

CONTROL OPTIMO Y JUEGOS DIFERENCIALES

VERA W. DE SPINADEL

1. Introducción

Se puede definir un juego diferencial como un modelo matemático diseñado para resolver una situación conflictiva variable con el tiempo. En la teoría de juegos diferenciales interviene la teoría de juegos, el cálculo de variaciones y la teoría de control óptimo.

Uno de los problemas que dio origen a la teoría de juegos diferenciales fue el problema de persecución-evasión, cuyos orígenes según Davis (1962) se remontan a Leonardo da Vinci (siglo XV). Sin embargo, recién comenzó el problema a despertar interés entre los matemáticos tres siglos después cuando P. Bouguer (1732) anunció haber hallado la curva de persecución, esto es, "la curva que describe una embarcación que se aproxima a otra que se mueve en línea recta, suponiendo que las velocidades de ambas embarcaciones están en relación constante". Años después, algunos autores consideraron el problema más difícil en el que la curva de persecución era una circunferencia. Estos trabajos previos tenían la característica común que tan sólo un jugador, el perseguidor, tenía libertad de movimiento, mientras que el otro, el evasor, estaba limitado a una trayectoria fija.

R. Isaacs (1954-56), en cuatro informes publicados por la Rand Corp., fue el primero en considerar el problema de persecución, con movimiento libre de ambos jugadores. Halló la solución de este problema usando las conocidas ecuaciones de Hamilton-Jacobi del cálculo de variaciones. Pero hasta que no apareció su famoso libro, "Differential Games", en 1965, el problema no despertó ningún interés particular, salvo el trabajo seminal de Bellman, Glicksberg y Gross (1956) sobre el llamado problema de control "bang-bang".

Desde entonces, el entusiasmo por el tema no ha decaído, dedicándose los matemáticos americanos, tanto como los soviéticos al estudio formal de los juegos diferenciales así como al análisis de ejemplos ilustrativos.

En particular, la escuela soviética se inició con Pontrjagin et al. (1962) que en su clásico libro usaron los mismos métodos que Isaacs, siendo su técnica la de un juego diferencial con un solo jugador. En 1961, Kelendzeridze extendió dicho método a dos jugadores y su trabajo puede considerarse como una continuación del de Isaacs.

Por último, en 1988, Krasovskii y Subbotin publicaron un libro: "Game-Theoretic Control Problems", que es una versión revisada de su monografía (1974). En este libro, el problema de control se plantea así:

Dado un sistema de control, cuya dinámica está descrita por un sistema de ecuaciones diferenciales ordinarias, se busca construir una estrategia, esto es, un control basado en el principio de realimentación, que asegure una calidad definida para el proceso controlado bajo condiciones especificadas cuando el proceso está sujeto a factores inciertos. Estos factores pueden ser una perturbación o bien una acción controlante proveniente del lado opositor. En el ámbito de la teoría de juegos diferenciales, este tipo de problemas se considera conjuntamente

con el problema inverso. El análisis de ambos problemas es útil, no sólo en situaciones en las cuales dos jugadores controlan el sistema sino también en casos en los cuales el sistema controlado (por un solo jugador) está sujeto a una perturbación incierta.

Las contribuciones americanas, en cambio, se orientaron hacia la fundamentación lógica del tema. Por ejemplo, Berkovitz (1961) aplicó el cálculo de variaciones a la estrategia de un jugador, quedando la estrategia del otro fijada de antemano. Fleming (1957), a su vez, introdujo una estrategia continua entre dos discretas, considerando distribuciones de probabilidad cuya esperanza era la certeza cuando el juego se convertía en continuo.

Contemporáneamente, la teoría del control óptimo, que en apariencia nada tenía que ver con los juegos diferenciales, comenzó a desarrollarse. La teoría de control óptimo fue estudiada principalmente por Bellman (1956), Lasalle (1960), Bushaw (1958), Gamkrelidze (1958), Krasovskii (1957), Roxin (1969) y otros. El problema de control óptimo es equivalente al de un juego diferencial con un solo jugador y los problemas de control óptimo pueden ser extendidos a juegos diferenciales agregando un oponente.

2. Ecuaciones de movimiento

Sea $x(t)$ un vector n -dimensional que indica el estado del sistema controlado al tiempo t . Su dinámica queda descrita por un sistema de n -ecuaciones diferenciales.

$$(1) \quad \frac{dx(t)}{dt} = f(t, x(t), u(t), v(t)),$$

con la condición inicial

$$(2) \quad x(t_0) = x_0.$$

$u(t)$ es un vector p -dimensional que caracteriza la acción de control y el vector $v(t)$, de dimensión w , corresponde a la perturbación actuante al tiempo t . Además

$$(3) \quad u(t) \in U \subset \mathbb{R}^p; \quad v(t) \in V \subset \mathbb{R}^q,$$

donde U y V son conjuntos compactos. Si se tratara de un juego diferencial, los vectores $u(\cdot)$ y $v(\cdot)$ serían los controles de los jugadores u y v , respectivamente. Supondremos que tanto $u(\cdot)$ como $v(\cdot)$ son funciones medibles, condición necesaria para que tenga sentido el sistema (1) - (2).

En lugar del sistema (1) - (2), podríamos usar una inclusión diferencial (Clarke, 1986) que no involucraría una perturbación específica sino que tan sólo indicaría las cotas de indeterminación generadas por una posible perturbación. En ese caso, el movimiento del sistema controlado debiera satisfacer la inclusión diferencial

$$(4) \quad \frac{dx(t)}{dt} \subset \{ f(t, x(t), u(t), v) : v \in V \}$$

y, en general, el conjunto de soluciones sería infinito.

