An Assessment of a Linear Quadratic Stochastic Control Algorithm

Ky M. Vu, PhD. AuLac Technologies Inc. ©2008
Email: kymvu@aulactechnologies.com

Abstract

The existing N finite steps optimal control algorithm of a discrete state space model, stochastic regulating control system is under review and compared with a new algorithm. The new algorithm is derived by the method of dynamic programming. The two algorithms give the same value for each controller of the N steps. The algorithms are physically implementable and must be used for applications with a small number of control steps. For a large or an infinite number of steps, a steady state controller, obtained by convergence of the algorithms, can be used. In this case, the controller has the physical meaning of a pseudo infinite steps optimal controller.

Keywords: Bellman equation, dynamic programming, linear quadratic control; state space model.

1 Introduction

Ever since the publication of the author's book (Vu, K. (2008)), the author received a good number of emails asking for a clarification of a control algorithm in the book. The algorithm is the N finite steps optimal control algorithm, the most difficult algorithm of linear quadratic stochastic control theory, of a discrete control system described by a state space model. The algorithm seems to be contradictory with an established algorithm in control literature. The established algorithm has appeared in good textbooks such as Åström, K.J. (1970), Åström, K.J. and Wittenmark, B. (1989) and Mosca, E. (1995), among others. The deterministic version of the algorithm is discussed in Kailath, T. (1980). In this paper, the established algorithm will be reviewed. Then the result of a new algorithm, suggested but not proved in Vu, K. (2008), is derived by the method of dynamic programming. The paper is organized as follows. Section one is the introduction section. In section two, the established algorithm is briefly described. In section three, the new algorithm is derived. In section four, the two algorithms are compared, and section five concludes the discussion of the paper.

2 The Existing N Steps Control Algorithm

For a simpler discussion, we consider the following discrete state space model of a stochastic control system

$$\begin{aligned} \mathbf{x}_{t+1} &= \mathbf{A}\mathbf{x}_t + \mathbf{b}u_t + \mathbf{w}_t, \\ y_t &= \mathbf{c}\mathbf{x}_t + v_t. \end{aligned}$$

The system is a single-input-single-output (SISO) system with no pure dead time. The inclusion of a pure dead time creates no difficulty for our discussion. However, since the control model in most textbooks that discuss the algorithm has no pure dead time, we also use a model with no pure dead time, to create less confusion. The control criterion for the N steps optimal control algorithm, in Åström, K.J. (1970), is

$$\underset{u_t,\cdots u_N}{Min} \quad E\{\mathbf{x}_{N+1}^T\mathbf{Q}_0\mathbf{x}_{N+1} + \sum_{s=t}^{N}\mathbf{x}_s^T\mathbf{Q}_1\mathbf{x}_s + u_s^T\mathbf{Q}_2u_s\}.$$

The control problem is understood as follows. The control algorithm is a stochastic control algorithm because of the white noises $\mathbf{w}_t$ and v_t, hence the expectation operator E. Even though it is not written in the reference, as it should, it is understood that the expectation is a conditional expectation. It is not an unconditional expectation, because an unconditional expectation will give us the variances of the state vector and the control input variable. The control algorithm will become the infinite steps optimal control algorithm because it has an infinite number of control input variable values as its solution.

The solution for the control problem, given in Åström, K.J. (1970), can be summarized as

$$\begin{aligned} u_t &= -\mathbf{L}_t\hat{\mathbf{x}}_{t|t}, \\ \mathbf{L}_t &= [\mathbf{Q}_2 + \mathbf{b}^T\mathbf{S}_{t+1}\mathbf{b}]^{-1}\mathbf{b}^T\mathbf{S}_{t+1}\mathbf{A}, \\ \mathbf{S}_t &= \mathbf{A}^T\mathbf{S}_{t+1}\mathbf{A} + \mathbf{Q}_1 - \frac{\mathbf{A}^T\mathbf{S}_{t+1}\mathbf{b}\mathbf{b}^T\mathbf{S}_{t+1}\mathbf{A}}{\mathbf{Q}_2 + \mathbf{b}^T\mathbf{S}_{t+1}\mathbf{b}} \quad (1) \\ \mathbf{S}_{N+1} &= \mathbf{Q}_0. \end{aligned}$$

3 The New N Steps Control Algorithm

For our discussion, we consider the control criterion or performance index

$$\underset{u_t}{Min} \quad E\{\sum_{t=k}^{N}\mathbf{x}_{t+1}^T\mathbf{Q}_1\mathbf{x}_{t+1} + \mathbf{Q}_2u_t^2 \mid \mathcal{Y}_t\}.$$

In the above equation, the condition $\mathcal{Y}_t$ is the condition of all available observations of the output variable y_t up to the time t. The performance index is a quadratic sum; therefore, its solution can be obtained by dynamic programming. Dynamic programming, pioneered by the American mathematician Richard Bellman (Bellman, R. (1997)), is a multistage optimization approach, ideal for a quadratic sum.

Its philosophy is similar to that of proving a theorem by induction. This means that we choose a starting point or an initial condition and prove that optimality is obtained for this case. Then, we obtain optimality for two recurrent stages. The result will be optimality for the whole sum.

