New Results for Stochastic Prediction and Filtering with Unknown Correlations

Uwe D. Hanebeck, Kai Briechle
Institute of Automatic Control Engineering
Technische Universität München
80290 München, Germany
Uwe.Hanebeck@ieee.org, Kai.Briechle@ei.tum.de

Abstract

This paper considers state estimation for dynamic systems in the case of nonwhite, mutually correlated noise processes. Here, the problem is complicated by the fact, that only the individual covariances are known; cross covariances between random variables obtained by taking individual noise processes at different time steps and between different noise processes are completely unknown. New estimator equations for solving this problem are derived in *feedback form* for both the prediction step and for the filtering step based on existing ideas known as covariance intersection. Solutions are given for the most general case of updating an N–dimensional state vector estimate based on M–dimensional observations. Furthermore, computationally efficient solutions for obtaining minimum covariance estimates are derived to avoid numerical optimization otherwise required.

1 Introduction

We consider the problem of state estimation for linear dynamic systems when nonwhite, mutually correlated noise processes are present. Solving this problem includes 1. sequentially updating the state estimate based on noisy observations and 2. propagating the state estimate through a dynamic system.

The main problem in deriving an appropriate filter is, that the assumed noise model leads to an unknown amount of correlation between a given state estimate and the noise processes. Hence, an appropriate estimator must yield results compatible with all possible levels of correlation. This is the key point of the proposed filter method, which is based on ideas known as covariance intersection [3, 1]: The unknown correlations are not neglected, but considered by producing conservative estimates for prediction and filtering step which are consistent with any level of correlation. Applying this kind of filter to complex estimation problems like multiple target tracking, robot localization, and decentralized fusion, often improves results significantly compared to a common Kalman filter based approach, especially if covariances cannot be estimated properly or the measurements are distorted by non-white noise [2, 4]. Good results may even be obtained in cases, where the Kalman filter based approach completely fails and the filter diverges, for example in decentralized control problems [1].

The main contribution of this paper is to derive new formulae for the prediction step and for the filtering step in *feedback form*, where the most general case of updating an N–dimensional state vector estimate based on M–dimensional observations is considered.

Furthermore, practically useful and computationally efficient solutions for obtaining minimum covariance estimates have been derived for both the prediction step and for the filtering step. This allows a simple and efficient overall implementation of the filter algorithm, because no numerical optimization routines are required.

We hope that the results are helpful for practitioners willing to implement this type of filter. Furthermore, since all the required derivations and proofs are included, it also provides the basis for further development along these lines.

In Sec. 2, a rigorous formulation of the problem of state estimation with unknown correlations is given. Section 3 then derives an algorithm for the prediction step (time update) in the presence of unknown correlations. An efficient solution for the optimal parameter for the prediction step is derived in Sec. 4. Subsequently, the filter step (measurement update) for vector measurements is derived in feedback form in Sec. 5, the solution for the corresponding optimal parameter is then given in Sec. 6. The scalar measurement case including a more explicit solution for obtaining minimum covariance estimates is treated separately in Sec. 7.

2 Problem Formulation

We consider a state space model with stochastic uncertainties according to

$$\underline{x}_{k+1} = \mathbf{A}_k \underline{x}_k + \mathbf{B}_k \underline{u}_k \text{ with } \underline{u}_k = \hat{\underline{u}}_k + \underline{e}_k^u \ , \quad (1)$$

$$\underline{\hat{y}}_k = \mathbf{H}_k \underline{x}_k + \underline{e}_k^y \ , \quad (2)$$

where the *nonwhite* additive input noise $\underline{e}_k^u$ is characterized by

$$\mathrm{Cov}\{\underline{e}_n^u, \underline{e}_m^u\} = \begin{cases} \tilde{\mathbf{C}}_n^{uu} & \text{for } n = m \\ \text{unknown} & \text{otherwise} \end{cases},$$

and the *nonwhite* additive output noise $\underline{e}_k^y$ by

$$\mathrm{Cov}\{\underline{e}_n^y, \underline{e}_m^y\} = \begin{cases} \tilde{\mathbf{C}}_n^{yy} & \text{for } n = m \\ \text{unknown} & \text{otherwise} \end{cases},$$

which is equivalent to unknown power spectra. Only upper bounds $\mathbf{C}_k^u$, $\mathbf{C}_k^y$ for the true covariance matrices are given by

$$\mathbf{C}_k^u \geq \tilde{\mathbf{C}}_k^{uu} ,$$
$$\mathbf{C}_k^y \geq \tilde{\mathbf{C}}_k^{yy} ,$$

where for two positive definite matrices $\mathbf{A}$ and $\mathbf{B}$, the expression $\mathbf{A} > \mathbf{B}$ is interpreted as $\mathbf{A} - \mathbf{B}$ positive definite. In addition, $\underline{e}_k^u$, $\underline{e}_k^y$ are possibly *mutually correlated*, i.e.,

$$\mathrm{Cov}\{\underline{e}_n^u, \underline{e}_m^y\} = \text{unknown for all } n, m ,$$

with an unknown amount of cross-correlation.

