

## Capítulo 7

# Introducción al Control Óptimo

### 7.1. El problema de control óptimo

En la formulación del problema de control óptimo hay dos ingredientes:

- 1.- Un sistema de control que en el caso continuo viene dado por un sistema de ecuaciones diferenciales:

$$\begin{cases} \dot{x}(t) = f(t, x(t), u(t)), & t \in [t_0, t_f] \subset \mathbb{R} \\ x(t_0) = x_0 \end{cases} \quad (7.1)$$

que cumple las condiciones de existencia y unicidad de las soluciones (véase la Sección 4.2.1). Y en el caso discreto por un sistema de ecuaciones en diferencias:

$$\begin{cases} x(t+1) = f_t(x(t), u(t)), & t = 0, 1, \dots, N-1 \\ x(0) = x_0 \end{cases} \quad (7.2)$$

En los dos casos  $x(t)$  es el vector de  $\mathbb{R}^n$  que para cada instante de tiempo  $t$  nos da el *estado* del sistema,  $u(t)$  es, para cada  $t$ , el vector de controles o entradas que toma valores en algún conjunto  $U \subseteq \mathbb{R}^m$  denominado conjunto de los valores admisibles de los controles,  $t_0$  es el tiempo inicial en cuyo instante el sistema se encuentra en el estado inicial  $x_0$ . En el caso discreto es habitual poner  $t = 0$  como instante inicial. Obsérvese que en este caso (el discreto) se puede pensar en la variable tiempo  $t$  como un índice; es por eso que la dependencia de  $f$  en  $t$  la indicamos como un subíndice  $f_t$ . Lo mismo haremos con las demás funciones que vayan apareciendo. Supondremos, además, que se dan las condiciones de existencia y unicidad de soluciones para cada sistema considerado.

- 2.- Un *coste funcional* que asocia un valor (coste) a cada función de control admisible. El coste funcional se suele denotar con  $J$  y tiene la siguiente forma en el caso continuo:

$$J(u(\cdot)) = K(t_f, x_f) + \int_{t_0}^{t_f} L(t, x(t), u(t)) dt, \quad (7.3)$$

y en el caso discreto

$$J(u(\cdot)) = K(x_f) + \sum_{t=0}^{N-1} L_t(x(t), u(t)). \quad (7.4)$$

En ambos casos  $L$ ,  $L_t$  ( $t = 0, 1, \dots, N-1$ ) y  $K$  son funciones dadas (*coste operativo* y *coste final*, respectivamente),  $t_f$  ( $N$  en el caso discreto) es el *tiempo final* o *terminal* que puede ser libre o fijo y  $x_f = x(t_f)$  ( $x(N) = x_f$  en el caso discreto) es el *estado final* o *terminal* que también puede ser libre o fijo. Al decir que  $t_f$  y/o  $x_f$  pueden ser libres o fijos se quiere indicar que  $t_f$  y  $x_f$  pueden haber sido fijados de antemano o se consideran como variables del problema. Nosotros supondremos que el tiempo final ha sido fijado de antemano aunque, como decimos, problemas más generales pueden ser considerados.

Debe observarse que  $u(\cdot)$  es una función de modo que  $J$  es una función con valores reales pero definida en un espacio de funciones. Es por ello, que decimos que  $J$  es un *funcional* (véase Apéndice A). El espacio de funciones donde puede elegirse  $u(\cdot)$  puede ser bastante restrictivo. Lo denotaremos por  $\mathcal{U}$  y, tal y como se dijo en la Sección 4.2.1, suele ser el espacio de funciones continuas a trozos,  $\mathcal{U} = PC([t_0, t_f], U)$ , siendo  $U$  el conjunto de los valores admisibles de los controles, o, más generalmente, el espacio de funciones localmente integrables (véase la Sección A.4 del Apéndice A). Lo llamaremos genéricamente *conjunto de funciones admisibles*. También los valores de los estados pueden estar sometidos a restricciones que, en cada caso y si fuera necesario, se enunciarán de forma explícita.

Con todas estas consideraciones en mente, podemos definir lo que se entiende por un *problema de control óptimo*: Encontrar un control  $u \in \mathcal{U}$  que minimiza el funcional  $J$  sobre todos los controles admisibles. Un poco más explícitamente, en el caso continuo se trataría de encontrar, si existe, una función  $u^*(\cdot) \in \mathcal{U}$  tal que la solución  $x^*(t)$  del problema de condiciones iniciales (7.1) con  $u(t) = u^*(t)$ ,  $t \in [t_0, t_f]$ , cumpla:

$$K(t_f, x^*(t_f)) + \int_{t_0}^{t_f} L(t, x^*(t), u^*(t)) dt \leq K(t_f, x(t_f)) + \int_{t_0}^{t_f} L(t, x(t), u(t)) dt \quad (7.5)$$

para todo control  $u(\cdot) \in \mathcal{U}$  y la correspondiente solución  $x(\cdot)$  de (7.1). Escribiremos brevemente que  $u^*(\cdot) \in \mathcal{U}$  es la solución del problema:

$$\begin{aligned} &\text{minimizar } J(u(\cdot)) = K(t_f, x(t_f)) + \int_{t_0}^{t_f} L(t, x(t), u(t)) dt \\ &\text{sujeto a } \begin{cases} \dot{x}(t) = f(t, x(t), u(t)), & t \in [t_0, t_f] \subset \mathbb{R} \\ x(t_0) = x_0. \end{cases} \end{aligned} \quad (7.6)$$

En lo sucesivo, y por simplicidad notacional, denotaremos a las funciones de  $\mathcal{U}$  por  $u$  en vez de  $u(\cdot)$ ; y haremos lo mismo con las soluciones  $x$  del correspondiente problema de condiciones iniciales. Entonces  $u(t)$  y  $x(t)$  son los valores que toman  $u$  y  $x$  en  $t \in [t_0, t_f]$ .

En el caso discreto, el problema de control óptimo consiste en encontrar, si existe, una sucesión de controles  $\{u^*(t)\}_{t=0}^{N-1} \subset U$  tal que la sucesión solución  $\{x^*(t)\}_{t=0}^{N-1}$  del problema de condiciones iniciales (7.2) cumple:

$$K(x^*(N)) + \sum_{t=0}^{N-1} L_t(x^*(t), u^*(t)) \leq K(x(N)) + \sum_{t=0}^{N-1} L_t(x(t), u(t)), \quad (7.7)$$

para toda sucesión  $\{u(t)\}_{t=0}^{N-1} \subset U$  y la correspondiente sucesión solución  $\{x(t)\}_{t=0}^{N-1}$  de (7.2). Al igual que en el caso continuo escribiremos que  $u^*$  es la solución del problema

$$\begin{aligned} &\text{minimizar } J(u) = K(x(N)) + \sum_{t=0}^{N-1} L_t(x(t), u(t)) \\ &\text{sujeto a } \begin{cases} x(t+1) = f_t(x(t), u(t)), & t = 0, 1, \dots, N-1 \\ x(0) = x_0 \end{cases} \end{aligned} \quad (7.8)$$

Para ambos tipos de problemas hay dos técnicas clásicas de resolución del problema de control óptimo: la *programación dinámica* desarrollada por Richard Bellman y la del *principio del máximo de Pontriagin* desarrollada por el matemático ruso L. S. Pontriagin. En las dos próximas secciones estudiaremos la primera de ellas para cada tipo de sistemas.

## 7.2. Programación dinámica: caso discreto

La Programación dinámica se basa en el llamado **principio de optimalidad de Bellman**: *Una política óptima tiene la propiedad de que cualquiera que sean el estado inicial y la decisión inicial, las restantes decisiones deben constituir una política óptima en relación al estado resultante de la primera decisión.*

Enseguida aplicaremos este principio a nuestro problema de control óptimo, pero conviene analizarlo de cerca para comprender su significado. En la Figura 7.1 se muestra un

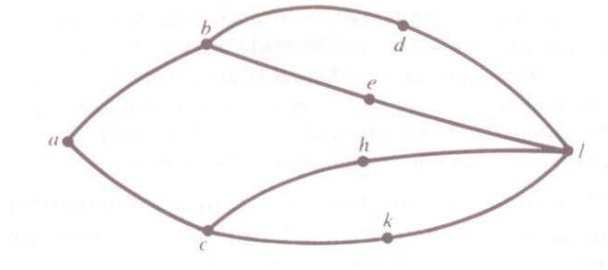


Figura 7.1: Grafo de recorridos de  $a$  a  $l$ .

grafo que representa las posibles formas de ir de un origen (estado inicial)  $a$  a un destino (estado final)  $l$ . Se puede ir pasando por ciertos puntos intermedios (estados posibles) pero ir de un punto (estado) a otro tiene un coste (se puede pensar en consumo de combustible,

tiempo, etc.). Si  $x$  y  $z$  son dos estados cualesquiera, representamos con  $J_{xz}$  el coste de ir del estado  $x$  al  $z$ . Así mismo  $J_x^*$  es el coste mínimo de ir del estado  $x$  al  $l$  y  $J^*$  el coste mínimo de  $a$  a  $l$ .

Una búsqueda exhaustiva nos da:

$$J^* = \min\{J_{ab} + J_{bd} + J_{dl}, J_{ab} + J_{be} + J_{el}, J_{ac} + J_{ch} + J_{hl}, J_{ac} + J_{ck} + J_{kl}\},$$

para lo que deberíamos realizar 12 sumas. El principio de optimalidad nos dice que si el trayecto óptimo para ir de  $a$  a  $l$  es, digamos, ir de  $a$  a  $c$ , de  $c$  a  $h$  y de  $h$  a  $l$ , entonces, necesariamente, el trayecto óptimo de ir de  $c$  a  $l$  es ir de  $c$  a  $h$  y entonces a  $l$ . ¿Cómo lo aplicamos? Moviéndonos hacia atrás en el grafo: Calculamos los trayectos óptimos de  $b$  y  $c$  a  $l$ :

$$J_b^* = \min\{J_{bd} + J_{dl}, J_{be} + J_{el}\}, \quad J_c^* = \{J_{ch} + J_{hl}, J_{ck} + J_{kl}\},$$

y, aplicando el principio de optimalidad,

$$J^* = \min\{J_b^* + J_{ab}, J_c^* + J_{ac}\}.$$

El número de sumas a realizar es, en este caso, exactamente la mitad de las que hay que hacer mediante la búsqueda exhaustiva.

