LA ECUACION DE HAMILTON- JACOBI BELLMAN Y ALGUNAS APLICACIONES.

Por

DIEGO BRICIO HERNANDEZ *

Se presenta la Programación Dinámica (PD), aplicada a la resolución de problemas de optimización en Teoría de Control. Se trabaja a tiempo contínuo y se consideran problemas determinísticos y no. El formalismo de la PD conduce a la ecuación en derivadas parciales conocida como de Hamilton-Jacobi-Bellman (HJB), satisfecha por la función de costo óptimo para condiciones iniciales variables. Esta ecuación es no lineal, hiperbólica de primer orden en el caso determinístico y parabólica de segundo orden en el no determinístico.

Los diferentes conceptos instroducidos se motivan en dos áreas de aplicación de gran interés en nuestro medio y en las que se requiere tomar decisiones en forma racional: a) irrigación de cultivos en tiempo de secas y b) administración del agua de una presa. Se formulan sendos modelos matemáticos con el fin de diseñar las políticas requeridas y, finalmente, se resuelven las correspondientes ecuaciones de HJB para obtener las políticas óptimas requeridas.

* Universidad Autónoma Metropolitana

1. DOS PROBLEMAS DE CONTROL.

En [7] hemos construido un modelo matemático para la variación temporal del contenido de humedad de una planta, que relaciona esta variable con la humedad atmosférica y la del suelo. Dicho modelo supone una variación aleatoria de la humedad atmosférica en torno a un valor promedio y supone además que la humedad del suelo se origina exclusivamente por la irrigación. Por ambas consideraciones, es de esperarse que el modelo en cuestión sea aplicable exclusivamente a cultivos en zonas áridas, como las que tanto abundan en el norte de México. Brevemente, su ingrediente básico es la ecuación diferencial estocástica

$$dx_t = (-\alpha x_t + \beta u_t + \delta)dt + \sigma dB_t \tag{1}$$

En esta ecuación, x y u son procesos estocásticos, que representan la variación temporal de la humedad de la planta y la humedad del suelo, respectivamente, ambas debidamente normalizadas. El proceso B no es otro que el movimiento browniano estándar unidimensional, en tanto que α, β, δ y σ son constantes positivas conocidas. A reserva de consultar las varias hipótesis subyacentes a ese modelo en la referencia [7], diremos que la ecuación (1) supone que las fluctuaciones aleatorias de la humedad atmosférica pueden representarse como un ruido blanco gaussiano y el movimiento browniano resulta precisamente de dicha suposición.

Por otro lado, sobre este modelo se plantea el problema de elegir el nivel óptimo de humedad en el suelo en cada instante, es decir, u. Para ello se define un índice

$$J(u) := E \int_0^T \{\lambda x_t^2 + \mu u_t^2\} dt \tag{2}$$

para cuantificar la bondad de un tal control u, de tal manera que un control u será óptimo si minimiza J sobre todos los otros controles que compiten y que constituyen una familia $\mathcal{U}$. En otras palabras, un control óptimo queda caracterizado por la condición

$$J(u^*) = \inf_{u \in \mathcal{U}} J(u) \tag{3}$$

En dicho ejemplo, la clase de los controles admisibles $\mathcal{U}$ está constituída por todos aquellos que pueden darse en la forma de una POLITICA: se requiere que

$$u_t = \phi(t, x_t) \tag{4}$$

donde ϕ es una función Borel medible de sus argumentos.

En otros trabajos nos hemos referido al problema de diseñar estrategias de control para una presa, encaminadas a tomar decisiones en cuanto al caudal de agua a desalojar en cada momento. Aquí el modelo matemático es aún mas simple y su ingrediente básico es la ecuación diferencial ordinaria

$$\dot{z}_t = x_t - u_t \tag{5}$$

En esta ecuación, z_t es el volumen de agua en la presa en el instante t. Por otro lado, x_t y u_t son los caudales de ingreso a y salida de la presa, respectivamente, ambos en el instante t.

Es razonable pensar que el caudal del rio que alimenta a la presa no sea facilmente predecible y de hecho suele ser bastante errático. Por otro lado, es razonable también que sí haya un cierto grado de dependencia entre los caudales que acarrea el rio en instantes cercanos; de hecho, es dable esperar variaciones estacionales más o menos periódicas. Por lo tanto, parece sensato proponer un modelo del tipo de

$$\ddot{x}_t + \omega^2 x_t = \omega^2 b + \sigma \xi_t$$

para el caudal que ingresa al embalse. Aquí, ω es la frecuencia con que se suceden las estaciones de aguas y de secas, b es el caudal promedio a lo largo del año y ξ es un ruido blanco gaussiano estándar. Procediendo de la manera usual, definimos la nueva variable de estado $y := x$, con lo que el modelo para las avenidas del río puede darse en la forma del sistema

$$dx_t = y_t dt \qquad dy_t = \omega^2 (b - x_t) dt + \sigma dB_t \tag{6}$$

En (6), como en (1), se utiliza la notación del cálculo estocástico de Ito [10] para legalizar el uso de ruido blanco en el modelo.

