Risk-Sensitive Filtering and Parameter Estimation

Jason Ford

DSTO-TR-0764

RISK-SENSITIVE FILTERING AND PARAMETER ESTIMATION

Jason Ford

Weapons Systems Division
Aeronautical and Maritime Research Laboratory

DSTO-TR-0764

ABSTRACT

This paper investigates the use of risk-sensitive filtering for state and parameter estimation in systems with model uncertainties. Modelling uncertainties arise from imperfectly known input process and noise characteristics as well as system model errors such as uncertain or time varying parameters of the system description. No new convergence results are given in this paper but simulation examples demonstrate that, in some situations, risk-sensitive filtering and estimation techniques allow for system uncertainties better than optimal techniques such as Kalman filtering.

APPROVED FOR PUBLIC RELEASE

DEPARTMENT OF DEFENCE
———————————— ◇ ————————————
DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION

19990413 051

AQF99-07-1286

DSTO–TR–0764

APPROVED FOR PUBLIC RELEASE

Risk-Sensitive Filtering and Parameter Estimation

EXECUTIVE SUMMARY

In control applications, including the control and guidance loops of modern guided missiles, filtering and system identification are two techniques for estimating unknown system information. Filtering provides information about dynamic missile states such as position and velocity, while system identification provides information about approximately constant quantities (known as the system model) that describe the missile's behaviour.

Traditional filtering techniques, such as Kalman filters, rely on assumptions about the system's structure. It is well known that a Kalman filter's performance can be dramatically reduced by errors in the system model on which the Kalman filter design is based. For example, if a missile is damaged or a missile is operating away from its nominal flight conditions than the system model for the missile will be incorrect and this may result in poor performance by the Kalman filter. This paper is concerned with techniques for relaxing some of the system model assumptions in a way that allows the performance of filters to degrade gracefully when faced with system modelling errors.

The main technique investigated in this paper is risk-sensitive filtering. It has been argued in the literature that, as the name suggests, risk-sensitive filters are sensitive to the risk (or uncertainty) in a system model and are better able to allow for system uncertainty (or possible errors in the system model) than so called "optimal" methods such as the Kalman filter.

This paper concludes that risk-sensitive filtering offers advantages over more traditional methods such as the Kalman filter when the system is not known with complete certainty (which is commonly the case). Additionally, this paper suggests that a new system identification technique known as risk-sensitive parameter estimation may offer advantages over existing system identification techniques. A more complete investigation and theoretical basis for risk-sensitive parameter estimation is required.

In a defence context, this paper suggests that risk-sensitive filters and risk-sensitive parameter estimation techniques may improve the robustness of a missile's control loops. Improved control loop robustness may enable reasonable missile performance when a missile is damaged or the missile is operating away from its nominal flight condition.

DSTO–TR–0764

THIS PAGE IS INTENTIONALLY BLANK

Author

Jason Ford
Weapons Systems Division

Jason Ford joined the Guidance and Control group in the Weapons Systems Division in February 1998. He received B.Sc. (majoring in Mathematics) and B.E. degrees from the Australian National University in 1995. He also holds the PhD (1998) degree from the Australian National University. His thesis presents several new on-line locally and globally convergent parameter estimation algorithms for hidden Markov models (HMMs) and linear systems.

His research interests include: adaptive signal processing, adaptive parameter estimation, risk-sensitive techniques and applications of HMM and Kalman filtering techniques.

THIS PAGE IS INTENTIONALLY BLANK

Contents

THIS PAGE IS INTENTIONALLY BLANK

1   Introduction

In control applications, filtering and system identification are two techniques for estimating unknown system information. System identification provides estimates of model parameters whilst filtering provides estimates of dynamic quantities such as state variables. A typical problem would involve using system identification to estimate the system model and then implementing filters based on the estimated system model. Filtering and system identification can therefore be seen as complementary techniques that can be used in tandem to achieve a desired objective.

Both system identification and filtering rely on assumptions about the system's structure. In filtering problems, system characteristics are assumed known (ie. the true system model is assumed known). While system identification can provide estimates of the unknown system dynamics, it is not possible to know the system with complete certainty. Likewise, system identification itself relies on assumptions such as: the true model is in the restricted class of models over which the identification is performed[22]. This paper is concerned with a technique for relaxing some of the assumptions made in both filtering and system identification.

The objective of the standard filtering problem is to find the state estimate for which the expected variance of the estimation error is minimized[1]. This minimum variance estimation is appealing in control (and other applications) because it can be seen as minimizing the "energy" in the estimation error. Unfortunately techniques which are designed assuming complete system knowledge, such as the Kalman filter, do not necessarily provide optimal estimates when there is system uncertainty[23].

Similarly, when system identification is used to reduce system uncertainty, it should be remembered that many simplifying assumptions underlie the identification process[22]. It should be noted that system identification is always performed over a restricted class of models, eg. linear systems of fixed order[22]. It is also often assumed that the true model is a member of the identification model set. In many applications, the objective of identification is to estimate the system model in the model set closest to the true system in an output error sense.

In control applications, the measure of "closeness" generally used is the prediction error variance[22]. That is, the estimated model is the model whose outputs (predictions of the system outputs) are closest in a variance sense to the real system outputs. In this way, system identification has an analogous performance index to the filtering problem.

This paper considers an alternative filtering problem known as *risk-sensitive filtering*[8]. Although this paper is motivated by control problems, the discussion is limited to zero input systems and this paper is a preliminary step towards applying risk-sensitive techniques to filtering for control problems. Risk-sensitive filters minimize an exponential of the error cost, which penalizes the higher order moments in the estimation error, that is, moments other than the variance[5]. It has been argued that system uncertainty appears in these higher moments and hence it is argued that risk-sensitive filters are more "robust" to system uncertainties than minimum variance estimators[5, 8] . This can be interpreted as meaning that the risk-sensitive filter can provide estimates that are better, in an error variance sense, than a Kalman filter, when both are based on the same model

assumptions[23]. A complete understanding of the type of uncertainties that lead to this situation has not yet been completed, but is the subject of continuing research[23].

