ROBUST MINIMAX ESTIMATION APPLIED TO KALMAN FILTERING

A THESIS

SUBMITTED TO THE DEPARTMENT OF ELECTRICAL AND

ELECTRONICS ENGINEERING

AND THE INSTITUTE OF ENGINEERING AND SCIENCES

OF BILKENT UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF SCIENCE

By

Bahadır Aybar

September 2008

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Prof. Dr. Orhan Arıkan(Supervisor)

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Asst. Prof. Dr. Sinan Gezici

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Asst. Prof. Dr. İbrahim Körpeoğlu

Approved for the Institute of Engineering and Sciences:

Prof. Dr. Mehmet Baray
Director of Institute of Engineering and Sciences

ABSTRACT

ROBUST MINIMAX ESTIMATION APPLIED TO KALMAN FILTERING

Bahadır Aybar
M.S. in Electrical and Electronics Engineering
Supervisor: Prof. Dr. Orhan Arıkan
September 2008

Kalman filtering is one of the most essential tools in estimating an unknown state of a dynamic system from measured data, where the measurements and the previous states have a known relation with the present state. It has generally two steps, prediction and update. This filtering method yields the minimum mean-square error when the noise in the system is Gaussian and the best linear estimate when the noise is arbitrary. But, Kalman filtering performance degrades significantly with the model uncertainty in the state dynamics or observations. In this thesis, we consider the problem of estimating an unknown vector x in a state-space model that may be subject to uncertainties. We assume that the model uncertainty has a known bound and we seek a robust linear estimator for x that minimizes the worst case mean-square error across all possible values of x and all possible values of the model matrix. Robust minimax estimation technique is derived and analyzed in this thesis, then applied to the state-space model and simulation results with different noise perturbation models are presented. Also, a radar tracking application assuming a linear state dynamics is also investigated.

Modifications to the James-Stein estimator are made according to the scheme we develop in this thesis, so that some of its limitations are dealt with. In our scheme, James-Stein estimation can be applied even if the observation equation is perturbed and the number of observations are less than the number of states, still yielding robust estimations.

Keywords: Mean-squared error estimation, minimax estimation, robust estimation, Kalman filter, maximum likelihood estimation, James-Stein estimation

ÖZET

GÜRBÜZ MİNİMUM-MAKSİMUM KESTİRİCİNİN KALMAN FİLTRESİNE UYARLANMASI

Bahadır Aybar
Elektrik ve Elektronik Mühendisliği Bölümü Yüksek Lisans
Tez Yöneticisi: Prof. Dr. Orhan Arıkan
Eylül 2008

Kalman filtreleme tekniği, ölçümlerin ve önceki durum değişkenlerinin şimdiki durum değişkenine bağlı olduğu dinamik bir sistemdeki durum değişkenlerini kestirme konusunda en önemli uygulamalardan biridir. Genel olarak tahmin ve güncelleme olarak iki aşamalıdır. Bu filtreleme, sistemdeki gürültü normal dağılıma sahipken asgari ortalama-kare hatası verirken, sistemde rastgele bir gürültü varken ise en iyi çizgisel tahmin sağlamaktadır. Fakat, sistem modelinde bir bilinmezlik olduğu koşulda Kalman filtreleme tekniği performansı oldukça düşmektedir. Bu tezde, belirsizliğe uğramış bir durum-uzay modelindeki bilinmeyen x vektörünün kestirilmesi problemi ele alınmıştır. Model belirsizliğinin bilinir bir sınırı olduğu kabul edilmiş, model matrisinin ve bilinmeyen x vektörünün olası her değeri çerçevesinde en kötü durumdaki ortalama-kare hatasını asgari düzeye çeken birinci dereceden gürbüz bir kestirici araştırılmıştır. Bu gürbüz minimum-maksimum kestirici, değişik gürültü modelleri ile birlikte durum-uzay modeline uygulanmış ve simülasyon sonuçları verilmiştir. Ek olarak, çizgisel bir dinamik modele sahip olduğu kabul edilmiş olan bir radar takip uygulaması incelenmiştir.

Bu tezde verilmiş olan şema çerçevesinde, James-Stein kestirme tekniğine değişiklikler yapılmış, böylece bu tekniğin bir takım kısıtlamalarının önüne geçilmiştir. Bu şemada, gözlem denkleminin belirsizliğe uğradığı ve gözlemlerin durum değişkenlerinin sayısından az olduğu durumlara da James-Stein tekniği uygulanmış ve gürbüz tahminler verdiği görülmüştür.

Anahtar Kelimeler: Ortalama-kare hata kestirme, minimum-maksimum kestirme, gürbüz kestirici, Kalman filtesi, azami yaklaşım kestirici, James-Stein kestiricisi

ACKNOWLEDGMENTS

I gratefully thank my supervisor Prof. Dr. Orhan Arikan for his supervision, guidance, suggestions and instructive comments throughout the development of this thesis.

I also would like to thank Asst. Prof. Dr. Sinan Gezici and Asst. Prof. Dr. İbrahim Körpeoğlu for accepting to read and review the thesis.

Finally I express my special thanks to my family for their permanent support and sincere love.

Contents

List of Figures

List of Tables

Dedicated to My Family. . .

Chapter 1

INTRODUCTION

In this thesis, state estimation from the state-space equation below is investigated:

$$\begin{aligned} x_{k+1} &= A_k x_k + B_k w_k \\ y_k &= C_k x_k + D_k v_k, \end{aligned} \tag{1.1}$$

where x vector denotes the states, y vector denotes the observations, w and v are noise vectors and A, B, C and D matrices are the parameters of the state-space equations. These equations are known as the state equation and the observation equation, respectively.

1.2 Review of Kalman Filtering Methods and Applications

Obtaining a minimum mean-square error (MSE) estimate for the state vector x is a very frequently confronted problem in control systems and avionics. Rudolf E. Kalman delivered a solution for this estimation problem which yields the minimum mean-square error estimate when the system is exposed to Gaussian noise and the best linear estimate when the noise is non-Gaussian [1]. Many applications involve this kind of estimation problem, such as target tracking [2, 3, 4]. A vehicle with varying speed and acceleration has to be tracked and where the system dynamics can be formulated as (1.1). Kalman filtering techniques are successfully used in the control of induction motors and motors with many poles such as hybrid stepper motors [5, 6]. In these applications, observation equation can be nonlinear. In such cases, a modified version of Kalman filter, extended Kalman filter, can be used in order to linearize the nonlinear observation so that Kalman filtering technique can be applied [7, Chapter 9]. Speech enhancement via Kalman filter is also an important application area of Kalman filtering [8, 9]. Also in [10], an adaptive Kalman filtering approach is used for the equalization of digital communication channels. Some other applications are detailed in [11, 12, 13, 14, 15, 16, 17, 18]. As an alternative, an interactive multiple model consisting of multiple weighted Kalman filters for tracking manoeuvring targets are examined in [19, 20].

In real applications, there can be some parameter uncertainty on the matrices that describe that system. In such cases, Kalman filtering results in high mean-square error if it uses nominal values for the unknown or perturbed parameters. Various approaches have been developed to overcome the problem of robustness of the Kalman filtering. It is shown in [21] that steady-state solution to robust Kalman filtering problem is related to two algebraic Riccati equations (ARE's). An estimation technique with guaranteed cost bound for linear systems with parameter uncertainties were proposed in [22] where the parameter uncertainty model is assumed to be dependent on state and measurement noises. Also in [23], robust Kalman filtering problem for linear continuous-time systems with parametric uncertainty only in the state matrix was considered. An alternative approach is developed in [24], where another robust discrete-time minimum variance filtering tehcnique is introduced. Also, in [25, 26, 27, 28], some other robust Kalman filtering techniques are detailed.

1.3 Robust Minimax Estimation Applied To Kalman Filtering

Consider the problem of estimating x from the observations $y = Hx + w$ where H is the observation matrix and w is the process noise. Suppose that the main concern is to minimize the worst-case mean-square error rather than the average mean-square error. In [29], a robust mean-squared error estimation is developed in the presence of parameter uncertainty in the observation matrix H. It is shown in [29] that for an arbitrary choice of weighting, the optimal minimax MSE estimator can be formulated as a solution to a semidefinite programming problem (SDP), which can be solved very efficiently.

We combined the robust minimax estimation technique in [29] with the Kalman filtering in order to estimate the states of the dynamics in (1.1). Thus, for the cases where the worst-case mean-square error is the main concern, robust minimax technique can be applied to the state-space equations in (1.1).

1.4 Organization of the Thesis

The thesis is organized as follows: The next chapter involves the main steps of the Kalman filtering and its derivations, simulation results and its disadvantages in the presence of parameter uncertainty. Chapter 3 describes two methods that are developed to overcome the problem of robust Kalman filtering and their comparisons with simulation results. In Chapter 4, we develop robust minimax estimation and its application to the Kalman filtering with simulation results. Finally, the thesis is concluded in Chapter 5.

Chapter 2

BACKGROUND AND PROBLEM DEFINITION

2.1 Kalman Filtering

In this section of the thesis, a brief introduction to Kalman filtering, its formulations and some of its properties will be stated and the notations will be the same as [7, Chapter 8]. Kalman filtering is an essential and widely used tool for the estimation of the states of a dynamic system which is exposed to noise because of the fact that it yields the minimum mean square error. It has been used in many applications such as radar tracking, econometrics, motor driving, flight control and color noise problems as stated in the previous section. It was first stated by Rudolf E. Kalman in 1960 [1].

The main scheme that Kalman filter applies includes a state equation and an observation equation distorted by Gaussian noise at each time instants. The discrete time state equations are:

$$x_{k+1} = A_k x_k + B_k w_k \tag{2.1}$$

$$y_k = C_k x_k + D_k v_k. \tag{2.2}$$

The terms and notations in equations (2.1) and (2.2) are:

x_k = (n x 1) process state vector at time t_k

y_k = (m x 1) observation vector

w_k = (n x 1) zero-mean white Gaussian noise vector with known covariance

v_k = (m x 1) zero-mean white Gaussian noise vector with known covariance and having no crosscorrelation with w_k

A_k = (n x n) state transition matrix

B_k = (n x n) state noise matrix at time t_k

C_k = (m x n) observation matrix at time t_k

D_k = (m x m) observation noise matrix at time t_k.

In this thesis, we denote the noise statistics as $w_k \sim N(0,Q_k)$ and $v_k \sim N(0,R_k)$. Once these conditions are met and the system is completely described by (2.1) and (2.2), then the Kalman filter, which will be detailed next, yields the minimum mean square error estimates for the state x_k at each time step.

The covariance matrices of w_k and v_k are given by:

$$E[w_k w_i^T] = \begin{cases} Q_k, \text{ for } k = i \\ 0, \text{ for } k \neq i. \end{cases} \tag{2.3}$$

$$E[v_k v_i^T] = \begin{cases} R_k, \text{ for } k = i \\ 0, \text{ for } k \neq i. \end{cases} \tag{2.4}$$

$$E[w_k v_i^T] = 0, \text{ for all } k \text{ and } i. \tag{2.5}$$

The time index, k, usually starts at k=0 and we assume we have an initial estimate of the state sequence x_k at $t = t_0$ as $\hat{x}_0$. Then we define *a priori* estimate, which will be denoted as $\hat{x}_{k|k-1}$, representing the best mean-squared error (MSE) estimate prior to taking x_k into account. At this point, we define the *estimation error* as

$$e_{k|k-1} = x_k - \hat{x}_{k|k-1}. \tag{2.6}$$

In order to find the covariance matrix associated with the above estimation error, we investigate

$$P_{k|k-1} = E[e_{k|k-1}\, e_{k|k-1}^T] = E[(x_k - \hat{x}_{k|k-1})\, (x_k - \hat{x}_{k|k-1})^T], \tag{2.7}$$

assuming $e_{k|k-1}$ has zero mean.

With this assumption of the a priori estimate $\hat{x}_{k|k-1}$, we use the observation y_k at time $t = t_k$ to improve the a priori estimate. Now, we form a linear combination of the a priori estimate and the noisy measurement based on the following equation:

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k(y_k - C_k\hat{x}_{k|k-1}), \tag{2.8}$$

where $\hat{x}_{k|k}$ is the a posteriori (updated) estimate and K_k is the Kalman gain.

