STATE ESTIMATION OF LINEAR SYSTEMS WITH STATE EQUALITY CONSTRAINTS

Sangho Ko , Robert R. Bitmead [1]

Department of Mechanical and Aerospace Engineering, University of California, San Diego, 9500 Gilman Drive, La Jolla, California, 92093-0411, USA.
{sko,rbitmead}@ucsd.edu

Abstract: This paper deals with the state estimation problem for linear systems with state equality constraints. Using noisy measurements which are available from the observable system, we construct the optimal estimate which also satisfies linear equality constraints. For this purpose, after reviewing modeling problems in linear stochastic systems with state equality constraints, we formulate a projected system representation. By using the constrained Kalman predictor for the projected system and comparing its predictor Ricccati Equation with those of the unconstrained and the projected Kalman predictors, we reach the conclusion that the current constrained estimator outperforms other filters for estimating the constrained system.

Keywords: Estimation, Constraints, Kalman filters, Projection, Descriptor systems

1. INTRODUCTION

Control with constraints is increasingly applied in industry (Maciejowski, 2002) and the practicality of incorporating both state and input constraints into control problems via MPC (Model Predictive Control) methods is in large part responsible for the recent upsurge in interest in these latter methods. Using explicit constraints, in place of their implicit inclusion using penalty and barrier methods, simplifies the design specification to focus on the performance objective. Constraints in control, particularly optimal control, have a long history (Bryson and Ho, 1975) and have focused on full state feedback systems. The constraints can be of two basic types: *physical constraints* reflecting known limits to physical state variables, such as positivity of mass or pressure; and *design constraints* which represent desired operating limits which might otherwise be violated by the controlled system. Forbidden, as opposed to undesirable, state motions may be incorporated into system descriptions through descriptor system representations.

Estimation of systems having constraints has drawn many practitioners' attention in diverse engineering disciplines: radio surgery (Altman and Tombropoulos, 1994), attitude determination of spacecraft (Lefferts *et al.*, 1982; Peng *et al.*, 2000; Chiang *et al.*, 2002), robotic systems for multi-sensor data fusion (Wen and Durrant-Whyte, 1992) and locating objects (Geeter *et al.*, 1997), target tacking of air vehicles (Tahk and Speyer, 1990; Alouani and Blair, 1993) and land-vehicles (Simon and Chia, 2002), and adaptive beamforming (Chen and Chiang, 1993). Each of these works is an example of dealing with physical constraints in the form of *equality* state constraints in a probabilistic framework. In the mainstream MPC literature, Rao *et al.* (2003) has developed a deterministic constrained state estimate which is based on a finite-horizon optimization. This deterministic approach does not yield an estimate quality measure of closeness to the actual state.

[1] Research supported by USA National Science Foundation Grants ECS-0200449 and ECS-0225530.

The problem which we consider here is the concomitant state estimation problem in constrained systems. Noisy measurements are available from the observable system and we desire to use these optimally to reconstruct a state estimate, which also is known to satisfy linear equality constraints. For this purpose, we formulate linear stochastic systems with linear equality constraints firstly in a stochastic descriptor system representation from which the projected system representation will be derived. Then, we construct the Kalman filter for the projected system and show that the current constrained estimator outperforms the other estimators available for constrained systems.

When some relations between state variables are known exactly, this type of constraint is called a *hard* or *strong constraint* and can be incorporated into the conventional Kalman filter through *perfect* pseudo-measurements. But this case results in a singular measurement noise covariance increasing the possibility of numerical problems (Stengel, 1994). For linear equality state constraints, it is always possible to reduce the system model parametrization and use the reduced state equation and the conventional Kalman filter. On the other hand, there are several good reasons for not using a reduced state space for treating the constrained system. Firstly, the dimension of the reduced state space may vary between systems of different dimensionality, such as is the case for locomotion systems (Hemami and Wyman, 1979). Secondly, the reduction of the state equations makes their interpretation less natural and more difficult (Simon and Chia, 2002). The most recent hard constraints method with non-reduced form of state equation is based on the projection method in which the constrained estimate is obtained from the unconstrained estimate of the conventional Kalman filter by projecting onto the subspace or the surface described by the hard constraint. Recently, Mahata and Söderstöm (2004) used a similar approach in estimating deterministic parameters of viscoelastic materials. Here the additional linear constraints are imposed in the form of a partially known boundary condition to obtain better estimates.

Chia (1985) and Simon and Chia (2002) used this projection method with Kalman prediction and filtering, and proved that the state estimation error covariance of the projected estimate is smaller than that of the unconstrained estimate. Wen and Durrant-Whyte (1992) independently developed a similar method by firstly including the constraint as a perfect observation and then showing that their method is theoretically exactly the same as projecting the filtered estimate of the unconstrained Kalman filter onto the surface of the hard constraint. In conventional linear stochastic models with additive white process noise, for a state vector to be constrained *strongly* in a proper subspace of the whole state space, the process noise must have a singular covariance consistent with a linear constraint on the state. However, because these two authors first used a positive definite or bigger process noise covariance for estimating the constrained system and then projected onto the constraint surface, their method is not optimal, which we show in later sections.

In the next section we consider possible models for linear stochastic systems with equality state constraints and then represent the system in a descriptor form and also in a projected form. In Section 3 we deal with the unconstrained Kalman filter and in Section 4 we consider two constrained estimators and compare these in terms of error covariance.

In this paper, matrices will be denoted by upper case boldface (e.g., $\mathbf{A}$), linear spaces are denoted by calligraphic uppercase (e.g., $\mathcal{A}$), column matrices (vectors) will be denoted by lower case boldface (e.g., $\mathbf{x}$), and scalars will be denoted by lower case (e.g., y) or upper case (e.g., Y). For a matrix $\mathbf{A}$, $\mathbf{A}^T$ denotes its transpose and $\mathbf{A}^\dagger$ represents the Moore-Penrose inverse of $\mathbf{A}$. For a symmetric matrices $\mathbf{P} > \mathbf{0}$ or $\mathbf{P} \geq \mathbf{0}$ denotes the fact that $\mathbf{P}$ is positive definite or positive semi-definite, respectively. For a random vector $\mathbf{x}$, $\mathcal{E}\{\mathbf{x}\}$ represents the mathematical expectation of $\mathbf{x}$.