One application relevant to defence is the design of control loops for missiles. Kalman filters are commonly used in the control and guidance loops of modern guided missiles. Risk-sensitive filters and other techniques may improve the robustness of a missile's control loops. Improved control loop robustness may enable reasonable missile performance when a missile is damaged or the missile is operating away from its nominal flight condition.

This paper also proposes a *risk-sensitive parameter estimation* problem. When the true model is not in the model set we suggest that there is important information in the higher error moments of the prediction error (analogous to the risk-sensitive filtering problem). It is suggested in this paper, without proof, that a risk-sensitive parameter estimation approach allows for the inability of the model class to perfectly represent the true system.

Many of the definitions and results given in this paper are well known in the established literature on risk-sensitive filtering and control. We have tried to reference the sources of results as they appear in this paper.

The paper is organized as follows: In Section 2 the risk-sensitive filtering problem is presented. The risk-sensitive filter for linear systems is given and an example is presented which compares the Kalman filter with the risk-sensitive filter when the system is not known with complete certainty. This section is a review of existing results. In Section 3 a new research problem is proposed which we call risk-sensitive parameter estimation. An parameter estimation example is given that compares the use of Kalman filter state estimates with the use of risk-sensitive filter state estimates. Finally, in Section 4 some conclusions are presented.

2 Risk-Sensitive filtering

2.1 Minimum Variance Estimation

We proceed with the notation used in [6, 8]. Consider the following stochastic, discrete-time state space system (also known as a Gauss-Markov linear system) defined on a probability space $(\Sigma, \mathcal{F}, P)$:

$$
\begin{aligned}
x_{k+1} &= A x_k + B w_{k+1}, & x_0 &\in R^{N \times 1} \\
y_k &= C x_k + D v_k, & y_k &\in R^{P \times 1}
\end{aligned}
\tag{2.1}
$$

where $k \in Z^+$; $x_k, B \in R^{N \times 1}$; $y_k, D \in R^{P \times 1}$; $w_k, v_k \in R$; $A \in R^{N \times N}$; and $C \in R^{P \times N}$. Here, x_k denotes the state of the system, y_k denotes the measurement, and w_k and v_k are the process noise and the measurement noise, respectively. It is assumed that the noises are independently and identically distributed (*iid*), zero mean unit variance Gaussian random variables, ie $w_k, v_k \sim N[0, 1]$. We denote sequences by bold face letters subscripted by the index range, for example the sequence $\{w_0, \ldots, w_k\}$ is denoted by $\mathbf{w}_{0,k}$. It is assumed that $\mathbf{w}_{0,k}$, $\mathbf{v}_{0,k}$ and x_0 are mutually independent. We also assume that x_0 (or an *a priori* distribution for x_0) is given. For simplicity we have not considered time-varying matrices but these are not excluded by the theory.

The conditional mean estimate is defined as follows

$$\hat{x}_k^{mv} = E[x_k | \mathbf{y}_{0,k}, x_0], \tag{2.2}$$

where $E[.|.]$ denotes conditional expectation on the probability space $(\Sigma, \mathcal{F}, P)$ and $\hat{x}_k^{mv}$ denotes the conditional mean estimate at time k. The conditional mean estimate is equivalent to the minimum variance estimate given as follows

$$\hat{x}_k^{mv} = \arg\min_{\omega} \left\{ V_k(\omega) = E[(x_k - \omega)'(x_k - \omega) | \mathbf{y}_{0,k}, x_0] \right\}. \tag{2.3}$$

where $\arg\min_{\omega} F(\omega)$ is the value of the argument ω that minimizes the cost $F(\omega)$ and the prime symbol $'$ denotes the transpose.

It is well known [1] that for Gauss-Markov linear systems, eg. (2.1), when the system is known, that the optimal filter for conditional mean estimates is the Kalman filter which happens to be a finite-dimensional filter.

When the true model is not known, the Kalman filter implemented assuming a model estimate $\hat{\lambda}$, gives estimates

$$\hat{x}_{k|\hat{\lambda}}^{kf} = E[x_k | \mathbf{y}_{0,k}, x_0, \hat{\lambda}] \tag{2.4}$$

where $\hat{x}_{k|\hat{\lambda}}^{kf}$ denotes the Kalman filter estimate at time k based on the assumed model $\hat{\lambda}$. When $\hat{\lambda}$ is not the true system, the Kalman filter estimates $\hat{x}_{k|\hat{\lambda}}^{kf}$ are generally not minimum variance estimates[23].

2.1.1 Estimator Performance Index

The performance analysis of estimators given below follows the presentation given in [23].

To enable comparison of different filters (or estimators) we introduce a cost associated with a filter, termed the *expected estimator cost*, as follows,

$$W(\phi) = E_{\mathcal{A}}[\bar{W}(\phi)] \tag{2.5}$$

where

$$\bar{W}(\phi) = E\left[(x_k - \hat{x}_{k|\hat{\lambda}}^{\phi})'(x_k - \hat{x}_{k|\hat{\lambda}}^{\phi}) \Big| \mathbf{y}_{0,k}, x_0, \lambda \right] \tag{2.6}$$

Here ϕ is a particular filter and $\hat{x}_{k|\hat{\lambda}}^{\phi}$ denotes the estimate of x from the filter ϕ at time k based on an assumed model $\hat{\lambda}$. The cost $\bar{W}(\phi)$ is termed the *estimation cost* and for large k, if the system is ergodic, converges to the *measured estimation cost*,

$$\bar{W}^m(\phi) = \frac{1}{T} \sum_{k=1}^{T} (x_k - \hat{x}_{k|\hat{\lambda}}^{\phi})'(x_k - \hat{x}_{k|\hat{\lambda}}^{\phi}). \tag{2.7}$$

