'94 KACC (1994. 10. 17 ~ 20)

Control of a Stochastic Nonlinear System by the Method of Dynamic Programming

Wan Sik Choi

TT&C Section, Satellite Communication Division, ETRI
Yusong P.O. BOX 106, Taejon, 305-600, Korea
Tel : 82-42-860-5610, Fax : 82-42-860-6430
Email : wschoi@kepler.etri.re.kr

Abstract

In this paper, we consider an optimal control problem of a nonlinear stochastic system. Dynamic programming approach is employed for the formulation of a stochastic optimal control problem. As an optimality condition, dynamic programming equation so called the Bellman equation is obtained, which seldom yields an analytical solution, even very difficult to solve numerically. We obtain the numerical solution of the Bellman equation using an algorithm based on the finite difference approximation and the contraction mapping method. Optimal controls are constructed through the solution process of the Bellman equation.We also construct a test case in order to investigate the actual performance of the algorithm.

1. Introduction

The method of dynamic programming developed by Richard Bellman was originated especially from the sequential stochastic programming problems, say inventory control [5,11]. This method can be considered as a powerful method in the sense that it provides necessary and sufficient conditions for optimality [4], and it falls well outside the domain of Linear Quadratic Gaussian control method [12].

Computational solution of the optimality condition, i.e. Bellman equation, obtained by the method of dynamic programming is very difficult to solve because of the complexity and the dimensionality. Therefore developing efficient computer-implementable algorithm for this equation has been a significant problem so far. In this paper we are going to consider finite time horizon control problem. For this case we have to solve the Bellman equation of parabolic type which is a nonlinear parabolic partial differential equation. So far there has not been much works on the computationl solution of the finite time horizon problem as far as the dynamic programming is concerned. An algorithm based on finite difference approximation and contraction mapping method is used for the computational solution. Test case is also constructed for validating the performance of algorithm. Finally in oder to illustrate how the method of dynamic programming is used for the control of nonlinear system, simple attitude control problem of a satellite is considered.

2. Optimality Condition

The Bellman equation also known as the dynamic programming equation arises in the general classes of stochastic control problems such as optimal stopping, regulation and tracking. We consider the following stochastic dynamic system which can be sufficiently general to include nonlinear and time varying parameters.

$$d(y^u_{t,x})_i(s) = m^u_i(s, y^u_{t,x}(s))ds + \sum_{j=1}^{n} \sigma^u_{i,j}(s, y^u_{t,x}(s))dw_j(s) \qquad (1)$$

$$i = 1,2,\ldots,n$$

where

$y^u_{t,x}(t) = x = \{x_1, x_2, \ldots, x_n\}$, w_j is a standard Wiener process, and $y^u_{t,x}(s)$ represents the solution of (1) at time s evolved from (t,x) with control u.

Equation (1) is called the Itô stochastic differential equation if m_i^u and σ_{ij}^u satisfy the so called Itô conditions [8,9,10]. We take a pretty general form of a Bolza type cost functional with finite time horizon, which can be employed for the problem of regulation or tracking.

$$J(t,x,u)=$$
$$E_{t,x}\{\int_t^{\min(T,\tau)} f^{u(s)}(s,y_{t,x}^{u(s)}(s))\exp\{-\int_t^s c^{u(\sigma)}(\sigma,y_{t,x}^{u(\sigma)}(\sigma))d\sigma\}ds$$
$$+\phi(y_{t,x}^{u(T)}(T))\exp\{-\int_t^T c^{u(\sigma)}(\sigma,y_{t,x}^{u(\sigma)}(\sigma))d\sigma\}\chi(T<\tau)\}$$
(2)

where

$E_{t,x}$: conditional expectation for $\{t,x\}$
τ : stopping time or exit time from the domain
$U=\{u_1,u_2,u_3,\dots\}$: set of all possible control actions
$u(s)=u_i$
$c^{u(\sigma)}$: discount factor

The stochastic control problem with finite time horizon by the method of dynamic programming involves a parabolic nonlinear partial differential equation as its optimality condition. It is called the Bellman equation or the dynamic programming equation.

Let

$$v(t,x)=\inf_u J(t,x,u). \tag{3}$$

Now applying the dynamic programming approach [5,8] and Itô's lemma [8,9,10] to (1),(2),(3) yield the following Bellman equation [3,8,12].

