TRABAJOS DE INVESTIGACION OPERATIVA
Vol. 7, núm. 1, 1992, pp. 77 a 86

CARACTERIZACION DE LA FUNCION DE VALOR DE LOS JUEGOS ESTOCASTICOS CONTINUOS

M. ANGELES MURUAGA Y R. VÉLEZ
UNED

RESUMEN

Se establece una caracterización de la función de valor de los juegos estocásticos continuos, similar a la contenida en [2] y [3] para juegos matriciales y en [4] para juegos estocásticos discretos. Tras la formulación del problema se señalan algunas propiedades de la función de valor. Más adelante se prueba que tales propiedades son suficientes para identificar el funcional que asigna a cada juego su valor.

ABSTRACT

The aim this paper is to prove a characterization of the value function for continuous stochastic games, similar to those given in [2] and [3] for matrix games and in [4] for discrete stochastic games. After the formulation of the problem, we point out some properties of the function value. We show that these properties are sufficient to identify it.
Key words: Continuous stochastic games. Discount factor. Value functional.
AMS 1980-*Subject Classification*: 90D15-93E05.

1. PLANTEAMIENTO

En el trabajo anterior [1] hemos considerado juegos estocásticos continuos, compuestos por una familia $\Gamma(a)$, $a \in A$, de juegos bipersonales

Recibido: Marzo 1991.
Revisado: Septiembre 1991.

de suma nula, tales que A y los conjuntos de estrategias X e Y son espacios métricos compactos. La función de pago $M(a, x, y)$ se supone continua y la función de transición débilmente continua.

Cuando el horizonte es infinito y se considera un factor de descuento, $\beta < 1$, se prueba que el valor $v_\beta(a)$ verifica

$$v_\beta(a) = \text{Val}\ \{M(a, x, y) + \beta \int_A v_\beta(b)P(x, y, a, db)\} \tag{1}$$

También se establece que ambos jugadores diponen de estrategias óptimas estacionarias, que consisten en mantener fijas las estrategias óptimas, $\xi_\beta^*(\cdot)$ y $\eta_\beta^*(\cdot)$, del juego que figura en la ecuación anterior.

La función de valor de un juego de suma nula fue caracterizada por Vilkas (1963) y Tijs (1981), en el sentido de expresar las propiedades que identifican el funcional que asigna, a cada uno de estos juegos, su valor.

Posteriormente Vrieze (1987) generaliza esta caracterización en el caso de juegos estocásticos discretos. Nuestro objetivo es obtener resultados similares en el caso continuo.

Para ello, consideraremos el funcional ϕ_β que asigna a cada juego estocástico continuo Γ, con espacio de estados A y horizonte infinito, su función de valor con factor de descuento determinado, $\beta < 1$. Es decir

$$\phi_\beta(\Gamma) = v_\beta$$

Se trata, entonces, de determinar propiedades de ϕ_β, suficientes para identificar dicho funcional.

2. PROPIEDADES DE LA FUNCION DE VALOR

Las primeras propiedades del valor de un juego estocástico continuo son fáciles de establecer:

Propiedad 1: *Si Γ es un juego estocástico continuo, con espacio de estados A, tal que para algún estado $a_0 \in A$ se verifica*

$$M(a_0, x, y) = c \quad y \quad P(x, y, a_0, \{a_0\}) = 1 \quad para\ cada\ (x, y) \in X \times Y$$

entonces

$$\phi_\beta(\Gamma)(a_0) = c(1 - \beta)^{-1}$$

Propiedad 2: *Si* Γ *y* Γ' *son dos juegos estocásticos continuos que sólo se diferencian en la función de pago, verificándose*

$$M(a, x, y) \leqslant M'(a, x, y) \quad \text{para todo } (a, x, y) \in A \times X \times Y$$

entonces

$$\phi_\beta(\Gamma)(a) \leqslant \phi_\beta(\Gamma')(a) \quad \text{para cada } a \in A$$

