IOSR Journal of Mathematics (IOSR-JM)
e-ISSN: 2278-5728, p-ISSN:2319-765X. Volume 9, Issue 5 (Jan. 2014), PP 07-13
www.iosrjournals.org

Numerical Experiment with Dynamic Programming in Solving Continuous-Time Linear Quadratic Regulator Problems

F. M. ADERIBIGBE and K. J. ADEBAYO
Department of Mathematical Sciences, Ekiti State University, Ado Ekiti, Nigeria.

Abstract: *This paper investigates and discusses the method of dynamic programming in solving Bolza's cost form of Linear Quadratic Regulator Problems (LQRP). It is the desire of the authors of this paper to experiment numerically the solution of this class of problem using dynamic programming to solve for the optimal controls and the trajectories compared with other numerical methods with a view to further improving the results. The method uses the principle of optimality to reduce mathematically the number of calculations required to determine the optimal control law as well as the corresponding optimal cost functional.*
Keywords: *Continuous-Time Linear Regulator Problem, Optimal Control, Discretization and Dynamic Programming.*

I. Introduction

The performance measure to be minimized is a continuous-time linear quadratic regulator problem considered by [2] and [4] as Bolza problem:

Problem (P1):

$$J(x,t_0,t_f,u(.)) = \frac{1}{2}x^T(t_f)Hx(t_f) + \frac{1}{2}\int_0^{t_f}\{x^T(t)Q(t)x(t) + u^T(t)R(t)u(t)\}dt \qquad 1.1$$

Subject to the differential equation

$$\dot{x}(t) = A(t)x(t) + B(t)u(t), \quad 0 \le t \le t_f, \qquad 1.2$$

where H and $Q(t)$ are real symmetric positive semi-definite $n \times n$ matries. $R(t)$ is a real symmetric positive definite $m \times m$ matrix, the initial time and the final time, t_f, are specified. $x(t)$ is an *n*-dimensional state vector, $u(t)$ is the *m*-dimensional plant control input vector. $u(t)$ and $x(t)$ are not constrained by any boundaries. A and B are specified constants which are not necessarily positive. If $H = 0$, (1.1) is called a Lagrange problem, but if Q(t) and R(t) are both zero matrices, it is called a Mayer problem. The linear plant dynamics and quadratic performance criteria (1.1) and (1.2) are referred to as linear regulator problem.

According to [3], before the numerical procedure of dynamic programming can be applied, the state equation dynamic, (1.2), must be approximated by a difference equation, and the integral in the performance measure must be approximated by a summation. This can be done most conveniently by dividing the time interval $0 \le t \le t_f$ into N equal increments, Δt. Then, from (1.2), we have

$$\frac{x(t+\Delta t)-x(t)}{\Delta t} \approx A(t)x(t) + B(t)u(t) \qquad 1.3$$

$$x(t+\Delta t) - x(t) = \Delta tA(t)x(t) + \Delta tB(t)u(t)$$

$$x(t+\Delta t) = [1 + \Delta tA(t)]x(t) + \Delta tB(t)u(t). \qquad 1.4$$

Here, it will be assumed that Δt is small enough so that the control signal can be approximated by a piecewise-constant function that changes only at the instants

$$t = 0,\ \Delta t,\ 2\Delta t, \dots,\ (N-3)\Delta t,\ (N-2)\Delta t, (N-1)\Delta t \qquad 1.5$$

thus,

$$t = k\Delta t. \qquad 1.6$$

Putting (1.6) for *t* in (1.4), we obtain

$$x((k+1)\Delta t) = [1 + \Delta tA(k\Delta t)]x(k\Delta t) + \Delta tB(k\Delta t)u(k\Delta t), \forall\ k = 0,1, \dots, N-1 \qquad 1.7$$

where $x(k\Delta t)$ is referred to as the *kth* value of *x* and is denoted by $x(k)$.

