Scientific Research Publishing

Least Squares Solution for Discrete Time Nonlinear Stochastic Optimal Control Problem with Model-Reality Differences

Sie Long Kek[1], Jiao Li[2], Kok Lay Teo[3]

[1]Center for Research on Computational Mathematics, Universiti Tun Hussein Onn Malaysia, Batu Pahat, Malaysia
[2]School of Mathematics and Computing Science, Changsha University of Science and Technology, Changsha, China
[3]Department of Mathematics and Statistics, Curtin University of Technology, Perth, Australia
Email: slkek@uthm.edu.my, lijiaoflying@163.com, K.L.Teo@curtin.edu.au

Abstract

In this paper, an efficient computational approach is proposed to solve the discrete time nonlinear stochastic optimal control problem. For this purpose, a linear quadratic regulator model, which is a linear dynamical system with the quadratic criterion cost function, is employed. In our approach, the model-based optimal control problem is reformulated into the input-output equations. In this way, the Hankel matrix and the observability matrix are constructed. Further, the sum squares of output error is defined. In these point of views, the least squares optimization problem is introduced, so as the differences between the real output and the model output could be calculated. Applying the first-order derivative to the sum squares of output error, the necessary condition is then derived. After some algebraic manipulations, the optimal control law is produced. By substituting this control policy into the input-output equations, the model output is updated iteratively. For illustration, an example of the direct current and alternating current converter problem is studied. As a result, the model output trajectory of the least squares solution is close to the real output with the smallest sum squares of output error. In conclusion, the efficiency and the accuracy of the approach proposed are highly presented.

1. Introduction

Stochastic dynamical system is a practical system in modeling and simulating

the real-world problems. The behavior of the fluctuation, which is caused by the effect of noise disturbance in the dynamical system to represent the real situation, rises to the attention of many researchers. See for examples, [1]-[8]. As such, optimization and control of the stochastic dynamical system are a challenging topic. In the past decades, the research outcomes, both for theoretical results and development of the algorithms, are well-defined [9] [10] [11] [12] [13]. In general, the real processes, which are modeled into the stochastic optimal control problem, are mostly nonlinear process. Their actual models in natural could be necessary unknown. Applying the linear model, including the Kalman filtering theory, to solve the stochastic optimal control problems, in particular, for nonlinear case, give a simplification methodology instead [14] [15] [16] [17].

Recently, the integrated optimal control and parameter estimation (ICOPE) algorithm, which solves the linear model-based optimal control problem iteratively, is proposed in the literature [18] [19]. The purpose of this algorithm is to provide the optimal solution of the discrete time nonlinear stochastic optimal control problem with the different structure and parameters. In this algorithm, the adjusted parameters are added into the model used so as the differences between the real plant and the model used can be calculated, in turn, update the optimal solution of the model used during the computation procedure. The integration of system optimization and parameter estimation is the main feature of this algorithm, where the principle of model-reality differences and the Kalman filtering theory are incorporated. Once the convergence is achieved, the iterative solution approximates to the correct optimal solution of the original optimal control problem, in spite of the model-reality differences. However, increasing the accuracy of the model output to track the real output would be necessary required.

In this paper, an efficient matching scheme, which diminishes the adjusted parameters, is established deliberately. In our approach, the model-based optimal control problem is simplified from the discrete-time nonlinear stochastic optimal control problem. Then, this model-based optimal control problem is reformulated into the input-output equations. During the formulation of the input-output equations, the Hankel matrix and the observability matrix are constructed. These matrices capture the characteristic of the model used into the output measurement. By virtue of this, the least square optimization problem is introduced. From the validation of the first order necessary condition, the normal equation is resulted and the optimal control law is updated accordingly on the recursion formula. As a result of this, the sum squares of the output error could be minimized demonstratively. Hence, the efficiency of the algorithm proposed is highly presented.

The rest of the paper is organized as follows. In Section 2, the discrete time nonlinear stochastic optimal control problem and its simplified model-based optimal control problem are described. In Section 3, the system optimization with the least squares updating scheme is discussed. The computation procedure is summarized as the iterative algorithm. In Section 4, the current converter prob-

lem is illustrated and the result is demonstrated. Finally, some concluding remarks are made.

2. Problem Description

Consider a general class of dynamical system given by

$$x(k+1) = f\left(x(k), u(k), k\right) + G\omega(k) \tag{1a}$$

$$y(k) = h\left(x(k), k\right) + \eta(k) \tag{1b}$$

where $u(k) \in \mathfrak{R}^m$, $k = 0, 1, \cdots, N-1$, $x(k) \in \mathfrak{R}^n$, $k = 0, 1, \cdots, N$, and $y(k) \in \mathfrak{R}^p$, $k = 0, 1, \cdots, N$, are, respectively, the control sequence, the state sequence and the output sequence. The terms $\omega(k) \in \mathfrak{R}^q$, $k = 0, 1, \cdots, N-1$, and $\eta(k) \in \mathfrak{R}^p$, $k = 0, 1, \cdots, N$, are, respectively, process noise sequences and output noise sequences. Both of these noise sequences are the stationary Gaussian white noise sequences with zero mean and their covariance matrices are given by Q_ω and R_η, respectively, where Q_ω is a $q \times q$ positive definite matrix and R_η is a $p \times p$ positive definite matrix. In addition, G is an $n \times q$ process noise coefficient matrix, $f : \mathfrak{R}^n \times \mathfrak{R}^m \times \mathfrak{R} \to \mathfrak{R}^n$ represents the plant dynamics and $h : \mathfrak{R}^n \times \mathfrak{R} \to \mathfrak{R}^p$ is the output measurement channel.

