Guia 1 de Optimización Dinámica

Mathias A. Klapp
maklapp@ing.puc.cl

September 24, 2021

1 Programación Dinámica Determinística

1.1 Mochila bidimensional (T1-2019, T1-2018)

Considere el siguiente problema de mochila binaria bidimensional

$$\begin{aligned} v^* = \max_{x \in \{0,1\}^n} & \sum_{j=1}^{n} v_j x_j \\ s.t. & \sum_{j=1}^{n} a_{1,j} x_j \leq B_1 \\ & \sum_{j=1}^{n} a_{2,j} x_j \leq B_2, \end{aligned}$$

definido por los parámetros $B_1, B_2 \in \mathbb{Z}_+$ y $a_{1,j}, a_{2,j}, v_j \in \mathbb{Z}_+$ para cada $j \in [n]$.

Formule este problema como un modelo de Programación Dinámica (DP) que permita calcular v^*. Sea explícito definiendo etapas, estados, acciones y sus ecuaciones de optimalidad de Bellman.

Solución:

- Estados: (b_1, b_2)
- Acción: Meter o no el ítem i.
- Etapas: $i = 1, \ldots, n$.

Definiendo $V_i(b_1, b_2)$ como el mayor valor posible con b_1 recursos en la primera dimensión y b_2 en la segunda considerando los primeros i ítemes, las ecuaciones de Bellman quedan:

$$V_1(b_1, b_2) = \begin{cases} v_1 & \text{si } b_1 \geq a_{1,1} \wedge b_2 \geq a_{2,1} \\ 0 & \text{si } b_1 < a_{1,1} \vee b_2 < a_{2,2} \end{cases}$$

$\forall i = 2, \ldots, n$:

$$V_i(b_1, b_2) = \max\{\underbrace{v_i + V_{i-1}(b_1 - a_{1,i}, b_2 - a_{2,i})}_{\text{si } b_1 \geq a_{1,i} \wedge b_2 \geq a_{2,i}} \quad , \quad V_{i-1}(b_1, b_2)\} \qquad \forall b_i \in [B_1], \forall b_2 \in [B_2]$$

Finalmente, el óptimo será $V^* = V_n(B_1, B_2)$.

1.2 Ejemplo donde el principio de recursión no es válido (T1-2019)

Considere el principio de recursión discutido en clases. Busque un ejemplo de problema de optimización en donde el principio de recursión no es aplicable. Debe ser estructuralmente diferente al planteado en clases. Muestre, como falla mediante un ejemplo de juguete que no logre la solución óptima mediante recursión.

Solución: Para este problema, basta con encontrar una función $h(x_1, y)$ que no sea decreciente en y. Por ejemplo, $h(x_1, y) = \sin(c_1 x_1 + y)$. De esta forma, la recursión para muchas etapas quedaría

$$\sin\big(c_1 x_1 + \sin(c_2 x_2 + \ldots)\big)$$

, pero no cumpliría el principio de recursión.

Luego basta explicar como falla en un ejemplo simple (por ejemplo de dos etapas).

1.3 Juego Matemático (T1-2019)

Suponga un juego matemático en donde se debe buscar un conjunto de números enteros positivos $\{n_1, n_2, \ldots, n_p\}$ cuya multiplicación es máxima entre los conjuntos que suman B.

Por ejemplo, una respuesta para $B = 4$ es $\{4\}$, pues los posibles conjuntos que suman 4 son:

- $\{4\}$ con multiplicación 4,
- $\{1, 3\}$ con multiplicación 3,
- $\{2, 2\}$ con multiplicación 4,
- $\{1, 1, 2\}$ con multiplicación 2 y
- $\{1, 1, 1, 1\}$ con multiplicación 1.

De la misma forma la respuesta para $B = 5$ es $\{2, 3\}$, pues los posibles conjuntos que suman 5 son:

- $\{5\}$ con multiplicación 5,
- $\{1, 4\}$ con multiplicación 4,
- $\{2, 3\}$ con multiplicación 6,
- $\{1, 1, 3\}$ con multiplicación 3,
- $\{1, 2, 2\}$ con multiplicación 4,
- $\{1, 1, 1, 2\}$ con multiplicación 2 y
- $\{1, 1, 1, 1, 1\}$ con multiplicación 1.

Formule este problema como un modelo de Programación Dinámica (DP). Sea explícito definiendo etapas, estados, acciones y sus ecuaciones de optimalidad de Bellman. Luego implemente esta solución en el lenguaje de programación que usted desee y resuelva el juego para $B = 10$, $B = 100$, $B = 1000$ y $B = 10000$.