With this condition, the system difference equation (1.7) can now be written as:

$$x(k+1) = [1 + \Delta tA(k)]x(k) + \Delta tB(k)u(k). \qquad 1.8$$

In a similar way, the performance measure, (1.1), becomes

$$\begin{aligned} J = {} & \frac{1}{2}x^T(N\Delta t)Hx(N\Delta t) + \frac{1}{2}\Big[\int_0^{\Delta t}\{x^T(0)Q(0)x(0) + u^T(0)R(0)u(0)\}dt \\ & + \int_{\Delta t}^{2\Delta t}\{x^T(\Delta t)Q(\Delta t)x(\Delta t) + u^T(\Delta t)R(\Delta t)u(\Delta t)\}dt \\ & + \int_{2\Delta t}^{3\Delta t}\{x^T(2\Delta t)Q(2\Delta t)x(2\Delta t) + u^T(2\Delta t)R(2\Delta t)u(2\Delta t)\}dt + \dots \\ & + \int_{(N-1)\Delta t}^{N\Delta t}\left\{\begin{matrix} x^T(N-1)\Delta tQ(N-1)\Delta tx(N-1)\Delta t \\ +u^T(N-1)\Delta tR(N-1)\Delta tu(N-1)\Delta t \end{matrix}\right\}dt\Big] \end{aligned} \qquad 1.9$$

$$J = \frac{1}{2}x^T(N\Delta t)Hx(N\Delta t) + \frac{1}{2}[t\big(x^T(0)Q(0)x(0) + u^T(0)R(0)u(0)\big)|_0^{\Delta t}$$
$$+ t\big(x^T(\Delta t)Q(\Delta t)x(\Delta t) + u^T(\Delta t)R(\Delta t)u(\Delta t)\big)|_{\Delta t}^{2\Delta t}$$
$$+ t\big(x^T(2\Delta t)Q(2\Delta t)x(2\Delta t) + u^T(2\Delta t)R(2\Delta t)u(2\Delta t)\big)|_{2\Delta t}^{3\Delta t} + \dots$$
$$+ t(x^T(N-1)\Delta tQ(N-1)\Delta tx(N-1)\Delta t$$
$$+ u^T(N-1)\Delta tR(N-1)\Delta tu(N-1)\Delta t)|_{(N-1)\Delta t}^{N\Delta t}] \quad 1.10$$

$$J = \frac{1}{2}x^T(N\Delta t)Hx(N\Delta t) + \frac{1}{2}[\Delta t\big(x^T(0)Q(0)x(0) + u^T(0)R(0)u(0)\big)$$
$$+ \Delta t\big(x^T(\Delta t)Q(\Delta t)x(\Delta t) + u^T(\Delta t)R(\Delta t)u(\Delta t)\big)$$
$$+ \Delta t\big(x^T(2\Delta t)Q(2\Delta t)x(2\Delta t) + u^T(2\Delta t)R(2\Delta t)u(2\Delta t)\big) + \dots$$
$$+ \Delta t(x^T(N-1)\Delta tQ((N-1)\Delta t)x(N-1)\Delta t$$
$$+ u^T(N-1)\Delta tR((n-1)\Delta t)u(N-1)\Delta t] \quad 1.11$$

On applying the same condition as in (1.8) to (1.11), we obtain

$$J = \frac{1}{2}x^T(N)Hx(N) + \frac{1}{2}\Delta t[\big(x^T(0)Q(0)x(0) + u^T(0)R(0)u(0)\big)$$
$$+ \big(x^T(1)Q(1)x(1) + u^T(1)R(1)u(1)\big) + \big(x^T(2)Q(2)x(2) + u^T(2)R(2)u(2)\big)$$
$$+ \dots + (x^T(N-1)Q(N-1)x(N-1) + x^T(N-1)Rx(N-1))] \quad 1.12$$

Then, (1.12) can now be written using summative convention as:

$$J = \frac{1}{2}x^T(N)Hx(N) + \frac{1}{2}\Delta t\sum_{k=0}^{N-1}[x^T(k)Q(k)x(k) + u^T(k)R(k)u(k)] \quad 1.13$$

From the above, (1.13) and (1.8) is the discrete-time counterpart form of the state and performance measure of the continuous-time linear regulator problem (1.1) and (1.2) respectively. This will lead us to generating a recurrence relation for the dynamic programming in the next section.