El juego comienza a un cierto $x(t_0)$ dado y termina cuando $x(t)$ alcanza una variedad cerrada terminal dada M . Sea T el primer valor de $t > t_0$ para el cual $x(t) \in M$.

Supongamos que la funcional pago (payoff) esté dada por

$$(5) \quad \mu = \mu [T, x(t), u(t), v(t); t_0 \leq t \leq T].$$

Esta es la cantidad que el jugador v debe pagar al jugador u . Más comunmente, el pago consiste en la suma de dos términos del tipo

$$(6) \quad P(u,v) = g(T; x(T)) + \int_{t_0}^T h(t, x(t), u(t), v(t)) dt,$$

donde $g(\cdot)$ es el *pago terminal* y el segundo término es el *pago integral*.

El objetivo del jugador u es maximizar el valor de $P(\cdot)$ mientras que el del jugador v será el minimizarlo.

Con respecto a la *estrategia* a seguir, la misma debe indicar a cada jugador cómo jugar el juego, esto es, cómo elegir su propio control al tiempo t . Dicha elección de la variable control se basa en el conocimiento del pasado y posible presente del juego. Por supuesto, un jugador no puede predecir el comportamiento futuro de su oponente. Además, el conjunto de estrategias debe ser tal que una vez elegidas por ambos jugadores, los controles correspondientes $u(\cdot)$ y $v(\cdot)$ existan y estén unívocamente determinados. Pero aún en casos muy simples, esto no siempre sucede ya que, en general, la elección de $u(\cdot)$ depende de la de $v(\cdot)$ para el mismo t y recíprocamente. En un juego real, como se produce una cierta demora en la transmisión y procesamiento de la información, además de cierta inercia propia del sistema, ello ayuda a suavizar estas situaciones incompatibles, pero son muy importantes para la formulación del modelo matemático.

En resumen, el problema fundamental para cada jugador se reduce a optimizar sobre el conjunto de *estrategias admisibles*.

3. Estrategias admisibles

R. Isaacs introdujo el concepto de estrategia como una función de la variable de estado x . Por ejemplo, una estrategia para u es una función $u(t,x)$ tal que la elección del control sea $u(t) = u(t,x(t))$. Esto corresponde a un *control de realimentación*, de gran utilidad en muchos juegos de importancia práctica. A pesar de este hecho, desde un punto de vista matemático, pueden aparecer algunas situaciones de incompatibilidad o aún de indeterminación. Y lo que es peor aún, no hay ninguna garantía de que para cada elección de $u(t,x(t))$ y $v(t,x(t))$, el sistema diferencial (1)-(2) tenga una solución ni siquiera que dicha solución sea única. Esta es la razón por la cual, dependiendo las posibles elecciones de $u(\cdot)$ de las de $v(\cdot)$ y recíprocamente, es muy difícil definir precisamente el conjunto de trayectoria admisibles.

Por otra parte, ejemplos muy simples muestran que los controles óptimos (así como también las trayectorias óptimas) poseen discontinuidades. La existencia de *superficies singulares* es uno de los puntos clave en la formulación de la teoría.

Para evitar las dificultades que se presentan con las estrategias admisibles introducidas por Isaacs, Varaiya (1967) y Friedman (1970) han desarrollado un enfoque más abstracto. Este enfoque no invalida los resultados de Isaacs sino que trata de dar a la teoría fundamentos más sólidos. Está relacionado con las ideas de Fleming y Berkovitz de aproximar el juego continuo mediante algún proceso de discretización, pero en este caso no es necesario considerar elecciones estocásticas para cada jugador. Lo que hace Varaiya no es una verdadera discretización sino que introduce retardos temporales. Sin embargo, ambos se basan en la idea de usar transformaciones en espacios funcionales para aproximar y, en el límite, definir las estrategias.

Veamos el enfoque de Friedman. Sea un juego diferencial definido por las ecuaciones (1), (2), (3) y (6) donde $0 \leq t \leq T$. Supongamos que los conjuntos de control U y V están fijos y además que

(i) $f(t,x,u,v)$ es continua,

(ii) existe una constante k tal que

$$(7) \quad \|f(t, x_1(t), u(t), v(t)) - f(t, x_2(t), u(t), v(t))\| \leq k_1 \|x_1 - x_2\|,$$

(iii) existen constantes k_2 y k_3 tales que

$$(8) \quad \|f(t, x(t), u(t), v(t))\| \leq k_2 \|x\| + k_3.$$

Los controles admisibles serán todas aquellas funciones medibles $u(\cdot)$ y $v(\cdot)$ que satisfacen la ecuación (3) en $[0, T]$. Las condiciones (i) - (iii) aseguran que para cada par de controles admisibles, existe una única *trayectoria* $x(t)$ del sistema (1)-(2) en $[0, T]$.

(iv) u y v actúan "por separado" en el miembro derecho de (1), es decir,

$$(9) \quad f(t, x(t), u(t), v(t)) = f_1(t, x(t), u(t)) + f_2(t, x(t), v(t)),$$

(v) $g(T, x(T))$ y $h(t, x(t), u(t), v(t))$ son funciones continuas,

(vi) u y v están también separadas en el pago integral, esto es,

$$(10) \quad h(t, x(t), u(t), v(t)) = h_1(t, x(t), u(t)) + h_2(t, x(t), v(t))$$

Bajo estas condiciones, podemos definir de manera análoga a la teoría clásica de juegos, el valor V del juego diferencial, que es el mejor valor del pago, su minimax. Será el pago cuando ambos jugadores actúen óptimamente. Si uno de ellos no actúa óptimamente, su oponente tendrá la oportunidad de obtener un pago mejor que su valor.

