A STOCHASTIC OPTIMAL DIRECT CONTROLLER FOR DISCRETE LINEAR SYSTEMS

ATAIR RIOS NETO

Inpe, Divisão de Mecânica Espacial e Controle
ataírrn@uol.com.br

TAKASHI YONEYAMA

ITA, Divisão Engenharia Eletrônica
takashi@ita.br

Abstract— A new form of solution is developed for the problem of the stochastic optimal regulator of discrete dynamic linear systems. Initially, using the optimality principle of Dynamic Programming it is shown that the solution to this problem can alternatively be obtained employing optimal linear estimation in each stage or step of discretization. Next, a new controller is obtained, as the result of hypothetically assuming the dynamic system reversible and analogous to a virtual multistage process with inverted flux and of treating the solution in an alternative and equivalent form employing optimal linear estimation. This new controller is a direct discrete sequential one; that is, a controller whose gain only depends on past and present information about the system. To illustrate the method behavior and to allow comparison with a previously developed similar heuristic approach, preliminary tests are done using the same test problem of satellite attitude control. The proposed solution is expected to lead to a more adequate controller to be used in adaptive control schemes.

Keywords— Stochastic Optimal Control; Optimal Regulator; Dynamical Systems Control; Dynamic Programming

Resumo— Uma nova forma de solução é desenvolvida para o problema do regulador ótimo estocástico de sistemas dinâmicos lineares discretos. Inicialmente, a partir de abordagem pelo princípio da otimalidade de Programação Dinâmica, mostra-se que a solução do problema pode ser obtida de forma alternativa equivalente, pela utilização de estimação linear ótima, em cada estágio ou ponto de discretização. Em seguida, obtém-se um novo controlador como resultado de hipoteticamente se admitir o sistema dinâmico reversível e análogo a um processo de múltiplos estágios, que virtualmente evolui no sentido dos tempos decrescentes, e de se tratar a solução de forma alternativa equivalente pela utilização de estimação linear ótima. Este novo controlador é sequencial direto, isto é, um controlador cujo ganho só depende de informações passadas e presentes do sistema. Para ilustrar o comportamento do método e para permitir comparação com uma abordagem heurística semelhante, previamente desenvolvida, testes preliminares são feitos utilizando um mesmo problema teste de controle de atitude de satélite. Espera-se que a solução proposta leve a procedimento mais adequado para o emprego em esquemas adaptativos de controle.

Palavras-chave— Controle Ótimo Estocástico; Regulador Ótimo; Controle de Sistemas Dinâmicos; Programação Dinâmica.

1 Introduction

In the search for an optimized solution of a linear system control problem it is usual the adoption of a quadratic index of performance. The gains of the resulting controller for the usual optimal control approach rely on future system dynamics knowledge. If Dynamic Programming (e.g., Bryson and Ho, 1975) principle of optimality and parameter optimal linear estimation (e.g., Jazwinski, 1970) are applied it can be shown that the optimal control can be equivalently determined in each stage. Though with this approach a stochastic meaning of the inverse of covariance matrices is given to the weight matrices, the characteristic of relying on future system dynamic knowledge still remains. This is a serious limitation in the most common situations where the linear dynamics is only a local approximation of the dynamical behavior of the system and where adaptive schemes are necessary.

In this work a new sequential controller applicable to discrete linear systems is proposed which has the distinctive characteristic of having the generation of gains only dependent on past and present system information. Based on the fact that the transition matrix does always have an inverse, the discrete linear system is viewed as a process of multiple stages in the reversed sense. If the Dynamic Programming principle of optimality is applied, it is again possible to formally treat the control action determination in a typical stage as a problem of parameter optimal linear estimation. However, this can now can be done in the sense of progressing time, i.e., such as to have a direct sequential controller.

The new and original version presented is inspired and related to a previous work (Rios Neto and Cruz, 1990) in which a regulator type of controller also capable of operating in the direct and sequential form was heuristically proposed. This present version reconsiders, reviews and updates a previous preliminary version (Rios Neto, 1990) in a renewed effort of establishing theoretical basis for other previous heuristic approaches adopted to synthesize adaptive controllers applied to ship control (Rios Neto and Cruz, 1985) and satellite control (Ferreira et al, 1985). If it is considered the interpretation of modeling control actions magnitude and deviations in state or output as dispersions using variances, then the new controller proposed is related to minimum variance controllers (e.g.: Li and Evans, 1997, 2002; Kharrat et al, 2010) and to predictive control schemes where Kalman filtering is used to estimate the control actions (e.g.: Silva and Rios Neto, 2000; Silva, 2000; Silva and Rios Neto, 2011).

In what follows, section 2 presents the formulation and usual form of solution of the problem as presented in Bryson and Ho (1975). In section 3, the proposed approach is developed, keeping the use of Bryson and Ho (1975) notation. In order to allow comparison of performance with a similar approach, section 4 presents results of preliminary tests using the same problem of satellite attitude control as in Rios Neto and Cruz (1990). Finally, in section 5 a few conclusions are drawn.

2 Usual Formulation and Solution

Given the discrete dynamical system and the correspondent observations for $i=1,2,\ldots,n-1$:

$$x(i+1) = \phi(i+1,i)x(i) + \Gamma(i)u(i) + w(i) = f^i\big(x(i),u(i),w(i)\big) \quad (1)$$

$$y(i) = H(i)x(i) + v(i), \quad (2)$$

where $v(i)$, $w(i)$, $x(0)$ are zero mean Gaussian, such that for $j=1,2,\ldots,n-1$:

$$E[v(i)w^T(j)] = 0,\ E[x(0)w^T(i)] = 0,$$
$$E[x(0)v^T(i)] = 0,\ E[w(i)w^T(j)] = Q(i)\delta_{ij},$$
$$E[x(0)x^T(0)] = P(0),\ E[v(i)v^T(j)] = R(i)\delta_{ij}.$$

The objective is to control the system to get $x(i)$ as close to zero as possible, according to a given strategy, as, for example, that of minimizing the criterion of performance:

$$J = E[\tfrac{1}{2}x^T(n)S(n)x(n) + \tfrac{1}{2}\textstyle\sum_{i=0}^{n-1}(x^T(i)A(i)x(i) + u^T(i)B(i)u(i))], \quad (3)$$

where $S(n)$, $A(i)$, $B(i)$ are weight matrices, assumed to be positive definite.