II. Recurrence Relation of Dynamic Programming for LRP

Haven't been able to change the continuous-time linear regulator problem (1.1) and (1.2) into discrete-time linear regulator plant described by the *n*-dimensional equation

$$x(k+1) = [1 + \Delta tA(k)]x(k) + \Delta tB(k)u(k), \quad k = 0, 1, 2, \dots, N \quad 2.1$$
$$x(0) = x_0 \quad 2.2$$

At time k, we are left with finding the optimal control $u^*(x(k)k)$ that minimizes the performance measure

$$J = \frac{1}{2}x^T(N)Hx(N) + \frac{1}{2}\Delta t\sum_{k=0}^{N-1}[x^T(k)Q(k)x(k) + u^T(k)R(k)u(k)], \quad 2.3$$

where H and $Q(k)$ are real symmetric positive semi definite matrices, $R(k)$ is a real symmetric positive definite matrix, and N is a fixed integer $(N > 0)$.

The corresponding Hamiltonian is

$$\mathcal{H} = \frac{1}{2}x^T(k)Hx(k) + \frac{1}{2}x^T(k)Q(k)x(k) + \frac{1}{2}u^T(k)Ru(k)$$
$$+ P^T\big[\big(1 + A(k)\big)x(k) + B(k)u(k)\big], \quad 2.4$$

where the matrices $Q(k)$ and $R(k)$ are functions of the stage k. The weighting matrices $Q(k)$ and $R(k)$ represent the individual component weights on the state and control respectively over the sampling period and for most problems of interest, the sampling interval is constant. Without loss of generality and for sake of notational simplification, we assume that the matrices A, B, Q, and R are constant and begin by defining

$$J_{NN}\big(x(N)\big) = \frac{1}{2}x^T(N)Hx(N) \quad 2.5$$

where J_{NN} is the cost of reaching the final state value $x(N)$. Next, we define

$$J_{N-1,N}(x(N-1), u(N-1)) \triangleq \frac{1}{2}[x^T(N-1)Q(N-1)x(N-1)$$
$$+ u^T(N-1)R(N-1)u(N-1) + x^T(N)Hx(N)]. \quad 2.6$$

On putting (2.5) in (2.6), we obtain

$$J_{N-1,N}\big(x(N-1), u(N-1)\big) = \frac{1}{2}[x^T(N-1)Q(N-1)x(N-1)$$
$$+ u^T(N-1)R(N-1)u(N-1) + J_{NN}\big(x(N)\big)] \quad 2.7$$

which is the cost of operation during the interval $(N-1)\Delta t \le t \le N\Delta t$. It must be noted that, $J_{N-1,N}$ is dependent only on $x(N-1)$ *and* $u(N-1)$. Since $x(N)$ is related to $x(N-1)$*and* $u(N-1)$ through the state (2.1), so we write

$$J_{N-1,N}\big(x(N-1), u(N-1)\big) = \frac{1}{2}[x^T(N-1)Q(N-1)x(N-1)$$
$$+ u^T(N-1)R(N-1)u(N-1) + J_{NN}\big([1 + \Delta tA(k)]x(k) + \Delta tB(k)u(k)\big)]. \quad 2.8$$

The optimal cost is then

$$J^*_{N-1,N}\big(x(N-1)\big) = Min\{\tfrac{1}{2}[x^T(N-1)Q(N-1)x(N-1) + u^T(N-1)R(N-1)u(N-1) + J_{NN}\begin{pmatrix}[1+\Delta tA(k)]x(k)\\ +\Delta tB(k)u(k)]\end{pmatrix}]\}. \qquad 2.9$$

Since the optimal choice of $u(N-1)$ will depend on $x(N-1)$, so we denote the minimizing control by $u^*(x(N-1), N-1)$. The cost of operation over last two intervals is then given by

$$\begin{aligned}J_{N-2,N}(x(N-2), u(N-2), u(N-1)) = \tfrac{1}{2}[&x^T(N-2)Q(N-2)x(N-2)\\ &+ u^T(N-2)R(N-2)u(N-2)\\ &+x^T(N-1)Q(N-1)x(N-1)\\ &+ u^T(N-1)R(N-1)u(N-1) + x^T(N)Hx(N)]\end{aligned} \qquad 2.10$$

$$\begin{aligned}J_{N-2,N}(x(N-2), u(N-2), u(N-1)) = \tfrac{1}{2}[&x^T(N-2)Q(N-2)x(N-2)\\ &+ u^T(N-2)R(N-2)u(N-2)\\ &+ J_{N-1,N}(x(N-1), u(N-1))],\end{aligned} \qquad 2.11$$

where again we have used the dependence of $x(N)$ on $x(N-1)$ and $u(N-1)$.

