Fusion under Unknown Correlation - Covariance Intersection as a Special Case

Lingji Chen, Pablo O. Arambel and Raman K. Mehra
Scientific Systems Company Inc.
500 West Cummings Park, Suite 3000
Woburn, MA 01801
chen,pablo,rkm@ssci.com

Abstract – *This paper addresses the problem of fusing several random variables (RV's) with unknown correlations. A family of upper bounds on the resulting covariance matrix is given, and is shown to contain the upper bound offered by the Covariance Intersection (CI) Algorithm proposed by Julier and Uhlmann. For trace minimization, the optimal one in this family is better than the one obtained by CI except in some cases where they are equal. It is further proved that the best pair of combination gains that minimizes the above optimal-trace-in-the-family coincides with the one associated with the best value of omega in CI. Thus the CI Algorithm provides a convenient one-dimensional parameterization for the optimal solution in the n-square dimensional space. The results are also extended to multiple RV's and partial estimates.*

Keywords: Fusion of random variables, unknown correlation, Covariance Intersection, consistent estimation, Kalman Filter

1 Introduction

The need of fusing several random variables (RV's) into one RV often arises in fusion and estimation problems. Updating the current estimate with a new observation is one such example. Usually the target RV is defined as a linear combination of the source RV's. In the case of two uncorrelated RV's, for any given pair of combination gains, the covariance matrix of the target RV is readily available, and the pair that minimizes the trace of this covariance matrix (and thus the estimation is optimal in the least square sense) is given by the "Kalman Gain" as defined in a Kalman Filter. When the correlation is unknown, assuming it to be zero may still present satisfactory results for many applications. However, there are also many situations in which the assumption of independence may lead to serious problems. For example, in a distributed network, when a node A receives a piece of information from a node B, the topology of the network may be such that B is passing along the information it originally received from A. Thus, if A were to fuse the "new" information from B with the "old" one it has using the Kalman Gain under the assumption of independence, then the covariance matrix (as an indicator of the uncertainty about this information) would be reduced, when in fact it should remain at the same level.

In their seminal papers [1, 2], the Covariance Intersection (CI) Algorithm was proposed to deal with this problem. The objective is to obtain a *consistent* estimate of the covariance matrix when two random variables are linearly combined. By "consistent" we mean that the estimated covariance is always an "upper-bound" (in the positive definite sense; see Section 2) of the true covariance, even when the correlation among the original RV's is unknown. Thus in the example above, after node A combines the information, the covariance matrix will remain approximately the same, rather than incorrectly reduced. Judiciously combined with Kalman Filter and prior knowledge about the systems, the CI Algorithm has found wide applications [3, 4, 5, 6, 7, 8, 9, 10].

Nevertheless, there are questions that are not answered by the CI Algorithm. When the variables being combined are n dimensional vectors, the combination gain is a matrix with n^2 elements, thus can be chosen from an n^2 dimensional space. But the CI Algorithm parameterizes only a one dimensional curve in this space. In order to get a complete picture, we pose two separate problems in this paper. The first problem is to obtain a consistent estimate of the covariance matrix when *fixed* combination gains are used. We solve this problem by presenting a family of such estimates. The member in this family having the minimal trace can be determined analytically, and is referred to as the "family-optimal" estimate. The second problem is to find the *best* pair of gains that minimizes the trace of the above family-optimal estimate. In general this is an optimization problem in an n^2 dimensional space of combination gains. However, we prove that the global optimal solution is actually given by the CI Algorithm, even though it conducts the search only along a one dimensional curve.

The paper is organized as follows. First a statement of the problem is given. Following this, the CI Algorithm is reviewed. Section 4 presents the main results of this paper. The results are generalized to partial estimates and multiple variables in Section 5. Finally some conclusions are drawn. Part of the results presented in this paper will also appear in [11].

2 Problem Statement

To highlight the essence of the results, no dynamics is considered in this paper, and the problem is simply stated as combining two estimates of the mean value of a random variable when the correlation between the estimates is unknown. The basic notations in [1] are followed here, but for simplicity no distinction is made between a random variable and its observation. More specifically, let c^* be the mean value of some random variable to be estimated. Two sources of information are available: estimate a and estimate b. Define their estimation errors as

$$\tilde{a} = a - c^*, \ \tilde{b} = b - c^*$$

and assume that

$$\mathrm{E}\{\tilde{a}\} = 0, \ \mathrm{E}\{\tilde{a}\tilde{a}^T\} = \tilde{P}_{aa}$$
$$\mathrm{E}\{\tilde{b}\} = 0, \ \mathrm{E}\{\tilde{b}\tilde{b}^T\} = \tilde{P}_{bb}$$

The true values of $\tilde{P}_{aa}$ and $\tilde{P}_{bb}$ may not be known, but consistent estimates are known:

$$P_{aa} \geq \tilde{P}_{aa}, \ P_{bb} \geq \tilde{P}_{bb} \tag{1}$$

Here inequality is in the sense of matrix positive definiteness, *i.e.*, $A > B$ if and only if $A - B$ is positive definite. The correlation between the two estimates

$$\mathrm{E}\{\tilde{a}\tilde{b}^T\} = \tilde{P}_{ab}$$

is also unknown.