De cara a enunciar la tercera propiedad del funcional de valor, es necesario generalizar el concepto de estrategia superflua para un jugador, introducido por Vrieze:

Un conjunto medible, $F_0 \subset X$, *de estrategias del primer jugador se denomina superfluo para el estado* $a_0 \in A$, *en un juego estocástico continuo de función de valor* v_β, *si existe una estrategia* $\xi_0 \in \bar{X}$ *del primer jugador, con* $\xi_0(F_0) = 0$ *y tal que, para cada* $x \in F_0$ *y todo* $y \in Y$, *se verifique*

$$M(a_0, x, y) + \beta \int_A v_\beta(b) P(x, y, a_0, db) \leqslant$$

$$\leqslant M(a_0, \xi_0, y) + \beta \int_A v_\beta(b) P(\xi_0, y, a, db)$$

Simétricamente pueden definirse los conjuntos de estrategias superfluas $G_0 \subset Y$, para el segundo jugador, en un estado $a_0 \in A$, mediante la condición

$$M(a_0, x, y) + \beta \int_A v_\beta(b) P(x, y, a_0, db) \geqslant$$

$$\geqslant M(a_0, x, \eta_0) + \beta \int_A v_\beta(b) P(x, \eta_0, a, db)$$

para todo $x \in X$, $y \in G_0$ y para algún $\eta_0 \in \bar{Y}$ con $\eta_0(G_0) = 0$.

Se puede probar que, salvo en el caso en que la estrategia ξ_0 que hace superfluo a un conjunto de estrategias F_0, pondere exclusivamente estrategias puras de la frontera de F_0, la adherencia $\bar{F}_0$ será también superflua. De manera que con gran frecuencia los conjuntos superfluos de estrategias de cada jugador serán cerrados.

En el caso de juegos estocásticos discretos, en los que no hay ninguna estructura subyacente que conservar, puede demostrarse (véase

[4]) que se puede prescindir, en cada estado, de cada una de las estrategias puras superfluas sin alterar el valor del juego.

En el caso continuo, han de tomarse mayores precauciones para no alterar la continuidad de las funciones de pago y de transición del juego; y, simultáneamente, conservar homogéneo el conjunto de estrategias puras de cada jugador, en los diversos estados del sistema.

Nos limitaremos pues a probar que puede prescindirse de cualquier componente conexa del espacio de estrategias puras de cada jugador, en el caso de que constituya un conjunto superfluo de estrategias, simultáneamente para todos los estados del juego.

Propiedad 3.1: *Si en un juego estocástico continuo Γ, con función de valor v_β, $F_0 \subset X$ es un conjunto cerrado de estrategias superfluas del primer jugador, para todos los estados $a \in A$, tal que F_0^c también es cerrado, entonces el juego estocástico Γ' obtenido suprimiendo F_0 del conjunto de estrategias puras del primer jugador, tiene el mismo valor v_β. Es decir $\phi_\beta(\Gamma) = \phi_\beta(\Gamma')$.*

Demostración

Sea $\bar{X}_0$ el conjunto de estrategias mixtas del primer jugador tales que $\xi(F_0) = 0$. Por hipótesis, existe $\xi_0 \in \bar{X}_0$ tal que $H(a, \xi_0) \geqslant 0$ siendo $H(a, \xi)$ el ínfimo, para $x \in F_0$ e $y \in Y$, de

$$M(a, \xi, y) + \beta \int_A v_\beta(b)P(\xi, y, a, db) - M(a, x, y) - \beta \int_A v_\beta(b)P(x, y, a, db)$$

Es fácil probar que $H(a, \xi)$ es una función continua en $A \times \bar{X}_0$, y, como $\bar{X}_0$ es compacto, el teorema de selección de Dubins y Savage asegura la existencia de una función medible $\bar{\xi}(a)$, de A en $\bar{X}_0$, tal que

$$H(a, \bar{\xi}(a)) = \sup_{\xi \in \bar{X}_0} H(a, \xi)$$

de forma que, para todo $x \in F_0$ e $y \in Y$,

$$M(a, \bar{\xi}(a), y) + \beta \int_A v_\beta(b)P(\bar{\xi}(a), y, a, db) \geqslant$$

$$\geqslant M(a, x, y) + \beta \int_A v_\beta(b)P(x, y, a, db)$$

Es decir, que $\bar{\xi}(a)$ hace superfluo a F_0 en el estado a, para cada $a \in A$.