2. LINEAR STOCHASTIC SYSTEMS WITH EQUALITY CONSTRAINTS

We investigate a method of estimating the state of systems modeled by a linear stochastic difference equation of the form

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k \\ \mathbf{y}_k &= \mathbf{C}\mathbf{x}_k + \mathbf{v}_k \end{aligned} \tag{1}$$

where the state $\mathbf{x}_k \in \mathbb{R}^n$ is known to be constrained in the null space of $\mathbf{D}$

$$\mathcal{N}(\mathbf{D}) \triangleq \{\mathbf{x} : \mathbf{D}\mathbf{x} = \mathbf{d} = \mathbf{0}\}\ ^2 \tag{2}$$

and $\mathbf{y}_k \in \mathbb{R}^p$ represents the measurement. Here $\mathbf{w}_k \in \mathbb{R}^n$ and $\mathbf{v}_k \in \mathbb{R}^p$ are of zero-mean white gaussian distribution with

$$\mathcal{E}\left\{ \begin{bmatrix} \mathbf{w}_k \\ \mathbf{v}_k \end{bmatrix} \begin{bmatrix} \mathbf{w}_j \\ \mathbf{v}_j \end{bmatrix}^T \right\} = \begin{bmatrix} \mathbf{Q}^c & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{bmatrix} \delta_{kj} \tag{3}$$

where δ_{kj} represents the Kronecker delta ($\delta_{kj} = 1$ if $k = j$, 0 otherwise). The matrix $\mathbf{D} \in \mathbb{R}^{m \times n}$ is assumed to have a full row rank. If $\mathbf{D}$ does not have a full row rank, there exist redundant state constraints. In that case we can simply remove linearly dependent rows from $\mathbf{D}$.

Since the allowable space of $\mathbf{x}_k$ with constraint is $\mathcal{N}(\mathbf{D})$, $\mathbf{x}_{k+1}$ given by (1) also must satisfy $\mathbf{x}_{k+1} \in \mathcal{N}(\mathbf{D})$, for which we identify the following possible cases:

Case 1: $(\mathbf{A}\mathbf{x}_k,\ \mathbf{B}\mathbf{u}_k,\ \mathbf{w}_k) \notin \mathcal{N}(\mathbf{D})$
Since the sum of the three elements must satisfy

[2] For $\mathbf{d} \neq \mathbf{0}$ case, translation of the state space such that $\mathbf{x}_k = \bar{\mathbf{x}}_k + \mathbf{D}^\dagger \mathbf{d}$ yields the state equation form (1) with an additional term in the state equation which can be treated as a deterministic constant noise. Hence, without loss of generality, it suffices to consider $\mathbf{d} = \mathbf{0}$ case for the analysis of current state estimation problem.

$\mathbf{Ax}_k$+$\mathbf{Bu}_k$+$\mathbf{w}_k \in \mathcal{N}(\mathbf{D})$ for any $\mathbf{x}_k \in \mathcal{N}(\mathbf{D})$, this case causes a correlated noise $\mathbf{w}_k$ with current input $\mathbf{u}_k$ and state $\mathbf{x}_k$ and causes theoretical as well as practical problems. Hence it is not a proper Markovian system model.

Case 2: $(\mathbf{Ax}_k,\ \mathbf{Bu}_k) \notin \mathcal{N}(\mathbf{D})$ but $\mathbf{w}_k \in \mathcal{N}(\mathbf{D})$
Similarly, the sum of the first two elements must satisfy $\mathbf{Ax}_k + \mathbf{Bu}_k \in \mathcal{N}(\mathbf{D})$. This case allows uncorrelated noise sequences $\mathbf{w}_k$ with the input $\mathbf{u}_k$ or the state $\mathbf{x}_k$, but the system cannot maintain the state constraint without corrective action of the input $\mathbf{u}_k$. Hence, this model is suitable for modeling systems having *design constraints.*

Case 3: $(\mathbf{Ax}_k, \mathbf{Bu}_k,\ \mathbf{w}_k) \in \mathcal{N}(\mathbf{D})$
This case also allows uncorrelated noise sequences $\mathbf{w}_k$ with the input $\mathbf{u}_k$ or the state $\mathbf{x}_k$ and has a proper form for modeling systems with *physical constraints* since regardless of the corrective input $\mathbf{u}_k$ the state stays within the constraint subspace $\mathcal{N}(\mathbf{D})$. Since, for all $\mathbf{x}_k \in \mathcal{N}(\mathbf{D})$, it is required that $\mathbf{Ax}_k \in \mathcal{N}(\mathbf{D})$, $\mathcal{N}(\mathbf{D})$ is $\mathbf{A}$-invariant (Wonham, 1979).

As in the Cases 2 and 3, the process noise sequence $\mathbf{w}_k$ being constrained in the null space of $\mathbf{D}$ is a reasonable choice from the viewpoint of usual system modeling practice: we model a system without considering noise or model uncertainty and then add a noise sequence to compensate for uncertainty of the system. Since the noise is in a lower dimensional space than the whole state space $\mathbb{R}^n$, support of the probability density function must have a lower dimension than the state space, which means the covariance matrix of the noise must be singular.

2.1 Descriptor System Representation of Constrained Systems

The linear system (1) with the constraint (2) can be represented in the form of a descriptor system. To do this, first we define the following:

Definition 1. [Skelton *et al.* (1998)] For a given matrix $\mathbf{N} \in \mathbb{R}^{n\times m}$ with rank r, $\mathbf{N}^{\perp} \in \mathbb{R}^{(n-r)\times n}$ is defined as any matrix such that $\mathbf{N}^{\perp}\mathbf{N} = \mathbf{0}$ and $\mathbf{N}^{\perp}\mathbf{N}^{\perp^T} > \mathbf{0}$.

