PERGAMON

Automatica 37 (2001) 1619–1628

automatica

www.elsevier.com/locate/automatica

Brief Paper

Constrained linear state estimation—a moving horizon approach☆

Christopher V. Rao[a], James B. Rawlings[b,*], Jay H. Lee[c]

[a]*Department of Bioengineering, University of California at Berkeley, USA*
[b]*Department of Chemical Engineering, University of Wisconsin-Madison, 1415 Engineering Drive, Madison, WI 53706-1691, USA*
[c]*Department of Chemical Engineering, Georgia Institute of Technology, USA*

Received 9 July 1999; revised 16 March 2000; received in final form 1 February 2001

Abstract

This article considers moving horizon strategies for constrained linear state estimation. Additional information for estimating state variables from output measurements is often available in the form of inequality constraints on states, noise, and other variables. Formulating a linear state estimation problem with inequality constraints, however, prevents recursive solutions such as Kalman filtering, and, consequently, the estimation problem grows with time as more measurements become available. To bound the problem size, we explore moving horizon strategies for constrained linear state estimation. In this work we discuss some practical and theoretical properties of moving horizon estimation. We derive sufficient conditions for the stability of moving horizon state estimation with linear models subject to constraints on the estimate. We also discuss smoothing strategies for moving horizon estimation. Our framework is solely deterministic.

Keywords: Constraints; State estimation; Optimization; Stability

1. Introduction

The Kalman filter is the standard choice for estimating the state of a linear system when the measurements are noisy and the process disturbances are unmeasured. One reason for the popularity of the Kalman filter is that it possesses many important theoretical properties such as stability. Often additional insight about the process is available in the form of inequality constraints. With the addition of inequality constraints, however, general recursive solutions such as Kalman filtering are unavailable. One strategy for determining an optimal state estimate is to reformulate the estimation problem as a quadratic program. This formulation allows for the natural addition of inequality constraints. While there exist many strategies to solve efficiently quadratic programs with the particular structure of the linear estimation problem (cf. Biegler, 1998), the problem grows without bound as we collect more measurements.

Building on the success of receding horizon control (for recent reviews, see Mayne, 1997; Lee & Cooley, 1997; Mayne, Rawlings, Rao, & Scokaert, 2000), moving horizon estimation (MHE) has been suggested as a practical strategy to incorporate inequality constraints in estimation (cf. Muske, Rawlings, & Lee, 1993; Muske & Rawlings, 1995; Robertson, Lee, & Rawlings, 1996; Tyler, 1997; Rao & Rawlings, 2000b). The basic strategy of MHE is to reformulate the estimation problem as a quadratic program using a moving, fixed-size estimation window. The fixed-size estimation window is necessary to bound the size of the quadratic program. Because only a subset of the data is considered, stability questions arise. The contribution of this article is that we prove stability for moving horizon estimation. We also briefly discuss smoothing strategies. The central theme of our analysis is the relationship between the full information estimation problem and its moving horizon approximation. This relationship is analyzed using forward dynamic programming and allows us to derive sufficient conditions for stability. Our stability results build on some of the general results of Rao and Rawlings (2000b).

☆This paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor Kenko Uchida under the direction of Editor Tamer Basar.

* Corresponding author. Tel.: + 608-263-5969; fax: + 608-265-8795.
E-mail address: jbraw@bevo.che.wisc.edu (J.B. Rawlings).

0005-1098/01/$ - see front matter
PII: S0005-1098(01)00115-7

2. Problem statement

Let the system generating the data sequence $\{y_k\}$ be modeled by the following linear, time-invariant, discrete-time system

$$x_{k+1} = Ax_k + Gw_k, \tag{1a}$$

$$y_k = Cx_k + v_k, \tag{1b}$$

where it is known that the states and disturbances satisfy the following constraints:

$$x_k \in \mathbb{X}, \quad w_k \in \mathbb{W}, \quad v_k \in \mathbb{V}.$$

We assume $x_k \in \mathbb{R}^n$, $y_k \in \mathbb{R}^p$, and $w_k \in \mathbb{R}^m$ and the sets $\mathbb{X}$, $\mathbb{W}$, and $\mathbb{V}$ are polyhedral and convex (i.e. $\mathbb{X} = \{x: Dx \leqslant d\}$) with $0 \in \mathbb{W}$ and $0 \in \mathbb{V}$. Let $x(k; z, \{w_j\})$ denote the solution of model (1) at time k subject to the initial condition z and disturbance sequence $\{w_j\}_{j=0}^{k-1}$:

$$x(k; z, \{w_j\}) := A^k z + \sum_{j=0}^{k-1} A^{k-j-1} G w_j.$$

We formulate the constrained linear state estimation problem as the solution to the following quadratic problem:

$$\phi_T^* = \min_{x_0, \{w_k\}_{k=0}^{T-1}} \phi_T(x_0, \{w_k\}) \tag{2}$$

subject to constraints

$$x_k \in \mathbb{X}, \quad w_k \in \mathbb{W}, \quad v_k \in \mathbb{V}, \tag{3}$$

where the objective function is defined by

$$\phi_T(x_0, \{w_k\}) = \sum_{k=0}^{T-1} v_k' R^{-1} v_k + w_k' Q^{-1} w_k + (x_0 - \hat{x}_0)' \Pi_0^{-1} (x_0 - \hat{x}_0),$$

$x_k := x(k; x_0, \{w_j\})$, and $v_k := y_k - Cx(k; x_0, \{w_j\})$. We assume the matrices Q, R, and Π_0 are symmetric positive definite. The pair $(\hat{x}_0, \Pi_0)$ summarizes the prior information at time $T = 0$ and is part of the data of the state estimation problem. We refer to this problem as the *full information estimator*, because we consider all of the available measurements. The solution to (2) at time T is the unique pair $(\hat{x}_{0|T-1}, \{\hat{w}_{k|T-1}\}_{k=0}^{T-1})$, and the optimal pair yields the state estimate $\{\hat{x}_{k|T-1}\}_{k=0}^{T-1}$, where

$$\hat{x}_{k|T-1} := x(k, \hat{x}_{0|T-1}, \{\hat{w}_k\}).$$

To simplify notation, let $\hat{x}_j := \hat{x}_{j|j-1}$, where $\hat{x}_{0|-1} := \hat{x}_0$.

3. Moving horizon approximation

Efficient strategies exist for solving the quadratic program (2). However, the problem size grows with time as the estimator processes more data. As a result, the problem complexity scales at least linearly with T. To make the estimation problem tractable, we need to bound the problem size. One strategy to reduce the problem (2) to a fixed dimension quadratic program is to employ a moving horizon approximation. The basic strategy of the moving horizon approximation is to consider explicitly a fixed amount of data, while approximately summarizing the old data not explicitly accounted for by the estimator. The key to preserving stability and performance is how one *approximately* summarizes the old data.