The symbol $E_{\mathcal{A}}[.]$ denotes expectation on the probability space $(\mathcal{A}, \mathcal{F}_{\mathcal{A}}, P_{\mathcal{A}})$ where $\mathcal{A}$ is the set (or space) whose elements α denote the possible dynamics of the system, $\mathcal{F}_{\mathcal{A}}$ is a σ-algebra on $\mathcal{A}$, and $P_{\mathcal{A}}$ is a probability function on $\mathcal{F}_{\mathcal{A}}$ which denotes the probability of particular dynamics $\alpha \subset \mathcal{A}$. This probability space provides a probabilistic description of

the unknown system dynamics and enables comparison of different filters or estimators. It should be noted that we consider the unknown but fixed model, λ, as a random variable. This is a different approach used in [23].

The *minimum cost estimator* is defined as

$$\hat{\phi} = \arg \min_{\phi \in \Phi} W(\phi) \tag{2.8}$$

where $\hat{\phi}$ denotes the *minimum cost estimator* and Φ denotes the set of possible estimators.

When $\mathcal{A}$ has one member, a Gauss-Markov linear system of the form (2.1), then the Kalman filter is the minimum cost estimator as defined by (2.8).

The key point of this paper is that for other $\mathcal{A}$, a filter other than the Kalman filter may be the minimum cost estimator[23]. In [23], examples are given where the measured estimation cost of the risk-sensitive filter is less than the measured estimation cost of the Kalman filter.

2.2   Risk-Sensitive Filtering

The following description of risk-sensitive filtering comes from [8]. The results were first established in [5]. Motivated by the desire to improve filter performance when system uncertainties exist we consider a filtering problem which seeks to minimize an exponential of the error performance index.

Analogous to the minimum variance estimate definition, the risk-sensitive filter estimate, based on an assumed model, is defined as[8],

$$\hat{x}^{rs}_{k|\hat{\lambda}} = \arg \min_{\omega \in R^n} J_k(\omega), \tag{2.9}$$

where

$$J_k(\omega) = E\left[\exp(\theta \Phi_{0,k}(\omega)) \Big| \mathbf{y}_{0,k}, x_0, \hat{\lambda}\right], \quad \theta > 0. \tag{2.10}$$

Here,

$$\Phi_{0,k}(\omega) = \hat{\Phi}_{0,k-1} + \frac{1}{2}(x_k - \omega)' Q_k (x_k - \omega), \tag{2.11}$$

where

$$\hat{\Phi}_{m,n} = \frac{1}{2} \sum_{k=m}^{n} (x_k - \hat{x}^{rs}_{k|\hat{\lambda}})' Q_k (x_k - \hat{x}^{rs}_{k|\hat{\lambda}}), \tag{2.12}$$

where $\theta > 0$ is the risk sensitivity parameter and $Q_k > 0$ is a weighting matrix. Here $\hat{\lambda}$ is the assumed model and is not necessarily equal to the true model λ.

The risk parameter can be thought of as describing the amount of uncertainty in the system description. The larger the θ value the greater the model, $\hat{\lambda}$, is believed to be in error. Conversely, as $\theta \to 0$ the model is believed with more certainty and the risk-sensitive filter approaches the Kalman filter, see [8] for details.

We do not go into details here but finite dimensional solutions to the risk-sensitive problem for linear systems have been presented previously [6, 8]. We present the risk-sensitive filter for linear systems in a later section.

Remarks

1. The risk-sensitive state estimate does not have an interpretation as a conditional mean estimate.

2. This is perhaps not the obvious risk-sensitive cost but it is shown in [8] that

$$\hat{x}_{k|\hat{\lambda}} = \arg\min_{\omega} E\left[\exp\left(\frac{\theta}{2}(x_k - \omega)'Q(x_k - \omega)\right)\middle| y_{0,k}, x_0, \hat{\lambda}\right] \qquad (2.13)$$

results in the same solution as the minimum variance (or risk-neutral) problem.

2.2.1 The Risk-Sensitive Cost does Penalize Higher Moments

To see how risk-sensitive estimation penalizes higher moments consider a scalar system model and set $Q_k = 1$. In this case we can write the cost as

$$
\begin{aligned}
J_k(\omega) &= E\left[\exp\left(\frac{1}{2}\theta(\omega - x_k)^2 + \theta\hat{\Phi}_{0,k-1}\right)\middle| y_{0,k}, x_0, \hat{\lambda}\right] \\
&= E\left[F_{k-1,\theta} \times \exp\left(\frac{1}{2}\theta(\omega - x_k)^2\right)\middle| y_{0,k}, x_0, \hat{\lambda}\right], \quad \theta > 0
\end{aligned}
$$

where $F_{k-1,\theta} := \exp\left(\theta\hat{\Phi}_{0,k-1}\right)$ is a factor independent of ω. Now writing the second exponential as an infinite series we get

$$J_k(\omega) = \frac{1}{2}E\left[F_{k-1,\theta} \times \left(1 + \frac{\theta(\omega - x_k)^2}{2} + \frac{1}{2}\frac{\theta^2(\omega - x_k)^4}{2^2} + \frac{1}{3!}\frac{\theta^3(\omega - x_k)^6}{2^3} + \ldots\right)\middle| y_{0,k}, x_0, \hat{\lambda}\right].$$

The terms $(\omega - x_k)^{2p}$ for $p > 1$ are the higher order moments that are not considered in minimum variance estimation. That is, the risk-sensitive cost penalizes error contributions from these higher moments whenever $\theta > 0$.

2.2.2 Risk-Sensitive Filtering for Linear Systems

Consider the Gauss-Markov linear system given earlier (2.1). The following theorem holds.