Our objective is to construct a *linear, unbiased* estimator c that combines a and b:

$$c = K_1 a + K_2 b \tag{2}$$

Define

$$\tilde{c} = c - c^*$$

It follows that

$$\mathrm{E}\{\tilde{c}\} = 0$$

if and only if

$$K_1 + K_2 = I \tag{3}$$

The covariance

$$\mathrm{E}\{\tilde{c}\tilde{c}^T\} = \tilde{P}_{cc}$$

may not be known, but we want to find its consistent estimate P_{cc}:

$$P_{cc} \geq \tilde{P}_{cc} \tag{4}$$

Note that

$$\tilde{P}_{cc} = K_1\tilde{P}_{aa}K_1^T + K_2\tilde{P}_{bb}K_2^T + K_1\tilde{P}_{ab}K_2^T + K_2\tilde{P}_{ba}K_1^T \tag{5}$$

We formulate the following two problems:

> **Problem 1:** Determine a consistent estimate (upper bound) P_{cc} for $\tilde{P}_{cc}$ in (5) for any *given* pair of K_1 and K_2.
>
> **Problem 2:** Find the pair of K_1 and K_2 such that the upper bound P_{cc} is optimal in some sense, *e.g.*, minimal trace or determinant.

If

$$\tilde{P}_{ab} = 0$$

then for any given K_1 and K_2, the estimate

$$P_{cc} = K_1 P_{aa} K_1^T + K_2 P_{bb} K_2^T$$

will be consistent (*i.e.*, $P_{cc} \geq \tilde{P}_{cc}$) as a direct consequence of (1). The trace of the above P_{cc} is minimized by

$$P_{cc} = (P_{aa}^{-1} + P_{bb}^{-1})^{-1}$$
$$K_1 = P_{cc}P_{aa}^{-1} = P_{bb}(P_{aa} + P_{bb})^{-1}$$
$$K_2 = P_{cc}P_{bb}^{-1} = P_{aa}(P_{aa} + P_{bb})^{-1}$$

This corresponds to the derivation of the Kalman Gain in Kalman Filter.

3 The Covariance Intersection Algorithm

If $\tilde{P}_{ab} \neq 0$ but is *known*, then P_{cc} can be given by

$$P_{cc} = [K_1, K_2] \begin{bmatrix} P_{aa} & \tilde{P}_{ab} \\ \tilde{P}_{ab}^T & P_{bb} \end{bmatrix} \begin{bmatrix} K_1^T \\ K_2^T \end{bmatrix}$$

The best choice of K_1 and K_2 that minimizes the trace of P_{cc} can be obtained by solving the following constrained optimization problem:

$$\min_K \mathrm{tr}\{KPK^T\} \qquad \text{subject to} \qquad K \begin{bmatrix} I \\ I \end{bmatrix} = I$$

where

$$K \triangleq [K_1, K_2], \ P \triangleq \begin{bmatrix} P_{aa} & \tilde{P}_{ab} \\ \tilde{P}_{ab}^T & P_{bb} \end{bmatrix} \tag{6}$$

The optimal solution of K_1 and K_2 yields a P_{cc} in the following form

$$P_{cc}^{-1} = [I \quad I]P^{-1}\begin{bmatrix} I \\ I \end{bmatrix}$$
$$= P_{aa}^{-1} + (P_{aa}^{-1}\tilde{P}_{ab} - I)(P_{bb} - \tilde{P}_{ab}^T P_{aa}^{-1}\tilde{P}_{ab})^{-1} \cdot (\tilde{P}_{ab}^T P_{aa}^{-1} - I) \tag{7}$$

Covariance Ellipses, as defined below, is a convenient way of visualizing the relative size of covariance matrices. For a positive definite matrix Q, we define

$$\mathcal{B}_Q(l) \triangleq \{x : x^T Q^{-1} x < l\} \tag{8}$$

A Covariance Ellipse at level l is the boundary of $\mathcal{B}_Q(l)$. (We will omit "(l)" in the following discussions.) Thus, if $Q_1 < Q_2$, then $\mathcal{B}_{Q_1} \subset \mathcal{B}_{Q_2}$.

Now we show that

1. For a *given* $\tilde{P}_{ab}$, and hence the optimal P_{cc} in (7), we have
$$\mathcal{B}_{P_{cc}} \quad \subset \quad \mathcal{B}_{P_{aa}} \cap \mathcal{B}_{P_{bb}}$$
Proof: Since P in (6) is positive definite, by definition the second term in (7) is positive definite. Thus $x^T P_{cc}^{-1} x < l$ implies $x^T P_{aa}^{-1} x < l$, *i.e.*, $\mathcal{B}_{P_{cc}} \subset \mathcal{B}_{P_{aa}}$. Similarly, $\mathcal{B}_{P_{cc}} \subset \mathcal{B}_{P_{bb}}$. ■