As we mentioned before, Kalman filtering is optimal in the mean-square sense, since the Kalman gain K_k is chosen such that the a posteriori estimate has the minimum mean square error. To obtain the expression for the error covariance matrix associated with the a posteriori estimate, we investigate

$$P_k = E[e_k\, e_k^T] = E[(x_k - \hat{x}_{k|k})\, (x_k - \hat{x}_{k|k})^T]. \tag{2.9}$$

As we substitute Eq. (2.2) into Eq. (2.8), then

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k(C_k x_k + D_k v_k - C_k\hat{x}_{k|k-1}). \tag{2.10}$$

Now, substituting the above result into Eq. (2.9), then we obtain

$$\begin{aligned} P_k = E\{&\big[x_k - \hat{x}_{k|k-1} - K_k(C_k x_k + D_k v_k - C_k\hat{x}_{k|k-1})\big] \\ &\big[x_k - \hat{x}_{k|k-1} - K_k(C_k x_k + D_k v_k - C_k\hat{x}_{k|k-1})\big]^T\}. \end{aligned} \tag{2.11}$$

After rearrangements, the equation becomes

$$P_k = E\{\big[(I - K_kC_k)(x_k - \hat{x}_{k|(k-1)}) - K_kD_kv_k\big]$$
$$\big[(I - K_kC_k)(x_k - \hat{x}_{k|(k-1)}) - K_kD_kv_k\big]^T\}. \tag{2.12}$$

We can further simplify the above expectation by using Eq. (2.7) and by setting $E[v_kv_k^T] = R_k$, $E[v_k(x_k - \hat{x}_{k|(k-1)})] = 0$ since the a priori estimation error is uncorrelated with the measurement noise v_k. Then, the error covariance matrix can be given as:

$$P_k = (I - K_kC_k)P_{k|k-1}(I - K_kC_k)^T + K_kD_kR_kD_k^TK_k^T, \tag{2.13}$$

which is the most general expression of the a posteriori error covariance matrix. Since the diagonal elements of the matrix P_k gives the estimation error variances of each element in the state vector, we need to minimize the diagonal elements of P_k, which is the same as minimizing $trace(P_k)$, with respect to K_k in order to obtain minimum mean-square error. We will use differential calculus methods for this minimization problem. Here are the needed matrix differentiation formulas:

$$\frac{d[trace(AB)]}{dA} = B^T \quad \text{where AB is a square matrix} \tag{2.14}$$
$$\frac{d[trace(ACA^T)]}{dA} = 2AC \quad \text{where C is a symmetric matrix,}$$

where the derivative of a scalar with respect to a matrix is defined as

$$\frac{ds}{dA} = \begin{bmatrix} \frac{ds}{da_{11}} & \frac{ds}{da_{12}} & \cdots & \frac{ds}{da_{1n}} \\ \frac{ds}{da_{21}} & \frac{ds}{da_{22}} & \cdots & \frac{ds}{da_{2n}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{ds}{da_{n1}} & \frac{ds}{da_{n2}} & \cdots & \frac{ds}{da_{nn}} \end{bmatrix}. \tag{2.15}$$

For completeness, the proof of these equations are given in Appendix A of the thesis. Thus, we can minimize $trace(P_k)$ with respect to K_k if P_k is linear or quadratic in K_k. In order to use the given matrix differentiation formulas, Eq. (2.13) can be rearranged as:

$$P_k = P_{k|k-1} - K_kC_kP_{k|k-1} - P_{k|k-1}C_k^TK_k^T + K_k(C_kP_{k|k-1}C_k^T + D_kR_kD_k^T)K_k^T. \tag{2.16}$$

One can see that the first term is independent of K_k, the second and third terms are linear in K_k and the last term is quadratic in K_k. Now, using the argument that minimizing $trace(P_k)$ gives us the minimum sum of the mean-square errors of individual elements in the state vector, let us evaluate

$$\frac{d(trace(P_k))}{dK_k} = -2(C_k P_{k|k-1})^T + 2K_k(C_k P_{k|k-1} C_k^T + D_k R_k D_k^T), \tag{2.17}$$

since $C_k P_{k|k-1} C_k^T + D_k R_k D_k^T$ is a symmetric matrix and every square matrix and its transpose has the same trace. Equating the above differential to zero, we obtain the optimal gain as

$$K_k = P_{k|k-1} C_k^T (C_k P_{k|k-1} C_k^T + D_k R_k D_k^T)^{-1}. \tag{2.18}$$

The above representation yields the most general form of Kalman gain equation. The covariance matrix associated with the optimal estimate can be computed by substituting Eq. (2.18) into Eq. (2.16). After some arrangements,

$$\begin{aligned} P_k &= P_{k|k-1} - P_{k|k-1} C_k^T (C_k P_{k|k-1} C_k^T + D_k R_k D_k^T)^{-1} C_k P_{k|k-1} \\ &= P_{k|k-1} - K_k C_k P_{k|k-1} \\ &= (I - K_k C_k) P_{k|k-1}, \end{aligned} \tag{2.19}$$

which gives the relation between the covariance of the optimal estimate and the covariance of the a priori estimate. Now, since we have the optimal Kalman gain expression, we can use the Eq. (2.8) to compute the estimated state at time k from the a priori estimated state and the measurement at time k. As we see from the Eq. (2.8), we need $\hat{x}_{k|k-1}$ and $P_{k|k-1}$ for $\hat{x}_{k|k}$. Since w_k in Eq. (2.1) is uncorrelated with w's at any other time and zero mean, we can project $\hat{x}_{k-1|k-1}$ by a simple transition matrix A_k,

$$\hat{x}_{k+1|k} = A_k \hat{x}_{k|k}, \tag{2.20}$$

which shows the relation of the propagation that the state estimates have. Similarly, we should formulate the relation of the propagation that the error covariance has. The error covariance matrix associated with $\hat{x}_{k+1|k}$ is

$$P_{k+1|k} = E[e_{k+1|k} {e_{k+1|k}}^T] = E[(x_{k+1} - \hat{x}_{k+1|k})(x_{k+1} - \hat{x}_{k+1|k})^T]$$

$$
\begin{aligned}
&= E[((A_k x_k + B_k w_k) - A_k \hat{x}_{k|k})((A_k x_k + B_k w_k) - A_k \hat{x}_{k|k})^T] \\
&= E[(A_k e_k + B_k w_k)(A_k e_k + B_k w_k)^T] \\
&= A_k P_k A_k^T + B_k Q_k B_k^T. \qquad (2.21)
\end{aligned}
$$

The Eq. (2.21) can be combined with the Eq. (2.19) to obtain a relation between the a priori and a posteriori error covariance matrices:

$$
P_{k+1|k} = A_k(I - K_k C_k) P_{k|k-1} A_k^T + B_k Q_k B_k^T. \qquad (2.22)
$$

At this point, we have $P_{k+1|k}$ from Eq. (2.22), z_{k+1} from the measurements, $\hat{x}_{k+1|k}$ from Eq. (2.20), thus we can compute the new state estimate $\hat{x}_{k+1|k+1}$ by using Eq. (2.8). To summarize, there are four main Kalman filtering equations:

$$
\begin{aligned}
\hat{x}_{k|k} &= \hat{x}_{k|k-1} + K_k(y_k - C_k \hat{x}_{k|k-1}) \qquad (2.23) \\
K_k &= P_{k|k-1} C_k^T (C_k P_{k|k-1} C_k^T + D_k R_k D_k^T)^{-1} \qquad (2.24) \\
\hat{x}_{k+1|k} &= A_k \hat{x}_{k|k} \qquad (2.25) \\
P_{k+1|k} &= A_k(I - K_k C_k) P_{k|k-1} A_k^T + B_k Q_k B_k^T. \qquad (2.26)
\end{aligned}
$$

This set of Kalman filtering equations provides optimal mean-squared estimator when the noises w_k and v_k in state-space equations (2.1) and (2.2) are Gaussian. It also provides best linear mean-squared estimator for the non-Gaussian case. Commonly, Kalman filter is implemented based on the architecture shown in the following block diagram [7].

Prior estimate $\hat{x}_{k|k-1}$ and its error covariance $P_{k|k-1}$ for $k = 0$ are the initial assumptions of the Kalman filter. Then, the algorithm computes the Kalman gain K_k using Eq. (2.24). Using this Kalman gain, the prior estimate is updated according to the Eq. (2.23) and error covariance for the updated estimate is obtained as $P_k = (I - K_k C_k) P_{k|k-1}$. Finally, the state estimate and its error covariance is projected ahead using (2.25) and (2.26).

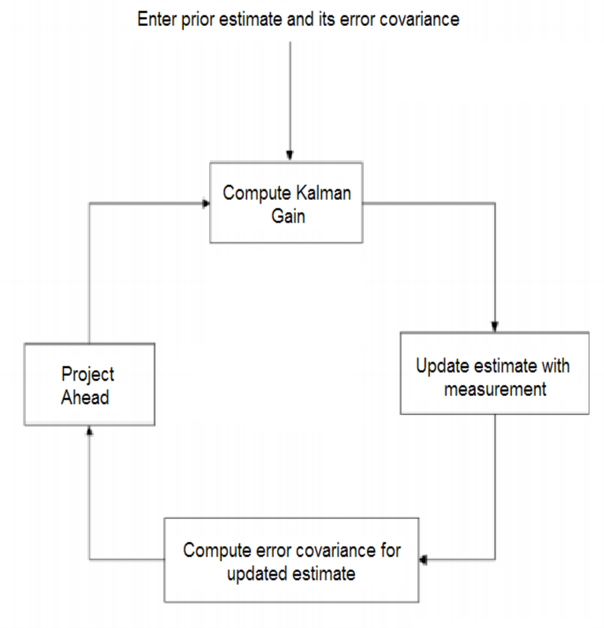

Figure 2.1: Kalman filter architecture.

2.2 Perturbation Model

As already discussed, the Kalman filter is optimal in the mean-square sense for linear state-space systems. One of the crucial assumptions underlying the performance of the Kalman filter is that the state-space model parameters are known precisely. If this is not the case and there are uncertainties in the state-space model, then it is observed that the Kalman filter may degrade significantly. Therefore, in applications where state-space model parameters are not known precisely, we are at risk in using the standard Kalman filter. Since this is the case in many applications including radar target tracking, a robust version of the Kalman filter should be used.

In this thesis, two different perturbation models will be used for making comparisons between the alternatives of robust Kalman filter. First model represents the case when all the entries of the matrix A_k in Eq (2.1) are subjected to some independent and identically distributed Gaussian noise with known distribution.

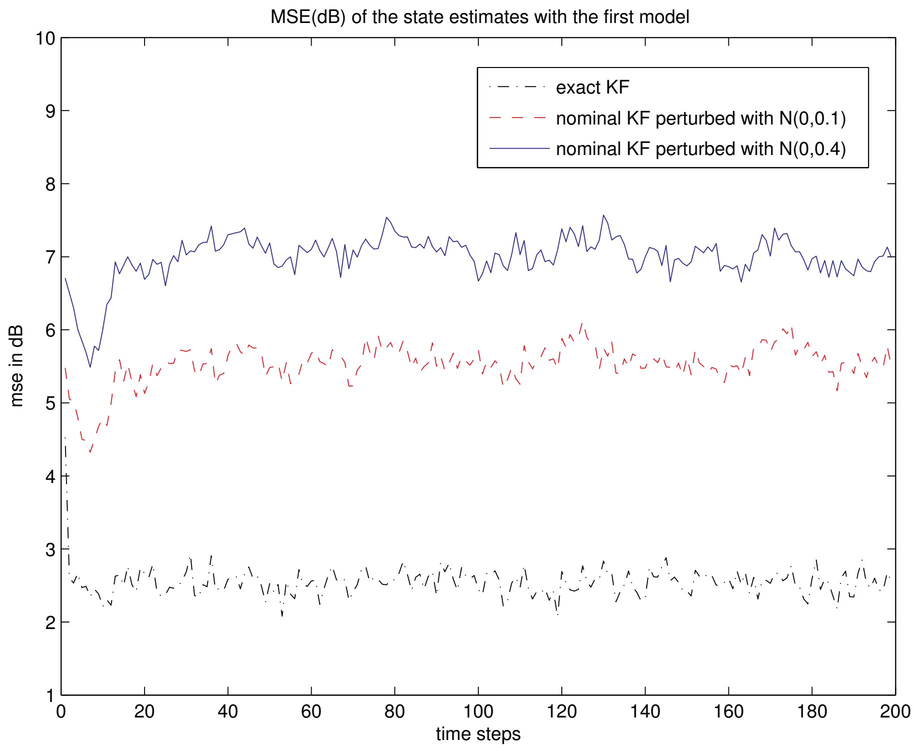

Figure 2.2: Kalman filter performances in dB when there is no parameter perturbation, when parameters are perturbed with $N(0, 0.1)$ and with $N(0, 0.4)$.

In the second model, we will assume that one of the entries of A_k has uncertainty on it, whereas the rest is exactly known.

For the simulation of the first model, we use the scheme in [28] as

$$A_k = \begin{bmatrix} 1.0 & -0.1 & -0.1 \\ 0.2 & 0.9 & -0.1 \\ 0.1 & 0.2 & 0.7 \end{bmatrix},$$

$$B_k = C_k = D_k = Q_k = R_k = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{2.27}$$

where the A matrix is exposed to a Gaussian noise with N(0,0.1) and N(0,0.4).