Finalmente, el cambio de variable

$$x' := x - b \qquad u' := u - b$$

nos permite dar el modelo global bajo la forma del sistema lineal de ecuaciones diferenciales estocásticas

$$\begin{bmatrix} dx_t \\ dy_t \\ dz_t \end{bmatrix} = \left\{ \begin{bmatrix} 0 & 1 & 0 \\ -\omega^2 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} u_t \right\} dt + \begin{bmatrix} 0 \\ \sigma \\ 0 \end{bmatrix} dB_t \tag{7}$$

Ahora bien, hay una serie de objetivos a cumplir para lograr una operación satisfactoria de este sistema de almacenamiento de agua. Entre otros, podemos listar los siguientes:

i) Debe satisfacerse la demanda de agua corriente abajo lo mejor posible.

ii) Tanto los derrames como los vaciados de la prensa deben evitarse, también dentro de lo posible

Este último requisito puede modelarse facilmente eligiendo una escala de medición tal que la capacidad total de la presa sea 2M unidades, siendo -M el mínimo volumen de agua tolerable en la presa y +M el nivel al cual ya es de temerse un derrame. Por consiguiente, se desea que el proceso z se mantenga confinado al intervalo $[-M,M]$ Si denotamos por τ_3 al instante en que primero se violan estas restricciones, es decir,

$$\tau_3 := \inf\{t \geq 0 : |z_t| > M\} \tag{8}$$

nos interesa la distribución de esta nueva variable aleatoria. En particular, quisieramos que la probabilidad

$$P(\tau_3 \leq T) \tag{8)''$$

fuera lo más pequeña posible. Aquí T es el "horizonte de planificación", es decir, el lapso de tiempo durante el cual vamos a implantar nuestras decisiones.

Por otro lado, el peligro de un exceso de agua en la presa o de un nivel de agua demasiado bajo ahí puede evitarse más facilmente si se lleva un registro del mismo caudal de ingreso. Aquí resulta importante asegurarse de que dicho caudal se mantenga dentro de ciertos límites, que sin pérdida de generalidad denotaremos como -K y +K, respectivamente. Finalmente, dicho caudal no debe variar en forma demasiado rápida, pues de ser así se dificultaría la acción de control; es razonable fijar límites también para el proceso y, que deberá estar confinado al intervalo $[-L,L]$ Por consiguiente, el estado del sistema deberá restringirse a un cubo Q en el espacio de tres dimensiones para garantizar una operación satisfactoria.

En cuanto al requisito que pide satisfacer la demanda, supondremos que ésta puede modelarse mediante un proceso estocástico d conocido, independientemente de los tres procesos estocásticos x, y y z. Aquí resulta razonable requerir que la deficiencia $|d_t \quad u_t|$ tome valores tan bajos como sea posible. Puede entonces sugerirse la conveniencia de elegir el control u de tal manera que se minimice la esperanza

$$E \int_0^T (d_t \quad u_t)^2 dt$$

Sin embargo, conviene también evitar variaciones de nivel demasiado bruscas, por lo que es conveniente también considerar el objetivo de minimizar la esperanza

$$E \int_0^T (x_t \quad u_t)^2 dt$$

Finalmente, observamos que pueden combinarse los requisitos anteriores en uno solo si consideramos como objetivo a minimizar la esperanza

$$J(u) := E \int_0^{\tau} \{\lambda(d_t - u_t)^2 + \mu(x_t \quad u_t)^2\} dt \tag{9}$$

donde λ y μ son dos números positivos que suman 1. Además

$$\tau := \inf\{t \geq 0 : (x_t, y_t, z_t) \notin Q\} \tag{10}$$

Aquí también resulta conveniente considerar controles que puedan darse bajo la forma de políticas, como en (4), es decir,

$$u_t = \phi(t, x_t, y_t, z_t)$$

Cabe también considerar restricciones en la capacidad de desagüe de la presa, por lo cual el control toma valores en algún intervalo compacto U. Así pues, cada política de control estará dada por una función Borel medible de $[0,T]\times Q$ en U. Los controles así obtenidos serán llamados ADMISIBLES, constituyendo así la familia $\mathcal{U}$ sobre la cual queda definido el índice J en (9).

Finalmente, el problema requiere encontrar un control óptimo u*, que debe desde luego satisfacer la condición (3) para serlo.

