PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
ESCUELA DE INGENIERÍA
Departamento de Ingeniería Industrial y de Sistemas
Departamento de Ingeniería de Transporte y Logística

Curso: Optimización Dinámica
Sigla: ICS3105 / ICT3473
Semestre: 02'2020
Profesor: Mathias Klapp

EXAMEN

PLAZO DE ENTREGA: JUEVES 17 DE DICIEMBRE 2020, 12:00 PM

- *La prueba contiene 60 puntos entre todas sus partes. El orden de las partes no tiene relación con su dificultad.*
- *Al desarrollarla, cada estudiante puede acceder a los apuntes y videos del curso, utilizar calculadora y su PC como fuente de cómputo offline.*
- *Este examen es 100 % personal. Está prohibido comunicarse con compañeros u otra persona al durante su desarrollo. También se prohibe acceder a material en internet fuera de la página del curso.*
- *El sistema de evaluación coloca la confianza en el estudiante. Por lo mismo, una falta a la ética se considerará gravísima.* ***Un estudiante que sea descubierto faltando a la ética será evaluado con un 1.1 en el curso y procesado por ello con el mayor rigor.***
- *Se permite hacer preguntas en el foro de CANVAS, pero solo los ayudantes y el profesor estarán autorizados para responderlas.* ***El estudiante que responda una pregunta de otro estudiante en el foro será sancionado****.*

Pregunta 1: Optimizando el consumo (15 puntos)

Un millonario decide no volver a trabajar y vivir de los intereses de sus inversiones en el mercado financiero. Al comienzo de este año su patrimonio es de $\$B$ pesos.

Cada año, el millonario busca determinar cuanto de su dinero gastar y cuanto ahorrar e invertir un año más. Todo el capital ahorrado en el año t genera en el año $t+1$ una ganancia $r > 0$ por peso invertido, es decir, invertir \$100 pesos en el año t genera $(1+r)\cdot 100$ pesos en el año $t+1$.

Suponga que consumir $\$x_t$ pesos durante el año t le brinda al millonario un bienestar evaluado en $\sqrt{x_t}$ utiles (*i.e.*, unidades de utilidad) ese año t. Esta función de bienestar es cóncava en el consumo anual, pues la utilidad marginal es decreciente en función del nivel de consumo. Adicionalmente, para comparar bienestar obtenido en diferentes momentos del tiempo, el millonario establece que una unidad de utilidad obtenido en el año $t+1$ está valorada en $\lambda \in (0,1)$ utiles en el año t, es decir la utilidad de diferentes años se puede sumar descontada a tasa λ.

Finalmente, asuma también que el millonario no posee otras fuentes de ingreso y que firmemente cree que vivirá T años más (incluyendo el actual).

- (5 puntos) Modele el problema que permite decidir al millonario cuanto consumir de su capital cada año con el objetivo de maximizar la utilidad obtenida durante su vida. Explicite las etapas de decisión, el espacio de estados en cada etapa, espacio de acciones en cada estado y etapa y retornos. Luego, enuncie las ecuaciones de Bellman.

  Respuesta:

  - Etapas: $t \in \mathscr{T} := \{0,\ldots,T-1\}$
  - Estados: $b_t \geq 0$, dinero al comienzo del año $t \in \mathscr{T}$
  - Acciones: $x_t \leq b_t$, consumo en el año t

- Retorno: $r_t(b_t, x_t) := \sqrt{x_t}$

Ecuaciones de Belmann:

- $V_t(b_t) := \text{máx}_{\{x_t \in [0,b_t]\}} \left\{ \sqrt{x_t} + \lambda \cdot V_{t+1}((1+r)\cdot(b_t - x_t)) \right\}, \qquad \forall t \in \mathscr{T} : t < T-1, \quad b_t \geq 0$
- $V_{T-1}(b_{T-1}) := \text{máx}_{\{x_T \in [0,b_{T-1}]\}} \left\{ \sqrt{x_{T-1}} \right\}, \qquad b_{T-1} \geq 0$

- (5 puntos) Obtenga una expresión cerrada en función de B, r, λ y T para la política óptima de consumo en cada año y el valor total descontado de la utilidad del millonario al comienzo de este año.

