Model Predictive Control with Unscented Kalman Filter for Nonlinear Implicit Systems

Takashi Shimizu, Tomoaki Hashimoto

Abstract—A class of implicit systems is known as a more generalized class of systems than a class of explicit systems. To establish a control method for such a generalized class of systems, we adopt model predictive control method which is a kind of optimal feedback control with a performance index that has a moving initial time and terminal time. However, model predictive control method is inapplicable to systems whose all state variables are not exactly known. In other words, model predictive control method is inapplicable to systems with limited measurable states. In fact, it is usual that the state variables of systems are measured through outputs, hence, only limited parts of them can be used directly. It is also usual that output signals are disturbed by process and sensor noises. Hence, it is important to establish a state estimation method for nonlinear implicit systems with taking the process noise and sensor noise into consideration. To this purpose, we apply the model predictive control method and unscented Kalman filter for solving the optimization and estimation problems of nonlinear implicit systems, respectively. The objective of this study is to establish a model predictive control with unscented Kalman filter for nonlinear implicit systems.

Keywords—Model predictive control, unscented Kalman filter, nonlinear systems, implicit systems.

I. INTRODUCTION

Model predictive control is a feedback control approach that optimizes control performance over a finite horizon, and its performance index has moving initial and terminal time. So far, several MPC methods have been proposed for fluid systems [1]-[4], spatiotemporal dynamic systems [5]-[9], Schrödinger systems [10], [11], stochastic systems [12]-[14], and probabilistic constrained systems [15]-[17]. Although the aforementioned studies have achieved tremendous progress in controlling various kinds of systems, all systems addressed in the above studies belong to a class of so-called explicit systems. On the other hand, the control problem of so-called implicit systems that belong to a more generalized class of systems than a class of explicit systems is even more challenging to study. In [18], model predictive control method has been proposed for a class of discrete-time nonlinear implicit systems. In this study, we examine the problem of model predictive control for such a class of implicit systems.

This work was supported in part by Japan Society for the Promotion of Science, Grant-in-Aid for Scientific Research (C), Grant Number JP18K04216. T. Hashimoto is with the Department of Mechanical Engineering, Osaka Institute of Technology, Asahikou 535-8585, Japan (e-mail: info@thashi.net).

Based on the model predictive control method, the present control input is determined by solving a finite-horizon open-loop optimal control problem using the present state of the system as the initial state, and this procedure is repeated at each sampling time. Hence, the current state of the system is assumed to be exactly known in this approach. In fact, it is assumed in the aforementioned result [18] that all state variables are accessible for designing a controller. However, it is usual that the state variables of systems are measured through the outputs and hence only limited parts of them can be used directly. The control method proposed in [18] is inapplicable to systems with limited measurable states.

The design problem of output feedback model predictive control is still unsolved for nonlinear implicit systems with limited measurable states. Therefore, the objective of this study is to provide a generalized framework for designing an output feedback model predictive control for nonlinear implicit systems with limited measurable states. For this purpose, we establish a state estimation method based on an unscented Kalman filter [19] with taking the process noise and sensor noise into consideration.

This paper is organized as follows. In Section II, we define the system model and notations. In Section III, we consider the problem of model predictive control for a class of discrete-time nonlinear implicit systems. Using the variational principle, we derive the stationary conditions that must be satisfied for a performance index to be optimized. In Section IV, we provide a state estimation method based on an unscented Kalman filter for a class of discrete-time nonlinear implicit systems. Finally, some concluding remarks are given in Section V.

II. NOTATION AND SYSTEM MODEL

Let \( \mathbb{R} \) denote a set of real numbers. Let \( \mathbb{R}_+ \) and \( \mathbb{Z}_+ \) denote the sets of nonnegative real numbers and integers, respectively. Let \( \mathbb{N} \) denote the sets of natural numbers (positive integers). For a matrix \( A \in \mathbb{R}^{n \times n} \), the transpose and the inverse of \( A \) are denoted by \( A^T \) and \( A^{-1} \), respectively. The determinant and rank of a matrix \( A \) are denoted by \( \det(A) \) and \( \text{rank}(A) \), respectively. Let \( I \) denote the identity matrix.

