdot & \cdot & 0 & \cdot & \cdot & 0 \\ 0 & c_2 & & & \cdot & & & \cdot \\ \cdot & & \cdot & & \cdot & & & \cdot \\ \cdot & & & & \cdot & & & \cdot \\ 0 & & & & c_r & & & 0 \end{bmatrix}}} \qquad \text{...........(6.7)}
$$

The block diagram for the system and measurements in vector forms is the same as for the scalar case, figs.(5.1) and (5.2) respectively. But only the notation changes to vectors, and the system and observation parameters become matrices as represented in the, figs.(6.1) and (6.2).

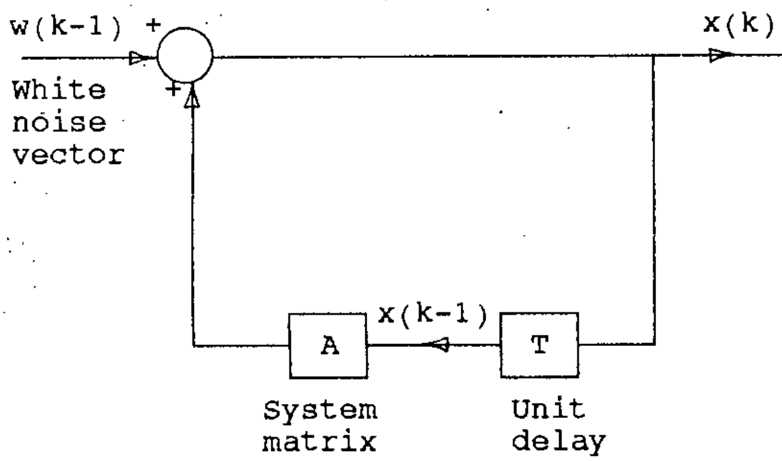

Fig. 6.1 System model

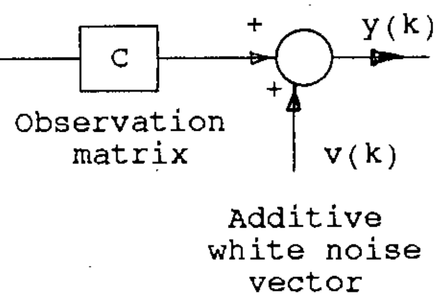

Fig. 6.2 Measurement model

6.3 Vector Problem Formulation

Here we look for the simultaneous estimates (filtered or predicted) of q-signal components. We have from the discussion in section 6.1, a signal vector $\mathbf{x}(k)$ obeying a known first-order vector dynamical equation

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{w}(k) \qquad \text{..........(6.8)}$$

which is to be extracted from a noisy measurement vector $\mathbf{y}(k)$

$$\mathbf{y}(k) = \mathbf{C}\mathbf{x}(k) + \mathbf{v}(k) \qquad \text{..........(6.9)}$$

The problem is how to form $\hat{\mathbf{x}}(k)$, the 'best' linear estimate (filtered value) of $\mathbf{x}(k)$ and its 'best' predicted value $\hat{\mathbf{x}}(k|k\text{-}1)$. By 'best' we now mean estimators that minimize the mean-square error of each signal component simultaneously. In the filtering operation each mean-square error

$$E[x_j(k)-\hat{x}_j(k)]^2 \quad \text{where } j=1,2,\ldots,q \qquad \text{........(6.10)}$$

is to be minimized. For the multidimensional signal, it is required to reformulate all the scalar Kalman equations into vector form and apply a matrix minimization procedure to obtain the optimum solutions. Here we shall not carry out the derivations of the optimum estimators. Since we already have the solutions for the one-dimensional (scalar) cases we can extend them to the multidimensional systems, using the equivalence of scalar and matrix operations given in Table 6.1 in which the subscript T stands for transpose of a matrix, and -1 for the inverse of a matrix.

Table 6.1 Transformation of scalar to matrix

Scalar →	**Matrix**
a+b	A+B
ab	AB
1/(a+b)	$(A+B)^{-1}$
a^2b	ABA^T

We have already seen that in transition from the single signal to vector signal, the system parameter a changed into the system matrix A, and the data coefficient c changed into the observation matrix C. We now consider the transition of other relevant quantities.

The transition from the observation noise variance to the observation noise covariance matrix (common variance of a number of signals) is written as

$$\sigma_v^2 = \sigma_{v1,1}^2 = E[v_1^2(k)] \rightarrow R(k) = E[v(k)v^T(k)] \qquad \text{....(6.11)}$$

Similarly, for the system noise, we have

$$\sigma_w^2 = \sigma_{w1,1}^2 = E[w_1^2(k)] \rightarrow Q(k) = E[w(k)w^T(k)] \qquad \text{....(6.12)}$$

where Q(k) represents the system noise covariance matrix. If there is no correlation between noise processes, the off-diagonal terms are zero.

The transition from the mean-square error for the single signal to the error covariance matrix is written as

$$p(k) = p_{1,1}(k) = E[e_1^2(k)] \rightarrow P(k) = [e(k)e^T(k)] \qquad \text{....(6.13)}$$

The diagonal terms of the error covariance matrix are the individual mean-square errors as formulated by eqn.(6.10).

6.4 Vector Kalman Filter

We now transform the scalar Kalman filter algorithm, given by equations 5.18 to 5.21, into the corresponding vector Kalman filter. With reference to these equations and the discussions in previous sections we can write directly the vectors and matrix equations in the following tabulated form:

Estimator:

$$\hat{\mathbf{x}}(k) = \mathbf{A}\hat{\mathbf{x}}(k-1) + \mathbf{K}(k)[\mathbf{y}(k) - \mathbf{CA}\hat{\mathbf{x}}(k-1)] \qquad \text{.....(6.14)}$$

Filter gain:

$$\mathbf{K}(k) = \mathbf{P}_1(k)\mathbf{C}^T[\mathbf{CP}_1(k)\mathbf{C}^T + \mathbf{R}(k)]^{-1} \qquad \text{.....(6.15)}$$

$$\text{where } \mathbf{P}_1(k) = \mathbf{AP}(k-1)\mathbf{A}^T + \mathbf{Q}(k-1) \qquad \text{.....(6.16)}$$

Error covariance matrix:

$$P(k) = \mathbf{P}_1(k) - \mathbf{K}(k)\mathbf{CP}_1(k) \qquad \text{.....(6.17)}$$

The above set of equations constitute the **vector Kalman filter** for the model described by the state eqn.(6.8) and measurement eqn.(6.9). Here we have used K(k) in place of b(k), because this is a commonly used notation for the gain matrix in the Kalman filter. Other quantities have been arranged, within the rules of the scalar-matrix equivalence Table 6.1, to obtain the standard form of Kalman equations used. In eqn.(6.16), we have used Q(k-1) since σ_w^2 in fact represents $E[w^2(k-1)]$. For time-varying systems and time-varying observations, matrices A and C are obviously functions of time.

One of the most significant features of the Kalman filter is its recursive form property that makes it extremely useful in processing measurements to obtain the optimum estimate, utilizing a digital computer. The processing of signal by using Kalman filter, it is not necessary to store any measurement data. The measurement will be processed automatically as the new data is available. But it is necessary to store $\hat{x}(k-1)$ in proceeding from time (k-1) to time k. Moreover, the physical model must be defined (A,C), and the statistics of the random processes must be known (Q,R).

The block diagram, fig.(6.3), is the representation of eqn.(6.14). The information flow in the filter can be explained by this figure. If we assume that $\hat{x}(k-1)$ is known for some k and we want to determine $\hat{x}(k)$, given y(k). The computational cycle would proceed as follows:

(1) the estimate $\hat{x}(k-1)$ is premultiplied by the system matrix A and propagated forward, denoted as $\hat{x}'(k)$;

(2) $\hat{x}'(k)$ is premultiplied by the observation matrix C giving $\hat{y}(k)$ which is subtracted from the actual measurement y(k) to obtain the error e(k);

(3) e(k) is premultiplied by the gain matrix K(k) and the result is added to $\hat{x}'(k)$ to give $\hat{x}(k)$;

(4) $\hat{x}(k)$ is stored until the time of the next measurement, when the cycle is repeated.

The filter operates in a 'predict-correct' fashion, i.e. the correction term K(k)e(k) is added to the predicted estimate $\hat{x}'(k)$ to determine the filtered estimate. The correction term involves the Kalman gain matrix K(k).

It is apparent that the filter shown in fig.(6.3) consists of the model of the dynamic process, which performs the function of prediction, and a feedback correction scheme in which the product of Kalman gain and the error term is applied to the model as a forcing function that brings the estimate to the steady state condition after some iteration. Since the gain matrix K(k) does not depend at all on the measurements, it can be calculated before the estimation is carried out. This approach requires storing the calculated vectors for each recursion and feeding them out as needed. Equations 6.15 to 6.17 define the algorithm for the recursive computation of the filter gain matrix K(k). At the same time, we obtain values for $P_1(k)$ and P(k), i.e. the variances of the components of the prediction and filtering errors respectively. It is obvious that the gain term is updated recursively as the estimation proceeds and hence there is no need to store all gain values, i.e. the previous value is the only one required.

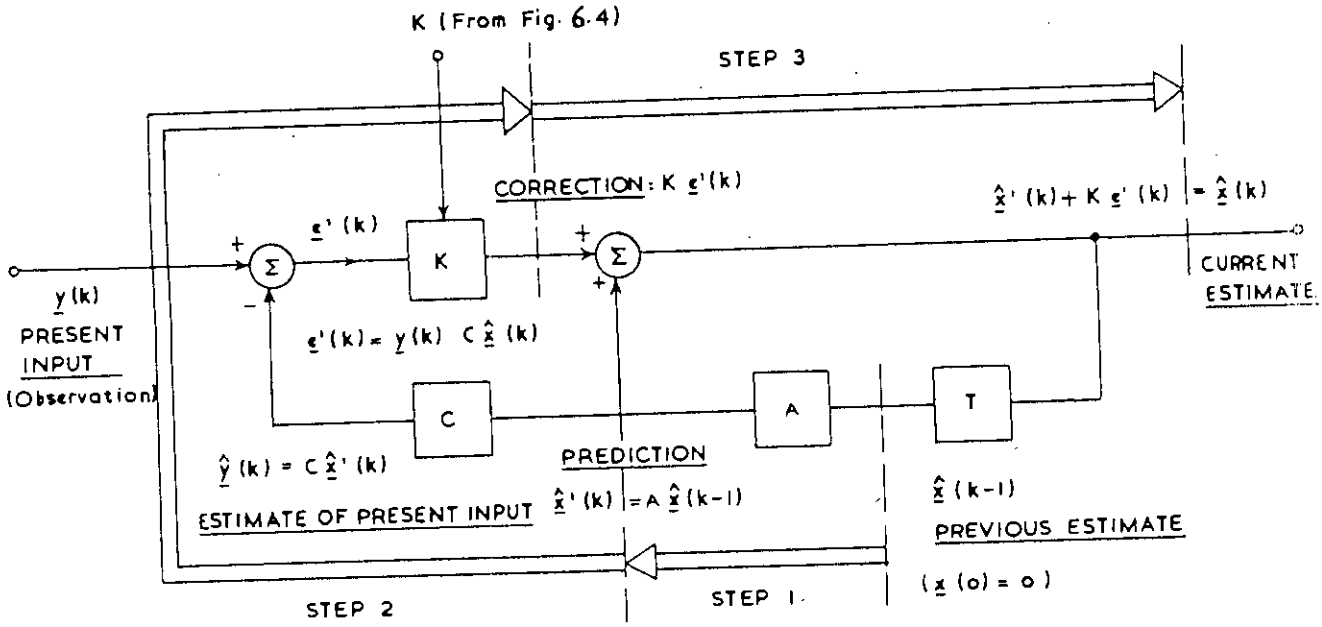

Fig. 6.3 Computational steps in Kalman filter

Fig.(6.4) represents the subroutine computational diagram and the computational cycle would proceed as follows:

(1) $P_1(k)$ is computed from the given value P(k-1), Q(k-1), A using eqn.(6.16);

(2) K(k) is obtained by substituting $P_1(k)$, R(k) and C into eqn.(6.15) which is used in step three of the filter computation;

(3) P(k) is determined by substituting $P_1(k)$, K(k) and C into eqn.(6.17) which is stored until the time of the next measurement, when the cycle is repeated.

For better understanding we can arrange the Kalman filter equations 6.14 to 6.17 as

(1) Prediction

$$\hat{x}(k|k-1) = A\hat{x}(k-1|k-1) \quad \text{......(6.18)}$$

$$P(k|k-1) = AP(k-1|k-1)A^T + Q(k-1) \quad \text{......(6.19)}$$

(2) Updating (or correction)

$$\hat{x}(k|k) = \hat{x}(k|k-1) + K(k)[y(k)-C\hat{x}(k|k-1)] \quad \text{......(6.20)}$$

$$P(k|k) = P(k|k-1) - K(k)CP(k|k-1) \quad \text{......(6.21)}$$

$$K(k) = P(k|k-1)C^T[CP(k|k-1)C^T + R(k)]^{-1} \quad \text{......(6.22)}$$

where P(k|k-1) and P(k|k) correspond to $P_1(k)$ and P(k) respectively in the Kalman filter equations 6.15 to 6.17. The first stage is prediction based on the state eqn.(6.8), and the second stage is the updating or correction based on the measurement eqn.(6.9). Both stages are illustrated in fig.(6.5).

