CONTROL ÓPTIMO: TIEMPO DISCRETO

EDUARDO MARTÍNEZ

En multitud de ocasiones la ley de control de un determinado sistema de control se decide por criterios de maximización de beneficios o minimización de esfuerzos. A tal fin, se define una función en el espacio de trayectorias del sistema y se aplican técnicas de optimización, especialmente adaptadas a la estructura de nuestro problema. Veremos dos de estas técnicas: la programación dinámica y el principio del máximo de Pontryagin, que proporcionan visiones complementarias del problema.

1. Programación dinámica

Se considera un sistema de control en tiempo discreto

$$x_{t+1} = f_t(x_t, u_t), \qquad t = 0, 1, \ldots, N-1, \tag{Σ}$$

y una función de coste

$$\mathcal{J}\big(x_0, \{u_t\}_{t=0}^{N-1}\big) = K_N(x_N) + \sum_{t=0}^{N-1} K_t(x_t, u_t).$$

Fijado un punto inicial x_0, buscamos los controles $u_0, u_1, \ldots, u_{N-1}$ que hacen que se minimice el valor de $\mathcal{J}$, donde $x_0, x_1, \ldots, x_N$ es la trayectoria asociada a dichos controles mediante (Σ).

En las expresiones anteriores, $f_t(x,u)$ y $K_t(x,u)$ son funciones definidas en algún subconjunto $\mathbb{R}^n \times \mathbb{R}^p$, para $t = 0, 1, \ldots, N-1$, con valores en $\mathbb{R}^n$ y $\mathbb{R}$, respectivamente, mientras que $K_N(x)$ es una función real definida en algún subconjunto $\mathbb{R}^n$.

Denotemos por $u_0^*, u_1^*, \ldots, u_{N-1}^*$ dichos controles óptimos, que supondremos únicos, y por $x_0^*, x_1^*, \ldots, x_N^*$ la trayectoria asociada, donde obviamente $x_0^* = x_0$.

La clave del método de programación dinámica está en el siguiente razonamiento, que se conoce con el nombre de ***principio de optimalidad*** de Bellman. Consideramos para cada $0 \leq i < N$ el problema

$$\begin{aligned} \text{minimizar} \quad & \mathcal{J}_i\big(x_i^*, \{u_t\}_{t=i}^{N-1}\big) = K_N(x_N) + \sum_{t=i}^{N-1} K_t(x_t, u_t) \\ \text{sujeto a} \quad & x_{t+1} = f_t(x_t, u_t), \qquad t = i, i+1, \ldots, N-1 \\ & x_i = x_i^*. \end{aligned}$$

Date: Febrero 2011, Febero 2014.

Es evidente que la solución a dicho problema es $u_i^*, u_{i+1}^*, \ldots, u_{N-1}^*$, siendo la trayectoria asociada $x_i^*, x_{i+1}^*, \ldots, x_N^*$, ya que en caso contrario si $u_i^+, u_{i+1}^+, \ldots, u_{N-1}^+$ minimiza $\mathcal{J}_i$ entonces $u_0^*, u_1^*, \ldots, u_{i-1}^*, u_i^+, u_{i+1}^+, \ldots, u_{N-1}^+$ nos daría un menor valor de $\mathcal{J}$ que la sucesión de controles óptimos $u_0^*, u_1^*, \ldots, u_{N-1}^*$.

Teniendo esto en cuenta, para utilizar el método de programación dinámica definimos la sucesión de funciones reales $V_t\colon \mathbb{R}^n \to \mathbb{R}$, para $t = 0, \ldots, N$, por medio de la recurrencia regresiva

$$V_N(x) = K_N(x), \qquad\qquad V_t(x) = \min_u \{K_t(x,u) + V_{t+1}(f_t(x,u))\}.$$

En ocasiones escribiremos $J_t(x,u)$ para la función a minimizar en cada paso

$$J_t(x,u) = K_t(x,u) + V_{t+1}(f_t(x,u)).$$

Teorema 1.1: *Supongamos que para cada $t = N-1, N-2, \ldots, 2, 1, 0$ y para cada $x \in \mathbb{R}^n$, el problema*

$$\min_u \{K_t(x,u) + V_{t+1}(f_t(x,u))\}$$

tiene solución única $u = h_t(x)$, y sea $V_t(x)$ el valor mínimo de dicho problema. Entonces la sucesión dada por

$$\begin{aligned} x_0^* &= x_0 \\ u_t^* &= h_t(x_t^*) \qquad\qquad\qquad t = 0, 1, \ldots, N-1, \\ x_{t+1}^* &= f_t(x_t^*, u_t^*) \end{aligned}$$

minimiza la función de coste total $\mathcal{J}\big(x_0, \{u_t\}_{t=0}^{N-1}\big)$ y su valor mínimo es $V_0(x_0)$.

