Rekursywna użyteczność w decyzyjnych procesach Markowa

Model rekursywnej użyteczności zasugerowany przez Krepsa i Porteusa (1978) oraz opisany dokładnie przez Epsteina i Zina (1989) jest powszechnie stosowany w dynamice gospodarczej. W 1978 roku D. Assaf udowodnił, że w przypadku szczególnym można policzyć analitycznie politykę optymalną i stacjonarną.
Moją motywacją było zmodyfikowanie równania Bellmana do nowego równania wrażliwego na ryzyko i zbadanie jego własności w szczególnych przypadkach, bazując na pracy D. Assafa.

Model

- Przestrzeń stanów X - standardowy zbiór borelowski, $X \neq \emptyset$,
- A(x) - zbiór borelowski, nazywany zbiorem akcji,
- D jest niepustym borelowskim podzbiorem $X \times A$,
- $\forall x \in X \ A(x) = \{a \in A : (x, a) \in D\}$ jest zwarty.
- Funkcja użyteczności $r : D \to R$ - ograniczona, borelowsko-mierzalna, $r(x, \cdot)$ jest ciągła na $A(x)$ dla każdego x,
- q jest regularnym rozkładem warunkowym z D do X.

Opis modelu

- Niech $x \in X$ będzie stanem początkowym,
- Decydent wybiera akcję $a \in A(x)$,
- Wtedy decydent czerpie chwilową użyteczność $r(x, a)$,
- Następny stan x' generowany jest z rozkładu $q(\cdot|x, a)$.
- Później procedura rozpoczyna się od nowa ze stanem x'.
- Stany i decyzje tworzą historię (trajektorię procesu) $h = (x_n, a_n)_{n\in\mathbb{N}} \in H := D^{\infty}$.

Standardowe programowanie dynamiczne

W standardowym programowaniu dynamicznym

- decydent stara się maksymalizować użyteczność od trajektorii:

$$U(x, \sigma) := E_x^{\sigma} \sum_{t=1}^{k} \beta^{t-1} \cdot r(x_t, a_t),$$

  gdzie $x \in X$ oznacza stan początkowy, a $\sigma \in \Sigma$, Σ oznacza zbiór wszystkich polityk (niezrandomizowanych lub zrandomizowanych), E_x^{σ} oznacza wartość oczekiwaną na zbiorze H wyznaczoną przez stan początkowy x i politykę σ.

- Optymalna wartość to:

$$V(x) = \sup_{\sigma\in\Sigma} U(x, \sigma)$$

- W standardowym programowaniu dynamicznym wyznaczenie optymalnej wartości polityki stacjonarnej sprowadza się do rozwiązania równania Bellmana, które jest postaci

$$V(x) = \sup_{a\in A(x)} \Big(r(x, a) + \beta \int_X V(y) q(dy \mid x, a) \Big).$$

Entropijna miara ryzyka

Niech $(\Omega, \mathbb{F}, P)$ będzie przestrzenią probabilistyczną i niech $f \in L^\infty(\Omega, \mathbb{F}, P)$ będzie losową wypłatą. Entropijną miarą ryzyka nazywamy

$$\rho(f) = -\frac{1}{\gamma} ln \int_\Omega e^{-\gamma f(\omega)} P(d\omega),$$

gdzie $\gamma > 0$ jest współczynnikiem ryzyka.
Zachodzą następujące własności ρ:

- monotoniczność, tzn. $f \leqslant g \Rightarrow \rho(f) \leqslant \rho(g)$
- translacja, tzn. $\rho(f + x) = \rho(f) + x \ \forall x \in \mathbb{R}$
- nierówność Jensena, tzn. $\rho(f) \leqslant Ef$
- miara nie jest dodatnio jednorodna, tzn. $\rho(\alpha f) \neq \alpha\rho(f)$ dla $\alpha > 0$

Jeśli $\gamma > 0$ jest dostatecznie bliskie 0, to $\rho(f) \approx E(f) - \frac{\gamma}{2} Varf$

Zamiana wartości oczekiwanej na entropijną miarę ryzyka.