Consider again the full information problem (2). We can rearrange the objective function $\phi_T(\cdot)$ by breaking the time interval into two pieces: $t_1 = \{k: 0 \leqslant k \leqslant T - N - 1\}$ and $t_2 = \{k: T - N \leqslant k \leqslant T - 1\}$.

$$\phi_T(x_0, \{w_k\}_{k=0}^{T-1}) = \phi_{T-N}(x_0, \{w_k\}) + \sum_{k=T-N}^{T-1} v_k' R^{-1} v_k + w_k' Q^{-1} w_k.$$

By the Markov property of the system (1), the quantity

$$\sum_{k=T-N}^{T-1} v_k' R^{-1} v_k + w_k' Q^{-1} w_k$$

depends implicitly through the model (1) only on the state x_{T-N} and the decision variables w_k in the second time interval t_2. Exploiting this relation using forward dynamic programming, we can establish the equivalence between a full information problem and an estimation problem with a fixed size estimation window.

Consider the reachable set of states at time T generated by a feasible initial condition x_0 and disturbance sequence $\{w_k\}_{k=0}^{T-1}$:

$$\mathscr{R}_T = \left\{ x(T; x_0, \{w_j\}): \begin{array}{c} x_0 \in \mathbb{X}, \\ x(k; x_0, \{w_j\}) \in \mathbb{X} \text{ for } k = 0, \ldots, T, \\ w_k \in \mathbb{W} \text{ for } k = 0, \ldots, (T-1) \end{array} \right\}.$$

For $z \in \mathscr{R}_T$, we define the *arrival cost*[1] as

$$\theta_T(z) := \min_{x_0, \{w_k\}_{k=0}^{T-1}} \{\phi_T(x_0, \{w_j\}): x(T; x_0, \{w_k\}) = z\},$$

where the minimization is subject to the constraints (3). It follows that $\theta_0(z) = (z - \hat{x}_0)' \Pi_0^{-1} (z - \hat{x}_0)$. Arrival cost is a fundamental concept in MHE, because the following equivalence can be established simply using forward dynamic programming

$$\min_{x_0, \{w_k\}_{k=0}^{T-1}} \phi_T(x_0, \{w_k\}) \equiv \min_{z, \{w_k\}_{k=T-N}^{T-1}} \sum_{k=T-N}^{T-1} v_k' R^{-1} v_k + w_k' Q^{-1} w_k + \theta_{T-N}(z),$$

[1] Other researchers have used the term *cost to come* (cf. Başar & Bernhard, 1995) or *cost to arrive* (cf. Verdu & Poor, 1987).

where the minimizations are subject to the constraints (3), $x_k := x(k-T-N; z,\{w_j\})$, and $v_k := y_k - Cx(k-T-N; z,\{w_j\})$.

The arrival cost compactly summarizes the effect of the data $\{y_k\}_{k=0}^{T-N-1}$ on the state x_{T-N}, thereby allowing us to fix the dimension of the optimization. We can view arrival cost as the analogue to the *cost to go* in standard backward dynamic programming. Loosely speaking, in probabilistic terms, the arrival cost generates the conditional density function $p(x_{T-N}|y_0,\dots,y_{T-N-1})$ and vice versa: the arrival cost is proportional to the negative logarithm of the conditional density function $p(x_{T-N}|y_0,\dots,y_{T-N-1})$.[2] Hence, we may view arrival cost as an equivalent statistic (Striebel, 1965) for the conditional density function $p(x_{T-N}|y_0,\dots,y_{T-N-1})$.

If we are able to construct analytic expressions for the arrival cost, then it is possible to develop recursive estimators. One example is Kalman filtering. Consider the unconstrained estimation problem (2). If we use the Kalman filter covariance update formula (Jazwinski, 1970)

$$\Pi_T = GQG' + A\Pi_{T-1}A' - A\Pi_{T-1}C'(R + C\Pi_{T-1}C')^{-1}C\Pi_{T-1}A' \tag{4}$$

subject to the initial condition Π_0, then, assuming the matrix Π_T is invertible, we can express the arrival cost explicitly as

$$\theta_T(z) = (z - \hat{x}_T)'\Pi_T^{-1}(z - \hat{x}_T) + \phi_T^*,$$

where $\hat{x}_T$ denotes the optimal estimate at time T given the measurements $\{y_k\}_{k=0}^{T-1}$ and ϕ_T^* denotes the optimal cost at time T. From the preceding arguments, we have

$$\min_{x_0,\{w_k\}_{k=0}^{T-1}} \phi_T(x_0,\{w_k\}) \equiv \min_{z,\{w_k\}_{k=T-N}^{T-1}} \sum_{k=T-N}^{T-1} v_k' R^{-1} v_k + w_k' Q^{-1} w_k + (z - \hat{x}_{T-N})'\Pi_{T-N}^{-1}(z - \hat{x}_{T-N}) + \phi_{T-N}^*.$$

We can extract the Kalman filter by considering a horizon of $N = 1$. For this scenario, we have

$$\phi_T(z, w_{T-1}) = v_{T-1}' R^{-1} v_{T-1} + w_{T-1}' Q^{-1} w_{T-1} + (z - \hat{x}_{T-1})'\Pi_{T-1}^{-1}(z - \hat{x}_{T-1}).$$

Substituting in the model equation (1), evaluating the minimum with respect to w_{T-1} and x_{T-1}, and using some algebra, we obtain the well-known result

$$\hat{x}_T = A\hat{x}_{T-1} + A\Pi_{T-1}C'(R + C\Pi_{T-1}C')^{-1}(y_T - CA\hat{x}_{T-1})$$

for the Kalman filter.

Unfortunately, for the constrained problem, we are unable to generate an analytic expression for the arrival cost. Inequality constraints make the problem combinatorial, so general analytic expressions for the arrival cost are unavailable. One reasonable solution then is to approximate the arrival cost for the constrained problem with the arrival cost for the unconstrained problem. This choice has the desirable property that when the inequality constraints are inactive, the approximation is exact. Because we consider an approximation of the arrival cost, stability questions arise: does a poor choice of an approximate arrival cost lead to instability? As demonstrated by examples in Rao (2000), Rao and Rawlings (2000a), the answer is yes. Instability may result for some systems if the arrival cost is improperly approximated. In the next section, we discuss the details of the stability arguments. As we demonstrate, it is not necessary to generate explicitly an analytic expression for the arrival cost. Rather, as discussed in Rao and Rawlings (2000b), the approximate arrival cost needs only to satisfy an inequality.