Sea $\xi_\beta^*(\cdot)$ la estrategia óptima del primer jugador en el juego Γ; de manera que, según (1), es

$$M(a,\ \xi_\beta^*(a),\ y) + \beta \int_A v_\beta(b) P(\xi_\beta^*(a),\ y,\ a,\ db) \geqslant v_\beta(a)$$

para cada $y \in Y$.

La medibilidad de $\bar{\xi}$ y ξ_β^* permite considerar la estrategia medible definida, para cada conjunto de Borel $B \subset X$, por:

$$\tilde{\xi}_\beta(a)(B) = \xi_\beta^*(a)(B \cap F_0^c) + \xi_\beta^*(a)(F_0)\bar{\xi}(a)(B)$$

que evidentemente verifica

$$\tilde{\xi}_\beta(a)(F_0^c) = \xi_\beta^*(a)(F_0^c) + \xi_\beta^*(a)(F_0) = 1$$

y con lo cual se tiene, por una parte

$$M(a, \tilde{\xi}_\beta(a), y) = \int_{F_0^c} M(a, x, y)\xi_\beta^*(a)(dx) + \xi_\beta^*(a)(F_0) \int_{F_0^c} M(a, x, y)\bar{\xi}(a)(dz) =$$

$$= M(a,\ \xi_\beta^*(a),\ y) - \int_{F_0} M(a,\ x,\ y)\xi_\beta^*(a)(dx) + \xi_\beta^*(a)(F_0)M(a,\ \bar{\xi}(a),\ y)$$

y, por otra parte,

$$P(\tilde{\xi}_\beta(a), y, a, B)\ y) = \int_{F_0^c} P((x, y, a, B)\xi_\beta^*(a)(dx) +$$

$$+ \xi_\beta^*(a)(F_0) \int_{F_0^c} P(z, y, a, B)\bar{\xi}(a)(dx) =$$

$$= P(\xi_\beta^*(a),\ y,\ a,\ B) - \int_{F_0} P(x,\ y,\ a,\ B)\xi_\beta^*(a)(dx) + \xi_\beta^*(a)(F_0)P(\bar{\xi}(a),\ y,\ a,\ B)$$

Resulta entonces

$$M(a,\ \tilde{\xi}_\beta(a),\ y) + \beta \int_A v_\beta(b)P(\tilde{\xi}(a),\ y,\ a,\ db) =$$

$$= M(a,\ \xi_\beta^*(a),\ y) + \beta \int_A v_\beta(b)P(\xi_\beta^*(a),\ y,\ a,\ db) -$$

$$-\int_{F_0} M(a, x, y)\xi_\beta^*(a)(dx) - \beta \int_A v_\beta(b) \int_{F_0} P(x, y, a, db)\xi_\beta^*(a)(dx) +$$

$$+ \xi_\beta^*(a)(F_0)\{M(a, \bar{\xi}(a), y) + \beta \int_A v_\beta(b)P(\bar{\xi}(a), y, a, db)\} \geqslant$$

$$\geqslant v_\beta(a) + \xi_\beta^*(a)(F_0)\left\{M(a, \bar{\xi}(a), y) + \beta \int_A v_\beta(b)P(\bar{\xi}(a), y, a, db) -\right.$$

$$\left. - \int_{F_0} \left[M(a, x, y) + \beta \int_A v_\beta(b)P(x, y, a, db)\right] \frac{\xi_\beta^*(a)(dx)}{\xi_\beta^*(a)(F_0)}\right\} \geqslant v_\beta(a)$$

En definitiva, en el juego de función de pago

$$M(a, x, y) + \beta \int_A v_\beta(b)P(x, y, a, db) \quad \text{con} \quad x \in F_0^c \quad \text{e} \quad y \in Y$$

el primer jugador puede asegurarse una ganancia superior o igual a $v_\beta(a)$ (utilizando únicamente estrategias de F_0^c). Mientras que, por supuesto, el segundo jugador puede garantizar que el pago no será superior a $v_\beta(a)$, mediante su estrategia óptima $\eta_\beta^*(a)$. El juego restringido anterior tiene pues valor igual a $v_\beta(a)$.