1. Dla decydenta w standardowym programowaniu dynamicznym, czyli obojętnego na ryzyko całkowita użyteczność wyraża się wzorem:

$$U(x,\sigma) = E(r(x_1,a_1) + \beta r(x_2,a_2) + \ldots + \beta^{n-1} r(x_n,a_n) + \ldots$$

$$= \lim_n Er(x_1,a_1) + \beta\left(E_2 r(x_2,a_2) + \beta\left(E_3 r(x_3,a_3) + \ldots + \beta(E_n r(x_n,a_n))\right)\right),$$

(w środku nawiasów mamy odpowiednie warunkowe wartości oczekiwane $E_2, E_3, ...$)

2. Gdy odpowiednie wartości oczekiwane zmienimy na ρ ($E := \rho$) otrzymamy użyteczność całkowitą dla decydenta wrzżliwego na ryzyko:

$$U^\gamma(x,\sigma) = \lim_n r(x_1,a_1) + \beta\rho\left(r(x_2,a_2) + \beta\rho\left(r(x_3,a_3)\ldots + \beta\left(\rho(r(x_n,a_n))\right)\right)\right).$$

Decydent wrażliwy na ryzyko maksymalizuje U^γ.

Zmodyfikowane równanie Bellmana

Dla decydentów wrażliwych na ryzyko

- Można zmodyfikować równania Bellmana w następujący sposób: znaleźć optymalną wartość $V(\cdot)$ tak aby rozwiązywała równanie

$$V(x) = \sup_{a \in A(x)} \left(r(x,a) - \frac{\beta}{\gamma} ln \int_X e^{-\gamma V(y)} q(dy \mid x,a) \right)$$

  gdzie $\gamma > 0$ jest miarą wrażliwości na ryzyko.

- Zauważmy, że

$$-\frac{\beta}{\gamma} ln \int_X e^{-\gamma V(y)} q(dy \mid x,a) \to \int_X V(y) q(dy \mid x,a) \text{ gdy } \gamma \downarrow 0,$$

- Dla $\gamma \approx 0$ mamy w przybliżeniu standardowe równanie Bellmana.

Rezultaty

Założenia:

- Zbiór $A(x)$ jest zwarty
- Dla każdego $x \in X$ i każdego zbioru borelowskiego $C \subset X$, funkcja $q(C \mid x, \cdot)$ jest ciągła na $A(x)$.
- Funkcja użyteczności $r(x, \cdot)$ jest górnie półciągła na $A(x)$ $\forall x \in X$.

Oznaczmy $U(X)$ jako zbiór wszystkich ograniczonych nieujemnych górnie półciągłych mierzalnych funkcji borelowskich na X.

Twierdzenie:

Istnieje dokładnie jedna funkcja $V \in U(X)$ i strategia $f^* \in F$ taka, że

$$V(x) = \sup_{a \in A(x)} \left(r(x,a) - \frac{\beta}{\gamma} ln \int_X e^{-\gamma V(y)} q(dy \mid x, a) \right)$$

dla każdego $x \in X$. Co więcej, f^* jest polityką optymalną i stacjonarną.

Schemat rozwiązania równania optymalnego dla problemu programowania dynamicznego

Niech $A(x) = A := a_0, a_1, ...a_n$ i $q_1(\cdot) := q(\cdot \mid a_1)$, $w(x) = e^{-\gamma V(x)}$ dla $x \in X$ i

$$g(x,a) = e^{-\gamma r(x,a)}, (x,a) \in D := X \times A.$$

Wtedy równanie Bellmana jest postaci

$$w(x) = \min_{a \in A} \Big[g(x,a) \Big(\int_X w(y) q(dy \mid a) \Big)^{\beta} \Big], x \in X.$$