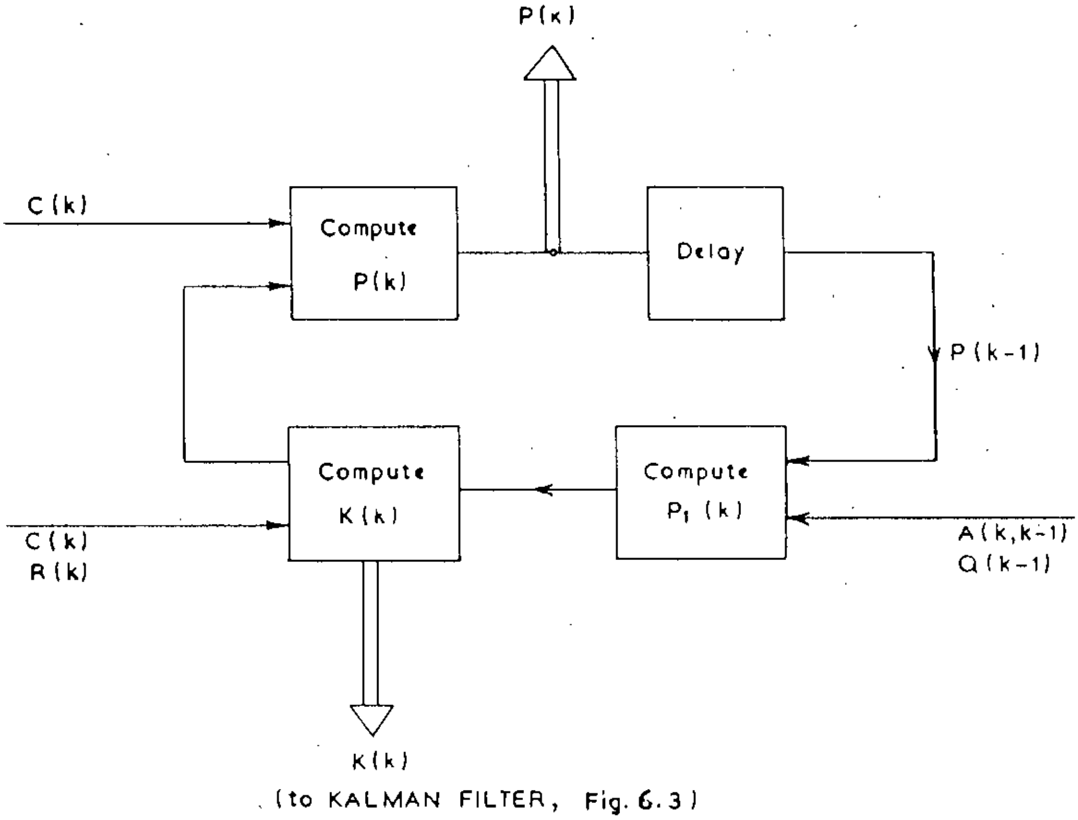

Fig. 6.4 Subroutine calculations for Kalman filter

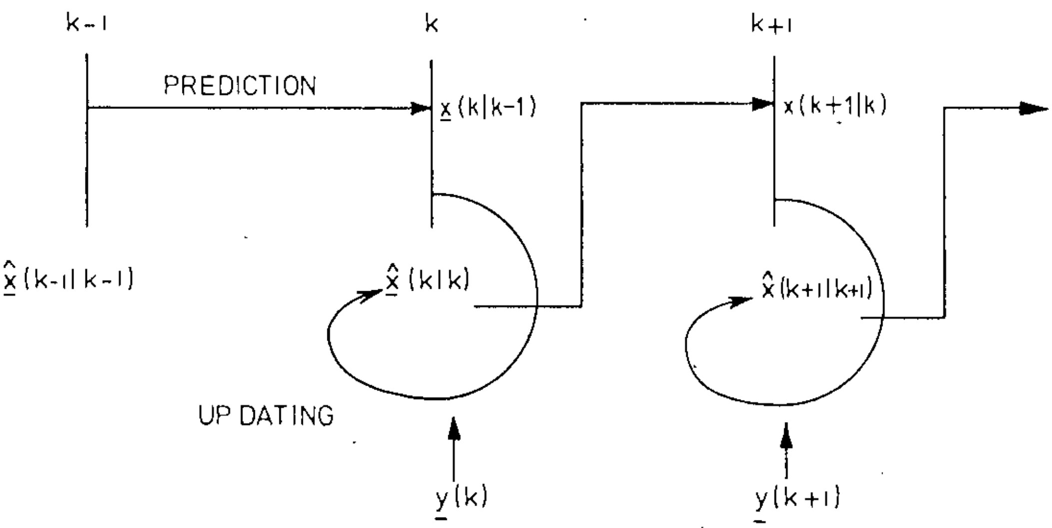

Fig. 6.5 Illustration for two stages Kalman computational cycle

6.5 Vector Kalman Predictor

We are now in a position to transform the scalar Kalman predictor algorithm, given by equations 5.32 to 5.34, into the corresponding vector Kalman predictor. With reference to these equations and the discussions in sections 6.1 to 6.4 we obtain the following vector and matrix set of equations.

Predictor equation:

$$\hat{\mathbf{x}}(k+1|k) = \mathbf{A}\hat{\mathbf{x}}(k|k-1) + \mathbf{G}(k)[\mathbf{y}(k)-\mathbf{C}\hat{\mathbf{x}}(k|k-1)] \qquad ....(6.23)$$

Predictor gain:

$$\mathbf{G}(k) = \mathbf{A}\mathbf{P}(k|k-1)\mathbf{C}^T[\mathbf{C}\mathbf{P}(k|k-1)\mathbf{C}^T + \mathbf{R}(k)]^{-1} \qquad ....(6.24)$$

Prediction mean-square error:

$$\mathbf{P}(k+1|k) = [\mathbf{A}-\mathbf{G}(k)\mathbf{C}]\mathbf{P}(k|k-1)\mathbf{A}^T + \mathbf{Q}(k) \qquad ....(6.25)$$

These equations constitute the **vector Kalman predictor** for the model described earlier by the state equations 6.8 and 6.9. We have introduced here G(k) as the predictor gain matrix in place of the previous time-varying scalar gain $\beta(k)$ in eqn.(5.32). Other quantities are the same as used in the previous section.

A relationship correspond to the scalar relationship $\beta(k)=ab(k)$ mentioned in section 5.4, eqn.(5.31) can be used directly as

$$G(k) = A\ K(k) \qquad ....(6.26)$$

We would calculate predictor gain matrix G(k) from eqn.(6.26), since we have found P(k|k-1) and hence K(k) from vector Kalman filter equations 6.19, 6.21 and 6.22 without using eqn.(6.24). Because it is more straight forward and less iterations are required in this case and hence, in developing computer program, it was actually done.

6.6 Summary

Here the first-order Kalman filter equations specified in chapter 5, are modified to take higher-order filters into account by replacing the scalars by vectors. Filtering, prediction and smoothing problems have been introduced in this chapter in the framework of estimation theory as applied to time-varying random variables. The theory developed in chapters 5 and 6 are of the same form and will be extensively applied in the next chapter, where the tracking functions are presented. The main results pertaining to the Kalman filter are :

(1) the filter has a feedback recursive structure that embedding the state model of the process,

(2) a central role is played by the Kalman gain, whose value depends on the model parameters and the prediction covariance,

(3) both the filter error covariance and the Kalman gain can be computed independently of the measurements,

(4) the unpredictable component of the measurements is a zero-mean white process,

(5) the convergence of the filter to a steady state solution is conditioned to the structural properties of the model (controllability, observability).

CHAPTER 7

KALMAN FILTER FORMULATION FOR RADAR TRACKING

7.0 Introduction

In the previous chapters we have developed the Kalman filter, a signal estimation technique. Now we show how one might apply the recursive techniques to the tracking function of an air traffic control radar. In chapter 2 we discussed the detection function of radars and in general discussion we point out that as the radars rotate, they continuously send out high-frequency pulses of electromagnetic energy into space. Pulses intercepting an aircraft in space are reflected back to the radar. The return pulses with noise mixed in must then be processed to show the presence of the aircraft. The time delay between transmission and reception of the pulses provides an estimate of aircraft range, while the location of the antenna beam at the time of detection provides the aircraft bearing (azimuth). A short range radar rotates typically at a scan rate of 15 rpm, while longer range radars rotate at 6 rpm. Therefore, we have for these cases new range and bearing estimates every 4 sec and every 10 sec respectively. This means that tracking filters are updated at such an interval denoted by T.

7.1 Radar Signal Processing

We consider a radar tracking problem that a vehicle being tracked is at range $R+\rho(k)$ at time k, and at range $R+\rho(k+1)$ at time k+1, T seconds later. We use T to represent the spacing between samples made one scan apart. The average range is denoted by R, and $\rho(k)$, $\rho(k+1)$ represent deviations from the average. We are interested in estimating these deviations, which are assumed to be statistically independent. To a first approximation, if the vehicle is travelling at a radial velocity $\dot{\rho}(k)$ and T is not too large,

$$\rho(k+1) = \rho(k) + T\dot{\rho}(k) \qquad \text{.....................(7.1)}$$

which is the range equation.

Similarly, if we consider the acceleration u(k) we have

$$T\,u(k) = \dot{\rho}(k+1) - \dot{\rho}(k) \qquad \text{.....................(7.2)}$$

which is the acceleration equation. Assuming that u(k) is zero-mean and uncorrelated from interval to interval, i.e. $E[u(k+1)u(k)] = 0$, but it has known variance $E[u^2(k)] = \sigma_u^2$. Such accelerations might be caused by sudden wind gusts or short-term irregularities in engine thrust. The quantity $u_1(k) = Tu(k)$ is also a white noise process, and we have in place of eqn.(7.2) the following one

$$\dot{\rho}(k+1) = \dot{\rho}(k) + u_1(k) \qquad \text{.....................(7.3)}$$

For the radar tracking problem, the signals to be estimated are the range $\rho(k)$, the radial velocity $\dot{\rho}(k)$, the bearing (azimuth) $\theta(k)$, and the angular velocity $\dot{\theta}(k)$. So **x**(k) in a two-dimensional plane is the four component vector

$$\mathbf{x}(k) = \begin{bmatrix} \rho(k) \\ \dot{\rho}(k) \\ \theta(k) \\ \dot{\theta}(k) \end{bmatrix} \qquad \text{...........(7.4)}$$

however, measurements of only range and bearing are made in the presence of additive noise $v_1(k)$ and $v_2(k)$ respectively. Here for the matrix **C** of eqn.(6.7), we have q=4 and r=2, and using eqn.(7.1), the velocities are found in terms of these quantities. The matrix **C** is given here by

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \qquad \text{...........(7.5)}$$

Now we have a four component signal vector **x**(k) with states $x_1(k)=\rho(k)$ for the range, $x_2(k)=\dot{\rho}(k)$ for the radial velocity, $x_3(k)=\theta(k)$ for the bearing, and $x_4(k)=\dot{\theta}(k)$ for the bearing rate (or angular velocity). The noise terms $u_1(k)$ and $u_2(k)$ represent the change in radial velocity and bearing rate respectively over interval T. Again, $u_1(k)$ and $u_2(k)$ are random with zero-mean and uncorrelated both with each other and individually from one interval to other. Therefore, we have the following system equation:

$$\underbrace{\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \\ x_4(k+1) \end{bmatrix}}_{\mathbf{x(k+1)}} = \underbrace{\begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \end{bmatrix}}_{\mathbf{x(k)}} + \underbrace{\begin{bmatrix} 0 \\ u_1(k) \\ 0 \\ u_2(k) \end{bmatrix}}_{\mathbf{w(k)}} \qquad \ldots(7.6)$$

Since the radar sensors provide the noisy estimates of the range $\rho(k)=x_1(k)$, and bearing $\theta(k)=x_3(k)$ at time intervals T. At time k, the two sensor outputs are

$$y_1(k) = x_1(k)+v_1(k)$$

$$y_2(k) = x_3(k)+v_2(k)$$

Therefore, the data vector can be written as

$$\underbrace{\begin{bmatrix} y_1(k) \\ y_2(k) \end{bmatrix}}_{\mathbf{y(k)}} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\mathbf{C}} \underbrace{\begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ x_4(k) \end{bmatrix}}_{\mathbf{x(k)}} + \underbrace{\begin{bmatrix} v_1(k) \\ v_2(k) \end{bmatrix}}_{\mathbf{v(k)}} \qquad \ldots.(7.7)$$

According to our assumption the additive noise, v(k), is to be gaussian with zero-mean and variances $\sigma_{\rho^2}(k)$ and $\sigma_{\theta^2}(k)$. Upto this we have established vector equations for the system model given by eqn.(7.6), and data model given by eqn.(7.7). Now we have to formulate noise covariance matrices **Q** for the system, and **R** for the measurement model. For the system noise covariance matrix, **Q** defined in eqn.(6.12), we have

$$\mathbf{Q}(k) = E[\mathbf{w}(k)\mathbf{w}^T(k)] = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \sigma_1^2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_2^2 \end{bmatrix} \qquad \ldots\ldots(7.8)$$

where $\sigma_1^2 = E(u_1^2)$ and $\sigma_2^2 = E(u_2^2)$ are the variances of T times the radial and angular acceleration respectively.

And for the observation noise covariance matrix, **R** using eqn.(6.11), we have

$$\mathbf{R}(k) = E[\mathbf{v}(k)\mathbf{v}^T(k)] = \begin{bmatrix} \sigma_\rho^2(k) & 0 \\ 0 & \sigma_\theta^2(k) \end{bmatrix} \qquad \ldots\ldots\ldots(7.9)$$

Specific values must be substituted for the variances σ_1^2 and σ_2^2, and also for $\sigma_\rho^2(k)$ and $\sigma_\theta^2(k)$ in order to define the Kalman filter numerically. To simplify, we assume that the probability density function (p.d.f.) of the acceleration in either direction (ρ or θ) to be of the form of fig.(7.1).

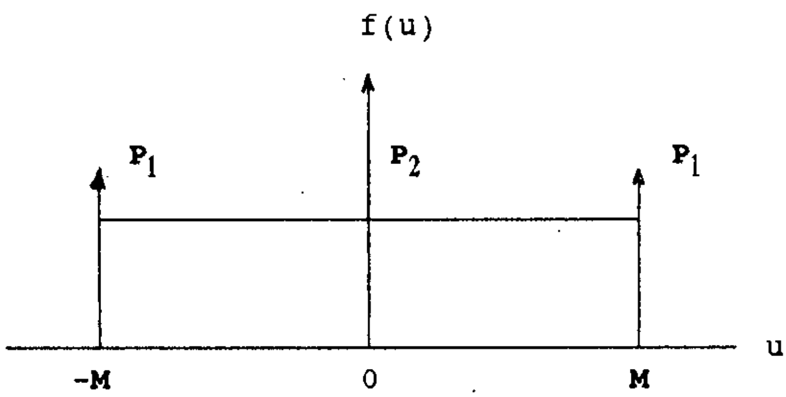

Fig.7.1 Typical probability distribution of u(k)

Three impulse functions representing discrete probabilities at ±M and 0 acceleration have been superimposed to represent that there is a probability P_2 that the aircraft will proceed at constant radial and angular velocities, while there is a probability P_1 that its acceleration (deceleration) in either direction is at the maximum value M. Now we can express the height of the uniform distribution as $P(u) = (1-2P_1-P_2)/2M$ and the variance of the random variable u is given by

$$\sigma_u^2 = \frac{M^2}{3}(1+4P_1-P_2) \qquad \text{.....................(7.10)}$$

The variances in eqn.(7.8) are then given by

$$\sigma_1^2 = T^2\sigma_u^2 \qquad \text{.....................(7.11)}$$

and

$$\sigma_2^2 = \frac{\sigma_1^2}{R^2} \qquad \text{.....................(7.12)}$$

The parameter σ_u represents the linear acceleration. Therefore, σ_1 is related to the change in linear velocity and σ_2 is related to the change in angular velocity [1,2,10].

7.2 Initialization of Kalman Filter

To start Kalman processing we have to initialize the gain matrix K(k). To do this the error covariance matrix P(k) has to be specified in some way. A reasonable initialization can be established using two measurements, range and bearing, at times k=1 and k=2. The first set of these two measurements, x(1) (k=1), can be used to estimate range and bearing but not range rate or bearing rate (unless we arbitrarily take these zero). The second pair of measurements x(2) (k=2) provide the additional two numbers required to make independent estimates of the four parameters of eqn.(7.4). Therefore the filter start only after obtaining the first two measurements x(1) and x(2). Using these, we can produce the following four-component signal vector estimates $\hat{\mathbf{x}}(2)$:

$$\hat{\mathbf{x}}(2) = \begin{bmatrix} \hat{x}_1(2) = \hat{\rho}(2) = y_1(2) \\ \hat{x}_2(2) = \hat{\dot{\rho}}(2) = 1/T[y_1(2)-y_1(1)] \\ \hat{x}_3(2) = \hat{\theta}(2) = y_2(2) \\ \hat{x}_4(2) = \hat{\dot{\theta}}(2) = 1/T[y_2(2)-y_2(1)] \end{bmatrix} \qquad \text{.....(7.13)}$$

We can use the general expression for P(k), eqn.(6.13) to calculate P(2) (for k=2), and therefore

$$P(2) = E\{[x(2)-\hat{x}(2)][x(2)-\hat{x}(2)]^T\} \qquad ......(7.14)$$

We obtain $\hat{x}(2)$ from eqn.(7.13), and using equations 7.6 and 7.7 for x(2), we have the following results:

$$\mathbf{x}(2)-\hat{\mathbf{x}}(2) = \begin{bmatrix} -v_1(2) \\ u_1(1) \; -(v_1(2)-v_1(1))/T \\ -v_2(2) \\ u_2(1) \; -(v_2(2)-v_2(1))/T \end{bmatrix} \qquad ......(7.15)$$

which is a (4x1) column vector.

