

Numerical Treatment Of The Bounded-Control LQR Problem By Updating The Final Phase Value

V. Costanza, P. S. Rivadeneira, *Member, IEEE* and J. A. G. Munera

Abstract— A novel approach has been developed for approximately solving the constrained LQR problem, based on updating the final state and costate of an unrestricted related regular problem, and the switching times (when the control meets the constraints). The main result is the expression of a suboptimal control in feedback form by using some corresponding Riccati equation. The gradient method is applied to reduce the cost via explicit algebraic formula for its partial derivatives with respect to the hidden final state/costate of the related regular problem. The numerical method results efficient because it does not involve integrations of states or cost trajectories and reduces the dimension of the relevant unknown parameters. All the relevant objects are calculated from a few auxiliary matrices, which are computed only once and kept in memory. The scheme is here applied to the “cheapest stop of a train” case-study whose optimal solution is already known.

Keywords— optimal control, restricted controls, LQR problem, gradient methods.

I. INTRODUCCIÓN

EL REGULADOR lineal cuadrático (LQR) es, probablemente, el problema más estudiado y utilizado en la literatura de control óptimo. Por otra parte, el formalismo Hamiltoniano ha sido central en el desarrollo de la teoría moderna de control óptimo [1-3]. El problema del regulador cuadrático (LQR) de horizonte finito n -dimensional con controles no acotados conduce a $2n$ ecuaciones diferenciales ordinarias (ODEs) con condiciones de frontera conocidas como las Ecuaciones Canónicas Hamiltonianas (HCEs). Existen métodos tradicionales (ver por ejemplo [4-5]) para resolver este problema de condiciones de frontera, a veces transformándolo a un sistema de condiciones iniciales al introducir ciertos objetos matemáticos adicionales.

En el contexto de los controles acotados, el problema de control óptimo es en general no regular, y su solución todavía no consta de un método estándar de resolución [5-7]. Desde los años sesenta, el Principio del Máximo de Pontryagin (PMP) ha sido el marco teórico necesario para tratar dichos problemas no regulares. En este artículo, para resolver el LQR con restricciones se toma ventaja de ciertas relaciones matemáticas que existen entre el PMP y el formalismo Hamiltoniano. El resultado teórico subyacente [8], base de este artículo, puede ser parafraseado de la siguiente manera: en general, la solución óptima de un problema LQR con controles acotados es la saturación de la solución de otro LQR

irrestringido, este último con la misma dinámica y la misma función de costo que el original, pero con diferentes condiciones iniciales \tilde{x}_0 , y sujeto a una penalización final cuadrática con matriz de penalización \tilde{S} diferente a la original.

En [9], se desarrollaron estrategias “off-line” y “on-line” (basadas en el método del gradiente) para detectar las “nuevas” condiciones iniciales \tilde{x}_0 y la matriz de penalización final \tilde{S} . Se daban allí fórmulas algebraicas para calcular las derivadas parciales del costo con respecto a \tilde{S} y \tilde{x}_0 , con el objetivo de evitar la integración numérica de trayectorias de estados, control y/o costo, y bajar el esfuerzo computacional.

La contribución principal del presente artículo viene de la reformulación del resultado teórico previo de [8]. Se muestra que en lugar de necesitar la matriz de penalización final \tilde{S} de $n \times n$ y las condiciones iniciales \tilde{x}_0 , es suficiente con encontrar el vector de estados y coestados finales (ρ, μ) de dimensión $2n$. Entonces, se consiguen nuevas fórmulas algebraicas para calcular las derivadas parciales con respecto a los nuevos parámetros. Desde el punto de vista práctico, buscar estos parámetros en lugar de los originales implica una reducción significativa en la dimensión de las incógnitas y por ende una disminución del esfuerzo computacional, que para el tratamiento de sistemas de gran dimensión puede resultar determinante en cuanto a la viabilidad de su instrumentación.

El esquema propuesto actualiza (ρ, μ) , y también los tiempos de “conmutación” τ_i donde se satura el control, mientras que el costo total es reducido vía el método del gradiente [10]. En general, se espera que la solución que produce el método sea sub-óptima, aunque en algunos casos, como el ilustrado aquí del “frenado óptimo de un tren”, es posible alcanzar la solución óptima numéricamente.

II. PROBLEMA LQR REGULAR

El problema LQR invariante en el tiempo, de horizonte finito, con condiciones finales libres y controles sin restricciones, minimiza el siguiente funcional cuadrático de costo

$$\mathcal{J}(u) = \int_0^{t_f} (x(t)'Qx(t) + u(t)'Ru(t))dt + x(t_f)'Sx(t_f), \quad (1)$$

con respecto a todas las trayectorias de control admisibles (aquí las continuas a tramos) $u: [0, t_f] \rightarrow \mathbb{R}^m$, sujeto a la restricción dinámica lineal

$$\dot{x} = Ax + Bu := f(x, u), \quad x(0) = x_0. \quad (2)$$

Este planteamiento se denotará como el problema $(A, B, Q, R, S, t_f, \mathbb{R}^m, x_0)$.