We take the average mean-square error of 1000 iterations for the first 200 data points (i.e. time steps) in dB. Thus, the formulation for the performance criteria is $10log(\sum_{k=1}^{N_r} J_k/N_r)$, where J_k is the risk at each run at time k, N_r is the

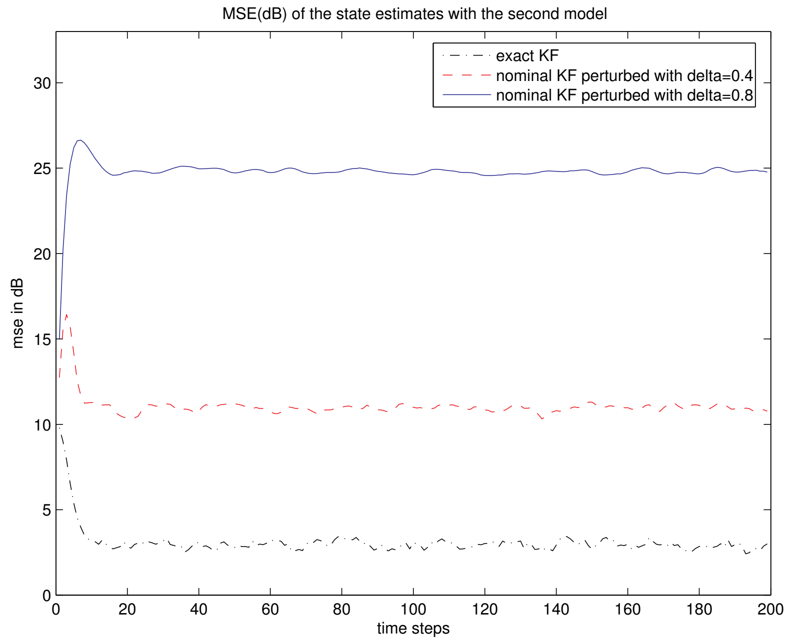

Figure 2.3: Kalman filter performances in dB when there is no parameter uncertainty, when $\delta = 0.4$ and when $\delta = 0.8$.

number of runs. This average mean-square errors for each time steps are drawn in Figures (2.2) and (2.3). In Fig. (2.2), the performance degradation can be observed in dB scale. It can be seen that as the perturbation in A increases, the Kalman filter degrades significantly.

As the second model, we use the following state-space parameters

$$\begin{aligned} A_k &= \begin{bmatrix} 0 & -0.5 \\ 1 & 0.6+\delta \end{bmatrix}, \quad B_k = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \\ C_k &= \begin{bmatrix} -100 & 10 \end{bmatrix}, \quad D_k = 1, \end{aligned} \tag{2.28}$$

where δ has a known bound. A close variant of this model was also used for a similar purpose in [24].

The figure (2.3) shows the degradation of the Kalman filter when the states are estimated with the exact knowledge about the parameters, with $|\delta| \leq 0.4$ and $|\delta| \leq 0.8$.

In the next chapter, we will introduce some of the known techniques that are proposed to overcome the robustness problem in Kalman filtering with imprecise parameters.

Chapter 3

A SHORT REVIEW OF ROBUST KALMAN FILTERING TECHNIQUES

3.1 Robust Kalman Filters

Various robust Kalman filter designs have been proposed to overcome this problem. One of the approaches used to solve the robust Kalman filter problem is that the uncertain parameter is replaced by a scaled version of the known or partially known nominal value, suggesting an upper bound on the estimation error covariance. Some other improvements are achieved by using Riccati equations in [21]. The system under consideration is subjected to time-varying norm-bounded parameter uncertainty in both state and observation equations. A linear filter is designed such that the error covariance is guaranteed to be within a certain bound for all uncertainties. In this section, we will introduce the theoretical background of this robust Kalman filtering technique and give some simulation results of its performance compared to the standard Kalman filter.

The system under consideration is the following uncertain discrete-time system:

$$\begin{aligned} x_{k+1} &= [A + \Delta A_k]x_k + w_k \\ y_k &= [C + \Delta C_k]x_k + v_k, \end{aligned} \tag{3.1}$$

where $x_k \in R^n$ is the state, $w_k \in R^n$ is the process noise, $y_k \in R^m$ is the measurement,$v_k \in R^m$ is the measurement noise, A and C are known nominal matrices, ΔA_k and ΔC_k are unknown matrices which represent parametric uncertainties. In this technique, it is assumed that the matrices ΔA_k and ΔC_k have the form of

$$\begin{bmatrix} \Delta A_k \\ \Delta C_k \end{bmatrix} = \begin{bmatrix} H_1 \\ H_2 \end{bmatrix} F_k E, \tag{3.2}$$

where $F_k \in R^{ixj}$ is an unknown matrix satisfying $F_k^T F_k \leq I$, $k = 0, 1, 2, ...$ and H_1, H_2 and E are known constant matrices. The dimensions of these matrices specify how the elements in matrices A and C are affected by the uncertainty in F_k.

In [21], the main objective is to design a state estimator of the form

$$\hat{x}_{k+1} = G\hat{x}_k + Ky_k, \tag{3.3}$$

where G and K are the matrices to be determined. Using this estimator, the estimation error dynamics can be kept asymptotically stable and a symmetric nonnegative definite matrix Q can be found such that as $k \to \infty$

$$E\left[(x_k - \hat{x}_k)^T(x_k - \hat{x}_k)\right] \leq trace(Q), \tag{3.4}$$

for all certainties. The upper bound $trace(Q)$ represents the guaranteed cost.

Let us define the estimation error $e_k = x_k - \hat{x}_k$. Then, from the state equation in (3.1) and the estimator (3.3), we have

$$e_{k+1} = Ge_k + (A - G - KC)x_k + [\Delta A_k - K\Delta C_k]x_k + w_k - Kv_k \tag{3.5}$$

and the composite system of (3.1) and (3.5) is given by

$$\eta_{k+1} = \left[\bar{A} + \bar{H} F_k \bar{E}\right] \eta_k + \bar{B} \xi_k \tag{3.6}$$

$$e_k = \begin{bmatrix} 0 & I \end{bmatrix} \eta_k \tag{3.7}$$

where $\eta = [x^T \ e^T]^T$, ξ_k is a zero mean white noise satisfying $E(\xi_k \ \xi_l^T) = \delta(k-l)I$ and

$$\bar{A} = \begin{bmatrix} A & 0 \\ A - G - KC & G \end{bmatrix}, \quad \bar{H} = \begin{bmatrix} H_1 \\ H_1 - KH_2 \end{bmatrix}, \tag{3.8}$$

$$\bar{E} = \begin{bmatrix} E & 0 \end{bmatrix}, \quad \overline{BB^T} = \begin{bmatrix} W & W \\ W & W + KVK^T \end{bmatrix}, \tag{3.9}$$

where w_k and v_k are assumed to satisfy the following conditions for all integers $k, l \geq 0$:

$$\begin{aligned} E(w_k) &= 0, \quad E(w_k w_l^T) = W\delta(k-l) \ \text{ for W}\geq 0; \\ E(v_k) &= 0, \quad E(v_k v_l^T) = V\delta(k-l) \ \text{ for V}\geq 0; \\ E(w_k v_l^T) &= 0 \end{aligned} \tag{3.10}$$

A solution to the above formulation in order to obtain the estimator in (3.3) is obtained by using a Riccati equation approach detailed in [21]. As a result, the estimator (3.3) can be given as

$$\hat{x}_{k+1} = \hat{A}\hat{x}_k + K(y_k - \hat{C}\hat{x}_k) \tag{3.11}$$

where

$$K = (\hat{A}Q\hat{C}^T + M)(\hat{R} + \hat{C}Q\hat{C}^T)^{-1}, \tag{3.12}$$

$$\hat{A} = A + \bar{W}(P^{-1} - \bar{W})^{-1}A, \quad \bar{W} = W + \frac{1}{\epsilon} H_1 H_1^T, \tag{3.13}$$

$$\hat{C} = C + \frac{1}{\epsilon} H_2 H_1^T (P^{-1} - \bar{W})^{-1} A, \tag{3.14}$$

$$\hat{R} = V + \frac{1}{\epsilon} H_2 H_2^T + \frac{1}{\epsilon^2} H_2 H_1^T (P^{-1} - \bar{W})^{-1} H_1 H_2^T, \tag{3.15}$$

$$M = \frac{1}{\epsilon}(I - \bar{W}P)^{-1} H_1 H_2^T \tag{3.16}$$

and ϵ is a scalar quantity. The matrices P and Q are the stabilizing solutions of two Riccati equations. Then, this estimator is stable quadratic with a guaranteed steady state cost given in (3.4).

3.1.1 Simulation Results

In [21], a close variant of (2.28) is used to compare the performances of the proposed estimator and the standard Kalman filter:

$$x_{k+1} = \begin{bmatrix} 0 & -0.5 \\ 1 & 1+\delta \end{bmatrix} x_k + \begin{bmatrix} -6 \\ 1 \end{bmatrix} w_k \tag{3.17}$$

$$y_k = \begin{bmatrix} -100 & 10 \end{bmatrix} x_k + v_k, \tag{3.18}$$

where $|\delta| \leq 0.3$. Applying the technique described above, we obtain the following:

$$\begin{aligned} H_1 &= \begin{bmatrix} 0 \\ 10 \end{bmatrix}, \; H_2 = 0, \; E = \begin{bmatrix} 0 & 0.03 \end{bmatrix}, \\ W &= \begin{bmatrix} 36 & -6 \\ -6 & 1 \end{bmatrix}, \; V = 1, \; \epsilon = 1.15, \\ \hat{A} &= \begin{bmatrix} 0.1349 & -0.4813 \\ 1.4169 & 1.2367 \end{bmatrix}, \; K = \begin{bmatrix} -0.0088 & 0.0079 \end{bmatrix}. \end{aligned} \tag{3.19}$$

Assuming that we are trying to estimate $z_k = [\ 1 \quad 0\ ]\hat{x}_k$, the actual costs for 3 different cases of δ are as in Table (3.1), which shows the simulation results in [21].

It can be seen that as the uncertainty increases in the system, standard Kalman filter performance degrades significantly whereas the designed filter is robust.

Filter Type	$\delta = 0$	$\delta = 0.3$	$\delta = -0.3$
Nominal Kalman Filter	36.0	8352.8	551.2
Proposed Filter	61.4	64.4	64.0

Table 3.1: Comparison Between The Nominal Kalman Filter and the Proposed Filter

3.2 James-Stein State Filter

In this section, we will describe the technique introduced in [28] and discuss its advantages and limitations. Consider a p-dimensional random vector X having a multivariate normal distribution with mean $\mu \in R^p$ and variance σ^2. Given the single realization X, the maximum likelihood estimate (MLE) of the mean is $\hat{\mu} = X$ and the risk (mean-square error) of this MLE is $E[\|\hat{\mu} - \mu\|^2] = p$. In [28], James and Stein proved that if the dimension p of X is greater than two, then the *James-Stein* estimator for μ

$$\hat{\mu}^{JS} = \left(1 - \frac{p-2}{\|X\|^2}\right) X \tag{3.20}$$

has a smaller mean-square error than the MLE for all values of μ. Various researches have been done in [30, 31] that study the applications and extensions of James-Stein estimator. This is a special case of *shrinkage* estimator, [32], which means that the second term in the Eq. (3.20) shrinks the MLE X to a mean. Also, some modifications on James-Stein estimator have been made so that the resulting estimator dominates the original one. Consider the observaton equation

$$z = Cx + Dw, \tag{3.21}$$

where $w \sim N(0, \sigma^2 I)$, C and D are known nxp and nxn real matrices, respectively. It is known that

$$\hat{x}^{ML} = (C^T(DD^T)^{-1}C)^{-1}C^T(DD^T)^{-1}z. \tag{3.22}$$

Using the above expression for the maximum likelihood estimator, the final form of James-Stein estimator can be expressed as:

$$\hat{x}^{JS} = \left(1 - \frac{\sigma^2 \left(\min\{(p-2), 2(\tilde{p}-2)\}\right)^+}{(\hat{x}^{ML})^T (C^T (DD^T)^{-1} C) \hat{x}^{ML}}\right)^+ \hat{x}^{ML}, \tag{3.23}$$

for the random vector $X \sim N(\mu, Q)$, where

$$\tilde{p} \triangleq \frac{tr(Q)}{\lambda_{max}(Q)} \quad \text{and} \quad x^+ \triangleq \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{3.24}$$

and $Q = (C^T(DD^T)^{-1}C)^{-1}$. The idea in the evolution of the James-Stein state filter is that when the observation model in Eq. (2.2) holds while there is uncertainty in the state dynamics in Eq. (2.1), or the process noise w_k in Eq. (2.1) is non-Gaussian, even the maximum likelihood estimator can give a better estimation than the Kalman filter. Since James-Stein estimator outperforms MLE in any case, then we can use James-Stein state filter (JSSF) instead of Kalman filter.