As before, we observe that $J_{N-2,N}$ is the cost of a two-stage process with initial state $x(N-2)$. The optimal policy during the last two intervals is found from

$$\begin{aligned}J^*_{N-2,N}\big(x(N-2)\big) = Min\{\tfrac{1}{2}[&x^T(N-2)Q(N-2)x(N-2)\\ &+ u^T(N-2)R(N-2)u(N-2)\\ &+ J_{N-1,N}(x(N-1), u(N-1)]\}.\end{aligned} \qquad 2.12$$

Going by the principle of optimality, for this two-stage process, whenever the initial state $x(N-2)$ and initial decision $u(N-1)$ must be optimal with respect to the value of $x(N-1)$ that results from application of $u(N-2)$; therefore,

$$\begin{aligned}J^*_{N-2,N}\big(x(N-2)\big) = Min\{\tfrac{1}{2}[&x^T(N-2)Q(N-2)x(N-2)\\ &+ u^T(N-2)R(N-2)u(N-2) + J^*_{N-1,N}(x(N-1))]\}.\end{aligned} \qquad 2.13$$

Also, since $x(N-1)$ is related to $x(N-2) and\, u(N-2)$ by the state (2.1), then (2.13) depends only on $x(N-2)$; and

$$\begin{aligned}J^*_{N-2,N}\big(x(N-2)\big) = Min\{\tfrac{1}{2}[&x^T(N-2)Q(N-2)x(N-2)\\ &+ u^T(N-2)R(N-2)u(N-2)\\ &+ J^*_{N-1,N}\big([1+\Delta tA(N-2)]x(N-2) + \Delta tB(N-2)u(N-2)\big)]\}.\end{aligned} \qquad 2.14$$

By considering the cost of operation over the final three stages, a three–stage process with initial state $x(N-3)$, we can follow exactly the same reasoning which led to (2.14) to obtain

$$\begin{aligned}J^*_{N-3,N}\big(x(N-3)\big) = Min\{\tfrac{1}{2}[&x^T(N-3)Q(N-3)x(N-3)\\ &+ u^T(N-3)R(N-3)u(N-3)\\ &+ J^*_{N-2,N}\big([1+\Delta tA(N-3)]x(N-3) + \Delta tB(N-3)u(N-3)\big)]\}.\end{aligned} \qquad 2.15$$

Continuing backward in this manner, we obtain for a k-stage process the result

$$J^*_{N-k,N}\big(x(N-k)\big) = min_{u(N-k),\dots,u(N-1)}\left\{\begin{matrix}\tfrac{1}{2}x^T(N)Hx(N) + \\ \tfrac{1}{2}\Delta t\sum_{k=N-k}^{N-1}(x^T(k)Q(k)x(k) + u^T(k)R(k)u(k))\end{matrix}\right\}. \qquad 2.16$$

And on applying the optimality principle to (2.16), it gives

$$J^*_{N-k,N}\big(x(N-k)\big) = min_{u(N-k)}\,\tfrac{1}{2}\left\{\begin{matrix}x^T(N-k)Q(N-k)x(N-k) + u^T(N-k)R(N-k)u(N-k)\\ + J^*_{N-k,N}\big([1+\Delta tA(N-k)]x(N-k) + \Delta tB(N-k)u(N-k)\big)\end{matrix}\right\}. \qquad 2.17$$

The equation (2.17) is the recurrence relation that we set out to obtain.

With the knowledge of $J^*_{N-(k-1),N}$, the optimal cost for a *(k-1)*-stage policy, $J^*_{N-(k-1),N}$ can be generated which is the optimal cost for a *k*-stage policy. And to begin the process, one simply starts with a zero-stage process and generate $J^*_{N,N} \triangleq J_{N,N}$. Next, the optimal cost can be found for a one-stage process by using $J^*_{N,N}$ and (2.17) and so on in which, beginning with a zero- stage process corresponds to starting at the terminal state and starting at the final time in the control.