The solution to this problem obey the principle of separation (see, for example, Bryson and Ho, 1975), and is as follows, supposing that $S(n)$ and $\bar{P}(0)$ are given.

$$u^*(i) = -C(i)\hat{x}(i), \quad (4)$$

$$\hat{x}(i) = \bar{x}(i) + K(i)\big(y(i) - H(i)\bar{x}(i)\big), \quad (5)$$

$$\bar{x}(i+1) = \phi(i+1,i)\hat{x}(i) + \Gamma(i)u^*(i), \quad (6)$$

$$C(i) = \big(\Gamma^T(i)S(i+1)\Gamma(i) + B(i)\big)^{-1}\Gamma^T(i) S(i+1)\phi(i+1,i), \quad (7)$$

$$K(i) = \bar{P}(i)H^T(i)(H(i)\bar{P}(i)H^T(i) + R(i))^{-1}, \quad (8)$$

$$S(i) = \phi^T(i+1,i)S(i+1)\phi(i+1,i) - C^T(i)\big(B(i) + \Gamma^T(i)S(i+1)\Gamma(i)\big)C(i) + A(i), \quad (9)$$

$$\bar{P}(i+1) = \phi(i+1,i)P(i)\phi^T(i+1,i) + Q(i), \quad (10)$$

$$P(i) = \big(I - K(i)H(i)\big)\bar{P}(i)\big(I - K(i)H(i)\big)^T + K(i)R(i)K^T(i). \quad (11)$$

The solution has the same structural form as that of the correspondent deterministic problem where there is a perfect knowledge of the state. The difference is that the control action is calculated using the value of the best estimate of the state, as given by the Kalman filter. The difficulties with this solution are those related to: (i) the choice of the weight matrices $S(n)$, $A(i)$, $B(i)$, (ii) the fact that the gains $C(i)$ depend on information which have to be generated from final to present time, complicating its application to the control of time variant systems with modeling errors.

3 Alternative Equivalent Form of Solution

The optimal control problem of Equations (1) and (3) of previous section can be viewed and treated considering a correspondent process of multiple stages as in Figure 1, where each stage corresponds to a step of discretization (Eq. (1)).

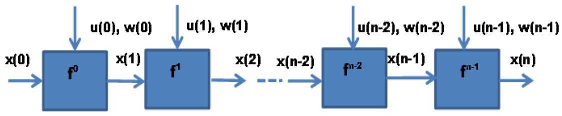


Figure 1: Dynamical System as a Process of Multiple Stages

The optimal control problem can then be solved considering the principle of certainty-equivalence, using the optimality principle of Dynamic Programming. With this approach, the optimal control can be determined starting in the last stage and progressing back to the first stage.

$$V_1(Y_{n-1}) = \min_{u(n-1)} E[1/2(x^T(n)S(n)x(n) + x^T(n-1)A(n-1)x(n-1) + u^T(n-1)B(n-1)u(n-1)) \mid Y_{n-1}], \quad (12)$$

where, $\{y(i): i = 0,1,2,\ldots,n-1\} \equiv Y_{n-1}$ is the set of known measurements. Considering,

$$\hat{x}(i) = x(i) + e(i),\ \ i = 0,1,2,\ldots,n-1, \quad (13)$$

where $e(i)$ is orthogonal to the estimate $\hat{x}(i)$ and not dependent on $u(i)$. Taking this in account and considering the dynamical constraint of Equation (1), together with the fact that $w(i)$ is a white noise, it results, unless of constants regarding the minimization with respect to $u(n-1)$, that, equivalently:

$$V_1(Y_{n-1}) = \min_{u(n-1)}[1/2(\phi(n,n-1)\hat{x}(n-1) + \Gamma(n-1)u(n-1))^T S(n)\big(\phi(n,n-1)\hat{x}(n-1) + \Gamma(n-1)u(n-1)\big) + u^T(n-1)B(n-1)u(n-1))]. \quad (14)$$

In the last stage, the solution of this minimization problem can then be shown to be formally equivalent to the optimal minimum variance solution of the parameter estimation problem that follows (see e.g. Bryson and Ho, 1975).

$$0 = u(n-1) + \varepsilon_u(n-1)\,, \quad (15)$$
$$0 = \bar{x}(n) + \varepsilon_x(n), \quad (16)$$

$\varepsilon_u(n-1) \sim N\big(0, B^{-1}(n-1)\big)$, $\varepsilon_x(n) \sim N\big(0, S^{-1}(n)\big)$, and $\bar{x}(n)$ is the estimated controlled state at last stage, given by:

$$\bar{x}(n) = \phi(n,n-1)\hat{x}(n-1) + \Gamma(n-1)u(n-1). \quad (17)$$

In fact, taking this value back in Equation (16), the minimum variance estimate (see e.g., Jazwinski, 1970) of $u(n-1)$ results as:

$$\hat{u}(n-1) = -\big(\Gamma^T(n-1)S(n)\Gamma(n-1) + B(n-1)\big)^{-1} \Gamma^T(n-1)S(n)\phi(n,n-1)\hat{x}(n-1). \quad (18)$$