Friedman ha probado que si se cumplen las condiciones (i) - (vi) y además:

(vii) $g(T, x(T))$ es una función Lipschitz uniformemente continua en x .

(viii) $f(t, x(t), u(t), v(t))$ y $h(t, x(t), u(t), v(t))$ son funciones Lipschitz uniformemente continuas en t ,

entonces existe el valor V del juego y es una función Lipschitz continua con respecto a t , y x .

Los resultados previos valen para juegos diferenciales con condición terminal $t = T$ fija. En el caso general, cuando $(T, x(T)) \in M$, surgen algunas dificultades. La primera es el problema de la *controlabilidad* que se puede expresar así: ¿puede un jugador forzar la terminación del juego contra la voluntad de su oponente? Esta es la pregunta clave en un juego de persecución. Siguiendo a Isaacs, éste es un *juego de clase*, donde el resultado es sí o no, o bien en términos del pago, $P(u, v) = +1$ o bien -1 . En cierto sentido, puede ser sustituido por un *juego de grado*, considerando el *tiempo de captura* como el pago en un juego de persecución, es decir, $P(u, v) = T$. Bajo condiciones apropiadas, este problema de tiempo óptimo posee una solución continua en la llamada *región de controlabilidad*, donde el perseguidor puede alcanzar al evasor; fuera de esta región, el valor del juego es $+\infty$. Esto significa una discontinuidad y, en realidad, el juego puede presentar discontinuidades (tanto del valor como de las estrategias óptimas) aún cuando el tiempo de captura sea finito.

Problemas similares aparecen cuando se consideran pagos más generales del tipo (6). En este caso, la noción de controlabilidad debe ser analizada con más cuidado. Desde el comienzo, no es fácil saber qué jugador (si es que hay uno) desea terminar el juego tan pronto como sea posible. Y esta circunstancia puede variar de una región a otra.

Isaacs ha introducido la noción de *parte usable* del conjunto terminal. Esta parte está determinada por aquellos puntos del contorno ∂M donde es factible que el juego (jugado

razonablemente) termine. Esto significa que desde cualquier punto x_0 de algún entorno de tal punto límite de M , cualquiera de los jugadores puede llevar $x(t)$ a la variedad terminal M en un cierto intervalo de tiempo que tiende a cero con la distancia entre x_0 y M .

En este enfoque funcional se ha puesto especial énfasis en la demostración de teoremas de existencia para el valor del juego. Pero aún en juegos diferenciales sin valor, Roxin (1970) ha probado que el juego se hace extremadamente sensible a cambios infinitesimales en la información que recibe cada jugador. En algunos casos particulares, es posible obtener una estimación del valor del juego usando funciones de Liapovnov y algunas inecuaciones diferenciales adecuadas.

4. Aplicaciones Económicas

Casi todos los resultados publicados sobre juegos diferenciales se han restringido a juegos de dos personas de suma cero, en los cuales la posibilidad de interés mutuo entre los jugadores en conflicto, está descartada. Para poder aplicar la teoría de juegos diferenciales al análisis de la competición económica, es necesario utilizar una formulación válida para N personas, de suma no cero. Ciletti y Starr (1970) han dado un modelo para considerar N firmas que fabrican un solo producto. En el mercado, estos N productos son sustituibles pero no idénticos. Aún haciendo hipótesis fuertemente simplificadoras, no es fácil dominar las dificultades analíticas. Y queda pendiente la pregunta: ¿Qué utilidad puede tener un modelo simplificado para aumentar la comprensión de una competencia oligopolística?

Recientemente, A. Mehlmann (1988) ha analizado el tema de juegos diferenciales de suma distinta de cero así como sus aplicaciones, presentando una tipología simple que refleja las propiedades estructurales de algunos juegos diferenciales cualitativos. Entre las aplicaciones seleccionadas, ha considerado en particular la dinámica de la oligopolía, tanto con horizonte temporal finito como infinito.

5. Juegos diferenciales lineales

Consideremos un juego diferencial lineal definido por el sistema

$$(11) \quad \dot{x} = f(x, u, v) = Ax + Bu + Cv,$$

donde $x \in \mathbb{R}^n$; u y v son variables de control tales que

$$(12) \quad u(t) \in U \subset \mathbb{R}^p; v(t) \in V \subset \mathbb{R}^q.$$

El juego termina cuando x alcanza una cierta variedad analítica $M \subset \mathbb{R}^n$.

El objetivo es encontrar una estrategia para u de manera que el juego se termine en el menor tiempo posible, suponiendo que conocemos el estado x y el valor de v a cierto tiempo t .

Siguiendo a Pontrjagin, podemos analizar la controlabilidad del sistema (11) de manera análoga a como lo haríamos para el correspondiente sistema de control

$$\dot{x} = Ax + Bu.$$

Para ello, supondremos que M sea lineal, por ejemplo, el plano $x_1 = 0$ y el control sujeto a las restricciones

$$\|u\| \leq 1, \|v\| \leq 1,$$

donde $\|u\| = \max |u_j|$. Al igual que en la teoría de control óptimo introduciremos un vector covariante $\Psi = (\Psi_1, \Psi_2, \dots, \Psi_n)$ y formamos el hamiltoniano

$$(13) \quad H(x, \Psi, u, v) = \Psi f(x, u, v) = \Psi Ax + \Psi Bu + \Psi Cv.$$

Supongamos que u trata de minimizar el tiempo y v de maximizarlo. Entonces los controles óptimos serán

$$(14) \quad \bar{u} : H \text{ máximo}; \Psi Bu \text{ máximo}; \quad \bar{u} = \text{sg}(\Psi B)^T$$

donde $\text{sg } x = \begin{matrix} \text{sg } x_1 \\ \text{sg } x_2 \\ \vdots \\ \text{sg } x_n \end{matrix}$ y T indica la matriz transpuesta.