Un ejemplo numérico es el siguiente:

**Ejemplo 7.1** Un motorista quiere minimizar el coste de ir desde el sitio en el que está a un destino determinado  $h$ . Sólo puede hacerlo siguiendo las direcciones indicadas en la Figura 7.2 (las calles son de dirección única) y el coste de ir de una intersección a la siguiente es el indicado; por ejemplo, el coste de ir del estado  $c$  al  $d$  es 3. Supongamos que

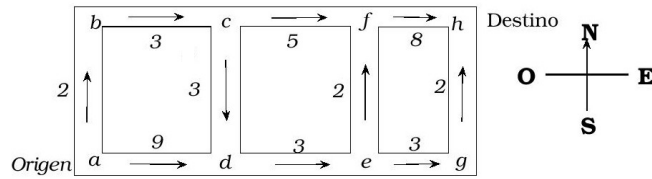


Figura 7.2: Mapa de las calles.

el motorista se encuentra en  $c$ . De ahí sólo se puede mover a  $d$  y a  $f$ . Supongamos también que sabe el coste mínimo de ir de  $d$  y de  $f$  a su destino final  $h$ . En este caso  $J_d^* = J_f^* = 8$ . Entonces, aplicando el principio de optimalidad, el coste mínimo de ir de  $c$  a  $h$  es

$$J_c^* = \min\{J_d^* + J_{cd}, J_f^* + J_{cf}\} = \min\{11, 13\} = 11.$$

Y el control óptimo sería la secuencia de decisiones: {S, E, E, N} donde S, E y N significan dirigirse hacia el Sur, hacia el Este y hacia el Norte, respectivamente.

¿Cómo sabe el motorista que  $J_d^* = 8$ ? Calculando primero  $J_e^*$ . Como para ir de  $e$  a  $h$  tiene dos opciones: a través de  $f$  o a través de  $g$ ; necesita primero calcular  $J_f^*$  y  $J_g^*$

cuyos valores son 8 y 2, respectivamente. Por lo tanto, de nuevo aplicando el principio de optimalidad

$$J_e^* = \min\{J_f^* + J_{ef}, J_g^* + J_{eg}\} = \min\{10, 5\} = 5.$$

En consecuencia,  $J_d^* = J_{de} + J_e^* = 8$ , tal y como habíamos dicho. En definitiva, para aplicar el principio de optimalidad a cada intersección, debemos calcular los valores óptimos empezando desde las intersecciones más próximas a  $h$  e ir hacia atrás. En este caso el orden sería:  $J_g^*$ ,  $J_f^*$ ,  $J_e^*$ ,  $J_d^*$ ,  $J_c^*$ ,  $J_b^*$  y  $J_a^*$ . En el Cuadro 7.1 se presentan los cálculos que hay que hacer aplicando el principio de optimalidad en cada intersección.

Intersección actual	Dirección	Próxima intersección	Coste mínimo de $\alpha$ a $h$ vía $x_i$	Coste mínimo de $\alpha$ a $h$	Dirección óptima en $\alpha$
$\alpha$	$u_i$	$x_i$	$J_{x_i h}^* + J_{\alpha x_i} = C_{\alpha x_i h}^*$	$J_{\alpha h}^*$	$u^*(\alpha)$
g	N	h	$0+2=2$	2	N
f	E	h	$0+8=8$	8	E
e	N	f	$8+2=10$		
	E	g	$2+3=5$	5	E
d	E	e	$5+3=8$	8	E
c	E	f	$8+5=13$		
c	S	d	$8+3=11$	11	S
b	E	c	$11+3=14$	14	E
a	N	b	$14+2=16$	16	N
a	E	d	$8+9=17$		

Tabla 7.1: Recorridos óptimos.

La notación empleada es la siguiente:

- $\alpha$  representa el estado del sistema en cada instante (intersección)
- $u_i$  son las direcciones admitidas (controles) en cada estado. En este ejemplo hay 3 posibles controles: N (norte), S (sur) y E (este).
- $x_i$  es el estado contiguo a  $\alpha$  que se puede alcanzar aplicando alguno de los controles  $u_i$  admitidos en  $\alpha$ .
- $J_{\alpha x_i}$  es el coste de ir de  $\alpha$  a  $x_i$
- $J_{x_i h}^* = J_{x_i}^*$  es el coste mínimo de alcanzar el estado final  $h$  desde  $x_i$ .
- $C_{\alpha x_i h}^*$  es el mínimo coste de ir de  $\alpha$  a  $h$  a través de  $x_i$ .
- $J_{\alpha h}^* = J_{\alpha}^*$  es el coste mínimo de ir de  $\alpha$  a  $h$  por cualquier trayecto admitido.
- $u^*(\alpha)$  es la decisión (control) óptima en  $\alpha$ .

Así, la trayectoria óptima para ir de  $a$  a  $h$  sería:  $a \rightarrow b \rightarrow c \rightarrow d \rightarrow e \rightarrow g \rightarrow h$  y el control óptimo:  $u^* = \{N, E, S, E, E, E, N\}$ . El coste mínimo  $J_a^* = 16$ .

Si pensamos en los cruces de las calles como puntos en un sistema de coordenadas cartesianas podemos asignar a cada cruce un par de coordenadas. Por ejemplo,  $a = (0, 0)$ ,  $b = (0, 1)$ ,  $d = (1, 0)$ ,  $c = (1, 1)$ ,  $e = (2, 0)$ ,  $f = (2, 1)$ ,  $g = (3, 0)$  y  $h = (3, 1)$ . El problema que estamos resolviendo se podría enunciar como : *calcular el coste mínimo de ir del punto  $(x, y)$  al punto  $(3, 1)$  usando las direcciones y el coste asignado a cada transición indicados en la Figura 7.2 siendo  $x = 0, 1, 2, 3$ ,  $y = 0, 1$ .*

A fin de establecer un modelo matemático para este problema observamos que las direcciones posibles en la Figura 7.2 son  $\{N, S, E\}$ . Además ir de un punto a sus adyacentes tiene un coste que podemos representar de la siguiente manera:

$$\begin{aligned} (x, y) &\xrightarrow{N} (x, y+1) \quad \text{coste: } c_N(x, y) \\ (x, y) &\xrightarrow{S} (x, y-1) \quad \text{coste: } c_S(x, y) \\ (x, y) &\xrightarrow{E} (x+1, y) \quad \text{coste: } c_E(x, y). \end{aligned}$$

Se debe entender que el coste de ir en la dirección contraria a la que apuntan las flechas de la Figura 7.2 es  $\infty$ . Se omite la dirección oeste porque en este ejemplo ninguna flecha apunta en dicha dirección. Incorporar la dirección oeste se podría hacer imponiendo que el coste en esa dirección desde cada punto es  $\infty$ . De la misma forma, podemos extender las definiciones de  $c_N$ ,  $c_S$  y  $c_E$  para todo par de puntos  $(x, y) \in \mathbb{Z}^2$  sin más que asignar coste infinito a cualquier transición entre dos puntos que no estén en  $\{0, 1, 2, 3\} \times \{0, 1\}$ .

Para  $(x, y) \in \mathbb{Z}^2$ , si definimos la *función valor óptimo* como la que da el coste del camino óptimo de  $(x, y)$  a  $(3, 1)$ , ésta sería  $J_{(x, y)}^*$ , que, por simplicidad notacional, la escribiremos  $J^*(x, y)$ . Y el principio de optimalidad establece que para todo  $(x, y) \in \mathbb{Z}^2$

$$J^*(x, y) = \min\{c_N(x, y) + J^*(x, y+1), c_S(x, y) + J^*(x, y-1), c_E(x, y) + J^*(x+1, y)\}.$$

Esta ecuación recibe el nombre de *ecuación de optimalidad* para el problema en estudio. Debemos exigir además la siguiente *condición de contorno*:  $J^*(3, 1) = 0$ . Esto tiene sentido porque en este problema sobreentendemos que si el motorista se encuentra en el punto de destino no hay ningún coste. Evidentemente, tal y como ha sido formulado el problema lo que se busca es el valor  $J^*(0, 0)$ .

Vistos los ejemplos precedentes, para aplicar el principio de optimalidad a los problemas de control óptimo debemos hacer la siguiente traducción: Una “decisión” es el valor del control  $u(t)$  para cada  $t = 0, 1, \dots, N-1$  y “la política” es la sucesión completa de controles  $\{u(t)\}_{t=0}^{N-1}$ . Así pues, el principio de optimalidad de Bellman quedaría: *Una sucesión de controles óptima  $\{u^*(t)\}_{t=0}^{N-1}$  tiene la propiedad de que para cualquier instante  $k \geq 0$  y cualesquiera que sean el estado y el control en ese instante  $x^*(k) = x_k^*$  y  $u^*(k)$ , la sucesión de los controles a partir de  $k$ ,  $\{u^*(t)\}_{t=k}^{N-1}$  debe ser la solución del problema de control óptimo con la condición inicial  $x(k) = x_k^*$ . En otras palabras,  $\{u^*(t)\}_{t=k}^{N-1}$  es la*

solución del problema:

$$\begin{aligned} & \text{minimizar } J(u) = K(x(N)) + \sum_{t=k}^{N-1} L_t(x(t), u(t)) \\ & \text{sujeto a } \begin{cases} x(t+1) = f_t(x(t), u(t)), & t = k, k+1, \dots, N-1 \\ x(k) = x_k \end{cases} \end{aligned} \quad (7.9)$$

Es evidente que la solución de este problema es  $u^*(k), u^*(k+1), \dots, u^*(N-1)$ , siendo los estados óptimos  $x^*(k), x^*(k+1), \dots, x^*(N)$ , porque si la solución fuera  $u^\dagger(k), u^\dagger(k+1), \dots, u^\dagger(N-1)$  entonces la sucesión de controles  $u^*(0), u^*(1), \dots, u^*(k-1), u^\dagger(k), u^\dagger(k+1), \dots, u^\dagger(N-1)$  daría un valor menor de  $J$  que  $u^*(0), u^*(1), \dots, u^*(N-1)$ . Es decir, ésta no sería la sucesión de controles que minimizaría la función de coste.