Here, the aim is to find the control sequence $u(k)$, $k = 0, 1, \cdots, N-1$, such that the following cost function

$$g_0(u) = E\left[\varphi\left(x(N), N\right) + \sum_{k=0}^{N-1} L\left(x(k), u(k), k\right)\right] \tag{2}$$

is minimized over the dynamical system in Equation (1), where $\varphi : \mathfrak{R}^n \times \mathfrak{R} \to \mathfrak{R}$ is the terminal cost and $L : \mathfrak{R}^n \times \mathfrak{R}^m \times \mathfrak{R} \to \mathfrak{R}$ is the cost under summation. The cost function g_0 is the scalar function and $E[\cdot]$ is the expectation operator. It is assumed that all functions in Equations (1) and (2) are continuously differentiable with respect to their respective arguments.

The initial state is

$$x(0) = x_0$$

where $x_0 \in \mathfrak{R}^n$ is a random vector with mean and variance given, respectively, by

$$E[x_0] = \bar{x}_0 \quad \text{and} \quad E\left[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^\mathrm{T}\right] = M_0.$$

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

This problem is regarded as the discrete-time nonlinear stochastic optimal control problem and is referred to as Problem (P). Notice that the structure of Problem (P) is complex, and the exact solution of Problem (P) is, in general, unable to be obtained. In view of these, Problem (P) is proposed to be solved via solving a simplified model-based optimal control problem iteratively. For this purpose, let this simplified model-based optimal control problem, which is referred to as Problem (M), be given below:

$$\min_{u(k)} g_1(u) = \frac{1}{2}\overline{x}(N)^{\mathrm{T}} S(N)\overline{x}(N) + \frac{1}{2}\sum_{k=0}^{N-1}\left(\overline{x}(k)^{\mathrm{T}} Q\overline{x}(k) + u(k)^{\mathrm{T}} Ru(k)\right)$$

subject to (3)

$$\overline{x}(k+1) = A\overline{x}(k) + Bu(k), \quad \overline{x}(0) = \overline{x}_0$$

$$\overline{y}(k) = C\overline{x}(k)$$

where $\overline{x}(k) \in \mathfrak{R}^n$, $k = 0,1,\cdots,N$, and $\overline{y}(k) \in \mathfrak{R}^p$, $k = 0,1,\cdots,N$, are, respectively, the expected state sequence and the expected output sequence. Here, 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, whereas $S(N)$ and Q are $n \times n$ positive semi-definite matrices, R is a $m \times m$ positive definite matrix and g_1 is the scalar cost function.

Notice that only solving Problem (M) actually would not give the optimal solution of Problem (P). However, by establishing an efficient matching scheme, it is possible to approximate the true optimal solution of Problem (P), in spite of model-reality differences. This could be done iteratively.

3. System Optimization with Least Square Updating Scheme

Now, let us define the Hamiltonian function for Problem (M) as follows:

$$H(k) = \frac{1}{2}\left(\overline{x}(k)^{\mathrm{T}} Q\overline{x}(k) + u(k)^{\mathrm{T}} Ru(k)\right) + p(k+1)^{\mathrm{T}}\left(A\overline{x}(k) + Bu(k)\right). \quad (4)$$

Then, from Equation (3), the augmented cost function becomes

$$g_1'(u) = \frac{1}{2}\overline{x}(N)^{\mathrm{T}} S(N)\overline{x}(N) + p(0)^{\mathrm{T}}\overline{x}(0)$$
$$- p(N)^{\mathrm{T}}\overline{x}(N) + \sum_{k=0}^{N-1}\left(H(k) - p(k)^{\mathrm{T}}\overline{x}(k)\right) \quad (5)$$

where $p(k) \in \mathfrak{R}^n$ is the appropriate multiplier to be determined later.

Applying the calculus of variation [20] [21] [22] to the augmented cost function in Equation (5), the following necessary conditions are resulted:

1) Stationary condition:

$$\frac{\partial H(k)}{\partial u(k)} = Ru(k) + B^{\mathrm{T}} p(k+1) = 0 \quad (6)$$

2) Costate equation:

$$\frac{\partial H(k)}{\partial \overline{x}(k)} = Q\overline{x}(k) + A^{\mathrm{T}} p(k+1) = p(k) \quad (7)$$

3) State equation:

$$\frac{\partial H(k)}{\partial p(k+1)} = A\overline{x}(k) + Bu(k) = \overline{x}(k+1) \quad (8)$$

with the boundary conditions $\overline{x}(0) = \overline{x}_0$ and $p(N) = S(N)\overline{x}(N)$.

3.1. State Feedback Optimal Control Law

The following theorem expresses the feedback control law based on the state that

can be used in solving Problem (M).

Theorem 1. For the given Problem (M), the optimal control law is the feedback control law defined by

$$u(k) = -K(k)\overline{x}(k) \tag{9}$$

where

$$K(k) = \left(B^T S(k+1) B + R\right)^{-1} B^T S(k+1) A \tag{10}$$

$$S(k) = A^T S(k+1)\left(A - BK(k)\right) + Q \tag{11}$$

with the boundary condition $S(N)$ given. Here, the feedback control law is a linear combination of the states. That is, the optimal control is linear state-variable feedback.

Proof. From Equation (6), the stationary condition can be rewritten as

$$Ru(k) = -B^T p(k+1). \tag{12}$$

Applying the sweep method [20] [21] [22],

$$p(k) = S(k)\overline{x}(k), \tag{13}$$

for $k = k+1$ in Equation (12) to yield

$$Ru(k) = -B^T S(k+1)\overline{x}(k+1). \tag{14}$$

Taking Equation (8) into Equation (14), we have

$$Ru(k) = -B^T S(k+1)\left(A\overline{x}(k) + Bu(k)\right).$$

After some algebraic manipulations, the feedback control law (9) is obtained, where Equation (10) is satisfied.