Correspondientemente

$$(15) \quad \bar{v} : H \text{ mínimo}; \Psi Cv \text{ mínimo} \quad \bar{v} = -\text{sg}(\Psi C)^T$$

Las ecuaciones canónicas resultan las siguientes

$$(16) \quad \dot{x} = \frac{\partial H}{\partial \Psi} = Ax + Bu + Cv.$$

$$(17) \quad \dot{\Psi} = -\frac{\partial H}{\partial x} = -\Psi f_x = -\Psi A.$$

Integrando esta última ecuación obtenemos

$$(18) \quad \Psi(t) = \Psi(0) e^{-At}$$

y usando (14) y (15) podemos calcular \bar{u} y \bar{v} y finalmente integrar (16) para tener $x(t)$, siempre que conozcamos $\Psi(0)$ y estemos seguros que *la solución óptima existe*.

Podemos generar soluciones óptimas comenzando en el punto terminal, sobre la variedad M . La "condición de transversalidad" entre M y la trayectoria se reduce, en el caso de tiempo mínimo, a que $\Psi(T)$ sea ortogonal a M .

De hecho, esto coincide con la condición de Isaacs ya que en este caso el valor del juego se reduce a $V(x)$ y

$$(19) \quad \Psi = \nabla V = \frac{\partial V}{\partial x}$$

siendo $V(x) = 0$ para $x \in M$. Entonces el vector $\Psi(T) = \frac{\partial V}{\partial x}$ resulta ortogonal a M .

Si la familia de trayectorias cubre todo el plano, entonces todo punto es un posible punto inicial para iniciar desde él la trayectoria que alcance a M .

Ejemplo 1

Consideremos el sistema de control

$$(20) \quad \begin{aligned} \dot{x}_1 &= x_2 + 2u \\ \dot{x}_2 &= -x_1 + v \end{aligned}$$

donde $|u| \leq 1$; $|v| \leq 1$. Sea M el plano $x_1 = 0$ y supongamos que u quiere minimizar el tiempo de llegada a M y v quiere maximizarlo. El hamiltoniano resulta

$$(21) \quad H = \Psi_1 x_2 - \Psi_2 x_1 + 2 \Psi_1 u + \Psi_2 v$$

de donde

$$(22) \quad \begin{aligned} \dot{\Psi}_1 &= -\frac{\partial H}{\partial x_1} = \Psi_2 \\ \dot{\Psi}_2 &= -\frac{\partial H}{\partial x_2} = -\Psi_1 \end{aligned}$$

Integrando tenemos

$$(23) \quad \begin{aligned} \Psi_1(t) &= K \cos(t + \varphi) \\ \Psi_2(t) &= -K \sin(t + \varphi) \end{aligned}$$

Para fijar la condición final podemos suponer que el tiempo final sea 0 tomando para el tiempo inicial algún valor negativo de t . Supongamos

$$\begin{aligned} x_1(0) &= 0 & x_2(0) & \text{arbitrario} \\ \Psi_2(0) &= 0 & \Psi_1(0) & \text{arbitrario, p.ej. } \Psi_1(0)=1 \end{aligned}$$

Reemplazando en (23) tenemos: $\varphi = 0$ o bien $\varphi = \pi$; $K = 1$. Las estrategias óptimas son las siguientes:

$$(24) \quad \begin{aligned} \bar{u} &= \text{sg } \Psi_1(t) = \text{sg } \cos t & \text{para } H \text{ máximo} \\ &= \text{sg } \cos(t + \pi) \end{aligned}$$

$$(25) \quad \begin{aligned} \bar{v} &= -\text{sg } \Psi_2(t) = -\text{sg } \sin t & \text{para } H \text{ mínimo} \\ &= -\text{sg } \sin(t + \pi) \end{aligned}$$

Las ecuaciones para las trayectorias son las siguientes:

$$(26) \quad \begin{aligned} \dot{x}_1 &= x_2 + 2 \text{sg } \cos t \\ &= x_2 + 2 \text{sg } \cos(t + \pi) \\ \dot{x}_2 &= -x_1 + \text{sg } \sin t \\ &= -x_1 + \text{sg } \sin(t + \pi) \end{aligned}$$

Consideremos cuál de las dos posibilidades: $\varphi = 0$ ó $\varphi = \pi$ es mejor. Antes de la terminación del juego, esto es, para t negativo pequeño, u trata de acelerar la terminación mientras que v trata de retrasarla. Entonces el punto x debe estar cerca del eje $x_1 = 0$, por ejemplo a la derecha:

$x_1 > 0$.

A partir de $\dot{x}_1 = x_2 + 2u$ resulta que u desea hacer el valor de x_1 tan negativo como sea posible y elige $u = -1$. En este caso, en (24) es mejor tomar

$$u = \text{sg } \cos(t + \pi)$$

Esto sucede si $x_2 < 2$ porque entonces u puede empujar el punto hacia el eje x_2 desde la derecha. Podemos asegurar que el semieje $x_2 < 2$ es alcanzable por trayectorias óptimas desde la derecha. Por simetría, $x_2 > -2$ será alcanzable por trayectorias óptimas desde la izquierda (eligiendo $\varphi = 0$).

