Los Procesos de Decisión de Markov Parcialmente Observable y su aplicación al Problema de Consumo e Inversión

Octavio Paredes Pérez[1], Víctor Hugo Vázquez Guevara[2]

Benemérita Universidad Autónoma de Puebla, Facultad de Ciencias Físico Matemáticas

Octavio.paredespe@alumno.buap.mx[1], vvazquez@fcfm.buap.mx[2]

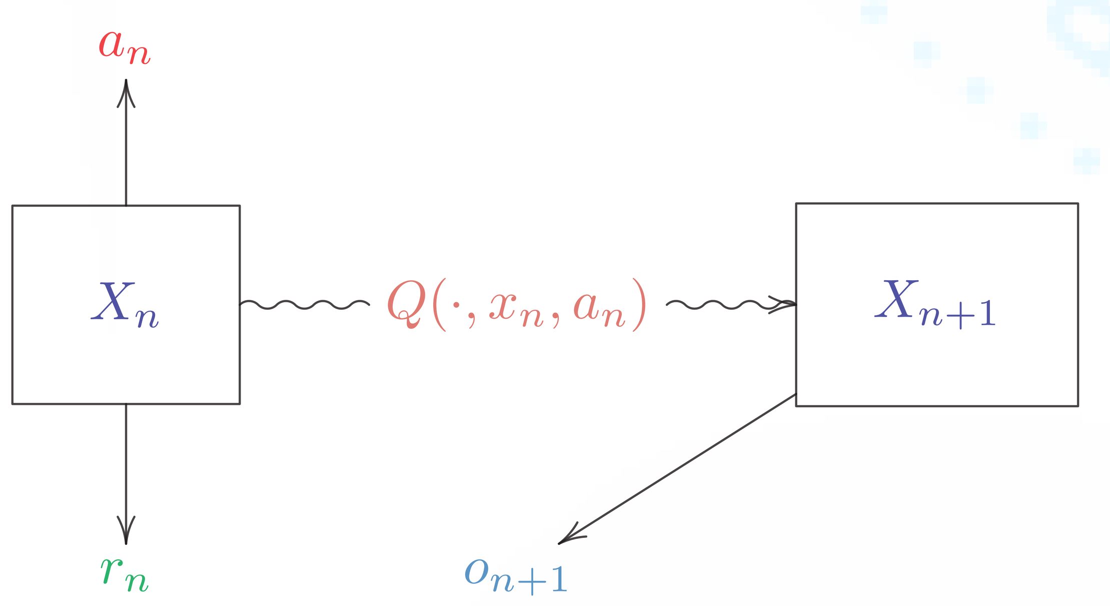
Introducción

Se tendrá el Problema de Consumo e Inversión, en el cual se presenta un cierto inversionista con cierta riqueza o capital inicial y al comienzo de cada uno de un número de N de periodos, este puede decidir que parte consumir y que parte invertir en el mercado financiero. Pero con la sorpresa de que después de cada parte consumida e invertida, este no sabe con exactitud cuanto capital tendrá realmente, pero se podrá deducir mediante una distribución de probabilidad lo que se podría tener en ese instante o momento.

- Para este tipo de problemas existe la teoría de los Procesos de Decisión de Markov Parcialmente Observables. Aquí el agente no puede observar directamente el estado subyacente del sistema. En cambio, puede ver una observación basada en la acción y el estado resultante.
- El objetivo del agente sigue siendo maximizar la recompensa futura acumulada esperada descontada.

PDMPO

Un PDMPO a tiempo discreto modela la relación entre un agente y su entorno. Formalmente, un PDMPO está constituido por $(E, A, D(x), Q, r_n, g_N, \ \Omega, O, \beta)$, para cualquier $n = 0, 1, \ldots, N$.



Problema de Consumo e inversión

Se tendrá:

1. Un inversionista con riqueza inicial $X > 0$,
2. En ciertos instantes se permite el consumo y la inversión de capital disponible,
3. El consumo se evalúa mediante una función de utilidad $U_c(c_n)$,
4. La riqueza restante es invertida en un activo riesgoso y en un bono sin riesgo,
5. La riqueza terminal g_N produce otra función de utilidad $U_p(X_N)$.

Propósito

¿Cómo debería Consumir e Invertir su capital el inversionista, con el fin de maximizar la suma de sus utilidades esperadas?.

Dinámica del proceso de riqueza

El proceso de la riqueza evolucionará de la siguiente forma:

$$X_{n+1} = (1 + i_{n+1})(X_n - c_n + a_n \cdot R_{n+1})$$

Modelo de Decisión de Markov

El Problema puede ser resuelto por el siguiente modelo de decisión de Markov:

- $E := [0, \infty)$,
- $B := [0, 1]$,
- $A := \mathbb{R}^+ \times \mathbb{R}^d$,
- $D(x) := \{(c_n, a_n) \in A | 0 \leq c_n \leq x_n \text{ y } (1 + i_{n+1})(x_n - c_n + a_n \cdot R_{n+1}) \geq 0, \ \mathbb{P} - c.s.\}$
- $\mathcal{Z} := [-1, \infty)^d$ donde $z_n \in \mathcal{Z}$ denota el riesgo relativo,
- $T(b_n, c_n, a_n, z_n) := (1 + i_{n+1})(b_n - c_n + a_n \cdot z_n)$, la función de transición,
- Q^Z la distribución de R_{n+1},
- $r_n(b_n, c_n, a_n) := \sum_{x_n \in E} b_n(x_n) U_c(c_n)$, la función de recompensa por etapa,
- $g_N(X_N) := U_p(X_N)$, la función de recompensa terminal.
- $\beta \in (0, 1]$ factor de descuento.