For a scalar function \( \phi(x) : \mathbb{R}^n \to \mathbb{R} \), the differentiation of \( \phi(x) \) with respect to \( x \in \mathbb{R}^n \) is defined by

\[
\frac{\partial \phi(x)}{\partial x} := \begin{bmatrix}
\frac{\partial \phi(x)}{\partial x_1} \\
\frac{\partial \phi(x)}{\partial x_2} \\
\vdots \\
\frac{\partial \phi(x)}{\partial x_n}
\end{bmatrix}
\]

The Jacobian matrix of a vector-valued function \( F(x) : \mathbb{R}^n \to \mathbb{R}^m \) is
\( \mathbb{R}^n \rightarrow \mathbb{R}^n \) is defined by
\[
\frac{\partial F(x)}{\partial x} := \begin{bmatrix}
\frac{\partial F_1(x)}{\partial x_1} & \cdots & \frac{\partial F_1(x)}{\partial x_n} \\
\vdots & \ddots & \vdots \\
\frac{\partial F_n(x)}{\partial x_1} & \cdots & \frac{\partial F_n(x)}{\partial x_n}
\end{bmatrix}
\]

In this study, we consider the following discrete-time nonlinear implicit systems:
\[
E(x(t))x(t+1) = F(x(t), u(t)), \tag{1}
\]
where \( t \in \mathbb{Z}_+ \), \( x(t) : \mathbb{Z}_+ \rightarrow \mathbb{R}^n \), and \( u(t) : \mathbb{Z}_+ \rightarrow \mathbb{R}^m \) denote a temporal variable, the state, and the control input, respectively. \( E(x(t)) : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n} \) and \( F(x(t), u(t)) : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n \) are continuously differentiable functions. For examples, the discretized equations for nonlinear diffusion process [20] and RLC network circuits [21] belong to a class of systems described by (1).

Here, we assume that \( E(x(t)) \) is not necessarily of full-rank, i.e., rank \( E(x(t)) \leq n \). In particular, we call system (1) the descriptor system when \( \det E(x(t)) = 0 \). Without loss of generality, we assume \( F(0, 0) = 0 \), that is, the origin \( x = 0 \) is the equilibrium point.

### III. Model Predictive Control

In this section, we consider the model predictive control problem of system (1). Using the variational principle, we analytically derive the stationary conditions that must be satisfied for a performance index to be optimized. The control input at each time \( t \) is determined so as to minimize the performance index given by
\[
J = \phi(x(t+N)) + \sum_{k=t}^{t+N-1} L(x(k), u(k)). \tag{2}
\]

Therein, \( N \in \mathbb{N} \) denotes the length of prediction horizon. \( \phi : \mathbb{R}^n \rightarrow \mathbb{R}_+ \) and \( L : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}_+ \) are so-called terminal cost function and stage cost function, respectively, and assumed to be continuously differentiable functions with \( \phi(0) = 0\) and \( L(0, 0) = 0 \).

The minimization problem of (2) subject to (1) can be reduced to the minimization of the following performance index introduced using the costate \( \lambda \in \mathbb{R}^n \) associated with system equation (1):
\[
\bar{J} = \phi(x(t+N)) + \sum_{k=t}^{t+N-1} \left[ L(x(k), u(k)) + \lambda^T(k+1) \{ F(x(k), u(k)) - E(x(k))x(k+1) \} \right]. \tag{3}
\]

Let \( H \in \mathbb{R} \) denote the Hamiltonian defined by
\[
H := L(x(k), u(k)) + \lambda^T(k+1)F(x(k), u(k)). \tag{4}
\]

Let \( \delta \bar{J}, \delta x, \delta \lambda, \) and \( \delta u \) denote the variations (infinitesimal changes) in \( \bar{J}, x, \lambda, \) and \( u \), respectively. Since the optimal solution must satisfy the stationary condition \( \delta \bar{J} = 0 \), we need to consider the variation \( \delta \bar{J} \) due to the variations \( \delta x, \delta \lambda, \) and \( \delta u \). Then, we need to calculate \( \delta \bar{J} = J(x + \delta x, \lambda + \delta \lambda, u + \delta u) - J(x, \lambda, u) \). Applying the Taylor expansion into \( J(x + \delta x, \lambda + \delta \lambda, u + \delta u) \) around \( (x, \lambda, u) \) and neglecting the high order terms of each variation, we can compute the variation in \( \bar{J} \). Thus, \( \delta \bar{J} \) can be described by
\[
\delta \bar{J} = \sum_{k=t}^{t+N} \frac{\partial \bar{J}}{\partial x(k)} \delta x(k) + \sum_{k=t}^{t+N-1} \frac{\partial \bar{J}}{\partial u(k)} \delta u(k)
\]
\[
+ \sum_{k=t}^{t+N} \frac{\partial \bar{J}}{\partial \lambda(k)} \delta \lambda(k). \tag{5}
\]