2. For any point $x \in \mathcal{B}_{P_{aa}} \cap \mathcal{B}_{P_{bb}}$, there exists a correlation matrix $\tilde{P}_{ab}$ such that (i) P as given in (6) is positive definite, and (ii) $x \in \mathcal{B}_{P_{cc}}$ where P_{cc} is given by (7).
Proof: First we assume that $l > x^T P_{aa}^{-1} x > x^T P_{bb}^{-1} x > 0$. Define
$$\lambda^2 \triangleq \frac{x^T P_{bb}^{-1} x}{x^T P_{aa}^{-1} x}$$
It follows that $\lambda < 1$. Since the vector $\lambda P_{aa}^{-\frac{1}{2}} x$ and $P_{bb}^{-\frac{1}{2}} x$ have the same length, one can be rotated to another by a unitary matrix U, *i.e.*,
$$\lambda P_{aa}^{-\frac{1}{2}} x = U P_{bb}^{-\frac{1}{2}} x, \; UU^T = U^T U = I$$
The matrix
$$\tilde{P}_{ab} \triangleq \lambda P_{aa}^{\frac{1}{2}} U P_{bb}^{\frac{1}{2}}$$
satisfies our requirements:
   (i) P as given in (6) is positive definite: $P_{aa} > 0$, and $P_{bb} - P_{ab}^T P_{aa}^{-1} P_{ab} = (1 - \lambda^2) P_{bb} > 0$.
   (ii) For P_{cc} as given in (7), we have
$$x^T P_{cc}^{-1} x = x^T P_{aa}^{-1} x + x^T (P_{aa}^{-1}\tilde{P}_{ab} - I) \cdot (P_{bb} - \tilde{P}_{ab}^T P_{aa}^{-1}\tilde{P}_{ab})^{-1}(\tilde{P}_{ab}^T P_{aa}^{-1} - I)x$$
$$= x^T P_{aa}^{-1} x < l$$
   and therefore, $x \in \mathcal{B}_{P_{cc}}$.

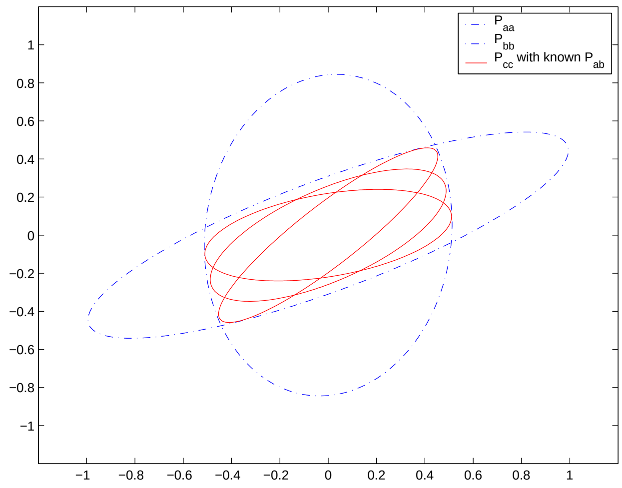

Figure 1: Covariance with known P_{ab} always lies inside the intersection

Other cases can be similarly proved by symmetry or by continuity argument (since we use strict inequality in (8)). ■

Based on the above observation, when $\tilde{P}_{ab}$ is not known, a consistent estimate of P_{cc} should be such that $\mathcal{B}_{P_{cc}} \supset \mathcal{B}_{P_{aa}} \cap \mathcal{B}_{P_{bb}}$, or loosely speaking, P_{cc} should include the intersection of P_{aa} and P_{bb}. This motivated the Covariance Intersection Algorithm [1]:

$$P_{cc}^{-1} = \omega P_{aa}^{-1} + (1 - \omega) P_{bb}^{-1} \tag{9}$$

$$K_1 = \omega P_{cc} P_{aa}^{-1}, \;\; K_2 = (1 - \omega) P_{cc} P_{bb}^{-1} \tag{10}$$

where $\omega \in [0, 1]$ is a parameter.

An illustration of the above discussion on intersection is shown in Figures 1 and 2. In Figure 1, 3 different known $\tilde{P}_{ab}$ are chosen, and the corresponding optimal covariance matrices P_{cc} are obtained, whose covariance ellipses at level 1 are shown in solid lines, while those for P_{aa} and P_{bb} are shown in dotted lines. In Figure 2, 3 different values of ω are chosen, and the corresponding covariance matrices are shown in the same fashion.

The CI Algorithm requires ω to be optimized, for example by minimizing the trace or the determinant of P_{cc}. Since the CI Algorithm computes the gains K_1 and K_2 on the fly, it does not provide a complete solution to Problem 1, where the gains are fixed *a priori*. For Problem 2, we show in Section 4.2 that the CI Algorithm does provide the global optimal solution, even though it searches only along a one-dimensional curve as shown by Equation (10), while the gains in the general case can be chosen from $\mathbb{R}^{n \times n}$.

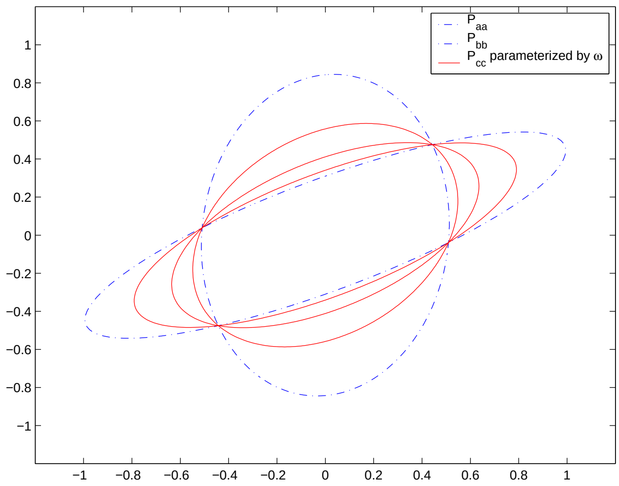

Figure 2: A family of covariance matrices containing the intersection.