  Respuesta:

  Etapa terminal: Clárament $x^*_{T-1} = b_{T-1}$ y $V_{T-1}(b_{T-1}) = \sqrt{b_{T-1}}$

  Etapa $T-2$: Se debe resolver:

$$V_{T-2}(b_{T-2}) := \max_{\{x_{T-2} \in [0,b_{T-2}]\}} \left\{ \sqrt{x_{T-2}} + \lambda \cdot \sqrt{((1+r)\cdot(b_{T-2} - x_{T-2}))} \right\}, \qquad \forall b_{T-2} \geq 0$$

  Sea $p^2 := \lambda^2 \cdot (1+r)$, el problema equivale a:

$$V_{T-2}(b_{T-2}) := \max_{\{x_{T-2} \in [0,b_{T-2}]\}} \left\{ \sqrt{x_{T-2}} + \sqrt{(p\cdot(b_{T-2} - x_{T-2}))} \right\}, \qquad \forall b_{T-2} \geq 0$$

  Derivando e igualando a cero, el óptimo se da en: $x^*_{T-2} := \frac{b_{T-2}}{1+p}$. Luego $V_{T-2} = \sqrt{b_{T-2}\cdot(1+p)}$

  Etapa $T-3$: Se debe resolver:

$$V_{T-3}(b_{T-3}) := \max_{\{x_{T-3} \in [0,b_{T-3}]\}} \left\{ \sqrt{x_{T-3}} + \sqrt{(p\cdot(b_{T-3} - x_{T-3})\cdot(1+p)))} \right\}, \qquad \forall b_{T-3} \geq 0$$

  Derivando e igualando a cero, el óptimo se da en: $x^*_{T-3} := \frac{b_{T-3}}{1+p+p^2}$. Luego $V_{T-3} = \sqrt{b_{T-3}\cdot(1+p+p^2)}$

  Etapa t:

  Es fácil notar que en general: $x^*_t := \frac{b_t}{S_{T-t-1}}$. Luego $V_t = \sqrt{b_t \cdot (S_{T-t-1})}$, donde: $S_t = \sum_{i=1}^{t} p^i$

  Luego, $V_0(B) = \sqrt{B\cdot(S_{T-1})}$

- (5 puntos) ¿Cómo es la política de consumo si $T \to \infty$ (*i.e.*, el millonario cree que tiene una larga vida por delante) en función de B, r y λ?

  Respuesta:

  Todo depende del valor de $p = \lambda \cdot \sqrt{1+r}$. Si es mayor o igual a 1, entonces el objetivo diverge y la política es nunca gastar un peso, pues $S_\infty \to \infty$, por lo que $X_t \to 0$ para todo $t \geq 0$. En este caso la ganancia de utilidad debido a la inversión supera el decuento temporal λ.

  Si $\lambda \cdot \sqrt{1+r} < 0$, entonces $S_\infty = \frac{1}{1-p}$ y la política de consumo es $x_t = b_t \cdot (1-p)$, es decir consumir una fracción $(1-p)$ del capital cada año.

Pregunta 2: Producción de trotadoras (18 puntos)

Este año aumentó la necesidad de ejercitarse en casa, por lo que creció la demanda de bicicletas estáticas, elípticas y trotadoras. Más aún, por las cuarentenas bajó el volumen de importaciones y los inventarios no se han repuesto (en su mayoría son importados).

Suponga que un empresario nacional decide invertir en una fábrica de trotadoras. Para ello, al comienzo del mes 1 dispone de una nueva línea de producción lista para producir. Una línea de producción "nueva" puede producir 200 trotadoras al mes que se venden inmediatamente en el mercado. Cada una genera una utilidad de \$1000 dólares descontando costos de producción. Sin embargo, la línea "nueva" se desgasta de forma aleatoria para el próximo mes. Con un 23 % de probabilidad la línea de producción seguirá "nueva" el próximo mes, con un 50 % estará "ok", con un 15 % estará "levemente desgastada", con un 10 % estará "gravemente desgastada" y con un 2 % terminará "destruida".

