Robust Multi-Sensor Fusion: A Decision-Theoretic Approach

MS-CIS-90-57
GRASP LAB 229

Gerda Kamberova
Max Mintz

Department of Computer and Information Science
School of Engineering and Applied Science
University of Pennsylvania
Philadelphia, PA 19104

August 1990

Robust Multi-Sensor Fusion: A Decision-Theoretic Approach

Gerda Kamberova and **Max Mintz***
GRASP Laboratory
Department of Computer and Information Science
University of Pennsylvania
Philadelphia, PA 19104-6389

Abstract

Many tasks in active perception require that we be able to combine different information from a variety of sensors that relate to one or more features of the environment. Prior to combining these data, we must test our observations for consistency. The purpose of this paper is to examine sensor fusion problems for linear location data models using statistical decision theory (SDT). The contribution of this paper is the application of SDT to obtain: (i) a robust test of the hypothesis that data from different sensors are consistent; and (ii) a robust procedure for combining the data that pass this preliminary consistency test. Here, robustness refers to the statistical effectiveness of the decision rules when the probability distributions of the observation noise and the a priori position information associated with the individual sensors are uncertain. The standard linear location data model refers to observations of the form: $Z = \theta + V$, where V represents additive sensor noise and θ denotes the "sensed" parameter of interest to the observer. While the theory addressed in this paper applies to many uncertainty classes, the primary focus of this paper is on asymmetric and/or multimodal models, that allow one to account for very general deviations from nominal sampling distributions. This paper extends earlier results in SDT and multi-sensor fusion obtained by [Zeytinoglu and Mintz, 1984], [Zeytinoglu and Mintz, 1988], and [McKendall and Mintz, 1988].

1 Introduction

Our research in active sensing is based on the theory and application of multiple sensors in the exploration of environments that are characterized by significant a priori uncertainties. In addition to uncertainty in the environment, the sensors themselves exhibit noisy behavior. While good engineering practice can reduce certain noise components, it is impractical if not impossible to eliminate them completely. Thus, all sensor measurements are uncertain. However, sensor errors can be modeled statistically, using both physical theory and empirical data. In developing these models, one recognizes that a single distribution is usually an inadequate description of sensor noise behavior. It is much more realistic and much safer to identify an envelope or class of distributions, one of whose members could represent the actual statistical behavior of the given sensor. This use of an uncertainty class (or equivalently: an envelope, set, or neighborhood) in distribution space, protects the system designer against the inevitable unpredictable changes that occur in sensor behavior. Reasons for uncertainty in statistical sensor models include: sporadic interference, drift due to aging, temperature variations, miscalibration, quantization, and other significant nonlinearities over the dynamic range of the sensor. The purpose of this paper is to examine a sensor fusion problem for linear location data models using statistical decision theory (SDT). The contribution of this paper is the application of SDT to obtain: (i) a robust test of the hypothesis that data from different sensors are consistent; and (ii) a robust procedure for combining the data that pass this preliminary consistency test. Here, robustness refers to the statistical effectiveness of the decision rules when the probability distributions of the observation noise and the a priori position information associated with the individual sensors are uncertain. The standard linear location data model refers to observations of the form: $Z = \theta + V$, where V represents additive sensor noise and θ denotes the "sensed" parameter of interest to the observer. The parameter θ is called a location parameter, since the distribution of Z is obtained from the distribution of V by a translation. While the location parameter fusion problem is only one of many possible fusion paradigms, it does provide a useful starting point for considering more complicated problems, e.g., nonlinear location sensor models of the form: $Z = h(\theta) + V$, where h denotes a given (nonlinear) function. It also provides a useful starting point for considering important generalizations of the location sensor model such as: $Z = h(\theta + V)$.

While the theory addressed in this paper applies to many uncertainty classes, the primary focus of this paper is on asymmetric and/or multimodal models, that allow one to account for very general deviations from

*Acknowledgement: Navy Contract N0014-88-K-0630; AFOSR Grants 88-0244, 88-0296; Army/DAAL 03-89-C-0031PRI; NSF Grants CISE/CDA 88-22719, IRI 89-06770; and the Dupont Corporation.

nominal sampling distributions. This paper extends earlier results in SDT and multi-sensor fusion obtained by [Zeytinoglu and Mintz, 1984], [Zeytinoglu and Mintz, 1988], and [McKendall and Mintz, 1988].

In the sequel we: (i) delineate several paradigms for robust fusion of multi-sensor linear location data; (ii) introduce some essential nomenclature and definitions from SDT; (iii) state the decision-theoretic results that this paper is based on; and (iv) present and discuss a methodology for robust fusion of multi-sensor linear location data.

Our presentation emphasizes the statement and application of the relevant theory. Proofs of theorems are omitted. The reader is referred to journal articles and reports for these details.