Theorem 1. *The optimal risk-sensitive estimate, $\hat{x}^{rs}_{k|\hat{\lambda}}$, defined in (2.9), can be expressed as*

$$\hat{x}^{rs}_{k|\hat{\lambda}} = A\hat{x}^{rs}_{k-1|\hat{\lambda}} + (R_k^{-1} + C'D^{-1}C)^{-1}C'D^{-1}(y_k - CA\hat{x}^{rs}_{k-1|\hat{\lambda}}) \qquad (2.14)$$

where $(R_k^{-1} + C'D^{-1}C - \theta Q) > 0$ for all k and R_k satisfies the following Riccati equation

$$R_{k+1} = B + A(R_k^{-1} + C'D^{-1}C - \theta Q)^{-1}A', \quad R_0 > 0 \qquad (2.15)$$

Proof: This was first proven in [5]. It is also shown in [8]. ☐

Remarks:

1. The risk-sensitive filter for Gauss-Markov linear systems is finite dimensional.

2. Note that the risk-sensitive filter is equivalent to the Kalman filter when $\theta = 0$.

3. Unlike the Kalman filter, the *a priori* and *a posteriori* estimates (or one-step-ahead predictions) for the risk-sensitive filtering problem are not simply related through A. See [3] for details of the predictive risk-sensitive filter.

Example 1. (Risk Sensitive Filtering.)

To demonstrate the possible improvement in state estimation consider the linear system, given earlier (2.1), with $A, C, B, D = 1$ and $x_0 = 0$. The state sequence $\mathbf{x}_{0,k}$ is measured indirectly via the observations $\mathbf{y}_{0,k}$.

The parameters B and C are known correctly, but A and D are not known. Consider the filtering problem where $\mathcal{A}$ is the set of three possible models with $(A = 0.8, D = 1.2)$, $(A = 0.9, D = 1.2)$ and $(A = 1, D = 1.2)$ respectively. Assume that the *a priori* probability of these models is equal.

We compare the performance of the Kalman filter and a risk-sensitive filter $(\theta = 0.5, Q = 1)$ on the basis of the expected estimator cost, ie. (2.5), and measured estimation cost, ie. (2.7).

Figure 1 shows both the risk-sensitive $(\theta = 0.5, Q_k = 1)$ and Kalman filter estimates against the true state value using the assumed model $\hat{A} = 0.9, \hat{D} = 1.2$. The risk-sensitive filter has smaller measured estimation cost than the Kalman filter. That is, $\bar{W}^m(\text{RS}) = 0.002442$ while $\bar{W}^m(\text{KF}) = 0.002480$.
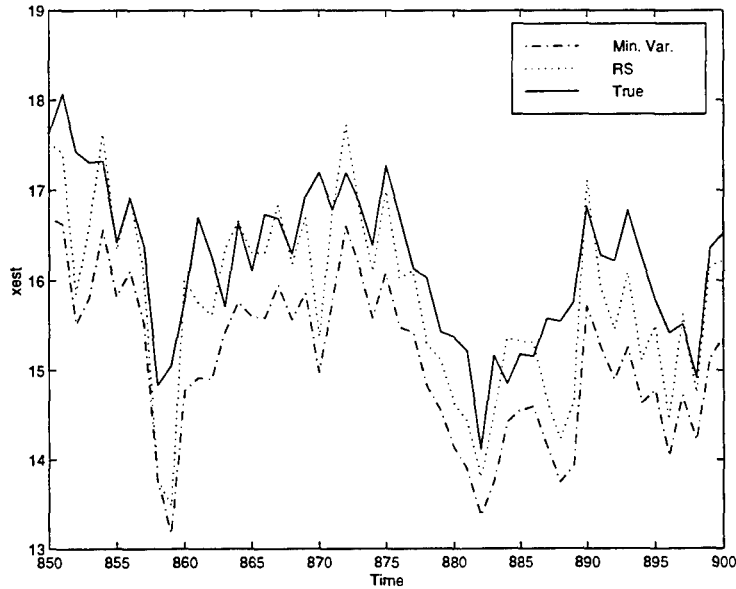

Figure 1 (U): Comparison of risk-sensitive filter and kalman filter estimates

We compared the performance of the filters is two ways. Firstly, we compare the measured estimator cost of the two filters for three different model assumptions when filtering data

from one unknown system $(A, B, C, D = 1)$. Secondly, we compared the estimator cost of the two filters (based on one model assumption $(\hat{A} = 0.9, \hat{D} = 1.2)$) when filtering data generated from the three systems in $\mathcal{A}$.

The first comparison examines the effect of varying model assumptions on the performance of the filters while the second comparison examines the ability of one fixed filter of each type (Kalman filter and risk-sensitive filter) to filter data generated from a variety of systems.

Table 1: The performance of filter for different assumed models when true model is $A, C, B, D = 1$.

Model (C,D known)	$\bar{W}^m(KF)$	$\bar{W}^m(RS)$
$\hat{A} = 0.8, \hat{B} = 1.2$	0.002515	0.002505
$\hat{A} = 0.9, \hat{B} = 1.2$	0.002480	0.002442
$\hat{A} = 1.0, \hat{B} = 1.2$	0.002618	0.002529

Table 2: The performance of filters on different systems when assumed model is $\hat{A} = 0.9, \hat{B}, \hat{C} = 1, \hat{D} = 1.2$

Model ($B, C, D = 1$)	$\bar{W}^m(KF)$	$\bar{W}^m(RS)$
$A = 0.8$	0.002550	0.002531
$A = 0.9$	0.002480	0.002442
$A = 1.0$	0.004887	0.003923
$W(\phi)$	0.003306	0.002965

Table 1 shows the results of the first comparison while Table 2 shows the results of the second comparison

The risk-sensitive filter performs better than the Kalman filter in all the situations presented in the tables. From Table 2, and using the fact that the systems are ergodic, a value for the expected estimator cost of the filters can be calculated and for this example the risk-sensitive filter has the lower expected estimator cost.