3 Time Update

At time step k, a state estimate $\underline{x}_k^s$ of the form

$$\underline{x}_k^s = \hat{\underline{x}}_k^s + \underline{e}_k^s$$

is given with mean $\hat{\underline{x}}_k^s$ and additive uncertainty $\underline{e}_k^s$. The joint covariance matrix of $\underline{e}_k^s$ and $\underline{e}_k^u$ is given by

$$\mathrm{Cov}\left\{\begin{bmatrix}\underline{e}_k^s \\ \underline{e}_k^u\end{bmatrix}\right\} = \begin{bmatrix}\tilde{\mathbf{C}}_k^{ss} & \tilde{\mathbf{C}}_k^{su} \\ \tilde{\mathbf{C}}_k^{us} & \tilde{\mathbf{C}}_k^{uu}\end{bmatrix} ,$$

where only upper bounds $\mathbf{C}_k^s$, $\mathbf{C}_k^u$ for the true covariances $\tilde{\mathbf{C}}_k^{ss}$, $\tilde{\mathbf{C}}_k^{uu}$ with

$$\mathbf{C}_k^s \geq \tilde{\mathbf{C}}_k^{ss} , \mathbf{C}_k^u \geq \tilde{\mathbf{C}}_k^{uu}$$

are known. The cross-covariances $\tilde{\mathbf{C}}_k^{su} = (\tilde{\mathbf{C}}_k^{us})^T$ are *completely unknown.*

When performing a time update at time step k by means of the system model (1), the mean $\hat{\underline{x}}_k^p$ of the predicted state $\underline{x}_k^p$ is simply given by the weighted sum of the mean of the last estimate $\hat{\underline{x}}_{k-1}^s$ and the estimate $\hat{\underline{u}}_{k-1}$ of the system input according to

$$\hat{\underline{x}}_k^p = \mathbf{A}_{k-1}\hat{\underline{x}}_{k-1}^s + \mathbf{B}_{k-1}\hat{\underline{u}}_{k-1} .$$

The problem is now to calculate the covariance of the predicted state $\underline{x}_k^p$, when the correlation between the previous estimate $\underline{x}_{k-1}^s$ and the system input $\underline{u}_{k-1}$ is unknown. Before solving this problem, we need the following two Theorems:

THEOREM 3.1 *Given a positive definite symmetric matrix* $\tilde{\mathbf{C}}$ *with*

$$\tilde{\mathbf{C}} = \begin{bmatrix}\mathbf{C}_{aa} & \mathbf{C}_{ab} \\ \mathbf{C}_{ba} & \mathbf{C}_{bb}\end{bmatrix} ,$$

a larger matrix $\mathbf{C} \geq \tilde{\mathbf{C}}$ *is given by*

$$\mathbf{C} = \begin{bmatrix}\frac{1}{0.5-\kappa}\mathbf{C}_{aa} & \mathbf{0} \\ \mathbf{0} & \frac{1}{0.5+\kappa}\mathbf{C}_{bb}\end{bmatrix} ,$$

with $\kappa \in (-0.5, 0.5)$.

PROOF. For $\mathbf{C} - \tilde{\mathbf{C}}$, we obtain

$$\mathbf{C} - \tilde{\mathbf{C}} = \begin{bmatrix}\left\{\frac{1}{0.5-\kappa} - 1\right\}\mathbf{C}_{aa} & -\mathbf{C}_{ab} \\ -\mathbf{C}_{ba} & \left\{\frac{1}{0.5+\kappa} - 1\right\}\mathbf{C}_{bb}\end{bmatrix}$$
$$= \begin{bmatrix}\frac{0.5+\kappa}{0.5-\kappa}\mathbf{C}_{aa} & -\mathbf{C}_{ab} \\ -\mathbf{C}_{ba} & \frac{0.5-\kappa}{0.5+\kappa}\mathbf{C}_{bb}\end{bmatrix} .$$

The quadratic form

$$Q = \begin{bmatrix}\underline{a}^T & \underline{b}^T\end{bmatrix}\left\{\mathbf{C} - \tilde{\mathbf{C}}\right\}\begin{bmatrix}\underline{a} \\ \underline{b}\end{bmatrix} = \begin{bmatrix}\underline{a}^T & \underline{b}^T\end{bmatrix}\begin{bmatrix}\lambda\mathbf{C}_{aa} & -\mathbf{C}_{ab} \\ -\mathbf{C}_{ba} & \frac{1}{\lambda}\mathbf{C}_{bb}\end{bmatrix}\begin{bmatrix}\underline{a} \\ \underline{b}\end{bmatrix}$$

with

$$\lambda = \frac{0.5+\kappa}{0.5-\kappa} \in [0, \infty)$$

and arbitrary vectors $\underline{a} \in \mathbb{R}^N$, $\underline{b} \in \mathbb{R}^M$ can be rewritten as

$$Q = \begin{bmatrix}\sqrt{\lambda}\,\underline{a}^T & -\frac{1}{\sqrt{\lambda}}\,\underline{b}^T\end{bmatrix}\underbrace{\begin{bmatrix}\mathbf{C}_{aa} & \mathbf{C}_{ab} \\ \mathbf{C}_{ba} & \mathbf{C}_{bb}\end{bmatrix}}_{\tilde{\mathbf{C}}}\begin{bmatrix}\sqrt{\lambda}\,\underline{a} \\ -\frac{1}{\sqrt{\lambda}}\,\underline{b}\end{bmatrix} .$$

