Índice general

Capítulo 1

Introducción

El trabajo realizado en la presente tesis está relacionado con la teoría de Procesos de Decisión y con la teoría de Juegos. Un Proceso de Decisión es una sucesión de decisiones realizadas en un tiempo determinado siguiendo una estrategia y pagando un costo por cada decisión realizada. Por otro lado, la teoría de juegos es una área de las matemáticas que estudia la interacción entre individuos (jugadores) que llevan a cabo procesos de decisión. En la tesis se busca optimizar un problema relacionado con la operación eficiente de máquinas, primero se resolverá por medio de los procesos de decisión usando programación dinámica y después por medio de teoría de juegos usando la definición de equilibrio de Nash.

En particular se estudiarán los procesos de decisión de Markov (PDM) a tiempo discreto (ver [5] y [10]). Un PDM es usado para modelar un sistema que es observado de forma discreta en un tiempo finito o infinito, y son muy útiles en áreas como economía, donde son usados para optimizar problemas de crecimiento económico (ver [11]).

El desarrollo de un PDM, puede ser descrito de la forma siguiente: suponga que en cada tiempo t, $t = 0, 1, \ldots, T$, donde T es el horizonte del problema, la política decide que acción será aplicada dependiendo del estado en el qué se encuentre el sistema. Entonces como consecuencia se paga un costo y el sistema cambia de estado en el instante $t+1$, de acuerdo a una ley de transición y se repite el procedimiento.

Una política es una regla mediante la cual se elige una acción en cada punto de observación del proceso (ver [10] y [14]). Para evaluar la eficiencia de cada política se cuenta con un criterio de rendimiento, en este trabajo se usará el criterio de costo total acumulado.

El problema de control óptimo consiste en encontrar una política que optimice el criterio de rendimiento. La política que optimiza el criterio de rendimiento es llamada política óptima y, al criterio de rendimiento evaluado en tal política, función de valor óptima. Una manera de resolver un PDM está basado en el principio de Bellman conocido como Programación Dinámica. El principio de Programación Dinámica permite resolver problemas en los que es necesario tomar decisiones en etapas sucesivas. Las decisiones tomadas en una etapa condicionan la evolución futura del sistema, afectando a las situaciones en las que el sistema se encontrará en el futuro (estados), y a las decisiones (acciones) que se plantearán en el futuro. En el presente trabajo, el principio de programación dinámica proporciona una técnica para determinar de manera eficiente las estrategias que optimizan un problema (ver [4] y [14]).

En lo que respecta a la parte de teoría de juegos se estudiará el caso de dos jugadores con estrategias finitas y se supondrá además que los juegos son de suma no nula, es decir, la ganancia de un jugador no necesariamente se corresponde con la pérdida del otro (ver [12]). De nueva cuenta se tratará de encontrar la política óptima para ambos jugadores, para ello se introducirá el concepto de equilibrio de Nash (ver [3] y [9]). Un equilibrio de Nash (formulado por John Nash, ver [9]) se define como un modo de obtener una estrategia óptima para juegos que involucren a dos o más jugadores con estrategias finitas. Nash demostró que las distintas soluciones que habían sido propuestas anteriormente para juegos tienen la propiedad de producir un equilibrio de Nash.

Una estrategia pura, para un juego, puede no tener equilibrios de Nash o tener más de uno. Nash demostró que si se permiten estrategias mixtas (en las que los jugadores pueden escoger estrategias al azar con una probabilidad dada), entonces todos los juegos de n jugadores en los que cada jugador puede escoger entre un número finito de estrategias tienen al menos un equilibrio de Nash. Para probar esto Nash se basa en la correspondencia de mejor respuesta y el teorema del punto fijo de Kakutani (ver [2]); sin embargo la prueba es de existencia, es decir, asegura la existencia de equilibrios de Nash, pero no muestran como calcularlos, se verá una manera de hallar puntos de equilibrio de Nash por medio del cálculo y haciendo un análisis por casos.

La tesis está organizada de la siguiente forma: en el Capítulo 2 se presentan los conceptos generales de los procesos de decisión, se trabajan problemas con horizonte finito y con el criterio de rendimiento de costo total acumulado. También se presentan dos ejemplo relacionados con el reemplazamiento de máquinas (ver [5]). En el primero de ellos, se considera una sóla máquina y se busca encontrar su función óptima, en particular se presenta un caso donde se da la solución explícita y se incluye uno más donde se trabaja un número mayor de periodos, para ello se incluye un programa en MATLAB (ver [6]) que permite resolver dicho problema de una manera más sencilla. En el segundo problema se trabajan dos máquinas cuyo funcionamiento de una no depende de la otra, análogamente se incluye un ejemplo particular con un número de estados y periodos pequeño, se anexa un programa en MATLAB para resolverlo en un número mayor de estados. En el Capítulo 3 se desarrolla la Teoría de Juegos para dos jugadores, se da la definición de un equilibrio de Nash y se proporcionan dos maneras de resolverlos, una por medio del cálculo y la otra resolviendo por casos a través de las matrices de pago. Se incluye un ejemplo del reemplazamiento de dos máquinas, consideradas como los jugadores, sólo que en este caso se supone que su funcionamiento es dependiente una de la otra, y se analiza dicho juego en un solo periodo y lo que se busca es encontrar un equilibrio de Nash para los jugadores. Se incluye un programa en MATLAB para resolver este juego.

Capítulo 2

Procesos de Decisión

Un Proceso de Decisión es una sucesión de acciones realizadas en un tiempo determinado mediante una regla y pagando por cada acción realizada un costo. En este capítulo se darán las definiciones básicas para la teoría de los procesos de decisión con horizonte finito. Además se introducirá el concepto de Proceso de Decisión de Markov (PDM).

2.1. Procesos de Decisión

Las siguientes definiciones serán usadas en los capítulos siguientes (ver [13]).

Definición 2.1.1 *Sean $x \in X$ y $f(x)$ una función definida para cada $x \in X$, donde X es un conjunto numerable. Entonces se define* argmax *por la siguiente relación*

$$x^* \in \arg\max_{x \in X} f(x) \Leftrightarrow f(x^*) = \max_{x \in X} f(x).$$

Observación 2.1.2 *Es decir,* $\arg\max f = \left\{ x^* \in X : f(x^*) = \max_{x \in X} f(x) \right\}$.

La selección de un comportamiento en un problema de decisión simple es llamada *acción*. El conjunto de todas las acciones posibles será denotado por A. En la tesis se supone que A es un subconjunto numerable.

Definición 2.1.3 *Un* pago *es una función $u : A \to \mathbb{R}$ que asocia un valor numérico a cada acción $a \in A$.*

Definición 2.1.4 *Una acción a^* es óptima si para cada $a \in A$*

$$u(a^*) \geq u(a)$$

o, equivalentemente,

$$a^* \in \arg\max_{a \in A} u(a).$$

Definición 2.1.5 *Una* estrategia *o* política *es una regla mediante la cual se elige una acción en cada punto de decisión. Una* estrategia pura *es aquella en la cual no hay aleatorización en su implementación y/o selección. El conjunto de todas las posibles estrategias es denotado por S.*

Definición 2.1.6 *Una* estrategia mixta *σ especifica la probabilidad $p(s)$ con la que cada estrategia pura $s \in S$ es usada.*

Se supondrá que el conjunto de estrategias es $S = \{s_1, s_2, \ldots\}$, donde de nueva cuenta se tiene que S es un conjunto numerable; entonces una estrategia mixta puede ser representada como un vector de probabilidades:

$$\sigma = (p(s_1), p(s_2), \ldots).$$

Una estrategia pura puede ser vista como un vector en el que todas las entradas son cero excepto una. Por ejemplo,

$$s_3 = (0, 0, 1, \ldots).$$

En este caso s_3 representa que se uso la estrategia 3 con probabilidad 1.

Las estrategias mixtas pueden ser representadas como una combinación lineal de estrategias puras:

$$\sigma = \sum_{s \in S} p(s)\, s.$$

Definición 2.1.7 *El soporte de una estrategia mixta σ es el conjunto $S(\sigma) \subseteq S$, formada por todas las estrategias puras para las cuales σ satisface que $p(s) > 0$.*

Lema 2.1.8 *Sea σ^* una estrategia mixta óptima con soporte S^*. Entonces $u(s) = u(\sigma^*)$, $s \in S^*$.*

Demostración. Si el conjunto de S^* contiene sólo una estrategia, entonces el lema es cierto. Suponga que el conjunto S^* contiene más de una estrategia. Si el lema no es verdadero, entonces existe al menos una estrategia que da un pago mayor que $u(\sigma^*)$; sea s' la estrategia que produce dicho pago. Entonces

$$\begin{aligned} u(\sigma^*) &= \sum_{s \in S^*} p^*(s)\, u(s) \\ &= \sum_{s \neq s'} p^*(s)\, u(s) + p^*(s')\, u(s') \\ &< \sum_{s \neq s'} p^*(s)\, u(s') + p^*(s')\, u(s') \\ &= u(s'). \end{aligned}$$

■

La historia parcial h es la sucesión de decisiones que han sido realizadas por un individuo en un tiempo t determinado. Al principio de un proceso de decisión (cuando ninguna decisión ha sido hecha) se tiene historia nula, $\hat{h}$. Una historia completa para una estrategia s es la sucesión de todas las decisiones que un individuo puede tomar siguiendo la estrategia s y se denotará por $H(s)$.

Observación 2.1.9 *Un individuo tiene memoria perfecta si recuerda todas sus decisiones pasadas, entonces cada punto de observación del proceso tiene historia única y cada historia da como resultado un único punto de observación.*

Sea $S(h) \in S$ el subconjunto de estrategias con historia h pero que difieren en las acciones que realizaran en el futuro. Entonces el pago óptimo que puede obtener un individuo, dada la historia h es:

$$u^*(s\,|h) = \max_{s \in S(h)} u(s).$$

Sea $A(h)$ el conjunto de acciones donde el individuo elige la acción que llevará a cabo. Después que la decisión ha sido hecha, a la historia h, del proceso se le añade la acción a. Esto será denotado por h, a.

Teorema 2.1.10 *(Principio de Optimalidad). Para cualquier individuo con memoria perfecta se tiene:*

1. $u^*(s\,|H(s)) = u(s)$,
2. $u^*(s\,|h) = \max_{a\in A(h)} u^*(s\,|h,a)$,
3. $u^* = \max_{s\in S(\hat{h})} u^*\left(s\,\middle|\hat{h}\right)$.

Demostración.

1. De acuerdo a la definición de $H(s)$, el individuo no tiene más acciones a elegir y el mejor pago que puede obtener es el que ya tiene usando la estrategia s.

2. Una estrategia pura es una sucesión de acciones $\{a_0, a_1, \ldots, a_h, a_{h+1}, \ldots, a_H\}$. Así,

$$u(s) = u(a_0, a_1, \ldots, a_h, a_{h+1}, \ldots, a_H).$$

   Sea h la historia parcial que da como resultado la sucesión $\{a_0, a_1, \ldots, a_h\}$ entonces

$$\begin{aligned} u^*(s\,|h) &= \max_{a_{h+1}} \max_{a_{h+2}} \cdots \max_{a_H} (a_0, a_1, \ldots, a_h, a_{h+1}, \ldots, a_H) \\ &= \max_{a_{h+1}} u^*(s\,|h, a_{h+1}). \end{aligned}$$

3. La historia $\hat{h}$ denota un problema de optimización empezando desde un tiempo $t = 0$. Entonces $S\left(\hat{h}\right) = S$ y

$$\begin{aligned} \max_{s\in S(\hat{h})} u^*\left(s\,\middle|\hat{h}\right) &= \max_{s\in S} u(s) \\ &= u^*. \end{aligned}$$

■

Definición 2.1.11 *Un proceso de decisión tiene la propiedad de Markov si satisface* $p(x_{t+1}\,|h_t, a_t) = p(x_{t+1}\,|x_t, a_t)$.

2.2. Procesos de Decisión de Markov

El objetivo de esta sección es introducir el concepto de procesos de decisión de Markov (PDM).

Los procesos de decisión de Markov proporcionan una herramienta para la toma de decisiones en las que los resultados dependen de una cierta probabilidad. Los PDM son muy útiles para el estudio de problemas de optimización resueltos a través de programación dinámica. En la tesis se trabajará con PDM a tiempo discreto y con espacio de estados finito (para casos más generales ver [7] y [10]).

Definición 2.2.1 *Un Modelo de Decisión de Markov (MDM), estacionario, a tiempo discreto, consiste de una quíntupla:*

$$(X, A, \{A(x)\,|x \in X\}\,, Q, c)\,,$$

donde:

a. X *es un conjunto numerable no vacío, llamado el espacio de* estados*;*

b. A *es un conjunto numerable no vacío, llamado el espacio de* acciones*;*

c. $\{A(x)\,|x \in X\}$ *es una familia de subconjuntos medibles, no vacíos* $A\,(x)$ *de* A*, donde* $A\,(x)$ *denota el conjunto de acciones admisibles cuando el sistema se encuentra en el estado* $x \in X$*. El conjunto* $\mathbb{K}$ *de parejas de estados acciones admisibles, está definido por*

$$\mathbb{K} = \{(x, a)\,|x \in X, a \in A\,(x)\}\,,$$

y se supondrá que es un conjunto medible del espacio producto $X \times A$*;*

d. Q *es una probabilidad de transición definida en* X *dado* $\mathbb{K}$*, llamada ley de transición, es decir, para cada* $(x, a) \in \mathbb{K}$*,* $Q\,(\cdot\,|x, a)$ *es una medida de probabilidad en* X*, y para cada* $B \subset X$*, medible,* $Q\,(B\,|\cdot)$ *es una función medible;*

e. $c : \mathbb{K} \to \mathbb{R}$ *es una función medible y se llama la función de costo de un paso.*

Se puede pensar en un MDM estacionario a tiempo discreto con horizonte finito T, donde T es el horizonte del problema de optimización, como un sistema estocástico controlado que se observa de manera periódica en los tiempos $t = 0, 1, 2, \ldots, T$. La dinámica que describe este sistema estocástico funciona de la forma siguiente: si el sistema al tiempo t se encuentra en el estado $x_t = x \in X$ y $a_t = a \in A(x)$ es aplicada, entonces ocurren dos cosas:

a) se paga un costo $c(x, a)$; y

b) el sistema se traslada a un nuevo estado x_{t+1}, mediante la ley de transición (medida de probabilidad) $Q(\cdot\,|x, a)$ sobre X, es decir,

$$Q(B\,|x, a) = \Pr(x_{t+1} \in B\,|x_t = x, a_t = a),$$

$B \in \mathcal{B}(X)$, donde $\mathcal{B}(X)$ denota la σ-álgebra de Borel de X.

Una vez hecha esta transición a un nuevo estado, se elige una nueva acción y la dinámica anteriormente descrita se repite.

Suposición 2.2.2 *Se supondrá que $\mathbb{K}$ contiene la gráfica de una función medible de X a A, es decir, existe $f : X \to A$ medible, tal que $f(x) \in A(x)$, para toda $x \in X$. El conjunto de estas funciones es denotado por $\mathbb{F}$ y sus elementos son llamados acciones de la multifunción $x \to A(x)$.*

2.3. Políticas

Para introducir el concepto de estrategia o política, considérese un MDM y defina $\mathbb{H}_t$, el espacio de las historias observadas del proceso de control hasta el tiempo t, como

$$\begin{aligned} \mathbb{H}_0 &= X, \\ \mathbb{H}_t &= \mathbb{K} \times \mathbb{H}_{t-1}, \end{aligned}$$

para $t = 1, 2, \ldots, T$. Un elemento h_t de $\mathbb{H}_t$ llamado t-historia es un vector de la forma

$$(x_0, a_0, x_1, a_1, \ldots, a_{t-1}, x_t),$$

donde $(x_i, a_i) \in \mathbb{K}$ para $i = 0, \ldots, t-1$ y $x_t \in X$.

Obsérvese que, para cada t, $\mathbb{H}_t$ es un subespacio de $\mathbf{H}_t := (X \times A)^t \times X$ y $\mathbf{H}_0 := X$.