Specifically, in this case P(2) is a (4x4) matrix given by

$$\mathbf{P}(2) = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \\ p_{41} & p_{42} & p_{43} & p_{44} \end{bmatrix} \qquad ......(7.16)$$

Taking into account the independence of noise sources u and v, and also the independence between individual noise samples, we obtain the following simplified form of P(2)

$$\mathbf{P}(2) = \begin{bmatrix} p_{11} & p_{12} & 0 & 0 \\ p_{21} & p_{22} & 0 & 0 \\ 0 & 0 & p_{33} & p_{34} \\ 0 & 0 & p_{43} & p_{44} \end{bmatrix} \qquad ....(7.17)$$

where $p_{11} = \sigma_\rho^2$

$p_{12} = p_{21} = \sigma_\rho^2/T$

$p_{22} = 2\sigma_\rho^2/T^2+\sigma_1^2$

$p_{33} = \sigma_\theta^2$ ....(7.18)

$p_{34} = p_{43} = \sigma_\theta^2/T$

$p_{44} = 2\sigma_\theta^2/T^2+\sigma_2^2$

We now present a numerical example that explain clearly how to initialize the Kalman filter. We take for range R=160 km, scan time T=15 s, and maximum acceleration M=3.27 m/s^2. The two parameters P_1 and P_2, the probabilities of maximum and zero acceleration are 0.1 and 0.3 respectively. Let the r.m.s. noise σ_ρ in the range sensor be equivalent to 10^3 m and the r.m.s. noise σ_θ in the bearing sensor be 0.017 rad. These two figures define numerically the noise covariance matrix R. From these numerical values we can calculate noise variances in the Q matrix, eqn.(7.8), as σ_1^2= 882.1643 and σ_2^2= 3.446×10^{-8}. Therefore the initial value of the estimation covariance matrix P(2), or in the alternative notation P(2|2), can be obtained from eqn.(7.17) as

$$\mathbf{P}(2|2) = \begin{bmatrix} 10^6 & 6.67\times10^4 & 0 & 0 \\ 6.67\times10^4 & 9.0\times10^3 & 0 & 0 \\ 0 & 0 & 2.890\times10^{-4} & 1.927\times10^{-5} \\ 0 & 0 & 1.927\times10^{-5} & 2.603\times10^{-6} \end{bmatrix} \qquad (7.19)$$

Since it is error matrix at k=2 , we could try to use it to calculate the predictor gain G(3) at k=3, which is given by

$$G(3) = AP(3|2)C^T[CP(3|2)C^T+R]^{-1} \qquad \text{........(7.20)}$$

where all quantities (A,C,R) are known except P(3|2). We might try to calculate P(3|2) using eqn.(6.25) as

$$P(3|2) = [A-G(2)C]P(2|1)A^T+Q$$

but G(2) and P(2|1) are not known. Since eqn.(6.25) has been derived from eqns.(6.16) and (6.17), we can use eqn.(6.16) instead of eqn.(6.25) as

$$P(k|k-1) = AP(k-1|k-1)A^T+Q \qquad \text{........(7.21)}$$

For k=3 the above equation becomes

$$P(3|2) = AP(2|2)A^T+Q \qquad \text{........(7.22)}$$

where P(2|2) is known from eqn.(7.19). Using the given numerical values we obtain

$$P(3|2) = \begin{bmatrix} 5.2\text{x}10^{6} & 2.13\text{x}10^{5} & 0 & 0 \\ 2.13\text{x}10^{5} & 1.065\text{x}10^{4} & 0 & 0 \\ 0 & 0 & 1.453\text{x}10^{-3} & 5.832\text{x}10^{-5} \\ 0 & 0 & 5.832\text{x}10^{-5} & 2.638\text{x}10^{-6} \end{bmatrix} \qquad \text{...(7.23)}$$

The diagonal elements represent the prediction errors. The first and third diagonal elements are mean-square range and bearing prediction errors respectively for k=3.

Using eqn.(7.23) in eqn.(7.20) we can calculate the predictor gain G(3) and the result is

$$G(3) = \begin{bmatrix} 1.355 & 0 \\ 3.44\text{x}10^{-2} & 0 \\ 0 & 1.336 \\ 0 & 3.35\text{x}10^{-2} \end{bmatrix} \qquad \text{.....(7.24)}$$

Now we have to determine P(3|3) using eqn.(6.17) which gives (for k=3)

$$P(3|3) = P(3|2)-K(3)CP(3|2)$$

where $K(3)=A^{-1}G(3)$

The process is now repeated by finding P(4|3), G(4) etc. [1-4].

7.3 Results of Computer Calculations

Results of computer calculations for the data obtained from the radar of **Zia International Airport** have been introduced in the following pages. In addition, a set of curves have been produced for different values of the parameters used in the model as the characteristics of the processor. The resulting performance figures, that are presented here, are the mean-square range prediction error, mean-square bearing prediction error and gain setting of Kalman filter.

Given, T = 4.000 sec, M = 3.270 m/s^2
P_1 = 0.100 , P_2 = 0.300
R = 160.000 km , σ_ρ = 1000.000 m
σ_θ = 0.017 rad

Step no.	Kalman gain	Mean-square range prediction errror	Mean square bearing prediction error
k	$K_{11}(k)$	$P_{11}(k+1\|k)$	$P_{33}(k+1\|k)$
3	0.833361	2334783.057548	0.000674
4	0.700130	1502257.761495	0.000434
5	0.600361	1103420.073580	0.000318
6	0.524584	871645.459881	0.000251
7	0.465711	721267.772347	0.000207
8	0.419033	616619.282450	0.000176
9	0.381425	540281.207764	0.000153
10	0.350768	482763.431677	0.000136
11	0.325584	438461.296653	0.000122
12	0.304813	403849.063034	0.000110
13	0.287673	376589.344683	0.000101
14	0.273567	355059.346466	0.000094
15	0.262025	338083.377154	0.000087
16	0.252662	324774.862234	0.000082
17	0.245155	314440.106693	0.000077
18	0.239220	306518.722382	0.000073
19	0.234607	300546.699869	0.000069
20	0.231093	296133.755819	0.000066
21	0.228475	292949.616994	0.000063
22	0.226575	290715.651157	0.000061
23	0.225236	289199.384648	0.000059
24	0.224325	288210.270751	0.000057
25	0.223729	287595.721492	0.000056
26	0.223359	287236.924260	0.000054
27	0.223142	287044.334651	0.000053
28	0.223026	286952.968656	0.000052
29	0.222971	286917.725347	0.000052
30	0.222950	286908.983994	0.000051
31	0.222944	286908.672264	0.000050
32	0.222944	286906.927607	0.000050
33	0.222943	286899.397291	0.000049
34	0.222938	286885.158896	0.000049
35	0.222930	286865.198823	0.000049
36	0.222918	286841.361794	0.000048
37	0.222903	286815.675952	0.000048
38	0.222888	286789.961572	0.000048
39	0.222872	286765.641906	0.000048
40	0.222858	286743.688716	0.000048
41	0.222844	286724.649843	0.000048
42	0.222833	286708.719978	0.000048
43	0.222823	286695.827788	0.000048
44	0.222816	286685.722185	0.000048
45	0.222809	286678.047934	0.000048

Given, T = 4.000 sec, M = 3.270 m/s^2
P_1 = 0.100 , P_2 = 0.300
R = 370.000 km , σ_ρ = 1200.000 m
σ_θ = 0.017 rad

Step no.	Kalman gain $K_{11}(k)$	Mean-square range prediction error $P_{11}(k+1\|k)$	Mean-square bearing prediction error $P_{33}(k+1\|k)$
3	0.833353	3361449.747010	0.000674
4	0.700091	2162257.938346	0.000434
5	0.600251	1587420.857255	0.000318
6	0.524348	1252981.390462	0.000251
7	0.465277	1035560.615671	0.000206
8	0.418314	883779.226951	0.000176
9	0. 38032	772540.351712	0.000153
10	0.349164	688170.202824	0.000135
11	0.323362	622597.712344	0.000121
12	0.301851	570754.856727	0.000110
13	0.283851	529295.650584	0.000100
14	0.268774	495914.698435	0.000092
15	0.256166	468961.839168	0.000086
16	0.245663	447213.284383	0.000080
17	0.236970	429730.396082	0.000075
18	0.229835	415770.101811	0.000070
19	0.224042	404727.122310	0.000067
20	0.219397	396096.531088	0.000063
21	0.215728	389449.633746	0.000060
22	0.212878	384418.633912	0.000057
23	0.210707	380686.994291	0.000055
24	0.209090	377983.305061	0.000053
25	0.207914	376077.100833	0.000050
26	0.207082	374775.552309	0.000049
27	0.206513	373920.354370	0.000047
28	0.206139	373384.453008	0.000045
29	0.205905	373068.497917	0.000044
30	0.205766	372897.073406	0.000043
31	0.205691	372814.852189	0.000042
32	0.205655	372782.847001	0.000040
33	0.205641	372774.921156	0.000039
34	0.205638	372774.679680	0.000039
35	0.205638	372772.813233	0.000038
36	0.205637	372764.919505	0.000037
37	0.205633	372749.787709	0.000036
38	0.205627	372728.104278	0.000036
39	0.205617	372701.521583	0.000035
40	0.205606	372672.024912	0.000035
41	0.205593	372641.533556	0.000034
42	0.205579	372611.677263	0.000034
43	0.205566	372583.697429	0.000033
44	0.205554	372558.431537	0.000033
45	0.205543	372536.348494	0.000033

Given, $T = 4.000$ sec, $M = 30.000$ m/s^2
$P_1 = 0.100$, $P_2 = 0.300$
$R = 160.000$ km , $\sigma_\rho = 1000.000$ m
$\sigma_\theta = 0.017$ rad

Step no.	Kalman gain	Mean-square range prediction error	Mean-square bearing prediction error
k	$K_{11}(k)$	$P_{11}(k+1\|k)$	$P_{33}(k+1\|k)$
3	0.835647	2454838.683076	0.000679
4	0.710551	1686139.746469	0.000441
5	0.627719	1371531.455742	0.000329
6	0.578332	1236391.391102	0.000267
7	0.552851	1183744.982860	0.000229
8	0.542071	1167342.522005	0.000205
9	0.538605	1163997.844967	0.000190
10	0.537892	1163795.050576	0.000181
11	0.537849	1163719.361654	0.000176
12	0.537833	1163410.287183	0.000173
13	0.537767	1163078.734395	0.000172
14	0.537696	1162856.716082	0.000171
15	0.537649	1162748.001667	0.000171
16	0.537625	1162708.337924	0.000171
17	0.537617	1162698.451156	0.000171
18	0.537615	1162697.317494	0.000171
19	0.537614	1162697.305987	0.000171
20	0.537614	1162696.911622	0.000171
21	0.537614	1162696.357132	0.000171
22	0.537614	1162695.932799	0.000171
23	0.537614	1162695.702828	0.000171
24	0.537614	1162695.609613	0.000171
25	0.537614	1162695.582586	0.000171
26	0.537614	1162695.578130	0.000171
27	0.537614	1162695.578059	0.000171
28	0.537614	1162695.577650	0.000171
29	0.537614	1162695.576781	0.000171
30	0.537614	1162695.576001	0.000171
31	0.537614	1162695.575531	0.000171
32	0.537614	1162695.575320	0.000171
33	0.537614	1162695.575251	0.000171
34	0.537614	1162695.575236	0.000171
35	0.537614	1162695.575235	0.000171
36	0.537614	1162695.575235	0.000171
37	0.537614	1162695.575234	0.000171
38	0.537614	1162695.575232	0.000171
39	0.537614	1162695.575231	0.000171
40	0.537614	1162695.575231	0.000171
41	0.537614	1162695.575231	0.000171
42	0.537614	1162695.575231	0.000171
43	0.537614	1162695.575231	0.000171
44	0.537614	1162695.575231	0.000171
45	0.537614	1162695.575231	0.000171

Given, $T = 4.000$ sec, $M = 30.000\ m/s^2$

$P_1 = 0.100$, $P_2 = 0.300$

$R = 370.000$ km , $\sigma_\rho = 1000.000$ m

$\sigma_\theta = 0.017$ rad

Step no.	Kalman gain	Mean-square range prediction error	Mean-square bearing prediction error
k	$K_{11}(k)$	$P_{11}(k+1\|k)$	$P_{33}(k+1\|k)$
3	0.835647	2454838.683076	0.000675
4	0.710551	1686139.746469	0.000435
5	0.627719	1371531.455742	0.000320
6	0.578332	1236391.391102	0.000254
7	0.552851	1183744.982860	0.000211
8	0.542071	1167342.522005	0.000181
9	0.538605	1163997.844967	0.000160
10	0.537892	1163795.050576	0.000145
11	0.537849	1163719.361654	0.000133
12	0.537833	1163410.287183	0.000124
13	0.537767	1163078.734395	0.000118
14	0.537696	1162856.716082	0.000113
15	0.537649	1162748.001667	0.000110
16	0.537625	1162708.337924	0.000107
17	0.537617	1162698.451156	0.000106
18	0.537615	1162697.317494	0.000105
19	0.537614	1162697.305987	0.000104
20	0.537614	1162696.911622	0.000103
21	0.537614	1162696.357132	0.000103
22	0.537614	1162695.932799	0.000103
23	0.537614	1162695.702828	0.000103
24	0.537614	1162695.609613	0.000103
25	0.537614	1162695.582586	0.000103
26	0.537614	1162695.578130	0.000103
27	0.537614	1162695.578059	0.000103
28	0.537614	1162695.577650	0.000103
29	0.537614	1162695.576781	0.000103
30	0.537614	1162695.576001	0.000103
31	0.537614	1162695.575531	0.000103
32	0.537614	1162695.575320	0.000103
33	0.537614	1162695.575251	0.000103
34	0.537614	1162695.575236	0.000103
35	0.537614	1162695.575235	0.000103
36	0.537614	1162695.575235	0.000103
37	0.537614	1162695.575234	0.000103
38	0.537614	1162695.575232	0.000103
39	0.537614	1162695.575231	0.000103
40	0.537614	1162695.575231	0.000103
41	0.537614	1162695.575231	0.000103
42	0.537614	1162695.575231	0.000103
43	0.537614	1162695.575231	0.000103
44	0.537614	1162695.575231	0.000103
45	0.537614	1162695.575231	0.000103

Given, T = 4.000 sec, M = 7.500 m/s^2
P_1 = 0.100 , P_2 = 0.300
R = 160.000 km , σ_ρ = 1000.000 m
σ_θ = 0.020 rad

