ReinforcedLearningClass

Norah Jones

2024-09-09

Table of contents

Preface 3

1 Tarea 1 (Fecha de Entrega 20 Septiembre 2024 12:00:00) 4

2 Tarea 2 14

References 15

Preface

This is a Quarto book.

To learn more about Quarto books visit https://quarto.org/docs/books.

1 Tarea 1 (Fecha de Entrega 20 Septiembre 2024 12:00:00)

Exercise 1.1. Read (Sec 1.1, pp 1-2 Sutton and Barto 2018) and answer the following. Explain why Reinforcement Learning differs for supervised and unsupervised.

Exercise 1.2. See the first Brunton's youtube about Reinforced Learning. Then acordingly to its presentation explain what is the meaning of the following expression.

$$V_\pi(s) = E\left(\sum_t \gamma^t r_t \mid s_0 = s\right)$$

La expresión presentada en el video Reinforcement Learning.

$$V_\pi(s) = E\left[\sum_t \gamma^t r_t \middle| s_0 = s\right]$$

hace referencia a la función de valor del problema de optimización representada por la recompensa esperada dado la politica π y el estado inicial s. Aquí γ es el factor de descuento y r_t es la recompensa por etapa t.

Exercise 1.3. Form (see Sutton and Barto 2018, sec. 1.7) obtain a time line pear year from 1950 to 2012.

Exercise 1.4. Consider the following comsuption-saving problem with dynamics

$$x_{k+1} = (1+r)(x_k - a_k), k = 0, 1, \dots, N-1$$

and utility function

$$\beta^N (x_N)^{1-\gamma} + \sum_{k=0}^{N-1} \beta^k (a_k)^{1-\gamma}.$$

Show that the value functions of the DP alghorithm the form

$$J_k(x) = A_k k\beta^k x^{1-\gamma},$$

where $A_N = 1$ and for $k = N-1, \dots, 0$,

$$A_k = \left[1 + \left((1+r)\beta A_{k+1}\right)^{1/\gamma}\right]^{\gamma}.$$

Show also that the optimal policies are $h_k(x) = A^{-1/\gamma}x$, for $k = N-1, \dots, 0$.

Considerando J_N como sigue

$$J_N(x) = \beta^N x^{1-\gamma} K_N,$$

con $K_N = 1$ bajo la hipótesis de que

$$c_k(x,a) = \beta^k a^{1-\gamma}$$

calculamos J_{N-1}.

$$\begin{aligned} J_{N-1}(x) &= \max_{a\in A(x)} \left\{c_{N-1}(x,a) + J_N\left((1+i)(x-a)\right)\right\} \\ &= \max_{a\in A(x)} \left\{\beta^{N-1}a^{1-\gamma} + \beta^N \left((1+i)(x-a)\right)^{1-\gamma}\right\} \end{aligned}$$

Definimos el argumento como una función q.

$$\begin{aligned} q(x,a) &= \beta^{N-1}a^{1-\gamma} + \beta^N \left((1+i)(x-a)\right)^{1-\gamma} \\ &= C_1 a^{1-\gamma} + C_2 \left(x-a\right)^{1-\gamma}, \end{aligned}$$

donde $C_1 = \beta^{N-1}$ y $C_2 = \beta^N(1+i)^{1-\gamma}K_N$. Como q es continua en (x,a). Podemos calcular el máximo mediante el gradiente.

$$\partial_a q = C_1\left(1-\gamma\right)a^{-\gamma} - C_2(1-\gamma)\left(x-a\right)^{-\gamma}.$$

Igualando, $\partial_a q = 0$.

$$\begin{aligned}
C_1 a^{-\gamma} &= C_2 \left(x-a\right)^{-\gamma} \\
\frac{C_1}{C_2} &= \left(\frac{x-a}{a}\right)^{-\gamma} \\
\left(\frac{C_1}{C_2}\right)^{-\frac{1}{\gamma}} &= \frac{x}{a} - 1 \\
\left(\frac{C_1}{C_2}\right)^{-\frac{1}{\gamma}} + 1 &= \frac{x}{a} \\
a &= \frac{x}{\left(\frac{C_1}{C_2}\right)^{-\frac{1}{\gamma}} + 1}
\end{aligned}$$