4 A Family of Solutions

4.1 Problem 1

In order to obtain an upper bound for

$$\tilde{P}_{cc} = K_1\tilde{P}_{aa}K_1^T + K_2\tilde{P}_{bb}K_2^T + K_1\tilde{P}_{ab}K_2^T + K_2\tilde{P}_{ba}K_1^T$$

when the correlation $\tilde{P}_{ab}$ is unknown, the following inequality is utilized

$$\mathrm{E}\{(\sqrt{\gamma}K_1\tilde{a} - \frac{1}{\sqrt{\gamma}}K_2\tilde{b})(\sqrt{\gamma}K_1\tilde{a} - \frac{1}{\sqrt{\gamma}}K_2\tilde{b})^T\} \geq 0 \quad (11)$$

where $\gamma > 0$ is a scalar parameter. It follows that

$$\gamma K_1\tilde{P}_{aa}K_1^T + \frac{1}{\gamma}K_2\tilde{P}_{bb}K_2^T \geq K_1\tilde{P}_{ab}K_2^T + K_2\tilde{P}_{ba}K_1^T \quad (12)$$

Therefore, from (1) and (12), a consistent estimate P_{cc} of $\tilde{P}_{cc}$ is characterized by the family

$$P_{cc} = (1+\gamma)K_1P_{aa}K_1^T + (1+\frac{1}{\gamma})K_2P_{bb}K_2^T, \quad \gamma > 0 \quad (13)$$

Values of γ can be chosen to optimize various performance criteria, and we will refer to this type of optimality as "family-optimal". To minimize the trace of P_{cc}, note that

$$\begin{aligned} \mathrm{tr}\{P_{cc}\} &= \mathrm{tr}\{K_1P_{aa}K_1^T\} + \mathrm{tr}\{K_2P_{bb}K_2^T\} + \\ &\quad \gamma\mathrm{tr}\{K_1P_{aa}K_1^T\} + \frac{1}{\gamma}\mathrm{tr}\{K_2P_{bb}K_2^T\} \\ &\geq \mathrm{tr}\{K_1P_{aa}K_1^T\} + \mathrm{tr}\{K_2P_{bb}K_2^T\} + \\ &\quad 2\sqrt{\mathrm{tr}\{K_1P_{aa}K_1^T\}\mathrm{tr}\{K_2P_{bb}K_2^T\}} \\ &= \left(\sqrt{\mathrm{tr}\{K_1P_{aa}K_1^T\}} + \sqrt{\mathrm{tr}\{K_2P_{bb}K_2^T\}}\right)^2 \quad (14) \end{aligned}$$

where equality holds when

$$\gamma = \sqrt{\frac{\mathrm{tr}\{K_2P_{bb}K_2^T\}}{\mathrm{tr}\{K_1P_{aa}K_1^T\}}} \quad (15)$$

Therefore we have the following.

Theorem 1 (fixed gains) *For any given K_1 and K_2, a family of consistent estimates of the covariance matrix is given by (13). When γ is chosen by (15), the trace is minimized and its value is given by the last equality in (14).*

Note: Theorem 1 holds for *any* pair of K_1 and K_2, *i.e.*, it does not require the equality constraint $K_1 + K_2 = I$. This constraint is posed simply because we are interested in an *unbiased* estimator. It is not necessary for fusing two RV's in general.

It is worth pointing out that for a *given* value of ω, and the corresponding gains K_1 and K_2 in (10), the bound (9) is contained in the above family (13). More specifically, the P_{cc} given by (9) satisfies equation (13) for a particular choice of γ, *i.e.*, if

$$K_1 = \omega P_{cc}P_{aa}^{-1}, \; K_2 = (1-\omega)P_{cc}P_{bb}^{-1}$$

and

$$\gamma = \frac{1-\omega}{\omega}$$

equation (13) becomes

$$\begin{aligned} P_{cc} &= \frac{1}{\omega}\omega P_{cc}P_{aa}^{-1}P_{aa}P_{aa}^{-1}P_{cc}\omega + \\ &\quad \frac{1}{1-\omega}(1-\omega)P_{cc}P_{bb}^{-1}P_{bb}P_{bb}^{-1}P_{cc}(1-\omega) \\ &= \omega P_{cc}P_{aa}^{-1}P_{cc} + (1-\omega)P_{cc}P_{bb}^{-1}P_{cc} \\ &= P_{cc}(\omega P_{aa}^{-1} + (1-\omega)P_{bb}^{-1})P_{cc} \end{aligned}$$

which is satisfied by

$$P_{cc}^{-1} = \omega P_{aa}^{-1} + (1-\omega)P_{bb}^{-1}$$

which is (9).