The parabolic Bellman equation :

$$\max_{u\in U}\{\frac{\partial v(t,x)}{\partial t}+L^u(t,x)v(t,x)-f^u(t,x)\}=0 \tag{4}$$

$$\text{for } (t,x)\in(0,T)\times\Omega$$
$$v(t,x)=0 \text{ for } (t,x)\in(0,T)\times\partial\Omega$$
$$v(0,x)=\phi(x) \text{ for } x\in\Omega$$
t : backward in time

$$L^u(t,x)=$$
$$-\frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n\sum_{k=1}^n\sigma_{ik}^u(t,x)\sigma_{jk}^u(t,x)\frac{\partial^2}{\partial x_i\partial x_j}-\sum_{i=1}^2 m_i^u(t,x)\frac{\partial}{\partial x_i}+c^u(t,x)$$

3. Algorithm for the Bellman equation

Finite difference discretization in both space and time for the Bellman equation yields the following equation [3].

$$\frac{v_{i,k}-v_{i,k-1}}{t_s}+\max_{u\in U}\{\sum_{j=1}^N D_{ij,k}^u v_{j,k}-f_{i,k}^u\}=0$$

where
$k=1,2,\dots,n_k$
$i=1,2,\dots,N$

In the above discrete Bellman equation, matrix $D_{ij,k}^u$ is obtained by applying the finite difference approximation to the operator $L^u v$ [3,12]. Modifying the discrete Bellman equation to the form of a fixed point iteration yields

$$v_{i,k}=(Iv)_{i,k}^u \tag{5}$$

where

$$v_{i,o}=\phi_{i,0}$$
$$(Iv)_{i,k}^u=\min_{u\in U}\{\sum_{j=1}^N H_{ij,k}^u v_{j,k}+\tilde{H}_{ii,k}^u v_{i,k-1}+\tilde{f}_{i,k}^u\}$$

$$H_{ij,k}^u=\begin{cases}\dfrac{-d_{ij,k}^u}{\frac{1}{t_s}+d_{ii,k}^u} & \text{if } i\neq j\\ 0 & \text{if } i=j\end{cases}$$

$$\tilde{H}_{ij,k}^u=\begin{cases}\dfrac{\frac{1}{t_s}}{\frac{1}{t_s}+d_{ii,k}^u} & \text{if } i=j\\ 0 & \text{if } i\neq j\end{cases}$$

$$\tilde{f}_{i,k}^u=\frac{f_{i,k}^u}{\frac{1}{t_s}+d_{ii,k}^u}$$

Since we are going to use 2-D case in the test case and the application of the algorithm, we consider the operator $L^u v$ of the following form.

$$L^u v(t,x_1,x_2)=$$
$$-(a_{11}^u)\frac{\partial^2 v}{\partial x_1^2}-2(a_{12}^u)\frac{\partial^2 v}{\partial x_1\partial x_2}-(a_{22}^u)\frac{\partial^2 v}{\partial x_2^2}+b_1^u\frac{\partial v}{\partial x_1}+b_1^u\frac{\partial v}{\partial x_1}+c^u v \tag{6}$$

where

$$a_{ij}^u=\frac{1}{2}\sum_{k=1}^2\sigma_{ik}^u\sigma_{jk}^u,\quad b_i^u=-m_i^u$$

4. Construction of A Test Case

Although an algorithm is verified mathematically, it is essential to do validation through the numerical implementation. This is because the actual performance of an algorithm might be different what we expected in term of mathematical reasoning.

In this section, we construct a mathematical example which can solve the Bellman equation exactly [12]. Thus, comparing the numerical solutions with the exact solutions, the performance of the algorithm can be examined rigorously.

By considering the boundary and initial conditions of the Bellman equation, we assume the following is the exact solution.

$$v(t,x_1,x_2) = \exp(\gamma t)\phi(x_1,x_2) + \lambda\sin(t)\xi(x_1,x_2) \qquad (7)$$

where

$$\phi(x_1,x_2) = x_1(x_1-\bar{a})x_2(x_2-\bar{b})$$
$$\xi(x_1,x_2) = \{x_1(x_1-\bar{a})\}^{\alpha}\{x_2(x_2-\bar{b})\}^{\beta}$$

Let the domain Ω of this test case be a rectangle (fig. 1). Then, $\Omega = \{(x_1,x_2) : 0 < x_1 < \bar{a},\ 0 < x_2 < \bar{b}\}$.