1.4 DP multiplicativo 1 (T1-2020)

Considere el siguiente problema de optimización no lineal entera con parámetro $B \geq 0$:

$$\begin{aligned} z(B) = \max_{\mathbf{x}}\ & x_1 \cdot x_2^2 \cdot x_3^3 \cdot x_4^4 \\ s.t.\ & x_1 + 2x_2 + 4x_3 + 6x_4 \leq B \\ & x_1, x_2, x_3, x_4 \in \mathbb{Z}_+ \end{aligned}$$

- Formule un modelo de Programación Dinámica que calcule $z(B)$. Sea explícito definiendo etapas, estados y ecuaciones de recursión de Bellman para cada estado y etapa. Justifique la validez de su recursión DP.
- Resuelva el problema para $B = 50$ y $B = 100$ mediante programación dinámica.
- Resuelva el problema para $B = 100$ suponiendo adicionalmente que $x_1 \leq 2$.

Solución:

- Primero definimos las etapas de decisión {1,2,3,4} y el espacio de estados $\mathbb{S} = [0, B]$. Las funciones de recursión para cada estado son:

$$C_4(b) = \max_{6x_4 \leq b}(x_4^4), \qquad \forall b \in \mathbb{S} \tag{1}$$
$$C_3(b) = \max_{4x_3 \leq b}(x_3^3 \cdot C_4(b - 4x_3)), \qquad \forall b \in \mathbb{S} \tag{2}$$
$$C_2(b) = \max_{2x_2 \leq b}(x_2^2 \cdot C_3(b - 2x_2)), \qquad \forall b \in \mathbb{S} \tag{3}$$
$$C_1(b) = \max_{x_1 \leq b}(x_1 \cdot C_2(b - x_1)), \qquad \forall b \in \mathbb{S} \tag{4}$$
$$\tag{5}$$

  Esta recursión cumple con el teorema de validez ya que existe una función $h_t(x_t, y) = x_t \cdot y$, con $t = \{1, 2, 3, 4\}$, la cual es continua y no decreciente en y. Finalmente, lo que se busca obtener es $z(B) = C_1(B)$

- A continuación se muestra el valor óptimo y solución para los problemas.

Problema	Valor óptimo	Solución
50	777.600	(6,5,4,3)
100	823.543.000	(10,10,7,7)

- Para $x_1 \leq 2$ la solución óptima es $\mathbf{x} = (2, 12, 8, 7)$ y el valor óptimo es 354.041.856

1.5 DP min–max 1 (T1-2020)

Considere el siguiente problema:

$$\min\{\max\{r(x_1), r(x_2), \ldots, r(x_n)\}\} \tag{6}$$
$$s.t. \sum_{i=1}^{n} x_j \geq B \tag{7}$$
$$x_j \in \mathbb{R}_+, \qquad \forall j \in \{1, ..., n\} \tag{8}$$

, donde $r(x)$ es una función creciente, $B \in \mathbb{R}_+$ y n entero positivo.

- Inspeccione el problema e identifique su valor y solución óptima en función de B y n.

  Solución: La solución óptima es $x_k = B/n$ para todo $k \in \{1, \ldots, n\}$. El valor óptimo es $z = r\left(\frac{B}{n}\right)$. Cualquier solución que tenga holgura en la restricción (7) o diferentes valores en variables se puede mejorar.

- Muestre un esquema válido de recursión en este problema. Plantee ecuaciones de optimalidad de Bellman dividiendo el problema en n etapas, donde la etapa $k \in \{1, \ldots, n\}$ determina la decisión de la variable x_k. Considere que el estado s en cada etapa k corresponde al consumo de la restricción (7) desde la etapa k en adelante.

Solución: Se puede aplicar la siguiente función de recursion continua y creciente en y

$$h(x_k, y) := \max\left(r(x_j), y\right).$$

Luego, se puede resolver definiendo la función de *cost-to-go* $C_k(s)$, donde s es el recurso utilizado en la restricción desde la etapa k hasta la n.

- Costo terminal: $C_n(s) = r(s), \qquad \forall s \geq 0$
- Recursión $\forall k \in \{1, \ldots, n-1\}$:

$$C_k(s) = \min_{x \in [0,s]} \max\left(r(x), C_{k+1}(s-x)\right), \qquad \forall s \geq 0$$

- Inicial: $V^* = C_1(B)$

- En cada etapa $k \in \{1, \ldots, n\}$, derive una expresión para $C_k(s)$ y para una política óptima $x_k^*(s)$ en función del estado s. Verifique que su política es óptima y que logra la solución óptima identificada en la primera parte.

