

Kalman Filtering with State Equality Constraints

DAN SIMON, Member, IEEE
Cleveland State University

TIEN LI CHIA, Member, IEEE
ControlSoft, Inc.

Kalman filters are commonly used to estimate the states of a dynamic system. However, in the application of Kalman filters there is often known model or signal information that is either ignored or dealt with heuristically. For instance, constraints on state values (which may be based on physical considerations) are often neglected because they do not fit easily into the structure of the Kalman filter. A rigorous analytic method of incorporating state equality constraints in the Kalman filter is developed here. The constraints may be time varying. At each time step the unconstrained Kalman filter solution is projected onto the state constraint surface. This significantly improves the prediction accuracy of the filter. The use of this algorithm is demonstrated on a simple nonlinear vehicle tracking problem.

Manuscript received June 29, 2000; accepted August 17, 2000; released for publication October 15, 2001.

IEEE Log No. T-AES/38/1/02578.

Refereeing of this contribution was handled by J. L. Leva.

Authors' current addresses: D. Simon, Dept. of Electrical and Computer Engineering, Fenn College of Engineering, 1960 East 24th St., Cleveland, OH 44115-2425, E-mail: (d.j.simon@csuohio.edu); T. L. Chia, ControlSoft Inc., Suite 200, 14077 Cedar Rd., Cleveland, OH 44118, E-mail: (tchia@controlsoftinc.com).

0018-9251/02/\$17.00 © 2002 IEEE

I. INTRODUCTION

For linear dynamic systems with white process and measurement noise, the Kalman filter is known to be an optimal estimator. In the application of Kalman filters there is often known model or signal information that is either ignored or dealt with heuristically [1]. This work presents a way to generalize the Kalman filter in such a way that known relations among the state variables (i.e., state constraints) are satisfied by the state estimate.

Some researchers have treated state constraints by reducing the system model parameterization [2, 3], but there are a couple of disadvantages with this approach. First, it may be desirable to maintain the form and structure of the state equations due to the physical meaning associated with each state. The reduction of the state equations makes their interpretation less natural and more difficult. Second, the equality constraint solution presented here can be extended to inequality constraints by checking the inequality constraints at each time step of the filter [4]. If the inequality constraints are satisfied at a given time step, then the inequality constrained problem is solved. If the inequality constraints are not satisfied, then the constrained solution presented here can be used to enforce the constraints.

Some researchers treat state constraints as perfect measurements [5, 6]. This results in a singular covariance matrix but does not present any theoretical problems [7]. In fact, Kalman's original paper [8] presents an example that uses perfect measurements (i.e., no measurement noise). However, there are a couple of considerations that indicate against the use of perfect measurements in a Kalman filter implementation. First of all, although the Kalman filter does not formally require a nonsingular covariance matrix, in practice a singular covariance increases the possibility of numerical problems [9, p. 249], [10, p. 365]. Secondly, the incorporation of state constraints as perfect measurements increases the dimension of the problem, which in turn increases the size of the matrix that needs to be inverted in the Kalman gain computation. For instance, if we have a Kalman filtering problem with m measurements, then we need to invert an $m \times m$ matrix in order to compute the Kalman gain. If in addition we have s state constraints that we treat as perfect measurements, then we need to instead invert an $(m + s) \times (m + s)$ matrix. An alternative method, a constrained Kalman filter, which incorporates the state constraints into the state estimation framework is proposed. The proposed solution does not have any numerical problems, does not increase the dimension of the problem, and has known and proven statistical properties. At each time step the unconstrained Kalman filter solution is projected onto the constraint surface. This is similar to

the solution of the constrained parameter identification problem [11].

Section II presents a brief summary of the Kalman filter solution without any state constraints. Section III shows three different ways that the Kalman filter solution can be projected onto the state constraint surface. Section IV develops a number of attractive theoretical properties of the constrained Kalman filter. Section V discusses nonlinear Kalman filtering in the presence of state constraints, and Kalman filtering in the presence of nonlinear state constraints. Section VI presents some simulation results, and Section VII closes with some concluding remarks and suggestions for further work.

II. UNCONSTRAINED STATE ESTIMATION

This section reviews unconstrained state estimation via the Kalman filter, along with some important properties of the Kalman filter that is used later in this paper. The results and notation are taken from [12].

Consider the discrete linear time-invariant system given by

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + w_k \\ y_k &= Cx_k + e_k \end{aligned} \quad (1)$$

where k is the time index, x is the state vector, u is the known control input, y is the measurement, and $\{w_k\}$ and $\{e_k\}$ are noise inputs. The problem is to find an estimate \hat{x}_{k+1} of x_{k+1} given the measurements $\{y_0, y_1, \dots, y_k\}$. We use the symbol Y_k to denote the column vector that contains the measurements $\{y_0, y_1, \dots, y_k\}$. We assume that the following conditions are satisfied:

$$E[x_0] = \bar{x}_0 \quad (2)$$

$$E[w_k] = 0 \quad (3)$$

$$E[e_k] = 0 \quad (4)$$

$$E[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T] = \Sigma_0 \quad (5)$$

$$E[w_k w_m^T] = Q\delta_{km} \quad (6)$$

$$E[e_k e_m^T] = R\delta_{km} \quad (7)$$

$$E[w_k e_m^T] = 0 \quad (8)$$

$$E[x_k e_m^T] = 0 \quad (9)$$

$$E[x_k w_m^T] = 0 \quad (10)$$

where $E[\cdot]$ is the expectation operator, \bar{x} is the expected value of x , and δ_{km} is the Kronecker delta function ($\delta_{km} = 1$ if $k = m$, 0 otherwise). Q and R are positive semidefinite covariance matrices. The Kalman filter equations are given by