Depending on the problem, the starting optimization in a dynamic programming solution can be at either end of the sum. For our problem, it must be at the high end (k=N) of the sum, because the solution at the lower end affects the variables at the higher end. This means that we must obtain optimality for the case k=N first. This is to say that we must solve the following control problem

$$\underset{u_N}{Min} \quad E\{\mathbf{x}_{N+1}^T\mathbf{Q}_1\mathbf{x}_{N+1} + \mathbf{Q}_2 u_N^2 \mid \mathcal{Y}_N\}.$$

By changing the state vector

$$\mathbf{x}_{N+1} \quad = \quad \mathbf{A}\mathbf{x}_N + \mathbf{b}u_N + \mathbf{w}_N,$$

we can write the previous equation as

$$\begin{aligned}\underset{u_N}{Min} \quad & E\{\mathbf{x}_N^T\mathbf{A}^T\mathbf{Q}_1\mathbf{A}\mathbf{x}_N + 2\mathbf{x}_N^T\mathbf{A}^T\mathbf{Q}_1\mathbf{b}u_N + 2\mathbf{x}_N^T\mathbf{A}^T\mathbf{Q}_1\mathbf{w}_N \\ & +2\mathbf{b}^T\mathbf{Q}_1\mathbf{w}_N u_N + \mathbf{w}_N^T\mathbf{Q}_1\mathbf{w}_N + [\mathbf{Q}_2 + \mathbf{b}^T\mathbf{Q}_1\mathbf{b}]u_N^2 \mid \mathcal{Y}_N\}.\end{aligned}$$

By taking the conditional expectation with the condition $\mathcal{Y}_N$, we have the state vector $\mathbf{x}_N$ as $\hat{\mathbf{x}}_{N|N}$. The products of $\mathbf{w}_N$ with other variables will be zero; the product of this white noise with itself gives its variance, $\mathbf{R}_w$, because $\mathcal{Y}_N$ contains up to $\mathbf{w}_{N-1}$. u_N will be unchanged. The first term can be obtained as follows. We write

$$\begin{aligned}\mathbf{x}_N^T\mathbf{A}^T\mathbf{Q}_1\mathbf{A}\mathbf{x}_N = & [\mathbf{x}_N - \hat{\mathbf{x}}_{N|N}]^T\mathbf{A}^T\mathbf{Q}_1\mathbf{A}[\mathbf{x}_N - \hat{\mathbf{x}}_{N|N}] + \\ & 2\mathbf{x}_N^T\mathbf{A}^T\mathbf{Q}_1\mathbf{A}\hat{\mathbf{x}}_{N|N} - \hat{\mathbf{x}}_{N|N}^T\mathbf{A}^T\mathbf{Q}_1\mathbf{A}\hat{\mathbf{x}}_{N|N}\end{aligned}$$

then take the conditional expectation. The variable inside the square brackets is the error vector of the state estimator $\hat{\mathbf{x}}_{N|N}$. The first term on the right hand side of the last equation can be obtained as

$$\begin{aligned}& E\{[\mathbf{x}_N - \hat{\mathbf{x}}_{N|N}]^T\mathbf{A}^T\mathbf{Q}_1\mathbf{A}[\mathbf{x}_N - \hat{\mathbf{x}}_{N|N}] \mid \mathcal{Y}_N\} \\ & \quad = E\{[\mathbf{x}_N - \hat{\mathbf{x}}_{N|N}]^T\mathbf{A}^T\mathbf{Q}_1\mathbf{A}[\mathbf{x}_N - \hat{\mathbf{x}}_{N|N}]\} \\ & \quad = E\{tr([\mathbf{x}_N - \hat{\mathbf{x}}_{N|N}]^T\mathbf{A}^T\mathbf{Q}_1\mathbf{A}[\mathbf{x}_N - \hat{\mathbf{x}}_{N|N}])\} \\ & \quad = E\{tr(\mathbf{A}^T\mathbf{Q}_1\mathbf{A}[\mathbf{x}_N - \hat{\mathbf{x}}_{N|N}][\mathbf{x}_N - \hat{\mathbf{x}}_{N|N}]^T)\} \\ & \quad = tr(\mathbf{A}^T\mathbf{Q}_1\mathbf{A}E\{[\mathbf{x}_N - \hat{\mathbf{x}}_{N|N}][\mathbf{x}_N - \hat{\mathbf{x}}_{N|N}]^T\}) \\ & \quad = tr(\mathbf{A}^T\mathbf{Q}_1\mathbf{A}\mathbf{P}_0).\end{aligned}$$

In the above equations, the condition disappears because the white noises $\mathbf{w}_t$ and v_t in the condition are uncorrelated to the error vector; the conditional expectation becomes the unconditional expectation. Therefore, we can write

$$\begin{aligned}E\{\mathbf{x}_N^T\mathbf{A}^T\mathbf{Q}_1\mathbf{A}\mathbf{x}_N \mid \mathcal{Y}_N\} \quad = \quad & \hat{\mathbf{x}}_{N|N}^T\mathbf{A}^T\mathbf{Q}_1\mathbf{A}\hat{\mathbf{x}}_{N|N} \\ & +tr(\mathbf{A}^T\mathbf{Q}_1\mathbf{A}\mathbf{P}_0).\end{aligned}$$

And we have the performance index as

$$\begin{aligned}\underset{u_N}{Min} \quad & \hat{\mathbf{x}}_{N|N}^T\mathbf{A}^T\mathbf{Q}_1\mathbf{A}\hat{\mathbf{x}}_{N|N} + 2\hat{\mathbf{x}}_{N|N}^T\mathbf{A}^T\mathbf{Q}_1\mathbf{b}u_N \\ & +[\mathbf{Q}_2 + \mathbf{b}^T\mathbf{Q}_1\mathbf{b}]u_N^2 + tr(\mathbf{Q}_1\mathbf{R}_w) + tr(\mathbf{A}^T\mathbf{Q}_1\mathbf{A}\mathbf{P}_0).\end{aligned}$$

The above equation is a quadratic function in the variable u_N. Since the quadratic coefficient is positive, the function has a minimum. And its minimum occurs at

$$u_N \quad = \quad -[\mathbf{Q}_2 + \mathbf{b}^T\mathbf{Q}_1\mathbf{b}]^{-1}\mathbf{b}^T\mathbf{Q}_1\mathbf{A}\hat{\mathbf{x}}_{N|N}. \qquad (2)$$

