Aprendizaje Reforzado

David Peña Peralta

2024-02-09

Table of contents

Preface

Ejercicio 2

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.

```
import numpy as np

x = 10
```

Ejercicio 4

Considerando J_N como sigue

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

con $K_N = 1$ bajo la hipótesis de que

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

calculamos J_{N-1}.

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

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

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

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

entonces

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

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

Entonces

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

Ahora calculamos J_{N-2}

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

donde

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

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

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

entonces

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

y

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

Por lo tanto, tenemos que

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

con $K_N = 1$.

Obteniendo así

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

Ejercicio 5

Para el caso infinito. Estamos considerando

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

Entonces

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

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

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

definimos

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

entonces

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

Si $\partial_a q = 0$. Entonces

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

Por lo tanto

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

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.

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

Ejercicio 6

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

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

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

Notemos que

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

Entonces

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

Definimos

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

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

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

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

$$b=\frac{\beta ad}{1-\beta},$$

entonces queda pendiente calcular a

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

1 Introduction

This is a book created from markdown and executable code.

See Knuth (1984) for additional discussion of literate programming.

2 Summary

In summary, this book has no content whatsoever.

References

Knuth, Donald E. 1984. "Literate Programming." *Comput. J.* 27 (2): 97–111. https://doi.org/10.1093/comjnl/27.2.97.