Remark 1. Note that the matrix $\mathbf{N}^{\perp}$ defined in Definition 1 exists if and only if $\mathbf{N}$ has linearly dependent rows ($n > r$), and the set of all such matrices can be captured by $\mathbf{N}^{\perp} = \mathbf{T}\mathbf{U}_2^T$, where $\mathbf{T}$ is an arbitrary nonsingular matrix and $\mathbf{U}_2$ is from the singular value decomposition (SVD)

$$\mathbf{N} = [\mathbf{U}_1\ \mathbf{U}_2]\begin{bmatrix}\mathbf{\Sigma}_r & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{V}_1^T\\ \mathbf{V}_2^T\end{bmatrix}. \tag{4}$$

In this paper, we consider only $\mathbf{T} = \mathbf{I}$ case and hence $\mathbf{N}^{\perp} = \mathbf{U}_2^T$.

Since $\mathbf{D}$ is of full row rank, $\mathbf{D}^T$ has dependent rows and therefore $\mathbf{D}^{T^{\perp}}$ can be defined. Then, by multiplying both sides of the state equation (1) by $\mathbf{D}^{T^{\perp}}$ and using the constrained equation (2) we obtain the following descriptor system representation

$$\mathbf{E}\mathbf{x}_{k+1} = \bar{\mathbf{A}}\mathbf{x}_k + \bar{\mathbf{B}}\mathbf{u}_k + \mathbf{E}\mathbf{w}_k \tag{5a}$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{v}_k \tag{5b}$$

where

$$\mathbf{E} = \begin{bmatrix}\mathbf{D}^{T^{\perp}}\\ \mathbf{0}\end{bmatrix},\ \bar{\mathbf{A}} = \begin{bmatrix}\mathbf{D}^{T^{\perp}}\mathbf{A}\\ \mathbf{D}\end{bmatrix},\ \bar{\mathbf{B}} = \begin{bmatrix}\mathbf{D}^{T^{\perp}}\mathbf{B}\\ \mathbf{0}\end{bmatrix}. \tag{6}$$

2.2 Projected system representation of constrained systems

In a descriptor system representation such as (5a), $\mathbf{x}_k$ is termed *descriptor* vector. Luenberger (1977) defined a *state* as a vector $\mathbf{z}_k = \mathbf{\Gamma}\mathbf{x}_k$, whose dimension is smaller than the descriptor vector $\mathbf{x}_k$ with a matrix $\mathbf{\Gamma}$ having a specific property, for a set of dynamic equations if knowledge of its value and the value of input and the noise sequence are sufficient to uniquely determine the descriptor vector $\mathbf{x}_k$. Furthermore, an equivalent condition was obtained for both a time-invariant descriptor system to be *regular* [3] and for a vector $\mathbf{z}_k = \mathbf{\Gamma}\mathbf{x}_k$ to be a *state* for the descriptor system.

The following Lemma 1 can be easily verified.

Lemma 1. $\begin{bmatrix}\mathbf{D}^{T^{\perp}}\\ \mathbf{D}\end{bmatrix}$ is invertible and $\begin{bmatrix}\mathbf{D}^{T^{\perp}}\\ \mathbf{D}\end{bmatrix}^{-1} = \begin{bmatrix}\mathbf{D}^{T^{\perp}T} & \mathbf{D}^{\dagger}\end{bmatrix}$.

By Lemma 1 above and Theorems 6 and 7 in Luenberger (1977), the following theorem is obtained.

Theorem 1. The following statements (i) and (ii) are equivalent:

(i) The descriptor system (5a) is regular and the vector $\mathbf{z}_k = \mathbf{D}^{T^{\perp}}\mathbf{x}_k$ is a state.
(ii) The matrix $\begin{bmatrix}\mathbf{D}^{T^{\perp}}\\ \mathbf{D}\end{bmatrix}$ is square and nonsingular.

By Lemma 1 and Theorem 1, the descriptor system (5a) describing the constrained system (1) and (2) is regular and $\mathbf{z}_k = \mathbf{D}^{T^{\perp}}\mathbf{x}_k$ is a state, from which the descriptor system (5a) can be further simplified through combining the state expression $\mathbf{z}_k = \mathbf{D}^{T^{\perp}}\mathbf{x}_k$ with the state constraint $\mathbf{D}\mathbf{x}_k = \mathbf{0}$, leading to

$$\begin{bmatrix}\mathbf{D}^{T^{\perp}}\\ \mathbf{D}\end{bmatrix}\mathbf{x}_k = \begin{bmatrix}\mathbf{z}_k\\ \mathbf{0}\end{bmatrix}. \tag{7}$$

Applying Lemma 1 to (7) yields

$$\mathbf{x}_k = \mathbf{D}^{T^{\perp}T}\mathbf{z}_k \quad \text{or} \quad \mathbf{x}_{k+1} = \mathbf{D}^{T^{\perp}T}\mathbf{z}_{k+1}. \tag{8}$$

[3] A set of dynamic equation is said to be *regular* if there is an initial condition vector which when propagated forward serves as a state vector for every time period.