The above equation is the equation of the one step optimal or myopic controller. The minimum value is

$$\underset{u_N}{Min} \quad E\{\mathbf{x}_{N+1}^T\mathbf{Q}_1\mathbf{x}_{N+1} + \mathbf{Q}_2 u_N^2 \mid \mathcal{Y}_N\} = \hat{\mathbf{x}}_{N|N}^T\mathbf{S}_N\hat{\mathbf{x}}_{N|N} + q_N$$

with

$$\begin{aligned}\mathbf{S}_N \quad & = \quad \mathbf{A}^T\mathbf{Q}_1\mathbf{A} - \frac{\mathbf{A}^T\mathbf{Q}_1\mathbf{b}\mathbf{b}^T\mathbf{Q}_1\mathbf{A}}{\mathbf{Q}_2 + \mathbf{b}^T\mathbf{Q}_1\mathbf{b}}, \qquad (3) \\ q_N \quad & = \quad tr(\mathbf{Q}_1\mathbf{R}_w) + tr(\mathbf{A}^T\mathbf{Q}_1\mathbf{A}\mathbf{P}_0).\end{aligned}$$

The last result gives us a starting point. Now, like the approach of proving a theorem by induction, we assume that we have optimality for the stage k with the minimum value given by the last equation with the letter k in place of the letter N, then we set up an optimization problem with two consecutive stages as

$$\begin{aligned}& \underset{u_{k-1},\, u_k}{Min} \quad E\{\sum_{t=k-1}^{k} \mathbf{x}_{t+1}^T\mathbf{Q}_1\mathbf{x}_{t+1} + \mathbf{Q}_2 u_t^2 \mid \mathcal{Y}_t\} = \\ & \underset{u_{k-1}}{Min} \left(\underset{u_k}{Min} \quad E\{\sum_{t=k-1}^{k} \mathbf{x}_{t+1}^T\mathbf{Q}_1\mathbf{x}_{t+1} + \mathbf{Q}_2 u_t^2 \mid \mathcal{Y}_t\} \right) = \\ & \underset{u_{k-1}}{Min} \quad E\{\mathbf{x}_k^T\mathbf{Q}_1\mathbf{x}_k + \mathbf{Q}_2 u_{k-1}^2 + \hat{\mathbf{x}}_{k|k}^T\mathbf{S}_k\hat{\mathbf{x}}_{k|k} + q_k \mid \mathcal{Y}_{k-1}\}.\end{aligned}$$

The last equation is called the *Bellman* equation. At this stage, we want to look for an optimal value of u_{k-1}. The state estimator $\hat{\mathbf{x}}_{k|k}$ contains this variable as shown below

$$\hat{\mathbf{x}}_{k|k} \quad = \quad \mathbf{A}\hat{\mathbf{x}}_{k-1|k-1} + \mathbf{b}u_{k-1} + \mathbf{k}^* a_k, \qquad (4)$$

and this fact gives an interconnection between two consecutive optimization stages in the Bellman equation.

Now, from previous result we can say that

$$\begin{aligned}& E\{\mathbf{x}_k^T\mathbf{Q}_1\mathbf{x}_k + \mathbf{Q}_2 u_{k-1}^2 \mid \mathcal{Y}_{k-1}\} = \\ & \hat{\mathbf{x}}_{k-1|k-1}^T\mathbf{A}^T\mathbf{Q}_1\mathbf{A}\hat{\mathbf{x}}_{k-1|k-1} + 2\hat{\mathbf{x}}_{k-1|k-1}^T\mathbf{A}^T\mathbf{Q}_1\mathbf{b}u_{k-1} \\ & \quad +[\mathbf{Q}_2 + \mathbf{b}^T\mathbf{Q}_1\mathbf{b}]u_{k-1}^2 + tr(\mathbf{Q}_1\mathbf{R}_w) + tr(\mathbf{A}^T\mathbf{Q}_1\mathbf{A}\mathbf{P}_0).\end{aligned}$$

And we have

$$\begin{aligned}& E\{\hat{\mathbf{x}}_{k|k}^T\mathbf{S}_k\hat{\mathbf{x}}_{k|k} + q_k \mid \mathcal{Y}_{k-1}\} \\ & = E\{\left[\mathbf{A}\hat{\mathbf{x}}_{k-1|k-1} + \mathbf{b}u_{k-1} + \mathbf{k}^* a_k\right]^T \mathbf{S}_k \\ & \quad \left[\mathbf{A}\hat{\mathbf{x}}_{k-1|k-1} + \mathbf{b}u_{k-1} + \mathbf{k}^* a_k\right] + q_k \mid \mathcal{Y}_{k-1}\} \\ & = \hat{\mathbf{x}}_{k-1|k-1}^T\mathbf{A}^T\mathbf{S}_k\mathbf{A}\hat{\mathbf{x}}_{k-1|k-1} + \mathbf{b}^T\mathbf{S}_k\mathbf{b}u_{k-1}^2 + \\ & \quad {\mathbf{k}^*}^T\mathbf{S}_k\mathbf{k}^*\sigma_a^2 + 2\hat{\mathbf{x}}_{k-1|k-1}^T\mathbf{A}^T\mathbf{S}_k\mathbf{b}u_{k-1} + q_k.\end{aligned}$$

Except for the term with the squared white noise a_k, the cross-products of this variable with other variables vanish with the conditional expectation because the condition does not contain this white noise, which does not cross-correlate to $\hat{\mathbf{x}}_{k-1|k-1}$ or u_{k-1}. The squared white noise becomes the unconditional variance σ_a^2. The matrix $\mathbf{k}^*$ is the Kalman filter matrix for the conditional simultaneous state estimator. Eq (4) can be obtained in Vu, K. (2008) or any standard textbook on Kalman filtering.