Now, substituting Equation (13) for $k = k+1$ into Equation (7), the costate equation is written as

$$p(k) = Q\overline{x}(k) + A^T S(k+1)\overline{x}(k+1) \tag{15}$$

and considering the state Equation (8) in Equation (15), the costate equation becomes

$$p(k) = Q\overline{x}(k) + A^T S(k+1)\left(A\overline{x}(k) + Bu(k)\right). \tag{16}$$

Hence, by applying the feedback control law (9) in Equation (16), and doing some algebraic manipulations, it could be seen that Equation (11) is satisfied after comparing the manipulation result to Equation (13). This completes the proof.

By substituting Equation (9) into Equation (8), the state equation becomes

$$\overline{x}(k+1) = \left(A - BK(k)\right)\overline{x}(k) \tag{17}$$

and the model output is measured from

$$\overline{y}(k) = C\overline{x}(k). \tag{18}$$

In view of this, the calculation procedure for obtaining the feedback control law for Problem (M) is summarized below:

Algorithm 1: Feedback control algorithm

Data   Given $A, B, C, Q, R, S(N), x_0, N$.

Step 0 Calculate $K(k)$, $k = 0, 1, \cdots, N-1$, and $S(k)$, $k = 0, 1, \cdots, N$, from Equations (10) and (11), respectively.

Step 1 Solve Problem (M) that is defined by Equation (3) to obtain $u(k)$, $k = 0, 1, \cdots, N-1$, and $\overline{x}(k)$, $\overline{y}(k)$, $k = 0, 1, \cdots, N$, respectively, from Equations (9), (17) and (18).

Step 2 Evaluate the cost function g_1 from Equation (3).

Remarks:

1) Data *A*, *B*, *C* can be obtained by the linearization of the real plant f and the output measurement h from Problem (P).

2) In Step 0, the offline calculation is done for $K(k)$, $k = 0, 1, \cdots, N-1$, and $S(k)$, $k = 0, 1, \cdots, N$.

3) The solution procedure, which the dynamical system is solved in Step 1, and the cost function is evaluated in Step 2, is known as system optimization.

3.2. Least Squares Updating Scheme

Now, we define the output error $r : \mathfrak{R}^m \to \mathfrak{R}^p$ given by

$$r(u) = y(k) - \overline{y}(k),\tag{19}$$

where the model output (18) is reformulated as

$$\overline{y}(k) = C\overline{x}(k) + Du(k).\tag{20}$$

where D is a $p \times m$ coefficient matrix. Formulate Equation (20) as the following input-output equations [23] for $k = 0, 1, \cdots, N$:

$$\begin{bmatrix} \overline{y}(0) \\ \overline{y}(1) \\ \overline{y}(2) \\ \vdots \\ \overline{y}(N) \end{bmatrix} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^N \end{bmatrix} \overline{x}_0 + \begin{bmatrix} D & 0 & 0 & \cdots & 0 \\ CB & D & 0 & \cdots & 0 \\ CAB & CB & D & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ CA^{N-1}B & CA^{N-2}B & CA^{N-3}B & \cdots & D \end{bmatrix} \begin{bmatrix} u(0) \\ u(1) \\ u(2) \\ \vdots \\ u(N-1) \end{bmatrix}.\tag{21}$$

For simplicity, we have

$$\overline{y} = E\overline{x}_0 + Fu\tag{22}$$

where

$$E = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^N \end{bmatrix} \quad \text{and} \quad F = \begin{bmatrix} D & 0 & 0 & \cdots & 0 \\ CB & D & 0 & \cdots & 0 \\ CAB & CB & D & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ CA^{N-1}B & CA^{N-2}B & CA^{N-3}B & \cdots & D \end{bmatrix}.$$

In addition, define the objective function $g_2 : \mathfrak{R}^m \to \mathfrak{R}$, which represents the sum squares of error (SSE), given by

$$g_2(u) = r(u)^{\mathrm{T}} r(u).\tag{23}$$

Then, an optimization problem, which is referred to as Problem (O), is stated as follows:

Problem (O):

Find a set of the control sequence $u(k)$, $k = 0, 1, \cdots, N-1$, such that the objective function g_2 is minimized.

It is obviously noticed that for solving Problem (O), Taylor's theorem [24] [25] is applied to write the objective function g_2 as the second-order Taylor expansion about the current $u^{(i)}$ at iteration *i* as follows:

$$g_2\left(u^{(i+1)}\right) \approx g_2\left(u^{(i)}\right) + \left(u^{(i+1)} - u^{(i)}\right)^{\mathrm{T}} \nabla g_2\left(u^{(i)}\right) \\ + \frac{1}{2}\left(u^{(i+1)} - u^{(i)}\right)^{\mathrm{T}} \left(\nabla^2 g_2\left(u^{(i)}\right)\right)\left(u^{(i+1)} - u^{(i)}\right) \tag{24}$$

where the higher-order terms are ignored and the notation ∇ represents the differential operator. The first-order condition in Equation (24) with respect to $u^{(i+1)}$ is expressed by

$$0 \approx \nabla g_2\left(u^{(i)}\right) + \left(\nabla^2 g_2\left(u^{(i)}\right)\right)\left(u^{(i+1)} - u^{(i)}\right). \tag{25}$$

Rearrange Equation (25) to yield the normal equation,

$$\left(\nabla^2 g_2\left(u^{(i)}\right)\right)\left(u^{(i+1)} - u^{(i)}\right) = -\nabla g_2\left(u^{(i)}\right). \tag{26}$$