4.2 Problem 2

Our solution to Problem 2 posed in Section 2 was motivated by the following observation. For any *given* ω, a pair of combination gains K_1 and K_2 are determined by (10). We have just shown that a consistent estimate of the resulting covariance matrix is given by the family (13), of which the matrix given in (9) is a special case. In terms of trace, the best in the family (13) is given by choosing γ as in (15), and this should yield a trace less than or equal to that of (9). How is the comparison when ω ranges over $[0,1]$? As an example, let

$$P_{aa} = \begin{bmatrix} 1.0 & 0.4 \\ 0.4 & 0.3 \end{bmatrix}, \; P_{bb} = \begin{bmatrix} 0.3 & 0.03 \\ 0.03 & 0.7 \end{bmatrix}$$

Figure 3 shows the trace of P_{cc} as a function of ω, with the solid line given by (9) in the CI Algorithm and the dashed line given by the family-optimal choice (14). As can be seen, the trace offered by the best ω matches the family-optimal one for that *particular* pair of K_1 and K_2. We next present an even stronger result: it matches the best family-optimal one for *all* pairs of K_1 and K_2.

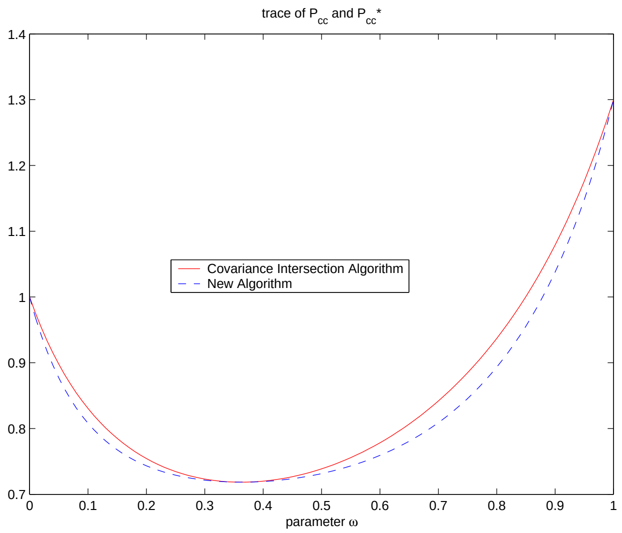

Figure 3: Trace as a function of ω: solid – CI, dashed – family-optimal

Note that a general solution to Problem 2 takes the form

$$K_1, K_2 = \underset{K_1,K_2}{\operatorname{argmin}} \min_{\gamma} J(P_{cc}(K_1, K_2, \gamma)), \quad K_1 + K_2 = I$$

where J represents a performance criteria such as trace or determinant. In the case of trace minimization, we show that the above optimal solution is actually given by the CI Algorithm (with trace minimization).

Theorem 2 (family-optimal gains) *There exists $\omega^* \in [0,1]$ such that*

$$h \triangleq \sqrt{\mathrm{tr}\{K_1 P_{aa} K_1^T\}} + \sqrt{\mathrm{tr}\{K_2 P_{bb} K_2^T\}}, \quad K_1 + K_2 = I \tag{16}$$

is minimized by

$$P_{cc} = (\omega^* P_{aa}^{-1} + (1-\omega^*) P_{bb}^{-1})^{-1}$$
$$K_1 = \omega^* P_{cc} P_{aa}^{-1}, \; K_2 = (1-\omega^*) P_{cc} P_{bb}^{-1}$$

Proof: For the case when h in (16) is minimized by either $K_1 = 0$ or $K_2 = 0$, ω^* can be chosen as 0 or 1. In the following we will assume that $K_1 \neq 0$ and $K_2 \neq 0$.

Define the following Lagrange function

$$L = \sqrt{\mathrm{tr}\{K_1 P_{aa} K_1^T\}} + \sqrt{\mathrm{tr}\{K_2 P_{bb} K_2^T\}} - \sum_{i,j} \Big(\Lambda \cdot (K_1 + K_2 - I) \Big)_{i,j}$$

where $\sum_{i,j} M_{i,j}$ is the summation of all the elements of the matrix M, the operator "$\cdot$" denotes elementwise product of two matrices, and Λ is a matrix of Lagrange multipliers. Using the identity

$$\frac{\partial \mathrm{tr}\{XPX^T\}}{\partial X} = 2XP, \; P = P^T$$

the stationary points are given by the following equations:

$$\frac{\partial L}{\partial K_1} = \frac{K_1 P_{aa}}{\sqrt{\mathrm{tr}\{K_1 P_{aa} K_1^T\}}} - \Lambda = 0 \tag{17}$$

$$\frac{\partial L}{\partial K_2} = \frac{K_2 P_{bb}}{\sqrt{\mathrm{tr}\{K_2 P_{bb} K_2^T\}}} - \Lambda = 0 \tag{18}$$

$$\frac{\partial L}{\partial \Lambda} = K_1 + K_2 - I = 0 \tag{19}$$

Let

$$\alpha \triangleq \sqrt{\mathrm{tr}\{K_1 P_{aa} K_1^T\}}, \;\; \beta \triangleq \sqrt{\mathrm{tr}\{K_2 P_{bb} K_2^T\}}$$