$$K_k = A\Sigma_k C^T (C\Sigma_k C^T + R)^{-1} \quad (11)$$

$$\hat{x}_{k+1} = A\hat{x}_k + Bu_k + K_k(y_k - C\hat{x}_k) \quad (12)$$

$$\Sigma_{k+1} = (A\Sigma_k - K_k C\Sigma_k)A^T + Q. \quad (13)$$

Note that this is the prediction form of the Kalman filter equations, so x_k is estimated on the basis of measurements up to and including time $k-1$. The filter is initialized with $\hat{x}_0 = \bar{x}_0$, and Σ_0 given above. It can be shown [12] that the Kalman filter has several attractive properties. For instance, if x_0 , $\{w_k\}$, and $\{e_k\}$ are jointly Gaussian, the Kalman filter estimate \hat{x}_{k+1} is the conditional mean of x_{k+1} given the measurements Y_k ; i.e., $\hat{x}_{k+1} = E[x_{k+1} | Y_k]$. Even if x_0 , $\{w_k\}$, and $\{e_k\}$ are not jointly Gaussian, the Kalman filter estimate is the best affine estimator given the measurements Y_k ; i.e., of all estimates of x_{k+1} that are of the form $FY_k + g$ (where F is a fixed matrix and g is a fixed vector), the Kalman filter estimate is the one that minimizes the variance of the estimation error. It can be shown [12, pp. 92 ff.] that the Kalman filter estimate (i.e., the minimum variance estimate) can be given by

$$\hat{x}_{k+1} = \bar{\bar{x}}_{k+1} \equiv \bar{x}_{k+1} + \Sigma_{xy} \Sigma_{yy}^{-1} (Y_k - \bar{Y}_k) \quad (14)$$

where $\bar{\bar{x}}_{k+1}$ is the expected value of x_{k+1} , \bar{Y}_k is the expected value of Y_k , Σ_{xy} is the variance matrix of x_{k+1} and Y_k , Σ_{yy} is the covariance matrix of Y_k , and $\bar{\bar{x}}_{k+1}$ is the conditional mean of x_{k+1} given the measurements Y_k . In addition, from [12, p. 93] we know that the Kalman filter estimate \hat{x}_{k+1} and Y_k are jointly Gaussian, in which case \hat{x}_{k+1} is conditionally Gaussian given Y_k . The conditional probability density function of x_{k+1} given Y_k is

$$P(x | Y) = \frac{\exp[-(x - \bar{\bar{x}})^T \Sigma^{-1} (x - \bar{\bar{x}}) / 2]}{(2\pi)^{n/2} |\Sigma|^{1/2}} \quad (15)$$

where n is the dimension of x and

$$\Sigma = \Sigma_{xx} - \Sigma_{xy} \Sigma_{yy}^{-1} \Sigma_{yx} \quad (16)$$

where Σ_{xx} is the covariance matrix of x_k . The Kalman filter estimate is that value of x that maximizes the conditional probability density function $P(x | Y)$, and Σ is the covariance of the Kalman filter estimate.

III. CONSTRAINED STATE ESTIMATION

This section extends the well-known results of the previous section to cases where there are known relationships among the state components. That is, given the dynamic system of (1), we are given the additional constraint

$$Dx_k = d_k \quad (17)$$

where D is a known $s \times n$ constant matrix, d_k is a known $s \times 1$ vector, s is the number of constraints, n is the number of states, and $s \leq n$. It is assumed in this paper that D is full rank, i.e., that D has rank s . This is an easily satisfied assumption. If D is not full rank that means we have redundant state constraints. In that case we can simply remove linearly dependent rows

from D (i.e., remove redundant state constraints) until D is full rank.

A weaker but more general constraint may exist on the expected value of the state

$$DE[x_k] = d_k. \quad (18)$$

For both constraints (17) and (18), we can require that the constrained state estimate \tilde{x}_k satisfy the constraint at each point in time. The derivation for the Kalman filter with constraint (17) is the same as the derivation with constraint (18), but the statistical properties of the two estimates are different [13]. Here we concentrate on the Kalman filter with constraint (17).

Three different approaches to the constrained state estimation problem are given in this section. The time index k is omitted in the remainder of this section for ease of notation.

A. Maximum Probability Method

In this section we derive the constrained Kalman filter by using the maximum probability method of the Kalman filter. From [12, pp. 93 ff.] we know that the Kalman filter estimate is that value of x that maximizes the conditional probability density function $P(x | Y)$, which is given in (15). The constrained Kalman filter can be derived by finding an estimate \tilde{x} such that the conditional probability $P(\tilde{x} | Y)$ is maximized and \tilde{x} satisfies the constraint (17). Maximizing $P(\tilde{x} | Y)$ is the same as maximizing its natural logarithm. So the problem we want to solve can be given by

$$\begin{aligned} \max \ln P(\tilde{x} | Y) &\implies \min(\tilde{x} - \bar{x})^T \Sigma^{-1} (\tilde{x} - \bar{x}) \\ \text{such that } D\tilde{x} &= d. \end{aligned} \quad (19)$$

To solve this problem we form the Lagrangian

$$L = (\tilde{x} - \bar{x})^T \Sigma^{-1} (\tilde{x} - \bar{x}) + 2\lambda^T (D\tilde{x} - d). \quad (20)$$

