



**Universidad Nacional de La Plata**  
**Facultad de Ingeniería**  
**Departamento de Electrotecnia**  
**Cátedra de Control Moderno**

---

# **Introducción al control óptimo**

**Ricardo Julián Mantz**

**Año 2003**

# 1. Introducción.

En estas notas se introducen ideas básicas de la Teoría del Control Óptimo. Algunas de éstas presentan diferencias significativas con respecto a las previamente vistas y resulta natural preguntarse:

*¿Por qué es necesario plantear un nuevo enfoque para el diseño de controladores?*

Hemos visto que los métodos de control basados en realimentación de estados son más versátiles y potentes que los correspondientes a la teoría del control clásico. En particular, permiten la asignación de los autovalores de lazo cerrado a ‘voluntad’. También se ha discutido como estos métodos pueden ser complementados con técnicas de observación de estados para sobrellevar limitaciones (económicas y/o físicas) en la medida de los mismos. Luego, es razonable cuestionarse acerca de la necesidad o conveniencia de introducir un nuevo marco teórico, de hecho más complejo, para el diseño de controladores. Con respecto a este posible cuestionamiento puede argumentarse que:

- Ni en los métodos de diseño de control clásico, ni en los de realimentación de estados para la asignación de autovalores, se pone explícitamente de manifiesto el compromiso que existe entre las especificaciones dinámicas y el ‘costo’ para poder cumplirlas (por ejemplo, entre la velocidad de respuesta de la variable controlada y la acción de control necesaria).
- Restricciones en el control<sup>1</sup> pueden imposibilitar un diseño de polos dominantes. Esto puede dificultar seriamente la selección de los autovalores de lazo cerrado para cumplir determinadas especificaciones temporales (tener presente que las conocidas expresiones que vinculan el sobrepaso, el tiempo de establecimiento, etc. con la ubicación de los polos, sólo tienen validez en sistemas de segundo orden puro). Adicionalmente, si existe un espacio no controlable, sólo algunos autovalores podrán ser asignados, y por consiguiente, difícilmente pueda hacerse un diseño con polos dominantes.
- La presencia de ceros en la función de transferencia pueden dificultar el diseño por asignación de autovalores, aún en aquellos casos en que las restricciones no sean fuertes. *Ver ejemplo 1*
- En sistemas MIMO no existe una correspondencia entre la respuesta temporal y la localización de los polos. Efectivamente, en sistemas MIMO, la misma asignación de polos puede hacerse con distintos juegos de ganancias, y de hecho, dan lugar a distintas respuestas temporales. *Ver ejemplo 2*
- Por otra parte, uno podría preguntarse, ¿porqué conformarse con ganancias de realimentación constantes?. ¿No podría mejorarse la respuesta de seguimiento de un sistema si la(s) ganancia(s) de realimentación se asocia(n) a la amplitud del error?.

Los argumentos previos, entre otros, sirvieron de motivación para el desarrollo de la teoría del control óptimo.

---

<sup>1</sup>todo sistema de control está sujeto a restricciones, por ejemplo, debidas al actuador de potencia, o a razones de seguridad, o debidas al ancho de banda necesario para rechazar determinadas perturbaciones, etc.

*Ejemplo 1.* Considere los siguientes sistemas:

- sistema 1

$$\begin{aligned}\dot{x} &= \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \\ y &= \begin{bmatrix} 1 & 0 \end{bmatrix} x\end{aligned}\quad (1)$$

- sistema 2

$$\begin{aligned}\dot{x} &= \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \\ y &= \begin{bmatrix} 1 & 2 \end{bmatrix} x.\end{aligned}\quad (2)$$

El sistema 2, solo difiere del 1, en que presenta un cero en  $-0,5$ . Se diseña un control por realimentación de estados que permite la misma asignación de autovalores de lazo cerrado en  $-0,7 \pm j 0,715$  (esto es posible realimentando los estados con un juego de ganancias  $[-0,4, -1]$ ). La figura 1, muestra la salida de los dos sistemas con la misma asignación de polos. Claramente, se distingue el efecto del cero del sistema 2, y por consiguiente la pérdida de correspondencia entre la ubicación de los polos de lazo cerrado y la respuesta temporal.

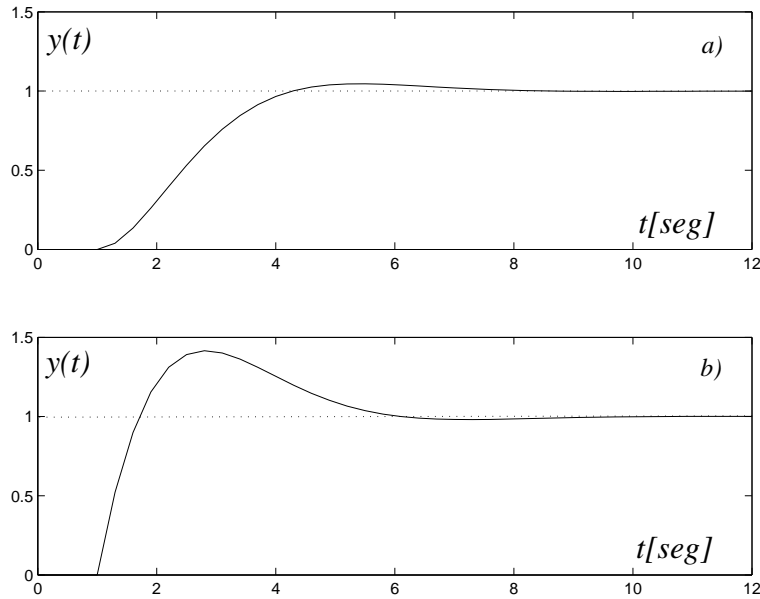


Figura 1: Respuesta de los sistemas 1 y 2 (partes a y b respectivamente), con la misma asignación de autovalores de lazo cerrado.

*Ejemplo 2.* Considere el control de un reactor (Kwakernaak & Sivan 1972), donde se pretende controlar volumen y concentración de una mezcla. El modelo de estados del sistema MIMO es:

$$\dot{x} = \begin{bmatrix} -0,01 & 0 \\ 0 & -0,02 \end{bmatrix} x + \begin{bmatrix} 1 & 1 \\ -0,25 & 0,75 \end{bmatrix} u, \quad (3)$$

siendo los estados las variables controladas (apartamientos del volumen y de la concentración con respecto a los valores de equilibrio).