Finalmente

$$a = h(x) = \frac{x}{(\beta(1+i)^{1-\gamma})^{\frac{1}{\gamma}} + 1}$$

Definiendo $\eta = (\beta(1+i)^{1-\gamma})^{\frac{1}{\gamma}} + 1$, $\eta - 1 = (\beta(1+i)^{1-\gamma})^{\frac{1}{\gamma}}$

entonces

$$h(x) = \frac{x}{\eta},$$

$$\begin{aligned}
J_{N-1}(x) &= \beta^{N-1}\left(\frac{x}{\eta}\right)^{1-\gamma} + \beta^{N}\left((1+i)\left(x - \frac{x}{\eta}\right)\right)^{1-\gamma} \\
&= \beta^{N-1} x^{1-\gamma}\left(\eta^{\gamma-1} + \beta\left(1+i\right)^{1-\gamma}\left(\frac{\eta-1}{\eta}\right)^{1-\gamma}\right) \\
&= \beta^{N-1} x^{1-\gamma}\eta^{\gamma-1}\left(1 + \beta\left(1+i\right)^{1-\gamma}\left(\eta-1\right)^{1-\gamma}\right) \\
&= \beta^{N-1} x^{1-\gamma}\eta^{\gamma-1}\left(1 + \beta\left(1+i\right)^{1-\gamma}\left(\eta-1\right)^{1-\gamma}\right) \\
&= \beta^{N-1} x^{1-\gamma}\eta^{\gamma-1}\left(1 + \beta\left(1+i\right)^{1-\gamma}\left((\beta(1+i)^{1-\gamma})^{\frac{1}{\gamma}}\right)^{1-\gamma}\right) \\
&= \beta^{N-1} x^{1-\gamma}\eta^{\gamma-1}\left(1 + \beta\left(1+i\right)^{1-\gamma}(\beta(1+i)^{1-\gamma})^{\frac{1}{\gamma}-1}\right) \\
&= \beta^{N-1} x^{1-\gamma}\eta^{\gamma-1}\left(1 + \beta^{\frac{1}{\gamma}}(1+i)^{(1-\gamma)\left(\frac{1}{\gamma}-1\right)+1-\gamma}\right) \\
&= \beta^{N-1} x^{1-\gamma}\eta^{\gamma-1}\left(1 + \beta^{\frac{1}{\gamma}}(1+i)^{\left(\frac{1}{\gamma}-1\right)}\right) \\
&= \beta^{N-1} x^{1-\gamma}\eta^{\gamma},
\end{aligned}$$

Entonces

$$K_{N-1} = \eta^{\gamma}, h_{k-1}(x) = \frac{x}{\left(K_{N-1}\right)^{1/\gamma}}$$

Ahora calculamos J_{N-2}

$$\begin{aligned} J_{N-2}(x) &= \max_{a \in A(x)} \left\{ \beta^{N-2} a^{1-\gamma} + \beta^{N-1} \left[(1+i)(x-a) \right]^{1-\gamma} \eta^{\gamma} \right\} \\ &= \max_{a \in A(x)} \left\{ q(x,a) \right\}, \end{aligned}$$

donde

$$q(x,a) = C_1 a^{1-\gamma} + C_2 (x-a)^{1-\gamma},$$

con $C_1 = \beta^{N-2}$ y $C_2 = \beta^{N-1} (1+i)^{1-\gamma} K_{N-1}$. Obteniendo, por recursividad

$$\begin{aligned} h_{N-2} &= \frac{x}{\left(\frac{C_1}{C_2}\right)^{-\frac{1}{\gamma}} + 1} \\ &= \frac{x}{\left(\frac{1}{\beta (1+i)^{1-\gamma} K_{N-1}}\right)^{-\frac{1}{\gamma}} + 1} \\ &= \frac{x}{\left(\beta (1+i)^{1-\gamma} K_{N-1}\right)^{\frac{1}{\gamma}} + 1} \end{aligned}$$