2 Paradigms for Sensor Fusion of Location Data

In this section we delineate several paradigms for robust fusion of location data. We restrict our attention to observations of one-dimensional location parameters. The results of this one-dimensional analysis can be applied to the multi-dimensional case by doing a component by component analysis. Alternatively, one can pursue a formal multi-dimensional extension of the methodology presented in this paper. This extension is part of our current research in sensor fusion.

The general one-dimensional paradigm is delineated as follows. We assume that we are given the sampled outputs of r sensor systems $\{\mathcal{S}_i : 1 \leq i \leq r\}$. We denote the k^{th} sampled output of $\mathcal{S}_i, 1 \leq k \leq N_i$ by:

$$Z_{ik} = \mu_i + W_i + \theta_i + V_{ik}, \tag{2.1}$$

where:

- $a_i \leq \theta_i \leq b_i$, denotes an unknown location parameter with known bounds a_i and b_i. [The bounds a_i and b_i may assume infinite values.] In many applications there is a common interval of location parameter uncertainty for all sensors. However, there is no need to make this assumption in the following mathematical developments.
- μ_i, denotes a known constant (offset) associated with the position of sensor $\mathcal{S}_i$ with respect to a common origin.
- V_{ik}, denotes the additive observation noise associated with the k^{th} observation (sample) from $\mathcal{S}_i$. The random variables $\{V_{ik} : 1 \leq k \leq N_i\}$ are assumed to be independent and identically distributed (i.i.d.). We further assume that the noise process associated with $\mathcal{S}_i$ is independent of the noise process associated with $\mathcal{S}_j$, when $i \neq j$. Finally, we assume that the probability distribution of V_{ik} belongs to a given uncertainty class of distributions, $\mathcal{F}_i$. We do *not* assume that the noise processes associated with different sensors are identically distributed.
- W_i, denotes the uncertainty in the position of sensor $\mathcal{S}_i$ with respect to a common origin. We consider two cases: (i) the position uncertainty of $\mathcal{S}_i$ can be expressed by a known interval $[l_i, u_i]$ — with no a priori probabilistic description; or (ii) the position uncertainty of $\mathcal{S}_i$ can be expressed by an unknown probability distribution from a given uncertainty class $\mathcal{P}_i$. In each case, we assume that the position uncertainty of $\mathcal{S}_i$ is independent of the observation noise $\{V_{ik} : 1 \leq k \leq N_i\}$, and independent of the observation noise and position uncertainty of the other sensors.

Remark 2.1 Without loss of generality, we can assume that the known offsets $\{\mu_i : 1 \leq i \leq r\}$ are each zero, since nonzero values can be subtracted from the observations $\{Z_{ik} : 1 \leq k \leq N_i\}$. Further, if the known, generally asymmetric, interval of uncertainty $[a_i, b_i]$ in θ_i is finite, then the observations $\{Z_{ik} : 1 \leq k \leq N_i\}$ can be shifted and the interval of uncertainty $[a_i, b_i]$ can be replaced by $[-d_i, d_i]$, where $d_i = (b_i - a_i)/2$. Similarly, we can assume the interval of sensor position uncertainty (where applicable) is also symmetric. Thus, (2.1) can be replaced by:

$$Z_{ik} = W_i + \theta_i + V_{ik}, \tag{2.2}$$

where: $1 \leq k \leq N_i$, $\mid \theta_i \mid \leq d_i$, and (where applicable) $\mid W_i \mid \leq \eta_i$, $1 \leq i \leq r$.

The uncertainty classes $\mathcal{F}_i$ and (where applicable) $\mathcal{P}_i$, $1 \leq i \leq r$, denote subsets in the space of probability distributions that are deemed to characterize the uncertainty in the specifications of the sampling distributions. Models for several uncertainty classes are described in Sections 4 and 6.

As stated in the introduction, the purpose of this paper is to examine a sensor fusion problem for location information using SDT. The contribution of this paper is the application of SDT to obtain: (i) a robust test of the hypothesis that data from different sensors are consistent, i.e., testing the hypothesis that $\theta_i = \theta_j$, $1 \leq i < j \leq r$; and (ii) a robust procedure for combining the data that pass this preliminary consistency test. Again, robustness refers to the statistical effectiveness of the decision rules when the probability distributions of the observation noise and the a priori position information of the individual sensors are uncertain.

In the following section, we introduce the notions of robust minimax decision rules and robust confidence procedures. These concepts provide the basis for the developments in the remainder of this paper.

3 Nomenclature and Definitions from SDT

The standard statement of a minimax location parameter estimation problem includes as given: a parameter space Ω; a space of actions $\mathcal{A}$; a loss function L defined on $\mathcal{A} \times \Omega$; and a CDF F. If the underlying CDF is imprecisely known, then this standard minimax decision model must be reformulated to account for this additional uncertainty. Statistical decision rules that are applicable in this more general problem setting are called robust procedures.