Fácilmente puede verificarse que cualquiera de los siguientes juegos de ganancias de realimentación

$$\begin{aligned} K_a &= \begin{bmatrix} 1,1 & 3,7 \\ 0 & 0 \end{bmatrix} \\ K_b &= \begin{bmatrix} 0 & 0 \\ 1,1 & -1,233 \end{bmatrix} \\ K_c &= \begin{bmatrix} 0,1 & 0 \\ 0 & 0,1 \end{bmatrix} \end{aligned} \quad (4)$$

conducen a la misma asignación de polos en  $-0,1 \pm j 0,049$ . Sin embargo, como puede verse en la figura 2, las respuestas del sistema son completamente distintas (los transitorios corresponden a condiciones iniciales  $[0, 0, 1]^T$ ). Las curvas ponen en evidencia la poca practicidad de las técnicas de asignación de polos en sistemas MIMO.

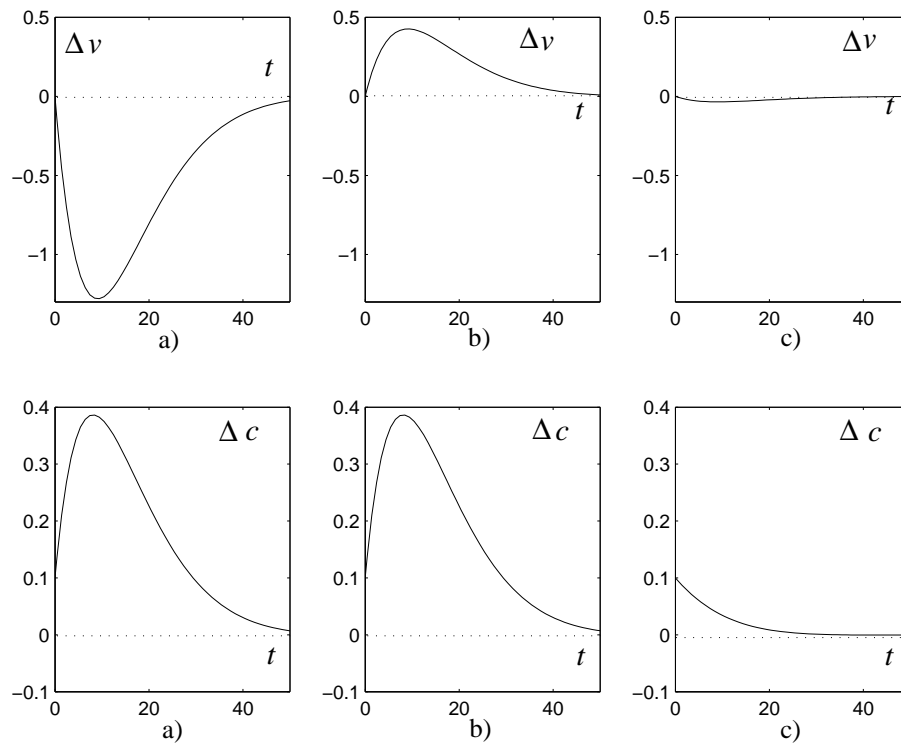


Figura 2: Variables controladas (volumen y concentración) a partir de la realimentación de los estados con los juegos de ganancias  $K_a$ ,  $K_b$  y  $K_c$ .

## 2. Significado de la palabra óptimo en el contexto de la teoría del control de sistemas.

Cuando se habla de la solución óptima de un problema, intuitivamente se piensa en que ésta es ‘la mejor solución’, es decir ‘insuperable’. De hecho, éste es el significado que puede encontrarse en el Diccionario de la Real Academia Española:

**óptimo:** forma procedente del superlativo latino *optimus*, que significa ‘bueno en grado sumo’, que no puede ser mejor. Por tanto, es incorrecto su empleo en

combinación con muy, más, menos o tan: *\*muy óptimo, \*más óptimo; \*menos óptimo, \*tan óptimo.*

Sin embargo, como muchos otros adjetivos, la palabra óptimo tiene un alto grado de subjetividad. Efectivamente, un pésimo control desde el punto de vista del comportamiento dinámico podría ser óptimo desde el punto de vista económico y viceversa. Luego, para calificar la bondad de un control (en particular para poder decir que es óptimo) es necesario asociarlo a un ‘índice’ de performance. En términos de control diremos que un control es óptimo si minimiza un funcional de costo en el que claramente se manifiesta un compromiso entre distintas especificaciones y restricciones. A este funcional lo llamaremos índice de performance y normalmente lo indicaremos con  $J$ . Obviamente, el mismo control evaluado con otro índice de performance  $J_1$  no será óptimo.

### 3. Índice de performance.

Considere la Figura 3. Tanto la parte a) como la parte b) de la figura muestran la variable controlada de un determinado proceso en dos casos diferentes, cuando es controlada empleando los controladores  $C_1$  y  $C_2$  respectivamente. En ambos casos se produce una perturbación en  $t = 12\text{seg}$ , la cual es rechazada con los transitorios mostrados. Aceptemos que la calidad de un determinado producto depende de esta variable. Luego, no parece desacertado pensar que la bondad de dicho producto en  $t_0$  dependerá de la amplitud del error  $e(t_0) = y_{sp} - y(t_0)$ , siendo  $y_{sp}$  el valor de referencia.