Notice that the gradient ∇g_2 is calculated from

$$\nabla g_2\left(u^{(i)}\right) = 2\nabla r\left(u^{(i)}\right)^{\mathrm{T}} r\left(u^{(i)}\right) \tag{27}$$

and the Hessian of g_2 is computed from

$$\nabla^2 g_2\left(u^{(i)}\right) = 2\left(\nabla^2 r\left(u^{(i)}\right)^{\mathrm{T}} r\left(u^{(i)}\right) + \nabla r\left(u^{(i)}\right)^{\mathrm{T}} \nabla r\left(u^{(i)}\right)\right) \tag{28}$$

where $\nabla r\left(u^{(i)}\right)$ is the Jacobian matrix of $r\left(u^{(i)}\right)$, and its entries are denoted by

$$\left(\nabla r\left(u^{(i)}\right)\right)_{ij} = \frac{\partial r_i}{\partial u_j}\left(u^{(i)}\right) = F, \ i = 1, 2, \cdots, N-1; \ j = 1, 2, \cdots, N-1. \tag{29}$$

From Equations (27) and (28), Equation (26) can be rewritten as

$$\left(\nabla^2 r\left(u^{(i)}\right)^{\mathrm{T}} r\left(u^{(i)}\right) + \nabla r\left(u^{(i)}\right)^{\mathrm{T}} \nabla r\left(u^{(i)}\right)\right)\left(u^{(i+1)} - u^{(i)}\right) = -\nabla r\left(u^{(i)}\right)^{\mathrm{T}} r\left(u^{(i)}\right). \tag{30}$$

By ignoring the second-order derivative term, that is, the first term at the left-hand side of Equation (30), and define

$$\Delta u^{(i)} = u^{(i+1)} - u^{(i)}, \tag{31}$$

the normal equation given by Equation (26) is simplified to

$$\left(\nabla r\left(u^{(i)}\right)^{\mathrm{T}} \nabla r\left(u^{(i)}\right)\right)\Delta u^{(i)} = -\nabla r\left(u^{(i)}\right)^{\mathrm{T}} r\left(u^{(i)}\right). \tag{32}$$

Then, we obtain the following updating recurrence relation,

$$u^{(i+1)} = u^{(i)} + \Delta u^{(i)} \tag{33}$$

with the initial $u^{(0)}$ given, where

$$\Delta u^{(i)} = -\left(\nabla r\left(u^{(i)}\right)^{\mathrm{T}} \nabla r\left(u^{(i)}\right)\right)^{-1}\left(\nabla r\left(u^{(i)}\right)^{\mathrm{T}} r\left(u^{(i)}\right)\right). \tag{34}$$

Hence, Equations (33) and (34) are known as the least squares recursive equations, which are based on the Gauss-Newton recursion formula.

From the discussion above, the least-squares updating scheme for the control sequence is summarized below:

Algorithm 2: Least-squares updating scheme

Step 0 Given an initial $u^{(0)}$ and the tolerance ε. Set $i = 0$.

Step 1 Evaluate the output error $r\left(u^{(i)}\right)$ and the Jacobian matrix $\nabla r\left(u^{(i)}\right)$ from Equations (19) and (29), respectively.

Step 2 Solve the normal equation from Equation (32) to obtain $\Delta u^{(i)}$.

Step 3 Update the control sequence by using Equation (33). If $u^{(i+1)} = u^{(i)}$, within a given tolerance ε, stop; else set $i = i+1$ and repeat from Step 1 to Step 3.

Remarks:

1) In Step 1, the calculation of the output error $r\left(u^{(i)}\right)$ and the Jacobian matrix $\nabla r\left(u^{(i)}\right)$ are done online, however, the Jacobian matrix $\nabla r\left(u^{(i)}\right)$ could be done offline if it is independent from $u^{(i)}$.

2) In Step 2, the inverse of $\nabla r\left(u^{(i)}\right)^{\mathrm{T}} \nabla r\left(u^{(i)}\right)$ must be exist, and the value of $\Delta u^{(i)}$ represents the step-size for the control set-point.

3) In Step 3, the initial $u^{(0)}$ is calculated from (9). The condition $u^{(i+1)} = u^{(i)}$ is required to be satisfied for the converged optimal control sequence. The following 2-norm is computed and it is compared with a given tolerance to verify the convergence of $u(k)$:

$$\left\| u^{(i+1)} - u^{(i)} \right\| = \left(\sum_{k=0}^{N-1} \left\| u(k)^{(i+1)} - u(k)^{(i)} \right\| \right)^{1/2}. \tag{35}$$

4) In order to provide a convergence mechanism for the state sequence, a simple relaxation method is employed:

$$\bar{x}^{(i+1)} = \bar{x}^{(i)} + k_x \left(x - \bar{x}^{(i)} \right) \tag{36}$$

where $k_x \in (0,1]$, and x is the state sequence of the real plant.

4. Illustrative Example

In this section, an example of the direct current and alternating current (DC/AC) converter model is illustrated [26] [27] [28]. In this model, the real plant is a nonlinear dynamics structure and the output channel is a single output measurement, where the stationary Gaussian white noises are taken into the model. Since this model is a continuous time model, the simple discretization scheme is required. We define this model as Problem (P), while the simplified model, which is an expectation model, is employed as Problem (M). The cost function for both problems is a quadratic criterion function. The true solution of the original optimal control problem would be obtained by using the approach proposed and the solution procedure is implemented in the MATLAB environment.