Definiujemy

$$g_i(x) := g(x, a_i), W_i = \int_X w(y) q_i(dy)$$

dla $i = 0, ..., n, x \in X$ i podzbioru

$$X_i := \Big\{ x \in X : g_i(x) W_i^{\beta} < g_j(x) W_j^{\beta} \text{ dla } j = 1, ..., i-1, \text{ i } g_i(x) W_i^{\beta} \leqslant g_j(x) W_j^{\beta} \text{ dla } j = i+1, ..., n \Big\}$$

dla każdego i=0,...,n. Równanie Bellmana sprowadza się do postaci

$$w(x) = \sum_{i=0}^{n} g_i(x) W_i^{\beta} 1_{X_i}(x)$$

Korzystając z nierówności w odniesieniu do q_j, dostajemy

$$W_j = \sum_{i=0}^{n} G_{ij} W_i^{\beta}, j = 0, ..., n,$$

gdzie

$$G_{ij} := \int_{X_i} g_i(x) q_j(dx).$$

Otrzymujemy

$$V(x) = -\frac{1}{\gamma} ln \Big(\sum_{i=0}^{n} g_i(x) W_i^{\beta} 1_{X_i}(x) \Big),$$

polityka optymalna jest postaci

$$f^*(x) = \begin{cases} a_0 & x \in X_0 \\ \vdots & \\ a_n & x \in X_n. \end{cases}$$

Przykład

Założenia:

- $A := \{a_0, a_1, a_2\}$, $X = [0,1]$,
- $r(x, a_0) := r_0(x) = 4 - x$,
- $r(x, a_1) := r_1(x) = 5 - 3x$,
- $r(x, a_2) := r_2(x) = 5 - 5x$,
- $q(\cdot|x, a_0) := q_0(\cdot)$ - rozkład jednostajny na $[0,1]$ dla wszystkich x,
- $q(\cdot|x, a_1) := q_1(\cdot)$ dla wszystkich x gdzie $q_1(0) = q_1(0.5) = 0.5$,
- $q(\cdot|x, a_2) := q_2(\cdot)$, gdzie $q_2(0) = 1$.

Niech $\gamma = 1.$ i oznaczmy

$$K_0 = W_0/W_1, \quad K_1 = W_0/W_2, \quad K_2 = W_1/W_2.$$

Niech $\beta = 0.5$. Można udowodnić, że $1/2 \in X_1$ $0 \in X_2$ i $1 \in X_0$.

$$X_0 = \{x \geqslant \ln\sqrt{K_0^\beta e}\}, \quad X_1 = \{\ln\sqrt{K_2^\beta} \leqslant x < \ln\sqrt{K_0^\beta e}\}, \quad X_2 = \{x < \ln\sqrt{K_2^\beta}\}$$

Obliczamy G_{ij}:

$$\begin{aligned}
G_{00} =& e^{-4}(e - \sqrt{K_0^\beta e}) = e^{-4}(e - \sqrt{(W_0/W_1)^\beta e}),\\
G_{10} =& \frac{1}{3}e^{-5}\Big((K_0^\beta e)^{\frac{3}{2}} - (K_2^\beta)^{\frac{3}{2}}\Big),\\
G_{20} =& \frac{1}{5}e^{-5}\Big((K_2^\beta)^{\frac{5}{2}} - 1\Big),\\
G_{01} =& 0, \quad G_{11} = 1/2e^{-7/2}, \quad G_{21} = 1/2e^{-5}, \quad G_{02} = G_{12} = 0, \quad G_{22} = e^{-5}
\end{aligned}$$

Z równania opisującego W_j dostajemy:

$$W_0 = 0.00053384, W_1 = 0.0002715, W_2 = 0.0000454.$$

Dostajemy

$$\begin{aligned} X_0 =& \Big\{x \geqslant 0.669034\Big\}, \\ X_1 =& \Big\{0.447113 \leqslant x < 0.669034\Big\}, \\ X_2 =& \Big\{x < 0.447113\Big\}. \end{aligned}$$

Polityka optymalna to

$$f^*_{0.50}(x) = \begin{cases} a_0, & x \in [0.669034; 1] \\ a_1, & x \in [0.447113; 0.669034) \\ a_2, & x \in [0; 0.447113), \end{cases}$$

natomiast rozwiązaniem równania Bellmana jest

$${}_{0.50}V(x) = \begin{cases} -7.767707 + x, & x \in X_0 \\ -9.105774 + 3x, & x \in X_1 \\ -10 + 5x, & x \in X_2. \end{cases}$$