Las matrices (invariantes en el tiempo) de las Ecs. (1, 2), se asumen con las siguientes propiedades: Q y S de dimensión

V. Costanza, Grupo de Sistemas No Lineales, INTEC (CONICET-UNL), Santa Fe, Argentina, tsinoli@santafe-conicet.gov.ar

P. S. Rivadeneira, Grupo de Sistemas No Lineales, INTEC (CONICET-UNL), Santa Fe, Argentina, tsinoli@santafe-conicet.gov.ar

J. A. G. Múnera, Grupo de Sistemas No Lineales, INTEC (CONICET-UNL), Santa Fe, Argentina, tsinoli@santafe-conicet.gov.ar

$n \times n$ definidas semi-positivas, R de $m \times m$ y definida positiva, A de $n \times n$, B de $n \times m$, y el par (A, B) controlable. En este artículo, se mantendrá la dimensión de la entrada como $m = 1$. La generalización para mayores variables manipuladas, aunque no presenta obstáculos esenciales, es engorrosa y se deja para otro artículo. La expresión cuadrática bajo la integral es llamada usualmente el “Langrangiano” $L(x, u) := x'Qx + u'Ru$. Bajo estas consideraciones el Hamiltoniano del problema, definido como $H(x, \lambda, u) := L(x, u) + \lambda'f(x, u)$ es una función regular $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$, y la teoría exige que H se minimice de manera única con respecto a u para cada par (x, λ) , en cuyo caso el control toma la siguiente forma explícita

$$u^0(x, \lambda) = \frac{1}{2}R^{-1}B'\lambda. \quad (3)$$

(independiente de x), el cual es llamado “el control de H -mínimo”. Entonces, encontrar el control óptimo de un problema regular (ver [5-6]) requiere resolver las HCEs en el espacio de fases (x, λ) :

$$\dot{x} = Ax - \frac{1}{2}W\lambda; \quad x(0) = x_0. \quad (4)$$

$$\dot{\lambda} = -2Qx - A'\lambda; \quad \lambda(t_f) = 2Sx(t_f), \quad (5)$$

donde $W := BR^{-1}B'$. La solución de las ecuaciones HCEs para un sistema lineal e invariante resulta

$$\begin{pmatrix} x(t) \\ \lambda(t) \end{pmatrix} = e^{-\mathbf{H}(t_f-t)} \begin{pmatrix} x(t_f) \\ \lambda(t_f) \end{pmatrix}, \quad (6)$$

donde $\mathbf{H} = \begin{pmatrix} A & -W/2 \\ -2Q & -A' \end{pmatrix}$ está asociado a la matriz fundamental del sistema (6) $\mathbf{U}(t) := e^{-\mathbf{H}(t)}$. Dado que $x(t_f)$ y $\lambda(t_f)$ son generalmente desconocidos, un método alternativo para obtener la solución del problema regular sin restricciones, planteado anteriormente, se basa en la solución $P(\cdot)$ de la “Ecuación diferencial de Riccati” (RDE)

$$\dot{P} = PWP - PA - A'P - Q; \quad P(t_f) = S, \quad (7)$$

la cual establece una relación muy útil entre los estados óptimos $x^*(\cdot)$ y coestados $\lambda^*(\cdot)$ óptimos, concretamente, $\lambda^*(t) = 2P(t)x^*(t)$. De la Ec. (3), se puede obtener la trayectoria de control óptima

$$u^*(t) = u^0(x^*(t), \lambda^*(t)) = -R^{-1}B'P(t)x^*(t), \quad (8)$$

o equivalentemente la ley de control óptima por retroalimentación de estados (o feedback) es

$$u_f(t, x) = -R^{-1}B'P(t)x. \quad (9)$$

En resultados previos del problema LQR en la forma Hamiltoniana (ver [1, 5]) aparecen otros objetos matemáticos que serán útiles en la comprensión de la solución con restricciones, por ejemplo, las matrices $\alpha(t_f, S)$ y $\beta(t_f, S)$:

$$\begin{pmatrix} \alpha(t_f - t, S) \\ \beta(t_f - t, S) \end{pmatrix} \triangleq \mathbf{U}(t_f - t) \begin{pmatrix} I \\ 2S \end{pmatrix}. \quad (10)$$

Estas matrices auxiliares nos permiten calcular, para un problema LQR sin restricciones, la solución de la RDE

$P(\cdot, t_f, S) \forall t \in [0, t_f]$ a través de la fórmula algebraica [1]

$$P(t, t_f, S) = \frac{1}{2}\beta(t_f - t, S)[\alpha(t_f - t, S)]^{-1}. \quad (11)$$

Además, las matrices α y β están relacionadas a las condiciones de frontera de las HCEs a través de: $x(0) = \alpha(t_f, S)x(t_f)$, $\lambda(0) = \beta(t_f, S)x(t_f)$ y $\lambda(t_f) = 2Sx(t_f)$.

III. CONTROL CON RESTRICCIONES

Las características cualitativas de la solución óptima de problemas con controles acotados difieren sustancialmente de la solución del respectivo problema irrestricto [2]. En la teoría quedan aún preguntas abiertas como: cuánto realmente difieren dichas soluciones en el costo alcanzado, cuál clase de problemas generan soluciones “bang-bang”, o cuándo las soluciones restrictas son sólo saturaciones de trayectorias óptimas del mismo problema sin restricciones. En lo que sigue, la variable manipulada quedará restringida por $u(t) \in \mathbb{U} := [u_{\min}, u_{\max}]$.

La búsqueda de soluciones a los problemas restrictos frecuentemente conduce al Principio del Máximo de Pontryagin (PMP) [2]. Sin embargo, aún cuando sea computable, la solución dada por el PMP no es suficientemente flexible para tratar perturbaciones en los estados, es decir, el PMP genera sólo trayectorias de control en lazo abierto (no se consiguen soluciones del tipo feedback). En términos precisos, en este artículo se explota el siguiente resultado teórico [8]:

Asumir que existe un instante de tiempo $t \in (0, t_f)$ para el cual se verifica: $u_{x_0}^*(t) \in \mathbb{U} := [u_{\min}, u_{\max}]$. Entonces, existirá un intervalo de tiempo $\mathfrak{I} \subset (0, t_f)$ que contiene a t , y la trayectoria de fase óptima $\{x_{x_0}^*, \lambda_{x_0}^*\}$ del problema original $(A, B, Q, R, S, t_f, \mathbb{U}, x_0)$ coincidirá en \mathfrak{I} con la trayectoria $\{\hat{x}, \hat{\lambda}\}$, quien a su vez es óptima para cierto (en general distinto) problema $(A, B, Q, R, \hat{S}, t_f, \mathbb{R}, \hat{x}_0)$.