Formally comparing this expression with that of Equations (3) and (7), it is seen that:

$$\hat{u}(n-1) = -C(n-1)\hat{x}(n-1) \equiv u^*(n-1). \quad (19)$$

Now using again the principle of optimality from the stage previous to the last one, there results that:

$$V_2(Y_{n-2}) = \min_{u(n-2)} E[(V_1(Y_{n-1}) + 1/2(\,x^T(n-2) A(n-2)x(n-2) + u^T(n-2)B(n-2)u(n-2)) \mid Y_{n-2}], \quad (20)$$

But, the optimal $V_1(Y_{n-1})$ is given by, unless of a constant,

$$V_1(Y_{n-1}) = [\tfrac{1}{2}((\phi(n,n-1)\hat{x}(n-1) -\Gamma(n-1)C(n-1)\,\hat{x}(n-1))^T\, S(n) (\phi(n,n-1)\hat{x}(n-1) - \Gamma(n-1)C(n-1)\hat{x}(n-1)) +\big(C(n-1)\hat{x}(n-1)\big)^T B(n-1)\big(C(n-1)\hat{x}(n-1)\big) +\hat{x}(n-1)^T A(n-1)\hat{x}(n-1))]. \quad (21)$$

If is taken in account that: $\hat{x}(n-1) = \bar{x}(\,n-1) + \bar{e}(n-1)$, and that $\bar{e}(n-1)$ depends on the observation $y(n-1)$, it finally results that:

$$V_2(Y_{n-2}) = \min_{u(n-2)} [1/2((\phi(n,n-1)\bar{x}(n-1) -\Gamma(n-1)C(n-1)\,\bar{x}\,(n-1))^T S(n) \big(\phi(n,n-1)\bar{x}(n-1) - \Gamma(n-1)C(n-1)\bar{x}(-1)\big) +\big(C(n-1)\bar{x}(n-1)\big)^T B(n-1) (C(n-1)\bar{x}(n-1)) + \bar{x}(n-1)^T A(n-1)\bar{x}(n-1) +u^T(n-2)B(n-2)u(n-2))] + cte. \quad (22)$$

This minimization problem can then be shown to be formally equivalent to the optimal minimum variance solution of the parameter estimation problem that follows:

$$0 = \bar{x}(n-1) + \varepsilon_x^a(n-1), \quad \varepsilon_x^a(n-1) \sim N\big(0, A^{-1}(n-1)\big), \quad (23a)$$

$$0 = -C(n-1)\bar{x}(n-1) + \varepsilon_u(n-1), \quad \varepsilon_u(n-1) \sim N\big(0, B^{-1}(n-1)\big), \quad (23b)$$

$$0 = \phi(n,n-1)\bar{x}(n-1) - \Gamma(n-1)C(n-1)\bar{x}(n-1) + \varepsilon_x(n), \quad \varepsilon_x(n) \sim N\big(0, S^{-1}(n)\big), \quad (23c)$$

$$0 = u(n-2) + \varepsilon_u(n-2), \quad \varepsilon_u(n-2) \sim N\big(0, B^{-1}(n-2)\big). \quad (23d)$$

However, since the errors are independent random variables, the first three equations can be processed in batch, generating an a priori information to then process the last equation. But this equivalently means that for the stage previous to the last one, the following a priori condition has to be considered for $\bar{x}(n-1)$, the controlled state:

$$0 = \bar{x}(n-1) + \varepsilon_x^a(n-1). \quad (24)$$

In the a priori distribution as given by condition in Equation (24), the optimal trajectory has to satisfy the condition of optimal solution according to the principle of optimality, which requires that from an outcome of the random process $\bar{x}(n-1)$:

$$\bar{x}(n-1) = \phi(n,n-1)\bar{x}(n-1) + \Gamma(n-1)\bar{u}(n-1), \quad \bar{u}(n-1) = -C(n-1)\bar{x}(n-1). \quad (25)$$

And, at the same time, this outcome has to be compatible with the distributions of Equations (15) and (16), that is:

$$0 = -C(n-1)\bar{x}(n-1) + \varepsilon_u(n-1), \quad (26)$$

$$0 = \phi(n,n-1)\bar{x}(n-1) - \Gamma(n-1)C(n-1)\bar{x}(n-1) + \varepsilon_x(n). \quad (27)$$

The outcome compatible with the distributions of Equations (24), (26) and (27), and which is optimal, is equivalently given by the minimum variance estimate (Gauss-Markov), which can be expressed as:

$$\hat{\bar{x}}(n-1) = \bar{x}(n-1) + \varepsilon_x(n-1), \quad (28)$$

$$\hat{\bar{x}}(n-1) = 0, \quad (29)$$

where $\hat{\bar{x}}(n-1)$ is the expected value conditioned to the conditions of Equations (24), (26) and (27), and with an estimated zero mean error ε_x with distribution:

$$E[\varepsilon_x(n-1)\varepsilon_x(n-1)^T] = (A(n-1) + C^T(n-1) B(n-1)C(n-1) + (\phi^T(n,n-1) - C^T(n-1) \Gamma^T(n-1))S(n)(\,\phi(n,n-1) - \Gamma(n-1)C(n-1)))^{-1}$$
$$\equiv (A(n-1) + \phi^T(n,n-1)S(n)\phi(n,n-1) - C^T(n-1) \big(B(n-1) + \Gamma^T(n-1)S(n)\Gamma(n-1))C(n-1)\big)^{-1} \quad (30)$$

Comparing with Equation (9), it is seen that:

$$E[\varepsilon_x(n-1)\varepsilon_x(n-1)^T] = S^{-1}(n-1). \quad (31)$$

With this result, the cycle of analysis for the stage previous to the last one is done and the problem of determination of the optimal control in this stage is formally equivalent to solving the optimal parameter estimation problem:

$$0 = u(n-2) + \varepsilon_u(n-2), \tag{32}$$

$$0 = \bar{x}(n-1) + \varepsilon_x(n-1),$$
$$\bar{x}(n-1) = \phi(n-1, n-2)\hat{x}(n-2)$$
$$+\Gamma(n-2)u(n-2). \tag{33}$$

The solution of this problem gives the minimum variance estimate:

$$u(n-2) = -C(n-2)\hat{x}(n-2) = u^*(n-2), \tag{34}$$

where $C(n-2)$ is as in Equation (7), taking *i=n-2*.