Objetivo

Será optimizar el criterio de rendimiento que es el de la recompensa total esperada descontada o en su defecto el costo, que está definido por:

$$V_n(x) = \sup_{\pi \in \Pi} \mathbb{E}^{\pi}_{nx} \left[\sum_{K=0}^{N-1} \beta b_n(x) U_c(c_n(b_n)) \right. \\ \left. + \, b_N(x) U_p(X_N) \right].$$

Donde el supremo se tomará sobre todas las políticas $\pi = \{f_0, \ldots, f_{N-1}\}$ con $f_n(b_n) = (c_n(b_n), a_n(b_n))$.

Función de Utilidad

En este trabajo se asumirá que ambas funciones de utilidad U_c y U_p serán de la forma:

$$-\gamma^{-1} \exp(-\gamma x_n), \ \gamma > 0 \text{ y } x_n \in E,$$

Entonces será posible deducir el siguiente resultado, el cual nos dará un mecanismo para encontrar la solución óptima para el Problema de Consumo e Inversión y cuya función de utilidad será creciente y aversa al riesgo. Además, su prima de riesgo será invariante con respecto a la riqueza.

Solución

Teorema. *Se supondrá que ambas funciones $U_c(c_n)$ y $U_p(X_N)$ son funciones de utilidad exponencial con $\gamma > 0$. Entonces se tendrá que:*

$$V_{N+1}(b_{N+1}) = 0,$$
$$V_N(b_N) = -b_N(x)\gamma^{-1} \exp(-\gamma b_{N+1})$$
$$V_n(b_n) = -k_n \exp(-d_n b_{n+1}), \ n \in \{0, 1, \ldots, N-1\},$$

el consumo óptimo por etapa n es $c_n^* = \zeta_n^* b_{n+1}$, donde

$$\frac{\gamma(1 + i_{n+1}) b_{n+1} \ln |b_n(x)| - \ln(\beta \mathbb{P}(o|b_n, a) b_{n+1}(x)(1 + i_{n+1}) v_n|}{\gamma + \gamma(1 + i_{n+1})}$$

y la inversión óptima por etapa n es

$$a_n^* = \frac{S_N^0}{S_n^0(1 + i_{n+1})} \alpha_n^*,$$

donde α_n^* es la solución de v_n, que se puede encontrar a través de $\alpha_n^* = \frac{1}{\gamma} \frac{S_n^0}{S_N^0} \tilde{\alpha}_n$ donde $\tilde{\alpha}_n$ es el mínimo de

$$\alpha \mapsto \mathbb{E}[exp(-\alpha \cdot R_{n+1})], \ \alpha \in \mathbb{R}^d.$$

Conclusiones

Los Procesos de Decisión de Markov Parcialmente Observables sirven de marco sólido para la toma de decisiones en entornos caracterizados por la incertidumbre y la información incompleta. Mientras que los Procesos de Decisión de Markov (MDP) con información completa nos darán un marco matemático para modelar decisiones en el cual se muestra un sistema con una serie de estados y proporciona acciones al tomador de decisiones basadas en dichos estados. Por lo que los PDMPO se basan también en ese concepto, pero con la diferencia de que cómo un sistema puede hacer frente a los desafíos cuando se tiene la observación limitada.

Se obtuvo lo siguiente:

1. Se resolvió el problema de Consumo e Inversión a partir de información incompleta.
2. Se obtuvo un mecanismo para la obtención de la política óptima para el Problema de Consumo e Inversión.

Referencias

[1] Bäuerle N., Rieder U., *Markov Decision Processes with Aplications to Finance. Springer Verlang*, ISBN9783642183232, 2011.

[2] Bertsekas, D. P. *Dynamic Programming nd Stochastic Control*, College of Engineering Coordinated Science Laboratory University of Illinois Urbana, Illinois. Academic Press, ISBN:978-0120932504,

[3] Kaelbling, L.P., Littman, M.L., & Cassandra, A.R. *Planning and acting in partially observable stochastic domains*, Artificial Intelligence, Vol. 101. pp. 99–134. doi:10.1016/S0004-3702(98)00023-X, 1998.

[4] Sondik, E.J. *The optimal control of partially observable Markov processes over the infinite horizon: discounted cost*, Operations Research. Vol. 26. pp. 282–304. doi:10.1287/opre.26.2.282, 1978.

[5] Vikram Krishnamurthy. *Partially Observed Markov Decision Processes*, Cornell University/Cornell Tech. Cambridge University Press, ISBN:9781316471104, 2016.