We formulate the first-order conditions necessary for a minimum as

$$\frac{\partial L}{\partial \tilde{x}} = 0 \implies \Sigma^{-1} (\tilde{x} - \bar{x}) + D^T \lambda = 0 \quad (21)$$

$$\frac{\partial L}{\partial \lambda} = 0 \implies d = D\tilde{x}. \quad (22)$$

This gives the solution

$$\lambda = (D\Sigma D^T)^{-1} (D\bar{x} - d) \quad (23)$$

$$\tilde{x} = \bar{x} - \Sigma D^T (D\Sigma D^T)^{-1} (D\bar{x} - d). \quad (24)$$

Since \bar{x} (the conditional mean of x) is the Kalman filter estimate, the constrained Kalman estimate \tilde{x} can be derived from the unconstrained estimate as

$$\tilde{x} = \hat{x} - \Sigma D^T (D\Sigma D^T)^{-1} (D\hat{x} - d). \quad (25)$$

Note that this constrained estimate depends on the conditional Gaussian nature of \hat{x} , which in turn depends on the Gaussian nature of x_0 , $\{w_k\}$, and $\{e_k\}$ in (1). This derivation is a generalization of the approach taken in [14, pp. 166 ff.] for the integration of robot sensor measurements in a partially known environment. The present paper is based on research in [13] that was conducted prior to [14].

B. Mean Square Method

In this section we derive the constrained Kalman filter by using a mean square minimization method. We seek to minimize the conditional mean square error subject to the state constraints.

$$\min_{\tilde{x}} E(\|x - \tilde{x}\|^2 | Y) \quad \text{such that } D\tilde{x} = d \quad (26)$$

where $\|\cdot\|$ denotes the vector two-norm. If we assume that x and Y are jointly distributed, the mean square error can be written as

$$\begin{aligned} E(\|x - \tilde{x}\|^2 | Y) &= \int (x - \tilde{x})^T (x - \tilde{x}) P(x | Y) dx \quad (27) \\ &= \int x^T x P(x | Y) dx \\ &\quad - 2\tilde{x}^T \int x P(x | Y) dx + \tilde{x}^T \tilde{x}. \end{aligned} \quad (28)$$

The Lagrangian of the constrained problem is

$$L = E(\|x - \tilde{x}\|^2 | Y) + 2\lambda^T (D\tilde{x} - d) \quad (29)$$

$$\begin{aligned} &= \int x^T x P(x | Y) dx - 2\tilde{x}^T \int x P(x | Y) dx \\ &\quad + \tilde{x}^T \tilde{x} + 2\lambda^T (D\tilde{x} - d). \end{aligned} \quad (30)$$

Noting that the conditional mean of x can be written as

$$\hat{x} = \int x P(x | Y) dx \quad (31)$$

we formulate the first order conditions necessary for a minimum as

$$\frac{\partial L}{\partial \tilde{x}} = 0 \implies -2\hat{x} + 2\tilde{x} + 2D^T \lambda = 0 \quad (32)$$

$$\frac{\partial L}{\partial \lambda} = 0 \implies D\tilde{x} - d = 0. \quad (33)$$

This gives the solution

$$\lambda = (DD^T)^{-1} (D\hat{x} - d) \quad (34)$$

$$\tilde{x} = \hat{x} - D^T (DD^T)^{-1} (D\hat{x} - d). \quad (35)$$

Note that the \hat{x} given in (31) is the Kalman filter estimate of x only if x_0 , $\{w_k\}$, and $\{e_k\}$ are jointly Gaussian. However, even if they are not jointly Gaussian, we can still use the conditional mean given in (31) to define an unconstrained estimate of x ,

which in turn leads to the constrained estimate given in (35).

C. Projection Method

In this section we derive the constrained Kalman filter by directly projecting the unconstrained state estimate \hat{x} onto the constraint surface. That is, we solve the problem

$$\min_{\tilde{x}} (\tilde{x} - \hat{x})^T W (\tilde{x} - \hat{x}) \quad \text{such that} \quad D\tilde{x} = d \quad (36)$$

where W is any symmetric positive definite weighting matrix. The solution of this problem is given by

$$\tilde{x} = \hat{x} - W^{-1} D^T (D W^{-1} D^T)^{-1} (D \hat{x} - d). \quad (37)$$

The constrained estimates derived by the maximum probability method (24) and the mean square method (35) can be obtained from this equation by setting $W = \Sigma^{-1}$ and $W = I$, respectively. Note that this derivation of the constrained estimate does not depend on the conditional Gaussian nature of \hat{x} ; i.e., x_0 , $\{w_k\}$, and $\{e_k\}$ in (1) are not assumed to be Gaussian.

If D is a square matrix (i.e., the number of constraints is equal to the number of states) then the state is fully constrained. In this case, remembering our assumption from the beginning of this section that D is full rank, (37) reduces to

$$\begin{aligned} \tilde{x} &= \hat{x} - W^{-1} D^T D^{-T} W D^{-1} (D \hat{x} - d) \\ &= D^{-1} d. \end{aligned} \quad (38)$$

IV. PROPERTIES OF CONSTRAINED STATE ESTIMATE

In this section we study some of the statistical properties of the constrained Kalman filter. We use \hat{x} to denote the state estimate of the unconstrained Kalman filter, and \tilde{x} to denote the state estimate of the constrained Kalman filter as given by (37), recalling that (24) and (35) are special cases of (37). In the development below, if A and B are square matrices with the same dimension, we use the notation $A > B$ to indicate that $A - B$ is positive definite, and $A \geq B$ to indicate that $A - B$ is positive semidefinite.