Problem (P):

$$\min_u g_0(u) = E\left[\frac{1}{2} x(N)^{\mathrm{T}} S(N) x(N) + \frac{1}{2} \sum_{k=0}^{N-1} \left(x(k)^{\mathrm{T}} Q x(k) + u(k)^{\mathrm{T}} R u(k) \right) \right]$$

subject to

$$\dot{x}_1(t) = \frac{(x_2(t))^2}{x_1(t)} - 5x_1(t) + 5u(t) + \omega_1(t)$$

$$\dot{x}_2(t) = \frac{(x_2(t))^3}{(x_1(t))^2} - 7x_2(t) + \left(5\frac{x_2(t)}{x_1(t)} + 2x_1(t)\right)u(t) + \omega_2(t)$$

$$y(t) = x_2(t) + \eta(k)$$

with the initial $x(0) = (0.1,0)^{\mathrm{T}}$. Here, $\omega(k) = (\omega_1(k)\ \omega_2(k))^{\mathrm{T}}$ and $\eta(k)$ are Gaussian white noise sequences with their respective covariance given by $Q_\omega = 10^{-2}$ and $R_\eta = 10^{-1}$. The weighting matrices in the cost function are $S(N) = I_{2\times2}$, $Q = I_{2\times2}$ and $R = 1$.

Problem (M):

$$\min_u g_1(u) = \frac{1}{2}\overline{x}(N)^{\mathrm{T}} S(N)\overline{x}(N) + \frac{1}{2}\sum_{k=0}^{N-1}\left(\overline{x}(k)^{\mathrm{T}} Q\overline{x}(k) + u(k)^{\mathrm{T}} Ru(k)\right)$$

subject to

$$\begin{pmatrix}\overline{x}_1(k+1)\\\overline{x}_2(k+1)\end{pmatrix} = \begin{pmatrix}1-5\cdot T & 0\\0 & 1-7\cdot T\end{pmatrix}\begin{pmatrix}\overline{x}_1(k)\\\overline{x}_2(k)\end{pmatrix} + \begin{pmatrix}5\cdot T\\0.2\cdot T\end{pmatrix}u(k)$$

$$\overline{y}(k) = \overline{x}_2(k)$$

for $k = 0,1,\cdots,80$, with the sampling time $T = 0.01$ second.

The simulation result is shown in Table 1. The implementation of the algorithm proposed takes four iteration numbers to converge. The initial cost of 0.0429 unit is the value of cost function belong to Problem (M) before the iteration begins. After the convergence is achieved, the final cost of 116.9461 units is obtained. The original cost of 1.0881×10^3 is the value of cost function of the original optimal control problem. It is noticed that the cost reduction is 89.28 percent. The value of SSE, which is 8.046097×10^{-12}, shows the smallest differences between the real output and the model output.

The trajectories of final control and real control are shown in Figure 1 and Figure 2, respectively. The final control takes into the consideration of noise sequence during the calculation, while the real control is free from the noise disturbance. Figure 3 shows the trajectories of final output and real output. It can be seen that the final output tracks closely to the real output. Figure 4 shows the trajectories of final state and real state, where they are statistically identical. Figure 5 and Figure 6 show, respectively, the output errors after and before iterations. The accuracy of the model output is increased due on the decrease of the output error.

Table 1. Simulation result.

| Number of Iterations	Initial Cost	Final Cost	Original Cost	SSE
4	0.0429	116.9461	1.0881×10^3	8.046097×10^{-12}
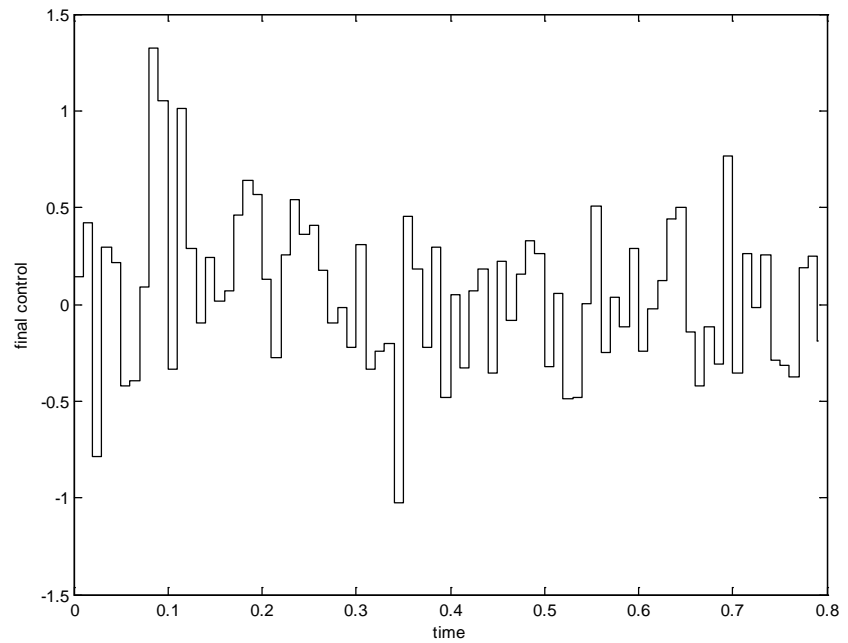
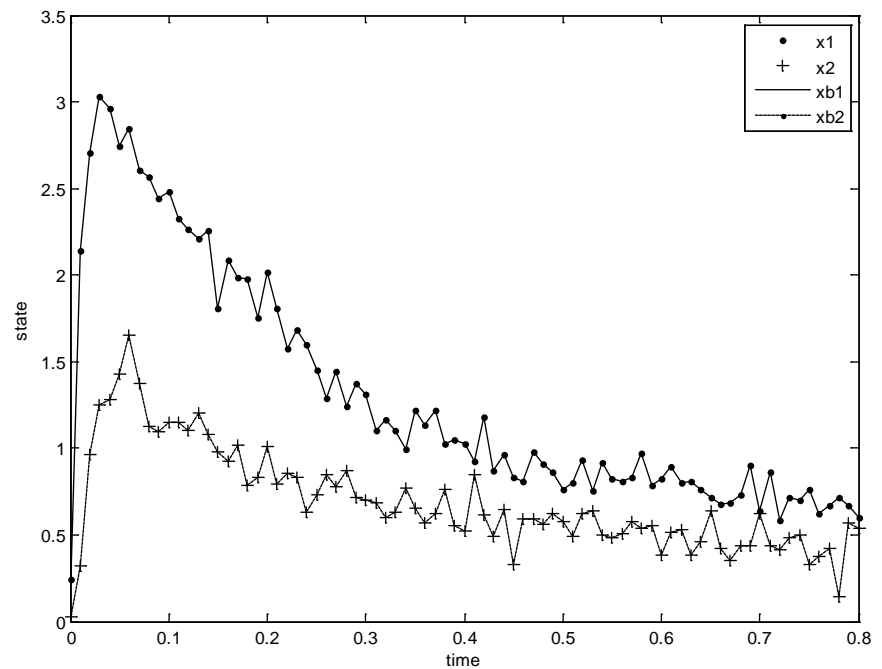**Figure 1.** Trajectory of final control.
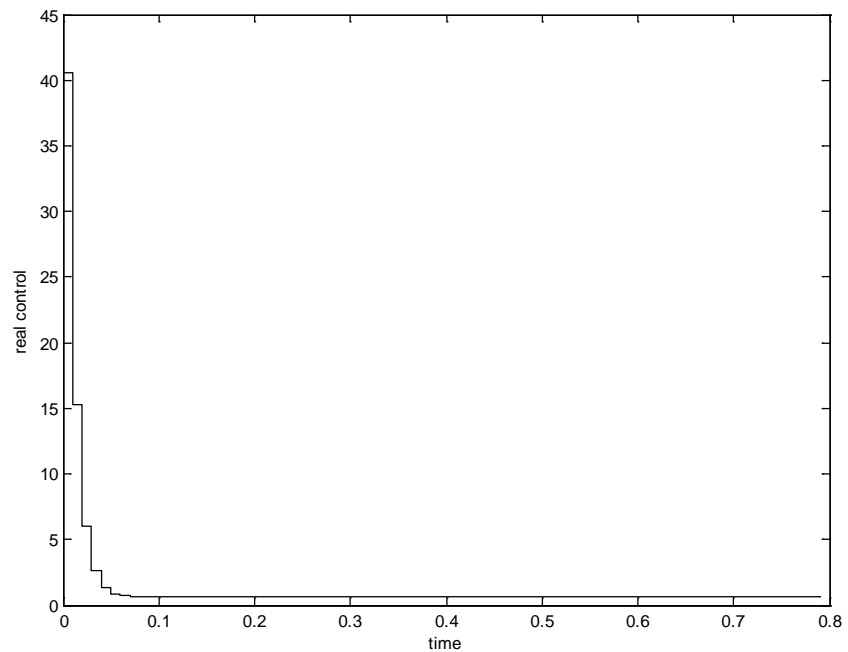
Figure 2. Trajectory of real control.