La fórmula (7.9) nos da la **función valor óptimo** del problema:

$$J_k^*(x) = \min_{u(k), \dots, u(N-1) \in U} \left\{ K(x(N)) + \sum_{t=k}^{N-1} L_t(x(t), u(t)) \mid x(k) = x, x(t+1) = f_t(x(t), u(t)), t = k, k+1, \dots, N-1 \right\}$$

Para obtener la ecuación de optimalidad vamos a desarrollar esta expresión. En primer lugar, el coste de alcanzar el estado final  $x_f = x_N$  es  $K(x_N)$ . Por ello convenimos que el valor de  $J_N^*(x)$  es constante e igual a

$$J_N^*(x) = K(x_N), \quad x = x_0, x_1, \dots, x_N.$$

En particular,  $J_N^*(x_N) = K(x_N)$ . A continuación calculamos  $J_{N-1}^*(x)$  que es el coste mínimo de alcanzar el estado  $x(N)$  cuando el sistema se encuentra en el estado  $x$  en el instante  $N-1$ :

$$\begin{aligned} J_{N-1}^*(x) &= \min_{u(N-1) \in U} \left\{ K(x(N)) + L_{N-1}(x(N-1), u(N-1)) \mid x(N-1) = x, x(N) = f_{N-1}(x(N-1), u(N-1)) \right\} \\ &= \min_{u(N-1) \in U} \{ J_N^*(f_{N-1}(x, u(N-1))) + L(x, u(N-1)) \}, \end{aligned} \quad (7.10)$$

donde hemos usado que  $x(N-1) = x$  y que, siendo  $x(N) = f_{N-1}(x(N-1), u(N-1)) = f_{N-1}(x, u(N-1))$  y  $J_N^*(x_N) = K(x(N))$ , se tiene  $K(x(N)) = J_N^*(f_{N-1}(x, u(N-1)))$ . Ahora,

$$\begin{aligned} J_{N-2}^*(x) &= \min_{u(N-2), u(N-1) \in U} \left\{ K(x(N)) + L_{N-1}(x(N-1), u(N-1)) \right. \\ &\quad \left. + L_{N-2}(x(N-2), u(N-2)) \mid x(N-2) = x, x(t+1) = f_t(x(t), u(t)), \right. \\ &\quad \left. t = N-2, N-1 \right\} \end{aligned}$$

Pero si  $x(N-2) = x$  entonces  $x(N-1) = f_{N-2}(x, u(N-2))$ . Así pues, teniendo en cuenta (7.10),

$$\begin{aligned} J_{N-2}^*(x) &= \min_{u(N-2) \in U} \left\{ \min_{u(N-1) \in U} \left\{ K(x(N)) + L_{N-1}(x(N-1), u(N-1)) \right| \right. \\ &\quad \left. x(N-1) = f_{N-2}(x(N-2), u(N-2)), x(N) = f_{N-1}(x(N-1), u(N-1)) \right\} \\ &\quad \left. + L_{N-2}(x(N-2), u(N-2)) \right| x(N-2) = x, x(N-1) = f_{N-2}(x(N-2), u(N-2)) \Big\} \\ &= \min_{u(N-2) \in U} \left\{ J_{N-1}^*(f_{N-2}(x(N-2), u(N-2))) + L_{N-2}(x(N-2), u(N-2)) \right| x(N-2) = x \Big\} \\ &= \min_{u(N-2) \in U} \left\{ J_{N-1}^*(f_{N-2}(x, u(N-2))) + L_{N-2}(x, u(N-2)) \right\}. \end{aligned}$$

Usando los mismos argumentos para  $t = N-2, N-1, N$  obtendríamos

$$J_{N-3}^*(x) = \min_{u(N-3) \in U} \left\{ J_{N-2}^*(f_{N-3}(x, u(N-3))) + L_{N-3}(x, u(N-3)) \right\}. \quad (7.11)$$

Continuando de esta manera hacia atrás obtendríamos la **ecuación de optimalidad**: Para  $t = N-1, N-2, \dots, 0, 1$ :

$$J_t^*(x) = \min_{u(t) \in U} \left\{ J_{t+1}^*(f_t(x, u(t))) + L_t(x, u(t)) \right\}, \quad (7.12)$$

donde  $f_{N-1}(x(N-1), u(N-1)) = x(N) = x_f$  es un dato del problema y  $J_N^*(x(N)) = K(x(N))$ .

Todo el proceso anterior se puede resumir en el siguiente teorema.

**Teorema 7.2** Sean  $x_0, x_f \in \mathbb{R}^n$  vectores dados y consideremos el problema de control óptimo en tiempo discreto (7.8) donde  $x(N) = x_f$ . Definamos  $J_N^*(x(N)) = K(x(N))$  y supongamos que para  $t = N-1, N-2, \dots, 1, 0$ , el problema

$$\min_{u(t) \in U} \left\{ L_t(x(t), u(t)) + J_{t+1}^*(f_t(x(t), u(t))) \right\},$$

tiene una única solución  $u(t) = h_t(x(t))$  y sea  $J_t^*(x(t))$  el valor mínimo de dicho problema. Entonces la sucesión dada por

$$\begin{aligned} x^*(0) &= x_0 \\ u^*(t) &= h_t(x(t)), & t &= 0, 1, \dots, N-1 \\ x^*(t+1) &= f_t(x^*(t), u^*(t)), & t &= 0, 1, \dots, N-1 \end{aligned} \quad (7.13)$$

minimiza la función de coste  $J(u)$  y su valor mínimo es  $J_0^*(x_0)$ .

Si el problema consiste en maximizar en vez de minimizar, se puede cambiar máximo por mínimo en todo lo anterior o, simplemente, cambiar el signo de la función y minimizar (haciendo los cambios de signo apropiados al terminar el proceso).



**Ejemplo 7.3** Se posee una cantidad  $A$  de recursos que se pueden asignar a  $N$  distintas actividades. La cantidad asignada a la  $k + 1$ -ésima actividad es  $u(k)$ ,  $k = 0, 1, \dots, N - 1$ , y el beneficio proporcionado por dicha actividad es  $g_k(u(k))$ . Se pretende maximizar el beneficio total. Supondremos que  $g_k(u) = \sqrt{u}$ .

Tenemos que resolver el problema:

$$\text{máx} \sum_{k=0}^{N-1} \sqrt{u(k)}, \text{ sujeto a } \sum_{k=0}^{N-1} u(k) = A.$$

Reformulamos el problema de la siguiente forma: supondremos que vamos asignando recursos a cada actividad sucesivamente, siendo  $x(t)$  la cantidad de recursos que nos quedan después de asignar recursos a las  $t$  primeras actividades. De esta manera, tenemos un sistema en diferencias finitas (tiempo discreto):

$$x(t+1) = x(t) - u(t), \quad t = 0, 1, 2, \dots, N-1$$

con condiciones de frontera  $x(0) = A$ ,  $x(N-1) = 0$  y función objetivo que se quiere maximizar:

$$J(u) = \sum_{t=0}^{N-1} \sqrt{u(t)}.$$

Además lo que se asigna a cada actividad nunca puede ser superior a los recursos disponibles en cada momento; i.e.  $0 \leq u(t) \leq x(t)$ .

Dado que  $K = 0$  es la función cero,  $J_N^*(x(N)) = K(x(N)) = 0$ . Pasamos entonces a calcular

$$\begin{aligned} J_{N-1}^*(x(N-1)) &= \text{máx}_{u(N-1)} \{L_{N-1}(x(N-1), u(N-1)) + J_N^*(x(N-1) - u(N-1))\} = \\ &= \text{máx}_{u(N-1)} \sqrt{u(N-1)} = \sqrt{x(N-1)} \end{aligned}$$

porque  $0 \leq u(N-1) \leq x(N-1)$ . Así  $J_{N-1}^*(x(N-1)) = \sqrt{x(N-1)}$  y se alcanza para  $u(N-1) = x(N-1)$ .

A continuación calculamos  $J_{N-2}^*(x(N-2))$ . En este caso,  $L_{N-2}(x(N-2), u(N-2)) = \sqrt{u(N-2)}$  y hay que maximizar

$$\sqrt{u(N-2)} + J_{N-1}^*(x(N-2) - u(N-2)) = \sqrt{u(N-2)} + \sqrt{x(N-2) - u(N-2)}.$$

De nuevo, como  $0 \leq u(N-2) \leq x(N-2)$ , concluimos que (mediante un simple cálculo o derivando respecto  $u(N-2)$  e igualando la derivada a 0):

$$J_{N-2}^*(x(N-2)) = \sqrt{2x(N-2)},$$

que se alcanza para  $u(N-2) = x(N-2)/2$ .

Siguiendo este proceso iterativamente, es fácil ver que  $J_t^*(x(t)) = \sqrt{(N-t)x(t)}$  y que dicho valor se obtiene para  $u(t) = x(t)/(N-t)$ ,  $t = N-1, N-2, \dots, 0$ . Así pues, el valor

óptimo de la función de coste es  $J_0^*(x(0)) = \sqrt{Nx_0} = \sqrt{NA}$  y la secuencia de controles y correspondientes estados que proporcionan dicho valor óptimo es:

$$\begin{aligned} x^*(0) &= A, & u^*(0) &= x^*(0)/N = A/N \\ x^*(1) &= x^*(0) - u^*(0) = A - A/N = (N-1)A/N, & u^*(1) &= x^*(1)/(N-1) = A/N \\ x^*(2) &= x^*(1) - u^*(1) = (N-2)A/N, & u^*(2) &= x^*(2)/(N-2) = A/N \\ &\dots \end{aligned}$$

Es decir,  $u(t) = A/N$  y  $x(t) = A(N-t)/N$  para  $t = 0, 1, \dots, N-1$ . Resultado que era esperable dado que la simetría del problema conduce a pensar que la estrategia óptima consiste en asignar la misma cantidad de recursos a cada actividad.

### 7.2.1. El problema $LQR$ discreto

El problema de control óptimo se llama  $LQR$  (Linear Quadratic Regulator problem) si el sistema de control es lineal

$$x(k+1) = A(k)x(k) + B(k)u(k) \quad (7.14)$$

y la función de coste tiene la siguiente forma:

$$J(u) = \frac{1}{2}x(N)^T H x(N) + \frac{1}{2} \sum_{k=0}^{N-1} [x(k)^T Q(k)x(k) + u(k)^T R(k)u(k) + 2u(k)^T S(k)x(k)] \quad (7.15)$$

donde todas las matrices que aparecen en la expresión de  $J$  son reales de tamaño  $n \times n$  y

- $H$  y  $Q(k)$  son simétricas semidefinidas positivas para todo  $k = 0, 1, \dots, N-1$ .
- $R(k)$  es simétrica definida positiva para todo  $k = 0, 1, \dots, N-1$ .
- $\begin{bmatrix} Q(k) & S^T(k) \\ S(k) & R(k) \end{bmatrix}$  es semidefinida positiva para todo  $k = 0, 1, \dots, N-1$ .
- $N$  es un entero positivo fijo.
- $^T$  significa trasposición.