Geoméricamente, en este ejemplo, u y v actúan en direcciones opuestas. Sin embargo, el sistema de control con $u(v=0)$ ó $v(u=0)$ es controlable.

Si ambos jugadores usan fuerza máxima, el sistema de control oscila en torno a un centro desplazado cuyas coordenadas son las siguientes:

$$\begin{array}{llll} u = 1 & ; & v = 1 & : & x_{2c} = -2 & ; & x_{1c} = 1 \\ u = -1 & ; & v = -1 & : & x_{2c} = 2 & ; & x_{1c} = -1 \\ u = 1 & ; & v = -1 & : & x_{2c} = -2 & ; & x_{1c} = -1 \\ u = -1 & ; & v = 1 & : & x_{2c} = 2 & ; & x_{1c} = 1 \end{array}$$

En la Fig.1 se han graficado los cuatro círculos con centro $(1;-2)$; $(-1;2)$; $(-1;-2)$ y $(1;2)$. Teniendo en cuenta que las oscilaciones son isócronas y los ángulos equivalentes a tiempo, podemos analizar cuáles son las mejores elecciones para u y v . Por ejemplo, partiendo del semieje $x_1 = 0$ para $t = 0$, $x_2 < 2$ para $t < 0$ con $u = -1$; $v = 1$. De esta manera, se cubre la región A. Después de rotar 90° debemos cambiar v a -1 . Los puntos de conmutación están sobre la recta $x_2 = 3$. Con centro en $(-1; -1)$ se cubre tanto la región B como la C. Desde un punto de la zona rayada existen dos caminos para alcanzar el eje de coordenadas: 1) $u = -1$; $v = 1$; 2) $u = -1$; $v = -1$.

En este ejemplo, integrando las ecuaciones de las trayectorias para atrás, obtenemos no sólo la estrategia óptima sino también otra trayectoria que satisface la condición de optimalidad.

Ejemplo 2

Sea el sistema de control definido por

$$(27) \quad \begin{aligned} \dot{x}_1 &= 2u_1 + v_1 \\ \dot{x}_2 &= 2u_2 + v_2 \end{aligned}$$

donde $u_1^2 + u_2^2 \leq 1$; $v_1^2 + v_2^2 \leq 1$. La variedad terminal es ahora el origen de coordenadas. El hamiltoniano resulta

$$H = 2 \Psi_1 u_1 + \Psi_1 v_1 + 2 \Psi_2 u_2 + \Psi_2 v_2$$

de donde $\dot{\Psi}_1 = \dot{\Psi}_2 = 0$ lo que significa $\Psi_1 = \Psi_2 = \text{const.}$ Las estrategias óptimas están dadas por las ecuaciones

$$\begin{aligned} \bar{u} &= \text{sg } \Psi && \text{para } H \text{ máximo} \\ \bar{v} &= -\text{sg } \Psi && \text{para } H \text{ mínimo.} \end{aligned}$$

La ecuación vectorial para la trayectoria se puede poner en la forma

$$\dot{x} = 2 \bar{u} + \bar{v} = 2 \bar{u} - \bar{u} = \bar{u} = \text{sg } \Psi .$$

En consecuencia, para alcanzar el origen, Ψ debe estar en la dirección del origen; v tiene una estrategia opuesta pero lo único que consigue es producir un retraso en la velocidad. Sin embargo, veremos que existe una estrategia para v que evita que u alcance el origen. Siendo la acción de v más débil que la de u , lo aconsejable para v es jugar de manera aleatoria. Entonces u sólo puede responder "con retraso" a la acción de v . Matemáticamente, este proceso es el "movimiento browniano", es decir, un proceso estocástico similar a la difusión de una partícula.

Es obvio que u puede llegar a hacer $x_1 = 0$ y también $x_2 = 0$ pero la probabilidad de que ambas condiciones se satisfagan simultáneamente, es nula. En general, cuando las trayectorias estocásticas corten al eje x_1 , no cortarán al eje x_2 y viceversa.

Como u no conoce en forma instantánea la acción de v , no puede mantener por ejemplo $x_1 = 0$ durante un intervalo finito de tiempo. La conclusión es que el juego diferencial nunca termina y la solución de Pontrjagin es falsa. Sin embargo, si escribimos las ecuaciones dinámicas en forma polar, esto es

$$\dot{\rho} = 2 u_p + v_p \quad \dot{\theta} = \rho^{-1} (2u_t + v_t)$$

donde u_p y u_t son las componentes radial y tangencial del vector u , respectivamente; lo mismo para el vector v . Suponiendo que podamos elegir siempre

$$u_p = -1; u_t = 0$$

esto es, u apuntando siempre hacia el origen, resulta

$$\dot{\rho} = -2 + v_p \quad \dot{\theta} = \rho^{-1} v_t$$

Siendo $|v_p| \leq 1$ se deduce que $\dot{\rho} \leq 1$ y entonces ρ disminuye hasta cero, alcanzando el origen en tiempo finito. Aún cuando este resultado es matemáticamente correcto, es evidente que físicamente carece de sentido ya que cerca del origen implicaría grandes cambios de dirección en tiempos muy cortos.

La explicación es que el problema de alcanzar el origen es relativamente artificial y difícilmente aparece en juegos diferenciales. Lo que sí es típico, es plantear como objetivo el alcanzar una variedad de dimensión arbitraria.

