OUTPUT REGULATION FOR DISCRETE-TIME NONLINEAR STOCHASTIC OPTIMAL CONTROL PROBLEMS WITH MODEL-REALITY DIFFERENCES

SIE LONG KEK
Department of Mathematics and Statistics
Universiti Tun Hussein Onn Malaysia
86400 Parit Raja, Malaysia

MOHD ISMAIL ABD AZIZ
Department of Mathematical Sciences
Universiti Teknologi Malaysia
81310 UTM, Skudai, Malaysia

Abstract. In this paper, we propose an output regulation approach, which is based on principle of model-reality differences, to obtain the optimal output measurement of a discrete-time nonlinear stochastic optimal control problem. In our approach, a model-based optimal control problem with adding the adjustable parameters is considered. We aim to regulate the optimal output trajectory of the model used as closely as possible to the output measurement of the original optimal control problem. In doing so, an expanded optimal control problem is introduced, where system optimization and parameter estimation are integrated. During the computation procedure, the differences between the real plant and the model used are measured repeatedly. In such a way, the optimal solution of the model is updated. At the end of iteration, the converged solution approaches closely to the true optimal solution of the original optimal control problem in spite of model-reality differences. It is important to notice that the resulting algorithm could give the output residual that is superior to those obtained from Kalman filtering theory. The accuracy of the output regulation is therefore highly recommended. For illustration, a continuous stirred-tank reactor problem is studied. The results obtained show the efficiency of the approach proposed.

1. Introduction. Recently, an integrated optimal control algorithm for solving discrete-time nonlinear stochastic optimal control problems has been proposed, see for examples [3], [9], [10] and [4]. The developed algorithm is an iterative approach, where the model-based optimal control problem is solved repeatedly in order to approximate the true optimal solution of the original optimal control problem. With the adjustable parameters that are introduced in the model, the differences between the real plant and the model used could be measured. The repetitive solution is then converged to the real optimal solution within a given tolerance in spite of model-reality differences [11], [12], [1]. On the other hand, because of the present of
the random disturbances, an optimal filtering solution of the nonlinear stochastic optimal control problem in discrete-time is obtained, where the modified linear quadratic Gaussian optimal control problem is solved repeatedly [5]. In addition to this, a least-square output residual is introduced in the cost functional such that the output error is further minimized [6]. However, minimizing the output error would not give a minimum value of the cost function due to the weighted parameter that is selected for the least-square output residual in the model. In this paper, we propose an efficient computation approach to improve this limitation. In our approach, the linear quadratic regulator optimal control model is considered, where the trajectories of state and control are smoothed in expectation manner. Moreover, an adjustable parameter is introduced to the model output, which is measured from the expected state trajectory. The aim of this adjustable parameter is to regulate the expected output as closely as possible to the real output, as such giving the smallest minimum output error. Note that the Kalman filtering theory is not applied here. It is remarked that the proposed approach gives both of the optimal expected solution and the optimal regulated output at the end of iteration computation procedure despite model-reality differences. Hence, the accuracy of output solution is highly recommended.

The rest of the paper is organized as follows. In Section 2, a general class of discrete-time nonlinear stochastic optimal control problem is described. In Section 3, the model-based optimal control problem with the adjustable parameters is discussed. The expectation optimal solution is obtained and then the expected output is regulated approximately to the real output in spite of model-reality differences. In Section 4, an illustrative example of continuous stirred-tank reactor problem is presented to show the efficiency of the proposed approach. Finally, some concluding remarks are made.