The validity for the remaining stages can be demonstrated in a completely analogous way as that done for the stage previous to the last stage.

4 Reverse Virtual Process of Multiple Stages and Dynamic Programming Approach

Consider now the system of multiple stages of Figure 2, where it is assumed the hypothetical and virtual situation, in which the evolution of the states $x(i)$ is considered in a reverse sense, from final to initial state. Though a hypothetical and virtual situation, this is numerically possible in the case of discretized linear dynamic systems, due to the fact that the transition matrix does always have an inverse.

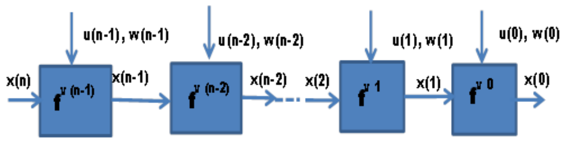


Figure 2: Process of Multiple Stages with Reverse Flux

Under this assumption, the following optimization problem is posed:

$$\textit{Minimize: } J = E[\tfrac{1}{2}(x^d(0) - x(0))^T S(0)(x^d(0) - x(0))$$
$$+\tfrac{1}{2}\textstyle\sum_{i=n}^{1}((x^d(i) - x(i))^T A(i)(x^d(i) - x(i)) +$$
$$u^T(i-1)B(i-1)u(i-1))], \tag{35}$$

$$\textit{Subject to: } x(i-1) = \phi^{-1}(i, i-1)x(i) -$$
$$\phi^{-1}(i, i-1)\big(\Gamma(i-1)u(i-1) + w(i-1)\big)$$
$$= f^{v(i-1)}(x(i), u(i-1), w(i-1)), \tag{36}$$

where, *i=n,n-1,…,1, S(0), A(i), B(i)* are weight matrices, assumed to be positive definite, and $x^d(i)$ is the desired trajectory to be followed.

As previously done, a Dynamic Programming approach is taken in a total similar way as before, but applied to a virtual reality. The optimal control can be determined starting in the first stage and virtually progressing back to the last stage.

$$V_1(Y_1^v) = \min_{u(0)} E[1/2((x^d(0) - x(0))^T S(0)(x^d(0)$$
$$-x(0)) + (x^d(1) - x(1))^T A(1)(x^d(1) - x(1)) +$$
$$u^T(0)B(0)u(0)) \mid Y_1^v], \tag{37}$$

Where, now $\{y^v(i): i = n, \dots, 2, 1\} \equiv Y_1^v$ is the set of known virtual measurements. Considering

$$\hat{x}^v(i) = x(i) + e^v(i),\ i = 1, 2, \dots, n, \tag{38}$$

where $e^v(i)$ is orthogonal to the optimal virtual estimate $\hat{x}^v(i) = E[x^v(i) \mid Y_i^v]$ and not dependent on $u(i-1)$. Taking this in account, considering the dynamical constraint of Equation (36), together with the fact that $w(i-1)$ is a white noise, regarding that the minimization in Equation (37) is with respect to $u(0)$ and that the conditioned expectation is with respect to Y_1^v, it results that, unless of constants, equivalently:

$$V_1(Y_1^v) = \min_{u(0)}[1/2((x^d(0) - \phi^{-1}(1,0)(\hat{x}^v(1) -$$
$$\Gamma(0)u(0)))^T S(0)(x^d(0)$$
$$-\phi^{-1}(1,0)(\hat{x}^v(1) - \Gamma(0)u(0))) + u^T(0)B(0)u(0))]. \tag{39}$$

Thus, in the last stage the solution of the associated problem of control is formally equivalent to the optimal minimum variance solution of the parameter estimation problem that follows.

$$0 = u(0) + \varepsilon_u(0), \tag{40}$$

$$0 = x^d(0) - \bar{x}^v(0) + \varepsilon_x(0), \tag{41}$$

where $\varepsilon_u(0) \sim N(0, B^{-1}(0))$, $\varepsilon_x(0) \sim N(0, S^{-1}(0))$, and $\bar{x}^v(0)$ is the controlled however virtual state at last stage, given by:

$$\bar{x}^v(0) = \phi^{-1}(1,0)\big(\hat{x}^v(1) - \Gamma(0)u(0)\big). \tag{42}$$

The solution to this problem is given by the minimum variance Gauss-Markov estimator:

$$\hat{u}(0) = \big(\Gamma^T(0)\bar{S}(0)\Gamma(0) + B(0)\big)^{-1}\Gamma^T(0)\bar{S}(0)$$
$$\big(\hat{x}^v(1) - \phi(1,0)x^d(0)\big)$$
$$= C^v(0)\big(\hat{x}^v(1) - \phi(1,0)x^d(0)\big), \tag{43}$$

where $\bar{S}^{-1}(0)$ is the covariance matrix of $\phi(1,0)\varepsilon_x(0)$.

Notice that $C^v(0) \neq C(0)$, and that $\hat{u}(0) = \hat{u}^v(0)$ is a virtual control. At this point it is already necessary to point out that in order to give practical use to this result and come back to the real world of progressing time, one has to heuristically specify $\hat{x}^v(1)$ such as to get regulation and notice that $x^d(0)$ can only be chosen to be $\hat{x}(0)$, in a certainty equivalent way. It all happens as if one had a predictive control with one step ahead.