5. Concluding Remarks

A computational algorithm, which is equipped with the efficient matching scheme, was discussed in this paper. The linear model-based optimal control problem, which is simplified from the discrete time nonlinear stochastic optimal control problem, was solved iteratively. During the calculation procedure, the model used was reformulated into the input-output equations. Then, the least squares optimization problem was introduced. By satisfying the first order necessary condition, the normal equation was solved such that the least squares
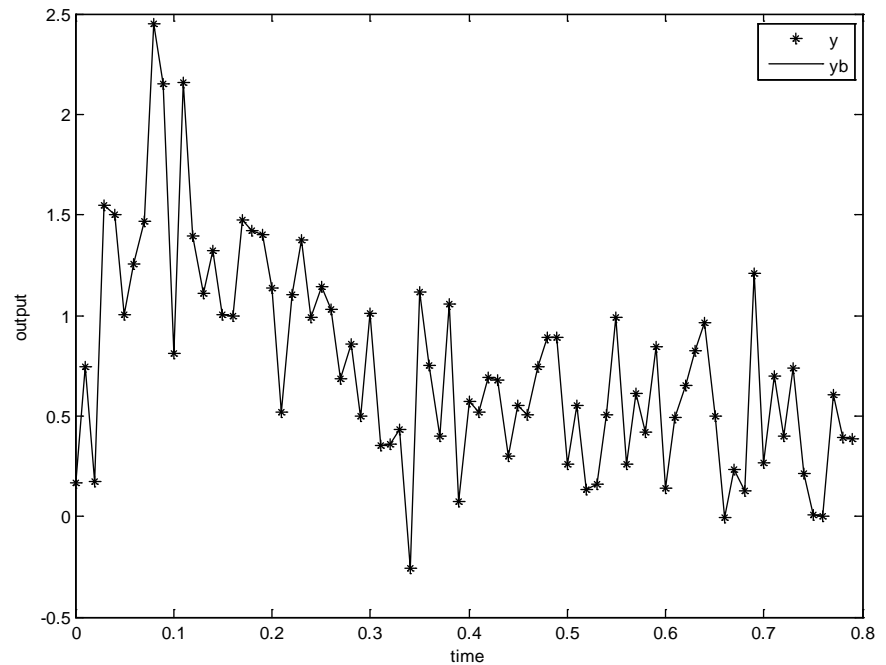
Figure 3. Trajectories of final output (−) and real output (∗).