From (17) we have

$$K_1 = \alpha \Lambda P_{aa}^{-1}$$

Similarly,

$$K_2 = \beta \Lambda P_{bb}^{-1}$$

Substituting into (19) we have

$$\Lambda = (\alpha P_{aa}^{-1} + \beta P_{bb}^{-1})^{-1} \tag{20}$$

Note that $\Lambda^T = \Lambda$. Now the definition of α and β yields

$$\alpha = \sqrt{\mathrm{tr}\{\alpha \Lambda P_{aa}^{-1} P_{aa} P_{aa}^{-1} \Lambda \alpha\}} = \alpha \sqrt{\mathrm{tr}\{\Lambda P_{aa}^{-1} \Lambda\}}$$

or

$$\mathrm{tr}\{\Lambda P_{aa}^{-1} \Lambda\} = 1 \tag{21}$$

Similarly

$$\mathrm{tr}\{\Lambda P_{bb}^{-1} \Lambda\} = 1 \tag{22}$$

Equations (20), (21), and (22) lead to polynomial equations of order $2n$ in the variables α and β, where n is the dimension of the problem. Our objective here is to parameterize the solutions using a one dimensional variable. Recall that the minimum trace h^2 (where h is defined in the theorem) is achieved by the choice of (15) in the family (13). The parameter γ now becomes

$$\gamma = \sqrt{\frac{\mathrm{tr}\{K_2 P_{bb} K_2^T\}}{\mathrm{tr}\{K_1 P_{aa} K_1^T\}}} = \sqrt{\frac{\beta^2 \mathrm{tr}\{\Lambda P_{bb}^{-1}\Lambda\}}{\alpha^2 \mathrm{tr}\{\Lambda P_{aa}^{-1}\Lambda\}}} = \frac{\beta}{\alpha}$$

Thus the family-optimal covariance matrix is

$$\begin{aligned} P_{cc} &= (1+\frac{\beta}{\alpha})K_1 P_{aa} K_1^T + (1+\frac{\alpha}{\beta})K_2 P_{bb} K_2^T \\ &= (\alpha+\beta)\alpha\Lambda P_{aa}^{-1}\Lambda + (\alpha+\beta)\beta\Lambda P_{bb}^{-1}\Lambda \\ &= (\alpha+\beta)\Lambda(\alpha P_{aa}^{-1} + \beta P_{bb}^{-1})\Lambda \\ &= (\alpha+\beta)\Lambda \\ &= (\frac{\alpha}{\alpha+\beta}P_{aa}^{-1} + \frac{\beta}{\alpha+\beta}P_{bb}^{-1})^{-1} \end{aligned}$$

and the gains are

$$K_1 = \frac{\alpha}{\alpha+\beta} P_{cc} P_{aa}^{-1}, \quad K_2 = \frac{\beta}{\alpha+\beta} P_{cc} P_{bb}^{-1}$$

The theorem is proved by setting

$$\omega^* = \frac{\alpha}{\alpha+\beta}$$

■

According to the theorem, the n^2 dimensional optimization problem can be reduced to a one-dimensional one.

5 Generalizations

In this section we will generalize the previous results to the case where partial estimates are to be fused, and to the case with multiple RV's.

5.1 Partial Estimates

The Kalman Filter usually takes the form of combining a state estimate $\hat{x}$ with a measurement $y = Hx$. The results presented so far can be generalized to the case with partial observations in a straightforward fashion. If

$$\mathrm{E}\{a\} = H_a c^*, \ \mathrm{E}\{b\} = H_b c^*$$

then a linear unbiased estimator can be defined as [1]

$$c = K_1 a + K_2 b, \ K_1 H_a + K_2 H_b = I$$

[1]For $K_1 H_a + K_2 H_b = I$ to have a solution, $\begin{bmatrix} H_a \\ H_b \end{bmatrix}$ has to have full row rank.

The family of consistent estimates of the covariance matrix is again given by

$$P_{cc} = (1+\gamma)K_1 P_{aa} K_1^T + (1+\frac{1}{\gamma})K_2 P_{bb} K_2^T, \quad \gamma > 0 \tag{23}$$

When

$$\gamma = \sqrt{\frac{\mathrm{tr}\{K_2 P_{bb} K_2^T\}}{\mathrm{tr}\{K_1 P_{aa} K_1^T\}}}$$

P_{cc} has the minimal trace

$$\left(\sqrt{\mathrm{tr}\{K_1 P_{aa} K_1^T\}} + \sqrt{\mathrm{tr}\{K_2 P_{bb} K_2^T\}}\right)^2 \tag{24}$$

With a constraint on H_a and H_b:

$$H_a^T P_{aa}^{-1} H_a + H_b^T P_{bb}^{-1} H_b > 0$$

the CI Algorithm can be generalized [3]:

$$P_{cc}^{-1} = \omega H_a^T P_{aa}^{-1} H_a + (1-\omega) H_b^T P_{bb}^{-1} H_b \tag{25}$$

$$K_1 = \omega P_{cc} H_a^T P_{aa}^{-1}, \ K_2 = (1-\omega) P_{cc} H_b^T P_{bb}^{-1} \tag{26}$$

where $\omega \in [0, 1]$.

It can be shown in a similar fashion that for any given ω and hence K_1 and K_2, the family (23) contains (25) for a particular choice of $\gamma = \frac{1-\omega}{\omega}$. It can also be proved along the same line as in Theorem 2 that the trace (24) is minimized by K_1^* and K_2^* where the latter can be parameterized by the generalized CI Algorithm as in (25) and (26).

5.2 Multiple RV's

In this section we will extend the results for two RV's to multiple RV's. Notations will be changed from a and b to $x_1, x_2, x_3, ...$, and from c to z. To illustrate the generalization, we will describe in detail the case of three RV's, and the case of multiple RV's will follow in a straightforward fashion.