Step no.	Kalman gain	Mean-square range prediction error	Mean-square beaing prediction error
k	$K_{11}(k)$	$P_{11}(k+1\|k)$	$P_{33}(k+1\|k)$
3	0.833480	2340957.936749	0.000934
4	0.700685	1511864.011394	0.000600
5	0.601889	1117934.126381	0.000441
6	0.527842	892669.234060	0.000348
7	0.471646	750502.687012	0.000287
8	0.428736	655783.683887	0.000245
9	0.396056	590991.038013	0.000214
10	0.371461	546389.627915	0.000190
11	0.353332	515984.738053	0.000171
12	0.340363	495740.970762	0.000157
13	0.331435	482745.404955	0.000145
14	0.325575	474806.314889	0.000135
15	0.321945	470259.450453	0.000127
16	0.319848	467866.345550	0.000121
17	0.318739	466744.580362	0.000115
18	0.318218	466303.175838	0.000111
19	0.318013	466176.650788	0.000107
20	0.317954	466161.300111	0.000105
21	0.317947	466159.566295	0.000102
22	0.317946	466136.449753	0.000100
23	0.317935	466088.835917	0.000099
24	0.317913	466026.221632	0.000098
25	0.317884	465960.148353	0.000097
26	0.317853	465899.499182	0.000096
27	0.317825	465849.260003	0.000096
28	0.317802	465811.002366	0.000096
29	0.317784	465783.979704	0.000096
30	0.317771	465766.229504	0.000095
31	0.317763	465755.417748	0.000095
32	0.317758	465749.367394	0.000095
33	0.317755	465746.315477	0.000095
34	0.317754	465744.979508	0.000095
35	0.317753	465744.512783	0.000095
36	0.317753	465744.410580	0.000095
37	0.317753	465744.407821	0.000095
38	0.317753	465744.390157	0.000095
39	0.317753	465744.326936	0.000095
40	0.317753	465744.226312	0.000095
41	0.317753	465744.108628	0.000095
42	0.317753	465743.992967	0.000095
43	0.317753	465743.892062	0.000095
44	0.317753	465743.811811	0.000095
45	0.317753	465743.752842	0.000095

Given, T = 4.000 sec, M = 7.500 m/s^2
P_1 = 0.100 , P_2 = 0.300
R = 370.000 km , σ_ρ = 1000.000 m
σ_θ = 0.020 rad

Step no.	Kalman gain	Mean-square range prediction error	Mean-square bearing prediction error
k	$K_{11}(k)$	$P_{11}(k+1\|k)$	$P_{33}(k+1\|k)$
3	0.833480	2340957.936749	0.000933
4	0.700685	1511864.011394	0.000600
5	0.601889	1117934.126381	0.000440
6	0.527842	892669.234060	0.000347
7	0.471646	750502.687012	0.000286
8	0.428736	655783.683887	0.000243
9	0.396056	590991.038013	0.000212
10	0.371461	546389.627915	0.000187
11	0.353332	515984.738053	0.000168
12	0.340363	495740.970762	0.000153
13	0.331435	482745.404955	0.000140
14	0.325575	474806.314889	0.000129
15	0.321945	470259.450453	0.000120
16	0.319848	467866.345550	0.000112
17	0.318739	466744.580362	0.000105
18	0.318218	466303.175838	0.000100
19	0.318013	466176.650788	0.000095
20	0.317954	466161.300111	0.000090
21	0.317947	466159.566295	0.000086
22	0.317946	466136.449753	0.000083
23	0.317935	466088.835917	0.000080
24	0.317913	466026.221632	0.000077
25	0.317884	465960.148353	0.000075
26	0.317853	465899.499182	0.000073
27	0.317825	465849.260003	0.000071
28	0.317802	465811.002366	0.000069
29	0.317784	465783.979704	0.000068
30	0.317771	465766.229504	0.000067
31	0.317763	465755.417748	0.000066
32	0.317758	465749.367394	0.000065
33	0.317755	465746.315477	0.000064
34	0.317754	465744.979508	0.000063
35	0.317753	465744.512783	0.000063
36	0.317753	465744.410580	0.000062
37	0.317753	465744.407821	0.000062
38	0.317753	465744.390157	0.000061
39	0.317753	465744.326936	0.000061
40	0.317753	465744.226312	0.000061
41	0.317753	465744.108628	0.000061
42	0.317753	465743.992967	0.000061
43	0.317753	465743.892062	0.000061
44	0.317753	465743.811811	0.000060
45	0.317753	465743.752842	0.000060

Given, T = 4.000 sec, M = 2.100 m/s^2
P_1 = 0.000 , P_2 = 0.000
R = 160.000 km , σ_ρ = 1000.000 m
σ_θ = 0.020 rad

Step no.	Kalman gain	Mean-square range prediction error	Mean-square bearing prediction error
k	$K_{11}(k)$	$P_{11}(k+1\|k)$	$P_{33}(k+1\|k)$
3	0.833344	2333876.896177	0.000933
4	0.700049	1500846.638585	0.000600
5	0.600135	1101282.890097	0.000440
6	0.524100	868535.360135	0.000347
7	0.464821	716908.974513	0.000286
8	0.417558	610708.993819	0.000243
9	0.379155	532494.124017	0.000211
10	0.347469	472758.502497	0.000187
11	0.321002	425890.777713	0.000168
12	0.298684	388369.811262	0.000152
13	0.279731	357876.089790	0.000139
14	0.263556	332819.358720	0.000128
15	0.249711	312071.605086	0.000119
16	0.237846	294808.303509	0.000111
17	0.227685	280409.982152	0.000104
18	0.219000	268399.023457	0.000098
19	0.211605	258397.920500	0.000092
20	0.205339	250101.096516	0.000088
21	0.200065	243255.594163	0.000084
22	0.195660	237647.741415	0.000080
23	0.192016	233093.943387	0.000076
24	0.189032	229434.367619	0.000073
25	0.186618	226528.666169	0.000071
26	0.184691	224253.114424	0.000068
27	0.183175	222498.703648	0.000066
28	0.182003	221169.837146	0.000064
29	0.181113	220183.368183	0.000062
30	0.180451	219467.791523	0.000060
31	0.179970	218962.463593	0.000059
32	0.179630	218616.779586	0.000058
33	0.179397	218389.278588	0.000056
34	0.179244	218246.679389	0.000055
35	0.179148	218162.870157	0.000054
36	0.179092	218117.885746	0.000053
37	0.179061	218096.908758	0.000052
38	0.179047	218089.327118	0.000052
39	0.179042	218087.873862	0.000051
40	0.179041	218087.866377	0.000050
41	0.179041	218086.553754	0.000050
42	0.179040	218082.573441	0.000049
43	0.179038	218075.512405	0.000049
44	0.179033	218065.563843	0.000049
45	0.179026	218053.267866	0.000048

Given, T = 4.000 sec, M = 2.100 m/s^2
P_1 = 0.000 , P_2 = 0.000
R = 370.000 km , σ_ρ = 1000.000 m
σ_θ = 0.020 rad

Step no.	Kalman gain	Mean-square range prediction error	Mean-square bearing prediction error
k	$K_{11}(k)$	$P_{11}(k+1\|k)$	$P_{33}(k+1\|k)$
3	0.833344	2333876.896177	0.000933
4	0.700049	1500846.638585	0.000600
5	0.600135	1101282.890097	0.000440
6	0.524100	868535.360135	0.000347
7	0.464821	716908.974513	0.000286
8	0.417558	610708.993819	0.000243
9	0.379155	532494.124017	0.000211
10	0.347469	472758.502497	0.000187
11	0.321002	425890.777713	0.000167
12	0.298684	388369.811262	0.000152
13	0.279731	357876.089790	0.000139
14	0.263556	332819.358720	0.000128
15	0.249711	312071.605086	0.000118
16	0.237846	294808.303509	0.000110
17	0.227685	280409.982152	0.000103
18	0.219000	268399.023457	0.000097
19	0.211605	258397.920500	0.000091
20	0.205339	250101.096516	0.000087
21	0.200065	243255.594163	0.000082
22	0.195660	237647.741415	0.000078
23	0.192016	233093.943387	0.000075
24	0.189032	229434.367619	0.000071
25	0.186618	226528.666169	0.000069
26	0.184691	224253.114424	0.000066
27	0.183175	222498.703648	0.000063
28	0.182003	221169.837146	0.000061
29	0.181113	220183.368183	0.000059
30	0.180451	219467.791523	0.000057
31	0.179970	218962.463593	0.000055
32	0.179630	218616.779586	0.000053
33	0.179397	218389.278588	0.000052
34	0.179244	218246.679389	0.000050
35	0.179148	218162.870157	0.000049
36	0.179092	218117.885746	0.000048
37	0.179061	218096.908758	0.000047
38	0.179047	218089.327118	0.000045
39	0.179042	218087.873862	0.000044
40	0.179041	218087.866377	0.000043
41	0.179041	218086.553754	0.000042
42	0.179040	218082.573441	0.000042
43	0.179038	218075.512405	0.000041
44	0.179033	218065.563843	0.000040
45	0.179026	218053.267866	0.000039

Given, T = 15.000 sec, M = 2.100 m/s^2
P_1 = 0.000 , P_2 = 0.000
R = 160.000 km , σ_ρ = 1000.000 m
σ_θ = 0.017 rad

Step no.	Kalman gain	Mean-square range prediction error	Mean-square bearing prediction error
k	$K_{11}(k)$	$P_{11}(k+1\|k)$	$P_{33}(k+1\|k)$
3	0.835375	2440421.875813	0.000679
4	0.709338	1664370.019028	0.000440
5	0.624677	1340788.109305	0.000328
6	0.572793	1196798.104161	0.000265
7	0.544792	1137263.016525	0.000226
8	0.532112	1116804.078032	0.000202
9	0.527590	1111786.227676	0.000186
10	0.526467	1111219.830121	0.000176
11	0.526340	1111211.386237	0.000170
12	0.526338	1110986.172784	0.000167
13	0.526288	1110665.085340	0.000165
14	0.526216	1110411.231257	0.000165
15	0.526159	1110267.223461	0.000164
16	0.526126	1110205.067344	0.000164
17	0.526112	1110185.238596	0.000164
18	0.526108	1110181.274970	0.000164
19	0.526107	1110181.072229	0.000164
20	0.526107	1110180.939459	0.000164
21	0.526107	1110180.474331	0.000164
22	0.526107	1110179.978085	0.000164
23	0.526107	1110179.638437	0.000164
24	0.526107	1110179.465633	0.000164
25	0.526107	1110179.398838	0.000164
26	0.526107	1110179.380508	0.000164
27	0.526107	1110179.377840	0.000164
28	0.526107	1110179.377835	0.000164
29	0.526107	1110179.377370	0.000164
30	0.526107	1110179.376540	0.000164
31	0.526107	1110179.375820	0.000164
32	0.526107	1110179.375385	0.000164
33	0.526107	1110179.375186	0.000164
34	0.526107	1110179.375118	0.000164
35	0.526107	1110179.375103	0.000164
36	0.526107	1110179.375101	0.000164
37	0.526107	1110179.375101	0.000164
38	0.526107	1110179.375100	0.000164
39	0.526107	1110179.375099	0.000164
40	0.526107	1110179.375098	0.000164
41	0.526107	1110179.375097	0.000164
42	0.526107	1110179.375097	0.000164
43	0.526107	1110179.375097	0.000164
44	0.526107	1110179.375097	0.000164
45	0.526107	1110179.375097	0.000164

Given, $T = 15.000$ sec, $M = 2.100$ m/s^2
$P_1 = 0.000$, $P_2 = 0.000$
$R = 370.000$ km , $\sigma_\rho = 1200.000$ m
$\sigma_\theta = .0.020$ rad

Step no.	Kalman gain	Mean-square range prediction error	Mean-square bearing prediction error
k	$K_{11}(k)$	$P_{11}(k+1\|k)$	$P_{33}(k+1\|k)$
3	0.834757	3467211.298556	0.000934
4	0.706554	2325285.512761	0.000601
5	0.617559	1828524.941993	0.000442
6	0.559434	1588879.782922	0.000349
7	0.524577	1475408.139004	0.000289
8	0.506073	1427015.090931	0.000248
9	0.497735	1409960.290613	0.000218
10	0.494730	1405642.417137	0.000195
11	0.493963	1405135.171843	0.000178
12	0.493873	1405126.452445	0.000165
13	0.493871	1404890.301099	0.000155
14	0.493829	1404523.091110	0.000147
15	0.493764	1404199.689532	0.000141
16	0.493706	1403990.086058	0.000136
17	0.493669	1403882.728000	0.000133
18	0.493650	1403839.149940	0.000130
19	0.493642	1403825.974994	0.000128
20	0.493640	1403823.609922	0.000127
21	0.493639	1403823.557079	0.000126
22	0.493639	1403823.336525	0.000126
23	0.493639	1403822.777224	0.000125
24	0.493639	1403822.187753	0.000125
25	0.493639	1403821.760495	0.000125
26	0.493639	1403821.520273	0.000125
27	0.493639	1403821.412665	0.000125
28	0.493639	1403821.375457	0.000125
29	0.493639	1403821.366759	0.000125
30	0.493639	1403821.365991	0.000125
31	0.493639	1403821.365892	0.000125
32	0.493639	1403821.365140	0.000125
33	0.493639	1403821.364128	0.000125
34	0.493639	1403821.363292	0.000125
35	0.493639	1403821.362774	0.000125
36	0.493639	1403821.362519	0.000125
37	0.493639	1403821.362421	0.000125
38	0.493639	1403821.362393	0.000125
39	0.493639	1403821.362389	0.000125
40	0.493639	1403821.362389	0.000125
41	0.493639	1403821.362388	0.000125
42	0.493639	1403821.362386	0.000125
43	0.493639	1403821.362385	0.000125
44	0.493639	1403821.362383	0.000125
45	0.493639	1403821.362383	0.000125

Given, T = 10.000 sec, M = 2.100 m/s^2
P_1 = 0.000 , P_2 = 0.000
R = 160.000 km , σ_ρ = 1000.000 m
σ_θ = 0.017 rad

Step no.	Kalman gain	Mean-square range prediction error	Mean-square bearing prediction error
k	$K_{11}(k)$	$P_{11}(k+1\|k)$	$P_{33}(k+1\|k)$
3	0.833741	2354550.699120	0.000675
4	0.701897	1532951.639093	0.000435
5	0.605204	1149585.232138	0.000320
6	0.534794	937942.653955	0.000253
7	0.483989	812161.344124	0.000210
8	0.448173	735859.106094	0.000181
9	0.423916	690349.949047	0.000160
10	0.408407	664471.087025	0.000144
11	0.399209	650843.827015	0.000132
12	0.394249	644417.990291	0.000123
13	0.391882	641836.678893	0.000117
14	0.390926	641037.586494	0.000112
15	0.390629	640896.156729	0.000108
16	0.390577	640895.530934	0.000106
17	0.390577	640858.940321	0.000104
18	0.390563	640765.928399	0.000103
19	0.390529	640646.643519	0.000102
20	0.390484	640532.874896	0.000101
21	0.390442	640442.608164	0.000101
22	0.390408	640380.334876	0.000101
23	0.390385	640342.348220	0.000101
24	0.390371	640321.858382	0.000101
25	0.390364	640312.246535	0.000101
26	0.390360	640308.493329	0.000101
27	0.390359	640307.401453	0.000101
28	0.390358	640307.239871	0.000101
29	0.390358	640307.239084	0.000101
30	0.390358	640307.156393	0.000101
31	0.390358	640306.981214	0.000101
32	0.390358	640306.770740	0.000101
33	0.390358	640306.577254	0.000101
34	0.390358	640306.427713	0.000101
35	0.390358	640306.326790	0.000101
36	0.390358	640306.266503	0.000101
37	0.390358	640306.234711	0.000101
38	0.390358	640306.220206	0.000101
39	0.390358	640306.214766	0.000101
40	0.390358	640306.213298	0.000101
41	0.390358	640306.213126	0.000101
42	0.390358	640306.213114	0.000101
43	0.390358	640306.212936	0.000101
44	0.390358	640306.212610	0.000101
45	0.390358	640306.212241	0.000101