From the above results, we can write

$$\begin{aligned}
&\underset{u_{k-1}}{Min} \quad E\{\mathbf{x}_k^T\mathbf{Q}_1\mathbf{x}_k + \mathbf{Q}_2u_{k-1}^2 + \hat{\mathbf{x}}_{k|k}^T\mathbf{S}_k\hat{\mathbf{x}}_{k|k} + q_k \mid \mathcal{Y}_{k-1}\} \\
&= \underset{u_{k-1}}{Min} \quad \hat{\mathbf{x}}_{k-1|k-1}^T\mathbf{A}^T(\mathbf{S}_k+\mathbf{Q}_1)\mathbf{A}\hat{\mathbf{x}}_{k-1|k-1} + \\
&2\hat{\mathbf{x}}_{k-1|k-1}^T\mathbf{A}^T(\mathbf{S}_k+\mathbf{Q}_1)\mathbf{b}u_{k-1} + [\mathbf{Q}_2 + \mathbf{b}^T(\mathbf{S}_k+\mathbf{Q}_1)\mathbf{b}]u_{k-1}^2 \\
&\quad +q_k + \mathbf{k}^{*^T}\mathbf{S}_k\mathbf{k}^*\sigma_a^2 + tr(\mathbf{Q}_1\mathbf{R}_w) + tr(\mathbf{A}^T\mathbf{Q}_1\mathbf{A}\mathbf{P}_0).
\end{aligned}$$

The optimal value for u_{k-1} will then be

$$u_{k-1} = -[\mathbf{Q}_2 + \mathbf{b}^T(\mathbf{S}_k+\mathbf{Q}_1)\mathbf{b}]^{-1}\mathbf{b}^T(\mathbf{S}_k+\mathbf{Q}_1)\mathbf{A}\hat{\mathbf{x}}_{k-1|k-1}.$$

By putting this optimal value into the criterion equation, we get

$$\begin{aligned}
&\underset{u_t}{Min} \; E\{\sum_{t=k-1}^{k}\mathbf{x}_{t+1}^T\mathbf{Q}_1\mathbf{x}_{t+1} + \mathbf{Q}_2u_t^2 \mid \mathcal{Y}_t\} \\
&\quad = \hat{\mathbf{x}}_{k-1|k-1}^T\mathbf{S}_{k-1}\hat{\mathbf{x}}_{k-1|k-1} + q_{k-1}.
\end{aligned}$$

with

$$\mathbf{S}_{k-1} = \mathbf{A}^T(\mathbf{S}_k+\mathbf{Q}_1)\mathbf{A} - \frac{\mathbf{A}^T(\mathbf{S}_k+\mathbf{Q}_1)\mathbf{b}\mathbf{b}^T(\mathbf{S}_k+\mathbf{Q}_1)\mathbf{A}}{\mathbf{Q}_2 + \mathbf{b}^T(\mathbf{S}_k+\mathbf{Q}_1)\mathbf{b}},$$

$$q_{k-1} = q_k + \mathbf{k}^{*^T}\mathbf{S}_k\mathbf{k}^*\sigma_a^2 + tr(\mathbf{Q}_1\mathbf{R}_w) + tr(\mathbf{A}^T\mathbf{Q}_1\mathbf{A}\mathbf{P}_0).$$

The minimum value of a two-stage sum has a similar form to that of a one-stage sum. Now, if we let k=N, then we have optimality for the sum of two stages N-1 and N as

$$\begin{aligned}
&\underset{u_t}{Min} \; E\{\sum_{t=N-1}^{N}\mathbf{x}_{t+1}^T\mathbf{Q}_1\mathbf{x}_{t+1} + \mathbf{Q}_2u_t^2 \mid \mathcal{Y}_t\} \\
&\quad = \hat{\mathbf{x}}_{N-1|N-1}^T\mathbf{S}_{N-1}\hat{\mathbf{x}}_{N-1|N-1} + q_{N-1}
\end{aligned}$$

with the optimal control actions:

$$\begin{aligned}
u_N &= -[\mathbf{Q}_2 + \mathbf{b}^T\mathbf{Q}_1\mathbf{b}]^{-1}\mathbf{b}^T\mathbf{Q}_1\mathbf{A}\hat{\mathbf{x}}_{N|N}, \\
u_{N-1} &= -[\mathbf{Q}_2 + \mathbf{b}^T(\mathbf{S}_N+\mathbf{Q}_1)\mathbf{b}]^{-1} \\
&\qquad \mathbf{b}^T(\mathbf{S}_N+\mathbf{Q}_1)\mathbf{A}\hat{\mathbf{x}}_{N-1|N-1}.
\end{aligned}$$

Similarly, we can determine the optimal value for u_{N-2} of the sum of the last three stages from the following equation

$$\begin{aligned}
&\underset{u_{N-2}}{Min} \quad E\{\mathbf{x}_{N-1}^T\mathbf{Q}_1\mathbf{x}_{N-1} + \mathbf{Q}_2u_{N-2}^2 + \\
&\qquad \hat{\mathbf{x}}_{N-1|N-1}^T\mathbf{S}_{N-1}\hat{\mathbf{x}}_{N-1|N-1} + q_{N-1} \mid \mathcal{Y}_{N-2}\}.
\end{aligned}$$