Una línea de producción "ok" también puede producir 200 trotadoras al mes. Sin embargo, con un 25 % de probabilidad seguirá "ok" el próximo mes, con un 50 % estará "levemente desgastada", con un 20 % estará "gravemente desgastada" o terminará "destruida".

Una línea de producción "levemente" desgastada reduce su capacidad de producción a 150 trotadoras mensuales y, además, su estado en el próximo mes sera: "levemente desgastada" con 30 % de probabilidad, "gravemente desgastada" con 50 % o terminará "destruida".

Una línea de producción "gravemente desgastada" solo puede producir 100 trotadoras al mes. Para el siguiente mes se mantendrá "gravemente desgastada" con un 70 % de probabilidad o terminará "destruida".

Finalmente, una línea de producción "destruida" no es capaz de producir.

Cada mes, después de ejecutar la producción y solo si la línea no está "destruida", el empresario puede realizar mantenciones a línea para que comience como "nueva" el próximo mes. Si la línea está "nueva" u "ok" la mantención cuesta \$150 mil dólares, si está "levemente desgastada" cuesta \$200 mil dólares y si está "gravemente desgastada" cuesta \$400 mil dólares. En el caso de una línea "destruida" solo se puede encargar o no una nueva línea que cuesta \$2.5 millones de dólares y que estará lista para operar al comienzo del mes siguiente.

Suponga que se opera a perpetuidad mes a mes, y que el empresario desea maximizar su retorno esperado descontado a una tasa mensual $\lambda = 97\,\%$. Desarrolle las siguiente preguntas:

1. (6 puntos) Modele un MDP que permita decidir si mantener o no la línea de producción en cada uno de sus posibles estados y decidir si comprar o no una línea de producción nueva si es que termina "destruida". Explicite el espacio de estados, espacio de acciones en cada estado, retornos y probabilidades de transición. Luego, enuncie cada ecuación de Bellman.

2. (4 puntos) Evalúe la utilidad de la política "siempre mantener la línea, pero nunca comprarla nueva".

3. (8 puntos) Resuelva el problema a una precisión de dos decimales (como usted desee). Desarrolle claramente su algoritmo de solución y verifique sus condiciones de optimalidad al terminar. Determine la política de decisión óptima. ¿Cuánto vale comenzar con una línea "nueva" al comienzo del primer mes? ¿Vale la pena reemplazar la línea si termina "destruida"?

Respuesta:

a) Primero definimos el espacio de estados $\mathbb{S} = \{0, 1, 2, 3, 4\}$ que representa los estados "Nueva", "OK", "Levemente desgastada", "Gravemente desgastada" y "Destruida", respectivamente.

   Las decisiones en cada etapa son $\mathbb{X}(s) = \{0, 1\}$, donde 0 representa "no mantener" y 1 representa la opción de "mantener".

El retorno inmediato del estado $s \in \mathbb{S}$ cuando se toma la decisión $x \in \mathbb{X}$ es: $r(s,x) = K_s - G_s x$, donde:

$$K = [200000, 200000, 150000, 100000, 0]^T$$
$$G = [150000, 150000, 200000, 400000, 2500000]^T$$

Finalmente, las ecuaciones de Bellman están dadas por:

$$V(s) = \max_{x \in \mathbb{X}(s)} \{K_s - G_s x + \lambda \sum_{j \in \mathbb{S}} p(j|s,x)V(j)\} \qquad \forall s \in \mathbb{S}$$

donde $p(j|s,x)$ es la probabilidad de transición al estado $j \in \mathbb{S}$ dado que el estado actual es $s \in \mathbb{S}$ y se ha tomado la decisión $x \in \mathbb{X}$.

b) Para evaluar esta política debemos resolver:

$$V^d = (I - \lambda P_d)^{-1} r_d$$

donde,

$$P_d = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$r_d = \begin{bmatrix} 50000 \\ 50000 \\ -50000 \\ -300000 \\ 0 \end{bmatrix}$$

Finalmente,

$$V^d = \begin{bmatrix} 1666666,\overline{6} \\ 1666666,\overline{6} \\ 1566666,\overline{6} \\ 1316666,\overline{6} \\ 0 \end{bmatrix}$$

