Connections and Differences between Covariance-Based and Bound-Based Information Fusion

Norton, J. P. [1, 2]

[1] Integrated Catchment Assessment & Management Centre, Fenner School of Environment & Society, The Australian National University, Canberra, ACT
[2] Mathematical Sciences Insitiute, The Australian National University, Canberra, ACT

Email: john.norton@anu.edu.au

Keywords: *Data fusion, set-membership estimation, ellipsoids*

EXTENDED ABSTRACT

In fields like image processing, target tracking and battlefield monitoring, the established use of Kalman filters has led to covariance-based information fusion. An almost equally longstanding, but non-probabilistic, alternative is to translate observations with known error bounds into bounds on state variables or model parameters. This yields the feasible set of all values compatible with all the observations. State or parameter bounding (henceforth just called bounding) may be motivated by its directness and simplicity, by a wish to avoid dubious probabilistic assumptions, or by a need to identify worst cases as a basis for decision-making. Bounding has potential in environmental applications, where data are often sparse and their errors are hard to characterise probabilistically, yet bounds on the errors can be specified.

Bounding can proceed recursively, imposing bounds derived from each new observation on the feasible set due to earlier ones. Bounded-error scalar observations linear in the state or parameters yield hyperplane bounds which define a polytope. The most popular recursive bounding algorithms update an ellipsoidal outer approximation to the polytope. The observation update fits an ellipsoid tightly round the intersection of the current bounding ellipsoid and the bounded region due to the new observation. The time update linearly transforms the current bounding ellipsoid to account for linear dynamics, vector-adds an ellipsoid which bounds the uncertain forcing, then ellipsoidally outer-bounds the result. The whole algorithm has a striking resemblance to the Kalman filter and some theoretical connections have been noted. For example, ellipsoidal parameter bounding (consisting of observation updates alone) can be viewed as recursive least squares with a dead band applied to the innovations. Another theoretical connection is that ellipsoidal bounds imply bounds on covariance.

Covariance intersection (CI) for information fusion has even stronger connections with ellipsoidal bounding. The family of ellipsoids, parameterised by a scalar, from which CI selects its approximation to the observation-updated covariance ellipsoid is identical to that used in bounding. Finding the smallest ellipsoid by minimising the trace or determinant of its describing matrix is an idea common to both. For bounding, efficient algorithms have been derived for both criteria. However, bounding addresses a more general problem in that the optimal new ellipsoid is determined by the centres, as well as the describing matrices, of the ellipsoids which it replaces. Bounding has also considered the time update, which is not part of CI. This paper notes the minimum-volume (minimum-determinant) ellipsoidal state-bounding algorithm of Maksarov and Norton (1996), stemming from initial work by Schweppe (1968, 1973), which provides efficient computation for minimum-volume CI. Results which allow checking of the compatibility of the prior state-error-covariance ellipsoid and the covariance ellipsoid resulting from a new vector observation are also noted.

1. INTRODUCTION

In estimation based on a model of a complex process, where the uncertainties due to modelling and measurement error are not fully understood and quantifiable, it is desirable to avoid distributional assumptions which are hard to test and likely to be idealised. For example, the Central Limit Theorem is often cited to justify an assumption that error variates are Gaussian, yet in practice errors may well be skewed (e.g. streamflow error), truncated (*e.g.* discretisation error or error in a non-negative variable) or dominated by other non-Gaussian effects. It is notable that the original derivation of the Kalman filter (Kalman, 1960) makes no distributional assumptions, but achieves orthogonality between estimates and errors.

Two well established ways to avoid assuming anything about distributions are complementary and strongly analogous to each other. The better-known way is to describe uncertainty solely through bias and covariance. Specifically, this paper considers minimum-covariance, linear, unbiased (MCLU) estimation, operating on means and covariances and providing one interpretation of Kalman filtering. The second distribution-free approach is state or parameter bounding (henceforth just called bounding), also known as set-membership estimation (Schweppe, 1973; Norton, 1987a; Walter and Piet-Lahanier, 1990; Norton, 1994, 1995; Milanese *et al.,* 1996; Walter, 2003). The idea is to translate bounded-error observations into bounds on state variables or model parameters, yielding the *feasible set* of all values compatible with all observations. Bounding not only avoids dubious distributional assumptions but also provides information about worst cases, a useful basis for decision or design. Bounding has potential for environmental applications (Norton, 1996), where data are often sparse and hard to characterise probabilistically, yet their error bounds can be specified, even if only in a "what if?" context. Moreover, bounding is suited to predictive modelling as an aid to environmental management, showing the range of credible outcomes of actions.