Since $\tilde{\mathbf{C}}$ is positive definite, we have $Q \geq 0$, which implies that $\mathbf{C} - \tilde{\mathbf{C}}$ is positive semi-definite. □

REMARK 3.1 $\mathbf{C} \geq \tilde{\mathbf{C}}$ in Theorem 3.1 implies that

$$\mathrm{trace}(\mathbf{C}) \geq \mathrm{trace}(\tilde{\mathbf{C}}) ,$$
$$|\mathbf{C}| \geq |\tilde{\mathbf{C}}| ,$$
$$\mathbf{C}_{ii} \geq \tilde{\mathbf{C}}_{ii} \text{ for } i = 1 \ldots N+M .$$

LEMMA 3.1 *If matrix* $\mathbf{C}$ *is larger (or equal) than matrix* $\tilde{\mathbf{C}}$, *i.e.,* $\mathbf{C} \geq \tilde{\mathbf{C}}$, *then* $\mathbf{T}\mathbf{C}\mathbf{T}^T \geq \mathbf{T}\tilde{\mathbf{C}}\mathbf{T}^T$ *for an arbitrary matrix* $\mathbf{T}$.

THEOREM 3.2 *Given two correlated random vectors* $\underline{a}$, $\underline{b}$, *with means* $\hat{\underline{a}}$, $\hat{\underline{b}}$ *and true covariance matrix*

$$\mathrm{Cov}\left\{\begin{bmatrix}\underline{a} \\ \underline{b}\end{bmatrix}\right\} = \begin{bmatrix}\tilde{\mathbf{C}}_{aa} & \tilde{\mathbf{C}}_{ab} \\ \tilde{\mathbf{C}}_{ba} & \tilde{\mathbf{C}}_{bb}\end{bmatrix} ,$$

where, however, only upper bounds $\mathbf{C}_{aa}$, $\mathbf{C}_{bb}$ *for the true individual covariances* $\tilde{\mathbf{C}}_{aa}$, $\tilde{\mathbf{C}}_{bb}$ *with*

$$\mathbf{C}_{aa} \geq \tilde{\mathbf{C}}_{bb} \ ,$$
$$\mathbf{C}_{aa} \geq \tilde{\mathbf{C}}_{bb} \ ,$$

are known. The cross-covariances $\tilde{\mathbf{C}}_{ab} = \tilde{\mathbf{C}}_{ba}^T$ *are unknown. Then, an upper bound* $\mathbf{C}_{cc}$ *for the true covariance* $\tilde{\mathbf{C}}_{cc}$ *of the random vector*

$$\underline{c} = \mathbf{A}\underline{a} + \mathbf{B}\underline{b} \tag{3}$$

is given by

$$\mathbf{C}_{cc} = \frac{1}{0.5-\kappa}\mathbf{A}\mathbf{C}_{aa}\mathbf{A}^T + \frac{1}{0.5+\kappa}\mathbf{B}\mathbf{C}_{bb}\mathbf{B}^T \ .$$

PROOF. Writing (3) as

$$\underline{c} = \begin{bmatrix}\mathbf{A} & \mathbf{B}\end{bmatrix}\begin{bmatrix}\underline{a} \\ \underline{b}\end{bmatrix} \ ,$$

the true covariance $\tilde{\mathbf{C}}_{cc}$ of $\underline{c}$ is obviously given by

$$\tilde{\mathbf{C}}_{cc} = \begin{bmatrix}\mathbf{A} & \mathbf{B}\end{bmatrix}\begin{bmatrix}\tilde{\mathbf{C}}_{aa} & \tilde{\mathbf{C}}_{ab} \\ \tilde{\mathbf{C}}_{ba} & \tilde{\mathbf{C}}_{bb}\end{bmatrix}\begin{bmatrix}\mathbf{A}^T \\ \mathbf{B}^T\end{bmatrix} \ .$$

Replacing the covariance matrix by a larger one according to Theorem 3.1 gives

$$\mathbf{C}_{cc} = \begin{bmatrix}\mathbf{A} & \mathbf{B}\end{bmatrix}\begin{bmatrix}\frac{1}{0.5-\kappa}\mathbf{C}_{aa} & \mathbf{0} \\ \mathbf{0} & \frac{1}{0.5+\kappa}\mathbf{C}_{bb}\end{bmatrix}\begin{bmatrix}\mathbf{A}^T \\ \mathbf{B}^T\end{bmatrix} \ .$$

With Lemma 3.1, we have $\mathbf{C}_{cc} \geq \tilde{\mathbf{C}}_{cc}$, which concludes the proof. □

Based on this Theorem, the solution to calculating the covariance of the predicted state $\underline{x}_k^p$, when the correlation between the previous estimate $\underline{x}_{k-1}^s$ and the system input $\underline{u}_{k-1}$ is unknown, is given by

$$\mathbf{C}_k^p = \frac{1}{0.5-\kappa_{k-1}}\mathbf{A}_{k-1}\mathbf{C}_{k-1}^s\mathbf{A}_{k-1}^T + \frac{1}{0.5+\kappa_{k-1}}\mathbf{B}_{k-1}\mathbf{C}_{k-1}^u\mathbf{B}_{k-1}^T$$

for $\kappa_{k-1} \in (-0.5, 0.5)$. The time update is started with the initial state estimate $\hat{\underline{x}}_0^s$ with initial covariance $\mathbf{C}_0^s$.