JSE in Eq. (3.23) can be modified to allow for shrinkage toward any prior estimate $\hat{x}_k$. Specifically, given the estimate $\hat{x}_k^{JS}$, one can obtaing an estimate $\hat{x}_{k+1}^{JS}$ by shrinking $\hat{x}_{k+1}^{ML}$ toward $A_k \hat{x}_k^{JS}$. The mean-square error obtained by this method will be less than MLE. We can formulate this as:

$$\begin{aligned} \hat{x}_{k|k}^{JS} &= \hat{x}_{k|k-1}^{JS} + G_{JS}(\hat{x}_k^{ML} - \hat{x}_{k|k-1}^{JS}) \\ \hat{x}_{k+1|k}^{JS} &= A_k \hat{x}_{k|k}^{JS}, \end{aligned} \tag{3.25}$$

where the James-Stein state filter gain G_{JS} is

$$G_{JS} = \left(1 - \frac{\sigma^2 \left(\min\{(p-2), 2(\tilde{p}-2)\}\right)^+}{(\hat{x}_k^{ML} - \hat{x}_{k|k-1}^{JS})^T (C_k^T (D_k D_k^T)^{-1} C_k)(\hat{x}_k^{ML} - \hat{x}_{k|k-1}^{JS})}\right)^+. \tag{3.26}$$

However, there are limitations of JSSF. The most drastic limitation is that the state dimension of the system must be no greater than the observation dimension, in other words $n \geq p$ must be satisfied. So, if the number of states exceeds the number of the observations, then the JSSF must be applied to a reduced state-space system, therefore satisfying this condition. In addition, JSSF improves the overall mean-square error, but does not improve individual risks of each state and this estimator is a biased one. These limitations may be an advantage or a disadvantage depending on different systems.

3.2.1 Simulation Results

This section presents simulation results of the performances of James-Stein state filter (3.25), maximum likelihood estimator (MLE) and the Kalman filter where the expression for MLE of x_k is

$$\hat{x}_k^{ML} = \left(C_k^T \left(D_k D_k^T\right)^{-1} C_k\right)^{-1} C_k^T \left(D_k D_k^T\right)^{-1} z_k. \tag{3.27}$$

500 data points are generated by using the state-space model (2.1) and (2.2) with parameters given in the Eq (2.27). For the estimation of the states, three different models are used: the exact model, a perturbed model where the elements of A_k are exposed to a Gaussian noise with N(0,0.1) and a completely incorrect model where

$$A_k = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}. \tag{3.28}$$

As performance criterion, the average of mean-square errors (MSE's) of MLE, KF and JSSF techniques over 500 independent runs in dB is used. Figures (3.1) - (3.3) show the MSE's of each technique for the first 200 data points, and Table (3.2) presents the average MSE's of each estimate.

	MLE(dB)	KF(dB)	JSSF(dB)
Exact	4.77	2.54	3.98
Perturbed	4.77	5.47	4.37
Incorrect	4.77	7.09	4.75

Table 3.2: MSE of each technique in dB under different models

From the simulations, it can be seen that even for small perturbations in model parameters, Kalman filtering results in a larger MSE than MLE and JSSF. Also, we can see that JSSF never performs worse than MLE. Thus, JSSF suggests a robust Kalman filtering method when the model parameters are not exactly known.

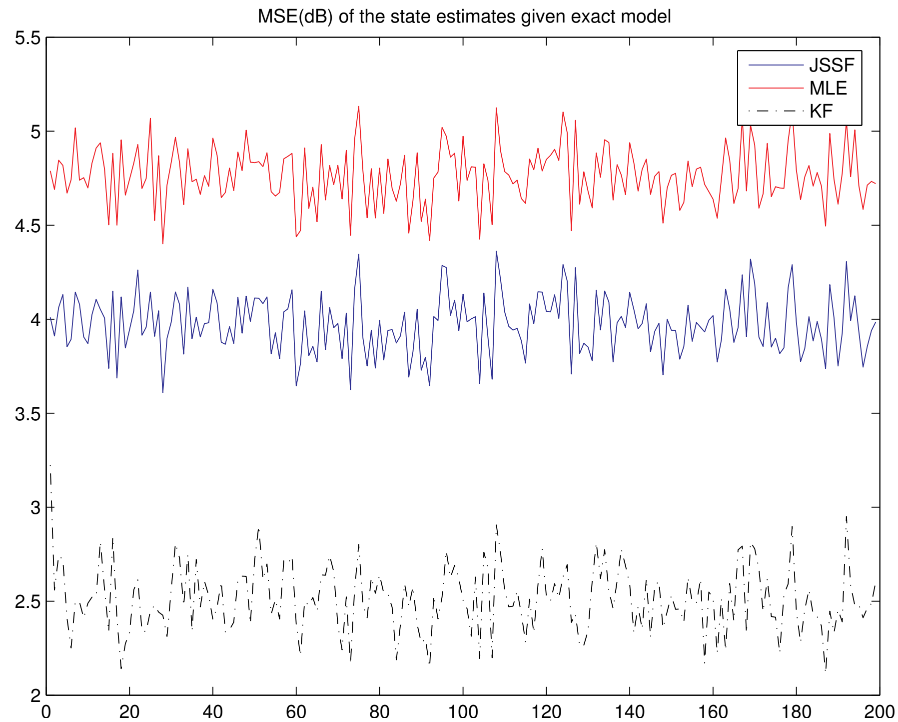

Figure 3.1: JSSF, KF and MLE under exact model.

In the next chapter, we will detail our proposed min-max sense optimal approach for robust Kalman filtering. We will also provide comparison results between the James-Stein state filter presented in this chapter and our approach as well as the standard Kalman filter.

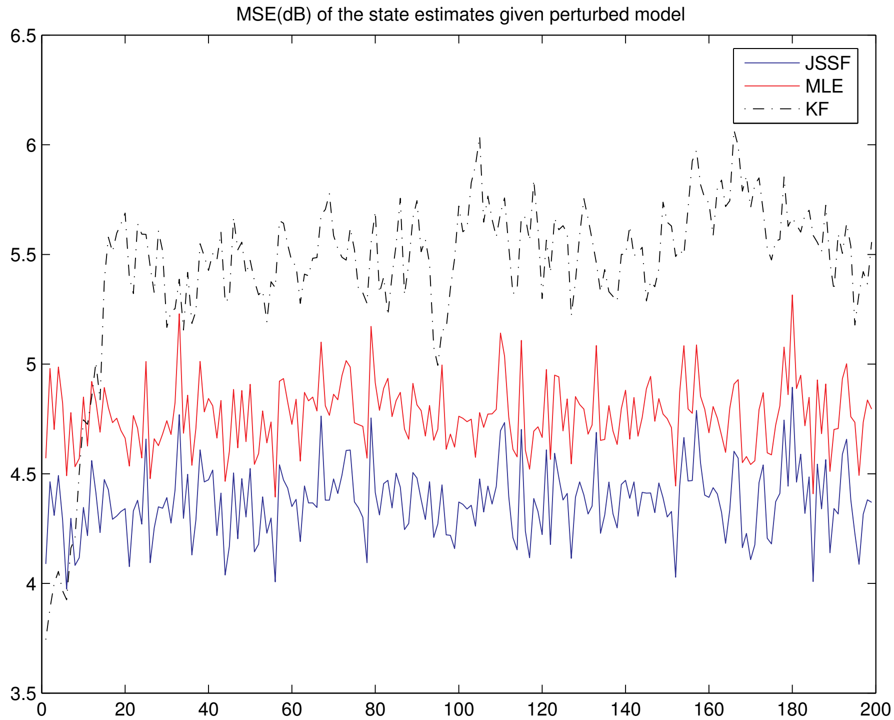

Figure 3.2: JSSF, KF and MLE under perturbed model.

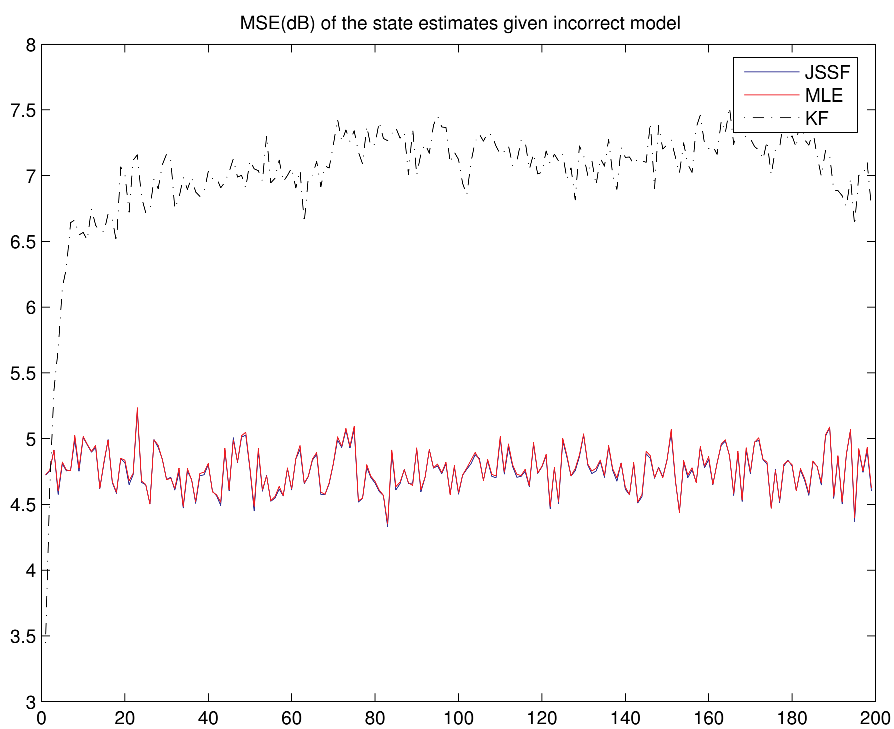

Figure 3.3: JSSF, KF and MLE under incorrect model.

Chapter 4

Robust Minimax Estimation

4.1 Minimax Estimation Background

The technique we will introduce is proposed as a solution to the problem of estimating an unknown deterministic parameter x observed through a linear transformation H and corrupted by noise w, which can be expressed as

$$y = Hx + w. \tag{4.1}$$

This problem comes up very frequently in many areas such as signal processing, control and economics. The method in [29] will be followed and detailed in this section.

We will first consider the case in which H is exactly known, and develop a minimax robust estimator, minimizing the maximum MSE across all the possible values of the states in a known bound, e.g., $x^T T x \leq L^2$ for some weighting matrix T and some constant L. The solution will be obtained after a semi-definite programming, which is a convex optimization problem.

Secondly, we will formulate the case in which the matrix H is not exactly known, but there is an uncertainty in it so that the observation matrix will be given by $H + \delta H$. In this formulation, we will assume that the nominal value, H, of this matrix is known, and δH has a known bound, $\|\delta H\| \leq \rho$. Under this model, we will seek a robust linear estimator that minimizes the maximum MSE across all the possible values of the states in a known bound. Again, the solution will be obtained using semi-definite programming algorithms.

4.1.1 Minimax MSE Estimation With Known H

Consider the problem of estimating the unknown x in the model given in Eq (4.1), where H is a known nxm matrix with full rank m, and w is a zero-mean random vector with covariance C_w. We assume that x satisfies the weighted norm constraint $\|x\|_T \leq L$ for some positive matrix T and scalar $L > 0$, where we define $\|x\|_T^2 = x^T T x$.