Solución: Calculemos alguna iteraciones:

- El *cost-to-go* terminal es $C_n(s) = r(s)$ y la política óptima $x_n^*(s) = s$.
- El *cost-to-go* en la etapa $k = n-1$ es

$$C_{n-1}(s) = \min_{x \in [0,s]} \max\left(r(x), r(s-x)\right)$$

  y se maximiza cuando $x_n^*(s) = s/2$. Luego: $C_{n-1}(s) = r(s/2)$.
- El *cost-to-go* en la etapa $k = n-2$ es

$$C_{n-2}(s) = \min_{x \in [0,s]} \max\left(r(x), r((s-x)/2)\right)$$

  y se maximiza cuando $x_n^*(s) = s/3$. Luego: $C_{n-1}(s) = r(s/3)$.

Por inducción, se puede demostrar que $x_k^*(s) = \frac{s}{n-k+1}$ y $C_k(s) = r\left(\frac{s}{n-k+1}\right)$.

Luego, $V^* = C_1(B) = r\left(\frac{B}{n}\right)$

Capturemos la solución óptima hacia adelante:

- $x_1^*(B) = \frac{B}{n}$. Luego, $s_2 = B - x_1^* = \frac{(n-1)B}{n}$
- $x_2^*\left(\frac{(n-1)B}{n}\right) = \frac{\frac{(n-1)B}{n}}{n-1} = \frac{B}{n}$. Luego, $s_3 = s_2 - x_2^* = \frac{(n-1)B}{n} - \frac{B}{n} = \frac{(n-2)B}{n}$.
- $x_3^*\left(\frac{(n-2)B}{n}\right) = \frac{\frac{(n-2)B}{n}}{n-2} = \frac{B}{n}$. Luego, $s_4 = s_3 - x_3^* = \frac{(n-2)B}{n} - \frac{B}{n} = \frac{(n-3)B}{n}$.
- $\ldots$

1.6 Optimizando el consumo (Ex-2020)

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.

  Solución:

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

  Ecuaciones de Belmann:

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

- (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.

  Solución:

  Etapa terminal: Cláramente $x^*_{T-1} = b_{T-1}$ 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:

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

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?

  Solución:

  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.

2 Conceptos Básicos de Optim bajo Incertidumbre

2.1 (2018) Min VE es mayor que PIR

Considere una función $f(x,s)$ que depende de una decisión $x \in \mathbb{X}$ y de un estado aleatorio $s \in \mathbb{S}$. El problema que toma la decisión de antemano que maximiza el valor esperado de f es:

$$v_1 = \max_{x \in \mathbb{X}} \mathbb{E}_s[f(x,s)]$$

El problema que calcula el valor esperado de la mejor decisión posible para cada s es:

$$v_2 = \mathbb{E}_s[\max_{x_s \in \mathbb{X}} f(x_s,s)]$$

Demuestre matemáticamente que $v_1 \leq v_2$.

Solución: Se tiene que:

$$f(x,s) \leq \max_{x_s \in \mathbb{X}} f(x_s,s), \qquad \forall x \in \mathbb{X},\ \forall s \in \mathbb{S}.$$

Aplicando esperanza sobre s a ambos lados queda

$$\mathbb{E}_s[f(x,s)] \leq \mathbb{E}_s[\max_{x_s \in \mathbb{X}} f(x_s,s)], \qquad \forall x \in \mathbb{X}.$$

En particular, se debe cumplir para $x = \text{argmax}_{x \in \mathbb{X}} \mathbb{E}_s[f(x,s)]$, por lo que:

$$\max_{x \in \mathbb{X}} \mathbb{E}_s[f(x,s)] \leq \mathbb{E}_s[\max_{x_s \in \mathbb{X}} f(x_s,s)]$$

.

3 Decision Trees

3.1 (2019) La importante decisión del estudiante

Supongamos que usted acaba de licenciarse en un programa de Ingeniería y debe tomar una importante decisión. Una opción es terminar sus estudios hasta obtener el título profesional de Ingeniero Civil. Sin embargo, usted también podría postular al Masters en Ciencias de la Ingeniería en la PUC y graduarse algunos meses más tarde incluyendo un postgrado.