We formulate MHE as the solution to the following quadratic program:

$$\hat{\phi}_T^* = \min_{z,\{w_k\}_{k=T-N}^{T-1}} \hat{\phi}_T(z,\{w_k\}) \tag{5}$$

subject to the constraints (3) where the objective function is defined by

$$\hat{\phi}_T(z,\{w_k\}) := \sum_{k=T-N}^{T-1} v_k' R^{-1} v_k + w_k' Q^{-1} w_k + (z - \hat{x}_{T-N}^{\mathrm{mh}})'\Pi_{T-N}^{-1}(z - \hat{x}_{T-N}^{\mathrm{mh}}) + \hat{\phi}_{T-N}^*,$$

$x_k := x(k-(T-N); z,\{w_j\})$ and $v_k := y_k - Cx(k-(T-N); z,\{w_j\})$. The MHE cost $\hat{\phi}_T^*$ approximates the full information cost ϕ_T^* by replacing the arrival cost $\theta_{T-N}(z)$ with the quadratic approximation $(z - \hat{x}_{T-N}^{\mathrm{mh}})'\Pi_{T-N}^{-1}(z - \hat{x}_{T-N}^{\mathrm{mh}}) + \hat{\phi}_{T-N}^*$. The pair $(\hat{x}_{T-N}^{\mathrm{mh}}, \Pi_{T-N})$ summarizes the prior information at time $T-N$. The vector $\hat{x}_{T-N}^{\mathrm{mh}}$ is the moving horizon state estimate at time $T-N$ and the matrix Π_{T-N} is the solution to (4) subject to the initial condition Π_0. For $T \leqslant N$, MHE is equivalent to the full information estimator: $\hat{\phi}_T(\cdot) = \phi_T(\cdot)$. We assume at this point that the matrix Π_{T-N} is invertible; conditions for nonsingularity are discussed later. The solution to (2) at time T is the unique pair $(z^*,\{\hat{w}_{k|T-1}^{\mathrm{mh}}\}_{k=T-N}^{T-1})$, and the optimal pair yields the state estimate $\{\hat{x}_{k|T-1}^{\mathrm{mh}}\}_{k=T-N}^{T-1}$, where

$$\hat{x}_{k|T-1}^{\mathrm{mh}} := x(k-(T-N); z^*,\{\hat{w}_j^{\mathrm{mh}}\}).$$

[2] For example, if the conditional density function is normally distributed (i.e. $p(x_{T-N}|y_0,\dots,y_{T-N-1}) \sim N(\hat{x}_{T-N},\Pi_{T-N})$), then $-\log(p(x_{T-N}|y_0,\dots,y_{T-N-1})) \propto (x_{T-N} - \hat{x}_{T-N})'\Pi_{T-N}^{-1}(x_{T-N} - \hat{x}_{T-N})$.

To simplify notation, let $\hat{x}_j^{\mathrm{mh}} := \hat{x}_{j|j-1}^{\mathrm{mh}}$, where $\hat{x}_{0|-1}^{\mathrm{mh}} := \hat{x}_0$. This formulation of MHE was first proposed by Muske et al. (1993) and Robertson et al. (1996).

4. Stability analysis

When the inequality constraints (3) are not present, the solution to the quadratic program (2) may be obtained analytically, yielding the Kalman filter. The relationship between least squares and the Kalman filter is well known (cf. Bryson & Frazier, 1963; Rauch, Tung, & Striebel, 1965). Even with the addition of constraints, the estimator enjoys analogous stability properties. In particular, the constrained estimator is stable in the sense of an observer. The following discussion of observer stability is premised on classical Lyapunov stability theory for dynamical systems. The concepts are completely analogous to their classical counterpart. To account for constraints, we have modified the definition of stability in an analogous manner to Keerthi and Gilbert (1988).

Definition 1. The estimator is an *asymptotically stable observer* for the system

$$x_{k+1} = Ax_k, \quad y_k = Cx_k \tag{6}$$

if for any $\varepsilon > 0$ there corresponds a number $\delta > 0$ and a positive integer $\bar{T}$ such that if $\|x_0 - \hat{x}_0\| \leqslant \delta$ and $\hat{x}_0 \in \mathbb{X}$, then $\|\hat{x}_T - A^T x_0\| \leqslant \varepsilon$ for all $T \geqslant \bar{T}$ and $\hat{x}_T \to A^T x_0$ as $T \to \infty$.

The implications of constraints on the estimator are more subtle than for the regulator. In particular, the estimator has no control over the evolution of the state of the system. A poor choice of constraints may prevent convergence to the true state of the system (6). For a more detailed discussion of constraints, see Rao (2000). One solution is to require that the evolution of the system (6) respects the constraint $\mathbb{X}$ (i.e. $A^k x_0 \in \mathbb{X}$ for $k \geqslant 0$). While this assumption is reasonable, the constraints need to satisfy only the following weaker assumption to prove stability.

(I) Suppose the system (6) with initial condition x_0 generates the data (i.e. $y_k = CA^k x_0$). We assume there exists $x_{0|\infty}$, $\{w_{k|\infty}\}_{k=0}^{\infty}$, and $\sigma > 0$ such that

$$\sum_{k=0}^{\infty} v_{k|\infty}' R^{-1} v_{k|\infty} + w_{k|\infty}' Q^{-1} w_{k|\infty} + (x_{0|\infty} - \hat{x}_0)' \Pi_0^{-1} (x_{0|\infty} - \hat{x}_0) \leqslant \sigma \|x_0 - \hat{x}_0\|^2$$

and

$$x_{k|\infty} \in \mathbb{X}, \quad w_{k|\infty} \in \mathbb{W}, \quad v_{k|\infty} \in \mathbb{V},$$

where $x_{k|\infty} := x(k; x_{0|\infty}, \{w_{j|\infty}\})$ and $v_{k|\infty} := y_k - Cx(k; x_{0|\infty}, \{w_{j|\infty}\})$.

Assumption **I** states that if we consider an infinite amount of data generated by the system (6), then there exists a feasible state and disturbance trajectory that yields bounded cost. It is straightforward to demonstrate that assumption **I** is a weaker assumption: if we choose $\sigma = \lambda_{\max}(\Pi_0^{-1})$, then assumption **I** follows if we assume the evolution (6) respects the constraints $\mathbb{X}$. Recall, by assumption, $0 \in \mathbb{W}$ and $0 \in \mathbb{V}$.

Assumption **I** is also a sufficient condition for the existence of a solution to the quadratic programs (2) and (5). The upper bound σ is necessary to prove stability. Without this bound, we have no reference with which to construct a Lyapunov function. Unlike regulation where we have a strictly monotone nonincreasing cost function that is bounded below by zero, we have a strictly monotone nondecreasing cost function in estimation that is not necessarily bounded above (e.g. consider the case when assumption **I** is violated). The role of σ is to provide this upper bound when constraints prevent the estimator from tracking the system perfectly. Otherwise, without constraints, we can readily generate the upper bound with $\sigma = \lambda_{\max}(\Pi_0^{-1})$ (i.e. the cost of tracking the system perfectly).

Before discussing the stability of the MHE, we first state the following stability result for the full information estimator.