En este caso, es preciso reemplazar el origen por un entorno circular de radio ϵ , reduciendo el problema al caso anterior. Lo importante sería entonces la "ε-controlabilidad", que significa alcanzar el entorno ϵ del origen en tiempo finito.

Ejemplo 3

Consideremos el problema de control

$$\dot{x}_1 = x_1 + 2u$$

$$\dot{x}_2 = v$$

donde $|u| \leq 1$; $|v| \leq 1$ y M es el plano $x_1 = 0$. Entonces

$$H = \Psi_1 x_1 + 2 \Psi_1 u + \Psi_2 v$$

siendo

$$\dot{\Psi}_1 = -\Psi_1 \quad \Psi_1(t) = K \cos(t + \phi)$$

$$\dot{\Psi}_2 = 0 \quad \Psi_2(t) = \text{const.}$$

Tomando como condiciones finales

$$\begin{array}{ll} x_1(0) = 0 & x_2(0) = \text{arbitrario} \\ \Psi_2(0) = 0 & \Psi_1(0) = 1 \end{array}$$

resulta $K = 1$; $\varphi = 0$ ó $\varphi = \pi$

Las estrategias óptimas son las siguientes:

$$\bar{u} = -\operatorname{sg} \Psi_1(t) = -\operatorname{sg} \cos t = \operatorname{sg} \cos(t + \pi) \quad \text{H máximo}$$

$$\bar{v} = -\operatorname{sg} \Psi_2(t) = -\operatorname{sg} \Psi_2 \quad \text{H mínimo}$$

Estas dos ecuaciones son independientes y en realidad es un problema de control para u (alcanzar $x_1 = 0$ en tiempo mínimo) y v no puede interferir. En este caso, se prueba que usando las ecuaciones de Pontrjagin para $x_1 < 2$, la estrategia óptima no sirve para alcanzar el origen. Si suponemos $x_1 > 0$, entonces u busca hacer x_1 tan negativo como sea posible y elige $u = -1$ y entonces $x_1 < 2$ pudiendo u forzar al punto hacia el eje vertical desde la derecha.

Ejemplo 4

Sea el sistema de control

$$\dot{x}_1 = x_2 + 2u$$

$$\dot{x}_2 = -x_1 + v$$

donde $|u| \leq 1$; $|v| \leq 1$ siendo M el plano $x_2 = 0$. El hamiltoniano resulta

$$H = -\Psi_1 x_2 - \Psi_2 x_1 + 2 \Psi_1 u + \Psi_2 v$$

siendo

$$\dot{\Psi}_1 = -\Psi_2 \quad \dot{\Psi}_2 = -\Psi_1$$

de donde

$$\Psi_1(t) = K \cos(t + \varphi) \quad \Psi_2(t) = -K \operatorname{sen}(t + \varphi)$$

con condiciones finales

$$\begin{array}{ll} x_2(0) = 0 & x_1(0) = \text{arbitrario} \\ \Psi_1(0) = 0 & \Psi_2(0) = \text{arbitrario p.ej: } \Psi_2(0) = 1 \end{array}$$

Así resulta $K = 1$; $\varphi = \pi$ o $\varphi = 3\pi/2$

Las estrategias óptimas son las siguientes:

$$\bar{u} = \operatorname{sg} \Psi_1(t) = \operatorname{sg} \cos(t + \pi/2) \quad \text{para H máximo}$$

$$= \operatorname{sg} \cos(t + 3\pi/2)$$

$$\bar{v} = -\operatorname{sg} \Psi_2(t) = \operatorname{sg} \operatorname{sen}(t + \pi/2) \quad \text{para H mínimo}$$

$$= \operatorname{sg} \operatorname{sen}(t + 3\pi/2)$$

y las trayectorias están dadas por las ecuaciones

$$\begin{array}{ll} \dot{x}_1 = x_2 + 2 \operatorname{sg} \cos(t + \pi/2) & \dot{x}_2 = -x_1 + \operatorname{sg} \operatorname{sen}(t + \pi/2) \\ = x_2 + 2 \operatorname{sg} \cos(t + 3\pi/2) & = -x_1 + \operatorname{sg} \operatorname{sen}(t + 3\pi/2) \end{array}$$

Veamos cuál de ambas posibilidades $\varphi = \pi/2$ ó $\varphi = 3\pi/2$ es mejor. Antes de finalizar el juego, si $x_2 > 0$ o $\dot{x}_2 < 0$, v trata de posponer la finalización y elige por ello $v = 1$. Entonces el juego puede terminar desde arriba ($x_2 > 0$) si y sólo si $-x_1 + v = -x_1 + 1 < 0$ lo que implica $x_1 > 1$. Análogamente, el juego puede terminar desde abajo sólo si $x_1 < -1$. Las trayectorias óptimas están graficadas en la Fig. 2. En el cuadrante $x_1 > -1$, $x_2 > 0$ es $u = v = 1$ y las trayectorias son círculos de centro (1; -2). Estas trayectorias alcanzan el eje x_1 para $x_1 > 1$. En el cuadrante $x_1 < -1$, $x_2 > 0$ es $u = 1$; $v = -1$ y las trayectorias resultan círculos de centro (1; 2). Tales trayectorias comienzan sobre el eje x_1 y se unen a las trayectorias del caso anterior. En el semiplano $x_2 < 0$, el comportamiento es similar.

Conclusiones

El análisis de estos juegos diferenciales lineales bidimensionales, aplicando metodologías de la teoría de control óptimo, muestra que aún cuando no se puede negar que el enriquecimiento mutuo de ambos campos resulta fructífero, hay que ser muy cuidadoso al trasladar métodos y resultados de uno a otro dominio.