Desde luego la situación es simétrica para el segundo jugador, en el sentido de que se verifica:

Propiedad 3.2: *Si en un juego estocástico continuo, Γ, con función de valor v_β, $G_0 \subset Y$ es un conjunto cerrado de estrategias superfluas del segundo jugador, para todos los estados $a \in A$, tal que G_0^c es también cerrado, entonces el juego estocástico continuo Γ'' obtenido suprimiendo G_0 del conjunto de estrategias puras del segundo jugador, tiene el mismo valor v_β. Es decir, $\phi_\beta(\Gamma) = \phi_\beta(\Gamma'')$.*

3. CARACTERIZACION DE LA FUNCION DE VALOR

Una vez establecidas las propiedades (1), (2), (3.1) y (3.2) del funcional de valor ϕ_β, el resultado de mayor interés es establecer que tales propiedades caracterizan el valor. En el sentido siguiente:

Teorema: *Si un funcional ϕ sobre el conjunto de los juegos estocásticos* *continuos con espacio de estados A, verifica las propiedades* (1), (2), (3.1) *y* (3.2), *entonces asigna a cada juego su valor con factor de descuento* β. *Es decir*, $\phi = \phi_\beta$.

Demostración

a) Consideremos en primer lugar un juego estocástico continuo Γ, tal que en X e Y existen puntos aislados, x_0 e y_0 respectivamente, para los cuales se verifica:

$$P(x_0, y, a, \{a\}) = 1 = P(x, y_0, a, \{a\}) \quad \text{para cada} \quad x \in X, y \in Y \quad \text{y} \quad a \in A$$

y

$$M(a, x_0, y_0) = \inf_{y \in Y} M(a, x_0, y) = \sup_{x \in X} M(a, x, y_0) \quad \text{para cada} \quad a \in A$$

El valor del juego estocástico Γ es entonces $(1-\beta)^{-1}M(a, x_0, y_0)$.

En efecto, puesto que el primer jugador puede elegir la estrategia pura x_0, en cualquier estado $a \in A$,

$$\sup_{\{\xi_n(\cdot)\}_{n=1}^\infty \in U^\infty} \inf_{\{\eta_n(\cdot)\}_{n=1}^\infty \in V^\infty} m_\beta(a, \{\xi_n(\cdot)\}_{n=1}^\infty, \{\eta_n(\cdot)\}_{n=1}^\infty) \geqslant$$

$$\geqslant \inf_{\{\eta_n(\cdot)\}_{n=1}^\infty \in V^\infty} \sum_{n=1}^\infty \beta^{n-1} M(a, x_0, \eta_n(a)) \geqslant (1-\beta)^{-1} M(a, x_0, y_0)$$

y simétricamente

$$\inf_{\{\eta_n(\cdot)\}_{n=1}^\infty \in V^\infty} \sup_{\{\xi_n(\cdot)\}_{n=1}^\infty \in U^\infty} m_\beta(a, \{\xi_n(\cdot)\}_{n=1}^\infty, \{\eta_n(\cdot)\}_{n=1}^\infty) \leqslant (1-\beta)^{-1} M(a, x_0, y_0)$$