Dem. Los dos últimos sumandos de la función $\mathcal{J}$ dependen solo de u_{N-1} mientras que el resto dependen de todas las variables u_t. De esta forma

$$\begin{aligned} \mathcal{J}\big(x_0, \{u_t\}_{t=0}^{N-1}\big) &= \sum_{t=0}^{N-2} K_t(x_t, u_t) + K_{N-1}(x_{N-1}, u_{N-1}) + K_N(x_N) \\ &= \sum_{t=0}^{N-2} K_t(x_t, u_t) + K_{N-1}(x_{N-1}, u_{N-1}) + V_N(x_N) \\ &= \sum_{t=0}^{N-2} K_t(x_t, u_t) + K_{N-1}(x_{N-1}, u_{N-1}) + K_N(f_{N-1}(x_{N-1}, u_{N-1})) \\ &= \sum_{t=0}^{N-2} K_t(x_t, u_t) + J_{N-1}(x_{N-1}, u_{N-1}) \end{aligned}$$

Separando dichos términos obtenemos

$$\begin{aligned}
&\min_{u_0,u_1,\ldots,u_{N-1}} \mathcal{J}(x_0,\{u_t\}_{t=i}^{N-1}) = \\
&\qquad = \min_{u_0,u_1,\ldots,u_{N-2}} \Big\{ \sum_{t=0}^{N-2} K_t(x_t,u_t) + \min_{u_{N-1}} J_{N-1}(x_{N-1},u_{N-1}) \Big\} \\
&\qquad = \min_{u_0,u_1,\ldots,u_{N-2}} \Big\{ \sum_{t=0}^{N-2} K_t(x_t,u_t) + V_{N-1}(x_{N-1}) \Big\},
\end{aligned}$$

donde el mínimo del último sumando se obtiene para $u_{N-1} = h_{N-1}(x_{N-1})$, es decir, $V_{N-1}(x_{N-1}) = J_{N-1}(x_{N-1}, h_{N-1}(x_{N-1}))$.

Repitiendo el mismo razonamiento,

$$\begin{aligned}
&\sum_{t=0}^{N-3} K_t(x_t,u_t) + K_{N-2}(x_{N-2},u_{N-2}) + V_{N-1}(x_{N-1}) = \\
&\qquad = \sum_{t=0}^{N-3} K_t(x_t,u_t) + J_{N-2}(x_{N-2},u_{N-2})
\end{aligned}$$

y procediendo iterativamente de la misma manera

$$\begin{aligned}
&\min_{u_0,u_1,\ldots,u_{N-1}} \mathcal{J}(x_0,\{u_t\}_{t=i}^{N-1}) = \\
&\qquad = \min_{u_0,u_1,\ldots,u_{N-2}} \Big\{ \sum_{t=0}^{N-2} K_t(x_t,u_t) + V_{N-1}(x_{N-1}) \Big\} \\
&\qquad = \min_{u_0,u_1,\ldots,u_{N-2}} \Big\{ \sum_{t=0}^{N-3} K_t(x_t,u_t) + J_{N-2}(x_{N-2},u_{N-2}) \Big\} \\
&\qquad = \min_{u_0,u_1,\ldots,u_{N-3}} \Big\{ \sum_{t=0}^{N-3} K_t(x_t,u_t) + \min_{u_{N-2}} J_{N-2}(x_{N-2},u_{N-2}) \Big\} \\
&\qquad = \min_{u_0,u_1,\ldots,u_{N-3}} \Big\{ \sum_{t=0}^{N-3} K_t(x_t,u_t) + V_{N-2}(x_{N-2}) \Big\} \\
&\qquad \vdots \\
&\qquad = \min_{u_0} \Big\{ K_0(x_0,u_0) + V_1(x_1) \Big\} \\
&\qquad = \min_{u_0} J_0(x_0,u_0) \\
&\qquad = V_0(x_0),
\end{aligned}$$

donde, en cada paso, el mínimo en el último sumando se obtiene para $u_t = h_t(x_t)$, obteniéndose el valor mínimo $V_t(x_t) = J_t(x_t, h_t(x_t))$.

Por tanto, el valor mínimo de la función de coste es $V_0(x_0)$ y se obtiene para los valores de u_t indicados en el enunciado. □

Si el problema consiste en maximizar en vez de minimizar, simplemente podemos cambiar máximo por mínimo en todo lo que se ha discutido. Alternativamente, podemos cambiar el signo a la función y minimizar.

Ejemplo 1.2 (Asignación de recursos): Se posee una cantidad A de recursos que se puede asignar a diversas actividades (N actividades). La cantidad asignada a la actividad k-ésima es u_k, y el beneficio proporcionado por dicha actividad es $g_k(u_k)$. Se pretende maximizar el beneficio total.

Supondremos para simplificar que $g_k(u) = \sqrt{u}$. Tenemos que maximizar

$$\max \sum_{k=0}^{N-1} \sqrt{u_k} \qquad \text{sujeto a } \sum_{k=0}^{N-1} u_k = A.$$

Este problema es equivalente al problema de control que tiene por ecuaciones del sistema[1]

$$x_{t+1} = x_t - u_t,$$

con puntos inicial y final fijos

$$x_0 = A, \qquad x_N = 0,$$

y función objetivo

$$\mathcal{J}(x_0, \{u_t\}_{t=0}^{N-1}) = \sum_{t=0}^{N-1} \sqrt{u_t}.$$

Además tenemos las restricciones $0 \leq u_t \leq x_t$, para todo t (no puedo asignar más recursos que los que tengo).