Bounding can be performed recursively, by predictor-corrector algorithms closely resembling the Kalman filter. Bounded-error observations linearly related to the state or parameters impose piecewise linear exact state or parameter bounds. The most popular recursive bounding algorithms (stemming from work by Schweppe (1968, 1973), Fogel and Huang (1982), Chenous'ko (1981) and others) use ellipsoidal outer approximations to those bounds. The observation update fits an ellipsoid tightly around the intersection of the current bounding ellipsoid and the bounded region implied by the error bounds of the new observation, as shown in Figure 1.

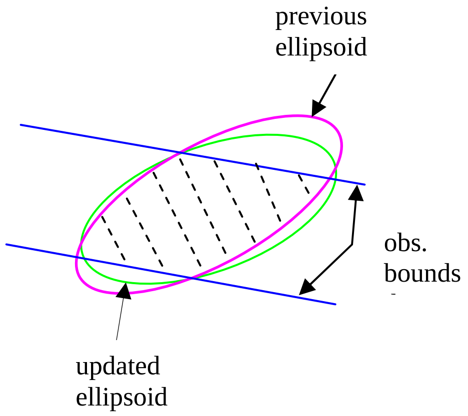

Figure 1. Ellipsoid updating on receipt of new bounded-error observation, linearly related to state or parameters. Feasible set is outer-bounded by hatched intersection, which is outer-bounded by updated ellipsoid.

The time update, for a linear dynamical model with additive forcing, linearly transforms the current bounding ellipsoid, vector-adds the bounded set of all possible effects of uncertain forcing, then ellipsoidally outer-bounds the result. Much of the algebra of the updates closely resembles Kalman filtering.

Some theoretical connections between bounding and covariance-based estimation have been noted. For example, the observation update of ellipsoidal parameter bounding can be viewed as a step of recursive least squares applying a dead band to the prediction error (Favier and Arruda, 1996). Another example, noted in Section 4, is that ellipsoidal bounds imply bounds on covariance.

Experiments (Maksarov and Norton, 1996) have shown ellipsoidal recursive state-bounding to have comparable performance to Kalman filtering, typically with mean-square error a little higher but with performance degraded less by asymmetry in the distribution of observation or process noise. Moreover, bounding does not require any assumption of whiteness and so can handle time-structured process or observation noise without requiring an auxiliary noise model. State bounding is thus an appealing alternative to conventional state estimation in some situations.

Information fusion in fields like image processing, target tracking and battlefield monitoring tends to be covariance-based, because state estimation is usually by Kalman filtering. Covariance intersection (CI) for information fusion (Julier and Uhlmann, 1997) has strong connections with ellipsoidal bounding. The family of matrices, parameterised by a scalar, from which CI selects its conservative estimate of the observation-updated state covariance is identical to that in the observation update of ellipsoidal bounding. In both cases an algorithm for minimising the trace of the updated matrix has been published. Minimising the determinant has also been suggested in both.

Even so, there are differences. Ellipsoidal bound updating solves a more general problem than CI, as discussed in Section 4, and analytical (but implicit) solutions have been derived for both the minimum-trace and minimum-determinant problems, allowing straightforward numerical solution. The paper summarises, in Section 3, the minimum-volume ellipsoidal state-bounding algorithm of Maksarov and Norton (1996). This provides minimum-determinant observation and time updates, the former giving minimum-volume CI as a special case.

The next two sections summarise MCLU estimation (as in Kalman filtering and CI) and bounding in terms of ellipsoids. Section 4 then points to further similarities and points of contact between the two approaches, and notes some possible extensions to MCLU estimation offered by existing bounding techniques.