Analicemos ahora la función que un funcional ϕ, que cumpla las propiedades establecidas, asigna al juego estocástico Γ. Para ello consideremos los juegos estocásticos continuos, Γ' y Γ'', idénticos a Γ excepto en la función de pago, que definiremos respectivamente, como

$$M'(a, x, y) = \begin{cases} M(a, x, y) & \text{si} \quad x = x_0 \text{ ó } y = y_0 \\ M(a, x, y) - k & \text{si} \quad x \neq x_0 \text{ e } y \neq y_0 \end{cases}$$

$$M''(a, x, y) = \begin{cases} M(a, x, y) & \text{si} \quad x = x_0 \text{ ó } y = y_0 \\ M(a, x, y) + k & \text{si} \quad x \neq x_0 \text{ e } y \neq y_0 \end{cases}$$

Según la propiedad (2), tendremos $\phi(\Gamma') \leqslant \phi(\Gamma) \leqslant \phi(\Gamma'')$.

Ahora bien, el mismo razonamiento realizado para Γ, muestra que el valor de Γ' y Γ'' es independiente del valor de la constante k. Sin embargo, si k es suficientemente grande, todas las estrategias de $X - \{x_0\}$ son superfluas, en el juego Γ', frente a x_0; pues basta para ello que sea, para cualquier $y \neq y_0$ y $x \neq x_0$,

$$M(a, x, y) - k + \frac{\beta}{1-\beta} \int_A M(b, x_0, y_0) P(x, y, a, db) \leqslant$$

$$\leqslant M(a, x_0, y) + \frac{\beta}{1-\beta} M(a, x_0, y_0)$$

lo cual se puede conseguir tomando $k \geqslant (1-\beta)^{-1} 2M$ (siendo M la cota de $M(a, z, y)$).

Según la propiedad (3.1), el juego $\bar{\Gamma}'$ obtenido suprimiendo todas las estrategias de X excepto x_0, cumple $\phi(\bar{\Gamma}') = \phi(\Gamma')$.

Además en $\bar{\Gamma}'$, todas las estrategias de $Y - \{y_0\}$ son superfluas frente a y_0, puesto que $M(a, x_0, y) \geqslant M(a, x_0, y_0)$ para todo $y \neq y_0$. Luego el juego $\bar{\bar{\Gamma}}'$ obtenido suprimiendo todas las estrategias de Y, excepto y_0, verifica $\phi(\bar{\bar{\Gamma}}') = \phi(\bar{\Gamma}') = \phi(\Gamma')$.

Pero, al juego $\bar{\bar{\Gamma}}'$, se le puede aplicar la propiedad (1) para deducir

$$\phi(\bar{\bar{\Gamma}}') = (1-\beta)^{-1} M(\cdot, x_0, y_0)$$

Un razonamiento similar con el juego Γ'' permite establecer que

$$\phi(\Gamma'') = (1-\beta)^{-1} \mathrm{M}(\cdot, x_0, y_0)$$

y en definitiva resulta $\phi(\Gamma) = (1-\beta)^{-1} \mathrm{M}(\cdot, \mathrm{x}_0, \mathrm{y}_0)$. De manera que el funcional ϕ asigna a los juegos estocásticos del tipo particular considerado, su función de valor.

b) Sea ahora Γ un juego estocástico continuo arbitrario, con espacio de estados A, cuya función de valor designaremos por v_β.

Construyamos a partir de Γ un nuevo juego estocástico continuo $\tilde{\Gamma}$, añadiendo a los espacios de estrategias puras, X e Y, sendos puntos aislados, x_0 e y_0, para los cuales se tenga

$$\begin{cases} M(a, x_0, y) = M(a, x, y_0) = (1-\beta) v_\beta(a) \\ P(x_0, y, a, \{a\}) = P(x, y_0, a, \{a\}) = 1 \end{cases}$$

cualquiera que sean $a \in A$, $x \in X$ e $y \in Y$.