Definición 2.3.1 *Una política aleatorizada o simplemente política, es una sucesión* $\pi = \{\pi_t, t = 0, 1, 2, \ldots, T-1\}$ *de probabilidades de transición definidas sobre* A*, dada la historia del proceso* $\mathbb{H}_t$ *y satisface que:* $\pi_t(A(x_t)|h_t) = 1$ *para toda* $h_t \in \mathbb{H}_t$ *y* $t = 0, 1, 2 \ldots, T$*.*

El conjunto de todas las políticas será denotado por Π.

De acuerdo con esta definición, una política $\pi = \{\pi_t\}$ puede interpretarse como una sucesión $\{a_t\}$ de variables aleatorias sobre A, tales que, para cada t-historia y $t = 0, 1, 2, \ldots, T$, la distribución de a_t es $\pi_t(\cdot|h_t)$, la cual está concentrada en el conjunto de acciones admisibles $A(x_t)$. En otras palabras, cuando se usa una política arbitraria, la acción en cualquier tiempo t es una variable aleatoria y depende de todas las t-historias.

Se denotará a la familia de probabilidades condicionales sobre A dado X, como $P(A|X)$.

Sea Φ el conjunto de todas las probabilidades condicionales φ en $P(A|X)$ tales que para toda $x \in X$ se tiene $\varphi(A(x)|x) = 1$.

Observación 2.3.2 *Por la suposición 2.2.2 se tiene que* $\mathbb{F} \subset \Phi$*.*

Definición 2.3.3 *Una política* $\pi \in \Pi$ *es:*

Markoviana Aleatorizada (Π_{RM})*. Si existe una sucesión* $\{\varphi_t\} \subset \Phi$ *(definidas sobre* A *dado* X*), tales que,* $\pi_t(\cdot|h_t) = \varphi_t(\cdot|x_t)$ *para toda* $h_t \in \mathbb{H}_t$ *y* $t = 0, 1, 2, \ldots, T$*.*

Markoviana Aleatorizada Estacionaria (Π_{RS})*. Si existe* $\varphi \in \Phi$*, tal que:* $\pi_t(\cdot|h_t) = \varphi(\cdot|x_t)$ *para toda* $h_t \in \mathbb{H}_t$ *y* $t = 0, 1, 2, \ldots, T$*.*

Determinista (Π_D)*. Si existe una sucesión* $\{g_t\}$ *de funciones medibles* $g_t : \mathbb{H}_t \to A$*, tales que, para cada* $h_t \in \mathbb{H}_t$ *y* $t = 0, 1, 2, \ldots, T$*, se tiene que* $g_t(h_t) \in A(x_t)$ *y* $\pi_t(\cdot|h_t)$ *está concentrada en* $g_t(h_t)$*.*

Determinista Markoviana (Π_{DM})*. Si existe una sucesión* $\{f_t\}$ *de funciones medibles* $f_t : X \to A$ *(o* $f_t \in \mathbb{F}$*), tales que,* $f_t(x_t) \in A(x_t)$ *y* $\pi_t(\cdot|h_t)$ *está concentrada en* $f_t(x_t)$ *para cada* $h_t \in \mathbb{H}_t$ *y* $t = 0, 1, 2, \ldots, T$*.*

Determinista Markoviana Estacionaria (Π_{DS})*. Si existe una función medible* $f : X \to A$ *(o* $f \in \mathbb{F}$*), tal que,* $f(x_t) \in A(x_t)$ *y* $\pi_t(\cdot|h_t)$ *está concentrada en* $f(x_t)$ *para cada* $h_t \in \mathbb{H}_t$ *y* $t = 0, 1, 2, \ldots, T$*.*

Se dice que $\pi(\cdot|h)$ está concentrada en $g(h)$, si $\pi(C|h) = I_C(g(h))$ para cada $C \in \mathcal{B}(A)$. Donde I_C es la función indicadora del conjunto C.

Observación 2.3.4 *Obsérvese que,* $\Pi_{RS} \subset \Pi_{RM} \subset \Pi$ *y* $\Pi_{DS} \subset \Pi_{DM} \subset \Pi_D \subset \Pi$.

Sea $(\Omega, \mathcal{F})$ un espacio medible que consiste del espacio muestral canónico $\Omega := \bar{H}_\infty = (X \times A)^\infty$ y $\mathcal{F}$ su correspondiente σ-álgebra producto. Los elementos de Ω son de la forma $\omega = (x_0, a_0, x_1, a_1, \ldots)$ con $x_t \in X$ y $a_t \in A$ para toda $t = 0, 1, 2, \ldots, T$, las proyecciones, x_t y a_t, de Ω sobre X y A son llamados estado y acción respectivamente.

Obsérvese que $H_\infty = \mathbb{K}^\infty \subset \Omega$ es el conjunto de parejas estado acción admisible. Sean $\pi \in \Pi$ una política arbitraria y $x_0 = x \in X$. Entonces por el Teorema de Ionescu-Tulcea (ver Apéndice A), existe una única medida de probabilidad P_x^π sobre $(\Omega, \mathcal{F})$. Además, para cada $C \in \mathcal{B}(A)$, $B \in \mathcal{B}(X)$, $h_t \in \mathbb{H}_t$ y $t = 0, 1, 2, \ldots, T$, se tiene que

$$P_x^\pi \left(a_t \in C \,|h_t \right) = \pi_t \left(C \,|h_t \right), \tag{2.1}$$

$$P_x^\pi \left(x_{t+1} \in B \,|h_t, a_t \right) = Q \left(B \,|x_t, a_t \right). \tag{2.2}$$

El proceso estocástico $\left(\left(\Omega, \mathcal{F}, P_x^\pi \right), \{x_t\} \right)$ es llamado un *Proceso de Control de Markov a tiempo discreto* o *Proceso de Decisión de Markov* (PDM).

La esperanza con respecto a P_x^π será denotada por E_x^π.

Criterio de Rendimiento. Cada PDM estará dotado de una función real, llamada criterio de rendimiento, que medirá de alguna manera la calidad de cada política, a través de la sucesión de costos que genera.

Considérese un modelo de decsión de Markov fijo y un conjunto de políticas Π. Se define para cada $x \in X$ y $\pi \in \Pi$

$$V(\pi, x) = E_x^\pi \left[\sum_{t=0}^{T-1} c\left(x_t, a_t\right) + c_T\left(x_T\right) \right].$$

$V(\pi, x)$ es conocido como el **Costo Total Acumulado** y $c_T(x_T)$ es el costo terminal. Al entero positivo T se le conoce como horizonte del problema, el cual representa el número de etapas en el cual el sistema está operando y puede ser finito o infinito.

Definición 2.3.5 *Para cada* $x \in X$ *se define*

$$V^*(x) = \inf_{\pi \in \Pi} V(\pi, x),$$

V^* *se le llama funciones de valores óptimos o valor óptimo.*

Definición 2.3.6 *Una política $\pi^* \in \Pi$, es óptima, si*

$$V(\pi^*, x) = \inf_{\pi \in \Pi} V(\pi, x),$$

$x \in X$.

El problema de control óptimo consiste en encontrar una política óptima, es decir, minimizar la función $\pi \to V(\pi, x)$ sobre Π, para toda $x \in X$.

2.4. Programación Dinámica

Considere como criterio de rendimiento el costo total acumulado con horizonte finito, es decir,

$$V(\pi, x) := E_x^\pi \left[\sum_{t=0}^{T-1} c(x_t, a_t) + c_T(x_T) \right],$$

donde c_T es una función real valuada definida sobre X llamada la función de costo terminal.

El objetivo de esta sección es proveer una técnica de solución para el problema de control óptimo, conocida como Programación Dinámica (ver [4]). Dicha técnica permite encontrar tanto a la función de valor óptimo V^*, como a la política óptima π^*. El principio de Bellman está basado en el principio de Optimalidad. La prueba del siguiente teorema se presenta en el caso continuo en [14].

Teorema 2.4.1 *Sean $V_0, V_1, \ldots, V_T$ funciones sobre X definidas por*

$$V_T(x) := c_T(x), \tag{2.3}$$

y para cada $t = 0, 1, \ldots, T-1$.

$$V_t(x) := \min_{a \in A(x)} \left[c(x, a) + \sum_{x' \in X} Q(x' | x, a) V_{t+1}(x') \right]. \tag{2.4}$$

Se supone que estas funciones son medibles y que para cada $t = 0, 1, 2, \ldots, T-1$, existe una acción $f_t \in \mathbb{F}$ con $f_t(x) \in A(x)$, tal que

$$V_t(x) = c(x, f_t(x)) + \sum_{x' \in X} Q(x' | x, f_t(x)) V_{t+1}(x').$$

Entonces, la política determinista de Markov $\pi^* = \{f_0, f_1, \ldots, f_{T-1}\}$ *es óptima y la función de valor óptimo* V^* *es* V_0*, es decir, para* $x \in X$ *se tiene que* $V^*(x) = V(\pi^*, x) = V_0(x)$*.*

La relación (2.4) es conocida como Ecuación de Programación Dinámica (EPD) junto con su condición inicial (2.3) y su nombre es debido a Bellman (ver [4]).

Demostración. Sea $\pi = \{\pi_t\}$ una política arbitraria, y sea

$$\begin{aligned} C_t(\pi, x) &:= E^\pi \left[\sum_{n=t}^{T-1} c(x_n, a_n) + c_T(x_T) \,|\, x_t = x \right], \\ C_T(\pi, x) &:= c_T(x), \end{aligned}$$

$t = 0, 1, 2, \ldots, T-1$. $C_t(\pi, x)$ es llamado el costo total del instante t a $T-1$ cuando se usa la política π y $x_t = x$. En particular note que

$$V(\pi, x) = C_0(\pi, x).$$

Para demostrar este teorema, se debe mostrar que para todo $x \in X$ y $t = 0, 1, 2, \ldots, T-1$, se tiene que

$$C_t(\pi, x) \geq V_t(x),$$

con igualdad cuando $\pi = \pi^*$, es decir,

$$C_t(\pi^*, x) = V_t(x).$$

En particular, si $t = 0$, se tiene que

$$C_0(\pi, x) = V(\pi, x) \geq V_0(x)$$

y si se utiliza a π^*:

$$C_0(\pi^*, x) = V(\pi^*, x) = V^*(x) = V_0(x).$$

Para probar las relaciones anteriores, primero obsérvese que

$$C_T(\pi, x) = V_T(x) = c_T(x).$$

Ahora, suponga que para alguna $t \in \{0, 1, \ldots, T-1\}$ se tiene que

$$C_{t+1}(\pi, x) = V_{t+1}(x).$$

Entonces por (2.1) y (2.2),

$$\begin{aligned}
C_t(\pi, x) &= E^{\pi}\left[\sum_{n=t}^{T-1} c(x_n, a_n) + c_T(x_T) \,|x_t = x\right], \\
&= E^{\pi}\left[c(x_t, a_t) + \sum_{n=t+1}^{T-1} c(x_n, a_n) + c_T(x_T) \,|x_t = x\right], \\
&= E^{\pi}\left[c(x_t, a_t) \,|x_t = x\right] + E^{\pi}\left[\sum_{n=t+1}^{T-1} c(x_n, a_n) + c_T(x_T) \,|x_t = x\right], \\
&= \sum_{a\in A} c(x, a)\, \pi(a\,|x) + E^{\pi}\left[\sum_{n=t+1}^{T-1} c(x_n, a_n) + c_T(x_T) \,|x_t = x\right],
\end{aligned}$$

y por propiedades de la esperanza condicional, se llega a que

$$\begin{aligned}
& E^{\pi}\left[\sum_{n=t+1}^{T-1} c(x_n, a_n) + c_T(x_T) \,|x_t = x\right] \\
= \;& E^{\pi}\left[E^{\pi}\left(\sum_{n=t+1}^{T-1} c(x_n, a_n) + c_T(x_T) \,|x_{t+1} = y\right) |x_t = x\right].
\end{aligned}$$

Por otro lado,

$$E^{\pi}\left[\sum_{n=t+1}^{T-1} c(x_n, a_n) + c_T(x_T) \,|x_{t+1} = y\right] = C_{t+1}(\pi, x).$$

Así,

$$\begin{aligned}
C_t(\pi, x) &= \sum_{a\in A}\left[c(x, a) + \sum_{y\in X} C_{t+1}(\pi, y)\, Q(y\,|x, a)\right] \pi_t(a\,|x), \\
&\geq \sum_{a\in A}\left[c(x, a) + \sum_{y\in X} V_{t+1}(y)\, Q(y\,|x, a)\right] \pi_t(a\,|x), \\
&\geq \min_{a\in A(x)}\left[c(x, a) + \sum_{y\in X} V_{t+1}(y)\, Q(y\,|x, a)\right] \\
&= V_t(x),
\end{aligned}$$

y bajo la hipótesis de inducción se tiene que

$$C_{t+1}\left(\pi^*, x\right) = V_{t+1}\left(x\right).$$

■

Se denotará por $P_{ij}\left(a,t\right)$ a la probabilidad de estar en el estado j, dado que nos encontramos en el estado i y se eligió la acción a, en el tiempo t, es decir,

$$P_{ij}\left(a,t\right) = P\left[x_{t+1} = j \left| x_t = i, a_t = a\right.\right],$$

2.5. Ejemplo

Ejemplo 2.5.1 *Considere un problema de operación de una máquina en T periodos. El espacio de estados es finito, es decir, $X = \{1, 2, \ldots, n\}$. Suponga que encontrarse en el estado i es mejor que estar en el estado $i+1$, y el estado 1 es el de condiciones perfectas en las que la máquina puede operar.*

Sea $g\left(i\right)$ el costo de operación por periodo cuando se está en el estado i y se usa la acción a, suponga que

$$g\left(1\right) \leq g\left(2\right) \leq \cdots \leq g\left(n\right).$$

Durante un periodo de operación, el estado de la máquina puede empeorar o quedarse igual.

Se considera la probabilidad de transición como:

$$p_{ij} = \begin{cases} p\left(j \left| i\right.\right), & j \geq i, \\ 0, & j < i. \end{cases}$$

Supóngase también que al comenzar cada periodo conocemos el estado de la máquina y tenemos las opciones siguientes:

1. *Que la máquina opere un periodo más, en el estado en que se encuentra.*

2. *Reparar la máquina hasta estar en perfectas condiciones, es decir, hasta que se encuentre en el estado 1, pagando un costo R.*

Con esto se tiene que el espacio de acciones A es también finito y tiene dos elementos (repararla y dejarla trabajar un periodo más).

Suponga que la máquina, una vez reparada, permanecerá en el estado 1 por al menos un periodo. En los periodos siguientes, puede deteriorarse para los estados $j > 1$ según las probabilidades de transición p_{ij}.

Así el objetivo es decidir sobre el nivel de deterioro (estado) en el cual vale la pena pagar el costo de reparación de la máquina, de tal modo que se obtengan costos de operación más bajos en el futuro. Obsérvese que la decisión también está afectada por el periodo en el que se encuentra la máquina, por ejemplo, se estaría menos inclinado por reparar la máquina cuando hay pocos periodos a la izquierda. Así, considerando el criterio de costo total acumulado,

$$V(\pi, x) := E_x^\pi \left[\sum_{t=0}^{T-1} g_t(x_t, a_t) + g_T(x_T) \right].$$

Por lo tanto, en este caso se tiene que la EPD

$$V_T(x) = g_T(x_T) = 0,$$

$$V_t(i) = \min_{a \in A(i)} \left[g(i) + \sum_{j=i}^{n} P_{ij}(a) V_{t+1}(j) \right].$$

Así, en general

$$\begin{aligned} V_k(i) &= \min \left\{ g(i) + \sum_{j=i}^{n} p_{ij} V_{k+1}(j), R + g(1) + V_{k+1}(1) \right\}, \\ k &= \{0, 1, 2, \ldots, T-1\}, \end{aligned}$$

esto significa dejar un periodo más trabajando o reparar la máquina.