Refer the end of this section for parameters $\gamma, \lambda, \bar{a}, \bar{b}\ \alpha, \beta$.

Then $v(t,x_1,x_2)$ satisfies the boundary and initial conditions of the following Bellman equation.

$$\max_{u\in U}\{\frac{\partial v(t,x_1,x_2)}{\partial t} + L^u(t,x_1,x_2)v(t,x_1,x_2) - f^u(t,x_1,x_2)\} = 0 \qquad (8)$$
$$\text{for } (t,x_1,x_2)\in(0,T)\times\Omega$$

$$v(t,x_1,x_2) = 0 \text{ for } (t,x_1,x_2)\in(0,T)\times\partial\Omega$$
$$v(0,x_1,x_2) = \phi(x_1,x_2) \text{ for } (x_1,x_2)\in\Omega$$

$$L^u(t,x_1,x_2) = -\frac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2}\sigma^u_{i,k}\sigma^u_{j,k}\frac{\partial^2}{\partial x_i\partial x_j} - \sum_{i=1}^{2}m^u_i\frac{\partial}{\partial x_i} + c^u$$

We still have to define $\sigma^u_{i,j}$ and m^u_i for the numerical solution. For this, the following is obtained from the satellite dynamics [7].

$$\sigma^u_{ij} = \begin{bmatrix} 0 & 0 \\ 0 & -\delta_0(\sin x_1 + \eta_0 x_2) \end{bmatrix},$$
$$m^u_i = \begin{bmatrix} x_2 \\ -\sin x_1 + l\sin(2x_1) - \eta_0 x_2 + u \end{bmatrix} \qquad (9)$$

Refer the end of this section for parameters δ_0, η_0, l.

In order to avoid the degeneracy of the problem, we take the operator $L^u(t,x_1,x_2)$ in the following form. ε guarantees uniform parabolicity of the operator $L^u(t,x_1,x_2)$ and subsequently uniqueness of solution [8].

$$L^u(t,x_1,x_2) = -(a^u_{11}+\varepsilon)\frac{\partial^2}{\partial x_1^2} - (a^u_{22}+\varepsilon)\frac{\partial^2}{\partial x_2^2} + b^u_1\frac{\partial}{\partial x_1} + b^u_1\frac{\partial}{\partial x_1} + c^u$$

(10)

where

$$a^u_{ii} = \frac{1}{2}(\sigma^u_{ii})^2,\quad b^u_i = -m^u_i$$

Since $v(t,x_1,x_2)$ and $L^u(t,x_1,x_2)$ can be obtained from (7) and (10) respectively, every term in the Bellman equation(8) except $f^u(t,x_1,x_2)$ is known. Thus if we choose $f^u(t,x_1,x_2)$ so as to satisfy the equation (8), $v(t,x_1,x_2)$ becomes the exact solution of the Bellman equation. Let u^* be the optimal control and take only two values.

$$u^* \in U = \{u_1, u_2\}$$

Assume

$$u^* = u_1 \text{ for } 0 < x_2 \le \bar{l}, \text{ and } u^* = u_2 \text{ for } \bar{l} < x_2 \le b \text{ (fig. 1)}$$

If the optimal control is u_1, the Bellman equation becomes

$$\frac{\partial v(t,x_1,x_2)}{\partial t} + L^{u_1}(t,x_1,x_2)v(t,x_1,x_2) - f^{u_1}(t,x_1,x_2) = 0 \text{ for } u_1$$

(11)

and

$$\frac{\partial v(t,x_1,x_2)}{\partial t} + L^{u_2}(t,x_1,x_2)v(t,x_1,x_2) - f^{u_1}(t,x_1,x_2) \le 0 \text{ for } u_2$$

(12)

Let

$$g^u(t,x_1,x_2) = \frac{\partial v(t,x_1,x_2)}{\partial t} + L^u(t,x_1,x_2)v(t,x_1,x_2)$$

Then

$$f^u(t,x_1,x_2) = g^{u_1}(t,x_1,x_2) \text{ for } (11)$$
$$f^u(t,x_1,x_2) = g^{u_2}(t,x_1,x_2) + \zeta \text{ for any } \zeta > 0 \text{ for } (12)$$

For the case that u_2 is the optimal control, $f^u(t,x_1,x_2)$ can be chosen similarly. As a consequence $f^u(t,x_1,x_2)$ can now be determined explicitly using the following expressions.