2. PROGRAMACION DINAMICA.

Cada uno de los ejemplos dados en la sección anterior condujo a un problema de optimización estocástica del siguiente tipo:

Sea dado un proceso de difusión x, solución de la ecuación diferencial estocástica

$$dx_t = b(t,x_t,u_t)dt + \sigma(t,x_t)dB_t \qquad (11)$$

para cada control u, a su vez un proceso estocástico que toma valores en un conjunto U. En general, suponemos que x es n-dimensional y que el movimiento browniano estándar B es d-dimensional, por lo que b y σ tienen dimensiones n y nxd, respectivamente. Precisando aún más, sea $\mathcal{U}$ la clase de todos los controles admisibles, a los que pediremos dos condiciones:

i) Para cada control admisible u, el problema de valores iniciales asociado a (11) tiene solución única, definida en todo $[0,T]$

ii) Cada control admisible u está dado en forma de política, es decir, se cumple (4) para $\phi:[0,T]\times\mathbb{R} \to U$ adecuado.

Por otra parte, la "vida útil" de nuestro sistema se da mientras que las trayectorias de las soluciones de (11) se mantengan en el interior de un conjunto restricción B. El instante en que las trayectorias de x abandonan B por primera vez es, precisamente,

$$\tau := \inf\{t \geq 0: x_t \notin B\},$$

con la convención extra de que $\inf \emptyset := T$ Sea x_τ el punto de la frontera de B por el que las trayectorias de x abandonan el interior de la restricción. Es bien sabido que tanto τ como x_τ son variables aleatorias, con valores en $[0,T]$ y ∂B, respectivamente.

Para cada trayectoria, sea

$$\int_0^\tau L(t,x_t,u_t)dt + \Phi(\tau,x_\tau) \tag{12}$$

el correspondiente costo de operación del sistema, sobre la vida útil del mismo. Aquí, L y Φ son funciones no negativas sobre $[0,T]\times B\times U$ y $\{T\}\times B + [0,T]\times\partial B$, respectivamente; estas funciones son datos del problema. El costo especificado en (12) es una variable aleatoria, cuyo valor esperado queremos minimizar.

Precisando, definamos J(u) para cada control admisible u mediante

$$J(u) := E\{\int_0^\tau L(t,x_t,u_t)dt + \Phi(\tau,x_\tau)\}$$

Se desea encontrar un control admisible u* con la propiedad

$$J(u^*) = \inf\{J(u) : u\varepsilon\mathcal{U}\} \tag{13}$$

Tal control se calificará de OPTIMO.

Observemos que el problema planteado se vuelve determinístico si el coeficiente σ en (11) es idénticamente cero. En ese caso tanto los controles u como las trayectorias x serán funciones, soluciones de la ecuación diferencial ordinaria

$$\dot{x}_t = b(t,x_t,u_t) \tag{14}$$

Un control será admisible si garantiza una sola trayectoria x para cada valor de x_0 Por ejemplo, podemos tomar $\mathcal{U}$ como la clase de todas las funciones $u:[0,T] \to U$ tales que

i) son continuas por tramos

ii) están dadas mediante políticas, como en (4)

Dado lo anterior, la función de costo $J:\mathcal{U} \to \mathbb{R}$ está dada por

$$J(u) := \int_0^{\tau} L(t,x_t,u_t)dt + \Phi(\tau,x_\tau) \tag{15}$$

Igual que en el caso aleatorio, un control $u^* \in \mathcal{U}$ será óptimo si cumple la condición (13).

Para resolver este problema determinístico, la Programación Dinámica comienza por sumergirlo en una familia infinita de problemas semejantes, uno para cada punto $(s,y) \in [0,T] \times B$. En efecto, el problema planteado supone que tanto los controles como las trayectorias comienzan en $t = 0$, siendo el valor inicial x_0 un punto bien definido de B. Aquí estamos variando tanto el instante inicial como el valor inicial de la trayectoria. Así pues, consideramos el problema de valores iniciales

$$\dot{x}_t = b(t,x_t,u_t) \qquad x_s = y \tag{16}$$

y restringimos los controles al intervalo $[s,T]$. Sea $\mathcal{U}_s$ la correspondiente clase de controles admisibles.

Hecho lo anterior, definimos el índice $J(u|x_s = y)$ mediante

$$J(u|x_s = y) := \int_s^{\tau} L(t,x_t,u_t)dt + \Phi(\tau,x_\tau)$$

Huelga la aclaración de que ahora τ es el primer instante DESPUES DE s en que la trayectoria x toca la frontera de B. Sea ahora $V(s,y)$ el valor del costo óptimo que corresponde a la condición inicial $x_s = y$, es decir

$$V(s,y) := \inf_{u \in \mathcal{U}_s} J(u|x_s = y) \tag{17}$$

Es claro que $V(0,x_0) = J(u^*)$, con u^* óptimo a partir de $t = 0$. Supondremos en todo lo que sigue que L es una función continua de sus argumentos además de que V es continuamente diferenciable.