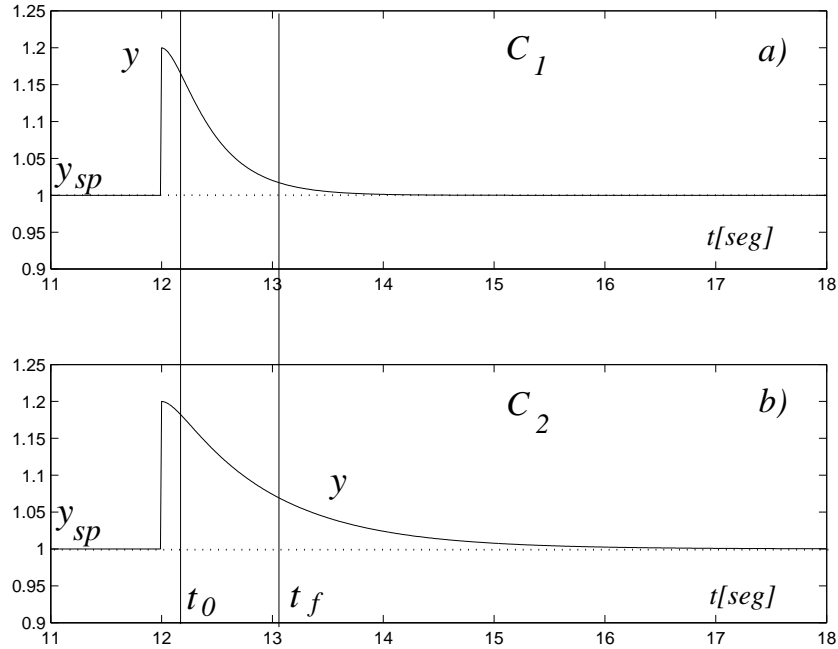


Figura 3: Comparación de performance de dos controladores.

También podríamos encontrar varios ejemplos (particularmente en procesos químicos) en los que la calidad del producto fabricado en el intervalo  $t_0 - t_f$  dependerá de

$$\int_{t_0}^{t_f} e(t) dt. \quad (5)$$

Si el valor de esta integral es usado para comparar la performance de los controladores  $C_1$  y  $C_2$  podría decirse que el controlador  $C_1$  es mejor que el  $C_2$ .

En realidad, la selección de (5) como índice de performance no sería muy afortunada. Efectivamente, un sistema con respuesta muy subamortiguada podría dar lugar a una cancelación de áreas bajo la curva del error, y así aparentar buena performance. Para evitar este problema, en todo índice de optimización solo se consideran variables positivas. Así son comunes índices de performance que consideran el modulo del error, por ejemplo:

$$\text{IAE} \quad J = \int_{t_0}^{t_f} |e(t)| dt, \quad (6)$$

$$\text{ITAE} \quad J = \int_{t_0}^{t_f} |e(t)| t dt, \quad (7)$$

o los conocidos como índices cuadráticos, por ejemplo:

$$J = \int_{t_0}^{t_f} e^2(t) dt. \quad (8)$$

Nuestra atención estará centrada exclusivamente en estos últimos, los índices cuadráticos.

Ahora bien, el índice (8) aún no pone de manifiesto ninguna restricción en el control, y de hecho si diseñáramos un control que minimice dicho índice, seguramente se presentarían problemas de incompatibilidad entre la acción de control requerida y la disponible. Para contemplar el compromiso entre la respuesta transitoria y las restricciones en el control  $u(t)$  puede considerarse el índice

$$J = \int_{t_0}^{t_f} (e^2(t) + \rho u^2(t)) dt + e^2(t_f), \quad (9)$$

donde  $\rho$  es simplemente un factor de ponderación de  $u^2(t)$  frente a  $e^2(t)$ , y el término  $e^2(t_f)$  penaliza el error en  $t = t_f$ .

Resolver un problema de control óptimo implicará obtener un controlador que minimice un funcional como (9).

## 4. Índice de performance para el problema del regulador óptimo.

Las presentes notas introducen las nociones de control óptimo a través de la resolución del problema conocido como LQR (linear quadratic regulator), que es un problema de *regulación*, que se resuelve a partir de la optimización de un índice de performance *cuadrático*, dando lugar a una solución *lineal*. El índice de performance es la generalización del funcional (9):<sup>2</sup>

$$J = x^T(t_f) P_0 x(t_f) + \int_{t_0}^{t_f} (x^T Q x + u^T R u) dt, \quad (10)$$

donde  $x$  son los estados del sistema elegidos como errores con respecto a los valores de estado estacionario. Las matrices de peso  $Q$ ,  $R$  y  $P_0$  son matrices reales, simétricas, constantes y definidas positivas <sup>3</sup>.

<sup>2</sup>de aquí en más, omitiremos el tiempo  $t$  en la notación de las variables, es decir  $x = x(t), \dots$

<sup>3</sup>En realidad,  $Q$  y  $P_0$  pueden ser definidas o semidefinidas positivas, pero no pueden ser simultáneamente nulas, ya que esto implicaría minimizar el índice:

$$J = \int_{t_0}^{t_f} u^T R u dt, \quad (11)$$

Probablemente, la forma más intuitiva de elegir<sup>4</sup> estas matrices sea la forma diagonal:

$$Q = \begin{bmatrix} q_{11} & 0 & \cdots & 0 \\ 0 & q_{22} & 0 & 0 \\ \vdots & 0 & \ddots & \\ 0 & 0 & 0 & \ddots \end{bmatrix} \quad R = \begin{bmatrix} r_{11} & 0 & \cdots & 0 \\ 0 & r_{22} & 0 & 0 \\ \vdots & 0 & \ddots & \\ 0 & 0 & 0 & \ddots \end{bmatrix} \quad (12)$$

En muchos problemas se considera  $t_f \rightarrow \infty$ , en este caso el índice de performance se reduce a

$$J = \int_{t_0}^{\infty} (x^T Q x + u^T R u) dt, \quad (13)$$

ya que, para que  $J$  sea finito, necesariamente  $x(t_f) \rightarrow 0$ .

Para el caso de sistemas de tiempo discreto, el índice de performance equivalente a (10) es

$$J = x^T(N) P_0 x(N) + \sum_{k=0}^{N-1} x^T(k) Q x(k) + u^T(k) R u(k), \quad (14)$$

mientras que el equivalente a (13), cuando  $N \rightarrow \infty$ , es

$$J = \sum_{k=0}^{\infty} x^T(k) Q x(k) + u^T(k) R u(k). \quad (15)$$

## 5. Problema del LQR. Principio de Optimalidad de Bellman.

En base a los comentarios previos, puede decirse que en esta nueva concepción del control (control óptimo) el establecimiento del problema requiere de:

1. la descripción del sistema a ser controlado,
2. la descripción de las restricciones y de los objetivos de control,
3. la definición de un criterio que permita juzgar el grado de optimización.