THEOREM 1 *The constrained state estimate \tilde{x} as given by (37) is an unbiased state estimator for the system (1) subject to the constraint (17) for any known symmetric positive definite weighting matrix W . That is,*

$$E(\tilde{x}) = E(x). \quad (39)$$

PROOF We know from (37) that for any symmetric positive definite weighting matrix W

$$x - \tilde{x} = x - \hat{x} + W^{-1} D^T (D W^{-1} D^T)^{-1} [D \hat{x} - d - (D x - d)] \quad (40)$$

$$= x - \hat{x} + W^{-1} D^T (D W^{-1} D^T)^{-1} (D \hat{x} - D x) \quad (41)$$

$$= (I - W^{-1} D^T (D W^{-1} D^T)^{-1} D) (x - \hat{x}). \quad (42)$$

If W is a known matrix that is independent of x , we can take the expectation of both sides to obtain

$$E(x - \tilde{x}) = (I - W^{-1} D^T (D W^{-1} D^T)^{-1} D) E(x - \hat{x}). \quad (43)$$

But the right side of the above equation is zero, because we know that $E(\hat{x}) = E(x)$ (i.e., the unconstrained Kalman filter provides an unbiased estimate of x). Therefore the left side of the above equation is also zero, which implies that the constrained state estimate \tilde{x} is an unbiased estimate of x . Q.E.D.

THEOREM 2 *The constrained state estimate \tilde{x} as given by (37) with $W = \Sigma^{-1}$, where Σ is the covariance of the unconstrained estimate given in (13) and (16), has a smaller error covariance than that of the unconstrained state estimate. That is,*

$$\text{Cov}(x - \tilde{x}) < \text{Cov}(x - \hat{x}). \quad (44)$$

PROOF From (42) we see that if $W = \Sigma^{-1}$ then

$$x - \tilde{x} = (I - \Sigma D^T (D \Sigma D^T)^{-1} D) (x - \hat{x}) \quad (45)$$

$$= (I - J) (x - \hat{x}) \quad (46)$$

where J is defined by the above equation; i.e., $J = \Sigma D^T (D \Sigma D^T)^{-1} D$. So the covariance of the constrained estimation error is given by

$$\text{Cov}(x - \tilde{x}) = E[(x - \tilde{x})(x - \tilde{x})^T] \quad (47)$$

$$= E\{[(I - J)(x - \hat{x})][(I - J)(x - \hat{x})]^T\} \quad (48)$$

$$= (I - J) \text{Cov}(x - \hat{x}) (I - J)^T \quad (49)$$

$$= \Sigma - J \Sigma - \Sigma J^T + J \Sigma J^T. \quad (50)$$

But it can be easily shown from the definition of J that $\Sigma J^T = J \Sigma J^T$. Therefore the previous equation reduces to

$$\text{Cov}(x - \tilde{x}) = \Sigma - J \Sigma. \quad (51)$$

We know that $J \Sigma$ is positive definite [12, p. 335]. (Here we are assuming that D is full rank as stated at the beginning of Section III.) Therefore we conclude that

$$\text{Cov}(x - \tilde{x}) - \Sigma = -J \Sigma < 0 \quad (52)$$

which means that

$$\text{Cov}(x - \tilde{x}) < \Sigma. \quad (53)$$

In other words, the covariance of the error of the constrained state estimate \tilde{x} with $W = \Sigma^{-1}$, where Σ is the covariance of the unconstrained estimate, is smaller than Σ . Q.E.D.

THEOREM 3 *Among all the constrained Kalman filters of (37), the filter that uses $W = \Sigma^{-1}$ has the smallest estimation error covariance.*

PROOF We use the symbol Λ_W to indicate the estimation error covariance of the constrained Kalman filter of (37). That is,

$$\Lambda_W = E[(x - \tilde{x})(x - \tilde{x})^T] \quad (54)$$

where \tilde{x} is given by (37). We can use (42) to derive

$$\begin{aligned} \Lambda_W &= (I - W^{-1}D^T(DW^{-1}D^T)^{-1}D) \\ &\quad \times \Sigma(I - W^{-1}D^T(DW^{-1}D^T)^{-1}D)^T \\ &= \Sigma - W^{-1}D^T(DW^{-1}D^T)^{-1}D\Sigma \\ &\quad - \Sigma D^T(DW^{-1}D^T)^{-1}DW^{-1} \\ &\quad + W^{-1}D^T(DW^{-1}D^T)^{-1}D\Sigma D^T(DW^{-1}D^T)^{-1}DW^{-1}. \end{aligned} \quad (55)$$

If we replace W with Σ^{-1} in the above equation we obtain

$$\Lambda_{\Sigma^{-1}} = \Sigma - \Sigma D^T(D\Sigma D^T)^{-1}D\Sigma. \quad (57)$$

Taking the difference of the two preceding equations gives

$$\Lambda_{\Sigma^{-1}} - \Lambda_W = -M\Sigma M^T \quad (58)$$

where $M = \Sigma D^T(D\Sigma D^T)^{-1}D - W^{-1}D^T(DW^{-1}D^T)^{-1}D$. We know from [12, p. 335] that $M\Sigma M^T$ is positive semidefinite for all M , which gives

$$\Lambda_{\Sigma^{-1}} \leq \Lambda_W. \quad (59)$$

Thus we have shown that $W = \Sigma^{-1}$ in (37) gives the constrained state estimate with the minimum error covariance. Q.E.D.

THEOREM 4 *The constrained Kalman filter estimate \tilde{x} of (37) with $W = I$ satisfies the inequality*

$$\|x - \tilde{x}\| \leq \|x - \hat{x}\| \quad (60)$$

where $\|\cdot\|$ is the vector two-norm and \hat{x} is the unconstrained Kalman filter estimate.