LEMA I

Sea u admisible a partir de $t = 0$ y sea x la correspondiente trayectoria, para x_0 dado. Entonces,

$$\frac{d}{dt} V(t,x_t) + L(t,x_t,u_t) \geq 0, \quad 0 \leq t \leq T \tag{18}$$

DEMOSTRACION

Sean t, $t+h \in (0,T)$ y sea v un control admisible a partir de $t+h$; extendamos este último a un control $\tilde{u}$ admisible a partir de t poniendo

$$\tilde{u}_s := \begin{cases} u_s & \text{si } t \leq s \leq t+h \\ v_s & \text{si } t+h < s \leq T \end{cases}$$

Entonces,

$$V(t,x_t) \leq \int_t^{t+h} L(s,x_s,u_s)ds + J(v|x_{t+h})$$

Podemos tomar ínfimo sobre v para obtener

$$V(t,x_t) \leq \int_t^{t+h} L(s,x_s,u_s)ds + V(t+h,x_{t+h})$$

De aquí se obtiene el resultado deseado al dividir por h y tomar el límite cuando $h \to 0$. ||

Conviene interpretar el resultado anterior de manera global, en términos de la función de costo. Para ello observemos que la conclusión del Lema 1 puede reescribirse en la forma de

$$\frac{d}{dt}\{V(t,x_t) - \int_t^T L(s,x_s,u_s)ds\} \geq 0$$

es decir, la función

$$t \to V(t,x_t) \quad \int_t^T L(s,x_s,u_s)ds \tag{19}$$

es no decreciente, cualquiera que sea el control u   En otras palabras, a lo largo de las trayectorias del sistema controlado (16), la diferencia entre "el costo óptimo a partir de un punto dado" y "el costo acumulado desde ese punto hasta el final de la trayectoria" simplemente no puede decrecer.

LEMA 2

Sea u^* óptimo a partir de $t=0$ y sea x^* la correspondiente trayectoria, para x_0^* dado. Entonces,

$$\frac{d}{dt} V(t,x_t^*) + L(t,x_t^*,u_t^*) = 0, \quad 0 \leq t \leq T$$

DEMOSTRACION

Es claro que la restricción de u* a $[t,T]$ es admisible a partir de t, por lo que

$$V(t,x_t^*) \leq \int_t^\tau L(s,x_s^*,u_s^*)ds + \Phi(\tau,x_\tau^*)$$

En otras palabras

$$V(t,x_t^*) - \int_t^\tau L(s,x_s^*,u_s^*)ds \leq \Phi(\tau,x_\tau^*), \quad 0 \leq t \leq T$$

Por otro lado, la optimalidad de u* implica que

$$V(0,x_0^*) - \int_0^\tau L(s,x_s^*,u_s^*)ds = \Phi(\tau,x_\tau^*)$$

En virtud del Lema 1, las últimas dos relaciones implican que

$$V(t,x_t^*) - \int_t^\tau L(s,x_s^*,u_s^*)ds = \Phi(\tau,x_\tau^*), \quad 0 \leq t \leq T$$

Basta diferenciar la función

$$t \to V(t,x_t^*) - \int_t^\tau L(s,x_s^*, u_s^*)ds$$

para obtener la conclusión deseada. ||

Así pues, a lo largo de una trayectoria óptima, la diferencia entre "el costo óptimo a partir de un punto dado" y "el costo a partir de dicho punto y a lo largo de la trayectoria" se mantiene constante.

Por otro lado, la demostración del lema anterior permite ver que u* restringido a $[t,T]$ es óptimo a partir de t, sin que intervengan para nada las acciones de control tomadas antes de t, pues

$$V(t,x_t^*) = \int_t^T L(s,x_s^*,u_s^*)ds + \Phi(T,x_T^*), \quad 0 \leq t \leq T \quad (20)$$

Tenemos así la importante conclusión que enunciamos enseguida, la cual se conoce como

PRINCIPIO DE OPTIMALIDAD

Sea u óptimo para el problema de control que empieza a partir de $t=0$ en el estado inicial x y sea u^s su restricción a $[s,T]$, para $s\varepsilon(0,T)$. Entonces, u^s es óptimo para (16), sin importar las decisiones tomadas sobre $[0,s)$.

Ahora bien, observemos que los problemas de optimización que hemos encontrado hasta ahora son del tipo "con restricciones de igualdad". Para este tipo de problemas, el método de los multiplicadores de Lagrange prescribe formar el HAMILTONIANO $H:[0,T]\times B\times U\times\mathbb{R} \to \mathbb{R}$, dado por

$$H(t,x,u,\lambda):= L(t,x,u) + \lambda^T b(t,x,u) \quad (21)$$

En términos de este Hamiltoniano, las conclusiones de los dos lemas anteriores pueden muy bien expresarse de la manera siguiente:

Sea $u\varepsilon\mathcal{U}$ y sea x la correspondiente trayectoria. Entonces,

$$\frac{\partial V}{\partial t}(t,x_t) + H(t,x_t,u_t, \frac{\partial V}{\partial x}(t,x_t)) \geq 0, \quad 0 \leq t \leq T$$

Si u* es óptimo y x* es la correspondiente trayectoria, entonces

$$\frac{\partial V}{\partial t}(t,x_t^*) + H(t,x_t^*,u_t^*,\frac{\partial V}{\partial t}(t,x_t^*)) = 0, \quad 0 \leq t \leq T$$

Podemos ahora combinar ambos resultados en la observación de que

$$\min_{v \in U} H(t,x,v, \frac{\partial V}{\partial x}(t,x)) = \frac{\partial V}{\partial t}(t,x),$$

además de que el mínimo del Hamiltoniano se obtiene precisamente en el valor u_t^* del control que resulta óptimo a partir de t.