En particular la resolución del problema del regulador óptimo LQR consiste en: dado un modelo del sistema a controlar en variables de estado y un índice de performance de la forma (10) (o de la forma (14) para el caso discreto), encontrar el control  $u(t) = u^*(t)$  que minimiza el citado funcional  $J(= J^*)$ .

La solución del problema LQR puede ser abordada desde distintos caminos, en estas notas emplearemos el principio de optimalidad de Bellman para obtener la solución.

### Principio de Optimalidad de Bellman.

Considere un control genérico  $u$ . Su aplicación sobre el sistema a controlar dará lugar a una evolución de los estados que define una trayectoria. Diremos que esta trayectoria es óptima si el valor del control  $u(t)$  minimiza  $J$ .

---

que daría lugar a la solución trivial  $u = 0$  (e.d. sin control).

<sup>4</sup>obviamente esto depende de cada problema.

Según Bellman (*Principio de optimalidad*): ‘si una trayectoria entre dos estados  $x(t_0)$  y  $x(t_f)$  es óptima, entonces cualquier tramo de esta trayectoria, definido desde un punto  $x(t_i)$  (perteneciente a dicha trayectoria) y el punto  $x(t_f)$ , es en sí una trayectoria óptima’ (Figura 4).

Este principio (que podría decirse que es casi obvio), está íntimamente ligado a los conceptos de programación dinámica que permiten simplificar problemas de optimización de distintas especialidades (economía, ciencias físicas, ingeniería, ...).

Por razones de simplicidad, primero se resuelve el problema del regulador óptimo para el caso discreto y con horizonte de optimización finito ( $N$  finito) (sección 6). Posteriormente, en la sección 7 se considera el caso continuo (con horizonte de optimización finito e infinito), y el caso discreto con horizonte de optimización infinito es citado en el Apéndice B.

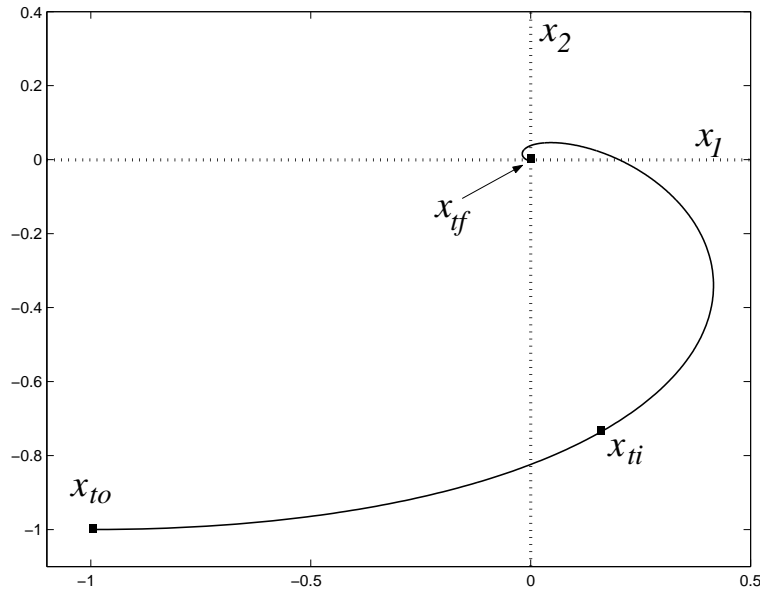


Figura 4: Trayectoria en el espacio de estados.

## 6. Solución del problema de regulación óptima. Caso discreto y horizonte finito.

En este caso el sistema a controlar está descrito por:

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) \end{aligned} \quad (16)$$

sujeto al índice de performance (14):

$$J = x^T(N) P_0 x(N) + \sum_{k=0}^{N-1} x^T(k) Q x(k) + u^T(k) R u(k). \quad (17)$$



## Optimización del control en el período definido por $(N - 1)T \rightarrow NT$ .

Mas allá de que existen  $N$  decisiones (una por cada período de muestreo) para optimizar el control del sistema (16)-(17), comencemos por deducir la ‘forma’ del control que corresponde a la última decisión, es decir la decisión a tomar en el instante  $(N - 1)T$ . Como veremos, la generalización de esta solución parcial nos permitirá obtener la ley de control óptima que buscamos<sup>5</sup>. Con esta finalidad supondremos que, a partir de las  $N - 1$  primeras decisiones se ha alcanzado el estado genérico  $x(N - 1)$ .

Para optimizar la última decisión consideramos el índice de performance

$$J_{N-1,N} = x^T(N) P_0 x(N) + x^T(N - 1) Q x(N - 1) + u^T(N - 1) R u(N - 1), \quad (18)$$

expresión en la que puede reemplazarse  $x(N)$  por las ecuaciones del sistema (16):

$$J_{N-1,N} = [Ax(N - 1) + Bu(N - 1)]^T P_0 [Ax(N - 1) + Bu(N - 1)] + x^T(N - 1) Q x(N - 1) + u^T(N - 1) R u(N - 1). \quad (19)$$

Para minimizar  $J_{N-1,N}$  con respecto al control  $u(N - 1)$  (constante en el intervalo  $(N - 1)T < t < NT$ ) es necesario que:

$$\frac{\partial J_{N-1,N}}{\partial u(N - 1)} = 0. \quad (20)$$

Es importante notar que el estado en el instante  $(N - 1)T$  no depende de la decisión a tomar en ese mismo instante, es decir es independiente de  $u(N - 1)$ . Luego, cálculos algebraicos triviales<sup>6</sup> conducen a

$$R u^*(N - 1) + B^T P_0 [Ax(N - 1) + B u^*(N - 1)] = 0 \quad (21)$$

(detalles de los pasos intermedios pueden encontrarse en el Apéndice A).

Luego, el control óptimo a aplicar en el instante  $(N - 1)T$  es<sup>7</sup>:

$$u^*(N - 1) = - [R + B^T P_0 B]^{-1} B^T P_0 A x(N - 1), \quad (22)$$

que podemos reescribirlo como

$$u^*(N - 1) = K(N - 1) x(N - 1). \quad (23)$$

Es decir que la decisión óptima a tomar en el instante  $(N - 1)T$ , es una realimentación de estados con la ganancia

$$K(N - 1) = - [R + B^T P_0 B]^{-1} B^T P_0 A, \quad (24)$$

---

<sup>5</sup>Este fraccionamiento en el proceso de optimización está basado en el principio de Bellman.