2. MCLU ESTIMATION, COVARIANCE INTERSECTION AND ELLIPSOIDS

The canonical problem underlying MCLU estimation is to find, from unbiased estimates $\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2$ with error covariances $\mathbf{P}_1, \mathbf{P}_2$ and error cross-covariance $\mathbf{Q} \equiv E[\tilde{\mathbf{x}}_1 \tilde{\mathbf{x}}_2^T]$, an unbiased estimate $\hat{\mathbf{x}}$ linear in $\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2$ and with minimal error covariance $\mathbf{P} \equiv \mathrm{cov}(\hat{\mathbf{x}}) \equiv E[\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T]$. The solution is found by first imposing linearity and unbiasedness on the updating gains producing $\hat{\mathbf{x}}$, then finding the condition for a smooth minimum of the resulting $\mathbf{P}$:

$$\hat{\mathbf{x}} = \left(\mathbf{I} - (\mathbf{P}_1 - \mathbf{Q})\mathbf{R}\right)\hat{\mathbf{x}}_1 + \left(\mathbf{I} - (\mathbf{P}_2 - \mathbf{Q}^T)\mathbf{R}\right)\hat{\mathbf{x}}_2 \quad (1)$$

$$\begin{aligned}\mathbf{P} &= \mathbf{P}_1 - (\mathbf{P}_1 - \mathbf{Q})\mathbf{R}(\mathbf{P}_1 - \mathbf{Q}^T) \\ &= \mathbf{P}_2 - (\mathbf{P}_2 - \mathbf{Q}^T)\mathbf{R}(\mathbf{P}_2 - \mathbf{Q})\end{aligned} \quad (2)$$

where

$$\mathbf{R} \equiv (\mathbf{P}_1 + \mathbf{P}_2 - \mathbf{Q} - \mathbf{Q}^T)^{-1} \quad (3)$$

The information-matrix equivalent of (2) (Chen *et al.*, 2002) is

$$\begin{aligned}\mathbf{P}^{-1} &= \mathbf{P}_1^{-1} + \mathbf{P}_1^{-1}(\mathbf{P}_1^{-1}\mathbf{Q} - \mathbf{I})\ldots \\ &\ldots(\mathbf{P}_2 - \mathbf{Q}^T\mathbf{P}_1^{-1}\mathbf{Q})^{-1}(\mathbf{Q}^T\mathbf{P}_1^{-1} - \mathbf{I})\end{aligned} \quad (4)$$

When $\mathbf{Q}$ is unknown, *covariance intersection* (CI) (Julier and Uhlmann, 1997) replaces (4) by

$$\mathbf{P}'^{-1} = \omega\mathbf{P}_1^{-1} + (1-\omega)\mathbf{P}_2^{-1}, \quad (5)$$

with scalar $0 \le \omega \le 1$. The corresponding state estimate, still linear and unbiased, is found from

$$\mathbf{P}'^{-1}\hat{\mathbf{x}}' = \omega\mathbf{P}_1^{-1}\hat{\mathbf{x}}_1 + (1-\omega)\mathbf{P}_2^{-1}\hat{\mathbf{x}}_2 . \quad (6)$$

Julier and Uhlmann (1997) show that CI does not underestimate the error covariance of $\hat{\mathbf{x}}$, *i.e.* that $\mathbf{P}' \ge \mathbf{P}$, but a simpler proof is possible. First note that the information-matrix update (5) replaces

$$\mathbf{P}_0 \equiv \mathrm{cov}\left(\begin{bmatrix}\tilde{\mathbf{x}}_1 \\ \tilde{\mathbf{x}}_2\end{bmatrix}\right) \equiv \begin{bmatrix}\mathbf{P}_1 & \mathbf{Q} \\ \mathbf{Q}^T & \mathbf{P}_2\end{bmatrix} \quad (7)$$

in (4) by

$$\mathbf{P}'_0 \equiv \begin{bmatrix}\omega\mathbf{P}_1 & \mathbf{0} \\ \mathbf{0} & (1-\omega)\mathbf{P}_2\end{bmatrix} \quad (8)$$

and that any linear, unbiased updated state estimate must have the form

$$\hat{\mathbf{x}} = \mathbf{K}\hat{\mathbf{x}}_1 + (\mathbf{I} - \mathbf{K})\hat{\mathbf{x}}_2 \quad (9)$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{K}$ is any conforming matrix. The exact error covariance is then

$$\mathbf{P} = \begin{bmatrix}\mathbf{K} & \mathbf{I} - \mathbf{K}\end{bmatrix}\mathbf{P}_0\begin{bmatrix}\mathbf{K}^T \\ \mathbf{I} - \mathbf{K}^T\end{bmatrix}. \quad (10)$$