Proposition 2. *Suppose the matrices Q, R, and Π_0 are positive definite, (C, A) is observable, and assumption* **I** *holds. Then, the constrained full information estimator is an asymptotically stable observer for the system* (6).

Proof. See Muske et al. (1993). □

To establish asymptotic stability for MHE, we require the following lemmas.

Lemma 3. *Suppose (C, A) is observable and $N \geqslant n$. If*

$$\sum_{k=T-N}^{T-1} \hat{v}_{k|T-1}' R^{-1} \hat{v}_{k|T-1} + \hat{w}_{k|T-1}' Q^{-1} \hat{w}_{k|T-1} \to 0$$

then $\|\hat{x}_T - x_T\| \to 0$.

Proof. The proof is omitted for brevity. See Rao (2000). □

Lemma 4. *The Kalman filter covariance matrix Π_T satisfies the following inequality for all $p \in \mathscr{R}_T$:*

$$(p - \hat{x}_T^{\mathrm{mh}})' \Pi_T^{-1} (p - \hat{x}_T^{\mathrm{mh}}) + \hat{\phi}_T^* \leqslant \min_{z, \{w_k\}_{k=T-N}^{T-1}} \{\hat{\phi}_T(z, \{w_k\}) : x(N; z, \{w_j\}) = p\},$$
$$:= \hat{\theta}_T(p),$$

where the minimization is subject to the constraints (3).

Proof. The proof is in Appendix A. □

Before we establish stability, we state conditions that guarantee the matrix Π_T is positive definite (invertible). If we assume that (C, A) is detectable and $(A, GQ^{-1/2})$ is controllable, then

$$\lim_{T \to \infty} \Pi_T = \Pi_\infty,$$

where $\Pi_\infty > 0$ is the unique steady-state solution to the Riccati equation (4) (de Souza, Gevers, & Goodwin, 1986). If we choose $\Pi_0 \geqslant \Pi_\infty$, then Π_k is positive definite for all $k \geqslant 0$ (Bitmead, Gevers, Petersen, & Kaye, 1985). As an alternative, if the matrix G is nonsingular (in which case GQG^T is positive definite), then Π_k is also positive definite for all $k \geqslant 0$.

Proposition 5. *Suppose the matrices Q, R, and Π_0 are positive definite, (C, A) is observable, assumption* **I** *holds, $N \geqslant n$, and either*

(i) *The matrix G is nonsingular, or*
(ii) *$(A, GQ^{-1/2})$ is controllable and $\Pi_0 \geqslant \Pi_\infty$.*

Then the constrained moving horizon estimator is an asymptotically stable observer for the system (6).

Proof. We begin by demonstrating convergence. An optimal solution to (5) exists (Frank & Wolfe, 1956), because the problem (5) is a convex quadratic program and the feasible region is not empty: the pair $x_{T-N|\infty}$ and $\{w_{k|\infty}\}_{k=T-N}^{T-1}$ is feasible. By definition

$$\hat{\phi}_T^* - \hat{\phi}_{T-N}^* \geqslant \sum_{k=T-N}^{T-1} \hat{v}_{k|T-1}^{\mathrm{mh}\prime} R^{-1} \hat{v}_{k|T-1}^{\mathrm{mh}} + \hat{w}_{k|T-1}^{\mathrm{mh}\prime} Q^{-1} \hat{w}_{k|T-1}^{\mathrm{mh}},$$

where $\hat{v}_{k|T-1}^{\mathrm{mh}} := y_k - C\hat{x}_{k|T-1}^{\mathrm{mh}}$. To demonstrate $\sigma\|x_0 - \hat{x}_0\|^2$ is a uniform bound, we proceed using an induction argument. For $T \leqslant N$, we have by optimality

$$\hat{\phi}_T^* \leqslant \theta_T(x_{T|\infty}) \leqslant \sigma\|x_0 - \hat{x}_0^{\mathrm{mh}}\|^2.$$

For $T \geqslant N$, Lemma 4 guarantees

$$\theta(x_{T|\infty}) = \hat{\theta}(x_{T|\infty}) \geqslant (x_{T|\infty} - \hat{x}_T)'\Pi_T^{-1}(x_{T|\infty} - \hat{x}_T) + \hat{\phi}_T^*.$$

Let us now assume, for $T > N$,

$$\theta(x_{T-N|\infty}) \geqslant (x_{T-N|\infty} - \hat{x}_{T-N})'\Pi_{T-N}^{-1} \times (x_{T-N|\infty} - \hat{x}_{T-N}) + \hat{\phi}_{T-N}^*$$

for the induction argument. Utilizing the optimality principle, the induction assumption, and properties of the arrival cost, for all $T \geqslant N$,

$$\begin{aligned}
&\sigma\|x_0 - \hat{x}_0\|^2 \\
&\geqslant \min_{z,\{w_k\}_{k=T-N}^{T-1}} \Big\{ \sum_{k=T-N}^{T-1} v_k' R^{-1} v_k + w_k' Q^{-1} w_k \\
&\quad + \hat{\theta}_{T-N}(z) : x(N; z, \{w_j\}) = x_{T|\infty} \Big\} \\
&\quad \text{(by optimality and assumption } \mathbf{I}) \\
&\geqslant \min_{z,\{w_k\}_{k=T-N}^{T-1}} \Big\{ \sum_{k=T-N}^{T-1} v_k' R^{-1} v_k + w_k' Q^{-1} w_k \\
&\quad + (z - \hat{x}_{T-N}^{\mathrm{mh}})'\Pi_{T-N}^{-1}(z - \hat{x}_{T-N}^{\mathrm{mh}}) \\
&\quad + \hat{\phi}_{T-N}^* : x(N; z, \{w_j\}) = x_{T|\infty} \Big\} \\
&\quad \text{(by the induction assumption)} \\
&\geqslant (x_{T|\infty} - \hat{x}_T^{\mathrm{mh}})'\Pi_T^{-1}(x_{T|\infty} - \hat{x}_T^{\mathrm{mh}}) + \hat{\phi}_T^* \\
&\quad \text{(by Lemma 4)} \\
&\geqslant \hat{\phi}_T^*,
\end{aligned}$$

where both minimizations are subject to the constraints (3). Hence, the sequence $\{\hat{\phi}_T^*\}$ is monotone nondecreasing and bounded above by $\sigma\|x_0 - \hat{x}_0^{\mathrm{mh}}\|^2$. Convergence implies

$$\sum_{k=T-N}^{T-1} \hat{v}_{k|T-1}^{\mathrm{mh}\prime} R^{-1} \hat{v}_{k|T-1}^{\mathrm{mh}} + \hat{w}_{k|T-1}^{\mathrm{mh}\prime} Q^{-1} \hat{w}_{k|T-1}^{\mathrm{mh}} \to 0$$

as $T \to \infty$. Lemma 3 guarantees the estimation error $\|\hat{x}_T^{\mathrm{mh}} - A^T x_0\| \to 0$ as $T \to \infty$.