Si concluye los estudios con el título de Ingeniería, entonces hay un 40% de probabilidad de obtener un trabajo normal que posee un VAN (valor actual neto) de sus ingresos futuros igual a US$1,5M. También, existe un 50% de probabilidad de obtener un trabajo existoso el cual posee un VAN (valor actual neto) de sus ingresos futuros igual a US$3M. Finalmente, existe una probabilidad del 10% de que obtenga un trabajo mediocre, en cuyo caso el VAN de los ingresos de los próximos 10 años sería de US$800 mil.

Si postula al post-grado, el profesor Klapp le haría una carta de referencia y tendría un 90% de probabilidad de quedar aceptado. Si no es aceptado, entonces simplemente termina sus estudios con el título profesional. Si es aceptado, usted tendrá que pagar el costo adicional del programa (incluyendo el tiempo en que deja de percibir salario). Para ello, usted posee dos alternativas de financiamiento. La primera es pedir un préstamo cuyo costo equivale a un VAN de US$250mil.

La segunda es pedirle un préstamo a su tío malvado quien le ha ofrecido pagar el costo de su postgrado, pero a cambio del 10% del total de sus ingresos futuros.

Un graduado del Masters debiese mejorar sus posibilidades de trabajo e ingresos futuros. Aproximadamente un 10% de sus graduados se convierten en emprendedores exitosos, con un VAN de sus ingresos de los próximos 10 años de US$5M. Un 60% obtiene un trabajo exitoso de empleado con un VAN de US$3M y el 30% restante posee un trabajo normal con un VAN de US$1,5M.

- (10p) En el contexto de esta pregunta, asumiendo que usted maximiza valor esperado y que es neutral al riesgo: ¿Postularía o no al Masters en Ciencias de la Ingeniería? ¿Cómo lo financiaría? Explique su razonamiento.

  Solución: falta pauta

- (4p) ¿En qué rango de porcentajes le aceptaría la oferta a su tío malvado?

  Solución: falta pauta

3.2 La decisión de Carlos (EX-2020)

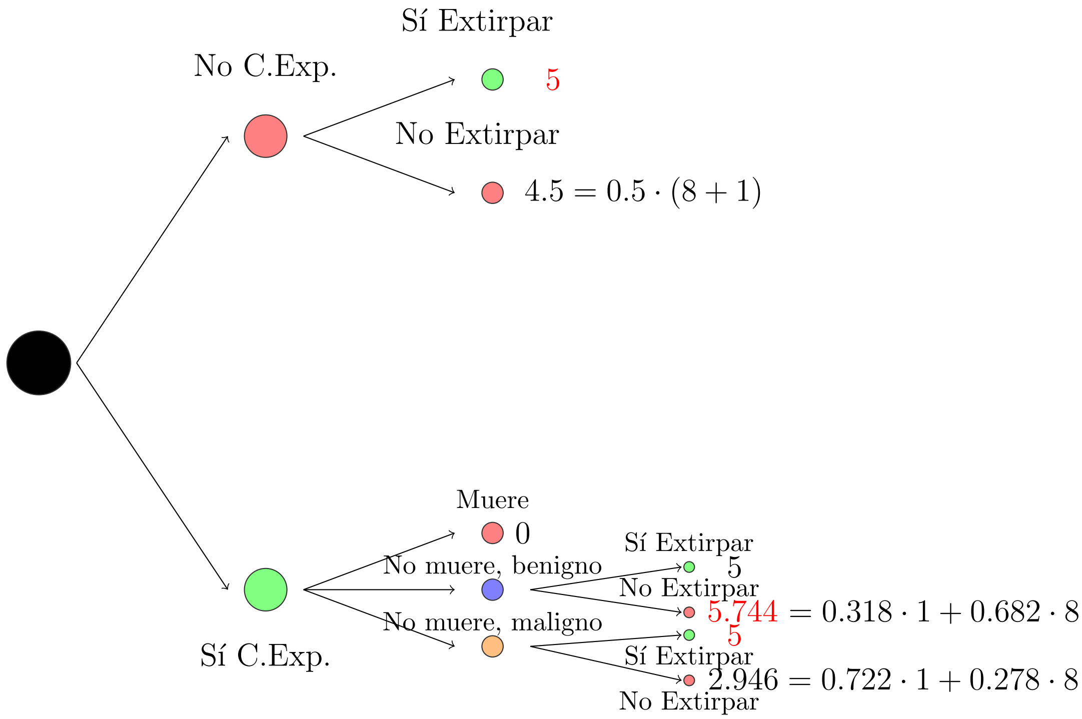
Carlos fue diagnosticado de un tumor cerebral y referido al médico jefe de cirugía del hospital clínico de la universidad. Este tipo de tumor es benigno en un 50% de los casos y maligno en el otro 50%. El tiempo de vida restante para Carlos depende de si el tumor es benigno o maligno y de la decisión 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.