El juego estocástico $\tilde{\Gamma}$ sigue teniendo función de valor v_β ya que, si ξ_0 representa la estrategia del primer jugador que elige x_0 en cualquier estado $a \in A$, se tiene, para cualquier $\{\eta_n(\cdot)\}_{n=1}^{\infty} \in V^{\infty}$

$$m_\beta(a, \{\xi_0\}_{n=1}^{\infty}, \{\eta_n(\cdot)\}_{n=1}^{\infty}) = \sum_{n=1}^{\infty} \beta^{n-1} M(a, x_0, \eta_n(a)) = v_\beta(a)$$

de manera que

$$\sup_{\{\xi_n(\cdot)\}_{n=1}^{\infty} \in U^{\infty}} \inf_{\{\eta_n(\cdot)\}_{n=1}^{\infty} \in V^{\infty}} m_\beta(a, \{\xi_n(\cdot)\}_{n=1}^{\infty}, \{\eta_n(\cdot)\}_{n=1}^{\infty}) \geqslant v_\beta(a)$$

y análogamente

$$\inf_{\{\eta_n(\cdot)\}_{n=1}^{\infty} \in V^{\infty}} \sup_{\{\xi_n(\cdot)\}_{n=1}^{\infty} \in U^{\infty}} m_\beta(a, \{\xi_n(\cdot)\}_{n=1}^{\infty}, \{\eta_n(\cdot)\}_{n=1}^{\infty}) \leqslant v_\beta(a)$$

Además $\tilde{\Gamma}$ es un juego estocástico del tipo considerado en *a*), de forma que se tiene $\phi(\tilde{\Gamma}) = v_\beta$.

Sea ahora $\xi_\beta^*(\cdot)$ la estrategia óptima para el primer jugador en el juego Γ, que cumplirá, de acuerdo con (1),

$$M(a, \xi_\beta^*(a), y) + \beta \int_A v_\beta(b) P(\xi_\beta^*(a), y, a, db) \geqslant v_\beta(a)$$

para todo $y \in Y$ y cada $a \in A$.

Podemos ver que, en el juego $\tilde{\Gamma}$, la estrategia $\xi_\beta^*(a)$ hace superflua la estrategia pura x_0 en cada estado $a \in A$. En efecto, para cada y (incluido y_0)

$$M(a, x_0, y) + \beta \int_A v_\beta(b) P(x_0, y, a, db) = (1-\beta) v_\beta(a) + \beta v_\beta(a) = v_\beta(a)$$

mientras que

$$M(a, \xi_\beta^*(a), y_0) + \beta \int_A v_\beta(b) P(\xi_\beta^*(a), y_0, a, db) =$$

$$= (1-\beta) v_\beta(a) + \beta v_\beta(a) = v_\beta(a)$$

De acuerdo con la propiedad (3.1) el juego $\tilde{\Gamma}'$, que resulta de $\tilde{\Gamma}$ suprimiendo la estrategia pura x_0, verifica $\phi(\tilde{\Gamma}') = \phi(\tilde{\Gamma}) = v_\beta$.

De forma análoga, la estrategia óptima $\eta_\beta^*(\cdot)$ del segundo jugador en el juego Γ, hace superflua la estrategia pura y_0, para cualquier estado $a \in A$.

Suprimiendo la estrategia y_0 del espacio de estrategias puras, el juego $\tilde{\Gamma}'$ queda reducido al juego original Γ, y según la propiedad (3.2) se verifica

$$\phi(\Gamma) = \phi(\tilde{\Gamma}') = v_\beta$$

Ello demuestra el teorema y establece una caracterización axiomática de la función de valor de los juegos estocásticos continuos.

REFERENCIAS

1. MURUAGA, M. A., y VELEZ, R. (1991): «Juegos estocásticos continuos: valor y estrategias óptimas», *Trabajos de Investigación Operativa*, vol. 7.
2. TIJS, S. H. (1981): «Characterization of the value of zero-sum two-person games», *Nav. Res. Log. Quart.*, vol. 28, 153-156.
3. VILKAS, E. I. (1963): «Axiomatic definition of the value of a matrix game», *Theory of Prob. and its Appl.*, vol. 8, 304-307.
4. VRIEZE, O. J. (1987): «Stochastic games with finite state and action spaces», *Centrum voor Wiskunde en Informatica*, 33.