And the optimal control action u_{N-2} can be obtained in the same way as the optimal control action u_{k-1} discussed above. By reasoning in the same fashion, we can expand the range of optimality of the sum to more stages and to eventually the original number of stages in the sum of the performance index. This means that we can prove, for any number of stages, the solution of our N finite steps optimal control algorithm is as follows. The performance index is

$$\underset{u_t}{Min} \; E\{\sum_{t=k}^{N}\mathbf{x}_{t+1}^T\mathbf{Q}_1\mathbf{x}_{t+1} + \mathbf{Q}_2u_t^2 \mid \mathcal{Y}_t\} = \hat{\mathbf{x}}_{k|k}^T\mathbf{S}_k\hat{\mathbf{x}}_{k|k} + q_k$$

with

$$\mathbf{S}_t = \mathbf{A}^T(\mathbf{S}_{t+1}+\mathbf{Q}_1)\mathbf{A} - \frac{\mathbf{A}^T(\mathbf{S}_{t+1}+\mathbf{Q}_1)\mathbf{b}\mathbf{b}^T(\mathbf{S}_{t+1}+\mathbf{Q}_1)\mathbf{A}}{\mathbf{Q}_2 + \mathbf{b}^T(\mathbf{S}_{t+1}+\mathbf{Q}_1)\mathbf{b}},$$

$$q_t = q_{t+1} + \mathbf{k}^{*^T}\mathbf{S}_{t+1}\mathbf{k}^*\sigma_a^2 + tr(\mathbf{Q}_1\mathbf{R}_w) + tr(\mathbf{A}^T\mathbf{Q}_1\mathbf{A}\mathbf{P}_0).$$

and the initial conditions

$$\begin{aligned}
\mathbf{S}_N &= \mathbf{A}^T\mathbf{Q}_1\mathbf{A} - \frac{\mathbf{A}^T\mathbf{Q}_1\mathbf{b}\mathbf{b}^T\mathbf{Q}_1\mathbf{A}}{\mathbf{Q}_2 + \mathbf{b}^T\mathbf{Q}_1\mathbf{b}}, \\
q_N &= tr(\mathbf{Q}_1\mathbf{R}_w) + tr(\mathbf{A}^T\mathbf{Q}_1\mathbf{A}\mathbf{P}_0).
\end{aligned}$$

The control actions to achieve this result are given as

$$\begin{aligned}
u_N &= -[\mathbf{Q}_2 + \mathbf{b}^T\mathbf{Q}_1\mathbf{b}]^{-1}\mathbf{b}^T\mathbf{Q}_1\mathbf{A}\hat{\mathbf{x}}_{N|N}, \\
u_t &= -[\mathbf{Q}_2 + \mathbf{b}^T(\mathbf{S}_{t+1}+\mathbf{Q}_1)\mathbf{b}]^{-1}\mathbf{b}^T(\mathbf{S}_{t+1}+\mathbf{Q}_1)\mathbf{A}\hat{\mathbf{x}}_{t|t}, \\
&= -\mathbf{L}_t\hat{\mathbf{x}}_{t|t}.
\end{aligned}$$

4 Comparison and Discussion

With a statistical misconception and many unclear mathematical maneuvers in Åström, K.J. (1970), the author of this paper was about to claim that the existing algorithm is flawed. It turns out that the important parts of the algorithm are saved from this negative claim. If we add the matrix $\mathbf{Q}_1$ to both sides of the iterative equation for $\mathbf{S}_t$ of the new algorithm, we get

$$\begin{aligned}
\mathbf{S}_t + \mathbf{Q}_1 &= \mathbf{A}^T(\mathbf{S}_{t+1}+\mathbf{Q}_1)\mathbf{A} + \mathbf{Q}_1 - \\
&\qquad \frac{\mathbf{A}^T(\mathbf{S}_{t+1}+\mathbf{Q}_1)\mathbf{b}\mathbf{b}^T(\mathbf{S}_{t+1}+\mathbf{Q}_1)\mathbf{A}}{\mathbf{Q}_2 + \mathbf{b}^T(\mathbf{S}_{t+1}+\mathbf{Q}_1)\mathbf{b}}.
\end{aligned}$$

Now, if we compare the matrix given by left hand side of the above equation with the matrix $\mathbf{S}_t$ in Eq. (1) and the equation for the controller $\mathbf{L}_t$ of the existing algorithm, we can say that the Riccati equation, Eq. (1), of the existing algorithm is the same as that of the new algorithm. Now let us see if the two algorithms give the same controllers of the N steps.

For a valid comparison, let us say that the weighting matrix for the state vector $\mathbf{x}_{N+1}$ is $\mathbf{Q}_0$, then the existing algorithm gives the control action at step N, the last step, as

$$u_N = -[\mathbf{Q}_2 + \mathbf{b}^T\mathbf{Q}_0\mathbf{b}]^{-1}\mathbf{b}^T\mathbf{Q}_0\mathbf{A}\hat{\mathbf{x}}_{N|N}.$$

The new algorithm also gives the same value if we replace the matrix $\mathbf{Q}_1$ in Eq. (2) with the matrix $\mathbf{Q}_0$. For the controller at step $N-1$, the matrix $\mathbf{S}_{t+1}$ in the controller of the existing algorithm is

$$\mathbf{S}_N = \mathbf{A}^T\mathbf{Q}_0\mathbf{A} + \mathbf{Q}_1 - \frac{\mathbf{A}^T\mathbf{Q}_0\mathbf{b}\mathbf{b}^T\mathbf{Q}_0\mathbf{A}}{\mathbf{Q}_2 + \mathbf{b}^T\mathbf{Q}_0\mathbf{b}}.$$

We can see that the new algorithm also gives the same controller if we replace all the matrices $\mathbf{Q}_1$'s in Eq. (3) with the matrix $\mathbf{Q}_0$ and add to both sides of the equation the matrix $\mathbf{Q}_1$. Since the two algorithms give the same values for u_N and u_{N-1}, they will give the same values for all the control actions. We can, therefore, say that the two algorithms are the same.