Dado que $K_N = 0$, tenemos $V_N(x) = 0$. El primer término a calcular es $V_{N-1}(x)$ que se obtiene maximizando $\sqrt{u} + V_N(x-u) = \sqrt{u}$, cuyo valor máximo se obtiene para $u = x$ (ya que $0 \leq u \leq x$) y su valor máximo es $V_{N-1}(x) = \sqrt{x}$. Para $N-2$ tengo que maximizar $\sqrt{u} + V_{N-1}(x-u) = \sqrt{u} + \sqrt{x-u}$ cuyo valor máximo es $V_{N-2}(x) = \sqrt{2x}$, que se obtiene $u = h_{N-2}(x) = x/2$.

Es fácil ver que, en general $V_t(x) = \sqrt{(N-t)x}$, obteniéndose dicho valor para $u_t = h_t(x) = x/(N-t)$. Por tanto, la mejor estrategia es asignar $u_t = x_t/(N-t)$, y el valor óptimo es $V_0(x_0) = \sqrt{Nx_0} = \sqrt{NA}$. La asignación a realizar es $u_0 = A/N$, de donde $x_1 = x_0 - u_0 = A - A/N = (N-1)/A$, $u_1 = x_1/(N-1) = A/N$, y así sucesivamente, es decir, $u_t = A/N$ y $x_t = A(N-t)/N$.

Como era de esperar, dada la simetría de problema, la estrategia óptima consiste en asignar la misma cantidad de recursos a cada una de las actividades. ◁

[1] Esta forma de entender el problema consiste es pensar que vamos asignando recursos a cada actividad sucesivamente, siendo x_t la cantidad de recursos que nos queda después de asignar recursos a las primeras $t-1$ actividades, y siendo u_t la cantidad de recursos que se le asigna a la t-ésima actividad.

Ejemplo 1.3 (Agente comercial: compraventa de bienes): Supongamos que disponemos de un cierto número de bienes g_t y una cierta cantidad de dinero m_t. Cada día t podemos vender una cantidad de bienes u_t (comprar si es negativa) y se conoce el precio π_t por unidad de dicho bien en dicho día. Además, mantener cada unidad de nuestro bien en nuestra cartera/almacén genera un gasto s cada día. Lo que busca nuestro agente de ventas es maximizar el capital al final de un cierto número N de días. Existe además una restricción en la cantidad de producto que puede comprar/vender diariamente, siendo esta no superior a k.

Las ecuaciones que rigen la evolución del sistema son

$$\begin{cases} m_{t+1} = m_t + \pi_t u_t - s g_t \\ g_{t+1} = g_t - u_t. \end{cases}$$

La función a maximizar es $\mathfrak{J}((m_0, g_0), \{u_t\}_{t=0}^{N-1}) = m_N + \pi_N g_N$, teniendo en cuenta que $|u_t| \leq k$.

Por tanto, nuestras variables de estado son (m, g), el control es u, los costes son $K_t(m, g, u) = 0$ para $0 \leq t \leq N-1$ y $K_N(m, g) = m + \pi_N g$, y la aplicación del sistema es

$$f_t((m, g), u) = (m + \pi_t u - sg, g - u), \qquad t = 0, 1, \ldots, N-1.$$

($t = N$) La función valor óptimo en el paso $t = N$ es claramente $V_N(m, g) = m + \pi_N g$.

($t = N-1$) En el paso $t = N-1$, tenemos que maximizar

$$\begin{aligned} K_{N-1}(m, g, u) + V_N(f_{N-1}(m, g, u)) &= 0 + V_N(m + \pi_{N-1}u - sg, g - u) \\ &= m + \pi_{N-1}u - sg + \pi_N(g - u) \\ &= m + (\pi_N - s)g + (\pi_{N-1} - \pi_N)u, \end{aligned}$$

que tiene el máximo para $u = k$ si $\pi_{N-1} - \pi_N > 0$ o para $u = -k$ si $\pi_{N-1} - \pi_N < 0$, siendo el valor máximo

$$V_{N-1}(m, g) = m + (\pi_N - s)g + k|\pi_{N-1} - \pi_N|.$$

($t = N-2$) En el paso $t = N-2$ tenemos que maximizar

$$\begin{aligned} &K_{N-2}(m, g, u) + V_{N-1}(f_{N-2}(m, g, u)) = \\ &\qquad = 0 + V_{N-1}(m + \pi_{N-2}u - sg, g - u) \\ &\qquad = m + \pi_{N-2}u - sg + (\pi_N - s)(g - u) + k|\pi_{N-1} - \pi_N| \\ &\qquad = m + (\pi_N - 2s)g + k|\pi_{N-1} - \pi_N| + (\pi_{N-2} - \pi_N + s)u \end{aligned}$$

que tiene el máximo cuando $u = k$, si $\pi_{N-2} - \pi_N + s > 0$, o cuando $u = -k$, si $\pi_{N-2} - \pi_N + s > 0$, siendo dicho valor máximo

$$V_{N-2}(m, g) = m + (\pi_N - 2s)g + k|\pi_{N-1} - \pi_N| + k|\pi_{N-2} - \pi_N + s|.$$

$(t = t)$ Si se realizan algunos pasos más, es fácil inducir la fórmula

$$V_t(m,g) = m + \big(\pi_N - s(N-t)\big)g + k\sum_{i=t}^{N-1} |\pi_i - \pi_N + (N-i-1)s|,$$

que se obtiene para $u = h_t(m,g)$, donde

$$h_t(m,g) = \begin{cases} k & \text{si } \pi_t - \pi_N + (N-t-1)s > 0 \\ -k & \text{si } \pi_t - \pi_N + (N-t-1)s < 0\,. \end{cases}$$

Como ejercicio, demuéstrese por inducción que dicha fórmula es válida.