A linear estimator $\hat{x} = Gy$ is computed for estimating x for some mxn matrix G. Then, the MSE of the estimator $\hat{x} = Gy$ is given by

$$\begin{aligned} E(\|\hat{x} - x\|^2) &= V(\hat{x}) + \|B(\hat{x})\|^2 \\ &= tr(GC_wG^T) + x^T(I - GH)^T(I - GH)x. \end{aligned} \tag{4.2}$$

The secong term in Eq (4.2) is dependent on x, so we can not directly minimize the MSE. Instead, we compute the linear estimator which minimizes the maximum MSE across all possible values of x satisfying the weighted norm constraint. Thus, we consider

$$\min_{\hat{x}=Gy} \max_{\|x\|_T \leq L} E(\|\hat{x}-x\|^2) = \min_{G} \max_{\|x\|_T \leq L} \left\{ tr(GC_wG^T) + x^T(I - GH)^T(I - GH)x \right\}. \tag{4.3}$$

We first compute the maximum of the term in paranthesis in Eq. (4.3) with respect to x, so only x-dependent term is to be maximized. Then, the inner problem is

$$\max_{\|x\|_T \leq L} x^T(I - GH)^T(I - GH)x. \tag{4.4}$$

Now, we define a new variable $z = T^{1/2}x$, then we obtain

$$\begin{aligned}
&\max_{x^T T x \leq L^2} x^T (I - GH)^T (I - GH) x \\
= &\max_{z^T z \leq L^2} \left\{ z^T T^{-1/2} (I - GH)^T (I - GH) T^{-1/2} z \right\} \\
= &L^2 \lambda_{max}, \qquad (4.5)
\end{aligned}$$

where λ_{max} is the largest eigenvalue of $T^{-1/2}(I - GH)^T(I - GH)T^{-1/2}$. Let the notation $A \preceq B$ denote $B - A$ is positive definite, then the Eq.(4.5) can be expressed as

$$\min_{\lambda} \lambda, \qquad (4.6)$$

subject to

$$T^{-1/2}(I - GH)^T(I - GH)T^{-1/2} \preceq \lambda I. \qquad (4.7)$$

From Eqs (4.5)-(4.7), the problem in Eq.(4.3) can be reformulated as

$$\min_{G,\lambda} \left\{ tr(GC_w G^T) + L^2 \lambda \right\}, \qquad (4.8)$$

subject to (4.7), which then can be rewritten as

$$\min_{\tau,G,\lambda} \tau, \qquad (4.9)$$

subject to

$$\begin{aligned}
tr(GC_w G^T) + L^2 \lambda &\leq \tau \\
T^{-1/2}(I - GH)^T(I - GH)T^{-1/2} &\preceq \lambda I. \qquad (4.10)
\end{aligned}$$

Since our problem is formulated as in Eq. (4.9) and (4.10), now we show that this problem can be solved using a standard semi-definite programming (SDP) algorithm, which can be used to minimize a linear objective subject to linear martix inequality (LMI) constraints. An LMI is a matrix constaint of the form $A(x) \succeq 0$, where A matrix linearly depends on x. We can use efficient computational methods to solve this SDP problem.

Let

$$g = vec(GC_w^{1/2}), \tag{4.11}$$

where $m = vec(M)$ denotes the vector obtained by stacking the columns of M. Using this notation, the constraints in (4.10) become

$$\begin{aligned} g^T g + L^2\lambda &\leq \tau \\ T^{-1/2}(I - GH)^T(I - GH)T^{-1/2} &\preceq \lambda I. \end{aligned} \tag{4.12}$$

The problem in constraints (4.12) is that the elements $G(i,j)$ of G don't appear linearly in $g^T g$ and $T^{-1/2}H^T G^T GHT^{-1/2}$. To rewrite the inequalities as LMI's in the variables $G(i,j)$,λ and τ, we use the following lemma [33, p. 472]:

Lemma 1 (Schur's Complement): Let

$$M = \begin{bmatrix} A & B^T \\ B & C \end{bmatrix} \tag{4.13}$$

be a Hermitian matrix with $C \succ 0$. Then, $M \succeq 0$ if and only if $\triangle_C \succeq 0$, where $\triangle_C$ is the Schur complement of C in M and is given by $\triangle_C = A - B^T C^{-1} B$.

Using Lemma 1, we can rewrite the constraints in (4.10) as

$$\begin{aligned} \begin{bmatrix} \tau - L^2\lambda & g^T \\ g & I \end{bmatrix} &\succeq 0, \\ \begin{bmatrix} \lambda I & T^{-1/2}(I - GH)^T \\ (I - GH)T^{-1/2} & I \end{bmatrix} &\succeq 0. \end{aligned} \tag{4.14}$$

Now, the constraints are LMIs in the variables G, λ and τ. We conclude that the problem of (4.3) is equivalent to the SDP defined by (4.9) and (4.14).

4.1.2 Minimax MSE Estimation With Unknown H

In the previous section, we derived a formulation for the optimal estimator that minimizes the maximum MSE across all values of x within the bound. Through the derivation, we assumed that the matrix H is exactly known, which is not the case in many engineering applications. Most of the time, H is subject to uncertainties, it might be estimated from noisy data or selected among different types of model matrices.

In our calculations, we will assume the true model matrix as $H + \delta H$ where δH is a perturbation matrix with known bound ρ and H is a known nominal value for the model matrix. In this section, we explicitly derive a formulation for the uncertain H case. Then our problem is

$$\min_{\hat{x}=Gy} \max_{\|x\|_T \leq L,\, \|\delta H\| \leq \rho} E(\|\hat{x} - x\|^2) = \min_G \{ \max_{\|x\|_T \leq L,\, \|\delta H\| \leq \rho} \left\{ x^T (I - G(H + \delta H))^T (I - G(H + \delta H))x \right\} + tr(GC_wG^T) \right\}. \quad (4.15)$$

We start with considering the inner maximization problem

$$\max_{\|x\|_T \leq L,\, \|\delta H\| \leq \rho} \{x^T (I - G(H + \delta H))^T (I - G(H + \delta H))x\} + tr(GC_wG^T). \quad (4.16)$$

Simply by changing variables, Eq (4.16) becomes

$$\begin{aligned} & \max_{\|\delta H\| \leq \rho} \max_{\|x\| \leq L} x^T T^{-1/2} (I - G(H + \delta H))^T (I - G(H + \delta H)) T^{-1/2} x \\ = & \max_{\|\delta H\| \leq \rho} L^2 \lambda_{max}(\delta H), \end{aligned} \quad (4.17)$$

where $\lambda_{max}(\delta H)$ is the largest eigenvalue of the matrix $T^{-1/2}(I - G(H + \delta H))^T (I - G(H + \delta H))T^{-1/2}$. So the problem in Eq. (4.17) can be expressed as

$$\min_\tau L^2 \tau \quad (4.18)$$

subject to

$$T^{-1/2}(I - G(H + \delta H))^T (I - G(H + \delta H))T^{-1/2} \preceq \tau I \quad \forall\, \delta H : \|\delta H\| \leq \rho. \quad (4.19)$$

Using Lemma 1, the constraint (4.19) can be rearranged as

$$\begin{bmatrix} \tau I & T^{-1/2}(I - G(H+\delta H))^T \\ (I - G(H+\delta H)) & I \end{bmatrix} \succeq 0 \quad \forall\, \delta H : \|\delta H\| \leq \rho \tag{4.20}$$

which is equivalent to

$$A(\tau, G) \succeq B^T(G)\delta H C + C^T(\delta H)^T B(G), \quad \forall\, \delta H : \|\delta H\| \leq \rho, \tag{4.21}$$

where

$$\begin{aligned} A(\tau, G) &= \begin{bmatrix} \tau I & T^{-1/2}(I - GH)^T \\ (I - GH)T^{-1/2} & I \end{bmatrix} \\ B(G) &= \begin{bmatrix} 0 & G^T \end{bmatrix} \\ C &= \begin{bmatrix} T^{-1/2} & 0 \end{bmatrix}. \end{aligned} \tag{4.22}$$

For further arrangements, we consider the following lemma [29]:

Lemma 2 : Given matrices P, Q, A with $A = A^T$.

$$A \succeq P^T X Q + Q^T X^T P, \quad \forall\, X : \|X\| \leq \rho \tag{4.23}$$

if and only if there exists a $\lambda \geq 0$ such that

$$\begin{bmatrix} A - \lambda Q^T Q & -\rho P^T \\ -\rho P & \lambda I \end{bmatrix} \succeq 0. \tag{4.24}$$

From Lemma 2, Eq (4.21) holds if and only if there exists a $\lambda \geq 0$ such that

$$\begin{bmatrix} \tau I - T^{-1} & T^{-1/2}(I - GH)^T & 0 \\ (I - GH)T^{-1/2} & I & -\rho G \\ 0 & -\rho G^T & \lambda I \end{bmatrix} \succeq 0. \tag{4.25}$$

Thus, the problem (4.15) can be expressed as

$$\min_{t, G, \lambda, \tau} t, \tag{4.26}$$

subject to the LMI (4.25) and

$$tr(G C_w G^T) + L^2 \tau \leq t. \tag{4.27}$$

Using Lemma 1, (4.27) can be rewritten as the LMI

$$\begin{bmatrix} t - L^2\tau & g^T \\ g & I \end{bmatrix} \succeq 0, \tag{4.28}$$

where $g = vec(GC_w^{1/2})$. Now we can formulate the problem of (4.26) subject to (4.25) and (4.27) as an SDP.

To summarize our results, we conclude that the problem

$$\min_{\hat{x}=Gy} \max_{\|x\|_T \le L,\ \|\delta H\| \le \rho} E(\|\hat{x} - x\|^2) \tag{4.29}$$

is equivalent to the SDP problem of

$$\min_{t,G,\lambda,\tau} t \tag{4.30}$$

subject to the constraints

$$\begin{bmatrix} t - L^2\tau & g^T \\ g & I \end{bmatrix} \succeq 0$$

$$\begin{bmatrix} \tau I - T^{-1} & T^{-1/2}(I - GH)^T & 0 \\ (I - GH)T^{-1/2} & I & -\rho G \\ 0 & -\rho G^T & \lambda I \end{bmatrix} \succeq 0,$$

where $g = vec(GC_w^{1/2})$.

4.2 Kalman Filter and Modified James-Stein State Filter Implementation

In this section, we develop a formulation for robust Kalman filter using the robust minimax estimation technique detailed in the previous section and the James-Stein state filter. Consider the state-space model

$$x_{k+1} = Ax_k + Bw_k \tag{4.31}$$

$$y_k = Cx_k + Dv_k, \tag{4.32}$$

where $w_k \sim N(0,Q)$ and $v_k \sim N(0,R)$. In Eq. (4.31), the next state is obtained from the current state. Let us rewrite this equation so that the previous state can be obtained from the current state. For example, in a target tracking problem, we can rewrite the Eq. (4.31) and use a new backward transition matrix A_B. This matrix gives the backward relation between current and previous state. So, we can write

$$x_{k-1} = A_B x_k + Bw_k. \tag{4.33}$$

In this equation, the previous state x_{k-1} is dependent on the current state x_k and obtained by using a backward state transition matrix A_B.

Now, let us define

$$\delta x_k \triangleq \hat{x}_{k|k} - A\hat{x}_{k-1|k-1} = \hat{x}_{k|k} - \hat{x}_{k|k-1}, \tag{4.34}$$

where $\hat{x}_{k|k-1}$ denotes the *a priori* estimate and $\hat{x}_{k|k}$ denotes the *a posteriori* estimate. Then, (4.31) and (4.32) can be arranged as

$$\begin{bmatrix} 0 \\ y_k - C\hat{x}_{k|k-1} \end{bmatrix} = \begin{bmatrix} A_B \\ C \end{bmatrix} \delta x_k + \begin{bmatrix} Bw_k \\ Dv_k \end{bmatrix}, \tag{4.35}$$

where w_k and v_k have the same distributions as before. We will apply the technique stated in the previous section in order to estimate x in the equation $z = Hx + n$ for unknown matrix H with the matrices

$$z = \begin{bmatrix} 0 \\ y_k - C\hat{x}_{k|k-1} \end{bmatrix}, \quad H = \begin{bmatrix} A_B \\ C \end{bmatrix} \quad \text{and} \quad n = \begin{bmatrix} Bw_k \\ Dv_k \end{bmatrix}. \tag{4.36}$$

Starting from an initial value of x_0, we compute a linear estimator G by semi-definite programming with proper parameters T, ρ and C_n. We use *Yalmip*, a Matlab toolbox, in order to obtain G for the problem in (4.30). Once G is computed, then using the a priori estimate $\hat{x}_{k|k-1}$ and the present data z_k, we compute the a posteriori estimate $\hat{x}_{k|k}$ at time k as:

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + Gz$$

$$= \quad A\hat{x}_{k-1|k-1} + Gz. \tag{4.37}$$

Modification on James-Stein State Filter: Using the new representation stated in (4.35), we can obtain a new James-Stein estimator which also takes the perturbations in A into account. The James-Stein estimator for x in the regression $z = Hx + Un$ is

$$\begin{aligned} \hat{x}^{JS}_{k|k} &= \hat{x}^{JS}_{k|k-1} + G_{JS}(\hat{x}^{ML} - \hat{x}^{JS}_{k|k-1}) \\ \hat{x}^{JS}_{k+1|k} &= A_k \hat{x}^{JS}_{k|k}, \end{aligned}$$

where

$$\begin{aligned} G_{JS} &= \left(1 - \sigma^2 \frac{(\min\{(p-2), 2(\tilde{p}-2)\})^+}{(\hat{x}^{ML} - \hat{x}^{JS}_{k|k-1})^T (H^T(UU^T)^{-1}H)(\hat{x}^{ML} - \hat{x}^{JS}_{k|k-1})}\right)^+ \\ \hat{x}^{ML}_k &= (H_k^T(U_kU_k^T)^{-1}H_k)^{-1}H_k^T(U_kU_k^T)^{-1}z_k \\ \tilde{p} &= \frac{tr(\tilde{Q})}{\lambda_{max}(\tilde{Q})} \\ \tilde{Q} &= (H^T(UU^T)^{-1}H)^{-1}, \end{aligned}$$

and p is the dimension of x. We can apply the James-Stein estimator above to the formulation in (4.35) by putting

$$z = \begin{bmatrix} 0 \\ y_k - C\hat{x}_{k|k-1} \end{bmatrix}, \quad H = \begin{bmatrix} A_B \\ C \end{bmatrix}. \tag{4.38}$$

In this formulation, $n \sim N(0, \sigma^2 I)$ is assumed. Obtaining the matrix U and the vector n might require a whitening process for this estimation to be used if the process and observation noises are dependent.