En un caso particular, suponga que $X = \{1, 2, 3, 4, 5\}$ y la máquina es revisada durante dos periodos. La matriz de transición esta dada por

$$\begin{bmatrix} \frac{2}{10} & \frac{1}{10} & \frac{3}{10} & \frac{2}{10} & \frac{2}{10} \\ 0 & \frac{5}{10} & \frac{3}{10} & \frac{1}{10} & \frac{1}{10} \\ 0 & 0 & \frac{7}{10} & \frac{2}{10} & \frac{1}{10} \\ 0 & 0 & 0 & \frac{4}{10} & \frac{6}{10} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

el costo esta definido como $g(i) = i + 1$*, para* $i = 1, 2, 3, 4, 5$*. Y el costo de reparación* $R = \frac{14}{5}$*.*

Utilizando programación dinámica

$$V_2(x) = 0.$$

En la etapa siguiente

$$\begin{aligned} V_1(i) &= \min\left\{ g(i) + \sum_{j=i}^{5} p_{ij} V_2(j), R + g(1) + V_2(1) \right\} \\ &= \min\left\{ i + 1, \frac{24}{5} \right\}. \end{aligned}$$

De esta forma,

$$V_1(i) = \begin{cases} i + 1, \text{ si } i = 1, 2, 3, \\ \frac{24}{5}, \text{ si } i = 4, 5; \end{cases}$$

es decir, para los tres primeros estados es mejor dejar trabajar un periodo más la máquina mientras que para los restantes es mejor mandar a reparar la máquina.

Para la siguiente etapa

$$\begin{aligned} V_0(i) &= \min\left\{ g(i) + \sum_{j=i}^{5} p_{ij} V_1(j), R + g(1) + V_1(1) \right\} \\ &= \min\left\{ i + 1 + \sum_{j=i}^{5} p_{ij} V_1(j), \frac{34}{5} \right\}. \end{aligned}$$

Ahora, resolviendo

$$\begin{aligned} V_0(1) &= \min\left\{ 1 + 1 + \sum_{j=1}^{5} p_{1j} V_1(j), \frac{34}{5} \right\} \\ &= \min\left\{ \frac{291}{50}, \frac{34}{5} \right\} \\ &= \frac{291}{50}, \end{aligned}$$

$$
\begin{aligned}
V_0\left(2\right) &= \min\left\{2+1+\sum_{j=2}^{5} p_{2j}V_1\left(j\right), \frac{34}{5}\right\} \\
&= \min\left\{\frac{333}{50}, \frac{34}{5}\right\} \\
&= \frac{333}{50},
\end{aligned}
$$

$$
\begin{aligned}
V_0\left(3\right) &= \min\left\{3+1+\sum_{j=3}^{5} p_{3j}V_1\left(j\right), \frac{34}{5}\right\} \\
&= \min\left\{\frac{206}{25}, \frac{34}{5}\right\} \\
&= \frac{34}{5},
\end{aligned}
$$

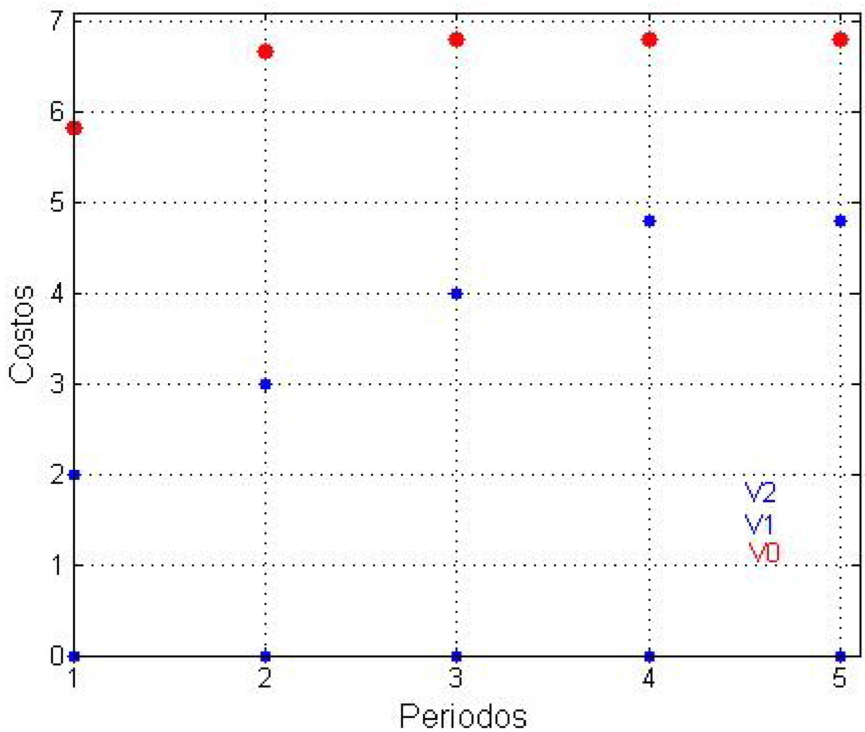
$$
\begin{aligned}
V_0\left(4\right) &= \min\left\{4+1+\sum_{j=4}^{5} p_{4j}V_1\left(j\right), \frac{34}{5}\right\} \\
&= \min\left\{\frac{49}{5}, \frac{34}{5}\right\} \\
&= \frac{34}{5},
\end{aligned}
$$

$$
\begin{aligned}
V_0\left(5\right) &= \min\left\{5+1+\sum_{j=5}^{5} p_{5j}V_1\left(j\right), \frac{34}{5}\right\} \\
&= \min\left\{\frac{54}{5}, \frac{34}{5}\right\} \\
&= \frac{34}{5}.
\end{aligned}
$$

Así,

$$V_0(i) = \begin{cases} \frac{291}{50},\ si\ i = 1, \\ \frac{333}{50},\ si\ i = 2, \\ \frac{34}{5},\ si\ i = 3, 4, 5. \end{cases}$$

Se concluye que en los dos primeros estados es mejor dejarla trabajar y en los siguientes reparar la máquina. La siguiente gráfica representa los costos de operación en cada periodo



Para el caso en el que se consideran un número mayor de estados y de periodos, se realizó un programa en MATLAB, cuyo algoritmo consiste en

1. *Asignar valores al costo de reparación* (R), *al costo de operación por periodo cuando se está en el estado* i, $(g(i))$, *al número de estados* (n) *y el número de periodos* (T),

2. *simular la matriz de transición p_{ij} y, generar la matriz de transición triangular inferior,*

3. *y finalmente, se implementa la ecuación de Bellman*

$$V_k(i) = \min\left\{ g(i) + \sum_{j=i}^{n} p_{ij} V_{k+1}(j), R + g(1) + V_{k+1}(1) \right\}.$$

Usando el algoritmo realizado en MATLAB, considere $n = 8$, $T = 10$, $R = 15$ y $g(i) = i + 1$. Se tiene que los costos son

	1	2	3	4	5	6	7	8
V_9	2	3	4	5	6	7	8	9
V_8	7.891	8.925	10.064	12.374	13.829	15.197	16.665	18
V_7	15.469	16.446	17.660	20.759	22.395	23.898	24.891	24.891
V_6	23.491	24.422	25.720	29.188	30.575	31.653	32.469	32.469
V_5	31.486	32.386	33.690	36.899	38.131	39.273	40.469	40.491
V_4	39.305	40.204	41.537	44.719	46.024	47.191	48.483	48.486
V_3	47.194	48.093	49.437	52.682	54.004	55.175	56.305	56.305
V_2	55.114	56.017	57.347	60.582	61.873	63.034	64.194	64.194
V_1	63.006	63.906	65.239	68.458	69.754	70.916	72.115	72.115
V_0	70.897	71.797	73.133	76.360	77.663	78.827	80.006	80.006

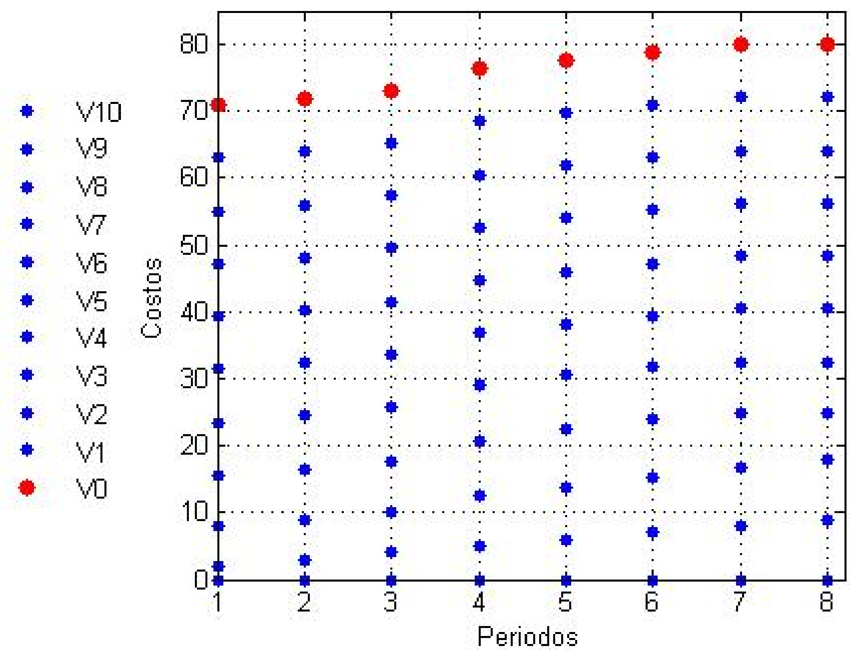
donde las filas representan los costos para cada periodo y las columnas los estados en que son revisados las máquinas; y las políticas a seguir son

	1	2	3	4	5	6	7	8
9	2	2	2	2	2	2	2	2
8	2	2	2	2	2	2	2	2
7	2	2	2	2	2	2	1	1
6	2	2	2	2	2	2	1	1
5	2	2	2	2	2	2	2	1
4	2	2	2	2	2	2	2	1
3	2	2	2	2	2	2	1	1
2	2	2	2	2	2	2	1	1
1	2	2	2	2	2	2	1	1
0	2	2	2	2	2	2	1	1

donde de nueva cuenta las filas representan los periodos y las columnas los estados, y además

$$\begin{array}{rcl} 2 & \equiv & \textit{Que la máquina trabajé un periodo más,} \\ 1 & \equiv & \textit{Reparar la máquina.} \end{array}$$

La gráfica de los costos de operación es



De igual forma se puede trabajar con un número mucho mayor de estados y periodos, por ejemplo $n = 40$, $T = 45$, $R = 15$ *y* $g(i) = i + 1$, *pero por falta de espacio sólo se incluye la función de valor óptimo,* V_0

$$V_0(i) = \begin{cases} 399{,}81211 \ si \ i = 1; \\ 401{,}48413 \ si \ i = 2; \\ 402{,}42412 \ si \ i = 3; \\ 403{,}69999 \ si \ i = 4; \\ 404{,}62182 \ si \ i = 5; \\ 405{,}68700 \ si \ i = 6; \\ 406{,}79460 \ si \ i = 7; \\ 407{,}81717 \ si \ i = 8, \ldots, 40; \end{cases}$$

y la política óptima a seguir es, permitir que la máquina trabaje para los primeros siete estados y a partir del estado 8 hasta el 40 pagar el costo de reparación.

Ejemplo 2.5.2 *Ahora considérese un problema de operación para dos máquinas (llamémosle máquina* 1 *y* 2 *respectivamente) en* T *periodos, cuyo funcionamiento es independiente. De nueva cuenta el espacio de estados es finito, es decir,* $X_1 = X_2 = \{1, 2, \ldots, n\}$*. Suponga que encontrarse en el estado* i *es mejor que estar en el estado* $i+1$*, para la máquina* 1*; y el estado* j *es mejor que el* $j+1$ *para la máquina* 2*, y el estado* 1 *es el de condiciones perfectas en las que ambas máquinas pueden operar.*

Sean $g_1(i)$ *y* $g_2(j)$ *el costo de operación para la máquina* 1 *y* 2 *respectivamente, suponga que*

$$\begin{aligned} g_1(1) &\leq \cdots g_1(n), \\ g_2(1) &\leq \cdots g_2(n). \end{aligned}$$

Durante un periodo de operación, el estado de las máquinas puede empeorar o quedarse igual.

Considere la probabilidad de transición como:

$$Q((i,j),(x,y)) = Q_1(i,x)\, Q_2(j,y),$$

(ver [8] p. 75). Donde

$$\begin{aligned} Q_1(i,x) &= \begin{cases} Q_1(x\,|i), & x \geq i, \\ 0, & x < i. \end{cases} \\ Q_2(j,y) &= \begin{cases} Q_2(y\,|j), & y \geq j, \\ 0, & y < j. \end{cases} \end{aligned}$$

Supóngase también que al comenzar cada periodo se conoce el estado de las máquinas. Entonces se tienen las opciones siguientes:

1. *Dejar trabajar ambas máquinas, por un periodo más pagando los costos de operación* $g_1(i)$*,* $g_2(j)$*;*

2. *Reparar la máquina* 1 *pagando un costo* R_1*, y dejar que la máquina* 2 *opere un periodo más, pagando el costo* $g_2(j)$*;*

3. *Dejar que la máquina 1 opere un periodo más, pagando $g_1(i)$ y reparar la máquina 2 hasta estar en perfectas condiciones, es decir, hasta que se encuentre en el estado 1, pagando un costo R_2.*

4. *Reparar ambas máquinas, hasta que estén en perfectas condiciones, pagando un costo $R_1 + R_2$.*

Con esto se tiene que el espacio de acciones A es también finito y tiene cuatro elementos.

Suponga que las máquinas, una vez reparadas, permanecerán en el estado 1 por al menos un periodo. En los periodos siguientes, puede deteriorarse para los estados $t > 1$ según la probabilidad de transición $Q((i,j),(x,y))$.

Así el objetivo es decidir sobre el nivel de deterioro (estado) en el cual vale la pena pagar el costo de reparación de las máquinas, de tal modo que obtengamos costos de operación más bajos en el futuro. Se tiene que la EPD está dada por

$$V_T(x,y) = 0,$$

$$\begin{aligned} V_k(i,j) &= \min \left\{ \begin{array}{c} g_1(i) + g_2(j) + \sum_{x=i}^{n} \sum_{y=j}^{n} Q_1(i,x) Q_2(j,y) V_{k+1}(x,y), \\ R_1 + g_1(1) + g_2(j) + \sum_{y=j}^{n} Q_1(i,1) Q_2(j,y) V_{k+1}(1,y), \\ R_2 + g_1(i) + g_2(1) + \sum_{x=i}^{n} Q_1(i,x) Q_2(j,1) V_{k+1}(x,1), \\ R_1 + R_2 + g_1(1) + g_2(1) + V_{k+1}(1,1) \end{array} \right\} \\ k &= \{0,1,2,\ldots,T-1\}. \end{aligned}$$

En un caso particular, suponga que $X = \{1,2\}$ y las máquinas son revisadas durante dos periodos. Las matrices de transición están dadas por

$$Q_1 = \begin{bmatrix} \frac{1}{5} & \frac{4}{5} \\ 0 & 1 \end{bmatrix}, \quad Q_2 = \begin{bmatrix} \frac{5}{7} & \frac{2}{7} \\ 0 & 1 \end{bmatrix}$$

los costos de operación están definidos por $g_1(i) = i$, $g_2(j) = j$, para $i,j = 1,2$. Y los costos de reparación $R_1 = R_2 = \frac{3}{2}$.