En consecuencia, el valor óptimo de la función objetivo es

$$V_0(m,g) = m + (\pi_N - sN)g + k\sum_{i=0}^{N-1} |\pi_i - \pi_N + (N-i-1)s|,$$

que se obtiene con la estrategia

$$u = h_t(m,g) = \begin{cases} k & \text{si } \pi_t - \pi_N + (N-t-1)s > 0 \\ -k & \text{si } \pi_t - \pi_N + (N-t-1)s < 0\,. \end{cases}$$

◁

Ejemplo 1.4 (Problemas LQ): Se denominan genéricamente problemas LQ aquellos problemas de control óptimo en los cuales el sistema de control es un sistema lineal (L)

$$x_{t+1} = Ax_t + Bu_t,$$

y la función objetivo es una función cuadrática (Q) en todas las variables (de estado y de control)

$$\mathcal{J}(x_0, \{u_t\}_{t=0}^{N-1}) = x_N^\mathsf{T} M x_N + \sum_{t=0}^{N-1}(x_t^\mathsf{T} Q x_t + u_t^\mathsf{T} R u_t + 2u_t^\mathsf{T} S x_t).$$

Se supondrá que la matriz R es simétrica y definida positiva, que la matriz M es simétrica y semidefinida positiva y que la matriz simétrica $[Q, S^\mathsf{T}; S, R]$ es también semidefinida positiva.

Teorema 1.5: *La ley de control óptimo para el problema LQ:*

$$\begin{aligned} &\textit{minimizar} \quad x_N^\mathsf{T} M x_N + \sum_{k=0}^{N-1}[x_t^\mathsf{T} Q x_t + u_t^\mathsf{T} R u_t + 2u_t^\mathsf{T} S x_t] \\ &\textit{sujeta a} \quad\;\; x_{t+1} = Ax_t + Bu_t, \end{aligned}$$

viene dada por el feedback lineal

$$u_t^* = h_t(x_t^*) = F_t x_t^*,$$

obteniéndose el valor mínimo de la función de coste

$$\mathcal{J}(x_0, \{u_t^*\}_{t=0}^{N-1}) = x_0^\mathsf{T} M_0 x_0,$$

donde las matrices F_t están definidas por

$$F_t = -(R + B^\top M_{t+1} B)^{-1}(S + B^\top M_{t+1} A),$$

y la sucesión de matrices M_t viene dada por la ley de recurrencia regresiva

$$\begin{cases} M_t = Q + A^\top M_{t+1} A - F_t^\top (R + B^\top M_{t+1} B) F_t \\ M_N = M. \end{cases}$$

Dem. Tenemos que

$$K_N(x) = x^\top M x \qquad \text{y} \qquad K_t(x,u) = x^\top Q x + u^\top R u + 2u^\top S x.$$

Probaremos que $V_t(x) = x^\top M_t x$, donde M_t son las matrices indicadas en el enunciado del teorema. La demostración la realizamos por inducción (regresiva) en t.

Para $t = N$ tenemos $V_N(x) = x^\top M_N x$ y $K_N(x) = x^\top M x$, que son iguales ya que $M_N = M$.

Supuesto cierto que $V_t(x) = x^\top M_t x$, veamos que $V_{t-1}(x) = x^\top M_{t-1} x$, con $u_{t-1} = F_{t-1} x$, y F_{t-1}, M_{t-1} definidas en términos de M_t como se indica en el enunciado.

Tenemos

$$\begin{aligned} J_{t-1}(x,u) &= K_{t-1}(x,u) + V_t(f_{t-1}(x,u)) \\ &= x^\top Q x + u^\top R u + 2u^\top S x + (Ax + Bu)^\top M_t (Ax + Bu) \\ &= u^\top (R + B^\top M_t B) u + 2u^\top (B^\top M_t A + S) x + x^\top (Q + A^\top M_t A) x. \end{aligned}$$

Cada una de las funciones $K_t(x,u)$ es semidefinida positiva, de donde se deduce que tanto $J_t(x,u)$ como $V_t(x)$ son también semidefinidas positivas[2]. Por tanto, cada una de las matrices M_t es semidefinida positiva. Como además R es estrictamente definida positiva, tenemos que $R + B^\top M_t B$ es definida positiva (en particular es invertible). De aquí se deduce que J_{t-1} tiene un único punto crítico (con respecto a la variable u) que es el punto de mínimo buscado. Para hallarlo, derivamos con respecto a u e igualamos a cero,

$$(R + B^\top M_t B) u + (S + B^\top M_t A) x = 0,$$

de donde se obtiene

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

Sustituyendo ahora u por $F_{t-1} x$ en J_t obtenemos el valor óptimo

$$\begin{aligned} V_{t-1}(x) &= x^\top \Big[F_{t-1}^\top (R + B^\top M_t B) F_{t-1} + 2F_{t-1}^\top (B^\top M_t A + S) + (Q + A^\top M_t A) \Big] x \\ &= x^\top \Big(-F_{t-1}^\top (R + B^\top M_t B) F_{t-1} + (Q + A^\top M_t A) \Big) x, \end{aligned}$$

[2] J_t es suma de semidefinidas positivas y V_t es el valor mínimo de una semidefinida positiva.

donde se ha tenido en cuenta que $(B^\top M_t A + S) = -(R + B^\top M_t B)F_{t-1}$. Por tanto, la matriz de la forma cuadrática V_{t-1} es M_{t-1}, lo que concluye la prueba. □

Nótese que de la demostración del teorema se deduce que las matrices M_t son todas semidefinidas positivas.