PROOF Since D is assumed to be full rank, the matrix $D^T(DD^T)^{-1}D$ is an $s \times n$ matrix with rank s . The range space of $D^T(DD^T)^{-1}D$ is s -dimensional, and the null space of $D^T(DD^T)^{-1}D$ is $(n-s)$ -dimensional. Now note that the unconstrained and constrained estimates can be written as

$$\hat{x} = x - \hat{\alpha} \quad (61)$$

$$\tilde{x} = x - \tilde{\alpha} \quad (62)$$

where $\hat{\alpha}$ is the unconstrained estimation error and $\tilde{\alpha}$ is the constrained estimation error. $\hat{\alpha}$ is an n -dimensional vector which can be written as

$$\hat{\alpha} = \hat{\alpha}_1 + \dots + \hat{\alpha}_s + \hat{\alpha}_{s+1} + \dots + \hat{\alpha}_n \quad (63)$$

where the $\hat{\alpha}_i$ s are all orthogonal to each other, $\{\hat{\alpha}_1, \dots, \hat{\alpha}_s\}$ are in the range space of $D^T(DD^T)^{-1}D$, and $\{\hat{\alpha}_{s+1}, \dots, \hat{\alpha}_n\}$ are in the null space of

$D^T(DD^T)^{-1}D$. From (37) we see that, if $W = I$,

$$\tilde{x} = \hat{x} - D^T(DD^T)^{-1}(D\hat{x} - d). \quad (64)$$

Subtracting both sides from x and substituting (17) for d gives

$$x - \tilde{x} = x - \hat{x} + D^T(DD^T)^{-1}(D\hat{x} - Dx) \quad (65)$$

$$\tilde{\alpha} = \hat{\alpha} - D^T(DD^T)^{-1}D\hat{\alpha}. \quad (66)$$

But the last term on the right side of the above equation is just the projection of $\hat{\alpha}$ onto the range space of $D^T(DD^T)^{-1}D$. Thus we can use (63) to obtain

$$\tilde{\alpha} = \hat{\alpha}_{s+1} + \dots + \hat{\alpha}_n \quad (67)$$

$$\|\tilde{\alpha}\|^2 = \|\hat{\alpha}_{s+1}\|^2 + \dots + \|\hat{\alpha}_n\|^2 \quad (68)$$

where the second equality follows from the orthogonality of the $\hat{\alpha}_i$ s. But (63) gives

$$\|\hat{\alpha}\|^2 = \|\hat{\alpha}_1\|^2 + \dots + \|\hat{\alpha}_n\|^2. \quad (69)$$

Combining the two preceding equations gives

$$\|\tilde{\alpha}\|^2 = \|\hat{\alpha}\|^2 - \|\hat{\alpha}_1\|^2 - \dots - \|\hat{\alpha}_s\|^2 \quad (70)$$

$$\|x - \tilde{x}\| \leq \|x - \hat{x}\|. \quad (71)$$

We have thus proved the inequality of the theorem. Note that equality holds only if

$$\hat{\alpha}_1 = \dots = \hat{\alpha}_s = 0 \quad (72)$$

which means that the unconstrained estimation error lies entirely in the null space of $D^T(DD^T)^{-1}D$, which in turn indicates that the unconstrained state estimate satisfies the state constraint. Q.E.D.

THEOREM 5 *The error of the constrained Kalman filter estimate of (37) with $W = I$ is smaller than the unconstrained estimation error in the sense that*

$$\text{Tr}[\text{Cov}(\tilde{x})] \leq \text{Tr}[\text{Cov}(\hat{x})] \quad (73)$$

where $\text{Tr}[\cdot]$ indicates the trace of a matrix, and $\text{Cov}(\cdot)$ indicates the covariance matrix of a random vector.

PROOF We can write the covariance of the unconstrained estimation error as

$$\text{Cov}(\hat{x}) = E[(\hat{x} - \bar{x})(\hat{x} - \bar{x})^T] \quad (74)$$

$$= E[(x - \hat{\alpha} - \bar{x} + \bar{\alpha})(x - \hat{\alpha} - \bar{x} + \bar{\alpha})^T] \quad (75)$$

$$= E(xx^T + \hat{\alpha}\hat{\alpha}^T - x\bar{x}^T + \bar{x}\bar{x}^T - \bar{x}x^T) \quad (76)$$

$$= E[(x - \bar{x})(x - \bar{x})^T] + E(\hat{\alpha}\hat{\alpha}^T) \quad (77)$$

where we have simplified the expression by using the fact that the unconstrained estimation error $\hat{\alpha}$ is zero-mean and uncorrelated with the state [12]. Now, for a general random vector z we can write [12, p. 93]

$$E(\|z\|^2) = E[\text{Tr}(zz^T)] = \text{Tr}[\text{Cov}(z)] + \text{Tr}(\bar{z}\bar{z}^T). \quad (78)$$

We can use this equation and (77) to write

$$\text{Tr}[\text{Cov}(\hat{x})] = \text{Tr}[\text{Cov}(x)] + E(\|\hat{\alpha}\|^2) \quad (79)$$

$$= \text{Tr}[\text{Cov}(x)] + \sum_{i=1}^n E(\hat{\alpha}_i^2) \quad (80)$$

where the $\hat{\alpha}_i$ s are the same as those given in (63). In a similar manner we can show that

$$\text{Tr}[\text{Cov}(\tilde{x})] = \text{Tr}[\text{Cov}(x)] + E(\|\tilde{\alpha}\|^2). \quad (81)$$

We can now use (67) to write

$$\text{Tr}[\text{Cov}(\tilde{x})] = \text{Tr}[\text{Cov}(x)] + \sum_{i=s+1}^n E(\hat{\alpha}_i^2). \quad (82)$$

Combining (80) and (82) gives the desired result

$$\text{Tr}[\text{Cov}(\tilde{x})] \leq \text{Tr}[\text{Cov}(\hat{x})]. \quad (83)$$

Note that equality holds only if

$$\hat{\alpha}_1 = \dots = \hat{\alpha}_s = 0 \quad (84)$$

which means that the unconstrained estimation error lies entirely in the null space of $D^T(DD^T)^{-1}D$, which in turn indicates that the unconstrained state estimate satisfies the state constraint. Q.E.D.