Solución: Es un árbol de decisión de dos etapas.



Solución: 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: $P(M|CB) = \dfrac{P(M) \cdot P(CB|M)}{P(CB)} = \dfrac{0.5 \cdot (1 - 0.65)}{P(CB|B) \cdot 0.5 + P(CB|M) \cdot 0.5} = \dfrac{0.175}{0.55} = 0.318$.
- 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: $P(M|CM) = \dfrac{P(M) \cdot P(CM|M)}{P(CM)} = \dfrac{0.5 \cdot 0.65}{P(CM|M) \cdot 0.5 + P(CM|B) \cdot 0.5} = \dfrac{0.325}{0.45} = 0.722$.

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:

$$P(CB) \cdot V^*(CB) + P(CM) \cdot V^*(CM) + P(Morir) \cdot V^*(Morir) = (0.95 \cdot 0.55) \cdot 5.744 + (0.95 \cdot 0.45) \cdot 5 + 0.0$$
$$= 5.13$$

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.

4 MDP finito

4.1 (I1-2018, I1-2019) Modelar un MDP: Banderas Chilenas

Este año habrá 5 días festivos para fiestas patrias. Suponga que usted desea emprender y comprará Q banderas chilenas antes del inicio de las fiestas que intentará vender durante los 5 días en una de las ramadas más importantes de la capital. Suponga que no se comprar más banderas después del inicio de las fiestas.

La organización de la ramada le cobrará $\$C$ diarios por instalarse con su puesto de venta; esta decisión la podrá tomar al comienzo de cada día de venta. Usted estima que la demanda diaria por banderas chilenas será una variable aleatoria D_t para cada $t \in \{1, \dots, 5\}$, cuya función de probabilidad discreta está dada por $f_t(d, p) := \mathbb{P}(D_t = d | p_t = p) \in [0, 1]$ conocida para todo $d \in \mathbb{D}$ y condicionada al precio de venta por bandera $p \in \{\underline{p}, \dots, \bar{p}\}$ a fijar por usted en cada día t dentro de un rango.

Considere que cualquier demanda no satisfecha en los 5 días será perdida, que al quedar sin stock se acaba su negocio y que al terminar las fiestas patrias las banderas dejan de tener valor.

Usted está interesado en estimar $g(Q)$, la ganancia máxima esperada que obtendría al comprar Q banderas al comienzo de fiestas patrias.

1. Diseñe un modelo con el que pueda calcular $g(Q)$ como un Proceso de Decisión Markoviana (MDP). Defina etapas, espacio de estados, espacio de decisiones, valor inmediato por etapa y obtenga la función de probabilidad de transición entre estados para este modelo.
2. Escriba las ecuaciones de Bellman que habría que resolver para calcular $g(Q)$.
3. Si cada bandera cuesta $\$v$ en ser producida: ¿Cómo podría estimar la compra óptima Q^* que debería hacer antes de las fiestas patrias?

Solución:

1. Este es un modelo:
   - Etapas: {1,2,3,4,5,6}. La etapa 6 es la terminal.

- Estado: $q \in \mathbb{S} := \{0, 1, \ldots, Q\}$, inventario de banderas al comienzo del día $t \in \{1, \ldots, 6\}$.
- Decisiones: $(z_t, p_t) \in \mathbb{X}$, donde $z_t \in \{0, 1\}$ (si vender o no) y p_t precio de venta. Luego $\mathbb{X} = \{(1, p), \forall p \in \{\underline{p}, \ldots, \bar{p}\}\} \cup \{(0, \cdot)\}$
- Valor inmediato: $r_t(q, z, p)$
  - $r_t(q, 0, \cdot) = 0, \qquad q \in \mathbb{S}$
  - $r_t(q, 1, p) = \min(q, \sum_{d \in \mathbb{D}} \{d \cdot f_t(d, p)\}) \cdot p - C, \qquad q \in \mathbb{S}, p \in \{\underline{p}, \ldots, \bar{p}\}$
- Probabilidad de transición: $w_t(q'|q, z, p)$
  - $w_t(q'|q, 0, \cdot) = \begin{cases} 1 & q' = q \\ 0 & q' \neq q \end{cases}$
  - $w_t(q'|q, 1, p) = \begin{cases} 0 & q' > q \\ f_t(q - q', p) & 0 < q' < q \\ \sum_{d \in \mathbb{D}: d \geq q} f_t(q, p) & q' = 0 \end{cases}$