2. Problem Description. Consider a general class of stochastic optimal control problem given below:

$$\min_{u(k)} J_0(u) = E[\varphi(x(N), N) + \sum_{k=0}^{N-1} L(x(k), u(k), k)]$$

subject to

$$x(k+1) = f(x(k), u(k), k) + G\omega(k)$$
$$y(k) = h(x(k), k) + \eta(k)$$

where $u(k) \in \mathbb{R}^m$, $k = 0, 1, \ldots, N-1$, $x(k) \in \mathbb{R}^n$, $k = 0, 1, \ldots, N$, and $y(k) \in \mathbb{R}^p$, $k = 0, 1, \ldots, N$, are, respectively, the control sequence, the state sequence, and the measured output. The terms $\omega(k) \in \mathbb{R}^q$, $k = 0, 1, \ldots, N-1$, and $\eta(k) \in \mathbb{R}^p$, $k = 0, 1, \ldots, N$, are the stationary Gaussian white noise sequences with zero mean and their covariance matrices are given by $Q_\omega$ and $R_\eta$, respectively, where $Q_\omega$ is a $q \times q$ positive definite matrix and $R_\eta$ is a $p \times p$ positive definite matrix. $G$ is an $n \times q$ process noise coefficient matrix, $f : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \rightarrow \mathbb{R}^n$ represents the real plant, and $h : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^p$ is the output measurement, whereas $\varphi : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$ is the terminal cost, $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \rightarrow \mathbb{R}$ is the cost under summation. Here, $J_0$ is the scalar cost function and $E[\cdot]$ is the expectation operator. It is assumed that all functions in (1) are continuously differentiable with respect to their respective arguments.
The initial state is

\[ x(0) = x_0 \]

where \( x_0 \in \mathbb{R}^n \) is a random vector with mean and covariance given, respectively, by

\[ E[x_0] = \bar{x}_0 \text{ and } E[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^\top] = M_0. \]

Here, \( M_0 \) is an \( n \times n \) positive definite matrix. It is assumed that the initial state, process noise and measurement noise are statistically independent.

This optimal control problem is regarded as the real optimal control problem, and is referred to as Problem (P). We notice that the exact solution of Problem (P) is, in general, unable to be obtained. Furthermore, applying the nonlinear filtering theory to estimate the state of the real plant is computationally demanding. In view of these, we propose to solve Problem (P) via solving a simplified model-based optimal control problem iteratively. Let this simplified model-based optimal control problem, which is referred to as Problem (M), be given below:

\[
\min_{u(k)} J_1(u) = \frac{1}{2} \bar{x}(N)^\top S(N) \bar{x}(N) + \gamma(N) \\
+ \sum_{k=0}^{N-1} \frac{1}{2} (\bar{x}(k)^\top Q \bar{x}(k) + u(k)^\top R u(k)) + \gamma(k]
\]

subject to

\[
\begin{align*}
\bar{x}(k+1) &= A \bar{x}(k) + B u(k) + \alpha_1(k), \quad \bar{x}(0) = \bar{x}_0 \\
\bar{y}(k) &= C \bar{x}(k) + \alpha_2(k) \\
\bar{\gamma}(k) &= \bar{y}(k) + \alpha_3(k)
\end{align*}
\]

where \( \bar{x}(k) \in \mathbb{R}^n, k = 0, 1, \ldots, N, \bar{y}(k) \in \mathbb{R}^p, k = 0, 1, \ldots, N, \) and \( \bar{\gamma}(k) \in \mathbb{R}^p, k = 0, 1, \ldots, N, \) are, respectively, the expected state sequence, the expected output sequence and the regulated output sequence. \( A \) is an \( n \times n \) state transition matrix, \( B \) is an \( n \times m \) control coefficient matrix, \( C \) is a \( p \times n \) output coefficient matrix, while \( S(N) \) and \( Q \) are \( n \times n \) positive semi-definite matrices and \( R \) is a \( m \times m \) positive definite matrix. Here, \( \gamma(k) \in \mathbb{R}, k = 0, 1, \ldots, N, \alpha_1(k) \in \mathbb{R}^n, k = 0, 1, \ldots, N - 1, \alpha_2(k) \in \mathbb{R}^p, k = 0, 1, \ldots, N, \) and \( \alpha_3(k) \in \mathbb{R}^p, k = 0, 1, \ldots, N, \) are introduced as adjustable parameters.

Notice that solving Problem (M) iteratively would give the optimal expected output solution of Problem (P), which is given by \( \bar{y}(k) \), and the optimal regulated output solution of Problem (P), which is represented by \( \bar{\gamma}(k) \). Since the Kalman filtering theory is not used here, the state error covariance is larger than the state error covariance as presented in [5]. However, the additional output measurement, which is added into the model, regulates the expected output sequence as closely as possible to the real output trajectory. In this regulation procedure, we aim to approximate the true output trajectory of Problem (P).