Entonces, sea

$$\eta' = \left(\beta (1+i)^{1-\gamma} K_{N-1}\right)^{\frac{1}{\gamma}} + 1.$$

Repitiendo, el caso anterior, tenemos que

$$
\begin{aligned}
J_{N-2}(x) &= \beta^{N-2}x^{1-\gamma}\eta_{,}^{\gamma-1}\left(1+K_{N-1}\beta\left(1+i\right)^{1-\gamma}\left(\left(\beta(1+i)^{1-\gamma}K_{N-1}\right)^{\frac{1}{\gamma}}\right)^{1-\gamma}\right)\\
&= \beta^{N-2}x^{1-\gamma}\eta_{,}^{\gamma-1}\left(1+K_{N-1}\beta\left(1+i\right)^{1-\gamma}\left(\left(\beta(1+i)^{1-\gamma}K_{N-1}\right)^{\frac{1}{\gamma}}\right)^{1-\gamma}\right)\\
&= \beta^{N-2}x^{1-\gamma}\eta_{,}^{\gamma-1}\left(1+K_{N-1}\beta\left(1+i\right)^{1-\gamma}\left(\beta(1+i)^{1-\gamma}K_{N-1}\right)^{\frac{1}{\gamma}-1}\right)\\
&= \beta^{N-2}x^{1-\gamma}\eta_{,}^{\gamma-1}\left(1+K_{N-1}\beta\left(1+i\right)^{1-\gamma}\left(1+i\right)^{(1-\gamma)\left(\frac{1}{\gamma}-1\right)}K_{N-1}^{\frac{1}{\gamma}-1}\right)\\
&= \beta^{N-2}x^{1-\gamma}\eta_{,}^{\gamma-1}\left(1+K_{N-1}\beta^{1/\gamma}\left(1+i\right)^{\frac{1}{\gamma}-1}K_{N-1}^{\frac{1}{\gamma}-1}\right)\\
&= \beta^{N-2}x^{1-\gamma}\eta_{,}^{\gamma-1}\left(1+\beta^{1/\gamma}\left(1+i\right)^{\frac{1}{\gamma}-1}K_{N-1}^{\frac{1}{\gamma}}\right)\\
&= \beta^{N-2}x^{1-\gamma}\eta'^{\gamma},
\end{aligned}
$$

entonces

$$K_{N-2} = \eta'^{\gamma},$$

y

$$h_{N-2} = \frac{x}{K_{N-2}^{1/\gamma}}$$

Por lo tanto, tenemos que

$$K_n = \left(\beta\left(1+i\right)^{1-\gamma}K_{n+1}\right)^{\frac{1}{\gamma}} + 1, n = 0, 1, 2, ..., N,$$

con $K_N = 1$.

Obteniendo así

$$
\begin{aligned}
J_n(x) &= \beta^n x^{1-\gamma}K_n\\
h_n(x) &= \frac{x}{K_n^{1/\gamma}}
\end{aligned}
$$

Exercise 1.5. Consider now the infinite-horizon version of the above compsumption problem.

1. Write down the associated Bellman equation.
2. Argue why a solution to the Bellman equation should be the form

$$v(x) = cx^{1-\gamma},$$

where c is constant. Find the constant c and the stationary optimal policy.

Para el caso infinito. Estamos considerando

$$c\left(x,a\right)=a^{1-\gamma}$$

Entonces

$$\nu\left(x\right)=\max_{a\in A(x)}\left\{a^{1-\gamma}+\beta\nu\left(\left(1+i\right)\left(x-a\right)\right)\right\},$$

considerando $\nu\left(x\right)=cx^{1-\gamma}$. Entonces

$$\nu\left(x\right)=\max_{a\in A(x)}\left\{a^{1-\gamma}+\beta c\left[\left(1+i\right)\left(x-a\right)\right]^{1-\gamma}\right\},$$

definimos

$$q\left(x,a\right)=a^{1-\gamma}+\beta c\left[\left(1+i\right)\left(x-a\right)\right]^{1-\gamma},$$

entonces

$$\partial_a q=\left(1-\gamma\right)a^{-\gamma}+\beta c\left(1-\gamma\right)\left(1+i\right)^{1-\gamma}\left(-1\right)\left(x-a\right)^{-\gamma}.$$

Si $\partial_a q=0$. Entonces

$$\begin{aligned}
a^{-\gamma} &= \beta c\left(1+i\right)^{1-\gamma}\left(x-a\right)^{-\gamma}\\
\left(\beta c\left(1+i\right)^{1-\gamma}\right)^{-1} &= \left(\frac{x-a}{a}\right)^{-\gamma}\\
\beta^{-1}c^{-1}\left(1+i\right)^{\gamma-1} &= \left(\frac{x}{a}-1\right)^{-\gamma}\\
\left[\beta^{-1}c^{-1}\left(1+i\right)^{\gamma-1}\right]^{-\frac{1}{\gamma}}+1 &= \frac{x}{a}
\end{aligned}$$

Por lo tanto

$$\begin{aligned}
a &= \frac{x}{\left[\beta^{-1}c^{-1}\left(1+i\right)^{\gamma-1}\right]^{-\frac{1}{\gamma}}+1}\\
&= \frac{x}{\left[\beta c\left(1+i\right)^{1-\gamma}\right]^{\frac{1}{\gamma}}+1}
\end{aligned}$$