En otras palabras, tenemos el siguiente resultado:

TEOREMA 1.

Si V es continuamente diferenciable, entonces satisface la ecuación en derivadas parciales

$$\frac{\partial V}{\partial t} + \min_{v \in U} H(t,x,v, \frac{\partial V}{\partial x}) = 0 \tag{22}$$

en el interior del cilindro $Q := [0,T] \times B$, además de las dos condiciones

$$V(T,x) = \Phi(T,x), \quad x \in \operatorname{Int} B \tag{23}$$

$$V(t,x) = \Phi(t,x), \quad 0 \leq t \leq T, \quad x \in \partial B \qquad (23)''$$

La ecuación (22) se conoce en la literatura de control como de HAMILTON-JACOBI-BELLMAN (HJB), por analogía con las ecuaciones diferenciales parciales que resultan en la teoría de Hamilton-Jacobi del Cálculo de Variaciones [1] [6] Las condiciones (23) definen un problema de valores finales y a la frontera (PVFF) asociado a la ecuación de HJB.

Además, resalta de lo anterior el hecho de que un control óptimo u* satisface necesariamente la relación

$$H(t,x,u_t^*, \frac{\partial V}{\partial x}(t,x)) = \min_{v\in U} H(t,x,v, \frac{\partial V}{\partial x}(t,x))$$

condición que expresa el hecho de que los valores de un control óptimo minimizan puntualmente el Hamiltoniano. De esta observación obtenemos condiciones necesarias para la optimalidad de un control, que se traducen en el siguiente procedimiento para la obtención de candidatos a controles óptimos:

i) Minimícese puntualmente el Hamiltoniano con respecto a los valores de control, para cada valor de las variables restantes. Se obtiene así una función v^*: $[0,T]\times B\times \mathbb{R}^n \to U$ que satisface la condición

$$H(t,x,v^*(t,x,\lambda),\lambda) = \min_{v\in U} H(t,x,v,\lambda)$$

ii) Obténgase un control u mediante

$$u_t^* := v^*(t,x_t^*, \frac{\partial V}{\partial x}(t,x_t^*)) \tag{24}$$

donde V resuelve el PVFF (22), (23)

Algo que debe señalarse acerca del control óptimo que arroja este método es que está dado por una política de realimentación del estado del sistema; en efecto, (24) lo muestra de manera por además evidente. Huelga decir que la aplicación práctica de un control de este tipo requiere que el estado del sistema controlado sea perfectamente observable, lo cual a menudo no se puede garantizar.

Otro problema de tipo práctico es el de contar con soluciones suficientemente suaves de la ecuación de HJB, lo cual a veces no puede garantizarse; además falta verificar a posteriori que V (la solución encontrada para la ecuación HJB) en efecto da el costo óptimo y que el control encontrado también es óptimo. Esto tampoco es automático en muchas aplicaciones.

3. CONTROL DE DIFUSIONES.

Regresemos ahora al problema de control planteado al inicio de la sección anterior, referente a la ecuación diferencial estocástica (11), con ánimo de aplicarle los conceptos que acerca de la Programación Dinámica desarrollamos en la segunda parte de dicha sección.

De nuevo, podemos sumergir el problema en una familia infinita de problemas similares, uno para cada condición inicial (s,y) Comenzamos por especificar las correspondientes dinámicas, a saber,

$$d\,x_t = b(t,x_t,u_t)dt + \sigma(t,x_t)dB_t \qquad (25)'$$

$$x_s = y \qquad (25)''$$

Hecho esto, podemos definir la función de "costo óptimo a partir de t = s" mediante

$$V(s,y) := \inf_{u\in\mathcal{U}_s} E\{\int_s^\tau L(t,x_t,u_t)dt + \Phi(\tau,x_\tau)\,|\,x_s = y\}$$

en completa analogía con (17).

La función V así definida tiene algunas propiedades que recuerdan las consignadas a los lemas 1 y 2, pero cuya demostración en esta formulación estocástica requiere de herramienta matemática más avanzada.

Comencemos por presentar el análogo del Lema 1, en el que de nuevo se establece el carácter no decreciente de la diferencia entre "el costo óptimo a partir de un punto dado de una trayectoria cualquiera" y "el costo de operación acumulado a partir de dicho punto de la trayectoria en cuastión". Sin embargo, en este caso las trayectorias son procesos estocásticos (sus valores son variables aleatorias) por lo que debemos proceder con mayor cuidado.