Utilizando la ecuación de programación dinámica

$$V_2(x,y) = 0.$$

En la etapa siguiente

$$V_1(i,j) = \min \left\{ \begin{array}{c} g_1(i) + g_2(j) + \sum_{x=i}^{2} \sum_{y=j}^{2} Q_1(i,x) Q_2(j,y) V_2(x,y), \\ R_1 + g_1(1) + g_2(j) + \sum_{y=j}^{2} Q_1(i,1) Q_2(j,y) V_2(1,y), \\ R_2 + g_1(i) + g_2(1) + \sum_{x=i}^{2} Q_1(i,x) Q_2(j,1) V_2(x,1), \\ R_1 + R_2 + g_1(1) + g_2(1) + V_2(1,1) \end{array} \right\}$$

De esta forma, se tiene

$$\begin{aligned}
V_1(1,1) &= \min\left\{\begin{array}{c} g_1(1)+g_2(1)+\sum_{x=1}^{2}\sum_{y=1}^{2}Q_1(1,x)\,Q_2(1,y)\,V_2(x,y)\,, \\ \frac{3}{2}+g_1(1)+g_2(1)+\sum_{y=1}^{2}Q_1(1,1)\,Q_2(1,y)\,V_2(1,y)\,, \\ \frac{3}{2}+g_1(1)+g_2(1)+\sum_{x=1}^{2}Q_1(1,x)\,Q_2(1,1)\,V_2(x,1)\,, \\ 3+g_1(1)+g_2(1)+V_2(1,1) \end{array}\right\} \\
&= \min\left\{2,\frac{7}{2},\frac{7}{2},5\right\} \\
&= 2,
\end{aligned}$$

$$\begin{aligned}
V_1(1,2) &= \min\left\{\begin{array}{c} g_1(1)+g_2(2)+\sum_{x=1}^{2}\sum_{y=2}^{2}Q_1(1,x)\,Q_2(2,y)\,V_2(x,y)\,, \\ \frac{3}{2}+g_1(1)+g_2(2)+\sum_{y=2}^{2}Q_1(1,1)\,Q_2(2,y)\,V_2(1,y)\,, \\ \frac{3}{2}+g_1(1)+g_2(1)+\sum_{x=1}^{2}Q_1(1,x)\,Q_2(2,1)\,V_2(x,1)\,, \\ 3+g_1(1)+g_2(1)+V_2(1,1) \end{array}\right\} \\
&= \min\left\{3,\frac{9}{2},\frac{7}{2},5\right\} \\
&= 3,
\end{aligned}$$

$$\begin{aligned}
V_1(2,1) &= \min\left\{\begin{array}{c} g_1(2)+g_2(1)+\sum_{x=2}^{2}\sum_{y=1}^{2}Q_1(2,x)\,Q_2(1,y)\,V_2(x,y)\,, \\ \frac{3}{2}+g_1(1)+g_2(1)+\sum_{y=1}^{2}Q_1(2,1)\,Q_2(1,y)\,V_2(1,y)\,, \\ \frac{3}{2}+g_1(2)+g_2(1)+\sum_{x=2}^{2}Q_1(2,x)\,Q_2(1,1)\,V_2(x,1)\,, \\ 3+g_1(1)+g_2(1)+V_2(1,1) \end{array}\right\} \\
& \min\left\{3,\frac{7}{2},\frac{9}{2},5\right\} \\
&= 3,
\end{aligned}$$

$$\begin{aligned}
V_1(2,2) &= \min\left\{\begin{array}{c} g_1(2)+g_2(2)+\sum_{x=2}^{2}\sum_{y=2}^{2}Q_1(2,x)\,Q_2(2,y)\,V_2(x,y)\,, \\ \frac{3}{2}+g_1(1)+g_2(2)+\sum_{y=2}^{2}Q_1(2,1)\,Q_2(2,y)\,V_2(1,y)\,, \\ \frac{3}{2}+g_1(2)+g_2(1)+\sum_{x=2}^{2}Q_1(2,x)\,Q_2(2,1)\,V_2(x,1)\,, \\ 3+g_1(1)+g_2(1)+V_2(1,1) \end{array}\right\} \\
&= \min\left\{4,\frac{9}{2},\frac{9}{2},5\right\} \\
&= 4,
\end{aligned}$$

Para la siguiente etapa

$$V_0(i,j) = \min \left\{ \begin{array}{c} g_1(i) + g_2(j) + \sum_{x=i}^{2} \sum_{y=j}^{2} Q_1(i,x)\, Q_2(j,y)\, V_1(x,y)\,, \\ R_1 + g_1(1) + g_2(j) + \sum_{y=j}^{2} Q_1(i,1)\, Q_2(j,y)\, V_1(1,y)\,, \\ R_2 + g_1(i) + g_2(1) + \sum_{x=i}^{2} Q_1(i,x)\, Q_2(j,1)\, V_1(x,1)\,, \\ R_1 + R_2 + g_1(1) + g_2(1) + V_1(1,1) \end{array} \right\}$$

Ahora, resolviendo

$$\begin{aligned} V_0(1,1) &= \min \left\{ \begin{array}{c} g_1(1) + g_2(1) + \sum_{x=1}^{2} \sum_{y=1}^{2} Q_1(1,x)\, Q_2(1,y)\, V_1(x,y)\,, \\ \frac{3}{2} + g_1(1) + g_2(1) + \sum_{y=1}^{2} Q_1(1,1)\, Q_2(1,y)\, V_1(1,y)\,, \\ \frac{3}{2} + g_1(1) + g_2(1) + \sum_{x=1}^{2} Q_1(1,x)\, Q_2(1,1)\, V_1(x,1)\,, \\ 3 + g_1(1) + g_2(1) + V_1(1,1) \end{array} \right\} \\ &= \min \left\{ \frac{174}{35}, \frac{277}{70}, \frac{11}{2}, 7 \right\} \\ &= \frac{277}{70}, \end{aligned}$$

$$\begin{aligned} V_0(1,2) &= \min \left\{ \begin{array}{c} g_1(1) + g_2(2) + \sum_{x=1}^{2} \sum_{y=2}^{2} Q_1(1,x)\, Q_2(2,y)\, V_1(x,y)\,, \\ \frac{3}{2} + g_1(1) + g_2(2) + \sum_{y=2}^{2} Q_1(1,1)\, Q_2(2,y)\, V_1(1,y)\,, \\ \frac{3}{2} + g_1(1) + g_2(1) + \sum_{x=1}^{2} Q_1(1,x)\, Q_2(2,1)\, V_1(x,1)\,, \\ 3 + g_1(1) + g_2(1) + V_1(1,1) \end{array} \right\} \\ &= \min \left\{ \frac{32}{5}, \frac{51}{10}, \frac{7}{2}, 7 \right\} \\ &= \frac{7}{2}, \end{aligned}$$

$$\begin{aligned} V_0(2,1) &= \min \left\{ \begin{array}{c} g_1(2) + g_2(1) + \sum_{x=2}^{2} \sum_{y=1}^{2} Q_1(2,x)\, Q_2(1,y)\, V_1(x,y)\,, \\ \frac{3}{2} + g_1(1) + g_2(1) + \sum_{y=1}^{2} Q_1(2,1)\, Q_2(1,y)\, V_1(1,y)\,, \\ \frac{3}{2} + g_1(2) + g_2(1) + \sum_{x=2}^{2} Q_1(2,x)\, Q_2(1,1)\, V_1(x,1)\,, \\ 3 + g_1(1) + g_2(1) + V_1(1,1) \end{array} \right\} \\ &= \min \left\{ \frac{43}{7}, \frac{7}{2}, \frac{93}{14}, 7 \right\} \\ &= \frac{7}{2}, \end{aligned}$$

$$
\begin{aligned}
V_0(2,2) &= \min \left\{ \begin{array}{c} g_1(2) + g_2(2) + \sum_{x=2}^{2} \sum_{y=2}^{2} Q_1(2,x) Q_2(2,y) V_1(x,y), \\ \frac{3}{2} + g_1(1) + g_2(2) + \sum_{y=2}^{2} Q_1(2,1) Q_2(2,y) V_1(1,y), \\ \frac{3}{2} + g_1(2) + g_2(1) + \sum_{x=2}^{2} Q_1(2,x) Q_2(2,1) V_1(x,1), \\ 3 + g_1(1) + g_2(1) + V_1(1,1) \end{array} \right\} \\
&= \min \left\{ \frac{15}{2}, \frac{9}{2}, \frac{9}{2}, 7 \right\} \\
&= \frac{9}{2}.
\end{aligned}
$$

Así, se tiene que la función óptima es:

$$
\begin{aligned}
V_0(1,1) &= \frac{277}{70}, \\
V_0(1,2) &= \frac{7}{2}, \\
V_0(2,1) &= \frac{7}{2}, \\
V_0(2,2) &= \frac{9}{2}.
\end{aligned}
$$

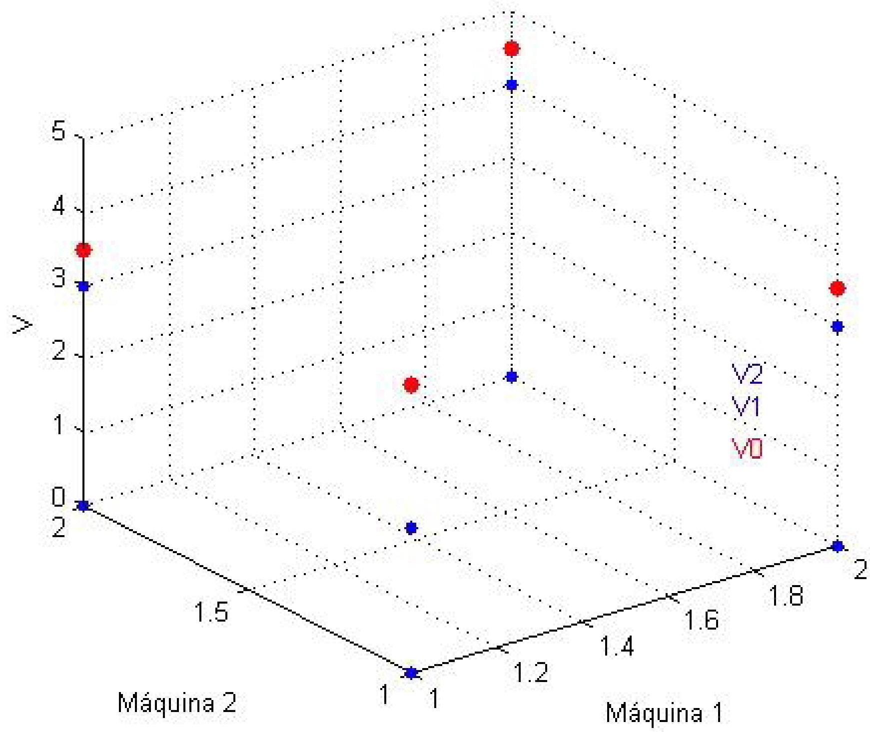
Donde la primera entrada representa los estados de la máquina 1 *y la segunda los de la máquina* 2*; se tiene que la política para cada periodo es:*

$$
\begin{aligned}
P_4(i;j) = (1,1), \;\; &para \;\; i, \; j = 1;2, \\
P_k(1,1) &= (0,1), \\
P_k(1,2) &= (1,0), \\
P_k(2,1) &= (0,1), \\
P_k(2,2) &= (0,1).
\end{aligned}
$$

Para $k = 0,1,2,3$ *y, donde, de nueva cuenta la primera entrada representa los estados de la máquina* 1 *y la segunda los de la máquina* 2 *y en el vector del lado derecho de la igualdad, la primera entrada representa las acciones para la máquina* 1 *y la segunda entrada las de la máquina* 2 *y, además*

$$
\begin{aligned}
1 &\equiv \textit{Dejar trabaja la máquina } i, \\
0 &\equiv \textit{Reparar la máquina } i.
\end{aligned}
$$

Para $i = 1,2$*. La gráfica siguiente representa los costos de operación en cada estdo y periodo*



Para el caso en el que se consideran un número mayor de estados y de periodos, nuevamente se realizó un programa en MATLAB (ver Apéndice B y [6]), cuyo algoritmo consiste en

1. *Asignar valores a los costos de reparación* (R_1 *y* R_2), *al costo de operación por periodo cuando se está en el estado* i, $g_1(i)$ *para la máquina* 1 *y* $g_2(j)$ *para la máquina* 2 *en el estado* j, *al número de estados* (n) *y al número de periodos* (T),
2. *simular las matrices de transición* $Q_1(i,x)$ *y* $Q_2(j,y)$,
3. *y finalmente se implementa la ecuación*

$$V_k(i,j) = \min \left\{ \begin{array}{c} g_1(i) + g_2(j) + \sum_{x=i}^{n} \sum_{y=j}^{n} Q_1(i,x)\, Q_2(j,y)\, V_{k+1}(x,y), \\ R_1 + g_1(1) + g_2(j) + \sum_{y=j}^{n} Q_1(i,1)\, Q_2(j,y)\, V_{k+1}(1,y), \\ R_2 + g_1(i) + g_2(1) + \sum_{x=i}^{n} Q_1(i,x)\, Q_2(j,1)\, V_{k+1}(x,1), \\ R_1 + R_2 + g_1(1) + g_2(1) + V_{k+1}(1,1) \end{array} \right\}$$

Ahora, usando el algoritmo de MATLAB, considere $n = 2$*,* $T = 5$*,* $R_1 = 25$*,* $R_2 = 39$ *,* $g_1(i) = i * 10 + 15$ *y* $g_2(j) = j * 8 + 20$*. Las matrices de transición están dadas por*

$$Q_1 = \begin{bmatrix} 0.3221 & 0.6779 \\ 0 & 1 \end{bmatrix}, \qquad Q_2 = \begin{bmatrix} 0.6983 & 0.3017 \\ 0 & 1 \end{bmatrix}$$

Se tiene que el costo de operación para ambas máquinas es

$$V_5(i,j) = 0 \quad \textit{para } i,j = 1,2$$

$$\begin{aligned} V_4(1,1) &= 53 \\ V_4(1,2) &= 61 \\ V_4(2,1) &= 63 \\ V_4(2,2) &= 71 \end{aligned}$$

$$\begin{aligned} V_3(1,1) &= 95.8496 \\ V_3(1,2) &= 92 \\ V_3(2,1) &= 78 \\ V_3(2,2) &= 86 \end{aligned}$$

$$\begin{aligned} V_2(1,1) &= 108.5004 \\ V_2(1,2) &= 92 \\ V_2(2,1) &= 78 \\ V_2(2,2) &= 86 \end{aligned}$$

$$\begin{aligned} V_1(1,1) &= 111.3460 \\ V_1(1,2) &= 92 \\ V_1(2,1) &= 78 \\ V_1(2,2) &= 86 \end{aligned}$$

y la función de valor óptimo es

$$\begin{aligned} V_0(1,1) &= 111.9860 \\ V_0(1,2) &= 92 \\ V_0(2,1) &= 78 \\ V_0(2,2) &= 86 \end{aligned}$$

y la políticas para cada periodo son:

$$P_5\left(i,j\right)=0,\ \textit{para}\ i,j=1,2,$$

$$P_4\left(i,j\right)=\left(1,1\right),\ \textit{para}\ i,j=1,2,$$

$$\begin{aligned}
P_3\left(1,1\right) &= \left(0,1\right), \\
P_3\left(1,2\right) &= \left(1,0\right), \\
P_3\left(2,1\right) &= \left(0,1\right), \\
P_3\left(2,2\right) &= \left(0,1\right),
\end{aligned}$$

$$\begin{aligned}
P_2\left(1,1\right) &= \left(0,1\right), \\
P_2\left(1,2\right) &= \left(1,0\right), \\
P_2\left(2,1\right) &= \left(0,1\right), \\
P_2\left(2,2\right) &= \left(0,1\right),
\end{aligned}$$