The new algorithm is, however, much easy to understand. It has a crystal clear derivation. The statistics of the derivation is correct and a user will have very little difficulty in the modification of the algorithm for the case of control systems with a dead time. The equation of the controllers is also a strength. It separates the control action into three distinctive components as can be shown below

$$[\mathbf{Q}_2 + \mathbf{b}^T(\mathbf{S}_{t+1} + \mathbf{Q}_1)\mathbf{b}]u_t = -\mathbf{b}^T(\mathbf{S}_{t+1} + \mathbf{Q}_1)\mathbf{A}\hat{\mathbf{x}}_{t|t}.$$

The component associated with $\mathbf{Q}_2$ is due to the penalty on the movement of the control action. The component associated with the matrix $\mathbf{Q}_1$ gives different weights to the state vector. The component associated with the matrix $\mathbf{S}_{t+1}$ accounts for the optimal effect on the future state vectors. The effect must be different for each control action because of the difference in the time to the end of the control period. At the last step N the component vanishes, $\mathbf{S}_{N+1} = \mathbf{0}$, because there is no more future state vectors to account for.

To verify the value of the performance index, we consider the special case

$$\mathbf{Q}_1 = \mathbf{c}^T\mathbf{c}, \;\; \mathbf{Q}_2 = 0.$$

Assume that we have a total of N steps. Now by adding the quantity $N\sigma_v^2$ to the performance index, we have

$$\underset{u_t}{Min} \;\; E\{\sum_{t=1}^{N} \mathbf{x}_{t+1}^T\mathbf{c}^T\mathbf{c}\mathbf{x}_{t+1} \mid \mathcal{Y}_t\} + N\sigma_v^2.$$

This is the case of minimum variance control of the output variable because we have

$$\begin{aligned} y_{t+1} &= \mathbf{c}\mathbf{x}_{t+1} + v_{t+1}, \\ E\{y_{t+1}^2 \mid \mathcal{Y}_t\} &= E\{\mathbf{x}_{t+1}^T\mathbf{c}^T\mathbf{c}\mathbf{x}_{t+1} \mid \mathcal{Y}_t\} + \sigma_v^2. \end{aligned}$$

In the last equation, we have a conditional expectation but unconditional variance σ_v^2 for the white noise v_t, because the condition does not contain v_{t+1} but only v_t. Since v_t is white, it is not correlated to v_{t+1}. The variable in the condition is not correlated to the variable in the stochastic quantity, the conditional expectation becomes the unconditional expectation.

In this case, we have

$$\begin{aligned} \mathbf{S}_N &= \mathbf{A}^T\mathbf{c}^T\mathbf{c}\mathbf{A} - \frac{\mathbf{A}^T\mathbf{c}^T\mathbf{c}\mathbf{b}\mathbf{b}^T\mathbf{c}^T\mathbf{c}\mathbf{A}}{\mathbf{b}^T\mathbf{c}^T\mathbf{c}\mathbf{b}}, \\ &= \mathbf{A}^T\left[\mathbf{c}^T\mathbf{c} - \frac{\mathbf{c}^T\mathbf{c}\mathbf{b}\mathbf{b}^T\mathbf{c}^T\mathbf{c}}{\mathbf{b}^T\mathbf{c}^T\mathbf{c}\mathbf{b}}\right]\mathbf{A}, \\ &= \mathbf{0}, \\ \mathbf{S}_t &= \mathbf{A}^T(\mathbf{0} + \mathbf{c}^T\mathbf{c})\mathbf{A} - \frac{\mathbf{A}^T(\mathbf{0} + \mathbf{c}^T\mathbf{c})\mathbf{b}\mathbf{b}^T(\mathbf{0} + \mathbf{c}^T\mathbf{c})\mathbf{A}}{\mathbf{b}^T(\mathbf{0} + \mathbf{c}^T\mathbf{c})\mathbf{b}}, \\ &= \mathbf{0}. \end{aligned}$$

All the controllers of the N steps are the same. The N steps optimal control algorithm is the same as the one step optimal control algorithm, a minimum variance control one. Since the control system has no dead time, the performance index value will be $N\sigma_a^2$, a multiple of the variance of the innovations white noise in the Kalman filter literature. This means that we can write

$$\begin{aligned} N\sigma_a^2 &= \hat{\mathbf{x}}_{1|1}^T\mathbf{S}_1\hat{\mathbf{x}}_{1|1} + q_1 + N\sigma_v^2, \\ &= 0 + q_1 + N\sigma_v^2, \\ &= \sum_{k=1}^{N-1} \mathbf{k}^{*T}\mathbf{S}_{k+1}\mathbf{k}^*\sigma_a^2 + N\sigma_v^2 + \\ &\qquad Ntr(\mathbf{c}^T\mathbf{c}\mathbf{R}_w) + Ntr(\mathbf{A}^T\mathbf{c}^T\mathbf{c}\mathbf{A}\mathbf{P}_0), \\ &= N\sigma_v^2 + Ntr(\mathbf{c}^T\mathbf{c}\mathbf{R}_w) + Ntr(\mathbf{A}^T\mathbf{c}^T\mathbf{c}\mathbf{A}\mathbf{P}_0). \end{aligned}$$

From the Kalman filter theory (Vu, K. (2007)), we have

$$\begin{aligned} \sigma_a^2 &= \mathbf{c}\mathbf{P}_1\mathbf{c}^T + \sigma_v^2, \\ &= \mathbf{c}[\mathbf{A}\mathbf{P}_0\mathbf{A}^T + \mathbf{R}_w]\mathbf{c}^T + \sigma_v^2, \\ &= \sigma_v^2 + tr(\mathbf{c}^T\mathbf{c}\mathbf{R}_w) + tr(\mathbf{A}^T\mathbf{c}^T\mathbf{c}\mathbf{A}\mathbf{P}_0). \end{aligned}$$

The matrix $\mathbf{P}_1$ is the variance matrix of the error vector of the one step ahead state estimator. From the above equations, we can say that the performance index value given in this paper is correct, for this special case.