As mentioned in Chapter 3, the most drastic limitation of the James-Stein state filter is that the number of observations should be no less than the number of states. Thus, for $C \in R^{n\text{x}p}$, the system should satisfy $n \geq p$ for the standart James-Stein state filter. After the above modification of taking the matrix A_B into account, the new observation matrix becomes $H \in R^{(n+p)\,\text{x}\,p}$. So, the limitation has automatically been overcome, which is an important modification on James-Stein state filter.

4.3 Simulation Results

Consider the system

$$A = \begin{bmatrix} 0 & -0.5 \\ 1 & 1+\delta \end{bmatrix}, \quad B = \begin{bmatrix} -6 & 0 \\ 0 & 1 \end{bmatrix},$$
$$C = \begin{bmatrix} 1 & -0.25 \\ -1 & 0.5 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \tag{4.39}$$

where $\delta = 0.3$ is assumed. In our simulation, in order to see the effect of the perturbations in matrix C, we added some zero mean Gaussian noise with different variances. The figure (4.1) shows the performances of James-Stein State filter detailed in the previous chapter, the modified James-Stein filter described in the previous section, the nominal Kalman filter and the minimax filter in dB scale drawn with respect to different variances of the noise added to the matrix C. Table (4.1) shows the MSE of each technique for different variances in dB.

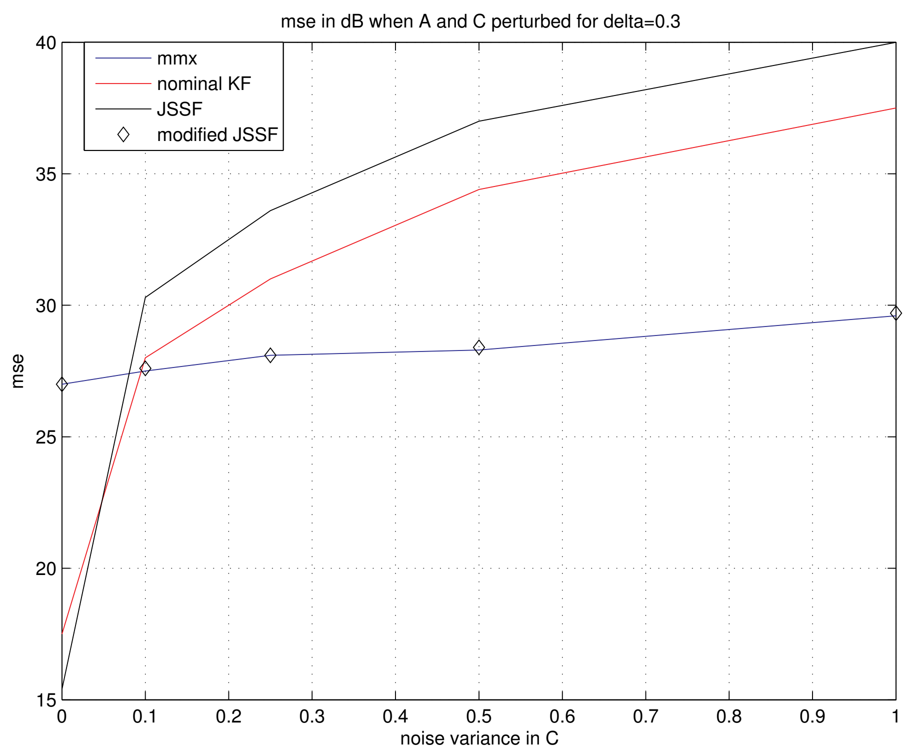

Figure 4.1: JSSF, modified JS, KF and minimax estimation performances.

	$\sigma_C^2 = 0$	$\sigma_C^2 = 0.1$	$\sigma_C^2 = 0.25$	$\sigma_C^2 = 0.5$	$\sigma_C^2 = 1$
Mmx	27	27.5	28.1	28.3	29.6
Nominal Kf	17.5	28	31	34.4	37.5
Jssf	15.4	30.3	33.6	37	40
Modified Jssf	27	27.6	28.1	28.4	29.7

Table 4.1: Performances for different noise variances in dB.

Since James-Stein state filter estimates the states using the observation equation only, it performs better compared to minimax filter and the modified James-Stein filter when there is no perturbation in matrix C. As the noise in C increases, the minimax filter and the modified James-Stein filter techniques perform better as these techniques takes the state equation information into account.

4.3.1 Radar Tracking Application

In this example, we consider a land-based vehicle moving with a constant speed and assume that we measure the range relative to two reference points $(0, 0)$ and $(R_x,\ R_y)$, where $R_x = 170$ km and $R_y = 100$ km and each reference point shows the easterly and northerly positions. A close variant of this scheme is also investigated in [2]. We define the state vector for the system in Fig. (4.2)

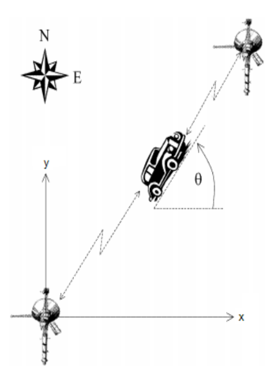

Figure 4.2: A system with 2 radars for tracking a land vehicle.

as $\underline{x}_k = [x_k\ y_k\ V_{x_k}\ V_{y_k}]^T$, where x_k and y_k denote the easterly and northerly positions of the vehicle respectively, and V_{x_k} and V_{y_k} denote the easterly and northerly velocities. Then the system is:

$$\underline{x}_{k+1} = \begin{bmatrix} 1 & 0 & T_s & 0 \\ 0 & 1 & 0 & T_s \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \underline{x}_k + Bw_k \tag{4.40}$$

$$\underline{y}_k = \begin{bmatrix} \sqrt{x_k^2 + y_k^2} \\ \sqrt{(x_k - R_x)^2 + (y_k - R_y)^2} \end{bmatrix} + De_k, \tag{4.41}$$

where w_k is the process noise caused by the potholes, wind, etc., e_k is the observation noise and T_s is the sampling period. Since the state dynamics are linear whereas the observations are nonlinear, we use an extended version of the Kalman filter.

Extended Kalman Filter : In the cases where state dynamics or observations are nonlinear, we have to use a linearized version of the Kalman filter, namely the extended Kalman filter. The system under investigation is:

$$\begin{aligned} x_k &= f(x_{k-1}) + w_k \\ y_k &= h(x_k) + v_k. \end{aligned}$$

The function f can be used to compute the predicted state from the previous estimate and similarly the function h can be used to compute the predicted measurement from the predicted state. However, f and h can not be applied to the covariance directly. Instead a matrix of partial derivatives, the Jacobian, is computed. Then, the extended Kalman filter equations that we use in the radar tracking example are:

Predict:

$$\begin{aligned} \hat{x}_{k|k-1} &= f(\hat{x}_{k-1|k-1}) \\ P_{k|k-1} &= F_k P_{k-1|k-1} F_k^T + Q_k, \end{aligned} \tag{4.42}$$

Update:

$$
\begin{aligned}
S_k &= H_k P_{k|k-1} H_k^T + R_k \\
K_k &= P_{k|k-1} H_k^T S_k^{-1} \\
\hat{x}_{k|k} &= \hat{x}_{k|k-1} + K_k(z_k - h(\hat{x}_{k|k-1})) \\
P_{k|k} &= (I - K_k H_k) P_{k|k-1},
\end{aligned} \tag{4.43}
$$

where z_k is the observation, Q_k and R_k are the state and observation noise covariances, respectively. The Jacobians are:

$$
\begin{aligned}
F_k &= \left.\frac{\partial f}{\partial x}\right|_{\hat{x}_{k-1|k-1}} \\
H_k &= \left.\frac{\partial h}{\partial x}\right|_{\hat{x}_{k|k-1}}
\end{aligned} \tag{4.44}
$$

In the radar tracking example, since the state dynamics are linear, we will linearize the observation equation only. Thus,

$$
H_k = \left.\left[\begin{array}{cc} \frac{x_k}{\sqrt{x_k^2+y_k^2}} & \frac{y_k}{\sqrt{x_k^2+y_k^2}} \\ \frac{x_k-R_x}{\sqrt{(x_k-R_x)^2+(y_k-R_y)^2}} & \frac{y_k-R_y}{\sqrt{(x_k-R_x)^2+(y_k-R_y)^2}} \end{array} \right]\right|_{\hat{x}_{k|k-1}} . \tag{4.45}
$$

In the simulations, we compared the extended Kalman filter with the minimax filter obtained by using the criteria in (4.9) and (4.14) for the case that there is no parameter uncertainty, and the minimax filter obtained by using the criteria in (4.30) for the case that there is parameter uncertainty. The sampling period in the simulations is 1 second and the covariances of both process and observation noises are

$$
\begin{aligned}
Q &= \mathrm{Diag}(\ 10m^2,\ 10m^2,\ 3(m/s)^2,\ 3(m/s)^2\ ) \\
R &= \mathrm{Diag}(\ 10m^2,\ 10m^2\ ).
\end{aligned} \tag{4.46}
$$

Initially, the vehicle is assumed to be at the point (20 km, 10 km) and moving at a constant speed of 100 km/h with $\theta = 30$ deg. Extended Kalman filter and the minimax filter with no parameter uncertainty and with parameter uncertainty

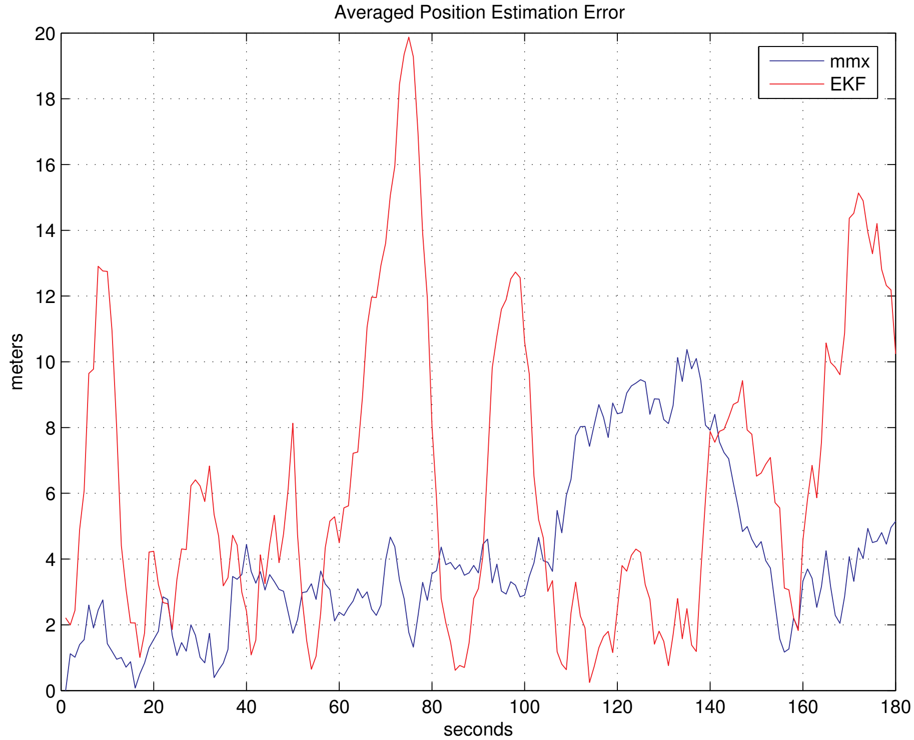

Figure 4.3: Average root mean-squared error in both directions with no uncertainty.

are run in Matlab separately for 180 seconds. The uncertainty in the observations caused by the angle and increasing range is represented with a maximum of (10 m, 10 m) error in the measured range of both radars.

In (4.3) and (4.4), we can see that as the parameter uncertainty is increased, minimax filtering results in much more accurate results than the extended Kalman filtering. The extended Kalman filtering gives an error of 70 m after 3 minutes whereas the minimax filtering error is around 30 m.

Figures (4.5) and (4.6) show the average position error of minimax filtering with no uncertainty and with parameter uncertainty, respectively. Similarly, (4.7) and (4.8) show the average position error of extended Kalman filtering with no uncertainty and with parameter uncertainty, respectively. Extended Kalman filtering estimation performance in both directions is highly degraded compared to the robust minimax estimation performance after we add the parameter uncertainty to the observations.