3. Output Regulation with Model-Reality Differences. Now, let us introduce an expanded optimal control problem, which is referred to as Problem (E),
given below:

\[
\begin{align*}
\min_{u(k)} J_2(u) &= \frac{1}{2} \bar{x}(N)^\top S(N)\bar{x}(N) + \gamma(N) \\
&\quad + \sum_{k=0}^{N-1} \frac{1}{2} (\bar{x}(k)^\top Q\bar{x}(k) + u(k)^\top Ru(k)) + \gamma(k) \\
&\quad + \frac{1}{2} r_1 \| u(k) - v(k) \|^2 + \frac{1}{2} r_2 \| \bar{x}(k) - z(k) \|^2 \\
&\quad + \frac{1}{2} r_3 \| \hat{y}(k) - \hat{y}(k) \|^2
\end{align*}
\]

subject to

\[
\begin{align*}
\bar{x}(k+1) &= A\bar{x}(k) + Bu(k) + \alpha_1(k), \quad \bar{x}(0) = \bar{x}_0 \\
\hat{y}(k) &= C\bar{x}(k) + \alpha_2(k) \\
\hat{y}(k) &= \hat{y}(k) + \alpha_3(k) \\
\frac{1}{2} z(N)^\top S(N)z(N) + \gamma(N) &= \varphi(z(N), N) \\
\frac{1}{2} (z(k)^\top Qz(k) + v(k)^\top Rv(k)) + \gamma(k) &= L(z(k), v(k), k) \\
Az(k) + Bv(k) + \alpha_1(k) &= f(z(k), v(k), k) \\
Cz(k) + \alpha_2(k) &= h(z(k), k) \\
\hat{y}(k) + \alpha_3(k) &= \hat{y}(k) \\
v(k) &= u(k) \\
z(k) &= \bar{x}(k) \\
\hat{y}(k) &= \hat{y}(k)
\end{align*}
\]

where \( v(k) \in \mathbb{R}^m \), \( k = 0, 1, \ldots, N-1 \), \( z(k) \in \mathbb{R}^n \), \( k = 0, 1, \ldots, N \), and \( \hat{y}(k) \in \mathbb{R}^p \), \( k = 0, 1, \ldots, N \), are introduced to separate the sequences of control, expected state and estimated output in the optimization problem from the respective signals in the parameter estimation problem, and \( \| \cdot \| \) denotes the usual Euclidean norm. The terms \( \frac{1}{2} r_1 \| u(k) - v(k) \|^2 \), \( \frac{1}{2} r_2 \| \bar{x}(k) - z(k) \|^2 \) and \( \frac{1}{2} r_3 \| \hat{y}(k) - \hat{y}(k) \|^2 \) with \( r_1 \in \mathbb{R} \), \( r_2 \in \mathbb{R} \) and \( r_3 \in \mathbb{R} \) are introduced to improve convexity and to facilitate convergence of the resulting iterative algorithm. It is important to note that the algorithm is designed such that the constraints \( v(k) = u(k) \), \( z(k) = \bar{x}(k) \) and \( \hat{y}(k) = \hat{y}(k) \) are satisfied upon termination of the iterations, assuming that convergence is achieved. The state constraint \( z(k) \), the output constraint \( \hat{y}(k) \) and the control constraint \( v(k) \) are used for the computation of the parameter estimation and matching schemes, while the corresponding state estimate constraint \( \bar{x}(k) \), output estimate constraint \( \hat{y}(k) \) and control constraint \( u(k) \) are used in the optimization of the model-based optimal control problem. On this basis, the system optimization and the parameter estimation are mutually interactive.

Note that Problem (E) is equivalent to the estimation of Problem (P), which is in terms of expectation and regulation.

### 3.1. Necessary optimality conditions.

To solve Problem (E), let us define the Hamiltonian function as follows:
where $\lambda(k) \in \mathbb{R}^n, k = 0, 1, \ldots, N-1$, and $\beta(k) \in \mathbb{R}^n, k = 0, 1, \ldots, N-1$, are modifiers. Then, the augmented cost function becomes

$$J_2(k) = \frac{1}{2}(\dot{x}(N)^T S(N) \dot{x}(N) + \gamma(N) + p(0)^T \dot{x}(0) - p(N)^T \dot{x}(N))$$

$$+ \xi(N)(\varphi(z(N), N) - \frac{1}{2} \dot{x}(N)^T S(N) \dot{x}(N) - \gamma(N))$$

$$+ \Gamma(\dot{x}(N) - z(N)) + \theta_1(N)^T (\dot{y}(N) - \dot{y}(N))$$

$$+ \sum_{k=0}^{N-1} [H_e(k) - p(k)^T \dot{x}(k) + \lambda(k)^T v(k) + \beta(k)^T z(k)$$