El problema $(A, B, Q, R, \hat{S}, t_f, \mathbb{R}, \hat{x}_0)$ no tiene restricciones sobre los valores de control ($\mathbb{U} \equiv \mathbb{R}$), y lo llamaremos el problema “oculto”, o el problema (\wedge). En lo que sigue, se asumirá que existe al menos un intervalo “regular” $(\tau_1, \tau_2) \subset (0, t_f)$ donde el control toma valores en (u_{\min}, u_{\max}) . La generación de estrategias de control sub-óptimas basadas en la aproximación de los parámetros desconocidos $\{\hat{S}, \hat{x}_0\}$ ha sido publicada recientemente, también en [8].

En [9] se desarrollaron fórmulas algebraicas para calcular las derivadas parciales del costo total con respecto a los coeficientes de la matriz \hat{S} y las condiciones iniciales \hat{x}_0 , sin necesidad de generar las trayectorias de estado ni evaluar el costo. En dicho caso, el número de variables a actualizar fue $n(n+3)/2$, que para aplicaciones de gran dimensión puede resultar poco práctico.

En este artículo se desarrollan fórmulas algebraicas para la nueva situación, evitando todavía la integración numérica de ODEs, y con un número menor de incógnitas que en los resultados presentados en [9]. Una consecuencia directa es la reducción del esfuerzo computacional necesario para minimizar el costo total.

A. Un hecho teórico relevante

Sea $(A, B, Q, R, \hat{S}, t_f, \mathbb{R}, \hat{x}_0)$ el problema sin restricción oculto, y sea $\{\hat{x}, \hat{\lambda}\}$ la trayectoria del estado/coestado óptima de dicho problema. Después de definir $\rho := \hat{x}(t_f)$ y $\mu := \hat{\lambda}(t_f)$, \tilde{S} se propone como

$$\tilde{S} := \frac{1}{2} \frac{\mu\mu'}{\rho'\mu} \quad (13)$$

(siempre que $\rho'\hat{S}\rho > 0$, si no $\tilde{S} := \hat{S}$).

Teorema 1. El problema $(A, B, Q, R, \tilde{S}, t_f, \mathbb{R}, \hat{x}_0)$ tiene las mismas trayectorias óptimas del estado y coestado que el problema $(A, B, Q, R, \hat{S}, t_f, \mathbb{R}, \hat{x}_0)$, o sea $\{\hat{x}, \hat{\lambda}\}$.

Para probarlo, denotaremos a la trayectoria $\{\tilde{x}, \tilde{\lambda}\}$ como la solución a las HCEs de cualquiera de los problemas (dichas ODEs no dependen de \tilde{S} o \hat{S}) con condiciones de frontera $\tilde{x}(t_f) = \rho$ y $\tilde{\lambda}(t_f) = 2\tilde{S}\rho = \mu$. Asimismo nótese

$$\tilde{\lambda}(t_f) = 2\tilde{S}\rho = \mu = 2\hat{S}\rho, \quad (14)$$

y por lo tanto $\{\tilde{x}, \tilde{\lambda}\}$ es la trayectoria óptima del problema $(A, B, Q, R, \hat{S}, t_f, \mathbb{R}, \hat{x}_0)$, es decir

$$\{\tilde{x}, \tilde{\lambda}\} \equiv \{\hat{x}, \hat{\lambda}\}. \quad (15)$$

Pero ya que ambos pares son soluciones de las ecuaciones Hamiltonianas, ellas deben también satisfacer $\tilde{\lambda} = 2\tilde{P}\tilde{x}$, $\hat{\lambda} = 2\hat{P}\hat{x}$ con \tilde{P} y \hat{P} dos soluciones de la RDE (7) con condiciones finales \tilde{S} y \hat{S} , respectivamente. Por otra parte la solución x de

$$\dot{x} = (A - W\tilde{P})x, \quad x(0) = \hat{x}_0 \quad (16)$$

es óptima para el problema $(A, B, Q, R, \tilde{S}, t_f, \mathbb{R}, \hat{x}_0)$. Ahora,

definiendo $\delta := x - \hat{x}$, se obtiene

$$\begin{aligned} \dot{\delta} &= A\delta - W(\tilde{P}x - \hat{P}\hat{x}) = A\delta - W(\tilde{P}x - \tilde{P}\hat{x}) \\ &= A\delta - W(\tilde{P}x - \tilde{P}\hat{x}) = (A - W\tilde{P})\delta. \end{aligned} \quad (17)$$

Dado que $\delta(0) = 0$, esto significa que $x \equiv \hat{x} = \tilde{x}$. \square

Corolario 1: $\tilde{P}\hat{x} = \tilde{P}\tilde{x} = \hat{P}\hat{x}$, y por lo tanto el feedback óptimo (9) es el mismo para los problemas \wedge y \sim . \square