4 Minimum Covariance Time Update

THEOREM 4.1 *The parameter* κ_k^{min} *for the minimum covariance matrix of the prediction result is given as the unique zero of*

$$\sum_{i=1}^{N} \frac{\bar{\mu}_k^i\kappa_k^2 + \kappa_k + 0.25\bar{\mu}_k^i}{\kappa_k + 0.5\bar{\mu}_k^i} \stackrel{!}{=} 0$$

in the interval $\kappa_k \in (-0.5, 0.5)$, *where* $\bar{\mu}_k^i$ *is given by*

$$\bar{\mu}_k^i = \frac{1+\mu_k^i}{1-\mu_k^i} \ ,$$

and μ_k^i, $i = 1, \ldots, N$, *are the generalized eigenvalues of the matrix pair* $\mathbf{A}_k\mathbf{C}_k^S\mathbf{A}_k^T$ *and* $\mathbf{B}_k\mathbf{C}_k^u\mathbf{B}_k^T$.

PROOF. Transforming the matrix pair $\mathbf{A}_k\mathbf{C}_k^S\mathbf{A}_k^T$ and $\mathbf{B}_k\mathbf{C}_k^u\mathbf{B}_k^T$ as $\mathbf{T}^T\mathbf{A}_k\mathbf{C}_k^S\mathbf{A}_k^T\mathbf{T} = \mathbf{I}$ and $\mathbf{T}^T\mathbf{B}_k\mathbf{C}_k^u\mathbf{B}_k^T\mathbf{T} = \operatorname{diag}\left[\mu_k^1, \ldots, \mu_k^N\right]$ results in

$$\det\{\mathbf{C}_k^p\} = \det\left\{\frac{1}{0.5-\kappa_k}\mathbf{I} + \frac{1}{0.5+\kappa_k}\operatorname{diag}\left[\mu_k^1, \ldots, \mu_k^N\right]\right\} \ ,$$

which is equivalent to

$$\det\{\mathbf{C}_k^p\} = \left\{\frac{1}{0.5-\kappa_k}\right\}^N \det\left\{\mathbf{I} + \frac{0.5-\kappa_k}{0.5+\kappa_k}\operatorname{diag}\left[\mu_k^1, \ldots, \mu_k^N\right]\right\} \ ,$$

where N is the dimension of the state space. Bilinear transformation of κ_k according to

$$\lambda_k = \frac{0.5-\kappa_k}{0.5+\kappa_k}$$

and differentiating with respect to λ_k gives the necessary condition

$$-N\left\{\frac{1+\lambda_k}{\lambda_k}\right\}^{N-1}\frac{1}{\lambda_k^2}\prod_{i=1}^{N}\left\{1+\mu_k^i\lambda_k\right\} + \left\{\frac{1+\lambda_k}{\lambda_k}\right\}^N\sum_{i=1}^{N}\left\{\mu_k^i\prod_{\substack{j=1\\j\neq i}}^{N}(1+\mu_k^j\lambda_k)\right\} = 0 \ ,$$

which can be simplified to

$$-N + \lambda_k(1+\lambda_k)\sum_{i=1}^{N}\frac{\mu_k^i}{1+\lambda_k\mu_k^i} = 0 \ .$$

With $N = \sum_{i=1}^{N} 1$ we obtain

$$\sum_{i=1}^{N}\frac{\mu_k^i\lambda_k^2 - 1}{1+\mu_k^i\lambda_k} = 0 \ .$$

Resubstituting κ_k gives

$$\sum_{i=1}^{N}\frac{\mu_k^i(0.5-\kappa_k)^2 + (0.5+\kappa_k)^2}{0.5+\kappa_k+(0.5-\kappa_k)\mu_k^i} = 0 \ .$$

Applying a bilinear transformation of the eigenvalues according to

$$\bar{\mu}_k^i = \frac{1+\mu_k^i}{1-\mu_k^i}$$

concludes the proof. □

5 Measurement Update: Vector Observations

We consider the measurement equation (2) with (uncertain) vector observations $\hat{\underline{y}}_k$ at time k. $\underline{x}_k$ denotes the state vector, $\underline{e}_k^y$ denotes the additive uncertainty. Furthermore, there exists a prior estimate $\hat{\underline{x}}_k^p$ of the state vector, which also suffers from additive uncertainty $\underline{e}_k^p$ according to

$$\underline{x}_k^p = \hat{\underline{x}}_k^p + \underline{e}_k^p \ .$$

$\underline{e}_k^p$, $\underline{e}_k^y$ are assumed to be correlated according to

$$\mathrm{Cov}\begin{pmatrix}\underline{e}_k^p \\ \underline{e}_k^y\end{pmatrix} = \begin{bmatrix}\tilde{\mathbf{C}}_k^{pp} & \tilde{\mathbf{C}}_k^{py} \\ \tilde{\mathbf{C}}_k^{yp} & \tilde{\mathbf{C}}_k^{yy}\end{bmatrix} \ . \tag{4}$$