Luego, el valor que se pide es $V_0^d = 1666666,\overline{6}$ ya que la máquina comienza nueva.

c) Resolviendo las ecuaciones de Bellman obtenemos política óptima y valores óptimos:

$$\pi = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$V^\pi = \begin{bmatrix} 2659472,97 \\ 2629688,78 \\ 2529688,78 \\ 2279688,78 \\ 79688,78 \end{bmatrix}$$

Finalmente, como la máquina parte nueva, el valor óptimo es $V_0^\pi = 2659472,97$

Pregunta 3: Conceptos (12 puntos)

Responda cada una de las siguientes preguntas:

1. (2p) En máximo tres líneas: ¿Qué diferencia conceptual existe entre una solución y una política en MDPs?
   Respuesta: Una solución es una serie específica de decisiones para una realización del sistema y una política es una función que recomienda decisiones para cada estado y etapa del sistema.

2. (2p) En máximo tres líneas: ¿Qué diferencia conceptual hay entre tomar decisiones dinámicas reactivas y una proactivas?

   Respuesta: Una política reactiva toma en cuenta el estado del sistema al momento de tomar decisiones, pero no considera información probable del futuro. Un política proactiva si lo hace.

3. (2p) ¿Es la siguiente frase verdadera para un MDP de maximización?: "El valor total esperado al comienzo del horizonte al ejecutar una política de *rollout* sobre otra política base π_B simulada hacia el futuro mediante métodos de Monte Carlo es siempre mayor o igual al valor de la política π_B." Justifique.

   Respuesta: Falso. Al simular, ambas son heurísticas sin garantía.

4. (2p) En máximo tres líneas: ¿Cuál es la principal ventaja al aplicar una política de horizonte rodante determinístico (HR) frente a una política basada en *Value Function Aproximation* (VFA) sobre un problema que sufre de la maldición de la dimensionalidad en el espacio de estados?

   Respuesta: Un HR es online y es capaz de adaptarse mucho mejor a cada estado (desagregado) del sistema. Es recomendable cuando el valor puede poseer mucha variabilidad en función del estado.

5. (2p) En máximo tres líneas: ¿Por qué ADP se concentra en estimar el *value-to-go* en el estado de post-decision $Q_t(y)$? ¿Por qué no estimar el *value-to-go* en el estado de pre-decision dado por $V_t(s)$?

   Respuesta: Para evitar simular. Usamos directamente $Q_t(y(s,x))$ en el paso de optimización.

   Si se estima $V_t(s)$, luego se debe calcular $\mathbb{E}_{s'}[V_t + 1(s')|s,x]$ de alguna forma.

6. (2p) Considere un MDP de maximización y horizonte finito. Suponga un escenario $\omega \in \Omega$, donde Ω es el conjunto de todos los posibles escenarios de la incertidumbre con probabilidad mayor a cero. Un estudiante del curso asegura que el valor de una política heurística cualquiera $V_T^H(\omega)$ sobre esta realización ω es menor o igual que el valor obtenido para ese escenario por la política óptima del MDP dado por $V_T^*(\omega)$. ¿Está en lo correcto? Justifique.

   Respuesta: No lo está. La política óptima del MDP optimiza el valor esperado, pero no cada escenario de la incertidumbre.

Pregunta 4: Superaditividad (5 puntos)

Considere una función superaditiva $g : \mathbb{S} \times \mathbb{X} \to \mathbb{R}$ tal que $\text{mín}_{x\in\mathbb{X}} g(s,x)$ es finito $\forall s \in \mathbb{S}$. Demuestre que $f(s) = \text{mín}\{y \in \text{argmín}_{x\in\mathbb{X}} g(s,x)\}$ es no creciente en $s \in \mathbb{S}$.

Respuesta: Sean $s_1 \leq s_2$ dos valores en $\mathbb{S}$.