B. Fórmulas algebraicas

B1. Objetos auxiliares: Los siguientes controles serán utilizados a menudo en las próximas subsecciones para obtener el control sub-óptimo del problema

$$\tilde{u}(t) := \begin{cases} u_{\min}, \forall t \in [0, \tau_1) \\ -R^{-1}B'P(t, \rho, \mu)x(t), \forall t \in [\tau_1, \tau_2), \\ u_{\max}, \forall t \in [\tau_2, t_f] \end{cases} \quad (18)$$

donde $\tilde{u}(t)$ es una notación simplificada para $\tilde{u}_{\rho, \mu, \tau_1, \tau_2}(t)$ (el feedback está asociado a los parámetros $(\rho, \mu, \tau_1, \tau_2)$). La nueva estrategia propone variaciones de las variables ρ y μ para construir la matriz \tilde{S} , teniendo en cuenta las siguientes consideraciones:

(i) La matriz inicial \tilde{S}_0 será construida a partir de los valores finales de las trayectorias óptimas del estado y coestado del problema irrestricto (1)-(2), aplicando el feedback (9). El coestado verifica en t_f que $\lambda(t_f) = 2P(t_f)x(t_f) = 2Sx(t_f)$, y entonces denotando $\rho_{seed} := x(t_f)$ y $\mu_{seed} := \lambda(t_f)$, la matriz semilla \tilde{S}_0 se puede calcular como

$$\tilde{S}_0 := \frac{1}{2} \frac{\mu_{seed}\mu'_{seed}}{\rho'_{seed}\mu_{seed}} \quad (19)$$

(ii) Se adopta la siguiente trayectoria “semilla” para iniciar el método iterativo

$$u_{seed}(t) := \begin{cases} u_{\min} & \text{si } -R^{-1}B'P(t, \tilde{S}_0)x(t) \leq u_{\min} \\ u_{\max} & \text{si } -R^{-1}B'P(t, \tilde{S}_0)x(t) \geq u_{\max} \\ -R^{-1}B'P(t, \tilde{S}_0)x(t) & \text{otro caso.} \end{cases} \quad (20)$$

La trayectoria de estados correspondiente al control u_{seed} , empezando en x_0 , es decir $x_{u_{seed}}$, se denotará como x_{seed} . Nótese que a partir del control semilla y de las trayectorias de estado, simuladas para \tilde{S}_0 , se pueden detectar, en caso de que existan, los primeros valores de los tiempos de conmutación denotados como $\tau_{1,0} \leq \tau_{2,0}$.

(iii) En lo que sigue $P(t, \rho, \mu)$ denotará la solución de la RDE (7), con condición final $P(t_f) = \tilde{S}$. Cuando sea claro en el texto la dependencia de P con ρ y μ , se utilizará la notación simplificada $P(t)$. De la Ec. (11) se obtiene la siguiente identidad, que será de utilidad para $\eta := \rho$ ó μ

$$\begin{aligned} \frac{\partial P(t, \eta)}{\partial \eta} &= \frac{\partial \left[\frac{1}{2} \beta(t_f - t, \eta) [\alpha(t_f - t, \eta)]^{-1} \right]}{\partial \eta} \\ &= \frac{1}{2} [\beta_\eta \alpha^{-1} - \beta \alpha^{-1} \alpha_\eta \alpha^{-1}] (t_f - t, \eta) \\ &= \frac{1}{2} [\beta_\eta - 2P(t, \eta) \alpha_\eta] \alpha^{-1}, \end{aligned} \quad (21)$$

donde α , $\alpha_\eta := \frac{\partial \alpha}{\partial \eta}$, y $\beta_\eta := \frac{\partial \beta}{\partial \eta}$ deben ser evaluadas en $(t_f - t, \rho, \mu)$. Ya que la matriz $\mathbf{U}(t)$ es $2n \times 2n$, por conveniencia será particionada en 4 sub-matrices de $n \times n$

$$\mathbf{U}(t) := \begin{pmatrix} \mathbf{U}_1(t) & \mathbf{U}_2(t) \\ \mathbf{U}_3(t) & \mathbf{U}_4(t) \end{pmatrix}. \quad (22)$$

Además, a partir de la Ec. (10) las derivadas parciales α_η y β_η resultan

$$\alpha_\eta = 2\mathbf{U}_2 \frac{\partial S}{\partial \eta}, \quad \beta_\eta = 2\mathbf{U}_4 \frac{\partial S}{\partial \eta}. \quad (23)$$

(iv) La matriz fundamental “saturada” es

$$\psi(t, \tau) := \int_{\tau}^t e^{A(t-\sigma)} d\sigma = e^{At} \int_{\tau}^t e^{-A\sigma} d\sigma, \quad (24)$$

y sus matrices relacionadas

$$\check{\psi}(t, \tau) := \int_{\tau}^t e^{A'(\sigma-\tau)} Q e^{A(\sigma-\tau)} d\sigma, \quad (25)$$

$$\hat{\psi}(t, \tau) := \int_{\tau}^t \psi'(\sigma, \tau) Q e^{A(\sigma-\tau)} d\sigma, \quad (26)$$

también serán necesarias en lo que sigue. En conjunto con la matriz $\mathbf{U}(t)$, estas matrices son calculadas, interpoladas off-line y guardadas en memoria. Entonces se asume que estarán disponibles como funciones de dos variables (t, τ) , en el rango $[0, t_f] \times [0, t_f]$.