Figure 4. Trajectories of final state (−) and real state (+).

recursion formula could be established. In this way, the control policy could be updated iteratively. As a result of this, the sum squares of the output errors was minimized, which indicates the equivalent of the model output to the real output. For illustration, an example of the direct current and alternating current converter model was studied. The result obtained showed the accuracy of the algorithm proposed. In conclusion, the efficiency of the algorithm proposed was highly presented.
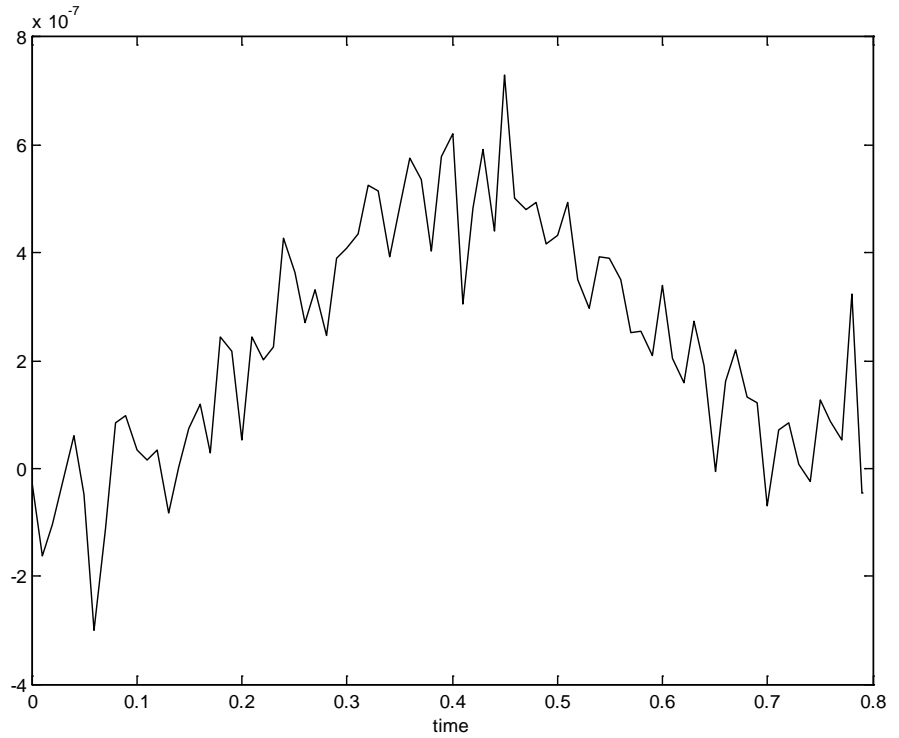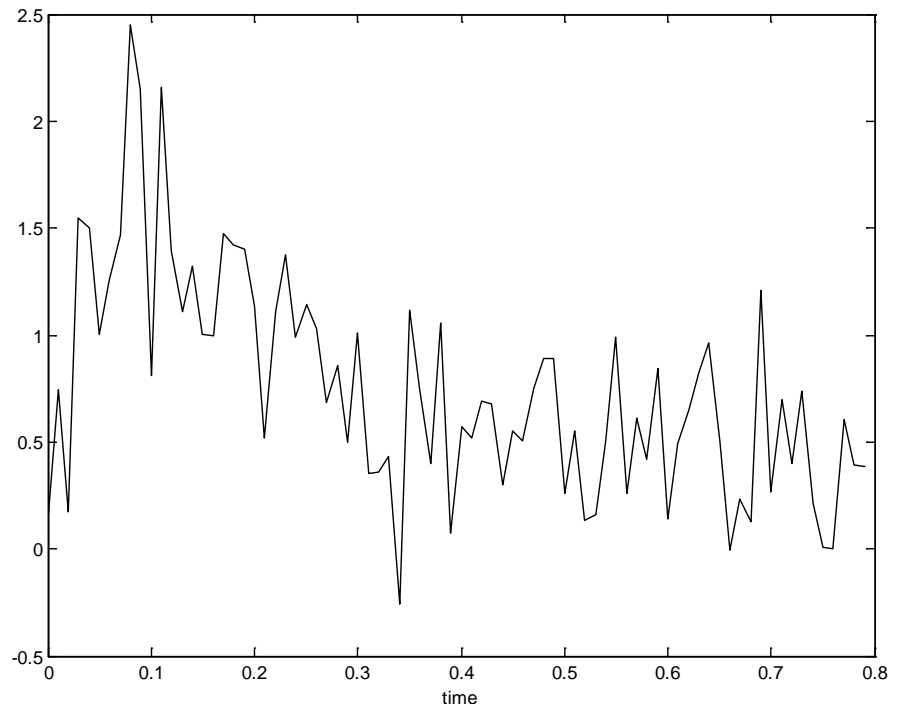
Figure 5. Output error after iteration.

Figure 6. Output error before iteration.

Acknowledgements

Scientific Research Publishing

NSF (11501053) of China and the fund (15C0026) of the Education Department of Hunan Province.

References

[1] Ahmed, N.U. (1973) Optimal Control of Stochastic Dynamical Systems. *Information and Control*, **22**, 13-30. https://doi.org/10.1016/S0019-9958(73)90458-0

[2] Ahmed, N.U. and Teo, K.L. (1974) Optimal Feedback Control for a Class of Stochastic Systems. *International Journal of Systems Science*, **5**, 357-365. https://doi.org/10.1080/00207727408920104

[3] Kliemann, W. (1983) Qualitative Theory of Stochastic Dynamical Systems—Applications to Life Sciences. *Bulletin of Mathematical Biology*, **45**, 483-506.

[4] Adomian, G. (1985) Nonlinear Stochastic Dynamical Systems in Physical Problems, *Journal of Mathematical Analysis and Applications*, **111**, 105-113. https://doi.org/10.1016/0022-247X(85)90203-3

[5] Li, X., Blount, P.L., Vaughan, T.L. and Reid, B.J. (2011) Application of Biomarkers in Cancer Risk Management: Evaluation from Stochastic Clonal Evolutionary and Dynamic System Optimization Points of View. *PLOS Computational Biology*, **7**, Article ID: e1001087. https://doi.org/10.1371/journal.pcbi.1001087

[6] Lopez, R.H., Ritto, T.G., Sampaio, R. and de Cursi, J.E.S. (2014) Optimization of a Stochastic Dynamical System. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, **36**, 257-264. https://doi.org/10.1007/s40430-013-0067-1

[7] Lozovanu, D. and Pickl, S. (2014) Optimization of Stochastic Discrete Systems and Control on Complex Networks: Computational Networks, Volume 12 of Advances in Computational Management Science. Springer, Cham Heidelberg, Switzerland.