Given,	T =	10.000 sec,	M =	2.100	m/s^2
	P_1 =	0.000 ,	P_2 =	0.000	
	R =	370.000 km ,	σ_ρ =	1200.000	m
	σ_θ =	0.020 rad			

Step no.	Kalman gain	Mean-square range prediction error	Mean-square bearing prediction error
k	$K_{11}(k)$	$P_{11}(k+1 \mid k)$	$P_{33}(k+1 \mid k)$
3	0.833616	3381222.236473	0.000933
4	0.701321	2192989.144631	0.000600
5	0.603632	1633748.785831	0.000440
6	0.531517	1319802.516358	0.000347
7	0.478224	1127825.942192	0.000286
8	0.439214	1006084.242147	0.000244
9	0.411304	928619.542959	0.000212
10	0.392051	880412.691293	0.000188
11	0.379421	851720.368428	0.000169
12	0.371651	835740.484936	0.000154
13	0.367239	827620.272288	0.000142
14	0.364973	823988.836575	0.000132
15	0.363954	822650.501226	0.000123
16	0.363578	822304.133144	0.000116
17	0.363481	822271.880777	0.000110
18	0.363472	822262.822195	0.000104
19	0.363469	822193.974178	0.000100
20	0.363450	822073.113666	0.000096
21	0.363416	821933.327477	0.000093
22	0.363376	821803.416924	0.000091
23	0.363340	821699.207604	0.000088
24	0.363311	821624.762462	0.000087
25	0.363290	821576.783841	0.000085
26	0.363276	821548.852928	0.000084
27	0.363268	821534.309059	0.000083
28	0.363264	821527.707433	0.000082
29	0.363262	821525.242700	0.000081
30	0.363262	821524.591434	0.000081
31	0.363261	821524.525818	0.000080
32	0.363261	821524.512537	0.000080
33	0.363261	821524.395271	0.000080
34	0.363261	821524.184135	0.000080
35	0.363261	821523.937143	0.000080
36	0.363261	821523.705990	0.000080
37	0.363261	821523.519618	0.000080
38	0.363261	821523.385915	0.000080
39	0.363261	821523.299412	0.000080
40	0.363261	821523.248854	0.000080
41	0.363261	821523.222409	0.000080
42	0.363261	821523.210335	0.000080
43	0.363261	821523.205786	0.000080
44	0.363261	821523.204562	0.000080
45	0.363261	821523.204430	0.000080

Given, $T = 5.000$ sec, $M = 5.000$ m/s^2
$P_1 = 0.100$, $P_2 = 0.400$
$R = 160.000$ km , $\sigma_\rho = 1100.000$ m
$\sigma_\theta = 0.019$ rad

Step no. k	Kalman gain $K_{11}(k)$	Mean-square range prediction error $P_{11}(k+1\|k)$	Mean-square bearing prediction error $P_{33}(k+1\|k)$
3	0.833453	2830854.822021	0.000843
4	0.700558	1826705.886398	0.000542
5	0.601542	1348703.517870	0.000398
6	0.527104	1074358.344898	0.000314
7	0.470311	900124.643679	0.000259
8	0.426574	782888.908689	0.000221
9	0.392841	701519.907463	0.000193
10	0.366996	644356.898507	0.000172
11	0.347483	604307.212190	0.000155
12	0.333079	576671.178922	0.000142
13	0.322763	558099.838940	0.000131
14	0.315650	546078.043750	0.000123
15	0.310965	538667.304091	0.000116
16	0.308044	534375.950266	0.000110
17	0.306342	532084.725264	0.000105
18	0.305430	530989.784900	0.000101
19	0.304993	530546.872909	0.000098
20	0.304816	530413.892524	0.000096
21	0.304763	530395.501030	0.000094
22	0.304756	530394.537078	0.000092
23	0.304755	530373.448964	0.000091
24	0.304747	530326.516481	0.000090
25	0.304728	530261.780484	0.000089
26	0.304702	530190.607217	0.000089
27	0.304674	530122.604725	0.000089
28	0.304647	530063.881791	0.000088
29	0.304623	530017.118826	0.000088
30	0.304605	529982.417318	0.000088
31	0.304591	529958.312535	0.000088
32	0.304581	529942.640375	0.000088
33	0.304575	529933.147788	0.000088
34	0.304571	529927.848553	0.000088
35	0.304569	529925.177520	0.000088
36	0.304568	529924.009841	0.000088
37	0.304567	529923.604856	0.000088
38	0.304567	529923.519212	0.000088
39	0.304567	529923.517799	0.000088
40	0.304567	529923.497681	0.000088
41	0.304567	529923.430673	0.000088
42	0.304567	529923.324348	0.000088
43	0.304567	529923.198335	0.000088
44	0.304567	529923.071840	0.000088
45	0.304567	529922.958490	0.000088

Given, T = 5.000 sec, M = 5.000 m/s^2
P_1 = 0.100 , P_2 = 0.400
R = 370.000 km , σ_ρ = 1100.000 m
σ_θ = 0.019 rad

Step no.	Kalman gain	Mean-square range prediction error	Mean-square bearing prediction error
k	$K_{11}(k)$	$P_{11}(k+1\|k)$	$P_{33}(k+1\|k)$
3	0.833453	2830854.822021	0.000842
4	0.700558	1826705.886398	0.000542
5	0.601542	1348703.517870	0.000397
6	0.527104	1074358.344898	0.000313
7	0.470311	900124.643679	0.000258
8	0.426574	782888.908689	0.000220
9	0.392841	701519.907463	0.000191
10	0.366996	644356.898507	0.000169
11	0.347483	604307.212190	0.000152
12	0.333079	576671.178922	0.000138
13	0.322763	558099.838940	0.000126
14	0.315650	546078.043750	0.000116
15	0.310965	538667.304091	0.000108
16	0.308044	534375.950266	0.000101
17	0.306342	532084.725264	0.000095
18	0.305430	530989.784900	0.000090
19	0.304993	530546.872909	0.000086
20	0.304816	530413.892524	0.000082
21	0.304763	530395.501030	0.000078
22	0.304756	530394.537078	0.000075
23	0.304755	530373.448964	0.000072
24	0.304747	530326.516481	0.000070
25	0.304728	530261.780484	0.000068
26	0.304702	530190.607217	0.000066
27	0.304674	530122.604725	0.000065
28	0.304647	530063.881791	0.000063
29	0.304623	530017.118826	0.000062
30	0.304605	529982.417318	0.000061
31	0.304591	529958.312535	0.000060
32	0.304581	529942.640375	0.000059
33	0.304575	529933.147788	0.000059
34	0.304571	529927.848553	0.000058
35	0.304569	529925.177520	0.000058
36	0.304568	529924.009841	0.000057
37	0.304567	529923.604856	0.000057
38	0.304567	529923.519212	0.000057
39	0.304567	529923.517799	0.000056
40	0.304567	529923.497681	0.000056
41	0.304567	529923.430673	0.000056
42	0.304567	529923.324348	0.000056
43	0.304567	529923.198335	0.000056
44	0.304567	529923.071840	0.000056
45	0.304567	529922.958490	0.000056

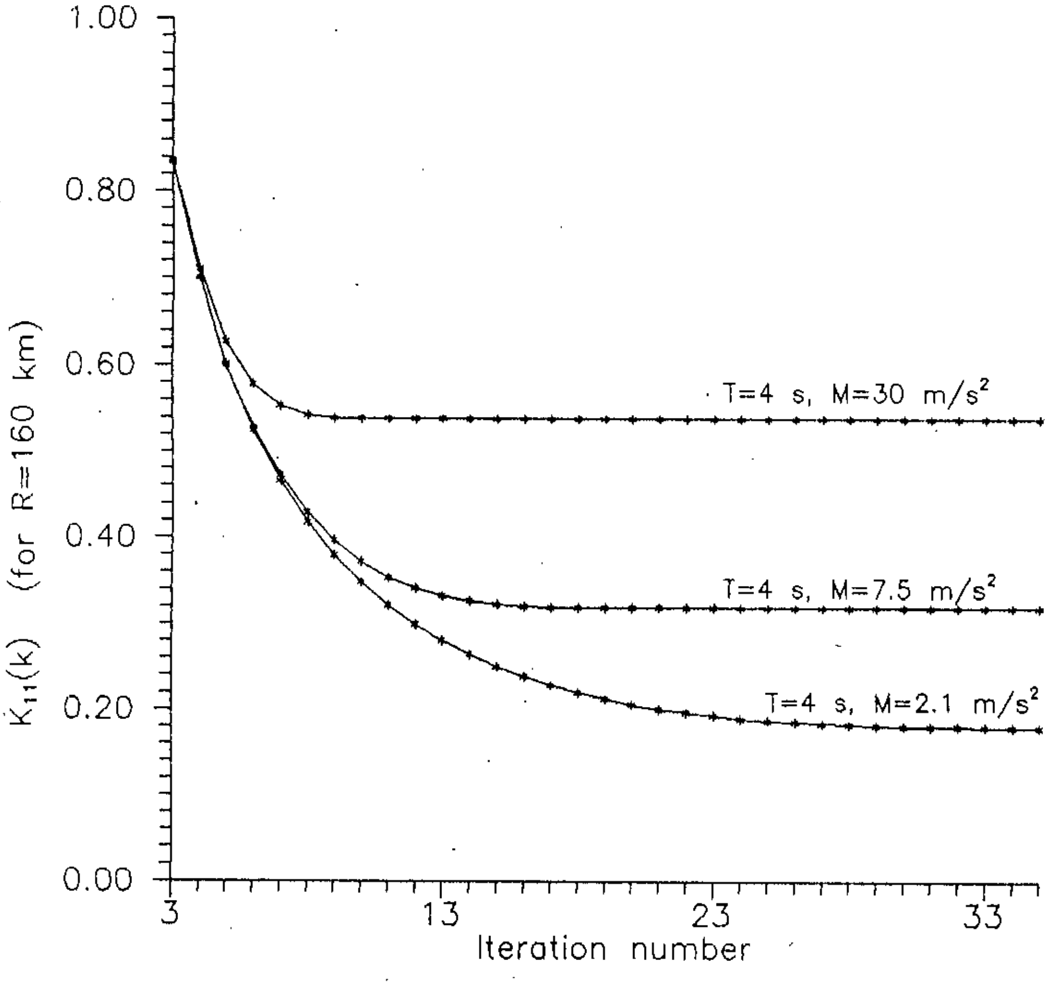

Fig.(7.2) Gain setting of Kalman filter
(for the same scan time)

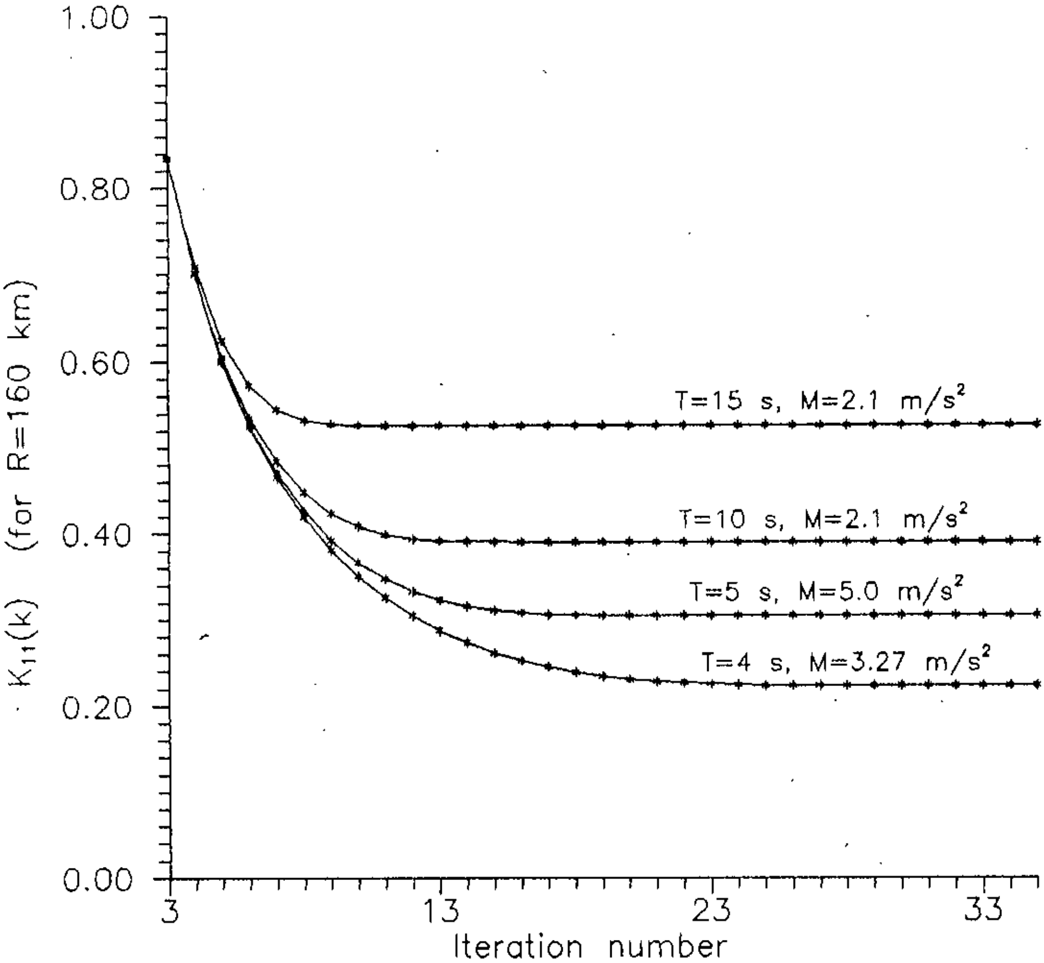

Fig.(7.3) Gain setting of Kalman filter (for different scan time)