From the definition of $\mathbf{z}_k$ and (1),

$$\mathbf{z}_{k+1} = \mathbf{D}^{T^\perp}\mathbf{x}_{k+1} = \mathbf{D}^{T^\perp}(\mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k). \quad (9)$$

Substituting (9) into (8), we have

$$\mathbf{x}_{k+1} = \mathbf{D}^{T^{\perp T}}\mathbf{D}^{T^\perp}\Big(\mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k\Big) \quad (10)$$
$$= P_{\mathcal{N}}\left(\mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k\right) \quad (11)$$

where $P_{\mathcal{N}} \triangleq \mathbf{D}^{T^{\perp T}}\mathbf{D}^{T^\perp} = \mathbf{U}_2\mathbf{U}_2^T$ is the orthogonal projector onto the null space of $\mathbf{D}$ since $\mathbf{U}_2$ spans $\mathcal{N}(\mathbf{D})$. Therefore, owing to $P_{\mathcal{N}}$, the system representation (11) is termed a *projected system.* Furthermore, for Case 3-constrained systems, one important consequence can be drawn. From the fact that, for any $\mathbf{x}_k \in \mathcal{N}(\mathbf{D})$, also $\mathbf{A}\mathbf{x}_k \in \mathcal{N}(\mathbf{D})$ holds, we have

$$\mathbf{A}\mathbf{x}_k = \mathbf{P}\mathbf{A}\mathbf{x}_k = \mathbf{A}\mathbf{P}\mathbf{x}_k \quad (12)$$

where $\mathbf{P}$ is any projection matrix onto the null space of $\mathbf{D}$. By taking a conditional expectation for given any measurements $\mathcal{Y}$ on both sides of (12), we obtain also

$$\mathbf{P}\mathbf{A}\ \mathcal{E}\{\mathbf{x}_k|\mathcal{Y}\} = \mathbf{A}\mathbf{P}\ \mathcal{E}\{\mathbf{x}_k|\mathcal{Y}\}. \quad (13)$$

Using (12) and (13) yields

$$\begin{aligned} \mathbf{A}\mathbf{P}\mathbf{\Sigma} = \mathbf{P}\mathbf{A}\mathbf{\Sigma} = \mathbf{A}\mathbf{\Sigma} \quad \text{and} \\ \mathbf{A}\mathbf{P}\mathbf{\Sigma}\mathbf{P}^T\mathbf{A}^T = \mathbf{P}\mathbf{A}\mathbf{\Sigma}\mathbf{A}^T\mathbf{P}^T = \mathbf{A}\mathbf{\Sigma}\mathbf{A}^T \end{aligned} \quad (14)$$

where $\mathbf{\Sigma} = \mathcal{E}\left\{[\mathbf{x}_k - \mathcal{E}(\mathbf{x}_k|\mathcal{Y})]\,[\mathbf{x}_k - \mathcal{E}\,(\mathbf{x}_k|\mathcal{Y})]^T\right\}$. The equations in (14) are crucial relations in subsequent analysis for comparing performance between different estimators which can be used for estimating the constrained system.

3. UNCONSTRAINED KALMAN FILTER

As discussed in Section 2, for the constrained system given by (1)-(2) to have an uncorrelated process noise sequence $\mathbf{w}_k$ with both current state $\mathbf{x}_k$ and input $\mathbf{u}_k$, it is required that the process noise be constrained in the null space $\mathcal{N}(\mathbf{D})$ of the constraint matrix $\mathbf{D}$, which means the covariance matrix of $\mathbf{w}_k$ must be singular. Hence, in order to design the "correct" Kalman filter for such a constrained system, the exact (singular) covariance matrix of the constrained noise must be used. When one does not know accurately the noise covariance, it is common practice (Anderson and Moore, 1979) to use the upper bound of the noise covariance or simply any bigger noise covariance than the (expected) correct value, causing, in some sense, a worst case design. Instead of the true singular (positive semi-definite) covariance matrix $\mathbf{Q}^c$, if a positive definite process noise covariance $\mathbf{Q}$ ($\geq \mathbf{Q}^c$) is used for the constrained system (1), then bigger or *unconstrained* estimation error covariances will be resulted. The corresponding Kalman predictor is given by the following equations:

$$\begin{aligned} \hat{\mathbf{x}}^u_{k+1|k} &= (\mathbf{A} - \mathbf{M}^u_k\mathbf{C})\hat{\mathbf{x}}^u_{k|k-1} + \mathbf{B}\mathbf{u}_k + \mathbf{M}^u_k\mathbf{y}_k \\ \mathbf{M}^u_k &= \mathbf{A}\mathbf{\Sigma}^u_{k|k-1}\mathbf{C}^T(\mathbf{C}\Sigma^u_{k|k-1}\mathbf{C}^T + \mathbf{R})^{-1} \\ \mathbf{\Sigma}^u_{k+1|k} &= \mathbf{A}\mathbf{\Sigma}^u_{k|k-1}\mathbf{A}^T + \mathbf{Q} \\ &\quad -\mathbf{A}\mathbf{\Sigma}^u_{k|k-1}\mathbf{C}^T(\mathbf{C}\mathbf{\Sigma}^u_{k|k-1}\mathbf{C}^T + \mathbf{R})^{-1}\mathbf{C}\mathbf{\Sigma}^u_{k|k-1}\mathbf{A}^T \end{aligned} \quad (15)$$

4. CONSTRAINED KALMAN FILTER

In this section, we consider the two different constrained predictors for the Case 3 constrained system described in Section 2. For a simple notation, the subscripts $(\cdot)_k$ will be used for denoting the Kalman predictor instead of $(\cdot)_{k|k-1}$.

4.1 Projected Kalman Filter

Chia (1985) and Simon and Chia (2002) derived a constrained Kalman predictor by directly projecting the unconstrained state estimate $\hat{\mathbf{x}}^u_k$ onto the constrained subspace $\mathcal{N}(\mathbf{D})$. Let us name it *projected* estimator which will be denoted by the superscript $(\cdot)^p$. They solved the problem, for any symmetric positive definite weighting matrix $\mathbf{W}$,

$$\begin{aligned} \min_{\hat{\mathbf{x}}^p_k}\ &(\hat{\mathbf{x}}^p_k - \hat{\mathbf{x}}^u_k)^T\mathbf{W}(\hat{\mathbf{x}}^p_k - \hat{\mathbf{x}}^u_k) \\ &\text{subject to } \mathbf{D}\hat{\mathbf{x}}^p_k = \mathbf{0} \end{aligned} \quad (16)$$

and obtained

$$\hat{\mathbf{x}}^p_k = P^W_{\mathcal{N}(\mathbf{D})}\hat{\mathbf{x}}^u_k \quad (17)$$

where $P^W_{\mathcal{N}(\mathbf{D})} \triangleq \mathbf{I} - \mathbf{W}^{-1}\mathbf{D}^T(\mathbf{D}\mathbf{W}^{-1}\mathbf{D}^T)^{-1}\mathbf{D}$ which is a projector to the constraint subspace $\mathcal{N}(\mathbf{D})$ with a weighting matrix $\mathbf{W}$.