Hence

$$\mathbf{P}' - \mathbf{P} \equiv [\mathbf{K}\;\; \mathbf{I}-\mathbf{K}] \begin{bmatrix} \frac{1-\omega}{\omega}\mathbf{P}_1 & -\mathbf{Q} \\ -\mathbf{Q}^T & \frac{\omega}{1-\omega}\mathbf{P}_2 \end{bmatrix} \begin{bmatrix} \mathbf{K}^T \\ \mathbf{I}-\mathbf{K}^T \end{bmatrix}$$

$$= E\left[\begin{array}{l} \left(\mathbf{K}\sqrt{\frac{1-\omega}{\omega}}\tilde{\mathbf{x}}_1 - (\mathbf{I}-\mathbf{K})\sqrt{\frac{\omega}{1-\omega}}\tilde{\mathbf{x}}_2\right)\ldots \\ \ldots\left(\sqrt{\frac{1-\omega}{\omega}}\tilde{\mathbf{x}}_1^T\mathbf{K}^T - \sqrt{\frac{\omega}{1-\omega}}\tilde{\mathbf{x}}_2^T\left(\mathbf{I}-\mathbf{K}^T\right)\right) \end{array}\right]$$

$$\geq \mathbf{0}\;\; \forall\;\; 0 \leq \omega \leq 1 \text{ and } \forall\, \mathbf{K}. \tag{11}$$

Julier and Uhlmann (1997) indicate, with a figure, that $\mathbf{P}' \geq \mathbf{P}$ for all $\mathbf{Q}$ can be visualised as $\mathsf{E}(\mathbf{0}, \mathbf{P}')$ containing the envelope of all ellipsoids

$$\mathsf{E}(\mathbf{0}, \mathbf{P}(\mathbf{Q})) \equiv \left\{\mathbf{x} \mid \mathbf{x}^T\mathbf{P}(\mathbf{Q})^{-1}\mathbf{x} \leq 1\right\} \tag{12}$$

which is the intersection of $\mathsf{E}(\mathbf{0}, \mathbf{P}_1)$ and $\mathsf{E}(\mathbf{0}, \mathbf{P}_2)$. The parameter ω can be optimised to tighten $\mathsf{E}(\mathbf{0}, \mathbf{P}')$, and they suggest minimising either the trace of $\mathbf{P}'$ (sum of squares of half-axis lengths of $\mathsf{E}(\hat{\mathbf{x}}, \mathbf{P}')$) or det $\mathbf{P}'$ (proportional to square of hypervolume of $\mathsf{E}(\mathbf{0}, \mathbf{P}')$). They do not produce an analytical solution for either, but Chen *et al*. (2002) give the minimum-trace solution.

3. STATE BOUNDING AND ELLIPSOIDS

Schweppe (1968) and others (Kurzhanski, 1977; Chernous'ko, 1981) introduced the idea of a recursive algorithm to update ellipsoidal bounds on the state of a scalar-output model conforming to the ordinary Kalman filter model. The specified means and covariances of the initial state error and process noise are replaced by ellipsoidal bounds, and the observation-noise covariance by bounds. In this way uncertainty is handled without probabilistic assumptions (not even whiteness).

Each time update linearly transforms the state-bounding ellipsoid by pre- and post-multiplying by the state-transition matrix, then expands it by vector-adding an ellipsoid which bounds the unknown additive forcing. Chernous'ko (1981) solved the time-update problem by finding the smallest ellipsoid which contains the vector sum, from a family of scalar-weighted convex combinations of the summed ellipsoids. Maksarov and Norton (1996) provide a simpler derivation.