2. Periodo terminal:

$$V_6(q) = 0, \qquad q \in \{0, \ldots, Q\}$$

   Recursión para $t = \{5, 4, 3, 2, 1\}$ y $q \in \{0, \ldots, Q\}$:

$$V_t(q) = \max \left\{ V_{t+1}(q), \max_{p \in \{\underline{p}, \ldots, \bar{p}\}} \left(r_t(q, 1, p) + \sum_{d \in \mathbb{D}: d < q} f(d, q) V_{t+1}(q - d) + \sum_{d \in \mathbb{D}: d \geq q} f(d, q) \cdot V_{t+1}(0) \right) \right\}$$

   Luego $g(Q) = V_1(Q)$

3. Habría que resolver $Q^* = \text{argmax}_{Q \in \mathbb{Z}} \{g(Q) - v \cdot Q\}$.

4.2 (I1-2018, T1-2019)Resolver un MDP de riesgo

Considere un Proceso de Decisión Markoviana (MDP) con dos estados $\mathbb{S} = \{0, 1\}$ y cuatro etapas $t = \{1, 2, 3, 4\}$, donde las primeras tres etapas son de decisión y la cuarta es una etapa terminal con valores terminales nulos. Las decisiones en cada etapa $t = 1, 2, 3$ son las siguientes:

- En el estado 0 se puede tomar la decisión L (libre de riesgo) o la decisión R (riesgosa). Si se toma la decisión libre de riesgo se genera una ganancia inmediata de 5 y el sistema queda en el mismo estado para la siguiente etapa. Si se toma la decisión riesgosa, con probabilidad 25% el sistema termina en el estado 1 y genera una ganancia inmediata de 32, de lo contrario se queda en el mismo estado sin generar ganancia.

- En el estado 1 no hay decisiones y siempre se salta al estado 0 generando una pérdida de 12.

Resuelva este MDP indicando la política óptima de decisión en el estado 0 para cada etapa y la ganancia máxima esperada V^* al comenzar en el estado 0 en $t = 1$.

Solución:

- Etapa terminal $t = 4$: Por enunciado $V_4(0) = V_4(1) = 0$
- Etapa $t = 3$: $V_3(0) = \max\{5+V_4(0); \frac{32+V_4(1)}{4}+\frac{3V_4(0)}{4}\} = \max\{5; 8\} = 8$, $V_3(1) = -12+V_4(0) = -12$. Luego:

$$x_3^*(0) = R$$

- Etapa $t = 2$: $V_2(0) = \max\{5 + V_3(0); \frac{32+V_3(1)}{4} + \frac{3V_3(0)}{4}\} = \max\{5 + 8; 8 - 12/4 + 3 * 8/4\} = \max\{13; 11\} = 13$, $V_2(1) = -12 + V_3(0) = -4$. Luego:

$$x_2^*(0) = L$$

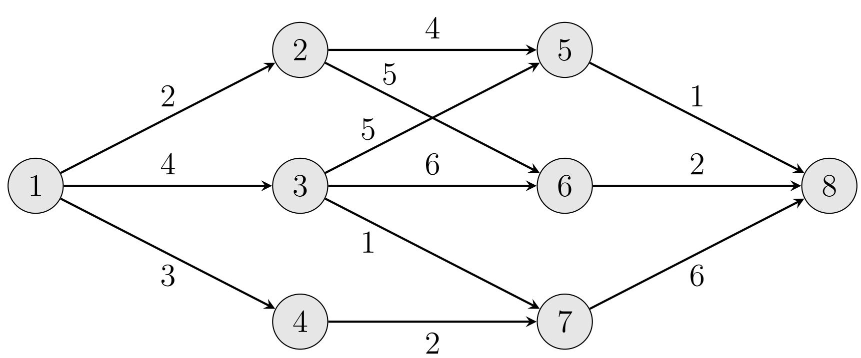
- Etapa $t = 1$: $V_1(0) = \max\{5 + V_2(0); \frac{32+V_2(1)}{4} + \frac{3V_2(0)}{4}\} = \max\{5 + 13; 8 - 4/4 + 3 * 13/4\} = \max\{18; 16,75\} = 16,75$, $V_1(1) = -12 + V_2(0) = 1$. Luego:

$$x_1^*(0) = L$$

La ganancia máxima esperada sería $V^* = 20,5$ y la política óptima sería $\pi^* = \{L, L, R\}$.