V. NONLINEARITIES

The development of the preceding sections has been restricted to linear systems (1) and linear state constraints (17). The constrained Kalman filter can be applied to nonlinear systems and nonlinear state constraints by linearizing about the current estimate. This is analogous to the extension of the standard Kalman filter to nonlinear systems via linearization of the system around the estimated state.

Consider the discrete nonlinear time-invariant system given by

$$\begin{aligned} x_{k+1} &= f(x_k) + b(u_k) + w_k \\ y_k &= h(x_k) + e_k \end{aligned} \quad (85)$$

where $f(\cdot)$, $b(\cdot)$, and $h(\cdot)$ are arbitrary nonlinear functions. As before, $\{w_k\}$ and $\{e_k\}$ are assumed to be white noise with covariance matrices Q and R , respectively. If this nonlinear system is linearized about the current state estimate \hat{x} , we obtain the extended Kalman filter [12, p. 195].

$$K_k = \Sigma_k H_k (H_k^T \Sigma_k H_k + R)^{-1} \quad (86)$$

$$\hat{x}_{k+1} = f(\hat{x}_k + K_k(y_k - h(\hat{x}_k))) + b(u_k) \quad (87)$$

$$\Sigma_{k+1} = F_k(I - K_k H_k^T) \Sigma_k F_k^T + Q \quad (88)$$

where H_k is the partial derivative of $h(\cdot)$ with respect to x evaluated at \hat{x}_k , and F_k is the partial derivative of $f(\cdot)$ with respect to x evaluated at \hat{x}_k . In this case we cannot make very many claims about the optimality

of the filter, but in practice the extended Kalman filter works quite nicely.

Now consider the nonlinear system of (85) with the state constraint of (17). The development of the constrained state estimator as given in the preceding sections is still valid. Dropping the k subscript for ease of notation, we obtain the constrained state estimate

$$\tilde{x} = \hat{x} - W^{-1}D^T(DW^{-1}D^T)^{-1}(D\hat{x} - d) \quad (89)$$

where W is any symmetric positive definite matrix. For linear systems, $W = \Sigma^{-1}$ gives the best constrained estimator, but we cannot make that claim for the nonlinear estimator.

Now consider the nonlinear state constraint

$$g(x_k) = d_k. \quad (90)$$

Dropping the k subscript for ease of notation, we note that this can be linearized about the current constrained state estimate \tilde{x} as

$$g(\tilde{x}) + g'(\tilde{x})(x - \tilde{x}) \approx d \quad (91)$$

which indicates that

$$g'(\tilde{x})x \approx d - g(\tilde{x}) + g'(\tilde{x})\tilde{x}. \quad (92)$$

So now we have an approximately linear constraint that is of the form of (17) where D is replaced with $g'(\tilde{x})$ and d is replaced with $d - g(\tilde{x}) + g'(\tilde{x})\tilde{x}$.

VI. SIMULATION RESULTS

In this section we present a simple example to illustrate the efficacy of the constrained Kalman filter. Consider a land-based vehicle that is equipped to measure its range relative to two reference points, (r_{n1}, r_{e1}) and (r_{n2}, r_{e2}) , where each reference point is specified by its northerly and easterly positions. The vehicle dynamics and measurements can be approximated by the equations

$$\begin{aligned} x_{k+1} &= \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} x_k + \begin{bmatrix} 0 \\ 0 \\ T \sin \theta \\ T \cos \theta \end{bmatrix} u_k + w_k \\ y_k &= \begin{bmatrix} (x_1 - r_{n1})^2 + (x_2 - r_{e1})^2 \\ (x_1 - r_{n2})^2 + (x_2 - r_{e2})^2 \end{bmatrix} + e_k \end{aligned}$$

where the first two elements of x are the northerly and easterly positions, the last two elements of x are the northerly and easterly velocities, w represents process disturbances due to potholes and the like, e represents measurement errors, and u is the commanded acceleration. T is the sample period of the position estimator, θ is the heading angle (measured counterclockwise from due east), and e is the measurement error. The dynamic equations are

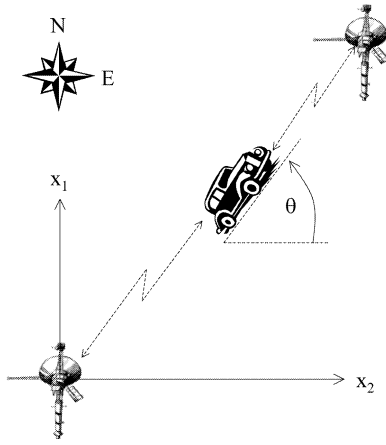


Fig. 1. Land vehicle and 2 transponders.