Supondremos además (sobre todo para simplificar la notación) que las matrices  $A$ ,  $B$ ,  $Q$  y  $R$  son constantes. Aplicaremos la técnica de programación dinámica y el principio de optimalidad para dar una expresión cerrada de la solución del problema  $LQR$ . Para ello aplicaremos el Teorema 7.2. Comenzamos identificando las funciones que definen el problema:

$$K(x) = \frac{1}{2}x^T H x, \quad \text{y} \quad L_t(x, u) = \frac{1}{2}(x^T Q x + u^T R u + 2u^T S x)$$

Demostraremos por inducción regresiva sobre  $t = N, N-1, \dots, 1, 0$ , que la ecuación de optimalidad se reduce a la expresión

$$J_t^*(x) = \frac{1}{2}x^T H_t x \quad (7.16)$$

donde  $H_N = H$  y

$$H_t = Q + A^T H_{t+1} A - F_t^T (R + B^T H_{t+1} B) F_t, \quad t = 0, 1, \dots, N-1, \quad (7.17)$$

y

$$F_t = -(R + B^T H_{t+1} B)^{-1} (S + B^T H_{t+1} A). \quad (7.18)$$

Obsérvese que siendo  $H_N = H$  simétrica, todas las matrices  $H_t$  son simétricas. Además, como  $H = H_N$  y  $Q$  son semidefinidas positivas, tenemos que  $Q + A^T H_{t+1} A$  y  $B^T H_{t+1} B$  son semidefinidas positivas. Finalmente, como  $R$  es definida positiva,  $R + B^T H_{t+1} B$  es también definida positiva. Esto implica que  $F_t$  está bien definida porque  $R + B^T H_{t+1} B$  tiene inversa.

En primer lugar, para  $t = N$  tenemos  $J_N^*(x) = \frac{1}{2} x^T H_N x = \frac{1}{2} x^T H x = K(x)$  porque  $H_N = H$ . Supongamos que  $J_t^*(x) = \frac{1}{2} x^T H_t x$  y demostremos que  $J_{t-1}^*(x) = \frac{1}{2} x^T H_{t-1} x$ . Por el Teorema 7.2 tenemos:

$$\begin{aligned} J_{t-1}^*(x) &= \min_u \{L_{t-1}(x, u) + J_t^*(f_{t-1}(x, u))\} = \\ &= \min_u \{L_{t-1}(x, u) + \frac{1}{2} f_{t-1}(x, u)^T H_t f_{t-1}(x, u)\} = \\ &= \frac{1}{2} \min_u \{x^T Q x + u^T R u + 2u^T S x + (Ax + Bu)^T H_t (Ax + Bu)\} \end{aligned}$$

Escribamos

$$\begin{aligned} J_{t-1}(x, u) &= x^T Q x + u^T R u + 2u^T S x + (Ax + Bu)^T H_t (Ax + Bu) = \\ &= u^T (R + B^T H_t B) u + 2u^T (B^T H_t A + S) x + x^T (Q + A^T H_t A) x, \end{aligned} \quad (7.19)$$

donde hemos usado que  $(Ax)^T H_t B u + (Bu)^T H_t A x = x^T A^T H_t B u + u^T B^T H_t A x = (u^T B^T H_t A x)^T + u^T B^T H_t A x = 2u^T B^T H_t A x$  porque  $H_t$  es simétrica.

Teniendo en cuenta que los controles admisibles  $u = u(t)$  para cada instante  $t$  son vectores arbitrarios de  $\mathbb{R}^n$  (al contrario de lo que pasaba en el Ejemplo 7.3, en el problema  $LQR$  no hay ninguna restricción sobre los posibles controles admisibles), para buscar los puntos críticos de  $J_{t-1}(x, u)$ , respecto a  $u$ , debemos calcular los valores  $u$  para los que se verifica

$$\frac{\partial J_{t-1}}{\partial u} = \begin{bmatrix} \frac{\partial J_{t-1}}{\partial u_1} \\ \frac{\partial J_{t-1}}{\partial u_2} \\ \vdots \\ \frac{\partial J_{t-1}}{\partial u_n} \end{bmatrix} = 0.$$

Ahora bien, para cualquier matriz  $G$  real de tamaño  $n \times n$ ,  $u^T G u = \sum_{i,j=1}^n g_{ij} u_i u_j$ . Si, además

$G$  es simétrica, se comprueba fácilmente que  $\frac{\partial u^T G u}{\partial u_i} = 2[g_{i1} \ g_{i2} \ \dots \ g_{in}] u$ . Es decir,

$\frac{\partial u^T G u}{\partial u} = 2Gu$ . De forma parecida se ve que  $\frac{\partial u^T G x}{\partial u} = Gx$  (en este caso no es necesario que  $G$  sea simétrica). En consecuencia,

$$\frac{\partial J_{t-1}}{\partial u} = 2(R + B^T H_t B)u + 2(B^T H_t A + S)x. \quad (7.20)$$

Dado que  $R + B^T H_t B$  es invertible, el vector que soluciona la ecuación  $\frac{\partial J_{t-1}}{\partial u} = 0$  es

$$u^* = -(R + B^T H_t B)^{-1}(B^T H_t A + S)x = F_{t-1}x.$$

Además, la matriz de las segundas derivadas parciales se calcula sin ninguna dificultad:

$$\frac{\partial^2 J_{t-1}}{\partial u^2} = \left[ \frac{\partial^2 J_{t-1}}{\partial u_i \partial u_j} \right]_{i,j=1}^n = R + B^T H_t B,$$

que, tal y como hemos visto más arriba, es definida positiva. Es decir,  $J_{t-1}(x, u)$  alcanza en  $u^* = F_{t-1}x$  un mínimo relativo. Pero es un resultado bien conocido que para una función cuadrática el mínimo relativo es también el mínimo absoluto o global. Sustituyendo  $u^* = F_{t-1}x$  en (7.19) obtenemos:

$$\begin{aligned} J_{t-1}^*(x) &= \frac{1}{2} \min_u \{x^T Q x + u^T R u + 2u^T S x + (Ax + Bu)^T H_t (Ax + Bu)\} = \\ &= \frac{1}{2} \left( x^T F_{t-1}^T (R + B^T H_t B) F_{t-1} x + 2x^T F_{t-1}^T (B^T H_t A + S)x + x^T (Q + A^T H_t A)x \right) = \\ &= \frac{1}{2} x^t \left( Q + A^T H_t A + F_{t-1}^T (R + B^T H_t B) F_{t-1} + 2F_{t-1}^T (-(R + B^T H_t B) F_{t-1}) \right) x \end{aligned}$$

donde hemos usado que  $-(R + B^T H_t B)^{-1}(B^T H_t A + S) = F_{t-1}$ . En definitiva,

$$J_{t-1}^*(x) = \frac{1}{2} x^t \left( Q + A^T H_t A - F_{t-1}^T (R + B^T H_t B) F_{t-1} \right) x = \frac{1}{2} x^T H_{t-1} x,$$

tal y como se deseaba demostrar.

Resumimos el resultado obtenido en el siguiente teorema

**Teorema 7.4** *La ley de control óptimo para el problema LQR viene dada por el feedback lineal*

$$u^*(t) = F_t x(t) \quad (7.21)$$

obteniéndose el valor mínimo de la función de coste

$$J(u^*) = x(0)^T H_0 x(0), \quad (7.22)$$

donde las matrices  $F_t$  están definidas por la fórmula (7.18) y la sucesión de matrices  $H_t$  por la fórmula (7.17) siendo  $H_N = H$ .

El resultado del teorema muestra que el controlador se implementa mediante un feedback de estados lineal que se puede construir recursivamente pero que, en general, es dependiente del tiempo. Se puede demostrar, no obstante, que si el sistema es controlable,  $H = 0$  y  $R$  y  $Q$  son constantes, entonces el feedback que proporciona el control óptimo tiende a ser constante. Es decir,  $F_{N-k} \rightarrow F$  (matriz constante) cuando  $N \rightarrow \infty$ .

**Ejemplo 7.5** Consideremos el siguiente sistema lineal discreto

$$x(k+1) = \begin{bmatrix} 0,9974 & 0,0539 \\ -0,1078 & 1,1591 \end{bmatrix} x(k) + \begin{bmatrix} 0,0013 \\ 0,0539 \end{bmatrix} u(k)$$

con la condición inicial  $x(0) = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ . Se quiere calcular una ley de control que minimice el funcional de coste

$$J = \frac{1}{2} \sum_{k=0}^{N-1} (0,25x_1(k)^2 + 0,05x_2(k)^2 + 0,05u(k)^2)$$

Para calcular la ley de control óptimo identificamos las matrices  $A$ ,  $B$ ,  $H$ ,  $Q$ ,  $R$  y  $S$ :

$$A = \begin{bmatrix} 0,9974 & 0,0539 \\ -0,1078 & 1,1591 \end{bmatrix}, \quad B = \begin{bmatrix} 0,0013 \\ 0,0539 \end{bmatrix}$$

$$H = 0, \quad Q = \begin{bmatrix} 0,25 & 0 \\ 0 & 0,05 \end{bmatrix}, \quad R = 0,05, \quad S = 0$$

Hacer las operaciones a mano es muy penoso a poco grande que sea  $N$ . Se puede diseñar una función de MATLAB como la que sigue para hacerlo. Alternativamente se puede usar el comando `dlqr` disponible en el Toolbox de Control. Una función simple de MATLAB podría ser la siguiente:

```
function [F,J]=ilqrd(A,B,H,Q,R,S,N)
% ILQR es una version rudimentaria para implementar el feedback y matriz de
% coste minimo en el problema LQR. N es el numero de etapas

F=[];
J=H;
Hk=H;
for k=N-1:-1:0
    Fk=-(R+B'*Hk*B)^(-1)*(S+B'*Hk*A);
    F=[Fk; F];
    Hk=Q+A'*Hk*A-Fk'*(R+B'*Hk*B)*Fk;
    J=[Hk J];
end
```

Con  $N = 200$  la matriz que proporciona el coste mínimo,  $H_0$ , y el valor mínimo de  $J$  son

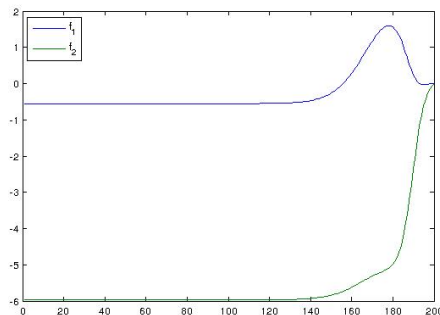


Figura 7.3: Gráfica de la evolución de las matrices de feedback del sistema del Ejemplo 7.5.