LEMA 3.

Para cualquier política de control u, admisible a partir de $t=0$,

$$V(t,x_t) \leq E\{\int_t^{t+h} L(s,x_s,u_s)ds + V(t+h,x_{t+h})|x_t\}$$

con probabilidad 1 para cada t.

LEMA 4.

Si u es óptimo, entonces

$$V(t,x_t^*) = E\{\int_t^{t+h} L(s,x_s^*,u_s^*)ds + V(t+h,x_{t+h}^*)|x_t\}$$

con probabilidad 1 para cada t.

Estos dos resultados se demuestran de una manera muy semejante a la utilizada para demostrar los lemas 1 y 2, respectivamente. Véase [2, VI.2] para los detalles. También en el caso estocástico podemos enunciar el importantísimo Principio de Optimalidad, exactamente en los mismos términos de la sección anterior.

Ahora bien, un resultado muy importante de la teoría de ecuaciones diferenciales estocásticas es la llamada Fórmula de Dynkin |5,Vol. 1|

En dicha fórmula aparece la familia de operadores diferenciales de segundo orden dada por

$$A^v\phi := \frac{1}{2}\, a(t,x)\ \frac{\partial^2\phi}{\partial x^2} + b(t,x,v)\ \frac{\partial\phi}{\partial x}$$

para $v\in U$. Aquí, $a(t,x)$ es una función matricial de orden nxn, dada por

$$a := \sigma\sigma^T$$

y denota el producto interior

$$A\cdot B := \mathrm{tr} A\, B^T$$

definido para matrices del mismo orden.

En términos de dichos operadores, la fórmula de Dynkin dice que

$$E\{V(t+h)|x_{t+h}\} = V(t,x_t) + E\{\int_t^{t+h}\left[\frac{\partial V}{\partial t}(s,x_s) + A^{u_s}V(s,x_s)\right]ds|x_t\}$$

siempre y cuando V sea suficientemente suave. De ser así, los dos lemas anteriores implican que

$$E\{\int_t^{t+h}\left[\frac{\partial V}{\partial t}(s,x_s) + A^{u_s}V(s,x_s) + L(s,x_s,u_s)\right]ds|x_t\} \geq 0$$

para cualquier control admisible u; más aún,

$$E\{\int_t^{t+h}\left[\frac{\partial V}{\partial t}(s,x^*_s) + A^{u^*_s}V(s,x^*_s) + L(s,x^*_s,u^*_s)\right]ds|x^*_t\} = 0$$

si u^* es óptimo. En ambos casos la relación en cuestión ocurre con probabilidad 1 para cada $t\in[0,T]$

Basta dividir entre h y tomar el límite cuando $h \to 0$ para obtener el siguiente resultado:

TEOREMA 2.

Supongamos que V es continuamente diferenciable, dos veces con respecto a x. Si u es admisible a partir de t = 0, entonces

$$\frac{\partial V}{\partial t}(t,x_t) + A^{u_t}V(t,x_t) + L(t,x_t,u_t) \geq 0\,.$$

Si u es óptimo vale la igualdad; en cada caso, la conclusión es con probabilidad 1 para cada $t\epsilon[0,T]$

Igual que en el caso determinístico, podemos expresar convenientemente el resultado anterior en términos del Hamiltoniano definido en (21),

$$H(t,x,v,\lambda) := L(t,x,v) + \lambda^T b(t,x,v)$$

En efecto, el resultado anterior afirma que

$$\inf_{v\epsilon U} H(t,x,v,\frac{\partial V}{\partial x},(t,x)) = -\left[\frac{\partial V}{\partial t}(t,x_t) + \frac{1}{2}\,a(t,x)\cdot\frac{\partial^2 V}{\partial x^2}(t,x)\right]$$

y, además

$$\inf_{v\epsilon U} H(t,x,v,\frac{\partial V}{\partial x}(t,x)) = H(t,x,u_t^*,\frac{\partial V}{\partial x}(t,x))$$

donde u* es óptimo. De nueva cuenta, si v es la función que minimiza puntualmente el Hamiltoniano, se obtiene un candidato a control óptimo mediante la fórmula

$$u_t^* = v^*(t,x_t^*,\frac{\partial V}{\partial x}(t,x_t^*))$$

Así pues, se aplica de nuevo el procedimiento para la resolución de problemas de control óptimo bosquejado al final de la sección anterior.