This paper considers robust fixed size confidence procedures for a restricted parameter space. These robust confidence procedures are based, in turn, on the solution of a related robust minimax decision problem:

Basic Minimax Decision Problem (MDP): Let $\mathbf{Z}$ denote a vector of N i.i.d. observations of a scalar random variable with CDF $F(z-\theta)$, where $F \in \mathcal{F}$, a given uncertainty class. Let $\Omega = \mathcal{A} = [-d, d]$, and define a zero-one loss function L on $\mathcal{A} \times \Omega$:

$$L(a, \theta) = \begin{cases} 0, & |a-\theta| \leq e; \\ 1, & |a-\theta| > e; \end{cases} \qquad (3.1)$$

where $e > 0$, is given. Further, let $R(\delta, \theta, F) = E[L(\delta, \theta) \mid \theta, F]$ denote the risk function of the decision rule δ given $\theta \in \Omega$ and $F \in \mathcal{F}$.

Definition 3.1 An estimator δ^* is said to be a robust minimax estimator for θ, if for all δ:

$$\sup_{\substack{\theta\in\Omega \\ F\in\mathcal{F}}} R(\delta^*, \theta, F) \leq \sup_{\substack{\theta\in\Omega \\ F\in\mathcal{F}}} R(\delta, \theta, F).$$

Based on these definitions and assumptions, we seek a robust minimax estimator δ^* for θ. For brevity, we restrict our consideration to the case when d/e is an integer ≥ 2.

Observation 3.1 The connection between the robust minimax rule $\delta^*(\mathbf{Z})$ and a robust fixed size confidence procedure is obtained by noting that:

$$C^*(\mathbf{Z}) = [\delta^*(\mathbf{Z}) - e, \delta^*(\mathbf{Z}) + e]$$

can be interpreted as a robust confidence procedure of size 2e that has the highest confidence coefficient $\inf_{\theta,F} P_{\theta,F}[\theta \in C^*(\mathbf{Z})]$.

Sections 4, 5, and 6 of this paper are organized as follows:

Section 4 presents solutions of two related single-sample minimax estimation problems where F is given. These results provide the basis for the solutions to the robust minimax estimation problems where $F \in \mathcal{F}$.

Section 5 extends the results of Section 4 to the multi-sample case.

Section 6 develops a theory and methodology for robust sensor fusion of location information based on the theory presented in Sections 4 and 5.

4 Minimax and Robust Minimax Rules

Throughout Section 4 we consider the single-sample decision problem MDP ($N = 1$).

4.1 Minimax Rules

Minimax problems are special cases of robust minimax problems in the sense that $\mathcal{F}$ contains a single CDF F. We begin with two minimax estimation problems that are defined by the zero-one loss function L (3.1). The solutions to these single-sample estimation problems provide the basis for solutions to both the single-sample and multi-sample robust minimax estimation problems. These preliminary results require Definitions 4.1-4.2 and are summarized by Theorems 4.1-4.2.

Definition 4.1 Let $\mathcal{C}_a$ denote the class of nonrandomized, monotone nondecreasing decision rules $\delta: E^1 \rightarrow \mathcal{A}$, where: $\mathcal{A} = [-d, d]$. Let $\Delta_a \subset \mathcal{C}_a$ denote the set of rules $\delta(t)$, defined for $t \in (-\infty, \infty)$ by (4.1), where: $i = 1, 2, \ldots, n$ and $-\infty < a_{-n} \leq \cdots \leq a_{-2} \leq a_{-1} \leq a_0 \leq a_1 \leq a_2 \leq \cdots \leq a_n < \infty$, $d = (2n+1)e + c$, and c equals zero (e) if d is an odd (even) multiple of e. Note that the parameter a_0 is relevant only when c equals e.

Observation 4.1 Let L denote the zero-one loss function (3.1). If the CDF F is continuous, then for each $\delta \in \Delta_a$, the risk function $R(\delta, \theta, F)$ is (4.2), where $i = 1, 2, \ldots, n-1$. For each $\delta \in \Delta_a$, $R(\delta, \theta, F)$ is a piecewise constant function of θ over the sets of a finite partition of Ω, and the maximum of $R(\delta, \theta, F)$ occurs at one or more of the nondegenerate intervals. The risk expression (4.2) can be readily modified to include CDF's F that are discontinuous. The generalized risk function $R(\delta, \theta, F)$ is again a piecewise constant function of θ over the sets of a finite partition of Ω expressed in (4.2).

Theorem 4.1 Let L denote the zero-one loss function (3.1), and T be a scalar random variable with given CDF $F(t-\theta)$. If F is absolutely continuous with respect to Lebesgue measure, has convex support, and possesses a (strictly) monotone likelihood ratio, then there exists a globally minimax (admissible) Bayes rule $\delta^* \in \Delta_a$ and a least favorable prior distribution λ^*.

Proof: See [Kamberova and Mintz, 1990].