4.3 (T 2019, 2018) Rutas mínimas estocásticas

Suponga que se le acerca un ebrio a la salida de un bar y le dice que anda perdido. Al ver el estado del sujeto, usted decide ayudarlo a llegar a casa. Considere el siguiente grafo que modela las diferentes intersecciones (nodos) y calles (arcos) que puede tomar esta persona para caminar desde el bar (ubicado en 1) hasta su casa (ubicada en 8). En cada arco se detalla tiempo de caminata en minutos.



1. (4 puntos) Formule un programa dinámico (DP) que pueda recomendar la ruta (secuencia de nodos) de menor tiempo de caminata desde 1 a 8. Indique sus estados, acciones y correspondientes ecuaciones de Bellman.

   Solución:

   - Estados: $\mathbb{S} = \{1, 2, \ldots, 8\}$.
   - Acciones: $\mathbb{X} = \delta_i^+$, es decir, los arcos salientes de i.

Las ecuaciones de Bellman quedan de la forma $C_i = \min_{(i,j)\in\delta_i^+}\{c_{ij} + C_j\}$. Es decir,

$$\begin{aligned}
C_1 &= \min\{2 + C_2, 4 + C_3, 3 + C_4\}\\
C_2 &= \min\{4 + C_5, 5 + C_6\}\\
C_3 &= \min\{5 + C_5, 6 + C_6, 1 + C_7\}\\
C_4 &= 2 + C_7 = 8\\
C_5 &= 1\\
C_6 &= 2\\
C_7 &= 6
\end{aligned}$$

2. (4 puntos) Ayude al ebrio y recomiéndele la ruta de menor tiempo de caminata.

   Solución:

$$\begin{aligned}
&d_3 \in \text{argmin}\{6, 8, 7\} = (3, 5) \quad \to C_3 = 6\\
&d_2 \in \text{argmin}\{5, 7\} = (2, 5) \quad \to C_2 = 5\\
&d_1 \in \text{argmin}\{2 + 5, 4 + 6, 3 + 8\} = (1, 2) \quad \to C_1 = 7
\end{aligned}$$

   Así, la ruta es $1 - 2 - 5 - 8$.

3. (7 puntos) Ahora suponga que el ebrio, al no estar en su sano juicio, es incapaz de seguir sus instrucciones con certeza. Específicamente, si el ebrio está en un nodo dado y usted le recomienda tomar un arco saliente desde ese nodo, entonces esta persona tomará ese camino con las siguientes probabilidades: Si el nodo tiene tres arcos salientes, entonces seguirá su recomendación con probabilidad 60% y tomará un arco de los restantes con probabilidad 20%. Si el nodo tiene dos arcos salientes, entonces el ebrio tomará su recomendación con probabilidad 60% y el otro con 40%. Finalmente, si hay solo un arco saliente, entonces el borracho seguirá ese camino con certeza. Considerando este nuevo antecedente, recomiende al ebrio un arco para tomar en cada nodo que pueda visitar con el objetivo de minimizar su tiempo esperado de caminata a casa.

   Solución:

$$C_8 = 0 \quad C_7 = 6 \quad C_6 = 2 \quad C_5 = 1 \quad C_4 = 2 + C_7 = 8$$

$$C_3 = \min\{0.6(6) + 0.2(8) + 0.2(7), 0.6(8) + 0.2(6) + 0.2(7), 0.6(7) + 0.2(8) + 0.2(6)\} = 6.6$$

$$C_2 = \min\{0.6(5) + 0.4(7), 0.6(7) + 0.4(5)\} = 5.8$$

$$C_1 = \min\{0.6(7.8)+0.2(10.6)+0.2(11), 0.6(10.6)+0.2(11)+0.2(7.8), 0.6(11)+0.2(7.8)+0.2(10.6)\} = 9$$

   Así, $C^* = 9$. La política será:

$$d_1 = (1, 2) \quad d_2 = (2, 5) \quad d_3 = (3, 5) \quad d_4 = (4, 7)$$

$$d_5 = (5, 8) \quad d_6 = (6, 8) \quad d_7 = (7, 8)$$

4. (1 punto) En promedio: ¿Cuál es el tiempo adicional de caminata para el ebrio por no ser capaz de seguir sus instrucciones?