The original version of the observation update was for scalar observations. It finds an ellipsoid containing the intersection of the current state-bounding ellipsoid and the strip of state values defined by the two linear bounds on state resulting from given upper and lower bounds on the observation. A vector observation is processed by a sequence of scalar-observation updates. The minimum-trace or minimum-determinant updated state-bounding ellipsoid is found. Treating the strip as a degenerate ellipsoid with all but one axis infinite, Fogel and Huang (1982) gave an analytical solution for the minimum-volume ellipsoid in a family of scalar-weighted convex combinations of the two intersecting sets. Maksarov and Norton (1996) generalised the observation update to handle an ellipsoidal error bound on a vector observation; for simplicity the case of equal observation and state dimensions will be discussed, but no fundamental difficulty arises in other cases. If the state-bounding ellipsoid prior to imposing the observation-derived bound is

$$\mathsf{E}(\hat{\mathbf{x}}_1, \mathbf{P}_1) \equiv \left\{\mathbf{x} \mid (\mathbf{x}-\hat{\mathbf{x}}_1)^T\mathbf{P}_1^{-1}(\mathbf{x}-\hat{\mathbf{x}}_1) \leq 1\right\} \tag{13}$$

and the observation-derived state-bounding ellipsoid is $\mathsf{E}(\hat{\mathbf{x}}_2, \mathbf{P}_2)$ defined similarly, all points in their intersection are contained in any ellipsoid of the family

$$\mathsf{E}(\hat{\mathbf{x}}, \mathbf{P}) \equiv \left\{\mathbf{x} \mid (\mathbf{x}-\hat{\mathbf{x}})^T\mathbf{P}(\mathbf{Q})^{-1}(\mathbf{x}-\hat{\mathbf{x}}) \leq 1\right\}$$

$$\equiv \left\{\begin{array}{l} \mathbf{x} \mid \omega(\mathbf{x}-\hat{\mathbf{x}}_1)^T\mathbf{P}_1^{-1}(\mathbf{x}-\hat{\mathbf{x}}_1) + \ldots \\ \ldots(1-\omega)(\mathbf{x}-\hat{\mathbf{x}}_2)^T\mathbf{P}_2^{-1}(\mathbf{x}-\hat{\mathbf{x}}_2) \leq 1 \end{array}\right\} \tag{14}$$

with $0 \leq \omega \leq 1$. Equating the linear-in-**x** terms in the two expressions in (14) yields

$$\mathbf{P}'^{-1}\hat{\mathbf{x}} = \omega\mathbf{P}_1^{-1}\hat{\mathbf{x}}_1 + (1-\omega)\mathbf{P}_2^{-1}\hat{\mathbf{x}}_2 \tag{15}$$

with $\mathbf{P}'$ given by (5). This is exactly as for CI, with $\mathbf{x}$ identified with $\mathbf{x}'$, but, equating the constant and quadratic-in-**x** terms in (14), the defining matrix $\mathbf{P}$ of $\mathsf{E}(\hat{\mathbf{x}}, \mathbf{P})$ turns out to be not $\mathbf{P}'$ but $\mathbf{P} = \rho\mathbf{P}'$ where

$$\rho = 1 - (\hat{\mathbf{x}}_1 - \hat{\mathbf{x}}_2)^T\left(\frac{\mathbf{P}_1}{\omega} + \frac{\mathbf{P}_2}{1-\omega}\right)^{-1}(\hat{\mathbf{x}}_1 - \hat{\mathbf{x}}_2)$$

(16)

except in the degenerate cases $\omega = 0, \omega = 1$. With $\mathbf{P}_1, \mathbf{P}_2$ positive definite, $0 \le \rho \le 1$; $\rho = 1$ only if $\hat{\mathbf{x}}_1$ coincides with $\hat{\mathbf{x}}_2$.

Maksarov and Norton (1996) give analytical (but implicit) solutions for the ω's minimising the trace and determinant of **P**. Each is a polynomial equation with a unique root between 0 and 1. For the minimum-volume $\mathsf{E}(\hat{\mathbf{x}}, \mathbf{P})$, ω satisfies

$$\sum_{i=1}^{n} \frac{1}{\left(1 - \lambda_i(\mathbf{P}_1^{-1}\mathbf{P}_2)\right)\omega - 1} = n\left(\frac{\omega}{\rho}\frac{\partial \rho}{\partial \omega} + 1\right) \qquad (17)$$

where n is the state dimension and { λ_i } are the eigenvalues of $\mathbf{P}_1^{-1}\mathbf{P}_2$.