Remark 4.1 $R(\delta^*, \theta, F)$ and λ^* have the following characteristics:

- The minimax rule δ^* is an "almost" equalizer rule, in the sense that the nondegenerate piecewise constant segments of the risk function are equalized to the minimax risk by a suitable choice of the parameter vector $\mathbf{a} = (a_{-n}, \ldots, a_n)^T$.
- The least favorable prior distribution λ^* is defined by a density function that is piecewise constant.

Remark 4.2 Theorem 4.1 extends the basic minimax results of [Zeytinoglu and Mintz, 1984] by allowing the inclusion of CDF's F that are asymmetric.

Definition 4.2 A rule is (robust) $\mathcal{D}$-minimax if it is (robust) minimax within the class $\mathcal{D}$. A rule is $\mathcal{D}$-Bayes if it is Bayes within the class $\mathcal{D}$. A rule is $\mathcal{D}$-admissible if it is admissible within the class $\mathcal{D}$.

In the following theorem we weaken the hypothesis of Theorem 4.1 by dropping the monotone likelihood ratio condition, and obtain a $\mathcal{C}_a$-minimax result.

Theorem 4.2 Let L denote the zero-one loss function (3.1), and T be a scalar random variable with given CDF $F(t-\theta)$. If F is absolutely continuous with respect to Lebesgue measure and has convex support, then there exists a $\mathcal{C}_a$-minimax rule $\delta^* \in \Delta_a$.

Proof: See [Kamberova and Mintz, 1990].

Remark 4.3 $R(\delta^*, \theta, F)$ has the following characteristic:

- The $\mathcal{C}_a$-minimax rule δ^* is an "almost" equalizer rule in the sense of Remark 4.1.

Remark 4.4 Theorem 4.2 extends the basic $\mathcal{C}$-minimax results of [Zeytinoglu and Mintz, 1984] by allowing the inclusion of CDF's F that are asymmetric and/or multimodal.

4.2 Robust Minimax Rules

In this section we define two uncertainty classes $\mathcal{F}$, and delineate the solutions to the corresponding robust minimax and robust $\mathcal{C}_a$-minimax estimation problems. These results require Definitions 4.3-4.4 and are summarized by Theorems 4.3-4.6.

$$\delta(t) = \begin{cases} d-e, & c+a_n+2ne \le t ; \\ t-a_i, & c+a_i+2(i-1)e \le t < c+a_i+2ie; \\ 2(i-1)e+c, & c+2(i-1)e+a_{i-1} \le t < c+a_i+2(i-1)e; \\ t-a_0, & -c+a_0 < t < c+a_0; \\ -2(i-1)e-c, & -c+a_{-i}-2(i-1)e < t \le -c+a_{-i+1}-2(i-1)e; \\ t-a_{-i}, & -c+a_{-i}-2ie < t \le -c+a_{-i}-2(i-1)e; \\ -d+e, & t \le -c+a_{-n}-2ne; \end{cases} \tag{4.1}$$

$$R(\delta,\,\theta,\,F) = \begin{cases} F(a_n-e), & d-2e < \theta \le d; \\ F(a_{n-1}-e), & \theta = d-2e; \\ F(a_i-e)+1-F(a_{i+1}+e), & c+(2i-1)e < \theta \le c+(2i+1)e; \\ F(a_{-1+c/e}-e)+1-F(a_1+e), & \theta = c+e; \\ F(a_0-e)+1-F(a_1+e), & -c+e < \theta < c+e; \\ F(a_{-1}-e)+1-F(a_1+e), & \theta = -c+e; \\ F(a_{-1}-e)+1-F(a_1+e), & c-e < \theta < -c+e; \\ F(a_{-2+c/e}-e)+1-F(a_{-1+c/e}+e), & \theta = c-e; \\ F(a_{-1}-e)+1-F(a_0+e), & -c-e < \theta < c-e; \\ F(a_{-2}-e)+1-F(a_{-1}+e), & \theta = -c-e; \\ F(a_{-(i+1)}-e)+1-F(a_{-i}+e), & -c-(2i+1)e \le \theta < -c-(2i-1)e; \\ 1-F(a_{-n+1}+e), & \theta = -d+2e; \\ 1-F(a_{-n}+e), & -d \le \theta < -d+2e; \end{cases} \tag{4.2}$$

Definition 4.3 Let $\mathcal{F}$ denote an uncertainty class with upper-envelope F_u:

$$\mathcal{F} = \{F\colon F(x^-) \le F_u(x),\ x \le s;\ F(x) \ge F_u(x),\ x > s\}, \tag{4.3}$$

where F_u is absolutely continuous with respect to Lebesgue measure and has convex support.

Remark 4.5 The CDF F_u defines the upper-envelope of $\mathcal{F}$ (4.3) in the sense that: $F(x) \le F_u(x)$ for all $F \in \mathcal{F}$, and $x < s$. The upper-envelope CDF F_u is permitted to be substochastic, i.e., F_u can have less than unit probability mass. Thus, all ϵ-contamination models can be represented by a simple generalization of $\mathcal{F}$ (4.3).