<sup>6</sup>Tener presente que  $\frac{d(x^T b)}{dx} = b$  y que  $\frac{d(x^T A x)}{dx} = (A^T + A) x$ .

<sup>7</sup>En realidad para asegurar que  $u^*(N - 1)$  minimiza el índice de performance (18) ( y que no lo maximice) debe verificarse que la segunda derivada de  $J$  sea ‘positiva’ (definida positiva). Efectivamente  $\frac{\partial^2 J_{N-1,N}}{\partial u^2(N-1)} = R + B^T P_0 B$  es definida positiva ya que  $R$  es definida positiva y  $P_0$  es definida o semidefinida positiva.

Para obtener el valor del índice de performance óptimo (el costo mínimo de la última decisión)  $J_{N-1,N}^*$  se reemplaza  $u^*(N-1)$  en (18), obteniéndose:

$$J_{N-1,N}^* = x^T(N-1) P_1 x(N-1), \quad (25)$$

donde

$$P_1 = [A + BK(N-1)]^T P_0 [A + BK(N-1)] + K^T(N-1)RK(N-1) + Q. \quad (26)$$

Como el control óptimo  $u^*(N-1)$  es obtenido a partir de consideraciones completamente generales, en particular, a partir de una condición inicial genérica, la solución obtenida tiene que estar contenida en la solución de optimizar el control en más de una decisión (principio de Bellman).

### Optimización del control en el período definido por $(N-2)T \rightarrow NT$ .

Procedamos ahora a la optimización en el período  $(N-2)T < t < NT$  que involucra dos decisiones (es decir, los controles  $u(N-2)$  y  $u(N-1)$ ). Para este fin, debemos considerar:

$$J_{N-2,N} = x^T(N) P_0 x(N) + \sum_{k=N-2}^{N-1} x^T(k) Q x(k) + u^T(k) R u(k), \quad (27)$$

es decir,

$$J_{N-2,N} = x^T(N-2) Q x(N-2) + u^T(N-2) R u(N-2) + J_{N-1,N}. \quad (28)$$

Como ya sabemos cual es el mínimo costo de  $J_{N-1,N}$  ( $= J_{N-1,N}^*$ ) podemos reemplazarlo en (28)

$$J_{N-2,N} = x^T(N-2) Q x(N-2) + u^T(N-2) R u(N-2) + J_{N-1,N}^*. \quad (29)$$

Luego, considerando las ecuaciones del sistema, resulta:

$$J_{N-2,N} = x^T(N-2) Q x(N-2) + u^T(N-2) R u(N-2) + [Ax(N-2) + Bu(N-2)]^T P_1 [Ax(N-2) + Bu(N-2)]. \quad (30)$$

Esta expresión tiene la forma de la ecuación (19), es decir, que al anular su derivada con respecto a  $u(N-2)$

$$\frac{\partial J_{N-2,N}}{\partial u(N-2)} = 0, \quad (31)$$

se obtiene una solución análoga a las ecuaciones (22)-(23)

$$u^*(N-2) = -[R + B^T P_1 B]^{-1} B^T P_1 A x(N-2) \quad (32)$$

$$u^*(N-2) = K(N-2) x(N-2). \quad (33)$$

Nuevamente, la decisión a tomar en el instante  $(N-2)T$  es una realimentación de estados, pero con un juego de ganancias distinto al correspondiente a la decisión a tomar en el instante  $(N-1)T$ .

Reemplazando el control óptimo  $u^*(N-2)$  en (29) se obtiene el mínimo costo de  $J_{N-2,N}$ :

$$J_{N-2,N}^* = x^T(N-2) P_2 x(N-2), \quad (34)$$

donde

$$P_2 = [A + BK(N-2)]^T P_1 [A + BK(N-2)] + K^T(N-2)RK(N-2) + Q. \quad (35)$$

### Optimización del control en el período definido por $0 \rightarrow NT$ .

Reiterando el procedimiento podemos encontrar, por inducción, la forma general del regulador óptimo. Éste resulta de la aplicación iterativa de las ecuaciones (36), (37), (38) y (39).

$$u^*(N-j) = K(N-j) x(N-j) \quad (36)$$

$$K(N-j) = -[R + B^T P_{j-1} B]^{-1} B^T P_{j-1} A \quad (37)$$

$$P_j = [A + BK(N-j)]^T P_{j-1} [A + BK(N-j)] + K^T(N-j) R K(N-j) + Q \quad (38)$$

$$J_{N-j,N}^* = x^T(N-j) P_j x(N-j). \quad (39)$$

Resulta conveniente observar que

- El regulador óptimo consiste en una realimentación lineal de los estados.
- Tanto el regulador como el sistema de lazo cerrado son variantes en el tiempo.
- Una vez calculadas las ganancias, la implementación del regulador resulta simple. *¿Cómo lo haría?* Si los estados no fueran accesibles podría complementarse el regulador con un observador de estados (obviamente el sistema debería satisfacer las condiciones de observabilidad).
- A partir de la última ecuación el costo mínimo  $J^*$  que resulta de aplicar el control óptimo  $u^*$ , resulta:

$$J^* = J_{0,N}^* = x^T(0) P_N x(0). \quad (40)$$

*Ejemplo 3.* Considere el sistema muestreado:

$$x(k+1) = \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} x(k) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(k). \quad (41)$$

Calcular el regulador óptimo que minimiza el índice

$$J = x_{(10)}^T \begin{bmatrix} 2,5 & 0 \\ 0 & 2,5 \end{bmatrix} x_{(10)} + \sum_{k=1}^9 \left( x_{(k)}^T \begin{bmatrix} 1 & 0 \\ 0 & ,05 \end{bmatrix} x_{(k)} + u_{(k)}^2 \right), \quad (42)$$

y simular la evolución de los estados a partir de la condición inicial  $x_{(0)} = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$ .