La ecuación de HJB para el caso estocástico no es otra que

$$\frac{\partial V}{\partial t} + \frac{1}{2}\, a(t,x) \quad \frac{\partial^2 V}{\partial x^2} + \min_{v \in U} H(t,x,v,\frac{\partial V}{\partial x}) = 0, \tag{26}$$

la cual debe resolverse junto con las condiciones finales y a la frontera

$$V(T,x) = \Phi(T,x), \quad x \varepsilon B \tag{27}$$

$$V(t,x) = \Phi(t,x), \quad 0 \leq t \leq T \quad x \varepsilon \partial B \tag*{(27)''}$$

Igual que en el caso determinístico, un paso importante del algoritmo es mostrar que existen soluciones de la ecuación de HJB con el suficiente grado de suavidad, lo cual no es automático. Por supuesto, luego hay que calcular la solución del PVF especificado por (26) y (27)

Un caso en el que "todo sale bien" es el problema de control de humedad en agricultura que fue planteado en la primera parte de la sección 1 Ahí el Hamiltoniano está dado por

$$H(t,x,u,\pi) = \frac{1}{2}\, \lambda x^2 + \frac{1}{2}\, \mu u^2 + \pi(-\alpha x + \beta u + \delta)$$

y la función que lo minimiza es

$$v^*(t,x,\pi) = \quad \frac{\beta}{\mu}\, \pi$$

Así pues, una condición necesaria para la optimalidad de un control es la de que

$$u_t^* = \quad \frac{\beta}{\mu} \; \frac{\partial V}{\partial x}\, (t,x_t^*)$$

donde V satisface la ecuación de HJB del problema:

$$\frac{\partial V}{\partial t} + \frac{\sigma^2}{2} \frac{\partial^2 V}{\partial x^2} \quad \frac{\beta^2}{2} \left(\frac{\partial V}{\partial x}\right)^2 + (\delta - \alpha x) \frac{\partial V}{\partial x} + \frac{\lambda}{2} x^2 = 0$$

junto con

$$V(T,x) = 0, \quad x \in \mathbb{R}$$

Véase [7, sección 3] para los detalles. Ahí se hace ver la factibilidad de tener soluciones de la forma

$$V(t,x) = \frac{1}{2} p(t) x^2 + q(t) x + r(t),$$

por lo que un control óptimo será

$$u_t^* = - \frac{\beta}{\mu} \left[p(t) x_t^* + q(t) \right]$$

4. SOLUCIONES DE LA ECUACION DE HJB.

Al final de la sección anterior revisamos someramente un ejemplo de control estocástico en el que el formalismo de la Programación Dinámica conduce facilmente a una solución. Previamente, se anotaron algunos de los principales problemas comunes a toda aplicación de este tipo, en las que se debe verificar a posteriori el enfoque utilizado, amén de garantizar la existencia de soluciones suficientemente suaves de la ecuación de HJB y de hecho encontrarlas. Resumiendo, hay tres tareas inherentes a toda aplicación de la Programación Dinámica a la solución de este tipo de problemas:

i) Verificar la validez del enfoque, probando un teorema idóneo.

ii) Establecer la existencia y uncidad de soluciones de la ecuación HJB con el suficiente grado de suavidad, y

iii) Encontrar dicha solución.

Existe una gama de teoremas de verificación de la Programación Dinámica y todos requieren de un cierto grado de suavidad en la solución de la ecuación de HJB. Por ejemplo, en la página 159 de |2| se prueba el siguiente

TEOREMA DE VERIFICACION.

Supongamos que V es continuamente diferenciable, dos veces con respecto a x, además de que

$$|V(t,x)| \leq C(1+|x|^k), \quad 0 \leq t \leq T \quad x \in B \tag{28}$$

para k natural y C positivo. Entonces,

a) Para cualquier control u, admisible a partir de t = s,

$$V(s,x_s) \leq E\{\int_s^\tau L(t,x_t,u_t)dt + \Phi(\tau,x_\tau) \; x_s\}$$

b) Si u* es admisible y se obtiene minimizando puntualmente el Hamiltoniano, entonces

$$V(s,x_s^*) = E\{\int_s^\tau L(t,x_t^*)dt + \Phi(\tau^*,x_\tau^*) | x_s^*\},$$

por lo que u* es óptimo.

De (28) se dice que constituye una condición de crecimiento polinominal para V. Abreviando, se dirá de V como en el Teorema de Verificación que "está en $C_p^{1,2}$".

¿ Qué condiciones garantizan la existencia de soluciones en $C_p^{1,2}$ para la ecuación de HJB ?

Hay un gran número de teoremas de existencia para ecuaciones diferenciales parabólicas semilineales, como es el caso de la de HJB [4] En todos ellos se obtienen soluciones generalizadas, cuya suavidad depende grandemente del hecho de que la matriz a(t,x) satisfaga una condición de la forma

$$\lambda|z|^2 \leq \sum_{i,j=1}^{n} a_{ij}(t,x)z_i z_j \leq \mu|z|^2; \quad z \in \mathbb{R}^n \tag{29}$$

para $0 < \lambda < \mu < \infty$. Esta condición significa que el espectro de a(t,x)

-necesariamente real- está contenido en $[\lambda,\mu]$, uniformemente. Se dice entonces que el operador diferencial

$$\sum_{i,j=1}^{n} a_{ij}(t,x) \frac{\partial^2}{\partial x_i \partial x_j}$$

es UNIFORMEMENTE ELIPTICO.