Cuando $u(t) \equiv u_{min}$ en $[0, \tau_1]$, el estado en τ_1 resulta

$$x(\tau_1) = e^{A\tau_1}x_0 + \psi(\tau_1, 0)Bu_{min}. \quad (27)$$

(v) En este punto es importante notar que el estado en τ_2 podría ser calculado utilizando el flujo Hamiltoniano. Ya que para cada $t \in [\tau_1, \tau_2]$ el control es $-R^{-1}B'P(t, \rho, \mu)x(t)$ y el coestado (correspondiente a este subconjunto de trayectoria regular), denotado como $\tilde{\lambda}(t)$, es $\tilde{\lambda}(t) = 2P(t, \rho, \mu)x(t)$, y

$$\begin{pmatrix} x(t) \\ \tilde{\lambda}(t) \end{pmatrix} = \mathbf{U}(\tau_1 - t) \begin{pmatrix} x(\tau_1) \\ \tilde{\lambda}(\tau_1) \end{pmatrix}, \quad (28)$$

el estado finalmente resulta en

$$\begin{aligned} x(\tau_2) &= \mathbf{U}_1(\tau_1 - \tau_2)x(\tau_1) + \mathbf{U}_2(\tau_1 - \tau_2)\tilde{\lambda}(\tau_1) \\ &= (\mathbf{U}_1(\tau_1 - \tau_2) + 2\mathbf{U}_2(\tau_1 - \tau_2)P(\tau_1, \eta))x(\tau_1). \end{aligned} \quad (29)$$

En lo que sigue el argumento t de $\mathbf{U}(t)$ será omitido cuando el contexto sea claro y lo permita. Cuando $u(t) \equiv u_{max}$ en $[\tau_2, t_f]$, el estado en t_f resultará

$$x(t_f) = e^{A(t_f - \tau_2)}x(\tau_2) + \psi(t_f, \tau_2)Bu_{max}. \quad (30)$$

B2. Derivadas parciales del costo: Se sabe de [11], que el costo total $J(\tilde{u})$ es diferenciable con respecto a las variables (η, τ_1, τ_2) . Para obtener las derivadas, por conveniencia el costo total será dividido en 4 partes:

$$J(\rho, \mu, \tau_1, \tau_2) := J(\tilde{u}) = J_1 + J_2 + J_3 + J_4, \quad (31)$$

entre $[0, \tau_1]$, $[\tau_1, \tau_2]$, $(\tau_2, t_f]$ y en t_f , respectivamente, y entonces utilizando resultados clásicos [5],

$$J_1 = Ru_{min}^2\tau_1 + \int_0^{\tau_1} x'(t)Qx(t)dt, \quad (32)$$

$$J_2 = x'(\tau_1)P(\tau_1)x(\tau_1) - x'(\tau_2)P(\tau_2)x(\tau_2), \quad (33)$$

$$J_3 = Ru_{max}^2(t_f - \tau_2) + \int_{\tau_2}^{t_f} x'(t)Qx(t)dt, \quad (34)$$

$$J_4 = x'(t_f)Sx(t_f). \quad (35)$$

Las derivadas D_η de J_1 y J_2 con respecto a las variables ρ y μ (denotadas indistintamente por la variable η) resultan:

$$D_\eta J_1 = 0, D_\eta J_2 = x'(\tau_1) \left[\frac{\partial P(\tau_1)}{\partial \eta} \dots \right. \quad (36)$$

$$\left. -4(\mathbf{U}_1 + 2\mathbf{U}_2P(\tau_1))'P(\tau_2)\mathbf{U}_2 \frac{\partial P(\tau_1)}{\partial \eta} - \dots \right.$$

$$\left. (\mathbf{U}_1 + 2\mathbf{U}_2P(\tau_1))' \frac{\partial P(\tau_2)}{\partial \eta} (\mathbf{U}_1 + 2\mathbf{U}_2P(\tau_1)) \right] x(\tau_1),$$

las derivadas parciales de P se calculan a partir de las Ecs. (21-23), y

$$\frac{\partial \tilde{S}_{kl}}{\partial \rho_i} = \frac{1}{2}(\mu_k \mu_l) \left(-\frac{\mu_i}{(\rho' \mu)^2} \right), \quad (37)$$

$$\frac{\partial \tilde{S}_{kl}}{\partial \mu_j} = \frac{1}{2} \left[\frac{Z_{j(kl)}}{(\rho' \mu)} - \frac{(\mu_k \mu_l) \rho_j}{(\rho' \mu)^2} \right], \quad (38)$$

donde

$$Z_{j(kl)} := \frac{\partial \mu_k \mu_l}{\partial \mu_j} = \begin{cases} 2\mu_j & \text{si } k = l = j \\ \mu_k & \text{si } l = j \text{ y } k \neq j \\ \mu_l & \text{si } k = j \text{ y } l \neq j \\ 0 & \text{si } k \neq j \text{ y } l \neq j \end{cases} \quad (39)$$

Las derivadas de J_3 y J_4 son

$$D_\eta J_3 = \int_{\tau_2}^{t_f} 2x'(t)Q \frac{\partial x(t)}{\partial \eta} dt = 4[x'(\tau_2)\check{\psi}(t_f, \tau_2) + \dots \quad (40)$$

$$u_{max}B'\hat{\psi}(t_f, \tau_2)]\mathbf{U}_2(\tau_1 - \tau_2) \frac{\partial P(\tau_1, \eta)}{\partial \eta} x(\tau_1),$$

después de reemplazar $x'(t) = x'(\tau_2)e^{A(t-\tau_2)} + B'\psi'(t, \tau_2)u_{max}$, y expandiendo $\frac{\partial x(t)}{\partial \eta} = e^{A(t-\tau_2)} \frac{\partial x(\tau_2)}{\partial \eta} = e^{A(t-\tau_2)} 2\mathbf{U}_2(\tau_1 - \tau_2) \frac{\partial P(\tau_1, \rho, \mu)}{\partial \eta} x(\tau_1)$. Finalmente,

$$D_\eta J_4 = 2x'(t_f)S \frac{\partial x(t_f)}{\partial \eta} = \quad (41)$$

$$= 4x'(t_f)Se^{A(t_f - \tau_2)}\mathbf{U}_2(\tau_1 - \tau_2) \frac{\partial P(\tau_1, \eta)}{\partial \eta} x(\tau_1).$$