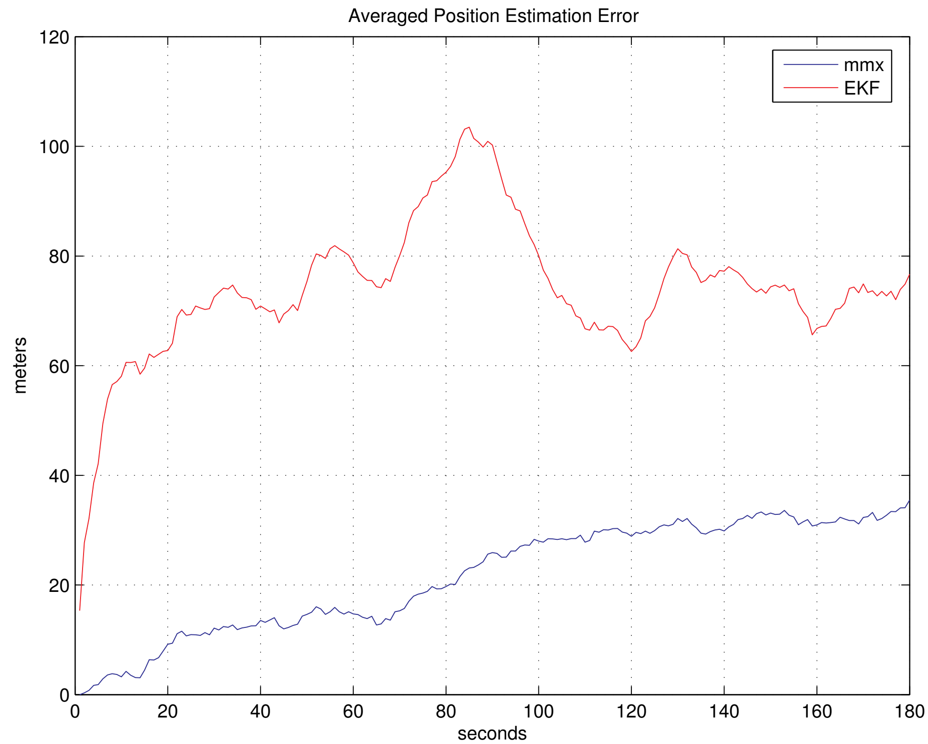

Figure 4.4: Average root mean-squared error in both directions with parameter uncertainty.

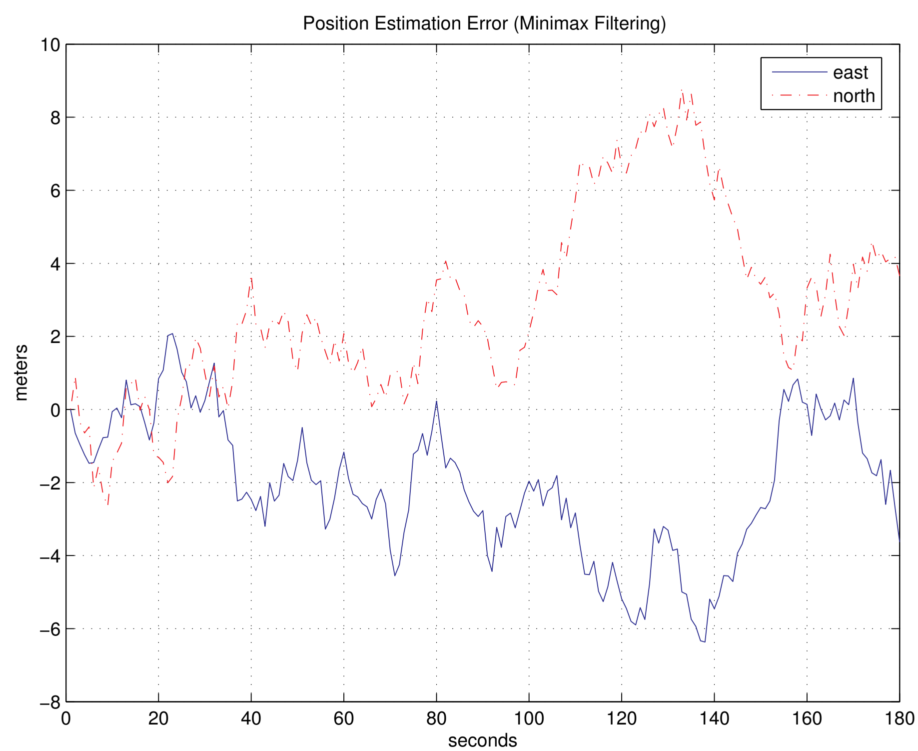

Figure 4.5: Minimax filter position error with no uncertainty.

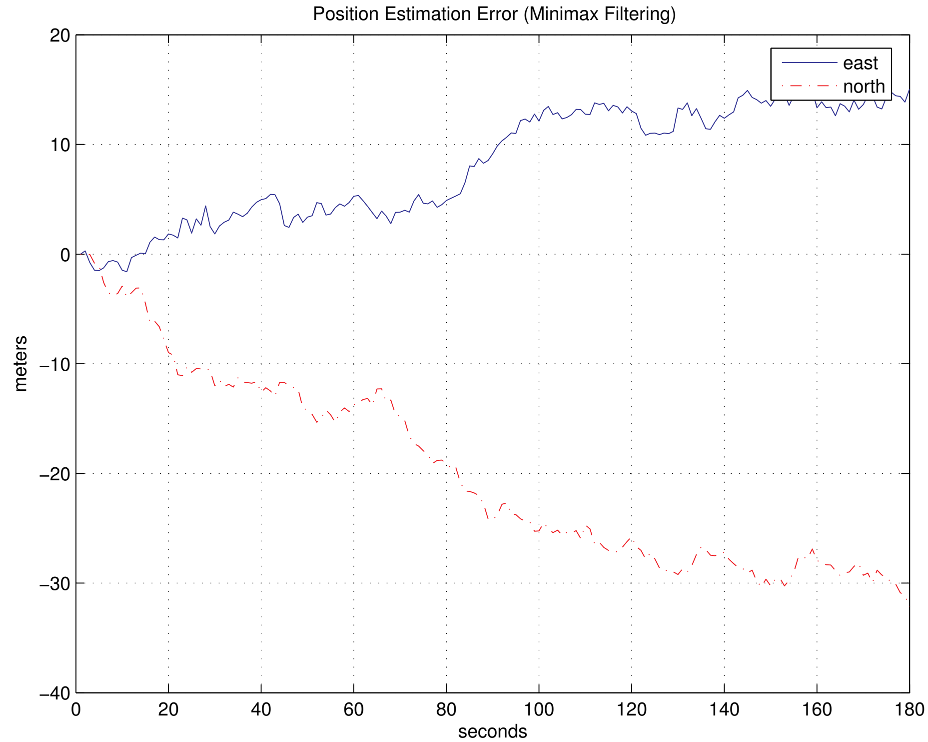

Figure 4.6: Minimax filter position error with parameter uncertainty

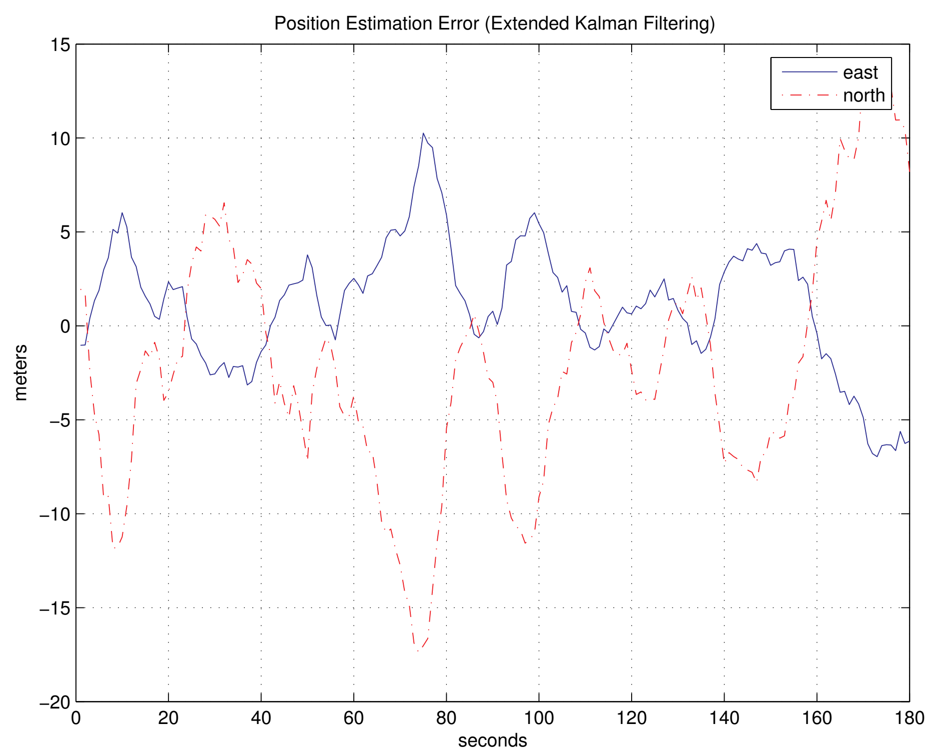

Figure 4.7: Extended Kalman filter position error with no uncertainty.

Figure 4.8: Extended Kalman filter position error with parameter uncertainty.

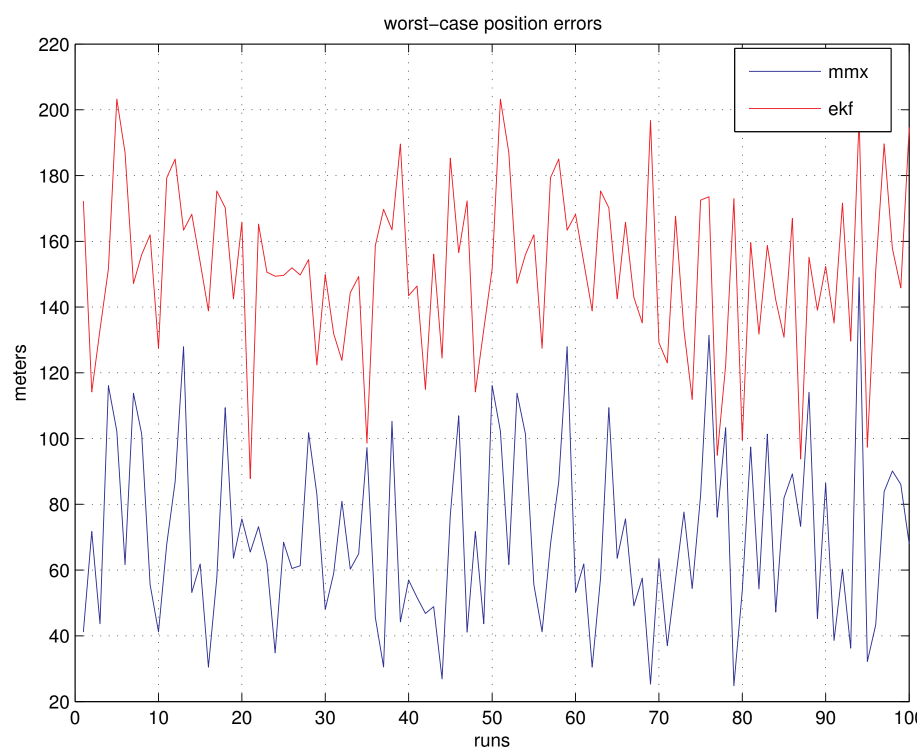

Figure 4.9: Worst-case position errors.

Figure (4.9) shows the worst-case position errors over all runs. Clearly, minimax filtering estimation yields much smaller worst-case estimation errors than the extended Kalman filter. Extended Kalman filter yields a worst-case estimation error that varies between 100 m and 200 m whereas the minimax filtering worst-case estimation error is between 40 m and 120 m.

Chapter 5

Conclusions and Future Work

5.1 Conclusions

In this thesis, we investigated some deficiencies of Kalman filtering in the presence of parameter uncertainties in the state-space model and proposed a new solution to this problem. We showed in the simulation results that Kalman filtering yields high mean-square error as the parameter uncertainty increases. Thus, in real applications, well-known Kalman filtering itself is not a good technique when there is a risk of noise disturbances in the state-space model.

We reviewed some techniques described in [21] and [28] that were developed to overcome the robustness problem of the Kalman filtering. The approach of solving algebraic Riccati equations [21] and James-Stein state filtering approach [28] are both good techniques, since they yield small average mean-square error for the cases in which their limitations and assumptions hold. We focused on minimizing the worst-case mean-square error of the state estimation, so we investigated the minimax approach to the estimation problem.

The main contribution of this study is that we applied the minimax approach to the state-space model described in (1.1) using semi-definite programming as an alternative to the standart Kalman filter. We carried out simulations and tested our technique with arbitrary matrices and noise levels besides a real application involving radar target tracking in two dimensions.