[8] Mellodge, P. (2015) A Practical Approach to Dynamical Systems for Engineers. Woodhead Publishing, Swaston, Cambridge, UK.

[9] Craine, R. and Havenner, A. (1977) A Stochastic Optimal Control Technique for Models with Estimated Coefficients. *Econometrica*, **45**, 1013-1021. https://doi.org/10.2307/1912689

[10] Teo, K.L. Goh, C.J. and Wong, K.H. (1991) A Unified Computational Approach to Optimal Control Problems. Longman Scientific and Technical, New York.

[11] Zhang, K., Yang, X.Q., Wang, S. and Teo, K.L. (2010) Numerical Performance of Penalty Method for American Option Pricing. *Optimization Methods and Software*, **25**, 737-752. https://doi.org/10.1080/10556780903051930

[12] Rawlik, K., Toussaint, M. and Vijayakumar, S. (2012) On Stochastic Optimal Control and Reinforcement Learning by Approximate Inference. *Proceeding of* 2012 *Robotics*: *Science and Systems Conference*, Sydney, 9-13 July 2012, 3052-3056.

[13] Kek, S.L., Mohd Ismail, A.A. and Teo, K.L. (2015) A Gradient Algorithm for Optimal Control Problems with Model-Reality Differences. *Numerical Algebra, Control and Optimization*, **5**, 252-266.

[14] Socha, L. (2000) Application of Statistical Linearization Techniques to Design of Quasi-Optimal Active Control of Nonlinear Systems. *Journal of Theoretical and Applied Mechanics*, **38**, 591-605.

[15] Li, W. and Todorov, E. (2007) Iterative Linearization Methods for Approximately optimal Control and Estimation of Nonlinear Stochastic System. *International Journal of Control*, **80**, 1439-1453. https://doi.org/10.1080/00207170701364913

[16] Zitouni, F. and Lefebvre, M. (2014) Linearization of a Matrix Riccati Equation Associated to an Optimal Control Problem. *International Journal of Differential Equa-*

tions, **2014**, Article ID: 741390. https://doi.org/10.1155/2014/741390

[17] Hernandez-Gonzalez, M. and Basin, M.V. (2014) Discrete-Time Optimal Control for Stochastic Nonlinear Polynomial Systems. *International Journal of General Systems*, **43**, 359-371. https://doi.org/10.1080/03081079.2014.892256

[18] Kek, S.L., Teo, K.L. and Mohd Ismail, A.A (2010) An Integrated Optimal Control Algorithm for Discrete-Time Nonlinear Stochastic System. *International Journal of Control*, **83**, 2536-2545. https://doi.org/10.1080/00207179.2010.531766

[19] Kek, S.L., Teo, K.L. and Mohd Ismail, A.A. (2012) Filtering Solution of Nonlinear Stochastic Optimal Control Problem in Discrete-Time with Model-Reality Differences. *Numerical Algebra, Control and Optimization*, **2**, 207-222. https://doi.org/10.3934/naco.2012.2.207

[20] Bryson, A. and Ho, Y.C. (1975) Applied Optimal Control: Optimization, Estimation, and Control. Taylor & Francis, Abingdon.

[21] Lewis, F.L., Vrabie, D.L. and Syrmos, V.L. (2012) Optimal Control. 3rd Edition, John Wiley & Sons, Hoboken. https://doi.org/10.1002/9781118122631

[22] Kirk, D.E. (2004) Optimal Control Theory: An Introduction. Dover Publication Inc., Mineola.

[23] Verhaegen, M. and Verdult, V. (2007) Filtering and System Identification: A Least Squares Approach. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511618888

[24] Minoux, M. (1986) Mathematical Programming: Theory and Algorithms. John Wiley & Sons, Paris.

[25] Chong, K.P. and Zak, S.H. (2013) An Introduction to Optimization. 4th Edition, John Wiley & Sons, Hoboken.

[26] Poursafar, N., Taghirad, H.D. and Haeri, M. (2010) Model Predictive Control of Nonlinear Discrete Time Systems: A Linear Matrix Inequality Approach. *IET Control Theory and Applications*, **4**, 1922-1932. https://doi.org/10.1049/iet-cta.2009.0650

[27] Haeri, M. (2004) Improved EDMC for the Processes with High Variations and/or Sign Changes in Steady-State Gain. *COMPEL*, **23**, 361-380. https://doi.org/10.1108/03321640410510541

[28] Passino, K.M. (1989) Disturbance Rejection in Nonlinear Systems: Example. *IEE Proceedings D*, **136**, 317-323. https://doi.org/10.1049/ip-d.1989.0042

Scientific Research Publishing

Submit or recommend next manuscript to SCIRP and we will provide best service for you:

Accepting pre-submission inquiries through Email, Facebook, LinkedIn, Twitter, etc.
A wide selection of journals (inclusive of 9 subjects, more than 200 journals)
Providing 24-hour high-quality service
User-friendly online submission system
Fair and swift peer-review system
Efficient typesetting and proofreading procedure
Display of the result of downloads and visits, as well as the number of cited articles
Maximum dissemination of your research work

Submit your manuscript at: http://papersubmission.scirp.org/
Or contact am@scirp.org

Scientific Research Publishing