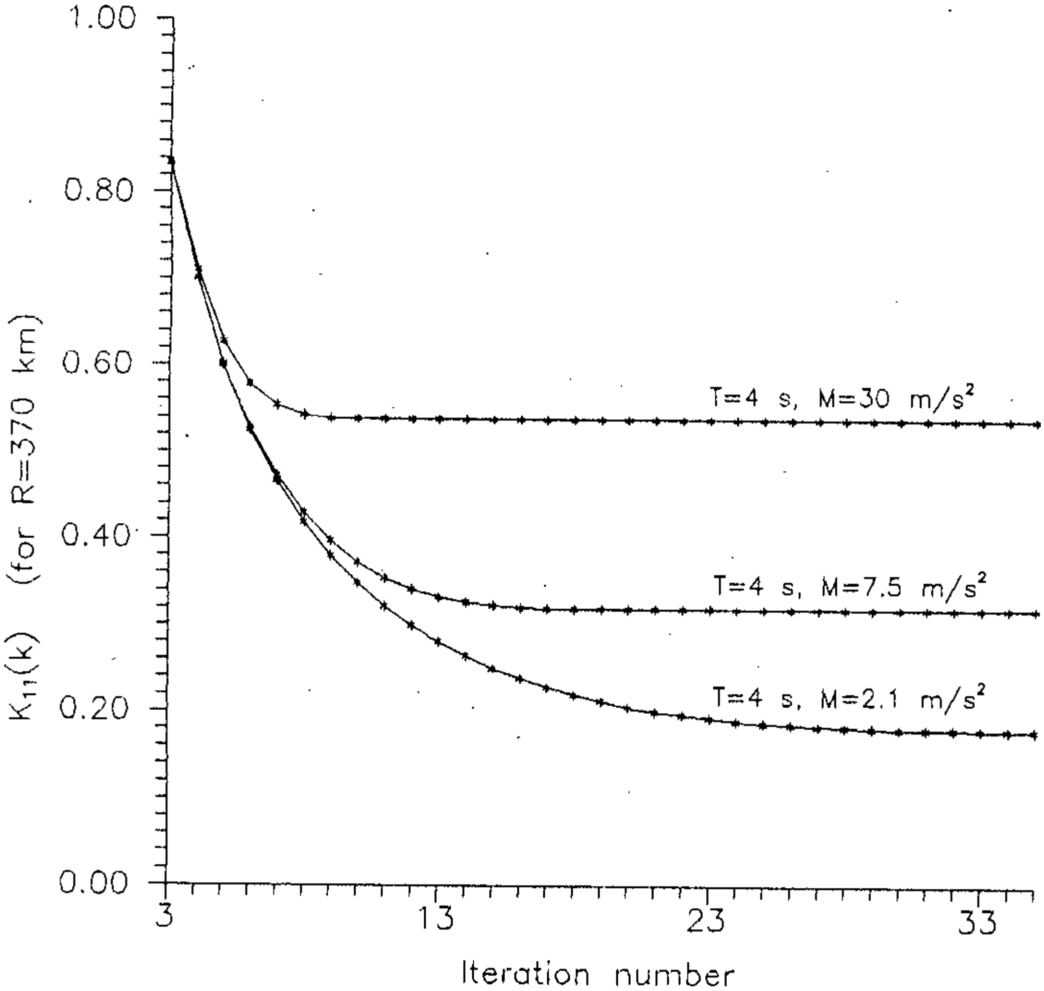

Fig.(7.4) Gain setting of Kalman filter (for the same scan time)

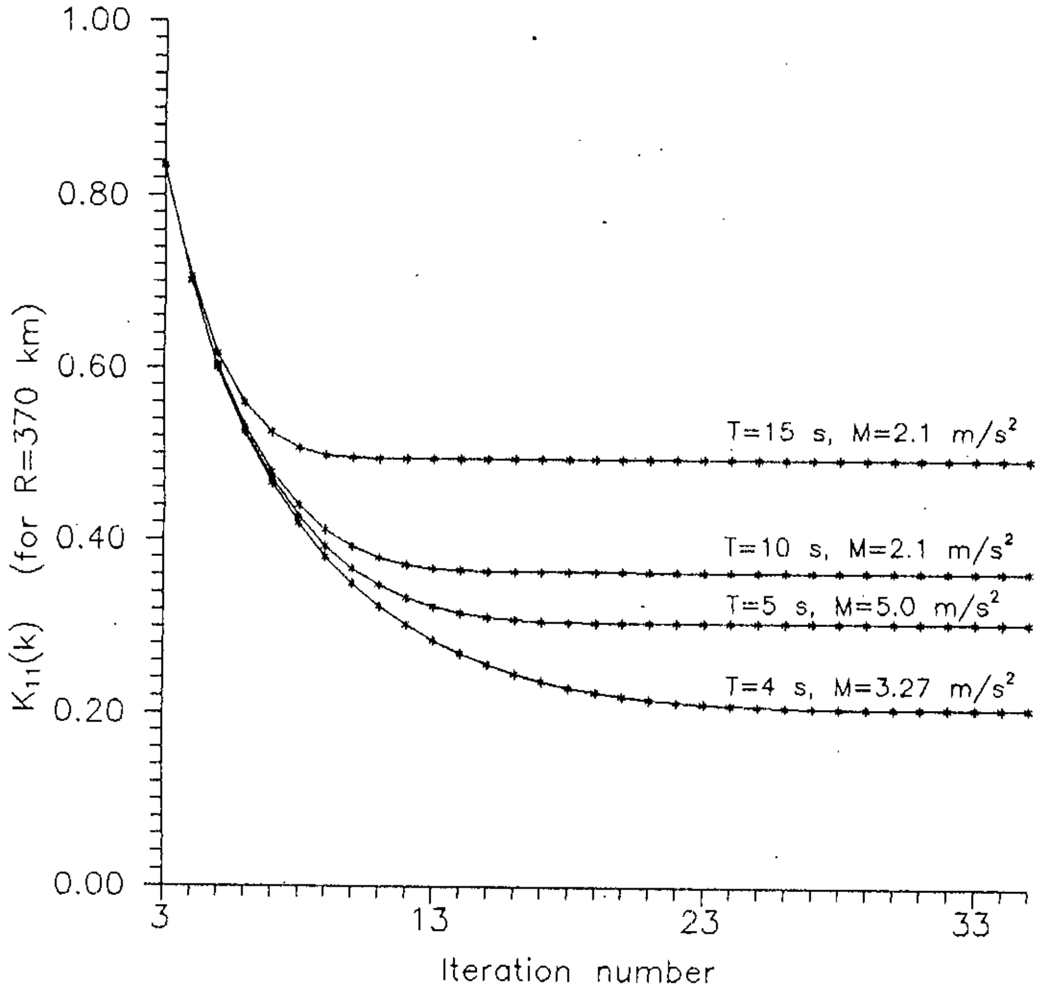

Fig.(7.5) Gain setting of Kalman filter
(for different scan time)

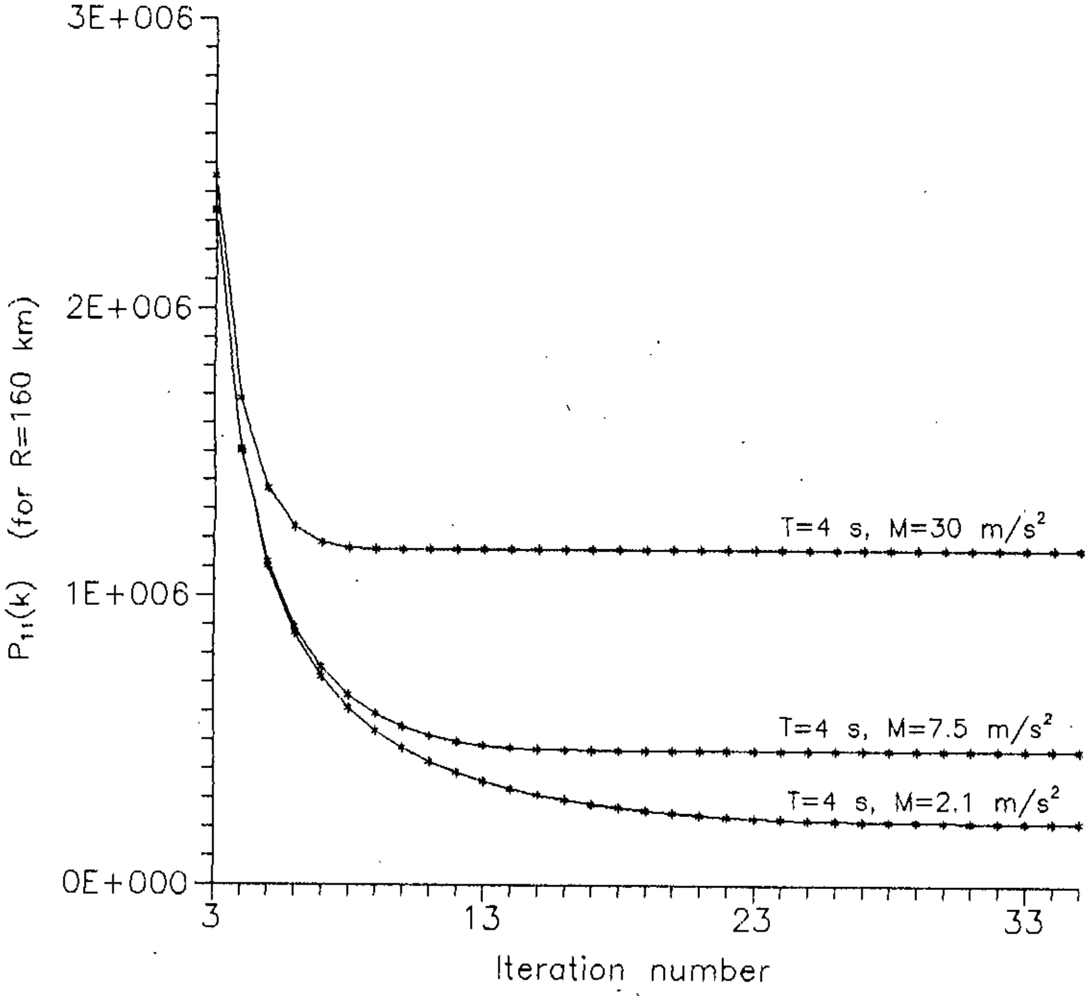

Fig.(7.6) Mean-square range prediction error
(for the same scan time)

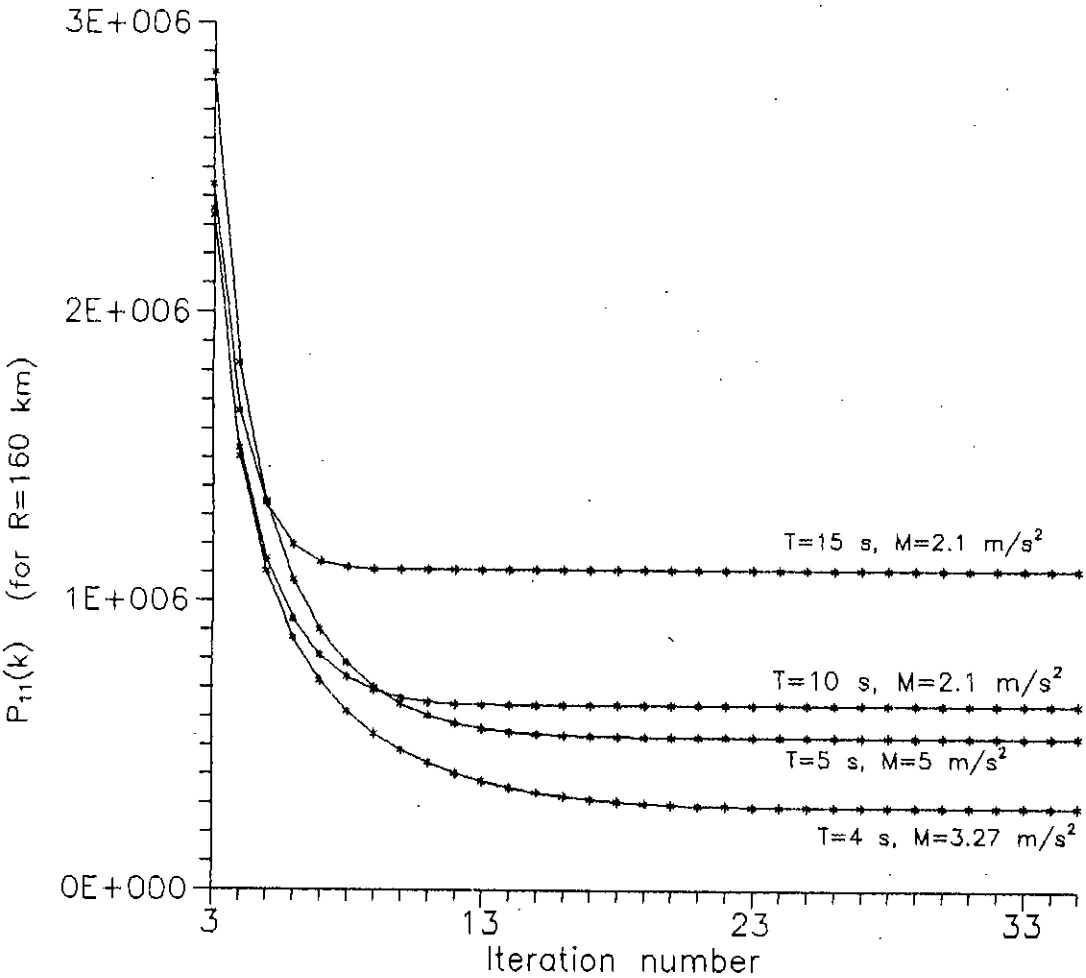

Fig.(7.7) Mean-square range prediction error (for different scan time)

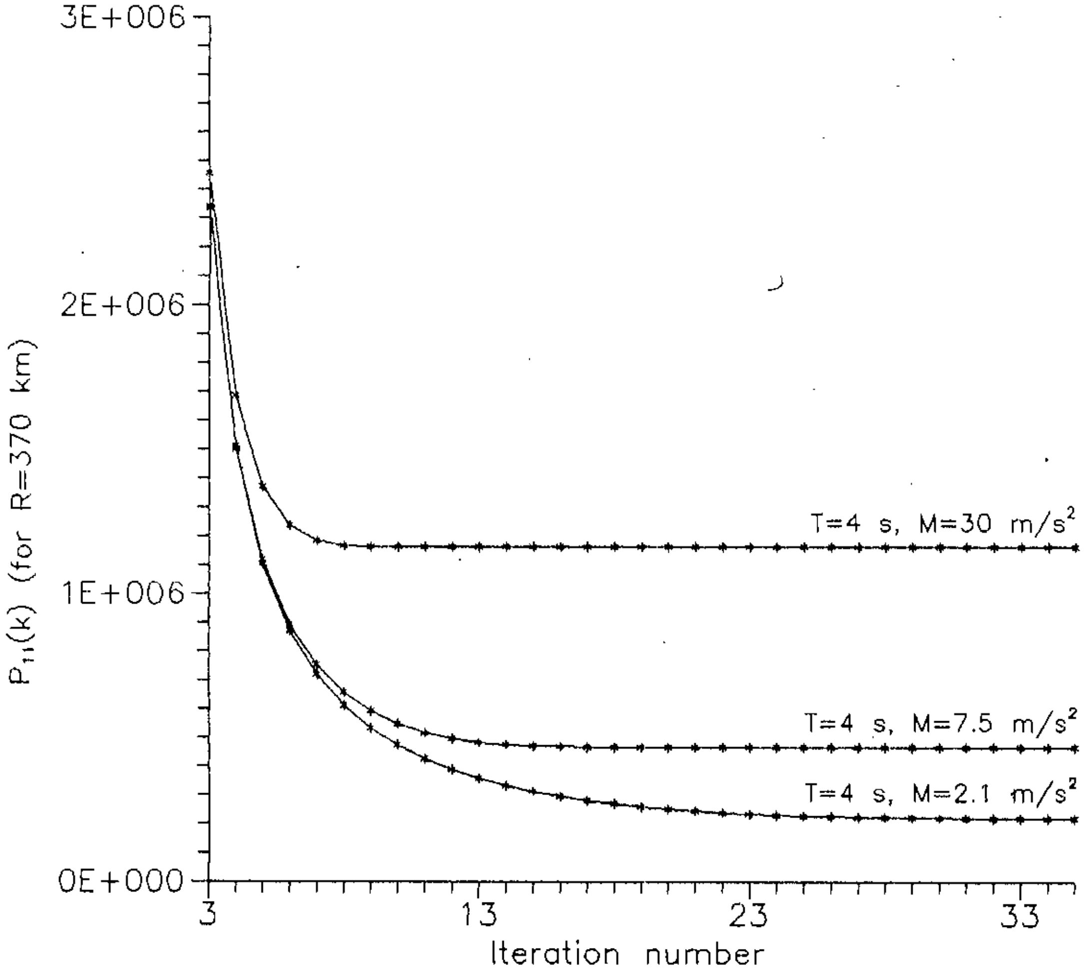

Fig.(7.8) Mean−square range prediction error (for the same scan time)