Only upper bounds $\mathbf{C}_k^p$, $\mathbf{C}_k^y$ for the true covariances $\tilde{\mathbf{C}}_k^{pp}$, $\tilde{C}_k^{yy}$ are known according to

$$\mathbf{C}_k^p \geq \tilde{\mathbf{C}}_k^{pp} \ ,$$
$$\mathbf{C}_k^y \geq \tilde{\mathbf{C}}_k^{yy} \ .$$

The cross-covariances $\tilde{\mathbf{C}}_k^{py} = (\tilde{\mathbf{C}}_k^{yp})^T$ are *completely unknown.*

THEOREM 5.1 *A conservative state estimate based on a given vector observation $\hat{\underline{y}}_k$ is given by*

$$\begin{aligned}\hat{\underline{x}}_k^s &= \hat{\underline{x}}_k^p + \lambda_k \mathbf{C}_k^p \mathbf{H}_k^T \\ &\quad \left\{\mathbf{C}_k^y + \lambda_k \mathbf{H}_k \mathbf{C}_k^p \mathbf{H}_k^T\right\}^{-1} \left(\hat{\underline{y}}_k - \mathbf{H}_k \hat{\underline{x}}_k^p\right) \ , \\ \mathbf{C}_k^s &= (1+\lambda_k)\mathbf{C}_k^p - (1+\lambda_k)\lambda_k \mathbf{C}_k^p \mathbf{H}_k^T \\ &\quad \left\{\mathbf{C}_k^y + \lambda_k \mathbf{H}_k \mathbf{C}_k^p \mathbf{H}_k^T\right\}^{-1} \mathbf{H}_k \mathbf{C}_k^p \ ,\end{aligned} \tag{5}$$

with scaling parameter $\lambda_k \in [0, \infty)$.

PROOF. Consider the random vector

$$\underline{z}_k = \begin{bmatrix}\underline{x}_k^p \\ \hat{\underline{y}}_k\end{bmatrix} \ .$$

The mean of $\underline{z}_k$ is

$$E\{\underline{z}_k\} = \begin{bmatrix}\hat{\underline{x}}_k^p \\ \mathbf{H}_k \hat{\underline{x}}_k^p\end{bmatrix} \ .$$

The covariance matrix of $\underline{z}_k$ is given by

$$\begin{aligned}&\mathrm{Cov}\{\underline{z}_k\} = \mathrm{Cov}\left\{\begin{bmatrix}\underline{x}_k^p \\ \hat{\underline{y}}_k\end{bmatrix}\right\} \\ &= E\left\{\begin{bmatrix}\underline{x}_k^p - \hat{\underline{x}}_k^p \\ \hat{\underline{y}}_k - \mathbf{H}_k \hat{\underline{x}}_k^p\end{bmatrix}\begin{bmatrix}\underline{x}_k^p - \hat{\underline{x}}_k^p \\ \hat{\underline{y}}_k - \mathbf{H}_k \hat{\underline{x}}_k^p\end{bmatrix}^T\right\} \\ &= E\left\{\begin{bmatrix}\underline{e}_k^p (\underline{e}_k^p)^T & \underline{e}_k^p \{(\underline{e}_k^p)^T \mathbf{H}_k^T + (\underline{e}_k^y)^T\} \\ \{\mathbf{H}_k \underline{e}_k^p + \underline{e}_k^y\}\underline{e}_k^p & \{\mathbf{H}_k \underline{e}_k^p + \underline{e}_k^y\}\{(\underline{e}_k^p)^T \mathbf{H}_k^T + (\underline{e}_k^y)^T\}\end{bmatrix}\right\} \\ &= \begin{bmatrix}\tilde{\mathbf{C}}_k^{pp} & \tilde{\mathbf{C}}_k^{pp}(\mathbf{H}_k)^T + \tilde{\mathbf{C}}_k^{py} \\ \mathbf{H}_k \tilde{\mathbf{C}}_k^{pp} + \tilde{\mathbf{C}}_k^{yp} & \mathbf{H}_k \tilde{\mathbf{C}}_k^{pp} \mathbf{H}_k^T + \mathbf{H}_k \tilde{\mathbf{C}}_k^{py} + \tilde{\mathbf{C}}_k^{yp} \mathbf{H}_k^T + \tilde{\mathbf{C}}_k^{yy}\end{bmatrix} \ .\end{aligned}$$

Bounding the true covariance matrix in (4) from above by

$$\mathrm{Cov}\left\{\begin{bmatrix}\underline{e}_k^p \\ \underline{e}_k^y\end{bmatrix}\right\} \leq \begin{bmatrix}\frac{1}{0.5-\kappa}\mathbf{C}_k^p & \mathbf{0} \\ \mathbf{0} & \frac{1}{0.5+\kappa}\mathbf{C}_k^y\end{bmatrix}$$

gives

$$\mathrm{Cov}\{\underline{z}_k\} \leq \begin{bmatrix}\frac{1}{0.5-\kappa_k}\mathbf{C}_k^p & \frac{1}{0.5-\kappa_k}\mathbf{C}_k^p \mathbf{H}_k^T \\ \frac{1}{0.5-\kappa_k}\mathbf{H}_k \mathbf{C}_k^p & \frac{1}{0.5+\kappa_k}\mathbf{C}_k^y + \frac{1}{0.5-\kappa_k}\mathbf{H}_k \mathbf{C}_k^p \mathbf{H}_k^T\end{bmatrix} \ .$$