Let z^* be the mean value to be estimated. Define the estimation errors

$$\tilde{x}_i \triangleq x_i - z^*, \quad i = 1, 2, 3$$

Assume that the estimations are unbiased

$$\mathrm{E}\{\tilde{x}_i\} = 0$$

and the upper bounds on the covariance matrices

$$\mathrm{E}\{\tilde{x}_i \tilde{x}_i^T\} \triangleq \tilde{P}_i \leq P_i$$

are known. Assume that the correlations

$$\mathrm{E}\{\tilde{x}_i \tilde{x}_j^T\} \triangleq \tilde{P}_{ij}$$

are unknown. Our objective is to obtain a linear unbiased estimate

$$z = K_1 x_1 + K_2 x_2 + K_3 x_3$$

where

$$K_1 + K_2 + K_3 = I$$

and an upper bound on its covariance matrix

$$\mathrm{E}\{\tilde{z}\tilde{z}^T\} \triangleq \tilde{P}_z \leq P_z$$

where $\tilde{z} \triangleq z - z^*$.

For fixed gains K_1, K_2 and K_3, there are more than one way to obtain the upper bound P_z. We will present a family of such bounds based on our results discussed earlier. First we define

$$y \triangleq K_1 x_1 + K_2 x_2$$

then z can be re-written as

$$z = y + K_3 x_3$$

Thus P_z can be given by the family

$$(1+\gamma_2)P_y + (1+\frac{1}{\gamma_2})K_3 P_3 K_3^T, \;\; \gamma_2 > 0$$

where the covariance P_y of the estimation error of y can be chosen from the family

$$(1+\gamma_1)K_1 P_1 K_1^T + (1+\frac{1}{\gamma_1})K_2 P_2 K_2^T, \;\; \gamma_1 > 0$$

Combine the above two, and we have

$$P_z = (1+\gamma_2)\left((1+\gamma_1)K_1 P_1 K_1^T + (1+\frac{1}{\gamma_1})K_2 P_2 K_2^T\right) + (1+\frac{1}{\gamma_2})K_3 P_3 K_3^T, \; \gamma_1, \gamma_2 > 0 \tag{27}$$

Next we will find the minimum trace in this family. Define

$$t_i \triangleq \mathrm{tr}\{K_i P_i K_i^T\}, \; i = 1,2,3$$

Then

$$\mathrm{tr}\{P_z\} = (1+\gamma_2)((1+\gamma_1)t_1 + (1+\frac{1}{\gamma_1})t_2) + (1+\frac{1}{\gamma_2})t_3$$

Setting

$$\frac{\partial \mathrm{tr}\{P_z\}}{\partial \gamma_1} = 0, \; \frac{\partial \mathrm{tr}\{P_z\}}{\partial \gamma_2} = 0$$

we have

$$\gamma_1 = \frac{\sqrt{t_2}}{\sqrt{t_1}}, \; \gamma_2 = \frac{\sqrt{t_3}}{\sqrt{t_1}+\sqrt{t_2}} \tag{28}$$

and

$$\mathrm{tr}\{P_z\} = \left(\sqrt{t_1}+\sqrt{t_2}+\sqrt{t_3}\right)^2 \tag{29}$$

Thus we have shown that for given K_1, K_2 and K_3, a family of upper bounds (27) is available, and by choosing the parameters as in (28), the one with the minimum trace (29) can be chosen. To find K_1, K_2 and K_3 that minimizes this family-optimal trace, we proceed as follows. Define the following Lagrange function

$$\begin{aligned} L &= \sqrt{t_1}+\sqrt{t_2}+\sqrt{t_3} - \\ &\quad \sum_{i,j} \left(\Lambda \cdot (K_1+K_2+K_3-I)\right)_{i,j} \end{aligned}$$

as in Section 4. Setting partial derivatives of L with respect to K_i to zero yields

$$\frac{K_i P_i}{\sqrt{t_i}} - \Lambda = 0, \; i = 1,2,3$$

or

$$K_i = \sqrt{t_i}\Lambda P_i^{-1}, i = 1,2,3 \tag{30}$$

The constraint $K_1 + K_2 + K_3 = I$ gives

$$\Lambda = \left(\sqrt{t_1}P_1^{-1} + \sqrt{t_2}P_2^{-1} + \sqrt{t_3}P_3^{-1}\right)^{-1}$$

Recall that the family-optimal trace is given by the parameters chosen as in (28) and the covariance matrix is given by (27). With K_i chosen as in (30), we have