To prove stability, let $\varepsilon > 0$ and choose $\varrho > 0$ sufficiently small for $T = N$ as specified by Lemma 3. If we choose $\delta > 0$ such that $\sigma\delta^2 < \varrho$, then we obtain the following inequality for all $T \geqslant N$:

$$\begin{aligned}
\sigma\delta^2 &\geqslant \sum_{k=T-N}^{T-1} \hat{v}_{k|T-1}^{\mathrm{mh}\prime} R^{-1} \hat{v}_{k|T-1}^{\mathrm{mh}} + \hat{w}_{k|T-1}^{\mathrm{mh}\prime} Q^{-1} \hat{w}_{k|T-1}^{\mathrm{mh}} \\
&\quad + (\hat{x}_{T-N|T-1}^{\mathrm{mh}} - \hat{x}_{T-N}^{\mathrm{mh}})'\Pi_{T-N}^{-1}(\hat{x}_{T-N|T-1}^{\mathrm{mh}} - \hat{x}_{T-N}^{\mathrm{mh}}) \\
&\quad + \hat{\phi}_{T-N}^* \\
&\geqslant \sum_{k=T-N}^{T-1} \hat{v}_{k|T-1}^{\mathrm{mh}\prime} R^{-1} \hat{v}_{k|T-1}^{\mathrm{mh}} + \hat{w}_{k|T-1}^{\mathrm{mh}\prime} Q^{-1} \hat{w}_{k|T-1}^{\mathrm{mh}}.
\end{aligned}$$

Hence, if the initial estimation error $\|x_0 - \hat{x}_0^{\mathrm{mh}}\| \leqslant \delta$, then the estimation error $\|\hat{x}_T^{\mathrm{mh}} - A^T x_0\| \leqslant \varepsilon$ for all $T \geqslant \tilde{T} = N$ as claimed. □

Remark 6. When inequality constraints are not included, MHE is equivalent to the Kalman filter. Proposition 5, therefore, establishes that the Kalman filter is stable under the stated conditions.

We may also formulate the constrained steady-state MHE where the objective function is now defined as

$$\hat{\phi}^{\infty}_T(z,\{w_k\}) := \sum_{k=T-1}^{T-N} v_k' R^{-1} v_k + w_k' Q^{-1} w_k + (z - \hat{x}_{T-N})' \Pi_\infty^{-1} (z - \hat{x}_{T-N}) + \hat{\phi}^*_{T-N}.$$

For $T \leqslant N$, we choose $\hat{\phi}^{\infty}_T(\cdot) = \phi_T(\cdot)$ with $\Pi_0 = \Pi_\infty$. Demonstrating the stability of steady-state MHE is immediate. In Proposition 5, we proved stability for all $\Pi_0 > 0$. If we choose $\Pi_0 = \Pi_\infty$, then $\Pi_T = \Pi_\infty$ for all T. We state this result as the following corollary to Proposition 5.

Corollary 7. *Suppose the matrices Q and R are positive definite, (C, A) is observable, $(A, GQ^{-1/2})$ is controllable, assumption* **I** *holds, and $N \geqslant n$. Then the constrained steady-state moving horizon estimator is an asymptotically stable observer for the system.*

5. Smoothing update

In our development of the MHE, we use a filter update to summarize the past information. With the filter update we transfer the prior information to current estimate window by conditioning the estimates at time T using $\hat{x}_{T-N}$. The conditioning is the result of the approximate arrival cost

$$(x_{T-N} - \hat{x}_{T-N})' \Pi_{T-N}^{-1} (x_{T-N} - \hat{x}_{T-N})$$

achieving its minimum at $\hat{x}_{T-N}$. A schematic of the filter update strategy is shown in Fig. 1.

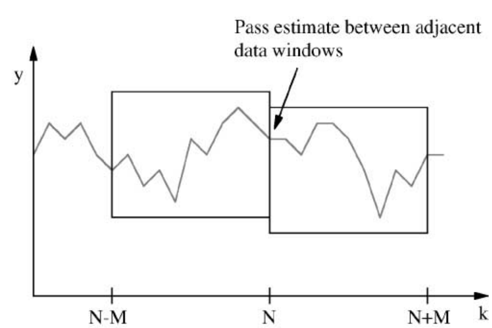


Fig. 1. A diagram of the filter update strategy for passing information forward in time.

Rather than conditioning the estimate at time T on $\hat{x}_{T-N}$, we may also condition the estimate on $\hat{x}_{T-N|T-2}$. With the filter update, we ignore the influence of the data $\{y_k\}_{T-N}^{T-2}$ on our knowledge of x_{T-N}. A diagram of the smoothing update strategy is shown in Fig. 2. This problem was first studied by Findeisen (1997).

For unconstrained systems, we have an algebraic expression for the "smoothed" arrival cost.

Lemma 8. *Suppose the matrix $\Pi_{j|T-1}$ is positive definite. Then, we have for $j < T$*

$$(z - \hat{x}_{j|T-1})' \Pi_{j|T-1}^{-1} (z - \hat{x}_{j|T-1}) + \phi_T^* = \min_{x_0, \{w_k\}_{k=0}^{T-1}} \{\phi_T(x_0, \{w_k\}) : x(j; x_0, \{w_k\} = z\}, \tag{7}$$

where

$$\Pi_{k|T} = \Pi_{k|k} + \Pi_{k|k} A_k' \Pi_{k+1|T}^{-1} \times (\Pi_{k+1|T} - \Pi_{k+1|k}) \Pi_{k+1|T}^{-1} A_k \Pi_{k|k},$$

$$\Pi_{k|k} = \Pi_k - \Pi_k C' (R + C \Pi_k C')^{-1} C \Pi_k \tag{8}$$

and $\Pi_{k|k-1} := \Pi_k$.