III. Computational Ingredients for the Discretized Linear Regulator Problems

In this section, we shall focus on the discrete system described by the state

$$x(k+1) = A(k)x(k) + B(k)u(k). \qquad 3.1$$

The states and controls are not constrained by any boundaries. The problem is to find an optimal policy $u^*(x(k), k)$ that minimizes the performance measure

$$J = \frac{1}{2}x^T(N)Hx(N) + \frac{1}{2}\sum_{k=0}^{N-1}(x^T(k)Q(k)x(k) + u^T(k)R(k)u(k)), \qquad 3.2$$

where $H\ and\ Q(k)$ are real symmetric positive semi definite $n \times n$ matrices. $R(k)$ is a real symmetric positive definite $m \times\ m$ matrix and N is a fixed integer greater than 0. (3.1) and (3.2) is the discrete counterpart of the continuous linear regulator problem considered in the section above. To simplify the notation in the derivation that follows, we assume that A, B, R, and Q are constant matrices. The approach we will take is to solve the functional equation (2.17).

To begin, [3] and [4] define

$$J_{N,N}\big(x(N)\big) = \frac{1}{2}x^T(N)Hx(N) = J_{N,N}^*\big(x(N)\big) \triangleq \frac{1}{2}x^T(N)P(0)x(N) \qquad 3.3$$

where $P(0) \triangleq H$. The cost over the final interval is given by

$$J_{N-1,N}\big(x(N-1), u(n-1)\big) = \frac{1}{2}[x^T(N-1)Qx(N-1) + u^T(N-1)Ru(N-1)$$
$$+x^T(N)P(0)x(N)] \qquad 3.4$$

And the minimum cost is given as

$$J_{N-k,N}^*\big(x(N-1)\big) = min_{u(N-1)}\{J_{N-1,N}\big(x(N-1), u(N-1)\big)\}. \qquad 3.5$$

Now, x(N) is related to $u(N-1)$ by the state equation, so

$$J_{N-1,N}^*\big(x(N-k)\big) = min_{u(N-1)}\{\tfrac{1}{2}[x^T(N-1)Qx(N-1) + u^T(N-1)Ru(N-1)]$$
$$+\frac{1}{2}[Ax(N-1) + Bu(N-1)]^T P(0)[Ax(N-1) + Bu(N-1)]. \qquad 3.6$$

Also by [1] and [3], it is assumed that, the admissible controls are not bounded: therefore, to minimize $J_{N-1,N}$ with respect to $u(N-1)$, we need to consider only those control values for which

$$\begin{bmatrix} \frac{\partial J_{N-1,N}}{\partial u_1(N-1)} \\ \frac{\partial J_{N-1,N}}{\partial u_2(N-1)} \\ \vdots \\ \frac{\partial J_{N-1,N}}{\partial u_m(N-1)} \end{bmatrix} \triangleq \frac{\partial J_{N-1,N}}{\partial u(N-1)} = 0. \qquad 3.7$$

On evaluating the indicated partial derivative gives

$$Ru(N-1) + B^T P(0)[Ax(N-1) + Bu(N-1)] = 0. \qquad 3.8$$

The control values that satisfy (3.8) may yield a minimum of $J_{N-1,N}$, a maximum, or neither. To further investigate this, we form the matrix of second partials given as:

$$\begin{bmatrix} \frac{\partial^2 J_{N-1,N}}{\partial u_1^2(N-1)} & \frac{\partial^2 J_{N-1,N}}{\partial u_2(N-1)\partial u_1(N-1)} & \cdots & \frac{\partial^2 J_{N-1,N}}{\partial u_1(N-1)\partial u_m(N-1)} \\ \frac{\partial^2 J_{N-1,N}}{\partial u_2(N-1)\partial u_1(N-1)} & \frac{\partial^2 J_{N-1,N}}{\partial u_2(N-1)} & \cdots & \frac{\partial^2 J_{N-1,N}}{\partial u_2(N-1)\partial u_m(N-1)} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 J_{N-1,N}}{\partial u_m(N-1)\partial u_1(N-1)} & \frac{\partial^2 J_{N-1,N}}{\partial u_m(N-1)\partial u_2(N-1)} & \cdots & \frac{\partial^2 J_{N-1,N}}{\partial u_m^2(N-1)} \end{bmatrix} \triangleq \frac{\partial^2 J_{N-1,N}}{\partial u^2(N-1)} = R + B^T P(0)B. \qquad 3.9$$