The following theorem extends the results of Theorem 4.1 to the single-sample robust minimax estimation problem.

Theorem 4.3 Let $\mathcal{F}$ denote the uncertainty class (4.3) with upper-envelope F_u. Assume F_u possesses a (strictly) monotone likelihood ratio. Let δ^* denote the minimax rule obtained through Theorem 4.1 based on CDF F_u. There exists a bound $B(d/e,\, F_u)$, such that if $e \ge B$, then δ^* is a robust minimax (admissible) Bayes rule.

Proof: See [Kamberova and Mintz, 1990].

The following theorem extends the results of Theorem 4.2 to the single-sample robust $\mathcal{C}_a$-minimax estimation problem.

Theorem 4.4 Let $\mathcal{F}$ denote the uncertainty class (4.3) with upper-envelope F_u. Let δ^* denote the $\mathcal{C}_a$-minimax rule obtained through Theorem 4.2 based on CDF F_u. There exists a bound $B(d/e,\, F_u)$, such that if $e \ge B$, then δ^* is a robust $\mathcal{C}_a$-minimax rule.

Proof: See [Kamberova and Mintz, 1990].

Definition 4.4 Let $\mathcal{F}$ denote the uncertainty class:

$$\mathcal{F} = \{F(\cdot) = F_0((\cdot - \tau)/\sigma)\colon |\,\tau\,| \le \eta;\ \ \sigma \le \sigma_u\}, \tag{4.4}$$

where: $\eta > 0$ and $\sigma_u > 0$ denote given bounds, and F_0 denotes a given CDF that is symmetric about zero, and absolutely continuous with respect to Lebesgue measure.

Remark 4.6 The uncertainty class $\mathcal{F}$ (4.4) models underlying uncertainty in both location and scale for a symmetric distribution F_0. Without loss of generality, we can assume $\sigma_u = 1$.

Remark 4.7 The delineation of robust minimax rules and robust $\mathcal{C}_a$-minimax rules for the estimation problem defined by the zero-one loss function L (3.1), and the uncertainty class $\mathcal{F}$ (4.4) is obtained by determining the joint worst-case behavior of the parameters: θ, τ, and σ. By worst case, we mean those combinations of parameter values that lead to maximum risk. In carrying out this worst-case analysis, it is necessary to consider two cases: d/e is odd, and d/e is even. For brevity, we restrict our analysis to the even case. The complete analysis appears in [Kamberova and Mintz, 1990].

Observation 4.2 Let $d = (2n+2)e$, $n \ge 0$. There exist bounds $B_1(d/e, \sigma_u, F_0)$ and $B_2(d/e, \sigma_u, F_0)$ such that if $\eta \le B_1$ and $e \ge B_2$, then the joint worst-case behavior of θ, τ, and σ is: $\tau = -\eta$ when $\theta > 0$; $\tau = \eta$ when $\theta < 0$; and $\sigma = \sigma_u$ for all θ.

Observation 4.3 As a consequence of the underlying even and odd symmetry in this decision problem, which is reflected by the worst-case analysis, we can restrict our attention to rules $\delta \in \Delta_a$ that possess odd symmetry about zero ($a_0 = 0$ and $a_{-i} = -a_i$). We denote this subset of Δ_a by Δ.

Observation 4.4 If the relation between the parameters θ, τ, and σ is defined by the worst-case analysis of Observation 4.2, then for any $\delta \in \Delta$, the worst-case risk (for $\theta > 0$) is (4.5), where: $\sigma_u = 1$, and $d = (2n+2)e$, $n \ge 0$. We can restrict our attention to the domain $\theta > 0$ due to the even and odd symmetry in this decision problem.

$$R(\delta, \theta, F_0) = \begin{cases} F_0(a_n + \eta - e), & d - 2e < \theta \leq d; \\ F_0(a_{n-1} + \eta - e), & \theta = d - 2e; \\ F_0(-a_n - \eta - e) + F_0(a_{n-1} + \eta - e), & d - 4e < \theta < d - 2e; \\ \cdot & \\ \cdot & \\ \cdot & \\ F_0(-a_2 - \eta - e) + F_0(a_1 + \eta - e), & 2e < \theta < 4e; \\ F_0(-a_2 - \eta - e) + F_0(\eta - e), & \theta = 2e; \\ F_0(-a_1 - \eta - e) + F_0(\eta - e), & 0 < \theta < 2e; \end{cases} \tag{4.5}$$

Lemma 4.1 If F_0 is absolutely continuous with respect to Lebesgue measure and has convex support, then there exists a choice of parameters $\{a_i : 1 \leq i \leq n\}$ that equalize the nondegenerate piecewise constant segments of the risk function (4.5). The corresponding rule δ^* is an "almost" equalizer rule.

The following theorem delineates the existence and structure for single-sample robust minimax rules in the case of the joint location-scale uncertainty class $\mathcal{F}$ (4.4).