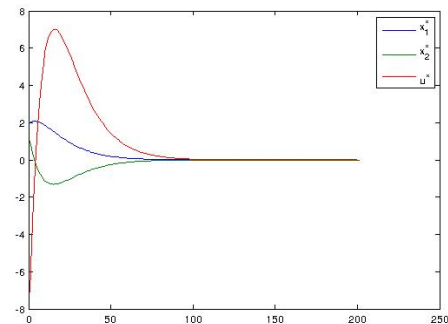


Figura 7.4: Trayectorias y control óptimo del problema en el Ejemplo 7.5.

```
>> J=J(1:2,1:2)
J =
    16.522469030619234    1.017694075776405
    1.017346891801629    6.509984904837332
>> coste_min=x0'*J*x0
coste_min =
    76.669942962470344
```

Una vez obtenida las matrices de feedback se pueden obtener fácilmente los controles y trayectoria óptimas, así como sus gráficas. Para el ejemplo que estamos trabajando, las gráficas proporcionadas se muestran en las Figuras 7.3 y 7.4. La siguiente función de MATLAB proporciona todo ello:

```
function [x u]=contraopt(A,B,H,Q,R,S,N,x0)
%CONTRAOPT proporciona los valores y graficas de los controles y
%trayectorias \óptimos que dan la solución del problema LQR discreto.

[F J]=ilqrd(A,B,H,Q,R,S,N);
[m,n]=size(B);
xk=x0;
x=[xk];
u=[];
for k=1:N
    uk=F((k-1)*n+1:k*n,:)*xk;
    xk=A*xk+B*uk;
    u=[u uk];
    x=[x xk];
end
subplot(2,1,1)
```

```

plot(1:N+1, x(1,:), 1:N+1, x(2,:), 1:N, u(1,:));
legend('x^ast_1', 'x^ast_2', 'u^ast')
subplot(2,1,2)
plot(1:N, F(:,1), 1:N, F(:,2));
legend('Location', 'NorthWest', 'f_1', 'f_2')

```

### 7.3. Programación dinámica: caso continuo

Recordemos que el problema de control óptimo en este caso es

$$\begin{aligned} & \text{minimizar } J(u) = K(t_f, x(t_f)) + \int_{t_0}^{t_f} L(t, x(t), u(t)) dt \\ & \text{sujeto a } \begin{cases} \dot{x}(t) = f(t, x(t), u(t)), & t \in [t_0, t_f] \subset \mathbb{R} \\ x(t_0) = x_0 \end{cases} \end{aligned} \quad (7.23)$$

Al igual que en el caso discreto, para aplicar la programación dinámica, definimos primero la **función valor óptimo**. Teniendo en cuenta que ahora  $t$  es una variable continua y no discreta, la función valor óptimo será:

$$J^*(t, x) = \inf_{u_{[t, t_f]}} J(t, x, u). \quad (7.24)$$

donde  $t \in [t_0, t_f]$ ,  $x \in \mathbb{R}^n$ ,  $u_{[t, t_f]}$  significa que  $u(\cdot) \in \mathcal{U}$  se restringe al intervalo  $[t, t_f]$  y donde

$$J(t, x, u) = K(t_f, x(t_f)) + \int_t^{t_f} L(s, x(s), u(s)) ds \quad (7.25)$$

siendo la función  $x(\cdot)$ , en la parte de la derecha de la ecuación, la solución del problema de condición inicial

$$\begin{cases} \dot{x}(\tau) = f(\tau, x(\tau), u(\tau)), & \tau \in [t, t_f] \subset \mathbb{R} \\ x(t) = x \end{cases}$$

Aquí hay un abuso de notación:  $x$  como argumento de  $J(t, x, u)$  es un vector cualquiera de  $\mathbb{R}^n$ , mientras que como argumento de  $L$  es una función vectorial. Además  $u$ , como argumento de  $J(t, x, u)$  es una función  $u(\cdot) \in \mathcal{U}$ ; i.e.  $u : [t, t_f] \rightarrow U \subseteq \mathbb{R}^m$ . Es sobre estas funciones sobre las que se calcula el ínfimo en (7.24). Hablando vagamente  $J^*(t, x)$  es la función de coste mínimo desde  $(t, x)$  a  $(t_f, x(t_f))$  (también llamada “optimal cost-to-go function”). Es claro que la función  $J^*(t, x)$  debe satisfacer la condición de frontera

$$J^*(t_f, x) = K(t_f, x), \quad \forall x \in \mathbb{R}^n. \quad (7.26)$$

Con estas definiciones, el **principio de optimalidad** dice lo siguiente: *Para cada  $(t, x) \in [t_0, t_f] \times \mathbb{R}^n$  y cada  $\Delta t \in (0, t_f - t]$  el valor de la función  $J^*$  definido en (7.24) cumple la siguiente condición:*

$$J^*(t, x) = \inf_{u_{[t, t+\Delta t]}} \left\{ \int_t^{t+\Delta t} L(s, x(s), u(s)) ds + J^*(t + \Delta t, x(t + \Delta t)) \right\} \quad (7.27)$$

El parecido con la ecuación (7.12) del caso discreto es manifiesto. En realidad que la función  $J^*(t, x)$  de (7.24) coincide con la de (7.27) requiere una demostración (sobre todo teniendo en cuenta que no estamos suponiendo la existencia de un mínimo). Denotemos con  $\widehat{J}^*(t, x)$  la expresión de la derecha de la igualdad (7.27):

$$\widehat{J}^*(t, x) = \inf_{u_{[t, t+\Delta t]}} \left\{ \int_t^{t+\Delta t} L(s, x(s), u(s)) ds + J^*(t + \Delta t, x(t + \Delta t)) \right\}. \quad (7.28)$$

Veamos en primer lugar que  $J^*(t, x) \geq \widehat{J}^*(t, x)$ . En efecto, por la definición de ínfimo, para cada  $t \in [t_0, t_f]$  y  $x \in \mathbb{R}^n$  y para cada  $\varepsilon > 0$  existe un control  $u_\varepsilon$  en  $[t, t_f]$  tal que

$$J^*(t, x) + \varepsilon \geq J(t, x, u_\varepsilon). \quad (7.29)$$

Si  $x_\varepsilon$  es la correspondiente solución del problema de condiciones iniciales con control  $u_\varepsilon$  y condición inicial  $x_\varepsilon(t) = x$ , tenemos

$$\begin{aligned} J(t, x, u_\varepsilon) &= \int_t^{t+\Delta t} L(s, x_\varepsilon(s), u_\varepsilon(s)) ds + J(t + \Delta t, x_\varepsilon(t + \Delta t), u_\varepsilon) \geq \\ &\geq \int_t^{t+\Delta t} L(s, x_\varepsilon(s), u_\varepsilon(s)) ds + J^*(t + \Delta t, x_\varepsilon(t + \Delta t)) \geq \widehat{J}^*(t, x). \end{aligned}$$

Dado que (7.29) se verifica para  $\varepsilon$  arbitrariamente pequeño, concluimos que  $J^*(t, x) \geq \widehat{J}^*(t, x)$ .

Recíprocamente, se sigue de la definición de  $\widehat{J}^*(t, x)$ , (7.28), que fijados  $t \in [t_0, t_f]$ ,  $\Delta t \in (0, t_f - t]$  y  $x \in \mathbb{R}^n$ , para todo  $\varepsilon > 0$  arbitrariamente pequeño existe un control  $u_\varepsilon$  en  $[t, t + \Delta t]$  tal que si  $x_\varepsilon(\cdot)$  es la solución del P.C.I.  $\begin{cases} \dot{x} = f(\tau, x, u) \\ x(t) = x \end{cases}$ ,  $\tau \in [t, t + \Delta t]$ , entonces

$$\widehat{J}^*(t, x) + \varepsilon \geq \int_t^{t+\Delta t} L(s, x_\varepsilon(s), u_\varepsilon(s)) ds + J^*(t + \Delta t, x_\varepsilon(t + \Delta t)). \quad (7.30)$$

Ahora bien,  $J^*(t + \Delta t, x_\varepsilon(t + \Delta t)) = \inf_{u_{[t+\Delta t, t_f]}} J(t + \Delta t, x_\varepsilon(t + \Delta t), u)$ . Por consiguiente, para  $\delta > 0$  arbitrariamente pequeño existe un control  $u_\delta$  en  $[t + \Delta t, t_f]$  tal que si  $x_\delta(\cdot)$  es la solución del P.C.I.  $\begin{cases} \dot{x} = f(\tau, x, u) \\ x(t + \Delta t) = x_\varepsilon(t + \Delta t) \end{cases}$ ,  $\tau \in [t + \Delta t, t_f]$ , entonces

$$J^*(t + \Delta t, x_\varepsilon(t + \Delta t)) + \delta \geq J(t + \Delta t, x_\varepsilon(t + \Delta t), u_\delta). \quad (7.31)$$

Nótese que  $x_\varepsilon(t + \Delta t) = x_\delta(t + \Delta t)$ .

Recordemos ahora que  $u_\varepsilon, u_\delta \in \mathcal{U} = PC([t_0, t_f], U)$  son funciones continuas a trozos. Para  $\tau \in [t_0, t_f]$  definimos

$$u(\tau) = \begin{cases} u_\varepsilon(\tau) & \text{si } \tau \in [t, t + \Delta t] \\ u_\delta(\tau) & \text{si } \tau \in [t + \Delta t, t_f] \end{cases}.$$

De esta forma  $u \in \mathcal{U}$  y el P.C.I.  $\begin{cases} \dot{x} = f(\tau, x, u) \\ x(t) = x \end{cases}$  admite una única solución en  $[t, t_f]$ , digamos  $x(\cdot)$ . Esta función es solución en  $[t, t + \Delta t]$  del sistema  $\dot{x} = f(\tau, x, u_\varepsilon)$  con la



condición inicial  $x(t) = x$ . Por lo tanto, debido a la unicidad de la solución,  $x(\tau) = x_\varepsilon(\tau)$  para  $\tau \in [t, t + \Delta t)$ . Un argumento similar demuestra que  $x(\tau) = x_\delta(\tau)$  para  $\tau \in [t + \Delta t, t_f]$ . Como, recordemos,  $x_\varepsilon(t + \Delta t) = x_\delta(t + \Delta t)$ , tenemos que  $x(\cdot)$  es continua en  $t + \Delta t$ . Así pues, de (7.30) y (7.31) obtenemos

$$\begin{aligned} \widehat{J}^*(t, x) + \varepsilon &\geq \int_t^{t+\Delta t} L(s, x(s), u(s)) ds + J^*(t + \Delta t, x(t + \Delta t)) \geq \\ &\geq \int_t^{t+\Delta t} L(s, x(s), u(s)) ds + J(t + \Delta t, x(t + \Delta t), u) - \delta \geq \\ &\geq J(t, x, u) - \delta \geq J^*(t, u) - \delta \end{aligned}$$

Es decir,  $\widehat{J}^*(t, x) + \varepsilon + \delta \geq J^*(t, u)$ . Y como esta desigualdad se verifica para  $\varepsilon > 0$  y  $\delta > 0$  arbitrariamente pequeños, debemos concluir que  $\widehat{J}^*(t, x) \geq J^*(t, u)$ . Por lo tanto,  $\widehat{J}^*(t, x) = J^*(t, u)$  tal y como se deseaba demostrar.