Existen multitud de ejemplos de sistemas físicos que se describen de esta manera. El resultado del teorema muestra que el controlador se implementa por medio de realimentación lineal de estados. Más aún, las matrices F_t, aunque dependan del tiempo, pueden construirse recursivamente e independientemente de la observación del sistema. ◁

2. Principio del máximo de Pontryagin

El segundo de los métodos que vamos a estudiar se obtiene al aplicar el conocido método de los multiplicadores de Lagrange al problema en cuestión. Dados dos puntos x_0 y x_N, arbitrarios pero fijos, consideremos el problema de control óptimo

$$\begin{aligned} &\text{minimizar} && \sum_{k=0}^{N-1} K_t(x_t, u_t) \\ &\text{sujeto a} && x_{t+1} = f_t(x_t, u_t). \end{aligned}$$

Para simplificar supondremos inicialmente que el dominio es un abierto en el que solo hay un punto crítico, que es el punto de mínimo buscado. Además supondremos que las ligaduras son independientes en el punto de óptimo. Veremos más adelante cómo proceder cuando esto no ocurre.

La función de Lagrange definida a partir de los datos anteriores es

$$\begin{aligned} \mathcal{L} = K_0(x_0, u_0) + K_1(x_1, u_1) + \cdots + K_{N-1}(x_{N-1}, u_{N-1}) + \\ + p_0^\top [x_1 - f_0(x_0, u_0)] + p_1^\top [x_2 - f_1(x_1, u_1)] + \cdots \\ \cdots + p_{N-1}^\top [x_N - f_{N-1}(x_{N-1}, u_{N-1})]. \end{aligned}$$

Dado que los puntos x_0 y x_N están prefijados, las variables de las que depende la función de Lagrange son $(x_1, \ldots, x_{N-1}, p_0, \ldots, p_{N-1}, u_0, \ldots, u_{N-1})$.

La derivada de $\mathcal{L}$ con respecto a u_t da

$$\frac{\partial K_t}{\partial u}(x_t, u_t) - p_t^\top \frac{\partial f_t}{\partial u}(x_t, u_t) = 0.$$

La derivada de $\mathcal{L}$ con respecto a x_t da

$$\frac{\partial K_t}{\partial x}(x_t, u_t) + p_{t-1} - p_t^\top \frac{\partial f_t}{\partial x}(x_t, u_t) = 0,$$

y finalmente, la derivada de $\mathcal{L}$ con respecto a p_t da

$$x_{t+1} - f_t(x_t, u_t) = 0.$$

Estas ecuaciones pueden reescribirse de forma adecuada como sigue. Definimos la función

$$H_t(x,p,u) = p^{\mathsf{T}} f_t(x,u) - K_t(x,u),$$

que se denomina ***Hamiltoniano de Pontryagin***. En términos de ésta, las ecuaciones para los puntos críticos son

$$\begin{aligned} 0 &= \frac{\partial H_t}{\partial u}(x_t,p_t,u_t) && t=0,\dots,N-1 \\ p_{t-1} &= \frac{\partial H_t}{\partial x}(x_t,p_t,u_t) && t=1,\dots,N-1 \\ x_{t+1} &= \frac{\partial H_t}{\partial p}(x_t,p_t,u_t) && t=0,\dots,N-1\,. \end{aligned}$$

que se denominan ***ecuaciones de Hamilton-Pontryagin***.

Es conveniente definir p_{-1} por medio de

$$p_{-1} = \frac{\partial H_0}{\partial x}(x_0,p_0,u_0) = p_0^{\mathsf{T}} \frac{\partial f_0}{\partial x}(x_0,u_0) - \frac{\partial K_0}{\partial x}(x_0,u_0).$$

De esta manera, los tres bloques de ecuaciones de Hamilton-Pontryagin se deben satisfacer para $t = 0,\dots,N-1$. Asumiremos este convenio en los que sigue.

Nota 2.1: Nótese que no se obtiene ninguna condición sobre el valor del multiplicador p_{N-1}, que debe obtenerse a partir de las ecuaciones anteriores. Por otro lado, nótese que, mientras que las ecuaciones para x son ecuaciones en diferencias progresivas, las ecuaciones para p son ecuaciones en diferencias regresivas. ◇

Nota 2.2: En el caso general de un dominio no abierto, o cuando existen otros puntos críticos de $\mathcal{L}$, un análisis más detallado del problema muestra que la condición $\frac{\partial H_t}{\partial u}(x_t,p_t,u) = 0$ debe sustituirse por la siguiente condición de optimalidad del Hamiltoniano: el valor de u_t es la solución del problema de optimización

$$\max_u H(x_t,p_t,u),$$

para los valores actuales de x_t y p_t. Las ecuaciones a resolver son entonces

$$\begin{aligned} u_t &= \arg\max_u H_t(x_t,p_t,u) \\ p_{t-1} &= \frac{\partial H_t}{\partial x}(x_t,p_t,u_t) \\ x_{t+1} &= \frac{\partial H_t}{\partial p}(x_t,p_t,u_t), \end{aligned} \tag{PMP}$$

para $t = 0,\dots,N-1$.