Ahora remplazamos en q

$$\begin{aligned}
\nu(x) &= \left(\frac{x}{\left[\beta c(1+i)^{1-\gamma}\right]^{\frac{1}{\gamma}}+1}\right)^{1-\gamma} + \beta c\left[(1+i)\left(x-\frac{x}{\left[\beta c(1+i)^{1-\gamma}\right]^{\frac{1}{\gamma}}+1}\right)\right]^{1-\gamma} \\
&= x^{1-\gamma}\left(\frac{1}{\left[\beta c(1+i)^{1-\gamma}\right]^{\frac{1}{\gamma}}+1}\right) + x^{1-\gamma}(1+i)^{1-\gamma}\beta c\left(1-\frac{1}{\left[\beta c(1+i)^{1-\gamma}\right]^{\frac{1}{\gamma}}+1}\right)^{1-\gamma} \\
&= x^{1-\gamma}\left[\left(\frac{1}{\left[\beta c(1+i)^{1-\gamma}\right]^{\frac{1}{\gamma}}+1}\right) + (1+i)^{1-\gamma}\beta c\left(1-\frac{1}{\left[\beta c(1+i)^{1-\gamma}\right]^{\frac{1}{\gamma}}+1}\right)^{1-\gamma}\right].
\end{aligned}$$

Entonces

$$\begin{aligned}
c &= \left(\frac{1}{\left[\beta c(1+i)^{1-\gamma}\right]^{\frac{1}{\gamma}}+1}\right)^{1-\gamma} + (1+i)^{1-\gamma}\beta c\left(1-\frac{1}{\left[\beta c(1+i)^{1-\gamma}\right]^{\frac{1}{\gamma}}+1}\right)^{1-\gamma} \\
&= \left(\frac{1}{\left[\beta c(1+i)^{1-\gamma}\right]^{\frac{1}{\gamma}}+1}\right)^{1-\gamma} + (1+i)^{1-\gamma}\beta c\left(\frac{\left[\beta c(1+i)^{1-\gamma}\right]^{\frac{1}{\gamma}}}{\left[\beta c(1+i)^{1-\gamma}\right]^{\frac{1}{\gamma}}+1}\right)^{1-\gamma} \\
&= \left(\frac{1}{\left[\beta c(1+i)^{1-\gamma}\right]^{\frac{1}{\gamma}}+1}\right)^{1-\gamma}\left(1+(1+i)^{1-\gamma}\beta c\left(\left[\beta c(1+i)^{1-\gamma}\right]^{\frac{1}{\gamma}}\right)^{1-\gamma}\right) \\
&= \left(\frac{1}{\left[\beta c(1+i)^{1-\gamma}\right]^{\frac{1}{\gamma}}+1}\right)^{1-\gamma}\left(1+(1+i)^{1-\gamma}\beta c\left[\beta c(1+i)^{1-\gamma}\right]^{\frac{1}{\gamma}-1}\right) \\
&= \left(\frac{1}{\left[\beta c(1+i)^{1-\gamma}\right]^{\frac{1}{\gamma}}+1}\right)^{1-\gamma}\left(1+\left[\beta c(1+i)^{1-\gamma}\right]^{\frac{1}{\gamma}}\right) \\
&= \left(\left[\beta c(1+i)^{1-\gamma}\right]^{\frac{1}{\gamma}}+1\right)^{\gamma-1}\left(1+\left[\beta c(1+i)^{1-\gamma}\right]^{\frac{1}{\gamma}}\right) \\
c &= \left(\left[\beta c(1+i)^{1-\gamma}\right]^{\frac{1}{\gamma}}+1\right)^{\gamma}
\end{aligned}$$

Ahora, nos queda despejar c.

$$\begin{aligned} c^{\frac{1}{\gamma}} &= \beta^{\frac{1}{\gamma}} c^{\frac{1}{\gamma}} (1+i)^{\frac{1}{\gamma}-1} + 1 \\ 1 &= \beta^{\frac{1}{\gamma}} (1+i)^{\frac{1}{\gamma}-1} + c^{-\frac{1}{\gamma}} \\ c^{-\frac{1}{\gamma}} &= 1 - \beta^{\frac{1}{\gamma}} (1+i)^{\frac{1}{\gamma}-1} \\ c &= \left(1 - \beta^{\frac{1}{\gamma}} (1+i)^{\frac{1}{\gamma}-1}\right)^{-\gamma} \end{aligned}$$