Now using again the principle of optimality from the stage previous to the last one, there results:

$$V_2(Y_2^v) = \min_{u(1)} E[(V_1(Y_1^v) + 1/2((x^d(2) -$$
$$x(2))^T A(2)(x^d(2) - x(2)) + u^T(1)B(1)u(1)) \mid Y_2^v], \tag{44}$$

But, the optimal $V_1(Y_1^v)$ is given by, unless of a constant,

$$V_1(Y_1^v) = [1/2((x^d(0) - \phi^{-1}(1,0)(\hat{x}^v(1) - \Gamma(0)\hat{u}(0)))^T S(0)(x^d(0) - \phi^{-1}(1,0)(\hat{x}^v(1) - \Gamma(0)\hat{u}(0))) + (x^d(1) - \hat{x}^v(1))^T A(1)(x^d(1) - \hat{x}^v(1)) + \hat{u}^T(0)B(0)\hat{u}(0))].$$

If it is considered that: $\hat{x}^v(1) = \bar{x}^v(1) + \bar{e}^v(1)$, and that $\bar{e}^v(1)$ depends on the observation $y^v(1)$, it results:

$$V_2(Y_2^v) = \min_{u(1)}[1/2((x^d(0) - \phi^{-1}(1,0)(\bar{x}^v(1) - \Gamma(0)\bar{u}(0)))^T S(0)(x^d(0) - \phi^{-1}(1,0)(\bar{x}^v(1) - \Gamma(0)\bar{u}(0))) + (x^d(1) - \bar{x}^v(1))^T A(1)(x^d(1) - \bar{x}^v(1)) + \bar{u}^T(0)B(0)\bar{u}(0) + u^T(1)B(1)u(1))]. \quad (45)$$

Where, the condition of optimal solution, according to the principle of optimality, requires that from an outcome of $\bar{x}^v(1)$ the optimal trajectory control has to be such as to satisfy:

$$\bar{u}(0) = C^v(0)\left(\bar{x}^v(1) - \phi(1,0)x^d(0)\right), \quad (46)$$

As done in Section 3.1, this minimization can equivalently be viewed and interpreted as an optimal minimum variance parameter estimation problem as follows.

For the a priori condition:

$$0 = x^d(1) - \bar{x}^v(1) + \varepsilon_x^a(1), \quad (47)$$

where $\varepsilon_x^a(1) \sim N(0, A^{-1}(1))$, at the same time the controlled $\bar{x}^v(1)$ has to be an outcome compatible with:

$$0 = C^v(0)\left(\bar{x}^v(1) - \phi(1,0)x^d(0)\right) + \varepsilon_u(0),$$
$$0 = (x^d(0) - \phi^{-1}(1,0)(\bar{x}^v(1) - \Gamma(0)C^v(0)(\bar{x}^v(1) - \phi(1,0)x^d(0))) + \varepsilon_x(0). \quad (48)$$

Or, in a more compact notation:

$$y^c(1) = H^c(1)\bar{x}^v(1) + \bar{\varepsilon}_x^c(1), \quad (49)$$

$$y^{c^T}(1) \triangleq ((C^v(0)\phi(1,0)x^d(0))^T : ((\Gamma(0)C^v(0) - I_n)\phi(1,0)x^d(0))^T),$$

$$H^{c^T}(1) \triangleq (C^{v^T}(0) : (\Gamma(0)C^v(0) - I_n)^T),$$

$$\bar{\varepsilon}_x^{c^T}(1) \triangleq \left(\varepsilon_u{}^T(0) : \bar{\varepsilon}_x{}^T(0)\right),\ \bar{\varepsilon}_x(0) = \phi(1,0)\varepsilon_x(0),$$

$$\bar{\varepsilon}_x^c(1) \sim N\left(0, S^{c^{-1}}(1)\right).$$

The outcomes compatible with the distributions of Equations (46) and (49),

$$\widehat{\bar{x}^v}(1) = \bar{x}^v(1) + \varepsilon_x(1), \quad (50)$$

are given by the Gauss Markov estimator, which in the Kalman form is given by:

$$\widehat{\bar{x}^v}(1) = x^d(1) + K(1)(y^c(1) - H^c(1)x^d(1)) \quad (51)$$

$$K(1) = A^{-1}(1)H^{c^T}(1)\left(H^c(1)A^{-1}(1)H^{c^T}(1) + S^{c^{-1}}(1)\right)^{-1} \equiv \left(I_n + A^{-1}(1)H^{c^T}(1)S^c(1)H^c(1)\right)^{-1} H^{c^T}(1)S^c(1), \quad (52)$$

$$E[\varepsilon_x(1)\varepsilon_x{}^T(1)] \triangleq S^{-1}(1) = A^{-1}(1) - K(1)H^c(1)A^{-1}(1). \quad (53)$$

With this result the cycle of analysis for the previous to last stage is complete and the problem of determination of the optimal control is formally equivalent to the parameter optimal estimation problem that results combining Equation (50), repeated bellow, with the condition in the control dispersion:

$$\widehat{\bar{x}^v}(1) = \bar{x}^v(1) + \varepsilon_x(1),$$
$$0 = u(1) + \varepsilon_u(1), \quad (54)$$

where $\varepsilon_u(1) \sim N\left(0, B^{-1}(1)\right)$, $\varepsilon_x(1) \sim N\left(0, S^{-1}(1)\right)$, and the dynamical constraint has to be considered:

$$\bar{x}^v(1) = \phi^{-1}(2,1)\hat{x}^v(2) - \phi^{-1}(2,1)\Gamma(1)u(1). \quad (55)$$

The solution to this problem is again that given by the minimum variance estimator:

$$\hat{u}(1) = C^v(1)\left(\hat{x}^v(2) - \phi(2,1)\widehat{\bar{x}^v}(1)\right), \quad (56)$$

$$C^v(1) = \left(\Gamma^T(1)\bar{S}(1)\Gamma(1) + B(1)\right)^{-1}\Gamma^T(1)\bar{S}(1), \quad (57)$$

where $\bar{S}^{-1}(1)$ is the covariance matrix of $\phi(2,1)\varepsilon_x(1)$.