- Por definición de $f(s)$ se cumple que $g(s_2, f(s_2)) \leq g(s_2,x)$ para todo $x \in \mathbb{X}$. En particular para un $x' \leq f(s_2)$ se cumple que $g(s_2, f(s_2)) \leq g(s_2,x')$.
- Por superaditivad se cumple además que: $g(s_1,x') + g(s_2, f(s_2)) \geq g(s_1, f(s_2)) + g(s_2,x')$

Combinando ambos, debe ocurrir que $g(s_1,x') \geq g(s_1, f(s_2))$ para todo $x' \leq f(s_2)$. Es decir

$$f(s_2) \leq \underset{x\in\mathbb{X}}{\text{argmín}}\, g(s_1,x) \leq f(s_1).$$

Pregunta 5: La decisión de Carlos (6 puntos)

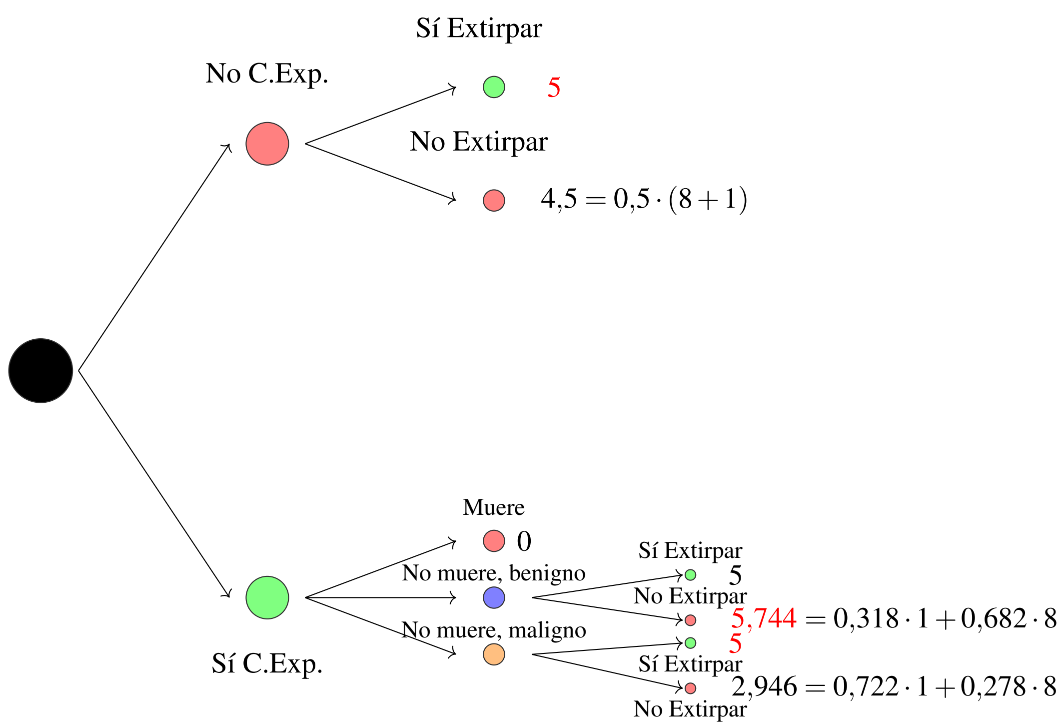
Carlos fue diagnosticado de un tumor cerebral y referido al médico jefe de cirugía del hospital clínico de la universidad. Este tipo de tumor es benigno en un 50 % de los casos y maligno en el otro 50 %. El tiempo de vida restante para Carlos depende de si el tumor es benigno o maligno y de la decisión de sacarlo o no. En la siguiente tabla se detallan estimaciones del tiempo restante de vida de Carlos (en años) en base a los posibles escenarios para este tipo de tumor.

Tipo de tumor	Sacar tumor	Hacer nada
Benigno	5	8
Maligno	5	1

El médico podría hacer una cirugía exploratoria antes de tomar la decisión de si sacar o no el tumor, para tener una mejor evaluación del estado del tumor. La cirugía exploratoria no es infalible y podría realizar una recomendación errada. Se sabe que una cirugía exploratoria sugiere un 75 % del tiempo que un tumor es benigno cuando de verdad es benigno, y sugiere que el tumor es maligno un 65 % del tiempo cuando de verdad es maligno. Esta cirugía exploratoria también tiene algo de riesgo: en un 5 % de los casos los pacientes con características similares a Carlos no sobreviven.