By assumption, H [and hence P(0)] is a positive semi definite matrix, and R is a positive definite matrix. It can be shown that since P(0) is positive semi-definite, so is $B^T P(0)B$. This means that $R + B^T P(0)B$ is the sum of a positive definite matrix and a positive semi-definite matrix, and this implies that, $R + B^T P(0)B$ is positive definite. Solving (3.8) for the optimal control gives

$$u^*(N-1) = -[R + B^T P(0)B]^{-1} B^T P(0)Ax(N-1) = F(N-1)x(N-1). \qquad 3.10$$

Since $R + B^T P(0)B$ is positive definite, the indicated inverse is guaranteed to exist. Substituting (3.10) into (3.4) gives $J_{N-1,N}^*$, which after the terms have been collected becomes

$$J_{N-1,N}^*\big(x(N-1)\big) = \frac{1}{2}x^T(N)-1)\{[A + BF(N-1)]^T P(0)[A + BF(N-1)]$$
$$+F^T(N-1)RF(N-1) + Q\}x(N-1)$$
$$= \frac{1}{2}x^T(N-1)P(1)x(N-1). \qquad 3.11$$

With

$$P(1) = [A + BF(N-1)]^T P(0)[A + BF(N-1)] + F^T(N-1)RF(N-1) + Q, \qquad 3.12$$

It is important to note that $J_{N-1,N}^*$ is exactly the same form of $J_{N,N}^*$, which means that, when we continue the process one stage further backward, the results will have will be exactly the same form; i.e.

$$u^*(N-2) = -[R + B^T P(1)B]^{-1} B^T P(0)Ax(N-2) = F(N-2)x(N-2). \qquad 3.13$$

And

$$J_{N-2,N}^*\big(x(N-2)\big) = \frac{1}{2}x^T(N)-2)\{[A + BF(N-2)]^T P(1)[A + BF(N-2)]$$
$$+F^T(N-2)RF(N-2) + Q\}x(N-2)$$
$$= \frac{1}{2}x^T(N-2)P(2)x(N-2). \qquad 3.14$$

By induction, for the *kth* stage of the general time-varying case, the same derivation gives

$$u^*(N-k) = -[R(N-k) + B^T(N-k)P(k-1)B(N-k)]^{-1} \times B^T(N-k)P(k-1)A(N-k)x(N-k) = F(N-k)x(N-k). \quad 3.15$$

$$J^*_{N-k,N}\big(x(N-k)\big) = \tfrac{1}{2}x^T(N-k)\{[A(N-1) + B(N-k)F(N-k)]^T P(k-1)[A(N-k) + B(N-k)F(N-k)] + F^T(N-k)R(N-k)F(N-k) + Q(N-k)\}x(N-k) = \tfrac{1}{2}x^T(N-k)P(k)x(N-k). \quad 3.16$$

Another important result of the derivation is that, the minimum cost for an *N*-stage process with initial state x_0 is given by

$$J^*_{0,N}(x_0) = \tfrac{1}{2}x_0^T P(N)x_0, \quad 3.17$$

which follows directly from the definition of $P(N-k)$. By implication, it means that, storage of the $P(N-k)$ matrices for $k = 1, 2, \dots N$ provides us with a means of determining the minimum costs for processes from 1 to N stages.

IV. Computational Procedure for Discrete Linear Regulator Problem

Step 1: Compute the optimal cost functional at the final stage, k = 0, with N specified using the relation:

$$J^*_{N,N}\big(x(N)\big) = \tfrac{1}{2}x^T(N)P(k)x(N) \quad 4.1a$$

where $P(k) = H$. 4.1b

Step 2: Compute the optimal cost law at k = 1, 2… Nth stages from the final stage using the relation:

$$u^*(N-k) = F(N-k)x(N-k) \quad 4.1c$$

where the feedback gains

$$F(N-k) = -[R(N-k) + B^T(N-k)P(k-1)B(N-k)]^{-1} \times B^T(N-k)P(k-1)A(N-k). \quad 4.1d$$

Step 3: Update the value of $P(k)\ at\ k = 1, 2, \dots, Nth$ stages from the final stage with

$$P(k) = V^T(N-k)P(k-1)V(N-k) + F^T(N-k)P(k-1)F(N-k) + Q(N-k) \quad 4.1e$$

where $V(N-k) = A(N-k) + B(N-k)F(N-k)$. 4.1f

Step 4: If $k = N$, then stop and update the state variable with

$$x(k+1) = A(k)x(k) + B(k)u(k) \quad 4.1g$$

using the initial state variable value $x(0) = x_0$ and work backward until $k = N-1$. Else, go back to Step 2 through 4 until $k = N$.