$$\begin{aligned} P_z &= \left(1+\frac{\sqrt{t_3}}{\sqrt{t_1}+\sqrt{t_2}}\right)\left(\left(1+\frac{\sqrt{t_2}}{\sqrt{t_1}}\right) t_1 \Lambda P_1^{-1}\Lambda + \right. \\ &\quad \left.\left(1+\frac{\sqrt{t_1}}{\sqrt{t_2}}\right) t_2 \Lambda P_2^{-1}\Lambda\right) + \left(1+\frac{\sqrt{t_1}+\sqrt{t_2}}{\sqrt{t_3}}\right) t_3 \Lambda P_3^{-1}\Lambda \\ &= \left(\sqrt{t_1}+\sqrt{t_2}+\sqrt{t_3}\right)\left(\sqrt{t_1}\Lambda P_1^{-1}\Lambda + \sqrt{t_2}\Lambda P_2^{-1}\Lambda + \right. \\ &\quad \left. \sqrt{t_3}\Lambda P_3^{-1}\Lambda\;\right) \\ &= \left(\sqrt{t_1}+\sqrt{t_2}+\sqrt{t_3}\right)\Lambda \\ &= \left(\frac{\sqrt{t_1}}{\sqrt{t_1}+\sqrt{t_2}+\sqrt{t_3}}P_1^{-1} + \frac{\sqrt{t_2}}{\sqrt{t_1}+\sqrt{t_2}+\sqrt{t_3}}P_2^{-1} + \right. \\ &\quad \left. \frac{\sqrt{t_3}}{\sqrt{t_1}+\sqrt{t_2}+\sqrt{t_3}}P_3^{-1}\right)^{-1} \\ &\triangleq \left(\omega_1 P_1^{-1} + \omega_2 P_2^{-1} + \omega_3 P_3^{-1}\right)^{-1} \end{aligned}$$

and

$$K_1 = \omega_1 P_z P_1^{-1}, \; K_2 = \omega_2 P_z P_2^{-1}, \; K_3 = \omega_3 P_z P_3^{-1}$$

where

$$\omega_1, \omega_2, \omega_3 \in [0,1], \; \omega_1 + \omega_2 + \omega_3 = 1$$

Thus the optimization can be conducted in $\mathbb{R}^2$ rather than in $\mathbb{R}^{2n^2}$.

6 Conclusions

The Covariance Intersection Algorithm is reexamined in this paper, in the general framework of obtaining a consistent estimate of the covariance matrix when combining two random variables with unknown correlation. For the case when the gains are chosen, a family of consistent estimates is given. For the case when optimal gains are to be found in order to minimize the trace of the estimated covariance, it is proved that the solution is given by the CI Algorithm, which conducts the search on a one-dimensional curve rather than in the whole parameter space, and thus the optimization problem can be solved very efficiently. The results have also been extended to the case of combining more than two variables, and to the case of partial observations where only $y = Hx$ is available, x being the quantity of interest. It is straightforward to extend the results to the case with dynamical equations. It is the authors' belief that with the newly gained understanding, Covariance Intersection Algorithm will find more applications in the areas of distributed filtering and estimation and data fusion.

7 Acknowledgment

This work was supported by NASA-JPL under contract NAS 3-01021.

References

[1] Simon J. Julier and Jeffrey K. Uhlmann. Non-divergent estimation algorithm in the presence of unknown correlations. In *Proceedings of the American Control Conference*, volume 4, pages 2369–2373, Piscataway, NJ, USA, 1997. IEEE.

[2] Simon Julier and Jeffrey Uhlmann. General decentralized data fusion with Covariance Intersection (CI). In D. Hall and J. Llians, editors, *Handbook of multisensor data fusion*, chapter 12, pages 12–1 to 12–25. CRC Press, 2001.

[3] Pablo O. Arambel, Constantino Rago, and Raman K. Mehra. Covariance intersection algorithm for distributed spacecraft state estimation. In *Proceedings of the 2001 American Control Conference*, volume 6, pages 4398 –4403, 2001.

[4] X. Xu and S. Negahdaripour. Application of extended covariance intersection principle for mosaic-based optical positioning and navigation of underwater vehicles. In *Proceedings of IEEE International Conference on Robotics and Automation*, volume 3, pages 2759 –2766, 2001.

[5] E T Baumgartner, H Aghazarian, A Trebi-Ollennu, T L Huntsberger, and M S. Garrett. State estimation and vehicle localization for the fido rover. In *Proceedings of SPIE - The International Society for Optical Engineering*, volume 4196, pages 329–336, Bellingham, WA, USA, 2000. Society of Photo-Optical Instrumentation Engineers.

[6] David Nicholson and Rob. Deaves. Decentralized track fusion in dynamic networks. In *Proceedings of SPIE - The International Society for Optical Engineering*, volume 4048, pages 452–460, Bellingham, WA, USA., 2000. Society of Photo-Optical Instrumentation Engineers.

[7] Simon J. Julier and Jeffrey K. Uhlmann. Real time distributed map building in large environments. In *Proceedings of SPIE - The International Society for Optical Engineering*, volume 4196, pages 317–328, Bellingham, WA, USA., 2000. Society of Photo-Optical Instrumentation Engineers.

[8] Jeffrey Uhlmann, Simon Julier, Behzad Kamgar-Parsi, Marco Lanzagorta, and Haw-Jye. Shyu. NASA mars rover: a testbed for evaluating applications of covariance intersection. In *Proceedings of Spie - the International Society for Optical Engineering*, volume 3693, pages 140–149, 1999.

[9] Ronald Mahler. Optimal/robust distributed data fusion: a unified approach. In *Proceedings of SPIE - The International Society for Optical Engineering*, volume 4052, pages 128–138, Bellingham, WA, USA, 2000. Society of Photo-Optical Instrumentation Engineers.

[10] C. Y. Chong and Shozo Mori. Convex combination and covariance intersection algorithms in distributed fusion. In *Proceedings of Fusion 2001*, 2001.

[11] Lingji Chen, Pablo A. Orambel, and Raman K. Mehra. Estimation under unknown correlation: Covariance Intersection revisited. *IEEE Transactions on Automatic Control*, 2002. to appear.