Risk-Sensitive Filtering Solutions Summary

Finite dimensional solutions for this problem have be found in particular situations including: linear systems [5, 8], bilinear systems [7] and zero process noise case ($w_k = 0$) [7]. It has also been shown that finite-dimensional filters exist for a class of discrete-time nonlinear systems[10]. Finite dimensional solutions can been obtained for more general nonlinear systems by using a generalized risk-sensitive cost index which is chosen to absorb the contribution from the nonlinear terms[9]. Solutions for the corresponding continuous-time problem are also available for linear systems [11].

The foundations for the risk-sensitive problem where introduced in the following papers that focus on the risk-sensitive control problem[15, 18, 19], see the remark below. Applications for the technique are described in [2, 13]. Recently, a book on risk-sensitive

control has been published[20]. Later work on the risk-sensitive problem can be found in [5, 16, 21].

Remark:

1. The risk sensitive control, u^{RS}, is defined as minimizing the risk-sensitive control cost:

$$\bar{J}(u) = E\left[\theta \exp\left(\frac{1}{2}\sum_{k=1}^{T-1}(x_k'Q_k x_k + u_k'R_k u_k) + x_T'Q_T x_T\right)\right],$$

where $u = \{u_1, ..u_{T-1}\}$ is a control sequence, $\theta > 0$ is the risk parameter and T is the input length. It is generally assumed that $Q_k \geq 0$ and $R_k > 0$ for all k.

3 Risk-Sensitive Parameter Estimation

In this section we discuss parameter estimators and propose a risk-sensitive parameter estimation problem. Before introducing the parameter estimation problem we will discuss the sources of model uncertainty that effect parameter estimation.

3.1 Model Uncertainty

Parameter estimation or system identification can be viewed as a technique to allow for model uncertainty. In the broadest sense, the objective of system identification is find the system that created the observed data. However, in practice, the class of models over which the search is performed needs to be restricted for computational and complexity reasons. Assumptions about the underlying system need to be made. The objective of system identification then becomes to find the model within the model class that best describes the observed data. To enable searching, model classes are parameterized and the system identification problem becomes a parameter estimation problem.

3.2 Adaptive Estimation

The need to use adaptive estimation arises in situations where all the quantities needed to estimate a parameter are not directly available, but the required quantities themselves can be estimated. For example, consider again the linear system,

$$\begin{aligned} x_{k+1} &= Ax_k + Bw_{k+1}, & x_0 &\in R^N \\ y_k &= Cx_k + Dv_k, & y_k &\in R \end{aligned} \tag{3.1}$$

where $k \in \mathcal{Z}^+$; $x_k, B \in R^{N\times 1}$; y_k, D, v_k and $w_k \in R$; $A \in R^{N\times N}$ and $C \in R^{1\times N}$. Also, the sequences $\mathbf{w}_{0,k}$ and $\mathbf{v}_{0,k}$ are sequences of *iid*, zero mean, unit variance Gaussian random variables. It is assumed that $\mathbf{w}_{0,k}, \mathbf{v}_{0,k}$ and x_0 are mutually independent random variables. Also, it is assumed that x_0 is given. Here, x_k denotes the state of the system which is observed via the observations, y_k.

Assume A, B and D are known and that we are interested in estimating C.

The recursive least squares algorithm for estimating C requires that x_k is known (which in this problem it is not). However, x_k can be estimated via a Kalman filter. If the estimates $\hat{x}_{k|\hat{\lambda}}^{kf}$ from a Kalman filter are substituted into the least squares algorithm in lieu of x_k then we have an adaptive estimation algorithm known as Extended Least Squares (ELS). That is,

$$
\begin{aligned}
\hat{C}_{k+1} &= \hat{C}_k + \frac{1}{k} P_k (y_k - \hat{C}_k \hat{x}_{k|\hat{\lambda}}^{kf}) \\
P_k^{-1} &= \frac{k-1}{k} P_{k-1}^{-1} + \frac{1}{k} (\hat{x}_{k|\hat{\lambda}}^{kf})^2
\end{aligned}
\tag{3.2}
$$

where $\hat{x}_{k|\hat{\lambda}}^{kf}$ are Kalman filter estimates.

In general, convergence results for ELS algorithms can not be established; however, there are many adaptive estimation algorithms for which strong convergence results have been established.

Consider again the linear system (3.1). Assuming that B and D are known, it is possible to estimate A and C as follows:

$$
\begin{aligned}
\hat{A}_k &= \hat{J}_k \hat{O}_k^{-1}, \qquad \hat{A}_0 \in R^{N \times N} \\
\hat{C}_k &= \hat{T}_k \hat{O}_k^{-1}, \qquad \hat{C}_0 \in R^{1 \times N}
\end{aligned}
\tag{3.3}
$$

when $\hat{O}_k^{-1}$ exists, where $\hat{A}_0$ and $\hat{C}_0$ are initial guesses for the parameters and

$$
\begin{aligned}
\hat{J}_k &= E[J_k | \mathbf{y}_{0,k}, x_0, \hat{\mathbf{A}}_{0,k-1}, \hat{\mathbf{C}}_{0,k-1}], \\
\hat{O}_k &= E[O_k | \mathbf{y}_{0,k}, x_0, \hat{\mathbf{A}}_{0,k-1}, \hat{\mathbf{C}}_{0,k-1}] \qquad \text{and} \\
\hat{T}_k &= E[T_k | \mathbf{y}_{0,k}, x_0, \hat{\mathbf{A}}_{0,k-1}, \hat{\mathbf{C}}_{0,k-1}].
\end{aligned}
\tag{3.4}
$$

Here,

$$
J_k := \sum_{\ell=1}^{k} x_{\ell+1} x_\ell', \qquad O_k := \sum_{\ell=1}^{k} x_\ell x_\ell' \qquad \text{and}
$$

$$
T_k := \sum_{\ell=1}^{k} y_\ell x_\ell'.
\tag{3.5}
$$

Note the notation $\hat{\mathbf{A}}_{0,k-1}$ denotes the sequence $\{\hat{A}_0, \hat{A}_1, \ldots, \hat{A}_{k-1}\}$. Filters for $\hat{J}_k$, $\hat{O}_k$ and $\hat{T}_k$ are given in [12].