For the other stages the situation is analogous and it is only necessary to adequate the indices of Equations (49) to (57), resulting for $i=1,\ldots,n-1$:

$$\bar{S}^{-1}(i) = \phi(i+1,i)S^{-1}(i)\phi^T(i+1,i) \quad (58)$$

$$\widehat{\bar{x}^v}(i) = x^d(i) + K(i)(y^c(i) - H^c(i)x^d(i)) \quad (59)$$

$$K(i) = A^{-1}(i)H^{c^T}(i)\left(H^c(i)A^{-1}(i)H^{c^T}(i) + S^{c^{-1}}(i)\right)^{-1} \equiv \left(I_n + A^{-1}(i)H^{c^T}(i)S^c(i)H^c(i)\right)^{-1} H^{c^T}(i)S^c(i), \quad (60)$$

$$E[\varepsilon_x(i)\varepsilon_x{}^T(i)] \triangleq S^{-1}(i) = A^{-1}(i) - K(i)H^c(i)A^{-1}(i). \quad (61)$$

$$y^{c^T}(i) \triangleq (\left(C^v(i-1)\phi(i,i-1)x^d(i-1)\right)^T : ((\Gamma(i-1)C^v(i-1) - I_n)\phi(i,i-1)x^d(i-1))^T),$$
$$H^{c^T}(i) \triangleq (C^{v^T}(i-1) : (\Gamma(i-1)C^v(i-1) - I_n)^T),$$
$$\bar{\varepsilon}_x^{c^T}(i) \triangleq \left(\varepsilon_u{}^T(i-1) : \bar{\varepsilon}_x{}^T(i-1)\right),$$
$$\bar{\varepsilon}_x(i-1) = \phi(i,i-1)\varepsilon_x(i-1),$$
$$\bar{\varepsilon}_x^c(i) \sim N(0, S^{c^{-1}}(i)).$$

$$\hat{u}(i) = C^v(i)\left(\hat{x}^v(i+1) - \phi(i+1,i)\widehat{\bar{x}^v}(i)\right), \quad (62)$$

$$C^v(i) = \left(\Gamma^T(i)\bar{S}(i)\Gamma(i) + B(i)\right)^{-1}\Gamma^T(i)\bar{S}(i), \tag{63}$$

Where, $\hat{u}(i) = \hat{u}^v(i)$ is a virtual control. Thus, as already pointed out before, to give practical use to this result and come back to the real world of progressing time, one has to heuristically specify $\hat{x}^v(i+1)$ such as to get regulation and $x^d(i)$ (Eq. (59)) can only be chosen to be $\hat{x}(i)$, in a certainty equivalent way. The use of the controller in order to stabilize the system, in the direct sense of real world, from t_0 to t_n, has to necessarily presuppose that a strategy is adopted to condition decreasing $\hat{x}^v(i+1)$ values. Thus, it seems reasonable that (Eq. (59)):

$$x^d(i) = \hat{x}(i). \tag{64}$$

And since there can be a choice for conditioning what is going to happen in the future, it also seems reasonable that (Eq. (62)):

$$\hat{x}^v(i+1) = \propto x^d(i), \ \propto < 1. \tag{65}$$

The choice for $x^d(i)$ as being equal to $\hat{x}(i)$ implies that the dispersion of the error $\varepsilon_x^a(i)$ has to be compatible with the order of magnitude of the error in the implementation of the control $\hat{u}(i)$. The inferior limit of this error $\varepsilon_x^a(i)$ dispersion, in the ideal situation of not having software and hardware imprecisions in the implementation of the control, would be of the order of the dispersion of the state estimation error. However, in practical situations the dispersion provoked by the controller is usually one order of magnitude bigger than that of the state estimator. Thus, in practical situations where state estimation errors are present it is reasonable to assume:

$$E[(\varepsilon_{x\,j}^a(i))^2] = \beta P_{jj}(i), \tag{66}$$

where $P_{jj}(i)$ is the jth term of the diagonal of the matrix of covariances of the errors in the state estimation and $\beta \gg 1$, and usually $\beta \geq 100$.

5. Preliminary Testing

For the sake of comparison, the example taken for preliminary testing is the same as in Rios Neto and Cruz (1990). It is based on a model of double-gimbaled momentum wheel for the attitude control of a geostationary satellite, as given in Kaplan (1976). The satellite has the following characteristics: mass of 716 kg; moments of inertia I_x=I_z=2000 N.m.s^2, I_y=400 N.m.s^2; nominal wheel momentum of 200 N.m.s; and orbital frequency of 7.28E-05 rad/s. It is equipped with sensors capable of observing roll and pitch with accuracy of 5.8E-05 rad as expressed by their standard deviations. The satellites axes x, y, and z are respectively in correspondence with roll, pitch and yaw; the wheel axis coincides with the y axis. In discrete-time form, based on a discretization time interval of 0.1 s, the state model is as follows:

$$\phi(i+1,i) = \begin{bmatrix} 1 & 0 & 1.21E-10 & 0.1 & 0 & -5.0E-04 \\ 0 & 1 & 0 & 0 & 0.1 & 0 \\ -1.21E-10 & 0 & 1 & 5.0E-04 & 0 & 0.1 \\ -7.28E-07 & 0 & 3.64E-09 & 1 & 0 & -0.01 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ -3.64E-09 & 0 & -7.28E-07 & 0.01 & 0 & 1 \end{bmatrix}$$