$$+ \xi(k)(Lz(k), v(k), k) - \frac{1}{2}(z(k)^T Qz(k) + v(k)^T Rv(k)) - \gamma(k))$$

$$+ \mu(k)^T (f(z(k), v(k), k) - Az(k) - Bv(k) - \alpha_1(k))$$

$$+ \theta_1(k)^T \dot{y}(k) + \theta_2(k)^T (C \ddot{x}(k) + \alpha_2(k) - \ddot{y}(k))$$

$$+ \theta_3(k)^T (\dot{y}(k) + \alpha_3(k) - \ddot{y}(k))$$

$$+ \pi_1(k)^T (h(z(k), k) - Cz(k) - \alpha_2(k))$$

$$+ \pi_2(k)^T (y(k) - \ddot{y}(k) - \alpha_3(k))]$$

(5)

where $p(k), \gamma(k), \xi(k), \mu(k), \Gamma, \lambda(k), \beta(k), \pi_i(k), i = 1, 2$, and $\theta_i(k), i = 1, 2, 3$, are the appropriate multipliers to be determined later.

Applying the calculus of variation [4], [5], [6], [2], [8], the following necessary optimality conditions are obtained:

(a) Stationary condition:

$$Ru(k) + B^T p(k + 1) - \lambda(k) + r_1(u(k) - v(k)) = 0$$

(6a)

(b) Co-state equation:

$$p(k) = Q\ddot{x}(k) + A^T p(k + 1) - \beta(k) + r_2(\dot{x}(k) - z(k))$$

(6b)

(c) State equation:

$$\dot{x}(k + 1) = A\ddot{x}(k) + Bu(k) + \alpha_1(k)$$

(6c)

(d) Boundary conditions:

$$p(N) = S(N)\ddot{x}(N) + \Gamma and \ddot{x}(0) = \ddot{x}_0$$

(e) Output equations:

$$\ddot{y}(k) = C\ddot{x}(k) + \alpha_2(k) and \ddot{y}(k) = \ddot{y}(k) + \alpha_3(k)$$

(7)
Adjustable parameter equations:

\[ \varphi(z(N), N) = \frac{1}{2} z(N)^\top S(N) z(N) + \gamma(N) \]  

\[ L(z(k), v(k), k) = \frac{1}{2} (z(k)^\top Q z(k) + v(k)^\top R v(k)) + \gamma(k) \]  

\[ f(z(k), v(k), k) = A z(k) + B v(k) + \alpha_1(k) \]  

\[ h(z(k), k) = C z(k) + \alpha_2(k) \]  

\[ y(k) = \hat{y}(k) + \alpha_3(k) \]  

Multiplier equations:

\[ \Gamma = \nabla z(k) \varphi - S(N) z(N) \]  

\[ \lambda(k) = -(\nabla v(k) L - R v(k)) - \left( \frac{\partial f}{\partial v(k)} - B \right)^\top \hat{p}(k + 1) \]  

\[ \beta(k) = -(\nabla z(k) L - Q z(k)) - \left( \frac{\partial f}{\partial z(k)} - A \right)^\top \hat{p}(k + 1) \]  

\[ \theta_1(k) = r_3(\bar{y}(k) - \bar{y}(k)) \]  

with \( \xi(k) = 1, \mu(k) = \hat{p}(k + 1), \pi_1(k) = \theta_2(k) = \theta_3(k) = 0, \) and \( \pi_2(k) = \theta_3(k) = 0. \)

Separable variables:

\[ z(k) = \tilde{x}(k), v(k) = u(k), \hat{p}(k) = p(k), \tilde{y}(k) = \bar{y}(k) \]  

In view of these optimality conditions, the multipliers are computed from (9), the parameter estimation problem is defined by (8), where the adjustable parameters are calculated, and the modified model-based optimal control problem, which satisfies the optimality conditions in (6) and (7), is given below. This modified model-based optimal control problem is referred to as Problem (MM).