It has been shown in [14] that if the output data was generated by (3.1) and the model order, N, is known then the estimates $\hat{A}_k$ and $\hat{C}_k$ almost surely converge to the true A and C model parameters.

3.3   Risk-Sensitive Parameter Estimation

There are two situations in which a risk-sensitive approach may be appropriate, firstly, when the true model, λ, is not in the model set, Λ, and secondly, from poor initializations

$\hat{\lambda}_0$. The adaptive estimation algorithm (3.3),(3.4) assumes that the true system model is in the model set. If the model is not in the model set and the state sequence is not measured then estimation may not be optimal in a prediction error sense.

Even if the true model is in the model set, convergence close to the true model may be slow from poor initial guesses. It has been suggested from simulation evidence that a risk-sensitive approach to estimating the quantities J_k, O_k and T_k may improve convergence from poor initializations[17]. These two situations motivate an investigation of risk-sensitive parameter estimation.

3.3.1   Risk-Sensitive Adaptive Estimation

In this subsection we propose a risk-sensitive estimation algorithm without study. Consider again the linear system (2.1).

Assuming that B and D are known, it is possible to estimate A and C as follows:

$$
\begin{aligned}
\hat{A}_k^{RS} &= \hat{J}_k^{RS} \left(\hat{O}_k^{RS} \right)^{-1}, && \hat{A}_0^{RS} \in R^{N \times N} \\
\hat{C}_k^{RS} &= \hat{T}_k^{RS} \left(\hat{O}_k^{RS} \right)^{-1}, && \hat{C}_0^{RS} \in R^{1 \times N}
\end{aligned}
\tag{3.6}
$$

when $\left(\hat{O}_k^{RS} \right)^{-1}$ exist, where $\hat{A}_0^{RS}$ and $\hat{C}_0^{RS}$ are initial guesses for the parameters and $\hat{J}_k^{RS}$, $\hat{O}_k^{RS}$ and $\hat{T}_k^{RS}$ are risk-sensitive estimates for the quantities J_k, O_k and T_k respectively.

The following example examines the use of risk-sensitive filter estimates in a parameter estimation problem.

Example 2. (Using Risk-Sensitive Filter Estimates for Parameter Estimation.)

Consider the following linear system

$$
\begin{aligned}
x_{k+1} &= A x_k + B w_{k+1}, && x_0 \in R \\
y_k &= C x_k + D v_k, && y_k \in R
\end{aligned}
$$

where $A = 0.9$, $C = 1$; $B, D = 0.1$; $x_0 = 0$ and w_k, v_k are *iid*, zero mean unit variance Gaussian random variables. Here, the state sequence $x_{0,k}$ is measured indirectly via the observations $y_{0,k}$ and we are interested in estimation of A.

The true model is not known and the following system parameters are assumed: $\tilde{B} = 0.1$, $\bar{C} = 0.6$ and $\bar{D} = 0.1$. Our initial guess for A is $\hat{A}^0 = 0.6$. Here, estimation of A is performed over the model set ($\bar{B} = 0.1$, $\bar{C} = 0.6$ and $\bar{D} = 0.1$) which does not contain the true system.

If the state sequence $x_{0,k}$ was measured then the least squares estimate of A would be

$$
\hat{A} = \left(\sum_{\ell=1}^{T} x_\ell x_{\ell-1} \right) / \left(\sum_{\ell=1}^{T} x_{\ell-1}^2 \right)
$$

where T is the number of data points.

However, when the state sequence is not measured then a multi-pass missing data approach [14, 24] can be used. Here filter estimates of the state are used in lieu of the true state and A estimated on pass ℓ as follows:

$$\hat{A}^\ell = \left(\sum_{k=1}^T \hat{x}_{k|\hat{A}^{\ell-1}} \hat{x}_{k-1|\hat{A}^{\ell-1}} \right) / \left(\sum_{k=1}^T \hat{x}^2_{k-1|\hat{A}^{\ell-1}} \right).$$

where $\hat{x}_{k|\hat{A}^{\ell-1}}$ is a estimate of the state at time k based on model assumptions ($\bar{B} = 0.1$, $\bar{C} = 0.6$ and $\bar{D} = 0.1$) and $\hat{A}^{\ell-1}$, either from the Kalman filter or from a risk-sensitive filter. Passes though the data are performed until $\hat{A}^\ell$ converges to some value.

We compare parameter estimation using Kalman filter estimates $\hat{x}^{kf}_{k|\hat{\lambda}}$ with estimation using risk-sensitive filter estimates $\hat{x}^{rs}_{k|\hat{\lambda}}$.

A data set of 1000 points was generated with the above parameter values. First, Kalman filter estimates were used and after 10 passes A was estimate as 0.9474. Then risk-sensitive filter estimates ($\theta = 35$, $Q = 1$) were used and after 10 passes A was estimated as 0.9049. This corresponds to an improvement in model performance, as measured by filtered output error (that is, $E[(y_k - \hat{y}^\phi_k)^2 | \mathbf{y}_{0,k}, x_0] \approx \frac{1}{k} \sum_{\ell=1}^k (y_\ell - \bar{C}\hat{x}^\phi_\ell)^2)$ from 0.002964 for the Kalman filter estimate to 0.001879 for the risk-sensitive filter estimate.