Theorem 4.5 Let $\mathcal{F}$ denote the location-scale uncertainty class (4.4) based on the symmetric CDF F_0. Assume F_0 possesses a (strictly) monotone likelihood ratio and has convex support. Let δ^* denote the rule obtained through Lemma 4.1. There exists bounds $B_1(d/e, \sigma_u, F_0)$, and $B_2(d/e, \sigma_u, F_0)$ such that if $\eta \leq B_1$, and $e \geq B_2$, then δ^* is a robust minimax (admissible) Bayes rule.

Proof: See [Kamberova and Mintz, 1990].

In the following theorem we weaken the hypothesis of Theorem 4.5 by dropping the monotone likelihood ratio condition, and obtain a robust $\mathcal{C}_a$-minimax result.

Theorem 4.6 Let $\mathcal{F}$ denote the location-scale uncertainty class (4.4) based on the CDF F_0. Assume F_0 has convex support. Let δ^* denote the rule obtained through Lemma 4.1. There exists a bound $B(d/e, \sigma_u, F_0)$ such that if $e \geq B$, then δ^* is a robust $\mathcal{C}_a$-minimax rule.

Proof: See [Kamberova and Mintz, 1990].

5 The Multi-Sample Case

This section extends the robust minimax results of Theorems 4.3-4.6 to the multi-sample problem ($N > 1$) by restricting the class of estimators to rules of the form $\delta(T(\mathbf{Z}))$, where: $\delta \in \mathcal{C}_a$, T is a real-valued function of $\mathbf{Z}$, and $T(\mathbf{Z})$ possesses a CDF that depends on θ as a location parameter, is absolutely continuous with respect to Lebesgue measure, and has convex support. Examples of candidate T statistics include: the sample mean, the sample median, and other linear combinations of order statistics. In the remainder of this section we consider the sample median.

Definition 5.1 Let Z_M denote the median of the N observations $\mathbf{Z}$. [If N is even, $Z_M = (Z_{[N/2]} + Z_{[(N/2)+1]})/2$.] The decision rule $\delta^*(Z_M)$, defined by the composition $\delta^* \circ Z_M$, is said to be a median-minimax estimator for θ, if δ^* is a minimax rule in the usual sense. The respective definitions of robust median-minimax rules, $\mathcal{C}_a$-median-minimax rules, and robust $\mathcal{C}_a$-median-minimax rules are obtained as before.

The median statistic $T(\mathbf{Z}) = Z_M$ possesses several properties that are used in obtaining Theorems 5.1-5.4. These properties are stated in Observations 5.1-5.2.

Observation 5.1 The centered median statistic $Z_M - \theta$ preserves the upper-envelope of the uncertainty class $\mathcal{F}$ (4.3). Further, the CDF of $Z_M - \theta$ preserves absolute continuity with respect to Lebesgue measure and convex support.

Observation 5.2 The median statistic Z_M preserves location ordering for fixed scale, and scale ordering for fixed location in the uncertainty class $\mathcal{F}$ (4.4). Further, the CDF of Z_M preserves absolute continuity with respect to Lebesgue measure and convex support.

The following theorem extends the results of Theorem 4.3 to the multi-sample robust minimax estimation problem.

Theorem 5.1 Let $N > 1$ and $\mathcal{F}$ denote the uncertainty class (4.3) with upper-envelope F_u. Let F_{uM} denote the CDF of the centered sample median $Z_M - \theta$, where the underlying common CDF is F_u. Assume F_{uM} possesses a (strictly) monotone likelihood ratio. Let δ^* denote the minimax rule obtained through Theorem 4.1 based on CDF F_{uM}. There exists a bound $B(d/e, N, F_u)$, such that if $e \geq B$, then δ^* is a robust median-minimax (median-admissible) median-Bayes rule.

Proof: See [Kamberova and Mintz, 1990].

The following theorem extends the results of Theorem 4.4 to the multi-sample robust $\mathcal{C}_a$-minimax estimation problem.

Theorem 5.2 Let $N > 1$ and $\mathcal{F}$ denote the uncertainty class (4.3) with upper-envelope F_u. Let F_{uM} denote the CDF of the centered sample median $Z_M - \theta$, where the underlying common CDF is F_u. Let δ^* denote the $\mathcal{C}_a$-minimax rule obtained through Theorem 4.2 based on CDF F_{uM}. There exists a bound $B(d/e, N, F_u)$, such that if $e \geq B$, then δ^* is a robust $\mathcal{C}_a$-median-minimax rule.

Proof: See [Kamberova and Mintz, 1990].

The following theorem extends the results of Theorem 4.5 to the multi-sample robust minimax estimation problem.