Por desgracia hay problemas de control estocástico que dan lugar a operadores (30) que NO son uniformemente elípticos; es un problema abierto el de dar condiciones bajo las cuales la ecuación de HJB tiene soluciones en $C_p^{1,2}$ en estos casos llamados "singulares".

A manera de ejemplo, regresemos al problema de control de presas que fue planteado en la segunda parte de la sección 1 Para este problema, el Hamiltoniano está dado por

$$H = \lambda|d(t) \quad u|^2 + \mu|x-u|^2 + py \qquad \omega^2 x q + (x \quad u)r$$

donde p, q y r son los multiplicadores de Lagrange y d(t) es la demanda de agua río abajo en el instante t. Véanse las ecuaciones (7) y (9) El control que minimiza el Hamiltoniano es, entonces,

$$u_t^* = \frac{\lambda}{\lambda+\mu} d(t) + \frac{\mu}{\lambda+\mu} x_t^* + \frac{1}{\lambda+\mu} \frac{\partial V}{\partial x} (t, x_t^*, y_t^*, z_t^*)$$

Por lo tanto, la ecuación de HJB asociada a este problema es

$$\frac{\partial V}{\partial t} + \frac{\sigma^2}{2} \frac{\partial^2 V}{\partial y^2} + y \frac{\partial V}{\partial x} - \omega^2 x \frac{\partial V}{\partial y} + x \frac{\partial V}{\partial z} - \left[\frac{1}{\lambda+\mu} \lambda d(t) + \mu x + \frac{1}{2} \frac{\partial V}{\partial z} \right]^2 = 0,$$

de la cual interesan las soluciones que se anulan en las caras lateral y final del cilindro $[0,T] \times Q$.

Finalmente, todavía queda el problema de aproximar adecuadamente las soluciones de la ecuación de HJB. Aquí puede resultar muy conveniente el enfoque a base de perturbaciones singulares, que reemplace la ecuación de HJB dada por otra cuya parte elíptica esté dada por la matriz

$$\begin{bmatrix} \varepsilon^2/2 & 0 & 0 \\ 0 & \sigma^2/2 & 0 \\ 0 & 0 & \varepsilon^2/2 \end{bmatrix}$$

Se resuelve este problema perturbado y se examina el comportamiento de la solución cuando $\varepsilon \to 0$. Este enfoque se sigue en numerosos estudios entre los que podemos mencionar la sección 3.3 de [8], que mucho debe al artículo [3]

Sin embargo, el teorema antes citado contiene hipótesis demasiado restrictivas en cuanto a la estructura de la ecuación de HJB. Por ello, en la sección 4.5 de la misma referencia [8] se prueba un teorema de existencia de soluciones continuas y acotadas de la ecuación de HJB en el caso singular por métodos probabilísticos y bajo suposiciones menos restrictivas. A su vez, dicho desarrollo debe mucho al artículo [9] y de hecho está emparentando con los métodos empleados en el volumen 2 de [5].

Finalmente, en la sección 4 de [8] se prueba un teorema de verificación para sistemas singulares de estructura semejante al que nos ocupa y que requiere unicamente soluciones continuas y acotadas para la ecuación HJB. Así pues, podría pensarse que se dispone de herramienta suficiente para lidiar con este tipo de problemas, pero aún queda mucho por hacer en cuanto al diseño de procedimientos efectivos de cálculo para resolver la ecuación de HJB en los casos singulares.

BILIOGRAFIA.

[1] CARATHEODORY, C., "Calculus of Variations and Partial Differential Equations of the First Order, Vol 1", Holden Day.

[2] FLEMING, W.H. y RISHEL, R.W., "Deterministic and Stochastic Optimal Control". Springer Verlag, Nueva York, 1975.

[3] FLEMING, W., "The Cauchy problem for degenerate parabolic equations", J. Math. Mech., 15(1964), 987-1008.

[4] FRIEDMAN, A., "Partial differential equations of parabolic type", Prentice Hall, Inc., NJ, 1964.

[5] FRIEDMAN, A., "Stochastic differential equations" (2 Vols.), Academic Press, Nueva York, 1974.

[6] GOLDSTEIN, H., "Classical Mechanics", Addison Wesley, Reading MA, 1950.

[7] HERNANDEZ, D.B., "Políticas óptimas para la irrigación: cultivo en tiempo de secas", manuscrito.

[8] HERNANDEZ, D.B., "on the verification of Dynamic Programming for Singular Stochastic Control Problems", Ph. D. Thesis, Imperial College of Science and Technology, University of London, 1974.

[9] RISHEL, R.W., "Weak solutions of a partial defferential equation of Dynamic Programming", SIAM J. Control 9 (1971), 519-528.

[10] WONG, E., "Stochastic Processes in Information and Dynamical Systems", Mc Graw Hill Book Co., Nueva York, 1971.