The second contribution of the thesis is that we modified the James-Stein estimator so that it can be applied to the Kalman filtering in the presence of uncertainties in both of the state transition and observation matrices whether the number of observations is smaller than the number of states or not. A robust filtering can be achieved by this way when we do not know the state and observation equations exactly.

5.2 Future Work

In this study, a time-invariant version of the Eq. (1.1) is under consideration. Model parameters are constant for each run, thus computational complexity is kept minimum. When we apply our technique to the radar tracking application, the linearization leads to a time-varying estimation, thus increasing the complexity and the run time of minimax estimation. One possible extension of this thesis is the optimal evaluation of a time-varying case of the state-space model.

Also, for robust estimation of the state-space model where both the average mean-square error and the worst-case mean-square error are of importance, a hybrid model consisting of scaled versions of Kalman filter and minimax estimation outputs can be developed as a future work.

APPENDIX A

Proof of Theorem 1

A.1 Matrix Differentiation

In this section of the thesis, we prove the matrix equations in (2.15) hold for the given matrix constraints.

Definition 1. Let us assume that X is an arbitrary (n x n) square matrix. Then, the trace of the matrix X is

$$trace(X) = \sum_{i=1}^{n} x_{ii}, \tag{A.1}$$

where x_{ii} is the i^{th} diagonal element of the matrix X.

Proposition 1. Let us assume that A and B are arbitrary matrices where A is (n x m) and B is (m x n) so that the matrix multiplication AB is a square matrix. The derivative of the AB with respect to A is equal to B^T.

Proof. Let us write A and B in the form of row and column vectors. Then,

$$trace(AB) \quad = \quad trace \begin{bmatrix} \longleftarrow \vec{a}_1 \longrightarrow \\ \longleftarrow \vec{a}_2 \longrightarrow \\ \vdots \\ \longleftarrow \vec{a}_n \longrightarrow \end{bmatrix} \begin{bmatrix} \vec{b}_1 & \vec{b}_2 & \cdots & \vec{b}_n \end{bmatrix},$$

$$
\begin{aligned}
&= \begin{bmatrix} \vec{a}_1\vec{b}_1 & \vec{a}_1\vec{b}_2 & \cdots & \vec{a}_1\vec{b}_n \\ \vec{a}_2\vec{b}_1 & \vec{a}_2\vec{b}_2 & \cdots & \vec{a}_2\vec{b}_n \\ \vdots & \vdots & \ddots & \vdots \\ \vec{a}_n\vec{b}_1 & \vec{a}_n\vec{b}_2 & \cdots & \vec{a}_n\vec{b}_n \end{bmatrix}, \\
&= \sum_{i=1}^{m} a_{1i}b_{i1} + \sum_{i=1}^{m} a_{2i}b_{i2} + \cdots + \sum_{i=1}^{m} a_{ni}b_{in}, \qquad \text{(A.2)}
\end{aligned}
$$

where a_{ij} belongs to i^{th} row and jth column of the matrix A. Using the differentiation formula in (2.15), we obtain

$$\frac{\partial trace(AB)}{\partial a_{ij}} = b_{ji}, \qquad \text{(A.3)}$$

which implies

$$\frac{\partial trace(AB)}{\partial A} = B^T. \qquad \text{(A.4)}$$

Proposition 2. Assume an arbitrary matrix A. The derivative of a function of the matrix A with respect to A is the same as the transpose of the derivative of the same function with respect to A^T, so that

$$[\nabla_{A^T} f(A)]^T = [\nabla_A f(A)] \qquad \text{(A.5)}$$

Proof. Let us write the differentiation elementwise:

$$
\begin{aligned}
[\nabla_{A^T} f(A)] &= \begin{bmatrix} \frac{\partial f(A)}{\partial A_{11}} & \frac{\partial f(A)}{\partial A_{21}} & \cdots & \frac{\partial f(A)}{\partial A_{n1}} \\ \frac{\partial f(A)}{\partial A_{12}} & \frac{\partial f(A)}{\partial A_{22}} & \cdots & \frac{\partial f(A)}{\partial A_{n2}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f(A)}{\partial A_{1n}} & \frac{\partial f(A)}{\partial A_{2n}} & \cdots & \frac{\partial f(A)}{\partial A_{nn}} \end{bmatrix} \\
&= [\nabla_A f(A)]^T \qquad \text{(A.6)}
\end{aligned}
$$

Proposition 3. Consider a function $f : R^n \to R$. Then,

$$\nabla_x f(Ax) = A^T \nabla f(Ax), \qquad \text{(A.7)}$$

where the matrix $A \in R^{nxm}$ and the vector $x \in R^m$.

Proof. Using the chain rule, we have

$$\begin{aligned}
\frac{\partial f(Ax)}{x_i} &= \sum_{k=1}^{n} \frac{\partial f(Ax)}{\partial (Ax)_k} . \frac{\partial (Ax)_k}{\partial x_i} = \sum_{k=1}^{n} \frac{\partial f(Ax)}{\partial (Ax)_k} . \frac{\partial \tilde{a}_k^T x}{\partial x_i} \\
&= \sum_{k=1}^{n} \frac{\partial f(Ax)}{\partial (Ax)_k} . a_{ki} = \sum_{k=1}^{n} \partial_k f(Ax) a_{ki} \\
&= a_i^T \nabla f(Ax) \qquad \text{(A.8)}
\end{aligned}$$

which can be generalized as

$$\nabla_x f(Ax) = A^T \nabla f(Ax). \qquad \text{(A.9)}$$

Proposition 4. Assume B is an arbitrary symmetric matrix. Then,

$$\frac{d[trace(ABA^T)]}{dA} = 2AB. \qquad \text{(A.10)}$$

Proof. Let us replace $AB = f(A)$. Then,

$$\begin{aligned}
\nabla_A tr(ABA^T) &= \nabla_A tr(f(A)A^T) \\
&= \nabla_\varphi tr(f(\varphi)A^T) + \nabla_\varphi tr(f(A)(\varphi)^T) \\
&= (A^T)^T f'(\varphi) + \left(\nabla_{(\varphi)^T} tr(f(A)\varphi^T)\right)^T \\
&= AB^T + \left(\nabla_{(\varphi)^T} tr((\varphi)^T f(A))\right)^T \\
&= AB + (f(A)^T)^T \\
&= AB + AB = 2AB \qquad \text{(A.11)}
\end{aligned}$$

This completes the proof of (2.15).

Bibliography

[1] R. E. Kalman, “A new approach to linear filtering and prediction problems,” *Transactions of the ASME–Journal of Basic Engineering*, vol. 82, no. Series D, pp. 35–45, 1960.

[2] D. Simon and T. L. Chia, “Kalman filtering with state equality constraints,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 38, pp. 128–136, 2002.

[3] A. E. Nordsjo and S. B. Dynamics, “Target tracking based on kalman type filters combined with recursive estimation of model disturbances,” *IEEE International Radar Conference*, pp. 115–120, 2005.

[4] P. Mookerjee and F. Reifler, “Reduced state estimators for consistent tracking of maneuvering targets,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 41, pp. 608–619, 2005.

[5] A. D. Aguila, F. Cupertino, L. Salvatore, and S. Stasi, “Kalman filter estimators applied to robust control of induction motor drives,” *Proceedings of the 24th Annual Conference of the IEEE*, vol. 4, pp. 2257–2262, 1998.

[6] J. Persson and Y. Perriard, “Steady state kalman filtering for sensorless control of hybrid stepper motors,” *Electric Machines and Drives Conference*, vol. 2, pp. 1174–1177, 2003.

[7] R. G. Brown and P. Y. C. Hwang, *Introduction to Random Signals and Applied Kalman Filtering*. USA: John Wiley and Sons, Inc., 1992.

[8] N. Ma, M. Bouchard, and R. A. Goubron, “A perceptual kalman filtering based approach for speech enhancement,” *Signal Processing and Its Applications, Proceedings Seventh International Symposium*, vol. 1, pp. 373–376, 2003.

[9] C. Li and S. V. Andersen, “Integrating kalman filtering and multi pulse coding for speech enhancement with a non stationary model of the speech signal,” *Conference Record of the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers*, vol. 2, pp. 2300–2304, 2004.

[10] R. Lawrence and H. Kaufman, “The kalman filter for the equalization of a digital communications channel,” *IEEE Transactions on Communication Technology*, vol. 19, pp. 1137–1141, 1971.

[11] R. Patton and A. Killen, “Discrete adaptive filters for short lived dynamic systems,” *IEEE Transactions on Aerospace and Electronic Systems*, pp. 323–327, 1993.

[12] G. Chen, J. Wang, and L. S. Shieh, “Interval kalman filtering,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 33, pp. 250–259, 1997.

[13] P. D. Hanlon and P. S. Maybeck, “Characterization of kalman filter residuals in the presence of mismodeling,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 36, pp. 114–131, 2000.

[14] X. Shenshui, Z. Zhaoying, Z. Limin, X. Changsheng, and Z. Wendong, “Adaptive filtering of color noise using the kalman filter algorithm,” *Instrumentation and Measurement Technology Conference*, vol. 2, pp. 1009–1012, 2000.

[15] V. B. Tadic and V. Krishnamurthy, “Mean square asymptotic anaiysis of cross-coupled kalman filter state estimation algorithm for bilinear systems,” *American Control Conference*, vol. 2, pp. 881–886, 2002.

[16] T. H. Lee, W. S. Ra, Yoon, and J. B. Park, “Robust kalman filtering via krein space estimation,” *Control Theory and Applications, IEE Proceedings*, vol. 151, pp. 59–63, 2004.

[17] X. Lui and A. Goldsmith, “Kalman filtering with partial observation losses,” *43rd IEEE Conference on Decision and Control*, vol. 4, pp. 4180–4186, 2004.

[18] M. B. Luca, S. Azou, G. Burel, and A. Serbanescu, “On exact kalman filtering of polynomial systems,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 53, pp. 1329–1340, 2006.

[19] M. Farooq, S. Bruder, T. Quach, and S. S. Lim, “Adaptive filtering techniques for manoeuvring targets,” *Proceedings of the 34th Midwest Symposium on Circuits and Systems*, pp. 31–34, 1991.

[20] G. Soysal and M. Efe, “An adaptive kalman filter for tracking maneuvering targets,” *IEEE 14th Signal Processing and Communications Applications*, pp. 1–4, 2006.

[21] L. Xie, Y. C.Soh, and C. E. de Souza, “Robust kalman filtering for uncertain discrete time systems,” *IEEE Transactions on Automatic Control*, vol. 39, pp. 1310–1314, 1994.

[22] W. M. Haddad and D. S. Bernstein, “The optimal projection equations for reduced-order discrete-time state estimation for linear systems with multiplicative white noise,” *Syst. Contr. Lett.*, vol. 8, pp. 381–388, 1987.

[23] I. R. Petersen and D. C. McFarlane, “Robust state estimation for uncertain systems,” *Proc. 30th IEEE Conf Decision Contr*, December 1991.

[24] Y. Theodor and U. Shaked, “Robust discrete time minimum variance filtering,” *IEEE Transactions on Signal Processing*, vol. 44, pp. 181–189, 1996.

[25] M. Grimble, “Robust filter design for uncertain systems defined by both hard and soft bounds,” *IEEE Transactions on Signal Processing*, vol. 44, pp. 1063–1071, 1996.

[26] X. Lu, W. Wang, H. Zhang, and L. Xie, “Robust kalman filtering for discrete time systems with measurement delay,” *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 54, pp. 522–526, 2007.

[27] T. Perala and R. Piche, “Robust extended kalman filtering in hybrid positioning applications,” *WPNC 4th Workshop on Positioning, Navigation and Communication*, pp. 55–63, 2007.

[28] J. H. Manton, V. Krishnamurthy, and H. V. Poor, “James stein state filtering algorithms,” *IEEE Transactions on Signal Processing*, vol. 46, pp. 2431–2447, 1998.

[29] Y. C. Eldar, A. B. Tal, and A. Nemirovski, “Robust mean squared error estimation in the presence of model uncertainties,” *IEEE Transactions on Signal Processing*, vol. 53, pp. 168–181, 2005.

[30] J. Manton and Y. Hua, “Rank reduction and james-stein estimation,” *IEEE Transactions on Signal Processing*, vol. 47, pp. 3121–3125, November 1999.

[31] Y. Y. Guo and N. Pal, “A sequence of improvements over the james stein estimator,” *J. Multivariate Anal.*, vol. 42, pp. 302–317, 1992.

[32] E. L. Lehmann and G. Casella, *Theory of Point Estimation.* New York: Springer-Verlag, Inc., 1998.

[33] R. A. Horn and C. R. Johnson, *Matrix Analysis.* Cambridge, UK: Cambridge University Press, 1985.