Theorem 5.3 Let $N > 1$ and $\mathcal{F}$ denote the location-scale uncertainty class (4.4) based on the symmetric CDF F_0. Assume F_0 has convex support. Let F_{0M} denote the CDF of the sample median, where the underlying common CDF is F_0. Assume F_{0M} possesses a (strictly) monotone likelihood ratio. Let δ^* denote the rule obtained through Lemma 4.1 based on the CDF F_{0M}. There exists bounds $B_1(d/e, N, \sigma_u, F_0)$, and $B_2(d/e, N, \sigma_u, F_0)$ such that if $\eta \leq B_1$, and $e \geq B_2$, then δ^* is a robust median-minimax (median admissible) median-Bayes rule.

Proof: See [Kamberova and Mintz, 1990].

The following theorem extends the results of Theorem 4.6 to the multi-sample robust $\mathcal{C}_a$-minimax estimation problem.

Theorem 5.4 Let $N > 1$ and $\mathcal{F}$ denote the location-scale uncertainty class (4.4) based on the symmetric CDF F_0. Assume F_0 has convex support. Let F_{0M} denote the CDF of the sample median, where the underlying common CDF is F_0. Let δ^* denote the rule obtained through Lemma 4.1 based on the CDF F_{0M}. There exists a bound $B(d/e, N, \sigma_u, F_0)$ such that if $e \geq B$, then δ^* is a robust $\mathcal{C}_a$-median-minimax rule.

Proof: See [Kamberova and Mintz, 1990].

6 Robust Fusion of Location Information

6.1 Preliminary Remarks

In this section we develop a theory and methodology for robust fusion of multi-sensor location information based on Sections 4 and 5. Our approach contains two distinct phases:

- **Phase I** provides a test of the hypothesis $\theta_i = \theta_j$, that the location data (2.2) from sensor $\mathcal{S}_i$ are consistent with the location data from sensor $\mathcal{S}_j$, where $i < j$.
- **Phase II** provides a means of combining the location data from the individual data sets that "pass" the Phase I test, i.e., those deemed to be consistent.

In both phases of this process, we seek procedures that are robust to heavy-tailed deviations from the nominal sampling distribution, such as exhibited in ϵ-contamination uncertainty classes. Our usage of "robust" is also intended to imply that the procedures have satisfactory behavior when the actual sampling distribution coincides with the nominal, e.g., a given Gaussian distribution.

6.2 Sample Sizes and Uncertainty Classes

In developing suitable consistency tests, there are three domains of sample sizes to address: (i) the single sample case, $N = 1$; (ii) the small sample case, $1 < N \leq 20$; and (iii) the large sample case, $N > 20$. In defining these classes, it is important to observe that the transition ($N = 20$) between the small sample and large sample cases is not a precise threshold value — the appropriate selection of this threshold is dependent on the uncertainty classes that define the given decision problem. The sample size for each sensor $\mathcal{S}_i$ is denoted by $N_i, 1 \leq i \leq r$. The sample sizes N_i and N_j can belong to different sample size domains.

The selection of appropriate sensor noise uncertainty classes $\{\mathcal{F}_i : 1 \leq i \leq r\}$ is an important issue in the development of a methodology for robust fusion of multi-sensor location information. Since, at the minimum, we seek to account for the occurrence of noise distributions with heavy tails, it is appropriate to consider both ϵ-contamination uncertainty classes as well as joint location-scale uncertainty classes. We consider two cases:

Case 1: We adopt an ϵ-contamination model $\mathcal{F}_{\epsilon_i}$ for each sensor $\mathcal{S}_i$, $1 \leq i \leq r$; in particular, the ϵ_i-contaminated non-Gaussian model for sensor $\mathcal{S}_i$ that is defined by:

$$\mathcal{F}_{\epsilon_i} = \{F : F = (1 - \epsilon_i)\Psi + \epsilon_i H\}, \tag{6.1}$$

where: (i) Ψ denotes a given asymmetric, (possibly) multi-modal CDF that is absolutely continuous with respect to Lebesgue measure, and has convex support, and (ii) the CDF H is arbitrary, and $0 < \epsilon_i < 1$. This uncertainty class is a simple generalization of the uncertainty class (4.3).

Case 2: We adopt a joint location-scale uncertainty class for each sensor $\mathcal{S}_i$, $1 \leq i \leq r$; in particular, the joint location-scale uncertainty class defined by (4.4), where F_0 is the $N(0,1)$ CDF, and the location-scale bounds are η_i and σ_{u_i}.

6.3 Phase I — Robust Consistency Tests

Analysis of Case 1: The following procedure provides a robust test of the hypothesis that $\theta_i = \theta_j$, $i < j$.