linear but the measurement equations are nonlinear, so the extended Kalman filter of Section V can be used to estimate the state vector. The navigation reference points are at $(0,0)$ and $(173210,100000)$ meters, while the covariances of the process and measurement noise are

$$Q = \text{Diag}(4 \text{ m/s}, 4 \text{ m/s}, 1 \text{ m/s}^2, 1 \text{ m/s}^2)$$

$$R = \text{Diag}(900 \text{ m}^2, 900 \text{ m}^2).$$

We can use a Kalman filter to estimate the position of the vehicle. During certain times the vehicle may be travelling off-road, or on an unknown road, in which case the problem is unconstrained. At other times it may be known that the vehicle is travelling on a given road, in which case the state estimation problem is constrained. For instance, if it is known that the vehicle is travelling on a road with a heading of θ then the matrix D and the vector d of (17) can be given by

$$D = \begin{bmatrix} 1 & -\tan\theta & 0 & 0 \\ 0 & 0 & 1 & -\tan\theta \end{bmatrix}$$

$$d = [0 \quad 0]^T$$

The sample period T is 3 s and the heading θ is set to a constant 60 deg. The commanded acceleration is alternately set to $\pm 1 \text{ m/s}^2$, as if the vehicle was alternately accelerating and decelerating in traffic. Note that with the 60 deg heading, the vehicle and the two reference points form a straight line, which makes the state estimation problem more difficult. This scenario is depicted in Fig. 1. The initial conditions are set to

$$x = \hat{x}_0 = [0 \quad 0 \quad 17 \quad 10]^T$$

$$P_0 = \text{Diag}[900 \quad 900 \quad 4 \quad 4]^T.$$

The unconstrained and constrained Kalman filters were simulated using MATLAB for 300 s. The figures in this paper show typical simulation results. Fig. 2

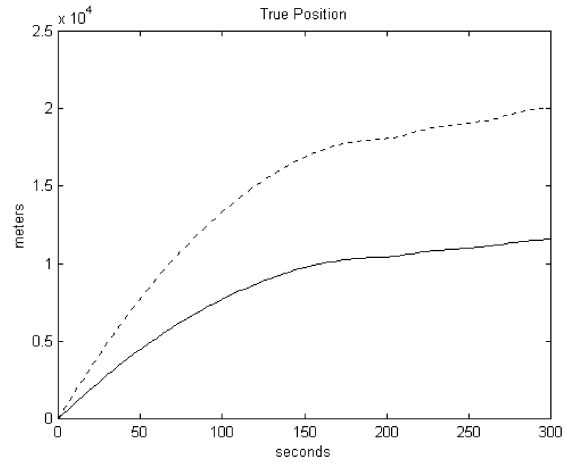


Fig. 2. True position. Dotted line is north position, solid line is east position.

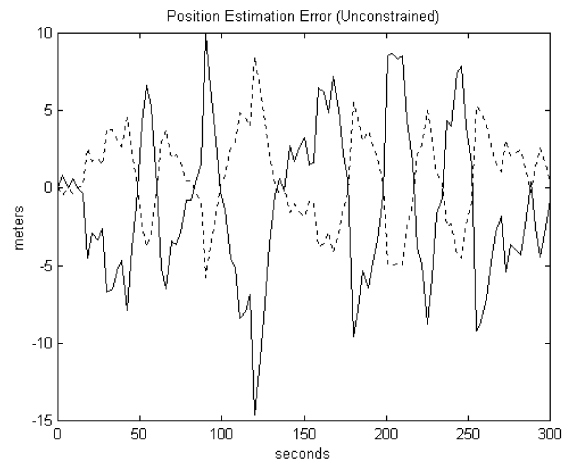


Fig. 3. Unconstrained filter position error. Dotted line is north position, solid line is east position.

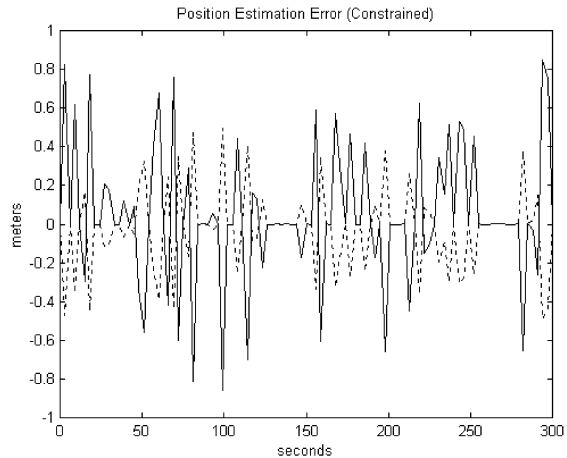


Fig. 4. Constrained filter position error ($W = I$). Dotted line is north position, solid line is east position.

shows the true vehicle position. Fig. 3 shows the position estimation error of the unconstrained Kalman filter, and Fig. 4 shows the position estimation error of the constrained Kalman filter. It can be seen

that the constrained filter results in much more accurate estimates than the unconstrained filter. The unconstrained filter results in average position errors of about 5 m, while the constrained filter results in position errors of about 0.2 m. In addition, for this particular example, the Kalman filter performs identically whether we incorporate state constraints using perfect measurements, or whether we incorporate state constraints using the approach presented here with either $W = I$ or $W = \Sigma^{-1}$. A MATLAB m-file that implements the algorithms in this paper and that was used to produce these simulation results can be downloaded from the World Wide Web page <http://academic.csuohio.edu/simond/kalmanconstrained/>.

VII. CONCLUSION

We have presented a method for incorporating linear state constraints in a Kalman filter. Simulation results demonstrate the effectiveness of this method. If the state constraints are nonlinear they can be linearized, although this may result in convergence problems. Further work along these lines could focus on combining our work with [7] in order to guarantee convergence in the presence of nonlinear constraints. Other efforts could explore the incorporation of state constraints for optimal smoothing, or the use of state constraints in H_∞ filtering [15].