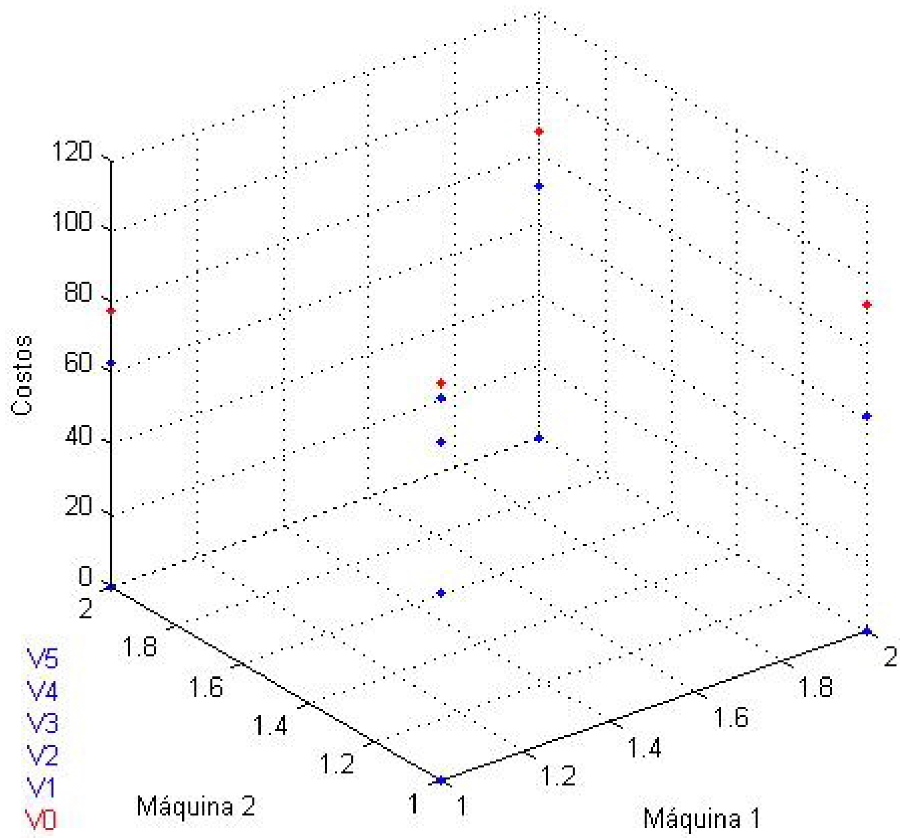
$$\begin{aligned}
P_1\left(1,1\right) &= \left(0,1\right), \\
P_1\left(1,2\right) &= \left(1,0\right), \\
P_1\left(2,1\right) &= \left(0,1\right), \\
P_1\left(2,2\right) &= \left(0,1\right).
\end{aligned}$$

$$\begin{aligned}
P_0\left(1,1\right) &= \left(0,1\right), \\
P_0\left(1,2\right) &= \left(1,0\right), \\
P_0\left(2,1\right) &= \left(0,1\right), \\
P_0\left(2,2\right) &= \left(0,1\right).
\end{aligned}$$

Además

$$\begin{aligned}
1 &\equiv \textit{Dejar trabaja la máquina } i, \\
0 &\equiv \textit{Reparar la máquina } i.
\end{aligned}$$

Para $i=1,2$*. La siguiente gráfica representa los costos de operación de las máquinas.*



En el siguiente capítulo se realizará una generalización del ejemplo de operación eficiente de dos máquinas en una sola etapa considerando el caso en que son dependientes, para ello se usara la teoría de juegos.

Capítulo 3

Teoría de Juegos

La teoría de juegos es un área de la matemática que utiliza modelos para estudiar interacciones (juegos) y realizar procesos de decisión. Un juego consiste en una serie de jugadores, un conjunto de estrategias para cada jugador, y un pago para cada jugador.

Un problema de decisión interactivo involucra dos o más individuos que llevan a cabo una decisión en una situación en la que el pago de cada uno depende (al menos en principio) de cada decisión individual. Usando la teoría de los juegos, los cuales son un ejemplo de los problemas de decisión interactivos, llamaremos a los problemas juegos y a los individuos que llevan a cabo las decisiones jugadores.

3.1. Equilibrio de Nash

En la tesis se considerará un juego para dos personas, y se supone que cada jugador tiene su propia matriz de pagos. Se consideran las matrices de pago como:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}, B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ \vdots & \vdots & \vdots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nm} \end{bmatrix}.$$

Por ejemplo, si el jugador 1 usa la fila 2 y el jugador 2 la columna 1, entonces el pago para el jugador 1 es a_{21}, y el pago del jugador 2 es b_{21}. En general, el pago cuando el jugador 1 elige la fila i y el jugador 2 la columna j, es una pareja de números (a_{ij}, b_{ij}), donde la primera componente es el pago del jugador 1 y la segunda el del jugador 2. Las filas y columnas son llamadas estrategias puras para cada jugador.

Una estrategia mixta para el jugador 1, es $X = (x_1, \ldots, x_n)$ donde $x_i \geq 0$ representa la probabilidad de que el jugador 1 use la fila i y $x_1 + \ldots + x_n = 1$. Similarmente, una estrategia mixta para el jugador 2 es $Y = (y_1, \ldots, y_m)$ donde $y_j \geq 0$ y $y_1 + \ldots + y_m = 1$. Después de que cada jugador elige su estrategia mixta, cada uno obtiene un pago dado por:

$$\begin{aligned} E_1(X,Y) &= XAY^T \text{ para el jugador 1,} \\ E_2(X,Y) &= XBY^T \text{ para el jugador 2.} \end{aligned}$$

El conjunto de las estrategias mixtas del jugador 1 será denotado por S_n y para el jugador 2 por S_m.

Definición 3.1.1 *Una pareja de estrategias mixtas* $(X^* \in S_n, Y^* \in S_m)$ *es un* equilibrio de Nash, *si* $E_1(X, Y^*) \leq E_1(X^*, Y^*)$ *para cada* $X \in S_n$ *y* $E_2(X^*, Y) \leq E_2(X^*, Y^*)$, *para cada* $Y \in S_m$. *En forma matricial,* (X^*, Y^*) *es un equilibrio de Nash si*

$$\begin{aligned} E_1(X^*,Y^*) &= X^*AY^{*T} \geq XAY^{*T} = E_1(X,Y^*)\,, \text{ para cada } X \in S_n, \\ E_2(X^*,Y^*) &= X^*BY^{*T} \geq X^*BY^T = E_2(X^*,Y)\,, \text{ para cada } Y \in S_m. \end{aligned}$$

Definición 3.1.2 *Una estrategia* $\hat{X} \in S_n$ *es la* mejor respuesta *para una estrategia del jugador* 2, $\hat{Y} \in S_m$, *si*

$$E_1\left(\hat{X}, \hat{Y}\right) = \max_{X \in S_n} E_1\left(X, \hat{Y}\right).$$

Similarmente, una estrategia $\hat{Y} \in S_m$ *es la* mejor respuesta *para una estrategia del jugador* 1, $\hat{X} \in S_n$, *si*

$$E_2\left(\hat{X}, \hat{Y}\right) = \max_{Y \in S_m} E_2\left(\hat{X}, Y\right).$$

Para el caso de las estrategias puras una fila i^* y una columna j^* pueden ser un equilibrio de Nash si satisfacen que

$$a_{ij^*} \leq a_{i^*j^*} \text{ y } b_{i^*j} \leq b_{i^*j^*}, i = 1, \ldots, n, j = 1, \ldots, m.$$

Así, $a_{i^*j^*}$ es el número más grande en la columna j^* y $b_{i^*j^*}$ es el número más grande en la fila i^*.

En adelante se considerará que $n = m = 2$, es decir, cada jugador cuenta con dos estrategias.

Sean $X = (x, 1-x)$, $Y = (y, 1-y)$, $0 \leq x \leq 1$, $0 \leq y \leq 1$ estrategias mixtas para el jugador 1 y 2, respectivamente. El pago esperado para cada jugador es:

$$
\begin{aligned}
E_1(X,Y) &= XAY^T = (x, 1-x)\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} y \\ 1-y \end{bmatrix}, \\
E_2(X,Y) &= XBY^T = (x, 1-x)\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}\begin{bmatrix} y \\ 1-y \end{bmatrix}.
\end{aligned}
$$

Proposición 3.1.3 *Una condición necesaria y suficiente para que $X^* = (x^*, 1-x^*)$ y $Y^* = (y^*, 1-y^*)$ sea un equilibrio de Nash para el juego con matrices de pagos (A, B) es*

$$E_1(1, Y^*) \leq E_1(X^*, Y^*) \tag{3.1}$$

$$E_1(2, Y^*) \leq E_1(X^*, Y^*) \tag{3.2}$$

$$E_2(X^*, 1) \leq E_2(X^*, Y^*) \tag{3.3}$$

$$E_2(X^*, 2) \leq E_2(X^*, Y^*) \tag{3.4}$$

Demostración. Para ver que esto es cierto, primero note que si (X^*, Y^*) es un equilibrio de Nash entonces las desigualdades se cumplen por definición. Sólo resta probar que las desigualdades son suficientes.

Suponga que las desigualdades son válidas para (X^*, Y^*). Sean $X = (x, 1-x)$, $Y = (y, 1-y)$ cualesquiera estrategias mixtas. Se tiene que multiplicando (3.1) por $x \geq 0$ y (3.2) por $(1-x) \geq 0$

$$
\begin{aligned}
xE_1(1, Y^*) &= x(1,0)AY^{*T} = x\left(a_{11}y^* + a_{12}(1-y^*)\right) \\
&\leq xE_1(X^*, Y^*)
\end{aligned}
$$

y

$$
\begin{aligned}
(1-x)E_1(2, Y^*) &= (1-x)(0,1)AY^{*T} \\
&= (1-x)\left(a_{21}y^* + a_{22}(1-y^*)\right) \\
&\leq (1-x)E_1(X^*, Y^*).
\end{aligned}
$$

Sumando ambos resultados

$$\begin{aligned} xE_1(1,Y^*) + (1-x)E_1(2,Y^*) &= x(a_{11}y^* + a_{12}(1-y^*)) + \\ &\quad (1-x)(a_{21}y^* + a_{22}(1-y^*)) \\ &\leq xE_1(X^*,Y^*) + (1-x)E_1(X^*,Y^*) \\ &= E_1(X^*,Y^*) \end{aligned}$$

pero

$$x(a_{11}y^* + a_{12}(1-y^*)) + (1-x)(a_{21}y^* + a_{22}(1-y^*)) = XAY^{*T},$$

entonces

$$\begin{aligned} (x,(1-x))AY^{*T} &= XAY^{*T} \\ &= E_1(X,Y^*) \leq E_1(X^*,Y^*). \end{aligned}$$

Como $X \in S_n$ es cualquier estrategia mixta para el jugador 1, se tiene X satisface la primera parte de la definición 3.1.1, para que (X^*,Y^*) sea un equilibrio de Nash. Ahora multiplicando (3.3) y (3.4) por $y \geq 0$ y $(1-y) \geq 0$, respectivamente

$$\begin{aligned} yE_2(X^*,1) &= yX^*B(1,0)^T = y(b_{11}x^* + b_{21}(1-x^*)) \\ &\leq yE_2(X^*,Y^*) \end{aligned}$$

y

$$\begin{aligned} (1-y)E_2(X^*,2) &= (1-y)X^*B(0,1)^T \\ &= (1-y)(b_{12}x^* + b_{22}(1-x^*)) \\ &\leq (1-y)E_2(X^*,Y^*) \end{aligned}$$

así,

$$\begin{aligned} yE_2(X^*,1) + (1-y)E_2(X^*,2) &= y(b_{11}x^* + b_{21}(1-x^*)) \\ &\quad + (1-y)(b_{12}x^* + b_{22}(1-x^*)) \\ &= X^*BY^T = E_2(X^*,Y) \\ &\leq yE_2(X^*,Y^*) + (1-y)E_2(X^*,Y^*) \\ &= E_2(X^*,Y^*). \end{aligned}$$

De lo cual se deduce que

$$E_2\left(X^*,Y\right) \leq E_2\left(X^*,Y^*\right),$$

como $Y \in S_m$ es cualquier estrategia para el jugador 2, por la Definición 3.1.1 se tiene que (X^*,Y^*) es un equilibrio de Nash. ■

Definición 3.1.4 *Sean* $X = (x, 1-x)$*,* $Y = (y, 1-y)$*,* $0 \leq x \leq 1$*,* $0 \leq y \leq 1$ *estrategias y sean* $f(x,y) = E_1(X,Y)$*, y* $g(x,y) = E_2(X,Y)$*. El* conjunto de reacción *para el jugador* 1 *es el conjunto de puntos*

$$R_1 = \left\{ (x,y) \middle| 0 \leq x, y \leq 1, \max_{0 \leq z \leq 1} f(z,y) = f(x,y) \right\},$$

y el conjunto de reacción *para el jugador* 2 *es el conjunto de puntos*

$$R_2 = \left\{ (x,y) \middle| 0 \leq x, y \leq 1, \max_{0 \leq w \leq 1} g(x,w) = g(x,y) \right\}.$$

Un punto $(x^*,y^*) \in R_1$ implica que $X^* = (x^*, 1-x^*)$ es la mejor respuesta a $Y^* = (y^*, 1-y^*)$. Similarmente, si $(x^*,y^*) \in R_2$, entonces $Y^* = (y^*, 1-y^*)$ es la mejor respuesta a $X^* = (x^*, 1-x^*)$. En consecuencia, un punto (X^*,Y^*) en R_1 y R_2 es un equilibrio de Nash.

Para simplificar la notación, consideremos $X^* = (x^*, 1-x^*)$ y $Y^* = (y^*, 1-y^*)$ como $(x, 1-x)$ y $(y, 1-y)$, y supongamos que es un equilibrio de Nash. Como $E_1(1,Y) \leq E_1(X,Y)$ y $E_1(2,Y) \leq E_1(X,Y)$, se tienen las desigualdades siguientes

$$\begin{aligned} & (a_{11} - a_{12} - a_{21} + a_{22})\, y + (a_{12} - a_{22}) + (a_{21} - a_{22})\, y + a_{22} \\ \leq\ & (a_{11} - a_{12} - a_{21} + a_{22})\, xy + (a_{12} - a_{22})\, x + (a_{21} - a_{22})\, y + a_{22} \end{aligned}$$

y

$$\begin{aligned} & (a_{11} - a_{12} - a_{21} + a_{22})\, xy + (a_{12} - a_{22})\, x + (a_{21} - a_{22})\, y + a_{22} \\ \geq\ & (a_{21} - a_{22})\, y + a_{22}. \end{aligned}$$

Simplificando, se tiene que

$$M(1-x)\,y - m(1-x) \leq 0 \text{ y } Mxy - mx \geq 0, \tag{3.5}$$

donde

$$M = a_{11} - a_{12} - a_{21} + a_{22} \text{ y } m = a_{22} - a_{12}.$$

Así, esto significa que cualquier (x, y) puede ser un equilibrio de Nash, $X = (x, 1-x)$ y $Y = (y, 1-y)$, si satisface (3.5), y además $0 \leq x, y \leq 1$. Las desigualdades son no lineales y no fáciles de resolver sin algunas hipótesis sobre m y M.

Se consideran los siguientes casos:

1. $M = m = 0$. En este caso $M(1-x)y - m(1-x) = 0$, y $Mxy - mx = 0$, para cualquier $x, y \in [0, 1]$. Este es el caso trivial, porque si $M = m = 0$, entonces $a_{22} = a_{12}$ y $a_{11} = a_{21}$. Es decir, no importa lo que el jugador 1 haga.

2. $M = 0$, $m > 0$. Entonces $-m(1-x) \leq 0$ y $-mx \geq 0$, implica que $x = 0$ y y es cualquiera en $[0, 1]$.

3. $M = 0$, $m < 0$. Entonces $(1-x) \leq 0$, y $x \geq 0$. La solución es $x = 1$, $0 \leq y \leq 1$.