Es de esperar que las relaciones entre los conjuntos alcanzables, los conjuntos de "espera" y conjuntos límite para sistemas de control (ver Roxin y Spinadel, 1986), puedan ser extendidos de algún modo a juegos diferenciales.

Referencias

- [1] DAVIS H.T.(1962). *Introduction to Nonlinear Differential and Integral Equations*, Dover Publications Inc., New York, pp. 113-27.
- [2] BOUGUER P. (1732). *Sur de nouvelles courbes ausquelles on peut donner de nom de Lignes de Poursuite*, Histoire de l'Academie Royale des Sciences, París, pp. 1-14.
- [3] ISAACS R. (1953-1956). *Differential Games I, II, III y IV*. Rand Corp. Research Memorandum RM-1391, 1399, 1411 y 1468.
- [4] ISAACS R. (1956). *Differential Games*, John Wiley & Sons.
- [5] BELLMAN R.I., GLICKSBERG y GROSS O. (1956). *On the Bang-bang Control Problem* Quart. Appl. Math., 14, pp. 11-18.
- [6] ROXIN E. Y SPINADEL V. (1959). *Reachable Zones in Autonomous Differential Systems* Contr. to Diff. Eqs., 1, No.3, pp. 275-315.
- [7] PONTRJAGIN L.S., BOLTJANSKII V.G., GAMKRELIDZE R.V. Y MISCHENKO E.F. (1962). *The Mathematical Theory of Optimal Processes*, Interscience, New York.
- [8] KELENDZERIDZE D.L. *Theory of an Optimal Pursuit Strategy*, Doklad. Akad. Nauk USSR (traducido en Sov. Math., 1961).
- [9] KRASOVSKII N.N. y SUBBOTIN A.I. (1988). *Game-Theoretic Control Problems*. Springer-Verlag.
- [10] BERKOVITZ L.D. (1961). *A Variational Approach to Differential Games*, Rand Report RM-2772.
- [11] FLEMING W.H. (1957). *A Note on Differential Games of Prescribed Duration*, Contr. to the Theory of Games III, Annals of Mathem. Studies 39, Princeton Univ., Princeton, N.J.
- [12] LASALLE J.P. (1960) *The Time Optimal Control Problem*, Contr. to the Theory of Nonlinear Oscillations, V, pp. 1-24, Princeton.
- [13] BUSHAW D.W. (1958) *Optimal Discontinuous Forcing Terms*, Contr. to the Theory of Nonlinear Oscillations, IV, Princeton.
- [14] GAMKRELIDZE R.V. (1958). *Theory of Time-Optimal Processes for Linear Systems*, Izv. Akad. Nauk. USSR. Ser Mathe. 22, pp. 449-474.
- [15] KRASOVSKII N.N. (1957). *Concerning the Theory of Optimal Control*. Avtomat. i Telemekh., 18, pp. 960-970.

- [16] ROXIN E. (1969). *Axiomatic Approach in Differential Games*. Journ. of Optimiz.Theory and Appl., 3,3.
- [17] CLARKE F.H. Y LOEWEN P.D. (1986). *The Value Function in Optimal Control : Sensivity, Controllability and Time-Optimality*. Siam J. Control and Optimiz., 24, No. 2, pp. 243-263.
- [18] VARAIYA P.P. (1967). *The Existence of Solutions to a Differential Game*. J. Soc. Ind. Appl. Math., Ser. A, Control 5, pp. 153-162.
- [19] FRIEDMAN A. (1970). *Existence of Value and Saddle Points for Differential Games of Pursuit and Evasion*, J. Diff. Eq., 7, pp. 69-91.
- [20] FRIEDMAN A. (1970). *Existence of Value and Saddle Points for Differential Games of Survival*. J. Diff. Eq.,7, pp. 111-125.
- [21] FRIEDMAN A. (1970). *Existence of Value and Saddle Points for Differential Games of Pursuit and Evasion*. J.Diff. Eq., 7, pp. 92-110.
- [22] ROXIN E. (1970). *On Differential Games without Value*, Symposium on the American Automatic Control Council.
- [23] CILETTI M.D. Y STARR A.W. (1970). *Differential Games, a Critical View*. Simposium on The American Automatic Control Council.
- [24] MEHLMANN A. (1988). *Applied Differential Games*, Plenum Publish. Corp.
- [25] PONTRJAGIN L.S. (1964). *On some Differential Games* (traducido en Soviet Math. 5, No. 3).
- [26] ROXIN E. Y SPINADEL V. (1986). *On Limit Sets of Linear Control Systems*. SIAM Conf. on Linear Algebra in Signals, Systems and Control, Boston, Massachusetts.

Departamento de Matemática.
Facultad de Ciencias Exactas
Universidad de Buenos Aires

Recibido por la U.M.A. el 16 de Mayo de 1989.