Al contrario de lo que sucedía en el caso discreto, la fórmula de  $J^*(t, x)$  en (7.27) no es operativa. Por ello, vamos a trabajar sobre ella para obtener una condición suficiente de existencia de un coste mínimo para el problema de control óptimo continuo. En primer lugar desarrollamos la función  $x(\cdot)$  en serie de Taylor en  $t$  hasta primer orden. Recordando que  $x(t) = x$  y  $\dot{x}(t) = f(t, x(t), u(t))$ :

$$x(t + \Delta t) = x + f(t, x, u(t))\Delta t + o(\Delta t),$$

donde  $o(\Delta t)$  representa los términos de la serie de Taylor que contienen potencias de  $\Delta t$  superiores a 1. Esto nos permite escribir  $J^*(t + \Delta t, x(t + \Delta t))$  de la siguiente forma (bajo la suposición, que mantendremos en adelante, de que  $J^*$  es de clase  $\mathcal{C}^1$ ):

$$\begin{aligned} J^*(t + \Delta t, x(t + \Delta t)) &= J^*(t, x) + \left[ \frac{\partial J^*}{\partial t}(t, x) \right] \Delta t + \\ &\quad + \left[ \frac{\partial J^*}{\partial x}(t, x) \right]^T (x(t + \Delta t) - x(t)) + o(\Delta t) = \\ &= J^*(t, x) + J_t^*(t, x)\Delta t + \langle J_x^*(t, x), f(t, x, u(t))\Delta t \rangle + o(\Delta t) \end{aligned} \quad (7.32)$$

donde

- $J_t^*(t, x) = \frac{\partial J^*}{\partial t}(t, x)$ .
- $J_x^*(t, x) = \frac{\partial J^*}{\partial x}(t, x) = \left[ \frac{\partial J^*}{\partial x_1}(t, x) \quad \frac{\partial J^*}{\partial x_2}(t, x) \quad \dots \quad \frac{\partial J^*}{\partial x_n}(t, x) \right]^T$ .
- $\langle \cdot, \cdot \rangle$  es el producto escalar habitual en  $\mathbb{R}^n$ . En particular,  $\langle J_x^*(t, x), f(t, x, u(t))\Delta t \rangle = J_x^*(t, x)^T f(t, x, u(t))\Delta t$ .

También tenemos (desarrollando de nuevo en serie de Taylor)

$$\int_t^{t+\Delta t} L(s, x(s), u(s)) ds = L(t, x(t), u(t))\Delta t + o(\Delta t). \quad (7.33)$$

Sustituyendo (7.32) y (7.33) en (7.27) obtenemos

$$J^*(t, x) = \inf_{u_{[t, t+\Delta t]}} \{L(t, x, u(t))\Delta t + J^*(t, x) + J_t^*(t, x)\Delta t + \langle J_x^*(t, x), f(t, x, u(t))\Delta t \rangle + o(\Delta t)\}$$

Dado que  $J^*(t, x)$  no depende de  $u$  podemos “sacarlo fuera del ínf” y simplificarlo con el  $J^*(t, x)$  que está a la izquierda de la igualdad:

$$0 = \inf_{u_{[t, t+\Delta t]}} \{L(t, x, u(t))\Delta t + J_t^*(t, x)\Delta t + \langle J_x^*(t, x), f(t, x, u(t))\Delta t \rangle + o(\Delta t)\} \quad (7.34)$$

Observamos que para valores  $s > t$  el único sumando que depende de  $u(s)$  en la expresión del mínimo es  $o(\Delta t)$ . Recordamos también que, por definición,  $\lim_{\Delta t \rightarrow 0} \frac{o(\Delta t)}{\Delta t} = 0$ . Por lo tanto, dividiendo por  $\Delta t$  y tomando el límite cuando  $\Delta t \rightarrow 0$  resulta que el mínimo se toma sobre “valores instantáneos” de  $u$  en  $t$ ; es decir, sobre los vectores  $u(t) \in U$  que abreviaremos suprimiendo el argumento  $t$ :  $u \in U$  ya que, en definitiva,  $u(t)$  puede ser cualquier vector de  $U \subseteq \mathbb{R}^m$ . Todo esto nos conduce a la siguiente ecuación que debe cumplir  $J^*(t, x)$  de acuerdo con el principio de optimalidad:

$$-J_t^*(t, x) = \inf_{u \in U} \{L(t, x, u) + \langle J_x^*(t, x), f(t, x, u) \rangle\} \quad (7.35)$$

Esta ecuación que la función  $J^*(t, x)$  debe cumplir para todo  $t \in [t_0, t_f]$  y todo  $x \in \mathbb{R}^n$  se conoce con el nombre de **ecuación de Hamilton-Jacobi-Bellman**. Es una ecuación en derivadas parciales con la condición de frontera (7.26):  $J^*(t_f, x) = K(t_f, x)$ ,  $\forall x \in \mathbb{R}^n$ .

Podemos aplicar todavía algunas transformaciones adicionales para expresar la ecuación de Hamilton-Jacobi-Bellman (HJB) de una forma más compacta. Se define el *Hamiltoniano* de las funciones  $h: \mathbb{R} \times \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n$  y  $g: \mathbb{R} \times \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$  como la siguiente función de 4 variables  $t \in \mathbb{R}$ ,  $x \in \mathbb{R}^n$ ,  $u \in \mathbb{R}^m$  y  $p \in \mathbb{R}^n$ :

$$\mathcal{H}(t, x, u, p) = \langle p, h(t, x, u) \rangle + g(t, x, u),$$

La ecuación HJB puede escribirse usando el Hamiltoniano como sigue: dado que (7.35) es equivalente

$$-J_t^*(t, x) = \inf_{u \in U} \{\langle J_x^*(t, x), f(t, x, u) \rangle + L(t, x, u)\}$$

la ecuación de Hamilton-Jacobi-Bellman puede escribirse como

$$-J_t^*(t, x) = \inf_{u \in U} \mathcal{H}(t, x, u, J_x^*(t, x)) \quad (7.36)$$

Hasta ahora no hemos asumido que la existencia de un control óptimo (es decir, que el ínfimo de la ecuación (7.35), o (7.36), se alcanza). Cuando un control óptimo (global),  $u^*$ , existe, el ínfimo en todo lo anterior se puede reemplazar por mínimo sustituyendo  $u$  por  $u^*$ . En particular, el principio de optimalidad quedaría:

$$\begin{aligned} J^*(t, x) &= \inf_{u_{[t, t_f]}} \left\{ \int_t^{t+\Delta t} L(s, x(s), u(s)) ds + J^*(t + \Delta t, x(t + \Delta t)) \right\} \\ &= \int_t^{t+\Delta t} L(s, x^*(s), u^*(s)) ds + J^*(t + \Delta t, x^*(t + \Delta t)) \end{aligned}$$

donde  $x^*(\cdot)$  y  $x(\cdot)$  son las trayectorias correspondientes a los controles  $u^*(\cdot)$  y  $u(\cdot)$ , respectivamente, que pasan, ambas, por el mismo punto inicial  $x^*(t) = x(t)$  en el instante  $t$ . En estas condiciones tenemos

$$\begin{aligned} -J_t^*(t, x) &= \min_{u \in U} \{L(t, x^*, u) + \langle J_x^*(t, x^*), f(t, x^*, u) \rangle\} \\ &= L(t, x^*(t), u^*(t)) + \langle J_x^*(t, x^*(t)), f(t, x^*(t), u^*(t)) \rangle. \end{aligned} \quad (7.37)$$

O, usando el Hamiltoniano,

$$\min_{u \in \mathcal{U}} \mathcal{H}(t, x^*, u, J_x^*(t, x^*)) = \mathcal{H}(t, x^*(t), u^*(t), J_x^*(t, x^*(t))) \quad (7.38)$$

**Ejemplo 7.6** Consideremos la ecuación  $\dot{x} = u$  ( $x, u \in \mathbb{R}$ ) y sea  $L(x, u) = x^4 + u^4$ . La ecuación HJB en este caso es

$$-J_t^*(t, x) = \inf_{u \in \mathbb{R}} \{x^4 + u^4 + J_x^*(t, x)u\} \quad (7.39)$$

Como la expresión en el interior del  $\inf$  es un polinomio en  $u$ , podemos encontrar fácilmente el control donde se alcanza el ínfimo: derivando respecto a  $u$  obtenemos  $4u^3 + J_x^*(t, x) = 0$ . Así,

$$u = -\left(\frac{1}{4}J_x^*(t, x)\right)^{1/3}. \quad (7.40)$$

Sustituyendo el valor de  $u$  en la ecuación HJB (7.39):

$$-J_t^*(t, x) = x^4 - 3\left(\frac{1}{4}J_x^*(t, x)\right)^{4/3}. \quad (7.41)$$

Esta es una ecuación diferencial en derivadas parciales que, suponiendo que la pudiéramos resolver, nos proporcionaría una función  $J^*(t, x)$  con la que obtendríamos un control óptimo  $u^*(t) = -\left(\frac{1}{4}J_x^*(t, x^*(t))\right)^{1/3}$  como una *feedback* de estados. Más explícitamente: una vez resuelta la ecuación (7.41) obtendríamos una función solución  $V(t, x)$ . Derivándola respecto a  $x$  y sustituyéndola la función obtenida en el lugar de  $J_x^*(t, x)$  en (7.40) conseguiríamos una función  $u(t, x)$ . Ésta se insertaría en la ecuación diferencial dada  $\dot{x}(t) = V_x(t, x)$  que una vez resuelta con la condición inicial que nos dieran  $x(t_0) = x_0$ , nos proporcionaría la trayectoria óptima  $x^*(t)$ . El control óptimo sería, como se ha dicho,  $u^*(t) = -\left(\frac{1}{4}J_x^*(t, x^*(t))\right)^{1/3}$ .

Hasta ahora hemos establecido que la función de coste del problema de control óptimo debe satisfacer necesariamente la condición (7.35) (o equivalentemente (7.36)) y que si existe una función de control para la que el ínfimo de estas ecuaciones se alcanza, entonces el Hamiltoniano debe cumplir (7.38). Vamos a dar ahora una condición suficiente de optimalidad en el siguiente teorema.