The property of the projected Kalman predictor is summarized in the following theorem.

Theorem 2. [Chia (1985);Simon and Chia (2002), Projected Kalman predictor]

(i) The projected state estimate $\hat{\mathbf{x}}^p_k$ given by (17) with $\mathbf{W} = (\mathbf{\Sigma}^u_k)^{-1}$ has a smaller state error covariance than that of the unconstrained state estimate $\hat{\mathbf{x}}^u_k$. That is

$$\mathbf{\Sigma}^p_k \triangleq \mathbf{Cov}(\mathbf{x}_k - \hat{\mathbf{x}}^p_k) \leq \mathbf{Cov}(\mathbf{x}_k - \hat{\mathbf{x}}^u_k) = \mathbf{\Sigma}^u_k \quad (18)$$

and the covariance of the projected estimator is given by

$$\mathbf{\Sigma}^p_k = \mathbf{P}_k\mathbf{\Sigma}^u_k\mathbf{P}^T_k = \mathbf{P}_k\mathbf{\Sigma}^u_k, \quad (19)$$

where $\mathbf{P}_k = \mathbf{I} - \mathbf{\Sigma}^u_k\boldsymbol{D}^T(\boldsymbol{D}\mathbf{\Sigma}^u_k\boldsymbol{D}^T)^{-1}\boldsymbol{D}$ is a projection matrix onto the null space of $\mathbf{D}$.

(ii) Among all the projected Kalman predictors of (17), the predictor that uses $\mathbf{W} = (\mathbf{\Sigma}^u_k)^{-1}$ has the smallest estimation error covariance.

Therefore, at $(k+1)$th-stage the state estimation error covariance of the projected Kalman predictor is, from (15) and (19), given by

$$\begin{aligned} \mathbf{\Sigma}^p_{k+1} &= \mathbf{P}_{k+1}\mathbf{\Sigma}^u_{k+1}\mathbf{P}^T_{k+1} \\ &= \mathbf{P}_{k+1}\mathbf{A}\mathbf{\Sigma}^u_k\mathbf{A}^T\mathbf{P}^T_{k+1} + \mathbf{P}_{k+1}\mathbf{Q}\mathbf{P}^T_{k+1} \\ &\quad - \mathbf{P}_{k+1}\mathbf{A}\mathbf{\Sigma}^u_k\mathbf{C}^T(\mathbf{C}\mathbf{\Sigma}^u_k\mathbf{C}^T + \mathbf{R})^{-1}\mathbf{C}\mathbf{\Sigma}^u_k\mathbf{A}^T\mathbf{P}^T_{k+1}. \end{aligned} \quad (20)$$

4.2 Constrained Kalman Filter for Projected System

In Section 2.2, we have proved that the original state equation (1) with the constraint (2) can be reduced to the projected system (11). Therefore, for the Case 3 projected system, the Kalman predictor is given by

$$\begin{aligned}\hat{\mathbf{x}}_{k+1}^c &= (\mathbf{P}_\mathcal{N}\mathbf{A} - \mathbf{M}_k^c\mathbf{C})\hat{\mathbf{x}}_k^c + \mathbf{B}\mathbf{u}_k + \mathbf{M}_k^c\mathbf{y}_k \\ \mathbf{M}_k^c &= \mathbf{P}_\mathcal{N}\mathbf{A}\boldsymbol{\Sigma}_k^c\mathbf{C}^T(\mathbf{C}\boldsymbol{\Sigma}_k^c\mathbf{C}^T + \mathbf{R})^{-1} \\ \boldsymbol{\Sigma}_{k+1}^c &= \mathbf{P}_\mathcal{N}\mathbf{A}\boldsymbol{\Sigma}_k^c\mathbf{A}^T\mathbf{P}_\mathcal{N} + \mathbf{P}_\mathcal{N}\mathbf{Q}\mathbf{P}_\mathcal{N} \\ &\quad - \mathbf{P}_\mathcal{N}\mathbf{A}\boldsymbol{\Sigma}_k^c\mathbf{C}^T(\mathbf{C}\boldsymbol{\Sigma}_k^c\mathbf{C}^T + \mathbf{R})^{-1}\mathbf{C}\boldsymbol{\Sigma}_k^c\mathbf{A}^T\mathbf{P}_\mathcal{N}\end{aligned} \tag{21}$$

where it is assumed that $\mathbf{Q}^c = \mathbf{P}_\mathcal{N}\mathbf{Q}\mathbf{P}_\mathcal{N}$.

Remark 2. In Section 2.2, a reduced state $\mathbf{z}_k = \mathbf{D}^{T^\perp}\mathbf{x}_k$ was used in the middle of deriving the projected system (11). We can also construct the Kalman predictor $\hat{\mathbf{z}}_k$ for estimating this reduced state $\mathbf{z}_k$ and use the relation $\hat{\mathbf{x}}_k = \mathbf{D}^{T^{\perp T}}\hat{\mathbf{z}}_k$ from (8) which will yield the Kalman predictor given in (21).