4. SIMILARITIES AND DIFFERENCES

The observation-update equations of CI and bounding differ only by the presence of the scale factor ρ in the **P** update of the latter. It arises because bounding requires the observation-updated ellipsoid to contain the intersection of the pre-observation ellipsoid $\mathsf{E}(\hat{\mathbf{x}}_1, \mathbf{P}_1)$ and the ellipsoid $\mathsf{E}(\hat{\mathbf{x}}_2, \mathbf{P}_2)$ due to the observation-error bound, whereas CI only requires that $\mathsf{E}(\mathbf{0}, \mathbf{P}')$ contains the intersection of $\mathsf{E}(\mathbf{0}, \mathbf{P}_1)$ and $\mathsf{E}(\mathbf{0}, \mathbf{P}_2)$. This is easily seen geometrically to maximise the size of $\mathsf{E}(\hat{\mathbf{x}}_1, \mathbf{P}_1) \cap \mathsf{E}(\hat{\mathbf{x}}_2, \mathbf{P}_2)$ with respect to both $\hat{\mathbf{x}}_1$ and $\hat{\mathbf{x}}_2$, so CI yields a larger ellipsoid in general than does bounding.

The volume of the ellipsoid produced by CI can be minimized simply by solving (17) for ω with the right-hand side simplified to n; to the author's knowledge this result has not been obtained elsewhere. Its practical utility depends on the relative computing loads of optimizing det(**P**) (maximizing det($\mathbf{P}^{-1}$)) directly by use of (5) and of searching for the unique root of (17) between 0 and 1. Even allowing for finding the eigenvalues of $\mathbf{P}_1^{-1}\mathbf{P}_2$, one would expect solving (17) to be quicker.

On a more fundamental level, bounding contrasts with CI in that the value of the observation, as well as the size of the ellipsoid indicating its uncertainty, influences how much new information the new observation (vector) adds to the previous information on state. The centre of the observation-derived ellipsoid depends on the value of the vector observation, and if the ellipsoid intersects the prior state-bounding ellipsoid only slightly, the updated ellipsoid $\mathsf{E}(\hat{\mathbf{x}}, \mathbf{P})$ bounding the intersection is small, *i.e.* the new observation sharpens the bounds on state greatly, even if the observation-noise-derived ellipsoid is large. The essential difference is that covariances are ensemble properties, whereas bounds are specific to each and every sample. The latter fact renders bounding vulnerable to mis-specified bounds (that is, to outliers (Norton and Veres, 1993)). On the other hand, it offers unequivocal indication of a clash between previous and new information when the prior and observation-derived ellipsoids do not intersect.

The question arises whether compatibility of prior and new information could be checked similarly in MCLU observation updates. In cases where the observation and state dimensions are equal and the excitation matrix is invertible, compatibility could be checked heuristically and straightforwardly by use of confidence ellipsoids, at the cost of assuming normality. One might decide to modify the noise covariances or reject the prior state estimate or the latest observation, depending on what is known of the reliability of the model, if the z% confidence ellipsoids (for some suitable z) around the prior state estimate and the value implied by the new observation do not intersect. Norton (2005) gives algorithms to check for the intersection of two ellipsoids, for determining by how much one must be expanded to intersect the other, and for checking whether one contains the other.

An interesting further link between ellipsoidal bounding and CI (and covariance-updating methods generally) is that the bound implies a bound on covariance. If an ellipsoidal bound

$$(\mathbf{x} - \hat{\mathbf{x}})^T \mathbf{S}^{-1} (\mathbf{x} - \hat{\mathbf{x}}) \le 1 \qquad (18)$$

on **x** exists, *i.e.* $\mathbf{x} \in \mathsf{E}(\hat{\mathbf{x}}, \mathbf{S})$ with $\hat{\mathbf{x}}, \mathbf{S}$ known, and with the covariance of **x** defined as

$$\mathbf{P} = \mathrm{cov}(\mathbf{x}) \equiv E[(\mathbf{x} - \bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}})^T] \qquad (19)$$

where $\bar{\mathbf{x}} \equiv E\mathbf{x}$, then for any real n-vector **a**,

$$\mathbf{a}^T\mathbf{Pa} = E[\left(\mathbf{a}^T(\mathbf{x}-\bar{\mathbf{x}})\right)^2] \le \max_{\mathbf{x}\in E(\hat{\mathbf{x}},\mathbf{S})} \{\left(\mathbf{a}^T(\mathbf{x}-\bar{\mathbf{x}})\right)^2\} \quad (20)$$