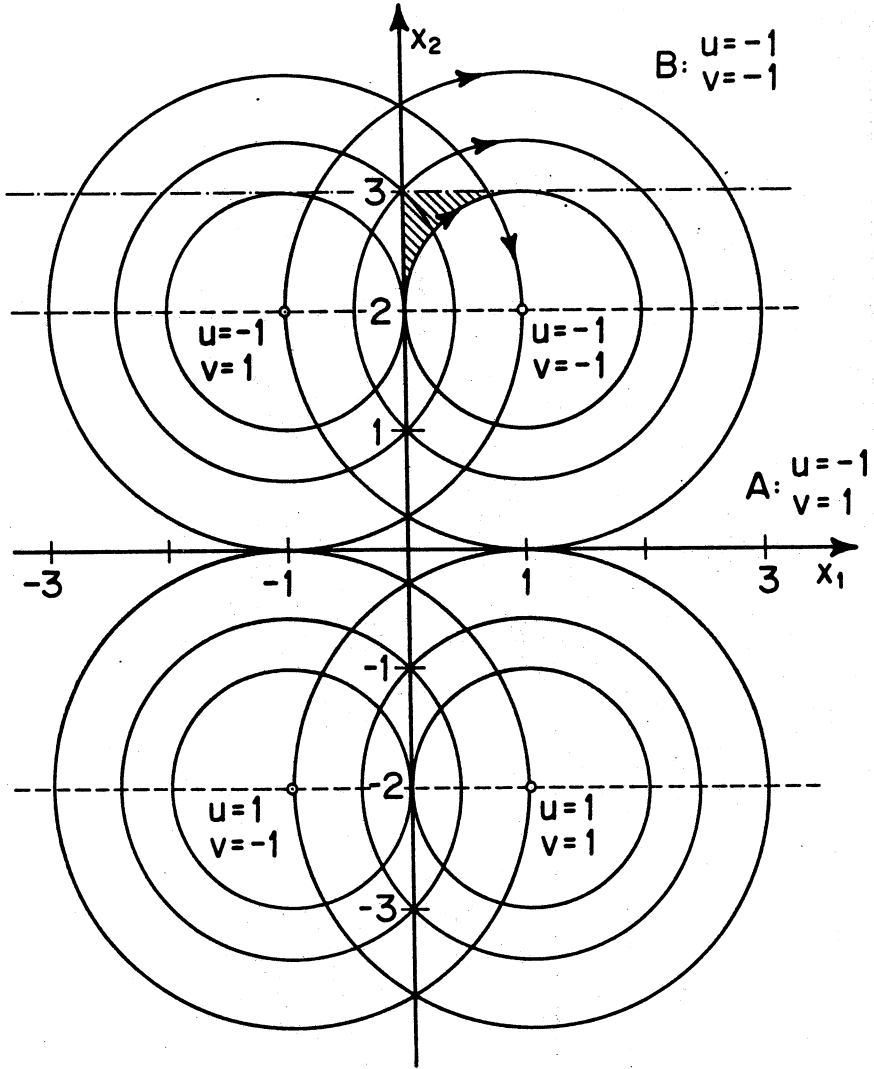


Figura 1

Estrategias óptimas

Para $x_1 > 0$ Región A: $u = -1; v = +1$
 Regiones B y C: $v = -1; v = -1$
 Para $x_1 < 0$ las soluciones son simétricas.

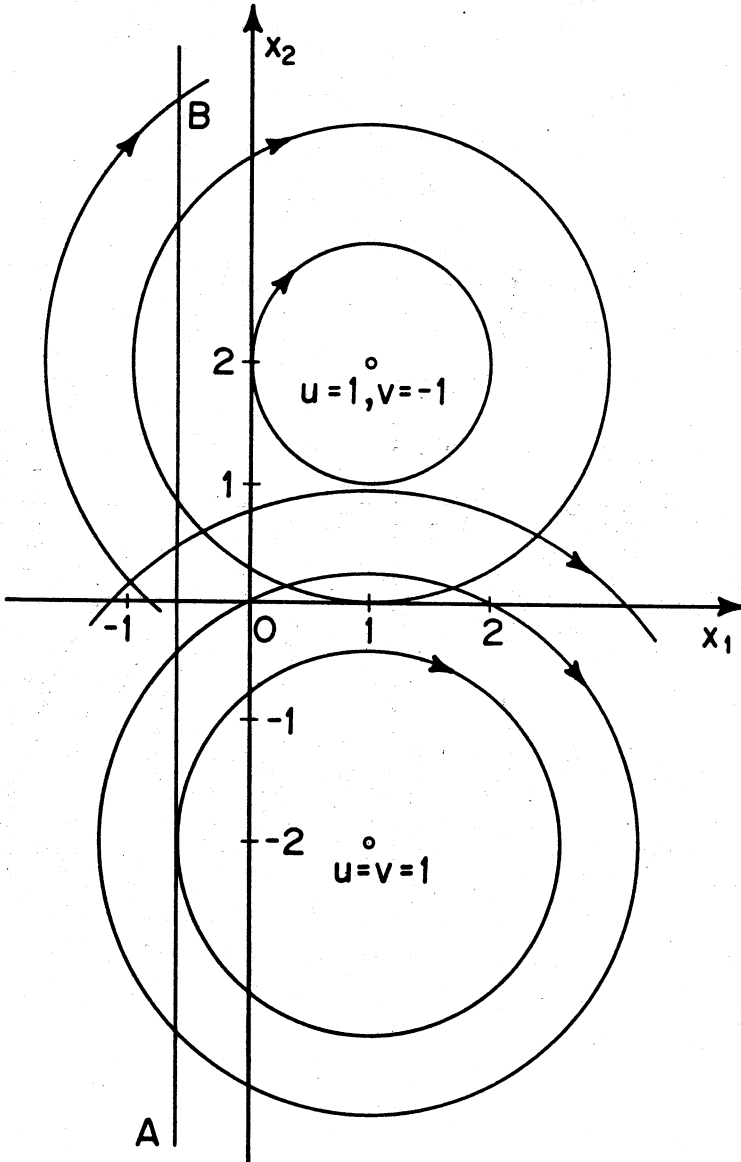


Figura 2