4.3 Comparison of Constrained Kalman Filters

From optimality of the Kalman predictor, we expect that the Kalman predictor derived from the projected state equation (11) is the optimal filter for the original constrained system modeled by (1) and (2). But, we need to know whether the projected estimator of Section 4.1 is also optimal or not. To compare the error covariances given by (20) and (21), let us express (20) as

$$\begin{aligned}\boldsymbol{\Sigma}_{k+1}^p &= \mathbf{P}_{k+1}\boldsymbol{\Sigma}_{k+1}^u\mathbf{P}_{k+1}^T \\ &= \mathbf{P}_{k+1}\Big\{\mathbf{A}\boldsymbol{\Sigma}_k^u\mathbf{A}^T + \mathbf{Q} \\ &\quad -\mathbf{A}\boldsymbol{\Sigma}_k^u\mathbf{C}^T\big(\mathbf{C}\boldsymbol{\Sigma}_k^u\mathbf{C}^T + \mathbf{R}\big)^{-1}\mathbf{C}\boldsymbol{\Sigma}_k^u\mathbf{A}^T\Big\}\mathbf{P}_{k+1}^T \\ &= \mathbf{P}_k\Big\{\mathbf{A}\boldsymbol{\Sigma}_k^p\mathbf{A}^T + \mathbf{Q} \\ &\quad -\mathbf{A}\boldsymbol{\Sigma}_k^p\mathbf{C}^T\big(\mathbf{C}\boldsymbol{\Sigma}_k^p\mathbf{C}^T + \mathbf{R}\big)^{-1}\mathbf{C}\boldsymbol{\Sigma}_k^p\mathbf{A}^T\Big\}\mathbf{P}_k^T \\ &\quad + \Delta\mathbf{Q}_k\end{aligned} \tag{22}$$

where

$$\begin{aligned}\Delta\mathbf{Q}_k \triangleq\ & \mathbf{P}_{k+1}\Big\{\mathbf{A}\boldsymbol{\Sigma}_k^u\mathbf{A}^T + \mathbf{Q} \\ & -\mathbf{A}\boldsymbol{\Sigma}_k^u\mathbf{C}^T\big(\mathbf{C}\boldsymbol{\Sigma}_k^u\mathbf{C}^T + \mathbf{R}\big)^{-1}\mathbf{C}\boldsymbol{\Sigma}_k^u\mathbf{A}^T\Big\}\mathbf{P}_{k+1}^T \\ & - \mathbf{P}_k\Big\{\mathbf{A}\boldsymbol{\Sigma}_k^p\mathbf{A}^T + \mathbf{Q} \\ & -\mathbf{A}\boldsymbol{\Sigma}_k^p\mathbf{C}^T\big(\mathbf{C}\boldsymbol{\Sigma}_k^p\mathbf{C}^T + \mathbf{R}\big)^{-1}\mathbf{C}\boldsymbol{\Sigma}_k^p\mathbf{A}^T\Big\}\mathbf{P}_k^T.\end{aligned} \tag{23}$$

Since $\boldsymbol{\Sigma}_k^p$ is constrained and using (14), we have

$$\begin{aligned}\mathbf{P}_k\mathbf{A}\boldsymbol{\Sigma}_k^p\mathbf{A}^T\mathbf{P}_k^T &= \mathbf{P}_{k+1}\mathbf{P}_k\mathbf{A}\boldsymbol{\Sigma}_k^p\mathbf{A}^T\mathbf{P}_k^T\mathbf{P}_{k+1}^T \\ &= \mathbf{P}_{k+1}\mathbf{A}\mathbf{P}_k\boldsymbol{\Sigma}_k^p\mathbf{P}_k^T\mathbf{A}^T\mathbf{P}_{k+1}^T \\ &= \mathbf{P}_{k+1}\mathbf{A}\boldsymbol{\Sigma}_k^p\mathbf{A}^T\mathbf{P}_{k+1}^T\end{aligned} \tag{24}$$

and also

$$\begin{aligned}&\mathbf{P}_k\mathbf{A}\boldsymbol{\Sigma}_k^p\mathbf{C}^T\left(\mathbf{C}\boldsymbol{\Sigma}_k^p\mathbf{C}^T + \mathbf{R}\right)^{-1}\mathbf{C}\boldsymbol{\Sigma}_k^p\mathbf{A}^T\mathbf{P}_k^T \\ &= \mathbf{P}_{k+1}\mathbf{A}\boldsymbol{\Sigma}_k^p\mathbf{C}^T\left(\mathbf{C}\boldsymbol{\Sigma}_k^p\mathbf{C}^T + \mathbf{R}\right)^{-1}\mathbf{C}\boldsymbol{\Sigma}_k^p\mathbf{A}^T\mathbf{P}_{k+1}^T.\end{aligned} \tag{25}$$

Therefore, we have

$$\Delta\mathbf{Q}_k = \mathbf{P}_{k+1}\boldsymbol{\Delta}_k\mathbf{P}_{k+1}^T + \mathbf{P}_{k+1}\mathbf{Q}\mathbf{P}_{k+1}^T - \mathbf{P}_k\mathbf{Q}\mathbf{P}_k^T \tag{26}$$

where

$$\begin{aligned}\boldsymbol{\Delta}_k =\ & \mathbf{A}\boldsymbol{\Sigma}_k^u\mathbf{A}^T - \mathbf{A}\boldsymbol{\Sigma}_k^u\mathbf{C}^T\left(\mathbf{C}\boldsymbol{\Sigma}_k^u\mathbf{C}^T + \mathbf{R}\right)^{-1}\mathbf{C}\boldsymbol{\Sigma}_k^u\mathbf{A}^T \\ & - \mathbf{A}\boldsymbol{\Sigma}_k^p\mathbf{A}^T + \mathbf{A}\boldsymbol{\Sigma}_k^p\mathbf{C}^T\left(\mathbf{C}\boldsymbol{\Sigma}_k^p\mathbf{C}^T + \mathbf{R}\right)^{-1}\mathbf{C}\boldsymbol{\Sigma}_k^p\mathbf{A}^T.\end{aligned} \tag{27}$$

To prove that $\boldsymbol{\Delta}_k \geq \mathbf{0}$, we need the following lemma about a monotonicity property of the Riccati Difference Equation.