**Teorema 7.7** Sea  $J^* : [t_o, t_f] \times \mathbb{R}^n \rightarrow \mathbb{R}$  una función de clase  $\mathcal{C}^1$  (con derivadas parciales continuas) que satisface la ecuación de Hamilton-Jacobi-Bellman (7.35) (o equivalentemente (7.36)) y la condición de frontera (7.26)  $J^*(t_f, x) = K(t_f, x)$ . Supongamos que para cada  $t \in \mathbb{R}$  y para cada  $x \in \mathbb{R}^n$  el mínimo de la ecuación HJB (7.35) con la condición de frontera dada se alcanza para una función  $u = h(t, x)$  y sea  $x^*(t)$  la única solución del problema de condición inicial

$$\begin{cases} \dot{x}(t) = f(t, x, h(t, x)) \\ x(t_0) = x_0 \end{cases}$$

Si  $u^*(t) = h(t, x^*(t))$ , entonces  $u^*(\cdot)$  es la función de control óptimo,  $x^*(\cdot)$  es la trayectoria óptima y  $J^*(t_0, x_0)$  es el coste óptimo; i.e., el valor óptimo de  $J(u)$ .

**Ejemplo 7.8** Consideremos el sistema  $\dot{x}(t) = x(t) + u(t)$  ( $x, u \in \mathbb{R}$ ) y supongamos que queremos minimizar el funcional  $J(u) = \frac{1}{4}x(T)^2 + \int_0^T \frac{1}{4}u(t)^2 dt$  siendo  $T$  el tiempo final.

Así,  $K(t_f, x) = \frac{1}{4}x(T)^2$  es el coste final (constante) y  $L(t, x, u) = \frac{1}{4}u(t)^2$  y la función de coste de  $(t, x)$  a  $(T, x(T))$  para cada  $(t, x) \in \mathbb{R} \times \mathbb{R}$  debe satisfacer la ecuación HJB (7.36):

$$-J_t^*(t, x) = \min_{u \in \mathbb{R}} \mathcal{H}(t, x, u, J_x^*(t, x)) = \min_{u \in U} \left\{ (x + u)J_x^*(t, x) + \frac{1}{4}u^2 \right\}$$

Para hallar el mínimo de  $\mathcal{H}$  derivamos respecto a  $u$  e igualamos a 0:

$$\frac{\partial \mathcal{H}}{\partial u} = \frac{1}{2}u + J_x^*(t, x) = 0.$$

Como, además,  $\frac{\partial^2 \mathcal{H}}{\partial u^2} = \frac{1}{2} > 0$ , tenemos que el mínimo de  $\mathcal{H}(t, x, u, J_x^*)$  se alcanza para

$$u^*(t) = -2J_x^*(t, x).$$

Sustituyendo esta función en la ecuación HJB vemos que la función  $J^*(t, x)$  debe satisfacer la ecuación en derivadas parciales

$$0 = J_t^* + \frac{1}{4}(-2J_x^*)^2 + J_x^*(x - 2J_x^*) = J_t^* - J_x^{*2} + xJ_x^* \quad (7.42)$$

con la condición  $J^*(T, x(T)) = \frac{1}{4}x(T)^2$ .

El problema que estamos resolviendo es un ejemplo de la versión continua del problema LQR de la Sección 7.2.1. Recordando que las matrices de coste en aquel caso tenían la forma  $J^*(t, x) = \frac{1}{2}x^T H(t)x$  vamos a probar con  $J^*(t, x) = \frac{1}{2}x^T P(t)x$  para (7.42). Tenemos entonces

$$J_t^*(t, x) = \frac{1}{2}\dot{P}(t)x, \quad J_x^*(t, x) = P(t)x,$$

que sustituidos en (7.42) produce:

$$0 = \frac{1}{2}\dot{P}(t)x^2 - P(t)^2x^2 + P(t)x^2$$

para todo  $x$ . Esto implica que  $P(t)$  debe ser solución de la ecuación diferencial

$$\frac{1}{2}\dot{P} - P^2 + P = 0 \quad (7.43)$$

y cumplir la condición  $J^*(T, x(T)) = \frac{1}{4}x(T)^2$ . Es decir,  $P(T) = \frac{1}{2}$ . La ecuación (7.43) es de tipo Ricatti y éstas, cuando son de coeficientes variables, no son fácilmente resolubles. En este caso, sin embargo, es de coeficientes constantes y se puede resolver separando las variables. En este ejemplo, la solución general de la ecuación es

$$P(t) = \frac{1}{1 - ce^{2t}}$$

donde  $c$  es una constante no nula. Imponiendo la condición  $P(T) = \frac{1}{2}$  llegamos a

$$P(t) = \frac{1}{1 + e^{2(t-T)}}.$$

El control óptimo es entonces

$$u^*(t) = -2J_x^*(t, x)x(t) = -2P(t)x(t) = -\frac{2}{1 + e^{2(t-T)}}x(t).$$

Se trata de un feedback de estados lineal que varía con  $t$ . El coste óptimo será  $J(u^*) = \frac{1}{2}P(t_0)x_0^2 = \frac{1}{2(1 + e^{2(t_0-T)})}x_0^2$  siendo  $x(t_0) = x_0$  la condición inicial de la ecuación.

Tal y como se ha mencionado, el Ejemplo 7.8 es un caso particular de la versión continua del problema  $LQR$  visto en la Sección 7.2.1. En la siguiente sección lo desarrollamos en completa generalidad.

## 7.4. El problema $LQR$ continuo

Al igual que en el caso discreto, el problema de control óptimo continuo se llama  $LQR$  si el sistema de control es lineal:

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) \quad (7.44)$$

y el funcional de coste es (por sencillez notacional suprimimos los  $\frac{1}{2}$  habituales (ver (7.15))):

$$J(u) = x(t_f)^T H x(t_f) + \int_{t_0}^{t_f} (x(t)^T Q(t)x(t) + u(t)^T R(t)u(t)) dt, \quad (7.45)$$

donde, para todo  $t \in [t_0, t_f]$ ,  $x(t) \in \mathbb{R}^n$ ,  $u(t) \in \mathbb{R}^m$ ,  $H, Q(t) \in \mathbb{R}^{n \times n}$ ,  $R(t) \in \mathbb{R}^{m \times m}$ ,  $H$  y  $Q(t)$  son simétricas semidefinidas positivas y  $R(t)$  es simétrica definida positiva.

Si  $Q(t) = I_n$ ,  $x(t)^T x(t) = \|x(t)\|_2^2$  indicaría lo lejos que está el estado del sistema en el instante  $t$  del estado cero ( $x(t) = 0$ ). Siendo  $Q(t) \neq I_n$ ,  $x(t)^T Q(t) x(t)$  representa el mismo criterio pero asignando a cada componente del estado en cada instante  $t$  un determinado peso. La integral  $\int_{t_0}^{t_f} x(t)^T Q(t) x(t) dt$  sería una medida de la desviación de  $x(t)$  del estado cero a lo largo del intervalo  $[t_0, t_f]$  habiendo asignado a cada componente del estado en cada instante  $t$  un determinado peso. Encontrar un control del sistema (7.44) que minimice esta cantidad equivale a hacer que el sistema se aproxime al estado cero lo más rápidamente posible; es decir, instantáneamente. Recordamos que el control de norma  $L_2$  mínima que lleva el estado de la posición  $x(t_0) = x_0$  al estado  $x(t_1) = 0$  es (véase (6.9)):

$$u(t) = -B(t)^T \Phi(t_1, t)^T W(t_1, t_0)^{-1} \Phi(t_1, t_0) x_0.$$

Un simple cálculo muestra que

$$\|u(\cdot)\|_2^2 = \int_{t_0}^{t_1} u(t)^T u(t) dt = x_0^T \Phi(t_1, t_0)^T W(t_1, t_0)^{-1} \Phi(t_1, t_0) x_0.$$

Ahora bien, por una parte la matriz Grammiana de controlabilidad (véase (6.7))  $W(t_1, t_0) = \int_{t_0}^{t_1} \Phi(t_1, t) B(t) B(t)^T \phi(t_1, t)^T dt$  es una matriz simétrica definida positiva para todo  $t_1 > t_0$ . Esto implica que existe una matriz ortogonal  $V$  tal que  $V^T W(t_1, t_0) V = D = \text{Diag}(\lambda_1, \dots, \lambda_n)$  siendo  $\lambda_1 \geq \dots \geq \lambda_n > 0$  los valores propios de  $W(t_1, t_0)$ . Por otra parte, siempre se puede escoger  $t_1 \approx t_0$  de modo que los elementos de  $W(t_1, t_0)$  sean muy pequeños. Esto conlleva (por la continuidad de los valores propios de una matriz respecto a los elementos de ésta) que los valores propios de  $W(t_1, t_0)$  serían tan pequeños como se quisiera para valores de  $t_1$  suficientemente próximos a  $t_0$ . Por lo tanto, los valores propios de  $W(t_1, t_0)^{-1}$  serían muy grandes y para ciertos valores del estado inicial  $x_0$  tendríamos que

$$\|u(\cdot)\|_2^2 = x_0^T \Phi(t_1, t_0) V^T D^{-1} V \Phi(t_1, t_0) x_0$$

podría llegar a ser tan grande como se quisiera (a base de aproximar  $t_1$  a  $t_0$ ). En otras palabras llevar el estado inicial del sistema al estado cero instantáneamente puede requerir controles de amplitud indefinidamente grande. El término  $\int_{t_0}^{t_f} u(t)^T R(t) u(t) dt$  en la función de coste  $J(u)$  representa una limitación en la amplitud de los controles admisibles.

Para el problema  $LQR$  tenemos que

$$L(t, x, u) = x(t)^T Q(t) x(t) + u(t)^T R(t) u(t), \quad K(t_f) = x(t_f)^T H x(t_f)$$

y el Hamiltoniano es

$$\begin{aligned} \mathcal{H}(t, x, u, J_x^*) &= L(t, x, u) + \langle J_x^*, (A(t)x + B(t)u) \rangle = \\ &= x^T Q(t) x + u^T R(t) u + J_x^{*T} A(t) x + J_x^{*T} B(t) u \end{aligned} \quad (7.46)$$

Para calcular el vector  $u \in \mathbb{R}^m$  donde  $\mathcal{H}$  alcanza el mínimo, derivamos respecto de  $u$  (recuérdese el desarrollo que condujo a (7.20)):

$$\frac{\partial \mathcal{H}}{\partial u} = 2R(t)u + B(t)^T J_x^*$$

Además

$$\frac{\partial^2 \mathcal{H}}{\partial u^2} = 2R(t)^T = 2R(t) > 0$$

porque  $R(t)$  es simétrica definida positiva. Por lo tanto,  $\mathcal{H}$  tiene un mínimo en

$$u^*(t) = -\frac{1}{2}R(t)^{-1}B(t)^T J_x^*(t, x)$$

Al igual que se hizo en la ecuación (??),  $u^*$  depende tanto de  $t$  como de  $x$ , aunque omitiremos esta dependencia en la notación. Sustituyendo  $u^*$  en la ecuación HJB (7.36) obtenemos

$$\begin{aligned} 0 &= J_t^* + x^T Q(t)x + u^*(t)^T R(t)u^*(t) + J_x^{*T} A(t)x + J_x^{*T} B(t)u^*(t) = \\ &= J_t^* + x^T Q(t)x - \frac{1}{4} J_x^{*T} B(t)R(t)^{-1}B(t)^T J_x^* + J_x^{*T} A(t)x \end{aligned} \quad (7.47)$$

que es una ecuación diferencial en derivadas parciales que debe satisfacer la función  $J^*(t, x)$  junto a la condición de frontera  $J^*(t_f, x(t_f)) = x(t_f)^T H x(t_f)$ .