Remark: It must be noted that, the minimum cost for an *N*-stage process with initial state x_0 is given by

$$J^*_{0,N}(x_0) = \tfrac{1}{2}x_0^T P(N)x_0. \quad 4.1h$$

V. Computational Results

In testing the efficiency and robustness of Dynamic Programming method in solving LQRP, the following problems were considered:

Problem (P2): Determine the optimal control law that will cause the linear discrete system

$$x(k+1) = \begin{pmatrix} 0.9974 & 0.0539 \\ -0.1078 & 1.1591 \end{pmatrix} x(k) + \begin{pmatrix} 0.0013 \\ 0.0539 \end{pmatrix} u(k),$$

with $x(0) = \begin{pmatrix} 2 \\ 1 \end{pmatrix}\ and\ u(0) = 1$ to minimize the performance measure

$$J = \tfrac{1}{2}\textstyle\sum_{k=0}^{N-1}\{0.25x_1^2(k) + 0.05x_2^2(k) + 0.05u^2(k)\}.$$

Problem (P3): Determine the optimal control law that will cause the linear discrete system

$$x(k+1) = \begin{pmatrix} 0 & 1 \\ -1 & -2 \end{pmatrix} x(k) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u(k),$$

with $x(0) = \begin{pmatrix} 1 \\ -1 \end{pmatrix}\ and\ u(0) = 1$ to minimize the performance measure

$$J = \tfrac{1}{2}x^T(N)\begin{pmatrix} 20 & 0 \\ 0 & 0 \end{pmatrix} x(N) + \textstyle\sum_{k=0}^{N-1}\left\{\tfrac{1}{2}x_1^2(k) + x_2^2(k) + 0.5u^2(k)\right\}.$$

Problem (P4): The first order linear system

$$\dot{x}_1(t) = 2x_2(t)$$
$$\dot{x}_2(t) = -x_1(t) - 3x_2 + u(t)$$

is to be controlled to minimize the performance measure

$$J = \tfrac{1}{2}x^2(t_f)Hx(t_f) + \tfrac{1}{2}\int_0^{t_f}\{x^T(t)Qx(t) + Ru^2(t)\}dt,$$

with $x(0) = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, t_f = 1, R = 1, Q = \begin{pmatrix} 1 & 0 \\ 0 & 10 \end{pmatrix}\ and\ H = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$. The admissible state and control values are not constrained by any boundaries. Find the optimal control law using Dynamic Programming.

VI. Conclusion

In this paper, we applied the dynamic programming method to solving discretized Bolza's cost form of Linear Quadratic Regulator Problems (LQRP). It has been observed that, the optimal control at each stage is a linear combination of the states; therefore, the optimal policy is linear state variable feedback.

Also, the minimum cost for an N-stage process with initial state x_0 is given by $J_{0,N}^*(x_0) = \frac{1}{2}x_0^T P(N)x_0$, which follows directly from the definition of $P(N-k)$. This means that, storage of the $P(N-k)$ matrices for $k = 1, 2, \dots, N$ provides us with a means of determining the minimum costs for processes from $1\ to\ N$ stages. In each of the problems considered, we have taken N = 10 and the optimal cost law and the optimal cost functional at each stage $u^*(N-k)\ and\ J_{N-k,N}^*(x(N-k))$ respectively are reported in the tables 1 to 3.

References

[1] Boudarel, J., et al, (1971), Dynamic Programming and Its Application to Optimization Theory, Academic Press, Inc., 111 Fifth Avenue, New York, New York 10003.
[2] David, G. Hull, (2003), Optimal Control Theory for Applications, Mechanical Engineering Series, Springer-Verlag, New York, Inc., 175 Fifth Avenue, New York, NY 10010.
[3] Kirk, E. Donald, (2004), Optimal control theory: An introduction, Prentice-Hall, Inc., Englewood Cliff, New Jersey.
[4] George M. Siouris, (1996), An Engineering Approach To Optimal Control And Estimation Theory, John Wiley & Sons, Inc., 605 Third Avenue, New York, 10158- 00 12.