Convergence to these values occurred for a range of choices for $\hat{A}^0$. Similar improvements in estimation of A occurs using risk-sensitive filters if the assumed model had $\bar{C} = 0.8$ or $\bar{C} = 0.9$.

□

Remarks

1. The parameter estimation problem and the approach presented in the above example is admittedly contrived and unlikely to occur in practice. Estimation of both A and C using standard techniques would be an obvious approach and would result in a better model estimate. However, the success of the risk-sensitive approach in this artificial problem motivates investigation of risk-sensitive approaches in more complicated problems.

2. The more usual measure of model performance is the prediction error but the risk-sensitive filter shown in this paper can not be used to generate predictions (see early comment and see [3] for the risk-sensitive predictor). Hence, for convenience the filtered output error has been used for comparison instead. There is a similar improvement in the prediction error of the risk-sensitive predictor over the Kalman filter predictor when they are based on the models estimated in this example.

3. The missing data approach used in this problem can be considered an example of an adaptive estimator (3.6) where $\hat{J}^{rs}_k = \sum_{\ell=1}^k \hat{x}^{rs}_{\ell|\hat{\lambda}} \hat{x}^{rs}_{\ell-1|\hat{\lambda}}$ etc. Convergence results or properties have not yet been established for the presented risk-sensitive estimation algorithm.

4 Conclusion

Unknown model dynamics can make the control problem difficult. Filtering and system identification are two techniques used to handle system uncertainties.

This paper repeated the known risk-sensitive filtering problem and the known solution for Gauss-Markov linear systems. An example was presented which compared the Kalman filter with a risk-sensitive filter in a situation where the system parameters were not known completely.

The key contribution of this paper is the proposal of the risk-sensitive parameter estimation problem. An example was presented which demonstrates a possible application of a risk-sensitive approach to the parameter estimation problem. In this example, it was shown that a better model (in an output error sense) could be found by using risk-sensitive state estimate than by using Kalman filter state estimates. This suggest that a more theoretical and complete investigation of risk-sensitive parameter estimation may be worthwhile.

References

1. B.D.O. Anderson and J.B. Moore, *Optimal Filtering*, Prentice Hall, New Jersey, 1979.

2. J.L. Speyer and K.D. Kim. "Passive homing missile guidance law based on new target maneuver models," *AIAA Jour. of Guidance, Control and Dynamics*, Vol. 13, No. 5, pp. 803-812, 1990.

3. R.N. Banavar and J.L. Speyer, "Properties of risk-sensitive filters/estimators", *IEE Proc.-Control Theory Appl*, Vol. 145, No. 1, pp. 106-112, Jan. 1998.

4. M.R. James, J.S. Baras and R.J. Elliott, "Risk-sensitive control and dynamics games for partially observed discrete-time nonlinear systems", *IEEE Trans. on Automatic Control*, AC-39, No. 4, pp. 780-792, April 1994.

5. J.L. Speyer, C. Fan and R.N. Banavar, "Optimal stochastic estimation with exponential cost criteria", *Proc. of Conference on Decision and Control, CDC92*, pp. 2293-2298, Arizona, December 1992.

6. S. Dey and J.B. Moore, "Risk-sensitive filtering and smoothing via reference probability methods," *Proc. of American Control Conference, ACC95*, Vol. 1, pp. 129-133, Seattle, June 1995.

7. S. Dey and J.B. Moore, "On finite-dimensional risk-sensitive estimation," *Proc. of International Symposium on Signal Processing and its Applications, ISSPA96*, Vol. 3, pp. 849-852, Gold Coast, August 1996.

8. S. Dey and J.B. Moore, "Risk-sensitive filtering and smoothing via reference probability methods," *IEEE Trans. on Automatic Control*, AC-42, No. 11, November 1997.

9. J.B. Moore and J.S. Baras, "Finite-dimensional filters for a class of nonlinear discrete-time systems," *System and Control Letters*, Vol. 26, pp. 223-230, 1995.

10. J. Levine and G. Pignie, "Exact finite-dimensional filters for a class of nonlinear discrete-time systems," *Stochastics*, Vol. 18, pp. 97-132, 1986.

11. V.E. Benes and R.J. Elliott, "Finite-dimensional risk-sensitive information states," *Proc. of the IFAC Symposium on Nonlinear Control System Design (NOLCOS)*, Vol. 2, pp. 471-476, June 1995.

12. R.J. Elliott and V. Krishnamurthy, "Finite Dimensional Filters for ML Estimation of Discrete-time Gauss-Markov Models," in *Proc. of CDC'97*, San Deigo, Dec. 1997.

13. J.L. Speyer, "An adaptive terminal guidance scheme based on an exponential cost criterion with application to homing missile guidance, " *IEEE Trans. on Automatic Control*, AC-26, pp. 371-375, 1976.

14. R.J. Elliott, J.J. Ford and J.B. Moore, "Almost -sure Parameter Estimation for Partially Observed Discrete-time Linear Systems," Submitted to *International Journal of Adaptive Control and Signal Processing*, May 1998.

15. D.H. Jacobsan, "Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games," *IEEE Trans. on Automatic Control*, AC-18, pp. 124-131, 1973.

16. P. Whittle, "Risk-sensitive maximum principle: The case of imperfect state observation," *IEEE Trans. on Automatic Control*, AC-36, pp. 793-801, 1991.

17. S. Athuraliya, J.J. Ford and J.B. Moore, "Faster parameter estimation using risk-sensitive filters", to appear in *CDC98* and submitted to *International Journal of Adaptive Control and Signal Processing* 1998.

18. P. Whittle, "Risk-sensitive linear/quadratic /Gaussian control," *Adv. App. Probability 13*, pp. 746-777, 1981.

19. J. Speyer, J. Deyst and D.H. Jacobson, "Optimization of stochastic linear systems with additive measurement and process noise using exponential performance criteria,", *IEEE Trans. on Automatic Control*, AC-19, pp. 358-366, August 1974.