Hence, according to the appendix we have

$$\begin{aligned}\hat{\underline{x}}_k^s &= \hat{\underline{x}}_k^p + \frac{1}{0.5-\kappa_k}\mathbf{C}_k^p \mathbf{H}_k^T \left\{\frac{1}{0.5+\kappa_k}\mathbf{C}_k^y \right. \\ &\quad \left. + \frac{1}{0.5-\kappa_k}\mathbf{H}_k \mathbf{C}_k^p \mathbf{H}_k^T\right\}^{-1} (\hat{\underline{y}}_k - \mathbf{H}_k \hat{\underline{x}}_k^p) \ , \\ \mathbf{C}_k^s &= \frac{1}{0.5-\kappa_k}\mathbf{C}_k^p - \frac{1}{(0.5-\kappa_k)^2}\mathbf{C}_k^p \mathbf{H}_k^T \left\{\frac{1}{0.5+\kappa_k}\mathbf{C}_k^y \right. \\ &\quad \left. + \frac{1}{0.5-\kappa_k}\mathbf{H}_k \mathbf{C}_k^p \mathbf{H}_k^T\right\}^{-1} \mathbf{H}_k \mathbf{C}_k^p \ ,\end{aligned}$$

or equivalently

$$\begin{aligned}\hat{\underline{x}}_k^s &= \hat{\underline{x}}_k^p + \frac{0.5+\kappa_k}{0.5-\kappa_k}\mathbf{C}_k^p \mathbf{H}_k^T \left\{\mathbf{C}_k^y \right. \\ &\quad \left. + \frac{0.5+\kappa_k}{0.5-\kappa_k}\mathbf{H}_k \mathbf{C}_k^p \mathbf{H}_k^T\right\}^{-1} (\hat{\underline{y}}_k - \mathbf{H}_k \hat{\underline{x}}_k^p) \ , \\ \mathbf{C}_k^s &= \frac{1}{0.5-\kappa_k}\mathbf{C}_k^p - \frac{1}{0.5-\kappa_k}\frac{0.5+\kappa_k}{0.5-\kappa_k}\mathbf{C}_k^p \mathbf{H}_k^T \left\{\mathbf{C}_k^y \right. \\ &\quad \left. + \frac{0.5+\kappa_k}{0.5-\kappa_k}\mathbf{H}_k \mathbf{C}_k^p \mathbf{H}_k^T\right\}^{-1} \mathbf{H}_k \mathbf{C}_k^p \ .\end{aligned}$$

With the bilinear transformation

$$\lambda_k = \frac{0.5+\kappa_k}{0.5-\kappa_k} \tag{6}$$

we have

$$\frac{1}{0.5-\kappa_k} = 1 + \lambda_k \ , \tag{7}$$

which achieves the proof. □

6 Minimum Covariance Measurement Update

THEOREM 6.1 *The scaling parameter λ_k^{min} for a minimum covariance matrix of the filtering result is given by*

$$\lambda_k^{min} = \frac{0.5+\kappa_k^{min}}{0.5-\kappa_k^{min}} \ ,$$

where κ_k^{min} is the unique zero of the expression

$$\sum_{i=1}^{M} \frac{\kappa_k - 0.5}{\kappa_k + 0.5\bar{\mu}_k^i} \stackrel{!}{=} M - N$$

with $\kappa_k \in (-0.5, 0.5)$. Note that M is the dimension of the measurement vector and N is the dimension of the state vector. $\bar{\mu}_k^i$ are given by the bilinear transformation

$$\bar{\mu}_k^i = \frac{1 + \mu_k^i}{1 - \mu_k^i} \ .$$

where μ_k^i, $i = 1, \ldots, M$, are the generalized eigenvalues of the real, symmetric, positive definite matrix pair $\mathbf{C}_k^y$ and $\mathbf{H}_k \mathbf{C}_k^p \mathbf{H}_k^T$.

PROOF. We have

$$\det\{\mathbf{C}_k^s\} = \frac{(1+\lambda_k)^N \det\{\mathbf{C}_k^p\} \det\{\mathbf{C}_k^y\}}{\det\{\mathbf{C}_k^y + \lambda_k \mathbf{H}_k \mathbf{C}_k^p \mathbf{H}_k^T\}} \ .$$

Transforming the matrix pair $\mathbf{C}_k^y$, $\mathbf{H}_k \mathbf{C}_k^p \mathbf{H}_k^T$ according to $\mathbf{T}^T \mathbf{C}_k^y \mathbf{T} = \mathbf{I}$, $\mathbf{T}^T \mathbf{H}_k \mathbf{C}_k^p \mathbf{H}_k^T \mathbf{T} = \mathrm{diag}[\mu_k^1, \ldots, \mu_k^M]$ we obtain

$$\det\{\mathbf{C}_k^s\} = c_k \frac{(1+\lambda_k)^N}{\prod_{i=1}^{M} (1 + \mu_k^i \lambda_k)}$$

with $c_k = \det\{\mathbf{C}_k^p\} \det\{\mathbf{C}_k^y\}$ and hence

$$\det\{\mathbf{C}_k^s\} = c_k \prod_{i=1}^{M} \frac{1}{1 + \mu_k^i \lambda_k} (1+\lambda_k)^N \ .$$