Now it is easy to show that $\left(\mathbf{a}^T(\mathbf{x}-\bar{\mathbf{x}})\right)^2$ reaches its maximum, subject to $\mathbf{x}\in E(\hat{\mathbf{x}},\mathbf{S})$, on the boundary of $\mathbf{x}\in E(\hat{\mathbf{x}},\mathbf{S})$. Defining the Lagrange cost function

$$L \equiv \left(\mathbf{a}^T(\mathbf{x}-\hat{\mathbf{x}})\right)^2 + \lambda\left((\mathbf{x}-\hat{\mathbf{x}})^T\mathbf{S}^{-1}(\mathbf{x}-\hat{\mathbf{x}})-1\right) \quad (21)$$

and setting $\partial L/\partial\mathbf{x} = \mathbf{0}$, $\partial L/\partial\lambda = 0$, the maximum is at

$$\mathbf{x}^* = \bar{\mathbf{x}} + \frac{\mathbf{Sa}}{\mathbf{a}^T\mathbf{x}^*} \quad (22)$$

where it is

$$\left(\mathbf{a}^T(\mathbf{x}^*-\bar{\mathbf{x}})\right)^2 = \mathbf{a}^T\mathbf{Sa} \quad (23)$$

Hence $\mathbf{a}^T\mathbf{Pa} \le \mathbf{a}^T\mathbf{Sa}\ \forall$ real $\mathbf{a}$. In other words, $\mathbf{P}\le\mathbf{S}$: **S** is an upper bound on cov(**x**).

It might be argued that this bound is usually loose, as over the ensemble of errors in **x, x** is not concentrated near the boundary of $E(\hat{\mathbf{x}},\mathbf{S})$. This will apply whenever several sources contribute significantly to the errors, implying that an error is near one of its bounds only when all the contributions are near their bounds (as when the Central Limit Theorem is relevant). However, as the state dimension *n* rises, a higher and higher proportion of the hypervolume of $E(\hat{\mathbf{x}},\mathbf{S})$ is near its boundary. For example, for *n* = 5, 40.95% of the hypervolume is more than 90% of the way from $\hat{\mathbf{x}}$ to the boundary, and for *n* = 20 the figure is 87.84%. The conclusion is that at higher state dimensions an ellipsoidal bound becomes an increasingly good bound on the covariance.

5. CONCLUSIONS

Similarities and differences between two complementary, distribution-free approaches to state estimation and information fusion, namely minimum-covariance, linear, unbiased estimation (encompassing Kalman filtering and CI) and ellipsoidal bounding, have been reviewed. Bounding has proceeded further than CI in some respects and offers some additional algorithmic possibilities for CI. The potential for checking compatibility between prior and new information, using ellipsoid-checking algorithms from bounding, has also been noted.

Two practical factors in bounding should be noted. First, in cases where the model is correct, so one can speak of the correct state or parameter values, avoidance of distributional assumptions does forfeit any improvement to be gained by knowing more than just the bounds. For instance, if the bounded observation errors are markedly skewed, the resulting state or parameter bounds are likely to be asymmetrical about the correct values. This is a deterministic counterpart of bias, although the bounds remain valid. Analogous asymmentry arises also when correlation between the variables being estimated and the observation error is unrecognised (Norton, 1987b), paralleling the "errors in variables" bias problem of MCLU estimation. Second, the bounding algorithm must take account of any known, physical constraints, such as non-negativity. They can be explicitly included, at every time update, if in a form compatible with the algorithm (linear or ellipsoidal bounds). If not, they have to be approximated, *e.g*. as piecewise linear.

It is worth mentioning that there is a good deal, not covered here, from the control-engineering literature on the ensemble properties (including convergence) of estimators presented with bounded-error observations, and on the worst-case performance of bounding algorithms. Given that for poor data and dubious models it is often easier to specify worst-case behaviour (beyond which gross errors could be detected and results rejected) than ensemble properties, the wider use of bounds, instead of or in addition to means and covariances, seems likely to be beneficial.