Obsérvese que, mientras que el problema original es un problema de minimización, en cada paso el valor del Hamiltoniano debe maximizarse con respecto a la variable u. ◇

Ejemplo 2.3 (Asignación de recursos): Volvemos a considerar el problema de la asignación de recursos, que ya estudiamos en el ejemplo 1.2, y planteamos de la forma

$$\begin{aligned} \text{minimizar} \quad & -\sum_{k=0}^{N-1} \sqrt{u_t} \\ \text{sujeto a} \quad & x_{t+1} = x_t - u_t, \end{aligned}$$

con puntos inical y final fijos $x_0 = A$ y $x_N = 0$.

La función Hamiltoniana es

$$H_t(x, p, u) = p(x - u) + \sqrt{u}.$$

La derivada con respecto a u, igualada a cero da

$$-p + \frac{1}{2\sqrt{u}} = 0.$$

La derivada con respecto a p nos da simplemente la ecuación del sistema de control

$$x_{t+1} = \frac{\partial H_t}{\partial p}(x_t, u_t) = x_t - u_t,$$

y la derivada con respecto a x nos da la evolución de las variables p_t

$$p_{t-1} = \frac{\partial H_t}{\partial x}(x_t, p_t) = p_t.$$

De esta última obtenemos que p_t es constante (no depende de t) y la ecuación para u implica que u_t también es constante $u_t \equiv u$. La solución de $x_{t+1} = x_t - u_t$ es entonces $x_t = x_0 - tu = A - tu$, e imponiendo $x_N = 0$ obtenemos $A = Nu$. Por tanto, la asignación debe ser $u = A/N$ para cada actividad. El valor óptimo es $\sum_{t=0}^{N-1} \sqrt{A/N} = \sqrt{NA}$. ◁

Nota 2.4: Hemos supuesto que las funciones de ligadura tienen, en el punto de óptimo, diferenciales linealmente independientes. Para detectar cuando esto no ocurre, podemos proceder como sigue. Una combinación lineal de las ligaduras se obtiene de la expresión de $\mathcal{L}$ si suponemos que todas las funciones K_t son nulas. Por tanto, las ecuaciones obtenidas entonces son válidas, de donde se deducen las condiciones

$$\begin{aligned} 0 &= p_t^{\mathsf{T}} \frac{\partial f_t}{\partial u}(x_t, u_t) \\ p_{t-1} &= p_t^{\mathsf{T}} \frac{\partial f_t}{\partial x}(x_t, u_t) \\ x_{t+1} &= f_t(x_t, u_t), \end{aligned}$$

para $(p_0, p_1, \ldots, p_{N-1}) \neq (0, \ldots, 0)$.

De las ecuaciones anteriores, se deduce que si $p_\tau = 0$, para algún tiempo τ, entonces $p_t = 0$ para todo $t \leq \tau$. En particular, $p_{N-1} \neq 0$ equivale a la condición anterior $(p_0, p_1, \ldots, p_{N-1}) \neq (0, \ldots, 0)$.

Estas condiciones pueden re-escribirse en términos de la función

$$H_t^0(x, p, u) = p^\mathsf{T} f_t(x, u),$$

$$\begin{aligned}
0 &= \frac{\partial H_t^0}{\partial u}(x_t, p_t, u_t) && t = 0, \dots, N-1 \\
p_{t-1} &= \frac{\partial H_t^0}{\partial x}(x_t, p_t, u_t) && t = 1, \dots, N-1 \\
x_{t+1} &= \frac{\partial H_t^0}{\partial p}(x_t, p_t, u_t) && t = 0, \dots, N-1\,,
\end{aligned}$$

con $p_{N-1} \neq 0$. Las soluciones de estas ecuaciones que minimizan nuestra función objetivo se denominan ***soluciones anormales***. ◇

Condiciones en la frontera. Debemos destacar que, en el problema general que hemos resuelto, hemos impuesto que los extremos x_0 y x_N son fijos. En muchos problemas se deben imponer otras condiciones distintas a dichos puntos o bien se añade un coste terminal a la función objetivo que penalice ciertos valores del punto final. Estás condiciones no cambian las ecuaciones obtenidas anteriormente (PMP), sino que solamente añaden condiciones a los multiplicadores inicial p_0 y final p_{N-1}.