Carlos debe decidir si hace o no la cirugía exploratoria y, luego, si sacar o no el tumor. Naturalmente, Carlos desea vivir el máximo tiempo esperado y quiere tomar la mejor decisión. ¿Cuál es la mejor política de decisión? Justifique mediante un análisis técnico.

Respuesta: Es un árbol de decisión de dos etapas.



Respuesta: Cabe destacar que el primer nivel de decisión es si hacer o no cirugía exploratoria, y el segundo es si extirpar o no el tumor. Las probabilidades y valores de cada nodo se sacaron de la siguiente forma:

- Si no se hace cirugía y se extirpa el tumor, vivirá 5 años más.

- Si no se hace cirugía y no se extirpa el tumor, hay un 50 % de probabilidad de que sea benigno (y viva 8 años) y 50 % de probabilidad de que sea maligno (y viva 1 año).

- Si se hace la cirugía, la probabilidad de morir es 5 % (0 años más).

- Si se hace la cirugía, no muere y sale benigno, y decide extirparlo, vivirá 5 años más (sin importar si era o no benigno).

- Si se hace la cirugía, no muere y sale benigno (CB), y decide no extirparlo, podría ser benigno o no. Así, hay que considerar el caso que sea efectivamente benigno (vivirá 8 años) y el que fuese maligno (vivirá 1 año). Tendremos: $\mathbb{P}(M|CB) = \dfrac{\mathbb{P}(M)\cdot\mathbb{P}(CB|M)}{\mathbb{P}(CB)} = \dfrac{0{,}5\cdot(1-0{,}65)}{\mathbb{P}(CB|B)\cdot 0{,}5+\mathbb{P}(CB|M)\cdot 0{,}5} = \dfrac{0{,}175}{0{,}55} = 0{,}318$.

- Si se hace la cirugía, no muere y sale maligno, y decide extirparlo, vivirá 5 años más (sin importar si era o no maligno).

- Si se hace la cirugía, no muere y sale maligno (CM), y decide no extirparlo, podría ser maligno o no. Así, hay que considerar el caso que sea efectivamente maligno (vivirá 1 año) y el que fuese benigno (vivirá 8 años). Tendremos: $\mathbb{P}(M|CM) = \dfrac{\mathbb{P}(M)\cdot\mathbb{P}(CM|M)}{\mathbb{P}(CM)} = \dfrac{0{,}5\cdot 0{,}65}{\mathbb{P}(CM|M)\cdot 0{,}5+\mathbb{P}(CM|B)\cdot 0{,}5} = \dfrac{0{,}325}{0{,}45} = 0{,}722$.

Con esto, la política óptima será:

- Si no se realizó cirugía exploratoria, debe extirpar ($V^* = 5$).
- Si se realizó cirugía exploratoria y sale maligna, debe extirpar ($V^* = 5$).
- Si se realizó cirugía exploratoria y sale benigna, debe no extirpar ($V^* = 5{,}744$).

Así, el nodo de sí hacer cirugía exploratoria quedará con un valor de:

$$\begin{aligned}\mathbb{P}(CB)\cdot V^*(CB)+\mathbb{P}(CM)\cdot V^*(CM)+\mathbb{P}(Morir)\cdot V^*(Morir) &= (0{,}95\cdot 0{,}55)\cdot 5{,}744+(0{,}95\cdot 0{,}45)\cdot 5+0{,}05\cdot 0\\ &= 5{,}13874\end{aligned}$$

Por lo anterior, la recomendación es sí hacer la cirugía exploratoria ($V^* = 5{,}13874$), y si sale maligno extirpar, y si sale benigno no extirpar.

Pregunta 6: Análisis de proyectos (4 puntos)

Escoja **dos** proyectos de grupos diferentes al suyo y comente brevemente:

- (2p) ¿Qué decisiones ayudó el grupo a planificar? ¿Cuál fue su objetivo? (sea específico)
- (2p) ¿Qué algoritmo de solución utilizaron? (exacto o no, de qué tipo)

Respuesta: Acá cualquier respuesta razonable y que diga la verdad es válida.