\[ \min_{u(k)} J_3(u) = \frac{1}{2} \tilde{x}(N)^\top S(N) \tilde{x}(N) + \Gamma^\top \tilde{x}(N) + \gamma(N) + \theta_1(N)^\top \tilde{y}(N) \]  

\[ + \sum_{k=0}^{N-1} \frac{1}{2} (\tilde{x}(k)^\top Q \tilde{x}(k) + u(k)^\top R u(k)) + \gamma(k) \]  

\[ + \frac{1}{2} r_1 \| u(k) - v(k) \|^2 + \frac{1}{2} r_2 \| \tilde{x}(k) - z(k) \|^2 \]  

\[ + \frac{1}{2} r_3 \| \bar{y}(k) - \bar{y}(k) \|^2 \]  

subject to

\[ \tilde{x}(k + 1) = A \tilde{x}(k) + B u(k) + \alpha_1(k), \quad \tilde{x}(0) = \tilde{x}_0 \]  

\[ \tilde{y}(k) = C \tilde{x}(k) + \alpha_2(k) \]  

\[ \tilde{y}(k) = \bar{y}(k) + \alpha_3(k) \]  

with the specified \( \alpha_1(k), \alpha_3(k), \alpha_3(k), \gamma(k), \Gamma, \lambda(k), \beta(k), \theta_1(k), v(k) \) and \( z(k) \), where the boundary conditions \( \tilde{x}(0) \) and \( p(N) \) are given with the specified modifier \( \Gamma \).

Note that it is essential to include the modification terms \( \lambda(k)^\top u(k), \beta(k)^\top \tilde{x}(k) \) and \( \theta_1(k)^\top \bar{y}(k) \) in the cost function of Problem (MM). Otherwise, the correct solution estimate of Problem (P) cannot be obtained by simply iterating the solution of
Problem (M) and performing parameter estimation at every iteration step. In addition, to obtain the solution of Problem (MM), it is necessary to solve the two-point boundary-value problem (TPBVP) that is defined by (6b) and (6c).

3.2. Feedback control law. The solution method for solving Problem (MM) is described in Theorem 3.1, where the feedback control law is resulted.

**Theorem 3.1.** Suppose the optimal control law for Problem (E) exists. Then, this control law is the feedback control law for Problem (MM) given by

\[ u(k) = -K(k)\bar{x}(k) + u_{ff}(k) \tag{12} \]

where

\[
\begin{align*}
    u_{ff}(k) &= (B^\top S(k+1)B + R_a)^{-1}(-B^\top s(k + 1) \\
    &\quad -B^\top S(k+1)\alpha_1(k) + \lambda_a(k)) \tag{13a} \\
    K(k) &= (B^\top S(k+1)B + R_a)^{-1}B^\top S(k+1)A \tag{13b} \\
    S(k) &= Q_a + A^\top S(k+1)(A - BK(k)) \tag{13c} \\
    s(k) &= (A - BK(k))^\top (s(k + 1) + S(k+1)\alpha_1(k) \\
    &\quad -\beta_a(k) + K(k)^\top \lambda_a(k) \tag{13d}
\end{align*}
\]

with the boundary conditions \( S(N) \) given and \( s(N) = 0 \), and

\[ R_a = R + r_1 I_m, Q_a = Q + r_2 I_n, \lambda_a(k) = \lambda(k) + r_1 v(k), \beta_a(k) = \beta(k) + r_2 z(k). \]

**Proof.** From the necessary optimality condition (6a), we obtain

\[ R_a u(k) = -B^\top p(k + 1) + \lambda_a(k) \tag{14} \]

Applying sweep method [2], [8],

\[ p(k) = S(k)\bar{x}(k) + s(k) \tag{15} \]

we substitute (15) for \( k = k + 1 \) into (14), which yields

\[ R_a u(k) = -B^\top S(k + 1)\bar{x}(k + 1) - B^\top s(k + 1) + \lambda_a(k) \tag{16} \]

Then, substitute the expected state equation (6c) into (16). After some algebraic manipulations, the feedback control law (12) is obtained, where (13a) and (13b) are satisfied.

From the co-state equation (6b), we substitute (15) for \( k = k + 1 \) to give

\[ p(k) = Q_a\bar{x}(k) + A^\top S(k+1)\bar{x}(k + 1) + A^\top s(k + 1) - \beta_a(k) \tag{17} \]

Consider the expected state equation (6c) in (17), we obtain

\[ p(k) = Q_a\bar{x}(k) + A^\top S(k+1)(A\bar{x}(k) + Bu(k) + \alpha_1(k)) + A^\top s(k + 1) - \beta_a(k) \tag{18} \]

Substitute the feedback control law (12) into (18), and doing some algebraic manipulations, it is found that (13c) and (13d) are satisfied after comparing to (15). This completes the proof.