   Solución: 9 - 7 = 2

5 Secretary Problem

5.1 Estacionamiento del Mall (T1-2018)

Considere que una persona realiza un viaje al Mall en auto este fin de semana para ir de compras. Para ello debe encontrar un espacio en el estacionamiento del mall, estacionar y caminar hacia la entrada de la tienda. Suponga que la persona no está en buen estado físico y que estacionar más cerca de la entrada les produce una inmensa satisfacción, pues caminaría menos.

Específicamente, considere la siguiente figura:

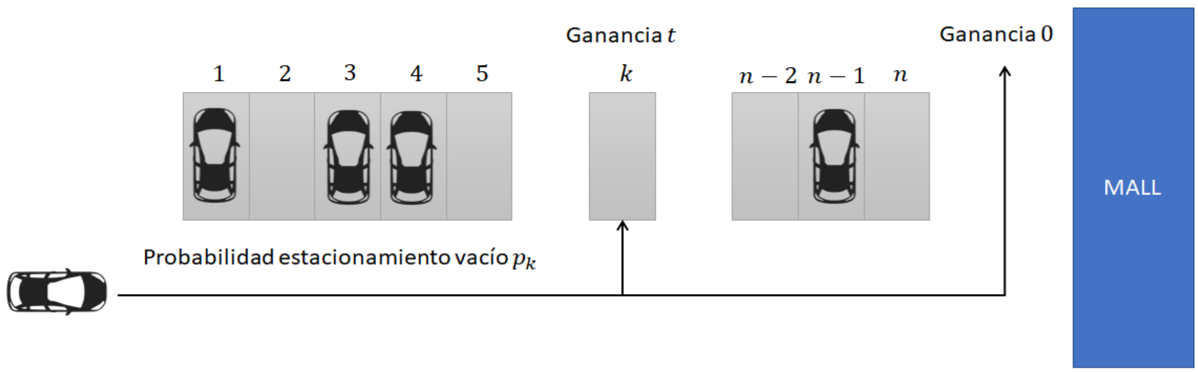


Figure 1: Esquema del estacionamiento

El estacionamiento del Mall se puede modelar como un pasillo que posee n lotes en línea para estacionar; el lote 1 es el más alejado del Mall y el lote n es el que está más cerca. La persona entra manejando al estacionamiento por el lote número 1 y avanza lote por lote. Si se encuentra frente al lote k y este se encuentra vacío entonces puede decidir estacionar en él y obtener una ganancia igual a k (más alta mientras más cerca del Mall), pero también puede decidir avanzar y probar suerte con el siguiente espacio. Si el lote está ocupado, está obligado a avanzar (no puede poner reversa). Si termina al final del pasillo se va para su casa sin haber podido estacionar y sin ganancia. Considere que la persona no puede ver los estacionamientos que todavía no ha visitado. Es decir, al estar frente al estacionamiento k no posee información de los estacionamientos $k+1, k+2, \ldots, n$. Sin embargo, sabe que la probabilidad de que el estacionamiento k esté vacío es igual a $p_k \in [0, 1]$.

- 3p Formule el problema que maximiza la ganancia esperada como un proceso de decisión markoviana.
- 3p Demuestre que existe una política de umbral τ en la cual usted dejará pasar la opción de estacionar para todo lote $k \leq \tau$. Para ello, muestre que si es que es óptimo no estacionar cuando el lote k está disponible, entonces también es óptimo no estacionar cuando el lote $t < k$ está disponible.
- 3p Demuestre que si $N \geq 2$ y $\prod_{k=2}^{n} (1 - p_k) \leq \frac{1}{2}$, entonces nunca es óptimo estacionar en el primer lote si está disponible,
- 3p Sea $\tau < n$ el último lote de estacionamiento en el cual es óptimo esperar. Calcule el *value-to-go* si es que el lote τ no está disponible en función de p_k, τ y n.
- 4p Si $p_k = p$ (independiente de k), derive un algoritmo para obtener τ en función del resultado anterior. Luego compute τ para $n = 100$ y para los casos: $p = 20\%, 40\%$ y 60%. Discuta su resultado. Si lo desea, puede apoyarse de un software computacional.

- 4p Si $p_k = 1/k$, derive un algoritmo para obtener τ. Luego compute τ para $n = 100$.