As a necessary condition for $\lambda_k^{\min}$ we obtain

$$\frac{\partial}{\partial \lambda_k} \det\{\mathbf{C}_k^s\} = c_k \left\{ \prod_{i=1}^{M} (1 + \mu_k^i \lambda_k)^2 \right\}^{-1} \left\{ N(1+\lambda_k)^{N-1} \right.$$
$$\left. \prod_{i=1}^{M} (1 + \mu_k^i \lambda_k) - (1+\lambda_k)^N \sum_{i=1}^{M} \left[\mu_k^i \prod_{\substack{j=1 \\ j \neq i}}^{M} (1 + \mu_k^j \lambda_k) \right] \right\} \stackrel{!}{=} 0$$

Because of $\mu_k^i > 0$ for $i = 1, \ldots, M$ and $\lambda_k \in [0, \infty)$ the relations

$$(1+\lambda_k)^{N-1} \geq 1 \ ,$$
$$\prod_{i=1}^{M} (1 + \mu_k^i \lambda_k) \geq 1 \ ,$$
$$\prod_{i=1}^{M} (1 + \mu_k^i \lambda_k)^2 \geq 1$$

hold and the above condition is simplified to

$$\sum_{i=1}^{M} \frac{1 - \mu_k^i}{1 + \mu_k^i \lambda_k} \stackrel{!}{=} M - N \ .$$

Applying bilinear transformations according to

$$\lambda_k^{\min} = \frac{0.5 + \kappa_k^{\min}}{0.5 - \kappa_k^{\min}} \ , \quad \bar{\mu}_k^i = \frac{1 + \mu_k^i}{1 - \mu_k^i}$$

gives the desired result. □

7 Measurement Update: Scalar Observations

The case of estimating a vector state from scalar observations is treated separately from the general case, since more explicit results can be obtained. We consider an (uncertain) scalar observation $\hat{y}_k$ at time k, with the associated measurement equation

$$\hat{y}_k = \underline{H}_k^T \underline{x}_k + e_k^y \ . \tag{8}$$

$\underline{x}_k$ denotes the state vector, e_k^y denotes the additive uncertainty. Furthermore, there exists a prior estimate $\underline{\hat{x}}_k^p$ of the state vector. $\underline{\hat{x}}_k^p$ also suffers from additive uncertainty $\underline{e}_k^p$ according to

$$\underline{x}_k^p = \underline{\hat{x}}_k^p + \underline{e}_k^p \ .$$

$\underline{e}_k^p$, e_k^y are assumed to be correlated according to

$$\mathrm{Cov} \begin{bmatrix} \underline{e}_k^p \\ e_k^y \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{C}}_k^{pp} & \underline{\tilde{C}}_k^{py} \\ (\underline{\tilde{C}}_k^{yp})^T & \tilde{C}_k^{yy} \end{bmatrix} \ .$$

Again, only upper bounds $\mathbf{C}_k^p$, C_k^y for the true covariances $\tilde{\mathbf{C}}_k^{pp}$, $\tilde{C}_k^{yy}$ are known according to

$$\mathbf{C}_k^p \geq \tilde{\mathbf{C}}_k^{pp} \ , \ C_k^y \geq \tilde{C}_k^{yy}$$

and the cross-covariances $\underline{\tilde{C}}_k^{py} = \underline{\tilde{C}}_k^{yp}$ are *completely unknown.*

THEOREM 7.1 *A conservative estimate for the state in the linear measurement equation according to (8) with a given scalar observation $\hat{y}_k$ is given by*

$$\underline{\hat{x}}_k^s = \underline{\hat{x}}_k^p + \lambda_k \frac{\mathbf{C}_k^p \underline{H}_k}{C_k^y + \lambda_k \underline{H}_k^T \mathbf{C}_k^p \underline{H}_k} \left(\hat{y}_k - \underline{H}_k^T \underline{\hat{x}}_k^p \right) \ , \tag{9}$$

$$\mathbf{C}_k^s = (1+\lambda_k)\mathbf{C}_k^p - (1+\lambda_k)\lambda_k \frac{\mathbf{C}_k^p \underline{H}_k \underline{H}_k^T \mathbf{C}_k^p}{C_k^y + \lambda_k \underline{H}_k^T \mathbf{C}_k^p \underline{H}_k} \tag{10}$$

with scaling parameter $\lambda_k \in [0, \infty)$.