20. P. Whittle, *Risk-sensitive optimal control*, John Wiley and Sons, New York, 1990.

21. C.R. Jaensch and J. Speyer, "Stochastic control problems with an exponential cost criteria", *Proc. of Conference on Decision and Control, CDC88*, pp. 286-291, December 1988.

22. L. Ljung, *System Identification: Theory for the user*, Prentice-Hall, Upper Saddle River, New Jersey, 1987.

23. R.K. Boel, M.R. James and I.R. Petersen, "Robust and Risk-Sensitive Filtering", *Proc. of Conference on Decision and Control, CDC97*, pp. 2273-2278, San Diego, California, Dec. 1997.

24. G.J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*, John Wiley & Sons, New York, 1997.

DISTRIBUTION LIST

Risk-Sensitive Filtering and Parameter Estimation

Jason J. Ford

AUSTRALIA

DEFENCE ORGANISATION

S&T Program

Chief Defence Scientist

FAS Science Policy ⎫ shared copy   1

AS Science Corporate Management ⎭

Director General Science Policy Development   2

Counsellor Defence Science, London (Doc Data Sheet only)

Counsellor Defence Science, Washington (Doc Data Sheet only)

Scientific Adviser to MRDC Thailand (Doc Data Sheet)

Director General Scientific Advisers and Trials ⎫ shared copy   3

Scientific Adviser Policy and Command ⎭

Navy Scientific Adviser (Doc Data Sheet and distribution list only)

Scientific Adviser - Army (Doc Data Sheet and distribution list only)

Air Force Scientific Adviser (Doc Data Sheet and distribution list only)

Director Trials (Doc Data Sheet and distribution list only)

Aeronautical and Maritime Research Laboratory

Director, AMRL   4

Chief, Weapons Systems Division   5

Research Leader Air Weapons Systems   6

Head, Guidance and Control   7

Task Manager, Dr Peter Simms WSD   8

Reviewer: Jeremy Thorne, RSISE, Australian National University   9

Author(s): Dr J. Ford   10

DSTO Library

Library Fishermens Bend   11

Library Maribyrnong   12

Library Salisbury (2 copies)   13,14

Australian Archives   15

Library, MOD, Pyrmont (Doc Data sheet only)

Capability Development Division

Director General Maritime Development (Doc Data Sheet only)

Director General Land Development (Doc Data Sheet only)

Director General C3I Development (Doc Data Sheet only)

Intelligence Program

DGSTA Defence Intelligence Organisation   16

Corporate Support Program (libraries)

OIC TRS, Defence Regional Library, Canberra   17
Officer in Charge, Document Exchange Centre
    (Doc Data Sheet and distribution list only)
*US Defence Technical Information Centre, 2 copies   18,19
*UK Defence Research Information Center, 2 copies   20,21
*Canada Defence Scientific Information Service, 1 copy   22
*NZ Defence Information Centre, 1 copy   23
National Library of Australia, 1 copy   24

UNIVERSITIES AND COLLEGES

Australian Defence Force Academy
    Library   25
    Head of Aerospace and Mechanical Engineering   26
Deakin University, Serials Section (M list), Deakin University Library   27
Senior Librarian, Hargrave Library, Monash University   28
Librarian, Flinders University   29

OTHER ORGANISATIONS

NASA (Canberra)   30
AGPS   31

OUTSIDE AUSTRALIA

ABSTRACTING AND INFORMATION ORGANISATIONS

INSPEC: Acquisitions Section Institution of Electrical Engineers   32
Library, Chemical Abstracts Reference Service   33
Engineering Societies Library, US   34
Materials Information, Cambridge Scientific Abstracts, US   35
Documents Librarian, The Center for Research Libraries, US   36

INFORMATION EXCHANGE AGREEMENT PARTNERS

Acquisitions Unit, Science Reference and Information Service, UK   37
Library - Exchange Desk, National Institute of Standards and Technology   38

SPARES (5 copies)

Total number of copies: **43**

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION DOCUMENT CONTROL DATA	1. CAVEAT/PRIVACY MARKING

2. TITLE Risk-Sensitive Filtering and Parameter Estimation	3. SECURITY CLASSIFICATION Document (U) Title (U) Abstract (U)

4. AUTHOR(S) Jason Ford	5. CORPORATE AUTHOR Aeronautical and Maritime Research Laboratory PO Box 4331, Melbourne, Victoria, Australia 3001

6a. DSTO NUMBER DSTO-TR-0764	6b. AR NUMBER AR-010-820	6c. TYPE OF REPORT Technical Report	7. DOCUMENT DATE January, 1999

8. FILE NUMBER 9505-15-139	9. TASK NUMBER DST 96/095	10. SPONSOR DSTO	11. No OF PAGES 13	12. No OF REFS 24

13. DOWNGRADING / DELIMITING INSTRUCTIONS Not Applicable	14. RELEASE AUTHORITY Chief, Weapons Systems Division

15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT

Approved For Public Release

OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE CENTRE, DIS NETWORK OFFICE, DEPT OF DEFENCE, CAMPBELL PARK OFFICES, CANBERRA, ACT 2600

16. DELIBERATE ANNOUNCEMENT

No Limitations

17. CITATION IN OTHER DOCUMENTS

No Limitations

18. DEFTEST DESCRIPTORS

Filters   State estimation
Parameter estimation

19. ABSTRACT

This paper investigates the use of risk-sensitive filtering for state and parameter estimation in systems with model uncertainties. Modelling uncertainties arise from imperfectly known input process and noise characteristics as well as system model errors such as uncertain or time varying parameters of the system description. No new convergence results are given in this paper but simulation examples demonstrate that, in some situations, risk-sensitive filtering and estimation techniques allow for system uncertainties better than optimal techniques such as Kalman filtering.