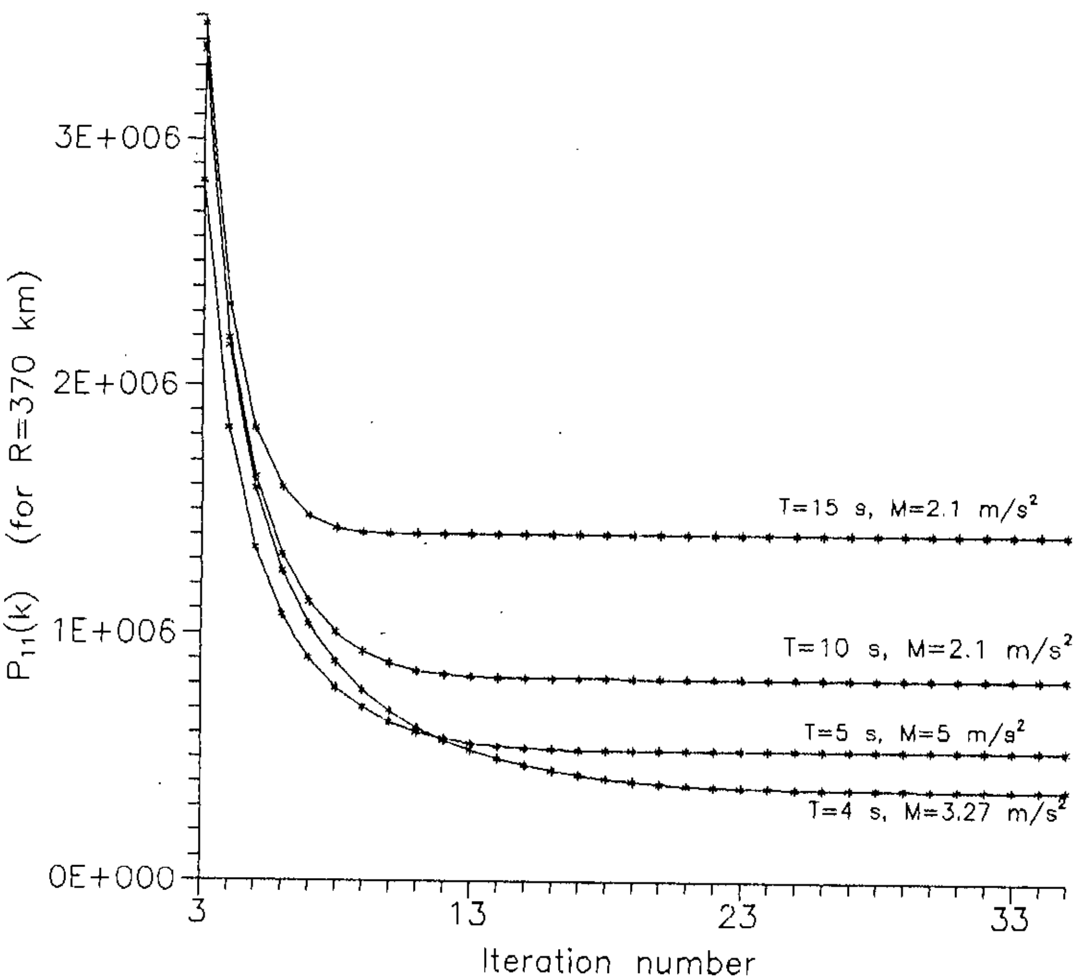

Fig.(7.9) Mean-square range prediction error (for different scan time)

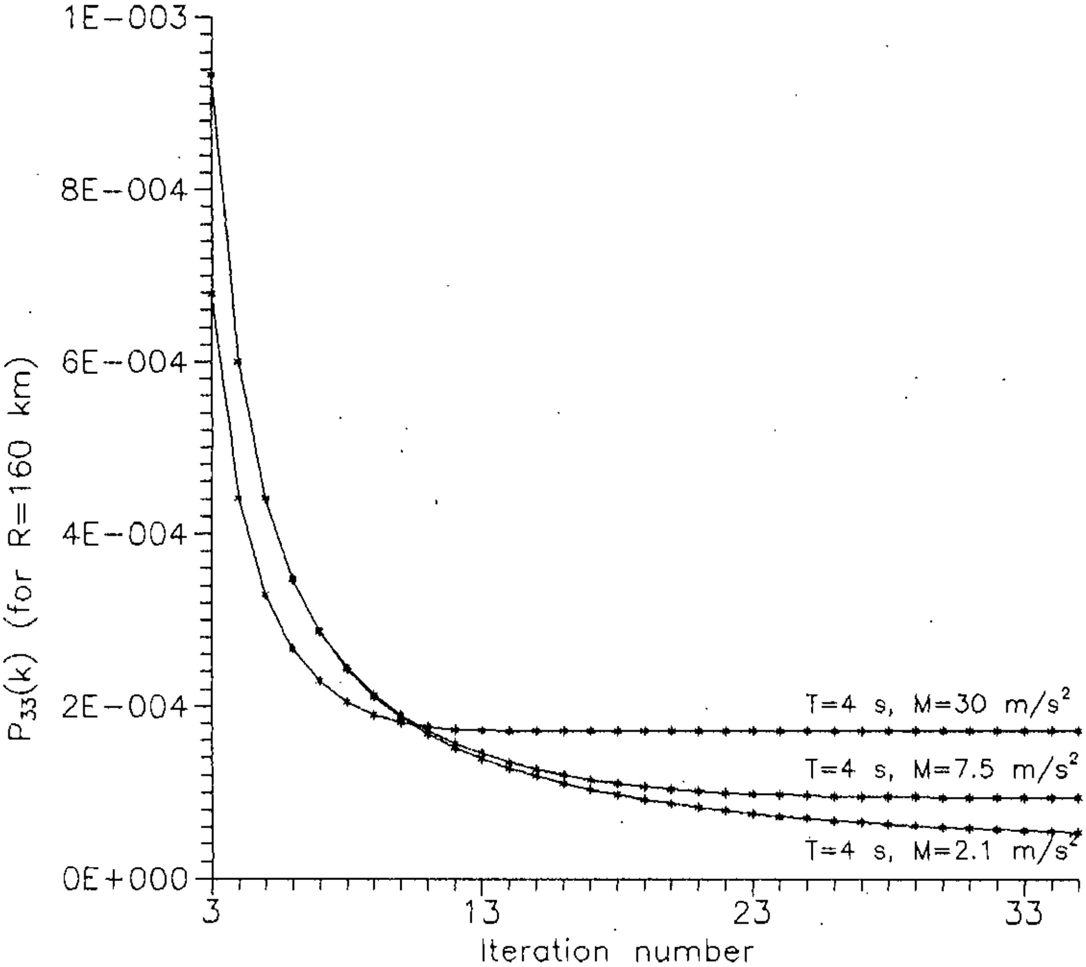

Fig.(7.10) Mean-square bearing prediction error (for the same scan time)

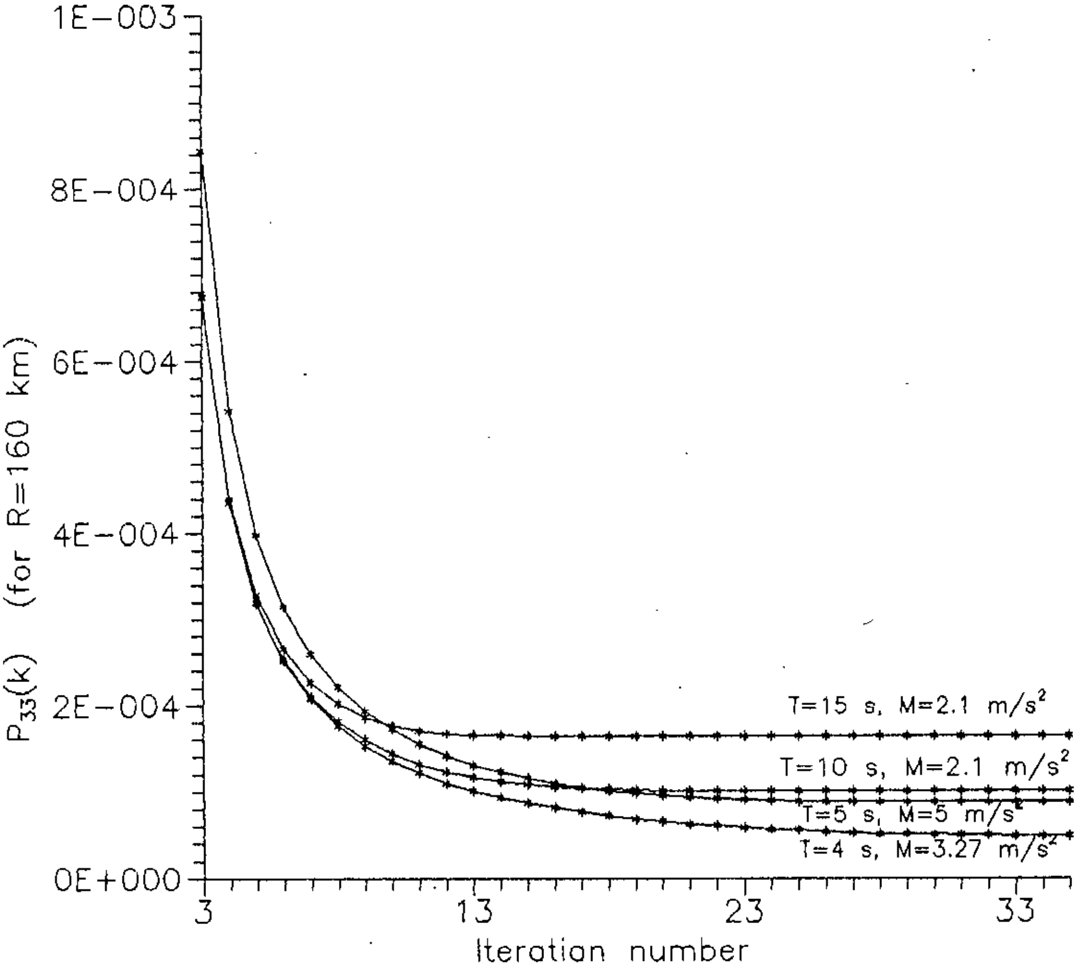

Fig.(7.11) Mean-square bearing prediction error (for different scan time)

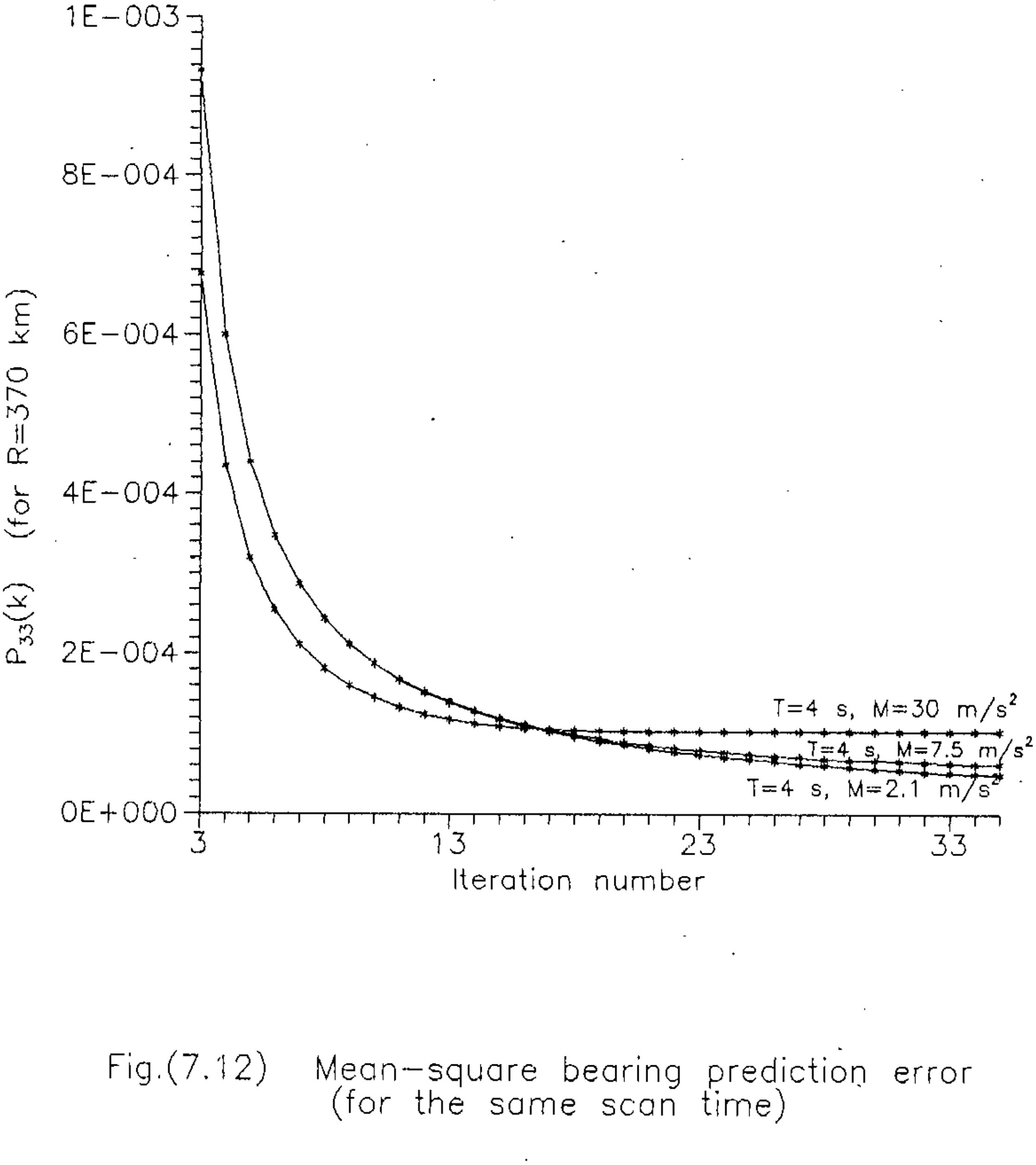

Fig.(7.12) Mean-square bearing prediction error (for the same scan time)