Empleando iterativamente las ecuaciones (37) y (38) pueden calcularse la matriz  $P$  y las ganancias de realimentación. Esta tarea es realizada a partir de la condición inicial  $P_0 = \begin{bmatrix} 2,5 & 0 \\ 0 & 2,5 \end{bmatrix}$  (correspondiente a  $k = 10$ ) en la ‘dirección de tiempos decrecientes’ hasta alcanzar  $N = 0$ .

La figura 4 muestra los elementos de la matriz  $P$  que permiten el cálculo de las ganancias de realimentación correspondientes a cada período de muestreo. Observe que las ganancias corresponden a  $k = 1, 2, \dots, 9$ , el valor para  $k = 10$  queda excluido *¿Por qué?*. Por último, en la parte c) de la figura se muestra la evolución de los estados del sistema a lazo cerrado.

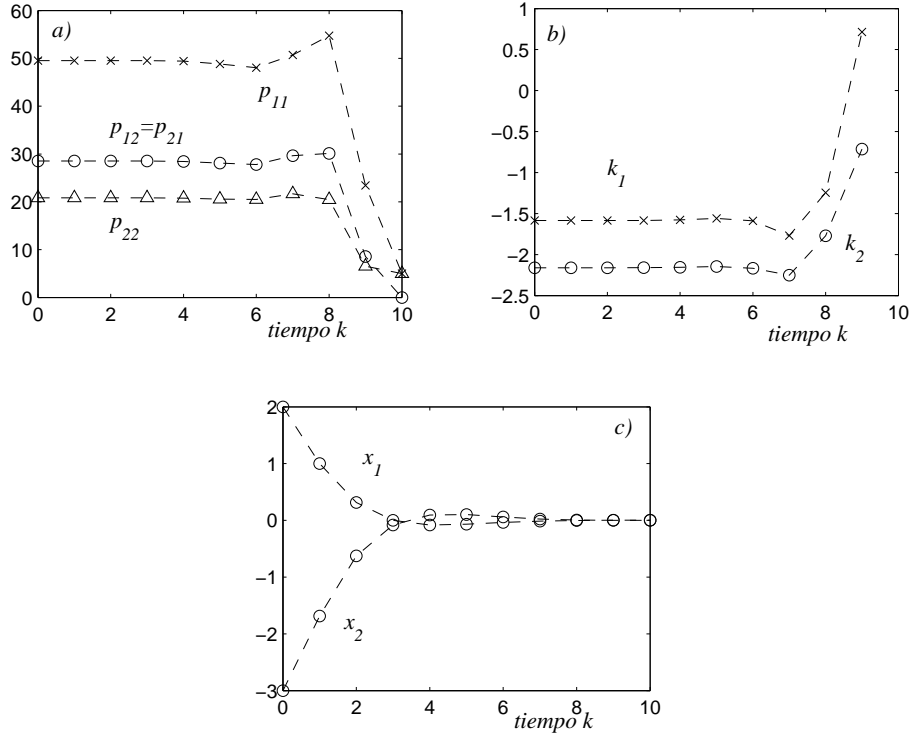


Figura 5: *a*- Elementos de la matriz  $P$  en función de  $k$ . *b*- Ganancias del regulador. *c*- Estados.

## 7. Solución del problema de regulación óptima. Caso continuo.

### 7.1. Horizonte de optimización finito ( $t_f - t_0$ finito).

La solución del problema de regulación óptima de un sistema de tiempo continuo

$$\begin{aligned}\dot{x} &= Ax + Bu \\ y &= Cx\end{aligned}\tag{43}$$

donde el índice de performance es de la forma

$$J = x^T(t_f) H x(t_f) + \int_{t_0}^{t_f} (x^T Q x + u^T R u) dt,\tag{44}$$

es similar a la solución del caso discreto correspondiente (para ser sinceros, ligeramente más complicada). Para ver los detalles de la deducción recomendamos el libro Linear State Space Systems, John Bay, 1999.

La solución a este problema también es una realimentación de estados

$$u^* = K(t) x,\tag{45}$$

pero ahora las ganancias son funciones continuas del tiempo

$$K(t) = -R^{-1} B^T P(t),\tag{46}$$

donde  $P(t)$  debe ser calculada a partir de la ecuación diferencial

$$\frac{dP}{dt} + Q - P(t) B R^{-1} B^T P(t) + P(t) A + A^T P(t) = 0,\tag{47}$$

conocida como **ecuación diferencial de Riccati**. La condición inicial para resolver (47) es  $P(t_f) = P_0 = H$ .

Como habrá notado la ecuación diferencial de Riccati, a mas de matricial, es no lineal. Su resolución puede ser complicada aún empleando métodos numéricos. Este hecho, podría desanimar a más de un diseñador. Sin embargo, nos sirve de motivación para proponer un controlador ‘cuasióptimo’ (es decir con performance muy próxima a la del control óptimo) de muy fácil implementación.

Para ser más claros, consideremos el siguiente ejemplo.

*Ejemplo 4.* Considere el sistema de primer orden

$$\dot{x}(t) = 2x(t) + u(t). \quad (48)$$

Se desea encontrar la ley de control que minimiza el siguiente índice de performance:

$$J = \int_0^{t_f} (3x^2 + \frac{1}{4}u^2) dt, \quad (49)$$

siendo  $t_f = 1 \text{ seg}$ .

Sabemos que la ley de control óptima responde a:

$$u^* = -R^{-1}B^T P(t) x(t) \quad (50)$$

donde  $P(t)$  es la solución de la ecuación diferencial de Riccati:

$$\dot{P}(t) + Q - P(t)BR^{-1}B^T P(t) + P(t)A + A^T P(t) = 0 \quad (51)$$

El presente ejemplo es lo suficientemente simple para obtener una solución analítica de  $P(t)$  sin mayor dificultad. En casos de mayor complejidad son recomendables métodos de integración numérica, desde  $t = t_f$  hasta  $t = t_0$  con la condición  $P(t_f) = H$ .

Fácilmente, se identifican las ‘matrices’:  $A = 2, B = 1, H = 0, Q = 3, R = 1/4$ . Luego la ecuación diferencial de Riccati resulta:<sup>8</sup>

$$\dot{p}(t) + 3 - 4p^2(t) + 4p(t) = 0 \quad (52)$$

con la condición  $p(t_f) = 0$ .