Las derivadas con respecto a los tiempos de conmutación τ_1 y τ_2 se mantienen sin cambios con respecto a las presentadas en [9]

$$D_{\tau_1} J_1 + D_{\tau_1} J_2 = x'(\tau_1)[2P(\tau_1)Bu_{min} + \dots \quad (42)$$

$$\dots + P(\tau_1)WP(\tau_1)]x(\tau_1) + Ru_{min}^2,$$

$$D_{\tau_1} J_3 = \int_{\tau_2}^{t_f} 2x'(t)Qe^{A(t-\tau_2)}dt \frac{\partial x(\tau_2)}{\partial \tau_1}, \quad (43)$$

$$D_{\tau_1} J_4 = 2x'(t_f)Se^{A(t_f - \tau_2)} \frac{\partial x(\tau_2)}{\partial \tau_1}, \quad (44)$$

donde $\frac{\partial x(\tau_2)}{\partial \tau_1} = \mathbf{U}_1(\tau_1 - \tau_2)(Ax(\tau_1) + Bu_{min}) -$

$$2\mathbf{U}_2(\tau_1 - \tau_2)(A'P + Q)x(\tau_1) - (A - WP(\tau_2))x(\tau_2).$$

Similarmente, las derivadas con respecto a τ_2 son $D_{\tau_2} J_1 = 0$, y

$$D_{\tau_2} J_2 + D_{\tau_2} J_3 = -x'(\tau_2)[2P(\tau_2)Bu_{max} + \dots \quad (45)$$

$$\dots + P(\tau_2)WP(\tau_2)]x(\tau_2) - Ru_{max}^2,$$

$$D_{\tau_2} J_4 = -2x'(t_f)Se^{A(t_f - \tau_2)}(WP(\tau_2)x(\tau_2) + Bu_{max}) \quad (46)$$

B3. Actualizando los parámetros: las primeras aproximaciones de los puntos de saturación óptimos denotadas $\tau_{1,0}, \tau_{2,0}$, se calculan después de simular la trayectoria de estado x_{seed} a partir de la trayectoria semilla de control

$$\tilde{u}_0(t) \equiv u_{seed}(t). \quad (47)$$

Los parámetros $(\rho, \mu, \tau_1, \tau_2)$ son actualizados para construir sucesivas estrategias de control \tilde{u}_j ; $j = 1, 2 \dots$

$$\mathcal{J}(\tilde{u}_{j+1}) \leq \mathcal{J}(\tilde{u}_j) \leq \dots \leq \mathcal{J}(\tilde{u}_{seed}) ; j = 1, 2 \dots \quad (48)$$

de acuerdo al método del gradiente:

$$\rho_j = \rho_{j-1} - \gamma_j \frac{\partial \mathcal{J}}{\partial \rho}(\rho_{j-1}, \mu_j, \tau_{1,j}, \tau_{2,j}) \quad (49)$$

$$\mu_j = \mu_{j-1} - \gamma_j \frac{\partial \mathcal{J}}{\partial \mu}(\rho_j, \mu_{j-1}, \tau_{1,j}, \tau_{2,j})$$

$$\tau_{i,j} = \tau_{i,j-1} - \gamma_j \frac{\partial \mathcal{J}}{\partial \tau_i}(\rho_j, \mu_j, \tau_{i,j-1}), i = 1, 2.$$

hasta que se alcanza la convergencia o se detiene el proceso por alguna anomalía o incremento en el costo.

IV. EL CASO DE ESTUDIO

El caso de estudio de este artículo es el “frenado óptimo de un tren”, un problema LQR, que se define a través de los siguientes objetos:

$$\begin{aligned} \dot{x}_1(t) &= x_2(t); & x_1(0) &= 1, \\ \dot{x}_2(t) &= u(t); & x_2(0) &= -1, \end{aligned} \quad (50)$$

$u(t) \in [0,3]$; $0 \leq t \leq t_f = 1$,
 $Q = 10$, $R = 0.5$ y $S = 100$. La solución óptima ya fue hallada analíticamente en [8]. El resultado de la actualización de los parámetros $(\rho, \mu, \tau_1, \tau_2)$ utilizando el método del gradiente se muestra en las Figs. 1-4. En la Fig. 1 se grafican las trayectorias semilla y óptima. En la Fig. 2 se muestra la reducción del costo total contra el número de iteraciones hasta obtener un valor mínimo, $J^* = 15.1632$.

Se realizaron dos simulaciones para estudiar el comportamiento del parámetro γ del método del gradiente: en la primera, el valor de γ fue mantenido constante, $\gamma = 0.005$ (línea negra a rayas en Fig. 2), y en la segunda, fue actualizado cuando resultó necesario para lograr que \mathcal{J} siguiera decreciendo. Es claro que al principio del método γ puede tener un valor grande, siempre y cuando el costo siga disminuyendo. Sin embargo, en alguna iteración este valor grande de γ genera grandes gradientes y por lo tanto podría “saltar” el valor óptimo, interrumpiendo la minimización del costo. En este momento γ tiene que ser reducido. Para este caso de estudio, el primer valor fue puesto en $\gamma = 0.01$, y el último valor alcanzado fue $\gamma = 0.005$ (línea sólida en Fig 2). De esta manera, cambiando el parámetro γ , se muestra que se obtiene un mejor desempeño del método.