Niech $\beta = \{0.51, 0.52, 0.53\}$. Wtedy polityki optymalne są następującej postaci:

$$f^*_{0.51} = \begin{cases} a_0, & x \in [0.673155; 1] \\ a_1, & x \in [0.463895; 0.673155) \\ a_2, & x \in [0; 0.463895), \end{cases}$$

$$f^*_{0.52} = \begin{cases} a_0, & x \in [0.677106; 1] \\ a_1, & x \in [0.481337; 0.677106) \\ a_2, & x \in [0; 0.481337), \end{cases}$$

$$f^*_{0.53} = \begin{cases} a_0, & x \in [0.680865; 1] \\ a_1, & x \in [0.499455; 0.680865) \\ a_2, & x \in [0; 0.499455), \end{cases}$$

a rozwiązania równania Bellmana (minimalizacja kosztów) to:

$${}_{0.51}V(x) = \begin{cases} -7.92998 + x, & x \in X_0 \\ -9.27629 + 3x, & x \in X_1 \\ -10.2041 + 5x, & x \in X_2, \end{cases}$$

$${}_{0.52}V(x) = \begin{cases} -8.09978 + x, & x \in X_0 \\ -9.45399 + 3x, & x \in X_1 \\ -10.4167 + 5x, & x \in X_2, \end{cases}$$

$$
{}_{0.53}V(x) = \begin{cases} -8.27766 + x, & x \in X_0 \\ -9.63939 + 3x, & x \in X_1 \\ -10.6383 + 5x, & x \in X_2. \end{cases}
$$

Łatwo zauważyć, że dla $\beta = 0.54$ potrzebujemy zmiany założeń. Z tego powodu zakładamy, że $1/2 \in X_2$ (tzn. $1/2 \notin X_1$). Chcemy wyznaczyć ponownie $G_{i,j}$:

$$
\begin{aligned}
G_{00} =& e^{-4}(e - \sqrt{K_0^\beta e}), \\
G_{10} =& \frac{1}{3} e^{-5}\Big((K_0^\beta e)^{\frac{3}{2}} - (K_2^\beta)^{\frac{3}{2}}\Big), \\
G_{20} =& \frac{1}{5} e^{-5}\Big((K_2^\beta)^{\frac{5}{2}} - 1\Big), \\
G_{01} =& 0, \quad G_{11} = 0, \quad G_{21} = 1/2e^{-5} + 1/2e^{\frac{-5}{2}}, \quad G_{02} = G_{12} = 0, \quad G_{22} = e^{-5}
\end{aligned}
$$

Mamy teraz

$$
\begin{aligned}
X_0 =& \Big\{x \geqslant 0.691905\Big\}, \\
X_1 =& \Big\{0.509150 \leqslant x < 0.691905\Big\}, \\
X_2 =& \Big\{x < 0.509150\Big\}.
\end{aligned}
$$

Niech $\beta = 0.75$. Wtedy $0.5 \in X_2$. Otrzymujemy

$$
W_0 = 0.0000000467043, W_1 = 0.00000001358557, W_2 = 0.000000002061154.
$$

Stąd, polityka optymalna wynosi

$$
f^*_{0.75}(x) = \begin{cases} a_0, & x \in [0.963061; 1] \\ a_1, & x \in [0.7071533; 0.963061) \\ a_2, & x \in [0; 0.7071533). \end{cases}
$$

Rozwiązanie równania Bellmana jest postaci

$$
{}_{0.75}V(x) = \begin{cases} -16.6596 + x, & x \in X_0 \\ -18.5857 + 3x, & x \in X_1 \\ -20 + 5x, & x \in X_2. \end{cases}
$$

Bibliografia

- Assaf, D., Invariant problems in discounted dynamic programming, Adv. Appl. Prob. 10 (1978).
- Becker, R.A., J.H. Boyd, Capital Theory, Equilibrium Analysis and Recursive Utility. Blackwell Publisher, New York, 1997.
- Epstein, L.G., S.E. Zin, Substitution, risk aversion, and the temporal behavior of consumption and asset returns: a theoretical framework. Econometrica 57 (1989).
- Jaśkiewicz, A., A note on risk-sensitive control of invariant models, Systems Control Lett. 56 (2007).
- Kreps, D.M., E.L. Porteus, Temporal resolution of uncertainty and dynamic choice, Econometrica 46 (2000.
- Miao, J., Economic Dynamics in Discrete Time. MIT Press, 2014.