Veamos algunos ejemplos:

- Punto final libre y coste terminal. A la función de coste $\mathfrak{J}$ se le añade $K_N(x_N)$ (como se hizo inicialmente en el estudio mediante programación dinámica). Por tanto, la función de Lagrange es ahora $\bar{\mathfrak{L}} = \mathfrak{L} + K_N(x_N)$, donde $\mathfrak{L}$ es la función de Lagrange que usamos anteriormente, y donde ahora x_N es una variable. La derivada de $\bar{\mathfrak{L}}$ con respecto a x_N igualada a cero resulta

$$p_{N-1} = -\frac{\partial K_N}{\partial x}(x_N),$$

  que nos proporciona el valor final del multiplicador p_t.
- Punto final ligado. Se impone que el punto final satisfaga una cierta ecuación $\psi(x_N) = 0$. Dado que ésta es una ligadura nueva, se debe multiplicar por un nuevo multiplicador ν y sumar a la función de Lagrange, que resulta $\bar{\mathfrak{L}} = \mathfrak{L} - \nu\psi(x_N)$, donde de nuevo x_N es una variable. La derivada de $\bar{\mathfrak{L}}$ con respecto a x_N igualada a cero da

$$p_{N-1} = \nu\frac{\partial \psi}{\partial x}(x_N),$$

  que de nuevo nos proporciona el valor final del multiplicador p_t, aunque ν debe ser determinado. En particular, si se permite que el punto final x_N sea completamente arbitrario, entonces $p_{N-1} = 0$.

- Punto inicial ligado. Si es al punto inicial a quien le imponemos que satisfaga una cierta ecuación $\psi(x_0) = 0$, el análisis es similar. Esta nueva ligadura se debe multiplicar por un nuevo multiplicador ν y sumar a la función de Lagrange, que resulta $\bar{\mathcal{L}} = \mathcal{L} - \nu\psi(x_0)$, donde x_0 no está fijado y es ahora una variable. La derivada de $\bar{\mathcal{L}}$ con respecto a x_0 igualada a cero da
$$p_{-1} = \nu\frac{\partial\psi}{\partial x}(x_0).$$
En particular, si se permite que el punto inicial x_0 sea completamente arbitrario, entonces $p_{-1} = 0$.

Ejercicio 2.5: Hallar las condiciones que se obtienen al imponer varias de las condiciones anteriores: ligadura inicial y final, ligadura final y coste terminal, etc..

Ejemplo 2.6 (Compraventa de bienes): Consideremos de nuevo el ejemplo 1.3 de un agente comercial. Como el problema original es de maximización, para adecuarlo a nuestra teoría, debemos cambiar el signo del coste terminal, de forma que $K_N(m, g) = -m - \pi_N g$, junto con $K_t = 0$ para $t = 0, \ldots, N-1$. Denotando por (p, q) a los multiplicadores, el Hamiltoniano es
$$H_t(m, g, p, q, u) = p(m + \pi_t u - sg) + q(g - u).$$
La evolución del multiplicador p viene dada por
$$p_{t-1} = \frac{\partial H_t}{\partial m}(m_t, g_t, p_t, q_t, u_t) = p_t,$$
por lo que p_t es constante. Como el valor terminal es
$$p_{N-1} = -\frac{\partial K_N}{\partial m}(m_N, g_N) = 1,$$
se deduce que $p_t = 1$ para todo $t = 0, 1, \ldots, N$. Para el momento q tenemos
$$q_{t-1} = \frac{\partial H_t}{\partial g}(m_t, g_t, p_t, q_t, u_t) = q_t - p_t s = q_t - s.$$
Como el valor terminal es
$$q_{N-1} = -\frac{\partial K_N}{\partial g}(m_N, g_N) = \pi_N$$
se deduce que $q_t = \pi_N - s(N - t - 1)$.

Sustituyendo estos valores en el Hamiltoniano se obtiene
$$\begin{aligned} H_t(m_t, g_t, p_t, q_t, u) &= m_t + \pi_t u - sg_t + (g_t - u)(\pi_N - s(N - t - 1)) \\ &= m_t + (\pi_N - s(N - t))g_t + (\pi_t - \pi_N + s(N - t - 1))u, \end{aligned}$$
que tiene un máximo con respecto a u en
$$u_t^* = \begin{cases} k & \text{si } \pi_t - \pi_N + s(N - t - 1) > 0 \\ -k & \text{si } \pi_t - \pi_N + s(N - t - 1) < 0\,. \end{cases}$$

Obteniendo, por tanto, la misma respuesta que por medio del método de programación dinámica. Para obtener el valor óptimo se deben resolver las ecuaciones de recurrencia originales con el valor de u_t^* obtenido. $\triangleleft$

Ejemplo 2.7 (Problemas LQ): Consideremos de nuevo un problema LQ como en ejemplo 1.4. El Hamiltoniano de Pontryagin es

$$H(x,p,u) = p^\intercal(Ax+Bu) - \frac{1}{2}(x^\intercal Qx + u^\intercal Ru + 2u^\intercal Sx).$$

La derivada con respecto a u es

$$B^\intercal p - Ru - Sx = 0,$$

de donde obtenemos $u = R^{-1}(B^\intercal p - Sx)$. Sustituyendo en la aplicación del sistema obtenemos

$$x_{t+1} = [A - BR^{-1}S]x_t + BR^{-1}B^\intercal p_t.$$

Derivando con respecto a x obtenemos

$$p_{t-1} = A^\intercal p_t - Qx_t - Su_t = A^\intercal p_t - Qx_t - Su_t = [A^\intercal - SR^{-1}B^\intercal]p_t + [SR^{-1}S - Q]x_t.$$