En las Figs. 3 y 4 se muestra la evolución de ρ , μ , τ_1 y τ_2 respectivamente. Luego de 700 iteraciones los parámetros convergen a sus valores óptimos. Los valores finales fueron: $\tau_1 = 0.5753$, $\tau_2 = 0.7764$, $\tilde{S}_{11} = 456.13$, $\tilde{S}_{12} = -134.475$, y $\tilde{S}_{22} = 39.6015$, calculados utilizando la Ec. (13) y el valor de ρ y μ que se indican en la Fig. 4. Se debe destacar que en este ejemplo el método que actualiza ρ y μ llega a la solución óptima, lo que no se consigue con el método de S, x_0 .

También se efectuó una comparación del tiempo de cómputo necesario por el método descrito en [9], y el método de este artículo que actualiza ρ y μ .

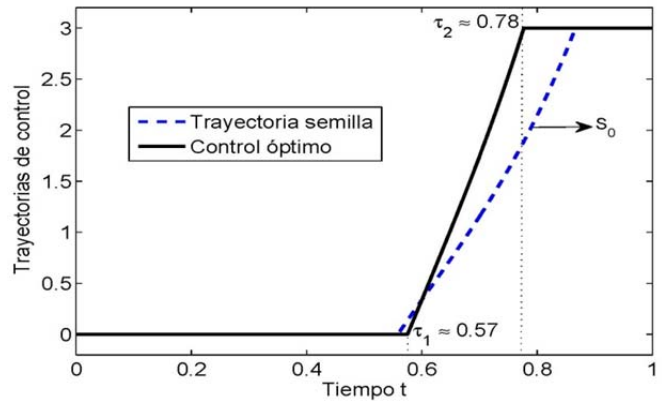


Figura 1. Estrategia de control de la aplicación del método del gradiente.

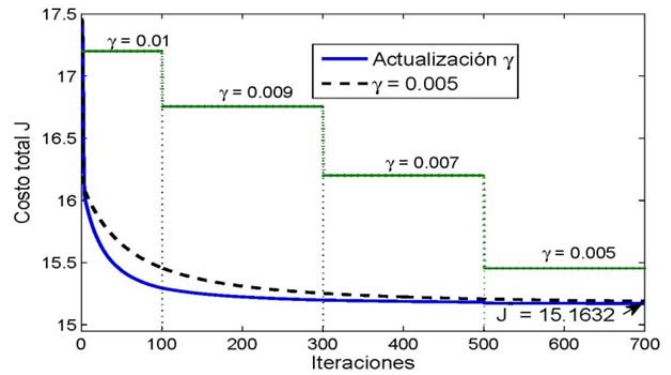


Figura 2. Evolución del costo total.

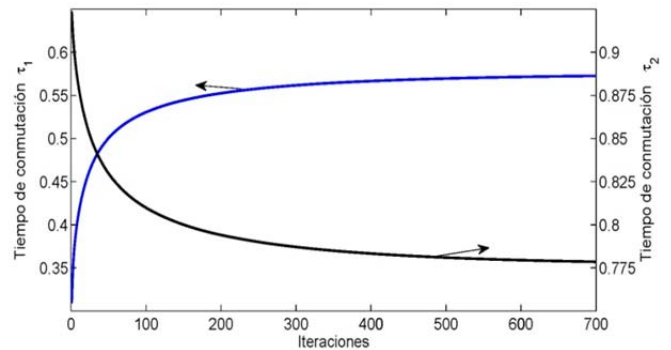


Figura 3. Evolución de los tiempos de conmutación τ_1 y τ_2 .

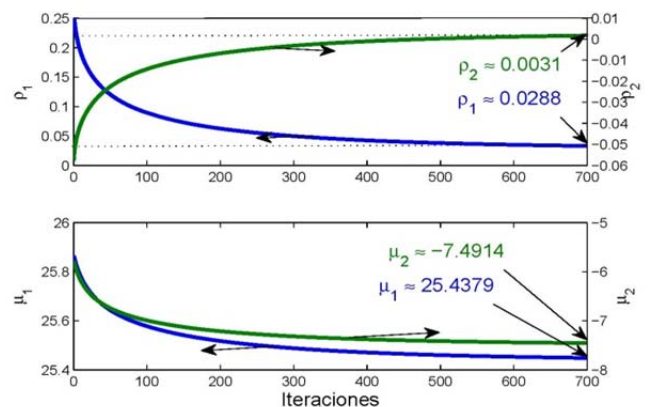


Figura 4. Evolución de ρ y μ .

Se obtuvieron tiempos de cómputo del mismo orden para los dos métodos, lo cual no es sorprendente ya que el sistema es de dimensión 2. Las variables actualizadas en el método S, x_0 fueron 3, mientras que para el método con ρ, μ son 4. Para $n = 10$ las variables a actualizar por el primer método serían 55 mientras que en el segundo se reducirían a 20. Sin embargo, sigue siendo un problema computacional difícil de resolver on-line, por lo que se está considerando el uso de computación en paralelo para los sistemas de gran dimensión [12].

V. CONCLUSIONES

Aunque los controles del tipo feedback son preferidos en la práctica cuando se espera que haya perturbaciones, un control en lazo cerrado es en general sub-óptimo cuando se deben respetar restricciones en la variable manipulada. Tomando estas limitaciones en cuenta, se desarrolló un algoritmo, eficiente para aproximar el control óptimo de lazo abierto vía feedback, basado en resultados teóricos recientes. Las estrategias resultantes difieren bastante de la forma del control óptimo saturado correspondiente al problema irrestricto con los mismos parámetros y condiciones iniciales, la cual es usada por el algoritmo como la primera aproximación, o trayectoria “semilla”.