$$\Gamma(i) = \begin{bmatrix} 0.25E-05 & 0 & -8.33E-09 \\ 0 & 1.25E-05 & 0 \\ 8.33E-09 & 0 & 2.5E-07 \\ 5.0E-05 & 0 & -2.5E-07 \\ 0 & 2.5E-04 & 0 \\ 2.5E-07 & 0 & 5.0E-05 \end{bmatrix}$$

$$Q(i) = diag(1.73E-28; 1.09E-24; 1.75E-29; 6.90E-27; 4.34E-22; 7.01E-27)$$

The initial conditions and controller parameters adjustment were as follows:

$$\hat{x}^T(0) = [2.D-02;\ 2.D-02;\ 2.D-02;\ -2.D-06;\ -2.D-07;\ -6.D-07],$$

$$A^{-1}(i) = diag(2.D-10; 2.D-10; 2.D-12; 1.D-06; 1.D-06; 1.D-06),$$

$$B(i) = diag(1.; 10.; 1.).$$

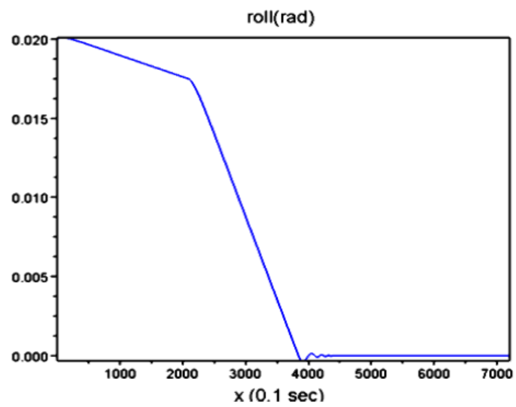
Where, the diagonal terms in $A^{-1}(i)$ were chosen according to its meaning of the a priori variances in the errors modeling the dispersions of controlled states relative to the desired solution. In the case of the variances modeling the dispersions in angles the values were chosen in correspondence to the accuracies to be attained; and in the case of angle rates the choice was such as to guarantee margins that allow variations that may be necessary for convergence in position. The matrix $B^{-1}(i)$ was chosen according to the same meaning, to represent the allowed dispersions in control action.

To guarantee convergent increments in state variables compatible with control action limits, the following choices were made:

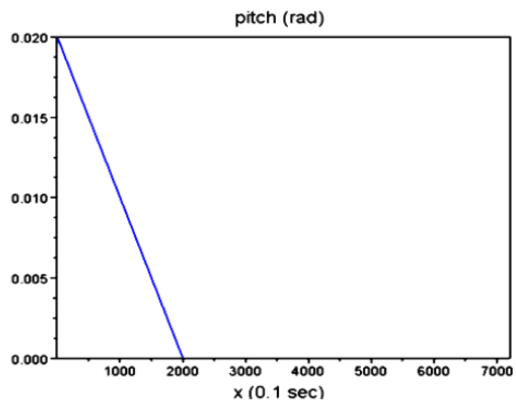
$$\hat{x}_j^v(i+1) = \hat{x}_j(i) - sign\left(\hat{x}_j(i)\right) * 1.\mathrm{D}-05, \qquad \mathrm{j} = 1,2,3,$$
$$\hat{x}_j^v(i+1) = -sign\left(\hat{x}_{j-3}(i)\right) * 1.\mathrm{D}-04, \qquad \mathrm{j} = 4,5,6. \tag{69}$$

To avoid near singularity numerical behavior, in analogy with what is done in state estimation with the addition of process noise, the diagonal terms of the matrix $\bar{S}^{-1}(i)$ (Eq. (54)) were saturated from below as follows:

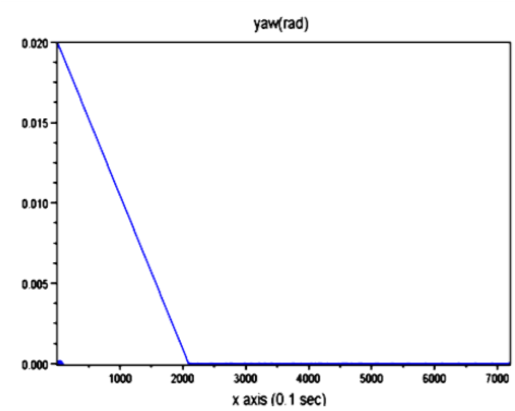
$$\bar{S}^{-1}(i) = \bar{S}^{-1}(i) + A^{-1}(i). \tag{70}$$



(a)



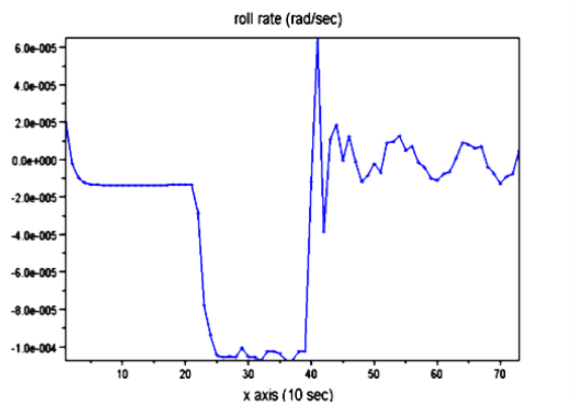
(b)