It follows from \( \delta \bar{J} \rightarrow 0 \) that we can obtain stationary conditions that must be satisfied for a performance index to be optimized. The detailed computation on \( \delta \bar{J} \rightarrow 0 \) can be found in [18]. On the basis of the variational principle, we obtain the necessary conditions for a stationary value of \( \bar{J} \) over the horizon \( t \leq k \leq t + N \) as follows.
\[
E(x(k))x(k+1) = F(x(k), u(k)) \tag{6a}
\]
\[
\lambda^T(t+N)E(x(t+N)) = \frac{\partial \phi(x(t+N))}{\partial x(t+N)} \tag{6b}
\]
\[
\lambda^T(k)E(x(k+1)) = \frac{\partial H(x(k), \lambda(k+1), u(k))}{\partial x(k)}
- \lambda^T(k+1) \frac{\partial \{ E(x(k))x(k+1) \}}{\partial x(k)} \tag{6c}
\]
\[
\frac{\partial H(x(k), \lambda(k+1), u(k))}{\partial u(k)} = 0. \tag{6d}
\]

Note that if \( E(x(k)) = 1 \), then the obtained stationary conditions (6) can be reduced to the well-known Euler–Lagrange equations. Hence, we see that the stationary conditions (6) are natural extensions to the Euler–Lagrange equations. Accordingly, we call the stationary conditions (6) the generalized Euler–Lagrange equations.

A well-known difficulty in solving nonlinear optimal control problems is that the obtained stationary conditions cannot be solved analytically in general. Thus, the computational algorithm for numerically solving the generalized Euler–Lagrange equations (6) has been proposed in [18]. The remainder of this section is devoted to a brief description of the algorithm.

Let \( U(t) \in \mathbb{R}^N \) be defined by
\[
U(t) := [u^T(t), u^T(t+1), \ldots, u^T(t+N-1)]^T.
\]

For a given initial optimal solution \( U(t) \) and the present state \( x(t) \), we first determine the state over the prediction horizon by using (6a), that is, \( x(k) \) for \( k = t, t+1, \ldots, t+N \) is calculated recursively from \( k = t \) to \( k = t + N \) by (6a). Next, the terminal costate \( \lambda(t+N) \) is determined from the obtained terminal state \( x(t+N) \) by (6b). Consequently, the costate over the prediction horizon is also determined by using (6c), that is, \( \lambda(k) \) for \( k = t+1, \ldots, t+N \) is calculated recursively from \( k = t+N \) to \( k = t+1 \) by (6c). Fig. 1 shows that the procedure for solving the equation of \( x \) is forward, whereas the one for solving the equation of \( \lambda \) is backward.

Because \( x(k) \) and \( \lambda(k) \) for \( k = t, t+1, \ldots, t+N \) are determined by \( U(t) \) and \( x(t) \) through (6a)-(6c), the remaining
If been developed such that \( \parallel \) the stationary conditions are satisfied. Several algorithms have on arguments of \( \Delta \) we obtain (6a)-(6c) for given \( U \). The Jacobian differential equation for \( U \) is a positive constant introduced to \( x \) is nonsingular, we obtain the following (7):

\[
D(U(t), x(t), t) := \begin{bmatrix}
\frac{\partial H(x(t), \lambda(t+1), w(t))}{\partial x(t+1)} \\
\vdots \\
\frac{\partial H(x(t+N-1), \lambda(t+N), w(t+N-1))}{\partial x(t+N-1)}
\end{bmatrix}
\]  

(7)

Because \( x(k) \) and \( \lambda(k) \) are uniquely determined through (6a)-(6c) for given \( U(t) \) and \( x(t) \), \( x(k) \) and \( \lambda(k) \) depend on \( U(t) \) and \( x(t) \). Hence, it is reasonable to consider the arguments of \( D \) as \( U(t), x(t), t \).