La idea principal del trabajo consiste en obtener una reducción total del costo utilizando fórmulas algebraicas a través de la actualización de los parámetros (ρ, μ) y los tiempos de conmutación τ_i . La estabilidad del método está garantizada ya que (i) al costo no se le permite aumentar, y (ii) está acotado inferiormente.

Algunas características positivas del nuevo método iterativo propuesto son:

- El método está basado en resultados teóricos que aseguran que (ρ, μ) y τ_1, τ_2 del problema “oculto” sean los parámetros críticos a ser determinados.
- Calcular la matriz S en términos de (ρ, μ) reduce la cantidad de parámetros a actualizar. En el método de [9], la cantidad mínima de variables a actualizar es $\frac{n(n+1)}{2}$, mientras que el método descrito aquí se reducen a $2n$, una ganancia significativa para sistema de gran dimensión.
- Se toma ventaja de la disponibilidad de (α, β) como funciones de $(t_f - t, \tilde{S})$, y consecuentemente en la posibilidad de generar las matrices de Riccati $P(t, t_f, \rho, \mu)$ sólo por operaciones algebraicas mientras se actualizan (ρ, μ) .
- La actualización de parámetros $(\rho, \mu, \tau_1, \tau_2)$ se hace por el método del gradiente y gracias a las fórmulas algebraicas desarrolladas. No se utiliza la integración numérica de ODEs para la predicción de trayectorias de estado, coestado, y costo.

La velocidad de cómputo es el asunto central cuando se piensa en aumentar la complejidad de las estrategias de control, sobre todo debido a los requerimientos de tiempo real. El incremento de variables a actualizar en sistemas de gran dimensión genera la necesidad de cambiar los paradigmas

actuales de computación. Es por eso que se analiza la instrumentación del algoritmo aquí propuesto mediante herramientas de computación en paralelo, como una alternativa promisoría, para alcanzar el control óptimo en tiempo real.

REFERENCIAS

- [1] V. Costanza y P.S. Rivadeneira, “Enfoque Hamiltoniano al control óptimo de sistemas dinámicos”. Saarbrücken, Alemania: OmniScriptum, 2014.
- [2] L.S. Pontryagin, V.G. Boltyanski, R.V. Gamkrelidze y E.F. Mishchenko. The Mathematical Theory of Optimal Processes. New York, USA: Macmillan, 1964.
- [3] J.L. Troutman, Variational Calculus and Optimal Control. New York, USA: Springer, 1996.
- [4] V. Costanza y C. E. Neuman, “Partial Differential Equations for Missing Boundary Conditions in the Linear-Quadratic Optimal Control Problem”, Latin American Applied Research, vol. 39, pp. 207-212, 2009.
- [5] E.D. Sontag, Mathematical Control Theory. 2nd ed. New York, USA: Springer, 1998.
- [6] M. Athans y P.L. Falb, Optimal Control: An Introduction to the Theory and its Applications. New York, USA: Dover, 2006.
- [7] J. L. Speyer y D. H. Jacobson. Primer on Optimal Control Theory. Philadelphia, USA: SIAM Books, 2010.
- [8] V. Costanza y P.S. Rivadeneira, “Optimal saturated feedback laws for LQR problems with bounded controls”, Computational and Applied Mathematics, vol. 32, pp. 355-371, 2013.
- [9] V. Costanza y P.S. Rivadeneira, “Online suboptimal control of linearized models”, System Science & Control Engineering, vol. 2, pp. 379-388, 2014.
- [10] P. Pardalos, R. Pytlak. “Conjugate Gradient Algorithms In Nonconvex Optimization. New York, USA: Springer, 2008.
- [11] V. Dhano y F. Tröltzsch, “Some aspects of reachability for parabolic boundary control problems with control constraints”, Computational Optimization and Applications, vol. 50, pp. 75-110, 2011.
- [12] J. A. Gómez Múnera y V. Costanza, “Computación en paralelo para control óptimo basado en modelos”, AAECA, Buenos Aires, Argentina, Oct. 27-29, 2014.



Vicente Costanza was born in Rosario, Argentina in 1950. He studied Chemical Engineering at the Universidad Tecnológica Nacional, in Rosario, until 1973; and then at Princeton University, USA, where he obtained his Ph.D. degree on 1980. After a postdoctoral stay at Caltech, Pasadena, he returned to Argentina, where he initiated research work on Control of Nonlinear Systems and graduate teaching in subjects related to Applied Mathematics at INTEC, Santa Fe, where he holds a full professorship. He has been a visiting professor at ETH-Zürich and EPF-Lausanne, at Universidad de Buenos Aires, and other institutions. In 2008 he was in sabbatical leave at the Centro de Matemática Aplicada, Universidad Nacional de San Martín, Argentina as an invited professor. He is at present Investigador Principal at Conicet, Argentina.



Pablo S. Rivadeneira was born in Ipiales, Colombia in 1981. He received the Control Engineering degree from the Universidad Nacional de Colombia, at Medellín, in 2005, and the doctorate in Chemical Technology degree from the Universidad del Litoral, at Santa Fe, Argentina in 2010. After that, he did a postdoctoral stage at IRRCyN, Nantes, France during 2010-2012. From 2013, he is Investigador Asistente at INTEC –CONICET, Santa Fe, Argentina. His main research interests include aspects of nonlinear systems and control, with applications to industrial and biological processes.



John A. Gómez Múnera was born in Medellín, Colombia in 1988. He received the Electronic Engineering degree from the Universidad de Antioquia, at Medellín, in 2011. He is current student of the doctorate in Chemical Technology degree from the Universidad del Litoral, at Santa Fe, Argentina. His main research interests include aspects of nonlinear systems, optimal control and parallel computing.