Various applications of this work are possible. For example, this work has been extended to state inequality constraints along the lines of [16] and applied to health parameter estimation for aircraft engines [4]. Another possible aerospace application might be the estimation of navigation parameters for an aircraft where it is known from physical laws that the vehicle is operating within some known flight envelope. A fuzzy logic related application of this method is the extension of [17, 18] to optimize fuzzy logic systems in such a way that the membership functions remain sum normal [19].

REFERENCES

- [1] Massicotte, D., Morawski, R., and Barwicz, A. (1995) Incorporation of a positivity constraint into a Kalman-filter-based algorithm for correction of spectrometric data. *IEEE Transactions on Instrumentation and Measurement*, **44**, 1 (Feb. 1995), 2–7.
- [2] Wen, W., and Durrant-Whyte, H. (1991) Model based active object localisation using multiple sensors. Intelligent Systems and Robotics, Osaka, Japan, Nov. 1991.
- [3] Wen, W., and Durrant-Whyte, H. (1992) Model-based multi-sensor data fusion. In *Proceedings of the IEEE International Conference on Robotics and Automation*, Nice, France, May 1992, 1720–1726.
- [4] Simon, D., and Simon, D. L. Kalman filtering with inequality constraints for turbofan health estimation. Submitted for publication.
- [5] Hayward, S. (1998) Constrained Kalman filter for least-squares estimation of time-varying beamforming weights. In *Mathematics in Signal Processing IV* (J. McWhirter and I. Proudler (Eds.)). New York: Oxford University Press, 1998, pp. 113–125.
- [6] Porrill, J. (1988) Optimal combination and constraints for geometrical sensor data. *International Journal of Robotics Research*, **7**, 6 (Dec. 1988), 66–77.
- [7] De Geeter, J., Van Brussel, H., and De Schutter, J. (1997) A smoothly constrained Kalman filter. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **19**, 10 (Oct. 1997), 1171–1177.
- [8] Kalman, R. (1960) A new approach to linear filtering and prediction problems. *Transactions of the ASME Journal of Basic Engineering* (Mar. 1960), 35–45.
- [9] Maybeck, P. (1979) *Stochastic Models, Estimation, and Control*, Vol. 1. New York: Academic Press, 1979.
- [10] Stengel, R. (1994) *Optimal Control and Estimation*. New York: Dover Publications, 1994.
- [11] Chia, T., Chow, P., and Chizek, H. (1991) Recursive parameter identification of constrained systems: An application to electrically stimulated muscle. *IEEE Transactions on Biomedical Engineering*, **38**, 5 (May 1991), 429–442.
- [12] Anderson, B., and Moore, J. (1979) *Optimal Filtering*. Englewood Cliffs, NJ: Prentice Hall, 1979.
- [13] Chia, T. (1985) Parameter identification and state estimation of constrained systems. Ph.D. dissertation, Case Western Reserve University, Cleveland, OH, 1985.
- [14] Durrant-Whyte, H. (1988) *Integration, Coordination and Control of Multi-Sensor Robot Systems*. Boston: Kluwer Academic Publishers, 1988.
- [15] Simon, D., and El-Sherief, H. (1996) Hybrid Kalman/minimax filtering in phase-locked loops. *Control Engineering Practice*, **4**, 5 (Oct. 1996), 615–623.
- [16] Timmons, W., Chizek, H., Casas, F., Chankong, V., and Katona, P. (1997) Parameter-constrained adaptive control. *Industrial Engineering Chemical Research*, **36**, 11 (Nov. 1997), 4894–4905.
- [17] Simon, D. (2000) Fuzzy membership function optimization via the extended Kalman filter. In *Proceedings of the North American Fuzzy Information Processing Society Conference*, Atlanta, GA, July 2000, 311–315.
- [18] Simon, D. Training fuzzy systems with the extended Kalman filter. *Fuzzy Sets and Systems*, to be published.
- [19] Simon, D. Sum normal optimization of fuzzy membership functions. Submitted for publication.



Dan Simon (S'89—M'90—SM'01) received his B.S., M.S., and Ph.D. degrees from Arizona State University, Tempe, the University of Washington, Seattle, and Syracuse University, Syracuse, NY, all in electrical engineering.

He has 14 years of industrial experience in the aerospace, automotive, biomedical, process control, and software engineering fields. He has been an Assistant Professor at Cleveland State University since 1999, where he teaches and conducts research in the fields of control theory, computational intelligence, and embedded systems.

Dr. Simon has over 30 publications in refereed journals and conference proceedings.

Tien Li Chia (S'84—M'85) received the B.S. degree from Tan Kung University, Taipei, Taiwan, ROC, in 1977 and the M.S. and Ph.D. degrees in systems and control engineering from Case Western Reserve University, Cleveland, OH, in 1982 and 1985, respectively.

He is the President of ControlSoft, Inc., and adjunct faculty at Cleveland State University, Cleveland, OH, in the Electrical Engineering and Computer Engineering Department. He is experienced in the design and development of advanced control-based control solutions, including adaptive control, expert-based solutions, and multivariable control, for various industrial and manufacturing processes. At ControlSoft, he has invented, designed, and directed the development of many control software solutions that are now widely used in industry, of which the core control technologies have also been licensed by many leading control system vendors.

Dr. Chia has been invited to speak in several ISA and AIChE meetings and to lecture in some industrial short courses. He is the author of more than a dozen advanced control related articles that have been published in both academic and industrial journals.