$$f^u(t,x_1,x_2) =$$

$$\begin{cases} g^u(t,x_1,x_2) & \text{if } (x_2 \le \bar{l} \text{ and } u = u_1) \text{ or } (x_2 > \bar{l} \text{ and } u = u_2) \\ g^u(t,x_1,x_2) + \zeta & \text{if } (x_2 \le \bar{l} \text{ and } u = u_2) \text{ or } (x_2 > \bar{l} \text{ and } u = u_1) \end{cases}$$

With the above construction of test case, the Bellman equation can be solved not only exactly but also numerically . The following parameters are employed for the computational solution of the test case.

$$\bar{a} = 1 = \bar{b},\ \bar{l} = 0.5,\ \delta_0 = 0.3,\ \eta_0 = 0.5,\ l = 1,$$
$$\varepsilon = 3 = c = \alpha = \beta,\ \gamma = -2,\ \lambda = 2$$

5. Application to A Satellite Example

We consider an attitude control problem of a satellite of which orbit is possibly in Low Earth Orbit. Such a satellite experiences an oscillatory motion about its desired attitude because of the effects of gravity gradient and aerodynamic torques. This oscillation, called libration, is detrimental to the proper function of the satellite, such as pointing accuracy of a satellite. In general, the librational motion which mainly resulted from gravity gradient and aerodynamic torques is characterized by 3-D motion of roll, pitch and yaw. A parameter contributing most to the aerodynamic torque is the local atmospheric density. Since the density fluctuates and its fluctuation increases with altitude, it can be modelled as a stochastic process. As a simple example, we take only the plane pitch motion which is described by the following equation [7].

$$\theta'' + \sin\theta - l\sin 2\theta + \eta_0\theta' - u + \delta_0(\sin\theta + \eta_0\theta')w' = 0 \tag{13}$$

Let

$$\{y^u_{t,x}(s)\}_1 = \theta$$
$$\{y^u_{t,x}(s)\}_2 = \theta' = d\{y^u_{t,x}(s)\}_1$$

Then the above equation can be written as

$$d\begin{bmatrix}(y^u_{t,x})_1(s)\\(y^u_{t,x})_2(s)\end{bmatrix} = \begin{bmatrix}m^u_1(s,y^u_{t,x}(s))\\m^u_2(s,y^u_{t,x}(s))\end{bmatrix}ds + \begin{bmatrix}\sigma^u_{11}(s,y^u_{t,x}(s)) & \sigma^u_{12}(s,y^u_{t,x}(s))\\ \sigma^u_{21}(s,y^u_{t,x}(s)) & \sigma^u_{22}(s,y^u_{t,x}(s))\end{bmatrix} d\begin{bmatrix}w_1(s)\\w_2(s)\end{bmatrix} \tag{14}$$

where

$$m^u_1(s,y^u_{t,x}(s)) = (y^u_{t,x})_2(s)$$
$$m^u_2(s,y^u_{t,x}(s)) = -\sin\{(y^u_{t,x})_1(s)\} + l\sin\{2(y^u_{t,x})_1(s)\} - \eta_0(y^u_{t,x})_2(s) + u$$

$$\sigma^u_{11}(s,y^u_{t,x}(s)) = 0 = \sigma^u_{12}(s,y^u_{t,x}(s)) = \sigma^u_{21}(s,y^u_{t,x}(s))$$
$$\sigma^u_{22}(s,y^u_{t,x}(s)) = -\delta_0\{\sin((y^u_{t,x})_1(s)) + \eta_0(y^u_{t,x})_2(s)\}$$

Parameters for the above nonlinear pitch motion of equation can be obtained by applying a stochastic Liapunov function method [7,10]. The following stability condition is used to choose parameters(i.e. δ_0, η_0, l) for the numerical experiments.

$$\delta_0^2 < \frac{2\eta_0(1-2l)}{1+\eta_0^2(1-2l)} \tag{15}$$

The following set Ω is taken as the domain of this satellite example.

$$\Omega = \{(x_1,x_2) :\ -1 < x_1 < 1,\ -1 < x_2 < 1\}$$

Physically x_1 and x_2 represent pitch angle and its angular velocity respectively. By the domain we implicitly assume that x_1, x_2 vary within this domain. We take the following quadratic cost functional for the problem of minimizing pitch error.

$$J(t,x,u) = E_{t,x}[\int_t^{\min(T,\tau)} \exp\{-c(s-t)\}\{Y'(s)QY(s) + ru^2\}ds + \exp\{-c(T-t)\}Y'(T)SY(T)\chi(T<\tau)] \tag{16}$$

where

$$Y(s) = \begin{bmatrix}(y^u_{t,x})_1(s)\\(y^u_{t,x})_2(s)\end{bmatrix}$$

Take $Q = I = S$ for simplicity.