Lemma 2. [De Souza (1989);Bitmead and Gevers (1991)] Consider two RDEs with the same $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{R}$ matrices but possibly different $\mathbf{Q}^1$ and $\mathbf{Q}^2$. Denote their solution matrices $\boldsymbol{\Sigma}_k^1$ and $\boldsymbol{\Sigma}_k^2$ respectively. Suppose that $\mathbf{Q}^1 \geq \mathbf{Q}^2$, and, for some k, we have $\boldsymbol{\Sigma}_k^1 \geq \boldsymbol{\Sigma}_k^2$, then for all $i > 0$

$$\boldsymbol{\Sigma}_{k+i}^1 \geq \boldsymbol{\Sigma}_{k+i}^2. \tag{28}$$

Theorem 2 tells us that $\boldsymbol{\Sigma}_k^u \geq \boldsymbol{\Sigma}_k^p$ and then for (27) we can deduce that $\boldsymbol{\Delta}_k \geq \mathbf{0}$ from Lemma 2.

Hence, (22) can be written as

$$\begin{aligned}\boldsymbol{\Sigma}_{k+1}^p =\ & \mathbf{P}_k\Big\{\mathbf{A}\boldsymbol{\Sigma}_k^p\mathbf{A}^T \\ & -\mathbf{A}\boldsymbol{\Sigma}_k^p\mathbf{C}^T\Big(\mathbf{C}\boldsymbol{\Sigma}_k^p\mathbf{C}^T + \mathbf{R}\Big)^{-1}\mathbf{C}\boldsymbol{\Sigma}_k^p\mathbf{A}^T\Big\}\mathbf{P}_k^T \\ & + \mathbf{P}_{k+1}\boldsymbol{\Delta}_k\mathbf{P}_{k+1}^T + \mathbf{P}_{k+1}\mathbf{Q}\mathbf{P}_{k+1}^T \\ =\ & \mathbf{A}\boldsymbol{\Sigma}_k^p\mathbf{A}^T + \mathbf{P}_{k+1}\boldsymbol{\Delta}_k\mathbf{P}_{k+1}^T + \mathbf{P}_{k+1}\mathbf{Q}\mathbf{P}_{k+1}^T \\ & - \mathbf{A}\boldsymbol{\Sigma}_k^p\mathbf{C}^T\left(\mathbf{C}\boldsymbol{\Sigma}_k^p\mathbf{C}^T + \mathbf{R}\right)^{-1}\mathbf{C}\boldsymbol{\Sigma}_k^p\mathbf{A}^T.\end{aligned} \tag{29}$$

Recall that from Section 4.2 and using (14) we have, with $\mathbf{Q}^c = \mathbf{P}_\mathcal{N}\mathbf{Q}\mathbf{P}_\mathcal{N}$,

$$\begin{aligned}\boldsymbol{\Sigma}_{k+1}^c &= \mathbf{A}\boldsymbol{\Sigma}_k^c\mathbf{A}^T + \mathbf{Q}^c \\ &\quad - \mathbf{A}\boldsymbol{\Sigma}_k^c\mathbf{C}^T(\mathbf{C}\boldsymbol{\Sigma}_k^c\mathbf{C}^T + \mathbf{R})^{-1}\mathbf{C}\boldsymbol{\Sigma}_k^c\mathbf{A}^T.\end{aligned} \tag{30}$$

The following Theorem 3 summarizes what we have shown regarding performance of three Kalman predictors that can be used for estimating the constrained state.

Theorem 3. For Case 3 constrained systems with the assumptions that the pair $[\mathbf{A}, \mathbf{Q}^{c^{1/2}}]$ is stabilizable and $[\mathbf{A}, \mathbf{C}]$ is detectable, the following hold:

(i) If $\mathbf{Q} > \mathbf{Q}^c$, with the initial conditions such that $\boldsymbol{\Sigma}_{0|-1}^c \leq \boldsymbol{\Sigma}_{0|-1}^p \leq \boldsymbol{\Sigma}_{0|-1}^u$, we have

$$\boldsymbol{\Sigma}_{k+1|k}^c \leq \boldsymbol{\Sigma}_{k+1|k}^p \leq \boldsymbol{\Sigma}_{k+1|k}^u \tag{31a}$$

$$\lim_{k\to\infty}\boldsymbol{\Sigma}_{k+1|k}^c \leq \lim_{k\to\infty}\boldsymbol{\Sigma}_{k+1|k}^p \leq \lim_{k\to\infty}\boldsymbol{\Sigma}_{k+1|k}^u. \tag{31b}$$

(ii) If $\mathbf{Q} = \mathbf{Q}^c$, with the initial conditions such that $\boldsymbol{\Sigma}_{0|-1}^c = \boldsymbol{\Sigma}_{0|-1}^p = \mathbf{P}_0\boldsymbol{\Sigma}_{0|-1}^u\mathbf{P}_0^T$, we have

$$\boldsymbol{\Sigma}_{k+1|k}^c \leq \boldsymbol{\Sigma}_{k+1|k}^p \leq \boldsymbol{\Sigma}_{k+1|k}^u \tag{32a}$$

$$\lim_{k\to\infty}\boldsymbol{\Sigma}_{k+1|k}^c = \lim_{k\to\infty}\boldsymbol{\Sigma}_{k+1|k}^p = \lim_{k\to\infty}\boldsymbol{\Sigma}_{k+1|k}^u. \tag{32b}$$

PROOF.

(i) Using $\mathbf{P}_{k+1}\mathbf{Q}\mathbf{P}_{k+1}^T \geq \mathbf{P}_{\mathcal{N}}\mathbf{Q}\mathbf{P}_{\mathcal{N}} = \mathbf{Q}^c$ from the fact that $\mathbf{P}_{\mathcal{N}}$ is the orthogonal projector, and $\mathbf{P}_{k+1}\mathbf{\Delta}_k\mathbf{P}_{k+1}^T \geq \mathbf{0}$, it can be shown that, again through Lemma 2 and Theorem 2, we obtain (31a) and also (31b). The existence of limit matrices is guaranteed from the assumptions of stabilizability of $[\mathbf{A}, \mathbf{Q}^{c^{1/2}}]$ (which implies also stabilizable $[\mathbf{A}, \mathbf{Q}^{1/2}]$, since $\mathbf{Q} > \mathbf{0}$) and detectability of $[\mathbf{A}, \mathbf{C}]$.