PROOF. Similar to vector case. □

THEOREM 7.2 *With $N \geq 2$ the dimension of the state space and $G_k = \underline{H}_k^T \mathbf{C}_k^p \underline{H}_k$, the minimum size of $\mathbf{C}_k^s$ in (10) is attained for λ_k^{min} given by*

$$\lambda_k^{min} = \frac{G_k - NC_k^y}{(N-1)G_k} \ .$$

PROOF. Minimizing the volume of

$$\mathbf{C}_k^s = (1+\lambda_k)\mathbf{C}_k^p - (1+\lambda_k)\lambda_k \frac{\mathbf{C}_k^p \underline{H}_k \underline{H}_k^T \mathbf{C}_k^p}{C_k^y + \lambda_k \underline{H}_k^T \mathbf{C}_k^p \underline{H}_k}$$

is equivalent to minimizing

$$\det(\mathbf{C}_k^s) = \det(\mathbf{C}_k^p) \det\left((1+\lambda_k)\mathbf{I} - (1+\lambda_k)\lambda_k \frac{\underline{H}_k \underline{H}_k^T \mathbf{C}_k^p}{C_k^y + \lambda_k \underline{H}_k^T \mathbf{C}_k^p \underline{H}_k} \right) .$$

From basic linear algebra, we have

$$\det(c\mathbf{I} + \underline{a}\underline{b}^T) = c^{L-1}\left(c + \underline{a}^T \underline{b}\right)$$

with L the dimension of the vectors $\underline{a}$, $\underline{b}$ and scalar c. Hence, we obtain

$$\det(\mathbf{C}_k^s) = \det(\mathbf{C}_k^p)(1+\lambda_k)^{N-1}\left(1 + \lambda_k - (1+\lambda_k)\lambda_k \frac{\underline{H}_k^T \mathbf{C}_k^p \underline{H}_k}{C_k^y + \lambda_k \underline{H}_k^T \mathbf{C}_k^p \underline{H}_k}\right) ,$$

which can be simplified to

$$\det(\mathbf{C}_k^s) \sim (1+\lambda_k)^N \frac{C_k^y}{C_k^y + \lambda_k G_k} \ .$$

Differentiation with respect to λ_k yields

$$\frac{\partial}{\partial \lambda_k} \det(\mathbf{C}_k^s) \sim (1+\lambda_k)^{N-1} C_k^y \frac{N(C_k^y + \lambda_k G_k) - (1+\lambda_k)G_k}{(C_k^y + \lambda_k G_k)^2} \ .$$

Setting the result to zero gives the necessary condition

$$\lambda_k^{\min}(N-1)G_k + NC_k^y - G_k = 0 \ .$$

With

$$\frac{\partial^2}{\partial \lambda_k^2} \det\{\mathbf{C}_k^s(\lambda_k^{\min})\} > 0 \ ,$$

this is the desired result. □

The special case of scalar states in the case of scalar measurements must be treated separately and yields simpler results.

THEOREM 7.3 *For the case $N = 1$, the scaling parameter λ_k^{min} for minimum variance of the filtering result is given by*

$$\lambda_k^{min} = \begin{cases} 0 & C_k^y \geq H_k^2 C_k^p \\ \infty & \text{otherwise.} \end{cases}$$

PROOF. The variance of the filtering result can be written as

$$C_k^s = \frac{C_k^p \frac{C_k^y}{H_k^2}(1+\lambda_k)}{\frac{C_k^y}{H_k^2} + \lambda_k C_k^p} \ .$$

The result follows by inspection. □

8 Conclusions

This article provides a self-contained derivation for both the prediction and filtering step for state estimation in the case of unknown correlations. The filtering step is based on existing ideas [1], but has been further extended. Moreover, an efficient algorithm for the time update step has been developed. For both the time update and the filtering step closed-form solutions for the calculation of minimum covariance estimates have been derived.

References

[1] Covariance Intersection Working Group, "A culminating advance in the theory and practice of data fusion, filtering, and decentralized estimation", Available at `http://www.ait.nrl.navy.mil/people/uhlmann/CovInt.html`, 1997.

[2] M. Csorba and H. Durrant-Whyte, "Map building using relative estimates", Proceedings of the 1997 SPIE Aerosense Conference, Orlando, Florida, 1997.

[3] S. Julier and J. Uhlmann, "A nondivergent estimation algorithm in the presence of unknown correlations", Proceedings of the 1997 American Control Conference (ACC'97), 1997.

[4] J. Uhlmann, S. Julier, and M. Csorba, "Nondivergent simultaneous map building and localization using covariance intersection", Proceedings of the 1997 SPIE Aerosense Conference, Orlando, Florida, 1997.

Appendix

If $\underline{a}$, $\underline{b}$ are jointly Gaussian with mean and covariance

$$\begin{bmatrix} \hat{\underline{a}} \\ \hat{\underline{b}} \end{bmatrix} \text{ and } \begin{bmatrix} \mathbf{C}_{aa} & \mathbf{C}_{ab} \\ \mathbf{C}_{ba} & \mathbf{C}_{bb} \end{bmatrix} ,$$

a given observation $\tilde{\underline{b}}$ yields the conditional (Gaussian) density of the random variable $\underline{a}$ conditioned on $\tilde{\underline{b}}$ according to

$$E\left\{\underline{a} \,|\, \underline{b} = \tilde{\underline{b}}\right\} = \hat{\underline{a}} + \mathbf{C}_{ab}\mathbf{C}_{bb}^{-1}\left(\tilde{\underline{b}} - \hat{\underline{b}}\right) \ ,$$

$$E\left\{\underline{a}\underline{a}^T \,|\, \underline{b} = \tilde{\underline{b}}\right\} = \mathbf{C}_{aa} - \mathbf{C}_{ab}\mathbf{C}_{bb}^{-1}\mathbf{C}_{ba} \ .$$