4. $M > 0$. Entonces $M(1-x)y - m(1-x) \leq 0$, y $Mxy - mx \geq 0$, se tiene que:

   Si $x = 1$

$$\begin{aligned} M(1-x)y - m(1-x) &= 0 \leq 0 \text{ y} \\ Mxy - mx &= My - m \geq 0 \\ &\Rightarrow y \geq \frac{m}{M}. \end{aligned}$$

   Si $x = 0$

$$\begin{aligned} M(1-x)y - m(1-x) &= My - m \leq 0 \text{ y} \\ Mxy - mx &= 0 \geq 0 \\ &\Rightarrow y \leq \frac{m}{M}. \end{aligned}$$

   Si $0 < x < 1$

$$\begin{aligned} M(1-x)y - m(1-x) &\leq 0 \Rightarrow My - m \leq 0 \\ Mxy - mx &\geq 0 \Rightarrow My - m \geq 0 \\ &\Rightarrow y = \frac{m}{M}. \end{aligned}$$

Por tanto, la solución es:

$$\begin{array}{ll} \text{si } x = 0, & 0 \leq y \leq \dfrac{m}{M}, \\ \text{si } 0 < x < 1, & y = \dfrac{m}{M}, \\ \text{si } x = 1, & \dfrac{m}{M} \leq y \leq 1. \end{array}$$

5. $M < 0$. Entonces $M\,(1-x)\,y - m\,(1-x) \leq 0$, y $Mxy - mx \geq 0$, analizando esto para los posibles valores de x se tiene que

   Si $x = 1$

$$\begin{aligned} M(1-x)y - m(1-x) &= 0 \leq 0 \text{ y} \\ Mxy - mx &= My - m \geq 0 \\ &\Rightarrow y \leq \frac{m}{M}. \end{aligned}$$

   Si $x = 0$

$$\begin{aligned} M\,(1-x)\,y - m\,(1-x) &= My - m \leq 0 \text{ y} \\ Mxy - mx &= 0 \geq 0 \\ &\Rightarrow y \geq \frac{m}{M}. \end{aligned}$$

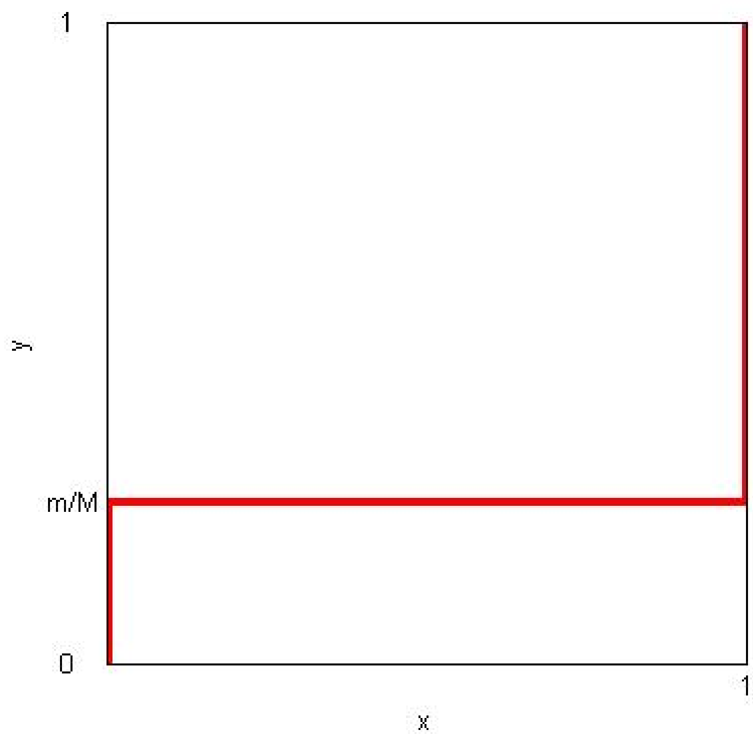
   Si $0 < x < 1$

$$\begin{aligned} M\,(1-x)\,y - m\,(1-x) &\leq 0 \Rightarrow My - m \leq 0 \\ Mxy - mx &\geq 0 \Rightarrow My - m \geq 0 \\ &\Rightarrow y = \frac{m}{M}. \end{aligned}$$

   Así la solución es

$$\begin{array}{ll} \text{si } x = 1, & 0 \leq y \leq \dfrac{m}{M}, \\ \text{si } 0 < x < 1, & y = \dfrac{m}{M}, \\ \text{si } x = 0, & \dfrac{m}{M} \leq y \leq 1. \end{array}$$

La siguiente figura representa el *conjunto de reacción* para el jugador 1 en el caso $M > 0$. Los ejes son las variables x y y. La línea roja representa los puntos (x, y) tales que $X = (x, 1-x)$, $Y = (y, 1-y)$ son solución para las ecuaciones $E_1(1, Y) \leq E_1(X, Y)$ y $E_1(2, Y) \leq E_1(X, Y)$.



Así, para $0 \leq y \leq 1$, la mejor respuesta para el jugador 1 es $X = (x, 1-x)$, donde x intersecta la línea roja para y dada. La línea roja representa el *conjunto de reacción* para el jugador 1 con Y dada.

El *conjunto de reacción* para el jugador 1, en el caso $M > 0$, está dado por:

$$R_1 = \left\{(0, y) \middle| 0 \leq y \leq \frac{m}{M}\right\} \cup \left\{\left(x, \frac{m}{M}\right) |0 < x < 1\right\} \cup \left\{(1, y) \middle| \frac{m}{M} \leq y \leq 1\right\}.$$

Ahora, considere las ecuaciones $E_2(X, 1) \leq E_2(X, Y)$ y $E_2(X, 2) \leq E_2(X, Y)$. Se tienen las desigualdades siguientes

$$\begin{aligned} & (b_{11} - b_{12} - b_{21} + b_{22})\, x + (b_{12} - b_{22})\, x + (b_{21} - b_{22}) + b_{22} \\ \leq \quad & (b_{11} - b_{12} - b_{21} + b_{22})\, xy + (b_{12} - b_{22})\, x + (b_{21} - b_{22})\, y + b_{22} \end{aligned}$$

y

$$\begin{aligned} & (b_{11} - b_{12} - b_{21} + b_{22})\, xy + (b_{12} - b_{22})\, x + (b_{21} - b_{22})\, y + b_{22} \\ \geq\ & (b_{12} - b_{22})\, x + b_{22}. \end{aligned}$$

Así, la desigualdad que se tiene que resolver es

$$Rx\,(1-y) - r\,(1-y) \leq 0 \text{ y } Rxy - ry \geq 0.$$

Donde

$$R = b_{11} - b_{12} - b_{21} + b_{22} \text{ y } r = b_{22} - b_{21}.$$

De nuevo, se consideran distintos casos, así

1. $R = 0$, $r = 0$. La solución es $0 \leq x \leq 1$, $0 \leq y \leq 1$.
2. $R = 0$, $r > 0$. Entonces $y = 0$, $x \in [0, 1]$.
3. $R = 0$, $r < 0$. La solución es $y = 1$, $0 \leq x \leq 1$.
4. $R > 0$. Entonces $Rx\,(1-y) - r\,(1-y) \leq 0$, y $Rxy - ry \geq 0$.
   Si $y = 0$

$$\begin{aligned} Rx(1-y) - r\,(1-y) &= Rx - r \leq 0 \text{ y} \\ Rxy - ry &= 0 \\ &\Rightarrow x \leq \frac{r}{R}. \end{aligned}$$

   Si $y = 1$

$$\begin{aligned} Rx(1-y) - r\,(1-y) &= 0 \leq 0 \text{ y} \\ Rxy - ry &= Rx - r \geq 0 \\ &\Rightarrow x \geq \frac{r}{R}. \end{aligned}$$

   Si $0 < y < 1$

$$\begin{aligned} Rx(1-y) - r\,(1-y) &\leq 0 \Rightarrow Rx - r \leq 0 \\ Rxy - ry &\geq 0 \Rightarrow Rx - r \geq 0 \\ &\Rightarrow x = \frac{r}{R}. \end{aligned}$$

Se tiene que la solución es:

$$\text{si } y = 0 \Rightarrow 0 \leq x \leq \frac{r}{R},$$

$$\text{si } 0 < y < 1 \Rightarrow x = \frac{r}{R},$$

$$\text{si } y = 1 \Rightarrow \frac{r}{R} \leq x \leq 1.$$

5. $R < 0$. Entonces $Rx\,(1-y) - r\,(1-y) \leq 0$, y $Rxy - ry \geq 0$.

   Si $y = 0$

$$\begin{aligned} Rx(1-y) - r\,(1-y) &= Rx - r \leq 0 \text{ y} \\ Rxy - ry &= 0 \geq 0 \\ &\Rightarrow x \geq \frac{r}{R}. \end{aligned}$$

   Si $0 < y < 1$

$$\begin{aligned} Rx(1-y) - r\,(1-y) &\leq 0 \Rightarrow Rx - r \leq 0 \\ Rxy - ry &\geq 0 \Rightarrow Rx - r \geq 0 \\ &\Rightarrow x = \frac{r}{R}. \end{aligned}$$

   Si $y = 1$

$$\begin{aligned} Rx(1-y) - r\,(1-y) &= 0 \leq 0 \text{ y} \\ Rxy - ry &= Rx - r \geq 0 \\ &\Rightarrow x \leq \frac{r}{R}. \end{aligned}$$

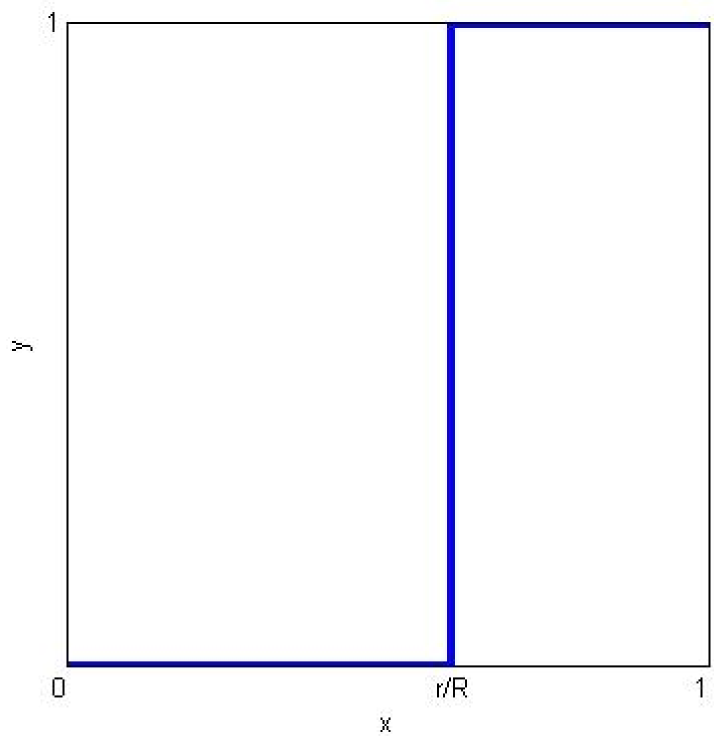
   Se tiene que la solución es:

$$\text{si } y = 0 \Rightarrow \frac{r}{R} \leq x \leq 1,$$

$$\text{si } y = 1 \Rightarrow 0 \leq x \leq \frac{r}{R},$$

$$\text{si } 0 < y < 1 \Rightarrow x = \frac{r}{R}.$$

La siguiente gráfica representa el *conjunto de reacción* para el jugador 2 en el caso $R > 0$ para un x dado.



El *conjunto de reacción* para el jugador 2, en el caso $R > 0$, está dado por:

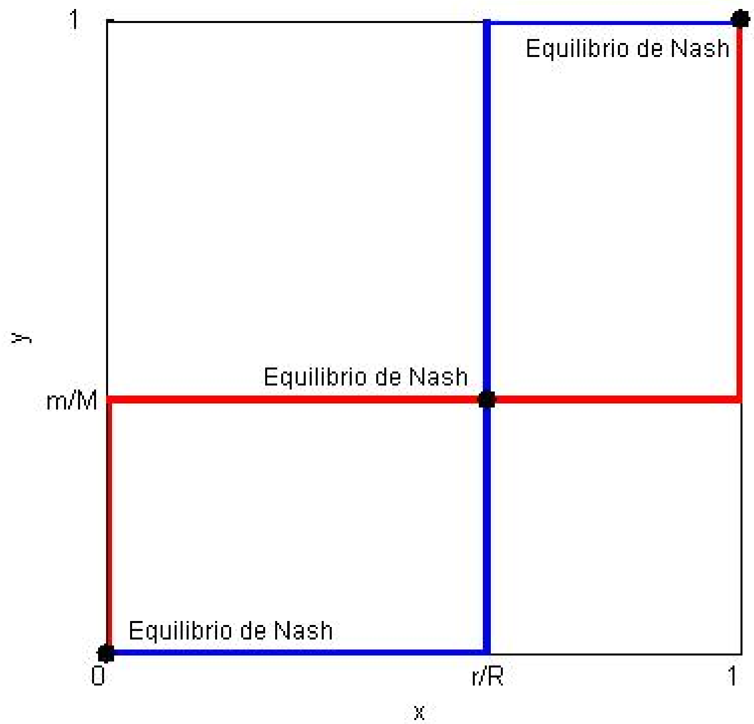
$$R_2 = \left\{(x,0)\,\Big|\, 0 \leq x \leq \frac{r}{R}\right\} \cup \left\{\left(\frac{r}{R}, y\right) |0 < y < 1\right\} \cup \left\{(x,1)\,\Big|\,\frac{r}{R} \leq x \leq 1\right\}.$$

Si se dibujara la gráfica de R_1 sobre la gráfica de R_2, se tiene que el punto (o puntos) en la intersección $R_1 \cap R_2$ es un equilibrio de Nash. En otras palabras, un equilibrio de Nash es un punto que se encuentra en ambos *conjuntos de reacción* de cada jugador.

Finalmente, para $M \neq 0$, $R \neq 0$, se tiene un equilibrio de Nash $X^* = (x, 1-x)$, $Y^* = (y, 1-y)$ para

$$x = \frac{r}{R} = \frac{b_{22} - b_{21}}{b_{11} - b_{12} - b_{21} + b_{22}} \text{ y } y = \frac{m}{M} = \frac{a_{22} - a_{12}}{a_{11} - a_{12} - a_{21} + a_{22}}.$$

Un equilibrio de Nash puro es un punto en las esquinas de la intersección de los *conjuntos de reacción*.



3.2. Puntos de Equilibrio de Nash Interiores

Cuando se buscan equilibrios de Nash se está tratando de maximizar la función de pago esperado suponiendo que el otro jugador hace lo mismo, así que resulta conveniente utilizar las herramientas de máximos y mínimos del cálculo. El procedimiento es el siguiente:

1. Las matrices de pago son $A_{2\times2}$ para el jugador 1 y $B_{2\times2}$ para el jugador 2. El pago esperado para el jugador 1 es $E_1(X,Y) = XAY^T$ y $E_2(X,Y) = XBY^T$ para el jugador 2.

2. Considere $X = (x, 1-x)$, $Y = (y, 1-y)$ así cada pago es una función de una sola variable. Se tiene que

$$\begin{aligned} E_I(x,y) &= E_1(X,Y), \\ E_{II}(x,y) &= E_2(X,Y). \end{aligned}$$

3. Encontrar las derivadas parciales y resolver el sistema de ecuaciones $\frac{\partial E_I}{\partial x} = 0$, $\frac{\partial E_{II}}{\partial y} = 0$.

4. Si hay solución para las ecuaciones y satisfacen que $0 \leq x \leq 1$, $0 \leq y \leq 1$, entonces (x, y) es un equilibrio de Nash mixto.

Ejemplo 3.2.1 *Considere un juego con matrices de pagos dada por*

$$A = \begin{bmatrix} 4 & 3 \\ 2 & 6 \end{bmatrix}, \; B = \begin{bmatrix} -5 & 2 \\ 2 & -7 \end{bmatrix}.$$

Sean $X = (x, 1 - x)$, $Y = (y, 1 - y)$. *Se tiene que*

$$\begin{aligned} E_I(x, y) &= [2x + 2]\, y + [6 - 3x](1 - y), \\ E_{II}(x, y) &= [2 - 7x]\, y + [9x - 7](1 - y). \end{aligned}$$

El jugador 1 *busca maximizar* E_I, *para* y *fija, así que*

$$\begin{aligned} \frac{\partial E_I}{\partial x} &= 2y - 3(1 - y) \\ &= 5y - 3 = 0 \\ &\Rightarrow \; y = \frac{3}{5}. \end{aligned}$$

De manera análoga, el jugador 2 *desea maximizar* E_{II}, *para un* x *fijo, por tanto*

$$\begin{aligned} \frac{\partial E_{II}}{\partial y} &= 2 - 7x - (9x - 7) \\ &= 9 - 16x = 0 \\ &\Rightarrow \; x = \frac{9}{16}. \end{aligned}$$

De acuerdo a lo anterior $X^* = \left(\frac{9}{16}, \frac{7}{16}\right)$, $Y^* = \left(\frac{3}{5}, \frac{2}{5}\right)$ *es un equilibrio de Nash para el juego.*

Se incluye un programa en MATLAB para resolver este problema (ver Apéndice B).