Taking (12) into (6c), the expected state equation becomes

\[ \bar{x}(k + 1) = (A - BK(k))\bar{x}(k) + Bu_{ff}(k) + \alpha_1(k) \tag{19} \]
whereas the expected output is measured from
\[ \hat{y}(k) = C\tilde{x}(k) + \alpha_2(k) \quad (20a) \]
and the regulated output is obtained from
\[ \bar{y}(k) = \bar{g}(k) + \alpha_3(k) \quad (20b) \]

3.3. **Iterative computation procedure.** From the discussion above, the result is summarized as an iterative algorithm, where the computation procedure is given below.

**The iterative computation procedure**

Data \( A, B, C, G, Q, R, Q_s, S(N), M_0, \bar{x}_0, N, r_1, r_2, k_e, k_z, k_p, k_y, f, L, \varphi, y \) and \( h \). Note that \( A \) and \( B \) may be chosen based on the linearization of \( f \), and \( C \) is obtained from the linearization of \( h \).

Step 0 Compute a nominal solution. Assuming that \( \alpha_1(k) = 0, k = 0, 1, \ldots, N-1 \), \( \alpha_2(k) = 0, k = 0, 1, \ldots, N \), \( \alpha_3(k) = 0, k = 0, 1, \ldots, N \), and \( r_1 = r_2 = r_3 = 0 \), solve Problem (M) that is defined by (2) to obtain \( u(k)^0, k = 0, 1, \ldots, N-1 \), and \( \tilde{x}(k)^0, p(k)^0, \hat{y}(k)^0 \), \( k = 0, 1, \ldots, N \). Then, with \( \alpha_1(k) = 0, k = 0, 1, \ldots, N-1 \), \( \alpha_2(k) = 0, \alpha_3(k) = 0, k = 0, 1, \ldots, N \), and using \( r_1, r_2, r_3 \) from the data, compute \( K(k) \) and \( S(k) \), respectively, from (13b) and (13c).

Set \( i = 0, v(k)^0 = u(k)^0, z(k)^0 = \tilde{x}(k)^0, \hat{p}(k)^0 = p(k)^0 \) and \( \hat{y}(k)^0 = \bar{y}(k)^0 \).

Step 1 Compute the parameters \( \gamma(k)^i, k = 0, 1, \ldots, N, \alpha_1(k)^i, k = 0, 1, \ldots, N-1 \), and \( \alpha_2(k)^i, \alpha_3(k)^i, k = 0, 1, \ldots, N \), from (8). This is called the parameter estimation step.

Step 2 Compute the modifiers \( \Gamma^i, \lambda(k)^i, \beta(k)^i, k = 0, 1, \ldots, N-1 \), and \( \theta_1(k)^i, k = 0, 1, \ldots, N \), from (9). Note that this step requires taking the derivatives of \( f, h \) and \( L \) with respect to \( v(k)^i \) and \( z(k)^i \).

Step 3 Using \( \alpha_1(k)^i, \alpha_2(k)^i, \alpha_3(k)^i, \gamma(k)^i, \Gamma^i, \lambda(k)^i, \beta(k)^i, \theta_1(k)^i, v(k)^i \) and \( z(k)^i \), solve Problem (MM) that is defined by (11) using the result that is presented in Theorem 3.1. This is called the system optimization step.

3.1 Solve (13d) backward to obtain \( s(k)^i, k = 0, 1, \ldots, N \), and solve (13a), either backward or forward, to obtain \( u_{ff}(k)^i, k = 0, 1, \ldots, N-1 \).

3.2 Use (12) to obtain the new control \( u(k)^i, k = 0, 1, \ldots, N-1 \).

3.3 Use (19) to obtain the new state \( \tilde{x}(k)^i, k = 0, 1, \ldots, N \).

3.4 Use (15) to obtain the new costate \( p(k)^i, k = 0, 1, \ldots, N \).

3.5 Use (20) to obtain the new output \( \hat{y}(k)^i \) and \( \bar{y}(k)^i, k = 0, 1, \ldots, N \).

Step 4 Test the convergence and update the optimal expectation solution and the optimal output regulation of Problem (P). In order to provide a mechanism for regulating convergence, a simple relaxation method is employed:

\[ v(k)^{i+1} = v(k)^i + k_v(u(k)^i - v(k)^i) \quad (21a) \]
\[ z(k)^{i+1} = z(k)^i + k_z(\tilde{x}(k)^i - z(k)^i) \quad (21b) \]
\[ \hat{p}(k)^{i+1} = \hat{p}(k)^i + k_p(p(k)^i - \hat{p}(k)^i) \quad (21c) \]
\[ \hat{y}(k)^{i+1} = \hat{y}(k)^i + k_y(\hat{y}(k)^i - \hat{y}(k)^i) \quad (21d) \]

where \( k_v, k_z, k_p, k_y \in (0, 1] \) are scalar gains. If \( v(k)^{i+1} = v(k)^i, k = 0, 1, \ldots, N-1, z(k)^{i+1} = z(k)^i, k = 0, 1, \ldots, N, \) and \( \hat{y}(k)^{i+1} = \hat{y}(k)^i, k = 0, 1, \ldots, N \), within a given tolerance, stop; else set \( i = i + 1 \), and repeat the procedure starting with Step 1.
Remarks:

(a) The off-line computation is done, as stated in Step 0, to compute $K(k)$, $k = 0, 1, \ldots, N - 1$, and $S(k), k = 0, 1, \ldots, N$, for the control law design. Then, these parameters are used for solving Problem (M) in Step 0 and for solving Problem (MM) in Step 3, respectively.

(b) The parameters $\alpha_1(k)^i, \alpha_2(k)^i, \alpha_3(k)^i, \alpha_4(k)^i, \gamma(k)^i, \lambda(k)^i, \beta(k)^i, \theta_1(k)^i$ and $s(k)^i$ are zero in Step 0. Their calculated values, where $\alpha_1(k)^i, \alpha_2(k)^i, \alpha_3(k)^i$ and $\gamma(k)^i$ in Step 1, $\lambda(k)^i, \beta(k)^i$ and $\theta_1(k)^i$ in Step 2, and $s(k)^i$ in Step 3, change from iteration to iteration.

(c) The driving input $u_{ff}(k)$ in (13a) corrects the differences between the real plant and the model used, and it also derives the controller given in (12).

(d) Problem (P) is not necessary to be linear or to have a quadratic cost function.

(e) The conditions $v(k)^{i+1} = v(k)^i, z(k)^{i+1} = z(k)^i$, and $\hat{y}(k)^{i+1} = \hat{y}(k)^i$ are required to be satisfied for the converged optimal control sequence, the converged state estimate sequence and the converged output estimate sequence, respectively. The following averaged 2-norms are computed, and then they are compared with a given tolerance to verify the convergence of $v(k), z(k)$ and $\hat{y}(k)$:

\[
\|v^{i+1} - v^i\|_2 = \left(\frac{1}{N-1} \sum_{k=0}^{N-1} \|v(k)^{i+1} - v(k)^i\|\right)^{1/2}
\]

\[
\|z^{i+1} - z^i\|_2 = \left(\frac{1}{N} \sum_{k=0}^{N} \|z(k)^{i+1} - z(k)^i\|\right)^{1/2}
\]

\[
\|\hat{y}^{i+1} - \hat{y}^i\|_2 = \left(\frac{1}{N} \sum_{k=0}^{N} \|\hat{y}(k)^{i+1} - \hat{y}(k)^i\|\right)^{1/2}
\]

(f) The relaxation scalars $(k_x, k_z, k_p, k_y)$ are step-size that regulates the convergence mechanism. They are normally chosen from the interval $(0, 1)$, but this choice may not result in an optimal number of iterations. It is important to note that the optimal choice of $k_x, k_z, k_p, k_y \in (0, 1]$ is problem dependent, requiring that the proposed algorithm is run several times from Step 1 to Step 4. These values are initially set as $k_x = k_z = k_p = k_y = 1$ for the first run of the algorithm from Step 1 to Step 4, and then the algorithm is run with different values ranging from 0.1 to 0.9. The value that provides the optimal number of iterations can then be determined. The parameters $r_1, r_2$ and $r_3$ are to enhance convexity, leading to the improvement of the convergence of the algorithm.