Now, let us evaluate the performance index value given in Åström, K.J. (1970). This reference divided the problem into two cases: complete state information where $\mathbf{c} = \mathbf{I}$ and $v_t = 0$ and the general case called incomplete state information. It was the second case that the author of the reference did not give the correct information. On page 259, the case implied that *"the control signal at time t is a function of all observed outputs up to time t"*. This is supposed to be so and the reason for the state estimator $\hat{\mathbf{x}}_{t|t}$ in the controller. But on page 281, it considered *"$u(t)$ is a function of $\mathcal{Y}_{t-1}$"*; and on page 282, this fact was confirmed with the following equation

$$\hat{\mathbf{x}}(t) = E\left[\mathbf{x}(t)|\mathcal{Y}_{t-1}\right],$$

which is equation (6.20) in the reference. The controller for this case, given as equation (6.24) in the reference, is

$$u(t) = -\mathbf{L}_t\hat{\mathbf{x}}(t) = -\mathbf{L}_t E\left[\mathbf{x}(t)|\mathcal{Y}_{t-1}\right].$$

The author of reference Åström, K.J. (1970) did mention the case "$u(t)$ *is a function of* $\mathcal{Y}_t$" again on page 283 but did not give any information about it. In Åström, K.J. and Wittenmark, B. (1989), on page 196, the same author had the Riccati and controller equations correct with the right state estimator. However, the performance index equation and the equation for the state estimator $\hat{\mathbf{x}}_{t|t}$ seemed to be dubious. And no value for the performance index was given. Since this information was not given, we cannot fairly give an assessment of this case.

We can, however, give an assessment of the case of complete information. For this case, reference Åström, K.J. (1970) gives, on page 281, the following equation for the performance index value

$$Min\ E\left[\mathbf{x}^T(N)\mathbf{Q}_0\mathbf{x}(N)+\sum_{t=t_0}^{N-1}\mathbf{x}^T(t)\mathbf{Q}_1\mathbf{x}(t)+u(t)^T\mathbf{Q}_2u(t)\right]$$

$$= m^TS(t_0)m + trS(t_0)\mathbf{R}_0 + \sum_{t=t_0}^{N-1} trS(t+1)\mathbf{R}_1(t).$$

The matrix $\mathbf{R}_1(t)$ is the same as the matrix $\mathbf{R}_w$ and m and $\mathbf{R}_0$ are the mean and variance of the state vector $\mathbf{x}(t_0)$. Obviously, we can see that this performance index value cannot be true because the variance matrix $\mathbf{P}_0$ of the error vector of the state estimator $\hat{\mathbf{x}}_{t|t}$ is not present in the performance index value.

For the case of control systems with a dead time, the model of the control system is

$$\begin{aligned} \mathbf{x}_{t+1} &= \mathbf{A}\mathbf{x}_t + \mathbf{b}u_{t-f} + \mathbf{w}_t, \\ y_t &= \mathbf{c}\mathbf{x}_t + v_t. \end{aligned}$$

The N finite steps optimal control algorithm, given in Vu, K. (2008), is as follows. The performance index value for this case is

$$\underset{u_t}{Min}\quad E\{\sum_{t=k}^{N}\mathbf{x}_{t+f+1}^T\mathbf{Q}_1\mathbf{x}_{t+f+1} + \mathbf{Q}_2u_t^2 \mid \mathcal{Y}_t\} \quad = \quad \hat{\mathbf{x}}_{k+f|k}^T\mathbf{S}_k\hat{\mathbf{x}}_{k+f|k} + q_k$$

with

$$\mathbf{S}_t = \mathbf{A}^T(\mathbf{S}_{t+1}+\mathbf{Q}_1)\mathbf{A} - \frac{\mathbf{A}^T(\mathbf{S}_{t+1}+\mathbf{Q}_1)\mathbf{b}\mathbf{b}^T(\mathbf{S}_{t+1}+\mathbf{Q}_1)\mathbf{A}}{\mathbf{Q}_2 + \mathbf{b}^T(\mathbf{S}_{t+1}+\mathbf{Q}_1)\mathbf{b}},$$

$$q_t = q_{t+1} + \mathbf{k}^T\mathbf{S}_{t+1}\mathbf{k}\sigma_a^2 + tr(\mathbf{A}^{f+1^T}\mathbf{Q}_1\mathbf{A}^{f+1}\mathbf{P}_0) + tr(\sum_{i=0}^{f}\mathbf{A}^{i^T}\mathbf{Q}_1\mathbf{A}^i\mathbf{R}_w).$$

and the initial conditions

$$\begin{aligned} \mathbf{S}_N &= \mathbf{A}^T\mathbf{Q}_1\mathbf{A} - \frac{\mathbf{A}^T\mathbf{Q}_1\mathbf{b}\mathbf{b}^T\mathbf{Q}_1\mathbf{A}}{\mathbf{Q}_2 + \mathbf{b}^T\mathbf{Q}_1\mathbf{b}}, \\ q_N &= tr(\mathbf{A}^{f+1^T}\mathbf{Q}_1\mathbf{A}^{f+1}\mathbf{P}_0) + tr(\sum_{i=0}^{f}\mathbf{A}^{i^T}\mathbf{Q}_1\mathbf{A}^i\mathbf{R}_w). \end{aligned}$$