Table 1: Computational Results for Solution of Problem (P2)

k	$F_1(N-k)$	$F_2(N-k)$	$x_1(k)$	$x_2(k)$	$u^*(N-k)$	$J_{N-k,N}^*(x(N-k))$
0	0	0	2	1	0.089254831	2271.267178
1	-0.67072573e-3	-0.62643319e-1	2.021301552	-0.192481813	0.130476228	2238.320038
2	-0.69124457e-2	-0.14696912	1.997972009	-0.760230483	0.273906332	825.0076642
3	-0.1892190e-1	-0.43060243	1.955233517	-0.954241264	0.631634837	306.3492589
4	-0.29116282e-1	-1.23398267	1.90702315	-0.97242069	1.467881515	114.8623145
5	-0.1649150e-1	-3.05856165	1.856803457	-1.036175256	3.138584426	42.96820265
6	-0.14099478e-1	-5.68525413	1.800206082	-1.232024452	5.501570713	16.06857392
7	-0.286160828	-7.28263538	1.731027674	-1.542982945	6.389880003	6.020809844
8	-1.264155036	-6.79563955	1.644181347	-1.941031197	2.640505972	2.274632176
9	-3.45491054	-5.51121792	1.535640972	-2.099285678	-5.922606819	0.824886132
10	-7.81819049	-5.43936856	1.418666427	-2.591791457	-21.07572936	0.419510988

Table 2: Computational Results for Solution of Problem (P3)

k	$F_1(N-k)$	$F_2(N-k)$	$x_1(k)$	$x_2(k)$	$u^*(N-k)$	$J_{N-k,N}^*(x(N-k))$
0	0	0	1	-1	0.070179618	3.811009052e15
1	0.97674419	1.95348837	-1	0.67041752	-0.114657122	8.094809459e13
2	0.66899391	1.32295023	0.67041752	-0.439874698	0.092280721	1.903917291e11
3	0.67985926	1.17533354	-0.439874698	0.282059946	-0.052568104	550347617.9
4	0.87410267	1.32525671	0.282059946	-0.175327656	0.041902617	20795519.778
5	0.95490138	1.37591297	-0.175327656	0.668202884	0.75196846	16880064.79
6	0.97701187	1.36936918	0.668202884	-0.409109652	0.035487626	2728.666113
7	0.98581874	1.35628707	-0.409109652	0.191919037	-0.051082462	57.03464117
8	0.99034065	1.34405007	0.191919037	-0.027296526	0.07272807	29.17652412
9	0.99298524	1.33341629	-0.027296526	-0.045045264	-0.099039658	0.193181275
10	0.99466828	1.32450763	-0.045045264	0.002729932	-0.32958248	0.2089105783e-2

Table 3: Computational Results for Solution of Problem (P4)

k	$F_1(N-k)$	$F_2(N-k)$	$x_1(k)$	$x_2(k)$	$u^*(N-k)$	$J^*_{N-k,N}(x(N-k))$
0	0	0	2	-1	1	43271.260173
1	2.270670566	-0.270670566	1.135335283	-0.135335283	0.927067056	32538.125038
2	1.048171086	1.347305166	0.524085543	-0.253414473	0.783677455	928.0076642
3	1.019904468	0.170001406	0.509952234	0.0859997037	0.688309568	372.3190581
4	1.287519274	0.066426266	0.437596376	0.0332131339	0.612835984	201.83623214
5	1.293700636	-0.020839452	0.364685035	-0.010419726	0.685090339	74.96320165
6	1.314568066	0.4477956125e-2	0.165728403	0.223897806e-2	0.616133509	28.41065792
7	1.078367745	0.2783677485e-2	0.783677455e-1	0.178367745e-2	-1.583677455	3.2080401845
8	0.968309568	0.6883095681e-3	0.238130956e-2	0.968530956e-3	-2.690130956	1.254612176
9	0.723671455	0.1283677455e-3	0.478536707e-3	0.783677455e-3	-2.951712763	0.682981613
10	0.658830956	0.2685839568e-4	0.168530956e-3	0.703309568e-3	-3.742916948	0.261939258e-2