3. REFERENCES

Chen., L., P. O. Arambel and R. K. Mehra (2002) Estimation under unknown correlation: covariance intersection revisited, *IEEE Trans. on Automatic Control*, 47(11), 1879-1882.

Chernous'ko, F. (1981), Optimal guaranteed estimates of indeterminacies with the aid of ellipsoids, I *Eng. Cybernetics,* 18(3), 109; II 18(4), 1-9.

Favier, G. and L. V. R. Arruda (1996), Review and comparison of ellipsoidal boundingtechniques, in Bounding Approaches to System Identification, Eds. M. Milanese, J. Norton, H. Piet-Lahanier and E. Walter, Plenum Press, New York and London, 43-68.

Fogel, E. and Y. F. Huang (1982), On the value of informaion in system identification - bounded noise case, *Automatica* 18, 229-238.

Julier, S. K. and J. K. Uhlmann (1997), A non-divergent estimation algorithm in the presence of unknown correlations, *Proc. American Control Conference*, Albuquerque, New Mexico, USA, 4, 2369-2373.

Kalman, R. E. (1960) A new approach to linear filtering and prediction, *Trans. ASME J. of Basic Eng.* 82, 34-45.

Kurzhanski, A. B. (1977), Control and Observation under Conditions of Uncertainty, Nauka, Moscow (in Russian).

Kurzhanski, A. B. and I. Valyi (1991), Guaranteed state estimates for dynamic systems: beyond the overviews, *Proc. 9th IFAC/IFORS Symposium on Identification & System Parameter Estimation,* Budapest, 1033-1037.

Maksarov, D. G. and J. P. Norton (1996), State bounding with ellipsoidal set description of the uncertainty, *International Journal of Control* 65 (5), 847-866.

Milanese, M., J. Norton, H. Piet-Lahanier and E. Walter (Eds.) (1996), Bounding Approaches to System Identification Plenum Press, New York and London.

Norton, J. P. (1987a), Identification and application of bounded-parameter models, *Automat*ica 23(4), 497-507.

Norton,. J. P. (1987b) Identification of parameter bounds for ARMAX models from records with bounded noise, *Int. J. of Control* **45** 2, 375-390.

Norton, J. P. (2005), Results to aid applications of ellipsoidal state bounds, *Mathematical & Computer Modelling of Dynamical Systems* 11(2) (special issue on Set-Membership Modelling of Uncertainties in Dynamical Systems, eds. F. L. Chernousko & B. T. Polyak), 209-224.

Norton, J. P. (Ed.) (1994), Special issue of *Int. J. of Adaptive Control & Signal Processing* on Bounded-Error Estimation, 8(1), 1-118, 1994.

Norton, J. P. (Ed.) Special issue of *Int. J. of Adaptive Control & Signal Processing* on Bounded-Error Estimation, 9(1), 1-134, 1995.

Norton, J. P. (1996), Roles for deterministic bounding in environmental modelling, *Ecological Modelling* 6,157-161.

Norton, J. P. and S. M. Veres (1993) Outliers in bound-based state estimation and identification, *Proc. IEEE Int. Symp. on Circuits & Systems, Chicago,* 790-793.

Schweppe, F. C. (1968) Recursive state estimation: unknown but bounded errors and systen inputs, *IEEE Trans. on Automatic Control*, 13, 22-28.

Schweppe, F. C. (1973), Uncertain Dynamic systems, Prentice-Hall, Englewood Cliffs, NJ, USA.

Walter, E. and H. Piet-Lahanier (1990), Estimation of parameter bounds from bounded-error data: a survey, *Math. and Comput. in Simul.*, 32, 449-468.

Walter, E. (2003) Bound-based Identification, in Theme 6.43, Control Systems, Robotics and Automation, Part B: Modeling, Identification and Parameter Estimation, edited by H. Unbehauen, in Encyclopedia of Life Support Systems (EOLSS), Developed under the auspices of the UNESCO, Eolss Publishers, Oxford, UK, [http://www.eolss.net].