For given \( U(t) \) and \( x(t) \), \( D \) is not necessarily equal to zero, so \( \|D\| \) is used to evaluate the optimality performance. If \( \|D\|=0 \) is satisfied for the given \( U(t) \) and \( x(t) \), then the stationary conditions are satisfied. Several algorithms have been developed such that \( \|D\| \) can be decreased by suitably updating \( U(t) \), as discussed below.

Instead of solving \( D(U(t), x(t), t) = 0 \) itself at each time by an iterative method such as the steepest descent method or Newton’s method, we find the derivative of \( U(t) \) with respect to time so that \( D(U(t), x(t), t) = 0 \) is satisfied identically. Namely we determine \( U(t) \) such that

\[
\dot{D}(U(t), x(t), t) = -\xi D(U(t), x(t), t)
\]  

(8)

is satisfied, where \( \xi \) is a positive constant introduced to stabilize \( D = 0 \). If we choose \( \xi = 1/\Delta t \), then the stability of (8) with forward difference approximation is guaranteed, where \( \Delta t \) denotes the sampling period. By total differentiation, we obtain

\[
\frac{\partial D}{\partial x(t)} \dot{x}(t) = -\xi D - \frac{\partial D}{\partial x(t)} \dot{x}(t) - \frac{\partial D}{\partial t}.
\]  

(9)

This equation can be regarded as a linear algebraic equation with the coefficient matrix \( \frac{\partial D}{\partial U(t)} \), which can be used to determine \( \dot{U}(t) \) for the given \( U(t), x(t), \dot{x}(t) \), and \( t \). Then, if the Jacobian \( \frac{\partial D}{\partial U} \) is nonsingular, we obtain the following differential equation for \( U(t) \):

\[
\dot{U}(t) = \left( \frac{\partial D}{\partial U(t)} \right)^{-1} \left( -\xi D - \frac{\partial D}{\partial x(t)} \dot{x}(t) - \frac{\partial D}{\partial t} \right).
\]  

(10)

We can update the solution \( U(t) \) of \( D(U(t), x(t), t) = 0 \) without using an iterative optimization method by integrating (10) in real time as, for example, \( U(t+\Delta t) = U(t) + \dot{U}(t) \Delta t \). This approach is a type of continuation method [22] in the sense that the solution curve \( U(t) \) is traced by integrating a differential equation. From the computational viewpoint, the differential equation (10) still involves expensive operations, i.e., solving the Jacobians \( \frac{\partial D}{\partial U(t)} \), \( \frac{\partial D}{\partial x(t)} \), and \( \frac{\partial D}{\partial t} \) and linear algebraic equation associated with \( \frac{\partial D}{\partial U} \)\(^{-1} \). To reduce the computational cost of the Jacobians and linear equation, we employ two techniques: the forward difference approximation for the products of the Jacobians and vectors and the GMRES method [23] for the linear algebraic equation. Using the forward difference approximation, we can obtain a linear equation with respect to \( \dot{U} \). Thereafter, we can apply the GMRES algorithm to find the solution \( \dot{U}(t) \) of the linear equation. Consequently, \( U \) can be updated so that \( D = 0 \) is stabilized.

Here, it is important to note that the above numerical algorithm is inapplicable to systems whose current state \( x(t) \) is not exactly known. In other words, model predictive control method is inapplicable to systems with limited measurable states.

In fact, it is usual that the state variables of systems are measured through outputs, hence, only limited parts of them can be used directly. It is also usual that output signals are disturbed by process and sensor noises. Hence, it is important to establish a state estimation method for nonlinear implicit systems with taking the process noise and sensor noise into consideration.

IV. ESTIMATION BASED ON UNSCENTED KALMAN FILTER

In this section, we design an observer using unscented Kalman filter [19] for estimating the state \( x(t) \) of system (1). For this purpose, we introduce the following observer or system with taking the process noise and observation noise into consideration.

\[
E(\hat{x}(t)) \bar{x}(t+1) = F(\hat{x}(t), u(t)) + v(t),
\]  

(11a)

\[
y(t) = G(x(t), u(t)) + w(t),
\]  

(11b)

where \( \hat{x}(t) \) denotes the estimation of \( x(t) \) and \( y(t) \) is the output. \( E \) and \( F \) are given in (1). \( v(t) \) and \( w(t) \) denote the process noise and the observation noise, respectively, which can be caused by disturbances.