Proof. This equality follows from the smoothing results for linear discrete-time system (cf. Rauch et al., 1965; Bryson & Ho, 1975). □

We formulate MHE with the smoothing update by using the objective function

$$\hat{\phi}_T(z, \{w_k\}) = \sum_{k=T-N}^{T-1} w_k' Q^{-1} w_k + v_k' R^{-1} v_k + \Gamma_{T-N}(z) + \hat{\phi}^*_{T-N},$$

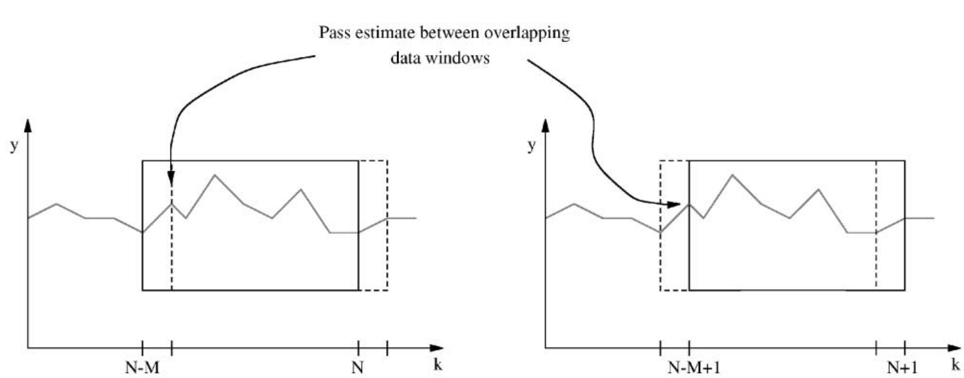


Fig. 2. A diagram of the smoothing strategy for passing information forward in time.

where

$$\Gamma_{T-N}(z) = (z - \hat{x}_{T-N|T-2})' \Pi^{-1}_{T-N|T-2} (z - \hat{x}_{T-N|T-2})$$
$$- (\mathscr{Y}_T^{N-2} - \mathcal{O}_{N-2} z)' W^{-1}_{N-2} (\mathscr{Y}_T^{N-2} - \mathcal{O}_{N-2} z)$$
$$+ (\hat{\phi}^*_{T-1} - \hat{\phi}^*_{T-N})$$

and

$$\mathscr{Y}_T^{N-2} := [y'_{T-N}, \quad y'_{T-N-1}, \quad \ldots, \quad y'_{T-2}]'.$$

Expressions for $\mathcal{O}_{N-2}$ and W_{N-2} are given in Appendix B.

To prove stability, it suffices in light of Proposition 5 to demonstrate that $\Gamma_T(\cdot)$ satisfies the inequality in Lemma 4.

Lemma 9. *Suppose the matrix $\Pi_{T-N|T-2}$ is positive definite. Then, for all $p \in \mathscr{R}_T$ and $j < T$*

$$\Gamma_T(p) + \hat{\phi}^*_T \leqslant \min_{z, \{w_k\}_{k=T-N}^{T-1}} \{\hat{\phi}_T(z, \{w_k\}) : (N; z, \{w_k\}) = p\},$$
$$:= \hat{\theta}_T(p).$$

where the minimization is subject to the constraints (3).

Proof. The proof is available in the Appendix C. □

Corollary 10. *Suppose the matrices Q, R, Π_0, and $\Pi_{T-N|T-2}$ for all $T \geqslant N$ are positive definite, (C, A) is observable, assumption* **I** *holds, $N \geqslant n$. Then the constrained moving horizon estimator with a smoothing update is an asymptotically stable observer for the system* (6).

6. Conclusion

We have demonstrated that moving horizon estimation (MHE) is a practical strategy for constrained state estimation. Three separate formulations were presented. The key result of this work is that if the full information estimator is stable, then MHE is also stable provided one does not introduce extra bias with the prior information. To characterize this condition, we analyzed the estimation problem using forward dynamic programming and the notion of arrival cost.

We believe one of the strengths of moving horizon estimation is that the cost function is designed to address the practical engineering tradeoff between following the model forecast and tracking the measurement. The novelty is that MHE allows further information to be included in the estimator in the form of constraints, which may provide a useful design feature that practitioners can exploit. Nominal estimator stability, although not a primary design goal, is guaranteed automatically by proper choice of arrival cost approximation and system observability. This fact leaves the designer free to adjust the tuning parameters to achieve other objectives such as rapid state reconstruction, low sensitivity to sensor noise, optimality for various assumed probability distributions, and robustness to various types of model error. These properties should make MHE useful to practicing engineers.

The strength and weakness of MHE is the use of quadratic programming. For reasonable models, the optimization problems can be solved in less than 1 s on desktop computers using standard software. However, for some problems this performance is insufficient. With the increasing power of computers and the ability to solve quadratic programs in polynomial time, MHE will become an alternative for an expanding class of estimation problems in the near future.

Acknowledgements

The authors gratefully acknowledge the financial support of the industrial members of the Texas-Wisconsin Modeling and Control Consortium and NSF through grant #CTS-9708497. The authors express their thanks to Professor D.Q. Mayne for helpful discussions and feedback about this work. The authors also acknowledge P.K. Findeisen, who studied some of the issues presented in this paper as part of a Master's research project (Findeisen 1997) at the University of Wisconsin-Madison. We are grateful to four referees of the original version of this paper, whose insightful comments improved the paper. An extended version of this paper is available at http://www.che.wisc.edu/ ~ rao.

Appendix A. Proof of Lemma 4

Before proving Lemma 4, we first establish the following lemma concerning general quadratic programs.

Lemma 11. *Let $\theta(z) = z'Qz$ where the matrix Q is symmetric positive definite and the sets Γ and Ω are closed and convex with $\Gamma \subseteq \Omega$. If a solution exists to the following quadratic programs $\theta(\hat{z}) = \min_{z \in \Omega} \theta(z)$, and $\theta(\bar{z}) = \min_{z \in \Gamma} \theta(z)$, then $\theta(\bar{z}) \geqslant \theta(\hat{z}) + \theta(\Delta z)$ where $\Delta z = \bar{z} - \hat{z}$.*

Proof. Substituting in for $\bar{z}$, we obtain

$$\theta(\bar{z}) = \theta(\hat{z} + \Delta z)$$
$$= \theta(\hat{z}) + \langle \nabla\theta(\hat{z}), \Delta z \rangle + \theta(\Delta z).$$

Optimality implies $\langle \nabla\theta(\hat{z}), z - \hat{z}\rangle \geqslant 0$ for every $z \in \Omega$.[3] This inequality implies $\theta(\bar{z}) \geqslant \theta(\hat{z}) + \theta(\Delta z)$ as claimed. □

Proof (*Lemma* 4). Without loss of generality, we take $\hat{x}_{T-N} = 0$. Consider an arbitrary $p \in \mathscr{R}_T$. Let

$$(\bar{x}_{T-N|T-1}, \{\bar{w}_{k|T-1}\}_{k=T-N}^{T-1}) = \arg\min_{z,\{w_k\}_{k=T-N}^{T-1}} \{\hat{\phi}_T(z,\{w_k\}) : x(N;z,\{w_j\}) = p\},$$

where the minimization is subject to the constraints (3). If

$$\Delta x_{T-N|T-1} := \bar{x}_{T-N|T-1} - \hat{x}_{T-N|T-1},$$

$$\Delta w_{k|T-1} := \bar{w}_{k|T-1} - \hat{w}_{k|T-1}$$

then, by Lemma 11, we have

$$\hat{\theta}_T(p) \geqslant \hat{\phi}_T^* + \hat{\phi}_T(\Delta x_{T-N|T-1}, \{\Delta w_{k|T-1}\}).$$