Let $\mathcal{M}_i$ denote the class of CDF's defined by the centered sample median Z_{M_i} of N_i i.i.d. samples with CDF $F \in \mathcal{F}_{\epsilon_i}$ (6.1), $1 \leq i \leq r$. Let $\mathcal{M}_{ij}$ denote the class of CDF's defined by the difference of the centered sample medians $(Z_{M_i} - \theta_i) - (Z_{M_j} - \theta_j)$, where the CDF's of the centered sample medians $(Z_{M_i} - \theta_i)$ and $(Z_{M_j} - \theta_j)$ belong, respectively, to $\mathcal{M}_i$ and $\mathcal{M}_j$, $1 \leq i < j \leq r$. It follows from these definitions that the class $\mathcal{M}_{ij}$ is a set of distributions of the form (4.3). Further,

$$Z_{M_i} - Z_{M_j} = \theta_i - \theta_j + \nu_{ij}, \tag{6.2}$$

where: the CDF of ν_{ij} belongs to $\mathcal{M}_{ij}$; and the a priori uncertainty in $\theta_i - \theta_j$ is given by the interval $[-d_{ij}, d_{ij}]$, where $d_{ij} = d_i + d_j$.

Hence, we can construct a robust fixed size ($2e$) confidence procedure for $\theta_i - \theta_j$. The parameter e is selected by the decision maker: (i) it defines the decision maker's tolerance to small errors between θ_i and θ_j; and (ii) it is used to select the size of the statistical test. The desired procedure $[\delta^* - e, \ \delta^* + e]$ is obtained via Theorem 5.2. Finally, the test of the hypothesis $\theta_i = \theta_j$ is obtained as follows: we reject $\theta_i = \theta_j$ if $0 \notin [\delta^* - e, \ \delta^* + e]$. From this test we also obtain the minimum probability that $\theta_i - \theta_j \in [\delta^* - e, \ \delta^* + e]$. Examples of applications of this class of robust consistency tests appears in [Kamberova *et al.*, 1990].

Analysis of Case 2: We follow the basic approach described in the analysis of case 1, but we replace the sample median statistics by the sample means. Here, the sample mean is useful, since the underlying uncertainty classes contain only Gaussian distributions. The robust consistency test is obtained via Theorem 5.3. The details appear in [Kamberova *et al.*, 1990].

6.4 Phase II — Robust Fusion of Consistent Multi-Sensor Location Information

The following procedure provides a robust estimate of the common location parameter θ of r sensor data sets, $r \geq 3$. We observe at the outset that, when V_1 and V_2 possess very heavy tails, in general, it is not useful to attempt to combine two observations of the form:

$$Z_1 = \theta + V_1$$

$$Z_2 = \theta + V_2$$

by convex combination. For example, if V_1 and V_2 are independent Cauchy $C(0,1)$ random variables, then any convex combination of Z_1 and Z_2 will be a $C(\theta, 1)$ random variable. Further, there are random variables with continuous unimodal symmetric density functions whose sample mean, for any sample size $N > 1$, has greater variability then any of its N i.i.d. components.

Analysis of Case 1: Let $\{Z_{M_i} : 1 \leq i \leq r\}$ denote the sample medians of r consistent data sets with common location parameter θ. To simplify the exposition, we further assume that the r sample medians are identically distributed. Let Z_{M_A} denote the median of the $\{Z_{M_i} : 1 \leq i \leq r\}$. Let $\mathcal{M}_A$ denote the uncertainty class of the centered sample median $Z_{M_A} - \theta$. The uncertainty class $\mathcal{M}_A$ is of the form (4.3). Thus, we can apply Theorem 5.2 to obtain a robust fixed size confidence procedure $[\delta^* - e,\ \delta^* + e]$ for θ. Examples of applications of this class of confidence procedures for the robust fusion of consistent multi-sensor location information appears in [Kamberova *et al.*, 1990].

Analysis of Case 2: We follow the basic approach described in the analysis of case 1, but we replace the sample median statistics by the sample means. Here, the sample mean is useful, since the underlying uncertainty classes contain only Gaussian distributions. A robust estimate of location is obtained via Theorem 5.3. The details appear in [Kamberova *et al.*, 1990].

References

[Kamberova and Mintz, 1990] Gerda Kamberova and Max Mintz. *Decision Theory for Robust Multi-Sensor Fusion*. Technical Report GRASP Lab, University of Pennsylvania, Philadelphia, PA, 1990.

[Kamberova *et al.*, 1990] Gerda Kamberova, Raymond McKendall, and Max Mintz. *Robust Multi-Sensor Fusion*. Technical Report GRASP Lab, University of Pennsylvania, Philadelphia, PA, 1990.

[McKendall and Mintz, 1988] Raymond McKendall and Max Mintz. Robust fusion of location information. In *Proceedings of the 1988 IEEE International Conference on Robotics and Automation*, pages 1239–1244, IEEE, Philadelphia, April 1988.

[Zeytinoglu and Mintz, 1984] Mehmet Zeytinoglu and Max Mintz. Optimal fixed size confidence procedures for a restricted parameter space. *Ann. Statist.*, 12(3):945–957, September 1984.

[Zeytinoglu and Mintz, 1988] Mehmet Zeytinoglu and Max Mintz. Robust fixed size confidence procedures for a restricted parameter space. *Ann. Statist.*, 16(3):1241–1253, September 1988.