3.3. Existencia de Equilibrios de Nash

En la sección 3.1 se afirma que un equilibrio de Nash es un punto que se encuentra en la intersección de los *conjuntos de reacción*, R_1 y R_2, pero ¿cómo asegurar que tal intersección es no vacía? En esta sección se vera que tal afirmación es cierta.

Definición 3.3.1 *El* conjunto de mejor respuesta *para cada jugador se define como*

$$
\begin{aligned}
MR_1(Y) &= \left\{ X \in S_n \,\middle|\, E_1(X,Y) = \max_{p \in Sn} E_1(p,Y) \right\}, \\
MR_2(X) &= \left\{ Y \in S_m \,\middle|\, E_2(X,Y) = \max_{q \in Sm} E_2(X,q) \right\}.
\end{aligned}
$$

La diferencia entre el *conjunto de reacción* y el *conjunto de mejor respuesta* es que el *conjunto de reacción* R_1 consiste de parejas de estrategias (X,Y), para la cual $E_1(X,Y) = \max_p E_1(p,Y)$, mientras que el conjunto $MR_1(Y)$ consiste de una estrategia (o colección de estrategias) X para el jugador 1 que es la mejor respuesta a una estrategia Y fija.

El siguiente teorema garantiza la existencia de al menos un equilibrio de Nash.

Teorema 3.3.2 *(de Nash). Existen $X^* \in S_n$ y $Y^* \in S_m$ tales que*

$$
\begin{aligned}
E_1(X^*,Y^*) &= X^* A Y^{*T} \geq E_1(X,Y^*), \\
E_2(X^*,Y^*) &= X^* B Y^{*T} \geq E_2(X^*,Y),
\end{aligned}
$$

para cualesquiera estrategias mixtas $X \in S_n$, $Y \in S_m$.

Demostración. La idea de la prueba se basa en el teorema del punto fijo de Kakutani (ver Apéndice A).

Primero se vera que $S_n \times S_m$ es un conjunto cerrado, acotado y convexo.

Se tiene que

$$
\begin{aligned}
S_n &= \{(x,\tilde{x}) \,|0 \leq x \leq 1, 0 \leq \tilde{x} \leq 1\} \\
&= \{(x,1-x) \,|0 \leq x \leq 1\} \\
&\subseteq [0,1] \times [0,1].
\end{aligned}
$$

Así, S_n es un conjunto acotado. Considere $\left\langle \bar{X}_k \right\rangle \subset S_n$ una sucesión convergente a $\bar{X}$, para ver que S_n en un conjunto cerrado basta probar que $\bar{X} \in S_n$. Se tiene que:

$$X_k = (x_k, 1 - x_k) \Rightarrow X = (x, 1 - x)$$

con $0 \leq x_k \leq 1$ para toda k y $0 \leq x \leq 1$. Por tanto, $X \in S_n$, así S_n es un conjunto cerrado. Como S_n es cerrado y acotado, se sigue que es compacto.

Ahora veamos que es un conjunto convexo, para esto sean $X = (x, 1 - x)$, $\tilde{X} = (\tilde{x}, 1 - \tilde{x}) \in S_n$ y $\lambda \in [0, 1]$

$$\begin{aligned}
\lambda X + (1 - \lambda) \tilde{X} &= \lambda (x, 1 - x) + (1 - \lambda) (\tilde{x}, 1 - \tilde{x}) \\
&= (\lambda x + (1 - \lambda) \tilde{x}, 1 - (\lambda x + (1 - \lambda) \tilde{x})) \\
&\in S_n.
\end{aligned}$$

Por tanto S_n es convexo. De manera similar se puede probar que S_m es convexo y acotado.

Luego $S_n \times S_m$ es un conjunto compacto y convexo.

Ahora para cada par de estrategias (X, Y), se pueden considerar las estrategias de *mejor respuesta*, $MR_2(X)$ del jugador 2 a X y $MR_1(Y)$ del jugador 1 a Y.

Siempre que se maximiza una función continua, lo cual es cierto para $E_1(X, Y)$, sobre un conjunto cerrado y acotado, se tiene que la función alcanza su máximo. Por tanto, se sabe que $MR_1(Y) \neq \emptyset$. De manera análoga se tiene que $MR_2(X) \neq \emptyset$.

Se define la multifunción

$$\varphi : (X, Y) \in S_n \times S_m \rightarrow MR_1(Y) \times MR_2(X) \subset S_n \times S_m$$

tal que a cada pareja (X, Y) le asigna la estrategia de mejor respuesta $\left(\hat{X}, \hat{Y}\right) \in \varphi(X, Y)$ con $\hat{X} \in MR_1(Y)$ y $\hat{Y} \in MR_2(X)$.

Es natural que el equilibrio de Nash buscado se encuentre entre las estrategias de *mejor respuesta* del oponente, debido a la definición dada del equilibrio de Nash.

Se probará que si φ satisface las condiciones del teorema del punto fijo de Kakutani (ver Apéndice A), entonces se tiene la existencia del equilibrio de Nash.

Solo resta probar que la gráfica de φ es cerrada, (ver [12] p. 51), donde

$$graf\,(\varphi) = \{(Z, \varphi\,(Z))\,|Z \in S_n \times S_m\}\,.$$

Sea $(Z_k, \varphi\,(Z_k)) \in graf\,(\varphi)$ tal que $(Z_k, \varphi\,(Z_k)) \to (Z, \varphi\,(Z))$; se quiere probar que $(Z, \varphi\,(Z)) \in graf\,(\varphi)$. Como $S_n \times S_m$ es un conjunto cerrado, entonces $Z \in S_n \times S_m$: Resta probar que $\varphi\,(Z) \in MR_1\,(Y) \times MR_2\,(X)$. Se tiene que

$$\varphi\,(Z_k) = \left(U_k^1, U_k^2\right) \subset S_n \times S_m$$

y

$$\left(Z_k^1, Z_k^2\right) \to \left(U_k^1, U_k^2\right)$$

donde $U_k^1 \in MR_1\,(Z_k^2)$ y $U_k^2 \in MR_1\,(Z_k^1)$, es decir,

$$\begin{aligned} E_1(U_k^1, Z_k^2) &= \max_{p \in S_n} E_1\left(p, Z_k^2\right), \\ E_2(Z_k^1, U_k^2) &= \max_{q \in S_m} E_2\left(Z_k^1, q\right) \end{aligned}$$

tomando límite

$$\begin{aligned} E_1(U^1, Z^2) &= \max_{p \in S_n} E_1\left(p, Z^2\right) \\ E_2(Z^1, U^2) &= \max_{q \in S_m} E_2\left(Z^1, q\right). \end{aligned}$$

Las relaciones anteriores son debidas a la linealidad de la esperanza y al teorema del Máximo (ver Apéndice A y [11]).

Por tanto, $\varphi\,(Z) = (U^1, U^2) \in MR_1\,(Y) \times MR_2\,(X)$. Asi $graf\,(\varphi)$ es un conjunto cerrado y así φ es compacta, convexa y superiormente semicontinua. Esto significa que φ cumple los requerimientos del Teorema de Kakutani y entonces tiene un punto fijo. Considere una pareja de estrategias $(X^*, Y^*) \in \varphi\,(X^*, Y^*)$ con $X^* \in MR_1\,(Y^*)$ y $Y^* \in MR_2\,(X^*)$. Es decir,

$$E_1\,(X^*, Y^*) = \max_{p \in Sn} E_1\,(p, Y^*) \geq E_1\,(X; Y^*)\,,$$

para toda $X \in S_n$, y

$$E_2\,(X^*, Y^*) = \max_{q \in Sm} E_2\,(X^*, q) \geq E_2\,(X^*; Y)\,,$$

para toda $Y \in S_m$. Así, se tiene que (X^*, Y^*) es un equilibrio de Nash. ∎

3.4. Ejemplo

Ejemplo 3.4.1 *Suponga que se analiza un proceso de manufactura que depende de dos máquinas. Este proceso puede ser visto como un juego de dos jugadores, considere las estrategias para cada jugador como:*

$$\begin{aligned} S_2^1 &= (\textit{dar mantenimiento a la máquina } 1, \textit{ que la máquina } 1 \textit{ trabaje}), \\ S_2^2 &= (\textit{dar mantenimiento a la máquina } 2, \textit{ que la máquina } 2 \textit{ trabaje}). \end{aligned}$$

Suponga que observa el juego en una sola etapa y además que, si la máquina 1 *no funciona la máquina* 2 *paga un costo de operación más un costo de almacenaje; y si la máquina* 2 *no funciona la máquina* 1 *tampoco. Así las matrices de pago están dadas por:*

$A =$

	0	1
0	$c.m.m\ 1$	$c.m.m\ 1$
1	0	$c.o.m\ 1$

,

$B =$

	0	1
0	$c.m.m\ 2$	$c.o.m\ 2\ +\ c.a$
1	$c.m.m\ 2$	$c.o.m\ 2$

.

Donde las filas de A y B representan las estrategias del jugador 1 *y las columnas las estrategias del jugador jugador* 2 *y,*

$$\begin{aligned} 0 &\equiv \textit{dar mantenimiento a la máquina}, \\ 1 &\equiv \textit{permitir que la máquina continue trabajando}, \\ c.m.m\ 1 &\equiv \textit{costo de mantenimiento máquina } 1, \\ c.o.m\ 1 &\equiv \textit{costo de operación máquina } 1, \\ c.m.m\ 2 &\equiv \textit{costo de mantenimiento máquina } 2, \\ c.o.m\ 2 &\equiv \textit{costo de operación máquina } 2, \\ c.a &\equiv \textit{costo de almacenaje}. \end{aligned}$$

Por tanto, las posibles jugadas son

1. *Dar mantenimiento a ambas máquinas, pagando el costo de mantenimiento,*

2. *Dar mantenimiento a la máquina* 1*, pagando un costo de mantenimiento y que la máquina* 2 *continúe trabajando pagando el costo de operación más el costo de almacenaje,*

3. *Que la máquina* 1 *trabaje perfectamente sin un costo y dar mantenimiento a la máquina* 2 *pagando el costo de mantenimiento,*

4. *Que ambas máquinas funcionen en condiciones óptimas, pagando un costo de operación,*

Además, suponga que el costo de operación de la máquina 1 *es mayor que el de mantenimiento; e inversamente para la máquina* 2 *y, además que el costo de mantenimiento de la máquina* 2 *es menor que el costo de operación más el costo de almacenaje.*

Lo que se desea es minimizar los costos de funcionamiento, es decir, hallar un equilibrio de Nash.

Un ejemplo de la situación anterior, es

$$A = \begin{bmatrix} -85 & -85 \\ 0 & -128 \end{bmatrix},\ B = \begin{bmatrix} -97 & -100 \\ -97 & -83 \end{bmatrix}. \tag{3.6}$$

Resolviendo este juego a través del análisis por casos dado en la sección 3.1 se tiene que

$$\begin{aligned} M &= a_{11} - a_{12} - a_{21} + a_{22} \\ &= -85 + 85 - 0 - 128 \\ &= -128 < 0 \end{aligned}$$

$$\begin{aligned} m &= a_{22} - a_{12} \\ &= -128 + 85 \\ &= -43 < 0 \end{aligned}$$

$$\Rightarrow \quad \frac{m}{M} = \frac{-43}{-128} = 0.3359.$$

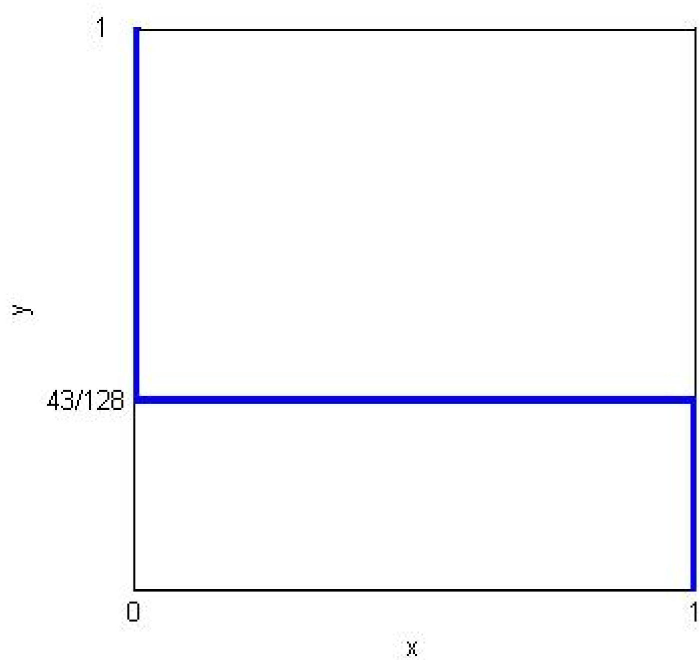
Así

$$si\ x = 1 \Rightarrow 0 \leq y \leq \frac{43}{128},$$

$$si\ 0 < x < 1 \Rightarrow y = \frac{43}{128},$$

$$si\ x = 0 \Rightarrow \frac{43}{128} \leq y \leq 1.$$

Se tiene que el conjunto de reacción *para el jugador* 1 *es*



Ahora, considerando al jugador 2*, se tiene que*

$$\begin{aligned} R &= b_{11} - b_{12} - b_{21} + b_{22} \\ &= -97 + 100 + 97 - 83 \\ &= 17 > 0 \end{aligned}$$

$$\begin{aligned} r &= b_{22} - b_{21} \\ &= -83 + 97 \\ &= 14 > 0 \end{aligned}$$

$$\Rightarrow \frac{r}{R} = \frac{14}{17} = 0.82359.$$

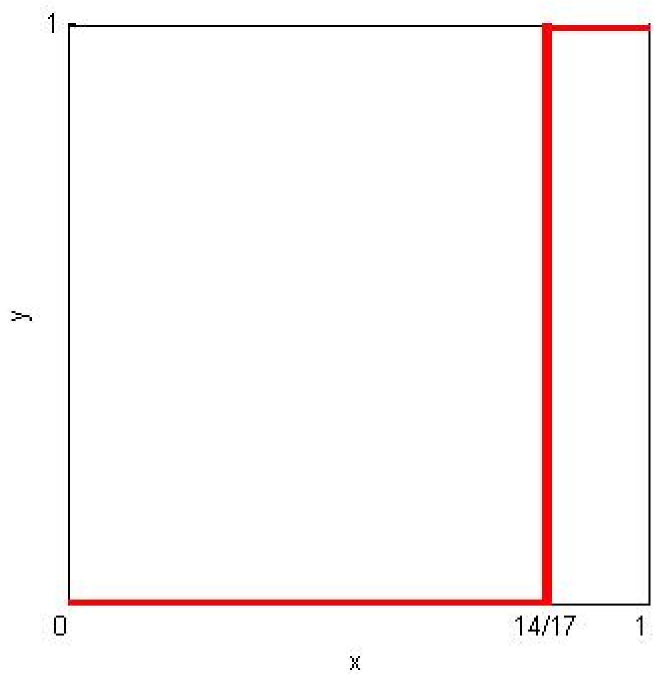
Por tanto

$$si\ y = 0 \Rightarrow 0 \le x \le \frac{14}{17},$$

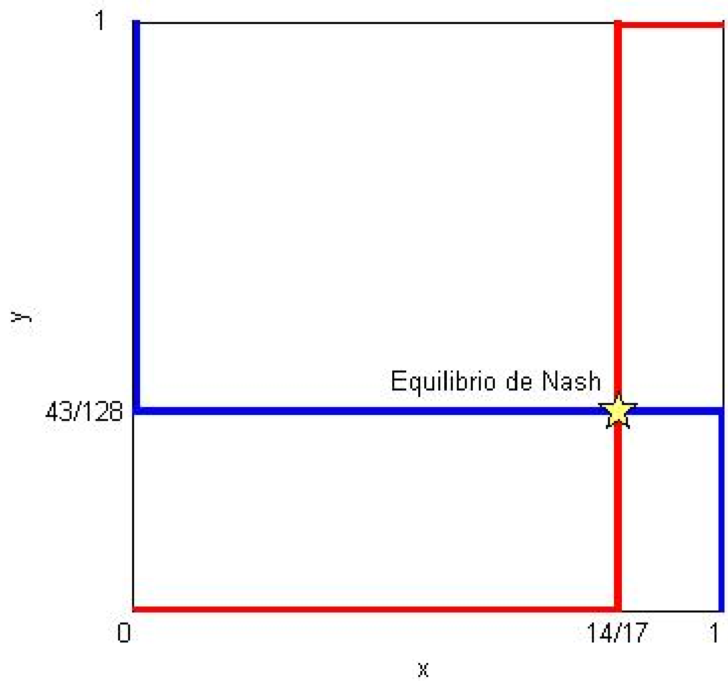
$$si\ 0 < y < 1 \Rightarrow x = \frac{14}{17},$$

$$si\ y = 1 \Rightarrow \frac{14}{17} \le x \le 1.$$

Y su conjunto de reacción *es*



La siguiente gráfica representa la intersección de los conjuntos de reacción *de cada jugador*



Así, se tiene que un equilibrio de Nash para el juego se encuentra en los puntos $X^* = \left(\frac{14}{17}, \frac{3}{17}\right)$ *y* $Y^* = \left(\frac{43}{128}, \frac{85}{128}\right)$*, por tanto con probabilidad* 0,82 *el jugador* 1 *elige dar mantenimiento a la máquina y con probabilidad* 0,66 *para el jugador* 2 *es mejor permitir que la máquina continúe trabajando.*

Capítulo 4

Conclusiones

En la tesis se trabajó con la teoría de procesos de decisión, en particular, con los Procesos de Decisión de Markov (PDM) a tiempo discreto con horizonte finito, y también con teoría de juegos en el caso de suma no nula para dos jugadores. El criterio de rendimiento usado fue el de costo total acumulado. Las estrategias óptimas fueron caracterizadas usando la ecuación de Bellman y en el caso de teoría de juegos, por medio de equilibrios de Nash.