The control actions to achieve this result are:

$$u_t = -[\mathbf{Q}_2 + \mathbf{b}^T(\mathbf{S}_{t+1}+\mathbf{Q}_1)\mathbf{b}]^{-1}\mathbf{b}^T(\mathbf{S}_{t+1}+\mathbf{Q}_1)\mathbf{A}\hat{\mathbf{x}}_{t+f|t},$$
$$u_N = -[\mathbf{Q}_2 + \mathbf{b}^T\mathbf{Q}_1\mathbf{b}]^{-1}\mathbf{b}^T\mathbf{Q}_1\mathbf{A}\hat{\mathbf{x}}_{N+f|N}.$$

The state estimator $\hat{\mathbf{x}}_{t+f|t}$ is, of course, obtained from Kalman filter theory as

$$\begin{aligned} \hat{\mathbf{x}}_{t+f|t} &= \mathbf{A}^{f-1}\hat{\mathbf{x}}_{t+1|t} + \sum_{i=1}^{f-1}\mathbf{A}^{i-1}\mathbf{b}u_{t-i}, \\ \hat{\mathbf{x}}_{t+1|t} &= \mathbf{A}\hat{\mathbf{x}}_{t|t-1} + \mathbf{b}u_{t-f} + \mathbf{k}[y_t - \mathbf{c}\hat{\mathbf{x}}_{t|t-1}], \\ &= [\mathbf{A} - \mathbf{k}\mathbf{c}]\hat{\mathbf{x}}_{t|t-1} + \mathbf{b}u_{t-f} + \mathbf{k}y_t \end{aligned}$$

with the matrix $\mathbf{k}$ as the Kalman filter matrix for the one step ahead state estimator.

By first look, one would think that the N finite steps optimal control algorithm is not physically implementable. Because at the time t when u_t is calculated, we have to know $\mathbf{S}_{t+1}$, a quantity with a time reference in the future of the time of calculation. In fact, $\mathbf{S}_{t+1}$'s and $\mathbf{L}_t$'s are calculated off-line and stored in the computer memory. The controller $\mathbf{L}_t$ is called up at the time t to calculate u_t. This means that the parameter N must be known. If N is large and storage for $\mathbf{L}_t$'s becomes a problem, e.g. limited memory or N is not known, control engineers can use the steady state version of the algorithm. This means that one obtains the controller as

$$\mathbf{L}_\infty = [\mathbf{Q}_2 + \mathbf{b}^T(\mathbf{S}_\infty+\mathbf{Q}_1)\mathbf{b}]^{-1}\mathbf{b}^T(\mathbf{S}_\infty+\mathbf{Q}_1)\mathbf{A}$$

with the matrix $\mathbf{S}_\infty$ obtained by iterating

$$\mathbf{S}_t = \mathbf{A}^T(\mathbf{S}_{t+1}+\mathbf{Q}_1)\mathbf{A} - \frac{\mathbf{A}^T(\mathbf{S}_{t+1}+\mathbf{Q}_1)\mathbf{b}\mathbf{b}^T(\mathbf{S}_{t+1}+\mathbf{Q}_1)\mathbf{A}}{\mathbf{Q}_2 + \mathbf{b}^T(\mathbf{S}_{t+1}+\mathbf{Q}_1)\mathbf{b}}$$

until convergence. And the control action is calculated as

$$u_t = -\mathbf{L}_\infty\hat{\mathbf{x}}_{t+f|t}.$$

In this case, we can call the control algorithm a pseudo infinite steps optimal control algorithm. Note that the algorithm is not infinite steps optimal, because this algorithm will require a spectral factorization and a spectral separation equations to obtain the controller (Vu, K. (2008)). The control algorithm can also be viewed as a myopic control algorithm but with the weighting matrix $\mathbf{S}_\infty + \mathbf{Q}_1$ on the state vector.

5 Conclusion

In this paper, we have presented a new algorithm for the N steps optimal control algorithm. The new algorithm gives the correct performance index value of the control algorithm. The new algorithm is derived by the method of dynamic programming, as the existing algorithm. Because of this, the two algorithms have equivalent Riccati and controller equations. However, the existing algorithm has a dubious statistic foundation. This is the reason for it not to have a verifiable performance index value and a confusion of its author as to which state estimator ($\hat{\mathbf{x}}_{t|t}$ or $\hat{\mathbf{x}}_{t|t-1}$) to be used in the controller. The N steps optimal control algorithm is physically implementable and is suitable for control systems where a definite small number of optimal control steps is required. When the number of steps N is very large, one can use the steady state version of the algorithm.

References

Åström, K.J. (1970) *Introduction to Stochastic Control Theory*. Academic Press, New York, NY, USA. ISBN 0-120-65650-7.

Åström, K.J. and Wittenmark, B. (1989) *Adaptive Control*. Addison Wesley Publishing Company, Reading, MA, USA. ISBN 0-201-09720-6.

R. Bellman (1997) *Introduction to Matrix Analysis*. Second Edition. SIAM, Philadelphia, PA, USA. ISBN 0-898-71399-4.

T. Kailath (1980) *Linear Systems*. Prentice-Hall Inc., Upper Saddle River, NJ, USA. ISBN 0-13-536961-4.

E. Mosca (1995) *Optimal, Predictive and Adaptive Control*. Prentice-Hall Inc., Englewood Cliffs, NJ, USA. ISBN 0-13-847609-8.

Ky M. Vu (2008) *Optimal Discrete Control Theory: The Rational Function Structure Model*. AuLac Technologies Inc., Ottawa, ON, Canada, ISBN 978-0-9783996-0-3.

Ky M. Vu (2007) *The ARIMA and VARIMA Time Series: Their Modelings, Analyses and Applications*. AuLac Technologies Inc., Ottawa, ON, Canada, ISBN 978-0-9783996-1-0.