Comparing the above cost functional with the standard form (2), we obtain

$$f^{u(s)}(s,y^{u(s)}_{t,x}) = Y'(s)QY(s) + ru^2$$
$$\phi(y^{u(T)}_{T,x}(T)) = Y'(T)QY(T).$$

With the above $f^{u(s)}$ and ϕ, the Bellman equation (4) for the case of satellite example can now be solved using the algorithm (5) described in section 3.

6. Computational Results and Discussions

In this section, computational results for both of the test case and the application problem are presented and discussed. In order to measure the performance of the algorithm and also check the correctness of computer results, two kinds of errors are introduced [12].

Absolute error : $E^{\bar{n}}_{abs} = \max\limits_{i,t} |v^{\bar{n}}_{i,t} - v_{i,t}|$

Relative error : $E^{\bar{n}}_{rel} = \max\limits_{i,t} |v^{\bar{n}}_{i,t} - v^{\bar{n}-1}_{i,t}|$

where

$v^{\bar{n}}_{i,t}$ and $v_{i,t}$ represent the numerical solution at $\bar{n}_{th}$ iteration and exact solution respectively.

Since the algorithm is based on contraction mapping, the relative error has to decrease rapidly. Figure 2 shows this property. It also shows that the absolute error stays constant after certain number of iterations, which we call the steady state error. This gap of error is probably due to the Taylor series truncation and the number of grid points taken for finite difference approximation of the operator $L^u(t,x)$. Table 1 shows that error decreases as the number of grid points increases. This is reasonable in terms of finite difference approximation that more grid points give better solution.. Figure 3 shows the map of optimal controls obtained for the test case, which exactly matches with the a priori map (Figure 1). Based on the above discussions, it can be concluded that the algorithm gives correct solutions.

For the application problem, it is also observed that the convergence rate is fast. Figure 4 shows typical four different kinds of the map of optimal controls for the control problem of a satellite influenced by the disturbance torques. Some of the maps for certain times are omitted because they have small variations from these maps. Since the Bellman equation is solved backward in time, we have to reverse the time in order to interprete these maps. From the map it can be seen that near the final time (t = 0.001 ~ 0.004), there are more varieties in the map. This can be interpreted that near the final time, more varieties of control actions are required in order to achieve the desired goal of controls.

7. Conclusions

In this paper it is shown that how the dynamic programming method can be applied for the control of nonlinear system. The Bellman equation of parabolic type is solved by using the algorithm based on finite difference approximation and contraction mapping method. We show how a test case is constructed for testing an algorithm. Computational results of the test case show that the algorithm gives reliable solutions. As a result of computational solution, maps of optimal controls are obtained for the nonlinear attitude control problem of a satellite.

References

1. S. A. Belbas, " Numerical Solution of Certain Nonlinear Parabolic Partial Differential Equations," In T. L. Gill and W. W. Zachary, editors, Nonlinear Semigroups, Partial Differential Equations and Attractors, pp. 5-14, Berlin Springer Verlag, 1987. Lecture Notes in Mathematics, Vol. 1248.
2. S. A. Belbas and I. D. Mayergoyz, "Application of Fixed-Point Methods to Discrete Variational and Quasi-Variational Inequalities," Numer. Math., 51, pp631-654, 1987.
3. S. A. Belbas and I. D. Mayergoyz, "Iterative Schemes for Certain Time-Dependent Problems of Stochastic Optimal Control," Int. J. Control, Vol. 50 (No.5), pp1681-1698, 1989.
4. R. Rishel, "Necessary and Sufficient Dynamic Programming Conditions for Continuous Time Stochastic Optimal Control," SIAM J. Control, Vol. 8, No. 4, Nov. 1970, pp559-571.
5. R. Bellman, "Adaptive Control Processes : A Guided Tour," Princeton University Press, 1961, pp51-59.
6. E. V. Denardo, "Contraction Mapping in the Theory Underlying Dynamic Programming," SIAM Rev. 9, 1967, pp165-177
7. P. Sagirow, "Stochastic Methods in the Dynamics of Satellite," CISM Courses and Lectures - No. 57, Springer-Verlag, Udine 1970, pp48-57, pp67-73.
8. W. H. Fleming and R. W. Rishel, "Deterministc and Stochastic Optimal Control," Springer-Verlag, 1975, p114, p128.
9. K. Itô, "On a Formula Concerning Stochastic Differentials," Nagoya Math. Journal, Mathematical Insitute, Nagoya University, pp55-65, 1951.
10. H. Kushner, "Stochastic Stability and Control," Academic Press, 1967, p13.
11. G. Hadley, "Nonlinear and Dynamic Programming," Addison-Wesley Publishing Company, 1972, p350.
12. W. S. Choi, "Numerical Solution of Two Classes of Stochastic Optimal Control Problems," Ph.D. Disseetation, The University of Alabama, Tuscaloosa, U.S.A., 1992, p2, pp26 - 28.