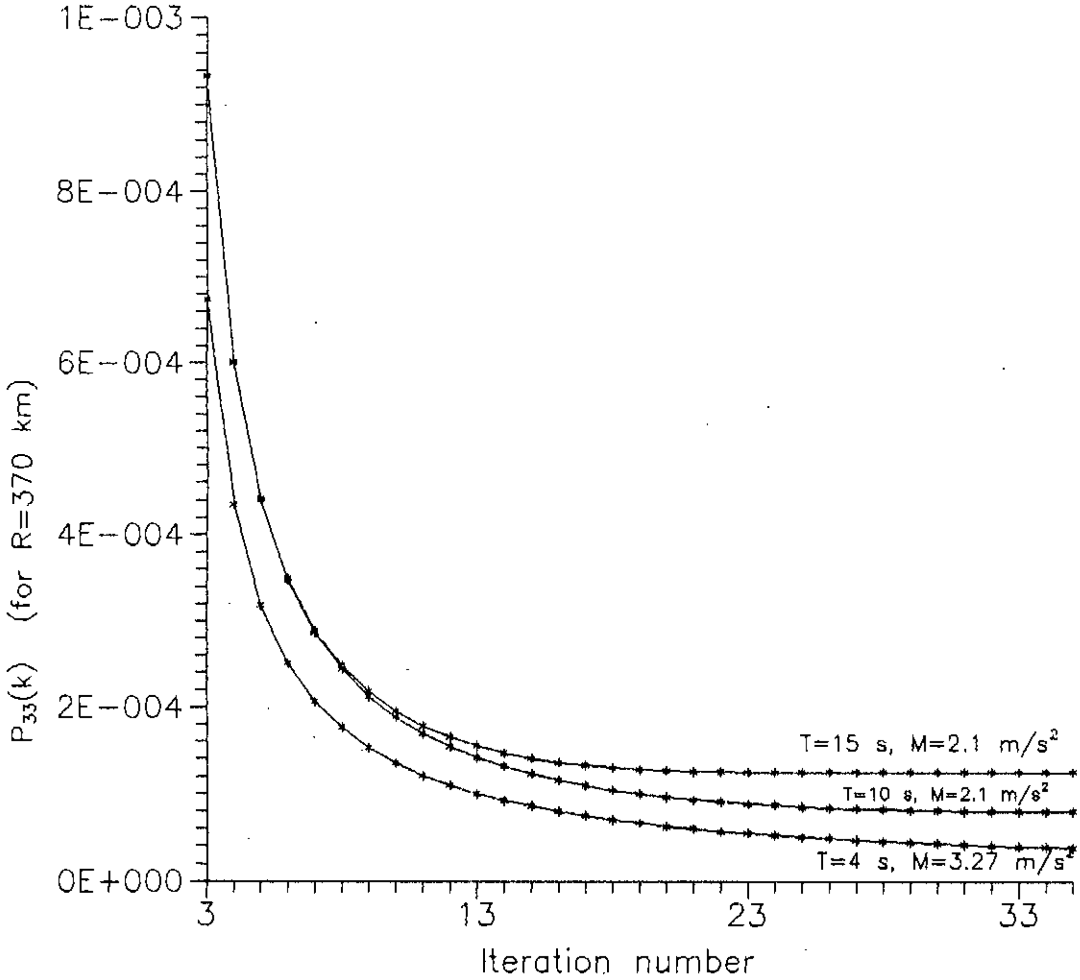

Fig.(7.13) Mean-square bearing prediction error (for different scan time)

CHAPTER 8

CONCLUSIONS AND RECOMMENDATIONS

8.1 Conclusions

In chapter 2 we focus on the detection of the presence of a signal in noise and in doing so we used the Neyman-Pearson criterion to assess the performance of air traffic control radar where a priori probability of receiving a signal is not known. From the performance curve as shown in fig.(2.6), wc have found that for the probability of error P_e less than 10^{-4}, we need $\sqrt{m}A/\sigma_v > 7.3$. For $\sqrt{m}A/\sigma_v = 10$, $P_e = 2.87\text{x}10^{-6}$, for $\sqrt{m}A/\sigma_v = 11$, $P_e = 1.9\text{x}10^{-7}$ and for $\sqrt{m}A/\sigma_v = 12$, $P_e = 1.0\text{x}10^{-9}$. So small changes in A/σ_v, the signal-to-noise ratio, result in significant changes in the probability of error. For this reason, increasing the signal-to-noise ratio are so important in signal processing systems, i.e. for the improvement of detection capability we should make A/σ_v as large as possible. To do this the number of samples m may also be increased. Since the samples are spaced a fixed time interval apart to ensure statistical independence, the time for detection increases with m.

The signal detection characteristics of the primary and secondary radars of **Zia International Airport** are also described. From the results given in chapter 2, it is seen that since the secondary radar is equipped with the transponder, radar range extensively increase with lower transmitter power as the transponder acts as a midway receiving transmitting station. We have found that though the primary radar has 1400 times the transmitted power of the secondary radar, the secondary radar can cover almost 2.3 times longer distance that of the primary radar as the secondary radar is equipped with the transponder. As the wider pulse used for the secondary radar, target detection capability of it increase correspondingly since for the wider pulse more energy is received. At the same time minimum radial spacing of two targets required to distinguish them also increase. The signal-to-noise ratio, E/n_0 as calculated for the two radars are

23.43dB for primary radar and 82.24dB for secondary radar (at the transponder) for maximum radial distances.

Both the primary and the secondary radar rotate at a rate of 15 rpm, completing a 360° scan in 4s. The two radar beams have azimuth bandwidths of 1.5° and 1.35°, respectively, while the number of pulses emitted are 633 and 316 pulses per second, respectively. This corresponds to 10.55 pulses reflected per target in the primary radar and 4.7 pulses per target in the secondary radar, providing that many more samples in the signal processing or, equivalently, augmenting the received energy by the same amount. Thus the primary radar energy received should be 10.55E and the secondary radar energy 4.7E. Alternately, this results in E/n_0 improvement, of 10.23dB and 6.72dB, respectively.

Fig.(2.10) represents a family of graphs relating probability of detection with signal-to-noise ratio for different values of probability of false-alarm. The signal-to-noise ratio needed to achieve a specified probability of detection without exceeding a specified false-alarm probability can be calculated from these graphs. It is found that for small values of false-alarm probability we need large signal-to-noise ratio to achieve a specific probability of detection, i.e. to maintain a certain probability of detection, the required value of signal-to-noise ratio increases correspondingly with decease of false-alarm probability. From the graphs we have seen that to attain maximum probability of detection, P_d for specified values of P_n (the probability of false-alarm) the required signal-to-noise ratios are as follows:

For the maximum P_d		
P_n	E/n_0	E/n_0 (in dB)
10^{-2}	14	11.46
10^{-4}	21	13.22
10^{-6}	28	14.47
10^{-8}	35	15.44
10^{-10}	40	16.02
10^{-12}	65	18.13

We have also established a frame work for the analysis of random signals using Kalman filter which is applied in radar tracking problem in chapter 7. We have found that the Kalman filter provides the best linear estimates of radar tracking. The resulting performance figures are presented, these being the mean-square range prediction error, mean-square bearing prediction error and gain setting of Kalman filter. In figs. 7.2 to 7.13 we have plotted the results of computer calculations of the appropriate matrices for the set of data obtained from the radars at **Zia International Airport** and for other relevant sets of data that explain the characteristics of the processor elaborately. We have also presented an example for the range R=160 km, scan time T=15 sec, maximum acceleration M=3.27 m/sec^2 and the probabilities of maximum and zero acceleration P_1 and P_2 to be equal to 0.1 and 0.3 respectively which explain the computational procedures of the filter. We have plotted only the first diagonal element, $K_{11}(k)$, of the gain matrix and the first and third diagonal elements of the one-step prediction covariance matrix $P(k+1|k)$ which correspond to the mean-square errors in the prediction of range and bearing.

The mean-square bearing prediction error converges to its steady-state value in 8 to 12 iterations, while the range prediction error converges in 5 to 12 iterations. The gain settings of the Kalman filter reach to its steady state value after 5 to 10 iterations. The variation of Kalman filter gain setting, the mean-square range prediction error and the mean-square bearing prediction error for different values of the parameters of the model are presented both in graphical and numerical form so that we can compare the results.

Graphs 7.2-7.5 show the gain settings of the Kalman filter, the first diagonal element of the gain matrix K(k) and we have found that the gain setting drops quickly to a final steady state value. It can be seen that gain setting decreases with the decrease of scan time. But this value increases with the increase of maximum acceleration. The results are very similar for primary and secondary radar.

Graphs 7.6-7.9 show the complete results of the mean-square range prediction error and we have found that mean-square range prediction error increases correspondingly with the increase of maximum acceleration M for the same scan time (here T=4s). For higher value of scan time (keeping σ_1^2 constant), the error also increases. Both the cases are shown in graphs.

Graphs 7.10-7.13 show the variation of the mean-square bearing prediction error for those sets of data. It is apparent from these graphs that it follow trend that is quite reasonable since the error decreases with the decrease of T for constant σ_1^2 and increases with M.

From these results we can say that this algorithm (i.e. Kalman filter) is suitable for radar tracking problem. We have seen that as the scan time T is decreased from 15 to 4 s, for the same σ_1^2, the gain setting and the mean-square errors decrease. Since the time between filter iterations has been decreased, the variation in range and bearing estimates decreases as well. Therefore, new estimates carry less information and can be weighted less. If the maximum acceleration M is increased (increasing σ_1^2) for the same scan time T, the mean-square errors and gain go up again, since then range and bearing estimates will vary more from iteration to iteration.

Finally, as a general conclusion, the Kalman filter provides the minimum variance estimate of the process even in a noisy observation case, and provides the best balance between the a priori information and measurements, and hence it is an efficient estimator for the state of estimation problem. The Kalman filter is best suited to analysis of continuous time series and it is therefore can be applied in areas such as radar tracking, navigation systems, weather forecasting, satellite communications, automatic traffic control system etc.

8.2 Suggestions for Further Study

So far we discussed the application of Kalman filtering to air traffic control. Also, throughout this thesis we have restricted our study to simple radar tracking problem, i.e. we have used a simple and moderate tracking method. We have found that in the linear target tracking system, the Kalman filter can be employed directly. But in the nonlinear tracking problem, some information about the target is lost because of the linearisation and approximation procedures required to make the problem trackable. When the measurement frequency is much lower than the error bandwidth, the successive errors are essentially uncorrelated and can be treated as white noise. This method assumes that the measurement noise is white. But in practice, the measurement noise may be sequentially correlated, and this is often referred to as colored noise.

For sequentially correlated measurement noise, the parameters can not be estimated with sufficient accuracy to obtain the desired tracking performance. By reformulating the measurement equation, the noise may be decorrelated so that the Kalman filter can be directly applied. By considering the effect of noise correlation, one can undertake a task of designing a modified computationally efficient method.

Many important publications are available in the references. The reader is invited to consult references [1-3], [7-10] and [16] for complementary information on radar tracking problem, and references [2,9,25] for random signal processing.

BIBLIOGRAPHY

[1] Schwartz, M. & Shaw, L. "Signal Processing - Discrete Spectral Analysis, Detection, and Estimation" McGraw Hill Book Company, 1975

[2] Bozic, S. M. "Digital and Kalman Filtering - An Introduction to Discrete-Time Filtering and Optimum Linear Estimation" Edward Arnold (Publishers) Ltd., 1979

[3] McLaughlin, S. "Adaptive equalization via Kalman filtering techniques" IEE Proceedings-F, Vol.138, No. 4, pp. 388-396, August 1991

[4] Wang, T. C. & Varshney, P. K. "Measurement preprocessing for nonlinear target tracking" IEE Proceedings-F, Vol.140, No. 5, pp. 316-322, October 1993

[5] Chan, Y. T. & Couture, F. "Manoeuvre detection and track correction by input estimation" IEE Proceedings-F, Vol.140, No. 1, pp. 21-27, February 1993

[6] Munu, M., Harrison, I., Wilkin, D. & Woolfson, M. S. "Comparison of adaptive target-tracking algorithms for phased-array radar" IEE Proceedings-F, Vol.139, No. 5, pp. 336-342, October 1992

[7] Skolnik, M. I. "Introduction to Radar Systems" McGraw Hill Book Company, 1980

[8] Chaumette, E., Comon, P. & Muller, D. "ICA-based technique for radiating sources estimation: application to airport surveillance" IEE Proceedings-F, Vol. 140, No. 6, pp. 395-401, December 1993

[9] Candy, J. M. "Signal Processing - The Model Based Approach" McGraw Hill Book Company, 1986

[10] Song, L. & Uhm, T. S. "Multiplicative noise model and composite signal detection" IEE Proceedings-F, Vol. 138, pp. 531-534, December 1991

[11] Feijoo, C., Asensio, A. & Perez, F. "New practical method for measurement error determination in radar systems under real traffic conditions" IEE Proceedings- F, Vol. 138, No. 6, pp. 525-530, December 1991

[12] Oppenheim, A. V. & Schafer, R. W. "Digital Signal Processing" Prentice Hall, 1975

[13] Schwartz, M. "Information Transmission, Modulation and Noise" McGraw Hill Book Company, 1990

[14] Felhauer, T. "Digital Signal Processing for optimum wideband channel estimation in the presence of noise" IEE Proceedings-F, Vol. 140, No. 63 pp. 179-185, June 1993

[15] Schwab, C. E. & Rost, D. P. "Airport Surface Detection Equipment" Proc. IEEE, Vol.138, pp. 388-396, August 1991

[16] Berkowitz, R. S. "Modern Radar - Analysis, Evaluation and System Design" John Wiley & Sons Inc., 1965

[17] Anderson, S.J. "Adaptive remote sensing with HF skywave radar" IEE Proceedings-F, Vol.139, No.2, pp. 182-191, april 1992

[18] Shrader, W. W. "Radar Technology Applied to Air Traffic Control" IEEE Trans. Commun., vol.- COM 21, May '73

[19] Farina, A. & Studer, F. A. "Radar Data Processing (Vol. I, Introduction and Tracking)" John Wiley & Sons Inc., 1985

[20] Guu, J.A. & Wei, C. H. "Tracking technique for manoeuvering target with correlated measurement noises and unknown parameters" IEE Proceedings-F, Vol. 138, No. 3, pp. 278-286, June 1991

[21] Dall, J., Jorgensen, J. H., Christensen, E. L. & Madsen, S. N. "Real-time Processor for the Danish airborne SAR" IEE Proceedings-F, Vol. 139, No.2, pp. 115-121, April 1992

[22] Dogan, M. C. & Mendel, J. M. "Single sensor detection and classification of multiple sources by higher-order spectra" IEE Proceedings-F, Vol. 140, No. 6, pp. 350-354, December 1993

[23] Taylor, J. W. Jr. & Brunins, G. "Design of a New Airport Surveillance Radars" Proceedings of the IEEE, Vol. 73, pp. 284-289, February 1985

[24] Reilly, J. P. & Snelling, W. E. "Short-range sea clutter in horizon search radar systems: Antenna pattern effects" IEE Proceedings-F, Vol. 140, No. 5, pp. 309-315, October 1993

[25] Gold, B. & Rader, C.M. "Digital Processing of Signals" McGraw Hill Book Company, 1969

[26] Haykin, S. S. "A Unified Treatment of Recursive Digital Filtering" IEEE Trans. on Automatic Control, pp. 113-116,

[27] Currie, A. & Brown, M. A. "Wide-swath SAR" IEE Proceedings-F, Vol. 139, No. 2, pp. 122-125, April 1992

[28] Ito, K. "Ranging and positioning experiments for aircraft using Engineering Test Satellite V" IEE Proceedings-F, Vol. 138, No. 6, pp. 571-575, December 1991

[29] Tugnait, J. K. "Two-channel tests for common non-gaussian signal detection" IEE Proceedings-F, Vol. 140, No. 6, pp. 343-345, December 1993

[30] Lim, C. H. & Lee, H. S. "Performance of order-statistics CFAR detector with noncoherent integration in homogeneous situations" IEE Proceedings-F, Vol. 140, No. 5, pp. 291-300, October 1993

[31] Honold, P. "Secondary Radar-Fundamentals and Instrumentation" Siemens AG., Berlin and Heyden, M. & Son Ltd., London, 1976