(ii) We observe that the RDEs (15) and (30) are the same except the different initial conditions such that $\mathbf{\Sigma}_{0|-1}^u \geq \mathbf{\Sigma}_{0|-1}^c = \mathbf{P}_0\mathbf{\Sigma}_{0|-1}^u\mathbf{P}_0^T$, from which, in combination with Lemma 2, we have $\mathbf{\Sigma}_k^p = \mathbf{P}_k\mathbf{\Sigma}_k^u\mathbf{P}_k^T \geq \mathbf{P}_k\mathbf{\Sigma}_k^c\mathbf{P}_k^T = \mathbf{\Sigma}_k^c$. Therefore, we obtain (32a). With the assumptions of stabilizable $[\mathbf{A}, \mathbf{Q}^{c^{1/2}}]$ and detectable $[\mathbf{A}, \mathbf{C}]$, the initial condition effects fade away as $k \to \infty$ and thus we obtain (32b). □

Remark 3. It can be shown that for Case 3 constrained systems, the *filter* version of Theorem 3 also holds.

5. CONCLUDING REMARKS

In this paper, we have analyzed the three estimators that can be used for estimating linear systems with known state equality constraints. Among them, it was proved that the current constrained estimator is optimal and thus outperforms the unconstrained and the projected estimators. The procedures used for discrete-time system can be similarly extended to the continuous-time case, which can be found in Ko (2005).

REFERENCES

Alouani, A. T. and W. D. Blair (1993). Use of a kinematic constraint in tracking constant speed, maneuvering targets. *IEEE Trans. Automat. Contr.* **38**, 1107–1111.

Altman, Russ B. and Rhea Tombropoulos (1994). Probabilistic constraint satisfaction. Reports SMI-94-0531. Stanford Medical Informatics.

Anderson, Brian D. O. and John B. Moore (1979). *Optimal Filtering.* Prentice-Hall. Englewood Cliffs, NY.

Bitmead, Robert R. and Michel Gevers (1991). Riccati difference and differential equations: Convergence, monotonicity and stability. In: *The Riccati Equation* (Sergio Bittanti, Alan J. Laub and Jan C. Willems, Eds.). pp. 263–291. Springer-Verlag. Berlin, Germany.

Bryson, Arthur E. Jr. and Yu-Chi Ho (1975). *Applied optimal control : optimization, estimation, and control.* Hemisphere Pub. Corp.. New York, NY.

Chen, Yuan-Hwang and Ching-Tai Chiang (1993). Adaptive beamforming using the constrained Kalman filter. *IEEE Trans. Antennas Propagat.* **41**, 1576–1580.

Chia, Tien-Li (1985). Parameter Identification and State Estimation of Constrained Systems. PhD thesis. Case Western Reserve Univ.. Cleveland, OH.

Chiang, Y. T., L. S. Wang, F. R. Chang and H. M. Peng (2002). Constrained filtering method for attitude determination using GPS and Gyro. *IEE Proc.-Radar Sonar Navig.* **149**, 258–264.

De Souza, Carlos E. (1989). On stabilizing properties of solutions of the riccati difference equation. *IEEE Trans. Automat. Contr.* **34**, 1313–1316.

Geeter, Jan De, Hendrik Van Brussel and Joris De Schutter (1997). A smoothly constrained Kalman filter. *IEEE Trans. Pattern Anal. Machine Intell.* **19**(10), 1171–1177.

Hemami, Hooshang and Bostwick F. Wyman (1979). Modeling and control of constrained dynamic systems with application to biped locomotion in the frontal plane. *IEEE Trans. Automat. Contr.* **24**(4), 526–535.

Ko, Sangho (2005). Performance Limitations in Linear Estimation and Control: Quantization and Constraint Effects (being prepared). PhD thesis. Univ. of California, San Diego. La Jolla, CA.

Lefferts, E. J., F. L. Markley and M. D. Shuster (1982). Kalman filtering for spacecraft attitude estimation. *J. Guid. Control Dyn.* **5**, 417–429.

Luenberger, David G. (1977). Dynamic equations in descriptor form. *IEEE Trans. Automat. Contr.* **22**(3), 312–321.

Maciejowski, Jan Marian (2002). *Predictive Control with Constraints.* Pearson Education Ltd.. Essex, England.

Mahata, Kaushik and Törsen Söderstöm (2004). Improved estimation performance using known linear constraints. *Automatica* **40**, 1307–1318.

Peng, H. M., Y. T. Chiang, F. R. Chang and L. S. Wang (2000). Maximum-likelihood-based filtering for attitude determination via GPS carrier phase. In: *Proc. of IEEE Position Location and Navigation Symposium.* San Diego, California. pp. 480–487.

Rao, Christopher V., James B. Rawlings and David Q. Mayne (2003). Constrained state estimation for nonlinear discrete-time systems: Stability and moving horizon approximations. *IEEE Trans. Automat. Contr.* **48**, 246–258.

Simon, Dan and Tien Li Chia (2002). Kalman filtering with state equality constraints. *IEEE Trans. Aerosp. Electron. Syst.* **38**, 128–136.

Skelton, R. E., T. Iwasaki and K. Grigoriadis (1998). *A Unified Algebraic Approach to Linear Control Design.* Taylor and Francis. London, UK.

Stengel, Robert F. (1994). *Optimal Control and Estimation.* Dover Publications, Inc.. Mineola, NY.

Tahk, Minjea and Jason L. Speyer (1990). Target tracking problems subject to kinematic constraints. *IEEE Trans. Automat. Contr.* **35**, 324–326.

Wen, W. and H. F. Durrant-Whyte (1992). Model-based multi-sensor data fusion. In: *Proc. 1992 IEEE International Conference on Robotics and Automation.* Nice, France. pp. 1720–1726.

Wonham, W. Murray (1979). *Linear multivariable control: a Geometric Approach.* Springer-Verlag. New York, NY.