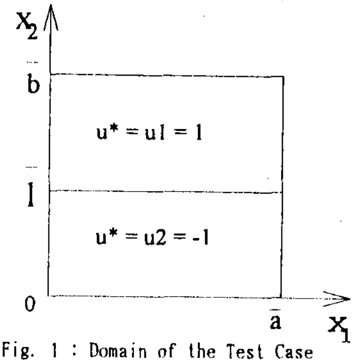

Fig. 1 : Domain of the Test Case

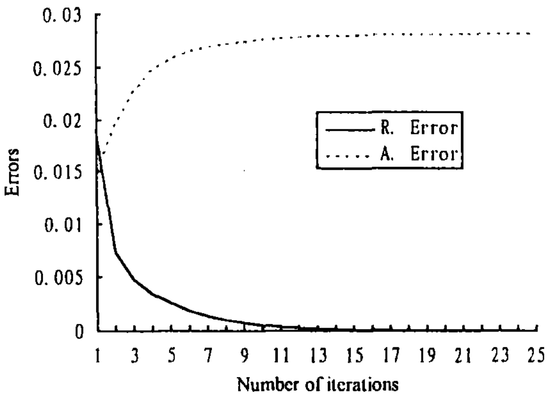

Figure 2 : Errors versus number of iterations

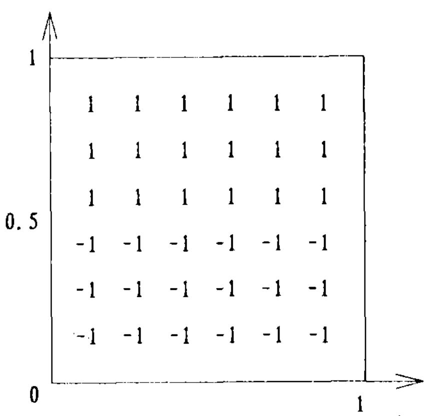

Fig. 3 : Map of optimal controls (6x6)

1	-1	-1	-1	-1	-1
1	1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1
1	1	1	1	1	1
1	1	1	1	1	1
1	1	1	1	1	-1

t = 0.001 ~ 0.004

1	-1	-1	-1	-1	-1
1	1	1	1	1	1
1	-1	-1	-1	-1	1
1	1	1	1	1	1
1	1	1	1	1	1
1	1	1	1	-1	-1

t = 0.006 ~ 0.008

1	1	-1	-1	-1	-1
1	1	1	1	1	1
1	1	1	1	1	1
1	1	1	1	1	1
1	1	1	1	1	-1
-1	-1	-1	-1	-1	-1

t = 0.018 ~ 0.027

1	1	1	1	1	1
1	1	1	1	1	1
1	1	1	1	1	1
1	1	1	1	1	1
-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1

t = 0.036 ~ 0.1

Figure 4. : Map of optimal controls

Table 1. : Effect of number of grid points

Grid	t_s	n_k	S. S. E.	Parameters
4 x 4	0.02	5	0.2804E-01	$\varepsilon = 3 = c = \alpha = \beta$
6 x 6	0.01	10	0.2236E-01	$\gamma = -2,\ \lambda = 2$
8 x 8	0.006	17	0.1831E-01	

Note : S. S. E. : Steady State Error (absolute)