The optimal estimate in the minimum mean-squared error sense is given by the conditional mean. Let \( \hat{x}(i) \) be the mean of \( \hat{x}(i) \) conditioned on all of the observations up to and including time \( j \), i.e., \( \hat{x}(i) = E[\hat{x}(i)|Y^j] \), where \( Y^j := \{y(1), y(2), \ldots, y(j)\} \).

It is assumed that the means of \( v(t) \) and \( w(t) \) are zero for all time \( t \). Let \( Q^v(t) \) and \( Q^w(t) \) be the covariances of \( v(t) \) and \( w(t) \), respectively.

The UKF [19] first predicts the mean and covariance of a future state using the process model and weighted sigma
points as follows:
\[
E(\chi^i(t))\chi^i(t+1|t) = F(\chi^i(t), u(t)),
\]
(12)
\[
\hat{x}(t+1|t) = \sum_{i=0}^{2n} W^i \chi^i(t+1|t),
\]
(13)
\[
Q^i(t+1) = Q^i(t+1) + \sum_{i=0}^{2n} W^i \left( \chi^i(t+1|t) - \hat{x}(t+1|t) \right) \left( \chi^i(t+1|t) - \hat{x}(t+1|t) \right)^T,
\]
(14)

where \( W^i \) and \( \chi^i \) denote the weight and sigma point, respectively. The definitions of \( W^i \) and \( \chi^i \) can be found in [19]. \( \chi^i(t+1|t) \) can be determined from (12). Then, \( \hat{x}(t+1|t) \) and \( Q^i(t+1|t) \) are determined form (13) and (14), respectively.

After we redraw a new set of sigma points \( \chi^i \) to incorporate the effect of the additive process noise, the predicted observation is calculated by
\[
\tilde{y}(t+1|t) = \sum_{i=0}^{2n} W^i G(\chi^i(t+1|t), u(t+1)).
\]
(15)

Moreover, the cross covariance \( P \) and innovation covariance \( R \) are determined by
\[
P(t+1|t) = \sum_{i=0}^{2n} W^i \left( \chi^i(t+1|t) - \hat{x}(t+1|t) \right) \left( \chi^i(t+1|t) - \hat{x}(t+1|t) \right)^T,
\]
(16)
\[
R(t+1|t) = \sum_{i=0}^{2n} W^i \left( G(\chi^i(t+1|t), u(t+1)) - \tilde{y}(t+1|t) \right) \left( G(\chi^i(t+1|t), u(t+1)) - \tilde{y}(t+1|t) \right)^T + Q^i(t+1),
\]
(17)

Consequently, the estimate at time \( t+1 \) is obtained by updating the prediction by the linear update rule:
\[
K(t+1) = P(t+1|t)R^{-1}(t+1|t),
\]
(18a)
\[
\hat{x}(t+1|t+1) = \hat{x}(t+1|t) + K(t+1) \left( y(t+1) - \tilde{y}(t+1|t) \right),
\]
(18b)
\[
Q^i(t+1|t+1) = Q^i(t+1|t) - K(t+1)R(t+1|t)K^T(t+1),
\]
(18c)

The model predictive control method proposed in [18] for implicit nonlinear systems is inapplicable to systems with limited measurable states. It is usual that the state variables of systems are measured through outputs, hence, only limited parts of them can be used directly. It is also usual that output signals are disturbed by process and sensor noises. Motivated by the above fact, this paper proposes a state estimation method for nonlinear implicit systems with taking the process noise and sensor noise into consideration. Consequently, we establish a model predictive control method with a state estimation method based on unscented Kalman filter for nonlinear implicit systems. It is known that time delays may cause instabilities of the closed-loop system and lead to more complex analysis [25]-[30]. The control problem of implicit nonlinear systems with time delays is a possible future work.

V. CONCLUSION

The model predictive control method proposed in [18] for implicit nonlinear systems is inapplicable to systems with limited measurable states. It is usual that the state variables of systems are measured through outputs, hence, only limited parts of them can be used directly. It is also usual that output signals are disturbed by process and sensor noises. Motivated by the above fact, this paper proposes a state estimation method for nonlinear implicit systems with taking the process noise and sensor noise into consideration. Consequently, we establish a model predictive control method with a state estimation method based on unscented Kalman filter for nonlinear implicit systems. It is known that time delays may cause instabilities of the closed-loop system and lead to more complex analysis [25]-[30]. The control problem of implicit nonlinear systems with time delays is a possible future work.

REFERENCES