Exercise 1.6. Let $\{\xi_k\}$ be a dynamics of iid random variables such that $E[\xi] = 0$ and $E[\xi^2] = d$. Consider the dynamics

$$x_{k+1} = x_k + a_k + \xi_k, k = 0, 1, 2, \ldots,$$

and the discounted cost

$$E\left[\sum \beta^k \left(a_k^2 + x_k^2\right)\right]$$

1. Write down the associated Bellman equation.
2. Conjecture that the solution to the Bellman equation takes the form $v(x) = ax^2 + b$, where a and b are constant.
3. Determine the constants a and b.
4. Conjecture that the solution to the Bellman equation takes the form $v(x) = ax^2 + b$, where a and b are constant. Determine the constants a and b.

$$\begin{aligned} \nu(x) &= \max_{a \in A(x)} \{c(x,a) + \nu(f(x,a))\} \\ &= \max_{a \in A(x)} \left\{a^2 + x^2 + E[\nu(x+a+\xi)]\right\} \end{aligned}$$

Para $\nu(x) = ax^2 + b$

$$\begin{aligned} \nu(x) &= \max_{a \in A(x)} \{c(x,a) + \beta E[\nu(f(x,a))]\} \\ &= \max_{a \in A(x)} \left\{A^2 + x^2 + \beta\left(E\left[a\left(f^2(x,a)\right)\right] + b\right)\right\} \\ &= \max_{a \in A(x)} \left\{A^2 + x^2 + \beta\left(aE\left[f^2(x,a)\right] + b\right)\right\} \end{aligned}$$

Notemos que

$$
\begin{aligned}
E\left[f^2(x,A)\right] &= E\left[(x+A+\xi)^2\right] \\
&= E\left[x^2+A^2+\xi^2+2xA+2x\xi+2\xi A\right] \\
&= x^2+A^2+E\left[\xi^2\right]+2xA+2xE\left[\xi\right]+2AE\left[\xi\right] \\
&= x^2+A^2+d+2xA
\end{aligned}
$$

Entonces

$$
\begin{aligned}
ax^2+b &= \max_{a\in A(x)}\left\{A^2+x^2+\beta\left[a\left(x^2+A^2+d+2xA\right)+b\right]\right\} \\
&= \max_{a\in A(x)}\left\{A^2+x^2+\beta a\left(x^2+A^2+d+2xA\right)+\beta b\right\} \\
&= \max_{a\in A(x)}\left\{A^2+x^2+\beta ax^2+\beta aA^2+\beta ad+2\beta axA+\beta b\right\} \\
&= \max_{a\in A(x)}\left\{A^2\left(\beta a+1\right)+2\beta axA+x^2+\beta ax^2+\beta ad+\beta b\right\}
\end{aligned}
$$

Definimos

$$
w(x,A)=A^2\left(\beta a+1\right)+2\beta axA+x^2+\beta ax^2+\beta ad+\beta b,
$$

entonces

$$
\partial_A w = 2A\left(\beta a+1\right)+2\beta ax.
$$

Si $\partial_A w=0$, entonces

$$
A=-\frac{\beta ax}{\beta a+1}
$$

Entonces

$$
\begin{aligned}
\nu(x) &= \left(\beta ax\right)^2-2\frac{\left(\beta ax\right)^2}{\beta a+1}+x^2+\beta ax^2+\beta ad+\beta b \\
&= x^2\left(\left[\beta a\right]^2-2\frac{\left(\beta a\right)^2}{\beta a+1}+1+\beta a\right)+\beta ad+\beta b
\end{aligned}
$$

Entonces

$$
\begin{aligned}
a &= \left[\beta a\right]^2-2\frac{\left(\beta a\right)^2}{\beta a+1}+1+\beta a \\
b &= \beta ad+\beta b,
\end{aligned}
$$

de forma rapida

$$b = \frac{\beta a d}{1-\beta},$$

entonces queda pendiente calcular a

$$\begin{aligned}
a &= [\beta a]^2 - 2\frac{(\beta a)^2}{\beta a + 1} + 1 + \beta a. \\
0 &= (\beta a)^2 \left(1 - \frac{2}{\beta a + 1}\right) + 1 + (\beta - 1)\, a \\
&= (\beta a)^2 (\beta a + 1 - 2) + \beta a + 1 + (a\beta - a)(\beta a + 1) \\
&= (\beta a)^2 (\beta a - 1) + \beta a + 1 + \left[(a\beta)^2 + a\beta - \beta a^2 - a\right] \\
&= (\beta a)^3 + 2a\beta + 1 - \beta a^2 - a \\
&= \beta^3 a^3 - \beta a^2 + (2\beta - 1)\, a + 1
\end{aligned}$$

2 Tarea 2

References