La solución  $p(t)$  puede ser obtenida por el método de integración de separación de variables

$$\dot{p}(t) = 4p^2(t) - 4p(t) - 3 = 4(p(t) - 3/2)(p(t) + 1/2) \quad (53)$$

$$\int_{t_f}^t \frac{dp}{4(p - \frac{3}{2})(p + \frac{1}{2})} = \int_{t_f}^t dt \quad (54)$$

$$\frac{1}{4} \int_{t_f}^t \frac{1/2}{(p - \frac{3}{2})} dp + \frac{1}{4} \int_{t_f}^t \frac{-1/2}{(p + \frac{1}{2})} dp = t - t_f \quad (55)$$

$$\frac{1}{8} \left\{ \ln\left(\frac{p(t) - 3/2}{p(t_f) - 3/2}\right) - \ln\left(\frac{p(t) + 1/2}{p(t_f) + 1/2}\right) \right\} = t - t_f \quad (56)$$

---

<sup>8</sup>como el sistema (48) tiene un sólo estado, la ecuación de Riccati resulta escalar, e.d.  $P(t) = p(t)$ .

resultando,

$$p(t) = \frac{\frac{3}{2}(1 - e^{8(t-t_f)})}{1 + 3e^{8(t-t_f)}}. \quad (57)$$

Luego, la acción de control resulta:

$$u^* = -4p(t) x(t). \quad (58)$$

Puede observarse que, aún para este ejemplo simple, la implementación del regulador óptimo con horizonte de optimización finito, resulta muy complicada (por favor, *No Desesperar!*).

## 7.2. Horizonte de optimización infinito.

El regulador óptimo de tiempo infinito es aquel que minimiza el índice de performance:

$$J = \int_0^\infty (x^T(t)Qx(t) + u^T Ru(t)) dt \quad (59)$$

Cabe destacar que:

- a diferencia del regulador de tiempo finito, ahora es necesario imponer condiciones sobre la controlabilidad del sistema para asegurar que el índice de performance sea finito.
- si  $t_f \rightarrow \infty$ ,  $x(t_f)$  también debe tender a 0. Si no fuera así  $J$  sería infinito. Luego, en el índice (59) se ha omitido el término  $x^T(t_f)Hx(t_f)$  presente en (44), ya que se ha perdido el sentido que justificaba su inclusión.

Empleemos el ejemplo 4 a los efectos de introducir la solución al problema de optimización cuando  $t_f$  es muy grande (regulador óptimo de tiempo infinito).

Sabemos que la ecuación diferencial de Riccati se resuelve ‘hacia atrás’ con el tiempo inicial  $t = t_f$  correspondiente a la condición inicial  $P(t_f)$ . En el ejemplo previo, la solución fue hallada en forma analítica sin mayor complejidad. En la figura 5 se grafica la variable  $p(t)$  que define la ganancia de realimentación. Puede observarse que  $p(t)$  presenta dos componentes, una transitoria y otra permanente. Para tiempos  $t = t_f - \varepsilon$  con  $\varepsilon$  positivo y pequeño, se distingue la respuesta transitoria debida a la condición inicial  $p(t_f)$ . Agotado el transitorio ( $t \ll t_f$ ),  $p(t)$  se corresponde con la solución de estado estacionario de la ecuación de Riccati, la cual puede calcularse anulando la derivada  $\dot{p}(t)$  en la expresión (51).

En la figura (6), también se muestra como influye un aumento de  $t_f$  sobre  $p(t)$ .

Obviamente, en la medida que  $t_f \rightarrow \infty$ , la respuesta de estado estacionario predomina sobre la transitoria. Luego, para el caso límite, podría tomarse  $p(t)$  constante e igual al valor de estado estacionario

$$p_{ss} = \lim_{t \rightarrow \infty} p(t) = 1,5 \quad (60)$$

como una muy buena aproximación al control óptimo. Así, puede implementarse un controlador extremadamente simple a través de una ley de realimentación invariante en el tiempo:

$$u^* = -6x(t). \quad (61)$$

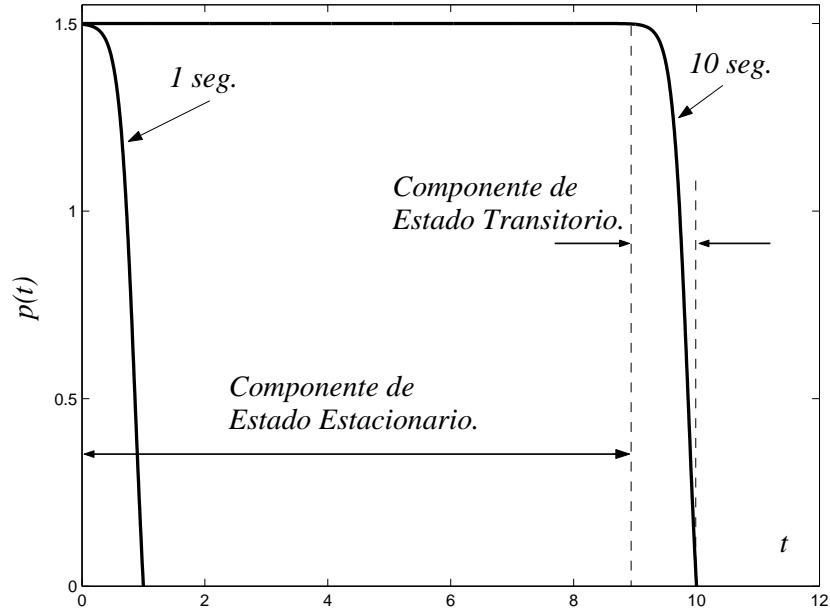


Figura 6:  $p(t)$  para distintos horizontes de optimización.

Vemos que la simplificación de (61) con respecto a (58) es significativa.

Más allá del ejemplo particular, la solución cuasióptima resulta particularmente motivante. Observe que, generalmente, el problema de optimización con horizonte de optimización infinito es más frecuente que el caso finito.