4. Illustrative Example. Consider a continuous stirred-tank reactor problem [7]. The real plant is given by

\[
x_1(k + 1) = x_1(k) - 0.02(x_1(k) + 0.25) + 0.01(x_2(k) + 0.5) \exp\left(\frac{25x_1(k)}{x_1(k) + 2}\right)
\]

\[-0.01(x_1(k) + 0.25)u(k) + \omega_1(k)
\]

\[
x_2(k + 1) = 0.99x_2(k) - 0.005 - 0.01(x_2(k) + 0.5) \exp\left(\frac{25x_1(k)}{x_1(k) + 2}\right) + \omega_2(k)
\]

for $k = 0, 1, \ldots, 77$, with initial condition
and the output measurement is $y(k) = x_1(k) + \eta(k)$, where the expected cost function

$$\min_{u(k)} J_0(u) = 0.01 \sum_{k=0}^{N-1} E[(x_1(k))^2 + (x_2(k))^2 + 0.1(u(k))^2]$$

is to be minimized. Here, $\omega(k) = [\omega_1(k) \quad \omega_2(k)]^T$ and $\eta(k)$ are Gaussian white noise sequences with their respective covariance given by $Q_\omega = 10^{-3}I_2$ and $R_\eta = 10^{-3}$.

This problem is referred to as Problem (P).

To obtain an optimal output solution, which is close enough to the real output, we simplify Problem (P) and propose the following model-based optimal control problem as Problem (M) given below:

$$\min_{u(k)} J_1(u) = \frac{1}{2} \sum_{k=0}^{N-1} \left[(\bar{x}_1(k))^2 + (\bar{x}_2(k))^2 + 0.1(u(k))^2 + 2\gamma(k)\right]$$

subject to

$$\begin{bmatrix} \bar{x}_1(k+1) \\ \bar{x}_2(k+1) \end{bmatrix} = \begin{bmatrix} 1.0895 & 0.0184 \\ -0.1095 & 0.9716 \end{bmatrix} \begin{bmatrix} \bar{x}_1(k) \\ \bar{x}_2(k) \end{bmatrix} + \begin{bmatrix} -0.003 \\ 0.000 \end{bmatrix} u(k) + \begin{bmatrix} \alpha_{11}(k) \\ \alpha_{12}(k) \end{bmatrix}$$

with the initial condition $\bar{x}(0) = [0.05 \quad 0]^T$, and the adjusted parameters $\gamma(k), \alpha_3(k), \alpha_2(k),$ and $\alpha_1(k) = [\alpha_{11}(k) \quad \alpha_{12}(k)]^T$.

After running the iterative algorithm, the simulation result is shown in Table 1, where the iteration number is 19 with the final cost 0.0164. It is almost 99% of the reduction to obtain the optimal cost.

**Table 1. Algorithm performance**

<table>
<thead>
<tr>
<th>Iteration number</th>
<th>Elapsed time (s)</th>
<th>Initial cost $J^*_i$</th>
<th>Final cost $J^*_f$</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
<td>2.233</td>
<td>2.0847</td>
<td>0.0164</td>
</tr>
</tbody>
</table>

The output residual of the proposed approach is 0.016748, which is smaller than the output residual of the filtering solution given by 0.034731; see [5]. Figures 1, 2 and 3 show, respectively, the trajectories of final control, final state, and final output. From these figures, the trajectories of control and state are smoothly free from disturbance, and the trajectory of output is regulated closely to the real output. These trajectories are then compared to the filtering trajectories of final control, final state and final output as shown in Figures 4, 5 and 6, respectively. It is concluded that the regulated output trajectory tracks the real output trajectory efficiently as well as giving the smallest output residual.

5. **Concluding Remarks.** An output regulation, which is added into the model-based optimal control problem, was discussed in this paper. The proposed iterative approach with adjustable parameters is for solving the discrete-time nonlinear stochastic optimal control problem in spite of model-reality differences. The expected trajectories of state and control were obtained and the output was measured deterministically. By introducing an adjustable parameter to the expected output, the
Figure 1. Final control $u(k)$ – regulation case

Figure 2. Final state $\bar{x}(k)$ – regulation case
Figure 3. Final output $\bar{y}(k)$ and real output $y(k)$ – regulation case

Figure 4. Final control $u(k)$ – filtering case
Figure 5. Final state $\bar{x}(k)$ – filtering case

Figure 6. Final output $\hat{y}(k)$ and real output $y(k)$ – filtering case
regulated output approximates closely to the real output with the smallest output error compared to the filtering solution. It is highly recommended, without applying the Kalman filtering theory, the output regulation solution is prior to the filtering solution and the efficiency of the proposed approach is shown.

REFERENCES


Received May 2014; 1st revision January 2015; final revision March 2015.

E-mail address: slkek@uthm.edu.my
E-mail address: mismail@utm.my