If we choose $p = \hat{x}_T$, then $\Delta x_{T-N|T-1} = 0$, $\Delta w_{k|T-1} = 0$, and

$$\hat{\phi}_T(\Delta x_{T-N|T-1}, \{\Delta w_{k|T-1}\}) = 0.$$

Let $\Delta p := p - \hat{x}_T$. We obtain, therefore, the following inequality:

$$\hat{\phi}_T(\Delta x_{T-N|T-1}, \{\Delta w_{k|T-1}\}) \geqslant \min_{\Delta z,\{\Delta w_j\}} \{\hat{\phi}(\Delta z,\{\Delta w_j\}) : x(N;\Delta z,\{\Delta w_j\}) = \Delta p\}, = (\Delta p)'\Pi_T^{-1}(\Delta p)$$

and the lemma follows as claimed. □

Appendix B. Formulae for smoothing covariance

For notational simplicity, we make the following identities:

$$\mathscr{O}_{N-2} := [C' \quad A'C' \cdots A^{(N-2)'}C']',$$

$$\mathscr{Y}_T^{N-2} := \begin{bmatrix} y_{T-N} \\ y_{T-N+1} \\ \cdots \\ y_{T-2} \end{bmatrix}$$

and

$$W_{N-2} = \begin{bmatrix} R & 0 & 0 & \cdots & 0 \\ 0 & CGQG'C' + R & CGQG'A'C' & \cdots & CGQG'A^{(N-3)'}C' \\ 0 & CAGQG'C' & C(GQG' + AGQG'A')C' + R & & \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & CA^{(N-3)}GQG'C' & CA^{(N-3)}GQG'A'C' & \cdots & C(\sum_{k=0}^{N-3} A^k GQG'A^{k'})C' + R \end{bmatrix}.$$

Appendix C. Proof of Lemma 9

Proof. In light of Lemma 4, it suffices to demonstrate

$$\Gamma_{T-N}(p) = (p - \hat{x}_{T-N})'\Pi_{T-N}^{-1}(p - \hat{x}_{T-N}).$$

From Lemma 8, we have the following equality:

$$\begin{aligned} &(p - \hat{x}_{T-N|T-2})'\Pi_{T-N|T-2}^{-1}(p - \hat{x}_{T-N|T-2}) + \phi_{T-1}^* \\ &= \min_{\{w_k\}_{k=T-N}^{T-2}} \Bigg\{ \sum_{k=T-N}^{T-2} v_k'R^{-1}v_k + w_k'Q^{-1}w_k \\ &\quad + (p - \hat{x}_{T-N})'\Pi_{T-N}^{-1}(p - \hat{x}_{T-N}) : v_k \\ &= y_k - Cx(k-(T-N);p,\{w_j\})\Bigg\} + \hat{\phi}_{T-N}^* \\ &= \min_{\{w_j\}_{T-N}^{T-2}} \Bigg\{ \sum_{k=T-N}^{T-2} v_k'R^{-1}v_k + w_k'Q^{-1}w_k : v_k \\ &\quad = y_k - Cx(k-(T-N);p,\{w_j\})\Bigg\} \\ &\quad + (p - \hat{x}_{T-N})'\Pi_{T-N}^{-1}(p - \hat{x}_{T-N}) + \hat{\phi}_{T-N}^*. \end{aligned}$$

Let

$$D_{N-2}(p) := \min_{\{w_j\}_{T-N}^{T-2}} \Bigg\{ \sum_{k=T-N}^{T-2} v_k'R^{-1}v_k + w_k'Q^{-1}w_k : v_k = y_k - Cx(k-(T-N);p,\{w_j\})\Bigg\}.$$

We may evaluate $D_{N-2}(p)$ using induction. Consider

$$D_1(p) = \min_{\{w_j\}_{T-3}^{T-2}} \Bigg\{ \sum_{k=T-3}^{T-2} v_k'R^{-1}v_k + w_k'Q^{-1}w_k : v_k = y_k - Cx(k-(T-3);p,\{w_j\})\Bigg\}.$$

[3] A proof by contradiction is immediate—assume there exists a z in Ω that violates the above condition and consider a convex combination between z and $\hat{z}$, which lies in Ω, and calculate cost; it decreases from $\theta(\hat{z})$ along the line, contradicting optimality. In other words, $-\nabla\theta(\hat{z}) \in T_\Omega(\hat{z})$ where $T_\Omega(\hat{z})$ denotes the normal cone to Ω at $\hat{z}$:

$$T_\Omega(\hat{z}) = \{z : \langle z^1 - \hat{z}, z\rangle \leqslant 0, \forall z^1 \in \Omega\}.$$

Evaluating the minimization analytically, we obtain

$$D_1(p) = \left(\begin{bmatrix} y_{T-3} \\ y_{T-2} \end{bmatrix} - \begin{bmatrix} C \\ CA \end{bmatrix} p\right)' \begin{bmatrix} R & 0 \\ 0 & R + CGQG'C' \end{bmatrix}^{-1} \times \left(\begin{bmatrix} y_{T-3} \\ y_{T-2} \end{bmatrix} - \begin{bmatrix} C \\ CA \end{bmatrix} p\right)$$

$$= \left(\begin{bmatrix} y_{T-3} \\ y_{T-2} \end{bmatrix} - \begin{bmatrix} C \\ CA \end{bmatrix} p\right)' W_1^{-1} \left(\begin{bmatrix} y_{T-3} \\ y_{T-2} \end{bmatrix} - \begin{bmatrix} C \\ CA \end{bmatrix} p\right).$$

Now assume

$$D_{N-3}(p) = (\mathscr{Y}_T^{N-3} - \mathscr{O}_{N-3} p)' W_{N-3}^{-1} (\mathscr{Y}_T^{N-3} - \mathscr{O}_{N-3} p)$$

for the induction hypothesis and consider $D_{N-1}(p)$. A standard dynamic programming decomposition leads to the following reformulation:

$$D_{N-2}(p) = \min_{w_{T-N}} \left\{ D_{N-3}(z) + v'_{T-N} R^{-1} v_{T-N} + w'_{T-N} Q^{-1} w_{T-N} : \begin{array}{c} v_{T-N} = Cp \\ z = x(1; z; w_{T-N}) \end{array} \right\}.$$

From the result concerning $D_1(p)$, we have the expression

$$D_{N-2}(p) = (\mathscr{Y}_T^{N-2} - \mathscr{O}_{N-2} p)' W_{N-2}^{-1} (\mathscr{Y}_T^{N-2} - \mathscr{O}_{N-2} p).$$

So

$$\Gamma_{T-N}(p) = (p - \hat{x}_{T-N|T-2})' \Pi_{T-N|T-2}^{-1} (p - \hat{x}_{T-N|T-2}) - D_{N-2}(p) + (\hat{\phi}^*_{T-1} - \hat{\phi}^*_{T-N}).$$

By inspection

$$\Gamma_{T-N}(p) = (p - \hat{x}_{T-N})' \Pi_{T-N}^{-1} (p - \hat{x}_{T-N})$$

and the lemma follows as claimed. □

References

Başar, T., & Bernhard, P. (1995). *H^∞—Optimal control and related minimax design problems: A dynamic game approach*. Boston: Birkauser.