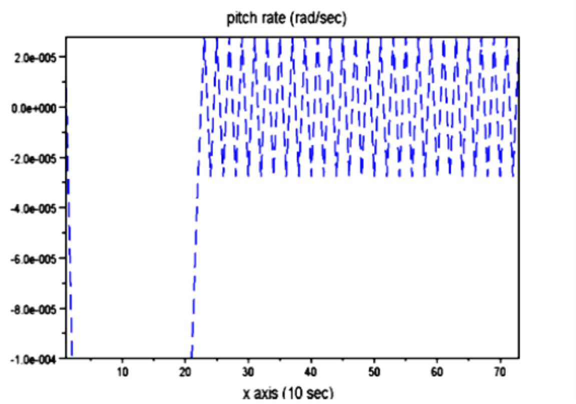
(c)

Figure 3: Three Axis Attitude Control Angles

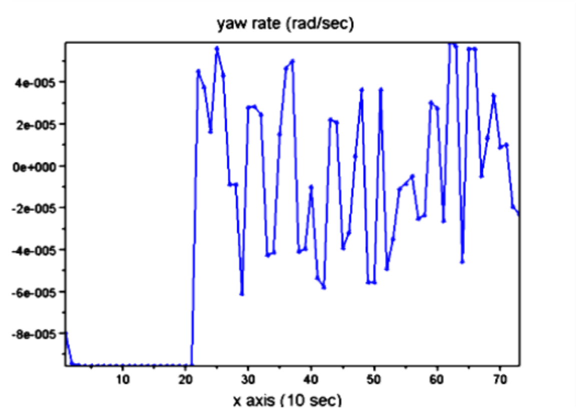
Under ideal simulation conditions of perfect knowledge of state and no process perturbations, the results obtained with the proposed controller are quite satisfactory, as depicted in Figures 3 and 4. The controller response could be taken in the limit of the control action capacity for the problem at hand, reaching a condition of satisfactory errors levels (of the order of magnitude of 1.E-05 rad in roll and yaw, 1.E-06 rad in pitch, and angles rates below 5.E-05 rad/s).



(a)



(b)



(c)

Figure 4: Three Axis Attitude Control Angle Rates

6. Conclusions

A new optimal direct sequential regulator for discrete linear systems was developed. It is optimal in the sense of minimizing a quadratic index of performance in the deviation errors relative to a desired dynamic system trajectory. It is direct and sequential in the sense that the gains to calculate control actions are obtained in each progressing time step.

Using the Dynamic Programming principle of optimality, it was first shown that employing a formally equivalent estimation approach a solution can be obtained which is identical to the usual Maximum Principle optimal solution. With this approach, a stochastic meaning was attached to the weight matrices in the index of performance. It was concluded that control actions order of magnitude and accepted deviations in state can be interpreted as dispersions using variances.

Assuming the dynamic system reversible and analogous to a virtual multistage process with inverted flux, the same approach was then used to develop a direct discrete sequential regulator. The resulting controller is one where the gain depends only on past and present information about the system and where the desired behavior of the controlled state is prescribed one step ahead. The minimization of control action and of state deviation as a formally equivalent problem of optimal estimation allowed again to interpret the weight matrices as error covariance matrices and to empirically use formally equivalent noise process to compensate for the bad numerical behavior in the calculations of the control gain. Thus, this formal equivalence opens the possibility of using results already available in state estimation to improve the numerical performance of the developed controller.

The preliminary test results are encouraging. The fact that the control strategy can be established by prescribing a virtual state one step ahead and that future system dynamics knowledge is not needed to calculate the present control action makes this controller suitable for application on nonlinear systems control where linearized and discrete local approximations can be used combined with adaptive schemes.

References

Bryson, A.E., Ho, Y.C. (1975). *Applied Optimal Control: Optimization, Estimation and Control,* John Wiley, New York.

Ferreira, L.D.D., Rios Neto, A. and Venkataraman, N.S. (1985). Stochastic Control of Pitch Motion of Satellites Using Stabilizing Flaps, *Acta Astronautica,* 12(11), 899-905.

Jazwinski, A.H. (1970). *Stochastic Processes and Filtering Theory,* Academic Press, New York.

Kharrat, M., Kamon, S., Chaari, A. (2010). Weighted minimum-variance self-tuning regulation of stochastic time-varying systems: application to a heat transfer process, *IJ-STA*, 4(1), 1138-1157.

Li, Z., Evans, R.J. (1997). Minimum Variance Control of Linear Time Varying Systems, *Automatica,* 33, 1531-1537.

Li, Z, and Evans, RJ. (2002). Generalized Minimum Variance Control of Linear Time-Varying Systems, *IEE Proceedings - Control Theory and Applications*, 149(1), 111-116.

Rios Neto, A. (1990). Controlador Sequencial Direto para Sistemas Dinâmicos Lineares, (INPE-5167- RPI/239).

Rios Neto, A., Cruz, J.J. (1985). A Stochastic Rudder Control Law for Ship Steering, *Automática*, 21(4), 371-384.

Rios Neto, A., Cruz, J.J. (1990). Design of a Stochastic Linear Discrete Regulator Based on an Optimal State Estimation Approach, *SBA: Controle e Automação*, 2(3), 151 – 157.

Silva, J. A., Rios Neto, A. (2000). Neural Predictive Satellite Attitude Control Based on Kalman Filtering Algorithms. *Proceedings of the 15th International Symposium in Space Flight Dynamics*, CNES, Biarritz, France.

Silva, J. A. (2001). *Controle Preditivo Utilizando Redes Neurais Artificiais Aplicado a Veículos Aeroespaciais*, Doctorate Dissertation, Instituto de Pesquisas Espaciais, São José dos Campos, SP, Brazil (INPE-8480 - TDI/778).

Silva, J. A., Rios Neto (2011). Three-Axis Stabilized Geostationary Satellite Attitude Control Using Neural Predictive Algorithms. *Proceedings of the 10th Brazilian Congress on Computational Intelligence (CBIC´2011)*, Fortaleza, Ceará, Brazil.

ISBN: 978-85-8001-069-5