Para la parte de PDM se incluyeron dos ejemplos que tienen que ver con el reemplazamiento de máquinas. El análisis se hace de la forma siguiente, primero se presenta el estudio para una sola máquina y se consideran dos acciones a seguir, reparar o dejar funcionar la máquina. Usando Programación Dinámica se encontró la política óptima para un problema de T periodos. Posteriormente, se hace un análisis considerando dos máquinas, para lo cual se supone que el funcionamiento de cada una de ellas es independiente y se considera que las acciones admisibles son: reparar ambas máquinas, dejar funcionar ambas o reparar alguna de las máquinas. Nuevamente, usando Programación Dinámica se caracteriza a la estrategia óptima. En ambos casos se dieron ejemplos particulares y se elaboraron algoritmos en MATLAB, los cuales permiten resolverlos para un número mayor de estados.

Finalmente, se dieron los resultados principales de la teoría de juegos en el caso de dos jugadores, se introdujo la definición de un equilibrio de Nash y se estudiaron procedimientos de solución para juegos basados en cálculo y por medio del análisis de las matrices de pagos. También, se incluyó un ejemplo de reemplazamiento de máquinas, donde las máquinas son consideradas como los jugadores, sólo que en este caso se supone dependencia en su funcionamiento y se analizó en un sólo periodo de tiempo. En este problema se encontró un equilibrio de Nash y, al igual que en el caso de PDMs, se elaboró un algoritmo en MATLAB para su análisis.

Los problemas que podrían derivar de esta tesis, y que se esperan continuar trabajando son:

i) Hacer un estudio detallado de la teoría de juegos en el contexto de sistemas aleatorios.

ii) Plantear y resolver el ejemplo de máquinas dependientes para un número mayor de estados y periodos.

iii) Elaborar un algoritmo en MATLAB que permita resolver el problema numéricamente, permitiendo aumentar el número de etapas.

iv) Buscar un conjunto de datos para aplicar la teoría desarrollada.

Capítulo 5

Apéndices

5.1. Apéndice A : Resultados Auxiliares

Definición 5.1.1 *Sea (X, τ) un espacio topológico, la mínima $\sigma-$álgebra que contiene a τ es la $\sigma-$álgebra de Borel, es decir, la $\sigma-$álgebra generada por τ. La denotaremos por $\mathcal{B}(X)$.*

Definición 5.1.2 *Sean X y A espacios de Borel. Una multifunción φ de X a A es una función tal que para toda $x \in X$ su imagen $\varphi(x)$ es un subconjnto no vacío de A, es decir, $\varphi : X \to \mathcal{P}(A)$, donde $\mathcal{P}$ denota al conjunto potencia de A. La gráfica de φ es el subconjunto de $X \times A$ definido como*

$$graf(\varphi) = \{(x, a) | x \in X, a \in \varphi(x)\}.$$

Definición 5.1.3 *Una multifunción $\Gamma : X \to Y$ es semicontinua superiormente (u.s.c.) en x si $\Gamma(x)$ es no vacía y si, para cada sucesión x_n convergente a x y cada sucesión $\{y_n\}$ tal que $y_n \in \Gamma(x_n)$, para toda n, existe una subsucesión convergente de $\{y_n\}$ cuyo punto límite y está en $\Gamma(x)$.*

Sean $X \subseteq \mathbb{R}^l$ y $Y \subseteq \mathbb{R}^m$, $f : X \times Y \to \mathbb{R}$ una función; y $\Gamma : X \to Y$ una correspondencia no vacía. Se tiene interes en trabajar con problemas de la forma $\sup_{y \in \Gamma(x)} f(x, y)$. Si para cada x, $f(x, \cdot)$ es continua en y y el conjunto $\Gamma(x)$ es no vacío y compacto, entonces para cada x se alcalza el máximo de la función. En esta caso la función

$$h(x) = \max_{y \in \Gamma(x)} f(x, y) \tag{5.1}$$

está bien definida, al igual que el conjunto no vacío

$$G(x) = \{y \in \Gamma(x) \,|\, f(x, y) = h(x)\} \tag{5.2}$$

de los valores de y donde el máximo se alcanza.

Teorema 5.1.4 *(del Máximo). Sean $X \subseteq \mathbb{R}^l$, $Y \subseteq \mathbb{R}^m$, $f : X \times Y \to \mathbb{R}$ una función continua y $\Gamma : X \to Y$ una multifunción compacta y continua. Entonces la función $h : X. \to \mathbb{R}$ definida en 5.1 es continua, y $G : X \to Y$ definida en 5.2 es una multifunción no vacía, compacta y semicontinua superiormente.*

Demostración. Ver [11]. ■

Teorema 5.1.5 *(de Kakutani). Sean $S \subseteq \mathbb{R}^l$ un conjunto compacto y convexo, una función $f : S \to \mathrm{P}(S)$ semicontinua superiormente, con $f(S)$ compacto y convexo. Entonces f tiene un punto fijo, es decir, existe $s \in S$ tal que $f(s) = s$.*

Demostración. Ver [2]. ■

Teorema 5.1.6 *(de Ionescu-Tulcea). Sea $X_0, X_1, \ldots,$ una sucesión de espacios de Borel y, para $n = 0, 1, \ldots,$ definimos $Y_n := X_0 \times X_1 \times \cdots \times X_n$ y $Y := \Pi_{n=0}^{\infty} X_n$. Sea ν una medida de probabilidad arbitraria sobre X_0 y para cada $n = 0, 1, \ldots,$ $P_n(dx_{n+1} \,|y_n)$ es una probabilidad condicional sobre X_{n+1} dada Y_n. Entonces existe una única medida de probabilidad $P\nu$ sobre Y tal que, para cada rectángulo medible $B_0 \times B_1 \times \cdots \times B_n$ en Y_n,*

$$\begin{aligned} P_\nu(B_0 \times B_1 \times \cdots \times B_n) = & \int_{B_0} \upsilon(dx_0) \int_{B_1} P_0(dx_1 \,|x_0) \\ & \int_{B_2} P_1(dx_2 \,|x_0, x_1) \cdots \\ & \int_{B_n} P_{n-1}(dx_n \,|x_0, \ldots, x_{n-1}) \,. \end{aligned}$$

Además, para cualquier función u medible y no negativa sobre Y, la función

$$x \to \int u(y)\, P_x(dy)$$

es medible en X_0, donde P_x representa a P_ν cuando ν es la probabilidad concentrada en $x \in X_0$.

Demostración. Ver [1]. ■

5.2. Apéndice B: Programas

A continuación se anexan los códigos en MATLAB para los ejemplos 2.5.1, 2.5.2, 3.2.1 y 3.4.1, respectivamente:

Algoritmo 5.2.1 *(Ejemplo 2.5.1)*

```
n=40; N=45; R=15;
g=zeros(1,n);
x=zeros(1,2);
V=zeros(N+1,n);
pol=zeros(N+1,n);
P=zeros(n,n);
A=zeros(n,n);
B=zeros(n,n);
c=zeros(1,n);
A=rand(n);
B=triu(A);
for i=1:n
c(i)=sum(B(i,:));
g(i)=i+1;
end
for i= 1:n
for j=1:n
P(i,j)=B(i,j)/c(i);
end
end
for k=2:N+1
for i=1:n
for j=i:n
x=[R+g(1)+ V(k-1,1), g(i)+ sum(P(i,:).*V(k-1,:))];
[xm,im]=min(x);
pol(k,i)=im;
V(k,i)=min(R+g(1)+ V(k-1,1), g(i)+ sum(P(i,:).*V(k-1,:)));
end
end
end
plot(V', '')
xlabel('Periodos', 'FontSize',12)
```

```
ylabel('Costos','FontSize',12)
```

Algoritmo 5.2.2 *(Ejemplo 2.5.2)*

```
n=10; N=12; R1=5; R2=7;
g=zeros(n,N);
x2=zeros(1,4);
z=zeros(n,n,n+1);
V2=zeros(n,n,N+1);
pol2=zeros(n,n,N+1);
Q1=zeros(n,n);
Q2=zeros(n,n);
A1=zeros(n,n);
A2=zeros(n,n);
B1=zeros(n,n);
B2=zeros(n,n);
c1=zeros(1,n);
c2=zeros(1,n);
A1=rand(n);
A2=rand(n);
B1=triu(A1);
B2=triu(A2);
for i=1:n
c1(i)=sum(B1(i,:));
c2(i)=sum(B2(i,:));
end
for i=1:n
for j=1:n
Q1(i,j)=B1(i,j)/c1(i);
Q2(i,j)=B2(i,j)/c2(i);
g(i,j)=i+j+20;
end
end
for k=2:N+1
for i=1:n
for j=1:n
x2=[g(i,j)+ sum(Q1(i,:).*Q2(j,:)*V2(:,:,k-1)),...
R1+g(1,j)+sum(Q1(i,1).*Q2(j,:).*V2(1,:,k-1)),...
R2+g(i,1)+sum(Q2(j,1).*Q1(i,:)*V2(:,1,k-1)),...
```

```
R1+R2+g(1,1)+V2(1,1,k-1)];
[xm,im]=min(x2);
pol2(i,j,k)=im;
V2(i,j,k)=min([g(i,j)+ sum(Q1(i,:).*Q2(j,:)*V2(:,:,k-1)),...
R1+g(1,j)+sum(Q1(i,1).*Q2(j,:).*V2(1,:,k-1)),...
R2+g(i,1)+sum(Q2(j,1).*Q1(i,:)*V2(:,1,k-1)),...
R1+R2+g(1,1)+V2(1,1,k-1)]);
end
end
end
for t=1:N+1
x=1:1:n;
y=1:1:n;
[x_grid,y_grid]=meshgrid(x,y);
z(x,y) = V2(x,y,t);
mesh(x,y,z(x,y));
hold on
end
```

Algoritmo 5.2.3 *(Ejemplo 3.2.1)*

```
A=[4 3; 2 6];
B=[-5 2; 2 -7];
X=sym('[x,1-x]');
Y=sym('[y; 1-y]');
E1=X*A*Y;
E2=X*B*Y;
de1=diff(E1,'x');
de2=diff(E2,'y');
xsol=solve(de2);
ysol=solve(de1);
EX=[xsol, 1-xsol];
EY=[ysol, 1-ysol];
```

Algoritmo 5.2.4 *(Ejemplo 3.4.1)*

```
A=[-85 -85; 0 -128];
B=[-97 -97; -100 -83];
a=zeros(1,1);
x=zeros(1,1);
```

```
y=zeros(1,1);
X=sym('[x,1-x]');
Y=sym('[y; 1-y]');
M=A(1,1)-A(1,2)-A(2,1)+A(2,2)
m=A(2,2)-A(1,2)
a=m/M
R=B(1,1)-B(1,2)-B(2,1)+B(2,2)
r=B(2,2)-B(2,1)
b=r/R
hold on
ezplot('0','y', [0 1])
ezplot('1','y',[0 1])
ezplot('x','0',[0 1])
ezplot('x','1',[0 1])
if (M==0 & m==0)
ezplot('x','y',[0 1])
elseif (M==0 & m>0)
ezplot('0','y',[0 1])
elseif (M==0 & m<0)
ezplot('1','y',[0 1])
elseif M>0
ezplot('0','y',[0 43/128])
ezplot('x','43/128',[0 1])
ezplot('1','y',[43/128 1])
elseif M<0
ezplot('1','y',[0 43/128])
ezplot('x','43/128',[0 1])
ezplot('0','y',[43/128 1])
end
if (R==0 & r==0)
ezplot('x','y',[0 1])
elseif (R==0 & r>0)
ezplot('x','0',[0 1])
elseif (R==0 & r<0)
ezplot('x','1',[0 1])
elseif R>0
ezplot('x','0',[0 14/17])
ezplot('14/17','y',[0 1])
```

```
ezplot('x','1',[14/17 1])
elseif R<0
ezplot('x','0',[14/17 1])
ezplot('x','1',[0 14/17])
ezplot('14/17','y',[0 1])
end
```

Bibliografía

[1] R. B. Ash, C. A. Doléans-Dade. Probability and Measure Theory. Academic Press Elsevier (2005).

[2] J. P. Aubin, I. Ekeland. Applied Nonlinear Analysis. John Wiley & Sons, Inc., Publication (1984).

[3] E. N. Barron. Game Theory: An Introduction. John Wiley & Sons, Inc., Publication (2008).

[4] R. Bellman. Dynamic Programming. Dover (2003).

[5] D. P. Bertsekas. Dynamic Programming: Deterministic and Stochastic Models. Prentice-Hall, Inc. (1987).

[6] J. García de Jalón, J. I. Rodríguez, J. Vidal. Aprenda Matlab 7.0 como si estuviera en primero. Universidad Politécnica de Madrid (2005).

[7] O. Hernández-Lerma, J. B. Lasserre. Discrete-Time Markov Control Processes Basic Optimality Criteria. Springer (1996).

[8] P. G. Hoel, S. C. Port, C. J. Stone. Introduction to Stochastic Processes. Waveland Press, Inc. (1987).

[9] J. Nash. Non-Cooperative Games. The Annals of Mathematics, Second Series, Volume 54, No. 2, pp. 286-295 (1951).

[10] M. L. Puterman. Markov Decision Processes: Discrete Stochastic Dynamic Programming. Wiley, New York (1994).

[11] N. L. Stokey, R. E. Lucas, Jr. with E. C. Prescott. Recursive Methods in Economic Dynamics. Harvard University Press (1989).

[12] F. Vega-Redondo. Economics and the theory of games. Cambridge University Press, New York (2003).

[13] J. N. Webb. Game Theory: Decisions, Interaction and Evolution. Springer (2007).

[14] G. Zacarías Espinoza. Procesos de Decisión de Markov Descontados. Tesis Licenciatura, BUAP (2007).