Por tanto las ecuaciones quedan

$$\begin{cases} x_{t+1} = [A - BR^{-1}S]x_t + BR^{-1}B^\intercal p_t \\ p_{t-1} = [SR^{-1}S - Q]x_t + [A^\intercal - SR^{-1}B^\intercal]p_t. \end{cases}$$

El punto final es libre, por lo que $p_{N-1} = 0$. $\triangleleft$

Relación con la programación dinámica. Comparando la condición de punto crítico para el Hamiltoniano $u \mapsto H_t(x_t, p_t, u)$ con la condición de punto crítico para la función $u \mapsto J_t(x_t, u) = K_t(x_t, u) + V_{t+1}(f_t(x_t, u))$:

$$-p_t^\intercal \frac{\partial f_t}{\partial u}(x_t, u_t) + \frac{\partial K_t}{\partial u}(x_t, u_t) = 0$$

frente a

$$\frac{\partial V_{t+1}}{\partial u}(f_t(x,u))\frac{\partial f_t}{\partial u}(x_t, u_t) + \frac{\partial K_t}{\partial u}(x_t, u_t) = 0,$$

se deduce que la relación entre ambos métodos debe venir dada por medio de

$$p_t = -\frac{\partial V_{t+1}}{\partial x}(x_{t+1}).$$

Veamos que, bajo condiciones adecuadas de regularidad de las funciones que han ido apareciendo, esto es cierto.

Teorema 2.8: *Supongamos que la función valor óptimo $V_t(x)$ es diferenciable, que el punto de máximo $u = h_t(x)$ se sitúa siempre en el* *interior de su dominio, siendo también la función h_t diferenciable. Sea (x_t^*, u_t^*) la sucesión de estados y controles óptimos, y definamos*

$$p_t^* = -\frac{\partial V_{t+1}}{\partial x}(x_{t+1}^*).$$

Entonces (x_t^*, p_t^*, u_t^*) satisface las ecuaciones de Hamilton-Pontryagin (PMP).

DEM. En efecto, por un lado, es evidente que $x_{t+1}^* = f(x_t^*, u_t^*)$. Por otro lado, u_t^* es un punto crítico de $J_t(x,u) = K_t(x,u) + V_{t+1}(f_t(x,u))$, por lo que

$$\begin{aligned}
0 &= \frac{\partial J_t}{\partial u}(x_t^*, u_t^*)\\
&= \frac{\partial K_t}{\partial u}(x_t^*, u_t^*) + \frac{\partial V_{t+1}}{\partial x}(x_{t+1}^*)\frac{\partial f_t}{\partial u}(x_t^*, u_t^*)\\
&= \frac{\partial K_t}{\partial u}(x_t^*, u_t^*) - p_t^*\frac{\partial f_t}{\partial u}(x_t^*, u_t^*)\\
&= \frac{\partial H_t}{\partial u}(x_t^*, u_t^*).
\end{aligned}$$

Antes de verificar la última ecuación, notemos lo siguiente. Para x fijo, el mínimo de $J_t(x,u)$ se obtiene en $u = h_t(x)$, por lo que $J_t(x, h_t(x)) = V_t(x)$. Derivando esta relación en $x = x_t^*$ y teniendo en cuenta que $h_t(x_t^*) = u_t^*$ obtenemos

$$\frac{\partial V_t}{\partial x}(x_t^*) = \frac{\partial J_t}{\partial x}(x_t^*, u_t^*) + \frac{\partial J_t}{\partial u}(x_t^*, u_t^*)\frac{\partial h_t}{\partial x}(x_t^*) = \frac{\partial J_t}{\partial x}(x_t^*, u_t^*).$$

Teniendo este resultado en cuenta, la ecuación para la evolución del momento es

$$\begin{aligned}
\frac{\partial H_t}{\partial x}(x_t^*, p_t^*, u_t^*) - p_{t-1}^* &= -p_{t-1}^* + p_t^*\frac{\partial f_t}{\partial x}(x_t^*, u_t^*) - \frac{\partial K_t}{\partial x}(x_t^*, u_t^*)\\
&= \frac{\partial V_t}{\partial x}(x_t^*) - \frac{\partial V_{t+1}}{\partial x}(x_{t+1}^*)\frac{\partial f_t}{\partial x}(x_t^*, u_t^*) - \frac{\partial K_t}{\partial x}(x_t^*, u_t^*)\\
&= \frac{\partial V_t}{\partial x}(x_t^*) - \frac{\partial}{\partial x}[V_{t+1}\circ f_t + K_t](x_t^*, u_t^*)\\
&= \frac{\partial V_t}{\partial x}(x_t^*) - \frac{\partial J_t}{\partial x}(x_t^*, u_t^*)\\
&= 0.
\end{aligned}$$

Esto prueba el teorema. □

EJERCICIO 2.9: Los problemas LQ fueron estudiados en el ejemplo 1.4 utilizando programación dinámica y en el ejemplo 2.7 utilizando el principio del máximo de Pontryagin. Utilizando el teorema 2.8 relacionar ambos resultados, comprobando en dicho caso que se obtiene la misma solución por ambos métodos.

Eduardo Martínez: IUMA and Departamento de Matemática Aplicada, Facultad de Ciencias, Universidad de Zaragoza, 50009 Zaragoza, Spain

E-mail address: emf@unizar.es