Al igual que hicimos en el Ejemplo 7.8 intentaremos ver bajo qué condiciones existe una solución de la forma

$$J^*(t, x) = x^T P(t)x$$

con  $P(t)$  una matriz simétrica definida positiva para todo  $t \in [t_0, t_f]$  y  $x \in \mathbb{R}^n$ . Tenemos  $J_t^*(t, x) = x^T \dot{P}(t)x$  y  $J_x^*(t, x) = 2P(t)x$ , que sustituidas en (7.47) producen:

$$0 = x^T \dot{P}(t)x + x^T Q(t)x + 2x^T P(t)A(t)x - x^T P(t)B(t)R(t)^{-1}B(t)^T P(t)x \quad (7.48)$$

Ahora bien, si  $S, Z \in \mathbb{R}^{n \times n}$  podemos escribir  $SZ = \frac{1}{2}[SZ + (SZ)^T] + \frac{1}{2}[SZ - (SZ)^T]$  y entonces

$$\begin{aligned} x^T SZx &= \frac{1}{2}x^T [SZ + (SZ)^T]x + \frac{1}{2}[x^T SZx - x^T (SZ)^T x] = \\ &= \frac{1}{2}x^T [SZ + (SZ)^T]x + \frac{1}{2}[x^T SZx - (x^T SZx)^T] = \frac{1}{2}x^T [SZ + (SZ)^T]x, \end{aligned}$$

donde hemos usado que  $(x^T SZx)^T = x^T SZx$  por ser un número real. Por consiguiente, (7.48) es equivalente a (recordando que  $P(t)$  es simétrica):

$$\begin{aligned} 0 &= x^T \dot{P}(t)x + x^T Q(t)x + x^T P(t)A(t)x + x^T A(t)^T P(t)x - \\ &\quad - x^T P(t)B(t)R(t)^{-1}B(t)^T P(t)x = \\ &= x^T (\dot{P}(t) + Q(t) + P(t)A(t) + A(t)^T P(t) - P(t)B(t)R(t)^{-1}B(t)^T P(t))x \end{aligned} \quad (7.49)$$

Como se debe cumplir esta identidad para todo  $x \in \mathbb{R}^n$  concluimos que la función  $J^*(t, x) = x^T P(t)x$  es solución de (7.47) si y sólo si  $P(t)$  es solución de la siguiente ecuación diferencial ordinaria:

$$\dot{P}(t) + Q(t) + P(t)A(t) + A(t)^T P(t) - P(t)B(t)R(t)^{-1}B(t)^T P(t) = 0. \quad (7.50)$$

Esta es una ecuación de tipo Ricatti y se le llama **Ecuación de Ricatti** para el problema de control óptimo  $LQR$ . Su solución queda determinada de forma única imponiendo la condición  $J^*(t_f, x(t_f)) = x(t_f)^T H x(t_f)$ ; es decir,  $P(t_f) = H$ . En general, para hallar la solución de (7.50) se deben emplear métodos numéricos, pero si todos los coeficientes ( $A$ ,  $B$ ,  $Q$  y  $R$ ) son constantes entonces, en ocasiones, se puede integrar simbólicamente.

**Ejemplo 7.9** En el Ejemplo 7.8 el funcional a minimizar era  $J(u) = \frac{1}{4}x(T)^2 + \int_0^T \frac{1}{4}u(t)^2 dt$  sujeto a  $\dot{x}(t) = x(t) + u(t)$ . En este caso  $H = \frac{1}{4}$ ,  $Q = 0$  y  $R = \frac{1}{4}$ ,  $A = 1$  y  $B = 1$ . La ecuación de Ricatti correspondiente ((7.50)) es

$$\dot{P}(t) + 0 + P(t) \cdot 1 + 1 \cdot P(t) - P(t) \cdot 1 \cdot \frac{1}{4} \cdot 1 \cdot P(t) = 0.$$

es decir,

$$\dot{P}(t) + 2P(t) - \frac{1}{4}P(t)^2 = 0. \quad (7.51)$$

Teniendo en cuenta que la función de coste  $J^*(t, x)$  en el desarrollo de esta sección es  $x^T P(t)x$  mientras que la elegida en el Ejemplo 7.8 fué  $\frac{1}{2}x^T P(t)x$ , vemos que las ecuaciones de Ricatti (7.51) y (7.43) son equivalentes.

**Ejemplo 7.10** Consideramos para terminar una variante del problema  $LQR$ . El objetivo es calcular un control óptimo  $u^*(\cdot)$  para que la correspondiente trayectoria  $x^*(\cdot)$  (solución del sistema  $\dot{x} = Ax + Bu^*$ ) se aproxime a (persiga la trayectoria de) una función dada  $\eta(t)$  ("tacking problem"). En este caso entonces, queremos encontrar un control que haga que la desviación de  $x(t)$  respecto a  $\eta(t)$  sea mínima a lo largo del intervalo  $[t_0, t_f]$ . Es decir,  $\eta(t)$  juega el papel de la función 0 del problema  $LQR$ . Así pues, el funcional a minimizar será:

$$J(u) = x(t_f)^T H x(t_f) + \int_{t_0}^{t_f} [(x(t) - \eta(t))^T Q(t)(x(t) - \eta(t)) + u(t)^T R(t)u(t)] dt. \quad (7.52)$$

En este caso, la forma de la función  $J^*(t, x)$  que se busca para resolver el sistema de ecuaciones en derivadas parciales resultante de la fórmula de Hamilton-Jacobi-Bellman es

$$J^*(t, x) = x^T P(t)x + x^T v(t) + z(t)$$

con  $P(t)$  simétrica,  $v(t) \in \mathbb{R}^{n \times 1}$  y  $z(t) \in \mathbb{R}$  para cada  $t \in [t_0, t_f]$ .

Veamos la aplicación de este procedimiento en un ejemplo: encontrar el control óptimo que minimiza

$$J(u) = \frac{1}{2} \int_0^1 [(x(t) - 1)^2 + u(t)^2] \quad (7.53)$$



sujeto a  $\dot{x}(t) = -x + u$  con  $x(0) = 0$ . Así pues, se trata de calcular el control óptimo  $u^*(\cdot)$  para que la trayectoria  $x^*(\cdot)$  solución del problema  $\dot{x} = -x + u^*$ ,  $x(0) = 0$ , se desvíe lo mínimo posible de la función  $\eta(t) = 1$  en el intervalo  $[0, 1]$ .

Para este problema  $L(t, x, u) = \frac{1}{2}(x-1)^2 + \frac{1}{2}u^2$  y  $\mathcal{H}(t, x, u, J_x^*) = \frac{1}{2}(x-1)^2 + \frac{1}{2}u^2 + J_x^*(-x+u)$ . Derivamos esta función respecto a  $u$  para calcular dónde se alcanza el mínimo:

$$\frac{\partial \mathcal{H}}{\partial u} = u + J_x^* = 0 \implies u^* = -J_x^*.$$

Como  $\frac{\partial^2 \mathcal{H}}{\partial u^2} > 0$ ,  $\mathcal{H}$  alcanza el mínimo (global) en  $u^* = -J_x^*$ . Sustituyendo este valor en la ecuación HJB:

$$\begin{aligned} 0 &= J_t^* + \frac{1}{2}(x-1)^2 + \frac{1}{2}J_x^{*2} + J_x^*(-x - J_x^*) = \\ &= J_t^* + \frac{1}{2}(x-1)^2 - \frac{1}{2}J_x^{*2} - xJ_x^*. \end{aligned}$$

Dado que estamos en el caso escalar ( $n = 1$ ), la función con la que intentar una solución es  $J^*(t, x) = p(t)x^2 + v(t)x + z(t)$  que sustituida en la ecuación de más arriba produce:

$$\begin{aligned} 0 &= \left( \dot{p}(t) - 2p(t)^2 - 2p(t) + \frac{1}{2} \right) x^2 + (\dot{v}(t) - 2p(t)v(t) - v(t) - 2)x + \\ &\quad + \left( \dot{z}(t) - \frac{1}{2}v(t)^2 + 1 \right). \end{aligned}$$

Como  $x \in \mathbb{R}$  es un número arbitrario, podemos ver esta identidad como dos expresiones del polinomio cero. Esto significa que los coeficientes del polinomio de la derecha son idénticamente 0. Es decir, se deben cumplir simultáneamente las tres siguientes ecuaciones diferenciales:

$$\dot{p}(t) - 2p(t)^2 - 2p(t) + \frac{1}{2} = 0 \quad (7.54)$$

$$\dot{v}(t) - (2p(t+1))v(t) - 2 = 0 \quad (7.55)$$

$$\dot{z}(t) - \frac{1}{2}v(t)^2 + 1 = 0 \quad (7.56)$$

La primera es una ecuación de Riccati de coeficientes constantes que se puede resolver por separación de variables. La condición de contorno para esta primera ecuación es  $p(1) = 0$ . Una vez resuelta, la solución se sustituye en (7.55) que es una ecuación lineal no homogénea. De nuevo la condición de contorno para esta ecuación es  $v(1) = 0$ . Una vez obtenida la solución, ésta se sustituye en (7.56) y se integra. También en este caso la condición inicial es  $z(1) = 0$ .

En todos los casos la condición de contorno es 0 porque la función de coste final  $K(t_f x(t_f)) = 0$  es este ejemplo. En general  $K(t_f x(t_f)) = (x(t_f) - \eta(t_f))^T H (x(t_f) - \eta(t_f)) = x(t_f)^T H x(t_f) - (x(t_f)^T H \eta(t_f) + \eta(t_f)^T H x(t_f)) + \eta(t_f)^T H \eta(t_f)$ . Cada sumando de esta expresión es la condición de contorno para cada una de las ecuaciones (7.56), (7.55) y (7.56), respectivamente.

Una vez obtenida  $J^*(t, x) = p(t)x^2 + v(t)x + z(t)$ , se sustituye en  $u^*(t) = -J_x^*(t, x) = 2p(t)x + v(t)$ . En el ejemplo que nos ocupa:

$$u^*(t) = \frac{1}{2} - \frac{1}{4}(1 + \sqrt{2})e^{\sqrt{2}t} - \frac{1}{4}(1 - \sqrt{2})e^{-\sqrt{2}t}.$$