Podemos resumir la **solución del regulador óptimo de tiempo infinito** en las siguientes ecuaciones:

$$u^* = K_c x, \quad (62)$$

con

$$K_c = -R^{-1} B^T P_c, \quad (63)$$

y donde  $P_c$  es conocido como la solución de la **ecuación algebraica de Riccati**

$$Q - P_c B R^{-1} B^T P_c + P_c A + A^T P_c = 0. \quad (64)$$

Observación: la solución de la ecuación algebraica de Riccati puede no ser única. Sin embargo, si el sistema es controlable, puede forzarse una solución única exigiendo que ésta sea la que corresponde a  $P_c$  definida positiva. En este caso, puede demostrarse, que el lazo cerrado es siempre estable<sup>9</sup>.

*Ejemplo 5.* Considere el sistema a controlar:

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u. \quad (65)$$

Se desea encontrar el control óptimo que permita minimizar:

$$J = \int_{t_0}^{\infty} (x_1^2 + u^2) dt \quad (66)$$

---

<sup>9</sup>se asume  $Q$  definida positiva

Fácilmente pueden identificarse:

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad Q = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}, \quad R = 2. \quad (67)$$

Luego, la ecuación algebraica de Riccati resulta:

$$\begin{aligned} & \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \frac{1}{2} \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} + \\ & + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} + \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \end{aligned} \quad (68)$$

es decir,

$$\begin{aligned} -\frac{p_{12}^2}{2} + 2 &= 0 \\ p_{11} - \frac{p_{12}p_{22}}{2} &= 0 \\ -\frac{p_{22}^2}{2} + 2p_{12} &= 0 \end{aligned} \quad (69)$$

La solución de estas ecuaciones conduce a la matriz definida positiva  $P_c$ :

$$P_c = \begin{bmatrix} 2\sqrt{2} & 2 \\ 2 & 2\sqrt{2} \end{bmatrix}, \quad (70)$$

Resultando el control óptimo

$$u^* = -R^{-1}B^T P_c x \quad (71)$$

$$u^* = -x_1 - \sqrt{2}x_2 \quad (72)$$



## Apéndice A. Obtención de la expresión (21)

$$\frac{\partial J_{N-1,N}}{\partial u(N-1)} = 2Ru(N-1) + \frac{\partial \left\{ [Ax(N-1) + Bu(N-1)]^T P_0 [Ax(N-1) + Bu(N-1)] \right\}}{\partial u(N-1)} \quad (\text{A-1})$$

$$\frac{\partial J_{N-1,N}}{\partial u(N-1)} = \begin{cases} 2Ru(N-1) + \\ + \frac{\partial}{\partial u(N-1)} \{ [Ax(N-1)]^T P_0 [Ax(N-1)] \\ + [Bu(N-1)]^T P_0 [Bu(N-1)] + \\ + \underbrace{[Ax(N-1)]^T P_0 [Bu(N-1)]}_{\text{escalar}} + \\ + [Bu(N-1)]^T P_0 [Ax(N-1)] \} \end{cases} \quad (\text{A-2})$$

$$[Ax(N-1)]^T P_0 [Bu(N-1)] = \left[ [Ax(N-1)]^T P_0 [Bu(N-1)] \right]^T \quad (\text{A-3})$$

$$[Ax(N-1)]^T P_0 [Bu(N-1)] = [Bu(N-1)]^T P_0 [Ax(N-1)] \quad (\text{A-4})$$

$$\frac{\partial J_{N-1,N}}{\partial u(N-1)} = \begin{cases} 2Ru(N-1) + \\ + \frac{\partial}{\partial u(N-1)} \{ [Bu(N-1)]^T P_0 [Bu(N-1)] + \\ + 2[Bu(N-1)]^T P_0 [Ax(N-1)] \} \end{cases} \quad (\text{A-5})$$

$$\frac{\partial J_{N-1,N}}{\partial u(N-1)} = \begin{cases} 2Ru(N-1) + \frac{\partial}{\partial u(N-1)} \{ [Bu(N-1)]^T P_0 [Bu(N-1)] + \\ + 2[Bu(N-1)]^T P_0 [Ax(N-1)] \} \end{cases} \quad (\text{A-6})$$

$$\frac{\partial J_{N-1,N}}{\partial u(N-1)} = 2Ru(N-1) + 2B^T P_0 Bu(N-1) + 2B^T P_0 Ax(N-1) \quad (\text{A-7})$$

$$\frac{\partial J_{N-1,N}}{\partial u(N-1)} = 2Ru(N-1) + 2B^T P_0 [Ax(N-1) + Bu(N-1)] \quad (\text{A-8})$$

$$\frac{\partial J_{N-1,N}}{\partial u(N-1)} = 2Ru^*(N-1) + 2B^T P_0 [Ax(N-1) + Bu(N-1)] = 0 \quad (\text{A-9})$$

## Apéndice B. Sistemas de tiempo discreto. Horizonte Infinito.

Dado el sistema muestreado:

$$x(k+1) = Ax(k) + Bu(k) \quad (\text{B-1})$$

$$y(k) = Cx(k), \quad (\text{B-2})$$

el control óptimo  $u^*(k)$  que minimiza el índice de performance:

$$J = \sum_{k=0}^{\infty} x^T(k) Q x(k) + u^T(k) R u(k) \quad (\text{B-3})$$

para dado apartamiento inicial de los estados  $x(0)$ , resulta invariante en el tiempo. Su valor se calcula a partir de la solución de estado estacionario de  $P$ , resultando:

$$u^*(k) = K x(k) \quad (\text{B-4})$$

con

$$K = - \left[ R + B^T P_{cd} B \right]^{-1} B^T P_{cd} A \quad (\text{B-5})$$

$$P_{cd} = [A + BK]^T P_{cd} [A + BK] + K^T R K + Q, \quad (\text{B-6})$$

siendo el mínimo valor de  $J$ :

$$J^* = x^T(0) P_{cd} x(0) \quad (\text{B-7})$$

## Referencias

- Bay, J. (1999). *Linear State Space Systems*, McGraw-Hill.
- Brogan, W. (1991). *Modern Control Theory*, Prentice Hall.
- Friedland, B. (1986). *Control System Design*, McGraw-Hill.
- Kwakernaak, H. & Sivan, R. (1972). *Linear Optimal Control Systems*, Willey-Intercience.
- Nagrath, I. & Gopal, M. (1982). *Control System Engineering*, Wiley.