Biegler, L. T. (1998). Efficient solution of dynamic optimization and NMPC problems. *Proceedings of the international symposium on nonlinear model predictive control*, Switzerland: Ascona.

Bitmead, R. R., Gevers, M. R., Petersen, I. R., & Kaye, R. J. (1985). Monotonicity and stabilizability properties of solutions of the Riccati difference equation: Propositions, lemmas, theorems, fallacious conjectures and counterexamples. *System Control Letters*, *5*, 309–315.

Bryson, A. E., & Frazier, M. (1963). Smoothing for linear and nonlinear dynamic systems. *Proceedings of the Optimum Systems Synthesis Conference*, U.S. Air Force Technical Report ASD-TDR-063-119.

Bryson, A. E., & Ho, Y. (1975). *Applied optimal control*. New York: Hemisphere Publishing.

de Souza, C. E., Gevers, M. R., & Goodwin, G. C. (1986). Riccati equation in optimal filtering of nonstabilizable systems having singular state transition matrices. *IEEE Transactions on Automatic Control*, *31*(9), 831–838.

Findeisen, P. (1997). Moving horizon state estimation of discrete time systems. Master's thesis, University of Wisconsin-Madison.

Frank, M., & Wolfe, P. (1956). An algorithm for quadratic programming. *Naval Research Logistics Quarterly*, *3*, 95–110.

Jazwinski, A. H. (1970). *Stochastic processes and filtering theory*. New York: Academic Press.

Keerthi, S. S., & Gilbert, E. G. (1988). Optimal infinite-horizon feedback laws for a general class of constrained discrete-time systems: Stability and moving-horizon approximations. *Journal of Optimization Theory and Applications*, *57*(2), 265–293.

Lee, J. H., & Cooley, B. (1997). Recent advances in model predictive control and other related are as. In J.C. Kantor, C.E. García, & B. Carnahan (Eds.), *Proceedings of Chemical Process Control—V* (pp. 201–216). CACHE, A.I.Ch.E.

Mayne, D. Q. (1997). Nonlinear model predictive control: An assessment. In J.C. Kantor, C.E. García, & B. Carnahan (Eds.), *Proceedings of chemical process control—V* (pp. 217–231). CACHE, A.I.Ch.E.

Mayne, D. Q., Rawlings, J. B., Rao, C. V., & Scokaert, P. O. M. (2000). Constrained model predictive control: Stability and optimality. *Automatica*, *36*(6), 789–814.

Muske, K. R., & Rawlings, J. B. (1995). Nonlinear moving horizon state estimation. In R. Berber (Ed.), *Methods of model based process control*, Nato Advanced Study Institute series: E Applied Sciences, Vol. 293 (pp. 349–365). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Muske, K. R., Rawlings, J. B., & Lee, J. H. (1993). Receding horizon recursive state estimation. *Proceedings of the 1993 American control conference*, (pp. 900–904). June 1993.

Rao, C. V. (2000). Moving horizon strategies for the constrained monitoring and control of nonlinear discrete-time systems. Ph.D. thesis, University of Wisconsin-Madison.

Rao, C. V., & Rawlings, J. B. (2000a). Constrained process monitoring: A moving horizon approach. *A.I.Ch.E. Journal*, submitted for publication.

Rao, C. V., & Rawlings, J. B. (2000b). Nonlinear moving horizon estimation. In F. Allgöwer, & A. Zheng (Eds.), *Nonlinear model predictive control, Progress in systems and control theory*, Vol. 26 (pp. 45–69). Basel: Birkhäuser.

Rauch, H. E., Tung, F., & Striebel, C. T. (1965). Maximum likelihood estimates of linear dynamic systems. *AIAA Journal*, *3*(8), 1445–1450.

Robertson, D. G., Lee, J. H., & Rawlings, J. B. (1996). A moving horizon-based approach for least-squares state estimation. *A.I.Ch.E. Journal*, *42*(8), 2209–2224.

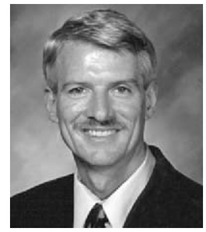
Striebel, C. (1965). Sufficient statistics in the optimum control of stochastic systems. *Journal of Mathematical Analysis and Application*, *12*, 576–592.

Tyler, M. L. (1997). Performance monitoring and fault detection in control systems. Ph.D. thesis. California Institute of Technology.

Verdu, S., & Poor, H. V. (1987). Abstract dynamic programming models under commutativity conditions. *SIAM Journal on Control and Optimization*, *25*, 990–1006.

Christopher V. Rao received the B.S. from Carnegie Mellon University in 1994 and the Ph.D. from the University of Wisconsin in 2000, both in Chemical Engineering. He is currently a post doctoral researcher in the Department of Bioengineering at the University of California, Berkeley and Lawrence Berkeley National Laboratory. His current research interests are in the areas of computational molecular biology and signal transduction.

Jay H. Lee was born in Seoul, Korea, in 1965. He obtained his B.S. degree in Chemical Engineering from the University of Washington, Seattle, in 1986, and his Ph.D. degree in Chemical Engineering from California Institute of Technology, Pasadena, in 1991. From 1991 to 1998, he was with the Department of Chemical Engineering at Auburn University, AL, as an Assistant Professor and an Associate Professor. From 1998 to 2000, he was with School of Chemical Engineering at Purdue University, West Lafayette, as an Associate Professor. Currently, he is a Professor in the School of Chemical Engineering at Georgia Institute of Technology, Atlanta. He has held visiting appointments at E.I. Du Pont de Nemours, Wilmington, in 1994 and at Seoul National University, Seoul, Korea, in 1997. He was a recipient of the National Science Foundation's Young Investigator Award in 1993. His research interests are in the areas of system identification, robust control, model predictive control and nonlinear estimation.



James B. Rawlings was born in Gary, Indiana, USA in 1957. He received the B.S. from the University of Texas in 1979 and the Ph.D. from the University of Wisconsin in 1985, both in Chemical Engineering. He spent one year at the University of Stuttgart as a NATO post-doctoral fellow and then joined the faculty at the University of Texas. He moved to the University of Wisconsin in 1995 and is currently the Chair and Paul A. Elfers Professor of the Department of Chemical Engineering and the co-director of the Texas-Wisconsin Modeling and Control Consortium (TWMCC). His research interests are in the areas of chemical process modeling and control, nonlinear model predictive control, moving horizon state estimation and monitoring, particle technology and crystal engineering.